Genetic divergence and phenotypic plasticity contribute to variation in cuticular hydrocarbons in the seaweed fly *Coelopa frigida*

Emma Berdan[1*], Swantje Enge[2,3], Göran M. Nylund[3], Maren Wellenreuther[4,5], Gerrit A. Martens[6], and Henrik Pavia[3]

[1]Department of Marine Sciences, University of Gothenburg, Göteborg, Sweden
[2]Institute for Chemistry and Biology of the Marine Environment, Carl-von-Ossietzky University Oldenburg, Wilhelmshaven, Germany
[3]Department of Marine Sciences – Tjärnö, University of Gothenburg, Strömstad, Sweden
[4]The New Zealand Institute for Plant & Food Research Limited, Nelson, New Zealand
[5]School of Biological Sciences, the University of Auckland, New Zealand
[6]Hochschule Bremen, City University of Applied Sciences, Bremen, Germany

* - Corresponding author: Emma Berdan, Department of Marine Sciences, Box 461, 405 30 Göteborg, Sweden. Email: emma.berdan@gmail.com

Cuticular hydrocarbons (CHCs) form the boundary between insects and their environments and often act as essential cues for species, mate and kin recognition. This complex polygenic trait can be highly variable both among and within species, but the causes of this variation, especially the genetic basis, are largely unknown. In this study, we investigated phenotypic and genetic variation of CHCs in the seaweed fly, *C. frigida,* and found that composition was affected by both genetic (sex and population) and environmental (larval diet) factors. We subsequently conducted behavioral trials that

show CHCs are likely used as a sexual signal. We identified general shifts in CHC chemistry as well as individual compounds and found that the methylated compounds, mean chain length, proportion of alkenes, and normalized total CHCs differed between sexes and populations. We combined this data with whole genome re-sequencing data to examine the genetic underpinnings of these differences. We identified 11 genes related to CHC synthesis and found population level outlier SNPs in 5 that are concordant with phenotypic differences. Together these results reveal that the CHC composition of *C. frigida* is dynamic, strongly affected by the larval environment, and likely under natural and sexual selection.

Keywords: cuticular hydrocarbons, *Coelopa frigida*, sexual signal, population differentiation, diet.

INTRODUCTION

In insects, cuticular hydrocarbons (CHCs) are a primary adaptation to life on land because they protect against desiccation (Wigglesworth 1945; Blomquist and Bagnères 2010). However, in many solitary and social insects, cuticular hydrocarbons are also used as one of the primary cues to recognize, and possibly discriminate between species, sexes, and among kin (Bagneres et al. 1996; Ferveur 2005; Blomquist and Bagnères 2010; Van Oystaeyen et al. 2014). The multifarious use of CHCs in adaptation and communication means that the composition is frequently under both natural (Howard and Blomquist 2005; Blomquist and Bagnères 2010; Foley and Telonis-Scott 2011;

47 Rajpurohit et al. 2017) and sexual selection (Ferveur 2005; Howard and Blomquist 2005;
48 Peterson et al. 2007; Thomas and Simmons 2009; Blomquist and Bagnères 2010; Steiger
49 et al. 2013).

50 *De novo* synthesis of CHCs is well defined and conserved across insects (Howard and
51 Blomquist 2005; Blomquist and Bagnères 2010): Synthesis takes place in oenocytes and
52 starts with acetyl-CoA which is then elongated by a fatty acid synthase (FAS) forming a
53 long-chained fatty acyl-CoA. FASs may also add methyl groups to these fatty acids (de
54 Renobales et al. 1986; Blomquist et al. 1994; Juarez et al. 1996; Chung et al. 2014).
55 Elongases further lengthen these fatty acyl-CoAs and double bonds or triple bonds are
56 added by desaturases (Howard and Blomquist 2005; Blomquist and Bagnères 2010).
57 These are reduced to aldehydes by reductases and then these aldehydes are converted to
58 hydrocarbons by a decarboxylation reaction, which is mediated by a cytochrome P450
59 (Qiu et al. 2012). Coding or expression variation in any number of these enzymes can
60 lead to significant changes in CHC composition (ex: de Renobales et al. 1986; Reed et al.
61 1995; Qiu et al. 2012; Chung et al. 2014; Dembeck et al. 2015; Chen et al. 2016).

62 Changes in gene expression (Dallerac et al. 2000a; Chertemps et al. 2007; Shirangi et al.
63 2009; Feldmeyer et al. 2014) and/or coding sequences (Keays et al. 2011; Badouin et al.
64 2013; Kulmuni et al. 2013) can impact both the overall composition of CHCs and the
65 amount of specific compounds. Multiple environmental factors may also influence CHC
66 composition. For example, CHCs can vary under different temperature regimes (Toolson
67 and Kupersimbron 1989; Savarit and Ferveur 2002; Rouault et al. 2004), diets (Stennett
68 and Etges 1997; Liang and Silverman 2000; Stojkovic et al. 2014) or other climatic
69 variables (Howard et al. 1995; Kwan and Rundle 2010). Moreover, multiple studies have

70 found that CHC composition varies between natural populations (ex: Haverty et al. 1990;
71 Etges and Ahrens 2001; Frentiu and Chenoweth 2010; Perdereau et al. 2010), but it is
72 unknown how much of this variation is due to plasticity versus genetic divergence. The
73 contribution of different factors can be teased apart using laboratory studies. The role of
74 genetic factors can, for example, be assessed by raising different populations in the same
75 environment. Likewise, plasticity can be assessed by raising the same population in
76 multiple environments (i.e. common garden and reciprocal transplant experiments).
77 Determining the causes and consequences of CHC variation is an important general step
78 towards understanding the evolution of CHCs and their co-option as sexual signals.

79 The seaweed fly *Coelopa frigida* presents an attractive system in which to investigate
80 causes and consequences of variation in CHCs. These flies live in highly dynamic
81 environments (Berdan et al. 2018) making it likely that both genetic changes as well as
82 phenotypic plasticity affect CHC composition. *Coelopa frigida* lives in "wrackbeds",
83 accumulations of decomposing seaweed on shorelines that act as both habitat and food
84 source for larvae and adults. Local adaptation in response to wrackbed composition has
85 been demonstrated in Swedish *C. frigida* (Wellenreuther et al. 2017). Furthermore, CHCs
86 may be under sexual selection as it is possible that CHCs are used as a mating signal in
87 this system. Mating in *C. frigida* is characterized by intense sexual conflict with no
88 courtship behaviour but evidence of female choice. Males will approach and forcefully
89 mount females. Females are reluctant to mate and use a variety of responses to dislodge
90 the male, such as downward curling of the abdomen (to prevent contact), kicking of the
91 legs, and shaking from side to side (Day et al. 1990; Blyth and Gilburn 2011). It has been
92 consistently reported that mating attempts by large males are more likely to result in

93 copulation (Butlin et al. 1982; Gilburn et al. 1992, 1993) but the signals used for mate
94 choice are unknown.

95 Here we investigate the cuticular hydrocarbons of *C. frigida.* Specifically, we examine
96 variation in CHC signatures due to sex (possible signature of sexual selection),
97 population (possible signature of natural selection), and larval environment/diet
98 (phenotypic plasticity). We combine this with behavioural trials to test the significance of
99 CHCs in a mate choice context. We also investigate genetic variation that may underlie
100 this phenotypic variation. We discuss our findings in light of sexual selection and
101 adaptation in this species and others and outline future research areas that deserve
102 attention.

103 METHODS

104 Study Species

105 *Coelopa frigida* belongs to the group of acalyptrate flies which exclusively forage on
106 decomposing seaweed (Cullen et al. 1987). This species is found along the seashores of
107 Northern Europe (Mcalpine 1991), plays a vital role in coastal environmental biodiversity
108 (Griffin et al. 2018) and health by accelerating the decomposition of algae, allowing for
109 faster release of nutrients (Cullen et al. 1987).

Sampling

Fly larvae were collected in April and May 2017 from two Norwegian populations, Skeie (58.69733, 5.54083) and Østhassel (58.07068, 6.64346), and two Swedish populations, Stavder (57.28153, 12.13746) and Ystad (55.425, 13.77254). These populations are situated along an environmental cline in salinity, wrackbed composition, and wrackbed microbiome (Berdan et al., in prep; Day et al. 1983) from the North Sea to the Baltic Sea (Figure 1). Larvae were transported to Tjärnö Marine Laboratory of the University of Gothenburg where they completed development in an aerated plastic pot filled with their own field wrack in a temperature controlled room at 25 °C with a 12h/12h light-dark cycle. After adults eclosed in the lab, they were transferred to a new pot filled with standard wrack consisting of 50 % *Fucus spp*. and 50 % *Saccharina latissima* which had been chopped, frozen, and then defrosted. The exception to this is a subset of Ystad adults that were allowed to mate on the same material they had been collected in (field wrack). Thus we had 5 treatment groups: All four populations (Skeie, Østhassel, Stavder, Ystad) on standard wrack, and Ystad also on field wrack. Adult flies were allowed to lay eggs on the provided wrack and this second generation was raised entirely in a temperature controlled room at 25 °C with a 12 h/12 h light dark cycle. As larvae pupated they were transferred to individual 2 mL tubes with a small amount of cotton soaked in 5 % (w/v) glucose to provide moisture and food for the eclosing adult. This ensured virginity in all eclosing flies. The tubes were checked every day and the date of eclosure of each fly was noted. Two days after eclosure adult flies were frozen at -80 °C and kept there until CHC extraction.

132 ## CHC Analysis

133 Frozen flies were placed in 12-well porcelain plates to defrost and dry, as moisture can
134 affect lipid dissolution. Each fly was then placed in a 1.5 ml high recovery vial
135 containing 300 µl of *n*-hexane, vortexed at a low speed for 5 second, and extracted for 5
136 minutes. Afterwards flies were removed from the vial and allowed to air dry on the
137 porcelain plates before they were weighed (Sartorius Quintix 124-1S microbalance) to
138 the nearest 0.0001 grams and sexed. Extracts were evaporated until dry under nitrogen
139 gas and then stored at -20°. Before analysis 20 µl of *n*-hexane containing 1 µg/ml *n*-
140 nonane was added as an internal standard to each vial, which was then vortexed at
141 maximum speed for 10 seconds.

142 The extracts from 20 male and 20 female flies from every treatment group were analyzed
143 on a GC (Agilent GC 6890) coupled to a MS (Agilent 5973 MSD) using a HP-5MS
144 capillary column (Agilent) (see Table S1 in supplementary materials for instrument
145 settings).

146 The total ion chromatograms were quality checked, de-noised and Savitzky-Golay
147 filtered before peak detection and peak area integration in OpenChrom®. Peaks were
148 then aligned using the R-package 'GCalignR'(Ottensmann et al. 2018) prior to statistical
149 analysis. Peaks were tentatively identified by their mass spectra and retention time index
150 based on a C21-C40 *n*-alkane standard solution (Sigma-Aldrich).

Statistical Analysis

The effects of sex and population on CHC profile were assessed using flies raised on standard wrack using a balanced sampling design (N=13 for each combination) and 127 peaks were identified by the R package “GCalignR”. The effects of sex and diet (i.e. wrack type) during the larval stage on CHC profiles were analyzed using a balanced design of flies from the Ystad population (N=12 for each combination) and 111 peaks were identified by R package “GCalignR”. Peak areas were normalized on the peak area of the internal standard and the weight of the fly before being auto-scaled prior to statistical analysis. Clustering of samples was visualized with a PCA, and group differences were analyzed using a PERMANOVA followed by multiple group comparisons using the PRIMER-E 7 software. We also analyzed group differences via (O)PLS-DA using the “ropls”-packages in R (Thevenot et al. 2015). Candidate compounds for group differentiation were determined from variables of importance for OPLS-DA projection (VIP scores) (Galindo‐Prieto et al. 2014), which were further assessed by univariate analysis using a false discovery adjusted significance level of $\alpha = 0.05$ using the Benjamini and Hochberg’s FDR-controlling procedure (Benjamini & Hochberg, 1995).

We examined sex and country effects in more depth to examine shifts in chemistry that could be linked to our genetic data (see below). Our results above indicated that the differences in CHC profiles between sexes and populations of *C. frigida* were due to quantitative rather than qualitative differences (i.e. changes in relative amounts rather than presence/absence of certain compounds). Thus, we looked for general shifts in CHC

chemistry by examining 1. The total normalized peak area (i.e. the sum of all peaks used in analysis), 2. The proportion of alkenes, and 3. The proportion of methylated compounds. For #1 we used all 127 peaks used in the previous analysis and for #2 and #3 we used the subset of these peaks that could be accurately categorized (116 peaks). For the total peak area we transformed the data with a log transformation and then used the 'dredge' function from the MuMIn package (Barton 2017) to compare AIC values for nested glm (Gaussian distribution, link=identity) models with the terms sex, country, and country*sex. If there were two models that differed in AIC by less than 1, we chose the simpler model. For alkenes and methylated compounds our data was proportional and did not include 1 or 0 so we used the 'betareg' function from the betareg package (Grun et al. 2012). We again used the 'dredge' function to compare AIC values for nested models with the terms sex, country, and country*sex and took the simpler model in case of an AIC difference of less than 1.

We also analysed the distribution of chain lengths in our samples. To do this, we subsetted all peaks where we had accurate length information (107 peaks) and then calculated the weighted mean chain length as

$$N = \Sigma\, l_i p_i$$

where l_i is the length of a carbon chain and p_i is the proportional concentration of all chains of that length. We also calculated dispersion around this mean as

$$D = \Sigma\, p_i (l_i - N)^2$$

194 We analyzed these in the same manner as described above, using AIC values to conduct
195 model choice from glm models (Gaussian distribution, link=identity).

196 Behavioral Trials

197 Given experimental restrictions, we were unable to accurately test for the use of CHCs in
198 female choice and thus we tested for the use CHCs in male choice. Behavioral trials were
199 conducted with flies from the Østhassel population that were collected in March 2018
200 and cultured in the lab for several generations. To collect adults for testing, we selected
201 pupae and kept them in individual Eppendorf tubes (to ensure virginity) containing a
202 small piece of cotton soaked in a 0.5% mannitol solution. When a fly eclosed, the
203 Eppendorf was moved to the refrigerator to slow down aging. In all behavioural trials,
204 five males (marked individually with nail polish) were introduced to a small petri dish (Ø
205 55mm) with a small amount of minced *Saccharina* (≈1g) and two dead flies mounted on
206 insect pins. We conducted two different types of trials: I. Males were able to choose
207 between a dead male and a dead female and II. Males were able to choose between two
208 dead females one of which had been extracted in hexane (see above for methods). Type 1
209 trials were used to confirm that our methodology was working and type II trials were
210 used to test if CHCs play a role in male choice. For type II trials we checked that
211 extracted females were mostly devoid of CHCs by re-extracting five females and
212 performing GC-MS analysis as described above. The hexane extraction reduced the total
213 amount of CHCs on the flies by 72% ±6. All trials lasted for 30 minutes and we recorded
214 all mountings and how long they lasted. We used a paired Wilcoxon rank-sum test on our

215 data, with male as our unit of replication, to determine if males preferentially mounted
216 one of the two proffered flies.

217 Candidate gene analysis

218 Based on a literature search for genes involved in CHC synthesis, we identified 37
219 reference protein sequences, which were downloaded from Flybase (flybase.org) in
220 October 2017. We used tblastx with the default settings to identify orthologs in the *C.*
221 *frigida* transcriptome (Berdan & Wellenreuther, unpublished). We then used gmap (Wu
222 and Watanabe 2005) to find areas of the *C. frigida* genome (Wellenreuther, *et al*.,
223 unpublished) that might contain these genes. We output all alignments as a gff3 file. We
224 have a previously made a VCF file created from 30x whole genome re-sequencing of 46
225 *C. frigida* across five populations in Scandinavia (Berdan *et. al*, unpublished). This VCF
226 file was made by aligning raw reads to the *C. frigida* genome using bwa-mem (Li and
227 Durbin 2010), duplicate reads were marked using 'picard' (http://picard.sourceforge.net).
228 SNPs were called with the Genome Analysis Toolkit (GATK; DePristo *et al.* 2011; Van
229 der Auwera *et al.* 2013) specifically the GATK-module 'UnifiedGenotyper' (Van der
230 Auwera *et al.* 2013). VCF filtering was done as described in Berdan *et al.* 2015. We used
231 this VCF file and the gff3 output from gmap in SNPeff (Cingolani et al. 2012) to annotate
232 the SNPs in these genes.

233 To find which SNPs are diverged between populations we used vcftools (Danecek et al.
234 2011) to calculate Cockerham and Weir's F_{ST} estimate (Weir and Cockerham 1984) for
235 all SNPs in our reference genome. We retained the top 5% of this distribution ($F_{ST} \geq$

236 0.142, mean F_{ST} 0.024) as SNPs that were potentially divergent. We then subset the VCF
237 file to include only outlier SNPs located in our previously identified genomic regions.

238 RESULTS

239 CHC analysis

240 A diverse blend of hydrocarbons with more than 100 different compounds was found in
241 the cuticular extract of *C. frigida*. The majority of CHCs ranged in chain length from 23
242 to 33 carbon atoms. They contained odd numbered *n*-alkanes, methyl-, dimethyl- and
243 trimethyl-alkanes, as well as odd numbered alkenes. Some even numbered alkanes were
244 also present but in lower quantities.

245 Sex and population effects

246 Results of the PERMANOVA indicated significant differences between females and
247 males and between populations but also a significant interaction between sex and
248 population (Table 1A). Pairwise tests on the interaction showed a significant difference
249 between males and females for all populations except Ystad (Table S2). Multiple
250 comparisons of populations showed that the Norwegian populations (Skeie and
251 Østhassel) were similar while the Swedish populations (Ystad and Stavder) differed for
252 males but not for females. In general the Norwegian populations tended to differ from the
253 Swedish and a similar pattern was also shown in the results of the PCA and PLS-DA
254 (Figure 2A, C, Table S3). After combining the Swedish and Norwegian populations, the
255 PERMANOVA indicated highly significant differences between females and males as

256 well as between Sweden and Norway (Table 1B). OPLS-DA models for both sex and
257 country were significant and lead to separate clusters of females and males, and Sweden
258 and Norway, respectively (Figure 2B, D, Table S3). The OPLS-DA model revealed 41
259 peaks that contributed with a VIP score >1 to the differentiation between females and
260 males. Of these, 31 were also significant in t-tests with a FDR adjusted p-values (Table
261 2). Generally compounds were more abundant in males with the exception of only 3
262 peaks that were more abundant in females. Looking at country differentiation between
263 Swedish and Norwegian populations we found 43 compounds with a VIP score >1, of
264 which 29 were significant in t-tests with FDR adjusted p-values and consisted of an equal
265 mix of compounds more abundant in either of the countries (Table 2). Differences in
266 CHC profiles between sexes and populations were due to quantitative differences,
267 affecting a wide range of compounds instead of qualitative differences, i.e. the presence/
268 absence of a few distinct compounds

269 Sex and diet effects

270 PERMANOVA demonstrated a significant difference between females and males in the
271 Ystad population and also a significant effect of the larval diet (wrack type) (Table 1C).
272 Both OPLS-DA models for either sex or wrack type were significant and lead to separate
273 clusters of females and males and flies raised on standard or field wrack (Figure 3B,C,
274 Table S3). The OPLS-DA model revealed 38 peaks that contributed to the differentiation
275 between females and males with a VIP score >1. Of these, 24 compounds were also
276 significant in a t-test with FDR adjusted p-values (Table 2). All of these, except for 3
277 peaks, were more abundant in males. Overall 65% (26 compounds) were shared between
278 this analysis and the previous analysis investigating sex and population effects. Looking

279 at differentiation between flies raised on a diet consisting of standard or field wrack we
280 found 33 compounds with a VIP score >1. Of these, 27 compounds were significant
281 using a t-test with FDR adjusted p-values, all with higher amounts in flies reared on field
282 wrack (Table 2). Flies reared on field wrack had more CHCs per gram body weight than
283 flies reared on standard wrack (normalized totals: Females field wrack 7.96 ± 0.94,
284 females standard wrack 4.09 ± 0.47, males field wrack 11.24 ± 1.03, males standard
285 wrack 7.09 ± 1.21). Differences in CHC profiles between sexes and diets were
286 exclusively due to quantitative differences, affecting a wide range of compounds instead
287 of qualitative differences, i.e. the presence/ absence of a few distinct compounds

288 Shifts in chemistry

289 As our results above indicated an overall effect of country rather than population we used
290 country and sex for our terms in the glm models. For total normalized peak area, the best
291 model contained only sex as a main effect (for a full AIC comparison for all models see
292 Table S4A-E). In general, males had more CHCs per gram body weight than females
293 (Figure 4A). For the proportion of alkenes, the best model contained country and sex as
294 main effects (for type II analysis of deviance tables for all best models see Table S5A-E).
295 Females tended to have more alkenes than males and individuals from Norway had more
296 alkenes than individuals from Sweden (Figure 4B). The best model for the proportion of
297 methylated compounds contained sex and country as the main effects. As with alkenes,
298 females tended to have more methylated compounds than males and individuals from
299 Norway tended to have more methylated compounds than individuals from Sweden
300 (Figure 4C).

301 The best model for weighted mean chain length was country and sex. Overall, males had
302 longer mean chain lengths than females and individuals from Sweden had longer mean
303 chain lengths than individuals from Norway (Figure 5A). The best model for dispersion
304 around mean chain length was also country and sex. Males had higher dispersion than
305 females and individuals from Norway had higher dispersion than individuals from
306 Sweden (Figure 5B).

307 Behavioural trials

308 We conducted 20 trials of type I (female vs. male) and 12 trials of type II (female vs.
309 extracted female). We only retained data from males that mounted at least one of the
310 proffered flies leaving us with data from 28 males (type I) and 33 males (type II). In type
311 I trials, males preferred the female to the male and on average spent 73% (± 7%) of their
312 "mating time" mounting her (Wilcoxon signed rank test with continuity correction, V =
313 220, $P = 0.0132$). In type 1I trials males preferred the control (i.e. non-extracted) female
314 and on average spent 82% (± 6%) of their "mating time" mounting her (Wilcoxon signed
315 rank test with continuity correction, V = 493, $P = 0.0002$).

316 Genetic analysis

317 We identified 15 isoforms from 13 transcripts in our transcriptome with matches to our
318 query proteins. These transcripts mapped to 16 unique areas of the genome assembly. We
319 blasted these areas against the NCBI nr database (accessed in April 2018) to confirm our
320 annotation. Five of these areas had either no match or matched to an unrelated protein
321 and were thus removed. The remaining genes include 2 putative fatty acid synthases, 2
322 putative desaturases, 5 putative elongases, 1 putative cytochrome P450 reductase, and 1

323 putative cytochrome P450-4G1 (Table 3). We found 1042 SNPs within these genes
324 (including 5K upstream and 5K downstream), of these 11 had an F_{ST} value > 0.142 (i.e.
325 in the top 5% of the F_{ST} distribution, Table 4).

326 DISCUSSION

327 Here we explored causes of variation in cuticular hydrocarbons (CHCs) in the seaweed
328 fly *C. frigida*. We detected considerable phenotypic CHC variation attributable to sex,
329 population, and diet and were able to describe genetic variation in candidate genes
330 involved in CHC synthesis that was attributable to population. We show that this trait is
331 likely used as a signal in male mate choice indicating that CHC composition may be
332 under both natural and sexual selection in this system. We discuss these results and their
333 consequences below.

334 Males and females differed in their CHC composition. Not only did the composition
335 between males and females differ (Table 2, Figure 2-5) but, in addition to that, males
336 consistently had more CHC per gram body weight than females (Figure 4A). Our
337 behavioural trials demonstrate that males preferentially mount females with an intact
338 signal (i.e. CHC cocktail) indicating that CHCs are likely used in male mate choice.
339 Although we were unable to test female choice, it is likely that females also use CHCs in
340 mate choice. A role for CHCs in sexual selection in *C. frigida* would be in line with the
341 general finding that CHCs play a major role in communication in insects, specifically in
342 Diptera (Howard and Blomquist 2005; Blomquist and Bagnères 2010; Ferveur and Cobb
343 2010). Differences in CHCs have been shown to be involved in a wide variety of

344 behaviours in diptera such as courtship, aggregation, and dominance (Ferveur and Cobb
345 2010). For instance, 11-*cis*-vaccenyl acetate, a male specific hydrocarbon in *Drosophila*
346 *melanogaster*, increases male-male aggression while suppressing male mating (Wang and
347 Anderson 2010). The same compound is also involved in aggregation behaviour in both
348 *D. simulans* and *D. melanogaster* (Bartelt et al. 1985; Schaner et al. 1987). Other
349 aggregation hydrocarbons, such as (Z)-10-heneicosene, in *D. virilis* have been shown to
350 attract males and females of certain ages (Bartelt and Jackson 1984). Further behavioural
351 tests in *C. frigida* can determine if CHCs are additionally used in aggregation and male-
352 male interactions.

353 We found a shift in the composition between sexes, populations, and diets rather than
354 differences in the presence/absence of specific compounds. This type of pattern differs
355 from some dipterans where sex and population differences are mostly qualitative
356 (Carlson and Schlein 1991; Ferveur and Sureau 1996; Dallerac et al. 2000b; Gomes et al.
357 2008; Everaerts et al. 2010) although many dipterans also show quantitative differences
358 (ex: Jallon and David 1987; Byrne et al. 1995). This discrepancy is likely to be partially
359 explained by the distribution of compounds that make up the CHCs in *C. frigida*. While
360 *C. frigida* chain lengths and compound classes are comparable with other dipteran
361 species (Ferveur and Cobb 2010) the distribution of these compounds is somewhat
362 different. Many dipterans have principal CHC components that make up a large
363 proportion of the CHC composition (Jallon and David 1987; Ferveur and Sureau 1996;
364 Etges and Jackson 2001; Gomes et al. 2008), while the most abundant compound (the
365 C25 alkane) observed in this study only made up on average 15 % of the CHCs in *C.*
366 *frigida*.

367

368 We found strong geographic signatures in CHC composition when flies were raised in a
369 common garden, which indicates a shift in genetic variation. We examined coding
370 variation to determine if we could tie population level phenotypic differences to genetic
371 differences. We observed country level variation with strong differences between
372 Norway and Sweden (Figure 2D, Figure 4, Figure 5). This pattern mirrors a strong
373 genetic split between the two countries (representing approximately 14% of genetic
374 variation, Berdan *et al*, unpublished). We found potential outlier SNPs in desaturases,
375 elongases, a cytochrome P450-4G1, and a cytochrome P450 reductase. Desaturases add
376 double bonds or triple bonds to alkanes (Howard and Blomquist 2005; Blomquist and
377 Bagnères 2010). We found an outlier missense variant in a putative Desaturase 1 as well
378 as a missense and two downstream outlier variants in a putative Desaturase (Table 4).
379 This aligns with our phenotypic data showing differences in the proportion of alkanes vs.
380 alkenes between Norway and Sweden. Elongases lengthen fatty acyl-CoAs and are
381 necessary for chains longer than 16 carbon atoms (Blomquist and Bagnères 2010). We
382 found both synonymous and downstream outlier variants in putative *C. frigida* elongases
383 (Table 4) and corresponding differences in both the mean chain length and variance of
384 the chain length (dispersion around the mean) due to country of origin. P450 reductases
385 are responsible for reducing fatty acids to aldehydes (Blomquist and Bagnères 2010).
386 Knockdown of P450 reductases in *D. melanogaster* leads to a striking reduction of
387 cuticular hydrocarbons (Qiu et al. 2012). P450-4G genes, which are unique to insects,
388 encode an oxidative decarboxylase that catalyzes the aldehyde to hydrocarbon reaction.
389 Work in other insect species has found that a knockdown or knockout of this gene leads

390 to a decrease in overall hydrocarbon levels (Reed et al. 1995; Qiu et al. 2012; Chen et al.
391 2016). We found a downstream outlier variant in our cytochrome P450-4G1 and multiple
392 kinds of outlier variants in a cytochrome P450 reductase. However, the total amount of
393 CHCs did not differ between populations, only sexes. As males and females share a
394 genome (with the exception of the sex chromosome) it is unlikely that these SNPs affect
395 this difference. Finally, we also found variation between Sweden and Norway in the
396 proportion of methylated compounds. Other studies suggest that differences in the
397 proportion of methylated compounds are often caused by either variation in Fatty Acid
398 Synthase (FAS, de Renobales et al. 1986; Blomquist et al. 1994; Juarez et al. 1996) or
399 multiple copies of FAS with different functions (Chung et al. 2014). We found two
400 different FAS' in our genome (Table 3) but neither of them contained divergent SNPs.
401 Future studies are needed to examine if these genetic and phenotypic changes are related.
402 Specifically, investigations into both RNA expression and tests of function of these loci
403 will be necessary to provide a direct link. However, the association between the known
404 function of several of these genes (elongases, desaturases) and the corresponding
405 differences in populations are marked. Given that we see shifts in multiple chemical
406 aspects (length, methylation, and alkene/alkane ratio) it is likely that multiple loci are
407 responsible for these patterns.

408 We also detected plasticity in CHC composition due to the environment; the wrackbed
409 itself had a strong influence on the CHC composition (Figure 3C). Wrackbed
410 composition and microbiome (i.e. the food source for *C. frigida* larvae) varies across *C.*
411 *frigida* populations in Europe (Berdan, et al., in prep, Day et al. 1983; Butlin and Day
412 1989; Wellenreuther et al. 2017). Consequently, the CHC composition of *C. frigida* is

413 likely to differ between natural populations in accordance with other research showing
414 that larval diet impacts CHC composition (Liang and Silverman 2000; Rundle et al. 2005;
415 Etges and de Oliveira 2014; Stojkovic et al. 2014).

416 Together our results indicate that both genetic and environmental factors influence the
417 CHC composition in *C. frigida*. This variation may lead to several potential evolutionary
418 scenarios: 1. Natural selection may select for a different CHC composition in different
419 environments and natural polymorphism in CHC genes may be under direct selection due
420 to this. For example, *Drosophila serrata* and *D. birchii* differ in the environment they
421 inhabit (habitat generalist vs. habitat specialist for humid rainforests). These two species
422 also differ in the concentration of methyl branched CHCs (especially important in
423 desiccation resistance) caused, in part, by changes in the cis-regulatory sequence likely
424 under natural selection (Chung et al. 2014). 2. Natural populations differ strongly in their
425 CHC composition due to environmental effects (i.e plasticity). If assortative mating or
426 selection on preferences is present, this could lead to locally distinct traits and
427 preferences (Chung and Carroll 2015). This would reduce effective migration between
428 populations, as foreign males would be at a disadvantage. Differences in CHC
429 composition and corresponding preferences have been shown to cause behavioral
430 isolation between many different *Drosophila* species or populations (Coyne et al. 1994;
431 Rundle et al. 2005; Etges and de Oliveira 2014).

432 In conclusion, the analysis of CHC extracts from male and female *C. frigida* from
433 multiple populations revealed a complex and variable mix of more than 100 different
434 compounds. We confirmed that CHCs are likely used in mate choice and found extensive
435 phenotypic variation attributable to diet, sex, and country as well as associated genetic

436 variation attributable to the latter. This reveals that CHC composition is dynamic,
437 strongly affected by the larval environment, and most likely under natural and sexual
438 selection. Further work is needed to explore the evolutionary consequences of these
439 differences.

440

441

442 **Acknowledgements**: We thank the members of the Marine Chemical Ecology group at
443 Gothenburg University for helpful comments on experimental design. E.B. was supported
444 by a Marie Skłodowska-Curie fellowship 704920 — ADAPTIVE INVERSIONS —
445 H2020-MSCA-IF-2015. MW and the research were supported by grants from
446 Vetenskapsrådet to MW (2012-3996), from the Mistra Foundation to HP (AquaAgriKelp)
447 and from the Swedish Foundation for Strategic Research to HP (Sweaweed).

Tables

Table 1– PERMANOVA and PERMDISP results for A. Sex and Population, B. Sex and Country, and C. Sex and Wrack type. * < 0.05 ** < 0.01 *** < 0.001

A.

	PERMANOVA: Sex and Population						PERMDISP
	df	SS	MS	F	p-value		
Sex	1	658.1	658.1	5.86	<0.001	***	see S2
Population	3	1150.3	383.4	3.41	<0.001	***	
Sex:Population	3	486.9	162.3	1.44	0.015	*	
Residuals	96	10786.0	112.4				
Total	103	13081					

B.

	PERMANOVA: Sex and Country						PERMDISP		
	df	SS	MS	F	p-value		F	p-value	
Sex	1	658.0	658.1	5.60	<0.001	***	8.94	0.003	**
Country	1	517.8	517.8	4.41	<0.001	***	1.19	0.279	
Sex:Country	1	154.1	154.1	1.31	0.128				
Residuals	100	11751.0	117.5						
Total	103	13081							

C.

	PERMANOVA: Sex and Wrack type						PERMDISP	
	df	SS	MS	F	p-value		F	p-value
Sex	1	360.9	360.9	3.73	<0.001	***	2.36	0.129
Wrack	1	481.3	481.3	4.98	<0.001	***	2.29	0.139
Sex:Wrack	1	122.3	122.3	1.27	0.1629			
Residuals	44	4252.5	96.7					
Total	47	5217						

Table 2 – List of compounds that are significantly important for differentiation of at least factor (sex, diet, or country). If compounds are significant for differentiation of any category in any comparison a * appears under that factor. Color indicates the direction of the differentiation (red- higher in females, blue – higher in males, grey – higher in Norwegian populations, green – higher in Swedish populations, yellow – higher in flies raised on standard wrack, purple – higher in flies raised on field wrack).

				Importantance in differentiation for factor			
Compound number	RI	Identified as	Methyl group position	Sex in Sex & Country	Country in Sex & Country	Sex in Sex & Diet	Wrack type in Sex & Diet
1	2096	unknown alkane					*
2	2101	C21			*		
3	2176	unknown alkane			*		
5	2202	C22		*			
7	2300	C23		*	*	*	
8	2344	unknown alkane		*		*	
9	2365	2-ME-C23			*		
11	2400	C24		*		*	*
14b	2478	25:x-C25			*		
16	2500	C25					*
17	2524	unknown alkane					*
20	2564	2-ME-C25					*
21	2576	3-ME-C25					*
22	2584	5,X-Me-C25	X=9,11,13,15	*		*	
23	2600	C26					*
24	2638	unknown alkane		*			
26	2665	2-Me-C26					*
27	2680	27:x-C27		*		*	
29	2701	C27			*		*
30	2712	unknown alkane					*
31	2735	11+13-Me-C27			*		*
33	2765	2-Me-C27 + 11,15-Me-C27			*		*
34	2775	3-ME-C27					*
37a	2807	3,X-Me-C27	X=15,13,11,7				*
40	2850	29:x,x-C29					*
41	2857	29:x,x-C29			*		
42	2865	2-ME-C28 or x:C29			*		*
44	2880	29:x-C29				*	*
45	2882	29:x-C29			*		
46	2891	4,12-Me-C28					*
47	2902	C29			*		*
48	2933	15+13+11-Me-C29			*		*
50	2961	11,X-ME-C29	X=15,17	*	*		
51	2972	7,X-ME-C29	X=11,13,15	*	*	*	
52	2982	5,13-Me-C29		*	*		
53	3007	3,13-ME-C29			*		*
54	3033	3,X,15-Me-C29	X= 11,13	*		*	
55	3042	3,7,15-Me-C29			*		
56	3054	31:x,x-C31		*			
57	3058	12,16-Me-C30			*		
60	3074	31:x-C31 or 6,16-Me-C30			*		
61	3082	31:x-C31			*	*	
62	3090	31:x-C31 or 4,16-Me-C30		*			
63	3101	C31		*	*	*	*
64	3118	unknown alkane		*	*		*
65	3131	15+13+11+9-Me-C31		*		*	*
66	3157	15,19-Me-C31		*		*	
67	3163	9,X-Me-C31	X=13,15	*	*	*	
68	3170	7,X-Me-C31	X=15,17	*		*	
69	3180	5,15-Me-C31		*		*	
71	3194	unknown alkane		*		*	
72	3205	3,X-Me-C31	X=9,11,13	*		*	
73a	3231	3,X,Y-Me-C31	X=9,11 ; Y=13,15	*		*	
74	3246	33:x,x,x,x-C33		*			
75	3256	33:x,x-C33			*		
76	3272	33:x-C33			*		
77	3284	33:x-C33			*		
78	3300	unknown alkane		*	*	*	*
79	3315	unknown alkene					*
80	3330	17+15+13+11-Me-C33		*		*	*
81	3353	11,X-Me-C33	X=13,15,17	*			
82	3359	9,X-Me-C33	X=11,13,15		*		
83a	3365	7,X-Me-C33	X=11,13,15,17			*	
83b	3368	7,X-Me-C33	X=11,13,15,17	*		*	
84	3377	5,X-Me-C33	X=15,17	*		*	
85a	3391	7,13,17-Me-C33		*		*	
85b	3393	7,13,17-Me-C33		*			

higher in females
higher in males
higher in Swedish populations
higher in Norwegian populations
higher in flies raised on standard wrack
higher in flies raised on field wrack
* significant t-test with FDA α=0.05

Table 3 – Putative genes involved in CHC synthesis in *C. frigida*

Scaffold[1]	Gene Position[1]	Best Blast Hit[2]
scaffold_129	185034-197861	Desaturase 1
scaffold_155	109640-112144	Cytochrome P450-4G1
scaffold_206	225893-231312	Cytochrome P450 reductase
scaffold_480	98147-106166	CG5326 (elongase family)
scaffold_480	2453-12223	CG31522 (elongase family)
scaffold_545	1553-3217	CG5326 (elongase family)
scaffold_340	1141-16368	Baldspot (elongase family)
scaffold_149	7122-16098	Fatty Acid Synthase 3
scaffold_51	250672-259608	Fatty Acid Synthase 1
scaffold_716	36957-55451	CG2781 (elongase family)
scaffold_994	14376-17314	CG9743 (contains Fatty acid desaturase domain and Acyl-CoA desaturase)

1. From unpublished genome assembly (Wellenreuther *et. al*)
2. From a BLAST against *D. melanogaster* annotated nucleotides on flybase.org

Table 4 – Outlier SNPs putatively involved in geographic differences

Scaffold[1]	Position[1]	Variant type[2]	Best Blast Hit[3]	F_{ST}[4]
scaffold_129	195074	missense variant	Desaturase 1	0.16
scaffold_155	116876	downstream gene variant	Cytochrome P450-4G1	0.22
scaffold_206	225645	downstream gene variant	Cytochrome P450 reductase	0.17
scaffold_206	226804	missense variant	Cytochrome P450 reductase	0.17
scaffold_206	229444	missense variant	Cytochrome P450 reductase	0.20
scaffold_206	230032	intron variant	Cytochrome P450 reductase	0.30
scaffold_480	105896	synonymous variant	CG5326 (elongase family)	0.22
scaffold_480	110918	downstream gene variant	CG5326 (elongase family)	0.16
scaffold_994	16976	missense variant	CG9743 (contains Fatty acid desaturase domain and Acyl-CoA desaturase)	0.16
scaffold_994	18844	downstream gene variant	CG9743 (contains Fatty acid desaturase domain and Acyl-CoA desaturase)	0.20
scaffold_994	19574	downstream gene variant	CG9743 (contains Fatty acid desaturase domain and Acyl-CoA desaturase)	0.20

1. From unpublished genome assembly (Wellenreuther *et. al*)
2. Annotated in SnpEff (see text for details)
3. From a BLAST against *D. melanogaster* annotated nucleotides on flybase.org
4. Calculated from a separate unpublished data set of WGS from 46 *C. frigida*

REFERENCES

Badouin, H., K. Belkhir, E. Gregson, J. Galindo, L. Sundström, S. J. Martin, R. K. Butlin, and C. M. Smadja. 2013. Transcriptome characterisation of the ant *Formica exsecta* with new insights into the evolution of desaturase genes in social hymenoptera. PloS one 8:e68200.

Bagneres, A. G., M. C. Lorenzi, G. Dusticier, S. Turillazzi, and J. L. Clement. 1996. Chemical usurpation of a nest by paper wasp parasites. Science 272:889-892.

Bartelt, R. and L. Jackson. 1984. Hydrocarbon component of the *Drosophila virilis* (Diptera: Drosophilidae) aggregation pheromone:(Z)-10-heneicosene. Annals of the Entomological Society of America 77:364-371.

Bartelt, R. J., A. M. Schaner, and L. L. Jackson. 1985. cis-Vaccenyl acetate as an aggregation pheromone in *Drosophila melanogaster*. Journal of chemical ecology 11:1747-1756.

Barton, K. 2017. MuMIn. https://CRAN.R-project.org/package=MuMIn

Berdan, E., H. Rosenquist, K. Larson, and M. Wellenreuther. 2018. Inversion frequencies and phenotypic effects are modulated by the environment: insights from a reciprocal transplant study in *Coelopa frigida*. Evolutionary Ecology:1-16.

Berdan, E. L., C. J. Mazzoni, I. Waurick, J. T. Roehr, and F. Mayer. 2015. A population genomic scan in *Chorthippus* grasshoppers unveils previously unknown phenotypic divergence. Molecular Ecology 24:3918-3930.

Blomquist, G. J. and A.-G. Bagnères. 2010. Insect hydrocarbons: biology, biochemistry, and chemical ecology. Cambridge University Press.

Blomquist, G. J., L. Guo, P. Gu, C. Blomquist, R. C. Reitz, and J. R. Reed. 1994. Methyl-branched fatty acids and their biosynthesis in the housefly, *Musca domestica* L.(Diptera: Muscidae). Insect Biochem Molec 24:803-810.

Blyth, J. E. and A. S. Gilburn. 2011. The function of female behaviours adopted during premating struggles in the seaweed fly, *Coelopa frigida*. Anim Behav 81:77-82.

Butlin, R. and T. Day. 1989. Environmental correlates of inversion frequencies in natural populations of seaweed flies (*Coelopa frigida*). Heredity 62:223.

Butlin, R. K., I. L. Read, and T. H. Day. 1982. The effects of a chromosomal inversion on adult size and male mating success in the seaweed fly, *Coelopa frigida*. Heredity 49:51-62.

Byrne, A., M. Camann, T. Cyr, E. Catts, and K. Espelie. 1995. Forensic implications of biochemical differences among geographic populations of the black blow fly, *Phormia regina* (Meigen). Journal of Forensic Science 40:372-377.

Carlson, D. and Y. Schlein. 1991. Unusual polymethyl alkenes in tsetse flies acting as abstinon in *Glossina morsitans*. Journal of chemical ecology 17:267-284.

Chen, N., Y. L. Fan, Y. Bai, X. D. Li, Z. F. Zhang, and T. X. Liu. 2016. Cytochrome P450 gene, CYP4G51, modulates hydrocarbon production in the pea aphid, *Acyrthosiphon pisum*. Insect Biochem Molec 76:84-94.

Chertemps, T., L. Duportets, C. Labeur, R. Ueda, K. Takahashi, K. Saigo, and C. Wicker-Thomas. 2007. A female-biased expressed elongase involved in long-chain hydrocarbon biosynthesis and courtship behavior in *Drosophila melanogaster*. P Natl Acad Sci USA 104:4273-4278.

Chung, H. and S. B. Carroll. 2015. Wax, sex and the origin of species: Dual roles of insect cuticular hydrocarbons in adaptation and mating. Bioessays 37:822-830.

Chung, H., D. W. Loehlin, H. D. Dufour, K. Vaccarro, J. G. Millar, and S. B. Carroll. 2014. A single gene affects both ecological divergence and mate choice in *Drosophila*. Science 343:1148-1151.

Cingolani, P., A. Platts, L. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Y. Lu, and D. M. Ruden. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w(1118); iso-2; iso-3. Fly 6:80-92.

Coyne, J. A., A. P. Crittenden, and K. Mah. 1994. Genetics of a pheromonal difference contributing to reproductive isolation in *Drosophila*. Science 265:1461-1464.

Cullen, S. J., A. M. Young, and T. H. Day. 1987. Dietary requirements of seaweed flies (*Coelopa frigida*). Estuarine, Coastal and Shelf Science 24:701-710.

Dallerac, R., C. Labeur, J.-M. Jallon, D. C. Knipple, W. L. Roelofs, and C. Wicker-Thomas. 2000a. A Δ9 desaturase gene with a different substrate specificity is responsible for the cuticular diene hydrocarbon polymorphism in *Drosophila melanogaster*. Proceedings of the National Academy of Sciences 97:9449-9454.

Dallerac, R., C. Labeur, J. M. Jallon, D. C. Knippie, W. L. Roelofs, and C. Wicker-Thomas. 2000b. A Delta 9 desaturase gene with a different substrate

specificity is responsible for the cuticular diene hydrocarbon polymorphism in *Drosophila melanogaster*. P Natl Acad Sci USA 97:9449-9454.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, and G. P. A. Grp. 2011. The variant call format and VCFtools. Bioinformatics 27:2156-2158.

Day, T. H., C. Dawe, T. Dobson, and P. C. Hillier. 1983. A chromosomal inversion polymorphism in Scandinavian populations of the seaweed fly, *Coelopa frigida*. Hereditas 99:135-145.

Day, T. H., S. P. Foster, and G. Engelhard. 1990. Mating behavior in seaweed flies (*Coelopa frigida*). Journal of Insect Behavior 3:105-120.

de Renobales, M., T. S. Woodin, and G. J. Blomquist. 1986. *Drosophila melanogaster* fatty acid synthetase: characteristics and effect of protease inhibitors. Insect biochemistry 16:887-894.

Dembeck, L. M., K. Boroczky, W. Huang, C. Schal, R. R. H. Anholt, and T. F. C. Mackay. 2015. Genetic architecture of natural variation in cuticular hydrocarbon composition in *Drosophila melanogaster*. Elife 4.

Etges, W. J. and M. A. Ahrens. 2001. Premating isolation is determined by larval-rearing substrates in cactophilic *Drosophila mojavensis*. V. Deep geographic variation in epicuticular hydrocarbons among isolated populations. The American Naturalist 158:585-598.

Etges, W. J. and C. C. de Oliveira. 2014. Premating isolation is determined by larval rearing substrates in cactophilic *Drosophila mojavensis*. X. Age-specific dynamics of adult epicuticular hydrocarbon expression in response to different host plants. Ecol Evol 4:2033-2045.

Etges, W. J. and L. L. Jackson. 2001. Epicuticular hydrocarbon variation in *Drosophila mojavensis* cluster species. Journal of Chemical Ecology 27:2125-2149.

Everaerts, C., J. P. Farine, M. Cobb, and J. F. Ferveur. 2010. *Drosophila* cuticular hydrocarbons revisited: Mating status alters cuticular profiles. Plos One 5.

Feldmeyer, B., D. Elsner, and S. Foitzik. 2014. Gene expression patterns associated with caste and reproductive status in ants: Worker‐specific genes are more derived than queen‐specific ones. Molecular ecology 23:151-161.

Ferveur, J.-F. 2005. Cuticular hydrocarbons: Their evolution and roles in *Drosophila* pheromonal communication. Behavior genetics 35:279.

Ferveur, J.-F. and M. Cobb. 2010. Behavioral and evolutionary roles of cuticular hydrocarbons in Diptera. Insect hydrocarbons: biology, biochemistry, and chemical ecology:325-343.

Ferveur, J. F. and G. Sureau. 1996. Simultaneous influence on male courtship of stimulatory and inhibitory pheromones produced by live sex-mosaic *Drosophila melanogaster*. P Roy Soc B-Biol Sci 263:967-973.

Foley, B. and M. Telonis-Scott. 2011. Quantitative genetic analysis suggests causal association between cuticular hydrocarbon composition and desiccation survival in *Drosophila melanogaster*. Heredity 106:68.

Frentiu, F. D. and S. F. Chenoweth. 2010. Clines in cuticular hydrocarbons in two *Drosophila* species with independent population histories. Evolution 64:1784-1794.

Galindo‐Prieto, B., L. Eriksson, and J. Trygg. 2014. Variable influence on projection (VIP) for orthogonal projections to latent structures (OPLS). Journal of Chemometrics 28:623-632.

Gilburn, A., S. Foster, and T. Day. 1992. Female mating preference for large size in *Coelopa frigida* (seaweed fly). Heredity 69:209-216.

Gilburn, A., S. Foster, and T. Day. 1993. Genetic correlation between a female mating preference and the preferred male character in seaweed flies (*Coelopa frigida*). Evolution 47:1788-1795.

Gomes, C. C. G., J. R. Trigo, and Á. E. Eiras. 2008. Sex pheromone of the American warble fly, *Dermatobia hominis*: The role of cuticular hydrocarbons. Journal of chemical ecology 34:636.

Griffin, C., N. Day, H. Rosenquist, M. Wellenreuther, N. Bunnefeld, and A. Gilburn. 2018. Tidal range and recovery from the impacts mechanical beach grooming. Ocean and Coastal Management 154:66-71.

Grun, B., I. Kosmidis, and A. Zeileis. 2012. Extended beta regression in R: Shaken, stirred, mixed, and partitioned. J Stat Softw 48:1-25.

Haverty, M. I., L. J. Nelson, and M. Page. 1990. Cuticular hydrocarbons of four populations of *Coptotermes formosanus* Shiraki in the united states. Journal of chemical ecology 16:1635-1647.

Howard, R. W. and G. J. Blomquist. 2005. Ecological, behavioral, and biochemical aspects of insect hydrocarbons. Annu. Rev. Entomol. 50:371-393.

Howard, R. W., C. D. Howard, and S. Colquhoun. 1995. Ontogenetic and environmentally induced changes in cuticular hydrocarbons of *Oryzaephilus*

surinamensis (Coleoptera: Cucujidae). Annals of the Entomological Society of America 88:485-495.

Jallon, J. M. and J. R. David. 1987. Variations in cuticular hydrocarbons among the eight species of the *Drosophila melanogaster* subgroup. Evolution 41:294-302.

Juarez, M. P., S. Ayala, and R. R. Brenner. 1996. Methyl-branched fatty acid biosynthesis in *Triatoma infestans*. Insect Biochem Molec 26:599-605.

Keays, M. C., D. Barker, C. WICKER‐THOMAS, and M. G. Ritchie. 2011. Signatures of selection and sex‐specific expression variation of a novel duplicate during the evolution of the *Drosophila* desaturase gene family. Molecular ecology 20:3617-3630.

Kulmuni, J., Y. Wurm, and P. Pamilo. 2013. Comparative genomics of chemosensory protein genes reveals rapid evolution and positive selection in ant-specific duplicates. Heredity 110:538-547.

Kwan, L. and H. D. Rundle. 2010. Adaptation to desiccation fails to generate pre- and postmating isolation in replicate *Drosophila melanogaster* laboratory populations. Evolution 64:710-723.

Li, H. and R. Durbin. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589-595.

Liang, D. and J. Silverman. 2000. "You are what you eat": diet modifies cuticular hydrocarbons and nestmate recognition in the Argentine ant, *Linepithema humile*. Naturwissenschaften 87:412-416.

Mcalpine, D. K. 1991. Review of the Australian kelp flies (Diptera, Coelopidae). Syst Entomol 16:29-84.

Ottensmann, M., M. A. Stoffel, H. J. Nichols, and J. I. Hoffman. 2018. GCalignR: An R package for aligning gas-chromatography data for ecological and evolutionary studies. Plos One 13.

Perdereau, E., F. Dedeine, J.-P. Christidès, and A.-G. Bagnères. 2010. Variations in worker cuticular hydrocarbons and soldier isoprenoid defensive secretions within and among introduced and native populations of the subterranean termite, *Reticulitermes flavipes*. Journal of chemical ecology 36:1189-1198.

Peterson, M. A., S. Dobler, E. L. Larson, D. Juárez, T. Schlarbaum, K. J. Monsen, and W. Francke. 2007. Profiles of cuticular hydrocarbons mediate male mate choice and sexual isolation between hybridising *Chrysochus* (Coleoptera: Chrysomelidae). Chemoecology 17:87-96.

Qiu, Y., C. Tittiger, C. Wicker-Thomas, G. Le Goff, S. Young, E. Wajnberg, T. Fricaux, N. Taquet, G. J. Blomquist, and R. Feyereisen. 2012. An insect-specific P450 oxidative decarbonylase for cuticular hydrocarbon biosynthesis. P Natl Acad Sci USA 109:14858-14863.

Rajpurohit, S., R. Hanus, V. Vrkoslav, E. L. Behrman, A. O. Bergland, D. Petrov, J. Cvačka, and P. S. Schmidt. 2017. Adaptive dynamics of cuticular hydrocarbons in *Drosophila*. Journal of evolutionary biology 30:66-80.

Reed, J. R., D. R. Quilici, G. J. Blomquist, and R. C. Reitz. 1995. Proposed mechanism for the cytochrome P450-catalyzed conversion of aldehydes to hydrocarbons in the house fly, *Musca domestica*. Biochemistry 34:16221-16227.

Rouault, J. D., C. Marican, C. Wicker-Thomas, and J. M. Jallon. 2004. Relations between cuticular hydrocarbon (HC) polymorphism, resistance against desiccation and breeding temperature; A model for HC evolution in *D. melanogaster* and *D. simulans*. Genetica 120:195-212.

Rundle, H. D., S. F. Chenoweth, P. Doughty, and M. W. Blows. 2005. Divergent selection and the evolution of signal traits and mating preferences. Plos Biology 3:1988-1995.

Savarit, F. and J. F. Ferveur. 2002. Temperature affects the ontogeny of sexually dimorphic cuticular hydrocarbons in *Drosophila melanogaster*. Journal of Experimental Biology 205:3241-3249.

Schaner, A. M., R. J. Bartelt, and L. L. Jackson. 1987. (Z)-11-octadecenyl acetate, an aggregation pheromone in *Drosophila simulans*. Journal of chemical ecology 13:1777-1786.

Shirangi, T. R., H. D. Dufour, T. M. Williams, and S. B. Carroll. 2009. Rapid evolution of sex pheromone-producing enzyme expression in *Drosophila*. PLoS biology 7:e1000168.

Steiger, S., G. D. Ower, J. Stökl, C. Mitchell, J. Hunt, and S. K. Sakaluk. 2013. Sexual selection on cuticular hydrocarbons of male sagebrush crickets in the wild. Proceedings of the Royal Society B: Biological Sciences 280:20132353.

Stennett, M. D. and W. J. Etges. 1997. Premating isolation Is determined by larval rearing substrates in cactophilic *Drosophila mojavensis*. III. Epicuticular hydrocarbon variation is determined by use of different host plants in *Drosophila mojavensis* and *Drosophila arizonae*. Journal of Chemical Ecology 23:2803-2824.

Stojkovic, B., U. Savkovic, M. Dordevic, and N. Tucic. 2014. Host-shift effects on mating behavior and incipient pre-mating isolation in seed beetle. Behav Ecol 25:553-564.

Thevenot, E. A., A. Roux, Y. Xu, E. Ezan, and C. Junot. 2015. Analysis of the human adult urinary metabolome variations with age, body mass index, and gender by implementing a comprehensive workflow for univariate and OPLS statistical analyses. J Proteome Res 14:3322-3335.

Thomas, M. L. and L. W. Simmons. 2009. Sexual selection on cuticular hydrocarbons in the Australian field cricket, *Teleogryllus oceanicus*. BMC Evolutionary Biology 9:162.

Toolson, E. C. and R. Kupersimbron. 1989. Laboratory evolution of epicuticular hydrocarbon composition and cuticular permeability in *Drosophila pseudoobscura* - Effects on sexual dimorphism and thermal-acclimation ability. Evolution 43:468-473.

Van Oystaeyen, A., R. C. Oliveira, L. Holman, J. S. van Zweden, C. Romero, C. A. Oi, P. d'Ettorre, M. Khalesi, J. Billen, and F. Wäckers. 2014. Conserved class of queen pheromones stops social insect workers from reproducing. Science 343:287-290.

Wang, L. M. and D. J. Anderson. 2010. Identification of an aggression-promoting pheromone and its receptor neurons in *Drosophila*. Nature 463:227-U114.

Weir, B. S. and C. C. Cockerham. 1984. Estimating F-Statistics for the analysis of population structure. Evolution 38:1358-1370.

Wellenreuther, M., H. Rosenquist, P. Jaksons, and W. Larson. 2017. Local adaptation along an environmental cline in a species with an inversion polymorphism. Journal of Evolutionary Biology 30:1068-1077.

Wigglesworth, V. 1945. Transpiration through the cuticle of insects. Journal of Experimental Biology 21:97-114.

Wu, T. D. and C. K. Watanabe. 2005. GMAP: A genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics 21:1859-1875.

Figures

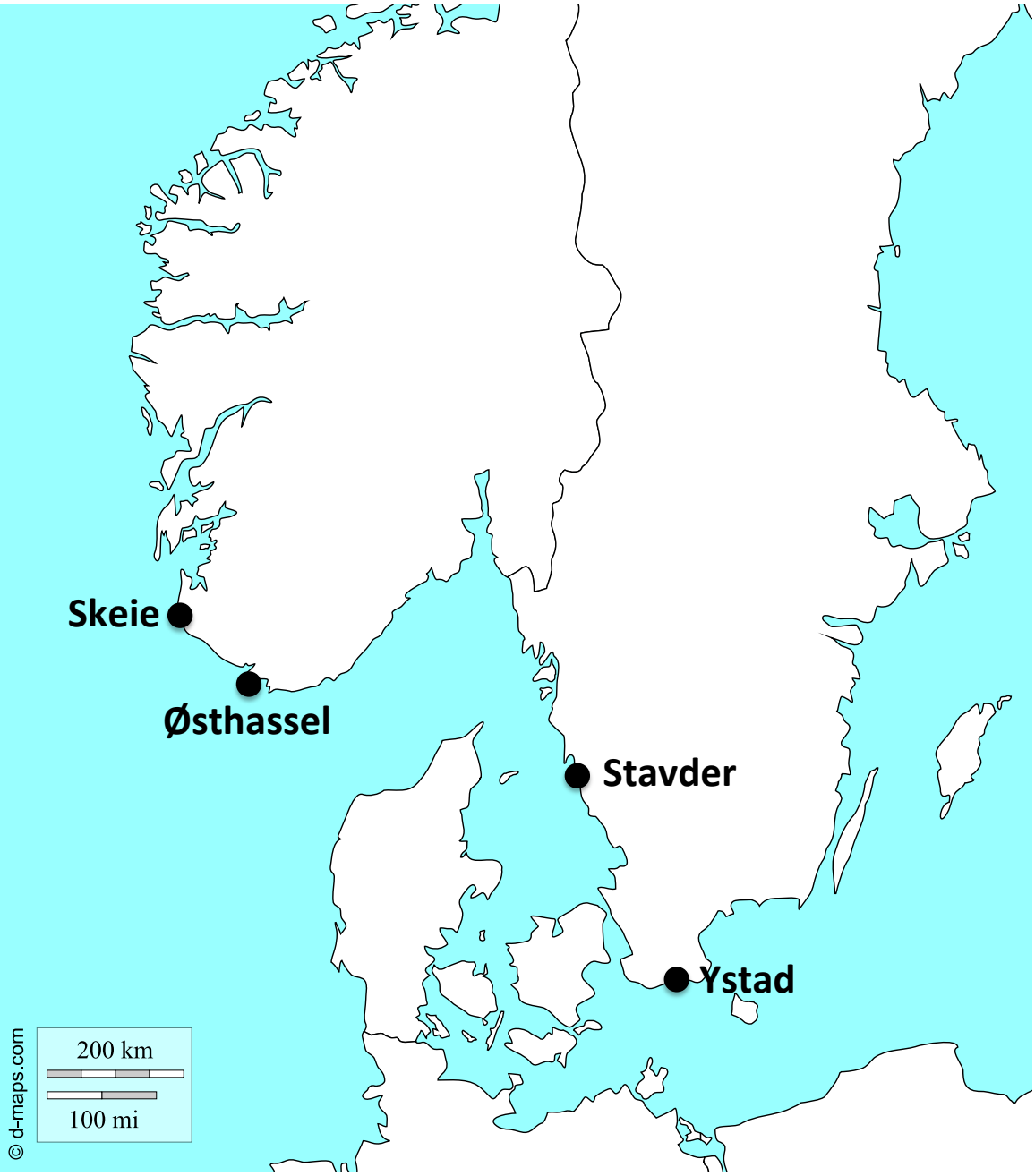

Figure 1 - Map of *C. frigida* populations used in this study. Base map from d-maps (http://d-maps.com/carte.php?num_car=5972)

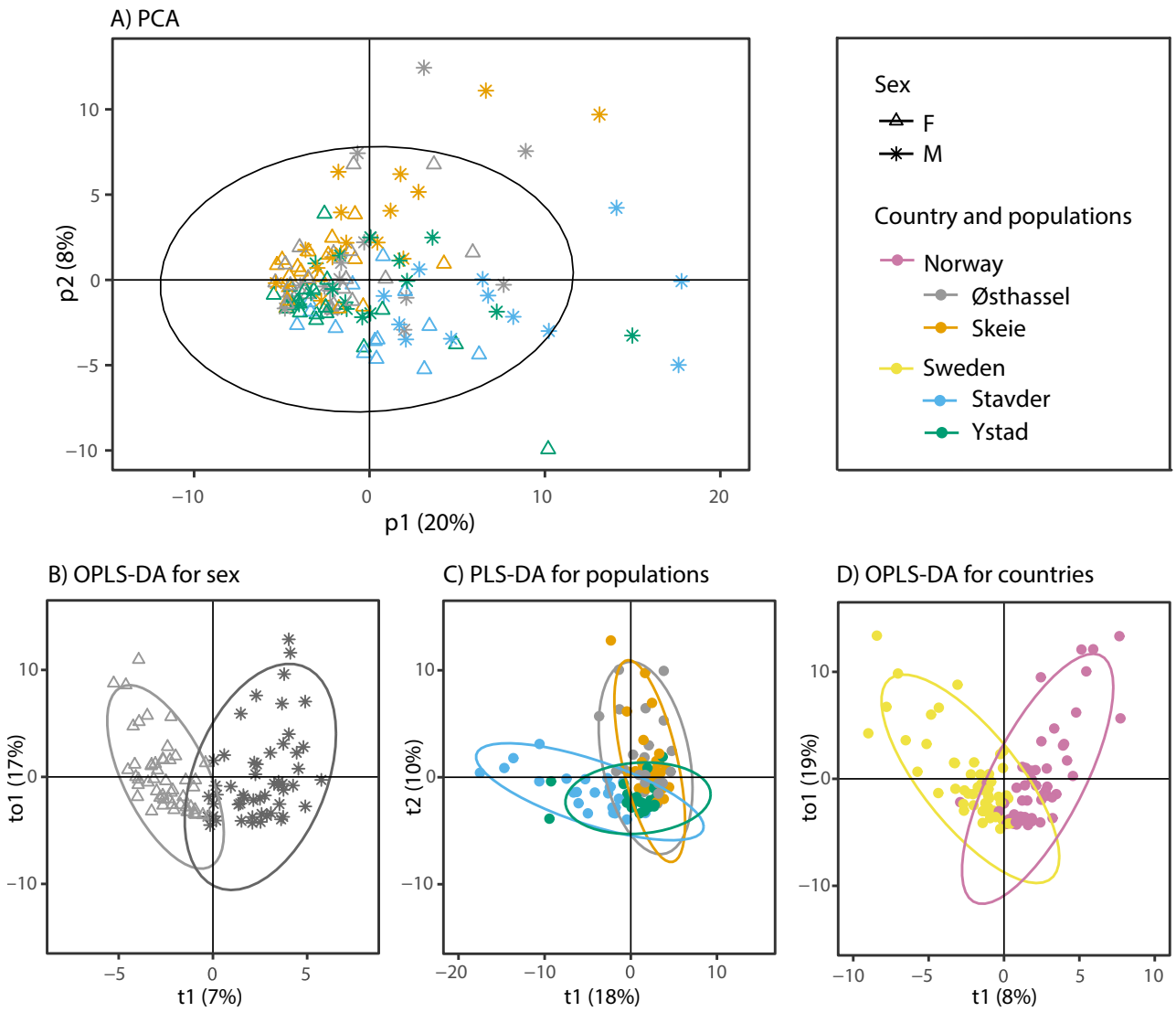

Figure 2 – A. PCA of all samples from the balanced data set indicating population (grey – Østhassel, orange – Skeie, blue – Stavder, green – Ystad) and sex (triangle – female, star – male). B. OPLS-DA for sex using all samples from the balanced data set, C. PLS-DA for population using all samples from the balanced data set, and D. OPLS-DA for country using all samples from the balanced data set (yellow – Sweden, purple – Norway). Circles indicate 95% confidence regions.

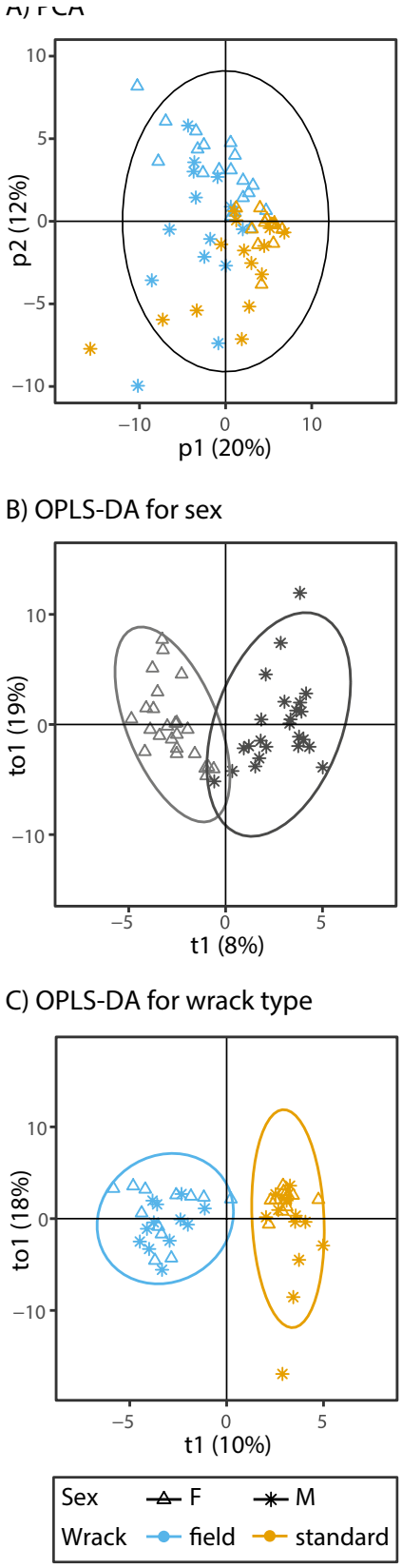

Figure 3 – A. PCA of samples from Ystad indicating diet (blue – field, yellow – standard) and sex (triangle – female, star – male). B. OPLS-DA for sex using only Ystad samples. C. OPLS-DA for diet using only Ystad samples. Circles indicate 95% confidence regions.

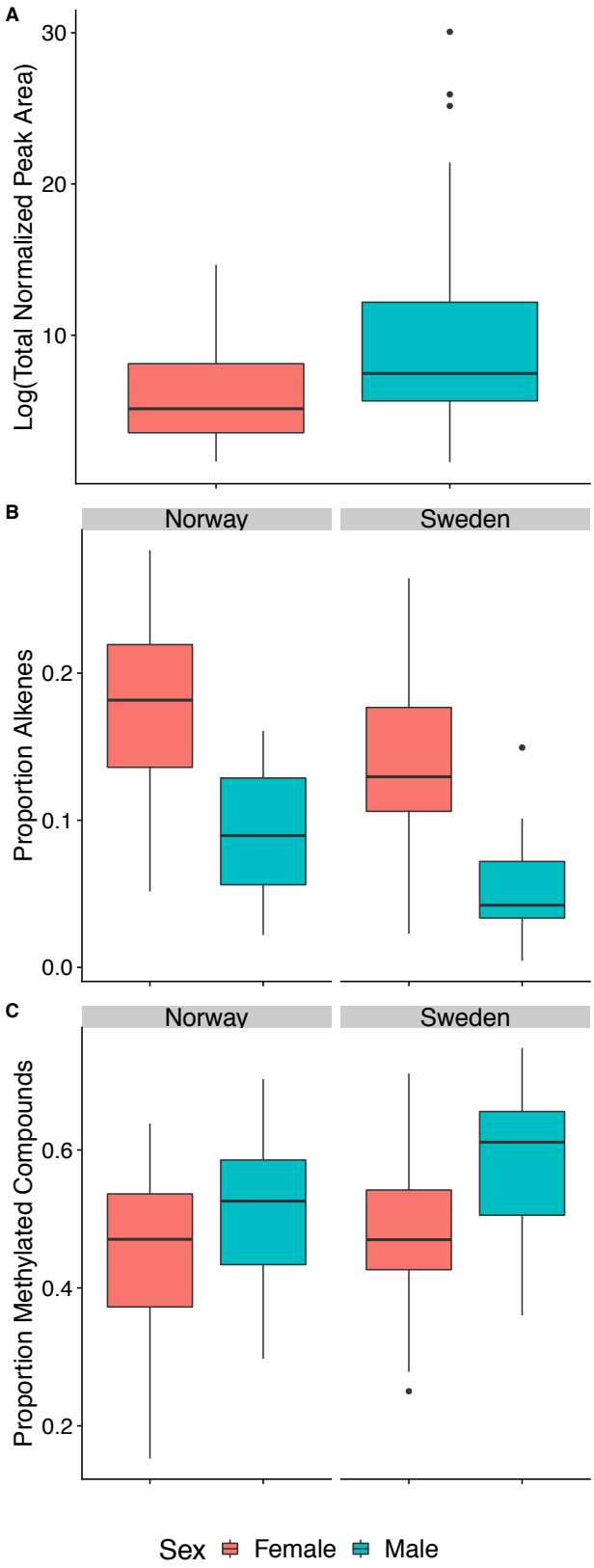

Figure 4 – Boxplot showing differences in A. Log transformed total peak area, B. proportion of compounds that are alkenes, C. proportion of methylated compounds. Females are in red, males are in blue.

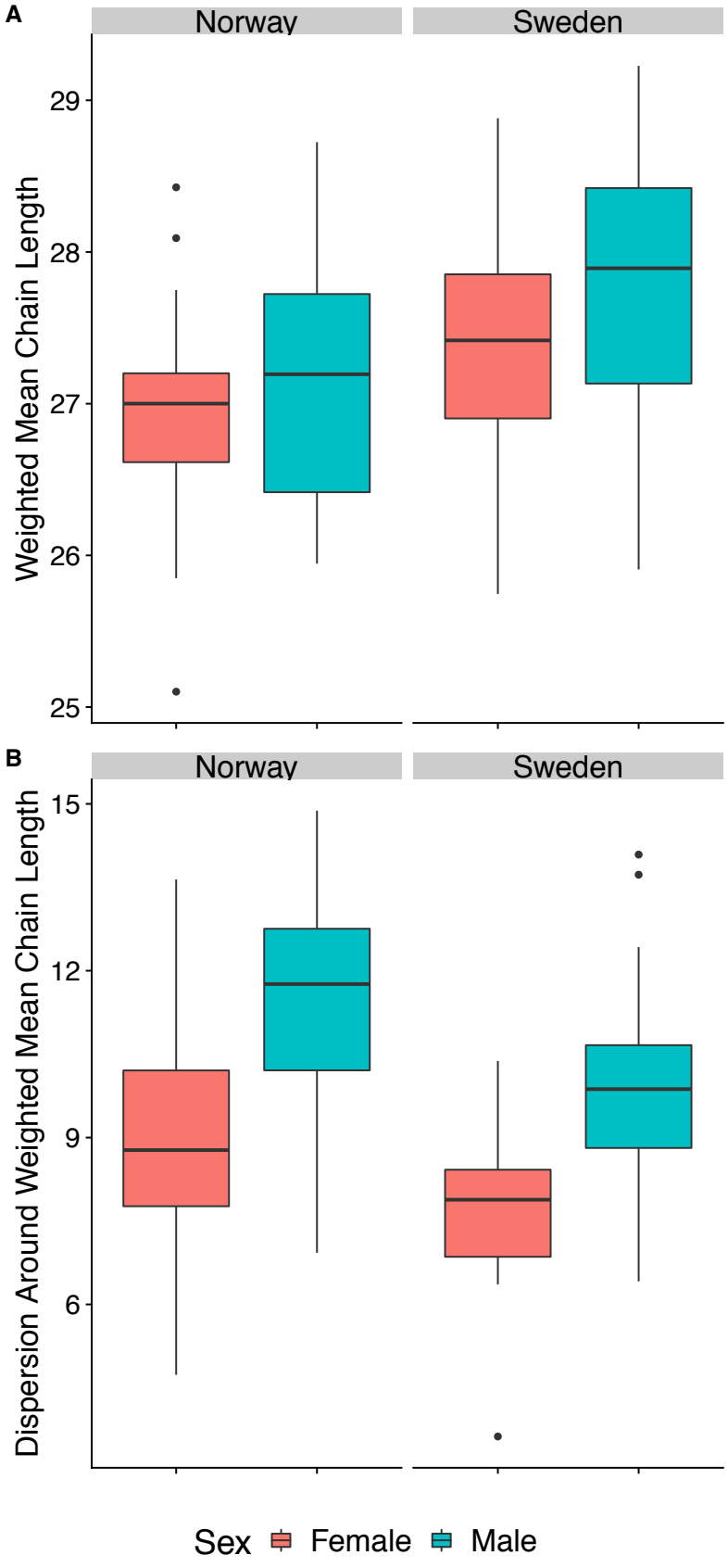

Figure 5 – Boxplot showing differences in A. Weighted mean chain length and B. Dispersion around the mean. Females are in red, males are in blue.