



HAL
open science

Sequence-based genome-wide association studies reveal the polygenic architecture of *Varroa destructor* resistance in Western honey bees *Apis mellifera*

Sonia Eynard, Fanny Mondet, Benjamin Basso, Olivier Bouchez, Tabatha Bulach, Yves Le Conte, Benjamin Dainat, Axel Decourtye, Lucie Genestout, Matthieu Guichard, et al.

► To cite this version:

Sonia Eynard, Fanny Mondet, Benjamin Basso, Olivier Bouchez, Tabatha Bulach, et al.. Sequence-based genome-wide association studies reveal the polygenic architecture of *Varroa destructor* resistance in Western honey bees *Apis mellifera*. 2024. hal-04493727

HAL Id: hal-04493727

<https://hal.inrae.fr/hal-04493727v1>

Preprint submitted on 21 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sequence-based genome-wide association studies reveal the polygenic architecture of *Varroa destructor* resistance in Western honey bees *Apis mellifera*

Sonia E. Eynard^{†1,2,11}, Fanny Mondet^{†3,11}, Benjamin Basso^{3,5,11}, Olivier Bouchez⁶, Yves Le Conte^{3,11}, Benjamin Dainat⁸, Axel Decourtye^{5,11}, Lucie Genestout², Matthieu Guichard^{8,9}, François Guillaume¹⁰, Emmanuelle Labarthe^{1,11}, Barbara Locke⁷, Rachid Mahla², Joachim de Miranda⁷, Markus Neuditschko⁹, Florence Phocas^{4,11}, Kamila Tabet^{1,11}, Alain Vignal^{§1,11}, and Bertrand Servin^{§1,11}

¹ GenPhySE, Université de Toulouse, INRAE, INP, ENVT, Castanet-Tolosan, 31320, France

² LABOGENA DNA, Palaiseau, 91120, France

³ INRAE, UR 406 Abeilles et Environnement, Avignon, 84914, France

⁴ Université Paris-Saclay, INRAE, AgroParisTech, GABI, Jouy-en-Josas, 78350, France

⁵ ITSAP, Avignon, 84914, France

⁶ GeT-PlaGe, Genotoul INRAE, Castanet-Tolosan, 31320, France

⁷ Department of Ecology, Swedish University of Agricultural Sciences, 750-07, Uppsala, Sweden

⁸ Agroscope Swiss Bee Research Centre, Bern 3003, Switzerland

⁹ Animal GenoPhenomics, Agroscope, Posieux 1725, Switzerland

¹⁰ SYNETICS, Noyal-sur-Vilaine 35538, France

¹¹ UMT PrADE, Avignon, 84914, France

Corresponding authors: sonia.eynard@inrae.fr, bertrand.servin@inrae.fr

†These authors contributed equally to the work

§These authors jointly supervised the work

Running title: Genetics of varroa resistance in honey bees

Abstract

Honey bees, *Apis mellifera*, have experienced the full impacts of globalisation, including the recent invasion by the parasitic mite *Varroa destructor* which has become one of the main causes of colony losses worldwide. Despite its lethal effects, some colonies have developed defence strategies conferring colony resistance and, assuming non-null heritability, selective breeding of naturally resistant bees could be a sustainable way to fight infestations. Here we report on the largest genome-wide association study performed on honey bees to understand the genetic basis of multiple phenotypes linked to varroa resistance. This study was performed on whole genome sequencing of more than 1,500 colonies belonging to different ancestries and combined in a meta-analysis. Results show that varroa resistance is polygenic. A total of 60 genetic markers were identified as having a significant impact in at least one of the tested populations pinpointing several regions of the honey bee genome. Our results also support strategies for genomic selection in honey bee breeding.

Keywords: *Apis mellifera*; *Varroa destructor*; varroa resistance; GWAS; mite non reproduction; recapping; pool sequencing; host-parasite interaction

1 Introduction

2 The honey bee, *Apis mellifera*, is a crucial contributor to sustainable food production [1].
3 However, for the past two decades beekeepers have been experiencing dramatic colony losses
4 [2, 3]. Such losses are not sustainable for the beekeepers and the agroecosystems relying on
5 services provided by honey bees. Extensive research has shown that honey bees are threatened
6 by multiple factors: both abiotic factors with the loss of natural resources and the impact of
7 pesticides due to agriculture intensification, and biotic factors with the infection by a diversity
8 of pests and parasites that impair their survival [4]. Among biotic factors, the ectoparasite
9 *Varroa destructor* is currently considered as the main threat to honey bee health and beekeeping
10 worldwide [5]. In most regions of the world, colony losses have dramatically increased since
11 its introduction in *Apis mellifera* populations in the early 80s [5, 6]. Originating from Asia,
12 where a stable host-parasite relationship exists with its historical host *Apis cerana*, varroa now
13 infests most *Apis mellifera* colonies worldwide. Varroa infests multiple compartments of the
14 honey bee colony: it reproduces in the brood, feeds on adult honey bee haemolymph and fat
15 body and favours virus infections [5, 7, 8]. Combined, these effects on individual bees lead to
16 colony collapse within a few months if no actions are applied to control mite infestations [9].
17 To date, managing varroa infestation presents many constraints that offer beekeepers only a few
18 unsustainable solutions to fight the deadly mite [7, 10].

19 However, since the beginning of the 1990's, colonies naturally surviving varroa infestation
20 without treatment have been observed in several regions of the world and raised hope for bee-
21 keepers to overcome the problems caused by varroa infestation. In these surviving colonies,
22 undergoing beekeeping activities, honey bees often display behavioural and physiological de-
23 fences against the varroa parasite. These collective responses are expected to contribute to the
24 limitation of parasite population growth and provide colonies with social immunity [11], a sus-
25 tainable long-term adaptation to counter the immense damage caused by varroa. However, thus
26 far such long term adaptation is not always reached, even in colonies expressing resistance traits
27 (e.g. colonies in Gotland, Sweden [12, 13]). The defence repertoire against varroa includes
28 hygiene behaviour targeted towards varroa parasitised brood cells, in the form of varroa sensi-
29 tive hygiene (VSH [14]) or recapping behaviour [15, 16, 17], and mechanisms suppressing mite
30 reproduction (SMR [18]). Expression of these different traits lead to an increase in mite non
31 reproduction within brood cells (MNR [15, 18], also called Decreased Mite Reproduction DMR
32 [19]). This triggers lower mite population growth and infestation [10]. Such mechanisms have
33 gained a major interest within the beekeeping sector and interest is growing to help decipher
34 the genetic mechanisms underlying varroa resistance in the honey bee.

35 As for common livestock species, the most obvious way to induce varroa resistance into
36 the honey bee population appears to be through selection and spreading of the most resistant
37 lines. Since the late 90's efforts have been put into the selective breeding of such resistant
38 honey bee lines [20, 21]. However, selecting for complex traits in honey bees has been hindered

39 by features that strongly differentiate honey bees from other typical livestock species. Due to
40 the social nature of honey bees, many phenotypic traits of interest to beekeeping (including
41 varroa resistance traits) are expressed at the group (i.e. colony) level and thus can not be
42 addressed by classical GWAS approaches where phenotypes are determined at the individual
43 level (e.g. individual bees, in this case). In addition, many of these traits, are linked to
44 group behaviour and are thus difficult to phenotype and can display rather low repeatability and
45 heritability [22, 23, 24, 25, 26]. At the breeding level, managing honey bee reproduction is
46 difficult due to polyandry [27] (queen bees mate with multiple males) and to the fact that sex
47 determination is governed by an haplodiploid mechanism at a single locus [28, 29] at which
48 diploid homozygosity is effectively lethal [30], limiting drastically the possibility of inbreeding.
49 Finally, honey bee populations bred for beekeeping encompass an important genetic diversity,
50 with strong regional clustering, which impedes the identification of genetic markers that are
51 valid outside the population where they are identified. Developing molecular tools to assess the
52 genetic make-up of honey bee colonies could help in honey bee breeding by coping with some
53 of these issues. To list a few possibilities, genomic-enabled prediction of resistance traits could
54 reduce the amount of complex phenotyping required by identifying promising colonies early in
55 life and genetic assessment of colonies could help identifying the genetic group of colonies [31]
56 and allow to design crosses limiting inbreeding in selection programs.

57 In addition to their use in selection programs, molecular tools can help to identify pathways
58 involved in a specific mechanism by performing genome-wide association studies, assessing
59 the statistical effects of polymorphisms on the variability of complex phenotypes. This offers
60 opportunities to develop new approaches that take into account the specific genetic determinism
61 of honey bees and to open avenues for genomic selection on traits such as varroa resistance.
62 However, tools to perform genetic association studies are so far not tailored to encompass such
63 genetic characteristics, which limits the power of genomic studies performed on honey bees
64 and their transferability into breeding tools. Some markers of interest have been identified (see
65 [16] for a review) but most genomic studies performed on honey bee traits so far were built on
66 a limited number of samples (10 to 200 hundred individuals or colonies [32, 33, 34, 35, 36, 37,
67 38, 39], restricting the power of the analyses to come out as a general breeding tool. To date,
68 the use of such markers has been limited, mostly due to a lack of easily accessible genotyping
69 tools, leaving beekeepers with very limited access to varroa resistant stock.

70 In this study, we took advantage of the unique situation represented by France. Geographical
71 crossroad, the French territory has the advantage to present a large variety of landscapes,
72 environments and ecosystems, where honey bee populations with different genetic background
73 coexist, together with a large variety of hybrid colonies [40]. To overcome the limitations in
74 sample size, we performed one of the largest genomic study applied to honey bees, with the
75 phenotyping and complete genome sequencing of more than 1,500 colonies. Using uniquely
76 tailored genetic and genomic tools, such as queen genotype reconstruction from pool sequence
77 data [31] and GWAS and meta-GWAS analysis [41, 42], we investigated the genetic bases of

78 three major traits linked to varroa resistance: overall varroa infestation of the colony, mite non
79 reproduction (MNR) and recapping of varroa infested brood cells. We identified multiple genetic
80 markers of interest, spread out on the whole genome and heterogeneous across the different
81 populations. This large-scale effort provides a new understanding of the genetic mechanism
82 underlying honey bee resistance to its main parasite, the *Varroa destructor* mite.

83 Results

84 Genetic and phenotypic diversity of honey bee colonies

85 Using allele frequencies estimated from pool sequence data for the 1513 sampled honey bee
86 colonies, we identified three groups of different genetic ancestries: 703 colonies were identified
87 as having more than 80% *Apis mellifera ligustica & carnica* genetic background, 407 having
88 more than 80% *Apis mellifera mellifera* genetic background and 382 as hybrids (Fig. 1). An
89 additional 21 colonies were found to be of pure *Apis mellifera caucasia* ancestry, but due to this
90 reduced sample size of this category they were not analysed further.

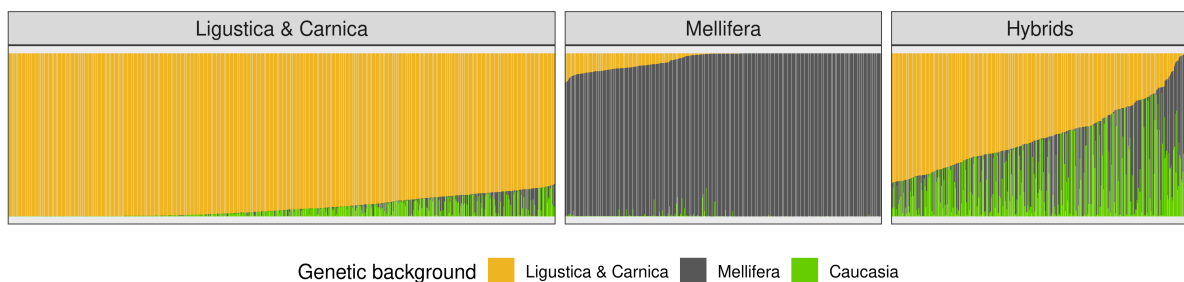


Fig. 1: **Genetic background for each colony, per group.** Proportion of the three main genetic background for each of the group analysed in our study *Apis mellifera ligustica & carnica*, *Apis mellifera mellifera* and hybrids.

91 For the purpose of this study we collected six phenotypes on these colonies. Out of the
92 six phenotypes initially available four were associated with varroa infestation (on the adult
93 bees: phoretic infestation rate v_pho and varroa mitochondrial sequence reads v_mito , inside
94 the brood: brood infestation rate v_brood and overall in the colony: varroa load v_load). They
95 were highly positively correlated with each other and drove the first dimension of the principal
96 component analysis (PCA) across all colonies (Fig. 2), and also within each group (supplemen-
97 tary figure S1), with about 60% of the variance explained in each case. Therefore, coordinates
98 of colonies on the first component of the PCA were used as the varroa infestation phenotype in
99 the genome wide association study (GWAS) ($varroa_inf$). The two remaining phenotypes were
100 expected to be linked to resistance to varroa infestation, either through mechanisms repressing
101 varroa reproduction, and thus varroa population growth, within the colony (so called mite non
102 reproduction) or through the cleaning of brood cells infested by varroa. Mite non reproduction

103 (*MNR*) was slightly positively correlated with the recapping of infested brood cells, as expected
 104 under the assumption that *recap* contributes to overall *MNR*. Both *MNR* and *recap* were slightly
 105 negatively correlated with varroa infestation. This observation is consistent with mechanisms
 106 being linked to a reduction in varroa infestation within the colony. *MNR* and *recap* contributed
 107 both to the second dimension of the PCA, explaining about 20% of the variance. They are
 108 separated on the third axis of the PCA, which explains about 12% of the variance (Fig. 2).
 109 GWAS were performed on each of these three phenotypes (*varroa_inf*, *MNR* and *recap*).

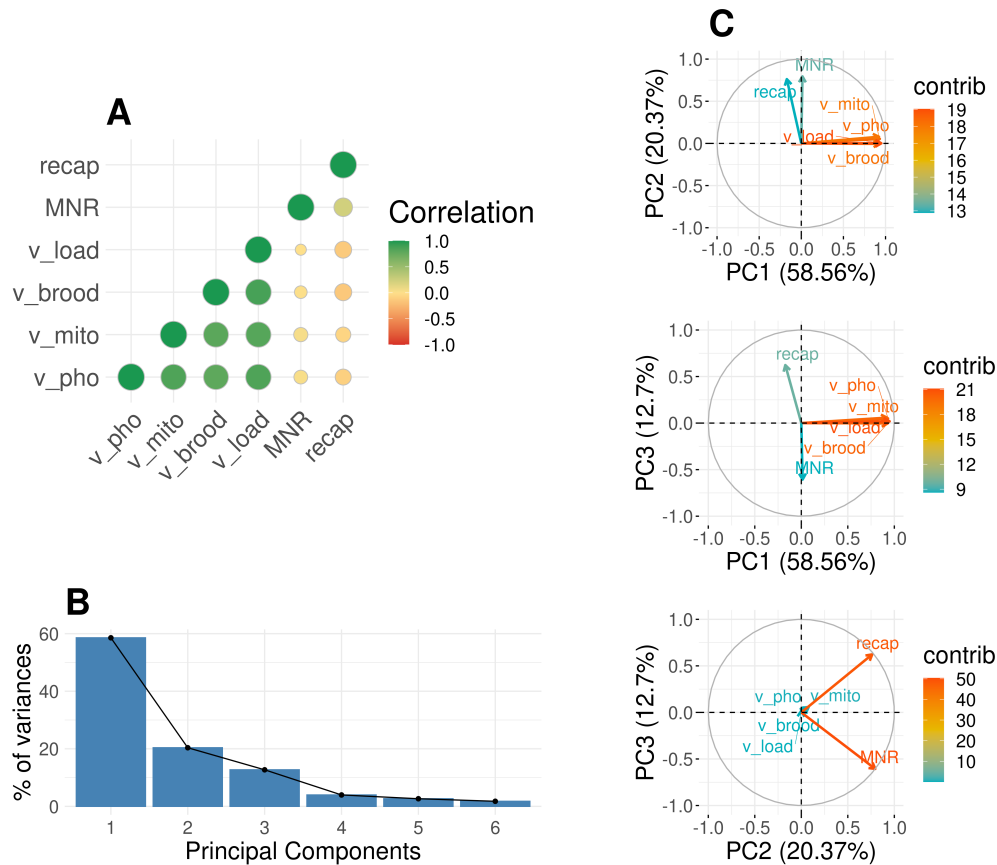


Fig. 2: **Correlation and principal component analysis.** Description of the correlation between phenotypes and principal component analysis. (A) gives the correlation between our original phenotypes. (B) summarises the percentage of variance explained by each of the principal component analysis from axis 1 to 5. (C) shows our phenotype on principal component analysis for axis 1, 2 and 3, the colour gives the contribution of each variable to the axis, the closer to red, the higher. The correlation and PCA estimates are based on an analysis of the whole dataset, including all colonies from the three groups.

110 Meta-analysis of varroa resistance

111 We aim to identify markers significantly associated with our traits of interest across the three
 112 main genetic backgrounds found in Europe. Within population we used standard linear mixed
 113 model equations, implemented in GEMMA [43], significant markers were identified relative to
 114 their local false discovery and false sign rates estimates from adaptive shrinkage [44]. Population
 115 level analyses were combined into a meta-analysis, using two Bayesian methods: a co-ancestry

116 specific program MANTRA [41] and the more generalist Mash [42], to increase power to detect
117 significant markers across the three genetic backgrounds.

118 **Associated variants**

119 The individual GWAS for each genetic background group and each phenotype allowed to iden-
120 tify 8 regions (n SNPs=9).

121 In details:

122 (i) For *varroa_inf*, we found one variant (1:10080627:C>T) in *A. m. ligustica & carnica*, with a
123 positive alternative allele effect, and one (4:11665460:G>A) in the hybrid group, with a negative
124 alternative allele effect (a detailed list of the single nucleotide polymorphisms, SNPs, can be
125 found in supplementary table SR1).

126 (ii) For *recap* we identified four variants significant in *A. m. mellifera*, two of which the alter-
127 native alleles had negative effects on the trait (2:2729874:T>C and 3:1059430:T>C) and two
128 having positive effects (4:4327611:G>A, 15:8485332:G>A). In *A. m. ligustica & carnica* a
129 region, with two SNPs, both of which alternative alleles had positive effects (2:12025610:A>G,
130 2:12025647:A>G) were found. Finally in the hybrid group, one variant for which the alternative
131 allele had a negative effect on the trait (13:9483955:C>T) was found.

132 Out of the 9 SNPs found significant in one trait for one group, eight fell inside genes of the
133 honey bee annotation [45] and one fell 7kb upstream of its closest gene (for more details see
134 supplementary table SR1).

135

136 The meta-analysis permitted to identify 51 regions, containing 56 significant SNPs across
137 the three traits of interest. From these 14 regions (n SNPs=14) were significant for *varroa_inf*,
138 14 regions (n SNPs=15) for *MNR* and 23 regions (n SNPs=27) for *recap*. They distributed
139 across the whole genome, on almost every chromosome (a detailed list of the SNPs can be
140 found in supplementary table SR2).

141 In details:

142 (i) For the *varroa_inf* trait: the 14 SNPs having a significant effect on the trait were distributed
143 on the chromosomes 7 (n=4), 1, 5 and 8 (n=2) and 4, 6, 11 and 12 (n=1). Thirteen of these
144 SNPs were considered significant in the MANTRA meta-analysis with log₁₀(BF) ranging from
145 7.59 to 5.09 and one was significant in the mash analysis with log₁₀(BF)=1.16. For six of
146 these variants, the alternative allele had a positive effect on the trait, in at least one of the
147 group (5:9190579:A>G, 7:5772089:A>T, 7:6738985:T>A, 7:11806658:G>A, 8:9799408:C>T
148 and 12:10734707:A>G), for seven alternative allele had a negative effect on the trait, in at
149 least one of the group (1:20960056:C>T, 1:25184394:C>T, 4:11665460:G>A, 5:75369:C>T,
150 6:10450971:C>T, 7:5762037:T>C and 8:2468335:C>T) and for one the effect depended on the
151 group (11:9369229:T>C). Eight SNPs fell inside genes found in the honey bee annotation, six
152 of them are intronic variation, one leads to a change in 5' UTR and one in 3' UTR. Three fell

153 downstream, within 54kb of their closest genes. The remaining three fell up upstream, within
154 11kb of their closest genes.

155 (ii) For *MNR*, the 15 SNPs having a significant effect on the trait were distributed on the chromo-
156 somes 1 (n=4), 8, 10 and 12 (n=2) and 2, 3, 5, 11 and 15 (n=1). All these SNPs were considered
157 significant in the MANTRA meta-analysis with log₁₀(BF) ranging from 5.12 to 6.44. In three of
158 these variants, the alternative allele had a positive effect on the trait, in at least one of the group
159 (2:4437645:G>A, 12:10153855:A>G and 15:4853529:C>T), for seven it had a negative effect
160 on the trait, in at least one of the group (1:16327085:C>T, 1:21374478:G>A, 1:24201224:C>T,
161 3:6206342:C>T, 8:1150346:C>T, 11:9527267:G>A and 12:136634:G>C) and for five the effect
162 depended on the group (1:2891204:G>A, 5:2008472:A>C, 8:9557205:C>T, 10:5359169:T>A
163 and 10:5359173:C>T). Out of the 15 significant SNPs, 12 fell inside genes, 10 are intronic
164 variation, one into 3' UTR and one causes a missense variation. The three remaining SNPs fell
165 within 17kb downstream of their closest genes.

166 (iii) Finally, for the *recap* trait SNPs were distributed on the chromosomes 2 and 7 (n=4), 1, 5 and
167 14 (n=3), 4 and 15 (n=2) and 3, 8, 9, 10, 11 and 16 (n=1). From these SNPs 26 were considered
168 significant based on their log₁₀(BF) values for the MANTRA meta-analysis with log₁₀(BF)
169 ranging from 5.01 to 7.61 and 6 SNPs add log₁₀(BF) values from mash higher than one, their
170 log₁₀(BF) ranged between 1.41 and 2.53. Differently from the other trait five SNPs were found
171 significant in both MANTRA and mash meta-analysis for *recap*. Their log₁₀(BF_MANTRA)
172 ranged between 5.03 and 7.61 and their log₁₀(BF_mash) between 1.41 and 2.53. They were lo-
173 cated on the chromosomes 2 (n=2), 3, 11 and 15. For eight of these variants, the alternative allele
174 had a positive effect on the trait, in at least one of the group (1:7448807:A>T, 1:7448811:T>C,
175 4:7321246:T>A, 4:7321247:G>T, 14:6686131:A>G, 14:8481541:A>G, 15:2081876:A>G and
176 15:8485332:G>A), for 18 the alternative allele had a negative effect on the trait, in at least
177 one of the group (1:15280956:G>A, 2:2729874:T>C, 2:8350714:G>A, 2:12025610:A>G,
178 2:16060868:G>A, 3:1059430:T>C, 5:6736534:T>C, 5:6761414:T>A, 5:8737386:G>A, 7:7028040:G>A,
179 7:7051965:A>G, 7:7078376:C>T, 8:1551638:C>T, 9:11564671:A>C, 10:2026877:C>G, 11:14369154:G>C,
180 14:3782741:G>A and 16:1812909:C>T) and for one the effect depended on the group (7:8466948:A>G).
181 For this trait, 18 variants fell inside genes with 16 intronic variation, one in a region coding
182 for long non coding RNA and one in 5' UTR, being also a missense variant. Three SNPs fell
183 between 52 and 102kb downstream of their closest genes. Six SNPs fell between 5 and 129kb
184 upstream of their closest genes (details are available in supplementary tables SR2 and SR3).

185

186 We identified chromosome regions smaller than 1Mb, sharing significant SNPs between
187 multiple traits. Between *varroa_inf* and *MNR* we identified two regions on chromosome 1 (20.9-
188 21.4Mb and 24.2-25.2Mb), a region on chromosome 8 (9.5-9.8Mb), a region on chromosome 11
189 (9.3-9.6Mb) and a region on chromosome 12 (10.1-10.8Mb) having significant SNPs for both
190 traits. Between *varroa_inf* and *recap* we identified a region on chromosome 5 (8.7-9.2Mb),
191 a region on chromosome 7 (6.7-7.1Mb) and a region on chromosome 8 (1.5-2.5Mb). Finally

192 between *MNR* and *recap* we identified a region on chromosome 8 (1.1-1.6Mb) (Fig. 3, for the
 193 details of each region see supplementary table SR4).

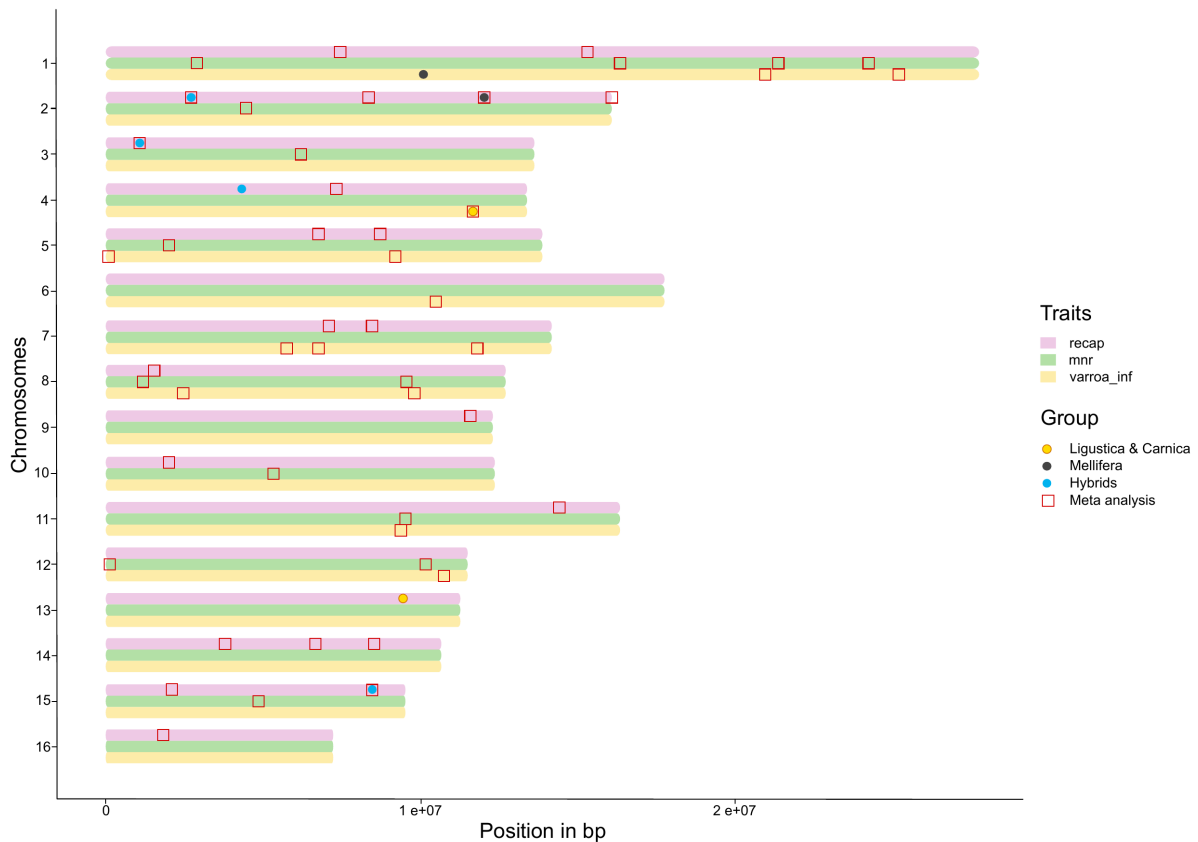


Fig. 3: **Overlap between traits across significant regions.** Position of each of the significant SNPs on their chromosomes. A coloured dot means that this position has been identified in either *A. m. ligustica & carnica*, *A. m. mellifera* or hybrids. A red square means that it has been identified in the meta analysis. The colour bars represent the phenotypes of interest *varroa_inf*, *MNR* and *recap*. This figure allows to see overlapping window containing significant marker across the phenotypes.

194 Heterogeneity of effects

195 We analysed three genetic groups of colonies in this study, with two groups having relatively
 196 pure genetic backgrounds, corresponding to the two main lines of honey bee subspecies in
 197 Europe, with the third group consisting of varying degrees of hybridization between these
 198 two main groups. The F_{st} , measure of population differentiation, between groups were 0.26,
 199 between *A. m. mellifera* and *A. m. ligustica & carnica*, 0.20 between *A. m. mellifera* and
 200 hybrids and 0.08 between *A. m. ligustica & carnica* and hybrids.

201 As detailed earlier, from the meta analysis we observed 17 variants ($n_{varroa_inf}=6$, $n_{MNR}=3$ and $n_{recap}=8$) for which alternative allele had significant positive effects for at least
 202 one of groups, 32 ($n_{varroa_inf}=7$, $n_{MNR}=7$ and $n_{recap}=18$) having significant negative
 203 effects for at least one of the group, and seven ($n_{varroa_inf}=1$, $n_{MNR}=5$ and $n_{recap}=1$)
 204 having divergent effect depending on the group (supplementary table SR3). Correlations
 205

206 between the effects of significant SNPs between different groups were null therefore showing
207 great heterogeneity in SNPs having an effect, their impact on the trait of interest and their
208 magnitude across the different genetic groups, especially for the *MNR* trait.

209 About 80% of the SNPs identified significant fell into intronic regions, there was no differ-
210 ences between the annotation of the significant SNPs on the genome and the annotation of all
211 tested SNPs. Five SNPs were found significant in both individual GWAS and meta-analysis,
212 one for *varroa_inf* identified significant for the hybrid group, and four for *recap*, three for *A. m.*
213 *mellifera* and one for *A. m. ligustica & carnica*. They were located on chromosome 2 (n=2), 3
214 (n=1), 4 (n=1) and 15 (n=1). They all located inside the genes, LOC102655235, LOC410853
215 (chromosome 2), LOC409402 (chromosome 3), LOC408787 (chromosome 4) and LOC726948
216 (chromosome 15).

217 **Example of associations**

218 As an example of a region associated with two traits, we focus on the region between 9.5 and
219 9.8Mb on the chromosome 8 where two SNPs appear significant in the meta analysis, one for
220 *MNR* at 9,557,205 bp and one for *varroa_inf* at 9,799,408 bp. In this region we identified
221 a couple of SNPs in close vicinity to the gene *Ecr*, ecdysone receptor (Fig. 4). In addition,
222 we also identified a SNP, located at 9,696,277bp so within *Ecr*, in high linkage disequilibrium
223 (LD) with the significant SNP 8:2468335:C>T also found in the meta-analysis for *varroa_inf*
224 (supplementary table SR6).

225

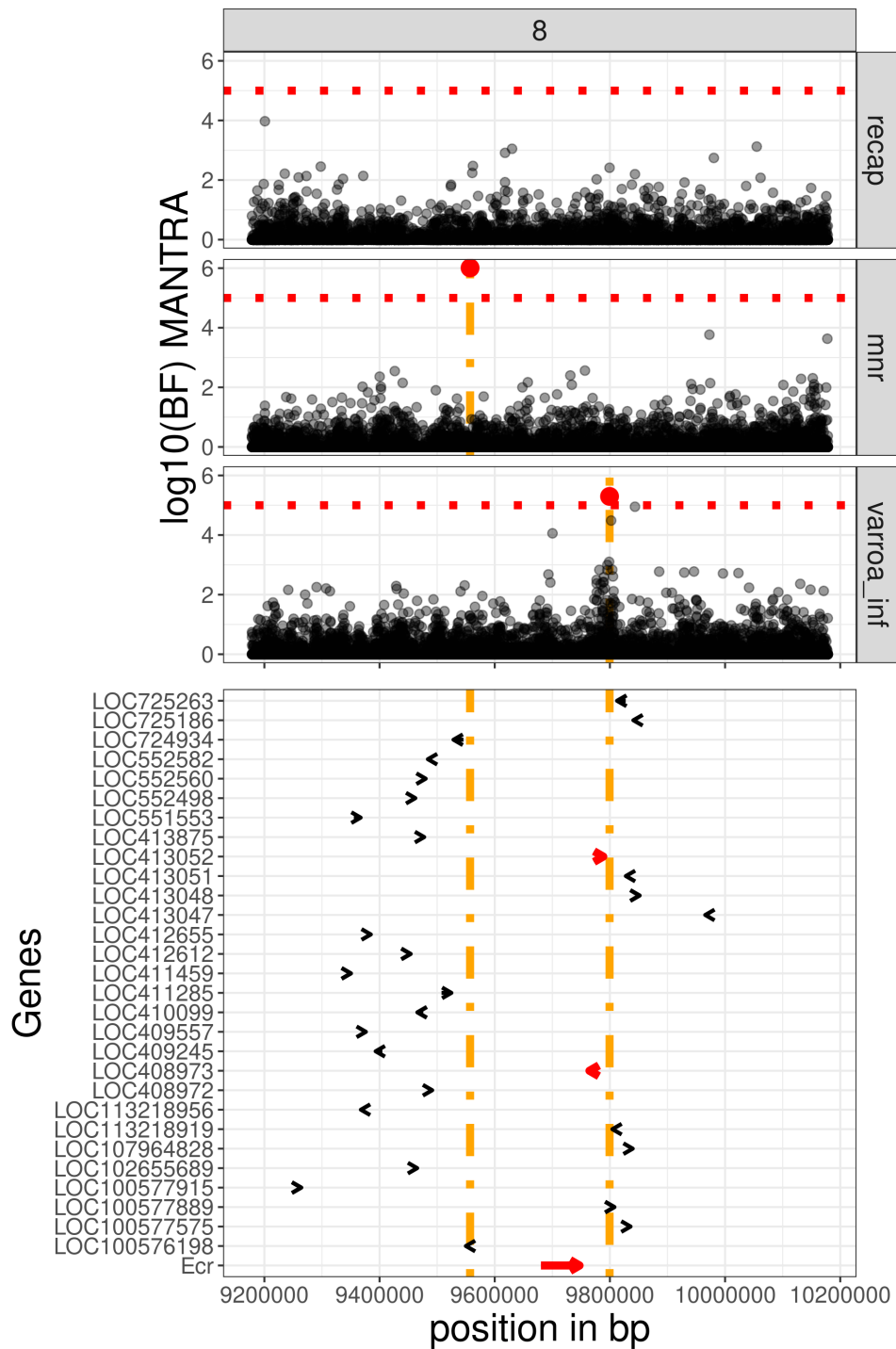


Fig. 4: **Region on chromosome 8.** From top to bottom we represented results from genome wide association studies for chromosome 8, region going from 9.2Mb to 10.2Mb, for *varroa_inf*, *MNR* and *recap*. We represented \log_{10} bayes factor for MANTRA, with the significant threshold as a dotted red line, under which we listed the genes identified within this region of chromosome 8. The orange lines represent the positions of the significant markers, the genes falling in the region between these markers are highlighted in red.

226 Next, we turn to chromosomes 15, where a significant SNP for *recap* was found at position
 227 2,081,876. Nearby we also observed two suggestive SNPs, close to the significant threshold set
 228 for this analysis, in positions 2,021,142 and 2,081,914 bp. The markers 15:2081876:A>G and

229 15:2081914:A>G were in full linkage disequilibrium in *A. m. ligustica & carnica* and in the
230 hybrid group and in high linkage disequilibrium ($r^2 > 0.8$) in *A. m. mellifera* and they fell within
231 the same haplotype block of 1.6kbp (as identified by Wragg et al. (2022) [40], supplementary
232 table SR7). This haplotype block did not seem to contain annotated genes. The marker
233 15:2021142:C>T fell in a short haplotype block (0.175kbp) containing the gene LOC413200
234 which has been identified as putative immune related gene by Ryabov et al. (2014) [46].
235 Interestingly they were located within less than 1Mbp downstream from a group of eight genes
236 coding for odorant binding proteins. For the first SNP (15:2021142:C>T), alternative allele had
237 a negative effect for all groups whereas the second (15:2081876:A>G), mostly significant in *A.*
238 *m. ligustica & carnica*, the alternative allele had a positive effect for all groups (Fig. 5). Marker
239 effects estimated with GEMMA lack precision, whereas the ash and mash methods apply a
240 strong shrinkage to the estimates. Going from individual analysis to meta-GWAS improved the
241 power to detect associations and improve our ability to estimate markers effects.

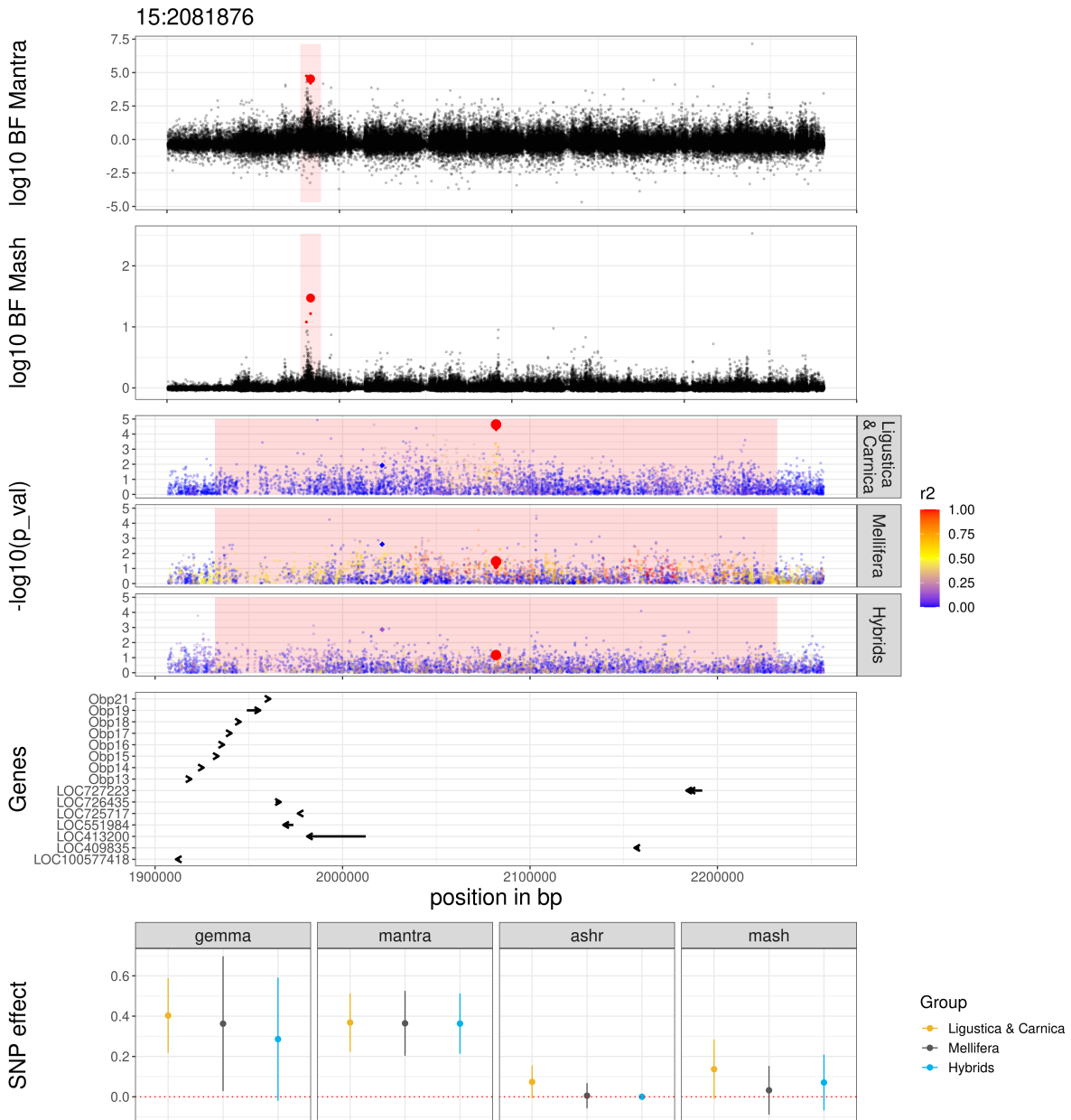


Fig. 5: **Region on chromosome 15.** We plotted manhattan plots for the marker positioned at 2,081,876 bp on chromosome 15 and the surrounding 0.3Mbp region. The marker 2,081,876 is represented by a large red dot when the close by markers identified as near significant thresholds are represented as smaller. From top to bottom we plotted $\log_{10}(\text{bays factor})$ for MANTRA, then for mash, then $-\log_{10}(\text{p-values})$, estimated using GEMMA, for the groups *A. m. ligustica & carnica*, *A. m. mellifera* and hybrids. Each marker is coloured based on its linkage disequilibrium r^2 with the marker of interest. The genes falling in the region of interest are plotted and then the effects of the significant marker extracted from genome wide association studies ran with GEMMA, ash, MANTRA and mash.

242 Polygenic architecture of varroa resistance

243 The program GEMMA [43], used to perform individual GWAS, provides estimates for the
 244 proportion of phenotypic variance explained by the SNPs (pve) and proportion of genetic
 245 variance explained by the sparse effects (pge) for each trait in each group (Tab. 1 and Fig. 6).

246 PVEs ranged between 0.20 and 0.82 (se=[0.09 ; 0.19]). PVEs, estimated using the Bayesian
 247 Sparse linear mixed model (BSLMM-GWA), were close to the lmm estimates (LMM-GWA),
 248 and ranged between 0.14 and 0.82 (se=[0.07; 0.15]). The 95% confidence and credible intervals
 249 for PVEs from both LMM-GWA and BSLMM-GWA appeared exclusively positive, except for
 250 *MNR* in the hybrids group. However, PGEs were much lower, ranging between 0.06 to 0.33
 251 (se=[0.08; 0.25]). Their 95% confidence/credible intervals often included zero. The only traits
 252 and groups having PGEs different from zero were *MNR* and *recap* for the group of colonies
 253 of *A. m. ligustica* type. The estimates for the GWAS on hybrids and *A. m. mellifera* always
 254 showed larger standard error, as expected due to the smaller sample sizes of these groups.
 255 Interestingly it appears that PVE estimate is slightly higher for *A. m. mellifera* for the *MNR*
 256 phenotype compared to the two other groups, whereas they seem similar between the three
 257 groups for the two other phenotypes, *varroa_inf* and *recap* (a complete summary can be found
 258 in supplementary table SR8).

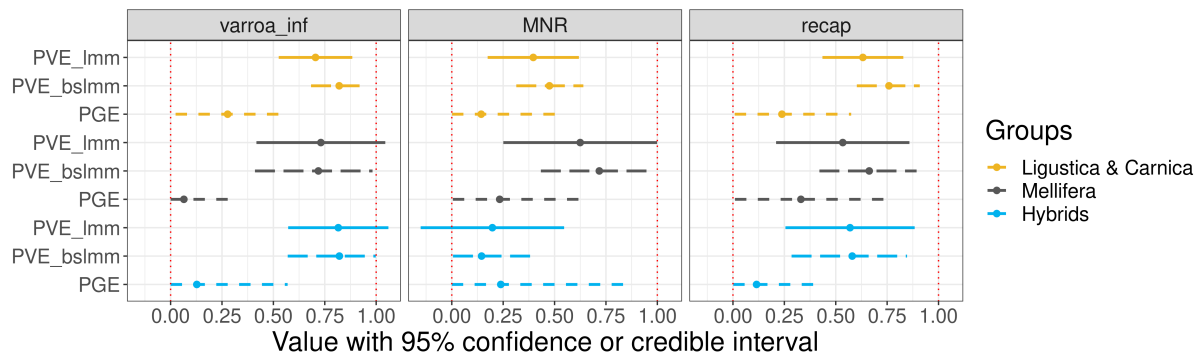


Fig. 6: **Genome wide association estimates.** Confidence and credible intervals for percentage of variance explained by linear mixed model (full line), Bayesian sparse linear mixed model (long dotted line) and percentage of genetic variance explained by Bayesian sparse linear mixed model (short dotted line) for the three group analysed, in yellow *Apis mellifera ligustica & carnica*, in dark grey for *Apis mellifera mellifera* and in blue for hybrid.

Tab. 1: Summary of PVE (Phenotypic Variance Explained) and PGE (Proportion of Genetic variance Explained).

	Varroa infestation	Resistance to varroa infestation		
	<i>Varroa_inf</i>	<i>MNR</i>		
		<i>Recap</i>		
PVE(se) from lmm	<i>Ligustica & Carnica</i>	0.705 (0.091)	0.396 (0.113)	0.632 (0.100)
	<i>Mellifera</i>	0.731 (0.160)	0.625 (0.191)	0.534 (0.165)
	<i>Hybrids</i>	0.816 (0.124)	0.197 (0.178)	0.570 (0.160)
PVE[95% ci] from bs1mm	<i>Ligustica & Carnica</i>	0.820 [0.683; 0.953]	0.475 [0.314; 0.640]	0.759 [0.602; 0.910]
	<i>Mellifera</i>	0.719 [0.410; 0.982]	0.718 [0.433; 0.981]	0.663 [0.420; 0.893]
	<i>Hybrids</i>	0.821 [0.570; 0.998]	0.145 [0.006; 0.393]	0.580 [0.285; 0.847]
PGE[95% ci] from bs1mm	<i>Ligustica & Carnica</i>	0.277 [0.024; 0.555]	0.143 [0.001; 0.503]	0.238 [0.008; 0.575]
	<i>Mellifera</i>	0.065 [0; 0.286]	0.233 [0.005; 0.666]	0.331 [0.010; 0.751]
	<i>Hybrids</i>	0.128 [0; 0.569]	0.238 [0; 0.880]	0.115 [0.001; 0.425]

259 Correlation between SNP effects, across the whole genome, estimated using GEMMA and
 260 **ash**, for individual GWAS analysis, on the one hand and MANTRA and mash, for meta-GWAS,
 261 on the other hand were positive. However, as expected, between the different groups there was no
 262 correlation when estimated from individual GWAS whereas there was some positive correlation
 263 when estimated from meta-GWAS. Across the different phenotypes, *recap* and *MNR* appear
 264 to be slightly positively genetically correlated, as one could expect knowing that recapping
 265 behaviour, performed by adult bees, is potentially a component of the *MNR* phenotype. The
 266 phenotypes *varroa_inf* and *MNR* were not genetically correlated. Finally, *varroa_inf* and *recap*
 267 were mostly negatively correlated for *A. m. mellifera* and *A. m. ligustica & carnica* whereas
 268 there was no to very little positive genetic correlation between these two phenotypes in the
 269 hybrids group (Tab. 2, the detailed correlations can be found in Supplementary table SR5).

Tab. 2: **Genetic correlations.** Range of the estimates from Pearson correlations on the different GWAS analyses methods used in our study.

		<i>Varroa_inf</i>	<i>MNR</i>	<i>Recap</i>
	<i>Varroa_inf</i>	1		
<i>Ligustica & Carnica</i>	<i>MNR</i>	[0.042; 0.049]	1	
<i>Mellifera</i>		[-0.023; 0.005]		
<i>Hybrids</i>		[0.050; 0.073]		
<i>Ligustica & Carnica</i>	<i>Recap</i>	[-0.304; -0.213]	[0.131; 0.176]	1
<i>Mellifera</i>		[-0.132; -0.058]	[0.088; 0.197]	
<i>Hybrids</i>		[0.021; 0.173]	[0.021; 0.119]	

270 Discussion

271 In this study we performed the largest genome wide association study on the resistance of honey
 272 bees to their current biggest biotic threat, the parasite *Varroa destructor*. We combined an
 273 extensive genotyping and phenotyping effort with meta-analyses methods to identify genetic
 274 markers and associated genes harbouring a significant effect on varroa resistance. This leveraged
 275 both multiple traits associated to varroa resistance and the complex genetic structure found in
 276 honey bee colonies. Our results show that we were able to pinpoint significant effects of
 277 some regions of the honey bee genome, within specific genetic types and across the whole
 278 meta-population, offering insights into the biological mechanisms involved in varroa resistance.
 279 These genomic regions only explain a small portion of the genetic variation which remains
 280 mostly polygenic. However we found the contribution of genetics to varroa resistance to be
 281 substantial, offering positive perspectives to a possible adaptation and selection of honey bee
 282 populations to this relatively recent threat.

283 Based on 1,500 sampled honey bee colonies, sequenced in pool and genotyped for around

284 3 million SNPs, this study benefits from the largest sample size (phenotyped and sequenced)
285 known thus far to perform genome wide association study in honey bees: it is the most global
286 association study for honey bee varroa resistance traits to date. The set of colonies studied
287 is highly representative of many honey bee populations worldwide, measured in terms of
288 the genetic backgrounds described in the diversity panel of Wragg et al. (2022) [40]. This
289 study stands out from the previous honey bee quantitative genetic studies that mostly focused on
290 specific genetic backgrounds, often in small experiments, not representative of the field situation
291 and for really specific phenotypes ([38] for a review). In addition to the experimental effort to
292 gather the raw data, this analysis benefited from dedicated statistical methods for reconstructing
293 the honey bee queen genotype from a pool sequencing experiment [31] and meta-analyses
294 approaches to increase statistical power to detect significant associations [41, 42].

295 **Phenotyping varroa resistance**

296 Resistance to varroa is a complex trait, involving many different aspects of the biology of honey
297 bee colonies. A recent extensive review of varroa resistance traits [26] reveals that there is no
298 clear evidence for significant correlations between the standard traits measured as proxy for
299 varroa resistance, such as varroa sensitive hygiene (VSH), grooming, mite non reproduction
300 (MNR), hygienic behaviour or uncapping-recapping as cleaning behaviour. Depending on the
301 population on which the trait has been measured, in particular their evolutionary history (*e.g.*
302 natural or artificial selection) the correlations between these traits range from 'apparent link' to
303 'no link'.

304 Here, we observe a slightly positive correlations between traits linked to varroa resistance,
305 *MNR* and *recap* overall or at the population level, and a small negative correlation between these
306 two traits and *varroa_inf* (Fig. 2 and in supplementary figure SM1). This finding fits with the
307 hypothesis that there is a panel of mechanisms allowing honey bee to resist varroa and that these
308 mechanisms do not seem to be completely shared across populations or genetic ancestries.

309 An important experimental aspect that limits varroa resistance studies, including this one,
310 is that most currently applied measures of varroa resistance are difficult to scale to large
311 samples. For example, they can imply tedious and potentially subjective scoring, induced or
312 artificial varroa infestation, multiple measures in time, estimation of ratios and applying heuristic
313 thresholds for minimum detection. However, direct estimates of varroa infestation are the most
314 simple traits to measure to summarise varroa resistance. Indeed, a low varroa infestation can be
315 explained by multiple phenomena either linked to the environment, beekeeping practices, varroa
316 biology or an action from the honey bee colony itself. In this study, in addition to the classical
317 measures of varroa infestation, we proposed to measure varroa infestation indirectly, using the
318 ratio of reads mapped to the varroa mitochondrial DNA over those mapped to the honey bee
319 genome (*varroa_mito*). We believe this new measure offers specific advantages for the study of
320 varroa resistance: (i) there is a high correlation between this estimate and the phoretic varroa

321 infestation (supplementary methods 1), a trait that is more complex to measure, (ii) both colony
322 sequence and varroa infestation information come from a unique biological sample, (iii) there
323 is no potential bias due to the collector, (iv) no specific technical skills are needed on the field
324 and (v) it is comparable across studies. Using it as a phenotype for varroa infestation would
325 facilitate the establishment of a large collection of standardised phenotypic records, necessary
326 to be able to build up information through time, follow phenotypic progress in a surveyed
327 populations, perform genetic meta-analyses and potentially breed honey bee populations for
328 varroa resistance.

329 **Insights into biological mechanisms underlying varroa resistance**

330 Varroa resistance mechanisms can be partitioned into two types of traits: first, traits related to
331 hygiene (including VSH, recapping and MNR, but also more broadly grooming behaviour) that
332 involve the accurate detection by workers of varroa infested cells and second, their subsequent
333 inspection/destruction. It has been shown that VSH bees target more specifically cells with
334 highly compromised brood, which is related to the level of infestation in the cells [47, 14]. As
335 a result, cells with fewer mites or mites that are not effectively reproducing are more likely to
336 stay intact, thus increasing the level of mite non reproduction in the colony (MNR). The second
337 type of trait is a trait expressed by either the workers or the brood, that would disrupt mite
338 reproduction within capped cells (and thus increase MNR). Both trait types can reduce mite
339 infestation in the colony, thus increasing varroa resistance of honey bee colonies. Interestingly,
340 in this study we found markers associated with genes that relate to these two categories.

341 **Impairment of mite reproduction** We took advantage of the recent review by Mondet et al.
342 (2020) [16] that summarises the results from previous association studies on varroa resistance.
343 We performed a liftover to obtain markers and regions coordinates over the latest genome
344 assembly (HAV3.1, [45]) (supplementary table SM3). Two of our significant SNPs, for *recap*
345 fell into a region around 7Mb on chromosome 1 described as a potential quantitative trait
346 locus (QTL) for VSH, in an association study by Tsuruda et al. (2012) [48]. Two other
347 significant markers 1:2891204:G>A and 1:21374478:G>A fell inside genes LOC410758 and
348 LOC413968 respectively. These genes have been identified in a study by Saelao et al. (2020)
349 [49] looking for selection signal in hygienic honey bee populations of the US. In addition, the
350 marker 11:14369154:G>C, significant for *recap*, is located close to the *TpnCi* gene, coding for
351 troponin C type. This gene has been shown to be over-expressed in non-hygienic Africanized
352 bee lines, when compared with hygienic lines, by Teixeira et al. (2021) [50]. Finally markers
353 4:10789077:T>C and 8:1551638:C>T fell into genes found by Ament et al. (2011) [51] as
354 linked to protein abundance in fat body and haemolymph of the adult honey bee. We know
355 that varroa, while infesting the colony, survives by feeding on these bee biological fluids, the
356 fat bodies when on adult bees and the haemolymph when on pupae [52]. These markers might

357 signal either an impact of varroa infestation, or the resilience to such infestation, on biological
358 pathways linked to composition of the honey bee haemolymph and fat body.

359 In the region on chromosome 8, we identified multiple significant markers close to the
360 gene *Ecr*, ecdysone receptor. This gene has been identified as crucial for the reproduction of
361 *Varroa destructor*, within the vitellogenin production pathway, and only produced by honey
362 bees. Hence it stands a key in the interaction between honey bees and varroa [53, 35, 54].

363 **Detection of varroa infested cells by honey bees** If we look more into the general biology
364 of the honey bee, we saw that two markers, significant for *MNR*, on chromosome 10, fell into
365 the 5-HT2beta gene. This gene is known to be a serotonin receptor, thus involved in olfactory
366 pathways in a large number of insects [55, 56]. The biological hypothesis for its role on the
367 resistance to varroa infestation by honey bee would be through cues for the adult bees to perform
368 behaviours leading to better colony resistance to the mite. Olfactory biological pathways thus
369 appear crucial to such response to the infestation. We also noticed that the significant markers
370 3:12973246:A>G and 12973248:A>G fell into LOC413503 (alias GB41230) also found by
371 Mondet et al. (2015) [47], as being differentially expressed in antennae of honey bees expressing
372 VSH behaviour.

373 In the region on chromosome 15, we identified multiple significant markers located within
374 less than 1 Mb downstream from a group of genes coding for odorant binding proteins. These
375 genes are found in two major clusters on the honey bee genome, seven on chromosome 9, nine
376 on chromosome 15 (monophyletic group called C-minus subfamily); and two on chromosome
377 10 and one respectively on chromosome 2 and 12, summing to 21 genes [57, 58]. The cluster
378 located on chromosome 15 contains OBPs 13 to 21, 6 of which have already been mentioned in
379 genomic studies looking at varroa resistance related traits (obp 14, 15, 16, 17, 18 and 21; [16,
380 59]). In particular, Obp14 has been identified as up-regulated in two studies looking at gene
381 and protein expression in VSH bees [47, 60, 61]. Obp18 has been identified by two proteomic
382 studies looking at the VSH and hygienic behaviour traits [62, 61, 60, 63]. Even though these
383 genes were not declared significant in the dedicated analysis we can still hypothesise that they
384 might be involved in some resistance mechanism targeting varroa infestation as they play a
385 major role in sensory pathways. These genes might be relevant for marker assisted selection, as
386 suggested by Guarna et al. (2017) [64] in the case of a tool dedicated to Canadian honey bees
387 selection for hygienic behaviour.

388 **Genetic architecture of varroa resistance**

389 The review by Guichard et al. (2020) [26] reported heritabilities ranging from close to 0 to up
390 to 0.85, with large standard errors, for traits associated with resistance to varroa. More recently
391 Gabel et al. (2023) [65] estimated the heritability of *MNR* to be close to 0.4, which is in the
392 same range as our estimates. Most 95% confidence/credible intervals for heritability estimates

393 found in literature included zero, while those estimates that did not were mostly modest (<
394 0.2). In addition, some repeatability estimates of these traits are low (e.g. [25, 21]). More
395 importantly, estimates based on different populations, e.g. *A. m. mellifera* or *A. m. ligustica* &
396 *carnica* show discrepancy. This could be explained by different genetic architectures involved
397 in these traits.

398 The heritability estimates from this study, for honey bee resistance traits, seem high compared
399 to standard traits measured on livestock species, which could be due to not being able to fully
400 disentangle genetic from environmental stratification. When intending to estimate heritabilities
401 in honey bee one faces the challenge to integrate the potentially large impact of environmental
402 variation. Part of the population structure is possibly associated with such variation, in addition
403 to genetic variation. In this study, we aim to correct for such structure in our sample by
404 thoroughly accounting for covariates, principal components of the genomic relationship matrix.
405 We computed correlations (using Mantel test) between genetic, geographic and environmental
406 distances between colonies and found that environmental variables are slightly correlated with
407 population structure, whereas their correlation with our phenotypes of interest is not significantly
408 different from zero (supplementary table SR9). Environmental variables seemed to have some
409 impacts that we do not take into account in our analysis (supplementary table SR10), and that
410 might have affected slightly our heritability estimates (upward bias). We can measure the
411 extent of this confounding by comparing the GWAS results obtained by performing single locus
412 GWAS (LMM-GWA) or multi-loci GWAS (BSLMM-GWA), because the former corrects for
413 confounding using principal components of the GRM while the latter does not. Consistent
414 with a small effect of phenotype / genotype confounding, we found the PVE estimates with
415 BSLMM-GWA to be usually larger than those obtained with LMM-GWA (Fig. 6, Tab. 1)
416 but the difference was always very small. Overall, we cannot rule out some inflation of PVE
417 estimates due to remaining confounding but it is not likely to affect our general.

418 The proportion of genetic variance explained by large effects (PGE), estimated with the
419 BSLMM-GWA [43] was generally low and included 0, except for *varroa_inf* in *A. m. ligustica*
420 & *carnica*. These estimates support our hypothesis that varroa resistance traits are highly
421 polygenic and not simply driven by a few markers with large effects. The traits linked to
422 varroa infestation and resistance seem to have a small yet significant part of genetic heritability,
423 and thus can be passed on from one generation to next through selection. This is consistent
424 with the few examples of the efficiency of artificial selection for honey bee resistance [14,
425 24]. Our results obtained in more diverse populations imply that genetic selection, natural or
426 artificial, has the potential to drive increased resistance in other contexts, a positive perspective
427 for honey bee populations worldwide. However, and even though we identified genetic markers
428 with significant effects, it is unlikely that large causal mutations, explaining a big part of the
429 phenotypic variance can contribute significantly to this adaptive response.

430 **The future of genome-wide association studies in honey bees**

431 The honey bee genome harbours some specificities. First, it is known to experience a large num-
432 ber of recombination events, with an average recombination rate of 37cM/Mbp [45]. Secondly,
433 the effective population size (i.e. the number of actively reproducing individuals) of a local pop-
434 ulation, or represented in a particular colony, is expected to be rather large, due to the polyandric
435 nature of the natural honey bee mating system, with each queen mating with multiple males
436 from numerous bee colonies within the local mating area, thus avoiding over-representation
437 of a specific animal or genetic line. Such particularities cause low linkage disequilibrium, as
438 seen in the supplementary methods 2, between genomic markers, making it harder to identify
439 candidate loci (QTLs) linked to specific traits and to further select for them.

440 In addition, the honey bee population exhibits a complex genetic diversity. In this study we
441 provide a better understanding of the genetic background behind varroa infestation and resistance
442 in honey bees in general and the French population in particular. Many honey bee colonies are
443 hybrids of various proportions from the three main *Apis mellifera* subspecies found in Europe,
444 *ligustica & carnica*, *mellifera* and *caucasica*. In this study we took advantage of this admixed
445 population to identify genetic markers linked with our traits of interest within genetic types,
446 in hybrids and across these populations, making it possible to see differences in significance
447 and effect depending on the genetic type. Knowledge of linkage disequilibrium (LD) and
448 associated estimation of local haplotypes for regions of interest combined with knowledge
449 on SNP effects for each sub-species can increase our prediction accuracy for different traits.
450 A better understanding of the local genetic background of the hybrid population could help
451 predict effects of specific SNPs. Studies focusing on hybrid colonies, describing their genetic
452 background throughout the genome and comparing different genetic make up could be highly
453 valuable to identify relevant genetic patterns, especially in the context of genomic selection. As
454 an example, multiple genome regions were flagged with more than one significant marker for
455 the trait *recap* but evidence linking these regions to honey bee biology is lacking. In addition,
456 we identified some markers with opposing effects across the different genetic backgrounds,
457 especially *A. m. ligustica & carnica* and *A. m. mellifera*. It appears relevant to short-list these
458 regions as potential regions of interest for future studies geared towards a better honey bee
459 genome annotation and understanding of underlying biological pathways.

460 **Selection on honey bee resistance to varroa** One practical perspective of our work would be
461 to integrate the identified variants into genomic selection program aiming to breed for resistant
462 honey bee colonies. Genomic selection is commonly used in mainstream livestock species. In
463 some countries, such as Germany, beekeepers are grouped in breeding organisations, that make
464 extensive use of artificial insemination to track their mating making the construction of breeding
465 schemes easier [66]. In this context, some studies [67, 68, 69, 70, 71] described the statistical
466 models and sampling strategies that can be successfully applied to implement genomic selection

467 in honey bees. One limitation to the widespread use of these methods is that so far genomic
468 selection has been proven successful when the focus is on a unique honey bee subspecies, e.g.
469 *A. m. carnica*, and concentrate their efforts on phenotypes linked to production. In the context
470 of the French honey bee population, as described in this study, the hybrid stock as well as the
471 complex phenotypes of interest, make it less straightforward to directly apply genomic selection.
472 Highly polygenic traits come with challenges when considering selection and we expect that the
473 markers identified in this study will not be sufficient to contribute alone to the establishment of a
474 genomic selection scheme even though the abundant estimated variances for the phenotypes still
475 supports the possibility of selection, as it has been pursued in the US [18]. For future studies, a
476 primary focus should be put on increasing the sample size, in terms of number of phenotyped
477 and genotyped colonies, to boost the precision, detection power and replication capacity of
478 association studies. A way to improve the robustness of our marker contributions to selection
479 decisions would be, for example, to deepen our information on pedigree [71], as well as having
480 access to a large number of individual queen genotypes [72], rather than reconstructed ones.
481 Note however, that accessing this information comes today with a greater experimental burden,
482 potentially limiting study sample sizes. Hence, the right balance to optimize statistical power
483 needs to be evaluated further. In addition, there is a need for standardised biological samples
484 in terms of genotype and phenotype. The genotypes could either come from large SNP panels,
485 covering the whole genome [73, 71] or come from whole genome sequencing experiments
486 characterised using a genetic diversity panel [40]. The phenotypes could be more automated
487 and obtained with reduced sampling variability, as we propose with a novel, sequence-derived
488 infestation measure (*varroa_mito* trait).
489 Finally, to fully understand the genetic architecture behind varroa resistance one needs to
490 broaden his horizon and look into additional phenomena, not only associated with genetics. For
491 example, it would be necessary to better describe maternal effects, disentangling queen from
492 drones genetic contribution, to our traits.

493 **Materials and Methods**

494 **Honey bee colonies and sampling strategy**

495 The sampling strategy was established to represent the diversity (both in terms of genetic
496 background and beekeeping practices) of honey bee colonies maintained by French beekeepers.
497 A total of 97 beekeepers, located in France, Switzerland, Luxembourg, the Netherlands, Sweden
498 and New Zealand participated in this study (Fig. 7). Foreign colonies were included as they show
499 similar genetic background to the French honey bee population, mostly because of historical
500 or ongoing trade between beekeepers. A total of 1,513 colonies were sampled to go through
501 sequencing. Out of these 1,513, 1,441 *Apis mellifera* colonies were phenotyped, under the
502 condition that each beekeeper contributed at least 6 colonies (from 6 to 125 with on average

503 19.25 colonies). These colonies were phenotyped for multiple traits known to be related to
504 varroa resistance (described in details below), once per colony at the end of the beekeeping
505 season (summer and fall in Europe, *i.e.* typically between July and September).

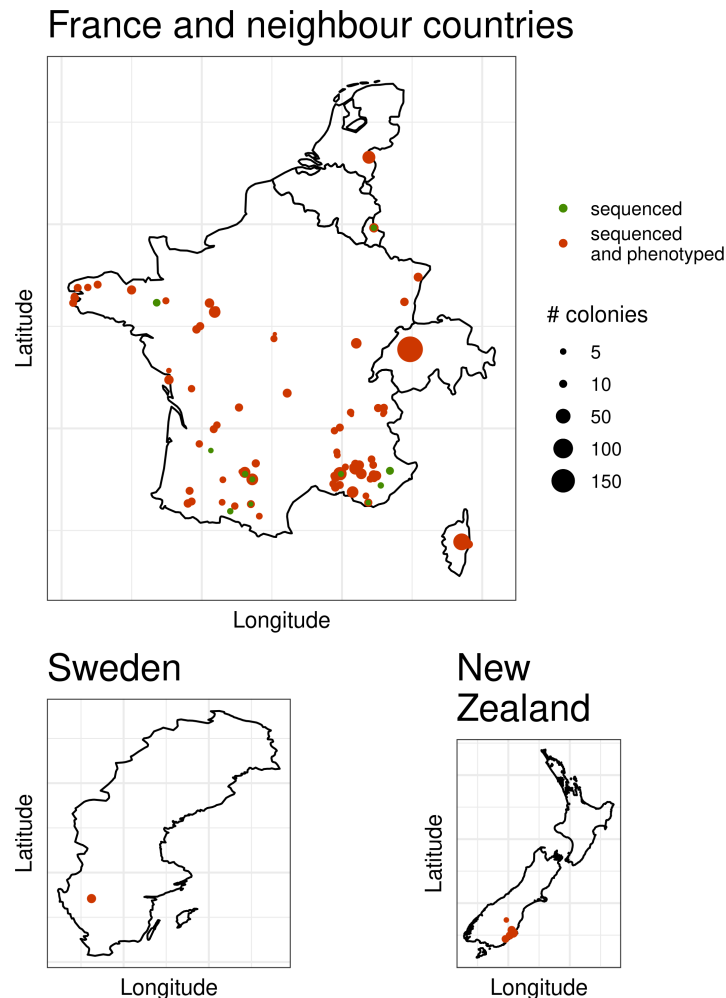


Fig. 7: Geographic distribution of the sampled colonies. Geographical locations of colonies that were whole genome sequenced (in green) and both phenotyped and sequenced (in red). The size of the dot represents the number of honey bee colonies per location and category.

506 **Genetic characterisation of honey bee colonies**

507 All colonies were genotyped from whole genome sequencing of pools of workers using the
508 strategy described in Eynard et al. (2022) [31]. This strategy was decomposed in three
509 steps: first allele counts at Single Nucleotide Polymorphisms (SNPs) were obtained from whole
510 genome sequence, then the genetic background of each colony was estimated from a set of well
511 differentiated markers and finally the genotype of the queen was predicted among colonies of
512 similar genetic background. These different steps are detailed below.

513 **Whole Genome Sequencing** For each colony approximately five hundred honey bee workers
514 were ground in 100 mL of TNE buffer. Fifteen mL of ground sample was then collected and

515 centrifuged for 15 min at 3400 rcf (relative centrifugal force). A volume of 200 μ L of supernatant
516 was lysed overnight at 56°C, with a solution of proteinase K (Eurobio GEXPRK01-B5) and
517 DTT (1,4 Dithiothreitol). Automated DNA extraction was done with a Qiasymphony[®]-Qiagen.
518 DNA concentrations for each sample were estimated with Infinit200[®]-Tecan. Thereafter, pool
519 sequencing was done on a NovaSeq6000 platform in order to obtain about 30X genome-wide raw
520 sequencing coverage. Sequencing reads were aligned on the honey bee reference genome Amel
521 HAv3.1 (Genbank accession GCA_003254395.2, [45]), using BWA-MEM [74]. In addition, it
522 was expected that the biological sample contained adult varroa, present in phoretic phase on the
523 bees, within the pool of sequenced honey bee workers. The obtained reads were therefore also
524 aligned on the varroa mitochondria (genome Vdes_3.0, Genbank accession GCA_002443255.1,
525 [75]) using the same procedure.

526 **SNP Genotyping** Genotypes were estimated at each of the 7,023,976 SNPs identified in
527 Wragg et al. (2022) [40]. Pool sequences were analysed using Samtools mpileup [76] with the
528 recommended parameters: -C minimum mapping quality for reads with excessive mismatches of
529 50, -q minimum mapping quality for an alignment of 20, -Q minimum base quality of 20. Then
530 Pileup files were interpreted by the PoPoolation2 utility mpileup2sync [77], with a minimum
531 quality of 20 and were finally converted to allele counts and sequencing depth files, filtering
532 out real tri-allelic and potential sequencing error. This procedure led to a set of 6,831,074
533 SNPs that were used in all downstream analyses. Colonies were sequenced on average with
534 27.4X coverage, each SNP was on average sequenced with 29.9X coverage, as planned during
535 experimental design (Supplementary table SM1).

536 **Population Structure** For each colony, we ran the model presented in Eynard et al. (2022)
537 [31] to estimate the genetic background on 48,589 SNPs selected so that they differentiate the
538 three main genetic background of honey bees in Europe [78, 40, 79]: the **C** lineage comprising
539 the lowly differentiated *Apis mellifera ligustica* and *Apis mellifera carnica*, the **M** lineage of
540 Western Europe *Apis mellifera mellifera* and the **O** lineage of Eastern Europe/ South-Western
541 Asia *Apis mellifera caucasia*. Specifically, the SNPs were chosen based on the following criteria:
542 (i) there is a maximum of two polymorphic sites within a 100 base pair window, (ii) only one
543 representative marker per linkage disequilibrium block with r^2 higher than 0.8, (iii) the variance
544 between allele frequencies in the three main European genetic background is higher than zero,
545 to allow for population identification and (iv) so that minor allele frequencies (MAF) within
546 the selected markers follows a uniform distribution. The list of selected markers is provided in
547 supplementary table SM2. For step (ii) above, LD was estimated from the reference diversity
548 panel from Wragg et al. (2022) [40] using the plink software version 1.9 [80, 81] with options
549 `-r2 -ld-window 100 -ld-window-kb 10 -ld-window-r2 0.8` and a unique marker was selected
550 manually as the median point for each LD block. The Admixture model from Eynard et al.
551 (2022) [31] was used to estimate the genetic background of each colony (*i.e.* the admixture

552 proportions [82] of each of the three genetic backgrounds).

553 **Queen genotype reconstruction** Following the procedure described in Eynard et al. (2022)
554 [31] we grouped the colonies based on their genetic background (*A. m. ligustica & carnica*, *A.*
555 *m. mellifera* and hybrid) to obtain homogeneous populations. Colonies were if they harboured
556 more than 80% of the same genetic background. Colonies that could not be assigned an
557 homogeneous background colonies were assigned to the 'hybrid' group. Due to a lack of pure
558 *A. m. caucasia* colonies this group was not considered further in the study. Once homogeneous
559 groups are defined it is possible to perform honey bee queen genotype inference using the
560 homogeneous model described in Eynard et al. (2022) [31]. In short, the method is based on
561 the likelihood of the queen genotype, written as

$$x_l^c \vee d_l^c, f_l, g_l^c \sim \text{Binomial} \left(\frac{f_l + g_l^c}{2}, d_l^c \right) \quad (1)$$

562 where g_l^c is the (unknown) queen genotype, f_l is the unknown reference allele frequency in
563 the population, d_l^c and x_l^c the sequencing depth and allele counts obtained from pool sequencing
564 experiments for locus l and colony c . By considering all colonies of the same genetic background
565 jointly, f_l can be estimated by maximum likelihood and the posterior probabilities of the three
566 possible genotypes of the queen computed.

567 **Varroa resistance phenotypes**

568 **Varroa Infestation** Varroa infestation was quantified with four different measures: phoretic
569 mite infestation (on adult bees using two different methods), brood infestation, and total mite
570 load.

571 *Phoretic mite infestation* (v_pho) was measured using the detergent method [83]. In brief,
572 a sample of approximately 300 adult honey bees was collected in each colony, on a frame
573 containing uncapped brood. After weighing of this sample, the number of mites falling as a
574 consequence of washing with a detergent solution was counted, and the proportion of mites
575 within the sample expressed as the number of varroa per 100 honey bees (assuming the weight
576 of 1 single bee to be 140 mg). An alternative measure of phoretic varroa (v_mito) was obtained
577 from the pool sequencing data by calculating the ratio of the number of reads mapping varroa
578 mitochondria on the number of reads mapping the honey bee genome.

579 *Brood infestation* (v_brood) was expressed as the proportion of varroa infesting brood cells
580 in the colony. This proportion was estimated among 300 randomly sampled brood cells on a
581 single frame containing capped brood aged 7 to 11 days post-capping (P5 to P8 stages).

582 *Total mite infestation* (v_load) was estimated by combining the phoretic and brood infesta-
583 tions:

$$v_{load} = v_{brood} \times n_{brood} + v_{pho} \times \frac{n_{bee}}{100} \quad (2)$$

584 where the total number of brood cells (n_{brood}) and adult bees n_{bee} in the colonies were
585 estimated using the ColEval method [84].

586 **Recapping of infested cells** The uncapping and further recapping of varroa infested brood
587 cells by adult honey bees is a behavioural trait that has been shown to be associated with
588 varroa resistance [17]. It can be estimated by dissecting brood cells to detect the presence
589 (non-recapped cell) or absence (recapped cell) of the silk cocoon built by the larvae [85]. This
590 was measured on the colony at the same time as measuring mite non reproduction (MNR). This
591 trait is expressed as the proportion of recapped cells within the infested cells.

592 **Mite non reproduction** Mite Non Reproduction (MNR), originally known as Suppressed
593 Mite reproduction (SMR), was estimated as detailed in Mondet et al. (2020) [15]. In brief,
594 this estimates infers varroa reproductive status for each brood cell infested by a single varroa
595 foundress and provides a proportion of reproductive mites in the colony. It was estimated on
596 about 300 brood cells (some also used to determine mite brood infestation) with the aim to
597 reach at least 35 single mite infested cells.

598 All phenotypes were recorded by technicians having followed an extensive training period
599 prior to sampling. Moreover, for statistical analysis they were corrected to fit the assumption of
600 Normality underlying genome wide association study models. Details on the transformations
601 can be found in Supplementary methods 1.

602 **Phenotypic Characterisation of colonies** The correlation between varroa-associated phe-
603 notypes within and across groups were estimated using the traditional Pearson's method. A
604 Principal Component Analysis (PCA) was performed using the R package FactoMineR [86],
605 after imputation of the missing phenotypes using the R package MissMDA [87]. Missing
606 phenotypes appeared due to sampling difficulties occurring when performing the scoring in
607 the field. The PCA was used to extract uncorrelated synthetic phenotypes to test for genetic
608 association.

609 **Genome wide association studies and meta-analyses**

610 **Genomic relationship matrix** For each group of genetic background identified earlier, *A.*
611 *m. ligustica & carnica*, *A. m. mellifera* and the hybrids, only SNPs with a MAF above
612 0.01 and missing rate below 5% were kept. A genomic relationship matrix (GRM) between
613 colonies of the group was estimated on pool sequencing experiment allele frequencies taking
614 SNP linkage disequilibrium into account through the SNP weights produced by LDAK [88],
615 see supplementary methods 2 for more details. Additionally, in order to describe further genetic

616 structure within the group, a PCA on the GRM was performed, using LDAK [89]. The Horn's
617 parallel analysis [90] was used to decide on the number of principal components kept to explain
618 the variance. 20, 12 and 16 first components were kept from this PCA for *A. m. ligustica* &
619 *carnica*, *A. m. mellifera* and the hybrids respectively.

620 **Genome wide association** Genome wide association studies (GWAS) were performed for
621 three traits: varroa infestation (thereafter called *varrao_inf*), MNR and recapping of varroa
622 infested cells (*recap*). Each GWAS tested the association between the reconstructed queen
623 genotypes and the phenotype using the univariate linear mixed model (lmm) as proposed in
624 GEMMA [43] at each SNP in turn (LMM-GWA), resulting for each SNP in an estimate of its
625 effect and associated standard error, as well as a p-value. In the GWAS, polygenic effects were
626 accounted for with the GRM described above. In addition, further correction was performed
627 by adding the principal components from the PCA on GRM, selected as explained above,
628 as covariates. This was done to correct for the effects of unmeasured confounders with the
629 genetic structure on the phenotype variation (such as apiaries, beekeeper, year ... effects). In
630 supplementary methods 2 we illustrate how the structures of the GRMs correlate somewhat
631 to different environmental structures in the data. Association studies were run for all markers
632 initially available, after filtering for minor allele frequency (MAF) above 0.01 and missing rate
633 below 5%. This led to retain a total of 3,084,335; 2,729,072 and 3,185,994 SNPs for the *A. m.*
634 *ligustica* & *carnica*, *A. m. mellifera* and hybrid individuals respectively.

635 To assess the effectiveness of the correction for population structure, the genomic inflation
636 factor λ_{gc} was estimated as the median of the chi-squared test statistics divided by the expected
637 median of the chi-squared distribution under the null hypothesis. λ_{gc} ranged between 1.02 and
638 1.08 for the GWAS on *A. m. ligustica* & *carnica*, between 0.98 and 1.03 for the GWAS on *A. m.*
639 *mellifera* and between 0.99 and 1.04 for the GWAS on the hybrid colonies therefore showing
640 really little inflation or deflation of the p-values associated to the tested SNPs (supplementary
641 figures S9).

642 To call SNPs significant, a False Discovery Rate procedure was applied, using the adaptive
643 shrinkage method [44] as implemented in the *ashr* R package. Specifically, SNPs with a
644 local false discovery rate (lfdr) and local false sign rate (lfsr) < 0.1 were deemed significant.
645 The lfdr is the probability, knowing the observed data, that an effect is declared significant
646 erroneously and lfsr is the probability, knowing observed data, that the sign of an effect declared
647 significant is wrong [44]. The proportion of phenotypic variance explained by the SNPs (*pve*),
648 and its standard error, was estimated by the univariate linear model and, using the Bayesian
649 Sparse linear mixed model (bslmm, BSLMM-GWA), with default parameters of 0 t 300 SNPs,
650 1,000,000 sampling steps and 100,000 burn-in iterations, proposed by GEMMA [91]. This
651 model was fitted with the GRM and associated covariates and we estimated the proportion of
652 genetic variance explained by the sparse effects (*pge*) of the trait as well as 95% credible interval
653 from empirical estimates.

654 **Meta-analysis** GWAS results for the same phenotype on the three genetic backgrounds were
655 combined with two meta GWAS methods: (i) MANTRA [41] a meta-analysis method dedicated
656 to combine GWAS results from different genetic ancestries (ii) Mash [42] a general purpose data-
657 driven Bayesian meta-analysis method modelling SNP effects with a mixture of multivariate
658 gaussian distributions with different correlation matrices.

659 MANTRA and mash were run on all SNPs, using effects (β) and associated standard errors
660 estimated with the lmm model of GEMMA. For mash inferences, canonical and data-driven
661 covariance matrices were used. The canonical matrices were estimated automatically by mash.
662 Data-driven matrices were: (i) estimated based on extreme deconvolution from PCA matrices,
663 (ii) based on F_{st} values between populations, similar to MANTRA and (iii) based on correlation
664 between SNP or gene effects in the different groups. Mash includes an estimation of the
665 residual correlation. In this analysis the simple residual correlation estimation model was
666 preferred, as it outperformed more complex residual correlation estimation models. SNPs with
667 a $\log_{10}(\text{Bayes Factor}) > 5$ with MANTRA were called significant, a threshold which was shown
668 to be conservative by Wang et al. (2013) [92]. Mash automatically assigns significance to each
669 marker. In our study the corresponding $\log_{10}(\text{BF})$ threshold varied from 1.16 to 1.4 depending
670 on the trait.

671 Genetic correlations were estimated using Pearson's correlation coefficients on the allele
672 effects, for each SNP, calculated by our different GWAS methods (individually with GEMMA
673 and ash, and across co-ancestries with MANTRA and mash).

674 **Gene prioritisation** The variant effect predictor (VEP) tool from Ensembl [93] was used
675 to identify, based on the honey bee genome annotation, for each of the significant SNPs, its
676 impact on the annotation (stop, gained or lost, missense, frameshift...), its closest genes and
677 their locations (upstream, downstream, intronic region ...). Additionally, we identified genes,
678 located elsewhere in the genome in LD regions, containing variants in high LD ($r^2 > 0.8$) with
679 the significant SNP. Linkage disequilibrium was computed, for each group, using data from the
680 genetic diversity reference panel [40], for each significant SNP with all other variants on the
681 same chromosome using the plink software version 1.9 [80, 81] with options `-ld-window-r2`
682 `0.8 -r2 inter-chr`.

683 **Acknowledgements**

684 This study was performed thanks to the consent of more than a hundred private beekeepers
685 and beekeeping groups to take part in this project, including Arista Bee Research (NL) and
686 Betta Bees Research (NZ). This study was performed with the support of the LABOGENA
687 DNA team (Yannick Poquet and Christina Sann) and the ITSAP team (Anne-Laure Guirao and
688 Antoine Sudan) for the maintenance of the honey bee colonies from the ITSAP experimental
689 station and Benoît Droz for the maintenance of the Swiss colonies, the multiple technicians

690 employed on the project (Baptiste Biet, Marius Bredon, Samuel Favier, H el ene Gerwis, Coralie
691 Guerry, Alb eric Delamotte, Pierre Lamy, Antonin Leclercq, Julie Petit, Diane Roriz, Emilien
692 Rottier, Justine Roug e, Agathe Valette, Chlo e Vasseur and Alice Revel) for visiting beekeepers
693 and collecting the data for the project. We thank the sequencing platform GeT-PlaGe Toulouse
694 (France), in particular Tabatha Bulach and R emy-F elix Serre, a partner of the National Infras-
695 tructure France G enomique, thanks to support by the Commissariat aux Grands Investissements
696 (ANR-10-INBS-0009), for the sequencing. Bioinformatics analyses were performed on the
697 computing facility Genotoul. The authors thank LABOGENA DNA for carrying this project.
698 Thanks to Fran ois Gerster for initial discussion and his help to build the consortium and target
699 appropriate funding.

700 **Funding:**

701 MOSAR RT 2015-776 project, FranceAgriMer, the Minist ere de l'Agriculture de l'Agroalimentaire
702 et de la For t.

703 BeeStrong PIA P3A, Minist ere de l'Agriculture de l'Agroalimentaire et de la For t and In-
704 vestissement d'avenir.

705 Bundesamt f ur Landwirtschaft BLW grant no. 627000708, Swiss Federal Office for Agricul-
706 ture.

707 **Author contributions:**

708 Conceptualization: AV, FM, LG, YLC, BB, BS, FP, AD

709 Data collection: AV, FM, YLC, BB, BS, AD

710 Data sampling: FM, BB, BD, MG, the BeeStrong technicians, BL, JM

711 Laboratory: KT, EL, RM

712 Sequencing: OB

713 Data curation: SEE, FM, BB, BD, MG, MN

714 Methodology: BS, SEE, AV, FG

715 Investigation: SEE, FM, AV, BS

716 Visualization: SEE, FM, AV, BS

717 Results interpretation: SEE, BS, AV, FM, FG, FP, YLC

718 Supervision: FM, YLC, AD, AV, BS

719 Project lead: LG

720 Funding acquisition: AV, FM, LG, YLC, BB, BS, FP, AD

721 Writing original draft: SEE, FM, AV, BS

722 Writing review & editing: BB, BL, OB, YLC, BD, AD, LG, MG, JM, FG, MN, FP, KT

723 **Competing interests:** The authors declare that they have no competing interests.

724 **Data and materials availability:** Raw sequences are made available under the bioproject ac-
725 cession PRJNA1083455 on NCBI ([https://www.ncbi.nlm.nih.gov/bioproject?term=](https://www.ncbi.nlm.nih.gov/bioproject?term=PRJNA1083455&cmd=DetailsSearch)
726 [PRJNA1083455&cmd=DetailsSearch](https://www.ncbi.nlm.nih.gov/bioproject?term=PRJNA1083455&cmd=DetailsSearch)). Additional data: count files (from popoolation2), raw
727 phenotypes, accession numbers for the sequences, summary statistics from GWAS analysis
728 and scripts are available in the following repository: <https://doi.org/10.57745/HHE4CZ>.

729 The scripts to perform the analysis are available on github (<https://github.com/seynard/>
730 **gwas_beestrong**)

731

References

- [1] A.M. Klein et al. “Importance of pollinators in changing landscapes for world crops”. In: *Proceedings of the Royal Society B: Biological Sciences* 274.1608 (2006), pp. 303–313. DOI: 10.1098/rspb.2006.3721. URL: <https://doi.org/10.1098/rspb.2006.3721>.
- [2] S. Bruckner et al. “A national survey of managed honey bee colony losses in the USA: results from the Bee Informed Partnership for 2017–18, 2018–19, and 2019–20”. In: *Journal of Apicultural Research* (2023), pp. 1–15. ISSN: 0021-8839. DOI: 10.1080/00218839.2022.2158586. URL: <https://doi.org/10.1080/00218839.2022.2158586>.
- [3] S. G. Potts et al. “Global pollinator declines: trends, impacts and drivers”. In: *Trends in Ecology and Evolution* 25.6 (2010), pp. 345–353. ISSN: 0169-5347. DOI: <https://doi.org/10.1016/j.tree.2010.01.007>. URL: <https://www.sciencedirect.com/science/article/pii/S0169534710000364>.
- [4] D. Goulson et al. “Bee declines driven by combined stress from parasites, pesticides, and lack of flowers”. In: *Science* 347.6229 (2015), p. 1255957. DOI: 10.1126/science.1255957. URL: <https://doi.org/10.1126/science.1255957>.
- [5] K. Traynor et al. “*Varroa destructor*: A Complex Parasite, Crippling Honeybees Worldwide”. In: (2020). DOI: 10.20944/preprints202002.0374.v1.
- [6] Y. Le Conte, M. Ellis, and W. Ritter. “*Varroa* mites and honey bee health: can *Varroa* explain part of the colony losses?” In: *Apidologie* 41.3 (2010), pp. 353–363. URL: <https://doi.org/10.1051/apido/2010017>.
- [7] P. Rosenkranz, P. Aumeier, and B. Ziegelmann. “Biology and control of *Varroa destructor*”. In: *Journal of Invertebrate Pathology* 103.Supplement (2010), S96–S119. URL: <http://www.sciencedirect.com/science/article/pii/S0022201109001906>.
- [8] J. de Miranda et al. “Honey bee viruses and their effect on bee and colony health”. In: 2011, pp. 71–102.
- [9] I. Fries, A. Imdorf, and P. Rosenkranz. “Survival of mite infested (*Varroa destructor*) honey bee (*Apis mellifera*) colonies in a Nordic climate”. In: *Apidologie* 37 (2006). DOI: 10.1051/apido:2006031. URL: <http://dx.doi.org/10.1051/apido:2006031>.
- [10] A. Noël, Y. Le Conte, and F. Mondet. “*Varroa destructor*: how does it harm *Apis mellifera* honey bees and what can be done about it?” In: *Emerging Topics in Life Sciences* 4.1 (2020), pp. 45–57. ISSN: 2397-8554. DOI: 10.1042/ETLS20190125. URL: <https://doi.org/10.1042/ETLS20190125>.

- 766 [11] S. Cremer, S. A. O. Armitage, and P. Schmid-Hempel. “Social Immunity”. In: *Current*
767 *Biology* 17.16 (2007), R693–R702. DOI: [https://doi.org/10.1016/j.cub.](https://doi.org/10.1016/j.cub.2007.06.008)
768 [2007.06.008](https://doi.org/10.1016/j.cub.2007.06.008). URL: [https://www.sciencedirect.com/science/article/pii/](https://www.sciencedirect.com/science/article/pii/S0960982207015035)
769 [S0960982207015035](https://www.sciencedirect.com/science/article/pii/S0960982207015035).
- 770 [12] S. Thaduri et al. “Disentangling host-parasite-pathogen interactions in a varroa-resistant
771 honeybee population reveals virus tolerance as an independent, naturally adapted survival
772 mechanism”. In: *Scientific Reports* (2019). ISSN: 2045-2322. DOI: [10.1038/s41598-](https://doi.org/10.1038/s41598-019-42741-6)
773 [019-42741-6](https://doi.org/10.1038/s41598-019-42741-6). URL: <https://doi.org/10.1038/s41598-019-42741-6>.
- 774 [13] J. J. M. van Alphen and B. J. Fernhout. “Natural selection, selective breeding, and the
775 evolution of resistance of honeybees (*Apis mellifera*) against Varroa”. In: *Zoological*
776 *Letters* 6 (2020). ISSN: 2056-306X. DOI: [10.1186/s40851-](https://doi.org/10.1186/s40851-020-00158-4)
777 [020-00158-4](https://doi.org/10.1186/s40851-020-00158-4). URL: <https://doi.org/10.1186/s40851-020-00158-4>.
- 778 [14] J. R. Harbo and J. W. Harris. “Suppressed mite reproduction explained by the behaviour
779 of adult bees”. In: *Journal of Apicultural Research* 44.1 (2005), pp. 21–23. ISSN: 0021-
780 8839. DOI: [10.1080/00218839.2005.11101141](https://doi.org/10.1080/00218839.2005.11101141). URL: [https://doi.org/10.](https://doi.org/10.1080/00218839.2005.11101141)
781 [1080/00218839.2005.11101141](https://doi.org/10.1080/00218839.2005.11101141).
- 782 [15] F. Mondet et al. “Evaluation of Suppressed Mite Reproduction (SMR) Reveals Potential
783 for Varroa Resistance in European Honey Bees (*Apis mellifera* L.)” In: *Insects* 11.9
784 (2020). ISSN: 2075-4450. DOI: [10.3390/insects11090595](https://doi.org/10.3390/insects11090595).
- 785 [16] F. Mondet et al. “Honey bee survival mechanisms against the parasite *Varroa destructor*: a
786 systematic review of phenotypic and genomic research efforts”. In: *International Journal*
787 *for Parasitology* 50.6 (2020), pp. 433–447. ISSN: 0020-7519. DOI: [https://doi.org/](https://doi.org/10.1016/j.ijpara.2020.03.005)
788 [10.1016/j.ijpara.2020.03.005](https://doi.org/10.1016/j.ijpara.2020.03.005). URL: [https://www.sciencedirect.com/](https://www.sciencedirect.com/science/article/pii/S002075192030093X)
789 [science/article/pii/S002075192030093X](https://www.sciencedirect.com/science/article/pii/S002075192030093X).
- 790 [17] M. Oddie, B. Dahle, and P. Neumann. “Reduced Postcapping Period in Honey Bees
791 Surviving *Varroa destructor* by Means of Natural Selection”. In: *Insects* 9.4 (2018).
- 792 [18] J. R. Harbo and J. W. Harris. “Heritability in honey bees (Hymenoptera: Apidae) of char-
793 acteristics associated with resistance to *Varroa jacobsoni* (Mesostigmata: Varroidae)”.
794 In: *Journal of Economic Entomology* 92.2 (1999), pp. 261–265. DOI: [10.1093/jee/](https://doi.org/10.1093/jee/92.2.261)
795 [92.2.261](https://doi.org/10.1093/jee/92.2.261). URL: <http://dx.doi.org/10.1093/jee/92.2.261>.
- 796 [19] A. von Virag et al. “Decreased Mite Reproduction to Select *Varroa destructor* (Acari:
797 Varroidae) Resistant Honey Bees (Hymenoptera: Apidae): Limitations and Potential
798 Methodological Improvements”. In: *Journal of Economic Entomology* 115.3 (Apr. 2022),
799 pp. 695–705. ISSN: 0022-0493. DOI: [10.1093/jee/toac022](https://doi.org/10.1093/jee/toac022). eprint: [https://](https://academic.oup.com/jee/article-pdf/115/3/695/43977501/toac022.pdf)
800 academic.oup.com/jee/article-pdf/115/3/695/43977501/toac022.pdf.
801 URL: <https://doi.org/10.1093/jee/toac022>.

- 802 [20] T. E. Rinderer et al. “Breeding for resistance to *Varroa destructor* in North America”.
803 In: *Apidologie* 41.3 (2010), pp. 409–424. URL: [https://doi.org/10.1051/apido/](https://doi.org/10.1051/apido/2010015)
804 [2010015](https://doi.org/10.1051/apido/2010015).
- 805 [21] R. Büchler et al. “Evaluation of Traits for the Selection of *Apis Mellifera* for Resistance
806 against *Varroa Destructor*”. In: *Insects* 11.9 (2020). ISSN: 2075-4450. DOI: [10.3390/](https://doi.org/10.3390/insects11090618)
807 [insects11090618](https://doi.org/10.3390/insects11090618).
- 808 [22] W. C. Rothenbuhler. “Behavior genetics of nest cleaning in honey bees. IV. Responses of
809 F1 and backcross generations to disease-killed blood”. In: *American Zoologist* 4 (1964),
810 pp. 111–23. ISSN: 0003-1569 (Print) 0003-1569. DOI: [10.1093/icb/4.2.111](https://doi.org/10.1093/icb/4.2.111).
- 811 [23] K. Bienefeld and F. Pirchner. “Heritabilities for several colony traits in the honeybee
812 (*Apis mellifera carnica*)”. In: *Apidologie* 21.3 (1990), pp. 175–183. URL: [https://](https://doi.org/10.1051/apido:19900302)
813 doi.org/10.1051/apido:19900302.
- 814 [24] J. R. Harbo and J. W. Harris. “Responses to Varroa by honey bees with different levels
815 of Varroa Sensitive Hygiene”. In: *Journal of Apicultural Research* 48.3 (2009), pp. 156–
816 161. ISSN: 0021-8839. DOI: [10.3896/IBRA.1.48.3.02](https://doi.org/10.3896/IBRA.1.48.3.02). URL: [https://doi.org/](https://doi.org/10.3896/IBRA.1.48.3.02)
817 [10.3896/IBRA.1.48.3.02](https://doi.org/10.3896/IBRA.1.48.3.02).
- 818 [25] S. E. Eynard et al. “Descriptive Analysis of the Varroa Non-Reproduction Trait in Honey
819 Bee Colonies and Association with Other Traits Related to Varroa Resistance”. In: *Insects*
820 11.8 (2020). ISSN: 2075-4450 (Print) 2075-4450. DOI: [10.3390/insects11080492](https://doi.org/10.3390/insects11080492).
- 821 [26] M. Guichard et al. “Advances and perspectives in selecting resistance traits against the
822 parasitic mite *Varroa destructor* in honey bees”. In: *Genetics Selection Evolution* 52.1
823 (2020), p. 71. ISSN: 1297-9686. DOI: [10.1186/s12711-020-00591-1](https://doi.org/10.1186/s12711-020-00591-1). URL: [https://](https://doi.org/10.1186/s12711-020-00591-1)
824 doi.org/10.1186/s12711-020-00591-1.
- 825 [27] D. R. Tarpy, S. Hatch, and D. J. C. Fletcher. “The influence of queen age and quality during
826 queen replacement in honeybee colonies”. In: *Animal Behaviour* 59.1 (2000), pp. 97–101.
827 ISSN: 0003-3472. DOI: <https://doi.org/10.1006/anbe.1999.1311>. URL: [https://](https://www.sciencedirect.com/science/article/pii/S0003347299913118)
828 www.sciencedirect.com/science/article/pii/S0003347299913118.
- 829 [28] A. Zayed. “Effective population size in Hymenoptera with complementary sex determi-
830 nation”. In: *Heredity* 93.6 (2004), pp. 627–630. ISSN: 1365-2540. DOI: [10.1038/sj.](https://doi.org/10.1038/sj.hdy.6800588)
831 [hdy.6800588](https://doi.org/10.1038/sj.hdy.6800588). URL: <https://doi.org/10.1038/sj.hdy.6800588>.
- 832 [29] M. Beye et al. “The Gene *csd* Is the Primary Signal for Sexual Development in the Hon-
833 eybee and Encodes an SR-Type Protein”. In: *Cell* 114.4 (2003), pp. 419–429. ISSN: 0092-
834 8674. DOI: [https://doi.org/10.1016/S0092-8674\(03\)00606-8](https://doi.org/10.1016/S0092-8674(03)00606-8). URL: [https://](https://www.sciencedirect.com/science/article/pii/S0092867403006068)
835 www.sciencedirect.com/science/article/pii/S0092867403006068.
- 836 [30] J. Woyke. “Drones from fertilized eggs and the biology of sex determination in the honey
837 bee, *Apimondia*, Prague 12-17.08.1963”. In: Aug. 1963. DOI: [10.13140/2.1.4890.](https://doi.org/10.13140/2.1.4890.8161)
838 [8161](https://doi.org/10.13140/2.1.4890.8161).

- 839 [31] S. E. Eynard et al. “Reconstructing queen genotypes by pool sequencing colonies in
840 eusocial insects: Statistical Methods and their application to honeybee”. In: *Molecular*
841 *Ecology Resources* 22.8 (2022), pp. 3035–3048. ISSN: 1755-098X. DOI: <https://doi.org/10.1111/1755-0998.13685>. URL: <https://doi.org/10.1111/1755-0998.13685>.
- 844 [32] J. R. Shorter et al. “A Genetic Analysis of the Stinging and Guarding Behaviors of the
845 Honey Bee”. In: *Behavior Genetics* 42.4 (2012), pp. 663–674. ISSN: 1573-3297. DOI:
846 [10.1007/s10519-012-9530-5](https://doi.org/10.1007/s10519-012-9530-5). URL: <https://doi.org/10.1007/s10519-012-9530-5>.
- 848 [33] Y. Liu et al. “Larva-mediated chalkbrood resistance-associated single nucleotide poly-
849 morphism markers in the honey bee *Apis mellifera*”. In: *Insect Molecular Biology* 25.3
850 (2016), pp. 239–250. ISSN: 0962-1075. DOI: <https://doi.org/10.1111/imb.12216>.
851 URL: <https://doi.org/10.1111/imb.12216>.
- 852 [34] B. R. Southey et al. “Characterization of Genomic Variants Associated with Scout and
853 Recruit Behavioral Castes in Honey Bees Using Whole-Genome Sequencing”. In: *PLOS*
854 *ONE* 11.1 (2016), e0146430. DOI: [10.1371/journal.pone.0146430](https://doi.org/10.1371/journal.pone.0146430). URL: <https://doi.org/10.1371/journal.pone.0146430>.
- 856 [35] B. H. Conlon et al. “A gene for resistance to the Varroa mite (Acari) in honey bee (*Apis*
857 *mellifera*) pupae”. In: *Molecular Ecology* 0.ja (2019). ISSN: 0962-1083. DOI: [10.1111/mec.15080](https://doi.org/10.1111/mec.15080). URL: <https://doi.org/10.1111/mec.15080>.
- 859 [36] A. Avalos et al. “Genomic regions influencing aggressive behavior in honey bees are de-
860 fined by colony allele frequencies”. In: *Proceedings of the National Academy of Sciences*
861 (2020), p. 201922927. DOI: [10.1073/pnas.1922927117](https://doi.org/10.1073/pnas.1922927117). URL: <http://www.pnas.org/content/early/2020/07/01/1922927117.abstract>.
- 863 [37] M. Guichard et al. “Identification of quantitative trait loci associated with calmness and
864 gentleness in honey bees using whole-genome sequences”. In: *Animal genetics* (2021).
865 ISSN: 0268-9146. DOI: [10.1111/age.13070](https://doi.org/10.1111/age.13070). URL: <https://doi.org/10.1111/age.13070>.
- 867 [38] M. Guichard et al. “Exploring Two Honey Bee Traits for Improving Resistance Against
868 *Varroa destructor*: Development and Genetic Evaluation”. In: *Insects* 12.3 (2021). ISSN:
869 2075-4450. DOI: [10.3390/insects12030216](https://doi.org/10.3390/insects12030216).
- 870 [39] M. Guichard et al. “Two quantitative trait loci are associated with recapping of *Varroa*
871 *destructor*-infested brood cells in *Apis mellifera mellifera*”. In: *Animal genetics* 1365-2052
872 (Electronic) (2022). DOI: [10.1111/age.13150](https://doi.org/10.1111/age.13150).

- 873 [40] D. Wragg et al. “Complex population structure and haplotype patterns in the Western
874 European honey bee from sequencing a large panel of haploid drones”. In: *Molecular*
875 *Ecology Resources* 22.8 (2022), pp. 3068–3086. ISSN: 1755-098X. DOI: <https://doi.org/10.1111/1755-0998.13665>. URL: <https://doi.org/10.1111/1755-0998.13665>.
- 878 [41] A. P. Morris. “Transethnic meta-analysis of genomewide association studies”. In: *Ge-*
879 *netic epidemiology* 35.8 (2011), pp. 809–822. ISSN: 1098-2272/0741-0395. DOI: [10.1002/gepi.20630](https://doi.org/10.1002/gepi.20630). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3460225/>.
- 882 [42] S. M. Uribut et al. “Flexible statistical methods for estimating and testing effects in
883 genomic studies with multiple conditions”. In: *Nature Genetics* 51.1 (2019), pp. 187–
884 195. ISSN: 1546-1718. DOI: [10.1038/s41588-018-0268-8](https://doi.org/10.1038/s41588-018-0268-8). URL: <https://doi.org/10.1038/s41588-018-0268-8>.
- 886 [43] X. Zhou and M. Stephens. “Genome-wide efficient mixed-model analysis for association
887 studies”. In: *Nature genetics* 44.7 (2012), pp. 821–824. URL: <https://www.ncbi.nlm.nih.gov/pmc/PMC3386377/>.
- 889 [44] M. Stephens. “False discovery rates: a new deal”. In: *Biostatistics* 18.2 (2017), pp. 275–
890 294. ISSN: 1465-4644. DOI: [10.1093/biostatistics/kxw041](https://doi.org/10.1093/biostatistics/kxw041). URL: <https://doi.org/10.1093/biostatistics/kxw041>.
- 892 [45] A. Wallberg et al. “A hybrid de novo genome assembly of the honeybee, *Apis mellifera*,
893 with chromosome-length scaffolds”. In: *BMC Genomics* 20.1 (2019), p. 275. ISSN: 1471-
894 2164. DOI: [10.1186/s12864-019-5642-0](https://doi.org/10.1186/s12864-019-5642-0). URL: <https://doi.org/10.1186/s12864-019-5642-0>.
- 896 [46] E. V. Ryabov et al. “A Virulent Strain of Deformed Wing Virus (DWV) of Honeybees
897 (*Apis mellifera*) Prevails after Varroa destructor-Mediated, or In Vitro, Transmission”.
898 In: *PLOS Pathogens* 10.6 (2014), e1004230. DOI: [10.1371/journal.ppat.1004230](https://doi.org/10.1371/journal.ppat.1004230).
899 URL: <https://doi.org/10.1371/journal.ppat.1004230>.
- 900 [47] F. Mondet et al. “Antennae hold a key to Varroa-sensitive hygiene behaviour in honey
901 bees”. In: *Scientific Reports* 5.1 (2015), p. 10454. ISSN: 2045-2322. DOI: [10.1038/srep10454](https://doi.org/10.1038/srep10454). URL: <https://doi.org/10.1038/srep10454>.
- 903 [48] J. M. Tsuruda et al. “High-resolution linkage analyses to identify genes that influence
904 Varroa sensitive hygiene behavior in honey bees”. In: *PLOS ONE* 7.11 (2012), e48276.
905 URL: <https://doi.org/10.1371/journal.pone.0048276>.
- 906 [49] P. Saelao et al. “Genome-wide patterns of differentiation within and among U.S. com-
907 mercial honey bee stocks”. In: *BMC Genomics* 21.1 (2020), p. 704. ISSN: 1471-2164.
908 DOI: [10.1186/s12864-020-07111-x](https://doi.org/10.1186/s12864-020-07111-x). URL: <https://doi.org/10.1186/s12864-020-07111-x>.
- 909

- 910 [50] E. W. Teixeira et al. “Transcriptomic analysis suggests candidate genes for hygienic
911 behavior in African-derived *Apis mellifera* honeybees”. In: *Apidologie* (2021). ISSN:
912 1297-9678. DOI: 10.1007/s13592-020-00834-6. URL: [https://doi.org/10.](https://doi.org/10.1007/s13592-020-00834-6)
913 [1007/s13592-020-00834-6](https://doi.org/10.1007/s13592-020-00834-6).
- 914 [51] S. A. Ament et al. “Mechanisms of stable lipid loss in a social insect”. In: *Journal of*
915 *Experimental Biology* 214.22 (2011), pp. 3808–3821. ISSN: 0022-0949. DOI: 10.1242/
916 [jeb.060244](https://doi.org/10.1242/jeb.060244). URL: <https://doi.org/10.1242/jeb.060244>.
- 917 [52] B. Han et al. “Life-history stage determines the diet of ectoparasitic mites on their
918 honey bee hosts”. In: *Nature Communications* 15.1 (2024), p. 725. ISSN: 2041-1723. DOI:
919 10.1038/s41467-024-44915-x. URL: [https://doi.org/10.1038/s41467-024-](https://doi.org/10.1038/s41467-024-44915-x)
920 [44915-x](https://doi.org/10.1038/s41467-024-44915-x).
- 921 [53] M. F. Feldlaufer and K. Hartfelder. “Relationship of the neutral sterols and ecdysteroids of
922 the parasitic mite, *Varroa jacobsoni* to those of the honey bee, *Apis mellifera*”. In: *Journal*
923 *of Insect Physiology* 43.6 (1997), pp. 541–545. ISSN: 0022-1910. DOI: [https://doi.](https://doi.org/10.1016/S0022-1910(97)00005-X)
924 [org/10.1016/S0022-1910\(97\)00005-X](https://doi.org/10.1016/S0022-1910(97)00005-X). URL: [https://www.sciencedirect.](https://www.sciencedirect.com/science/article/pii/S002219109700005X)
925 [com/science/article/pii/S002219109700005X](https://www.sciencedirect.com/science/article/pii/S002219109700005X).
- 926 [54] C. M. Aurori et al. “Juvenile hormone pathway in honey bee larvae: A source of possible
927 signal molecules for the reproductive behavior of *Varroa destructor*”. In: *Ecology and*
928 *Evolution* 11.2 (2021), pp. 1057–1068. ISSN: 2045-7758. DOI: [https://doi.org/10.](https://doi.org/10.1002/ece3.7125)
929 [1002/ece3.7125](https://doi.org/10.1002/ece3.7125). URL: <https://doi.org/10.1002/ece3.7125>.
- 930 [55] A. S. French et al. “The role of serotonin in feeding and gut contractions in the hon-
931 eybee”. In: *Journal of insect physiology* 61.100 (2014), pp. 8–15. DOI: 10.1016/j.
932 [jinsphys.2013.12.005](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3969292/). URL: [https://www.ncbi.nlm.nih.gov/pmc/articles/](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3969292/)
933 [PMC3969292/](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3969292/).
- 934 [56] H. Verlinden, R. Vleugels, and J. Vanden Broeck. “Serotonin, serotonin receptors and
935 their actions in insects”. In: *Neurotransmitter* 2 (2015), e314. DOI: 10.14800/nt.314.
- 936 [57] S. Forêt and R. Maleszka. “Function and evolution of a gene family encoding odorant
937 binding-like proteins in a social insect, the honey bee (*Apis mellifera*)”. In: *Genome*
938 *research* 16.11 (2006), pp. 1404–1413. DOI: 10.1101/gr.5075706. URL: [https:](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1626642/)
939 [//www.ncbi.nlm.nih.gov/pmc/articles/PMC1626642/](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1626642/).
- 940 [58] B. Mam, S. D. Karpe, and R. Sowdhamini. “Minus-C subfamily has diverged from
941 Classic odorant-binding proteins in honeybees”. In: *Apidologie* 54.1 (2023), p. 12. ISSN:
942 1297-9678. DOI: 10.1007/s13592-022-00988-5. URL: [https://doi.org/10.](https://doi.org/10.1007/s13592-022-00988-5)
943 [1007/s13592-022-00988-5](https://doi.org/10.1007/s13592-022-00988-5).

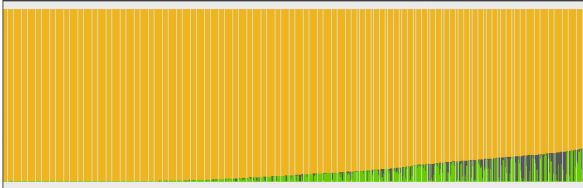
- 944 [59] H. Gebremedhn et al. “Expression of Molecular Markers of Resilience against *Varroa*
945 *destructor* and Bee Viruses in Ethiopian Honey Bees (*Apis mellifera simensis*) Focussing
946 on Olfactory Sensing and the RNA Interference Machinery”. In: *Insects* 14.5 (2023).
947 DOI: 10.3390/insects14050436.
- 948 [60] S. Jiang et al. “Differential gene expression of two extreme honey bee (*Apis mellifera*)
949 colonies showing varroa tolerance and susceptibility”. In: *Insect Molecular Biology* 25.3
950 (2016), pp. 272–282. ISSN: 0962-1075. DOI: <https://doi.org/10.1111/imb.12217>.
951 URL: <https://doi.org/10.1111/imb.12217>.
- 952 [61] H. Hu et al. “Proteome Analysis of the Hemolymph, Mushroom Body, and Antenna
953 Provides Novel Insight into Honeybee Resistance against *Varroa* Infestation”. In: *Journal*
954 *of Proteome Research* 15.8 (2016), pp. 2841–2854. ISSN: 1535-3893. DOI: 10.1021/
955 [acs.jproteome.6b00423](https://doi.org/10.1021/acs.jproteome.6b00423). URL: <https://doi.org/10.1021/acs.jproteome.6b00423>.
- 957 [62] M. Guarna et al. “A search for protein biomarkers links olfactory signal transduction to
958 social immunity”. In: *BMC Genomics* 16 (2015). DOI: 10.1186/s12864-014-1193-6.
- 959 [63] A. McAfee et al. “A death pheromone, oleic acid, triggers hygienic behavior in honey
960 bees (*Apis mellifera* L.)” In: *Scientific Reports* 8.1 (2018), p. 5719. ISSN: 2045-2322. DOI:
961 10.1038/s41598-018-24054-2. URL: <https://doi.org/10.1038/s41598-018-24054-2>.
- 963 [64] M. Marta Guarna et al. “Peptide biomarkers used for the selective breeding of a complex
964 polygenic trait in honey bees”. In: *Scientific Reports* 7.1 (2017), p. 8381. ISSN: 2045-2322.
965 DOI: 10.1038/s41598-017-08464-2. URL: <https://doi.org/10.1038/s41598-017-08464-2>.
- 967 [65] M. Gabel et al. “Heritability of *Apis mellifera* recapping behavior and suppressed mite
968 reproduction as resistance traits towards *Varroa destructor*”. In: *Frontiers in Insect Sci-*
969 *ence* 3 (2023). ISSN: 2673-8600. URL: [https://www.frontiersin.org/articles/](https://www.frontiersin.org/articles/10.3389/finsc.2023.1135187)
970 [10.3389/finsc.2023.1135187](https://www.frontiersin.org/articles/10.3389/finsc.2023.1135187).
- 971 [66] S. Lorenz. “The endangerment of bees and new developments in beekeeping: a social
972 science perspective using the example of Germany”. In: *International Journal of Environ-*
973 *mental Studies* 73.6 (2016), pp. 988–1005. ISSN: 0020-7233. DOI: 10.1080/00207233.
974 2016.1220703. URL: <https://doi.org/10.1080/00207233.2016.1220703>.
- 975 [67] M. Du et al. “Consequences of incorrect genetic parameter estimates for single-trait
976 and multi-trait genetic evaluations in honeybees”. In: *Journal of Animal Breeding and*
977 *Genetics* 139.6 (2022), pp. 666–678. ISSN: 0931-2668. DOI: <https://doi.org/10.1111/jbg.12728>.
978 URL: <https://doi.org/10.1111/jbg.12728>.

- 979 [68] M. Du et al. “Influence of model selection and data structure on the estimation of genetic
980 parameters in honeybee populations”. In: *G3 Genes|Genomes|Genetics* (2022), jkab450.
981 ISSN: 2160-1836. DOI: 10.1093/g3journal/jkab450. URL: [https://doi.org/10.
982 1093/g3journal/jkab450](https://doi.org/10.1093/g3journal/jkab450).
- 983 [69] A. Hoppe et al. “Heritability of disease resistance to chronic bee paralysis, chalkbrood and
984 nose mites in the honeybee (*A.m.carnica*)”. In: *Proceedings of 12th World Congress on
985 Genetics Applied to Livestock Production (WCGALP)*. Wageningen Academic Publishers,
986 2022.
- 987 [70] A. Uzunov et al. “606. Evaluating the potential for mating control in honey bee breeding
988 in three SE European countries (preliminary results)”. In: *Proceedings of 12th World
989 Congress on Genetics Applied to Livestock Production (WCGALP)*. Wageningen Aca-
990 demic Publishers, 2022. Chap. 606, pp. 2508–2511. DOI: doi : 10 . 3920 / 978 - 90 -
991 8686 - 940 - 4 _ 60610 . 3920 / 978 - 90 - 8686 - 940 - 4 _ 606 . URL: [https://doi.org/
992 10.3920/978-90-8686-940-4_606](https://doi.org/10.3920/978-90-8686-940-4_606).
- 993 [71] R. Bernstein et al. “First large-scale genomic prediction in the honey bee”. In: *Heredity*
994 (2023). ISSN: 1365-2540. DOI: 10.1038/s41437-023-00606-9. URL: [https://doi.
995 org/10.1038/s41437-023-00606-9](https://doi.org/10.1038/s41437-023-00606-9).
- 996 [72] R. Bernstein et al. “Simulation studies to optimize genomic selection in honey bees”.
997 In: *Genetics Selection Evolution* 53.1 (2021), p. 64. ISSN: 1297-9686. DOI: 10.1186/
998 s12711-021-00654-x. URL: <https://doi.org/10.1186/s12711-021-00654-x>.
- 999 [73] J. C. Jones et al. “Tool for genomic selection and breeding to evolutionary adaptation:
1000 Development of a 100K single nucleotide polymorphism array for the honey bee”. In:
1001 *Ecology and Evolution* 10.13 (2020), pp. 6246–6256. ISSN: 2045-7758. DOI: 10.1002/
1002 ece3.6357. URL: <https://doi.org/10.1002/ece3.6357>.
- 1003 [74] H. Li. “Aligning sequence reads, clone sequences and assembly contigs with BWA-
1004 MEM”. In: *arXiv preprint arXiv:1303.3997* (2013).
- 1005 [75] M. A. Techer et al. “Divergent evolutionary trajectories following speciation in two
1006 ectoparasitic honey bee mites”. In: *Communications Biology* 2.1 (2019), p. 357. DOI:
1007 10.1038/s42003-019-0606-0. URL: [https://doi.org/10.1038/s42003-019-
1008 0606-0](https://doi.org/10.1038/s42003-019-0606-0).
- 1009 [76] H. Li and R. Durbin. “Fast and accurate short read alignment with Burrows–Wheeler
1010 transform”. In: *Bioinformatics* 25.14 (2009), pp. 1754–1760. ISSN: 1367-4803. DOI: 10.
1011 1093/bioinformatics/btp324. URL: [https://doi.org/10.1093/bioinformatics/
1012 btp324](https://doi.org/10.1093/bioinformatics/btp324).

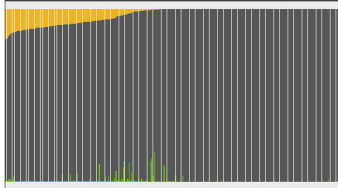
- 1013 [77] R. Kofler, R. V. Pandey, and C. Schlötterer. “PoPoolation2: identifying differentiation
1014 between populations using sequencing of pooled DNA samples (Pool-Seq)”. In: *Bioinform-*
1015 *atics* 27.24 (2011), pp. 3435–3436. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/
1016 btr589. URL: <https://doi.org/10.1093/bioinformatics/btr589>.
- 1017 [78] K. A. Dogantzis et al. “Thrice out of Asia and the adaptive radiation of the western honey
1018 bee”. In: *Science Advances* 7.49 (2021), eabj2151. DOI: 10.1126/sciadv.abj2151.
1019 URL: <https://doi.org/10.1126/sciadv.abj2151>.
- 1020 [79] J. Obšteter et al. “Inferring whole-genome tree sequences and population and demo-
1021 graphic parameters of the Western honeybee”. In: *Proceedings of 12th World Congress*
1022 *on Genetics Applied to Livestock Production (WCGALP)*. Wageningen Academic Pub-
1023 lishers, 2022.
- 1024 [80] Purcell S. and Chang C. *PLINK 1.9*. URL: www.cog-genomics.org/plink/1.9/.
- 1025 [81] S. Purcell et al. “PLINK: a tool set for whole-genome association and population-based
1026 linkage analyses”. In: *American journal of human genetics* 81.3 (2007), pp. 559–575.
1027 ISSN: 0002-9297/1537-6605. DOI: 10.1086/519795. URL: [https://www.ncbi.nlm.](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1950838/)
1028 [nih.gov/pmc/articles/PMC1950838/](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1950838/).
- 1029 [82] J. K. Pritchard, M. Stephens, and P. Donnelly. “Inference of population structure using
1030 multilocus genotype data”. In: *Genetics* 155.2 (2000), pp. 945–959. ISSN: 0016-6731.
1031 URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1412715/>.
- 1032 [83] V. Dietemann et al. “Standard methods for varroa research”. In: *Journal of Apicultural*
1033 *Research* 52 (2012), pp. 1–54. DOI: 10.3896/IBRA.1.52.1.09.
- 1034 [84] J. Hernandez et al. “ColEval: Honeybee COLony Structure EVALuation for Field Sur-
1035 veys”. In: *Insects* 11.1 (2020). DOI: 10.3390/insects11010041.
- 1036 [85] R. Büchler et al. “Screening for low Varroa mite reproduction (SMR) and recapping in
1037 European honey bees”. In: (2017).
- 1038 [86] S. Lê, J. Josse, and F. Husson. “FactoMineR: An R Package for Multivariate Analysis”. In:
1039 *Journal of Statistical Software* 25.1 (2008), pp. 1–18. DOI: 10.18637/jss.v025.i01.
1040 URL: <https://www.jstatsoft.org/index.php/jss/article/view/v025i01>.
- 1041 [87] J. Josse and F. Husson. “missMDA: A Package for Handling Missing Values in Multi-
1042 variate Data Analysis”. In: *Journal of Statistical Software* 70.1 (2016), pp. 1–31. DOI:
1043 10.18637/jss.v070.i01. URL: [https://www.jstatsoft.org/index.php/jss/](https://www.jstatsoft.org/index.php/jss/article/view/v070i01)
1044 [article/view/v070i01](https://www.jstatsoft.org/index.php/jss/article/view/v070i01).
- 1045 [88] D. Speed et al. “Reevaluation of SNP heritability in complex human traits”. In: *Nature*
1046 *Genetics* 49.7 (2017), pp. 986–992. ISSN: 1546-1718. DOI: 10.1038/ng.3865. URL:
1047 <https://doi.org/10.1038/ng.3865>.

- 1048 [89] D. Speed et al. “Improved heritability estimation from genome-wide SNPs”. In: *American*
1049 *journal of human genetics* 91.6 (2012), pp. 1011–1021. ISSN: 1537-66050002-9297. DOI:
1050 10.1016/j.ajhg.2012.10.010. URL: [https://www.ncbi.nlm.nih.gov/pmc/](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3516604/)
1051 [articles/PMC3516604/](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3516604/).
- 1052 [90] A. Dinno. “Implementing Horn’s Parallel Analysis for Principal Component Analy-
1053 sis and Factor Analysis”. In: *Stata Journal* 9 (2009), pp. 291–298. DOI: 10.1177/
1054 1536867X0900900207.
- 1055 [91] X. Zhou, P. Carbonetto, and M. Stephens. “Polygenic Modeling with Bayesian Sparse
1056 Linear Mixed Models”. In: *PLOS Genetics* 9.2 (2013), e1003264. DOI: 10.1371/
1057 journal.pgen.1003264. URL: [https://doi.org/10.1371/journal.pgen.](https://doi.org/10.1371/journal.pgen.1003264)
1058 [1003264](https://doi.org/10.1371/journal.pgen.1003264).
- 1059 [92] X. Wang et al. “Comparing methods for performing trans-ethnic meta-analysis of genome-
1060 wide association studies”. In: *Human Molecular Genetics* 22.11 (Feb. 2013), pp. 2303–
1061 2311. ISSN: 0964-6906. DOI: 10.1093/hmg/ddt064. eprint: [https://academic.](https://academic.oup.com/hmg/article-pdf/22/11/2303/17258469/ddt064.pdf)
1062 [oup.com/hmg/article-pdf/22/11/2303/17258469/ddt064.pdf](https://academic.oup.com/hmg/article-pdf/22/11/2303/17258469/ddt064.pdf). URL: [https:](https://doi.org/10.1093/hmg/ddt064)
1063 [//doi.org/10.1093/hmg/ddt064](https://doi.org/10.1093/hmg/ddt064).
- 1064 [93] W. McLaren et al. “The Ensembl Variant Effect Predictor”. In: *Genome Biology* 17.1
1065 (2016), p. 122. ISSN: 1474-760X. DOI: 10.1186/s13059-016-0974-4. URL: [https:](https://doi.org/10.1186/s13059-016-0974-4)
1066 [//doi.org/10.1186/s13059-016-0974-4](https://doi.org/10.1186/s13059-016-0974-4).

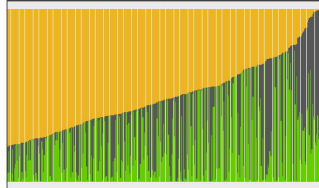
Ligustica & Carnica



Mellifera



Hybrids



Genetic background



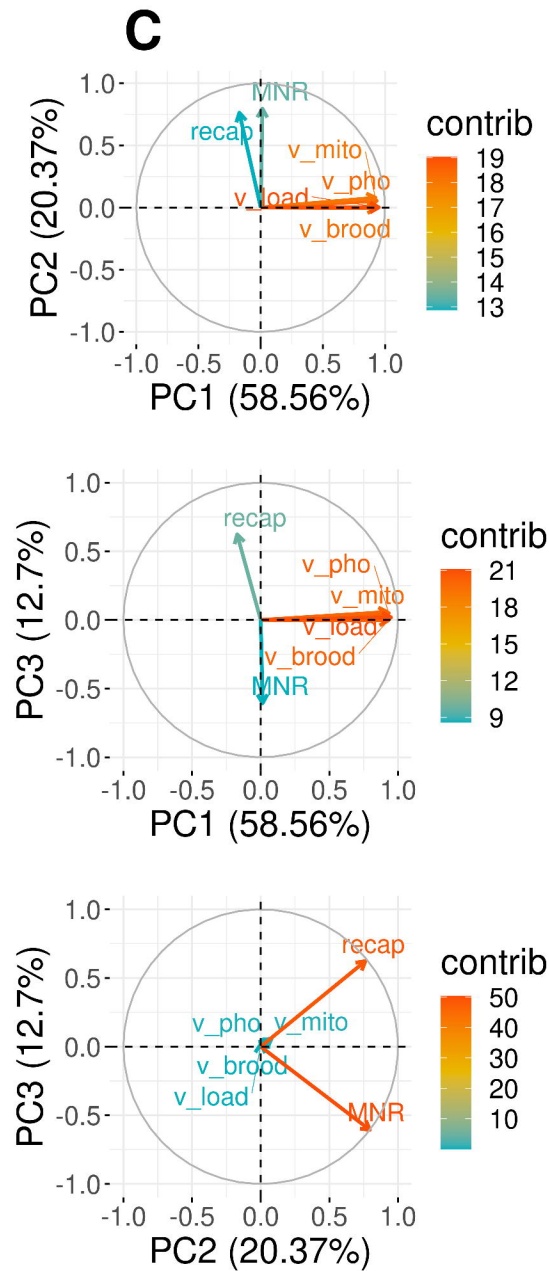
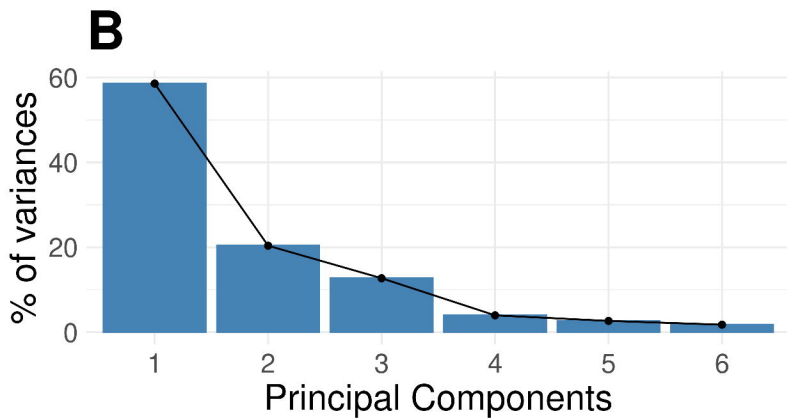
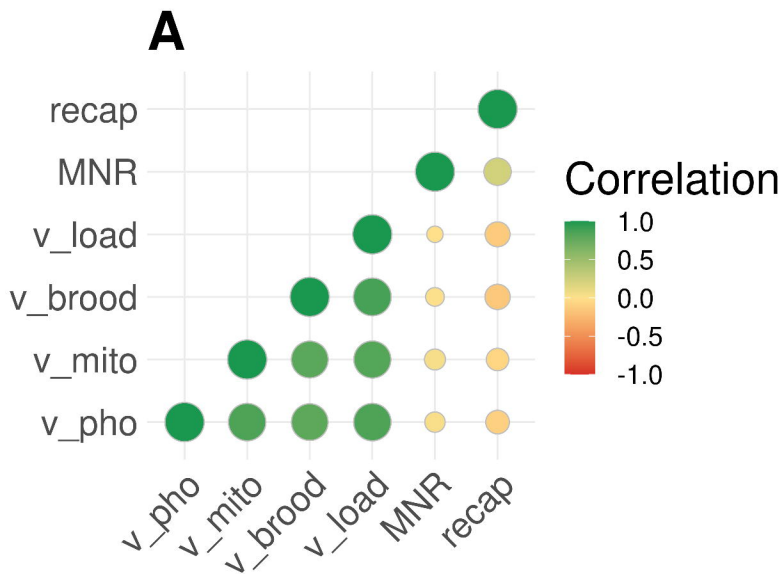
Ligustica & Carnica

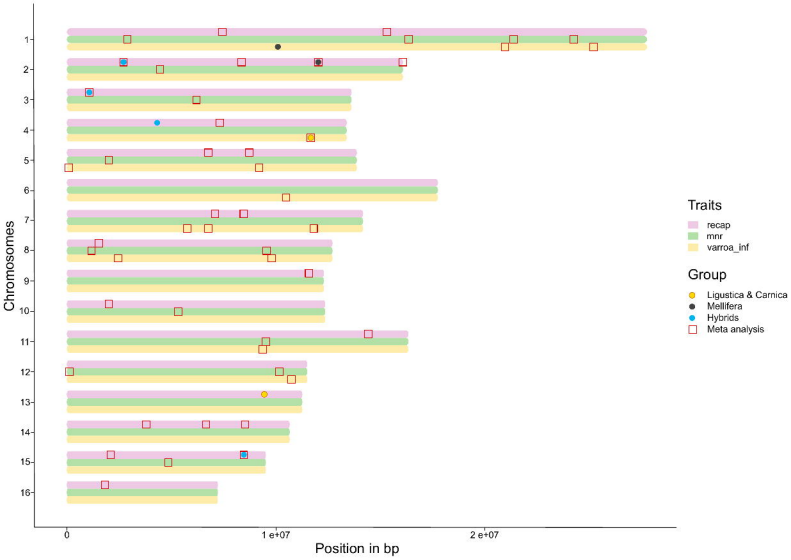


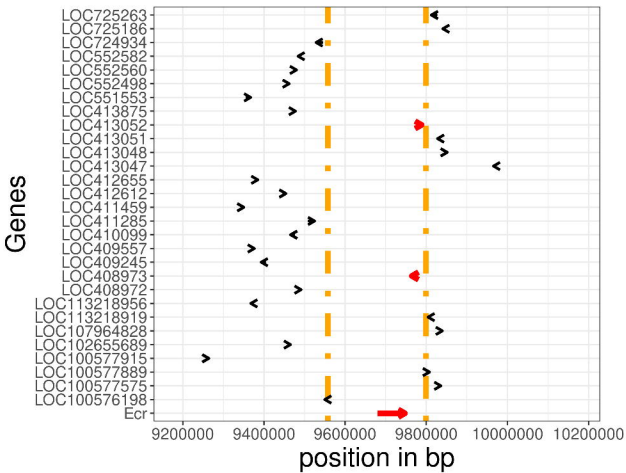
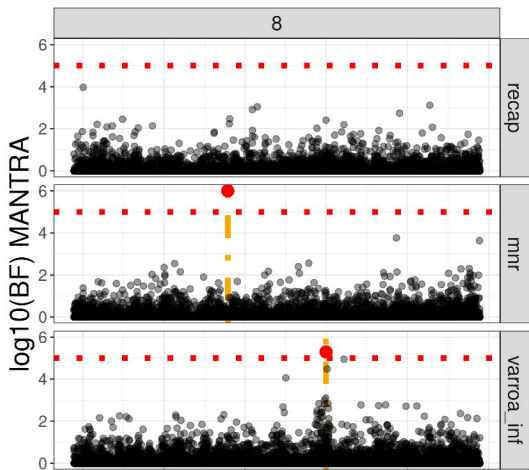
Mellifera



Caucasia

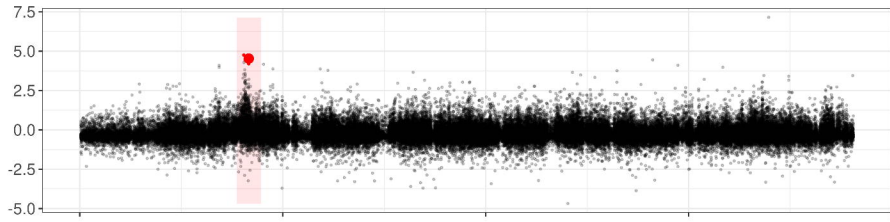




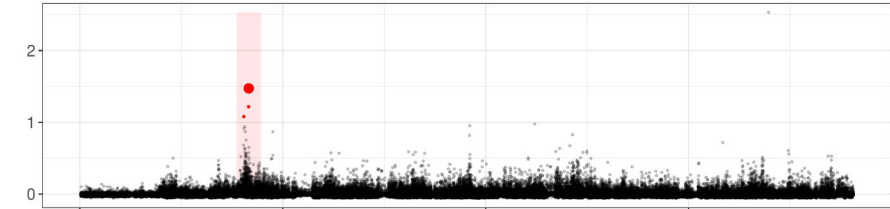


15:2081876

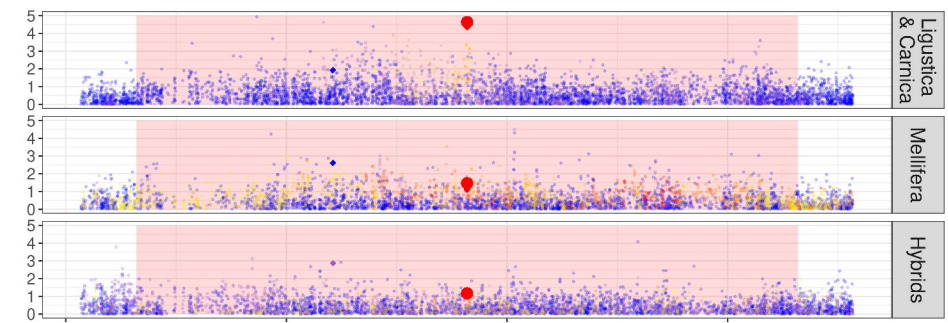
log10 BF Mantra



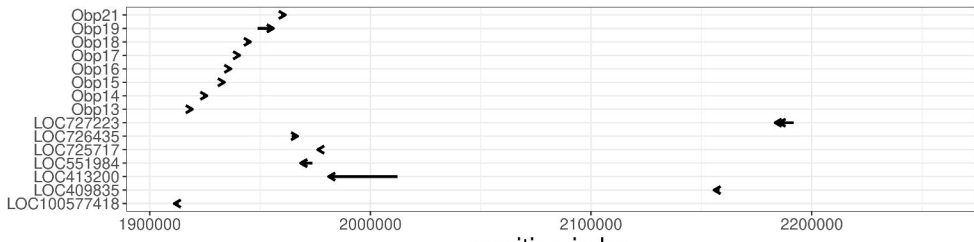
log10 BF Mash



$-\log_{10}(p_val)$

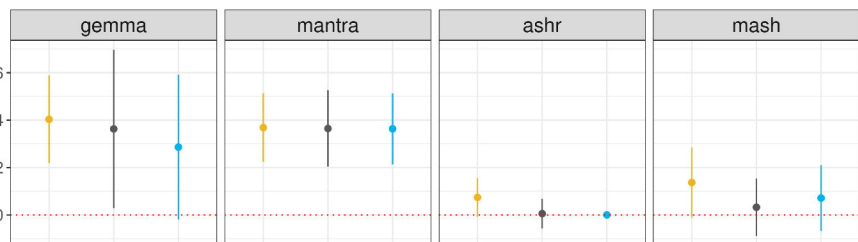


Genes

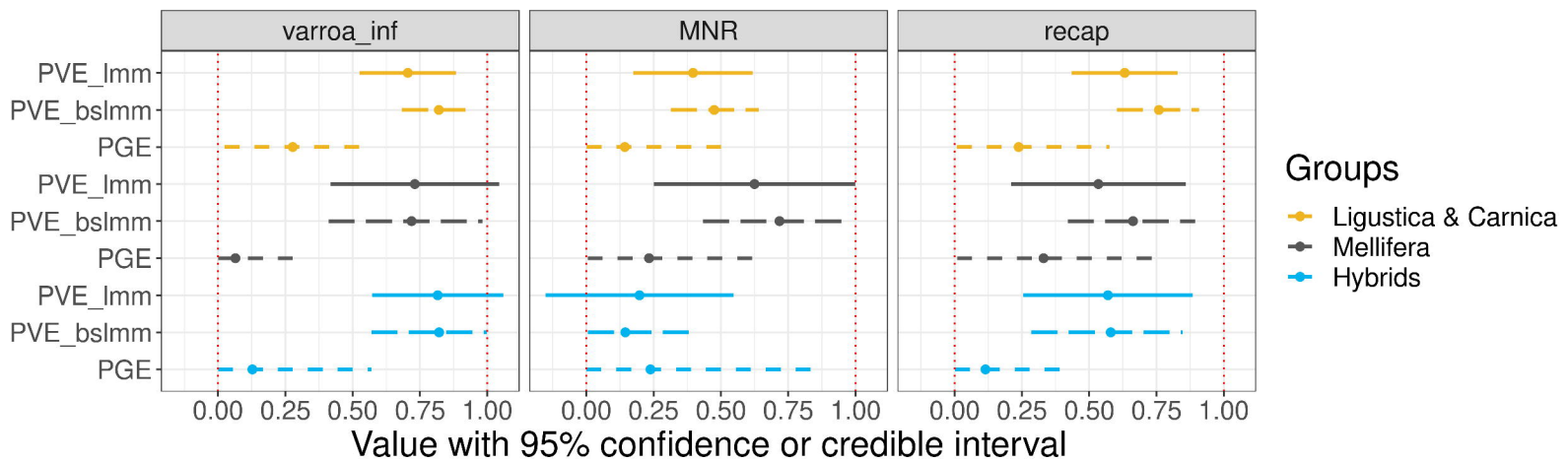


position in bp

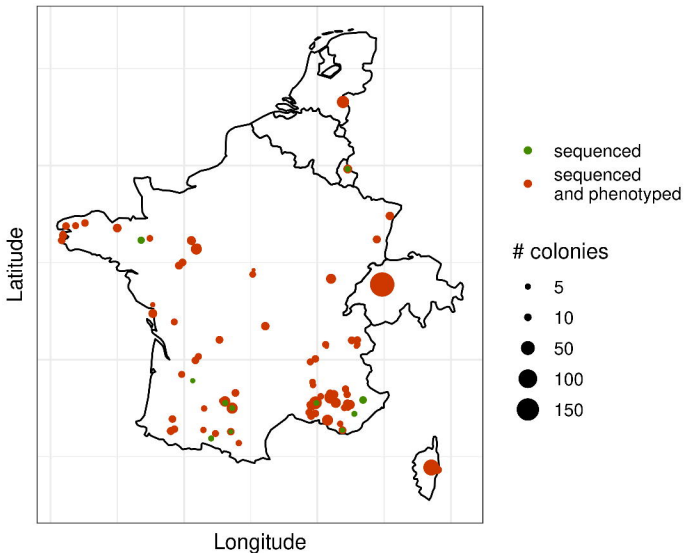
SNP effect



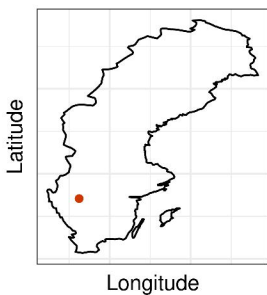
Group
 ● Ligustica & Carnica
 ● Mellifera
 ● Hybrids



France and neighbour countries



Sweden



New Zealand

