Titel der Arbeit:

ACOUSTIC COMMUNICATION, SEXUAL SELECTION, AND SPECIATION IN FIELD CRICKETS

Dissertation
zur Erlangung des akademischen Grades:

Doctor rerum naturalium (*Dr. rer. nat.*)

eingereicht an der Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin von:

Thomas Blankers

Präsident der Humboldt-Universität zu Berlin:

Prof. Dr. Jan-Hendrik Olbertz

Dekan der Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin :

Prof. Dr. Richard Lucius

Gutachter/innen:

1. Prof R. Matthias Hennig
2. Prof Michael G. Ritchie
3. Prof Kerry L. Shaw

Tag der mündlichen Prüfung:

20. Mai 2016

TABLE OF CONTENTS

CHAPTER 1

GENERAL INTRODUCTION

"The case of the male Argus pheasant is eminently interesting, because it affords good evidence that the most refined beauty may serve as a sexual charm, and for no other purpose"

Charles Darwin – *The Descent of Man, and Selection in Relation to Sex*, 1871

The diversity in sexual displays and behaviours, among species as well as between the two sexes of a single species, is truly fascinating. Striking ornamentation such as display feathers in birds and horns and weaponry in mammals, intriguingly complex courtship rituals seen in exotic birds, and the overwhelming choruses of acoustic insects and frogs calling for mates testify to the fundamental role of sexual advertisement in the life and times of animals everywhere. Almost a century and a half since Darwin's second famous book (1871) where he introduced his concept of sexual selection, we have greatly advanced our understanding of how and why sexual signals are so diverse (*e.g.* Andersson 1994). However, the role of sexual selection in establishing the primary reproductive barriers between populations and thus driving speciation has hitherto remained a topic of active debate (Fisher 1930, Dobzhanksy 1940, West-Eberhard 1983, Andersson 1994, Panhuis *et al.* 2001, Ritchie 2007, Servedio 2015).

Variation in sexually selected traits is directly linked to mating success and thus predisposed to drive reproductive isolation between populations that differ in sexual traits and corresponding mate preferences. Until the early 1980s the role of sexual selection in establishing reproductive barriers and instigating speciation was problematic conceptually because it was unclear how sexual selection could establish and maintain reproductive isolation between closely related species (Ritchie 2007). Following the influential work by Fisher (1930), Lande (1980, 1981), and West-Eberhard (1983) species recognition and sexual selection were reconciled giving rise to a conceptual model for (arbitrary) trait-preference co-evolution and speciation by sexual selection. Since then, empirical evidence supporting a strong role for sexual selection in facilitating speciation has accumulated for a wide range of taxa (Arbuthnott 2009, Kraaijeveld *et al.* 2011).

However, there are still many challenges to a comprehensive synthesis of sexual selection and speciation. One of the biggest conundrums concerns the case where populations are geographically overlapping and/or exchange migrants. Parapatric or sympatric speciation was once thought to be rare and unlikely, but the contemporary view is that speciation frequently unfolds in the face of gene flow (Bolnick & Fitzpatrick 2007, Bird *et al.* 2012). At the same time, theoretical research predicts that sexual selection has only very limited or potential adverse effects on speciation when individuals migrate readily between diverging populations (van Doorn *et al.* 2004, Weissing *et al.* 2011, Servedio 2015). This is in strong contrast to observations for several species rich radiations where divergence rates are estimated to be extremely high among geographically overlapping species that have parted primarily in mating behaviour (*e.g.* cichlid fish: Seehausen 2000, Wagner *et al.* 2012; swordtail crickets: Mendelson & Shaw 2005).

Whether or not currently sympatric species have at some point exchanged genetic material is often unclear. Thus, one of the main challenges towards understanding the role of sexual selection in speciation is determining how gene flow affects interspecific genomic variation in sexually divergent species and whether genomic regions harbouring loci involved in mating behaviour are less homogenic between species compared to other regions (Feder *et al.* 2013, Sousa & Hey 2013). A second challenge is determining how mate choice drives variation among individuals within species. Mate choice can be a purely aesthetic and arbitrary process (Darwin 1871, Fisher 1930, Lande 1981) in which case a strong association between the genetic architecture of traits and preferences (to allow for trait-preference co-evolution) is expected. Alternatively, sexual traits are used by the choosing sex to gain information on the vigour and fitness of the sender (Wallace 1895, Zahavi 1975, Hamilton & Zuk 1982).

We thus need to unravel the genetic architecture of signals and preferences and examine the genomic ramifications of the evolution of reproductive isolation between sexually divergent, ecologically similar, and geographically proximate taxa. This first requires a profound understanding of the behavioural basis of mate choice and quantitative genetic insights in the evolution of signals and preferences (*i.e.* the extent to which preferences match signals, the genomic distribution and location of

the loci controlling preference and signal variation, and the selection response constraining phenotypic evolution). Second, to understand the role of sexual signals in speciation we must unravel the evolutionary historical context in which they evolved. Few study systems allow for such a broad and comprehensive assessment of the behavioural, quantitative genetics, and genomic implications of sexual selection and sexual signal evolution.

Crickets (Orthoptera: Gryllidae) are at the forefront of neurobiological, evolutionary, and behavioural research and are a model system for acoustic communication (Greenfield 2002, Gerhardt & Huber 2002). The most widely studied field cricket genera (*Gryllus*, *Laupala*, and *Teleogryllus*) have been used to unravel neuronal mechanisms of signal production and evaluation (*e.g.* Bennet-Clark 2003, Hedwig & Poulet 2004, Hennig *et al.* 2004, Schöneich and Hedwig 2012, Schöneich *et al.* 2015), to study the evolution of mating behaviour (*e.g.* Simmons 2004, Gray 2011), and to inform the genetic architecture of sexual signals and corresponding preferences (*e.g.* Hoy 1974, Shaw *et al.* 2007). Closely related species of crickets often occur sympatrically and are morphologically and ecologically cryptic, but show strong differentiation in the male acoustic mating signals and corresponding female preferences (Alexander 1962, Otte 1992, Otte 1994) Recently, genomic and transcriptomic resources for several species of field crickets and other orthopterans have become available (*e.g.* Andres *et al.* 2013, Bailey *et al.* 2013 Berdan *et al.* 2015, Berdan *et al.* in review). Thus, a substantial amount of neurological, behavioural, and genetic resources combined with geographically dense distributions of closely related species and nonetheless strongly divergent acoustic mating signals render crickets an ideal system for the study of mate choice behaviour, genetic control of sexual signals, and the role of divergence in mating behaviour in speciation.

In this thesis I examine acoustic sexual communication in North-American field crickets (*Gryllus*). Using a combination of behavioural, quantitative genetics, and genomic approaches, this thesis aims at (1) understanding the evolution of the calling song in crickets and the role of song traits in reproductive isolation, and (2) unravelling the contribution of demography, gene flow, and (sexual) selection to genome-wide patterns of genetic divergence during speciation. Furthermore, this thesis provides insight in multivariate phenotypic evolution and the mechanistic basis of acoustic behaviour in field crickets. Below I will briefly introduce these two topics.

Phenotypic evolution and spherical cows

Phenotypic and genetic distributions of biological traits are multivariate and thus form a complex trait space. This space is described by the orthogonal phenotypic dimensions in which traits vary. Because these dimensions consist of combinations of traits (which are thus correlated) and are not of equal variance, we cannot assume that the trait space is spherical: the evolution of a single trait can only be understood when the variation in other traits that covary with the focal trait are taken into account (Lande 1979, Lande & Arnold 1983).

If two traits are strongly correlated and the greatest phenotypic dimension of variation does not align with the direction of selection the response to selection can be biased away from the selection optimum (Steppan 2002, Blows 2007; Fig. 1a). This is commonly referred to as an evolutionary constraint on selection. When only one of two traits is under selection, a strong correlation between the traits will result in indirect selection acting on the trait that is not under selection (Fig. 1b). Despite a comprehensive framework developed by Lande and Arnold (Lande 1979, Lande & Arnold 1983) and advances in analytical tools (*e.g.* Aguirre *et al.* 2014) many studies of quantitative phenotypic evolution largely ignore the genetic constraints and pleiotropic effects that can result from covariance among traits. This will strongly overestimate the selection response (Fig. 1a versus Fig. 1c) or neglect the effects of indirect selection (Fig. 1b versus Fig. 1d).

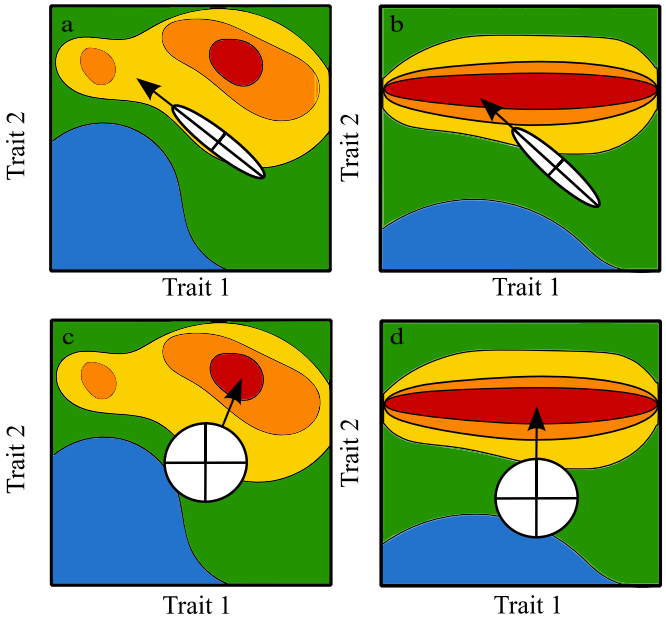

Figure 1 Multivariate response to selection (adapted from Steppan *et al.* 2002). The white ellipse represents a strong correlation between two traits in a population in such a way that the response to selection is constrained (a). Due to the patterns of covariation, the population in (a) cannot evolve towards the fitness optimum in the top-right corner. In (b), the fitness peak is much broader because trait 1 is not under selection directly. The population can respond to selection acting on trait 2, but due to the strong correlation between trait 1 and 2, the traits will covary in response to this selection and trait 1 will be dragged along by indirect selection. In (c) and (d) the covariance is effectively zero (spherical) and the selection response is neither constrained nor results in indirect selection.

The multivariate phenotype and the additive genetic (co)variation for multivariate traits are described by the phenotypic and genetic variance-covariance matrix, denoted **P** and **G**, respectively. These are symmetric *n* x *n* matrices with the variances of the *n* traits on the diagonal and the covariance between each trait pair in the off-diagonal matrix elements. Both **P** and **G** are vital to the mathematical description of evolution and are used to study genetic constraints, pleiotropy, phenotypic plasticity, and predict the evolutionary response to natural and sexual selection (Lande 1979, Lande & Arnold 1983, Lynch & Walsh 1998).

Similar to morphological traits, sexual signals like the calling song of crickets are often multicomponent (Scheuber *et al.* 2004, Hebets & Papaj 2005, Higham & Hebets 2013). Despite the caveats to univariate examinations of sexual signal evolution (Blows 2007), a comprehensive overview of multivariate divergence in sexual signals is missing and our understanding is limited to insights from only a handful of studies (*e.g.*, Blows *et al.* 2004, Bentsen *et al.* 2006, Bertram *et al.* 2011). This thesis provides a detailed examination of interspecific divergence in multivariate song phenotypes across several cricket species as well as in hybrids between species. Simultaneously the song preference space is explored in a multivariate context to gain insight into the dimensionality and strength of sexual selection. One of the powerful features of studying sexual signalling behaviour in crickets is that these results can be coupled with several decades' worth of neuro-ethological insights into the mechanisms of song production and song perception which are discussed briefly below.

Acoustic mate choice in crickets

Cricket males produce songs by rubbing their forewings (*i.e.* stridulation). With each closing movement of the wing, the plectrum or scraper of one of the wings (typically the left wing) excites the stridulatory teeth of the file on the bottom of the other wing. The number of excitations of the teeth and the size and resonant properties of the wing (*i.e.* the harp) determine the carrier frequency (*i.e.* the pitch) of the song; the temporal parameters of the song (*i.e.* pulse and chirp rhythm) are determined by the rate and number of opening and closing movements of the wing (reviewed in Gerhardt & Huber 2002). In this thesis I will frequently discriminate between two song types: trilled songs, which are long trains of pulses (the convention is >20 pulses, Alexander 1962) irregularly interspaced with pauses, and chirped

songs, which are more regular groups of between 3 and 10 pulses. Throughout this thesis, species producing trills and species producing chirps will be referred to as 'trillers' and 'chirpers', respectively.

The central nervous system governs both the activation and patterning of the acoustic signals (reviewed in Gerhardt & Huber 2002). Chemical (*i.e.* by injecting acetylcholine [ACh] or gamma-aminobutyric acid [GABA]) or electrical stimulation of specific locations in the brain elicit or inhibit calling (and courtship/aggressive) song production. The excitatory regions are concentrated in the protocerebrum and adjacent to the mushroom body. The fine-scale temporal properties of the song depend on neuronal networks called pattern generators that control the movement of stridulatory organs. Neuromuscular properties and intrinsic properties of the muscles affect the temporal rhythm of the song. Movement of the wings for the purpose of both flight and stridulation involve, at least in part, the same muscles and motor neurons, but stridulation and flight are controlled by distinct neuronal networks (Hennig 1990). Although the neuronal control of song productions is hardwired early on in development, likely as early as during the final stages of embryogenesis, neurogenesis is continuous throughout the life of crickets (Cayre *et al.* 1994).

Acoustic signals are processed by females first in the auditory periphery, where the mechanical properties of the ear (tympanum) determine the sensitivity to a specific carrier frequency (Gerhardt & Huber 2002). A neuronal representation of the stimulus is then passed on to the central auditory pathway, where a complex network of ascending neurons and local brain neurons with varying filter properties is used in evaluation of the sound signal (Gerhardt & Huber 2002, Hennig *et al.* 2014, Schöneich *et al.* 2015).

Female crickets respond to the male song and move towards the sound source (phonotaxis) either in mid-air or while walking. A single pulse is sufficient to induce steering behaviour in female crickets (Hedwig & Poulet 2004), but integration of the sensory output over longer timescales determines the direction in which they walk or fly. Preferences are usually finely tuned to conspecific songs and combining the information from multiple traits is expected to facilitate accurate discrimination against heterospecifics. However, preference functions for specific song traits can also differ in shape, *i.e.* a function can be open-ended ('linear' preference) or closed or unimodal ('concave' preference). The shape of the preference function depends on the filter properties of the neurons of the auditory pathway and is potentially related to the information a trait conveys (*i.e.* 'essential' versus 'motivational' traits [Popov & Shuvalov 1976] or 'static' versus 'dynamic' traits [Gerhardt 1991]).

A**ims**

The aims of this thesis are twofold. First, intra- and interspecific variation in songs and preferences are analysed to provide insight into the role of song traits in mate choice and in the evolution of reproductive isolation. **Chapter 2** elucidates the preference mechanisms for acoustic communication in three trilling field crickets and elaborates on the role of song traits in sexual communication; **Chapter 3 and 4** focus on the evolvability (*sensu lato*) of the song by disentangling the genetic architecture and comparing multivariate song divergence to female preference behaviour. Second, genome-wide nucleotide polymorphism data from two closely related and sympatrically occurring species will be used to discuss patterns of genetic variation and footprints of selection in an evolutionary historical context (**Chapter 5**). In the general discussion (**Chapter 6**) I will then elaborate on the proximate basis and ultimate consequences of acoustic mate choice behaviour and on the role of sexual selection on speciation with gene flow.

REFERENCES

Arbuthnott, D. 2009. The genetic architecture of insect courtship behavior and premating isolation. *Heredity*. **103**: 15–22.

Aguirre, J.D., Hine, E., McGuigan, K. & Blows, M.W. 2014. Comparing G: multivariate analysis of genetic variation in multiple populations. *Heredity*. **112**: 21–9.

Alexander, R. 1962. Evolutionary change in cricket acoustical communication. *Evolution* **6**: 443–467.

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

Andrés, J.A., Larson, E.L., Bogdanowicz, S.M. & Harrison, R.G. 2013. Patterns of transcriptome divergence in the male accessory gland of two closely related species of field crickets. *Genetics* **193**: 501–513.

Bailey, N.W., Veltsos, P., Tan, Y.F., Millar, A.H., Ritchie, M.G. & Simmons, L.W. 2013. Tissue-specific transcriptomics in the field cricket *Teleogryllus oceanicus*. *G3 Genes|Genomes|Genetics* **3**: 225–230.

Bennet-Clark, H.C. 2003. Wing resonances in the Australian field cricket *Teleogryllus oceanicus*. *J. Exp. Biol.* **206**: 1479–96.

Bentsen, C.L., Hunt, J., Jennions, M.D. & Brooks, R. 2006. Complex multivariate sexual selection on male acoustic signaling in a wild population of *Teleogryllus commodus*. Am. Nat. 167: E102–E116.
Berdan, E.L., Blankers, T., Waurick, I., Mazzoni, C.J. & Mayer, F. in Review. A genes eye view of ontogeny: De novo assembly and profiling of a *Gryllus rubens* transcriptome. *Mol. Ecol. Resour.*
Berdan, E.L., Mazzoni, C.J., Waurick, I., Roehr, J.T. & Mayer, F. 2015. A population genomic scan in *Chorthippus* grasshoppers unveils previously unknown phenotypic divergence. *Mol. Ecol.* **24**: 3918–30.
Bertram, S.M., Fitzsimmons, L.P., McAuley, E.M., Rundle, H.D. & Gorelick, R. 2011. Phenotypic covariance structure and its divergence for acoustic mate attraction signals among four cricket species. *Ecol. Evol.* **2**: 181–195.
Bird, C.E., Fernandez-Silva, I., Skillings, D.J. & Toonen, R.J. 2012. Sympatric Speciation in the Post "Modern Synthesis" Era of Evolutionary Biology. *Evol. Biol.* **39**: 158–180.
Blows, M.W. 2007. A tale of two matrices: multivariate approaches in evolutionary biology. *J. Evol. Biol.* **20**: 1–8.
Blows, M.W., Chenoweth, S.F. & Hine, E. 2004. Orientation of the genetic variance-covariance matrix and the fitness surface for multiple male sexually selected traits. *Am. Nat.* **163**: 329–340.
Bolnick, D.I. & Fitzpatrick, B.M. 2007. Sympatric Speciation : Models and Empirical Evidence. *Annu. Rev. Ecol. Evol. Syst.* **38**: 459–487.
Cayre, M., Strambi, C. & Strambi, A. 1994. Neurogenesis in an adult insect brain and its hormonal control. *Nature* **368**: 57–59.
Darwin, C. 1871. *The Descent of man, and Selection in Relation to Sex*. D. Appleton and Company, New York.
Dobzhansky, T. 1940. Speciation as a Stage in Evolutionary Divergence. *Am. Nat.* **74**: 312.
Doorn, G.S. Van, Dieckmann, U. & Weissing, F.J. 2004. Sympatric Speciation by Sexual Selection : A Critical Reevaluation. *Am. Nat.* **163**: 709–725.
Feder, J.L., Flaxman, S.M., Egan, S.P., Comeault, A.A. & Nosil, P. 2013. Geographic Mode of Speciation and Genomic Divergence. *Annu. Rev. Ecol. Evol. Syst.* **44**: 73–97.
Fisher, R. 1930. *The genetical theory of natural selection*. Oxford University Press, New York.
Gerhardt, H.C. 1991. Female choice in treefrogs: static and dynamic acoustic criteria. *Anim. Behav.* **42**: 615–636.
Gerhardt, H.C. & Huber, F. 2002. *Acoustic communication in insects and anurans*. The University of Chicago Press, Chicago.
Gray, D.A. 2011. Speciation, divergence, and the origin of *Gryllus* rubens: behavior, morphology, and molecules. *Insects* **2**: 195–209.
Greenfield, M.D. 2002. *Signalers and Receivers: Mechanisms and Evolution of Arthropod Communication*. Oxford University Press, New York.
Hamilton, W. & Zuk, M. 1982. Heritable true fitness and bright birds: a role for parasites? *Science (80-.).* **218**: 384–387.
Hebets, E.A. & Papaj, D.R. 2005. Complex signal function: developing a framework of testable hypotheses. *Behav. Ecol. Sociobiol.* **57**: 197–214.
Hedwig, B. & Poulet, J.F. 2004. Complex auditory behaviour emerges from simple reactive steering. *Nature* **430**: 781–785.
Hennig, R.M. 1990. Neuronal control of the forewings in two different behaviours: Stridulation and flight in the cricket, *Teleogryllus* commodus. *J. Comp. Physiol. A* **167**: 617–627.
Hennig, R.M., Franz, A. & Stumpner, A. 2004. Processing of Auditory Information in Insects. *Microsc. Res. Tech.* **63**: 351–374.
Higham, J.P. & Hebets, E.A. 2013. An introduction to multimodal communication. *Behav. Ecol. Sociobiol.* **67**: 1381–1388.
Hoy, R.R. 1974. Genetic Control of Acoustic Behavior in Crickets. *Am. Zool.* **14**: 1067–1080.
Kraaijeveld, K., Kraaijeveld-Smit, F.J.L. & Maan, M.E. 2011. Sexual selection and speciation: the comparative evidence revisited. *Biol. Rev.* **86**: 367–377.
Lande, R. 1981. Models of speciation by sexual selection on polygenic traits. *Proc. Natl. Acad. Sci.* **78**: 3721–3725.
Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* **33**: 402–416.
Lande, R. 1980. The genetic covariances between characters maintained by pleiotropic mutations. *Genetics* **94**: 203–215.
Lande, R. & Arnold, S.J. 1983. The measurement of selection on correlated characters. *Evolution* **37**: 1210–1226.
Lynch, M. & Walsh, B. 1998. *Genetics and analysis of quantitative Traits*. Sinauer, Sunderland, MA.
Mendelson, T.C. & Shaw, K.L. 2005. Rapid speciation in an arthropod. *Nature* **433**: 375–376.
Otte, D. 1992. Evolution of cricket songs. *J. Orthoptera Res.* **1**: 25–49.
Otte, D. 1994. *The Crickets of Hawaii: Origin, Systematics, and Evolution*. Orthoptera Society/Academy of Natural Sciences of Philadelphia, Philadelphia, PA.
Panhuis, T.M., Butlin, R., Zuk, M. & Tregenza, T. 2001. Sexual selection and speciation. *Trends Ecol. Evol.* **16**: 364–371.
Popov, A. V. & Shuvalov, V.F. 1977. Phonotactic behavior of crickets. *J. Comp. Physiol. A* **119**: 111–126.
Prum, R.O. 2012. Aesthetic evolution by mate choice: Darwin's really dangerous idea. *Philos. Trans. R. Soc. B Biol. Sci.* **367**: 2253–2265.
Ritchie, M.G. 2007. Sexual Selection and Speciation. *Annu. Rev. Ecol. Evol. Syst.* **38**: 79–102.
Servedio, M.R. 2015. Geography, assortative mating, and the effects of sexual selection on speciation with gene flow. Evol. Appl. doi: 10.1111/eva.12296.
Scheuber, H., Jacot, A. & Brinkhof, M.W. 2004. Female preference for multiple condition-dependent components of a sexually selected signal. *Proc Biol Sci* **271**: 2453–2457.
Schöneich, S. & Hedwig, B. 2012. Cellular basis for singing motor pattern generation in the field cricket (*Gryllus bimaculatus* DeGeer). Brain Behav. 2: 707–725.
Schoneich, S., Kostarakos, K. & Hedwig, B. 2015. An auditory feature detection circuit for sound pattern recognition. *Sci. Adv.* **1**: e1500325–e1500325.
Seehausen, O. 2000. Explosive speciation rates and unusual species richness in haplochromine cichlid fishes: effects of sexual selection. *Adv. Ecol. Res.* **31**: 237–274.
Shaw, K.L., Parsons, Y.M. & Lesnick, S.C. 2007. QTL analysis of a rapidly evolving speciation phenotype in the Hawaiian cricket Laupala. *Mol. Ecol.* **16**: 2879–2892.
Simmons, L.W. 2004. Genotypic variation in calling song and female preferences of the field cricket *Teleogryllus* oceanicus. *Anim. Behav.* **68**: 313–322.
Sousa, V. & Hey, J. 2013. Understanding the origin of species with genome-scale data : modelling gene flow. *Nature* **14**: 404–414.

Steppan, S.J., Phillips, P.C. & Houle, D. 2002. Comparative quantitative genetics: evolution of the G matrix. *Trends Ecol. Evol.* **17**: 320–327.
Wagner, C.E., Harmon, L.J. & Seehausen, O. 2012. Ecological opportunity and sexual selection together predict adaptive radiation. *Nature* **487**: 366–369.
Wallace, A.R. 1895. *Natural selection and tropical nature: Essays on descriptive and theoretical biology*. Macmillan, London.
Weissing, F.J., Edelaar, P. & van Doorn, G.S. 2011. Adaptive speciation theory: a conceptual review. *Behav. Ecol. Sociobiol.* **65**: 461–480.
West-Eberhard, M.J. 1983. Sexual selection, social competition, and speciation. *Q. Rev. Biol.* **58**: 155–183.
Zahavi, A. 1975. Mate selection—A selection for a handicap. *J. Theor. Biol.* **53**: 205–214.

CHAPTER 2

CONSERVATION OF MULTIVARIATE FEMALE PREFERENCE FUNCTIONS AND PREFERENCE MECHANISMS IN THREE SPECIES OF TRILLING FIELD CRICKETS

Based on: Blankers, T, Hennig RM. Gray DA. 2015. Conservation of multivariate female preference functions and preference mechanisms in three species of trilling field crickets. *Journal of Evolutionary Biology* **28**: 630-641.

Abstract. Divergence in mate recognition systems among closely related species is an important contributor to assortative mating and reproductive isolation. Here we examine divergence in male song traits and female preference functions in three cricket species with songs consisting of long trills. The shape of female preference functions appears to be mostly conserved across species and follows the predictions from a recent model for song recognition. Multivariate preference profiles, combining the pulse and trill parameters, demonstrate selectivity for conspecific pulse rates and high trill duty cycles. The rules for integration across pulse and trill timescales were identical for all three species. Generally, we find greater divergence in male song traits than in associated female preferences. For pulse rate we find a strong match between divergent male traits and female peak preferences. Preference functions for trill parameters and carrier frequency are similar between species and show less congruence between signal and preference. Differences among traits in the degree of trait-preference (mis)match may reflect the strength of preferences and the potential for linkage disequilibrium, selective constraints, and alternative selective pressures, but appear unrelated to selection for species recognition *per se*.

INTRODUCTION

Divergence in sexual signal traits is thought to result from selection imposed by divergent mate preferences (Lande 1981; West-Eberhard 1983; Higashi *et al.* 1999) and co-evolution of signals and preferences helps maintain species specific communication (Gerhardt & Huber 2002; Greenfield 2002). Therefore, male sexual signalling traits are expected to be closely congruent with female mating preferences (Rodriguez *et al.* 2006). However, the degree of congruence between signals and preferences is likely to vary (1) with the strength of those preferences (Rodriguez *et al.* 2006; Rodriguez *et al.* 2013), (2) with the specific genetic mechanisms that may link changes in signals with changes in preferences (Andersson & Simmons 2006; Chenoweth & Blows 2006), and (3) depending upon whether the preferences themselves are under selection because they lead to high quality matings (*i.e.* preferences for traits indicative of direct or indirect benefits) or are free to vary (*i.e.* preferences for arbitrary 'Fisherian' traits).

Because sexual signals are often multimodal and/or multicomponent (Scheuber *et al.* 2004; Hebets & Papaj 2005; Higham & Hebets 2013), understanding the evolution of male signal traits and female preferences requires a multidimensional framework (Blows *et al.* 2003; Gerhardt & Brooks 2009). Multivariate preference functions reflect the complexity of signal processing and allow different selection pressures on different traits to be compared (Bentsen *et al.* 2006; Gerhardt & Brooks 2009; Oh & Shaw 2013). In addition to adopting a multivariate framework, researchers should strive to examine the neuro-physiological mechanisms underlying signal recognition and preference (Römer *et al.* 2002; Bass & McKibben 2003; Fisher *et al.* 2005, Kostarakos & Hedwig 2012). This is critical for understanding the degree of mechanistic change required in order to produce species-level divergence. That is, does the evolution of assortative mating require major changes to mechanisms of signal recognition, or are minor changes to the "tuning" of the same underlying preference mechanisms sufficient to facilitate divergence (Clemens & Hennig 2013; Hennig *et al.* 2014)? Here, we address this question by comparing the variation in the calling song, used by male crickets in long distance mate attraction (Alexander 1962), and corresponding female preferences between three species of field crickets.

A**coustic communication and song recognition in crickets**

Cricket calling songs vary among species in two features: the carrier frequency and the temporal pattern of pulses and chirps/trills (long chirps exceeding 20 pulses, Alexander 1962, Weissman *et al.* 2009, Walker 2014). The temporal pattern of pulses and chirps/trills can be described by the duration, pause, period (or, inversely, rate) and duty cycle (Fig. 1a). Female crickets are known to evaluate carrier frequency and varying combinations of the parameters governing the temporal pattern of a song over short and long timescales (Popov & Shuvalov 1977, Doherty 1985, Grobe *et al.* 2012). In this study we will refer to the evaluation of song parameters, or, matching of the signal with an internal, neuronal template as 'song recognition' or 'signal recognition'. A recent model of song recognition in crickets suggests that females evaluate the song signal with a short, innate template that responds best to a particular pulse parameter (*e.g.* pulse rate). The output of this pulse detector is then integrated (*i.e.*, the accumulation of sensory evidence for decision making, Gold & Shadlen 2007) over a given time window (Clemens & Hennig 2013, Hennig *et al.* 2014). The model predicts that (1) females evaluate the temporal pattern of the pulses, (2) following integration of the pulse detector output, female preference is independent of the temporal pattern of the trill, *i.e.* depends only on the duty cycle (equivalent to song energy over time), and (3) a trill pattern without a pulse rate modulation (*i.e.*, continuous tones at a repetition rate equivalent to the trill rate) is ineffective in eliciting a positive female response.

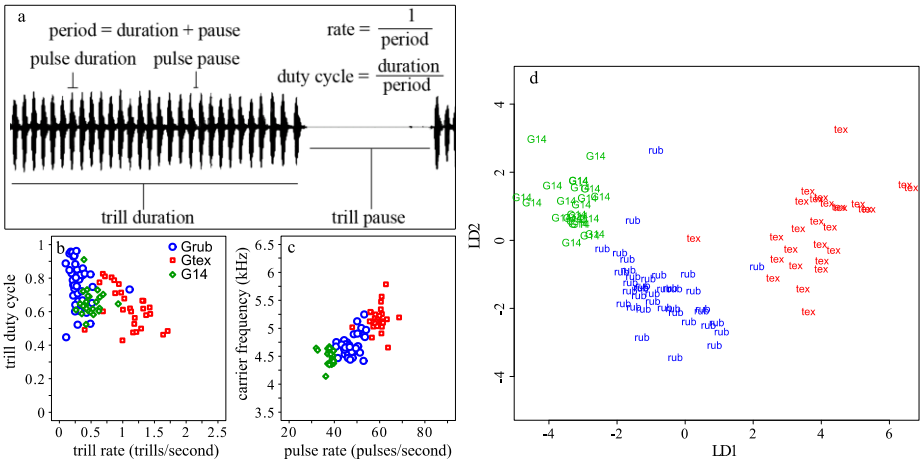

Figure 1 Signals for long distance mate attraction in crickets. A schematic overview of the typical structure of the calling song of *G. rubens*, *G. texensis*, and G#14 is given in (a), along with a definition of all temporal song parameters discussed in this study. In (b) and (c) bivariate plots constrained by trill duty cycle and trill rate, and carrier frequency and pulse rate, respectively, show the intra- and interspecific variation in the calling song. In (d) variation along the first two linear discriminants (LD1 and LD2) is shown. The coefficients of the variables on the linear discriminants and the proportion of the trace can be found in Table 1. Song recordings from 40 *G. rubens* (blue), 37 *G. texensis* (red), and 24 G#14 (green) males were analysed.

Here we adopt a comparative mechanistic approach to investigate whether these three predictions hold in three species of *Gryllus* field crickets that produce long trills (hypothesis 1) and whether interspecific divergence in the multivariate female preference functions is manifested on both timescales and in the spectral content of the song (*i.e.,* carrier frequency; hypothesis 2). We then examine the degree of mismatch between signals and preferences (hypothesis 3) and interpret our results in comparison with other cricket species, and in light of the putative mechanisms that influence the degree of mismatch between male signals and female preferences.

We characterized preference functions of three trilling species of crickets in the genus *Gryllus*: *G. rubens* Scudder 1902, *G. texensis* Cade & Otte 2000, and an undescribed species which we refer to here as 'G#14'. The former two species are partially sympatric sister species (Gray *et al.* 2008; Gray 2011). *Gryllus rubens* ranges across the south-eastern USA and southern gulf states to eastern Texas; *G. texensis* ranges from central Texas across the southern gulf states to western Florida; G#14 has a limited distribution in central Arizona apparently disjunct from populations further south in Mexico. G#14 is entirely allopatric with respect to *G. rubens* and *G. texensis*. It is not currently clear if G#14 is sister to either species or not. All three species produce long trills, in strong contrast with the short chirps found in most *Gryllus* species.

METHODS

Collecting and rearing

Gryllus texensis were collected in Austin (TX), Lancaster (TX) and Round Rock (TX); *G. rubens* were collected in Gainesville (FL), Lake City (FL), and Live Oak (FL); G#14 were collected in Agua Fria National Monument (AZ). The crickets were housed in 19L containers at an average temperature of 25.3 ◦C (+/- 2.73 SD) with gravel, shelter, and water and food *ad libitum*. Males and females were separated before reaching sexual maturity and acoustically isolated during the experimental period. Second generation laboratory offspring was used in the experiments.

Male song recording and analysis

Individuals were placed in separate boxes (mean temperature 24.9 ◦C +/- 0.98 SD) for a 16-24 hour period. Each box was equipped with a microphone and isolating foam to ensure acoustic isolation. Using customized software (LabVIEW, 2007), each microphone (TCM 141 Conrad) was scanned for 800 ms and a male was recorded for 20 seconds if it produced sound during that 800 ms interval. The dominant carrier frequency was determined from the spectral peak of the real time signal. For analysis of the temporal pattern, the normalized envelope of the song signal was computed after signal rectification by squaring and low-pass filtering at 200 Hz (equivalent to a temporal resolution of 2.5 ms). Temporal parameters such as pulse and pause duration were calculated when the envelope crossed or fell below a threshold value at 10 - 15% of the signal envelope. Individual mean values were based on at least two 10 second windows, typically containing around 400 pulses and 2-10 trills each, from two different recordings.

Female preference functions

Female preference was tested using a trackball system similar to that used by Hedwig and Poulet (2004). Females, mounted to a metal rod, were placed on a hollow Styrofoam sphere (diameter: 100 mm, weight: 1.2 to 1.8 g) supported by an air stream between two perpendicularly placed loudspeakers (Piezo Horntweeter PH8) in a wooden box with sound absorbing foam. Each loudspeaker was calibrated with a Bruel and Kjaer 2231 sound level meter and a half-inch condenser microphone (Bruel and Kjaer 4133 relative to 2 x 10^{-5} Pa, fast reading) at the top of the sphere where the female cricket was placed during experiments.

Digitally stored signals were transmitted from a hard disk by a D/A-board (update rate: 100.000 kHz, PCI 6221, National Instruments, Texas) to a digitally controlled attenuator (PA5, Tucker-Davis, Florida), amplified (Raveland, Conrad Electronics) and broadcast through the speakers. The longitudinal and lateral movements of the sphere were recorded by either a single optical sensor (Agilent ADNS-2051) at the bottom of the half-sphere or by two sensors (ADNS-5050, Avago Technologies) with a focussing lens positioned laterally at an angle of 90 degrees.

A silent control was used to monitor baseline activity and a continuous tone was used to control for motivation and selectivity of female responses. At the beginning and the end of each test session, a species-specific, attractive song signal was presented to control for possible changes in phonotactic motivation during a session. For each test signal, the lateral deviation of a female during signal presentation for each of the two speakers was averaged and normalized with respect to the attractive control signal. The resulting phonotactic scores were therefore typically between 0 (no orientation towards the sound signal) and 1 (strong orientation towards the signal), although negative scores (orientation away from the signal) and scores higher than 1 (orientation towards signal stronger compared to control stimulus) were possible.

Test signals and controls were presented at 80 dB sound pressure level (SPL). The responses of 8 to 30 females per species were evaluated for each test pattern (for sample sizes per test see figure captions). All tests included the four control stimuli (silent, continuous tone, and an attractive stimulus at the beginning and end of a test) and eight test stimuli (total duration was 29 minutes per test), and were performed at 24°C-26°C.

Statistical analyses

All statistical analyses were performed in R 2.15.1 (R Core Team 2012). A linear discriminant analysis (MASS package, Venables & Ripley 2002) was used to separate males in multivariate space. Within-species coefficients of variation (CV) and intra-class correlation coefficients were calculated to examine variability of traits within and between species. The intra-class correlation coefficient (ICC) between males of the same species was calculated using the ICCbare function from the ICC package (Wolak *et al.* 2012) and indicates the amount of variation partitioned within rather than between groups (in this case species).

First, we tested on which combination of temporal traits female preferences were dependent and whether this dependence was linear or nonlinear by using a linear regression model including linear and polynomial terms (hypothesis 1). The regression curves were fitted using the lm function from the Stats package and the model with the best fit was selected using Akaike's information Criterion (AIC, Akaike 1973).

Second, we tested whether the general features of song recognition were conserved across species (hypothesis 2). We compared the individual-level peak preferences (most attractive stimulus for each individual) for carrier frequency using Mann-Whitney U tests. The relative weight of preferences for the pulse and the trill timescale and a bivariate preference space constrained by pulse rate and duty cycle were compared between the species. Third, the strength of trait-preference mismatch for pulse rate and duty cycle, carrier frequency, and trill duty cycle was assessed for each species (hypothesis 3). To this end, a population-level peak preference was determined. Instead of simply taking the highest preferred stimulus, we quantified which stimuli were not statistically different from the highest preferred stimulus. This way, peak preference could be calculated as the average of all stimuli that elicit a very high response, rather than by a single stimulus that elicited a slightly higher response than similarly attractive stimuli by chance (and might stochastically bias peak preference away from or closer to the male mean). Pairwise t-tests were performed to compare the response for the highest preferred stimulus with the response for the stimulus with the second highest preference, followed by comparing the response for the latter stimulus with that for the stimulus with third highest preference, and so on, until the P-value associated with the t-test was lower than a Bonferroni-corrected significance level ($\alpha = 0.05$). The peak preference was either the most attractive stimulus or the average of the most attractive stimuli that were not significantly different from each other. Mismatch was then calculated as the mean of the male signal

minus the peak preference, divided by the mean of the male signal means (grand mean). Therefore, mismatch was scaled to allow for comparison between song traits differing in magnitude. Negative mismatch indicated a lower value for the mean of the male signal compared to the preference.

RESULTS

Male song divergence

The majority of song parameters differed among species, but pulse rate and carrier frequency were most divergent (Table 1; Fig. 1d). Pulse rate alone was sufficient to discriminate among species, but the discriminative potential strongly increased when trill rate and carrier frequency were also considered (Fig. 1b,c). The variation in carrier frequency and pulse parameters was mostly partitioned between species (low CVs and high ICCs); the exception being pulse duty cycle, which was similar between species. For trill parameters, more variation was partitioned within species rather than among (higher CVs and lower ICCs); the exception being a high ICC for trill rate and low CV for trill duty cycle (Table 1).

Table 1 Male song variation and the potential for species discrimination. The mean, standard deviation (SD) and coefficient of variation (CV = SD/mean) of the carrier frequency, and of the period, rate, duration, pause duration, and duty cycle at both pulse and trill timescales are shown. Additionally, the intra-class correlation coefficient (ICC) and the coefficients of the first and second linear discriminants (LD1 and LD2, proportion of trace in parentheses) for the parameters included in the linear discriminant analysis are given.

	G. rubens n=40	*G. texensis* n=37	G#14 n=24			
	Mean ± SD, CV	Mean ± SD, CV	Mean ± SD, CV	ICC	LD1 (0.823)	LD2 (0.124)
Frequency (kHz)	4.7 ± 0.2, 0.04	5.2 ± 0.2, 0.04	4.5 ± 0.1, 0.03	0.766	0.654	0.328
Pulse						
period (ms)	21.3 ± 1.6, 0.08	16.9 ± 1.1, 0.06	26.5 ± 1.7, 0.06	0.899		
rate (s^{-1})	47.3 ± 3.6, 0.08	59.3 ± 3.6, 0.06	38.0 ± 2.0, 0.05	0.904	2.205	-1.557
duration (ms)	9.0 ± 1.6, 0.17	7.1 ± 1.0, 0.14	11.9 ± 1.6, 0.13	0.715		
pause (ms)	12.3 ± 2.3, 0.18	9.9 ± 1.3, 0.13	14.6 ± 1.5, 0.10	0.596		
duty cycle	0.43 ± 0.08, 0.19	0.43 ± 0.06, 0.14	0.45 ± 0.05, 0.11	0.012	-0.218	0.337
Trill						
period (s)	4.1 ± 1.8, 0.45	1.1 ± 0.4, 0.38	2.5 ± 0.8, 0.33	0.605		
rate (s^{-1})	0.31 ± 0.17, 0.54	1.08 ± 0.31, 0.28	0.49 ± 0.14, 0.29	0.784	0.277	1.647
duration (s)	3.2 ± 1.6, 0.50	0.7 ± 0.3, 0.46	1.7 ± 0.7, 0.44	0.579		
pause (s)	0.92 ± 0.83, 0.90	0.36 ± 0.20, 0.46	0.79 ± 0.29, 0.37	0.193		
duty cycle	0.76 ± 0.13, 0.17	0.63 ± 0.12, 0.19	0.66 ± 0.07, 0.11	0.298	-0.150	-0.282

Preference functions: carrier frequency

Preferences for carrier frequencies between 2.5 and 7 kHz were tested using stimuli with a constant, attractive temporal pattern. First, we compared the responses to these stimuli between species by examining the shape and peak of the preference functions (hypothesis 2). No species-specific tuning of the preference functions to a particular frequency was found (Fig. 2a). Rather, all three species responded well to signals with carrier frequencies from 4.5 to 6 kHz (Fig. 2a). No significant differences in individual-level peak preference were found between *G. rubens* and *G. texensis* (W_{29} = 208; P = 0.912) and between *G. rubens* and G#14 (W_{12} = 146, P = 0.1738). Peak preference was slightly but significantly higher for G#14 compared to *G. texensis* (5.3 vs. 5.0 kHz; W_{12} = 72, P = 0.0134).

Second, we examined the congruence between the range of variation in the males and the optimum and width of the preference functions (hypothesis 3). The carrier frequency of male songs was restricted to a smaller range and located below the peak of the female preference (Table 1, Fig. 2a). The strength of mismatch between population-level peak preferences and the male signal varied between 0.04 for *G. texensis* and -0.20 for G#14 (Table 3). The divergence in the male signal was not mirrored by female preference (Fig. 2a).

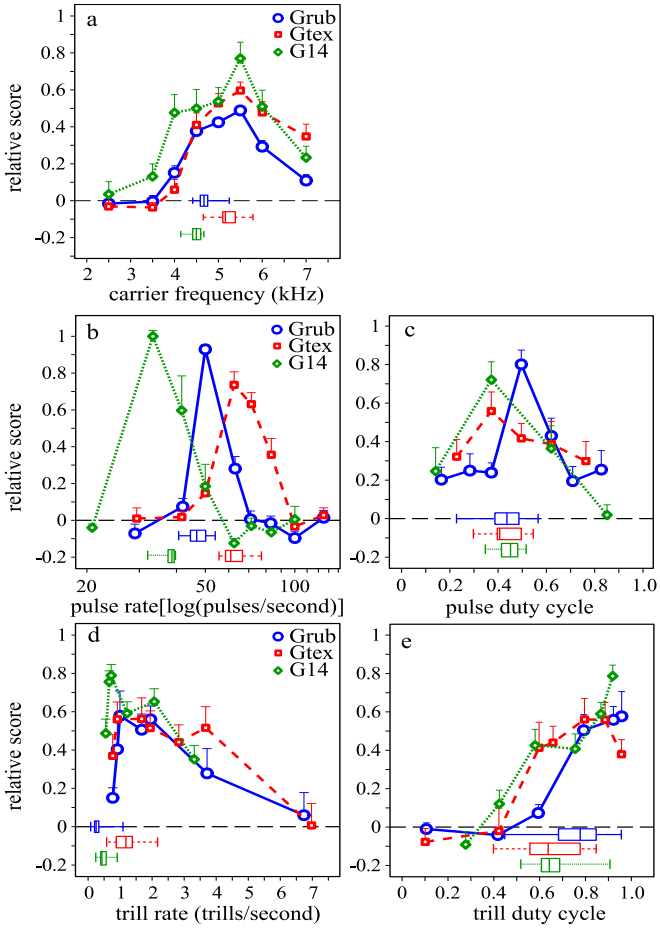

Figure 2 Female preferences for variation of spectral and temporal song traits. The preference functions are based on the average responses to test patterns that vary in the trait shown on the x-axis. Colours are as in Fig. 1. Error bars show the standard error. The black, dashed lines represent a phonotactic score of zero. The boxes (1st and 3th quantile and median) and whiskers (full range) show variation in the male signal. (a) The preference functions for carrier frequency (n = 29 for *G. rubens* and *G. texensis*, and n = 12 for G#14). (b) The preference curves for pulse rate (additional test performed at constant duty cycle = 0.5; n = 82, 31, 8). The curves for pulse duty cycle in (c) result from diagonal transects taken from the pulse profile from upper left to lower right in Fig. 3a-c (n=20, 14, 9) at approximately constant pulse rate (*G. rubens*: 54.5 pps; *G. texensis*: 72.5 pps; G#14: 29.9 pps). The curves in (d) and (e) represent diagonal transects across the trill profile in Fig. 3d-f (n = 16, 14, 14) taken at approximately constant trill duty cycle (*G. rubens*: >0.74; *G. texensis*: >0.72; G#14: >0.65) or trill rate (*G. rubens*: 1.45 tps; *G. texensis*: 1.29 tps; G#14: 1.02 tps), respectively.

Preference functions: bivariate pulse profiles and peak pulse rate preferences

Responses to variation in pulse pattern were tested using bivariate pulse tests. This bivariate preference space is constrained by 16 combinations of pause and pulse durations (Fig. 3a-c). Along a diagonal line across one of the panels in Fig. 3a-c, either the sum of pause and pulse duration is constant (from the top left to the bottom right corner) in which case the pulse duty cycle varies but not the pulse rate, or the ratio of pulse duration to pulse period remains constant (from the bottom left to the top right corner) in which case the pulse rate varies, but not the pulse duty cycle.

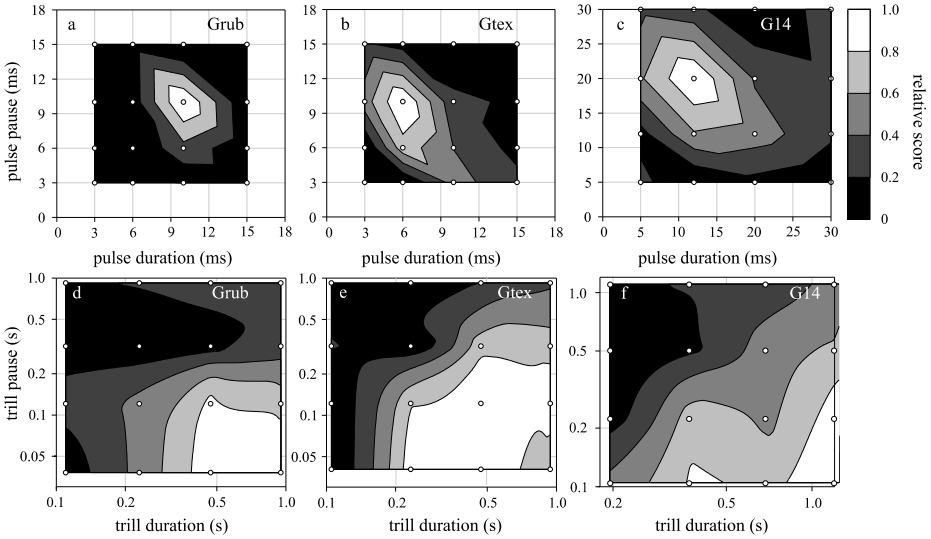

Figure 3 Preference functions for the pulse (a-c) and trill pattern (d-f). Within each panel, pulse and trill rates increase from upper right to lower left; pulse and trill duty cycles increase from upper left to lower right. Note the differences in scale between (a, b) and (c) and between (d, e) and (f). The grayscale indicates the differences in relative scores as per the legend to the right of (c). Scores were normalized with respect to the response to the most attractive stimulus for each species. The contour plots show the preference space obtained using Delaunay triangulation and linear interpolation using the TriScatteredInterp function in MATLAB. The small white circles represent the 16 test points (stimuli). Sample sizes are as in Fig. 2c for the pulse profile plots (top three panels), and as in Fig. 2d, e for the trill profile plots (bottom three panels).

First we examined the features for song recognition to test for the relevant temporal parameters of the pulse pattern (hypothesis 1). All three species had a single most attractive test pattern (Fig. 3a-c). The range of preferred song patterns was constrained by a small range of pulse rates across a wide range of pulse duty cycles (from top left to bottom right) rather than by a small range of pulse duty cycles across a wide range of pulse rates. However, this pulse rate dependency was weaker for *G. rubens* (Fig. 3a). We performed a linear regression on the preference for all 16 test points (dependent variable) shown in Fig. 3a-c. A model including quadratic terms for pulse rate, pulse duty cycle, and their interaction had the best fit according to the AIC score (Table 2). The coefficients of the quadratic terms for both pulse rate and duty cycle were significant (Table S1).

Second, we looked for differences between the species (hypothesis 2). The shape of the preference function was similar between species, but they differed in the range of preferred pulse rates (Fig. 3a-c). In addition to the bivariate pulse profile tests, we also conducted a univariate test varying the pulse rate at a constant pulse duty cycle of 0.5. The purpose of this univariate test was to examine species differences and compare the male signal and female preference at a higher resolution. All three species showed unimodal, sharply tuned preference functions (Fig. 2b). The univariate preference function for pulse duty cycle, which is derived from the bivariate pulse test by taking a cross section along a diagonal with constant pulse rate, was similar between species and centred at low to intermediate values, corresponding to the male signal (Fig. 2c). For pulse rate, the strength of mismatch between population-level peak preference and the male signal (hypothesis 3) varied between -0.06 (*G. rubens*) and -0.16 (*G. texensis*). For pulse duty cycle the strength of mismatch ranged from 0.09 for G#14 to -0.32 for *G. texensis* (Table 3).

Table 2 AIC based model selection for the pulse and trill preference tests (Fig. 3). Four different linear models were tested including (1) only linear terms for pulse/trill rate and pulse/trill duty cycle and their interaction ("linear"), (2) a quadratic term for rate and a linear term for duty cycle and their interaction ["(rate)^2"], (3) a linear term for rate and a quadratic term for duty cycle and their interaction ["(duty cycle)^2"], or (4) only quadratic terms and their interaction ("quadratic"). The lowest AIC score is in bold face type and for the corresponding model the summary statistics are shown to the right.

		Models				Model with lowest AIC		
species	timescale	linear	(rate)^2	(duty cycle)^2	quadratic	R^2	F-statistic	P-value
G. rubens	pulse	204.79	210.18	208.09	**201.19**	0.11	$F_{8,183} = 3.856$	0.0003
	trill	**102.66**	108.90	106.34	115.01	0.17	$F_{3,108} = 7.286$	0.0001
G. texensis	pulse	100.93	94.64	107.63	**75.78**	0.31	$F_{8,103} = 7.117$	< 0.0001
	trill	**141.73**	144.67	149.97	157.23	0.19	$F_{3,156} = 13.19$	< 0.0001
G#14	pulse	227.40	218.63	226.55	**216.33**	0.14	$F_{8,215} = 5.706$	< 0.0001
	trill	**445.85**	457.00	456.10	468.36	0.30	$F_{3,284} = 41.48$	< 0.0001

Preference functions: bivariate trill profiles and linear selection on trill duty cycle

Similar to the analysis of the preference for parameters of the pulse pattern, we conducted two bivariate tests for trill preferences with 16 combinations of trill duration and pause (Fig. 3d-f). Here too we first examined the features for song recognition to test the predictions based on the model for song recognition (hypothesis 1). The observed preference spaces in Fig. 3d-f indicated that females responded well to a range of different trill durations as long as the pause between trills was shorter than the trill duration. This is equivalent to a response to songs with a high trill duty cycle. At very low trill rates, higher duty cycles (bottom left in Fig. 3d-f) also yielded low responses. We compared several models for the relationship between the preference for a stimulus (dependent variable) and different combinations of linear and quadratic terms for trill rate and trill duty cycle (independent variables). We found that a linear term for trill rate and trill duty cycle and an interaction term of the two yielded the lowest AIC (Table 2). However, only the trill duty cycle term was significant (Table S2). The linear model confirms the pattern observed in Fig. 3d-f: trill duty cycle is the main cue for the attractiveness of the long timescale.

To examine the divergence in the preference functions for the trill pattern (hypothesis 2) we compared preferences in bivariate space (Fig. 3d-f) as well as unidimensional representations of these data (Fig. 2d-e).The preference functions in Fig. 2d-e correspond to cross sections across the contour plots in Fig. 3d-f. Females of all three species preferred faster trill rates and higher trill duty cycles than males tended to produce, and the shapes of the preference functions were similar, especially for *G. rubens* and G#14 (Fig. 2d-e, Fig. 3d-f). *Gryllus texensis* females discriminated against the highest trill duty cycles and showed a broader preference function favouring higher trill rates than the other species (Fig. 3e). There was considerable mismatch between the male signal and female preferences for trill duty cycle (ranging from -0.36 for *G. texensis* to -0.2 for G#14, Table 3). However, the range of variation in trill duty cycle in males overlapped with the range of trill duty cycles preferred by the females (Fig. 2e).

Table 3 Trait-preference mismatch. Peak preference (range, mean) was calculated as the average of the stimuli for which we measured significantly higher phonotactic scores than the next highest stimulus. Mismatch was calculated as the mean of the male signal minus the peak preference divided by the mean of the male signal means (grand mean). The male signal means are shown for comparison.

Trait	parameter	*G. rubens*	*G. texensis*	G#14
carrier frequency	peak preference	4.5-5.5, 5.0	4.5-5.5, 5	5.5
(kHz)	male signal (mean ± SD)	4.7 ± 0.2	5.2 ± 0.2	4.5 ± 0.1
	mismatch	-0.06	0.04	-0.21
pulse rate	peak preference	50.0	62.5-71-4, 67.0	31.0-39.0, 35.0
(pulses/s)	male signal (mean ± SD)	47.3 ± 3.6	59.3 ± 3.6	38.0 ± 2.0
	mismatch	-0.06	-0.16	0.06
pulse duty cycle	peak preference	0.50-0.63, 0.56	0.38-0.77, 0.57	0.20-0.63, 0.41
	male signal (mean ± SD)	0.43 ± 0.08	0.43 ± 0.06	0.45 ± 0.05
	mismatch	-0.30	-0.32	0.09
trill duty cycle	peak preference	0.86-0.96, 0.91	0.75-0.96, 0.85	0.65-0.92, 0.79
	male signal (mean ± SD)	0.76 ± 0.13	0.63 ± 0.12	0.66 ± 0.07
	mismatch	-0.22	-0.32	-0.19

Preference functions: integration of pulse and trill timescales

Template matching on the short timescale for features of the pulse pattern should be independent of the overall trill pattern of the song (Clemens & Hennig 2013). The trill pattern on the other hand, should, in the absence of modulation by pulses at conspecific pulse rates, not result in positive phonotaxis (hypothesis 1). To test this prediction for our data, a series of tests were conducted to measure the relative weight of the pulse and trill pattern. Stimuli were presented as trains of pulses (with a constant duty cycle of 0.5) without trill modulation (*i.e.*, continuous trains of pulses with no trill pause, corresponding to a trill duty cycle of 1). At very low pulse rates (*i.e.*, longer pulse durations and pauses) the pulses represented trills without any pulse structure (*i.e.*, pure continuous tones, followed by a pause with an equivalent duration). Any increased response at these low rates would indicate an important role for the trill timescale in song recognition bypassing any template operating on the pulse timescale. None of the species showed a response at low rates corresponding to conspecific trill rates without pulse modulation (Fig. 4) nor did they respond to long, continuous tones as used as a negative control (see Methods). However, the responses to species-specific pulse rates without any trill modulation were as strong as seen in Fig. 2b for all three species.

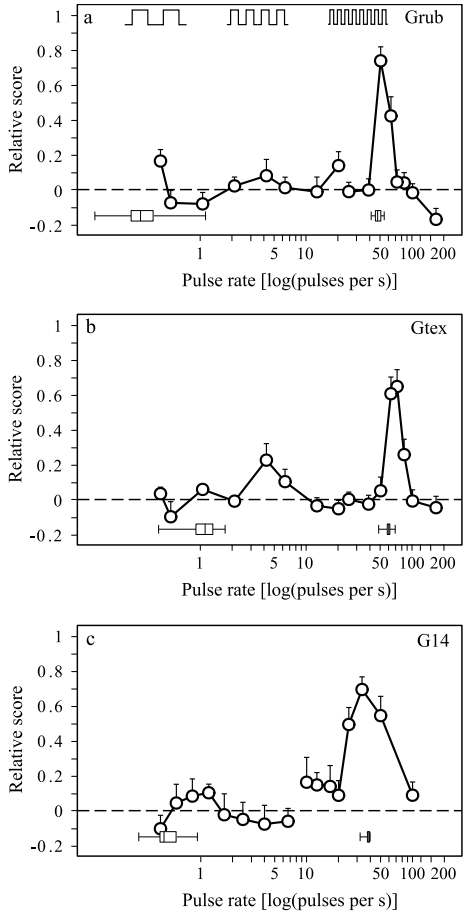

Figure 4 Integration of short and long timescales. The curves show responses to continuous trains of pulses varying in the rate from 0.33 pulses per second to a rate of 167 pulses per second at constant duty cycle of 0.5. Corresponding pulse durations ranged from 1500 ms to 3 ms. In (a) a schematic representation of three test patterns played back to the females is shown (top). Error bars represent standard error, the dashed line shows a phonotactic score of zero, and the box-and-whiskers show the median, 1st and 3rd quantile, and range of male pulse and trill rate [(a): n= 20, (b): n= 16, (c): n = 9].

Song recognition

The results from the previous tests (pulse pattern, trill pattern, and weighing of the timescales) suggest that song recognition was primarily dependent on pulse rate and trill duty cycle and interspecific differences in preference functions were mostly manifested in different peak preferences for pulse rate. To visualize each of the three species in the proper parameter space (*i.e.*, defined by pulse rate and trill duty cycle) the results from both bivariate profile tests (Fig. 3) were combined. Additionally, data from another test with eight unique combinations of trill duty cycle and pulse rate complemented the parameter space for song recognition (Fig. 5). The three species showed divergent preferences for pulse rate with little to no overlap. *Gryllus rubens* responded only to very high trill duty cycles (>0.75), whereas G#14 and *G. texensis* also responded well to lower duty cycles (>0.6). *Gryllus texensis* females also showed a reduced response to the highest duty cycle (~0.96)

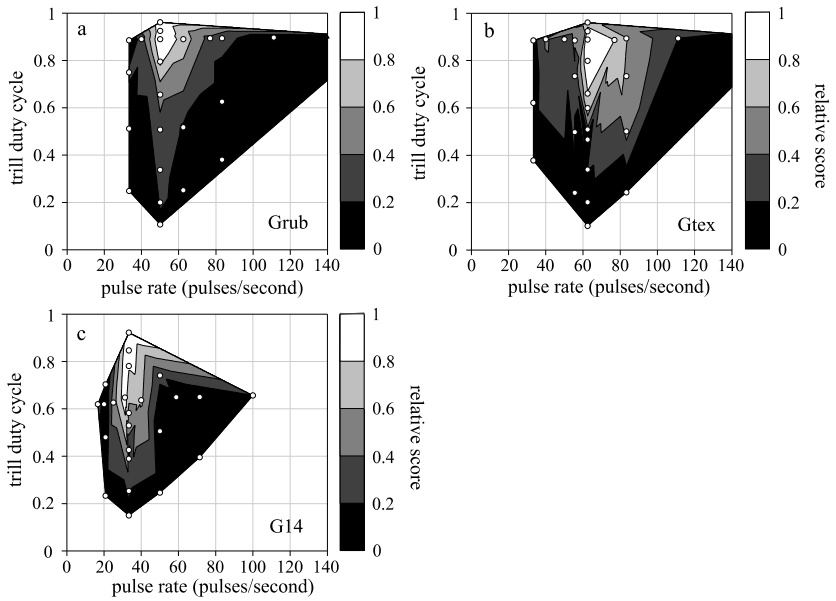

Figure 5 Illustration of the preference space constrained by trill duty cycle and pulse rate. The interpolation was done using the same algorithm as in Fig. 3 and included the test points for both the pulse and trill pattern tests in Fig. 3, as well as an additional test encompassing stimuli with other combinations of pulse rates and trill duty cycle [for the latter test: n=12 (a), n=13 (b), n=8 (c)]. Scores were normalized with respect to the response score for the most attractive stimulus in each species. The test patterns are displayed in the contour plots as white points, with the patterns for the pulse test (Fig. 3a-c) at constant trill duty cycle (*i.e.*, along a horizontal line), the patterns for the trill test (Fig. 3e-f) at constant pulse rate (*i.e.*, along a vertical line) and the patterns from the additional test outside either the horizontal or vertical array of points.

DISCUSSION

Female preference in three trilling *Gryllus* species was primarily constrained by pulse rate and trill duty cycle, the latter of which was largely independent of temporal measures such as trill rate (hypothesis 1). Between species we found a shift in the peak preference for pulse rate, small differences between trill duty cycle preferences, and no divergence in any other preference functions (hypothesis 2). Pulse rate also revealed the strongest degree of congruence between the male signal and female preference; for pulse duty cycle, carrier frequency and trill duty cycle we observed mismatch (hypothesis 3).

Conservation of preferences and underlying mechanisms

Preference for a high trill duty cycle discriminates against males producing songs with short chirps or trills separated by longer pauses, *i.e.* song patterns with a low to intermediate chirp/trill duty cycle. From a computational point of view, the observed selectivity follows the recently proposed, general scheme of song pattern recognition in crickets (Clemens & Hennig 2013). The tympanic ear of crickets constitutes an initial, peripheral filter for the carrier frequency (Gerhardt & Huber 2002) and generates a neuronal representation of the received stimulus. Then, the song pattern is first compared to a short template for evaluation of the pulse pattern. The result of this sensory computation is manifested here as the pulse rate preference of females (Fig. 2b, 3a-c). Second, the output of this sensory filter at the pulse timescale is integrated over a longer time window. The observed preference of females for high trill duty cycles and the rejection of low trill duty cycles (Fig. 2e, Fig. 3e-f) correspond to such an integration process that does not evaluate the specific timing of pulses or trills. Rather, females respond stronger with increasing song energy over time, equivalent to an increase in trill duty cycle (Clemens & Hennig 2013). Mechanistically, the computational components for the evaluation of calling songs with long trills appear to be the most basic and parsimonious required: a simple pulse rate detector combined with a linear preference for high trill duty cycles (Fig. 5a-c). Song recognition by crickets with short chirps would similarly require a detector for the pulse pattern followed by integration of that detector's output over time. However, additional selectivity for intermediate chirp rates and chirp duty cycles (Hennig 2009; Grobe *et al.* 2012; Rothbart & Hennig 2012) needs additional computational components (Hennig *et al.* 2014). Because trills are considered ancestral in Gryllidae (Alexander 1962), the data from the present study suggest that these species' preference functions are also of an ancestral type, although perhaps representing a reversal to an ancestral type within *Gryllus* (Desutter-Grandcolas & Robillard 2003).

Selection manifested in preference functions

Despite differences in the carrier frequency of males, we found broadly-tuned, unimodal preference functions for carrier frequency in each species with overlapping peak preferences (Fig. 2a, Table 3). Preferences for carrier frequency are typically characterized by closed functions (Gerhardt & Huber, 2002) and in crickets likely depend to a large degree on the mechanical properties of the ear (Michelsen & Löhe, 1995; Kostarakos *et al.* 2009). Consequently, divergence in carrier frequency preferences is possibly constrained by covariation in the morphological module in which the tympanum is embedded and largely independent of divergent selection for song recognition. This would explain the observation that closely related, morphologically cryptic species largely share their preferences for carrier frequency despite divergence in preference functions across other axes of song variation.

In contrast to carrier frequency, the pulse rate of the male calling song was strongly divergent between the species and closely matched the interspecific differences in the sharply-tuned preference functions (Fig. 2b,c; Fig. 4a-c). Pulse rate is typically the most divergent trait among species (Gerhardt & Huber 2002) and considered to be of primary importance in song recognition (Gerhardt & Huber 2002; Hennig 2009; Kostarakos & Hedwig 2012; Rothbart & Hennig 2012; Pollack & Kim 2013). In earlier studies with *G. rubens* and *G. texensis*, no evidence for character displacement in pulse rate was found (Gray & Cade 2000; Izzo & Gray, 2004). Therefore, although our data confirm a strong role for species-specific selection on pulse rate, we consider selection for species recognition *per se* unlikely.

Two lines of evidence indicated that evaluation of the trill pattern is different from that of the pulse pattern, corresponding to the observed differences in the extent to which interspecific variation in the male signal tracked variation in female preferences. First, we find no evidence for a significant contribution of trill rate to the variation in preference across test stimuli. Rather, a linear filter favouring higher trill duty cycles was characteristic for all three species (Fig. 2d,e; Fig. 5a-c). Second, the experiment in which the relative weight of the pulse and trill timescales was assessed indicated a strong effect of variation in the pulse but not the trill pattern on the attractiveness of a stimulus. We did find that *G. rubens* responded to a smaller range of trill duty cycles than *G. texensis* and, especially, G#14, corresponding to the variation in the male signal and supported by earlier measurements for *G. texensis* (Gray & Cade 1999a, b). Linear selection for higher chirp/trill duty cycle may be a common property of female preferences among crickets (Gray 1997; Hedrick & Weber 1998; Shaw & Herlihy 2000; Wagner & Reiser 2000), especially for those species in which males produce long trains of pulses in irregularly interrupted trills.

Preference-trait co-evolution in *Gryllus*

One notable feature of our study is that different traits showed different levels of correspondence to female peak preferences. In particular, the strength of mismatch for pulse rate and carrier frequency was relatively low, whereas trill and pulse duty cycle were mismatched with their respective preferences. Additionally, interspecific differences in carrier frequency were clearly not a result of divergent preferences: although the strength of mismatch we found was relatively small, the variation in the males did not track the variation between the preference functions. Several, non- mutually exclusive explanations for variation in trait-preference mismatch are discussed in the following section.

The degree of mismatch may be related to the strength of the female preference, supported by formal analysis in *Enchenopa* treehoppers (Rodriguez *et al.* 2006). We found that pulse rate closely matched the unimodal and finely-tuned preference functions, whereas preferences for other traits were either unimodal but broadly tuned or linear and for these traits we found relatively strong mismatch.

Alternatively, linkage disequilibrium of loci contributing to the signal and preference correlates negatively with the magnitude of mismatch. The genetic architecture of pulse rate and pulse rate preference in *Laupala* crickets suggests common genetic effects, either pleiotropy or close physical linkage of genes (Wiley & Shaw 2010; Wiley *et al.* 2012). Strong genetic non-independence of traits and preferences would lead to coordinated evolution. A quantitative genetic study in *G. texensis* (Gray & Cade 1999b) found that a significant genetic correlation between male song and female preference for trill duration (at constant trill pause, hence varying the trill duty cycle) was due to assortative mating and not physical linkage. This would also facilitate trait-preference matching, albeit with a somewhat higher probability (compared to physical linkage or pleiotropy) of mismatch.

Lastly, other selective forces and evolutionary constraints such as natural selection, phenotypic integration, or low evolvability of some but not other male signal traits can constrain the potential for signal-preference co-evolution. For example, natural or (calling song independent) sexual selection for larger body size may result in lower mean carrier frequencies in the population, as size correlates with the dominant frequency in crickets (Gerhardt & Huber 2002). Phenotypic covariance can similarly constrain the evolutionary liability of traits that are under selection (Lande & Arnold 1983) and for

several cricket species covariance between calling song traits has been demonstrated (Bertram *et al.* 2011, Pitchers *et al.* 2013). A formal analysis of the phenotypic covariance among field cricket calling song traits can provide insight in the evolvability of the calling song. That is, the hypothesis that some calling song traits that do not appear to be under selection by differential mate preferences are likely to have diverged by their correlation with, for example, pulse rate or trill duty cycle can be tested explicitly.

Summary

In conclusion, we found divergence across most song traits in the males of *G. rubens*, *G. texensis*, and G#14, but female preferences were relatively similar between species including the computational algorithm for the fusion of different timescales of the pulse and trill pattern. The major exception was female preference for pulse rate, for which species-specific, closed preference functions closely matching the conspecific male signal were found. Comparing the preferences between different dimensions of the calling song revealed that the pulse pattern and spectral component were characterized by nonlinear selection, whereas linear selection was found for trill duty cycle. The extent of trait-preference (mis)match likely reflects both the strength of the female preferences and effects from other mechanisms such as linkage disequilibrium between and natural selection and selective constraints acting on signal traits and preferences. Overall, our results suggest that relatively small changes to the tuning of female preference functions can promote species divergence, without requiring fundamental changes to preference functions or the neural mechanisms that integrate multivariate stimuli across different timescales.

ACKNOWLEDGEMENTS

We much appreciate support for the trackball system by Hansjürgen Dahmen and the help with experiments by Elisa Becker, Darja Hahn and Vivienne Kremling. The manuscript was strongly improved by comments from Emma Berdan, Jonas Finck, Michael Reichert, and peer review from Mike Ritchie (Editor) and two anonymous reviewers. The performed experiments comply with the "Principles of animal care", publication No. 86-23, revised 1985 of the National Institute of Health, and also with the current laws of Germany. The authors declare no conflict of interest. Funded by: DFG/SFB 618, 'Theoretical Biology', and Leibniz Institute for Evolution and Biodiversity Science.

REFERENCES

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. In: *Breakthroughs in statistics* (B. N. Petrov & F. Csaki, eds), pp. 267–281.

Alexander, R.D. 1962. Evolutionary change in cricket acoustical communication. *Evolution* **16**: 443–467.

Andersson, M. & Simmons, L.W. 2006. Sexual selection and mate choice. *Trends Ecol. Evol.* **21**: 296–302.

Bass, A.H. & McKibben, J.R. 2003. Neural mechanisms and behaviors for acoustic communication in teleost fish. *Prog. Neurobiol.* **69**: 1–26.

Bentsen, C.L., Hunt, J., Jennions, M.D. & Brooks, R. 2006. Complex multivariate sexual selection on male acoustic signaling in a wild population of *Teleogryllus commodus*. *Am. Nat.* **167**: E102–16.

Bertram, S.M., Fitzsimmons, L.P., McAuley, E.M., Rundle, H.D. & Gorelick, R. 2011. Phenotypic covariance structure and its divergence for acoustic mate attraction signals among four cricket species. *Ecol. Evol.* **2**: 181–195.

Blows, M.W., Brooks, R. & Kraft, P.G. 2003. Exploring complex fitness surfaces: multiple ornamentation and polymorphism in male guppies. *Evolution.* **57**: 1622–1630.

Chenoweth, S.F. & Blows, M.W. 2006. Dissecting the complex genetic basis of mate choice. *Nat. Rev. Genet.* **7**: 681–692.

Clemens, J. & Hennig, R.M. 2013. Computational principles underlying the recognition of acoustic signals in insects. *J. Comput. Neurosci.* **35**: 75–85.

Desutter-Grandcolas, L. & Robillard, T. 2003. Phylogeny and the evolution of calling songs in *Gryllus* (Insecta, Orthoptera, Gryllidae). *Zool. Scr.* **32**: 173–183.

Doherty, J.A. (1985) Trade-off phenomena in calling song recognition and phonotaxis in the cricket, *Gryllus bimaculatus* (Orthoptera, Gryllidae). *J. Comp. Physiol.* **156**: 787 - 801

Fisher, H., Aron, A. & Brown, L.L. 2005. Romantic love: an fMRI study of a neural mechanism for mate choice. *J. Comp. Neurol.* **493**: 58–62.

Gerhardt, H.C. & Brooks, R. 2009. Experimental analysis of multivariate female choice in gray treefrogs (Hyla versicolor): evidence for directional and stabilizing selection. *Evolution* **63**: 2504–2512.

Gerhardt, H.C. & Huber, F. 2002. *Acoustic communication in insects and anurans*. The University of Chicago Press, Chicago.

Gold, J.I. & Shadlen, M.N. 2007. The neural basis of decision making. *Annu. Rev. Neurosci.* **30**: 535-574.

Gray, D.A. 1997. Female house crickets, Acheta domesticus, prefer the chirps of large males. *Anim. Behav.* **54**: 1553–1562.

Gray, D.A. 2011. Speciation, divergence, and the origin of *Gryllus* rubens: behavior, morphology, and molecules. *Insects* **2**: 195–209.

Gray, D.A. & Cade, W.H. 1999a. Quantitative genetics of sexual selection in the field cricket, *Gryllus* integer. *Evolution* **53**: 848–854.

Gray, D.A. & Cade, W.H. 1999b. Sex, death and genetic variation: natural and sexual selection on cricket song. *Proc. R. Soc. London, Ser. B Biol. Sci.* **266**: 707–709.

Gray, D.A. & Cade, W.H. 2000. Sexual selection and speciation in field crickets. *Proc. Natl. Acad. Sci.* **97**: 14449–14454.

Greenfield, M.D. 2002. *Signalers and Receivers: Mechanisms and Evolution of Arthropod Communication*. Oxford University Press, New York.

Grobe, B., Rothbart, M.M., Hanschke, A. & Hennig, R.M. 2012. Auditory processing at two timescales by the cricket *Gryllus* bimaculatus. *J. Exp. Biol.* **215**: 1681–1690.
Hebets, E.A. & Papaj, D.R. 2005. Complex signal function: developing a framework of testable hypotheses. *Behav. Ecol. Sociobiol.* **57**: 197–214.
Hedrick, A. V & Weber, T. 1998. Variance in female responses to the fine structure of male song in the field cricket, *Gryllus* integer. *Behav. Ecol.* **9**: 582–591.
Hedwig, B. & Poulet, J.F. 2004. Complex auditory behaviour emerges from simple reactive steering. *Nature* **430**: 781–785.
Hennig, R.M. 2009. Walking in Fourier's space: algorithms for the computation of periodicities in song patterns by the cricket *Gryllus* bimaculatus. *J. Comp. Physiol. A* **195**: 971–987.
Hennig, R.M., Heller, K.-G. & Clemens, J. 2014. Time and timing in the acoustic recognition system of crickets. *Front. Physiol.* **5**.
Higashi, M., Takimoto, G. & Yamamura, N. 1999. Sympatric speciation by sexual selection. *Nature* **402**: 523–526.
Higham, J.P. & Hebets, E.A. 2013. An introduction to multimodal communication. *Behav. Ecol. Sociobiol.* **67**: 1381–1388.
Izzo, A.S. & Gray, D.A. 2004. Cricket song in sympatry: species specificity of song without reproductive character displacement in *Gryllus rubens*. *Ann. Entomol. Soc. Am.* **97**: 831–837.
Kostarakos, K. & Hedwig, B. 2012. Calling song recognition in female crickets: temporal tuning of identified brain neurons matches behavior. *J. Neurosci.* **32**: 9601–9612.
Kostarakos, K., Hennig, M.R. & Römer, H. 2009. Two matched filters and the evolution of mating signals in four species of cricket. *Front. Zool.* **6**: 22.
Lande, R. 1981. Models of speciation by sexual selection on polygenic traits. *Proc. Natl. Acad. Sci.* **78**: 3721–3725.
Lande, R. & Arnold, S.J. 1983. The measurement of selection on correlated characters. *Evolution* **37**: 1210–1226.
Michelsen, A. & Löhe, G. 1995. Tuned directionality in cricket ears. *Nature* **375**: 639.
Oh, K.P. & Shaw, K.L. 2013. Multivariate sexual selection in a rapidly evolving speciation phenotype. *Proc. R. Soc. B Biol. Sci.* **280**: 20130482.
Pitchers, W.R., Brooks, R., Jennions, M.D., Tregenza, T., Dworkin, I. & Hunt, J. 2013. Limited plasticity in the phenotypic variance-covariance matrix for male advertisement calls in the black field cricket, *Teleogryllus commodus*. *J. Evol. Biol.* **26**: 1060–1078.
Pollack, G.S. & Kim, J.S. 2013. Selective phonotaxis to high sound-pulse rate in the cricket *Gryllus* assimilis. *J. Comp. Physiol. A* **199**: 285–293.
Popov, A.V. & Shuvalov, V.F. 1977. Phonotactic behavior of crickets. *J. Comp. Physiol.* **119**: 111 – 126.
Rodríguez, R.L., Hallett, A.C., Kilmer, J.T. & Fowler-Finn, K.D. 2013. Curves as traits: genetic and environmental variation in mate preference functions. *J. Evol. Biol.* **26**: 434–442.
Rodríguez, R.L., Ramaswamy, K. & Cocroft, R.B. 2006. Evidence that female preferences have shaped male signal evolution in a clade of specialized plant-feeding insects. *Proc. R. Soc. B Biol. Sci.* **273**: 2585–2593.
Römer, H., Hedwig, B. & Ott, S.R. 2002. Contralateral inhibition as a sensory bias: the neural basis for a female preference in a synchronously calling bushcricket, *Mecopoda elongata*. *Eur. J. Neurosci.* **15**: 1655–1662.
Rothbart, M.M. & Hennig, R.M. 2012. Calling song signals and temporal preference functions in the cricket *Teleogryllus leo*. *J. Comp. Physiol. A* **198**: 817–825.
Scheuber, H., Jacot, A. & Brinkhof, M.W. 2004. Female preference for multiple condition-dependent components of a sexually selected signal. *Proc. R. Soc. B Biol. Sci.* **271**: 2453–2457.
Shaw, K.L. & Herlihy, D.P. 2000. Acoustic preference functions and song variability in the Hawaiian cricket *Laupala cerasina*. *Proc. R. Soc. B Biol. Sci.* **267**: 577–584.
Venables, W.N. & Ripley, B.D. 2002. *Modern applied statistics with S*. Springer, New York, NY.
Wagner Jr., W.E. & Reiser, M.G. 2000. The importance of calling song and courtship song in female mate choice in the variable field cricket. *Anim. Behav.* **59**: 1219–1226.
Walker, T.J. 2014. Crickets. in Singing Insects of North America. http://entnemdept.ifas.ufl.edu/walker/Buzz/crickets.htm.
Weissman, D.B., Walker, T.J. & Gray, D.A. 2009. The Jamaican field cricket *Gryllus assimilis* and two new sister species (Orthoptera: Gryllidae). *Ann. Entomol. Soc. Am.* **102**: 367–380.
West-Eberhard, M.J. 1983. Sexual selection, social competition, and speciation. *Q. Rev. Biol.* **58**: 155–182.
Wiley, C., Ellison, C.K. & Shaw, K.L. 2012. Widespread genetic linkage of mating signals and preferences in the Hawaiian cricket *Laupala*. *Proc. R. Soc. B Biol. Sci.* **279**: 1203–1209.
Wiley, C. & Shaw, K.L. 2010. Multiple genetic linkages between female preference and male signal in rapidly speciating Hawaiian crickets. *Evolution* **64**: 2238–2245.
Wolak, M.E., Fairbairn, D.J. & Paulsen, Y.R. 2012. Guidelines for estimating repeatability. *Methods Ecol. Evol.* **3**: 129–137.

SUPPORTING INFORMATION

Table S1. Linear model with the response on all 16 test patterns in the pulse test (Fig. 3a-c) as dependent variables and linear and quadratic terms for pulse rate (pps) and pulse duty cycle (pdc). The model shown for each species is the model with the lowest AIC (see Table 2 for comparison of alternative models). Poly(*trait*, 2) refers to the two orthogonal polynomials, the first being the linear and the second being the quadratic. The R-squared (R^2) is the adjusted R-squared. Degrees of freedom (DF) are shown as numerator;denominator. For the terms, the regression coefficient, the standard deviation (SD), the t-value, and the P-value are shown. Terms with a P-value below 0.05 are in bold face type.

G. rubens				
Model	R^2	DF	F-statistic	P-value
lm(pref ~ poly(pps, 2) * poly(pdc, 2), data = rub.pfpul)	0.1068	8;183	3.856	3.215E-04
Term	Coefficient	SD	t-value	P-value
(Intercept)	0.0369	0.0635	0.5810	0.5622
poly(pps, 2)1	-7.5085	1.8225	-4.1200	0.0001
poly(pps, 2)2	-5.6802	1.5163	-3.7460	0.0002
poly(pdc, 2)1	0.3017	0.8124	0.3710	0.7108
poly(pdc, 2)2	-4.8033	1.1694	-4.1080	0.0001
poly(pps, 2)1:poly(pdc, 2)1	-7.2151	20.8404	-0.3460	0.7296
poly(pps, 2)2:poly(pdc, 2)1	-14.1427	18.3563	-0.7700	0.4420
poly(pps, 2)1:poly(pdc, 2)2	-110.3123	28.0504	-3.9330	0.0001
poly(pps, 2)2:poly(pdc, 2)2	-54.9730	19.6388	-2.7990	0.0057
G. texensis				
Model	R^2	DF	F-statistic	P-value
lm(pref ~ poly(pps, 2) * poly(pdc, 2), data = tex.pfpul)	0.306	8;103	7.117	1.84E-07
Term	Coefficient	SD	t-value	P-value
(Intercept)	-0.0087	0.0591	-0.1470	0.8830
poly(pps, 2)1	-7.3295	1.2678	-5.7810	<0.0001
poly(pps, 2)2	-7.9311	1.1587	-6.8450	<0.0001
poly(pdc, 2)1	0.1713	0.5728	0.2990	0.7650
poly(pdc, 2)2	-5.0056	0.8599	-5.8210	<0.0001
poly(pps, 2)1:poly(pdc, 2)1	3.1502	10.7334	0.2930	0.7700
poly(pps, 2)2:poly(pdc, 2)1	5.9407	10.6452	0.5580	0.5780
poly(pps, 2)1:poly(pdc, 2)2	-84.7119	15.2408	-5.5580	<0.0001
poly(pps, 2)2:poly(pdc, 2)2	-64.1314	11.4224	-5.6150	<0.0001
G#14				
Model	R^2	DF	F-statistic	P-value
lm(pref ~ poly(pps, 2) * poly(pdc, 2), data = g14.pfpul)	0.1444	8;215	5.706	1.38E-06
Term	Coefficient	SD	t-value	P-value
(Intercept)	0.1066	0.0568	1.8790	0.0616
poly(pps, 2)1	-5.6937	1.9696	-2.8910	0.0042
poly(pps, 2)2	-6.6114	1.6422	-4.0260	0.0001
poly(pdc, 2)1	-0.0029	0.8762	-0.0030	0.9974
poly(pdc, 2)2	-3.9143	1.1331	-3.4540	0.0007
poly(pps, 2)1:poly(pdc, 2)1	32.0109	27.8527	1.1490	0.2517
poly(pps, 2)2:poly(pdc, 2)1	30.2891	23.6006	1.2830	0.2007
poly(pps, 2)1:poly(pdc, 2)2	-88.4540	32.5122	-2.7210	0.0070
poly(pps, 2)2:poly(pdc, 2)2	-66.0778	22.8936	-2.8860	0.0043

Table S2. Linear model with the response on all 16 test patterns in the trill test (Fig. 3e-f) as dependent variables and linear and quadratic terms for trill rate (tps) and trill duty cycle (tdc). The model shown for each species is the model with the lowest AIC (see Table 2 for comparison of alternative models). The R-squared (R^2) is the adjusted R-squared. Degrees of freedom (DF) are shown as numerator;denominator. For the terms, the regression coefficient, the standard deviation (SD), the t-value, and the P-value are shown. Terms with a P-value below 0.05 are in bold face type.

G. rubens				
Model	R^2	DF	F-statistic	P-value
lm(formula = pref ~ tps * tdc, data = rub.pftrill)	0.1452	3;108	7.286	0.00017
Term	Coefficient	SD	t-value	P-value
(Intercept)	-0.07433	0.14735	-0.504	0.615
tps	-0.04732	0.0879	-0.545	0.5867
tdc	0.5549	0.22465	2.47	0.0151
tps:tdc	0.02069	0.12728	0.163	0.8712
G. texensis				
Model	R^2	DF	F-statistic	P-value
lm(formula = pref ~ tps * tdc, data = tex.pftrill)	0.1869	3;156	13.19	1.02E-07
Term	Coefficient	SD	t-value	P-value
(Intercept)	0.01303	0.12709	0.102	0.9185
tps	-0.04657	0.07569	-0.615	0.53931
tdc	0.59221	0.19565	3.027	0.00289
tps:tdc	0.05258	0.11154	0.471	0.638
G#14				
Model	R^2	DF	F-statistic	P-value
lm(pref ~ tps * tdc, data = g14.pftrill)	0.2973	3;284	41.48	<2.20E-16
Term	Coefficient	SD	t-value	P-value
(Intercept)	0.3083	0.1897	1.625	0.10523
tps	-0.4734	0.1771	-2.674	0.00794
tdc	0.871	0.312	2.791	0.0056
tps:tdc	0.5431	0.2847	1.907	0.05747

CHAPTER 3

PHENOTYPIC VARIATION AND COVARIATION INDICATE HIGH EVOLVABILITY OF ACOUSTIC COMMUNICATION IN CRICKETS

Based on: Blankers, T, Lübke, AK, Hennig RM. 2015. Phenotypic variation and covariation indicate high evolvability of acoustic communication in crickets. *Journal of Evolutionary Biology* **28**: 1656-1669

Abstract. Studying the genetic architecture of sexual traits provides insight into the rate and direction at which traits can respond to selection. Traits associated with few loci and limited genetic and phenotypic constraints tend to evolve at high rates typically observed for secondary sexual characters. Here, we examined the genetic architecture of song traits and female song preferences in the field crickets *Gryllus rubens* and *G. texensis*. Song and preference data were collected from both species and interspecific F_1 and F_2 hybrids. We first analysed phenotypic variation to examine interspecific differentiation and trait distributions in parental and hybrid generations. Then, the relative contribution of additive and additive-dominance variation was estimated. Finally, phenotypic variance-covariance (**P**) matrices were estimated to evaluate the multivariate phenotype available for selection. Song traits and preferences had unimodal trait distributions and hybrid offspring were intermediate with respect to the parents. We uncovered additive and dominance variation in song traits and preferences. For two song traits we found evidence for X-linked inheritance. On one hand, the observed genetic architecture does not suggest rapid divergence, although sex-linkage may have allowed for somewhat higher evolutionary rates. On the other hand, **P** revealed that multivariate variation in song traits aligned with major dimensions in song preferences, suggesting a strong selection response. We also found strong covariance between the main traits that are sexually selected and traits that are not directly selected by females, providing an explanation for the striking multivariate divergence in male calling songs despite limited divergence in female preferences.

INTRODUCTION

The striking diversity in sexual traits in the animal kingdom has long fascinated biologists. Many studies in the past decades have discussed the high evolutionary rates observed for sexual traits (Lande 1981, West-Eberhard 1983, Arnqvist 1998, Wiens 2001, Svensson & Gosden 2007) and the potential contribution of sexual selection to divergence and speciation in the presence of gene flow (Higashi *et al.* 1999, Van Doorn *et al.* 2004).

The rate and direction at which traits respond to selection depend on the genetic architecture (Templeton, 1981, Gavrilets 2003). The theoretical models used to examine the potential for (sympatric) speciation by sexual selection typically assume that the traits under selection are controlled by few loci, which are unlinked and inherit purely additively (Gourbiere 2004). However, empirical studies indicate that some sexual traits have relatively complex genetic architectures with significant genetic constraints on the evolvability of traits and preferences (Chenowith & McGuigan 2010). For many sexual traits and, especially, preferences for those traits, very little is known about the underlying genetics and as a consequence we have limited knowledge of the magnitude and direction of trait change that can result from sexual selection.

The genetic architecture can be studied by distinct quantitative genetic measurements (Hansen 2006, Chenowith & McGuigan 2010). First, speciation is more likely to occur when traits are underlain by few loci, rather than by a polygenic architecture (Gourbiere 2004, Gavrilets & Vose 2007). Courtship traits are commonly associated with a major-effect loci genetic architecture (Arbuthnott 2009), although it has been suggested that variation in qualitative traits is typically explained by few loci and that quantitative traits tend to be polygenic (Ritchie & Phillips 1998, Arbuthnott 2009). Second, evolvability is limited by the amount of additive genetic variance. Dominance and epistasis variation are not uncommon for behavioural phenotypes (Meffert *et al.* 2002) and potentially mask additive genetic variation. Third, sex-linkage of traits is associated with higher evolutionary rates of these traits (Coyne & Orr 1989, Kirkpatrick & Hall 2004) and loci associated with sexual traits have been found linked to sex chromosomes (*e.g.* Wolfenbarger & Wilkinson 2001, Shaw & Lesnick 2009).

Finally, the evolvability of traits also depends on the degree to which traits can vary independently. Traits and preferences are often multivariate or multidimensional (Scheuber *et al.* 2004, Hebets & Papaj 2005, Higham & Hebets 2013) and loci controlling traits in multivariate mating signals can be coupled, for example through genetic linkage or pleiotropic interactions (*e.g.*, Foley *et al.* 2007). Estimating the genetic variance-covariance (**G**) matrix for multivariate signals and preferences can shed light on the rate and direction at which multivariate traits can respond to selection (Blows *et al.* 2003, McGuigan & Blows 2007, Chenowith & McGuigan 2010). Alternatively, the **P** matrix sets the upper limit to the correlated response to selection and can be used as surrogate for G (Cheverud *et al.* 1989, McGuigan & Blows 2007, but see Willis *et al.* 1991). Additionally, analysing the phenotypic variance-covariance matrix (**P**) reveals the multivariate phenotype available for selection (Lande & Arnold 1983).

Here, we examine the patterns of inheritance (hybrid versus parental trait distributions, additive versus additive-dominance variation, X-linked versus autosomal inheritance) and phenotypic covariation of calling song traits and song preferences in field crickets. Field crickets offer an interesting system for studying the genetic architecture of mate choice behaviour. Males produce a calling song, which is a multivariate signal used in long distance mate attraction (Alexander 1962). Most of the differentiation between different species of field crickets is manifested in song and song preference divergence (Otte 1992, Wilkins *et al.* 2013) and closely related species are often strongly behaviourally isolated through acoustic mate choice in the absence of other significant reproductive barriers (Veen *et al.* 2013).

The study species *Gryllus rubens* and *G. texensis* are widely distributed across the gulf states in the south of the US and have a large area of overlap (Gray *et al.* 2008; Gray 2011). Males of both species produce long trills (chirps with > 20 pulses, Alexander 1962) and differ in pulse rate as well as in trill rate, trill duty cycle (the duration of a trill relative to the trill period), and carrier frequency (Gray & Cade 1999, Gray 2011, Blankers *et al.* 2015). Female preferences are strongly divergent for pulse rate (Gray & Cade 2000) and, to a smaller extent, for trill duty cycle (*G. rubens* strongly prefers high trill duty cycles, however *G. texensis* also accepts lower trill duty cycles; Blankers *et al.* 2015). Here, we produced F_1 and F_2 hybrids and examined the segregation of traits and preferences. We then tested whether

genetic variation controlling song traits and preferences was mostly additive or whether additionally dominance and epistatic variation could be uncovered. We discuss our results in the light of previous findings and putative speciation models. Lastly, we estimated the phenotypic variance-covariance matrix to evaluate the multivariate phenotype available for selection as well as potential constraints (*i.e.*, highly correlated trait pairs and strong integration, see Methods) in the response to selection.

MATERIAL AND METHODS

Breeding

Crickets were collected in the field [*G. texensis*: 84 females from Austin (TX), Lancaster (TX), and Round Rock (TX); *G. rubens*: 76 females from Gainesville (FL), Lake City (FL), and Live Oak (FL)] and housed in 19L containers (mean temperature: 25.3 ^{o}C $\pm$ 2.73 SD) with gravel, shelter, and water and food ad libitum. Females, which had already mated in the field before collecting them, were divided into groups of maximally 20 individuals to oviposit in containers with moist vermiculite. All laboratory stocks tested for this study derived from individuals collected during the fall of 2013 and offspring from females from all populations were used in the experiments. Male and female offspring were separated before reaching sexual maturity and acoustically isolated during the experimental period. Male and female *G. rubens* and *G. texensis* were crossed to obtain twelve full-sib F_1 hybrid families in both reciprocal directions and we tested a total of 50 male and 33 female F_1 hybrids. Second generation hybrids were obtained by crossing between 8 and 12 F_1 hybrid males and similar numbers of F_1 hybrid females. All four possible F_2 hybrid cross types (both reciprocal crosses with each of the two F_1 cross types) were generated, yielding 35 males and 37 females which were used in the experiments. To minimize the effect of breeding and developmental conditions, we only used laboratory-bred individuals. We included all parents of the F_1 hybrid families, as well as individuals that were tested and raised simultaneously with the F_2 hybrid generation.

Song recording and analysis

The methods to obtain calling song data have been described in detail in a previous study (Blankers *et al.* 2015). In short, individuals were placed in separate boxes (mean temperature 24.9 ^{o}C +/- 0.98 SD) and recorded in the dark for a 16-24 hour period. For each recording, the duration of individual pulses and trills, and pauses between pulses and trills were measured. From the period (sum of duration and pause of a pulse/trill) and the duration of a pulse/trill, the rate (inverse of the period) and duty cycle (duration divided by the period) were calculated. Individual mean values used in the statistical analyses described below were based on at least two 10 second windows (typically containing around 400 pulses and 2-10 trills each) from separate calling bouts. These calling bouts were typically recorded on the same day, but repeated measurements were also collected from different days, which were then maximally 6 days apart.

Female preference testing

Female preference was tested using a trackball system similar to that used by Hedwig & Poulet (2004). The methods have been described in detail elsewhere (Blankers *et al.* 2015). In short, females were mounted to a metal rod and placed on a sphere supported by an air stream between two perpendicularly placed loudspeakers. The trackball system was housed in a closed wooden box with sound absorbing foam to minimize the influence from visual stimuli, internal echoes and external acoustic stimuli. Optical sensors recorded the longitudinal and lateral movements of the sphere relative to the loudspeaker from which the stimulus was broadcast.

A silent control and a continuous tone control were used to control for baseline activity and selectivity of responding females. At the beginning and the end of each test session, a species-specific, attractive song signal was presented to control for possible changes in phonotactic motivation during a session. The lateral movement of a female during signal presentation was averaged between the consecutive playbacks from the two speakers and normalized with respect to the response to the attractive control signal.

Preference functions were measured for three components of the calling song, which are known to show the largest interspecific variation in the male song (Blankers *et al.* 2015). A pulse rate test varied the pulse duration and pause duration simultaneously (keeping the pulse duty cycle constant at 0.5). Eight stimuli ranging from 29.4 pulses/s (well below the lowest male *G. rubens* pulse rate), to 125 pulses/s (well above the highest male *G. texensis* pulse rate) were played back at constant carrier frequency (5.0 kHz) and with a constant trill pattern (trill rate = 1.0, trill duty cycle = 0.9). A test for carrier frequency measured responses of females across 5 stimuli (3.5, 4.5, 5.0, 5.5, and 7.0 kHz, respectively) with an attractive, species specific pulse pattern (pulse duty cycle 0.5 for all lines, pulse rate 50.0 for *G. rubens*, 71.4 for *G. texensis*, 62.5 for F_1 hybrids and, depending on the pulse rate preference, any of the former for an F_2 hybrid). A third test alternated between stimuli with varying trill duration and pause and included four combinations of short (~130 ms) and long (~530 ms) trill duration and a short (~50 ms) and long (~390 ms) pause. The exact duration of each trill or pause depended on the pulse rate (corresponding to the preferred stimulus in the pulse rate test) but variation between tests was minimal (ca. 10 ms differences). The carrier frequency was kept constant at 5.0 kHz. All tests were performed at a temperature between 24°C and 26°C and test patterns were played back in random order. Repeated measurements were obtained only for the pulse rate test, and were measured always on different days, less than a week apart.

Quantifying female preference

Female preference functions were compared between species and hybrid cross lines using a function-valued trait approach (Meyer & Kirkpatrick 2005, see also below). In addition, for meaningful application of some of the quantitative genetic methods we sought to obtain a single measure describing the dominant interspecific differences between the preference functions. Therefore, we fitted a linear discriminant function ('lda' function in the R-package 'MASS'; Venables & Ripley 2002), where the parental lines, *G. rubens* and *G. texensis*, were used as the grouping factor to separate between the responses to the stimuli. We then predicted the scores along the discriminant axis (LD scores) for the parental lines and F_1 and F_2 hybrid cross lines. This approach accurately describes the linear combinations of test stimuli that separate species and is repeatable between females (see Results). Further details on the use of this method for quantifying female preferences are found in the Supporting Information.

Trait and preference divergence

All statistical analyses were performed in R version 3.1.1 (R Core Team, 2014). First, to test for the repeatability of traits, we used the MCMC approach implemented in the rptR package (Nakagawa & Schielzeth 2010) with default priors. Using the methods described below, we then tested whether parental lines were significantly different from each other, whether F_1 and F_2 hybrid lines were intermediate to and significantly different from the two parental lines, and whether the two reciprocal F_1 lines and four F_2 lines had overlapping distributions.

For the males, a MANOVA was used, followed by post-hoc two-sample t-tests for pairwise comparisons of the 5 song parameters (carrier frequency, pulse rate, pulse duty cycle, trill rate, and trill duty cycle) between the lines. For the female preferences, we first compared the preference functions (Fig. S1) between *G. rubens* and *G. texensis* using random effects mixed models. The model included a linear and quadratic fixed effect term for the test stimulus and for species (2-level factor) as well as the interaction between stimulus and species, and the individual identification as a random effect. A significant effect of the interaction between the quadratic term for the test stimulus and the species term indicated a significant difference between the preference functions. We then decomposed the preference function. Mann-Whitney U tests were used to compare the peak preferences, defined as the stimulus yielding the highest response for each individual, for pulse rate and carrier frequency (reciprocal F_1 hybrid crosses and F_2 cross types were also compared separately). An ANOVA with post-hoc pairwise t-tests was used to test for differences in preference strength (or discrimination strength). The strength of discrimination was calculated following Gray & Cade (1999), as the number of standard

deviations between the average preference across all tested stimuli and the score for the highest preferred stimulus.

Joint-scaling test for additive versus dominance variation

To test for the contribution of additive and additive-dominance variation we used the joint-scaling method (Cavalli 1952, Mather & Jinks 1971). The observed trait means, $\boldsymbol{z}$, are compared to predicted trait means, $\hat{\boldsymbol{z}}$, estimated from the matrix multiplication

$$\hat{\boldsymbol{z}} = \mathbf{M}\hat{\boldsymbol{a}} \qquad \text{(Eq. 1)}$$

where $\mathbf{M}$ is the matrix, $\mathbf{M_A}$, with the coefficients of additive variation

$$\mathbf{M_A} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}$$

or the matrix, $\mathbf{M_{AD}}$, with the coefficients of additive and dominance variation

$$\mathbf{M_{AD}} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

and $\hat{\boldsymbol{a}}$ is the vector of parameter estimates $(\mu, \quad \alpha, \quad \delta)$ for the trait mean among all lines, and the additive genetic and dominance effects, respectively. The coefficients in the matrices $\mathbf{M_A}$ and $\mathbf{M_{AD}}$ represent the standardized coefficients for the global mean (μ, equal for all lines), the deviation due to additive effects (of magnitude α, positive for one of the parental lines, negative for the other parental line, and intermediate for the hybrids), and the dominance deviation (of magnitude δ, no heterozygotes in the parental lines, exclusively heterozygotes in the F_1 and intermediate amounts of heterozygotes in the F_2 generation). The vector $\hat{\boldsymbol{a}}$ is obtained from the weighted least squares regression

$$\hat{\boldsymbol{a}} = (\mathbf{M}^{\mathrm{T}}\mathbf{M}\,\mathbf{V}^{-1})^{-1}\mathbf{M}^{\mathrm{T}}\boldsymbol{z}\,\mathbf{V}^{-1} \qquad \text{(Eq. 2)}$$

where $\boldsymbol{z}$ is the vector of observed trait means and $\mathbf{V}^{-1}$ is the inverse of the matrix with diagonal weights (square of the standard error of the observed trait means),

$$\mathbf{V} = \begin{bmatrix} [SE(\mu_{P1})]^2 & 0 & 0 & 0 \\ 0 & [SE(\mu_{P2})]^2 & 0 & 0 \\ 0 & 0 & [SE(\mu_{F1})]^2 & 0 \\ 0 & 0 & 0 & [SE(\mu_{F2})]^2 \end{bmatrix}.$$

The predicted trait means, $\hat{\boldsymbol{z}}$, can be compared with the observed trait means, $\boldsymbol{z}$, using the χ^2 test for goodness of fit

$$\chi^2 = \sum_{i=1}^{l} \frac{(z_i - \hat{z}_i)^2}{SE(z_i)^2} \qquad \text{(Eq. 3)}$$

where l is the number of lines (generations) and z_i and $\hat{z}_i$ are the observed and predicted trait mean of the *i*th line. A significant χ^2 statistic indicates that the lower model (*i.e.*, additive-only) has insufficient explanatory power and that a higher-order model (*i.e.*, additive-dominance) should be considered. We applied this method to the four male song traits for which we found interspecific differences (see Results), as well as for female preference for carrier frequency, pulse rate and the trill pattern.

P matrix analysis

Individual level **P** matrices were retained from Bayesian linear mixed models (package 'MCMCglmm', Hadfield 2010). Prior to the model fit, the data were standardized, for each species separately, to have zero mean and unit variance. The model fitted to the male song data included carrier frequency, pulse rate and duty cycle, and trill rate and duty cycle as response variables, and individual as the random effect. The rationale for using the rate and duty cycle of pulse and trill rather than pause and duration, is that rate and duty cycle are the critical parameters under sexual selection (Blankers *et al.* 2015). For the female preference data, linear mixed models were fitted for the LD scores for pulse rate, carrier frequency, and trill pattern preference functions, for the peak preferences of those functions, and for the preference strength (D).

For all the models (male and female data) parameter expanded, inverse-Wishart priors (Gelman 2006, Hadfield 2012) were used. All models were ran for 500,000 iterations (burn-in = 50,000) and the trace and posterior distributions as well as the autocorrelation of the fit were evaluated to assess model fit. The models were tolerant to different prior distribution as a flat, REML-like prior or a gamma distributed prior gave similar results (not shown).

The posterior mode of the variance-covariance (vcv) matrix and the correlation matrix were retained for each species and cross line independently to evaluate the pairwise correlations between traits. Eigenanalysis was used to reveal the orthogonal dimensions of linear combinations of traits (eigenvectors) and the total amount of phenotypic variation represented by each of those dimensions (eigenvalues) for each of the four matrices separately. Additionally, the magnitude of phenotypic integration was calculated as the relative variance of the eigenvalues of the correlation matrix (Pavlicev *et al.* 2009).

The pairwise correlations in the **P** matrix are informative about the extent to which traits tend to covary within an individual. However, insight in the eigenvectors and eigenvalues of the matrix is critical for the understanding of the constraints in selection response. The trait loadings on the eigenvectors provide information about which traits vary in the same dimension and which traits vary in independent (orthogonal) dimensions. The variance in the eigenvalues (*i.e.*, magnitude of integration) indicate to what extent the leading eigenvector(s) constrains the variation in orthogonal directions (*i.e.*, in the direction of other eigenvectors). Low variance among eigenvalues (*i.e.*, weak phenotypic integration) would result in a near spherical covariance structure that can respond to selection acting in any direction. On the other extreme, strong integration would be represented by a strongly elliptical covariance structure, which can respond to selection in the direction of the first eigenvector, but not in the direction of any of the remaining eigenvectors (see examples in Steppan *et al.* 2002). Therefore, the pairwise correlations, the eigenvectors and eigenvalues, and the variance among eigenvalues are all informative about the potential multivariate response to selection.

To estimate phenotypic integration, the correlation matrix was calculated for each of 1000 random samples from the posterior distribution and the variance of the eigenvalues associated with the 5 (for males) or 3 (for females) eigenvectors (corresponding to the 5 song traits and 3 preference traits, respectively) was calculated. The eigenvalue variance was scaled to the maximum eigenvalue variance, N-1, where N is the number of variables, returning the relative eigenvalue variance between 0 (random, uncorrelated matrix) and 1 (all except the first eigenvalue are zero, indicating maximal integration). We repeated this routine 1000 times to calculate the mean and 95% confidence intervals.

RESULTS

Trait and preference divergence

The song traits were generally repeatable, although considerably lower repeatability was found for pulse duty cycle and trill duty cycle compared to pulse rate, trill rate, and carrier frequency (Table 1). We found unimodal distributions for all traits in the parental lines and intermediate values for F_1 and F_2 hybrids (Fig. 1, Table 1). Variation in the male calling song depended significantly on the lines (MANOVA with line as the independent variable [5 levels: *G. rubens*, *G. texensis*, two reciprocal F_1 lines and combined F_2 lines]: Wilk's $\lambda = 0.1306$, $F_{20,638} = 27.02$, $P < 0.0001$). With the exception of pulse duty

cycle (t_{43}= -0.3435; P = 0.7321) song traits were significantly different between *G. rubens* and G. texenis (Table S1).

Table 1 Means ± standard deviation and repeatability (R) with lower and upper threshold of the 95% confidence interval for male song traits. The sample sizes per line are in parentheses; note that for the combined F_1 hybrid lines this number refers to the number of families. "rubtex" F_1 hybrids result from a cross between a *G. rubens* female and a *G. texensis* male and the reciprocal cross is referred to as "texrub".

♂ signal		*G. rubens* (73)	*G. texensis* (44)	rubtex (22)	texrub (28)	F_1 hybrid (12*)	F_2 hybrid (35)
carrier frequency [kHz]	mean (±SD)	4.73 (0.27)	5.18 (0.22)	4.82 (0.27)	5.01 (0.17)	4.97 (0.17)	4.86 (0.29)
	R (lower, upper)	0.76 (0,74,0.82)	0.86 (0.83,0.91)	-	-	0.87 (0.83,0.90)	0.79 (0.75,0.80)
pulse rate [pulses/s]	mean (±SD)	45.34 (3.86)	66.88 (5.40)	55.27 (3.96)	61.96 (2.79)	59.01 (2.79)	56.26 (4.813)
	R (lower, upper)	0.52 (0.38,0.60)	0.87 (0.81,0.90)	-	-	0.76 (0.71,0.81)	0.84 (0.80,0.86)
pulse duty cycle	mean (±SD)	0.43 (0.06)	0.44 (0.08)	0.46 (0.06)	0.48 (0.08)	0.47 (0.07)	0.40 (0.07)
	R (lower, upper)	0.20 (0.07,0.36)	0.04 (0.00,0.29)	-	-	0.05 (0.00,0.20)	0.41 (0.32,0.56)
trill rate [trills/s]	mean (±SD)	0.28 (0.15)	1.28 (0.39)	0.59 (0.25)	0.68 (0.35)	0.64 (0.24)	0.63 (0.24)
	R (lower, upper)	0.74 (0.70,0.80)	0.46 (0.23,0.65)	-	-	0.44 (0.29,0.55)	0.73 (0.67,0.76)
trill duty cycle	mean (±SD)	0.78 (0.10)	0.62 (0.11)	0.71 (0.10)	0.73 (0.14)	0.72 (0.10)	0.71 (0.12)
	R (lower, upper)	0.08 (0.00,0.25)	0.24 (0.01,0.43)	-	-	0.15 (0.02,0.31)	0.80 (0.76,0.82)

* number of families

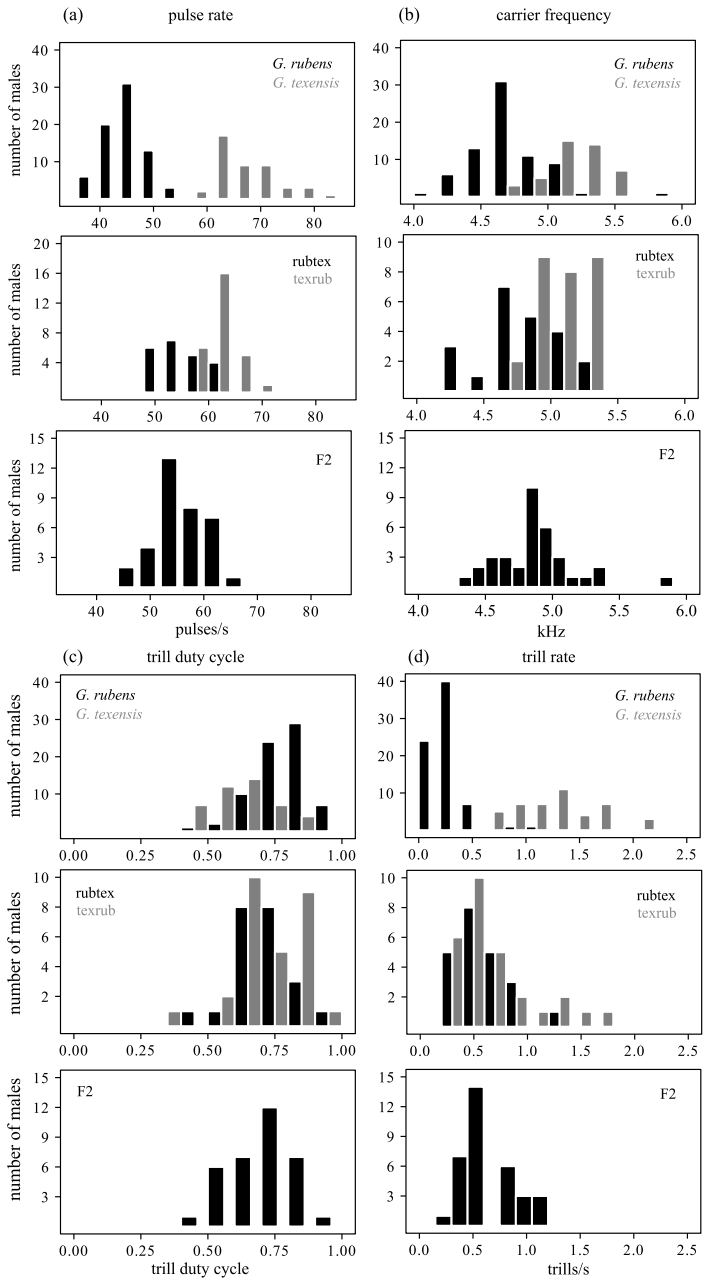

Figure 1 Distribution of male song traits. The distribution of pulse rate (a), carrier frequency (b), trill duty cycle (c), and trill rate (d) are given for *G. rubens* (n = 73) and *G. texensis* (n = 44) in the top panels, for the reciprocal F_1 hybrid crosses (rubtex n = 22, texrub n = 28) in the middle panels, and for the F_2 hybrids (n = 35) in the bottom panels. "rubtex" F_1 hybrids result from a cross between a *G. rubens* female and a *G. texensis* male and the reciprocal cross is referred to as "texrub".

The reciprocal F_1 hybrid crosses differed significantly in pulse rate (*G. rubens* dam: 55.27, *G. texensis* dam: 61.96; $t_{22} = -6.72$; $P < 0.0001$) and carrier frequency (*G. rubens* dam: 4.82, *G. texensis* dam: 5.01; $t_{22} = -3.9146$; $P = 0.0004$) and the distributions of both traits within the F_1 generation were bimodal (Fig. 1a,b). No differences were found between the F_2 cross types (Table S1).

For female preference, significant repeatability was found within each of the cross lines for peak preference, strength of discrimination, and the LD scores for the pulse rate test (Table 2). Female preference functions for pulse rate differed strongly between *G. rubens* and *G. texensis* (LME model interaction term [stimulus]^2 * Species: $F_{2,494} = 11.60$, $P < 0.0001$, Fig. S1). The preference functions did not differ between the two F_1 or four F_2 cross types (LME model interaction term [stimulus]^2 * Cross; F_1: $F_{6,298} = 1.29$, $P = 0.2767$; F_2: $F_{6,443} = 0.99$, $P = 0.4333$). Comparing the peak preference for pulse rate indicated strong significant differences between all lines except between F_1 and F_2 hybrids ($U_{32} = 512.5$, $P = 0.2219$), and between reciprocal F_1 hybrid crosses ($U_{13} = 123.5$, $P = 0.7221$; Table S1). Strength of discrimination was marginally significantly dependent on the line (ANOVA with cross line as the independent variable [5 levels: *G. rubens*, *G. texensis*, two reciprocal F_1 lines and the combined F_2 lines]: $F_{3,166} = 2.23$, $P = 0.0868$). The strength of discrimination, D, the peak preferences and the LD scores for the pulse rate test were intermediate for the F_1 and F_2 hybrids with respect to the parental lines (Table 2, Fig. 2).

Preference functions for the trill pattern were significantly different between *G. rubens* and *G. texensis* ($F_{2,110} = 7.18$, $P = 0.0012$), but preference functions for carrier frequency were not ($F_{2,204} = 0.71$, $P = 0.4952$; Fig. S1). We failed to find differences between species or cross types when comparing peak preferences for carrier frequency or the trill pattern (Table 2; Table S1). Strength of discrimination was also not significantly associated with cross line for either preference function (carrier frequency: $F_{3,94} = 1.20$, $P = 0.3155$, trill pattern: $F_{3,78} = 1.41$, $P = 0.2401$).

Table 2 Means ± standard deviation and repeatability (R) with lower and upper threshold of the 95% confidence interval for female preference. The sample sizes per line are in parentheses; note that for the combined F_1 hybrid lines this number refers to the number of families.

♀ preference		*G. rubens* (43)	*G. texensis* (26)	rubtex (14)	texrub (19)	F_1 hybrid (7*)	F_2 hybrid (37)
Pulse rate peak preference	mean (±SD)	51.95 (5.06)	69.97 (8.49)	63.00 (8.14)	62.21 (7.53)	62.55 (7.68)	59.84 (8.93)
	median	50.00	66.95	62.50	62.50	62.50	62.50
	R (lower, upper)	0.99 (0.94,1)	0.99 (0.96,1)	-	-	0.04 (0,0.87)	0.55 (0.16,0.81)
Frequency peak preference	mean (±SD)	5.05 (0.45)	5.22 (0.69)	5.21 (0.69)	5.33 (0.58)	5.27 (0.63)	5.63 (0.77)
	median	5.0	5.0	5.0	5.0	5.0	5.5
Trill pattern peak preference	median	3	3	3	3	3	3
LD1 pulse rate	mean (±SD)	-4.54 (0.72)	4.54 (1.43)	0.25 (3.74)	0.04 (3.73)	0.13 (3.68)	-1.27 (4.06)
	R (lower, upper)	0.08 (0.02,0.72)	0.97 (0.14,0.99)	-	-	0.01 (0.00,0.70)	0.76 (0.49,0.89)
LD1 frequency	mean (±SD)	-0.39 (1.00)	0.39 (0.99)	0.38 (0.76)	0.76 (1.01)	0.57 (0.90)	-0.02 (1.12)
LD1 trill pattern	mean (±SD)	-0.63 (0.96)	0.63 (1.06)	0.76 (1.32)	-0.00 (0.85)	0.42 (1.18)	-0.35 (1.12)
Discrimination (D) pulse rate	mean (±SD)	1.85 (0.36)	1.71 (0.15)	1.63 (0.05)	1.83 (0.17)	1.75 (0.11)	1.77 (0.13)
	R (lower, upper)	0.42 (0.14,0.76)	0.79 (0.29,0.95)	-	-	0.50 (0.17,0.77)	0.25 (0.13,0.58)
Discrimination (D) frequency	mean (±SD)	1.23 (0.30)	1.34 (0.23)	1.22 (0.26)	1.15 (0.27)	1.19 (0.26)	1.25 (0.27)
Discrimination (D) trill pattern	mean (±SD)	1.24 (0.22)	1.05 (0.25)	1.15 (0.23)	1.13 (0.29)	1.14 (0.25)	1.14 (0.23)

* Number of families

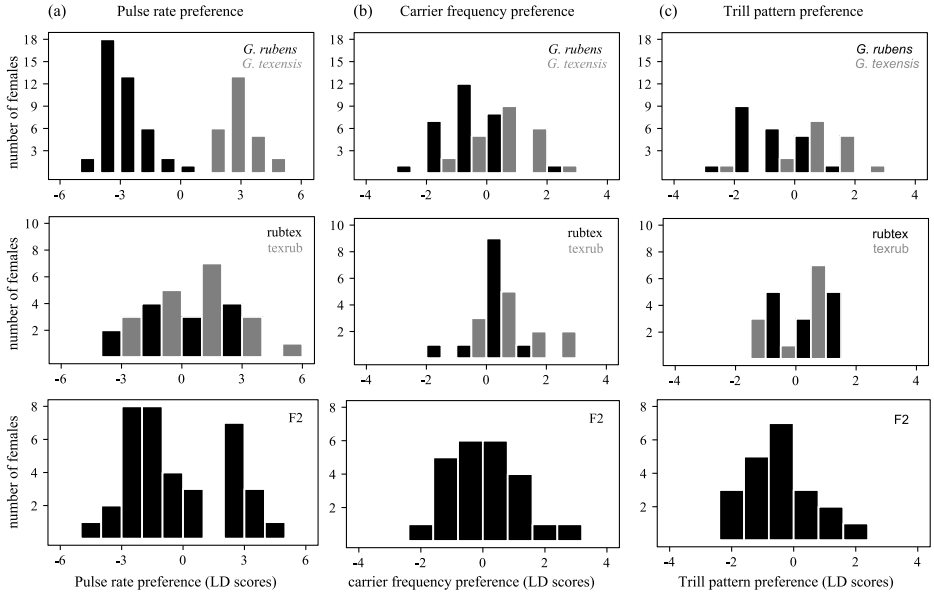

Figure 2 Distribution of female preferences. Preference is measured as the score on the first linear discriminant (LD1) for the pulse rate (a), carrier frequency (b), and trill tests (c). Upper panels: *G. rubens* (n = 43) and *G. texensis* (n = 26); Middle panels: the reciprocal F_1 hybrid lines (n = 14, n = 19, respectively). Bottom panels: the F_2 hybrid line (n = 37).

Joint-scaling test for additive versus dominance variation

The joint-scaling method suggested dominance effects for pulse rate and trill rate of the male song following the poor fit of the additive-only model (pulse rate: $\chi^2_2 = 25.73$, $P < 0.0001$; trill rate: $\chi^2_2 = 12.78$, $P = 0.0008$). The coefficient for dominance effects was, however, small compared to the coefficient of additive genetic variation (Table 3). For other male traits, *i.e.* carrier frequency and trill duty cycle, the additive-only model explained the observed variation in trait means much better, although the observed values differed marginally significantly from the expected values for carrier frequency (Table 3).

The estimated trait means under the additive-only model differed marginally significant from the observed trait means for pulse rate preference ($\chi^2_2 = 3.53$, $P = 0.0856$). Strong significant differences were found for the additive-only model for carrier frequency preference ($\chi^2_2 = 9.48$, $P = 0.0044$) and trill pattern preference ($\chi^2_2 = 7.79$, $P = 0.0102$) and also for the additive-dominance model for trill pattern preference ($\chi^2_1 = 6.47$, $P = 0.0062$; Table 3).

Table 3 Joint-scaling method to estimate additive and additive-dominance effects (Eq. 1 – 3). The estimated grand mean of the traits, μ, the additive genetic and dominance effects (α and δ, respectively), and the χ^2 value are shown for carrier frequency, pulse rate, trill rate, and trill duty cycle of the male signal and female preference for pulse rate, carrier frequency, and the trill pattern (LD scores). Degrees of freedom (DF) for the χ^2 test of the null hypothesis that the model explains the data sufficiently are given as the number of model parameters minus the number of lines. Values of χ^2 below the $P = 0.05$ threshold are in bold face type. Trends ($P < 0.1$) are indicated with an asterisk.

	Carrier frequency	Pulse rate	Trill rate	Trill duty cycle	Pulse rate preference	Frequency preference	Trill preference
Additive							
μ	4.95	57.61	0.69	0.71	-0.05	0.17	-0.00
α	0.22	11.56	0.43	0.08	4.50	0.43	0.63
χ^2 (2 DF)	4.10 *	**25.73 ‡**	**12.87 ‡**	1.34	3.53 *	**9.48‡**	**7.79†**
Dominance							
μ	4.95	57.40	0.69	0.71	-0.20	0.22	0.04
α	0.22	10.67	0.49	0.08	4.52	0.38	0.60
δ	0.01	1.48	-0.07	0.01	-0.17	0.27	0.13
χ^2 (1 DF)	**3.92 †**	2.24 *	2.95 *	0,00	3.22 *	2.12*	**6.47‡**

* $P < 0.1$, † $P < 0.05$, ‡ $P < 0.01$

P matrix analysis

We estimated the individual-level **P** matrices for *G. rubens*, *G. texensis*, and F_1 and F_2 hybrids. Among male song traits, several parameters showed strong correlations in the parental lines (Fig. 3a-b), most notably trill duty cycle and trill rate (strong, negative in both species) and pulse and trill rate (strong, positive in *G. texensis* but not in *G. rubens*). In the F_1 and F_2 hybrid lines, strong, positive correlations were observed among carrier frequency, pulse rate, and trill duty cycle (Fig. 3c-d). Phenotypic integration (relative variance of the eigenvalues, $Var_{rel}[\lambda]$) was much stronger in *G. texensis* ($Var_{rel}[\lambda]$= 0.21) compared to *G. rubens* ($Var_{rel}[\lambda]$= 0.08, Table 4a) and not significantly lower among F_1 ($Var_{rel}[\lambda]$= 0.18) or F_2 ($Var_{rel}[\lambda]$= 0.17) hybrids compared to the parental lines (Table 4a). The eigenvectors of **P** indicate that pulse rate, carrier frequency, and trill duty cycle can vary independently; variation in the trill rate is coupled to variation either in carrier frequency (*G. rubens*) or pulse rate (*G. texensis*; Table S3).

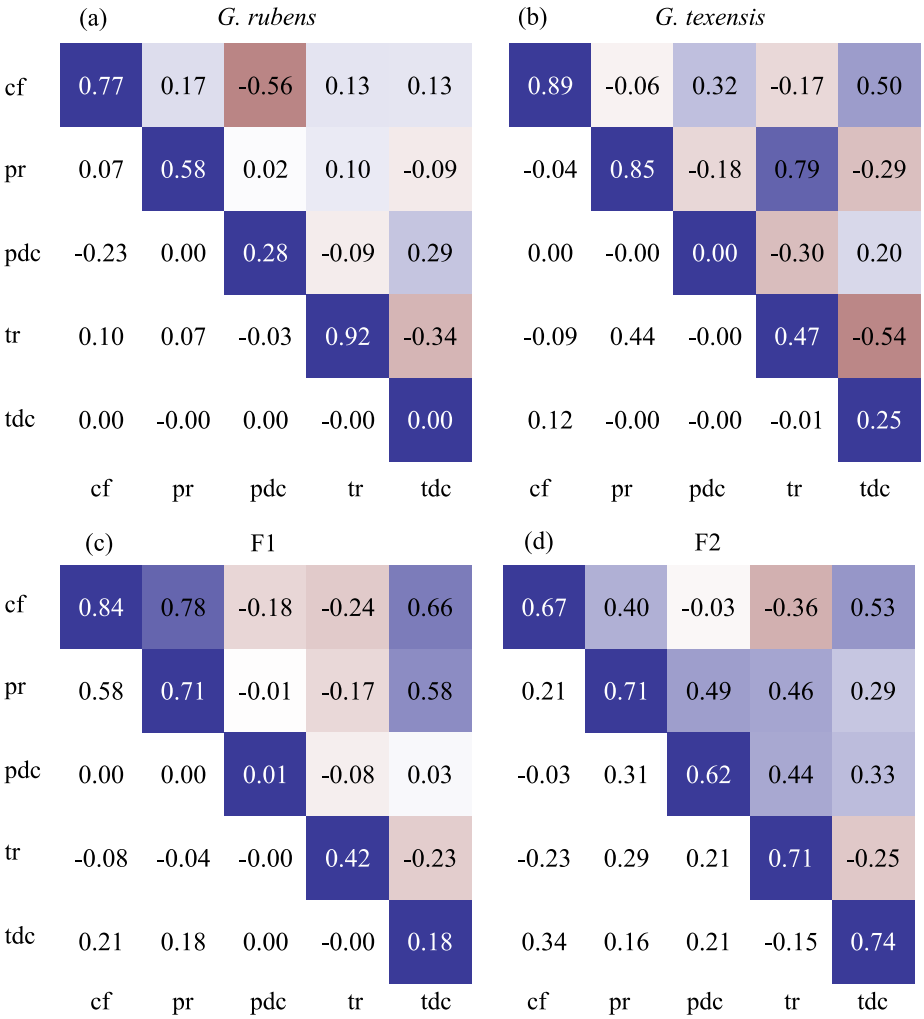

Figure 3 **P** matrices for the male song. The correlation among carrier frequency (cf), pulse rate (pr), pulse duty cycle (pdc), trill rate (tr), and trill duty cycle (tdc) are given for *G. rubens* (a), *G. texensis* (b), and the F_1 (c) and F_2 (d) hybrids. The intensity of the colour of the heat plots indicates the strength of the positive (blue) or negative (red) correlation. The diagonals show the variances of the standardized variables (in white). Correlations are shown above and covariances below the diagonal.

Table 4 Magnitude of phenotypic integration. The mean and lower and upper boundaries of the 95% confidence intervals for the relative variance among eigenvalues of the correlation matrix are given for each line. For the female preference data, the eigenvalue variance is given for the variance-covariance matrix from the LD scores, the peak preference, and the preference strength, respectively. The relative variance among eigenvalues is calculated following Pavlicev *et al.* (2009), as $Var_{rel}(\lambda) = Var(\lambda)/(N-1)$. One thousand random samples were taken from the posterior distribution of the correlation matrix to calculate the eigenvalue variance and this was bootstrapped 1000 times to obtain the confidence intervals.

(a) ♂ signal	lower	mean	upper
G. rubens	0.06	0.08	0.10
G. texensis	0.17	0.21	0.24
F_1 hybrid	0.16	0.18	0.21
F_2 hybrid	0.16	0.17	0.19
(b) ♀ preference	lower	mean	upper
LD Scores			
G. rubens	0.00	0.02	0.07
G. texensis	0.00	0.05	0.13
F_1 hybrid	0.01	0.08	0.16
F_2 hybrid	0.00	0.05	0.12
Peak			
G. rubens	0.00	0.02	0.07
G. texensis	0.00	0.07	0.17
F_1 hybrid	0.00	0.03	0.08
F_2 hybrid	0.00	0.02	0.08
Strength			
G. rubens	0.00	0.02	0.08
G. texensis	0.00	0.02	0.07
F_1 hybrid	0.00	0.02	0.07
F_2 hybrid	0.00	0.02	0.08

The magnitude of phenotypic integration among female preference functions was very low and none of the traits showed particularly strong correlations with other traits, the only exception being pulse rate preference and carrier frequency preference in *G. texensis* (0.43; Fig. 4, Table 4b). The lower bound of the 95% confidence interval was very close to zero and the mean was close to the expectation under random correlation matrices (equal to the inverse of the sample size, Pavlicev *et al.* 2009). This was regardless of whether integration was estimated using the LD scores or using the peak or discrimination strength of the preference function.

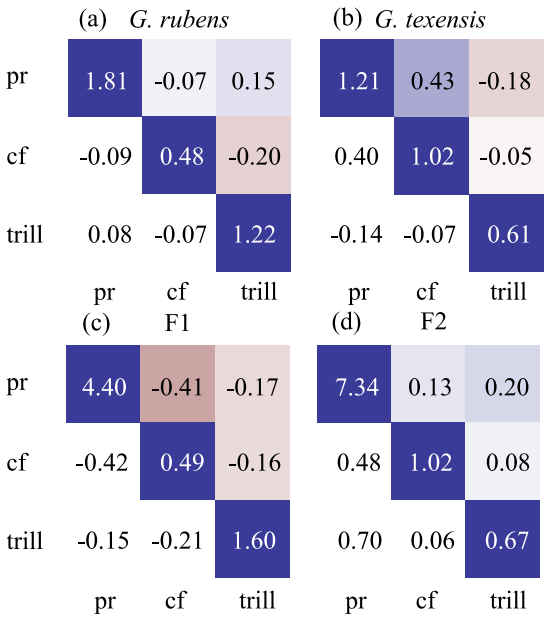

Figure 4 **P** matrices for female preferences. The correlations between the scores on the first linear discriminant (LD) retained from the parental preferences for the pulse rate, carrier frequency, and trill test are given for *G. rubens* (a), *G. texensis* (b), and the F_1 (c) and F_2 (d) hybrids. The intensity of the colours in each heat plots indicates the strength of the positive (blue) or negative (red) correlation. The diagonals show the variances (in white), correlations are given above and covariances below the diagonal.

DISCUSSION

We studied the genetic architecture of interspecific mate signalling divergence between *G. texensis* and *G. rubens*. The genetic architecture of mate signalling informs us about the potential response to selection, as well as about the evolutionary history of divergent mate communication. Traits and preferences of F_1 and F_2 hybrids were intermediate with respect to the parents and showed a wide, unimodal distribution (Fig. 1, Fig. 2) with increasing variance between successive generations, indicative of additive polygenic inheritance. However, for two male song traits, pulse rate and carrier frequency, the F_1 hybrid distribution was distinctly bimodal and biased towards the maternal parent species. Using the joint-scaling method, we also uncovered dominance (pulse rate, trill rate, frequency and trill preference) and epistatic variation (trill preference). Below we discuss the observed patterns of inheritance in the light of previous findings for acoustic mate signalling systems and with respect to putative speciation models. In the next section, we elaborate on the finding that the multivariate phenotype (the **P** matrix) aligns with the major axes of song discrimination and allows for a strong, correlated response to multivariate selection.

Inheritance and speciation models

The observed trait distributions with gradually increasing variance between successive generations are suggestive of a polygeny rather than a single major effect locus. However, in this study we present no formal test for the amount of loci associated with song and preference traits. The best-known method for biometric estimates of the number of genes is the Castle-Wright method (Wright 1968). However, this method is highly sensitive to the presence of non-additive genetic variation and even more to the sample size of the F_2 hybrid or backcross generation (which should be over 100 following Lande 1981) and is likely to provide biologically meaningless results given our data (Zeng *et al.* 1990, Zeng 1992). Nevertheless, based on our observations here, the studies of genetic control of song production in Australian *Teleogryllus* species (Bentley & Hoy 1972, Hoy 1974) and biometric and QTL mapping experiments in Hawaiian *Laupala* crickets (Shaw 1996, Shaw *et al.* 2007, Oh *et al.* 2012), we infer that several unlinked loci (each potentially containing up to several genes) underlie the variation in song and preference traits.

Models of sympatric speciation often assume few loci controlling traits targeted by sexual selection and suggest that speciation is less likely to occur when the number of loci increases (Gourbiere 2004). However, our results support the observation that quantitative traits in general and acoustic mating signals in particular are more likely to have a polygenic basis (Ritchie & Phillips 1998, Shaw &

Parsons 2002). Traits associated with a polygenic architecture are predicted to evolve gradually rather than by large, rapid changes (Templeton 1981). Classically, this has led to the assumption that divergence of polygenic traits is less likely to occur (Templeton 1981), but evidence from, for example, parallel radiations in Hawaiian *Laupala* crickets showed that this need not be true (Mendelson & Shaw 2005). Below, we discuss two insights from this and other studies highlighting mechanisms that potentially facilitated divergence between *G. texensis* and *G. rubens:* X-linkage of calling song traits and peripatric divergence.

X-linkage of loci controlling traits involved in pre-mating isolation increases the likelihood of divergence because of reduced recombination rates on sex chromosomes (Coyne & Orr 1989, Ritchie & Phillips 1998). Sex-linked traits also tend to show higher evolutionary rates because they are more frequently expressed and thus more frequently exposed to selection (Qvarnström & Bailey 2009). In line with earlier findings for *Teleogryllus* (Bentley & Hoy 1972, Hoy 1974) and *Laupala* crickets (Shaw 1996, Shaw *et al.* 2007, Oh *et al.* 2012), we observed that the pulse rate and carrier frequency of reciprocal F_1 hybrid cross lines resembled the maternal parent more than the paternal parent. The four F_2 lines did not differ in their mean phenotype indicating that maternal effects are unlikely.

In addition to the genetic mechanisms constraining the process of divergence and speciation, information on the demographic history of species can be used to infer past evolutionary dynamics. Molecular analysis of the genetic divergence between *G. texensis* and *G. rubens* indicated a peripatric origin of *G. rubens* from one lineage of two historical haplotype partitions in *G. texensis* (Gray *et al.* 2008). Under this scenario, divergence initially took place by selection acting on *G. rubens*, geographically isolated from ancestral *G. texensis* lineages (Gray *et al.* 2008). The authors proposed that demographic subpartitioning may have facilitated divergence in the mating signal. However, these results need confirmation from additional, multi-locus analyses in which demographic scenarios as well as population specific parameters can be assessed more accurately.

Response to selection

The genetic architecture is informative about potential genetic constraints limiting divergence. Here, we analysed the phenotypic variance-covariance matrix to quantify the variation in the multivariate phenotype available to selection. Preferences for the different song traits were largely uncoupled and the magnitude of phenotypic integration (*i.e.*, variance among the eigenvalues of the **P** matrix) was effectively zero (Fig. 4, Table 4b). In contrast, analysis of the **P** matrix of the calling song revealed several strong pairwise correlations and higher levels of integration among orthogonal dimensions. Weak phenotypic integration indicates that the multivariate response to selection is not limited to one or few phenotypic dimensions (*i.e.*, the leading eigenvectors). Strong phenotypic integration can be indicative of limited evolvability when the orthogonal dimensions of song variation do not align with the direction of selection vectors. However, the main song traits that are divergently selected by female crickets (pulse rate and trill duty cycle; Blankers *et al.* 2015, this study) can vary independently (Table S3), suggesting overall high evolvability of the calling song in directions corresponding to predominant selection gradients (Hansen & Houle 2008). Below we discuss how the observed patterns of covariation may potentially affect multivariate responses to selection.

A strong negative correlation was observed between trill rate and trill duty cycle (-0.54/-0.34 for *G. rubens*/*G. texensis*; Fig. 3). Although both trill duty cycle and trill rate can be derived from the trill period mathematically, they can be varied independently through the (independent) variation of the trill duration and trill pause. Potentially, the negative correlation between trill rate and trill duty cycle represents a trade-off between trill duration and trill rate, as has been shown for the grey tree frog (Wells & Taugen 1986, Reichert & Gerhardt 2012). In this case, selection for higher trill duty cycles would result in lower trill rates, which would affect *G. rubens* more strongly than *G. texensis* (because the strength of directional selection on trill duty cycle is lower in *G. texensis* and the correlation between trill duty cycle and trill rate is higher in *G. rubens*, see Blankers *et al.* 2015 and Fig. S1 in this study).

In *G. texensis* (but not *G. rubens*) we also observed a strong correlation between pulse rate and trill rate (0.80) and both traits had strong vector loadings on the first eigenvector of the covariance matrix (Fig. 3, Table S3). In an earlier study comparing **P** among four cricket species (including *G.*

texensis), the authors reported strong dependence of pulse duration on trill duration as well as a (somewhat weaker) correlation between pulse pause and trill pause (Bertram *et al.* 2011). Although the authors analysed different traits of the calling song (duration and pause versus rate and duty cycle), a positive correlation between the temporal pattern of the pulse and that of the trill supports our finding of a positive correlation between pulse rate and trill rate. Interestingly, the other three *Gryllus* species analysed by Bertram *et al.* (2011) showed the same pattern, further strengthening the hypothesis of correlated evolution of pulse rate and trill (or chirp) rate in crickets.

On a proximate level, a correlation between trill rate and pulse rate may result from shared neuronal mechanisms. For example, for the field crickets *Teleogryllus oceanicus* and *G. campestris* it was shown that a single command neuron was associated with the temporal pattern on both the long and the short timescale of the calling song (Bentley 1977). However, it is unclear whether the central pattern generators for pulse and trill rate are physiologically coupled by thoracic neuronal networks for the two timescales of the song rhythm (Hedwig 2000, Schöneich & Hedwig 2011).

On an ultimate level, a correlation between pulse rate and trill rate, as well as correlated evolution of trill rate and trill duty cycle provide insight in the constraints in the evolution of the calling song. On one hand, selection for higher pulse rates may be constrained by limitations in maximum energy expenditure when singing at higher pulse rates co-occurs with singing at higher (energetically expensive) trill rates. On the other hand, in the absence of such a limitation, or when singing at higher rates trades-off with singing at lower duty cycles, trill rate will be indirectly selected following direct selection on pulse rate and trill duty cycle. In the case of *G. rubens* and *G. texensis*, a negative correlation between trill rate and trill duty cycle and a positive correlation between trill rate and pulse rate would thus lead to the trill rate being indirectly selected due to selection on pulse rate (lower in *G. rubens*) and trill duty cycle (higher in *G. rubens*) in the observed direction of divergence (lower pulse rate, lower trill rate, and higher trill duty cycle in *G. rubens*).

In summary, the patterns of covariation highlight a possible mechanism for direct and indirect selection to augment evolvability of acoustic mating signals rather than constrain it. The leading dimensions of the multivariate signal align with selection gradients. Whether and how rapidly the observed selection response could have resulted in divergence between these sibling species also strongly depends on the genetic architecture. On one hand, a polygenic genetic architecture and the presence of dominance and epistatic variation may have resulted in slow and gradual evolution of reproductive isolation in sympatry. On the other hand, ancestral demography and X-linkage may have facilitated divergence. Together, these results are informative about the potential evolutionary mechanisms as well as the rate and direction of divergence in acoustic mate signalling under sexual selection.

ACKNOWLEDGEMENTS

We greatly appreciate support for the trackball system by Hansjürgen Dahmen. We thank D.A. Gray for help with collecting the crickets in Texas and Florida. The manuscript strongly benefitted from comments by Emma Berdan, Jonas Finck, and Michael Reichert. Additionally, for reviewing and commenting on the manuscript we thank Mike Ritchie (Editor), Susan Bertram, and one anonymous reviewer. The performed experiments comply with the "Principles of animal care", publication No. 86-23, revised 1985 of the National Institute of Health, and also with the current laws of Germany. The authors declare no conflict of interest. This study is part of the GENART project funded by the Leibniz Association (SAW-2012-MfN-3)

REFERENCES

Alexander, R.D. 1962. Evolutionary change in cricket acoustical communication. *Evolution.* **16**: 443–467.
Andersson, M. 1994. *Sexual Selection*. Princeton University Press.
Arbuthnott, D. 2009. The genetic architecture of insect courtship behavior and premating isolation. *Heredity.* **103**: 15–22.
Arnqvist, G. 1998. Comparative evidence for the evolution of genitalia by sexual selection. *Nature* **393**: 784–786.
Bentley, D. 1977. Control of cricket song patterns by descending interneurons. *J. Comp. Physiol. A.* **116**: 19–38.
Bentley, D.R. & Hoy, R.R. 1972. Genetic control of the neuronal network generating cricket (*Teleogryllus Gryllus*) song patterns. *Anim. Behav.* **20**: 478–492.

Bertram, S.M., Fitzsimmons, L.P., McAuley, E.M., Rundle, H.D. & Gorelick, R. 2011. Phenotypic covariance structure and its divergence for acoustic mate attraction signals among four cricket species. *Ecol. Evol.* **2**: 181–195.
Blankers, T., Hennig, R.M. & Gray, D.A. 2015. Conservation of multivariate female preference functions and preference mechanisms in three species of trilling field crickets. *J. Evol. Biol.* **28**: 630–641.
Blows, M.W., Brooks, R. & Kraft, P.G. 2003. Exploring complex fitness surfaces: multiple ornamentation and polymorphism in male guppies. *Evolution.* **57**: 1622–1630.
Cavalli, L.L. 1952. An Analysis of Linkage in Quantitative Inheritance. In: *Quantitative Inheritance* (E. C. R. Reeve & C. H. Waddington, eds), pp. 135–144. HM Stationery Office, London.
Chenoweth, S.F. & McGuigan, K. 2010. The genetic basis of sexually selected variation. *Annu. Rev. Ecol. Evol. Syst.* **41**: 91–101.
Cheverud, J.M., Wagner, G.P. & Dow, M.M. 1989. Methods for the comparative analysis of variation patterns. *Syst. Zool.* **38**: 201–213.
Coyne, J.A. & Orr, H.A. 1989. Two rules of speciation. In: *Speciation and its consequences* (J. A. Endler & D. Otte, eds). Sinauer Associates, Sunderland, MA.
Doorn, G.S. Van, Dieckmann, U. & Weissing, F.J. 2004. Sympatric speciation by sexual selection : a critical reevaluation. *Am. Nat.* **163**: 709–725.
Foley, B., Chenoweth, S.F., Nuzhdin, S. V. & Blows, M.W. 2007. Natural genetic variation in cuticular hydrocarbon expression in male and female Drosophila melanogaster. *Genetics* **175**: 1465–1477.
Gavrilets, S. 2003. Perspective: models of speciation: what have we learned in 40 years? *Evolution* **57**: 2197–2215.
Gavrilets, S. & Vose, A. 2007. Case studies and mathematical models of ecological speciation. 2. Palms on an oceanic island. *Mol. Ecol.* **16**: 2910–2921.
Gelman, A. 2006. Prior distributions for variance parameters in hierarchical models. *Bayesian Anal.* **1**: 515–533.
Gourbiere, S. 2004. How do natural and sexual selection contribute to sympatric speciation ? *J. Evol. Biol.* **17**: 1297–1309.
Gray, D. A., Huang, H. & Knowles, L.L. 2008. Molecular evidence of a peripatric origin for two sympatric species of field crickets (*Gryllus* rubens and *G. texensis*) revealed from coalescent simulations and population genetic tests. *Mol. Ecol.* **17**: 3836–3855.
Gray, D.A. 2011. Speciation, divergence, and the origin of *Gryllus* rubens: behavior, morphology, and molecules. *Insects* **2**: 195–209.
Gray, D.A. & Cade, W.H. 1999. Quantitative genetics of sexual selection in the field cricket, *Gryllus* integer. *Evolution.* **53**: 848–854.
Gray, D.A. & Cade, W.H. 2000. Sexual selection and speciation in field crickets. *Proc Natl Acad Sci* **97**: 14449–14454.
Hadfield, J.D. 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Softw.* **33**: 1–22.
Hadfield, J.D. 2012. MCMCglmm Course Notes.
Hansen, T.F. 2006. The evolution of genetic architecture. *Annu. Rev. Ecol. Evol. Syst.* **37**: 123–157.
Hansen, T.F. & Houle, D. 2008. Measuring and comparing evolvability and constraint in multivariate characters. *J. Evol. Biol.* **21**: 1201–1219.
Hebets, E.A. & Papaj, D.R. 2005. Complex signal function: developing a framework of testable hypotheses. *Behav. Ecol. Sociobiol.* **57**: 197–214.
Hedwig, B. 2000. Control of cricket stridulation by a command neuron: efficacy depends on the behavioral state. *J. Neurophysiol.* **83**: 712–722.
Hedwig, B. & Poulet, J.F. a. 2004. Complex auditory behaviour emerges from simple reactive steering. *Nature* **430**: 781–785.
Higashi, M., Takimoto, G. & Yamamura, N. 1999. Sympatric speciation by sexual selection. *Nature* **402**: 523–526.
Higham, J.P. & Hebets, E.A. 2013. An introduction to multimodal communication. *Behav. Ecol. Sociobiol.* **67**: 1381–1388.
Hoy, R.R. 1974. Genetic control of acoustic behavior in crickets. *Am. Zool.* **14**: 1067–1080.
Kirkpatrick, M. & Hall, D.W. 2004. Sexual selection and sex linkage. *Evolution.* **58**: 683–691.
Lande, R. 1981. Models of speciation by sexual selection on polygenic traits. *Proc Natl Acad Sci* **78**: 3721–3725.
Lande, R. & Arnold, S.J. 1983. The measurement of selection on correlated characters. *Evolution.* **37**: 1210–1226.
Mather, K. & Jinks, J.L. 1971. *Biometrical Genetics.* Chapman and Hall, NY.
McGuigan, K. & Blows, M.W. 2007. The phenotypic and genetic covariance structure of drosphilid wings. *Evolution.* **61**: 902–911.
Meffert, L.M., Hicks, S.K. & Regan, J.L. 2002. Nonadditive genetic effects in animal behavior. *Am. Nat.* **160**: S198–S213.
Mendelson, T.C. & Shaw, K.L. 2005. Rapid speciation in an arthropod. *Nature* **433**: 375–376.
Meyer, K. & Kirkpatrick, M. 2005. Up hill , down dale : quantitative genetics of curvaceous traits. *Philos. Trans. R. Soc. B Biol. Sci.* **360**: 1443–1455.
Nakagawa, S. & Schielzeth, H. 2010. Repeatability for Gaussian and non-Gaussian data: A practical guide for biologists. *Biol. Rev.* **85**: 935–956.
Oh, K.P., Fergus, D.J., Grace, J.L. & Shaw, K.L. 2012. Interspecific genetics of speciation phenotypes: Song and preference co-evolution in Hawaiian crickets. *J. Evol. Biol.* **25**: 1500–1512.
Otte, D. 1992. Evolution of cricket songs. *J. Orthoptera Res.* 25–49.
Pavlicev, M., Cheverud, J.M. & Wagner, P. 2009. Measuring morphological integration using eigenvalue variance. *Evol. Biol.* **36**: 157–170.
Qvarnström, A. & Bailey, R.I. 2009. Speciation through evolution of sex-linked genes. *Heredity.* **102**: 4–15.
Reichert, M.S. & Gerhardt, H.C. 2012. Trade-offs and upper limits to signal performance during close-range vocal competition in gray tree frogs Hyla versicolor. *Am. Nat.* **180**: 425–437.
Ritchie, M.G. & Phillips, S.D.F. 1998. The genetics of sexual isolation. In: *Endless Forms: Species and Speciation* (D. J. Howard & S. H. Berlocher, eds), p. 291. Oxford University Press, New York.
Scheuber, H., Jacot, A. & Brinkhof, M.W. 2004. Female preference for multiple condition-dependent components of a sexually selected signal. *Proc Biol Sci* **271**: 2453–2457.
Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. *Evolution.* **50**: 1766–1774.
Schöneich, S. & Hedwig, B. 2011. Neural basis of singing in crickets: Central pattern generation in abdominal ganglia. *Naturwissenschaften* **98**: 1069–1073.

Shaw, K.L. 1996. Polygenic inheritance of a behavioral phenotype: interspecific genetics of song in the Hawaiian cricket genus Laupala. *Evolution.* **50**: 256–266
Shaw, K.L. & Lesnick, S.C. 2009. Genomic linkage of male song and female acoustic preference QTL underlying a rapid species radiation. *Proc. Natl. Acad. Sci.*. **106**: 9737–9742.
Shaw, K.L. & Parsons, Y.M. 2002. Divergence of mate recognition behavior and its consequences for genetic architectures of speciation. *Am. Nat.* **159 Suppl** : S61–S75.
Shaw, K.L., Parsons, Y.M. & Lesnick, S.C. 2007. QTL analysis of a rapidly evolving speciation phenotype in the Hawaiian cricket Laupala. *Mol. Ecol.* **16**: 2879–2892.
Steppan, S.J., Phillips, P.C. & Houle, D. 2002. Comparative quantitative genetics: evolution of the G matrix. *Trends Ecol. Evol.* **17**: 320–327.
Svensson, E.I. & Gosden, T.P. 2007. Contemporary evolution of secondary sexual traits in the wild. *Funct. Ecol.* **21**: 422–433.
Templeton, A.R. 1981. Mechanisms of speciation--A population genetic approach. *Annu. Rev. Ecol. Syst.* **12**: 23–48.
Veen, T., Faulks, J., Tyler, F., Lloyd, J. & Tregenza, T. 2013. Diverse reproductive barriers in hybridising crickets and maintenance of isolation. *Evol. Ecol.* **27**: 993–1015.
Venables, W.N. & Ripley, B.D. 2002. *Modern applied statistics with S*. Springer, New York, NY.
Walker, T.J. 2015. Crickets. *In* Singing Insects of North America. http://entnemdept.ifas.ufl.edu/walker/Buzz/crickets.htm.
Wells, K.D. & Taigen, T.L. 1986. The effect of social interactions on calling energetics in the gray treefrog (Hyla versicolor). *Behav. Ecol. Sociobiol.* **19**: 9–18.
West-Eberhard, M.J. 1983. Sexual selection, social competition, and speciation. *Q. Rev. Biol.* **58**: 155–183.
Wiens, J.J. 2001. Widespread loss of sexually selected traits : how the peacock lost its spots. *Trends Ecol. Evol.* **16**: 517–523.
Willis, J.H., Coyne, J.A. & Kirkpatrick, M. 1991. Can one predict the evolution of quantitative characters without genetics. *Evolution.* **45**: 441–444.
Wilkins, M.R., Seddon, N. & Safran, R.J. 2013. Evolutionary divergence in acoustic signals : causes and consequences. *Trends Ecol. Evol.* **28**: 156–166.
Wolfenbarger, L.L. & Wilkinson, G.S. 2001. Sex-linked expression of a sexually selected trait in the stalk-eyed fly, Cyrtodiopsis dalmanni. *Evolution.* **55**: 103–110.
Wright, S. 1968. *Evolution and genetics of populations. Vol. 1. Genetic and biometric foundation*. University of Chicago Press, Chicago.
Zeng, Z. 1992. Correcting the bias of Wright's estimates of the number of genes affecting a quantitative character: a further improved method. *Genetics* **131**: 987–1001.
Zeng, Z., Houle, D. & Cockerham, C.C. 1990. How informative is Wright' s estimator of the number of genes affecting a quantitative character? *Genetics* **126**: 235–247.

SUPPORTING INFORMATION

Table S1: pairwise comparisons of song traits and preferences (peak and strength).
A short discussion on using linear discriminant analysis to quantify differences in preference functions.
Fig. S1: the preference functions for pulse rate, carrier frequency, and the trill pattern.
Table S2: the coefficients of the trace of LD 1.
Table S3: eigenvectors of the variance-covariance matrices.

Table S1 Pairwise t-tests (male traits) and Mann-Whitney U tests (female preferences).

t-tests pulse rate			
	rub	tex	F1
tex	t_{43} = -23.40; P < 0.0001		
F1	t_{49} = -16.96; P < 0.0001	t_{43} = 7.14; P < 0.0001	
F2	t_{34} = -11.41; P < 0.0001	t_{34}= 9.33; P < 0.0001	t_{34}= 2.7775; P = 0.0070
rubtex vs texrub	t_{22} = -6.72; P < 0.0001		
	F2A	F2B	F2C
F2B	t_{1} = 0.6086; P = 0.6193		
F2C	t_{6} = 0.2124; P = 0.8356	t_{1} = -0.5305; P = 0.6769	
F2D	t_{3}= -0.7718; P = 0.4625	t_{1} = -1.1018; P = 0.4125	t_{3} = -1.1394; P = 0.3011
t-tests pulse duty cycle			
	rub	tex	F1
tex	t_{43}= -0.3435; P = 0.7321		
F1	t_{49} = -2.7356; P = 0.0073	t_{43} = 2.0002; P = 0.0483	
F2	t_{34}= 1.9458; P = 0.0560	t_{34}= 2.0064; P = 0.0485	t_{34}= 4.0485; P = 0.0001
rubtex vs texrub	t_{22}= -0.8392; P = 0.4055		
t-tests carrier frequency			
	rub	tex	F1
tex	t_{43} = -9.9019; P < 0.0001		
F1	t_{49} = -5.0015; P < 0.0001	t_{43} = 4.0932; P < 0.0001	
F2	t_{34}= -2.19999; P = 0.0315	t_{34}= 5.082; P < 0.0001	t_{34} = 1.7369; P = 0.0865
rubtex vs texrub	t_{22}= -3.9146; P = 0.0004		
t-tests trill rate			
	rub	tex	F1
tex	t_{43} = -16.3497; P < 0.0001		
F1	t_{49} = -7.5851; P < 0.0001	t_{43} = 8.7341; P < 0.0001	

F2	t_{34}= -7.7691; P < 0.0001	t_{34}= 9.1786; P < 0.0001	t_{34} = 0.1576; P = 0.8752
rubtex vs texrub	t_{22}= -1.0907; P = 0.2809		
t-tests trill duty cycle			
	rub	tex	F1
tex	t_{43} = 7.3622; P < 0.0001		
F1	t_{49} = 2.6851; P = 0.0086	t_{43} = -4.0415; P = 0.0001	
F2	t_{34}= 2.8056; P = 0.0068	t_{34}= -3.376; P = 0.0012	t_{34} = 0.3493; P = 0.7279
rubtex vs texrub	t_{22}= -0.4633; P = 0.6453		
Mann-Whitney U-tests Pulse Rate			
	rub	tex	F1
tex	U_{25} = 49 ; P < 0.0001		
F1	U_{32} = 232.5; P < 0.0001	U_{25} = 269; P = 0.0084	
F2	U_{36} = 418; P < 0.0001	U_{25} = 243; P = 0.0005	U_{32} = 512.5; P = 0.2219
rubtex vs texrub	U_{13} = 123.5; P = 0.7221		
	F2A	F2B	F2C
F2B	U_{8} = 26.5; P =0.0227		
F2C	U_{14} = 66; P =0.1231	U_{8} = 65.5; P =0.315	
F2D	U_{1} = 14.5; P =0.9999	U_{1} = 13; P =0.2085	U_{1} = 17.5; P =0.4699
Mann-Whitney U-tests carrier frequency			
	rub	tex	F1
tex	U_{22} = 308; P = 0.6253		
F1	U_{23} = 300; P = 0.3684	U_{22} = 293.5; P = 0.7025	
F2	U_{23} = 207.5; P = 0.0578	U_{23} = 349.5; P = 0.0510	U_{23} = 374; P = 0.0605
rubtex vs texrub	U_{11} = 123.5; P = 0.5349		
Mann-Whitney U-tests trill pattern			
	rub	tex	F1
tex	U_{15} = 155.5; P = 0.4985		
F1	U_{21} = 277; P = 0.9705	U_{15} = 219.5; P = 0.5892	
F2	U_{20} = 182.5; P = 0.1003	U_{15} = 161.5; P = 0.8364	U_{20} = 217; P = 0.2613
T-tests preference strength (D) pulse rate			
	rub	tex	F1
tex	t_{33} = 1.5531; P = 0.1251		
F1	t_{37} = 1.3576; P = 0.1787	t_{37} = -0.2783; P = 0.782	
F2	t_{40} = -0.3761; P = 0.7084	t_{40} = -0.5691; P = 0.5718	t_{37} = -0.212; P = 0.8327
T-tests preference strength (D) carrier frequency			
	rub	tex	F1
tex	t_{22} = -1.4688; P = 0.1481		
F1	t_{23}= 0.6927; P = 0.4916	t_{22} = 1.2621; P = 0.2135	
F2	t_{23} = -0.1992; P = 0.8429	t_{23} = 2.2437; P = 0.0298	t_{23} = -0.8985; P = 0.3736
T-tests preference strength (D) trill pattern			
	rub	tex	F1
tex	t_{15} = 2.3334; P = 0.02661		
F1	t_{21} = 1.3974; P = 0.1691	t_{15} = -1.0946; P = 0.2819	
F2	t_{20} = 1.4759; P = 0.1477	t_{15} = -0.9926; P = 0.3285	t_{20} = 0.1001; P = 0.9207

Quantifying female preference

The LD scores give an overly simplified description of the preference function. However, our data indicate that this measure accurately reflects the main differences in the shape and peak values of the preference functions and is repeatable among individuals (Table S2 vs Fig. S1). For pulse rate the biggest difference was between preferences for stimuli with conspecific versus heterospecific pulse rates (Fig. S1). The linear discriminant used to predict scores for the parental and hybrid individuals was also correlated with those stimuli: negative scores indicated high preferences for *G. rubens* pulse rates (50.0 pulses/s), whereas positive scores were associated with high preferences for stimuli with *G. texensis* pulse rates (71.4 pulses/s; Table S2).

The differences in trill pattern preferences were also accurately described by variation in the LD scores. *Gryllus rubens* had a strong preference for the third trill pattern stimulus (corresponding to the highest trill duty cycle), whereas *G. texensis* were overall less selective (supported by a significant difference [t_{15} = 2.3334; P = 0.02661] in the discrimination strength, D, between the species; Table S1, Fig. S1), but also had a median peak preference for stimulus 3. The LD scores reflected this difference in that positive scores (*G. texensis*-like) were associated with higher preferences for all stimuli except stimulus 3 and negative scores (*G. rubens*-like) with lower preferences for stimuli 1,2, and 4, and higher preferences for stimulus 3 (Table S2).

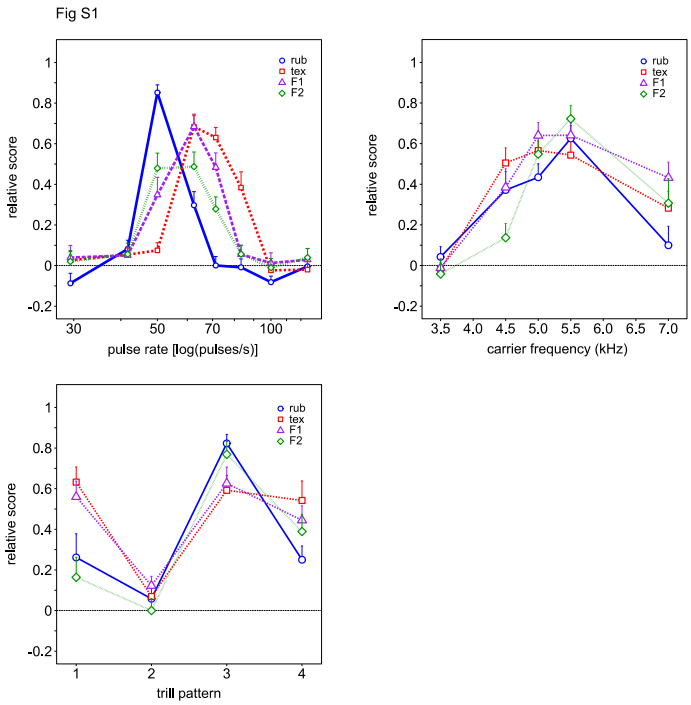

Figure S1 Preference functions for the pulse rate, carrier frequency and trill tests. Preferences and standard error of the mean are shown for *G. rubens* (rub, blue solid lines and circles), *G. texensis* (tex, red dashed lines and squares), F_1 hybrids (F_1, purple long-dashed lines and triangles), and F_2 hybrids (F_2, green dot-dashed lines and diamonds). The black, horizontal, dashed line shows the relative phonotactic score of zero. The trill parameter values corresponding to the 4 stimuli in the trill test are (1): duration=136, pause=54, duty cycle=0.72, rate=5.3; (2): duration=136, pause=388, duty cycle=0.26, rate=1.9; (3): duration=536, pause=54, duty cycle=0.91, rate=1.7; (4): duration=536, pause=388, duty cycle=0.57, rate=1.1.

Table S2 Coefficients of the trace for the linear discriminant analysis on the preference functions for pulse rate (a), carrier frequency (b), and the trill pattern (c).

(a)	pulse rate	LD1.PR
PR1	125.0	-0.73
PR2	100.0	0.57
PR3	83.3	-0.14
PR4	71.4	1.28
PR5	62.5	-0.26
PR6	50.0	-9.12
PR7	41.7	-0.57
PR8	29.4	-0.73
(b)	carrier frequency	LD1.CF
Freq1	3.50	-1.50
Freq2	4.50	0.87
Freq3	5.00	1.65
Freq4	5.50	-0.75
Freq5	7.00	1.63
(c)	trill duty cycle	LD1.trill
Trill1	0.72	0.95
Trill2	0.26	0.75
Trill3	0.91	-2.14
Trill4	0.57	1.09

Table S3 Eigenvectors (ev1-ev5) for the scaled variance-covariance matrices of the male song traits (Fig. 3). The eigenvalues, λ, in italics, and the vector loadings for carrier frequency, pulse rate, pulse duty cycle, trill rate, and trill duty cycle are shown. High loadings (>0.5) are in bold face type.

G. rubens	ev1	ev2	ev3	ev4	ev5
λ	*1.02*	*0.78*	*0.56*	*0.18*	*0.00*
cf	**0.57**	**0.72**	-0.11	0.38	0.00
pr	0.22	0.01	**0.97**	-0.07	0.00
pdc	-0.20	-0.30	0.12	**0.92**	-0.01
tr	**0.77**	**-0.62**	-0.16	-0.01	0.00
tdc	0.00	0.00	0.00	0.01	**1.00**
G. texensis					
λ	*1.17*	*0.89*	*0.22*	*0.18*	*0.00*
cf	0.31	**0.93**	-0.17	-0.09	0.01
pr	**-0.79**	0.30	-0.08	**0.53**	0.00
pdc	0.00	0.00	-0.01	0.00	**-1.00**
tr	**-0.53**	0.12	0.11	**-0.83**	0.00
tdc	0.05	0.17	**0.98**	0.12	-0.01
F_1 hybrid					
λ	*1.42*	*0.42*	*0.20*	*0.12*	*0.01*
cf	**0.73**	0.01	**0.64**	0.25	0.00
pr	**0.65**	-0.12	**-0.75**	0.05	0.00
pdc	0.00	0.00	0.00	0.00	**1.00**
tr	-0.08	**-0.99**	0.08	0.07	0.00
tdc	0.21	-0.07	0.14	**-0.96**	0.00
F_2 hybrid					
λ	*1.32*	*1.16*	*0.51*	*0.29*	*0.17*
cf	-0.39	**0.49**	**0.50**	-0.01	**0.59**
pr	**-0.59**	-0.27	0.45	-0.26	**-0.56**
pdc	-0.45	-0.29	**-0.55**	**-0.50**	0.40
tr	-0.16	**-0.68**	0.16	**0.62**	0.33
tdc	**-0.52**	0.39	-0.46	**0.55**	-0.27

CHAPTER 4

CONTRASTING SEXUAL SELECTION REGIMES DRIVE INTERSPECIFIC DIVERGENCE IN THE PHENOTYPIC VARIANCE-COVARIANCE MATRIX OF AN ACOUSTIC MATING SIGNAL

Based on: Blankers, T, Gray, DA, Hennig RM. 2015. Contrasting sexual selection regimes drive interspecific divergence in the phenotypic variance-covariance matrix of an acoustic mating signal. In review

Abstract. The hypothesis that species exposed to similar selection regimes acting on multivariate traits also have patterns of trait covariation in common and thus evolve along similar phenotypic dimensions has rarely been addressed empirically, in particular not for behavioural mating signals. Here, we compared the multivariate calling song among seven species of field crickets (*Gryllus*) encompassing two selection regimes: (i) three species producing trilled songs (*i.e.* with long trains of pulses) for which preference functions for acoustic energy (chirp duty cycle) are linear, and (ii) four species producing chirped songs for which preference functions for chirp duty cycle are concave. Comparison of common subspaces and tensor analysis of the set of **P** matrices revealed interspecific differentiation. Traits describing the chirp structure, *i.e.* chirp rate and chirp duty cycle, had a strong contribution to the divergence in **P**. Variation in **P** was found to affect the predicted effects of selection mostly through indirect selection on chirp rate. The comparisons of **P** strongly contrasted species experiencing different selection regimes. This study highlights the potential for sexual selection to shape **P** and shows that the *shape* of the selection landscape can alter the evolutionary trajectory of mating signals, even among closely related species.

INTRODUCTION

Phenotypic evolution and the response of traits to selection is a central topic in evolutionary biology. Predicting selection responses is also of practical value both to agricultural breeders and to evolutionary ecologists inferring species' responses to environmental change (Lande 1979, Hazel *et al.* 1994, Steppan *et al.* 2002, Arnold *et al.* 2008, Laughlin & Messier 2015). For a single trait, the breeder's equation, $R = h^2S$, describes the expected change in trait mean in response to selection as a function of the phenotypic and additive genetic variance (Lynch & Walsh 1998). However, the ubiquity of genetic and phenotypic correlations between traits underlines the importance of the multivariate extension of the breeder's equation, $\boldsymbol{\Delta z = GP^{-1}s}$. Here, **G** and **P** are the genetic and phenotypic (co)variance, respectively, which determine the correlated response to $\boldsymbol{s}$, a vector of selection differentials (Lande 1979, Lande & Arnold 1983, Lynch & Walsh 1998).

The **G** matrix reveals whether additive genetic variation is available in the direction of selection and whether the selection response is expected to be biased away from that direction (Lande 1979, Steppan *et al.* 2002, Arnold *et al.* 2008). The **P** matrix represents the multivariate phenotype on which selection acts and comparing **P** among species can thus shed light on phenotypic divergence across phylogenetic, environmental, and selective gradients (*e.g.* Steppan 1997, Kolbe *et al.* 2011). Compared to **G**, **P** is easier to obtain and estimated with smaller error (Willis *et al.* 1991). Additionally, **P** and **G** are expected to be correlated and as such **P** can inform the genetic architecture (Cheverud 1988, Roff *et al.* 1999). However, similarity between **G** and **P** can only be assumed in the presence of additional data (*e.g.* on heritability, Willis *et al.* 1991). Thus, although variation in **P** may also provide some information about variation in **G**, a comparative analysis of **P** primarily offers detailed insights into multivariate phenotypic evolution and can be used to infer the role of selection in shaping phenotypic covariation and the evolutionary trajectory of traits.

Traditionally, patterns of covariance were assumed to be constant over time, but natural and experimentally selected populations have been found to rapidly diverge in their (co)variance structure (Turelli 1988, Blows & Higgie 2003, Hine *et al.* 2009). Simulation studies have shown that variation in the adaptive landscape, in particular the onset of directional selection, may significantly alter the covariance structure of multivariate traits (Jones *et al.* 2003, Arnold *et al.* 2008, Melo & Marroig 2014). This implies a strong role for selection in shaping **G** and **P**. However, the hypothesis that species experiencing similar selection regimes also share patterns of covariance and are more similar in the direction of multivariate phenotypic evolution compared to species with different selection regimes remains largely untested empirically (but see Blows *et al.* 2004, Hine *et al.* 2009, Berner *et al.* 2010, Kolbe *et al.* 2011, Roff & Fairbairn 2012).

Here, we studied multivariate phenotypic evolution in light of variation in the selection regime and tested the extent to which divergence in **P** between selection regimes is expected to affect evolutionary trajectories of a behavioural mating signal. The majority of studies of multivariate phenotypic evolution are limited to pairwise comparisons and are concentrated on morphological traits (Steppan *et al.* 2002, Arnold *et al.* 2008, Aguirre *et al.* 2014). Using the available tools to compare matrices among multiple populations simultaneously (Hine *et al.* 2009, Aguirre *et al.* 2014) and studying patterns of phenotypic covariance for behavioural traits can significantly improve our understanding of phenotypic evolution beyond pairwise interspecific differences in allometric integration.

Of particular interest are among species comparisons of sexually selected (mating) signals, because sexual signalling traits play a central role in divergence and speciation (Ritchie 2007). Acoustic mating signals such as those common to frogs and crickets, are typically comprised of multiple, correlated quantitative traits (Gerhardt & Huber 2002) and for many of these signals detailed information on the corresponding (multivariate) preferences is available (*e.g.* Bentsen *et al.* 2006, Gerhardt & Brooks 2009, Rodriguez *et al.* 2013). This opens the door to comprehensive, among species comparisons of **P** across selective regimes.

The selection regime of a mating signal depends on the shape of the preference function, which can be analysed in a quantitative framework using selection analysis (Lande & Arnold 1983), and the phenotypic distribution of the signal in relation to that preference function (Ryan & Rand 1993). Linear (*i.e.* open-ended) preference functions will always result in directional selection, but concave (*i.e.*, unimodal, closed) preference functions effect either stabilizing or directional selection depending on the distance between the distribution of the signal and the peak of the preference function (Rodriguez *et al.* 2013). However, regardless of whether concave preferences result in directional or stabilizing selection, they are very different from linear preference functions from a proximate perspective, *i.e.* in terms of the neuronal filter properties that accommodate signal perception and evaluation (*e.g.* Hennig *et al.* 2014).

We compared **P** among seven species of *Gryllus* field crickets. The calling song varies considerably among all of these species and interspecific variation is highest in pulse rate (*i.e.* the

repetition rate of the individual sound pulses, Alexander 1962, Otte 1992, Walker 2015, see Results). We assigned the seven species to either of two groups (see also Results): (1) species producing long, continuous trills (chirps with more than 20 consecutive pulses, Alexander 1962): *Gryllus texensis*, *G. rubens*, and an undescribed congener referred to here as G#14; and (2) species that produce short chirps of only a few (3-10) pulses: *G. lineaticeps*, *G. personatus*, *G. firmus*, and another undescribed congener G#15 (Sakaguchi & Gray 2011; Gray *et al.* 2015). The major distinction between these two groups is in the acoustic energy in the song (*i.e.*, chirp duty cycle: the relative amount of active singing during each chirp period; see Results), which is low in the chirping species and high in the trilling species.

The variation in the calling song within and between chirpers and trillers is reflected in female preference function shape. Pulse rate preference functions are concave in all seven species, with peak preferences closely tuned to the mean value among males (Blankers *et al.* 2015, R.M. Hennig, T. Blankers, D.A. Gray, in review). However, chirp duty cycle is associated with linear preference functions in trillers (Blankers *et al.* 2015), whereas preference functions for chirp duty cycle in the chirping species are concave with peak preferences offset above the mean values produced by males (R. M. Hennig, T. Blankers, D. Gray, in review; see also Wagner & Basolo 2007 for *G. lineaticeps* chirp preference). These data thus provide a rare opportunity to examine the variability in the covariance structure of a mating signal in relation to variation in preference function shape on a genus-wide level.

Using three methods from the framework presented by Aguirre *et al.* (2014) while accounting for phylogenetic non-independence of **P** using Mantel tests (Swenson 2014), we addressed the following questions: (1) Does **P** vary among species? (2) Is divergence in **P** among species associated with song traits for which variation exists in the shape of the preference function (*i.e.*, the traits describing the temporal pattern of the chirp)? (3) Is variation in **P** largest between species experiencing different selection regimes (*i.e.* between chirpers and trillers)? (4) Does variation in **P** among species translate to variation in their evolutionary trajectories?

We tested for variation in **P** among species using the multi-population extension of Krzanowski's subspace comparison and analysis of **P** in tensor form (Hine *et al.* 2009, Aguirre *et al.* 2014) but also employed traditional pairwise methods (Common Principal Components analysis, Krzanowski's subspace comparison and Mantel tests). The multi-population extension of Krzanowski's subspace comparison provides a metric indicating similarity (*i.e.* common subspaces) of covariance structures across multiple populations (Aguirre *et al.* 2014). For the analysis of **P** in tensor form we estimated the variance-covariance tensor, which quantifies (co)variation among second-order tensors (*i.e.* matrices). To examine whether traits that are associated with variation in the shape of the preference function contribute more strongly to divergence in **P** than traits for which preference functions are similar in shape, the variance-covariance tensor was decomposed in its eigentensors to reveal in which dimensions the **P** matrices differ and which combination of traits contributed to the divergence in **P** (Hine *et al.* 2009, Aguirre *et al.* 2014). The coordinates of the species' **P** matrices on the eigentensor were used to test if species that have similar selection regimes are less divergent in **P** compared to species that differ in the selection regime. To examine whether variation in **P** translated to variation in the effects of sexual selection and whether selection effects were most divergent between selection regimes we used the selection response decomposition method described in Aguirre *et al.* (2014). Through their Bayesian nature, these methods allow for quantifying the uncertainty (*e.g.*, based on the quantile-based Highest Posterior Density [HPD] interval) in the estimation of **P** and the uncertainty in the response to the linear transformations during the analyses.

METHODS

Collecting and rearing

Laboratory populations were initiated from field caught individuals: *Gryllus texensis* were collected in Austin (TX), Lancaster (TX) and Round Rock (TX); *G. rubens* and *G. firmus* in Gainesville (FL), Lake City (FL), and Live Oak (FL); G#14 and G#15 in Agua Fria National Monument (AZ); *G. lineaticeps* in Santa Monica Mountains (CA); and *G. personatus* in Winslow (AZ). The crickets were brought back to the lab and housed in 19L containers at an average temperature of 25.3 °C (± 2.73 SD) with gravel, shelter, and water and food *ad libitum*. At least 30 males and 30 females were used for breeding each generation. First, second, and third generation laboratory offspring were used in the experiments.

Song recording

The methods to obtain calling song data have been described in detail in a previous study (Blankers *et al.* 2015). In short, individuals were placed in separate boxes (mean temperature 24.9 °C +/- 0.98 SD) and recorded in the dark for a 16-24 hour period. For each recording, the duration of individual pulses and chirps, and pauses between pulses and chirps were measured. From the period (sum of duration and

pause of a pulse or chirp) and the duration of a pulse or chirp, the rate (inverse of the period) and duty cycle (duration divided by the period) were calculated. Individual mean values used in the statistical analyses described below were based on multiple 10 second windows from separate calling bouts. Some of these calling bouts were recorded on the same day, but repeated measurements were also collected from different days, which were maximally 6 days apart. A total of 637 song recordings were used in the analyses (median per individual = 2, range = 1 – 7; see Table 1 for the sample size per species).

Interspecific song divergence

All analyses were conducted in R 3.1.2 (R core team 2015). In our analyses we included the five song traits that are the most likely candidates to be under stabilizing and/or directional selection: carrier frequency, pulse rate, pulse duty cycle, chirp rate, and chirp duty cycle (Gray & Cade 2000, Wagner & Basolo 2007, Grobe *et al.* 2012, Rothbart & Hennig 2012, Blankers *et al.* 2015). Since these traits are potentially collinear, the variance inflation factor (VIF) was calculated by including all five traits in a linear regression. The VIF was lower than 5 for all song traits, indicating low levels of collinearity (Heiberger & Holland 2004).

Prior to the analyses described below, the data were standardized globally by the trait means. To compare among traits that are on different measurement scales (*e.g.*, pulse rate ranging from ~12 to 125 pulses per second versus pulse or chirp duty cycle ranging from 0 to 1) standardization is necessary. Standardization can have significant effects on univariate and multivariate analyses (Houle *et al.* 2011). If traits are standardized by the global mean (scaled but not centred) among species differences in the mean and variance are preserved (but proportional rather than on an absolute scale; Hansen & Houle 2008). After standardization of the data, we explored which traits had the largest contribution (on a standardized scale) to interspecific divergence using linear discriminant analysis (LDA, MASS package, Venables & Ripley 2002). We also explored the distribution of the seven species in multivariate phenotypic space using principal component analysis.

Species-specific P matrices

A Bayesian linear mixed model (MCMCglmm package, Hadfield 2010) was fitted to the song recordings. The model included carrier frequency, pulse rate, pulse duty cycle, chirp rate, and chirp duty cycle as response variables, and individual identification as a random effect. We used flat, REML-like priors here, but using an inverse-Wishart prior produced qualitatively similar results (data not shown). Convergence of the MCMC chain and independence (autocorrelation) of successive values in the output of the MCMC chain were used to assess model fit. We retained 45,000 (chain length = 500,000, burn-in= 50,000, thinning interval = 10) posterior samples of **P** for each species. The 7 x 45,000 **P** matrices were analysed to address the variability in **P** among the seven *Gryllus* species.

We used two approaches to test for variation among multiple matrices (described below) in conjunction with randomization to create a measure for the variation among matrices due to random sampling error. The variation in random matrices provided a null distribution against which the variation among the observed **P** matrices was compared. Random **P** matrices were obtained by taking a random sample (without replacement) from the (globally) standardized data across all seven species, then dividing the random data in 7 subsets corresponding to the sample sizes of the species and calculating the covariance matrix for each of the seven subsets (using the 'cov' function from the 'stats' package) 45,000 times. The seven arrays of 45,000 random **P** matrices were thus all from the same global "population" with sample sizes corresponding to actual the number of song recordings in each species.

Krzanowski multi-population comparison

We first tested the hypothesis that the primary subspace of **P** was similar among species. Using Krzanowski's multi-population subspace comparison (Aguirre *et al.* 2014) this hypothesis was addressed by comparing the eigenvalues (λ) of a common subspace among species, denoted **H**, from the observed data ($\mathbf{H}_{\mathbf{obs}}$) with those based on randomly generated data ($\mathbf{H}_{\mathbf{rand}}$). $\mathbf{H}_{\mathbf{obs}}$ was calculated following the methods in Aguirre *et al.* (2014) using the first two eigenvectors of each of the 45,000 posterior samples of each species **P**. Similarly, $\mathbf{H}_{\mathbf{rand}}$ was calculated using eigenvector one and two of each of the 45,000 variance-covariance matrices calculated for each of seven randomly drawn data sets. When the HPD_{95} (95% Highest Posterior Density) interval of either eigenvalue of the first two eigenvectors of $\mathbf{H}_{\mathbf{obs}}$ (λ_{obs}) did not overlap with the HPD_{95} interval of the corresponding eigenvalue of $\mathbf{H}_{\mathbf{ran}}$, it was concluded that the most common dimensions across **P** matrices was not part of all species-specific two-dimensional subspaces. To quantify the deviation between the common subspace and each species' **P**, the angle

between the eigenvectors for which $\lambda_{obs} < \lambda_{rand}$ and each species-specific subspace was calculated following Aguirre *et al.* (2014).

Eigentensor analysis

We then tested whether there was significant variation in the covariance structure of **P** and whether this variation could be attributed to specific song traits. To this end, the variance among the seven **P** matrices was quantified through tensor analysis (Hine *et al.* 2009, Aguirre *et al.* 2014). A first-order tensor is a vector that describes the linear relationships of its components in a single dimension. Similarly, a **P** or **G** matrix is a second-order tensor describing the variation and covariation in multiple dimensions (Hine *et al.* 2009, Aguirre *et al.* 2014). Variation among variance-covariance matrices can be described by a fourth-order covariance tensor. The tensor was obtained by calculating the (co)variance among each of the elements of the seven **P** matrices. The significance of the variation among **P** matrices was assessed by decomposing the covariance tensor and comparing the eigenvalues of the tensor of the observed **P** matrices to the eigenvalues of the tensor of **P** matrices constructed from data randomized among species. A significantly non-zero eigenvalue implied significant variation in **P** in the dimensions described by the corresponding eigentensor. This is analogous to non-zero eigenvalues of **G** or **P** matrices indicating significant additive genetic (or phenotypic) variation in that dimension (Aguirre *et al.* 2014). The eigentensor was then decomposed to examine the contribution of specific song traits to the variation described by the eigenvector of that eigentensor. The eigenvalues of an eigentensor indicate how the eigenvectors of that eigentensor vary with respect to one another across the **P** matrices. The strength of the correlation of the individual traits on each eigenvector reveals which linear combinations of traits are varying among **P** matrices (Aguirre *et al.* 2014).

Pairwise comparisons

For comparison with previous studies, we also included three methods to test for pairwise differences between matrices: Common Principal Components analysis (CPC, Flury 1988, Phillips & Arnold 1999), Krzanowski's original subspace comparison (Krzanowski 1979), and matrix correlation via Mantel tests. Details on these three methods are in the Supporting Information.

Selection response decomposition

The tests described above are informative about matrix similarity, but do not provide any information about how differences in eigenstructure of **P** may affect the evolutionary trajectory of the calling song. To provide insight in whether divergence in **P** has resulted in variation in the effects of (sexual) selection, we used the methods for selection response decomposition described in Aguirre *et al.* (2014). Selection response decomposition is commonly used to decompose the Lande equation ($\boldsymbol{\Delta z = G\beta}$; Lande 1979). Comparing the predicted selection response among species provides insight in interspecific variation in the response of the traits contained in a **G** matrix due to the effects of direct and indirect selection. This method can also be used to assess the among species variation in the direct and indirect effects of selection in changing the means of the traits captured in **P** by decomposing $\boldsymbol{s = P\beta}$ (Lande & Arnold 1983). The directional selection gradients $\boldsymbol{\beta}$ were obtained from the linear regression coefficients of linear models fitted to female preference data for the three trilling species (*G. texensis*, *G. rubens*, and G#14) in Blankers *et al.* (2015) and for three of the four chirping species (G#15, *G. lineaticeps*, *G. personatus*; Hennig, RM, Blankers T, Gray, DA, in review). The variation in the linear regression coefficients was accounted for in four selection vectors. Vector 1 and vector 2 ("trill-like" and "chirp-like") represented the 'typical' linear preferences for trilling and chirping species respectively. Two other vectors represented alternative preferences: low linear coefficients for carrier frequency, pulse rate, pulse duty cycle, and (with a negative sign) for chirp duty cycle (vector 3) and higher linear coefficients for chirp duty cycle preference but no significant linear selection on pulse rate (vector 4, Table 3). For each species the average change in each of the five traits along with the 95% HPD interval was calculated. When the HPD intervals of two species did not overlap for a particular trait, the selection response of that trait was considered significantly different between the two species.

RESULTS

The calling song has diverged in all song traits, but most strongly in pulse rate (Table 1). The seven species can be categorized coarsely into two groups based on the chirp duty cycle: there are the three trilling species, *G. texensis*, *G. rubens*, and G#14 on the one side and on the other side are the species producing short chirps followed by relatively long pauses, G#15, *G. lineaticeps*, *G. personatus*, and *G. firmus* (Table 1). The contribution of linear combinations of the song traits to interspecific divergence was examined using a linear discriminant analysis. The first linear discriminant (LD1, proportion of the

trace: 0.75) correlated strongly with variation in the pulse rate (LD coefficients in Table 1), indicating that pulse rate was the most divergent trait among species. The secondary axis of divergence (LD2, 0.24) was positively correlated with chirp duty cycle and negatively correlated with chirp rate (Table 1). Similarly, principal component analysis revealed the two major directions of phenotypic divergence among individuals. The first PC (46.0%) was strongly correlated with pulse rate (0.62) and carrier frequency (0.58) (Fig. 1a). The second PC (34.5%), strongly correlated with chirp duty cycle (0.60) and chirp rate (-0.71), separated the seven species in two groups corresponding to the chirpers (high scores on PC2) and the trillers (low scores on PC2).

Species-specific P matrices

The **P** matrices for the seven species are visualized in Fig. 1b-h as covariance-correlation matrices (covariances above and correlations below the diagonal). Although pairwise correlations between song traits varied between the species, two main patterns could be distinguished. First, pulse rate and chirp rate are always positively correlated, excepting in *G. firmus* (-0.03), but the strength of this correlation varies. Second, the trilling species always showed a negative correlation (ranging between -0.10 and -0.46) between chirp rate and chirp duty cycle, while the chirpers showed a strong, positive correlation between these traits (0.57 – 0.75). Eigenanalysis indicated that most of the phenotypic variation was in two orthogonal dimensions (*i.e.*, the two leading eigenvectors explaining 34-52% and 20-29% of the variance, respectively, Table S1). One eigenvector was typically strongly loaded by pulse rate and chirp rate, whereas the other was strongly loaded by carrier frequency, but substantial variation in the phenotypic dimension with the highest trait variance was observed (Table S1).

Table 1 Interspecific differentiation in song parameters. Mean ± standard deviation (SD), and coefficient of variation (CV, standard deviation divided by the mean, in parentheses) of the carrier frequency (cf in kHz), pulse duration (pdur in ms), pulse pause (ppau in ms), pulse rate (pr in pulses per second), pulse duty cycle (pdc as proportion), chirp duration (cdur in ms), chirp pause (cpau in ms), chirp rate (cr in chirps per second), and chirp duty cycle (cdc as proportion) of the male signal of *G. texensis*, *G. rubens*, G#14, G#15, *G. lineaticeps*, *G. personatus*, and *G. firmus* are shown. The linear discriminant (LD) coefficients are shown in the last column as (LD1 coefficient, LD2 coefficient). The proportions of the trace for LD1 and LD2 were 0.75 and 0.24, respectively. Sample size per species is given as (number of individuals, number of recordings). The song traits used in the **P** matrices are in bold.

	G. texensis (47, 85)	*G. rubens* (70, 122)	G#14 (24, 86)	G#15 (26, 70)	*G. lineaticeps* (24, 54)	*G. personatus* (10, 29)	*G. firmus* (55, 191)	LD (1,2) coefficients
cf	**5.17 ± 0.22 (0.04)**	**4.73 ± 0.27 (0.06)**	**4.49 ± 0.12 (0.03)**	**5.52 ± 0.28 (0.05)**	**5.28 ± 0.38 (0.07)**	**4.21 ± 0.24 (0.06)**	**4.15 ± 0.22 (0.05)**	0.09, -0.31
pdur	6.64 ± 1.21 (0.18)	9.59 ± 1.72 (0.18)	11.76 ± 1.43 (0.12)	5.21 ± 0.84 (0.16)	6.02 ± 1.20 (0.19)	6.91 ± 0.81 (0.12)	20.89 ± 5.29 (0.25)	
ppau	8.53 ± 1.32 (0.16)	12.53 ± 1.91 (0.15)	14.53 ± 1.46 (0.10)	5.34 ± 1.13 (0.21)	8.26 ± 1.81 (0.22)	8.03 ± 0.96 (0.17)	43.51 ± 6.49 (0.15)	
pr	**66.38 ± 5.57 (0.08)**	**45.48 ± 3.41 (0.07)**	**38.18 ± 1.73 (0.05)**	**95.28 ± 10.88 (0.11)**	**71.27 ± 9.88 (0.14)**	**67.07 ± 2.88 (0.04)**	**15.66 ± 0.8 (0.05)**	6.85, -0.01
pdc	**0.44 ± 0.08 (0.18)**	**0.43 ± 0.07 (0.16)**	**0.45 ± 0.05 (0.11)**	**0.5 ± 0.08 (0.16)**	**0.42 ± 0.08 (0.19)**	**0.46 ± 0.06 (0.13)**	**0.33 ± 0.08 (0.24)**	-0.03, -0.15
cdur	581.8 ± 286.2 (0.49)	3646 ± 1802 (0.49)	1697 ± 748 (0.44)	79.79 ± 16.21 (0.20)	86.04 ± 17.61 (0.20)	89.9 ± 14.97 (0.17)	182.3 ± 28.79 (0.16)	
cpau	304.9 ± 115.1 (0.38)	972.5 ± 753.9 (0.78)	797.7 ± 289.7 (0.36)	469.9 ± 258.5 (0.55)	382.0 ± 134.7 (0.35)	763.1 ± 279.6 (0.37)	689.3 ± 196.7 (0.29)	
cr	**1.27 ± 0.38 (0.30)**	**0.28 ± 0.16 (0.57)**	**0.48 ± 0.14 (0.29)**	**2.24 ± 0.54 (0.24)**	**2.38 ± 0.67 (0.28)**	**1.34 ± 0.45 (0.34)**	**1.24 ± 0.29 (0.23)**	-0.70, -1.48
cdc	**0.63 ± 0.11 (0.17)**	**0.78 ± 0.10 (0.13)**	**0.65 ± 0.07 (0.11)**	**0.19 ± 0.05 (0.26)**	**0.2 ± 0.03 (0.15)**	**0.13 ± 0.05 (0.38)**	**0.23 ± 0.05 (0.22)**	0.42, 2.82

.

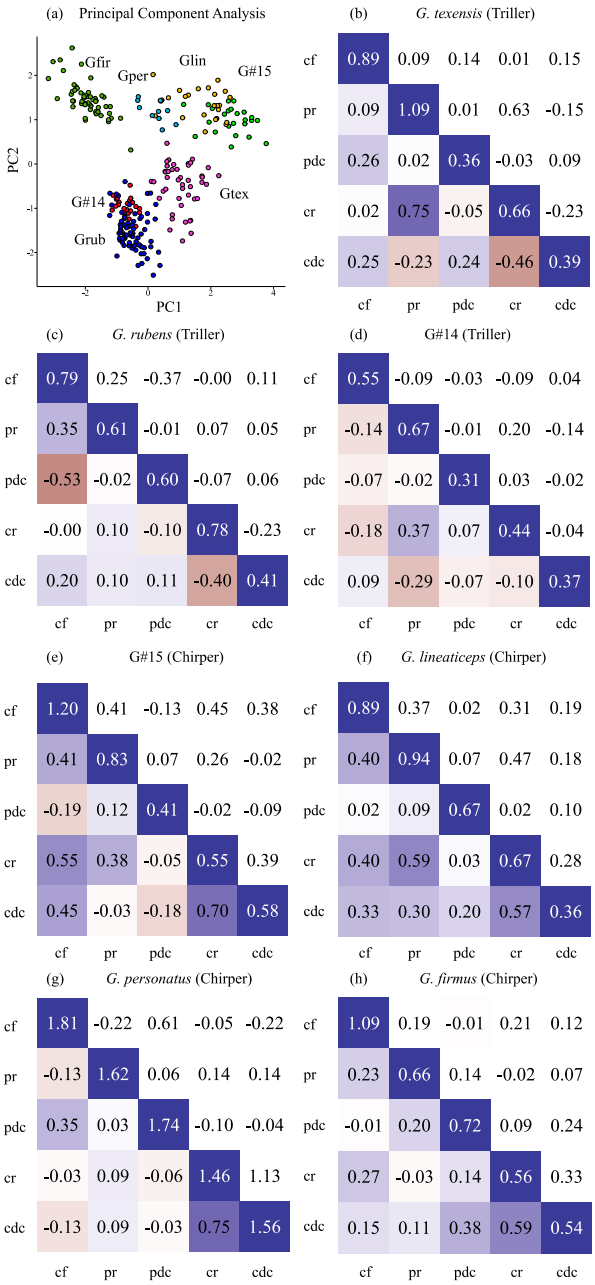

Figure 1 Variability in song traits and song **P** matrices. (a) Divergence in song traits in multivariate space. The scores of each male along the first and second principal components (PC1 and PC2, respectively) are shown for *G. texensis* (purple), *G. rubens* (blue), G#14 (red), G#15 (light green), *G. lineaticeps* (orange), *G. personatus* (light blue), *G. firmus* (dark green). (b-h) **P** matrices are shown as the correlation (below the diagonal), the variance of the standardized data (diagonal), and the covariance (above the diagonal). The intensity of the colours indicates the strength of the positive (blue) and negative (red) correlation between the song traits: carrier frequency (cf), pulse rate (pr), pulse duty cycle (pdc), chirp rate (cr), and chirp duty cycle (cdc).

Krzanowski multi-population comparison

To test whether the two leading eigenvectors were shared among all species, a multi-population extension of Krzanowski's comparison of subspaces was used (Aguirre *et al.* 2014). Comparing the first two eigenvectors of $\mathbf{H}_{\mathbf{obs}}$ (explaining 60% of the variance) with the eigenvectors for $\mathbf{H}_{\mathbf{ran}}$, indicated significant variation across the subspaces ($\mathbf{H}_{\mathbf{obs}}$: λ_{h1} = 5.05, HPD_{95}: 4.11-6.00; $\mathbf{H}_{\mathbf{ran}}$: λ_{h1} = 6.97, HPD_{95}: 6.94-7.00; Fig. 2a). By calculating the angle between the eigenvectors of $\mathbf{H}_{\mathbf{obs}}$ and the subset of the eigenvectors of each population, we observed how close the subspaces of the populations are to shared subspace $\mathbf{H}_{\mathbf{obs}}$. The strongest deviation was seen for *G. personatus* and G#14 (mean angle: 50.7° and 46.0°, respectively; Table S2). Angles were generally lower for the other species, but HPD_{95} intervals overlapped between all species. Therefore, no single species drives the significant deviation of the eigenvalues of $\mathbf{H}_{\mathbf{ran}}$ from those of $\mathbf{H}_{\mathbf{obs}}$.

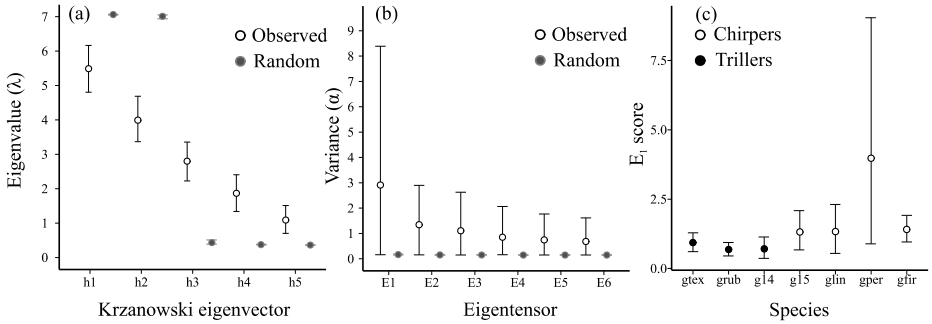

Figure 2 Multi-population comparison of **P**. (a) Krzanowski subspace comparison. The eigenvalues (mean and 95% HPD interval) of each of the five eigenvectors of **H** are shown for the observed (black, open circles) and randomized (grey, closed circles) **P**; (b) Eigentensor analysis. The variance (mean and 95% HPD interval) associated with each of the 6 eigentensors is shown for the observed (black, open circles) and randomized (grey, closed circles) **P**. (c) The scores of each of the species on the first eigentensor (**E1**) are shown for each of the seven species. Open circles indicate chirpers, closed circles indicate trillers.

Eigentensor analysis

Variance-covariance tensors can indicate whether or not significant (co)variation exists among **P** matrices. To correct for sampling error, the HPD_{95} interval of the variance accounted for by the eigentensors, α, corresponding to the observed **P** matrices was compared with randomly generated matrices. When the lower bound of the HPD_{95} for an observed eigentensor was higher than the upper bound of the HPD_{95} for a random eigentensor, we inferred that there was significant variation in **P** matrices in the direction of that eigentensor. Only the first eigentensor, E_1, showed a trend towards variation in **P** ($E_{1\_observed}$: α= 3.00, HPD_{95} = 0.02-8.95; E_{1_random}: α= 0.02, HPD_{95} = 0.00-0.04; Fig. 2b) but HPD_{95} intervals of α of the random **P** matrices overlapped with those of observed **P** matrices. The leading eigenvector of eigentensor E_1, e_{11}, strongly correlated with variation in chirp rate and chirp duty cycle (Table 2, Table S3). Furthermore, the trillers and chirpers were teased apart along the first eigentensor (E_1 scores < 1 for trillers, > 1 for chirpers), but note that HPD_{95} intervals of the coordinates on E_1 overlapped among the species (Fig. 2c). Removing *G. personatus*, which was somewhat of an outlier in this analysis, from the analysis strongly decreased the total variance along the first eigentensor, but resulted in significant variation in this dimension ($E_{1\_observed}$: α= 0.54, HPD_{95} = 0.04-1.37; E_{1_random}: α= 0.02, HPD_{95} = 0.00-0.03).

Table 2 The eigenstructure of the first eigentensor (eigenvalue = 3.00). The loadings and eigenvalues of the five eigenvectors are shown.

Eigenvector	Vector eigenvalue	cf	pr	pdc	cr	cdc
e11	-0.77	-0.21	0.12	-0.14	0.64	0.72
e12	-0.57	0.65	-0.02	0.70	0.19	0.16
e13	-0.24	0.27	-0.88	-0.33	0.16	0.02
e14	-0.10	-0.65	-0.44	0.60	-0.06	0.05
e15	0.02	-0.11	0.06	0.06	0.72	-0.67

In summary, the comparisons of **P** among multiple species showed that the subspace with the highest phenotypic variance differed (Fig. 2a). Divergence in **P** was limited and associated with high error (Fig. 2b) but contrasted chirpers and trillers (Fig. 2c). The major axis of variation among **P** matrices (*i.e.* the first eigentensor) strongly correlated with chirp rate and chirp duty cycle (Table 2). These results thus indicated that the **P** matrices have diverged somewhat among species and that the variation in the long timescale of the song rhythm (for which selection regimes differ between chirpers and trillers) rather than variation on the short timescale (for which selection regimes are similar, but among species differentiation is largest) contributed to this divergence.

Pairwise matrix comparisons

For comparison with previous studies, we also employed three pairwise methods for comparison of matrices. Krzanowski's subspace comparison (mean similarity S: 0.96 ± 0.33SD, with a possible maximum S of 2) and CPC analysis (common principal components or higher hierarchical level of similarity) showed that the principal components of **P** were shared among species, but also revealed differentiation (Table S4). Mantel tests indicated that most **P** matrices were uncorrelated (Table S5).

We then tested whether divergence in **P** was a consequence of genetic distance or whether divergence in **P** tracked divergence in song trait means. We compared the genetic distance matrix (Table S5), a Euclidean distance matrix calculated from the species averages of the songs traits, and a Euclidean

distance matrix of the species scores on the first eigentensor (Fig. 2c) pairwise using Mantel test. Neither divergence in the song traits nor in **P** were a function of genetic distance ($r = 0.353$, $p = 0.0948$; $r = 0.046$, $p = 0.3384$, respectively). Similarly, divergence in **P** was not associated with divergence in the song traits means ($r = -0.0496$, $p = 0.6021$).

Selection response decomposition

In addition to divergence in **P** among species, we were interested in possible interspecific variation in the effects of selection (Lande & Arnold 1983) on the five song traits. We used four different selection vectors which were based on the song preference data in Blankers *et al.* (2015) and in a currently unpublished study (Hennig, RM, Blankers, T, Gray, DA, in review). There were significant (*i.e.* non-zero) linear selection coefficients for carrier frequency (of magnitude 0.05-0.09), pulse rate (of magnitude 0.04-0.06), pulse duty cycle (only for chirping species, 0.05-0.09) and chirp duty cycle (0.05-0.18; Table 3). Based on these data we applied the selection response decomposition in Aguirre *et al.* (2014) using four selection vectors that encompass the full range of among species variation in directional selection gradients (Table 3). We found significant variation in the effects of selection on the chirp rate (26 pairwise comparisons) and pulse duty cycle (12 pairwise comparisons, Table 3). Differential selection effects were observed mostly between chirping and trilling species. Among the trilling species, the effects on pulse rate differed between *G. texensis* and *G. rubens* (Table 4, Table S4). Among chirping species, four comparisons indicated a significant difference in the selection effects on pulse duty cycle. The remaining 34 cases were all in comparisons contrasting chirping with trilling species. In more than 75% of those cases (26) the chirp rate was affected differently between species (Table 4, Table S4).

DISCUSSION

Comparative studies of **P** are vital to our understanding of phenotypic evolution. Very few studies have examined among species variation in the covariance structure of mating signals (but see Roff *et al.* 1999, Bertram *et al.* 2011, Roff & Fairbairn 2012, Pitchers *et al.* 2013) and this has been discussed in relation to observed selection pressures only for male cuticular hydrocarbon (CHC) variation in fruit flies (*e.g.* Blows *et al.* 2004, Hine *et al.* 2009). However, insight in multivariate phenotypic evolution of mating signals in relation to variation in sexual selection can strongly help our understanding of the magnitude and direction of mating signal divergence and the role of sexual selection in speciation.

Here we analysed the calling song of seven closely related field cricket species and revealed significant variation in **P** (Fig. 2a,b)**.** Divergence in **P** was not simply a function of genetic distance nor did it track divergence in the song means. Instead, we found that among species variation in **P** depended on the covariance patterns between the traits governing the long timescale of the song and contrasted chirpers and trillers (Fig. 2c). The observed divergence in **P** was also predicted to influence the effects of selection and, again, underlined that the largest differences in **P** were in comparisons of species with different selection regimes (Table 3, Table 4). The data presented here thus show that the **P** was more divergent between species with different selection regimes compared to species for which preference functions are of similar shape. This study provides a rare empirical example of divergence in the covariance structure of a mating signal with significant effects on the evolutionary trajectory of the signal and highlights a role for sexual selection in shaping the patterns of phenotypic covariation. Below we discuss the patterns of multivariate song divergence, the evolutionary and mechanistic factors that have contributed to divergence, and the evolutionary consequences of variation in the covariance structure.

Table 3 Decomposition of the effects of selection. The four selection vectors are based on the linear selection coefficients from a regression model fitted to song preference data. The terms in that model that had significant regression coefficients are in bold face type. Below each selection vector, the number of pairwise comparisons between the seven cricket species (21 comparisons per song trait per selection vector, 420 comparisons in total) that resulted in non-overlapping 95% HPD intervals of the predicted trait change of all posterior samples is shown. The last row shows the sum of differential responses over all four vectors.

Species	cf	pr	pdc	cr	cdc
G. texensis (triller)	**0.07**	0.00	0.00	0.02	**0.10**
G. rubens (triller)	**0.07**	**-0.04**	0.02	-0.02	**0.09**
G#14 (triller)	**0.05**	**-0.06**	-0.04	-0.01	**0.18**
G#15 (chirper)	**0.09**	0.00	**0.09**	0.02	-0.02
G. lineaticeps (chirper)	**0.06**	**0.04**	**0.06**	-0.02	**-0.05**
G. personatus (chirper)	**0.06**	0.00	**0.05**	-0.03	-0.03
Vector 1 (trill-like)	0.05	-0.05	0	0	0.15
Differential response	0	1	5	10	0
Vector 2 (chirp-like)	0.05	0	0.05	0	0
Differential response	1	0	2	4	0
Vector 3	0.05	0.05	0.05	0	-0.05
Differential response	0	0	1	2	0
Vector 4	0.05	0	0	0	0.15
Differential response	0	1	4	10	0
Total	1	2	12	26	0

Table 4 Effects of selection in chirpers versus trillers. The 21 pairwise comparisons of the predicted selection effects on each of the five traits in response to the four selection vectors were subdivided in comparisons contrasting chirping species (G#15, *G. lineaticeps*, *G. personatus*, *G. firmus*) with other chirping species (6 comparisons in total), those contrasting chirping species with trilling species (12 in total), and those contrasting trilling species (*G. texensis*, *G. rubens*, and G#14) with other trilling species. The number of pairwise comparisons that had non-overlapping 95% HPD intervals is shown. The last column contains the sum over all five traits.

Comparison	cf	pr	pdc	cr	cdc	sum
Chirp vs. Chirp (6)	0	0	4	0	0	4
Chirp vs. Trill (12)	1	0	7	26	0	34
Trill vs. Trill (3)	0	2	1	0	0	3

Variation in P

The **P** matrices differed significantly among species. We found divergence in the subspace of **P** described by the first two eigenvectors (Fig. 2a) and high variation among the **P** matrices in the direction of the first eigentensor (Fig. 2b). The pairwise methods for matrix comparisons further supported the finding that **P** has diverged among species to some extent, but also showed that most **P** matrices shared principal components (within trillers and among trillers and chirpers) or varied proportionally (among chirpers; Table S4, below the diagonal). The overall similarity of subspaces was approximately 50% of the maximum similarity possible for the two eigenvectors (Table S4, above the diagonal).

Univariate and most other multivariate analyses of variation among cricket species indicated that pulse rate is the most divergent trait among cricket species (Alexander 1962, Gerhard & Huber 2002, see also Table 1). Conversely, here the among species variation in **P** depended strongly on chirp rate and chirp duty cycle (Table 2), which were negatively correlated in species producing trills and positively correlated in chirping species (Fig. 1b-d versus Fig. 1e-h). Our results thus emphasize that variation in and selection on traits that show lower levels of interspecific divergence in mean trait values (*i.e.* chirp duty cycle in *Gryllus*) can still substantially influence the evolutionary response of the mating signal as a whole. However, we acknowledge that the selection response ultimately depends on the availability of additive genetic variation in the direction of selection.

In addition to variation in **P** among species, our results also indicate that some patterns of covariance were shared. Most notably, our results revealed a positive correlation between pulse rate and chirp rate across all species, excepting *G. firmus* (Fig. 1a, Table S1). The generality of this interaction between the short and the long timescale in crickets is further supported by the findings in Bertram *et al.* (2011), where a strong, positive correlation between pulse duration and chirp duration (note that the rate is the inverse of the sum of duration and pause) was found in three out of four species (including *G. texensis*). Coupling of pulse and chirp rate can affect the evolutionary response in different ways. In

the absence of direct selection, the chirp rate can be dragged along when females select males for higher or lower pulse rates. This would provide an explanation for the observed divergence in the chirp rate among males, despite a lack of direct selection on chirp rate in trilling species (Blankers *et al.* 2015). Alternatively, when higher chirp rates are selected against, for example because signalling at high rates is energetically costly (Hoback & Wagner 1997), divergence in pulse rate can be constrained by its covariance with chirp rate.

The role of sexual selection in shaping P

There are numerous factors that can contribute to variation in the phenotypic covariance patterns among species, including environmental variation, small population sizes, neutral processes such as mutation and drift, and selection (Roff 2000, Jones *et al.* 2003, Melo & Marroig 2014). Most of these factors are unlikely to play a major role here. Potential environmental effects were likely minimal because (i) all species were raised under the same breeding conditions, (ii) data were collected across multiple laboratory generations, and (iii) mature animals used for the song recordings were in individual containers and kept in the same room. Several observations indicate that neutral processes are also unlikely to have resulted in the observed patterns of variation. First, effective population sizes are expected to be large in crickets (Gray *et al.* 2008, Broughton & Harrison 2003) and swarms of crickets can easily consist of tens of thousands of individuals. Second, variation in **P** as a result of drift is expected to be proportional (*i.e.*, change the relative strength of the covariance, but not the orientation of the matrix; Roff 2000), which is not consistent with the patterns of variation observed here. Third, the observed variation in **P** reflects the variation in the chirp structure, which is closely linked to female preference variation and thus unlikely to be driven by neutral processes alone. Fourth, we found no evidence for phylogenetic effects on the divergence in **P**, although **P** was not compared in an explicit phylogenetic comparative context nor did the comparison encompass deep phylogenetic nodes within *Gryllus*. Comparing **P** or **G** across larger taxonomic scales (*e.g.* Haber 2014) or in a comprehensive phylogenetic framework (*e.g.*, Begin & Roff 2004, Haber 2014) can improve a cross-species comparison of phenotypic evolution. Presently, phylogenetic information for *Gryllus* is limited to a comparison of mitochondrial DNA (*16s ribosomal RNA* and *Cytochrome b* genes, 1536 combined bp) across only a subset of the extant species (Huang *et al.* 2000) and information is lacking for three of the species studied here (*G. personatus*, G#14, and G#15, but see Gray *et al.* 2015).

Because the selection landscape is less curved in the presence of directional selection, **G** (and, by extension, **P**) becomes less stable (Jones *et al.* 2003, Melo & Marroig 2014). The role of selection in shaping the patterns of covariation is particularly likely to be revealed when the populations among which **P** is compared differ in either the presence or direction of linear selection. Evidence for selection effects on variation in **P** has been shown previously for morphological traits when comparing lake-dwelling populations of the three-spined stickleback that differed in the strength of directional selection (*e.g.*, Berner *et al.* 2010, although divergence in **P** was mostly a consequence of fluctuations in the variance in one trait) and among replicate populations of Anolis lizards that differ in microhabitat use (Kolbe *et al.* 2011).

Presently, not much is known about variation in the patterns of covariance or about the role of selection in driving **P**/**G** variation for sexual signalling traits. Sexual selection is expected to deplete genetic variation in the direction of selection (Kokko *et al.* 2003) and multivariate comparative studies of CHC profile variation in *Drosophila* showed that genetic variation was strongly directed away from the major direction of linear selection (Blows *et al.* 2004, Hine *et al.* 2009). Selection poorly explained divergence in the (orientation of the) genetic covariance structure, likely because linear selection acted in the same direction in different populations (Hine *et al.* 2009).

Among the seven species analysed here linear selection coefficients varied in magnitude within trillers, but were either not significantly different from zero or of opposite sign in chirpers (Table 3) and we showed that the phenotypic covariance structure differed accordingly. Based on the higher similarity within than between selection regimes and the strong dependency of the major direction of variation in **P** (*i.e.*, the first eigentensor) on variation in a trait targeted by linear selection in one group but not the other, we conclude that the differences in the selection regime likely play a strong role in driving the divergence in **P.** To our knowledge, this study provides the first empirical evidence that the direction and presence of linear selection on mating signals can cause shifts in the orientation of **P**. Note, however, that these results are informative about the shape of the preference function (or selection landscape) and not necessarily about the type of selection: in the absence of non-zero linear selection coefficients, directional selection can still be manifested in concave preferences depending on the phenotypic distribution and breeding values of the selected trait (Ryan & Rand 1993).

To elucidate the mechanistic basis of variation in **P** as a result of directional selection, we briefly reflect on the mechanisms underlying signal production (males) and signal evaluation (females). The temporal pattern of the song in males is manifested in the central pattern generator in the nervous system (Hedwig 2000, Schöneich & Hedwig 2012). The tempo of the short (pulse) and long (chirp) timescale varies depending on independent variation in the duration of the opening (pulse pause) and closing (pulse duration) movements of the wings as well as the duration of sequential wing stroke movements (chirp duration) and the interruptions of the chirps (chirp pause). Female preferences depend on evaluation of the song signal with a short, innate template that responds best to a specific pulse rate. The output of this pulse detector is integrated over a time window constrained by the chirp duty cycle (Clemens and Hennig 2013, Hennig *et al.* 2014). This is the reason we decided to analyse the patterns of covariation in the rate and duty cycle, rather than in the duration and pauses of pulses and chirps.

However, changes in the variance and covariance of the rate and duty cycle in response to selection can only come about by changes in duration and pause of pulses and chirps. Based on the coefficients of variation (CVs) in the chirp duration and chirp pause in chirping versus trilling species (see Table 1), we present a hypothesis for the mechanisms underlying the observed differentiation in the covariance between chirp duty cycle and chirp rate. For chirpers, song pattern generation is constrained by constant chirp duration as indicated by the low CVs (Table 1). Variation in chirp rate or chirp duty cycle will therefore result mostly from variation in the chirp pause. At constant chirp duration, variation in chirp rate (inverse of the sum of duration and pause) and chirp duty cycle (duration divided by the sum of duration and pause) will therefore be positively correlated for chirping species (Fig. 3b,c). For trillers, the CVs for the trill pause are slightly lower than those for the chirp duration (except for *G. rubens*, where the chirp pause is more variable). Directional selection on chirp duty cycle in trilling species has favoured crickets singing with increasingly longer chirp duration. Different from chirpers, variation in the chirp rate and chirp duty cycle of the songs of trilling species will then result mostly from variation in the chirp duration (and not, or less, in the chirp pause). The chirp rate and duty cycle will therefore be negatively correlated (Fig. 3b,d).

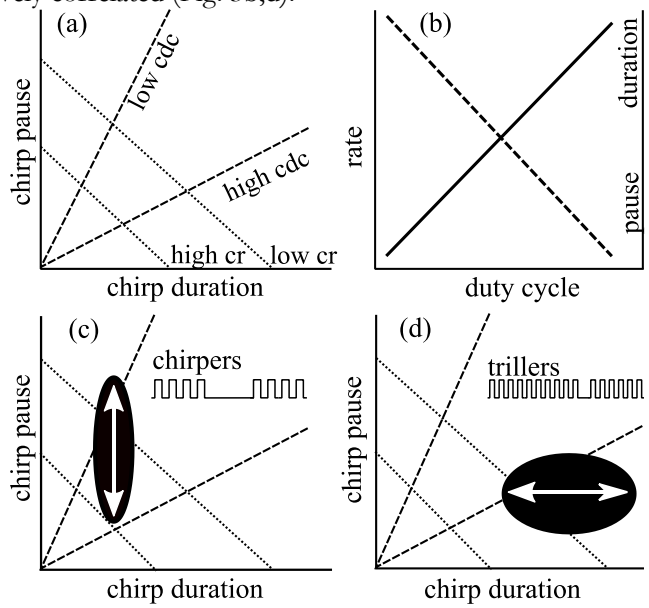

Figure 3 Schematic representation of the change in covariance in chirping versus trilling cricket species. The signal space of chirp duration and chirp pause (a) can be roughly subdivided in areas of high or low chirp rate (to the left or right of the dotted lines, respectively) and of high or low chirp duty cycle (below and above the dashed lines, respectively). Panel (b) shows how the duty cycle and rate are positively correlated (solid black line) when the duration is constant and the pause varies or negatively correlated (dashed grey line) when the pause is constant and the duration varies. In cricket species producing short chirps (c) variation is constrained by the chirp duration. The rate and duty cycle are therefore positively correlated. In trilling species (d) the chirp duration has become much more variable following directional selection for higher chirp duty cycles. The chirp pause has become more constant, resulting in a negative correlation between the rate and the duty cycle. For chirpers and trillers, schematic representations of two consecutive chirps are given in (c) and (d).

Evolutionary trajectory

Without information or stringent assumptions about **G**, the variation in **P** cannot be used to predict the response to selection. However, in combination with information about the selection gradients, variation in **P** provides insight into the predicted effects of selection (Lande & Arnold 1983). The selection response decomposition method in Aguirre *et al.* (2014) showed strong differentiation in the predicted selection effects between chippers and trillers due to the indirect effects on chirp rate (Table 3, Table 4). Our results thus indicated that the observed variation in **P** caused different species to experience similar selection gradients differently. As we used standardized variables, the effects of selection estimated here can be used to compare among species but do not predict changes in units in which the traits were originally measured.

The effects of selection were compared across 4 selection vectors and 5 signalling traits between 21 species pairs (420 pairwise comparisons). Forty-one comparisons showed no overlap in 95% HPD intervals. It is conceivable that random variation among the **P** matrices results in differences between posterior distributions simply by chance, inflating the probability of a type I error. However, we note that the selection effects differed non-randomly. Non-overlapping HPD intervals were observed almost exclusively when comparing trillers with chirpers and the vast majority revealed differential selection effects for chirp rate, rather than any of the other song traits. Additionally, most of these differences in selection were quite large (Table S6). Therefore, even if some of the cases where we observed differential selection effects between two species were simply by chance and represented false positives, we expect that the major outcome of this analysis (contrasting chirping and trilling species, highlighting the indirect effects on chirp rate) is not the result of stochasticity.

CONCLUSION

This study is one of few empirical endeavours addressing whether covariance matrices vary in a way that affects evolutionary trajectories (Aguirre *et al.* 2014) and whether co-evolutionary patterns of traits across multiple species vary in relation to selection within those species (Arnold *et al.* 2008). We provide a comprehensive overview of divergence in the multivariate calling song in crickets with respect to variation in the selection regime across seven species. Our analyses revealed a positive interaction between the long and short timescale of the mating signal, suggesting pleiotropic evolution of temporal rhythm across timescales. We also uncovered significant variation in **P** among species and showed that variation in **P** affected the predicted effects of selection and contrasted species experiencing different (sexual) selection regimes. This is a rare example of how sexual selection not only drives divergence in trait means, but alters the multivariate phenotype available for selection thereby changing the evolutionary trajectory of a behavioural mating signal.

ACKNOWLEGEMENTS

The manuscript strongly benefitted from comments by Emma Berdan, Jonas Finck, and Michael Reichert and peer review from the associate editor Derek A. Roff and two anonymous reviewers. The performed experiments comply with the "Principles of animal care", publication No. 86-23, revised 1985 of the National Institute of Health, and also with the current laws of Germany. The authors declare no conflict of interest. This study is part of the GENART project funded by the Leibniz Association (SAW-2012-MfN-3).

REFERENCES

Aguirre, J.D., Hine, E., McGuigan, K. & Blows, M.W. 2014. Comparing G: multivariate analysis of genetic variation in multiple populations. *Heredity.* **112**: 21–9.

Alexander, R.D. 1962. Evolutionary change in cricket acoustical communication. *Evolution.* **16**: 443–467.

Arnold, S.J., Bürger, R., Hohenlohe, P.A., Ajie, B.C. & Jones, A.G. 2008. Understanding the evolution and stability of the G-matrix. *Evolution.* **62**: 2451–2461.

Bégin, M. & Roff, D.A. 2004. From micro- to macroevolution through quantitative genetic variation: positive evidence from field crickets. *Evolution* **58**: 2287–2304.

Bentsen, C.L., Hunt, J., Jennions, M.D. & Brooks, R. 2006. Complex multivariate sexual selection on male acoustic signaling in a wild population of *Teleogryllus commodus*. *Am. Nat.* **167**: E102–E116.

Berner, D., Stutz, W.E. & Bolnick, D.I. 2010. Foraging trait (co)variances in stickleback evolve deterministically and do not predict trajectories of adaptive diversification. *Evolution.* **64**: 2265–2277.

Bertram, S.M., Fitzsimmons, L.P., McAuley, E.M., Rundle, H.D. & Gorelick, R. 2011. Phenotypic covariance structure and its divergence for acoustic mate attraction signals among four cricket species. *Ecol. Evol.* **2**: 181–195.

Blankers, T., Hennig, R.M. & Gray, D.A. 2015. Conservation of multivariate female preference functions and preference mechanisms in three species of trilling field crickets. *J. Evol. Biol.* **28**: 630–641.

Blows, M.W., Chenoweth, S.F. & Hine, E. 2004. Orientation of the genetic variance-covariance matrix and the fitness surface for multiple male sexually selected traits. *Am. Nat.* **163**: 329–340.

Blows, M.W. & Higgie, M. 2003. Genetic constraints on the evolution of mate recognition under natural selection. *Am. Nat.* **161**: 240–253.

Broughton, R.E. & Harrison, R.G. 2003. Nuclear gene genealogies reveal historical, demographic and selective factors associated with speciation in field crickets. *Genetics* **163**: 1389–1401.
Cheverud, J.M. 1988. A comparison of genetic and phenotypic correlations. *Evolution.* **42**: 958–968.
Clemens, J. & Hennig, R.M. 2013. Computational principles underlying the recognition of acoustic signals in insects. *J. Comput. Neurosci.* **35**: 75–85.
Flury, B. 1988. *Common principal components and related multivariate models*. Wiley, New York.
Gerhardt, H.C. & Brooks, R. 2009. Experimental analysis of multivariate female choice in gray treefrogs (Hyla versicolor): Evidence for directional and stabilizing selection. *Evolution.* **63**: 2504–2512.
Gerhardt, H.C. & Huber, F. 2002. *Acoustic communication in insects and anurans*. The University of Chicago Press, Chicago.
Gray, D.A. & Cade, W.H. 2000. Sexual selection and speciation in field crickets. *Proc Natl Acad Sci* **97**: 14449–14454.
Gray, D.A., Huang, H. & Knowles, L.L. 2008. Molecular evidence of a peripatric origin for two sympatric species of field crickets (Gryllus rubens and G. texensis) revealed from coalescent simulations and population genetic tests. *Mol. Ecol.* **17**: 3836–3855.
Grobe, B., Rothbart, M.M., Hanschke, A. & Hennig, R.M. 2012. Auditory processing at two timescales by the cricket Gryllus bimaculatus. *J. Exp. Biol.* **215**: 1681–1690.
Haber, A. 2014. The Evolution of Morphological Integration in the Ruminant Skull. *Evol. Biol.* **42**: 99–114.
Hadfield, J.D. 2010. MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *J. Stat. Softw.* **33**: 1–22.
Hansen, T.F. & Houle, D. 2008. Measuring and comparing evolvability and constraint in multivariate characters. *J. Evol. Biol.* **21**: 1201–1219.
Hazel, L.N., Dickerson, G.E. & Freeman, A.E. 1994. The selection index--then, now, and for the future. *J. Dairy Sci.* **77**: 3236–3251.
Hedwig, B. 2000. Control of cricket stridulation by a command neuron: efficacy depends on the behavioral state. *J. Neurophysiol.* **83**: 712–722.
Heiberger, R.M. & Holland, B. 2004. *Statistical analysis and data display: an intermediate course with examples in S-plus, R, and SAS*. Springer Texts in Statistics, Springer, New York.
Hennig, R.M., Heller, K.-G. & Clemens, J. 2014. Time and timing in the acoustic recognition system of crickets. *Front. Physiol.* **5**.
Hine, E., Chenoweth, S.F., Rundle, H.D. & Blows, M.W. 2009. Characterizing the evolution of genetic variance using genetic covariance tensors. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **364**: 1567–1578.
Hoback, W.W. & Wagner Jr, W.E. 1997. The energetic cost of calling in the variable field cricket, Gryllus lineaticeps. *Physiol. Entomol.* **22**: 286–290.
Houle, D., Pelabon, C., Wagner, G.P. & Hansen, T.F. 2011. Measurement and meaning in Biology. *Q. Rev. Biol.* **86**: 3–34.
Huang, Y., Ortí, G., Sutherlin, M., Duhachek, A. & Zera, A. 2000. Phylogenetic relationships of north American field crickets inferred from mitochondrial DNA data. *Mol. Phylogenet. Evol.* **17**: 48–57.
Jones, A.G., Arnold, S.J. & Bürger, R. 2003. Stability of the g -matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution.* **57**: 1747–1760.
Kokko, H., Brooks, R., Jennions, M.D. & Morley, J. 2003. The evolution of mate choice and mating biases. *Proc. Biol. Sci.* **270**: 653–664.
Kolbe, J.J., Revell, L.J., Szekely, B., Brodie III, E.D. & Losos, J.B. 2011. Convergent evolution of phenotypic integration and its alignment with morphological diversification in caribbean Anolis ecomorphs. *Evolution.* **65**: 3608–3624.
Krzanowski, W.J. 1979. Between groups comparision of principal components. *J Am Stat Assoc* **74**: 703–707.
Krzanowski, W.J. 2000. *Principles of multivariate analysis: a user's perspective*. Oxford University Press.
Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution.* **33**: 402–416.
Lande, R. & Arnold, S.J. 1983. The measurement of selection on correlated characters. *Evolution.* **37**: 1210–1226.
Laughlin, D.C. & Messier, J. 2015. Fitness of multidimensional phenotypes in dynamic adaptive landscapes. *Trends Ecol. Evol.* **30**: 1–10.
Lynch, M. & Walsh, B. 1998. *Genetics and analysis of quantitative Traits*. Sinauer, Sunderland, MA.
Melo, D. & Marroig, G. 2014. Directional selection can drive the evolution of modularity in complex traits. *Proc. Natl. Acad. Sci.* **112**.
Otte, D. 1992. Evolution of cricket songs. *J. Orthoptera Res.* **1**: 25–49.
Phillips, P.C. & Arnold, S.J. 1999. Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury hierarchy. *Evolution.* **53**: 1506–1515.
Pitchers, W.R., Brooks, R., Jennions, M.D., Tregenza, T., Dworkin, I. & Hunt, J. 2013. Limited plasticity in the phenotypic variance-covariance matrix for male advertisement calls in the black field cricket, *Teleogryllus* commodus. *J. Evol. Biol.* **26**: 1060–1078.
R Development Core Team, R. 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.
Ritchie, M.G. 2007. Sexual Selection and Speciation. *Annu. Rev. Ecol. Evol. Syst.* **38**: 79–102. Annual Reviews.
Rodríguez, R.L., Hallett, A.C., Kilmer, J.T. & Fowler-Finn, K.D. 2013. Curves as traits: genetic and environmental variation in mate preference functions. *J. Evol. Biol.* **26**: 434–442.
Roff, D. 2000. The evolution of the G matrix: Selection or drift? *Heredity.* **84**: 135–142.
Roff, D., Mousseau, T. & Howard, D. 1999. Variation in genetic architecture of calling song among populations of Allonemobius socius, A. fasciatus, and a hybrid population: drift or selection? *Evolution.* **53**: 216–224.
Roff, D.A. & Fairbairn, D.J. 2012. The evolution of trade-offs under directional and correlational selection. *Evolution.* **66**: 2461–2474.
Rothbart, M.M. & Hennig, R.M. 2012. The Steppengrille (*Gryllus spec./assimilis*): Selective Filters and Signal Mismatch on Two Timescales. *PLoS One* **7**: e43975.
Ryan, M.J. & Rand, A.S. 1993. Species Recognition and Sexual Selection as a Unitary Problem in Animal Communication. *Evolution.* **47**: 647–657.
Sakaguchi, K.M. & Gray, D.A. 2011. Host song selection by an acoustically orienting parasitoid fly exploiting a multispecies assemblage of cricket hosts. *Anim. Behav.* **81**: 851–858.

Schöneich, S. & Hedwig, B. 2012. Cellular basis for singing motor pattern generation in the field cricket (Gryllus bimaculatus DeGeer). *Brain Behav.* **2**: 707–725.
Steppan, S.J. 1997. Phylogenetic analysis of phenotypic covariance structure. I. Contrasting results from matrix correlation and common principal component analysis. *Evolution.* **51**: 571–586.
Steppan, S.J., Phillips, P.C. & Houle, D. 2002. Comparative quantitative genetics: Evolution of the G matrix. *Trends Ecol. Evol.* **17**: 320–327.
Swenson, N.G. 2014. *Functional and Phylogenetic Ecology in R.* Springer, New York, NY, U.S.A.
Turelli, M. 1988. Phenotypic Evolution, Constant Covariances, and the Maintenance of Additive Variance. *Evolution.* **42**: 1342–1347.
Venables, W.N. & Ripley, B.D. 2002. *Modern applied statistics with S*. Springer, New York, NY.
Wagner, W.E. & Basolo, A.L. 2007. The relative importance of different direct benefits in the mate choices of a field cricket. *Evolution.* **61**: 617–622.
Walker, T.J. 2015. Crickets. In Singing Insects of North America. http://entnemdept.ifas.ufl.edu/walker/Buzz/crickets.htm.
Willis, J.H., Coyne, J.A. & Kirkpatrick, M. 1991. Can One Predict the Evolution of Quantitative Characters Without Genetics. *Evolution.* **45**: 441–444.

SUPPORTING INFORMATION

- **Supplementary methods for pairwise comparisons of matrices**
- **Supplementary tables S1 – S6**
- **Supplementary references**

Supplementary methods for pairwise comparisons of matrices

Common principal component analysis

To examine matrix similarity between the species pairs, we used two common methods for pairwise matrix comparison. First, the eigenstructure of **P** was compared pairwise between the species using Flury's method of hierarchical subspace comparison as implemented in Common Principal Component analysis (CPC, Phillips & Arnold 1991). The 'CPC program' (Phillips 1998) was used to compare the **P** matrices pairwise in a 'jump-up' approach': each model of higher hierarchy (*i.e.*, 'common principal components', 'proportionality', 'equality', respectively) was compared against the lowest model ('unrelated') with a χ-square test. This test is informative about the eigenstructure of **P**, but does not take into account the amount of variation each eigenvector represents (*i.e.*, the eigenvalue): similarity between two matrices can be inferred even if the common principal components represent different magnitudes of variation (Blows *et al.* 2004).

Krzanowski's subspace comparison

The second method we employed for pairwise comparison of phenotypic covariance was Krzanowski's geometric subspace comparison (Krzanowski 1979). The overall similarity, the sum of the eigenvalues of a matrix S was calculated as

$$S = A^T B B^T A \quad (1)$$

The matrices A and B represented the first two eigenvectors of each of the two species in a pairwise comparison. If the two eigenvectors are perfectly aligned (*i.e.*, the angle between the vectors is 0), the sum of the eigenvalues of *S* takes a maximum value of 2. The differences between each of the subspaces of a species pair were examined by calculating the angle between the eigenvectors, α, as

$$\alpha = cos^{-1}\sqrt{k_A^T k_B k_A k_B^T} \quad (2)$$

where k_A and k_B represent nearest eigenvectors in the matrices A and B, respectively. For more detail concerning this method see Blows *et al.* (2004).

Mantel test for phylogenetic comparison

Mantel tests were performed using the zt software with exact permutation tests of significance (Bonnet & Van de Peer, 2002). Tamura-Nei genetic distances were calculated in Geneious v 6.1.8 on a single concatenated alignment of 3041 bp with a single consensus sequence per species per gene region (D. A. Gray, unpublished data; alignment available upon request, dave.gray@csun.edu).

Supplementary references

Blows, M. W., S. F. Chenoweth, and E. Hine. 2004. Orientation of the genetic variance-covariance matrix and the fitness surface for multiple male sexually selected traits. *Am. Nat.* 163:329–340.
Bonnet, E. & Van de Peer, Y. 2002. zt: a software tool for simple and partial Mantel tests. *J. Stat Soft.* 7: 1-12.

Krzanowski, W. J. 1979. Between groups comparision of principal components. *J Am Stat Assoc* 74:703–707.
Phillips, P. C. 1998. *CPC: Common Principal Components Analysis*. Wiley, NY, USA
Phillips, P. C., and S. J. Arnold. 1999. Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury hierarchy. *Evolution*. 53:1506–1515.

Table S1 Eigenvalues and vectors for *G. texensis*, *G. rubens*, G#14, G#15, *G. lineaticeps*, *G. personatus*, and *G. firmus*. The eigenvalue (λ, in italics) of each of the five eigenvectors and the loadings of the song traits (cf: carrier frequency; pr: pulse rate; pdc: pulse duty cycle; cr: chirp/trill rate, cdc: chirp duty cycle) are shown. High loadings (> 0.40) are shown in bold face type.

		G. texensis	*G. rubens*	G#14	G#15	*G. lineaticeps*	*G. personatus*	*G. firmus*
ev1	λ	*1.60*	*1.17*	*0.88*	*1.87*	*1.72*	*2.81*	*1.33*
	cf	0.07	**-0.77**	0.36	**-0.74**	**-0.53**	0.39	**0.75**
	pr	**0.79**	-0.37	**-0.76**	-0.39	**-0.62**	-0.20	0.29
	pdc	-0.01	0.51	-0.04	0.08	-0.08	0.29	0.26
	cr	**0.58**	-0.10	**-0.46**	**-0.43**	**-0.51**	**-0.58**	0.39
	cdc	-0.20	-0.07	0.28	-0.34	-0.26	**-0.62**	0.37
ev2	λ	*0.98*	*0.90*	*0.51*	*0.80*	*0.69*	*2.27*	*0.92*
	cf	**0.92**	0.14	**0.90**	-0.05	-0.26	**0.61**	**0.57**
	pr	0.06	0.00	0.39	**0.79**	0.04	0.00	-0.02
	pdc	0.26	0.10	-0.13	0.28	**0.95**	**0.61**	**-0.66**
	cr	-0.08	**-0.88**	-0.01	-0.14	-0.02	0.36	-0.23
	cdc	0.29	**0.45**	-0.12	**-0.53**	0.15	0.34	**-0.43**
ev3	λ	*0.37*	*0.61*	*0.36*	*0.47*	*0.57*	*1.64*	*0.70*
	cf	0.39	-0.03	0.15	**0.53**	**0.79**	0.17	0.01
	pr	-0.35	**0.79**	-0.10	-0.16	**-0.49**	**-0.91**	**-0.75**
	pdc	**-0.54**	**0.57**	-0.01	**-0.50**	0.23	-0.32	-0.32
	cr	0.21	0.14	**0.70**	**-0.51**	-0.28	0.14	**0.52**
	cdc	**-0.63**	0.17	**0.69**	**-0.43**	0.03	0.12	0.26
ev4	λ	*0.28*	*0.30*	*0.33*	*0.32*	*0.40*	*1.11*	*0.42*
	cf	0.04	-0.28	-0.19	**-0.42**	-0.15	**0.65**	0.33
	pr	0.28	**0.42**	0.29	0.35	**-0.57**	0.35	**-0.58**
	pdc	**-0.80**	-0.30	**-0.79**	**-0.82**	-0.10	**-0.65**	**0.60**
	cr	-0.25	**-0.43**	-0.33	0.18	**0.57**	0.10	-0.34
	cdc	**0.47**	**-0.68**	**0.40**	0.08	**0.57**	-0.10	-0.27
ev5	λ	*0.15*	*0.21*	*0.26*	*0.11*	*0.17*	*0.36*	*0.19*
	cf	0.05	**0.56**	0.03	-0.03	-0.05	0.12	0.07
	pr	**0.41**	-0.24	**-0.41**	0.29	0.22	0.03	-0.11
	pdc	0.08	**0.56**	**-0.60**	0.05	-0.15	-0.08	-0.19
	cr	**-0.75**	-0.12	**0.44**	**-0.71**	**-0.59**	**-0.71**	**-0.64**
	cdc	**-0.51**	**-0.55**	**-0.52**	**0.64**	**0.76**	**0.69**	**0.73**

Table S2 Angles between the first two eigenvectors of each species **P** and the posterior mean of Krzanowski's H_{obs}. The posterior mean and 95% HPD intervals are shown for each angle between the two leading eigenvectors of **P** for each species and H_{obs}.

	Ev1 versus H			Ev2 versus H		
	lower	mean	upper	lower	mean	upper
G. texensis	14.3	19.2	82.0	19.7	25.0	75.9
G. rubens	30.4	33.5	70.2	32.7	53.2	71.7
G#14	38.0	46.0	67.1	17.2	29.0	78.7
G#15	14.4	17.0	82.3	9.40	51.7	85.6
G. lineaticeps	8.50	22.8	86.1	26.8	51.1	75.5
G. personatus	29.6	50.7	63.6	17.6	51.2	84.7
G. firmus	22.3	32.4	74.8	23.3	60.8	81.2

Table S3 Eigentensors with corresponding eigenvalues and eigenvectors. For each of the 6 non-zero eigentensors the eigenvalues and the loadings and eigenvalues of the five eigenvectors are shown.

Tensor	vector	Tensor eigenvalue	Vector eigenvalue	cf	pr	pdc	cr	cdc
1	e11	3.00	-0.77	-0.21	0.12	-0.14	0.64	0.72
	e12	3.00	-0.58	0.66	-0.02	0.71	0.20	0.16
	e13	3.00	-0.25	0.27	-0.89	-0.33	0.16	0.02
	e14	3.00	-0.10	-0.66	-0.44	0.60	-0.07	0.05
	e15	3.00	0.02	-0.11	0.06	0.06	0.72	-0.68
2	e21	1.30	-0.80	-0.63	-0.24	-0.48	-0.41	-0.38
	e22	1.30	0.60	0.51	-0.18	0.33	-0.54	-0.56
	e23	1.30	0.04	-0.25	0.89	0.22	-0.29	-0.11
	e24	1.30	0.04	0.14	-0.06	-0.15	-0.67	0.71
	e25	1.30	-0.01	0.51	0.33	-0.77	0.07	-0.17
3	e31	1.04	-0.74	0.70	-0.07	0.70	-0.09	-0.10
	e32	1.04	0.65	-0.15	-0.06	-0.05	-0.68	-0.71
	e33	1.04	-0.14	-0.23	0.91	0.31	0.09	-0.14
	e34	1.04	-0.04	-0.11	0.10	0.13	-0.70	0.68
	e35	1.04	0.03	0.65	0.40	-0.62	-0.15	0.01
4	e41	0.76	0.82	0.12	-0.80	-0.03	-0.51	-0.29
	e42	0.76	-0.56	-0.02	-0.55	-0.21	0.56	0.58
	e43	0.76	-0.11	-0.34	0.11	-0.92	-0.14	-0.11
	e44	0.76	-0.03	0.93	0.14	-0.33	-0.05	0.10
	e45	0.76	0.03	-0.11	0.11	0.07	-0.64	0.75
5	e51	0.65	-0.59	-0.55	0.62	-0.41	-0.20	-0.32
	e52	0.65	0.55	-0.41	-0.60	-0.62	0.17	0.24
	e53	0.65	0.47	0.44	0.45	-0.44	0.52	0.36
	e54	0.65	-0.36	0.58	-0.20	-0.48	-0.38	-0.50
	e55	0.65	-0.05	0.07	0.12	-0.10	-0.72	0.68
6	e61	0.58	0.66	0.18	-0.37	-0.75	-0.30	-0.41
	e62	0.58	-0.57	0.73	-0.04	-0.23	0.46	0.44
	e63	0.58	-0.41	0.35	-0.70	0.57	-0.26	-0.07
	e64	0.58	0.28	-0.56	-0.61	-0.18	0.41	0.34
	e65	0.58	0.00	-0.05	0.03	-0.15	-0.67	0.72

.

Table S4 Pairwise comparison of P matrices for male calling song traits. Below the diagonal, the Flury hierarchical decomposition (equal, proportional, common principal components [CPC], 3/2/1 common principal components [CPC3,CPC2,CPC1], unrelated; Phillips & Arnold 1999) is shown for each pairwise comparison. Above the diagonal, the Krzanowski comparison of subspaces (Krzanowski 2000) is shown as the degree of similarity based on the first two eigenvectors with the angles between the first eigenvectors and the second eigenvectors of each species in parentheses.

	Gtex	Grub	G#14	G#15	Glin	Gper	Gfir
Gtex		0.97 (43.0; 48.7)	1.73 (11.5; 28.9)	1.30 (10.3; 54.8)	0.96 (13.1; 83.7)	1.04 (32.7; 54.6)	1.00 (11.9; 77.7)
Grub	CPC2		1.02 (30.0; 59.0)	0.70 (33.1; 88.9)	0.93 (15.8; 89.8)	0.10 (73.7; 81.6)	0.83 (24.3; 90.0)
G#14	CPC	CPC		1.28 (13.0; 54.8)	0.77 (31.3; 78.0)	0.73 (35.6; 75.4)	1.01 (17.8; 71.6)
G#15	CPC1	CPC	CPC		1.03 (14.1; 72.9)	0.62 (41.8; 74.6)	0.96 (13.3; 84.6)
Glin	CPC	CPC3	CPC	Equal		0.72 (37.0; 73.3)	1.40 (21.3; 43.0)
Gper	CPC	CPC3	Prop	Prop	Prop		1.07 (18.2; 66.1)
Gfir	CPC1	CPC3	CPC3	CPC2	CPC3	Equal	

Table S5 Matrix correlations and genetic distances among the seven *Gryllus* species studied here. Below the diagonal are the pairwise values of the matrix correlation, *r*, of the song trait phenotypic correlation matrices, followed by the associated Mantel test *p* value from exact permutation; values in bold show significant or nearly significant correlations among species' phenotypic correlation matrices. Above the diagonal are Tamura-Nei genetic distances based on a concatenated alignment of partial sequences from the mitochondrial 16s gene (545 aligned bp), and three nuclear genes, Internal-transcribed Spacer 2 (710 aligned bp, no data for *G. rubens* or *G. firmus*), 28s ribosomal rna gene (1229 aligned bp), and Elongation factor 1-alpha (588 aligned bp).

	Gfir	G#15	Glin	Gper	G#14	Grub	Gtex
Gfir		0.01003	0.01136	0.00916	0.01223	0.00871	0.01485
G#15	0.380 (0.1167)		0.00471	0.00675	0.02545	0.01346	0.02753
Glin	0.280 (0.2083)	**0.763 (0.0667)**		0.00572	0.02682	0.01348	0.02820
Gper	0.135 (0.375)	0.022 (0.4833)	0.144 (0.2917)		0.02542	0.01215	0.02610
G#14	-0.350 (0.1833)	0.239 (0.2000)	0.073 (0.4250)	-0.094 (0.4667)		0.01215	0.00980
Grub	-0.113 (0.3667)	0.195 (0.2667)	0.331 (0.1750)	**-0.768 (0.0083)**	0.119 (0.3917)		0.01126
Gtex	**-0.660 (0.0167)**	-0.126 (0.3500)	-0.037 (0.5000)	-0.238 (0.3167)	**0.681 (0.0667)**	0.262 (0.2667)	

Table S6 Selection response decomposition. The mean and (lower, upper) 95% HPD interval is shown for each species and each trait. Trait abbreviations as in Table 1.

	cf	pr	pdc	cr	cdc
vector 1	0.05	-0.05	0	0	0.15
G. texensis	0.06 (0.02,0.1)	-0.07 (-0.12,-0.03)	0.02 (0,0.04)	-0.07 (-0.1,-0.03)	0.07 (0.05,0.1)
G. rubens	0.04 (0.03,0.06)	-0.01 (-0.03,0)	-0.01 (-0.02,0)	-0.04 (-0.07,-0.01)	0.06 (0.04,0.09)
G#14	0.04 (0.01,0.07)	-0.06 (-0.12,-0.01)	0 (-0.02,0.01)	-0.02 (-0.07,0.02)	0.06 (0.03,0.1)
G#15	0.1 (0.04,0.16)	-0.02 (-0.1,0.04)	-0.02 (-0.05,0)	0.07 (0.03,0.11)	0.11 (0.06,0.16)
G. lineaticeps	0.05 (0.02,0.1)	0 (-0.05,0.05)	0.01 (0,0.02)	0.03 (0,0.08)	0.06 (0.03,0.09)
G. personatus	0.07 (-0.15,0.31)	-0.07 (-0.27,0.07)	0.02 (-0.19,0.24)	0.16 (0.01,0.41)	0.22 (0.05,0.48)
G. firmus	0.06 (0.03,0.1)	-0.01 (-0.04,0.01)	0.03 (0.01,0.05)	0.06 (0.04,0.09)	0.08 (0.06,0.11)
vector 2	0.05	0	0.05	0	0
G. texensis	0.05 (0.03,0.08)	0 (0,0.01)	0.02 (0.02,0.04)	0 (0,0)	0.01 (0,0.02)
G. rubens	0.02 (0.01,0.03)	0.01 (0,0.02)	0.01 (0,0.02)	0 (-0.01,0)	0.01 (0,0.02)
G#14	0.03 (0.01,0.04)	0 (-0.02,0.01)	0.01 (0,0.03)	0 (-0.01,0.01)	0 (-0.01,0.01)
G#15	0.05 (0.03,0.09)	0.02 (0.01,0.04)	0.01 (0,0.03)	0.02 (0.01,0.04)	0.01 (0,0.03)
G. lineaticeps	0.05 (0.02,0.08)	0.02 (-0.01,0.06)	0.03 (0.02,0.06)	0.02 (-0.01,0.05)	0.01 (0,0.03)
G. personatus	0.12 (0.02,0.29)	-0.01 (-0.05,0.03)	0.12 (0.02,0.27)	-0.01 (-0.15,0.12)	-0.01 (-0.15,0.11)
G. firmus	0.05 (0.03,0.08)	0.02 (0,0.03)	0.04 (0.03,0.05)	0.01 (0,0.03)	0.02 (0.01,0.03)
vector 3	0.05	0.05	0.05	0	-0.05
G. texensis	0.04 (0.03,0.06)	0.07 (0.03,0.1)	0 (0,0.01)	0.04 (0.02,0.07)	-0.02 (-0.03,-0.01)
G. rubens	0.05 (0.03,0.07)	0.04 (0.03,0.06)	-0.02 (-0.04,-0.01)	0.01 (0.01,0.02)	-0.01 (-0.02,0)
G#14	0.02 (-0.01,0.05)	0.04 (0.01,0.06)	0 (-0.01,0.01)	0.01 (-0.01,0.02)	-0.02 (-0.04,-0.01)
G#15	0.06 (0.03,0.1)	0.06 (0.03,0.1)	0 (-0.01,0.02)	0.02 (-0.01,0.04)	-0.01 (-0.04,0.02)
G. lineaticeps	0.05 (0.02,0.1)	0.06 (0.02,0.11)	0 (-0.02,0.01)	0.03 (-0.01,0.07)	0 (-0.03,0.03)
G. personatus	0.09 (0.01,0.22)	0.06 (-0.03,0.2)	0.04 (-0.06,0.16)	-0.05 (-0.18,0.03)	-0.08 (-0.22,0)
G. firmus	0.06 (0.04,0.08)	0.04 (0.03,0.05)	-0.01 (-0.02,0.01)	-0.01 (-0.02,0.01)	-0.02 (-0.03,0)
vector 4	0.05	0	0	0	0.15
G. texensis	0.07 (0.03,0.11)	-0.02 (-0.04,0)	0.02 (0,0.04)	-0.03 (-0.05,-0.02)	0.07 (0.04,0.09)
G. rubens	0.06 (0.04,0.08)	0.02 (0,0.04)	-0.01 (-0.03,0.01)	-0.03 (-0.06,-0.01)	0.07 (0.05,0.09)
G#14	0.03 (0.01,0.06)	-0.03 (-0.07,0.01)	0 (-0.02,0.01)	-0.01 (-0.05,0.03)	0.06 (0.03,0.09)
G#15	0.12 (0.04,0.2)	0.02 (-0.04,0.08)	-0.02 (-0.04,0)	0.08 (0.04,0.13)	0.11 (0.06,0.16)
G. lineaticeps	0.07 (0.02,0.13)	0.05 (-0.02,0.12)	0.02 (0,0.03)	0.06 (0.01,0.12)	0.06 (0.03,0.1)
G. personatus	0.06 (-0.18,0.31)	0.01 (-0.09,0.11)	0.02 (-0.18,0.25)	0.17 (0,0.44)	0.22 (0.05,0.51)
G. firmus	0.07 (0.03,0.12)	0.02 (0,0.04)	0.04 (0.02,0.06)	0.06 (0.04,0.09)	0.09 (0.06,0.11)

CHAPTER 5

THE EVOLUTIONARY HISTORY OF TWO SYMPATRIC CRICKET SPECIES WITH DIVERGENT MATING SIGNALS: GENETIC DIVERGENCE AND THE ROLE OF GENE FLOW AND SEXUAL SELECTION IN SPECIATION

Based on: Blankers, T, Berdan, EL, Vilaça, ST, Waurick, I, Gray, DA, Mazzoni, CJ, Mayer, F. 2015. The evolutionary history of two sympatric cricket species with divergent mating signals: genetic divergence and the role of gene flow and sexual selection in speciation. In preparation.

Abstract. The role of sexual selection and geography in speciation is not trivial and we generally lack insight into genome-wide genetic variation, population size variation and the extent and duration of gene flow during divergence in model systems for sexual selection research. Here we used RNA-seq to access functionally relevant genomic SNP data and studied the patterns of genetic divergence, the presence and timing of bottlenecks and gene flow, and signatures of selection in the sympatric field crickets *Gryllus rubens* and *G. texensis*. These species have primarily diverged in mating behaviour and are ecologically and morphologically cryptic. We thus expect a strong contribution to reproductive isolation of genes controlling the production of acoustic and chemical sexual signals. We found a slightly skewed allele frequency distribution with a pronounced right tail. Contigs showing elevated levels of divergence were enriched for biological processes that play a role in acoustic and chemical mate signalling. cuticular hydrocarbon synthesis, wing vein morphogenesis, muscle development, and neuromuscular junction development. A custom pipeline was used to retain haplotype sequences from phased SNP data which were then used to infer the demographic history under the Approximate Bayesian Computation framework. We found strong support for a long history of ancient gene flow, which ceased some time during the last Pleistocene glacial cycles, and a bottleneck in the derived species (*G. rubens*). A strong role for sexual selection in facilitating speciation with gene flow was further supported by a high rate of nonsynonymous substitutions in contigs showing similarities to genes underlying neuronal mechanisms used in song generation. These results provide unprecedented insight into the evolutionary history of a classic model system for sexual selection research and contribute to a synthesis of sexual selection and speciation in the face of gene flow.

INTRODUCTION

Speciation is a complex process in which a single population splits into two groups that become reproductively isolated from each other. Fundamental to our understanding of speciation is (1) determining the processes and (2) finding the genes or genomic regions associated with the establishment of reproductive isolation (RI; Gavrilets 2003, Feder *et al.* 2013). Of central interest are populations that diverge in the face of gene flow (Sousa & Hey 2013), because gene flow tends to homogenize genetic variation throughout the genome except in regions harbouring loci involved in RI (Wu 2001). In order to understand the evolutionary history and interpret the patterns of genetic variation between species, the effects of selection and demography have to be disentangled and the relative contribution of the two towards divergence addressed.

The use of next generation sequencing (NGS) and development of statistical tools for genome-wide data analysis are revolutionizing speciation research (Metzker 2010). Genomic scans have been used to identify candidate genes involved in the evolution of RI by searching for genes or genomic regions showing elevated divergence (Ellegren *et al.* 2012, Ruegg *et al.* 2014) or by exploring those genes that have an exceptionally high rate of non-synonymous single nucleotide polymorphisms (SNPs) (Nielsen 2005). Meanwhile, combining genome-wide data with coalescent simulations is advancing our knowledge of the historical demography of species and the role of gene flow during divergence (*e.g.* Morgan *et al.* 2010, Wegmann & Excoffier 2010, Li & Durbin 2011, Nadachowska-Brzyska *et al.* 2013).

One of the most flexible approaches based on coalescent theory is approximate Bayesian computation (ABC, Beaumont *et al.* 2002). In the ABC framework, users define demographic scenarios, simulate the observed genetic data under each of these scenarios, and assess the proximity of simulations to the observed data using summary statistics (Bertorelle *et al.* 2010). So far, analyses implementing genome-wide data in coalescent theory have mostly been restricted to organisms for which ample genomic resources were available. However, decreased costs and improved computational power have rendered the use of NGS data for demographic inference possible for non-model organisms (Metzker 2010).

In field crickets, closely related species often occur sympatrically and are morphologically and ecologically cryptic, but show strong differentiation in the male acoustic mating signals and corresponding female preferences (Alexander 1962, Otte 1992). This makes them ideal subjects for the study of speciation with historic and/or secondary gene flow (Maroja *et al.* 2009, Larson *et al.* 2013). Genomic resources for crickets are still scarce (Danley *et al.* 2007) but rapidly expanding through *de novo* transcriptome assembly (Zeng *et al.* 2013, Bailey *et al.* 2013). While *de novo* assembly of a genome is still relatively laborious for many non-model organisms, RNA-seq and the use of transcriptomic data for downstream analyses offer profound benefits, because of the reduced data load and focus on functional regions (Wang *et al.* 2009). We recently completed the transcriptome of the south-eastern field cricket *Gryllus rubens* (Berdan, E.L., Blankers, T., Waurick, I., Mazzoni, C., Mayer, F., Mol. Ecol. Res., in review). Now, we used RNA-seq data to (1) analyse genome-wide divergence, (2) infer the demographic history to provide insight into the influence of gene flow and population size variation during speciation, and (3) search for candidate genes contributing to barriers to RI between *G. rubens* and its closest relative, the Texas field cricket *G. texensis*.

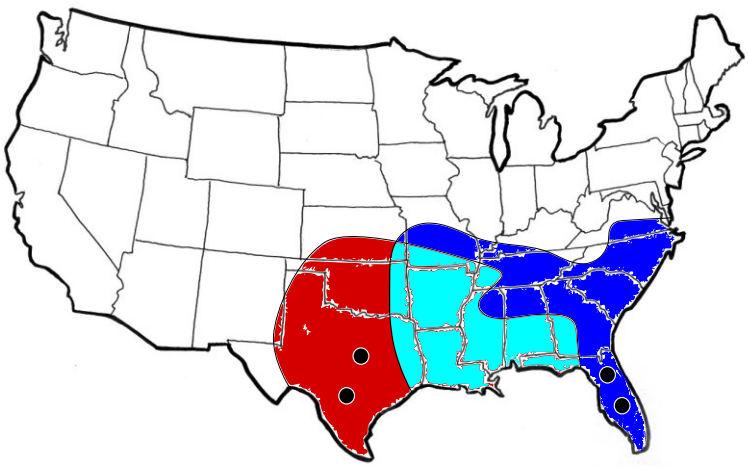

Figure 1 Geographic distributions for *G. texensis* (red) and *G. rubens* (blue). The sympatric zone is marked with turquoise. The distributions are based on Walker 1974 and on Walker 2015 (SINA, http://entnemdept.ufl.edu/walker/buzz/index.htm). The black dots in Texas and Florida represent the sampling locations for *G. texensis* and *G. rubens*, respectively.

Gryllus rubens and *G. texensis* are widely distributed across the southern Gulf and mid-Atlantic states in North America, from western Texas to Maryland, with a broad sympatric region from eastern Texas through western Florida. Males are morphologically cryptic (Gray *et al.* 2008, Walker 2015, Fig. 1) and there is no documented ecological divergence. However, females differ in the length of the ovipositor (Gray *et al.* 2001), which tentatively reflects ecological adaptation to different soil types (Bradford *et al.* 1993). There is strong premating isolation between the species through species-specific long-distance mate attraction songs (Walker 1998, Gray & Cade 2000, Blankers *et al.* 2015a) and close-range courtship songs (Gray 2005, Izzo & Gray 2011). There is potentially also premating isolation due to differences in cuticular hydrocarbons (CHCs, Gray 2005), which are known to be used in chemical mate signalling in field crickets (Tregenza & Wedell 1997, Thomas & Simmons 2009, 2010, Maroja *et al.* 2014). Naturally occurring interspecific hybrids have not been documented and all of the offspring of field-caught females appear to be pure species (Walker 1998, Izzo & Gray 2004, pers. observations). In the laboratory, hybrid offspring are easily obtained from single-pair interspecific matings (Gray 2011, Blankers *et al.* 2015b), but during interspecific crossing experiments in a recent study (Blankers *et al.* 2015b), we observed significant break-down in second generation hybrid offspring. This indicates that also some post-mating RI is present. A prior molecular study of *G. texensis* and *G. rubens* suggested a peripatric mode of speciation with divergence occurring between 0.25 and 2 million years ago (Gray *et al.* 2008), but that study was based only on a single mitochondrial locus.

In this study, we sequenced the transcriptomes of 40 *G. rubens* and *G. texensis* individuals across multiple allopatric populations. We quantified transcriptomic differentiation and used ABC to infer the influence of gene flow and population size variation in their evolutionary history. We then examined contigs (transcripts) showing elevated divergence or high rates of nonsynonymous substitutions. We expected a strong role for gene flow, since these species have overlapping distributions and are differentiated phenotypically predominantly in sexual communication. Therefore, we also expected that genetic differentiation was enriched for genes that play a role in chemical and acoustic mating behaviour. Our results provide insight into the speciation processes during divergence of two currently sympatric cricket species with unprecedented detail and show that sexual selection can be a powerful process in genetic divergence even in the face of gene flow.

METHODS

Sampling

Gryllus texensis were collected in Lancaster (TX; 60 females) in early September and in Austin (TX; 20 females) in early October 2013, and *G. rubens* in Lake City (FL; 18 females), and Ocala (FL; 20 females) in mid-September 2013 (Fig. 1). While collecting animals in the field, crickets were housed in containers in groups of a maximum of 15 individuals. Each container had gravel substrate and was provided with shelter, water, and food. Each container also contained a cup with vermiculite for oviposition. During two weeks, eggs were collected and transferred to a new container. Hatchlings were then reared to adulthood in groups of maximally 40 individuals in 19L plastic containers with gravel, shelter, and water and food *ad libitum*. We used laboratory-raised offspring of the field-caught females between 1 and 3 weeks after their final moult rather than field-caught specimens in order to standardize rearing conditions across all samples. Prior to preserving samples for RNA isolation, all animals (males and females) were played back an artificial stimulus resembling the conspecific male song for 10 minutes. The rationale here was that one of our primary interests was to look at genetic divergence in relation to mating behaviour polymorphism. If specific genes involved in female preference behaviour were only expressed upon hearing a male song signal, this could potentially be overcome by a brief play back 30 – 120 minutes prior to sample preservation. Stimulus play back occurred for females and males to also standardize our RNA sampling method across sexes. Within two hours of stimulus presentation, we sacrificed the cricket, removed the gut and then preserved the body in RNAlater following the manufacturer's instructions; samples were then stored at -80^{o}C until RNA isolation. A total of 5 males and 5 females were used from each of the two populations for each species (40 individuals in total).

RNA isolation

Whole animal samples (minus two legs removed for a DNA sample) were individually homogenized in TriFast (PEQLAB, Erlangen, Germany) using the Precellys ceramic kit 1.4/2.8 mm (PEQLAB, Erlangen, Germany) with a Minilys homogenizer. Total RNA was extracted from the homogenized samples following the manufacturer's instructions except that samples were precipitated with isopropanol diluted 1:2 with nuclease free water. A NanoDrop spectrophotometer (NanoDrop Products, Wilmington, Delaware) and a 2100 BioAnalyzer (Agilent Technologies, Santa Clara, California) were used to check for purity and quality. Samples with a 260/280 ratio appreciably lower

than 2.0, a 230/280 ratio lower than their 260/280 ratio, or severe degradation visible on an Agilent RNA 6000 Pico Assay electropherogram were not further processed and replaced with a new sample. To decrease the concentration of ribosomal RNA, an mRNA enrichment was performed using the Dynabeads mRNA Purification Kit (Life Technologies, Carlsbad, California) following manufacturer's instructions.

Illumina library preparation

The NEXTflex Directional RNA-seq Kit (dUTP based; Bioo Scientific, Austin, Texas) was used to make directional, strand specific RNA libraries. Library quality was determined using an Agilent High Sensitivity DNA Chip on a 2100 BioAnalyzer and library concentration was determined using a Qubit 2.0 fluorometer (Life Technologies, Carlsbad, California). Only libraries with a distinct band / product peak at approximately 350bp and a concentration >10nM were sequenced. Libraries were sequenced on a HiSeq 2000 (Illumina, San Diego, California) at the Max-Delbrück-Centrum (Berlin, Germany) to generate 100-bp paired-end reads. Libraries were individually barcoded using NEXTflex RNA-seq Barcodes from Bioo Scientific (Austin, Texas) and the 40 libraries were sequenced at a depth of 8 libraries per lane.

SNP calling

Raw reads were processed using Flexbar (Dodt *et al.* 2012) to remove sequencing primers, adaptors, and low quality bases on the 3' end. Samples were mapped to a reference transcriptome (Berdan, E.L., Blankers, T., Waurick, I., Mazzoni, C., Mayer, F., Mol. Ecol. Res., in review) using Bowtie2 (Langmead & Salzberg 2012) with default parameters but specifying read groups to mark reads as belonging to a specific individual. Duplicate reads were marked using 'picard' (http://picard.sourceforge.net) and then files from all individuals were merged prior to SNP calling. The Genome Analysis Toolkit (GATK; DePristo *et al.* 2011; Van der Auwera *et al.* 2013) was used for SNP calling. SNPs were called using the GATK-module 'UnifiedGenotyper' (Van der Auwera *et al.* 2013). Several levels of filtering were used to only retain high quality SNPs: (i) genotypes with a Phred-scaled confidence score less than 30 were removed; (ii) we filtered out variants if more than 10% of the total reads had a mapping quality of zero, if the quality score normalized by the amount of coverage was less than 2, if the root mean square of mapping quality across all samples was less than 40, if the Phred-scaled *P*-value for the Fisher exact test for strand bias was greater than 60, if the u-based z-approximation of the Mann-Whitney mapping quality rank sum test (testing different mapping qualities between reads with the reference allele versus the alternate allele) was less than -12.5, or if the u-based z-approximation of the Mann-Whitney read position rank sum test (testing if the reference and alternate allele are consistently present in different positions of the read) was less than -8; (iii) variants that did not have at least 10x coverage per individual for 90% of the individuals were also removed; (iv) we used VCFtools (Danecek *et al.* 2011) to filter out all variants with a minor allele frequency (MAF) less than 0.025 (at least two copies of the minor allele must be present), all variants with more than 2 alleles, and all variants with genotypes missing for more than 15% of individuals. After quality control, the final set of 175,244 SNPs across 8835 contigs was output as a VCF file.

Transcriptomic divergence

We used VCFtools to calculate Tajima's D and the allele frequency difference (*D*), quantified for each SNP as the absolute difference between the major allele frequencies of the two species, to explore genomic variation between *G. rubens* and *G. texensis*. We chose to analyse the distribution of *D*, rather than the F_{ST} distribution, for two reasons. First, F_{ST} depends on variation between as well as within populations and is thus a relative measure of among population divergence. Compared to *D*, F_{ST} is more sensitive to demographic changes (Holsinger & Weir 2009) and genomic regions with high F_{ST} are often the result of reduced diversity (within population) rather than reduced gene flow (Strasburg *et al.* 2012, Cruickshank & Hahn 2014). Second, we wanted to compare our results with a recent study of genomic variation between two other closely related, partially sympatric cricket species *G. pennsylvanicus* and *G. firmus* (Andrés *et al.*, 2013).

We then examined in more detail contigs that contained at least one fixed SNP (4828 SNPs across 1693 contigs), the 5% contigs with more than 3 SNPs that showed the highest number of fixed differences between species (433 contigs), and the contigs with more than 3 SNPs that showed high divergence ($D > 0.95$; 10 contigs). To place the respectively 4828 and 433 outlier contigs in a meaningful phenotypic context we searched for enriched functional categories using Gene Ontology (GO) enrichment implemented in Blast2go (Conesa *et al.* 2005). We further assessed gene function by inspecting

experimentally proven functions of the *D. melanogaster* genes that matched our outlier contigs using FlyBase (flybase.org).

Haplotype reconstruction

It is generally not recommended to use solely polymorphism data for coalescent simulations as the allele frequency spectrum can be strongly biased, especially in the case of historic population size variation (*i.e.* bottlenecks, Morin *et al.* 2004). We therefore chose to obtain haplotypes for 50 loci representative of the full data set and used these for the demographic analysis. To exclude contigs with exceptionally high SNP density, which can be indicative of sequencing errors, we filtered the 8835 contigs with polymorphic nucleotides by removing the top 5% of contigs with the highest number of SNPs per bp (measured between the first and the last SNP), thereby excluding 260 contigs. Also, we filtered for length (contigs > 3000 bp were excluded, totalling 3628 contigs) to exclude contigs potentially spanning across many exons. Finally, a subset of 50 contigs was randomly drawn from the remaining 4947 contigs. We confirmed that the random subset was representative of the original data set (no appreciable difference in F_{ST}, Tajima's D, and nucleotide diversity π). We created separate VCF files for each of our 50 contigs and then used fastPhase (Scheet & Stephens 2006) to phase them. Afterwards we merged these files and then separated them by individual. Finally, we used VCFlibs (part of the FreeBayes package; https://github.com/ekg/vcflib), a custom Python script (written by Marie Jeschek), and a custom R script (written by Thomas Blankers) to obtain haplotype sequences for all 50 contigs for all 40 individuals. The pipeline to convert the VCF and reference FASTA files into haplotypes for transcriptomic sequence data is available in the Supporting Information.

Demographic inference

To look for population substructure we first inspected allele frequency variation within and between species and population within species using principal component analysis on all 175,244 SNPs. We then used a Bayesian approach to infer population substructure as implemented in the program STRUCTURE (Falush *et al.* 2003) for a random subset of our SNP loci (*G. rubens*: 4, 976 SNPs, *G. texensis*: 4,974 SNPs). We used the admixture model with sampling location as prior information. We ran STRUCTURE with an MCMC chain length of 100,000 and with 10,000 burn-in length for K=1 through K=5 with three repetitions for each K-value. Results were analysed using STRUCTURE HARVESTER (Earl 2012).

We then determined whether the 50 haplotype loci had evolved in an approximately neutral way using BEAST 1.8.0 (Drummond *et al.* 2012). Selective constraints can change the gene genealogy thus biasing coalescent reconstruction of the demographic history (Ho & Shapiro 2011). It is therefore important to assure an approximate neutral evolution when using transcriptomic data. We first chose a relaxed molecular clock under a constant prior for both species (chain length: $5 \cdot 10^7$, burn-in: $5 \cdot 10^6$, thinning interval: $1 \cdot 10^4$) and found that we could not reject strict clock-like behaviour based on the 'ucld.stdev' parameter (*G. texensis* ucld.stdev = 0.435, *G. rubens* ucld.stdev = 0.302; values closer to 0 than to 1 indicate that loci do not deviate strongly from a constant, strict molecular clock; Drummond *et al.* 2007). We then used a strict clock in another series of simulations to evaluate the effective population size variation using the Extended Bayesian Skyline Plot (Heled & Drummond 2008). Two runs of $5 \cdot 10^8$ generations, 10% burn-in and logged every $1 \cdot 10^4$ were combined using LogCombiner. Because *Gryllus* has no precise estimation of mutation rates, a uniform mutation rate of 1 – 3% per site per million years was used based on average substitution rates in insect nuclear loci (Lin & Danforth 2004). Each locus was considered as a different partition, and we used the GTR+Γ substitution model for all partitions. Convergence (ESS>200) was checked in Tracer v1.6 (Rambaut *et al.* 2014).

To investigate the demographic history of *G. rubens* and *G. texensis* we used an ABC framework (Beaumont *et al.* 2002). We used ABCtoolbox (Wegmann *et al.* 2010), which has convenient wrapper functions linking several programs and calling them sequentially from the command line. We set up ABCsampler to simulate our data under several demographic scenarios in fastsimcoal v2.5.2.3 (Excoffier & Foll 2011, Excoffier *et al.* 2013) and then calculate summary statistics using arlsumstat v.3.5.1.3 (Arlequin v 3.5, Excoffier & Lischer 2010). We initially calculated a rather large subset of the summary statistics supported by arlsumstat (28 in total: mean, sd, and species-specific number of haplotypes, heterozygosity, the number of polymorphic sites in each species, the number of private polymorphic sites, Tajima's D, Fu's F_S, and nucleotide diversity (π), as well as pairwise π and F_{ST}). Then, using partial least squares regression (PLS), we retained the summary statistics with the highest predictive power (*i.e.* those with high factor loadings on the PLS components that lead to an increase in the predictive power of parameter estimates) for demographic estimates. Our final selection of summary statistics for model selection and parameter estimation included the between-species mean and standard

deviation of the number of polymorphic sites (mean_S & sd_S), the number of private polymorphic sites (PrS_1 & PrS_2), Tajima's D (D_1 & D_2), and nucleotide diversity π for each species (Pi_1 & Pi_2), as well as pairwise (between species) F_{ST} (FST_2_1) and π (Pi_2_1).

The demographic scenarios we compared are given in Fig. 2. We intentionally considered only relatively simple models with few parameters to avoid the risk of overparameterisation (Box 1979, Csilléry *et al.* 2010). We first ran 200,000 iterations of a simple divergence model [DIV, Fig. 2a], three gene-flow scenarios [Fig. 2b, continuous gene flow (CGF), ancient gene flow (AGF), and recent gene flow (*i.e.*, secondary contact, RGF)], and three bottleneck models [Fig. 2c, bottleneck for *G. rubens* (RB), for *G. texensis* (TB), and for both species (BB)]. Prior ranges for population sizes and time points were chosen on a log-uniform scale spanning across several orders of magnitude and for bottleneck size and migration rates on a uniform scale not overlapping zero.

After simulating the scenarios, model selection and posterior predictive checks were performed in R (R Core Team 2015). For model selection, we used a nested procedure. The three gene flow models (Fig. 2b) as well as the three bottleneck models (Fig. 2c) are nested, so the first model selection step was to determine the best gene flow and best bottleneck model, respectively. We reduced the posterior samples used in model selection to the 1% samples with the smallest Euclidean distance between the summary statistics of the simulated data and the observed data ('1% nearest posterior samples' from hereon) for each scenario separately. We then rotated summary statistics of the 1% nearest posterior samples across models to obtain a set of linear discriminants ('MASS' package, Venables & Ripley 2002) that maximized the distance among models within the nested categories. With these linear combinations of summary statistics, the 'postpr' function (multinomial logistics regression, 'abc' package, Csilléry *et al.* 2012) was used to calculate the posterior model probabilities. The two models (one gene flow and one bottleneck model) with the highest posterior probability ('best model' from hereon) were retained. Model selection was validated by visually inspecting the degree of overlap of the 1% nearest posterior samples of the models in multivariate summary statistic space. We also performed a leave-one-out cross validation with logistic regression ('cv4postpr' function, 'abc' package). After we had retained a single best gene flow model and a single best bottleneck model, we repeated model selection to select among a simple divergence scenario (DIV), the best gene flow (AGF) and bottleneck (RB) scenarios, and a scenario combining the best gene flow and the best bottleneck scenario (AGFRB, Fig. 2d).

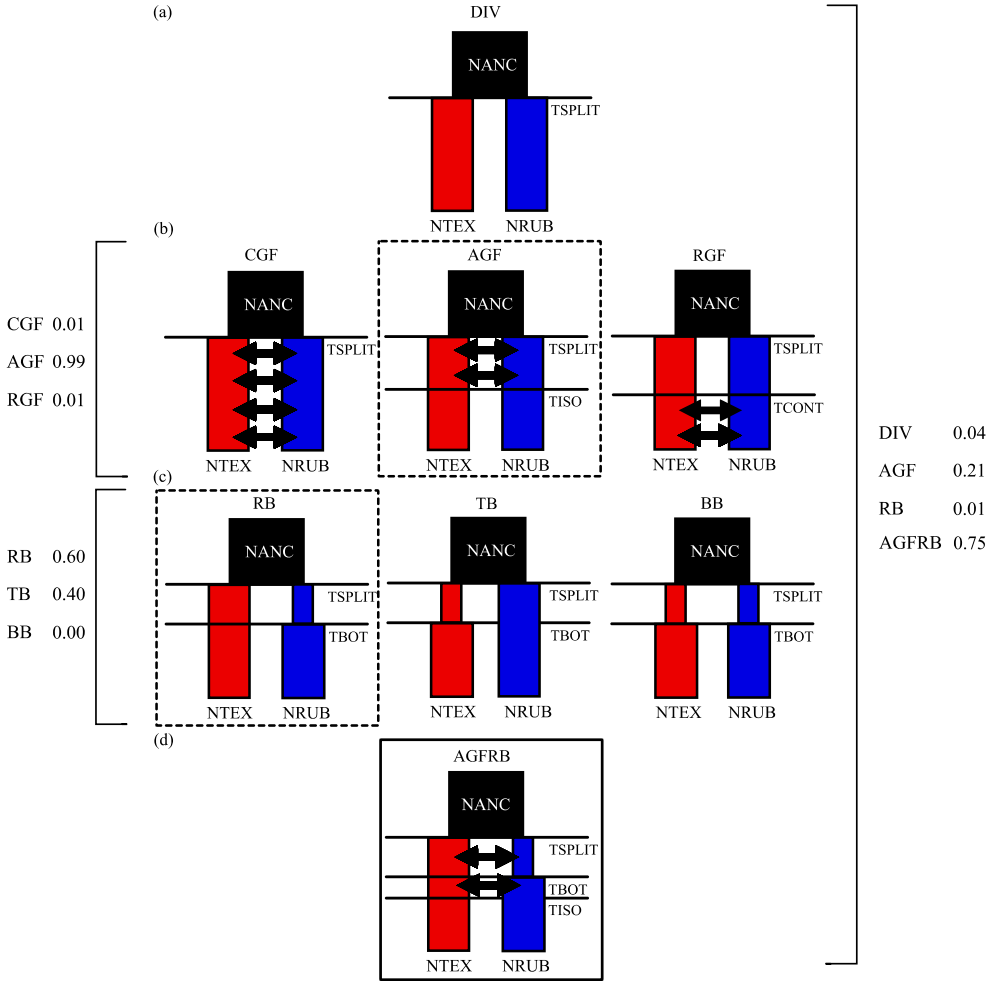

Figure 2 Demographic scenarios for Approximate Bayesian Computation. Eight scenarios were simulated under the ABC framework. (a) A simple divergence scenario (DIV) with a log uniform prior on the divergence time (T_{SPLIT}), the ancestral population size (N_{ANC}) and the current effective population sizes for *G. rubens* and *G. texensis* (N_{RUB}, N_{TEX}). (b) Three different gene flow models with either continuous gene flow (CGF), ancient gene flow (AGF), or recent gene flow (RGF) were described by the additional parameters describing migration rates ($M_{TEX>>RUB}$, $M_{RUB>>TEX}$; uniform priors not overlapping zero) and the time point of cessation of gene flow (T_{ISO}) or secondary contact (T_{CONT}), both with log uniform priors. (c) Three bottleneck models described by the time point of recovery to current population sizes (T_{BOT}; log uniform prior) and the relative population size reduction (BOTSIZE; uniform prior not overlapping zero) for *G. rubens* (RB), *G. texensis* (TB), or both (BB). (d) An additional model (AGFRB) combining the best gene flow (AGF) and best bottleneck (RB) model, marked by the black rectangles. The posterior probabilities for model selection are given left of the square (opening) brackets for the three gene flow and the three bottleneck models, and right of the square (closing) bracket for the final model selection step (see text for details).

We then ran 1,000,000 new simulations under the model with the highest posterior probability to estimate demographic parameters. Posterior predictive checks were performed by calculating the predicted R^2 and root mean squared error (RMSEP) using the 'pls' package (Mevik & Wehrens 2007). We also used the 'cv4abc' function from the 'abc' package to evaluate prediction error. We estimated the demographic parameters with the 'abc' function using non-linear regression and a tolerance rate of 0.05.

Signatures of selection

We scanned our sequences for loci that are potentially under positive selection using the ratio of non-synonymous to synonymous SNPs (dN/dS). We pooled our reads by species *in silico* and then called SNPs using GATK pipeline described above (with ≥ 200x coverage). Then, two "species-specific" transcriptomes were created using the FastaAlternateReferenceMaker from GATK (GATK; DePristo *et al.* 2011; Van der Auwera *et al.* 2013). We then used 'transdecoder' (part of the TRINITY package, Haas *et al.* 2013) to estimate all possible Open Reading Frames (ORFs) and estimated dN/dS following two different likelihood-based methods (NG: Nei & Gojobori 1986; YN: Yang & Nielsen 2000) using KaKs Calculator (Zhang *et al.* 2006). Contigs with only non-synonymous SNPs were left out because this

potentially results from an error in the estimation procedure. We obtained 7052 reading frames across 6419 contigs. All contigs with $dN/dS > 1$ using both the NG and the YN methods were considered contigs under selection. We searched for enriched functional categories across these contigs using Gene Ontology (GO) enrichment implemented in Blast2go (Conesa *et al.* 2005) and inspected experimentally proven gene functions in *D. melanogaster* on FlyBase.

RESULTS

Transcriptomic divergence

At a MAF cut-off of 0.025 we found a total of 175,244 SNPs. The average transition-transversion ratio was 1.6:1. Median *D* was 0.07 (first quantile: 0.05, third quantile 0.20) and 2.7% of the SNPs (4828) was fixed between the species (Fig. 3a). At a higher MAF cut-off (0.1; totalling 49.5 thousand SNPs), median *D* was 0.40 and almost ten percent of the SNPs was fixed between the species (Fig. S1a). Highly diverged contigs ($D > 0.85$) comprise about 1% and 11% of the contigs for low and high MAF cut-off, respectively (Fig. S1b,c). Average Tajima's D was negative for both species, but the distribution across loci showed substantial variation with both positive and negative values (Fig. 3b,c).

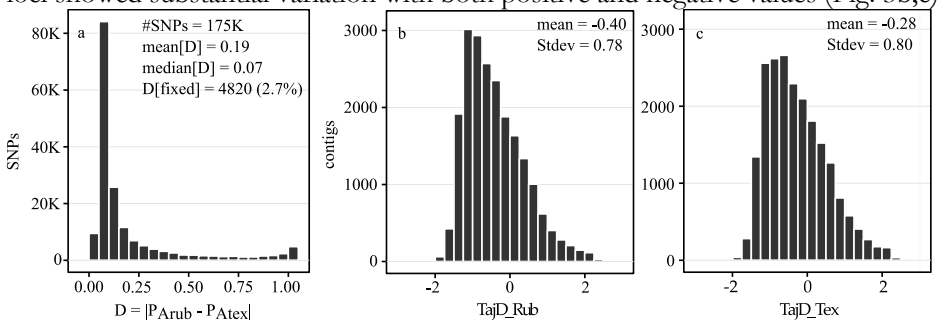

Figure 3 Genome-wide divergence. The distribution of the interspecific allele frequency difference (*D*) across SNPs (a), and of Tajima's D across contigs for *G. rubens* (b) and *G. texensis* (c) are shown.

To search for genes with the highest levels of interspecific divergence we first took all the contigs with at least one fixed SNP (1694 contigs) and performed a GO enrichment analysis in Blast2Go. We found 42 enriched GO terms (Table S1). In addition to terms related to cellular transport, transport across membranes, post-translational modification, RNA binding, binding of precursors for catalytic and signalling pathways, and regulation of gene expression, we found significant enrichment for 'actin filament binding'. This term included genes involved in muscle development and contraction (*sacromere length short*, 3 fixed SNPs; *coro*, 5 fixed SNPs) that control locomotor behaviour and activity of the direct and indirect flight muscle in *D. melanogaster* (Bai *et al.* 2007, Schnorrer *et al.* 2010). Additionally, two GO terms that are important in synthesis of cuticular hydrocarbons (CHCs; 'lipid particle' and 'cuticle hydrocarbon biosynthetic process') were enriched and included *pale* (three fixed SNPs) and *fatty acid desaturase 1* (*desat1*; we found six contigs matching *Desat1*, five of which had multiple fixed SNPs). *pale* and *desat1* are both involved in CHC biosynthesis and CHC-based mate choice behaviour in *D. melanogaster* (Marciallac *et al.* 2005, Liu *et al.* 2009). Lastly, the GO term 'adult behaviour' was also enriched and included *period* (six fixed SNPs), which is well known for its role in the love song of *D. melanogaster* (Lagisz *et al.* 2012) as well as its effects on the mating rhythms of males and females (Sakai & Ishida 2001).

We then looked into enriched functional gene categories for the 5% contigs with the highest number of fixed differences. This highlighted many contigs matching genes involved in wing morphogenesis and neurogenesis in *D. melanogaster*. Two of the 14 significant GO terms ('imaginal disc-derived wing vein morphogenesis' and 'synaptic growth at neuromuscular junction'; Table 1) contained contigs matching genes involved in wing vein morphogenesis, such as *split ends* and *osa* (Mace & Tugores 2004, Terriente-Felix & de Celis 2009), (wing) muscle development such as *heartless*, Butler 2003), and neuromuscular junction development (*akt 1* [Natarajan *et al.* 2013], *atlastin* [Lee *et al.* 2009], *Abl* [Lin *et al.* 2009], *shot* [Valakh *et al.* 2013], and *Scr64B* [Tsai *et al.* 2008]). The latter two genes are also involved in mushroom body development (Reuter *et al.* 2003, Nicolai *et al.* 2003). We also found relatively high divergence in *moesin* (11 fixed SNPs, $D = 0.30$) and *calmodulin* (8 fixed SNPs, $D = 0.38$). In addition to its role in endocytosis, *moesin* plays a role in *D. melanogaster* song production (Moran & Kyriacou 2009) and *calmodulin* is calcium-binding protein that mediates many important processes such as muscle contraction and memory, and is also involved in the perception of sound and smell (Senthilan *et al.* 2012, Mukunda *et al.* 2014).

Table 1 The GO enrichment terms for the 5% contigs with the highest number of fixed SNPs between *G. texensis* and *G. rubens*. The GO term ID number, the GO term, the category (cellular component, CC; molecular function, MF; biological process, BP), and the FDR corrected P-value are shown.

GO-ID	Term	Category	P(FDR)
GO:0019897	extrinsic to plasma membrane	CC	0.0103
GO:0000166	nucleotide binding	MF	0.0105
GO:0043168	anion binding	MF	0.0403
GO:0031177	phosphopantetheine binding	MF	0.0469
GO:0048149	behavioural response to ethanol	BP	0.0052
GO:0006897	endocytosis	BP	0.0103
GO:0035160	maintenance of epithelial integrity, open tracheal system	BP	0.0116
GO:0051124	synaptic growth at neuromuscular junction	BP	0.0183
GO:0000165	MAPK cascade	BP	0.0186
GO:0046673	negative regulation of compound eye retinal cell programmed cell death	BP	0.0276
GO:0051641	cellular localization	BP	0.0363
GO:0033036	macromolecule localization	BP	0.0389
GO:0042058	regulation of epidermal growth factor receptor signalling pathway	BP	0.0450
GO:0008586	imaginal disc-derived wing vein morphogenesis	BP	0.0450

Finally, we inspected the annotation of the contigs with $D > 0.95$ (11 contigs). High divergence was observed in contigs matching genes involved in mushroom body development (*RN-tre*), motor neuron activity (*convoluted*, Kurusu *et al.* 2008), and muscle attachment (*Ced-12*, Liu *et al.* 2013; Table 2).

Table 2 Contigs with high genetic divergence ($D \geq 0.95$). For each of the 11 contigs with the largest allele frequency differences the number of SNPs (number of fixed SNPs in parentheses), the annotation, and the experimentally proven function(s) in *D. melanogaster* (Flybase.org) are shown.

Contig	D	#SNPs(fixed)	Annotation	Function
c218723_g5_i3	0.96	3(0)	CG31743	carbohydrate biosynthetic process
c204595_g2_i2	0.97	3(2)	Ubiquitin conjugating enzyme E2H	lateral inhibition
c114076_g1_i1	1.00	3(3)	Ced-12	CNS development / muscle attachment
c142606_g1_i1	1.00	3(3)	cg7477	-
c154500_g1_i2	1.00	3(3)	convoluted	motor neuron axon guidance / open tracheal system development
c200075_g6_i2	1.00	3(3)	rn-tre	GTP-ase activation / Mushroom Body development
c203808_g16_i4	1.00	3(3)	PP4R1	protein phosphatase
c208057_g2_i3	1.00	4(4)	cg6178	fatty-acyl-CoA synthase activity
c220375_g1_i5	1.00	4(4)	-	-
c220530_g3_i1	1.00	4(4)	cg6982	establishment/maintenance of cell polarity
c222173_g1_i1	1.00	3(3)	-	-

Demographic inference

We were first interested in the presence of population substructure within species, as this affects the interpretation of genetic divergence metrics comparing variation between species. We found no substantial evidence for population substructure within either species. Rotating the variation in allele frequencies in a principal component analysis revealed that the major axis of variation (23.9%) teased apart the species (Fig. 4). Subordinate axes separated populations within *G. texensis* (PC2, 6.13%) and *G. rubens* (PC4, 4.35%). Clustering of individuals within species along PC1 was considerably stronger than clustering of individuals within populations along the other PCs.

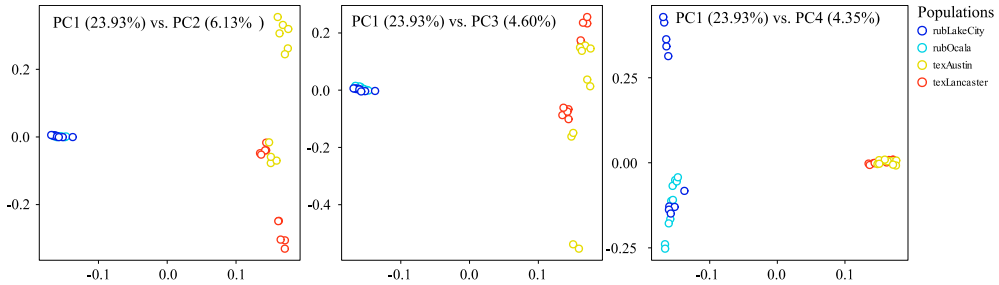

Figure 4 Population substructure in *G. rubens* and *G. texensis*. Variation in allele frequencies between species and between populations within species (Lake City and Ocala for *G. rubens*; Lancaster and Austin for *G. texensis*) is shown. The allele frequency variation in all 175,244 SNPs is summarized in the first four principal components teasing apart the species (PC1), and the populations in *G. texensis* (PC 2) and *G. rubens* (PC 4). Note that clustering along the PCs explaining within species between populations variation is much weaker compared to clustering of the species along PC1 and variation in PC2-4 represents much less variation compared to variation in PC1.

The Bayesian test for population structure implemented in STRUCTURE further supported our expectations that neither of the species was strongly differentiated geographically. The optimal K equalled 2 when we ran STRUCTURE with both species included (Fig. S2). Examining population structure within species, we found for *G. texensis* that K = 2 had the highest probability using the Delta K method (Evanno *et al.* 2005), but the probability of K = 2 differed only marginally from other values for K (Fig. S2); for *G. rubens* the optimal K was 4 but again only very small differences in probability with other values for K were observed (Fig. S2). Given the marginal differences among values for K within species and the strong support for K = 2 across species, we conservatively accept K = 1 for each species.

To gain insight in the role of demography during divergence we tested for population size variation and strength and timing of gene flow. Extended Bayesian skyline plots (Fig. S3) showed population growth over the last 15,000 years for both species. There were small differences between the species in population size (larger in *G. texensis*) and in the onset of population growth (*G. texensis*: ca. 15,000 years ago, *G. rubens*: ca. 12,000 years ago), but these estimates were associated with large standard deviations.

To examine the demographic history of *G. rubens* and *G. texensis* in more detail and compare predicted patterns of genetic variation under hypothetical demographic scenarios against observed genetic variation, we followed the ABC framework. We used a nested rejection procedure to select the best model out of eight different models varying in the presence and timing of bottlenecks and gene flow (Fig. 2). The best gene flow model was a model with ancient gene flow (AGF, $P_{posterior}$ = 0.98, CGF: $P_{posterior}$ = 0.01, RGF: $P_{posterior}$ =0.01). Model choice was strongly supported by cross validation. In this cross validation step, 1000 independent model selection steps are performed for each of the demographic scenarios. In each step it is assumed that a given model is the 'true' model. The number of times (out of 1000) that the 'true' model was chosen thus gives a measure of the accuracy with which one hypothetical scenario can be chosen in favour of other scenarios. Note that although this measure of accuracy is expected to be higher for models that fit the data better, this method provides no support for the absolute or relative fit of the model to the observed data. Rather, it indicates whether models are sufficiently different from each other (*i.e.*, simulations under each scenario produce significantly different patterns of genetic variation). The 'true' model was chosen 99%, 72%, and 50% of time for AGF, CGF, and RGF, respectively. The best bottleneck model included a bottleneck for *G. rubens* but not for *G. texensis* (RB: $P_{posterior}$ = 0.60, TB: $P_{posterior}$ = 0.40, BB: $P_{posterior}$ = 0.00). Again, model choice was strongly supported by the cross validation (75%, 76%, and 86% of the time for RB, TB, and BB respectively). Since the models with ancient gene flow and with a bottleneck for *G. rubens* were the best models for their respective nested categories, we defined an additional model including both ancient migration and a bottleneck for *G. rubens* (AGFRB). The final model selection step thus compared four models: a simple divergence model (DIV), and the AGF, RB, and AGFRB models. Cross validation indicated that we could distinguish between the models (66%, 63%, 74%, 53% for DIV, AGF, RB, AGFRB, respectively). However, model selection between AGF and AGFRB and between DIV and RB was associated with higher uncertainty, given that the distribution of posterior samples in multivariate summary statistic space showed strong overlap (Fig. 5). The model with the highest posterior probability was AGFRB ($P_{posterior}$ = 0.75) followed by AGF ($P_{posterior}$ = 0.21). Both were significantly stronger supported than DIV ($P_{posterior}$ = 0.04) and RB ($P_{posterior}$ = 0.01).

Because there was significant overlap between the posterior of AGF and AGFRB (Fig. 5), both models were used to estimate demographic parameters. Since the AGFRB model differs from the AGF model exclusively in the addition of parameters setting the relative population size (BOTSIZE) and timing (T_{BOT}) of the bottleneck we only discuss the results for the AGFRB model. However, the demographic parameter estimates for the AGF model were very similar (Fig. 6a-c versus Fig. 6d-f; Table 3 versus Table S2).

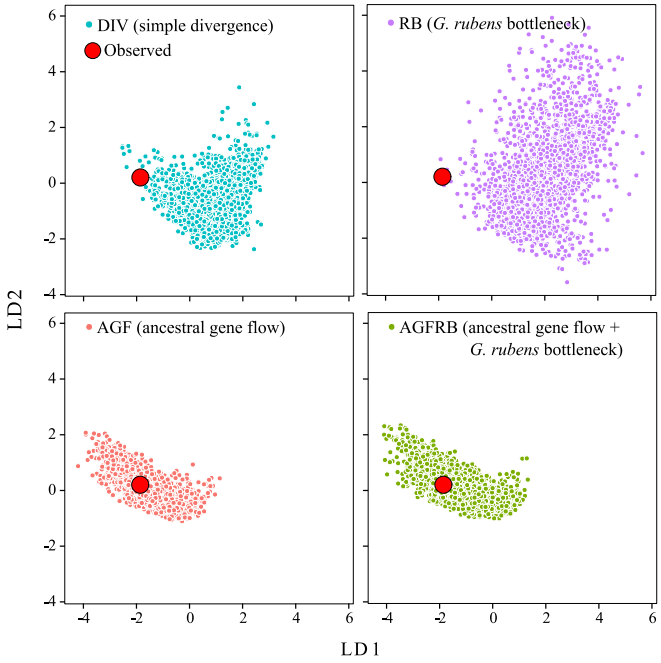

Figure 5 Distribution of observed and simulated data sets in multivariate summary statistic space. For each of the four models used in the final model selection step (see also Fig. 2) the distribution of the 1% posterior samples with the smallest Euclidean distance to the observed data is shown relative to the coordinates of the observed data. The multivariate summary statistic space is constrained by the first two linear discriminants (see text for details) representing linear combinations of the summary statistics used in model selection.

Table 3 Prior distributions, posterior predictive checks and posterior parameter estimates for the model with ancestral gene flow and a bottleneck for *G. rubens* (AGFRB). For each of the parameters the minimum and maximum of the uniform (u) or log-uniform (lu) prior distributions, the power (R^2), the root mean squared error prediction (RMSEP), and the 95% confidence interval and median are shown. Parameter estimates for the AGF model are in Table S2. Abbreviations for the parameters: Ancestral population size (N_{ANC}), effective population sizes for *G. rubens* (N_{RUB}) and *G. texensis* (N_{TEX}), time since divergence (T_{SPLIT}), time since cessation of gene flow (T_{ISO}), time since recovery from the bottleneck (T_{BOT}), relative population size during the bottleneck (BOTSIZE), rate of migration from *G. texensis* into *G. rubens* populations ($M_{TEX>>RUB}$), and vice versa ($M_{RUB>>TEX}$).

	Prior		Validation		Posterior		
Parameter	minimum	maximum	R^2	RMSEP	2.5%	median	97.5%
$LOG_{10}(N_{ANC})$	4.0	6.0 (lu)	0.05	0.974	4.94	5.32	5.72
$LOG_{10}(N_{RUB})$	3.0	6.0 (lu)	0.89	0.333	4.70	4.79	4.87
$LOG_{10}(N_{TEX})$	3.0	6.0 (lu)	0.88	0.346	4.73	4.85	4.94
$LOG_{10}(T_{SPLIT})$	5.0	7.0 (lu)	0.01	0.997	5.49	6.23	6.74
$LOG_{10}(T_{ISO})$	3.0	7.0 (lu)	0.81	0.438	4.27	4.53	4.72
$LOG_{10}(T_{BOT})$	5.0	7.0 (lu)	0.02	0.990	5.14	5.19	5.32
BOTSIZE	0.01	0.5 (u)	0.01	0.995	0.09	0.15	0.23
$M_{TEX>>RUB}$	0.01	0.5 (u)	0.12	0.938	0.05	0.12	0.18
$M_{RUB>>TEX}$	0.01	0.5 (u)	0.12	0.938	0.01	0.18	0.75

The observed summary statistics fell well within the range of the simulated summary statistics under the AGF and AGFRB models (Fig. 5). For some demographic parameters (current population sizes for *G. rubens* [N_{RUB}] and *G. texensis* [N_{TEX}], and time since cessation of gene flow [T_{ISO}) we found high predictive power ($R^2 > 0.81$) and relatively low root mean squared error predictions (RMSEP < 0.44; Table 3). For other parameters (migration rates [$M_{TEX>>RUB}$ and $M_{RUB>>TEX}$], divergence time [T_{SPLIT}], ancestral population size [N_{ANC}], and the relative size [BOTSIZE] and timing [T_{BOT}] during the bottleneck for *G. rubens*) the error was appreciably higher and the predictive power was much lower

(Table 3). For all parameters, except the estimate of the divergence time and $M_{RUB>>TEX}$, the 95% confidence intervals of the distributions were narrow. The ancestral effective population size was estimated between 87,096 and 524,807 and about an order of magnitude higher than the model estimates for current effective population sizes in *G. rubens* (50,118 – 74,131) or *G. texensis* (53,703 – 87,096; Table 3, Fig. 6a). The divergence time estimate was between 0.31 and 5.50 million generations ago (mga), with a median divergence time of 1.70 mga (Fig. 6b; Table 3). Since both species have two generations annually (Alexander, 1968), divergence was estimated to have taken place 0.85 million years ago (mya). A bottleneck for *G. rubens* was estimated at 9% to 23% of the current effective population size (Table 3, Fig. 6c) and recovery to current population sizes was achieved between 69,019 and 104,464 years ago (Table 3, Fig. 6b). Ancestral gene flow was bi-directional (0.12 and 0.18 for gene flow from *G. texensis* into *G. rubens* and vice versa, respectively; Table 3, Fig. 6c) and ceased between 9,310 and 26,240 years ago (Table 3, Fig. 6b).

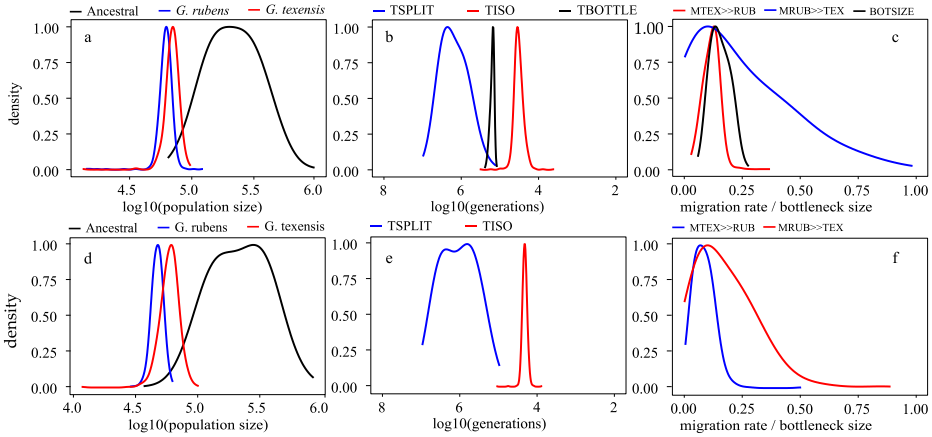

Figure 6 Demographic parameter estimation. The density distribution under the AGFRB (a-c) and the AGF (d-f) are shown for the ancestral and current population sizes (a,d), the time point for divergence, cessation of gene flow, and recovery to current population sizes after the bottleneck (b,e), and the migration rates and bottleneck size (c,f) are shown. The density lines have been trimmed to the existent parameter distribution (*i.e.*, no density extrapolation) and have been smoothed by adjusting the bandwidth. For lines within one panel the same bandwidth has been used.

Signature of selection

To look for genes that are potentially under positive selection we explored the contigs with high rates of nonsynonymous SNPs. We found no enriched GO categories among the 123 contigs with $dN/dS > 1$. Sixty-three had either no significant match to annotated genes in the Blast2Go data base or matched genes with no experimentally proven functions in *D. melanogaster* (Table S3). Of the genes matching the remaining 60 contigs (*i.e.* all with experimentally proven functions in *D. melanogaster*) 26 were involved in neurogenesis and neuro-system development, four in muscle development, three in wing morphogenesis, and two in the sensory perception of sound and smell, respectively (Table S3). Additionally, we found high dN/dS for a calcium and a potassium channel, for serotonergic and GABA-ergic neurotransmitter transport, and for *takeout* which interacts with *fruitless* in *D. melanogaster* courtship behaviour (Dauwalder *et al.* 2002). Lastly, the CHC biosynthesis gene *fas2* (which is also involved in locally adaptive desiccation resistance in the *D. birchii* – *D. serrata* speciation event; Chung *et al.* 2014), had among the highest ratios of nonsynonymous to synonymous SNPs ($dN/dS = 6.64$).

DISCUSSION

We investigated genomic divergence and demographic history in *Gryllus rubens* and *G. texensis*, two closely related sympatric species of field crickets that are isolated most strongly by divergent mating signals and mate preferences. Genomic divergence across ~175k SNPs showed a bimodal and slightly right-skewed distribution of allele frequency differences (*D*) (Fig. 3), with many highly diverged contigs matching genes involved in neuromuscular development, wing morphogenesis, and CHC biosynthesis (Table 1, Table 2). The observed transition:transversion ratio of 1.6:1 suggests that sequencing error did not contribute unduly to SNP discovery. Also, the observed ratio compares well with a recent estimate (1.55) from the *G. firmus/pennsylvanicus* species pair (Andrés *et al.* 2013). Bayesian analysis in STRUCTURE and principal component analysis of the allele frequencies suggested strong structure between species with weak structure within species. Our demographic analyses indicated that these species went through a long period of ancient gene flow before they were completely reproductively

isolated between about 9,000 and 26,000 years ago (Fig. 6, Table 3). Similar to the loci showing elevated levels of genetic divergence, many of the contigs with putative footprints of selection matched genes involved in neurogenesis, muscle development, information processing and neuronal function, and mating behaviour.

Transcriptomic divergence

We reported variable levels of *D* across the transcriptome of *G. rubens* and *G. texensis*. The vast majority of loci were shared polymorphisms (median $D = 0.08$; Fig. 3a), but a substantial proportion (2.7% for MAF cut-off = 0.025, 9.7% for MAF cut-off = 0.1) was fixed between species. Contrary to the distribution for Tajima's D, which was unimodal, the bimodal distribution of *D* with the higher MAF cut-off (Fig. S1a, where more recent, low frequency mutations are filtered out) suggests that the transcriptomes of these crickets are heterogeneously affected by introgression (Wu 2001) and ancestral (linked) and recent selection (Cruikshank & Hahn 2014).

Interestingly, the observed distribution of *D* is similar to the distribution of allele frequency differences between two congeneric species *G. firmus* and *G. pennsylvanicus* (Andrés *et al.* 2013). This despite the fact that *G. rubens* and *G. texensis* appear to have had a long period of ancient but not recent gene flow whereas the species in the Andrés *et al.* study diverged in allopatry (and so presumably without ancient gene flow) before secondary contact led to contemporary gene flow (Broughton & Harrison 2003, Andrés *et al.* 2013). Both species pairs show clear bimodality in the distribution of *D* with a large number of fixed differences. An initial period of allopatry for *G. pennsylvanicus* and *G. firmus* is a potential reason why despite a more recent divergence (ca. 200,000 years ago), the number of fixed differences is much higher (11.6% at MAF cut-off = 0.01 versus 2.7% at MAF cut-off = 0.025 in this study). Gene flow at the onset of divergence, as occurred in *G. rubens/texensis*, is expected to be stronger than gene flow following secondary contact, as occurred in *G. firmus/pennsylvanicus*. Alternatively, the difference in the proportion of fixed differences may be due to different sampling approaches (whole body tissue in this study versus only the accessory glands in Andrés *et al.* 2013).

Due to the long history of bidirectional hybridization, fixed SNPs potentially represent polymorphisms in genes contributing to RI (Nosil & Schluter 2011). As *G. rubens* and *G. texensis* have most conspicuously diverged in mating behaviour, we expected that genetic divergence would be high in contigs matching genes potentially involved in acoustic and chemical mating behaviour. We found direct evidence supporting this hypothesis, with high divergence in contigs matching *period*, *moesin, pale*, *desat1*, *fas2*, and *calmodulin*. Although part of the role of *period* in the courtship song rhythm of *D. melanogaster* has been revoked (Stern 2014), there is still substantial evidence for a correlation between genetic variation in *period* and variation in the fruit fly courtship song (*e.g.* Lagisz *et al.* 2012). Interestingly, *period* was also divergent between the Hawaiian swordtail crickets *Laupala cerasina* and *L.paranigra*, sister species for which a significant relationship between the pulse rate of the song and locomotion rhythm and between circadian singing and locomotion rhythm was shown (Fergus & Shaw 2013), indicating a pleiotropic basis of rhythmic behaviours.

Other evidence for genetic variation in genes involved in acoustic communication in crickets was more indirect. Examining all contigs with at least one fixed SNP, the 5% contigs with the highest number of fixed SNPs, or the contigs with $D < 0.95$, we found many indications for elevated divergence in genes controlling wing morphogenesis, neurogenesis, and nervous system and neuromuscular development. The spectral and temporal properties of the song produced by male crickets to attracted mates depend on (i) the stridulatory organ morphology (*i.e.* plectrum and the teeth on the file) and the resonant properties of the wing (*i.e.* harp) that determine in the carrier frequency (*i.e.* the pitch) of the song (Bennet-Clark 1999), (ii) neural networks called central pattern generators that control the stridulatory organs, and (iii) neuromuscular (synaptic) junctions and intrinsic properties of the muscles that affect the temporal rhythm of the song (reviewed in Gerhardt & Huber 2002). Similarly, song recognition and preference in females are controlled by a complex network of neurons (*e.g.* Hennig *et al.* 2014, Schöneich *et al.* 2015), that are located in mesothoracic ganglia and in the brain. Variation in song signals and preferences is thus expected to be manifested in changes in the properties of muscles, neuromuscular junctions, and channels mediating excitatory and inhibitory stimuli from within the nervous system (*i.e.* properties of calcium and potassium channels and synaptic neurotransmitter transporters). Despite the fact that acoustic singing and preference behaviour is hardwired early-on in development (Gerhard & Huber 2002), neurogenesis is continuous in crickets (Cayre *et al.* 1994), therefore making it possible to uncover these genes using RNA-seq.

Demographic inference

We used the ABC framework (Beaumont *et al.* 2002) to test for the effects of gene flow and population size variation and found that a scenario including a bottleneck for *G. rubens* and bi-directional ancient migration had the best support. Model selection was strongly supported by the cross-validation approaches, although the model with only historical gene flow (and no bottleneck) largely overlapped in multivariate summary statistic space with the model including historical gene flow and a bottleneck for *G. rubens*. Importantly, we note that these models gave very similar estimates for the parameters (Table 3 versus Table S2). We conclude that ancient gene flow occurred throughout most of the species' history before they became fully reproductively isolated between 9,000 and 26,000 years ago, *i.e.* during the last two Pleistocene glacial cycles, which may have imposed a period of allopatry.

A previous study based on mitochondrial DNA (mtDNA) suggested a peripatric origin for *G. rubens* from one of two ancestral haplotype clusters in *G. texensis* followed by both demographic and range expansion of *G. rubens* (Gray *et al.* 2008). Our finding here of a bottleneck in *G. rubens*, reducing the population size to about 10% to 25% of the current population size, supports this scenario. The STRUCTURE results in this study show little or no support for the idea that *G. texensis* consists of two admixed ancestral genotypic clusters (Gray *et al.* 2008), but note that the PCA revealed some clustering of the alleles in *G. texensis* which was unrelated to the sampling location. Despite these similar results, this study differs in important ways from the Gray *et al.* (2008) study other than just transcriptomic representation of nuclear DNA (nDNA) versus mtDNA. In this study, the genomic sampling is massively extensive, ca. 175k SNPs, however the species sampling is limited to 40 individuals (some of which may be siblings) from two allopatric populations per species. Conversely, the prior study was very limited in genomic resources, using only a single mitochondrial locus, *Cytochrome Oxidase I*, but the species sampling was extensive, consisting of 365 unrelated individuals from 48 localities throughout the species' ranges (Gray *et al.* 2008). Despite these differences, the concordance of the results is striking: both studies suggest parapatric (or peripatric) divergence on the order of ~1 mya (0.25 to 2 mya for mtDNA and 0.85 mya from RNAseq data) and both studies strongly suggest that *G. rubens* underwent a major bottleneck.

Our estimate for the ancient population size (median: 208,929) was almost an order of magnitude higher than the current estimated effective population sizes (*G. rubens*: 61,659; *G. texensis*: 70,795). The estimated effective population sizes are surprisingly low given regular observations of late summer/fall outbreaks of many 100,000s of *G. texensis* (see *e.g.* https://www.youtube.com/watch?v=MTsY4jtxLEQ); the expected census population size for *G. texensis* is in the millions (Gray *et al.* 2008). However, both the extended Bayesian skyline plots (Fig. S3) and negative values for Tajima's D support a scenario of recent population expansion (recent meaning end-Pleistocene, ca. 12,000 to 15,000 years ago). Recent population expansion typically results in effective population sizes far smaller than census sizes (*e.g.*, Nadachowska-Brzyska *et al.* 2013; Ptak & Przeworski 2002). Furthermore, wild field cricket populations show substantial inter-individual variation in mating success (Ritz & Köhler 2010, Rodriguez-Muñoz *et al.* 2010) considerably lowering effective population size relative to census estimates (Lande & Barrowclough 1987).

Signatures of selection

Similar to our findings for elevated levels of genetic divergence, many contigs with high rates of nonsynonymous SNPs ($dN/dS > 1$) matched genes involved in sensory perception and motor patterns. Examples include genes involved in neurogenesis, nervous system and muscle development, and sensory perception of sound and smell. Our analyses also uncovered a gene involved in courtship behaviour in *D. melanogaster* (*takeout*) and genes involved in serotonergic and GABA-ergic neurotransmitter transport (in invertebrates, these neurotransmitters drive the initiation and possibly modulation of rhythmic motor activity, including stridulatory behaviour in crickets [Gerhardt & Huber 2002]). These results thus provide further support for the hypothesis that divergent mating behaviour has strongly contributed to the establishment of RI.

Both our scan for highly divergent loci and for loci with high rates of nonsynonymous SNPs are sensitive for the discovery of false positives. However, the fact that the different analyses gave similar results (but not necessarily highlighting the same contigs) shows that our findings are robust. However, the applied sampling method has potentially somewhat biased our results. We sequenced samples from first generation laboratory offspring rather than animals directly from the field. The different populations have thus been raised in similar milieus and on the same food source. On the one hand this brings the practical advantage that conditions are standardized across species and populations. On the other hand, despite the fact that no differences between *G. texensis* and *G. rubens* in ecology,

microhabitat use, or feeding behaviour have been described, the laboratory conditions have potentially limited our potential to detect genetic differences related to local adaptation in the transcriptomic data.

Sexual selection and speciation

From a theoretical standpoint, speciation with gene flow by sexual selection alone is unlikely and can occur only in the presence of other disruptive selective forces (van Doorn *et al.* 2004, Weissing *et al.* 2011, Servedio 2015). However, explosive radiations of *Laupala* crickets and African cichlid fish, both characterized by sympatrically occurring but sexually divergent species (Seehausen 2000, Mendelson & Shaw 2005), provide compelling empirical counterexamples. Our data suggest that for *G. texensis* and *G. rubens* sexual selection on acoustic and chemical mating traits has played a major role in the establishment of RI and divergent took place, for the most part, in the face of gene flow. We also inferred relatively recent divergence (less than one million years ago).

One of the major challenges is thus to integrate these findings with the theoretical conundrums of speciation by sexual selection. First, a more detailed study of the way multiple selective forces act simultaneously is warranted, as is a greater interest in how (sexual) selection interacts with the ecology (Scordato *et al.* 2014). For example, long-distance and short-distance mating behaviour in crickets is reasonably well studied, but we have very little knowledge on the degree to which this behaviour depends on the environment, on male-male aggressive behaviour in crickets (Alexander 1961), and on the polyandrous mating strategies of female Gryllids (Bretman & Tregenza 2005). Second, unravelling the genetic architecture can reveal if traits and preferences can evolve rapidly and unconstrained, which would facilitate an early onset of reproductive barriers due to mating polymorphisms. For example, X-linkage, which has been widely documented for crickets (*e.g.* Hoy 1974, Shaw 1996) including for the *G. rubens* and *G. texensis* system (Blankers *et al.* 2015b), and pleiotropic genetic effects linking traits and preferences as is possibly the case in swordtail crickets (Shaw & Lesnick 2009, Shaw *et al.* 2011) can strongly enhance the evolution and divergence of sexual traits. Overall, our data suggest that sexual selection was a major factor in the initial phases of speciation during which bi-directional gene flow was inferred to be likely. Although this seemingly contradicts predictions from recent and earlier speciation models, we stress that many biotic and abiotic factors, including the genetic architecture of possible speciation traits, can strongly affect the reach of (sexual) selection but that these factors and the breath of their effects usually remain poorly understood.

SUMMARY

Our results are informative about the evolution of reproductive isolation in the presence of gene flow and give unprecedented detail about the divergence process in the field crickets *G. rubens* and *G. texensis*. We sequenced the transcriptomes of 40 individuals and showed limited interspecific divergence in the majority of polymorphic loci while also uncovering a substantial number of contigs harbouring SNPs with high interspecific divergence. Using haplotype sequences obtained through a custom pipeline we inferred that a long period of ancient gene flow and a potential bottleneck in the derived species (*G. rubens*) preceded completion of RI. Contigs with high sequence divergence or contigs showing footprints of positive selection highlighted genes with experimentally proven functions in neuromuscular development, courtship behaviour, and chemical mating behaviour. To our knowledge, this is the first study to use genome-wide variation to successfully infer genes contributing to divergence in acoustic and chemical mating behaviour in crickets. Our study advances our understanding of genomic divergence and the role of secondary sexual characters in the early phases of speciation in the presence of gene flow.

ACKNOWLEDGEMENTS

We thank Marie Jeschek and Harald Detering from the Berlin Center for Genomics and Biodiversity (BeGenDiv) for providing scripts that were used in haplotype reconstruction and Ka/Ks (dN/dS) estimation, respectively. Sample collection and processing comply with the "Principles of animal care", publication No. 86-23, revised 1985 of the National Institute of Health, and also with the current laws of Germany. The authors declare no conflict of interest. This study is part of the GENART project and funded by the Leibniz Association (SAW-2012-MfN-3). STV is currently supported by an Alexander von Humboldt Foundation fellowship.

REFERENCES

Alexander, R. 1968. Life cycle origins, speciation, and related phenomena in crickets. *Q. Rev. Biol.* **43**.

Alexander, R.D. 1961. Aggresiveness, Territoriality, and Sexual Behaviour in Field Crickets (Orthoptera: Gryllidae). *Behaviour* **17**: 130–223.

Alexander, R.D. 1962. Evolutionary change in cricket acoustical communication. *Evolution.* **16**: 443–467.

Andrés, J.A., Larson, E.L., Bogdanowicz, S.M. & Harrison, R.G. 2013. Patterns of transcriptome divergence in the male accessory gland of two closely related species of field crickets. *Genetics* **193**: 501–513.
Bai, J., Hartwig, J.H. & Perrimon, N. 2007. SALS, a WH2-domain-containing protein, promotes sarcomeric actin filament elongation from pointed ends during Drosophila muscle growth. *Dev. Cell* **13**: 828–42.
Bailey, N.W., Veltsos, P., Tan, Y.F., Millar, A.H., Ritchie, M.G. & Simmons, L.W. 2013. Tissue-specific transcriptomics in the field cricket *Teleogryllus oceanicus*. *G3 Genes|Genomes|Genetics* **3**: 225–230.
Beaumont, M.A., Zhang, W. & Balding, D.J. 2002. Approximate Bayesian Computation in Population Genetics. *Genetics* **162**: 2025–2035.
Bennet-Clark, H.C. 1999. Resonators in insect sound production: how insects produce loud pure-tone songs. *J. Exp. Biol.* **202**: 3347–3357.
Bertorelle, G., A, B. & S, M. 2010. ABC as a flexible framework to estimate demography over space and time : some cons , many pros. *Mol. Ecol.* **19**: 2609–2625.
Blankers, T., Hennig, R.M. & Gray, D.A. 2015a. Conservation of multivariate female preference functions and preference mechanisms in three species of trilling field crickets. *J. Evol. Biol.* **28**: 630–641.
Blankers, T., Lübke, A.K. & Hennig, R.M. 2015b. Phenotypic variation and covariation indicate high evolvability of acoustic communication in crickets. *J. Evol. Biol.*, doi: 10.1111/jeb.12686.
Box, G.E.P. 1979. Science and Statistics. *J. Am. Stat. Assoc.* **71**: 791–799.
Bradford, M.J., Guerette, P.A. & Roff, D.A. 1993. Testing hypotheses of adaptive variation in cricket ovipositor lengths. *Oecologia* **93**: 263–267.
Bretman, A. & Tregenza, T. 2005. Measuring polyandry in wild populations: a case study using promiscuous crickets. *Mol. Ecol.* **14**: 2169–79.
Broughton, R.E. & Harrison, R.G. 2003. Nuclear gene genealogies reveal historical, demographic and selective factors associated with speciation in field crickets. *Genetics* **163**: 1389–1401.
Butler, M.J., Jacobsen, T.L., Cain, D.M., Jarman, M.G., Hubank, M., Whittle, J.R.S., *et al.* 2003. Discovery of genes with highly restricted expression patterns in the Drosophila wing disc using DNA oligonucleotide microarrays. *Dev.* **130** : 659–670.
Cayre, M., Strambi, C. & Strambi, A. 1994. Neurogenesis in an adult insect brain and its hormonal control. *Nature* **368**: 57–59.
Chung, H., Loehlin, D.W., Dufour, H.D., Vaccarro, K., Millar, J.G. & Carroll, S.B. 2014. A single gene affects both ecological divergence and mate choice in Drosophila. *Science.* **343**: 1148–51.
Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M. & Robles, M. 2005. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**: 3674–3676.
Cruickshank, T.E. & Hahn, M.W. 2014. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol. Ecol.* **23**: 3133–3157.
Csilléry, K., Blum, M.G.B., Gaggiotti, O.E. & François, O. 2010. Approximate Bayesian Computation (ABC) in practice. *Trends Ecol. Evol.* **25**: 410–418.
Csilléry, K., François, O. & Blum, M. 2012. Approximate Bayesian Computation (ABC) in R: A Vignette. *202.162.217.53* 1–21.
Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., *et al.* 2011. The variant call format and VCFtools. *Bioinformatics* **27** : 2156–2158.
Danley, P.D., Mullen, S.P., Liu, F., Nene, V., Quackenbush, J. & Shaw, K.L. 2007. A cricket Gene Index: a genomic resource for studying neurobiology, speciation, and molecular evolution. *BMC Genomics* **8**: 109.
Dauwalder, B., Tsujimoto, S., Moss, J. & Mattox, W. 2002. The Drosophila takeout gene is regulated by the somatic sex-determination pathway and affects male courtship behavior. *Genes Dev.* **16**: 2879–92.
DePristo, M.A., Banks, E., Poplin, R., Garimella, K. V, Maguire, J.R., Hartl, C., *et al.* 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**: 491–498.
Dodt, M., Roehr, J.T., Ahmed, R. & Dieterich, C. 2012. FLEXBAR—flexible barcode and adapter processing for next-generation sequencing platforms. *Biology.* **1**: 895–905.
Doorn, G.S. Van, Dieckmann, U. & Weissing, F.J. 2004. Sympatric Speciation by Sexual Selection : A Critical Reevaluation. *Am. Nat.* **163**: 709–725.
Drummond, A.J., Ho, S.Y.W., Rawlence, N. & Rambaut, A. 2007. A Rough Guide to BEAST 1.4.
Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. 2012. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29** : 1969–1973.
Earl, D. & vonHoldt, B. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**: 359–361. Springer Netherlands.
Ellegren, H., Smeds, L., Burri, R., Olason, P.I., Backström, N., Kawakami, T., *et al.* 2012. The genomic landscape of species divergence in Ficedula flycatchers. *Nature* 1–5.
Evanno, G., Regnaut, S. & Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology. *Mol. Ecol.* **14**: 2611–2620.
Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V.C. & Foll, M. 2013. Robust Demographic Inference from Genomic and SNP Data. *PLoS Genet.* **9**.
Excoffier, L. & Foll, M. 2011. fastsimcoal: a continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics* **27**: 1332–1334.
Excoffier, L. & Lischer, H.E.L. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**: 564–567.
Falush, D., Stephens, M. & Pritchard, J.K. 2003. Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. *Genet.* **164** : 1567–1587.
Feder, J.L., Flaxman, S.M., Egan, S.P., Comeault, A.A. & Nosil, P. 2013. Geographic Mode of Speciation and Genomic Divergence. *Annu. Rev. Ecol. Evol. Syst.* **44**: 73–97.
Fergus, D.J. & Shaw, K.L. 2013. Circadian rhythms and period expression in the Hawaiian cricket genus Laupala. *Behav. Genet.* **43**: 241–253.
Gavrilets, S. 2003. Perspective: models of speciation: what have we learned in 40 years? *Evolution* **57**: 2197–2215.
Gerhardt, H.C. & Huber, F. 2002. *Acoustic communication in insects and anurans*. The University of Chicago Press, Chicago.
Gray, D.A. 2005. Does courtship behavior contribute to species-level reproductive isolation in field crickets? *Behav. Ecol.* **16**: 201–206.

Gray, D.A. 2011. Speciation, Divergence, and the Origin of Gryllus rubens: Behavior, Morphology, and Molecules. *Insects* **2**: 195–209.
Gray, D.A. & Cade, W.H. 2000. Sexual selection and speciation in field crickets. *Proc Natl Acad Sci* **97**: 14449–14454.
Gray, D.A., Huang, H. & Knowles, L.L. 2008. Molecular evidence of a peripatric origin for two sympatric species of field crickets (Gryllus rubens and G. texensis) revealed from coalescent simulations and population genetic tests. *Mol. Ecol.* **17**: 3836–3855.
Gray, D.A., Walker, T.J., Conley, B.E. & Cade, W.H. 2001. A morphological means of distinguishing females of the cryptic field cricket species, Gryllus rubens and G. texensis (Orthoptera : Gryllidae). *Florida Entomol.* **84**: 314–315.
Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., *et al.* 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**: 1494–1512.
Heled, J. & Drummond, A.J. 2008. Bayesian inference of population size history from multiple loci. *BMC Evol. Biol.* **8**: 289.
Hennig, R.M., Heller, K.-G. & Clemens, J. 2014. Time and timing in the acoustic recognition system of crickets. *Front. Physiol.* **5**.
Ho, S.Y. & Shapiro, B. 2011. Skyline-plot methods for estimating demographic history from nucleotide sequences. *Mol. Ecol. Resour.* **11**: 423–434.
Holsinger, K.E. & Weir, B.S. 2009. Genetics in geographically structured populations: defining, estimating and interpreting FST. *Nat. Rev. Genet.* **10**: 639–650.
Hoy, R.R. 1974. Genetic Control of Acoustic Behavior in Crickets. *Am. Zool.* **14**: 1067–1080.
Izzo, A.S. & Gray, D. a. 2011. Heterospecific courtship and sequential mate choice in sister species of field crickets. *Anim. Behav.* **81**: 259–264.
Izzo, A.S. & Gray, D.A. 2004. Cricket song in sympatry: Species specificity of song without reproductive character displacement in Gryllus rubens. *Ann. Entomol. Soc. Am.* **97**: 831–837.
Kurusu, M., Cording, A., Taniguchi, M., Menon, K., Suzuki, E. & Zinn, K. 2008. A screen of cell-surface molecules identifies leucine-rich repeat proteins as key mediators of synaptic target selection. *Neuron* **59**: 972–85.
Lagisz, M., Wen, S.-Y., Routtu, J., Klappert, K., Mazzi, D., Morales-Hojas, R., *et al.* 2012. Two distinct genomic regions, harbouring the period and fruitless genes, affect male courtship song in *Drosophila montana*. *Heredity.* **108**: 602–8.
Lande, R.L. & Barrowclough, G.F. 1987. Effective population size, genetic variation, and their use in population management. In: *Viable populations for conservation* (M. E. Soule, ed), pp. 87–124. Cambridge University Press, Cambridge, NY.
Langmead, B. & Salzberg, S.L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**: 357–359.
Larson, E.L., Andrés, J.A., Bogdanowicz, S.M. & Harrison, R.G. 2013. Differential introgression in a mosaic hybrid zone reveals candidate barrier genes. *Evolution.* **67**: 3653–3661.
Lee, M., Paik, S.K., Lee, M.-J., Kim, Y.-J., Kim, S., Nahm, M., *et al.* 2009. Drosophila Atlastin regulates the stability of muscle microtubules and is required for synapse development. *Dev. Biol.* **330**: 250–62.
Li, H. & Durbin, R. 2011. Inference of human population history from individual whole-genome sequences. *Nature* **475**: 493–496.
Lin, C.P. & Danforth, B.N. 2004. How do insect nuclear and mitochondrial gene substitution patterns differ? Insights from Bayesian analyses of combined datasets. *Mol. Phylogenet. Evol.* **30**: 686–702.
Lin, T.-Y., Huang, C.-H., Kao, H.-H., Liou, G.-G., Yeh, S.-R., Cheng, C.-M., *et al.* 2009. Abi plays an opposing role to Abl in Drosophila axonogenesis and synaptogenesis. *Development* **136**: 3099–107.
Liu, T., Dartevelle, L., Yuan, C., Wei, H., Wang, Y., Ferveur, J.F., *et al.* 2009. Reduction of dopamine level enhances the attractiveness of male Drosophila to other males. *PLoS One* **4**: 1–5.
Liu, Z.C., Odell, N. & Geisbrecht, E.R. 2013. Drosophila importin-7 functions upstream of the Elmo signaling module to mediate the formation and stability of muscle attachments. *J. Cell Sci.* **126**: 5210–23.
Mace, K. & Tugores, A. 2004. The product of the split ends gene is required for the maintenance of positional information during Drosophila development. *BMC Dev. Biol.* **4**: 15.
Marcillac, F., Grosjean, Y. & Ferveur, J.-F. 2005. A single mutation alters production and discrimination of Drosophila sex pheromones. *Proc. Biol. Sci.* **272**: 303–309.
Maroja, L.S., Andrés, J.A. & Harrison, R.G. 2009. Genealogical discordance and patterns of introgression and selection across a cricket hybrid zone. *Evolution* **63**: 2999–3015.
Maroja, L.S., McKenzie, Z.M., Hart, E., Jing, J., Larson, E.L. & Richardson, D.P. 2014. Barriers to gene exchange in hybridizing field crickets: the role of male courtship effort and cuticular hydrocarbons. *BMC Evol. Biol.* **14**: 65.
Mendelson, T.C. & Shaw, K.L. 2005. Rapid speciation in an arthropod. *Nature* **433**: 375–376.
Metzker, M.L. 2010. Sequencing technologies - the next generation. *Nat. Rev. Genet.* **11**: 31–46.
Mevik, B.-H. & Wehrens, R. 2007. The pls Package: Principal Component and Partial Least Squares Regression in R. *J. Stat. Softw.* **18**: 1–24.
Moran, C.N. & Kyriacou, C.P. 2009. Functional neurogenomics of the courtship song of male Drosophila melanogaster. *Cortex* **45**: 18–34.
Morgan, K., Linton, Y.M., Somboon, P., Saikia, P., Dev, V., Socheat, D., *et al.* 2010. Inter-specific gene flow dynamics during the Pleistocene-dated speciation of forest-dependent mosquitoes in Southeast Asia. *Mol. Ecol.* **19**: 2269–2285.
Morin, P.A., Luikart, G. & Wayne, R.K. 2004. SNPs in ecology, evolution and conservation. *Trends Ecol. Evol.* **19**: 208–216.
Mukunda, L., Miazzi, F., Kaltofen, S., Hansson, B.S. & Wicher, D. 2014. Calmodulin modulates insect odorant receptor function. *Cell Calcium* **55**: 191–9.
Nadachowska-brzyska, K., Burri, R., Olason, P.I., Kawakami, T. & Ellegren, H. 2013. Demographic Divergence History of Pied Flycatcher and Collared Flycatcher Inferred from Whole-Genome Re-sequencing Data. *PLoS Genet.* **9**: e1003942.
Natarajan, R., Trivedi-Vyas, D. & Wairkar, Y.P. 2013. Tuberous sclerosis complex regulates Drosophila neuromuscular junction growth via the TORC2/Akt pathway. *Hum. Mol. Genet.* **22**: 2010–23.
Nei, M. & Gojobori, T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3** : 418–426.
Nicolaï, M., Lasbleiz, C. & Dura, J.-M. 2003. Gain-of-function screen identifies a role of the Src64 oncogene in Drosophila mushroom body development. *J. Neurobiol.* **57**: 291–302.
Nielsen, R. 2005. Molecular signatures of natural selection. *Annu. Rev. Genet.* **39**: 197–218.
Nosil, P. & Schluter, D. 2011. The genes underlying the process of speciation. *Trends Ecol. Evol.* **26**: 160–167.

Otte, D. 1992. Evolution of cricket songs. *J. Orthoptera Res.* **1**: 25–49.
Ptak, S.E. & Przeworski, M. 2002. Evidence for population growth in humans is confounded by fine-scale population structure. *Trends Genet.* **18**: 559–563.
R Development Core Team, R. 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.
Rambaut, A., Suchard, M.A., Xie, D. & Drummond, A. 2014. Tracer v1.6.
Reuter, J.E., Nardine, T.M., Penton, A., Billuart, P., Scott, E.K., Usui, T., *et al.* 2003. A mosaic genetic screen for genes necessary for Drosophila mushroom body neuronal morphogenesis. *Development* **130**: 1203–13.
Ritz, M.S. & Köhler, G. 2010. Natural and sexual selection on male behaviour and morphology, and female choice in a wild field cricket population: Spatial, temporal and analytical components. *Evol. Ecol.* **24**: 985–1001.
Rodriguez-Munoz, R., Bretman, A., Slate, J., Walling, C.A. & Tregenza, T. 2010. Natural and sexual selection in a wild insect population. *Science.* **79**: 1269–1272.
Ruegg, K., Anderson, E.C., Boone, J., Pouls, J. & Smith, T.B. 2014. A role for migration-linked genes and genomic islands in divergence of a songbird. *Mol. Ecol.* **23**: 4757–4769.
Sakai, T. & Ishida, N. 2001. Circadian rhythms of female mating activity governed by clock genes in Drosophila. *Proc. Natl. Acad. Sci.* **98**: 9221–9225.
Scheet, P. & Stephens, M. 2006. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *Am. J. Hum. Genet.* **78**: 629–644.
Schnorrer, F., Schönbauer, C., Langer, C.C.H., Dietzl, G., Novatchkova, M., Schernhuber, K., *et al.* 2010. Systematic genetic analysis of muscle morphogenesis and function in Drosophila. *Nature* **464**: 287–91.
Schöneich, S., Kostarakos, K. & Hedwig, B. 2015. An auditory feature detection circuit for sound pattern recognition. *Sci. Adv.* **1**: e1500325–e1500325.
Scordato, E.S.C., Symes, L.B., Mendelson, T.C. & Safran, R.J. 2014. The role of ecology in speciation by sexual selection: a systematic empirical review. *J. Hered.* **105 Suppl** : 782–94.
Seehausen, O. 2000. Explosive speciation rates and unusual species richness in haplochromine cichlid fishes: effects of sexual selection. *Adv. Ecol. Res.* **31**: 237–274.
Senthilan, P.R., Piepenbrock, D., Ovezmyradov, G., Nadrowski, B., Bechstedt, S., Pauls, S., *et al.* 2012. Drosophila auditory organ genes and genetic hearing defects. *Cell* **150**: 1042–54.
Servedio, M.R. 2015. Geography, assortative mating, and the effects of sexual selection on speciation with gene flow. Evol. Appl. doi: 10.1111/eva.12296.
Shaw, K.L. 1996. Polygenic Inheritance of a Behavioral Phenotype: Interspecific Genetics of Song in the Hawaiian Cricket Genus *Laupala*. *Evolution.* **50**: 256–266. Society for the Study of Evolution.
Shaw, K.L., Ellison, C.K., Oh, K.P. & Wiley, C. 2011. Pleiotropy, "sexy" traits, and speciation. *Behav. Ecol.* **22**: 1154–1155.
Shaw, K.L. & Lesnick, S.C. 2009. Genomic linkage of male song and female acoustic preference QTL underlying a rapid species radiation. *Proc. Natl. Acad. Sci.*. **106**: 9737–9742.
Sousa, V. & Hey, J. 2013. Understanding the origin of species with genome-scale data : modelling gene flow. *Nat. Publ. Gr.* **14**: 404–414.
Stern, D.L. 2014. Reported Drosophila courtship song rhythms are artifacts of data analysis. *BMC Biol.* **12**: 38.
Terriente-Félix, A. & de Celis, J.F. 2009. Osa, a subunit of the BAP chromatin-remodelling complex, participates in the regulation of gene expression in response to EGFR signalling in the Drosophila wing. *Dev. Biol.* **329**: 350–61.
Thomas, M.L. & Simmons, L.W. 2010. Cuticular hydrocarbons influence female attractiveness to males in the Australian field cricket, *Teleogryllus oceanicus*. *J. Evol. Biol.* **23**: 707–714.
Thomas, M.L. & Simmons, L.W. 2009. Sexual selection on cuticular hydrocarbons in the Australian field cricket, *Teleogryllus oceanicus*. *BMC Evol. Biol.* **9**: 162.
Tregenza, T. & Wedell, N. 1997. Definitive evidence for cuticular pheromones in a cricket. *Anim. Behav.* **54**: 979–84.
Tsai, P.-I., Kao, H.-H., Grabbe, C., Lee, Y.-T., Ghose, A., Lai, T.-T., *et al.* 2008. Fak56 functions downstream of integrin alphaPS3betanu and suppresses MAPK activation in neuromuscular junction growth. *Neural Dev.* **3**: 26.
Valakh, V., Walker, L.J., Skeath, J.B. & DiAntonio, A. 2013. Loss of the spectraplakin short stop activates the DLK injury response pathway in Drosophila. *J. Neurosci.* **33**: 17863–73.
Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., *et al.* 2013. From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Curr. Protoc. Bioinforma.* **43**: 1–11. J
Venables, W.N. & Ripley, B.D. 2002. *Modern applied statistics with S*. Springer, New York, NY.
Walker, T. 1998. Trilling field crickets in a zone of overlap (Orthoptera: Gryllidae: *Gryllus*). *Ann. Entomol. Soc. Am.* **91**: 175–184.
Walker, T.J. 2015. Crickets. In Singing Insects of North America. http://entnemdept.ifas.ufl.edu/walker/Buzz/crickets.htm.
Wang, Z., Gerstein, M. & Snyder, M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**: 57–63.
Wegmann, D. & Excoffier, L. 2010. Bayesian inference of the demographic history of chimpanzees. *Mol. Biol. Evol.* **27**: 1425–1435.
Wegmann, D., Leuenberger, C., Neuenschwander, S. & Excoffier, L. 2010. ABCtoolbox: a versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics* **11**: 116–122.
Weissing, F.J., Edelaar, P. & van Doorn, G.S. 2011. Adaptive speciation theory: a conceptual review. *Behav. Ecol. Sociobiol.* **65**: 461–480.
Wu, C.I. 2001. The genic view of the process of speciation. *J. Evol. Biol.* **14**: 851–865.
Yang, Z. & Nielsen, R. 2000. Estimating Synonymous and Nonsynonymous Substitution Rates Under Realistic Evolutionary Models. *Mol. Biol. Evol.* **17** : 32–43.
Zeng, V., Ewen-Campen, B., Horch, H.W., Roth, S., Mito, T. & Extavour, C.G. 2013. Developmental gene discovery in a hemimetabolous insect: de novo assembly and annotation of a transcriptome for the cricket Gryllus bimaculatus. *PLoS One* **8**: e61479.
Zhang, Z., Li, J., Zhao, X.-Q., Wang, J., Wong, G.K.-S. & Yu, J. 2006. KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. *Genomics. Proteomics Bioinformatics* **4**: 259–63.

SUPPORTING INFORMATION

- **Pipeline for obtaining haplotype sequences from RNA-seq data**
- **Supplementary tables S1-S3**
- **Supplementary figures S1-S3**

Pipeline for obtaining haplotype sequences from RNA-seq data

1. Create separate VCF files for each contig. Run vcftools with –-chr xxx where xxx is the desired transcript id. Example:

```
vcftools --vcf /home/rubtex.vcf --out /home/rubtex_subset1.1 --chr
c197865_g1_i1 --recode
```

2. Change from vcf to fastphase format by running vcf2fastphase.pl (from Laurie Stevison with permission: https://github.com/lstevison/vcf-conversion-tools/blob/master/vcf2fastPHASE.pl).

   Run as:
```
./vcf2fastPHASE.pl –input.vcf –outputfile1 –outputfile2 –samplesize
```

   The input file is the contig specific vcf file, the output files have to be specified. Example (40 individuals):

```
./vcf2fastphase.pl /home/rubtex_subset1.1.recode.vcf rubtex_subset1.1
rubtex_subset1.1a 40
```

3. Run fastphase for every file.

   Run as:
```
./fastPHASE_Linux -T20 –usubpops –outputfile –inputfile
```

   The input file is the first of the two output files in step 2. The output file gets the extension hapguess_switch.out. Example:

```
./fastPHASE_Linux -T20 -usubpops -oresults_subset1.1 -rubtex_subset1.1
```

4. Change fastphase files back to vcf by running fastphase2vcf.pl (from Laurie Stevison with permission: https://github.com/lstevison/vcf-conversion-tools/blob/master/fastPHASE2VCF.pl)

   Run as:
```
./fastphase2vcf.pl -inputfile1 –inputfile2 –outputfile.vcf –blocks
```

   "inputfile1" is the first output file in step 2, "inputfile2" is the hapguess_switch.out file from step 3. "–outputfile.vcf" specifies a new vcf file, "blocks" is the number of contigs (in this case 1). Example:

```
./fastphase2vcf.pl rubtex_subset1.1 results_subset1.1_hapguess_switch.out
/home/rubtex_subset1.1.phased.vcf 1
```

5. Merge all vcf files together using the vcf-concat tool from vcftools. Example:

```
./vcf-concat /home/rubtex_subset1.1.phased.vcf /home/rubtex_subset1.2.phased.vcf
/home/rubtex_subset1.3.phased.vcf > /home/rubtex_finalhaps_merged.vcf
```

6. Split merged vcf file by individual using vcftools. Example:

```
./vcftools --vcf /home/rubtex_finalhaps_merged.vcf --recode --out
/home/rub30038 --indv rub30038 --recode-INFO-all
```

7. Change vcf files to haplotypes with vcflib's vcfgeno2haplo program. Example:

```
./vcfgeno2haplo -r /home/reference_transcriptome.fasta -w 2000
/home/rub30038.recode.vcf > rub30038_haps.txt
```

8. Reformat haplotypes using custom script by Marie Jeschek. Script is attached below. Example:

```
python vcf2hap.py reference_haplotypes2.fasta rub30038_haps.txt
rub30038.haps.
```

Here the fasta file contains only the transcripts that you are making haplotypes for.

Python script (depends: BioPython):

```
from Bio import SeqIO



# parse arguments
parser = argparse.ArgumentParser(description="Read haplotype sequences from VCF
file.", formatter_class=argparse.RawTextHelpFormatter)
parser.add_argument("-r", required=True, metavar="referenceFile",
type=argparse.FileType('r'),help="Fasta file with reference.")
parser.add_argument("-v", required=True, metavar="vcffile",
type=argparse.FileType('r'),help="VCF file.")
parser.add_argument("-o", required=True, metavar="outputFile",
type=argparse.FileType('w'),help="New table with complete haplotype sequences.")
args = parser.parse_args()

# parse vcf and collect data
print("Parsing vcf ...")
snpData = {}
for line in args.v:
       if line.startswith("#"):
              continue

              (CHROM,POS,ID,REF,ALT,QUAL,FILTER,INFO,FORMAT,RUB) =
              line.strip("\n").split("\t")
       if CHROM in snpData:
              print("Warning: Contig is there twice: %s" % CHROM)
              alleles=[REF] + ALT.split(",")

       snpData[CHROM] = (alleles[int(RUB[0])] , alleles[int(RUB[2])])


# get reference sequence and write haplotypes
print("Reading reference ...")
reference={}
args.o.write("Contig\tHaplotype_1\tHaplotype_2\n")
for sR in SeqIO.parse(args.r,"fasta"):
       print("\t%s"%sR.id)
       if sR.id not in snpData:
              args.o.write("%s\t%s\t%s\n" % (sR.id,str(sR.seq),str(sR.seq)))
       else:
              args.o.write("%s\t%s\t%s\n" %
              (sR.id,snpData[sR.id][0],snpData[sR.id][1]))

print("Done.")
```

9. Place the phased sequences back into the original contigs using custom R script.

R script:

```
library(seqinr)

ref<-read.fasta("fastafilename.fasta", seqtype="DNA", as.string=TRUE,
forceDNAtolower = FALSE)
#this is the fasta file with the complete list of contigs

haplotypes_ref<-read.fasta("haplotype_ref.fasta", seqtype="DNA", as.string=TRUE,
forceDNAtolower = FALSE)
```

```
#this is the fasta file with only the contigs for which #haplotypes have been
reconstructed

hapfiles=dir(pattern="*txt.haps")
#hapfiles summarizes the file names of all haplotype output #files (with the
extension txt.haps)

old_haplotypes<-list()
new_haplotypes<-list()
haplotypes<-NULL

#first, concatenate the output haplotypes of all individuals
for(i in 1:length(hapfiles)) {

       old_haplotypes[[i]]<-as.data.frame(read.delim(hapfiles[i], sep="\t",
       header=TRUE))
       }

names(old_haplotypes)<-hapfiles


#Then, place these haplotypes back into the original contigs so #that each
individual has two phased copies of the full sequence #of each contig
for(j in 1:length(old_haplotypes)) {

       haplotypes<-old_haplotypes[[j]]
       for(i in seq(1:length(haplotypes_ref))) {

              contig<-match(names(haplotypes_ref[i]),names(ref))
              overlap<-regexpr(as.character(haplotypes_ref[i]),
              as.character(ref[contig]), fixed=TRUE)
              begin<-overlap[1]
              fin<-begin+attributes(overlap)$match.length
              haplotypes[i,4]<-paste(substring(ref[contig], first=1, last=begin-
              1),haplotypes[i,2],substring(ref[contig], first=fin,
              last=nchar(ref[contig])),sep="")
              haplotypes[i,5]<-paste(substring(ref[contig], first=1, last=begin-
              1),haplotypes[i,3],substring(ref[contig], first=fin,
              last=nchar(ref[contig])),sep="")
              colnames(haplotypes)[4:5]<-c("full_haplo_1","full_haplo_2")
              }

       new_haplotypes[[j]]<-haplotypes
       }

names(new_haplotypes)<-hapfiles

#From the list new_haplotypes the new haplotypes can be #extracted in any desired
way. For example, simply write a text #file for each individual
#with the haplotype sequences for each contig:

for(i in 1:length(new_haplotypes)) {
       write.table(new_haplotypes[[i]],
       paste("new_haplotypes_",names(new_haplotypes)[i],".txt",sep=""), sep="\t",
       quote=FALSE, row.names=FALSE)
       }

#Or create a fasta file for each contig that contains the two #haplotype sequence
for each individual, as would be the case #for BEAST.

#In our case we had two species (rub and tex) and we separated #out the sequences
of the two species first, then wrote them to #a fasta file.

#Each entry was labelled with the individual name (two entries #per individual)
new_haplotypes_rub<-list()
new_haplotypes_tex<-list()
n_rub<-20
n_tex<-20

for(i in 1:n_rub) {
       new_haplotypes_rub[[i]]<-new_haplotypes[[i]]
       }
```

```
for(i in 1:n_tex) {
       new_haplotypes_tex[[i]]<-new_haplotypes[[i+n_rub]]
       }
for(i in 1:length(new_haplotypes_rub)) {

       ind_haplotypes<-new_haplotypes_rub[[i]]
       for(j in 1:nrow(ind_haplotypes)) {

              temp_fasta<-list()
              temp_fasta[[1]]<-as.character(ind_haplotypes[j,4])
              names(temp_fasta)[1]<-strsplit(names(new_haplotypes)[i],"_")[[1]][3]
              temp_fasta[[2]]<-as.character(ind_haplotypes[j,5])
              names(temp_fasta)[2]<-strsplit(names(new_haplotypes)[i],"_")[[1]][3]

              filename=paste("rub_",as.character(ind_haplotypes[j,1]),".fasta",sep
              ="")

              write.fasta(sequences=temp_fasta, names=names(temp_fasta),
              file.out=filename, nbchar=60, open="a")
              }
       }


for(i in 1:length(new_haplotypes_tex)) {

       ind_haplotypes<-new_haplotypes_tex[[i]]
       for(j in 1:nrow(ind_haplotypes)) {

              temp_fasta<-list()
              temp_fasta[[1]]<-as.character(ind_haplotypes[j,4])
              names(temp_fasta)[1]<-
              strsplit(names(new_haplotypes)[i+20],"_")[[1]][3]
              temp_fasta[[2]]<-as.character(ind_haplotypes[j,5])
              names(temp_fasta)[2]<-
              strsplit(names(new_haplotypes)[i+20],"_")[[1]][3]

              filename=paste("tex_",as.character(ind_haplotypes[j,1]),".fasta",sep
              ="")

              write.fasta(sequences=temp_fasta, names=names(temp_fasta),
              file.out=filename, nbchar=60, open="a")
              }
       }
```

Table S1 The GO enrichment terms for the SNPs fixed between *G. texensis* and *G. rubens*. The GO term ID number, the GO term, the category (cellular component, CC; molecular function, MF; biological process, BP), and the FDR corrected P-value are shown.

GO-ID	Term	Category	P(FDR)
GO:0005783	endoplasmic reticulum	CC	2.10E-03
GO:0005811	lipid particle	CC	5.60E-03
GO:0005681	spliceosomal complex	CC	1.73E-02
GO:0012505	endomembrane system	CC	1.88E-02
GO:0035183	female germline ring canal inner rim	CC	2.72E-02
GO:0005746	mitochondrial respiratory chain	CC	3.23E-02
GO:0008250	oligosaccharyltransferase complex	CC	3.34E-02
GO:0030117	membrane coat	CC	3.34E-02
GO:0003729	mRNA binding	MF	5.46E-05
GO:0051015	actin filament binding	MF	1.50E-02
GO:0015036	disulfide oxidoreductase activity	MF	2.33E-02
GO:0032553	ribonucleotide binding	MF	3.27E-02
GO:0005525	GTP binding	MF	3.27E-02
GO:0004576	oligosaccharyl transferase activity	MF	3.34E-02
GO:0008187	poly-pyrimidine tract binding	MF	4.72E-02
GO:0006396	RNA processing	BP	9.92E-04
GO:0016071	mRNA metabolic process	BP	2.54E-03
GO:0071840	CC organization or biogenesis	BP	3.39E-03
GO:0051641	cellular localization	BP	6.76E-03
GO:0007265	Ras protein signal transduction	BP	8.27E-03
GO:0043462	regulation of ATPase activity	BP	9.64E-03
GO:0032879	regulation of localization	BP	1.14E-02
GO:0008103	oocyte microtubule cytoskeleton polarization	BP	1.20E-02
GO:0006909	phagocytosis	BP	1.28E-02
GO:0006886	intracellular protein transport	BP	1.49E-02
GO:0000381	regulation of alternative mRNA splicing, via spliceosome	BP	1.50E-02
GO:0046130	purine ribonucleoside catabolic process	BP	1.63E-02
GO:0072657	protein localization to membrane	BP	1.72E-02
GO:0031445	regulation of heterochromatin assembly	BP	1.72E-02
GO:0006184	GTP catabolic process	BP	2.36E-02
GO:0008340	determination of adult lifespan	BP	2.56E-02
GO:0042775	mitochondrial ATP synthesis coupled electron transport	BP	2.74E-02
GO:0006412	translation	BP	3.25E-02
GO:0008652	cellular amino acid biosynthetic process	BP	3.98E-02
GO:0006613	cotranslational protein targeting to membrane	BP	4.06E-02
GO:0009070	serine family amino acid biosynthetic process	BP	4.06E-02
GO:0045047	protein targeting to ER	BP	4.06E-02
GO:0010608	posttranscriptional regulation of gene expression	BP	4.28E-02
GO:0030534	adult behavior	BP	4.41E-02
GO:0006723	cuticle hydrocarbon biosynthetic process	BP	4.72E-02
GO:0042254	ribosome biogenesis	BP	4.76E-02

Table S2 Parameter estimation for the model with ancestral gene flow (AGF). For each of the parameters the power (R^2), the root mean squared error prediction (RMSEP), and the 95% confidence interval and median are shown. Abbreviations for the parameters: Ancestral population size (NANC), effective population sizes for G. rubens (NRUB) and G. texensis (NTEX), time since divergence (TSPLIT), time since cessation of gene flow (TISO), rate of migration from G. texensis into G. rubens populations (MTEX>>RUB), and vice versa (MRUB>>TEX).

	Prior		Validation		Posterior		
Parameter	minimum	maximum	R^2	RMSEP	2.5%	Median	97.5%
LOG_{10}(NANC)	4.0	6.0 (lu)	0.01	0.967	4.96	5.41	5.82
LOG_{10}(NRUB)	3.0	6.0 (lu)	0.92	0.276	4.60	4.71	4.79
LOG_{10}(NTEX)	3.0	6.0 (lu)	0.92	0.280	4.64	4.81	4.94
LOG_{10}(TSPLIT)	5.0	7.0 (lu)	0.01	0.998	5.23	6.01	6.82
LOG_{10}(TISOLATE)	3.0	7.0 (lu)	0.75	0.503	4.21	4.32	4.42
$M_{TEX>>RUB}$	0.01	0.5 (u)	0.06	0.969	0.02	0.08	0.17
$M_{RUB>>TEX}$	0.01	0.5 (u)	0.06	0.968	0.02	0.15	0.49

Table S3 Annotation of the 288 contigs with d*N*/d*S* > 1. For each of the contigs the d*N*/d*S* value, number of synonymous and nonsynonymous subtitutions, the annotation and the experimentally proven functions in *D. melanogaster* (Flybase.org) are shown.

Contig	d*N*/d*S*	# SNPs (fixed)	*D*	Annotation	Function
c214514_g1_i5	1.01	22 (0)	0.22	eip74ef	autophagy
c192755_g1_i2	1.71	2 (0)	0.23	scb	axon guidance and negative regulation of growth at neuromuscular junction
c214303_g1_i11	1.73	50 (0)	0.22	sdc	axon guidance, glial cell development , motor neuron axon guidance, regulation of synaptic growth and neuromuscular junction
c211780_g1_i3	1.35	47 (2)	0.34	spoonbill	border follicle cell formation/long-term memory/thermosensory behaviour
c222405_g1_i1	3.46	496 (0)	0.10	rya-r44f	calcium channel
c209955_g2_i5	1.08	17 (0)	0.12	nlaz	cell surface glycoprotein on neurons http://dev.biologists.org/content/121/1/135.short
c219219_g3_i8	1.15	3 (0)	0.08	cg18801	cellular response to gamma radiation/double strand break repair
c219887_g3_i3	1.75	27 (0)	0.22	clip-190	cellularization
c222209_g1_i5	1.58	58 (0)	0.09	cg6735	centrosome organization/sensory perception of sound/embryonic morphogenesis
c210654_g1_i6	1.17	21 (0)	0.19	cg6803	chaeta development/wing disc development
c209013_g1_i1	1.39	14 (0)	0.30	cg2467	chitine based embryonic cuticle biosynthetic process
c211056_g2_i1	1.53	3 (0)	0.23	cg15637	chitine based embryonic cuticle biosynthetic process
c221412_g1_i1	1.24	7 (0)	0.04	htl	CNS development/muscle development/wing morphogenesis
c211687_g2_i3	2.78	29 (0)	0.12	cnn	CNS development/peripheral NS development
c220628_g1_i2	1.14	7 (0)	0.22	eg:	CNS development/serotonergic neuron differentiation
c216158_g1_i3	2.70	39 (8)	0.37	a	compound eye development
c222058_g1_i2	5.20	31 (0)	0.21	yelow-2d	cutcicle pigment/yellow pathway
c221500_g1_i3	6.64	49 (1)	0.19	FAS2	cuticular hydrocarbon synthesis
c213075_g4_i3	1.02	3 (0)	0.14	cyp301a1	Cytochrome P450
c218351_g1_i2	1.05	27 (0)	0.11	ankyrin	cytoskeletal anchoring to plasma membrane
c221519_g3_i2	1.92	53 (0)	0.15	cg2056	defence response to bacteria/fungi
c220213_g3_i1	1.14	33 (0)	0.21	toll-7	defence response to virus
c216204_g4_i1	1.31	6 (1)	0.51	cg1732	GABA transporter, sodium symporter activity

c220200_g3_i1	1.72	43 (0)	0.17	gst2	glutathione metabolic process
c213887_g3_i2	1.33	18 (0)	0.09	ald	glycolysis
c221585_g1_i3	1.08	127 (0)	0.13	hml	hemostasis
c201296_g4_i1	1.71	3 (0)	0.18	su 3-9	histone methyltransferase
c205635_g1_i2	1.27	23 (0)	0.17	bm-40 sparc	larval feeding behaviour
c218223_g1_i1	1.46	25 (0)	0.25	cg9518	lateral inhibition
c204340_g1_i3	1.02	53 (3)	0.19	klg	learning/memory/olfactory learning
c220415_g1_i2	1.10	41 (0)	0.16	takeout	male courtship behavior
c211945_g6_i1	2.55	25 (0)	0.30	vkg	Malpighian tube morphogenesis/skeletal muscle development
c219331_g1_i1	2.50	29 (0)	0.12	h	mechano-sensory bristle morphogenesis
c215240_g1_i1	3.50	73 (1)	0.16	tequila	memory
c205585_g1_i2	1.08	4 (0)	0.20	trol	motor neuron axon guidance, defasciculation of motor neuron axon
c211178_g1_i6	1.25	11 (0)	0.11	cg4118	mRNA transport
c211176_g5_i1	3.62	36 (0)	0.30	cg11326 Thrombospondin	Muscle attachment/development
c222088_g1_i3	1.43	76 (0)	0.12	sp4	negative regulation of protein processing
c221584_g1_i4	1.81	47 (0)	0.08	cactus	nervous system development
c200506_g1_i2	1.38	18 (0)	0.08	cg11583	neurogenesis
c221491_g1_i2	2.42	11 (0)	0.06	cg9938	neurogenesis
c222235_g1_i3	2.68	23 (0)	0.11	cg8782	neurogenesis
c212371_g2_i1	4.51	40 (0)	0.25	pxd	neurogenesis/phagocytosis
c211575_g3_i2	2.78	136 (1)	0.16	scb	neuromuscular junction regulation/axon guidance/short-term memory
c221965_g1_i1	2.14	13 (0)	0.16	cg9778	neuron projection morphogenesis/neurotransmitter secretion
c195331_g1_i6	1.20	31 (6)	0.32	mael	oogenesis
c210435_g8_i4	1.24	8 (3)	0.51	cg13284	phagocytosis
c207900_g2_i3	1.23	13 (2)	0.32	ogt	polycomb gene
c206039_g3_i1	1.11	4 (1)	0.55	psc	polycomb gene/ axon guidance
c221934_g2_i1	1.66	3 (0)	0.24	cg6747	potassium ion transport
c219793_g2_i1	2.13	45 (0)	0.10	ptp61f	regulates two different embryonic signalling pathways
c214322_g1_i1	1.32	24 (0)	0.25	l efl	regulation of translational initiation by eIF2 alpha phosphorylation
c220756_g1_i3	1.11	20 (0)	0.22	spdo	sensory organ development/peripheral NS development
c201535_g1_i2	1.09	33 (0)	0.27	pgk	synaptic transmission
c218723_g1_i1	1.56	18 (0)	0.20	Diap-1	Negative regulation of apoptosis, antennal morphogenesis, sensory organ development
c191856_g1_i1	1.74	1 (0)	0.00	rna-directed dna polymerase from mobile element jockey-like	transposable element
c209168_g3_i1	1.23	7 (0)	0.15	cg5926	transposable element
c208708_g1_i1	2.46	58 (0)	0.12	slit	Axon guidance, dendrite morphogenesis, gonad development, chemotaxis, neuron differentiation
c216222_g2_i2	3.91	9 (0)	0.14	slit	Axon guidance, dendrite morphogenesis, gonad development, chemotaxis, neuron differentiation

c218802_g1_i1	5.02	44 (0)	0.18	blistery	wing morphogenesis
c201103_g3_i2	4.79	41 (0)	0.14	hemolymph lipopolysaccharide-binding protein	
c202325_g1_i5	1.00	26 (0)	0.11	cyclase-like precursor	
c204201_g7_i3	1.20	17 (1)	0.26	---NA---	
c207043_g4_i2	1.00	19 (1)	0.17	apolipoprotein d	
c210411_g2_i4	1.12	1 (1)	1.00	cbp80 20-dependent translation initiation factor	
c213122_g1_i1	1.95	18 (5)	0.37	hypothetical protein L798_08690	
c215101_g1_i1	1.35	1 (0)	0.68	hypothetical protein L798_14783	
c215778_g3_i1	1.44	42 (4)	0.29	trans-golgi network integral membrane protein 2	
c216200_g2_i1	1.66	12 (0)	0.29	collagen alpha-1 chain	
c217662_g1_i7	1.36	8 (0)	0.19	hypothetical protein L798_14783	
c217965_g1_i1	3.63	19 (0)	0.06	intracellular protein transport protein uso1-like isoform x1	
c218657_g5_i1	1.01	37 (0)	0.22	probable gpi-anchored adhesin-like protein pga55 isoform x1	
c220223_g1_i1	1.47	38 (0)	0.09	hypothetical protein LOTGIDRAFT_238761	
c220237_g1_i2	1.04	2 (0)	0.11	hemolymph lipopolysaccharide-binding protein	
c220753_g1_i3	3.47	63 (0)	0.20	hypothetical protein L798_13577	
c221118_g1_i3	1.75	4 (0)	0.19	serine protease p146	
c221213_g1_i6	1.91	135 (0)	0.14	cug-bp- and etr-3-like factor 2	
c222015_g1_i2	1.21	132 (0)	0.19	vitellogenin partial	
c222277_g1_i1	1.07	95 (0)	0.10	reticulocyte-binding protein 2-like protein a	
c222427_g1_i1	1.15	78 (0)	0.06	histone-lysine n-methyltransferase setmar-like	
c189536_g1_i1	1.60	9 (6)	0.73	cg4553	---NA---
c191246_g2_i1	2.27	1 (0)	0.16	cg9952	---NA---
c193362_g1_i1	1.13	10 (0)	0.07	cg7038	---NA---
c198306_g2_i1	1.98	4 (1)	0.31	cg1583	---NA---
c200758_g1_i1	1.07	8 (3)	0.45	cg15561	---NA---
c201659_g1_i1	1.23	15 (0)	0.26	cg12177	---NA---
c206346_g2_i3	1.44	49 (0)	0.09	cg9915	---NA---
c206972_g3_i1	3.15	9 (0)	0.17	cg17292	---NA---
c207631_g2_i2	1.26	52 (0)	0.22	cg8483	---NA---
c207659_g2_i1	1.36	50 (0)	0.11	cg8112	---NA---
c208488_g4_i1	1.14	22 (0)	0.40	cg7453	---NA---
c210706_g2_i3	1.28	19 (0)	0.15	p5cr	---NA---
c211987_g1_i1	1.16	26 (3)	0.36	cg9524	---NA---
c212269_g3_i1	1.00	22 (0)	0.12	bg:	---NA---
c212269_g3_i1	1.05	22 (0)	0.12	bg:	---NA---
c212666_g3_i2	2.78	28 (1)	0.20	cg17698	---NA---
c212983_g5_i2	5.19	10 (1)	0.25	accessory gland protein	---NA---
c213430_g1_i1	1.94	45 (0)	0.25	cg13458	---NA---

c213531_g3_i7	1.77	14 (0)	0.14	cg7047	---NA---
c213713_g1_i4	2.01	17 (2)	0.21	cg13124	---NA---
c213857_g1_i1	2.43	30 (0)	0.11	kek3	---NA---
c213857_g1_i1	1.49	30 (0)	0.11	kek3	---NA---
c215061_g3_i1	3.57	45 (0)	0.10	bg:	---NA---
c215901_g1_i5	1.40	7 (0)	0.15	cg1282	---NA---
c216448_g1_i1	1.30	52 (1)	0.40	cg1869	---NA---
c217450_g1_i3	1.34	9 (0)	0.10	cg4925	---NA---
c217582_g1_i1	2.42	72 (0)	0.16	cg7509	---NA---
c217619_g4_i2	1.53	1 (0)	0.05	cg3950	---NA---
c218209_g1_i3	1.33	62 (0)	0.14	cg10175	---NA---
c218475_g5_i2	1.46	50 (0)	0.17	cg5397	---NA---
c218475_g5_i2	1.25	50 (0)	0.17	cg5397	---NA---
c218622_g2_i1	6.03	43 (0)	0.20	cg7896	---NA---
c218978_g1_i1	1.83	1 (0)	0.05	cg9219	---NA---
c219034_g1_i1	4.14	6 (0)	0.30	cg5278	---NA---
c219113_g2_i2	1.39	28 (0)	0.19	cg14613	---NA---
c219929_g1_i1	1.30	66 (0)	0.12	cg7896	---NA---
c220099_g4_i1	2.17	47 (0)	0.10	cg3246	---NA---
c220845_g1_i5	2.00	2 (0)	0.15	cg17068	---NA---
c221131_g1_i7	1.74	58 (0)	0.16	bg:	---NA---
c221548_g1_i1	1.27	34 (0)	0.22	alpha-est9	---NA---
c222060_g1_i1	2.26	115 (0)	0.12	cg11883	---NA---
c222114_g3_i2	2.18	36 (0)	0.12	cg17061	---NA---
c222138_g1_i2	1.50	9 (0)	0.06	cg11079	---NA---

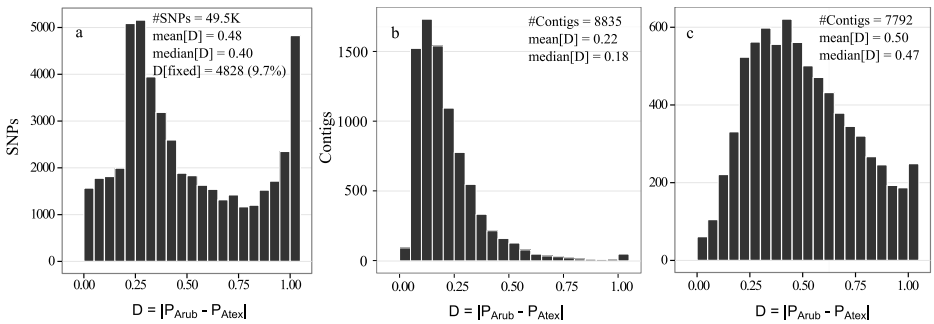

Figure S1 Genome-wide divergence. The distribution of the allele frequency difference (*D*) across SNPs with MAF cut-off = 0.1 (a), and distribution of *D* across contigs for MAF cut-off = 0.025 (b) and MAF cut-off = 0.1 (c).

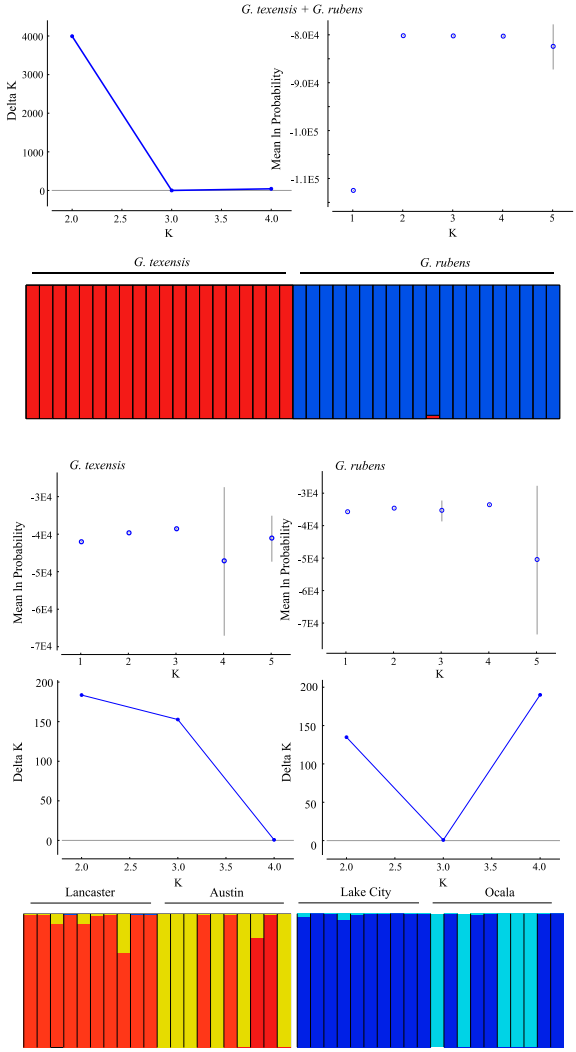

Figure S2 STRUCTURE results. For each of the species, STRUCTURE was ran for 100,000 iterations at values for K=1 through K=5. The mean natural logarithm of the probability and the delta K (increase or decrease in likelihood between consecutive runs for different values of K) were inspected to determine the most likely predicted number of populations. A run of *G. rubens* and *G. texensis* separately showed in both cases that, although the highest likelihood was for K>1, that the differences were only marginal with K=1 and that a defined pattern in population substructure was absent (as per the bar plots). The run for the species combined shows minimal introgression of *G. texensis* genes into the *G. rubens* samples and none vice versa.

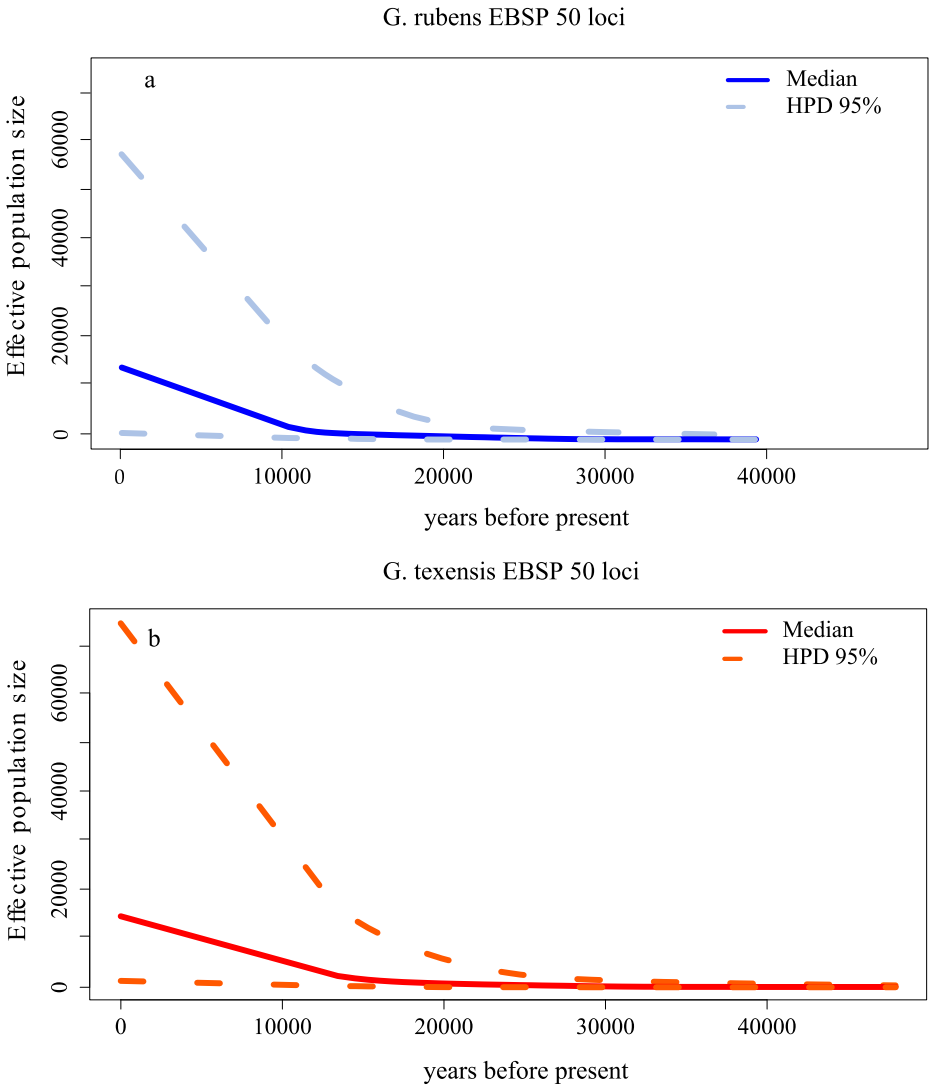

Figure S3 Extended Bayesian Skyline Plots. The estimate size change throughout the history of *G. rubens* (a) and *G. texensis* (b) are shown. A gradual increase after the last glacial period is seen for both species.

CHAPTER 6

GENERAL DISCUSSION

Signals used in mate choice are expected to change over time because some individuals produce signals that are more attractive than signals from other individuals of the same population; these individuals will have higher reproductive success causing the frequency of phenotypes in each subsequent generation to change relative to previous generations. However, the rate and magnitude of phenotypic evolution as well as whether sexual selection within populations eventually leads to reproductive isolation between populations depends on multiple factors. This thesis provides a broad, cross-disciplinary study of mate choice behaviour and acoustic divergence in North-American field crickets, combining behavioural, quantitative genetic, and population genomic analyses. The major results are: (1) Female crickets examine two key traits for song recognition. Interspecific divergence in the male song extends to traits that are not under selection by females and exceeds among species divergence in female preferences (Chapter 2); (2) Song and preference variation is associated with a complex, polygenic genetic architecture (Chapter 3), but comparing patterns of covariance among song traits with female preferences indicates high evolvability (Chapter 3 and 4); (3) The songs differ more strongly in patterns of covariation between species for which preference functions differ in shape than between species with similarly shaped preference functions. The different selection regimes between chirpers and trillers also strongly contrast the effects of sexual selection on the calling song (Chapter 3); (4) *Gryllus texensis* and *G. rubens* diverged in the face of gene flow but nevertheless show high functionally relevant divergence in loci matching *Drosophila melanogaster* genes involved in courtship behaviour, chemical mate choice, and neuromuscular properties that are potentially related to song production (Chapter 5). In the following sections, I will use these insights to discuss the role of song traits in reproductive isolation, the evolution of the calling song in crickets, and the role of sexual selection in speciation with gene flow.

Levels of song recognition and their contribution to reproductive isolation

In Chapter 2 we found that female preferences in three species of trilling field crickets largely followed the predictions from a recent mechanistic model for song recognition (Clemens & Hennig 2013, Hennig *et al.* 2014). In this model the selectivity for the temporal pattern of cricket songs by female preferences is described by a neuronal template that evaluates the timing of the pulses and an integration window that monitors the duration of the preferred pulse pattern on a longer timescale (Clemens & Hennig 2013, Hennig *et al.* 2014). There are thus three 'levels of song recognition': the peripheral filter for carrier frequency and the filters for pulse rate and chirp duty cycle in the central nervous system. Interestingly, the properties of the preference functions, the degree of trait-preference matching, and the relative contribution of the independent song traits to reproductive isolation varies across these three levels.

Tuning in on the right frequency

Preferences for the carrier frequency of the song constitute the first (and a peripheral) level of signal evaluation. As there is little interspecific divergence between the preference functions for carrier frequency as well as a disconnect between variation in preferences and variation in signals (Chapter 2, Fig. 2a therein) we concluded that it is unlikely that carrier frequency evolves as a result of (direct) sexual selection (Fig. 1a). Instead, it is more likely that interspecific variation in the male song and the lack of variation across female preferences result from covariation with other traits. The preferences for carrier frequency depend on the mechanical properties of the tympanic ear (Michelsen & Löhe 1995; Kostarakos *et al.* 2009) and are thus potentially constrained as a result of allometric integration. Variation in the carrier frequency of the male song may also depend on morphological constraints. The remarkably low levels of intraspecific variation in carrier frequency (coefficients of variation around 3%) are certainly more reminiscent of a morphological structure than a behavioural trait. Previous studies have provided evidence for a correlation between body size and carrier frequency (Gerhard & Huber 2002). Selection on male body size, *e.g.* via higher reproductive success of larger males (Saleh *et al.* 2014, but see Cade & Cade 1992), would then drive (co)variation in the carrier frequency of the song signal.

An alternative, but not mutually exclusive explanation for the diversity in the carrier frequency of the male signal is genetic linkage between loci controlling pulse rate and carrier frequency, resulting in correlated evolution of temporal and spectral traits. Both traits are, at least in part, controlled by sex-linked loci (Chapter 3, Fig. 1a,b therein). Although a positive correlation between pulse rate and carrier frequency is not a clear and consistent observation across species-specific **P** matrices, traits covary in all species (Chapter 4, Fig. 1 therein) and correlations are exceptionally strong in F_1 and F_2 *G. rubens*/*G. texensis* hybrid lines (Chapter 3, Fig. 3 therein). Additionally, interspecific divergence in the carrier frequency among both trilling (Chapter 2, Fig. 2a therein) and chirping species (RM Hennig, T Blankers, DA Gray, J. Comp. Phys, in review) tracks divergence in pulse rate. Therefore, variation in carrier

frequency potentially results from indirect selection effects through covariance with other song traits or through morphological integration, but is not driven by variation in mate choice preferences.

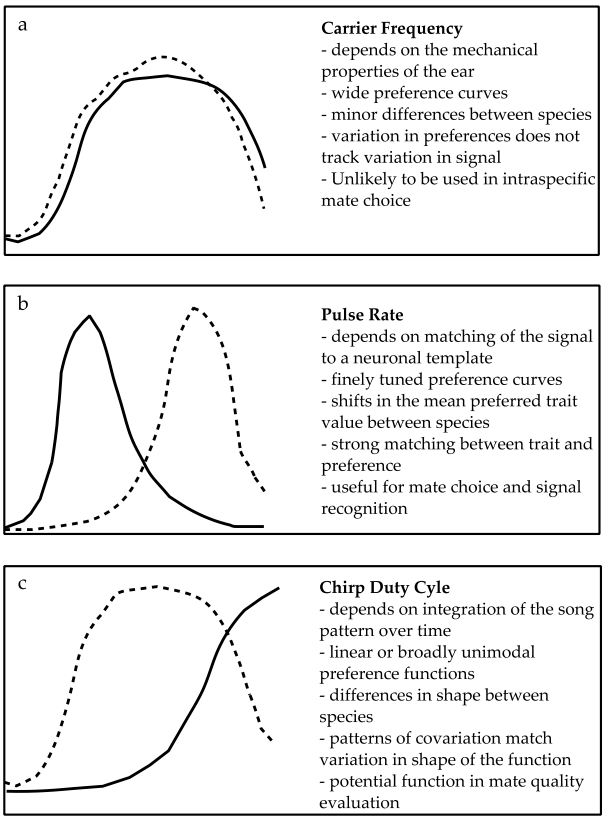

Figure 1 Schematic preference functions for the three levels of song recognition (Hennig *et al.* 2014): carrier frequency (a), pulse rate (b), chirp duty cycle (c). The solid and dashed lines represent average phonotactic scores for two (hypothetical) species across a range of stimuli varying in the respective trait.

Discriminating mates based on pulse rate

The second level of song recognition is manifested in pulse rate preference: the comparison of a neuronal representation of the song signal generated by the neurons in the ear to a short template for evaluation of the pulse pattern (Chapter 2, Fig. 3a-c therein). Preference functions are strongly tuned to conspecific pulse rates (Chapter 2, Fig. 2b therein) and it is the trait with the highest discriminatory power across Gryllids (Chapter 4, Fig. 1a therein, Alexander 1962, Otte 1992, Otte 1994, Mendelson & Shaw 2005; Fig. 1b).

The genetic architecture of pulse rate suggests rapid phenotypic evolution. We inferred a substantial role for the X-chromosome in the genetic control of pulse rate (Chapter 3, Fig. 1a therein), similar to studies on *Teleogryllus* (Bentley & Hoy 1972, Hoy 1974) and *Laupala* (Shaw 1996, Shaw *et al.* 2007, Oh *et al.* 2012) and predicted a high response to selection (Chapter 4). Additionally, a quantitative trait locus (QTL) for pulse rate variation in *Laupala* co-localizes with a QTL for pulse rate preference (Shaw & Lesnick 2009), suggesting that signal and preference can co-evolve through genetic coupling. High evolvability and strongly species-specific signals and preferences suggest that pulse rate plays a central role in the establishment and maintenance of reproductive isolation through acoustic signalling behaviour.

Integrating acoustic energy over time

An integration window evaluating the output of the pulse pattern filter over longer timescales constitutes the third level of signal recognition. In contrast to evaluation of the temporal pattern on the short timescale, the integration step depends on the duration (length of the time window), rather than the specific timing of the acoustic events (Hennig *et al.* 2014). The key trait for evaluation of the song on the long timescale is thus the chirp duty cycle (Fig. 1c). Differences between closely related species in chirp duty cycle are small and variation within species is high, indicating that it is unlikely to play a major role in reproductive isolation. However, in combination with pulse rate, chirp duty cycle contributes to separating species in phenotypic space (Fig. 2).

A genetic correlation between chirp duration and preferences for chirp duration (measured at constant chirp pause, thus corresponding to variation in the chirp duty cycle) was due to assortative mating (Gray & Cade 1999). This indicates there is some potential for trait-preference co-evolution for chirp duty cycle. The orchestrated divergence of chirp duty cycle and chirp duty cycle preferences between trilling and chirping species strongly discriminates species producing chirps from trilling species even when they sing at similar pulse rates (Fig. 2). Additionally, as will be discussed below, linear selection on the chirp duty cycle in species producing long trills has resulted in strong multivariate phenotypic divergence with potentially far-reaching consequences for the evolutionary trajectory of the song.

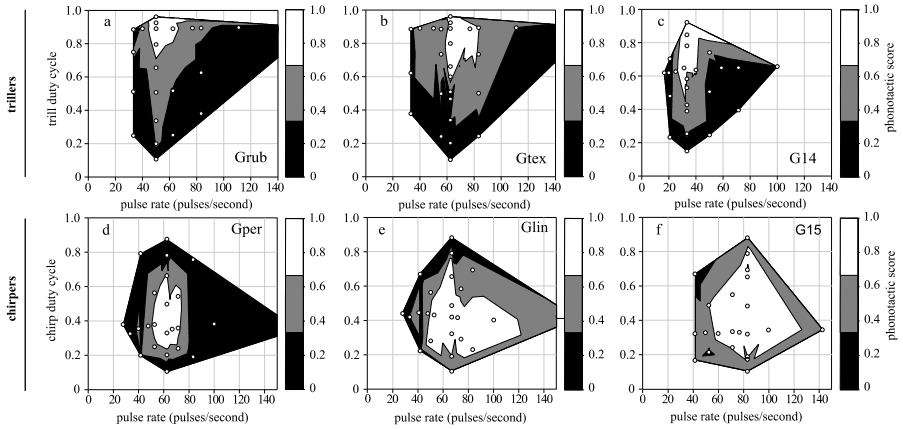

Figure 2 Song preference space. The phonotactic score is shown as a function of pulse rate and chirp duty cycle variation among artificial song stimuli (white circles). The preferences of three trillers [(a) *G. rubens,* (b) *G. texensis*, and (c) G#14; adapted from Blankers *et al.* 2015, see Chapter 1] and three chirpers [(d) *G. personatus*, (e) *G. lineaticeps*, and (f) G#15, adapted from Hennig *et al.* in review] are shown.

In the contemporary views on sexual selection, the assumption that traits convey some additional information about the sender or that preferences are under natural selection is commonplace (Prum 2012). In Chapter 2 we showed variation in the shape of the preference functions among song traits. Potentially, mate choice behaviour in cricket can thus encompass Fisherian or otherwise arbitrary selection and selection for 'good genes' or 'sexy sons' simultaneously by combining the output of preferences for different traits. Pulse rate and chirp duty cycle differ in both the shape of the preference function and the distribution of the male traits. Pulse rate is associated with finely tuned concave preference functions. Conversely, chirp duty cycle, which is a measure of acoustic energy of the song, is associated with higher variance in males and open ended preference functions, at least in trillers. Linear preferences for acoustic energy are a common feature across insects and anurans (*i.e.* 'motivational' [Popov & Shuvalov 1976] or 'dynamic' [Gerhardt 1991] traits, reviewed in Ryan & Keddy-Hector 1992). By contrast, closed preference functions are expected when preferences are selected to favour choosiness (a potential direct benefit of mate choice when finding and courting mates is costly, Wagner 2011) or to be recognized with high confidentiality (*i.e.* 'essential' [Popov & Shuvalov 1976] or 'static' [Gerhardt 1991] traits). One interpretation of mate choice behaviour in crickets across song components is thus that females evaluate different traits for different purposes and combine arbitrary and non-arbitrary (direct or indirect benefits) selection in a single signalling modality.

Pleiotropic effects and the role of selection in shaping the multivariate phenotype

Recognition of the temporal patterns of the calling song is thus constrained by pulse rate and chirp duty cycle (Chapter 2; RM Hennig, T Blankers, DA Gray, J. Comp. Phys, in review). Preference functions were similarly shaped across six species of field crickets (Fig. 2) with among species differences mostly limited to shifts in the mean preferred trait. However, the shape of the preference function for chirp duty cycle differed strongly between trillers and chirpers: females of trilling species showed linear preferences (Fig. 2a-c; Chapter 4, Table 3 therein), while females of chirpers showed concave preferences (Fig. 2d-f; RM Hennig, T Blankers, DA Gray, J. Comp. Phys, in review). These findings raise the question how songs have become so divergent among species in relatively short timeframes (Gray *et al.* 2008, Andrés *et al.* 2013, Chapter 5), despite limited divergence in female preferences. Two important findings that were discussed in chapter 3 and Chapter 4 provide an explanation for rapid divergence in song traits.

First, the temporal pattern of the pulse timescale appears to be, at least phenotypically, coupled to the temporal pattern of the chirp timescale (Chapter 4, see Bertram *et al.* 2011 for similar results).

This is shown by the strong covariation between pulse rate and chirp rate. Similarly, between several other pairs of song traits high correlations were found, although not as consistent across species as for pulse rate and chirp rate. Genetic or phenotypic correlations between song traits can result in evolutionary constraints when selection acts conflictingly on each of the traits in a correlated trait pair. Alternatively, when one of the traits is under strong directional selection while the other trait is not constrained in its response to selection, trait correlation can result in *indirect* selection effects (Chapter 1, Fig. 1 therein). Significant linear and quadratic preferences for chirp rate were virtually absent across species, indicating that chirp rate can get dragged along following selection for higher or lower pulse rates. We thus hypothesized that a strong positive correlation results in indirect selection on chirp rate as a consequence of direct selection on pulse rate, thus orchestrating divergence across timescales.

The reason for coupling of the timescales is unclear. Physiologically, neurons involved in generating the song pattern may be coupled or shared (Bentley 1977), although it has never been shown that the central pattern generators for the pulse and chirp are correlated physiologically despite a wealth of studies into the neuronal control of sound production in crickets (*e.g.* Hennig 1990, Hedwig 2000, Hedwig 2006, Schöneich & Hedwig 2011,2012). Alternatively, song variation across timescales may be coupled genetically. Both pulse rate and chirp rate show signs of X-linkage, although this is much weaker for the latter trait due to high levels of inter- and intra-individual variation (Chapter 3).

There is also some potential for pleiotropic genetic effects driving divergence in the temporal rhythm on the short and long timescale simultaneously. Screening the individual transcriptomes across populations of *G. texensis* and *G. rubens* highlighted elevated levels of divergence and signatures of selection on several genes involved in the circadian rhythm (*e.g.*, *period*; Chapter 5). The function of many clock genes extends beyond day-night rhythm regulation as we know from, the courtship song rhythm in *Drosophila melanogaster* (Kyriacou & Hall 1980, Lagisz *et al.* 2012, but see Stern 2014) and social learning in mammals (Nicholas *et al.* 2007). In *Laupala cerasina* and *L. paranigra* the pulse rate of the song correlates with song and locomotor rhythm (Fergus & Shaw 2013). This study also found interspecific coding divergence in *period* between the species that further differed in pulse rate and in song and locomotor activity period. Pleiotropic effects on rhythmic behaviours could potentially result in the co-evolution of the temporal pattern across timescales thus explaining how songs diverge strongly despite only minor differences in preferences.

A second explanation for high rates of divergence in song signals is that linear selection on chirp duty cycle has apparent strong effects on (co)variation in song traits, extending well beyond merely driving change in mean trait values (Chapter 4). Divergence in **P** is high between chirpers and trillers, but variation in **P** among trillers or chirpers is limited (Chapter 4, Fig. 2c therein) and mostly proportional (*i.e.*, the *relative* change in variances across eigenvectors is more or less similar; Chapter 4, Table 4 therein). This means that the orientation of the covariances has remained more or less constant across species that share a selection regime, but differs between species experiencing different selection regimes. Importantly, divergence in **P** translates to variation in the evolutionary trajectory of the songs, as chirpers and trillers differ strongly in the predicted effects of selection. This variation is mostly a consequence of the direction of covariation between chirp rate and chirp duty cycle. Thus, although pulse rate shows the largest univariate divergence among cricket species, variation in the temporal rhythm of the long timescale has profound effects on multivariate phenotypic evolution.

Theoretical studies suggest that selection may significantly alter the covariance structure (Jones *et al.* 2003, Arnold *et al.* 2008, Melo & Marroig 2014), but the role of selection in shaping **G** and **P** remains largely untested empirically (but see Blows *et al.* 2004, Hine *et al.* 2009, Berner *et al.* 2010, Kolbe *et al.* 2011, Roff & Fairbairn 2012). Chapter 4 provides a rare example of the extent to which directional sexual selection can alter the covariance structure of a behavioural sexual signal. However, although interesting on its own account, a comparative analysis of divergence in **P** primarily provides information about variation in the multivariate phenotype on which selection acts. Whether the variation in selection regimes translates to variation in selection responses requires additional insight into among species divergence in **G**. Additionally, to understand the full extent of the relationship between linear preference functions and variation in **G** and **P** matrices more comparative studies are vital.

Speciation with gene flow and sexual selection

In addition to the divergence in long-range acoustic communication discussed above, at least two other (close range) reproductive barriers exist between field cricket species (for a review see Veen *et al.* 2013, their Fig. 7). When a female has approached a male, he produces a species-specific courtship song which substantially reduces the success rate of heterospecific matings (Alexander 1962, Gray 2005). Additionally, a role for cuticular hydrocarbons in mate choice (in addition to desiccation resistance) has been implicated for many insect species (Blomquist & Bagnères 2010) including the cricket genera

Teleogryllus (Balakrishnan & Pollack 1997, Simmons *et al.* 2013) and *Gryllus* (Maroja *et al.* 2014). Mating behaviour in crickets, both on the short and the long range and across modalities is thus highly divergent among species and the data presented in this thesis suggest rapid evolution of acoustic mating signals in response to mating preferences.

In contrast, cricket species show very limited divergence in ecological characters (the one documented exception being ovipositor length) and the distributions of closely related species that are highly diverse in their mating signals are often overlapping (Alexander 1962, Otte 1992). Similar to African cichlid fish (Seehausen 2000) and Hawaiian swordtail crickets (Mendelson & Shaw 2005) the data here indicate a strong role for sexual selection in driving divergence. The results in Chapter 5 further support this hypothesis by suggesting relatively recent speciation in the order of less than 1 million years ago of *G. texensis* and *G. rubens*, a long period of bi-directional gene flow, and high divergence and signs of positive selection in genes that are involved in courtship behaviour, chemical (CHC) mate signalling, neuromuscular development (potentially involved in song production), and wing morphogenesis.

Although speciation with gene flow has traditionally been received with scepticism, more recent evidence suggests it may be much more common than previously thought (Bolnick & Fitzpatrick 2007, Bird *et al.* 2012). However, theoretical studies of parapatric or sympatric speciation by sexual selection indicate that additional forces of disruptive selection are required to maintain mating polymorphisms during the early stages of speciation (van Doorn *et al.* 2004, Bolnick & Fritzpatrick 2007, Weissing *et al.* 2011). Potential sources of disruptive selection that can interact with sexual selection to drive speciation include, but are not limited to, natural selection, phenology, geographic history, and disruptive selection on female mate preferences.

Natural selection can interact with sexual selection if assortative mating evolves with respect to an ecologically relevant character. In such a scenario, speciation can occur rapidly even in the presence of gene flow (Dieckmann & Doebeli 1999). This may be true for, for example, the *Heliconius* radiation, where species show extensive range overlap suggesting sympatric speciation (Rosser *et al.* 2015). The most divergent trait among these butterfly species is the wing colour pattern which is involved in Müllerian mimicry but also used in mate choice. For crickets it is however unclear how natural selection would contribute to divergence as crickets are extreme generalists and show only limited divergence in ecological characters.

Variation in the timing of reproduction can form strong prezygotic barriers to gene flow and there is a substantial amount of phenological variation among cricket species (Alexander 1968). However, closely related species typically display very similar life cycles. The timing of emerging adults and egg-laying are synchronous and the stage during diapause (either egg or early nymphal stages) is similar in (but differs among) each of the sister species pairs *G. veletis* and *G. fultoni*, *G. firmus* and *G. pennsylvanicus*, and *G. rubens* and *G. texensis* (Alexander 1968, see Huang *et al.* 2000 for a phylogeny). The latter two species also both have a bivoltine life cycle and adults emerge more or less synchronically both during the early spring and late summer mating season. Given that closely related species have very similar life cycles and variation is substantial only among deeper nodes of the phylogeny (Huang *et al.* 2000), it is unlikely that phenological variation contributes strongly to reproductive isolation in the earliest stages of speciation.

The geographic history can strongly influence the reach of gene flow, but is often largely unknown for extant species. The current distribution of *G. texensis* and *G. rubens* spans across a wide geographic range and includes a large zone of overlap. The genetic data indicate a long history of ancient gene flow thus suggesting sympatric or parapatric speciation (Chapter 5). However, our models in Chapter 5 also support a bottleneck in *G. rubens* thus corroborating an earlier study examining genetic divergence in *G. texensis* and *G. rubens*, where the authors suggested that the latter species derived from the former in peripatry (Gray *et al.* 2008). Exchange of migrants and the build-up of reproductive isolation may have thus been highly asymmetrical in the earliest stages of speciation because of the strongly reduced population sizes in *G. rubens*. Furthermore, when species diverge in peripatry the strength of sexual selection may be (temporarily) relaxed because at low densities selection on female preferences favours less stringent choosiness (Kaneshiro 1980, Ödeen & Florin 2002).

A final source of variation that theoretically affects the likelihood of speciation with gene flow by sexual selection is disruptive selection on female mating preferences (van Doorn *et al.* 2004). When access to males is in no way limited, as is often the assumption in theoretical models of sexual selection, female mate preferences are not expected to be under frequency-dependent selection. However, when males can invest a limited amount of energy or time into mating, when males compete for access to females (or shelter to attract females into), and females compete with other females for mating, selection on female mate preferences becomes frequency-dependent. In this scenario, speciation with gene flow

by sexual selection becomes possible (van Doorn *et al.* 2004). However, it is unclear how general this scenario would be in the natural world, as it requires the simultaneous action of independent biological processes (mate choice, male-male competition, and female-female competition) targeting the same traits (van Doorn *et al.* 2004).

Interestingly, field crickets do seem to meet these requirements to some extent. Both mate choice and male-male antagonistic interactions (*i.e.*, stepwise escalating fights for access to burrows; Hoffman & Stevenson 2000, Hofmann & Schildberger 2001) are in one way or another dependent on male song signals (*i.e.* mate attraction and courtship songs as well as aggressive songs used during and after aggressive encounters). Additionally, cricket females are polyandrous (Bretman & Tregenza 2005), generating competition for mates among females and resulting in selection on mating preferences themselves (Kvarnemo & Simmons 2013).

In summary, the results presented in this thesis, both from a behavioural and quantitative genetics angle and by examining genome-wide genetic variation, exemplify the extent to which sexual selection can drive divergence on relatively short timescales. Even in the face of gene flow, phenotypic and genetic variation between closely related species is biased towards traits related to mating behaviour in crickets. Furthermore, the expected evolutionary responses of the song are high, in part because of the genetic architecture facilitating high evolvability, in part because selection on mating signals can serve multiple purposes simultaneously.

Concluding remarks

In this thesis I aimed at integrating insights in the proximate mechanisms and ultimate consequences of mate choice behaviour in crickets. The behavioural and quantitative genetics studies in Chapter 2, 3, and 4 underscore that simply measuring single traits and their preference to infer evolutionary change can give a strongly distorted view. In these studies we have accomplished a very detailed, multivariate study of phenotypic evolution in response to female mate choice preference, but some important questions remain unanswered. First, although the **P** matrix gives important insight into the phenotype on which selection acts and can be useful in across species comparisons to understand phenotypic evolution, the selection response is ultimately constrained by the genetic constraints described by the **G** matrix. More important even is the question of whether genetic correlations between traits and preferences exist. In a forthcoming QTL mapping study, the genetic architecture of traits and preferences will be examined using genome-wide genetic markers in a large number of backcrossed progeny. We will directly address whether traits and preferences co-localize, thus indicating genetic linkage or pleiotropy. A pleiotropic genetic basis for traits and preferences is expected to greatly enhance co-evolution of signals and responses and facilitates speciation by rapid divergence of mating traits unrelated to ecological fitness (*i.e.* through the Fisherian runaway process, see also Shaw & Lesnick 2009 and Shaw *et al.* 2011 for examples).

Chapter 5 showcased the great potential for next-generation sequence data to revolutionize our understanding of the genomic ramifications of sexual selection in an evolutionary historical context. More studies in this quickly developing field, especially for model systems of sexual selection, are greatly needed. Chapter 5 also underlined the importance of a thorough understanding of the mechanistic basis of the traits that are expected to be influenced by selection. Especially for behavioural traits, it is unlikely that bottom-up (*e.g.* QTL mapping studies) or top-down (*e.g.* population genomics scans) approaches will promptly pin down the loci under selection that underlie the phenotype of interest; the complexity of traits involved in (mating) behaviour prohibits easy access to information about their genetic basis. However, by integrating behavioural, quantitative genetics, and genomic studies as well as by keeping track of the ecological context in which divergence takes place we can advance our knowledge on the role of mate choice and sexual selection in speciation. This will be of great importance in manging our ever changing environment, in the conservation of biodiversity, and in understanding the fascinating diversity of secondary sexual characters.

REFERENCES

Alexander, R. 1962. Evolutionary change in cricket acoustical communication. *Evolution* **16**: 443–467.

Alexander, R. 1968. Life cycle origins, speciation, and related phenomena in crickets. *Q. Rev. Biol.* **43**.

Andrés, J.A., Larson, E.L., Bogdanowicz, S.M. & Harrison, R.G. 2013. Patterns of transcriptome divergence in the male accessory gland of two closely related species of field crickets. *Genetics* **193**: 501–513.

Arnold, S.J., Bürger, R., Hohenlohe, P. a, Ajie, B.C. & Jones, A.G. 2008. Understanding the evolution and stability of the G-matrix. *Evolution* **62**: 2451–61.

Balakrishnan, R. & Pollack, G. 1997. The role of antennal sensory cues in female responses to courting males in the cricket *Teleogryllus oceanicus*. *J. Exp. Biol.* **200**: 511–22.

Bentley, D. 1977. Control of cricket song patterns by descending interneurons. *J. Comp. Physiol. A* **116**: 19–38.

Bentley, D.R. & Hoy, R.R. 1972. Genetic control of the neuronal network generating cricket (*Teleogryllus Gryllus*) song patterns. *Anim. Behav.* **20**: 478–492.
Berner, D., Stutz, W.E. & Bolnick, D.I. 2010. Foraging trait (co)variances in stickleback evolve deterministically and do not predict trajectories of adaptive diversification. *Evolution* **64**: 2265–2277.
Bertram, S.M., Fitzsimmons, L.P., McAuley, E.M., Rundle, H.D. & Gorelick, R. 2011. Phenotypic covariance structure and its divergence for acoustic mate attraction signals among four cricket species. *Ecol. Evol.* **2**: 181–195.
Bird, C.E., Fernandez-Silva, I., Skillings, D.J. & Toonen, R.J. 2012. Sympatric Speciation in the Post "Modern Synthesis" Era of Evolutionary Biology. *Evol. Biol.* **39**: 158–180.
Blomquist, G.J. & Bagnères, A.-G. 2010. *Insect Hydrocarbons: Biology, Biochemistry, and Chemical Ecology*. Cambridge University Press, Cambridge, NY.
Blows, M.W., Chenoweth, S.F. & Hine, E. 2004. Orientation of the genetic variance-covariance matrix and the fitness surface for multiple male sexually selected traits. *Am. Nat.* **163**: 329–340.
Bolnick, D.I. & Fitzpatrick, B.M. 2007. Sympatric Speciation : Models and Empirical Evidence. *Annu. Rev. Ecol. Evol. Syst.* **38**: 459–487.
Bretman, A. & Tregenza, T. 2005. Measuring polyandry in wild populations: a case study using promiscuous crickets. *Mol. Ecol.* **14**: 2169–79.
Cade, W.H. & Cade, E.S. 1992. Male mating success, calling and searching behaviour at high and low densities in the field cricket, *Gryllus* integer. *Anim. Behav.* **43**: 49–56.
Clemens, J. & Hennig, R.M. 2013. Computational principles underlying the recognition of acoustic signals in insects. *J. Comput. Neurosci.* **35**: 75–85.
Dieckmann, U. & Doebeli, M.O. 1999. On the origin of species by sympatric speciation. *Nature* **400**: 354–357.
Doorn, G.S. Van, Dieckmann, U. & Weissing, F.J. 2004. Sympatric Speciation by Sexual Selection : A Critical Reevaluation. *Am. Nat.* **163**: 709–725.
Fergus, D.J. & Shaw, K.L. 2014. Cricadian rhythms and period expression in the Hawaiian cricket genus Laupala. *Behav. Genet.* **43**: 241–253.
Gerhardt, H.C. 1991. Female choice in treefrogs: static and dynamic acoustic criteria. *Anim. Behav.* **42**: 615–636.
Gerhardt, H.C. & Huber, F. 2002. *Acoustic communication in insects and anurans*. The University of Chicago Press, Chicago.
Gray, D. a., Huang, H. & Knowles, L.L. 2008. Molecular evidence of a peripatric origin for two sympatric species of field crickets (*Gryllus rubens* and *G. texensis*) revealed from coalescent simulations and population genetic tests. *Mol. Ecol.* **17**: 3836–3855.
Gray, D. & Cade, W. 1999. Quantitative genetics of sexual selection in the field cricket, *Gryllus integer*. *Evolution* **53**: 848–854.
Gray, D.A. 2005. Does courtship behavior contribute to species-level reproductive isolation in field crickets? *Behav. Ecol.* **16**: 201–206.
Hedwig, B. 2000. Control of cricket stridulation by a command neuron: efficacy depends on the behavioral state. *J. Neurophysiol.* **83**: 712–722.
Hedwig, B. 2006. Pulses , patterns and paths : neurobiology of acoustic behaviour in crickets. J. Comp. Physiol. A 192: 677–689.
Hennig, R.M., Blankers, T. & Gray, D.A. n.d. Divergence in male cricket song and female preference functions in three allopatric sister species. *J. Comp. Physiol. A*.
Hennig, R.M., Heller, K.-G. & Clemens, J. 2014. Time and timing in the acoustic recognition system of crickets. *Front. Physiol.* **5**.
Hine, E., Chenoweth, S.F., Rundle, H.D. & Blows, M.W. 2009. Characterizing the evolution of genetic variance using genetic covariance tensors. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **364**: 1567–1578.
Hofmann, H.A. & Schildberger, K. 2001. Assessment of strength and willingness to fight during aggressive encounters in crickets. *Anim. Behav.* **62**: 337–348.
Hofmann, H.A. & Stevenson, P.A. 2000. Flight restores fight in crickets. *Nature* **403**: 613.
Hoy, R.R. 1974. Genetic Control of Acoustic Behavior in Crickets. *Am. Zool.* **14**: 1067–1080.
Huang, Y., Ortí, G., Sutherlin, M., Duhachek, A. & Zera, A. 2000. Phylogenetic relationships of north American field crickets inferred from mitochondrial DNA data. *Mol. Phylogenet. Evol.* **17**: 48–57.
Jones, A.G., Arnold, S.J. & Bürger, R. 2003. Stability of the g -matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution* **57**: 1747–1760.
Kaneshiro, K.Y. 1980. Sexual isolation, speciation and the direction of evolution. *Evolution* **34**: 437–444.
Kolbe, J.J., Revell, L.J., Szekely, B., Brodie III, E.D. & Losos, J.B. 2011. Convergent evolution of phenotypic integration and its alignment with morphological diversification in caribbean Anolis ecomorphs. *Evolution* **65**: 3608–3624.
Kostarakos, K., Hennig, M.R. & Römer, H. 2009. Two matched filters and the evolution of mating signals in four species of cricket. *Front Zool* **6**: 22.
Kvarnemo, C. & Simmons, L.W. 2013. Polyandry as a mediator of sexual selection before and after mating. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **368**: 20120042.
Kyriacou, C.P. & Hall, J.C. 1980. Circadian rhythm mutations in *Drosophila melanogaster* affect short-term fluctuations in the male's courtship song. *Proc. Natl. Acad. Sci.* **77**: 6729–6733.
Lagisz, M., Wen, S.-Y., Routtu, J., Klappert, K., Mazzi, D., Morales-Hojas, R., *et al.* 2012. Two distinct genomic regions, harbouring the period and fruitless genes, affect male courtship song in *Drosophila montana*. *Heredity* **108**: 602–8.
Maroja, L.S., McKenzie, Z.M., Hart, E., Jing, J., Larson, E.L. & Richardson, D.P. 2014. Barriers to gene exchange in hybridizing field crickets: the role of male courtship effort and cuticular hydrocarbons. *BMC Evol. Biol.* **14**: 65.
Melo, D. & Marroig, G. 2014. Directional selection can drive the evolution of modularity in complex traits. *Proc. Natl. Acad. Sci.* **112**.
Mendelson, T.C. & Shaw, K.L. 2005. Rapid speciation in an arthropod. *Nature* **433**: 375–376.
Michelsen, A. & Löhe, G. 1995. Tuned directionality in cricket ears. *Nature* **375**: 639.
Nicholas, B., Rudrasingham, V., Nash, S., Kirov, G., Owen, M.J. & Wimpory, D.C. 2007. Association of Per1 and Npas2 with autistic disorder: support for the clock genes/social timing hypothesis. *Mol. Psychiatry* **12**: 581–592.
Ödeen, A. & Florin, A.B. 2002. Sexual selection and peripatric speciation: The Kaneshiro model revisited. *J. Evol. Biol.* **15**: 301–306.

Oh, K.P., Fergus, D.J., Grace, J.L. & Shaw, K.L. 2012. Interspecific genetics of speciation phenotypes: Song and preference co-evolution in Hawaiian crickets. *J. Evol. Biol.* **25**: 1500–1512.
Otte, D. 1992. Evolution of cricket songs. *J. Orthoptera Res.* **1**: 25–49.
Otte, D. 1994. *The Crickets of Hawaii: Origin, Systematics, and Evolution*. Orthoptera Society/Academy of Natural Sciences of Philadelphia, Philadelphia, PA.
Popov, a. V. & Shuvalov, V.F. 1977. Phonotactic behavior of crickets. *J. Comp. Physiol. A* **119**: 111–126.
Prum, R.O. 2012. Aesthetic evolution by mate choice: Darwin's really dangerous idea. *Philos. Trans. R. Soc. B Biol. Sci.* **367**: 2253–2265.
Roff, D.A. & Fairbairn, D.J. 2012. The evolution of trade-offs under directional and correlational selection. *Evolution* **66**: 2461–2474.
Rosser, N., Kozak, K.M., Phillimore, A.B. & Mallet, J. 2015. Extensive range overlap between heliconiine sister species: evidence for sympatric speciation in butterflies? *BMC Evol. Biol.* **15**: 125.
Ryan, M. & Keddy-Hector, A. 1992. Directional patterns of female mate choice and the role of sensory biases. *Am. Nat.* **139**: S4–S35.
Saleh, N.W., Larson, E.L. & Harrison, R.G. 2014. Reproductive Success and Body Size in the Cricket *Gryllus firmus*. *J. Insect Behav.* **27**: 346–356.
Schöneich, S. & Hedwig, B. 2011. Neural basis of singing in crickets: Central pattern generation in abdominal ganglia. *Naturwissenschaften* **98**: 1069–1073.
Schöneich, S. & Hedwig, B. 2012. Cellular basis for singing motor pattern generation in the field cricket (Gryllus bimaculatus DeGeer). Brain Behav. 2: 707–725.
Seehausen, O. 2000. Explosive speciation rates and unusual species richness in haplochromine cichlid fishes: effects of sexual selection. *Adv. Ecol. Res.* **31**: 237–274.
Shaw, K.L. 1996. Polygenic Inheritance of a Behavioral Phenotype: Interspecific Genetics of Song in the Hawaiian Cricket Genus Laupala. *Evolution* **50**: 256–266.
Shaw, K.L. & Lesnick, S.C. 2009. Genomic linkage of male song and female acoustic preference QTL underlying a rapid species radiation. *Proc. Natl. Acad. Sci.* **106**: 9737–9742.
Shaw, K.L., Parsons, Y.M. & Lesnick, S.C. 2007. QTL analysis of a rapidly evolving speciation phenotype in the Hawaiian cricket Laupala. *Mol. Ecol.* **16**: 2879–2892.
Shaw, K.L., Ellison, C.K., Oh, K.P. & Wiley, C. 2011. Pleiotropy, "sexy" traits, and speciation. Behav. Ecol. 22: 1154–1155.
Simmons, L.W., Thomas, M.L., Simmons, F.W. & Zuk, M. 2013. Female preferences for acoustic and olfactory signals during courtship: Male crickets send multiple messages. *Behav. Ecol.* **24**: 1099–1107.
Stern, D.L. 2014. Reported Drosophila courtship song rhythms are artifacts of data analysis. *BMC Biol.* **12**: 38.
Veen, T., Faulks, J., Tyler, F., Lloyd, J. & Tregenza, T. 2013. Diverse reproductive barriers in hybridising crickets and maintenance of isolation. *Evol. Ecol.* **27**: 993–1015.
Wagner, C.E., Harmon, L.J. & Seehausen, O. 2012. Ecological opportunity and sexual selection together predict adaptive radiation. *Nature* **487**: 366–369.
Wagner, W.E. 2011. Direct Benefits and the Evolution of Female Mating Preferences. Conceptual Problems, Potential Solutions, and a Field Cricket. *Adv. Study Behav.* **43**: 273–319.
Weissing, F.J., Edelaar, P. & van Doorn, G.S. 2011. Adaptive speciation theory: a conceptual review. *Behav. Ecol. Sociobiol.* **65**: 461–480.

ABSTRACT

Almost a century and a half since Darwin's second book on evolutionary theory, *The descent of man, and selection in relation to sex*, we still lack a comprehensive understanding of the role of sexual selection in speciation. A major caveat to the synthesis of sexual selection and speciation is the limited knowledge on the geographic and demographic historical context in which populations diverge. On the one hand, many speciation events occur in sympatry or parapatry, thus involving gene flow. Simultaneously we observe that among closely related species mating signals are often the most divergent characteristics, thus implying a strong role for sexual selection in driving the initial barrier to gene exchange between populations. On the other hand, theoretical research predicts that sexual selection alone is unlikely to facilitate speciation in the presence of gene flow.

This thesis integrates insights from neuro-ethological, behavioural, quantitative genetics, and genomic approaches to provide novel insights in the role of sexual selection in speciation, in particular focussing on speciation with gene flow. Few organisms are as suitable for such interdisciplinary endeavours as field crickets (*Gryllus*). Many species of the speciose genus *Gryllus* have geographically overlapping distributions, differ in mating behaviour, but are largely morphologically and ecologically similar. Crickets use acoustic signals in mate attraction and are strongly reproductively isolated by variation in mating behaviour. The neural and behavioural basis of acoustic mating behaviour in crickets is well understood. The signal produced by males is a multicomponent calling song which varies between species in the pitch (carrier frequency) and in the temporal pattern on two timescales: (i) pulses, each produced by a single closing movement of the forewings, and (ii) chirps, groups of pulses that can either be relatively short and regularly interspaced by pauses or consist of long trains of pulses which are then also referred to as trills. The temporal pattern is described by the repetition rate of single pulses and chirps, *i.e.* pulse rate and chirp rate, and the acoustic energy, *i.e.* pulse or chirp duty cycle, which is the duration of a pulse or chirp relative to the period. The well-described and highly polymorphic sexual behaviour and the limited divergence in ecological characters among geographically overlapping species render crickets an ideal system for the study of sexual selection and speciation with gene flow.

The aims in this thesis are twofold: (1) understanding the evolution of the acoustic mating signal in crickets and its role in reproductive isolation, and (2) unravelling the contribution of demography, gene flow, and (sexual) selection to genetic divergence and speciation. An explicit multivariate context was used to analyse female mate choice behaviour as well as assess the effects of selection on song traits resulting from mate choice. In **Chapter 2** we studied the behavioural basis of song recognition and compared the shape of female preference functions among different male traits and among species. This illuminated the role of specific song traits in reproductive isolation and the mechanistic basis of mate choice. In **Chapter 3** the genetic architecture of song traits and preferences was examined by generating interspecific hybrid crosses between the two sympatric sister species *G. rubens* and *G. texensis*. This provided insight into the expected response to selection thus revealing whether the calling song and song preferences are expected to evolve rapidly. In **Chapter 4** the multivariate examination of song variation was expanded by quantifying interspecific divergence in the phenotypic variance-covariance (**P**) matrix of the song and evaluating multivariate phenotypic evolution in the cricket mating signal. Specifically, we questioned whether the shape of the preference landscape could influence the orientation of **P** and thus affect the evolutionary trajectory of the song in addition to driving directional variation in trait means. In **Chapter 5** we then embarked on a molecular approach using RNA-seq data. Using the sympatric sister species *G. rubens* and *G. texensis* we examined the genome-wide patterns of functionally relevant genetic variation, inferred the demographic history to study the role of gene flow, and looked for footprints of selection across the genome.

In three species of field crickets that produce long trills we found that the shape of female preference functions appears to be mostly conserved across species (**Chapter 2**). The first level of song recognition is manifested in a peripheral filter that is sensitive to a range of carrier frequencies and is highly similar among species. The evaluation of the temporal pattern follows the predictions from a recent model for song recognition: Multivariate preferences demonstrate selectivity for conspecific pulse rates, thus indicating that on the short timescale the timing of pulses is an important predictor of mate choice. On the long timescale females show a preference for high chirp duty cycles, indicating that evaluation of the chirp timescale is independent of the timing of chirps. Results for three species producing short chirps corroborated these findings to a large extent. However, in the chirping species the preference functions for the chirp duty cycle are concave, in strong contrast to the linear preference functions observed in trillers (**Chapter 6**). Song preferences thus depend on few traits and showed strong similarities across species while the male song has diverged strongly in all traits (**Chapters 2 and 4**). Because the different levels of song recognition are characterized by different types of preferences functions, it is conceivable that multivariate preferences can extract various cues for mate quality from

the different traits simultaneously (*i.e.* pulse rate as an arbitrary 'recognition' trait and chirp duty cycle as an indicator of male fitness).

To explore the evolutionary effects of the selection resulting from the female preferences studied in **Chapter 2**, we examined the genetic architecture of song and preference variation. On the one hand, the polygenic genetic architecture found for song traits and preferences would probably limit selection responses and thus divergence rates. Sex-chromosomal inheritance of several song traits may have allowed for somewhat higher rates (**Chapter 3**). On the other hand, the major traits under selection, *i.e.* pulse rate and chirp duty cycle, could vary freely and were not constrained in their response to selection. Importantly, strong covariance was found between traits that are under sexual selection and traits that are not directly selected by females. This indicates that *indirect* selection may be responsible in part for the striking multivariate divergence in the male calling song despite limited divergence in female preferences (**Chapters 3 and 4**). Furthermore, comparing **P** between species producing chirps (with concave preference functions for chirp duty cycle) and species producing trills (with linear preference functions for chirp duty cycle) showed that the shape of the preference function can affect the orientation of trait covariance and the predicted response of the male song to selection (**Chapter 4**). These findings highlight the potential for sexual selection to shape the multivariate phenotype available to selection. They also show that the shape of the selection landscape can alter the evolutionary liability, even among closely related species. This is expected to strongly enhance the predicted effects of sexual selection on the evolution of reproductive isolation.

Finally, an RNA-seq approach was taken to access functionally relevant genomic SNP data and to examine patterns of genetic divergence, the potential effects of gene flow, and signatures of selection in the sympatric sister species *G. rubens* and *G. texensis* (**Chapter 5**). Using coalescent simulations we found strong support for a bottleneck in the derived species (*G. rubens*) and a long history of ancient gene flow that ceased some time during the last Pleistocene glacial cycles. Many of the outlier contigs in the analyses of overall genetic divergence and the analyses of the rates of amino acid changing substitutions matched *Drosophila melanogaster* genes with experimentally proven functions in neuromuscular development (potentially related to stridulatory movement), courtship behaviour, and chemical mating behavior. This observation revealed a potential strong role for sexual selection in driving divergence with gene flow. Together, the findings in **Chapters 2, 3, 4, and 5** underline the potential for sexual selection to drive reproductive isolation. Importantly, they make clear that for a comprehensive understanding of the role of sexual selection in speciation, interdisciplinary approaches are vital. However, in **Chapter 6** I also emphasize there are many other aspects of mating behaviour in crickets that have not been addressed in this thesis, but can alter the effects of sexual selection. The ecological context in which populations diverge, variation in life history traits, and male-male and female-female competition can change the dynamics of mate choice and affect the likelihood with which mating behavior polymorphisms successfully enter a population. Future research should therefore strive to integrate 'classic' (behavioural and quantitative genetic) and modern genomic approaches, combining insights in proximate mechanisms and ultimate evolutionary consequences to understand the fascinating diversity in sexual signals in the natural world.

ZUSAMMENFASSUNG

Fast eineinhalb Jahrhunderte nach Darwins zweitem evolutionsbiologischen Buch, *The descent of man, and selection in relation to sex*, fehlt uns immer noch ein umfassendes Wissen über die Rolle der sexuellen Selektion bei Artbildungprozessen. Eine wichtige Bedeutung auf die Synthese von sexueller Selektion und Artbildung hat der historische geografische und demografische Kontext, in dem divergierenden Populationen sich trennen. Auf der einen Seite erfolgt Artbildung oft in sympatrie oder parapatrie, das heißt Artbildung mit Genfluss, wo in der Regel Paarungssignale die am meisten divergenten Eigenschaften zwischen nah verwandten Arten sind. Das deutet auf eine wichtige Rolle der sexuelle Selektion bei der Entstehen von primären Barriere zum Genfluss hin. Auf der anderen Seite sagt die theoretische Forschung voraus, dass Artbildung mit Genfluss nur durch Effekte von sexueller Selektion nicht stattfinden kann.

Die vorliegende Dissertation verbindet Ergebnisse aus neuroethologischen, Verhaltens, quantitativ genetische und genomische Ansätze, um neue Erkenntnisse über die Rolle der sexuellen Selektion bei der Artbildungsprozessen zu erlangen. Nur wenige Organismen sind für solche interdisziplinäre Forschung ähnlich gut geeignet wie Feldgrillen (*Gryllus*). Viele Arten der artenreiche Gattung *Gryllus* zeigen geografisch überlappende Verbreitungen, unterscheiden sich im Paarungsverhalten, und sind größtenteils morphologisch und ökologisch kryptisch. Grillen verwenden akustische Signale für das Anlocken von Paarungspartner und sind durch die Variation im Paarungsverhalten reproduktiv isoliert. Das Verhalten und die neuronale Grundlage der akustischen Paarungssignale sind in Grillen gut untersucht. Das von Männchen erzeugte Signal ist ein multivariater Paarungsgesang, der zwischen den Arten in der Tonhöhe (Trägerfrequenz) sowie in zwei Zeitebenen des zeitlichen Musters variiert: (i) Pulse, die jeweils von einer einzigen Schließbewegung der Vorderflügel erzeugt werden, und (ii)Chirps, die aus Gruppen von Pulsen bestehen, die entweder relativ kurz oder regelmäßig durch Pausen unterbrochen sein können. Darüber hinauswerden sehr lange Abschnitte von Pulsen als Trills bezeichnet. Das zeitliche Muster eines akustischen Signals ist beschrieben durch die Wiederholungsrate der einzelnen Pulsen und Chirps (Pulsrate bzw. Chirprate), und die akustische Energie (Pulse bzw. Chirp Duty Cycle: der Dauer eines Pulses oder eines Chirps relativ zur Periode). Das umfassend beschriebene und hoch polymorphe Sexualverhalten zusammen mit der begrenzten ökologischer Divergenz unter geographisch überlappenden Arten, macht Grillen zu einem idealem System für die Untersuchung von sexueller Selektion und Artbildung mit Genfluss.

Die Ziele dieser Arbeit sind: (1) die Evolution des Grillengesangs und die Rolle der einzelnen Gesangsmerkmalen in der reproduktiven Isolation verstehen, sowie (2) den Beitrag von demographische Geschichte, Genfluss und (sexuelle) Selektion an genetischen Divergenz und Artbildung erläutern. Es wurde ein multivariater Kontext verwendet, um die Auswirkungen von Weibchenpräferenz auf männliche Gesangsmerkmale, sowie die aus der Partnerwahl resultierenden Effekte der Selektion auf den Gesang zu analysieren. In **Kapitel 2** haben wir die Verhaltensgrundlage der Gesangserkennung erforscht und die Form der weiblichen Präferenzfunktionen zwischen den verschiedenen männlichen Gesangsmerkmalen sowie zwischen Arten verglichen. Dies lieferte Information über die Rolle von bestimmten Gesangsmerkmalen in der reproduktiven Isolation zwischen den Arten und ermöglichte Einblicke in die mechanistische Grundlage der Partnerwahl. In **Kapitel 3** ist der genetischen Architektur des Gesangs und der Gesangspräferenz untersucht worden durch Erzeugung von Hybriden zwischen den sympatrischen Schwesterarten *G. rubens* und *G. texensis*. Dies bot Einblick in die zu erwarteten Antwort der Selektion und zeigte, ob sich Gesang und Gesangspräferenz schnell evolvieren können. In **Kapitel 4** wurde die multivariate Untersuchung der Gesangsvariation erweitert durch Quantifizierung der interspezifischen Divergenz, in der phänotypischen Varianz-Kovarianz (**P**) Matrix. Außerdem wurde die phänotypische Evolution des Paarungssignals der Grillen untersucht. Insbesondere sprechen wir hier die Frage an, ob die Form der Präferenzlandschaft neben Änderungen in den Mittelwerten der Merkmale auch die Orientierung der **P** Matrix, und damit die evolutionären Trajektorie des Gesangs, beeinflussen kann. In **Kapitel 5** verwenden wir RNA-seq Daten als molekularen Ansatz. Unter Verwendung der sympatrischen Schwesterarten *G. rubens* und *G. texensis* ist das genomweite Muster der funktionell relevante, genetische Variation, die demographische Geschichte und die Rolle des Genflusses untersucht worden. Zusätzlich wurde das Transkriptom auf durch Selektion beeinflusste genomische Fußabdrücke überprüft.

In drei Feldgrillenarten, die lange Trills erzeugen, haben wir festgestellt, dass die Form der weiblichen Präferenzfunktionen zwischen Arten konserviert ist (**Kapitel 2**). Die erste Ebene der Gesangserkennung ist manifestiert in einem peripheren Filter, der empfindlich ist für eine Reihe von Trägerfrequenzen und sich zwischen den Arten stark ähnelt. Die Auswertung des zeitlichen Musters folgt der Vorhersage von einem rezenten Modell für Gesangserkennung. Multivariate Präferenzprofile zeigten Selektivität für konspezifischen Pulsraten, das heißt auf der kurzen Zeitebene ist das Timing des Pulses ein wichtiger Indikator für die Partnerwahl. Auf der längeren Zeitebene zeigten Weibchen eine

Präferenz für höhe Chirp Duty Cycle, was darauf hinweist, dass das Evaluieren des zeitlichen Musters des Chirps unabhängig von dem Timing eines Chirps ist. Zusätzlich werden diese Ergebnisse bei drei Spezies, die kurze Chirps erzeugen, weitgehend bestätigt. Allerdings waren die Präferenzfunktionen für Chirp Duty Cycle, bei Arten die Chirps erzeugen, im Gegensatz zu den linearen Präferenzfunktionen bei Trillern, konkav (**Kapitel 6**). Gesangspräferenzen sind somit abhängig von wenigen Merkmalen und ähneln sich zwischen den Arten, während die Männchengesänge sich in allen Merkmalen unterschieden (**Kapitel 2 und 4**). Dadurch dass die unterschiedlichen Ebenen der Gesangerkennung durch unterschiedliche Präferenzunktionen charakterisiert sind, ist es möglich, dass multivariate Präferenzen gleichzeitlich verschiedene Indikationen für Paarungspartnerqualität aus den Gesangsmerkmalen extrahieren (das heißt Pulsrate als ein arbiträres "Erkennungsmerkmal" und Chirp Duty Cycle als Indikator für männliche Fitness).

Um die aus der weiblichen Präferenz resultierende evolutionären Auswirkungen der Selektion zu erforschen (siehe **Kapitel 2)**, untersuchten wir die genetische Architektur des Gesangs und der Gesangspräferenz. Auf der einen Seite weist die polygene genetische Architektur der Gesangsmerkmale und der Präferenz auf eine eher langsamere Divergenz hin, obwohl die sex-chromosomale Vererbung mehrerer wichtige Gesangsmerkmale höhere Evolutionsraten (**Kapitel 3**) zulässt. Auf der anderen Seite könnten die wichtigsten Merkmale für die Weibchenpräferenzen, das heißt Pulsrate und Chirp Duty Cycle, frei variieren ohne Begrenzungen in der Selektionsantwort, was auf hohe Evolvierbarheit hinweist. Interessanterweise wurde starke Kovarianz zwischen den Merkmalen die direkt unter sexuelle Selektion stehen und Merkmale, die nicht direkt von Weibchen gewählt werden beobachtet. Dies zeigt an, dass teilweise *indirekte* Selektion verantwortlich sein könnte für die markante multivariate Divergenz in männliche Gesänge, trotz begrenzter Divergenz der weiblichen Präferenzen (**Kapitel 3 und 4**). Ferner zeigte ein Vergleich der **P** Matrizen zwischen Arten die Chirps erzeugen (mit konkaven Präferenzfunktionen für Chirp Duty Cycle) und Arten die Trills erzeugen (mit linearen Präferenzfunktionen für Chirp Duty Cycle) das die Form der Präferenzfunktion, die Ausrichtung der Kovarianzen von Merkmale und die erwartete Selektionsantwort der männlichen Gesänge beeinflussen kann (**Kapitel 4**). Diese Ergebnisse unterstreichen das Potenzial der sexuellen Selektion den multivariaten Phänotyp das Seletkion zur Verfügung steht zu formen. Es zeigt auch, dass die Form der Selektionslandschaft auch unter nah verwandten Arten die evolutionäre Tendenz der Paarungssignale ändern kann. Dies lässt vermuten, dass die vorhergesagten Auswirkungen der sexuellen Selektion die Entstehung der reproduktiven Isolation begünstigt.

Abschließend wurde ein RNA-seq Ansatz gewählt, um mit Hilfe von funktionell relevante genomische SNP Daten und dem Muster der genetischen Divergenz, die möglichen Auswirkungen des Genflusses und des Fußabdruckes der Selektion in den sympatrischen Schwesterarten *G. Rubens* und *G. texensis* zu untersuchen (**Kapitel 5**). Mittels coalescent Simulationen fanden wir starke Hinweise für einen evolutionären Flaschenhals in der evolutionsbiologischen jüngeren Art (*G. rubens*). Außerdem zeigten diese Ergebnisse eine lange Genflussperiode zwischen den Arten, bis es zur Aufspaltung während der letzten pleistozänen Eiszeiten kam. Viele der Ausreißercontigs in den Analysen, die sich auf genetische Divergenz und den Anteil von Aminosäureänderungen in den Polymorphismen bezogen, hatten Übereinstimmungen mit *Drosophila melanogaster* Gene die experimentell nachgewiesenen Funktionen in neuromuskulärer Entwicklung, allgemeines und chemisches Paarungsverhalten haben. Diese Beobachtung deutet damit auf eine starke Rolle für sexuelle Selektion bei Divergenz mit Genfluss hin. Zusammen zeigen die Ergebnisse in den **Kapiteln 2, 3, 4 und 5** das Potenzial von sexueller Selektion bei der Entstehung und Aufrechterhaltung von reproduktiver Isolation zwischen Arten. Die vorliegende Arbeit verdeutlicht, dass für ein umfassendes Verständnis von sexueller Selektion bei der Artbildung interdisziplinäre Ansätze unerlässlich sind. In **Kapitel 6** diskutiere ich weitere Aspekten des Paarungsverhaltens bei Grillen, die die sexuelle Selektion beeinflussen können, aber in dieser Dissertation nicht untersucht worden sind. Der ökologischen Kontext, in dem Populationen divergieren, die Variation in Life History Merkmale sowie Männchen-Männchen bzw. Weibchen-Weibchen Interaktion können die Dynamik der Paarungswahl verändern und die Wahrscheinlichkeit das ein Paarungsverhaltenpolymorphismus in einer Population eintritt beeinflussen. Zukünftige Forschung sollte versuchen "klassische" (Verhaltens und quantitative genetische) Ansätze und moderne genomische Ansätze zu integrieren und Erkenntnisse in proximalen Mechanismen und ultimativen evolutionären Konsequenzen zu kombinieren, um die faszinierende Vielfalt der sexuellen Signale in der Natur zu verstehen.

SELBSTÄNDIGKEITSERKLÄRUNG

Hiermit versichere ich, dass ich die Arbeit selbständig und nur unter Zuhilfenahme der aufgeführten Hilfsmittel und Literaturquellen angefertigt habe.

Berlin, den ……

Thomas Blankers

ACKNOWLEDGEMENTS

I am thankful first and foremost to **Matthias Hennig**, my 'Doktorvater' and scientific beacon from day one onwards. **Matthias**, I am thankful for your endless patience when I was stubborn and dragged you through lengthy discussions about the acoustic taste of female crickets. I am thankful for the freedom you gave me to wander onto the scientific paths of my choosing and I am thankful that you nevertheless always remained vigilant that I did not lose focus. You have been a great advisor, mentor, and inspiration to me.

I also have the greatest appreciation for **Frieder Mayer**. **Frieder** you are not only the spiritual father of the GENART project which brought me to Berlin in the first place, you have also always been available for scientific advice, inquiries on the administrative idiosyncrasies in the museum, or just some colloquial chit-chat.

There is not enough space here to express my appreciation for my colleagues, friends, and family. **Emma**, I would need a whole additional thesis to describe my gratitude for your help and friendship. Your sharp criticism and wealth of scientifically apt comments on my writing and ideas have been crucial to every manuscript in this thesis and all that will follow. Additionally, your enthusiasm and sympathy have made every day in Berlin a happier one. I am sad not to be able to work with you on a daily basis anymore, but I have no doubt that our friendship and scientific relationship will overcome any geographic distance imaginable (including *e.g.* Sweden – Thomasland). You are an awesome and caring person, just about the right amount of crazy, and among the most scientifically talented people I know.

Dave, I highly appreciate your help with collecting and mailing crickets, your many intellectual contributions to the manuscripts in this thesis, and your pleasant company on a short road trip through Florida with a trunk full of crickets.

Jonas it was a pleasure sharing an office, a fair amount of very memorable moments [enter some exhaustive list of bars and BBQs here], and a passion for driving the motorcycle. I am also very thankful for your help with speaking, writing, and understanding the German language and its eccentricities.

Isabelle, you are an amazing technician that strongly contributed to the achievements within GENART and an extremely friendly and caring person. It was a pleasure working with you as well as being a guest the day you became Misses Waurick.

The AG VhPhys: **Michael**, for your intelligent remarks on my research ideas and manuscripts and for your keen sense of humour; **Flo**, you completed our office, our digital infrastructure, and most if not every BBQ party; **Eileen, Fleur, Jenny, Monika, Sarah, Stef** thank you for your pleasant scientific and social intermezzo's; **Regina,** the technician and Orthoptera Godmother of the AG; and of course **Bernhard Ronacher** for keeping this wild bunch together, pursuing excellent science and, as legend has it, spotting the tiniest inconsistency in a scatterplot from miles away.

The students that helped out with experiments – **Darja Hahn, Elisa Becker,** and **Vivienne Kremling** – and those that put in a lot of hard work to write excellent bachelor theses – **Katharina Lübke**, **Rafael Block**, and **Carolin Geelhaar.**

The members of the GENART project not mentioned before: **Camila Mazzoni, Christoph Dieterich, Frank Kirschbaum, Johannes Röhr, Linh Duy, Francesco Lamanna,** and **Ralph Tiedemann.** It was an honour and a pleasure partaking in the interdisciplinary meetings with you. Also highly appreciated are **Roger Butlin** and **Mike Ritchie** for providing comments during the starting and closing meetings of the GENART project.

The commission that will assemble to review and criticise my thesis, including those mentioned before: **Bernhard Ronacher, Frieder Mayer, Kerry Shaw, Matthias Hennig,** and **Mike Ritchie**.

The BeGenDiv for hosting excellent seminars on next-generation sequencing and in particular **Camila Mazzoni** and **Sibelle Vilaça**, who mixed time and patience in a big bowl of Brazilian friendliness to explain me the basics of approximate Bayesian demographic analyses.

The Leibniz Institut for funding and the Museum für Naturkunde and Humboldt Universität for (financial and other) support in field trips, conferences, etc.

[Hier bitte weiter lesen mit "New Kids" Akzent] Der AG Mayer: **Simon**, meine Zeit in Berlin könnte niemals so schön gewesen sein ohne deine Humor oder deine Panne12 Dachparty's. Ohne dich würden die Mensa und eigentlich ganz Berlin recht langweilig sein. **Nora**, durch deine grenzlose Lust auf Spaß (ein kleines geht ja immer…) und dein weitreichende und übertragbares Enthusiasmus war ich seit ich dir kennengelernt habe fast nie ‚Bocklos'. Ich danke dir auch für deinen etwa neurotische doch scharfe blick auf grafischen Inkonsistenz in meinen Abbildungen. **Christiane, Linus, Sonja, Steffen** es war immer schön euch zu treffen zum Mittagessen, zum Feiern oder einfach auf 'nen Bierchen nach der Arbeit.

The British Lions FC, lads you have treated me on the best football season of my life. There is no way I would have enjoyed each and every Saturday morning traveling half hungover to the most obscure corners of East and West Berlin (this does not include the Almas-arena of course) if it would have not been for your spirit, your "composure" [enter sarcastic laugh and a wink to coach Ben here], your enthusiasm and the legendary 2014/2015 season when we won it all.

Vrienden en familie, die me kwamen opzoeken in Berlijn of die er thuis voor me waren en me keer op keer vergaven voor langdurige afwezigheid. In het bijzonder bedank ik mijn lieve **papa, mama** en **broer** voor al hun liefde en onvoorwaardelijke steun als ik me kiplekker voelde of als ik er helemaal doorheen zat.

Lieve **Eline**, ik kan niet onder woorden brengen hoe blij ik ben dat wij op de valreep, een week voordat ik naar Berlijn ging, verliefd werden. Omdat je altijd in mijn hoofd zit, maar we vaak lang moesten wachten tot we elkaar weer zagen, waren de afgelopen drieëneenhalf jaar tegelijkertijd de mooiste en de zwaarste jaren uit mijn leven. Nu ik bij je woon, klopt alles. Ik hou van je.