

RESEARCH ARTICLE

Genetic variation for rectal gland volatiles among recently collected isofemale lines and a domesticated strain of Queensland fruit fly, *Bactrocera tryoni* (Diptera: Tephritidae)

Cynthia Castro-Vargas^{1,2,3}, Gunjan Pandey^{1,2*}, Heng Lin Yeap^{1,4}, Shirleen S. Prasad^{1,2,3}, Michael J. Lacey⁵, Siu Fai Lee^{1,2,3}, Soo J. Park^{2,3}, Phillip W. Taylor^{2,3}, John G. Oakeshott^{1,2}

1 Environment, Black Mountain, Commonwealth Scientific and Industrial Research Organisation, Acton, ACT, Australia, **2** Applied BioSciences, Macquarie University, North Ryde, NSW, Australia, **3** Australian Research Council Centre for Fruit Fly Biosecurity Innovation, Macquarie University, North Ryde, NSW, Australia, **4** Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, VIC, Australia, **5** National Collections and Marine Infrastructure, Black Mountain, Commonwealth Scientific and Industrial Research Organisation, Acton, ACT, Australia

* Gunjan.Pandey@csiro.au



OPEN ACCESS

Citation: Castro-Vargas C, Pandey G, Yeap HL, Prasad SS, Lacey MJ, Lee SF, et al. (2023) Genetic variation for rectal gland volatiles among recently collected isofemale lines and a domesticated strain of Queensland fruit fly, *Bactrocera tryoni* (Diptera: Tephritidae). PLoS ONE 18(4): e0285099. <https://doi.org/10.1371/journal.pone.0285099>

Editor: Volker Nehring, Albert-Ludwigs-Universitat Freiburg, GERMANY

Received: October 7, 2022

Accepted: April 14, 2023

Published: April 28, 2023

Copyright: © 2023 Castro-Vargas et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its [Supporting Information](#) files.

Funding: This research was supported by the Australian Research Council Industrial Transformation Training Centre (ITTC) for Fruit Fly Biosecurity Innovation (Project IC50100026), funded by the Australian Government, including a fellowship to S.J.P. and PhD scholarship to C.C-V. This research was also conducted as part of the

Abstract

Divergence between populations in mating behaviour can function as a potent premating isolating mechanism and promote speciation. However, very few cases of inherited intraspecific variation in sexual signalling have been reported in tephritid fruit flies, despite them being a highly speciose family. We tested for such variation in one tephritid, the Queensland fruit fly, *Bactrocera tryoni* (Qfly). Qfly mating behaviour depends on volatiles secreted from male rectal glands but no role for the volatiles from female rectal glands has yet been reported. We previously detected over 100 volatile compounds in male rectal glands and identified over 30 of them. Similar numbers were recorded in females. However, many compounds showed presence/absence differences between the sexes and many others showed quantitative differences between them. Here we report inherited variation among 24 Qfly lines (23 isofemale lines established from recent field collections and one domesticated line) in the abundance of three esters, two alcohols, two amides, an aldehyde and 18 unidentified volatiles in male rectal glands. We did not find any compounds in female rectal glands that varied significantly among the lines, although this may at least partly reflect lower female sample numbers. Most of the 26 male compounds that differed between lines were more abundant in the domesticated line than any of the recently established isofemale lines, which concurs with other evidence for changes in mating behaviour during domestication of this species. There were also large differences in several of the 26 compounds among the isofemale lines, and some of these differences were associated with the regions from which the lines were collected. While some of the variation in different compounds was correlated across lines, much of it was not, implicating involvement of multiple genes. Our findings parallel reports of geographic variation in other Qfly traits and point to inherited differences in reproductive physiology that could provide a basis for evolution of premating isolation between ecotypes.

SITplus collaborative fruit fly program. Project Raising Q-fly Sterile Insect Technique to World Standard (HG14033) is funded by the Hort Frontiers Fruit Fly Fund, part of the Hort Frontiers strategic partnership initiative developed by Hort Innovation, with coinvestment from Macquarie University and contributions from the Australian Government. We also acknowledge CSIRO for its generous support of this project. Australian Research Council, <https://www.arc.gov.au/> Hort Innovation, <https://www.horticulture.com.au/> Macquarie University, <https://www.mq.edu.au/> CSIRO, <https://www.csiro.au/> The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Pheromones involved in mating behaviour have been found to function as effective premating isolating mechanisms among sympatric populations of many closely related insect species; in some cases quite subtle differences in pheromone composition can be sufficient to deter inter-specific mating [1–4]. However, much less is known about intraspecific variation in pheromone composition, and less still about its genetic basis. Two major exceptions involve moths; in one case racial differences in pheromone content in *Ostrinia nubilalis* were largely attributed to allelic differences at a single locus affecting a fatty acyl reductase activity [5], whereas in the other case quantitative intra-population variation in an aldehyde component of a *Heliothis virescens* pheromone was found to be polygenically inherited, with the major locus only explaining 11–21% of the variation [6]. Although the genetic architectures were very different, both cases exemplify a situation where genetic polymorphism in pheromone composition within a species might potentially provide a basis for the development of a premating isolating mechanism. Variations in sex pheromone composition between geographic populations of other moths [7–10] and a sandfly [11, 12] have also been reported but the genetic bases of these differences are not yet known.

Tephritid fruit flies are a promising model to further investigate intraspecific polymorphism in pheromone composition and its potential to drive non-random mating. Several of the key genera are highly speciose [13, 14], and some now-sympatric sibling species pairs have been shown to be effectively isolated by differences in mating behaviours [15, 16]. Moreover many are important horticultural pests [17, 18], and there has been considerable research into biologically based control methods such as the Sterile Insect Technique (SIT) and Male Annihilation Technique (MAT) which exploit and/or intervene in aspects of mating or more general sex-specific behaviours [19–25]. There is also increasing interest in the biochemistries underpinning those behaviours, which are often found to involve semiochemicals [26–34]. Male rectal glands have been shown to produce many of these compounds in various tephritids, including the Mexican fruit fly, *Anastrepha ludens* [35, 36], Mediterranean fruit fly, *Ceratitis capitata* [37–40], Oriental fruit fly, *Bactrocera dorsalis* [41, 42] and Queensland fruit fly (Qfly), *Bactrocera tryoni* [31, 43–46]. Comparative analyses have shown substantial qualitative and quantitative divergence in rectal gland volatiles among *Bactrocera* species [47, 48].

Qflies are a major horticultural pest in northern and eastern Australia and parts of Melanesia [14, 17, 49–51], and is one of the most studied tephritids with respect to pheromone gland chemistry. Sexually mature Qfly males release a blend of volatiles from their rectal glands and disperse them via rapid wing vibrations (‘calling’) to attract females [31, 52–56]. Over 100 volatile compounds have been found in the rectal glands of Qfly males, and over 30 of these have been identified [47, 57]. The most abundant compounds are aliphatic amides [43, 58, 59]. However it is not yet known which of the compounds are responsible for the short range attraction of females observed in experiments with crushed male rectal glands.

Qfly females also secrete diverse volatiles from their rectal glands, including lesser amounts of some of the amides found in males, with their profiles instead dominated by fatty acid esters [47, 57, 60, 61]. Overall a majority of the volatiles secreted from Qfly rectal glands show at least quantitative and in many cases presence/absence differences between the sexes [47]. Less is known about the functions of the volatile compounds emitted by female Qflies [62, 63], although work on other tephritids has implicated their female rectal gland volatiles in some mating and other reproductive behaviours [41, 64–66].

Qualitative and quantitative geographic variation in male pheromone composition has been described in two tephritid species, specifically between three different *C. capitata* populations in Greece and Argentina [67, 68] and seven different *Anastrepha fraterculus* populations

in Brazil and Argentina [69]. However, the genetic basis of this variation, and any consequences for its behaviour, are not yet known. Nor is anything yet known about geographic variation in Qfly pheromone composition, although the species does show heritable geographic variation in climate stress tolerance [70] and molecular population genetic studies show clear genetic divergence between geographic populations, including geographically proscribed introgression from the closely related taxon *Bactrocera aquilonis* [50, 71].

Non-genetic factors could also generate geographic variation in rectal gland volatiles in fruit flies such as Qfly. For example, geographic variation in host plant utilisation could play a role, because host plant chemicals can be translocated to rectal glands in some tephritids, and, in certain of those cases, transformed to other bioactive volatiles [28, 31, 72–74]. Many of the fatty acid esters found in rectal glands of Qfly females are also found in host fruits [75–78] and some are attractive to Qflies of both sexes [79]. One well studied example of such a process involves cuelure (4-(4-acetoxyphenyl)-2-butanone), an analogue of the more common naturally occurring plant phenylpropanoid raspberry ketone [80]. Cuelure is a strong attractant for Qfly males (as is raspberry ketone) and is a widely used lure for monitoring this species [19, 22, 81, 82]. When ingested, cuelure is transported through the haemolymph to the rectal gland, where it is stored in its hydrolysed form as raspberry ketone and released together with the endogenously produced volatiles [73].

Testing for a genetic basis to any geographical variation in rectal gland volatiles therefore requires a ‘common garden’ approach which controls for dietary and environmental effects.

Here we report our application of gas chromatography-flame ionization detection (GC-FID) techniques to analyse rectal gland extracts of male and female Qflies from a long domesticated laboratory colony and 23 isofemale lines that were recently established using flies sourced from across this species’ geographic range. The analyses were conducted in a common garden laboratory environment with carefully defined diet to control for dietary effects. We found differences between lines across generations in the abundance of 26 compounds in males but none in females. Most of the 26 compounds were more abundant in the long domesticated stock than the isofemale lines, which is consistent with the genetic responses to selection on various other traits which have been reported to occur during domestication of this and other insect species [83–85]. However, several of the compounds also showed differences among the isofemale lines, and these included differences both within and between source populations. Some of the variation was correlated across lines, suggesting coordinated genetic control, but much of it was not, implicating variation in several different genes. Four of the 26 varying compounds were identified against authentic standards and another four were tentatively identified from the mass spectra of peaks with equivalent Kovats Indices (KIs) in parallel solid phase microextraction-coupled gas chromatography-mass spectrometry (SPME GC-MS).

Results

GC-FID detection of significantly varying peaks

Twenty-three isofemale lines were established from seven collections of Qflies from widely separated locations in northern and eastern Australia in 2017 and 2018 (Table 1). Sexually mature flies from single sex (i.e., virgin) and mixed sex cohorts were assayed for the nine lines collected latest in this period, referred to hereafter as the major set, whereas only males from mixed sex cohorts were assayed for the 14 earlier collected lines, hereafter referred to as the minor set. The two sets of lines were assayed in different batches but the long-domesticated S06 line, which was originally collected from Sydney in 2006 [85], was also included in both sets (albeit using different S06 generations in the two sets). The flies in the mixed sex cohorts

Table 1. *B. tryoni* lines used in the GC-FID analyses. Column heading “Screened generations” refers to the generations in the laboratory post-collection when the GC-FID analyses were carried out. Precise numbers for these generations could not be provided for S06, because records for its first few years in the laboratory were not available.

Strain type	Source population	Lines	Location		Screened generations	Collection date
			Latitude	Longitude		
Isofemale lines (major set)	Alice Springs	AS09, AS19, AS36	-23.69	133.89	6, 7, 8, 10	Nov 2017
	Cape Tribulation	CT07, CT38, CT60	-16.09	145.46	4, 5, 6, 8	Aug 2018
	Sydney	SY13, SY18, SY53	-33.90	151.14	6, 7, 8, 10	Sep 2017
Isofemale lines (minor set)	Brisbane	BR05, BR34, BR39, BR50	-27.41	152.90	6, 8, 10, 12	Mar 2017
	Mareeba	MB09, MB46, MB50, MB64	-17.01	145.43	6, 8, 10, 12	Feb 2017
	Narrabri	NB02, NB11, NB28	-30.33	149.78	6, 8, 10, 12	Mar 2017
	Utchee Creek	UT01, UT03, UT43	-17.60	145.99	6, 8, 10, 12	Mar 2017
Mass-bred laboratory strain	Sydney	S06	-33.90	151.14	~ 116–122	2006

<https://doi.org/10.1371/journal.pone.0285099.t001>

(hereafter denoted mixed males and mixed females, respectively) were obtained from cages containing both sexes, 90–95% of females from which had been mated [47].

Our GC-FID methods were the same as those of Castro-Vargas et al. [47], whose major focus was on the S06 strain. Some of their data are also used herein (Table 2). They reliably distinguished over 180 rectal gland peaks across the two sexes and identified 45 of them in parallel solid phase microextraction gas chromatography-mass spectrometry (SPME GC-MS) analyses. The relatively large number of identified peaks enabled them to identify three retention time (Rt) intervals, 4.60–10.39, 10.40–13.59 and 13.60–21 min (hereafter short, mid and long Rt peaks), which contained disproportionately high numbers of the identified short chain alcohols, esters and ketones, aliphatic amides, and fatty acid esters, respectively. These intervals corresponded to Kovats Indices (KIs) of 840–1095, 1096–1240 and 1241–2175 respectively. Castro-Vargas et al. [47] further classified their peaks as major, intermediate and minor in respect of abundance if they contributed $\geq 1\%$, $0.1\text{--}0.99\%$ or $< 0.1\%$ of total peak area in the sex/mating history category where they were most abundant, and also rated them as sex specific (only present in one sex), sex selective ($>\log_{10}$ difference in average peak area between the sexes) or no strong sex bias ($<\log_{10}$ difference in average peak area between the sexes) within each mating history category.

We recovered 184 of the peaks of Castro-Vargas et al. [47] in our data sets and their Rt ranges and abundance and sex bias classifications are summarised in Table 2. Seventeen additional peaks not reported by Castro-Vargas et al. [47] were also detected in our minor set of lines; these were all rated as minor in abundance in the mixed males we analysed from these lines but in the absence of data for females from these lines they could not be classified for sex bias. The four abundance data sets (i.e., for virgin and mixed flies of each sex) for the major set of lines are in S1 Dataset of Castro-Vargas et al. [47], and the one abundance data set (i.e. for mixed males only) for the minor set of lines is in S1 Dataset of this article. S1 Fig shows illustrative GC-FID chromatograms.

Linear modelling (see Materials and Methods) of abundance data, as represented by peak areas, from each of the five data sets found no significant variation among the constituent lines in any of the peaks in either female data set (i.e., the major virgin and mixed female sets; 85 and 75 peaks tested, respectively) but found a total of 30 cases of significant variation among lines across the three male data sets (17 cases in major set virgin males, two in major set mixed males and 11 in minor set mixed males; 127, 153 and 120 peaks tested, respectively). S1 Table shows the linear modelling results for all peaks tested and Figs 1 and 2 summarise the abundance data for each line for the 30 cases of significant variation, with the corresponding

Table 2. Peaks detected in the CG-FID analysis of virgin and mixed flies from the major set of lines, categorised as per Table 2 of Castro-Vargas et al. [47] by sex specificity and selectivity. Peaks shown as bolded, underlined or plain text are classified as major, intermediate or minor in abundance, respectively, in the category in which they were most abundant. “~” refers to peaks with no sex specificity or selectivity. “nd” refers to peaks not detected in the corresponding category. Seventeen other peaks only detected in minor set mixed males were not classified for sex specificity because minor set females were not assayed (see S1 Dataset).

Type	Single sex	Mixed sex	Peaks		
			Short Rt	Mid Rt	Long Rt
Female biased (34)	Female specific (21)	Female specific (12)	8.82	11.27	15.51, 18.40, 18.46 (Ethyl tetradecanoate) , <u>18.73</u> , <u>18.78</u> , <u>18.83</u> , 19.59, <u>20.05</u> , <u>20.11</u> , 20.25
		Female selective (3)			<u>19.14</u> , 19.51 (Ethyl (E)-9-octadecenoate) , <u>20.55</u>
		~ (2)			<u>18.61</u> , 19.76
		n.d. (4)	<u>9.63</u>		<u>18.85</u> , 19.70, 19.99
		Female selective (10)		10.85	17.55, 17.69, 18.88, 18.98 (Ethyl (Z)-9-hexadecenoate) , <u>19.01</u> , <u>19.39</u>
	Female selective (10)	Female selective (7)			18.42 (Ethyl (Z)-9-tetradecenoate) , 18.93
		~ (2)			<u>20.52</u>
		n.d. (1)			<u>20.52</u>
	~ (3)	Female specific (2)			<u>18.37</u> , <u>20.31</u>
		Female selective (1)			<u>18.21 (Methyl tetradecanoate)</u>
Male biased (110)	Male specific (63)	Male specific (49)	<u>4.61 (n-propyl 2-methylpropanoate)</u> , <u>5.20</u> , <u>5.41</u> , <u>5.99</u> , <u>6.30</u> , <u>6.64</u> , <u>6.94</u> , <u>8.53</u> , <u>9.43</u> , <u>9.58</u>	<u>11.46</u> , 11.93, 13.59 (N-(3-Methylbutyl)-2-methylpropanamide)	13.79, 13.95, 15.22, 15.38, 15.67, 15.98, 15.99, 16.18, 16.52, 16.59, 16.66, 16.85, 16.89, <u>16.94</u> , <u>17.01</u> , <u>17.07</u> , <u>17.17</u> , <u>17.26</u> , <u>17.41</u> , <u>17.64</u> , <u>17.75</u> , <u>17.84</u> , <u>18.01</u> , <u>18.48</u> , <u>18.68</u> , <u>18.80</u> , 18.89 , 19.08, <u>19.12</u> , <u>19.62</u> , <u>19.72</u> , <u>19.83</u> , <u>20.33</u> , <u>20.59</u> , <u>20.64</u> , <u>20.98</u>
		n.d. (10)	10.05	13.33	<u>16.73</u> , <u>17.72</u> , <u>18.76</u> , <u>19.04</u> , <u>19.65</u> , <u>20.19</u> , <u>20.35</u> , <u>20.81</u>
		~ (4)			<u>15.59</u> , <u>18.72</u> , <u>20.01</u> , <u>20.13</u>
	Male selective (4)	Male specific (2)		<u>10.41 (N-(2-Methylbutyl)acetamide)</u> , <u>11.04</u>	
		Male selective (2)		10.59 (N-(3-Methylbutyl)acetamide) , 12.96 (N-(3-Methylbutyl)propanamide)	
	~ (14)	Male specific (12)	4.82 (4-Heptanone), 4.96, 6.21	12.27	15.87, 17.15, <u>17.32</u> , <u>17.52</u> , <u>18.12</u> , 18.53, 19.48 , 20.07
		Male selective (2)		12.63 (N-(2-Methylbutyl)propanamide)	19.43
	n.d. (29)	Male specific (29)	4.07, 7.79, 8.73, 9.79, <u>9.98</u>	11.7, 11.79 (Diethyl succinate), <u>11.97</u> , <u>12.45</u> , <u>13.34</u>	13.88, 14.54, 15.07, 15.45, 16.14, 16.70, 16.78, <u>18.02</u> , <u>19.15</u> , <u>19.20</u> , <u>19.31</u> , <u>19.78</u> , <u>19.92</u> , <u>19.94</u> , <u>19.97</u> , <u>20.22</u> , <u>20.23</u> , <u>20.40</u> , <u>20.42</u>
	Changed specificity (4)	Male specific (2)			<u>18.39</u> , <u>20.62</u>
		Female specific (2)			15.63, 20.15
No sex bias (36)	~ (31)	~ (31)	<u>4.23</u> , <u>6.06</u> , <u>7.66 (2-Ethyl-1-hexanol)</u> , <u>8.38</u>	13.47 (N-(2-Methylbutyl)-2-methylpropanamide)	<u>14.37</u> , <u>14.86</u> , <u>15.77</u> , 16.11 , <u>16.32</u> , <u>16.45</u> , <u>17.21</u> , <u>17.37 (Methyl dodecanoate)</u> , <u>17.45</u> , <u>17.63 (Ethyl dodecanoate)</u> , <u>17.92</u> , <u>17.96</u> , <u>18.10</u> , <u>18.17</u> , <u>18.27</u> , <u>18.32</u> , <u>18.64</u> , <u>19.25</u> , <u>19.29</u> , <u>19.35</u> , <u>19.56</u> , <u>19.66</u> , <u>19.80</u> , <u>19.86</u> , <u>20.50</u> , <u>20.84</u>
	n.d. (5)	~ (5)			<u>16.50</u> , <u>20.37</u> , <u>20.68</u> , <u>20.88</u> , <u>20.90</u>

<https://doi.org/10.1371/journal.pone.0285099.t002>

A) Major set virgin males												
Peak	Specificity/ selectivity	F	AS09	AS19	AS36	CT07	CT38	CT60	SY13	SY18	SY53	S06
4.61 ¹	Msp/Msp	8.87***	1207 ab	1602 bc†	0 a†	1483 abc	202 a†	414 a†	823 ab	838 ab	1322 ab	3616 c
5.20	Msp/Msp	8.38***	1190 b†	891 ab	389 abc	0 ab	0 ab	515 ab	1254 b†	96 a†	926 ab	2124 c
5.41	Msp/Msp	5.10**	5875 bc†	4797 ab	4642 abc	8874 bc†	1280 a†	2569 ab	5852 bc†	4295 ab	4335 ab	13884 c
5.99	Msp/Msp	3.30*	548 ab	364 a	476 ab	548 ab	227 a	542 ab	545 ab	493 ab	425 a	1456 b
6.30	Msp/Msp	4.97**	1446 bc†	1262 ab	1073 abc	1789 bc†	236 a†	668 ab	1462 bc†	105 ab	1028 ab	2993 c
6.64	Msp/Msp	5.19**	1010 ab	810 ab	315 ab	1043 ab	0 a	172 a	139 a	565 a	790 ab	1961 b
10.05	Msp/nd	4.66 **	653 ab	223 a	0 ab	180 ab	0 a	193 a	1086 ab	141 a	499 ab	1618 b
10.85	Fse/Fse	3.40*	112 a†	59 a†	284 ab	109 ab	67 ab	422 ab	421 ab	1109 b†	683 ab	355 ab
13.47 ²	~	5.50**	4399 ab	3391 a	3113 ab	5138 ab	1471 a	2727 a	2175 a	5723 ab	2045 a	12427 b
16.18	Msp/Msp	9.56***	2758 bc†	3152 c†	691 abc	919 abc	82 a†	597 ab†	587 ab†	211 a†	286 a†	4477 c
17.07	Msp/Msp	3.18*	418 ab	296 ab	196 ab	634 ab	81 a	272 ab	417 ab	334 ab	254 ab	793 b
17.37 ³	~	5.73**	1488 ab	781 a	399 ab	577 ab	312 a	397 a	354 a	1258 a	1916 ab	6988 b
17.69	Fse/Fse	3.53*	756 ab	354 a	245 ab	824 ab	243 a	259 a	301 a	833 ab	266	3477 b
20.35	Msp/nd	3.22*	1813 b†	57 a†	1342 ab	704 ab	59 ab	34 a†	991 ab	43 a†	42 a†	50 a
20.50	~	3.54*	781 ab	61 a	0 ab	1369 ab	48 ab	553 ab	185 ab	210 ab	128 ab	1255 b
20.59	Msp/Msp	3.80*	2552 b†	730 ab	916 ab	2904 b†	221 ab	1166 ab	1630 ab	1660 ab	185 a†	2132 b
20.98	Msp/Msp	4.83**	566 ab	134 a	0 a	423 ab	61 a	304 a	312 a	361 a	525 ab	866 b
B) Major set mixed males												
18.68	Msp/Msp	26.71*	19464 c†	15350 bc†	-	4529 a†	-	-	6800 a†	6603 a†	9517 ab	11685 b
20.15	Fsp/Msp	24.05*	7964 cd†	4425 abc	-	6256 bcd	-	-	1678 ab†	12062 d†	5838 abcd	0 a

Fig 1. Back-transformed emmeans (estimated marginal means) for peak areas for peaks showing significant differences among isofemale lines in the major set of virgin (A) and mixed (B) males. As per Table 2, peaks are bolded, underlined or neither if they are classified as major, intermediate or minor in abundance. Known peak identities are indicated with superscripts as follows: ¹*n*-propyl 2-methylpropanoate; ²*N*-(2-methylbutyl) 2-methylpropanamide; ³ methyl dodecanoate. Sex bias (also taken from Table 2) is indicated as follows: Msp = male specific, Mse = male selective, Fse = female specific, Fse = female selective, nd = not detected and ~ = not sex biased, with the bias in virgin and mixed sex cohorts indicated before and after the slash, respectively. Bold lines between rows separate peaks in the different Rt ranges. Confidence limits are given in S2 Table (see Materials and Methods). FDA-corrected significance values for F statistics are indicated with asterisks, with * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$. Letter codes are used to show the results of *post hoc* pairwise contrasts between lines, where lines with the same letters are not significantly different from one another. Lines with a single letter only sit in a single group of lines and non-overlapping lines/groups of lines are coloured, with blue, green, purple and pink representing groups with progressively higher values. A red † indicates a value in an isofemale line which differs significantly from the value(s) in at least one other isofemale line(s). Some apparent inconsistencies between the groupings reflect variation in the between-replicate variances for the particular means, as evidenced in S2 Table. A dash indicates the peak was not analysed in that line. Replicate numbers are indicated in S1 Dataset of Castro-Vargas et al. [47].

<https://doi.org/10.1371/journal.pone.0285099.g001>

A) Minor set mixed males																	
Peak	Sex bias	F	BR05	BR34	BR39	BR50	MB09	MB46	MB50	MB64	NB02	NB11	NB28	UT01	UT03	UT43	S06
4.17	-	2.72*	0 ab	26 ab	141 ab	37 ab	436 b	439 ab	274 ab	399 ab	374 ab	371 ab	425 ab	95 ab	115 ab	344 ab	0 a
4.61 ¹	Msp/Msp	3.01*	615 ab	477 a	312 ab	316 a	311 a	306 a	68 a	336 a	203 a	286 ab	316 a	285 a	312 a	628 ab	1827 b
5.41	Msp/Msp	3.18*	1387 ab	1223 a	1720 ab	696 a	1399 ab	1232 a	641 a	1183 a	1070 a	1412 ab	1014 a	1203 ab	1138 a	1733 ab	4195 b
6.06	~/-	3.03*	4547 ab	5504 ab	7547 ab	5765 ab	3151 a	3137 a	996 ab	3166 a	4075 ab	3547 ab	3225 a	5486 ab	4090 ab	4425 ab	9194 b
6.21	~Msp	2.73*	2760 ab	3059 ab	5204 ab	3104 ab	1682 a	1685 a	659 ab	1694 a	2297 ab	1928 ab	1761 a	2953 ab	3063 ab	2531 ab	5260 b
6.30	Msp/Msp	3.05*	453 b†	189 ab	60 ab	50 a†	334 b†	265 ab	182 ab	259 ab	183 ab	364 ab	239 ab	122 ab	234 ab	612 ab	616 b
8.06	-	2.94*	324 ab	188 ab	256 ab	281 ab	145 a	145 a	98 ab	167 a	177 a	146 ab	145 a	247 ab	260 ab	217 ab	408 b
10.41 ²	Mse/Msp	2.54*	4970 ab	3026 ab	1820 ab	2034 ab	2862 ab	2706 ab	920 ab	2015 ab	1402 a	3255 ab	2227 ab	1404 ab	2324 ab	4776 ab	5295 b
10.85	Fse/Fse	3.18*	196 ab	719 b†	810 ab	495 b†	267 ab	226 ab	175 ab	287 ab	170 a†	282 ab	582 b†	264 ab	251 ab	0 ab	331 ab
19.80	~/-	2.61*	120 ab	4040 ab	11082 ab	6201 b†	3606 ab	378 ab	2181 ab	98 a†	4215 ab	1815 ab	6232 ab	2356 ab	5480 ab	0 ab	58 ab
19.83	Msp/Msp	2.64*	2055 ab	2261 ab	4250 ab	4833 b	4117 b	1903 ab	1250 ab	1759 ab	1915 ab	2114 ab	3067 ab	2427 ab	4720 b	1251 ab	787 a

Fig 2. Back-transformed emmeans for peak areas for peaks showing significant differences among minor set lines for mixed males. Formatting is as per Fig 1 and confidence limits are given in S2 Table. Known peak identities are indicated with superscripts as follows: ¹*n*-propyl 2-methylpropanoate; ²*N*-(2-methylbutyl) acetamide. Note that sex bias data are not shown for peaks 4.17 and 8.06 because they did not meet the inclusion criterion (detected in > 50% of samples in at least one sex/mating history category) for analysis in the major set of lines and therefore lacked data for females. As per Fig 1, some apparent inconsistencies between the groupings reflect variation in the between-replicate variances for the particular means. Replicate numbers are indicated in S1 Data set attached.

<https://doi.org/10.1371/journal.pone.0285099.g002>

confidence limits given in [S2 Table](#). Twenty six peaks were involved in the 30 cases of significant differences, with four peaks (4.61, 5.41, 6.30 and 10.85) showing significant differences in both the major set virgin and minor set mixed males. Greater statistical power due to higher sampling intensity likely contributed to the higher numbers of significant differences seen in the major set virgin males and minor set mixed males but it is not clear whether the absence of any significant differences in the two female data sets was solely due to their lower sample sizes, since two differences were found in the similarly sized sample of major set mixed males (see [Materials and Methods](#)). The 26 significantly varying peaks represented 14% of all peaks tested across the three male data sets.

GC-FID characterisation of significantly varying peaks

Four of the 26 peaks involved in significant lines differences had been identified in Castro-Vargas et al. [47]: 4.61 (*n*-propyl 2-methylpropanoate), 10.41 (*N*-(2-methylbutyl)acetamide), 13.47 (*N*-(2-methylbutyl)-2-methylpropanamide) and 17.37 (methyl dodecanoate). Sex specificity was likewise already known for 24 of the 26 (all but two that were significant in minor set mixed males but not detected in the major set) and over half of these (14) were male-specific. Three of the 26 were ranked major in abundance, 13 as intermediate and ten as minor, which was similar to the ratio of the three abundance categories among all peaks tested across the five data sets (21: 75: 105, respectively; $\chi^2 = 2.36$, $df = 2$, $p > 0.05$).

However, the ratio of significant peaks in the short, mid and long Rt ranges, 11: 3: 12, respectively, differed from that of 31: 23: 147, respectively, seen among all the peaks tested across the five data sets ($\chi^2 = 14.77$, $df = 2$, $p < 0.001$). In particular there was a marked excess of significant peaks in the short Rt range, and a corresponding deficit of those in the long Rt range. Thus, 35% of the short Rt peaks tested differed significantly between lines and, given that three of the four peaks that varied significantly in two data sets had short Rt's, they accounted for 47% of the 30 cases of significant differences among lines.

Post hoc pairwise comparisons using emmeans (Figs 1 and 2, [S2 Table](#) and see [Material and Methods](#)) showed that the long domesticated line, S06, had the highest abundance value for 20 of the 30 cases of significant line differences. In particular, S06 had the highest value for 13 of the 14 cases involving short Rt peaks. In some cases, the values for S06 were about two-fold, and in a few cases three- or four-fold, higher than the highest value among the isofemale lines. Interestingly, in four cases, including the one exception among the short Rt cases where S06 did not yield the highest value, it actually recorded amongst the lowest values. In two of these, including the exceptional short Rt case, the peak was undetectable in S06.

In addition to the differences involving S06, there were also significant differences among the isofemale lines in 13 of the 30 cases, five involving the short Rt range, two the mid Rt range and six the long Rt range (Figs 1 and 2 and [S2 Table](#)). A total of eleven peaks contributed to these cases, with two peaks (6.30 and 10.85) differing in both the major set virgin and minor set mixed male data sets. Six of the 13 cases involved four- to ten-fold differences between the highest and lowest lines for these peaks, the rest all being greater than 20 fold. Two peaks otherwise classified as intermediate in abundance were undetectable in certain line(s). While there was some trend for particular lines to have either relatively high (e.g., AS09 and CT07) or relatively low (e.g., CT38) abundances across multiple peaks, most of the lines showed a mix of relatively high and low abundance values. The correlations across peaks are assessed in more detail below.

All nine of the isofemale lines in the major set showed significant differences between them in pairwise contrasts in at least one peak in the virgin males and all but one of those tested also did among the mixed males (red †s in [Fig 1](#) and see also [S1 Table](#)). One of these major set

Table 3. Summary of PERMANOVAs assessing overall cross-peak differences in each of the three Rt ranges in rectal gland volatiles in the five data sets. Analysis I assessed the differences across all 24 lines. If that term was significant, Analysis II, which excluded S06, then assessed the differences due to locality of origin and lines within localities. F ratios are shown with degrees of freedom in brackets. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

	Major set				Minor set
	Virgin males	Mixed males	Virgin females	Mixed females	Mixed males
Analysis I: Differences across all 24 lines for each Rt range					
Short Rts; Lines	2.40 (9,37)***	2.60 (6,6)	1.21 (9,17)	4.5 (6,4)	1.60 (14,88)**
Mid Rts; Lines	2.07 (9,37)*	4.60 (6,6)*	0.90 (9,17)	2.20 (6,4)	1.60 (14,88)*
Long Rts; Lines	1.30 (9,37)	5.00 (6,6)*	0.95 (9,17)	0.60 (6,4)	1.30 (14,88)*
Analysis II: Differences across the 23 isofemale lines for each Rt range					
Short Rts; Localities	1.80 (2,32)	-	-	-	2.93 (3,83)***
Short Rts; Lines within Localities	1.44 (6,32)	-	-	-	0.81 (10,83)
Mid Rts; Localities	1.02 (2,32)	3.60 (2,4)	-	-	0.44 (3,83)
Mid Rts; Lines within Localities	1.41 (6,32)	0.09 (3,4)	-	-	1.52 (10,83)
Long Rts; Localities	-	8.90 (2,4)*	-	-	1.01 (3,83)
Long Rts; Lines within Localities	-	0.17 (3,4)	-	-	1.21 (10,83)

<https://doi.org/10.1371/journal.pone.0285099.t003>

lines, AS09, differed significantly from some of the others in nine different peaks. In contrast, only half of the 14 minor set isofemale lines showed significant differences in at least one peak and the highest number of peaks involved in a significant difference in a single minor set line was two, for BR50 (red †s in Fig 2 and see also S2 Table). Only one source location, Utchee Creek, in the minor set of lines, was not involved in any significant differences between any of the isofemale lines. Furthermore there were significant differences between isofemale lines from the same source location for all locations except Utchee Creek and Mareeba, also in the minor set of lines.

In order to assess higher level differences in the composition of the rectal gland volatiles among lines, we then used permutational multivariate analyses of variance (PERMANOVAs) to investigate the extent of overall differences between the lines across all the peaks in each of the three Rt ranges. Thus these analyses used the data for all 201 peaks, not just the 26 which had individually shown significant differences. The PERMANOVAs revealed significant differences among both virgin and mixed males in the major set of lines, as well as in mixed males from the minor set, but not in either virgin or mixed females from the major set (Table 3). Differences in the major set virgin males involved both the short and mid Rt ranges; those in the major set mixed males involved the mid and long Rt ranges; and those in the minor set mixed males involved all three Rt ranges. Several of these differences were largely due to S06 being different from the other lines because they became non-significant when the S06 data were excluded from the analyses (Table 3). However, there were also significant differences between isofemale lines from different source localities for major set mixed males for peaks in the long Rt range, and for minor set mixed males for peaks in the short Rt range (Table 3).

Linear Discriminant Analysis (LDA; Fig 3) showed the difference in the major set mixed males in the long Rt range was largely due to Cape Tribulation differing from Sydney and Alice Springs, but this result was difficult to interpret because only a single Cape Tribulation line had sufficient data to be included in this analysis. On the other hand, the LDA for the short Rt peaks in the minor set mixed males clearly separated the four populations into three groups, with Mareeba plus Narrabri at one extreme, Utchee Creek at the other, and Brisbane between them.

We also assessed the extent to which different peaks showing significant line variation showed correlated patterns of variation. To do this we focussed on a single data set, major set

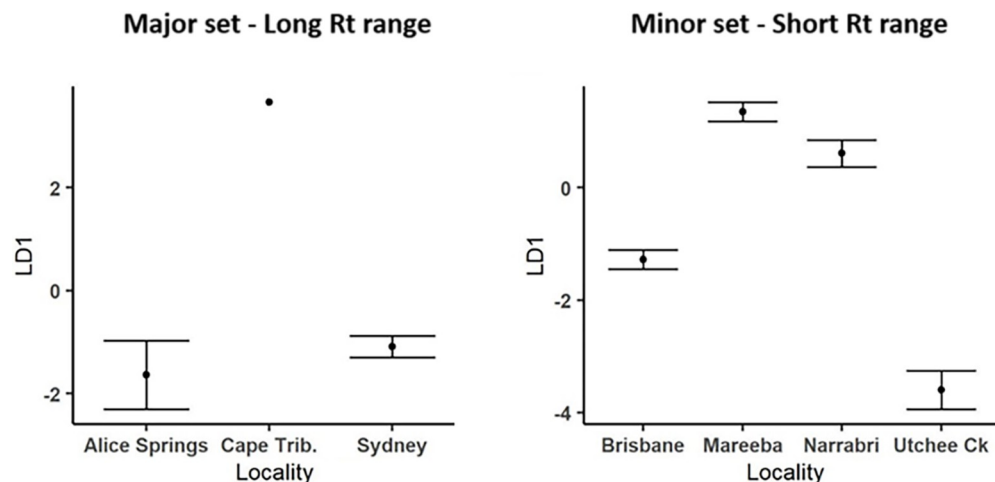


Fig 3. LDAs of significant cross-peak differences between localities in the PERMANOVAs in Table 3. Separate plots are shown for peaks from major set mixed males eluting in the long Rt range and those from minor set mixed males eluting in the short Rt range. Standard error bars are also shown. LD1 in each plot explained > 80% of the between-locality variance.

<https://doi.org/10.1371/journal.pone.0285099.g003>

virgin males, because it had the most peaks showing significant differences and we wished to avoid the possibility of producing spurious patterns that might arise if we combined data sets involving different mating histories. Overall, we found many significant correlations when S06 was included in the analysis, and some of these remained significant when S06 was excluded (Fig 4). The great majority of the correlations were positive, and many of them significantly so: 95 when S06 was included and 50 when it was not. Only four of the few negative correlations across the two analyses were significant, all four involving peak 10.05 and two also involving peak 16.18. Two peaks, 10.85 and 20.35, showed little correlation with other peaks in either analysis, suggesting independent genetic control of the variation in those peaks. The analysis excluding S06 also found several other peaks were largely uncorrelated with one another, but strong correlations remained among most of the six short Rt peaks analysed, among a few of the eight long Rt peaks, and, intriguingly, between two long Rt peaks (17.07 and 17.69) and four of the short Rt peaks. The strong correlations suggest some shared genetic basis for the differences in the peaks involved, both among the isofemale lines and between them and S06. Overall, however, low r^2 values show that most of the variation in different peaks was independent of those in other peaks.

GC-MS characterisation of significantly varying peaks

SPME GC-MS analysis of samples of virgin and mixed S06 males and females was then carried out in order to further characterise the 22 significantly varying peaks that had not already been identified against authentic standards in our GC-FID analysis. The gas-liquid extraction methods mediated by the SPME fibres yield cleaner mass spectra, albeit not always the same peaks, as compared to the liquid-liquid hexane extractions used in our original GC-MS analysis [47].

Our strategy for this work was first to establish the overall relationship between the Rt profile of the peaks detected by GC-FID and the profile of Kovats Indices (KIs) obtained for peaks detected by the GC-MS analysis, which we did by comparing the results of peaks identified against authentic standards with both technologies. This relationship could then be used to impute KI values for all the other GC-FID peaks. Further, by comparing the imputed KIs of the other GC-FID peaks with the observed KIs of the SPME GC-MS peaks, and also

	4.61	5.2	5.41	5.99	6.3	6.64	10.05	10.85	13.47	16.18	17.07	17.37	17.69	20.35	20.5	20.59	20.98
4.61		0.43**	0.66***	0.41**	0.65***	0.59***	0.34*	-0.28.	0.45**	0.52***	0.58***	0.63***	0.62***	-0.15	0.33*	0.17	0.45**
5.2	0.19		0.54***	0.57***	0.53***	0.37**	0.72***	-0.06	0.20	0.58***	0.43**	0.39**	0.26.	0.03	0.29*	0.22	0.30*
5.41	0.53***	0.36*		0.66***	0.95***	0.56***	0.50***	-0.13	0.36*	0.53***	0.79***	0.56***	0.56***	-0.04	0.47***	0.40**	0.53***
5.99	0.20	0.39*	0.56***		0.65***	0.42**	0.45**	0.21	0.52***	0.39**	0.78***	0.49***	0.55***	-0.08	0.65***	0.29*	0.38**
6.3	0.51***	0.32*	0.94***	0.54***		0.61***	0.54***	-0.05	0.42**	0.56***	0.74***	0.56***	0.58***	-0.05	0.40**	0.42**	0.44**
6.64	0.44**	0.13	0.42**	0.23	0.50***		0.31*	0.04	0.45**	0.57***	0.54***	0.61***	0.46***	0.02	0.40**	0.14	0.51***
10.05	0.05	0.59***	0.30.	0.23	0.33*	0.08		-0.01	0.28.	0.42**	0.41**	0.39**	0.26.	0.07	0.22	0.44**	0.26.
10.85	-0.33*	-0.09	-0.09	0.23	-0.01	0.06	0.01		0.23	-0.32*	0.00	0.08	-0.07	-0.23	-0.02	-0.20	-0.01
13.47	0.24	-0.12	0.17	0.39*	0.24	0.29.	0.00	0.27.		0.32*	0.45**	0.40**	0.48***	-0.14	0.44**	0.39**	0.42**
16.18	0.37*	0.47**	0.40**	0.24	0.42**	0.45**	0.25	-0.36*	0.15		0.46**	0.42**	0.38**	0.10	0.34*	0.40**	0.24.
17.07	0.45**	0.26.	0.72***	0.74***	0.66***	0.42**	0.23	0.08	0.31*	0.31*		0.55***	0.59***	0.00	0.70***	0.33*	0.58***
17.37	0.46**	0.12	0.38*	0.29.	0.36*	0.45**	0.12	0.12	0.17	0.23	0.40**		0.48***	-0.05	0.46***	0.19	0.57***
17.69	0.47**	-0.02	0.41**	0.40**	0.41**	0.26	-0.04	-0.11	0.31*	0.19	0.49**	0.27.		0.03	0.47***	0.43**	0.51***
20.35	-0.06	0.13	0.06	0.00	0.03	0.12	0.17	-0.34*	-0.09	0.20	0.09	0.04	0.14		0.15	0.29*	0.23
20.5	0.10	0.07	0.27.	0.55***	0.21	0.19	0.00	0.03	0.31*	0.16	0.59***	0.29.	0.34*	0.27.		0.46**	0.69***
20.59	-0.06	0.06	0.24	0.16	0.27.	-0.02	0.30.	-0.17	0.29.	0.32*	0.17	-0.01	0.33*	0.41**	0.37*		0.35*
20.98	0.22	0.01	0.34*	0.14	0.22	0.30.	-0.02	0.02	0.22	0.01	0.42**	0.39*	0.38*	0.40**	0.58***	0.22	

Fig 4. Pairwise Spearman correlation coefficients between the mean peak areas of peaks differing significantly among S06 and the major set of isofemale lines in virgin males. Raw data are taken from Fig 1. Top right and bottom left show the correlations with and without the S06 data, respectively. Bold lines between rows and columns separate peaks in the different Rt ranges. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Negative correlations are highlighted in grey.

<https://doi.org/10.1371/journal.pone.0285099.g004>

considering their relative abundances in virgin and mixed flies of each sex, we were able to pair more GC-FID peaks with GC-MS spectra than had originally been possible just using the original GC-MS data of Castro-Vargas et al. [47] and the relatively small number of shared authentic standards. These additional pairings enabled us to assign mass spectra to the respective GC-FID peaks and, in several cases, to assign identities to them. In some cases confirmed identification were possible, because of the greater number of authentic standards used in the GC-MS work. In other cases tentative identifications could be made in the absence of authentic standards because the spectra were of sufficient quality to recognise matches in the National Institute of Standards and Technology (NIST) database.

The SPME GC-MS analyses revealed a total of 97 peaks across a KI range from 674 to 2188 (S2 Dataset). This included all of the 37 peaks we had previously identified with authentic standards (S3 Table). Fifteen peaks (hereafter termed calibrator peaks) of these 37 had also been identified using the corresponding standards in the GC-FID analysis (Table 2). Calibrations of the Rt data from the GC-FID analysis against the KI data from the GC-MS analysis for these 15 (S2 Fig) then allowed us to impute KIs for all 201 of the GC-FID peaks analysed (S1 Table).

We then compared the KI profiles of the GC-FID and GC-MS peaks in the KI interval from 839 to 1783 which was common to both data sets (S4 Table). This region contained 105 GC-FID and 50 GC-MS peaks in male samples, including eight of the 15 calibrator peaks. The 105 GC-FID peaks included 14 of the previously unidentified peaks that had varied significantly between lines. Thirty five of the 105 male GC-FID peaks, including six of the latter 14, were provisionally assessed as being common to the two data sets on the basis of their KI values.

The identities of nine of these 35 (the eight calibrator peaks plus one other peak) were already known because they had been validated with standards in the GC-FID analyses in Castro-Vargas et al. [47]. Another seven of the 35 were discounted as potentially false matches because of differences between their relative abundances across the four sex/mating history categories as calculated from their peak areas in the GC-FID and SPME GC-MS analyses (S4 Table; note that quantitative data had not been collected in the original GC-MS experiment). The latter was a harsh filter because the differences in extraction techniques (the GC-FID, like

Table 4. Details of 19 pairs of GC-FID and GC-MS peaks matched for KI and sex-specific abundances in S06 males. Six of these peaks which showed significant differences among lines in Figs 1 or 2 are indicated with asterisks. The Rts, GC-FID KIs, GC-MS KIs and confirmed or tentative identifications are given. As per Table 2, Rts for peaks classified as intermediate in abundance are underlined while those rated minor in abundance are not. None of the 19 peaks were rated as major in abundance. See S3 Fig for the mass spectra for these 19 peaks and S1 Table for the imputed KIs of all the other GC-FID peaks. As S4 Table shows, a twentieth GC-FID peak, Rt 18.42 and imputed KI 1783, was matched to a GC-MS peak but it had already been identified from a standard in the GC-FID runs and was unavailable for use as a GC-MS standard.

Rts	KI(imp)	KI(obs)	Validated identifications	Tentative identifications
<u>5.20*</u>	885	883		2-Methyl-3-hexanol
6.21*	933	933	Ethyl 2-methylpentanoate	
6.30*	937	939	<i>x</i> -Octenal isomer 2	
<u>6.64</u>	953	955		<i>n</i> -Butylcyclopentane
7.08	974	972		Phenol
8.06*	1021	1019	<i>n</i> -Octen-1-ol	
<u>9.58</u>	1093	1094	N-(2-Methylpropyl) propanamide	
10.05*	1115	1115		
11.16	1161	1164		
<u>11.46</u>	1171	1171	2-Bornanone	
11.5	1172	1173		
11.7	1179	1179		Borneol isomer 1
11.93	1185	1187		Borneol isomer 2
11.97	1186	1188		
13.88	1252	1253		
<u>13.95</u>	1254	1256		
<u>14.37</u>	1267	1267		
17.07*	1463	1464		
17.17	1482	1484		

<https://doi.org/10.1371/journal.pone.0285099.t004>

the original GC-MS data, were based on liquid-liquid hexane extractions) would be expected to affect recovery efficiency anyway, and indeed such differences were seen in some of the eight shared calibrator compounds. Nevertheless, it was considered necessary in order to minimise the risk of false positive matches. It left us with 19 previously anonymous GC-FID peaks in males for which the GC-MS matching then enabled us to confidently assign mass spectral data (S3 Fig). This included all six of the peaks that had shown significant variation between lines above.

Five of the 19 previously anonymous GC-MS-matched GC-FID peaks could also then be identified because the corresponding standards had been included in the original GC-MS analysis (Table 4). They included a short chain alcohol, an aldehyde and an ester in the Short Rt range and an amide and a ketone in the Mid Rt range. Amongst these peaks were three of the six in the shared set that had shown significant variation between lines, namely *n*-octen-1-ol, an *x*-octenal isomer and ethyl 2-methyl pentanoate, none of which have previously been implicated in insect pheromone functions.

Detailed examination of the mass spectra for the other 14 previously anonymous GC-MS-matched GC-FID peaks and comparisons to the NIST library then yielded a further five identifications, which, in the absence of authentic standards, were nevertheless designated as tentative (Table 4). These comprised two more alcohols and an alkane in the short Rt range and two alcohols in the mid KI range. One of the short Rt alcohols, 2-methyl-3-hexanol, was another one of the six in the shared set that had shown significant variation between lines. This compound has not been implicated in insect pheromone functions in previous literature.

One of the other four newly identified compounds, the short Rt alcohol, phenol, may just be a trace metabolite of a more complex phenolic compound(s) because phenol itself is toxic.

However, various phenolics, and phenol itself, have been implicated in pheromone functions in other insects and ticks [86, 87]. The two newly identified mid Rt alcohols are also notable because both are isomers of borneol and closely related to 2-bornanone, a.k.a. camphor, which we had previously identified in S06 rectal glands [47]. Camphor and borneol have both been implicated in sex pheromone functions in another insect [88].

None of the GC-MS peaks had a KI (~1462) or mass spectrum resembling raspberry ketone. This concurs with previous literature which has only recorded the compound in the rectal glands of Qfly laboratory cultures when they are provided it in their diet [31, 73].

Discussion

This appears to be the first study to use isofemale lines to identify inherited intraspecific variation in any tephritid phenotype. The technique has been widely used to detect inherited trait variation in wild populations of *Drosophila* (e.g., [89, 90]) and recently also mosquitoes [83]. It effectively captures and preserves polymorphism in recently collected field populations that might otherwise be lost during domestication. It does this by trapping genetic variants in different lines each established from a different single mated female. In our case, we have then used repeat assays of these lines over multiple generations to confirm the genetic basis of differences in a phenotype which has previously been refractory to genetic analysis in tephritids. While our results clearly show genetic variation between lines reared and tested in a common garden environment, they do not of course exclude the possibility that diet can also contribute to variation in the field. Future work will be able to use crosses among isofemale lines to further dissect the genetics involved, including by modern Quantitative Trait Locus (QTL) analyses. We suggest the isofemale line approach can become an important tool for genetic analyses of diverse traits in tephritids, providing valuable information both for understanding their population biology as well as to guide breeding programs for, for example, SIT mass release strains (see below).

In our case, the limited replication possible in the early post-collection generations of some of the isofemale lines would have limited our power to detect significant line differences in abundances. Nevertheless, we have still found quantitative genetic variation in 26 individual male peaks, representing ~14% of all peaks scored in males, among our isofemale lines and the long domesticated S06 line. Some of the differences were manifest in virgin males, some in mixed males and four in both (albeit, in the latter case, not in the same lines). Notably, just over half of the variable male peaks were male-specific. By contrast, no differences were found among females from either single or mixed sex samples. This accords with the findings of Castro-Vargas et al. [47] who found presence/absence differences between the sexes in ~46% of the 184 peaks they analysed, and quantitative differences between the sexes in another ~34%, with many more of both categories being male-specific or -selective than female-specific or -selective (see also Table 2). The results of both studies thus suggest profoundly different biological functions for the rectal gland volatiles of the two sexes.

The 26 peaks showing significant differences in males covered a wide range of Rt values but a disproportionately high number, eleven, were in the short Rt range (Figs 1 and 2). Given only 31 short Rt peaks in total were assayed, over a third of them varied between the lines. Short Rt peaks would be expected to be more volatile than other peaks and therefore perhaps candidates for longer range sexual attraction or other pheromonal functions. Nine of the ten variable short Rt peaks whose sex specificity/selectivity had been determined were male-specific, which would support the idea that they have a role in male sex pheromone functions. None of the eleven were rated major but five were intermediate in abundance. We were able to identify five of the eleven, namely the esters *n*-propyl 2-methylpropanoate and ethyl 2-methyl

pentanoate, the alcohols 2-methyl 3-hexanol and *n*-octen-1-ol, and an octenal. The latter two are known natural products and all five have been used as fruity flavour and/or fragrance additives [91–95].

Only three compounds in the mid Rt range varied between lines, although two of them were identified, *N*-(2-methylbutyl) acetamide and *N*-(2-methylbutyl) 2-methylpropanamide. Neither of the amides nor the unidentified Rt 10.85 (imputed KI 1151) were male-specific. The first amide was intermediate in abundance but the other two compounds were both rated major in abundance. The two amides were among the six aliphatic amides which have been consistently detected in previous studies of Qfly rectal glands and, in large measure because of their abundance, the amides as a group have been proposed to be male sex pheromones [43–46]. Given that we found only one of the group, *N*-(3-methylbutyl) 2-methylpropanamide, to be male-specific, this appears unlikely to us. On the other hand, in light of the major abundances of one of the amides and Rt 10.85, their variation, which in the case of the latter was evident in both sets of lines, may impact on some other function(s).

Although twelve long Rt compounds were found to vary significantly among lines, they represented fewer than 10% of all long Rt compounds scored. One of the twelve, Rt 17.69, imputed KI 1606, was major, six others intermediate, and five minor in abundance. Only one of the twelve was identified, methyl dodecanoate, which showed no sex bias [47]. Methyl dodecanoate has been found in the labial glands of several species of bumblebees, where it acts as a male sex pheromone [96–99], and it is also present in cuticular profiles of southeast Asian stingless bees [100]. Notably seven of the other variable long Rt compounds were male-specific, whereas none of the eight fatty acid esters whose identity was established by Castro-Vargas et al. [47] in the long Rt ranges were male-specific. Thus at least some of the seven unidentified variable long Rt compounds may not have been fatty acid esters either. Long chain alcohols and alkanes would also be expected to have retention times in this range and have been implicated in various pheromone functions in other insects, either in themselves or as intermediates in the biosynthesis of the active compounds [101].

Totaled across the three Rt ranges, the proportions of the 26 varying peaks that were classified as major, intermediate and minor in abundance (3:13:10, respectively) were not significantly different from the proportions among all the peaks scored. While it is tempting to argue that their relative abundances make the variability of the major and intermediate compounds more relevant biologically, Jang et al. [38] found no relationship between the abundance of several rectal gland volatiles and their biological activity as measured by electroantennogram responses in *C. capitata*.

Both our individual volatile analyses and multivariate analyses found significant differences in male volatiles between the isofemale lines as a group and the domesticated S06 line. Most of the significantly varying peaks in the individual volatile analyses were either substantially more abundant in S06 than any of the isofemale lines or, in a smaller number of cases, amongst the very lowest abundances seen amongst the isofemale lines (and in two cases no longer detectable). Peaks in all three Rt ranges were implicated in the abundance differences between S06 and the isofemale lines. Notable also is the fact that three of the isofemale lines were from Sydney, the same locality as S06. Insomuch as two of the peaks that were higher in S06 included the two significantly varying amides, our data also bear out an earlier finding [58] that the amounts of some of the aliphatic amides increased during domestication. Taken together, all these results suggest a strong but complex pattern of directional selection has operated on many of the rectal gland volatiles during the domestication of S06. It is perhaps to be expected that the high densities in which most laboratory colonies are maintained would select for different male pheromone compositions [68]. The effect could potentially explain some of the

relatively low mating success of mass released sterile males in the field that has been observed in SIT programs for various tephritids [102].

Our analyses have also found differences between isofemale lines both within and across their source localities. The analysis of individual peaks showed peaks in all three Rt ranges and all three abundance categories were involved and the variation within localities in a few cases involved order of magnitude differences. Interestingly, most of the lines in the major set were involved but only about half the minor set lines were involved, and differences between localities were also more pronounced in the major than minor set lines. This may be because the three localities sampled in the major set isofemale lines included the most westerly (Alice Springs), northerly (Cape Tribulation) and southerly (Sydney) populations that we scored. Notably, however, the multivariate analyses of all peaks in the male data sets, whether or not they individually showed significant differences, was also able to differentiate lines from different localities in the minor set that had not been distinguished by the individually varying peaks. In general terms, our findings are consistent with earlier work showing significant geographic variation in desiccation resistance [70] and some genetic and genomic markers [50, 71, 103, 104] in Qflies. Our data thus also support the conclusion from those earlier studies that there is significant ecotypic variation among Qfly populations. Some of the variation we have found may reflect the significant but variable levels of introgression of genes from the sister taxon *B. aquilonis*, which is most prevalent in Qflies from the north-western parts of its range [71].

Significant positive correlations were found between several of the 17 peaks that showed significant variation across lines in virgin males from the major set (Fig 4). This was particularly the case when S06 data were included in the analysis but remained so in several cases when S06 data were excluded. Although r^2 values seldom exceeded 0.5, these correlations imply some shared genetic basis for the variation in the peaks involved. The correlations were strongest between some of the short Rt peaks, between some of the long Rt peaks and, in a small number of cases, between short Rt and long Rt peaks. Interestingly, some of these patterns may be interpretable in terms of the likely biosynthetic pathways for peaks in the different Rt ranges [47, 101]. For example, the long Rt peaks, many of which are known or likely to be fatty acid esters, alcohols, acids and alkanes, are likely to have been generated from acetyl CoA precursors, so some could have shared precursors. Conversely, the available data indicates some of the amides in the mid Rt range are produced by an entirely different process, involving condensation reactions between amino and other organic acids, which might explain why peaks in this Rt range showed relatively few correlations with those in the other ranges. On the whole, however, the conclusion from the correlational analysis is that most of the variation in most peaks had independent genetic bases, implying that variation in multiple genes contributed to the line differences in the volatile contents of virgin male rectal glands.

Overall our data are consistent with the hypothesis [47, 105] that inherited variation exists in male sex pheromones that may facilitate speciation in tephritids. More direct support for this hypothesis will now require a) further GC-MS work to identify some of the 18 still anonymous significantly varying peaks, b) electrophysiological and behavioural tests to assess the functional significance of particular variations, and c) mapping and identification of the key gene(s) involved. While beyond the scope of this article, all three of these objectives are feasible in the near term future.

Expanding the number of identified varying peaks could be achieved by more extensive GC-MS, including use of more extraction techniques (e.g., different SPME fibres [57]) and authentic standards. Our panel of standards was depauperate for long chain alkanes and alcohols, fatty acids and terpenes, which have recurrently been implicated in pheromone functions in various insects [101], as well as spiroacetals, which have been implicated in various such functions in tephritids in particular [32, 33, 106, 107].

Electrophysiological assays focusing on both the antennae and palp are now well established for various bactrocerans [32, 33, 38, 66, 107]. Likewise, behavioural assays in various formats and spatial scales have been developed for measuring different aspects of the attraction response. Results for a range of tephritids with these various formats suggest that short vs long distance attraction, attraction vs retention (arrestation) and opposite-sex vs same- or both-sex attraction/aggregation could each involve distinct pheromones [36, 38, 40, 106, 108, 109], and hence have distinct genetic bases.

The genetic aspects of this subsequent work would be facilitated by the availability of genome assemblies for *B. tryoni* (GenBank accessions GCA_000695345.1 [110] and GCA_016617805.2 (RefSeq)) and the rapidly decreasing costs of full genome resequencing of multiple individuals. Also useful in this respect, our results suggest that a single cross between a domesticated and wild/caught line could simultaneously capture variation in many different volatiles and enable joint mapping of multiple different encoding genes.

A further priority will be to investigate differences in the rectal gland compositions and the key genes controlling them in the two closest relatives of Qfly, namely the Northern Territory fruit fly *B. aquilonis*, with which some hybridisation has occurred [71], and the lesser Queensland fruit fly *B. neohumeralis*, for which there is as yet no evidence of hybridisation in the field, despite extensive sympatry and the fertility of crosses with *B. tryoni* in the laboratory [15]. Genome sequence is also available for *B. neohumeralis* (https://www.ncbi.nlm.nih.gov/data-hub/genome/GCF_024586455.1/).

Finally, the behavioural and genetic analyses which we argue are priorities for further work could also be important to future efforts to improve the efficiency and effectiveness of SIT. Specifically, such analyses could enable marker-assisted breeding during the development of the mass-release strains that minimises the loss of mating competitiveness and ecological competence when the males are released into the field.

Materials and methods

Construction of the isofemale lines

The seven mass-bred populations from which the isofemale lines were made were established from infested fruits brought back to the laboratory between 2017 and 2018 (Table 1). Three days after eclosion, flies identified as *B. tryoni* following the taxonomic key of Drew [13] were used to establish laboratory colonies. Table 1 summarises the collection details for these populations. Further information on them is available in Popa-Báez et al. [70] and Ahmed et al. [111].

Isofemale lines were set up from the mass-bred stocks after one (minor set) and five (major set) generations in the laboratory. Single gravid females from each stock were placed in individual translucent polystyrene containers (425 mL, Dowlings, Canberra, Australia) and allowed to oviposit into a 35 mL plastic sauce bottle (Sistema, New Zealand) that was pierced with small holes and contained a piece of mini capsicum as an attractant to encourage localized egg laying. This oviposition device proved to be critical in boosting egg laying sufficiently to sustain the lines. After 24 h, eggs were collected from the bottle walls by rinsing them with water and placed into 250 mL plastic containers (250 mL, International Scientific Group, UK) containing gel larval diet [112]. The progeny from each female were then used to establish an isofemale line. A total of 23 lines were set up and maintained as described previously [70], except for our use of the mini capsicum as an oviposition substrate.

Extracts of flies from up to four generations were assayed by GC-FID for both the major (generations 6, 7, 8 and 10 post-collection) and minor (generations 6, 8, 10 and 12 post-collection) set isofemale lines (Table 1, S1 Dataset from Castro-Vargas et al. [47], S1 Dataset for this

article). Generally only one replicate extract of glands from ten flies could be taken from each generation but occasionally up to four, all taken from different cages, were available for a particular line. On the other hand, for some lines, only one generation produced enough flies to generate a sample without risking the ongoing maintenance of the line. This was because the productivity of recently collected Qfly strains is generally low [84, 85, 113], likely exacerbated in our case by the inbreeding incurred by founding each line from a single female. The constraints were greatest for the major set isofemale lines where we were seeking four sex/mating history combinations but were reluctant to harvest females in particular until we could see larvae for the following generation. The average numbers of replicates assayed, summed across the different generations, from each isofemale line were 4.7, 1.6, 2.1, 1.1 and 7.0 for the major set virgin and mixed males, major set virgin and mixed females and minor set mixed males, respectively. The corresponding numbers for S06, being thoroughly domesticated and not isofemale, were more consistent, 6, 3, 2, 2 and 6 respectively, and, in the case of the minor set, only required two generations of sampling.

GC-FID methods

The methods for sample preparation for the GC-FID analysis were as in Castro-Vargas et al. [47]. Briefly, flies aged 15–20 days since emergence were frozen at -20°C and their rectal glands excised under a stereoscopic microscope (Leica Microsystems, Heidelberg, Germany). Groups of ten glands were collected in 2 mL glass vials (Shimadzu Corporation, Kyoto, Japan) and immersed in 200 μL of high performance liquid chromatography grade *n*-hexane for 10 min. The hexane was then carefully transferred to 2 mL GC vials with 250 μL glass inserts (Shimadzu Corporation, Kyoto, Japan). The vials were then capped and stored at -80°C until used.

The GC-FID analyses were also carried out as per Castro-Vargas et al. [47]. Briefly, 1 μL samples were injected into an Agilent 7890A gas chromatograph (Agilent Technologies, Santa Clara, USA) fitted with a HP-5ms UI column (30 m x 0.250 mm i.d. x 0.25 μm film thickness; Agilent Technologies) using a split mode of 1/10 and helium as carrier gas at a constant flow rate of 1 mL/min. The oven temperature was initially held at 60°C for 2 min and then increased, first to 100°C at $10^{\circ}\text{C}/\text{min}$ with a 3 min hold, then to 325°C at $5^{\circ}\text{C}/\text{min}$, with a final 10 min hold. The FID detector operated at 300°C . Empty vials with no rectal glands were injected as blanks after every 10 sample injections and at the beginning and end of each analysis sequence. The samples from each of the major and minor sets were run in four batches, with the contents of each batch determined according to the four generations scored (i.e., first generation collected in the first batch, etc). Twenty two authentic standards were also used.

GC-FID data analysis

Data from the blank samples were subtracted from the chromatograms to remove peaks not originating from the rectal glands. Peaks retained from all chromatograms were manually checked and corrected for retention time shift, following automatic alignment with the GCalignR package in R (version 1.0.2 [114]). The maximum accepted linear shift (max_linear_shift) in a peak's retention time (Rt) was 0.05 min, and the maximum and minimum accepted Rt difference from the peak mean (max_diff_peak2mean and min_diff_peak2mean, respectively) was 0.03 min in both. Peak areas were calculated using ACD/Spectrus Processor software version 2019.2.0 (Advanced Chemistry Development Inc., Canada) and the values for different batches were normalised to correct for (minor) batch differences in total peak areas. These values were used as indicators of relative abundances because internal standards were not included in all runs. Only peaks which were detected in at least 50% of samples in at least

one sex/mating history category in at least one line in the respective data set were retained for further analysis.

All statistical analyses of peak areas were performed in R Studio (version 1.3.1093 [115]).

Quantitative analyses to test for differences between S06 and isofemale lines in individual peaks were carried out using linear modelling (lm function, tidyverse package, version 1.3.0 [116]). Following the approaches of Hoffmann and Parsons [117], David et al. [118] and several other studies (e.g. [83, 89, 90]), significant F ratios for the variation between lines as compared to between replicate variation, both within and between generations, indicates heritable differences between the lines compared. For each peak showing significant line variation, we then investigated the contributions of the different lines to the overall difference by calculating the estimated marginal mean for each line, together with its confidence limits, using the emmeans package [119]. All these analyses were carried out after transforming peak area values appropriately to meet the assumptions of normality and homoscedacity required. The BestNormalize package was used to select the appropriate transformation for each peak [120]. OrderNorm, which normalises the ranks of the response variable to a standard normal distribution (i.e., zero mean and unit variance [121]) was selected in almost all cases, including all those where significant differences were found. Emmeans values were back-transformed for display purposes in Figs 1 and 2 using the predict function in the stats package [115, 122]. Significant differences in all analyses were identified after false discovery rate (FDR) adjustment of p values (p.adjust function, tidyverse package [116]).

To test for overall line and locality effects, we first performed PERMANOVAs (adonis function, vegan package [123] on all GC-FID peaks against line effects for each of the three Rt ranges. If line effect was significant, we then ran a PERMANOVA on the data for the respective Rt range against locality, with line nested within the locality effect. To visualise locality differences revealed in the latter analyses, we first performed principal component analyses (PCAs, using the prcomp function) of the GC-FID peaks to obtain a reduced number of variables (fewer than the sample size for each locality). Then linear discriminant analysis (LDA) was conducted on those principal components against locality (lda function, MASS package [124]).

Pairwise Spearman correlations were computed for all pairs of peaks showing significant line differences across all lines in the major set virgin males (with and without the S06 data) and all correlation values were tested for statistically significant differences from zero after FDR adjustment of p values as above.

SPME GC-MS methods

Sample preparation for the SPME GC-MS analysis was as per the GC-FID analysis above except that the dissected glands were placed in 10 ml screw-cap SPME headspace vials (Sigma Aldrich) before being stored at -80°C until analysed. As per Castro-Vargas et al. [47], the SPME vials had been pre-incubated at -400°C over the previous night to remove potential volatile contaminants. Three replicates from different cages of the S06 strain (~ generation 128) were analysed for each sex/mating history combination.

The GC-MS methods followed those of Castro-Vargas et al. [47], again utilising a Shimadzu TQ8050 gas chromatograph (Shimadzu Corporation) fitted with a fused silica DB5 column (30 m x 0.250 mm i.d. x 1 µm film thickness; Agilent Technologies). Briefly, each sample was preincubated for 5 min at 30°C before a 30 min adsorption into a conditioned polydimethylsiloxane (PDMS) SPME fibre (30 µm thickness; Restek Corporation, PA, USA) at 50°C with constant agitation at 750 rpm. Samples were then desorbed at 250°C for 60 s in an injector fitted with a Topaz inlet liner (Restek Corporation) in a splitless injection mode. The carrier gas was helium, at a total flow rate of 1 mL/min. An initial 40°C oven temperature was held for 1

min and then increased, firstly to 100°C at 10°C/min with a 3 min hold, then to 325°C at 30°C/min with a 10 min hold. The ion source was held at 230°C and detection was set for m/z values from 20 to 600, with a scan interval of 0.3 s. Blank controls were used as per the GC-FID methods above.

All chromatograms were analysed with ACD/Spectrus Processor software version 2019.2.0 (Advanced Chemistry Development Inc., Canada). Blank chromatograms were subtracted from sample chromatograms and peaks were manually aligned to correct for retention time drift. Peak areas were calculated as above and quantitative analysis was performed on peak areas for every peak that was present in at least 50% of the samples in at least one of the four sex/mating history categories. Thirty seven of the peaks had previously been identified by matching equivalent extracts against authentic standards [47, 57, 61, 125]. Other peaks were tentatively identified by detailed analysis of their spectra and comparison of their spectra with the NIST libraries of known compounds with similar KIs (MS Search version 2.3).

Supporting information

S1 Fig. Illustrative GC-FID chromatograms. Part A shows a complete chromatogram from a mixed S06 male sample (upper) and a segmented version of the same chromatogram (lower) in which the Y axes of certain segments have been adjusted to facilitate visualisation of smaller peaks. The 15 ‘calibrator’ peaks used in the cross referencing to GC-MS data (see text) are labelled, with abbreviations as per [S3 Table](#). **Part B** shows segmented, Y axes-adjusted, chromatograms of three samples in which the 26 significantly varying peaks are labelled according to their R_t values and, where known, the abbreviations for their names from [S3 Table](#). The upper, middle and lower samples are from NB28, MB64 and S06 mixed males, respectively. (DOCX)

S2 Fig. R_t -KI calibration curve for imputing KI values for GC-FID peaks. Observed KIs from the GC-MS analysis for the 15 “calibrator” peaks (see [S1 Fig](#)) whose identities had been determined with both technologies were plotted against the R_t s for those compounds in the GC-FID analyses. Linear relationships were evident for peaks up to R_t 13.59 (green), and from R_t 17.37 onwards (blue) and a multinomial function was fitted for the intervening region (red). Best fit equations for the three relationships are also shown. KI values for all the other GC-FID peaks were imputed by interpolation between the values for adjacent peaks with known KIs in the two linear ranges and using the polynomial function indicated for the intervening region. (DOCX)

S3 Fig. Mass spectra and, where identified, names of the 19 GC-MS peaks matched with GC-FID peaks in S06 males in [S4 Table](#). Peaks are labelled according to their GC-FID R_t s and imputed KIs. Names for the five compounds identified against authentic standards are given without asterisks while the five that were tentatively identified by reference to the NIST database are indicated with an asterisk. Also shown are extracts of the gas chromatograms from the GC-MS analysis showing the 19 peaks. R_t s on those chromatograms do not match those in the GC-FID analyses because of differences in the chromatographic procedures. (DOCX)

S1 Table. Linear modelling results for peaks in each of the five data sets and imputed KI values for all peaks. F values from the modelling are shown and significance after FDR correction is indicated with asterisks. * $p < 0.05$, * $p < 0.01$, * $p < 0.001$. A dash indicates the peak was not detected in that data set. Imputed KI values were determined from the calibration

curve in [S2 Fig](#).
(DOCX)

S2 Table. Emmeans for OrderNorm transformed peak areas, plus their 95% confidence limits, for peaks showing significant variation between lines. Panel A shows the results for virgin males from the major set of lines, Panel B shows the results for mixed males from the major set and Panel C the results for mixed males from the minor set of lines. As per [Table 2](#), peaks are bolded, underlined or neither if they are classified as major, intermediate or minor in abundance. Their sex bias (also taken from [Table 2](#)) in samples from single/mixed sex cohorts is indicated as follows: Msp = male specific, Mse = male selective, Fse = female specific, Fse = female selective, nd = not detected and ~ = not sex biased, and virgin and mixed sex mating history categories are indicated before and after the slash, respectively. FDR-corrected significance values for F statistics are indicated with asterisks (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). The OrderNorm transformation used to ensure normality and homoscedacity also scaled the values to zero mean and unit variance. Confidence limits are given in parentheses. Letter codes are used to show the results of *post hoc* pairwise contrasts between lines, where lines with the same letters are not significantly different from one another. Note that peaks 4.17 and 8.06 in Panel C do not appear in [Table 2](#) because they did not meet the inclusion criterion (detected in $> 50\%$ of samples in at least one sex/mating status category in at least one line) for analysis in the major set of lines and therefore were not classified for abundance or sex bias.

(DOCX)

S3 Table. Thirty seven previously identified rectal gland GC-MS peaks in Castro-Vargas et al. *. KIs given in that study and abbreviations used in [S1](#) and [S3](#) Figs are also shown. Four additional compounds tentatively identified by GC-MS in the current study were 2-methyl 3-hexanol, *n*-butylcyclopentane and two isomers of borneol (2M3H, NBCP, Bor1 and Bor2, respectively).

(DOCX)

S4 Table. Alignment of GC-FID and GC-MS profiles in S06 males. For each of the two technologies, data are given for KI, maximum abundance ("Abund.", i.e., relative abundance in the sex/mating history where it was highest; maj, int and min = major, intermediate and minor, respectively), sex-dependent abundances in virgin and mixed samples ("Sex differences"), and, where known, compound names. Rt values and significant line variation ("sig") are also indicated for the GC-FID peaks. Peaks are listed in ascending KI order, with the data for calibrator peaks with the two technologies on the same line. Also shown on the same line are 27 pairs of peaks whose imputed GC-FID KI and observed GC-MS KI were no more than four units apart (Best match) and clearly more similar to each other than the immediately adjacent peaks (Next before and Next after). These 27 pairs were then further filtered in respect of their sex-dependent abundances in virgin and mixed samples (before or after the comma respectively, where M or F = present in all male and female samples, respectively, m or f = present in at least one but not all of the respective samples, nd = not detected, na = not assayed, and $>$ or $<$ represent differences of at least 10 fold in the mean peak areas for the two sexes). Compatibility was rated high (YY) for 15 pairs of peaks whose sex-specific expression in each type of sample was the same with both technologies (whether present in all replicates of each type of sample or not, for peaks 5.20, 6.64, 9.58, 10.05, 11.46, 11.70, 11.93, 11.97, 13.95, 14.37, 18.42) or not assayed with one technology (minor data set peaks 7.08, 8.06, 11.16, 11.50). Compatibility was considered acceptable (Y) for five other pairs showing a presence/absence difference in which the presence involved an m or f in a peak whose maximum abundance was still

only minor (peaks 6.21, 13.88), or the presence still showed a greater than ten fold bias towards the sex in question (peaks 6.30, 17.07, 17.17). The remaining seven of the 27 KI-matched pairs were considered incompatible (N, for peaks 4.23, 11.04, 15.38, 15.59, 16.32, 16.94, 17.01) in respect of their sex-dependent abundances in either virgin or mixed samples. (DOCX)

S1 Dataset. Raw peak areas of peaks detected by GC-FID in the minor set of isofemale lines. This Data Set contains the peaks that were each present in at least half the samples in at least one of the four sex/mating history categories. The “Cage” column denotes the replicates for each of the lines in each generation. “Gen” refers to the generation in which the lines were screened. Headings for columns F—DU give the respective retention times (Rts) of the peaks screened and the data in body of the table are peak areas after baseline adjustment and removal of contaminants from the traces. (CSV)

S2 Dataset. Raw peak areas of S06 peaks detected by SPME GC-MS. This Data Set contains all the peaks that were present in at least half the S06 samples in at least one of the sex/mating history categories. Observed KIs are shown across the top and the data in the body of the table are peak areas after baseline adjustment and removal of contaminants from the traces. (CSV)

Acknowledgments

We thank Matt Taylor and Greg Dojchinov for advice and assistance with the analytical chemistry, Jeanneth Pérez, Vivian Mendez and Sally Noushini for their advice on tephritid semiochemistry, Constance Allegre for the Cape Tribulation sampling, Michele Schiffer for species identification and Valentina Colombo for assisting with the establishment of the isofemale lines.

Author Contributions

Conceptualization: Cynthia Castro-Vargas, Gunjan Pandey, Siu Fai Lee, Phillip W. Taylor, John G. Oakeshott.

Formal analysis: Cynthia Castro-Vargas, Gunjan Pandey, Heng Lin Yeap, Michael J. Lacey.

Funding acquisition: Phillip W. Taylor.

Investigation: Cynthia Castro-Vargas, Gunjan Pandey, Shirleen S. Prasad, Siu Fai Lee, Phillip W. Taylor, John G. Oakeshott.

Methodology: Cynthia Castro-Vargas, Gunjan Pandey, Siu Fai Lee, Soo J. Park.

Project administration: Phillip W. Taylor.

Resources: Siu Fai Lee, Phillip W. Taylor, John G. Oakeshott.

Supervision: Gunjan Pandey, Siu Fai Lee, Phillip W. Taylor, John G. Oakeshott.

Writing – original draft: Cynthia Castro-Vargas, Gunjan Pandey, John G. Oakeshott.

Writing – review & editing: Cynthia Castro-Vargas, Gunjan Pandey, Heng Lin Yeap, Michael J. Lacey, Siu Fai Lee, Soo J. Park, Phillip W. Taylor, John G. Oakeshott.

References

1. Billeter JC, Atallah J, Krupp JJ, Millar JG, Levine JD. Specialized cells tag sexual and species identity in *Drosophila melanogaster*. *Nature*. 2009; 461: 987–991. <https://doi.org/10.1038/nature08495> PMID: 19829381
2. Smadja C, Butlin RK. On the scent of speciation: The chemosensory system and its role in premating isolation. *Heredity*. 2009; 102: 77–97. <https://doi.org/10.1038/hdy.2008.55> PMID: 18685572
3. Dopman EB, Robbins PS, Seaman A. Components of reproductive isolation between North American pheromone strains of the European corn borer. *Evolution*. 2010; 64: 881–902. <https://doi.org/10.1111/j.1558-5646.2009.00883.x> PMID: 19895559
4. Shahandeh MP, Pischedda A, Turner TL. Male mate choice via cuticular hydrocarbon pheromones drives reproductive isolation between *Drosophila* species. *Evolution*. 2018; 72: 123–135. <https://doi.org/10.1111/evo.13389> PMID: 29098691
5. Lassance JM, Groot AT, Liénard MA, Antony B, Borgwardt C, Andersson F, et al. Allelic variation in a fatty-acyl reductase gene causes divergence in moth sex pheromones. *Nature*. 2010; 466: 486–489. <https://doi.org/10.1038/nature09058> PMID: 20592730
6. Groot AT, Schöfl G, Inglis O, Donnerhacke S, Classen A, Schmalz A, et al. Within-population variability in a moth sex pheromone blend: Genetic basis and behavioural consequences. *Proc R Soc B Biol Sci*. 2014; 281. <https://doi.org/10.1098/rspb.2013.3054> PMID: 24500170
7. McElfresh JS, Millar JG. Geographic variation in the pheromone system of the saturniid moth *Hemileuca eglanterina*. *Ecology*. 2001; 82: 3505–3518.
8. McElfresh JS, Millar JG. Geographic variation in sex pheromone blend of *Hemileuca electra* from southern California. *J Chem Ecol*. 1999; 25: 2505–2525. <https://doi.org/10.1023/A:1020826208715>
9. Groot AT, Inglis O, Bowdridge S, Santangelo RG, Blanco C, López JD, et al. Geographic and temporal variation in moth chemical communication. *Evolution*. 2009; 63: 1987–2003. <https://doi.org/10.1111/j.1558-5646.2009.00702.x> PMID: 19473383
10. Duménil C, Judd GJR, Bosch D, Baldessari M, Gemeno C, Groot AT. Intraspecific variation in female sex pheromone of the codling moth *Cydia pomonella*. *Insects*. 2014; 5: 705–721. <https://doi.org/10.3390/insects5040705> PMID: 26462935
11. Spiegel CN, Dias dos Santos DB, Araki AS, Hamilton JGC, Brazil RP, Jones TM. The *Lutzomyia longipalpis* complex: A brief natural history of aggregation-sex pheromone communication. *Parasites and Vectors*. 2016; 9: 1–15. <https://doi.org/10.1186/s13071-016-1866-x> PMID: 27842601
12. Hamilton JGC, Maingon RDC, Alexander B, Ward RD, Brazil RP. Analysis of the sex pheromone extract of individual male *Lutzomyia longipalpis* sandflies from six regions in Brazil. *Med Vet Entomol*. 2005; 19: 480–488. <https://doi.org/10.1111/j.1365-2915.2005.00594.x> PMID: 16336313
13. Drew RAI. The tropical fruit flies (Diptera: Tephritidae: Dacinae) of the Australasian and Oceanian regions. Brisbane, Australia: Memoirs of the Queensland Museum; 1989.
14. Hancock D, Drew R. Phylogeny of the tribe Dacini (Dacinae) based on morphological, distributional, and biological data. In: Aluja M, Norrbom AL, editors. *Fruit flies (Tephritidae): Phylogeny and evolution of behavior*. Boca Raton: CRC Press; 1999. pp. 491–504. <https://doi.org/10.1201/9781420074468.ch19>
15. Yeap HL, Lee SF, Robinson F, Mourant RG, Sved JA, Frommer M, et al. Separating two tightly linked species-defining phenotypes in *Bactrocera* with hybrid recombinant analysis. *BMC Genet*. 2020; 21: 1–16. <https://doi.org/10.1186/s12863-020-00936-1> PMID: 33339498
16. Perkins M V., Fletcher MT, Kitching W, Drew RAI, Moore CJ. Chemical studies of rectal gland secretions of some species of *Bactrocera dorsalis* complex of fruit flies (Diptera: Tephritidae). *J Chem Ecol*. 1990; 16: 2475–2487. <https://doi.org/10.1007/BF01017470> PMID: 24264212
17. Clarke AR, Powell KS, Weldon CW, Taylor PW. The ecology of *Bactrocera tryoni* (Diptera: Tephritidae): What do we know to assist pest management? *Ann Appl Biol*. 2011; 158: 26–54. <https://doi.org/10.1111/j.1744-7348.2010.00448.x>
18. White IM, Elson-Harris MM. *Fruit flies of economic significance: Their identification and bionomics*. Wallingford, UK: CAB International; 1992.
19. Dominiak BC, Clifford CS, Nielsen SG. Queensland fruit fly (*Bactrocera tryoni* Froggatt) attraction to and chemical analysis of male annihilation blocks using three concentrations of cuelure at Dubbo, New South Wales, Australia. *Plant Prot Q*. 2009; 24: 157–160.
20. Dominiak BC, Ekman J, Broughton S. Mass trapping and other management options for Mediterranean fruit fly and Queensland fruit fly in Australia. *Gen Appl Entomol*. 2016; 44: 1–8.
21. Vargas RI, Souder SK, Mackey B, Cook P, Morse JG, Stark JD. Field trials of solid triple lure (trime-dlure, methyl eugenol, raspberry ketone, and DDVP) dispensers for detection and male annihilation of

- Ceratitis capitata*, *Bactrocera dorsalis*, and *Bactrocera cucurbitae* (Diptera: Tephritidae) in Hawaii. J Econ Entomol. 2012; 105: 1557–1565. <https://doi.org/10.1603/ec12122> PMID: 23156150
22. Reynolds OL, Osborne T, Crisp P, Barchia IM. Specialized pheromone and lure application technology as an alternative male annihilation technique to manage *Bactrocera tryoni* (Diptera: Tephritidae). J Econ Entomol. 2016; 109: 1254–1260. <https://doi.org/10.1093/jeetow023> PMID: 27018434
 23. Vargas RI, Souder SK, Hoffman K, Mercogliano J, Smith TR, Hammond J, et al. Attraction and mortality of *Bactrocera dorsalis* (Diptera: Tephritidae) to STATIC Spinosad ME weathered under operational conditions in California and Florida: A reduced-risk male annihilation treatment. J Econ Entomol. 2014; 107: 1362–1369. <https://doi.org/10.1603/EC14121> PMID: 25195423
 24. Wang HL, Ding BJ, Dai JQ, Nazarens TJ, Borges R, Mafra-Neto A, et al. Insect pest management with sex pheromone precursors from engineered oilseed plants. Nat Sustain. 2022; 5: 981–990. <https://doi.org/10.1038/s41893-022-00949-x>
 25. Caravantes-Villatoro LA, Liedo P, Guillén-Navarro K, Rojas JC. Effect of a probiotic-enriched diet on sexual competitiveness, pheromone emission, and cuticular hydrocarbons of sterile and fertile *Anastrepha ludens* (Diptera: Tephritidae). J Econ Entomol. 2022; 115: 1–9. <https://doi.org/10.1093/jeet/toac105> PMID: 35848878
 26. Fletcher BS. The biology of dactine fruit flies. Annu Rev Entomol. 1987; 32: 115–144.
 27. Tan KH, Nishida R. Zingerone in the floral synomone of *Bulbophyllum baileyi* (Orchidaceae) attracts *Bactrocera* fruit flies during pollination. Biochem Syst Ecol. 2007; 35: 334–341. <https://doi.org/10.1016/j.bse.2007.01.013>
 28. Hee AKW, Tan KH. Transport of methyl eugenol-derived sex pheromonal components in the male fruit fly, *Bactrocera dorsalis*. Comp Biochem Physiol—Part C Toxicol Pharmacol. 2006; 143: 422–428. <https://doi.org/10.1016/j.cbpc.2006.04.009> PMID: 16750428
 29. Nishida R, Fukami H. Sequestration of distasteful compounds by some pharmacophagous insects. J Chem Ecol. 1990; 16: 151–164. <https://doi.org/10.1007/BF01021276> PMID: 24264904
 30. Tan KH, Nishida R. Sex pheromone and mating competition after methyl eugenol consumption in the *Bactrocera dorsalis* complex. Fruit Fly Pests. 1996; 147–153. <https://doi.org/10.1201/9780367812430-23>
 31. Kumaran N, Hayes RA, Clarke AR. Cuelure but not zingerone make the sex pheromone of male *Bactrocera tryoni* (Tephritidae: Diptera) more attractive to females. J Insect Physiol. 2014; 68: 36–43. <https://doi.org/10.1016/j.jinsphys.2014.06.015> PMID: 25010549
 32. Noushini S, Pérez J, Jean Park S, Holgate D, Mendez Alvarez V, Jamie I, et al. Attraction and electrophysiological response to identified rectal gland volatiles in *Bactrocera frauenfeldi* (Schiner). Molecules. 2020; 25: 1–14. <https://doi.org/10.3390/molecules25061275> PMID: 32168881
 33. Noushini S, Park SJ, Jamie I, Jamie J, Taylor P. Rectal gland exudates and emissions of *Bactrocera bryoniae*: Chemical identification, electrophysiological and pheromonal functions. Chemoecology. 2021; 31: 137–148. <https://doi.org/10.1007/s00049-020-00335-z>
 34. Noushini S, Pérez J, Park SJ, Holgate D, Jamie I, Jamie J, et al. Rectal gland chemistry, volatile emissions, and antennal responses of male and female banana fruit fly, *Bactrocera musae*. Insects. 2020; 11: 32. <https://doi.org/10.3390/insects11010032> PMID: 31906084
 35. Robacker DC, Hart WG. Courtship and territoriality of laboratory-reared Mexican fruit flies, *Anastrepha ludens* (Diptera: Tephritidae), in cages containing host and nonhost trees. Ann Entomol Soc Am. 1985; 78: 488–494. <https://doi.org/10.1093/aesa/78.4.488>
 36. Robacker DC. Behavioral responses of female Mexican fruit flies, *Anastrepha ludens*, to components of male-produced sex pheromone. J Chem Ecol. 1988; 14: 1715–1726. <https://doi.org/10.1007/BF01014639> PMID: 24276529
 37. Ohinata K, Nakagawa S, Fujimoto M, Higa H, Jacobson M. Mediterranean fruit fly: Laboratory and field evaluations of synthetic sex pheromones. J Environ Sci Heal—Part A Environ Sci Eng. 1977; 12: 67–78. <https://doi.org/10.1080/10934527709374736>
 38. Jang EB, Light DM, Flath RA, Nagata JT, Mon TR. Electroantennogram responses of the Mediterranean fruit fly, *Ceratitis capitata* to identified volatile constituents from calling males. Entomol Exp Appl. 1989; 63: 13–26. <https://doi.org/10.1111/j.1570-7458.1992.tb02415.x>
 39. Landolt PJ, Heath RR, Chambers DL. Oriented flight responses of female Mediterranean fruit flies to calling males, odor of calling males, and a synthetic pheromone blend. Entomol Exp Appl. 1992; 65: 259–266. <https://doi.org/10.1111/j.1570-7458.1992.tb00679.x>
 40. Jang EB, Light DM, Binder RG, Flath RA, Carvalho LA. Attraction of female Mediterranean fruit flies to the five major components of male-produced pheromone in a laboratory flight tunnel. J Chem Ecol. 1994; 20: 9–20. <https://doi.org/10.1007/BF02065987> PMID: 24241695

41. Baker R, Bacon AJ. The identification of spiroacetals in the volatile secretions of two species of fruit fly (*Dacus dorsalis*, *Dacus curcurbitae*). *Experientia*. 1985; 41: 1484–1485. <https://doi.org/10.1007/BF01950049>
42. Ohinata K, Jacobson M, Kobayashi RM, Chambers DL, Fujimoto MS, Higa HH. Oriental fruit fly and melon fly; Biological and chemical studies of smoke produced by males. *J Environ Sci Heal—Part A Environ Sci Eng*. 1982; 17: 197–216. <https://doi.org/10.1080/10934528209375028>
43. Bellas TE, Fletcher BS. Identification of the major components in the secretion from the rectal pheromone glands of the Queensland fruit flies *Dacus tryoni* and *Dacus neohumeralis* (Diptera: Tephritidae). *J Chem Ecol*. 1979; 5: 795–803.
44. Fletcher BS. The structure and function of the sex pheromone glands of the male Queensland fruit fly, *Dacus tryoni*. *Insect Physiol*. 1969; 15: 1309–1322.
45. Tychsen PH. Mating behaviour of the Queensland fruit fly, *Dacus tryoni* (Diptera: Tephritidae), in field cages. *Aust J Entomol*. 1978; 16: 459–465. <https://doi.org/10.1111/j.1440-6055.1977.tb00139.x>
46. Fletcher MT, Kitching W. Chemistry of fruit flies. *Chem Rev*. 1995; 95: 789–828. <https://doi.org/10.1021/cr00036a001>
47. Castro-Vargas C, Pandey G, Yeap HL, Lacey MJ, Lee SF, Park SJ, et al. Diversity and sex differences in rectal gland volatiles of Queensland fruit fly, *Bactrocera tryoni* (Diptera: Tephritidae). *PLoS One*. 2022; 17: e0273210. <https://doi.org/10.1371/journal.pone.0273210> PMID: 36001616
48. Symonds MRE, Moussalli A, Elgar MA. The evolution of sex pheromones in an ecologically diverse genus of flies. *Biol J Linn Soc*. 2009; 97: 594–603. <https://doi.org/10.1111/j.1095-8312.2009.01245.x>
49. Hancock DL, Hamacek EL, Lloyd AC, Elson-Harris MM. The distribution and host plants of fruit flies (Diptera: Tephritidae) in Australia. Brisbane, Australia: Information Series; Queensland Dept. of Primary Industries; 2000.
50. Popa-Báez ÁD, Lee SF, Yeap HL, Westmore G, Crisp P, Li D, et al. Tracing the origins of recent Queensland fruit fly incursions into South Australia, Tasmania and New Zealand. *Biol Invasions*. 2021; 23: 1–14. <https://doi.org/10.1007/s10530-020-02422-2>
51. Sultana S, Baumgartner JB, Dominiak BC, Royer JE, Beaumont LJ. Potential impacts of climate change on habitat suitability for the Queensland fruit fly. *Sci Rep*. 2017; 7: 1–10. <https://doi.org/10.1038/s41598-017-13307-1> PMID: 29026169
52. Fletcher BS. Storage and release of a sex pheromone by the Queensland fruit fly, *Dacus tryoni* (Diptera: Tephritidae). *Nature*. 1968; 219: 631–632. <https://doi.org/10.1038/219631a0> PMID: 5665720
53. Fletcher BS, Giannakakis A. Factors limiting the response of females of the Queensland fruit fly, *Dacus tryoni*, to the sex pheromone of the male. *J Insect Physiol*. 1973; 19: 1147–1155. [https://doi.org/10.1016/0022-1910\(73\)90038-3](https://doi.org/10.1016/0022-1910(73)90038-3)
54. Giannakakis A, Fletcher BS. An improved bioassay technique for the sex pheromone of male *Dacus tryoni* (Diptera: Tephritidae). *Can Entomol*. 1978; 110: 125–129. <https://doi.org/10.4039/Ent110125-2>
55. Monro J. Stridulation in the Queensland fruit fly *Dacus* (Strumeta) *tryoni* Frogg. *Aust J Sci*. 1953; 16: 60–62.
56. Mankin RW, Lemon M, Harmer AMT, Evans CS, Taylor PW. Time-pattern and frequency analyses of sounds produced by irradiated and untreated male *Bactrocera tryoni* (Diptera: Tephritidae) during mating behavior. *Ann Entomol Soc Am*. 2008; 101: 664–674. [https://doi.org/10.1603/0013-8746\(2008\)101\[664:tfaos\]2.0.co;2](https://doi.org/10.1603/0013-8746(2008)101[664:tfaos]2.0.co;2)
57. Noushini S, Park SJ, Jamie I, Jamie J, Taylor PW. Sampling technique biases in the analysis of fruit fly volatiles: A case study of Queensland fruit fly. *Sci Rep*. 2020; 10: 1–14. <https://doi.org/10.1038/s41598-020-76622-0> PMID: 33188282
58. Pérez J, Park SJ, Taylor PW. Domestication modifies the volatile emissions produced by male Queensland fruit flies during sexual advertisement. *Sci Rep*. 2018; 8: 1–10. <https://doi.org/10.1038/s41598-018-34569-3> PMID: 30405136
59. Lewis JA. Characterization and synthesis of dacine fruit fly pheromones. PhD thesis. University of Queensland, Australia. 1988.
60. Booth YK, Schwartz BD, Fletcher MT, Lambert LK, Kitching W, Voss JJD. A diverse suite of spiroacetals, including a novel branched representative, is released by female *Bactrocera tryoni* (Queensland fruit fly). *Chem Commun*. 2006; 3975–3977. <https://doi.org/10.1039/b611953k> PMID: 17003870
61. El-Sayed AM, Venkatesham U, Unelius CR, Sporle A, Pérez J, Taylor PW, et al. Chemical composition of the rectal gland and volatiles released by female Queensland fruit fly, *Bactrocera tryoni* (Diptera: Tephritidae). *Environ Entomol*. 2019; 48: 807–814. <https://doi.org/10.1093/ee/nvz061> PMID: 31145449

62. Scolari F, Valerio F, Benelli G, Papadopoulos NT, Vaničková L. Tephritid fruit fly semiochemicals: Current knowledge and future perspectives. *Insects*. 2021; 12: 1–56. <https://doi.org/10.3390/insects12050408> PMID: 33946603
63. Benelli G, Daane KM, Canale A, Niu CY, Messing RH, Vargas RI. Sexual communication and related behaviours in Tephritidae: Current knowledge and potential applications for Integrated Pest Management. *J Pest Sci*. 2014; 87: 385–405. <https://doi.org/10.1007/s10340-014-0577-3>
64. Baker R, Herbert R, Howse PE, Jones OT, Francke W, Reith W. Identification and synthesis of the major sex pheromone of the olive fly (*Dacus oleae*). *Chem Commun*. 1980; 52–53. <https://doi.org/10.1039/c39800000052>
65. Mazomenos BE, Haniotakis GE. A multicomponent female sex pheromone of *Dacus oleae* Gmelin: Isolation and bioassay. *J Chem Ecol*. 1981; 7: 437–444. <https://doi.org/10.1007/BF00995766> PMID: 24420489
66. Canale A, Germinara SG, Carpita A, Benelli G, Bonsignori G, Stefanini C, et al. Behavioural and electrophysiological responses of the olive fruit fly, *Bactrocera oleae* (Rossi) (Diptera: Tephritidae), to male- and female-borne sex attractants. *Chemoecology*. 2013; 23: 155–164. <https://doi.org/10.1007/s00049-013-0131-4>
67. Vaničková L. Chemical ecology of fruit flies genera *Ceratitis* and *Anastrepha*. PhD Thesis, Institute of Chemical Technology. 2012. <https://doi.org/10.13140/2.1.1221.0885>
68. Vaničková L, Do Nascimento RR, Hoskovec M, Ježková Z, Břizová R, Tomčala A, et al. Are the wild and laboratory insect populations different in semiochemical emission? The case of the medfly sex pheromone. *J Agric Food Chem*. 2012; 60: 7168–7176. <https://doi.org/10.1021/jf301474d> PMID: 22741541
69. Břizová R, Mendonça AL, Vaničková L, Mendonça AL, Eduardo Da Silva C, Tomčala A, et al. Pheromone analyses of the *Anastrepha fraterculus* (Diptera: Tephritidae) cryptic species complex. *Florida Entomol*. 2013; 96: 1107–1115. <https://doi.org/10.1653/024.096.0351>
70. Popa-Báez ÁD, Lee SF, Yeap HL, Prasad SS, Schiffer M, Mourant RG, et al. Climate stress resistance in male Queensland fruit fly varies among populations of diverse geographic origins and changes during domestication. *BMC Genet*. 2020; 21: 135. <https://doi.org/10.1186/s12863-020-00935-2> PMID: 33339509
71. Popa-Báez ÁD, Catullo R, Lee SF, Yeap HL, Mourant RG, Frommer M, et al. Genome-wide patterns of differentiation over space and time in the Queensland fruit fly. *Sci Rep*. 2020; 10: 1–14. <https://doi.org/10.1038/s41598-020-67397-5> PMID: 32612249
72. Wee SL, Tan KH. Temporal accumulation of phenylpropanoids in male fruit flies, *Bactrocera dorsalis* and *B. carambolae* (Diptera: Tephritidae) following methyl eugenol consumption. *Chemoecology*. 2007; 17: 81–85. <https://doi.org/10.1007/s00049-006-0363-7>
73. Tan KH, Nishida R. Incorporation of raspberry ketone in the rectal glands of males of the Queensland fruit fly, *Bactrocera tryoni* Froggatt (Diptera: Tephritidae). *Appl Entomol Zool*. 1995; 30: 494–497. <https://doi.org/10.1303/aez.30.494>
74. Nishida R, Iwahashi O, Tan KH. Accumulation of *Dendrobium superbum* (orchidaceae) fragrance in the rectal glands by males of the melon fly, *Dacus cucurbitae*. *J Chem Ecol*. 1993; 19: 713–722. <https://doi.org/10.1007/BF00985003> PMID: 24249012
75. Takeoka G, Buttery RG, Flath RA, Teranishi R, Wheeler EL, Wieczorek RL, et al. Volatile constituents of pineapple (*Ananas comosus* [L.] Merr.). Flavor chemistry: Trends and developments. 1989. pp. 223–237. <https://doi.org/10.1021/bk-1989-0388.ch018>
76. Schultz TH, Flath RA, Mon TR. Analysis of orange volatiles with vapor sampling. *J Agric Food Chem*. 1971; 19: 1060–1065. <https://doi.org/10.1021/jf60178a019>
77. Matich AJ, Rowan DD, Banks NH. Solid phase microextraction for quantitative headspace sampling of apple volatiles. *Anal Chem*. 1996; 68: 4114–4118. <https://doi.org/10.1021/ac9604548> PMID: 8946792
78. Jennings WG. Volatile components of figs. *Food Chem*. 1977; 2: 185–191. [https://doi.org/10.1016/0308-8146\(77\)90032-2](https://doi.org/10.1016/0308-8146(77)90032-2)
79. Cunningham JP, Carlsson MA, Villa TF, Dekker T, Clarke AR. Do fruit ripening volatiles enable resource specialism in polyphagous fruit flies? *J Chem Ecol*. 2016; 42: 931–940. <https://doi.org/10.1007/s10886-016-0752-5> PMID: 27586434
80. Weldon CW, Perez-Staples D, Taylor PW. Feeding on yeast hydrolysate enhances attraction to cue-lure in Queensland fruit flies, *Bactrocera tryoni*. *Entomol Exp Appl*. 2008; 129: 200–209. <https://doi.org/10.1111/j.1570-7458.2008.00768.x>
81. Dominiak BC, Nicol HI. Chemical analysis of male annihilation blocks used in the control of Queensland fruit fly *Bactrocera tryoni* (Froggatt) in New South Wales. *Plant Prot Q*. 2012; 27: 31–35.

82. Lloyd AC, Hamacek EL, Kopittke RA, Peek T, Wyatt PM, Neale CJ, et al. Area-wide management of fruit flies (Diptera: Tephritidae) in the Central Burnett district of Queensland, Australia. *Crop Prot.* 2010; 29: 462–469. <https://doi.org/10.1016/j.cropro.2009.11.003>
83. Ross PA, Endersby-Harshman NM, Hoffmann AA. A comprehensive assessment of inbreeding and laboratory adaptation in *Aedes aegypti* mosquitoes. *Evol Appl.* 2019; 12: 572–586. <https://doi.org/10.1111/eva.12740> PMID: 30828375
84. Bartlett AC. Genetic changes during insect domestication. In: King EG, Leppla NC, editors. *Advances and challenges in insect rearing.* New Orleans, LA, USA: USDA/ARS; 1984. pp. 2–8.
85. Gilchrist AS, Cameron EC, Sved JA, Meats AW. Genetic consequences of domestication and mass rearing of pest fruit fly *Bactrocera tryoni* (Diptera: Tephritidae). *J Econ Entomol.* 2012; 105: 1051–1056. <https://doi.org/10.1603/EC11421> PMID: 22812147
86. Wood WF, Leahy SMG, Galun R, Prestwich GD, Meinwald J, Purnell RE, et al. Phenols as pheromones of ixodid ticks: A general phenomenon? *J Chem Ecol.* 1975; 1: 501–509. <https://doi.org/10.1007/BF00988590>
87. Marshall DG, Jackson TA, Unelius CR, Wee SL, Young SD, Townsend RJ, et al. *Morganella morganii* bacteria produces phenol as the sex pheromone of the New Zealand grass grub from tyrosine in the colleterial gland. *Sci Nat.* 2016; 103: 1–6. <https://doi.org/10.1007/s00114-016-1380-1> PMID: 27352077
88. de Melo DJ, da Silva MR, Moliterno AAC, Zarbin PHG. Camphor and borneol as the male-produced sex pheromone of the shield bug, *Orsillochides leucoptera* (Hemiptera: Scutelleridae). *J Chem Ecol.* 2020; 46: 490–496. <https://doi.org/10.1007/s10886-020-01189-z> PMID: 32494930
89. Hoffmann AA, Hallas R, Sinclair C, Mitrovski P. Levels of variation in stress resistance in *Drosophila* among strains, local populations, and geographic regions: Patterns for desiccation, starvation, cold resistance, and associated traits. *Evolution.* 2001; 55: 1621–1630. <https://doi.org/10.1111/j.0014-3820.2001.tb00681.x> PMID: 11580021
90. Griffiths JA, Schiffer M, Hoffmann AA. Clinal variation and laboratory adaptation in the rainforest species *Drosophila birchii* for stress resistance, wing size, wing shape and development time. *J Evol Biol.* 2005; 18: 213–222. <https://doi.org/10.1111/j.1420-9101.2004.00782.x> PMID: 15669978
91. Mierendorff HG, Stahl-Biskup E, Posthumus MA, van Beek TA. Composition of commercial Cape chamomile oil (*Eriocephalus punctulatus*). *Flavour Fragr J.* 2003; 18: 510–514. <https://doi.org/10.1002/ffj.1259>
92. Champs AB. *Perfumery: Techniques in evolution. Part V. Perfum Flavorist.* 2004; 29: 42–62.
93. Propyl isobutyrate. Available: <http://www.thegoodscentscompany.com/data/rw1034031.html>
94. Manzanate (Givaudan). Available: <http://www.thegoodscentscompany.com/data/rw1006801.html>
95. Cook SM, Jönsson M, Skellern MP, Murray DA, Anderson P, Powell W. Responses of *Phradis* parasitoids to volatiles of lavender, *Lavendula angustifolia*—a possible repellent for their host, *Meligethes aeneus*. *BioControl.* 2007; 52: 591–598. <https://doi.org/10.1007/s10526-006-9057-x>
96. Kubo R, Ono M. Comparative analysis of volatile components from labial glands of male Japanese bumblebees (*Bombus* spp.). *Entomol Sci.* 2010; 13: 167–173. <https://doi.org/10.1111/j.1479-8298.2010.00375.x>
97. Bertsch A, Schweer H, Titze A, Tanaka H. Male labial gland secretions and mitochondrial DNA markers support species status of *Bombus cryptarum* and *B. magnus* (Hymenoptera, Apidae). *Insectes Soc.* 2005; 52: 45–54. <https://doi.org/10.1007/s00040-004-0761-1>
98. Valterová I, Urbanová K, Hovorka O, Kindl J. Composition of the labial gland secretion of the bumblebee males *Bombus pomorum*. *Zeitschrift für Naturforschung—Sect C J Biosci.* 2001; 56: 430–436. <https://doi.org/10.1515/znc-2001-5-618> PMID: 11421461
99. Coppée A, Terzo M, Valterova I, Rasmont P. Intraspecific variation of the cephalic labial gland secretions in *Bombus terrestris* (L.) (Hymenoptera: Apidae). *Chem Biodivers.* 2008; 5: 2654–2661. <https://doi.org/10.1002/cbdv.200890219> PMID: 19089823
100. Leonhardt SD, Blüthgen N, Schmitt T. Smelling like resin: Terpenoids account for species-specific cuticular profiles in Southeast-Asian stingless bees. *Insectes Soc.* 2009; 56: 157–170. <https://doi.org/10.1007/s00040-009-0007-3>
101. Yew JY, Chung H. Insect pheromones: An overview of function, form, and discovery. *Prog Lipid Res.* 2015; 59: 88–105. <https://doi.org/10.1016/j.plipres.2015.06.001> PMID: 26080085
102. Shelly T, McInnis D. Sterile Insect Technique and control of tephritid fruit flies: Do species with complex courtship require higher overflooding ratios? *Ann Entomol Soc Am.* 2016; 109: 1–11. <https://doi.org/10.1093/aesa/sav101>

103. Gilchrist AS, Dominiak B, Gillespie PS, Sved JA. Variation in population structure across the ecological range of the Queensland fruit fly, *Bactrocera tryoni*. Aust J Zool. 2006; 54: 87–95. <https://doi.org/10.1071/ZO05020>
104. Sved JA, Yu H, Dominiak B, Gilchrist AS. Inferring modes of colonization for pest species using heterozygosity comparisons and a shared-allele test. Genetics. 2003; 163: 823–831. <https://doi.org/10.1093/genetics/163.2.823> PMID: 12618417
105. Juárez ML, Devescovi F, Břízová R, Bachmann G, Segura DF, Kalinová B, et al. Evaluating mating compatibility within fruit fly cryptic species complexes and the potential role of sex pheromones in pre-mating isolation. Zookeys. 2015; 2015: 125–155. <https://doi.org/10.3897/zookeys.540.6133> PMID: 26798257
106. Canale A, Benelli G, Germinara GS, Fusini G, Romano D, Rapalini F, et al. Behavioural and electrophysiological responses to overlooked female pheromone components in the olive fruit fly, *Bactrocera oleae* (Diptera: Tephritidae). Chemoecology. 2015; 25: 147–157. <https://doi.org/10.1007/s00049-014-0183-0>
107. van der Pers JNC, Haniotakis GE, King BM. Electroantennogram responses from olfactory receptors in *Dacus oleae*. Entomol Hell. 2017; 2: 47–53. <https://doi.org/10.12681/eh.13901>
108. Mazomenos BE, Haniotakis GE. Male olive fruit fly attraction to synthetic sex pheromone components in laboratory and field tests. J Chem Ecol. 1985; 11: 397–405. <https://doi.org/10.1007/BF01411425> PMID: 24309970
109. Raptopoulos D, Haniotakis G, Koutsafikis A, Kelly D, Mavraganis V. Biological activity of chemicals identified from extracts and volatiles of male *Rhagoletis cerasi*. J Chem Ecol. 1995; 21: 1287–1297. <https://doi.org/10.1007/BF02027562> PMID: 24234627
110. Gilchrist AS, Shearman DCA, Frommer M, Raphael KA, Deshpande NP, Wilkins MR, et al. The draft genome of the pest tephritid fruit fly *Bactrocera tryoni*: Resources for the genomic analysis of hybridising species. BMC Genomics. 2014; 15: 1153. <https://doi.org/10.1186/1471-2164-15-1153> PMID: 25527032
111. Ahmed KA, Yeap HL, Pandey G, Lee SF, Taylor PW, Oakeshott JG. Population differences and domestication effects on mating and remating frequencies in Queensland fruit fly. Sci Rep. 2022; 12: 1–14. <https://doi.org/10.1038/s41598-021-04198-4> PMID: 34997097
112. Moadeli T, Taylor PW, Ponton F. High productivity gel diets for rearing of Queensland fruit fly, *Bactrocera tryoni*. J Pest Sci. 2017; 90: 507–520. <https://doi.org/10.1007/s10340-016-0813-0>
113. Meats A, Holmes HM, Kelly GL. Laboratory adaptation of *Bactrocera tryoni* (Diptera: Tephritidae) decreases mating age and increases protein consumption and number of eggs produced per milligram of protein. Bull Entomol Res. 2004; 94: 517–524. <https://doi.org/10.1079/ber2004332> PMID: 15541191
114. Ottensmann M, Stoffel MA, Nichols HJ, Hoffman JI. GCalinR: An R package for aligning gas-chromatography data for ecological and evolutionary studies. PLoS One. 2018; 13: 1–20. <https://doi.org/10.1371/journal.pone.0198311> PMID: 29879149
115. R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at <https://www.R-project.org/>.
116. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the Tidyverse. J Open Source Softw. 2019; 4: 1686. <https://doi.org/10.21105/joss.01686>
117. Hoffmann A, Parsons P. The analysis of quantitative variation in natural populations with isofemale strains. Genet Sel Evol. 1988; 20: 87–98. <https://doi.org/10.1186/1297-9686-20-1-87> PMID: 22879309
118. David JR, Gibert P, Legout H, Pétavy G, Capy P, Moreteau B. Isofemale lines in *Drosophila*: An empirical approach to quantitative trait analysis in natural populations. Heredity. 2005; 94: 3–12. <https://doi.org/10.1038/sj.hdy.6800562> PMID: 15329665
119. Lenth RV, Love J, Hervé M. Package ‘emmeans: Estimated marginal means, aka least-squares means’. Available: cran.r-project.org/web/packages/emmeans/emmeans.pdf. 2017.
120. Peterson RA. Finding optimal normalizing transformations via bestNormalize. R J. 2021; 13: 310–329. <https://doi.org/10.32614/rj-2021-041>
121. Peterson RA, Cavanaugh JE. Ordered quantile normalization: a semiparametric transformation built for the cross-validation era. J Appl Stat. 2020; 47: 2312–2327. <https://doi.org/10.1080/02664763.2019.1630372> PMID: 35707424
122. Chambers JM, Hastie TJ. Statistical models in S. Chambers JM, Hastie T, editors. Pacific Grove, CA, USA: Wadsworth & Brooks/Cole Advanced Books & Software; 1992.
123. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, et al. vegan: Community ecology Package. 2020. Available: <http://mirror.bjtu.edu.cn/cran/web/packages/vegan/vegan.pdf>

124. Venables WN, Ripley BD. Modern applied statistics with S. 4th ed., New York, USA: Springer-Verlag; 2002. https://doi.org/10.1007/978-1-4757-2719-7_14
125. Park SJ, Pandey G, Castro-Vargas C, Oakeshott JG, Taylor PW, Mendez V. Cuticular chemistry of the Queensland fruit fly *Bactrocera tryoni* (Froggatt). *Molecules*. 2020; 25: 1–18. <https://doi.org/10.3390/molecules25184185> PMID: 32932681