JOURNAL OF Evolutionary Biology

doi:10.1111/j.1420-9101.2010.02000.x

Divergent patterns of diversification in courtship and genitalic characters of *Timema* walking-sticks

D. ARBUTHNOTT*, M. G. ELLIOT*, M. A. McPEEK† & B. J. CRESPI*

*Department of Biological Sciences, Simon Fraser University, Burnaby, BC, Canada
†Department of Biological Sciences, Dartmouth College, Hanover, NH, USA

Keywords:

courtship evolution;
genitalia;
phylogenetics;
reproductive isolation;
speciation.

Abstract

Understanding the patterns of diversification in sexual traits and the selection underlying such diversification represents a major unresolved question in evolutionary biology. We examined the phylogenetic diversification for courtship and external genitalic characters across ten species of *Timema* walking-sticks, to infer the tempos and modes of character change in these sexual traits and to draw inferences regarding the selective pressures underlying speciation and diversification in this clade. Rates of inferred change in male courtship behaviours were proportional to speciation events, but male external genitalic structures showed a pattern of continuous change across evolutionary time, with divergence proportional to branch lengths. These findings suggest that diversification of courtship behaviour is mediated by processes that occur in association with speciation, whereas diversification of genitalia occurs more or less continuously, most likely driven by forces of sexual selection.

Introduction

Divergence in sexual traits often contributes to the evolution of reproductive isolation (Butlin & Ritchie, 1994; Panhuis *et al.*, 2001; Seddon, 2005). Such divergence may be underlain by a number of mechanisms, such as sexual selection and species recognition, which will influence the patterns of trait diversification. Understanding the tempo and mode of sexual trait divergence can lead to useful insights on the selective forces underlying phenotypic change. One particularly useful distinction to make is between continuous and speciational change (Mooers *et al.*, 1999); if diversification is continuous, change in sexual characters accumulate through time between speciation events, suggesting that there is little association between trait divergence and speciation. Examples of selective forces that may lead to continuous change are drift and sexual selection (Andersson, 1994; Polihronakis, 2006). Aspects of sexual selection, such as the runaway process, may mediate sexual trait diversification via mechanisms whereby differing intraspecific mate preferences or traits affecting intrasexual competition confer a reproductive advantage on certain sexual characters, pushing character values in different directions among populations.

By contrast, if diversification is speciational, the magnitude of phenotypic divergence is proportional to the number of speciation events between groups, rather than time since divergence, which suggests that change is concentrated at, and associated with, the process of speciation. Traits subject to stabilizing selection between speciation events would show such a pattern, such as traits involved in species or population recognition (Eberhard, 1985; Arnqvist, 1997). In this scenario, selection in the context of species recognition may drive the evolution of sexual characters when there are costs to mating with individuals from divergent populations (e.g., Blows & Allan, 1998; McPeek *et al.*, 2008; see Andersson, 1994 for review), which may also select for increased levels of prezygotic isolation via reinforcement (Butlin, 1989; Coyne & Orr, 2004). However, following speciation, there may be stabilizing selection for species-specific traits involved in mate discrimination (McPeek *et al.*, 2008, 2009). The patterns of phenotypic change and selection underlying diversification can be inferred by analysing character data from extant taxa in a phylogenetic framework, and interpreting these data in

Correspondence: Devin Arbuthnott, Department of Biology, University of Ottawa, Gendron Hall, room 160, 30 Marie Curie, Ottawa, ON, Canada K1N 6N5.
Tel.: +1 613 562 5800 ext. 6837; fax: +1 613 562 5486;
e-mail: devin.arbuthnott@uottawa.ca

the context of information on microevolutionary processes and character function (e.g. Mooers *et al.*, 1999; Cardoso & Mota, 2008; McPeek *et al.*, 2008, 2009; Prager & Andersson, 2009).

McPeek *et al.* (2008, 2009) used a combination of morphometric and phylogenetic data to infer that the majority of change in the external genitalia of *Enallagma* damselflies occurs during speciation events, a pattern concordant with the demonstrated species recognition function of these structures in this genus. Additional studies of the tempos and modes of diversification in sexual and species-recognition traits, across clades that differ in their mating systems, should lead to novel insights regarding the sexual and ecological factors that drive microevolutionary and macroevolutionary change in traits related to male–female interactions. In this study, we analysed patterns of phylogenetic diversification in male courtship behavioural displays, male external genitalic morphology, and male and female body morphology in *Timema* walking-sticks, a genus of phytophagous insects that exhibits notable interspecific variation in male genital structures (Vickery, 1993) and stereotyped courtship behaviours (Arbuthnott & Crespi, 2009). Our main goal is to evaluate alternative hypotheses for the tempos and modes of diversification in different forms of sexual and nonsexual traits in this genus and to draw inferences regarding the selective pressures that have generated the observed patterns of variation within and between species.

Materials and methods

Study system

The walking-stick genus *Timema* comprises 20 described species, 15 of which are sexual and five asexual. These species are distributed primarily in California, but also in regions of Oregon, Arizona, Nevada and northern Mexico (Vickery, 1993; Law & Crespi, 2002). The patterns of male–female interactions and mating in *Timema* involve a sequence of (i) initial contact between a male and female, (ii) pairing, whereby the male climbs onto the dorsal surface of the female, (iii) male courtship behaviour, during which males rapidly vibrate (wave) their hind legs or middle and hind legs, then their antennae, (iv) copulation, which involves the male twisting the end of his abdomen beneath the female, using his tripartite, asymmetrical clasping external genitalia to engage, hold and manipulate the female during insertion of the aedaegus, and (v) long-term postcopulatory mate guarding, whereby the male rides on the female's dorsal surface, not in copula, for up to 5 days (Arbuthnott & Crespi, 2009). The order of *Timema*'s behavioural courtship is consistent across all examined species. Our data on courtship is limited to male premating behaviours and does not include other possible components, such as chemical signals. Females are capable of resisting male copulation attempts by moving their abdomens away from male genitalia.

Sexual behaviour is integral to the evolution of reproductive isolation in *Timema*, given that premating behavioural isolation forms a major reproductive barrier between ecologically isolated populations within species (Nosil *et al.*, 2002; Nosil, 2004) and between species (Arbuthnott & Crespi, 2009). For example, *T. cristinae* exhibits a decreased incidence of courtship between individuals from populations on different host plants, whereas isolation between species (e.g. *T. cristinae*, *T. chumash* and *T. podura*) manifests as decreased incidence of pairing (the stage preceding courtship) between interspecific males and females (Arbuthnott & Crespi, 2009). Species recognition and premating isolation between ecologically divergent populations in *Timema* appear to involve chemical signals, given that mate discrimination occurs after antennal contact, but before pairing or courtship, although the specific cues involved in species recognition are currently unknown. Because isolation seems to occur before courtship is initiated, in both within- and between-species mate discrimination, observed interspecific differences in courtship (leg and antenna waving frequencies) apparently do not influence current reproductive isolation at either intermediate or complete levels of isolation.

Timema possess complex, asymmetric external genitalia comprised of three structures used to clasp and manoeuver females (Huber *et al.*, 2007). Traits associated with sperm transfer and use may be important in the mating system and evolution of reproductive isolation in *Timema*, as *T. cristinae* females lay fewer eggs if they have mated with a male from a different host plant (Nosil & Crespi, 2006).

Courtship observations

Timema were collected from March to June 2007 and 2008, and male courtship displays were recorded in the laboratory. Observation and recording procedures are described in detail by Arbuthnott & Crespi (2009).

Variation within species

To assess the level of intraspecific courtship variation, which can provide information on processes of intraspecific trait diversification, we quantified the courtship behaviour of multiple males (n = 4–9, except for *T. cristinae*) in each of ten species. One species (*T. cristinae*, 27 observations) was better represented in our behavioural data because it was the focus of another study (Arbuthnott & Crespi, 2009), whereas all other species were limited by population densities and ease of sampling. While our sample sizes of courtship observations are limited, these were adequate to detect species-level differences in courtship characters in a previous study (Arbuthnott & Crespi, 2009). We also collected samples from five populations of *T. cristinae* on different host

plants (three from *Adenostoma* and two from *Ceanothus* populations); we tested for differences among these populations by performing an analysis of variance on leg and antenna waving frequencies and used a Tukey's *post-hoc* test to determine the significance of each pairwise population difference. To assess the degree of intraspecific genitalic variation, we analysed genitalia characters of individuals from four separate *T. cristinae* populations.

Species specificity of courtship

To evaluate the species specificity of courtship, we performed a discriminant function analysis using four courtship characters: leg waving frequency, antenna waving frequency, per cent of active courtship spent on leg waving and the number of legs used during leg waving. This analysis allows quantification of the degree to which species can be separated in multivariate space, using linear combinations of traits involved in courtship.

Genitalic morphology

Timema external genitalia comprise two asymmetric cerci (claspers) and an intradextral process. The two cerci are used for clasping the female's abdomen during copulation, whereas the intradextral process (on the sinistral side of the right cercus) aids in the opening of the female subgenital plate, which in turn allows the male to insert his aedaegus (internal genitalia). We generated and analysed three-dimensional representations of genitalia using computer tomography (CT) for ten *Timema* species, following the protocols described in McPeek *et al.* (2008, see also Shen *et al.*, 2009). Two to eight males from each species were scanned. Digital image stacks for each were converted into three-dimensional representations using the programme AMIRA version 5.1 (Mercury Computer Systems, Chelmsford, MA, USA). For analysis, all structures were standardized to a centroid size of 1.0. Models were reconstructed as a triangular mesh with 20 000 data points and analysed using spherical harmonics (Shen & Makedon, 2006). Spherical harmonics analysis is an extension of classical Fourier analysis and reduces the complex, three-dimensional shape of the genitalic structures to a high dimensional set of spherical harmonics coefficients (Shen *et al.*, 2009). We used principal components (PC) analyses to reduce this high dimensional representation of the shapes to a small number of axes. All analyses were carried out for each of the three genitalic structures separately.

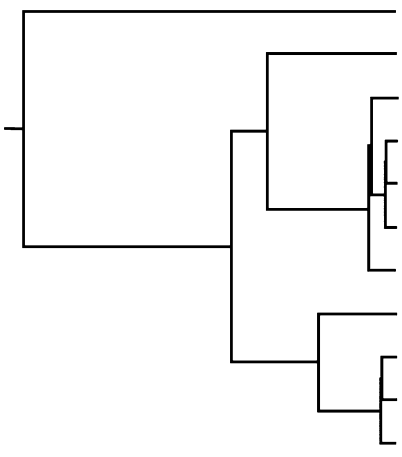
Morphological body characters

To compare the diversification patterns of courtship behaviour and genitalic morphology with the diversification patterns of morphological traits not involved in courtship or mating, we also analysed a set of linear body traits, including length of the right hind and middle tibia, head width (eye-to-eye distance), thorax width (width of the widest portion of the thorax) and body length (anterior-most point of the head to the posterior-most point of the last abdominal segment). These measurements were carried out for males and females and for 2–16 individuals per species depending on the availability of samples.

Phylogenetics of diversification

Phylogenetic trees describing the relationships between the ten *Timema* species for which courtship, morphological and genitalic data are available were constructed from mitochondrial COI sequences 789 bp in length. We used Mantophasmatodea COI (http://www.ncbi.nlm.nih.gov/nuccore/84488734) from Cameron *et al.* (2006) as an outgroup because of the availability of appropriate COI data. The best nucleotide substitution model, general time reversible model with gamma distributed site rate variation and a proportion of invariable sites, was selected using hierarchical likelihood ratio tests (Nylander, 2004). The model was fit to the sequence data using MrBayes version 3.1 (Ronquist & Huelsenbeck, 2003). Markov chains were run for five million generations after attaining stationarity, as judged from the lack of a directional trend in likelihood over time. Chains were sampled every 1000 steps. The harmonic mean log likelihood of the stationary chain was calculated first with free branch lengths and again under the constraint of a global clock; a likelihood ratio test was unable to reject the hypothesis that the sequences have evolved under a clock-like process (Clock-like mean log likelihood = −3322.4, free mean log likelihood = −3316.9, $\chi^2_9 = 11.102$, $P = 0.26$), such that branch lengths in this phylogeny can be interpreted as proportional to time. The relationship between mtDNA divergence and time has been investigated and calibrated previously through the combination of phylogeographic and genetic data (e.g. Papadopoulou *et al.*, in press); clock-like sequence evolution always involves nucleotide change proportional to branch length and time, rather than numbers of speciation events.

The inferred phylogeny (Fig. 1) is fully compatible with previous studies of *Timema* (Law & Crespi, 2002), in that *Timema* is split into a northern clade (*T. cristinae*, *T. landelsensis*, *T. knulli*, *T. poppensis*, *T. petita* and *T. californicum*) and a southern clade (*T. chumash*, *T. bartmani*, *T. boharti* and *T. podura*). Maximum parsimony (MP) and maximum likelihood (ML) analyses yielded the same topology for all well-supported nodes (nodes with Bayesian *a posteriori* values of 1, which exhibited ML and MP bootstrap values over 90%, for 200 replicates). Effects of uncertainty in tree topology on character change results were assessed by repeating analyses on the patterns of change of all courtship traits, and the PC1 scores of all genitalic structures using nine alternative



	Leg waving frequency	Antenna waving frequency	Left cercus PC1	Right cercus PC1	Intradextral process PC1
S. parasisense					
T. cristinae	3.4	1.5	0.005	0.002	–0.003
T. landelsensis	3.7	1.6	0.006	0.0008	0.001
T. petita	4.4	1.6	0.004	0.002	0.0009
T. poppensis	2.5	1.1	0.003	0.003	–0.002
T. knulli	4.6	1.0	0.004	0.001	0.0005
T. californicum	4.2	1.6	0.005	0.003	0.00003
T. chumash	3.9	2.0	–0.010	0.002	–0.002
T. podura	1.3	2.2	–0.009	–0.007	0.007
T. bartmani	3.9	2.2	–0.008	–0.004	0.004
T. boharti	5.3	1.7	–0.009	–0.002	0.003

Fig. 1 Clock-constrained tree for the ten examined *Timema* species based on Bayesian inference. All of the Bayesian a posteriori values are 100%, except for the branch basal to (*T. petita* + *T. poppensis*) and the branch basal to (*T. knulli* + *T. landelsensis* + *T. petita* + *T. poppensis*), which show values of 50% and are thus unresolved. Maximum parsimony (MP) and maximum likelihood (ML) bootstrap values (200 replicates) were all over 90% except for the branches basal to the clades (*T. knulli* + *T. poppensis*) and (*T. podura* + *T. bartmani*), which showed bootstraps under 60%; the clade [*T. californicum* ((*T. knulli* + *T. petita* + *T. poppensis*), *T. landelsensis*)] was similarly unresolved. Species averages of key courtship and genitalic characters are given alongside species.

trees, representing all permutations of branches exhibiting bootstrap or *a posteriori* support under 90%.

Multiple analytic methods were deployed to analyse the macroevolution of courtship traits, genitalic traits and body-morphology traits, to increase the robustness of hypothesis-testing. We used the programme CoMET (Lee *et al.*, 2006) whereby Akaike Information Criterion (AIC) values are calculated to determine which model of evolutionary change best characterizes the diversification of courtship, genitalia and body-morphology traits. CoMET compares nine models of change, represented by a 3 × 3 matrix of possible model combinations, where the best model of the first group (which we refer to as phylogenetic-signal models) is paired with the best model of the second group (which we refer to as tempo of change models). The first group of model classifications represents the phylogenetic pattern of phenotypic change, as pure-phylogenetic, nonphylogenetic or punctuational. Under the pure-phylogenetic model, phenotypic change occurs along all branches, and phenotype therefore shows positive phylogenetic autocorrelation. The nonphylogenetic model disregards branching points and essentially assumes a star phylogeny, such that phenotype shows no phylogenetic autocorrelation. The punctuational model assumes that at each bifurcation in the tree, one daughter branch retains the ancestral phenotype (zero change) and the other daughter branch changes.

The second group of model classifications represents the tempo of phenotypic change along branches as three possible models: distance, equal and free. In the distance model, change is proportional to genetic distance (branch length). The equal model assumes that all branch lengths are equal, and phenotypic change is therefore independent of time since divergence. In the free model, branch lengths are any nonnegative value calculated using ML on trait values. The free model thus calculates phenotypic branch lengths, where lengths are proportional to the amount of phenotypic change, rather than as a function of genetic divergence. Under this model, phenotypic change is not proportional to time since divergence, and lineages diversify under different rates of phenotypic change.

We also tested for phylogenetic autocorrelation of each courtship, morphological and genitalic trait using Phylogenetic Independence (Abouheif, 1999; see Pavoine *et al.*, 2008 for statistical validation of Phylogenetic Independence), shuffling nodes of each topology 1000 times and shuffling character data 1000 times. Phylogenetic independence outputs a correlation coefficient and the significance of this test statistic, and therefore gives a continuous measure of phylogenetic autocorrelation. CoMET, in contrast, estimates which discrete model of phylogenetic autocorrelation best represents the data.

To further assess the tempo of character diversification, we estimated Pagel's κ (Pagel, 1997), which involves raising branch lengths to a range of exponents (κ) and determining what value best characterizes phenotypic change along a phylogenetic tree for each trait. If branches are best described by an exponent of zero, then all branch lengths are equal to one, and diversification is

thus characterized by change only at speciation. If branches are best described by being raised to a power of one, then branch lengths are equal to genetic distances, and diversification is characterized as continuous change. If κ is greater than one, it suggests that trait diversification is continuous but progresses faster than genetic diversification. We used a ML search method to find the exponent that best characterizes phenotypic change along the *Timema* phylogeny, as described in McPeek *et al.* (2008). κ calculations determine what tempo of change best fits character divergence in a continuous way, by determining where in the continuum of speciational to continuous change a particular character lies, whereas CoMET assesses which end of this continuum best fits phenotypic data in a discrete way.

Results

Intraspecific variation in courtship

Some of the species examined exhibit substantial levels of intraspecific phenotypic variability in courtship traits (Fig. 2). *T. cristinae* showed significant interpopulation differences in antenna waving frequency (ANOVA; $F_{4,20} = 3.82$, $P = 0.018$), but not in leg waving frequency ($F_{4,18} = 0.53$, $P = 0.72$). In particular, one *Adenostoma* population and one *Cenothus* population differed from one another in antenna waving frequency ($P = 0.022$), and the differences between two *Adenostoma* populations approached significance ($P = 0.07$). These findings document the presence of among-population courtship differences within species of *Timema* that may be associated with speciation and macroevolutionary change in this genus.

Species specificity of courtship

Courtship characters showed a considerable degree of overlap among species (Fig. 2), such that these courtship traits are not highly species specific. Based on each individual's courtship characters, discriminant function analysis assigned individuals to their correct species 58.2% of the time across the genus as a whole. Courtship phenotypes (an individual's combination of all measured courtship characters) were very predictive of species identity in a few cases, as suggested by a high percentage of correct species assignments in the discriminant functions analysis (*T. cristinae*: 85.2%, *T. podura*: 77.8%, *T. poppensis*: 83.3%), but courtship phenotype displayed intermediate to low levels of species specificity in the remainder of the species (*T. bartmani*: 50.0%, *T. boharti*: 20%, *T. californicum*: 20%, *T. chumash*: 14.3%, *T. knulli*: 60%, *T. landelsensis*: 33.3%, *T. petita*: 20%). Given that we would only expect 10% correct assignment by chance with ten species if courtship possessed no species specificity, 58% correct assignment overall suggests an intermediate level of species specificity in courtship signals. We also note that our small sample sizes for courtship observations (< 10 for most species) could limit our power to detect species-level differences. However, these data suggest that the courtship traits quantified here do not show high levels of species specificity for several species, which is concordant with observations that reproductive isolation among species of *Timema* occurs prior to pairing and courtship (Arbuthnott & Crespi, 2009).

Genitalic morphology

The first two PCs for the *Timema* left cercus accounted for 68.5% of phenotypic variance among species. PC1 (57.1% of total variation explained) represents the relative thickness and the curvature of the clasper: claspers with more negative values were relatively straight and with a larger medial process near the middle of the structure, and claspers with positive values were more curved and had smaller medial processes at the base and top (Fig. 3b). By contrast, PC2 (11.4% of total

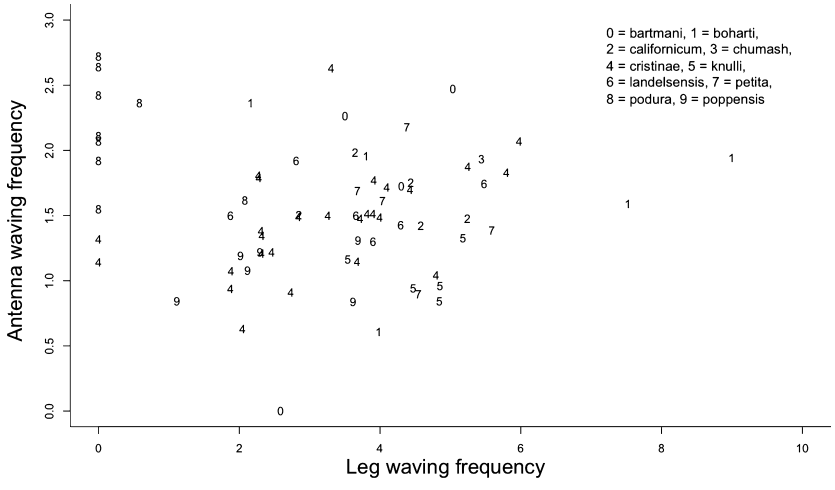


Fig. 2 Diversity of courtship among *Timema* species. Leg and antenna waving frequencies display species-level differences. Members of different species are denoted using numbers.

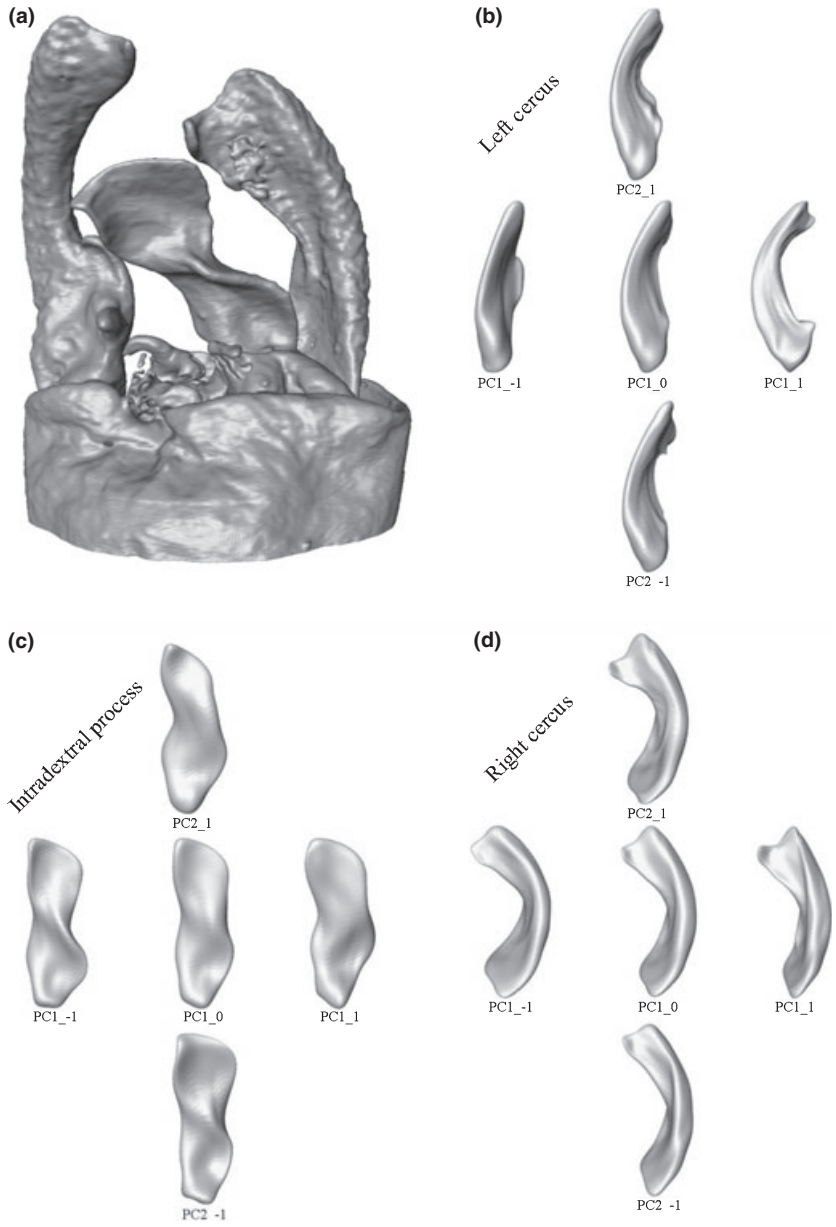


Fig. 3 Three-dimensional representations of *Timema* external genitalic structures. (a) Ventral view of complete genitalia of one *T. cristinae* individual, showing (left to right) the left cercus, intradextral process and right cercus. (b–d) Variation in genitalic shape along principal component axes for *Timema* left cercus, intradextral process and right cercus.

variation explained) represents the degree of twisting of the clasper (Fig. 3b). All other PCs for the left cercus explained < 7% of total variation each, and so we consider only the first two PCs.

For the right cercus, the first two PCs account for 59.1% of phenotypic variance. Increasing values for PC1 (42.6% of total variation) had more forked tips, whereas increasing values for PC2 (16.5% of total variation) had increasing relative thickness and a broader terminal tip (Fig. 3d). We considered the first two PCs for the right cercus because all other PCs explained < 7% of total variation each.

For the intradextral process, the first two PCs accounted for 52.7% of phenotypic variance. PC1 (31.9% of total variation) quantified twisting of the process, whereas PC2 (20.7% of total variation) quantified the curvature (Fig. 3c).

Genitalic shape of all structures clusters well by species, although intraspecific variation was large and many species overlapped in distribution (Fig. 4). PC1 of the left

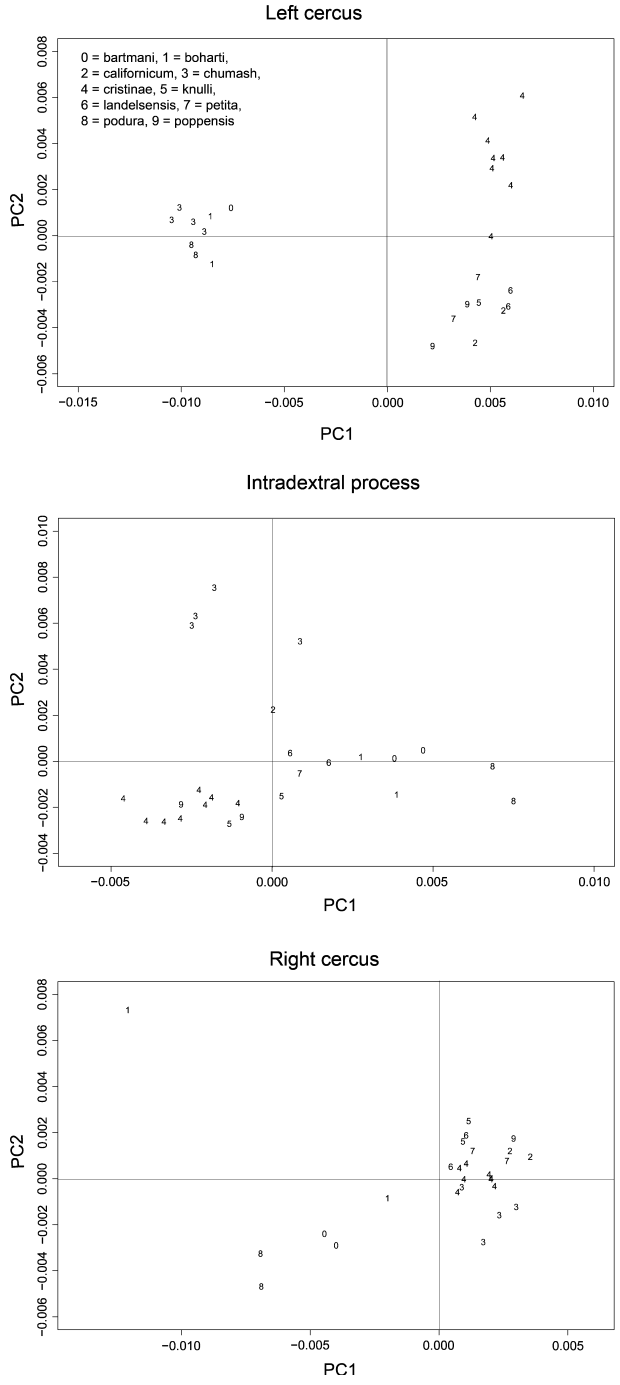


Fig. 4 Principal component representations of left cercus, intradextral process and right cercus morphology for all measured *Timema* individuals. Genitalic shape clusters by species and clade for *Timema* left cercus, intradextral process and right cercus morphology (southern clade represents *T. bartmani*, *T. boharti*, *T. chumash* and *T. podura*, northern clade represents all other species). Intraspecific variation can be seen in *T. cristinae*. Numbers denote species. Principal component axes correspond to those in Fig. 2. When the outlier (*boharti* right cercus data) is removed, all CoMET and κ results are qualitatively unchanged.

cercus separated the southern clade of *Timema* (*T. podura*, *T. chumash*, *T. bartmani* and *T. boharti*) from the northern clade (all other *Timema* species; Fig. 4). The northern clade clustered notably for right cercus morphology, whereas the southern clade was more varied (Fig. 4). The northern clade showed some clustering in intradextral process morphology, although the separation from southern clade species was not as extreme as the other two structures (Fig. 4).

Phylogenetics of trait diversification

The average and standard deviation of all courtship, morphological and genitalic traits for each species are given in Table 1. For each courtship characteristic, CoMET analyses provided evidence for the same model of tempo, equal change (change not proportional to time since divergence), that fits each trait better than any other model (AIC difference > 3; Table 2). Pagel's κ values for antenna waving frequency and per cent of active courtship spent leg waving were approximately 0, again suggesting that courtship divergence was not proportional to branch lengths, whereas the κ value for leg waving frequency was 0.3, which suggests an association between branch length and divergence, although most change seems to occur in association with speciation events. By contrast, the model representing phylogenetic signal differed between traits: antenna waving frequency was best described by a model of pure-phylogenetic change (positive phylogenetic autocorrelation), per cent of active courtship spent on leg waving was best described by a model of punctuational change and leg waving frequency was best described by a model of nonphylogenetic change (no phylogenetic signal). From autocorrelation analysis using Phylogenetic Independence (Abouheif, 1999), antenna waving frequency showed positive phylogenetic autocorrelation, whereas the per cent of active courtship spent on leg waving showed significant negative autocorrelation (characters of more closely related species are more different than characters of distantly related species) and leg waving frequency showed a negative phylogenetic autocorrelation that approached significance.

For all genitalic structures, CoMET analyses provided evidence for the same model of tempo, with change proportional to genetic distance that fit each trait better than any other model (AIC difference > 2; Table 2). Similarly, a Pagel's κ value of approximately 1 best fits phenotypic change in intradextral process shape, and a κ value of approximately 2 best fits change in left cercus shape. These values correspond to a model of genitalic character change as a more or less continuous process across macroevolutionary time. A κ value of 0.3 best fits change in right cercus shape, which suggests that there may have been accelerated change during speciation events, but that phenotypic change also accumulated along branches in proportion to time. A

Table 1 Average courtship and morphological values for ten *Timema* species.

Species	*bartmani*	*boharti*	*californicum*	*chumash*	*cristinae*	*knulli*	*landelsensis*	*petita*	*podura*	*poppensis*
Hosts*	F	E, F	G	D, E	D, E	B	C	D	D, E	A
Number of courtship observations	4	5	5	7	27	5	6	5	9	6
Legs used during leg waving	2 pairs	2 pairs	2 pairs	1 pair	1 pair	2 pairs	2 pairs	2 pairs	1 pair	2 pairs
Average leg waving frequency†	3. 858 (1.054)	5.292 (2.85)	4.15 (0.92)	3.861 (1.19)	3.412 (1.2)	4.578 (0.63)	3.67 (1.24)	4.447 (0.72)	1.333 (1.06)	2.476 (1.0)
Per cent active courtship spent leg waving‡	80 (14)	27 (14)	39.7 (19)	28.2 (18)	29.6 (18)	46.6 (16)	24.2 (12)	20.2 (10)	7.4 (15)	20.6 (10)
Average antenna waving frequency†	2.16 (0.38)	1.693 (0.66)	1.629 (0.24)	2.032 (0.3)	1.465 (0.41)	1.048 (0.19)	1.566 (0.23)	1.554 (0.47)	2.168 (0.44)	1.08 (0.2)
Number of females measured	4	4	4	5	8	3	4	1	5	5
Number of males measured	0	2	4	5	8	2	4	1	5	3
Female hind tibia length	3.52 (0.38)		3.47 (0.08)	3.77 (0.2)	2.85 (0.5)	4.91 (0.06)	3.65 (0.13)	3.19	2.91 (0.13)	4.25 (0.15)
Male hind tibia length			3.11 (0.06)	3.01 (0.1)	2.55 (0.14)	4.0 (0.19)	3.22 (0.25)	2.52	2.5 (0.12)	3.58 (0.09)
Female middle tibia length	2.44 (0.17)	2.31 (0.19)	2.44 (0.04)	2.65 (0.06)	2.34 (0.11)	3.3 (0.12)	2.64 (0.25)	2.15	2.21 (0.18)	2.95 (0.19)
Male middle tibia length		2.08	2.37 (0.08)	2.27 (0.18)	1.96 (0.15)	2.85 (0.22)	2.3 (0.13)	1.79	1.98 (0.08)	2.46 (0.05)
Female head width	2.04 (0.22)	1.95 (0.09)	2.04 (0.05)	2.29 (0.12)	1.79 (0.12)	2.31 (0.03)	2.1 (0.09)	1.75	1.92 (0.08)	2.12 (0.1)
Male head width		1.79 (0.24)	1.86 (0.05)	2.01 (0.06)	1.56 (0.11)	1.94 (0.03)	1.77 (0.07)	1.4	1.7 (0.08)	1.88 (0.11)
Female thorax width	3.41 (0.32)	4.13 (0.34)	3.79 (0.1)	4.49 (0.29)	3.9 (0.37)	4.41 (0.27)	4.15 (0.31)	3.58	3.82 (0.18)	3.95 (0.36)
Male thorax width		2.86 (0.32)	2.77 (0.02)	2.86 (0.08)	2.69 (0.15)	2.94 (0.56)	2.76 (0.16)	2.4	2.75 (0.1)	2.76 (0.01)
Female body length	13.93 (1.31)	20.11 (0.13)	18.51 (1.79)	14.72 (1.62)	18.85 (1.69)	20.01 (2.02)	20.17 (0.63)	12.78	16.42 (2.27)	19.95 (2.23)
Male body length		16.23 (1.43)	13.05 (1.07)	9.79 (0.82)	12.09 (1.56)	16.09 (2.24)	14.69 (0.57)	11.14	11.73 (0.97)	16.29 (1.48)
Number of male genitalia measured	2	2	2	4	8	2	2	2	2	2
Left cercus PC1 score	−0.0076	−0.0086	0.005	−0.0097	0.0053	0.0044	0.0059	0.0038	−0.0094	0.003
Left cercus PC2 score	0.0012	−0.0002	−0.004	0.00067	0.0034	−0.0029	−0.0027	−0.0027	−0.00062	−0.0039
Left cercus PC3 score	0.0015	0.0013	−0.00057	−0.0021	−0.00022	−0.0015	−0.003	0.004	0.0016	0.0018
Right cercus PC1 score	−0.0042	−0.002	0.0032	0.002	0.0015	0.001	0.00076	0.002	−0.0069	0.0029
Right cercus PC2 score	−0.0027	0.00085	0.0011	−0.0015	0.00005	0.0021	0.0012	0.00099	−0.004	0.0017
Right cercus PC3 score	0.0003	−0.0003	0.00099	−0.0039	0.0013	0.00003	0.00087	0.0004	0.0016	0.00009
Intradextral process PC1 score	0.0042	0.0033	0.00003	−0.0015	−0.0028	0.00052	0.0012	0.00086	0.0072	−0.0019
Intradextral process PC2 score	0.0003	−0.0006	0.0023	0.0062	−0.002	−0.0021	0.00015	−0.00053	−0.00097	−0.0022
Intradextral process PC3 score	0.00031	0.0028	0.00004	−0.00036	0.0013	−0.004	0.00057	0.00066	−0.00078	−0.0036

Standard deviations are given in brackets.

*Hosts are noted as in Law & Crespi (2002): A, *Pseudotsuga menziesii*; B, *Sequoia sempervirens*; C, *Arctostaphylos* species (manzanita); D, *Ceanothus* species; E, *Adenostoma fasiculatum*; F, *Abies concolor*; G, *Quercus* species.

†Waving frequencies are measured as number of waves per second.

‡% active courtship is calculated as time spent on behaviour/(time spent on leg waving + time spent on antenna waving).

Table 2 Phylogenetic patterns of change for all *Timema* characters. All courtship and most body morphology meet predictions of speciational change, whereas all genitalic structures meet predictions of continuous change.

Trait	Phylogenetic auto-correlation*	Phylogenetic auto-correlation *P*-value	CoMET model phylogenetic signal†	CoMET model tempo of trait change†	κ‡
Courtship traits					
Leg waving frequency	−0.2226	0.069	Nonphylogenetic	Equal	0.378 (logL = −14)
Per cent active courtship spent leg waving	−0.3102	0.035	Punctuated	Equal	0 (−1.5)
Antenna waving frequency	0.5444	0.011	Pure-phylogenetic	Equal	0 (6.7)
Body morphology					
Female hind tibia length	0.025	0.317	Nonphylogenetic, pure-phylogenetic	Equal	
Male hind tibia length	−0.139	0.48	Nonphylogenetic, pure-phylogenetic	Equal	
Female middle tibia length	−0.0459	0.468	Nonphylogenetic, pure-phylogenetic	Equal	0 (0.15)
Male middle tibia length	−0.1112	0.321	Nonphylogenetic	Equal	
Female head width	−0.1356	0.217	Nonphylogenetic	Equal	0.602 (5.4)
Male head width	−0.1958	0.178	Nonphylogenetic, pure-phylogenetic	Equal	
Female thorax width	−0.0085	0.482	Nonphylogenetic, pure-phylogenetic	Equal	0.287 (−1.6)
Male thorax width	−0.1193	0.224	Nonphylogenetic	Equal	
Female body length	−0.1263	0.256	Punctuated, nonphylogenetic	Equal	0.152 (−19.5)
Male body length	−0.1146	0.342	Punctuated, nonphylogenetic	Equal	
Genitalic morphology					
Overall left cercus shape			Pure-phylogenetic	Distance	2.25 (47.2)
Left cercus PC1	0.7523	0.004	Pure-phylogenetic	Distance	
Left cercus PC2	0.3741	0.02	Pure-phylogenetic	Distance	
Left cercus PC3	0.1163	0.283	Nonphylogenetic	Equal, distance	
Overall right cercus shape			Pure-phylogenetic, nonphylogenetic	Distance	0.32 (45.2)
Right cercus PC1	0.5514	0.013	Pure-phylogenetic	Distance	
Right cercus PC2	0.6999	0.002	Nonphylogenetic	Equal, distance	
Right cercus PC3	0.1599	0.115	Nonphylogenetic	Distance	
Overall intradextral process shape			Pure-phylogenetic, nonphylogenetic	Distance	1.277 (46.2)
Intradextral process PC1	0.4226	0.039	Pure-phylogenetic	Distance	
Intradextral process PC2	0.1211	0.223	Nonphylogenetic, pure-phylogenetic	Distance	
Intradextral process PC3	0.0408	0.418	Nonphylogenetic	Equal	

*Phylogenetic autocorrelation calculated with Phylogenetic Independence (Abouheif, 1999).

†Best model of evolutionary change as determined by CoMET. Models presented have AIC scores of 2 or more difference from any other model, and AIC scores are within 2 where two separate models are presented. CoMET pairs models, but we separate phylogenetic signal and tempo of change results for ease of reading.

‡Pagel's κ that best describes each character (with κ log likelihood values in brackets). κ values approximating zero indicate speciational change (equal model in CoMET), whereas values approximating one indicate continuous change (distance model in CoMET). κ values are calculated for genitalic shapes by analysing the distances between species in principal component (PC) space, which is why only one value is given for each genitalic structure. κ could not be calculated for characters with missing species data. Pagel's κ value was calculated for all characters for which data on all ten *Timema* species. κ was not calculated for those characters where data was missing, and a negative κ value was rounded up to 0, as the implications for a negative κ do not differ from those of $\kappa = 0$.

phylogenetic-signal model of pure-phylogenetic change fit most of the genitalic traits, although a model of nonphylogenetic change also fit right cercus and intradextral process morphology. PC shape axes differed to some degree within the three genitalic structures, but these one or two models always best explain the majority of variance. From Phylogenetic Independence tests, the morphology of both cerci shows significant positive phylogenetic autocorrelation.

For all nongenitalic morphological traits, CoMET provided one or two models that fit phenotypic changes better than other models (AIC difference > 3; Table 2). All morphological traits of both males and females were best fit by a tempo model of equal change. In agreement with the results from CoMET, κ values for all available nongenitalic (body) morphology were less than one. These results suggest a pattern of accelerated phenotypic change during speciation for courtship and body traits, although change may also have accumulated through evolutionary time for some morphological body traits. The majority of traits fit a phylogenetic-signal model of nonphylogenetic change, although a model of pure-phylogenetic change was not significantly worse for hind tibia length of both sexes, female middle tibia length,

male head width or female thorax width, and a model of punctuated change also fit changes in male and female body length.

There was generally good agreement between our two metrics that quantify the tempo of character change (κ values and results of the CoMET model tests for tempo of change) and between our two metrics that measure phylogenetic autocorrelation (Phylogenetic Independence and the phylogenetic-signal test results from CoMET). Moreover, the same models of change were chosen for all characters across nine alternative phylogenies, indicating that these results are robust with respect to uncertainties in the phylogeny.

Discussion

Courtship characters of male *Timema* walking-sticks demonstrate patterns of change primarily in proportion to speciation events, whereas genitalic characters change more or less continuously. Courtship characters also show high levels of intraspecific variation and low species specificity, which is concordant with behavioural observations showing that courtship is not used for species recognition among well-differentiated species in extant populations (Arbuthnott & Crespi, 2009). However, our intraspecific analyses of *T. cristinae* courtship provide evidence that among-population differences in courtship may be involved in the speciation process. In contrast to courtship traits, the pattern of change (proportionality to time) inferred for genitalic characters implicates continuous selection, such as sexual selection independent of speciation events, in genitalic diversification. These findings indicate that courtship and genitalia traits undergo notably different patterns of macroevolutionary change, apparently in association with different microevolutionary forces during and between speciation events.

Courtship diversification

Courtship behaviour effectively isolates populations and species in many taxa (e.g. Hoikkala & Welbergen, 1995; Henry *et al.*, 2002; Hoikkala *et al.*, 2002), although in other taxa, courtship does not play an obvious role in observed levels of reproductive isolation (e.g. Boake & Hoikkala, 1995; Price & Boake, 1995; Saarikettu *et al.*, 2005). The effects of courtship behaviour on reproductive isolation among pairs or small sets of related species have been studied extensively, but the dynamics of phylogenetic change in courtship and other sexual behaviours have yet to be investigated in sufficient detail for robust comparative inferences to be drawn. In this study, we examined the patterns of evolutionary change in *Timema* courtship characters to test alternative hypotheses, based on sexual selection and species recognition, regarding the evolutionary forces acting on this suite of behavioural traits.

Timema courtship traits show different patterns of phylogenetic autocorrelation, but both CoMET and κ-based analyses indicate that phenotypic change tends to be independent of time for all courtship traits. Such discontinuous diversification suggests that change in courtship may occur primarily in association with speciation events, which in turn indicates that more or less continuous selective processes, such as sexual selection, are unlikely to drive courtship diversification in *Timema* independently of speciation. The courtship traits that we measured also show notable intraspecific variation, and discriminant function analyses indicate that these courtship traits are not highly species specific for several species. These findings suggest a lack of strong stabilizing selection on courtship traits between speciation events, which is concordant with previous observations that species recognition occurs before males and females pair (Arbuthnott & Crespi, 2009), such that courtship does not mediate mate choice at the interspecific level among extant species. Chemical or visual cues may mediate species recognition, although the specific mechanisms of interspecific mate discrimination are currently unknown.

In contrast to our among-species findings, among-population differences in courtship, such as those found for *T. cristinae*, may still be involved in the speciation process. A role for courtship in *Timema* speciation is also supported by evidence for reinforcement of prezygotic isolation in this genus (Nosil *et al.*, 2003; Arbuthnott & Crespi, 2009), which might be expected to drive rapid, speciation-associated change in the behavioural characters that underlie mating. Given that reinforcement may be important in *Timema* speciation, we postulate that mate discrimination should be selected to occur earlier in male–female interactions as populations and species diverge, given benefits to both sexes from efficient avoidance of interspecific pairing and mating. In *Timema*, the mode of reproductive isolation has been inferred to change across evolutionary time scales, as isolation between ecologically divergent populations is mediated through a decreased probability of courtship, whereas species-level isolation occurs as decreased probability of pairing (Arbuthnott & Crespi, 2009).

Courtship *in Timema* may function in particular as a population-recognition cue, given that courtship traits show patterns consistent with change at speciation but do not appear to be used in species recognition or isolation between ecologically divergent populations (Arbuthnott & Crespi, 2009). By this scenario, females would use male courtship to recognize and preferentially mate with individuals of their own population. However, as populations and incipient species diverge, and the cost of interpopulation matings increases, reinforcement may promote the use of cues used earlier in male–female interactions, such as chemical or visual traits, to more efficiently avoid costly matings. Therefore, courtship may be important in the early stages of population divergence, when it is under stabilizing selection, but become less important during later stages of divergence. Testing this

hypothesis requires fine-scale analyses of courtship evolution across populations and related species.

Genitalic diversification

The evolutionary forces underlying the diversity of form and function in animal genitalia have been the subject of considerable debate for many years (e.g., see Alexander *et al.*, 1997; Eberhard, 1997). Alternate hypotheses such as lock-and-key, cryptic female choice, sperm competition and sexual conflict have been described (Arnqvist, 1997), but little data exist regarding patterns of macroevolutionary change in genitalia to make inferences about the relative strength of these alternative mechanisms on a broad scale. McPeek *et al.* (2008) highlight the need for information on the dynamics of change, rather than just quantifications of current diversity, in testing alternative hypotheses regarding genitalic diversification.

The majority of theory and empirical work on genitalia focuses on male intromittent organs, the structures that directly deliver sperm and seminal fluid. Such internal genitalia are commonly associated with complex, secondary sexual characters, such as structures (e.g. claspers) for manipulating components of female reproductive morphology. As these secondary sexual traits are indirectly involved in the transfer of sperm, it is not always clear whether they should be considered genitalia. A useful framework for studying genitalic evolution may be to consider genitalia as an integrated system of morphological structures that are used to transfer sperm and promote its use in females. Because *Timema*'s tripartite claspers are integral in the deployment of the aedeagus, the organ of intromission, and are closely associated with the aedeagus in a morphological context, we consider change in these claspers to be involved in the evolution of genitalia. However, the relationship between *Timema*'s internal and external genitalia has yet to be examined.

In *Timema*, both CoMET and κ-based analyses indicate that change in external genitalic structures is best fit by a model of change proportional to evolutionary time, which suggests that the evolution of *Timema* genitalic morphology is largely or entirely independent of speciation events. The κ value of the right cercus suggests that change may be accelerated to some degree during speciation events, but this calculation, coupled with CoMET models, suggest that change is primarily continuous. Therefore, continuous selection, such as sexual selection and/or sexual conflict, may underlie genitalic shape diversification in *Timema*. In contrast to such patterns for genitalic morphology, nonsexual morphological traits of *Timema*, such as leg length and thorax width, show a pattern of change primarily at speciation, similar to courtship characteristics (although κ indicates that these traits may also show some change between speciation events). This pattern of change may be indicative of the importance of body morphology in adaptation to different host plants via crypsis (e.g., Sandoval & Crespi, 2008), which is a strong, well-documented selective force in this genus (Nosil & Crespi, 2004).

The functions of *Timema* genitalia, aside from a general role in clasping the female's abdomen and prying the female's subgenital plate apart via the intradextral process, remain to be investigated. With respect to sexual selection, *Timema* external genitalia may stimulate females and encourage sperm use (copulatory courtship; Eberhard, 1985, 2004). Genitalic shape may also change in the context of males more effectively clasping onto intraspecific females, to facilitate efficient copulation or reduce copulation disruption from competing males. Finally, if genitalic structures are capable of carrying out their specific functions regardless of shape differences, genitalic shape may represent a neutral character. By this hypothesis, changes in shape represent effects of drift, which might be expected to generate change proportional to time if speciation does not involve an acceleration of drift, for example via founder effects. Further data on copulatory behaviour, sperm transfer and female sperm use are needed to discriminate between these and other hypotheses.

The inference that change in genitalic shape is best fit by a model of continuous divergence agrees with a considerable body of work suggesting that genitalic diversification is mediated by sexual selection; for example, Arnqvist (1998) found that phylogenetic groups with strong sexual selection show up to twice as much genitalic diversity as monogamous groups, and both Eberhard (1985) and Arnqvist (1997) conclude that sexual selection best fits pattern of genitalic diversity for most animals. Like waterstriders (which meet predictions of genitalic diversification through sexual conflict; Arnqvist & Thornhill, 1998; Hosken & Stockley, 2004), *Timema* genitalic morphology also shows intraspecific variation comparable to interspecific differences in some genitalic structures, which implicates continuous processes driving genitalic diversification.

The tempo of genitalic evolution has been investigated in only one other study system, *Enallagma* damselflies, using the combined phylogenetic and morphometric approach deployed here. In these damselflies, males do not court females, male genitalic structures are highly species specific and behavioural experiments provide strong evidence for a role of genitalic structures in species recognition (Paulson, 1974; Robertson & Paterson, 1982; Fincke, 1982; Fincke *et al.*, 2007). In agreement with these observations, phylogenetic analyses demonstrated that male genitalic shape change of *Enallagma* is accelerated during speciation events (McPeek *et al.*, 2008), an inference further corroborated by correlated changes in associated morphology of female mating structures, also evolving in accordance with a model of speciational change (McPeek *et al.*, 2009).

The diversification of genitalic morphology in *Enallagma* provides an interesting contrast to *Timema* (Table 3). In *Timema*, species recognition takes place prior

Table 3 Comparison of sexual trait diversification patterns and functions in *Enallagma* damselflies and *Timema* walking-sticks.

	Enallagma	*Timema*
Mating system	Females arrive at pond to oviposit. Males of several species compete for the opportunity to mate with females, with the successful male clasping on to the female. Females may accept or reject males at this stage (Fincke, 1982; Fincke *et al.*, 2007)	After initial contact, males and females pair with conspecifics. Males then court females, and female rejection at this stage is uncommon. Males and females copulate for several hours, after which males remain on the female mate guarding for up to several days (Arbuthnott & Crespi, 2009)
Species recognition	Females reject heterospecific males at clasping (Fincke *et al.*, 2007; McPeek *et al.*, 2008)	Males usually do not pair with heterospecific females (Arbuthnott & Crespi, 2009)
Sexual selection	Males of several species compete for control of female; males attempt to mate with any female (Fincke, 1982; Fincke *et al.*, 2007)	Males preferentially court females from the same host plant, females can reject males during copulation attempts (Arbuthnott & Crespi, 2009)
Genitalic function	Clasping, species recognition (McPeek *et al.*, 2008)	Clasping, possible sexual selection or sexual conflict functions
Pattern of genitalic change	Proportional to speciation events (McPeek *et al.*, 2008)	Primarily continuous
Courtship function	No courtship	In *T. cristinae* males preferentially court females from the same host plant (Arbuthnott & Crespi, 2009) possible role in early reproductive isolation
Pattern of courtship change	Not relevant	Proportional to speciation events

to copulation, which may remove genitalic structures from selective pressures underlying reproductive isolation. In comparison, *Enallagma* damselflies lack courtship, and males of several species harass females and compete to mate (Fincke, 1982; Fincke *et al.*, 2007), which may select for a species recognition function for both male and female genitalic morphology. Observed differences between these two genera suggest that species recognition systems may interact with systems underlying sexual selection, which directs and limits the diversification patterns of the traits involved in these processes. Because *Enallagma* do not show effective behavioural modes of species recognition, genitalia serve this role in male–female interactions, and change primarily at speciation. *Timema*, in contrast, exhibit a mating system which allows for behavioural isolation early in male–female interactions, which may predispose mating behaviour to changes primarily during speciation but removes genitalia from selective pressures specific to reproductive isolation. Additional studies that integrate phylogenetic pattern with microevolutionary process, across diverse clades of animals, are required to further evaluate the separate and joint roles of sexual selection, species recognition and other processes in the diversification of sexual traits.

Acknowledgments

We thank G. Arnqvist, Michael Jennions and an anonymous reviewer for comments on a previous version of this manuscript. We thank Tanja Schwander and Patrik Nosil for help in the field, and the FAB laboratory group at Simon Fraser University for helpful discussion. Tanja Schwander and Lisa Raeburn performed necessary laboratory work. Todd Oakley provided laboratory space. DA was supported by a NSERC CGS-M grant, and BJC was supported by a NSERC Discovery grant. MAM was supported by National Science Foundation grant IBN-0516104.

References

Abouheif, E. 1999. A method for testing the assumption of phylogenetic independence in comparative data. *Evol. Ecol. Res.* **1**: 895–909.

Alexander, R.D., Marshall, D.C. & Cooley, J.R. 1997. Evolutionary perspectives on insect mating. In: *The evolution of mating systems in insects and arachnids* (J.C. Choe & B.J. Crespi, eds), pp. 4–31. Cambridge University Press, Cambridge, UK.

Andersson, M. 1994. *Sexual Selection*. Princeton University Press, Princeton, NJ.

Arbuthnott, D. & Crespi, B.J. 2009. Courtship and mate discrimination within and between species of Timema walking-sticks. *Anim. Behav.* **78**: 53–59.

Arnqvist, G. 1997. The evolution of animal genitalia: distinguishing between hypotheses by single species studies. *Biol. J. Linn. Soc.* **60**: 365–379.

Arnqvist, G. 1998. Comparative evidence for the evolution of genitalia by sexual selection. *Nature* **393**: 784–786.

Arnqvist, G. & Thornhill, R. 1998. Evolution of animal genitalia: patterns of phenotypic and genotypic variation and condition dependence of genital and non-genital morphology in water strider (Heteroptera: Gerridae: Insecta). *Genet. Res.* **71**: 193–212.

Blows, M.W. & Allan, R.A. 1998. Levels of mate recognition within and between two *Drosophila* species and their hybrids. *Am. Nat.* **152**: 826–837.

Boake, C.R.B. & Hoikkala, A. 1995. Courtship behaviour and mating success of wild-caught Drosophila-silvestris males. *Anim. Behav.* **49**: 1303–1313.

Butlin, R.K. 1989. Reinforcement of premating isolation. In: *Speciation and Its Consequences* (D. Otte & J.A. Endler, eds), pp. 158–179. Sinauer, Sunderland, MA.

Butlin, R.K. & Ritchie, M.G. 1994. Behaviour and speciation. In: *Behaviour and Evolution* (P.J.B. Slater & T.R. Slater, eds), pp. 43–79. Cambridge University Press, Cambridge, UK.

Cameron, S.L., Barker, S.C. & Whiting, M.F. 2006. Mitochondrial genomics and the new insect order Mantophasmatodea. *Mol. Phylogenet. Evol.* **38**: 274–279.

Cardoso, G.C. & Mota, P.G. 2008. Speciational evolution of coloration in the genus *Carduelis*. *Evolution* **62**: 753–762.

Coyne, J.A. & Orr, A.H. 2004. *Speciation*. Sinauer Associates Publishers, Sunderland, MA.

Eberhard, W.G. 1985. *Sexual Selection and Animal Genitalia*. Harvard University Press, Cambridge, MA.

Eberhard, W.G. 1997. Sexual selection by cryptic female choice in insects and arachnids. In: *The evolution of mating systems in insects and arachnids* (J.C. Choe & B.J. Crespi, eds), pp. 32–57. Cambridge University Press, Cambridge, UK.

Eberhard, W.G. 2004. Rapid divergent evolution of sexual morphology: comparative tests of antagonistic coevolution and traditional female choice. *Evolution* **58**: 1947–1970.

Fincke, O.M. 1982. Lifetime mating success in a natural population of the damselfly *Enellagma hageri* (Walsh) (Odanata, Coenagrionidae). *Behav. Ecol. Sociobiol.* **10**: 293–302.

Fincke, O.M., Farevieille, A. & Schulz, T.D. 2007. Lack of innate preference for morph and species identity in mate-searching Enallagma damselflies. *Behav. Ecol. Sociobiol.* **61**: 1121–1131.

Henry, C.S., Wells, M.L.M. & Holsinger, K.E. 2002. The inheritance of mating songs in two cryptic, sibling lacewing species (Neuroptera : Chrysopidae : Chrysoperla). *Genetica* **116**: 269–289.

Hoikkala, A. & Welbergen, P. 1995. Signals and responses of females and males in successful and unsuccessful courtships of 3 Hawaiian lek-mating Drosophila species. *Anim. Behav.* **50**: 177–190.

Hoikkala, A., Saarikettu, M., Paallysaho, S. & Ritchie, M.G. 2002. Co-evolution of male courtship songs and female song preferences in Drosophila virilis group species. *Behav. Genet.* **32**: 470–470.

Hosken, D.J. & Stockley, P. 2004. Sexual selection and genital evolution. *Trends Ecol. Evol.* **19**: 87–93.

Huber, B.A., Sinclair, B.J. & Scmitt, M. 2007. The evolution of asymmetric genitalia in spiders and insects. *Biol. Rev.* **82**: 647–698.

Law, J.H. & Crespi, B.J. 2002. The evolution of geographic parthenogenesis in Timema walking-sticks. *Mol. Ecol.* **11**: 1471–1489.

Lee, C., Blay, S., Mooers, A.O., Singh, A. & Oakley, T.H. 2006. CoMET: a mesquite package for comparing models of continuous character evolution on phylogenies. *Evol. Bioinform. Online* **2**: 193–196.

McPeek, M.A., Shen, L., Torrey, J.Z. & Farid, H. 2008. The tempo and mode of three-dimensional morphological evolution in male reproductive structures. *Am. Nat.* **171**: E158–E178.

McPeek, M.A., Shen, L. & Farid, H. 2009. The correlated evolution of three-dimensional reproductive structures between male and female damselflies. *Evolution* **63**: 73–83.

Mooers, A.Ø., Vamosi, S.M. & Schluter, D. 1999. Using phylogenies to test macroevolutionary hypotheses of trait evolution in Cranes (Gruinae). *Am. Nat.* **154**: 249–259.

Nosil, P. 2004. Reproductive isolation caused by visual predation on migrants between divergent environments. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **271**: 1521–1528.

Nosil, P. & Crespi, B.J. 2004. Does gene flow constrain adaptive divergence or vice versa? A test using ecomorphology and sexual isolation in Timema cristinae walking-sticks. *Evolution* **58**: 102–112.

Nosil, P. & Crespi, B.J. 2006. Ecological divergence promotes the evolution of cryptic reproductive isolation. *Proc. R. Soc. B Biol. Sci.* **273**: 991–997.

Nosil, P., Crespi, B.J. & Sandoval, C.P. 2002. Host-plant adaptation drives the parallel evolution of reproductive isolation. *Nature* **417**: 440–443.

Nosil, P., Crespi, B.J. & Sandoval, C.P. 2003. Reproductive isolation driven by the combined effects of ecological adaptation and reinforcement. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **270**: 1911–1918.

Nylander, J.A.A. 2004. *MrModeltest v2*. Program distributed by the author. Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden.

Pagel, M. 1997. Inferring evolutionary processes from phylogenies. *Zool. Scr.* **26**: 331–348.

Panhuis, T.M., Butlin, R., Zuk, M. & Tregenza, T. 2001. Sexual selection and speciation. *Trends Ecol. Evol.* **16**: 364–371.

Papadopoulou, A., Anastasiou, I. & Vogler, A.P. in press. Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. *Mol. Biol. Evol.*

Paulson, D.R. 1974. Reproductive isolation in damselflies. *Syst. Zool.* **23**: 40–49.

Pavoine, S., Ollierb, S., Pontiera, D. & Chessela, D. 2008. Testing for phylogenetic signal in phenotypic traits: new matrices of phylogenetic proximities. *Theor. Popul. Biol.* **73**: 79–91.

Polihronakis, M. 2006. Morphometric analysis of intraspecific shape variation in male and female genitalia of Phyllophaga hirticula (Coleoptera : Scarabaeidae : Melolonthinae). *Ann. Entomol. Soc. Am.* **99**: 144–150.

Prager, M. & Andersson, S. 2009. Phylogeny and evolution of sexually selected tail ornamentation in widowbirds and bishops (*Euplectes* spp.). *J. Evol. Biol.* **22**: 2068–2076.

Price, D.K. & Boake, C.R.B. 1995. Behavioural reproductive isolation in Drosophila-silvestris, D-heteroneura, and their F1 hybrids (Diptera, Drosophilidae). *J. Insect Behav.* **8**: 595–616.

Robertson, H.M. & Paterson, H.E.H. 1982. Mate recognition and mechanical isolation in *Enallagma* damselflies (Odanata: Coenagrionidae). *Evolution* **36**: 243–250.

Ronquist, F. & Huelsenbeck, J.P. 2003. MRBAYES 3: Bayesion phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572–1574.

Saarikettu, M., Liimatainen, J.O. & Hoikkala, A. 2005. Intraspecific variation in mating behaviour does not cause sexual isolation between Drosophila virilis strains. *Anim. Behav.* **70**: 417–426.

Sandoval, C.P. & Crespi, B.J. 2008. Adaptive evolution of cryptic coloration: the shape of host plants and dorsal stripes in Timema walking-sticks. *Biol. J. Linn. Soc.* **94**: 1–5.

Seddon, N. 2005. Ecological adaptation and species recognition drives vocal evolution in neotropical suboscine birds. *Evolution* **59**: 200–215.

Shen, L. & Makedon, F. 2006. Spherical mapping for processing of 3D closed surfaces. *Image Vis. Comput.* **24**: 743–761.

Shen, L., Farid, H. & McPeek, M.A. 2009. Modeling three-dimensional morphological structures using spherical harmonics. *Evolution* **63**: 1003–1016.

Vickery, V.R. 1993. Revision of Timema Scudder (Phasmatoptera, Timematodea) including 3 new species. *Can. Entomol.* **125**: 657–692.

Received 19 December 2009; revised 16 March 2010; accepted 17 March 2010