The genetic architecture of sexually selected traits in two natural populations of *Drosophila montana*

Veltsos P, Gregson E, Morrissey B, Slate J, Hoikkala A, Butlin RK, & Ritchie MG

Centre for Biological Diversity, University of St Andrews, Dyers Brae House, KY169TH, St Andrews, UK

Animal & Plant Sciences, University of Sheffield, Alfred Denny Building, Sheffield, S10 2TN, UK

Department of Biological and Environmental Science, PO Box 35, 40014 University of Jyväskylä, Finland

Sven Lovén Centre – Tjärnö, University of Gothenburg, SE 452 96 Strömstad, Sweden

Corresponding author: Paris Veltsos parisveltsos@gmail.com 0041 21 692 5615

Running Title: Intra-specific song and CHC QTL in *Drosophila montana*.

Word count: 5027

Keywords: sexual selection, QTL, *Drosophila montana*, courtship song
Abstract
We investigated the genetic architecture of courtship song and cuticular hydrocarbon traits in two phylogenetically distinct populations of *Drosophila montana*. In order to study natural variation in these two important traits, we analysed within-population crosses among individuals sampled from the wild. Hence, the genetic variation analysed should represent that available for natural and sexual selection to act upon. In contrast to previous between-population crosses in this species, no major QTLs were detected, perhaps because the between-population QTL were due to fixed differences between the populations. Partitioning the trait variation to chromosomes suggested a broadly polygenic genetic architecture of within population variation, although some chromosomes explained more variation in one population compared to the other. Studies of natural variation provide an important contrast to crosses between species or divergent lines, but our analysis highlights recent concerns that segregating variation within populations for important quantitative ecological traits may largely consist of small effect alleles, difficult to detect with studies of moderate power.
Introduction

Quantitative trait locus (QTL) studies are typically performed by identifying lines or species that differ in a trait of interest, crossing them and correlating the phenotypic variation in the resulting offspring with genetic variation, to pinpoint genomic regions associated with the trait of interest. If the trait of interest is thought to be important to speciation, the initial cross is often made between inbred lines from different species or differentially adapted populations. However, using highly differentiated populations cannot indicate whether the same differences were involved in the initial isolation of the populations, because current differences may have arisen since the populations became isolated. In fact, it seems that the types of mutations providing an initial response to selection within species may differ from those eventually fixed between species (Stern and Orgogozo, 2008; 2009). QTL studies should, therefore, be performed within and between populations representing multiple divergence times.

More generally, the range of trait variation captured by the initial cross will influence both the ability to detect QTL and possibly the nature and the effect size of QTLs detected. Choosing very different individuals for the initial cross may identify large effect QTLs, which may not be segregating in wild populations. The effectiveness of the candidate gene approach is a related issue. Candidate genes are identified by the large effects and, if their function is evolutionarily conserved, they should be associated with variation in the same function in different species. Such loci have been found for a variety of traits, including behaviour patterns (Fitzpatrick et al., 2005; Martin and Orgogozo, 2013). In other cases a polygenic genetic architecture is reported, for example in recent QTL studies of ecologically important traits (Davies et al., 2011; Rockman, 2012; Travisano and Shaw, 2013). This has led some to question the value of looking for large effect loci for understanding ecological adaptation within species. Even where large effect loci are segregating, they may be at very low frequency leading to only a small contribution to the population level heritability of such traits (Scoville et al., 2011). It is therefore important to compare the genetic architecture of traits in different types of crosses, such as within vs between populations or lab- vs field- reared, and assess how often similar or different genetic effects are
Studies of the quantitative genetics of ecologically important traits in the field are proceeding well for some vertebrates with extensive pedigree data (Kruuk and Hill, 2008; Robinson et al., 2013) and some important segregating loci have been identified (Johnston et al., 2011) but there are very few studies for quantitative traits in species such as Drosophila, that are the source of much of the information on species differences. The circum-arctic species Drosophila montana has proved to be a good species in which to study adaptive traits, including cold tolerance, developmental time, juvenile body weight and reproductive diapause (Vesala and Hoikkala, 2011; Salminen et al., 2012; Parker et al., 2015) in the wild. Phylogeographic analysis of the species based on mtDNA and microsatellites suggests two clades, a North-American and European, that were isolated about 0.5 Myr ago (Mirol et al., 2007). North-American D. montana populations have been classified as Standard, Giant and Alaskan-Canadian based on their geographical origin, size and chromosome structure (Throckmorton, 1982), and Eurasian populations contain several unique inversions (Morales-Hojas et al., 2007). There is evidence for prezygotic isolation (assortative mating), post-mating, pre-zygotic isolation (successful sperm transfer but failure to fertilise eggs), and post-mating post-zygotic isolation (progeny production) between some populations (Jennings et al., 2011; Jennings, Snook, et al., 2014). The main reasons to classify these populations as belonging to the same species are that the strength of these different barriers to gene flow depends on the population pairs being crossed and that pre-mating isolation is only evident when females can choose between a male from their own population and another population. Population differences therefore fall into the rarely studied early part of the speciation continuum (Safran et al., 2013).

There have been extensive studies of within and between species phenotypic variation in the sexual behaviour of the virilis group to which D. montana, belongs. Male courtship song is essential for mating in D. montana (Liimatainen et al., 1992), has a species-specific component, inter-pulse interval (IPI) (Liimatainen and Hoikkala, 1998; Saarikettu, Liimatainen, and Hoikkala, 2005a), and crosses between closely related species have indicated X-linkage for genes influencing some song traits (Hoikkala et al., 2000;
Päällysaho et al., 2003). Courtship song is under sexual selection and seems to be an honest indicator of male quality: one song component, intra-pulse frequency (FRE), is condition-dependent, correlated with offspring survival and under strong and contrasting viability and sexual selection in the field (Aspi and Hoikkala, 1995; Hoikkala et al., 1998). Various song components influence mating latency and courtship duration (Veltos et al., 2012), and populations differ in both song and preference (Klappert et al., 2007; Ritchie et al., 2007). Recent studies of cuticular hydrocarbons (CHCs) in D. montana show that they influence mating decisions independently of song and also differ between populations (Veltos et al., 2012; Jennings, Etges, et al., 2014).

We previously analysed the influence of song and CHC variation in male mating success in D. montana through no-choice experiments where females were presented with a male from the same population (Oulanka, Finland, or Vancouver, Canada) and the mating occurrence, mating latency and female rejection song production were correlated with various components of male courtship song and CHCs (Veltos et al., 2012). In this paper, we complete quantitative genetic and QTL analyses of the same traits by investigating and comparing their genetic architecture. The same two independent populations, representing distinct phylogeographic clades (Oulanka and Vancouver) were used. The phenotypic data of both populations were jointly analysed with principal component analysis, so that the populations can be compared. We compare the within-population QTL analysis of song variation with previous work that identified QTL influencing courtship song by employing between-population outbred crosses of D. montana (from Oulanka, Finland, and Colorado, N. America). The detected QTL were on two chromosomes (X and 2) and one candidate gene was potentially associated with each QTL (per and fru, respectively) (Schafer et al., 2010; Lagisz et al., 2012), although it has not been confirmed if these or other genes under the peaks are responsible for the phenotypic effects.

Here we investigate whether within-population variation in the same traits involves the same genomic regions as the between-population analysis. We established two independent pedigrees from wild collected females, developed SNP markers by transcriptome sequencing and scored individuals from the pedigrees for SNPs and song and CHC traits. Most traits exhibited
moderate heritability. Our QTL and chromosome partitioning results reveal no large effects in the regions identified by the between-population crosses, and suggest differences in genetic architecture between the populations in terms of the proportion of additive genetic variation explained by each chromosome. Overall, they highlight the difficulty of gaining adequate power to detect loci of small effect in studies on the genetic architecture of wild populations.
**Materials and Methods**

**Sampling and phenotyping**

Field sampling, cross design and phenotyping have been described previously (Veltsos et al., 2012). Briefly, we constructed and collected phenotypic and genotypic information on 2 or 3 generation pedigrees of about 500 individuals each from the offspring of wild-caught females from Vancouver, Canada, and Oulanka, Finland. There were 30 isofemale lines from Vancouver, and 42 from Oulanka. Each line may have had more than one father, as the frequency of multiple mating in the wild has been estimated as 1.19±0.31 (Aspi, 1992). Song recording took place while setting up the crosses of the next generation. We measured four song parameters: intra-pulse frequency (FRE), inter-pulse interval (IPI), pulse number (PN) and cycle number per pulse (CN). We analysed these traits and the principal components of song variation (using data from both populations so that the PCs are directly comparable), including recording temperature as a covariate because it has a strong effect on song (Ritchie et al., 2001).

CHCs were extracted from whole individuals of both sexes, after the females had laid sufficient eggs to provide individuals for the next pedigree generation (>20 eggs). Principal components of CHC data were obtained from both sexes simultaneously because they are only moderately sexually dimorphic in *D. montana*, with no sex-specific compounds (Veltsos et al., 2012; Jennings, Etges, et al., 2014). The PCA analysis was on relative proportions of log transformed CHCs (Rundle et al. 2009) and used 18 compounds in total. Principal component analysis was performed on data from both populations simultaneously, to make them directly comparable in QTL analysis. A full description of the PCA analysis, characterisation of population differences in CHCs as well as their association with measures of fitness in the lab are presented in Veltsos et al. (2012).

**Transcriptome sequencing**

RNA was extracted from 120 heads of 2-3 week old virgin adults of both sexes from offspring of the last generation of each pedigree, using a Qiagen RNA extraction kit (Qiagen, Hilden, Germany), following the manufacturer’s instructions and including on-column DNA digestion. RNA samples were
processed by Genepool in Edinburgh (Vancouver flies) and the Centre for Genomics Research in Liverpool (Oulanka flies). cDNA libraries were constructed without normalisation, and were sequenced on half a 454 pyrosequencing plate each (Roche, FLX).

The sequencing reads were filtered to remove low quality sequence (threshold of 20) and trimmed to remove adaptor, primer sequence and poly-A tails using SeqMan NGen version 1.2 (DNASTAR Madison, WI, USA). 348,267/417,467 and 804,456/1,211,678 assembled/total reads were available for Vancouver and Oulanka respectively. Significantly more reads were available for Oulanka because two sequencing runs were combined (the first run yielded only a few reads). An initial assembly was made by combining all reads from both populations. The resulting master contigs were used to guide two assemblies, one for each transcriptome. This ensured easy comparison of SNP locations from each population. All assemblies were performed with SeqMan NGen version 1.2 (DNASTAR Madison, WI, USA).

SNP marker development

We aimed to obtain SNP markers that segregated in both populations. The best 384 SNPs to score were identified through custom scripts (on Dryad). Briefly, SNPs were called in each population using SeqMan NGen version 1.2 (DNASTAR Madison, WI, USA), and were categorised as being unique to a population, on the same contig, or identical between populations. Criteria for shortlisting SNPs were good coverage (at least 8 reads in total and 3 reads or 20% for the minor allele), long distance from other SNPs (at least 25 bp), scoring primer location outside of intron-exon boundary and, when possible, presence in both populations. Information on the first blastp hit against the non-redundant (nr) database of each contig was obtained using Blast2GO (Conesa et al., 2005) and used along with blast searches on a list of candidate genes to add SNPs that fell within candidate genes using less restrictive criteria than other SNPs, for example accepting SNPs unique to one population.

Blastclust (http://www.csc.fi/english/research/sciences/bioscience/programs/blast/blastclust) was also used to remove contigs aligning to potentially highly repeated sequences, which could cause SNP scoring problems. For more information on SNP development see Data Archiving.
Genotyping

DNA extractions were performed by squashing individual flies in 50 µl Squishing buffer (10mM TrisHCl pH 8.2, 1mM EDTA, 25mM NaCl, 200mg/ml proteinase K), incubating at 50°C for 2 h and boiling for 2 min to deactivate the proteinase K. DNA clean-up was performed with standard ethanol precipitation (Sambrook and Russell, 2001). The 384 SNPs were scored on the Illumina BeadXpress platform (Illumina Inc.) and analysed with GenomeStudio 2010.1 (Genotypic module 1.7.4, Illumina, San Diego, CA). Individuals were removed if they were associated with genotyping errors (non-Mendelian inheritance in multiple markers), as identified by GenotypeChecker (Paterson and Law, 2011). Genotyping errors cause too many spurious recombination events and inflate genetic maps. Individuals were also removed if they did not have relatives or were missing phenotypic data. After cleaning, 214 and 334 individuals (Vancouver) and 171 and 340 individuals (Oulanka) were included in the QTL analyses for song and CHC, respectively. SNPs were removed from further analysis if they resulted in any parent-offspring mismatches, or were obvious outliers in call rate (<0.80). In total 127 and 130 SNPs were considered of sufficient quality for genetic map construction for Vancouver and Oulanka, respectively, and 79 of them were common between the populations. All data used in the QTL analysis have been submitted to Dryad.

Genetic map construction

Linkage mapping was performed as in (Slate, 2008) using CRI-MAP v2.504a (Green et al., 1990; 2009) (http://www.animalgenome.org/tools/share/crimap/). The pedigrees were first split into subfamilies using the CRIGEN command, and linkage groups were initially defined by the AUTOGROUP command. Linked markers were identified using the TWOPOINT command, with all pairs of markers producing LOD scores in excess of 3.0 being regarded as linked. For each linkage group, the most parsimonious marker order was determined using the BUILD, FLIPS, FLIPS3 and FLIPS5 commands. The linkage groups were assigned to the five D. montana chromosomes by blasting the sequence of their markers against the D. virilis scaffolds which have been mapped to chromosomes in FlyBase (Altschul et al., 1990; St Pierre et al., 2014). When the
linkage maps had a different marker order between the populations (chromosomes 2 and 4), the difference in log likelihood of the collinear marker order and the parsimonious marker order was obtained for the shared markers using CHROMPIC and was always >3 (40.06 and 38.13 respectively), providing support that the markers are not collinear. The genetic maps were plotted using MAPCHART v2.2 (Voorrips, 2002).

Marker 12716_321_id was placed by CRI_MAP 100 cM away from the next marker at one end of the Oulanka X chromosome. We retained the marker, by reducing the distance to 49.5 cM, for three reasons: It mapped without problems on the Vancouver X, it has a homologue on the D. virilis X, and there was low marker coverage of the X.

Chromosome partitioning analysis

Genome-wide relatedness matrices (GWRMs), weighted for expected relatedness from pedigree information, were constructed using the methodology of (Robinson et al., 2013): For each autosome, pairs of GWRM matrices were made for its markers and those of the remaining autosomes. Each pair of GWRMs was used, in turn, to partition the variance explained by an autosome. To determine the additive genetic variation explained by an autosome, we compared the likelihood of a model with both GWRMs and a model with only the GWRM of the remaining autosomes, for each GWRM pair. A separate GWRM was constructed for the X chromosome and was only used to estimate the variance explained by the X, in a similar fashion. The models included temperature for song and sex for CHC data as fixed effects. Model fitting was performed in ASReml version 3 (Gilmour et al., 2009). Full details of the methodology are available in (Santure et al., 2013).

If a trait is polygenic, the variance explained by each chromosome should be proportional to its size. We considered three proxies of gene number per chromosome: the physical length of the D. virilis homologous chromosome and the genetic map length of the chromosome in either D. montana or D. virilis. We consider the D. virilis chromosome physical length to be the best proxy because loci on the X chromosome were underrepresented in our genetic maps, making them shorter than expected. Possible reasons for the low SNP density on the X include the fact that males, which provided about half the RNA used in sequencing, are hemizygous (rare SNPs would be less
likely to be retained by the SNP filtering pipeline). The *D. virilis* chromosome lengths were obtained from Flybase (Altschul *et al.*, 1990; St Pierre *et al.*, 2014), by summing up the lengths of all the *D. virilis* scaffolds, that have been mapped to each chromosome.

**QTL - animal model analysis**

Identity by descent (IBD) matrices were constructed at 10 cM intervals for all individuals of each pedigree, using all SNPs on the relevant chromosome with an MCMC approach implemented in LOKI v2.4.5 (Heath *et al.*, 1997; Heath, 1997), and a parameter setting of 10,000 iterations for each position. The X chromosome is hemizygous in males, but LOKI requires two alleles for all markers. We manually made the male X marker information compatible with the analysis by adding a novel allele for each X male genotype (i.e. all males carried one copy of a dummy allele that was absent in females). The phenotypic variance was partitioned into fixed and random effects in a restricted maximum-likelihood (REML) framework using ASReml version 3 (Gilmour *et al.*, 2009), as in Slate (2005). Heritability ($h^2$) was estimated as the ratio of additive genetic to total phenotypic variance from a polygenic model. Potential QTL effects were estimated as in Santure *et al.* (2013). Briefly, a LOD score was calculated by subtracting the log-likelihood of a polygenic model from the log-likelihood of a polygenic plus QTL model, the latter also including the IBD matrix, for each genomic interval. Significance thresholds were adjusted for multiple testing based on genome size, following Lander and Kruglyak (1995). Two genome-wide linkage thresholds were calculated for each population (File S1): the suggestive linkage threshold is expected once per genome scan, while the significant linkage threshold has a 1/20 false positive rate, i.e. it represents a genome-wide significance of $P < 0.05$ (Lander and Kruglyak, 1995; Nyholt, 2000).

We could not directly compare the two populations within a single analysis because not all of their markers were shared, and because the relative positions of the markers differed. Instead we indirectly compared the populations by comparing results from the chromosome partitioning analysis, which is not affected by marker order.
We performed simulations using custom scripts to estimate the probability of detecting, within the Vancouver pedigree, a QTL of similar magnitude to the one detected in the inter-population study (Lagisz et al., 2012). We used sex-averaged map distances and simulated a trait with heritability of 0.30, half the variance of which was caused by a single QTL in the middle of chromosome 2 ($V_e = 0.15$, $V_{qtl} = 0.15$, $V_{residual} = 0.70$). The simulations were run for 100 replicates, assumed all individuals were phenotyped and employed the same suggestive and significant linkage thresholds as the experimental QTL analysis.
Results

Phenotypic variation

Phenotypic variation for the two pedigrees has been described before and the populations differ in all traits (Veltsos et al., 2012). Means and standard deviations for all phenotyped traits are presented in Table 1.

Genetic maps

The total length of the map from the Vancouver pedigree was 789.8 cM, and that from the Oulanka pedigree 689.8 cM. Comparison of these maps clearly shows a different marker order for chromosomes 2 and 4 (Figure 1, File S2). Chromosomal rearrangements are known to have occurred within D. montana and we therefore believe that extensive changes in gene order are not unlikely. A population on the Western Coast of North America (around Vancouver) has at least one inversion on the 2nd chromosome and 10 inversions on the 4th chromosome that are not found in Finland, while the Finnish population has 3 inversions on the 2nd chromosome that are not found in the Vancouver region (Hsu, 1952; Patterson and Stone, 1952; Moorhead, 1954; Morales-Hojas et al., 2007). Information on the markers such as heterozygosities, their top blast hit, potential functions and sequence of the relevant contig is provided in File S3.

Heritability estimates

Heritability and standard error estimates based on the relatedness matrices calculated from all SNPs are shown in Table 1. The estimates are broadly consistent between the populations, suggesting some similarities between the traits (correlation coefficient between heritability estimates for PCs = 0.83).

Power analysis

Simulations suggest there was only moderate power to detect QTLs of an effect size similar to those detected in between-population crosses. Out of 100 simulations, 55 detected the simulated QTL at the suggestive threshold and only 13 detected it at the genome-wide significance threshold (File S4a). When detected, the Beavis Effect (Göring et al., 2001) was observed, i.e. the QTL
effect size was usually overestimated (mean effect size was 0.25 instead of simulated 0.15 proportion of total variance). When the QTL location was detected, it was accurately mapped (File S4b).

QTL results

Based on the total genetic map length, the suggestive and significant LOD scores for Vancouver were 1.12 and 2.62 respectively, and for Oulanka 1.05 and 2.56. We did not detect genome-wide significant results in the QTL analysis but observed suggestive linkage once, on chromosome 5, for CN in Oulanka (Figure 2). The result had a slightly higher LOD score when song principal components were used, (song_PC3 which is strongly influenced by CN (Veltzos et al., 2012)), but was otherwise qualitatively similar (File S6). Interestingly the song characters mostly influencing song_pc3 (FRE, IPI and CN (Veltzos et al., 2012)) had lower LRT (log likelihood ratio test) values individually than the LRT for the principal component, suggesting that multivariate analysis may increase detection power, presumably because a principal component axis reduces the noise of individual measurements.

Genetic variation partitioning

There was no relationship between percent variance explained (PVE) and chromosome length, regardless of the proxy of chromosome length used (Table 2, File S5). It is difficult to detect such a correlation because of the small number of data points as there are only 5 chromosomes in D. montana. In addition it is difficult to detect differences in PVE between the chromosomes because they have similar sizes, which may make the between chromosome variance small, relative to the within chromosome error in estimating variance. There were 18 cases where a single chromosome explained non-zero variance (Table 3). No formal test of significance was performed because the null hypothesis of no variation explained by any chromosome is false for polygenic heritable traits.

Population comparison

We compared the genetic architecture between the populations by plotting the percent variance explained by each chromosome for each trait, in the two populations (Files S7, S9). While the values were often similar, there
were occasions where a chromosome clearly explained more variance in one population than in the other population (e.g. for song PC1).
Discussion

We have attempted to identify the architecture of genetic variation available for natural and sexual selection to act upon, within natural populations, by performing heritability and QTL analysis on song and CHC phenotypes using pedigrees established from two populations of *D. montana*. Although we found significant heritability, we did not find evidence of large effect QTL for any trait. Our results are broadly suggestive of polygenic determination for intrapopulation phenotypic variation, though some individual chromosomes explained more phenotypic variation than others, in a population-specific manner.

Song and CHCs contribute to mating success in *D. montana*, and CHCs are probably also involved in ecological adaptation (Veltsos et al., 2012; Jennings, Etges, et al., 2014). Heritabilities were sometimes low (Table 1), which may well have mitigated against our ability to detect QTL (the only suggestive QTL found was for a song trait with one of the highest heritabilities, CN in Oulanka). Nevertheless, our heritability estimates for song were usually larger than previous estimates in *D. montana*, though sometimes they were not (Aspi and Hoikkala, 1993; Suvanto et al., 1999). The estimates reported here are probably more robust because of the larger sample size, the estimation of relatedness from pedigrees and the reduced measurement noise achieved by using different song analysis methods. The earlier estimation of heritabilities from overwintered flies and the finding that they increase compared to non-overwintered flies (Suvanto et al., 1999) remains important because sexual selection on courtship song often acts only after the flies overwinter as adults.

Simulations suggested that our power to detect QTL was only moderate. The simulated trait effect size was based on a QTL detected for song variation in crosses between North American and Finnish *D. montana* (Colorado and Oulanka, i.e. a different population from North America) (Mirol et al., 2007; Schafer et al., 2010; Lagisz et al., 2012). We failed to detect QTL for the same traits and genomic positions as the between-population study. These were for IPI and PN on the X chromosome, and FRE, CN and IPI on chromosome 2 (Lagisz et al., 2012). We then attempted to implicate the same chromosomes using the chromosome partitioning analysis. Of the above traits, a large
proportion of variance (31.8 ± 17.3 \%) was only found to be predicted for IPI on chromosome 2 in Oulanka (Files S5, S7). One possible explanation for the disparity between the studies is that large-effect alleles that differ between clades are not making a major contribution to genetic variation within populations. This supports recent studies suggesting that the genetic architecture of intra-population and inter-clade variation may differ for *Drosophila* song, (Gleason, 2005; Arbuthnott, 2009) and for sexually selected traits generally (Chenoweth and McGuigan, 2010). Similarly, some between-population QTL may be due to Colorado-specific alleles of large effect, which were not sampled in this study. An alternative explanation for the disparity between the within- and between- population studies may be the greater range in phenotypes amongst the crosses between two populations, compared to within population variation. Colorado flies are different from those from Vancouver, in both their songs and pheromones (Klappert *et al.*, 2007; Routtu *et al.*, 2007; Jennings, Etges, *et al.*, 2014), despite coming from the same phylogeographic clade.

A third possibility is that QTL of large effect are difficult to detect using intra-population analysis. When detected in crosses between lines, QTL may be caused by variants of large effect that are at low frequency within the population. Variants of large effect under selection would rapidly fix within a population during adaptation to new environments and are not likely to segregate within populations (Scoville *et al.*, 2011). Consequently they may be readily associated with large QTL in between-population crosses that do not explain much phenotypic variance at the population level.

A fourth possibility is that the between-population QTL seem large because fixed chromosomal rearrangements between the populations caused a reduction in recombination in some genomic regions, effectively combining the effects of many loci of small effect, and causing them to segregate as a single locus. Loci implicated in repeated evolution of adaptive traits are often associated with inversions (Martin and Orgogozo, 2013). When the genes captured by an inversion are involved in reproductive isolation, the inversion may become important in maintaining isolation in the face of gene flow (Noor *et al.*, 2001; Kirkpatrick and Barton, 2006). The IPI QTL in an inversion on chromosome 2 in the between-population study (Lagisz *et al.*, 2012) is
compatible with both ideas of capturing multiple genes and being associated
with reproductive isolation, since IPI is under sexual selection and differs
between species (Saarikettu, Liimatainen, and Hoikkala, 2005b; Veltos et al.,
2012). Within-population crosses would not be affected by such inversions,
which could explain the failure to detect QTL.

The QTL in sexually selected traits in within- and between- population
studies are usually different (Arbuthnott, 2009; Chenoweth and McGuigan,
2010). One exception, for a pre-mating signal, involves a pheromone QTL in
the moth species Heliothis subflexa and H. virescens (Groot et al., 2013). Few
studies have compared the magnitude of effect sizes within and between
populations in addition to assessing if the same QTLs are implicated, and it
may be that consistent genetic effects are associated with phenotypes where
large effect loci are more likely to be found, such as pheromone
polymorphisms or genes influencing melanism (Wittkopp et al., 2009).

Comparison between populations

One common method to demonstrate polygenic variation is through a
correlation of percent variance explained (PVE) with chromosome size. The
relatively low number of similarly-sized chromosomes makes this an
ineffective test for Drosophila, but most chromosomes explained some
variation which is compatible with polygenic genetic architecture. There were
cases where one chromosome explained more variation than others. This
sometimes happened in a population-specific manner (File S7), providing
evidence of a subtly different genetic architecture between the Oulanka and
Vancouver populations.

The X chromosome had the lowest number of markers in our study even
though it is the largest chromosome. The low X coverage would be
particularly unfortunate if there is a large X effect, that is if the X is
disproportionately influencing reproductive isolation (Presgraves, 2008) or
genes under sexual selection (Qvarnstrom and Bailey, 2009; Dean and Mank,
2014), which may have further hindered our ability to detect segregating
genetic variation for these traits.

Conclusion

In performing this study we have developed useful resources, including
a comparative transcriptomic data set and a set of SNP markers and 
associated genetic maps for this species. Like many studies of wild 
populations, our study suffers from limited statistical power. Investigating 
the genetic architecture of within-population variation probably requires 
sample sizes in the order of a few thousand, in order to ensure sufficient 
statistical power (Rockman, 2012; Slate, 2013). Despite significant heritabilities 
of the traits we found limited evidence of common genetic architecture in 
within-population phenotypic variation and interspecific differences, 
extending the trend found in other comparisons (Arbuthnott, 2009; 
Chenoweth and McGuigan, 2010). This may suggest that large effect alleles 
fixed between populations are not major contributors to variation within 
populations. It is important to study within-population variation despite 
such difficulties because it is the variation upon which natural and sexual 
selection act, making it the most relevant variation for immediate responses to 
selection (Prokop et al., 2012). Concentrating on QTL mapping populations 
from between-species crosses can reveal genetic mechanisms of species 
differences, but such genes may have diverged in frequency or become fixed 
after speciation. Understanding sources of genetic variation across the 
speciation continuum remains a major challenge of evolutionary genetics (The 
Marie Curie SPECIATION Network, 2012).

Acknowledgments

The work was supported by the National Environment Research Council 
(grant NE/E015255/1 to M.G.R and R.K.B.) and the Academy of Finland 
(project 132619 to A.H.). We would like to thank Juan Galindo and Anna 
Santure for bioinformatic and statistical help. We are grateful to Dario Beraldi 
for providing a script to streamline the pedigree analysis and parse files 
between PedigreeChecker, CRIMAP and ASREML, and Matthew Robinson 
for providing scripts allowing the chromosome partitioning analysis.

Data archiving

Phenotype, genotype and pedigree data as well as raw reads and the 
assembled transcriptome, are available from the Dryad Digital Repository 
Figures

Figure 1: Comparison of genetic maps from the two populations. The markers are represented by the name of the top *D. virilis* blast hit of the relevant contig. Chromosome 4 is clearly not collinear between the two populations. Candidate genes are indicated by shading: Dvir/GJ14444-Ebony, Dvir/GJ23648-Slowpoke, Dvir/Desat1-DsatF and Dvir/GJCy-tb5-r-CytB5.
**Figure 2:** QTL maps of song and CHC in the two populations. The dotted and uninterrupted horizontal lines indicate suggestive and significant linkage, respectively. Chromosomes X and 2, which showed QTL for song in the between-population study, have no QTL identified.
Table 1: Trait values (genotyped individuals from (Veltos et al., 2012)) and heritability estimates from polygenic models of song and CHC traits.

<table>
<thead>
<tr>
<th>Character</th>
<th>value ± sd</th>
<th>h² ± se</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRE (Hz)</td>
<td>253 ± 35</td>
<td>0.08 ± 0.11</td>
</tr>
<tr>
<td>IPI (msec)</td>
<td>40.18 ± 6.38</td>
<td>0.2 ± 0.14</td>
</tr>
<tr>
<td>PN</td>
<td>4.80 ± 0.95</td>
<td>0.14 ± 0.12</td>
</tr>
<tr>
<td>CN</td>
<td>9.40 ± 1.25</td>
<td>0.52 ± 0.18</td>
</tr>
<tr>
<td>Song_pc1</td>
<td>0.39 ± 1.16</td>
<td>0.14 ± 0.13</td>
</tr>
<tr>
<td>Song_pc2</td>
<td>-0.02 ± 1.05</td>
<td>0.28 ± 0.14</td>
</tr>
<tr>
<td>Song_pc3</td>
<td>0.16 ± 0.93</td>
<td>0.28 ± 0.14</td>
</tr>
<tr>
<td>Song_pc4</td>
<td>-0.03 ± 0.46</td>
<td>0.28 ± 0.14</td>
</tr>
<tr>
<td>CHC_MF_pc1</td>
<td>1.19 ± 2.19</td>
<td>0.42 ± 0.13</td>
</tr>
<tr>
<td>CHC_MF_pc2</td>
<td>-0.90 ± 0.93</td>
<td>0.56 ± 0.13</td>
</tr>
<tr>
<td>CHC_MF_pc3</td>
<td>-1.19 ± 1.08</td>
<td>0.35 ± 0.13</td>
</tr>
<tr>
<td>CHC_MF_pc4</td>
<td>0.38 ± 0.78</td>
<td>0.28 ± 0.13</td>
</tr>
</tbody>
</table>

Table 2: Correlation between PVE and D. virilis chromosome length. Each chromosome is represented by the combined PVE values from the two pedigrees, since they did not differ significantly (data not shown). P-values of negative correlations are labelled as ‘NS’ because only positive correlations are expected.

<table>
<thead>
<tr>
<th>character</th>
<th>correlation</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRE</td>
<td>-0.52</td>
<td>NS</td>
</tr>
<tr>
<td>IPI</td>
<td>0.47</td>
<td>0.17</td>
</tr>
<tr>
<td>PN</td>
<td>-0.04</td>
<td>NS</td>
</tr>
<tr>
<td>CN</td>
<td>-0.50</td>
<td>NS</td>
</tr>
<tr>
<td>song_pc1</td>
<td>0.51</td>
<td>0.13</td>
</tr>
<tr>
<td>song_pc2</td>
<td>-0.30</td>
<td>NS</td>
</tr>
<tr>
<td>song_pc3</td>
<td>-0.41</td>
<td>NS</td>
</tr>
<tr>
<td>song_pc4</td>
<td>-0.53</td>
<td>NS</td>
</tr>
<tr>
<td>CHC_pc1</td>
<td>-0.19</td>
<td>0.60</td>
</tr>
<tr>
<td>CHC_pc2</td>
<td>0.28</td>
<td>0.43</td>
</tr>
<tr>
<td>CHC_pc3</td>
<td>-0.52</td>
<td>NS</td>
</tr>
<tr>
<td>CHC_pc4</td>
<td>-0.3</td>
<td>NS</td>
</tr>
</tbody>
</table>

Table 3: Traits where a chromosome explains non-zero variance (percent variance explained - SE > 0). In all cases only one of the populations explains > 0 PVE.

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Character</th>
<th>PVE ± SE</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>song_pc1</td>
<td>0.30 ± 0.17</td>
<td>O</td>
</tr>
<tr>
<td>2</td>
<td>IPI</td>
<td>0.00 ± 0.00</td>
<td>V</td>
</tr>
<tr>
<td>2</td>
<td>CHC_pc2</td>
<td>0.10 ± 0.09</td>
<td>O</td>
</tr>
<tr>
<td>3</td>
<td>song_pc3</td>
<td>0.35 ± 0.16</td>
<td>O</td>
</tr>
<tr>
<td>3</td>
<td>PN</td>
<td>0.10 ± 0.12</td>
<td>O</td>
</tr>
<tr>
<td>3</td>
<td>FRE</td>
<td>0.00 ± 0.00</td>
<td>O</td>
</tr>
<tr>
<td>3</td>
<td>CN</td>
<td>0.19 ± 0.17</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>------------</td>
<td>--------</td>
<td>-----</td>
</tr>
<tr>
<td>3</td>
<td>CHC_pc3</td>
<td>0.03 ± 0.00</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>CN</td>
<td>0.22 ± 0.16</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>CHC_pc3</td>
<td>0.00 ± 0.00</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>CHC_pc1</td>
<td>0.00 ± 0.00</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>IPI</td>
<td>0.15 ± 0.13</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>song_pc2</td>
<td>0.15 ± 0.14</td>
<td>O</td>
</tr>
<tr>
<td>5</td>
<td>song_pc2</td>
<td>0.00 ± 0.00</td>
<td>O</td>
</tr>
<tr>
<td>5</td>
<td>CHC_pc3</td>
<td>0.07 ± 0.00</td>
<td>O</td>
</tr>
<tr>
<td>5</td>
<td>CHC_pc2</td>
<td>0.00 ± 0.00</td>
<td>O</td>
</tr>
<tr>
<td>5</td>
<td>IPI</td>
<td>0.07 ± 0.12</td>
<td>O</td>
</tr>
<tr>
<td>5</td>
<td>song_pc1</td>
<td>0.07 ± 0.12</td>
<td>O</td>
</tr>
</tbody>
</table>

**Supplementary files**

**File S1**: QTL significance thresholds per chromosome and genome (significant and suggestive linkage).

**File S2**: Visualisation of the *D. virilis* scaffold associated with each marker. A different colour was used for different scaffolds in each chromosome.

**File S3**: Marker information. The columns show the SNP name, the name of the contig it is associated with, the most likely *D. virilis* homologue of the contig, a description of the top blast hit based on Blast2GO annotation, the *D. virilis* scaffold it maps to, the *D. virilis* chromosome associated with that scaffold, the map distance estimated from the Oulanka and Vancouver pedigrees and the number of individuals genotyped and the heterozygosity of the marker for each of the pedigrees (obtained using Cervus v3.0.3 (Kalinowksi et al., 2007)). The last column can be used to make a SNP chip order from Illumina for the BeadExpress platform.

**File S4**: A. Simulated maximum LRT score distribution. Summary of the maximum LRT score over 100 simulations. Suggestive and significant linkage correspond to LRT values 5.16 and 12.08, respectively. B. Mean values of LRT across 100 simulations. The simulated QTL is at 80 cM (arrow). LRT values 5.16 and 12.08 correspond to the suggestive and significant LOD scores 1.12 and 2.62 respectively. The plot can be regarded as ‘conservative’, because the
same position does not yield the maximum test stat each time and because it
includes the replicates where the QTL was not found.

File S5: Chromosome partitioning results. PVE: Percent variance
explained. LRT: Log likelihood ratio test. Three proxies of chromosome
length are provided. The *D. virilis* physical and genetic map lengths are taken
from Flybase.

File S6: QTL map of song PCs. The result is qualitatively similar to the
analysis of individual song characters. The PCs are the same as in (Veltsos et
al., 2012).

File S7: Comparison of PVE by chromosome in the two populations.
The scale is the same in all plots. While the values often correlate, there are
clear occasions where one chromosome in one population explains more
variance than in the other population.
References

592
593 Arbuthnott D (2009). The genetic architecture of insect courtship behavior and premating
594
596
597 Aspi J, Hoikkala A (1993). Laboratory and natural heritabilities of male courtship song
598
599 Aspi J, Hoikkala A (1995). Male mating success and survival in the field with respect to size
and courtship song characters in Drosophila littoralis and D. montana (Diptera,
601
603
605 tool for annotation, visualization and analysis in functional genomics research.
607
association studies establish that human intelligence is highly heritable and polygenic.
609
610 Dean R, Mank JE (2014). The role of sex chromosomes in sexual dimorphism: discordance
611
613
615 Hempstead, UK.
616
617 Gleason JM (2005). Mutations and natural genetic variation in the courthsip song of
618
619 Goring HHH, Terwilliger JD, Blangero J (2001). Large upward bias in estimation of
locus-specific effects from genomewide scans. The American Journal of Human Genetics 69:
1357–1369.
620
622
624
quantitative trait locus for intra- and interspecific variation in a sex pheromone.
627
630
632 Epidemiology 14: 1011–1016.


Salminen T, Vesala L, Hoikkala A (2012). Photoperiodic regulation of life-history traits before and after eclosion: egg-to-adult development time, juvenile body mass and reproductive


Slate J (2013). From beavis to beak colour: A simulation study to examine how much QTL mapping can reveal about the genetic architecture of quantitative traits. *Molecular Ecology: no-no.*


