

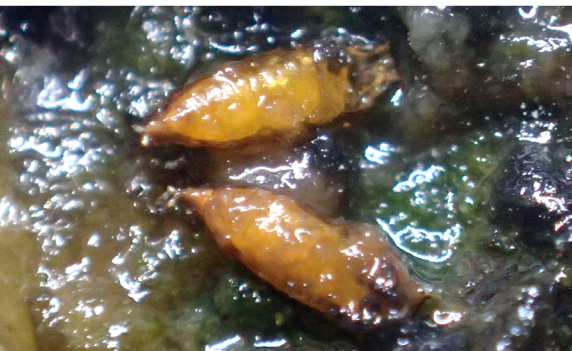
1014

Supplementary figures

1015

1016 Figure S1. Culturing species of the *Lordiphosa clarofinis* species complex. (A) Thick growth of
1017 *Galinsoga parviflora* Gav., one of the natural breeding substrates (Baihualing, Baoshan, Yunnan).
1018 (B) Leaves and flowers of the same plant. (C) Frozen leaves and stems of *G. parviflora* used for
1019 rearing *Lordiphosa* flies. (D) Glass vials containing *G. parviflora* as larval food, and filter paper as
1020 pupation substrate (pieces of apple, used as adult food, were added later). (E) Pupae of *L.*
1021 *clarofinis* reared in the lab on this substrate.

1022



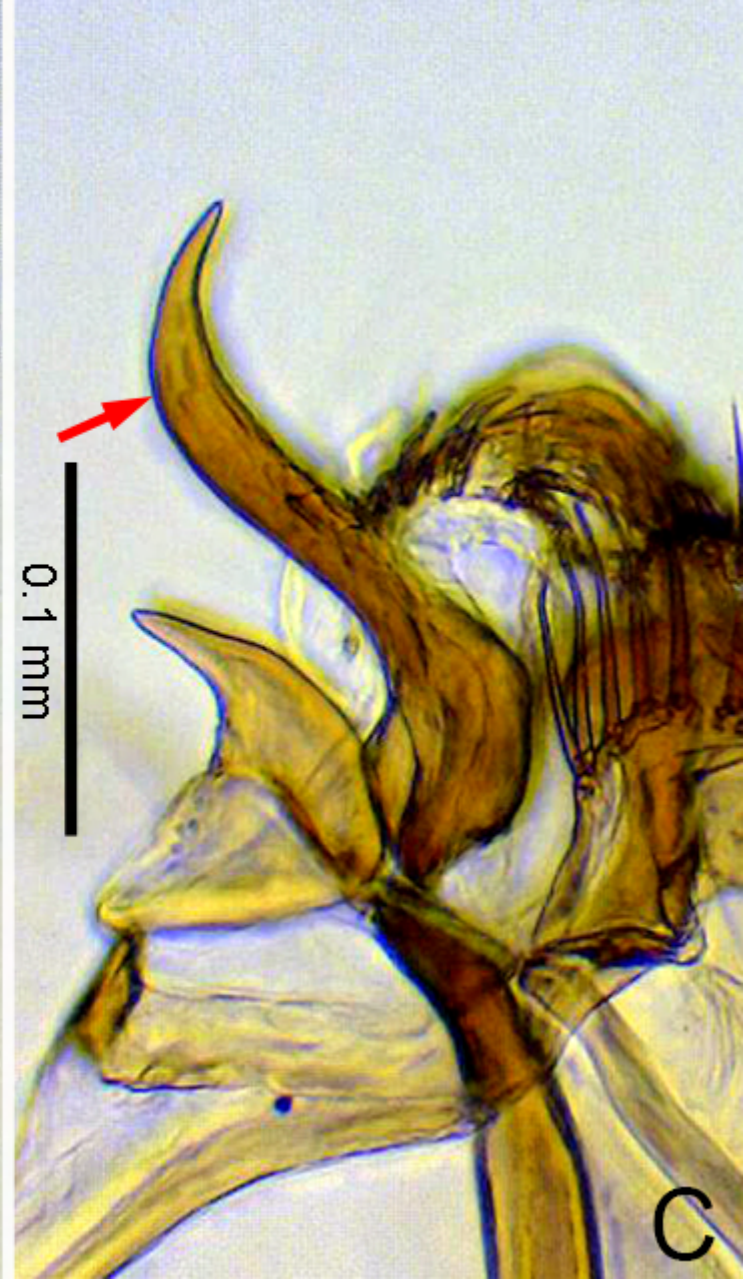
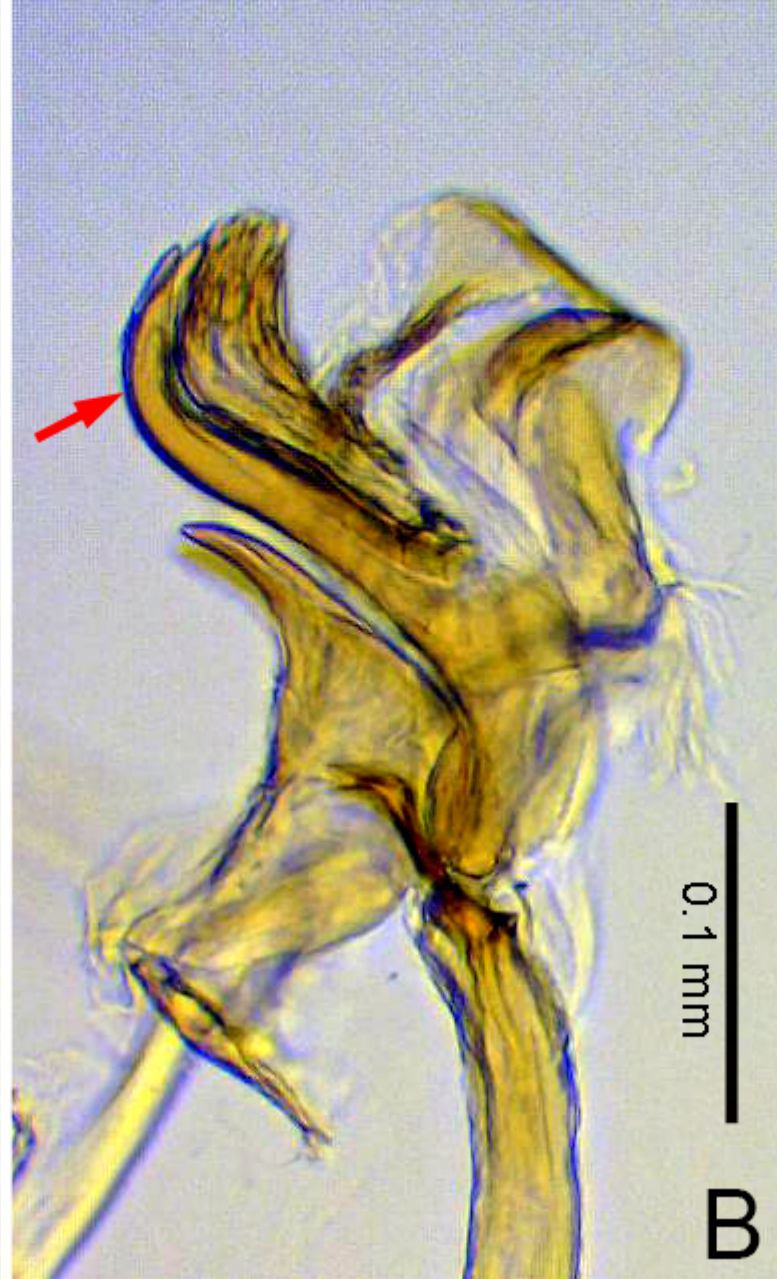
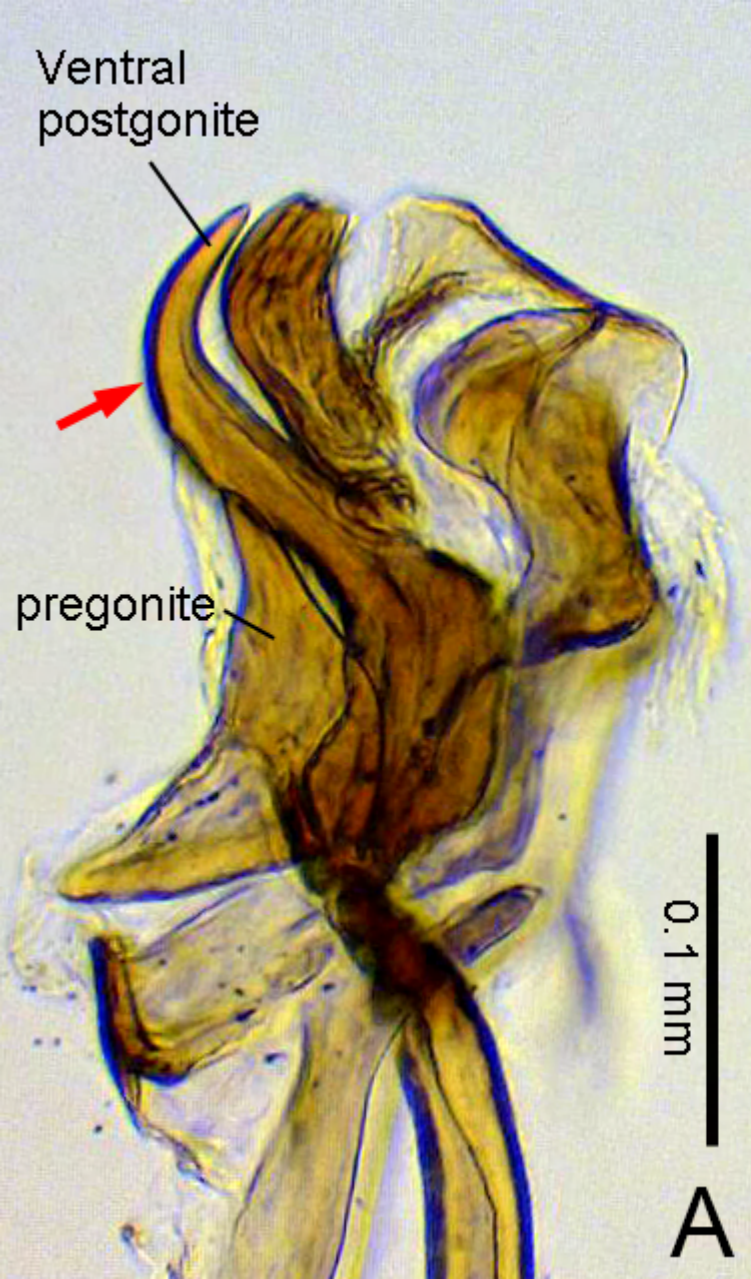
1023 Figure S2. Collection sites of four members of the *L. clarofinis* complex and two outgroup species.
1024 Map was created using Simplemappr (<http://www.simplemappr.net/>).
1025



- *L. clarofinis*
- *L. sp.1 aff. clarofinis*
- *L. sp.2 aff. clarofinis*
- *L. sp.3 aff. clarofinis*
- *L. magnipectinata*
- *L. stackelbergi*

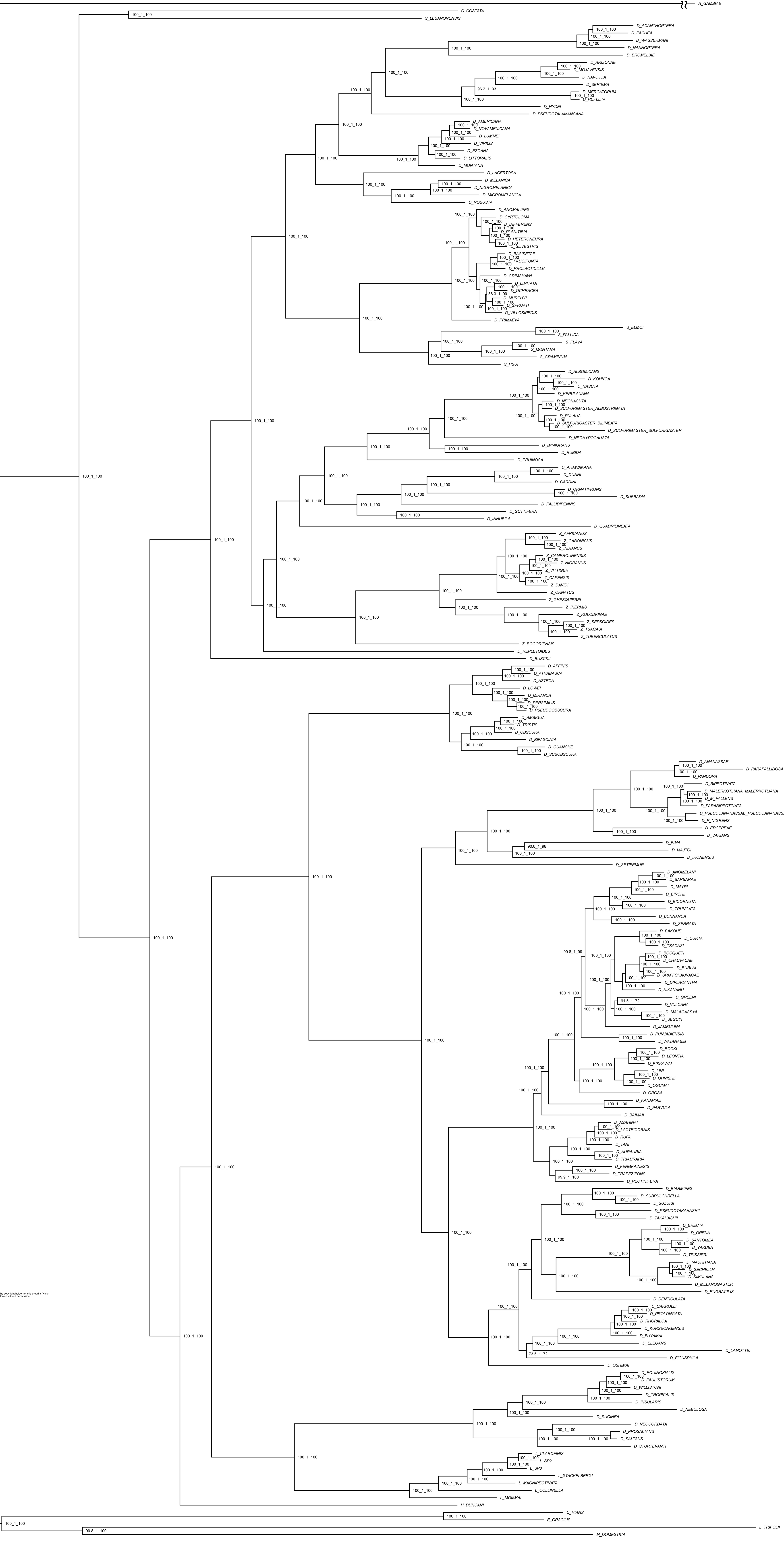
1026 Figure S3. Pregonite and ventral postgonite of *L. clarofinis* (all lateral views; A, specimen from
1027 Korea; B, specimen #01711 from Taibaishan, Shaanxi, China) and *L. sp.1* aff. *clarofinis* (C,
1028 specimen #01903 from Hokkaido, Japan). Red arrows indicate the bending position of the ventral
1029 postgonite in each specimen.

1030

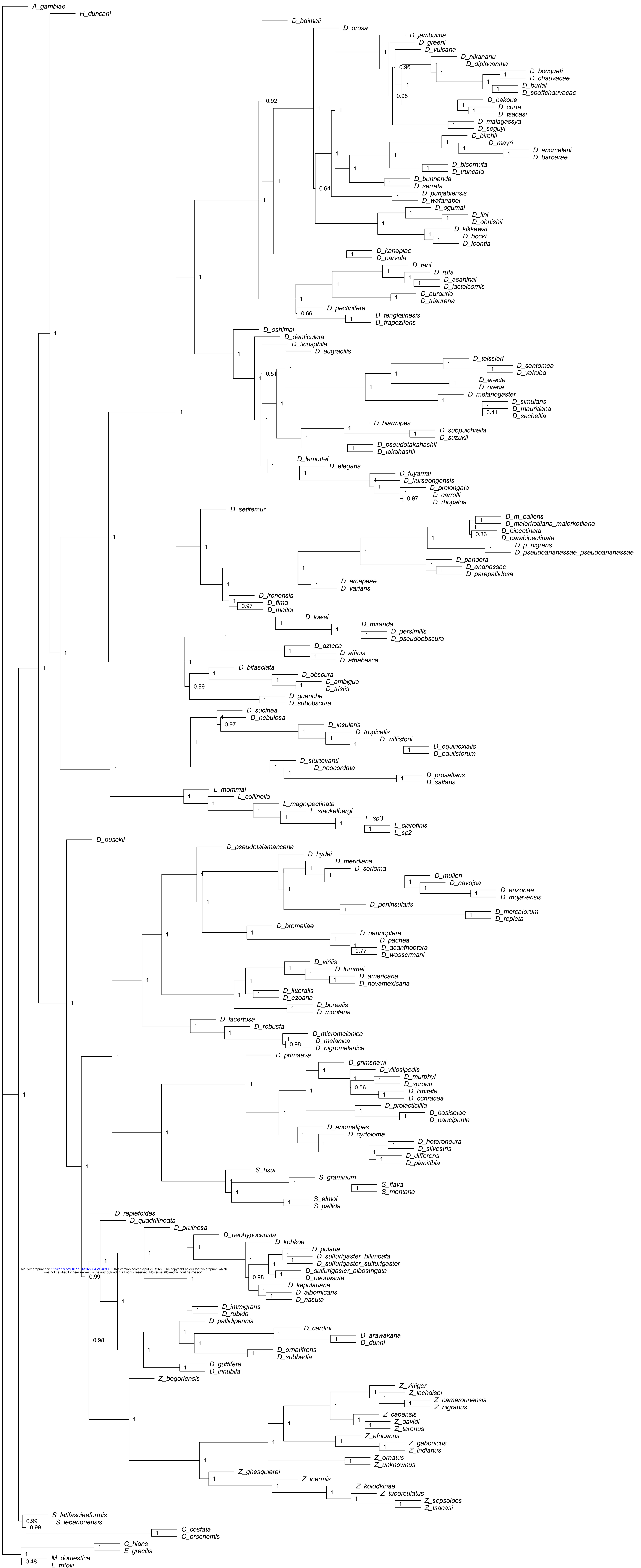


1031 Figure S4. **A.** A maximum likelihood tree of 207 species of *Drosophila*, *Lordiphosa*, and related
1032 genera reconstructed using IQ-TREE v1.6.5 from 250 single-copy BUSCO loci (572,343 total sites)
1033 (Table S4) and rooted with *Anopheles gambiae*. At each node, the three measures of node
1034 support are, in order: 1000 replicates of ultrafast bootstrap (Minh et al., 2013), a Bayesian-like
1035 transformation of approximate likelihood ratio test (aLRT) (Anisimova et al., 2011), and aLRT with
1036 the nonparametric Shimodaira–Hasegawa correction (SH-aLRT). The *A. gambiae* branch is not to
1037 scale. See Methods for details of phylogeny reconstruction. **B.** ASTRAL tree topology, with local
1038 posterior probability support measures for each node, estimated from the 250 individual gene
1039 trees. Branch lengths are reported in coalescent time units. ASTRAL only reports the length of
1040 internal branches; the length of each terminal branch was manually set to 1.

1041



bioRxiv preprint doi: <https://doi.org/10.1101/2022.04.01.484805>; this version posted April 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



bioRxiv preprint doi: <https://doi.org/10.1101/2022.04.21.489090>; this version posted April 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

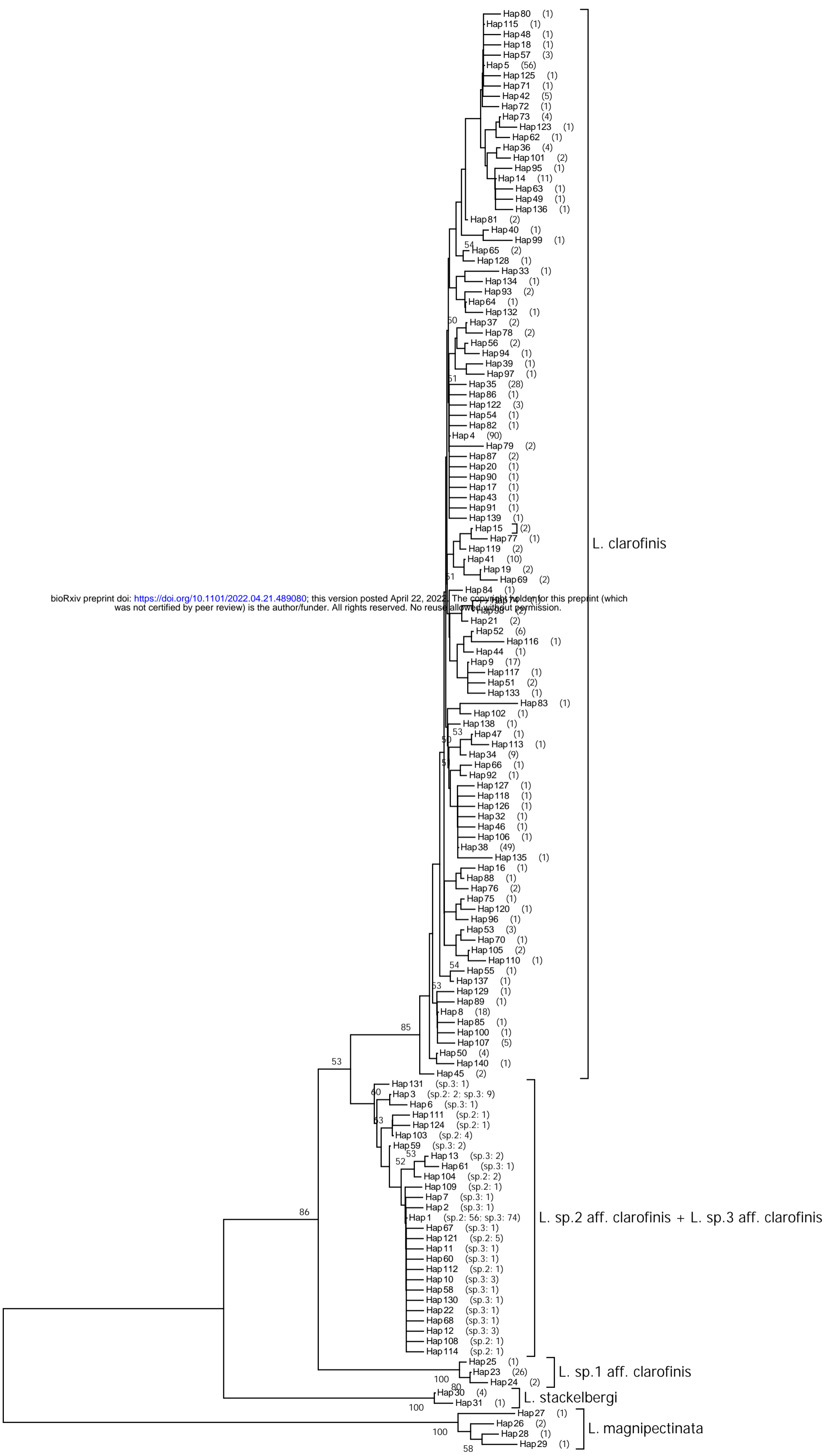
1042 Figure S5. Ancestral state reconstruction under five different models of trait evolution. The
1043 phylogenetic tree is based on 250 single-copy BUSCO genes extracted from whole genome
1044 sequences ([Fig. S4A](#)). Species are colored according to sex comb state (red = present, blue =
1045 absent). Pie charts at internal nodes reflect estimated probabilities of ancestral character states
1046 under five different models of trait evolution: (A) MK model with unequal rates and a strict
1047 molecular clock (Lewis, 2001); (B) MK model with unequal rates and a random local relaxed clock
1048 (Drummond and Suchard, 2010); (C) Hidden states variable rates model with two latent rate
1049 classes (Beaulieu et al., 2013); (D) Modified threshold model (Felsenstein, 2005) with 9 ordered
1050 latent states; (E) Approximation of a Dollo model, with rate of loss >300-fold higher than rate of
1051 gain. See Methods and [Table S5](#) for details of ancestral character reconstruction.

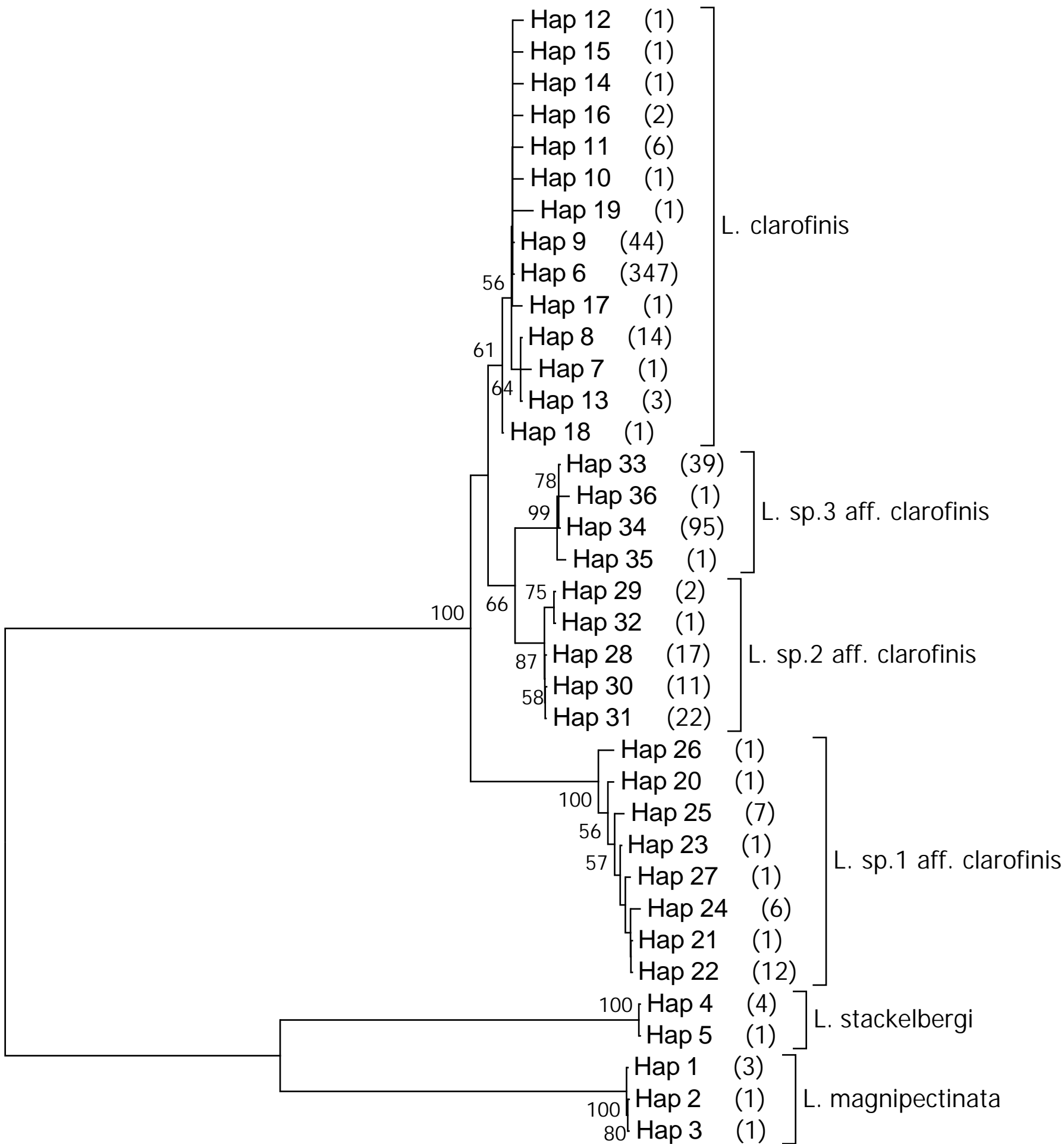
1052

1053 Figure S6. Haplotype trees of the *L. clarofinis* species complex. (A) Unrooted NJ tree of 140
1054 selected *COI* haplotypes. For each haplotype, the frequency is given in parentheses (with the
1055 name of the corresponding species indicated in the cluster of *L. sp.2* aff. *clarofinis* + *L. sp.3* aff.
1056 *clarofinis*). Note that haplotypes 1 and 3 are both shared by these two species. (B) Unrooted NJ
1057 tree of all 36 ITS1 haplotypes. Frequency is given in parentheses for each haplotype.

1058

bioRxiv preprint doi: <https://doi.org/10.1101/2022.04.21.489080>; this version posted April 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



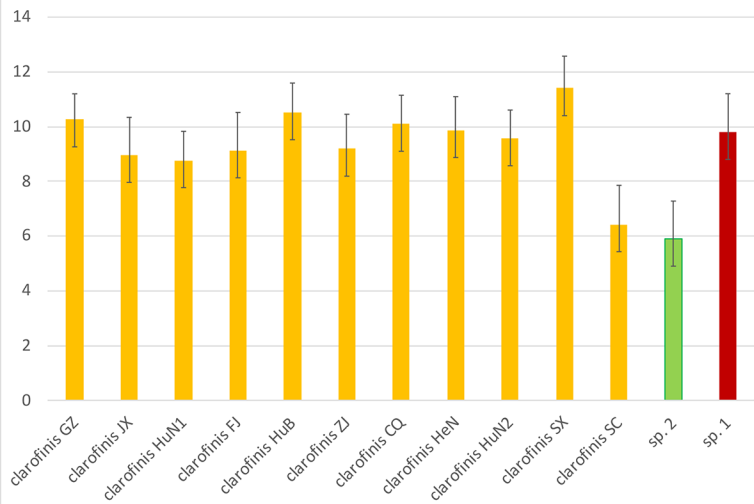


0.01

1059 Figure S7. Sex comb size (number of sex comb teeth) in the *clarofinis* complex. Species are coded
1060 by color: yellow = *L. clarofinis*, green = sp.2, red = sp.1. All strains of sp.3 lack sex combs. For *L.*
1061 *clarofinis*, different geographic populations are shown separately. Note that *L. clarofinis*
1062 population from Sichuan (SC) is more similar to sp.2 than to other *L. clarofinis* populations in the
1063 number of sex comb teeth on the first (most proximal) tarsal segment.

1064

tarsomere 1



tarsomere 2

