Genetic basis of cuticular hydrocarbon variation in the desert ant *Cataglyphis niger*

Shani Inbar, Besan Saied, Pnina Cohen, Zeev Frenkel, Yoann Pellen, Abraham B. Korol, Eyal Privman[#]
Department of Evolutionary and Environmental Biology, Institute of Evolution, University of Haifa, Haifa, Israel
[#] Corresponding author: eprivman@univ.haifa.ac.il

Abstract

Cuticluar hydrocarbons (CHCs) are a ubiquitous component of insect cuticles that are used for a wide range of chemical signaling functions, especially recognition. Recognition and other signals are vital for the maintenance of insularity and cooperation in social insect colonies. Therefore, we expect natural selection on the composition and variabitlity of social insect CHC profiles. Selection on these signals may result in the evolution of genetic polymorphism affecting variation in CHC profiles. Here we tested for a genetic basis of CHC variation in the desert ant *Cataglyphis niger*. We applied a genomic mapping appraoch to a cohort of brothers from the same nest to reduce noise from environmental effects and achieve a clear statistical signal for association between the variation of CHCs and quantitative trait loci (QTL). This analysis identified 19 QTLs associated with 8 out of the 31 CHCs identified, and one QTL associated with total CHC quantity. These QTLs are located on 11 different chromosomes, including two cases where QTLs of different CHCs overlap. Each QTL explains between 13-25% of the variation in a specific CHC. We highlight several candidate genes in the QTLs identified, including fatty acid elongase and reductase genes. Our results reveal a polygenic genomic architecture underlying CHC variation in a population of the desert ant and open new research avenues into the genetic basis and evolution of chemical signaling in social insects.

Introduction

Long-chain cuticular hydrocarbons (CHCs) are a major component of the cuticular lipids of insects, which function as touch pheromones communicating a range of chemical signals (Drijfhout, et al. 2009). For example, males and females display distinct, species-specific CHCs and other cuticular lipids that are used for mate choice in different species, including *Drosophila* fruit flies (Ferveur and Cobb 2010; Jezovit, et al. 2017; Khallaf, et al. 2021) and *Nasonia* parasitoid wasps (Steiner, et al. 2006; Niehuis, et al. 2013; Buellesbach, et al. 2022).

In social insects, CHCs are also used for nestmate recognition. The ability to discriminate nestmates from foreign intruders is essential for colony integrity and protection of colony resources. CHCs are the main chemical cues underlying nestmate recognition (Holldobler and Michener 1980; Singer 1998). In this study, we investigated the genetic basis of CHC variability in terms of the number of loci and their effect sizes and identified candidate genes that may affect this important phenotype.

CHCs are the main lipid constituents of the epicuiticular layer of insects (reviewed in Blomquist and Bagneres 2010; Blomquist, et al. 2020). It has been speculated that the ancestral role of CHCs was probably to provide a waterproof layer preventing desiccation (Hadley 1994; Blomquist and Bagnères 2010 p. 5). Subsequently, they evolved a secondary role as chemical cues for signaling, communication, and recognition. Insect CHCs have variable chain lengths (20 to 50 carbons) and variable number and position of double bonds and methyl branches, thus allowing a great diversity that affords high information content as chemical cues. CHCs function as sex pheromones in many insects, as attractants, as fertility signals, and as chemical cues for species recognition that facilitates mate choice (Pickett, et al. 2009; Ferveur and Cobb 2010). Different social insect castes generally display distinct CHC profiles, including distinction between fertile individuals (queens, gamergates, or egg laying workers) and their less fertile or non-reproductive nestmates (Howard, et al. 1982; Nunes, et al. 2009; Kleeberg, et al. 2017; Derstine, et al. 2018; Sprenger and Menzel 2020). Signaling by CHCs have been extensively studied in social insects, where they make up an important component of the elaborate communication system necessary for coordinating complex collective behaviors (Hefetz 2007).

The enzymatic pathways responsible for the biosynthesis of CHCs were not characterized in social insects, however, a large body of literature describes such enzymes in model insect species, most notably *Drosophila melanogaster* (Blomquist 2010; Holze, et al. 2021). CHCs are *de novo* synthesized from two- or three-carbon units. Short saturated fatty acids (up to 18 carbons) are first synthesized by fatty acid synthases. These precursors are then elaborated by a series of enzymes that elongate the carbon chain (elongases), introduce double bonds (desaturases), and the final enzymatic steps that convert the fatty acid to a hydrocarbon, including reductases and decarboxylases. Hymenopteran homologs of these enzymes have been identified and studied for their role in CHC biosynthesis in bees and wasps (e.g., Falcón, et al. 2014; Buellesbach, et al. 2022; Moris, et al. 2023). Homologs are also found in ant

genomes, and are the natural candidates for the genetic basis of ant CHCs (e.g., Sprenger, et al. 2021; Wen, et al. 2025; and our own observations). Some of these gene families have significantly expended in ants, indicating the evolution of greater complexity in CHC synthesis (e.g., Hartke, et al. 2019). Beyond the genes involved in synthesis of CHCs, other genes may be involved in regulation of the quantitates of CHCs produced and secreted through the cuticle. These genes may function in any stage of the long road from the site of CHC production to the external surface of the cuticle. CHCs can be either synthesized de novo or from diet-derived fatty acids (Piek 1964; Blomquist and Jackson 1973; Liang and Silverman 2000; Blomquist 2010). They are thought to be produced primarily in cells known as oenocytes in the fat body, and then transported through the hemolymph, secreted and deposited on the cuticle (Gu, et al. 1995; Fan, et al. 2003).

One of the best understood social functions of CHCs is in nestmate recognition, a fundamental and vital function in social insect colonies. Nestmate recognition is primarily based on chemical recognition, which was suggested to involve matching an olfactory perception of chemical cues to an experience-derived neural template (Ozaki and Hefetz 2014). Different species display different mixtures of CHCs (Martin and Drijfhout 2009a; d'Ettorre and Lenoir 2010), which suggests rapid evolution, possibly driven by strong selection. Intra-specific quantitative variation generates colony-specific CHC profiles that facilitate nestmate recognition, as demonstrated by experimental manipulation in multiple ant species (Bonavita-Cougourdan, et al. 1987; Nowbahari, et al. 1990; Akino, et al. 2004; Ozaki, et al. 2005; Martin, et al. 2008), including in our study species *Cataglyphis niger* (Lahav, et al. 1999). These studies transferred the hydrocarbon fraction of extracts from one ant to another, which caused ants to attack nestmates or to reduce aggression towards non-nestmates. Some studies went a step further: they synthesized specific hydrocarbons and tested their effect on nestmate recognition (Akino, et al. 2004; Martin, et al. 2008).

Recognition cues may be heritable and/or environmentally derived. Heritability of CHC profiles would imply that quantitative variations in CHCs are partially explained by genetic polymorphism. Both genetic and environmental effects were documented in *Drosophila*, and some of the underlying genes were identified (reviewed in Ferveur 2005). Multiple studies suggest that quantitative variation in CHCs is heritable in diverse insect species including crickets (Thomas and Simmons 2008), termites (Dronnet, et al. 2006), and parasitoid wasps (Niehuis, et al. 2011; Buellesbach, et al. 2022). Empirical studies in social insects suggest

105 that CHC variation is affected by both environmental and genetic factors: For example, a
106 strong effect of the diet was reported in the Argentine ant (Liang and Silverman 2000).
107 Pathogens may also have an effect on nestmate recognition in ants (Tranter, et al. 2015;
108 Leclerc and Detrain 2016). Studies in ants and bees reported the effect of parasites and
109 pathogens on CHCs and nestmate recognition (Trabalon, et al. 2000; Beros, et al. 2017;
110 Cappa, et al. 2019). Other potential factors may include abiotic effects such as temperature
111 and humidity (Buellesbach, et al. 2018; Menzel, et al. 2018; Sprenger, et al. 2018; Silva, et al.
112 2024). Nevertheless, genetic effects may also be involved, and multiple studies suggested that
113 ant CHCs are heritable to a certain degree: Several studies demonstrated that intercolonial
114 variation in CHC profiles and the ability for nestmate recognition are retained under
115 controlled conditions and uniform diet (Nowbahari, et al. 1990; Lahav, et al. 1999; Akino, et
116 al. 2004; Ozaki, et al. 2005; van Zweden, et al. 2009). Behavioral experiments on different
117 populations of the ant *Temnothorax longispinosus* revealed an association between
118 aggression and genetic dissimilarity (Stuart and Herbers 2000). Isolation of *Formica exsecta*
119 workers revealed a weak patriline effect on CHCs, indicating that different alleles inherited
120 from different fathers had an effect on the CHC profiles of their daughters (Martin, et al.
121 2012). A study in the ant *Monomorium pharaonis* tested for heritability of individual CHCs
122 by systematic intercrossing eight colonies over nine generations (Walsh, et al. 2020),
123 resulting in low to moderate estimated heritability: narrow-sense heritability h^2 between
124 0.006 and 0.36. These studies provide evidence for heritability of CHCs in diverse ant
125 species.
126
127 The vital function of nestmate recognition is expected to have large fitness consequences,
128 including "red queen" dynamics due to selection against social parasites (Blomquist and
129 Bagneres 2010). This hypothesis is supported by the great interspecific diversity of CHCs and
130 their rapid evolution (Martin and Drijfhout 2009a; d'Ettorre and Lenoir 2010). Some studies
131 found stronger differentiation between colonies by methyl-branched vs. straight alkanes (e.g.,
132 Dahbi, et al. 1996). The abundance of branched alkanes and unsaturated alkenes in ant CHCs,
133 which provide greater information content as cues while compromising protection from
134 desiccation, was suggested to be the result of selection for signaling functions (d'Ettorre and
135 Lenoir 2010; Sturgis and Gordon 2012). The exceptional complexity of ant CHC profiles
136 suggests selection for recognition accuracy. If genetic polymorphism arises such that it
137 affects nestmate recognition cues, selection for distinctness of colony profiles is expected to
138 increase the frequency of novel alleles. The requirement for distinctness of colony CHC

139 profiles should lead to selection in favor of rare alleles, i.e. negative frequency-dependent
140 selection resulting in balancing selection (Ayala and Campbell 1974; Jongepier and Foitzik
141 2016). Such processes result in a large number of alleles per locus and the maintenance of
142 highly diverged alleles in the population over long evolutionary time, as observed for
143 example in mammalian Major Histocompatibility Complex (MHC) genes (Hughes and Nei
144 1988) and the honeybee Complementary Sex Determination (CSD) locus (Cho, et al. 2006).
145 We propose that the critical role of CHCs in nestmate recognition is expected to result in such
146 evolutionary dynamics, which would lead to the accumulation of polymorphic loci that affect
147 the display of ant CHCs quantitatively, i.e., quantitative trait loci (QTLs).
148
149 Here, we conducted a QTL mapping study to test whether quantitative variation of CHCs is
150 associated with genetic polymorphism in the desert ant *Cataglyphis niger*. We produce the
151 genomic infrastructures necessary for mapping associated genomic loci – a chromosome-
152 level reference genome sequence. We conduct reduced representation genomic sequencing of
153 a cohort of brothers to identify polymorphic loci associated with CHC variation (QTLs) and
154 characterize the genomic architecture of these traits. The use of haploid males affords
155 valuable advantages in terms of the simplicity of the underlying genetics and minimal
156 environmental variation (see Methods). The QTLs we discovered in males are most likely to
157 have a functional role in shaping their CHC profiles that are used as cues for mate choice,
158 i.e., as sex pheromones. However, we propose that these QTLs are likely to also have effects
159 on CHCs in workers and queens because many of the same CHCs are observed in all castes
160 of the species (Hefetz 2007; Inbar and Privman 2019) and because the biosynthetic pathway
161 of a given CHC is likely to be similar in all castes. We highlight promising candidate genes in
162 these loci. The discovery of genetic loci associated with CHC variation serves as a first step
163 in the study of the molecular mechanisms underlying chemical signaling functions including
164 sex pheromones and potentially also nestmate recognition, and their evolution in a social
165 insect species.
166

167 # Methods

168 **Samples**

169 Our analysis was based on a sample of 117 brothers, collected on 3/5/2016 from a single *C.*
170 *niger* nest in Betzet Beach in the northern Israeli coastline (colony BZT50, 33°03'32N
171 35°06'15E) and frozen in -80°C. Each of these samples were used for chemical analysis and

genomic sequencing. This colony belongs to a population that was previously referred to as *Cataglyphis drusus* (Eyer, et al. 2017), but our species delimitation analysis showed that it is part of the species described as *C. niger* that is widespread across Israel (Brodetzki, et al. 2019). Our population structure analysis of nuclear genomic DNA sequence data revealed no sub-structure of genetically distinct population clusters –populations across Israel are not genetically differentiated (Brodetzki, et al. 2019). We note that in some areas in Israel *C. niger* colonies have multiple queens (polygyne) and that our recent study revealed that this social polymorphism is associated with a supergene on chromosome 2 (Lajmi, et al. 2025). However, the study population used here is purely monogyne, i.e., each colony is headed by a single queen (Brodetzki, et al. 2019). The supergene is always homozygous (MM) in monogyne colonies (the P haplotypes is only found in polygyne colonies), so we do not expect it to affect the QTL mapping analysis in this study.

Using a cohort of brothers for QTL mapping affords several advantages. Ants have haplodiploid sex determination, where haploid males develop from unfertilized eggs, and diploid females derive from fertilized eggs. Thus, males have no fathers. They only receive one of their mother's alleles in each locus. And so, a cohort of brothers resents the products of maternal meiosis. Such a cohort servses as a simple mapping population similar to a back-cross experiment that is traditionally used for QTL mapping studies. Moreover, recessive alleles are more readily revealed in haploid progeny because samples are hemizygous at all loci, whereas recessive effects might not be observable in diploid samples (in heterozygous samples). For the same reason, similar studies using haploid males were conducted in another hymenopteran species – the parasitoid wasp *Nasonia vitripennis* (Niehuis, et al. 2011). This approach successfully identified a large number of QTLs associated with many of the CHCs analyzed. Another advantage is that CHC variation among brothers sampled from the same nest is likely to be minimally affected by environmental factors because they all experienced a similar environment in the nest. For example, we can safely assume they were fed the same diet and exposed to the same microbes by the workers that cared for them. Males are also expected to be more uniform than workers because they do not have distinct task groups that are exposed to different environments (e.g. nurses vs. foragers).

Chemical analysis

After being kept in a -80 freezer, a random substet of 81 of the 117 samples were individually extracted in 100 µl hexane, containing 200 ng tetracosane (C24) as internal standard.

206 Compounds were identified by combined gas chromatography/mass spectrometry (GC/MS)
207 based on their fragmentation patterns, as compared to authentic standards (linear aliphatic
208 hydrocarbons from C10 to C40 supplied by Supelco), following the methods used in previous
209 studies of this and other populations of the species (e.g., Brodetzki, et al. 2019). GC/MS was
210 conducted using an Agilent 19091S-433UI instrument at EI and splitless modes, equipped
211 with an HP-5ms (Ultra Inert) column, temperature-programmed from 60°C to 300°C (with 1
212 min initial hold) at a rate of 5°C per min, with a final hold of 15 min. Each sample was then
213 quantified by gas chromatography. Peak integration was performed using the program Varian
214 Galaxie version 1.9 to obtain the relative quantities of CHCs. We chose the conservative
215 approach of using a baseline that connects the vallies on either side of a peak in order to
216 avoid overestimation due to the effect of large peaks on neihboring smaller peaks.
217 Conversely, this approach might result in underestimation. Absolute quantities were
218 calculated relative to the internal standard, however the QTL mapping analysis was done
219 using the relative quantities.
220
221 **Reference genome sequencing, assembly and annotation**
222 Identification of the genes inside QTLs requires a genome assembly of sufficient contiguity,
223 that is, genomic scaffolds should be large enough so that the markers of the linkage map can
224 be linked to the genes within the confidence interval for the position of a QTL. We already
225 had a reference genome assembly for *C. niger* that was assembled from Illumina short-read
226 sequencing of a single haploid male (version Cnig_gn1; Yahav and Privman 2019). However,
227 this assembly was rather fragmented, with a scaffold N50 size of only 17.9 Kb. Therefore, we
228 conducted Pacbio long-read sequencing to improve the assembly contiguity. High-molecular
229 weight DNA was extracted from a pool of ~30 pupae from a *C. niger* colony also collected
230 from the Betzet Beach population. A library was constructed and sequenced (CLR
231 sequencing) on a Pacbio Sequel II machine at the Roy J. Carver Biotechnology Center,
232 University of Illinois at Urbana-Champaign. The sequencing yielded 7,031,744 reads, with
233 169 Gbp in total, and an average subread length of 7,227 bp.
234
235 The Pacbio data was assembled into two versions of the genome, Cnig_gn3.1 and
236 Cnig_gn3a.1. Version Cnig_gn3.1 was built more stringently, aiming to include only
237 confidently assembled contigs, while version Cnig_gn3a.1 aimed to map as much of the
238 genome sequence to chromosomes. We note that some artefacts are likely in the latter
239 version, which may be harmful to various types of analyses. We only use this assembly for

240 identifying candidate genes in QTLs. For this purpose, this practice is unlikely to have
241 serious detrimental effect. To avoid technical artifacts in Pacbio sequencing process short (<5
242 kbp) and long (>50 kbp) reads were excluded from assembly Cnig_gn3.1. Reads were
243 assembled onto contigs using *Canu* version 2.1.1 (Koren, et al. 2017). Contigs of
244 Cnig_gn3a.1 were based on all of the PacBio reads. Contigs of both assembly versions were
245 polished by *Pilon* version 1.24 (Walker, et al. 2014) using the short Illumina reads from
246 assembly Cnig_gn1. Polished contigs were anchored to the linkage map (see below) based on
247 markers with a unique *blastn* hit to contigs. Contigs with no mapped markers were fitted into
248 chromosomes based on the local similarity between the genomes of *C. niger* and *Formica*
249 *selysi* (https://www.ncbi.nlm.nih.gov/assembly/GCA_009859135.1/) if the synteny of the two
250 genomes was colinear in the relevant chromosomal region based on *blast* hits of *C. niger*
251 genetic markers to *F. selysi* chromosomes and *blast* of *F. selysi* genes to *C. niger* contigs.
252 Chromosomes in the Cnig_gn3.1 assembly contain only contigs with unique position and
253 defined orientation. Cnig_gn3a.1 also includes contigs with non-unique mapping and/or
254 unknown orientation. Contigs with several (up to 5) possible positions were included in
255 Cnig_gn3a.1 several times. Therefore, the Cnig_gn3a.1 had greater total length (265 vs. 236
256 Mbp) and N50 contig size (3.8 vs. 1.1 Mbp) relative to Cnig_gn3.1. The Cnig_gn3a.1
257 assembly includes 83% of the total assembly length in chromosomes (i.e. 17% of the
258 sequence is in contigs that no mapped markers), whereas the Cnig_gn3.1 assembly includes
259 74%.
260
261 Gene annotation was conducted using the *GeMoMa* pipeline (Keilwagen, et al. 2016;
262 Keilwagen, et al. 2018) that incorporates three types of evidence: (1) RNA sequencing of
263 samples from all life stages and castes (NCBI accession PRJNA494690; Yahav and Privman
264 2019) aligned to the genome using *HISAT2* (Kim, et al. 2019); (2) homology to genes
265 annotated in other insect genomes (*Camponotus floridanus, Harpegnathos saltator,*
266 *Linepithema humile, Solenopsis invicta, Pogonomyrmex barbatus, Formica exsecta, Apis*
267 *mellifera, Nasonia vitripennis, Anopheles gambiae, Drosophila melanogaster, Tribolium*
268 *castaneum*); and (3) *ab initio* gene prediction using *Augustus* (Stanke, et al. 2006). The
269 annotation resulted in 14,889 and 16,260 putative protein-coding genes for Cnig_gn3.1 and
270 Cnig_gn3a.1, respectively. Functional annotation of genes was assigned based on homology
271 using the *Trinotate* tool (version 3.0.2; https://github.com/Trinotate/Trinotate.github.io/wiki)
272 from the *Trinity* pipeline (Grabherr, et al. 2011), which uses *blast* against the Swissprot
273 protein database and assigns genes to functional categories of the Gene Ontology (GO)

274 database (Harris, et al. 2004). Out of the 16,260 genes in Cnig_gn3a.1, 10,268 genes had
275 Swissprot hits, and GO terms were assigned 10,186 genes. To identify candidate genes in
276 QTLs, we searched both Swissprot gene description and GO terms for keywords related to
277 molecular functions and biological processes implicated in CHC synthesis, including fatty
278 acid metabolism (keyword "fatty") and development and function of the cuticle (keyword
279 "cuticle").
280
281 **RAD sequencing**
282 DNA was extracted individually from each sample using TRIzol® (Invitrogen Life
283 Technologies) and kept at -20°C. Quantity and quality of extracts were determined by
284 NanoDrop™ 2000 Spectrophotometer and run on a 1.5% agarose gel to verify consistency in
285 DNA yield, degradation, and phenol/salts residues. Reduced representation genomic libraries
286 were constructed according to a modified double-digest Restriction-site Associated DNA
287 sequencing (ddRAD-seq) protocol based on Brelsford, et al. (2016). Briefly, DNA was
288 digested by two different restriction enzymes (EcoR1 rare-cutter; Mse1 frequent-cutter; from
289 New England Biolabs) and ligated to barcoded adaptors for multiplexing (using custom
290 oligos with 96 barcodes of variable lengths of 4-8bp, paired with two different barcoded PCR
291 primers to allow multiplexing all 117 samples together). Amplified products were pooled and
292 purified (Beckman Coulter, Agencourt® AMPure® XP) so that 200-700bp fragments would
293 be selected. The library was paired-end sequenced on one lane of a HiSeq4000 Illumina
294 sequencer.
295
296 The sequencing resulted in 693,969,414 150b reads (346,984,707 pairs), of which
297 662,096,312 were associated with a sample (barcode). The reads were quality checked and
298 those in which the quality score drops below an average phred score of 10 (i.e. 1 in 10 chance
299 for sequencing error) in a sliding window of 22 bases were removed using *process_radtags*
300 from the *STACKS* pipeline (Catchen, et al. 2011). After this initial filtering a total of
301 630,112,110 reads remain, which is expected to give an average sequencing depth of ~20X
302 (based on 67,687 positions of the EcoRI restriction sites in the reference genome).
303
304 Reads were aligned to the reference genome using *Bowtie2* (Langmead and Salzberg 2012).
305 Aligned reads that have more than 2 mismatches to the reference genome and reads that had
306 multiple mappings were removed. The number of aligned reads retaind totaled 431,321,204,
307 62% of the original number of reads.

308

309 The aligned reads were analyzed together to create a catalogue of 88,439 SNPs using the
310 *STACKS* pipeline, with an average sequencing depth of 14X across SNPs and samples. These
311 SNPs were filtered using custom perl scripts to remove loci that had more than three samples
312 had a heterozygous genotype call (not expected for haploid samples) or less than four
313 samples with the rare allele (the brothers are expected to have roughly 50% of each of the
314 maternal alleles) or more than 50% missing data. 7,769 SNPs passed this filtering.
315

316 **Linkage map**

317 The program *MultiPoint ultra-dense* (Ronin, et al. 2017) was used to construct a linkage map,
318 based on the genotypes of 7769 SNPs in the 117 samples. These loci were further filtered by
319 *MultiPoint* for segregation distortion (based on Chi^2 test to reject the null hypothesis of 1:1
320 segregation with $p < 0.005$) and for missing data (no more than 23 missing genotypes),
321 leaving 5466 SNPs. Clustering of linkage groups resulted in 27 clusters of SNP markers
322 using a recombination frequency threshold of 0.28. Skeleton markers were selected in every
323 linkage group (giving priority to "twin markers") and ordered by an "evolutionary strategy"
324 heuristic optimization, as implemented in *MultiPoint*. Ordering was assessed by jackknife
325 resampling. We manually inspected the linkage groups and applied methods implemented in
326 *MultiPoint* to merge linkage groups, exclude markers violating local map monotonicity, and
327 complement the map by adding suitable singleton markers, resulting in a final map consisting
328 of 26 linkage groups. To verify that the map length is not inflated by marker ordering errors,
329 we checked the order by jackknifing and observed a very nearly perfect support across the
330 map (i.e., 100% jackknife repeats gave identical order among each marker neighborhood
331 along the linkage group length).
332

333 **QTL mapping**

334 The program *MultiQTL* (Korol, et al. 1995) was used to map 32 quantitative traits of 81
335 samples, including 31 specific CHC traits and the total CHC amount. We defined the 31
336 specific CHC traits as relative amounts, while total CHC is an absolute amount in ng. Many
337 of the traits had skewed, non-normal distributions, so we removed outliers and applied log-
338 transformation where necessary. We first identified QTLs using the SIM (Single Interval
339 Mapping) model in *MultiQTL*, which calculates a LOD score for every trait on each
340 chromosome. For every putative QTL with LOD score greater than 1.8, we conducted a
341 permutation test to obtain a *p*-value for the significance of association between genotypes and

342 phenotypes. QTLs with *p*-value < 0.1 were carried on to the next step. For traits with two
343 peaks of the LOD scores on a chromosome, we used the linked-QTL model to test the
344 hypothesis of having two QTLs on the same chromosome (likelihood ratio test). For traits
345 that had candidate QTLs on more than one chromosome, we applied the MIM (multi-interval
346 mapping) model to all chromosomes identified in the SIM analysis, which generally results in
347 higher LOD scores. The permutation test was repeated for every QTL using the MIM model.
348 The SIM, linked-QTL, or MIM models also provided an estimate of the PEV (percent
349 explained variation) for each QTL. We identified the location of each QTL by the peak of the
350 LOD score, and its 95% confidence interval was defined as the region around the maximum,
351 where the LOD score is greater than the maximum score minus 1.5 points. We applied the
352 false discovery rate (FDR) method implemented in R (Benjamini and Hochberg 1995; Team
353 2021) to correct for multiple testing (32 traits by 26 chromosomes = 832 tests). We mapped
354 the genetic position of the confidence interval of each QTL location to the genome assembly
355 relative to the physical positions of genetic map markers. We then listed the annotated genes
356 that reside within the scaffold regions contained in each QTL interval in assembly
357 Cnig_gn3a.1.
358

359 ## Results

360 **Reference genome and linkage map**

361 As a basis for QTL mapping, we generated a new genome assembly for *C. niger* using high-
362 coverage Pacbio sequencing that yielded improved contiguity compared to the first draft
363 (scaffold N50 size of 1.08Mb compared to 17.9Kb for the first draft). We then constructed a
364 linkage map using the genotypes of a cohort of 117 haploid brothers (Supp. File S1). We used
365 RAD sequencing to genotype 7769 SNP loci that clustered into 26 linkage groups. We
366 anchored the assembly contigs to the linkage map, yielding 26 chromosomal sequences that
367 together encompass 83% of the total assembly length. The total map length was 5041cM,
368 which is exceptional for animal species. Some of the chromosomes have a very long genetic
369 map length: chromosome 1 is 402cM long and five additional chromosomes are longer than
370 280cM (Fig. 1). Given an estimated physical genome size of 236Mb (the total assembly
371 length), we estimated the genome-wide average recombination rate at 21.3 cM/Mb.

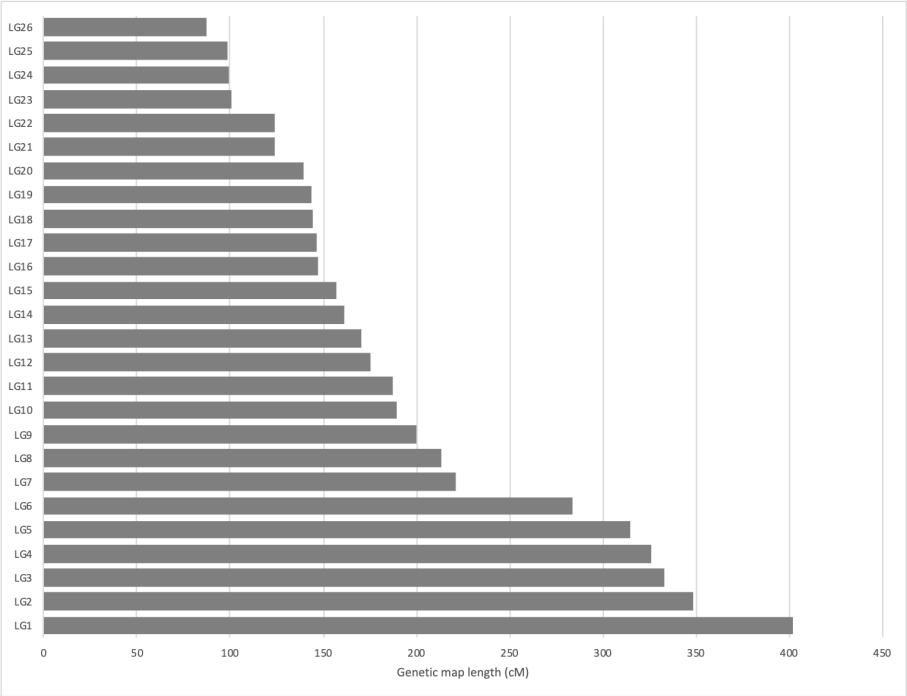
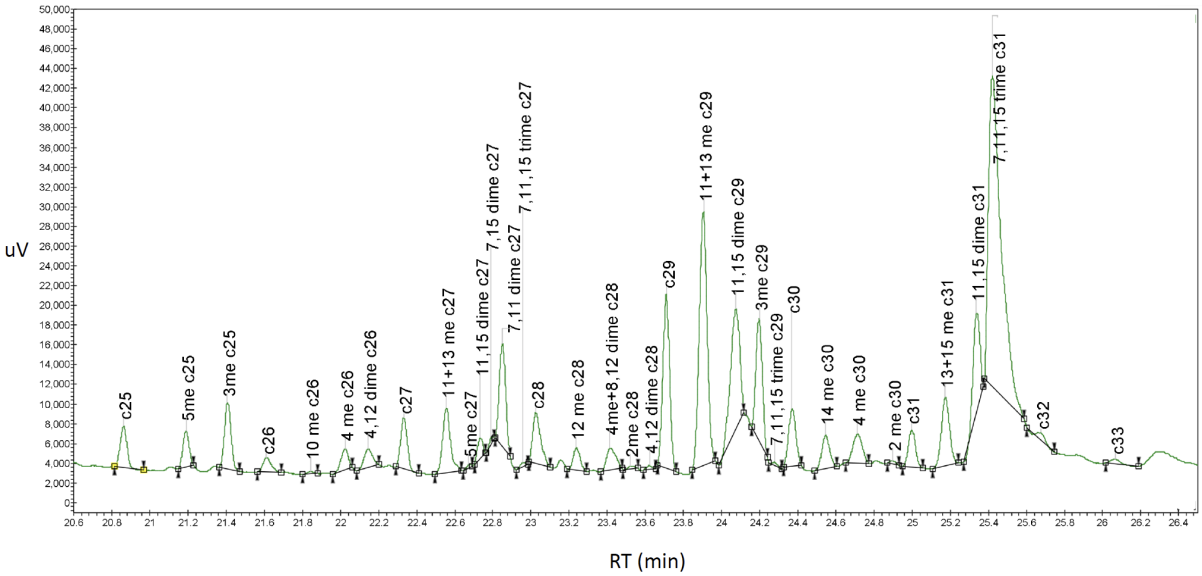


Figure 1: Distribution of genetic map lengths among the 26 chromosomes of *C. niger*.

Chemical analysis of CHCs

We analyzed the cuticular extracts of the same brothers used for linkage map construction, which were treated as quantitative traits for QTL mapping. Chemical analysis identified 32 differentiated peaks in the chromatogram (Fig. 2; Table 1; raw data are provided in Supp. File S2). We were able to identify 36 CHCs in 32 peaks (four peaks consisted of two CHCs that failed to separate). These were saturated alkanes with carbon chain length ranging from C25 to C31, many of them having a methyl group, and some with two and even three methyl branches. We did not observe any unsaturated alkenes. Note that this means that our list of candidate genes for CHC synthesis will not include fatty acid desaturases, which are known from the biosynthetic pathway of alkenes in other insects. The quantitative traits for QTL analysis were defined as the proportion of each CHC out of the total CHC mixture. We also included the total CHC amount as an additional trait, which varied from 228 to 4825ng; average and standard deviation 1478±762ng. There was extensive variation both among and within traits. The most abundant peak contained 11Me- and 13Me-C29, and the second most abundant was C29. These two peaks had averages and standard deviations of 226±122ng and 168±121ng, respectively. Conversely, the least abundant CHC, 5Me-C27, was only 1.1±1.8ng, and undetectable in 34 samples. Generally, the distributions of all CHCs were asymmetrical, with many low values and a long right-hand tail, which required transformation and removal of outliers before QTL analysis (see Methods and Supp. Fig. S1).

394 One of these peaks (7,11,15TriMe-C31) was excluded from the analysis because its
395 distribution was extremely non-normal, even after log transformation. Finally, 31 CHC traits
396 were carried forward to the QTL mapping analysis.



397

398 **Figure 2:** A chromatogram (uV absorption vs. retention time; RT) of the non-polar fraction (RT 20.5-
399 26.5) of cuticular extracts from a *C. niger* male. Marked peaks were annotated by mass spectrometry.
400

401 **Table 1: CHCs detected and their relative quantities in the cuticular extracts of *C. niger* males**

CHC	Average	Standard deviation
C25	0.03975	0.01111
3Me-C25	0.04626	0.01486
5Me-C25	0.01874	0.02621
C26	0.01057	0.00592
4,12DiMe-C26	0.00781	0.00660
10Me-C26	0.00103	0.00266
C27	0.03679	0.00911
11+13Me-C27	0.03846	0.00937
3Me-C27	0.04817	0.01595
5Me-C27	0.00132	0.00627
7,11,15TriMe-C27	0.00744	0.02473
13+15Me-C31	0.03752	0.01125
11,15DiMe-C27	0.00659	0.00279
7,11DiMe-C27	0.00234	0.00805
C28	0.08391	0.06570
4,12DiMe-C28	0.00070	0.00075
12Me-C28	0.01390	0.01524
2Me-C28	0.00140	0.00138
4Me+8,12DiMe-C28	0.01622	0.00360
C29	0.10508	0.03258
11,15DiMe-C29	0.08399	0.01314
11+13Me-C29	0.15324	0.02623
3Me-C29	0.07331	0.01762
7,11,15TriMe-C29	0.00098	0.00337
C30	0.03410	0.00748
14Me-C30	0.01734	0.00378

2Me-C30	0.00085	0.00101
4Me-C30	0.01969	0.00521
C31	0.02266	0.00877
11,15DiMe-C31	0.06731	0.01630
7,11,15TriMe-C31	0.00250	0.01576

402

403 **QTLs associated with variation in CHCs**

404 Each CHC (a peak in the GC chromatograms) was treated as a quantitative trait and tested for
405 association with the map markers (SNPs) in a QTL mapping analysis. The analysis included
406 the relative amounts of 31 CHCs, and the absolute total amount of CHCs as an additional
407 trait. A preliminary analysis identified 62 candidate genomic intervals with LOD scores
408 greater than 1.8. Each candidate was evaluated by permutation analysis, which yielded *p*-
409 values that were then corrected for multiple testing to control the false discovery rate (FDR).
410 This analysis resulted in 20 QTLs (FDR < 10%) for nine traits, including one QTL for total
411 CHC quantity (Table 2). The location and 95% confidence interval of each QTL was
412 estimated based on the LOD function (Fig. 3).

413

414 Interestingly, we did not detect statistically significant QTLs for most CHCs, but when we
415 did find QTLs for a certain CHC we found more than one (Supp. Fig. S2A). Out of the eight
416 CHCs with QTLs, four had three QTLs, five had two QTLs, on only one CHC had a single
417 QTL, which would be unlikely if all CHCs had equal probability to have QTLs (Poisson
418 dispersion test $p = 0.003$). These results indicate that the variation in some CHCs has a
419 greater genetic component than others. QTLs with large effect size are more likely to be
420 discovered with the limited statistical power given our sample size. Thus, other CHCs may
421 still be affected by QTLs with smaller effect sizes, which are below our detection threshold.
422 The portion of the phenotypic variation explained by QTLs (percent explained variation;
423 PEV) for each CHC ranged from 29.7% to 57.2% (Table 2). The PEV for single QTLs varied
424 between 12.8% and 25.4%. The most genetically determined CHC is C28, for which three
425 QTLs had PEV of 14.4%, 13.8%, and 21.9%, and together accounted for 57.2% PEV, which
426 implies that this trait is affected by genetic factors more than by environmental factors.

427

428 The 20 QTLs are located in 11 of the 26 chromosomes of *C. niger*, and the number of QTLs
429 per chromosome was not evenly distributed (Supp. Fig. S2B; between 1-5 QTLs per
430 chromosome; Poisson dispersion test $p = 0.06$). In some chromosomes the QTLs for different
431 CHCs appear to be clustered and even overlapping (Fig. 3). For example, QTLs for C28 and
432 2Me-C30 were found together in the tellomeric region of chromosome 4. Chromosome 18

433 harbors one locus associated with C26 and another associated with C27, 13Me- and 15Me-
434 C31 (the LOD plots are virtually overlapping). Each of the QTLs for C26 and C27 on
435 chromosome 17 are also weakly associated with the other CHC. These results suggest that
436 multiple CHCs may be regulated by the same gene or by different genes in a small genomic
437 region, which could harbor a cluster of co-regulated CHC genes.
438
439 **Table 2: QTLs underlying variation in CHCs**

CHC	LOD	Location (cM)	PEV (%)	*p*-value	*q*-value	MIM PEV(%)
C25	6.59*	**Chr6:** 0-43.2 **Chr6:** 76.9-88.8	19.4*	0.001*	0.069*	29.7
C26	4.215	**Chr6:** 56.4-67.4	16.2	0.001	0.069	46.1
	4.168	**Chr17:** 61.0-75.5	16.0	0.001	0.023	
	3.590	**Chr18:** 0-36.5	13.8	0.001	0.062	
C27	3.899	**Chr17:** 87.0-111.9	17.5	0.002	0.078	35.4
	3.981	**Chr18:** 102.1-113.5	18.0	0.001	0.023	
7,11,15TriMe-C27	4.139	**Chr15:** 122.9-154.7	15.0	0.001	0.064	49.4
	4.810	**Chr3:** 207.6-228.1	13.7	0.001	0.023	
	4.120	**Chr6:** 191.7-240.7	12.8	0.001	0.023	
C28	4.761	**Chr10:** 97.9-108.1	14.4	0.001	0.023	57.2%
	6.134	**Chr4:** 8.9-33.0	21.9	0.001	0.023	
	3.910	**Chr5:** 283.0-325.2	13.8	0.002	0.092	
4,12DiMe-C28	4.31*	**Chr20:** 5.0-28.5 **Chr20:** 57.9-84.4	24.9*	0.001*	0.023*	N/A
2Me-C30	3.823	**Chr21:** 10.9-34.2	14.1	0.001	0.064	46.1%
	4.083	**Chr13:** 17.3-31.7	15.6	0.001	0.023	
	3.838	**Chr4:** 0-17.2	16.4	0.002	0.092	
13+15Me-C31	4.837	**Chr18:** 105.9-120.3	25.4	0.001	0.023	N/A
Total CHC	3.886	**Chr15:** 89.3-102.5	15.1	0.001	0.064	44.0%

440 Results from the MIM model, showing only QTLs that passed the 10% FDR cutoff. The reported
441 MIM PEV is based on the analysis of all chromosomes for each trait. The full results for all tested
442 candidate QTLs are reported in Table S1, which includes results that did not pass the significance
443 threshold in the permutation test. LOD = log-of-odds score; Location = confidence interval of QTL
444 location; *q*-value = FDR corrected *p*-value; PEV = percent explained variation; MIM = multi-interval
445 model summarizing all QTLs for a given CHC over different chromosomes.
446 * Two QTLs for the same CHC on the same chromosome evaluated using the linked-QTL model of
447 *MultiQTL*, which gives the LOD, PEV, *p*-value, and *q*-value are for both QTLs together.
448

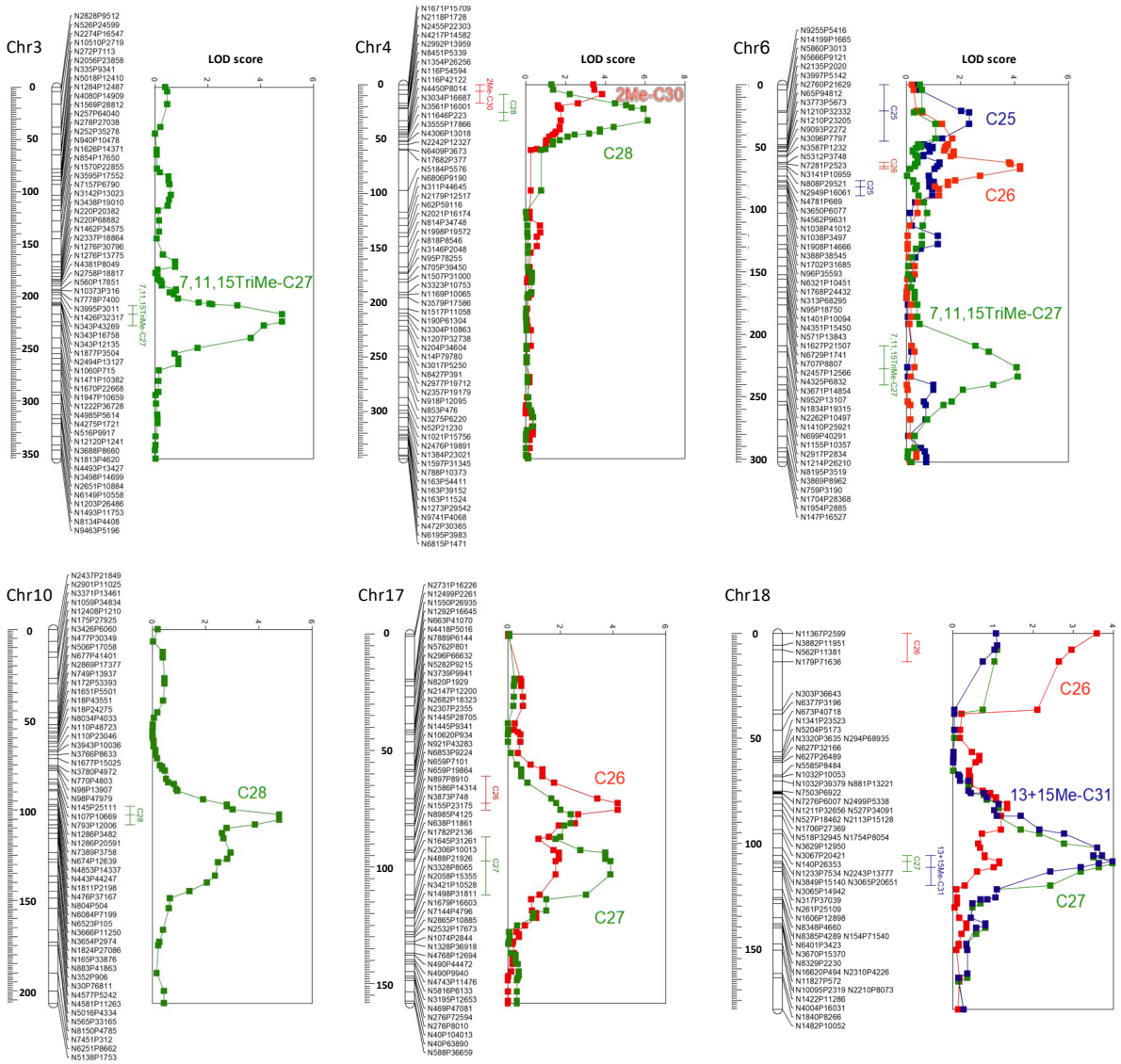


Figure 3: Location of CHC QTLs on six of the *C. niger* chromosomes. Map markers are listed with distances in cM along each linkage group. LOD scores are plotted for each trait, and 95% confidence intervals of QTL locations are indicated.

Candidate genes

Our sample size (*n*=81) limited the resolution of QTL mapping, resulting in 95% confidence intervals of genetic map length ranging between 10 and 49cM, and physical sizes ranging between 345 and 3,255 Kb. We note that three of the 20 QTLs included contigs that were mapped based on the conservation of synteny with *Formica* (150 out of 565 kbp of the QTL for C25 on chromosome 6; 167 / 3,571 kbp of 4,12DiMe-C28 on chr. 9; and 101 / 1,287 kbp of C26 on chr. 18). The mapping of these contigs was confirmed by the conservation of microsynteny (gene order) with orthologous genes in *Formica*. Each interval contained between 21 and 156 putative protein coding genes. Using our data, it is impossible to narrow

463 down these gene lists and identify the loci functionally responsible for the variation in the
464 phenotype. Nevertheless, we noted several candidate genes in these intervals that were
465 implicated in molecular functions and biological processes related to the synthesis of CHCs
466 or to the biology and development of the cuticle (Table 3). For example, these include a fatty
467 acyl-CoA reductase gene (Cnig_gene3669) in the QTL for C27 on chromosome 17 and a
468 long-chain fatty acid elongase gene (Cnig_gene6578) in the QTL for C26 on chromosome 18.
469
470 **Table 3: Candidate genes in QTLs underlying variation in CHCs**

QTL	Gene	Ortholog	Description
2Me-C30 chr. 13	Cnig_gene5601	FABP9_CAEEL	Fatty acid-binding protein homolog 9
2Me-C30 chr. 21	Cnig_gene2019	LIPI_HUMAN	Lipase member I
7,11,15TriMe-C27 chr. 06	Cnig_gene13278	PEX2_RAT	Peroxisome biogenesis factor 2
7,11,15TriMe-C27 chr. 15	Cnig_gene4696	ACADV_HUMAN	Very long-chain specific acyl-CoA dehydrogenase, mitochondrial (ECO:0000305|PubMed:7668252)
7,11,15TriMe-C27 chr. 15	Cnig_gene4695	KAPR1_DROME	cAMP-dependent protein kinase type I regulatory subunit
C26 chr. 18	Cnig_gene6578	ELO3_CAEEL	Putative fatty acid elongation protein 3
C27 chr. 17	Cnig_gene3669	WAT_DROME	Fatty acyl-CoA reductase *wat* (ECO:0000305)
Total chr. 15	Cnig_gene4381	FABD_DROME	Probable malonyl-CoA-acyl carrier protein transacylase, mitochondrial

471 QTL; gene ID for the *C. niger* candiate gene; gene ID of the best Swissprot blast hit and the
472 functional description of that gene.
473

474 ## Discussion

475 Here we applied a genomic mapping approach to test for genetic effects on CHCs, which
476 serve as major chemical cues for recognition and other functions in social insects. We had to
477 establish a chromosome-level reference genome assembly for the desert ant *Cataglyphis*
478 *niger* as the foundation for this genomic mapping study. The long-read sequencing combined
479 with the high-resolution linkage map resulted in draft assemblies for each of the 26 linkage
480 groups, which is in agreement with the chromosome number in *Cataglyphis hispanica*, where
481 one lineage has 26 chromosomes while the other has 27 (Darras, et al. 2022). The linkage
482 map and the assembly revealed an exceptionally high estimate of a genome-wide average
483 recombination rate of 21.3 cM/Mb. This rate is similar to the record for animals that is held
484 by honey bees, with estimates ranging between 19-22 cM/Mb (Beye, et al. 2006; Solignac, et

485 al. 2007; Meznar, et al. 2010; Korol and Rybnikov 2024). The total physical and genetic
486 lengths of the *C. niger* genome are both greater than those of *Apis mellifera*: 236Mb and
487 5041cM vs. 187Mb and 4,114cM (Solignac, et al. 2007). For the QTL mapping analysis, we
488 analyzed CHC profiles and genomic polymorphism in a sample of 81 haploid brothers. We
489 used RAD sequencing to obtain a high-density genetic map for QTL mapping based on 7769
490 markers. This density of markers is more than needed for this progeny sample. The QTL
491 detection and the resolution of QTL location is limited by the number of recombination
492 events in our sample rather than the number of genetic markers. We report associations
493 between 19 QTLs and 8 out of 31 CHCs tested, and one locus that is associated with the total
494 amount of CHCs. We identified rather large effect sizes that each explain between 13% and
495 25% of the variation in a given CHC. Between 30% and 57% of the variation of the eight
496 affected CHCs was explained by the combined genetic effects. Our study opens a window to
497 the genomic architecture underlying CHC variation in a social insect species.
498
499 Our results add to a large volume of literature on determinants of CHC profiles. Previous
500 studies demonstrated large environmental effects by various factors, including diet (Liang
501 and Silverman 2000), pathogens (Tranter, et al. 2015; Leclerc and Detrain 2016), and abiotic
502 environmental factors such as temperature and humidity (Silva, et al. 2024). Incorporating
503 our results allows a comprehensive understanding of the mechanisms shaping CHC profiles
504 by a combination of genetic and environmental effects. Our results reveal large variation in
505 the genetic effects on different CHCs. We discovered QTLs with intermediate effect sizes for
506 eight CHCs; each locus explaining between 12 and 25 percent of phenotypic variation. Up to
507 57% of the variation of C28 was explained by the combined effects of the QTLs we
508 identified, but for the other CHCs most of the phenotypic variation was not explained by the
509 genetic effects. CHCs with QTLs are expected to be heritable as suggested by previous
510 studies in multiple ant species (Nowbahari, et al. 1990; Lahav, et al. 1999; Akino, et al. 2004;
511 Ozaki, et al. 2005; van Zweden, et al. 2009; Walsh, et al. 2020). Our results are in line with
512 estimates of narrow-sense heritability (h^2) of individual CHCs in the ant *Monomorium*
513 *pharaonis* that ranged between 0 and 0.36 (Walsh, et al. 2020). We could not detect any
514 QTLs associated with 23 out of the 31 CHCs in our data. This may be because some CHCs
515 are only affected by small effect size QTLs that could not be detected with our sample size.
516 Generally, our estimates of the total genetic component for any of the studied CHCs should
517 be considered as a lower bound, because they may be affected by additional small effect size
518 QTLs that we could not detect. Nevertheless, our results indicate substantial variation among

519 CHCs in the magnitude of their genetic basis and that for most CHCs the genetic component
520 is smaller than the environmental component. Environmental effects may dominate the cues
521 for nestmate recognition because nestmates share the same abiotic environment of the nest,
522 the same diet, and other biotic effects such as pathogens.
523
524 The eight CHCs with associated QTLs include four straight-chain and four methyl-branched
525 alkanes. This may be surprising considering that methyl-branched CHCs have greater
526 information content, and thus are expected to be under greater selection as recognition cues,
527 and more likely to evolve a genetic basis. This hypothesis was supported by some studies that
528 reported stronger differentiation between colonies by methyl-branched vs. straight alkanes in
529 some species (e.g., in *Cataglyphis iberica*; Dahbi, et al. 1996) and stronger direct and indirect
530 genetic effects (e.g., in *Formica rufibarbis*; Van Zweden, et al. 2010). There may be other
531 reasons for the evolution of a genetic basis for the less informative straight-chain alkanes,
532 other than for nestmate recognition purposes. For example, these alkanes may be used for
533 simple, conserved signals that do not require high information content, such as differentiating
534 between foragers and nurses (Greene and Gordon 2003; Martin and Drijfhout 2009b) or to
535 identify a mated queen (Hora, et al. 2008). If alkanes are used for such a conserved function
536 they may evolve a genetic basis to better regulate their expression.
537
538 As discussed in the Introduction, the most obvious candidate genes are ant homologs of
539 enzymes that were extensively studied in model insects and implicated in the biosynthetic
540 pathway of long-chain fatty acids, which are the precursors of CHCs. These include fatty acid
541 synthases, fatty acid elongases, and fatty acid reductases. One such candidate,
542 Cnig_gene3669 in the QTL for C27 on chromosome 17, is a homolog of *Drosophila* fatty
543 acyl-CoA reductase (FAR) genes. The top blast hit in *Drosophila* is the gene *waterproof*
544 (*wat*; see a gene tree of ant and fly FAR genes in Supp. Fig. S3), which is involved in tracheal
545 development. Knocking out *wat* interfered with gas filling of the tracheal tubes during
546 *Drosophila* embryogenesis (Jaspers, et al. 2014) and *wat* homologs in planthoppers affect
547 CHC quantities (Li, et al. 2019; Li, et al. 2020). Generally, FAR genes are responsible for one
548 of the last steps in CHC biosynthetic pathways, where fatty-acyl CoA precursors are
549 converted to hydrocarbons. The FAR gene family was implicated in CHC synthesis in
550 *Drosophila* based on their expression in oenocytes and their role in regulation of oenocyte
551 development (Finet, et al. 2019), association with intraspecific CHC variation, and
552 knockdown experiments that affected CHC length (Dembeck, et al. 2015). Another

553 candidate, Cnig_gene6578 in the QTL for C26 on chromosome 18, is a homolog of
554 *Drosophila* long-chain fatty acid elongase genes. This gene family takes part in the CHC
555 synthesis pathway, in elongating the chain length of fatty acids initially synthesized by fatty
556 acid synthases beyond 20 carbons (Blomquist and Bagneres 2010). Knockdown of such
557 elongases in *Drosophila* and in the brown planthopper *Nilaparvata lugens* affected
558 abundances of different CHCs (Holze, et al. 2021). The QTL for 2Me-C30 on chromosome
559 21 contains a homolog of human lipase member I, which hydrolyzes phosphatidic acid (PA)
560 to 2-acyl lysophosphatidic acid (LPA; a potent bioactive lipid mediator) and a fatty acid
561 (Hiramatsu, et al. 2003). While the function of the insect homolog is unknown, it is a
562 promising candidate in lipid metabolism, which may be involved in the synthesis of CHC
563 precursors.
564
565 In this study we used a cohort of male brothers as the mapping population. Since this cohort
566 was collected from a single nest, it represent a sample of a single colony from the population.
567 As such, we cannot see how the genetic architecture of CHCs varies across the population.
568 Nevertheless, given the strong evidence for multiple QTLs for multiple CHCs in this colony
569 sample we expect this to generally characterize the genetic architecture of ant CHCs, even if
570 the specific loci and their effect sizes may vary. Our motivation for using males was that they
571 are haploid, which is advantageous for genomic mapping, and because they developed in the
572 same nest in uniform environmental conditions. Male CHCs are likely to have functional
573 roles for recognition and discrimination in mate choice during mating flights. CHCs were
574 extensively investigated in such roles in non-social insects. For example, male offspring from
575 a cross between the wasps *Nasonia vitripennis* and *N. giraulti* were used for a QTL mapping
576 study of CHCs (Niehuis, et al. 2011). CHCs were demonstrated to function as cues for mate
577 choice in this system, which allows females to discriminate between males from different
578 species (Buellesbach, et al. 2013). Similarly, CHCs of male ants are likely to function as cues
579 for mate choice by queens (Sprenger and Menzel 2020). In the worker caste of social insects,
580 CHCs are most consequential in their role as nestmate recognition cues for identifying
581 intruders from other colonies (van Zweden and d'Ettorre 2010). The genetic effects we
582 uncovered in males are likely to also affect CHC profiles of queens and workers, because
583 most CHCs are observed in more than one caste of the same species (Hefetz 2007). Genes
584 involved in the synthesis of these CHCs likely evolved in response to selection pressures on a
585 range of various functional roles in each of the different castes (e.g. nestmate recognition in
586 workers vs. mate choice in males). Natural selection on the ability to discriminate against

foreign workers may increase genetic polymorphism that contributes to CHC variation, and this polymorphism may also affect variation of the same CHC in males even if this CHC is not under selection in males. Therefore, some of the QTLs we detected in males may actually be functionally important in one or the other caste. Further studies in other castes will provide a more comprehensive understanding of this genetic system and its functional significance.

Acknowledgements

We thank Abraham Hefetz for advice and access to equipment for chemical analysis of CHCs. We thank Jessica Purcell and Alan Brelsford for advice on RAD sequencing. This study was funded by Israel Science Foundation grant number 646/15.

Data availability

All sequence data was deposited to NCBI: the new genome assembly Cnig_gn3.1 is under NCBI BioProject accession PRJNA1168153 and the raw Pacbio sequencing reads are under SRA accession SRR10302333. The RAD-seq data is under Bioproject accession PRJNA1000554. Chemical data is provided as supplementary materials to this paper.

References

Akino T, Yamamura K, Wakamura S, Yamaoka R. 2004. Direct behavioral evidence for hydrocarbons as nestmate recognition cues in Formica japonica (Hymenoptera : Formicidae). Applied Entomology and Zoology 39:381-387.

Ayala FJ, Campbell CA. 1974. Frequency-dependent selection. Annual Review of Ecology and Systematics 5:115-138.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B Met 57:289-300.

Beros S, Foitzik S, Menzel F. 2017. What are the mechanisms behind a parasite-induced decline in nestmate recognition in ants? Journal of Chemical Ecology 43:869-880.

Beye M, Gattermeier I, Hasselmann M, Gempe T, Schioett M, Baines JF, Schlipalius D, Mougel F, Emore C, Rueppell O, et al. 2006. Exceptionally high levels of recombination across the honey bee genome. Genome Res 16:1339-1344.

Blomquist GJ. 2010. Biosynthesis of cuticular hydrocarbons. In: Blomquist GJ, Bagnères A-G, editors. Insect hydrocarbons: biology, biochemistry and chemical ecology. p. 325-343.

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

Blomquist GJ, Bagnères A-G. 2010. Introduction: history and overview of insect hydrocarbons. In: Blomquist GJ, Bagnères A-G, editors. Insect hydrocarbons: biology, biochemistry and chemical ecology. p. 325-343.

Blomquist GJ, Jackson LL. 1973. Incorporation of labelled dietary n-alkanes into cuticular lipids of the grasshopper Melanoplus sanguinipes. Journal of insect physiology 19:1639-1647.

Blomquist GJ, Tittiger C, Jurenka R. 2020. Cuticular hydrocarbons and pheromones of arthropods. In: Wilkes H, editor. Hydrocarbons, oils and lipids: diversity, origin, chemistry and fate. Cham: Springer. p. 213-244.

Bonavita-Cougourdan A, Clément J, Lange C. 1987. Nestmate recognition: the role of cuticular hydrocarbons in the ant Camponotus vagus Scop.

Brelsford A, Dufresnes C, Perrin N. 2016. High-density sex-specific linkage maps of a European tree frog (Hyla arborea) identify the sex chromosome without information on offspring sex. Heredity (Edinb) 116:177-181.

Brodetzki TR, Inbar S, Cohen P, Aron S, Privman E, Hefetz A. 2019. the Interplay between Incipient species and social polymorphism in the Desert Ant Cataglyphis. Scientific reports 9:9495.

Buellesbach J, Gadau J, Beukeboom LW, Echinger F, Raychoudhury R, Werren JH, Schmitt T. 2013. Cuticular hydrocarbon divergence in the jewel wasp Nasonia: evolutionary shifts in chemical communication channels? Journal of evolutionary biology 26:2467-2478.

Buellesbach J, Holze H, Schrader L, Liebig J, Schmitt T, Gadau J, Niehuis O. 2022. Genetic and genomic architecture of species-specific cuticular hydrocarbon variation in parasitoid wasps. Proceedings of the Royal Society B 289:20220336.

Buellesbach J, Whyte BA, Cash E, Gibson JD, Scheckel KJ, Sandidge R, Tsutsui ND. 2018. Desiccation resistance and micro-climate adaptation: cuticular hydrocarbon signatures of different Argentine ant supercolonies across California. Journal of Chemical Ecology 44:1101-1114.

Cappa F, Petrocelli I, Dani FR, Dapporto L, Giovannini M, Silva-Castellari J, Turillazzi S, Cervo R. 2019. Natural biocide disrupts nestmate recognition in honeybees. Scientific reports 9:3171.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH. 2011. Stacks: building and genotyping Loci de novo from short-read sequences. G3 1:171-182.

Cho S, Huang ZY, Green DR, Smith DR, Zhang J. 2006. Evolution of the complementary sex-determination gene of honey bees: Balancing selection and trans-species polymorphisms. Genome Research 16:1366-1375.

d'Ettorre P, Lenoir A. 2010. Nestmate recognition. Ant ecology:194-209.

Dahbi A, Cerdá X, Hefetz A, Lenoir A. 1996. Social closure, aggressive behavior, and cuticular hydrocarbon profiles in the polydomous ant Cataglyphis iberica (Hymenoptera, Formicidae). Journal of Chemical Ecology 22:2173-2186.

Darras H, de Souza Araujo N, Baudry L, Guiglielmoni N, Lorite P, Marbouty M, Rodriguez F, Arkhipov I, Koszul R, Flot J-F. 2022. Chromosome-level genome assembly and annotation of two lineages of the ant Cataglyphis hispanica: steppingstones towards genomic studies of hybridogenesis and thermal adaptation in desert ants. bioRxiv.

Dembeck LM, Böröczky K, Huang W, Schal C, Anholt RR, Mackay TF. 2015. Genetic architecture of natural variation in cuticular hydrocarbon composition in Drosophila melanogaster. Elife 4:e09861.

Derstine N, Gries R, Zhai H, Jimenez S, Gries G. 2018. Cuticular hydrocarbons determine sex, caste, and nest membership in each of four species of yellowjackets (Hymenoptera: Vespidae). Insectes Sociaux 65:581-591.

Drijfhout FP, Kather R, Martin SJ. 2009. The role of cuticular hydrocarbons in insects. In: Zhang W, Liu H, editors. Behavioral and chemical ecology: Nova Science Publishers, Inc. p. 91-114.

676 Dronnet S, Lohou C, Christides J-P, Bagnères A-G. 2006. Cuticular hydrocarbon
677 composition reflects genetic relationship among colonies of the introduced termite
678 Reticulitermes santonensis Feytaud. Journal of Chemical Ecology 32:1027-1042.
679 Eyer PA, Seltzer R, Reiner-Brodetzki T, Hefetz A. 2017. An integrative approach to
680 untangling species delimitation in the Cataglyphis bicolor desert ant complex in Israel.
681 Mol Phylogenet Evol 115:128-139.
682 Falcón T, Ferreira-Caliman MJ, Nunes FMF, Tanaka ÉD, do Nascimento FS, Bitondi MMG.
683 2014. Exoskeleton formation in Apis mellifera: Cuticular hydrocarbons profiles and
684 expression of desaturase and elongase genes during pupal and adult development.
685 Insect Biochemistry and Molecular Biology 50:68-81.
686 Fan Y, Zurek L, Dykstra MJ, Schal C. 2003. Hydrocarbon synthesis by enzymatically
687 dissociated oenocytes of the abdominal integument of the German cockroach, Blattella
688 germanica. Naturwissenschaften 90:121-126.
689 Ferveur J-F. 2005. Cuticular hydrocarbons: their evolution and roles in Drosophila
690 pheromonal communication. Behavior genetics 35:279-295.
691 Ferveur J-F, Cobb M. 2010. Behavioral and evolutionary roles of cuticular hydrocarbons in
692 Diptera. In: Blomquist GJ, Bagnères A-G, editors. Insect hydrocarbons: biology,
693 biochemistry and chemical ecology. p. 325-343.
694 Finet C, Slavik K, Pu J, Carroll SB, Chung H. 2019. Birth-and-death evolution of the fatty
695 acyl-CoA reductase (FAR) gene family and diversification of cuticular hydrocarbon
696 synthesis in Drosophila. Genome biology and evolution 11:1541-1551.
697 Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L,
698 Raychowdhury R, Zeng Q. 2011. Trinity: reconstructing a full-length transcriptome
699 without a genome from RNA-Seq data. Nature biotechnology 29:644.
700 Greene MJ, Gordon DM. 2003. Cuticular hydrocarbons inform task decisions. Nature
701 423:32-32.
702 Gu X, Quilici D, Juarez P, Blomquist G, Schal C. 1995. Biosynthesis of hydrocarbons and
703 contact sex pheromone and their transport by lipophorin in females of the German
704 cockroach (Blattella germanica). Journal of insect physiology 41:257-267.
705 Hadley NF. 1994. The water relations of terrestrial arthropods: CUP Archive.
706 Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, Lewis S,
707 Marshall B, Mungall C, et al. 2004. The Gene Ontology (GO) database and informatics
708 resource. Nucleic acids research 32:D258-261.
709 Hartke J, Schell T, Jongepier E, Schmidt H, Sprenger PP, Paule J, Bornberg-Bauer E, Schmitt
710 T, Menzel F, Pfenninger M. 2019. Hybrid genome assembly of a neotropical
711 mutualistic ant. Genome biology and evolution 11:2306-2311.
712 Hefetz A. 2007. The evolution of hydrocarbon pheromone parsimony in ants (Hymenoptera:
713 Formicidae)—interplay of colony odor uniformity and odor idiosyncrasy.
714 Myrmecological News 10:59-68.
715 Hiramatsu T, Sonoda H, Takanezawa Y, Morikawa R, Ishida M, Kasahara K, Sanai Y,
716 Taguchi R, Aoki J, Arai H. 2003. Biochemical and molecular characterization of two
717 phosphatidic acid-selective phospholipase A1s, mPA-PLA1α and mPA-PLA1β. Journal
718 of Biological Chemistry 278:49438-49447.
719 Holldobler B, Michener CD. 1980. Mechanisms of identification and discrimination in social
720 Hymenoptera. In: Markl H, editor. Evolution of social behaviour: hypotheses and
721 empirical tests. p. 35-58.
722 Holze H, Schrader L, Buellesbach J. 2021. Advances in deciphering the genetic basis of
723 insect cuticular hydrocarbon biosynthesis and variation. Heredity 126:219-234.

Hora RR, Ionescu-Hirsh A, Simon T, Delabie J, Robert J, Fresneau D, Hefetz A. 2008. Postmating changes in cuticular chemistry and visual appearance in Ectatomma tuberculatum queens (Formicidae: Ectatomminae). Naturwissenschaften 95:55-60.
Howard RW, McDaniel C, Nelson DR, Blomquist GJ, Gelbaum LT, Zalkow LH. 1982. Cuticular hydrocarbons of Reticulitermes virginicus (Banks) and their role as potential species-and caste-recognition cues. Journal of Chemical Ecology 8:1227-1239.
Hughes AL, Nei M. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature 335:167-170.
Inbar S, Privman E. 2019. Differences in Long-Chained Cuticular Hydrocarbons between Males and Gynes in Cataglyphis Desert Ants. bioRxiv:538926.
Jaspers MH, Pflanz R, Riedel D, Kawelke S, Feussner I, Schuh R. 2014. The fatty acyl-CoA reductase Waterproof mediates airway clearance in Drosophila. Developmental biology 385:23-31.
Jezovit JA, Levine JD, Schneider J. 2017. Phylogeny, environment and sexual communication across the Drosophila genus. Journal of Experimental Biology 220:42-52.
Jongepier E, Foitzik S. 2016. Ant recognition cue diversity is higher in the presence of slavemaker ants. Behavioral Ecology 27:304-311.
Keilwagen J, Hartung F, Paulini M, Twardziok SO, Grau J. 2018. Combining RNA-seq data and homology-based gene prediction for plants, animals and fungi. BMC bioinformatics 19:1-12.
Keilwagen J, Wenk M, Erickson JL, Schattat MH, Grau J, Hartung F. 2016. Using intron position conservation for homology-based gene prediction. Nucleic acids research 44:e89-e89.
Khallaf MA, Cui R, Weißflog J, Erdogmus M, Svatoš A, Dweck HK, Valenzano DR, Hansson BS, Knaden M. 2021. Large-scale characterization of sex pheromone communication systems in Drosophila. Nature communications 12:4165.
Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nature biotechnology 37:907-915.
Kleeberg I, Menzel F, Foitzik S. 2017. The influence of slavemaking lifestyle, caste and sex on chemical profiles in Temnothorax ants: insights into the evolution of cuticular hydrocarbons. Proceedings of the Royal Society B: Biological Sciences 284:20162249.
Korol A, Rybnikov S. 2024. Recombination. In. Encyclopedia of Biodiversity: Elsevier. p. 375-396.
Korol AB, Ronin YI, Kirzhner VM. 1995. Interval mapping of quantitative trait loci employing correlated trait complexes. Genetics 140:1137-1147.
Lahav S, Soroker V, Hefetz A, Vander Meer RK. 1999. Direct behavioral evidence for hydrocarbons as ant recognition discriminators. Naturwissenschaften 86:246-249.
Lajmi A, Cohen P, Lee C-C, Frenkel Z, Pellen Y, Privman E. 2025. Repeated evolution of supergenes on an ancient social chromosome. bioRxiv, accepted to Current Biology:2024.2012. 2001.626239.
Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nature methods 9:357-359.
Leclerc J-B, Detrain C. 2016. Ants detect but do not discriminate diseased workers within their nest. The Science of Nature 103:1-12.
Li D-T, Chen X, Wang X-Q, Zhang C-X. 2019. FAR gene enables the brown planthopper to walk and jump on water in paddy field. Science China Life Sciences 62:1521-1531.
Li DT, Dai YT, Chen X, Wang XQ, Li ZD, Moussian B, Zhang CX. 2020. Ten fatty acyl-CoA reductase family genes were essential for the survival of the destructive rice pest, Nilaparvata lugens. Pest management science 76:2304-2315.

774 Liang D, Silverman J. 2000. “You are what you eat”: diet modifies cuticular hydrocarbons
775 and nestmate recognition in the Argentine ant, Linepithema humile.
776 Naturwissenschaften 87:412-416.
777 Martin S, Drijfhout F. 2009a. A review of ant cuticular hydrocarbons. Journal of Chemical
778 Ecology 35:1151-1161.
779 Martin S, Trontti K, Shemilt S, Drijfhout F, Butlin R, Jackson D. 2012. Weak patriline effects
780 are present in the cuticular hydrocarbon profiles of isolated Formica exsecta ants but
781 they disappear in the colony environment. Ecology and evolution 2:2333-2346.
782 Martin SJ, Drijfhout FP. 2009b. Nestmate and task cues are influenced and encoded
783 differently within ant cuticular hydrocarbon profiles. Journal of Chemical Ecology
784 35:368-374.
785 Martin SJ, Vitikainen E, Helantera H, Drijfhout FP. 2008. Chemical basis of nest-mate
786 discrimination in the ant Formica exsecta. Proceedings of the Royal Society B-
787 Biological Sciences 275:1271-1278.
788 Menzel F, Zumbusch M, Feldmeyer B. 2018. How ants acclimate: impact of climatic
789 conditions on the cuticular hydrocarbon profile. Functional Ecology 32:657-666.
790 Meznar ER, Gadau J, Koeniger N, Rueppell O. 2010. Comparative linkage mapping suggests
791 a high recombination rate in all honeybees. Journal of Heredity 101:S118-S126.
792 Moris VC, Podsiadlowski L, Martin S, Oeyen JP, Donath A, Petersen M, Wilbrandt J, Misof
793 B, Liedtke D, Thamm M. 2023. Intrasexual cuticular hydrocarbon dimorphism in a
794 wasp sheds light on hydrocarbon biosynthesis genes in Hymenoptera. Communications
795 Biology 6:147.
796 Niehuis O, Buellesbach J, Gibson JD, Pothmann D, Hanner C, Mutti NS, Judson AK, Gadau
797 J, Ruther J, Schmitt T. 2013. Behavioural and genetic analyses of Nasonia shed light on
798 the evolution of sex pheromones. Nature 494:345-348.
799 Niehuis O, Bullesbach J, Judson AK, Schmitt T, Gadau J. 2011. Genetics of cuticular
800 hydrocarbon differences between males of the parasitoid wasps Nasonia giraulti and
801 Nasonia vitripennis. Heredity (Edinb) 107:61-70.
802 Nowbahari E, Lenoir A, Clément J, Lange C, Bagneres A, Joulie C. 1990. Individual,
803 geographical and experimental variation of cuticular hydrocarbons of the ant
804 Cataglyphis cursor (Hymenoptera: Formicidae): their use in nest and subspecies
805 recognition. Biochemical Systematics and Ecology 18:63-73.
806 Nunes T, Turatti I, Mateus S, Nascimento F, Lopes N, Zucchi R. 2009. Cuticular
807 hydrocarbons in the stingless bee Schwarziana quadripunctata (Hymenoptera, Apidae,
808 Meliponini): differences between colonies, castes and age. Genetics and Molecular
809 Research 8:589-595.
810 Ozaki M, Hefetz A. 2014. Neural Mechanisms and Information Processing in Recognition
811 Systems. Insects 5:722-741.
812 Ozaki M, Wada-Katsumata A, Fujikawa K, Iwasaki M, Yokohari F, Satoji Y, Nisimura T,
813 Yamaoka R. 2005. Ant nestmate and non-nestmate discrimination by a chemosensory
814 sensillum. Science 309:311-314.
815 Pickett JA, Birkett MA, Woodcock CM, Zhou J-J. 2009. Scents and sex: Insect pheromones.
816 The Biochemist 31:28-32.
817 Piek T. 1964. Synthesis of wax in the honeybee (Apis mellifera L.). Journal of insect
818 physiology 10:563-572.
819 Ronin YI, Mester DI, Minkov DG, Akhunov E, Korol AB. 2017. Building ultra-high-density
820 linkage maps based on efficient filtering of trustable markers. Genetics 206:1285-1295.
821 Silva KL, Batista NR, Antonialli-Junior WF. 2024. Effect of Temperature on Nestmate
822 Recognition in the Ant Odontomachus chelifer. Sociobiology 71:e9944-e9944.

Singer TL. 1998. Roles of hydrocarbons in the recognition systems of insects. American Zoologist 38:394-405.

Solignac M, Mougel F, Vautrin D, Monnerot M, Cornuet J-M. 2007. A third-generation microsatellite-based linkage map of the honey bee, Apis mellifera, and its comparison with the sequence-based physical map. Genome Biol 8:1-14.

Sprenger PP, Burkert LH, Abou B, Federle W, Menzel F. 2018. Coping with the climate: cuticular hydrocarbon acclimation of ants under constant and fluctuating conditions. Journal of Experimental Biology 221:jeb171488.

Sprenger PP, Hartke J, Schmitt T, Menzel F, Feldmeyer B. 2021. Candidate genes involved in cuticular hydrocarbon differentiation between cryptic, parabiotic ant species. G3 11:jkab078.

Sprenger PP, Menzel F. 2020. Cuticular hydrocarbons in ants (Hymenoptera: Formicidae) and other insects: how and why they differ among individuals, colonies, and species. Myrmecological News 30.

Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. 2006. AUGUSTUS: ab initio prediction of alternative transcripts. Nucleic acids research 34:W435-W439.

Steiner S, Hermann N, Ruther J. 2006. Characterization of a female-produced courtship pheromone in the parasitoid Nasonia vitripennis. Journal of Chemical Ecology 32:1687-1702.

Stuart RJ, Herbers JM. 2000. Nest mate recognition in ants with complex colonies: within- and between-population variation. Behavioral Ecology 11:676-685.

Sturgis SJ, Gordon DM. 2012. Nestmate recognition in ants (Hymenoptera: Formicidae): a review. Myrmecological News 16:101-110.

Team RC. 2021. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Thomas M, Simmons L. 2008. Cuticular hydrocarbons are heritable in the cricket Teleogryllus oceanicus. Journal of evolutionary biology 21:801-806.

Trabalon M, Plateaux L, Péru L, Bagnères A-G, Hartmann N. 2000. Modification of morphological characters and cuticular compounds in worker ants Leptothorax nylanderi induced by endoparasites Anomotaenia brevis. Journal of insect physiology 46:169-178.

Tranter C, LeFevre L, Evison SE, Hughes WO. 2015. Threat detection: contextual recognition and response to parasites by ants. Behavioral Ecology 26:396-405.

Van Zweden J, Brask JB, Christensen JH, Boomsma JJ, Linksvayer TA, D'ettorre P. 2010. Blending of heritable recognition cues among ant nestmates creates distinct colony gestalt odours but prevents within-colony nepotism. Journal of evolutionary biology 23:1498-1508.

van Zweden JS, d'Ettorre P. 2010. estmate recognition in social insects and the role of hydrocarbons. In: Blomquist GJ, Bagnères A-G, editors. Insect hydrocarbons: biology, biochemistry and chemical ecology. p. 222-243.

van Zweden JS, Dreier S, d'Ettorre P. 2009. Disentangling environmental and heritable nestmate recognition cues in a carpenter ant. Journal of insect physiology 55:159-164.

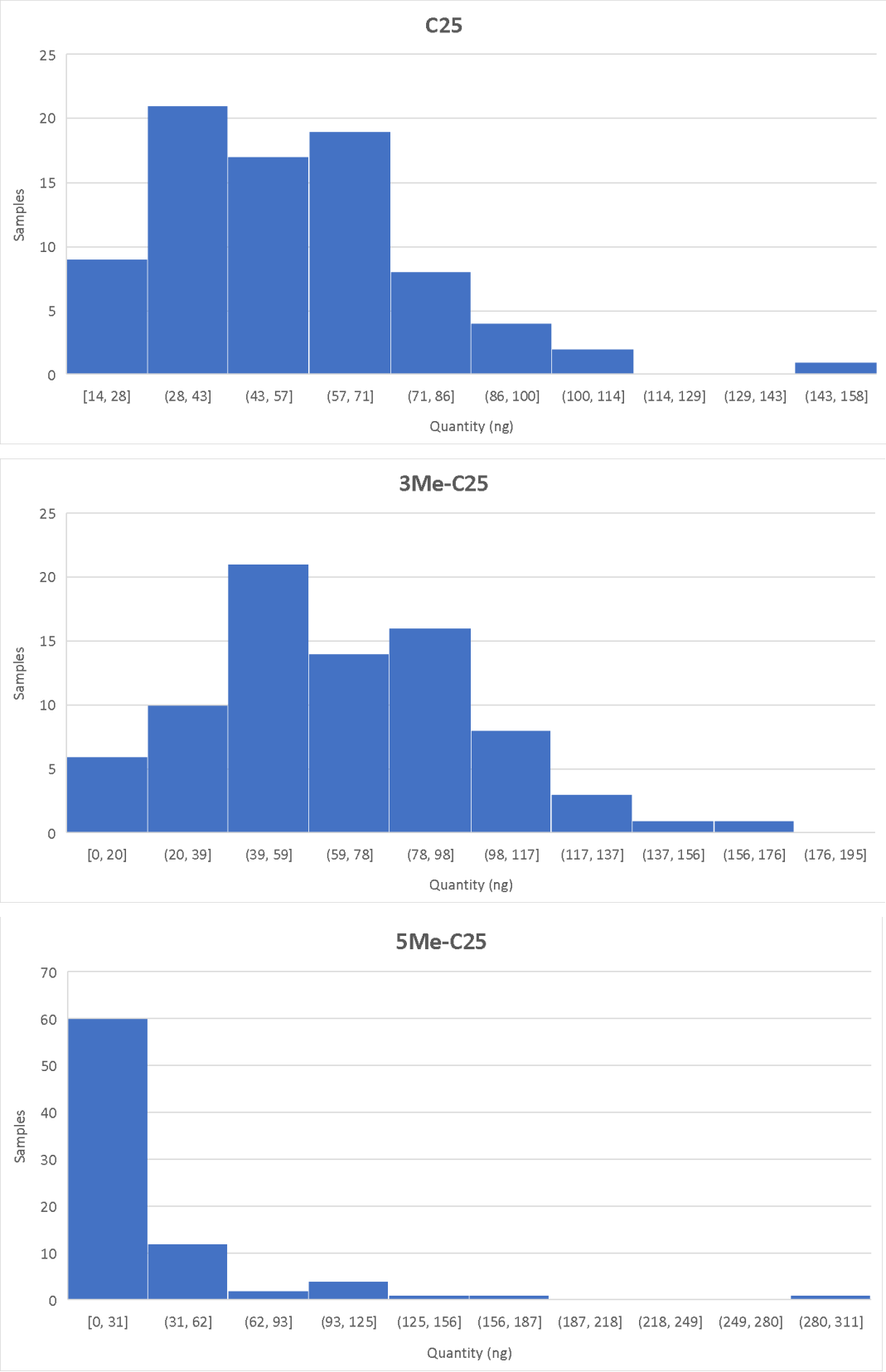
Walsh J, Pontieri L, d'Ettorre P, Linksvayer TA. 2020. Ant cuticular hydrocarbons are heritable and associated with variation in colony productivity. Proc Biol Sci 287:20201029.

Wen H, Ma X, Zhang W, Wang L, Li Y, Lv D, Sun J, Mang D. 2025. A fructose-sensitive gustatory receptor links nutrient sensing to lipid metabolism in the red imported fire ant. Insect Biochemistry and Molecular Biology:104440.

Yahav T, Privman E. 2019. A comparative analysis of methods for de novo assembly of hymenopteran genomes using either haploid or diploid samples. Sci Rep 9:6480.

875 **Supplementary**

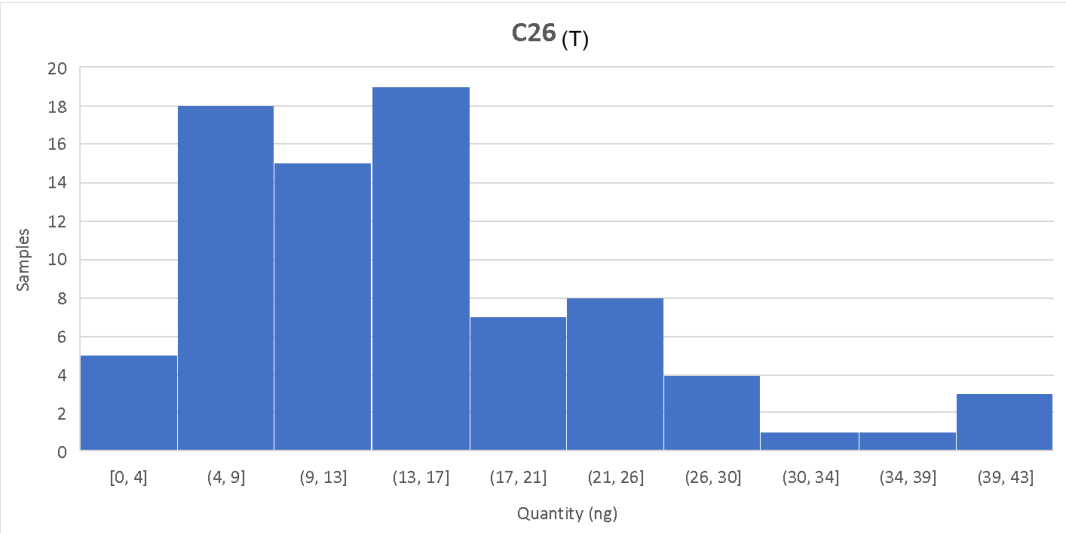
876 **Supplementary Figure S1: Distributions of the absolute quantity of each CHC in the**
877 **samples. CHCs marked with (T) were log-transformed.**



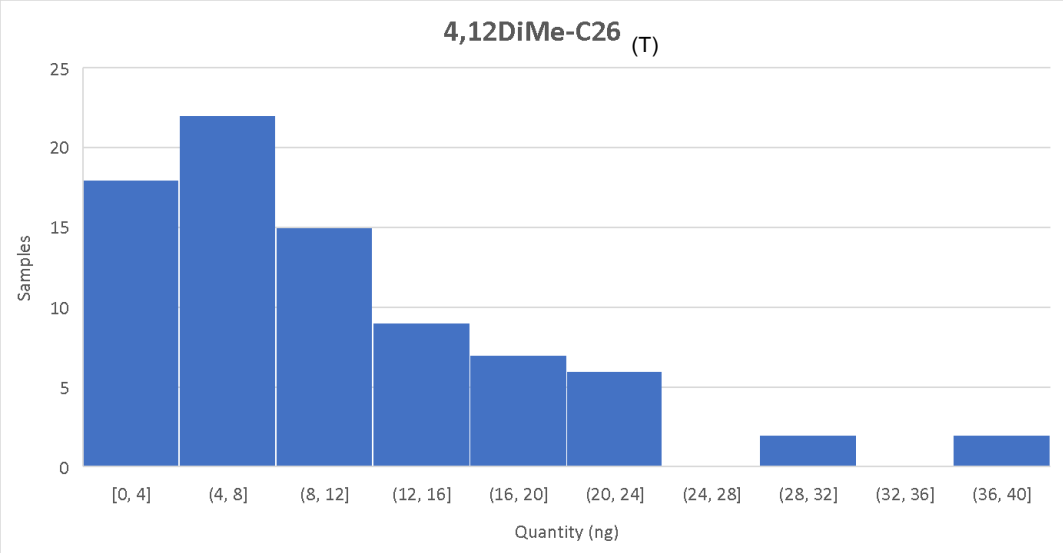
878

879

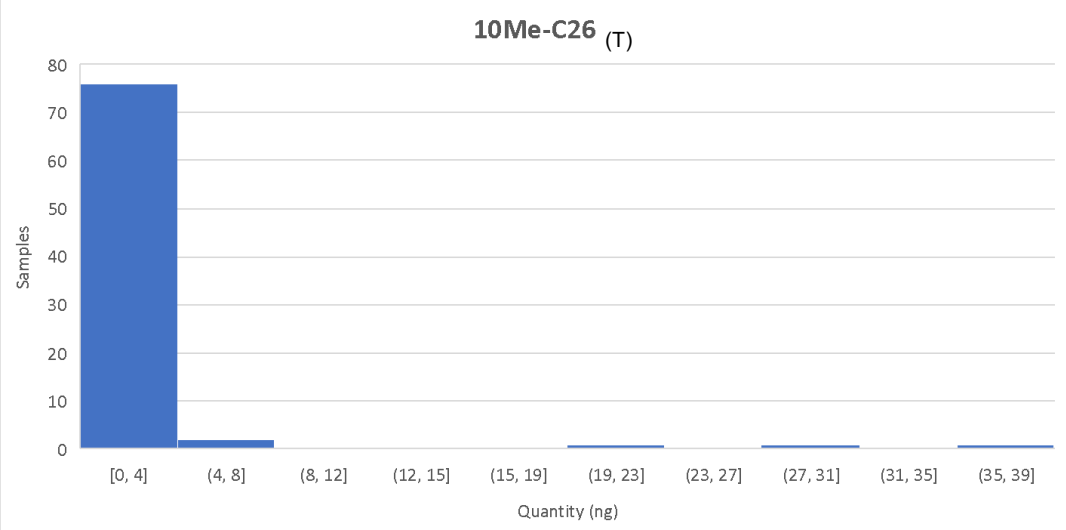
880


C26 (T)
Samples
20
18
16
14
12
10
8
6
4
2
0
[0, 4]
(4, 9]
(9, 13]
(13, 17]
(17, 21]
(21, 26]
(26, 30]
(30, 34]
(34, 39]
(39, 43]
Quantity (ng)

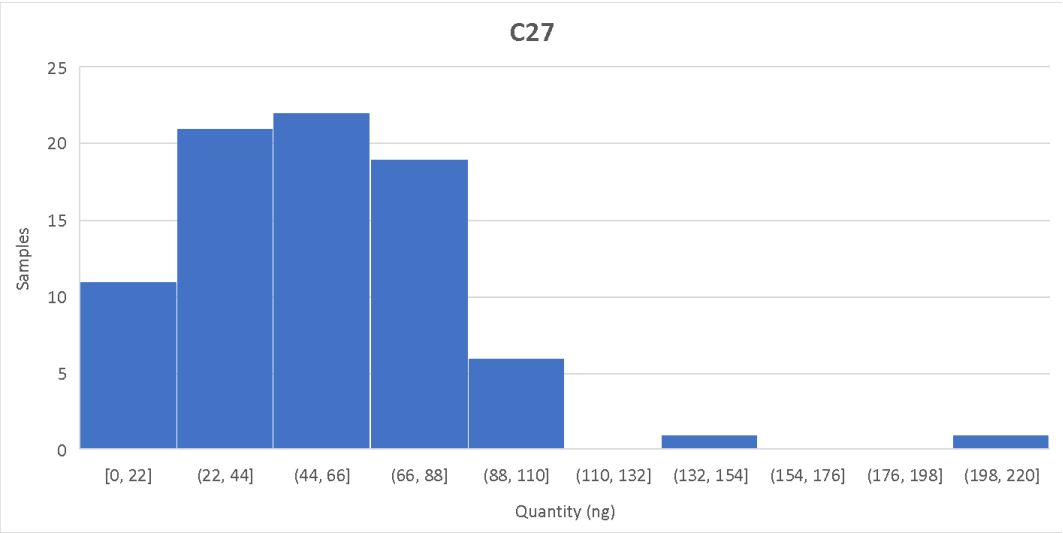
881


4,12DiMe-C26 (T)
Samples
25
20
15
10
5
0
[0, 4]
(4, 8]
(8, 12]
(12, 16]
(16, 20]
(20, 24]
(24, 28]
(28, 32]
(32, 36]
(36, 40]
Quantity (ng)

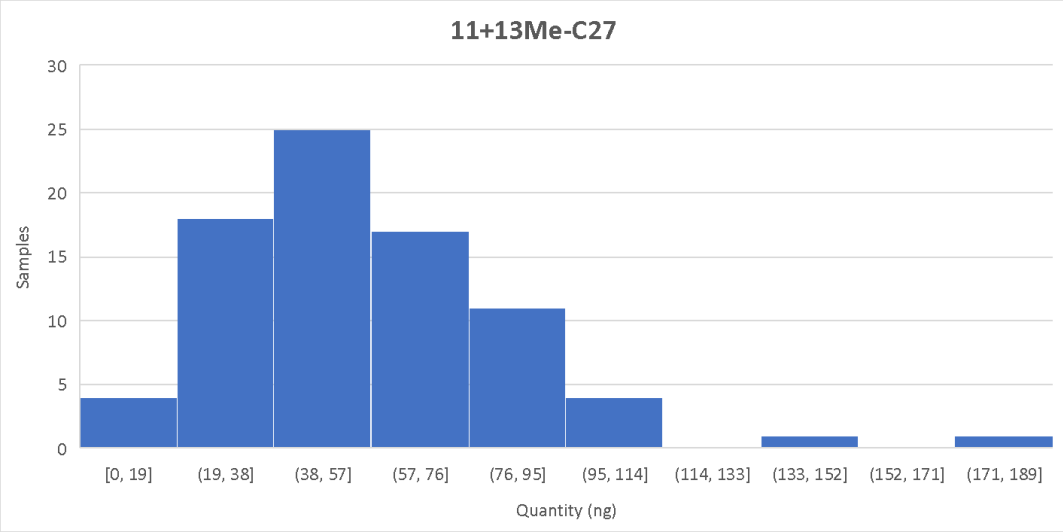
882


10Me-C26 (T)
Samples
80
70
60
50
40
30
20
10
0
[0, 4]
(4, 8]
(8, 12]
(12, 15]
(15, 19]
(19, 23]
(23, 27]
(27, 31]
(31, 35]
(35, 39]
Quantity (ng)

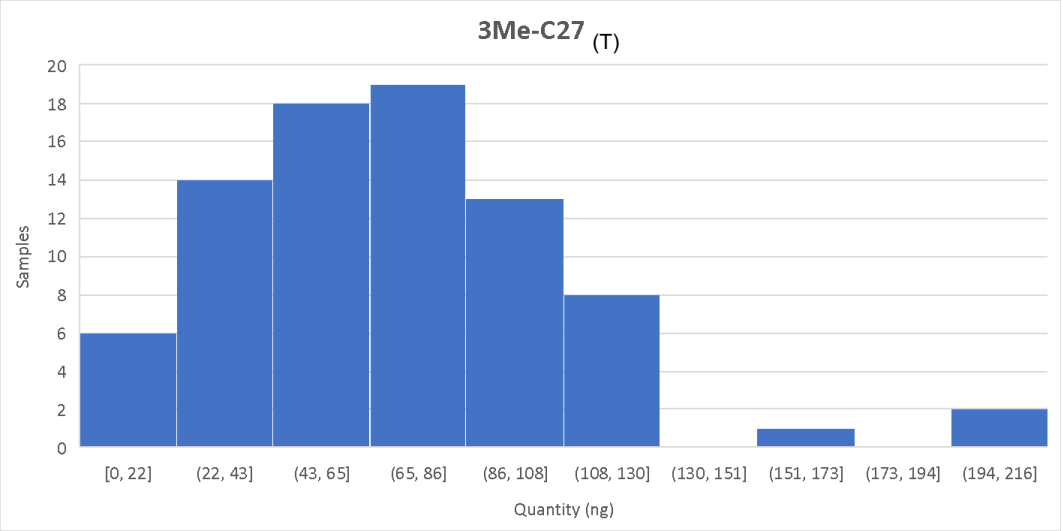
883


C27
Samples
25
20
15
10
5
0
[0, 22]
(22, 44]
(44, 66]
(66, 88]
(88, 110]
(110, 132]
(132, 154]
(154, 176]
(176, 198]
(198, 220]
Quantity (ng)

884

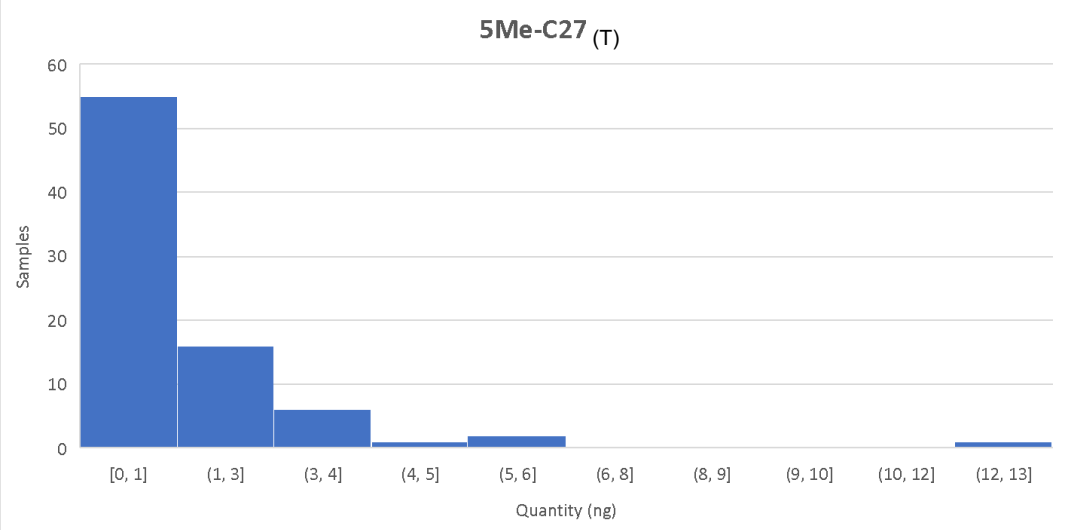

11+13Me-C27
Samples
30
25
20
15
10
5
0
[0, 19]
(19, 38]
(38, 57]
(57, 76]
(76, 95]
(95, 114]
(114, 133]
(133, 152]
(152, 171]
(171, 189]
Quantity (ng)

885

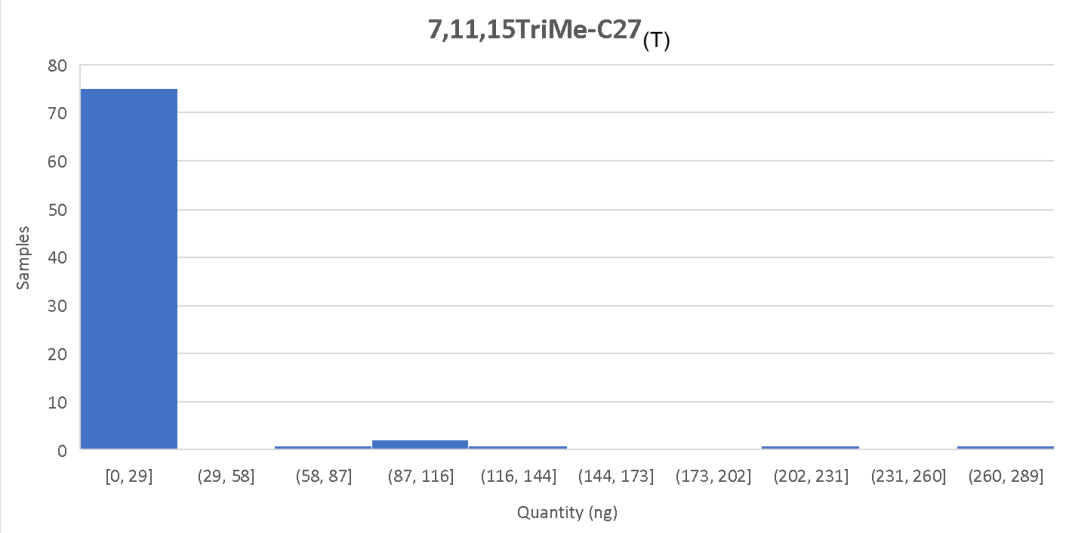

3Me-C27 (T)
Samples
20
18
16
14
12
10
8
6
4
2
0
[0, 22]
(22, 43]
(43, 65]
(65, 86]
(86, 108]
(108, 130]
(130, 151]
(151, 173]
(173, 194]
(194, 216]
Quantity (ng)

886

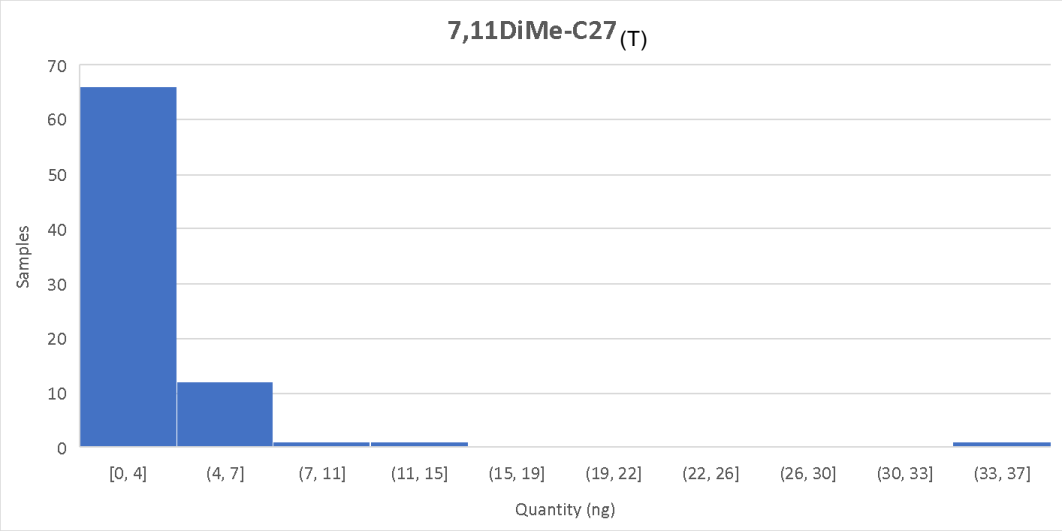
887

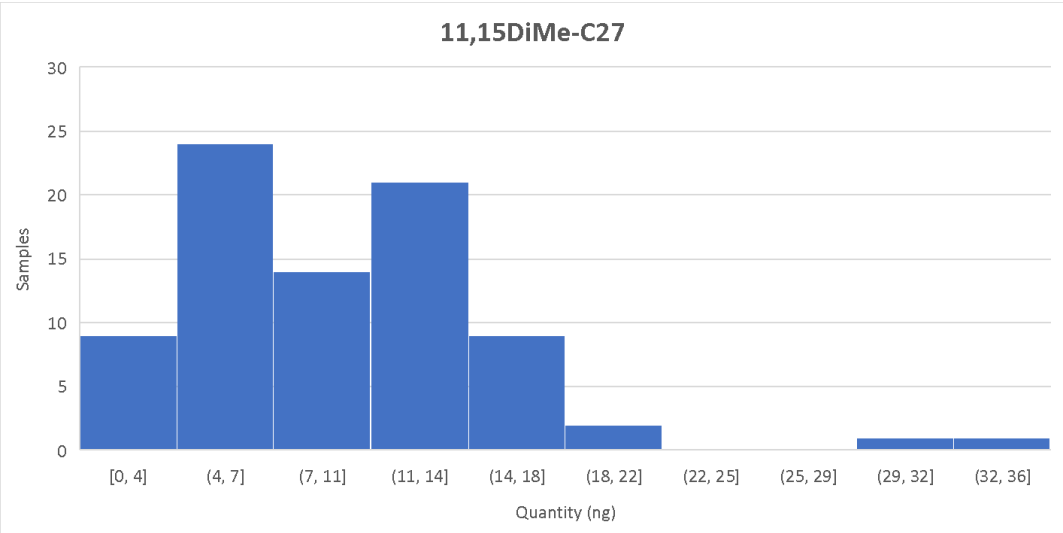

5Me-C27 (T)
Samples
60
50
40
30
20
10
0
[0, 1]
(1, 3]
(3, 4]
(4, 5]
(5, 6]
(6, 8]
(8, 9]
(9, 10]
(10, 12]
(12, 13]
Quantity (ng)

888

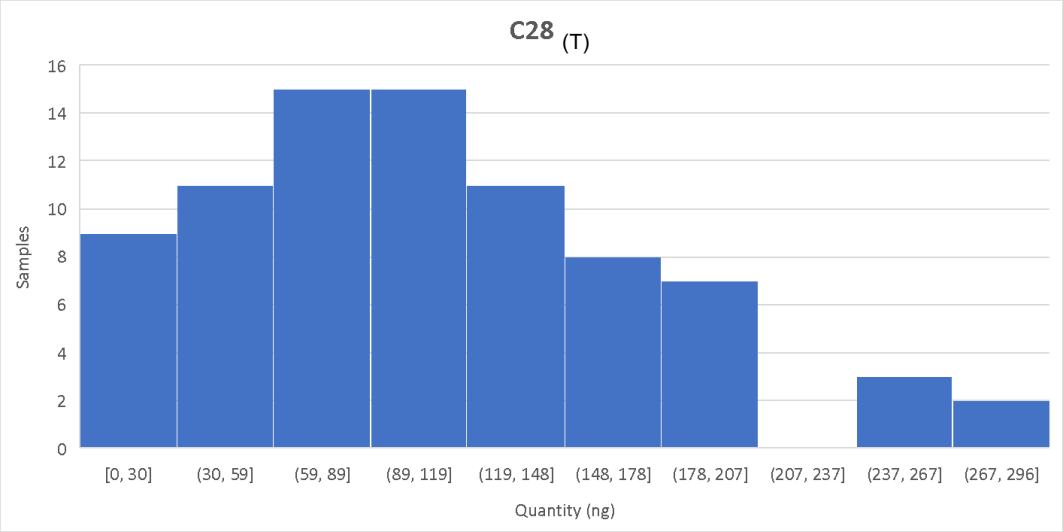

7,11,15TriMe-C27 (T)
Samples
80
70
60
50
40
30
20
10
0
[0, 29]
(29, 58]
(58, 87]
(87, 116]
(116, 144]
(144, 173]
(173, 202]
(202, 231]
(231, 260]
(260, 289]
Quantity (ng)

889

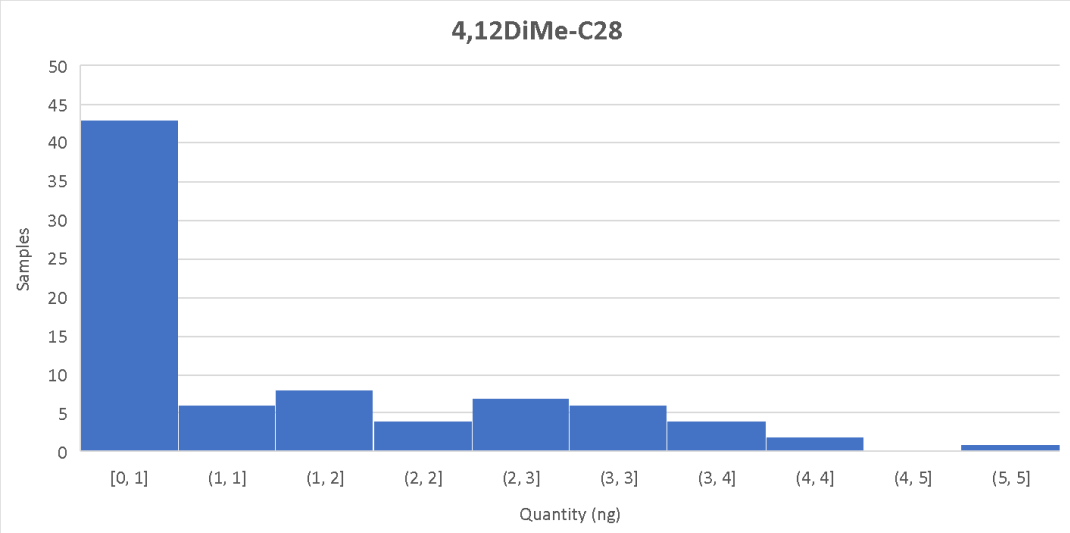

7,11DiMe-C27 (T)
Samples
70
60
50
40
30
20
10
0
[0, 4]
(4, 7]
(7, 11]
(11, 15]
(15, 19]
(19, 22]
(22, 26]
(26, 30]
(30, 33]
(33, 37]
Quantity (ng)


11,15DiMe-C27
Samples
30
25
20
15
10
5
0
[0, 4]
(4, 7]
(7, 11]
(11, 14]
(14, 18]
(18, 22]
(22, 25]
(25, 29]
(29, 32]
(32, 36]
Quantity (ng)

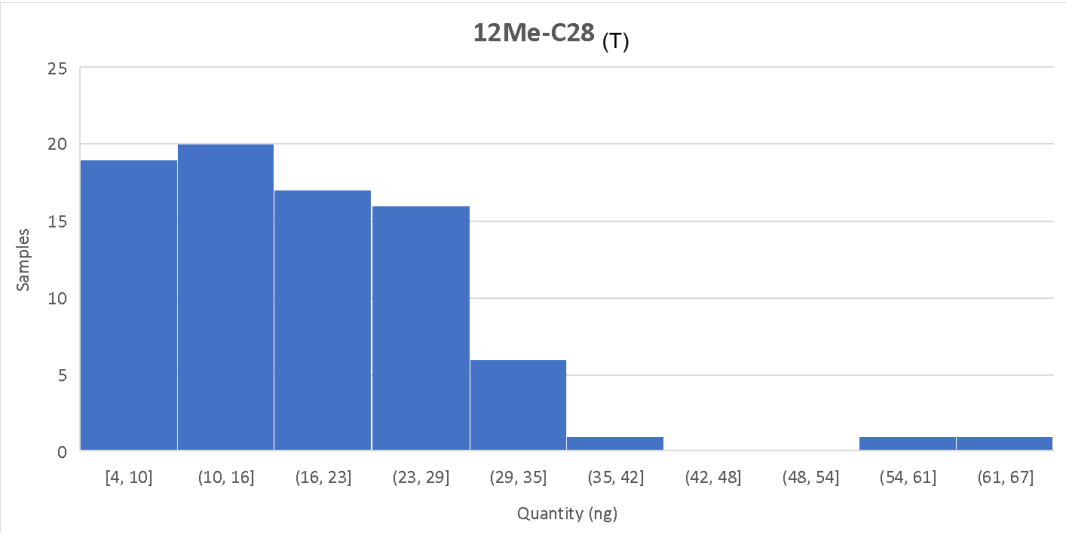
890

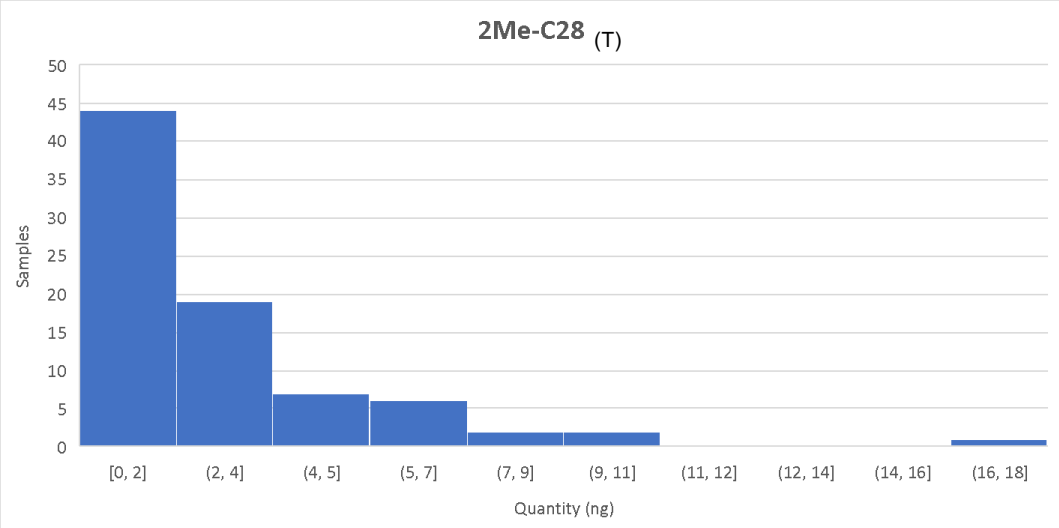

C28 (T)
Samples
16
14
12
10
8
6
4
2
0
[0, 30]
(30, 59]
(59, 89]
(89, 119]
(119, 148]
(148, 178]
(178, 207]
(207, 237]
(237, 267]
(267, 296]
Quantity (ng)

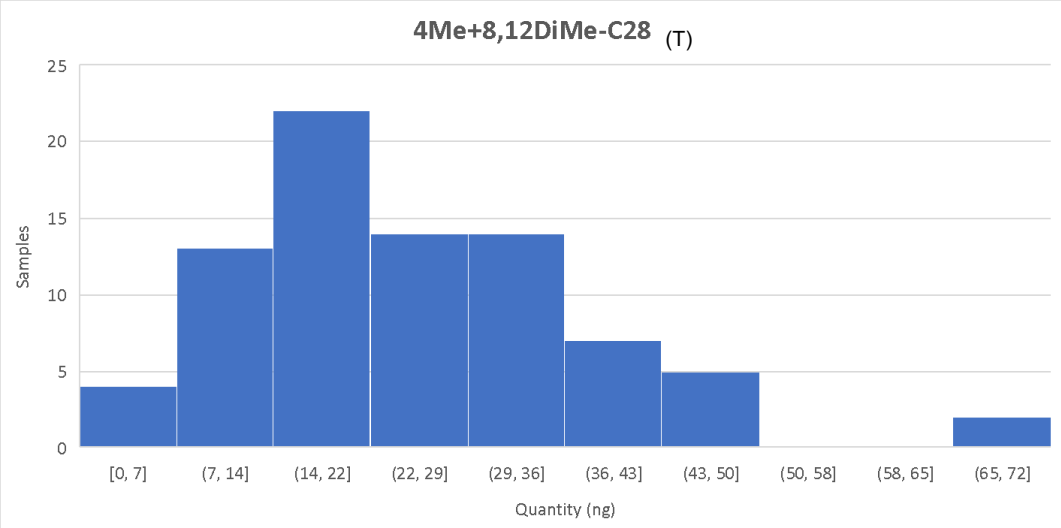
891

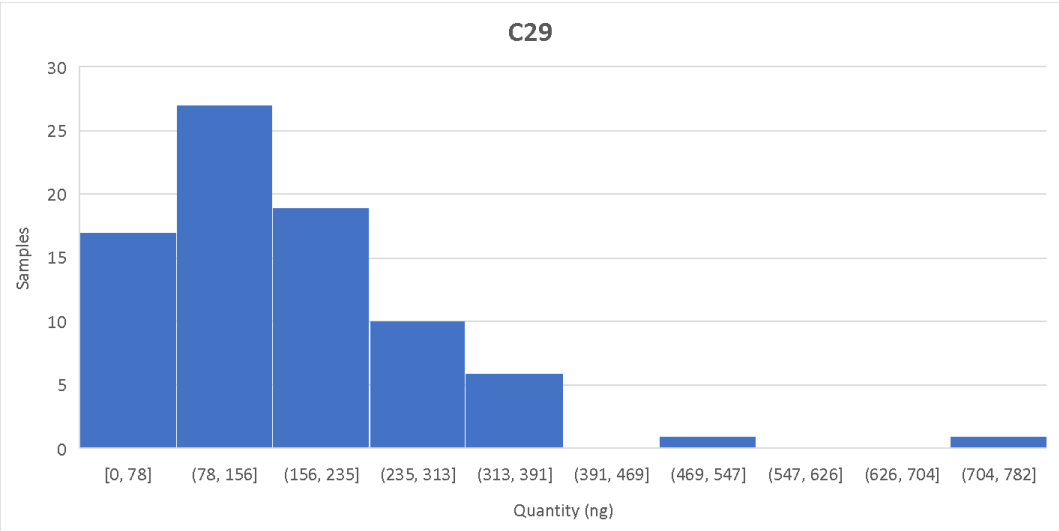

4,12DiMe-C28
Samples
50
45
40
35
30
25
20
15
10
5
0
[0, 1]
(1, 1]
(1, 2]
(2, 2]
(2, 3]
(3, 3]
(3, 4]
(4, 4]
(4, 5]
(5, 5]
Quantity (ng)

892

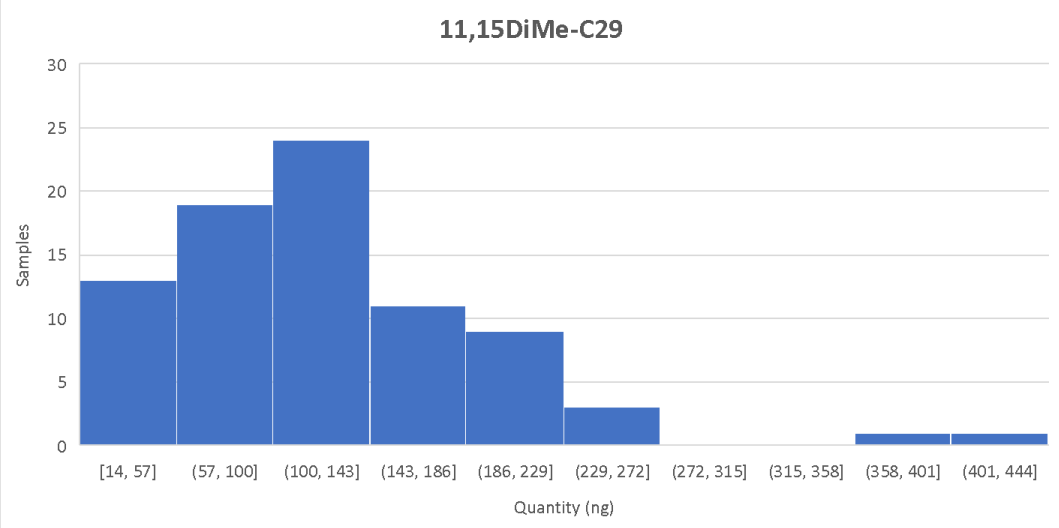

12Me-C28 (T)
Samples
25
20
15
10
5
0
[4, 10]
(10, 16]
(16, 23]
(23, 29]
(29, 35]
(35, 42]
(42, 48]
(48, 54]
(54, 61]
(61, 67]
Quantity (ng)


2Me-C28 (T)
Samples
50
45
40
35
30
25
20
15
10
5
0
[0, 2]
(2, 4]
(4, 5]
(5, 7]
(7, 9]
(9, 11]
(11, 12]
(12, 14]
(14, 16]
(16, 18]
Quantity (ng)

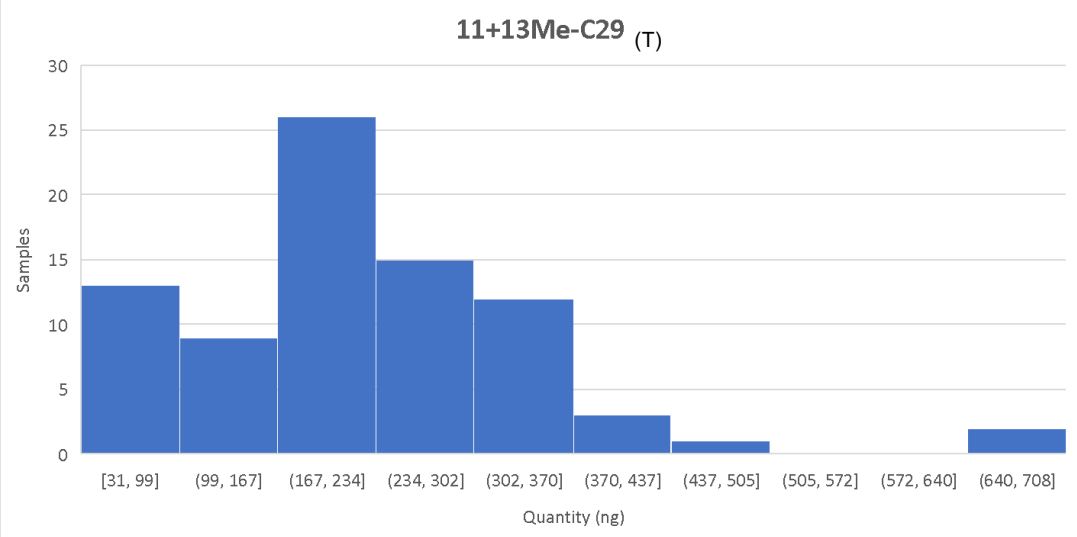

4Me+8,12DiMe-C28 (T)
Samples
25
20
15
10
5
0
[0, 7]
(7, 14]
(14, 22]
(22, 29]
(29, 36]
(36, 43]
(43, 50]
(50, 58]
(58, 65]
(65, 72]
Quantity (ng)


C29
Samples
30
25
20
15
10
5
0
[0, 78]
(78, 156]
(156, 235]
(235, 313]
(313, 391]
(391, 469]
(469, 547]
(547, 626]
(626, 704]
(704, 782]
Quantity (ng)

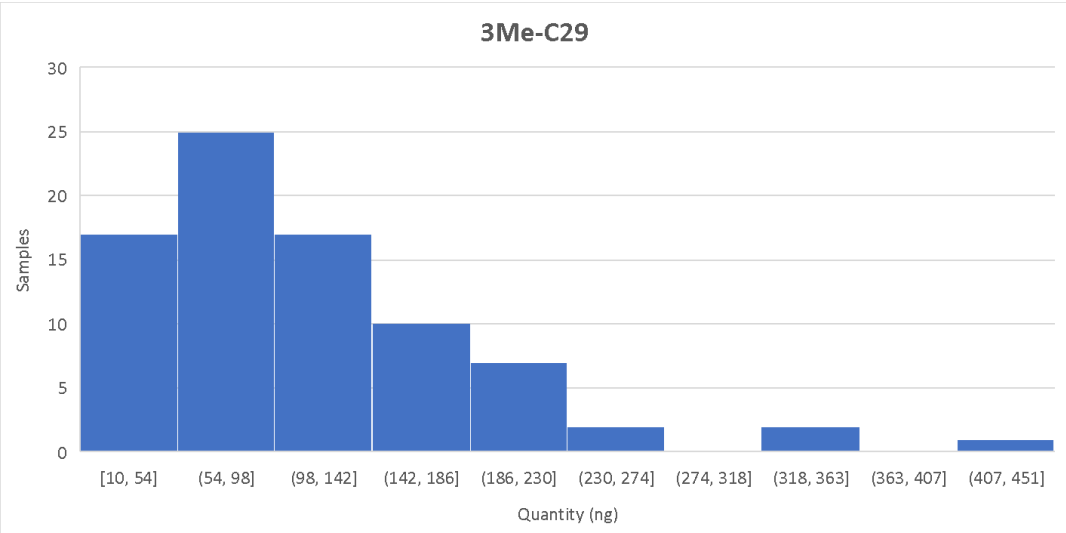
896

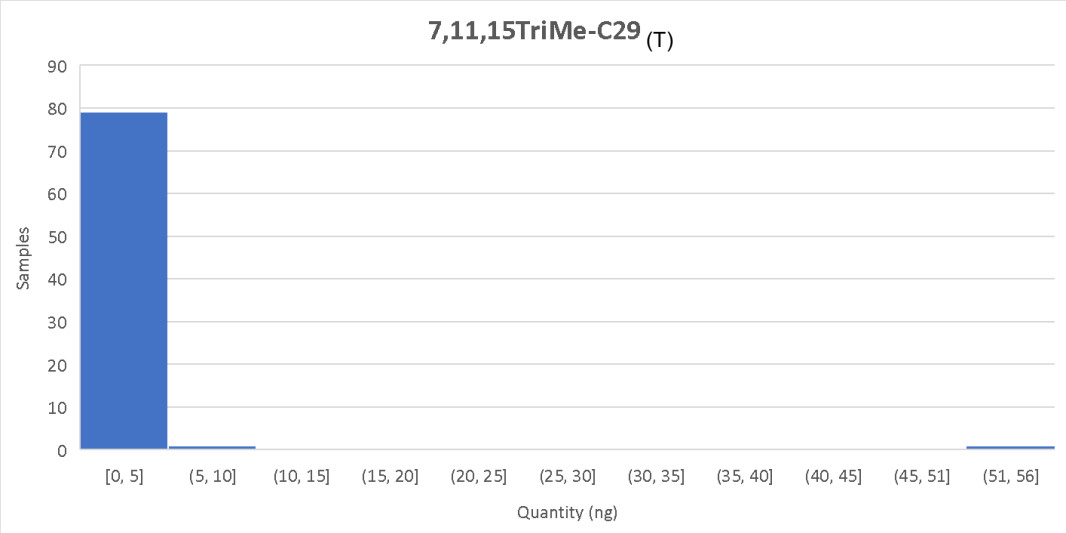

11,15DiMe-C29
Samples
30
25
20
15
10
5
0
[14, 57]
(57, 100]
(100, 143]
(143, 186]
(186, 229]
(229, 272]
(272, 315]
(315, 358]
(358, 401]
(401, 444]
Quantity (ng)

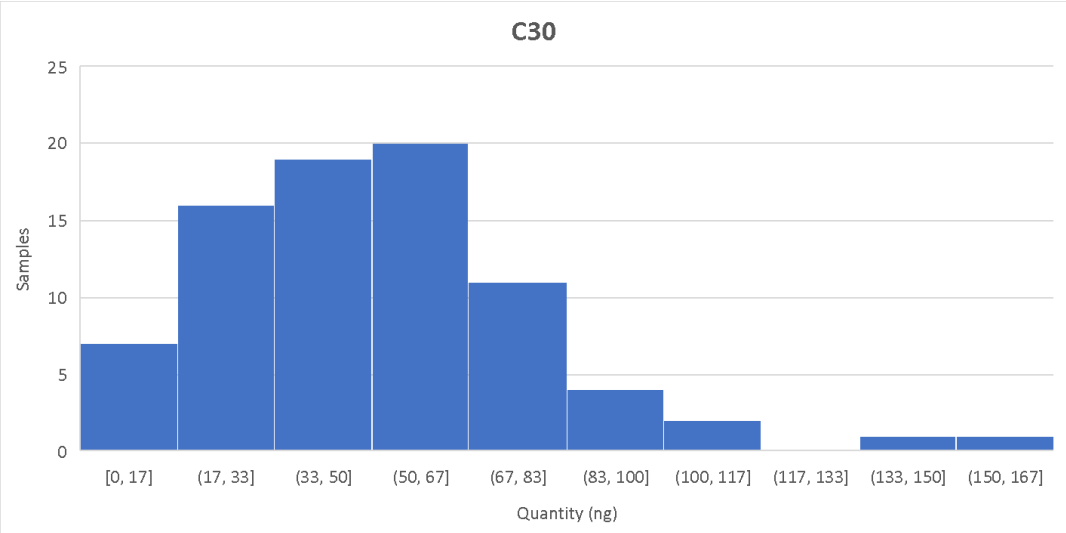
897

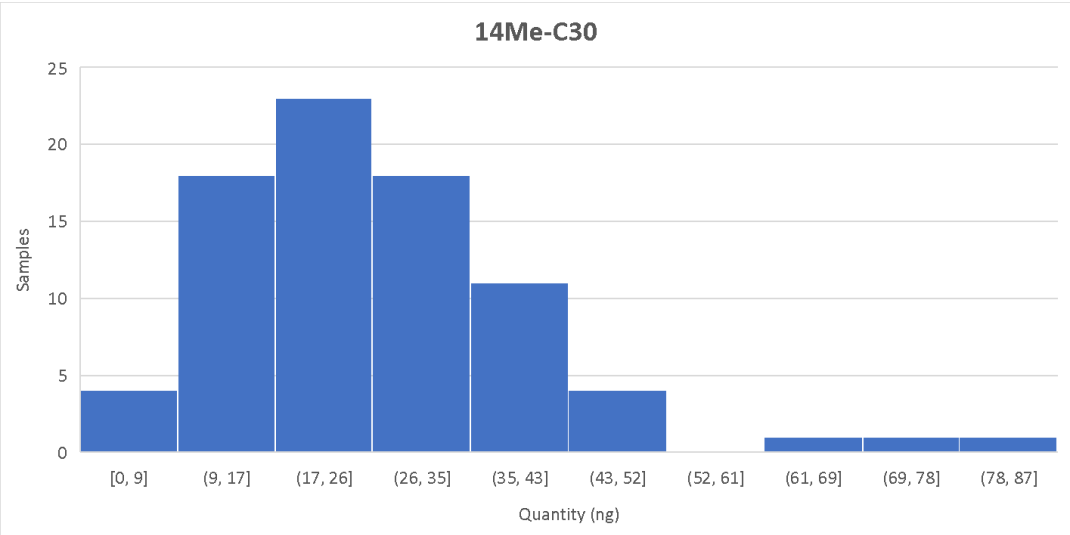

11+13Me-C29 (T)
Samples
30
25
20
15
10
5
0
[31, 99]
(99, 167]
(167, 234]
(234, 302]
(302, 370]
(370, 437]
(437, 505]
(505, 572]
(572, 640]
(640, 708]
Quantity (ng)

898

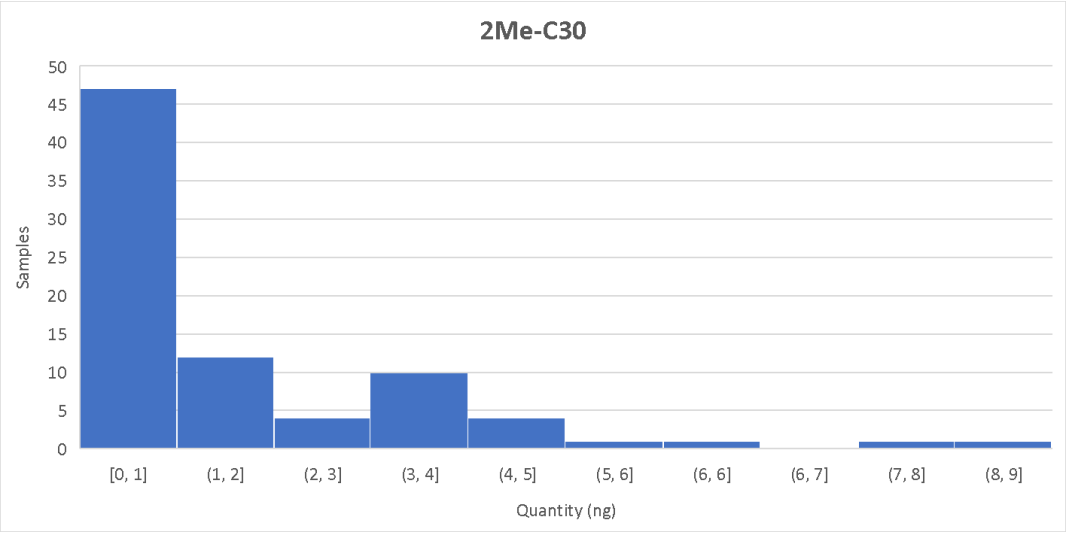

3Me-C29
Samples
30
25
20
15
10
5
0
[10, 54]
(54, 98]
(98, 142]
(142, 186]
(186, 230]
(230, 274]
(274, 318]
(318, 363]
(363, 407]
(407, 451]
Quantity (ng)


7,11,15TriMe-C29 (T)
Samples
90
80
70
60
50
40
30
20
10
0
[0, 5]
(5, 10]
(10, 15]
(15, 20]
(20, 25]
(25, 30]
(30, 35]
(35, 40]
(40, 45]
(45, 51]
(51, 56]
Quantity (ng)

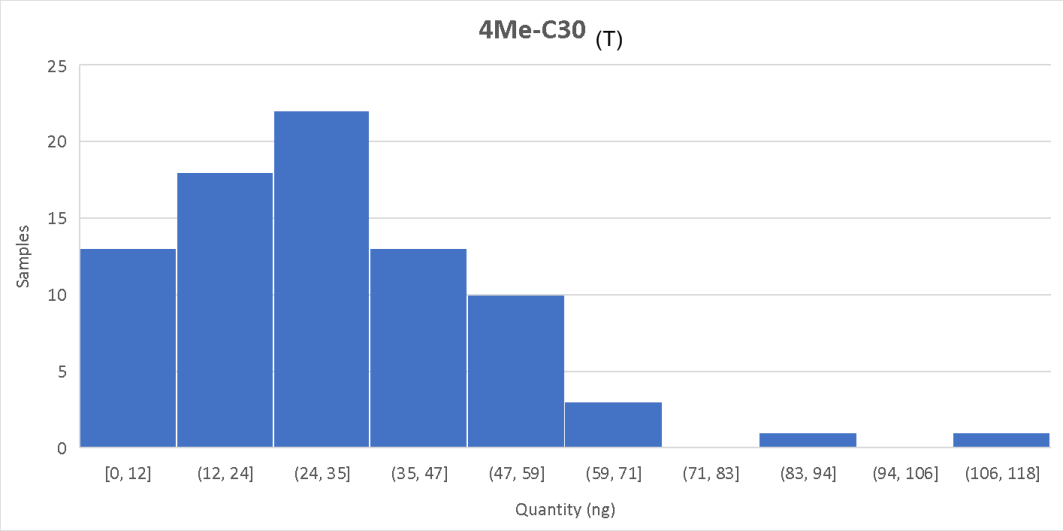

C30
Samples
25
20
15
10
5
0
[0, 17]
(17, 33]
(33, 50]
(50, 67]
(67, 83]
(83, 100]
(100, 117]
(117, 133]
(133, 150]
(150, 167]
Quantity (ng)


14Me-C30
Samples
25
20
15
10
5
0
[0, 9]
(9, 17]
(17, 26]
(26, 35]
(35, 43]
(43, 52]
(52, 61]
(61, 69]
(69, 78]
(78, 87]
Quantity (ng)

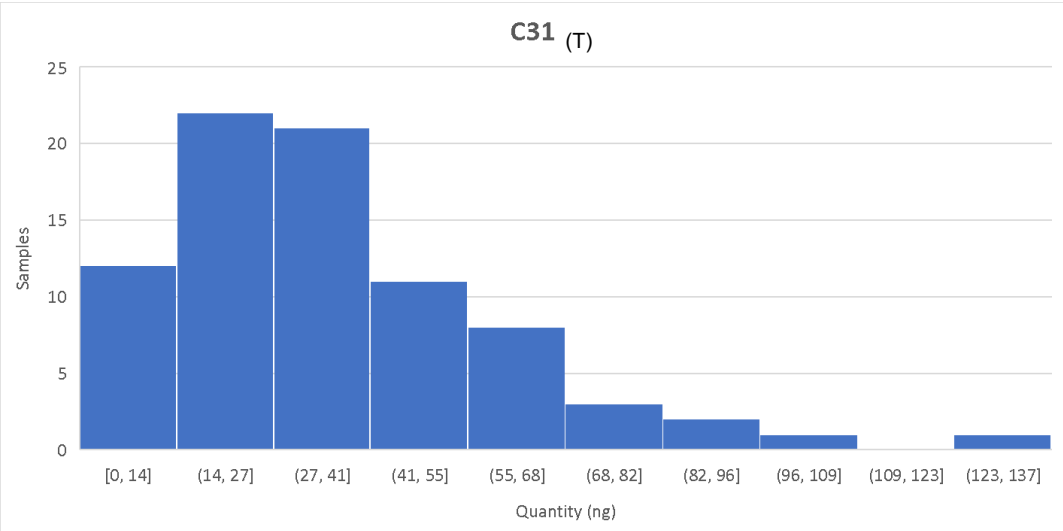
902


2Me-C30
Samples
50
45
40
35
30
25
20
15
10
5
0
[0, 1]
(1, 2]
(2, 3]
(3, 4]
(4, 5]
(5, 6]
(6, 6]
(6, 7]
(7, 8]
(8, 9]
Quantity (ng)

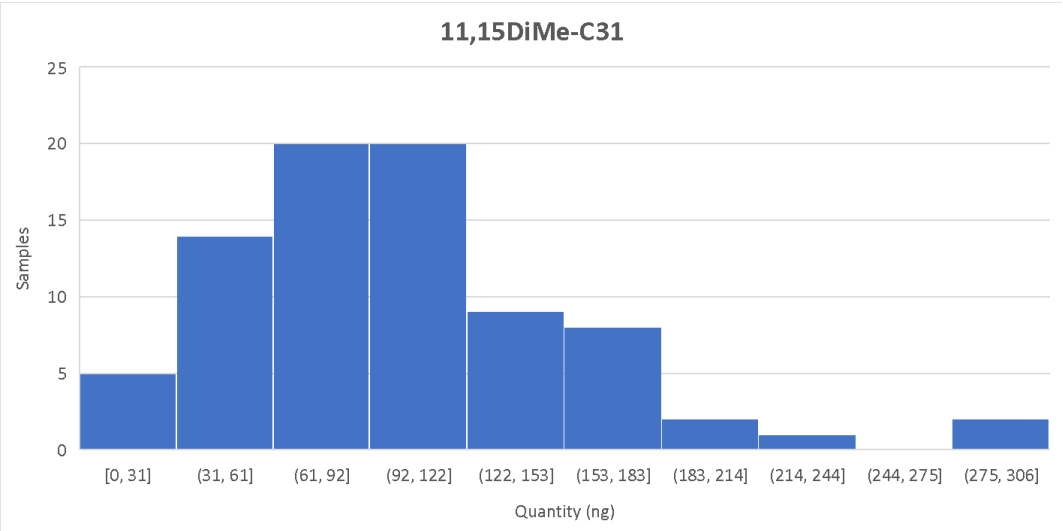
903


4Me-C30 (T)
Samples
25
20
15
10
5
0
[0, 12]
(12, 24]
(24, 35]
(35, 47]
(47, 59]
(59, 71]
(71, 83]
(83, 94]
(94, 106]
(106, 118]
Quantity (ng)

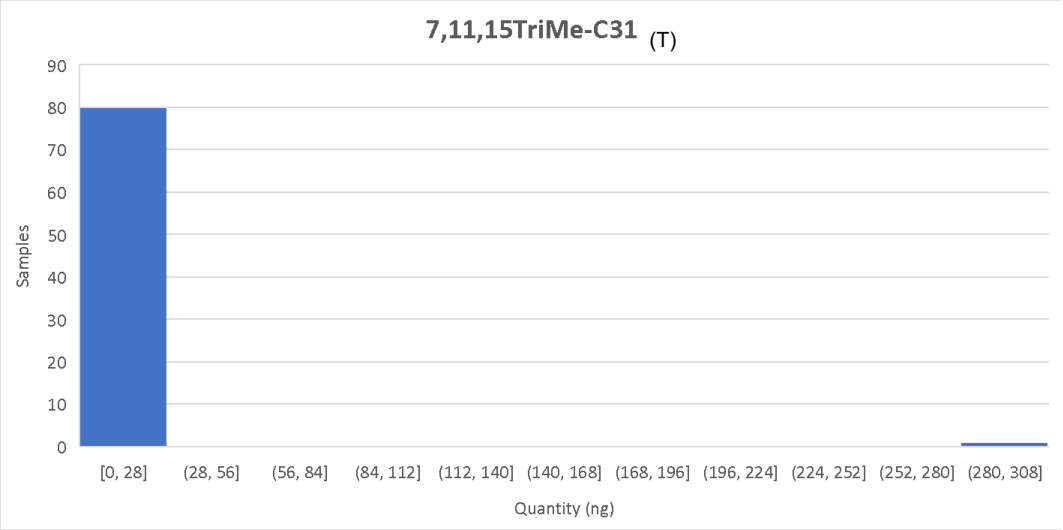
904


C31 (T)
Samples
25
20
15
10
5
0
[0, 14]
(14, 27]
(27, 41]
(41, 55]
(55, 68]
(68, 82]
(82, 96]
(96, 109]
(109, 123]
(123, 137]
Quantity (ng)

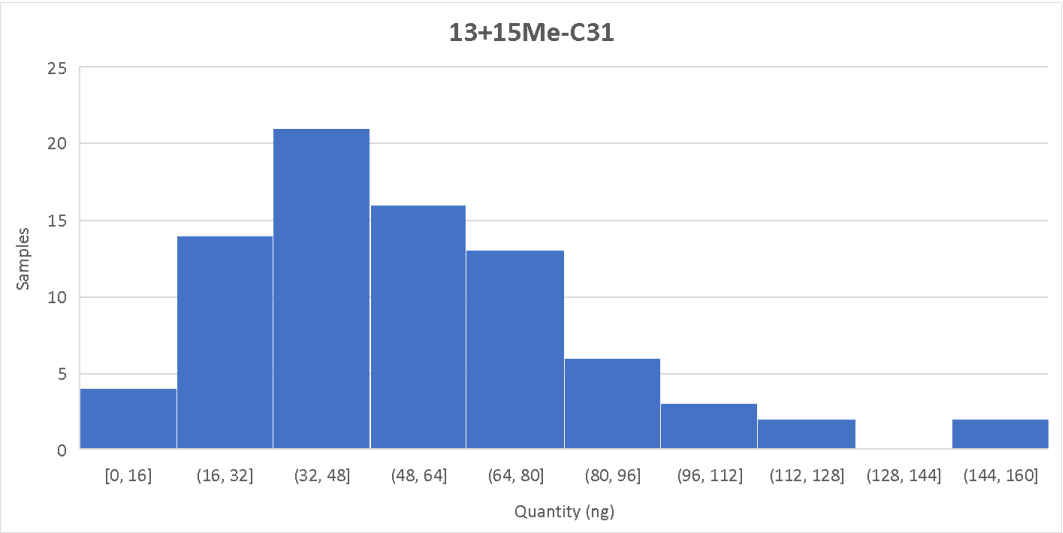
905

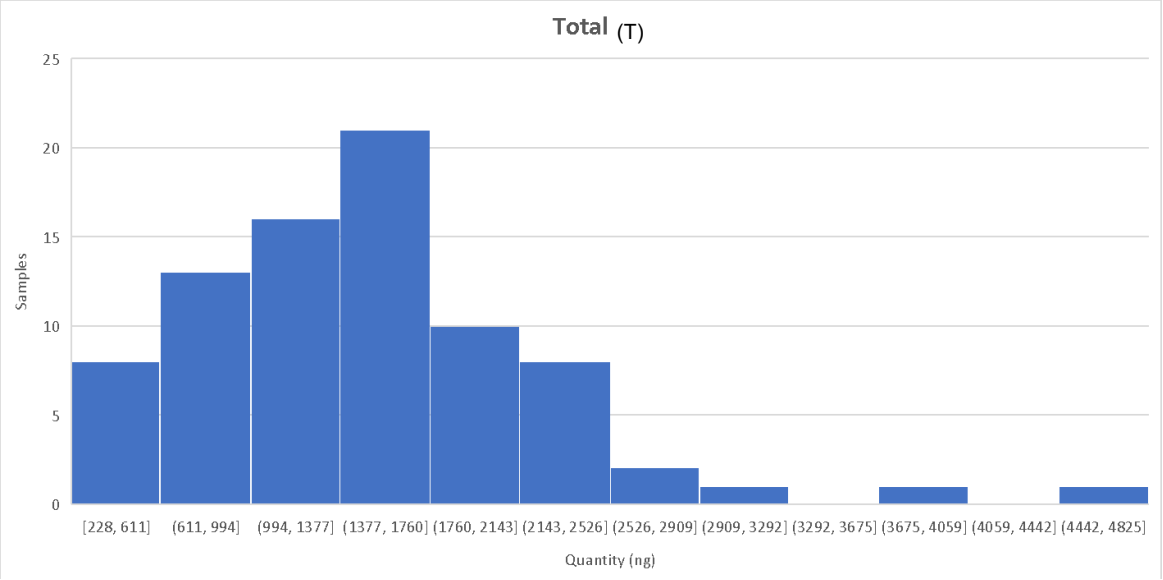

11,15DiMe-C31
Samples
25
20
15
10
5
0
[0, 31]
(31, 61]
(61, 92]
(92, 122]
(122, 153]
(153, 183]
(183, 214]
(214, 244]
(244, 275]
(275, 306]
Quantity (ng)

906

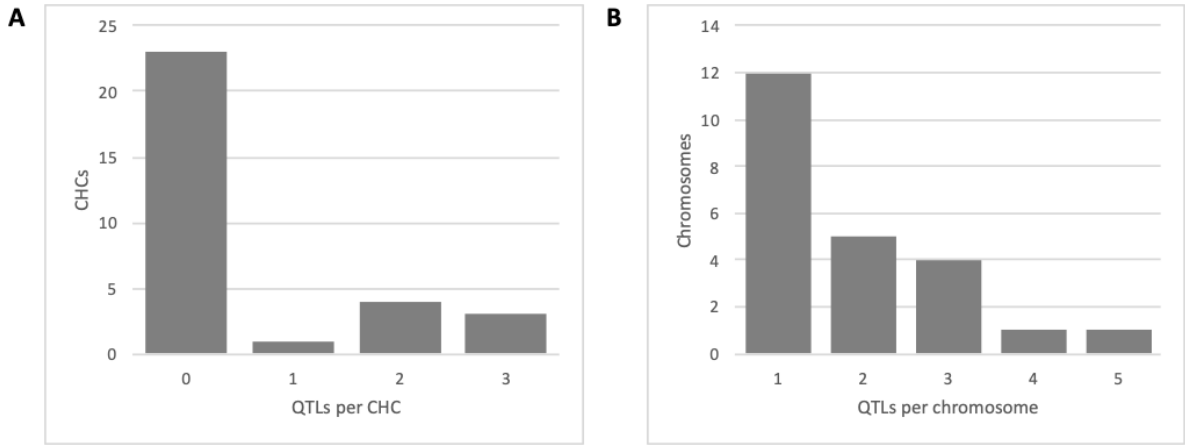

7,11,15TriMe-C31 (T)
Samples
90
80
70
60
50
40
30
20
10
0
[0, 28]
(28, 56]
(56, 84]
(84, 112]
(112, 140]
(140, 168]
(168, 196]
(196, 224]
(224, 252]
(252, 280]
(280, 308]
Quantity (ng)

907

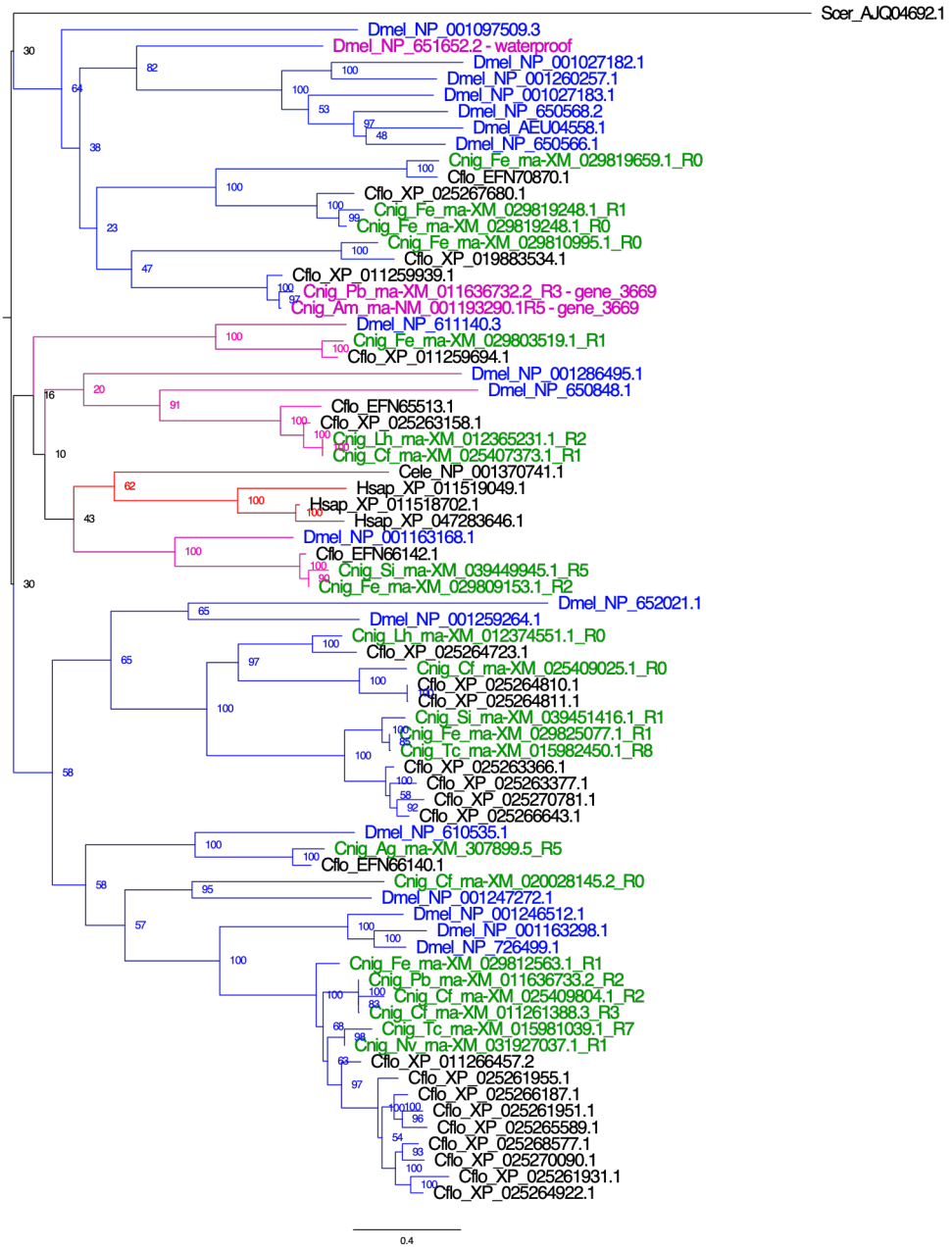




Supplementary Figure S2: Distribution of the number of QTLs per CHC (A) and QTLs per chromosome (B).



Supplementary Figure S3: Fatty acyl-CoA reductase gene tree. Including fly (*Drosophila melanogaster*) and ant (*Cataglyphis niger* and *Camponotus floridanus*) genes. The *Drosophila waterproof* gene and other fly paralogs are the closest homologs in *Drosophila* to *C. niger* gene 3669 in the QTL for C27. The gene names are colored by species and subfamilies that existed in the common ancestor of ants and flies are colored in distinct colors.



922 **Supplementary Table S1: QTLs underlying variation in CHCs (full table of tested**
923 **candidate QTLs, including non-significant results)**

CHC	LOD	Location (map intervals)	Location (cM)	P.E.V	*p*-value	*q*-value	MIM P.E.V
C25	6.589	Chr6: 1-3-5	0-21.06-43.20	19.41%	0.001	0.069	29.70%
	linked*	Chr6: 19-22-25	76.86-81.62-88.75	linked*	linked*	linked*	
	2.710	Chr8: 15-17-22	58.38-81.91-101.32	10.6%	0.013	0.318	
3Me-C25	2.834	Chr5: 4-8-16	30.89-46.11-66.91	16.8%	0.006	0.207	30.1%
	1.859	Chr1: 50-57-60	386.10-406.87-431.58	13.0%	0.091	1.000	
5Me-C25	2.523	Chr8: 49-50-50	221.43-244.92-244.92	14.9%	0.010	0.301	N/A
C26	4.215	Chr6: 13-16-18	56.40-65.36-67.36	16.2%	0.001	0.069	46.1%
	4.168	Chr17: 17-19-21	60.99-72.57-75.45	16.0%	0.000	0.023	
	3.590	Chr18: 1-1-5	0-0.0-36.50	13.8%	0.001	0.062	
4,12DiMe-C26	2.710	Chr18: 36-39-43	111.63-125.97-130.72	8.7%	0.012	0.318	48.6%
	7.069	Chr3: 1-2-4	0-12.84-37.91	39.9%	0.003	0.125	
	linked*	Chr3: 50-52-56	294.10-313.03-333.81	linked*	linked*	linked*	
4Me-C26	3.349	Chr11: 44-45-48	155.99-172.36-183.23	14.5%	0.003	0.109	35.5%
	2.402	Chr10: 38-43-49	108.13-127.65-154.33	10.4%	0.028	0.503	
	2.507	Chr18: 40-47-52	125.97-143.39-164.19	10.6%	0.018	0.398	
10Me-C26	3.691	Chr9: 45-50-54	190.30-213.93-225.26	16.0%	0.014	0.333	N/A
C27	3.899	Chr17: 25-29-32	87.01-97.50-111.93	17.5%	0.002	0.078	35.4%
	3.981	Chr18: 31-33-37	102.15-108.7-113.54	18.0%	0.000	0.023	
7,11DiMe-C27	3.160	Chr22: 1-1-7	0-0.0-11.61	18.2%	0.004	0.158	N/A
11,15DiMe-C27	2.453	Chr7: 3-4-7	20.58-32.56-44.27	12.4%	0.018	0.398	24.3%
	2.283	Chr22: 1-1-8	0-0.0-13/74	11.8%	0.022	0.447	
3Me-C27	2.250	Chr25: 20-23-28	81.78-106.66-141.64	12.8%	0.018	0.398	N/A
5Me-C27	2.341	Chr14: 1-1-6	0-2.709-19.28	13.6%	0.024	0.484	N/A
11+13Me-C27	2.966	Chr12: 41-43-47	123.83-137.66-145.54	14.8%	0.009	0.283	49.4%
	1.952	Chr10: 5-9-15	27.08-52.66-60.07	10.5%	0.064	0.951	
	1.922	Chr19: 44-46-49	124.68-139.08-143.78	10.5%	0.064	0.951	
	2.760	Chr22: 3-6-10	3.906-11.61-18.61	13.6%	0.011	0.301	
7,11,15TriMe-C27	4.139	Chr15: 43-47-48	122.95-147.06-154.68	15.0%	0.001	0.064	49.4%
	4.810	Chr3: 38-39-42	207.63-216.96-228.07	13.7%	0.000	0.023	
	4.120	Chr6: 41-44-46	191.72-227.46-240.65	12.8%	0.000	0.023	
	4.725	Chr4: 4-8-12	19.14-35.30-45.95	15.9%	0.011	0.301	
	linked*	Chr4: 34-37-46	195.36-209.88-268.45	linked*	linked*	linked*	

C28	2.512	Chr15: 11-15-20	45.45-66.39-71.96	7.1%	0.019	0.405	57.2%
	4.761	Chr10: 34-35-38	97.87-102.49-108.13	14.4%	0.000	0.023	
	6.134	Chr4: 3-7-8	8.909-25.69-33.02	21.9%	0.000	0.023	
	3.910	Chr5: 59-62-64	283.00-300.42-325.22	13.8%	0.002	0.092	
4,12DiMe-C28	4.310	Chr20: 1-4-9	4.967-12.71-28.54	24.9%	0.000	0.023	N/A
	linked*	Chr20: 20-24-30	57.93-73.72-84.35	linked*	linked*	linked*	
2Me-C28	3.001	Chr6: 5-12-16	43.20-53.65-63.35	17.0%	0.010	0.301	N/A
4Me+8,12DiMe-C28	2.119	Chr10: 40-42-49	112.81-127.09-154.33	11.3%	0.045	0.720	21.5%
	1.852	Chr22: 46-51-55	99.76-120.75-130.05	10.8%	0.050	0.785	
12Me-C28	1.86	Chr11: 14-19-23	64.44-83.04-88.70	9.7%	0.032	0.543	20.9%
	2.122	Chr24: 22-27-32	36.70-50.61-58.70	11.2%	0.013	0.318	
C29	3.279	Chr12: 39-43-48	121.89-135.12-153.78	19.0%	0.006	0.208	N/A
11,15DiMe-C29	1.864	Chr16: 15-21-26	70.20-83.69-90.70	10.7%	0.055	0.847	N/A
3Me-C29	2.013	Chr12: 22-29-35	73.42-103.8-116.09	12.0%	0.067	0.978	N/A
11+13Me-C29	4.667	Chr8: 27-31-37	116.69-124.23-143.22	25.2%	0.025	0.484	N/A
	linked*	Chr8: 38-39-42	152.77-158.45-170.69	linked*	linked*	linked*	
7,11,15TriMe-C29	2.766	Chr8: 38-43-45	143.22-173.42-187.41	15.3%	0.013	0.318	28.7%
	2.602	Chr5: 26-29-36	128.50-149.1-163.91	13.4%	0.026	0.484	
C30	2.095	Chr8: 26-31--38	106.23-124.23-143.22	12.3%	0.025	0.484	N/A
2Me-C30	3.823	Chr21: 6-13-17	10.95-26.48-34.17	14.1%	0.001	0.064	46.1%
	4.083	Chr13: 5-9-12	17.33-28.88-31.73	15.6%	0.000	0.023	
	3.838	Chr4: 1-2-4	0-6.219-17.23	16.4%	0.002	0.092	
4Me-C30	2.116	Chr12: 10-14-19	36.87-55.90-64.78	12.5%	0.043	0.716	N/A
14Me-C30	2.027	Chr6: 25-27-30	88.74-113.07-127.81	12.0%	0.078	1.000	25.5%
	2.398	Chr18: 27-36-39	87.11-113.54-122.12	13.9%	0.026	0.484	
C31	1.813	Chr3: 12-18-21	92.83-142.88-166.87	13.0%	0.150	1.000	26.1%
	2.386	Chr21: 46-52-52	112.22-135.98-139.79	13.6%	0.005	0.189	
11,15DiMe-C31	2.022	Chr23: 7-15-23	14.06-36.32-556.53	13.1%	0.032	0.543	N/A
13+15Me-C31	4.837	Chr18: 32-35-38	105.86-111.63-120.25	25.4%	0.000	0.023	N/A
Total	2.141	Chr14: 26-30-39	77.33-97.10-115.14	8.0%	0.045	0.720	44.0%
	3.886	Chr15: 26-29-31	89.34-98.91-102.52	15.1%	0.001	0.064	
	3.065	Chr3: 29-32-36	187.59-194.33-205.79	11.8%	0.011	0.305	
	2.413	Chr7: 13-17-21	65.37-76.84-83.48	9.1%	0.031	0.543	

CHC = cuticular hydrocarbon; LOD = log-of-odds score; Location = location of confidence interval start and end, and of the maximal LOD score; *q*-value = FDR corrected *p*-value; PEV = percent explained variation; MIM

926 = multi-interval model summarizing all QTLs for a given CHC over different chromosomes. Shades of red
927 highlight statistically significant results.
928 * Two QTLs for the same CHC on the same chromosome evaluated using the linked-QTL model of *MultiQTL*.
929 Therefore, the LOD, P.E.V., *p*-value, and *q*-value are for both QTLs together.
930