Expression analysis and phylogeny studies based on *de novo* assembled transcript libraries of leaf beetle larvae

Dissertation

zur Erlangung des akademischen Grades
„doctor rerum naturalium“ (Dr. rer. nat.)

vorgelegt dem Rat der Biologisch-Pharmazeutischen Fakultät
der Friedrich-Schiller-Universität Jena

von Dipl.-Bioinf. Magdalena Stock
geboren am 13.08.1980 in Jena

1. Gutachter: Prof. Dr. Wilhelm Boland (MPI für chemische Ökologie, Jena)
2. Gutachter: Prof. Dr. Reinhard Guthke (Hans-Knöll-Institut, Jena)
3. Gutachter: Prof. Dr. Karl Schmid (Universität Hohenheim)

Tag der öffentlichen Disputation: 11. Juni 2014

Für meine Familie, meine Eltern
und meine Schwestern.

Conducting data analysis is like drinking a fine wine.
It is important to swirl and sniff the wine, to unpack
the complex bouquet and to appreciate the experience.
Gulping the wine doesn't work.

[Wright, D. B. (2003). Making friends with your data:
Improving how statistics are conducted and reported.
Br J Educ Psychol., 73(1), 123-136.]

CONTENTS

1 INTRODUCTION

In the last ten years, next-generation sequencing (NGS) technologies have had a huge impact on (functional) genomics and transcriptomics. These methods are rapidly replacing the conventional Sanger strategy (Sanger et al., 1977) that has been the principal method of sequencing DNA since the late 1970s (Horner et al., 2010).

Not only NGS conducted recent genomic and transcriptomic research, but also RNA interference (RNAi) became the method of choice to study genes of interest. RNAi allows silencing of genes without making knockout strains, and the phenotype associated with the gene function can be denoted (March and Bentley, 2007, Huvenne and Smagghe, 2010).

The combination of these two methods provides a great opportunity to study non-model organisms and to gain new knowledge and insights into yet unknown genes, enzyme functions, metabolic pathways, signal transduction and the regulation of gene expression (Nakasugi et al., 2013). In this present dissertation, this combination of methods and methods to infer phylogenetic trees were used to investigate the transcriptomes of two different leaf beetle species with special focus on transport proteins that are involved in the chemical defense during the larval stage.

1.1 NEXT-GENERATION SEQUENCING TECHNOLOGIES

The continuous improvement in sequencing resulted both in an immense increase in number and length of reads and in a consequent reduction in cost per base sequenced. The Roche 454 technology (Droege and Hill, 2008) already provides a realistic substitute for many applications of conventional Sanger sequencing at greatly reduced cost, while the Illumina Genome Analyzer (Bennett, 2004) and ABI SOLiD (Porreca et al., 2006) platforms generate an many more reads of (relatively) reduced length. The Illumina Genome Analyzer and Roche 454 platforms use novel techniques for amplification and sequencing template molecules, though they still share the underlying principle of 'sequencing by extension' used in the Sanger methodology (Ansorge, 2009).

The systems described above require the emulsion PCR amplification step of DNA fragments to make the light signal strong enough for reliable base detection. A minimum use of biochemicals and no amplification bias would be achieved if the sequence could be determined directly from a single DNA molecule without the need for PCR amplification. This requires a very sensitive light detection system for identifying light from a single dye molecule (Ansorge, 2009). One of the first techniques for sequencing from a single DNA molecule was described by the

team of S. Quake (Braslavsky et al., 2003) and licensed by Helicos Biosciences. Other innovative DNA sequencing techniques are, on the one hand, the single-molecule real-time (SMRT) technology which was announced recently by Pacific Biosciences (http://www.pacific-biosciences.com/index.php). On the other hand, another single-molecule sequencing technique derives from the study of the translocation of DNA through various artificial nanopores. The modulation of the ionic current of the DNA is detected, revealing characteristics (diameter, length and conformation) of the molecule (Trepagnier et al., 2007).

In conclusion, the availability of optical instruments capable of reliably detecting millions of sources of light or fluorescence on the surface of small glass slides facilitated the development of the new massively parallel sequencing technologies. In addition to the conventional objectives of genome (re-)sequencing and the discovery of single nucleotide polymorphisms, these technologies can be efficiently applied to a number of other scopes, including comprehensive studies of transcriptomes, facilitation of gene annotation and identification of splice variants and novel mutations (Tucker et al., 2009, Ansorge, 2009, Horner et al., 2010, Mardis, 2008, Mardis, 2012, Koboldt et al., 2013). Furthermore, the immediate applications and relevance of NGS techniques in the medical field have been demonstrated already, for example by the ability to detect cancer alleles with sequencing genomic DNA in cancerous tissues (Ansorge, 2009, Tucker et al., 2009, Wheeler et al., 2008, Levy et al., 2007, Mardis, 2012). Another approach of NGS is ChIP-sequencing. The chromatin immunoprecipitation (ChIP) refers to the isolation of genomic fragments bound to proteins through the use of crosslinking agents. Furthermore, specific antibodies are used to identify genomic regions bound to histones. ChIP sequencing focuses also on the genome-wide mapping of many transcription factors (TF) and their corresponding DNA-binding sites (Cullum et al., 2011). To conclude, this technology is becoming the method of choice for large-scale identification of TF-DNA interactions and of the characterization of chromatin packaging.

1.1.1 RNA SEQUENCING AND TRANSCRIPTOME ASSEMBLY

NGS technologies enable the analysis of complex samples containing a mixture of a large number of nucleic acids. This is achieved by sequencing simultaneously the entire sample content and allows the detection of low abundance RNAs, small RNAs including siRNAs and ncRNAs, or the presence of rare cells contained in the sample (Marguerat and Bahler, 2010, Wang et al., 2009). Recently, NGS techniques have been applied to transcriptome profiling (Cloonan et al., 2008, Sultan et al., 2008) and RNA-seq studies (Mortazavi et al., 2008, Trapnell et al., 2010, Costa et al., 2010, Simon et al., 2009).

The sequencing of RNA, which is RNA-seq, and the *de novo* assembly of RNA-seq data enable researchers to study transcriptomes without the need for a genome sequence. This approach is usefully applied, for instance, in the research of 'non-model organisms' of ecological, economic and evolutionary relevance (Petzold et al., 2013, Nakasugi et al., 2013, Liang et al., 2013, Crawford et al., 2010, Mittapalli et al., 2010, Niu et al., 2012, Su et al., 2012, Wang et al., 2010, Gremski et al., 2010, Keeling et al., 2012, Novaes et al., 2008).

Computational methods that can assemble a transcriptome are required when a genome sequence is not available. Two basic methods exist to generate transcript sequences from raw RNA-seq data: through the guidance of genomic sequences assembled beforehand or *via de novo* assembly. Several tools are available for the *de novo* assembly of RNA-seq data (Gongora-Castillo and Buell, 2013, Yang and Smith, 2013). Trans-ABySS (Robertson et al., 2010), Velvet-Oases (Schulz et al., 2012) and SOAPdenovo-trans (http://soap.genomics.org.cn/SOAPdenovo-Trans.html) are all extensions of earlier developed genome assemblers. A novel alternative method for transcriptome assembly is Trinity (Haas et al., 2013). Trinity partitions RNA-seq data into many independent de Bruijn graphs (ideally one graph per expressed gene), and uses parallel computing to reconstruct transcripts from these graphs, including alternatively spliced isoforms (Grabherr et al., 2011, Haas et al., 2013, Zhao et al., 2011). To conclude, correct transcriptome assembly tools should produce one full-length contig per distinct transcript (isoform) and not per locus, and different transcripts should have different coverage, reflecting their different expression levels (Haas et al., 2013, Grabherr et al., 2011). The Trinity software was the method I chose to apply for two non-model organisms, namely the mustard leaf beetle *Phaedon cochleariae* and the poplar leaf beetle *Chrysomela populi* (Manuscripts 2, 3 and 4).

1.1.2 Analysis of RNA abundance levels

The analysis of RNA abundance levels and the determination of differentially expressed genes is an application well-suited to NGS technologies. Sequencing-based measurements of gene expression have the advantage of generating novel cDNA sequence data across the length of the transcripts, including identification of sequence and splice site variants (Simon et al., 2009, Wall et al., 2009, Horner et al., 2010) in contrast to microarrays which are restricted to distinct sequences.

For RNA-seq data, one strategy taken is to count the number of reads that fall into annotated genes and to perform statistical analysis on the table of counts to discover quantitative changes in expression levels between experimental groups (Anders and Huber, 2010, Anders et al., 2013). These experimental groups, in the case of Manuscripts 1 to 3, are

tissue samples that derived from leaf beetle larvae that have been either treated with interfering RNA or not treated.

Figure 1 displays the computational pipeline underlying RNA-seq experiments. First, reads are mapped to the genome or transcriptome. Second, mapped reads for each sample are 'assembled' into gene-level, exon-level or transcript-level expression summaries (depending on the aim of the experiment). Next, these count data are normalized together with the statistical testing of differential expression, leading to a ranked list of genes with associated *P*-values and fold-changes. Normalization methods are the subject of the discussion chapter 4.1.2. Finally, biological insight from this list can be gained by taking system biology approaches, such as determining gene ontology (GO) terms and associated pathways (Oshlack et al., 2010). Since, in our case, there was no genomic data available, the reference genome is substituted by *de novo* assembled transcriptomes (Costa et al., 2010).

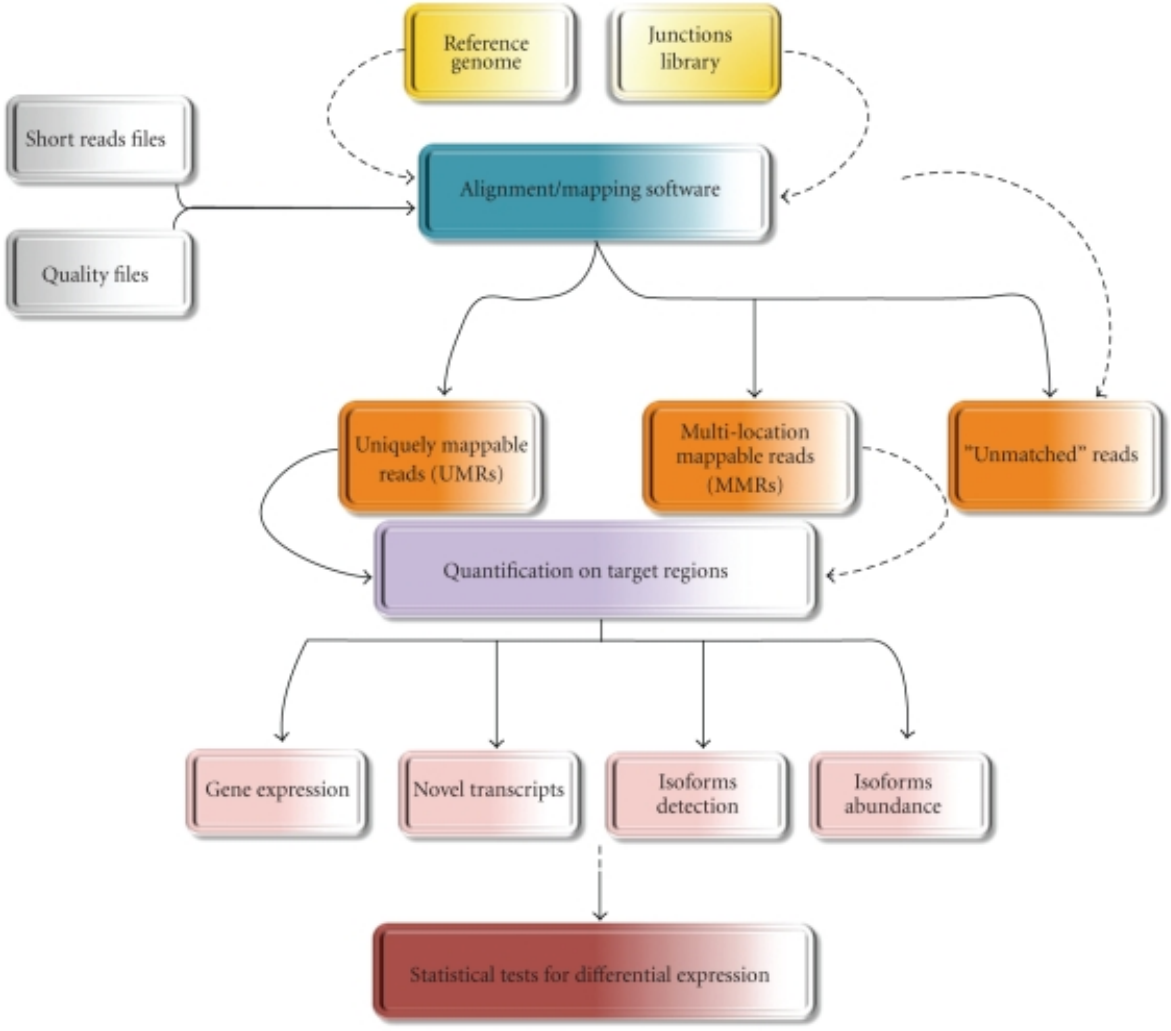


Figure 1: Computational pipeline for RNA-seq studies (adapted from (Costa et al., 2010)).

1.1.3 Normalization methods and differential expression Analysis

There are several approaches to estimating transcript-level expression.

On the one hand, FPKM is used to measure transcript abundances and is defined as the expected fragments per kilobase of transcript per million fragments mapped (Trapnell et al., 2010). On the other hand, the "RPKM prevalence information on a per-locus basis" can be considered. RPKM is the measure for transcript levels in reads per kilobase of exon model per million mapped reads which takes the sequence length into account and enables the comparison of transcript levels both within and between samples (Mortazavi et al., 2008).

In addition, there are methods that adjust the samples based on their transcript distributions among samples. Marioni *et al.* took into account the total number of reads (Marioni et al., 2008), and Bullard *et al*. proposed an upper-quartile normalization procedure which scales the expression level at the 75th percentile in each sample to the average across all samples (Bullard et al., 2010, Glusman et al., 2013). Robinson *et al*. applied a quantile-adjusted estimator for the negative binomial distribution (Robinson and Smyth, 2008) which has great success in very small samples. Two years later in 2010, Robinson *et al*. developed the trimmed mean of M-values (TMM) normalization method (Robinson and Oshlack, 2010) which is implemented in the edgeR Bioconductor package. Srivastava *et al*. proposed a two-parameter generalized Poisson model which is intended to "fit the position-level read counts more appropriately than a traditional Poisson-model" (Srivastava and Chen, 2010). They also observed that the standard RPKM can bias estimates of differential expression (Bullard et al., 2010). Glusman *et al.* developed a data-driven algorithm that computes sample-specific scaling factors. These factors lead to correct results only in the context of their comparison to other samples which means that, depending on the other samples, the examined sample it is compared to is scaled differently (Glusman et al., 2013). With this method they hope to provide accurate expression levels.

The counting approach by Anders *et al.* is "direct, flexible and can be used for many types of count data" (Anders et al., 2013) and was used to analyze the data upon which this thesis is based. This approach is implemented in the DESeq Bioconductor package and its application straightforward.

1.2 RNA INTERFERENCE (RNAi)

In general, RNAi occurs naturally but in recent years this method was established in laboratory research to study genes of interest by their knock-down. Hence, the changed phenotype can be observed, and the function associated with the gene can be assigned (March and Bentley, 2007).

RNAi can be divided in cell-autonomous and non-cell-autonomous RNAi. As the name suggests, in the case of cell-autonomous RNAi the silencing process is limited to the cell in which the dsRNA is introduced. This introduced dsRNA is then cleaved by a RNase III, called Dicer, into 21-25 nt-long siRNA duplexes. These are then incorporated into the RNA-induced silencing complex (RISC). Thereafter, the passenger strand is discarded and the RISC binds to a homologous mRNA, cleaves it and thereby hampers translation (Huvenne and Smagghe, 2010, Aigner, 2007).

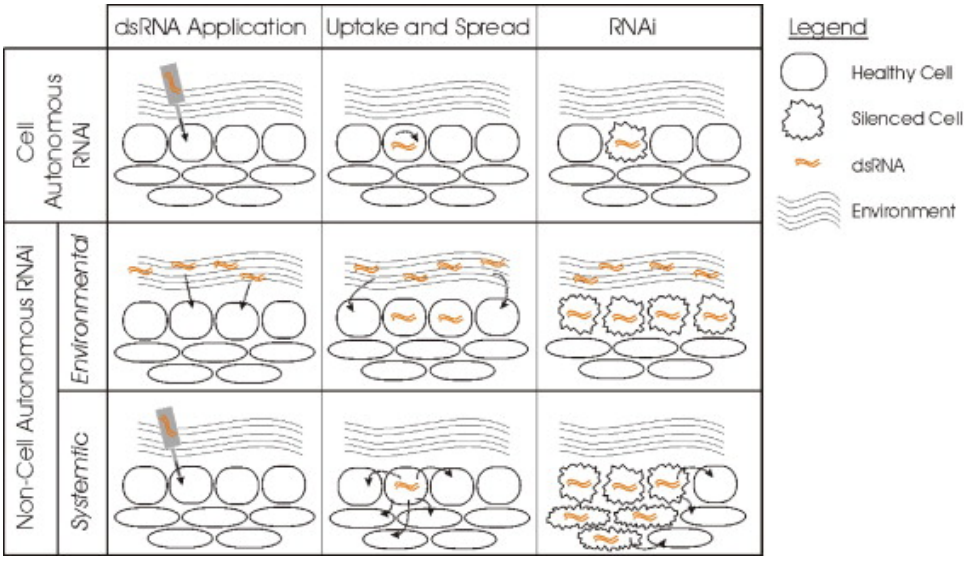


Figure 2: Cell-autonomous *versus* non-cell-autonomous RNAi (adapted from (Huvenne and Smagghe, 2010)).

In case of non-cell-autonomous RNAi, the interfering effect takes place in tissues/cells different from the location of the application of the dsRNA. There are two different kinds of non-cell-autonomous RNAi: environmental RNAi and systemic RNAi. Systemic RNAi can only take place in multicellular organisms. The target expression is silenced in the cell the dsRNA was exposed to. Thereafter, a silencing signal is transported from one cell to another or from one tissue type to another (Huvenne and Smagghe, 2010, Whangbo and Hunter, 2008). Systemic RNAi can be induced artificially by injection of dsRNA into the insect. The process of environmental RNAi includes the uptake of dsRNA from the intestinal lumen. In the intestinal

cells, the interfering RNA leads to a silencing of the target transcript. The dsRNA and dsRNA-derived silencing signals are then spread to cells throughout the animal by systemic RNAi. Finally, the silencing occurs also in distant cells (Whangbo and Hunter, 2008). Environmental RNAi – e.g. when feeding dsRNA by providing transgenic plants or, in the case of *Caenorhabditis elegans*, by soaking in dsRNA – is a more attractive approach than injection since it is non-invasive (Burand and Hunter, 2013).

1.2.1 RNAi in insects

Successful RNAi experiments by injection of dsRNA have been carried out in a number of lepidopteran and coleopteran species, mainly using *Tribolium castaneum* (Huvenne and Smagghe, 2010, Tomoyasu et al., 2008), *Bombyx mori* and *Manduca sexta*. Feeding of dsRNA has been applied with greatest effect in *Plutella xylostella*, *Spodoptera exigua*, *M. sexta* and *Ostrinia nubilalis* (Terenius et al., 2011). However, some organisms, such as the silkworm moth *Bombyx mori* and *Drosophila,* lack a robust systemic RNAi. Since *Tribolium* has the ability to respond to dsRNA systemically, it is an ideal model system for studying this process in insects (Tomoyasu et al., 2008).

Another aspect of RNAi, besides the functional annotation of specific enyzmes, is to control insect pest and disease prevention in insects. This includes research of the tsetse flies which promote Trypanosome species causing the African sleeping sickness and the cattle disease Nagana (Burand and Hunter, 2013). Hunter *et al.* applied RNAi in honey bees to suppress a virus infection (Hunter et al., 2010) and Abd-Alla *et al.* interfered a sterilizing virus in tsetse flies (Abd-Alla et al., 2011). Baum *et al.* triggered RNAi in several coleopteran species by feeding them an artificial diet (Baum et al., 2007). Their method of feeding the western corn rootworm (WCR) *Diabrotica virgifera virgifera* on transgenic corn plants led to a reduced WCR feeding damage.

1.2.2 Off-target prediction

In experiments based on the RNAi method, off-target effects can be induced by unintended cross-hybridization between small interfering RNA molecules (siRNAs) and endogenous RNA sequences other than the targeted sequences (Seinen et al., 2011, Kulkarni et al., 2006, Moffat et al., 2007, Jackson et al., 2003, Jackson et al., 2006).

Off-targets obscure the functional interpretation of gene silencing experiments and should therefore be avoided. Thus, the most important issue for siRNA selection is the validation that the designed dsRNA is specific to only the target mRNA (Cullen, 2006).

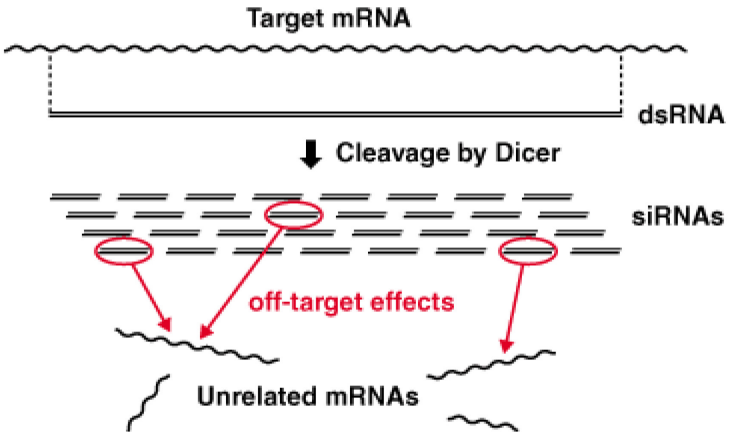


Figure 3: The dsRNA of the target sequence is injected or transferred into the organism of interest. Thereafter, it is cleaved into small interfering RNAs. These fragments can by chance be homologous to partial sequences of unrelated non-target mRNAs and cause off-target effects by diminishing those non-targeted mRNAs (adapted from (Naito et al., 2005)).

Sensitive tools were specifically designed to predict possible RNAi off-targets (Iyer et al., 2007, Naito et al., 2005), and others to enable the design of effective RNAi constructs (Chaudhary et al., 2011, Naito et al., 2009). It has been demonstrated that 79% similarity (15/19) between a siRNA sequence and mRNA with few as 11 complementary contiguous nucleotides is sufficient to trigger the silencing of non-targeted transcripts (Jackson et al., 2003). Additionally, many non-target transcripts that were silenced by siRNAs showed 3' UTR sequence complementarity to the seed region of the siRNA (Jackson et al., 2006). In summary, there are many issues that cumulate when exploring off-targets. Many off-target prediction tools are based on the genome sequences of studied organisms. Nevertheless, off-target prediction for non-model organisms with only transcript catalogues available should also be manageable, although no homology search can be conducted *versus* ncRNA or other small RNA which was not sequenced beforehand.

As published in Manuscript 1, I established an off-target prediction method which was successfully applied, not only in Manuscript 1, but also in the other embedded Manuscripts 2 to 4. This off-target prediction method first splits the target sequence into all possible 21 bp fragments (forward and reverse strand). Second, all those fragments are then compared to the *de novo* assembled transcriptomes of *Phaedon cochleariae* as well as *Chrysomela populi*.

1.3 PHYTOPHAGOUS LEAF BEETLES

The Chrysomelidae, comprising all leaf beetles, with about 50,000 species (Reid, 1995, Santiago-Blay, 1994) distributed all over the world, is the model taxon which is the subject of this thesis. Leaf beetles spend their whole life cycle on their host plant. They have to protect themselves against natural enemies, especially in early developmental stages, in order to reach fertility and to produce offspring to ensure the evolutionary survival of the species. Particularly, larvae of the subtribe Chrysomelina possess sophisticated strategies in terms of chemical defense (Termonia et al., 2001). In order to emit deterrent substances, the larvae developed specialized pair-wise exocrine glands on their dorsal thoracal and abdominal segments. In these glands, gland cells are attached to an impermeable reservoir which stores the deterrent compound. In case of predatory attack, the reservoirs release droplets of secretions containing the defensive compounds (Figure 4) (Hinton, 1951, Renner, 1970, Pasteels et al., 1990).

Figure 4: The left picture: *Phaedon cochleariae* (mustard leaf beetle) larvae with upended gland reservoirs containing defensive secretions. Right picture: *Chrysomela populi* (poplar leaf beetle) larvae dropping defensive secretions (left picture © Sindy Frick/Raimund Nagel, MPI for Chemical Ecology, Jena; left picture © Antje Burse, MPI for Chemical Ecology, Jena).

1.3.1 TRANSPORT PROCESSES

The source of the deterrent compound in the larval secretions depends on different synthetic strategies. *Chrysomela* species as well as *Phratora vitellinae,* for instance, developed an energy-saving but host plant-dependent defense mechanism. Those species sequester the phenolic glucoside salicin and convert it to the biologically active salicylaldehyde in the glandular reservoir (Kuhn et al., 2004, Discher et al., 2009, Pasteels et al., 1983a, Pasteels et al., 1983b, Rowell-Rahier and Pasteels, 1986). A second approach evolved in the so-called *interrupta* group of the Chrysomelina species (e.g. *Chrysomela lapponica*). They developed a mixed-mode

mechanism with *de novo* produced acids that are esterified in the defensive glands with leaf alcohols taken up as glycosides by sequestration (Schulz et al., 1997, Kuhn et al., 2007, Blum et al., 1972, Hilker and Schulz, 1994, Termonia and Pasteels, 1999, Tolzin-Banasch et al., 2011). The third approach in the Chrysomelina is the *de novo* production of iridoids (cyclopentanoid monoterpenes) (Soe et al., 2004, Meinwald et al., 1977, Blum et al., 1978, Pasteels et al., 1982, Burse et al., 2007, Lorenz et al., 1993, Burse et al., 2009). For example, the larvae of the leaf beetle *P. cochleariae* synthesize the iridoid chrysomelidial (Pasteels et al., 1982, Frick et al., 2013, Termonia et al., 2001). A natural advantage of the *de novo* group is their ability to defend themselves independently of their host plant-derived metabolites. They do not need to sequester compounds but do produce the glucosidically bound precursor *de novo* in the fat body tissue. Interestingly, feeding experiments elucidated that *de novo* producers already possess the ability to sequester (Feld et al., 2001, Kuhn et al., 2004, Kuhn et al., 2007, Soe et al., 2004, Discher et al., 2009, Kunert et al., 2008).

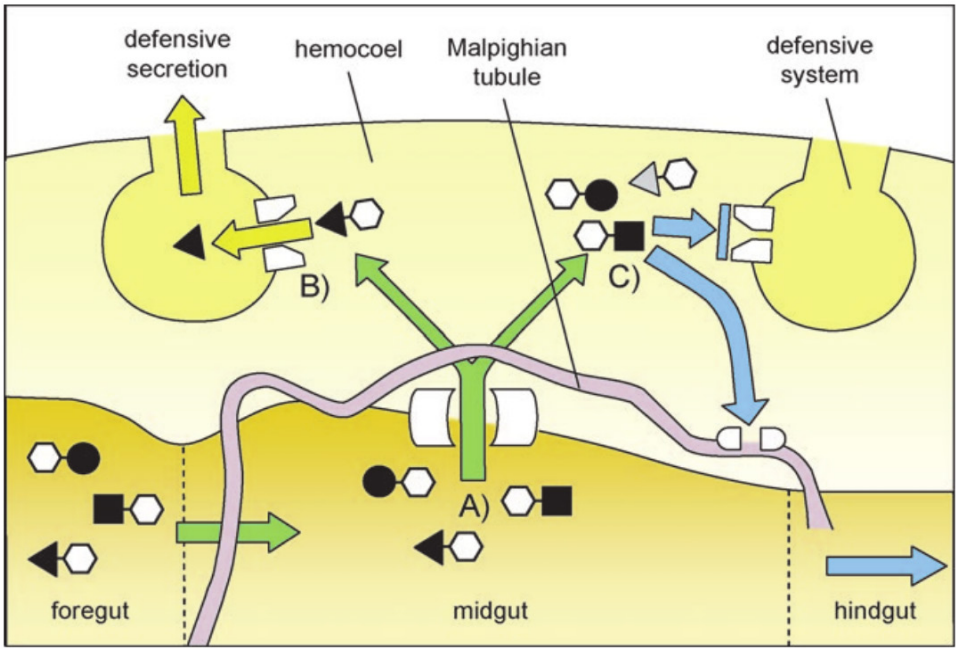


Figure 5: Transport of glucosides from the larval intestinal lumen to the defensive glandular reservoir. A) Unspecific transition of the glucosides through the gut membrane. B) Highly selective import of deterrent precursors into the glandular reservoir. C) Excess of structurally unrelated glucosides through the Malpighian tubules by transporters with low selectivity (adapted from (Discher et al., 2009)).

In order to establish the sequestration process, transport processes are required. On the one hand, there are substrate-specific and, on the other hand, there are substrate-unspecific transporters involved (Discher et al., 2009). Figure 5 demonstrates the transport processes in the interior of a leaf beetle larvae.

ABC TRANSPORTERS

Transporters are assumed to be involved in the sequestration processes, as the species-specific deterrent precursors need to be transported into the defensive glandular tissue. Transporter proteins are classified into main groups including ion channels, water channels, pumps, ATP-binding cassette (ABC) transporters and solute carriers (SLCs) (Fredriksson et al., 2008). ABC transporters, in general, are vital to any living system and are involved in the translocation of a wide variety of substrates ranging from ions, sugars, amino acids, vitamins, lipids, antibiotics and drugs to larger molecules (Hollenstein et al., 2007, Dean et al., 2001). Therefore, with regard to the sequestration process, those are eligible candidates for study. Furthermore, it is known that these transporters are involved in the glycoside transport in plants (Yazaki, 2006). Strauss *et al.* identified an ATP-binding cassette (ABC) transporter, which is present specifically in the cells of the defensive glands (Strauss et al., 2012, Strauss et al., 2013). ABC transporters are composed of two hydrophobic TMDs and two water soluble nucleotide-binding domains (NBDs) bound to the cytosolic face of the TMDs (Biemans-Oldehinkel et al., 2006, Oldham et al., 2008, Seeger and van Veen, 2009). In addition to these core domains, accessory domains or proteins can be part of the ABC transporter (Biemans-Oldehinkel et al., 2006). The "engines of ABC transporters are the NBDs as these power substrate translocation by ATP-binding and hydrolysis" (Biemans-Oldehinkel et al., 2006) (Hollenstein et al., 2007, Schneider and Hunke, 1998, Seeger and van Veen, 2009). A functional ABC transporter can either be a complete protein with two NBDs and two TMDs (a full transporter) or a dimer of two half transporters. Eight subfamilies are constituted for the ABC transporters (ABCA-ABCH). The ABCA and ABCC subfamilies are composed entirely of full transporters, whereas the ABCD, ABCG, and ABCH subfamilies have half transporters. The ABCB subfamily comprises both half and full transporters. Furthermore, the ABCE and ABCF subfamilies are composed of proteins with two NBFs and no TMDs (Dean and Annilo, 2005).

SUGAR TRANSPORTERS

SLCs are the largest group of transporters. In total, there are 52 SLC families (see also http://www.genenames.org/genefamilies/SLC, 12th Dec. 2013). All members of these families vary in their biochemical properties. The sugar transporters here focused on play an important role during the sequestration process and in providing sugar as energy source for, e.g., the mentioned ABC transporters (Wu et al., 2011), belong to the SLC family 2 (Fredriksson et al., 2008, Hoglund et al., 2011).

The protein family of facilitative GLUTs, which belongs to the SLC2 family (Mueckler and Thorens, 2013), comprises 14 isoforms that share common structural features such as 12 transmembrane domains (TMD). Based on their sequence homology, three classes can be distinguished: class I includes GLUT1-4 and GLUT14, class II comprises GLUT5, 7, 9, 11 and class

III GLUT6, 8, 10, 12 and the proton driven myoinositol transporter HMIT (or GLUT13) (Augustin, 2010). Further comprehensive phylogenetic studies of all mammalian GLUT isoforms revealed a segregation of the GLUTs into five distinct classes, with class III splitting into three separate groups. Remarkably, GLUT6 and GLUT8 build up a sister group with strong support, and GLUT10 and GLUT12 form a strongly supported distinct clade, respectively. Moreover, the HMIT sequences fall within a larger but separate clade of sequences (Wilson-O'Brien et al., 2010).

Until now, especially mammalian hexose or glucose transporters have been functionally characterized and observed in detail. Nonetheless, there are just a few functionally characterized sugar transporters in insects. Three of them are the hexose transporter 1 of *Nilaparvata lugens* (NlHT1) which specifically transports glucose (Price et al., 2007), the sugar transporter 6 (NlST6) which Kikuta *et al.* observed as a facilitated glucose/fructose transporter (Kikuta et al., 2010), and the sugar transporter 3 (Ap_ST3) of *Acyrthosiphon pisum* (Price et al., 2010) which results indicate a specificity for the transport of fructose over glucose.

Additionally, trehalose is the major hemolymph sugar in insects (Thompson, 2003). It is predominantly synthesized in the fat body and released into the hemolymph. Kanamori *et al.* proposed that the trehalose transporter 1 (TRET1) of *Polypedilum vanderplanki* is responsible for the incorporation of trehalose into tissues that require a carbon source (Kanamori et al., 2010). In addition they stated that *Tret1* orthologs constitute a subfamily of the human GLUTs, distinct from other insect sugar transporters (Kanamori et al., 2010). Furthermore, they suggest a divergence of *Tret1* orthologs from other sugar transporter sequences to be widespread among insects eventually.

Studying sugar transporters in general in *N. lugens*, Kikuta *et al.* observed, additionally to the already mentioned NlST6, that NlST8 encodes a trehalose transporter (Kikuta et al., 2012). Thus far, the number of functionally characterized trehalose transporters in insects is very small and has not received as much detailed attention as glucose transporters. In conclusion, trehalose and glucose transporters exhibit the same primary structure. Besides the 12 TMD, also the conserved amino acid motifs (Joost et al., 2002, Schurmann et al., 1997, Mueckler and Thorens, 2013, Uldry and Thorens, 2004, Zhao and Keating, 2007) are also present in the trehalose transporters. And it is not possible to predict the substrate specificity from this property alone (Kikawada et al., 2007, Kanamori et al., 2010)[Manuscript 2].

1.3.2 Phylogenetic studies with focus on transporters in insects

In general, phylogenies are built to reconstruct or propose a model of the evolution of life. Both transporter families studied in this thesis comprise many members and subfamilies with highly diverse functionality. On that account and to gain deeper insight into these families,

phylogenetic analyses have been carried out to identify relationships between already functionally characterized transporters and the putative transporters identified in *P. cochleariae* and *C. populi* and, thus, to propose potential substrate specificity.

A large number of methods exist that can be applied to infer phylogenetic trees. In the manuscripts 2 and 3 of this thesis, two different programs have been applied to generate phylogenetic trees. On the one hand, MrBayes which focuses on Bayesian inference of phylogeny, and on the other hand, RAxML which emphasizes maximum likelihood as underlying method were applied.

MrBayes was developed by Huelsenbeck and Ronquist (Huelsenbeck and Ronquist, 2001). A Markov chain Monte Carlo method is used to approximate the posterior probabilities (Metropolis 1953, Hastings 1970, Green 1995). This method works by first proposing a new state for one chain using a stochastic mechanism. Thereafter and second, the probability to accept this state is calculated. This probability involves the "summation over all possible trees" and, for each separate tree, the "integration over all possible combinations of branch length and substitution model parameter values" (Huelsenbeck et al., 2001). Third, a uniform random variable between 0 and 1 is drawn. "If this number is less than the acceptance probability, then the new state is accepted and the state of the chain is updated." Otherwise, it remains in the old state. This procedure is repeated thousands of times. "The proportion of the time any single tree is visited during the course of the chain is a valid approximation of its posterior probability" (Huelsenbeck and Ronquist, 2001). The new version by Ronquist *et al.* provides, in addition, a convergence diagnostic (Ronquist et al., 2012).

RAxML (randomized axelerated maximum likelihood) is a maximum likelihood-based inference program. This ML phylogeny inference is memory intensive, algorithmically complex and highly computationally costly (Stamatakis, 2006) and, therefore, a rapid bootstrap algorithm was implemented (Stamatakis et al., 2008). Bootstrapping is a statistical method used to estimate distributions that are difficult to calculate exactly (Whelan et al., 2001). The resulting bootstrap values give an indication for the robustness of an inferred tree. The optimal maximum likelihood "states that the phylogenetic tree that makes a given sequence data set most likely constitutes the maximum likelihood estimate of the phylogeny and is the preferred explanation" (Rokas, 2011). Having a bootstrap value of at least 90%, it can be confidently stated that the sequences of that specific branch are grouped correctly (Whelan et al., 2001).

Many parameters are involved in the model of sequence evolution. Among those are the substitution rate between amino acids, the frequencies of amino acids and the rate of heterogeneity across sites of the multiple sequence alignment. One approach that is often used in inferring phylogenetic trees is the gamma distribution to approximate the rate distribution in a protein multiple sequence alignment (Rokas, 2011) also was exerted in Manuscript 2 and 3.

1.4 Aims of the thesis

The defense strategy and the corresponding metabolism of deterrent compounds of the Chrysomelina larvae depict an excellent research field. Three different ways to produce the repellent glucosides are proposed, namely the *de novo* production, the sequestration of plant derived compounds and the most evolved strategy - the combination of sequestration and *de novo* production of compounds. Irrespective of the deterrent production, all larvae show common principles. Besides the uniform morphology, common transport mechanisms play a fundamental role. The glucosidically bound precursors are transferred *via* the hemolymph into the glandular reservoir. This uptake is highly selective and substrate specific. One transporter class, the ATP-binding cassette transporter superfamily, was shown to play a key part in this sequestration process (Strauss et al., 2013). Furthermore, the amount of deterrent compound precursors in the defensive glands' tissue has to be maintained on high level to ensure defense at any time. Therefore, sugar transporters are the special focus of this thesis.

In order to comprehensively study the defense metabolism with respect to transport proteins in various leaf beetle species, the transcriptomes of *Phaedon cochleariae* as well as *Chrysomela populi* were sequenced by applying next-generation sequencing technologies, particularly the Illumina sequencing technology.

Therefore, my first objective has been to *de novo* assemble these transcriptomes to provide essential transcript catalogues for further studies. Thereafter, those transcriptomes were annotated and transcripts were assigned putative functions not only by searching for sequence homologies but also by determining (binding) domains and motifs.

Second, by using methods to calculate phylogenetic trees, I was able to denote all identified putative sugar transporters with regard to glucose as well as putative trehalose transporters that have been functionally characterized to date (Manuscript 2).

Third, in order to characterize particular enzymes and proteins, RNA interference (RNAi) methodology was established. To circumvent the knock-down of non-target transcripts, I implemented an off-target prediction method that identifies target sequence sections (all 21 base pairs in length) that are similar to any other (partial) sequence in the transcript library (Manuscript 1).

2 Overview of manuscripts

Manuscript 1: Precise RNAi-mediated silencing of metabolically active proteins in the defense secretions of juvenile leaf beetles

René Roberto Bodemann, Peter Rahfeld, **Magdalena Stock**, Maritta Kunert, Natalie Wielsch, Marco Groth, Sindy Frick, Wilhelm Boland, Antje Burse

Status: published in *Proceedings of the Royal Society B: Biological Sciences* (2012), 279(1745), 4126-4134. doi:10.1098/rspb.2012.1342.

Summary:

An RNAi methodology was established in juvenile *Chrysomela populi* and *Phaedon cochleariae* to analyze *in vivo* functions of proteins involved in the biosynthesis of deterrent compounds. In particular the salicyl alcohol synthase (SAO) of *C. populi* and a novel protein from *P. cochleariae*, annotated as juvenile hormone-binding protein (JHBP), were studied. Both proteins, located in the glandular secretions, were successfully silenced. In order to ensure that the interfering RNA does not target RNA which is not supposed to be diminished, an off-target prediction method has been generated and applied. Altogether, the approach demonstrates clearly that RNAi is a suitable method to annihilate selectively enzymes of a distinct biosynthetic pathway.

Author contributions:

R.R.B. and P.R. established RNAi approach in leaf beetle larvae and performed research on *Cp*SAO and *Pc*To-like. M.S. established and performed off-target prediction and contributed to the interpretation of LC/MSE output data. M.K. designed GC/MS assays, synthesized standards and contributed to the interpretation of output data. N.W. performed LC/MSE analysis and contributed to the interpretation of the MSE data. S.F. generated transcriptome libraries of *P. cochleariae* and performed research to a small proportion. W.B. and A.B. contributed substantially to the interpretation of all data. R.R.B., P.R. and A.B. wrote first draft of the manuscript, and all authors contributed substantially to revisions.

MANUSCRIPT 2: PUTATIVE SUGAR TRANSPORTERS OF THE MUSTARD LEAF BEETLE *PHAEDON COCHLEARIAE*: THEIR PHYLOGENY AND ROLE FOR NUTRIENT SUPPLY IN LARVAL DEFENSIVE GLANDS

Magdalena Stock, René R. Gretscher, Marco Groth, Simone Eiserloh,
Wilhelm Boland, Antje Burse

Status: published in *PLoS ONE* (2013). 8(12): e84461. doi:10.1371/journal.pone.0084461

Summary:
This manuscript provides an overview of all identified sugar transporters in *P. cochleariae* annotated as either solute carrier 2 or trehalose transporter. Phylogenetic studies have been carried out to analyze sequence homology and to observe relationships to already functionally characterized glucose/fructose and trehalose transporters of other Metazoa. Furthermore, the effects of silencing the most abundant sugar transporters *via* RNAi and subsequent sequencing of the defensive glands' transcriptome were observed by applying R statistics (DESeq Bioconductor package) revealing a co-regulation of sugar transporters to ensure sugar homeostasis.

Author contributions:
M.S. designed and performed research, analyzed and interpreted the data including the *de novo* assembly, phylogenetic studies and determining differential expression and wrote a large part of the manuscript. R.R.G. performed RNAi injection, helped to collect samples and to interpret the results. S.E. performed the quantitative real-time PCR experiments. M.G. carried out the cDNA library preparation and Illumina sequencing. A.B. performed quantitative real-time PCR experiments and helped with the interpretation of data. A.B., M.S., and W.B. supervised the work, and all authors revised the manuscript.

MANUSCRIPT 3: TISSUE-SPECIFIC TRANSCRIPT PROFILING FOR ABC TRANSPORTERS IN THE SEQUESTERING LARVAE OF THE PHYTOPHAGOUS LEAF BEETLE *CHRYSOMELA POPULI*

Anja S. Strauß, Ding Wang, **Magdalena Stock**, Marco Groth,
Wilhelm Boland, Antje Burse

Status: in preparation for publication in *PLoS ONE*

Summary:

This publication represents a comprehensive study of all putative ABC transporters of *Chrysomela populi*, a phytophagous leaf beetle. All 65 identified ABC transporters were phylogenetically classified with regard to the subfamilies ABCA to ABCH. Furthermore, the distribution of all ABC transporters among the fat body, gut, Malpighian tubules and defensive glands was analyzed by quantitative real-time PCR experiments. Besides these observations, the most abundant ABC transporter localized in the defensive glands' tissue was silenced *via* RNAi, and the resulting transcriptomic changes were analyzed using the DESeq package. In summary, the ABCC transporter class seems to be responsible for excretion and sequestration processes, and the ABCH subfamily with still unknown function and absence in humans might possess a special function in the defensive glands.

Author contributions:

A.B., A.S.S., and M.S. designed the experiments, interpreted the results and wrote the manuscript. A.S.S. performed the quantitative real-time PCR experiments, interpreted the results and performed RNAi experiments. D.W. performed the identification of ABC transporters and phylogenetic analyses. M.S. *de novo* assembled the transcriptome and performed and interpreted the differential expression analysis. D.W., M.S. and A.B. interpreted the phylogenetic trees. M.G. carried out the cDNA library preparation and Illumina sequencing. R.R.G. dissected the larvae and collected the different tissues. A.B., A.S.S., M.S., and W.B. supervised the work, and all authors revised the manuscript.

MANUSCRIPT 4: INDEPENDENT RECRUITMENTS OF OXIDASES FROM THE GMC OXIDOREDUCTASE FAMILY ENABLED THE EVOLUTION OF CHEMICAL DEFENCE IN LEAF BEETLE LARVAE

Peter Rahfeld, Roy Kirsch, Susann Kugel, Natalie Wielsch, **Magdalena Stock**, Marco Groth, Wilhelm Boland, Antje Burse

Status: submitted to *Proceedings of the Royal Society B: Biological Sciences*

Summary:

Chrysomelina larvae display glandular secretions to repel their enemies in case of attack. The repellents can be produced *de novo* or by sequestering plant-derived precursors. Both biosynthetic strategies include an oxidation step. When using plant-derived precursors, such as salicin, this step is catalyzed by salicyl alcohol oxidases (SAOs) that belong to the GMC oxidoreductase superfamily. Phylogenetic analyses of chrysomeline GMCs showed that the oxidase in the *de novo* production originated from a GMC clade other than that of the SAOs. Thus, the host-dependent chemical defense in Chrysomelina larvae correlates with the independent recruitment of genes from different GMC oxidoreductase families, and the GMC multi-gene family seems to play an important role in the adaptive processes during plant-insect interactions.

Author contributions:

P.R., R.K. and A.B. designed the study. P.R. performed the identification of Pc8HGO, Pc8HGO411 like, the RNAi experiment, the heterologous expression, the resulting protein assays and the interpretation of all resulting data. R.K. extracted and manually annotated GMC encoding sequences, performed the phylogeny of larval chrysomeline glandular oxidases and related GMC oxidoreductases and made the interpretation. S.K. performed qPCR and contributed to the interpretation of output data. N.W. performed LC/MSE analysis, collected and contributed to the interpretation of output data. M.G. and M.S. generated transcriptome libraries, M.S. *de novo* assembled the transcriptome and carried out off-target prediction. W.B. and A.B. contributed substantially to the interpretation of all output data. P.R., R.K. and A.B. wrote the first draft of the manuscript, and all authors contributed substantially to revisions.

3 MANUSCRIPTS

3.1 MANUSCRIPT 1: PRECISE RNAI-MEDIATED SILENCING OF METABOLICALLY ACTIVE PROTEINS IN THE DEFENSE SECRETIONS OF JUVENILE LEAF BEETLES

PROCEEDINGS OF THE ROYAL SOCIETY B

CrossMark click for updates

Proc. R. Soc. B (2012) **279**, 4126–4134
doi:10.1098/rspb.2012.1342
Published online 8 August 2012

Precise RNAi-mediated silencing of metabolically active proteins in the defence secretions of juvenile leaf beetles

René Roberto Bodemann[1,†], Peter Rahfeld[1,†], Magdalena Stock[1], Maritta Kunert[1], Natalie Wielsch[1], Marco Groth[2], Sindy Frick[1], Wilhelm Boland[1] and Antje Burse[1,*]

[1]*Max Planck Institute for Chemical Ecology, Beutenberg Campus, Hans-Knoell-Str. 8, 07745 Jena, Germany*
[2]*Leibniz Institute for Age Research – Fritz Lipmann Institute, Beutenbergstr. 11, 07745 Jena, Germany*

PROCEEDINGS OF THE ROYAL SOCIETY B | BIOLOGICAL SCIENCES

Precise RNAi-mediated silencing of metabolically active proteins in the defence secretions of juvenile leaf beetles

René Roberto Bodemann, Peter Rahfeld, Magdalena Stock, Maritta Kunert, Natalie Wielsch, Marco Groth, Sindy Frick, Wilhelm Boland and Antje Burse

Proc. R. Soc. B 2012 **279**, doi: 10.1098/rspb.2012.1342 first published online 8 August 2012

Supplementary data "Data Supplement" http://rspb.royalsocietypublishing.org/content/suppl/2012/08/01/rspb.2012.1342.DC1.html

References This article cites 56 articles, 15 of which can be accessed free http://rspb.royalsocietypublishing.org/content/279/1745/4126.full.html#ref-list-1

Subject collections Articles on similar topics can be found in the following collections

biochemistry (31 articles)
physiology (53 articles)

Email alerting service Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click here

To subscribe to *Proc. R. Soc. B* go to: http://rspb.royalsocietypublishing.org/subscriptions

PROCEEDINGS OF THE ROYAL SOCIETY B

Proc. R. Soc. B (2012) 279, 4126–4134
doi:10.1098/rspb.2012.1342
Published online 8 August 2012

Precise RNAi-mediated silencing of metabolically active proteins in the defence secretions of juvenile leaf beetles

René Roberto Bodemann[1,†], Peter Rahfeld[1,†], Magdalena Stock[1], Maritta Kunert[1], Natalie Wielsch[1], Marco Groth[2], Sindy Frick[1], Wilhelm Boland[1] and Antje Burse[1,*]

[1]Max Planck Institute for Chemical Ecology, Beutenberg Campus, Hans-Knoell-Str. 8, 07745 Jena, Germany
[2]Leibniz Institute for Age Research – Fritz Lipmann Institute, Beutenbergstr. 11, 07745 Jena, Germany

Allomones are widely used by insects to impede predation. Frequently these chemical stimuli are released from specialized glands. The larvae of Chrysomelina leaf beetles produce allomones in gland reservoirs into which the required precursors and also the enzymes are secreted from attached gland cells. Hence, the reservoirs can be considered as closed bio-reactors for producing defensive secretions. We used RNA interference (RNAi) to analyse *in vivo* functions of proteins in biosynthetic pathways occurring in insect secretions. After a salicyl alcohol oxidase was silenced in juveniles of the poplar leaf beetles, *Chrysomela populi*, the precursor salicyl alcohol increased to 98 per cent, while salicyl aldehyde was reduced to 2 per cent within 5 days. By analogy, we have silenced a novel protein annotated as a member of the juvenile hormone-binding protein superfamily in the juvenile defensive glands of the related mustard leaf beetle, *Phaedon cochleariae*. The protein is associated with the cyclization of 8-oxogeranial to iridoids (methylcyclopentanoid monoterpenes) in the larval exudates made clear by the accumulation of the acylic precursor 5 days after RNAi triggering. A similar cyclization reaction produces the secologanin part of indole alkaloids in plants.

Keywords: RNAi; insects; leaf beetle; secretome; salicyl alcohol oxidase; monoterpene cyclization

1. INTRODUCTION

Insects are extraordinarily inventive when it comes to producing defensive compounds for repelling their enemies. To circumvent auto-intoxicative effects, these natural products frequently originate in the epidermis-derived exocrine glands [1]. The gland cells produce secretions that are fortified with defensive compounds [2,3]. It has been demonstrated that insects convert either intrinsic precursors or food-derived compounds into biologically active allelochemicals [4–7]. The precursors can be activated in the defensive glands or in the secretions. Immature leaf beetles of the subtribe Chrysomelina, for example, produce their deterrents in biphasic secretions, and store them in nine unique pairs of impermeable reservoirs in their backs [8,9]. The larval exudates containing salicyl aldehyde (**3**) have been of particular interest [10,11]. The hydrophobic aldehyde forms an organic layer, accounting for 15 per cent of the total discharge volume, while the aqueous phase constitutes 85 per cent [12]. The latter contains the precursor salicyl alcohol (**2**) and a flavine-dependent salicyl alcohol oxidase (SAO); the SAO uses molecular oxygen as an electron acceptor for alcohol oxidation, yielding the aldehyde and hydrogen peroxide [12–14] (figure 1). Salicyl aldehyde is considered as a potent repellent against generalist predators [11,15] and as an antimicrobial agent [16]. The larvae feed on salicaceous plants and sequester the secondary metabolite salicin (**1**) [17–19]. After shuttling salicin to the defensive glands, the glucoside is cleaved by a β-glucosidase into **2** and glucose for further metabolism [20] (figure 1). According to phylogenetic analyses, the synthesis of **3** from sequestered precursors has evolved from the de novo production of defensive iridoids (methylcyclopentanoid monoterpenes containing an iridane skeleton) [21]. Also the last steps of the iridoid pathway in the secretions are thought to be similar to those found in sequestering species [20] (figure 1). At first, the sugar moiety is cleaved from 8-hydroxygeraniol-8-*O*-β-D-glucoside (**4**), and an oxygen-dependent oxidase converts the aglucone into 8-oxogeranial (**6**) [20,22–24]. A subsequent cyclization reaction yields iridoids (**7**) [25].

Despite the many current genome- and transcriptome-sequencing projects, up to now it has only been shown for SAO sequences to be entangled in allomone production in the defensive secretions of the leaf beetle species *Chrysomela tremulae*, *Chrysomela populi*, *Chrysomela lapponica* and *Phratora vitellinae* [13,14,26]. To demonstrate the *in vivo* relevance of a target sequence, gene silencing by RNA interference (RNAi) is a suitable method. RNAi is an endogenous mechanism, derived from an anti-viral immune response [27], and can be found virtually in all eukaryotic species. It can be triggered artificially by double-stranded RNA (dsRNA), whose nucleotide sequence is identical to that of the target gene [28]. The

* Author for correspondence (aburse@ice.mpg.de).
† These authors contributed equally to the study.

Electronic supplementary material is available at http://dx.doi.org/10.1098/rspb.2012.1342 or via http://rspb.royalsocietypublishing.org.

Received 12 June 2012
Accepted 17 July 2012

This journal is © 2012 The Royal Society

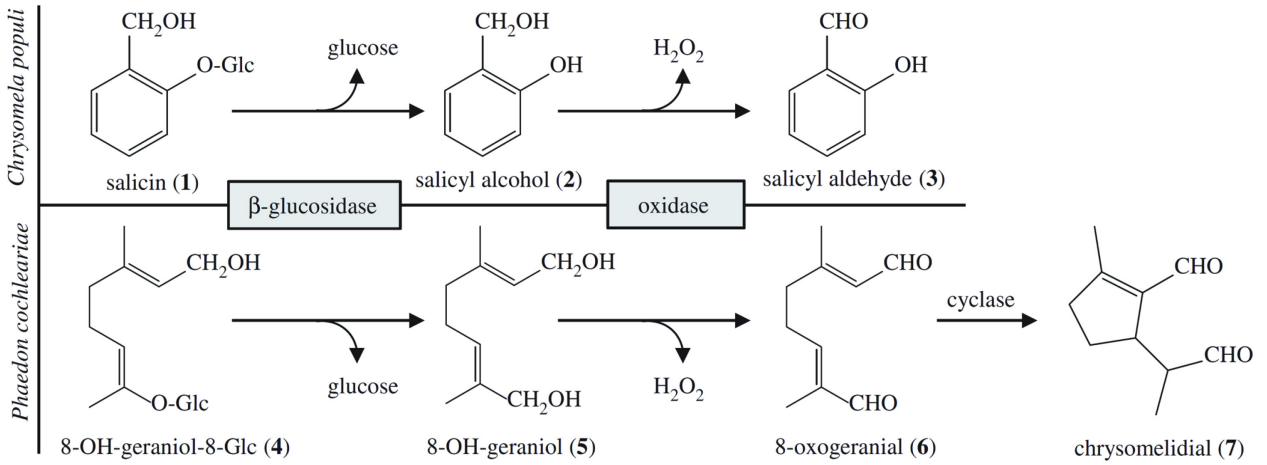


Figure 1. Enzymatic reactions in the defensive secretions of juvenile *C. populi* and *P. cochleariae* adapted from Michalski *et al.* [14]. Glc, Glucose.

RNAi effect is attended by decreased transcript and protein levels, and consequently by loss-of-function phenotypes. In addition to embryogenesis, pattern formation, reproduction and behaviour, RNAi allows biosynthetic pathways in insects to be successfully analysed [29–31].

Here, we describe how RNAi can be used to target the biosynthesis of discrete components in the defensive discharges of juvenile Chrysomelina. We first validated this technique by silencing the known *SAO* sequence in the sequestering species *C. populi* (*CpopSAO*). After knocking down the SAO, the alcohol precursor of **3** accumulated in the gland. This showed that we are able to interrupt the deterrent biosynthesis *in vivo*. Next, we extended the method to the related de novo iridoid-producing species, *P. cochleariae*. In the secretions of its larvae C_{10}-precursors are converted to the methylcyclopentanoid monoterpene chrysomelidial. Particularly, the cyclization mechanism is of importance because it occurs not only in insects but also in plants. Here, the cyclization leads ultimately to secologanin, one of the building blocks for more than 2500 indole alkaloids that have been isolated mainly from three plant families [32]. Although an enzyme with cyclase activity for secologanin biosynthesis has long been predicted, a corresponding sequence has yet to be published. In the *P. cochleariae* secretome, we identified a novel protein which is involved in the cyclization reaction of the monoterpenoid 8-oxogeranial to chrysomelidial.

2. MATERIAL AND METHODS

See electronic supplementary material for complete secretome analyses by data-independent liquid chromatography/mass spectrometry detection (LC/MS^E), cloning procedures, detailed quantitative real-time PCR procedure (qPCR), all primer sequences and accession numbers.

(a) *Beetle rearing and secretion analyses*

Chrysomela populi (L.) was collected near Dornburg, Germany (latitude 51.015, longitude 11.64), on *Populus maximowiczii* × *Populus nigra*. In the laboratory, beetles were kept in a 16 L : 8 D cycle, 18 ± 2°C in light and 13 ± 2°C in darkness. *Phaedon cochleariae* (F.) was laboratory-reared on *Brassica oleracea* convar. *capitata* var. *alba* (Gloria F1) in 16 L : 8 D cycle conditions and 15 ± 2°C. According to [33], we obtained the relative growth rate (RGR) of six biological replicates of each group of five larvae by RGR = [(final weight − weight of neonate larva)/(weight of neonate larva × developmental time (days))]. Each replicate group was weighed every 24 ± 3 h and data were compared with two-tailed *t*-test. Larval secretions were collected in glass capillaries (inner diameter, 0.28 mm; outer diameter, 0.78 mm, length 100 mm; Hirschmann, Eberstadt, Germany). Sealed capillaries containing samples were stored at −20°C until needed. Secretions were weighed in the sealed capillaries on an ultra-microbalance (Mettler-Toledo, Greifensee, Switzerland) three times; the weight of the capillaries was subtracted and the final weight was averaged.

(b) *Production of double-stranded RNA*

Sequenced plasmids pIB-*CpopSAO* (GeneBank: HQ245 154.1) and pIB-*PcTo-like* (GeneBank: JQ728549) were used to amplify a 1.5 kb *CpopSAO* fragment and a 450 bp *PcTo-like* fragment, respectively. The *gfp* sequence was amplified from pcDNA3.1/CT- GFP-TOPO (Life Technology, Darmstadt, Germany). The amplicons were subject to *in vitro* transcription assays according to instructions from the Ambion MEGAscript RNAi kit (Life Technologies, Darmstadt, Germany). The resulting dsRNA was eluted after nuclease digestion three times with 50 µl of injection buffer (3.5 mM Tris–HCl, 1 mM NaCl, 50 nM Na_2HPO_4, 20 nM KH_2PO_4, 3 mM KCl, 0.3 mM EDTA, pH 7.0). The concentration of dsRNA was calculated with A = 1 = 45 mg ml^{-1} and adjusted to 1 µg $µl^{-1}$. The quality of dsRNA was checked by TBE-agarose-electrophoresis.

(c) *Injection of double-stranded RNA*

First instar of *C. populi* with 5 mm body length was injected with 0.1–3 µg of dsRNA approximately 10 days after hatching. *Phaedon cochleariae* second instar with 4 mm body length was injected with 0.3 µg of dsRNA approximately 5 days after hatching. Injections were accomplished with ice-chilled larvae using a Nano2000 injector (WPI, Sarasota, FL, USA) directed by a three-axis micromanipulator. The larvae were injected parasagittally between the pro- and mesothorax.

(d) *Off-target prediction*

According to the mechanism of RNAi [28], the top and bottom strands of dsRNAs of *CpopSAO*, *PcTo-like* and *gfp* were diced *in silico* into all possible 21 bp fragments [34]. The resulting siRNAs were subjected to BLASTn

(stand-alone NCBI-BLAST) [35] by invoking BLASTALL v. 2.2.21 (parameters: -p blastn -e 1e-1 -G 7 -T -b 80 -v 80) searching against our in-house transcriptome databases of *C. populi* and *P. cochleariae*. Hits less than 20 nts in length were ignored and hits more than or equal to 20 nts were considered as putative off targets.

(e) CpopSAO *and* PcTo-like *transcript abundance*

Cq values of genes of interest from three biological replicates were normalized by *CpRPL45* and *CpActin* for *C. populi* and *PcRP-L8* and *PcRP-S18* for *P. cochleariae*, respectively. Real-time PCR data were acquired on an Mx3000P Real-Time PCR system using Brilliant II SYBR Green qPCR Master Mix (Agilent, Santa Clara, CA, USA).

(f) *Gas chromatography/mass spectrometry analysis of low-molecular-weight compounds in chrysomelid secretions*

Secretions of *C. populi* were diluted in 1 : 150 (w/v) ethyl acetate and secretions of *P. cochleariae* were diluted in 1 : 100 (w/v) dichloromethane. Of each diluted secretion, 1 µl was subjected to GC/EIMS analysis (ThermoQuest Finnigan Trace GC/MS 2000, Frankenhorst, Germany) equipped with Phenomenex (Aschaffenburg, Germany) ZB–5–W/Guardian–column, 25 m. Substances were separated using helium as a carrier (1.5 ml min^{-1}). Conditions for *C. populi* secretions: 50°C (1 min), 10°C per minute to 80°C, 60°C per minute to 280°C (1 min). Inlet temperature was 220°C, transfer line was 280°C. Substances were identified according to standard substances **2** and **3**. Conditions for *P. cochleariae* secretions: 50°C (2 min), 10°C per minute to 80°C, 5°C per minute to 200°C, 30°C per minute to 300°C (1 min). Inlet temperature was 220°C and transfer line was 280°C. Substances were identified according to [36] and the reference compounds 8-oxogeranial and chrysomelidial. The synthesis of 8-oxogeranial and chrysomelidial was carried out as in [25,37], respectively. Peak areas from GC-chromatograms were obtained using an ICIS-algorithm (XCALIBUR BUNDLE v. 2.0.7, Thermo Scientific).

(g) *Statistical analyses*

Two-tailed Student's *t*-tests for unequal variation were used to value the significance levels of transcript abundances and to weight differences comparing values of three different biological replicates from the non-injected control (NIC) group with those of either the RNAi group or the *gfp* control. Multi-dimensional ANOVA tests were carried out to validate significant differences in time series and between different RNAi treatments. The level of significance was reached at a *p*-value of 0.05. Calculations were done with R (http://www.r-project.org/).

3. RESULTS

(a) *Targeting the defensive glands of juvenile poplar leaf beetles by RNA interference*

Recently, a 1872-bp *CpopSAO* cDNA (Genbank/HQ245154.1) encoding a 69 kDa protein for conversion of **2** into **3** was identified from the larval defensive glands of *C. populi* [13,14] (figure 2*a*). It belongs to the glucose–methanol–choline (GMC) family of oxidoreductases [38]. Given that the expression of *CpopSAO* was detectable exclusively in glandular tissues (figure 2*b*), silencing this gene would affect only the process of glandular biosynthesis.

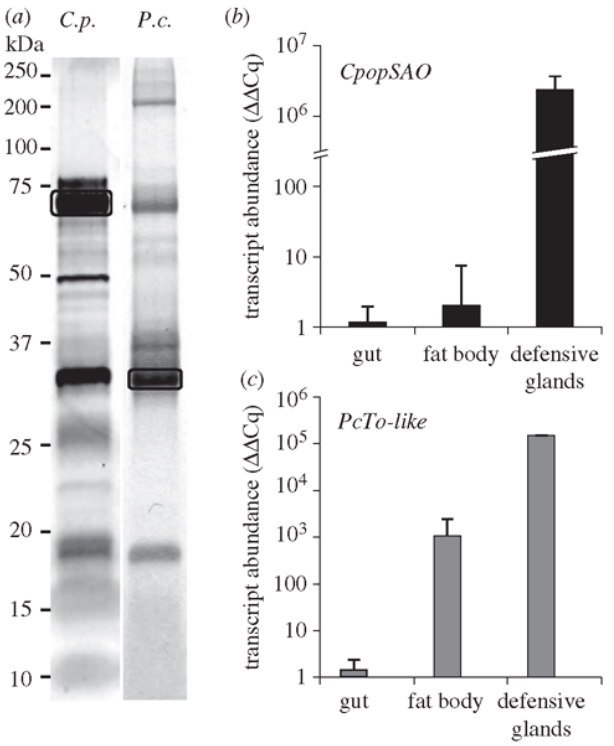


Figure 2. Protein and transcript abundance in juvenile leaf beetles. (*a*) Proteins in defensive exudates separated by one-dimensional SDS–PAGE. left: 1 mg secretions of *C. populi* (*C.p.*), silver stained, box marks *Cpop*SAO; right: 0.65 mg secretions of *P. cochleariae* (*P.c.*), Coomassie stained, box marks *Pc*To-like. (*b*) Expression pattern of *CpopSAO* ±s.d. in different *C. populi* tissues, $n = 3$. (*c*) Expression pattern of *PcTo-like* ±s.d. in different *P. cochleariae* tissues, $n = 2$. Both *y*-axes are in $\log_{10}$ scale.

To induce RNAi in *C. populi* larvae, we injected 1.0 µg of 1.5 kb *CpopSAO* dsRNA into late first instar. A 719-bp dsRNA fragment of *gfp* served as a control for effects caused by dsRNA; although the RNAi machinery will be induced, genes should not be silenced. Furthermore, we included an NIC group in our experiments. By monitoring the developmental traits and the secretion production in *C. populi* and comparing the results with those from control groups, we found that silencing *CpopSAO* did not influence either growth rate or pupae weight (see the electronic supplementary material, figure S1*a*). But the larvae treated with *CpopSAO* dsRNA produced slightly more secretions than did the larvae of the control groups (see the electronic supplementary material, figure S1*b*), which might be owing to the different osmotic characteristics between **2** and **3** [12]. Because we did not detect significant differences between NIC and *gfp* controls in any experiments delineated below, we continue showing only the data of the *gfp* controls.

Transcript abundance was measured in glandular tissue using qPCR after 1, 3 and 12 days. Comparing tissue from these samples to tissue from the *gfp* controls, we noticed significant reductions to 7.6 per cent mRNA level ($p = 0.002$) just 24 h after injection. After day 3, the transcript level was diminished to 1.6 per cent ($p = 0.004$), and after day 12, to 0.5 per cent (figure 3*a*).

In accordance with the literature, SAO corresponds to the dominant band at 70 kDa in the secretions of

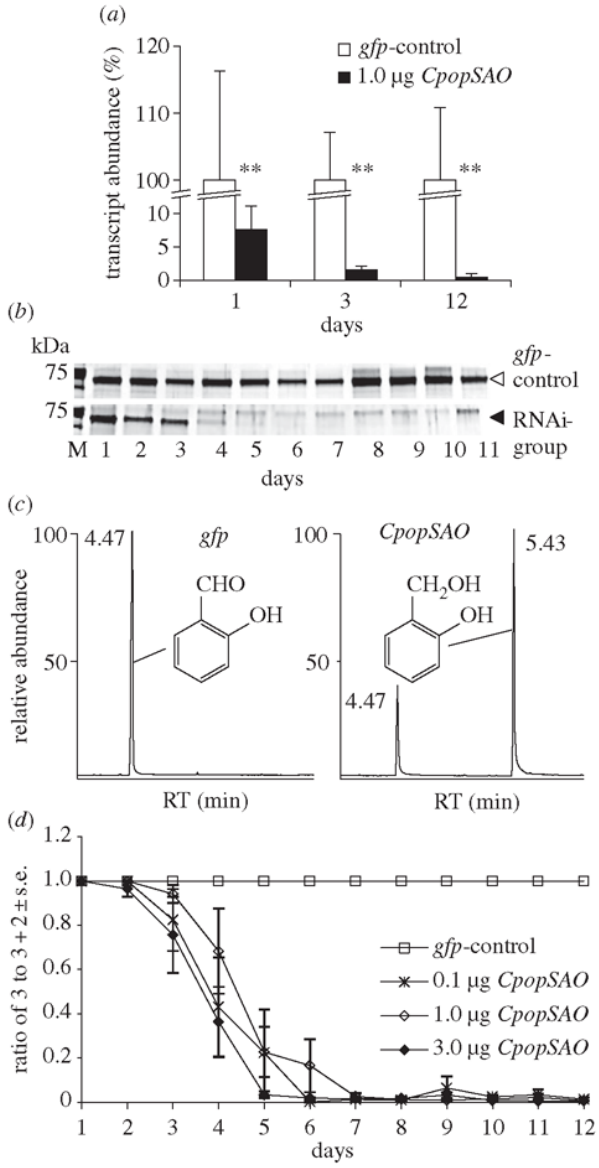


Figure 3. RNAi effects in juvenile *C. populi*. (*a*) White bars: transcript abundance of *CpopSAO* after injecting 1.0 µg *gfp* dsRNA, $n = 3 \pm$ s.d.. Black bars: transcript abundance of *CpopSAO* after injecting 1.0 µg *CpopSAO* dsRNA, $n = 3$. 100% = ΔCq of *gfp* control. Asterisks indicate level of significance: $^{**}p < 0.01$ (*b*) *Cpop*SAO protein abundance in defence secretions was monitored over time. A total of 0.85 mg secretions per lane was separated on silver-stained SDS gels. The 70-kDa band corresponds to *Cpop*SAO. Secretions originating from control treatment with 1 µg dsRNA of *gfp* (white arrowhead) and RNAi treatment with 1 µg dsRNA of *CpopSAO* (black arrowhead) are shown. (*c*) GC-chromatogram of secretions on day 5 after treatment; left: injecting 1.0 µg *gfp* dsRNA resulted in the production of **3**, right: injecting 1.0 µg *CpopSAO* dsRNA resulted in the production of **2** and **3**. (*d*) GC-chromatogram peak-area-based plot of secretions after dsRNA injection of *gfp* and different amounts of *CpopSAO* ($n = 5$).

C. populi (figure 2*a*) [12,14]. The composition of the secretome after dsRNA treatment was monitored in a time series in silver-stained one-dimensional SDS gels. Owing to the silencing effect, the quantity of *Cpop*SAO was apparently reduced just 2 days after dsRNA injection and the protein was barely visible after day 5 (figure 3*b*).

The effects on the biosynthesis of **3** in the defensive secretions were determined by GC/MS analysis. For these experiments, 0.1, 1.0 and 3.0 µg of *CpopSAO* dsRNA were injected into larvae from the same clutch. As in the protein reduction, we detected **2** in the defensive secretions just 2 days after the injection of 3.0 µg *CpopSAO* dsRNA (figure 3*c*). Compound **2** was not detectable in *gfp* control secretions. In addition, no unexpected chemical compound arose owing to the dsRNA treatments. By setting the peak area of **3** in ratio equalling the sum of the main peak areas, a diagram of the RNAi-dependent reduction of **3** can be plotted (figure 3*d*). We have tested dsRNA amounts ranging from 0.1 to 3.0 µg. After RNAi induction, significantly less aldehyde was observed for the 3.0 µg *CpopSAO* group ($p = 0.015$) on the 4th day and for all tested *CpopSAO* groups on the 5th day (0.1 µg, $p = 0.016$; 1 µg, $p = 0.002$; 3 µg, $p = \ll 0.001$). Biological variation prevented us from observing dose-dependent RNAi effects in these experiments; the amount of **2** did not differ significantly between the RNAi samples.

(b) *Off-target prediction and validation for* CpopSAO

Owing to strong sequence identities, co-silencing non-target genes can cause unintended side-effects [39,40]. Therefore, we performed off-target predictions for the desired dsRNA sequences of *CpopSAO* and *gfp*. Predicted off-target genes were validated by qPCR using cDNA derived from successful RNAi experiments. Because of a lack of genome sequences, the potential silencing effects of targeting the nucleus where fragments of the long dsRNA may bind to non-transcribed regulatory sequences [41] or introns [42] could be neither predicted nor validated.

For *gfp* dsRNA, no critical candidates were detected in the transcriptome library of *C. populi*. Off-target analyses in the *C. populi* sequence library, however, identified 25–21 bp contiguous regions of *CpopSAO* dsRNA that were identical to sequences of eight unique transcripts (see the electronic supplementary material, table ST2 for putative off-target hits). Three of them encode putative proteins having the GMC-oxidoreductase motif in the C-terminal region (*Cp*GMClike I-III) and five were annotated as hypothetical proteins (*Cp*COMP3092; *Cp*COMP6024; *Cp*COMP36289; *Cp*COMP38777; *Cp*COMP51471).

CpopSAO shares with *CpGMClike-I* two similar regions spanning 22 and 25 nucleotides (nts) each; these regions are interrupted by one mismatch (22/1 and 25/1) and, with *CpGMClike-III*, one similar sequence stretch without mismatch (24/0). *CpGMClike-II* and the five remaining transcripts possess sequence regions of 22/0 to 20/0 nts identical to *CpopSAO*. In all tissues, all putative off-target genes exhibited generally low expression levels with relative Cq values median less than 2×10^{-3}. qPCR assays were carried out for all eight targets 12 days after larvae were treated with 1.0 µg 1500-bp *CpopSAO* dsRNA; only for the *CpGMClike-I* and *CpGMClike-II* did these treatments reveal significant differences of transcript level in the gut tissue ($p = 0.049$; $p = 0.032$). No other tested transcripts showed changes of mRNA abundance in the examined tissues

(see the electronic supplementary material, figure S2). Since *C. populi* larvae transport the plant-derived precursor into the defensive glands for final transformation, we assume that the off-target effects on the putative GMC-oxidoreductases in gut tissue of unknown function do not distort the RNAi effects observed in the secretions.

(c) *Identification of unknown proteins in the defence-related secretome of* Phaedon cochleariae

After successfully introducing the 'lack-of-function approach' to the defensive secretions of *C. populi* by silencing an enzyme for which we had a clear expectation of the resulting phenotype, we used the method to identify proteins in unknown secretions. For this reason, we chose the larval exudates from the related de novo iridoid-producing species *P. cochleariae*. We assigned to the abundant 35-kDa band a putative protein whose deduced sequence contains 243 amino acids and a 22 amino acid signal peptide; the existence of such a sequence suggests that the mature protein is secreted (figure 2*a*). It possesses a conserved domain (pfam06585) characteristic for the juvenile hormone-binding protein (JHBP) superfamily. Sequence comparisons using the BLAST algorithm [35] revealed that the *P. cochleariae* amino acid sequence shares only very limited identity with functionally characterized insect proteins, for example, 12 per cent identity with the JHBP from *Bombyx mori* [43] and 16 per cent with the takeout (To) 1 from *Epiphyas postvittana* [44] (figure 2*a*). Higher identities up to 25 per cent were found only with insect proteins not yet fully characterized in their functions, such as those with the To-like protein (NP_001191952) from *Acyrtosyphon pisum* or the JHBP-like (XP_001359416) from *Drosophila pseudoobscura pseudoobscura*. None of the mentioned insect species is known to produce cyclic monoterpenoids.

The JHBP superfamily combines both the To protein family and the JHBP family. There are two major differences between the families: one is the number of disulphide bonds (To proteins have one and JHBPs have two) and the second are the conserved C-terminal sequence motives that are only present in To proteins. *In silico* analyses predicted in the *P. cochleariae* sequence seven N-glycosylation sites and only one disulphide bond. Along with identifying the two To-specific motives [45] (figure 4) in the C-terminal region, we conclude that our protein can be attributed to the To family. Therefore, we named it *Pc*To-like.

Despite the generally low sequence similarity, most To proteins and JHBPs are classified as ligand-binding proteins for juvenile hormones or similar hydrophobic terpenoids [44,46–49]. Because the precursor of the cyclic iridoid is also a terpenoid, we hypothesize that the putative protein could be involved in the iridoid biosynthesis in the defensive secretions. The assumption that the putative protein has relevance in the defensive glands is corroborated by the high transcript level which has been detected mainly in the glandular tissue of juvenile mustard leaf beetles (figure 2*c*). Low mRNA levels were also detectable in the fat body tissue.

(d) *RNA interference effects in larvae of* Phaedon cochleariae

A total of 0.3 µg of dsRNA derived from a 450-bp fragment of the *PcTo-like* sequence was injected into second instars of *P. cochleariae*. Transcript quantification 5 days after dsRNA injection confirmed a significant reduction of the mRNA in glandular tissue ($p = 0.043$) down to 1.0 per cent ($\pm 0.9\%$) compared with mRNA levels in *gfp* injections (figure 5*a*).

Phenotypic analyses after injection of dsRNA on the composition of low-molecular-weight compounds in the secretions were carried out by using GC/MS. The quality of the metabolites in samples collected 1 and 2 days after *Pc*To-like RNAi induction did not vary from the quality of the metabolites in those collected from *gfp* controls. In both treatments, we detected only the end-product **7**. The first deviation in the composition of the secretions was measured 3 days after dsRNA injection. Only in samples triggered by *PcTo-like* RNAi did minor amounts of the postulated intermediate **6** in addition to **7** emerge. After 5 days, however, **6** clearly accumulated in addition to **7** owing to the RNAi effect (figure 4*b*). Therefore, we conclude that the *Pc*To-like has to be involved in the cyclization of monoterpene precursors into iridoids.

Off-target effects were predicted using the method described for *CpopSAO*, and predicted off-target effects were avoided from final dsRNA sequence by choosing the template for dsRNA outside of areas of predicted off-target effects.

4. DISCUSSION

The results of our larval RNAi experiments clearly demonstrate selective excision of a component in a biosynthetic pathway. To the best of our knowledge, RNAi has never been used to target enzymes in insect defensive secretions. Owing to the silencing of *Cpop*SAO, the chemical composition in the larval exudates of *C. populi* was massively altered, starting as early as 48 h after treatment. This shows a distinct function of this enzyme *in vivo*. Before, Kirsch *et al.* [13] showed activity only in *in vitro* assays. Evidently, RNAi is a valuable technique for identifying *in vivo* relevance for unknown proteins in defensive glands. Although insects contain a large number of exocrine glands in which bioactive compounds are produced, to date few studies have relied on RNAi to provide evidence for the *in vivo* function of enzymes in insect glands. One example is the production of sex pheromones in special glands of the silkmoth *Bombyx mori*. By injecting the pupae with dsRNA, Ohnishi *et al.* [50,51] were able to dissect the components of the biosynthetic pathway as well as assign a function to a transport protein within the glands of adults. Another RNAi target was the production of pheromone in jewel wasps, *Nasonia vitripennis*. Silencing an epoxide hydrolase in these insects resulted in pheromone reduction by 55 per cent and suppressed the targeted gene transcripts by 95 per cent [52]. Freshly emerged males were injected and 2 days later levels of transcript and pheromones were analysed. As our results demonstrate, RNAi effects are easily detectable in exocrine glands. In the secretions of immature *P. cochleariae*, we were able to assign *in vivo* relevance to a cDNA encoding a protein which is important for the cyclization of iridoids. The iridoid pathway in insects was already proposed by using deuterium labelled precursors by Weibel *et al.* [53]. In his work, the stereospecifity of the cyclization was analysed and allocated to an enzymatic conversion. However, to date JHBPs and

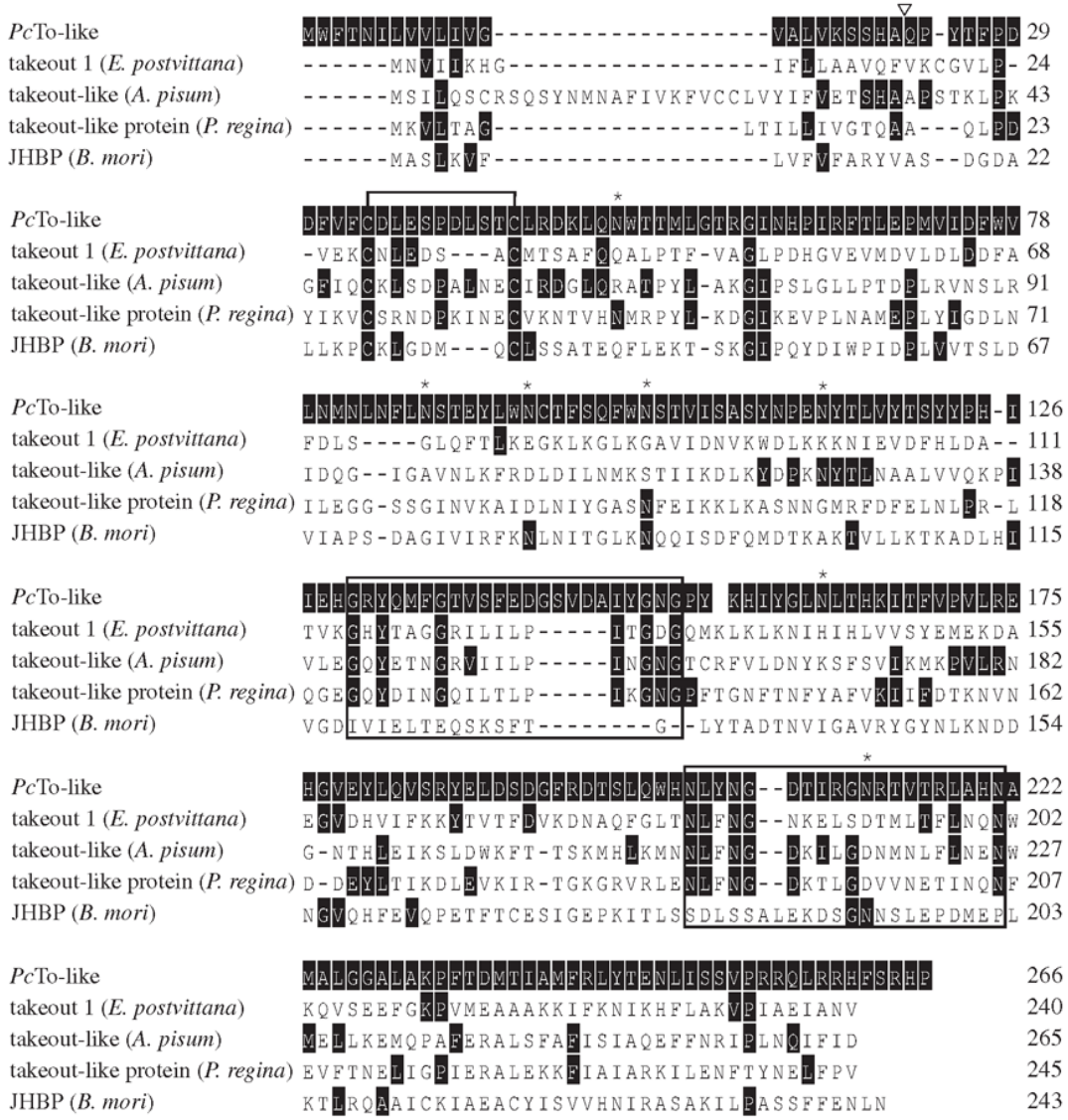


Figure 4. Amino acid alignments (ClustalW) of *Pc*To-like from *P. cochleariae* and other members from the To/JHBP family (*Epiphyas postvittana* To1: GeneBank: ACF39401; *Acyrthosiphon pisum* To-like: GeneBank: XP_001952685; *Phormia regina* To-like protein: GeneBank: BAD83405; *Bombyx mori* JHBP: GeneBank: AAF19267). Solid black shading depicts amino acids identical to *Pc*To-like sequence. White inverted triangles indicate the predicted signal peptide of *Pc*To-like. Asterisks mark the predicted N-glycosylation sites. Conserved cysteine residues that form disulphide bonds are marked with a bracket above the sequence. The two black boxes show the location of the To-typical motives.

To proteins have been established as being carriers of hydrophobic ligands [44,48]. Several lines of evidence indicate that JHBPs form complexes with juvenile hormones (JHs) which provide protection of the chemically labile JHs against nonspecific enzymatic degradation and/or adsorption to lipophilic surfaces during the delivery process from the production site to the target tissue [46,47,49]. Up to now only the crystal structure of To 1 from *E. postvittana* with ubiquinone provided direct evidence for ligand binding in To proteins [44]. Most of the putative To proteins await elucidation of their mode of action. Therefore, the actual mechanism how *Pc*To-like acts in the defensive exudates has to be analysed *in vitro* with purified recombinant protein. On-going experiments will reveal more functional enzymes in Chrysomelina and clarify the molecular machinery for the biosynthesis of deterrents in larval defence secretions.

To perform RNAi experiments, it is essential to ensure the specificity. Off-target effects can arise when siRNAs diced from long dsRNA fragments possess sufficient sequence similarity to non-target mRNA and thus triggering degradation of similar sequences [39]. For sequenced organisms, genome-scale off-target prediction programs are available [34]. These approaches are not suitable for organisms whose genomes have yet to be sequenced. In the last few years, several approaches have been used to detect off targets for those species, such as screening for target specificity by rapidly amplifying cDNA ends [54]. Another approach has used microarrays to compare the cDNAs from treated groups with those from non-treated groups; such comparisons offer proof of differentially expressed transcripts *via* qPCR [55]. Transcriptome sequences have rarely been used for approaches based on local alignment algorithms but represent an economical

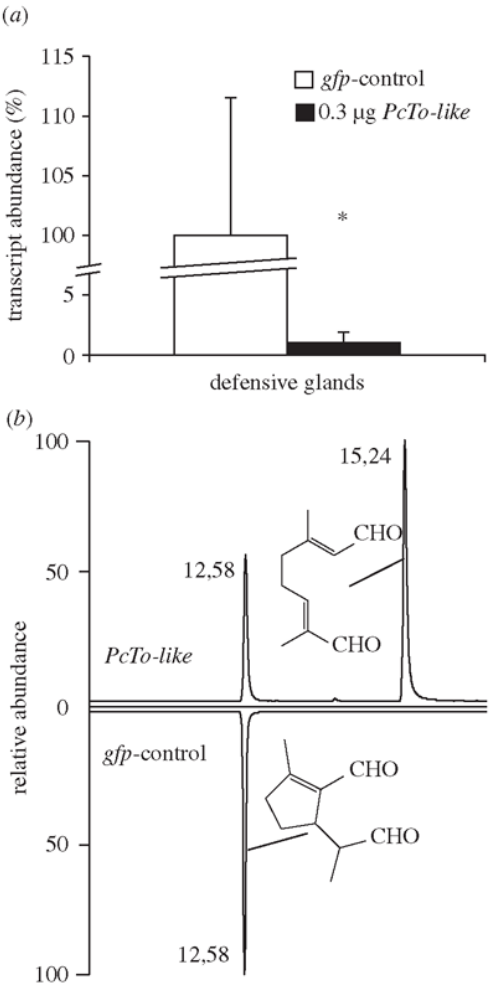


Figure 5. RNAi effects in juvenile *P. cochleariae*. (*a*) White bars: transcript abundance of *PcTo-like* after injecting 0.3 µg *gfp* dsRNA, $n = 3$ ±SD. Black bars: transcript abundance of *PcTo-like* after injecting 0.3 µg *PcTo-like* dsRNA, $n = 3$. 100% = ΔCq of *gfp*-control. Asterisks indicate level of significance: $*p < 0.05$. (*b*) GC-chromatogram of diluted secretions on day 5 after treatment, above: injecting 0.3 µg *PcTo-like* dsRNA resulted in the production of **6** and **7**; below: injecting 0.3 µg *gfp* dsRNA resulted in the production of **7**.

way to make off-target predictions [56]. In our case, we showed that *in silico* dicing of long dsRNA pieces to 21-bp fragments and subsequent BLASTn searches in our transcriptome libraries also lead to the identification of putative off-target transcripts. Subsequent qPCR analysis after successful RNAi induction revealed the co-silencing of predicted transcripts in *C. populi*. Two of eight mRNAs were significantly altered in gut tissue (see the electronic supplementary material, figure S2). But the observed off-target silencing could be assigned neither to the length of the fragments nor to the amount of pmol of the putative siRNAs (see the electronic supplementary material, table ST2). Furthermore, the composition and internal stability of the sequence fragments are supposed to have an impact on successful RNAi triggering [57] and could be included in the prediction. Although publications concerning off-target prediction have increased in the last 2 years, as yet no standard method is available. But as our results indicate, off-target validation is crucial for a realistic discussion of RNAi effects.

The authors would like to express their gratitude to Heiko Vogel for making available 454-sequences from *P. cochleariae*. We also gratefully acknowledge Angelika Berg, Kerstin Ploss, Gerhard Pauls and Antje Loele for technical assistance. We thank Gregor Bucher for helpful discussions on aspects of this work. Very special thanks are due to Prof. Jacques M. Pasteels and Emily Wheeler for their critical reading of the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft (BU 1862/2-1) and the Max Planck Society.

R.R.B., P.R., M.S., M.K., N.W., W.B. and A.B. designed study. R.R.B. established RNAi in leaf beetles and performed RNAi treatments of *Cpop*SAO and its control treatments, performed off-target validation, collected all corresponding data except the off-target prediction and analysed output data. P.R. identified *PcTo-like*, performed RNAi treatments of *PcTo-like* and control treatments, collected all corresponding data and analysed output data. S.F., M.S. and M.G. generated transcriptome libraries. M.S. established and performed off-target prediction and contributed to interpretation of LC/MS^E output data. M.K. designed GC/MS assays, synthesized **6** and **7** and contributed to interpretation of output data. N.W. performed LC/MS^E analysis, collected and contributed to interpretation of output data. W.B. and A.B. contributed substantially to interpretation of all output data. R.R.B., P.R. and A.B. wrote first draft of the manuscript, and all authors contributed substantially to revisions.

REFERENCES

1 Quennedey, A. 1998 Insect epidermal gland cells: ultrastructure and morphogenesis. In *Microscopic anatomy of invertebrates, vol. 11A: insecta* (eds W. H. Harrison & M. Locke) pp. 177–207. New York: Wiley-Liss.
2 Dettner, K. 1987 Chemosystematics and evolution of beetle chemical defenses. *Annu. Rev. Entomol.* **32**, 17–48. (doi:10.1146/annurev.en.32.010187.000313)
3 Pasteels, J. M., Gregoire, J. C. & Rowell-Rahier, M. 1983 The chemical ecology of defense in arthropods. *Annu. Rev. Entomol.* **28**, 263–289. (doi:10.1146/annurev.en.28.010183.001403)
4 Burse, A., Frick, S., Discher, S., Tolzin-Banasch, K., Kirsch, R., Strauss, A., Kunert, M. & Boland, W. 2009 Always being well prepared for defense: the production of deterrents by juvenile Chrysomelina beetles (Chrysomelidae). *Phytochemistry* **70**, 1899–1909. (doi:10.1016/j.phytochem.2009.08.002)
5 Duffey, S. S. 1980 Sequestration of plant natural-products by insects. *Annu. Rev. Entomol.* **25**, 447–477. (doi:10.1146/annurev.en.25.010180.002311)
6 Opitz, S. E. W. & Mueller, C. 2009 Plant chemistry and insect sequestration. *Chemoecology* **19**, 117–154. (doi:10.1007/s00049-009-0018-6)
7 Laurent, P., Braekman, J. C., Daloze, D. & Pasteels, J. 2003 Biosynthesis of defensive compounds from beetles and ants. *Eur. J. Org. Chem.* 2733–2743. (doi:10.1002/ejoc.200300008)
8 Hinton, H. E. 1951 On a little-known protective device of some chrysomelid pupae (Coleoptera). *Proc. R. Entomol. Soc. A* **26**, 67–73. (doi:10.1111/j.1365-3032.1951.tb00123.x)
9 Pasteels, J. M. & Rowell-Rahier, M. 1989 Defensive glands and secretions as taxonomical tools in the Chrysomelidae. *Entomography* **6**, 423–432.
10 Pasteels, J. M., Daloze, D. & Rowell-Rahier, M. 1986 Chemical defense in chrysomelid eggs and neonate larvae. *Physiol. Entomol.* **11**, 29–37. (doi:10.1111/j.1365-3032.1986.tb00388.x)

11 Pasteels, J. M., Rowell-Rahier, M., Braekman, J. C. & Dupont, A. 1983 Salicin from host plant as precursor of salicyl aldehyde in defensive secretion of chrysomeline larvae. *Physiol. Entomol.* **8**, 307–314.

12 Brueckmann, M., Termonia, A., Pasteels, J. M. & Hartmann, T. 2002 Characterization of an extracellular salicyl alcohol oxidase from larval defensive secretions of *Chrysomela populi* and *Phratora vitellinae* (Chrysomelina). *Insect Biochem. Mol. Biol.* **32**, 1517–1523. (doi:10.1016/S0965-1748(02)00072-3)

13 Kirsch, R., Vogel, H., Muck, A., Reichwald, K., Pasteels, J. M. & Boland, W. 2011 Host plant shifts affect a major defense enzyme in *Chrysomela lapponica*. *Proc. Natl Acad. Sci. USA* **108**, 4897–4901. (doi:10.1073/pnas.1013846108)

14 Michalski, C., Mohagheghi, H., Nimtz, M., Pasteels, J. M. & Ober, D. 2008 Salicyl alcohol oxidase of the chemical defense secretion of two chrysomelid leaf beetles—molecular and functional characterization of two new members of the glucose–methanol–choline oxidoreductase gene family. *J. Biol. Chem.* **283**, 19 219–19 228. (doi:10.1074/jbc.M802236200)

15 Wallace, J. B. & Blum, M. S. 1969 Refined defensive mechanisms in *Chrysomela scripta*. *Ann. Entomol. Soc. Am.* **62**, 503–506.

16 Gross, J., Podsiadlowski, L. & Hilker, M. 2002 Antimicrobial activity of exocrine glandular secretion of *Chrysomela* larvae. *J. Chem. Ecol.* **28**, 317–331. (doi:10.1023/a:1017934124650)

17 Rowell-Rahier, M. & Pasteels, J. M. 1986 Economics of chemical defense in Chrysomelinae. *J. Chem. Ecol.* **12**, 1189–1203. (doi:10.1007/bf01639004)

18 Kuhn, J., Pettersson, E. M., Feld, B. K., Burse, A., Termonia, A., Pasteels, J. M. & Boland, W. 2004 Selective transport systems mediate sequestration of plant glucosides in leaf beetles: a molecular basis for adaptation and evolution. *Proc. Natl Acad. Sci. USA* **101**, 13 808–13 813. (doi:10.1073/pnas.0402576101)

19 Discher, S., Burse, A., Tolzin-Banasch, K., Heinemann, S. H., Pasteels, J. M. & Boland, W. 2009 A versatile transport network for sequestering and excreting plant glycosides in leaf beetles provides an evolutionary flexible defense strategy. *ChemBioChem* **10**, 2223–2229. (doi:10.1002/cbic.200900226)

20 Pasteels, J. M., Duffey, S. & Rowell-Rahier, M. 1990 Toxins in chrysomelid beetles possible evolutionary sequence from de novo synthesis to derivation from food-plant chemicals. *J. Chem. Ecol.* **16**, 211–222. (doi:10.1007/bf01021280)

21 Termonia, A., Hsiao, T. H., Pasteels, J. M. & Milinkovitch, M. C. 2001 Feeding specialization and host-derived chemical defense in Chrysomeline leaf beetles did not lead to an evolutionary dead end. *Proc. Natl Acad. Sci. USA* **98**, 3909–3914. (doi:10.1073/pnas.061034598)

22 Daloze, D. & Pasteels, J. M. 1994 Isolation of 8-hydroxygeraniol-8-O-beta-D-glucoside, a probable intermediate in biosynthesis of iridoid monoterpenes, from defensive secretions of *Plagiodera versicolora* and *Gastrophysa viridula* (Coleoptera, Chrysomelidae). *J. Chem. Ecol.* **20**, 2089–2097. (doi:10.1007/BF02066245)

23 Soetens, P., Pasteels, J. M. & Daloze, D. 1993 A simple method for *in vivo* testing of glandular enzymatic activity on potential precursors of larval defensive compounds in *Phratora* species (Coleoptera, Chrysomelinae). *Experientia* **49**, 1024–1026. (doi:10.1007/BF02125653)

24 Veith, M., Oldham, N. J., Dettner, K., Pasteels, J. M. & Boland, W. 1997 Biosynthesis of defensive allomones in leaf beetle larvae—stereochemistry of salicylalcohol oxidation in *Phratora vitellinae* and comparison of enzyme substrate and stereospecificity with alcohol oxidases from several iridoid producing leaf beetles. *J. Chem. Ecol.* **23**, 429–443. (doi:10.1023/B:JOEC.0000006369.26490.c6)

25 Veith, M., Lorenz, M., Boland, W., Simon, H. & Dettner, K. 1994 Biosynthesis of iridoid monoterpenes in insects—defensive secretions from larvae of leaf beetles (Coleoptera, Chrysomelidae). *Tetrahedron* **50**, 6859–6874. (doi:10.1016/S0040-4020(01)81338-7)

26 Kirsch, R., Vogel, H., Muck, A., Vilcinskas, A., Pasteels, J. M. & Boland, W. 2011 To be or not to be convergent in salicin-based defence in chrysomeline leaf beetle larvae: evidence from *Phratora vitellinae* salicyl alcohol oxidase. *Proc. R. Soc. B* **278**, 3225–3232. (doi:10.1098/rspb.2011.0175)

27 Ding, S.-W., Li, H., Lu, R., Li, F. & Li, W.-X. 2004 RNA silencing: a conserved antiviral immunity of plants and animals. *Virus Res.* **102**, 109–115. (doi:10.1016/j.virusres.2004.01.021)

28 Fire, A., Xu, S. Q., Montgomery, M. K., Kostas, S. A., Driver, S. E. & Mello, C. C. 1998 Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811. (doi:10.1038/35888)

29 Belles, X. 2010 Beyond *Drosophila*: RNAi *in vivo* and functional genomics in insects. *Annu. Rev. Entomol.* **55**, 111–128. (doi:10.1146/annurev-ento-112408-085301)

30 Mito, T., Nakamura, T., Bando, T., Ohuchi, H. & Noji, S. 2011 The advent of RNA interference in entomology. *Entomol. Sci.* **14**, 1–8. (doi:10.1111/j.1479-8298.2010.00408.x)

31 Terenius, O. *et al.* 2011 RNA interference in Lepidoptera: an overview of successful and unsuccessful studies and implications for experimental design. *J. Insect Physiol.* **57**, 231–245. (doi:10.1016/j.jinsphys.2010.11.006)

32 Szabo, L. F. 2008 Rigorous biogenetic network for a group of indole alkaloids derived from strictosidine. *Molecules* **13**, 1875–1896. (doi:10.3390/molecules13081875)

33 Kuhnle, A. & Muller, C. 2011 Responses of an oligophagous beetle species to rearing for several generations on alternative host–plant species. *Ecol. Entomol.* **36**, 125–134. (doi:10.1111/j.1365-2311.2010.01256.x)

34 Naito, Y., Yamada, T., Matsumiya, T., Ui-Tei, K., Saigo, K. & Morishita, S. 2005 dsCheck: highly sensitive off-target search software for double-stranded RNA-mediated RNA interference. *Nucleic Acids Res.* **33**, W589–W591. (doi:10.1093/nar/gki419)

35 Altschul, S. F., Madden, T. L., Schaeffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402. (doi:10.1093/nar/25.17.3389)

36 Oldham, N. J., Veith, M., Boland, W. & Dettner, K. 1996 Iridoid monoterpene biosynthesis in insects—evidence for a de novo pathway occurring in the defensive glands of *Phaedon armoraciae* (Chrysomelidae) leaf beetle larvae. *Naturwissenschaften* **83**, 470–473. (doi:10.1007/BF01144016)

37 Uesato, S., Ogawa, Y., Doi, M. & Inouye, H. 1987 Biomimetic synthesis of (±)-chrysomelidial, (±)-dehydroiridoidal, and (±)-iridoidal. *J. Chem. Soc.-Chem. Commun.* 1020–1021. (doi:10.1039/c39870001020)

38 Cavener, D. R. 1992 GMC oxidoreductases: a newly defined family of homologous proteins with diverse catalytic activities. *J. Mol. Biol.* **223**, 811–814. (doi:10.1016/0022-2836(92)90992-s)

39 Jackson, A. L., Bartz, S. R., Schelter, J., Kobayashi, S. V., Burchard, J., Mao, M., Li, B., Cavet, G. & Linsley, P. S. 2003 Expression profiling reveals off-target gene

regulation by RNAi. *Nat. Biotechnol.* **21**, 635–637. (doi:10.1038/nbt831)
40 Seinen, E., Burgerhof, J. G. M., Jansen, R. C. & Sibon, O. C. M. 2011 RNAi-induced off-target effects in *Drosophila melanogaster*: frequencies and solutions. *Brief. Funct. Genom.* **10**, 206–214. (doi:10.1093/bfgp/elr017)
41 Morris, K. V., Chan, S. W. L., Jacobsen, S. E. & Looney, D. J. 2004 Small interfering RNA-induced transcriptional gene silencing in human cells. *Science* **305**, 1289–1292. (doi:10.1126/science.1101372)
42 Bosher, J. M., Dufourcq, P., Sookhareea, S. & Labouesse, M. 1999 RNA interference can target pre-mRNA: consequences for gene expression in a *Caenorhabditis elegans* operon. *Genetics* **153**, 1245–1256.
43 Vermunt, A. M. W., Kamimura, M., Hirai, M., Kiuchi, M. & Shiotsuki, T. 2001 The juvenile hormone binding protein of silkworm haemolymph: gene and functional analysis. *Insect Mol. Biol.* **10**, 147–154. (doi:10.1046/j.1365-2583.2001.00249.x)
44 Hamiaux, C., Stanley, D., Greenwood, D. R., Baker, E. N. & Newcomb, R. D. 2009 Crystal structure of *Epiphyas postvittana* takeout 1 with bound ubiquinone supports a role as ligand carriers for takeout proteins in insects. *J. Biol. Chem.* **284**, 3496–3503. (doi:10.1074/jbc.M807467200)
45 So, W. V., Sarov-Blat, L., Kotarski, C. K., McDonald, M. J., Allada, R. & Rosbash, M. 2000 Takeout, a novel *Drosophila* gene under circadian clock transcriptional regulation. *Mol. Cell. Biol.* **20**, 6935. (doi:10.1128/MCB.20.18.6935-6944.2000)
46 deKort, C. A. D. & Granger, N. A. 1996 Regulation of JH titers: the relevance of degradative enzymes and binding proteins. *Arch. Insect Biochem. Physiol.* **33**, 1–26. (doi:10.1002/(sici)1520-6327(1996)33:1<1::aid-arch1>3.0.co;2-2)
47 Prestwich, G. D., Wojtasek, H., Lentz, A. J. & Rabinovich, J. M. 1996 Biochemistry of proteins that bind and metabolize juvenile hormones. *Arch. Insect Biochem. Physiol.* **32**, 407–419. (doi:10.1002/(sici)1520-6327(1996)32:3/4<407::aid-arch13>3.0.co;2-g)
48 Suzuki, R., Fujimoto, Z., Shiotsuki, T., Tsuchiya, W., Momma, M., Tase, A., Miyazawa, M. & Yamazaki, T. 2011 Structural mechanism of JH delivery in hemolymph by JHBP of silkworm, *Bombyx mori*. *Sci. Rep.* **1**, 133. (doi:10.1038/srep00133)
49 Winiarska, B., Dwornik, A., Debski, J., Grzelak, K., Bystranowska, D., Zalewska, M., Dadlez, M., Ozyhar, A. & Kochman, M. 2011 N-linked glycosylation of *G. mellonella* juvenile hormone binding protein—Comparison of recombinant mutants expressed in *P. pastoris* cells with native protein. *Biochim. Biophys. Acta: Proteins Proteom.* **1814**, 610–621. (doi:10.1016/j.bbapap.2011.02.002)
50 Ohnishi, A., Hashimoto, K., Imai, K. & Matsumoto, S. 2009 Functional characterization of the *Bombyx mori* fatty acid transport protein (BmFATP) within the silkmoth pheromone gland. *J. Biol. Chem.* **284**, 5128–5136. (doi:10.1074/jbc.M806072200)
51 Ohnishi, A., Hull, J. J. & Matsumoto, S. 2006 Targeted disruption of genes in the *Bombyx mori* sex pheromone biosynthetic pathway. *Proc. Natl Acad. Sci. USA* **103**, 4398–4403. (doi:10.1073/pnas.0511270103)
52 Abdel-Latief, M., Garbe, L. A., Koch, M. & Ruther, J. 2008 An epoxide hydrolase involved in the biosynthesis of an insect sex attractant and its use to localize the production site. *Proc. Natl Acad. Sci. USA* **105**, 8914–8919. (doi:10.1073/pnas.0801559105)
53 Weibel, D. B., Oldham, N. J., Feld, B., Glombitza, G., Dettner, K. & Boland, W. 2001 Iridoid biosynthesis in staphylinid rove beetles (Coleoptera : Staphylinidae, Philonthinae). *Insect Biochem. Mol. Biol.* **31**, 583–591. (doi:10.1016/s0965-1748(00)00163-6)
54 Sabirzhanov, B., Sabirzhanova, I. B. & Keifer, J. 2011 Screening target specificity of siRNAs by rapid amplification of cDNA ends (RACE) for non-sequenced species. *J. Mol. Neurosci.* **44**, 68–75. (doi:10.1007/s12031-011-9514-6)
55 Lew-Tabor, A. E., Kurscheid, S., Barrero, R., Gondro, C., Moolhuijzen, P. M., Valle, M. R., Morgan, J. A. T., Covacin, C. & Bellgard, M. I. 2011 Gene expression evidence for off-target effects caused by RNA interference-mediated gene silencing of *Ubiquitin-63E* in the cattle tick *Rhipicephalus microplus*. *Int. J. Parasitol.* **41**, 1001–1014. (doi:10.1016/j.ijpara.2011.05.003)
56 Qiu, S., Adema, C. M. & Lane, T. 2005 A computational study of off-target effects of RNA interference. *Nucleic Acids Res.* **33**, 1834–1847. (doi:10.1093/nar/gki324)
57 Reynolds, A., Leake, D., Boese, Q., Scaringe, S., Marshall, W. S. & Khvorova, A. 2004 Rational siRNA design for RNA interference. *Nat. Biotechnol.* **22**, 326–330. (doi:10.1038/nbt936)

3.2 MANUSCRIPT 2: PUTATIVE SUGAR TRANSPORTERS OF THE MUSTARD LEAF BEETLE PHAEDON COCHLEARIAE: THEIR PHYLOGENY AND ROLE FOR NUTRIENT SUPPLY IN LARVAL DEFENSIVE GLANDS

OPEN ACCESS Freely available online

PLOS ONE

Putative Sugar Transporters of the Mustard Leaf Beetle *Phaedon cochleariae*: Their Phylogeny and Role for Nutrient Supply in Larval Defensive Glands

Magdalena Stock[1], René R. Gretscher[1], Marco Groth[2], Simone Eiserloh[1], Wilhelm Boland[1], Antje Burse[1]*

1 Department of Bioorganic Chemistry, Max Planck Institute for Chemical Ecology, Jena, Thuringia, Germany, **2** Genome Analysis Group, Leibniz Institute for Age Research – Fritz Lipmann Institute, Jena, Thuringia, Germany

OPEN ACCESS Freely available online

PLOS ONE

Putative Sugar Transporters of the Mustard Leaf Beetle *Phaedon cochleariae*: Their Phylogeny and Role for Nutrient Supply in Larval Defensive Glands

Magdalena Stock[1], René R. Gretscher[1], Marco Groth[2], Simone Eiserloh[1], Wilhelm Boland[1], Antje Burse[1]*

1 Department of Bioorganic Chemistry, Max Planck Institute for Chemical Ecology, Jena, Thuringia, Germany, **2** Genome Analysis Group, Leibniz Institute for Age Research – Fritz Lipmann Institute, Jena, Thuringia, Germany

Abstract

Background: Phytophagous insects have emerged successfully on the planet also because of the development of diverse and often astonishing defensive strategies against their enemies. The larvae of the mustard leaf beetle *Phaedon cochleariae*, for example, secrete deterrents from specialized defensive glands on their back. The secretion process involves ATP-binding cassette transporters. Therefore, sugar as one of the major energy sources to fuel the ATP synthesis for the cellular metabolism and transport processes, has to be present in the defensive glands. However, the role of sugar transporters for the production of defensive secretions was not addressed until now.

Results: To identify sugar transporters in *P. cochleariae*, a transcript catalogue was created by Illumina sequencing of cDNA libraries. A total of 68,667 transcripts were identified and 68 proteins were annotated as either members of the solute carrier 2 (SLC2) family or trehalose transporters. Phylogenetic analyses revealed an extension of the mammalian GLUT6/8 class in insects as well as one group of transporters exhibiting distinctive conserved motifs only present in the insect order Coleoptera. RNA-seq data of samples derived from the defensive glands revealed six transcripts encoding sugar transporters with more than 3,000 counts. Two of them are exclusively expressed in the glandular tissue. Reduction in secretions production was accomplished by silencing two of four selected transporters. RNA-seq experiments of transporter-silenced larvae showed the down-regulation of the silenced transporter but concurrently the up-regulation of other SLC2 transporters suggesting an adaptive system to maintain sugar homeostasis in the defensive glands.

Conclusion: We provide the first comprehensive phylogenetic study of the SLC2 family in a phytophagous beetle species. RNAi and RNA-seq experiments underline the importance of SLC2 transporters in defensive glands to achieve a chemical defense for successful competitive interaction in natural ecosystems.

Citation: Stock M, Gretscher RR, Groth M, Eiserloh S, Boland W, et al. (2013) Putative Sugar Transporters of the Mustard Leaf Beetle *Phaedon cochleariae*: Their Phylogeny and Role for Nutrient Supply in Larval Defensive Glands. PLoS ONE 8(12): e84461. doi:10.1371/journal.pone.0084461

Editor: Claude Wicker-Thomas, CNRS, France

Received September 26, 2013; **Accepted** November 22, 2013; **Published** December 31, 2013

Copyright: © 2013 Stock et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors' work was funded by the Max Planck Society (http://www.mpg.de/en) and the Deutsche Forschungsgemeinschaft (grand number BU1862/2-1; http://www.dfg.de/en/index.jsp). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: aburse@ice.mpg.de

Introduction

Sugar sweetens our life. Glucose is one of the major energy sources and an important substrate for both protein and lipid synthesis. Its catabolism fuels cellular respiration for ATP production. For glucose absorption in the mammalian small intestine two pathways are known: the passive, paracellular absorption which bats rely on for more than 70% of their total sugar uptake [1] and the transporter-mediated transcellular pathways which non-flying mammals use preferentially. They take up glucose from interstitial fluid by a passive, facilitative transport process driven by the downward concentration gradient across the plasma membrane [2]. Exclusively in the epithelial cell brush border of the small intestine and the kidney proximal convoluted tubules, glucose is absorbed or reabsorbed against its electrochemical gradient by a secondary active transport mechanism which uses the sodium concentration gradient established by Na^+/K^+/ATP pumps [3]. Unlike glucose in mammals, the major blood (hemolymph) sugar of insects is often the disaccharide trehalose (α-D-glucopyranosyl-α-D-glucopyranoside) [4–6]. It is synthesized from glucose phosphates in fat body tissue and serves as a source of carbohydrates for various tissues including flight muscles [7,8], intestinal tract [9], fat body [10] or ovaries [11]. Besides trehalose absorption, the absorption of other sugars, such as fructose, glucose, and galactose, has been shown for different insect tissues [12–17]. Only few examples of sugar transport proteins from insects have been functionally characterized to date [17–23]. Except for SCRT, which was classified as a member of family 49 of solute carriers (SLCs) [17], they all belong to the SLC2 family of glucose and polyol transporters [24–27]. In principle, SLC2 proteins are integral membrane proteins exhibiting a predicted twelve transmembrane domain topology. The so-called 'sugar transport signatures' that signify substrate binding and catalytic activity are also conserved in the SLC2 family. Usually, the expression of these proteins is tissue-specific and

responds to metabolic and hormonal regulation. Each of the transporters possesses different affinities for sugars [28].

Although beetles (Coleoptera) represent one of the most diversified lineages on earth with about 350,000 species described and total numbers probably an order of magnitude higher, SLC2 sugar transporters of beetles have not been in the focus of researchers so far. In particular, Chrysomelidae (leaf beetles) constitute, together with the Cerambycidae (longhorn beetles) and the Curculionoidea (weevils), the largest beetle radiation, namely "Phytophaga", which represent about 40% of all known species [29]. Among phytophagous beetles are many pests such as the mustard leaf beetle *Phaedon cochleariae* which causes yield losses on cruciferous crops in Europe [30]. This species is adapted to use host plants' leaves as its food source for the duration of its life [31,32]. Due to its life in an exposed environment, it has to be protected against both the noxious effect of plant secondary metabolites and the attacks by its predators and parasitoids.

When it comes to producing defensive compounds to repel their omnipresent enemies, insect in general are very innovative [33]. To circumvent auto-intoxicative effects, these natural products frequently are processed in exocrine glands [34–36]. *P. cochleariae* is known to produce defensive compounds in such glands, herein further referred to as defensive glands, in the larval as well as in the adult stage [37–41]. The juvenile beetles possess nine pairs of these glands on their back and release deterrent secretions upon disturbance [42,43]. Each of these glands is composed of several secretory cells which are attached to a large reservoir. The anti-predatory effect of the secretions can be attributed to cyclopentanoid monoterpenes (iridoids) which are synthesized within the reservoir by few enzymatic reactions from a pre-toxin which is made in the fat body and from there transferred into the defensive glands [44,45]. Previously it has been shown that defensive gland cells possess ATP-binding cassette (ABC) transporters which are crucial for the shuttling of pre-toxins into the reservoir [46]. Because ATP is used for the cellular metabolism and transport process into the defensive glands, sugars need to be delivered by transporters to drive ATP production in this tissue. Although defensive secretions from phytophagous insects are key players in trophic interactions found in terrestrial food webs [47], their production pathways are often not fully understood. Sugar transporters may be essential components to fuel energy in insect defensive glands, however, nothing is known about their *in vivo* relevance for deterrent production.

Here we focus on a first catalogue of putative members of the SLC2 family as well as trehalose transporters for a phytophagous leaf beetle species. By means of a *de novo*-assembled transcriptome created from the mRNA of *P. cochleariae*, we have performed comprehensive and statistically supported phylogenetic analyses of the identified sequences together with their orthologs selected from other insects and other Metazoa including the known human glucose transporter (GLUT) isoforms. Our data revealed an enormous expansion of the GLUT6/8 sister group in insects and a clade of sequences unique for beetles. Subsequent next generation sequencing-based expression studies revealed putative SLC2 transcripts highly expressed in the defensive glands of juvenile *P. cochleariae*. Single silencing of selected SLC2-candidates by RNA interference (RNAi) resulted only in two cases in a reduced production of defensive secretions. However, in two other cases silencing did not affect deterrent production suggesting an adaptive backup system which stabilizes the sugar level in the defensive glands. To prove the observed homeostasis, we subsequently sequenced the mRNA of silenced larvae to study the differential expression of putative SLC2 transporters not only on phenotypic but also on a transcriptional level.

Materials and Methods

Beetles

P. cochleariae (F.) was lab-reared on *Brassica rapa chinensis* in a light/dark cycle of 16 h light and 8 h darkness, 14°C±1°C in light and 12°C±1°C in darkness.

RNA Isolation, Library Construction and Sequencing

Total RNA was isolated from tissue samples from *P. cochleariae* larvae as described by Bodemann *et al.* [48]. Up to 5 µg of total RNA was then used for library preparation using TruSeq™ RNA Sample Prep Kit according to manufacturer's description. RNA sequencing (RNA-seq) was done using Illumina next-generation sequencing technique [49] on a HiSeq2000 (Illumina, Inc., San Diego, California USA) in 50 bp single-read mode (two or three samples multiplexed in one lane). Pooled total RNAs from the different tissues (such as defensive glands, gut, fat body, Malpighian tubules, and head) from larval *P. cochleariae* were used for paired-end sequencing. Therefore, the fragmentation step during library preparation of the pooled total RNAs was shortened to four minutes (seven minutes for all the remaining samples) to obtain longer fragments. This library was sequenced using a GAIIx (Illumina, Inc., San Diego, California USA) in 150 bp paired-end mode in one sample per lane. All reads were extracted in FastQ format using CASAVA v1.7 (HiSeq) and v1.8 (GAIIx) (Illumina, Inc., San Diego, California USA) for further analysis.

Subsequently to the RNAi experiments (described below), additional sequencing was carried out. Four samples of glandular tissue (one sample of *gfp* dsRNA-injected larvae and three samples of larvae injected with dsRNA of three highly abundant sugar transporters) – two biological replicates each, have been prepared as mentioned above and sequenced on a HiSeq2500 (Illumina, Inc., San Diego, California USA) in 50 bp single-read mode (multiplexed in one lane). All reads were extracted in FastQ format using bcl2fastq v.1.8.3 (Illumina, Inc., San Diego, California USA). The raw sequence data are listed in Table 1 and are stored in the Sequence Read Archive at NCBI (cDNA library 1: SRA100673; cDNA libraries 2: SRA106118, SRA106122, SRA106161; cDNA libraries 3: SRA108012, SRA108036; cDNA libraries 4: SRA108037, SRA108041; cDNA libraries 5: SRA109958, SRA109964; cDNA libraries 6: SRA109966, SRA109967). The corresponding BioProject for *P. cochleariae* can be accessed at NCBI homepage (BioProject ID: PRJNA210148).

Transcriptome *de novo* Assembly

The paired-end reads were assembled using Trinity, a RNA-seq *de novo* assembly software [50,51] with default parameters, minimal contig length of 300 bp and paired fragment length of 500 bp. Afterwards, the *de novo* assembled transcripts were reassembled using the TGI Clustering tools (TGICL), a software program to cluster large EST datasets [52]. The clustering step is performed by NCBI's megablast [53] and the resulting clusters are then assembled using CAP3 assembly program [54] with following parameters: minimum overlap length of 100 bp and sequence similarity of 90 percent.

Pfam Analysis

The *de novo* transcripts were translated into their possible protein sequences (all six reading frames) by applying the transeq script which is part of the EMBOSS package (http://imed.med.ucm.es/EMBOSS/). Thereafter, the script pfam-scan.pl (downloaded from the ftp://ftp.sanger.ac.uk/pub/databases/Pfam/Tools/site) was used with showing overlapping hits within clan member families in addition to default parameters to search the protein

Table 1. Overview of the raw sequence data.

cDNA library	Tissues for RNA isolation	Total number of reads	Sequencing mode	Remarks
1	gut, fat body, glands, Malpighian tubules	46,030,279	GAIIx 2×150 bp	–
2	glands (3 replicates)	101,383,127	50 bp, HiSeq2000	–
3	glands (2 replicates)	33,013,829	50 bp, HiSeq2500	dsRNA-*gfp* injected
4	glands (2 replicates)	50,926,062	50 bp, HiSeq2500	dsRNA-*Pcsut1* injected
5	glands (2 replicates)	59,678,392	50 bp, HiSeq2500	dsRNA-*Pcsut2* injected
6	glands (2 replicates)	53,204,574	50 bp, HiSeq2500	dsRNA-*Pcsut6* injected

The table exhibits the RNA derived specimens, number of reads, sequencing technology and sequencing mode.
doi:10.1371/journal.pone.0084461.t001

sequences against the Pfam-A database which consists of high quality protein families based on profile HMMs and clans. A clan is a collection of Pfam-A entries which are related by similarity of sequence, structure or profile-HMM [55,56].

Identification of SLC2 Sequences and Trehalose Transporters

The SLC2 sequences and trehalose transporters were identified by searching the Pfam results for Sugar_tr hits. Sugar (and other) transporters are part of the major facilitator superfamily [25] which is believed to function primarily in the uptake of sugars. All identified sugar transporters were searched *via* BLASTp with an E-value smaller than 1e-3 against the Swiss-Prot protein database (ftp://ftp.ncbi.nlm.nih.gov/blast/db/). Swiss-Prot is a high quality and manually annotated and reviewed, non-redundant protein sequence database [57,58]. For each sugar transporter the top ten hits were inspected, and sequences homologous to known SLC2 or trehalose transporters were identified. SLC2 and trehalose transporters of full-length transcripts or those having a coding sequence of at least 900 bp of length were chosen for further analysis. All studied sugar transporter transcripts are stored in the GenBank database at NCBI as either mRNA (accession numbers: KF803259–KF803269) or Transcriptome Shotgun Assembly (TSA) sequences. This TSA project has been deposited at DDBJ/EMBL/GenBank under the accession GAPU00000000. The version described in this paper is the first version, GAPU01000000. All those sugar transporters were then observed by applying TMHMM [59] and Memsat2 [60] to predict their 12 transmembrane (TM) domains.

Calculation of Phylogenetic Trees

Phylogenetic trees were calculated for identified SLC2 encoding sequences (see above). Amino acid sequences in multi-FASTA format were aligned using Probalign version 1.4 [61] with default parameters or MAFFT version 7.023b with following settings: –maxiterate 1000 and –localpair. Thereafter, phylogenetic trees were calculated using MrBayes [62,63] version 3.2.1 and RAxML [64] version 7.2.8 which use different methods. Namely, MrBayes uses Bayesian inference of phylogeny, and RAxML uses maximum likelihood estimation. Applying MrBayes the following settings were used: number of generations was set to 300,000, samplefreq and printfreq were set to 100, the number of runs was set to 2, and the type to calculate the consensus tree was set to allcompat. Applying RAxML the following parameters were used: the model of substitution was PROTGAMMAJTT (GAMMA model of rate heterogeneity), and 1000 rapid bootstrap inferences were done.

Table 2. Characterized trehalose as well as glucose/fructose transporters.

	Description	NCBI accession	Organism
AaSUT_XP_001664193	sugar transporter	XP_001664193.1	*Aedes aegypti*
AgTRET1_BAF96742	trehalose transporter AgTRET1	BAF96742.1	*Anopheles gambia*
AmTRET1_NP_001107211	trehalose transporter 1	NP_001107211.1	*Apis mellifera*
ApTret1like_XP_001950697	fac trehalose transporter Tret1-like	XP_001950697.1	*Acyrthosiphon pisum*
BmTRET1_NP_001108344	fac trehalose transporter Tret1	NP_001108344.1	*Bombyx mori*
DmTRET1–1A	trehalose transporter 1–1, isoform A	NP_610693.1	*Drosophila melanogaster*
DmTRET1–1B	trehalose transporter 1–1, isoform B	NP_725068.1	*Drosophila melanogaster*
DmTRET1–2	trehalose transporter 1–2, isoform A	NP_610694.1	*Drosophila melanogaster*
NlHT1_ABM01870	fac hexose transporter 1	ABM01870.1	*Nilaparvata lugens*
NlSUT1	sugar transporter 1	BAI83415.1	*Nilaparvata lugens*
NlSUT6	sugar transporter 6	BAI83420.1	*Nilaparvata lugens*
NlSUT8	sugar transporter 8	BAI83422.1	*Nilaparvata lugens*
PvTRET1_A5LGM7	fac trehalose transporter Tret1	A5LGM7.1	*Polypedilum vanderplanki*
SiGLUT8	glucose transporter 8	AAX92638.1	*Solenopsis invicta*

They are listed with their description, accession number and organism. Those transporters were added to *P. cochleariae*'s chosen sequences to calculate phylogenetic trees.
doi:10.1371/journal.pone.0084461.t002

Table 3. Number of assembled transcripts and average length after assembly and reassembly showing the usefulness of reassembling.

	Number of transcripts	Sum_length	25th_pc	75th_pc	Ave_length
After Trinity assembly	107323	85,475,541 bp	379 bp	841 bp	796 bp
After reassembly with TGICL	68667	63,815,627 bp	383 bp	1037 bp	929 bp

doi:10.1371/journal.pone.0084461.t003

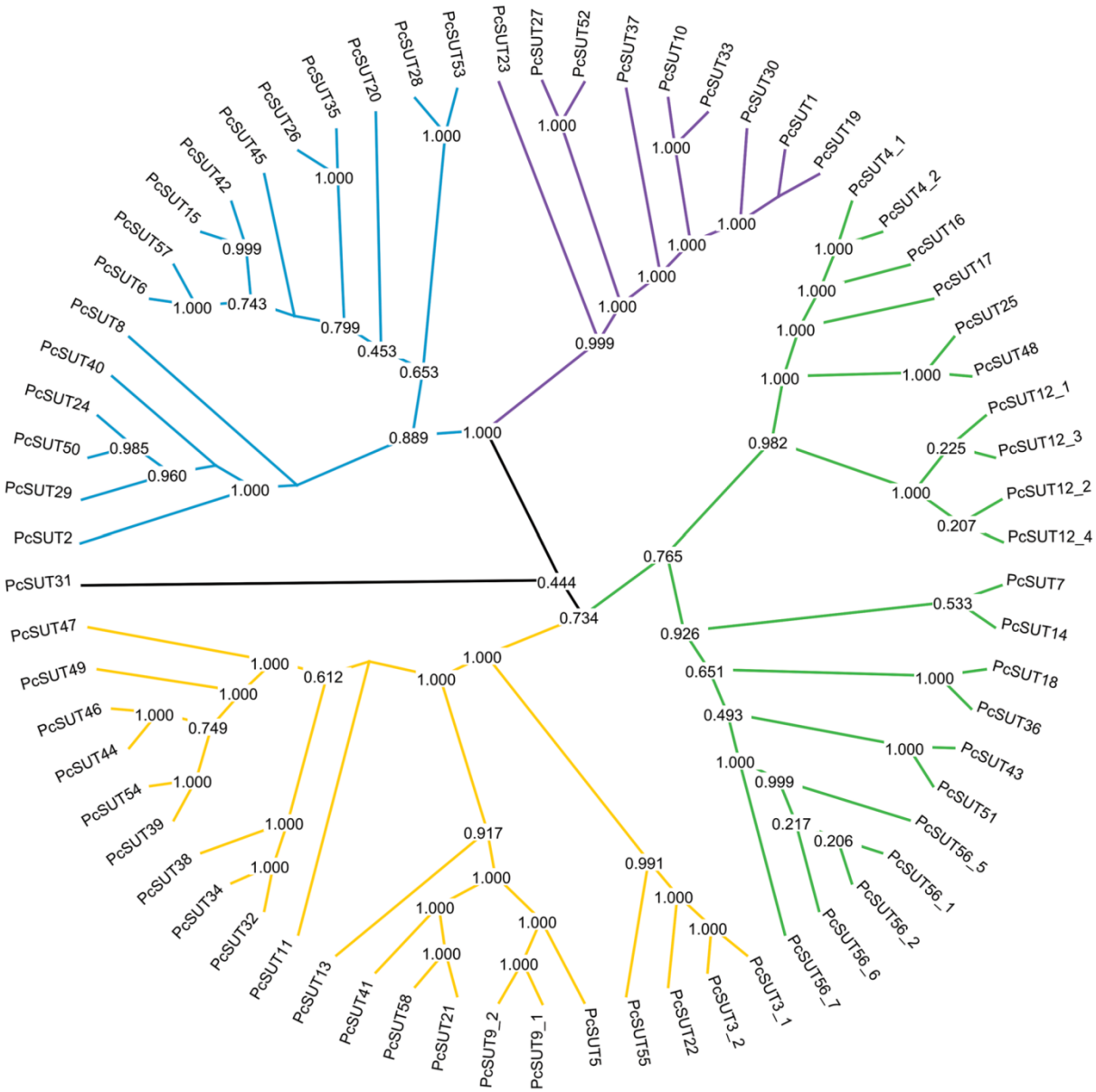


Figure 1. Phylogenetic tree of the 68 chosen sugar transporters derived from *P. cochleariae*. This circular phylogram shows the main 4 groups of chosen sugar transporters. Tree was calculated using MrBayes.
doi:10.1371/journal.pone.0084461.g001

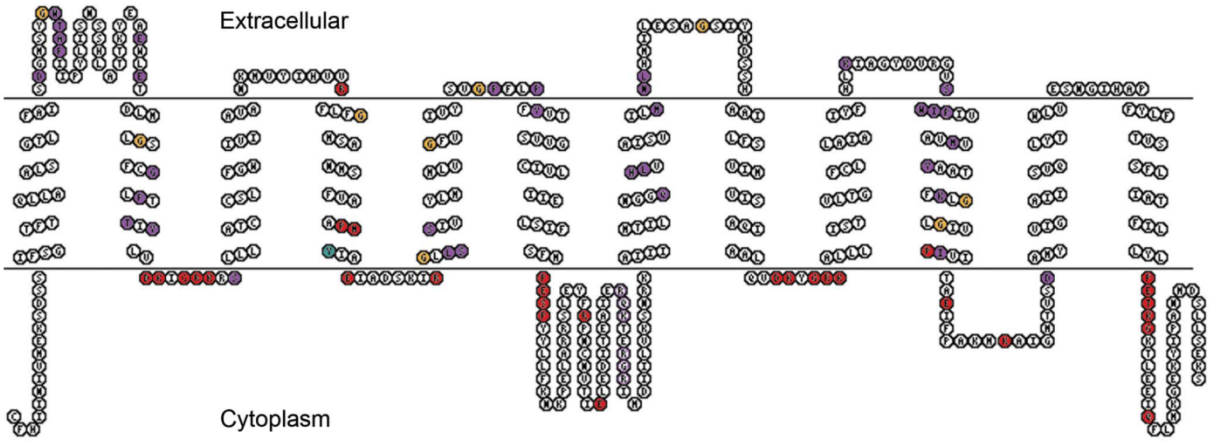


Figure 2. Schematic model for the structure of the putative SLC2 transporters derived from *P. cochleariae* by means of *Pc*SUT1. All 4 groups show the known and conserved facilitated sugar transporter motifs, such as DRxGRR/K in the second loop, PESPR/K in the sixth loop, E/DRxGRR/K in loop 8, and PETK/RGK/R in the carboxy terminal [2,15,18,76,86]. Furthermore, there are conserved amino acids, such as E and R in loop 4 and 10 (red). Those are needed for the glucose transport activity. Conserved tyrosines (turquois), such as the PMY motif mentioned by Chen *et al.* [15], can be found in our sequences in TMD 4. Additionally, conserved glycines (yellowish) in TMD 1, 2, 4, 5, 7, 8, and 10 as well as in loop 2 and 7 are present, characteristic for the mammalian glucose transporter family. The purple branch exhibits a GWTAP motif in loop 1, a PFYV motif in loop 5, and a VILMNLH motif in TMD 10 (purple colored amino acids).
doi:10.1371/journal.pone.0084461.g002

For every calculation of the phylogenetic tree only conserved parts were taken into account. Thus, the N-terminus, the loop between TM6 and TM7 as well as the C-terminus were excluded from calculating the phylogenetic tree.

First of all, the trees were calculated only for the chosen sequences derived from *P. cochleariae* to divide these into groups with respect to their sequence similarities. Second, sequences originating from other insects (*Nilaparvata lugens*, *Solenopsis invicta*, *Acyrthosiphon pisum*, *Polypedilum vanderplanki*, *Bombyx mori*) that have been functionally characterized were added to these sequences (see Table 2) [15,18,19,22]. Trees were re-calculated to identify sequences of *P. cochleariae* that are homologous to the characterized sequences. Third, the human GLUTs as well as other homologous sugar transporters from other Metazoa were selected. The chosen sequences derived from *P. cochleariae* were added to this selection and phylogenetic trees were calculated again to investigate the organisms' distribution in the trees. Furthermore, we calculated trees for one specific branch (which separated from the others with bootstrap percentage of 100%). For this, orthologous sequences derived from other beetles namely *Dendroctonus ponderosae* [65] and *Tribolium castaneum* were adjoined.

Gene Expression Profiling and Real-time PCR Validation of Putative Sugar Transporters in the Defensive Glands

Three replicates of the cDNA of *P. cochleariae*'s glands were prepared and sequenced as described above. All short reads of three replicates of glandular tissue were mapped onto the reassembled transcripts using Bowtie, an ultrafast short read aligner [66] with –best and –strata options. Bioconductor is an open source, open development software project to provide tools for the analysis and comprehension of high-throughput genomic data. It is based primarily on the R programming language [67]. The mapping results in bowtie format were loaded into R statistics using the ShortRead package [68] which is part of the Bioconductor package. For estimation of variance-mean dependence in count data, the DESeq package was used which is also part of the Bioconductor package (release 2.11) [69,70]. After analyzing the transcripts with DESeq, the sequences encoding the sugar transporters were selected and sorted regarding the number of sequence counts, beginning with the most abundant reads present in the glands. Furthermore, we normalized the normalized counts to the standards used for quantitative real-time PCR (*Pcrpl6* and *Pcrps3*) by first calculating the normalization factor [71]. The normalized counts were then divided by this normalization factor to get fold changes comparable to the values resulting from quantitative real-time PCR experiments.

Real-time PCR was employed for relative quantification [71]. RNA was isolated as described above. Up to 5 µg of the RNA was reverse transcribed at 50°C for 60 min using SuperScript III and Oligo(dT)$_{12–18}$ primer (life technologies, Darmstadt, Germany). Two technical replicates were performed from three biological replicates each. Technical replicates with a Cq difference of >1 were repeated. To normalize the PCRs for the amount of cDNA template added to the reactions, *Pcrpl6* and *Pcrps3* were chosen for *P. cochleariae* as reference genes. Primers were designed using primer3PLUS: http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi (see Table S1 for primer sequences). Quantitative real-time PCR data were acquired on the CFX96 Touch Real-Time PCR Detection System (Bio-Rad Laboratories GmbH, Munich, Germany) using SYBR Premix Ex Taq II (Tli RNase H Plus) (Takara Bio Inc., Otsu, Japan). Running conditions: 3′ 94°C, 40 cycles [30″ 94°C; 30″ 60°C], melting curve with 1°C increase 60–95°C. These assays were performed following the MIQE-guidelines [72].

RNA Interference in *P. cochleariae* Larvae

The open reading frames encoding putative transporters of four highly abundant transcripts (*Pcsut1*, *Pcsut2*, *Pcsut5* and *Pcsut6*) were cloned into T7-promotor site lacking TOPO-plasmids pIBV5/HIS (life technologies, Darmstadt, Germany). Plasmids were sequenced prior to further processing. For double stranded RNA (dsRNA) production, sequences of these targets were analyzed *in silico* to avoid sequenced related off-target effects according to [48]. Unique parts of the sequences were amplified with opposing T7-

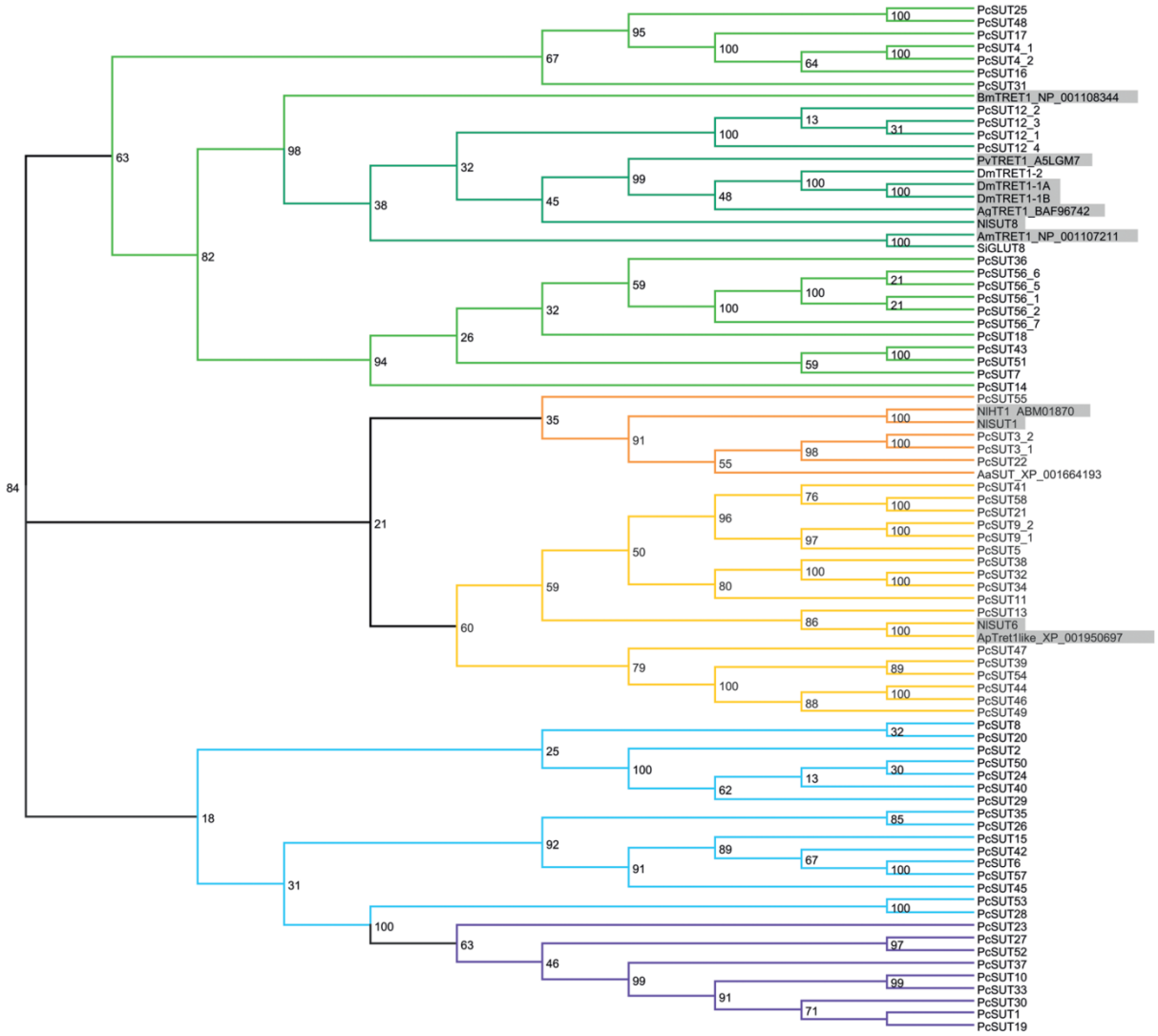


Figure 3. Phylogenetic tree of 68 chosen SLC2 transporters derived from *P. cochleariae* and chosen sugar transporters that have been functionally annotated in various insects (see Table 2). This tree was calculated by applying RAxML. The functionally characterized glucose/fructose transporters as well as trehalose transporters from insects are shaded in grey.
doi:10.1371/journal.pone.0084461.g003

promotor sequences attached to the 5′-end of each forward and reverse primer. The *gfp* sequence was amplified from pcDNA3.1/CT-GFP-TOPO (life technologies, Darmstadt, Germany) accordingly. The amplicons were subject to *in vitro* transcription assays according to instructions from the Ambion MEGAscript RNAi kit (life technologies, Darmstadt, Germany). The resulting dsRNA was eluted from silica membranes after nuclease digestion three times with 50 µl of injection buffer (3.5 mM Tris-HCl, 1 mM NaCl, 50 nM Na_2HPO_4, 20 nM KH_2PO_4, 3 mM KCl, 0.3 mM EDTA, pH 7.0). The concentration of dsRNA was calculated with $A_{260} = 1 = 45$ mg/ml and adjusted to 2 µg/µl. The quality of dsRNA was checked by TBE-agarose-electrophoresis.

P. cochleariae second instars with 4 mm body length, 0.5–0.7 mg body weight were injected with 0.4 µg of dsRNA about five days after hatching. Injections were accomplished with ice-chilled larvae using a Nano2010 injector (WPI, Sarasota, FL, USA) driven by a three-axis micromanipulator. The larvae were injected dorso-central between the pro- and mesothorax.

According to [45,73], we calculated the relative growth rate (RGR) of six biological replicates of each group of five larvae by RGR = [(final weight − weight of neonate larva)/(weight of neonate larva × developmental time [days])]. Each replicate group was weighed each 24 or 48 h ± 3 h and data were compared statistically.

GC/MS Analysis of Low-molecular-weight Compounds in Defense Secretions

Larval secretions were collected in glass capillaries (i.d.: 0.28 mm, o.d.: 0.78 mm, length 100 mm; Hirschmann, Eberstadt, Germany). Secretions were weighed in the sealed capillaries on an ultra-microbalance (Mettler-Toledo, Gießen, Germany) three times; the weight of the capillaries was subtracted and the final

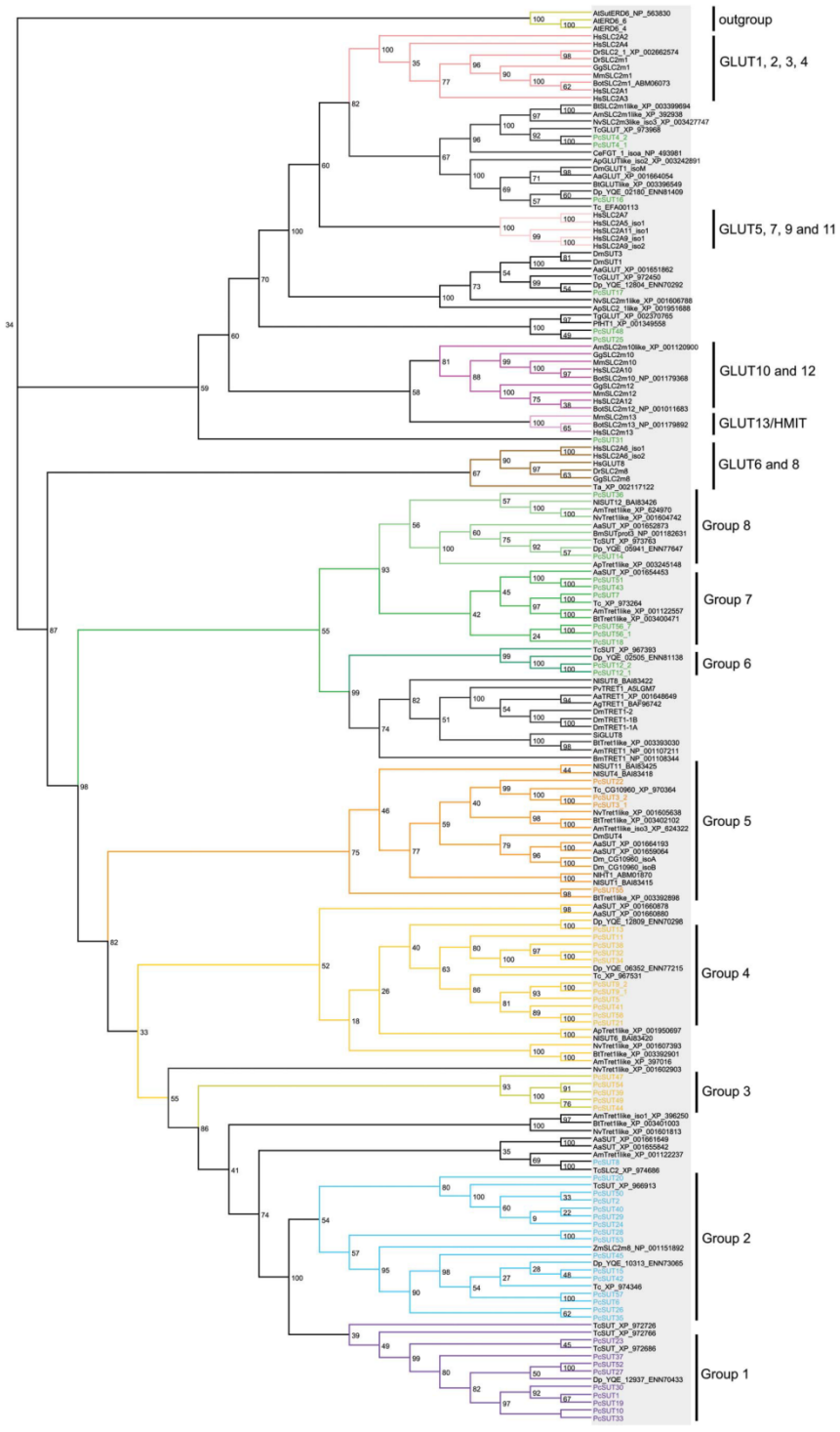


Figure 4. Phylogenetic tree of the *P. cochleariae* sequences and homologous sequences derived from the tree of life calculated using RAxML. Highlighted sequences regard to *P. cochleariae* and most similar sequences. Especially the green branch has to be subdivided into various subbranches, presenting all homologous sequences belonging to Metazoa. The tree significantly shows that the sugars (glucose) and trehalose transporters build up a huge tree in insects. Figure S5 shows the phylogeny of the selected organisms from the tree of life.
doi:10.1371/journal.pone.0084461.g004

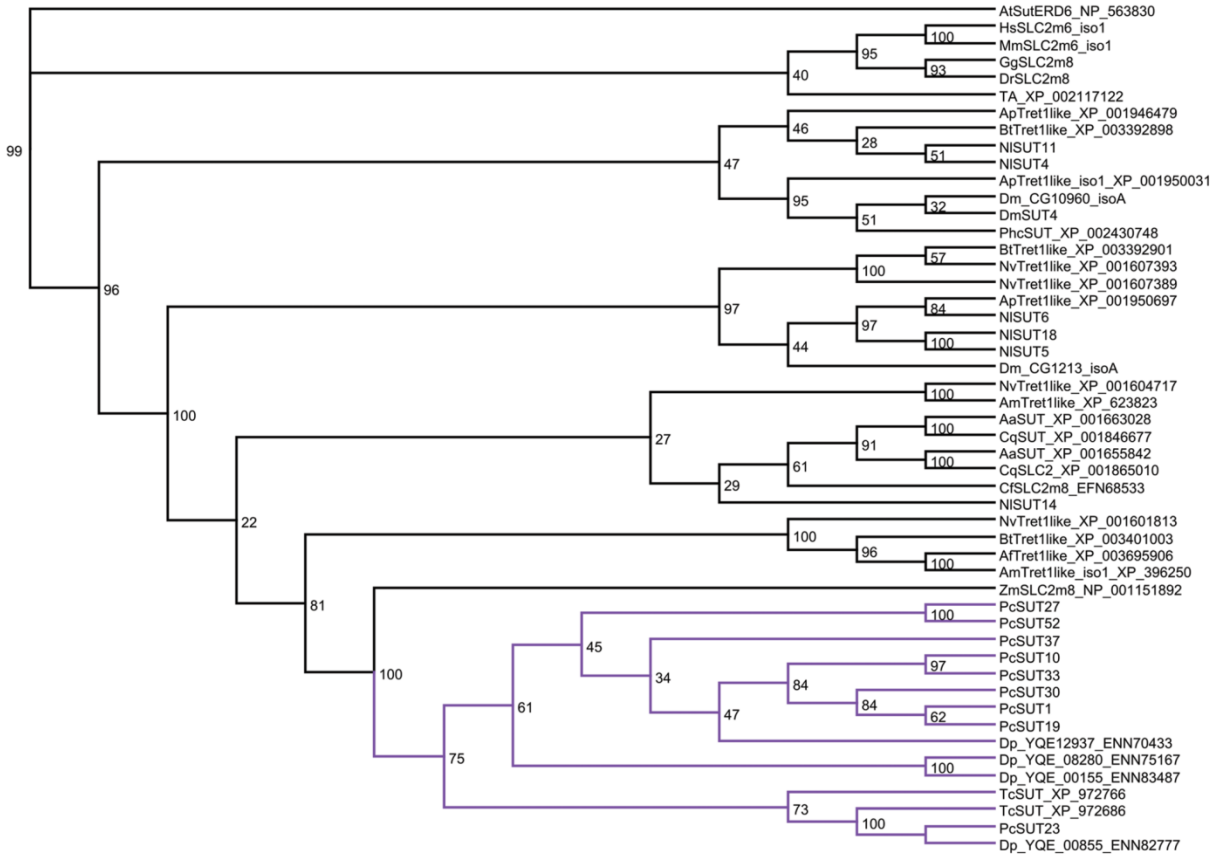


Figure 5. The phylogenetic tree of the nine sequences derived from *P. cochleariae* belonging to the purple branch (see Figure 3 and 4, Figure S4) and homologous sequences derived from the whole tree of life, especially from *Dendroctonus ponderosae* (Dc) as well as from *Tribolium castaneum* (Tc), was calculated using RaxML. Indicated in purple, it can be seen that the beetles' sequences build up a separate branch.
doi:10.1371/journal.pone.0084461.g005

weight was averaged. Sealed capillaries containing samples were stored at −80°C until needed.

According to [48] secretions of *P. cochleariae* were diluted in 1:200 (w/v) ethylacetate, supplemented with 100 µg/ml methylbenzoate as internal standard. Of each diluted secretion, 1 µl was subjected to GC/EIMS analysis (ThermoQuest Finnigan Trace GC/MS 2000, Frankenhorst, Germany) equipped with Phenomenex (Aschaffenburg, Germany) ZB–5–W/Guardian–column, 25 m. Substances were separated using helium as a carrier (1.5 ml/min). The column temperatures were set as followed: 50°C (2 min), 10°C min^{-1} to 80°C, 5°C mi^{-1} to 200°C, 30°C min^{-1} to 300°C (1 min). Inlet temperature was 220°C and transfer line was 280°C. Chrysomelidial was identified according to [36]. Peak areas of GC-chromatograms were obtained using the ICIS-algorithm (Xcalibur bundle vers. 2.0.7, Thermo Scientific).

Analysis of Differentially Expressed Genes in Glandular Tissue of RNAi Silenced *P. cochleariae* Larvae

The short reads (sequenced in 50 bp single-mode) from the glandular tissue (4 samples) of the RNAi-silenced *P. cochleariae* larvae have been mapped onto the studied sugar transporters of *P. cochleariae*'s transcriptome using bowtie [66]. The mappings results for the sugar transporters transcripts were subjected to DESeq statistical analysis by reading them into R statistics software, and transcript counts were normalized to the effective library size. Afterwards, the negative binomial testing was carried out to identify differentially expressed transcripts. All those sugar transporters were stringently determined as differentially expressed when having an adjusted p-value smaller than 0.1. Additionally, the normalized counts were stabilized according their variance as outlined in the DESeq package tutorial and heatmaps were generated [70].

Results

Identification of Sequences Encoding Putative Members of the SLC2 Family and Trehalose Transporters in the Transcript Catalogue of *Phaedon Cochleariae*

In order to predict sugar transporters in the larvae of *P. cochleariae* with special emphasis on the defensive glands, we performed a comprehensive analysis of transcriptomic data. For this purpose, cDNA libraries prepared from different tissues of juvenile *P. cochleariae* have been sequenced by using the Illumina technique. In addition to the sequencing of cDNA derived from a tissue pool in 150 bp paired-end mode, three biological replicates of the RNA of *P. cochleariae*'s defensive glands were extracted, processed and each sequenced in 50 bp single-read modes. The raw sequence data (in the following called reads) are listed in

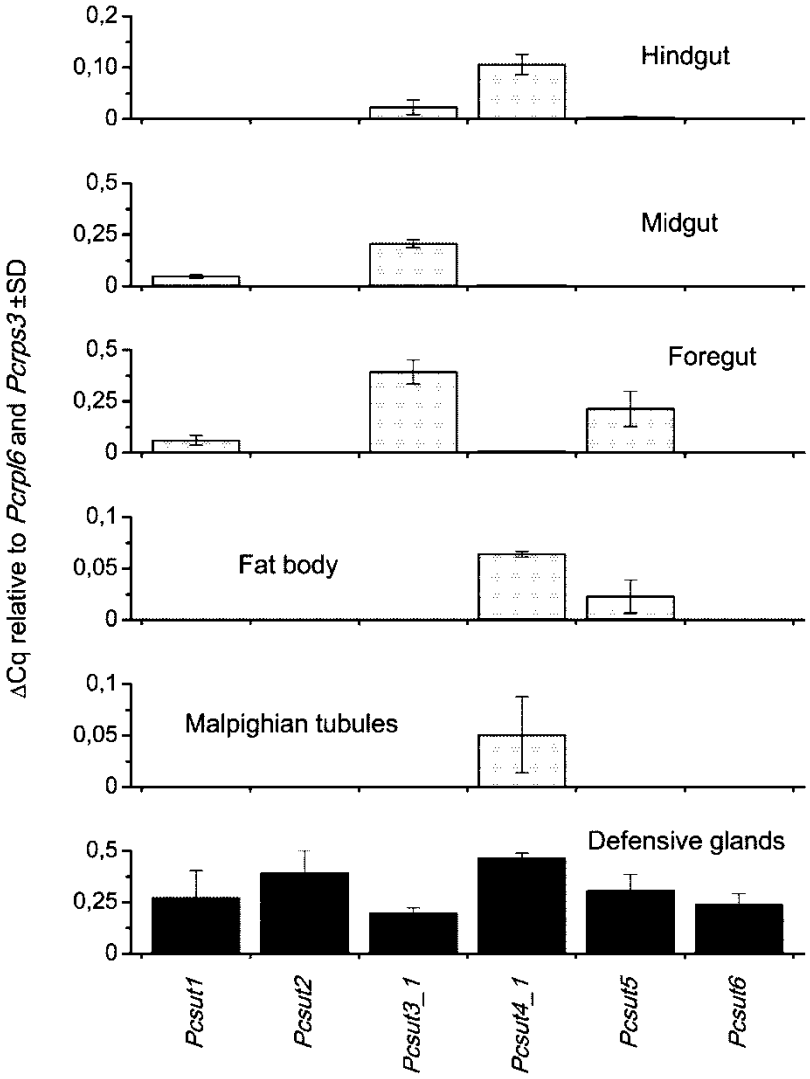


Figure 6. Distribution of mRNA levels of putative SLC2 transporters in various tissues of juvenile *P. cochleariae* by using quantitative real-time PCR.
doi:10.1371/journal.pone.0084461.g006

Table 1. The sequencing of the tissue pool resulted in 46,030,279 read pairs. The *de novo* assembly of those reads by using the Trinity software [50] resulted in 107,323 transcripts with an average sequence length of 796 bp. Reassembly by applying the TGI Clustering tool [52,74] reduced the number of transcripts to 68,667 with an average length of 929 bp (Table 3).

These 68,667 transcripts were then translated into possible protein sequences. The sequences encoding putative sugar transporters were identified by searching the Protein family database (Pfam) [55]. All sequences possessing a Sugar_tr domain (Sugar (and other) transporter family, PF00083) were selected. As a result, a total of 207 sugar transporters could be identified. Those hits were searched via BLASTp with an E-value threshold of 1e-3 against the Swiss-Prot protein database [58]. 68 predicted proteins were annotated as either SLC2 or trehalose transporters (Table S2, Figure S1). These sequences were given temporary designations as numbered series in the form of *Pc*SUTxx (Table S3). According to previous studies, transcripts with a minimum coding sequence length of 900 bp have been chosen for further phylogenetic analyses [20]. The prediction of putative transmembrane (TM) α-helices for those sequences by applying TMHMM [59] as well as Memsat2 [60] revealed that most of our sequences possess 12 TM regions (Table S3, columns 7 and 8). In total, Memsat2 was able to predict those 12 TM regions for 40 of all chosen 68 sequences, whereas TM1MM predicted these for just 33 of all 68 sequences.

Phylogenetic Analyses of Putative Sugar Transporter Sequences

The phylogenetic relationships of our selected 68 transporter sequences were analyzed by applying Probalign (as well as MAFFT [75] for the more complex multiple sequence alignments) for calculating the multiple alignments followed by calculating the phylogenetic trees. Two different methods namely MrBayes and

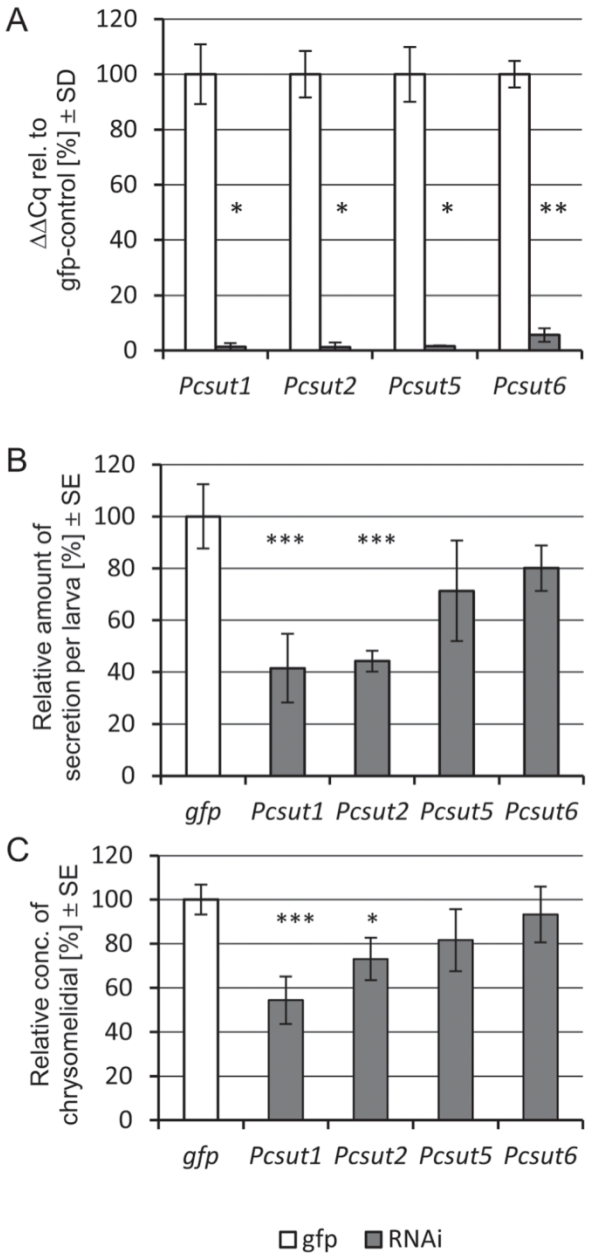


Figure 7. RNAi effects on transcript levels, amounts of defense secretions and chrysomelidial concentrations 10 days post RNAi induction in juvenile *P. cochleariae*. A, Relative expression of chosen transporters in glandular tissue, normalized internally to *Pcrpl6* and *Pcrps3* and externally to *gfp*-control, n = 5. B, Amounts of secretions produced by individual larvae were weighted and normalized to the control treatments. n = 5. C, Secretions samples of RNAi induced larvae were analyzed using GC/MS; Amounts of chrysomelidial were normalized to internal standard (methylbenzoate), values were calculated against control. n = 5. Asterisks indicate level of significance (T-test, 2-tailed; p-value< = 0.05 = *, < = 0.01 = **, < = 0.001 = ***).
doi:10.1371/journal.pone.0084461.g007

RAxML have been applied. In our case, we could show that both programs resulted in a division of the predicted transporters into at least 4 groups (Figure 1). Especially the purple branch is well supported by a bootstrap value of 100% (Figure S2).

In general, all 4 groups possess the conserved facilitated sugar transporter motifs important for transport activity and ligand binding, namely DRxGRR/K in the second loop, PESPR/K in the sixth loop, E/DRxGRR/K in loop 8, and PETK/RGK/R in the carboxy terminus (Figure 2, red colored amino acids) [2,15,18,22,76]. Additionally, conserved tryptophanes in TM domain (TMD) 4, 10 and 11, and loop 10 [77,78], and conserved tyrosines (turquois colored amino acids in TMD 4), such as the PMY [76], can be found. Furthermore, conserved glycines (yellow colored amino acids) in TMD 1, 2, 4, 5, 7, 8, and 10 as well as in loop 2 and 7 are present, characteristic for the mammalian glucose transporter family [76].

Besides the general conserved amino acid residues, we observed branch specific differences particularly for the purple branch (Figure 2, purple colored amino acids). The conserved motif QQLSG [2] which is present in all branches but substituted with QHXXG in the purple branch is important for the putative substrate binding. In addition, this purple branch exhibits several conserved motifs not present in other branches and not reported from any other transporter of the SLC2 family until now, such as a GWTAP motif in loop 1 (instead of GWTSP), a PFLPFY motif in loop 5, a VILMNLH motif in TMD 7, and a SWIP motif followed by a conserved methionine and tyrosine in TMD 10 (Figure S3, multiple sequence alignment).

By comparing the sequences of chosen sugar transporters to sugar transporters of other insect species that have been functionally characterized [15,18,19,22], we could show that these sequences fall exclusively into the green and yellow branch and not into the purple or blue branch (Figure 3, Figure S4). The functionally proven trehalose transporters such as TRET1 from *Polypedilum vanderplanki*, *Anopheles gambiae*, *Apis mellifera*, *Drosophila melanogaster* and *Bombyx mori* build up a subbranch within the green branch together with *Pc*SUT12_1–4 (with bootstrap percentage of 98%). The GLUT8 of *Solenopsis invicta* also belongs to this group. Therefore, we suggest (in agreement with Kanamori *et al.* [18]), that this fire ant GLUT8 is probably a TRET1 ortholog. The functionally characterized glucose and glucose/fructose transporters from *Nilaparvata lugens* and *Acyrthosiphon pisum* cluster in the yellow branch. However, functional analysis of *P. cochleariae*'s transcripts is required to confirm substrate spectra of the transporters clustering into different branches.

To study phylogeny of sugar transporters in a larger context, we selected the *P. cochleariae* sequences and its homologs from selected Metazoa from various branches of the tree of life (Figure S5), including the human SLC2 isoforms (GLUT1–12, H^+-*myo*-inositol transporter (HMIT)), for cladistic analyses by MrBayes and RAxML. Generally, Figure 4 displays that the sugar (glucose) and trehalose transporters build up a huge branch in insects, and the mammalian sugar porters form a separate branch. In accordance with the published phylogenetic analyses by Wilson-O'Brien *et al.* [79], the mammalian GLUTs isoforms segregate into five distinct classes, also showing that the mammalian GLUTs are separate from their insect orthologs. Some insect sequences also cluster into the mammalian clades, e.g. the class I clade (GLUT1, 2, 3, 4) also contains insect sequences branching at the base with strong support (with bootstrap percentage of 100%). *Pc*SUT31 clusters together with mammalian HMIT. The class II (GLUT5, 7, 9, 11) and the GLUT10/12 clade contain only vertebrate sequences. According to Wilson-O'Brien *et al.* [79], class II genes were most likely to arise after the divergence of this

Table 4. Differential expression analysis using DESeq package.

Differentially expressed transcripts with padj<0.1 after dsRNA injection of *Pcsut1*:						
Seq_Id	**baseMeanA**	**baseMeanB**	**foldChange**	**log_2FoldChange**	**pval**	**padj**
PcSUT1	4804.953616	240.8731589	0.050130174	−4.31817696	1.26E-16	7.31E-15
Differentially expressed transcripts with padj<0.1 after dsRNA injection of *Pcsut2*:						
Seq_Id	**baseMeanA**	**baseMeanB**	**foldChange**	**log_2FoldChange**	**pval**	**padj**
PcSUT2	8428.420254	2031.419654	0.241020214	−2.052773946	4.15E-08	2.49E-06
PcSUT25	22.79115587	394.3587471	17.30314817	4.112962644	0.003061159	0.091834771
Differentially expressed transcripts with padj<0.1 after dsRNA injection of *Pcsut6*:						
Seq_Id	**baseMeanA**	**baseMeanB**	**foldChange**	**log_2FoldChange**	**pval**	**Padj**
PcSUT6	3380.778995	95.65776606	0.028294593	−5.143329799	4.20E-14	2.56E-12
PcSUT25	22.79115587	995.628527	43.68486322	5.449061568	0.000272434	0.008309243

baseMeanA: mean of normalized counts value of ds*gfp*-injected samples. baseMeanB: mean of normalized counts values of dsRNA-*gfp*-injected, dsRNA-*Pcsut1*-injected, dsRNA-*Pcsut2*-injected, dsRNA-*Pcsut6*-injected samples. Fold-change: baseMeanA compared to baseMeanB. Log_2fold-change: logarithm (to base 2) of fold-change values. Pval: p-value for the statistical significance of this change. Padj: p-value adjusted for multiple testing with Benjamini-Hochberg procedure which controls false discovery rate.

doi:10.1371/journal.pone.0084461.t004

phylum, whereas the GLUT10/12 clade might has been lost in invertebrates secondarily. The majority of the tested insect sequences form a huge sister group of the GLUT6/8 class with strong support (100% bootstrap), suggesting an expansion of this class in insects. On inspection of the insect sequences, especially the green branch is remarkably large and can be subdivided into various subbranches. In addition to the five mammalian SCL2 classes, we suggest eight more classes in insects (each separated from the others with bootstrap percentages of at least 80%).

The bootstrap values of the tree of life support the notion that the purple branch may be restricted to the insect order Coleoptera. For deeper analysis of this branch, we have calculated trees with the sequences of this purple branch and its homologs/orthologs in insects. The MrBayes tree as well as the RAxML tree (Figure 5 and Figure S6) shows that only sequences derived from beetles such as *Dendroctonus ponderosae*, *Tribolium castaneum*, and *P. cochleariae* belong to this purple branch. Whereas, looking at Figure 3 and 4, all other sequences derived from *P. cochleariae* show homologies to Diptera, Apocrita and other insect orders. According to our cladistic analyses, the purple branch of the phylogenetic trees seems to be the most interesting. All changed and additionally conserved motifs together with the motif QH lead to the hypothesis that the sequences belonging to the purple branch may provide a new class of sugar transporters in insects, especially in the order Coleoptera, not before described. However, their biological function remains to be elucidated.

Expression Profiles of SLC2 and Trehalose Transporter Transcripts in the Defensive Glands of Immature *P. cochleariae*

The Illumina short reads derived from the glandular tissue sequenced using HiSeq2000 in 50 bp single-read mode (two or three samples multiplexed in one lane) resulted in 34,918,295 reads with length of 50 bp as well as 36,598,828 and 29,866,004 short reads, respectively. Those short reads have been mapped onto our transcriptome. 82.62% to 84.90% of all short reads could be mapped, whereas 14.28% to 16.31% of the reads failed to align (Table S4).

The mapping results were taken as input for R statistics. DESeq, belonging to the Bioconductor package, was applied to analyze the mapping results. The counts of transcripts are listed in Table S3 (columns 3 to 5). All observed putative sugar transporters of SCL2 and trehalose transporters could be identified in the glandular tissue, although most of them at a very low level. Interestingly, the six most abundant transcripts (inclusive isoforms) with more than 3,000 normalized counts are spread among all four major phylogenetic subtrees as highlighted in Figure S7. Therefore, no specific branch is particularly overrepresented in the defensive glands. As previously demonstrated in literature [80], evaluation of the RNA-seq data (standardized values shown in Table S5) with quantitative real-time PCR data shows also in our experiments the comparability of the two methods (Figure S8). Subsequently, the six transcripts have been analyzed regarding the distribution of their transcript levels in different larval tissues. Two sugar transporters namely *Pc*SUT2 and *Pc*SUT6, clustering into the blue branch, are exclusively expressed in defensive glands (Figure 6). The other four tested candidates were found to be also expressed in at least one more tissue.

Silencing of Putative Sugar Transporters in Immature *P. cochleariae* by RNAi

RNA interference (RNAi) was carried out, besides for *Pcsut2* and *Pcsut6*, also for *Pcsut1* (highly expressed in defensive glands and gut) and *Pcsut5* (highly expressed in defensive glands, fat body and gut) to get deeper insights into their biological function for the development of *P. cochleariae* larvae and the production of defensive secretions *in vivo*. Early second instars were injected with either dsRNA identical to a unique part of one of these transporters sequences or with dsRNA of *gfp*. As confirmed by quantitative real-time PCR, transcription of all targets was successfully silenced in the defensive glands 10 days after injection (Figure 7A). By monitoring the development of treated larvae, significant weight reduction was observed neither for the larvae nor the pupae by

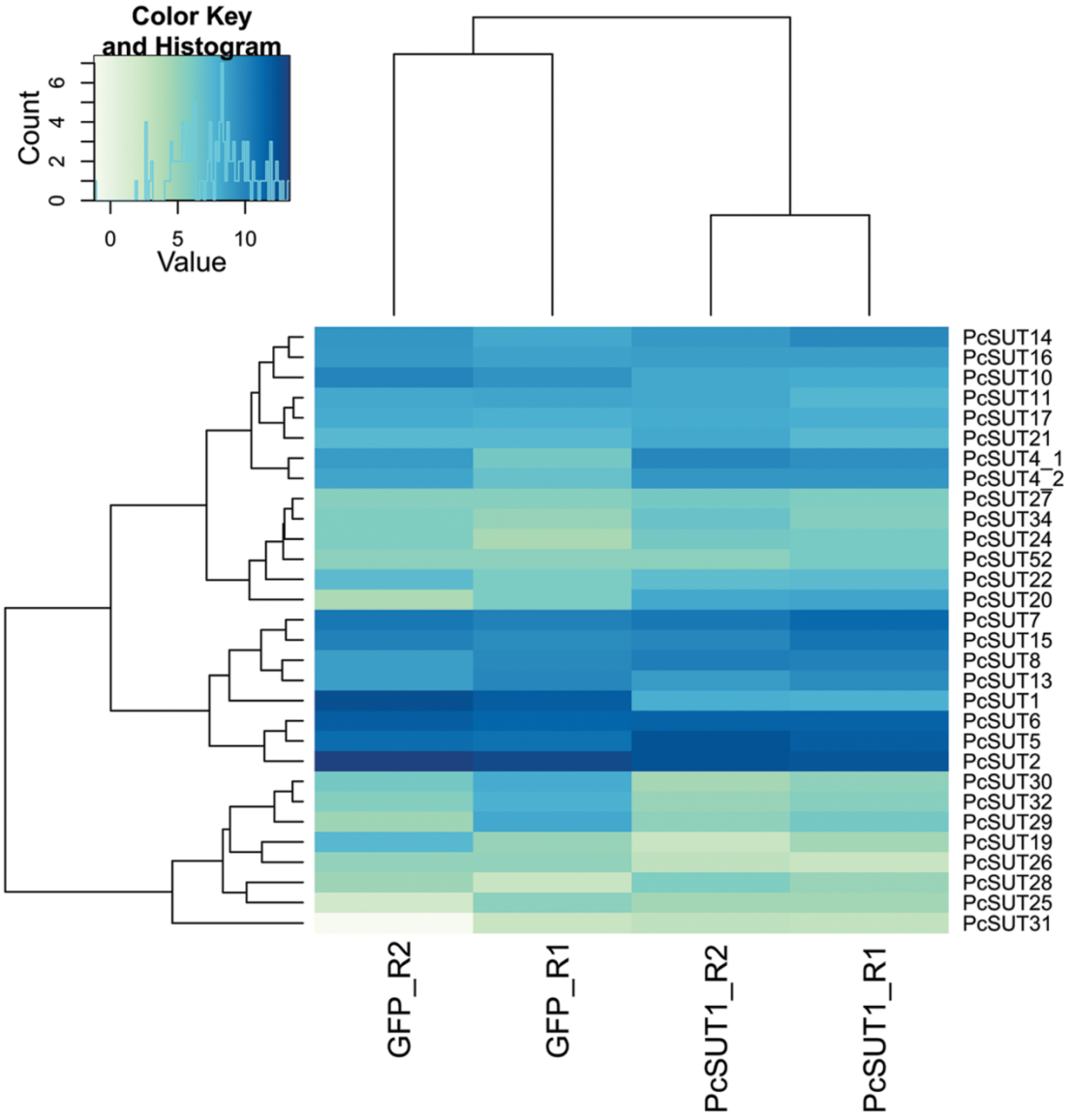


Figure 8. Heatmap of the variance stabilization transformed data (vsd) of ds*Pcsut1*-injected *vs.* ds*gfp*-injected samples. Samples derived from glandular tissue. For this, the transcript counts of the sugar transporters of each sample after dsRNA-injection have been normalized to the effective library size and the variance over all samples has been stabilized by applying the DESeq package. For each heatmap, the 30 most abundant sugar transporter transcripts are shown. Ds*gfp*-injected samples are the same in each heatmap.
doi:10.1371/journal.pone.0084461.g008

transporter silencing (Figure S9). To screen for the function of the transporters for the synthesis of secretions, GC/MS analysis was carried out for quantification of the amount of chrysomelidial relatively to the *gfp*-controls. Here, we could observe decreases in the amount of chrysomelidial by silencing *Pcsut1* ($p = 0.008$) and *Pcsut2* ($p = 0.001$) and also in the amount of secretions produced by targeting *Pcsut1* ($p = 0.007$) and *Pcsut2* ($p = 0.03$) (Figure 7B and 7C). Knocking down of neither *Pcsut5* nor *Pcsut6* resulted in alterations of the phenotype. We assume that *Pc*SUT1 and *Pc*SUT2 seem to be important for the production of defensive secretions. Their substrate selectivity, however, needs to be further studied *in vitro*.

In contrast to the silencing effect of the ABC transporter in the defensive glands which caused a total loss of defensive secretions [46], RNAi targeting SLC2 transporters could not shut down the production of defensive exudates completely. This may be due to the turn-over rate of the integral-membrane proteins which may diminish the silencing effect or other transporters may be expressed to take over the function to achieve homeostasis in the tissue. To test the latter hypothesis, we have gathered RNA-seq data from larvae silenced in *Pcsut2* or *Pcsut6* (as examples for SLC2 members exclusively expressed in glands) or *Pcsut1* (as example highly expressed in defensive glands and gut) and analyzed differential expression of SLC2 transporters in comparison to *gfp*-treated larvae.

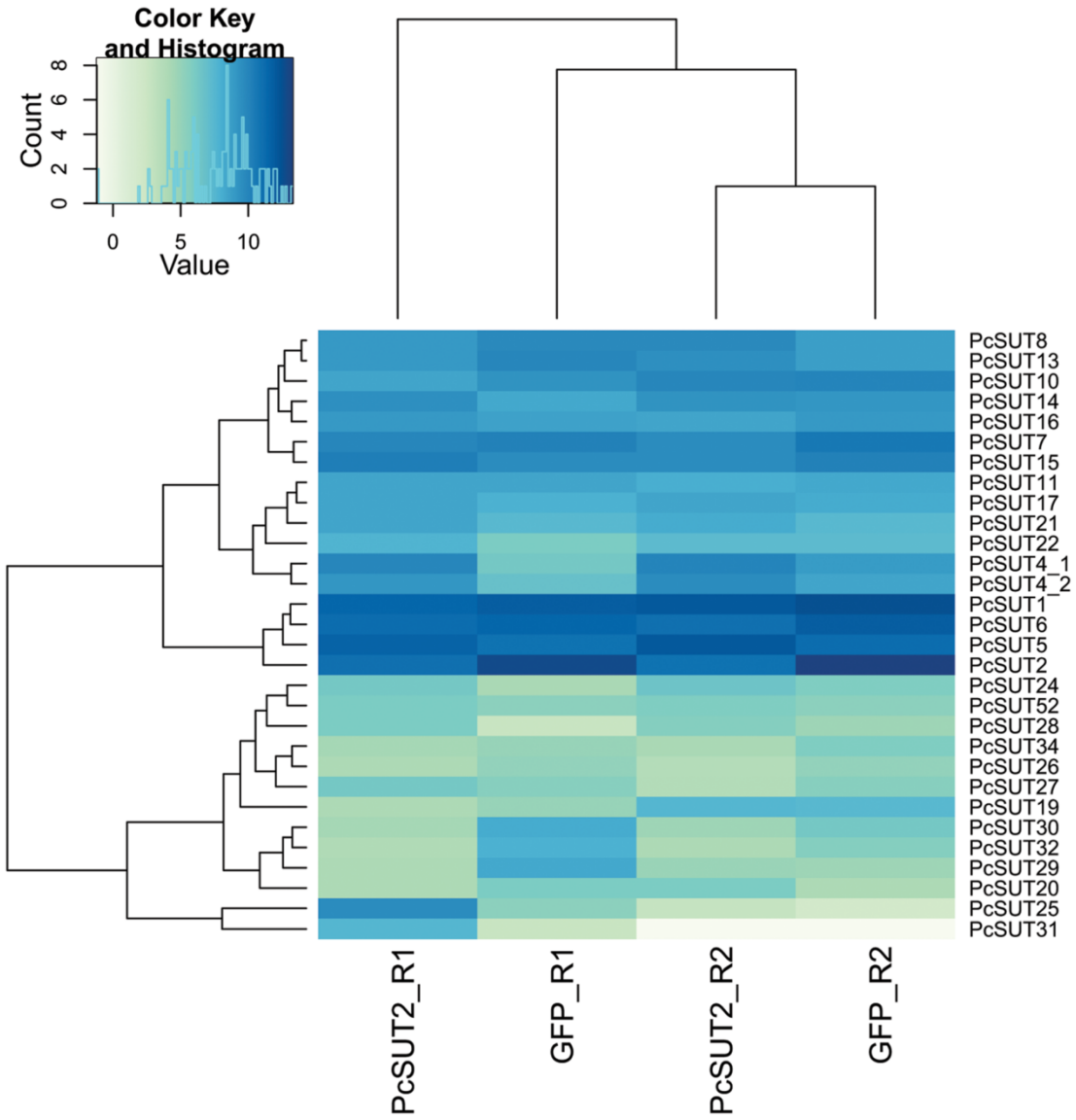


Figure 9. Heatmap of the variance stabilization transformed data (vsd) of ds*Pcsut2*-injected *versus* ds*gfp*-injected samples. Samples derived from glandular tissue. For further explanation see Figure 8.
doi:10.1371/journal.pone.0084461.g009

Analysis of Differentially Expressed Genes in the Glandular Tissue of RNAi-silenced *P. cochleariae* Larvae

Ten days after *Pcsut1*, *Pcsut2* and *Pcsut6*-silencing and ds*gfp*-injection, glandular tissues were dissected and two biological replicates for each treatment were sequenced. The normalized counts of all sugar transporters of all samples are listed in Table S6. The $\log_2$ fold-changes of the silenced transporters (ds*gfp*-injected samples as control) and adjusted p-values were determined using the DESeq package (see Table 4). In Figure S10 all sugar transporters exhibiting significantly different transcription levels are colored red (MA plot showing $\log_2$ fold-changes *vs.* mean values). In all samples (prepared in RNA-seq and quantitative real-time PCR experiments), we observed varying transcript levels of SCL2 transporters owing to the individual biological variance and diversity despite similar developmental stage or living conditions.

Pcsut1 knocking-down led to a significant decrease of its own transcript level (Table S6: adjusted p-value (padj) = 7.31E-15). Figure 8 exhibits a heatmap of the 30 most abundant sugar transporters. Besides *Pcsut1*, three more sugar transporters were co-silenced, namely *Pcsut10*, *Pcsut30* and *Pcsut32*. Those were not determined as differentially expressed, but have a $\log_2$ fold-change smaller −2 (Figure S10A, lower right quadrant). In contrast, *Pcsut4*, *Pcsut5*, *Pcsut20*, and *Pcsut22* were up-regulated ($\log_2$ fold-change of 1) suggesting counter-regulation to the silencing effect to ensure sugar homeostasis in the defensive glands. This up-regulation, however, did not fully compensate the silencing effect

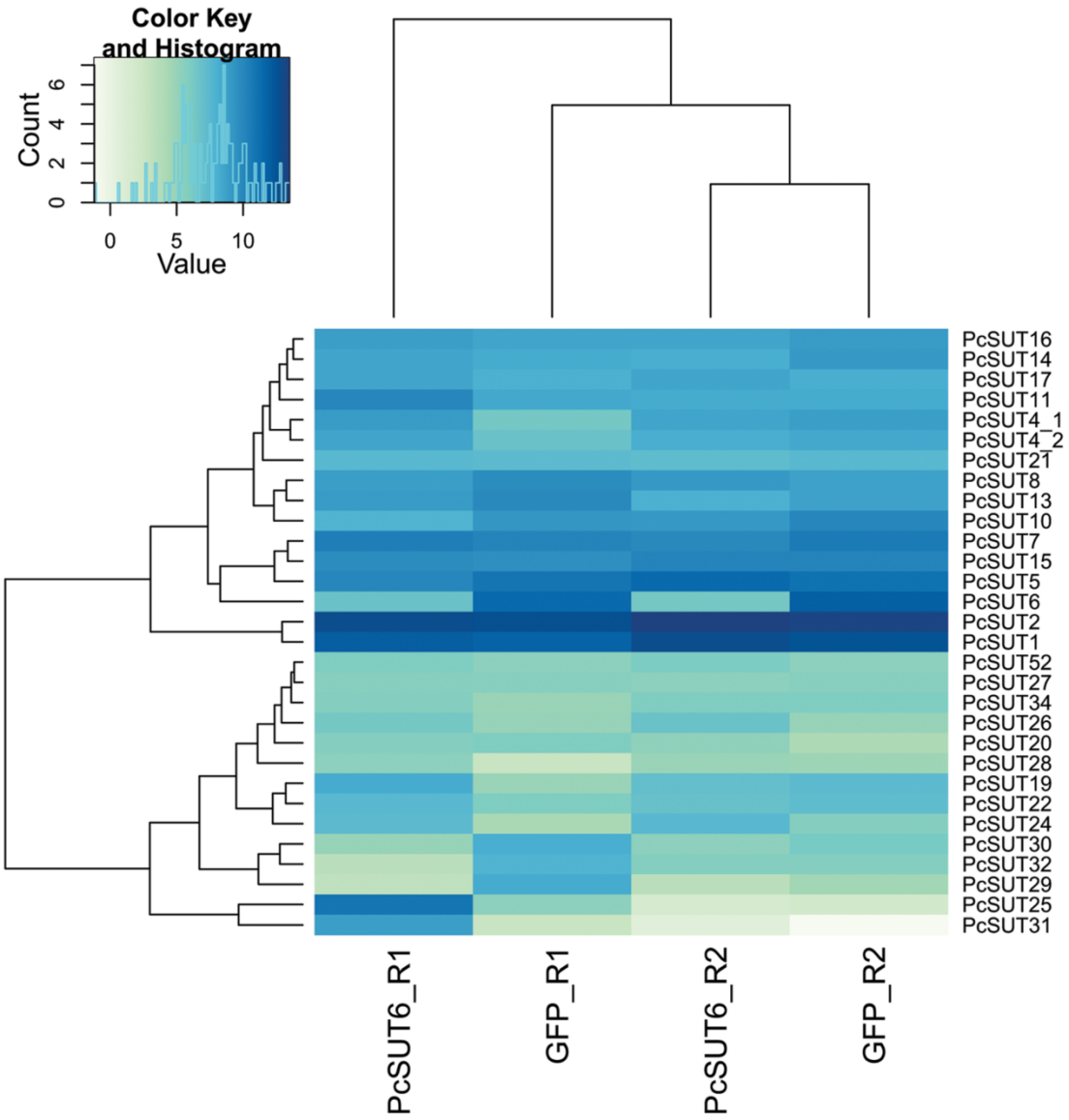


Figure 10. Heatmap of the variance stabilization transformed data (vsd) of ds*Pcsut6*-injected *vs.* ds*gfp*-injected samples. Samples derived from glandular tissue. For further explanation see Figure 8.
doi:10.1371/journal.pone.0084461.g010

of *Pcsut1* indicated by the decrease in the production of defensive secretions (as shown before).

The samples prepared after dsRNA-injection of *Pcsut2* showed, on the one hand, a significant down-regulation of *Pcsut2* itself (padj = 2.49E-06), but also a decrease of *Pcsut6*, *Pcsut26*, *Pcsut30*, *Pcsut32* and *Pcsut34* ($\log_2$ fold-change of −1, Figure S10B lower right section). On the other hand, *Pcsut4*, *Pcsut5*, *Pcsut14*, *Pcsut17*, *Pcsut21*, *Pcsut24* and *Pcsut28* were up-regulated (Figure 9). Additionally, the transcript level of *Pcsut25* was determined as significantly higher in the silenced samples than in the ds*gfp*-injected samples (padj = 0.092) (Table 4). But also here, the counter-regulated transporters did not compensate the silencing effect completely.

In the samples with *Pcsut6* silenced *via* dsRNA-injection, *Pcsut6* was drastically reduced (padj = 2.56E-12). Furthermore, *Pcsut13*, *Pcsut29*, *Pcsut30* and *Pcsut32* were also down-regulated ($\log_2$ fold-change of −1, Figure S10C). To establish the sugar homeostasis in this sample, *Pcsut1*, *Pcsut17*, *Pcsut24* and *Pcsut26* were transcribed at a higher level ($\log_2$ fold-change of 1) compared to the ds*gfp*-injected samples. Especially *Pcsut25* was significantly higher expressed (padj = 0.008) (Figure 10, Table 4).

Discussion

Sugars play an important role in all species' metabolism. Transporters of the SLC2 family are key elements involved in the adaptive response to sugar demand that have important physiological implications to cell survival and growth. They are expressed

in a tissue-specific manner with different affinity, specificity and capacity for substrate transport [81]. Recent phylogenetic analyses of genomic data available from sequenced insects suggest a remarkable expansion of the SLC2 family in insects compared to mammals [18,20,23,26,79]. Beetles, however, have not yet been addressed. Here we present the first comprehensive phylogenetic analysis of members of the SLC2 sugar transporter family identified in a leaf beetle species.

Phylogenetic Analyses

We created a transcript catalogue of juveniles of *P. cochleariae* in which we annotated 68 sequences as SLC2 transporters. The phylogenetic analyses of putative sugar transporters were performed with MrBayes and RAxML. MrBayes, on the one hand, is a program for Bayesian inference and model choice across a large space of phylogenetic and evolutionary models. On the other hand, RAxML is a maximum likelihood phylogeny estimation. In our case, we were able to show that both programs result in a division of the predicted transporters into 4 main groups. However, there are some differences in the results of the calculations with the two methods. For example, *Pc*SUT8 belongs to the green branch of the phylogenetic tree showing *P. cochleariae*'s sequences as well as the functionally characterized sequences of the other insects calculated by MrBayes; RAxML sorted it into the blue branch (Figure 3, Figure S4). Comparing these trees with the trees showing only sequences of *P. cochleariae* as well as with the trees which also include the sugar transporter sequences of mammals and other organisms, it can be seen that *Pc*SUT8 always belongs to the blue branch. To conclude, we propose that RAxML results in more stable trees when adding or deleting homologous sequences. This fact might be strengthened by Douady *et al.* [82] who concluded that the more conservative use of bootstrap percentages (as used by RAxML) might be less prone to supporting strongly a false phylogenetic hypothesis.

The mammalian GLUT proteins are well-studied, and phylogenetic studies have been carried out. In addition to the five distinct classes of mammalian GLUTs [79], we propose eight more groups of transporters including the trehalose transporters in the Insecta. The tree including mammalian GLUTs as well as homologous sequences of other organisms including insects shows that many transporters belonging to the SLC2 family derived from *P. cochleariae* and other insects constitute a huge subtree separate from the well-studied mammalian GLUTs. We suggest that the orange branch within this insect clade constitutes the fructose/glucose transporters. This suggestion is supported by, on the one hand, *Ap*ST3 from the pea aphid *A. pisum* which acts in the gut as a low-affinity uniporter for fructose and glucose [23], and on the other hand, *Nl*HT1 (*Nl*ST1) and *Nl*ST6 from the brown planthopper *N. lugens* which function as glucose and glucose/fructose transporters, respectively, in the gut [20,22]. The green branch within the huge insect clade most likely contains trehalose transporters including the transporter TRET1 from the sleeping chironomid *P. vanderplanki* and its orthologs from *A. gambiae*, *B. mori*, *A. mellifera* and *D. melanogaster* [18,19] and a proton-dependent transporter participating in trehalose reabsorption in Malpighian tubules of *N. lugens* [21]. Kinetic parameters show different affinities for trehalose among the TRET1 orthologs which mirror the trehalose:glucose ratio in the hemolymph of each species [18]. No functional assumptions can be made regarding all other subbranches, especially concerning the purple one which seems the most interesting to the study of beetles.

The large number of SLC2 transporters may result from gene duplications as has been suggested for *A. pisum*, whose genome contains a conspicuous number of genes encoding predicted sugar transporters [23]. While such an idea still needs experimental proof, it may contribute to the reflection on how insects are able to adapt to extreme dietary conditions and to the testing of substrates not yet in the focus for SLC2 members such as plant derived glucosides, for example, present in large amounts in the diet of phytophagous insects.

Interestingly, the mammalian GLUT6/8 form a sister group of the expanded insect clade. In accordance with the literature [76,79] we see that GLUT6 and 8 are more closely related to sugar transporters present in invertebrate species than to other mammalian GLUTs. Mueckler *et al.* [27] stated that the primary physiological substrate for mammalian GLUT6 and 8 have not been definitely identified. Therefore, knowledge of substrate selectivity of the SLC2 members in the insect clade may also contribute to a deeper understanding of the function of GLUT 6 and 8 in mammals.

Membrane Protein Topology

Membrane proteins, such as sugar transporters, seem to have a restricted range of folds than their water-soluble counterparts, making them more amenable to structural predictions [83]. α-Helical membrane proteins contain one or more transmembrane helices which consist predominantly of hydrophobic amino acids. In our study of sugar transporters, 12 TM helices are stated [26]. To predict the TM domains of *P. cochleariae*'s sugar transporters, TMHMM as well as Memsat2 were used. For 40 of 68 sequences, Memsat2 was able to predict all 12 TM domains. TMHMM predicted those 12 TM domains for just 33 sequences. But, for three sequences Memsat2 failed to predict at least 10 TM helices and proposed 1 (twice) and 2 (once) TM helices. For those three sequences, TMHMM predicted 10 (once) and 12 (twice) TM helices. This leads us to the conclusion that neither of these prediction methods is perfect, but applying both gives us the required information. Furthermore, Cuthbertson *et al.* [83] suggested that optimal prediction is obtained by the method that best reflects the biological and physical principles governing membrane protein architecture.

RNAi and Subsequent RNA-seq

After silencing *Pcsut1* and *Pcsut2*, phenotypic analyses revealed a reduction of defensive exudates in the larvae. Silencing of *Pcsut5* and *Pcsut6* did not result in a changed phenotype. By combining the RNAi experiments (silencing *Pcsut1*, *Pcsut2* and *Pcsut6*) with subsequent mRNA isolation and RNA sequencing, we could show the down-regulation as well as up-regulation of sugar transporters in the defensive glands. Silenced and induced sequences belong to separate branches in the phylogenetic trees which also suggest that insect transporters belonging to different clades can have comparable substrate preference. Anyhow, a direct correlation of counter-acting transporters was difficult to identify. All predicted off-targets according to Bodemann *et al.* [48] were excluded in the transporter sequences used for RNAi. Nevertheless, co-silencing effects could not be avoided. These effects could not be predicted most likely because of metabolic co-silencing which was already observed for two hexokinases in *T. castaneum* [84].

Homeostasis may not only be achieved by the induced expression of transporters but, for example, also by the induced trafficking of transport proteins within a cell. This trafficking is known, for example, from the mammalian GLUT4. The protein is hormonally induced to translocate from intra-cellular membranes to the plasma membrane for the absorption of excessive glucose from the blood [85]. In insects, however, an analogous phenomenon is not known and was not addressed in our analyses. In general, sugar homeostasis is not very well understood in insects,

and the role of SLC2 members in this process has not been addressed to date. By transporter silencing in the defensive glands of the mustard leaf beetle larvae, we conclude that there is a complex network of SLC2 transporters in which several transporters compensate the function of the silenced ones. We demonstrate clearly the potential of SLC2 transporters to respond adaptively to nutrient demand. This response may have paramount ecological implications for the survival of phytophagous beetles in plant-insect interactions.

Supporting Information

Figure S1 Bar plot showing the molecular functions assigned to the *Pc*SUTs. First, BLASTx was performed to annotate the studied sugar transporters of *P. cochleariae*. Thereafter, BLAST2GO was applied to those BLAST hits. The assigned molecular functions are displayed.
(EPS)

Figure S2 Phylogenetic tree of the 68 chosen sugar transporters derived from *P. cochleariae*. The phylogenetic tree was calculated with RAxML showing the main 4 groups. The purple subbranch is separated from the other branches with a bootstrap value of 100%.
(EPS)

Figure S3 Multiple sequence alignment of the 68 chosen sugar transporters derived from *P. cochleariae*. The multiple sequence alignment was calculated using Probalign. The purple-branch-specific amino acids are framed.
(TIF)

Figure S4 Phylogenetic tree of *Pc*SUTs derived from *P. cochleariae* and functionally characterized sugar transporters of other insects listed in Table 2 calculated with MrBayes.
(EPS)

Figure S5 Phylogenetic tree showing the selected organisms from the tree of life.
(EPS)

Figure S6 The phylogenetic tree of the nine sequences derived from *P. cochleariae* belonging to the purple branch (see Figure 3 and 4, Figure S4) and homologous sequences derived from the whole tree of life, especially from *Dendroctonus ponderosae* (Dc) as well as from *Tribolium castaneum* (Tc), was calculated using MrBayes. Indicated in purple, it can be seen that the beetles' sequences build up a separate branch.
(EPS)

Figure S7 Phylogenetic *tree* of the 68 chosen sugar transporters derived from *P. cochleariae*. The six most abundant glandular sugar transporters are circled and not branch-specific, but distributed all over the tree.
(EPS)

Figure S8 Relative mRNA levels of putative SLC2 transporters in the defensive glands of juvenile *P. cochleariae* determined by carrying out RNA-seq (A) and quantitative real-time PCR (B) experiments. The corresponding fold-changes of the RNA-seq samples are listed in Table S5.
(TIF)

Figure S9 RNAi effects on the development of the larvae from *P. cochleariae*. A, The development of larval weight was documented and measured in a 24 or 48 h±3 h period. B, In neither the relative growth rate nor in the weight of freshly emerged pupae significant differences could be observed between ds*gfp*- and ds*Pcsut1*-, ds*Pcsut2*-, ds*Pcsut5*-, ds*Pcsut6*-injected larvae, n = 30.
(TIF)

Figure S10 MvA-Plot showing normalized mean values *versus* $\log_2$fold-changes. The fold-changes (log-transformed) were computed for the comparison of ds*gfp*-injected and ds*Pcsut1*-injected samples. The transcript levels are significant at 10% FDR (padj<=0.1, padj: p-value adjusted for multiple testing with the Benjamini-Hochberg procedure which controls false discovery rate (FDR)). The differentially expressed transporters, the ones colored red, are listed in Table 4. (A) MvA-Plot of the comparison of ds*Pcsut1*-injected and ds*gfp*-injected samples. *Pcsut1*'s transcript level was significantly reduced by RNAi (red dot in the lower right quadrant). (B) MvA-Plot of the comparison of ds*Pcsut2*-injected and ds*gfp*-injected samples. *Pcsut2*'s transcript level was significantly reduced by RNAi (red dot in the lower right area). Additionally, the expression of *Pcsut25* was significantly induced (red dot in the upper right part). (C) MvA-Plot of the comparison of ds*Pcsut6*-injected and ds*gfp*-injected samples. *Pcsut6*'s transcript level was significantly reduced by RNAi (red dot in the lower right area). Additionally, the expression of *Pcsut25* was significantly induced (red dot in the upper right part).
(TIF)

Table S1 Primer sets used in quantitative real-time PCR and RNAi experiments.
(XLSX)

Table S2 BLAST2GO results for all identified sugar transporters. Table shows sequence description, sequence length, minimal e-value of BLAST search, mean similarity to all 5 hits for each query sequence respectively, the number of assigned gene ontology (GO) categories as well as assigned GO categories.
(XLSX)

Table S3 Prediction of 12 transmembrane regions using TMHMM as well as Memsat2. Furthermore, the counts of each replicate as well as sequence length and coding region's length are listed.
(XLSX)

Table S4 The short reads of glandular tissues were mapped onto the transcriptome using bowtie. The short reads of the three replicates of RNA derived from the glandular tissue were mapped onto the transcriptome by applying bowtie. The percentages of aligned reads and of reads that did not align or were suppressed (due to option m) are listed. Furthermore, the numbers of reads that have been mapped are shown.
(XLSX)

Table S5 *Pcsut* counts normalized to the effective library size, relatively to *Pcrps3* and *Pcrpl6* according to Livak and Schmittgen [71]. Samples were derived from glandular tissue from juvenile *P. cochleariae*.
(XLSX)

Table S6 Counts normalized to the effective library size of all sugar transporters after dsRNA-injection. dsRNA targeting *Pcsut1*, *Pcsut2* and *Pcsut6* as well as dsRNA targeting *gfp* were injected. Samples were derived from glandular tissue from juvenile *P. cochleariae*.
(XLSX)

Acknowledgments

We sincerely thank Dr. Marita Kunert and Angelika Berg for technical assistance. We are grateful for Dr. Lydia Gramzow's advices regarding the calculation of phylogenetic trees. Furthermore, we thank Andreas Petzold, Anja Strauß, Sindy Frick and Peter Rahfeld for all of their encouragement and fruitful discussions and Melinda Palmer-Kolb for critical reading of the manuscript.

Author Contributions

Conceived and designed the experiments: AB MS RRG. Performed the experiments: MS RRG AB MG SE. Analyzed the data: MS RRG MG AB. Contributed reagents/materials/analysis tools: MG WB. Wrote the paper: MS AB RRG.

References

1. Caviedes-Vidal E, Karasov WH, Chediack JG, Fasulo V, Cruz-Neto AP, et al. (2008) Paracellular absorption: A bat breaks the mammal paradigm. Plos One 3.
2. Augustin R (2010) The protein family of glucose transport facilitators: It's not only about glucose after all. Iubmb Life 62: 315–333.
3. Wright EM, Loo DDF, Hirayama BA (2011) Biology of human sodium glucose transporters. Physiol Rev 91: 733–794.
4. Elbein AD, Pan YT, Pastuszak I, Carroll D (2003) New insights on trehalose: a multifunctional molecule. Glycobiology 13: 17R–27R.
5. Becker A, Schloder P, Steele JE, Wegener G (1996) The regulation of trehalose metabolism in insects. Experientia 52: 433–439.
6. Thompson SN (2003) Trehalose - The insect 'blood' sugar. In: Simpson SJ, editor. Advances in Insect Physiology, Vol 31. London: Academic Press Ltd-Elsevier Science Ltd. 205–285.
7. Wegener G, Macho C, Schloder P, Kamp G, Ando O (2010) Long-term effects of the trehalase inhibitor trehazolin on trehalase activity in locust flight muscle. J Exp Biol 213: 3852–3857.
8. Candy DJ, Becker A, Wegener G (1997) Coordination and integration of metabolism in insect flight. Comp Biochem Physiol B Biochem Mol Biol 117: 497–512.
9. Silva MCP, Terra WR, Ferreira C (2010) The catalytic and other residues essential for the activity of the midgut trehalase from *Spodoptera frugiperda*. Insect Biochem Mol Biol 10: 733–741.
10. Bounias M, Bahjou A, Gourdoux L, Moreau R (1993) Molecular activation of a trehalase purified from the fat body of a coleopteran insect (*Tenebrio molitor*), by an endogenous insulin-like peptide. Biochem Mol Biol Int 31: 219–266.
11. Shimada S, Yamashita O (1979) Trehalose absorption related with trehalase in developing ovaries of the silkworm, *Bombyx mori*. J Comp Physiol 131: 333–339.
12. Bifano TD, Alegria TGP, Terra WR (2010) Transporters involved in glucose and water absorption in the *Dysdercus peruvianus* (Hemiptera: Pyrrhocoridae) anterior midgut. Comp Biochem Physiol B Biochem Mol Biol 157: 1–9.
13. Caccia S, Casartelli M, Grimaldi A, Losa E, de Eguileor M, et al. (2007) Unexpected similarity of intestinal sugar absorption by SGLT1 and apical GLUT2 in an insect (*Aphidius ervi*, Hymenoptera) and mammals. Am J Physiol Regul Integr Comp Physiol 292: R2284–R2291.
14. Giordana B, Milani A, Grimaldi A, Faraneti R, Casartelli M, et al. (2003) Absorption of sugars and amino acids by the epidermis of *Aphidius ervi* larvae. J Insect Physiol 49: 1115–1124.
15. Chen ME, Holmes SP, Pietrantonio PV (2006) Glucose transporter 8 (GLUT8) from the red imported fire ant, *Solenopsis invicta* Buren (Hymenoptera : Formicidae). Arch Insect Biochem Physiol 62: 55–72.
16. Pascual I, Berjon A, Lostao MP, Barber A (2006) Transport of D-galactose by the gastrointestinal tract of the locust, *Locusta migratoria*. Comp Biochem Physiol B Biochem Mol Biol 143: 20–26.
17. Meyer H, Vitavska O, Wieczorek H (2011) Identification of an animal sucrose transporter. J Cell Sci 124: 1984–1991.
18. Kanamori Y, Saito A, Hagiwara-Komoda Y, Tanaka D, Mitsumasu K, et al. (2010) The trehalose transporter 1 gene sequence is conserved in insects and encodes proteins with different kinetic properties involved in trehalose import into peripheral tissues. Insect Biochem Mol Biol 40: 30–37.
19. Kikawada T, Saito A, Kanamori Y, Nakahara Y, Ichi Iwata K, et al. (2007) Trehalose transporter 1, a facilitated and high-capacity trehalose transporter, allows exogenous trehalose uptake into cells. Proc Natl Acad Sci U S A 104: 11585–11590.
20. Kikuta S, Kikawada T, Hagiwara-Komoda Y, Nakashima N, Noda H (2010) Sugar transporter genes of the brown planthopper, *Nilaparvata lugens*: A facilitated glucose/fructose transporter. Insect Biochem Mol Biol 40: 805–813.
21. Kikuta S, Hagiwara-Komoda Y, Noda H, Kikawada T (2012) A novel member of the trehalose transporter family functions as an H^+-dependent trehalose transporter in the reabsorption of trehalose in Malpighian tubules. Front Physiol 3: 290–290.
22. Price DRG, Wilkinson HS, Gatehouse JA (2007) Functional expression and characterisation of a gut facilitative glucose transporter, NlHT1, from the phloem-feeding insect *Nilaparvata lugens* (rice brown planthopper). Insect Biochem Mol Biol 37: 1139–1148.
23. Price DRG, Tibbles K, Shigenobu S, Smertenko A, Russell CW, et al. (2010) Sugar transporters of the major facilitator superfamily in aphids; from gene prediction to functional characterization. Insect Mol Biol 19: 97–112.
24. Thorens B, Mueckler M (2010) Glucose transporters in the 21st Century. Am J Physiol Endocrinol Metab 298: E141–E145.
25. Pao SS, Paulsen IT, Saier MH (1998) Major facilitator superfamily. Microbiol Mol Biol Rev 62: 1–34.
26. Joost HG, Thorens B (2001) The extended GLUT-family of sugar/polyol transport facilitators: nomenclature, sequence characteristics, and potential function of its novel members. Mol Membr Biol 18: 247–256.
27. Mueckler M, Thorens B (2013) The SLC2 (GLUT) family of membrane transporters. Mol Aspects Med 34: 121–138.
28. Manolescu AR, Witkowska K, Kinnaird A, Cessford T, Cheeseman C (2007) Facilitated hexose transporters: New perspectives on form and function. Physiology 22: 234–240.
29. Gomez-Zurita J, Hunt T, Kopliku F, Vogler AP (2007) Recalibrated tree of leaf beetles (Chrysomelidae) indicates independent diversification of angiosperms and their insect herbivores. Plos One 2.
30. Uddin MM, Ulrichs C, Tokuhisa JG, Mewis I (2009) Impact of glucosinolate structure on the performance of the crucifer pest *Phaedon cochleariae* (F.). J Appl Bot Food Qual 82: 108–113.
31. Tremmel M, Mueller C (2013) The consequences of alternating diet on performance and food preferences of a specialist leaf beetle. J Insect Physiol 59: 840–847.
32. Termonia A, Hsiao TH, Pasteels JM, Milinkovitch MC (2001) Feeding specialization and host-derived chemical defense in Chrysomeline leaf beetles did not lead to an evolutionary dead end. Proc Natl Acad Sci USA 98: 3909–3914.
33. Whitman DW, Blum MS, Alsop DW (1990) Allomones: chemicals for defense. 289–351 p.
34. Noirot C, Quennedey A (1974) Fine-structure of insect epidermal glands. Annu Rev Entomol 19: 61–80.
35. Noirot C, Quennedey A (1991) Glands, gland-cells, glandular units - some comments on terminology and classification. Annales De La Societe Entomologique De France 27: 123–128.
36. Quennedey A (1998) Insect epidermal gland cells: Ultrastructure and morphogenesis; Harrison FWLM, editor. 177–207 p.
37. Laurent P, Braekman JC, Daloze D, Pasteels J (2003) Biosynthesis of defensive compounds from beetles and ants. Eur J Org Chem: 2733–2743.
38. Laurent P, Braekman JC, Daloze S (2005) Insect chemical defense. In: Schulz S, editor. Chemistry of Pheromones and Other Semiochemicals II. Schulz, S. ed: Springer, Berlin/Heidelberg. 167–229.
39. Pasteels JM, Braekman JC, Daloze D, Ottinger R (1982) Chemical defence in chrysomelid larvae and adults. Tetrahedron 38: 1891–1897.
40. Pasteels JM, Rowell-Rahier M, Braekman J-C, Daloze D (1994) Chemical defence of adult leaf beetles updated. In: Jolivet PH, Cox ML, Petitpierre E, editors. Novel aspects of the biology of Chrysomelidea. Dordrecht: Kluwer Academic Publishers. 289–301.
41. Deroe C, Pasteels JM (1982) Distribution of adult defense glands in chrysomelids (Coleoptera: Chrysomelidae) and its significance in the evolution of defense mechanisms within the family. J Chem Ecol 8: 67–82.
42. Pasteels JM, Rowell-Rahier M (1989) Defensive glands and secretions as taxonomical tools in the Chrysomelidae. Entomography 6: 423–432.
43. Renner K (1970) Über die ausstülpbaren Hautblasen der Larven von *Gastroidea viridula* De Geer und ihre ökologische Bedeutung (Coleoptera: Chrysomelidae). Beiträge zur Entomologie 20: 527–533.
44. Burse A, Schmidt A, Frick S, Kuhn J, Gershenzon J, et al. (2007) Iridoid biosynthesis in Chrysomelina larvae: Fat body produces early terpenoid precursors. Insect Biochem Mol Biol 37: 255–265.
45. Frick S, Nagel R, Schmidt A, Bodemann RR, Rahfeld P, et al. (2013) Metal ions control product specificity of isoprenyl diphosphate synthases in the insect terpenoid pathway. Proc Natl Acad Sci U S A 110: 1194–1199.
46. Strauss AS, Peters S, Boland W, Burse A (2013) ABC transporter functions as a pacemaker for sequestration of plant glucosides in leaf beetles. eLife: DOI:10.7554/eLife.01096.
47. Ibanez S, Gallet C, Després L (2012) Plant insecticidal toxins in ecological networks. Toxins 4: 228–243.
48. Bodemann RR, Rahfeld P, Stock M, Kunert M, Wielsch N, et al. (2012) Precise RNAi-mediated silencing of metabolically active proteins in the defence secretions of juvenile leaf beetles. Proc R Soc B 279: 4126–4134.
49. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, et al. (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456: 53–59.
50. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29: 644–652.

51. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, et al. (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc 8: 1494–1512.
52. Pertea G, Huang X, Liang F, Antonescu V, Sultana R, et al. (2003) TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. Bioinformatics 19: 651–652.
53. Zhang Z, Schwartz S, Wagner L, Miller W (2000) A greedy algorithm for aligning DNA sequences. J Comput Biol 7: 203–214.
54. Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. Genome Res 9: 868–877.
55. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, et al. (2012) The Pfam protein families database. Nucleic Acids Res 40: D290–D301.
56. Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, et al. (2008) The Pfam protein families database. Nucleic Acids Res 36: D281–D288.
57. Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, et al. (2003) The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res 31: 365–370.
58. O'Donovan C, Martin MJ, Gattiker A, Gasteiger E, Bairoch A, et al. (2002) High-quality protein knowledge resource: SWISS-PROT and TrEMBL. Brief Bioinform 3: 275–284.
59. Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305: 567–580.
60. Jones DT, Taylor WR, Thornton JM (1994) A model recognition approach to the prediction of all-helical membrane protein structure and topology. Biochemistry 33: 3038–3049.
61. Roshan U, Livesay DR (2006) Probalign: multiple sequence alignment using partition function posterior probabilities. Bioinformatics 22: 2715–2721.
62. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, et al. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61: 539–542.
63. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17: 754–755.
64. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688–2690.
65. Keeling CI, Henderson H, Li M, Yuen M, Clark EL, et al. (2012) Transcriptome and full-length cDNA resources for the mountain pine beetle, *Dendroctonus ponderosae* Hopkins, a major insect pest of pine forests. Insect Biochem Mol Biol 42: 525–536.
66. Langmead B (2010) Aligning short sequencing reads with Bowtie. Curr Protoc Bioinformatics Chapter 11: Unit 11.17.
67. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5: R80.
68. Morgan M, Anders S, Lawrence M, Aboyoun P, Pagès H, et al. (2009) ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. Bioinformatics 25: 2607–2608.
69. Anders S, Huber W (2010) Differential expression analysis for sequence count data. Genome Biol 11: R106.
70. Anders S, McCarthy DJ, Chen Y, Okoniewski M, Smyth GK, et al. (2013) Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. Nat Protoc 8: 1765–1786.
71. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25: 402–408.
72. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, et al. (2009) The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. Clin Chem 55: 611–622.
73. Kuhnle A, Muller C (2009) Differing acceptance of familiar and unfamiliar plant species by an oligophagous beetle. Entomol Exp Appl 131: 189–199.
74. Quackenbush J, Liang F, Holt I, Pertea G, Upton J (2000) The TIGR gene indices: reconstruction and representation of expressed gene sequences. Nucleic Acids Res 28: 141–145.
75. Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30: 3059–3066.
76. Zhao F Q, Keating AF (2007) Functional properties and genomics of glucose transporters. Curr Genomics 8: 113–128.
77. Garcia JC, Strube M, Leingang K, Keller K, Mueckler MM (1992) Amino acid substitutions at tryptophan 388 and tryptophan 412 of the HepG2 (Glut1) glucose transporter inhibit transport activity and targeting to the plasma membrane in *Xenopus* oocytes. J Biol Chem 267: 7770–7776.
78. Schuermann A, Keller K, Monden I, Brown FM, Wandel S, et al. (1993) Glucose transport activity and photolabelling with 3-[^{125}I]iodo-4-azidophenethylamido-7-*O*-succinyldeacetyl(LAPS)-forskolin of two mutants at tryptophan-388 and −412 of the glucose transporter GLUT1: dissociation of the binding domains of forskolin and glucose. Biochem J 290: 497–501.
79. Wilson-O'Brien AL, Patron N, Rogers S (2010) Evolutionary ancestry and novel functions of the mammalian glucose transporter (GLUT) family. BMC Evol Biol 10.
80. Petzold A, Reichwald K, Groth M, Taudien S, Hartmann N, et al. (2013) The transcript catalogue of the short-lived fish *Nothobranchius furzeri* provides insights into age-dependent changes of mRNA levels. BMC Genomics 14: 185.
81. Cura AJ, Carruthers A (2012) Role of monosaccharide transport proteins in carbohydrate assimilation, distribution, metabolism, and homeostasis. Compr Physiol 2: 863–914.
82. Douady CJ, Delsuc F, Boucher Y, Doolittle WF, Douzery EJP (2003) Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. Mol Biol Evol 20: 248–254.
83. Cuthbertson JM, Doyle DA, Sansom MS (2005) Transmembrane helix prediction: a comparative evaluation and analysis. Protein Eng Des Sel 18: 295–308.
84. Fraga A, Ribeiro L, Lobato M, Santos V, Silva JR, et al. (2013) Glycogen and glucose metabolism are essential for early embryonic development of the red flour beetle *Tribolium castaneum*. Plos One 8.
85. Leto D, Saltiel AR (2012) Regulation of glucose transport by insulin: traffic control of GLUT4. Nat Rev Mol Cell Biol 13: 383–396.
86. Price DRG, Karley AJ, Ashford DA, Isaacs HV, Pownall ME, et al. (2007) Molecular characterisation of a candidate gut sucrase in the pea aphid, *Acyrthosiphon pisum*. Insect Biochem Mol Biol 37: 307–317.

3.3 *Manuscript 3: Tissue-specific transcript profiling for ABC transporters in the sequestering larvae of the phytophagous leaf beetle Chrysomela populi*

Tissue-specific transcript profiling for ABC transporters in the sequestering larvae of the phytophagous leaf beetle *Chrysomela populi*

Anja S. Strauss[a,†], Ding Wang[a,†], Magdalena Stock[a,†], Rene R. Gretscher[a], Marco Groth[b], Wilhelm Boland[a], and Antje Burse[a*]

[a] Max Planck Institute for Chemical Ecology, Beutenberg Campus, Hans-Knoell-Str. 8, D-07745 Jena, Germany

[b] Leibniz Institute for Age Research – Fritz Lipmann Institute, Beutenbergstr. 11, D-07745 Jena, Germany

† These authors contributed equally to this study.

*Corresponding author. Phone: ++49(0) 3641-571265, Fax: ++49(0) 3641-571202, E-mail: aburse@ice.mpg.de

Abstract

Background:

Insects evolved ingenious adaptations to use extraordinary food sources. Particularly, the diet of herbivores enriched with noxious plant secondary metabolites requires detoxification mechanisms. Sequestration, which involves the uptake, transfer, and concentration of occasionally modified phytochemicals into specialized tissues or hemolymph, is one of the most successful detoxification strategies found in most insect orders. Due to the ability of ATP-binding cassette carriers to transport a wide range of molecules including phytochemicals and insecticides, it is highly likely that they play a role in this sequestration process. To shed light on the role of ATP-binding cassette carriers in sequestration, we describe an inventory of putative ABC transporters in various tissues in the sequestering juvenile poplar leaf beetle, *Chrysomela populi*.

Results: In the transcriptome of *C. populi*, we predicted 65 ABC transporters. To link the proteins with a possible function, we performed comparative phylogenetic analyses with ABC transporters of other insects and of humans. While tissue-specific profiling of each ABC transporter subfamily suggests that ABCB, D and G influence the plant metabolite absorption in the gut, ABCC with 14 members is the preferred subfamily responsible for the excretion of these metabolites *via* Malpighian tubules. Moreover, salicin, which is sequestered from poplar plants, is translocated into the defensive glands for further deterrent production. In these glands and among all identified ABC transporters, an exceptionally high transcript level was observed only for *Cpabc35* (*Cpmrp*). RNAi revealed the deficiency of other ABC pumps to compensate the function of *Cp*ABC35, demonstrating its key role during sequestration.

Conclusion: We provide the first comprehensive phylogenetic study of the ABC family in a phytophagous beetle species. RNA-seq data from different larval tissues propose the importance of ABC pumps to achieve a homeostasis of plant-derived compounds and offer a basis for future analyses of their physiological function in sequestration processes.

Introduction

Lipid bilayers form efficient barriers for cellular partitioning. The translocation across these membranous barriers is crucial for many aspects of cell physiology, including the uptake of nutrients, the elimination of waste products, or energy generation and cell signaling. The ATP-binding cassette (ABC) transporters constitute one of the largest families of membrane translocators [1]. The core functional unit of ABC proteins consists of four domains: two cytoplasmic domains containing the highly conserved nucleotide-binding domain (NBD), which is responsible for the ATP hydrolysis needed to provide energy for the transport cycle, and two transmembrane domains (TMD), each in most cases composed of six membrane-spanning helices which impart substrate specificity and translocation [2-4]. The NBD harbors several conserved sequence motifs from N- to C-terminus. These are the Walker A motif (also called P-loop) which is glycine-rich, a flexible loop with a conserved glutamine residue (Q-loop), the ABC signature (LSGGQ) motif (also called C-loop), the Walker B motif, and a conserved histidine residue (His-switch). The ABC signature motif is diagnostic for this family as it is present only in ABC transporters, while Walker A and B motifs are found in many other ATP-utilizing proteins. The domains are encoded by separate genes, either by genes encoding one NBD and one TMD whose products dimerize to form the functional transporter, or by genes encoding two NBDs and two TMDs on a single polypeptide.

71 In eukaryotic genomes, ABC genes are widely dispersed and highly conserved between
72 species, indicating that most of these genes have existed since the beginning of eukaryotic
73 evolution [5-8]. ABC transporters can be classified into subfamilies according to sequence
74 homology and domain topology. In eukaryotes eight major subfamilies have been defined:
75 ABCA to ABCH [9]. The eighth subfamily (H) was defined after the analysis of the genome of
76 the fruit fly *Drosophila melanogaster* [9]. Most ABC proteins transport a wide range of
77 compounds, either within the cell as part of a metabolic process into an intracellular
78 compartment (e.g. endoplasmic reticulum, mitochondria, and peroxisomes) or outside the cell for
79 transport processes to other organs. In humans, the known functions of ABC transporters include
80 cholesterol and lipid transport, multidrug resistance, antigen presentation, mitochondrial iron
81 homeostasis and the ATP-dependent regulation of ion channels [10-13]. Owing to the
82 importance of ABC transporters for cell functions, they are still extensively investigated in many
83 eukaryotes. In insects, one of the best studied ABC proteins is White, which is crucial for
84 pigment transfer in insect eyes [14-19]. As is known for *D. melanogaster*, ABC transporters
85 facilitate translocation of attractants for germ cell migration [20] or participate in the modulation
86 of the molting hormones' (ecdysteroids') signaling in insect tissues [21]. Furthermore, they seem
87 to be frequently implicated in insecticide resistance [22,23], such as in the DTT tolerance of the
88 *Anopheles* mosquitoes which transmit malaria agents [24] or in the tolerance against pest control
89 toxins from *Bacillus thuringiensis* which is reported of lepidopterans [25,26].
90 Although ABC transporters were previously analyzed in several insect species at
91 genome-wide level [27,28], profiles of the transcript levels of ABC transporters in non-model
92 insects are not available to date. For this study we analyzed the transcriptomic data with regard
93 to ABC transporters in a phytophagous leaf beetle species. Leaf beetles (Chrysomelidae *sensu*

94 *lato*; including the seed beetles Bruchidae) constitute together with the Cerambycidae (longhorn
95 beetles) and the Curculionoidea (weevils) the largest beetle radiation. These are known as
96 "Phytophaga" and represent roughly 40% of all the 350,000 described species [29]. Leaf beetles
97 mainly feed on green plant parts. The species of the leaf beetle taxon Chrysomelina, for example,
98 are adapted to use host plants' leaves as a food source during their whole life cycle [30].
99 Therefore, they have to be protected against both, the noxious effect of plant secondary
100 metabolites and attacks by their enemies. Some species evolved the ability to exploit the
101 phytochemicals for their own chemical defense [31-33]. The larvae of the poplar leaf beetle
102 *Chrysomela populi*, for example, take up the phenolglucoside salicin from salicaceaous food
103 plants. This precursor salicin is transported into nine pairs of exocrine, dorsal glands [34,35],
104 where the compound is converted into salicylaldehyde – a potent, volatile deterrent that repels
105 predators and prevents fatal microbial infections [33,36,37]. This process of sequestration
106 involves a complex influx-efflux transport network which guides plant-derived glucosides
107 through the insect body [38].

108 Although sequestration is a widespread phenomenon attributed to many insect orders, we
109 recently identified the first example of a transport protein essential for the translocation of
110 phytochemicals in insects [39]. The transporter belongs to the ATP-binding cassette transporter
111 family and functions in the defensive exocrine glands of juvenile poplar leaf beetles. Thus, the
112 comprehensive analysis of putative ABC transporters in the phytophagous *C. populi* larvae
113 provides implications for further studies on the predicted physiological functions of this
114 transporter class in sequestering insects, such as the incorporation and excretion mechanisms of
115 toxic compounds. For this reason, we present a complete inventory of ABC transporters based on
116 available *C. populi* transcriptome sequences. Detailed sequence comparisons of members of each

117 subfamily with those of the red flour beetle *Tribolium castanaeum*, the bark beetle *Dendroctonus*
118 *ponderosae*, the silk worm *Bombyx mori*, *D. melanogaster* and humans reveal their
119 correspondences. We, additionally, studied the expression profiles of ABC encoding transcripts
120 in various tissues by using next-generation sequencing in juvenile *C. populi* and propose a
121 function of ABC pumps in the sequestration process.
122
123 **Material and Methods**
124 **Rearing, maintaining and dissecting *Chrysomela populi***
125 *C. populi* (L.) was collected near Dornburg, Germany (+51°00'52.00", +11°38'17.00") on
126 *Populus maximowiczii x Populus nigra*. The beetles were kept in a light/dark cycle of 16 h light
127 and 8 h darkness (LD 16/8) at 18°C ± 2°C in light and 13°C ± 2°C in darkness.
128
129 **RNA isolation, library construction and sequencing**
130 Tissue samples from five *C. populi* larvae per biological replicate were collected as described by
131 Bodemann *et al.* [40]. Total RNA was extracted from defensive glands, fat body, Malpighian
132 tubules and gut tissue with the RNAqueous Micro Kit (Ambion, Life Technologies) according to
133 the manufacturers' instructions with the exception of 1% (v/v) ExpressArt NucleoGuard
134 (Amplification Technologies, Hamburg, Germany) added to the lysis buffer. The RNA integrity
135 was validated by electrophoresis on RNA 6000 Nano labchips on a Bioanalyzer 2100 (Agilent
136 Technologies). RNA concentrations were determined by employing a NanoView (GE-
137 Healthcare). Up to 5 µg of total RNA was then used for library preparation using TruSeq™ RNA
138 Sample Prep Kit according to the manufacturer's description. RNA sequencing (RNASeq) for
139 three biological samples per prepared tissue was done using Illumina next-generation sequencing

technique [41] on a HiSeq2000 (Illumina, Inc., San Diego, California USA) in 50-bp single read mode (two or three samples multiplexed in one lane).

Pooled total RNAs from adults (two males, two females), one pupa, and nine first-to-third-instars were used for paired-end sequencing. Therefore, the fragmentation step during library preparation of these RNAs was shortened to four minutes (seven minutes for all the remaining samples) to obtain longer fragments. This library was sequenced using a GAIIx (Illumina, Inc., San Diego, California USA) in 150-bp paired-end mode in one sample per lane. All reads were extracted in FastQ format and used for further analysis.

De novo assembly of _C. populi_'s transcriptome

The transcript catalogue of *C. populi* was generated as described in Rahfeld *et al.* [42]. Briefly, the paired-end reads were *de novo* assembled by applying the open source tool Trinity v2012-03-17 [43] with the following parameters: minimal contig length of 300 bp and the paired fragment length limited to 500 bp. In order to reconstruct full-length transcripts, we used the software TGICl (vJan.2009) [44] to reassemble the transcriptome output from Trinity with a minimum overlap length of 100 bp and sequence similarity of 90 percent. A summary of these results is given in Table S1.

Annotation of _de novo_ assembled transcript library and identification of ABC transporters

We annotated the above mentioned transcriptome by translating the cDNAs of the putative transcripts into all six possible open reading frames. This was achieved by applying transeq which is part of the EMBOSS package (v6.3.1). Afterwards, the protein sequences were searched against the Pfam database (update, Jan 2013) with an e-value cut-off of 1e-5 [45,46]. 102 hits

were obtained that belong to the protein family "PF00005" (ABC_tran domain). The exact NBDs of 90 ABC transporters were identified after the removal of 12 sequences highly similar to obligate intracellular Microsporida parasites found by BLASTx against the non-redundant protein sequence database (at NCBI). For the identification of NBDs, firstly, the highly conserved NBDs of the human (ABCA-ABCG, 48 amino acid sequences of NBD) and fruit fly (ABCH, 3 amino acid sequences of NBD) ABC transporters were retrieved from GenBank (NCBI) and chosen as 'homology search targets'. Then, the long coding sequence for each annotated beetle ABC transporter was determined by using getorf of the EMBOSS tools. Afterwards, these longest coding sequences and the chosen 'homology search targets' were aligned by applying the multiple sequence alignment program MAFFT v7.01 [47] (using option E-INS-i). Transcripts containing all five motifs of NBDs with roughly 170 amino acids were kept. Secondly, the remaining ABC transporter transcripts with incomplete motifs were checked again. Their six possible protein sequences were aligned to the chosen 'homology search targets' (with the same parameter E-INS-i, MAFFT). All sequences containing at least four motifs of NBDs and having a sequence length of more than 130 amino acids were selected and added to the other sequences for further studies. After removal of isoforms, the resulting beetle sequences were deposited as Transcriptome Shotgun Assembly project at DDBJ/EMBL/GenBank under the accession GARF00000000. The version described in this paper is the first version, GARF01000000.

Calculation of phylogenetic trees

The protein sequences were aligned by the G-INS-i methods from MAFFT with default parameters. To calculate the phylogenetic tree two programs were used: MrBayes v3.2.1 [48], a

186 program for Bayesian inference, and RAxML v7.2.8 [49], a program based on maximum-
187 likelihood inference. In RAxML, the best fit model of protein evolution was RTREVF with
188 gamma distribution for modeling rate heterogeneity. The best fit model was determined by the
189 best likelihood score under GAMMA (perl script ProteinModelSelection.pl, which was
190 downloaded from http://sco.h-its.org/exelixis/hands-On.html). The maximum-likelihood
191 phylogenetic tree was reconstructed with a bootstrap test of 1000 replicates in RAxML. In
192 addition, we used MrBayes with gamma-shaped rate variation and a proportion of invariable
193 sites to check the phylogenetic tree of RAxML.
194 For phylogenetic analysis of the ABC transporter subfamilies, we used the same methods along
195 with sequences of *T. castaneum*, the most closely related model species to *C. populi,* and
196 *H. sapiens*. The ABC transporter protein sequences of *T. castaneum* were retrieved after personal
197 communication with H. Merzendorfer (University of Osnabrück, Germany) [50].
198
199 **Expression profiling of putative ABC transporter transcripts**
200 Each 50-bp single-read dataset of four tissues (gut, defensive glands, fat body, Malpighian
201 tubules) contained three biological replicate samples. To compare the transcript expression levels
202 of the four tissues, we mapped the RNA-seq reads onto the (*de novo* assembled) transcriptome of
203 *C. populi* with the open source tool Bowtie v0.12.7 [51] using default parameters. Afterwards,
204 the R package DESeq [52,53] (which is part of the Bioconductor package [54]) was used to
205 detect differentially expressed transcripts in the four different tissues.
206 Based on the Lander/Waterman equation [55], the average coverage per base in each
207 transcript of each biological replicate was separately computed. The mean values of average
208 coverage of each replicate for each tissue, respectively, were compared to show the expression

209 levels of tissues. To compare these results with quantitative real-time PCR measurements, we
210 normalized the output from DESeq to the standards *CpelF4a* and *CpEF1alpha,* which were used
211 in quantitative real-time PCR, as described by Livak *et al*. [56]. The comparison of RNA-seq and
212 qRT-PCR values is shown in Figure S1.
213
214 **Quantitative real-time PCR (qPCR)**
215 Total RNA was extracted from larval tissue using an RNeasy MINI kit (Qiagen).
216 Complementary DNA was synthesized from DNA-digested RNA using the RNAqueous micro
217 kit (Life Technologies). Real-time PCR was performed using Brilliant II SYBR Green qPCR
218 Master Mix (Agilent) according to the manufacturer's instructions and the Mx3000P Real-Time
219 PCR system. *CpelF4a* and *CpEF1alpha* expression were used to normalize transcript quantities
220 (see Table S2 for primer sequences). Running conditions: 3' 94°C, 40 cycles [30" 94°C; 30"
221 60°C], melting curve with 1°C increase 60-95°C. Analyses were performed according to the
222 MIQE-guidelines [57].
223
224 **RNA interference of *Cpabc35 (Cpmrp)* in *C. populi* larvae**
225 The most abundant ABC transporter derived from the glandular tissue (*Cpabc35*) [39] was
226 analyzed *via* RNAi experiments. The sequence-verified plasmid pIB-*Cp*MRP was used to
227 amplify a 730bp fragment of *Cpmrp* dsRNA. As control, a *gfp* sequence was amplified from
228 pcDNA3.1/CT- GFP-TOPO (Invitrogen). The amplicons were subjected to *in vitro*-transcription
229 assays according to the instructions of the Ambion MEGAscript RNAi kit (Life Technologies;
230 see Table S2 for primer sequences). The resulting dsRNA was eluted three times with 50 µl of
231 injection buffer (3.5 mM Tris-HCl, 1 mM NaCl, 50 nM Na_2HPO_4, 20 nM KH_2PO_4, 3 mM KCl,

232 0.3 mM EDTA, pH 7.0) after nuclease digestion. The quality of dsRNA was checked by TBE-
233 agarose-electrophoresis.

234 First-instars of *C. populi* (3-4 days after hatching) with 3-5 mm body length (chilled on
235 ice) were injected with 0.25 µg of dsRNA by using a nanoliter microinjection system (WPI
236 Nanoliter 2000 Injector). Injections were made into the hemolymph next to the ventral side
237 between the pro- and mesothorax. Differential expression in the glandular tissue was analysed
238 10 days after RNAi treatment. Therefore, we carried out RNA sequencing (RNA-Seq). Two
239 biological replicates (pool of glandular tissue of 3 larvae, each) compared to two biological
240 replicates of gfp-control samples [41] were sequenced on a HiSeq2500 (Illumina, Inc., San
241 Diego, California USA) in 50-bp single read mode (two or three samples multiplexed in one
242 lane). All short reads again were extracted in FastQ format for further analysis.
243

244 **Analysis of differentially expressed genes in the glandular tissue of RNAi silenced *C. populi***
245 **larvae**

246 The short reads (sequenced in 50 bp single-mode) from the glandular tissue of the RNAi-
247 silenced (2 samples) as well as ds*gfp*-injected (2 samples) *C. populi* larvae were mapped onto
248 *C. populi*'s transcriptome using Bowtie [51]. The mapping results for the ABC transporter
249 transcripts were subjected to DESeq statistical analysis [52,53] by reading them into R statistics
250 software. Transcript counts were normalized to the effective library size. Afterwards, the
251 negative binomial testing was carried out to identify differentially expressed genes (DEGs). All
252 those genes were determined as differentially expressed when having an adjusted p-value less
253 than 0.1. From all DEGs, the annotated ABC transporters were selected and checked for co-
254 regulation.

Results and discussion

Identification of putative ABC transporters encoded in the transcript catalogue of *C. populi*

In our study, we focused on the distribution of ABC transporters in the different tissues of juvenile C. *populi* to functionally link each transcript to a certain tissue. For this purpose, we first identified potential ABC transporters in the *de novo* assembled transcript catalogue of the poplar leaf beetle. The transcriptome sequences were translated into all possible amino acid sequences and further processed as described in the method section. As a result, we predicted 65 ABC transporters for *C. populi*. This corresponds with previous studies on insects reporting, for example, 73 ABC transporter genes in the genome of *T. castanaeum* [50], 44 in *Anopheles gambiae* [28], 56 in *D. melanogaster*, 43 in *Apis mellifera*, or 51 in *Bombyx mori* [27]. The *C. populi* sequences were given temporary designations as numbered series in the form of *Cp*ABCxx (Table S3).

Phylogenetic analysis of the putative ABC transporters

Based on structural and functional similarity, ABC transporters in general can be grouped into subfamilies. In order to predict the subfamilies for the 65 identified ABC transporters in *C. populi*, we used their extracted NBDs for the multiple sequence alignments and then calculated the phylogenetic tree. Similarly to other insects and eukaryotes, we were able to show a division of the predicted transporters into 8 subfamilies (A-H) (Figure 1; Table 1). Members of ABCA, ABCE/F, ABCG and ABCH form distinct branches (bootstrap value ≥75 percent). ABCH forms a sister group of ABCA. The ABCC family segregates into two groups: ABCC1 contains NBDs1 and shows a similarity to the ABCD subfamily; ABCC2 contains NBDs2 and shows a similarity to the ABCB subfamily. Among the 65 putative ABC transporters from

278 *C. populi* we identified full, half and incomplete transporters. The distribution of domains in the
279 sequences is shown in detail for each subfamily in Table 2 and for each sequence in Table S3.
280 Next, we integrated human and other insect sequences into our phylogenetic trees. This
281 allowed us to group the putative *C. populi* ABC transporters with functionally characterized
282 proteins and, thus, to propose a substrate for the beetles´ proteins.
283 In the case of **subfamily A**, its members in humans are full transporters and implicated in
284 the transport processes of phospholipids, sterols, sphingolipids, bile salts, retinal derivatives
285 (restricted to ABCA4) and other lipid conjugates indispensable for many biological processes
286 [10,58-61]. In insects, both full and half transporters were identified whose physiological
287 function, however, is not yet understood [27]. In *C. populi* we predicted five transporters.
288 According to our phylogenetic analysis, ABCA proteins segregate into one branch containing
289 NBD1 and one branch with NBD2 (Figure 2). Human ABCAs form three groups (I, ABCA1-4,
290 7; II, ABCA5, 6, 8-10; III, ABCA12 and 13; bootstrap values of ≥75 percent particularly in the
291 NBD2 branch). Considering the beetles´ homologs, the tree shows that the majority of *C. populi*
292 and *T. castanaeum* sequences clusters to human ABCA3.
293 The **ABCB** subfamily contains ABCB1 (MDR1/P-glycoprotein) which is the first
294 characterized human ABC transporter to confer multidrug resistance (MDR) in cancer cells [62-
295 64] and which has been intensively studied ever since the discovery of cross-resistances after
296 selection with chemotherapeutics [65-69]. Later studies revealed additional ABCB transporters
297 as MDR proteins. Besides xenobiotic extrusion (ABCB1, 5, 8) [70-72], ABCB members are also
298 known in human biology for the translocation, for example, of phosphatidylcholine (ABCB4)
299 [73], bile salt (ABCB11) [74], peptides (TAP1 and 2, TAPL, and ABCB10) [75], porphyrins
300 (ABCB6) [76], or iron (ABCB7 and 8) [77-79]. In insects, several examples suggest the

301 involvement of P-glycoproteins in the resistance to insecticides used for crop protection
302 [22,23,80-88]. However, only few P-glycoprotein-like genes have been linked to a xenobiotic
303 substrate such as *Mdr49* and *Mdr65* of *D. melanogaster* with tolerance against colchicine and α-
304 amanitin [89,90]. Alternatively, Mdr49 can act as transporter for a germ cell attractant in fruit
305 flies [20].

306 Similar to other insects, the eight sequences from *C. populi* encode full and half
307 transporters. Bootstrapping of the ABCB phylogenetic tree in Figure 3 and Figure S1 (together
308 with ABCC) produced nodes with ≥69 bootstrap percentage supporting segregation of this
309 subfamily into three groups containing human and insect ABCB sequences. Group I forms a
310 clade with full transporters including the human ABCB1, 4, 5, 11 and the fruit fly Mdr proteins.
311 Group II encompasses the human ABCB8 and 10 as well as TAP1, 2 and TAPL. The human
312 TAP1:2 heterodimers pump antigenic peptides from the cytosol into the lumen of the
313 endoplasmic reticulum and have a key role in antigen processing in the adaptive immune system.
314 TAPL mediates peptide transfer in lysosomes, but its physiological function is still an open
315 question [91]. Based on our phylogenetic analysis, we found no homologs to TAP sequences
316 (bootstrap value 99 percent) in the insects, but homologs to TAPL and *Hs*ABC10 were
317 identified. In accordance with the literature, we can also speculate that TAPL is the ancestor of
318 the TAP family [91]. Group III in our phylogenetic tree of the subfamily B comprises the human
319 ABCB6 and 7.

320 Full transporters of striking functional diversity are found in the **ABCC** subfamily. In
321 humans thirteen ABCC members were identified, nine of which are referred to as multidrug
322 resistance proteins (MRP) ('short' MRPs (ABCC4, -5, -11 and -12); 'long' MRPs (ABCC1, -2, -
323 3, -6 and -10) [12,92,93]. Some ABCC members not considered as MRPs have unique functions.

The cystic fibrosis transmembrane conductance regulator (CFTR/ABCC7), for example, functions as an epithelial ATP-gated chloride channel [94,95]. ABCC8 and ABCC9 are assembled as sulfonylurea receptors into ATP-sensitive K^+ channels and are coupled to the gating mechanism of the ion-conducting pore [96]. In insects, ABCC members are thought to be involved in the translocation of xenobiotics and phytochemicals [22-25,39,88,97]. As observed in the red flour beetle *T. castanaeum* and the spider mite *Tetranychus urticae*, the ABCC subfamily in *C. populi* with 29 putative members has undergone an expansion (Table 1, 2). In our phylogenetic tree, the NBDs1 and NBDs2 form distinct clusters (bootstrap value of 100 percent; Figure 4, Figure S1). In each NBD-group, bootstrapping showed that the vast majority of insect sequences cluster together with CFTR and ABCC4. Generally, ABCC4 is a lipophilic anion pump whose physiological substrates include cyclic nucleotides, prostaglandins, and steroid conjugates. But it also contributes to drug resistance in certain types of human cancer cells [92]. This implies a broad substrate spectrum for the ABCC4-like insect transporters. Our phylogenetic analysis, including ABCC homologs from *D. melanogaster*, *B. mori* and *Dendroctonus ponderosae* exhibited high similarity to one another. However, we could reveal a statistically supported branch unique to beetles (bootstrap value 61 percent; NBD 1 in Figure S2).

Members of the **ABCD** subfamily are involved in the translocation of fatty acids into peroxisomes [98]. The ABC transporters are half-size and assemble mostly as a homodimer after posttranslational transport to peroxisomal membranes. ABCD4 is not a peroxisomal membrane protein but an ER-resident protein that mediates translocation of lipid molecules essential for lipid metabolism in the ER [99]. As in humans and other insects, *C. populi* contains two half transporters. Because they are homologous to the human peroxisomal and *T. castanaeum*

347 transporters, a similar function can be inferred in poplar leaf beetles. No orthologous insect
348 sequence could be grouped to ABCD4 (Figure S3).

349 The **ABCE** and **ABCF** proteins comprise a pair of linked NBDs but lack TMDs.
350 Therefore, they are not involved in molecule transport, but they are active in a wide range of
351 other functions pivotal for cell viability. For example, the human ABCE1 not only acts as a
352 ribonuclease L inhibitor, it also regulates RNA stability, viral infection, tumor cell proliferation,
353 anti-apoptosis, translation initiation, elongation, termination, and ribosome recycling [100]. In
354 *D. melanogaster*, the ABCE homolog Pixie plays a catalytic role in the assembly of protein
355 complexes required for translation initiation [101]. All genomes of multicellular eukaryotes
356 analyzed to date possess one ABCE gene [102]. In the transcript catalogue of *C. populi*, one
357 complete ABCE protein has been predicted. The NBDs of *Cp*ABC45 are highly conserved with
358 the respective NBDs of the human ABCE1 and *T. castanaeum* ABCE-3A (Figure S3). Among
359 the subfamily ABCF involved in translation initiation and elongation in humans [102], we found
360 three putative members each with two NBDs that are highly similar to the transporters of human
361 and *T. castanaeum* suggesting functional proteins used in similar physiological processes in the
362 cell.

363 The **ABCG** subfamily in humans is comprised of five half transporters. While the
364 homodimer ABCG2 is a multidrug transporter with a wide substrate specificity [70], the
365 homodimers ABCG1 and ABCG4 and the heterodimer ABCG5:ABCG8 translocate cholesterol
366 and other sterole derivatives [103-106]. In insects, ABCG transporters are essential for the
367 translocation of ommochromes for the pigmentation of eyes and body coloration. In
368 *D. melanogaster*, for example, the half transporter White forms heterodimers with Scarlet or
369 Brown, each of which is responsible for the transport of another type of ommochrome precursor

370 to pigment granules [14-16]. In silkworms, White-orthologs (Bm-ok) are responsible for the
371 translocation of uric acid for accumulation in urate granules in epidermal cells, resulting in
372 opaque white coloration of the larval skin [19,107]. In *D. melanogaster*, E23 encodes a
373 transporter capable of modulating the ecdysone response with consequences for the circadian
374 transcription of clock genes [21,108].

375 The phylogenetic analysis revealed that the majority of the chosen insect sequences,
376 including predicted ABCG proteins from *C. populi*, cluster together with the human ABCG1 and
377 ABCG4 (bootstrap value of 98 percent) (Figure 5). Several insect ABCG candidates form a
378 branch with the human ABCG5:ABCG8. Also E23 from *D. melanogaster* clusters in this branch
379 together with *Tc*ABCG-8A. Silencing of *Tc*ABCG-8A resulted in molting defects, premature
380 compound eye development, aberrant wing development and lethality, suggesting a function in
381 the regulation of ecdysteroid-mediated effects [50]. Because *Cp*ABC49 is homologous to
382 *Tc*ABCG-8A and *Dm*-E23, it allows the expectation of a similar function for this protein in *C.*
383 *populi*. In addition, the insect ABCG proteins (White, Brown, and Scarlet) involved in the
384 transfer of ommochrome precursors form a separate branch (bootstrap value of 93 percent). In
385 accordance with the observation of *T. castanaeum* [50], in *C. populi* a Brown ortholog is also
386 missing. Interestingly, not a single analyzed insect sequence clusters with the human multidrug
387 efflux transporter ABCG2.

388 The transporters of the **ABCH** subfamily were observed only in insects and not in
389 humans [6,8,109]. The ABCH subfamily of *C. populi* includes three putative ABC transporters
390 that are highly similar to those of *T. castaneum* (Figure 2).
391

392 **RNA-seq analyses reveal tissue-specific expression of ABC transporters in juvenile**
393 ***C. populi***

394 To link the above suggested functions for the *C. populi* ABC proteins to those which are
395 differentially expressed in the larval tissues of *C. populi*, we carried out a comprehensive
396 transcriptome sequencing of different tissues dissected from the poplar leaf beetle. All raw
397 sequence data (in the following called reads) are listed in Table S4. The resulting expression
398 patterns of all identified ABC transporters in intestinal tissue, Malpighian tubules, fat body and
399 defensive glands is depicted in Figure 6. It shows that among the 65 predicted ABC transporters,
400 43 are expressed at least in one of the tested tissues with more than 25 normalized read counts
401 per base (25-fold sequence coverage).

402 Five transcripts were found to be abundant in all tested tissues which suggest their
403 essential role in cellular processes. Among them is, for example, *Cp*ABC4 which was classified
404 as member of the ABCA subfamily. According to our phylogenetic analysis, the closest human
405 homologs, which are involved in lipid translocation, are clustered into group I of the NBD1
406 branch (Figure 2). Although the NBD2 of *Cp*ABC4 clusters to ABCA12 and 13, the sequence
407 comparison (using BLAST) of the complete sequence supported the homology of *Cp*ABC4 to
408 human ABCA members of group I. Additionally, *Cp*ABC49 as ABCD candidate was highly
409 expressed in all larval tissues, as well. It is homologous to the human ABCD1 and 2 and,
410 therefore, presumably linked to the transport of very long chains of fatty acids in peroxisomes
411 (Figure S3) [98]. Furthermore, we detected in all larval tissues abundantly expressed transcripts
412 encoding soluble ABC proteins: *Cp*ABC52 as a member of the ABCE and *Cp*ABC57 and
413 *Cp*ABC58 as members of the ABCF subfamily. Also, in the red flour beetle, the *Tc*ABCE-3A
414 and *Tc*ABCF-2A transcripts were abundant throughout all life stages and highly abundant in the

415 adult intestinal/excretory tissues and carcass [50]. Furthermore, the silencing of *Tc*ABCE-3A as
416 well as *Tc*ABCF-2A resulted in growth arrest and mortality of the beetles. Thus, ABCE and
417 ABCF proteins are essential for cellular functions in all insect tissues including initiation of
418 translation [100,101] and ribosome biogenesis [110].
419
420 In the following, we describe differential expression of putative ABC transporters in the different
421 larval tissues:
422 ***Gut tissue***
423 We found 17 transcripts abundant in the intestinal tissue of *C. populi* encoding members of the
424 following subfamilies: one sequence of ABCA, two of ABCB, eight of ABCC, five of ABCG
425 (Figure 6). The existence of ABC transporters in the gut influences the absorption and
426 bioavailability of nutrients, ions and plant derived compounds. The predicted ABCA-similar
427 sequence *Cp*ABC5 exhibits a high mRNA level only in the gut tissue. Its deduced protein
428 clusters together with *Tc*ABCA-9A/B of *T. castanaeum* (Figure 2). The silencing of these two
429 red flour beetle genes resulted in high mortality and severe defects in wing and elytra
430 development, depending on the developmental stage of treatment. This indicates an essential
431 function for cell physiology, but a ligand has not been identified for these proteins to date [50].
432 The closest homolog in humans is ABCA3 which is related to phospholipid transfer but also to
433 the modulation of cell susceptibility to chemotherapy of tumors [10,111,112]. Thus, *Cp*ABC5
434 may have a special function in this tissue, in addition to a role in lipid trafficking.
435 The highest transcript level of ABC transporters in the intestinal tissue was detected for
436 *CpABC12* which was classified into the subfamily B. It is also expressed in Malpighian tubules
437 but ten times less. *Cp*ABC12 is a full transporter, and together with human ABCB1 (MDR1, P-

438 glycoprotein), 4, 5 and 11 as well as the *D. melanogaster* Mdr proteins forms group I (Figure 3).
439 Though ABCB4 acts in humans as a transporter for phospholipids in the liver [73], it is involved
440 in the zebrafish's cellular resistance to noxious chemicals [113]. Except for ABCB11, which is a
441 bile salt transporter [74], all the vertebrate ABCB members of group II can confer multidrug
442 resistance [70-72]. Taking into account the Mdr proteins related to toxin tolerance in *D.*
443 *melanogaster*, we hypothesize a function in the translocation of phytochemicals for *Cp*ABC12 in
444 the gut of *C. populi*. *Cp*ABC7 is the second ABCB candidate with a high expression level in the
445 gut, albeit not as high as *Cp*ABC12. Moreover, *Cp*ABC7 is 3 times more highly expressed in the
446 Malpighian tubules than in the gut. Because *Cp*ABC7 is homologous to the human mitochondrial
447 ABCB6 in group III (Figure 3), which facilitates porphyrin transport [76], the beetle protein
448 could possess the specificity for structurally similar substrates.

449 All ABCC candidates highly expressed in the gut tissue cluster together with the
450 multidrug resistance associated protein ABCC4 in humans. Interestingly, the silencing of
451 members of the C-subfamily in *T. castanaeum* failed to result in detectable phenotypes, perhaps
452 due to overlapping substrate selectivity of the transporters [50].

453 All five ABCG candidates highly expressed in the larval gut tissue cluster together with
454 the human ABCG1 and ABCG4. These proteins are involved in sterol homeostasis. Among these
455 *C. populi* ABCG proteins, *Cpabc55* showed the most elevated transcript level. It is homologous
456 to *Tc*ABCG-4C whose involvement in the transport of lipids to the cuticle has been suggested
457 and, thus, that it is required for the formation of a waterproof barrier in the epicuticle [50].
458 *Cpabc55* is also highly expressed in glands and fat body tissue but not in the Malpighian tubules.
459 The expression of *Tcabcg-4c* was higher in intestinal/excretory tissues than in carcass tissue [50].
460 The function of the other four ABCG transporters cannot be predicted from our analyses.

461 However, it has been demonstrated recently that an ABCG1-homolog in the fungus *Grossmannis*
462 *clavigera* confers tolerance to monoterpenes which contributes to the fungus' ability to cope with
463 the chemical defence of its host plant [114]. Therefore, the ABCG proteins' specificity in insects
464 may not be limited to sterols or lipids but may have a broader substrate spectrum - that is not
465 known to date. Besides trafficking of physiological substrates, the identified ABC transporters in
466 the gut tissue may also play a critical role in regulating the absorption of plant secondary
467 metabolites or influence the effectiveness of pesticides in the phytophagous *C. populi*.
468
469 ***Malpighian tubules***
470 Insect Malpighian tubules are critical for osmoregulation. Moreover, the tubules have the
471 capability to excrete actively a broad range of organic solutes and xenobiotics, such as
472 insecticides. Additionally, they play a significant role in immunity by sensing bacterial infections
473 and mounting an effective killing response by secretion of antimicrobial peptides [115]. We
474 found 21 transcripts abundant in the Malpighian tubules of *C. populi* encoding members of the
475 following subfamilies known to contain multidrug resistance proteins: four of ABCB, 14 of
476 ABCC, three of ABCG (Figure 6).
477 Among the four predicted ABCB members displaying a high mRNA level in the
478 Malpighian tubules, two, *Cp*ABC7 and 12, were already described in the gut section above. The
479 third candidate, *Cp*ABC8, is most similar to human mitochondrial ABCB10 of group II (Figure
480 3). For ABCB10 different roles have been suggested, including protection against toxic reactive
481 oxygen species, heme synthesis, or peptide transport [75,116,117]. For this tissue, we speculate
482 that it is involved in antimicrobial peptide transfer. The forth ABCB protein, *Cp*ABC9, clusters
483 in group III together with the human mitochondrial ABCB7 which is involved in the iron-sulfur

484 cluster assembly essential for multiple metabolic pathways throughout the cell (Figure 3)
485 [77,78]. RNAi of the homologous *Tc*ABCB-5A demonstrated the pivotal function of this gene in
486 the red flour beetle: its down-regulation resulted in severe morphological defects and high
487 mortality depending on the developmental stage treated [50]. Hence, the three most likely
488 mitochondrial localized ABCB candidates, namely *Cp*ABCB7-9, are proposed to be of vital
489 importance in the cells. However, for *Cp*ABC12, which is a full transporter and clusters in group
490 I related to xenobiotic resistance, we can predict a similar function in the larval excretion system.
491 Most putative ABC transporter transcripts identified in *C. populi* are present at a high
492 level in the excretion system of the juvenile beetles compared to the other tissues. Particularly,
493 the 14 candidates belonging to the ABCC subfamily are the most highly transcribed when
494 compared to other subfamilies in this tissue. They all cluster to the human ABCC4 which has a
495 broad substrate spectrum. Remarkably, one of the highly expressed candidates, *Cp*ABC16,
496 clusters in our phylogeny together with *Cp*ABC35 which is involved in the accumulation of
497 plant-derived precursors such as salicin (Figure 4) [39]. Therefore, it is tempting to speculate a
498 role for *Cp*ABC16 in the excretion of phytochemicals in *C. populi* larvae.
499 Among the three candidates of the G-subfamily, two are highly transcribed only in the
500 Malpighian tubules: *Cp*ABC54 is a homolog of *Tc*ABCG-9B from the White group and
501 *Cp*ABC62 is homologous to *Tc*ABCG-9A from the Scarlet group (Figure 5). RNAi targeting
502 *Tcabcg-9a* or *b* resulted in both cases not only in white eyes but also in a whitish appearance of
503 the Malpighian tubules due to the absence of tryptophan metabolites/kynurenine and pteridines.
504 These eye pigment precursors are stored and processed in the larval tubules before being released
505 for further conversion into pigments in the developing adult eyes [118-120]. In addition, in
506 *D. melanogaster* White is expressed in intracellular vesicles in tubule principal cells, suggesting

that White participates in vesicular transepithelial transport of cGMP [121]. *Cp*ABC57 is the only ABCG candidate that is also expressed in the intestine and belongs to the human ABCG1 and ABCG4 branch (Figure 5). Taken together, the conspicuous overrepresentation of drug-resistance associated proteins, including the ABCC4-like proteins together with the members of the subfamily B (*Cp*ABC12) and G (*Cp*ABC57), in the excretion system suggests a role for these candidates in the extrusion of xenobiotics or phytochemicals from the larval body.

Fat body

The fat body of insects is a polymorphic tissue. It performs a vast array of fundamental activities in the intermediary metabolism and is involved in maintaining the homeostasis of hemolymph proteins, lipids, and carbohydrates [122]. Predominantly, the storage of lipid reserves in the form of glycogen and triglycerides is essential in the life of holometabolous insects, primarily in their survival of metamorphosis [123]. In humans, members of the subfamilies A, B, D and G are known to be involved in lipid transport [10,124]. In principal, we found the expression of ABC transporters in the larval fat body of *C. populi* to be low compared to the other tested tissues (Figure 6). From the ABCB subfamily, we identified in the fat body only *Cp*ABC8 exhibiting a low transcript level comparable to that of the Malpighian tubules. As described above, it clusters with the human mitochondrial ABCB10 which is associated with different functions, also described above, but not particularly with lipid transfer.

From ABCG we found *Cp*ABC51 and *Cp*ABC55 with high expression in the fat body, both clustering to human ABCG1 and ABCG4 (Figure 5). Only one sequence was exclusively expressed in this body part, namely *Cp*ABC41, a member of the ABCC4-like group (Figure 4).

529 Other ABCC4 members which are highly expressed in this tissue are the homologous *Cp*ABC16
530 and *Cp*ABC35. *Cp*ABC35 is known to translocate phytochemicals [39].
531 Noticeably, we found high expression of putative ABCH genes (*CpABC64, CpABC65)* in
532 the fat body tissue. Up to now the function of this insect specific subfamily has been unclear.
533 However, RNAi targeting *Tc*ABCH-9C in the flour beetle revealed a lethal, desiccated
534 phenotype similar to the silencing of *Tc*ABCG-4C mentioned above. This ABCH member also
535 seems to be involved directly or indirectly in the transport of lipids from epidermal cells to the
536 cuticle [50]. Based on our data we can hypothesize a role for ABC transporters in phytochemical
537 translocation (by members of the ABCC4-like group and the ABCG candidate), in cuticle
538 formation (by members of the ABCH subfamily) in the fat body, but not particularly in the lipid
539 storage of this tissue. Transporters which are important for this function might be lowly
540 expressed and therefore not detected in our analyses.
541
542 ***Defensive glands***
543 The nine pairs of defensive glands enable larvae of *C. populi* to chemically defend themselves
544 *via* deterrent secretions. Each of these dorsal glands is composed of several secretory cells which
545 are attached to a large reservoir. The anti-predatory effect of the secretions can be attributed to
546 salicylaldehyde synthesized within the reservoir by a few enzymatic reactions from the pre-toxin
547 salicin, which is sequestered from the host plant [34,36]. Recent studies have identified
548 *Cp*ABC35/*Cp*MRP which is essential for the sequestration of salicin [39]. It is associated with
549 the accumulation of the plant-derived metabolite in intracellular storage vesicles. Intriguingly,
550 *CpABC35* is the only predominant transcript in the defensive glands of *C. populi* (Figure 6). Its
551 expression level lies far beyond all other ABC transporters in all tissues. There are four

552 additional predicted ABCC proteins with high expression clustering to the human ABCC4 but
553 not to *Cp*ABC35. In *T. castanaeum* another member of the ABCC4 cluster has been identified as
554 playing a role in the production of secretions in odiferous stink glands. The silencing of
555 TC015346/*Tc*ABCC-6A in *T. castanaeum* resulted in a strong reduction of alkenes in the
556 secretions produced by abdominal and prothoracic glands (Figure 4) [125]. Thus, the hypothesis
557 can be advanced that ABC transporters functioning in the formation of secretions seem to be a
558 widespread phenomenon in insects.

559 Besides ABCC proteins, members of the subfamilies B, G and H also have elevated
560 mRNA levels in the defensive glands. *Cp*ABC13 is a member of the B-subfamily exclusively
561 expressed in the defensive glands. It clusters into the human transporter group II, particularly
562 with the human mitochondrial ABCB8 (Figure 3). ABCB8 is known to be responsible for iron
563 transport and doxorubicin resistance in melanoma cells *via* the protection of mitochondrial DNA
564 from doxorubicin-induced DNA damage [126].

565 Among the five candidates of the ABCG also possessing a high mRNA level in the
566 defensive glands, *Cp*ABC56, 59 and 61 are expressed only in this tissue. *Cp*ABC59 clusters to
567 the human ABCG5:ABCG8 that pump cholesterol and other sterol derivatives, and all of the four
568 other proteins cluster to human ABCG1 and 4, which may have a broader substrate spectrum
569 including xenobiotics (Figure 5).

570 Remarkably, the expression of putative ABCH genes (*Cpabc64, Cpabc65)* was almost 3
571 times higher in the glandular tissue compared to the fat body tissue. Owing to this, the two
572 ABCH proteins may have a special function as yet unknown in the defensive glands, but they
573 may also be associated with the formation of the cuticle reservoir for storage of secretions.

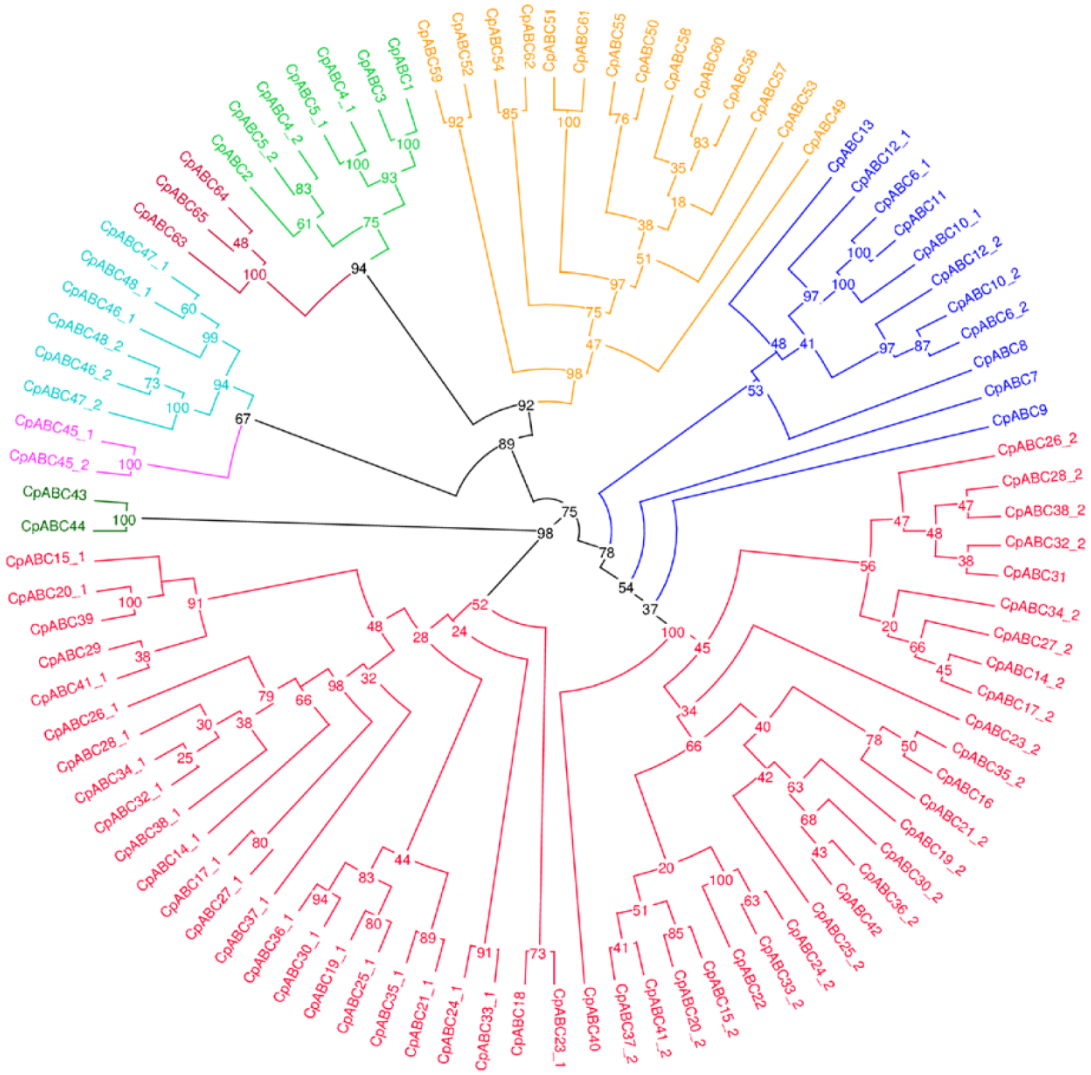
574 Furthermore, in this tissue there are also ABC candidates potentially associated with the
575 translocation of phytochemicals or other xenobiotics.
576
577 **RNAi with predominant ABC transporter – *Cpabc35* (*Cpmrp*)**
578 Conspicuously, only one ABC gene, namely *Cpabc35,* displays an exceedingly high transcript
579 level in the defensive glands of *C. populi*. As recently described [39], its function and key role in
580 the sequestration of defensive compound precursors has been demonstrated. In order to test
581 cooperative or compensation effects of other ABC genes, we performed RNAi silencing
582 experiments for *Cpabc35*. Ten days after the injection of *Cpabc35*-dsRNA and *gfp*-dsRNA,
583 glandular tissues were dissected and two biological replicates for each treatment were sequenced.
584 The normalized counts of all transcripts of all samples were calculated. Thereafter, the $\log_2$ fold-
585 changes of the silenced ABC transporter (*gfp*-injected samples as control) and adjusted p-values
586 were determined using the DESeq package. In all samples (either in RNA-seq or quantitative
587 real-time PCR experiments), we observed varying transcript levels corresponding to the
588 individual biological variance and diversity despite similar developmental stage or living
589 conditions during sample preparation.
590 The silencing of *Cpabc35* resulted in a significant decrease of its own transcript level
591 (adjusted p-value (padj)=7.31E-15). One additional ABC transporter, *Cpabc50*, belonging to
592 subfamily G, was determined as differentially expressed (slight upregulation). There is indication
593 of a positive regulation mechanism in response to the downregulation of *Cpabc35*. However,
594 *Cpabc50* could not compensate the function of the salicin translocation into storage vesicles,
595 and, hence, its function remains unclear. Overall, *Cpabc35* is an exclusive and highly specific

596 transporter used in the sequestration process, which explains its extraordinarily high transcript
597 level in the defensive glands.
598

599 ## Conclusion

600 Phytophagous beetles are adapted to cope with the chemical defense of their host plant. The
601 larvae of the poplar leaf beetle, *C. populi*, evolved the ability to sequester the plant-derived
602 compound salicin and to use it for their own defense against their enemies. The sequestration
603 process proceeds *via* barriers with different selectivity. While the uptake from the gut lumen into
604 the hemolymph together with the excretion by Malpighian tubules is non-selective, the
605 translocation into the defensive glands is selective. In these glands two barriers must be passed: a
606 selective membrane on the hemolymph side and a non-selective membrane on the side towards
607 the cuticle reservoir containing the defensive secretions. Based on our analyses, we predicted
608 specific ABC proteins that are related to the translocation of plant-derived compounds in the
609 larvae. In the gut of *C. populi*, genes of the subfamilies A, B, D, and G are predominantly
610 expressed. All these ABC candidates have been linked in our phylogenetic trees with proteins
611 known to be associated with xenobiotic or drug resistance and which may, therefore, contribute
612 to the non-selective translocation into the larval hemocoel. But depending on the localization of
613 the proteins in the intestinal cells, they may also take part in the detoxification of plant
614 metabolites or pesticides by back-exporting them into the gut lumen. The Malpighian tubules are
615 dominated by candidates of subfamilies B, C and G. In particular, members of the multidrug-
616 related ABCC4-group are present in great numbers in this tissue, which suggests a role in the
617 previously observed non-selective phytochemical extrusion in the excretion system.
618 Furthermore, in fat body we found an ABCC member homologous to a known salicin ABC

619 transporter, implying a similar function in this tissue as well. In the defensive glands this salicin-
620 transporting ABCC4-like protein *Cp*ABC35/*Cp*MRP is extraordinarily highly expressed in
621 comparison to the other tested tissues. It is localized intracellularly in storage compartments of
622 the gland cells and accumulates salicin in these vesicles for further exocytosis into the glandular
623 reservoir. *Cp*ABC35 has a broad substrate spectrum of phytochemicals and controls the non-
624 selective barrier into the reservoir. The differential expression analysis of *Cp*ABC35-silenced
625 defensive glands in comparison to control samples corroborated the observation that the function
626 cannot be compensated by any other ABC transporter with overlapping substrate selectivity in
627 this particular compartment of the glandular cells. The occurrence of other drug-resistant related
628 ABC transporters in the defensive glands may contribute to the selectivity in the membrane of
629 the hemolymph side of the glandular cells by extruding unused plant-derived compounds from
630 these cells. Thus, ABC transporters are key components in the homeostasis control of
631 phytochemicals in the sequestering poplar leaf beetle larvae.
632
633 **Author contributions**
634 A.B., A.S.S., and M.S. designed the experiments, interpreted the results and wrote the
635 manuscript. A.S.S. performed the quantitative real-time PCR experiments, interpreted the results
636 and performed the RNAi experiment. D.W. performed the identification of ABC transporters and
637 phylogenetic analyses. M.S. *de novo* assembled the transcriptome and performed and interpreted
638 the differential expression analysis. D.W., M.S. and A.B. interpreted the phylogenetic trees.
639 M.G. carried out the cDNA library preparation and Illumina sequencing. R.R.G. dissected the
640 larvae and prepared the various tissues for RNA sequencing. A.B., A.S.S., M.S., and W.B.
641 supervised the work, and all authors revised the manuscript.



642

643 **Figure 1: Eight subfamilies of 65 putative ABC transporters of *C. populi*.** Some transporters
644 contain two NBDs (NBD1 as *Cp*-ABCX_1 and NBD2 as *Cp*-ABCX_2), others contain only one
645 NBD. The software RAxML was used to calculate the phylogenetic tree under the RTREVF
646 amino acid substitution model (see methods). Numbers at nodes represent bootstrap values.

647

648

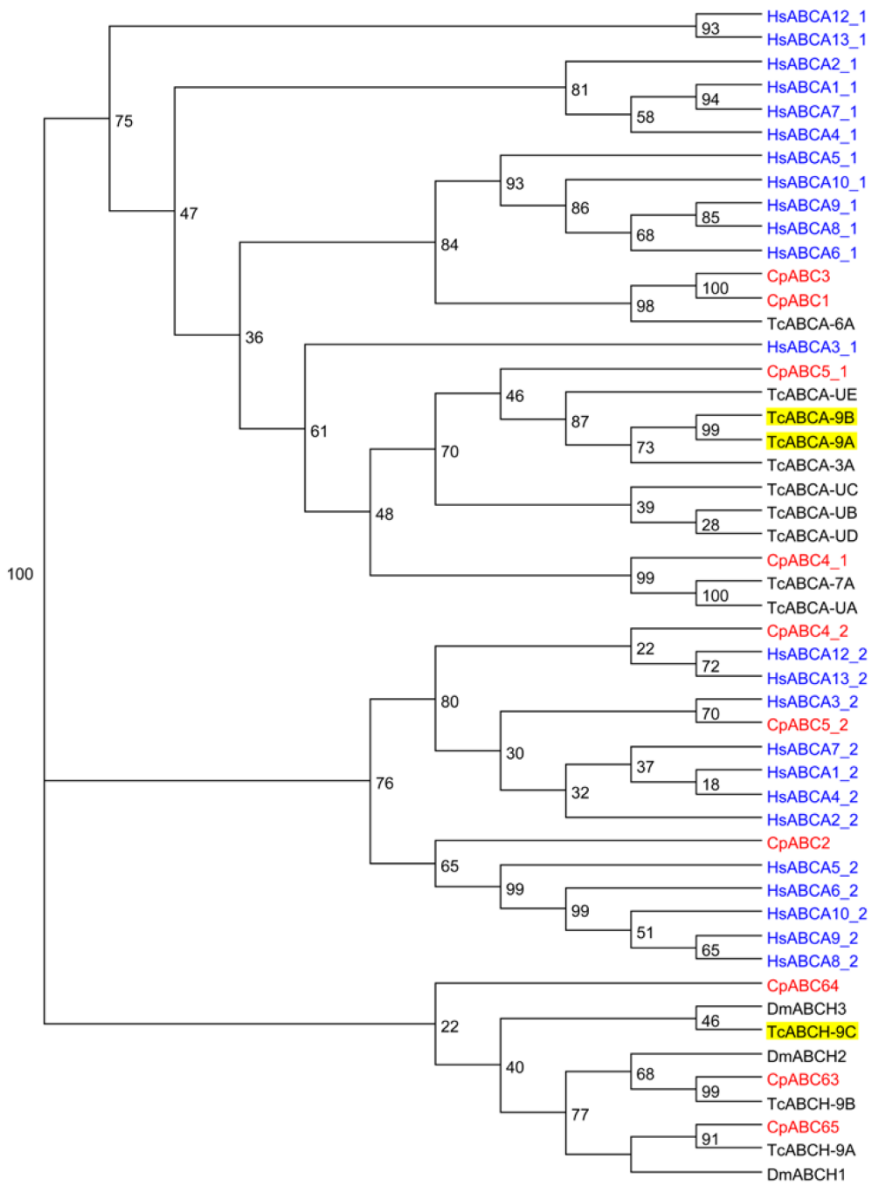
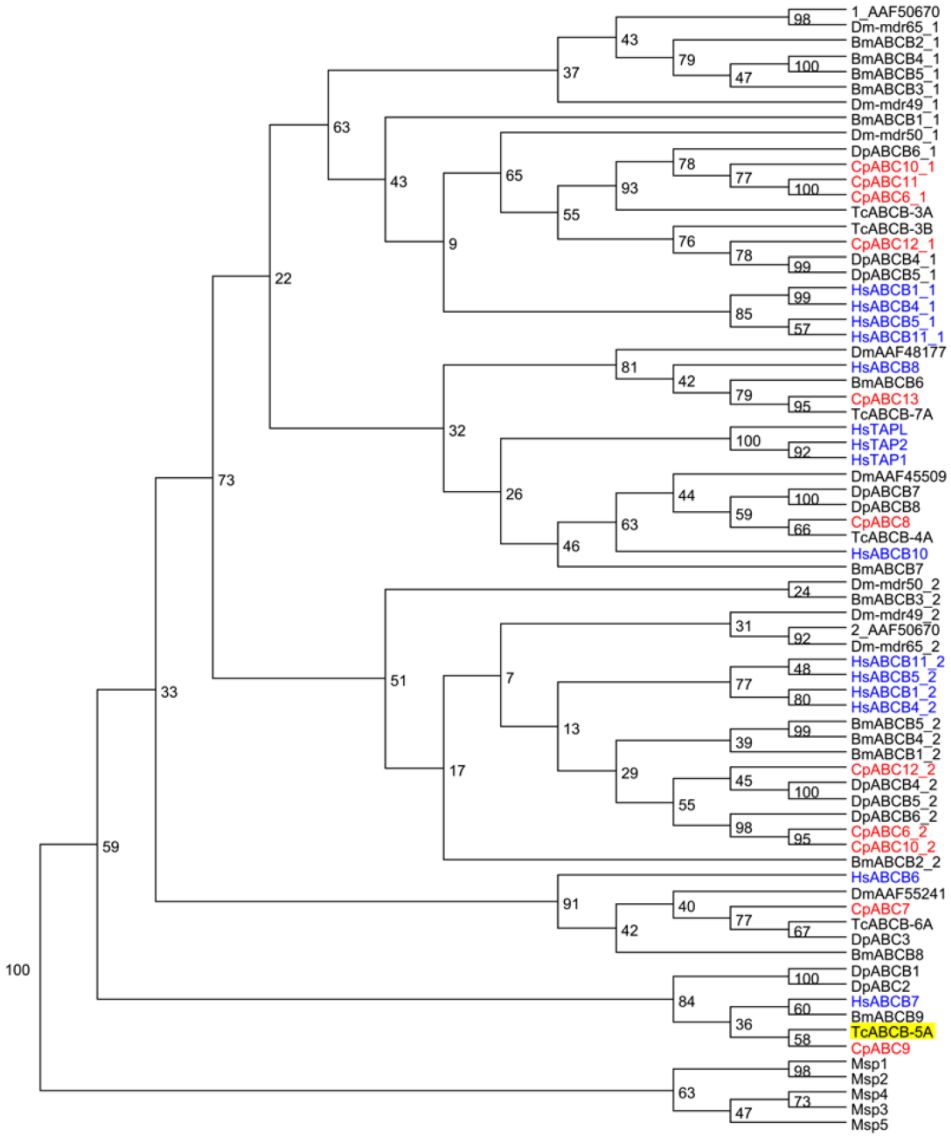


Figure 2: Phylogenetic tree of the proteins of subfamilies ABCA and ABCH of three species, *C. populi, H. sapiens* and *T. castaneum.* Red: *C. populi*; blue: *H. sapiens (*ABCA*)*, *D. melanogaster* (ABCH); black: *T. castaneum*; Numbers at nodes represent bootstrap values. Highlighted sequences: the transporters of *T. castaneum* with known function.



674 **Figure 3: Phylogenetic tree of ABCB of three species, *C. populi*, *H. sapiens* and**
675 ***T. castaneum.*** Red: *C. populi*; blue: *H. sapiens (*ABCA*)*; black: *T. castaneum*; Numbers at nodes
676 represent bootstrap values. Highlighted sequences: the ABC transporters of *T. castaneum* with
677 known function.

678

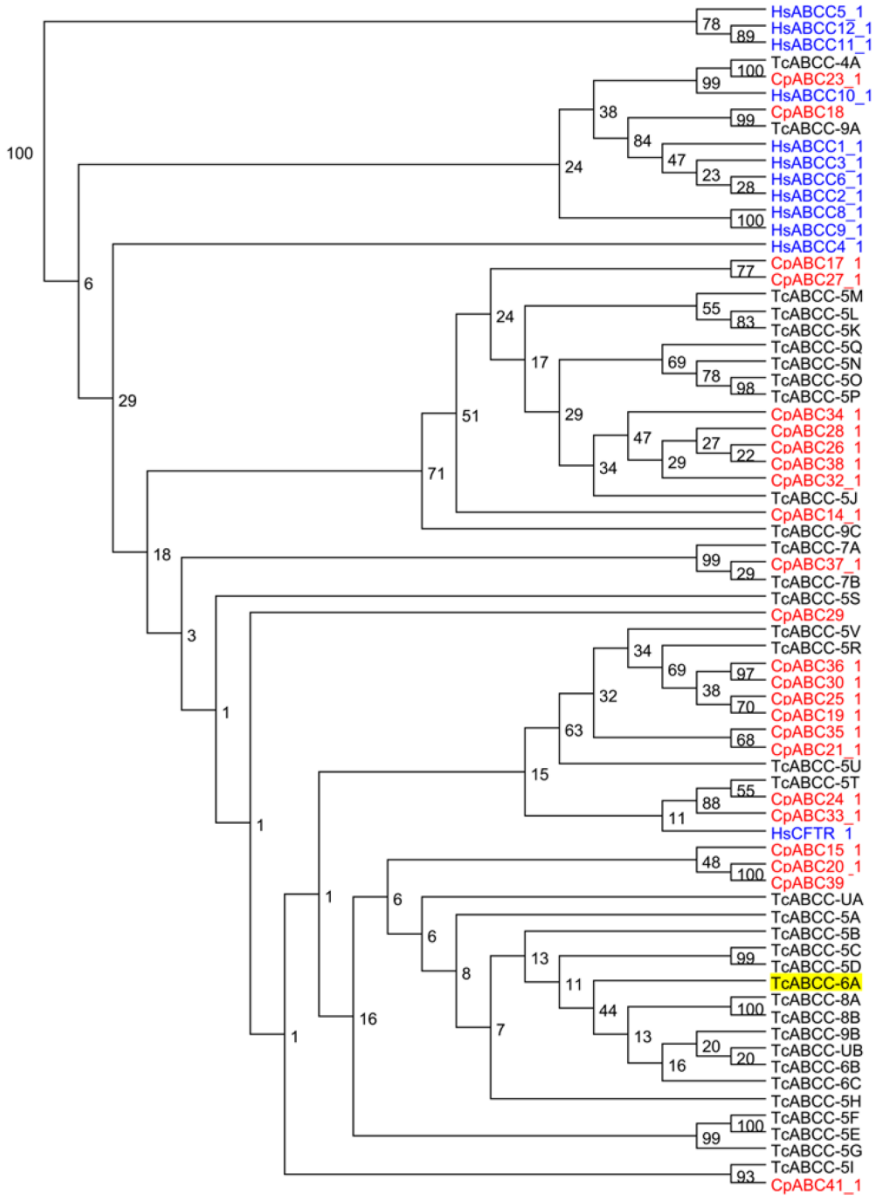


Figure 4: Phylogenetic tree of NBD1 from ABCC of three species, *C. populi, H. sapiens* and *T. castaneum*. Red: *C. populi*; blue: *H. sapiens*; black: *T. castaneum*; Numbers at nodes represent bootstrap values. Highlighted sequences: the ABC transporters with known function of *T. castaneum*.

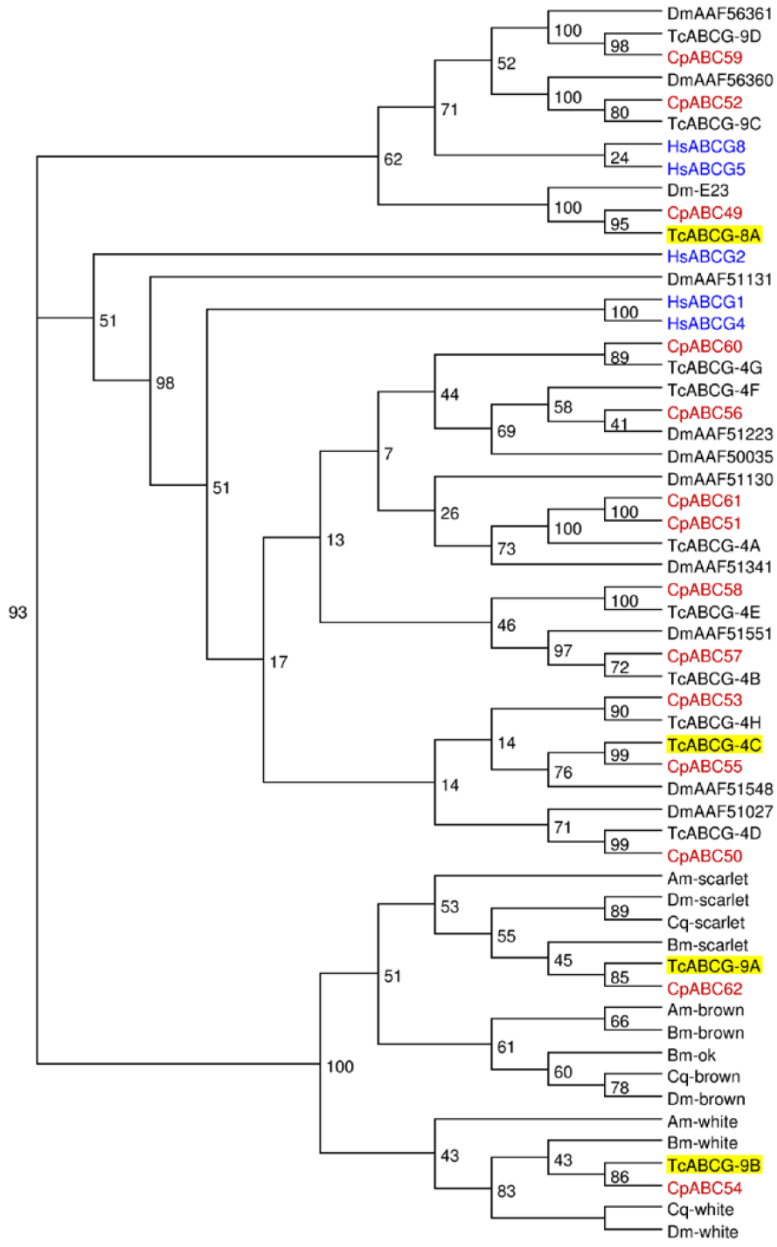


Figure 5: Phylogenetic tree of ABCG subfamily transporters of three species, *C. populi* (red)**, *H. sapiens* and *T. castaneum*.** blue: *H. sapiens*; black: *T. castaneum*; Numbers at nodes represent bootstrap values.

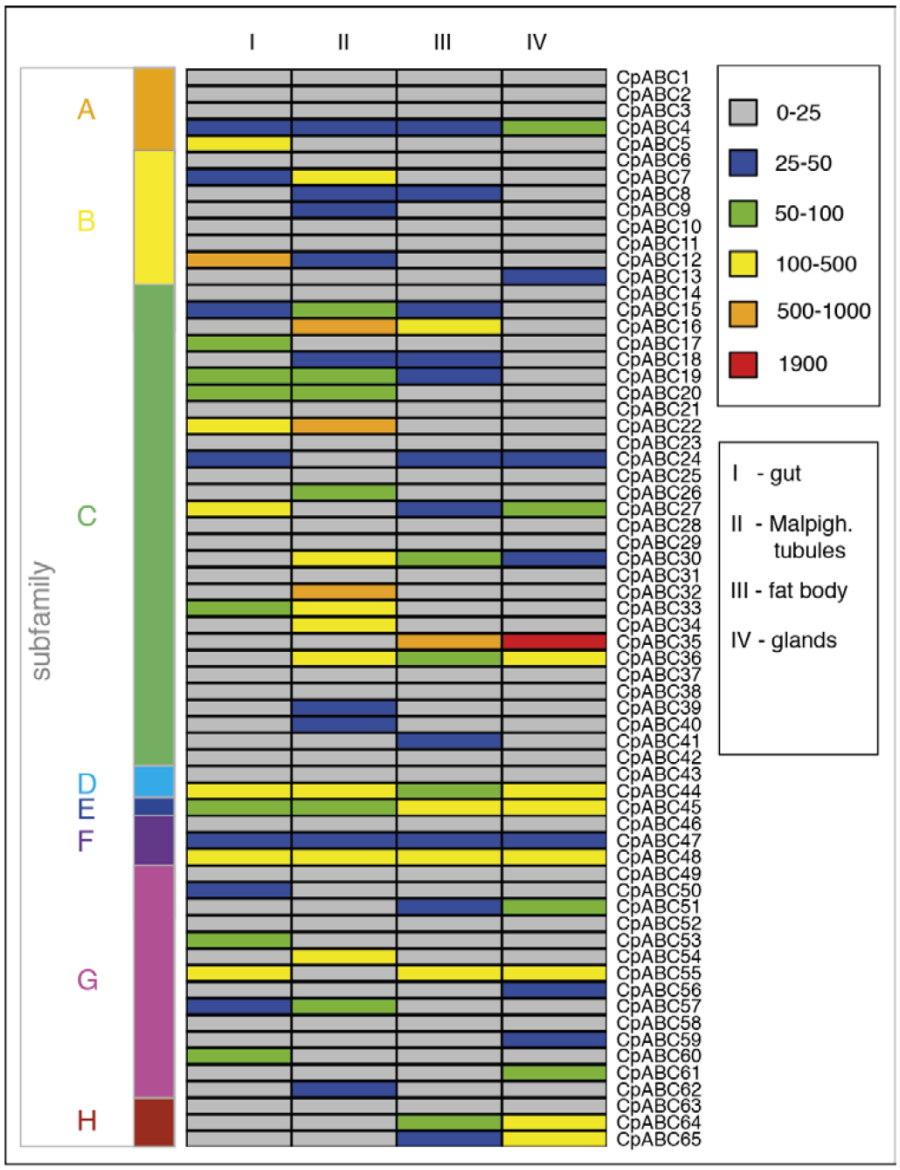


Figure 6: Heatmap of the expression profiles of the 65 putative ABC genes of *C. populi*. Values are shown for four different tissues: glands, gut, fat body and Malpighian tubules. Counts of RNA-seq reads (derived from three replicates for each tissue) normalized to the effective library size. Expression levels are illustrated by a six grade color scale representing the sequence coverage for each transcript for each tissue, respectively.

748

749 **Tables**

750 **Table 1. Subfamilies of ABC genes in 8 species** (Numbers were derived from [127]).
751

Species	ABCA	ABCB	ABCC	ABCD	ABCE	ABCF	ABCG	ABCH	total
S. cerevisiae	0	4	6	2	2	6	10	0	30
C. elegans	7	24	9	5	1	3	9	0	58
D. pulex	4	7	7	3	1	4	24	15	65
T. urticae	10	4	39	2	1	3	23	22	103
D. melanogaster	10	8	14	2	1	3	15	3	56
T. castaneum	10	6	35	2	1	3	13	3	73
C. populi	5	8	29	2	1	3	14	3	65
H. sapiens	12	11	12	4	1	3	5	0	48

752

753 **Table 2.** Distribution of the ABC transporter domains of *C. populi* in eight subfamilies. Full-
754 trans, full transporters; Half-trans, half transporters; NBD, nucleotide-binding domain; TMD,
755 transmembrane domain; 2*NBD+1*TMD, two NBDs and one TMD (example).

	Full-trans	**Half-trans**	**2*NBD**	**2*NBD+1*TMD**	**1*NBD**	**1*NBD+2*TMD**	**total**
ABCA	2	1	0	0	2	0	**5**
ABCB	3	5	0	0	0	0	**8**
ABCC	18	4	0	3	3	1	**29**
ABCD	0	2	0	0	0	0	**2**
ABCE	0	0	1	0	0	0	**1**
ABCF	0	0	3	0	0	0	**3**
ABCG	0	12	0	0	2	0	**14**
ABCH	0	3	0	0	0	0	**3**
total	23	27	4	3	7	1	**65**

756

757

758

759

760

761

762 **References**

763

764 1. Dassa E (2011) Natural history of ABC systems: not only transporters. In: Sharom FJ, editor. Essays
765 in Biochemistry: Abc Transporters. pp. 19-42.
766 2. Holland IB (2011) ABC transporters, mechanisms and biology: an overview. In: Sharom FJ, editor.
767 Essays in Biochemistry: Abc Transporters. pp. 1-17.
768 3. Jin MS, Oldham ML, Zhang QJ, Chen J (2012) Crystal structure of the multidrug transporter P-
769 glycoprotein from Caenorhabditis elegans. Nature 490: 566-+.
770 4. Al-Shawi MK (2011) Catalytic and transport cycles of ABC exporters. In: Sharom FJ, editor. Essays in
771 Biochemistry: Abc Transporters. pp. 63-83.
772 5. Dassa E, Bouige P (2001) The ABC of ABCs: a phylogenetic and functional classification of ABC
773 systems in living organisms. Res Microbiol 152: 211-229.
774 6. Dean M, Annilo T (2005) Evolution of the ATP-binding cassette (ABC) transporter superfamily in
775 vertebrates. Annual Review of Genomics and Human Genetics. pp. 123-142.
776 7. Wang B, Dukarevich M, Sun EI, Yen MR, Saier MH (2009) Membrane Porters of ATP-Binding
777 Cassette Transport Systems Are Polyphyletic. J Membr Biol 231: 1-10.
778 8. Annilo T, Chen ZQ, Shulenin S, Costantino J, Thomas L, et al. (2006) Evolution of the vertebrate ABC
779 gene family: analysis of gene birth and death. Genomics 88: 1-11.
780 9. Dean M, Rzhetsky A, Allikmets R (2001) The human ATP-binding cassette (ABC) transporter
781 superfamily. Genome Res 11: 1156-1166.
782 10. Quazi F, Molday RS (2011) Lipid transport by mammalian ABC proteins. In: Sharom FJ, editor.
783 Essays in Biochemistry: Abc Transporters. pp. 265-290.
784 11. Sharom FJ (2011) The P-glycoprotein multidrug transporter. In: Sharom FJ, editor. Essays in
785 Biochemistry: Abc Transporters. pp. 161-178.
786 12. Slot AJ, Molinski SV, Cole SPC (2011) Mammalian multidrug-resistance proteins (MRPs). In: Sharom
787 FJ, editor. Essays in Biochemistry: Abc Transporters. pp. 179-207.
788 13. Tamaki A, Ierano C, Szakacs G, Robey RW, Bates SE (2011) The controversial role of ABC
789 transporters in clinical oncology. In: Sharom FJ, editor. Essays in Biochemistry: Abc
790 Transporters. pp. 209-232.
791 14. Ewart GD, Howells AJ (1998) ABC transporters involved in transport of eye pigment precursors in
792 Drosophila melanogaster. Abc Transporters: Biochemical, Cellular, and Molecular Aspects 292:
793 213-224.
794 15. Mackenzie SM, Brooker MR, Gill TR, Cox GB, Howells AJ, et al. (1999) Mutations in the white gene
795 of Drosophila melanogaster affecting ABC transporters that determine eye colouration.
796 Biochimica Et Biophysica Acta-Biomembranes 1419: 173-185.
797 16. Ohare K, Murphy C, Levis R, Rubin GM (1984) DNA-SEQUENCE OF THE WHITE LOCUS OF
798 DROSOPHILA-MELANOGASTER. J Mol Biol 180: 437-455.
799 17. Quan GX, Kanda T, Tamura T (2002) Induction of the white egg 3 mutant phenotype by injection of
800 the double-stranded RNA of the silkworm white gene. Insect Mol Biol 11: 217-222.
801 18. Sumitani M, Yamamoto DS, Lee JM, Hatakeyama M (2005) Isolation of white gene orthologue of
802 the sawfly, Athalia rosae (Hymenoptera) and its functional analysis using RNA interference.
803 Insect Biochemistry and Molecular Biology 35: 231-240.
804 19. Tatematsu K-i, Yamamoto K, Uchino K, Narukawa J, Iizuka T, et al. (2011) Positional cloning of
805 silkworm white egg 2 (w-2) locus shows functional conservation and diversification of ABC
806 transporters for pigmentation in insects. Genes Cells 16: 331-342.

807 20. Ricardo S, Lehmann R (2009) An ABC Transporter Controls Export of a Drosophila Germ Cell
808 Attractant. Science 323: 943-946.
809 21. Hock T, Cottrill T, Keegan J, Garza D (2000) The E23 early gene of Drosophila encodes an ecdysone-
810 inducible ATP-binding cassette transporter capable of repressing ecdysone-mediated gene
811 activation. Proc Natl Acad Sci U S A 97: 9519-9524.
812 22. Buss DS, Callaghan A (2008) Interaction of pesticides with p-glycoprotein and other ABC proteins:
813 A survey of the possible importance to insecticide, herbicide and fungicide resistance. Pestic
814 Biochem Physiol 90: 141-153.
815 23. Labbe R, Caveney S, Donly C (2011) Genetic analysis of the xenobiotic resistance-associated ABC
816 gene subfamilies of the Lepidoptera. Insect Mol Biol 20: 243-256.
817 24. Jones CM, Toe HK, Sanou A, Namountougou M, Hughes A, et al. (2012) Additional Selection for
818 Insecticide Resistance in Urban Malaria Vectors: DDT Resistance in Anopheles arabiensis from
819 Bobo-Dioulasso, Burkina Faso. Plos One 7.
820 25. Gahan LJ, Pauchet Y, Vogel H, Heckel DG (2010) An ABC Transporter Mutation Is Correlated with
821 Insect Resistance to Bacillus thuringiensis Cry1Ac Toxin. PLoS Genet 6.
822 26. Atsumi S, Miyamoto K, Yamamoto K, Narukawa J, Kawai S, et al. (2012) Single amino acid
823 mutation in an ATP-binding cassette transporter gene causes resistance to Bt toxin Cry1Ab in
824 the silkworm, Bombyx mori. Proc Natl Acad Sci U S A 109: E1591-E1598.
825 27. Liu SM, Zhou S, Tian L, Guo EN, Luan YX, et al. (2011) Genome-wide identification and
826 characterization of ATP-binding cassette transporters in the silkworm, Bombyx mori. BMC
827 Genomics 12.
828 28. Roth CW, Holm I, Graille M, Dehoux P, Rzhetsky A, et al. (2003) Identification of the Anopheles
829 gambiae ATP-binding cassette transporter superfamily genes. Molecules and Cells 15: 150-
830 158.
831 29. Gomez-Zurita J, Hunt T, Kopliku F, Vogler AP (2007) Recalibrated tree of leaf beetles
832 (Chrysomelidae) indicates independent diversification of angiosperms and their insect
833 herbivores. Plos One 2.
834 30. Termonia A, Hsiao TH, Pasteels JM, Milinkovitch MC (2001) Feeding specialization and host-
835 derived chemical defense in Chrysomeline leaf beetles did not lead to an evolutionary dead
836 end. Proc Natl Acad Sci U S A 98: 3909-3914.
837 31. Opitz SEW, Mueller C (2009) Plant chemistry and insect sequestration. Chemoecology 19: 117-154.
838 32. Pasteels JM, Duffey S, Rowell-Rahier M (1990) Toxins in chrysomelid beetles possible evolutionary
839 sequence from *de novo* synthesis to derivation from food-plant chemicals. J Chem Ecol 16:
840 211-222.
841 33. Pasteels JM, Rowell-Rahier M, Braekman JC, Dupont A (1983) Salicin from host plant as precursor
842 of salicyl aldehyde in defensive secretion of chrysomeline larvae. Physiol Entomol 8: 307-314.
843 34. Kuhn J, Pettersson EM, Feld BK, Burse A, Termonia A, et al. (2004) Selective transport systems
844 mediate sequestration of plant glucosides in leaf beetles: A molecular basis for adaptation and
845 evolution. Proc Natl Acad Sci U S A 101: 13808-13813.
846 35. Smiley JT, Horn JM, Rank NE (1985) ECOLOGICAL EFFECTS OF SALICIN AT 3 TROPHIC LEVELS - NEW
847 PROBLEMS FROM OLD ADAPTATIONS. Science 229: 649-651.
848 36. Michalski C, Mohagheghi H, Nimtz M, Pasteels JM, Ober D (2008) Salicyl alcohol oxidase of the
849 chemical defense secretion of two chrysomelid leaf beetles - Molecular and functional
850 characterization of two new members of the glucose-methanol-choline oxidoreductase gene
851 family. J Biol Chem 283: 19219-11928.
852 37. Gross J, Podsiadlowski L, Hilker M (2002) Antimicrobial activity of exocrine glandular secretion of
853 Chrysomela larvae. J Chem Ecol 28: 317-331.

38. Discher S, Burse A, Tolzin-Banasch K, Heinemann SH, Pasteels JM, et al. (2009) A versatile transport network for sequestering and excreting plant glycosides in leaf beetles provides an evolutionary flexible defense strategy. ChemBioChem 10: 2223-2229.
39. Strauss AS, Peters S, Boland W, Burse A, Dicke M (2013) ABC transporter functions as a pacemaker for sequestration of plant glucosides in leaf beetles. eLife 2.
40. Bodemann RR, Rahfeld P, Stock M, Kunert M, Wielsch N, et al. (2012) Precise RNAi-mediated silencing of metabolically active proteins in the defence secretions of juvenile leaf beetles. Proceedings of the Royal Society B-Biological Sciences 279: 4126-4134.
41. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, et al. (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456: 53-59.
42. Rahfeld Peter, Kirsch Roy , Kugel Susann , Wielsch Natalie, Stock Magdalena, et al. (2014) Independent recruitments of oxidases from the GMC oxidoreductase family enabled the evolution of iridoid defence in leaf beetle larvae (subtribe Chrysomelina). Proceedings of the Royal Society B-Biological Sciences.
43. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29: 644-652.
44. Pertea G, Huang X, Liang F, Antonescu V, Sultana R, et al. (2003) TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. Bioinformatics 19: 651-652.
45. Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, et al. (2008) The Pfam protein families database. Nucleic Acids Res 36: D281-D288.
46. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, et al. (2012) The Pfam protein families database. Nucleic Acids Res 40: D290-D301.
47. Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30: 3059-3066.
48. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17: 754-755.
49. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688-2690.
50. Broehan G, Kroeger T, Lorenzen M, Merzendorfer H (2013) Functional analysis of the ATP-binding cassette (ABC) transporter gene family of Tribolium castaneum. BMC Genomics 14.
51. Langmead B (2010) Aligning short sequencing reads with Bowtie. Curr Protoc Bioinformatics Chapter 11: Unit 11.17.
52. Anders S, Huber W (2010) Differential expression analysis for sequence count data. Genome Biol 11: R106.
53. Anders S, McCarthy DJ, Chen Y, Okoniewski M, Smyth GK, et al. (2013) Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. Nature protocols 8: 1765-1786.
54. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5: R80.
55. Lander ES, Waterman MS (1988) Genomic mapping by fingerprinting random clones: a mathematical analysis. Genomics 2: 231-239.
56. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2-DELTADELTACT method. Methods (Orlando) 25: 402-408.
57. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, et al. (2009) The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. Clin Chem 55: 611-622.
58. Albrecht C, Viturro E (2007) The ABCA subfamily - gene and protein structures, functions and associated hereditary diseases. Pflugers Archiv-European Journal of Physiology 453: 581-589.

902 59. Broccardo C, Luciani MF, Chimini G (1999) The ABCA subclass of mammalian transporters.
903 Biochimica Et Biophysica Acta-Biomembranes 1461: 395-404.
904 60. Zarubica A, Trompier D, Chimini G (2007) ABCA1, from pathology to membrane function. Pflugers
905 Archiv-European Journal of Physiology 453: 569-579.
906 61. Tsybovsky Y, Orban T, Molday RS, Taylor D, Palczewski K (2013) Molecular Organization and ATP-
907 Induced Conformational Changes of ABCA4, the Photoreceptor-Specific ABC Transporter.
908 Structure 21: 854-860.
909 62. Juliano RL, Ling V (1976) A surface glycoprotein modulating drug permeability in Chinese hamster
910 ovary cell mutants. Biochim Biophys Acta 455: 152-162.
911 63. Ueda K, Cardarelli C, Gottesman MM, Pastan I (1987) Expression of a full-length cDNA for the
912 human "MDR1" gene confers resistance to colchicine, doxorubicin, and vinblastine. Proc Natl
913 Acad Sci U S A 84: 3004-3008.
914 64. Gros P, Croop J, Housman D (1986) Mammalian multidrug resistance gene: complete cDNA
915 sequence indicates strong homology to bacterial transport proteins. Cell 47: 371-380.
916 65. Gottesman MM, Fojo T, Bates SE (2002) Multidrug resistance in cancer: Role of ATP-dependent
917 transporters. Nature Reviews Cancer 2: 48-58.
918 66. Yu M, Ocana A, Tannock IF (2013) Reversal of ATP-binding cassette drug transporter activity to
919 modulate chemoresistance: why has it failed to provide clinical benefit? Cancer and
920 Metastasis Reviews 32: 211-227.
921 67. Zhou SF (2008) Structure, function and regulation of P-glycoprotein and its clinical relevance in
922 drug disposition. Xenobiotica 38: 802-832.
923 68. Aller SG, Yu J, Ward A, Weng Y, Chittaboina S, et al. (2009) Structure of P-glycoprotein reveals a
924 molecular basis for poly-specific drug binding. Science 323: 1718-1722.
925 69. Seeger MA, van Veen HW (2009) Molecular basis of multidrug transport by ABC transporters.
926 Biochimica Et Biophysica Acta-Proteins and Proteomics 1794: 725-737.
927 70. Sarkadi B, Homolya L, Szakacs G, Varadi A (2006) Human multidrug resistance ABCB and ABCG
928 transporters: Participation in a chemoimmunity defense system. Physiol Rev 86: 1179-1236.
929 71. Chen KG, Valencia JC, Gillet JP, Hearing VJ, Gottesman MM (2009) Involvement of ABC
930 transporters in melanogenesis and the development of multidrug resistance of melanoma.
931 Pigment Cell & Melanoma Research 22: 740-749.
932 72. Kawanobe T, Kogure S, Nakamura S, Sato M, Katayama K, et al. (2012) Expression of human ABCB5
933 confers resistance to taxanes and anthracyclines. Biochem Biophys Res Commun 418: 736-741.
934 73. Oude Elferink RP, Paulusma CC (2007) Function and pathophysiological importance of ABCB4
935 (MDR3 P-glycoprotein). Pflugers Archiv : European journal of physiology 453: 601-610.
936 74. Stieger B, Meier Y, Meier PJ (2007) The bile salt export pump. Pflugers Archiv : European journal of
937 physiology 453: 611-620.
938 75. Herget M, Tampe R (2007) Intracellular peptide transporters in human - compartmentalization of
939 the "peptidome". Pflugers Archiv-European Journal of Physiology 453: 591-600.
940 76. Krishnamurthy P, Xie T, Schuetz JD (2007) The role of transporters in cellular heme and porphyrin
941 homeostasis. Pharmacol Ther 114: 345-358.
942 77. Pondarre C, Antiochos BB, Campagna DR, Greer EL, Deck KM, et al. (2006) The mitochondrial ATP-
943 binding cassette transporter Abcb7 is essential in mice and participates in cytosolic iron-sulfur
944 cluster biogenesis. Hum Mol Genet 15: 953-964.
945 78. Ye H, Rouault TA (2010) Human Iron-Sulfur Cluster Assembly, Cellular Iron Homeostasis, and
946 Disease. Biochemistry 49: 4945-4956.
947 79. Ichikawa Y, Bayeva M, Ghanefar M, Potini V, Sun L, et al. (2012) Disruption of ATP-binding cassette
948 B8 in mice leads to cardiomyopathy through a decrease in mitochondrial iron export. Proc
949 Natl Acad Sci U S A 109: 4152-4157.

80. Podsiadlowski L, Matha V, Vilcinskas A (1998) Detection of a P-glycoprotein related pump in Chironomus larvae and its inhibition by verapamil and cyclosporin A. Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology 121: 443-450.
81. Aurade RM, Jayalakshmi SK, Sreeramulu K (2010) P-glycoprotein ATPase from the resistant pest, Helicoverpa armigera: Purification, characterization and effect of various insecticides on its transport function. Biochimica Et Biophysica Acta-Biomembranes 1798: 1135-1143.
82. Aurade RM, Jayalakshmi SK, Udikeri SS, Sreeramulu K (2012) MODULATION OF P-GLYCOPROTEIN ATPASE OF HELICOVERPA ARMIGERA BY CHOLESTEROL: EFFECTS ON ATPASE ACTIVITY AND INTERACTION OF INSECTICIDES. Arch Insect Biochem Physiol 79: 47-60.
83. Luo L, Sun YJ, Wu YJ (2013) Abamectin resistance in Drosophila is related to increased expression of P-glycoprotein via the dEGFR and dAkt pathways. Insect Biochemistry and Molecular Biology 43: 627-634.
84. Porretta D, Gargani M, Bellini R, Medici A, Punelli F, et al. (2008) Defence mechanisms against insecticides temephos and diflubenzuron in the mosquito Aedes caspius: the P-glycoprotein efflux pumps. Med Vet Entomol 22: 48-54.
85. Srinivas R, Udikeri SS, Jayalakshmi SK, Sreeramulu K (2004) Identification of factors responsible for insecticide resistance in Helicoverpa armigera. Comparative Biochemistry and Physiology C-Toxicology & Pharmacology 137: 261-269.
86. Lanning CL, Fine RL, Corcoran JJ, Ayad HM, Rose RL, et al. (1996) Tobacco budworm P-glycoprotein: Biochemical characterization and its involvement in pesticide resistance. Biochimica Et Biophysica Acta-General Subjects 1291: 155-162.
87. Murray CL, Quaglia M, Arnason JT, Morris CE (1994) A PUTATIVE NICOTINE PUMP AT THE METABOLIC BLOOD-BRAIN-BARRIER OF THE TOBACCO HORNWORM. J Neurobiol 25: 23-34.
88. Sorensen JS, Dearing MD (2006) Efflux transporters as a novel herbivore countermechanism to plant chemical defenses. J Chem Ecol 32: 1181-1196.
89. Begun DJ, Whitley P (2000) Genetics of alpha-amanitin resistance in a natural population of Drosophila melanogaster. Heredity 85: 184-190.
90. Tapadia MG, Lakhotia SC (2005) Expression of mdr49 and mdr65 multidrug resistance genes in larval tissues of Drosophila melanogaster under normal and stress conditions. Cell Stress Chaperones 10: 7-11.
91. Bangert I, Tumulka F, Abele R (2011) The lysosomal polypeptide transporter TAPL: more than a housekeeping factor? Biol Chem 392: 61-66.
92. Zhou SF, Wang LL, Di YM, Xue CC, Duan W, et al. (2008) Substrates and inhibitors of human multidrug resistance associated proteins and the implications in drug development. Curr Med Chem 15: 1981-2039.
93. Chen Z-S, Tiwari AK (2011) Multidrug resistance proteins (MRPs/ABCCs) in cancer chemotherapy and genetic diseases. FEBS J 278: 3226-3245.
94. Aleksandrov AA, Aleksandrov LA, Riordan JR (2007) CFTR (ABCC7) is a hydrolyzable-ligand-gated channel. Pflugers Archiv-European Journal of Physiology 453: 693-702.
95. Hunt JF, Wang C, Ford RC (2013) Cystic Fibrosis Transmembrane Conductance Regulator (ABCC7) Structure. Cold Spring Harbor Perspectives in Medicine 3.
96. Bryan J, Munoz A, Zhang X, Dufer M, Drews G, et al. (2007) ABCC8 and ABCC9: ABC transporters that regulate K(+) channels. Pflugers Archiv-European Journal of Physiology 453: 703-718.
97. Heckel DG (2012) Learning the ABCs of Bt: ABC transporters and insect resistance to Bacillus thuringiensis provide clues to a crucial step in toxin mode of action. Pestic Biochem Physiol 104: 103-110.
98. Morita M, Imanaka T (2012) Peroxisomal ABC transporters: Structure, function and role in disease. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease 1822: 1387-1396.

998 99. Kashiwayama Y, Seki M, Yasui A, Murasaki Y, Morita M, et al. (2009) 70-kDa peroxisomal
999 membrane protein related protein (P70R/ABCD4) localizes to endoplasmic reticulum not
1000 peroxisomes, and NH2-terminal hydrophobic property determines the subcellular localization
1001 of ABC subfamily D proteins. Exp Cell Res 315: 190-205.
1002 100. Tian Y, Han X, Tian D-l (2012) The biological regulation of ABCE1. IUBMB Life 64: 795-800.
1003 101. Andersen DS, Leevers SJ (2007) The essential Drosophila ATP-binding cassette domain protein,
1004 pixie, binds the 40 S ribosome in an ATP-dependent manner and is required for translation
1005 initiation. J Biol Chem 282: 14752-14760.
1006 102. Kerr ID (2004) Sequence analysis of twin ATP binding cassette proteins involved in translational
1007 control, antibiotic resistance, and ribonuclease L inhibition. Biochem Biophys Res Commun
1008 315: 166-173.
1009 103. Kerr ID, Haider AJ, Gelissen IC (2011) The ABCG family of membrane-associated transporters: you
1010 don't have to be big to be mighty. Br J Pharmacol 164: 1767-1779.
1011 104. Moitra K, Silverton L, Limpert K, Im K, Dean M (2011) Moving out: from sterol transport to drug
1012 resistance - the ABCG subfamily of efflux pumps. Drug Metabolism and Drug Interactions 26:
1013 105-111.
1014 105. Tarr PT, Tarling EJ, Bojanic DD, Edwards PA, Baldan A (2009) Emerging new paradigms for ABCG
1015 transporters. Biochimica Et Biophysica Acta-Molecular and Cell Biology of Lipids 1791: 584-
1016 593.
1017 106. Woodward OM, Koettgen A, Koettgen M (2011) ABCG transporters and disease. FEBS J 278: 3215-
1018 3225.
1019 107. Wang L, Kiuchi T, Fujii T, Daimon T, Li M, et al. (2013) Mutation of a novel ABC transporter gene is
1020 responsible for the failure to incorporate uric acid in the epidermis of ok mutants of the
1021 silkworm, Bombyx mori. Insect Biochemistry and Molecular Biology 43: 562-571.
1022 108. Itoh TQ, Tanimura T, Matsumoto A (2011) Membrane-bound transporter controls the circadian
1023 transcription of clock genes in Drosophila. Genes Cells 16: 1159-1167.
1024 109. Sturm A, Cunningham P, Dean M (2009) The ABC transporter gene family of Daphnia pulex. BMC
1025 Genomics 10: 170.
1026 110. Dong JS, Lai R, Jennings JL, Link AJ, Hinnebusch AG (2005) The novel ATP-binding cassette protein
1027 ARB1 is a shuttling factor that stimulates 40S and 60S ribosome biogenesis. Mol Cell Biol 25:
1028 9859-9873.
1029 111. Overbeck TR, Hupfeld T, Krause D, Waldmann-Beushausen R, Chapuy B, et al. (2013) Intracellular
1030 ATP-Binding Cassette Transporter A3 is Expressed in Lung Cancer Cells and Modulates
1031 Susceptibility to Cisplatin and Paclitaxel. Oncology 84: 362-370.
1032 112. Aung T, Chapuy B, Vogel D, Wenzel D, Oppermann M, et al. (2011) Exosomal evasion of humoral
1033 immunotherapy in aggressive B-cell lymphoma modulated by ATP-binding cassette
1034 transporter A3. Proc Natl Acad Sci U S A 108: 15336-15341.
1035 113. Fischer S, Kluver N, Burkhardt-Medicke K, Pietsch M, Schmidt AM, et al. (2013) Abcb4 acts as
1036 multixenobiotic transporter and active barrier against chemical uptake in zebrafish (Danio
1037 rerio) embryos. BMC Biol 11.
1038 114. Wang Y, Lim L, DiGuistini S, Robertson G, Bohlmann J, et al. (2013) A specialized ABC efflux
1039 transporter GcABC-G1 confers monoterpene resistance to Grosmannia clavigera, a bark
1040 beetle-associated fungal pathogen of pine trees. New Phytol 197: 886-898.
1041 115. Dow JAT, Davies SA (2006) The Malpighian tubule: Rapid insights from post-genomic biology. J
1042 Insect Physiol 52: 365-378.
1043 116. Bayeva M, Khechaduri A, Wu RX, Burke MA, Wasserstrom JA, et al. (2013) ATP-Binding Cassette
1044 B10 Regulates Early Steps of Heme Synthesis. Circul Res 113: 279-287.

1045 **117. Chloupkova M, LeBard LS, Koeller DM (2003) MDL1 is a high copy suppressor of ATM1: Evidence**
1046 **for a role in resistance to oxidative stress. J Mol Biol 331: 155-165.**
1047 **118. Sullivan DT, Sullivan MC (1975) TRANSPORT DEFECTS AS PHYSIOLOGICAL BASIS FOR EYE COLOR**
1048 **MUTANTS OF DROSOPHILA-MELANOGASTER. Biochem Genet 13: 603-613.**
1049 **119. Ryall RL, Howells AJ (1974) OMMOCHROME BIOSYNTHETIC-PATHWAY OF DROSOPHILA-**
1050 **MELANOGASTER - VARIATIONS IN LEVELS OF ENZYME-ACTIVITIES AND INTERMEDIATES**
1051 **DURING ADULT DEVELOPMENT. Insect Biochemistry 4: 47-61.**
1052 **120. Sullivan DT, Grillo SL, Kitos RJ (1974) SUBCELLULAR-LOCALIZATION OF FIRST 3 ENZYMES OF**
1053 **OMMOCHROME SYNTHETIC PATHWAY IN DROSOPHILA-MELANOGASTER. Journal of**
1054 **Experimental Zoology 188: 225-233.**
1055 **121. Evans JM, Day JP, Cabrero P, Dow JAT, Davies S-A (2008) A new role for a classical gene: White**
1056 **transports cyclic GMP. J Exp Biol 211: 890-899.**
1057 **122. Haunerland NH, Shirk PD (1995) REGIONAL AND FUNCTIONAL-DIFFERENTIATION IN THE INSECT**
1058 **FAT-BODY. Annual Review of Entomology 40: 121-145.**
1059 **123. Arrese EL, Soulages JL (2010) Insect Fat Body: Energy, Metabolism, and Regulation. Annual**
1060 **Review of Entomology. pp. 207-225.**
1061 **124. Tarling EJ, Vallim TQD, Edwards PA (2013) Role of ABC transporters in lipid transport and human**
1062 **disease. Trends Endocrinol Metab 24: 342-350.**
1063 **125. Li JW, Lehmann S, Weissbecker B, Naharros IO, Schutz S, et al. (2013) Odoriferous Defensive Stink**
1064 **Gland Transcriptome to Identify Novel Genes Necessary for Quinone Synthesis in the Red**
1065 **Flour Beetle, Tribolium castaneum. PLoS Genet 9.**
1066 **126. Elliott AM, Al-Hajj MA (2009) ABCB8 Mediates Doxorubicin Resistance in Melanoma Cells by**
1067 **Protecting the Mitochondrial Genome. Mol Cancer Res 7: 79-87.**
1068 **127. Dermauw W, Osborne EJ, Clark RM, Grbic M, Tirry L, et al. (2013) A burst of ABC genes in the**
1069 **genome of the polyphagous spider mite Tetranychus urticae. BMC Genomics 14.**
1070
1071

3.4 *MANUSCRIPT 4: INDEPENDENT RECRUITMENTS OF OXIDASES FROM THE GMC OXIDOREDUCTASE FAMILY ENABLED THE EVOLUTION OF CHEMICAL DEFENCE IN LEAF BEETLE LARVAE*

Independent recruitments of oxidases from the GMC oxidoreductase family enabled the evolution of chemical defence in leaf beetle larvae

Submitted to Proceedings of the Royal Society B: For Review Only

Independent recruitments of oxidases from the GMC oxidoreductase family enabled the evolution of chemical defence in leaf beetle larvae

Journal:	*Proceedings B*
Manuscript ID:	Draft
Article Type:	Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Rahfeld, Peter; Max Planck Institute for Chemical Ecology, Bioorganic Chemistry Kirsch, Roy; Max Planck Institute for Chemical Ecology, Entomology Kugel, Susann; Max Planck Institute for Chemical Ecology, Bioorganic Chemistry Wielsch, Natalie; Max Planck Institute for Chemical Ecology, Mass Spectrometry/Proteomics Stock, Magdalena; Max Planck Institute for Chemical Ecology, Bioorganic Chemistry Groth, Marco; Leibniz Institute for Age Research – Fritz Lipmann Institute, Genome Analysis Boland, Wilhelm; Max Planck Institute for Chemical Ecology, Bioorganic Chemistry Burse, Antje; Max Planck Institute for Chemical Ecology, Bioorganic Chemistry
Subject:	Biochemistry < BIOLOGY, Molecular Biology < BIOLOGY, Evolution < BIOLOGY
Keywords:	secretions, phylogeny, iridoid biosynthesis, RNA interference, chemical defence, insects
Proceedings B category:	Evolutionary Biology

SCHOLARONE™
Manuscripts

Independent recruitments of oxidases from the GMC oxidoreductase family enabled the evolution of chemical defence in leaf beetle larvae (subtribe Chrysomelina)

Peter Rahfeld[1], Roy Kirsch[2], Susann Kugel[1], Natalie Wielsch[3], Magdalena Stock[1], Marco Groth[4], Wilhelm Boland[1], Antje Burse[1*]

[1] Department of Bioorganic Chemistry, Max Planck Institute for Chemical Ecology, Jena, Germany

[2] Department of Entomology, Max Planck Institute for Chemical Ecology, Jena, Germany

[3] Research Group Mass Spectrometry/Proteomics, Max Planck Institute for Chemical Ecology, Jena, Germany

[4] Genome Analysis Group, Leibniz Institute for Age Research – Fritz Lipmann Institute, Jena, Germany

[*] Corresponding author. Phone: ++49(0)3641-571265, Fax: ++49(0)3641-571202, E-mail: aburse@ice.mpg.de

ABSTRACT

Juveniles of the leaf beetle subtribe Chrysomelina repel their enemies by displaying glandular secretions that contain defensive compounds. These repellents can be produced either *de novo* (iridoids) or by using plant-derived precursors (e.g. salicylaldehyde). The autonomous production of iridoids, as in *Phaedon cochleariae*, is the ancestral chrysomeline chemical defence and pre-dates the evolution of salicylaldehyde-based defence. Both biosynthesis strategies include an oxidative step of an alcohol intermediate. In salicylaldehyde-producing species, this step is catalysed by salicyl alcohol oxidases (SAOs) of the glucose-methanol-choline (GMC) oxidoreductase superfamily, but the enzyme oxidizing the iridoid precursor is unknown. Here we show by *in vitro* as well as *in vivo* experiments that *P. cochleariae* also uses an oxidase recruited from the GMC superfamily for defensive purposes. However, our phylogenetic analysis of chrysomeline GMCs revealed that the oxidase of the iridoid pathway originated from a GMC clade other than that of the SAOs. Thus, the evolution of a host-independent chemical defence followed by a shift to a host-dependent chemical defence in Chrysomelina beetles coincided with the independent recruitment of genes from different GMC subfamilies. These findings illustrate the importance of the GMC multi-gene family for adaptive processes in plant-insect interactions.

INTRODUCTION

Beetles (order Coleoptera) make up the largest order of animals with approximately 350,000 species and 40% of all insects [1]. Among them, leaf beetles (family Chrysomelidae) represent one of the most successful radiations; their success can be attributed to a long adaptive evolutionary history with plants [2, 3]. Several thousand species are external leaf chewers; due to their life on the surface of plants, they are permanently exposed to life-threatening predators and parasitoids. The ability to produce chemical defences is one of the most fascinating adaptations that have evolved in the leaf beetles of the subtribe Chrysomelina. These defences ensure that all developmental stages, from egg to adult, are protected. When disturbed, their larvae display droplets of defensive secretions on their backs by everting nine pairs of glandular reservoirs located under their dorsal cuticle [4]. The defensive droplets contain chemically diverse deterrents [5-10]. Phylogenetic analyses of chrysomeline species revealed that the composition of their secretions reflects a step-wise scenario of host-plant adaptation [11]. The evolutionary history of chrysomeline larval chemical defence started with the *de novo* production of deterrent iridoids (cyclopentanoic monoterpenoids) that does not rely on the secondary metabolites of their hosts [12, 13]. Derived from this autonomous biosynthesis, two lineages independently developed a defensive strategy that relies on the sequestration of salicin (Figure 1), a plant-derived precursor from Salicaceae used to produce the deterrent salicylaldehyde [11]. Although this sequestration narrows the host range, the caloric benefit may trade off the tight dependency on the plant chemistry [14].

Despite the different composition and origin of the defensive compounds in the secretions of chrysomeline larvae (*de novo* versus sequestration), the synthesis of these defensive compounds depends on common enzymatic steps. For example, the hydrolysis of the glucosidically bound precursors that are transported into the glandular reservoir is facilitated by beta-glucosidase activities [15, 16] (Figure 1). Subsequently, the released

alcohol is oxidised in iridoid- and salicylaldehyde-producing species [17, 18]. The common consecutive activity of beta-glucosidase and oxidase suggested that it was the acquisition of only a few amino acid substitutions in the ancestral enzymes of the iridoid- *de novo* producing species that enabled the sequestration-based salicylaldehyde biosynthesis [4].

However, among the predicted enzymes in the chrysomeline secretions, only the functionally characterised salicyl alcohol oxidases (SAOs) from salicin-sequestering species have been analysed with respect to their ancestry [18-20]. These SAOs belong to the glucose-methanol-choline (GMC) oxidoreductase multi-gene family [21, 22] and convert salicylalcohol into salicylaldehyde [18]. Comparative genome analyses show that the GMC oxidoreductase family harbours genes, most of which are in a cluster and unique with respect to their expansion in insects [23-25]. Phylogenetic analyses of the functional SAOs from salicin-sequestering Chrysomelina strongly support a common ancestry in the GMC$_{\iota}$ clade [20] irrespective of the evolutionary affiliation of the corresponding beetle species [11].

However, the evolutionary correlation between the SAO and the predicted oxidase of the iridoid pathway has not yet been resolved, as the latter enzyme has not been identified. Early biochemical investigations of secretions in iridoid- (*Phratora amoraciae, Phratora laticollis*) and salicylaldehyde- (*Phratora vitellinae*) producing species revealed side activities for the non-natural substrates salicyl alcohol and 8-hydroxygeraniol, respectively [26]. This result suggests that there has been a specific change of the oxidative enzyme according to the evolution of chrysomeline beetles [11, 18]. However, the recombinant SAO from the salicylaldehyde- (*Ph. vitellinae*) utilising species lacks the 8-hydroxygeraniol oxidase (8HGO) activity, suggesting another enzyme is responsible for this cross-reactivity in the secretions [20]. Moreover, proteome analyses of secretions from iridoid-producing larvae did not reveal an oxidase related to the GMC$_{\iota}$ clade [20]. Hence, these recent results point to an acquisition of SAOs in salicylaldehyde-producing larvae that is independent of the 8-hydroxygeraniol-converting oxidase in the ancestral species.

In the present paper, we identified a glandular-specific oxidase from the iridoid-producing larvae of the mustard leaf beetle, *Phaedon cochleariae*. Functional characterisation including the substrate specificity of this enzyme after heterologous expression revealed the selective oxidation of 8-hydroxygeraniol to 8-oxogeranial. The importance of this 8HGO activity for the formation of iridoid has been further verified by RNA interference (RNAi) *in vivo*. Phylogenetic analyses demonstrated that 8HGO had been recruited from a beetle-specific GMC clade and not from the GMC*1* clade, from which SAOs arose. These findings elucidate a key event in the evolution of glandular chemical defence in chrysomeline beetles. Moreover, they provide insights into adaptive mechanisms that enabled the transition from *de novo* biosynthesis to sequestration and, thus, the underlying evolutionary dynamics of host-plant affiliation.

MATERIAL & METHODS

See electronic supplementary material (ESM) for the complete proteome analyses of the secretions by data-independent liquid chromatography/mass spectrometry detection (LC/MS^E), cloning procedures, detailed quantitative PCR (qPCR) procedure, phylogenetic analysis, gas chromatography–mass spectrometry (GC-MS) analysis, all primer sequences (Table S1) and accession numbers (Table S3).

(a) The silencing of the *P. cochleariae* 8-hydroxygeraniol oxidase (*Pc*8HGO) and the *P. cochleariae* 8-hydroxygeraniol oxidase-like protein (*Pc*8HGO-like) by RNAi

The coding sequences of *Pc8HGO* and *Pc8HGO-like* were analysed for off-target prediction according to Bodemann and Rahfeld *et al.* [27]. This analysis revealed that *Pc8HGO* and *Pc8HGO-like* have a contiguous 26 bp fragment in common which is interrupted by only one dissimilar base at position 11 and which may be sufficient to trigger off-target effects (Figure

129 S4). Furthermore, no putative off-target effects with other transcripts were predicted with the
130 chosen dsRNA sequences for a critical value of at least 21 continuous nucleotides.
131 For the dsRNA-constructs, 200-bp-fragments (Table S1) from the coding sequences of
132 *Pc8HGO* and *Pc8HGO-like* were amplified by a Phusion High-Fidelity DNA Polymerase
133 (Fisher Scientific - Germany GmbH, Schwerte, Germany). After purification with a PCR-
134 purification kit (Roche, Basel, Switzerland), the resulting fragments were cloned into T7-
135 promotor site free pIB/V5-HIS-TOPO vectors (Life Technologies, Carlsbad, USA). For
136 dsRNA synthesis, templates with opposite T7-promotor sites were amplified out of sequenced
137 pIB-*200bpPc8HGO* as well as pIB-*200bpPc8HGO-like* and further processed as described in
138 Bodemann and Rahfeld *et al.* [27]. The concentration of dsRNA was adjusted to 1 µg/µl.
139 Early second-instar larvae of *P. cochleariae* were used for injections. They were collected 7 d
140 after hatching and treated with 100 nl (100 ng) dsRNA of *200bpPc8HGO* or *200bpPc8HGO-*
141 *like*. The dsRNA of *720bpeGFP* was used as described in Bodemann and Rahfeld *et al.* [27]
142 for control treatments.
143
144 **(b) Heterologous expression of *Pc*8HGO in HighFive insect cells and protein purification**
145 Heterologous expression was carried out in the insect cell line HighFive (Life Technologies,
146 Carlsbad, USA). The construct pIB-*Pc8HGO* was transfected with the FuGeneHD–Kit
147 (Promega GmbH, Fitchburg, USA) and MA Lipofection Enhancer (IBA GmbH, Göttingen
148 Germany) according to the manufacturer's instructions. After one day of incubation at 27°C,
149 the culture was supplied with 80 µg/ml blasticidin (Life Technologies, Carlsbad, USA) to
150 initiate the selection of stable cell lines. The insect cells were selected over three passages.
151 The cultivation of the stable cell lines for protein expression was carried out in six 75 cm^2 cell
152 culture flasks with each 15 ml culture media (expressfive™ (Life Technologies, Carlsbad,
153 USA), 20 µg/ml blasticidin, 1xProtease Inhibitor HP Mix (SERVA Electrophoresis GmbH,
154 Heidelberg, Germany)). After 3 days of growth, the supernatant was collected and the cells

were discarded (4000xg, 10 min, 4°C). The supernatant containing *Pc*8HGO was dialysed overnight at 4°C against 50 mM NaH_2PO_4, 10 mM imidazol (Pufferan), 5% (v/v) glycerol, pH 7.5.

The subsequent purification was done with HisPureCobalt (Life Technologies, Carlsbad, USA) according to the manufacturer's instructions with alterations to the elution buffer of 50 mM NaH_2PO_4, 150 mM imidazol (Pufferan), 5% (v/v) glycerol, pH 7.5. To confirm the identity of the purified protein, it was separated by any-kD gradient gels (Bio-Rad Laboratories, Munich, Germany) in 1D-SDS-PAGE and then analysed *via* Nano-UPLC-MS^E as described in the ESM.

(c) *Pc*8HGO activity assay

The purified proteins were dialysed overnight at 4°C against an assay buffer consisting of 50 mM NaH_2PO_4, pH 4.5 to support the protein with the proper pH-value. To confirm the catalytic activity 10 µl of purified protein, 10 µl 50 mM 8-hydroxygeraniol (end concentration 5 mM) or 10 µl 50 mM salicyl alcohol (end concentration 5 mM) and 80 µl assay buffer were incubated for 0, 30, 60 min at 30°C.

RESULTS

See ESM for results of the *de novo* assembly of *C. populi*'s transcriptome.

(a) Identification and sequence analysis of GMC oxidoreductases from the defensive secretions of *P. cochleariae*

Eleven protein bands were recovered after the separation of *P. cochleariae* larval secretions by 1D-SDS-PAGE (Figure S1). The resulting LC/MS^E data were searched against a *P. cochleariae* protein library derived from a *P. cochleariae* transcriptome as described in Stock *et al.* [28]. The analysis of the band of about 70 kDa revealed two proteins (*Pc*8HGO

and *Pc*8HGO-like) showing similarity to the GMC oxidoreductase family (GMC_oxred_N (PF00732)). Full-length amplification and sequencing of the corresponding transcripts led to coding sequences of 1672 bp (623 amino acids (aa)) and 1669 bp (622 aa) for *Pc*8HGO and *Pc*8HGO-like, respectively, with 77% sequence identity to each other on the amino acid level. Despite the sequence similarity, both proteins were unambiguously identified from the secretions as LC-MS^E-derived peptides matching *Pc*8HGO or *Pc*8HGO-like (Table S2). N-terminal signal peptides with a length of 16 aa (*Pc*8HGO) and 22 aa (*Pc*8HGO-like) were indicated by cleavage site predictions (SignalP 4.1: http://www.cbs.dtu.dk/services/SignalP/) [29].

Previous studies of larval secretions of salicin-sequestering Chrysomelina identified SAO proteins, oxidizing salicyl alcohol to the respective aldehyde, as members of the same GMC clade and of a similar molecular weight (69 kDa) [18]. But *Pc*8HGO and *Pc*8HGO-like show only a low degree of sequence identity, about 36% on the amino acid level, to these SAOs. Despite low sequence similarity, their protein alignment, including GMC oxidoreductases such as SAOs of the closely related Chrysomelina *C. populi* and *C. lapponica* as well as the aryl alcohol oxidase (AAO) from the bacterium *Arthrobacter globiformis* and the glucose oxidase (GOX) from the fungus *Aspergillus niger,* illustrates, at least a few conserved regions (Figure S2). Like other GMC oxidoreductase proteins, *Pc*8HGO and *Pc*8HGO-like possess the N-terminal beta-alpha-beta dinucleotide-binding motif ($GxGxxG(x)_{18}E$) necessary to bind the flavin adenine dinucleotide (FAD) co-factor [21, 30, 31]. Additionally, there are several blocks of conserved amino acid sequences common among GMC oxidoreductases (Figure S2) [23].

(b) Transcript localisation of *Pc8HGO* and *Pc8HGO-like*

We compared the *Pc8HGO* and *Pc8HGO-like* expression levels in different larval tissues by qPCR. Both genes are specifically expressed in the defensive glands with an at least ~130 fold

(*Pc8HGO*) and ~240 fold (*Pc8HGO-like*) higher transcript abundance compared to gut, Malpighian tubules, fat body or head (Figure S3). The qPCR products were cloned and sequenced to confirm their identity. Their specific expression in the glandular tissue is in accordance with the identification of the respective proteins in the glandular secretions revealing that both proteins possess gland-specific functions after being secreted into the corresponding reservoir.

(c) Functional importance of glandular GMC oxidoreductases identified by RNAi

RNAi was used to analyse the potential impact of *Pc*8HGO and *Pc*8HGO-like in the biosynthesis of the defensive iridoid chrysomelidial in the larval glandular secretions *in vivo*. The down-regulation of the corresponding transcripts in the glandular tissue was surveyed by qPCR (Figure S5). Comparing the treatment control *eGFP* and the non-injection-control (NIC), we found no significant difference in the transcript abundance of *Pc8HGO* ($p = 0.86$) and *Pc8HGO-like* ($p = 0.74$). In contrast, seven days after injecting the dsRNA of *Pc8HGO* or *Pc8HGO-like,* the down-regulation by ~98% and ~96% of the respective transcripts compared to the *eGFP* treatment control was detected. In addition, although off-target prediction has been taken into account while designing the dsRNA fragments for RNAi [27], non-targeted transcripts were silenced. The down-regulation of *Pc8HGO* led at the same time to a ~85% mRNA reduction of the non-targeted *Pc8HGO-like* and vice versa (Figure S5). This effect can be traced back to the high nucleotide sequence similarity of the targeted transcripts *Pc8HGO* and *Pc8HGO-like* of 83% that complicated the design of specific dsRNA constructs (Figure S4). Nonetheless, in both cases, the down-regulation of the targeted transcript was significantly more effective compared to non-targeted transcript (Table S4).

We collected glandular secretions for GC-MS analyses after silencing *Pc8HGO* or *Pc8HGO-like* to identify potential changes in the secretions' terpenoid composition. As observed in the *eGFP* larvae (Figure S6), chrysomelidial, the final product of the iridoid

pathway, accumulated in the secretions when *Pc8HGO-like* was knocked-down (Figure 2a). In contrast, in the secretions of the larvae injected with dsRNA-targeting *Pc8HGO*, chrysomelidial was no longer detectable (Figure 2b). Moreover, another substance accumulated in the secretions that could be identified as the chrysomelidial precursor 8-hydroxygeraniol [12]. Taken together, the RNAi experiments verified the importance of the *Pc*8HGO protein in the iridoid biosynthesis occurring in the glandular system of *P. cochleariae* larvae. The accumulation of 8-hydroxygeraniol indicates that this precursor is a substrate of the *Pc*8HGO enzyme, which, in turn, catalyses the oxidation to the chrysomelidial biosynthesis intermediate 8-oxogeranial (Figure 1). *Pc*8HGO-like was rejected as a potential 8-hydroxygeraniol oxidase as the glandular secretion of silenced larvae did not contain 8-hydroxygeraniol. The significance of *Pc*8HGO in the glandular context is additionally supported by a loss of the yellow colour of the secretions (Figure 2a) that cannot be observed in any other treatment. The silencing seems to have no effect if the targeted transcript is reduced to >10%.

(d) Catalytic activity of the purified *Pc*8HGO

To validate the results obtained from the RNAi experiments and to test for the oxidative capacity, *Pc*8HGO was heterologously expressed. Therefore, the plasmid pIB-*Pc8HGO* carrying the coding sequence (cds) of *Pc*8HGO was transfected into insect cells followed by a purification of the resulting protein. *Pc*8HGO was successfully expressed and purified with only a few impurities (Figure S7). The identity of the protein was certified through LC-MS^E analysis (Table S2).

The purified *Pc*8HGO was used for activity assays with 8-hydroxygeraniol as a substrate (Figure 3). The reaction was stopped after 0, 30 and 60 min, and GC-MS analyses revealed that *Pc*8HGO is able to metabolise 8-hydroxygeraniol, as the corresponding peak (retention time 10.9 min) disappeared over time. Whereas in the beginning only the substrate

was present, three new peaks were detectable after 30 min. Using a standard compound, one of the peaks with a retention time of 14 min could be identified as 8-oxogeranial. The other substances eluting at 12.1 and 12.7 min are probably the semi-aldehydes 8-hydroxygeranial and 8-oxogeraniol as described in previous studies of the oxidative capacity in *P. cochleariae* secretions [17]. After 60 min, nearly all of the substrate and intermediate peaks were oxidised to 8-oxogeranial. These assays coincide with the phenotype observed after *Pc8HGO* was silenced and the *Pc*8HGO enzyme was verified to be the oxidase in the glandular secretion of iridoid-producing *P. cochleariae* larvae converting 8-hydroxygeraniol to the respective aldehyde 8-oxogeranial.

In addition, the substrate specificity of *Pc*8HGO was tested by incubating the oxidase with salicyl alcohol, the substrate of chrysomeline SAOs. No enzyme-based conversion to salicylaldehyde could be detected (Figure S8), indicating this particular enzyme does not react with salicyl alcohol.

(e) Evolution of glandular oxidases in Chrysomelina

To uncover the evolutionary origin of the 8-hydroxygeraniol oxidase *Pc*8HGO from *P. cochleariae* and to test for a dependent versus independent origin of the SAOs already known from *Chrysomela* species, among others, we combined GMCs from both species in a phylogenetic analysis. BLAST searches against *P. cochleariae* and *C. populi* transcriptome libraries revealed 10 and six full-length coding sequences, respectively, each showing high sequence similarity to the query sequences *Pc8HGO* and *Pc8HGO-like*. Phylogenetic analyses, including those sequences, members of different insect GMC oxidoreductase subfamilies as well as chrysomeline SAOs and related sequences from previous work [19, 20], showed that *Pc8HGO* and chrysomeline SAOs had independent origins (Figure 4). As it has been shown earlier, SAOs and related sequences cluster in an insect GMC*ι* clade closest to *Tribolium castaneum* GMC*ι*5 [19, 20]. In contrast, the *Pc8HGO* is affiliated with GMC

285 oxidoreductases from *T. castaneum* (XM961538, XM961446, XM967481); according to a
286 global insect GMC analysis [23], these cluster separately from their GMC*i* counterparts in the
287 so-called beetle GMC clade. The origin of *Pc8HGO* within the beetle GMC clade is indicated
288 by a close relationship to the *T. castaneum* GMC (XM967481) and is supported by high
289 posterior probability and bootstrap values (1, 97, 92). The finding that *Pc8HGO* clusters with
290 three other *P. cochleariae* GMCs (including *Pc*8HGO-like) but just with a single *C. populi*
291 (*Cp*GMC*bl*6) and *T. castaneum* GMC (*Tcas*GMCXM967481) probably reflects gene
292 duplications restricted to the chrysomeline iridoid-producing lineage. However, the high
293 number of beetle GMC clade genes in *P. cochleariae* and *C. populi* in general and the
294 presence of four strict orthologs among those indicates that the chrysomeline ancestor already
295 possessed a diverse set of these genes. The same most likely holds true for the GMC*i* clade, as
296 we found three genes of *P. cochleariae* (*Pc*SAO-like 1 to 3) clustering with *Chrysomela spp.*
297 SAO counterparts and the single homolog of *T. castaneum* GMC*i*5. Concluding, our
298 phylogenetic analysis supports the hypothesis that 8HGO and SAO arose from two clades of
299 GMCs and started to diversify early in chrysomeline evolution.
300
301 **DISCUSSION**
302 Oxidation-reduction reactions are the most prevalent and fundamental reactions in the
303 metabolism of all organisms. In the defensive secretions of Chrysomelina larvae, these
304 reactions are implicated in the production of deterrent compounds. Enzymes that catalyse
305 such reactions often belong to the GMC oxidoreductase family. Here we have identified GMC
306 oxidoreductases (*Pc*8HGO and *Pc*8HGO-like) in the secretions of the juvenile *P. cochleariae.*
307 They share common features of GMC oxidoreductases; for example, in both the alpha-beta-
308 alpha ATP binding site is responsible for the covalent binding of FAD [21]. *Pc8HGO* and
309 *Pc8HGO-like* transcripts are abundant only in the cells of the glandular tissue and therefore
310 clearly intended for a role in defence machinery (Figure S3). The deduced protein sequences

311 of *Pc*8HGO and *Pc*8HGO-like carry signal peptides targeting the proteins to the secretory
312 pathway (Figure S2). RNAi experiments *in vivo* confirmed that *Pc*8HGO is an important
313 enzyme for iridoid production in the secretions, whereas the function for *Pc*8HGO-like
314 remains unclear, but its involvement in iridoid metabolism can be excluded (Figure 2).

315 To validate the function *in vitro*, *Pc*8HGO was heterologously expressed in insect cell
316 lines. The purified recombinant *Pc*8HGO has the catalytic ability to oxidise 8-
317 hydroxygeraniol to 8-oxogeranial without requiring the addition of $NADP^+$ (Figure 3),
318 because GMC oxidoreductases are known to use molecular oxygen as an electron acceptor
319 [22]. To sum up our results, *Pc*8HGO is an extracellular enzyme which converts 8-
320 hydroxygeraniol to the corresponding dialdehyde in the secretions of *P. cochleariae*.

321 By identifying a GMC oxidoreductase involved in the defensive metabolism from a *de*
322 *novo* producing species, we gain access to phylogenetic analyses that allow us to disentangle
323 the ancestry of SAOs and the origin of glandular oxidases in chrysomelines in general.
324 Although they are members of the same gene family, oxidases of the salicylaldehyde and
325 iridoid biosynthetic pathways have been recruited -- one from the GMC*ι* and one from the
326 beetle GMC clade, respectively -- during chrysomeline evolution. The shift to a
327 salicylaldehyde-based defence and also most likely the shift to salicin-containing host plants
328 has been made possible through the recruitment of a new glandular oxidase instead of
329 "recycling" of an old one. The evolutionary steps towards 8HGO and SAO activity remain
330 unknown as e.g. the function of the respective *T. castaneum* counterparts has not been
331 characterized. But the high copy number of GMC*ι* and beetle GMCs in *P. cochleariae* and
332 *Chrysomela spp.* indicate that gene duplication played a major role in the evolution of both
333 8HGO and SAO.

334 Despite comprehensive GMC gene analyses, phylogenetic relations of both clades are
335 not yet completely resolved [23, 24]. But irrespective of whether GMC*ι* and beetle GMCs
336 cluster separately in two subfamilies [23] or had an intertwined evolutionary history [24], the

337 corresponding *T. castaneum* GMCs are not close relatives. Thus, both analyses support our
338 findings of an independent origin of chrysomeline SAO and 8HGO similarly to their closely
339 related *T. castaneum* counterparts (*Tcas*GMC*t*5 and *Tcas*GMCXM967481 also do not
340 cluster).

341 Hence, these results suggest that it is a biochemical property of the *P. vitellinae*
342 secretions to oxidise salicyl alcohols as well as 8-hydroxygeraniols [26], which is not
343 attributable to one single oxidase [20]. This leads to the conclusion that *P. vitellinae* imply
344 another GMC oxidoreductase, probably from the beetle clade, a potential evolutionary relict
345 of their host plant/deterrent shift.

346 How widespread the recruitment of oxidases in other insects for iridoid biosynthesis is
347 remains to be elucidated. But the same secondary metabolites can also be identified in insects
348 that frequently use iridoids as chemical stimuli for communication or defence [32, 33]. One of
349 the first insects discovered to contain iridoids is the eponym ant *Iridomyrmex* spp. [34]. The
350 phasmid *Graeffea crouani* (coconut stick insect) and the pseudophasmid *Anisomorpha*
351 *buprestoides* (Southern walking-stick) also use *de novo* produced iridodials and
352 nepetalactones [35, 36]. In the stick insects, the biosynthetic origin of the compounds pointed
353 to acetate, mevalonate and mevalonate lactone [37], the precursors of terpene biosynthesis.
354 The same biosynthetic origin could be detected in the leaf beetle *Phaedon armoraciae* for
355 chrysomelidial [13]. Other beetles known to produce iridoids are the *Chloridolum*
356 *loochooanum* (long-horn beetle) [38] and the carnivorous feeding *Philonthus spp.* (rove
357 beetles) [39]. The oxidation of 8-hydroxygeraniol in the rove beetle is accompanied by the
358 appearance of hydrogen peroxide (H_2O_2) [39], which is a characteristic for GMC
359 oxidoreductases [40]. The ability to biosynthesise iridoids seems to have evolved
360 independently in different insect families. In contrast, the production of iridoid precursors
361 depends on the mevalonate pathway conserved in all insects e.g. for the production of juvenile
362 hormones. Any gene from the GMC gene pool may be recruited to accomplish 8-

363 hydroxygeraniol oxidations independently in the insect families. However, more sequences
364 need to be available to untangle the ancestry of oxidases implicated in insect iridoid
365 biosynthesis.

366 The highest variety of iridoids is encountered in the plant kingdom [41-43]. One
367 example is *Catharanthus roseus*. Here the iridoids are precursors for secologanin, which is
368 then processed into clinically important alkaloids such as vinblastine or vincristine [44].
369 Interestingly, plants use a completely different enzyme family to oxidise 8-hydroxygeraniol.
370 In plants, a P450 enzyme (CYP76B6) [45] works as a multi-functional geraniol-8-oxidase
371 which first oxidises the geraniol to 8-hydroxygeraniol and, later, to 8-oxogeranial. The
372 identification of a new protein family able to produce also intermediates of the iridoid
373 biosynthesis opens the possibility of using *Pc*8HGO as an additional tool for heterologously
374 engineering plants [46].

375 In addition to the GMC cluster conserved in known insect genomes, which has been
376 discussed as functional in the ecdysone metabolism [23, 47], some GMC genes exist outside
377 of this cluster and have frequently experienced large lineage-specific expansion [23, 24]. It
378 has been suggested that these expansions of gene families may be correlated with the
379 adaptation to different environmental issues or specific life strategies [48]. Because insects
380 have evolved to occupy a vast diversity of habitats on earth, it can be hypothesised that these
381 GMCs have expanded further in insects in order to adapt to different environmental
382 conditions. The adaptation of the immune response in insects is an important fitness factor. In
383 silkworms, the knockdown of several GMC oxidoreductases, for example, led to a reduction
384 of survival rates after treatment with *Bacillus bombyseptieus* or *Escherichia coli* [24]. This
385 effect can be explained by the by-products arising during the oxidation reaction: GMC
386 oxidoreductases produce H_2O_2 [40]. The action of H_2O_2 as a messenger or toxin in the
387 immune response to microbial infections is already well described [49, 50]. Interestingly, the
388 secretions of juvenile Chrysomelina also have an antimicrobial effect, which is not entirely

389 due to the defensive compounds [51, 52]. This effect may also be related to the other
390 extracellular GMC oxidoreductase, *Pc*8HGO-like. Although silencing *Pc8HGO-like* did not
391 affect the phenotype with respect to the composition of deterrents, it may be that *Pc*8HGO-
392 like is involved in the antimicrobial activity of the secretions.

393 With the elucidation of the catalytic activity of *Pc*8HGO, we provide the GMC
394 oxidoreductase family with an additional functional classified member in insects. This is the
395 first functionally characterised enzyme identified in insects involved in iridoid production. It
396 seems that the substrate diversity in redox reactions supplied by this multi-gene family equips
397 insects with a toolbox that allows them to adjust to the particular biotic and abiotic conditions
398 that may result, for example, when host plants shift. We believe that the characterisation of
399 additional GMC oxidoreductases will help clarifying the role of these enzymes in the
400 adaptation of insects to their environment.

401

402 **Acknowledgment**

403 The authors would like to express their gratitude to Heiko Vogel for making available 454-
404 sequences from *P. cochleariae*. We also gratefully acknowledge Angelika Berg, Gerhard
405 Pauls and Dr. Maritta Kunert for technical assistance. Very special thanks are due to Emily
406 Wheeler for her critical reading of the manuscript. This work was financially supported by the
407 Max Planck Society.

408

409 **Author contributions**:

410 P.R., R.K. and A.B. designed study. P.R. performed the identification of *Pc8HGO*, *Pc8HGO-*
411 *like*, the RNAi experiment, the heterologous expression, the resulting protein assays and the
412 interpretation of all resulting data. R.K. extracted and manually annotated GMC encoding
413 sequences, performed the phylogeny of larval chrysomeline glandular oxidases and related
414 GMC oxidoreductases and made the interpretation. S.K. performed qPCR and contributed to

415 the interpretation of output data. N.W. performed LC/MS^{E} analysis, collected and contributed
416 to the interpretation of output data. M.G. and M.S. generated transcriptome libraries, applied
417 OTP. W.B. and A.B. contributed substantially to the interpretation of all output data. P.R.,
418 R.K. and A.B. wrote the first draft of the manuscript, and all authors contributed substantially
419 to revisions.
420
421 **REFERENCES**
422

423 1. Hammond P. 1992 *Global Biodiversity: status of the Earth's living resources.*
424 2. Gomez-Zurita J., Hunt T., Kopliku F., Vogler A.P. 2007 Recalibrated tree of leaf
425 beetles (Chrysomelidae) indicates independent diversification of angiosperms and their insect
426 herbivores. *Plos One* **2**(4). (doi:ARTN e360 DOI 10.1371/journal.pone.0000360).
427 3. Farrell B.D., Mitter C. 1998 The timing of insect/plant diversification: might
428 Tetraopes (Coleoptera: Cerambycidae) and Asclepias (Asclepiadaceae) have co-evolved? *Biol*
429 *J Linn Soc* **63**(4), 553-577. (doi:DOI 10.1111/j.1095-8312.1998.tb00329.x).
430 4. Pasteels J.M., Duffey S., Rowell-Rahier M. 1990 Toxins in chrysomelid beetles -
431 possible evolutionary sequence from *de novo* synthesis to derivation from food-plant
432 chemicals. *J Chem Ecol* **16**(1), 211-222.
433 5. Meinwald J., Jones T.H., Eisner T., Hicks K. 1977 Defense mechanisms of arthropods.
434 New methylcyclopentanoid terpenes from larval defensive secretion of a chrysomelid beetle
435 (*Plagiodera versicolora*). *Proc Natl Acad Sci USA* **74**(6), 2189-2193. (doi:DOI
436 10.1073/pnas.74.6.2189).
437 6. Blum M.S., Wallace J.B., Duffield R.M., Brand J.M., Fales H.M., Sokoloski E.A.
438 1978 Chrysomelidial in defensive secretion of leaf beetle *Gastrophysa cyanea* Melsheimer. *J*
439 *Chem Ecol* **4**(1), 47-53. (doi:Doi 10.1007/Bf00988259).
440 7. Sugawara F., Matsuda K., Kobayashi A., Yamashita K. 1979 Defensive secretion of
441 chrysomelid larvae *Gastrophysa atrocyanea* (Motschulsky) and *Phaedon brassicae* (Baly). *J*
442 *Chem Ecol* **5**(5), 635-641.
443 8. Pasteels J.M., Braekman J.C., Daloze D., Ottinger R. 1982 Chemical defense in
444 chrysomelid larvae and adults. *Tetrahedron* **38**(13), 1891-1897.

445 9. Pasteels J.M., Rowell-Rrahier M., Braekman J.C., Daloze D. 1984 Chemical defenses
446 in leaf beetles and their larvae - the ecological, evolutionary and taxonomic significance.
447 *Biochem Syst Ecol* **12**(4), 395-406.
448 10. Termonia A., Pasteels J.M. 1999 Larval chemical defence and evolution of host shifts
449 in Chrysomela leaf beetles. *Chemoecology* **9**(1), 13-23. (doi:DOI 10.1007/s000490050029).
450 11. Termonia A., Hsiao T.H., Pasteels J.M., Milinkovitch M.C. 2001 Feeding
451 specialization and host-derived chemical defense in Chrysomeline leaf beetles did not lead to
452 an evolutionary dead end. *Proc Natl Acad Sci USA* **98**(7), 3909-3914.
453 12. Veith M., Lorenz M., Boland W., Simon H., Dettner K. 1994 Biosynthesis of iridoid
454 monoterpenes in insects - defensive secretions from larvae of leaf beetles (Coleoptera,
455 Chrysomelidae). *Tetrahedron* **50**(23), 6859-6874. (doi:Doi 10.1016/S0040-4020(01)81338-7).
456 13. Oldham N.J., Veith M., Boland W., Dettner K. 1996 Iridoid monoterpene biosynthesis
457 in insects: Evidence for a de novo pathway occurring in the defensive glands of *Phaedon*
458 *armoraciae* (Chrysomelidae) leaf beetle larvae. *Naturwissenschaften* **83**(10), 470-473.
459 14. Rowell-Rahier M., Pasteels J.M. 1986 Economics of chemical defense in
460 Chrysomelinae. *J Chem Ecol* **12**(5), 1189-1203.
461 15. Soetens P., Pasteels J.M., Daloze D. 1993 A simple method for in-vivo testing of
462 glandular enzymatic-activity on potential precursors of larval defensive compounds in
463 *Phratora*-species (Coleoptera, Chrysomelinae). *Experientia* **49**(11), 1024-1026. (doi:Doi
464 10.1007/Bf02125653).
465 16. Laurent P., Dooms C., Braekman J.C., Daloze D., Habib-Jiwan J.L., Rozenberg R.,
466 Termonia A., Pasteels J.M. 2003 Recycling plant wax constituents for chemical defense:
467 hemi-biosynthesis of triterpene saponins from beta-amyrin in a leaf beetle.
468 *Naturwissenschaften* **90**(11), 524-527. (doi:DOI 10.1007/s00114-003-0471-y).
469 17. Veith M., Dettner K., Boland W. 1996 Stereochemistry of an alcohol oxidase from the
470 defensive secretion of larvae of the leaf beetle *Phaedon armoraciae* (Coleoptera:
471 Chrysomelidae). *Tetrahedron* **52**(19), 6601-6612.
472 18. Bruckmann M., Termonia A., Pasteels J.M., Hartmann T. 2002 Characterization of an
473 extracellular salicyl alcohol oxidase from larval defensive secretions of *Chrysomela populi*
474 and *Phratora vitellinae* (Chrysomelina). *Insect Biochem Mol Biol* **32**(11), 1517-1523.
475 19. Kirsch R., Vogel H., Muck A., Reichwald K., Pasteels J.M., Boland W. 2011 Host
476 plant shifts affect a major defense enzyme in *Chrysomela lapponica*. *Proc Natl Acad Sci USA*
477 **108**(12), 4897-4901. (doi:10.1073/pnas.1013846108).

478 20. Kirsch R., Vogel H., Muck A., Vilcinskas A., Pasteels J.M., Boland W. 2011 To be or
479 not to be convergent in salicin-based defence in chrysomeline leaf beetle larvae: evidence
480 from *Phratora vitellinae* salicyl alcohol oxidase. *Proc R Soc Lond, Ser B: Biol Sci* **278**(1722),
481 3225-3232. (doi:10.1098/rspb.2011.0175).
482 21. Fernandez I.S., Ruiz-Duenas F.J., Santillana E., Ferreira P., Martinez M.J., Martinez
483 A.T., Romero A. 2009 Novel structural features in the GMC family of oxidoreductases
484 revealed by the crystal structure of fungal aryl-alcohol oxidase. *Acta Crystallogr D* **65**, 1196-
485 1205. (doi:Doi 10.1107/S0907444909035860).
486 22. Cavener D.R. 1992 GMC oxidoreductases - a newly defined family of homologous
487 proteins with diverse catalytic activities. *J Mol Biol* **223**(3), 811-814. (doi:Doi 10.1016/0022-
488 2836(92)90992-S).
489 23. Iida K., Cox-Foster D.L., Yang X., Ko W.Y., Cavener D.R. 2007 Expansion and
490 evolution of insect GMC oxidoreductases. *BMC Evol Biol* **7**, 75. (doi:10.1186/1471-2148-7-
491 75).
492 24. Sun W., Shen Y.H., Yang W.J., Cao Y.F., Xiang Z.H., Zhang Z. 2012 Expansion of
493 the silkworm GMC oxidoreductase genes is associated with immunity. *Insect Biochem Mol*
494 *Biol* **42**(12), 935-945. (doi:10.1016/j.ibmb.2012.09.006).
495 25. Zamocky M., Hallberg M., Ludwig R., Divne C., Haltrich D. 2004 Ancestral gene
496 fusion in cellobiose dehydrogenases reflects a specific evolution of GMC oxidoreductases in
497 fungi. *Gene* **338**(1), 1-14. (doi:10.1016/j.gene.2004.04.025).
498 26. Veith M., Oldham N.J., Dettner K., Pasteels J.M., Boland W. 1997 Biosynthesis of
499 defensive allomones in leaf beetle larvae: Stereochemistry of salicylalcohol oxidation in
500 *Phratora vitellinae* and comparison of enzyme substrate and stereospecificity with alcohol
501 oxidases from several iridoid producing leaf beetles. *J Chem Ecol* **23**(2), 429-443.
502 27. Bodemann R.R., Rahfeld P., Stock M., Kunert M., Wielsch N., Groth M., Frick S.,
503 Boland W., Burse A. 2012 Precise RNAi-mediated silencing of metabolically active proteins
504 in the defence secretions of juvenile leaf beetles. *Proc R Soc Lond, Ser B: Biol Sci* **279**(1745),
505 4126-4134. (doi:10.1098/rspb.2012.1342).
506 28. Stock M., Gretscher R., Groth M., Eiserloh S., Boland W., Burse A. 2013 Putative
507 sugar transporters of the mustard leaf beetle *Phaedon cochleariae*: their phylogeny and role
508 for nutrient supply in larval defensive glands. *Plos One*.
509 29. Petersen T.N., Brunak S., von Heijne G., Nielsen H. 2011 SignalP 4.0: discriminating
510 signal peptides from transmembrane regions. *Nat Methods* **8**(10), 785-786. (doi:Doi
511 10.1038/Nmeth.1701).

512 30. Wierenga R.K., Drenth J., Schulz G.E. 1983 Comparison of the 3-dimensional protein
513 and nucleotide structure of the FAD-binding domain of para-hydroxybenzoate hydroxylase
514 with the FAD-binding as well as NADPH-binding domains of glutathione-reductase. *J Mol*
515 *Biol* **167**(3), 725-739. (doi:Doi 10.1016/S0022-2836(83)80106-5).
516 31. Varela E., Martinet M.J., Martinez A.T. 2000 Aryl-alcohol oxidase protein sequence: a
517 comparison with glucose oxidase and other FAD oxidoreductases. *Bba-Protein Struct M*
518 **1481**(1), 202-208. (doi:Doi 10.1016/S0167-4838(00)00127-8).
519 32. Birkett M.A., Pickett J.A. 2003 Aphid sex pheromones: from discovery to commercial
520 production. *Phytochemistry* **62**(5), 651-656. (doi:Pii S0031-9422(02)00568-X Doi
521 10.1016/S0031-9422(02)00568-X).
522 33. Dobler S. 2001 Evolutionary aspects of defense by recycled plant compounds in
523 herbivorous insects. *Basic Appl Ecol* **2**(1), 15-26. (doi:10.1078/1439-1791-00032).
524 34. Cavill G.W.K., Ford D.L., Locksley H.D. 1956 The chemistry of ant. Terpenoid
525 constituents of some australian *Iridomyrmex* species. *Aust J Chem* **9**(2), 288-293.
526 35. Meinwald J., Chadha M.S., Hurst J.J., Eisner T. 1962 Defense mechanisms of
527 arthropods. Anisomorphal, the secretion of a phasmid insect. *Tetrahedron Lett* (1), 29-33.
528 36. Smith R.M., Brophy J.J., Cavill G.W.K., Davies N.W. 1979 Iridodials and
529 nepetalactone in the defensive secretion of the coconut stick insects, *Graeffea crouani*. *J*
530 *Chem Ecol* **5**(5), 727-735. (doi:Doi 10.1007/Bf00986557).
531 37. Meinwald J., Happ G.M., Labows J., Eisner T. 1966 Cyclopentanoid terpene
532 biosynthesis in a phasmid insect and in catmint. *Science* **151**(3706), 79-&. (doi:DOI
533 10.1126/science.151.3706.79).
534 38. Ohmura W., Hishiyama S., Nakashima T., Kato A., Makihara H., Ohira T., Irei H.
535 2009 Chemical composition of the defensive secretion of the longhorned beetle, *Chloridolum*
536 *loochooanum*. *J Chem Ecol* **35**(2), 250-255. (doi:DOI 10.1007/s10886-009-9591-y).
537 39. Weibel D.B., Oldham N.J., Feld B., Glombitza G., Dettner K., Boland W. 2001 Iridoid
538 biosynthesis in staphylinid rove beetles (Coleoptera : Staphylinidae, Philonthinae). *Insect*
539 *Biochem Mol Biol* **31**(6-7), 583-591.
540 40. Hernandez-Ortega A., Ferreira P., Martinez A.T. 2012 Fungal aryl-alcohol oxidase: a
541 peroxide-producing flavoenzyme involved in lignin degradation. *Appl Microbiol Biotechnol*
542 **93**(4), 1395-1410. (doi:10.1007/s00253-011-3836-8).
543 41. Bowers M.D. 1991 Iridoid glycosides. *Herbivores: their interactions with secondary*
544 *plant metabolites* **Vol. I, 2nd edn (G. A. Rosenthal & M. R. Berenbaum, eds)**, 297-325.

545 42. Jensen S.R. 1992 Systematic implications of the distribution of iridoids and other
546 chemical compounds in the Loganiaceae and other families of the Asteridae. *Annals of the*
547 *Missouri Botanical Garden*, 284-302.
548 43. Jensen S.R., Franzyk H., Wallander E. 2002 Chemotaxonomy of the Oleaceae:
549 iridoids as taxonomic markers. *Phytochemistry* **60**(3), 213-231.
550 44. Szabó L. 2008 Rigorous biogenetic network for a group of indole alkaloids derived
551 from strictosidine. *Molecules* **13**(8), 1875-1896.
552 45. Collu G., Unver N., Peltenburg-Looman A.M.G., van der Heijden R., Verpoorte R.,
553 Memelink J. 2001 Geraniol 10-hydroxylase, a cytochrome P450 enzyme involved in
554 terpenoid indole alkaloid biosynthesis. *FEBS Lett* **508**(2), 215-220. (doi:Doi 10.1016/S0014-
555 5793(01)03045-9).
556 46. Hofer R., Dong L., Andre F., Ginglinger J.F., Lugan R., Gavira C., Grec S., Lang G.,
557 Memelink J., Van Der Krol S., et al. 2013 Geraniol hydroxylase and hydroxygeraniol oxidase
558 activities of the CYP76 family of cytochrome P450 enzymes and potential for engineering the
559 early steps of the (seco)iridoid pathway. *Metab Eng*. (doi:10.1016/j.ymben.2013.08.001).
560 47. Takeuchi H., Rigden D.J., Ebrahimi B., Turner P.C., Rees H.H. 2005 Regulation of
561 ecdysteroid signalling during *Drosophila* development: identification, characterization and
562 modelling of ecdysone oxidase, an enzyme involved in control of ligand concentration.
563 *Biochem J* **389**, 637-645.
564 48. Yamanaka K., Fang L., Inouye M. 1998 The CspA family in *Escherichia coli*:
565 multiple gene duplication for stress adaptation. *Mol Microbiol* **27**(2), 247-255. (doi:DOI
566 10.1046/j.1365-2958.1998.00683.x).
567 49. Lin W.L., Hu X.Y., Zhang W.Q., Rogers W.J., Cai W.M. 2005 Hydrogen peroxide
568 mediates defence responses induced by chitosans of different molecular weights in rice. *J*
569 *Plant Physiol* **162**(8), 937-944. (doi:DOI 10.1016/j.jplph.2004.10.003).
570 50. Santos K.S., dos Santos L.D., Mendes M.A., de Souza B.M., Malaspina O., Palma
571 M.S. 2005 Profiling the proteome complement of the secretion from hypopharyngeal gland of
572 Africanized nurse-honeybees (*Apis mellifera* L.). *Insect Biochem Mol Biol* **35**(1), 85-91.
573 (doi:DOI 10.1016/j.ibmb.2004.10.003).
574 51. Gross J., Muller C., Vilcinskas A., Hilker M. 1998 Antimicrobial activity of exocrine
575 glandular secretions, hemolymph, and larval regurgitate of the mustard leaf beetle Phaedon
576 cochleariae. *J Invertebr Pathol* **72**(3), 296-303. (doi:DOI 10.1006/jipa.1998.4781).
577 52. Gross J., Podsiadlowski L., Hilker M. 2002 Antimicrobial activity of exocrine
578 glandular secretion of Chrysomela larvae. *J Chem Ecol* **28**(2), 317-331.

579

580 **FIGURES**

581

582 **Figure 1.** Steps of the deterrent biosynthesis in the larval gland reservoir of *C. populi*
583 (sequestration) and *P. cochleariae* (*de novo*). The common enzymatic activities are
584 highlighted with boxes. The occurrence of those biosynthetic pathways in various
585 chrysomeline leaf beetle genera is plotted onto their phylogeny SAO, salicyl alcohol
586 oxidoreductase. 8HGO, 8-hydroxygeraniol oxidoreductase. Glc, glucose. *Ph.* spp., *Phratora*
587 species. The figure is adapted from Kirsch *et al.* [20].

588

589 **Figure 2.** GC-MS analysis of larval secretions seven days after treatment with dsRNA-
590 *200bpPc8HGO-like* (a) and dsRNA-*200bpPc8HGO* (b). The picture shows the everted larval
591 glands after the different treatments. Mass range (+/-1): 67+79+105.

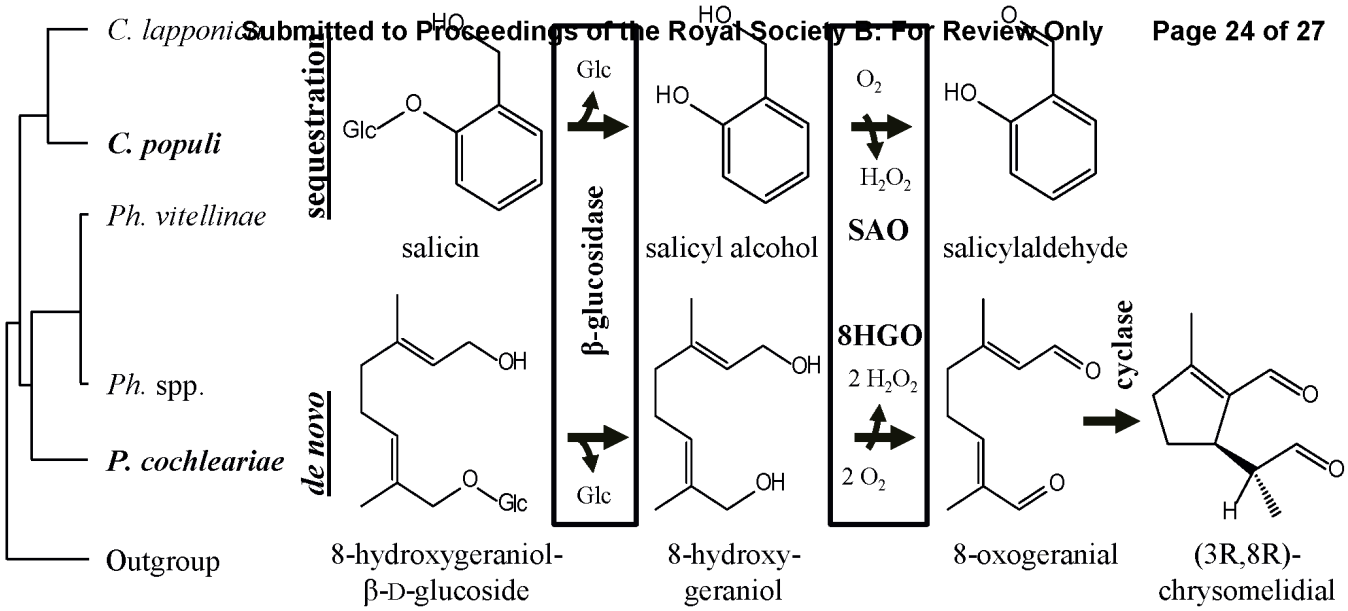
592

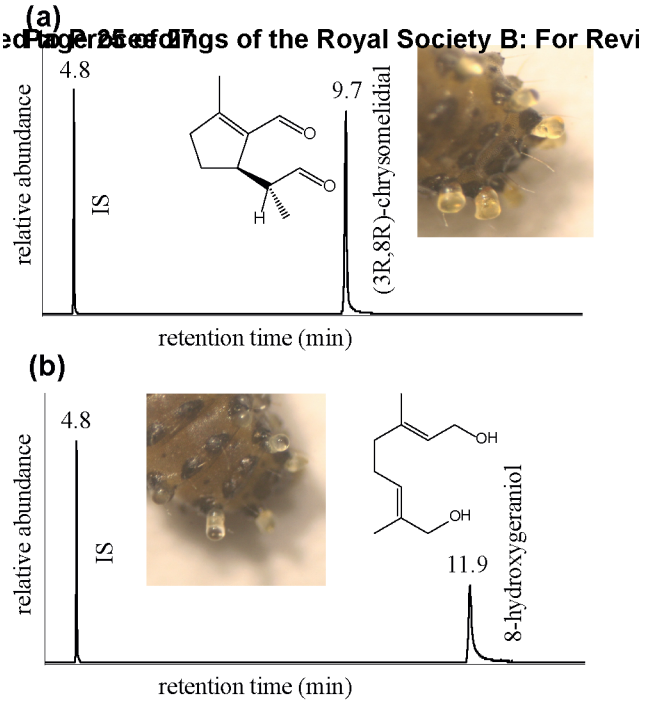
593 **Figure 3.** GC-MS analysis of activity assay with purified protein *Pc*8HGO from insect cell
594 culture medium. The chromatogram shows the conversion of 8-hydroxygeraniol (10.9 min) to
595 8-oxogeranial (14.0 min) after 0, 30 and 60 min. Methylbenzoate (6.2 min) is the internal
596 standard. Two intermediate substances, probably the semi-aldehydes (8-hydroxygeranial and
597 8-oxogeraniol), occur, with retention times of 12.1 and 12.7 min. Mass range (+/-1):
598 67+79+105. The elution fraction of a similarly treated empty vector control was used as the
599 control reaction.

600

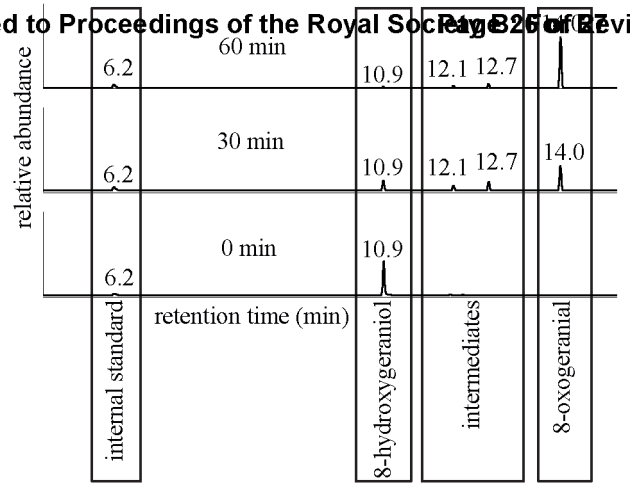
601 **Figure 4.** Phylogeny of chrysomeline glandular oxidases and related GMC oxidoreductases
602 including protein sequences of other insects. The phylogenetic tree was generated using a
603 Bayesian inference method. Posterior probability values are shown next to each node. The
604 second and third numbers, exemplarily indicated, represent bootstrap values based on a

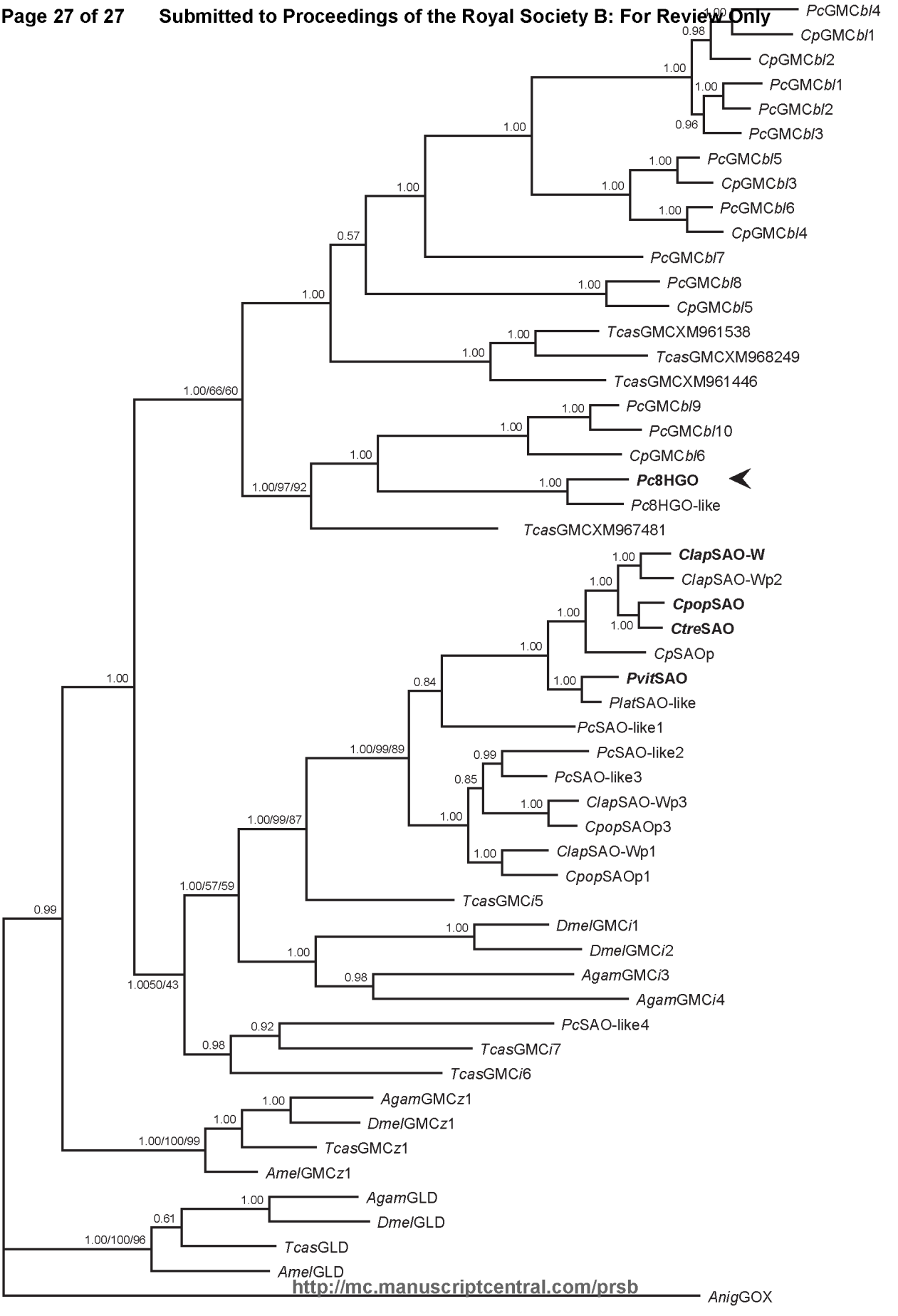
neighbour-joining algorithm and maximum likelihood estimation, respectively, using the same set of data. Abbreviations: *Cp* and *Cpop* (*C. populi*), *Ctre* (*C. tremulae*), *Clap* (*C. lapponica*), *Plat* (*Phratora laticollis*), *Pvit* (*Phratora vitellinae*), *Pc* (*Phaedon cochleariae*), *Tcas* (*Tribolium castaneum*), *Agam* (*Anopheles gambiae*), *Dmel* (*Drosophila melanogaster*), *Anig* (*Aspergillus niger*), SAO-W (salicyl alcohol oxidase of willow-feeder), p (paralogous), GMC (glucose-methanol-choline oxidoreductase), *bl*1 to *bl*10 (beetle-like), GLD (glucose dehydrogenase), GOX (glucose oxidase), 8HGO (8-hydroxygeraniol oxidase), 8HGO-like (8-hydroxygeraniol oxidase-like).


C. lapponica
C. populi
Ph. vitellinae
Ph. spp.
P. cochleariae
Outgroup
sequestration
de novo
salicin
salicyl alcohol
salicylaldehyde
β-glucosidase
Glc
Glc
SAO
O_2
H_2O_2
8HGO
2 H_2O_2
2 O_2
cyclase
8-hydroxygeraniol-β-D-glucoside
8-hydroxy-geraniol
8-oxogeranial
(3R,8R)-chrysomelidial


(a)
4.8
9.7
relative abundance
IS
(3R,8R)-chrysomelidial
retention time (min)
(b)
4.8
relative abundance
IS
OH
OH
11.9
8-hydroxygeraniol
retention time (min)

ed to Proceedings of the Royal Society B: Page 26 of 27




PcGMCbl4
CpGMCbl1
CpGMCbl2
PcGMCbl1
PcGMCbl2
PcGMCbl3
PcGMCbl5
CpGMCbl3
PcGMCbl6
CpGMCbl4
PcGMCbl7
PcGMCbl8
CpGMCbl5
TcasGMCXM961538
TcasGMCXM968249
TcasGMCXM961446
PcGMCbl9
PcGMCbl10
CpGMCbl6
Pc8HGO
Pc8HGO-like
TcasGMCXM967481
ClapSAO-W
ClapSAO-Wp2
CpopSAO
CtreSAO
CpSAOp
PvitSAO
PlatSAO-like
PcSAO-like1
PcSAO-like2
PcSAO-like3
ClapSAO-Wp3
CpopSAOp3
ClapSAO-Wp1
CpopSAOp1
TcasGMCi5
DmelGMCi1
DmelGMCi2
AgamGMCi3
AgamGMCi4
PcSAO-like4
TcasGMCi7
TcasGMCi6
AgamGMCz1
DmelGMCz1
TcasGMCz1
AmelGMCz1
AgamGLD
DmelGLD
TcasGLD
AmelGLD
AnigGOX
beetle GMC - clade
insect GMCi - clade
0.1

4 General Discussion

In the published manuscripts of my doctoral thesis, I investigated the *de novo* assembled transcriptomes of two different leaf beetle species, namely *Phaedon cochleariae*, the mustard leaf beetle, and *Chrysomela populi*, the poplar leaf beetle. Each manuscript includes a detailed discussion. Therefore, in the following I shall give an overview of specific topics not discussed in detail in the publications.

4.1 RNA sequencing and its applications

In many cases, no genomic data is available for organisms, especially non-model organisms such as leaf beetles. Sequencing of their transcript libraries is a comprehensive and cost-effective way to study them. For this, the total RNA of *P. cochleariae* and *C. populi* was extracted, prepared and sequenced using the Illumina sequencing technology. Subsequently, the RNA pools of each species were assembled *de novo* to reconstruct the individual transcript library for each leaf beetle species. Hence, several assembler software tools were applied and tested, among them CLC Genomics Workbench (http://www.clcbio.com) and Trinity (Grabherr et al., 2011).

4.1.1 Challenges in de novo assemblies of transcriptomes

Sequencing technologies are well prepared and developed, but challenges arise for handling and analyzing transcriptome sequences when there is no reference genome. Without a reference genome, is it possible to assess the quality of a *de novo* assembled transcriptome?

In order to reconstruct a transcript catalogue for *P. cochleariae* as well as *C. populi*, I applied the CLC Genomics Workbench, T-IDBA (Peng et al., 2011), ABySS (Simpson et al., 2009) and also Trinity which eventually proved to be the assembly tool of my choice.

Exerting the CLCbio *de novo* assembly tool with suitable parameter settings resulted in a reasonable number of transcripts (approximately 57,000 contiguous sequences), but only 70 percent of all read pairs were employed and many read pairs were split in order to be integrated into the assembly (unpublished data). Applying ABySS yielded many (over one million) very short transcript fragments. Changing the parameter settings did not improve the assembly. Thus, this assembly software tool also turned out to be unfeasible concerning the underlying raw sequence data. Utilizing T-IDBA failed to assemble the data, due to the exceeding of the virtual memory (unpublished data). Finally, applying Trinity, with subsequent reassembly (as

published in Manuscripts 2 to 4), resulted in the most feasible and acceptable leaf beetle transcript libraries including isoforms. This has been demonstrated with leaf beetle sequences that have been amplified and sequenced in earlier studies. Those have been compared to the *de novo* assembled transcripts to check their integrity. Thus, two studied *P. cochleariae* isoprenyl diphosphate synthases (accession number KC109782, and unpublished) which are half-part identical could be resolved only by the Trinity assembler software. Also the *C. populi* ABC transporter (KC112554, 4 kb) could be reconstructed to its full length by applying Trinity and by subsequent reassembly (unpublished data).

Two tools, namely Trinity and SOAPdenovo-trans, for *de novo* assembling transcriptomes have been examined by Vijay *et al*. They postulate that Trinity assemblies "appeared to be more contiguous than assemblies obtained from SOAPdenovo-trans" (Vijay et al., 2013). Additionally, they observed in Trinity assemblies that many isoforms were inferred inaccurately. But having no genomic data as reference, it is a challenge to distinguish correctly and falsely inferred isoforms. Regardless, "the true number of isoforms is not known, even in model-organisms" (Wang et al., 2008). Overall, transcriptome assemblies are assured to be robust and accurate and the information on isoforms is valuable. Regarding erroneously inferred isoforms, the data should be interpreted with caution.

Yang and Smith studied transcriptome assemblies with regard to phylogenomic purposes (Yang and Smith, 2013). In this case, only one representative transcript for each gene is needed and splice variants are not utilized since the detection of true paralogs would be hindered. Their results have not been used in the phylogenetic studies carried out for the incorporated manuscripts because they were not published by the time the investigations were carried out. For future phylogenetic studies, their proposals for parameter settings should be taken into consideration for the improvement of gene coverage of transcripts.

The depth of sequencing is another aspect during the assembly process that should be considered. If transcripts are very rare in specific samples it is likely that they are absent in the *de novo* assembly. And this is a critical fact for data interpretation. Hence, one should be aware of modest cost and data-handling when deeper sequencing is planned to increase the transcript coverage (Gongora-Castillo and Buell, 2013, Wall et al., 2009). Furthermore, some problems may be overcome by hybrid sequencing strategies using a mixture of sequencing methodologies, such as a combination of FLX and Illumina sequencing for optimal transcriptome coverage (Wall et al., 2009).

Additionally, also Lu *et al.* and Martin *et al.* observed several *de novo* assembly tools and genome-guided assemblers and propose a combination of several assembler tools to achieve more complete transcriptome assemblies (Lu et al., 2013, Martin and Wang, 2011).

4.1.2 Normalization and identification of differentially expressed genes

RNA sequencing, additionally, is not only used to study and reconstruct the transcriptome of an organism, but also provides a great opportunity to study the transcript level for each gene specifically. This can be achieved for many samples, either for tissue-specific or for treated samples. Treated insects in our case have been injected with interfering dsRNA to study proteins (with special focus on enzymes and transporters involved in sequestration) by looking at not only phenotypic but also transcriptional changes. In order to do statistical analysis for these samples, normalization has to be carried out beforehand.

Many crucial aspects have to be considered in such analyses, such as read counting, quality control, appropriate treatment of biological variability and appropriate statistical modeling. Anders *et al.* established a count based differential expression method that aims at the discovery of changes in transcript abundance between experimental groups (Anders et al., 2013). In most RNA-seq experiments, often not more than two or three biological replicates are designed. This requires statistical methods suitable for such small-sample situations. In the case for the R package DESeq which was applied in Manuscripts 2 and 3 the formal distribution assumed for the underlying data is approximated by the negative binomial model (Anders et al., 2013, Anders and Huber, 2010). This distribution has two parameters: namely the mean and the dispersion which are used to model mean-variance relationships. The Poisson distribution, in contrast, is used to describe technical replicates and is constrained in that it supposes the variance of the variable to be equal to the mean (Soneson and Delorenzi, 2013).

Normalization and isoform-specific expression

Recently, several approaches have been proposed to estimate transcript-level expression, but isoform-specific expression estimation remains enhanced.

Besides FPKM as measure for transcript abundances which is defined as expected fragments per kilobase of transcript per million fragments mapped (Trapnell et al., 2010) and RPKM as measure for transcript levels in reads per kilobase of exon model per million mapped reads which takes the sequence length into account (Mortazavi et al., 2008), Howard *et al*. proposed the "ability to reliably compute quantitative isoform expression values" by using a non-uniform read sampling distribution (Howard and Heber, 2010), and Li *et al*. observed a general distribution bias when inferring expression (Li et al., 2010). In addition, Wu *et al.* combined Howard *et al*.'s proposal and Li *et al*.'s work to effectively improve the expression inference (Hansen et al., 2011).

Furthermore, there are methods that adjust the samples according to their transcript distributions among samples. Marioni *et al.* take into account the total number of reads (Marioni

et al., 2008), and Bullard *et al*. also proposed an upper-quartile normalization procedure which scales the expression level at the 75th percentile in each sample to the average across all samples (Bullard et al., 2010, Glusman et al., 2013). Robinson *et al*. applied a quantile-adjusted estimator for the negative binomial distribution (Robinson and Smyth, 2008) which has a great success in very small samples. Two years later in 2010, Robinson *et al*. developed the trimmed mean of M-values (TMM) normalization method (Robinson and Oshlack, 2010) which is implemented in the edgeR Bioconductor package. Srivastava *et al*. proposed a two-parameter generalized Poisson model to "fit the position-level read counts more appropriately than a traditional Poisson-model" (Srivastava and Chen, 2010). They also observed that the standard RPKM can bias estimates of differential expression (Bullard et al., 2010). Glusman *et al.* developed a data-driven algorithm that computes sample-specific scaling factors. These factors lead to correct results only in the context of their comparison to other samples which means that, depending on the other samples to which the examined sample is compared, the examined sample is scaled differently (Glusman et al., 2013). With this method they aim to provide accurate expression levels.

Alternative splicing and the presence of isoforms have to be remembered when attempting to improve the assembly by correcting sequencing errors. MacManes and Eisen examined the effect of sequencing error correction on the assembly accuracy. They concluded, that because the assembly accuracy increases, this step should become routine in processing Illumina sequencing data (Macmanes and Eisen, 2013). In any case, it will remain a problem for error correctors to distinguish between differences in isoforms and sequencing errors. Therefore, one should be aware of 'missing' short reads to reconstruct isoforms that are masked as errors (Yang et al., 2013).

SEQUENCING DEPTH

Other important factors for normalizing expression data are the sequence length as well as the sequencing depth. Mortazavi *et al.* stated that the transcript length plays an important role in identifying correct isoforms and their corresponding expression levels (Mortazavi et al., 2008). In addition, Bullard *et al.* observed that longer genes are more likely declared differentially expressed (Bullard et al., 2010). Moreover, Tarazona *et al*. observed "that most existing methodologies suffer from a strong dependency on sequencing depth for their differential expression calls". The more reads there are, the higher the number of false positives that are declared differentially expressed (Tarazona et al., 2011). Therefore, they proposed a non-parametric method to model the noise distribution from actual data. This approach is robust with regard to sequencing depth changes (Chen et al., 2011). In addition to that, Francis *et al*. observed in their studies that a sequencing depth of 30 million reads leads to a good balance between transcript coverage and (sequencing) noise. When sequencing more than 60

million reads only a few more new genes are discovered, but the errors during the sequencing process are likely to increase (Francis et al., 2013). Beyond this, Vijay *et al.* stated a "high overall coverage is vital for successful RNA-seq experiments" (Vijay et al., 2013). Most sequencing runs carried out for all samples of various (larval) tissues derived from the leaf beetles *P. cochleariae* and *C. populi* resulted in at least 20 million reads that showed coverage for the majority of transcripts and were sufficient to conduct statistical analysis.

To conclude, many normalization methods and tools to determine differentially expressed genes are available. The counting approach by Anders *et al.* is "direct, flexible and can be used for many types of count data" (Anders et al., 2013) and was used in our studies. This approach is implemented in the DESeq Bioconductor package and straightforward to apply. Dillies *et al*. favor this package, in addition to the TMM method, since it uses a normalization method within the statistical model for differential analysis rather than the data themselves (Dillies et al., 2012). Additionally, Rapaport *et al.* conclude "that methods based on negative binomial modeling (DESeq, edgeR, and baySeq) have improved specificity and sensitivity as well as good control of false positive errors" (Rapaport et al., 2013). In any case, no method is optimal under all circumstances or conditions, and there is still a need for further research into developing statistical methods for mRNA-Seq.

BIOLOGICAL VARIANCE

Independently of all normalization methods and sequencing as well as library preparation effects, biological variance is observed among all samples studied, also in those published in this thesis. Significant results in studies with just few biological replicates may be due to biological variation and thus not reproducible. Furthermore, it is impossible to know whether expression patterns are characteristic for a specific individual or typical of the population used in the study. For example, there are large differences observed in the expression of genes which might likely be important since the expression of that specific gene varies only a little bit among the individuals, but this same expression variance shown by certain other genes is meaningless, because the expression of that gene is known to vary highly (Hansen et al., 2011). The only possibility to overcome this problem is sequencing many more biological replicates (Marioni et al., 2008, Bullard et al., 2010, Liu et al., 2013). Anyhow, no sequencing technology is able to "eliminate biological variability" (Hansen et al., 2011).

4.1.3 FUTURE APPLICATIONS OF NEXT-GENERATION SEQUENCING

Next to the quantitative aspect of sequencing, RNA sequencing is used to annotate transcripts including mutations, aberrant transcripts, gene regulations or splice isoforms, which

provides a basis for the in-depth observation of diseases (Marguerat and Bahler, 2010, Costa et al., 2010, Robertson et al., 2010, Horner et al., 2010, Simon et al., 2009). RNA-seq will lead to many more exciting discoveries within the next few years, and it permits researchers to analyze dynamics of transcriptomes, genome regulations and even evolutionary mechanisms (when sequencing is deep enough) independently of transcript size or knowledge of genomes (Marguerat and Bahler, 2010). This cannot be said of hybridization-based methods such as microarrays. Expression profiling using microarrays has been very successful. Nevertheless, microarrays are able to observe only the genes of interest that are spotted on the chip without distinguishing different isoforms. Furthermore, the required amount of RNA needed is very high in contrast to RNA-seq. Microarrays may remain the method of choice to probe specific subsets of genes being tested over many samples. The greatest current advantage of microarrays over sequencing is the comparatively low cost. Still, since the cost of sequencing will drop further in the near future, RNA-seq is expected to replace microarray analyses for many applications (Wang et al., 2009, Feng et al., 2010, Liu et al., 2013) and for broad genome- and transcriptome-wide experiments, including isoform identification. NGS technologies are rapidly gaining acceptance (Cullum et al., 2011). Detecting genes with low expression remains a problem for microarrays but also for sequencing technologies (Malone and Oliver, 2011).

New sequencing technologies, also called third-generation sequencing, such as single molecule sequencing provided by the Helicos™ Single Molecule Sequencing technology, directly sequence cellular nucleic acids in an unbiased manner. Most NGS technologies require an amplification step, and the target molecules fall within a specific range of size which is not the case for single molecule sequencing (Thompson and Steinmann, 2010). The high accuracy made possible by sequencing and the very small amounts of starting material (picogram quantities) required are advantageous for studying samples with low cell content, such as tumor samples, or samples with miRNAs. This cannot be achieved by any amplification-based sequencing system (Thompson and Steinmann, 2010, Hart et al., 2010). Taking these advantages into account, Ozsolak and Milos developed the first direct RNA sequencing (DRS) technology based on the Helicos™ Genetic Analysis System. “The DRS sample preparation step involves only polyadenylation of 3' ends of RNA molecules without the need of complicated and potentially biased steps such as ligation or PCR amplification of cDNAs” (Ozsolak and Milos, 2011b). DRS has the ability to sequence several femtomoles of RNA in a low-cost and high-throughput manner. This technology enables the research of polyA+RNA species, mapping of polyadenylation sites and mutation detection. To conclude, it is a valuable tool for various clinical diagnostic applications (Ozsolak and Milos, 2011a).

Another third-generation sequencer is PacBio (Eid et al., 2009) developed by Pacific Biosciences (http://www.pacific-biosciences.com/index.php). Their single-molecule real-time (SMRT) technology is based on zero mode waveguides, and the observation volume is only 20

zeptolitres. This technology is able to measure the fluorescence of nucleotides incorporated by a single DNA polymerase enzyme into a growing DNA strand in real time. Its RNA-seq reads are up to several kilobases long. These sequencers are therefore "capable of sequencing a single transcript to its full length in a single read". This denotes "no assembly required" transcriptomes in the future (Martin and Wang, 2011).

In conclusion, the availability of low-cost, efficient and accurate technologies to study gene expression will provide new knowledge and insights in pathology and common genetic disorders, and it will help to understand drug response and nutrient-gene interactions. This will lead "to the development of targeted therapies for many human diseases" (Costa et al., 2010) with special focus on single cell transcriptomics (McGettigan, 2013).

Beyond mRNA sequencing, research also focusses on proteomics and epigenetics. In future, orthogonal transcriptomic and proteomic datasets will be observed in order to measure linkages between RNA and protein levels in individual samples. Epigenetic modifications on a genome-wide scale and posttranslational modifications of histones can be examined and the essential role in cellular processes, in disease appearance and in oncogenetic development can be investigated. This will provide a deeper and more complete understanding of individual cells and different tissue types (McGettigan, 2013, Mutz et al., 2013).

4.2 RNAi-mediated silencing to identify enzyme functions

RNA interference (RNAi) has become a valuable and widely used research tool to knock down and analyze the function of genes, especially in non-model organisms. Most of the research using RNAi was done to investigate developmental processes and functions of enzymes in insects, such as Lepidoptera (Terenius et al., 2011) and Coleoptera (Tomoyasu et al., 2008, Tomoyasu and Denell, 2004, Baum et al., 2007, Bai et al., 2011, Alves et al., 2010). RNAi decreases transcript and protein levels, which enables the demonstration and proof of defined enzymes *in vivo,* also of particular metabolic pathways (Frick et al., 2013, Alves et al., 2010, Belles, 2010, Mito et al., 2011, Terenius et al., 2011, Ohnishi et al., 2006, Ohnishi et al., 2009).

4.2.1 RNAi in leaf beetles

This technique was established for leaf beetle larvae because of the interest in identifying enzymes involved in the biosynthesis of deterrent secretions. Furthermore, it is used to demonstrate *in vivo* relevance of target sequences which were already identified by either HPLC/MS analyses or by using degenerated primers for candidate genes (Manuscript 1). The

known SAO derived from *C. populi* (Kirsch et al., 2011a, Kirsch et al., 2011b) (Manuscript 1) was used as a target sequence to validate the applicability of this method. RNAi of the SAO leads to accumulation of salicyl alcohol because the production of the deterrent salicyl aldehyde was interrupted. Additionally, this method was also applied for the *de novo* producer *P. cochleariae*. In *P. cochleariae* the deterrent chrysomelidial is synthesized in the defensive glands *via* a cyclization step. And the specific knock-down of a putative cyclase resulted in the accumulation of another compound, namely 8-oxogeranial which is the non-cyclized precursor in addition to chrysomelidial (Manuscript 1).

In addition to enzymes involved in the sequestration and deterrent production process, the RNAi method was also applied to study sugar transporters (Manuscript 2) as well as a characterized ABC transporter (Strauss et al., 2013, Strauss et al., 2012)(Manuscript 3).

As published in Manuscript 2, RNAi is a powerful tool, but there are cases in which no phenotypic changes can be observed. In Manuscript 2 four sugar transporters highly expressed in the defensive glands were silenced *via* RNAi, but for two of them the chrysomelidial amount decreased only slightly. Subsequent RNA sequencing revealed a counter-regulation of other sugar transporters that compensates the decrease of targeted transcripts.

4.2.2 Off-target prediction to ensure correct silencing via interfering RNA

In order to ensure that only the target transcript will be knocked down, possible off-targets should be determined and sequence homologies clarified. Only dsRNA, that is unique for the target sequence, should be injected or taken up by the organism, and for which off-target prediction has been carried out before.

Several off-target prediction tools are available (Sabirzhanov et al., 2011, Amarzguioui and Prydz, 2004, Reynolds et al., 2004, Chalk et al., 2004, Henschel et al., 2004, Arziman et al., 2005, Luo et al., 2007, Gong et al., 2008). For this thesis I was able to establish a method that takes all possible 21 bp fragments of the target transcript and compares those against to the corresponding transcript catalogue by applying BLAST (Manuscript 1). Thereafter, only those sequence parts that are homologous to the target sequence are used to design the dsRNA that is to be injected or taken up by the organism. Sabirzhanov *et al*. (Sabirzhanov et al., 2011) proposed a method using rapid amplification of cDNA ends (RACE) which can be especially useful for organisms that have limited sequence data available. Due to the requirements for primers that are used in their RACE method (such as a high G/C content) it cannot be guaranteed that “all dsRNA that may cross-react with non-target mRNAs” are eliminated because of the fact that the designed dsRNAs may not fit these primer requirements. This circumstance is not existent in our proposed method published in Manuscript 1. Many software tools mentioned

above use several criteria for designing dsRNA such as the GC content or thermodynamic stability besides the sequence homology (which is the defining component of the off-target prediction method published in Manuscript 1). However, in our studies, including those published in Manuscript 2 and 3, almost no silencing of non-target sequences was observed. When studying the sugar transporters in *P. cochleariae*, a few other transporters were down-regulated or knocked-down *via* RNAi besides the silenced and targeted transporter. Since there is no obvious sequence homology, it is assumed that this co-silencing effect is not induced by RNAi but metabolically caused. This might be achieved by the larvae sensing the missing transporter and a feedback activation of gene expression of other (high-affinity) sugar transporters to ensure that the sugar level is maintained at a stable level (Kim et al., 2013). Irrespective of possible off-targets, during this work, it always was the case that the transcript level of the supposed target sequences decreased most significantly in the observed samples. Hence, the proposed method of Manuscript 1 seems suitable and applicable.

4.3 Phylogeny and application of sugar and ABC transporters

Leaf beetle species have developed the successful strategy of releasing defensive secretions *via* dorsal glands to defend themselves in case of an attack. To establish the production of these deterrents, the glucosidic precursors have to be transported *via* the hemolymph into the defensive glands where they are converted into the deterrents. The mustard leaf beetle *P. cochleariae*, on the one hand, is able to produce the precursor compounds *de novo* in the fat body, whereas, on the other hand, the poplar leaf beetle *C. populi* takes up the glucosidic compounds by feeding on its host plant.

The transport of the precursor compounds *via* the hemolymph indicates a complex mechanism involving many transport proteins. And it is known for all investigated larvae that the uptake of the precursors into the defensive glands is highly selective and substrate-specific.

Manuscript 2 illustrates phylogenetic studies of all putative sugar transporters of *P. cochleariae* that might play a role in the defensive process or provide energy for metabolic processes. In Manuscript 3 putative ABC transporters identified in *C. populi* larvae were investigated.

Molecular evolutionary studies are carried out in many research areas, particularly since the number of sequence databases has increased greatly. Concurrently, the number of applied methodologies has also progressed considerably. Powerful methods statistically infer probabilistic models of the biological processes (Whelan et al., 2001) which is then presented in phylogenetic trees. In Manuscripts 2 and 3 phylogenetic trees have been calculated by using, on the one hand, a Bayesian inference method, and on the other hand, a maximum likelihood (ML)

approach. For both of them, models of amino acid replacements have been applied (Manuscripts 2, 3 and 4). The best estimate for a phylogeny when using Bayesian inference is the tree with the highest posterior probability (PP). Using an ML approach, the ML bootstrap percentages (BP) are calculated for each node. But what kind of relationship is there between PP and BP? Douady *et al.* observed a strong correlation between PP and BP. Furthermore, they stated that Bayesian inference very efficiently estimates "substitution model parameters, branch lengths and topology", but also that BP "might be less prone to strongly supporting a false phylogenetic hypothesis". In conclusion, both methods have a great impact on phylogeny and state "potential upper and lower bound of node support, but they are surely not interchangeable and cannot be directly compared" (Douady et al., 2003). Carrying out both types of methodologies to infer phylogenetic trees, which was also performed in Manuscripts 2, 3 and 4, generates trustworthy phylogenetic constructs for interpretation.

Over the last few years sugar transporter classes have been studied regarding sugar transporter evolution. In 2001, Joost *et al*. (Joost and Thorens, 2001) stated that the GLUT 6 and 8 are the oldest GLUT isoforms. Later on, Wilson-O'Brien *et al.* (Wilson-O'Brien et al., 2010) extended this observation and proposed that this clade has "arisen after the divergence of the metazoans" which is supported with a bootstrap value of 100%. We could also clearly demonstrate (Manuscript 2) that this clade is also shared by invertebrates including a huge number of trehalose transporters as sister group.

It is highly likely that these many different sugar transporters including all isoforms of glucose transporters (GLUTs) play specific unique roles, which could be one of many different functions, such as different glucose handling in different cell types to regulate metabolism, gene expression, differentiation, or oncogenesis. Especially for the GLUTs 1-5 it was shown that they each have a specific role in the control of sugar homeostasis (Thorens and Mueckler, 2010, Wood and Trayhurn, 2003, Zhao and Keating, 2007). This is stated because of their tissue-specific expression, substrate specificity and role in control of glucose metabolism.

Might their transport mechanism be someday understood, and can it be adapted to pharmacological use in the treatment of diseases such as cancer? It has already been shown that the GLUT 4 is elusively regulated by insulin, and that the knock-down of this regulation is predominant in obesity and type 2 diabetes mellitus (Thorens and Mueckler, 2010).

Additionally, the initiation and progression of cancer and tumors is suspected for the GLUT 12 when glucose supply is low since its function is to provide energy (Wilson-O'Brien et al., 2010, Macheda et al., 2005). However, Mueckler *et al*. stated that the role of GLUT 12 under normal conditions still remains unknown (Mueckler and Thorens, 2013).

The usefulness of trehalose transporters has not been addressed before in detail, but a few proposals were put forward by Kikawada *et al*. (2007) and Kanamori *et al*. (2010)

(Kikawada et al., 2007, Kanamori et al., 2010). By studying trehalose transporters, the biological functions of trehalose can be determined. And thus, trehalose was investigated as a chemical chaperone, a radical scavenger and a signal molecule *in vivo*.

The observation of ABC transporters, has shown that genes encoding ABC transporters are conserved among vertebrates (Dean and Annilo, 2005). Furthermore, ABC transporter genes are also conserved in invertebrates, and, additionally, the expansion of ABC transporter subfamilies (ABCC, ABCG and ABCH) has been determined (Broehan et al., 2013, Xie et al., 2012, Sturm et al., 2009).

Studying the ABC transporters in terms of their phylogeny is difficult since this superfamily contains some of the most functionally diverse proteins known (Dean and Allikmets, 1995). Despite the difficulty, phylogenetic analysis was used to divide all known ABC transporters into eight distinct subfamilies of proteins (ABCA-H) (Dean and Annilo, 2005, Dassa, 2011). It is proposed that these subfamilies originated by gene duplication. Gene duplications occurred to provide specific functions in many metabolic processes and to cope with multicellularity (Dean and Annilo, 2005). Especially the "vertebrate evolution has been largely driven by the duplication of genes that allows for the acquisition of new functions" (Annilo et al., 2006). Annilo *et al.* identified multiple gene duplication and deletion events in different lineages. They observed that zebrafish is the only vertebrate with a member of the ABCH subfamily. ABCH genes had been thought present only in insects (Annilo et al., 2006). Additionally, studies of the ABC transporters in *Daphnia pulex* (planktonic crustacean) revealed a high number of gene duplications in the ABCG and ABCH subfamilies (Sturm et al., 2009), and Xie *et al.* observed in *Bombyx mori* that the ABCC subfamily has expanded more than in other species (Xie et al., 2012, Liu et al., 2011). This expansion of the ABCC subfamily was also observed in *Tribolium castaneum* (Broehan et al., 2013) and *C. populi* [Manuscript 3].

The ABC genes are one of the gene families with an enormous number of members in all eukaryotes. Since many of these transporters "play a role in the efflux of a wide variety of substrates" and the loss-of-function is crucial for specific diseases, such as Mendelian diseases, the ABC transporters are and will be in the focus of further research to obtain new insights into their genes´ contribution to many diseases (Dean and Annilo, 2005, Moitra and Dean, 2011, Silverton et al., 2011).

Wu *et al.* published on the coordinated mediation of SLCs and ABC transporters to move a huge variety of substrates across epithelial barriers (Wu et al., 2011). They observed that homeostasis is achieved by the coordination of transporters having overlapping substrate preferences. Furthermore, they found that those transporters are sensitive to environmental

changes and that remote sensing and signalling occurs. However, the sensing mechanism is still undetermined until now (Wu et al., 2011).

As future perspective and with regard to leaf beetles, combined studies of sugar as well as ABC transporters will result in comprehensive knowledge. Differential expression of about 500 genes was observed when knocking down the specific ABC transporter that is extraordinarily expressed in the defensive glands of *Chrysomela populi* larvae (unpublished data). Continuing analysis of these differentially expressed genes will provide deeper insights into involved metabolic pathways and also other (transporter) proteins associated with this gland-specific ABC transporter.

5 SUMMARY

Chrysomelina larvae possess a sophisticated strategy in terms of chemical defense eminently adapted to their natural habitat. In case of a predatory attack, the deterrent secretions are presented from dorsal thoracic and abdominal glands. The source of the deterrent compound for this ingenious mechanism depends on different biosynthetic strategies and subdivides the Chrysomelina into three different groups. The ancestral strategy represents the *de novo* production of iridoids. Species belonging to this group, e.g. the mustard leaf beetle *Phaedon cochleariae,* produce the deterrent compound independently from host plant derived precursors. More evolved species of the *Chrysomela* sequester phenol glucosides to produce salicylaldehyde. The third group includes several species of the *interrupta*-group belonging also to the genus *Chrysomela*. They exhibit a combinatorial biosynthetic strategy by using *de novo* and phytogenic derived precursors to produce butyrate-esters as defensive compounds.

Irrespective of the different deterrent substances, the defense systems of all groups exhibit a uniform architecture and morphology and constitute a research subject providing enzymatic steps underlying permanent evolutionary changes and adaptation which in this flexible manner is not found in any other organism.

Until quite recently, almost no genomic and transcriptomic data were available for leaf beetles. Therefore, because sequencing of the genomes of insects is still very cost-intensive and the assembly without genome templates complicated, transcript libraries of a variety of Chrysomelina larvae were established. This was achieved by the *de novo* assembly of short paired reads resulting from Illumina sequencing (in cooperation with the Fritz-Lipmann-Institute, Jena). Several assembly software tools were applied and tested. Trinity was the assembly tool of choice and used to create feasible and applicable assemblies of the leaf beetle transcriptomes, especially of the poplar leaf beetle *Chrysomela populi* and *P. cochleariae*. In addition to the establishment of transcript libraries, an off-target prediction method was developed to facilitate an injection of unique dsRNA suitable for the target sequence when performing RNA interference experiments. This off-target prediction is applied prior to injection and helps to design dsRNA to avoid the knock-down of off-targets. This methodology facilitates the identification and functional annotation of specific enzymes, membrane proteins and all other genes of interest within a convenient timeframe and with convenient effort.

This approach combining RNA-seq with RNAi promoted the observation of enzymes involved in the sequestration process. First, the total RNA of various tissues of *P. cochleariae*

as well as *C. populi* larvae was extracted, prepared, and finally sequenced resulting in many tens of thousands of short reads. Thereafter, those reads were mapped onto the corresponding transcriptome database, and the transcript counts were obtained for each present transcript. These steps were also performed after injecting the larvae with dsRNA unique for proteins of interest involved in sequestration. The obtained read counts for each transcript were then compared and statistically evaluated with regard to significant differential "expression".

The off-target prediction method was published in Manuscript 1 and was successfully applied for the silencing of two enzymes namely a salicyl alcohol oxidase in *C. populi* larvae and the juvenile hormone-binding protein superfamily identified in the defensive glands of *P. cochleariae* larvae. RNAi targeting the most abundant sugar transporters in *P. cochleariae* (Manuscript 2) resulted in a slight decrease in the amount of deterrent but subsequent RNA-seq analysis revealed a counter-regulation of other sugar transporters. This counter-regulation is needed to ensure the defense ability and nutrient supply in the larvae of mustard leaf beetles. In Manuscript 4, RNAi targeting the most abundant glandular ABC transporters in *C. populi* showed an immense decrease in the transcription level of itself and almost no counter-regulation of any other ABC transporter. This validates the significant function of this specific transporter (Strauss et al., 2013). Many more sequences are differentially expressed, besides this knocked down specific ABC transporter, though continuative analysis remains an open task.

Phylogenetic analyses accessorily carried out accounted for additional information and annotation of selected leaf beetle transcripts with regard to already functionally characterized enzymes and transporters. The transcripts could be grouped according to their sequence homologies and putative functions, or substrate specificity could be assigned. When considering the importance of the great number of sugar and ABC transporters and their diverse functions, the methods mentioned above promote a deeper knowledge and novel insights of known and unknown regulatory processes. Additionally, I was able to show for a small group of sugar transporters in *P. cochleariae* that they are leaf-beetle specific and not present in mammals, for example.

In conclusion, next-generation sequencing (especially focusing on RNA) and the RNA sequencing after triggering RNA interference (RNAi) holds great opportunities for the study of genes of interest. This approach is especially valuable for the observation of non-model organisms´ transcriptome in its entirety and for accomplished transcriptome-wide analyses, in this thesis, of the mustard leaf beetle *P. cochleariae* and of the poplar leaf beetle *C. populi.*

6 ZUSAMMENFASSUNG

Die Larven der Blattkäfergruppe Chrysomelina besitzen eine raffinierte und spezifisch an ihre Umwelt angepasste chemische Abwehrstrategie. Im Falle eines Angriffes von potentiellen Fraßfeinden können die Larven ein Abwehrsekret über dorsale Drüsen präsentieren, das die Angreifer abschreckt. Die Quelle des Abwehrstoffes für diesen ausgeklügelten Mechanismus ist abhängig von verschiedenen biosynthetischen Strategien, die die Gruppe der Chrysomelina in drei Gruppen aufteilt. Die erste Strategie umfasst die *de novo*-Herstellung der Iridoide. Spezies, die zu dieser Gruppe gehören, wie der Meerrettichblattkäfer *Phaedon cochleariae*, produzieren ihre Abwehrstoffe, z.B. das Iridoid Chrysomelidial, unabhängig von den Zuckerverbindungen der Wirtspflanze, auf der sie fressen. Weiter entwickelte Spezies der Gattung Chrysomela sequestrieren phenolische Glucoside, um Salicylaldehyd herzustellen. Weiterhin gibt es einige Spezies der Gattung Chrysomela, die in der interrupta-Gruppe zusammengefasst werden. Deren Eigenschaft ist es, *de novo* hergestellte und pflanzlich gewonnene Vorstufen synthetisch miteinander zu kombinieren, um so Butyrat-Ester als Abwehrmoleküle herzustellen.

Ungeachtet der verschiedenen Abwehrsubstanzen haben alle drei Gruppen der Chrysomelina-Blattkäfer ein Abwehrsystem mit uniformer Architektur und Morphologie. Dieses System eignet sich hervorragend zur Erforschung, da sich all die verschiedenen enzymatischen Reaktionen stetigen evolutionären Veränderungen angepasst haben wie es von keinem anderen Organismus bekannt ist.

Bis vor wenigen Jahren gab es kaum genomische und transkriptomische Sequenzdaten von Chrysomela-Spezies. Da das Sequenzieren von vollständigen eukaryotischen Genomen weiterhin kostenintensiv und das Assemblieren solcher Genome hochkomplex (ohne eine zugrundeliegende Genomvorlage) ist, wurden von mehreren Blattkäferarten Transkript-Bibliotheken erstellt. Dies erfolgte durch das *de novo*-Assemblieren von paarweisen kurzen Sequenzen (Reads genannt), die in Zusammenarbeit mit dem Fritz-Lipmann-Institut Jena die Illumina-Technologie nutzend sequenziert wurden. Für das Assemblieren wurden mehrere verschiedene Assemblier-Programme angewandt und getestet. Die Software Trinity wurde gewählt, um brauchbare und geeignete Transkript-Banken von Blattkäfern, vor allem von *P. cochleariae* und dem Pappelblattkäfer *Chrysomela populi*, zu erstellen. Zusätzlich zum Etablieren der Transkript-Bibliotheken, habe ich eine Nicht-Ziel-Vorhersage (off-target prediction) zur effektiven Anwendung von RNA-Interferenz (RNAi) in den Blattkäfer(-larven) entwickelt. Diese Vorhersage soll ermöglichen, dass die dsRNA, die in den Organismus injiziert wird, einzig zu der RNA passt, die abgebaut werden soll, und off-target-Effekte ausbleiben bzw. vermieden werden. Diese Herangehensweise erlaubt das gezielte Untersuchen und die funktionelle Analyse von

Enzymen, Transport- oder Membranproteinen oder anderer Proteine von Interesse in geeigneter Zeit und mit geeignetem Aufwand.

In dieser Arbeit wurde die RNA-Sequenzierung mit der RNAi-Methode kombiniert, um Enzyme des Sequestrierungsprozesses bei verschiedenen Blattkäferlarven zu untersuchen. Dazu wurde zuerst die totale RNA von verschiedenen Geweben der Blattkäferlarven von *P. cochleariae* und *C. populi* extrahiert, präpariert und schließlich sequenziert. Dies resultierte in mehreren zehntausend kurzen Reads, die anschließend mit ihrer entsprechenden Transkriptom-Datenbank verglichen wurden. Diese Daten wurden genutzt, um die „Counts" eines jeden vorhandenen Transkriptes zu bestimmen. Diese Schritte wurden auch ausgeführt, nachdem die Larven mit dsRNA von Proteinen aus dem Sequestrierungsprozess injiziert wurden. Die „Counts" dieser Experimente wurden mit den Kontrolldaten verglichen, um statistisch signifikant differentielle „Expression" zu bestimmen.

Die oben erwähnte Vorhersage von off-targets für RNAi-Experimente wurde in Manuskript 1 veröffentlicht und für zwei Enzyme, eine Salicylalkohol-Oxidase in *C. populi*-Larven und ein Protein, das zur Juvenilhormon-Bindeprotein-Superfamilie gehört und in den Abwehrdrüsen von *P. cochleariae*-Larven identifiziert wurde, erfolgreich gezeigt. RNAi-Experimente, die die Zuckertransporter inhibieren, die am höchsten in den Abwehrdrüsen von *P. cochleariae*-Larven vorhanden sind (Manuskript 2), resultierten in einer geringen Abnahme in der Menge der Abwehrsubstanz. Nachfolgende Analyse der vorhandenen RNA mittels RNA-Sequenzierung zeigte eine Gegenregulation von anderen Zuckertransportern. Diese Gegenregulation wird wohl benötigt, um die Fähigkeit, sich zu verteidigen, und die Nährstoffversorgung in den Larven des Meerrettichblattkäfers aufrecht zu erhalten.

In Manuskript 4 wurde mittels RNAi ein ABC Transporter in *C. populi* inhibiert, der mit Abstand die höchste Transkriptmenge aufweist. Dies führte zu einer immensen Abnahme der Transkriptmenge dieses Transporters. Zusätzlich wurde keine Gegenregulation eines anderen ABC Transporters beobachtet. Dieses Ergebnis verdeutlicht die wichtige Stellung dieses spezifischen Transporters im Sequestrierungsprozesses (Strauss et al., 2013). Viele weitere Sequenzen wurden in diesem RNAi Experiment differentiell exprimiert, doch die Auswertung und Annotation dieser verbleibt eine offene Aufgabe.

Phylogenetische Analysen, die zusätzlich ausgeführt wurden, ergaben weitere Informationen und Annotationen zu ausgewählten Transkripten, vor allem durch den direkten Vergleich mit schon funktionell charakterisierten Enzymen und Transportern. So konnten die Sequenzen jeweils mit ihren homologen charakterisierten Sequenzen gruppiert werden und putative Funktionen oder die Substratspezifität für die ausgewählten Sequenzen der Blattkäfer zugeordnet bzw. vorgeschlagen werden. Wenn man dabei die Wichtigkeit der riesigen Menge an Zucker- und ABC-Transporter mit ihrer diversen Funktionalität betrachtet, so können all die genannten Methoden zu fundiertem Wissen und neuen Erkenntnissen über bekannte und

unbekannte regulatorische Prozesse und Stoffwechselwege führen. Weiterhin konnte ich für eine kleine Gruppe an Zuckertransport-Proteinen in *P. cochleariae* zeigen, dass sie käferspezifisch sind und nicht, zum Beispiel, im Menschen vorkommen.

Zusammenfassend erkennt man, dass Sequenziermethoden der zweiten (oder dritten) Generation (oder next-generation sequencing) mit besonderer Hinsicht auf das Sequenzieren von RNA, vor allem auch das Sequenzieren von RNA nach dem Ausführen von RNAi Experimenten, wunderbar genutzt werden können, um Gene und Transkripte von Interesse zu erforschen. Diese Herangehensweise ist insbesondere für Nicht-Modellorganismen von großer Bedeutung, um deren Transkriptome in ihrer ganzen Breite zu untersuchen, und ermöglicht, transkriptomweite Analysen durchzuführen. Dies konnte in dieser Arbeit eindrucksvoll anhand des Merrettichblattkäfers *P. cochleariae* und des Pappelblattkäfers *C. populi* gezeigt werden.

BIBLIOGRAPHY

ABD-ALLA, A. M., PARKER, A. G., VREYSEN, M. J. & BERGOIN, M. 2011. Tsetse salivary gland hypertrophy virus: hope or hindrance for tsetse control? *PLoS Negl Trop Dis,* 5**,** e1220.

AIGNER, A. 2007. Applications of RNA interference: current state and prospects for siRNA-based strategies in vivo. *Appl Microbiol Biotechnol,* 76**,** 9-21.

ALVES, A. P., LORENZEN, M. D., BEEMAN, R. W., FOSTER, J. E. & SIEGFRIED, B. D. 2010. RNA interference as a method for target-site screening in the western corn rootworm, Diabrotica virgifera virgifera. *Journal of Insect Science,* 10.

AMARZGUIOUI, M. & PRYDZ, H. 2004. An algorithm for selection of functional siRNA sequences. *Biochem Biophys Res Commun,* 316**,** 1050-8.

ANDERS, S. & HUBER, W. 2010. Differential expression analysis for sequence count data. *Genome Biol,* 11**,** R106.

ANDERS, S., MCCARTHY, D. J., CHEN, Y., OKONIEWSKI, M., SMYTH, G. K., HUBER, W. & ROBINSON, M. D. 2013. Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat Protoc,* 8**,** 1765-86.

ANNILO, T., CHEN, Z. Q., SHULENIN, S., COSTANTINO, J., THOMAS, L., LOU, H., STEFANOV, S. & DEAN, M. 2006. Evolution of the vertebrate ABC gene family: analysis of gene birth and death. *Genomics,* 88**,** 1-11.

ANSORGE, W. J. 2009. Next-generation DNA sequencing techniques. *N Biotechnol,* 25**,** 195-203.

ARZIMAN, Z., HORN, T. & BOUTROS, M. 2005. E-RNAi: a web application to design optimized RNAi constructs. *Nucleic Acids Res,* 33**,** W582-8.

AUGUSTIN, R. 2010. The protein family of glucose transport facilitators: It's not only about glucose after all. *Iubmb Life,* 62**,** 315-333.

BAI, H., ZHU, F., SHAH, K. & PALLI, S. R. 2011. Large-scale RNAi screen of G protein-coupled receptors involved in larval growth, molting and metamorphosis in the red flour beetle. *BMC Genomics,* 12**,** 388.

BAUM, J. A., BOGAERT, T., CLINTON, W., HECK, G. R., FELDMANN, P., ILAGAN, O., JOHNSON, S., PLAETINCK, G., MUNYIKWA, T., PLEAU, M., VAUGHN, T. & ROBERTS, J. 2007. Control of coleopteran insect pests through RNA interference. *Nature Biotechnology,* 25**,** 1322-1326.

BELLES, X. 2010. Beyond Drosophila: RNAi In Vivo and Functional Genomics in Insects. *Annual Review of Entomology.* Palo Alto: Annual Reviews.

BENNETT, S. 2004. Solexa Ltd. *Pharmacogenomics,* 5**,** 433-8.

BIEMANS-OLDEHINKEL, E., DOEVEN, M. K. & POOLMAN, B. 2006. ABC transporter architecture and regulatory roles of accessory domains. *FEBS Lett,* 580**,** 1023-35.

BLUM, M. S., BRAND, J. M., WALLACE, J. B. & FALES, H. M. 1972. Chemical characterisation of devensive secretions of a *chrysomelid* larva. *Life Sciences Pt-2 Biochemistry General and Molecular Biology,* 11**,** 525-531.

BLUM, M. S., WALLACE, J. B., DUFFIELD, R. M., BRAND, J. M., FALES, H. M. & SOKOLOSKI, E. A. 1978. Chrysomelidial in the defensive secretion of the leaf beetle *Gastrophysa cyanea* Chrysomelidial in the defensive secretion of the leaf beetle *Gastrophysa cyanea*. *Journal of Chemical Ecology,* 4**,** 47-53.

BRASLAVSKY, I., HEBERT, B., KARTALOV, E. & QUAKE, S. R. 2003. Sequence information can be obtained from single DNA molecules. *Proc Natl Acad Sci U S A,* 100**,** 3960-4.

BROEHAN, G., KROEGER, T., LORENZEN, M. & MERZENDORFER, H. 2013. Functional analysis of the ATP-binding cassette (ABC) transporter gene family of Tribolium castaneum. *Bmc Genomics,* 14.

BULLARD, J. H., PURDOM, E., HANSEN, K. D. & DUDOIT, S. 2010. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics,* 11**,** 94.

BURAND, J. P. & HUNTER, W. B. 2013. RNAi: future in insect management. *J Invertebr Pathol,* 112 Suppl**,** S68-74.

BURSE, A., FRICK, S., DISCHER, S., TOLZIN-BANASCH, K., KIRSCH, R., STRAUSS, A., KUNERT, M. & BOLAND, W. 2009. Always being well prepared for defense: The production of deterrents by juvenile Chrysomelina beetles (Chrysomelidae). *Phytochemistry,* 70**,** 1899-1909.

BURSE, A., SCHMIDT, A., FRICK, S., KUHN, J., GERSHENZON, J. & BOLAND, W. 2007. Iridoid biosynthesis in Chrysomelina larvae: Fat body produces early terpenoid precursors. *Insect Biochemistry and Molecular Biology,* 37**,** 255-265.

CHALK, A. M., WAHLESTEDT, C. & SONNHAMMER, E. L. 2004. Improved and automated prediction of effective siRNA. *Biochem Biophys Res Commun,* 319**,** 264-74.

CHAUDHARY, A., SRIVASTAVA, S. & GARG, S. 2011. Development of a software tool and criteria evaluation for efficient design of small interfering RNA. *Biochem Biophys Res Commun,* 404**,** 313-20.

CHEN, G., WANG, C. & SHI, T. 2011. Overview of available methods for diverse RNA-Seq data analyses. *Sci China Life Sci,* 54**,** 1121-8.

CLOONAN, N., FORREST, A. R., KOLLE, G., GARDINER, B. B., FAULKNER, G. J., BROWN, M. K., TAYLOR, D. F., STEPTOE, A. L., WANI, S., BETHEL, G., ROBERTSON, A. J., PERKINS, A. C., BRUCE, S. J., LEE, C. C., RANADE, S. S., PECKHAM, H. E., MANNING, J. M., MCKERNAN, K. J. & GRIMMOND, S. M. 2008. Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nat Methods,* 5**,** 613-9.

COSTA, V., ANGELINI, C., DE FEIS, I. & CICCODICOLA, A. 2010. Uncovering the complexity of transcriptomes with RNA-Seq. *J Biomed Biotechnol,* 2010**,** 853916.

CRAWFORD, J. E., GUELBEOGO, W. M., SANOU, A., TRAORE, A., VERNICK, K. D., SAGNON, N. & LAZZARO, B. P. 2010. De novo transcriptome sequencing in Anopheles funestus using Illumina RNA-seq technology. *PLoS One,* 5**,** e14202.

CULLEN, B. R. 2006. Enhancing and confirming the specificity of RNAi experiments. *Nat Methods,* 3**,** 677-81.

CULLUM, R., ALDER, O. & HOODLESS, P. A. 2011. The next generation: using new sequencing technologies to analyse gene regulation. *Respirology,* 16**,** 210-22.

DASSA, E. 2011. Natural history of ABC systems: not only transporters. *In:* SHAROM, F. J. (ed.) *Essays in Biochemistry: Abc Transporters.*

DEAN, M. & ALLIKMETS, R. 1995. Evolution of ATP-binding cassette transporter genes. *Curr Opin Genet Dev,* 5**,** 779-85.

DEAN, M. & ANNILO, T. 2005. Evolution of the ATP-binding cassette (ABC) transporter superfamily in vertebrates. *Annual Review of Genomics and Human Genetics.*

DEAN, M., RZHETSKY, A. & ALLIKMETS, R. 2001. The human ATP-binding cassette (ABC) transporter superfamily. *Genome Research,* 11**,** 1156-1166.

DILLIES, M. A., RAU, A., AUBERT, J., HENNEQUET-ANTIER, C., JEANMOUGIN, M., SERVANT, N., KEIME, C., MAROT, G., CASTEL, D., ESTELLE, J., GUERNEC, G., JAGLA, B., JOUNEAU, L., LALOE, D., LE GALL, C., SCHAEFFER, B., LE CROM, S., GUEDJ, M., JAFFREZIC, F. & ON BEHALF OF THE FRENCH STATOMIQUE, C. 2012. A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Brief Bioinform*.

DISCHER, S., BURSE, A., TOLZIN-BANASCH, K., HEINEMANN, S. H., PASTEELS, J. M. & BOLAND, W. 2009. A versatile transport network for sequestering and excreting plant glycosides in leaf beetles provides an evolutionary flexible defense strategy. *Chembiochem,* 10**,** 2223-2229.

DOUADY, C. J., DELSUC, F., BOUCHER, Y., DOOLITTLE, W. F. & DOUZERY, E. J. P. 2003. Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. *Mol Biol Evol,* 20**,** 248-254.

DROEGE, M. & HILL, B. 2008. The Genome Sequencer FLX System--longer reads, more applications, straight forward bioinformatics and more complete data sets. *J Biotechnol,* 136**,** 3-10.

EID, J., FEHR, A., GRAY, J., LUONG, K., LYLE, J., OTTO, G., PELUSO, P., RANK, D., BAYBAYAN, P., BETTMAN, B., BIBILLO, A., BJORNSON, K., CHAUDHURI, B., CHRISTIANS, F., CICERO, R., CLARK, S., DALAL, R., DEWINTER, A., DIXON, J., FOQUET, M., GAERTNER, A., HARDENBOL, P., HEINER, C., HESTER, K., HOLDEN, D., KEARNS, G., KONG, X., KUSE, R., LACROIX, Y., LIN, S., LUNDQUIST, P., MA, C., MARKS, P., MAXHAM, M., MURPHY, D., PARK, I., PHAM, T., PHILLIPS, M., ROY, J., SEBRA, R., SHEN, G., SORENSON, J., TOMANEY, A., TRAVERS, K., TRULSON, M., VIECELI, J., WEGENER, J., WU, D., YANG, A., ZACCARIN, D., ZHAO, P., ZHONG, F., KORLACH, J. & TURNER, S. 2009. Real-time DNA sequencing from single polymerase molecules. *Science,* 323**,** 133-8.

FELD, B. K., PASTEELS, J. M. & BOLAND, W. 2001. *Phaedon cochleariae* and *Gastrophysa viridula* (Coleoptera : Chrysomelidae) produce defensive iridoid monoterpenes *de novo* and are able to sequester glycosidically bound terpenoid precursors. *Chemoecology,* 11**,** 191-198.

FENG, L., LIU, H., LIU, Y., LU, Z., GUO, G., GUO, S., ZHENG, H., GAO, Y., CHENG, S., WANG, J., ZHANG, K. & ZHANG, Y. 2010. Power of deep sequencing and agilent microarray for gene expression profiling study. *Mol Biotechnol,* 45**,** 101-10.

FRANCIS, W. R., CHRISTIANSON, L. M., KIKO, R., POWERS, M. L., SHANER, N. C. & SH, D. H. 2013. A comparison across non-model animals suggests an optimal sequencing depth for de novo transcriptome assembly. *BMC Genomics,* 14**,** 167.

FREDRIKSSON, R., NORDSTROM, K. J., STEPHANSSON, O., HAGGLUND, M. G. & SCHIOTH, H. B. 2008. The solute carrier (SLC) complement of the human genome: phylogenetic classification reveals four major families. *FEBS Lett,* 582**,** 3811-6.

FRICK, S., NAGEL, R., SCHMIDT, A., BODEMANN, R. R., RAHFELD, P., PAULS, G., BRANDT, W., GERSHENZON, J., BOLAND, W. & BURSE, A. 2013. Metal ions control product specificity of isoprenyl diphosphate synthases in the insect terpenoid pathway. *Proc Natl Acad Sci U S A,* 110**,** 4194-9.

GLUSMAN, G., CABALLERO, J., ROBINSON, M., KUTLU, B. & HOOD, L. 2013. Optimal scaling of digital transcriptomes. *PLoS One,* 8**,** e77885.

GONG, W., REN, Y., ZHOU, H., WANG, Y., KANG, S. & LI, T. 2008. siDRM: an effective and generally applicable online siRNA design tool. *Bioinformatics,* 24**,** 2405-6.

GONGORA-CASTILLO, E. & BUELL, C. R. 2013. Bioinformatics challenges in de novo transcriptome assembly using short read sequences in the absence of a reference genome sequence. *Nat Prod Rep,* 30**,** 490-500.

GRABHERR, M. G., HAAS, B. J., YASSOUR, M., LEVIN, J. Z., THOMPSON, D. A., AMIT, I., ADICONIS, X., FAN, L., RAYCHOWDHURY, R., ZENG, Q., CHEN, Z., MAUCELI, E., HACOHEN, N., GNIRKE, A., RHIND, N., DI PALMA, F., BIRREN, B. W., NUSBAUM, C., LINDBLAD-TOH, K., FRIEDMAN, N. & REGEV, A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol,* 29**,** 644-652.

GREMSKI, L. H., DA SILVEIRA, R. B., CHAIM, O. M., PROBST, C. M., FERRER, V. P., NOWATZKI, J., WEINSCHUTZ, H. C., MADEIRA, H. M., GREMSKI, W., NADER, H. B., SENFF-RIBEIRO, A. & VEIGA, S. S. 2010. A novel expression profile of the Loxosceles intermedia spider venomous gland revealed by transcriptome analysis. *Mol Biosyst,* 6**,** 2403-16.

HAAS, B. J., PAPANICOLAOU, A., YASSOUR, M., GRABHERR, M., BLOOD, P. D., BOWDEN, J., COUGER, M. B., ECCLES, D., LI, B., LIEBER, M., MACMANES, M. D., OTT, M., ORVIS, J., POCHET, N., STROZZI, F., WEEKS, N., WESTERMAN, R., WILLIAM, T., DEWEY, C. N., HENSCHEL, R., LEDUC, R. D., FRIEDMAN, N. & REGEV, A. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc,* 8**,** 1494-512.

HANSEN, K. D., WU, Z., IRIZARRY, R. A. & LEEK, J. T. 2011. Sequencing technology does not eliminate biological variability. *Nat Biotechnol,* 29**,** 572-3.

HART, C., LIPSON, D., OZSOLAK, F., RAZ, T., STEINMANN, K., THOMPSON, J. & MILOS, P. M. 2010. Single-molecule sequencing: sequence methods to enable accurate quantitation. *Methods Enzymol,* 472**,** 407-30.

HENSCHEL, A., BUCHHOLZ, F. & HABERMANN, B. 2004. DEQOR: a web-based tool for the design and quality control of siRNAs. *Nucleic Acids Res,* 32**,** W113-20.

HILKER, M. & SCHULZ, S. 1994. Composition of larval secretion of *Chrysomela lapponica* (Coleoptera, Chrysomelidae) and its dependence on host-plant. *Journal of Chemical Ecology,* 20**,** 1075-1093.

HINTON, H. E. 1951. On a little-known protective device of some Chrysomelid pupae (Coleoptera). *Proceedings of the Royal Entomological Society of London (A),* 26**,** 67-73.

HOGLUND, P. J., NORDSTROM, K. J., SCHIOTH, H. B. & FREDRIKSSON, R. 2011. The solute carrier families have a remarkably long evolutionary history with the majority of the human families present before divergence of Bilaterian species. *Mol Biol Evol,* 28**,** 1531-41.

HOLLENSTEIN, K., DAWSON, R. J. P. & LOCHER, K. P. 2007. Structure and mechanism of ABC transporter proteins. *Current Opinion in Structural Biology,* 17**,** 412-418.

HORNER, D. S., PAVESI, G., CASTRIGNANO, T., DE MEO, P. D., LIUNI, S., SAMMETH, M., PICARDI, E. & PESOLE, G. 2010. Bioinformatics approaches for genomics and post genomics applications of next-generation sequencing. *Briefings in Bioinformatics,* 11**,** 181-197.

HOWARD, B. E. & HEBER, S. 2010. Towards reliable isoform quantification using RNA-SEQ data. *BMC Bioinformatics,* 11 Suppl 3**,** S6.

HUELSENBECK, J. P. & RONQUIST, F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics,* 17**,** 754-5.

HUELSENBECK, J. P., RONQUIST, F., NIELSEN, R. & BOLLBACK, J. P. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. *Science,* 294**,** 2310-4.

HUNTER, W., ELLIS, J., VANENGELSDORP, D., HAYES, J., WESTERVELT, D., GLICK, E., WILLIAMS, M., SELA, I., MAORI, E., PETTIS, J., COX-FOSTER, D. & PALDI, N. 2010. Large-scale field application of RNAi technology reducing Israeli acute paralysis virus disease in honey bees (Apis mellifera, Hymenoptera: Apidae). *PLoS Pathog,* 6**,** e1001160.

HUVENNE, H. & SMAGGHE, G. 2010. Mechanisms of dsRNA uptake in insects and potential of RNAi for pest control: A review. *Journal of Insect Physiology,* 56**,** 227-235.

IYER, S., DEUTSCH, K., YAN, X. W. & LIN, B. Y. 2007. Batch RNAi selector: A standalone program to predict specific siRNA candidates in batches with enhanced sensitivity. *Computer Methods and Programs in Biomedicine,* 85**,** 203-209.

JACKSON, A. L., BARTZ, S. R., SCHELTER, J., KOBAYASHI, S. V., BURCHARD, J., MAO, M., LI, B., CAVET, G. & LINSLEY, P. S. 2003. Expression profiling reveals off-target gene regulation by RNAi. *Nature Biotechnology,* 21**,** 635-637.

JACKSON, A. L., BURCHARD, J., SCHELTER, J., CHAU, B. N., CLEARY, M., LIM, L. & LINSLEY, P. S. 2006. Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA,* 12**,** 1179-87.

JOOST, H. G., BELL, G. I., BEST, J. D., BIRNBAUM, M. J., CHARRON, M. J., CHEN, Y. T., DOEGE, H., JAMES, D. E., LODISH, H. F., MOLEY, K. H., MOLEY, J. F., MUECKLER, M., ROGERS, S., SCHURMANN, A., SEINO, S. & THORENS, B. 2002. Nomenclature of the GLUT/SLC2A family of sugar/polyol transport facilitators. *Am J Physiol Endocrinol Metab,* 282**,** E974-6.

JOOST, H. G. & THORENS, B. 2001. The extended GLUT-family of sugar/polyol transport facilitators: nomenclature, sequence characteristics, and potential function of its novel members. *Molecular Membrane Biology,* 18**,** 247-256.

KANAMORI, Y., SAITO, A., HAGIWARA-KOMODA, Y., TANAKA, D., MITSUMASU, K., KIKUTA, S., WATANABE, M., CORNETTE, R., KIKAWADA, T. & OKUDA, T. 2010. The trehalose transporter 1 gene sequence is conserved in insects and encodes proteins with different kinetic properties involved in trehalose import into peripheral tissues. *Insect Biochemistry and Molecular Biology,* 40**,** 30-37.

KEELING, C. I., HENDERSON, H., LI, M., YUEN, M., CLARK, E. L., FRASER, J. D., HUBER, D. P., LIAO, N. Y., DOCKING, T. R., BIROL, I., CHAN, S. K., TAYLOR, G. A., PALMQUIST, D., JONES, S. J. & BOHLMANN, J. 2012. Transcriptome and full-length cDNA resources for the mountain pine beetle, *Dendroctonus ponderosae* Hopkins, a major insect pest of pine forests. *Insect Biochem Mol Biol,* 42**,** 525-36.

KIKAWADA, T., SAITO, A., KANAMORI, Y., NAKAHARA, Y., ICHI IWATA, K., TANAKA, D., WATANABE, M. & OKUDA, T. 2007. Trehalose transporter 1, a facilitated and high-capacity trehalose transporter, allows exogenous trehalose uptake into cells. *Proc Natl Acad Sci U S A,* 104**,** 11585-11590.

KIKUTA, S., HAGIWARA-KOMODA, Y., NODA, H. & KIKAWADA, T. 2012. A novel member of the trehalose transporter family functions as an H^+-dependent trehalose transporter in the reabsorption of trehalose in Malpighian tubules. *Frontiers in Physiology,* 3**,** 290-290.

KIKUTA, S., KIKAWADA, T., HAGIWARA-KOMODA, Y., NAKASHIMA, N. & NODA, H. 2010. Sugar transporter genes of the brown planthopper, *Nilaparvata lugens*: A facilitated glucose/fructose transporter. *Insect Biochemistry and Molecular Biology,* 40**,** 805-813.

KIM, J. H., ROY, A., JOUANDOT, D. & CHO, K. H. 2013. The glucose signaling network in yeast. *Biochimica Et Biophysica Acta-General Subjects,* 1830**,** 5204-5210.

KIRSCH, R., VOGEL, H., MUCK, A., REICHWALD, K., PASTEELS, J. M. & BOLAND, W. 2011a. Host plant shifts affect a major defense enzyme in Chrysomela lapponica. *Proceedings of the National Academy of Sciences of the United States of America,* 108**,** 4897-4901.

KIRSCH, R., VOGEL, H., MUCK, A., VILCINSKAS, A., PASTEELS, J. M. & BOLAND, W. 2011b. To be or not to be convergent in salicin-based defence in chrysomeline leaf beetle larvae: evidence from *Phratora vitellinae* salicyl alcohol oxidase. *Proc Biol Sci,* 278**,** 3225-32.

KOBOLDT, D. C., STEINBERG, K. M., LARSON, D. E., WILSON, R. K. & MARDIS, E. R. 2013. The next-generation sequencing revolution and its impact on genomics. *Cell,* 155**,** 27-38.

KUHN, J., PETTERSSON, E. M., FELD, B. K., BURSE, A., TERMONIA, A., PASTEELS, J. M. & BOLAND, W. 2004. Selective transport systems mediate sequestration of plant glucosides in leaf beetles: A molecular basis for adaptation and evolution. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 13808-13813.

KUHN, J., PETTERSSON, E. M., FELD, B. K., NIE, L., TOLZIN-BANASCH, K., M'RABET, S. M., PASTEELS, J. & BOLAND, W. 2007. Sequestration of plant-derived phenolglucosides by larvae of the leaf beetle Chrysomela lapponica: Thioglucosides as mechanistic probes. *Journal of Chemical Ecology,* 33**,** 5-24.

KULKARNI, M. M., BOOKER, M., SILVER, S. J., FRIEDMAN, A., HONG, P., PERRIMON, N. & MATHEY-PREVOT, B. 2006. Evidence of off-target effects associated with long dsRNAs in Drosophila melanogaster cell-based assays. *Nat Methods,* 3**,** 833-8.

KUNERT, M., SOE, A., BARTRAM, S., DISCHER, S., TOLZIN-BANASCH, K., NIE, L., DAVID, A., PASTEELS, J. & BOLAND, W. 2008. De novo biosynthesis versus sequestration: A network of transport systems supports in iridoid producing leaf beetle larvae both modes of defense. *Insect Biochemistry and Molecular Biology,* 38**,** 895-904.

LEVY, S., SUTTON, G., NG, P. C., FEUK, L., HALPERN, A. L., WALENZ, B. P., AXELROD, N., HUANG, J., KIRKNESS, E. F., DENISOV, G., LIN, Y., MACDONALD, J. R., PANG, A. W., SHAGO, M., STOCKWELL, T. B., TSIAMOURI, A., BAFNA, V., BANSAL, V., KRAVITZ, S. A., BUSAM, D. A., BEESON, K. Y., MCINTOSH, T. C., REMINGTON, K. A., ABRIL, J. F., GILL, J., BORMAN, J., ROGERS, Y. H., FRAZIER, M. E., SCHERER, S. W., STRAUSBERG, R. L. & VENTER, J. C. 2007. The diploid genome sequence of an individual human. *PLoS Biol,* 5**,** e254.

LI, B., RUOTTI, V., STEWART, R. M., THOMSON, J. A. & DEWEY, C. N. 2010. RNA-Seq gene expression estimation with read mapping uncertainty. *Bioinformatics,* 26**,** 493-500.

LIANG, C., LIU, X., YIU, S. M. & LIM, B. L. 2013. De novo assembly and characterization of Camelina sativa transcriptome by paired-end sequencing. *BMC Genomics,* 14**,** 146.

LIU, S. M., ZHOU, S., TIAN, L., GUO, E. N., LUAN, Y. X., ZHANG, J. Z. & LI, S. 2011. Genome-wide identification and characterization of ATP-binding cassette transporters in the silkworm, Bombyx mori. *BMC Genomics,* 12.

LIU, Y., ZHOU, J. & WHITE, K. P. 2013. RNA-seq differential expression studies: more sequence, or more replication? *Bioinformatics*.

LORENZ, M., BOLAND, W. & DETTNER, K. 1993. Biosynthesis of iridodials in the defense glands of beetle larvae (Chrysomelinae). *Angewandte Chemie-International Edition in English,* 32**,** 912-914.

LU, B., ZENG, Z. & SHI, T. 2013. Comparative study of de novo assembly and genome-guided assembly strategies for transcriptome reconstruction based on RNA-Seq. *Sci China Life Sci,* 56**,** 143-55.

LUO, Q., KANG, Q., SONG, W. X., LUU, H. H., LUO, X., AN, N., LUO, J., DENG, Z. L., JIANG, W., YIN, H., CHEN, J., SHARFF, K. A., TANG, N., BENNETT, E., HAYDON, R. C. & HE, T. C. 2007. Selection and validation of optimal siRNA target sites for RNAi-mediated gene silencing. *Gene,* 395**,** 160-9.

MACHEDA, M. L., ROGERS, S. & BEST, J. D. 2005. Molecular and cellular regulation of glucose transporter (GLUT) proteins in cancer. *J Cell Physiol,* 202**,** 654-62.

MACMANES, M. D. & EISEN, M. B. 2013. Improving transcriptome assembly through error correction of high-throughput sequence reads. *PeerJ,* 1**,** e113.

MALONE, J. H. & OLIVER, B. 2011. Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biol,* 9**,** 34.

MARCH, J. C. & BENTLEY, W. E. 2007. Methods for gene silencing with RNAi. *Methods Mol Biol,* 388**,** 427-34.

MARDIS, E. R. 2008. The impact of next-generation sequencing technology on genetics. *Trends Genet,* 24**,** 133-41.

MARDIS, E. R. 2012. Applying next-generation sequencing to pancreatic cancer treatment. *Nat Rev Gastroenterol Hepatol,* 9**,** 477-86.

MARGUERAT, S. & BAHLER, J. 2010. RNA-seq: from technology to biology. *Cell Mol Life Sci,* 67**,** 569-79.

MARIONI, J. C., MASON, C. E., MANE, S. M., STEPHENS, M. & GILAD, Y. 2008. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res,* 18**,** 1509-17.

MARTIN, J. A. & WANG, Z. 2011. Next-generation transcriptome assembly. *Nat Rev Genet,* 12**,** 671-82.

MCGETTIGAN, P. A. 2013. Transcriptomics in the RNA-seq era. *Curr Opin Chem Biol,* 17**,** 4-11.

MEINWALD, J., JONES, T. H., EISNER, T. & HICKS, K. 1977. Defense-mechanisms of arthropods .56. New methylcyclopentanoid terpenes from larval defensive secretion of a chrysomelid beetle (*Plagiodera versicolora*). *Proceedings of the National Academy of Sciences of the United States of America,* 74**,** 2189-2193.

MITO, T., NAKAMURA, T., BANDO, T., OHUCHI, H. & NOJI, S. 2011. The advent of RNA interference in Entomology. *Entomological Science,* 14**,** 1-8.

MITTAPALLI, O., BAI, X., MAMIDALA, P., RAJARAPU, S. P., BONELLO, P. & HERMS, D. A. 2010. Tissue-specific transcriptomics of the exotic invasive insect pest emerald ash borer (Agrilus planipennis). *PLoS One,* 5**,** e13708.

MOFFAT, J., REILING, J. H. & SABATINI, D. M. 2007. Off-target effects associated with long dsRNAs in Drosophila RNAi screens. *Trends Pharmacol Sci,* 28**,** 149-51.

MOITRA, K. & DEAN, M. 2011. Evolution of ABC transporters by gene duplication and their role in human disease. *Biological Chemistry,* 392**,** 29-37.

MORTAZAVI, A., WILLIAMS, B. A., MCCUE, K., SCHAEFFER, L. & WOLD, B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods,* 5**,** 621-8.

MUECKLER, M. & THORENS, B. 2013. The SLC2 (GLUT) family of membrane transporters. *Molecular Aspects of Medicine,* 34**,** 121-138.

MUTZ, K. O., HEILKENBRINKER, A., LONNE, M., WALTER, J. G. & STAHL, F. 2013. Transcriptome analysis using next-generation sequencing. *Curr Opin Biotechnol,* 24**,** 22-30.

NAITO, Y., YAMADA, T., MATSUMIYA, T., UI-TEI, K., SAIGO, K. & MORISHITA, S. 2005. dsCheck: highly sensitive off-target search software for double-stranded RNA-mediated RNA interference. *Nucleic Acids Research,* 33**,** W589-W591.

NAITO, Y., YOSHIMURA, J., MORISHITA, S. & UI-TEI, K. 2009. siDirect 2.0: updated software for designing functional siRNA with reduced seed-dependent off-target effect. *BMC Bioinformatics,* 10**,** 392.

NAKASUGI, K., CROWHURST, R. N., BALLY, J., WOOD, C. C., HELLENS, R. P. & WATERHOUSE, P. M. 2013. De novo transcriptome sequence assembly and analysis of RNA silencing genes of Nicotiana benthamiana. *PLoS One,* 8**,** e59534.

NIU, J.-Z., DOU, W., DING, T.-B., SHEN, G.-M., ZHANG, K., SMAGGHE, G. & WANG, J.-J. 2012. Transcriptome analysis of the citrus red mite, Panonychus citri, and its gene expression by exposure to insecticide/acaricide. *{INSECT MOLECULAR BIOLOGY},* {21}**,** {422-436}.

NOVAES, E., DROST, D. R., FARMERIE, W. G., PAPPAS, G. J., GRATTAPAGLIA, D., SEDEROFF, R. R. & KIRST, M. 2008. High-throughput gene and SNP discovery in Eucalyptus grandis, an uncharacterized genome. *BMC Genomics,* 9**,** 312.

OHNISHI, A., HASHIMOTO, K., IMAI, K. & MATSUMOTO, S. 2009. Functional characterization of the *Bombyx mori* fatty acid transport protein (BmFATP) within the silkmoth pheromone gland. *Journal of Biological Chemistry,* 284**,** 5128-5136.

OHNISHI, A., HULL, J. J. & MATSUMOTO, S. 2006. Targeted disruption of genes in the *Bombyx mori* sex pheromone biosynthetic pathway. *Proceedings of the National Academy of Sciences of the United States of America,* 103**,** 4398-4403.

OLDHAM, M. L., DAVIDSON, A. L. & CHEN, J. 2008. Structural insights into ABC transporter mechanism. *Curr Opin Struct Biol,* 18**,** 726-33.

OSHLACK, A., ROBINSON, M. D. & YOUNG, M. D. 2010. From RNA-seq reads to differential expression results. *Genome Biol,* 11**,** 220.

OZSOLAK, F. & MILOS, P. M. 2011a. Single-molecule direct RNA sequencing without cDNA synthesis. *Wiley Interdiscip Rev RNA,* 2**,** 565-70.

OZSOLAK, F. & MILOS, P. M. 2011b. Transcriptome profiling using single-molecule direct RNA sequencing. *Methods Mol Biol,* 733**,** 51-61.

PASTEELS, J. M., BRAEKMAN, J. C., DALOZE, D. & OTTINGER, R. 1982. Chemical defence in chrysomelid larvae and adults. *Tetrahedron,* 38**,** 1891-1897.

PASTEELS, J. M., DUFFEY, S. & ROWELL-RAHIER, M. 1990. Toxins in chrysomelid beetles possible evolutionary sequence from *de novo* synthesis to derivation from food-plant chemicals. *Journal of Chemical Ecology,* 16**,** 211-222.

PASTEELS, J. M., GREGOIRE, J. C. & ROWELL-RAHIER, M. 1983a. The chemical ecology of defense in arthropods. *Annual Review of Entomology,* 28**,** 263-289.

PASTEELS, J. M., ROWELL-RAHIER, M., BRAEKMAN, J. C. & DUPONT, A. 1983b. Salicin from host plant as precursor of salicyl aldehyde in defensive secretion of chrysomeline larvae. *Physiological Entomology,* 8**,** 307-314.

PENG, Y., LEUNG, H. C. M., YIU, S. M. & CHIN, F. Y. L. 2011. T-IDBA: A de novo Iterative de Bruijn Graph Assembler for Transcriptome. *Research in Computational Molecular Biology,* 6577**,** 337-338.

PETZOLD, A., REICHWALD, K., GROTH, M., TAUDIEN, S., HARTMANN, N., PRIEBE, S., SHAGIN, D., ENGLERT, C. & PLATZER, M. 2013. The transcript catalogue of the short-lived fish *Nothobranchius furzeri* provides insights into age-dependent changes of mRNA levels. *BMC Genomics,* 14**,** 185.

PORRECA, G. J., SHENDURE, J. & CHURCH, G. M. 2006. Polony DNA sequencing. *Curr Protoc Mol Biol,* Chapter 7**,** Unit 7 8.

PRICE, D. R. G., TIBBLES, K., SHIGENOBU, S., SMERTENKO, A., RUSSELL, C. W., DOUGLAS, A. E., FITCHES, E., GATEHOUSE, A. M. R. & GATEHOUSE, J. A. 2010. Sugar transporters of the major facilitator superfamily in aphids; from gene prediction to functional characterization. *Insect Molecular Biology,* 19**,** 97-112.

PRICE, D. R. G., WILKINSON, H. S. & GATEHOUSE, J. A. 2007. Functional expression and characterisation of a gut facilitative glucose transporter, NlHT1, from the phloem-feeding insect *Nilaparvata lugens* (rice brown planthopper). *Insect Biochem Mol Biol,* 37**,** 1138-1148.

RAPAPORT, F., KHANIN, R., LIANG, Y., PIRUN, M., KREK, A., ZUMBO, P., MASON, C. E., SOCCI, N. D. & BETEL, D. 2013. Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. *Genome Biol,* 14**,** R95.

REID, C. A. M. 1995. A cladistic analysis of subfamilial relationships in the Chrysomelidae sensu lato (Chrysomeloidea). *In:* PAKALUK, J. A. S., S.A. (ed.) *In Biology, phylogeny and classification of Coleoptera: papers celebrating the 80th birthday of Roy A. Crowson.* Muzeum i Instytut Zoologii PAN: Warszawa.

RENNER, K. 1970. Über die ausstülpbaren Hautblasen der Larven von *Gastroidea viridula* De Geer und ihre ökologische Bedeutung (Coleoptera: Chrysomelidae). *Beitäge zur Entomologie,* 20**,** 527-533.

REYNOLDS, A., LEAKE, D., BOESE, Q., SCARINGE, S., MARSHALL, W. S. & KHVOROVA, A. 2004. Rational siRNA design for RNA interference. *Nature Biotechnology,* 22**,** 326-330.

ROBERTSON, G., SCHEIN, J., CHIU, R., CORBETT, R., FIELD, M., JACKMAN, S. D., MUNGALL, K., LEE, S., OKADA, H. M., QIAN, J. Q., GRIFFITH, M., RAYMOND, A., THIESSEN, N., CEZARD, T., BUTTERFIELD, Y. S., NEWSOME, R., CHAN, S. K., SHE, R., VARHOL, R., KAMOH, B., PRABHU, A. L., TAM, A., ZHAO, Y., MOORE, R. A., HIRST, M., MARRA, M. A., JONES, S. J., HOODLESS, P. A. & BIROL, I. 2010. De novo assembly and analysis of RNA-seq data. *Nat Methods,* 7**,** 909-12.

ROBINSON, M. D. & OSHLACK, A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol,* 11**,** R25.

ROBINSON, M. D. & SMYTH, G. K. 2008. Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics,* 9**,** 321-32.

ROKAS, A. 2011. Phylogenetic analysis of protein sequence data using the Randomized Axelerated Maximum Likelihood (RAXML) Program. *Curr Protoc Mol Biol,* Chapter 19**,** Unit19 11.

RONQUIST, F., TESLENKO, M., VAN DER MARK, P., AYRES, D. L., DARLING, A., HÖHNA, S., LARGET, B., LIU, L., SUCHARD, M. A. & HUELSENBECK, J. P. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol,* 61**,** 539-542.

ROWELL-RAHIER, M. & PASTEELS, J. M. 1986. Economics of chemical defense in Chrysomelinae. *Journal of Chemical Ecology,* 12**,** 1189-1203.

SABIRZHANOV, B., SABIRZHANOVA, I. B. & KEIFER, J. 2011. Screening target specificity of siRNAs by Rapid Amplification of cDNA Ends (RACE) for non-sequenced species. *Journal of Molecular Neuroscience,* 44**,** 68-75.

SANGER, F., NICKLEN, S. & COULSON, A. R. 1977. DNA Sequencing with Chain-Terminating Inhibitors. *Proceedings of the National Academy of Sciences of the United States of America,* 74**,** 5463-5467.

SANTIAGO-BLAY, J. A. 1994. Paleontology of leaf beetles. *In:* P. H. JOLIVET, M. L. C., E. PETITPIERRE (ed.) *Novel aspects of the biology of Chrysomelidae.* Kluwer Academic publishers.

SCHNEIDER, E. & HUNKE, S. 1998. ATP-binding-cassette (ABC) transport systems: functional and structural aspects of the ATP-hydrolyzing subunits/domains. *FEMS Microbiol Rev,* 22**,** 1-20.

SCHULZ, M. H., ZERBINO, D. R., VINGRON, M. & BIRNEY, E. 2012. Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics,* 28**,** 1086-92.

SCHULZ, S., GROSS, J. & HILKER, M. 1997. Origin of the defensive secretion of the leaf beetle Chrysomela lapponica. *Tetrahedron,* 53**,** 9203-9212.

SCHURMANN, A., DOEGE, H., OHNIMUS, H., MONSER, V., BUCHS, A. & JOOST, H. G. 1997. Role of conserved arginine and glutamate residues on the cytosolic surface of glucose transporters for transporter function. *Biochemistry,* 36**,** 12897-902.

SEEGER, M. A. & VAN VEEN, H. W. 2009. Molecular basis of multidrug transport by ABC transporters. *Biochim Biophys Acta,* 1794**,** 725-37.

SEINEN, E., BURGERHOF, J. G. M., JANSEN, R. C. & SIBON, O. C. M. 2011. RNAi-induced off-target effects in Drosophila melanogaster: frequencies and solutions. *Briefings in Functional Genomics,* 10**,** 206-214.

SILVERTON, L., DEAN, M. & MOITRA, K. 2011. Variation and evolution of the ABC transporter genes ABCB1, ABCC1, ABCG2, ABCG5 and ABCG8: implication for pharmacogenetics and disease. *Drug Metabol Drug Interact,* 26**,** 169-79.

SIMON, S. A., ZHAI, J., NANDETY, R. S., MCCORMICK, K. P., ZENG, J., MEJIA, D. & MEYERS, B. C. 2009. Short-read sequencing technologies for transcriptional analyses. *Annu Rev Plant Biol,* 60**,** 305-33.

SIMPSON, J. T., WONG, K., JACKMAN, S. D., SCHEIN, J. E., JONES, S. J. M. & BIROL, I. 2009. ABySS: A parallel assembler for short read sequence data. *Genome Research,* 19**,** 1117-1123.

SOE, A. R. B., BARTRAM, S., GATTO, N. & BOLAND, W. 2004. Are iridoids in leaf beetle larvae synthesized de novo or derived from plant precursors? A methodological approach. *Isotopes in Environmental and Health Studies,* 40**,** 175-180.

SONESON, C. & DELORENZI, M. 2013. A comparison of methods for differential expression analysis of RNA-seq data. *BMC Bioinformatics,* 14**,** 91.

SRIVASTAVA, S. & CHEN, L. 2010. A two-parameter generalized Poisson model to improve the analysis of RNA-seq data. *Nucleic Acids Res,* 38**,** e170.

STAMATAKIS, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics,* 22**,** 2688-2690.

STAMATAKIS, A., HOOVER, P. & ROUGEMONT, J. 2008. A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol,* 57**,** 758-71.

STRAUSS, A. S., PETERS, S., BOLAND, W. & BURSE, A. 2013. ABC transporter functions as a pacemaker for sequestration of plant glucosides in leaf beetles. *eLife***,** DOI:10.7554/eLife.01096.

STRAUSS, A. S., PETERS, S., BURSE, A., HEINEMANN, S. H. & BOLAND, W. 2012. ABC transporter facilitates sequestration of plant glucosides in leaf beetle larvae. *FEBS Journal,* 279**,** 263-263.

STURM, A., CUNNINGHAM, P. & DEAN, M. 2009. The ABC transporter gene family of Daphnia pulex. *Bmc Genomics,* 10**,** 170.

SU, Y.-L., LI, J.-M., LI, M., LUAN, J.-B., YE, X.-D., WANG, X.-W. & LIU, S.-S. 2012. Transcriptomic Analysis of the Salivary Glands of an Invasive Whitefly. *{PLOS ONE},* {7}.

SULTAN, M., SCHULZ, M. H., RICHARD, H., MAGEN, A., KLINGENHOFF, A., SCHERF, M., SEIFERT, M., BORODINA, T., SOLDATOV, A., PARKHOMCHUK, D., SCHMIDT, D., O'KEEFFE, S., HAAS, S., VINGRON, M., LEHRACH, H. & YASPO, M. L. 2008. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science,* 321**,** 956-60.

TARAZONA, S., GARCIA-ALCALDE, F., DOPAZO, J., FERRER, A. & CONESA, A. 2011. Differential expression in RNA-seq: a matter of depth. *Genome Res,* 21**,** 2213-23.

TERENIUS, O., PAPANICOLAOU, A., GARBUTT, J. S., ELEFTHERIANOS, I., HUVENNE, H., KANGINAKUDRU, S., ALBRECHTSEN, M., AN, C. J., AYMERIC, J. L., BARTHEL, A., BEBAS, P., BITRA, K., BRAVO, A., CHEVALIERI, F., COLLINGE, D. P., CRAVA, C. M., DE MAAGD, R. A., DUVIC, B., ERLANDSON, M., FAYE, I., FELFOLDI, G., FUJIWARA, H., FUTAHASHI, R., GANDHE, A. S., GATEHOUSE, H. S., GATEHOUSE, L. N., GIEBULTOWICZ, J. M., GOMEZ, I., GRIMMELIKHUIJZEN, C. J. P., GROOT, A. T., HAUSER, F., HECKEL, D. G., HEGEDUS, D. D., HRYCAJ, S., HUANG, L. H., HULL, J. J., IATROU, K., IGA, M., KANOST, M. R., KOTWICA, J., LI, C. Y., LI, J. H., LIU, J. S., LUNDMARK, M., MATSUMOTO, S., MEYERING-VOS, M., MILLICHAP, P. J., MONTEIRO, A., MRINAL, N., NIIMI, T., NOWARA, D., OHNISHI, A., OOSTRA, V., OZAKI, K., PAPAKONSTANTINOU, M., POPADIC, A., RAJAM, M. V., SAENKO, S., SIMPSON, R. M., SOBERON, M., STRAND, M. R., TOMITA, S., TOPRAK, U., WANG, P., WEE, C. W., WHYARD, S., ZHANG, W. Q., NAGARAJU, J., FFRENCH-CONSTANT, R. H., HERRERO, S., GORDON, K., SWELTERS, L. & SMAGGHE, G. 2011. RNA interference in Lepidoptera: An overview of successful and unsuccessful studies and implications for experimental design. *Journal of Insect Physiology,* 57**,** 231-245.

TERMONIA, A., HSIAO, T. H., PASTEELS, J. M. & MILINKOVITCH, M. C. 2001. Feeding specialization and host-derived chemical defense in Chrysomeline leaf beetles did not lead to an evolutionary dead end. *Proceedings of the National Academy of Sciences, USA,* 98**,** 3909-3914.

TERMONIA, A. & PASTEELS, J. M. 1999. Larval chemical defence and evolution of host shifts in Chrysomela leaf beetles. *Chemoecology,* 9**,** 13-23.

THOMPSON, J. F. & STEINMANN, K. E. 2010. Single molecule sequencing with a HeliScope genetic analysis system. *Curr Protoc Mol Biol,* Chapter 7**,** Unit7 10.

THOMPSON, S. N. 2003. Trehalose - The insect 'blood' sugar. *In:* SIMPSON, S. J. (ed.) *Advances in Insect Physiology, Vol 31.* London: Academic Press Ltd-Elsevier Science Ltd.

THORENS, B. & MUECKLER, M. 2010. Glucose transporters in the 21st Century. *American Journal of Physiology Endocrinology and Metabolism,* 298**,** E141-E145.

TOLZIN-BANASCH, K., DAGVADORJ, E., SAMMER, U., KUNERT, M., KIRSCH, R., PLOSS, K., PASTEELS, J. M. & BOLAND, W. 2011. Glucose and Glucose Esters in the Larval Secretion of Chrysomela Lapponica; Selectivity of the Glucoside Import System from Host Plant Leaves. *Journal of Chemical Ecology,* 37**,** 195-204.

TOMOYASU, Y. & DENELL, R. E. 2004. Larval RNAi in Tribolium (Coleoptera) for analyzing adult development. *Development Genes and Evolution,* 214**,** 575-578.

TOMOYASU, Y., MILLER, S. C., TOMITA, S., SCHOPPMEIER, M., GROSSMANN, D. & BUCHER, G. 2008. Exploring systemic RNA interference in insects: a genome-wide survey for RNAi genes in *Tribolium*. *Genome Biology,* 9.

TRAPNELL, C., WILLIAMS, B. A., PERTEA, G., MORTAZAVI, A., KWAN, G., VAN BAREN, M. J., SALZBERG, S. L., WOLD, B. J. & PACHTER, L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol,* 28**,** 511-5.

TREPAGNIER, E. H., RADENOVIC, A., SIVAK, D., GEISSLER, P. & LIPHARDT, J. 2007. Controlling DNA capture and propagation through artificial nanopores. *Nano Lett,* 7**,** 2824-30.

TUCKER, T., MARRA, M. & FRIEDMAN, J. M. 2009. Massively parallel sequencing: the next big thing in genetic medicine. *Am J Hum Genet,* 85**,** 142-54.

ULDRY, M. & THORENS, B. 2004. The SLC2 family of facilitated hexose and polyol transporters. *Pflugers Archiv-European Journal of Physiology,* 447**,** 480-489.

VIJAY, N., POELSTRA, J. W., KUNSTNER, A. & WOLF, J. B. 2013. Challenges and strategies in transcriptome assembly and differential gene expression quantification. A comprehensive in silico assessment of RNA-seq experiments. *Mol Ecol,* 22**,** 620-34.

WALL, P. K., LEEBENS-MACK, J., CHANDERBALI, A. S., BARAKAT, A., WOLCOTT, E., LIANG, H., LANDHERR, L., TOMSHO, L. P., HU, Y., CARLSON, J. E., MA, H., SCHUSTER, S. C., SOLTIS, D. E., SOLTIS, P. S., ALTMAN, N. & DEPAMPHILIS, C. W. 2009. Comparison of next generation sequencing technologies for transcriptome characterization. *Bmc Genomics,* 10**,** 347.

WANG, E. T., SANDBERG, R., LUO, S., KHREBTUKOVA, I., ZHANG, L., MAYR, C., KINGSMORE, S. F., SCHROTH, G. P. & BURGE, C. B. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature,* 456**,** 470-6.

WANG, X. W., LUAN, J. B., LI, J. M., BAO, Y. Y., ZHANG, C. X. & LIU, S. S. 2010. De novo characterization of a whitefly transcriptome and analysis of its gene expression during development. *BMC Genomics,* 11**,** 400.

WANG, Z., GERSTEIN, M. & SNYDER, M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet,* 10**,** 57-63.

WHANGBO, J. S. & HUNTER, C. P. 2008. Environmental RNA interference. *Trends Genet,* 24**,** 297-305.

WHEELER, D. A., SRINIVASAN, M., EGHOLM, M., SHEN, Y., CHEN, L., MCGUIRE, A., HE, W., CHEN, Y. J., MAKHIJANI, V., ROTH, G. T., GOMES, X., TARTARO, K., NIAZI, F., TURCOTTE, C. L., IRZYK, G. P., LUPSKI, J. R., CHINAULT, C., SONG, X. Z., LIU, Y., YUAN, Y., NAZARETH, L., QIN, X., MUZNY, D. M., MARGULIES, M., WEINSTOCK, G. M., GIBBS, R. A. & ROTHBERG, J. M. 2008. The complete genome of an individual by massively parallel DNA sequencing. *Nature,* 452**,** 872-6.

WHELAN, S., LIO, P. & GOLDMAN, N. 2001. Molecular phylogenetics: state-of-the-art methods for looking into the past. *Trends Genet,* 17**,** 262-72.

WILSON-O'BRIEN, A. L., PATRON, N. & ROGERS, S. 2010. Evolutionary ancestry and novel functions of the mammalian glucose transporter (GLUT) family. *BMC Evolutionary Biology,* 10.

WOOD, I. S. & TRAYHURN, P. 2003. Glucose transporters (GLUT and SGLT): expanded families of sugar transport proteins. *British Journal of Nutrition,* 89**,** 3-9.

WU, W., DNYANMOTE, A. V. & NIGAM, S. K. 2011. Remote communication through solute carriers and ATP binding cassette drug transporter pathways: an update on the remote sensing and signaling hypothesis. *Mol Pharmacol,* 79**,** 795-805.

XIE, X., CHENG, T., WANG, G., DUAN, J., NIU, W. & XIA, Q. 2012. Genome-wide analysis of the ATP-binding cassette (ABC) transporter gene family in the silkworm, Bombyx mori. *Mol Biol Rep,* 39**,** 7281-91.

YANG, X., CHOCKALINGAM, S. P. & ALURU, S. 2013. A survey of error-correction methods for next-generation sequencing. *Brief Bioinform,* 14**,** 56-66.

YANG, Y. & SMITH, S. A. 2013. Optimizing de novo assembly of short-read RNA-seq data for phylogenomics. *BMC Genomics,* 14**,** 328.

YAZAKI, K. 2006. ABC transporters involved in the transport of plant secondary metabolites. *FEBS Lett,* 580**,** 1183-91.

ZHAO, F.-Q. & KEATING, A. F. 2007. Functional properties and genomics of glucose transporters. *Current Genomics,* 8**,** 113-128.

ZHAO, Q. Y., WANG, Y., KONG, Y. M., LUO, D., LI, X. & HAO, P. 2011. Optimizing de novo transcriptome assembly from short-read RNA-Seq data: a comparative study. *BMC Bioinformatics,* 12 Suppl 14**,** S2.

ACKNOWLEDGEMENTS

At this point I would like to express my gratitude to Prof. Dr. Wilhelm Boland and Dr. Antje Burse for their supervision of my work and the confidence and liberty they offered me in following up on and realizing my scientific ideas. I especially want to thank Antje for all her support in establishing bioinformatic analyses in the Bioorganic Chemistry Department of the MPI for Chemical Ecology, Jena, and in helping to write the manuscripts.

Best thanks go to Andreas Petzold who always helped me when I struggled with assembly tools and for all the fruitful discussions of the many aspects of bioinformatics.

I am very grateful for Dr. Karin Groten´s and Dr. Jan Kellmann´s advice and encouragement to continue working on my doctoral thesis when I almost surrendered during maternal leave. Additionally, I want to thank Karin, and also Dr. Dorit Schmidt, for having a great time as associate in the IMPRS and for allowing me to attend many adjuvant soft skill courses.

I want to give many thanks to all members of the "beetle group", especially Dr. Sindy Frick, Peter Rahfeld, René Gretscher, Gerhard Pauls and also Dr. Sabrina Discher for worthwhile coffee breaks including innumerable enlightening and inspiring conversations and for a great working atmosphere. I cordially thank Sindy for providing me the MS Word template. I also thank René for the establishment of his RNAi injection method, because, together with his work, my studies became even more fascinating.

I want to acknowledge Dr. Karla Tolzin-Banasch who first told me about the opportunity to apply at the Bioorganic Chemistry Department to analyze their sequence data. Without her recommendation I would not have known of this available PhD position.

I want to appreciate Angelika Berg for always being around at the institute to take care of the beetles and larvae.

I am very thankful for Dr. Lydia Gramzow´s explanations with regard to calculating and interpreting phylogenetic trees and for Ewald Grosse-Wilde´s willingness to let me work a few days at his workstation to use the CLCbio Workbench installed there.

I want to express thanks to Dr. Sven Geiselhardt and Prof. Dr. Monika Hilker for the cooperation in the study of the antennal transcriptome of *P. cochleariae* which, unfortunately, in the end could not be finished successfully.

I express my gratefulness to Prof. Dr. Reinhard Guthke who assisted me during the last year of my doctoral studies and approved the manuscript on the sugar transporters of *P. cochleariae*.

I want to acknowledge Dr. Natalie Wielsch for all her MS analyses and processing, and I hope to stay in contact and to cooperate with her in the future.

Many thanks go to Dr. Marco Groth for doing a great job sequencing all samples derived from the beetles and to Bernd Senf for making possible the use of the computer network and installed software tools at the Fritz Lipmann Institute, Jena.

Furthermore, special thanks are given to the IT group, the administration and the library team of the MPI for Chemical Ecology, especially to Martin Niebergall and Hendrik Tilger for their support.

I thank Dr. Steffen Priebe for a pleasant time during the workshop “Advanced RNA-Seq and ChiP-Seq data analysis” at EMBL, Heidelberg, in 2009 and for help regarding R when needed.

Sincere thanks are given to Melinda Palmer Kolb who critically proofread my thesis and the sugar transporter manuscript and for helpful discussions concerning the English language.

I would like to thank all my friends for sharing an eventful time and supporting me during my doctoral research studies.

I am especially grateful for my family´s love, motivation and support over and above my doctoral thesis.

EHRENWÖRTLICHE ERKLÄRUNG

Hiermit erkläre ich, dass mir die geltende Promotionsordnung (§ 5 PromO) der Biologisch-Pharmazeutischen Fakultät der Friedrich-Schiller-Universität Jena bekannt ist und ich mich mit bestem Wissen an diese Ordnung gehalten habe. Die vorliegende Dissertation habe ich selbstständig und nur unter der Verwendung, Hilfe oder Benutzung der angegebenen Hilfsmittel, Daten und Quellen angefertigt. Übernommene Inhalte aus anderen Quellen und von anderen Personen, die in dieser Arbeit Verwendung fanden oder auf die direkt Bezug genommen wird, sind als solche eindeutig kenntlich gemacht. Unterstützung während meiner wissenschaftlichen Arbeit und zur Erstellung dieses Dissertationstextes habe ich nur von den genannten Co-Autoren oder den in der Danksagung genannten Personen erhalten. Ich habe keine Hilfe von externen Vermittlungs- und Beratungsdiensten in Anspruch genommen, und niemand hat unmittelbar noch mittelbar von mir geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die vorgelegte Dissertation wurde zu keinem früheren Zeitpunkt weder im Inland noch im Ausland in gleicher oder ähnlicher Form an einer anderen Prüfungsbehörde oder Hochschule zum Zweck einer Promotion oder eines anderen Prüfungsverfahrens eingereicht.

Magdalena Stock

Jena, den

CURRICULUM VITAE

PERSONAL DATA

Name:	Magdalena Stock
Current address:	Weißenfelser Str. 6
	07743 Jena, Germany
Date of Birth:	13.08.1980
City of Birth:	Jena, Germany
Sex:	Female
Marital status:	married, 3 children
Nationality:	German

ACADEMIC EDUCATION

02/2013 – current 04/2009 – 12/2011	Doctoral studies at the Max Plank Institute for Chemical Ecology, Jena Department of Bioorganic Chemistry First supervisor: Prof. Wilhelm Boland
10/2008 – 02/2009 05/2007 – 09/2007	Diploma student at the Hans-Knöll-Institute, Jena First supervisor: Prof. Reinhard Guthke Title of diploma thesis: "Inferring gene regulatory networks from microarray data of *Candida albicans*"
10/2000 – 02/2005 04/2006 – 02/2009	Studies of Bioinformatics (diploma course) at the Friedrich Schiller University, Jena

SCIENTIFIC PUBLICATIONS

JOURNAL ARTICLES

Stock, M., Gretscher, R. R., Groth, M., Eiserloh, S., Boland, W., Burse, A. (2013). Putative sugar transporters of the mustard leaf beetle *Phaedon cochleariae*: their phylogeny and role for nutrient supply in larval defensive glands. *PLoS ONE*. 8(12): e84461. doi:10.1371/journal.pone.0084461

Bodemann, R. R., Rahfeld, P., **Stock, M.**, Kunert, M., Wielsch, N., Groth, M., Frick, S., Boland, W., Burse, A. (2012). Precise RNAi mediated silencing of metabolic active proteins in defence secretions of juvenile leaf beetles. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 279(1745), 4126-4134. doi:10.1098/rspb.2012.1342.

ORAL PRESENTATIONS

Stock M. (2011) Annotation of *Phaedon cochleariae*'s transcriptome and establishment of BLAST servers. *10th IMPRS Symposium*, MPI for Chemical Ecology, Dornburg, DE

POSTER PRESENTATIONS

Gretscher R., Wielsch N., **Stock M.**, Boland W., Burse A. (2013) The non-model approach. Transcriptome guided identification of secretory proteins from leaf beetle defense secretions via LC-MSE. *7th Central and Eastern European Proteomics Conference (CEEPC)*, Max Planck Institute for Chemical Ecology, Jena, DE

Stock M., Bodemann R., Boland W., Burse A. (2013) Exploring the transcriptome of *Phaedon cochleariae*; putative sugar transporters involved in larval defense. *12th IMPRS Symposium*, Max Planck Institute for Chemical Ecology, Jena, DE

Stock M., Bodemann R., Boland W., Burse A. (2013) Exploring the transcriptome of *Phaedon cochleariae*: putative sugar transporters involved in larval defense. *iBeetle Symposium: New horizons in molecular zoology*, Göttingen, DE

Stock M., Burse A., Boland W. (2011) Exploring the transcriptomes of leaf beetles; candidate

genes involved in larval defense. *27th ISCE Meeting, International Society of Chemical Ecology*, Simon Fraser University, Burnaby, British Columbia, CA

Stock M., Burse A., Boland W. (2011) Exploring *Phaedon cochleariae*'s transcriptome to identify candidate genes involved in larval defense. *15th Annual International Conference on Research in Computational Molecular Biology*, Vancouver, CA

Burse A., Frick S., Bodemann RR., Kirsch R., Kunert M., Pauls G., Shaheed M., **Stock M**., Strauß A., Tolzin-Banasch K., Boland W. (2010) Leaf beetles – Adaptation to their host plants and exploitation of phytochemicals. *SAB Meeting 2010*, MPI for Chemical Ecology, Jena, DE

Stock M., Burse A., Boland W. (2010) Transcriptome Analysis of *Phaedon cochleariae* Using High-Throughput Sequencing. *26th ISCE Meeting, International Society of Chemical Ecology*, Tours, FR

Stock M., Burse A., Boland W. (2010) Transcriptome analysis of leaf beetles using high-throughput sequencing. *9th IMPRS Symposium*, MPI for Chemical Ecology, Dornburg, DE
Burse A., Frick S., Discher S., Kirsch R., Kunert M., **Stock M**., Strauss A., Tolzin K., Boland W. (2009) Chemical Defense in Leaf Beetle Larvae. *ICE Symposium*, MPI for Chemical Ecology, Jena, DE