

Research article

Open Access

## Phylogenetic relatedness and host plant growth form influence gene expression of the polyphagous comma butterfly (*Polygonia c-album*)

Hanna M Heidel-Fischer<sup>†1</sup>, Dalial Freitak<sup>†1</sup>, Niklas Janz<sup>\*†2</sup>, Lina Söderlind<sup>2</sup>, Heiko Vogel<sup>1</sup> and Sören Nylin<sup>2</sup>

Address: <sup>1</sup>Department of Entomology, Max Planck Institute for Chemical Ecology, Hans-Knöll-Straße 8, D-07745 Jena, Germany and <sup>2</sup>Department of Zoology, Stockholm University, 106 91 Stockholm, Sweden

Email: Hanna M Heidel-Fischer - hfischer@ice.mpg.de; Dalial Freitak - dalial.freitak@agrar.uni-giessen.de; Niklas Janz\* - niklas.janz@zoologi.su.se; Lina Söderlind - lina.soderlind@zoologi.su.se; Heiko Vogel - hvogel@ice.mpg.de; Sören Nylin - soeren.nylin@zoologi.su.se

\* Corresponding author †Equal contributors

Published: 31 October 2009

Received: 21 April 2009

BMC Genomics 2009, 10:506 doi:10.1186/1471-2164-10-506

Accepted: 31 October 2009

This article is available from: <http://www.biomedcentral.com/1471-2164/10/506>

© 2009 Heidel-Fischer et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** The mechanisms that shape the host plant range of herbivorous insect are to date not well understood but knowledge of these mechanisms and the selective forces that influence them can expand our understanding of the larger ecological interaction. Nevertheless, it is well established that chemical defenses of plants influence the host range of herbivorous insects. While host plant chemistry is influenced by phylogeny, also the growth forms of plants appear to influence the plant defense strategies as first postulated by Feeny (the "plant apparency" hypothesis). In the present study we aim to investigate the molecular basis of the diverse host plant range of the comma butterfly (*Polygonia c-album*) by testing differential gene expression in the caterpillars on three host plants that are either closely related or share the same growth form.

**Results:** In total 120 genes were identified to be differentially expressed in *P. c-album* after feeding on different host plants, 55 of them in the midgut and 65 in the restbody of the caterpillars. Expression patterns could be confirmed with an independent method for 14 of 27 tested genes. Pairwise similarities in upregulation in the midgut of the caterpillars were higher between plants that shared either growth form or were phylogenetically related. No known detoxifying enzymes were found to be differently regulated in the midgut after feeding on different host plants.

**Conclusion:** Our data suggest a complex picture of gene expression in response to host plant feeding. While each plant requires a unique gene regulation in the caterpillar, both phylogenetic relatedness and host plant growth form appear to influence the expression profile of the polyphagous comma butterfly, in agreement with phylogenetic studies of host plant utilization in butterflies.

### Background

Chemical defenses of plants influence the host range of herbivorous insects [1-4]. Although by no means the only factor involved in shaping insect-host associations [5],

few researchers would argue against its general importance. However, there is an ongoing argument over the reasons for similarities among plant defense chemicals. There are two ways for evolution to achieve similarity:

either through shared ancestry or through evolutionary convergence (or parallelism). Ehrlich and Raven [2] suggested that related insects tend to feed on related plants, and several other studies have continued to demonstrate a role of host plant phylogeny (shared ancestry) on patterns of host plant use [6-14]. Hence, there is strong support for a historical component in patterns of host plant use.

On the other hand, several authors have also pointed out that plant chemistry does not always follow phylogeny [10,15,16]. Feeny [17] suggested that plant defense strategies should differ depending on their "apparency"; plants that are long-lived and/or physically large will always be found by attacking insects and should possess constitutive chemical defenses such as tannins, terpenes, and flavanoids (see also [18-21]) that have a quantitative, dosage-dependent effect. Unapparent plants, with lower risk of detection by herbivores, should instead tend to utilize induced chemical defenses. Trees should be the most apparent of plants as they are both physically large and long-lived. According to the apparency hypothesis, trees should then tend to have more convergent constitutive defenses than herbs, and as a consequence we should see more host shifts involving trees than herbs. Despite some criticism and uncertainty regarding the mechanisms involved, the apparency theory has remained a useful framework for thinking about the evolution of plant chemical defense in the context of plant community heterogeneity [22].

In a phylogenetic reanalysis of Ehrlich and Raven's [2] study on butterfly and plant coevolution, Janz & Nylin [9] found strong effects of both plant phylogeny and growth form on patterns of host use among butterflies. An overwhelming majority of host shifts occurred while feeding on trees, giving support for Feeny's "plant apparency" hypothesis. Trees appeared to serve as a "bridge" that could facilitate host shifts between distantly related plants.

Hence, there is support for both shared ancestry and convergent evolution in the large-scale chemical structuring of insect-host associations, but the mechanistic basis remains largely unknown. However, recent years have seen great progress in understanding of the molecular mechanisms that enable insect to feed on certain host plants. In general it is assumed that insects apply phase I and phase II detoxifying enzymes to metabolize secondary plant compounds. Several studies have for example revealed the important role of the cytochrome P450 enzyme family for detoxification of plant secondary compounds as well as insecticides. [23-27]. Glutathione-S-transferases (GSTs) have also been shown to be induced in generalist and specialist lepidopteran larvae upon feed-

ing on their host plants [28,29]. Wittstock et al. [30] identified the Nitrile-specifier Protein (NSP) in *Pieris rapae*. NSP redirects the glucosinolate hydrolysis and by that enables the Pierinae butterflies to feed on the plant family Brassicaceae. Further research has been done on the evolution of NSP showing its evolution by domain and gene duplication from a gene of unknown function that is widespread in insect species [31,32]. *Plutella xylostella* is also feeding on glucosinolate containing plants. Here the Glucosinolate sulfatase (GSS) inhibits the hydrolysis of glucosinolates completely by forming a desulfo-glucosinolate. In spite of progress in recent years, much is still to be discovered in the detoxification mechanisms of insects.

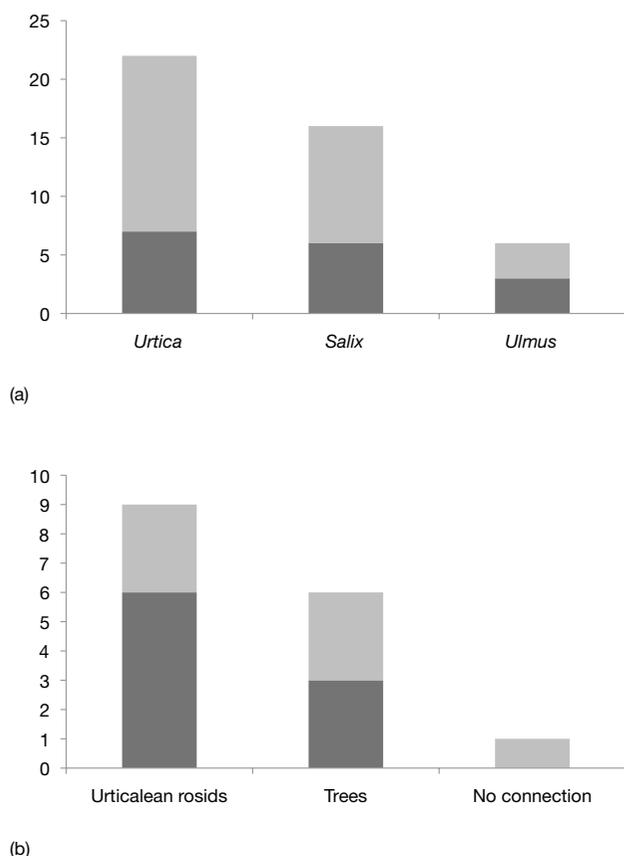
In the present study we aim to investigate the molecular basis of the diverse host plant range of the comma butterfly (*Polygonia c-album*, Lepidoptera: Nymphalidae). *P. c-album* is a widespread polyphagous butterfly species of the family Nymphalidae. It is found all over Eurasia, from England to Japan and from the central of Sweden to the northern tip of Africa. The larvae can be found on host plants from several taxa: the "urticalean rosids" *Urtica*, *Humulus* and *Ulmus* and the distantly related *Salix* (Salicaceae), *Ribes* (Grossulariaceae), *Betula* and *Corylus* (Betulaceae) [33]; hence the species is at the extreme end of polyphagy among butterflies, although by no means an indiscriminate generalist [9].

For this study, we used a test array with three host plants of *P. c-album* that are either closely related (Stinging nettle *Urtica dioica* and Wych Elm *Ulmus glabra* - both in Urticales) or share the same growth form (Great Sallow *Salix caprea* and *Ulmus glabra* - both trees). Following Feeny's "plant apparency" hypothesis and the phylogenetic patterns of host shifts found by Janz & Nylin [9] and Janz et al. [8], we expected to find more similarities in the gene expression profiles of caterpillars that have been feeding on plants that either have a shared ancestry (*U. dioica* and *U. glabra*) or belong to the same growth form (*S. caprea* and *U. glabra*).

## Results

### General patterns of gene expression

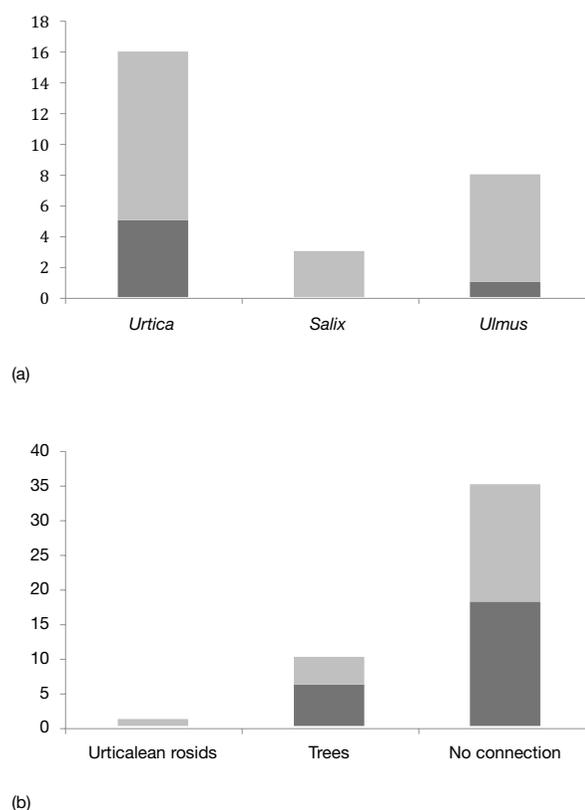
The pattern of expression in the midgut corresponded to the formed predictions that more similarities will be found in the gene expression profiles of caterpillars that have been feeding on plants that either have a shared ancestry (*U. dioica* and *U. glabra*) or belong to the same growth form (*S. caprea* and *U. glabra*) (Additional File 1; Table S1, Figure 1). There were significantly more uniquely upregulated genes in the larval midgut on the *Urtica* and *Salix* diets than on the *Ulmus* diet (likelihood-ratio  $\chi^2$  test, observed frequencies (*Urtica*, *Salix*, *Ulmus*) = 22, 16, 6 (expected frequencies equal),  $\chi^2 = 9.8990$ ,  $p = 0.007$ ). Pairwise similarities in upregulation were higher



**Figure 1**  
**General patterns of gene expression in the larval midgut of *Polygonia c-album*.** Dark grey represent qualitative differences or similarities (present/absent), light grey represent quantitative differences or similarities (more/less). **(a)** The number of uniquely upregulated genes on the three host plants used in the study, showing that significantly fewer genes were uniquely upregulated on *Ulmus*. **(b)** Pairwise similarities in upregulation between two of the three host plants. "Urticalean rosids" = *Urtica dioica* + *Ulmus glabra*; "Trees" = *Ulmus glabra* + *Salix caprea*; "No connection" = *Urtica dioica* + *Salix caprea*. Following predictions, more genes were concurrently upregulated on plants that shared either ancestry or growth form.

between plants that shared either growth form or were phylogenetically related (likelihood-ratio  $\chi^2$  test, observed frequencies (Urticalean rosids, trees, no connection) = 9, 6, 1 (expected frequencies equal),  $\chi^2 = 7.4839$ ,  $p = 0.024$ ).

There were no corresponding patterns in the restbodies (Additional File 1; Table S2, Figure 2). Here, more genes were uniquely upregulated on *Urtica* and *Ulmus* than on *Salix* (likelihood-ratio  $\chi^2$  test, observed frequencies (*Urtica*, *Salix*, *Ulmus*) = 16, 3, 8 (expected frequencies equal),  $\chi^2 = 9.9355$ ,  $p = 0.007$ ) and there were more sim-



**Figure 2**  
**General patterns of gene expression in the larval restbody of *Polygonia c-album*.** Dark grey represent qualitative differences or similarities (present/absent), light grey represent quantitative differences or similarities (more/less). The pattern seen in the restbody is distinctly different from the midgut. **(a)** The number of uniquely upregulated genes on the three host plants used in the study. **(b)** Pairwise similarities in upregulation between two of the three host plants. "Urticalean rosids" = *Urtica dioica* + *Ulmus glabra*; "Trees" = *Ulmus glabra* + *Salix caprea*; "No connection" = *Urtica dioica* + *Salix caprea*. Again, the pattern is different from the midgut; there are no effects of shared ancestry or growth form.

ilarities in upregulated genes between the "unrelated" plants than between the trees or urticalean rosid groups (likelihood-ratio  $\chi^2$  test, observed frequencies (Urticalean rosids, trees, no connection) = 1, 10, 35 (expected frequencies equal),  $\chi^2 = 43.7634$ ,  $p < 0.001$ ).

#### Function and patterns of differentially expressed genes

In total we identified 120 differentially expressed genes, 55 in the midgut of *P. c-album* and 65 in the restbody (Additional File 1; Table S1 and Table S2). In the midgut six sequences gave no hit and in the restbody 2 sequences gave no hit in BLAST searches. Many of the differentially expressed genes in the midgut are likely to be involved in

metabolism and digestion, ranging from protein degradation to starch and lipid breakdown for nutrient acquisition. We could also identify eight ribosomal genes and seven genes that are involved in translation regulation and the DNA structure in the cell nucleus. Furthermore, we found three genes involved in immunity and one transmembrane transport domain. There were also differentially expressed proteases, but we did not identify any differentially expressed known detoxifying enzymes in the midgut. Four of the differentially expressed genes were unknown (Additional File 1; Table S1).

In the restbody of the comma butterfly larvae again, the majority of the gene products are most likely involved in the metabolism of the caterpillar or translate into ribosomal proteins. Seven of the sequenced gene fragments appear to have a structural role, and are mostly involved in chitin binding. We could also identify translation regulation genes and some involved in cellular architecture (Additional File 1; Table S2). One sequence showed moderate similarities (e-value of 0,00000008) to potential detoxification genes, namely a cytochrome P450 of *Plutella xylostella* (Additional File 1; Table S2).

#### Confirmation of differentially expressed genes

Of the total 122 genes originating from GeneFishing, we picked 27 genes (18 from midguts and 9 from restbodies) to confirm differential gene expression pattern with qRT-PCR. We were able to see identical expression patterns in the GeneFishing and in the qRT-PCR results for 14 of the 27 genes (10 in midguts and 4 in restbodies). Partial similarity (same in relation to one diet) in expression patterns between two independent methods could be observed in 7 genes (4 in midguts and 3 in restbodies) and 5 genes (4 in midguts and 1 in restbodies) behaved differently (Additional File 1; Table S3).

In the midgut, genes that showed similar expression patterns as in the GeneFishing experiment included proteins involved in digestion, namely chymotrypsinogen-like protein 3, serine protease, chymotrypsin-like protease, alpha-amylase, trypsin-like protease, trypsin Ia precursor, long-chain acyl-CoA synthetase and short-chain dehydrogenase/reductase 2. In addition, also a ribosomal protein S16 and the immune response related protein cobatoxin showed same expression patterns in the qRT-PCR and in the differential gene expression study. In the restbodies alpha-amylase and the serine protease precursor were similarly expressed. The stress related heat shock cognate 70 protein and a potential cytochrome P450 also showed similar expression pattern in both methods.

#### Discussion

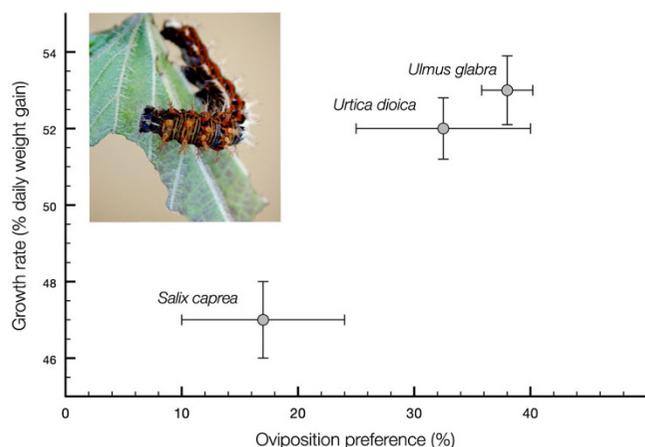
Our expression data suggest a complex interaction between the comma butterfly and its host plants. On the

one hand, each plant species appears to require a very specific subset of genes to be regulated in the midgut upon feeding. This involves digestion proteins, immunity related genes, general metabolism genes and ribosomal protein as well as translational and transport genes. In the restbody we also find cell signaling related domains, hormone synthesis genes and genes involved in silk production, suggesting a gene regulation, tuned specifically for each plant species. On the other hand, counting of upregulation of genes in the midgut shows a suggestive pattern (Figure 1). Both species of the Urticales family (*Ulmus* and *Urtica*), as well as both trees, have a higher agreement of gene regulation than do the stinging nettle and the great willow. Especially the digestion and ribosomal genes show clear differences here (Additional File 1; Table S1). This suggests that phylogenetic and/or growth form relatedness demand more similar expression profiles in the midgut of the caterpillars.

However, this pattern only holds true for the midgut of the comma butterfly; the restbody of the caterpillars show a completely different picture. Here the stinging nettle and great willow diet in contrast share the highest similarity in gene expression, while the wych elm appears to require less gene upregulation in general (Additional File 1; Table S2, Figure 2). The midgut is the place of contact with the food bolus, and the location where the first detoxifying and digestive actions will take place, whereas the restbody is only indirectly involved in this process by receiving the solubilized products and only sometimes toxic or toxin degradation products. Here the actual nutrient content is of more importance. This might explain the converse picture we observe in those two tissues. While the defense compounds of the phylogenetically or growth form-related plants might be more similar, our data suggests that the nutritional value does not follow this line.

Previous data on larval performance show that caterpillars of *P. c-album* perform differently on the plant species used in this study, suggesting different levels of adaptation and/or different nutritional values (e.g. [34]). The general pattern is that the plants in Urticales (*U. dioica* and *U. glabra*) support the highest growth rates, while larvae grow slower on *S. caprea* (but pupal weights are typically larger). These differences are also reflected in the preference hierarchy of the ovipositing females, which means that the preference-performance correlation is good on a population level [33,34] (Figure 3).

Oviposition preference and larval performance on *U. glabra* are more variable between years than the other plants, so that in some years it is considerably less preferred than *U. dioica*, with a correspondingly lower performance (Nylin, S. and Janz, N., unpublished data). This could be caused by higher between-year variability in *U. glabra*



**Figure 3**  
**Correlation between preference and performance in *Polytonia c-album*.** Correlation between relative oviposition preference (degree of preference in simultaneous choice trials) and larval performance (mean growth rate) for the plants used in the present study (means  $\pm$  SE). Data from [34].

compared to the other two hosts. Such unpredictable variation would make it more difficult to adapt to the host. Indeed, the lower amount of uniquely upregulated genes on *Ulmus* found in this study would suggest such insufficient adaptation to this host.

In the midgut we could not detect any differentially expressed known detoxifying genes, such as cytochromes P450s or glutathione-S-transferases (GST). However, those enzymes are with high probability active in the gut of the comma butterfly larvae when encountering plant material. The GeneFishing method identifies solely differentially expressed genes. Hence, *P. c-album* might very well express such Phase I and Phase II detoxifying genes, but in a constant manner independent of the diet. Being a highly polyphagous insect species, the comma butterfly might be endowed with very broad acting detoxifying enzymes that can cope with a wide variety of the compounds and hence are expressed continuously when feeding independent of the diet. The lack of differentially expressed detoxifying enzymes could also be due to overlaps of toxic chemistry on the three host plants.

With the exception of cytochrome P450s and GSTs not much is known about detoxifying enzymes applied by generalist insects. Hence such expressed enzymes could not be identified in databases. In the midgut we identified 4 unknown differentially expressed genes, which could very well be unknown genes involved in detoxification. We could, however, identify a differentially expressed gene homologous to a cytochrome P450 of *P. xylostella* in

the restbody of *P. c-album* larvae. Cytochrome P450's are also involved in the metabolism of many endogenous compounds, hence this P450 expressed in the restbody is possibly not involved in the detoxification of plant defense compounds, but in metabolism.

While the detoxifying mechanisms appear not to be differentially expressed we see genes belonging to other classes differentially expressed. Short-chain dehydrogenases (SCDH) for example are upregulated in the midguts of on *Urtica* and *Ulmus* feeding larvae. SCDH form a large protein family with highly different enzymes, which only share 15-30% identity among each other [35]. They are present in all the life forms studied so far, have a very wide substrate spectrum and are generally involved in cellular differentiation and signaling [36].

Not surprisingly, also many digestive enzymes are differentially expressed in the midgut (Additional File 1; Table S1). For example an alpha-amylase was down-regulated in both tree diets in comparison to the stinging nettle. Amylases are enzymes participating in carbohydrate digestion. It is known that insects possess different amylases for starches [37]. It is possible that the differential expression we observe is due to different starch contents in the stinging nettle. We also found several serine proteases being differentially expressed upon feeding on the host plants, namely trypsin and chymotrypsin. Plants possess proteinase inhibitors (PIs), which are insect inducible peptidases that can suppress insect proteinases and by that reduce the digestibility [38,39]. It has been shown that lepidopteran larvae adapt their proteinases expression profile to the PI content of their food plant, upregulating proteinases that are insensitive to the plant PIs [37]. Our expression patterns suggest different proteinase requirements for the three host plants. We also excised four bands that were identified as homologous to Bmlipase-1; three of them were highly expressed in larvae feeding on *Urtica* and one on *Salix*. Insect midgut lipases have been studied in few insects so far and little to nothing is known on different expression of lipases in insects [37]. They form in insects a gene family that underwent many duplication events [40] with resulting diverse and overlapping function. Bmlipase-1 from *Bombyx mori* shows a high antiviral activity against *Bombyx mori* nucleopolyhedrovirus, although it is not inducible by viral infection. Its main function is probably as a digestive enzyme, as it is exclusively expressed in the gut tissue and has lipase activity [41].

We also found many ribosomal proteins to be differentially regulated upon feeding on different host plants, namely seven in the midgut and twelve in the restbody of *P. c-album*. Ribosomal genes are considered to be stably expressed and have been suggested and used as house-

keeping genes in expression analysis. However, there are a number of ribosomal proteins that have been found to be differentially expressed upon different treatment and in different tissues [42], suggesting that there are major differences in expression patterns between different ribosomal proteins.

The complex pattern found in this study should perhaps not be surprising, considering that secondary metabolites, leaf texture and nutrient contents vary greatly between different plant species, and also depending on the plant age [43]. The three chosen *P. c-album* host plants have different leaf architecture, with the stinging nettle possessing thin hairy leaves, and both trees waxy leaves, that differ in shape. Besides the prominent stinging trichomes that only provide an effective defense against vertebrate herbivores [44], stinging nettle leaves also contain the phenolic compound caffeic acid, tannins, nicotine in measurable amounts and flavonoid glucosides [45,46]. *Salix* leaves has low levels of toxins but more condensed tannins [47], which is similar to the situation in elm [45,46].

Phenolic compounds, nicotine, tannins and flavonoids have been shown to affect insects feeding on them. Tannins for example are astringent, bitter plant polyphenols that either bind and precipitate or shrink proteins. Studies on the effect of tannins on the gut pH and redox potential of larvae found that many specialist and generalist insects have developed adaptations to cope with them [48,49]. Tannin metabolites are for example oxidized during the gut passage in the aquatic caterpillar *Acentria ephemerella*. Resistance against polyphenols appears to be correlated in general with better repair mechanisms in the gut tissue that enables to cope with free radical stress occurring during oxidation better [50]. Nicotine acts as agonist of the postsynaptic nicotinic acetylcholine receptors of the insect central nervous system. Metabolism of nicotine has been attributed to the action of cytochrome P450 [51,52]. Flavonoids in contrast can be sequestered from the diet and used for protection and pigmentation. In *B. mori* for example three flavonoid glycosides could be isolated from the cocoon shell [53]. The role of these secondary plant compounds in the interaction between *P. c-album* and its host plant and the adaptations *P. c-album* developed to them however, is not known.

Different plants also possess different microfloras [54], which affect immune response-related gene expression in the midgut of lepidopteran larvae [55]. We detected cobatoxin and gloverin homologous genes that are known to be inducible by bacterial challenge to be differentially expressed in the midgut tissue [56-58]. We see gloverin expression only in *Urtica* fed larvae and cobatoxin in *Salix* fed larvae, suggesting different microbial environments on those two species. In addition to differences in plant secondary metabolites *P. c-album* must therefore also face

different bacterial quantities and qualities on its various host plants.

## Conclusion

Our data suggest a complex picture of gene expression in response to host plants, but also revealed some general patterns. While each plant evidently requires a unique gene regulation in the caterpillar, both phylogenetic relatedness and host plant growth form appear to influence the expression profile of the polyphagous comma butterfly, in agreement with phylogenetic studies of host plant utilization in butterflies.

## Methods

### Larval rearing and preparation

Stock used in the experiments was the offspring of four female comma butterflies collected in early May 2007 in the near-Stockholm area of Sweden. The females had already mated in the wild with unknown males and were put into cages for oviposition. Each female was presented with the host plants stinging nettle (*U. dioica*) and Great Sallow (*S. caprea*). Eggs were counted in the beginning of each day and were incubated in small jars on a sun-lit windowsill until hatching.

Larvae of each female were evenly spread across the three host plants. They were raised on stinging nettle, great sallow or elm (*U. glabra*) in individual jars. Plant material was cut from mature foliage of naturally growing plants and placed in water culture. The jars were placed in a climate room (temperature 20°C, LD 12:12) where larvae were raised to the 4th instar before dissection of the midgut. Plants were changed when needed due to withering or feeding. To maintain humidity, water was sprayed over the jars twice a day. Jars were changed randomly to avoid position effects. Between 10 and 43 individuals from each family were dissected, for a total of 109 individuals across the three different diets. Midguts and the rest of the larval body were preserved separately in RNA later® (Ambion).

### RNA Isolation and Reverse Transcription

Larvae were dissected, and the midguts and restbodies were stored in RNA later® (Ambion). Tissue samples were pooled (10-13 individuals) according to the larval diet (*Salix*, *Urtica*, *Ulmus*). The guts were homogenized by an Ultra-Turrax homogenizer (Beckman Coulter Scientific) in TRIzol (Invitrogen) reagent and restbodies were crushed in liquid nitrogen. For all samples TRIzol Reagent was used to isolate the RNA according to the manufacturer's protocol with the following modifications. After adding chloroform to separate the phases, the tubes were stored for 15 minutes at 4°C before centrifugation. To precipitate the RNA the solution was stored at -20°C overnight. After precipitation the RNA solution was centrifuged for 30 min at 4°C. The obtained dried pellet was dissolved in 90 µl RNA storage solution (Ambion), and

any remaining genomic DNA contamination was removed by DNase treatment (TURBO DNase, Ambion). The DNase enzyme was removed and the RNA was further purified by using the RNeasy MinElute Clean up Kit (Qiagen) following the manufacturer's protocol and eluted in 20 µl of RNA storage solution (Ambion).

#### Differential gene expression

The GeneFishing (SeeGene) DEG Premix Kit essentially allows the amplification of the same set of genes from different samples due to a 10-mer core of arbitrary annealing control primers. By not exceeding the exponential phase of the PCR amplification, differentially expressed genes can be identified on an agarose gel.

To study differential gene expression between *P. c-album* larvae grown on different plants (*U. glabra*, *S. caprea* and *U. dioica*) the DEG GeneFishing Kit was used, following the manufacturer's protocol. In short 3 µg of DNA-free total RNA was converted into single-stranded cDNA using annealing control primer one (dTACP1) and a mixture of different reverse transcriptases (Array Script, Ambion; Power Script, Clontech; Bioscript, Bioline). Second-strand cDNA synthesis and subsequent PCRs were performed essentially as described in the DEG GeneFishing protocol. PCR products were separated and visualized on a 2% agarose gel. Differentially expressed bands were cut out from the agarose gels and PCR products extracted using Zymo-clean Gel DNA Recovery Kit™ (Zymo Research) according to the manufacturer's instructions. DNA fragments were cloned into the pCR II TOPO vector (Invitrogen). Eight clones were picked for each extracted band and further processed.

#### DNA Sequencing and Analysis

Plasmid minipreparations from bacterial colonies grown in 96 deep-well plates were performed using the 96 robot plasmid isolation kit (Eppendorf) on a Tecan Evo Freedom 150 robotic platform (Tecan). Single-pass sequencing of the 5' and 3' termini of individual clones was carried out on an ABI 3730 xl automatic DNA sequencer (PE Applied Biosystems).

Vector clipping, quality trimming and sequence assembly was done with the Lasergene software package (DNASar Inc.). BLAST searches were conducted on a local server using the National Center for Biotechnology Information (NCBI) blastall program and best hits were recorded. Identified sequences of this study were submitted to Genbank and can be found under Accession numbers [GT154770](#) - [GT155229](#).

When two independently identified differentially expressed sequences clustered in the assembly in the same contig, it was assumed to be a recent duplication.

#### Quantitative real-time PCR

500 ng of DNA-free total RNA was converted into single-stranded DNA using a mix of random and oligo-dT20 primers according to the ABgene protocol (ABgene). Real-time PCR oligonucleotide primers were designed using the online Primer3 internet based interface <http://frodo.wi.mit.edu>. Primers were designed by the rules of highest maximum efficiency and sensitivity rules were followed to avoid formation of self and hetero-dimers, hairpins and self-complementarity. Gene-specific primers were designed on the basis of sequences obtained for selected *P. c-album* genes and two additional genes as potential housekeeping genes (ribosomal protein subunit 18S and elongation initiation factor 4) to serve as the endogenous control (normalizer) [see Additional file 2].

Both house-keeping primers were thoroughly tested. RPS 18 was the most consistent gene, and this was then used for the further analysis. QRT-PCR was done in optical 96-well plates on a MX3000P Real-Time PCR Detection System (Stratagene) using the Absolute QPCR SYBR green Mix (ABgene) to monitor double-stranded DNA synthesis in combination with ROX as a passive reference dye included in the PCR master mix.

#### Scoring of expression similarities

Similarities in gene regulation were scored by counting upregulation (+/++) versus no expression (-). A semi-quantitative measure of upregulation was obtained by visually differentiating between strong (++) and weak (+) bands. The number of genes that were uniquely upregulated in any of the three plants, and the number of genes that were similarly upregulated in a pair of plants were then compared and analyzed with  $\chi^2$  goodness-of-fit tests (Figure 1 and 2).

#### Authors' contributions

DF and HH-F carried out the molecular laboratory work, with the help of HV, and performed the statistical analysis of the data with the help of NJ and SN. LS carried out the qRT-PCR experiments with the help of DF and HH-F. LS, SN and NJ set up the study and reared larvae. All authors contributed in the conception and design of the study. HH-F, DH and NJ participated in drafting the manuscript. All authors participated in the writing and approval of the final manuscript.

#### Additional material

##### Additional file 1

Supplemental Tables. Supplemental Tables S1 - S3.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-506-S1.doc>]

### Additional file 2

**qRT-PCR Primers.** List of primers used in the quantitative real-time PCR analysis.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-506-S2.pdf>]

### Acknowledgements

We thank D Schnabelrauch and H. Ringys-Beckstein for providing essential technical assistance for sequencing. This work was supported by the International Max Planck Research School (IMPRS) and the Max Planck Society and by research grants from the Swedish Research Council (VR) to NJ and SN.

### References

- Dethier VG: **Chemical factors determining the choice of food plants by *Papilio* larvae.** *Am Nat* 1941, **75**:61-73.
- Ehrlich PR, Raven PH: **Butterflies and plants: a study in coevolution.** *Evolution* 1964, **18**:586-608.
- Fraenkel GS: **The raison d'être of secondary plant substances.** *Science* 1959, **129**:1466-1470.
- Thorsteinson AJ: **Host selection in phytophagous insects.** *Annu Rev Entomol* 1960, **5**:193-218.
- Bernays EA: **Host range in phytophagous insects - the potential role of generalist predators.** *Evol Ecol* 1989, **3**:299-311.
- Becerra JX: **Insects on plants: macroevolutionary chemical trends in host use.** *Science* 1997, **276**:253-256.
- Futuyma DJ, Keese MC, Scheffer SJ: **Genetic constraints and the phylogeny of insect-plant associations - responses of *Ophraella communa* (Coleoptera, Chrysomelidae) to host plants of its congeners.** *Evolution* 1993, **47**:888-905.
- Janz N, Nyblom K, Nylin S: **Evolutionary dynamics of host-plant specialization: a case study of the tribe Nymphalini.** *Evolution* 2001, **55**:783-796.
- Janz N, Nylin S: **Butterflies and plants: a phylogenetic study.** *Evolution* 1998, **52**:486-502.
- Kergoat GJ, Delobel A, Fediere G, Le Ru B, Silvain JF: **Both host-plant phylogeny and chemistry have shaped the African seed-beetle radiation.** *Mol Phyl Evol* 2005, **35**:602-611.
- Lopez-Vaamonde C, Charles H, Godfray J, Cook JM: **Evolutionary dynamics of host-plant use in a genus of leaf-mining moths.** *Evolution* 2003, **57**:1804-1821.
- Menken SBJ: **Pattern and process in the evolution of insect-plant associations: Yponomeuta as an example.** *Entomol Exp Appl* 1996, **80**:297-305.
- Murphy SM, Feeny P: **Chemical facilitation of a naturally occurring host shift by *Papilio machaon* butterflies (Papilionidae).** *Ecological Monographs* 2006, **76**:399-414.
- Ronquist F, Liljeblad J: **Evolution of the gall wasp-host plant association.** *Evolution* 2001, **55**:2503-2522.
- Ohshima I, Yoshizawa K: **Multiple host shifts between distantly related plants, Juglandaceae and Ericaceae, in the leaf-mining moth *Acrocercops leucophaea* complex (Lepidoptera: Gracillariidae).** *Mol Phyl Evol* 2006, **38**:231-240.
- Wahlberg N: **The phylogenetics and biochemistry of host-plant specialization in Melitaeine butterflies (Lepidoptera: Nymphalidae).** *Evolution* 2001, **55**:522-537.
- Feeny P: **Plant apparency and chemical defence.** In *Recent Advances in Phytochemistry: Biological Interactions Between Plants and Insects Volume 10*. Edited by: Wallace JW, Mansell RL. Plenum Press, New York; 1976:1-40.
- Chew FS, Courtney SP: **Plant apparency and evolutionary escape from insect herbivory.** *Am Nat* 1991, **138**:729-750.
- Futuyma DJ: **Food plant specialization and environmental predictability in Lepidoptera.** *Am Nat* 1976, **110**:285-292.
- Miller AM, McArthur C, Smethurst PJ: **Effects of within-patch characteristics on the vulnerability of a plant to herbivory.** *Oikos* 2007, **116**:41-52.
- Wasserman SS: **Allelochemic diversity and plant apparency - evidence from the detoxification systems of caterpillars.** *American Midland Naturalist* 1979, **102**:401-403.
- Agrawal AA, Lau JA, Hamback PA: **Community heterogeneity and the evolution of interactions between plants and insect herbivores.** *Q Rev Biol* 2006, **81**: 349-376-EF.
- Berenbaum MR, Favret C, Schuler MA: **On defining "key innovations" in an adaptive radiation: Cytochrome P450s and papilionidae.** *Am Nat* 1996, **148**:S139-S155.
- Berenbaum MR, Feeny P: **Chemical mediation of host-plant specialization: The papilionid paradigm.** In *Specialization, Speciation, and Radiation: the Evolutionary Biology of Herbivorous Insects* Edited by: Tilmon KJ. Berkeley, California: University of California Press; 2008:3-19.
- Li XC, Baudry J, Berenbaum MR, Schuler MA: **Structural and functional divergence of insect CYP6B proteins: From specialist to generalist cytochrome P450.** *Proc Natl Acad Sci USA* 2004, **101**:2939-2944.
- Mao WF, Berhow MA, Zangerl AR, McGovern J, Berenbaum MR: **Cytochrome P450-mediated metabolism of xanthotoxin by *Papilio multicaudatus*.** *J Chem Ecol* 2006, **32**:523-536.
- Zangerl AR, Berenbaum MR: **Phenotype matching in wild parsnip and parsnip webworms: causes and consequences.** *Evolution* 2003, **57**:806-815.
- Wadleigh RW, Yu SJ: **Detoxification of Isothiocyanate allelochemicals by Glutathione transferase in 3 lepidopterous species.** *J Chem Ecol* 1988, **14**:1279-1288.
- Yu SJ: **Host plant induction of glutathione S-transferase in the fall armyworm.** *Pesticide Biochemistry and Physiology* 1982, **18**:101-106.
- Wittstock U, Agerbirk N, Stauber EJ, Olsen CE, Hippler M, Mitchell-Olds T, Gershenzon J, Vogel H: **Successful herbivore attack due to metabolic diversion of a plant chemical defense.** *Proc Natl Acad Sci USA* 2004, **101**:4859-4864.
- Fischer HM, Wheat CW, Heckel DG, Vogel H: **Evolutionary origins of a novel host plant detoxification gene in butterflies.** *Molecular Biology and Evolution* 2008, **25**:809-820.
- Wheat CW, Vogel H, Wittstock U, Braby MF, Underwood D, Mitchell-Olds T: **The genetic basis of a plant-insect coevolutionary key innovation.** *Proc Natl Acad Sci USA* 2007, **104**:20427-20431.
- Nylin S: **Host plant specialization and seasonality in a polyphagous butterfly, *Polygonia c-album* (Nymphalidae).** *Oikos* 1988, **53**:381-386.
- Janz N, Nylin S, Wedell N: **Host-plant utilization in the comma butterfly - sources of variation and evolutionary implications.** *Oecologia* 1994, **99**:132-140.
- Jornvall H, Persson B, Krook M, Atrian S, Gonzalez-Duarte R, Jeffery J, Ghosh D: **Short-Chain Dehydrogenase/Reductases (SDR).** *Biochemistry* 1995, **34**:6013.
- Kallberg Y, Oppermann U, Jornvall H, Persson B: **Short-chain dehydrogenase/reductase (SDR) relationships: A large family with eight clusters common to human, animal, and plant genomes.** *Protein Science* 2002, **11**:636-641.
- Terra WR, Ferreira AH: **Biochemistry of digestion.** In *Comprehensive Molecular Insect Science 4* Edited by: Gilbert LI, Iatrou K, Gill SS. Oxford: Elsevier; 2005:171-224.
- Steppuhn A, Baldwin IT: **Resistance management in a native plant: nicotine prevents herbivores from compensating for plant protease inhibitors.** *Ecol Lett* 2007, **10**:499-511.
- Zavala JA, Patankar AG, Gase K, Hui DQ, Baldwin IT: **Manipulation of endogenous trypsin proteinase inhibitor production in *Nicotiana attenuata* demonstrates their function as antiherbivore defenses.** *Plant Physiology* 2004, **134**:1181-1190.
- Horne I, Haritos VS: **Multiple tandem gene duplications in a neutral lipase gene cluster in *Drosophila*.** *Gene* 2008, **411**:27-37.
- Ponnuvel KM, Nakazawa H, Furukawa S, Asaoka A, Ishibashi J, Tanaka H, Yamakawa M: **A lipase isolated from the silkworm *Bombyx mori* shows antiviral activity against nucleopolyhedrovirus.** *Journal of Virology* 2003, **77**:10725-10729.
- Thorrez L, Van Deun K, Tranchevent LC, Van Lommel L, Engelen K, Marchal K, Moreau Y, Van Mechelen I, Schuit F: **Using ribosomal protein genes as reference: a tale of caution.** *PLoS ONE* 2008, **3**:e1854.
- Brenes-Arguedas T, Horton MW, Coley PD, Lokvam J, Waddell RA, Meizoso-O'Meara BE, Kursar TA: **Contrasting mechanisms of**

- secondary metabolite accumulation during leaf development in two tropical tree species with different leaf expansion strategies. *Oecologia* 2006, **149**:91-100.
44. Pullin AS: **Life history strategies of the butterflies, *Inachis io* and *Aglais urticae*, feeding on nettle, *Urtica dioica*.** In *PhD thesis* Oxford Polytechnic; 1986.
  45. Hegnauer R: **Chemotaxonomie der Pflanzen: eine Übersicht über die Verbreitung und die systematische Bedeutung der Pflanzenstoffe.** In *Bd 6, Dicotyledoneae: Rafflesiaceae-Zygophyllaceae* Basel: Birkhäuser; 1973.
  46. Hegnauer R: **Chemotaxonomie der Pflanzen: eine Übersicht über die Verbreitung und die systematische Bedeutung der Pflanzenstoffe.** In *Bd 9, Nachträge zu Band 5 und Band 6 (Magnoliaceae bis Zygophyllaceae)* Basel: Birkhäuser; 1990.
  47. Hallgren P, Ikonen A, Hjalten J, Roininen H: **Inheritance patterns of phenolics in F1, F2, and back-cross hybrids of willows: implications for herbivore responses to hybrid plants.** *J Chem Ecol* 2003, **29**:1143-1158.
  48. Johnson KS: **Plant phenolics behave as radical scavengers in the context of insect (*Manduca sexta*) hemolymph and midgut fluid.** *Journal of Agricultural and Food Chemistry* 2005, **53**:10120-10126.
  49. Johnson KS, Felton GW: **Physiological and dietary influences on midgut redox conditions in generalist lepidopteran larvae.** *Journal of Insect Physiology* 1996, **42**:191-198.
  50. Gross EM, Brune A, Walenciak O: **Gut pH, redox conditions and oxygen levels in an aquatic caterpillar: Potential effects on the fate of ingested tannins.** *Journal of Insect Physiology* 2008, **54**:462-471.
  51. Dowd PF, Smith CM, Sparks TC: **Detoxification of plant toxins by insects.** *Insect Biochemistry* 1983, **13**:453-468.
  52. Nauen R, Denholm I: **Resistance of insect pests to neonicotinoid insecticides: Current status and future prospects.** *Archives of Insect Biochemistry and Physiology* 2005, **58**:200-215.
  53. Hirayama C, Ono H, Tamura Y, Konno K, Nakamura M: **Regioselective formation of quercetin 5-O-glucoside from orally administered quercetin in the silkworm, *Bombyx mori*.** *Phytochemistry* 2008, **69**:1141-1149.
  54. Meyling NV, Eilenberg J: **Isolation and characterisation of *Beauveria bassiana* isolates from phylloplanes of hedgerow vegetation.** *Mycological Research* 2006, **110**:188-195.
  55. Freitag D, Wheat CW, Heckel DG, Vogel H: **Immune system responses and fitness costs associated with consumption of bacteria in larvae of *Trichoplusia ni*.** *BMC Biol* 2007, **5**:56.
  56. Lundstrom A, Liu G, Kang D, Berzins K, Steiner H: ***Trichoplusia ni* gloverin, an inducible immune gene encoding an antibacterial insect protein.** *Insect Biochem Mol Biol* 2002, **32**:795-801.
  57. Seitz V, Clermont A, Wedde M, Hummel M, Vilcinskas A, Schlatterer K, Podsiadlowski L: **Identification of immunorelevant genes from greater wax moth (*Galleria mellonella*) by a subtractive hybridization approach.** *Dev Comp Immunol* 2003, **27**:207-215.
  58. Volkoff AN, Rocher J, d'Alencon E, Bouton M, Landais I, Quesada-Moraga E, Vey A, Fournier P, Mita K, Devauchelle G: **Characterization and transcriptional profiles of three *Spodoptera frugiperda* genes encoding cysteine-rich peptides. A new class of defensin-like genes from lepidopteran insects?** *Gene* 2003, **319**:43-53.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

