G3: Genes|Genomes|Genetics Early Online, published on October 12, 2016 as doi:10.1534/g3.116.033092

A linkage map and QTL analysis for pyrethroid resistance in the bed bug *Cimex lectularius*

Fountain, T[1, *, †], Ravinet, M[2, ‡], Naylor, R[3], Reinhardt, K[4] & Butlin, R.K [3, **]

Authors in bold contributed equally

Present addresses:

[1]Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden.

[2]Centre for Ecological and Evolutionary Synthesis, University of Oslo, Oslo, Norway.

[3]Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK.

[4]Department of Biology, Applied Zoology, Technische Universitaet Dresden, Dresden, Germany

Other affiliations:

[*]Department of Biosciences, University of Helsinki, Helsinki, Finland

[†]Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK.

[‡]Ecological Genetics Division, National Institute of Genetics, Mishima, Japan.

[**]Department of Marine Sciences, University of Gothenburg, Gothenburg, Sweden.

Running head: QTL for bed bug pesticide resistance

Keywords: ectoparasite, household pest, insecticide, RAD-seq

Corresponding author: Mark Ravinet, Centre for Ecological and Evolutionary Synthesis, Department of Biosciences, University of Oslo, PO Box 1066 Blinder, NO-0316, Oslo, Norway, +47 921 556 81, mark.ravinet@ibv.uio.no

Abstract

The rapid evolution of insecticide resistance remains one of the biggest challenges in the control of medically and economically important pests. Insects have evolved a diverse range of mechanisms to reduce the efficacy of the commonly used classes of insecticides and finding the genetic basis of resistance is a major aid to management. In a previously unstudied population, we performed an F_2 resistance mapping cross for the common bed bug, *Cimex lectularius*, for which insecticide resistance is increasingly widespread. Using 334 SNP markers obtained through RAD-sequencing, we constructed the first linkage map for the species, consisting of 14 putative linkage groups (LG), with a length of 407 cM and an average marker spacing of 1.3 cM. The linkage map was used to reassemble the recently published reference genome, facilitating refinement and validation of the current genome assembly. We detected a major QTL on LG12 associated with insecticide resistance, occurring in close proximity (1.2 Mb) to a carboxylesterase encoding candidate gene for pyrethroid resistance. This provides another example of this candidate gene playing a major role in determining survival in a bed bug population following pesticide resistance evolution. The recent availability of the bed bug genome, complete with a full list of potential candidate genes related to insecticide resistance, in addition to the linkage map generated here, provides an excellent resource for future research on the development and spread of insecticide resistance in this resurging pest species.

Introduction

The common bed bug, *Cimex lectularius* L. (Heteroptera, Cimicidae), is re-emerging as a significant economic and public health pest, precipitated by a recent global resurgence in populations (Boase 2001; Doggett & Russell 2008; Potter *et al.* 2008;

Richards *et al.* 2009). Much of its recent success has been attributed to widespread resistance to insecticides (Romero *et al.* 2007; Romero & Anderson 2016) making pest control increasingly challenging and costly (Koganemaru & Miller 2013). Developing a more detailed understanding of the genetic and molecular basis of insecticide resistance is therefore of clear importance.

Previously, two point mutations, V419L and L925I, have been identified in the α-subunit gene of the voltage sensitive sodium channel (VSSC) that are functionally associated with resistance to the pyrethroid deltamethrin (Yoon *et al.* 2008). Pyrethroids are one of the most widely used insecticides, but as over 80% of sampled populations in the USA (Zhu *et al.* 2010) and >95% of sampled populations in Europe (Booth *et al.* 2015) contained the V419L and/or L925I mutation(s), it is likely that target-site-based pyrethroid resistance has become widespread. In addition, several candidate loci associated with metabolic and penetrative resistance have been identified in studies comparing resistant and non-resistant populations, with increased expression of genes coding for detoxifying metabolic enzymes (including P450s, glutathione-S-transferases and carboxylesterases), ATP-binding cassette (ABC) transporters and cuticular protein genes associated with pyrethroid resistance (Adelman *et al.* 2011; Mamidala *et al.* 2011; 2012; Zhu *et al.* 2012; Koganemaru *et al.* 2013).

The recent availability of the bed bug genome (Benoit *et al.* 2016; Rosenfeld *et al.* 2016) gives an ideal opportunity to further investigate the genetic basis of resistance. For example 58 genes and one pseudogene coding for P450 enzymes have been identified in the *C. lectularius* genome (Benoit *et al.* 2016), with four of these genes previously implicated in pyrethroid resistance (Zhu *et al.* 2012). The further

identification of genes coding for other metabolic enzymes, cuticular protein genes, and ABC transporters, allows the assessment of their contribution to resistance.

Although these genetic-association and genome annotation studies have pointed to a promising group of candidate genes for pyrethroid resistance, their correlative top-down approaches lack the ability to demonstrate a direct association between any of these genes and the resistance trait. In addition, most of these studies used only one susceptible strain. Here, we perform an F_2 mapping cross between a pyrethroid resistant and a susceptible bed bug population using RAD-sequencing. Our reduced-representation sequencing approach offers two advantages. Firstly we are able to reassemble >65% of the bed bug reference genome into 14 linkage groups: a valuable resource for the community in future genome-based applications. Secondly, we are able to identify a new QTL associated with pyrethroid resistance that strongly implicates a functional role for a carboxylesterase encoding gene in this resistance trait.

Materials and methods

Experimental cross design and phenotyping

An F_2 mapping cross was established through mating a pyrethroid resistant female from a field population, originating from London, UK, with a pyrethroid susceptible male from a lab stock population, originating from a culture from the London School of Hygiene and Tropical Medicine (more information on these populations, called Field UK and Lab Stock, is available in Fountain *et al.* (2015)). The field population was checked for the resistance phenotype prior to the crossing to ensure resistance had not been lost. Our experimental design for QTL analysis with a single family assumes that the grandparents used to initiate the cross were homozygous for QTL involved in

insecticide resistance and for loci linked to these genomic regions. Because the lines were not highly inbred, this may not have been the case for all loci. However, since natural populations tend to have low heterozygosity (Fountain *et al.* 2014) and the lines had been maintained in the lab for multiple generations, it is likely that they were homozygous at resistance loci and for the great majority of markers. One male and one female from the F_1 offspring were selected at random and mated. 90 F_2 offspring, along with the F_1 parents were subsequently phenotyped for pyrethroid resistance.

Pyrethroid resistance was tested using 40 mg per m^2 of alpha-cypermethrin (Sigma number: 45806-100MG). The insecticide was dissolved in acetone and pipetted onto Whatman 90mm Grade 1 cellulose filter paper (Sigma number - Z240079). Once the filter paper was dry, it was placed in a 90mm diameter sterile polystyrene Petri dish. Individuals were added in groups (no more than 10 individuals per trial) and knock-down/mortality was scored at 24 and 48 hours. Phenotyping was performed at 26+/-1°C and 70+/-5% relative humidity, with the knockdown/mortality score at 48 hours after exposure used as the resistance phenotype. Individuals were scored as susceptible (knocked down, unable to right themselves if turned over), partially resistant (able to right themselves, but walk with some difficulty), or resistant (walk normally, motor-control apparently unaffected).

DNA isolation and sequencing

Full body extractions (minus the head) were performed using DNAeasy Blood & Tissue Kit (Qiagen). RAD library preparation was performed as in Baird *et al.* (2008), using *Sbf*1. Following library preparation, sequencing was performed on a single Illumina HiSeq lane (100bp PE) at the Natural Environment Research Council Biomolecular Analysis Facility at the University of Edinburgh, UK.

155

156 *Quality filtering and reference mapping*

157 Following sequencing, library quality was checked using FASTQC (Babraham
158 Informatics; http://www.bioinformatics.babraham.ac.uk/projects/fastqc). All
159 downstream handling of sequencing data, with the exception of mapping to the
160 reference genome, was conducted using the Stacks v (1.35) pipeline (Catchen *et al.*,
161 2011, 2013). Based on the average quality scores per read generated by FASTQC, the
162 Stacks *process_radtags* module was used to remove any read where Phred quality
163 scores fell below 15 (i.e. 3% error rate) in a 5 bp window. The module was
164 additionally used to remove reads with traces of adapter sequence, remove any reads
165 with an uncalled base and demultiplex the pooled libraries. Following this initial
166 processing, the *clone_filter* module was used on the paired-end sequence data for each
167 individual in order to remove PCR duplicates, a major source of potential bias for
168 RAD-sequencing approaches (Andrews *et al.*, 2014).

169 Paired-end sequence reads filtered for duplicates were then aligned to one of
170 the recently published *Cimex lectularius* genomes (Clec_1.0; NCBI Accession number:
171 PRJNA167477; Benoit *et al.* 2016) using GSNAP 2014-12-29 (Wu & Nacu, 2010).
172 We allowed a maximum of 10 alignments per read, no terminal alignments and only
173 the optimal hit to be reported. A maximum of four single nucleotide mismatches was
174 allowed for each alignment.

175

176 *Stacks catalogue construction and SNP calling*

177 Aligned read data were processed using the reference-mapping branch of the Stacks
178 pipeline (*ref_map.pl*), specifying an F_2 cross and identifying both parents and offspring
179 using the –p and –s flags, respectively. We allowed a minimum of three reads to form a

stack in the *pstacks* module (i.e. minimum read depth for an allele, not a locus) and a single mismatch amongst loci during catalogue construction. Stacks construction was conducted with these values following sensitivity testing with both *de novo* and reference mapping pipelines, revealing them to be the optimal parameters (supplementary materials). SNPs were called using the Stacks default SNP calling method, i.e. maximum likelihood estimation based on a multinomial probability distribution derived from the nucleotide frequency at each read position (Catchen *et al.*, 2013). Following initial catalogue construction, we used the *rxstacks* module to reanalyse and correct the *de novo* assembly. We filtered confounded loci (i.e. loci within individuals matching multiple catalogue loci, indicative of repetitive regions), pruned excess haplotypes (i.e. removed potential erroneous haplotypes based on frequency), and recalled SNPs using Stacks' bounded error model with $\varepsilon = 0.1$. Following catalogue correction, the *genotypes* module was used to export SNP data from the main catalogue. We used the module to perform automatic corrections to the data and we only exported markers where at least eight F_2 progeny had genotype calls. To clarify, this cutoff did not represent our final threshold for missing data but was chosen to maximise the output loci for downstream filtering. Output genotype calls were filtered to include only loci that were heterozygote in the F_1 parents (*i.e.* AB/AB where A and B are alleles from the female and male grandparents respectively) then converted to R/qtl format using a custom R script (R Development Core Team, 2012).

Genetic map construction and QTL mapping

To perform QTL analysis, we used the R/qtl package (Broman *et al.*, 2003). Our first step was to perform additional data screening following the best-practice guidelines outlined on the R/qtl website (http://www.rqtl.org/tutorials/geneticmaps.pdf). We

removed all individuals with genotypes for fewer than 50% of markers and all markers with genotypes for fewer than 50% of individuals. We additionally screened for uninformative markers with duplicate information and any markers showing extreme segregation distortion (i.e. being nearly monomorphic). We then estimated a genetic map using the *est.rf* function. Following previously published information on *C. lectularius* karyotype (Sadílek *et al.* 2013), showing an average of 14 autosome pairs and one X chromosome, we varied the maximum recombination fraction and the minimum LOD score (i.e. 'logarithm of the odds score' – a $\log_{10}$ transformation of the likelihood ratio between a model with linkage and a null model) threshold in order to create approximately the same number of linkage groups as autosomes. We then checked our initial genetic map following the R/qtl guidelines (http://www.rqtl.org/tutorials/geneticmaps.pdf) and removed any problematic markers before reordering markers based on likelihood analysis of permuted marker orders using the *ripple* function. The R script used to produce our genetic map is available at Dryad (http://dx.doi.org/10.5061/dryad.d4r50).

Following map construction, we performed standard interval mapping with a single QTL model for pesticide resistance using R/qtl. In order to account for genotyping error in our QTL analysis, we applied a maximum likelihood-based estimate of error rate. QTL genotype probabilities were then calculated using a Kosambi mapping function and an error probability based on our maximum-likelihood estimation. We then used the *scanone* function to estimate QTL LOD scores using both the EM algorithim and Haley-Knott regression. To test the significance of our QTL and to estimate confidence limits on QTL positions, we reanalysed our dataset with *scanone* using 1000 permutations.

Testing for sex-linkage

A limitation of our F_2 mapping approach was that it did not allow for mapping of putatively sex-linked loci. Furthermore, as F_2 individuals were phenotyped and processed for DNA extraction as 4th instar nymphs, we were unable to determine their sex. In order to account for potential sex-linkage in our RAD dataset, we identified loci that were heterozygous in the female F_1 parent and homozygous for the grandmother's allele in the male F_1 (i.e. AB/AA). Our rationale for this was that, assuming no error, a cross using AA x BB grandparents should only result in homozygous genotypes for loci that occur on the sex chromosome in the heterogametic sex. To rule out error, we focused only on loci with an AA genotype in the female grandparent and with greater than 50% of individuals genotyped. Using this set of putatively sex-linked loci, we then performed a Chi-squared test of independence to test for an association with the resistance phenotype. False Discovery Rate (FDR) correction was used to account for multiple testing; since many loci are not independent (i.e. multiple loci map to the same scaffold), we used the number of scaffolds and the minimum *P*-value for each scaffold to perform this correction.

Identifying candidate genes in functional regions

RAD-seq loci are typically short (i.e. approximately 100 bp) and, since they only sample a relatively small proportion of the genome, are unlikely to occur within resistance genes themselves. Similarly, short consensus RAD loci are unlikely to be of much use in identifying candidate genes using a functional analysis such as BLAST. To identify candidate genes associated with QTL regions, we first used the calculated 95% Bayesian credible intervals around the QTL. Using the markers flanking the interval, we then located the corresponding physical position in the reference genome

and identified all candidate genes within this interval. We also searched for genes associated with pyrethroid resistance on the same scaffold as our identified QTL. Genes were identified from recently published annotations (Benoit *et al.* 2016) and extracted using custom R scripts.

Genome reassembly

In order to combine our genetic map with the recently published *C. lectularius* reference genome, we used Chromonomer (Amores *et al.* 2014). Chromonomer first removes markers that are inconsistent with local assembly order on the genetic map and then anchors genome scaffolds to linkage groups based on marker mapping position before finally reassembling the genome accordingly. Chromonomer was run using the default settings as described in the online manual (http://catchenlab.life.illinois.edu/chromonomer/manual/).

Data availability

Raw RAD-sequencing reads are archived at EMBL-ENA (PRJEB15267 - see Table S1 for individual accession numbers). All bash scripts for alignment, filtering, trimming and Stacks catalogue construction are archived on Dryad (http://dx.doi.org/10.5061/dryad.d4r50). All R scripts for R/qtl analysis are also archived on Dryad. The reassembled genome is archived in NCBI Genbank (GCA_000648675.2) and is hosted at https://i5k.nal.usda.gov/Cimex_lectularius.

Results

RAD sequence mapping and Stacks catalogue construction

279 Following filtering for quality and PCR duplicates, an average of 240 898 ± 150 591
280 (mean ± SD) reads was retained for each individual (see Table S1). A high proportion
281 of these reads (78.0% ± 5.61) mapped to the reference genome (see Table S1). Initial
282 RAD locus catalogue construction resulted in 12 992 unique RAD loci, which was
283 reduced to 12 962 tags following *rxstacks* correction. Of these corrected RAD loci,
284 1171 occurred in greater than 8 of the F_2 progeny; 430 of these loci were heterozygous
285 in both F_1 parents and were subsequently included for genetic map construction.
286
287 *Genetic map construction*
288 Prior to map construction, we filtered individuals with a high proportion of missing
289 markers, markers missing in a high proportion of individuals (both >50%), duplicate
290 markers (*i.e.* likely originating from either side of the same RAD locus) and markers
291 with a highly distorted segregation ratio. This resulted in a reduced dataset of 75
292 individuals and 357 high quality markers.
293 Initial recombination fraction estimates were strongly correlated with high
294 LOD scores (Fig S1). To account for this, we merged markers into 31 linkage groups;
295 i.e. approximately two linkage groups per chromosome (assuming $n = 15$),
296 representing correct and potentially misidentified alleles. Linkage groups were
297 combined based on high LOD scores but low recombination fractions among markers
298 (Fig S2). Following additional filtering, allele correction and removal of loci with
299 apparent genotyping errors and/or extreme segregation distortion, we re-estimated
300 linkage groups to ensure high LOD and low recombination fractions amongst markers
301 on the same chromosome (Fig S3). Our final map, based on 71 individuals after
302 filtering, was 407 cM long with an average spacing of 1.3 cM between each of the 334
303 markers and consisted of 14 linkage groups (Fig 1A, Fig S3, Table 1). Scaffold

positions of all mapped markers are given in Table S2. Given the genome size of 650.5 Mb, this implies an average recombination rate of 0.6 cM/Mb.

QTL analysis

Maximum-likelihood estimation indicated genotyping error rate was 0.0025, suggesting such error was not an issue in our filtered mapping dataset (Fig S4). Per-locus estimates of error rates suggested few consistent errors across loci, therefore QTL analysis was first conducted without manual genotype correction. QTL scans for pesticide resistance using both the Hayley-Knott and EM algorithms revealed a clear signal of a single QTL towards the end of linkage group 12, centred on RAD locus r449_NW_014465016 (LOD = 6.84, P = < 0.0005 based on 10000 permutations; see Fig 1B). No genotyping errors were present on this linkage group and repeating this analysis with manual corrections produced identical results. Examining phenotype counts at this locus clearly showed that AA homozygotes showed complete pesticide susceptibility whereas 90% of BB homozygotes were resistant and 10% partially resistant (Table 2). Heterozygotes at this locus were mainly susceptible (66%), although some showed partial resistance (24%) and a minority showed full resistance (10%, see Table 2). The LG12 QTL explained 64.2% of the variation in phenotype indicating pesticide resistance is not completely explained by this bi-allelic QTL.

Sex-linkage

Identifying loci that were AB/AA in the F_1 cross and with an AA genotype in the grandmother resulted in 106 putatively sex-linked loci that were not included in our genetic map construction. Chi-squared tests for independence identified four out of the 106 putatively sex-linked loci occurring on different scaffolds that showed an

329 association with the resistance phenotype ($P < 0.005$; Table S3), but none of these
330 associations remained significant following FDR correction.
331
332 *Candidate gene identification*
333 Our pyrethroid resistance QTL maps to Scaffold 2 (start position = 15 333 169) of the
334 reference genome. The LOD peak surrounding this QTL on LG12 of our linkage map
335 spanned a total 95% Bayesian credible interval of 11.1 cM (i.e. from 10.8 – 21.9 cM).
336 This corresponds to approximately the last 6.8 Mb of scaffold 2 and the last 4 Mb of
337 scaffold 18 in the reference genome (*N.B.* our linkage map suggests a reverse
338 orientation for some scaffolds, including scaffold 18). Searching for coding sequences
339 within these scaffold intervals, we identified 211 unique gene names (Table S4). The
340 most likely candidate for pyrethroid resistance was an ubiquitin carboxyl-terminal
341 hydrolase 15-like gene occurring 1.2 Mb downstream from our QTL peak. However,
342 we also identified two other putative candidates occurring further downstream (i.e. > 3
343 Mb) on scaffold 2: a VSSC protein para gene and a glutathione S-transferase (see
344 Table 3). Additionally, we identified a cytochrome P450 6B5-like gene on Scaffold 18
345 (Table 3). Finally, three of the four putatively sex-linked RAD-loci that showed an
346 association with resistance mapped to scaffolds containing three further candidate gene
347 classes – a glutathione-S transferase (Scaffold 6), a cuticular protein gene cluster
348 (Scaffold 24) and P450 genes (Scaffold 31).
349
350 *Genome reassembly*
351 Chromonomer pruned 130 markers from our genetic map that were inconsistent with
352 local assembly order, using 208 well-behaved markers to perform genome reassembly.
353 Of the 1402 scaffolds in the previously published reference genome, 69 were anchored

354 to our linkage map, whereas five aligned to more than one position and were split –
355 resulting in a total of 74 anchored scaffolds (mean size: 7.5 Mb, range: 0.06 – 33 Mb).
356 Chromonomer was thus able to reassemble 67% of the genome into 14 autosomal
357 linkage groups spanning 433 Mb. The newly reassembled genome
358 (GCA_000648675.2) is available at https://i5k.nal.usda.gov/Cimex_lectularius.
359
360 **Discussion**
361 We have constructed the first linkage map for the common bed bug, *Cimex lectularius,*
362 by performing an F_2 insecticide resistance mapping cross using 334 high quality SNP
363 markers identified with RAD-sequencing. Our final linkage map consisted of 14
364 putative linkage groups (LG), and was 407 cM in length with an average marker
365 spacing of 1.3 cM. We successfully demonstrated the ability of the linkage map to
366 order scaffolds from the newly available bed bug genome by anchoring 74 scaffolds to
367 linkage map positions. Therefore, we were able to reassemble 67% of the draft genome
368 into our putative linkage groups, facilitating refinement and validation of the current
369 genome assembly. In addition to constructing a genetic map, we detected a bi-allelic
370 QTL on LG 12 that explains 64% of variation in pyrethroid resistance, in very close
371 proximity (1.2 Mb) to a carboxylesterase encoding candidate gene for pyrethroid
372 resistance (Zhu *et al*. 2013; Benoit *et al*. 2016) and less than 10 cM (but still within the
373 95% Bayesian credible interval) from the voltage-sensitive sodium channel (VSSC),
374 another candidate strongly associated with insecticide survival in other studies (Yoon
375 *et al.* 2008, Zhu *et al.* 2010).
376 Our construction of a genetic map for *C. lectularius* should be considered a first
377 attempt to assemble the recently published reference genome into clusters of markers
378 related by linkage. We stress that our map estimates only linkage groups and genetic

379 distance. Its relationship to the actual physical map of the *C. lectularius* genome
380 remains uncertain because a considerable proportion of the genome remains
381 unassembled into any linkage group (~33%). Importantly, this also includes the sex
382 chromosome which we were unable to map due to our cross design, although we did
383 identify putatively sex-linked scaffolds (S3).

384 Bed bugs, like other Cimicidae, have received attention for their unusual
385 cytogenetic characteristics (e.g. Darlington 1939; Slack 1939; Ueshima 1967; Grozeva
386 *et al.* 2010; 2011). For example, one study showed that populations appeared to be
387 geographically variable for their karyotype across Europe with 2n chromosome
388 number varying from 29 to 47, which was further complicated by fragmentation of sex
389 chromosomes in some populations (Sadílek *et al.* 2013). Since the grandparents from
390 our cross were not karyotyped, the expected number of chromosomes in our F_2
391 generation is unknown and may even be variable amongst individuals. Excluding sex
392 chromosomes and assuming 2n = 28 autosomes, we would expect to identify at least
393 14 linkage groups in our analysis. Therefore it seems likely that the majority of our
394 linkage groups correspond to physical autosomes. Due to our cross design, we were
395 unable to ascertain the sex of F_2 individuals, preventing us from including sex as a
396 mappable trait or from clearly identifying sex-linked loci. Furthermore, by only
397 including loci heterozygous in both F_1 parents (i.e. using an AB x AB cross), we were
398 also unable to identify linkage groups putatively associated with sex. However, using
399 an independent analysis outside of our linkage map construction, we were able to
400 identify a proportion of potentially sex-linked loci and by extension, genome scaffolds
401 which may anchor to the sex chromosome. Additional crosses are necessary to identify
402 genomic regions specifically involved in sex-determination. Nonetheless, further work,

403 such as FISH-based mapping is now possible, and necessary, to physically map our
404 inferred linkage groups to *C. lectularius* chromosomes.
405 Using QTL mapping, we identified a clear signal of a single bi-allelic QTL
406 related to pyrethroid resistance on LG12, with all AA genotypes completely
407 susceptible, all BB genotypes showing resistance or at least partial resistance, and 66%
408 of heterozygotes being susceptible. These proportions suggest our QTL is partly
409 recessive.
410 Given the large LOD peak confidence intervals, it is unclear whether the QTL
411 identified here represents the actions of a single gene or a complex of multiple co-
412 adapted genes for pyrethroid resistance. Importantly the resistance QTL occurs in close
413 proximity to several previously identified candidate genes for pyrethroid resistance.
414 This suggests that despite our relatively high-density approach using reduced-
415 representation sequencing, we did not have adequate resolution to identify the exact
416 candidate gene involved in insecticide resistance. Additional higher resolution QTL
417 mapping, using a combination of high and low coverage whole-genome re-sequencing
418 of larger families may allow more fine scale identification of the exact resistance QTL
419 in this context (Glazer *et al.* 2015). Despite this, our RAD-seq QTL analysis has
420 identified a region containing several important known candidate genes for pyrethroid
421 resistance.
422 The first and closest of these candidates, an ubiquitin carboxyl-terminal
423 hydrolase 15-like gene occurs just 1.5 Mb from our inferred QTL. Carboxylesterases
424 are a gene family coding for esterase enzymes that hydrolyse ester bonds present in a
425 wide variety of insecticides, including pyrethroids (Montella *et al.* 2012). More
426 efficient metabolic breakdown, resulting in a decrease in insecticide concentration
427 following exposure has previously been implicated as a means of pyrethroid resistance

428 in bed bugs (Zhu *et al.* 2013). Metabolic breakdown genes are likely to contribute to
429 insecticide resistance via at least one of three mechanisms; 1) gene duplication, 2)
430 increased gene expression or 3) mutation in the enzyme-coding sequence (Montella *et*
431 *al.* 2012). Gene annotation reveals at least 30 carboxylesterase genes in the *C.*
432 *lectularius* genome with clustering on some scaffolds (Benoit *et al.* 2016).
433 Furthermore, an expression analysis of geographically widespread bed bug populations
434 indicated overexpression of a carboxylesterase gene in resistant samples (Zhu *et al.*
435 2013). However, the cluster of carboxylesterase genes found on genome scaffold 18
436 does not map to linkage group 12. In addition, the carboxylesterase candidate gene
437 close to the LOD peak of our QTL differs from the overexpressed gene reported
438 previously. The LG12 QTL may therefore represent a gene that is not overexpressed,
439 *e.g.* a transcription factor involved in expression regulation of multiple
440 carboxylesterase genes. Additionally, our QTL explains approximately 64% of the
441 variation in resistance phenotypes, meaning that other genes may be involved. Further
442 investigation is required to examine whether coding mutations in the ubiquitin
443 carboxyl-terminal hydrolase 15-like gene on genome scaffold 2 may result in more
444 efficient metabolic breakdown of pyrethroid insecticides.
445 In addition to the ubiquitin carboxyl-terminal hydrolase 15-like gene, the QTL
446 is located upstream of two other major pyrethroid resistance candidates, the VSSC and
447 the metabolic detoxifying enzyme coding glutathione S-transferase gene. Knockdown
448 (*kdr*) resistance to pyrethroids is increasingly widespread in bed bugs (e.g. Zhu *et al.*
449 2010) with *kdr* mutations at the target site of pyrethroids, the VSSC, identified as
450 major mechanism for resistance (Yoon *et al.* 2008). However, there is increasing
451 evidence for a more complex basis for this trait, with penetrative (Mamidala *et al.*
452 2012; Zhu *et al.* 2013) and metabolic mechanisms (Zhu *et al.* 2012) as well as

453 behavioural avoidance (Romero *et al.* 2009) associated with pyrethroid resistance. This
454 is not a unique feature of bed bugs, with evidence of interactions between multiple
455 insecticide resistance mechanisms found in a number of medically and economically
456 important pests e.g. German cockroach (Anspaugh *et al.* 1994), cotton bollworm
457 (Martin *et al.* 2002), houseflies (Georghio, 1969, Sawicki, 1970, Shono *et al.* 2002)
458 and mosquitoes (Perera *et al.* 2008; Hardstone *et al.* 2009; Awolola *et al.* 2009). These
459 and our results, therefore, support the view that understanding the interaction between
460 resistance loci will be an important part of developing new resistance management
461 strategies (Hardstone *et al.* 2009). For example, epistatic interactions between
462 resistance loci (e.g., Bohannan *et al.* 1999) may reduce their costs (Gordon *et al.*
463 2015), facilitating the maintenance and spread of resistant alleles. Interestingly, the
464 QTL identified in the present study occurs in close proximity to genes coding for
465 detoxifying metabolic enzymes as well as the VSSC. Future work should focus on
466 identifying the casual mutation(s) underlying this QTL and how they interact with
467 previously identified resistance loci in bed bugs.
468 In addition to multiple resistance mechanisms, bed bug metapopulation
469 structure (Fountain *et al.* 2014) may further promote the spread of resistance alleles.
470 For example, if an insecticide resistant individual enters a (usually inbred - Fountain *et*
471 *al.* 2014, Booth *et al.* 2015) bed bug population, both heterosis (Fountain *et al.* 2015)
472 and the introduction of resistance alleles (Saccheri & Brakefield 2002; Song *et al.*
473 2011) may lead to the rapid recovery of a population and spread of resistance. Rapid
474 selection to environmental disturbance can be prevalent in metapopulations (Reznick
475 & Ghalambor 2001; Bell & Gonzalez 2011) and this may also have contributed to the
476 rapid spread of resistance mutations in bed bug populations both in the US (Zhu *et al.*
477 2010) and Europe (Booth *et al.* 2015).

To conclude, our mapping cross identified a QTL in close proximity to a number of candidate genes related to pyrethroid resistance and thereby provides strong evidence that these candidate genes play a major role in determining survival following pesticide treatment. Functional assays and higher resolution QTL approaches should now investigate the exact mechanism by which these genes convey resistance. The recent availability of the bed bug genome (Benoit *et al.* 2016; Rosenfeld *et al.* 2016) complete with a full list of potential candidate genes related to insecticide resistance, in addition to the linkage map generated here, will provide an excellent resource for future research on the development and spread of insecticide resistance in the bed bug.

Acknowledgements

RAD library preparation and sequencing were performed at the NERC Biomolecular Analysis Facility at the University of Edinburgh, UK, supported by the Natural Environment Research Council (NERC), UK (NBAF-724). Bioinformatics analyses were performed using the National Institute of Genetics Super Computer, Mishima, Japan. TF was funded by a NERC postgraduate studentship and support from the Royal Entomological Society. MR was funded by a JSPS Standard Postdoctoral Fellowship. We are grateful for the DDBJ and NIG Super Computer facility for allocating cluster resources for RAD-seq analysis. We thank Andy Krupa, Gavin Horsburgh, and Karim Gharbi for valuable advice concerning library preparation and sequencing.

References

Adelman ZN, Kilcullen KA, Koganemaru R *et al.* (2011) Deep sequencing of pyrethroid-resistant bed bugs reveals multiple mechanisms of resistance within a

single population. *PLoS ONE*, **6**, e26228.

Anspaugh DD, Rose RL, Koehler PG (1994) Multiple mechanisms of pyrethroid resistance in the German cockroach, *Blattella germanica* (L.). *Pesticide Biochemistry and Physiology,* **50**, 138–148.

Awolola TS, Oduola OA, Strode C *et al.* (2009) Evidence of multiple pyrethroid resistance mechanisms in the malaria vector *Anopheles gambiae* sensu stricto from Nigeria. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **103**, 1139–1145.

Bell G, Gonzalez A (2011) Adaptation and evolutionary rescue in metapopulations experiencing environmental deterioration. *Science*, **332**, 1327–1330.

Benoit JB, Adelman ZN, Reinhardt K *et al.* (2016) Unique features of a global human ectoparasite identified through sequencing of the bed bug genome. *Nature Communications*, **7**, 10165.

Boase C (2001) Bedbugs - back from the brink. *Pesticide Outlook*. **12**, 159-162.

Bohannan B, Travisano M, Lenski RE (1999) Epistatic interactions can lower the cost of resistance to multiple consumers. *Evolution*, **53**, 292–295.

Booth W, Balvín O, Vargo EL, Vilímová J, Schal C (2015) Host association drives genetic divergence in the bed bug, *Cimex lectularius*. *Molecular Ecology*, **24**, 980–992.

Darlington CD (1939) The genetical and mechanical properties of the sex chromosomes. *Journal of Genetics*, **39**, 101–137.

Doggett SL, Russell RC (2008) The resurgence of bed bugs, *Cimex* spp. (Hemiptera: Cimicidae) in Australia. *Proceedings of the Sixth International Conference on Urban Pests* (ed. by W.H. Robinson & D. Bajomi). pp. 407-425. OOK-Press, Veszprem.

528 Fountain T, Butlin RK, Reinhardt K, Otti O (2015) Outbreeding effects in an
529 inbreeding insect, *Cimex lectularius*. *Ecology and Evolution*, **5**, 409–418.
530 Fountain T, Duvaux L, Horsburgh G, Reinhardt K, Butlin RK (2014) Human-
531 facilitated metapopulation dynamics in an emerging pest species, *Cimex*
532 *lectularius*. *Molecular Ecology*, **23**, 1071–1084.
533 Georghiou GP (1969) Genetics of resistance to insecticides in houseflies and
534 mosquitoes. *Experimental Parasitology*, **26,** 224–255.
535 Glazer AM, Killingbeck EE, Mitros T (2015) Genome assembly improvement and
536 mapping convergently evolved skeletal traits in sticklebacks with Genotyping-by-
537 Sequencing. *G3: Genes| Genomes| Genetics,* **5***,* 1463-1472.
538 Gordon JR, Potter MF, Haynes KF (2015) Insecticide resistance in the bed bug comes
539 with a cost. *Scientific Reports*, **5**, 10807.
540 Grozeva S, Kuznetsova V, Anokhin B (2010) Bed bug cytogenetics: karyotype, sex
541 chromosome system, FISH mapping of 18S rDNA, and male meiosis in *Cimex*
542 *lectularius* Linnaeus, 1758 (Heteroptera: Cimicidae). *Comparative Cytogenetics*,
543 **4**, 151–160.
544 Grozeva S, Kuznetsova V, Anokhin B (2011) Karyotypes, male meiosis and
545 comparative FISH mapping of 18S ribosomal DNA and telomeric (TTAGG)n
546 repeat in eight species of true bugs (Hemiptera, Heteroptera). *Comparative*
547 *cytogenetics*, **5**, 355–374.
548 Hardstone MC, Leichter CA, Scott JG (2009) Multiplicative interaction between the
549 two major mechanisms of permethrin resistance, kdr and cytochrome P450-
550 monooxygenase detoxification, in mosquitoes. *Journal of Evolutionary Biology*,
551 **22**, 416–423.
552 Koganemaru R, Miller DM (2013) The bed bug problem: past, present, and future

553 control methods. *Pesticide Biochemistry and Physiology*, **106**, 177–189.
554 Koganemaru R, Miller DM, Adelman ZN (2013) Robust cuticular penetration
555 resistance in the common bed bug (*Cimex lectularius* L.) correlates with increased
556 steady-state transcript levels of CPR-type cuticle protein genes. *Pesticide*
557 *Biochemistry and Physiology*, **106**, 190–197.
558 Mamidala P, Jones SC, Mittapalli O (2011) Metabolic resistance in bed bugs. *Insects*,
559 **2**, 36–48.
560 Mamidala P, Wijeratne AJ, Wijeratne S *et al.* (2012) RNA-Seq and molecular docking
561 reveal multi-level pesticide resistance in the bed bug. *BMC Genomics*, **13**:6
562 Martin T, Chandre F, Ochou OG, Vaissayre M (2002) Pyrethroid resistance
563 mechanisms in the cotton bollworm *Helicoverpa armigera* (Lepidoptera:
564 Noctuidae) from West Africa. *Pesticide Biochemistry and Physiology*, **74**, 17–26.
565 Montella IR, Schama R, Valle D (2012) The classification of esterases: an important
566 gene family involved in insecticide resistance--a review. *Memórias do Instituto*
567 *Oswaldo Cruz*, **107**, 437–449.
568 Perera MDB, Hemingway J, Karunaratne SP (2008) Multiple insecticide resistance
569 mechanisms involving metabolic changes and insensitive target sites selected in
570 anopheline vectors of malaria in Sri Lanka. *Malaria Journal*, **7**, 168.
571 Potter MF, Romero A, Haynes KF (2008) Battling bed bugs in the USA. *Proceedings*
572 *of the Sixth International Conference on Urban Pests* (ed. by W.H. Robinson & D.
573 Bajomi). pp. 401-406. OOK-Press, Veszprem.
574 R Core Team (2015) R: A language and environment for statistical computing. R
575 Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-
576 project.org/.
577 Reznick DN, Ghalambor CK (2001) The population ecology of contemporary

578 adaptations: what empirical studies reveal about the conditions that promote
579 adaptive evolution. *Genetica*, **112**, 183–198.
580 Richards L, Boase CJ, Gezan S, Cameron MM (2009) Are bed bug infestations on the
581 increase within Greater London. *Journal of Environmental Health Research*. **9**, 17-
582 24.
583 Romero A, Anderson TD (2016) High levels of resistance in the common bed bug,
584 *Cimex lectularius* (Hemiptera: Cimicidae), to neonicotinoid insecticides. *Journal*
585 *of Medical Entomology*. **53**, 727-731.
586 Romero A, Potter MF, Haynes KF (2009) Behavioral responses of the bed bug to
587 insecticide residues. *Journal of Medical Entomology*, **46**, 51–57.
588 Romero A, Potter MF, Potter DA, Haynes KF (2007) Insecticide resistance in the bed
589 bug: a factor in the pest's sudden resurgence? *Journal of Medical Entomology*, **44**,
590 175–178.
591 Rosenfeld JA, Reeves D, Brugler MR *et al.* (2016) Genome assembly and geospatial
592 phylogenomics of the bed bug *Cimex lectularius*. *Nature Communications*, **7**,
593 10164.
594 Saccheri IJ, Brakefield PM (2002) Rapid spread of immigrant genomes into inbred
595 populations. *Proceedings of the Royal Society B: Biological Sciences*, **269**, 1073–
596 1078.
597 Sadílek D, Šťáhlavský F, Vilímová J, Zima J (2013) Extensive fragmentation of the X
598 chromosome in the bed bug *Cimex lectularius* Linnaeus, 1758 (Heteroptera,
599 Cimicidae): a survey across Europe. *Comparative cytogenetics*, **7**, 253–269.
600 Sawicki RM (1970) Interaction between the factor delaying penetration of insecticides
601 and the desethylation mechanism of resistance in organophosphorus-resistant
602 houseflies. *Pesticide Science*, **1,** 84–87.

603 Shono T, Kasai S, Kamiya E, Kono Y, Scott JG (2002) Genetics and mechanisms of
604 permethrin resistance in the YPER strain of house fly. *Pesticide Biochemistry and*
605 *Physiology*, **73**, 27–36.
606 Slack HD (1939) Structural hybridity in cimex L. *Zeitschrift für Zellforschung und*
607 *Mikroskopische Anatomie Abt. B Chromosoma*, **1**, 104–118.
608 Song Y, Endepols S, Klemann N *et al.* (2011) Adaptive introgression of anticoagulant
609 rodent poison resistance by hybridization between old world mice. *Current*
610 *Biology*, **21**, 1296–1301.
611 Ueshima N (1967) Supernumerary chromosomes in the human bed bug, *Cimex*
612 *lectularius* Linn. (Cimicidae:Hemiptera). *Zeitschrift für Zellforschung und*
613 *Mikroskopische Anatomie Abt. B Chromosoma*, **20**, 311–331.
614 Vicoso B, Charlesworth B (2006) Evolution on the X chromosome: unusual patterns
615 and processes. *Nature Reviews Genetics*, **7**, 645–653.
616 Yoon KS, Kwon DH, Strycharz JP *et al.* (2008) Biochemical and molecular analysis of
617 deltamethrin resistance in the common bed bug (Hemiptera: Cimicidae). *Journal*
618 *of Medical Entomology*, **45**, 1092–1101.
619 Zhu F, Gujar H, Gordon JR *et al.* (2013) Bed bugs evolved unique adaptive strategy to
620 resist pyrethroid insecticides. *Scientific Reports*, **3**, 1456.
621 Zhu F, Sams S, Moural T *et al.* (2012) RNA interference of NADPH-cytochrome P450
622 reductase results in reduced insecticide resistance in the bed bug, *Cimex*
623 *lectularius*, *PLoS ONE*, **7**, e31037.
624 Zhu F, Wigginton J, Romero A *et al.* (2010) Widespread distribution of knockdown
625 resistance mutations in the bed bug, *Cimex lectularius* (Hemiptera: Cimicidae),
626 populations in the United States. *Archives of Insect Biochemistry and Physiology*,
627 **73**, 245-257.
628

629 **Tables**

630 **Table 1: Genetic map summary.** Summary of marker number, spacing, map
631 distance and physical size (from reassembled genome) for *C. lectularius* linkage
632 groups.

Linkage group	No. Markers	Length (cM)	Length (Mb)	Mean spacing (cM)	Max spacing (cM)
1	38	28.23	39.95	0.76	4.28
2	37	45.18	48.13	1.26	10.32
3	34	35.58	34.27	1.08	9.66
4	33	35.49	54.64	1.11	5.39
5	30	27.26	49.12	0.94	4.52
6	27	43.01	32.09	1.65	7.01
7	25	39.19	44.27	1.63	4.81
8	21	24.88	30.64	1.24	8.17
9	19	24.15	22.31	1.34	5.74
10	19	16.96	13.04	0.94	4.52
11	18	35.41	38.21	2.08	10.13
12	15	21.93	16.06	1.57	4.57
13	14	25.53	6.51	1.96	6.45
14	4	3.78	4.48	1.26	2.26
All	334	406.58	433.71	1.27	10.32

633

634

635

636

Table 2: Genotype-phenotype counts and percentages at the focal marker (r2020_s2) on chromosome 12.

Genotype	Partial resistance	Resistance	Susceptible
AA	0	0	22 (100.0%)
AB	7 (24.1%)	3 (10.3%)	19 (65.5%)
BB	2 (10%)	18 (90.0%)	0

Table 3: Putative pyrethroid resistance candidate genes. Gene name, original reference genome scaffold and position for putative pyrethroid resistance genes identified in LG12 resistance QTL 95% Bayesian probability interval.

Gene name	NCBI Gene ID	Scaffold	Start (bp)	End (bp)
ubiquitin carboxyl-terminal hydrolase 15-like	106668434	Scaffold 2	16617167	16638289
sodium channel protein para	106667833	Scaffold 2	18435119	18463718
glutathione S-transferase	106666926	Scaffold 2	21129642	21130758
cytochrome P450 6B1-like	106663981	Scaffold 18	3480876	3508711
cytochrome P450 6B1-like	106663982	Scaffold 18	3456406	3461345
cytochrome P450 6B5-like	106663983	Scaffold 18	3404836	3440903
probable cytochrome P450 6a14	106663984	Scaffold 18	3476053	3495873

Figures

Figure 1: Linkage map and QTL analysis. (A) Linkage map showing positions of SNP markers for *C. lectularius* F_2 cross over 15 inferred linkage groups (putative chromosomes). (B) LOD scores for markers across the genome reveals a strong and significant peak on LG 12.

Supplementary Material

Figure S1: LOD scores vs. recombination fraction prior to correction. Co-variation in LOD scores and recombination fractions between markers suggests the presence of switched alleles in the dataset.

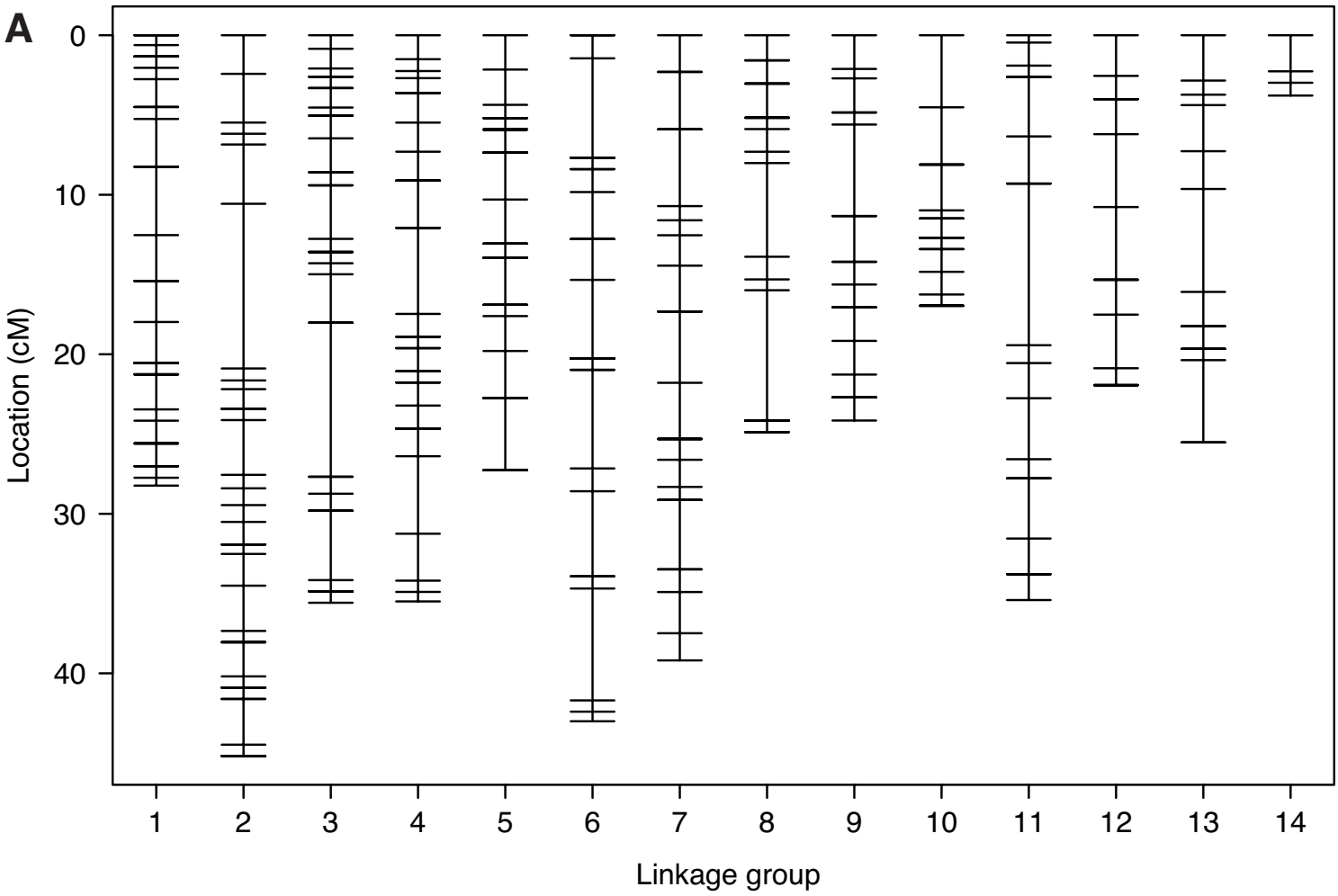
Figure S2: High LOD scores and low recombination fractions. High LOD scores between a single marker from LG1 and other markers on both LG2 and LG8 suggest a strong association, however recombination fraction is less than 0.5 on LG2 and >0.5 LG8, suggesting switched alleles.

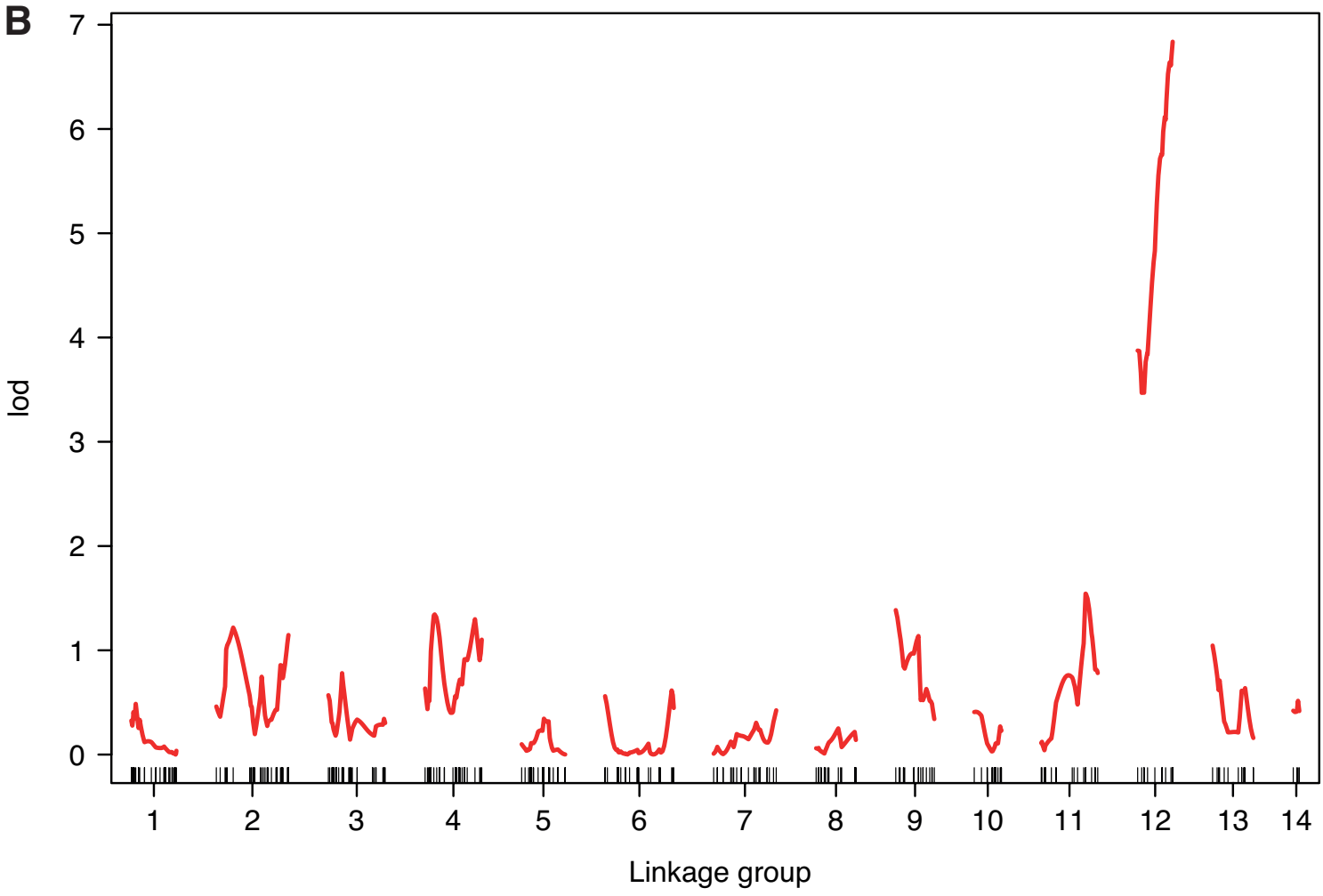
Figure S3: LOD scores vs. recombination fraction following correction and filtering. Heat map showing pairwise recombination fraction (upper diagonal) and LOD scores (lower diagonal) for the final linkage map following filtering. Purple = low LOD or large recombination fraction values, Yellow = high LOD and low recombination fraction (i.e. linked).

Figure S4: Maximum-likelihood estimation of genotyping error rate.

Table S1: Read statistics. Read, alignment and RAD locus statistics for each sequenced individual.

677 **Table S2: Marker information:** Reference genome scaffold, left-most base pair
678 position (i.e. alignment start position), linkage group and physical map position for
679 RAD markers used in genetic map construction.
680 **Table S3**: **Putative sex-linked loci and mapped scaffolds.**
681 **Table S4**: **Genes located in pyrethroid resistance QTL regions.**
682


A
Location (cM)
0
10
20
30
40
1
2
3
4
5
6
7
8
9
10
11
12
13
14
Linkage group


B
lod
0
1
2
3
4
5
6
7
1
2
3
4
5
6
7
8
9
10
11
12
13
14
Linkage group