Signatures of Positive Selection in the Genome of *Apis mellifera carnica*: A Subspecies of European Honeybees

Qiang Huang [1,2], Yong-Qiang Zhu [3], Bertrand Fouks [4], Xu-Jiang He [1,2], Qing-Sheng Niu [5], Hua-Jun Zheng [3,*] and Zhi-Jiang Zeng [1,2,*]

[1] Honeybee Research Institute, Jiangxi Agricultural University, Zhimin Ave. 1101, Nanchang 330045, China
[2] Jiangxi Province Key Laboratory of Honeybee Biology and Beekeeping, Jiangxi Agricultural University, Zhimin Ave. 1101, Nanchang 330045, China
[3] Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center at Shanghai and Shanghai Institute for Biomedical and Pharmaceutical Technologies, Shanghai 200237, China
[4] Institute for Evolution and Biodiversity, Molecular Evolution and Bioinformatics, Westfälische Wilhelms Universität, 48149 Münster, Germany
[5] Apiculture Science Institute of Jinlin Province, Yuanlin Rd., Jinlin 132108, China
* Correspondence: zhenghj@chgc.sh.cn (H.-J.Z.); bees1965@sina.com (Z.-J.Z.)

Abstract: The technology of long reads substantially improved the contingency of the genome assembly, particularly resolving contiguity of the repetitive regions. By integrating the interactive fragment using Hi-C, and the HiFi technique, a solid genome of the honeybee *Apis mellifera carnica* was assembled at the chromosomal level. A distinctive pattern of genes involved in social evolution was found by comparing it with social and solitary bees. A positive selection was identified in genes involved with cold tolerance, which likely underlies the adaptation of this European honeybee subspecies in the north hemisphere. The availability of this new high-quality genome will foster further studies and advances on genome variation during subspeciation, honeybee breeding and comparative genomics.

Keywords: honeybee; subspecies; selection; sociality; cold tolerance

1. Introduction

Insect pollination is essential for maintaining the balance of the ecosystem, which contributes approximately 35% to crop pollination [1]. The honeybee *Apis mellifera* is a key managed pollinator, with an estimated annual global economic value of $195 billion [2]. The recent decline of honeybee colonies has provoked serious concerns regarding the biodiversity, as well as food security [3]. A number of stressors have been proven to cause the honeybee collapses, including parasites, pesticide, climate change and habitat loss [4–8]. Among those stressors, parasites are a major cause leading to colony losses, with a synergistic effect with pesticide. It is known that honeybee strains showed variation in tolerance towards parasites and climates [9–13]. However, the genetic mechanism underlying that tolerance is yet not fully understood. In this study, we aim to reveal the genomic variation within *A. mellifera* species, which is essential to understand their evolution, adaptation to different climates, as well as to refine breeding strategies for disease-tolerant strains.

In this study, we provided a highly contiguous genome assembly of a valuable subspecies of the European honeybee, *Apis mellifera carnica*. Two complementary sequencing technologies of HIFI and Hi-C were used to generate a high-quality chromosome genome assembly. Comparative genomic analysis revealed gene families selected during the social evolution and climate adaptation.

2. Material and Methods

2.1. DNA Extraction and Sequencing Library Preparation

The honeybee *A. mellifera carnica* were collected from the national honeybee breeding center in Apicultural Science Institute of Jilin, China. Six *A. mellifera carnica* drone pupae were collected from a single colony, which were pooled for DNA extraction by AxyPrep™ Multisource Genomic DNA Miniprep Kit (Axygen, Irvine, CA, USA). The Qubitfluorimetry system was used to define the DNA concentration (Thermo Fisher, Waltham, MA, USA). Fragment size distribution was assessed using the Agilent 2100 Bioanalyzer with the 12,000 DNA kit (Agilent, Santa Clara, CA, USA). Then, 5 μg of high molecular weight genomic DNA was used to prepare the library. The DNA which uses g-Tube (Covaris, Woburn, MA, USA) to shear into 10 kb was used as input into the SMRTcell library preparation according to PacBio 10 kb library preparation protocol. The library was sequenced on a PacBio Sequel system using Sequencing Kit 3.0.

Additionally, five drone pupae from the same colony were fixed with formaldehyde and lysed for Hi-C library. The cross-linked genomic DNA was digested with Hind III overnight. Sticky ends of the genomic DNA were biotinylated and proximity-ligated to form chimeric junctions that were enriched for and then physically sheared to a size of 300–700 bp. Chimeric fragments representing the original cross-linked long-distance physical interactions were then processed into paired-end sequencing libraries and sequenced on the Illumina HiSeq X Ten platform.

2.2. Genome Assembly and Gene Annotation

The reads were filtered through Fastp with default parameters, and 2,602,139 clean Pacbio reads with a total size of 22,818,161,032 bp were obtained [14]. The reads were assembled by HGAP4 of SMRT Link (version 6.0) with default parameters. Additionally, 280,437,292 Hi-C reads with a total size of 84.13 Gb was obtained. The raw contigs were split into segments of 50 Kb on average. The Hi-C reads were aligned to the segments using BWA (version 0.7.10) with default parameters [15]. The uniquely mapped reads were retained to assemble the genome using LACHESIS package [16].

Augustus (v3.3.2) was used to train and predict the gene features with Amel_HAv3.1 gene set [17,18]. Based on the deduced amino acid sequences, the annotation was performed through BLASTP against the non-redundant peptide database with the cut-off E-value at 10^{-5} and RPS-BLAST against the Conserved Domain Database at E-value at 10^{-3} [19]. Gene ontology analysis was performed using BLASTP against the InterProScan [20]. The Pathway was constructed based on the KEGG database [19].

2.3. The Genome Completeness and Honeybee Evolution

The protein sequences were used to query the BUSCO arthropod ortholog set to evaluate the genome completeness. The genomes of honeybee subspecies *A. mellifera mellifera* and *A. mellifera* DH4 were further queried to *A. mellifera carnica* using minimap2 [21]. The alignment files were viewed using package pafr (https://github.com/dwinter/pafr, accessed on 7 January 2021). The microsatellite was identified using MISA (microsatellite identification tool) with default parameters [22]. The microsatellite markers primers were blasted against the three genomes and the congruent was analyzed using Chi-squared test, R [23].

2.4. Phylogenetic Analysis of Social Genes

To investigate the molecular mechanisms underlying sociality between social and solitary bees, the protein sequences of four insulin family proteins *insulin receptor substrate 1* (IRS1), *insulin receptor substrate 4* (IRS4), *insulin receptor like* (IR-like) and *insulin-like peptides* (IPR) were retrieved from Eastern honeybees (*Apis cerana*), dwarf honeybees (*Apis florea*), giant honeybees (*Apis dorsata*), bumble bees (*Bombus terrestris*), digger bees (*Habropoda laboriosa*), red mason bees (*Osmia bicornis bicornis*) and small carpenter bees

(*Ceratina calcarata*) by NCBI blast with E-value cutoff $\leq 1 \times 10^{-5}$. The protein sequences were aligned using Muscle with default parameters by MEGA X package (Version 10.0.2) [24]. The phylogenetic tree was constructed using the Neighbor joining model with 1000 bootstraps and the small carpenter bees were used to root the tree.

2.5. Identifying Genome Selection Pressures

Orthologs between 7 species (*Nasonia vitripennis*, *Bombus terrestris*, *Bombus impatiens*, *Apis florea*, *Apis dorsata*, *Apis laboriosa*, *Apis cerana*) and 4 subspecies (*Apis mellifera carnica*, *Apis mellifera caucasica*, *Apis mellifera ligustica*, *Apis mellifera mellifera*) were discovered using Orthofinder (v2.5.2) [25]. To optimize the number of single-copy orthologs, we categorized them as such if at least 3 *Apis mellifera* subspecies and *Apis cerana* had a single-copy gene for the given orthogroup, culminating at 6328 single-copy ortholog families. Phylogenetic tree reconstruction, including all species described above, was undertaken by OrthoFinder. For each single-copy ortholog family, the longest protein isoforms for each of the species' gene were used in multiple sequence alignment with MAFFT (using local-pair algorithm and 1000 iterations) [26] and unreliably aligned residues and sequences were masked with GUIDANCE (v2.02) [27]. To optimize alignment length without gaps, we ran a maxalign script and removed subsequent sequences leading to more than 30% of gapped alignment as long as it did not result in the removal of any *A. mellifera* subspecies and *A. cerana* [28]. The protein sequences were replaced with coding sequences in the multiple alignments using the pal2nal script [29]. Furthermore, sequences containing a stop codon or having a length inconsistency between protein and DNA coding sequences (after removal of undefined bases) were filtered out. Alignments regions, where gapped positions were present, were removed with a custom python script, as these are the most problematic for positive selection inference [30,31]. Finally, CDS shorter than 100 nucleotides were eliminated [32].

Phylogenetic tests of positive selection in protein-coding genes usually contrast substitution rates at non-synonymous sites to substitution rates at synonymous sites taken as a proxy to neutral rates of evolution. The adaptive branch-site random effects model (aBSREL) from Hyphy software package was used to detect positive selection experienced by a gene family in a subset of sites in a specific branch of its phylogenetic tree [33]. The test for positive selection was run only on the branches leading to the origin of *A. mellifera* and on each *A. mellifera* subspecies. Results from the adaptive branch-site random effects model were corrected for multiple testing as one series using False Discovery Rate (FDR) and set up our significant threshold at 10% [34].

2.6. Test for Functional Category Enrichment

Gene Ontology (GO) annotations for our gene families were taken from Hymenoptera Genome database [35]. The enrichment of functional categories was evaluated with the package topGO version 2.4 of Bioconductor [36,37]. To identify functional categories enriched for genes under positive selection, strengthened, and relaxed selection pressure, the SUMSTAT test was used [38,39]. The SUMSTAT test is more sensitive than other methods and minimizes the rate of false positives [40–43]. To be able to use the distribution of log-likelihood ratios of the aBSREL and RELAX tests as scores in the SUMSTAT test, a fourth root transformation was used [39]. This transformation conserves the ranks of gene families [44]. Gene Ontology categories mapped to less than 10 genes were discarded. The list of significant gene sets resulting from enrichment tests is usually highly redundant. We therefore implemented the "elim" algorithm from the Bioconductor package topGO, to decorrelate the graph structure of the Gene Ontology. To account for multiple testing, the final list of p-values resulting from this test was corrected with the FDR and set up our significant threshold at 20%. To cluster the long list of significant functional categories, we used REVIGO with the SimRel semantic similarity algorithm and medium size (0.7) result list [45,46].

3. Results and Discussion

3.1. The Genome Assembly Statistics of A. mellifera carnica

A robust genome of 226.02 Mbp comprised of 313 contigs was assembled, which were further collapsed into 169 scaffolds (GCA_013841245.2) (Table 1). By aligning the predicted protein sequences to 1066 core arthropod Benchmarking Universal Single-copy orthologs (BUSCOs) [47], 93.53% of complete BUSCOs were identified. The results suggest that the assembled genome and predicted gene set were complete. Phylogenetic tree reconstruction revealed the subspeciation events of the European honeybees and the topology agreed with the evolution from solitary to social living (Figure 1) [48,49].

Table 1. Statistics of the studied *A. mellifera carnica* genome and other recently assembled honeybee genomes. NA indicates that the genome was assembled at contig level.

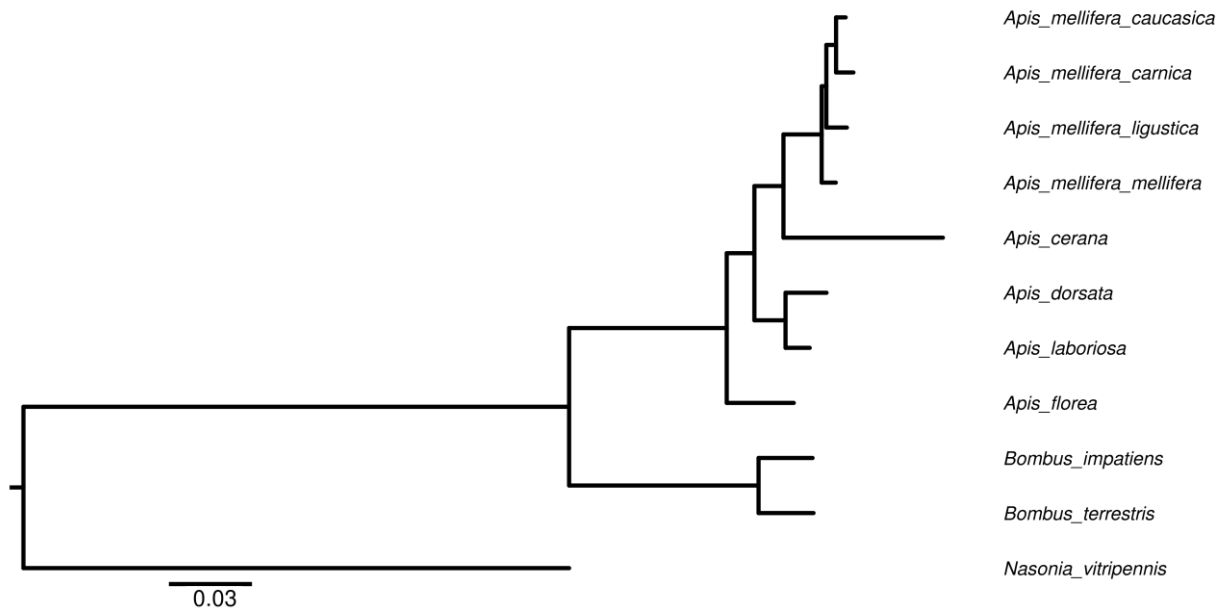|	*A. mellifera carnica*	*A. mellifera caucasica*	*A. mellifera* DH4	*A. mellifera mellifera*	*A. cerana cerana*
Technology	Pacbio, Hi-C	Pacbio	Pacbio, Hi-C	Pacbio, Hi-C	Pacbio, Hi-C
Coverage	101	112	192	100	134
Assembly size (Mbp)	226.02	224.7	225.25	227.03	215.6
Number of Contigs	313	224	227	199	214
N50 Scaffold size (Mbp)	13.4	NA	13.6	13.5	13.7
Releasing date	2020	2020	2018	2018	2020

Figure 1. Phylogenetic tree of the studies insects for the genome selection. All nodes were 100% bootstrap supported. *N. vitripennis* was used to root the tree. For *A. mellifera carnica*, 93.53% of complete BUSCOs were found, which suggests the assembly is complete.

3.2. Genome Alignment among Honeybee Subspecies

By pair-wised alignment, 83% and 89% of *A. mellifera carnica* genome can be perfectly aligned to *A. mellifera mellifera* and *A. mellifera* DH4 genome, respectively (Figures 2 and S1). The average length of the aligned region was 103 Kbp and 108 Kbp for *A. mellifera mellifera* and *A. mellifera* DH4, respectively (Figure S2). The inverted fragment may reflect natural structural variation among the genomes, reflecting local adaption [50,51]. Overall, 88,380 microsatellites with dinucleotides motif were identified in *A. mellifera carnica*. Comparatively, 89,099 and 89,569 were identified in *A. mellifera* DH4 and *A. mellifera mellifera*,

respectively. The relative abundance of microsatellites along the motifs were not significantly different among the three genomes (Pearson's Chi-squared test, df = 24, p = 0.26). However, the number of microsatellites decreased with the increasing number of repeats for all three genomes (Pearson's correlation coefficient, df = 9, p < 0.001, Figure 3A). A set of linkage map makers were further compared among the three genomes [52–54]. Out of 1081 paired microsatellite primers, 839 (77%) could be aligned to all three genomes (Figure S3), with an average density of 4.8 cM per locus (Figure 3B). For the remaining 242 markers, 162 were aligned to at least one genome. *A. mellifera* DH4 shared a higher number of markers with *A. mellifera carnica* compared with *A. mellifera mellifera*, which significantly deviated from random (Pearson's Chi-squared test, df = 2, p < 0.01) and is congruent with the genome phylogenetic tree in general (Figure S4).
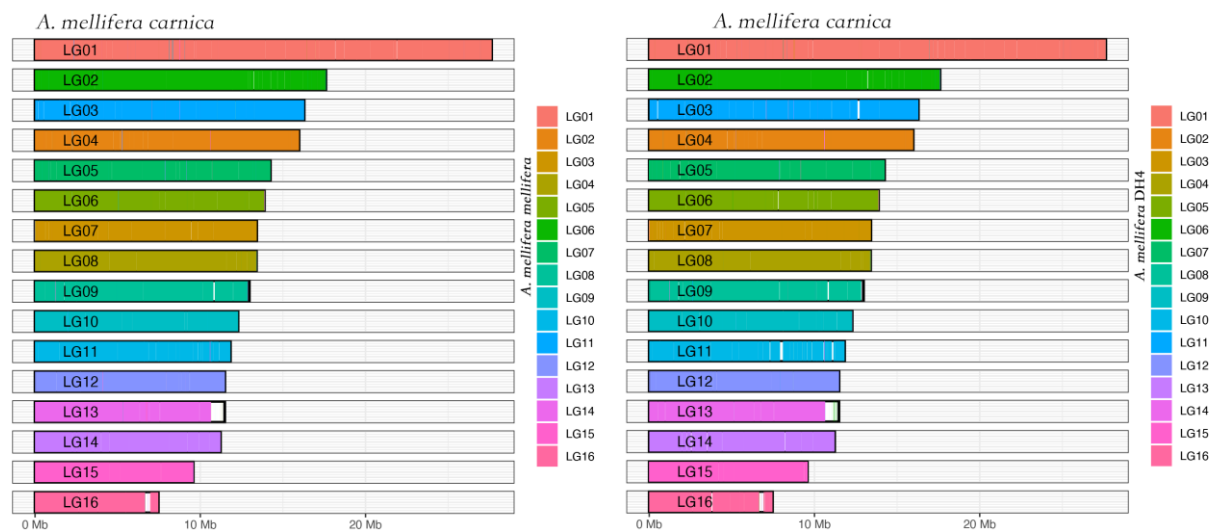

Figure 2. Coverage map of the honeybee subspecies. The genomes of *A. mellifera melllifera* and *A. mellifera* DH4 were compared with *A. mellifera carnica*. Overall, 80% of the nucleotides were aligned.
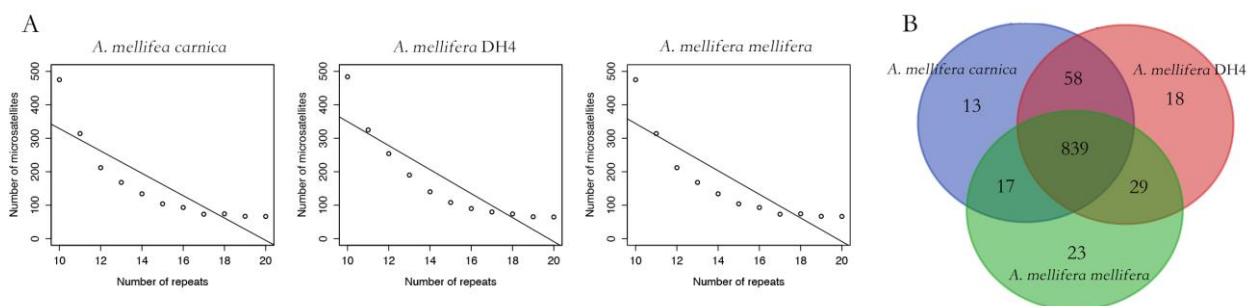

Figure 3. Microsatellite analysis of the three honeybee subspecies. (**A**): the distribution of microsatellites of the dinucleotide motif. The number of microsatellites decreased with the increasing number of repeats for all three genomes. Overall, the variation of the microsatellite distribution among the three genomes was not significant. (**B**): Venn diagram of the linkage map makers among the three honeybee genomes. *A. mellifera* DH4 shared significantly higher number of markers with *A. mellifera carnica* compared with *A. mellifera mellifera*.

3.3. Phylogenetic Analysis of Sociality-Related Proteins

The evolutionary process of bee sociality is fascinating, and highlights how genomes evolved to give rise to new and complex behaviors [49,55,56]. Insulin is an essential gene family regulating honeybee caste determination [57,58]. Hexamerin regulates the reproductive tissue development after honeybee caste differentiation [59–61]. The two gene families were selected to indicate the social evolution. The phylogenetic tree of insulin and hexamerin gene families clearly showed that solitary bees (digger bees and red mason

bees) were an early branch from the root (small carpenter bees), followed by bumble bees (Figure 4). The four honeybee species were clustered together, indicating that the a distinctive gene selection of sociality [62–64].
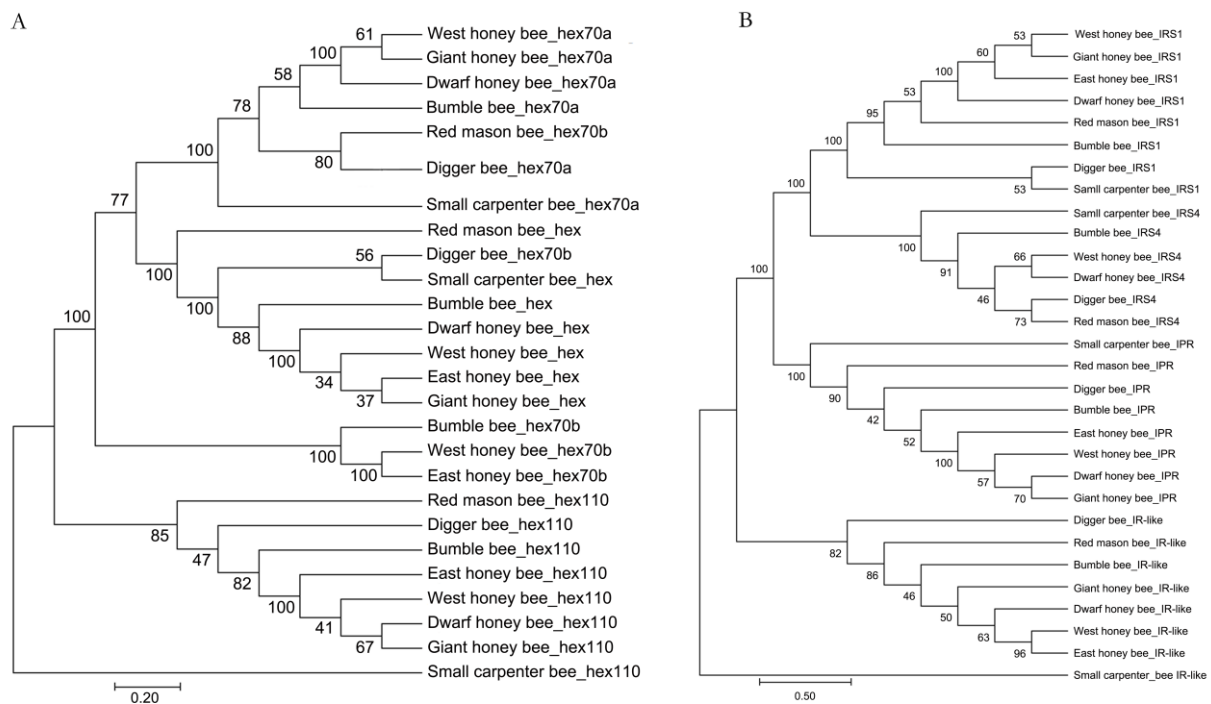

Figure 4. Phylogenetic tree of social genes. (**A**): phylogenetic tree of hexamerin family proteins. B: phylogenetic tree of insulin family. The orthologs were retrieved by BLAST in NCBI manually with E-value cutoff $\leq 1 \times 10^{-5}$. The sequences were aligned with Muscle and the phylogenetic tree was constructed using Neighbor joining model with 1000 bootstrap.

3.4. Signature of Positive Selection

Overall, 4897 single-copy orthologous groups were identified, out of which 245 orthologous groups displayed signs of positive selection in at least one branch test. The number of orthologous groups under positive selection within each branch varied significantly (Chi-squared test, $p < 2.2 \times 10^{-16}$) and ranged from 10 to 114 with 10, 27, 45, 78, and 114 for *A. mellifera caucasica*, *A. mellifera mellifera*, *A. mellifera* spp., *A. mellifera carnica*, and *A. mellifera ligustica*, respectively (Table S1). Such a variation in the number of genes under positive selection among the different *A. mellifera* subspecies may highlight a fast pace of adaptation or directed domestication in *A. mellifera carnica* and *A. mellifera ligustica* compared with others [65,66]. However, this result could also be the consequence of different population structures among *A. mellifera* subspecies with differential gene flow and introgression levels. Only one orthologous group, encoding for the protein obscurin involved in myogenesis and Hippo signaling pathway [67,68], was found to be under positive selection in three branches (*A. m. carnica*, *A. m. ligustica*, and *A. m.* spp., Table 2). Moreover, a few orthologous groups were found under positive selection in more than one branch (Table 2). Additionally, Hippo signaling pathway was involved in cold temperature adaptation in bees [13].

Table 2. Orthologous found to be under positive selection in more than one branch tested.

Gene/Protein Name	Putative Function	Branches
obscurin	Involved in myogenesis and the Hippo signaling pathway	*A. mellifera carnica* *A. mellifera ligustica* *A. mellifera* spp.
ATP-dependent RNA helicase ab-strakt	axonal growth (visual system)	*A. mellifera carnica* *A. mellifera ligustica*
bifunctional heparan sulfate N-deacetylase/N-sulfotransferase	involved in Wnt signaling	*A. mellifera carnica* *A. mellifera ligustica*
foxP protein	Central Nervous System/learning-memory	*A. mellifera carnica* *A. mellifera ligustica*
serine/threonine-protein kinase tri-corner	Post-transcriptional regulation, development (dendrite morphogenesis)	*A. mellifera carnica* *A. mellifera ligustica*
protein turtle homolog B	Establishing coordinated motor control, axonal targeting of the R7 photoreceptor	*A. mellifera carnica* *A. mellifera ligustica*
GTPase-activating Rap/Ran-GAP domain-like protein 3	signal transduction	*A. mellifera carnica* *A. mellifera ligustica*
zinc finger protein 512B	involved in transcriptional regulation	*A. mellifera carnica* *A. mellifera mellifera*
prohormone-2 precursor	neuropeptide/social behavior regulation	*A. mellifera carnica* *A. mellifera mellifera*
protein lingerer	copulation/short-term memory	*A. mellifera carnica* *A. mellifera mellifera*
retinoid-inducible serine carboxy-peptidase	involved in vascular wall	*A. mellifera carnica* *A. mellifera mellifera*
cAMP-specific 3′,5′-cyclic phos-phodiesterase	signaling, phsysiology, female fertility, learning/memory	*A. mellifera carnica* *A. mellifera mellifera*
7SK snRNA methylphosphate cap-ping enzyme	methylation, development	*A. mellifera carnica* *A. mellifera mellifera*
RB1-inducible coiled-coil protein 1	involved in autophagy	*A. mellifera carnica* *A. mellifera caucasica*
muscle LIM protein Mlp84B	cell differentiation late in myogenesis	*A. mellifera carnica* *A. mellifera* spp.
junctophilin-1	formation of junction membrane in sarcomere	*A. mellifera carnica* *A. mellifera* spp.
RNA binding protein fox-1 homolog 2	alternative splicing	*A. mellifera ligustica* *A. mellifera mellifera*
chitin synthase chs-2	chitin biosynthetic synthesis	*A. mellifera ligustica* *A. mellifera* spp.
ell-associated factor Eaf	regulation of transcription elongation	*A. mellifera ligustica* *A. mellifera* spp.
nuclear protein localization protein 4 homolog	ubiquitination	*A. mellifera ligustica* *A. mellifera* spp.
kinesin-like protein unc-104	synaptic vesicle transport/locomotion	*A. mellifera ligustica* *A. mellifera* spp.
actin-interacting protein 1	sarcomere organization/development/locomotion	*A. mellifera ligustica* *A. mellifera* spp.
ubiquitin carboxyl-terminal hydro-lase 36	stem cell line maintenance	*A. mellifera ligustica* *A. mellifera* spp.
teneurin-a	neural development	*A. mellifera ligustica* *A. mellifera* spp.
troponin I	muscle contraction	*A. mellifera ligustica* *A. mellifera* spp.

3.5. Functional Categories Enriched of Positively Selected Genes

We identified 11 significant functional categories in *A. m. carnica*, 28 in *A. m. ligustica*, 5 in *A. m. mellifera* and *caucasica*, and 6 in *A. mellifera* branch, which were enriched in genes under positive selection at 20% FDR. The long list of significant GO-terms found to be significantly enriched of positively selected genes in *A. m. ligustica* were mainly related to larval development (Figure 5), as demonstrated with clustering from REVIGO [45]. Interestingly, the two most significant enriched functional genes under positive selection in *A. mellifera* branch, chitin metabolism and mitochondrial translation (Figure 6), matched functional genes previously found in *A. mellifera* [34]. While functions found to be enriched of positively selected genes in *A. m. mellifera* were mainly related to the nervous system, in *A. m. caucasica.* They were mainly related to autophagy/cell death (Figure 6). In *A. m. carnica*, most of significant GO terms seems to be linked with stress tolerance, notably response to wounding, reactive oxygen species (ROS) metabolism, and larval midgut programmed cell death (Figure 6). More precisely, several significant GO terms are likely involved in cold resistance in honeybees, such as developmental growth [69], ROS metabolism [70], and hippo signaling [13]. Interestingly, the gene encoding the protein Dachsous, which was under positive selection in *A. m. carnica* and found in several significant GO terms, plays a key role in the adaptation to temperate climate in the *A. m. sinisxinyuan* [13]. Furthermore, the genes *Amel_mTOR* and *Amel_Imp*, encoding serine/threonine-protein kinase mTOR and insulin-like growth factor 2 mRNA-binding protein 1, respectively, might also play a role in cold tolerance as they both are involved in the insulin pathway regulating food intake, essential for cold tolerance [71,72]. Moreover, cold tolerance in *A. cerana* involved serine/threonine-protein kinases like *mTOR* [73]. The protein RB1-inducible coiled-coil protein 1, part of the GO term 'larval midgut cell programmed cell death', was found to be under positive selection in *A. m. caucasica* and involved in cold adaptation in amphipods [74].
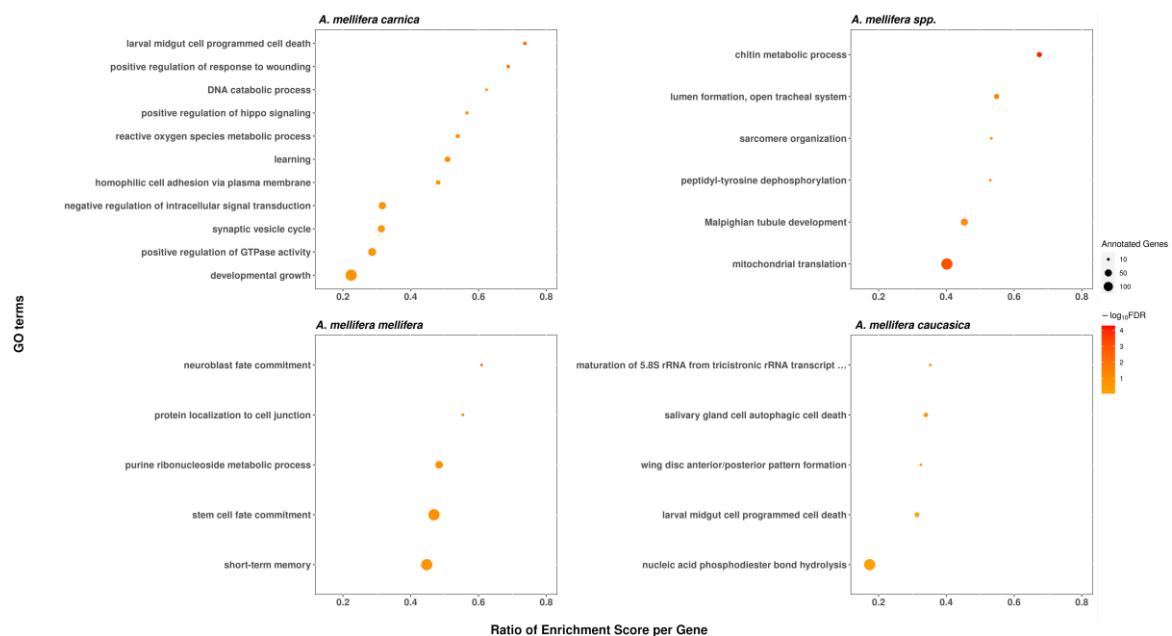

Figure 5. Functional categories significantly enriched of genes under positive selection in the branch *A. m. ligustica*, using REVIGO to cluster the long list of significant GO terms. The medium cluster size was 0.7 and the semantic similarity measure SimRel.
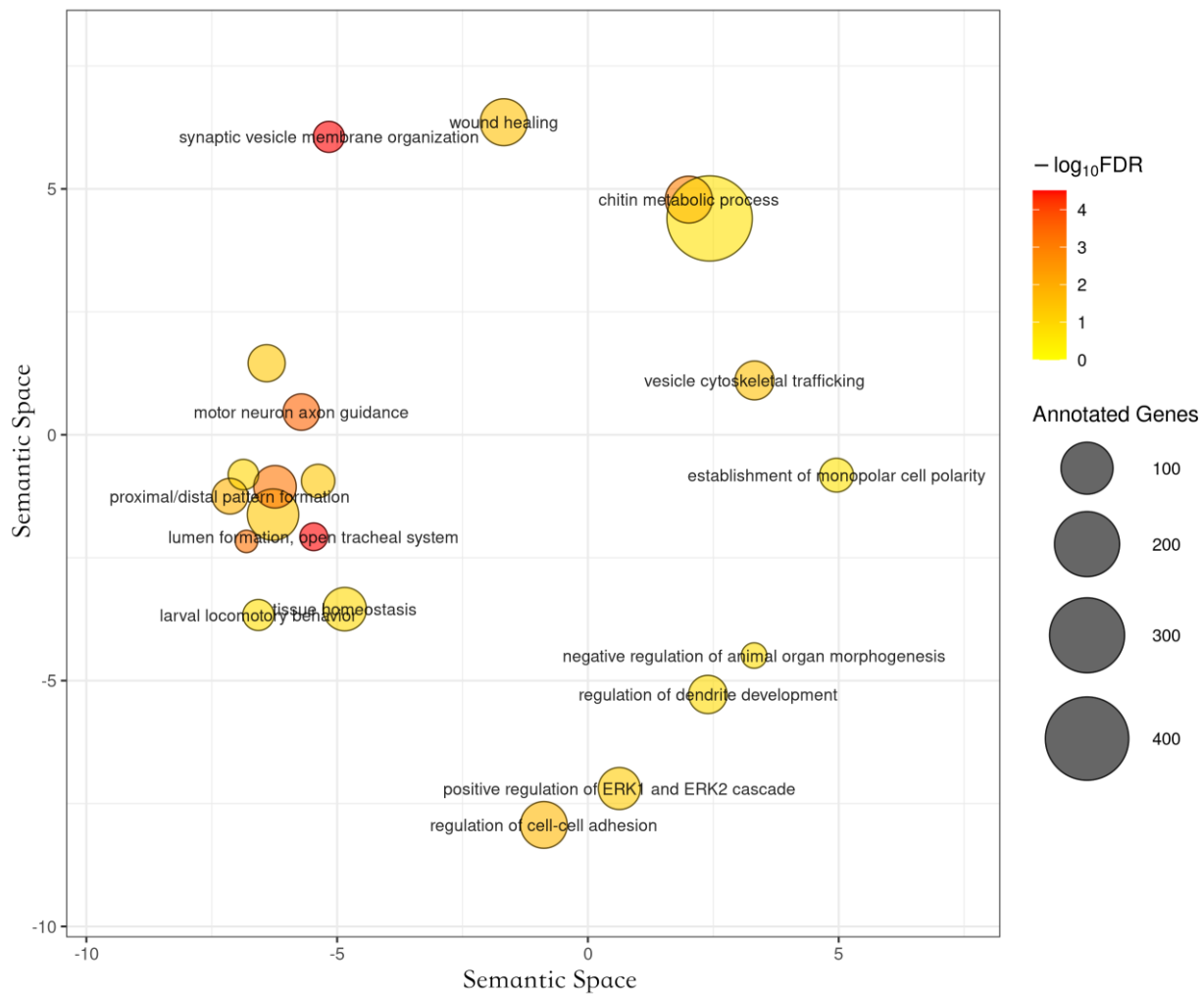
Figure 6. Functional categories significantly enriched of genes under positive selection in the branch *A. mellifera* spp.

Author Contributions: Z.-J.Z. conceived the experiment, Q.-S.N. collected the sample, Y.-Q.Z. and H.-J.Z. assembled the genome, Q.H. analyzed genome completeness, B.F. analyzed genome selection, X.-J.H. analyzed the phylogeny of sociality-related proteins, Q.H., Y.-Q.Z., B.F., X.-J.H., Q.-S.N., H.-J.Z. and Z.-J.Z. organized the manuscript. All authors have read and agreed to the published version of the manuscript.

Reference

1.　IPBES. *The Assessment Report of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services on Pollinators, Pollination and Food Production*; Potts, S.G., Imperatriz-Fonseca, V.L., Ngo, H.T., Eds.; Secretariat of the Intergovernmen: Kuala Lumpur, Malaysia, 2016.

2.　Bauer, D.M.; Sue Wing, I. The macroeconomic cost of catastrophic pollinator declines. *Ecol. Econ.* **2016**, *126*, 1–13.

3.　Parreño, M.A.; Alaux, C.; Brunet, J.-L.; Buydens, L.; Filipiak, M.; Henry, M.; Keller, A.; Klein, A.-M.; Kuhlmann, M.; Leroy, C.; et al. Critical links between biodiversity and health in wild bee conservation. *Trends Ecol. Evol.* **2022**, *37*, 309–321.

4.　Flores, J.M.; Gámiz, V.; Jiménez-Marín, Á.; Flores-Cortés, A.; Gil-Lebrero, S.; Garrido, J.J.; Hernando, M.D. Impact of *Varroa destructor* and associated pathologies on the colony collapse disorder affecting honeybees. *Res. Vet. Sci.* **2021**, *135*, 85–95.

5.　Anderson, D.; East, I.J. The latest buzz about colony collapse disorder. *Science* **2008**, *319*, 724–725.

6.　Van Engelsdorp, D.; Hayes, J.; Underwood, R.M.; Pettis, J. A Survey of Honeybee Colony Losses in the U.S., Fall 2007 to Spring 2008. *PLoS ONE* **2008**, *3*, e4071.

7.　Cox-Foster, D.L.; Conlan, S.; Holmes, E.; Palacios, G.; Evans, J.D.; Moran, N.A.; Quan, P.L.; Briese, T.; Hornig, M.; Geiser, D.M. A metagenomic survey of microbes in honeybee colony collapse disorder. *Science* **2007**, *318*, 283–287.

8.　Flores, J.M.; Gil-Lebrero, S.; Gámiz, V.; Rodríguez, M.I.; Ortiz, M.A.; Quiles, F.J. Effect of the climate change on honeybee colonies in a temperate Mediterranean zone assessed through remote hive weight monitoring system in conjunction with exhaustive colonies assessment. *Sci. Total Environ.* **2019**, *653*, 1111–1119.

9.　Kurze, C.; Le Conte, Y.; Dussaubat, C.; Erler, S.; Kryger, P.; Lewkowski, O.; Muller, T.; Widder, M.; Moritz, R.F.A. *Nosema* Tolerant Honeybees (*Apis mellifera*) Escape Parasitic Manipulation of Apoptosis. *PLoS ONE* **2015**, *10*, e0140174.

10.　Kurze, C.; Routtu, J.; Moritz, R.F.A. Parasite resistance and tolerance in honeybees at the individual and social level. *Zoology* **2016**, *119*, 290–297.

11.　Thaduri, S.; Stephan, J.G.; de Miranda, J.R.; Locke, B. Disentangling host-parasite-pathogen interactions in a *Varroa*-resistant honeybee population reveals virus tolerance as an independent, naturally adapted survival mechanism. *Sci. Rep.* **2019**, *9*, 6221.

12.　Li, X.; Ma, W.; Shen, J.; Long, D.; Feng, Y.; Su, W.; Xu, K.; Du, Y.; Jiang, Y. Tolerance and response of two honeybee species *Apis cerana* and *Apis mellifera* to high temperature and relative humidity. *PLoS ONE* **2019**, *14*, e0217921.

13.　Chen, C.; Liu, Z.; Pan, Q.; Chen, X.; Wang, H.; Guo, H.; Liu, S.; Lu, H.; Tian, S.; Li, R.; et al. Genomic Analyses Reveal Demographic History and Temperate Adaptation of the Newly Discovered Honeybee Subspecies *Apis mellifera sinisxinyuan* n. ssp. *Mol. Biol. Evol.* **2016**, *33*, 1337–1348.

14.　Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **2018**, *34*, i884–i890.

15.　Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760.

16.　Burton, J.N.; Adey, A.; Patwardhan, R.P.; Qiu, R.; Kitzman, J.O.; Shendure, J. Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. *Nat. Biotechnol.* **2013**, *31*, 1119–1125.

17.　Stanke, M.; Keller, O.; Gunduz, I.; Hayes, A.; Waack, S.; Morgenstern, B. AUGUSTUS: Ab initio prediction of alternative transcripts. *Nucleic Acids Res.* **2006**, *34*, 435–439.

18.　Wallberg, A.; Bunikis, I.; Pettersson, O.V.; Mosbech, M.-B.; Childers, A.K.; Evans, J.D.; Mikheyev, A.S.; Robertson, H.M.; Robinson, G.E.; Webster, M.T. A hybrid de novo genome assembly of the honeybee, *Apis mellifera*, with chromosome-length scaffolds. *BMC Genom.* **2019**, *20*, 275.

19.　Marchler-Bauer, A.; Lu, S.; Anderson, J.B.; Chitsaz, F.; Derbyshire, M.K.; DeWeese-Scott, C.; Fong, J.H.; Geer, L.Y.; Geer, R.C.; Gonzales, N.R.; et al. CDD: A Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* **2011**, *39*, D225–D229.

20.　Mulder, N.J.; Apweiler, R. The InterPro Database and Tools for Protein Domain Analysis. *Curr. Protoc. Bioinforma.* **2008**, *21*, 2–7.

21.　Li, H. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **2018**, *34*, 3094–3100.

22.　Beier, S.; Thiel, T.; Münch, T.; Scholz, U.; Mascher, M. MISA-web: A web server for microsatellite prediction. *Bioinformatics* **2017**, *33*, 2583–2585.

23.　R Core Team. *R: A Language and Environment for Statistical computing*; R Foundations for Statistical Computing: Vienna, Austria, 2013. Available online: http://www.R-project.org/ (accessed on 4 June 2021).

24. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797.
25. Emms, D.M.; Kelly, S. OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **2019**, *20*, 238.
26. Rozewicki, J.; Li, S.; Amada, K.M.; Standley, D.M.; Katoh, K. MAFFT-DASH: Integrated protein sequence and structural alignment. *Nucleic Acids Res.* **2019**, *47*, W5–W10.
27. Sela, I.; Ashkenazy, H.; Katoh, K.; Pupko, T. GUIDANCE2: Accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. *Nucleic Acids Res.* **2015**, *43*, W7–W14.
28. Gouveia-Oliveira, R.; Sackett, P.W.; Pedersen, A.G. MaxAlign: Maximizing usable data in an alignment. *BMC Bioinform.* **2007**, *8*, 312.
29. Suyama, M.; Torrents, D.; Bork, P. PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* **2006**, *34*, W609–W612.
30. Fletcher, W.; Yang, Z. The effect of insertions, deletions, and alignment errors on the branch-site test of positive selection. *Mol. Biol. Evol.* **2010**, *27*, 2257–2267.
31. Markova-Raina, P.; Petrov, D. High sensitivity to aligner and high rate of false positives in the estimates of positive selection in the 12 *Drosophila genomes*. *Genome Res.* **2011**, *21*, 863–874.
32. Hambuch, T.M.; Parsch, J. Patterns of synonymous codon usage in *Drosophila melanogaster* genes with sex-biased expression. *Genetics* **2005**, *170*, 1691–1700.
33. Pond, S.L.K.; Frost, S.D.W.; Muse, S. V HyPhy: Hypothesis testing using phylogenies. *Bioinformatics* **2005**, *21*, 676–679.
34. Fouks, B.; Brand, P.; Nguyen, H.N.; Herman, J.; Camara, F.; Ence, D.; Hagen, D.E.; Hoff, K.J.; Nachweide, S.; Romoth, L.; et al. The genomic basis of evolutionary differentiation among honeybees. *Genome Res.* **2021**, *31*, 1203–1215.
35. Walsh, A.T.; Triant, D.A.; Le Tourneau, J.J.; Shamimuzzaman, M.; Elsik, C.G. Hymenoptera Genome Database: New genomes and annotation datasets for improved go enrichment and orthologue analyses. *Nucleic Acids Res.* **2022**, *50*, D1032–D1039.
36. Mi, H.; Muruganujan, A.; Ebert, D.; Huang, X.; Thomas, P.D. PANTHER version 14: More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.* **2019**, *47*, D419–D426.
37. Alexa, A.; Rahnenfuhrer, J. *TopGO: Enrichment Analysis for Gene Ontology*, R Package Version 2.46.0; R Foundations for Statistical Computing: Vienna, Austria, 2021. Available online: https://bioconductor.org/packages/release/bioc/html/topGO.html (accessed on 31 August 2022).
38. Efron, B.; Tibshirani, R. On testing the significance of sets of genes. *Ann. Appl. Stat.* **2007**, *1*, 107–129.
39. Roux, J.; Privman, E.; Moretti, S.; Daub, J.T.; Robinson-Rechavi, M.; Keller, L. Patterns of positive selection in seven ant genomes. *Mol. Biol. Evol.* **2014**, *31*, 1661–1685.
40. Ackermann, M.; Strimmer, K. A general modular framework for gene set enrichment analysis. *BMC Bioinform.* **2009**, *10*, 47.
41. Daub, J.T.; Hofer, T.; Cutivet, E.; Dupanloup, I.; Quintana-Murci, L.; Robinson-Rechavi, M.; Excoffier, L. Evidence for polygenic adaptation to pathogens in the human genome. *Mol. Biol. Evol.* **2013**, *30*, 1544–1558.
42. Fehringer, G.; Liu, G.; Briollais, L.; Brennan, P.; Amos, C.I.; Spitz, M.R.; Bickeböller, H.; Wichmann, H.E.; Risch, A.; Hung, R.J. Comparison of pathway analysis approaches using lung cancer GWAS data sets. *PLoS ONE* **2012**, *7*, e31816.
43. Tintle, N.L.; Borchers, B.; Brown, M.; Bekmetjev, A. Comparing gene set analysis methods on single-nucleotide polymorphism data from Genetic Analysis Workshop 16. *BMC Proc.* **2009**, *3*, S96.
44. Canal, L. A normal approximation for the chi-square distribution. *Comput. Stat. Data Anal.* **2005**, *48*, 803–808.
45. Supek, F.; Bošnjak, M.; Škunca, N.; Šmuc, T. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLoS ONE* **2011**, *6*, e21800.
46. Rimal, R.; Almøy, T.; Sæbø, S. A tool for simulating multi-response linear model data. *Chemom. Intell. Lab. Syst.* **2018**, *176*, 1–10.
47. Seppey, M.; Manni, M.; Zdobnov, E.M. *BUSCO: Assessing Genome Assembly and Annotation Completeness BT—Gene Prediction: Methods and Protocols*; Kollmar, M., Ed.; Springer: New York, NY, USA, 2019; pp. 227–245; ISBN 978-1-4939-9173-0.
48. Peters, R.S.; Krogmann, L.; Mayer, C.; Donath, A.; Gunkel, S.; Meusemann, K.; Kozlov, A.; Podsiadlowski, L.; Petersen, M.; Lanfear, R.; et al. Evolutionary History of the Hymenoptera. *Curr. Biol.* **2017**, *27*, 1013–1018.
49. Kapheim, K.M.; Pan, H.; Li, C.; Salzberg, S.L.; Puiu, D.; Magoc, T.; Robertson, H.M.; Hudson, M.E.; Venkat, A.; Fischman, B.J.; et al. Social evolution. Genomic signatures of evolutionary transitions from solitary to group living. *Science* **2015**, *348*, 1139–1143.
50. Tichkule, S.; Cacciò, S.M.; Robinson, G.; Chalmers, R.M.; Mueller, I.; Emery-Corbin, S.J.; Eibach, D.; Tyler, K.M.; van Oosterhout, C.; Jex, A.R. Global Population Genomics of Two Subspecies of *Cryptosporidium hominis* during 500 Years of Evolution. *Mol. Biol. Evol.* **2022**, *39*, msac056.
51. Wallberg, A.; Schöning, C.; Webster, M.T.; Hasselmann, M. Two extended haplotype blocks are associated with adaptation to high altitude habitats in East African honeybees. *PLOS Genet.* **2017**, *13*, e1006792.
52. Solignac, M.; Mougel, F.; Vautrin, D.; Monnerot, M.; Cornuet, J.-M. A third-generation microsatellite-based linkage map of the honey bee, *Apis mellifera*, and its comparison with the sequence-based physical map. *Genome Biol.* **2007**, *8*, R66.
53. Behrens, D.; Huang, Q.; Geßner, C.; Rosenkranz, P.; Frey, E.; Locke, B.; Moritz, R.F.A.; Kraus, F.B. Three QTL in the honey bee *Apis mellifera* L. suppress reproduction of the parasitic mite *Varroa destructor*. *Ecol. Evol.* **2011**, *1*, 451–458.
54. Huang, Q.; Kryger, P.; Le Conte, Y.; Lattorff, H.M.G.; Kraus, F.B.; Moritz, R.F.A. Four quantitative trait loci associated with low *Nosema ceranae* (Microsporidia) spore load in the honeybee *Apis mellifera*. *Apidologie* **2014**, *45*, 248–256.

55. Saleh, N.W.; Ramírez, S.R. Sociality emerges from solitary behaviours and reproductive plasticity in the orchid bee *Euglossa dilemma*. *Proc. R. Soc. B Biol. Sci.* **2019**, *286*, 20190588.

56. Shell, W.A.; Steffen, M.A.; Pare, H.K.; Seetharam, A.S.; Severin, A.J.; Toth, A.L.; Rehan, S.M. Sociality sculpts similar patterns of molecular evolution in two independently evolved lineages of eusocial bees. *Commun. Biol.* **2021**, *4*, 253.

57. de Azevedo, S.V.; Hartfelder, K. The insulin signaling pathway in honeybee (*Apis mellifera*) caste development—Differential expression of insulin-like peptides and insulin receptors in queen and worker larvae. *J. Insect Physiol.* **2008**, *54*, 1064–1071.

58. Wang, Y.; Azevedo, S.V.; Hartfelder, K.; Amdam, G. V Insulin-like peptides (AmILP1 and AmILP2) differentially affect female caste development in the honeybee (*Apis mellifera* L.). *J. Exp. Biol.* **2013**, *216*, 4347–4357.

59. Lago, D.C.; Humann, F.C.; Barchuk, A.R.; Abraham, K.J.; Hartfelder, K. Differential gene expression underlying ovarian phenotype determination in honeybee, *Apis mellifera* L., caste development. *Insect Biochem. Mol. Biol.* **2016**, *79*, 1–12.

60. Martins, J.R.; Nunes, F.M.F.; Cristino, A.S.; Simões, Z.L.P.; Bitondi, M.M.G. The four hexamerin genes in the honey bee: Structure, molecular evolution and function deduced from expression patterns in queens, workers and drones. *BMC Mol. Biol.* **2010**, *11*, 23.

61. Martins, J.R.; Morais Franco Nunes, F.; Luz Paulino Simões, Z.; Maria Gentile Bitondi, M. A honeybee storage protein gene, hex 70a, expressed in developing gonads and nutritionally regulated in adult fat body. *J. Insect Physiol.* **2008**, *54*, 867–877.

62. Page, R.E., Jr.; Scheiner, R.; Erber, J.; Amdam, G.V. The development and evolution of division of labor and foraging specialization in a social insect (*Apis mellifera* L.). *Curr. Top. Dev. Biol.* **2006**, *74*, 253–286.

63. Wittwer, B.; Hefetz, A.; Simon, T.; Murphy, L.E.; Elgar, M.A.; Pierce, N.E.; Kocher, S.D. Solitary bees reduce investment in communication compared with their social relatives. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 6569–6574.

64. Smith, A.; Simons, M.; Bazarko, V.; Seid, M. The influence of sociality, caste, and size on behavior in a facultatively eusocial bee. *Insectes Soc.* **2019**, *66*, 153–163.

65. Oldroyd, B.P. Domestication of honeybees was associated with expansion of genetic diversity. *Mol. Ecol.* **2012**, *21*, 4409–4411.

66. Wallberg, A.; Han, F.; Wellhagen, G.; Dahle, B.; Kawata, M.; Haddad, N.; Simões, Z.L.P.; Allsopp, M.H.; Kandemir, I.; De la Rúa, P.; et al. A worldwide survey of genome sequence variation provides insight into the evolutionary history of the honeybee *Apis mellifera*. *Nat. Genet.* **2014**, *46*, 1081–1088.

67. Yamauchi, T.; Moroishi, T. Hippo Pathway in Mammalian Adaptive Immune System. *Cells* **2019**, *8*, 398.

68. Ma, S.; Meng, Z.; Chen, R.; Guan, K.-L. The Hippo Pathway: Biology and Pathophysiology. *Annu. Rev. Biochem.* **2019**, *88*, 577–604.

69. Ramirez, L.; Luna, F.; Mucci, C.A.; Lamattina, L. Fast weight recovery, metabolic rate adjustment and gene-expression regulation define responses of cold-stressed honeybee brood. *J. Insect Physiol.* **2021**, *128*, 104178.

70. Mucci, C.A.; Ramirez, L.; Giffoni, R.S.; Lamattina, L. Cold stress induces specific antioxidant responses in honeybee brood. *Apidologie* **2021**, *52*, 596–607.

71. Azevedo, S.V.; Caranton, O.A.M.; de Oliveira, T.L.; Hartfelder, K. Differential expression of hypoxia pathway genes in honeybee (*Apis mellifera L.*) caste development. *J. Insect Physiol.* **2011**, *57*, 38–45.

72. Liu, N.; Ren, Z.; Ren, Q.; Chang, Z.; Li, J.; Li, X.; Sun, Z.; He, J.; Niu, Q.; Xing, X. Full length transcriptomes analysis of cold-resistance of *Apis cerana* in Changbai Mountain during overwintering period. *Gene* **2022**, *830*, 146503.

73. Xu, K.; Niu, Q.; Zhao, H.; Du, Y.; Jiang, Y. Transcriptomic analysis to uncover genes affecting cold resistance in the Chinese honeybee (*Apis cerana cerana*). *PLoS ONE* **2017**, *12*, e0179922.

74. Lipaeva, P.; Vereshchagina, K.; Drozdova, P.; Jakob, L.; Kondrateva, E.; Lucassen, M.; Bedulina, D.; Timofeyev, M.; Stadler, P.; Luckenbach, T. Different ways to play it cool: Transcriptomic analysis sheds light on different activity patterns of three amphipod species under long-term cold exposure. *Mol. Ecol.* **2021**, *30*, 5735–5751.