JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.13079

Signals of selection in conditionally expressed genes in the diversification of three horned beetle species

M. H. PESPENI*† , J. T. LADNER‡ & A. P. MOCZEK*

*Department of Biology, Indiana University, Bloomington, IN, USA
†Department of Biology, University of Vermont, Burlington, VT, USA
‡Center for Genome Sciences, United States Army Medical Research Institute of Infectious Diseases, Fort Detrick, MD, USA

Keywords:

comparative genomics;
polyphenism;
positive selection;
relaxed selection;
sexual dimorphism.

Abstract

Species radiations may be facilitated by phenotypic differences already present within populations, such as those arising through sex-specific development or developmental processes biased towards particular reproductive or trophic morphs. We sought to test this hypothesis by utilizing a comparative transcriptomic approach to contrast among- and within-species differentiation using three horned beetle species in the genus *Onthophagus*. These three species exhibit differences along three phenotypic axes reflective of much of the interspecific diversity present within the genus: horn location, polarity of sexual dimorphism and degree of nutritional sensitivity. Our approach combined *de novo* transcript assembly, assessment of amino acid substitutions (dN/dS) across orthologous gene pairs and integration of gene function and conditional gene expression data. We identified 17 genes across the three species pairs related to axis patterning, development and metabolism with dN/dS > 1 and detected elevated dN/dS in genes related to metabolism and biosynthesis in the most closely related species pair, which is characterized by a loss of nutritional polyphenism and a reversal of sexual dimorphism. Further, we found that genes that are conditionally expressed (i.e. as a function of sex, nutrition or body region) within one of our focal species also showed significantly stronger signals of positive or relaxed purifying selection between species divergent along the same morphological axis (i.e. polarity of sexual dimorphism, degree of nutritional sensitivity or location of horns). Our findings thus reveal a positive relationship between intraspecific differentiation due to condition-specific development and genetic divergences among species.

Introduction

A fundamental goal in evolutionary biology is to identify the molecular mechanisms that enable the emergence of novel complex traits and facilitate their subsequent diversification (Carroll, 2008; Kratochwil & Meyer, 2014). While linking phenotypes to underlying genotypes has a long history in domesticated plants and animals (Sax, 1923), recent advances in sequencing technologies have made addressing evolutionary and ecological questions feasible in both model and emerging model organisms (Ekblom & Galindo, 2011; Ellegren, 2014). Furthermore, if applied to instances of parallel phenotypic evolution, for example across multiple sets of populations or closely related species, comparative genomic approaches provide powerful opportunities to reveal repeated targets of natural selection (Elmer & Meyer, 2011; Martin & Orgogozo, 2013), as for instance in stickleback fish (Hohenlohe *et al.*, 2010), stick insects (Soria-Carrasco *et al.*, 2014), or Darwin's finches (Lamichhaney *et al.*, 2015). In particular, genomic studies in highly speciose clades that exhibit morphological variation both between and within species have strong potential to provide important insights into the genes and pathways that facilitate and bias the origin and diversification of novel traits and trait

Correspondence: Melissa H. Pespeni, Department of Biology, University of Vermont, 109 Carrigan Drive, Burlington, VT 05403, USA.
Tel.: +802 656 0628; fax: +802 656 2914; e-mail: mpespeni@uvm.edu

© 2017 EUROPEAN SOCIETY FOR EVOLUTIONARY BIOLOGY. *J. EVOL. BIOL.* **30** (2017) 1644–1657
JOURNAL OF EVOLUTIONARY BIOLOGY © 2017 EUROPEAN SOCIETY FOR EVOLUTIONARY BIOLOGY

variants (Whitehead & Crawford, 2006; Wund *et al.*, 2008; Pfennig *et al.*, 2010).

Divergence between species is enabled by phenotypic variation present in natural populations (Gompel *et al.*, 2005; Wittkopp *et al.*, 2009; Sanger *et al.*, 2013; Carlson *et al.*, 2015) and biased along both genetic and developmental lines of least resistance (Schluter, 1996), such as those arising through sex-specific development or developmental processes biased towards particular reproductive or trophic morphs (Lorch *et al.*, 2003; West-Eberhard, 2003; Smith *et al.*, 2008; Aubin-Horth & Renn, 2009). For example, many invertebrate and vertebrate species have diverged most obviously in the expression of secondary sexual traits, which also account for the greatest phenotypic variation within populations (Emlen *et al.*, 2012). Similarly, many social insect species have diverged from each other in the disparity present among casts, which at the same time accounts for the greatest amount of phenotypic variation present within species (Wheeler, 1986). A growing body of work is beginning to clarify the genetic and developmental mechanisms underlying these biases at the molecular level. For example, numerous studies have found a positive relationship between biased gene expression and protein sequence divergence in a wide range of taxa (Conifers (Hodgins *et al.*, 2016), *Arabidopsis* (Gossmann *et al.*, 2014), *Neurospora* (Whittle *et al.*, 2014), *Drosophila* (Nuzhdin *et al.*, 2004; Lemos *et al.*, 2005; Larracuente *et al.*, 2008; Perry *et al.*, 2014), pea aphids (Purandare *et al.*, 2014), honeybee (Harpur *et al.*, 2014), ants (Hunt *et al.*, 2013; Roux *et al.*, 2014; Mikheyev & Linksvayer, 2015), urchins (Oliver *et al.*, 2010), spadefoot toads (Leichty *et al.*, 2012), mammals (Khaitovich *et al.*, 2005; Warnefors & Kaessmann, 2013)), lending broad support to a long-standing hypothesis that conditional gene expression may promote rapid evolution due to a reduction in pleiotropic constraint (West-Eberhard, 1989; Leichty *et al.*, 2012). Alternatively, the order of events may be reversed; that is, fast-evolving genes that experience low levels of purifying selection could be more readily co-opted into conditional gene expression patterns over time due to neutral rather than adaptive processes (Khaitovich *et al.*, 2005; Hunt *et al.*, 2011; Helantera & Uller, 2014). Just as important, several studies failed to find a relationship between biased gene expression and protein evolution (e.g. ants (Smith *et al.*, 2015), sunflowers (Moyers & Rieseberg, 2013), copepods (Barreto *et al.*, 2015), cichlids (Kavembe *et al.*, 2015)) or revealed a negative relationship between gene expression plasticity and adaptive population divergence (guppies (Ghalambor *et al.*, 2015)). However, few if any studies have tested for a positive relationship between conditional gene expression and protein sequence divergence in species where there exists extensive phenotypic variation in multiple traits both within and between species. This may be due to the difficulty in finding taxa that combine relevant focal traits and patterns of relatedness with experimental accessibility. Here, we take advantage of such an opportunity in the horned beetle genus *Onthophagus* to investigate the molecular mechanisms underlying repeated parallel phenotypic divergences at different levels of biological organization.

Horned beetles in the genus *Onthophagus* contain high levels of intra- and interspecific diversity in the expression of secondary sexual traits (Emlen *et al.*, 2005a,b), resulting in dramatic sexual dimorphisms as well as nutritionally cued alternative male morphs in many species. Previous studies have provided evidence that both morph- and sex-biased genes are more evolutionary labile than genes with unbiased expression (Snell-Rood *et al.*, 2011; Warren *et al.*, 2014). Specifically, Snell-Rood *et al.* found that morph-biased expression of different genes between horned (major, fighter) and hornless (minor, sneaker) male morphs of *O. taurus* was positively related to amino acid divergence (dA) between *O. taurus* and the distantly related flour beetle, *Tribolium castaneum* (Snell-Rood *et al.*, 2011). Using the same *O. taurus* data set, Warren *et al.*, 2014 detected a modest but significantly positive relationship between morph- and sex-biased expression and dN/dS for comparisons to the congener *O. nigriventris* as well as *Trypoxylus dichotomus*, a member of a separate scarab subfamily which independently evolved horns (Warren *et al.*, 2014). Although these studies provided a first indication of a possible link between conditional gene expression and accelerated protein evolution, they also had to contend with significant limitations, including modest orthologous gene sets (Snell-Rood *et al.*, 2011; Warren *et al.*, 2014) and differential transcript data obtained from first-generation custom microarrays (*O. taurus*, Snell-Rood *et al.*, 2011). As a consequence, existing studies remain limited in their ability to fully resolve the relationships (if any) between sex-biased, morph-biased or body region-biased expression to protein evolution. Here, we seek to assess these relative contributions by investigating and integrating transcriptome-wide sequence divergence across three *Onthophagus* species, *O. taurus, O. nigriventris* and *O. sagittarius*, with transcriptome-wide gene expression conditional upon sex, body region and nutritional status.

We selected these three focal species because phenotypic differences among them reflect the three axes of phenotypic diversification that dominate the morphological radiation across the genus *Onthophagus*: (i) *horn position*: *Onthophagus* horns differ in their position and develop either on the head (as in *O. taurus* males), thorax (*O. nigriventris* males) or both (*O. sagittarius* females); (ii) *orientation of sexual dimorphism*: horn formation is often restricted to or most exaggerated in males (*O. nigriventris, O. taurus*), but can also be sex-reversed and female-biased (*O. sagittarius*); and (iii) *nutritional responsiveness of horn formation within sexes*: *Onthophagus* species vary in the degree to which horn

growth is sensitive to nutritional conditions, resulting in lesser or greater disparity between alternative, nutritionally cued morphs among species (Emlen *et al.*, 2005a). Nutritionally cued male polyphenism is most extreme in *O. nigriventris*, followed by *O. taurus*, and has been secondarily lost in *O. sagittarius*. Importantly, the most phenotypically divergent species *O. taurus* and *O. sagittarius* shared a common ancestor as recent as 10 MYA, whereas *O. taurus* and *O. nigriventris* diverged approximately 30 MYA (Emlen *et al.*, 2005b; Fig. 1).

By comparing patterns of sequence divergence among these three species, we sought to identify and characterize the genomic underpinnings of species-level morphological divergence in this highly diverse genus. At the same time, we sought to use the degree of trait variation among these three species to test whether there was a positive relationship with biased gene expression between morphs, sexes and/or body regions within a species. To relate conditional gene expression to sequence evolution across these three species, we first generated a high coverage transcriptome assembly *de novo* for *O. sagittarius* and then contrasted these transcripts to those already existing for *O. taurus* and *O. nigriventris* (Snell-Rood *et al.*, 2011; Warren *et al.*, 2014). We calculated amino acid sequence divergence (dN/dS) and tested for elevated substitution rates as a

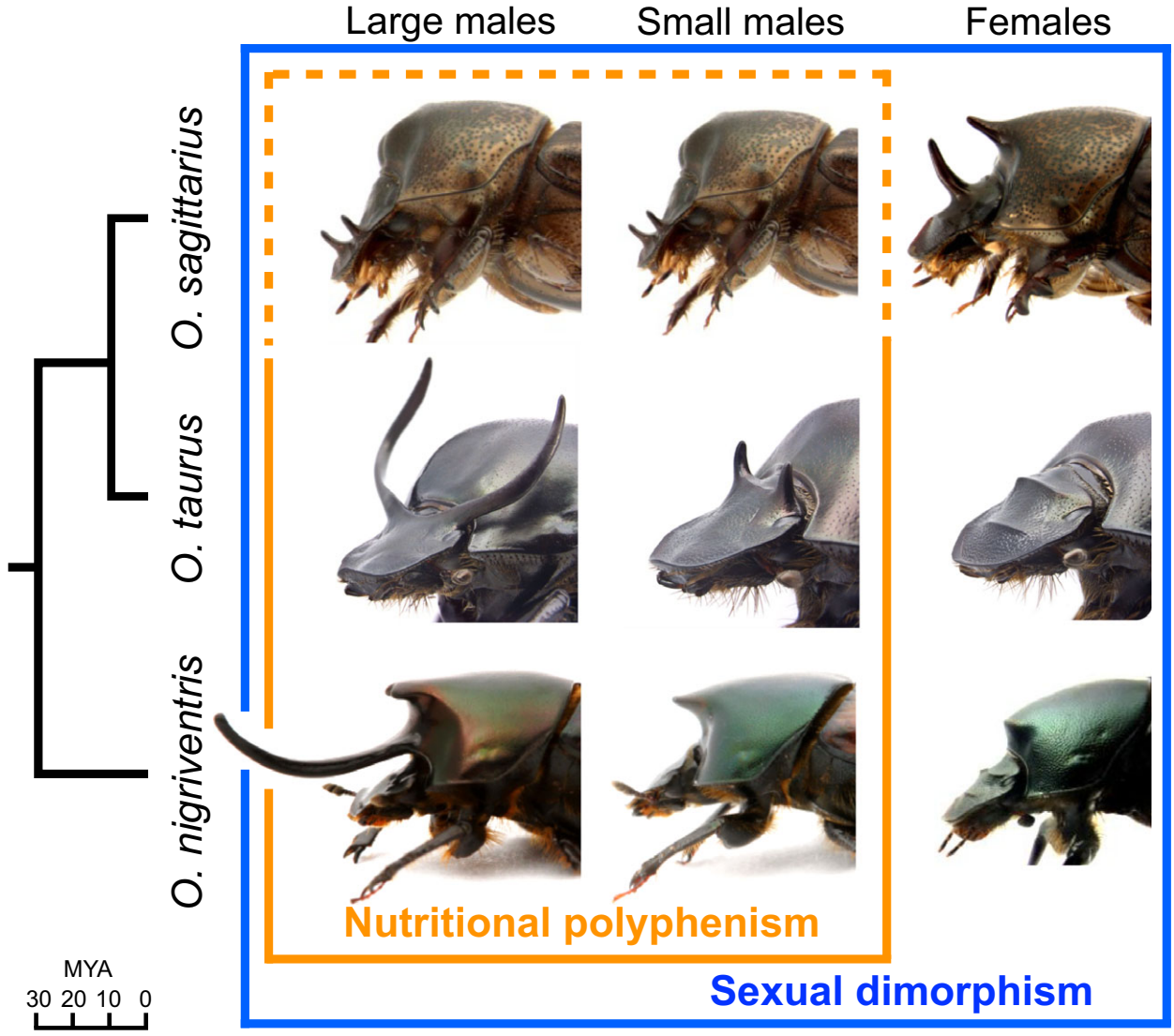

Fig. 1 Morphological diversity among three beetle species in the genus *Onthophagus*. Species differ along three axes of morphological differentiation emblematic for much of the diversity seen within the genus: horn position (head vs. thorax), degree of sensitivity to nutritional conditions during male development (male horn polyphenism) and polarity of the sexual dimorphism.

function of Gene Ontology (GO) and conditional gene expression. We defined conditional gene expression based on expression differences between sexes, nutritional states and body regions using our previously generated data set from *O. taurus* (Kijimoto *et al.*, 2014). This approach allowed us to test the hypothesis that common genes and pathways underlie morphological variation within and between species. In support of this hypothesis, we predicted a positive relationship between genes differentially expressed between sexes, morphs or body regions within a species and genes with elevated amino acid divergences between species that differ in the orientation of sexual dimorphism, degree of male dimorphism and horn position, respectively. Conversely, if the genetic basis of morphological variation within and between species is independent of one another, we expected no relationship between within-species gene expression differences and between-species sequence divergence.

Materials and Methods

Transcriptome sequencing and assembly

To explore genomic variation across our three focal species, we utilized two existing 454 transcriptome assemblies for *O. taurus* (Choi *et al.*, 2010) and *O. nigriventris* (Warren *et al.*, 2014) and generated a *de novo* transcriptome assembly for a previously uncharacterized third species, *O. sagittarius*, which exhibits a rare reversed sexual dimorphism and a secondary loss of the nutritional polyphenism in males (Fig. 1).

To generate a *de novo* transcriptome assembly for *O. sagittarius*, we collected live beetles from an introduced population on the island of Oahu, Hawai'i. We maintained the beetles in the laboratory using previously described culturing techniques (Moczek & Nagy, 2005). After two generations in laboratory conditions, we set up breeding containers to collect three females and three males at each of seven developmental stages: mid- and late third-instar larvae, day 1 and day 2 prepupae, day 1 and day 2 pupae and day 4 adults. We selected this range of developmental stages to maximize the breadth of gene coverage across the genome and to capture genes that may only be expressed at specific developmental stages (De Wit *et al.*, 2015). We did not use the developmental stage-specific expression data for downstream analyses because samples were comprised of multiple tissues. For each sample, we homogenized tissues using metal beads and mechanical agitation and extracted total RNA (Qiagen TissueLyser and RNeasy mini-prep kit). We prepared cDNA libraries for sequencing using TruSeq RNA Sample Prep Kit v2 (San Diego, CA) individually barcoding each sample. We sequenced paired-end 100-base pair reads on four Illumina HiSeq2000 lanes (Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA).

We processed raw reads for quality using FastX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) and Trimmomatic v. 0.25 (Bolger *et al.*, 2014) to remove Illumina adapter sequences, crop low-quality bases and exclude sequences with average quality scores less than 28 across 6 bp and removed any sequences shorter than 40 bp in length. We generated the *de novo* transcriptome assembly using Trinity v. 2012-10-05 with default parameters (Grabherr *et al.*, 2011). To evaluate assembly accuracy, we mapped left and right reads back to the assembly using bowtie v. 0.12.8 (Langmead *et al.*, 2009) and counted the number of reads correctly paired and in correct orientation using the SAM_nameSorted_to_uniq_count_stats.pl script (provided with Trinity package).

To control for partial sequence fragments potentially introduced by incomplete or inaccurate assembly or allelic variants from the 42 individuals contributing to the assembly, we collapsed sequences based on 90% sequence similarity using the program CD-HIT-EST (Li & Godzik, 2006). We performed the same clustering on the *O. nigriventris* assembly, but not on the *O. taurus* assembly as this step had already been executed (Choi *et al.*, 2010). We used two approaches to assess assembly completeness. First, we identified the number of core eukaryotic genes (CEGs) that were > 70% complete using the program CEGMA (Parra *et al.*, 2007). Second, we used results from BLASTP to the Swiss-Prot database (www.uniprot.org) and the analyze_blastPlus_topHit_coverage.pl script provided in the Trinity package to assess the fraction of nearly full-length contigs (> 80% alignment coverage).

Identification of orthologous gene sets

To identify orthologous gene sets, for each assembly, we predicted open reading frames (ORFs) requiring a start and stop codon and at least 100 bp in length using the Perl script transdecoder.pl (http://transdecoder.github.io/). We then concatenated the three resulting peptide files and used the BLASTP algorithm to compare all vs. all sequences with a cut-off of $e < 10^{-50}$ (higher e-values, i.e. less stringent cut-offs, did not appreciably increase the number of orthologous gene sets identified). From these BLASTP results, we used a clustering algorithm (*mcl*) (Enright *et al.*, 2002) to identify exclusive orthologous genes sets, or clusters, that had exactly one member sequence represented from each of the three species. We used exclusive clusters with just one gene member from each species in order to avoid including recent paralogous genes or pseudogenes in clusters of orthologous genes. We also focused on three-member orthologous gene sets rather than pairwise sets of orthologs so that the same genes would be present in all subsequent pairwise analyses.

Calculation of rates of evolution

We aligned nucleotide sequences of each orthologous gene set using MUSCLE (Edgar, 2004). For each orthologous gene set, we used these alignments to calculate pairwise rates of dS, dN and dN/dS using the PAML package version 4.6 (Yang, 2007). We used the following parameters in the control file for the CODONML program: runmode = −2 to calculate pairwise dN/dS among species using the Nei & Gojobori (1986) method (Nei & Gojobori, 1986) and NSsites = 0 so that dN/dS values were averaged across sites (Yang, 2007). With fewer than 10 species, it is recommended not to test for selection at individual codons in PAML (Yang & Swanson, 2002; Yang, 2007). For this reason, we calculated dN/dS averaged across all codons for each ortholog and species pair and integrated across functionally related genes to test for signatures of natural selection using enrichment analyses.

Integration of rates of evolution with gene function and gene expression data

To test for signals of natural selection in suites of functionally related genes that may not surpass the stringent dN/dS > 1 cut-off, we tested for enrichment of functional classes of genes for higher dN/dS values than expected by chance. To test whether signals of natural selection matched our predictions based on morphological differences between species pairs, we focused on pairwise rates of dN/dS between species rather than overall dN/dS (omega in PAML) across the three species. Given the primary phenotypic axes of diversification across these species, we *a priori* predicted genes related to the following processes to have been subject to selection during species divergence: (i) the sensing of nutritional conditions (critical for nutrition-dependent expression of horns as in male polyphenisms) along either branch in comparisons involving *O. sagittarius* which has lost the nutritional polyphenism, (ii) growth, metabolism and their regulation (critical for differential growth in sexual or male dimorphisms) also along either branch in species pairs involving *O. sagittarius*, (iii) developmental axis patterning (critical for the correct growth and positioning of horns) particularly between *O. taurus* and *O. nigriventris* that develop head vs. thoracic horns, respectively and (iv) somatic sex determination and differentiation, also along either branch in species pairs involving *O. sagittarius* which exhibits reversed sexual dimorphism.

To test for elevated dN/dS in functionally defined sets of genes, we implemented gene score resampling (GSR) using ErmineJ v3.0 (Lee *et al.*, 2005; Gillis *et al.*, 2010). GSR differs from most over-representation analyses in that the complete distribution of gene scores in a gene set is used rather than testing for enrichment in a subset of genes defined by a threshold cut-off. For each gene set (e.g. a specific GO category or custom-defined gene set), an aggregate score from the complete distribution of gene scores for that set is computed and significance of that score is determined by random sampling of all data (10 000 iterations). Given this approach, significant enrichment for elevated scores in a gene set can be driven by the collective of genes in the set with scores higher than expected by random chance (Gillis *et al.*, 2010; De Wit *et al.*, 2012), even if individual scores are insufficient to pass significance. The GSR approach can use a range of different metrics for gene scores including negative log-transformed *P*-values for differential expression, F_{ST} or dN/dS, as in the present study and others (Moyers & Rieseberg, 2013; Pespeni *et al.*, 2013; Dunning *et al.*, 2016; Franks *et al.*, 2016). Also, sets of genes can be custom defined in the program. Here, we defined gene sets in two ways: (i) based on membership in Gene Ontology (GO) categories and (ii) custom gene sets defined based on differential gene expression between large and small individuals (nutrition-sensitive), different sexes and different body regions (see below for details). Maximum and minimum gene set sizes were limited to 100 and 10 as recommended to avoid the use of overly general categories and those that would have low statistical power, respectively (Gillis *et al.*, 2010; De Wit *et al.*, 2012). This resulted in 146 gene sets tested, 143 Biological Process GO categories represented across the 2010 orthologous genes and three custom, expression-defined gene sets (described below). Significance was corrected for multiple testing using the Benjamini and Hochberg procedure (Benjamini & Hochberg, 1995). We considered categories significantly enriched with a FDR *P*-value < 0.05. This alpha value after false discovery rate correction was chosen to minimize false-positive results. We plotted dN/dS values for each species pair and the three expression-defined gene sets using the violin plot function (geom_violin) in the R package 'ggplot2' (Wickham, 2009).

To link genes to GO categories, we used BLASTx to the NCBI nonredundant (nr) database (cut-off of $e < 10^{-5}$) and the Blast2GO program (Conesa *et al.*, 2005) to match contigs with Gene Ontology (GO, Ashburner *et al.*, 2000) annotations. To define condition-dependent expressed gene sets, we used gene expression data from our previous study (Kijimoto *et al.*, 2014). Briefly, in this former study, we used a custom microarray to characterize the nutritional and sex-biased gene expression responses of four different body regions for 42 010 transcripts in *O. taurus* beetles (Kijimoto *et al.*, 2014). Using an ANOVA framework to compare the 16 conditions (2 sizes × 2 sexes × 4 body regions = 16 conditions, four biological replicates for each), we identified genes that were (i) nutritionally sensitive as those that were differentially expressed between large, high-nutrition and small, low-nutrition animals, (ii) sex-biased as those that were differentially

expressed between males and females of the same body size and (iii) body region-specific as those that were differentially expressed between four body regions: head horns, thoracic horns, legs and abdominal epidermis (FDR P-value < 0.05). Using this approach with the 2010 orthologous genes from the present study, we characterized 31 orthologs as nutrition-sensitive, 1448 orthologs as body region-specific, and 918 orthologs as sex-specific.

We wanted to test for elevated dN/dS in these conditionally expressed gene sets in the same statistical framework and analytical runs as the Gene Ontology-defined sets for two reasons: (i) to be able to compare enrichment results across category types (expression-defined vs. GO) and (ii) to be able to apply false discovery rate corrections to all categories tested simultaneously. To incorporate these three custom gene sets into ErmineJ analyses, we limited the gene set size of the expression-defined gene sets to 10–100 genes to match the size limits imposed on GO categories. As a result, the nutrition-sensitive gene set included all 31 genes significantly differentially expressed between large, high-nutrition and small, low-nutrition animals. The body region-specific and sex-specific gene sets included the 100 most significantly differentially expressed genes between tissues and sexes, respectively (based on the ANOVA results described above). We added these three custom gene sets to ErmineJ using the custom gene set function (Gillis *et al.*, 2010). In this way, we were able to perform three enrichment analyses total, one for each species pair. For each species pair, we tested for enrichment in each of the 146 gene sets (three custom, expression-defined and 143 GO Biological Process categories) and applied the false discovery rate correction for the 146 tests.

Results

Sequencing and assembly

For the *de novo* assembly of *O. sagittarius*, we assembled ~ 80 million paired, quality-trimmed reads to yield 76 764 contigs with a median length of 1126 base pairs (Table S1). To collapse allelic variants and remove fragment contigs, we clustered sequences based on sequence similarity reducing the number of contigs to 56 540. We also collapsed contigs based on sequence similarity for the *O. nigriventris* assembly, reducing the number of contigs from 70 295 to 52 327 (Table S1).

We found the *O. sagittarius* assembly to be ~ 98% complete with 243 of the 248 core eukaryotic genes (CEGs) present at > 70% coverage. More broadly, we found that over half of the contigs in the clustered assembly had greater than 80% alignment coverage to the best hit in the Sprot database. We found that 19 691 of the 23 012 ORF contigs had a BLAST hit to the nr database at an e-value cut-off of 10^{-5}. Of these, 5094 contigs (22%) were assigned to at least one Biological Process Gene Ontology category.

Comparative transcriptomics and rates of evolution

We identified 2010 orthologous gene sets with exactly one member from each species represented in each cluster. We found that rates of neutral evolution as estimated by the rate of synonymous substitutions (dS) matched the known evolutionary history of our three focal species. The two most closely related species, *O. taurus* and *O. sagittarius*, exhibited the lowest mean dS values (0.50, standard deviation 0.34, Table 1) with the most narrow distribution (KS test, $P < 0.0001$; Fig. 2a), followed by *O. taurus* and *O. nigriventris* (mean dS 0.66, 0.42 SD, Table 1), and *O. sagittarius* and *O. nigriventris* (mean dS 0.84, 0.47 SD, Table 1). We found that the two male polyphenic species that differ in the position of horns, *O. nigriventris* and *O. taurus*, showed the highest mean dN/dS values across all orthologs (0.13, 0.18 SD, Table 1) with the broadest distribution (KS test, $P < 0.0001$; Fig. 2b,c).

Across all three species pairs, and of the 2010 orthologs, we identified 17 genes with pairwise dN/dS values greater than one (Table 2). These include several genes related to development and metabolism, axes formation and positioning, including ADAMTS3, FGFR and Nodal (Fernandes *et al.*, 2001; Goriely *et al.*, 2005; Lapraz *et al.*, 2007; Muñoz-Chápuli, 2011; Rudolf *et al.*, 2013), as well as sulphide quinone reductase, a gene related to sulphur metabolism from bacteria to humans (Theissen *et al.*, 2003; Goubern *et al.*, 2007). Many of the remaining proteins exhibit sequence similarity to predicted proteins in the *Tribolium* flour beetle genome, but these proteins have not been functionally characterized (Table 2).

Functional enrichment for elevated dN/dS

Despite having the broadest distribution of dN/dS values (Fig. 2c), we found that high dN/dS was not concentrated in specific functional classes of genes when comparing *O. taurus* and *O. nigriventris* (gene score resampling (GSR), 143 GO categories tested, FDR $P > 0.05$). We found that the same was true in the

Table 1 Descriptive statistics for dN/dS calculations between three species pairs and overall.

Comparison	dN ± SD	dS ± SD	dN/dS ± SD
O. taurus vs. *O. nigriventris*	0.10 ± 0.18	0.66 ± 0.42	0.13 ± 0.18
O. sagittarius vs. *O. nigriventris*	0.08 ± 0.13	0.84 ± 0.47	0.09 ± 0.15
O. taurus vs. *O. sagittarius*	0.06 ± 0.14	0.50 ± 0.34	0.09 ± 0.14

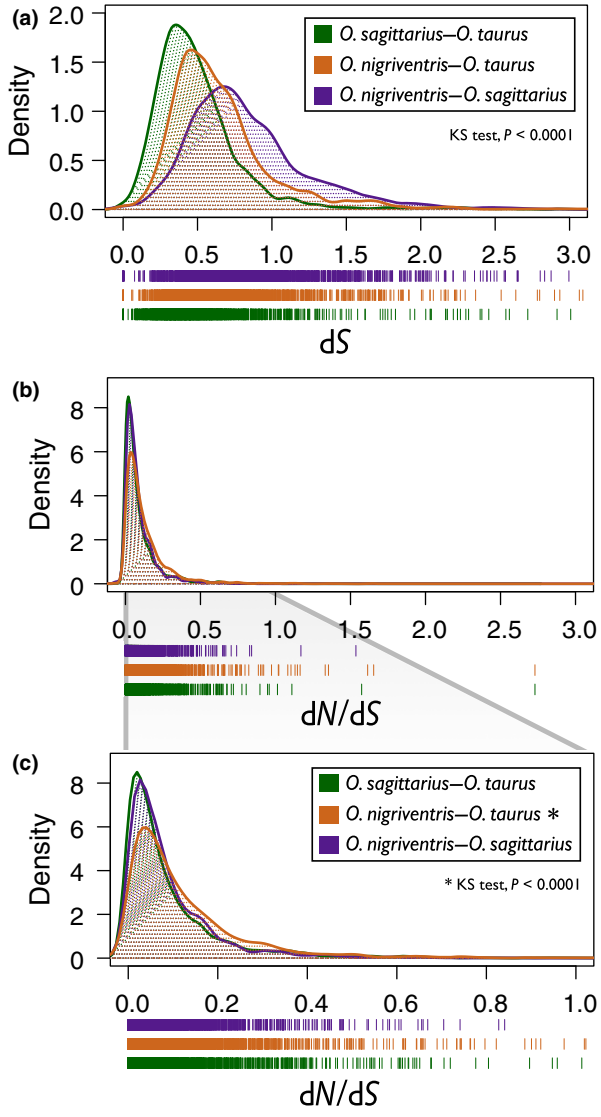

Fig. 2 Rates of evolution among the three species pairs. Transcriptome-wide density plots of (a) dS and (b) dN/dS for the 2010 orthologous gene sets. (c) dN/dS distributions between the values of 0 and 1. Tick marks under each density plot are rug plots that show the values for each ortholog.

O. nigriventris–*O. sagittarius* species pair (GSR, 143 GO categories tested, FDR $P > 0.05$). In contrast, comparing the two most closely related species, *O. taurus* and *O. sagittarius*, which exhibit reversed polarity in the direction of sexual dimorphism and presence (*O. taurus*) and secondary loss (*O. sagittarius*) of male polyphenism, we found three functional categories, related to lipid and nitrogen metabolism, enriched for elevated dN/dS suggesting positive selection or relaxed purifying selection acting in these gene classes (Table 3, GSR, 143 GO categories tested, FDR $P < 10^{-9}$). A total of 104 unique genes were represented across these three categories. Further, an additional 20 classes of genes related to metabolism and biosynthesis were enriched at the FDR $P < 0.05$ level (Table S2). Due to the overlap of genes across functional categories in Gene Ontology classification, the same suite of 104 unique high dN/dS genes present in the top three categories drove enrichment in these additional 20 gene categories.

Relationship between interspecific sequence divergence and intraspecific differential gene expression

We predicted that the most closely related species pair, *O. taurus* and *O. sagittarius*, that differs along all three focal phenotypic axes, presence (*O. taurus*) and secondary loss (*O. sagittarius*) of male polyphenism, reversed polarity in the direction of sexual dimorphism and position of the horn on the thorax (*O. sagittarius* females) and heads (*O. taurus* males and greatly reduced horns on *O. sagittarius* males), would show elevated divergence (dN/dS) in the genes that underlie differentiation between morphs, sexes and body regions within a species. Matching these predictions, we found that the most closely related species pair had higher-than-expected dN/dS values in all three sets of genes, those related to nutritional sensitivity (FDR $P < 10^{-10}$), genes differentially expressed between sexes (FDR $P < 10^{-10}$) and genes differentially expressed between body regions (FDR $P < 10^{-10}$, Fig. 3).

Between *O. nigriventris* and *O. taurus*, two species that differ in horn location and degree of nutrition-responsive growth (horn growth relative to body size is even more pronounced in *O. nigriventris* than in *O. taurus*), but not in the polarity of sexual dimorphism, we found elevated dN/dS in the genes that are most differentially expressed between body regions (FDR $P < 10^{-10}$) as well as several nutritionally sensitive genes (FDR $P < 0.05$), but no enrichment for elevated dN/dS in genes differentially expressed between sexes (FDR $P > 0.05$, Fig. 3), consistent with the similarity in polarity of sexual dimorphism in these species.

Finally, comparing *O. nigriventris* and *O. sagittarius*, two species that differ in the polarity of their sexual dimorphism and the presence (*O. nigriventris*) and secondary loss (*O. sagittarius*) of male nutritional polyphenism, as well as differences in the position of male horns, we found elevated dN/dS in genes differentially expressed between sexes and body regions (FDR $P < 0.05$, FDR $P < 10^{-10}$, respectively), but surprisingly, not in the nutritionally responsive gene set (FDR $P > 0.05$).

Taken together, we found a positive relationship between amino acid divergence between species and gene expression variation within species that matched predictions based on morphology in eight of nine tests, an unlikely result by chance (binomial probability, eight of nine significant with a conservative 0.5 probability of obtaining a significant test, $P = 0.0195$, Fig. 3).

Table 2 Genes putatively under positive selection based on pairwise dN/dS values greater than 1.

Comparison	Gene	Ortholog ID	dN	dS	dN/dS
O. taurus vs. *O. nigriventris*	Nodal modulator protein	Ortho_1917	0.21	0.07	3.26
	Uncharacterized protein	Ortho_0891	1.24	0.45	2.73
	ADAMTS-3 precursor, putative	Ortho_1947	0.86	0.52	1.65
	Hypothetical protein TcasGA2_TC006783	Ortho_0770	0.99	0.61	1.61
	Hypothetical protein TcasGA2_TC003503	Ortho_0620	0.99	0.73	1.35
	Sulphide quinone reductase	Ortho_0658	0.57	0.42	1.33
	Uncharacterized protein [Aedes aegypti]	Ortho_1008	0.97	0.83	1.16
	Uncharacterized protein [Tribolium castaneum]	Ortho_1422	0.91	0.80	1.15
	Protein disulphide-isomerase A3	Ortho_0974	1.27	1.12	1.13
	Fgfr1 oncogene partner	Ortho_1791	0.68	0.62	1.10
	Uncharacterized protein	Ortho_1296	0.17	0.16	1.08
	RecQ4	Ortho_1927	1.26	1.23	1.02
	Putative esterase	Ortho_1424	1.21	1.19	1.02
O. sagittarius vs. *O. nigriventris*	Sulphide quinone reductase	Ortho_0658	0.57	0.11	4.97
	Contactin-associated protein-like 2 [Habropoda laboriosa]	Ortho_1574	0.61	0.40	1.54
	Cation-transporting ATPase	Ortho_0007	0.25	0.22	1.17
O. taurus vs. *O. sagittarius*	Uncharacterized protein	Ortho_0891	1.24	0.45	2.73
	ADAMTS-3 precursor, putative	Ortho_1947	0.81	0.52	1.57
	Zinc finger CCCH domain-containing protein 14	Ortho_0599	0.65	0.58	1.11
	Golgin-84 isoform X2	Ortho_0041	0.70	0.69	1.01

Discussion

The tests for selection across three morphologically diverse *Onthophagus* species and integration with within-species gene expression variation data yielded three major findings. First, despite an overall strong signature of purifying selection, we identified a suite of 17 genes with dN/dS greater than one, with protein functions related to axis patterning, tissue development and metabolism. Second, integrating across the transcriptome-wide values of dN/dS, we found that the most closely related species pair, which differs in nutritional sensitivity, polarity of sexual dimorphism and horn location, had elevated dN/dS in proteins related to metabolism and biosynthesis. Lastly, we found that patterns of enrichment for elevated dN/dS in gene classes that underlie within-species differentiation between sexes, body regions and nutritional morphs match predictions based on morphological divergence between species. Below we discuss the most significant implications of our findings.

Evolutionary divergence and natural selection in three horned beetle species

Among the 17 genes with high rates of protein evolution, 11 (65%) play roles in tissue development, angiogenesis and axis patterning. For example, we identified a nodal modulator protein, a diffusible member of the TGF-β superfamily, shown to be essential for specification of germ layers, body axes and dorsal–ventral patterning in both invertebrates and vertebrates (Lapraz *et al.*, 2007). Although the function of this particular protein in horned beetles remains to be investigated, it

Table 3 Functional classes of genes enriched for elevated dN/dS between each species pair (146 categories tested: three custom gene expression-defined and 143 GO categories).

Species pair	Biological process GO category	Category ID	Num. genes	Corrected *P*-value
O. taurus vs. *O. sagittarius*	Body region-specific	Custom	100	2.1E-11
	Sex-biased	Custom	100	2.5E-11
	Nutrition-sensitive	Custom	31	3.0E-11
	Cellular lipid metabolic process	GO:0044255	19	4.8E-11
	Nitrogen compound metabolic process	GO:0006807	89	7.2E-11
	Cellular nitrogen compound metabolic process	GO:0034641	67	1.4E-10
O. nigriventris vs. *O. taurus*	Body region-specific	Custom	100	7.4E-11
	Nutrition-sensitive	Custom	31	0.0432
O. nigriventris vs. *O. sagittarius*	Body region-specific	Custom	100	4.9E-11
	Sex-biased	Custom	100	0.0414

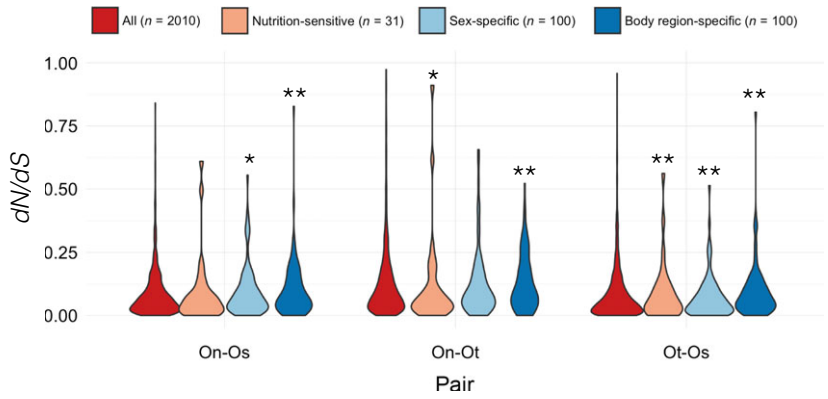

Fig. 3 Violin plots of dN/dS values for the three sets of conditionally expressed genes for each species pair. Asterisks indicate significant enrichment for genes with elevated dN/dS compared to random sampling of all orthologs; two asterisks indicate FDR $P < 0.0001$ and one asterisk indicates FDR $P < 0.05$.

is worth noting that previous work has independently implicated another member of the TGF-β signalling pathway, *decapentaplegic (dpp)* in both head and thoracic horn formation in *O. sagittarius* and *O. binodis* (Wasik & Moczek, 2011). Also included in this list is ADAMTS3, an extracellular matrix protein that plays an important role in developing epidermal tissues, collagen processing and angiogenesis in mammals and invertebrates (Fernandes *et al.*, 2001; Porter *et al.*, 2005). Similarly, we identified the fibroblast growth factor receptor (FGFR) which is also known to guide angiogenesis and wound healing in invertebrates from *Drosophila* (named *branchless* in *Drosophila*) to the colonial ascidian *Botryllus* (Muñoz-Chápuli, 2011), and has also been shown to be subject to strong positive selection in humans, affecting craniofacial and limb phenotypes (Goriely *et al.*, 2005). Interestingly, both ADAMTS3 and FGFR play key roles in angiogenesis – the process of blood vessel expansion in mammals – which has its evolutionary roots in vascular growth in invertebrates (Muñoz-Chápuli, 2011). In insects, orthologous pathways are responsible for tracheal system development, oxygen sensing and delivery and metabolism (Centanin *et al.*, 2010; Muñoz-Chápuli, 2011). Amino acid divergence in these genes between these beetle species may be related to differences in metabolic demands and differences in response to nutritional conditions. In fact, previous work has provided strong evidence linking tracheal system development to growth, moulting and metamorphosis (Callier & Nijhout, 2011) and at least one study demonstrated a positive relationship between beetle size and tracheal investment both within and between species (Kaiser *et al.*, 2007). A fourth gene of known function, sulphide quinone reductase, plays an important role in the metabolism of sulphur in the mitochondria of gut cells from bacteria to humans (Theissen *et al.*, 2003; Goubern *et al.*, 2007) and so could be related to differences in response to nutrition in beetles. The unknown roles of these genes in horned beetle morphological divergence present interesting opportunities for further functional studies (Kijimoto *et al.*, 2012).

In addition, we identified specific functional classes of genes that have higher dN/dS than expected by chance, potentially due to positive selection or relaxed purifying selection. We found that the most closely related species pair, *O. taurus* and *O. sagittarius*, characterized by a rare reversal of sexual dimorphism and loss of the nutritional polyphenism (both in *O. sagittarius*), exhibited elevated dN/dS in specific functional classes of genes related to metabolism and biosynthesis, matching our *a priori* predictions based on divergence in nutritional sensitivity and horn morphology. These results highlight the power of integrating nucleotide sequence and functional annotation data in the study of evolution in emerging model organisms.

The developmental genetic bases of intra- and interspecific divergences

We hypothesized that genes that underlie morphological differentiation between morphs and sexes *within* a species may be the same genes – or may belong to the same functional group – that underlie parallel divergences between species. For example, we sought to answer whether genes that exhibited nutrition-responsive expression within a species were the same as those that underlie evolved changes in nutritional sensitivity across species. Such a relationship may be expected, for instance, because genes that are expressed in a tissue-, sex- or nutrition-dependent context are subject to reduced selective constraints and therefore more free to diverge between species either via positive selection or relaxed constraint (Pál *et al.*, 2006). We predicted which gene sets we would expect to show stronger signals of positive selection for each species pair based on differences in the orientation of the sexual dimorphism, the degree of sensitivity to nutritional status and the position of horns in different body regions. These

results validate and expand upon findings from previous studies that found positive relationships between biased gene expression and purifying or positive selection in beetles across a broad phylogenetic context (Snell-Rood *et al.*, 2011; Warren *et al.*, 2014). By focusing on three key species that reflect much of the morphological diversity represented across the genus *Onthophagus*, integrating with expression data from along the same axes of variation and using gene expression and protein sequence data from an order of magnitude more genes than that available to previous studies, we found that conditionally expressed genes exhibited higher levels of protein sequence divergence in cases where species had diverged along those axes and no relationship where species showed no such divergence, that is the shared polarity of sexual dimorphism in *O. taurus* and *O. nigriventris* (Fig. 3). In particular, we found a strong positive relationship between biased gene expression and protein sequence divergence in the most closely related species pair that had diverged along all three axes: reversal of the sexual dimorphism, loss of the nutritional polyphenism and sex-specific location of head and thoracic horns (Fig. 3).

Additionally, all three species pairs showed elevated dN/dS for genes exhibiting significant body region specificity, the axis along which all species pairs have diverged from each other. These findings match results from previous studies showing that genes expressed in fewer tissues exhibit faster rates of evolution, likely due to relaxed pleiotropic constraint (Duret & Mouchiroud, 2000; Pál *et al.*, 2006). Similarly, we found strong enrichment for elevated dN/dS in nutrition-sensitive genes in the most closely related species pair (*O. taurus, O. sagittarius*) characterized by the recent loss of nutritional polyphenism in male *O. sagittarius* (Fig. 3), matching the Gene Ontology results for enrichment of high dN/dS for genes in metabolism categories in the same species pair (Table 3).

Regarding sex-biased genes, we sought to test the hypothesis that genes that are highly differentially expressed between sexes (within a species) also exhibit signatures of positive selection when orthologous gene sets are compared across species that have diverged in the polarity of the sexual dimorphism. Specifically, we predicted that genes that are associated with phenotypic differences between sexes should exhibit evidence of positive selection in species pairs where the polarity of sexual dimorphism is reversed as in *O. sagittarius* vs. *O. taurus* and/or *O. nigriventris*. Matching predictions, we found the strongest enrichment for elevated dN/dS in the species pairs involving *O. sagittarius* (Table 3), whereas the species pair with the same polarity in sexual dimorphism (*O. nigriventris, O. taurus)* showed rates of evolution in sex-biased genes on par with the rest of the transcriptome (Table 3). Many studies have shown that sex-biased genes are subject to strong natural and sexual selection facilitated by their reduced pleiotropy and strong potential for affecting fitness-related traits such as fecundity and mating success (Meiklejohn *et al.*, 2003; Ellegren & Parsch, 2007; Oliver *et al.*, 2010). Indeed, previous work in *O. sagittarius* showed a positive relationship between female horn size and fitness in the context of competition for dung as a breeding resource (Watson & Simmons, 2010). Our results thus raise the possibility that positive selection acting on sex-biased genes may have facilitated the unusual, female-specific development of a prominent horn phenotype in *O. sagittarius*.

In sum, we find that many of the same genes that are differentially expressed between sexes, body regions and nutritional states within a species are frequently subject to positive or relaxed purifying selection among our three focal species that exhibit canalized divergences along similar morphological axes. Such relationships may arise for at least two nonexclusive reasons. First, fast-evolving genes subject to low levels of purifying selection may simply be more readily co-opted into conditional gene expression patterns over time due to neutral rather than adaptive processes (Khaitovich *et al.*, 2005; Hunt *et al.*, 2011; Helantera & Uller, 2014). If so, the relative abundance and functional distribution of such genes may impose significant constraints and limits on the evolution of developmental plasticity mediated by environment-sensitive gene expression. Alternatively, the arrow of causation may be reversed, and conditional gene expression may – by virtue of reducing pleiotropic constraints – facilitate the relatively more rapid evolution of conditionally expressed genes (West-Eberhard, 1989; Leichty *et al.*, 2012). If correct, intraspecific variation via condition-specific development may thus contribute important developmental genetic substrate to facilitate divergences among species (West-Eberhard, 2003; Gomez-Mestre & Buchholz, 2006; Pfennig *et al.*, 2010), and, moreover, may encourage and bias adaptive radiations along certain morphological routes over others (Schluter, 1996; Wund *et al.*, 2008). More generally, our results may contribute to a better understanding of the molecular genetic mechanisms underlying genetic accommodation, a process whereby initially environment-induced phenotypes become heritably modified through selection on standing genetic variation or newly arriving mutations (Pfennig *et al.*, 2010; Moczek *et al.*, 2011). Conditional gene expression is by definition influenced by the internal and/or external environment, and our results suggest that in *Onthophagus*, genes whose expression is influenced by sex, body region and nutrition may be predisposed to contribute disproportionately to interspecific diversification in sexual dimorphisms, changes in the location of secondary sexual traits and the degree of male polyphenism, respectively. Exactly how differential expression due to condition may fuel interspecific divergence in protein

sequence, however, is not immediately clear and constitutes an exciting area for future research.

Acknowledgments

We thank Ian Warren and Doug Emlen for sharing the transcriptome assembly for *O. nigriventris* before it was published, Carrie Ganote and the National Center for Genome Analysis Support (DBI-1458641 and ABI-1062432) for genome analysis support, James Pease for computational assistance and Adam Moore, Alex Neufeld, Hailey Riggs and Justin Song for expert help with beetle husbandry. This research was made possible through National Science Foundation grants DBI 1103716 to M.H.P. and IOS 1120209 and 1256689 to A.P.M. Partial support also derived from a grant from the John Templeton Foundation. The opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by the U.S. Army, Department of Defense, the U.S. Government or the John Templeton Foundation.

Competing interests

The authors have no competing interests.

References

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M. *et al.* 2000. Gene ontology: tool for the unification of biology. *Nat. Genet.* **25**: 25–29.

Aubin-Horth, N. & Renn, S.C.P. 2009. Genomic reaction norms: using integrative biology to understand molecular mechanisms of phenotypic plasticity. *Mol. Ecol.* **18**: 3763–3780.

Barreto, F.S., Pereira, R.J. & Burton, R.S. 2015. Hybrid dysfunction and physiological compensation in gene expression. *Mol. Biol. Evol.* **32**: 613–622.

Benjamini, Y. & Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **51**: 289–300.

Bolger, A.M., Lohse, M. & Usadel, B. 2014. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Callier, V. & Nijhout, H.F. 2011. Control of body size by oxygen supply reveals size-dependent and size-independent mechanisms of molting and metamorphosis. *Proc. Natl. Acad. Sci. USA* **108**: 14664–14669.

Carlson, J.E., Holsinger, K.E. & Carlson, J.E. 2015. Extrapolating from local ecological processes to genus-wide patterns in colour polymorphism in South African Protea. *Proc. R. Soc. B Biol. Sci.* **282**: 20150583.

Carroll, S.B. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**: 25–36.

Centanin, L., Gorr, T.A. & Wappner, P. 2010. Tracheal remodelling in response to hypoxia. *J. Insect Physiol.* **56**: 447–454.

Choi, J.-H., Kijimoto, T., Snell-Rood, E.C., Tae, H., Yang, Y.-I., Moczek, A.P. *et al.* 2010. Gene discovery in the horned beetle Onthophagus taurus. *BMC Genom.* **11**: 703.

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M. & Robles, M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**: 3674–3676.

De Wit, P., Pespeni, M.H., Ladner, J.T., Barshis, D.J., Seneca, F., Jaris, H. *et al.* 2012. The simple fool's guide to population genomics via RNA-Seq: an introduction to high-throughput sequencing data analysis. *Mol. Ecol. Resour.* **12**: 1058–1067.

De Wit, P., Pespeni, M.H. & Palumbi, S.R. 2015. SNP genotyping and population genomics from expressed sequences -current advances and future possibilities. *Mol. Ecol.* **24**: 2310–2323.

Dunning, L.T., Hipperson, H., Baker, W.J., Butlin, R.K., Devaux, C., Hutton, I. *et al.* 2016. Ecological speciation in sympatric palms: 1. Gene expression, selection and pleiotropy. *J. Evol. Biol.* **29**: 1472–1487.

Duret, L. & Mouchiroud, D. 2000. Determinants of substitution rates in mammalian genes: expression pattern affects selection intensity but not mutation rate. *Mol. Biol. Evol.* **17**: 68–74.

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**: 1792–1797.

Ekblom, R. & Galindo, J. 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity.* **107**: 1–15.

Ellegren, H. 2014. Genome sequencing and population genomics in non-model organisms. *Trends Ecol. Evol.* **29**: 51–63.

Ellegren, H. & Parsch, J. 2007. The evolution of sex-biased genes and sex-biased gene expression. *Nat. Rev. Genet.* **8**: 689–698.

Elmer, K.R. & Meyer, A. 2011. Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends Ecol. Evol.* **26**: 298–306.

Emlen, D.J., Hunt, J. & Simmons, L.W. 2005a. Evolution of sexual dimorphism and male dimorphism in the expression of beetle horns: phylogenetic evidence for modularity, evolutionary lability, and constraint. *Am. Nat.* **166**(Suppl): S42–S68.

Emlen, D.J., Marangelo, J., Ball, B., Cunningham, C.W. & Rowe, L. 2005b. Diversity in the weapons of sexual selection: horn evolution in the beetle genus *Onthophagus* (Coleoptera: Scarabaeidae). *Evolution.* **59**: 1060–1084.

Emlen, D.J., Warren, I.A., Johns, A., Dworkin, I. & Corley Lavine, L. 2012. A mechanism of extreme growth and reliable signaling in sexually selected ornaments and weapons Douglas. *Science* **337**: 860–865.

Enright, A.J., Van Dongen, S. & Ouzounis, C.A. 2002. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* **30**: 1575–1584.

Fernandes, R.J., Hirohata, S., Engle, J.M., Colige, A., Cohn, D.H., Eyre, D.R. *et al.* 2001. Procollagen II amino propeptide processing by ADAMTS-3. Insights on dermatosparaxis. *J. Biol. Chem.* **276**: 31502–31509.

Franks, S.J., Kane, N.C., O'Hara, N.B., Tittes, S. & Rest, J.S. 2016. Rapid genome-wide evolution in *Brassica rapa* populations following drought revealed by sequencing of ancestral and descendant gene pools. *Mol. Ecol.* **25**: 3622–3631.

Ghalambor, C.K., Hoke, K.L., Ruell, E.W., Fischer, E.K., Reznick, D.N. & Hughes, K.A. 2015. Non-adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. *Nature* **525**: 372–375.

Gillis, J., Mistry, M. & Pavlidis, P. 2010. Gene function analysis in complex data sets using ErmineJ. *Nat. Protoc.* **5**: 1148–1159.

Gomez-Mestre, I. & Buchholz, D.R. 2006. Developmental plasticity mirrors differences among taxa in spadefoot toads linking plasticity and diversity. *Proc. Natl. Acad. Sci. USA* **103**: 19021–19026.

Gompel, N., Prud'homme, B., Wittkopp, P.J., Kassner, V.A. & Carroll, S.B. 2005. Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in Drosophila. *Nature* **433**: 481–487.

Goriely, A., McVean, G.A.T., van Pelt, A.M.M., O'Rourke, A.W., Wall, S.A., de Rooij, D.G. *et al.* 2005. Gain-of-function amino acid substitutions drive positive selection of FGFR2 mutations in human spermatogonia. *Proc. Natl. Acad. Sci. USA* **102**: 6051–6056.

Gossmann, T.I., Schmid, M.W., Grossniklaus, U. & Schmid, K.J. 2014. Selection-driven evolution of sex-biased genes is consistent with sexual selection in arabidopsis thaliana. *Mol. Biol. Evol.* **31**: 574–583.

Goubern, M., Andriamihaja, M., Nübel, T., Blachier, F. & Bouillaud, F. 2007. Sulfide, the first inorganic substrate for human cells. *FASEB J.* **21**: 1699–1706.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I. *et al.* 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**: 644–652.

Harpur, B.A., Kent, C.F., Molodtsova, D., Lebon, J.M.D., Alqarni, A.S., Owayss, A.A. *et al.* 2014. Population genomics of the honey bee reveals strong signatures of positive selection on worker traits. *Proc. Natl. Acad. Sci. USA* **111**: 2614–2619.

Helantera, H. & Uller, T. 2014. Neutral and adaptive explanations for an association between caste biased gene expression and rate of sequence evolution. *Front. Genet.* **5**: 1–27.

Hodgins, K.A., Yeaman, S., Nurkowski, K.A., Rieseberg, L.H. & Aitken, S.N. 2016. Expression divergence is correlated with sequence evolution but not positive selection in conifers. *Mol. Biol. Evol.* **33**: 1502–1516.

Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A. & Cresko, W.A. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet.* **6**: e1000862.

Hunt, B.G., Ometto, L., Wurm, Y., Shoemaker, D., Yi, S.V., Keller, L. *et al.* 2011. Relaxed selection is a precursor to the evolution of phenotypic plasticity. *Proc. Natl. Acad. Sci. USA* **108**: 15936–15941.

Hunt, B.G., Ometto, L., Keller, L. & Goodisman, M.A.D. 2013. Evolution at two levels in fire ants: the relationship between patterns of gene expression and protein sequence evolution. *Mol. Biol. Evol.* **30**: 263–271.

Kaiser, A., Klok, C.J., Socha, J.J., Lee, W.-K., Quinlan, M.C. & Harrison, J.F. 2007. Increase in tracheal investment with beetle size supports hypothesis of oxygen limitation on insect gigantism. *Proc. Natl. Acad. Sci. USA* **104**: 13198–13203.

Kavembe, G.D., Franchini, P., Irisarri, I., Machado-Schiaffino, G. & Meyer, A. 2015. Genomics of adaptation to multiple concurrent stresses: insights from comparative transcriptomics of a cichlid fish from one of earth's most extreme environments, the hypersaline soda lake Magadi in Kenya, East Africa. *J. Mol. Evol.* **81**: 90–109.

Khaitovich, P., Hellmann, I., Enard, W., Nowick, K., Leinweber, M., Franz, H. *et al.* 2005. Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science* **309**: 1850–1854.

Kijimoto, T., Pespeni, M., Beckers, O. & Moczek, A.P. 2012. Beetle horns and horned beetles: emerging models in developmental evolution and ecology. *Wiley Interdiscip. Rev. Dev. Biol.* **2**: 405–418.

Kijimoto, T., Snell-Rood, E.C., Pespeni, M.H., Rocha, G., Kafadar, K. & Moczek, A.P. 2014. The nutritionally responsive transcriptome of the polyphenic beetle *Onthophagus taurus* and the importance of sexual dimorphism and body region. *Proc. R. Soc. B* **281**: 20142084.

Kratochwil, C.F. & Meyer, A. 2014. Closing the genotype-phenotype gap: emerging technologies for evolutionary genetics in ecological model vertebrate systems. *BioEssays* **37**: 213–226.

Lamichhaney, S., Berglund, J., Almén, M.S., Maqbool, K., Grabherr, M., Martinez-Barrio, A. *et al.* 2015. Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* **518**: 371–375.

Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**: R25.

Lapraz, F., Duboc, V. & Lepage, T. 2007. A genomic view of TGF-β signal transduction in an invertebrate deuterostome organism and lessons from the functional analyses of Nodal and BMP2/4 during sea urchin development. *Signal Transduct.* **7**: 187–206.

Larracuente, A.M., Sackton, T.B., Greenberg, A.J., Wong, A., Singh, N.D., Sturgill, D. *et al.* 2008. Evolution of protein-coding genes in Drosophila. *Trends Genet.* **24**: 114–123.

Lee, H.K., Braynen, W., Keshav, K. & Pavlidis, P. 2005. ErmineJ: tool for functional analysis of gene expression data sets. *BMC Bioinformatics* **6**: 269.

Leichty, A.R., Pfennig, D.W., Jones, C.D. & Pfennig, K.S. 2012. Relaxed genetic constraint is ancestral to the evolution of phenotypic plasticity. *Integr. Comp. Biol.* **52**: 16–30.

Lemos, B., Meiklejohn, C.D., Cáceres, M. & Hartl, D.L. 2005. Rates of divergence in gene expression profiles of primates, mice, and flies: stabilizing selection and variability among functional categories. *Evolution* **59**: 126–137.

Li, W. & Godzik, A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**: 1658–1659.

Lorch, P.D., Proulx, S., Rowe, L. & Day, T. 2003. Condition-dependent sexual selection can accelerate adaptation. *Evol. Ecol. Res.* **5**: 867–881.

Martin, A. & Orgogozo, V. 2013. The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution* **67**: 1235–1250.

Meiklejohn, C.D., Parsch, J., Ranz, J.M. & Hartl, D.L. 2003. Rapid evolution of male-biased gene expression in Drosophila. *Proc. Natl. Acad. Sci. USA* **100**: 9894–9899.

Mikheyev, A.S. & Linksvayer, T.A. 2015. Genes associated with ant social behavior show distinct transcriptional and evolutionary patterns. *Elife* **2015**: 1–17.

Moczek, A.P. & Nagy, L.M. 2005. Diverse developmental mechanisms contribute to different levels of diversity in horned beetles. *Evol. Dev.* **7**: 175–185.

Moczek, A.P., Sultan, S., Foster, S., Ledón-Rettig, C., Dworkin, I., Nijhout, H.F., et al. 2011. The role of developmental plasticity in evolutionary innovation. *Proc. Biol. Sci.* **278**: 2705–2713.

Moyers, B.T. & Rieseberg, L.H. 2013. Divergence in gene expression is uncoupled from divergence in coding sequence in a secondarily woody sunflower. *Int. J. Plant Sci.* **174**: 1079–1089.

Muñoz-Chápuli, R. 2011. Evolution of angiogenesis. *Int. J. Dev. Biol.* **55**: 345–351.

Nei, M. & Gojobori, T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3**: 418.

Nuzhdin, S.V., Wayne, M.L., Harmon, K.L. & McIntyre, L.M. 2004. Common pattern of evolution of gene expression level and protein sequence in Drosophila. *Mol. Biol. Evol.* **21**: 1308–1317.

Oliver, T.A., Garfield, D.A., Manier, M.K., Haygood, R., Wray, G.A. & Palumbi, S.R. 2010. Whole-genome positive selection and habitat-driven evolution in a shallow and a deep-sea urchin. *Genome Biol. Evol.* **2**: 800–814.

Pál, C., Papp, B. & Lercher, M.J. 2006. An integrated view of protein evolution. *Nat. Rev. Genet.* **7**: 337–348.

Parra, G., Bradnam, K. & Korf, I. 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* **23**: 1061–1067.

Perry, J.C., Harrison, P.W. & Mank, J.E. 2014. The ontogeny and evolution of sex-biased gene expression in *Drosophila melanogaster*. *Mol. Biol. Evol.* **31**: 1206–1219.

Pespeni, M.H., Sanford, E., Gaylord, B., Hill, T.M., Hosfelt, J.D., Jaris, H.K. *et al.* 2013. Evolutionary change during experimental ocean acidification. *Proc. Natl. Acad. Sci. USA* **110**: 6937–6942.

Pfennig, D.W., Wund, M.A., Snell-Rood, E.C., Cruickshank, T., Schlichting, C.D. & Moczek, A.P. 2010. Phenotypic plasticity's impacts on diversification and speciation. *Trends Ecol. Evol.* **25**: 459–467.

Porter, S., Clark, I.M., Kevorkian, L. & Edwards, D.R. 2005. The ADAMTS metalloproteinases. *Biochem. J.* **386**: 15–27.

Purandare, S.R., Bickel, R.D., Jaquiery, J., Rispe, C. & Brisson, J.A. 2014. Accelerated evolution of morph-biased genes in pea aphids. *Mol. Biol. Evol.* **31**: 2073–2083.

Roux, J., Privman, E., Moretti, S., Daub, J.T., Robinson-Rechavi, M. & Keller, L. 2014. Patterns of positive selection in seven ant genomes. *Mol. Biol. Evol.* **31**: 1661–1685.

Rudolf, A., Hübinger, C., Hüsken, K., Vogt, A., Rebscher, N., Önel, S.F. *et al.* 2013. The hydra FGFR, Kringelchen, partially replaces the Drosophila heartless FGFR. *Dev. Genes. Evol.* **223**: 159–169.

Sanger, T.J., Sherratt, E., McGlothlin, J.W., Brodie, E.D., Losos, J.B. & Abzhanov, A. 2013. Convergent evolution of sexual dimorphism in skull shape using distinct developmental strategies. *Evolution* **67**: 2180–2193.

Sax, K. 1923. The association of size differences with seed-coat pattern and pigmentation in Phaseolus vulgaris. *Genetics* **8**: 552–560.

Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. *Evolution* **50**: 1766–1774.

Smith, C.R., Toth, A.L., Suarez, A.V. & Robinson, G.E. 2008. Genetic and genomic analyses of the division of labour in insect societies. *Nat. Rev. Genet.* **9**: 735–748.

Smith, C.R., Helms Cahan, S., Kemena, C., Brady, S.G., Yang, W., Bornberg-Bauer, E. *et al.* 2015. How do genomes create novel phenotypes Insights from the loss of the worker caste in ant social parasites. *Mol. Biol. Evol.* **32**: 2919–2931.

Snell-Rood, E.C., Cash, A., Han, M.V., Kijimoto, T., Andrews, J. & Moczek, A.P. 2011. Developmental decoupling of alternative phenotypes: insights from the transcriptomes of horn-polyphenic beetles. *Evolution* **65**: 231–245.

Soria-Carrasco, V., Gompert, Z., Comeault, A.A., Farkas, T.E., Parchman, T.L., Johnston, J.S. *et al.* 2014. Stick insect genomes reveal natural selection's role in parallel speciation. *Science* **344**: 738–742.

Theissen, U., Hoffmeister, M., Grieshaber, M. & Martin, W. 2003. Single eubacterial origin of eukaryotic sulfide:quinone oxidoreductase, a mitochondrial enzyme conserved from the early evolution of eukaryotes during anoxic and sulfidic times. *Mol. Biol. Evol.* **20**: 1564–1574.

Warnefors, M. & Kaessmann, H. 2013. Evolution of the correlation between expression divergence and protein divergence in mammals. *Genome Biol. Evol.* **5**: 1324–1335.

Warren, I.A., Vera, J.C., Johns, A., Zinna, R., Marden, J.H., Emlen, D.J. *et al.* 2014. Insights into the development and evolution of exaggerated traits using *de novo* transcriptomes of two species of horned scarab beetles. *PLoS ONE* **9**: e88364.

Wasik, B.R. & Moczek, A.P. 2011. Decapentaplegic (dpp) regulates the growth of a morphological novelty, beetle horns. *Dev. Genes. Evol.* **221**: 17–27.

Watson, N.L. & Simmons, L.W. 2010. Reproductive competition promotes the evolution of female weaponry. *Proc. Biol. Sci.* **277**: 2035–2040.

West-Eberhard, M.J. 1989. Phenotypic plasticity and the origins of diversity. *Annu. Rev. Ecol. Syst.* **20**: 249–278.

West-Eberhard, M.J. 2003. *Developmental Plasticity and Evolution*. Oxford University Press, New York.

Wheeler, D.E. 1986. Developmental and physiological determinants of caste in social Hymenoptera: evolutionary implications. *Am. Nat.* **128**: 13–34.

Whitehead, A. & Crawford, D.L. 2006. Variation within and among species in gene expression: raw material for evolution. *Mol. Ecol.* **15**: 1197–1211.

Whittle, C.A., Sun, Y. & Johannesson, H. 2014. Dynamics of transcriptome evolution in the model eukaryote Neurospora. *J. Evol. Biol.* **27**: 1125–1135.

Wickham, H. 2009. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York, NY.

Wittkopp, P.J., Stewart, E.E., Arnold, L.L., Neidert, A.H., Haerum, B.K., Thompson, E.M. *et al.* 2009. Intraspecific polymorphism to interspecific divergence: genetics of pigmentation in Drosophila. *Science* **326**: 540–544.

Wund, M.A., Baker, J.A., Clancy, B., Golub, J.L. & Foster, S.A. 2008. A test of the "flexible stem" model of evolution: ancestral plasticity, genetic accommodation, and morphological divergence in the threespine stickleback radiation. *Am. Nat.* **172**: 449–462.

Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**: 1586.

Yang, Z. & Swanson, W.J. 2002. Codon-substitution models to detect adaptive evolution that account for heterogeneous selective pressures among site classes. *Mol. Biol. Evol.* **19**: 49–57.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:
Table S1 Summary statistics for transcriptome assemblies of three horned beetle species.
Table S2 Functional classes of genes enriched for elevated dN/dS between each species pair at FDR $P < 0.05$ (146 categories tested: 3 custom gene expression-defined and 143 GO categories).

Data deposited at Dryad: https://doi.org/10.5061/dryad.rv08 j

Received 16 March 2016; accepted 5 March 2017