Genes involved in the convergent evolution of asexuality in stick insects

Darren J. Parker[1,2,*], Jens Bast[1], Kirsten Jalvingh[1], Zoé Dumas[1], Marc Robinson-Rechavi[1,2], and Tanja Schwander[1]

1. Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland
2. Swiss Institute of Bioinformatics, Lausanne, Switzerland

* Corresponding author: DarrenJames.Parker@unil.ch

Abstract

The ability to reproduce is one of the most fundamental traits that distinguishes living organisms from inorganic matter, yet, organisms use a panoply of strategies for reproduction. The evolution of these strategies, especially sexual and asexual reproduction, has been the focus of intensive study. By contrast, the molecular underpinnings of sexual and asexual reproduction remain relatively unknown. We investigated convergent gene expression changes and patterns of molecular evolution across five independent transitions to asexuality in stick insects. We compared gene expression of asexual females to those of females from close sexual relatives in whole-bodies and two tissues: the reproductive tract and legs. We identified a striking amount of convergent gene expression change, ranging from 5 to 8% of genes examined. Convergent changes were also tissue-specific, with most convergent genes changing in only one tissue type. Functional enrichment tests found that genes showing convergent changes in the reproductive tract were associated with meiotic spindle formation and centrosome organization. These genes are particularly interesting as they can influence the production of unreduced eggs, a key barrier to asexual reproduction. Changes in legs and whole-bodies were likely involved in female sexual trait decay, with enrichment in terms such as sperm-storage and pigmentation. By identifying changes occurring across multiple independent transitions to asexuality, our results provide a rare insight into the molecular basis of asexual phenotypes and suggest that the evolutionary path to asexuality is highly constrained, requiring repeated changes to the same key genes.

Keywords: convergent evolution, parthenogenesis, *Timema*

Introduction

Sexual reproduction is extremely costly. Sex is less efficient than asexuality for transmitting genes to future generations [1] and in order to outcross, an individual has to find a partner, forgo foraging, and risk contracting sexually transmitted diseases and predation while mating [2,3]. Yet, the overwhelming number of sexual, as compared to asexual, animal and plant species [4,5] indicates that sexual reproduction is highly advantageous. Identifying potential advantages conferred by sex has motivated decades of research and a rich body of work on the evolution and maintenance of sexual and asexual reproduction has been produced (reviewed in [2,6–9]). By contrast, little is known about the molecular underpinnings of transitions between reproductive systems [10]. Yet these molecular underpinnings have the potential to provide insights into the processes involved in the evolution of asexuality, and to help understand how sex is maintained. For example, sex is more easily maintained if asexuality evolves gradually in a sexual population than if it emerges suddenly via major effect mutations [11–13].

Some insight into the genetic basis of asexuality has been gained from studies of individual asexual lineages [14–17], but a broad comparative framework for exploring common principles of the molecular basis of asexuality is lacking. For example, a major unresolved question is whether independent transitions to asexuality involve similar or different molecular changes. To address these shortcomings, we explored the molecular underpinnings of asexuality in stick insects of the genus *Timema*, a genus of wingless, herbivorous insects native to the West coast of North America and the mountains of the Desert Southwest. This group is uniquely suited for comparative studies of asexuality, as asexuality has evolved at least seven times independently [18] (Fig. 1), allowing us to study convergence across replicate transitions from sexual to asexual reproduction. Furthermore, close sexual relatives are at hand for each asexual lineage for comparison. All asexual *Timema* species reproduce via obligate parthenogenesis [19], meaning that they evolved the ability to produce unreduced eggs which develop without fertilization by sperm. Additional phenotypic changes evolved convergently as adaptations to asexuality, including a reduced sperm storage organ, and reduced sexual pheromone production [20]. Thus, asexual *Timema* females are less attractive to sexual males [20], which use both airborne and contact signals to identify suitable mates [21–23], and even when copulations between sexual males and asexual females are forced under laboratory conditions, eggs are not fertilized [20].

70 To capture molecular changes associated with the evolution of asexuality we performed whole-
71 body and tissue-specific transcriptome sequencing (RNA-seq) on females from five sexual and
72 five asexual *Timema* species (Fig. 1). We chose two different tissues, the reproductive tract and
73 legs, to identify the molecular mechanisms underlying the production of asexual offspring
74 (reproductive tract), and adaptations to a celibate life (e.g. reduction of various different sexual
75 traits in the reproductive tract and legs). Note that the reproductive tract and leg samples actually
76 represent a collection of tissues, but we refer to them as tissues throughout for brevity. Whole-
77 body samples were included as they allow us to identify important changes that may be missing
78 in the tissue-specific transcriptomes. Using this approach, we identified convergent expression
79 changes which were likely driven by selection. We also observed changes specific to each sexual-
80 asexual species-pair which typically showed concerted changes across tissues, consistent with
81 being a product of drift [24,25]. Finally, to complement our expression analyses, we examined
82 patterns of molecular evolution in genes showing convergent expression changes following a
83 transition to asexuality.

Results

Transcriptomes and orthology

Reference transcriptome assemblies for each species were generated previously [26]. Bast et al. [26] also identified 3010 one-to-one orthologs, which were used as our transcriptome reference. For each tissue, orthologs with low expression (counts per million less than 0.5 in two or more libraries per species) were filtered prior to expression analyses. Thus, the final number of orthologs kept for analyses of whole-body, reproductive tract, and leg samples was 2984, 2753, and 2740, respectively.

Convergent gene expression changes

We identified convergent gene expression changes between sexual and asexual species by modelling gene expression as a function of species-pair (see Fig. 1), reproductive mode (sexual or asexual), and their interaction in edgeR [27]. In such a model, convergence is indicated by an overall effect of reproductive mode (FDR < 0.05), but no interaction (FDR > 0.05) (Supplemental Table 1). Approximately four times as many genes changed convergently in the reproductive tract (7%; 203/2754) and legs (8%; 206/2737) as compared to the whole-body (2%; 57/2985), perhaps reflecting the relative difficulty in identifying expression changes in complex tissue assemblies such as whole-bodies [28]. The amount of convergence we observe is considerable and approximately double what we would expect by chance, for all tissues (whole-body: $p = 0.0128$, reproductive tract: $p < 0.0001$, legs: $p < 0.0001$, Supplemental Fig. 1). The amount of change between sexual and asexual females was relatively small for convergent genes, with a mean fold change of approximately 1.4 (absolute $\log_2$ expression change for whole-body = 0.55, reproductive tract = 0.68, and legs = 0.46) (Fig. 2).

As expected for selection-driven gene expression changes [24,25], convergent changes between sexual and asexual species were highly tissue-specific. Only 22 of the convergent genes in the reproductive tract (203) and legs (206) overlapped between the two tissues, a value not significantly greater than expected by chance (Table 1). There was also little overlap between convergent genes in the two tissues and whole-bodies (Table 1, Supplemental Fig. 2). This supports the interpretation that the convergent changes in expression are driven by selection rather than by drift, as drift is more likely cause similar changes across multiple tissues. This

interpretation of selection also predicts that convergent genes will be involved in divergent functions in each tissue which, as we show below, is indeed the case.

Functional processes of convergently expressed genes

To detect convergence at the process level, we performed gene set enrichment analyses (GSEA). Briefly, we scored Gene Ontology (GO) terms according to the rank of convergent expression change of genes annotated to the terms; GO terms were then called significant if they had a better average rank than expected by chance (see Methods). More than 100 GO terms are enriched in each tissue studied (FDR < 0.05), providing strong support for convergence of biological processes between asexual species (Supplementary Tables 2-4). This signal is not dependent on any threshold at the gene level, and thus provides information on convergence at the process level due to small but consistent contributions from many genes. Consistent with the gene expression results, enriched GO terms were generally tissue specific; we found no significant overlap between GO-terms enriched in the legs and reproductive tract (11 shared terms, FDR = 0.123), between the legs and whole-body (4 shared terms, FDR = 0.799), or between whole-body and reproductive tract samples (10 shared terms, FDR = 0.064).

To reduce the number of enriched GO terms to examine we semantically clustered enriched GO terms using ReviGO [29] (Supplemental Tables 5-7). The annotations of convergent changes in the reproductive tract reflect the convergent evolution of parthenogenesis in asexual *Timema*, as they were linked to meiosis (meiotic spindle organization, meiosis II, centrosome duplication, meiosis I cytokinesis, meiosis II cytokinesis), and reproduction (growth of a germarium-derived egg chamber, sperm individualization, gamete generation). However, convergent changes were also linked to neuron development (neurogenesis, neuron development, neuron recognition), as well as several GO terms involved in development and metabolic processes for which the link to asexuality is less clear. In legs we identified GO terms involved in immune defence (response to fungus, regulation of production of molecular mediator of immune response, regulation of antimicrobial peptide production, regulation of humoral immune response), which may be because asexual females are no longer susceptible to the costs associated with diseases transmitted from sexual interactions, which can be considerable [30]. Convergent changes were also linked to sex determination (primary sex determination; soma, primary response to X:A ratio), which may control changes in the expression of sexual traits, and several metabolic processes. In whole-body samples we find some reproduction associated terms (courtship behavior, male

mating behavior, male courtship behavior, sperm storage, regulation of ovulation) as in the reproductive tract, and behavioral, and immune related terms (immune response-regulating cell surface receptor signaling pathway) as in legs, but also some unique terms relating to the cuticle (ecdysone, pupal chitin-based cuticle development).

Convergently expressed genes in whole-bodies show evidence for sexual trait decay

Several of the enriched functional processes described above are suggestive of sexual trait decay. Under this scenario we expect a reduction of purifying selection on genes underlying sexually dimorphic traits in asexual species, indicated by an increased accumulation of non-synonymous changes.

The power to detect differences in pN/pS or dN/dS between gene sets in asexuals is low, as genes are inherited as a single linkage group. Nevertheless, we found that genes showing convergent changes in expression in whole-bodies showed elevated pN/pS and dN/dS when compared to the genomic background (permuted t-test p-value for: pN/pS < 0.0001, dN/dS = 0.0084, Fig. 3, Supplemental Figs. 3 and 4), consistent with the idea of sexual trait decay. Sexual trait decay is further supported by the examination of functional annotations for such genes which include one gene (OG-2854) that is produced primarily in male accessory glands in *Drosophila*, and at least three other genes (OG-2197, OG-663, OG-1014) that are involved in pigment synthesis pathways (pigmentation is sexually dimorphic in *Timema* see Fig. 1B). In contrast, genes showing convergent changes in expression in the reproductive tract and legs did not show elevated pN/pS or dN/dS (Fig. 3, Supplemental Figs. 3 and 4), suggesting that convergent expression changes in these tissues do not coincide with reduced purifying selection acting on their sequences.

We conducted several additional analyses to check the robustness of our results and corroborate our interpretations. Firstly, we examined in detail the functions of candidate gene sets for which there was very strong evidence for convergent changes, and secondly, we used cross-species mapping to examine expression changes occurring across the whole transcriptome, rather than only in the subset of genes we identified as single copy orthologs between the 10 species. Both approaches support the results from our original analyses and are described below.

186 ***Strongly convergent candidate genes and their function***

187

188 Although all the convergent genes we identified showed an overall shift in expression across the
189 five species-pairs, often expression change in one or two of the pairs was small (<1.2 fold
190 change). We defined top candidate genes as convergent genes for which the absolute $\log_2$ fold
191 change in expression was more than 0.25 (~1.2 fold change) for all species-pairs. Most of these
192 top genes showed convergent shifts in the reproductive tract (36 genes, relative to 4 and 15 genes
193 for whole-body and legs, respectively) (Figure 4, Supplemental Table 1). The functions of these
194 candidate genes largely reflected the functional processes identified for the full set of convergently
195 expressed genes, and highlight a number of key genes potentially involved in producing asexual
196 phenotypes.

197

198 For the reproductive tract four genes are involved in meiotic spindle formation and centrosome
199 organization (OG-513, OG-1448, OG-1488, OG-314). In particular we find two genes (OG-1448,
200 OG-1488) belonging to a family of Elovl proteins that mediate elongation of very-long-chain fatty
201 acids, including an ortholog to *D. melanogaster* gene *bond*, which effects spindle formation and
202 has been shown to be important for meiotic, but not mitotic, cytokinesis [31]. In particular, *D.*
203 *melanogaster* males defective for *bond* commonly display two to four nuclei in spermatids causing
204 sterility. Female *bond* mutants are also infertile [31], although the mechanism is unknown. The
205 other two genes (OG-513, OG-314) have roles in centrosome function, including an ortholog to
206 *poc1* which is involved in centrosome formation [32]. Six genes (OG-758, OG-2002, OG-1478,
207 OG-1993, OG-2686, OG-148) were annotated with reproduction associated terms which may be
208 responsible for the convergent reproductive changes we observe between asexual and sexual
209 females. Interestingly, one gene, OG-511, is an ortholog to glucose dehydrogenase which is
210 important for sperm storage in female *D. melanogaster* [33]. Finally, we find that 11 genes (OG-
211 1195, OG-1478, OG-1841, OG-2197, OG-2808, OG-366, OG-445, OG-511, OG-705, OG-712,
212 OG-758, OG-810) have annotations to the nervous system. The majority of these appear to be
213 sensory in nature, and in particular seven are annotated with the GO term “sensory perception of
214 pain”. Changes in these genes may represent changes associated with female receptivity and
215 post-mating behaviour in asexual females, which are target of substances in the male ejaculate
216 [34–36].

217

218 For leg samples three genes (OG-1651, OG-2048 and OG-1081) are involved in immune defence.
219 In particular orthologs of both genes (*Trx-2* and *MP1*) are involved in the activation of melanisation

in response to fungal and bacterial infection in *D. melanogaster* [37,38]. Three genes are involved in cuticle development (OG-2221, OG-2738, and OG-2995). Orthologs of two other genes (OG-1371 and OG-2031) are involved in male specific behaviours (male courtship behavior and inter-male aggressive behavior) in *D. melanogaster* (*CaMKII* and *Fkbp14*) [39,40]. Since these genes are also expressed in females, changes to their expression may have resulted from the release of intralocus sexual conflict.

Whole-body samples had only four strong candidate genes, and all either have no annotation or have only broad GO-terms annotated. One potentially interesting gene, OG-2188, has an ortholog (*CG12237*) that has been associated with female sterility in *D. melanogaster* [41]. Finally, the remaining candidate genes across all tissues were either unannotated (12 genes) or only have very broad GO-terms annotated (10 genes).

Cross-species mapping

Using only the 3010 genes with 1-to-1 orthologs across all species could impact our ability to detect convergent changes since we only use a relatively small fraction of the total number of transcripts in each assembly (23435 to 37847; Supplemental Table 8). To investigate more genes, we mapped reads from all samples to genes from each species which had a reciprocal-best-blast-hit between species-pairs (which includes the 1-to-1 orthologs analysed above). This approach generated 10 different datasets (one for each species assembly), with between 15500 and 17583 genes. After filtering out genes with low expression (using cpm, see Methods) in each dataset, this approach allowed us to examine between 2.43-3.12 (dependent on species and tissue) times more genes than using the 1-to-1 orthologs (Supplemental Table 8). Results from this approach qualitatively confirmed the results found using only the 1-to-1 orthologs: the percentage of genes showing a convergent expression ranged from 4-5% for whole-body samples and 6-8% for the reproductive tract and leg samples, dependent on which of the species transcriptome was used (Supplemental Table 8), and GSEA produced similar enriched GO terms (Supplemental Tables 9-11)

Species-pair specific changes

The approach taken above allowed us to identify genes which showed convergent changes in expression across independent transitions to asexuality. This approach will not identify expression

changes confined to a single or few species-pairs. Such changes are clearly not convergent at the gene expression level, however, these changes could be convergent at the functional process level, whereby species-pair specific changes in gene expression are involved in common functional processes between species-pairs. To test this, we compared each asexual species to its closest sexual relative and called differentially expressed (DE) genes from each pairwise comparison.

The number of significantly DE genes between each pair varied greatly depending on species-pair and tissue (59 to 626, Supplementary Fig. 5), with a generally greater number of genes DE in leg tissue likely due to the smaller variation between replicates (common biological coefficient of variation was lowest for legs: whole-body = 0.314, reproductive tract = 0.340, legs = 0.238). There were no genes that showed overlap between all sexual-asexual species-pairs in any tissue (Fig. 5A). Examination of overlaps between pairs of sexual-asexual species-pairs found some overlapping genes, but these were close to the expectation by chance (Fig. 5B, for all levels see Supplemental Table 12). The majority of the DE genes also showed a significant interaction between species-pair and reproductive mode in the model used to identify convergently changing genes (whole-body = 69%, reproductive tract 66%, and legs = 81%), corroborating the finding that the vast majority of the DE genes are species-pair specific. Note the species-pair by reproductive mode interactions do not appear to be generated by one specific species-pair as generally genes DE between one species-pair were not DE between the other 4 species-pairs.

The DE genes of the different species pairs are not involved in convergent functional processes. Species-pair-specific genes were enriched for a number of GO terms, however no GO terms were found to overlap between all pairs in any tissue (Supplemental Fig. 6). Examination of overlaps between pairs of sexual-asexual species-pairs found some overlapping GO terms, but these were close to the expectation by chance (Supplemental Table 13). This pattern remained even when a more liberal approach, whereby related GO-terms were considered as a unit, was applied (Supplemental Fig. 7A). This overall lack of overlap suggests that species-pair-specific genes are not involved in producing convergent phenotypes but are instead the product of either lineage-specific selection or drift. These two processes are difficult to disentangle, but our results are more consistent with drift rather than lineage-specific selection. Indeed, species-pair specific genes showed similar changes in gene expression across tissues, in contrast to the mainly tissue-specific changes uncovered for convergently changing genes. The overlap of species-pair specific genes between tissues was significantly greater than expected by chance (Table 1).

288 Finally, these results were reproduced when examining a much larger set of genes (genes with
289 reciprocal-best-blast-hits between species-pairs, see above) as both genes DE between each
290 species-pair, and their enriched GO terms, showed little overlap (Supplemental Figs 8 and 9, and
291 Supplemental Tables 14 and 15).
292

Discussion

Asexuality has convergently evolved numerous times across the tree of life, and a large body of research focuses on the reasons why sexual reproduction persists in the face of competition from asexual lineages. By contrast, the molecular underpinnings of transitions from sexual to asexual reproduction remain largely unknown [10]. In this study we examined gene expression changes associated with transitions to asexuality across five independently evolved asexual lineages, in whole-bodies, reproductive tracts and legs. The changes we observe provide, for the first time, insights into the convergent evolution of asexuality at the molecular level.

We found evidence for convergent changes in gene expression in all three tissues. Three lines of evidence suggest that these changes are a product of selection. Firstly, parallel changes across multiple independent transitions represent strong evidence of selection and thus are unlikely to be due to drift [42,43]. Secondly, convergent changes were primarily tissue-specific. This finding is consistent with selection, because expression changes due to drift are likely to be correlated across tissues [24,25]. Indeed, the different functional roles of reproductive tracts and legs make it unlikely that selection would drive changes in the same genes in all tissues. Finally, the functional processes of convergently expressed genes mirror the changes observed at the phenotypic level, supporting the interpretation that these genes contribute to the convergent phenotypic changes observed between sexual and asexual females.

The overall amount of convergence is striking, particularly in the reproductive tract and legs with approximately 8% of genes showing a convergent shift in expression. Such a large amount of convergence suggests that the path from sexual reproduction to asexuality is strongly constrained, requiring changes to the same genes and biological processes in order to produce asexual phenotypes.

Convergent changes in gene expression reveal the mechanisms underlying the production of asexual offspring

Asexuality is a complex adaptation that includes two major components: the ability to produce viable asexual offspring, and secondary adaptive changes that would not have been selected for in sexual species (e.g. the reduction of costly sexual traits). A key change necessary for the production of asexual offspring is the ability to produce unreduced eggs [44]. Convergently

expressed genes in the reproductive tract were enriched for changes in meiosis, and in particular meiotic spindles, which are key for the proper division of cells during meiosis. Mutations in meiotic spindles have been shown to result in unreduced meiotic products in *D. melanogaster*, and specifically in two genes (*bond* [31] and *pelo* [45]) which show convergent changes in expression in asexual *Timema.* As such we suggest that these changes may underlie the non-reduction of eggs in asexual *Timema.* An alternative hypothesis is that since *Timema* reproduce parthenogenetically (and thus likely no longer recombine) changes in meiotic genes represent trait decay. Although possible, previous work has shown that, in fact, meiotic genes are not only retained in asexual lineages without damaging mutations, but often appear to be subject to selection for changes in expression, via duplication or differential upregulation of promoters [46–49]. Taken together with our results, we suggest that modifications to meiotic genes, specifically those that that disrupt meiotic cell division, are key in overcoming a major barrier to the evolution of asexuality: the production of unreduced eggs.

The production of unreduced eggs is not the only barrier to producing offspring asexually. In most species, sperm transfer essential components for the formation of a functioning centrosome [50,51]. This paternal contribution represents a second key barrier in the evolution of parthenogenesis in many systems [44]. However, in phasmids the centrosome is assembled without any contribution from sperm in both sexual and asexual species [52]. This may act as pre-adaptation for asexuality in stick insects and account, in part, for the large number of asexual stick insects.

A final barrier to asexual offspring production in many systems is egg activation. In many species mature oocytes are arrested at a specific stage (e.g. at metaphase II in mammals, and metaphase I in most insects), and must be activated by sperm to re-enter the cell-cycle [44,53,54]. In insects however, egg activation does not require sperm as activation is induced by the transit through the reproductive tract [55]. Despite this, ovulation and egg-laying rates are strongly tied to mating [56,57] meaning this signal must be modified in order for asexual insects to have normal levels of fecundity. In insects, the signal to a female that she has successfully mated is likely detected by sensory neurons in her reproductive tract [58]. Consistent with this, we find changes in gene expression linked to sensory neurons in the reproductive tract of asexual females, which may act to cue high levels of ovulation without mating. Alternatively, these changes may represent the decay of these neurons since they are no longer needed to detect mating events, or these changes may result from cessation of sexual conflict. Sensory neurons in the reproductive tract

are known targets of substances in the male ejaculate [34–36] to induce the release of eggs and to reduce female receptivity [57]. This manipulation is countered by female resistance adaptations which are likely costly, meaning that, following a transition to asexuality, there will be selection against them.

Convergent changes in gene expression show evidence for the decay of female sexual traits

Sexual traits in asexual females are often observed to be reduced or lost [59]. For instance, in insects, females typically produce pheromones as a sexual cue to attract males [60], and this cue has been repeatedly reduced or lost in several asexual species (see [59]), including *Timema* [20]. Such trait decay can be the result of either reduced purifying selection acting on traits that are now selectively neutral, or selection to reduce the cost of producing sexual traits. In asexual *Timema* reproductive decay has been primarily attributed to selection rather than reduced purifying selection, as reproductive trait decay in very young asexual lineages is as extensive as in old ones [20].

Convergent gene expression changes underlying the decay of reproductive traits are mostly observed in *Timema* whole-bodies. In particular, we find enrichment of terms associated with sperm storage and sexual behaviour. Changes in the legs were less obviously associated with reproductive trait decay, however we do find changes in genes involved in cuticle development, pigment biosynthesis, sensory perception of touch, and changes in sexual behavior. These changes could represent the reproductive decay of both sexual cues (e.g. cuticular hydrocarbons and pigmentation which are both important for mate choice in insects (reviewed in [61]), and their detection (via sensory receptors on the leg (reviewed in [62]). In addition, we also find changes in genes associated with sex determination in the soma, including *sex-lethal*, a master-feminizing switch in *Drosophila* [63] which may have a major influence the development of many sexual traits in the legs.

Although we focus on expression, it is possible that the decay of sexual traits is also evident at the sequence level. By examining the coding regions of genes, we found evidence for reduced purifying selection acting on the sequence of genes showing convergent expression changes in the whole-body. This suggests, that in some cases, the reduction of sexual traits may be

accomplished by both expression and sequence changes, which potentially act interactively to produce a phenotypic change.

Unexpectedly, we also find changes to immune function in the legs and whole-body, the majority of which show down-regulation in asexual females. A possible explanation for this is that asexual females are likely to face a reduced number of immune challenges compared to sexual females due to the elimination of sexually transmitted diseases, the costs of which can be considerable, even shaping the evolution of many aspects of an organism's life history, such as mate choice, mating rate, and sexual signal investment [30,64]. As such we suggest asexual females may be reducing the allocation of resources to immune function due to the absence sexually transmitted diseases. This effect may be particularly strong in solitary species such as *Timema*, where the majority of socially transmitted diseases come from sexual interactions.

Species-pair specific changes

In addition to convergent changes, we also identified many species-pair specific gene expression changes. In contrast to convergent genes, species-pair specific genes showed common shifts in expression across tissues, and inconsistent associations with functional processes between species-pairs, that were largely unrelated to asexual phenotypes. Taken together, these results suggest that the majority of changes we observe from a single sex-asex species-pair comparison are due to drift rather than selection. Our findings thus highlight the problem of drawing inferences on the causes or consequences of asexuality from the examination of only a single transition to asexuality, whereas examining several transitions allows us to disentangle adaptive changes and those due to drift.

Overall, we find evidence for a striking number of convergent changes across five transitions to asexuality. The amount of molecular convergence to expect, however, is dependent on several factors including the complexity of the phenotype, and the size of the mutational target [65]. For instance, here we find that a key change required for asexual reproduction, the production of unreduced eggs, likely requires changes to meiotic spindle regulation. The pathways that govern meiotic spindle regulation are relatively small in number [66], meaning that only a small minority of genes are likely able to confer the relevant changes, making the chance of convergence for this trait relatively high. In contrast, the observed reduction of sexual traits could be produced by changes to numerous genes and pathways (i.e. there is a large mutational target) making

428 convergent changes for these traits less likely. Despite this, our and previous studies examining
429 trait loss have also demonstrated a high amount of convergence [67–70], implying that certain
430 genes have a disproportionate role in not only the convergent evolution of novel phenotypes, but
431 also in their convergent loss [65,71].

Methods

Samples

Females for whole-body samples were collected from the field as juveniles in spring 2013. All individuals were then raised in common garden conditions (23°C, 12h:12h, 60% humidity, fed with Ceanothus cuttings) until eight days following their final molt. Prior to RNA extraction, individuals were fed with artificial medium for two days to avoid RNA contamination with gut content and then frozen at -80°C. Individuals used for tissue-specific samples were collected in spring 2014 as juveniles and raised in the same common-garden conditions as whole-body samples. For leg samples three legs were used from each individual (one foreleg, one midleg, and one hindleg). Reproductive tracts were dissected to consist of ovaries, oviducts and spermatheca. Note the same individuals were used for leg and reproductive tract samples. Collection locations for all samples are given in Supplemental Table 16.

RNA extraction and sequencing

The three biological replicates per species and tissue consisted of 1-9 individuals per replicate, which were combined prior to RNA extraction (207 individuals in 90 replicates in total; see Supplemental Table 16). RNA extraction was performed by freezing individuals in liquid nitrogen followed by addition of Trizol (Life Technologies) before being homogenized using mechanical beads (Sigmund Lindner). Chloroform and ethanol were then added to the samples and the aqueous layer transferred to RNeasy MinElute Columns (Qiagen). RNA extraction was then completed using an RNeasy Mini Kit following the manufacturer's instructions. RNA quantity and quality was measured using NanoDrop (Thermo Scientific) and Bioanalyzer (Agilent). Strand-specific library preparation and single-end sequencing (100 bp, HiSeq2000) were performed at the Lausanne Genomic Technologies Facility.

The 90 libraries produced a total of just over 3 billion single-end reads. Four whole-body and six tissue-specific libraries produced significantly more reads than the average for the other samples. To reduce any influence of this on downstream analyses, these libraries were sampled down to approximately the average number of reads for whole-body or tissue-specific libraries respectively using seqtk (https://github.com/lh3/seqtk Version: 1.2-r94).

466 ***Transcriptome references***

467

468 *De novo* reference transcriptome assemblies for each species were generated previously [26].
469 For our analyses we used the 3010 one-to-one orthologs present in all 10 *Timema* species as
470 identified by Bast et al. [26]. Identified ortholog sequences varied in length among different
471 species. Since length variation might influence estimates of gene expression, we aligned
472 orthologous sequences using PRANK (v.100802, default options) [72] and trimmed them using
473 alignment_trimmer.py [73] to remove overhanging gaps at the ends of the alignments. If the
474 alignment contained a gap of greater than 3 bases then sequence preceding or following the
475 alignment gap (whichever was shortest) was discarded. Three genes were discarded at this stage
476 as the trimmed length of sequence was <300 bp. These trimmed sequences were then used as
477 reference transcriptomes for read mapping. Note that genes with significant Blast hits to rRNA
478 sequences were removed prior to mapping.

479

480 ***Read trimming and mapping***

481

482 Raw reads were trimmed before mapping. Firstly CutAdapt [74] was used to trim adapter
483 sequences from the reads. Reads were then quality trimmed using Trimmomatic v 0.36 [75]: first
484 clipping leading or trailing bases with a phred score of <10 from the read, before using a sliding
485 window from the 5' end to clip the read if 4 consecutive bases had an average phred score of
486 <20. Following quality trimming any reads <80 bp in length were discarded. Surviving reads from
487 each library were then mapped separately to the reference transcriptome using Kallisto (v. 0.43.1)
488 [76] with the following options -l 210 -s 25 --bias --rf-stranded for whole-body samples and -l 370
489 -s 25 --bias --rf-stranded for tissue specific samples (the -l option was different for whole-body
490 and tissue specific samples as the fragment length for these libraries was different).

491

492 ***Differential expression analysis***

493

494 Expression analyses were performed using the Bioconductor package EdgeR (v. 3.18.1) [27] in
495 R (v. 3.4.1) [77]. Analyses were done separately for each tissue. Genes with counts per million
496 less than 0.5 in 2 or more libraries per species were excluded from expression analyses.
497 Normalization factors for each library were computed using the TMM method in EdgeR. To
498 estimate dispersion we then fit a generalized linear model (GLM) with negative binomial
499 distribution with the terms species-pair, reproductive mode and their interaction. We used a GLM

likelihood ratio test to determine significance of model terms for each gene by comparing appropriate model contrasts. P-values were corrected for multiple tests using Benjamini and Hochberg's algorithm [78], with statistical significance set to 5%. Using this approach, we classified genes as convergently differentially expressed when there was a significant effect of reproductive mode (FDR < 0.05) but no interaction effect of species-pair by reproductive mode (FDR > 0.05). Differentially expressed genes within each species-pair were identified using pairwise contrasts between each sexual and asexual pair.

To determine if genes DE within each species-pair and tissue show greater than expected number of overlapping genes we used the SuperExactTest package (v. 0.99.4) [79] in R which calculates the probability of multi-set intersections. When examining multiple intersections p-values were multiple test corrected using Benjamini and Hochberg's algorithm implemented in R.

To test if the observed number of convergent genes was significantly greater than expected by chance we performed a permutation test by whereby, for the read counts of each gene, we randomly switched the assignment of reproductive mode (sexual or asexual) within a species-pair. Note that all biological replicates from a particular group were always assigned to the same reproductive mode (i.e. In the event of a switch, all sexual replicates were assigned as asexual, and vice versa). This process was repeated to produce 10,000 permuted data sets, which were then ran through the gene expression pipeline described above to generate a distribution of the number of convergent genes we expect to find by chance.

Go term analysis

Genes were functionally annotated using Blast2GO (version 4.1.9) [80] as follows: sequences from each sexual species were compared with BlastX to either NCBI's nr-arthropod or *Drosophila melanogaster* (drosoph) databases, keeping the top 20 hits with e-values $<1 \times 10^{-3}$. Interproscan (default settings within Blast2GO) was then run for each sequence, and the results merged with the blast results to obtain GO terms. This produced two sets of functional annotations, one derived from all arthropods and one specifically from *Drosophila melanogaster*. The *D. melanogaster* GO term annotation generated around four times more annotations per sequence than NCBI's nr-arthropod database. We therefore conducted all subsequent analyses using the GO terms derived

from *D. melanogaster*, but note that results using the annotations from all arthropods were qualitatively the same (see Supplementary Fig. 7B).

We conducted gene set enrichment analyses using the R package TopGO (v. 2.28.0) [81] using the elim algorithm to account for the GO topology. Gene set enrichment analyses identify enriched GO terms in a threshold-free way, by finding GO-terms that are overrepresented at the top of a ranked list of genes. For comparisons within a species-pair, genes were ranked by FDR; to identify enrichment of convergent genes, genes were ranked by FDR value for reproductive mode, with the FDR value for genes that showed a significant lineage by reproductive mode set to 1. GO terms were considered to be significantly enriched when $p < 0.05$. Enriched GO terms were then semantically clustered using ReviGO [29] to aid interpretation.

The significance of overlapping GO terms was determined using SuperExactTest as described above. The hierarchical nature of GO terms generates a bias towards finding a significant amount of overlap, since enrichment terms are non-independent. It is however possible that the complexity of the GO term hierarchy could lead to convergent functional processes being overlooked. For instance if a GO term is enriched in one comparison, but its parent term is enriched in another comparison, then there would be no apparent overlap. To address this, we also looked at the amount of 'linked overlap' of GO terms, whereby significant GO terms were first clustered together based on parent or child terms.

For the GO term enrichment analyses of convergently differential expressed genes we used only the annotation from *T. bartmani* as it had the most number of sequences annotated. Annotations to each of the other species were very similar to those from *T. bartmani,* with 80% of annotations being identical across all 5 species annotations. The remaining 20% of sequences were typically characterized by an additional term in one or more of the species. For comparisons within a lineage we used the annotation of the sexual species in that lineage. Although the annotations are very similar across all ten species the small differences in annotation could create differences in the amount of overlap observed between contrasts (e.g. if a term is annotated to an ortholog in one annotation but not another). To examine this, we repeated the analysis using only annotations from *T. bartmani*. This produced a virtually identical result (Supplemental Fig. 7C) as when using the species-pair specific annotations.

Polymorphism and divergence

To test for differences in the rate of evolutionary divergence between gene categories, we used dN/dS ratios for each of the one-to-one orthologs from [26]. To obtain an estimate for pN/pS reads from the whole-body libraries for each asexual species were mapped to the reference using RSEM/bowtie2 with default parameters and fragment length mean = 200 fragment length sd = 100 [82,83]. Samtools v1.2 was then used to create an mpileup file, which was filtered with VarScan v2.3.2 (minimum coverage = 20, minor allele frequency = 10%, and minimum average phred quality = 20) to obtain SNPs. To identify nonsynonymous and synonymous segregating polymorphisms we identified the n-fold degenerate positions following Li et al. [84] from which pN, pS and (pN/pS) could be calculated per gene. Comparison of mean pN/pS and dN/dS between convergent and non-convergent (background) genes was conducted using a permutation t-test (number of permutations = 10000) in R.

Cross-species mapping

All of the above analyses used only the one-to-one orthologs. To examine a larger fraction of the transcriptome we produced species-pair references by using a reciprocal blast between the assemblies of sexual-asexual sister species (blastN, minimum e-val = 0.00001, minimum query coverage = 30%). Prior to this step potential contaminants were filtered from these by blasting transcripts to local versions of the nt (using blastN, default options except task blastn, max_target_seqs = 10) and nr (using blastX, default options except, max_target_seqs = 10) databases (downloaded: 07/08/2016) using NCBI's blast client (v. 2.2.30+). Blast hits with an e-value > 0.0000001 were discarded. The remaining blast hits were used to assign a phylum to sequences if >=50% of Blast hits came from one phylum (in the event of a tie, the taxa with the highest e-value was used as a tiebreaker). Transcripts that were assigned to a non-arthropoda phylum were discarded (note that transcripts with no Blast hits or that blasted to mixed phyla were retained). This filtering removed between 4-8% of transcripts (see Supplemental Table 8). Reads of each species were then mapped to each species-pair reference in the same way as for the 1-to-1 orthologs. Differential expression analyses and GO-term enrichment analyses were then repeated as described above.

Data

Raw reads have been deposited in the SRA. Accession codes are given in Supplementary Table 16. Scripts for the analyses in this paper are available at: https://github.com/DarrenJParker/Timema_convergent_gene_expression.

Acknowledgments

This study was supported by Swiss FNS grants PP00P3_170627, PP00P3_139013, and CRSII3_160723. We would like to thank Chloé Larose and Bart Zijlstra for their assistance in the field.

References

1. Maynard Smith J. The origin and maintenance of sex. In: Williams GC, editor. Group Selection. Chicago: Aldine Atherton; 1971. pp. 163–175.
2. Bell G. The Masterpiece of Nature: The Evolution and Genetics of Sexuality. CUP Archive; 1982.
3. Lehtonen J, Jennions MD, Kokko H. The many costs of sex. Trends Ecol Evol. 2012;27: 172–178.
4. Avise JC. Clonality: The Genetics, Ecology and Evolution of Sexual Abstinence in Vertebrate Animals. Oxford: Oxford University Press; 2008.
5. van der Kooi CJ, Matthey-Doret C, Schwander T. Evolution and comparative ecology of parthenogenesis in haplodiploid arthropods. Evolution Letters. Wiley Online Library; 2017;1: 304–316.
6. Lewis WM Jr. The cost of sex. In: Stearns SC, editor. The evolution of sex and its consequences. Verlag: Birkhauser; 1987. pp. 33–57.
7. West SA, Lively CM, Read AF. A pluralist approach to sex and recombination. J Evol Biol. Wiley Online Library; 1999;12: 1003–1012.
8. Otto SP. The evolutionary enigma of sex. Am Nat. journals.uchicago.edu; 2009;174 Suppl 1: S1–S14.
9. Neiman M, Lively CM, Meirmans S. Why Sex? A Pluralist Approach Revisited. Trends Ecol Evol. 2017; doi:10.1016/j.tree.2017.05.004
10. Neiman M, Sharbel TF, Schwander T. Genetic causes of transitions from sexual

631 reproduction to asexuality in plants and animals. J Evol Biol. 2014;27: 1346–1359.
632 11. Templeton AR. The Prophecies of Parthenogenesis. In: Dingle H, Hegmann JP, editors.
633 Evolution and Genetics of Life Histories. New York: Springer; 1982. pp. 75–101.
634 12. Burt A. Perspective: Sex, recombination, and the efficacy of selection—Was Weismann
635 right? Evolution. The Society for the Study of Evolution; 2000;54: 337–351.
636 13. Schwander T, Vuilleumier S, Dubman J, Crespi BJ. Positive feedback in the transition from
637 sexual reproduction to parthenogenesis. Proc Biol Sci. 2010;277: 1435–1442.
638 14. Jaquiéry J, Stoeckel S, Larose C, Nouhaud P, Rispe C, Mieuzet L, et al. Genetic Control of
639 Contagious Asexuality in the Pea Aphid. PLoS Genet. Public Library of Science; 2014;10:
640 e1004838.
641 15. Lynch M, Seyfert A, Eads B, Williams E. Localization of the genetic determinants of meiosis
642 suppression in Daphnia pulex. Genetics. 2008;180: 317–327.
643 16. Eads BD, Tsuchiya D, Andrews J, Lynch M, Zolan ME. The spread of a transposon
644 insertion in *Rec8* is associated with obligate asexuality in *Daphnia*. Proc Natl Acad Sci U S
645 A. National Acad Sciences; 2012;109: 858–863.
646 17. Innes DJ, Hebert PDN. The origin and genetic basis of obligate parthenogenesis in
647 *Daphnia pulex*. Evolution. [Society for the Study of Evolution, Wiley]; 1988;42: 1024–1035.
648 18. Schwander T, Henry L, Crespi BJ. Molecular evidence for ancient asexuality in *Timema*
649 stick insects. Curr Biol. Elsevier Ltd; 2011;21: 1129–1134.
650 19. Schwander T, Crespi BJ. Multiple direct transitions from sexual reproduction to apomictic
651 parthenogenesis in *Timema* stick insects. Evolution. 2009;63: 84–103.
652 20. Schwander T, Crespi BJ, Gries R, Gries G. Neutral and selection-driven decay of sexual
653 traits in asexual stick insects. Proc Biol Sci. 2013;280: 20130823.
654 21. Nosil P, Crespi BJ, Gries R, Gries G. Natural selection and divergence in mate preference
655 during speciation. Genetica. 2007;129: 309–327.
656 22. Arbuthnott D, Crespi BJ. Courtship and mate discrimination within and between species of
657 *Timema* walking-sticks. Anim Behav. Elsevier Ltd; 2009;78: 53–59.
658 23. Schwander T, Arbuthnott D, Gries R, Gries G, Nosil P, Crespi BJ. Hydrocarbon divergence
659 and reproductive isolation in *Timema* stick insects. BMC Evol Biol. BMC Evolutionary
660 Biology; 2013;13: 151.
661 24. Liang C, Musser JM, Cloutier A, Prum RO, Wagner GP. Pervasive correlated evolution in
662 gene expression shapes cell and tissue type transcriptomes. Genome Biol Evol.
663 academic.oup.com; 2018; doi:10.1093/gbe/evy016
664 25. Blekhman R, Oshlack A, Chabot AE, Smyth GK, Gilad Y. Gene regulation in primates

665 evolves under tissue-specific selection pressures. PLoS Genet. 2008;4: e1000271.
666 26. Bast J, Parker DJ, Dumas Z, Jalvingh KM, Tran Van P, Jaron KS, et al. Consequences of
667 asexuality in natural populations: insights from stick insects. Mol Biol Evol. 2018;
668 doi:10.1093/molbev/msy058
669 27. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential
670 expression analysis of digital gene expression data. Bioinformatics. 2010;26: 139–140.
671 28. Johnson BR, Atallah J, Plachetzki DC. The importance of tissue specificity for RNA-seq:
672 highlighting the errors of composite structure extractions. BMC Genomics. 2013;14: 586.
673 29. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO summarizes and visualizes long lists of
674 gene ontology terms. PLoS One. 2011;6: e21800.
675 30. Knell RJ, Webberley KM. Sexually transmitted diseases of insects: distribution, evolution,
676 ecology and host behaviour. Biol Rev Camb Philos Soc. 2004;79: 557–581.
677 31. Szafer-Glusman E, Giansanti MG, Nishihama R, Bolival B, Pringle J, Gatti M, et al. A role
678 for very-long-chain fatty acids in furrow ingression during cytokinesis in *Drosophila*
679 spermatocytes. Curr Biol. 2008;18: 1426–1431.
680 32. Blachon S, Cai X, Roberts KA, Yang K, Polyanovsky A, Church A, et al. A proximal
681 centriole-like structure is present in Drosophila spermatids and can serve as a model to
682 study centriole duplication. Genetics. 2009;182: 133–144.
683 33. Bloch Qazi MC, Heifetz Y, Wolfner MF. The developments between gametogenesis and
684 fertilization: ovulation and female sperm storage in *Drosophila melanogaster*. Dev Biol.
685 2003;256: 195–211.
686 34. Heifetz Y, Wolfner MF. Mating, seminal fluid components, and sperm cause changes in
687 vesicle release in the *Drosophila* female reproductive tract. Proc Natl Acad Sci U S A.
688 2004;101: 6261–6266.
689 35. Sakai T, Kasuya J, Kitamoto T, Aigaki T. The *Drosophila* TRPA channel, Painless,
690 regulates sexual receptivity in virgin females. Genes Brain Behav. 2009;8: 546–557.
691 36. Heifetz Y, Lindner M, Garini Y, Wolfner MF. Mating regulates neuromodulator ensembles at
692 nerve termini innervating the *Drosophila* reproductive tract. Curr Biol. 2014;24: 731–737.
693 37. Jin LH, Shim J, Yoon JS, Kim B, Kim J, Kim-Ha J, et al. Identification and functional
694 analysis of antifungal immune response genes in *Drosophila*. PLoS Pathog. 2008;4:
695 e1000168.
696 38. Tang H, Kambris Z, Lemaitre B, Hashimoto C. Two proteases defining a melanization
697 cascade in the immune system of *Drosophila*. J Biol Chem. 2006;281: 28097–28104.
698 39. Mehren JE, Griffith LC. Cholinergic neurons mediate CaMKII-dependent enhancement of

699 courtship suppression. Learn Mem. 2006;13: 686–689.
700 40. Edwards AC, Zwarts L, Yamamoto A, Callaerts P, Mackay TFC. Mutations in many genes
701 affect aggressive behavior in *Drosophila melanogaster*. BMC Biol. 2009;7: 29.
702 41. Sopko R, Foos M, Vinayagam A, Zhai B, Binari R, Hu Y, et al. Combining genetic
703 perturbations and proteomics to examine kinase-phosphatase networks in *Drosophila*
704 embryos. Dev Cell. 2014;31: 114–127.
705 42. Losos JB. Convergence, adaptation, and constraint. Evolution. Wiley Online Library;
706 2011;65: 1827–1840.
707 43. Natarajan C, Hoffmann FG, Weber RE, Fago A, Witt CC, Storz JF. Predictable
708 convergence in hemoglobin function has unpredictable molecular underpinnings. Science.
709 2016;354: 336–339.
710 44. Engelstädter J. Constraints on the evolution of asexual reproduction. Bioessays. 2008;30:
711 1138–1150.
712 45. Eberhart CG, Wasserman SA. The *pelota* locus encodes a protein required for meiotic cell
713 division: an analysis of G2/M arrest in *Drosophila* spermatogenesis. Development.
714 1995;121: 3477–3486.
715 46. Hanson SJ, Schurko AM, Hecox-Lea B, Welch DBM, Stelzer C-P, Logsdon JM Jr. Inventory
716 and phylogenetic analysis of meiotic genes in monogonont rotifers. J Hered. 2013;104:
717 357–370.
718 47. Srinivasan DG, Fenton B, Jaubert-Possamai S, Jaouannet M. Analysis of meiosis and cell
719 cycle genes of the facultatively asexual pea aphid, *Acyrthosiphon pisum* (Hemiptera:
720 Aphididae). Insect Mol Biol. 2010;19 Suppl 2: 229–239.
721 48. Raborn RT, Spitze K, Brendel VP, Lynch M. Promoter Architecture and Sex-Specific Gene
722 Expression in *Daphnia pulex*. Genetics. 2016;204: 593–612.
723 49. Warren WC, García-Pérez R, Xu S, Lampert KP, Chalopin D, Stöck M, et al. Clonal
724 polymorphism and high heterozygosity in the celibate genome of the Amazon molly. Nat
725 Ecol Evol. 2018; doi:10.1038/s41559-018-0473-y
726 50. Schatten G. The centrosome and its mode of inheritance: the reduction of the centrosome
727 during gametogenesis and its restoration during fertilization. Dev Biol. 1994;165: 299–335.
728 51. Manandhar G, Schatten H, Sutovsky P. Centrosome reduction during gametogenesis and
729 its significance. Biol Reprod. 2005;72: 2–13.
730 52. Marescalchi O, Zauli C, Scali V. Centrosome dynamics and inheritance in related sexual
731 and parthenogenetic *Bacillus* (Insecta Phasmatodea). Mol Reprod Dev. 2002;63: 89–95.
732 53. Nishiyama T, Tachibana K, Kishimoto T. Cytostatic arrest: post-ovulation arrest until

733 fertilization in metazoan oocytes. Oogenesis: The Universal Process. John Wiley & Sons,
734 Ltd; 2010; 357–387.
735 54. Stricker SA. Comparative biology of calcium signaling during fertilization and egg activation
736 in animals. Dev Biol. 1999;211: 157–176.
737 55. Sartain CV, Wolfner MF. Calcium and egg activation in *Drosophila*. Cell Calcium. 2013;53:
738 10–15.
739 56. Eberhard W. Female Control: Sexual Selection by Cryptic Female Choice. Princeton
740 University Press; 1996.
741 57. Gillott C. Male accessory gland secretions: modulators of female reproductive physiology
742 and behavior. Annu Rev Entomol. 2003;48: 163–184.
743 58. Yapici N, Kim Y-J, Ribeiro C, Dickson BJ. A receptor that mediates the post-mating switch
744 in *Drosophila* reproductive behaviour. Nature. 2008;451: 33–37.
745 59. van der Kooi CJ, Schwander T. On the fate of sexual traits under asexuality. Biol Rev
746 Camb Philos Soc. 2014;89: 805–819.
747 60. Greenfield MD. Signalers and Receivers. Oxford: Oxford University Press; 2002.
748 61. Hunt J, Sakaluk SK. Mate choice. The evolution of insect mating systems Oxford Univ
749 Press, Oxford, UK. books.google.com; 2014; 129–158.
750 62. Dahanukar A, Hallem EA, Carlson JR. Insect chemoreception. Curr Opin Neurobiol.
751 2005;15: 423–430.
752 63. Cline TW, Dorsett M, Sun S, Harrison MM, Dines J, Sefton L, et al. Evolution of the
753 Drosophila feminizing switch gene *Sex-lethal*. Genetics. 2010;186: 1321–1336.
754 64. Kokko H, Ranta E, Ruxton G, Lundberg P. Sexually transmitted disease and the evolution
755 of mating systems. Evolution. 2002;56: 1091–1100.
756 65. Martin A, Orgogozo V. The Loci of repeated evolution: a catalog of genetic hotspots of
757 phenotypic variation. Evolution. 2013;67: 1235–1250.
758 66. Bennabi I, Terret M-E, Verlhac M-H. Meiotic spindle assembly and chromosome
759 segregation in oocytes. J Cell Biol. 2016;215: 611–619.
760 67. Sackton TB, Grayson P, Cloutier A, Hu Z, Liu JS, Wheeler NE, et al. Convergent regulatory
761 evolution and the origin of flightlessness in palaeognathous birds [Internet]. bioRxiv. 2018.
762 p. 262584. doi:10.1101/262584
763 68. Gross JB, Borowsky R, Tabin CJ. A novel role for *Mc1r* in the parallel evolution of
764 depigmentation in independent populations of the cavefish *Astyanax mexicanus*. PLoS
765 Genet. 2009;5: e1000326.
766 69. Whittall JB, Voelckel C, Kliebenstein DJ, Hodges SA. Convergence, constraint and the role

of gene expression during adaptive radiation: floral anthocyanins in *Aquilegia*. Mol Ecol. 2006;15: 4645–4657.

70. Partha R, Chauhan BK, Ferreira Z, Robinson JD, Lathrop K, Nischal KK, et al. Subterranean mammals show convergent regression in ocular genes and enhancers, along with adaptation to tunneling. Elife. 2017;6. doi:10.7554/eLife.25884
71. Stern DL. The genetic causes of convergent evolution. Nat Rev Genet. 2013;14: 751–764.
72. Löytynoja A, Goldman N. An algorithm for progressive multiple alignment of sequences with insertions. Proc Natl Acad Sci U S A. 2005;102: 10557–10562.
73. Parker DJ. fasta_tools v1.2. Zenodo https://zenodo.org/record/59775. 2016; doi:10.5281/zenodo.162913
74. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011;17: 10–12.
75. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30: 2114–2120.
76. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol. 2016;34: 525–527.
77. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2017. Available: https://www.R-project.org/
78. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol. [Royal Statistical Society, Wiley]; 1995;57: 289–300.
79. Wang M, Zhao Y, Zhang B. Efficient test and visualization of multi-set intersections. Sci Rep. The Author(s); 2015;5: 16923.
80. Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, et al. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res. 2008;36: 3420–3435.
81. Alexa A, Rahnenfuhrer J. topGO: Enrichment Analysis for Gene Ontology. 2016.
82. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9: 357–359.
83. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011;12: 323.
84. Li WH, Wu CI, Luo CC. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. Mol Biol Evol. 1985;2: 150–174.

801 85. Riesch R, Muschick M, Lindtke D, Villoutreix R, Comeault AA, Farkas TE, et al. Transitions
802 between phases of genomic differentiation during stick-insect speciation. Nature Ecology &
803 Evolution. Nature Publishing Group; 2017;1: 0082.
804
805
806

807 **Table 1 |** Overlap of differentially expressed genes between different tissue types. Number of
808 genes expected by chance given in parentheses. P-values are from a fisher's exact test corrected
809 for multiple tests. Species names are abbreviated as follows: Tbi = *T. bartmani*, Tce = *T. cristinae*,
810 Tps = *T. poppensis*, Tcm = *T. californicum*, Tpa = *T. podura*, Tte = *T. tahoe*, Tms = *T. monikensis*,
811 Tdi = *T. douglasi*, Tsi = *T. shepardi*, and Tge = *T. genevievae*.

	Reproductive tract & legs	**Whole-body & legs**	**Whole-body & reproductive tract**
Convergent genes	22 (16) p = 0.14	7 (5) p = 0.21	4 (4) p = 0.65
Species-pair (Tbi-Tte)	62 (24) $p = 1.30 \times 10^{-14}$	68 (29) $p = 4.57 \times 10^{-14}$	25 (7) $p = 4.53 \times 10^{-9}$
Species-pair (Tce-Tms)	88 (47) $p = 1.36 \times 10^{-10}$	48 (9) $p = 6.19 \times 10^{-25}$	47 (14) $p = 3.86 \times 10^{-16}$
Species-pair (Tcm-Tsi)	42 (11) $p = 4.45 \times 10^{-15}$	24 (5) $p = 4.64 \times 10^{-11}$	19 (3) $p = 4.64 \times 10^{-11}$
Species-pair (Tpa-Tge)	153 (134) $p = 1.99 \times 10^{-2}$	102 (60) $p = 1.22 \times 10^{-9}$	86 (54) $p = 8.75 \times 10^{-07}$
Species-pair (Tps-Tdi)	56 (20) $p = 1.31 \times 10^{-15}$	59 (37) $p = 6.08 \times 10^{-5}$	32 (7) $p = 7.81 \times 10^{-14}$

812
813
814
815
816

Figures

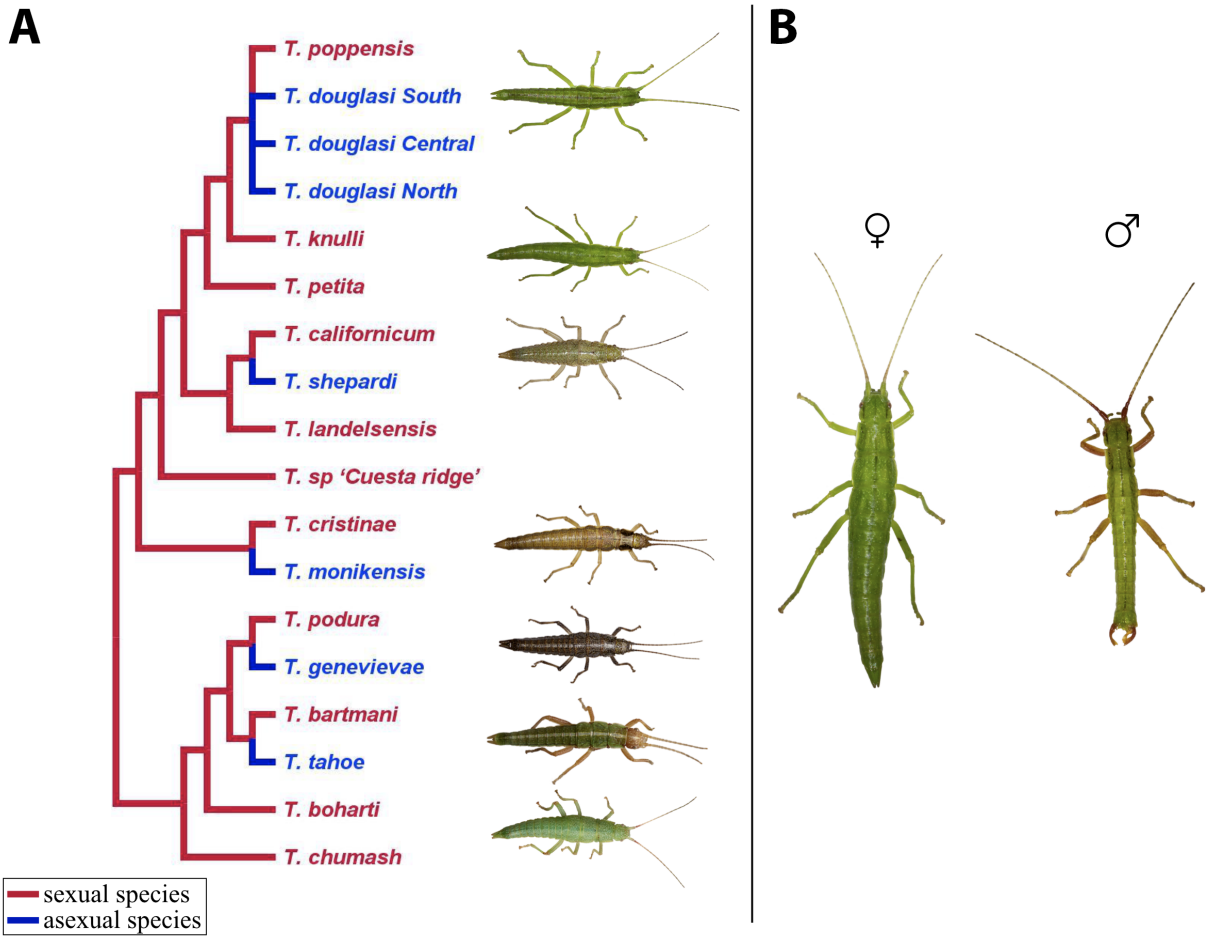


Figure 1 | **A.** Phylogeny of described *Timema* species (redrawn from [85] with asexual species added from [18]). Sexually reproducing species are shown in red, independently derived asexual lineages in blue. Our study used the five asexual species (for *T. douglasi* only the southern lineage was used) and their sexual sister species. **B.** Sexual dimorphism in Timema (*T. knulli*).

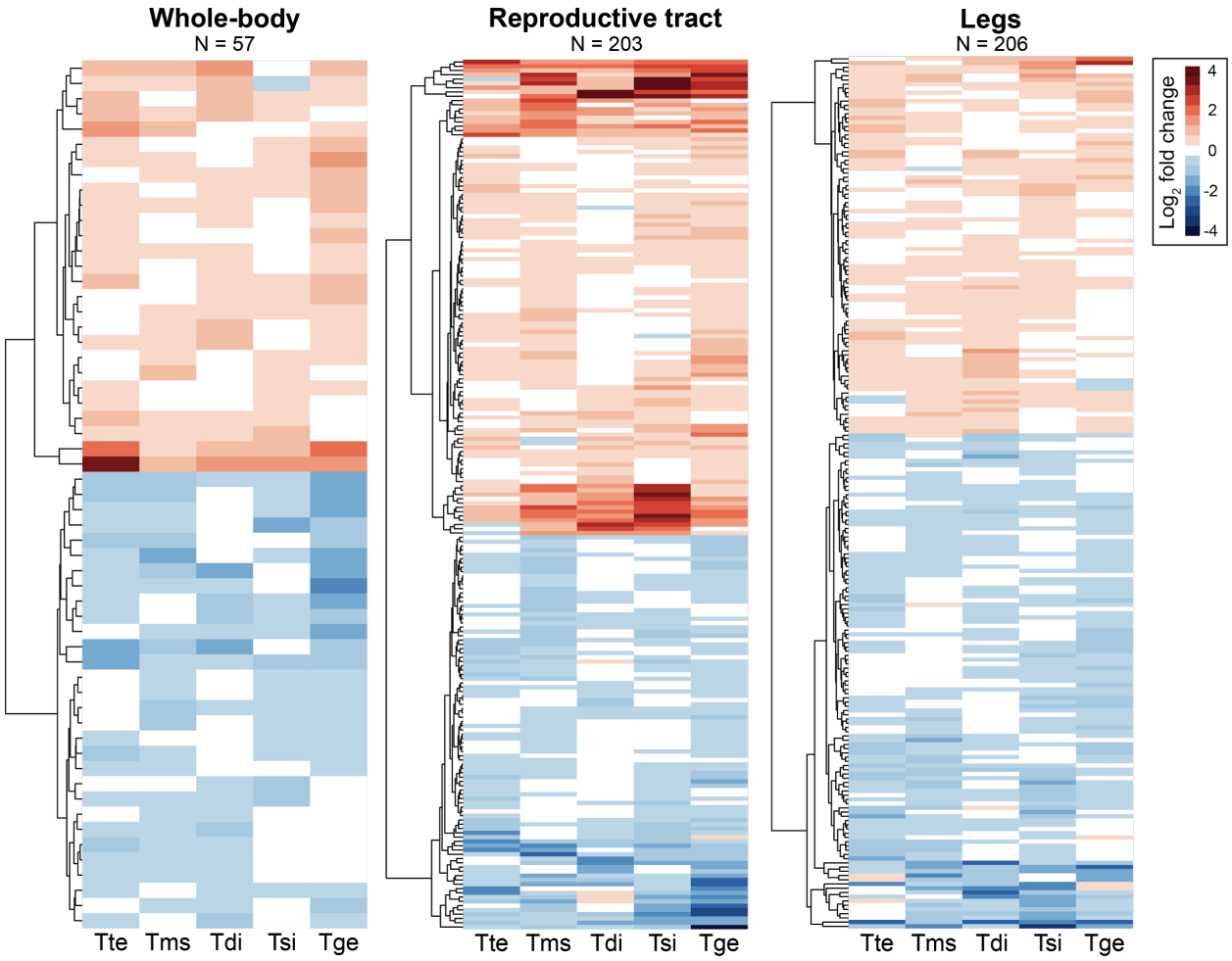


Figure 2 | Heatmaps of genes showing convergent gene expression changes between sexual and asexual females for whole-bodies, reproductive tract, and legs. Species names are abbreviated as follows: Tte = *T. tahoe*, Tms = *T. monikensis*, Tdi = *T. douglasi*, Tsi = *T. shepardi*, and Tge = *T. genevievae*.

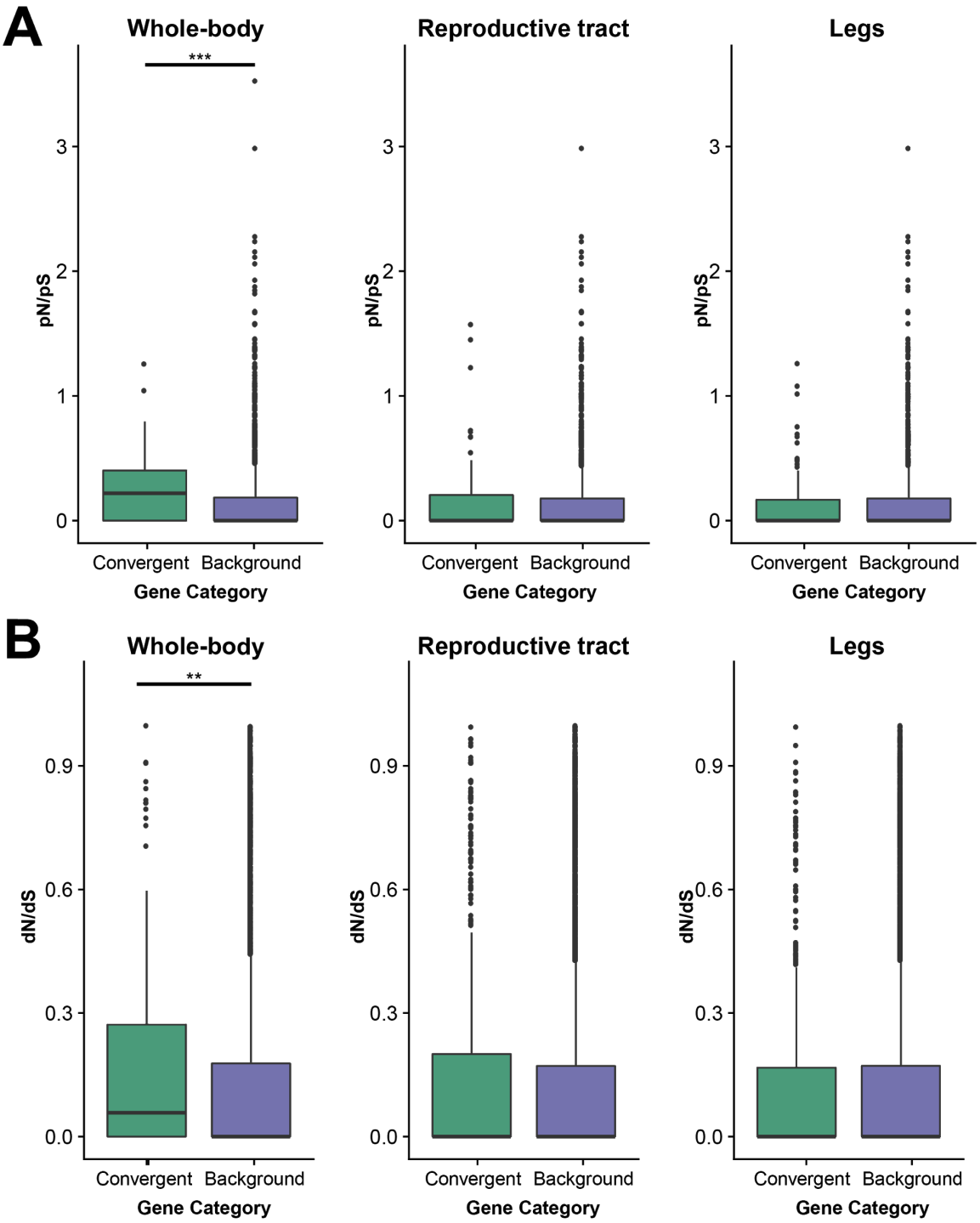


Figure 3 | pN/pS ratios (A) and dN/dS ratios (B) for convergently expressed genes versus all other genes expressed in that tissue for whole-bodies, reproductive tracts and legs. Significance is indicated by asterisks (** < 0.01, *** < 0.001).

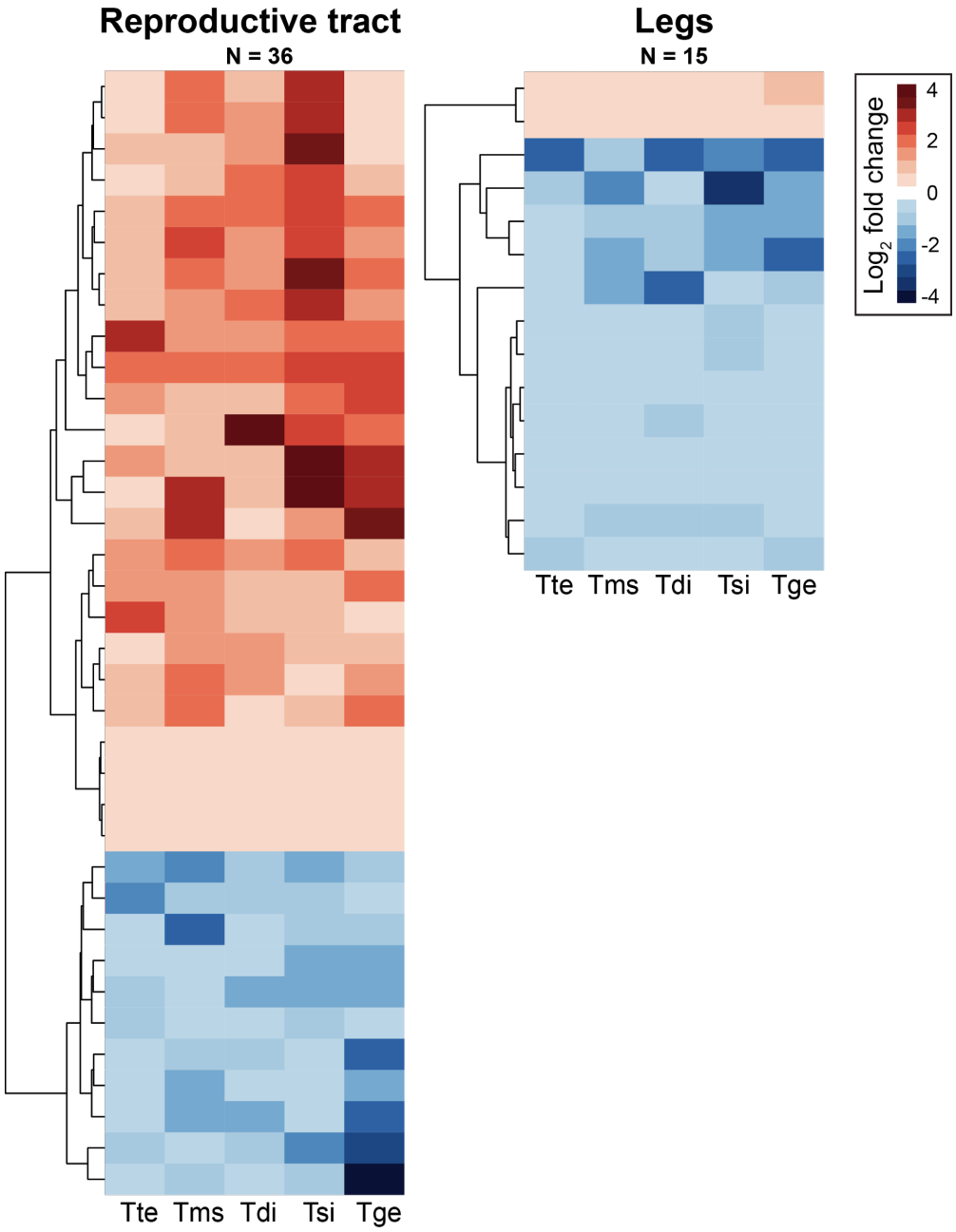


Figure 4 | Heatmaps of candidate convergent gene expression changes between sexual and asexual females for the reproductive tract and legs. Species names are abbreviated as follows: Tte = *T. tahoe*, Tms = *T. monikensis*, Tdi = *T. douglasi*, Tsi = *T. shepardi*, and Tge = *T. genevievae*.

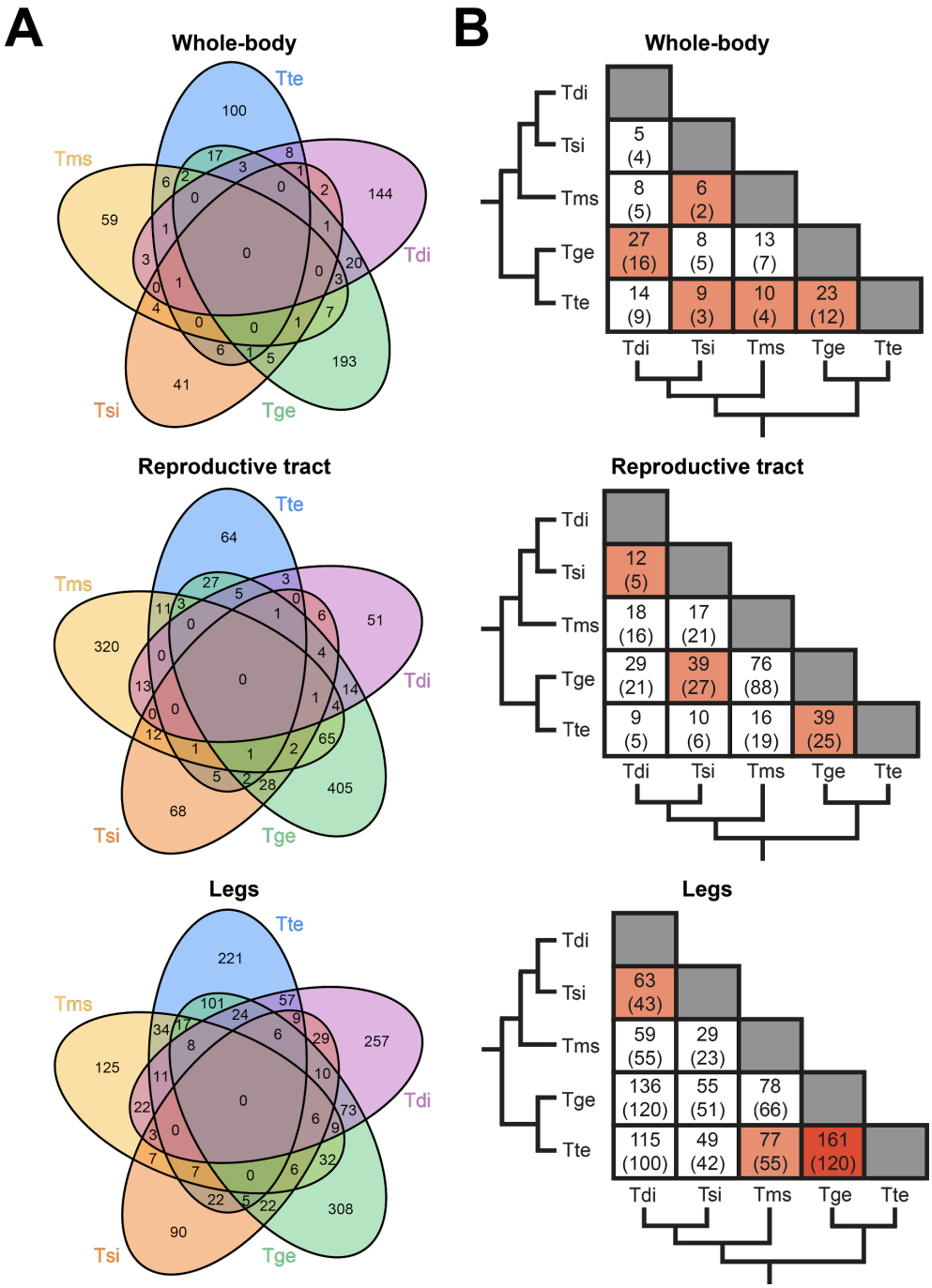
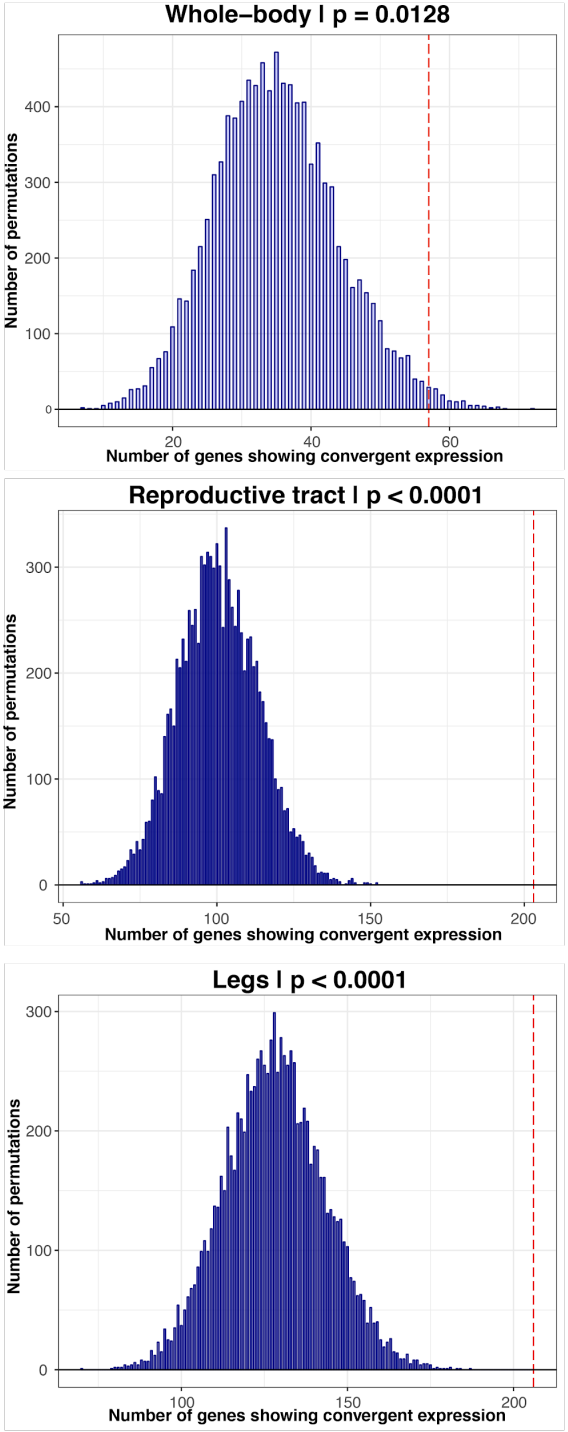


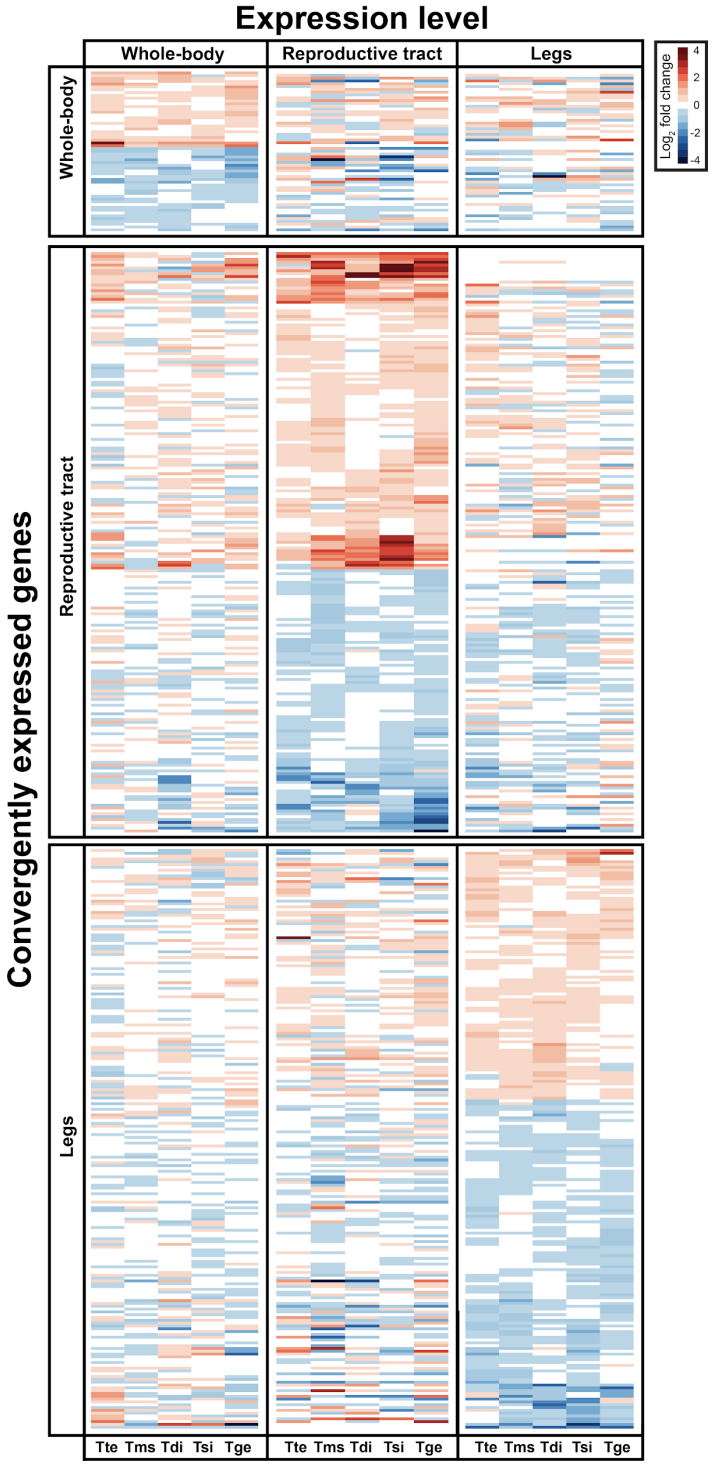
Figure 5 | A. Venn-diagrams showing the number differentially expressed (DE) genes between sexual and asexual females that are shared among species-pairs for whole-body, reproductive tract, and legs for 10 species orthologs (FDR < 0.05). B. Matrices showing pairwise overlap of DE genes between sex-asex species with the number of genes expected by chance given in parentheses. Colours represent a significantly greater overlap than expected by chance (red, FDR < 0.001, orange < 0.05). The phylogeny shows the relationships between asexual species (from [18]). Species-pair names are abbreviated as follows: Tte = *T. tahoe*, Tms = *T. monikensis*, Tdi = *T. douglasi*, Tsi = *T. shepardi*, and Tge = *T. genevievae*.

Supplementary Figures



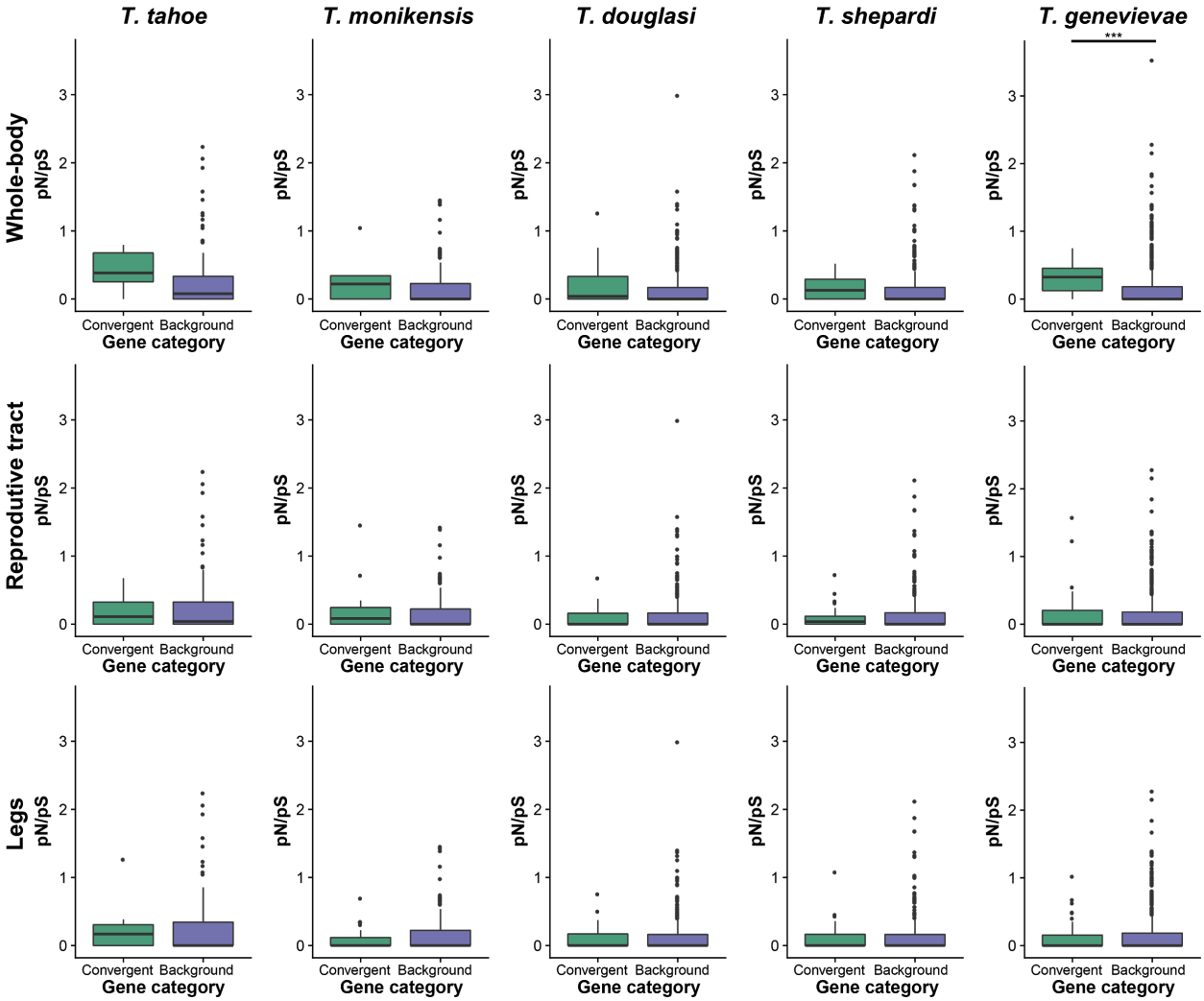
Supplementary Figure 1 | Number of genes expected to show a convergent expression pattern by chance (assessed by assigning reproductive mode randomly within species pairs for each gene for 10,000 permuted datasets). The observed number of convergent genes is indicated by

the red dashed line. P-values refer to the probability of observing a number of convergent genes greater-than or equal-to the observed value.

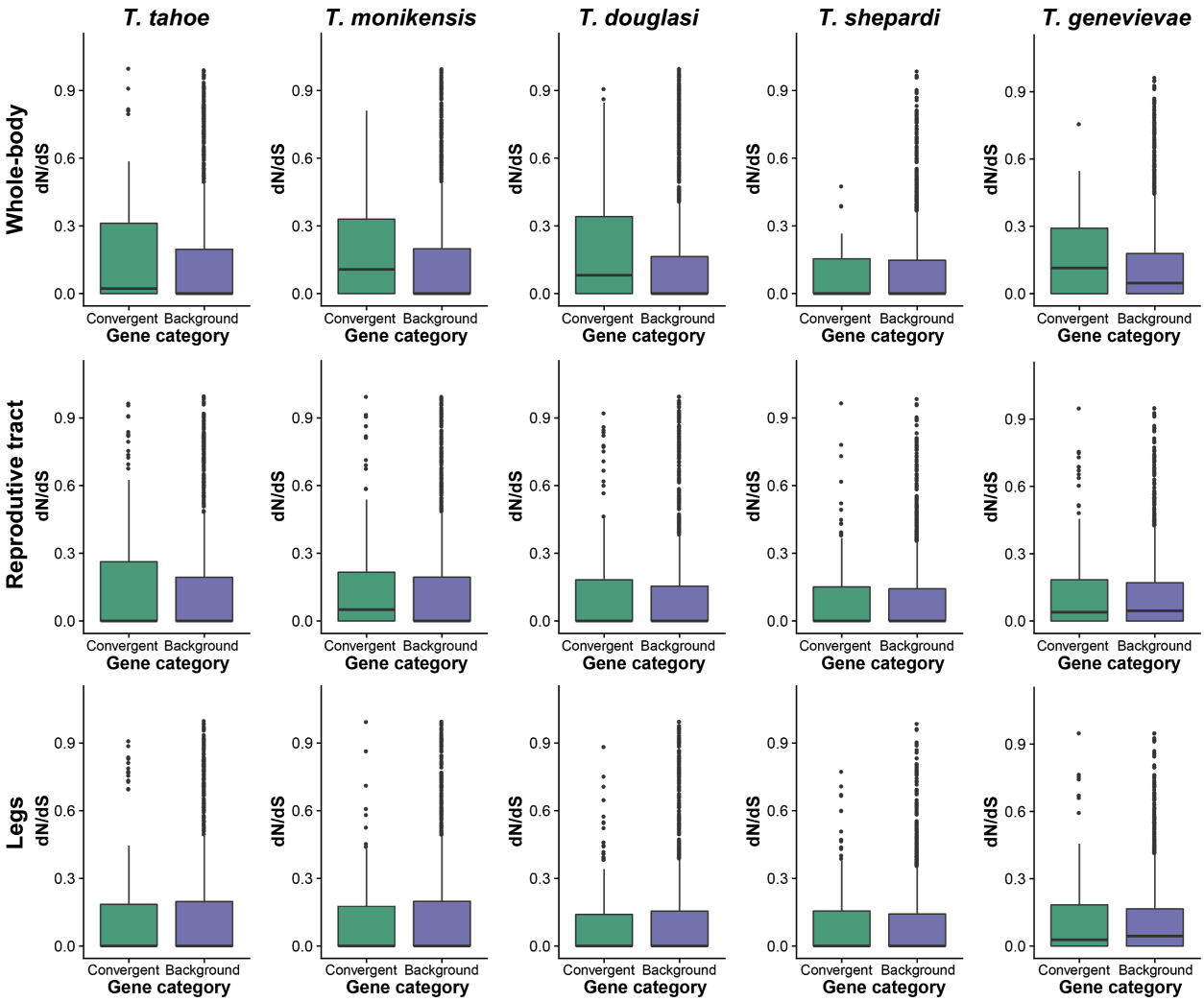


Supplementary Figure 2 | Convergently changing genes are largely tissue-specific. Heatmaps of genes showing convergent gene expression changes between sexual and asexual females for whole-bodies, reproductive tract, and legs, including their expression in other tissues.

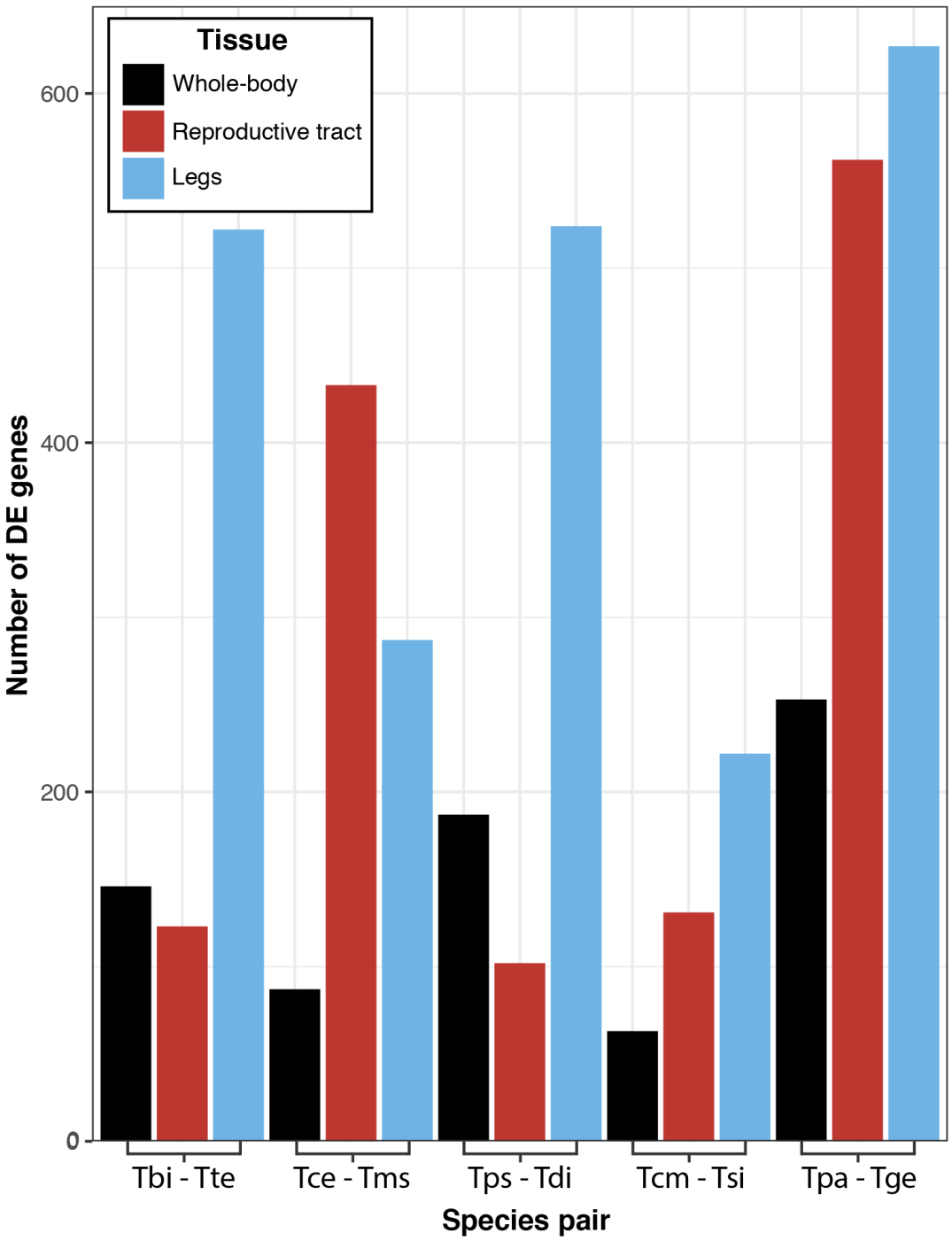
Species names are abbreviated as follows: Tte = *T. tahoe*, Tms = *T. monikensis*, Tdi = *T. douglasi*, Tsi = *T. shepardi*, and Tge = *T. genevievae*.



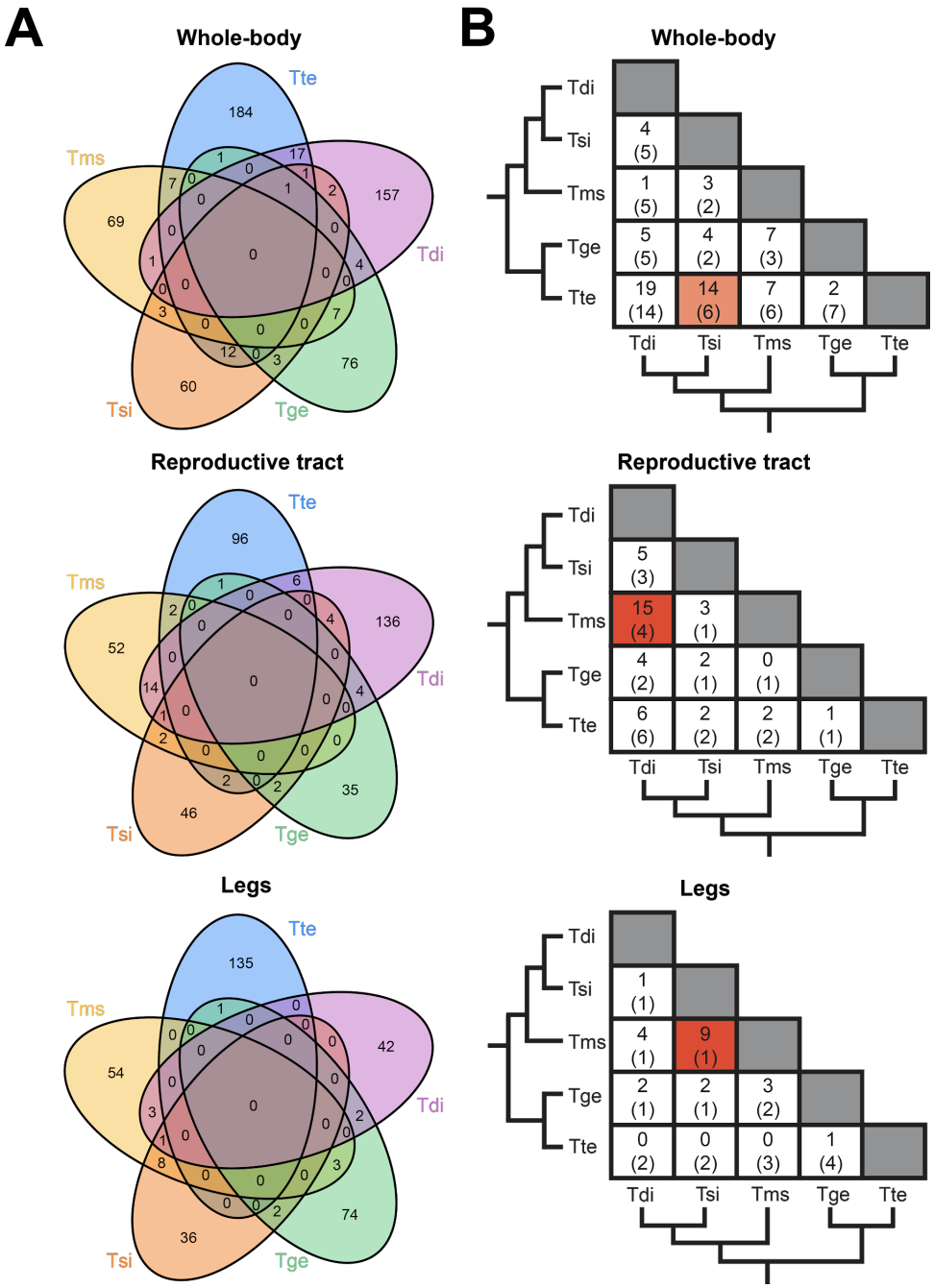
Supplementary Figure 3 | pN/pS ratios for convergently expressed genes versus all other genes expressed in that tissue for whole-bodies, reproductive tracts and legs for each asexual species. Significance is indicated by asterisks (*** < 0.001) from a Wilcoxon test.



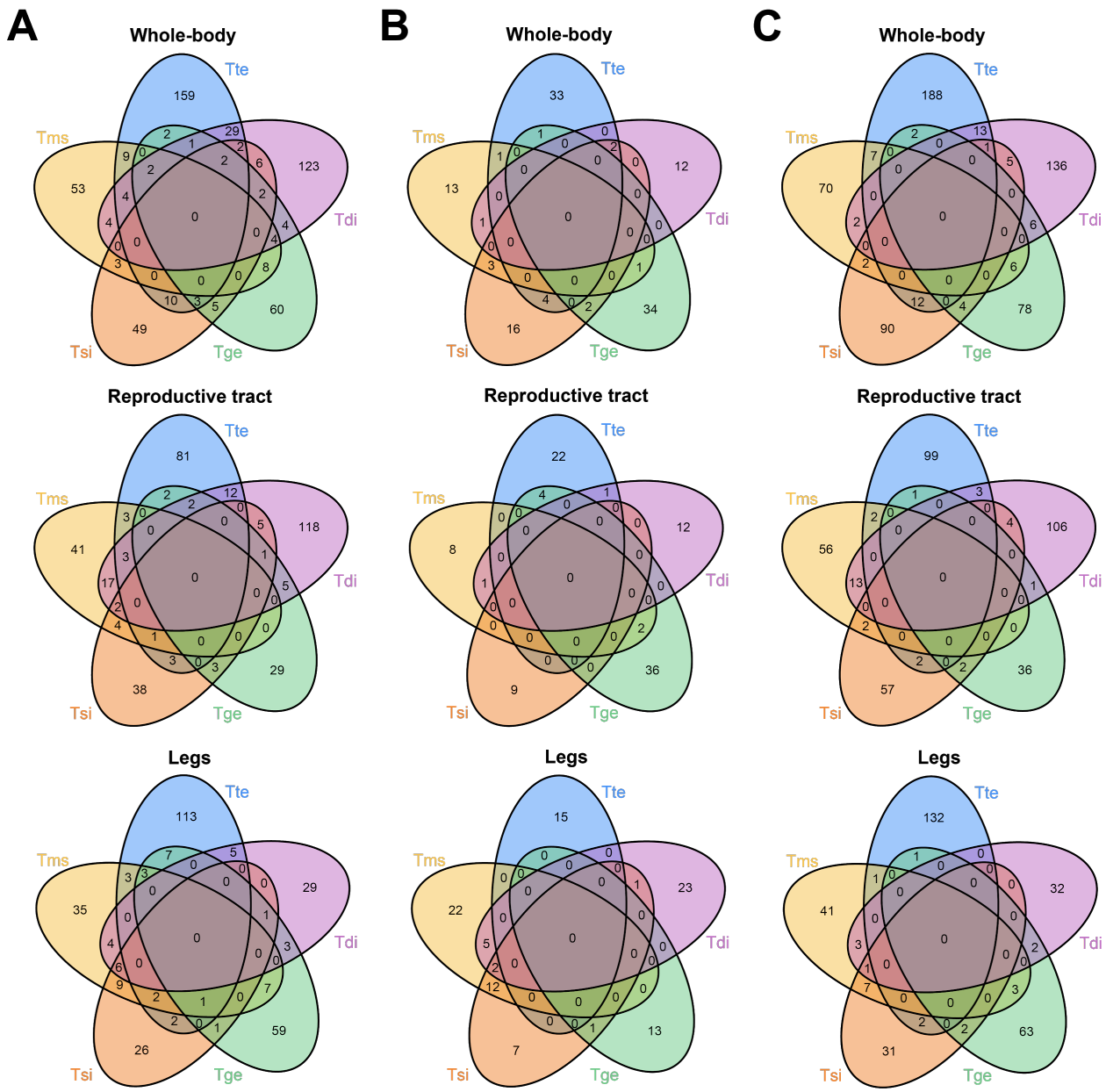
Supplementary Figure 4 | dN/dS ratios for convergently expressed genes versus all other genes expressed in that tissue for whole-bodies, reproductive tract and legs for each asexual species. No comparisons were significantly different (Wilcoxon test, $p > 0.05$).



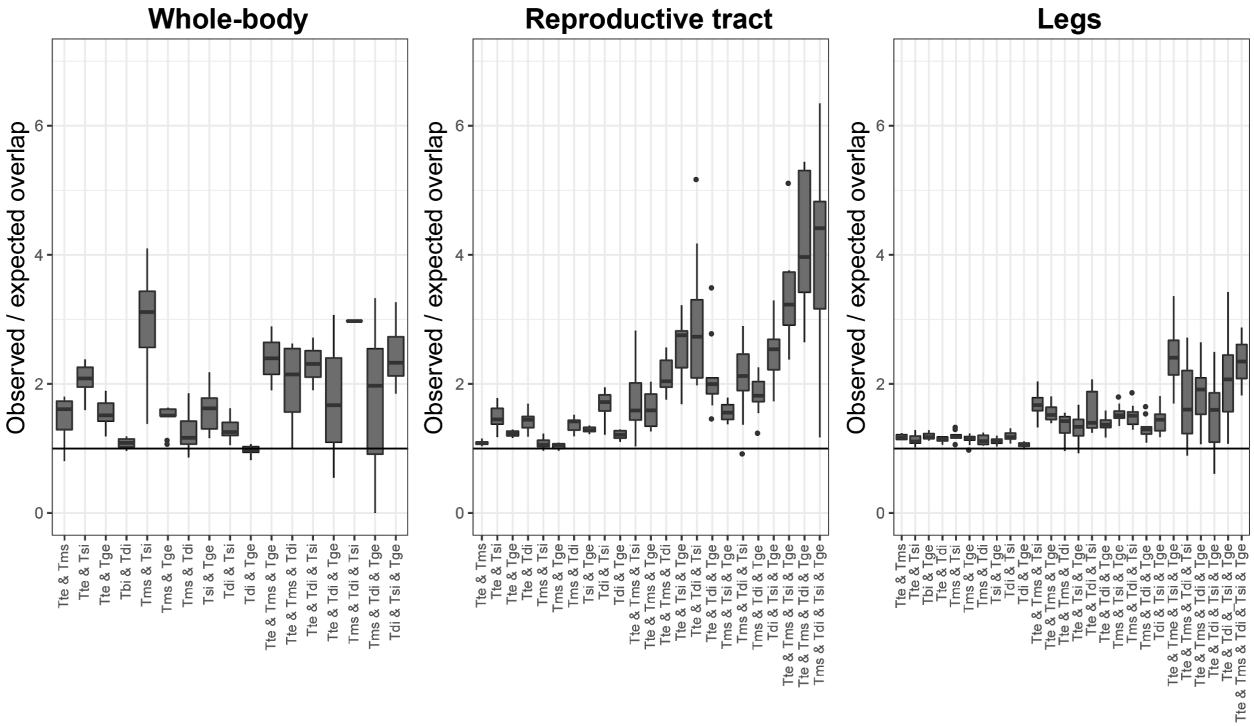
Supplementary Figure 5 | Number of DE genes (FDR < 0.05) between sexual and asexual females for each species-pair for whole-body, reproductive tract, and legs for the 10 species orthologs. Species names are abbreviated as follows: Tbi = *T. bartmani*, Tce = *T. cristinae*, Tps = *T. poppensis*, Tcm = *T. californicum*, Tpa = *T. podura*, Tte = *T. tahoe*, Tms = *T. monikensis*, Tdi = *T. douglasi*, Tsi = *T. shepardi*, and Tge = *T. genevievae*.



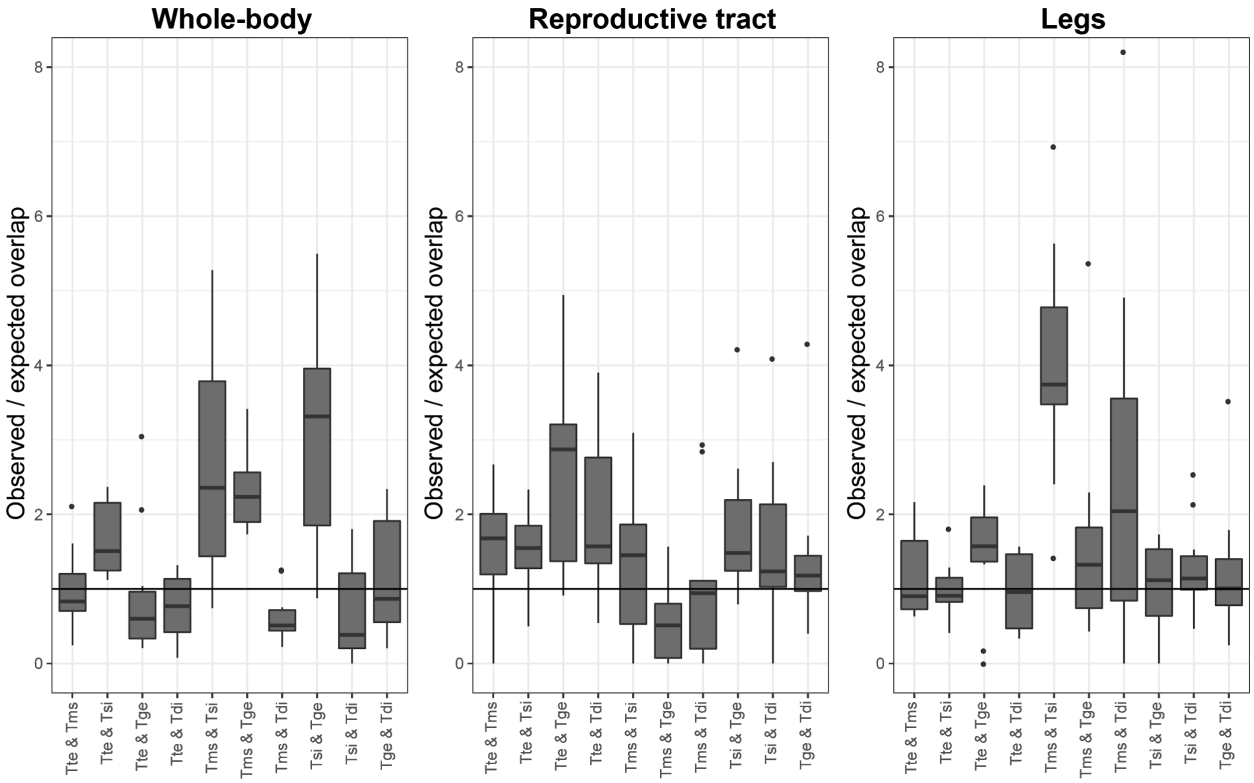
Supplementary Figure 6 | A. Venn-diagrams showing the number of enriched GO-terms between sexual and asexual females that are shared among species-pairs for whole-body, reproductive tracts, and legs for 10 species orthologs (FDR < 0.05). B. Matrices showing pairwise overlap of enriched GO-terms between sexual and asexual females with the number of GO terms expected by chance given in parentheses. Colours represent a significantly greater overlap than expected by chance (red, FDR < 0.001, orange < 0.05). The phylogeny shows the relationships between asexual species (from Schwander et al [18]). Species names are abbreviated as follows: Tte = *T. tahoe*, Tms = *T. monikensis*, Tdi = *T. douglasi*, Tsi = *T. shepardi*, and Tge = *T. genevievae*.



Supplementary Figure 7 | Venn-diagrams showing the number of enriched GO-terms ($p < 0.05$) for differences between sexual and asexual females that are shared among species-pairs in the whole-body, reproductive tract, and legs when: **A** GO terms were first clustered together based on parent or child terms, **B** only NCBI's nr annotation was used, **C** only the *T. bartmani* annotation. Species names are abbreviated as follows: Tte = *T. tahoe*, Tms = *T. monikensis*, Tdi = *T. douglasi*, Tsi = *T. shepardi*, and Tge = *T. genevievae*.



Supplementary Figure 8 | Ratio of observed to expected amount of overlap for genes when reads were mapped to the whole transcriptome for each species. Note boxes are only shown when the expected overlap was >= 1. Species names are abbreviated as follows: Tte = *T. tahoe*, Tms = *T. monikensis*, Tdi = *T. douglasi*, Tsi = *T. shepardi*, and Tge = *T. genevievae*.



Supplementary Figure 9 | Ratio of observed to expected amount of overlap of enriched GO-terms when reads were mapped to the whole transcriptome for each species. Note boxes are only shown when the expected overlap was >= 1. Species names are abbreviated as follows: Tte = *T. tahoe*, Tms = *T. monikensis*, Tdi = *T. douglasi*, Tsi = *T. shepardi*, and Tge = *T. genevievae*.