Gene expression and structural differences underpinning black and white colouration in spiders

Fabian C. Salgado-Roa*, Devi Stuart-Fox, Laura Ospina-Rozo, Iliana Medina

School of Biosciences, The University of Melbourne, Parkville, VIC 3010, Australia

*Email: fcsalgador@gmail.com

ABSTRACT

The presence of multiple colour variants within a single population, known as colour polymorphism, is a striking example of intraspecific phenotypic variation and a model for studying evolutionary processes. These processes depend on the underlying mechanisms such as the genes and pigments involved; but we currently lack such information for non-model groups such as arachnids. We examined the genes associated with colour polymorphism in the Australian Christmas spider, *Austracantha minax*, a widespread species with two female-restricted colour morphs: one that is completely black and the other presenting a black-and-white pattern on its dorsum. We sequenced the mRNA of individuals of both morphs at different developmental stages and assessed differences in gene expression between colour morphs based on a *de novo* transcriptome assembly for the species, and integrated this with high resolution electron microscopy observations. We found that gene expression patterns are primarily influenced by the stage of development, and within each stage we found differentially expressed genes between the black and black-and-white morphs, especially in the adult stage. Contrary to expectations, these genes were not associated with melanin or guanine crystal deposition (detected in the microscopy), but were associated with chitin and cuticle metabolic processes. By

exploring arachnid coloration at the transcriptomic level, our work reveals for the first time genetic mechanisms underpinning the evolution of colour polymorphism in spiders.

Keywords: Colour polymorphism, chitin metabolism, melanin, arachnids

INTRODUCTION

Colour variation is one of the most noticeable forms of biodiversity. Colour traits have multiple adaptive roles, and the study of their genetic and evolutionary mechanisms has generated influential hypotheses to explain how phenotypic variation changes and is maintained through time and space (Gray & McKinnon, 2007; Cuthill et al., 2017; Svensson, 2017). One of the most common patterns of colour variation in animals involves light and dark phenotypes, often referred to as melanism (Majerus, 1998). In vertebrates, the pigmentary, genetic, and structural mechanisms behind dark coloration are clear, and common mechanisms linked to melanin production have been identified across different lineages, suggesting similar evolutionary pathways to produce colour diversity (Hubbard et al., 2010; Caro & Mallarino, 2020; McNamara et al., 2021; Elkin et al., 2023). Despite arthropods being one of the largest and most colourful clades, there is limited information on the mechanisms related to their colour diversity. Most of the genetic and morphological mechanisms associated with colour production come from a handful of model organisms, and we are still far from understanding the mechanisms of colouration across diverse and speciose lineages such as crustaceans and arachnids (Campli et al., 2024).

Understanding the molecular and structural mechanisms behind coloration in poorly studied lineages is crucial to identify conserved evolutionary patterns and provide insights into repeated usage of genes and structures in divergent lineages under similar selective pressures (Nosil et al., 2018; Orteu & Jiggins, 2020). In arthropods, most of our understanding about the mechanisms behind dark coloration comes from model systems such as *Drosophila* and its melanisation (Wittkopp et al., 2003;

Wittkopp & Beldade, 2009). This process involves a group of enzymes (*pale, DDC, DAT, yellow, ebony, tan, black*) that catalyse multiple steps of the melanin pathway (Detailed pathway in True, 2003; Sugumaran & Barek, 2016). Many of these enzymes have been reported to play a key role in melanisation in bees (Rahman et al., 2021), butterflies (Zhang et al., 2017) and hemimetabolous insects (J. Liu et al., 2016), suggesting a conserved melanisation process across insects. Only a few of the core set of genes responsible for insect melanisation have been found in other arthropods like arachnids (Croucher et al., 2013), crustaceans (Lin et al., 2022) and myriapods (Wang et al., 2021). For these lineages, whether and how melanisation occurs is still unresolved (Campli et al., 2024). For instance, to our knowledge, there is a single transcriptomic study with spiders that identified only seven out of the 19 possible genes that are associated with melanin production in insects (Croucher et al., 2013), but there is no evidence that these genes are actually associated with dark phenotypes in spiders.

The pigmentary basis of dark and brown body colours of spiders has only been explored in a few studies, despite dark colours being present in many taxa (Salgado-Roa et al. 2024). Critically, melanin was reported in spiders for the first time only in 2015 (Hsiung et al., 2015) and the presence of melanosomes was confirmed later in jumping spiders (Hsiung et al., 2017). Current evidence of the presence of melanin comes mainly from leg hair samples, and its association with abdominal colour is not clear. For example, we do not know whether melanin is embedded within the cuticle, as is the case in dark insects, or whether it is present in the hypodermis under a transparent cuticle. The role of melanin in generating dark abdominal colours in spiders had been disregarded previously based on hypodermal dissection (not the

cuticle). Instead, ommochrome pigments (a combination of xanthommatin and ommins) were suggested as the compounds generating dark colours in spiders (Seligy, 1972). Therefore, to date we do not clearly understand the prevalence of melanin or ommochrome pigments in generating dark colouration in spiders. Understanding which pigments underlie colour variation could have functional implications, considering that melanin confers immune protection (Nosanchuk & Casadevall, 2003), UV protection (Kollias et al., 1991), and reduces abrasion (True, 2003) in different arthropod species.

Colour polymorphism is a common phenomenon in web-building spiders; in fact, polymorphisms that include dark and light morphs are among the most common in this group (Salgado-Roa et al., 2024). This makes them excellent models for studying the mechanisms producing dark and light coloration in arachnids. The Christmas spider, *Austracantha minax*, is a colour polymorphic species that is common and abundant in Australia. Females of this species present two colour variants that coexist throughout much of its southern distribution: one is completely dark on both its dorsal and ventral sides, and the other presents black and white on the dorsal side and a mosaic yellow pattern ventrally (Figure 1; Waldock, 1991). The mechanisms responsible for differences in the proportion of dark colours between the colour variants of this species likely involve the synthesis of melanin or ommochromes. Moreover, since these colour differences also involve white coloration, mechanisms controlling the presence and synthesis of guanine crystals must be active during the development of this species, as light colours in spiders have been clearly linked to the presence of disorganized guanine crystal deposits

under a translucent cuticle (Insausti & Casas, 2008; Levy-Lior et al., 2010; Wagner et al., 2022).

In this work, we explored the ontogeny, structures, and genes associated with light and dark phenotypes of a colour polymorphic spider (*Austracantha minax*). We first raised spider eggs under controlled conditions to observe how colour differentiation occurs across life stages. To assess differences in gene expression, we sequenced the mRNA of individuals of both morphs at different developmental stages and assessed differences between colour morphs based on a *de novo* transcriptome assembly for the species. Lastly, to explore the structure and composition of the black and white colours in spiders, we visualized and qualitatively assessed the structures of the different colour morphs using cryo-scanning electron microscopy (cryoSEM). We expected to find structural differences, such as the presence of guanine crystals, between the light and dark morphs. We predicted that genes involved in melanin, ommochrome, or guanine synthesis would differ in expression between the two colour morphs. These differences are expected to be stronger and clearer in the early stages of colour development.

METHODS

Microscopy of structures of colour variants

We used high resolution cryo-scanning electron microscopy (cryoSEM) to qualitatively assess whether there were structures associated with the differences in coloration between the black and black-and-white morphs. We studied two adult specimens for each morph collected in the field. We used a dissecting light microscope to remove as much haemolymph and gut content as possible without destroying the white patches or cuticle in the posterior end of the abdomen. After this, the samples were mounted in a universal cryo stub exposing the abdominal spines, and the sample was frozen in liquid nitrogen. We fractured the abdominal spines to remove their tip and expose the inner structures in cryogenic conditions. After coating with a standard electrically conductive platinum layer, samples were transferred to the cold stage in a SU7000 field emission scanning electron microscope (Hitachi, Japan). Images were obtained using a 2.00 kV electron beam and three different detectors. The resulting EM images were overlayed using Fiji, ImageJ 1.54f to determine the general shape, depth and topology of the structures in the posterior spines. We report any differential structure visualised.

Sampling and spider rearing for expression analyses

We collected egg sacs and adults of *Austracantha minax* from one location near Geelong, Victoria, Australia (-38.22, 144.211) and transported them live to controlled laboratory conditions at the University of Melbourne (24°C, 60% RH, 12-hour photoperiod). Given that the inheritance of coloration in this system is unknown, we reared individuals from egg sacs, keeping track of their developmental times and

separating each egg sac into an individual net cage. Spiderlings were fed *Drosophila melanogaster*.

To explore gene expression differences between morphs and across development, we selected individuals from three different stages (Fig. 1). We chose five unrelated individuals at the first instar, when spiderlings lack white or black coloration on their abdomen (Fig. 1). The first stage where we could differentiate between black-and-white and black juveniles was during the third instar, so we selected fifteen individuals from this stage from nine different egg sacs (Fig. 1; eight black-and-white and seven black). Additionally, we collected ten adults (Fig. 1; five black-and-white and five black) from the field, representing the stage where the colour phenotype is fully developed. All individuals were subjected to at least three days of starvation before RNA extraction.

RNA extraction and sequencing

To recover the maximum number of transcripts and RNA amount, we extracted total RNA from each individual using a hybrid protocol with TRIzol® and ethanol/chloroform. Because the individuals were so small (First instar: 0.25 mm abdominal width, juveniles: 0.42 mm abdominal width, adults: 8 mm abdominal width), we could not isolate cuticle from other tissues, so we first ground the entire body of every fresh individual in 500 µL of cold TRIzol and left the mixture overnight at 4°C. We added chloroform (200 µL for each mL of TRIzol®) and mixed the solution by inverting the tube. This mixture was then kept at 4°C overnight. Next, we centrifuged the sample at 14,000 rpm for 15 minutes, maintaining the temperature at

4°C. The upper aqueous phase, which contained the RNA, was then combined with an equal volume of cold 100% isopropanol and incubated overnight at 4°C. The RNA was pelleted by centrifugation for 10 min at 12,000 x g and 4°C. The RNA pellet was washed twice using 1 mL of 75% cold ethanol, and between washes, the samples were spun at 7,500 x g for 5 minutes at 4°C. After the final wash, the ethanol was removed, and the pellet was resuspended in 30 µL of ultrapure water. With the final samples, we followed the Thermo Scientific DNase I/RNase-free protocol.

RNA extractions were shipped to the Australian Genome Research Facility (AGRF, Melbourne, Australia) for quality checking, library preparation, and sequencing. We used a mRNA-Seq library preparation with standard poly-A tail purification with Illumina primers and individual barcoding. The samples were sequenced using 150 bp paired-end reads with Illumina NovaSeq X.

Transcriptome assembly and annotation

Before performing a transcriptome assembly, we processed the raw reads by applying quality control, which consisted of inspecting the data and subsequent correction or filtering when necessary (Raghavan et al., 2022). First, we detected and removed reads with erroneous k-mers using rCorrector v1.0.5 (Song & Florea, 2015). We then trimmed low-quality bases and adapters using Trim Galore v0.6.10 and cutadapt v4.2 (Martin, 2011) with the flags "--paired --phred33 --length 36 -q 5 --stringency 1 -e 0.1". The resulting trimmed sequences were mapped to the SILVA database (Quast et al., 2013) to remove reads that potentially did not correspond to eukaryotes using Bowtie v2.4.5 (Langmead & Salzberg, 2012). After this, we assessed the quality of the reads and identified overrepresented sequences using

FastQC v0.11.7 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Overrepresented sequences were removed using the Python script RemoveFastqcOverrepSequenceReads.py (https://github.com/harvardinformatics/TranscriptomeAssemblyTools).

We used the resulting filtered and inspected reads to assemble a *de novo* transcriptome using Trinity v2.15.1 (Grabherr et al., 2011) *de novo* transcriptome assembly algorithm with all the sequences and default parameters and a minimum contig length of 300. This set of parameters was optimal during the assembly quality evaluation. Because *de novo* transcriptome assemblers usually generate far more sequences than anticipated, considering the number of genes present in the genome (Raghavan et al., 2022), we thinned the assembly using the clustering tool CD-HIT with an identity cutoff of 95% (Fu et al., 2012). We evaluated the completeness of the assembly based on their BUSCO scores against the arachnid database arachnida_odb10 (Simão et al., 2015). We also evaluated the quality of the assemblies by calculating the read representation of the assembly (where at least ~80% of the input RNA-Seq reads should be represented in the transcriptome assembly) and the E90N50 transcript contig length. To functionally annotate the *de novo* assembled transcriptome, we used the EnTap pipeline v1.0.1 (Hart et al., 2020) using SwissProt, TrEMBL, EggNOG, and RefSeq Invertebrate databases as references.

Differential gene expression

We calculated transcript counts by creating a pseudoalignment of all the reads per sample against the generated reference transcriptome using Salmon (Patro et al., 2017). With the obtained raw transcript counts per sample, we explored differential gene expression using the R package DESeq2 (Love et al., 2014). To reduce possible contaminants in our analyses, we first discarded those transcripts that, after the transcriptome annotation, were not associated with Araneomorphae, Arthropoda, or Insecta (keeping an average of 74% of the expressed transcripts, SE=0.4%). We then summarized the counts of all isoforms that had the same annotation within the same gene cluster to avoid overrepresentation in the following analyses. We did this by creating a custom tx2gene file to import the counts with tximport (Soneson et al., 2016) into R. With these transcripts, we built three different models using likelihood ratio tests. The first model used developmental stage as a fixed term, the second explored differences between the black-and-white and black colour morphs in juveniles, and the third examined differences between the black-and-white and black colour morphs in adults. We were unable to build a model that included all our samples combining coloration and developmental stage as predictors, given the lack of colour variation at the first stage of development.

For each model, we fit a negative binomial generalized linear model, and tested significance with a Wald test. We visualised sample clustering by plotting the first two principal components (PCs) that summarized the variation of the top 1,000 differentially expressed transcripts. This allowed us to identify outlier samples that did not cluster with any of the treatments (Fig. S1). Outlier samples typically correspond to contaminated samples or technical failures, and removing them is essential for an accurate estimation of treatment variance (Chen et al., 2020). As a

result, five individuals of the third instar and two adult samples were removed given their clear mismatch from their experimental group (supplementary material). Our final sampling consisted of nine black-and-white and eight black juveniles, four black-and-white and four black adults. We calculated the fold changes of transcripts between stages or colour morphs using an adaptive shrinkage estimator from the R package *ashr* (Stephens, 2017). Transcripts were considered differentially expressed at a FDR-adjusted p- value <0.01. The pattern of differential gene expression between colour morphs was compared with a fold expression to P-value plot build with the *EnhancedVolcano* R package (https://github.com/kevinblighe/EnhancedVolcano). We also visualised transcript expression variation with hierarchical clustered heatmaps obtained with the R package *pheatmap* (https://github.com/raivokolde/pheatmap).

Gene ontology enrichment analysis

We inferred the biological function of the differentially expressed transcripts by performing a gene ontology (GO) enrichment analysis using GOSEQ v1.46 (Young et al., 2010) with a P-value cutoff of 0.05 and adjusted P-value of 0.001. P-values were adjusted for multiple testing using the Benjamini-Hochberg method (Benjamini & Hochberg, 1995). To account for selection bias, we considered the length and level of expression of the transcripts to identify overrepresented GO terms. Transcripts categorized under overrepresented GO terms of interest were identified using a BLASTX search (--evalue 1e-5 --outfmt 6 --max-target-seqs 1 --max-hsps 1) in DIAMOND (Buchfink et al., 2015) against the Invertebrate RefSeq database (O'Leary et al., 2016), selecting the best hit per transcript.

Detection and differential expression of melanin and ommochrome pathway candidate genes

Considering that the main colour difference between the morphs of the Christmas spider is the amount of dark colouring, we decided to specifically test whether there were differences in the expression of melanin-associated transcripts and/or ommochrome-associated transcripts. To accomplish this, we used two approaches: first, we explored if any of the transcripts linked with the melanin metabolic process GO term (GO:0006582) or the ommochrome metabolic process GO term (GO:0046152) were present in the list of differentially expressed transcripts. Second, we attempted to identify if ten melanin pathway genes (*ebony, black, aaNAT, yellow, tan, pale, DAT, nubbin, prophenoloxidase 3, peroxidase*) critical for the development of melanin in arthropods (Liu et al., 2016; Rahman et al., 2021) and six ommochrome synthesis genes (*Tryptophan 2,3-dioxygenase (TDO),Kynurenine formamidase (KFase),Kynurenine 3-monooxygenase (KMO),Phenoxazinone synthase (PHS),Heme peroxidase, KAT*) common among invertebrates (Figon & Casas, 2019), were present in our assembled transcriptome. We downloaded *Drosophila melanogaster* protein sequences of the 16 genes available on the NCBI and performed a TBLASTN against our reference transcriptome, selecting the best hit (-outfmt 6 -evalue 1e-8 -max_target_seqs 1 -max_hsps 1). For those transcripts significantly associated with the ten melanin or six ommochrome associated genes, we performed a reciprocal BLASTX search (--evalue 1e-5 --outfmt 6 --max-target-seqs 1 --max-hsps 1) in DIAMOND (Buchfink et al., 2015) against the Invertebrate RefSeq database (O'Leary et al., 2016) to validate their identity. If their identity was concordant, we checked if they were present in the list of differentially expressed

transcripts and visually explored the normalised count in both colour morphs. If any of the 16 crucial genes for melanin or ommochrome synthesis were absent in our transcriptome annotation, we explored if they were present in other spider genomes or transcriptomes available on the NCBI by performing a TBLASTN against all Araneomorphae (taxid:6905).

RESULTS

Structural differences of the colour variants

The transverse view of the posterior end of the abdomen in both colour morphs revealed that the two black-and-white individuals possess multiple disorganised prismatic crystal structures directly under the cuticle and only in the region where the white colouration was present (Fig. 2). This structure was completely absent in the black individuals. Considering the location and shape of the crystal-like structures, they likely correspond to guanine deposits (G. Oxford, 1997; Kariko et al., 2018).

Gene expression patterns

Our transcriptome assembly consisted of 502,752 transcripts that were reduced to 335,702 transcripts after filtering them by redundancy. All our samples had mapping rates above 90% back to this assembly, confirming the quality of the assembly. The average sequence length was 840 bp, with the longest sequence of 1,026 bp, and an N50 statistic of 395 bp. BUSCO match was 99.1% (single=58.6%, duplicate=40.5%), validating that most of our sequences corresponded to arachnids. We were able to annotate 44,432 (13.24%) of our transcripts, that were not filtered out through frame selection, with most of them corresponding to Araneomorph genes and high similarity with *Argiope bruennichi* (73.14% of the sequences matching to this species).

The PCA plot summarising the differential expression pattern of the model that used developmental stage as a fixed term shows three groups that correspond to the

stages of development we sampled (Fig. 3). The first instar appears to be the most distinct, separated from the developmental stages with abdominal colouration (principal component 1: 38% of variation). This model also shows that the black-and-white individuals in the adult stage differ from the black individuals (Fig. 3), a pattern that is not clear in the third or first instar. In the second model, comparing black-and-white and black individuals during the third instar, we found almost no differentiation (Fig. S2). There were only 10 differentially expressed transcripts for this instar, which are unlikely to be related to colour or structural changes according to our BLAST results (Fig. S3 & S4, Table S1).

When we compared the pattern of gene expression between black-and-white and black adults, we observed that 33% of the variation in gene expression separates the colour variants (Fig. S5). In total, we found 814 highly differentially expressed transcripts (Table S2), with the majority (63%) appearing to be overexpressed in the black-and-white morph compared to the black morph (Fig. 4).

We found that three out of the 814 highly differentially expressed transcripts in adults were also overexpressed in the juveniles. According to our BLAST results, one was 40S ribosomal protein S17-like, one was metalloprotease TIKI1-like, and one was protein turtle-like (Uniprot Q967D7).

Gene ontology (GO) enrichment analysis

Considering that adulthood was the stage that showed a clear pattern of differentiation between colour morphs, we restricted our GO enrichment analysis to the differentially expressed transcripts of these individuals. We found 24 statistically significant terms (adjusted P-value < 0.001; Table 1), of which six are linked to the cuticle: chitin metabolic process, chitin binding, structural constituent of cuticle, structural constituent of chitin-based cuticle, glucosamine-containing compound metabolic process, and aminoglycan metabolic process. These cuticle-associated GO terms group 64 unique transcripts, from which 45 were upregulated in black-and-white adults and 19 were upregulated in black individuals (Fig. 4 & 5).

Melanin and Ommochrome pathway candidate transcripts

In our annotation, we found 104 transcripts associated with the melanin metabolic process GO term (GO:0582). None of these transcripts were significantly differentially expressed in morphs in juveniles. Contrary, we found one transcript that was more common in black adults than in black-and white adults, we identified this transcript as *cation-independent mannose-6-phosphate receptor* (Fig. 4). We found that from the ten melanin pathway genes critical for the development of dark colour in insects (*ebony, black, aaNAT, yellow, tan, pale, DAT, nubbin, prophenoloxidase 3, peroxidase*), seven were present in our annotation (*ebony, black, pale, DAT, nubbin, prophenoloxidase 3, peroxidase*) and two of them (*yellow* and *tan*) were absent in our annotation and also in the available Araneomorphae genomes and transcriptomes on the NCBI. The seven melanin genes found were not significantly differentially expressed in juveniles; however, we found that *peroxidase* was

significantly upregulated in black adults (Fig. 4, Fig. S6 & S7). This is concordant with the visualisation of the normalised counts of these genes in both colour morphs, except for *peroxidase*, which seems to have higher counts in the black adults (Fig. 4; Fig. S6 & S7).

Our transcriptome annotation included 30 transcripts associated with the ommochrome metabolic process GO term (GO:0046152). In juveniles, we found that one transcript identified as *Polyadenylate-binding protein* was more common in black juveniles than in black-and-white juveniles (Log_2 fold change=-23.3, P-value adjusted= 8E-12; Fig. 4). Similarly, in adults we found one transcript that was significantly more common in black-and-white adults than in black individuals (Log_2 fold change=-7.16, P-value adjusted= 1.18E-6; Fig 4). We identified this transcript as *Serine/threonine-protein kinase*. From the six ommochrome synthesis genes common among invertebrates (Figon & Casas, 2019), we identified that five were present in our transcriptome annotation (*Tryptophan 2,3-dioxygenase (TDO), Kynurenine 3-monooxygenase (KMO), Phenoxazinone synthase (PHS), Heme peroxidase, KAT*) and one was absent (*KFase*). We also discarded the presence of *KFase* in the Araneomorphae genomes and transcriptomes available in the NCBI. None of the present genes was differentially expressed between the colour morphs of the juveniles; however, we discovered that *TDO-like* was upregulated in the black-and-white adults (Fig. 4; Log_2 fold change=1.55, P-value adjusted= 1.91E-4) and *PHS-like (peroxidase)* was significantly more abundant in black adults than in black-and-white adults (Fig. 4; Log_2 fold change=-3.20, P-value adjusted= 1.21E-12)

DISCUSSION

Through observations across ontogeny, developmental transcriptomics, and microscopy, we uncovered genetic and structural differences associated with dark and light colour morphs in a widespread species of spider. We found that black and black-and-white adults differ in their gene expression patterns, with transcripts associated with the production of dark colours (melanin and ommochrome metabolism) differentially expressed between morphs. However, our results suggest that cuticle metabolism genes have a bigger effect on morph differentiation, and this could be associated with the presence of guanine crystals in black-and-white adults—a structure completely absent in black individuals. Interestingly, we found that colour differentiation becomes clear during the third instar, when black individuals lose white colouration after moulting. However, at this stage, we did not observe a clear pattern of differential gene expression. Our work constitutes one of the first studies exploring gene regulation of dark and white colouration in arachnids.

Our ontogenetic observations of colour differentiation coincide, both in timing and progression, with previous spider studies where individuals are brown/yellow at hatching, and colour differentiation begins after two or three moults (G. S. Oxford, 1983; G. S. Oxford & Gillespie, 1996; Wagner et al., 2022). After this first stage of colour deposition, where all the spiderlings have black and white abdominal colours, we observed that colour differentiation became clearer; black juveniles lost abdominal white colours, while in black-and-white juveniles white colouration persisted. Because of these changes, we expected to find a clear pattern of differential gene expression in genes associated with either melanin or

ommochromes at this stage. However, this prediction was not met. Instead, we found a non-consistent group of differentially expressed transcripts (Fig. S3), which corresponded mostly to uncharacterized proteins or proteins without a clear function linked to colour or morphological structures (Table S1). This contrasts with reports in insects where crucial genes for melanisation, such as *black, pale, ebony,* and *nubbin*, persist in high numbers in early developmental stages of dark individuals after pigment deposition (Liu et al., 2016; Zhang et al., 2017; Rahman et al., 2021). It is possible that we did not capture differential gene expression between black and black-and-white juveniles, because in spiders colour differentiation occurs in a narrow time window during the previous instar where all the individuals have the same phenotype. This is similar to what is observed in some butterflies and bees, where crucial genes involved in dark coloration are upregulated in the stage before melanin deposition (Connahs et al., 2016; Rahman et al., 2021). We also acknowledge that our methodological approach of using whole body, instead of tissue-specific samples, could mask some patterns of gene expression. Whole body RNAseq is still a common practice in small organisms, but because it includes different cell types with specific expression patterns, results might represent average gene expression of all cell types that does not reflect specific phenotype development (Hoedjes et al., 2024). Future experimental approaches should include stages prior to colour morph differentiation, tissue specific transcriptomics or single-cell RNAseq (scRNAseq) to elucidate clearer patterns of colour developmental gene regulation in spiders (Freedman & Sackton, 2024).

Our analyses revealed differential expression patterns in adults, with 63% pf the transcripts showing higher abundance in black-and-white individuals compared to

black individuals. These differentially expressed transcripts were predominantly associated with cuticle and chitin metabolism, most of which encoded cuticular proteins or proteins with chitin-binding domains. This pattern suggests that cuticular proteins are crucial in maintaining adult colour phenotypic variation, at least in this species of spiders. Cuticular proteins play a fundamental role in determining insect cuticle structure and variation by cross-linking chitin fibres (Togawa et al., 2007; Qiao et al., 2014; Noh et al., 2016; W. Liu et al., 2022). It is possible that the differential expression patterns we observed here reflect cuticle structural variations where white and black colours are exhibited. During our dissections, we observed that the cuticle was transparent in white parts of the abdomen, but not in black areas. This coincides with previous descriptions of a single transparent cuticle layer in parts of the abdomen with guanine crystals (Insausti & Casas, 2008; Levy-Lior et al., 2010; Gawryszewski et al., 2015; Kariko et al., 2018). In contrast, regions with pigments but no guanine crystals show a two-layered structure: an external transparent layer and an internal pigmented layer (Kariko et al., 2018).

In insects, cuticle proteins and associated genes—particularly those with the GO term "structural constituent of cuticle" (GO:0042302)—are upregulated during dark phenotype development in both larvae and adults (Wu et al., 2016; He et al., 2016; Zhang et al., 2017; Wee et al., 2023). While we discovered some cuticular proteins upregulated in black adults, these proteins were more abundant in black-and-white individuals. A possible explanation for these results is that the black-and-white phenotype, which displays a heterogeneous colour palette, might require the upregulation of more genes compared to the black morph, which only has a single

colour. These findings suggest that cuticular protein expression plays an important role in determining colour phenotypes in spiders.

We confirmed the presence of 136 melanin and ommochrome-associated genes in the transcriptome of the Christmas spider, while noting the absence of some well-known genes for insect melanogenesis and ommochrome synthesis in our annotation and in the publicly available Araneomorphae genomic information. In adults, we found six differentially expressed transcripts linked to melanin or ommochrome synthesis when comparing black and black-and-white individuals.

Some of these genes are involved in multiple biological or molecular processes (e.g., Acetoacetyl-CoA synthetase is involved in 107 biological processes; or peroxidase, which participates in multiple oxidation-reduction reactions (Felton & Summers, 1995)), so we cannot conclusively infer that they correspond to differences in melanin or ommochrome synthesis. However, we consider them candidates for such differences, given the absence of other known morphological differences between colour morphs. In contrast, it is worth highlighting the absence of *yellow, tan* and *KFase* genes in the *A. minax* transcriptome. These genes were also reported as absent in an unrelated colour-polymorphic spider (Croucher et al., 2013) , and we also confirmed their absence in the spider genomes and transcriptomes available on the NCBI. *tan* and *yellow* are key regulators of melanin synthesis in many insects, for the synthesis of dopamine, and for converting dopamine and dihydroxyphenylalanine (DOPA) into dopamine melanin and DOPA melanin, respectively (Wittkopp et al., 2003; Wittkopp & Beldade, 2009; J. Liu et al., 2016; Zhang et al., 2017). *KFase,* on the other hand, is necessary for the synthesis of ommochromes, and coverts

formylkynurenine to kynurenine, which is a precursor of ommins and xanthommatin (Osanai-Futahashi et al., 2012; Figon & Casas, 2019). The absence of these key genes, but the presence of melanin and ommochromes in spiders (Seligy, 1972; B. Hsiung et al., 2015), suggests that melanogenesis and ommochrome synthesis in spiders must be different from that in insects.

Altogether, our results reveal a pattern of differential gene expression between two spider phenotypes that involves an overexpression of cuticular proteins in the black-and-white morph. Our microscopy observations suggest that this difference might result from the presence of guanine crystals in individuals with white colours, which are completely absent in dark individuals. Our study confirmed the presence of melanin and ommochrome synthesis genes in spiders. However, we reported the absence of important genes for the synthesis of these pigments not only in *A. minax*, but also in the available spider genomes. This suggests that melanin and ommochrome synthesis pathways in spiders differ from those reported in insects. Since light and dark morphs occur in multiple colour polymorphic spiders (Salgado-Roa et al., 2024), validating our results in other species might elucidate the prevalence and specific role of cuticular proteins in maintaining dark and white colouration. Additionally, future studies should apply more detailed, tissue-specific developmental transcriptomics to identify the gene expression pathway of melanin and ommochromes in spiders.

Acknowledgement

This study received funding from several organizations. We are grateful to the Ecological Society of Australia for providing support through the Holsworth Wildlife Research Endowment. We acknowledge the Society for the Study of Evolution (SSE) for their funding through the Graduate Research Excellence Grant R.C. Lewontin Award. Additionally, we thank the University of Melbourne for their financial support via the Georgina Sweet Scholarship, The Jasper Loftus-Hills Memorial Award, and MacBain Research Scholarship. We also appreciate the feedback on this manuscript from Rebecca Young and Claire McLean.

REFERENCES

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, *12*(1), 59–60. https://doi.org/10.1038/nmeth.3176

Campli, G., Volovych, O., Kim, K., Veldsman, W. P., Drage, H. B., Sheizaf, I., Lynch, S., Chipman, A. D., Daley, A. C., Robinson-Rechavi, M., & Waterhouse, R. M. (2024). The moulting arthropod: a complete genetic toolkit review. *Biological Reviews*, *n/a*(n/a). https://doi.org/https://doi.org/10.1111/brv.13123

Caro, T., & Mallarino, R. (2020). Coloration in Mammals. *Trends in Ecology & Evolution*, *35*(4), 357–366. https://doi.org/https://doi.org/10.1016/j.tree.2019.12.008

Chen, X., Zhang, B., Wang, T., Bonni, A., & Zhao, G. (2020). Robust principal component analysis for accurate outlier sample detection in RNA-Seq data. *BMC Bioinformatics*, *21*(1), 269. https://doi.org/10.1186/s12859-020-03608-0

Connahs, H., Rhen, T., & Simmons, R. B. (2016). Transcriptome analysis of the painted lady butterfly, Vanessa cardui during wing color pattern development. *BMC Genomics*, *17*(1), 270. https://doi.org/10.1186/s12864-016-2586-5

Croucher, P. J. P., Brewer, M. S., Winchell, C. J., Oxford, G. S., & Gillespie, R. G. (2013). De novo characterization of the gene-rich transcriptomes of two color-

polymorphic spiders, Theridion grallator and T. californicum (Araneae: Theridiidae), with special reference to pigment genes. *BMC Genomics*, *14*(1), 862. https://doi.org/10.1186/1471-2164-14-862

Cuthill, I. C., Allen, W. L., Arbuckle, K., Caspers, B., Chaplin, G., Hauber, M. E., Hill, G. E., Jablonski, N. G., Jiggins, C. D., Kelber, A., Mappes, J., Marshall, J., Merrill, R., Osorio, D., Prum, R., Roberts, N. W., Roulin, A., Rowland, H. M., Sherratt, T. N., … Caro, T. (2017). The biology of color. *Science*, *357*(6350). https://doi.org/10.1126/science.aan0221

Elkin, J., Martin, A., Courtier-Orgogozo, V., & Santos, M. E. (2023). Analysis of the genetic loci of pigment pattern evolution in vertebrates. *Biological Reviews*, *98*(4), 1250–1277. https://doi.org/https://doi.org/10.1111/brv.12952

Felton, G. W., & Summers, C. B. (1995). Antioxidant systems in insects. *Archives of Insect Biochemistry and Physiology*, *29*(2), 187–197. https://doi.org/https://doi.org/10.1002/arch.940290208

Figon, F., & Casas, J. (2019). Ommochromes in invertebrates: biochemistry and cell biology. *Biological Reviews*, *94*(1), 156–183. https://doi.org/https://doi.org/10.1111/brv.12441

Freedman, A. H., & Sackton, T. B. (2024). Rethinking eco-evo studies of gene expression for non-model organisms in the genomic era. *Molecular Ecology*, *n/a*(n/a), e17378. https://doi.org/https://doi.org/10.1111/mec.17378

Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics*, *28*(23), 3150–3152. https://doi.org/10.1093/bioinformatics/bts565

Gawryszewski, F. M., Birch, D., Kemp, D. J., & Herberstein, M. E. (2015). Dissecting the variation of a visual trait: the proximate basis of UV-Visible reflectance in crab spiders (Thomisidae). *Functional Ecology*, *29*(1), 44–54. https://doi.org/10.1111/1365-2435.12300

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., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883

Gray, S. M., & McKinnon, J. S. (2007). Linking color polymorphism maintenance and speciation. *Trends in Ecology and Evolution*, *22*(2), 71–79. https://doi.org/10.1016/j.tree.2006.10.005

Hart, A. J., Ginzburg, S., Xu, M. (Sam), Fisher, C. R., Rahmatpour, N., Mitton, J. B., Paul, R., & Wegrzyn, J. L. (2020). EnTAP: Bringing faster and smarter functional annotation to non-model eukaryotic transcriptomes. *Molecular Ecology Resources*, *20*(2), 591–604. https://doi.org/https://doi.org/10.1111/1755-0998.13106

He, S., Tong, X., Lu, K., Lu, Y., Luo, J., Yang, W., Chen, M., Han, M., Hu, H., Lu, C., & Dai, F. (2016). Comparative Analysis of Transcriptomes among Bombyx mori Strains and Sexes Reveals the Genes Regulating Melanic Morph and the Related Phenotypes. *PLOS ONE*, *11*(5), e0155061. https://doi.org/10.1371/journal.pone.0155061

Hoedjes, K. M., Grath, S., Posnien, N., Ritchie, M. G., Schlötterer, C., Abbott, J. K.,

Almudi, I., Coronado-Zamora, M., Durmaz Mitchell, E., Flatt, T., Fricke, C., Glaser-Schmitt, A., González, J., Holman, L., Kankare, M., Lenhart, B., Orengo, D. J., Snook, R. R., Yılmaz, V. M., & Yusuf, L. (2024). From whole bodies to single cells: A guide to transcriptomic approaches for ecology and evolutionary biology. *Molecular Ecology*, *n/a*(n/a), e17382. https://doi.org/https://doi.org/10.1111/mec.17382

Hsiung, B.-K., Blackledge, T. A., & Shawkey, M. D. (2015). Spiders do have melanin after all. *Journal of Experimental Biology*, *218*(22), 3632–3635. https://doi.org/10.1242/jeb.128801

Hsiung, B., Blackledge, T. A., & Shawkey, M. D. (2015). Spiders have melanin after all. *Journal of Experimental Biology*, *October*. https://doi.org/10.1242/jeb.128801

Hsiung, B. K., Justyn, N. M., Blackledge, T. A., & Shawkey, M. D. (2017). Spiders have rich pigmentary and structural colour palettes. *Journal of Experimental Biology*, *220*(11), 1975–1983. https://doi.org/10.1242/jeb.156083

Hubbard, J. K., Uy, J. A. C., Hauber, M. E., Hoekstra, H. E., & Safran, R. J. (2010). Vertebrate pigmentation: from underlying genes to adaptive function. *Trends in Genetics*, *26*(5), 231–239. https://doi.org/https://doi.org/10.1016/j.tig.2010.02.002

Insausti, T. C., & Casas, J. (2008). The functional morphology of color changing in a spider: development of ommochrome pigment granules. *Journal of Experimental Biology*, *211*(5), 780–789. https://doi.org/10.1242/jeb.014043

Kariko, S., Timonen, J. V. I., Weaver, J. C., Gur, D., Marks, C., Leiserowitz, L., Kolle, M., & Li, L. (2018). Structural origins of coloration in the spider Phoroncidia rubroargentea Berland, 1913 (Araneae: Theridiidae) from Madagascar. *Journal*

of The Royal Society Interface, *15*(139), 20170930. https://doi.org/10.1098/rsif.2017.0930

Kollias, N., Sayre, R. M., Zeise, L., & Chedekel, M. R. (1991). New trends in photobiology: Photoprotection by melanin. *Journal of Photochemistry and Photobiology B: Biology*, *9*(2), 135–160. https://doi.org/https://doi.org/10.1016/1011-1344(91)80147-A

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Levy-Lior, A., Shimoni, E., Schwartz, O., Gavish-Regev, E., Oron, D., Oxford, G., Weiner, S., & Addadi, L. (2010). Guanine-Based Biogenic Photonic-Crystal Arrays in Fish and Spiders. *Advanced Functional Materials*, *20*(2), 320–329. https://doi.org/https://doi.org/10.1002/adfm.200901437

Lin, S., Zhang, L., Wang, G., Huang, S., & Wang, Y. (2022). Searching and identifying pigmentation genes from Neocaridina denticulate sinensis via comparison of transcriptome in different color strains. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, *42*, 100977. https://doi.org/https://doi.org/10.1016/j.cbd.2022.100977

Liu, J., Lemonds, T. R., Marden, J. H., & Popadić, A. (2016). A Pathway Analysis of Melanin Patterning in a Hemimetabolous Insect. *Genetics*, *203*(1), 403–413. https://doi.org/10.1534/genetics.115.186684

Liu, W., Chang, T., Zhao, K., Sun, X., Qiao, H., Yan, C., & Wang, Y. (2022). Genome-wide annotation of cuticular protein genes in non-biting midge Propsilocerus akamusi and transcriptome analysis of their response to heavy metal pollution. *International Journal of Biological Macromolecules*, *223*, 555–

566. https://doi.org/https://doi.org/10.1016/j.ijbiomac.2022.10.279

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Majerus, M. E. N. (1998). *Melanism: Evolution in Action*. Oxford University Press.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal; Vol 17, No 1: Next Generation Sequencing Data Analysis*. https://doi.org/10.14806/ej.17.1.200

McNamara, M. E., Rossi, V., Slater, T. S., Rogers, C. S., Ducrest, A.-L., Dubey, S., & Roulin, A. (2021). Decoding the Evolution of Melanin in Vertebrates. *Trends in Ecology & Evolution*, *36*(5), 430–443. https://doi.org/https://doi.org/10.1016/j.tree.2020.12.012

Noh, M. Y., Muthukrishnan, S., Kramer, K. J., & Arakane, Y. (2016). Cuticle formation and pigmentation in beetles. *Current Opinion in Insect Science*, *17*, 1–9. https://doi.org/https://doi.org/10.1016/j.cois.2016.05.004

Nosanchuk, J. D., & Casadevall, A. (2003). The contribution of melanin to microbial pathogenesis. *Cellular Microbiology*, *5*(4), 203–223. https://doi.org/https://doi.org/10.1046/j.1462-5814.2003.00268.x

Nosil, P., Villoutreix, R., De Carvalho, C. F., Farkas, T. E., Soria-Carrasco, V., Feder, J. L., Crespi, B. J., & Gompert, Z. (2018). Natural selection and the predictability of evolution in timema stick insects. *Science*, *359*(6377), 765–770. https://doi.org/10.1126/science.aap9125

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R.,

Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., … Pruitt, K. D. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*, *44*(D1), D733–D745. https://doi.org/10.1093/nar/gkv1189

Orteu, A., & Jiggins, C. D. (2020). The genomics of coloration provides insights into adaptive evolution. *Nature Reviews Genetics*, *21*(8), 461–475. https://doi.org/10.1038/s41576-020-0234-z

Osanai-Futahashi, M., Tatematsu, K., Yamamoto, K., Narukawa, J., Uchino, K., Kayukawa, T., Shinoda, T., Banno, Y., Tamura, T., & Sezutsu, H. (2012). Identification of the Bombyx Red Egg Gene Reveals Involvement of a Novel Transporter Family Gene in Late Steps of the Insect Ommochrome Biosynthesis Pathway *♦ *Journal of Biological Chemistry*, *287*(21), 17706–17714. https://doi.org/10.1074/jbc.M111.321331

Oxford, G. (1997). Guanine as a colorant in spiders: development, genetics, phylogenetics and ecology. *Proceedings of the 17th European Colloquium of Arachnology, Edinburgh*, 121–131.

Oxford, G. S. (1983). Genetics of colour and its regulation during development in the spider Enoplognatha ovata (Clerck) (Araneae: Theridiidae). *Heredity*, *51*(3), 621–634. https://doi.org/10.1038/hdy.1983.74

Oxford, G. S., & Gillespie, R. G. (1996). Genetics of a colour polymorphism in Theridion grallator (Araneae: Theridiidae), the Hawaiian happy-face spider, from Greater Maui. *Heredity*, *76*, 238–248. https://doi.org/10.1038/hdy.1996.37

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, *14*(4), 417–419. https://doi.org/10.1038/nmeth.4197

Qiao, L., Xiong, G., Wang, R., He, S., Chen, J., Tong, X., Hu, H., Li, C., Gai, T., Xin, Y., Liu, X., Chen, B., Xiang, Z., Lu, C., & Dai, F. (2014). Mutation of a Cuticular Protein, BmorCPR2, Alters Larval Body Shape and Adaptability in Silkworm, Bombyx mori. *Genetics*, *196*(4), 1103–1115. https://doi.org/10.1534/genetics.113.158766

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1), D590–D596. https://doi.org/10.1093/nar/gks1219

Raghavan, V., Kraft, L., Mesny, F., & Rigerte, L. (2022). A simple guide to de novo transcriptome assembly and annotation. *Briefings in Bioinformatics*, *23*(2), 1–30. https://doi.org/10.1093/bib/bbab563

Rahman, S. R., Terranova, T., Tian, L., & Hines, H. M. (2021). Developmental Transcriptomics Reveals a Gene Network Driving Mimetic Color Variation in a Bumble Bee. *Genome Biology and Evolution*, *13*(6), evab080. https://doi.org/10.1093/gbe/evab080

Salgado-Roa, F. C., Stuart-Fox, D., White, T. E., & Medina, I. (2024). Colour polymorphism is prevalent on islands but shows no association with range size in web-building spiders. *Journal of Evolutionary Biology*, voae118. https://doi.org/10.1093/jeb/voae118

Seligy, V. L. (1972). Ommochrome pigments of spiders. *Comparative Biochemistry*

and Physiology Part A: Physiology, *42*(3), 699–709. https://doi.org/https://doi.org/10.1016/0300-9629(72)90448-3

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V, & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Soneson, C., Love, M. I., & Robinson, M. D. (2016). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences [version 2; peer review: 2 approved] . *F1000Research*, *4*(1521). https://doi.org/10.12688/f1000research.7563.2

Song, L., & Florea, L. (2015). Rcorrector: efficient and accurate error correction for Illumina RNA-seq reads. *GigaScience*, *4*(1), s13742-015-0089-y. https://doi.org/10.1186/s13742-015-0089-y

Stephens, M. (2017). False discovery rates: a new deal. *Biostatistics*, *18*(2), 275–294. https://doi.org/10.1093/biostatistics/kxw041

Sugumaran, M., & Barek, H. (2016). Critical Analysis of the Melanogenic Pathway in Insects and Higher Animals. In *International Journal of Molecular Sciences* (Vol. 17, Issue 10). https://doi.org/10.3390/ijms17101753

Svensson, E. I. (2017). Back to basics: Using colour polymorphisms to study evolutionary processes. *Molecular Ecology*, *26*(8), 2204–2211. https://doi.org/10.1111/mec.14025

Togawa, T., Augustine Dunn, W., Emmons, A. C., & Willis, J. H. (2007). CPF and CPFL, two related gene families encoding cuticular proteins of Anopheles

gambiae and other insects. *Insect Biochemistry and Molecular Biology*, *37*(7), 675–688. https://doi.org/https://doi.org/10.1016/j.ibmb.2007.03.011

True, J. R. (2003). Insect melanism: The molecules matter. *Trends in Ecology and Evolution*, *18*(12), 640–647. https://doi.org/10.1016/j.tree.2003.09.006

Wagner, A., Ezersky, V., Maria, R., Upcher, A., Lemcoff, T., Aflalo, E. D., Lubin, Y., & Palmer, B. A. (2022). The Non-Classical Crystallization Mechanism of a Composite Biogenic Guanine Crystal. *Advanced Materials*, *34*(31), 2202242. https://doi.org/https://doi.org/10.1002/adma.202202242

Wang, J., Bai, Y., Zhao, H., Mu, R., & Dong, Y. (2021). Reinvestigating the phylogeny of Myriapoda with more extensive taxon sampling and novel genetic perspective. *PeerJ*, *9*, e12691. https://doi.org/10.7717/peerj.12691

Wee, J. L. Q., Murugesan, S. N., Wheat, C. W., & Monteiro, A. (2023). The genetic basis of wing spots in Pieris canidia butterflies. *BMC Genomics*, *24*(1), 169. https://doi.org/10.1186/s12864-023-09261-0

Wittkopp, P. J., & Beldade, P. (2009). Development and evolution of insect pigmentation: Genetic mechanisms and the potential consequences of pleiotropy. *Seminars in Cell & Developmental Biology*, *20*(1), 65–71. https://doi.org/https://doi.org/10.1016/j.semcdb.2008.10.002

Wittkopp, P. J., Carroll, S. B., & Kopp, A. (2003). Evolution in black and white: genetic control of pigment patterns in Drosophila. *Trends in Genetics*, *19*(9), 495–504. https://doi.org/https://doi.org/10.1016/S0168-9525(03)00194-X

Wu, S., Tong, X., Peng, C., Xiong, G., Lu, K., hu, H., Tan, D., Li, C., Han, M., Lu, C., & Dai, F. (2016). Comparative analysis of the integument transcriptomes of the

black dilute mutant and the wild-type silkworm Bombyx mori. *Scientific Reports*, *6*(1), 26114. https://doi.org/10.1038/srep26114

Young, M. D., Wakefield, M. J., Smyth, G. K., & Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology*, *11*(2), R14. https://doi.org/10.1186/gb-2010-11-2-r14

Zhang, L., Martin, A., Perry, M. W., van der Burg, K. R. L., Matsuoka, Y., Monteiro, A., & Reed, R. D. (2017). Genetic Basis of Melanin Pigmentation in Butterfly Wings. *Genetics*, *205*(4), 1537–1550. https://doi.org/10.1534/genetics.116.196451

Table 1. Statistically significant overrepresented gene ontology (GO) terms in adults. GO terms that are more common than expected by chance, associated with the 814 differentially expressed transcripts between black-and-white vs. black adults.

Category	Term	Overrepresented P-value	Overrepresented adjusted P-value
GO:0005576	extracellular region	9.08E-21	1.04E-16
GO:0006030	chitin metabolic process	1.59E-19	9.11E-16
GO:0008061	chitin binding	3.91E-19	1.50E-15
GO:0042302	structural constituent of cuticle	7.55E-19	2.16E-15
GO:1901071	glucosamine-containing compound metabolic process	2.38E-17	5.46E-14
GO:0006040	amino sugar metabolic process	2.41E-16	4.60E-13
GO:0006022	aminoglycan metabolic process	4.94E-16	8.09E-13
GO:0097367	carbohydrate derivative binding	1.88E-15	2.70E-12
GO:0005198	structural molecule activity	5.43E-12	6.91E-09
GO:0036379	myofilament	7.92E-11	9.08E-08
GO:0030016	myofibril	1.31E-10	1.36E-07

GO:0043292	contractile fiber	1.71E-10	1.63E-07
GO:0030017	sarcomere	5.91E-10	4.84E-07
GO:0005214	structural constituent of chitin-based cuticle	1.00E-08	7.67E-06
GO:0004857	enzyme inhibitor activity	3.12E-07	2.24E-04
GO:0030239	myofibril assembly	3.70E-07	2.49E-04
GO:0005615	extracellular space	4.49E-07	2.71E-04
GO:0005865	striated muscle thin filament	5.70E-07	3.26E-04
GO:0015629	actin cytoskeleton	6.46E-07	3.38E-04
GO:0031032	actomyosin structure organization	6.49E-07	3.38E-04
GO:0044092	negative regulation of molecular function	7.19E-07	3.58E-04
GO:0004866	endopeptidase inhibitor activity	9.14E-07	4.36E-04
GO:0061135	endopeptidase regulator activity	1.07E-06	4.88E-04
GO:0030414	peptidase inhibitor activity	1.75E-06	7.73E-04

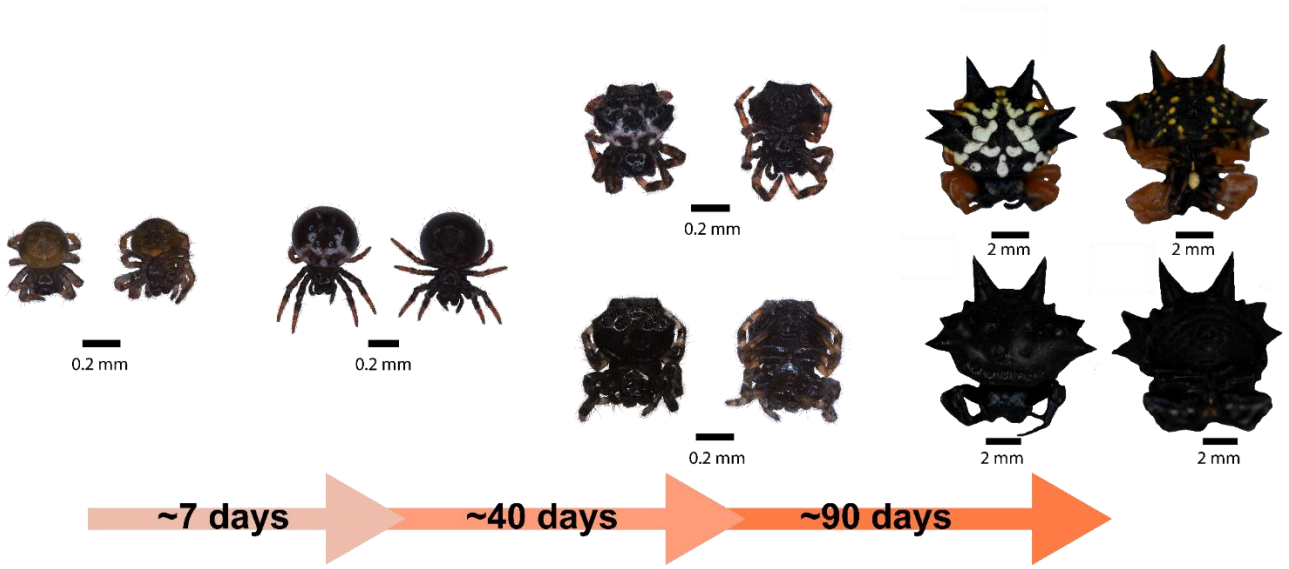

Figure 1. Ontogenetic colour changes in Austracantha minax. First instars lack apparent coloration (n=5). Second instars all exhibit similar black and white abdominal patterns. Third instars can be differentiated into black (n=8) and black-and-white (n=9) morphs. Adults shown for both black (n=5) and black-and-white morphs (n=5).

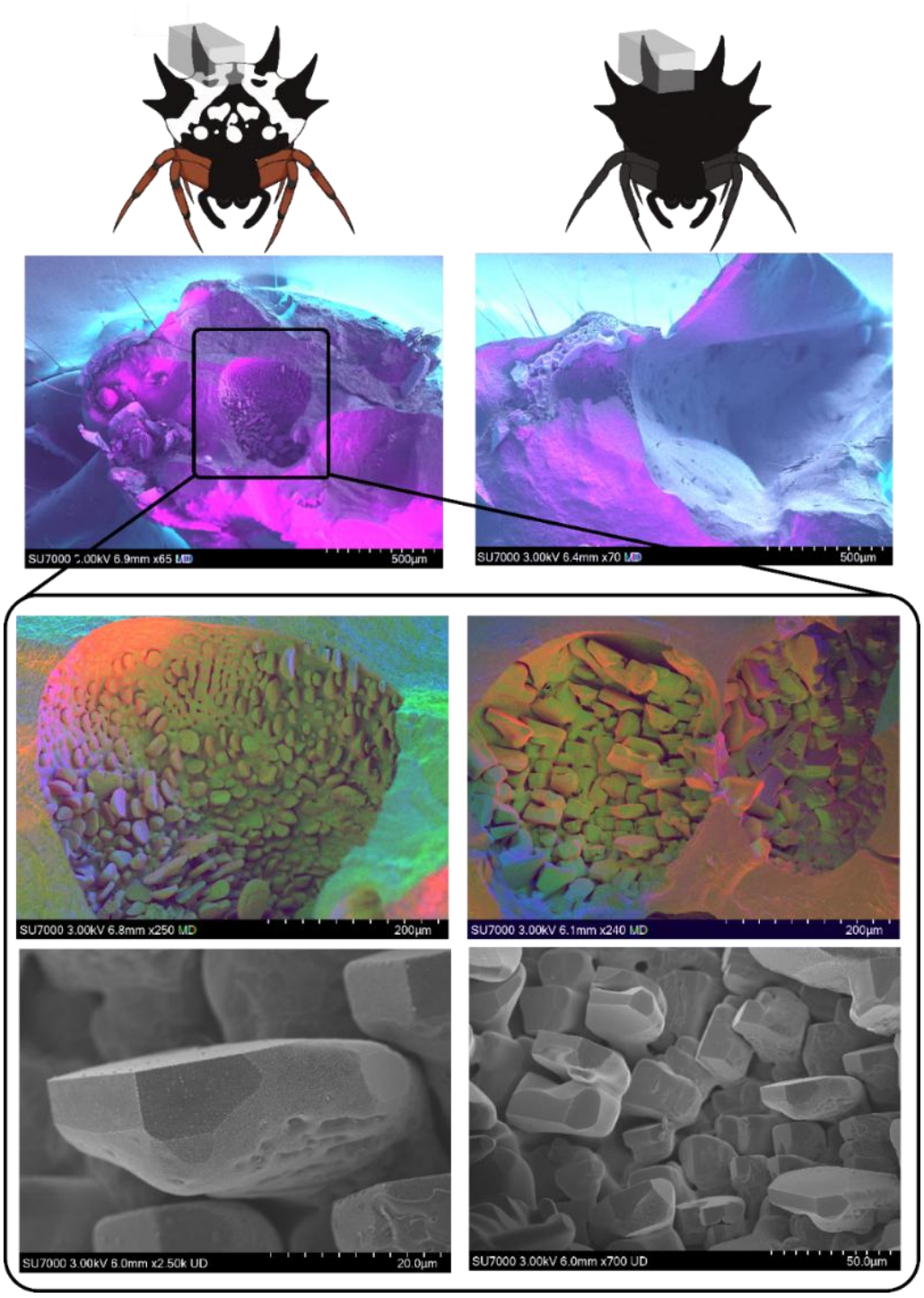

Figure 2. Cryo-SEM transverse view of the posterior part of the abdomen in black-and-white and black adults. The cubes in the first row highlight where the dissection was made. The second row highlights the presence of crystal-like structures (potentially guanine crystals) in the black-and-white individuals but their absence in the black individuals. The other images show different magnifications of these structures.

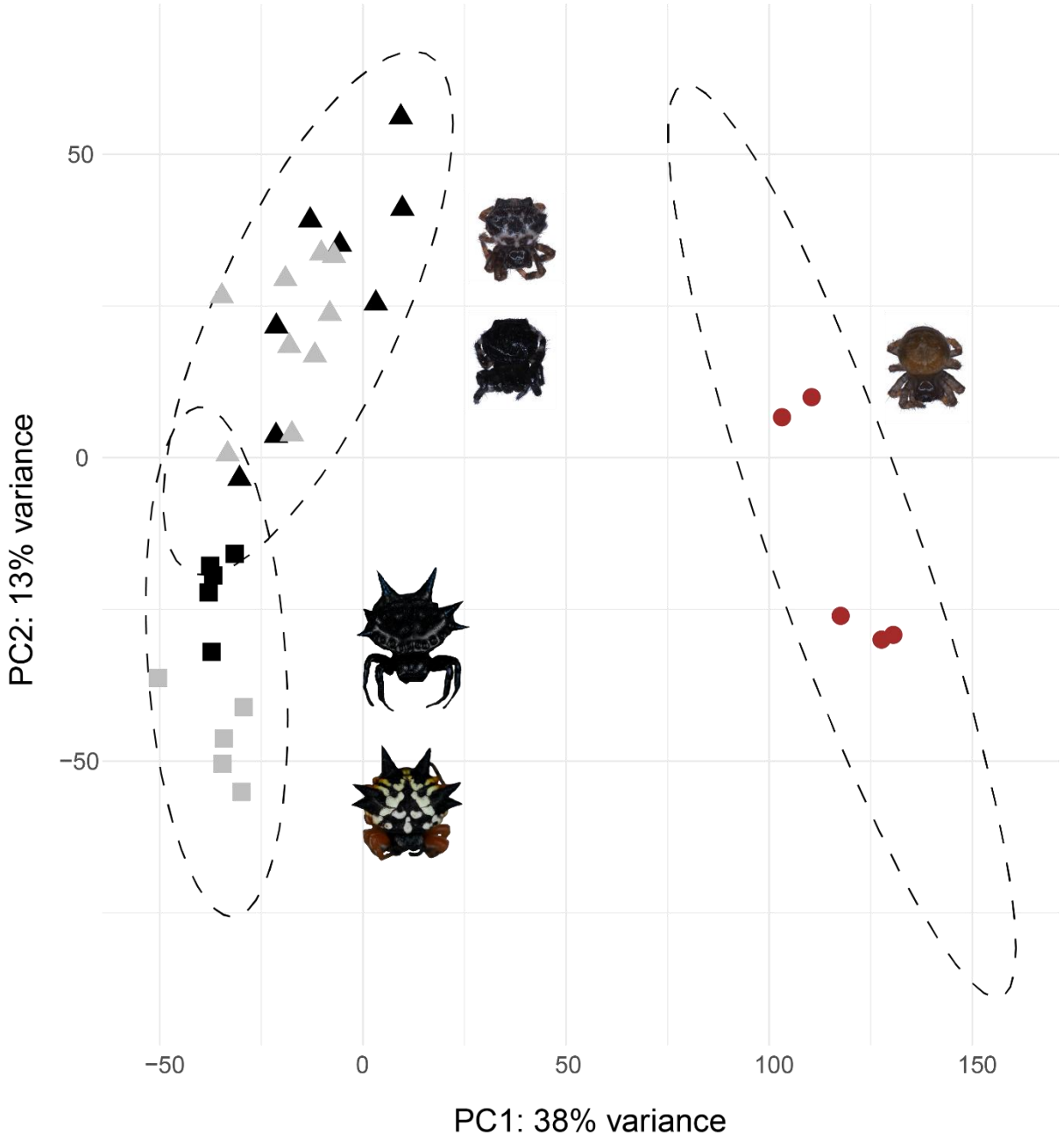

Figure 3. PCA plot summarising the variation of the top 10000 differentially expressed transcripts across three developmental stages. Red circles represent individuals of the first instar where the spiders do not have white or black colour. Triangles correspond to juveniles in the first stage where we can discriminate between colour morphs, third instar. Squares correspond to adults. Grey symbols symbolise black-and-white individuals. Black symbols symbolise black individuals. The ellipsoids delimit the three developmental stages.

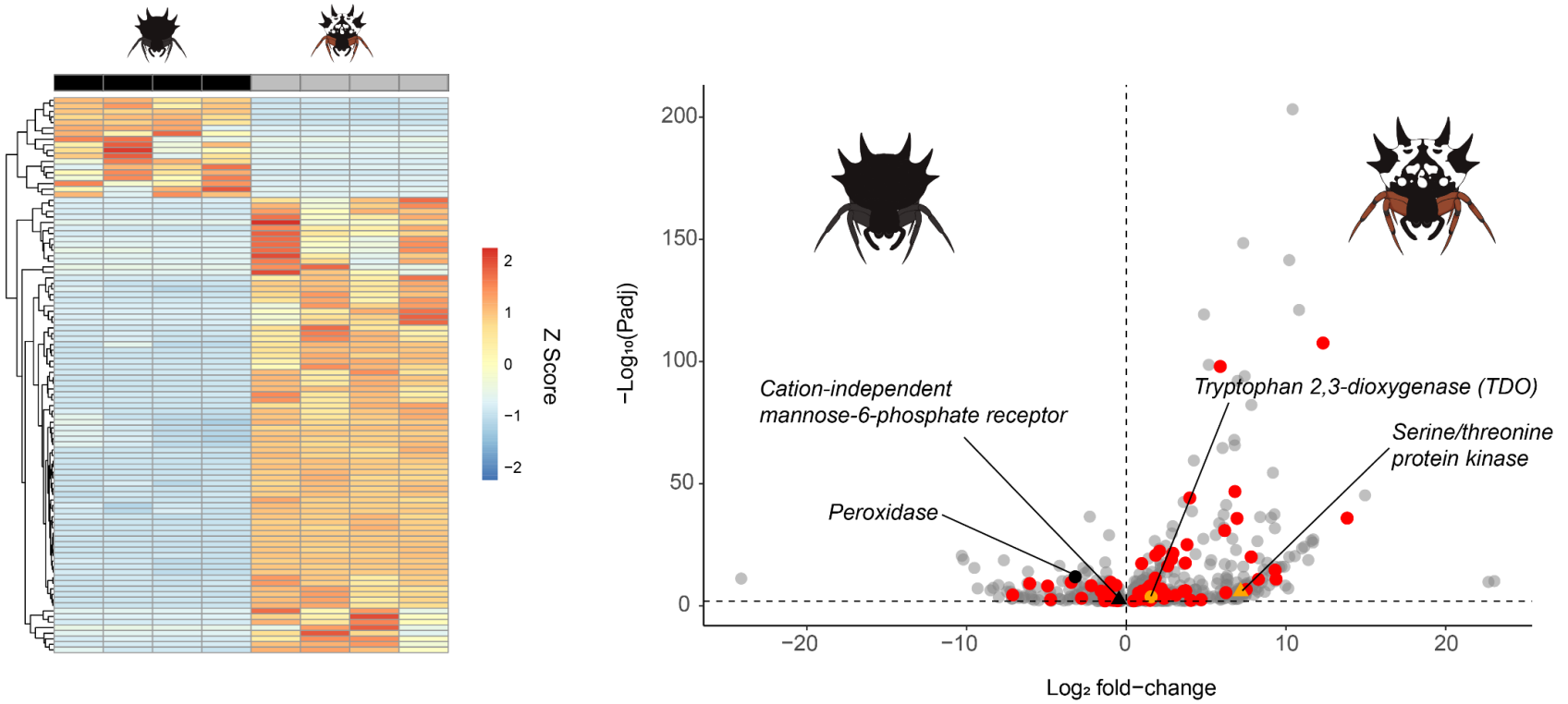

Figure 4. Differential gene expression in adults. The left panel shows the pattern of expression in the top 100 differentially expressed transcripts. Warm colours represent high expression, cold colours represent low expression. Grey columns symbolise black-and-white individuals. Black columns symbolise black individuals. The right panel is a volcano plot showing the association of statistical significance and fold change. Positive values on the x-axis show transcripts that are overexpressed in the black-and-white individuals, and negative values on the same axis represent transcripts that are overexpressed in the black morph. Gray dots are the transcripts that pass the statistically significant (FDR-adjusted p- value <0.01) and fold change thresholds ($0<\log_2 FC>0$). Red dots are the transcripts associate with chitin or cuticle metabolic processes. Black triangles and circles represent melanin-associated transcripts identified in our transcriptome annotation via GO term (GO:0006582) and specific BLAST search, respectively. Orange triangles and circles represent ommochrome-associated transcripts identified via GO term (GO:0046152) and specific BLAST search, respectively.

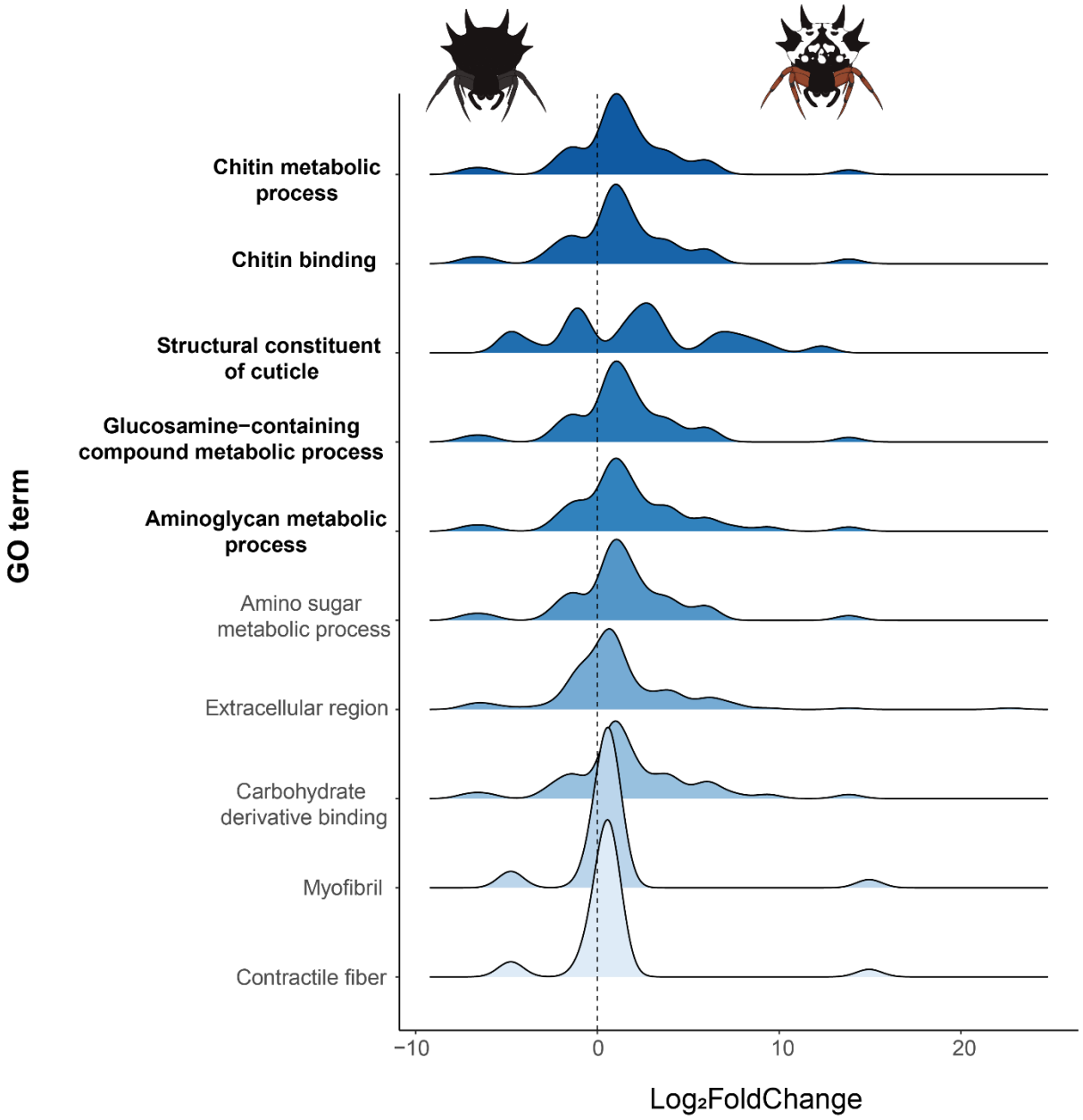

Figure 5. Relationship between the top 10 overrepresented gene ontology (GO) terms and the fold change between black-and-white and black individuals. Density plots show the distribution of fold change values of the differentially expressed transcripts in each GO term. Positive values symbolise higher expression in the black-and-white morph compared to the black morph, while negative values indicate higher expression in the black individuals than in the black-and-white individuals. The dotted line represents the limit between positive and negative values. GOterms in bold are associated with cuticle metabolic processes.