The origin of the odorant receptor gene family in insects

Philipp Brand[1,*], Hugh M. Robertson[2,*], Wei Lin[3], Ratnasri Pothula[4], William E. Klingeman[5], Juan Luis Jurat-Fuentes[4], and Brian R. Johnson[3]

[1] Department of Evolution and Ecology, Center for Population Biology, University of California, Davis, Davis, CA 95616, USA

[2] Department of Entomology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[3] Department of Entomology and Nematology, University of California, Davis, Davis, CA 95616, USA

[4] Department of Entomology and Plant Pathology, University of Tennessee, Knoxville TN 37996, USA

[5] Department of Plant Sciences, University of Tennessee, Knoxville TN 37996, USA

* Shared first authorship

Corresponding authors: Philipp Brand, pbrand@ucdavis.edu, and Hugh Robertson, hughrobe@uiuc.edu

Abstract

The sense of smell enables the detection and discrimination of airborne chemicals via chemosensory receptors that have evolved independently multiple times throughout the tree of life. In insects, the odorant receptor (OR) gene family is the major chemosensory gene family involved in olfaction and its origin has been hypothesized to coincide with the evolution of a terrestrial lifestyle in hexapods. Missbach et al. (2014) challenged this view and suggested that ORs evolved with an ancestral OR co-receptor (Orco) after the origin of terrestriality, hypothesizing that the OR/Orco system is an adaptation to winged flight in insects instead. Building upon this work, we investigated the genomes of basal hexapod and insect lineages including Collembola, Diplura, Archaeognatha, Zygentoma, Odonata, and Ephemeroptera in an effort to identify the origin of the insect OR gene family. While absent from all non-insect hexapod lineages analyzed, ORs are present in all insect genomes. Orco is absent only in the most ancient insect lineage Archaeognatha. A fully functional OR/Orco system was present in our newly generated genome data of the Zygentoma *Thermobia domestica*. We suggest that ORs did evolve as adaptation to a terrestrial lifestyle outside high-humidity habitats, and not winged flight, representing a key evolutionary novelty in the ancestor of all insects. The OR family is therefore the first known molecular synapomorphy for the Class Insecta.

Introduction

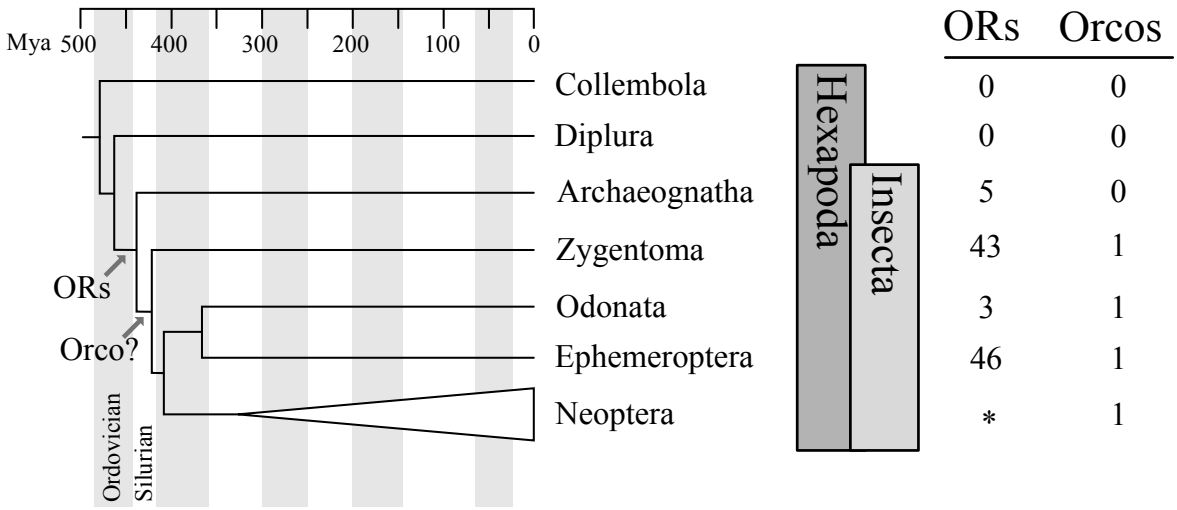
From bacteria to mammals, living organisms of all levels of complexity have evolved chemosensory receptors to detect and discriminate chemicals in the environment (Wuichet and Zhulin, 2010; Hansson and Stensmyr, 2011). The largest metazoan gene families for example encode tens to hundreds of odorant receptors (ORs) that interact with volatile chemicals at the sensory periphery underlying the sense of smell (Sánchez-Gracia et al., 2009; Niimura et al., 2014). OR gene families have evolved multiple times throughout the metazoans, including independent origins in vertebrates, nematodes, and insects (Hansson and Stensmyr, 2011). In insects, the OR gene family evolved from within the ancestral gustatory receptor (GR) gene family (Scott et al., 2001; Robertson et

al., 2003) that extends back to ancient metazoan lineages (Robertson, 2015; Saina et al., 2015; Eyun et al., 2017). ORs are absent from non-insect arthropod genomes (Peñalva-Arana et al., 2009; Almeida et al., 2015; Gulia-Nuss et al., 2016; Ngoc et al., 2016; Eyun et al., 2017), and have been hypothesized to have evolved concomitant with the evolution of terrestriality in hexapods (Robertson et al., 2003).

The lack of molecular resources for early branching hexapod and insect lineages has prevented the precise dating of the origin of insect ORs. Only recently, whole-genome sequencing efforts suggested that ORs are absent in non-insect hexapods such as Collembola (Wu et al., 2017) but present in early branching pterygote insects such as damselflies (Odonata; Ioannidis et al., 2017). Efforts to understand more precisely the origin of the OR gene family within hexapods were greatly advanced by the findings of Missbach et al. (2014) who sequenced transcriptomes of the chemosensory organs of two apterygote insects, the bristletail *Lepismachilis y-signata* (Archaeognatha) and the firebrat *Thermobia domestica* (Zygentoma). They identified three ORs in the firebrat, which they named TdomOrco1-3, with apparent similarity to the neopteran odorant receptor co-receptor (Orco; Vosshall and Hansson, 2011). Orco is a highly conserved single-copy gene present in all other insects studied to date and encodes a protein that is a partner with each of the other "specific" ORs (Benton et al., 2006) which is required for OR-based olfaction in insects (Larsson et al., 2004). In contrast, Missbach et al. (2014) could not find ORs or Orco relatives in their bristletail transcriptome, instead finding only members of the ionotropic receptor (IR) gene family. Given evidence that IRs serve olfactory roles in terrestrial crustaceans and insects (Rytz et al., 2013; Groh-Lunow et al., 2015; Rimal and Lee, 2018), they argued that olfaction in early-branching terrestrial hexapods and apterygote insects is entirely IR-dependent, with Orco evolving as ancestral OR from the GR lineage between the Archaeognatha and Zygentoma. Based on these findings, Missbach et al. (2014) suggested that the Orco/OR system evolved together with flight in pterygote insects and left off with the observation that "the existence of three Orco types remains mysterious".

93 Recently, phylogenetic analysis of the OR gene family of the damselfly *Calopteryx*
94 *splendens* suggested that at least one of the three Orco-like ORs from *T. domestica*,
95 TdomOrco3, might be a specific OR instead of an Orco (Ioannidis et al., 2017). If this is
96 correct, then the entire Orco/OR system evolved before winged insects, which would
97 explain the "mystery" of three apparent Orco types in Zygentoma. In an effort to identify
98 the origin of the insect OR gene family and the Orco/OR system, we investigated the
99 genome sequences of species belonging to multiple ancient terrestrial hexapod and insect
100 orders, including Collembola (springtails), Diplura (two-pronged bristletails),
101 Archaeognatha (jumping bristletails), Zygentoma (silverfish and firebrats), Odonata
102 (damselflies and dragonflies), and Ephemeroptera (mayflies).
103
104 **Results and Discussion**
105
106 **ORs were present in the ancestor of insects**
107
108 We detected no ORs in two non-insect hexapod lineages, Collembola and Diplura (a
109 genome sequence is not available for the third lineage, the Protura), despite extensive
110 annotation efforts. In contrast, we identified genes with similarity to known insect ORs in
111 all other genomes investigated (Figure 1; details in Supplemental Material, Tables S2 and
112 S3). These included one species of each of the earliest branching insect lineages, the
113 Archeognatha, Zygentoma, Ephemeroptera, and Odonata (Misof et al., 20014).
114 Accordingly, ORs were likely present in the ancestor of all insects but absent from all
115 non-insect hexapod lineages. This suggests that the origin of the OR gene family
116 coincided with the evolution of insects. Thus, our analysis does not support the
117 hypothesis that ORs evolved with the evolution of winged flight in insects (Missbach et
118 al., 2014) but is compatible with the hypothesis that they evolved with terrestriality in
119 insects (Robertson et al. 2003). Terrestrial hexapod lineages without ORs (*i.e.* Diplura
120 and Collembola) are confined to highly humid environments including moist soil and
121 leaf-litter. This is analogous to terrestrial crustaceans, which evolved OR-independent
122 modified sensory systems allowing the detection of airborne odors on land (Stensmyr et
123 al., 2005; Hansson et al., 2010). While terrestrial crustaceans employ the same olfactory

124 gene-families as their marine ancestors in combination with anatomical adaptations to
125 terrestrial olfaction, their sense of smell is highly dependent on humidity (Krång et al.,
126 2012). Although the molecular and physiological basis of dipluran and collembolan
127 olfaction is unexplored, they are known to have a highly sensitive sense of smell used in
128 pheromone communication and foraging behavior (Verhoef et al., 1977; Purrington et al.,
129 1991; Staaden et al., 2011). It is thus possible that ORs evolved as an adaptation to a
130 terrestrial lifestyle outside ancestrally humid environments in insects.
131



132
133 **Figure 1 Origin of the insect odorant receptor gene family.** The number of ORs and OR co-receptors (Orcos) for all
134 ancient insect and non-insect orders analyzed was mapped on the hexapod phylogeny sensu Misof et al. (2014). ORs
135 are present in all insects but absent from non-insect hexapod genomes, and thus likely represent a molecular
136 synapomorphy for the Clade Insecta. Orco is present in all but Archaeognatha, the earliest branching insect order. This
137 suggests two scenarios including either the loss of Orco in Archaeognatha or an Orco origin following the evolution of
138 ORs (as indicated). The OR gene family likely evolved as an adaptation to a terrestrial lifestyle outside ancestrally
139 humid environments in insects. *: The genomes of all neopteran insects analyzed to date encode ORs, ranging from 10
140 ORs in head lice (Kirkness et al., 2010) to more than 300 ORs in ants (Smith et al., 2011a; b).

141

142 **The *Thermobia domestica* genome harbors a full Orco/OR gene family repertoire**

143

144 With the exception of the Zygentoma, all lineages analyzed had genome data either
145 published (Faddeeva-Vakhrusheva et al., 2016; 2017; Wu et al., 2017) or available from
146 the i5k Pilot Project from the Baylor College of Medicine at the i5k NAL Workspace
147 (Poelchau et al., 2015). To complete taxon sampling of early branching hexapod and
148 insect lineages, we produced a draft genome assembly for *T. domestica* (Supplemental
149 Material; Figure S1; Table S1). This enabled direct comparison to Missbach et al. (2014)

150 and revealed that the *T. domestica* genome encodes far more than the three Orco-like OR
151 proteins. Our manual annotation revealed 43 ORs encoded by 31 genes including the
152 three previously identified genes (TdomOrco1-3; Missbach et al., 2014). Four genes are
153 modeled as exhibiting alternative splicing leading to the additional protein isoforms
154 (Supplemental Material). We used the antennal transcriptome of Missbach et al. (2014)
155 for support of intron-exon boundaries, however only a few transcriptome reads mapped
156 to the "specific" OR genes (Table S2), indicating that the RNA-seq analysis of Missbach
157 et al. (2014) did not sequence to a sufficient depth to reconstruct these low-expressed
158 transcripts.
159
160 Phylogenetic analyses of all ORs we annotated in the bristletail *Machilis hrabei* (5 ORs),
161 the dragonfly *Ladona fulva* (4 ORs), and the mayfly *Ephemera danica* (47 ORs), as well
162 as the previously annotated damselfly *C. splendens* (6 ORs; Ioannidis et al., 2017)
163 revealed that one of the *T. domestica* ORs (TdomOrco2) clustered confidently with the
164 Orco lineage in pterygote insects (Figure 2; Figure S2). We believe this is the sole Orco
165 relative because it shares unique features with the pterygote Orco proteins, such as a
166 TKKQ motif in the expanded intracellular loop 2 (positions 327-330 in DmelOrco), and
167 so we simply rename it TdomOrco. TdomOr1-8 represent a set of Orco-like proteins that
168 share a common gene structure with TdomOrco, with introns in phases 0-2-0-0-0. These
169 last four introns are present in all the other TdomOr genes, as well as those of the
170 bristletail, Odonata, and mayfly, and correspond to the four introns identified by
171 Robertson et al. (2003) as being ancestral to the OR family. The first phase-0 intron of
172 Orco and Or1-8 is the only additional intron shared by most pterygote Orco genes.
173
174 With the exception of the bristletail *M. hrabei*, all insect genomes analyzed have both
175 single genes with high similarity to Orco and multiple genes with similarity to specific
176 ORs. The *M. hrabei* genome did not encode an Orco, but instead contains 5 ORs of high
177 similarity that form a highly supported clade in the gene phylogeny (Figure 2). We also
178 could not find an Orco in the deep RNAseq transcriptome Missbach et al. (2014)
179 generated for their bristletail, *L. y-signata*. This finding leaves open two possibilities.
180 First, bristletails might have lost their Orco gene. Second, the OR family might have

originated with a few specific ORs like those of *M. hrabei*, with the Orco lineage evolving between Archaegnatha and Zygentoma. Phylogenetic analysis using various sets of GRs from other insects, arthropods, and animals as outgroup to root the OR family tree does not resolve this question confidently (data not shown). In any case, these five specific ORs in *M. hrabei*, at least one of which is present in *L. y-signata*, must function in the absence of Orco, perhaps alone or as dimers.

Finally, we note that while insects are defined by morphological and developmental synapomorphies (shared derived characters that are unique to a taxon), to the best of our knowledge presence of the OR gene family is the first molecular synapomorphy for the Class Insecta.

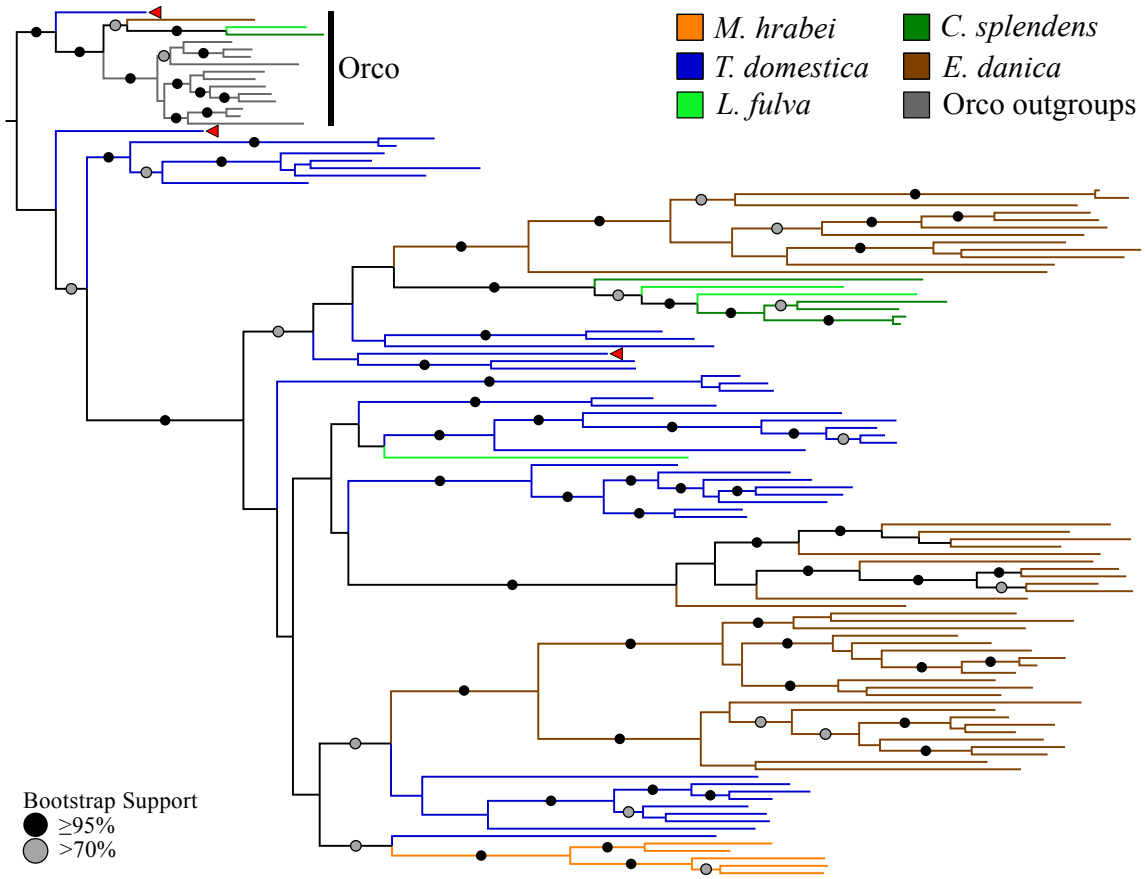


Figure 2 Odorant receptor (OR) gene family phylogeny including representatives of all apterygote and paleopteran insect orders. The Maximum Likelihood tree demonstrates monophyly of the single-copy insect Orco with high bootstrap support. The *M. hrabei* genome lacks Orco but encodes five highly similar ORs clustering in a single highly-supported clade. *T. domestica* has a fully developed functional OR/Orco system. The red arrowhead indicates the location of the three *T. domestica* ORs identified by Missbach et al. (2014).

Materials and Methods

A detailed Materials and Methods section can be found in the Supplementary Material.

Sequencing and assembly of the *Thermobia domestica* genome

Sequencing and assembly of the *T. domestica* genome followed Brand et al. (2018). Briefly, we sequenced the DNA extracted from a single *T. domestica* individual for assembly with the DISCOVAR v1 pipeline (Weisenfeld et al., 2014). Quality assessment of the resulting assembly was based on standard N statistics, k-mer distribution analysis, and the BUSCO v2 pipeline (Simão et al., 2015) as previously described (Brand et al., 2017).

Odorant receptor annotation

The genomes of the dipluran *Catajapyx aquinolaris,* the collembolans *Holacanthella duospinosa, Orchesella cincta,* and *Folsomia candida*, the firebrat *Thermobia domestica*, the bristletail *Machilis hrabei,* the dragonfly *Ladona fulva*, and the mayfly *Ephemera danica* were used for manual OR gene annotation (detailed in Supplementary Material).

Odorant receptor gene family analysis

OR protein alignments were produced with CLUSTALX v2 (Larkin et al., 2007) and trimmed using TrimAl (Capella-Gutiérrez et al., 2009). The resulting alignment was used for gene tree inference using RaxML (Stamatakis et al., 2005) under the JTT + G substitution model with 20 independent ML searches and 1000 bootstrap replicates as previously described (Brand and Ramírez, 2017).

Acknowledgements

We thank Stephen Richards for permission to examine unpublished genome sequences from the i5k pilot project, and Kimberly Walden for assistance with BLAST searches.

This work was funded by a National Science Foundation grant, IOS-1456678, to Brian Johnson and Juan Luis Jurat-Fuentes and a Hatch grant to Brian Johnson (CA-D-ENM 2161-H).

References

Almeida FC, Sánchez-Gracia A, Walden KKO, Robertson HM, Rozas J. 2015. Positive selection in extra cellular domains in the diversification of *Strigamia maritima* chemoreceptors. Front Ecol Evol 3:1669.

Benton R, Sachse S, Michnick SW, Vosshall LB. 2006. Atypical membrane topology and heteromeric function of *Drosophila* odorant receptors in vivo. PLoS Biol 4:e20.

Brand P, Lin W, Johnson BR. 2018. The Draft Genome of the Invasive Walking Stick, Medauroidea extradentata, Reveals Extensive Lineage-Specific Gene Family Expansions of Cell Wall Degrading Enzymes in Phasmatodea. G3: Genes, Genomes, Genetics:g3.200204.2018.

Brand P, Ramírez SR. 2017. The Evolutionary Dynamics of the Odorant Receptor Gene Family in Corbiculate Bees. Genome Biol Evol 9:2023–2036.

Brand P, Saleh N, Pan H, Li C, Kapheim KM, Ramírez SR. 2017. The Nuclear and Mitochondrial Genomes of the Facultatively Eusocial Orchid Bee *Euglossa dilemma*. G3: Genes, Genomes, Genetics 7:2891–2898.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972–1973.

Eyun S-I, Soh HY, Posavi M, Munro JB, Hughes DST, Murali SC, Qu J, Dugan S, Lee SL, Chao H, Dinh H, Han Y, Doddapaneni H, Worley KC, Muzny DM, Park E-O, Silva JC, Gibbs RA, Richards S, Lee CE. 2017. Evolutionary History of Chemosensory-Related Gene Families across the Arthropoda. Mol Biol Evol 34:1838–1862.

Faddeeva-Vakhrusheva A, Derks MFL, Anvar SY, Agamennone V, Suring W, Smit S, van Straalen NM, Roelofs D. 2016. Gene Family Evolution Reflects Adaptation to Soil Environmental Stressors in the Genome of the Collembolan *Orchesella cincta*. Genome Biol Evol 8:2106–2117.

Faddeeva-Vakhrusheva A, Kraaijeveld K, Derks MFL, Anvar SY, Agamennone V, Suring W, Kampfraath AA, Ellers J, Le Ngoc G, van Gestel CAM, Marien J, Smit S, van Straalen NM, Roelofs D. 2017. Coping with living in the soil: the genome of the parthenogenetic springtail *Folsomia candida*. BMC Genomics 18.

Groh-Lunow KC, Getahun MN, Grosse-Wilde E, Hansson BS. 2015. Expression of

266 ionotropic receptors in terrestrial hermit crab's olfactory sensory neurons. Front Cell
267 Neurosci 8:198.

268 Gulia-Nuss M, Nuss AB, Meyer JM, Sonenshine DE, Roe RM, Waterhouse RM, Sattelle
269 DB, la Fuente de J, Ribeiro JM, Megy K, Thimmapuram J, Miller JR, Walenz BP,
270 Koren S, Hostetler JB, Thiagarajan M, Joardar VS, Hannick LI, Bidwell S,
271 Hammond MP, Young S, Zeng Q, Abrudan JL, Almeida FC, Ayllon N, Bhide K,
272 Bissinger BW, Bonzon-Kulichenko E, Buckingham SD, Caffrey DR, Caimano MJ,
273 Croset V, Driscoll T, Gilbert D, Gillespie JJ, Giraldo-Calderon GI, Grabowski JM,
274 Jiang D, Khalil SMS, Kim D, Kocan KM, Koci J, Kuhn RJ, Kurtti TJ, Lees K, Lang
275 EG, Kennedy RC, Kwon H, Perera R, Qi Y, Radolf JD, Sakamoto JM, Sánchez-
276 Gracia A, Severo MS, Silverman N, Simo L, Tojo M, Tornador C, Van Zee JP,
277 Vazquez J, Vieira FG, Villar M, Wespiser AR, Yang Y, Zhu J, Arensburger P,
278 Pietrantonio PV, Barker SC, Shao R, Zdobnov EM, Hauser F, Grimmelikhuijzen
279 CJP, Park Y, Rozas J, Benton R, Pedra JHF, Nelson DR, Unger MF, Tubio JMC, Tu
280 Z, Robertson HM, Shumway M, Sutton G, Wortman JR, Lawson D, Wikel SK, Nene
281 VM, Fraser CM, Collins FH, Birren B, Nelson KE, Caler E, Hill CA. 2016. Genomic
282 insights into the *Ixodes scapularis* tick vector of Lyme disease. Nature
283 Communications 7.

284 Hansson BS, Harzsch S, Knaden M, Stensmyr M. 2010. The Neural and Behavioral Basis
285 of Chemical Communication in Terrestrial Crustaceans. In: Chemical
286 Communication in Crustaceans. New York, NY: Springer, New York, NY. p 149–
287 173.

288 Hansson BS, Stensmyr MC. 2011. Evolution of insect olfaction. Neuron 72:698–711.

289 Ioannidis P, Simão FA, Waterhouse RM, Manni M, Seppey M, Robertson HM, Misof B,
290 Niehuis O, Zdobnov EM. 2017. Genomic features of the damselfly *Calopteryx*
291 *splendens* representing a sister clade to most insect orders. Genome Biol
292 Evol:evx006.

293 Kirkness EF, Haas BJ, Sun W, Braig HR, Perotti MA, Clark JM, Lee SH, Robertson HM,
294 Kennedy RC, Elhaik E, Gerlach D, Kriventseva EV, Elsik CG, Graur D, Hill CA,
295 Veenstra JA, Walenz B, Tubío JMC, Ribeiro JMC, Rozas J, Johnston JS, Reese JT,
296 Popadic A, Tojo M, Raoult D, Reed DL, Tomoyasu Y, Kraus E, Krause E, Mittapalli
297 O, Margam VM, Li H-M, Meyer JM, Johnson RM, Romero-Severson J, VanZee JP,
298 Alvarez-Ponce D, Vieira FG, Aguadé M, Guirao-Rico S, Anzola JM, Yoon KS,
299 Strycharz JP, Unger MF, Christley S, Lobo NF, Seufferheld MJ, Wang N, Dasch GA,
300 Struchiner CJ, Madey G, Hannick LI, Bidwell S, Joardar V, Caler E, Shao R, Barker
301 SC, Cameron S, Bruggner RV, Regier A, Johnson J, Viswanathan L, Utterback TR,
302 Sutton GG, Lawson D, Waterhouse RM, Venter JC, Strausberg RL, Berenbaum MR,
303 Collins FH, Zdobnov EM, Pittendrigh BR. 2010. Genome sequences of the human
304 body louse and its primary endosymbiont provide insights into the permanent
305 parasitic lifestyle. Proc Natl Acad Sci USA 107:12168–12173.

306 Krång A-S, Knaden M, Steck K, Hansson BS. 2012. Transition from sea to land:

307 olfactory function and constraints in the terrestrial hermit crab *Coenobita clypeatus*.
308 Proc Biol Sci 279:3510–3519.

309 Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H,
310 Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG.
311 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948. Available
312 from: https://academic.oup.com/bioinformatics/article-
313 lookup/doi/10.1093/bioinformatics/btm404

314 Larsson MC, Domingos AI, Jones WD, Chiappe ME, Amrein H, Vosshall LB. 2004.
315 Or83b encodes a broadly expressed odorant receptor essential for *Drosophila*
316 olfaction. Neuron 43:703–714.

317 Misof B, Liu S, Meusemann K, Peters RS, Donath A, Mayer C, Frandsen PB, Ware J,
318 Flouri T, Beutel RG, Niehuis O, Petersen M, Izquierdo-Carrasco F, Wappler T, Rust
319 J, Aberer AJ, Aspöck U, Aspöck H, Bartel D, Blanke A, Berger S, Böhm A, Buckley
320 TR, Calcott B, Chen J, Friedrich F, Fukui M, Fujita M, Greve C, Grobe P, Gu S,
321 Huang Y, Jermiin LS, Kawahara AY, Krogmann L, Kubiak M, Lanfear R, Letsch H,
322 Li Y, Li Z, Li J, Lu H, Machida R, Mashimo Y, Kapli P, McKenna DD, Meng G,
323 Nakagaki Y, Navarrete-Heredia JL, Ott M, Ou Y, Pass G, Podsiadlowski L, Pohl H,
324 Reumont von BM, Schütte K, Sekiya K, Shimizu S, Slipinski A, Stamatakis A, Song
325 W, Su X, Szucsich NU, Tan M, Tan X, Tang M, Tang J, Timelthaler G, Tomizuka S,
326 Trautwein M, Tong X, Uchifune T, Walzl MG, Wiegmann BM, Wilbrandt J, Wipfler
327 B, Wong TKF, Wu Q, Wu G, Xie Y, Yang S, Yang Q, Yeates DK, Yoshizawa K,
328 Zhang Q, Zhang R, Zhang W, Zhang Y, Zhao J, Zhou C, Zhou L, Ziesmann T, Zou
329 S, Li Y, Xu X, Zhang Y, Yang H, Wang J, et al. 2014. Phylogenomics resolves the
330 timing and pattern of insect evolution. Science 346:763–767.

331 Missbach C, Dweck HK, Vogel H, Vilcinskas A, Stensmyr MC, Hansson BS, Grosse-
332 Wilde E. 2014. Evolution of insect olfactory receptors. Elife 3:e02115.

333 Ngoc PCT, Greenhalgh R, Dermauw W, Rombauts S, Bajda S, Zhurov V, Grbic M, Van
334 de Peer Y, Van Leeuwen T, Rouze P, Clark RM. 2016. Complex Evolutionary
335 Dynamics of Massively Expanded Chemosensory Receptor Families in an Extreme
336 Generalist Chelicerate Herbivore. Genome Biol Evol 8:3323–3339.

337 Niimura Y, Matsui A, Touhara K. 2014. Extreme expansion of the olfactory receptor
338 gene repertoire in African elephants and evolutionary dynamics of orthologous gene
339 groups in 13 placental mammals. Genome Res 24:1485–1496.

340 Peñalva-Arana DC, Lynch M, Robertson HM. 2009. The chemoreceptor genes of the
341 waterflea Daphnia pulex: many Grs but no Ors. BMC Evol Biol 9:79.

342 Poelchau M, Childers C, Moore G, Tsavatapalli V, Evans J, Lee C-Y, Lin H, Lin J-W,
343 Hackett K. 2015. The i5k Workspace@NAL--enabling genomic data access,
344 visualization and curation of arthropod genomes. Nucleic Acids Res 43:D714–9.

345 Purrington FF, Kendall PA, Bater JE, Stinner BR. 1991. Alarm Pheromone in a

346 Gregarious Poduromorph-Collembolan (Collembola, Hypogastruridae). Great Lakes
347 Entomologist 24:75–78.

348 Rimal S, Lee Y. 2018. The multidimensional ionotropic receptors of *Drosophila*
349 *melanogaster*. Insect Mol Biol 27:1–7.

350 Robertson HM, Warr CG, Carlson JR. 2003. Molecular evolution of the insect
351 chemoreceptor gene superfamily in *Drosophila melanogaster*. Proc Natl Acad Sci
352 USA 100.

353 Robertson HM. 2015. The Insect Chemoreceptor Superfamily Is Ancient in Animals.
354 Chem Senses:bjv046.

355 Rytz R, Croset V, Benton R. 2013. Ionotropic receptors (IRs): chemosensory ionotropic
356 glutamate receptors in *Drosophila* and beyond. Insect Biochem Mol Biol 43.

357 Saina M, Busengdal H, Sinigaglia C, Petrone L, Oliveri P, Rentzsch F, Benton R. 2015.
358 A cnidarian homologue of an insect gustatory receptor functions in developmental
359 body patterning. Nature Communications 6.

360 Sánchez-Gracia A, Vieira FG, Rozas J. 2009. Molecular evolution of the major
361 chemosensory gene families in insects. Heredity 103:208–216.

362 Scott K, Brady R, Cravchik A, Morozov P, Rzhetsky A, Zuker C, Axel R. 2001. A
363 chemosensory gene family encoding candidate gustatory and olfactory receptors in
364 *Drosophila*. Cell 104:661–673.

365 Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO:
366 assessing genome assembly and annotation completeness with single-copy orthologs.
367 Bioinformatics 31:3210–3212.

368 Smith CD, Zimin A, Holt C, Abouheif E, Benton R, Cash E, Croset V, Currie CR, Elhaik
369 E, Elsik CG, Favé M-J, Fernandes V, Gadau J, Gibson JD, Graur D, Grubbs KJ,
370 Hagen DE, Helmkampf M, Holley J-A, Hu H, Viniegra ASI, Johnson BR, Johnson
371 RM, Khila A, Kim JW, Laird J, Mathis KA, Moeller JA, Munoz-Torres MC, Murphy
372 MC, Nakamura R, Nigam S, Overson RP, Placek JE, Rajakumar R, Reese JT,
373 Robertson HM, Smith CR, Suarez AV, Suen G, Suhr EL, Tao S, Torres CW, van
374 Wilgenburg E, Viljakainen L, Walden KKO, Wild AL, Yandell M, Yorke JA,
375 Tsutsui ND. 2011a. Draft genome of the globally widespread and invasive Argentine
376 ant (*Linepithema humile*). Proc Natl Acad Sci USA 108:5673–5678.

377 Smith CR, Smith CD, Robertson HM, Helmkampf M, Zimin A, Yandell M, Holt C, Hu
378 H, Abouheif E, Benton R, Cash E, Croset V, Currie CR, Elhaik E, Elsik CG, Favé M-
379 J, Fernandes V, Gibson JD, Graur D, Gronenberg W, Grubbs KJ, Hagen DE,
380 Viniegra ASI, Johnson BR, Johnson RM, Khila A, Kim JW, Mathis KA, Munoz-
381 Torres MC, Murphy MC, Mustard JA, Nakamura R, Niehuis O, Nigam S, Overson
382 RP, Placek JE, Rajakumar R, Reese JT, Suen G, Tao S, Torres CW, Tsutsui ND,
383 Viljakainen L, Wolschin F, Gadau J. 2011b. Draft genome of the red harvester ant

384 *Pogonomyrmex barbatus*. Proc Natl Acad Sci USA 108:5667–5672.

385 Staaden S, Milcu A, Rohlfs M, Scheu S. 2011. Olfactory cues associated with fungal
386 grazing intensity and secondary metabolite pathway modulate Collembola foraging
387 behaviour. Soil Biology & Biochemistry 43:1411–1416.

388 Stamatakis A, Ludwig T, Meier H. 2005. RAxML-III: a fast program for maximum
389 likelihood-based inference of large phylogenetic trees. Bioinformatics 21:456–463.

390 Stensmyr MC, Erland S, Hallberg E, Wallén R, Greenaway P, Hansson BS. 2005. Insect-
391 like olfactory adaptations in the terrestrial giant robber crab. Current Biology
392 15:116–121.

393 Verhoef HA, Nagelkerke CJ, Joosse E. 1977. Aggregation Pheromones in Collembola.
394 Journal of Insect Physiology 23:1009–1013.

395 Vosshall LB, Hansson BS. 2011. A unified nomenclature system for the insect olfactory
396 coreceptor. Chem Senses 36:497–498.

397 Weisenfeld NI, Yin S, Sharpe T, Lau B, Hegarty R, Holmes L, Sogoloff B, Tabbaa D,
398 Williams L, Russ C, Nusbaum C, Lander ES, MacCallum I, Jaffe DB. 2014.
399 Comprehensive variation discovery in single human genomes. Nat Genet 46:1350–
400 1355.

401 Wu C, Jordan MD, Newcomb RD, Gemmell NJ, Bank S, Meusemann K, Dearden PK,
402 Duncan EJ, Grosser S, Rutherford K, Gardner PP, Crowhurst RN, Steinwender B,
403 Tooman LK, Stevens MI, Buckley TR. 2017. Analysis of the genome of the New
404 Zealand giant collembolan (*Holacanthella duospinosa*) sheds light on hexapod
405 evolution. BMC Genomics 18.

406 Wuichet K, Zhulin IB. 2010. Origins and diversification of a complex signal transduction
407 system in prokaryotes. Sci Signal 3:ra50–ra50.

408