Strains of Apis melliferaligustica honey bees artificially bred for apicultural traits are not consistently differentiated by mitochondrial DNA genome markers

Journal:	*Genome*
Manuscript ID	gen-2025-0033.R1
Manuscript Type:	Research Article
Date Submitted by the Author:	07-May-2025
Complete List of Authors:	Carr, Steven; Memorial University of Newfoundland, Department of Biology
Is the manuscript for consideration in a Special Issue or Collection?:	Not applicable (regular submission)
Keyword:	Apis mellifera ligustica, apiculture, Royal Jelly, Varoa resistance, mtDNA genomes

Strains of *Apis mellifera ligustica* honey bees artificially bred for apicultural traits are not consistently differentiated by mitochondrial DNA genome markers

Steven M. Carr, Department of Biology, 45 Arctic Drive,
Memorial University of Newfoundland, St John's NL A1C 5S7

Submitted 2025 April 02
Genome
Revision through 2025 May 04; submitted in revised form 2025 May 05

Address for correspondence and proofs: scarr@mun.ca / 1 709 687 7314 cell

Abstract

Strains of the Italian honey bee *Apis mellifera ligustica* Spinola, 1806 are selectively bred for desirable apiculture traits. Ma et al. (2022) compared SNP differences in mtDNA genomes between a strain bred for enhanced royal jelly production (**RJB**) and an unselected strain (**ITB**). Kim et al. (2019) compared SNP and intergenic repeats differences between a *Varroa* mite resistant strain bred for high-hygienic behavior (**HHB**) and an unselected low-hygienic strain (**LHB**). Phylogenetic comparison of 23 complete *A. m. ligustica* mtDNA sequences, including the two **HHB** and **LHB** strains and 14 **RJB** and **ITB** haplotypes and along with other subspecies, shows significant clade structure for SNP differences and amino acid substitutions, however this structure is not diagnostic of the strains under selection. Numbers of repeats show little or no phylogenetic signal: similarities are symplesiomorphic and differences convergent. Differences between the **HHB** and **LHB** strains arise from coincidental fixation of alternative SNPs in different clades. Lineages within the **RJB** and **ITB** strains are not reciprocally monophyletic, and are often identical or near identical to those in the other strain or to other variant strains. Evaluation of the diagnostic and (or) adaptive significance of mtDNA markers requires broad knowledge of within-subspecies polymorphism.

Keywords: *Apis mellifera ligustica*, apiculture, Royal Jelly, *Varoa* mite resistance, mtDNA genomes, intra-subspecific systematics.

Introduction

Strains of various subspecies of the European Honey Bee *Apis mellifera* L., including the Italian or Golden Honey Bee, *A. m. ligustica* Spinola, 1806 have for centuries been selectively bred to enhance traits commercially desirable to apiculturists (Seeley 2019). For example, infestation of colonies by the parasitic mite *Varroa destructor* is the major threat to apiculture worldwide (Johnson 2023). Colony survival is improved in strains bred for enhanced hygienic behavior, according to the efficiency whereby workers remove dead or diseased bee larvae. Other traits, such as enhanced production of “royal jelly” as a high-energy human food supplement, are also desirable and subject to artificial selection.

Kim et al. (2019) bred a high-hygienic behavior (**HHB**) strain of *A. m. ligustica* and compared aspects of its mitochondrial DNA (**mtDNA**) genome sequence against that of a low-hygienic (**LHB**) strain of similar geographic origin. The **HHB** strain was bred by artificial insemination. Colonies are subsequently monitored by the freezer-killed brood method, for which molecular markers would be valuable. They reported counts of the numbers of repeats in several intergenic regions, and SNP differences in protein-coding genes between the two strains, and suggested that these markers may be useful in distinguishing **HHB** from **LHB** strains.

Ma et al. (2022) bred an enhanced royal jelly (**RJB**) strain that produces 30 times as much royal jelly as an unselected Italian strain (**ITB**). Where oxidative metabolism in mitochondria contributes strongly to energy production (Ma & Li, 2021), Ma et al. (2022) ask if the increased energy demands for higher yield of royal jelly in the **RJB** strain might be due to adaptive changes in mtDNA sequences. They compared differences between the two strains.

Carr (2023) constructed a phylogeny of 22 subspecies of *A. mellifera* based on the complete set of protein-coding mtDNA sequences, including multiple representatives of some subspecies such as *A. m. ligustica*. The analysis identified a Eurasian clade, consisting of *A. m. ligustica* together with *A. m. carnica* and *A. m. carpatica* in a Southeast Europe clade, in contrast to an Asia Minor clade (*A. m. anatoliaca*, *A. m. caucasia*, and *A. m. meda*) (cf. Carr 2023, Figure 2). Pairwise differences among *A. m ligustica* sequences were greater than those between other subspecies in the same clade, as was the case for several other subspecies. Carr (2023) showed that a proper understanding of phylogenetic relationships and biogeographic evolution required knowledge of variation within as well as between subspecies.

Kim et al. (2019) included single mtDNA sequences from their two **HHB** and **LHB** strains (the replicates of each strain all being identical), and asked if any observed mtDNA differences are diagnostic of the **HHB** strain, and (or) might potentially contribute to hygienic behavior. Similar questions arise for the more polymorphic **RJB** and **ITB** strains of Ma et al. (2022), where it is plausible that royal jelly production is intimately linked to enhanced intracellular metabolism in the powerhouses of the cell.

Answers to these questions require consideration of the broader range of genetic polymorphism within *A. m. ligustica* and related subspecies in the same evolutionary clade. Here, I analyze differences in complete mtDNA genomes between the **HHB** / **LHB** and **RJB** / **ITB** strains within that broader context. I test the hypothesis that SNPs, amino acid substitutions, and (or) repeat numbers in artificially selected strains of *A. m. ligustica* are of diagnostic and (or) functional significance, or are instead alternative partitions of neutral phylogenetic markers.

Materials and Methods

I obtained both complete (~ 16.5 kb) and 13 protein-coding only (11,070 b) mtDNA sequences for 3,690 amino acids (cf. Table 2, Carr 2023) from 31 accessions in GenBank (Table 1), including one each from **LHB** and **HHB** strains of Kim et al. (2019), n = 6 and n = 8 from the **ITB** and **RJB** strains of Ma et al. (2022) including two inter-strain duplicates, n = 7 additional sequences referred to *A. m. ligustica*, n = 7 sequences referred to five subspecies in the Southeast Europe and Asia Minor clades described by Carr (2023), and an *A. m. mellifera* sequence as overall outgroup.

Phylogenetic analysis was done with MEGA11 (Tamura et al. 2021) on complete sequences without the 3' AT-rich region (~15.6 kbp), and on protein-coding DNA triplets and their amino acid sequences as above, by Maximum Parsimony with Subtree-Pruning and Regrafting (SPR) and 3,000 bootstrap replicates. Maximum Likelihood analysis of protein-coding DNA triplets was done with a General Time Reversible model, Gamma-distributed with invariant sites (five categories), and 3,000 bootstrap replicates with Nearest-Neighbor Interchange (NNI).

Counts of two- or three-base repeats were made from alignments of the complete sequences at regions near 6.1 kb (**AT**), 10.1 kb (**AAT**), and 14.6 kb (**TA**) as identified by Kim et al. (2019).

Results

A Maximum Parsimony tree for SNP variation at protein-coding nucleotide sequences was initially rooted by inclusion of *A. m. mellifera* (KY926884), which separated the Southeast Europe and Asia Minor clades as previously obtained (cf. Figure 2 of Carr 2023). After removal of *A. m. mellifera*, Figure 1 is rooted with the Asia Minor clade as the outgroup to the *A. m. ligustica*-inclusive clade. There are nine trees of length L = 112, which differ only in the arrangement of unresolved tritomys. The Maximum Likelihood network has an identical topology, and substantially identical bootstrap support (results not shown). The Maximum Parsimony tree for amino acid substitutions (Figure 2) recovers five trees of length L = 35, which again differ only in the arrangements of unresolved tritomys. The topology of the amino acid tree conforms to that of the SNP tree. The Maximum Parsimony tree for complete sequences is substantially the same as Figures 1 and 2, except for statistically unsupported resolution of some tritomys (results not shown).

The basic structure of the *A. m. ligustica* tree is a poorly-resolved basal group of five sequences, and a well-supported, derived clade of 18 sequences that comprises two subclades of nine sequences each (including duplicates from **ITB** and **RJB**). The basal group includes **LHB**, representatives of the **ITB** and **RJB** series, and a sequence referred to *A. m. carnica* (NC061380) weakly paired with an **ITB** sequence. NC061380 is cladistically and phenetically distinct from the other sequence (MN250878) referred to *A. m. carnica*, which along with *A. m. carpatica* (AP018430) occurs inside *A. m. ligustica* at the base of the derived clade. When *A. m. carnica* and *A. m carpatica* are removed, the major clade is supported by 10 SNPs (99%), and includes an **HHB**-inclusive subclade supported by five SNPs (99%) (Supplementary Figure S3).

HHB and **LHB** are distantly related cladistically and differ phenetically. **HHB** occurs in one of the two derived subclades, supported by seven **SNPs** (98% of bootstrap replicates) (Figure 1) and three amino acid substitutions (88%) (Figure 2), which clade excludes **LHB**. Analysis of phylogenetically informative amino acid substitutions shows that **HHB** differs from **LHB** by the four synapomorphies of the derived clade (Table 2), and that **LHB** converges on that clade by one parallel substitution. **HHB** and **LHB** differ by 26 pairwise SNP differences and six amino acid substitutions (Table 3), versus means of 12.7 ± 9.3 SNPs and 4.18 ± 3.59 amino acid substitutions among all *A. m. ligustica* sequences.

Counts of repeat numbers at three positions are given in the parsimony tree in Figure 1, shown in Figure 3 for **AT** and in Supplementary Figure S2 for **AAT** and **TA**. **AAT**$_4$ is plesiomorphic for *A. m. mellifera*, most other subspecies, and most *ligustica* lineages including **RJB13** and **HHB**. In the clade comprising **RJB13** / **ITB16** + **LHB**, the slip to **AAT**$_3$ in **LHB** is shared with **ITB16**. In the unresolved clade comprising **HHB** / **ITB08** / **Buckfast**, **AAT**$_{6,7}$ occur in the latter two. **AAT**$_{14}$ occurs in a separate lineage. Higher **TA** repeat counts in **HHB** (**TA**$_{29}$) and **LHB** (**TA**$_{20}$) are convergent, the most closely related lineages having lower numbers. The same is true for their lower **AT** counts, **AT**$_{12}$ and **AT**$_{16}$, respectively. Counts of the **TA** and **AT** repeats are similar in the **RJB13** / **ITB16** / **LHB** clade. Note that **HHB** and **ITB08** have identical mtDNA protein-coding sequences, but differ by both **AAT** and **TA** repeat counts.

Among the 13 distinct sequences identified in the **ITB** and **RJB** strains of Ma et al. (2022), **HHB** is identical to **ITB08** (OM203226). **LHB** is most closely related (92% bootstrap support) to **ITB16** (OM203234), from which it differs by 17 SNPs including 10 in protein-coding genes, and three amino acid substitutions. Recall that **ITB08** and **ITB16** are from the unselected strain.

Discussion

MtDNA sequences referred to *A. m. ligustica* are extremely polymorphic, with as many as 27 SNPs and 15 amino acid substitutions between individual GenBank accessions. Differences between individuals within the subspecies in many cases exceed those between other subspecies of *A. mellifera* (cf. Carr 2023). Recognition of any SNP and (or) amino acid substitution as diagnostic of a regional population or commercial strain must consider the larger range of polymorphism within the subspecies and related subspecies. In this expanded context, I have re-examined the difference between two strains of *A. m. ligustica* (Kim et al. 2019), one bred for high hygienic behavior against *Varroa* mites (**HHB**) and the other not (**LHB**), together with two strains (Ma et al. 2022), one bred for enhanced royal jelly production (**RJB**) and the other not (**IJB**), and other wild and commercial strains.

Counts of repeats provide little or no phylogenetic or diagnostic signal. Compared with Figure 1, the atypical **AAT$_3$** repeat in **LHB** is derived, compared with the basal **AAT$_4$** repeat shared by **HHB** and other *A. m. ligustica*. Neither the **AT** or **TA** repeats distinguish **HHB** from **LHB**, nor any other branches. In general, higher or lower repeat counts in these regions have arisen by convergence and do not correspond to clades.

A. m. ligustica sensu lato includes sequences referred to both *A. m. carnica* and *A. m carpatica*, but replicate *A. m. carnica* sequences are polyphyletic in the basal and derived groups. When these two sequences are removed, the major clade is supported by 10 SNPs

(99%), and includes an **HHB**-inclusive subclade supported by five SNPs (99%) (Supplementary Figure S3). In the data here, patristic SNP differences among subspecies of the Asia Minor outgroup are 4 ~ 7 (Figure 1), compared with 28 inter-subspecific SNPs between *A. m. meda* and **LHB** and 30 intra-specific SNPs between **LHB** and **HHB** within *A. m. ligustica* (Table 1). Carr (2023) discussed the taxonomic implications of smaller inter-subspecific than intra-subspecific genetic differences. Names matter: see Supplementary Figure S4 for inclusive analysis of all GenBank sequences referred to *A. m. ligustica* and the Eurasian clades.

No SNP differences or amino acid substitutions are diagnostic of the **HHB** strain with respect to **LHB** or any other strain or single *ligustica* sequence. The amino acid sequence of **HHB** is identical to both **ITB08**, an unselected Italian strain, and **Buckfast**, a widely-distributed commercial strain. **LHB** differs from the most closely related sequences, in the **ITB** and **RJB** strains, by two amino acid substitutions. Phylogenetic analysis shows that the 26 pairwise SNP and six amino acid differences between the two strains are coincidental fixations of standing polymorphism in the subspecies

MtDNA sequence variation between the uniform **HHB** and **LHB** strains can be compared with that among the much more variable **ITB** and **RJB** strains of Ma et al. (2022). Identical amino acid sequences of **ITB** and **RJB** lineages occur in the basal group, and also in the three major clades and subclades (Table 2). **ITB** and **RJB** are not reciprocally monophyletic, as might be expected if selection for the latter trait coincidentally restricted mtDNA variation to a smaller subset of lineages.

In conclusion, the high hygienic behavior selected for in the **HHB** strain of Kim et al. (2019) cannot be associated with any diagnostic changes in its mtDNA exome with respect to

LHB nor any other closely related but unselected strains. The enhanced royal jelly production of the **RJB** strain of Ma et al. (2022) is associated with multiple mtDNA lineages within that strain, and also within unselected **ITB** and other strains with identical SNP and (or) amino acid patterns. Absence of reciprocal monophyly between strains is evidence against the origin of the trait in any single genetic lineage, followed by loss of mtDNA variation through genetic drift during selection for the trait.

Mitochondrial DNA has a long history of utility in the delineation of regional variation within insect (Cameron 2014) and other species (Wilson et al. 1985; Carr 2020). MtDNA SNPs in *Apis* have been shown to be valuable tools in identification of subspecies and selected strains involved in apiculture. However, documentation of any diagnostic or functional association of such variation with production traits among commercial strains of *A. m. ligustica*, and likely other subspecies of *Apis mellifera* L., must be made with broad knowledge and appreciation of genetic polymorphism and its phylogenetic structure within the species *in toto*.

Acknowledgements

The author has benefited from discussions of apiculture and *Varroa* mites with Ms Paige Marchant, Provincial Apiculturist for Newfoundland and Labrador. I thank Prof Chuan Ma for previous discussion of genetic variation in his **ITB** / **RJB** series and other subspecies of *Apis mellifera*. The final draft of the MS was submitted while the author was in residence at the Atlanta Soto Zen Center: *gassho* to all friends and the many beings.

Disclosure Statement

The author reports no conflicts. The author is an Associate Editor for *Mitochondrial DNA*, but played no part in the review or publication of Kim *et al*. (2019) or Ma (2021).

Funding

The present work is unfunded.

Data availability Statement

The data that support the findings of this study are openly available in the GenBank database at https://www.ncbi.nlm.nih.gov/taxonomy/?term=Apis+mellifera, with accession numbers as given in the manuscript.

References

Cameron SL. 2014. Insect mitochondrial genomics: implications for evolution and phylogeny. Annu Rev Entomol. 59:95–117.

Carr SM. 2020. Evidence for the persistence of ancient Beothuk and Maritime Archaic mitochondrial DNA genome lineages among modern Native American peoples. Genome. 63:349-355.

Carr SM. 2023. Things Fall Apart: Multiple mitogenomes provide an alternative evolutionary phylogeography to Thrice Out of Africa or Asia hypotheses for Western honey bees (*Apis mellifera*). Nat Sci Reps. 13:9386-9398.

Kim JS, Wang AR, Kim MJ, Lee HL, Kim I. 2019. Single-nucleotide polymorphism markers in mitochondrial genomes for identifying *Varroa* destructor-resistant and -susceptible strains of *Apis mellifera* (Hymenoptera: Apidae), Mitochondrial DNA Part A 30:477-489.

Johnson BR. 2023. Honey Bee Biology. Princeton University Press. Xi + 481.

Ma C, Hu R, Costa C, Li J. 2022. Genetic Drift and Purifying Selection Shaped Mitochondrial Genome Variation in the High Royal Jelly-Producing Honeybee Strain (*Apis mellifera ligustica*). Front Genet. 13:835-967.

Ma C, Li J. 2021. Characterization of the Mitochondrial Genome of a High Royal Jelly-Producing Honeybee Strain (*Apis mellifera ligustica*). Mitochondrial DNA Part B 6:2939–2940.

Seeley TD. 2019. The Lives of Bees. Princeton University Press. Xiii + 351.

Tamura K, Stecher G, Kumar S. 2021. MEGA 11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol. 38:3022-3027.

Wilson AC Wilson, Cann RL, Carr SM, George M, Gyllensten UB, Helm-Bychowski KM, Higuchi RG, Palumbi SR, Prager EM, Sage RD, Stoneking M. 1985. Mitochondrial DNA and two perspectives on evolution. Biol J Linnaen Soc.26:375-400.

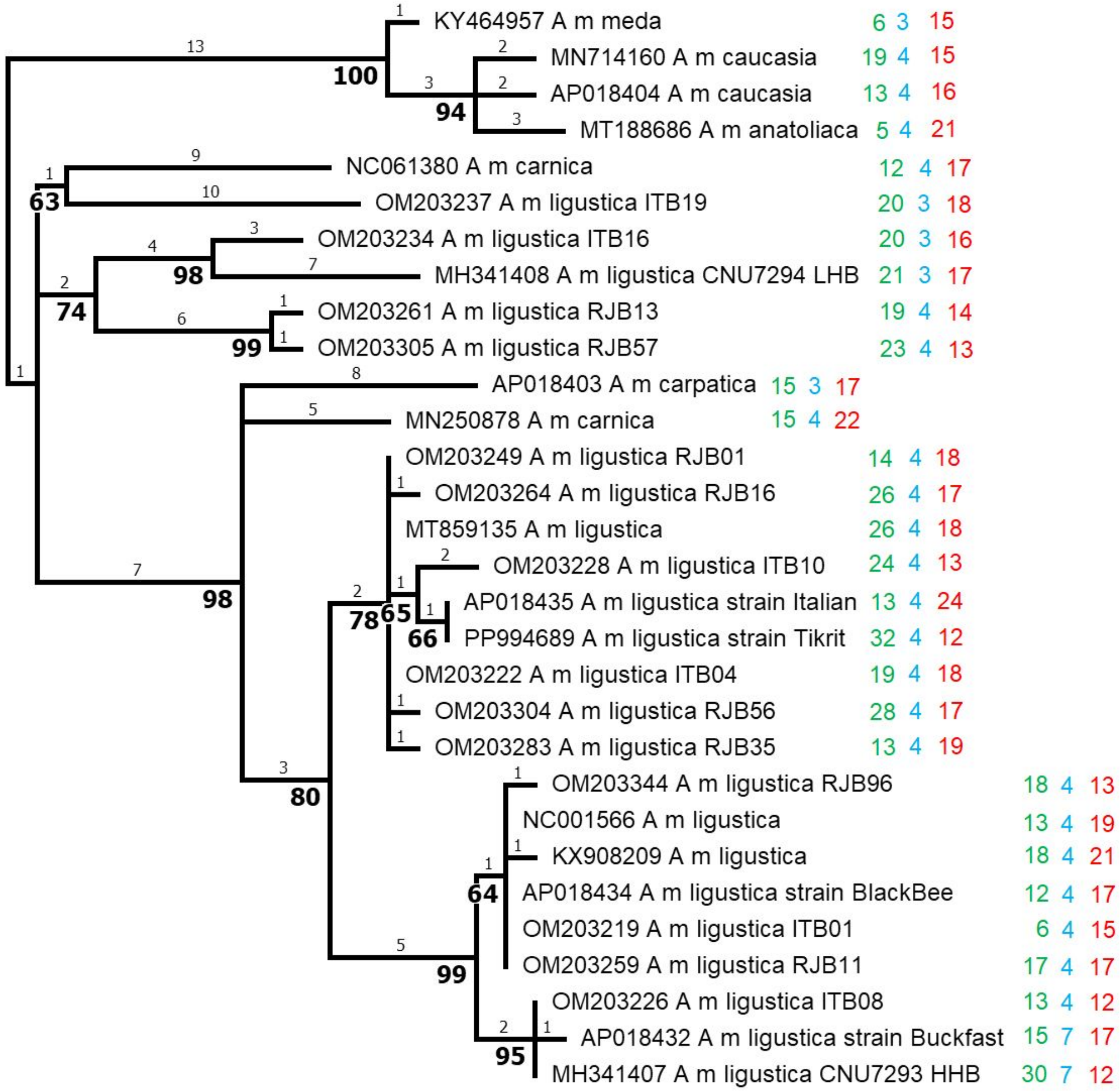


Figure 1 - Maximum Parsimony tree of 11,070 bp (3,690 protein-coding DNA triplets) in 23 sequences from *A. m. ligustica* and seven from five related subspecies. 3,000 bootstrap replicates below branches, SNP changes above. Tree rooted with four sequences from the Asia Minor clade of Carr (2023): see text. *A. m. carnica* and *A. m. carpatica* in the Southeast Europe clade are included within *A. m. ligustica* sensu lato. The first two clades include five *A. m.*

ligustica sequences including **LHB**, versus 18 in the major derived clade including **HHB** (98% support).

Color codes indicate repeat numbers at three positions, 6.1kb (**AT**), 10.1kb (**AAT**), and 14.7kb (**TA**). See Figure 3 for analysis of the **AT** repeats, and Supplementary Figure S2 for analysis of the **AAT** and **TA** repeats.

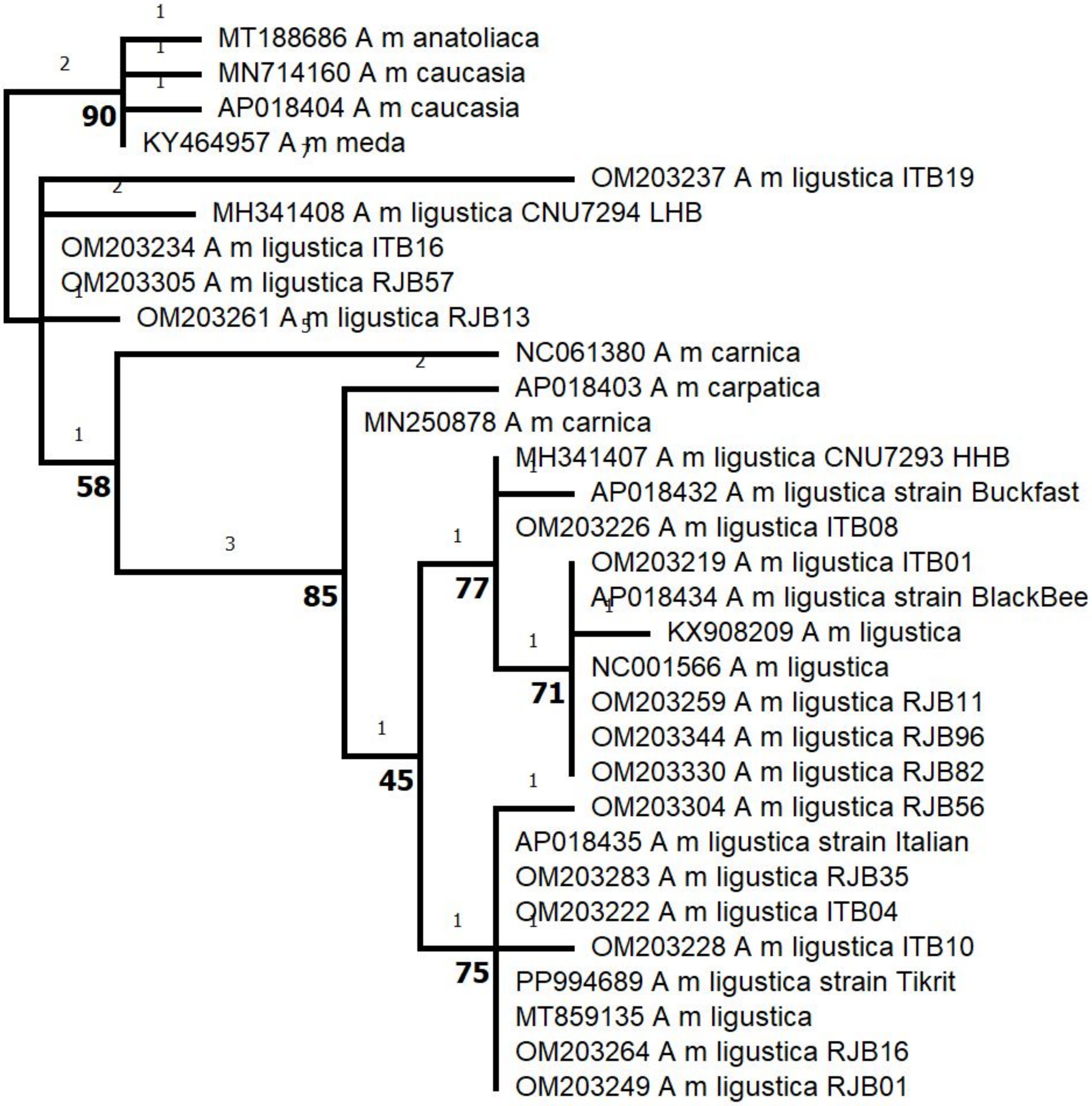


Figure 2 - Maximum Parsimony tree of 3,690 amino acid residues in protein coding regions.

Sequences, rooting, bootstrap support, and branch length as in Figure 1. The clades inclusive of **HHB** in the amino acid and SNP data (Figure 1) are essentially identical. Relationships of other *A. m. ligustica* lineages including **LHB** are unresolved by the amino acid data.

The basal group includes identical pairs of **ITB** and **RJB** sequences, as do the major clade (85% bootstrap) and subclades (77% and 75% bootstrap) homologous to those in the SNP tree (98%, 79%, and 99% bootstrap, respectively).

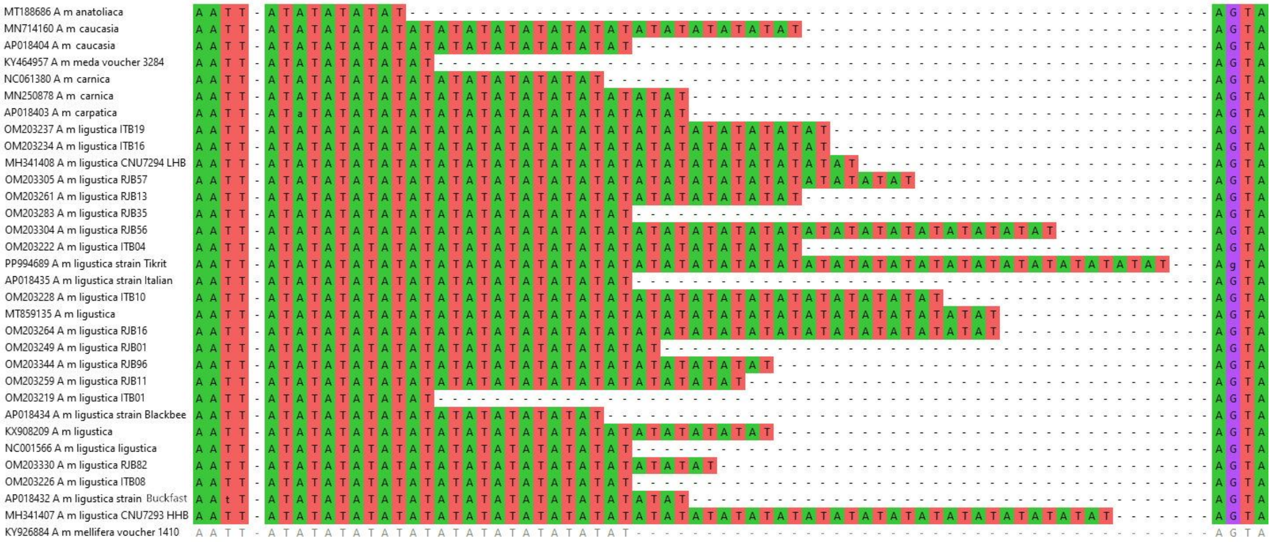
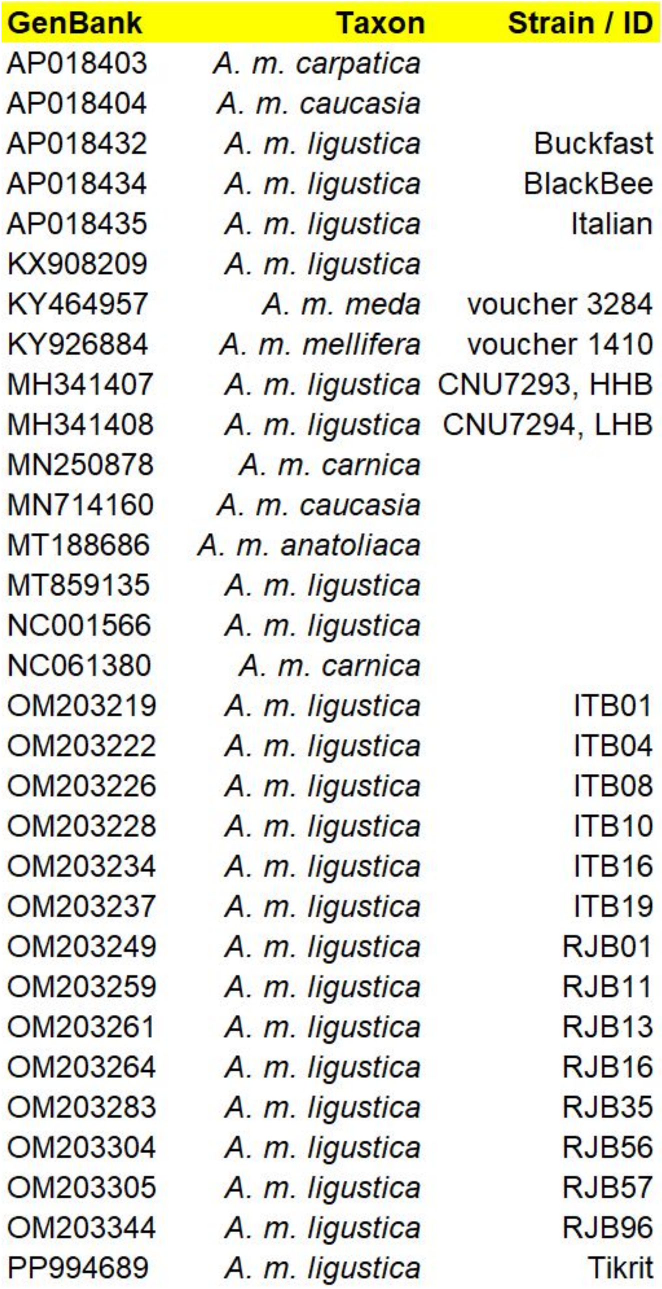
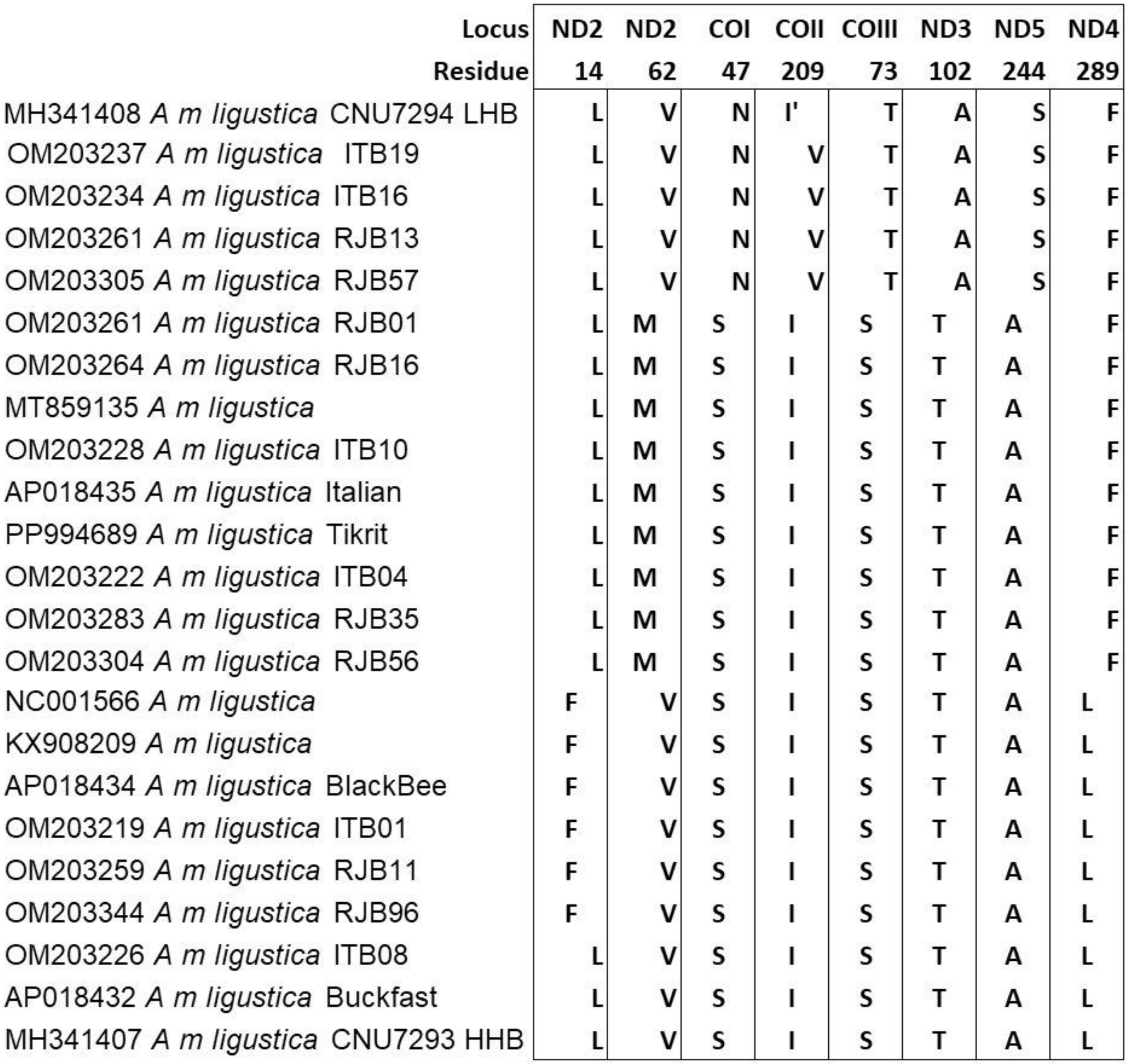


Figure 3. Variation in number of AT repeats near position 6.1 Kb in *A. m. ligustica* and related subspecies. See counts in Figure 1. *A. m. mellifera* is typical with **AT$_{12}$**. The shortest are **AT$_{4}$** in *A. m. anatoliaca* and **AT$_{5}$** in *A. m. meda* in the Asia Minor subspecies group. The longest is **AT$_{31}$** in **Tikrit**. Compare **AT$_{29}$** and **AT$_{20}$** in **HGB** and **LGB**, respectively: both have gained high numbers of repeats, in parallel (cf. Figure 1). See text and Supplementary Figure S2 for discussion and diagrams of **AAT** and **AT** repeat regions.



GenBank	Taxon	Strain / ID
AP018403	*A. m. carpatica*	
AP018404	*A. m. caucasia*	
AP018432	*A. m. ligustica*	Buckfast
AP018434	*A. m. ligustica*	BlackBee
AP018435	*A. m. ligustica*	Italian
KX908209	*A. m. ligustica*	
KY464957	*A. m. meda*	voucher 3284
KY926884	*A. m. mellifera*	voucher 1410
MH341407	*A. m. ligustica*	CNU7293, HHB
MH341408	*A. m. ligustica*	CNU7294, LHB
MN250878	*A. m. carnica*	
MN714160	*A. m. caucasia*	
MT188686	*A. m. anatoliaca*	
MT859135	*A. m. ligustica*	
NC001566	*A. m. ligustica*	
NC061380	*A. m. carnica*	
OM203219	*A. m. ligustica*	ITB01
OM203222	*A. m. ligustica*	ITB04
OM203226	*A. m. ligustica*	ITB08
OM203228	*A. m. ligustica*	ITB10
OM203234	*A. m. ligustica*	ITB16
OM203237	*A. m. ligustica*	ITB19
OM203249	*A. m. ligustica*	RJB01
OM203259	*A. m. ligustica*	RJB11
OM203261	*A. m. ligustica*	RJB13
OM203264	*A. m. ligustica*	RJB16
OM203283	*A. m. ligustica*	RJB35
OM203304	*A. m. ligustica*	RJB56
OM203305	*A. m. ligustica*	RJB57
OM203344	*A. m. ligustica*	RJB96
PP994689	*A. m. ligustica*	Tikrit

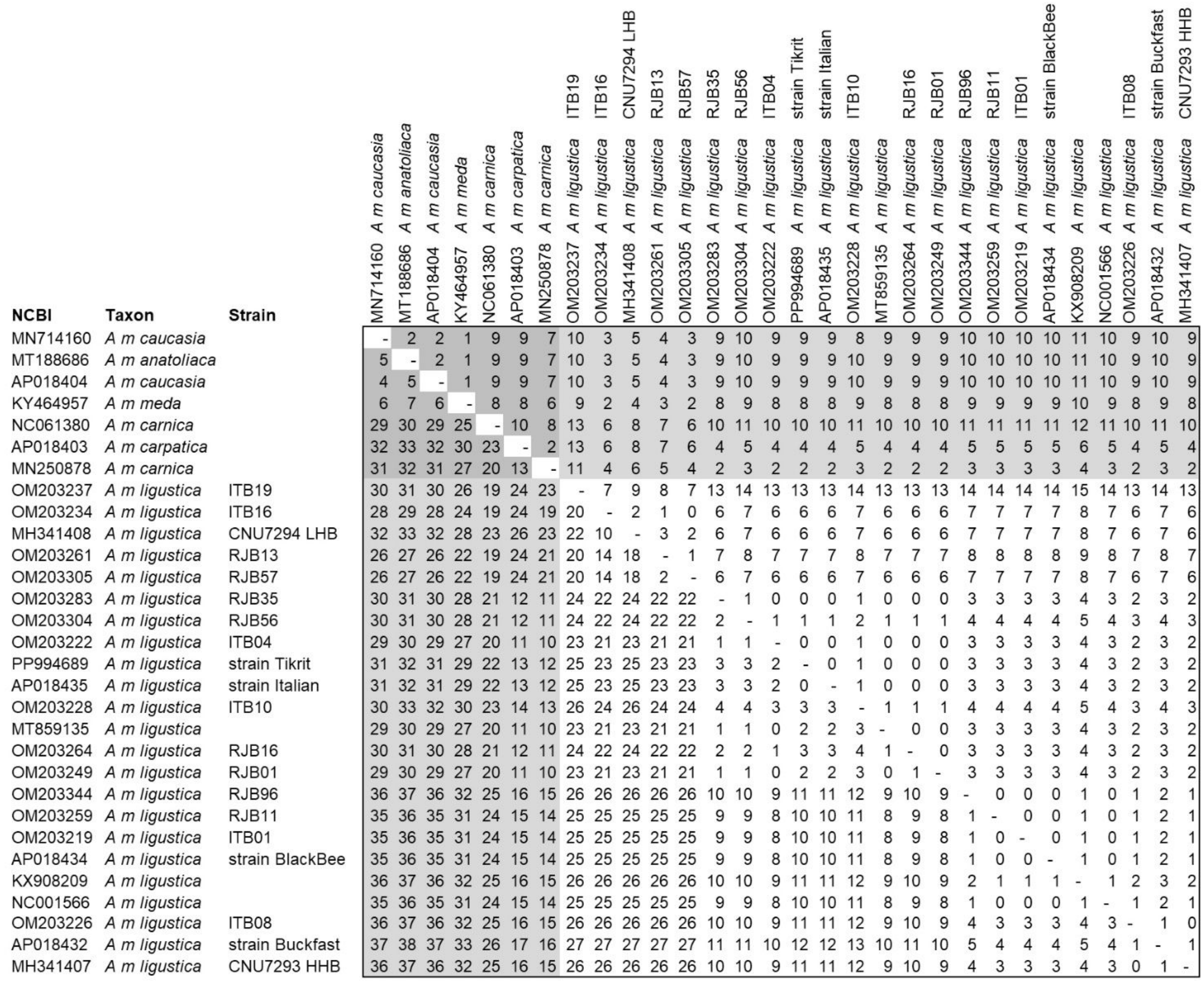
Table 1 – GenBank accession numbers, taxon referrals, and strain or series IDs for 31 mtDNA genome sequences used in this study. Note the single representatives of the HHB and LHB strains from Kim et al. (2019), 14 representatives of the ITB (n = 6) and RJB (n = 8) from Ma et al. (2022), n = 7 additional sequences referred to *A. m. ligustica*, and n = 8 sequences referred to six other subspecies.



Locus	ND2	ND2	COI	COII	COIII	ND3	ND5	ND4
Residue	14	62	47	209	73	102	244	289
MH341408 *A m ligustica* CNU7294 LHB	L	V	N	I'	T	A	S	F
OM203237 *A m ligustica* ITB19	L	V	N	V	T	A	S	F
OM203234 *A m ligustica* ITB16	L	V	N	V	T	A	S	F
OM203261 *A m ligustica* RJB13	L	V	N	V	T	A	S	F
OM203305 *A m ligustica* RJB57	L	V	N	V	T	A	S	F
OM203261 *A m ligustica* RJB01	L	M	S	I	S	T	A	F
OM203264 *A m ligustica* RJB16	L	M	S	I	S	T	A	F
MT859135 *A m ligustica*	L	M	S	I	S	T	A	F
OM203228 *A m ligustica* ITB10	L	M	S	I	S	T	A	F
AP018435 *A m ligustica* Italian	L	M	S	I	S	T	A	F
PP994689 *A m ligustica* Tikrit	L	M	S	I	S	T	A	F
OM203222 *A m ligustica* ITB04	L	M	S	I	S	T	A	F
OM203283 *A m ligustica* RJB35	L	M	S	I	S	T	A	F
OM203304 *A m ligustica* RJB56	L	M	S	I	S	T	A	F
NC001566 *A m ligustica*	F	V	S	I	S	T	A	L
KX908209 *A m ligustica*	F	V	S	I	S	T	A	L
AP018434 *A m ligustica* BlackBee	F	V	S	I	S	T	A	L
OM203219 *A m ligustica* ITB01	F	V	S	I	S	T	A	L
OM203259 *A m ligustica* RJB11	F	V	S	I	S	T	A	L
OM203344 *A m ligustica* RJB96	F	V	S	I	S	T	A	L
OM203226 *A m ligustica* ITB08	L	V	S	I	S	T	A	L
AP018432 *A m ligustica* Buckfast	L	V	S	I	S	T	A	L
MH341407 *A m ligustica* CNU7293 HHB	L	V	S	I	S	T	A	L

Table 2. Phylogenetic distribution of eight phylogenetically informative amino acid substitutions in *A. m. ligustica*. IUPAC single-letter codes. For each position, the inferred apomorphic state as inferred from Figures 1 & 2 is offset to the left. Residues 3 and 5 – 8 show the symplesiomorphic state in **LHB** (and four other sequences), versus the synapomorphic state inclusive of **HHB** (and 18 other sequences), or a subset at Residue 8. Positions 1 and 2 separate synapomorphic states in subclades within that clade not inclusive of **HHB**. Position 4 requires a

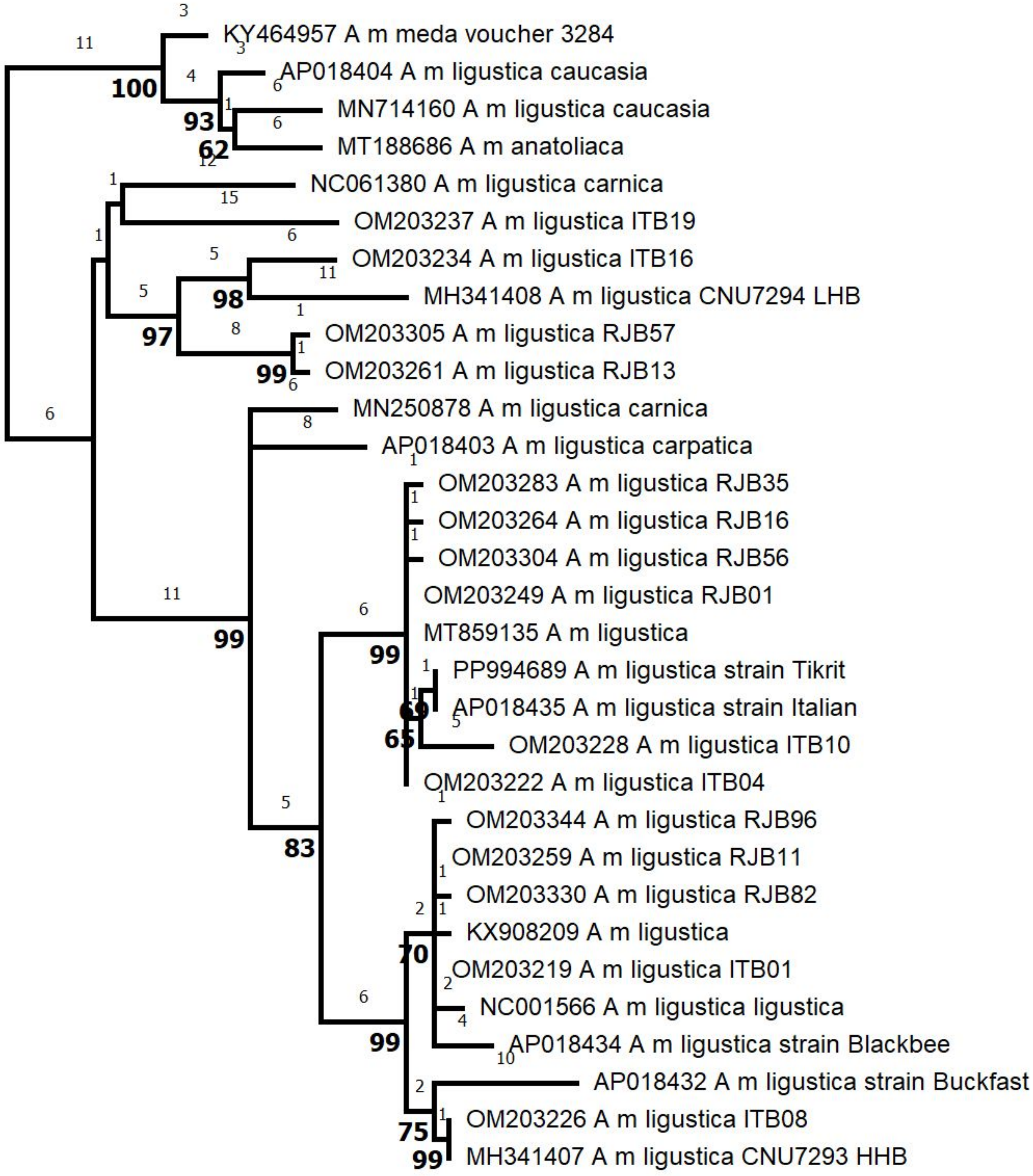
parallel change (2nd position **A** → **G**: **V** → **I**) in **LHB** with respect to **HHB** in the major derived clade.



NCBI	Taxon	Strain	MN714160 *A m caucasia*	MT188686 *A m anatoliaca*	AP018404 *A m caucasia*	KY464957 *A m meda*	NC061380 *A m carnica*	AP018403 *A m carpatica*	MN250878 *A m carnica*	OM203237 *A m ligustica* ITB19	OM203234 *A m ligustica* ITB16	MH341408 *A m ligustica* CNU7294 LHB	OM203261 *A m ligustica* RJB13	OM203305 *A m ligustica* RJB57	OM203283 *A m ligustica* RJB35	OM203304 *A m ligustica* RJB56	OM203222 *A m ligustica* ITB04	PP994689 *A m ligustica* strain Tikrit	AP018435 *A m ligustica* strain Italian	OM203228 *A m ligustica* ITB10	MT859135 *A m ligustica*	OM203264 *A m ligustica* RJB16	OM203249 *A m ligustica* RJB01	OM203344 *A m ligustica* RJB96	OM203259 *A m ligustica* RJB11	OM203219 *A m ligustica* ITB01	AP018434 *A m ligustica* strain BlackBee	KX908209 *A m ligustica*	NC001566 *A m ligustica*	OM203226 *A m ligustica* ITB08	AP018432 *A m ligustica* strain Buckfast	MH341407 *A m ligustica* CNU7293 HHB
MN714160	*A m caucasia*		-	2	2	1	9	9	7	10	3	5	4	3	9	10	9	9	9	8	9	9	9	10	10	10	10	11	10	9	10	9
MT188686	*A m anatoliaca*		5	-	2	1	9	9	7	10	3	5	4	3	9	10	9	9	9	10	9	9	9	10	10	10	10	11	10	9	10	9
AP018404	*A m caucasia*		4	5	-	1	9	9	7	10	3	5	4	3	9	10	9	9	9	10	9	9	9	10	10	10	10	11	10	9	10	9
KY464957	*A m meda*		6	7	6	-	8	8	6	9	2	4	3	2	8	9	8	8	8	9	8	8	8	9	9	9	9	10	9	8	9	8
NC061380	*A m carnica*		29	30	29	25	-	10	8	13	6	8	7	6	10	11	10	10	10	11	10	10	10	11	11	11	11	12	11	10	11	10
AP018403	*A m carpatica*		32	33	32	30	23	-	2	13	6	8	7	6	4	5	4	4	4	5	4	4	4	5	5	5	5	6	5	4	5	4
MN250878	*A m carnica*		31	32	31	27	20	13	-	11	4	6	5	4	2	3	2	2	2	3	2	2	2	3	3	3	3	4	3	2	3	2
OM203237	*A m ligustica*	ITB19	30	31	30	26	19	24	23	-	7	9	8	7	13	14	13	13	13	14	13	13	13	14	14	14	14	15	14	13	14	13
OM203234	*A m ligustica*	ITB16	28	29	28	24	19	24	19	20	-	2	1	0	6	7	6	6	6	7	6	6	6	7	7	7	7	8	7	6	7	6
MH341408	*A m ligustica*	CNU7294 LHB	32	33	32	28	23	26	23	22	10	-	3	2	6	7	6	6	6	7	6	6	6	7	7	7	7	8	7	6	7	6
OM203261	*A m ligustica*	RJB13	26	27	26	22	19	24	21	20	14	18	-	1	7	8	7	7	7	8	7	7	7	8	8	8	8	9	8	7	8	7
OM203305	*A m ligustica*	RJB57	26	27	26	22	19	24	21	20	14	18	2	-	6	7	6	6	6	7	6	6	6	7	7	7	7	8	7	6	7	6
OM203283	*A m ligustica*	RJB35	30	31	30	28	21	12	11	24	22	24	22	22	-	1	0	0	0	1	0	0	0	3	3	3	3	4	3	2	3	2
OM203304	*A m ligustica*	RJB56	30	31	30	28	21	12	11	24	22	24	22	22	2	-	1	1	1	2	1	1	1	4	4	4	4	5	4	3	4	3
OM203222	*A m ligustica*	ITB04	29	30	29	27	20	11	10	23	21	23	21	21	1	1	-	0	0	1	0	0	0	3	3	3	3	4	3	2	3	2
PP994689	*A m ligustica*	strain Tikrit	31	32	31	29	22	13	12	25	23	25	23	23	3	3	2	-	0	1	0	0	0	3	3	3	3	4	3	2	3	2
AP018435	*A m ligustica*	strain Italian	31	32	31	29	22	13	12	25	23	25	23	23	3	3	2	0	-	1	0	0	0	3	3	3	3	4	3	2	3	2
OM203228	*A m ligustica*	ITB10	30	33	32	30	23	14	13	26	24	26	24	24	4	4	3	3	3	-	1	1	1	4	4	4	4	5	4	3	4	3
MT859135	*A m ligustica*		29	30	29	27	20	11	10	23	21	23	21	21	1	1	0	2	2	3	-	0	0	3	3	3	3	4	3	2	3	2
OM203264	*A m ligustica*	RJB16	30	31	30	28	21	12	11	24	22	24	22	22	2	2	1	3	3	4	1	-	0	3	3	3	3	4	3	2	3	2
OM203249	*A m ligustica*	RJB01	29	30	29	27	20	11	10	23	21	23	21	21	1	1	0	2	2	3	0	1	-	3	3	3	3	4	3	2	3	2
OM203344	*A m ligustica*	RJB96	36	37	36	32	25	16	15	26	26	26	26	26	10	10	9	11	11	12	9	10	9	-	0	0	0	1	0	1	2	1
OM203259	*A m ligustica*	RJB11	35	36	35	31	24	15	14	25	25	25	25	25	9	9	8	10	10	11	8	9	8	1	-	0	0	1	0	1	2	1
OM203219	*A m ligustica*	ITB01	35	36	35	31	24	15	14	25	25	25	25	25	9	9	8	10	10	11	8	9	8	1	0	-	0	1	0	1	2	1
AP018434	*A m ligustica*	strain BlackBee	35	36	35	31	24	15	14	25	25	25	25	25	9	9	8	10	10	11	8	9	8	1	0	0	-	1	0	1	2	1
KX908209	*A m ligustica*		36	37	36	32	25	16	15	26	26	26	26	26	10	10	9	11	11	12	9	10	9	2	1	1	1	-	1	2	3	2
NC001566	*A m ligustica*		35	36	35	31	24	15	14	25	25	25	25	25	9	9	8	10	10	11	8	9	8	1	0	0	0	1	-	1	2	1
OM203226	*A m ligustica*	ITB08	36	37	36	32	25	16	15	26	26	26	26	26	10	10	9	11	11	12	9	10	9	4	3	3	3	4	3	-	1	0
AP018432	*A m ligustica*	strain Buckfast	37	38	37	33	26	17	16	27	27	27	27	27	11	11	10	12	12	13	10	11	10	5	4	4	4	5	4	1	-	1
MH341407	*A m ligustica*	CNU7293 HHB	36	37	36	32	25	16	15	26	26	26	26	26	10	10	9	11	11	12	9	10	9	4	3	3	3	4	3	0	1	-

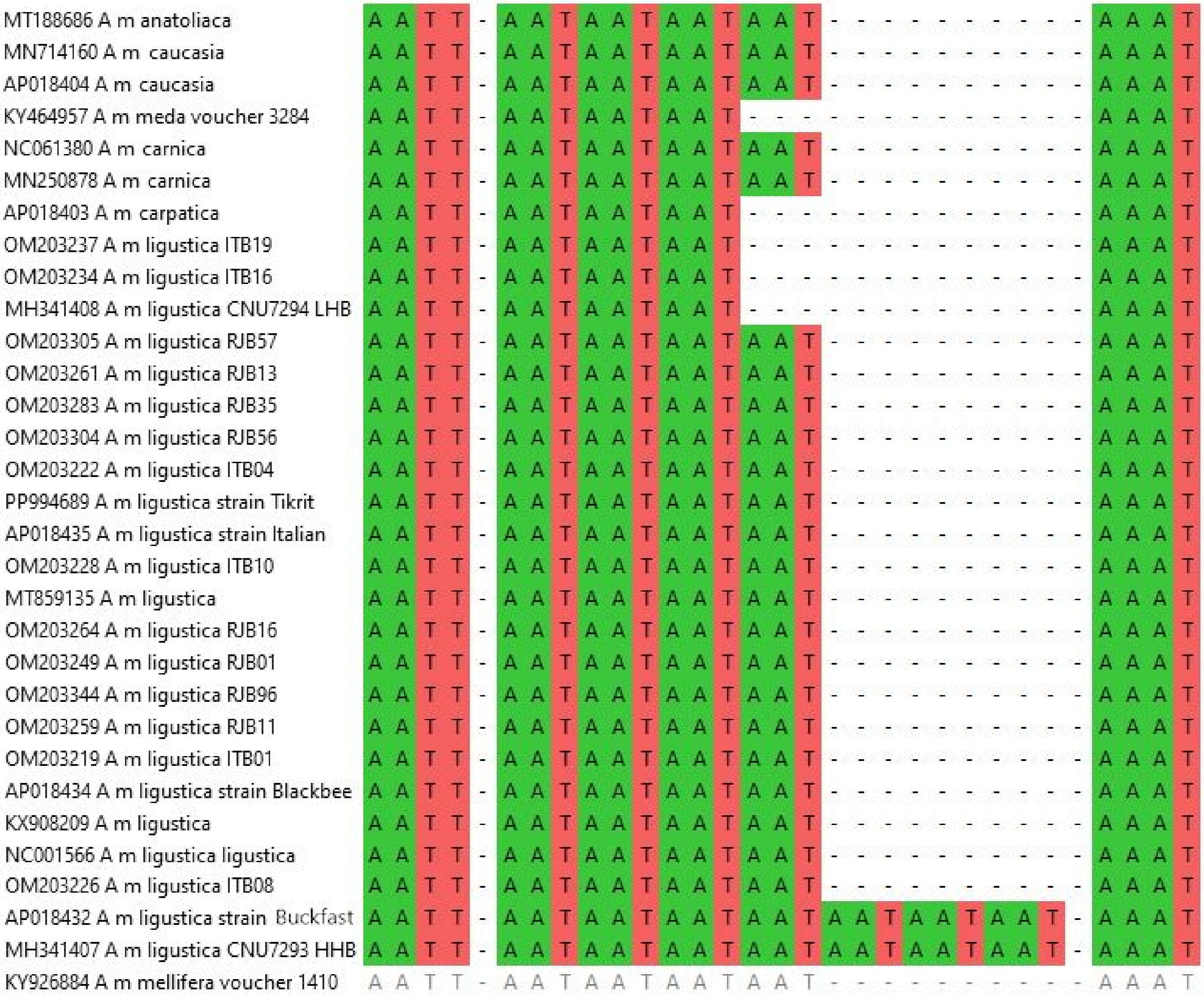
Table 3. Pairwise SNP differences and amino acid substitutions among protein-coding nucleotide regions of mitogenomes from *A. m. ligustica* and five related subspecies. Lower and upper triangular matrices, respectively. Inter-subspecies numbers in dark grey; inter-subspecies differences including *A. m. ligustica* in light grey. There are 26 pairwise SNP differences and six amino acid substitutions between **HHB** and **LHB**, versus intra-subspecies means of 12.7 ± 9.3 and 4.18 ± 3.59, respectively.

Supplementary Figures and Tables.



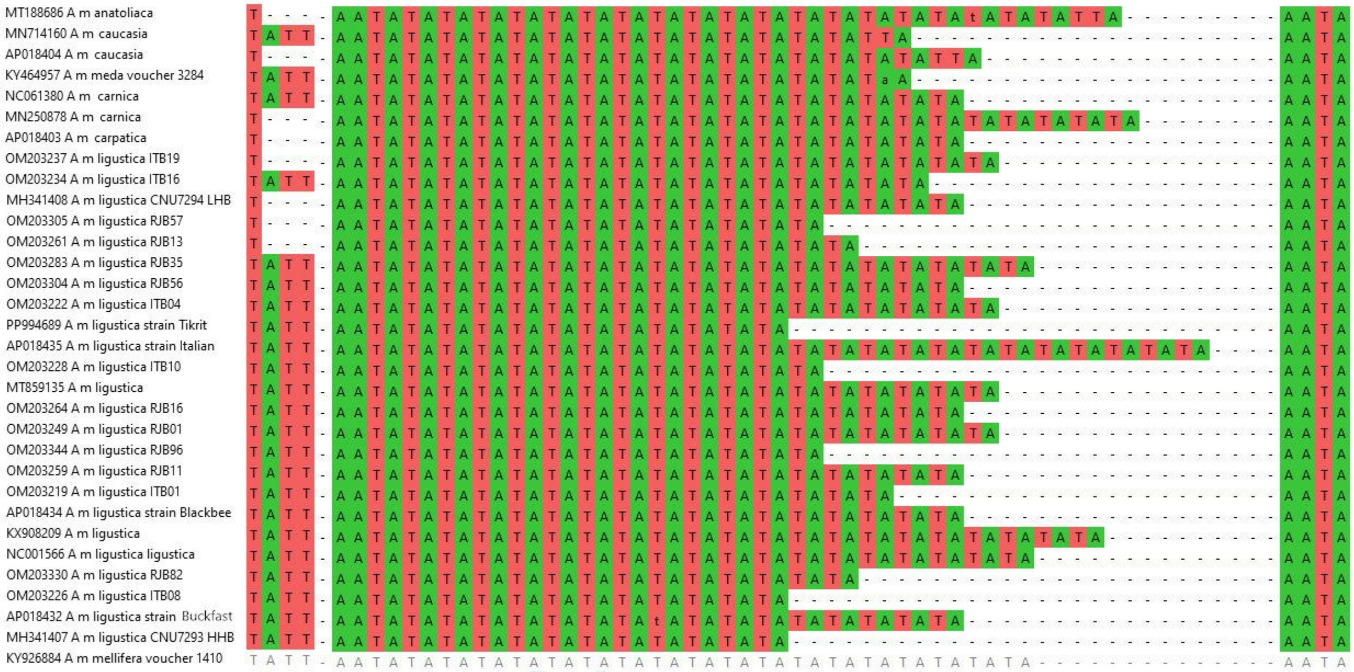
Supplementary Figure S1: **Maximum Parsimony analysis of complete mtDNA sequences.** Tritomys in Figure 1 are resolved without statistical support, and bootstrap support for various clades is increased.

A

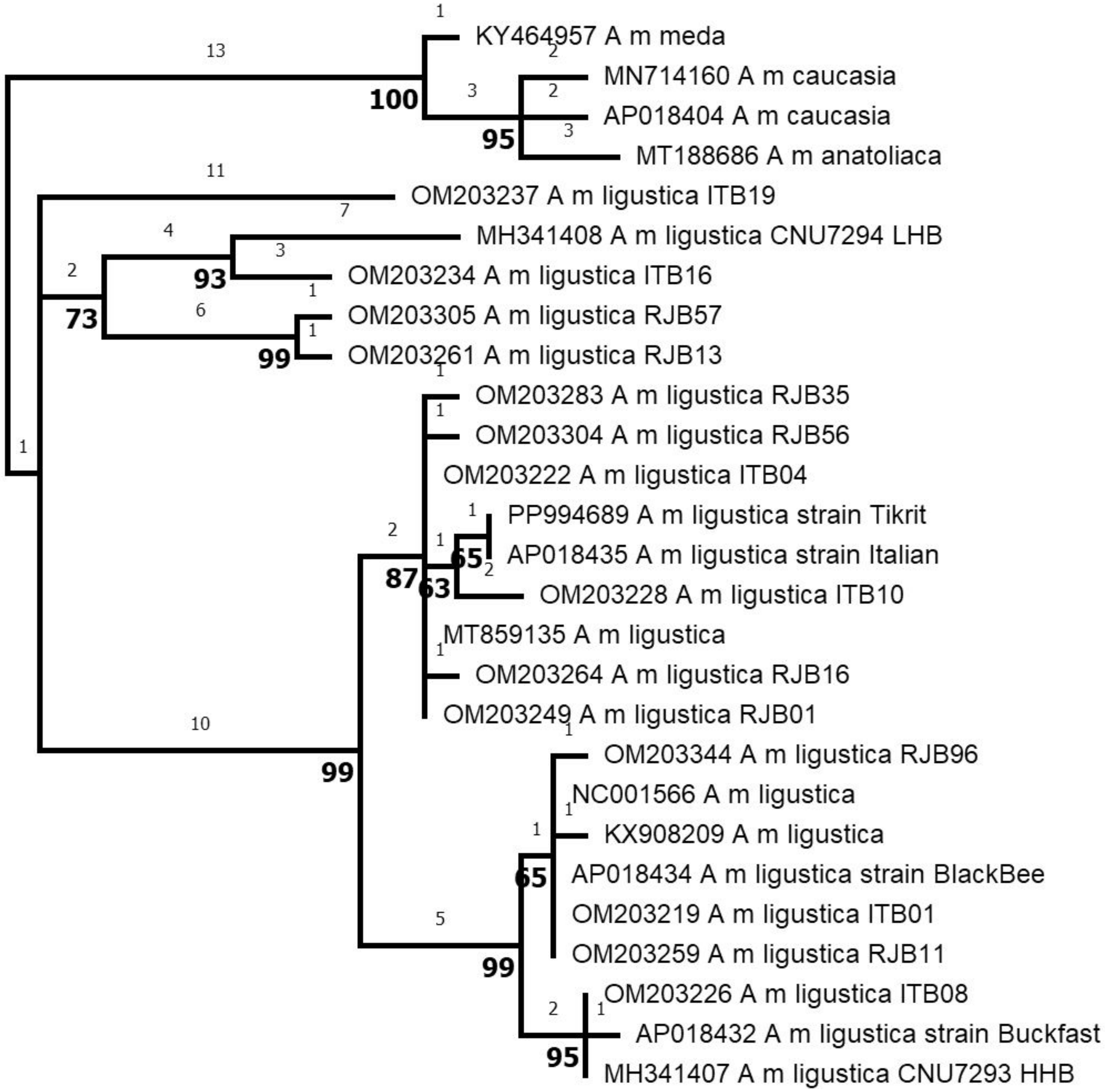


MT188686 A m anatoliaca	AATT-AATAATAATAAT----------AAAT
MN714160 A m caucasia	AATT-AATAATAATAAT----------AAAT
AP018404 A m caucasia	AATT-AATAATAATAAT----------AAAT
KY464957 A m meda voucher 3284	AATT-AATAATAAT-------------AAAT
NC061380 A m carnica	AATT-AATAATAATAAT----------AAAT
MN250878 A m carnica	AATT-AATAATAATAAT----------AAAT
AP018403 A m carpatica	AATT-AATAATAAT-------------AAAT
OM203237 A m ligustica ITB19	AATT-AATAATAAT-------------AAAT
OM203234 A m ligustica ITB16	AATT-AATAATAAT-------------AAAT
MH341408 A m ligustica CNU7294 LHB	AATT-AATAATAAT-------------AAAT
OM203305 A m ligustica RJB57	AATT-AATAATAATAAT----------AAAT
OM203261 A m ligustica RJB13	AATT-AATAATAATAAT----------AAAT
OM203283 A m ligustica RJB35	AATT-AATAATAATAAT----------AAAT
OM203304 A m ligustica RJB56	AATT-AATAATAATAAT----------AAAT
OM203222 A m ligustica ITB04	AATT-AATAATAATAAT----------AAAT
PP994689 A m ligustica strain Tikrit	AATT-AATAATAATAAT----------AAAT
AP018435 A m ligustica strain Italian	AATT-AATAATAATAAT----------AAAT
OM203228 A m ligustica ITB10	AATT-AATAATAATAAT----------AAAT
MT859135 A m ligustica	AATT-AATAATAATAAT----------AAAT
OM203264 A m ligustica RJB16	AATT-AATAATAATAAT----------AAAT
OM203249 A m ligustica RJB01	AATT-AATAATAATAAT----------AAAT
OM203344 A m ligustica RJB96	AATT-AATAATAATAAT----------AAAT
OM203259 A m ligustica RJB11	AATT-AATAATAATAAT----------AAAT
OM203219 A m ligustica ITB01	AATT-AATAATAATAAT----------AAAT
AP018434 A m ligustica strain Blackbee	AATT-AATAATAATAAT----------AAAT
KX908209 A m ligustica	AATT-AATAATAATAAT----------AAAT
NC001566 A m ligustica ligustica	AATT-AATAATAATAAT----------AAAT
OM203226 A m ligustica ITB08	AATT-AATAATAATAAT----------AAAT
AP018432 A m ligustica strain Buckfast	AATT-AATAATAATAATAATAATAAT-AAAT
MH341407 A m ligustica CNU7293 HHB	AATT-AATAATAATAATAATAATAAT-AAAT
KY926884 A m mellifera voucher 1410	AATT-AATAATAATAAT----------AAAT

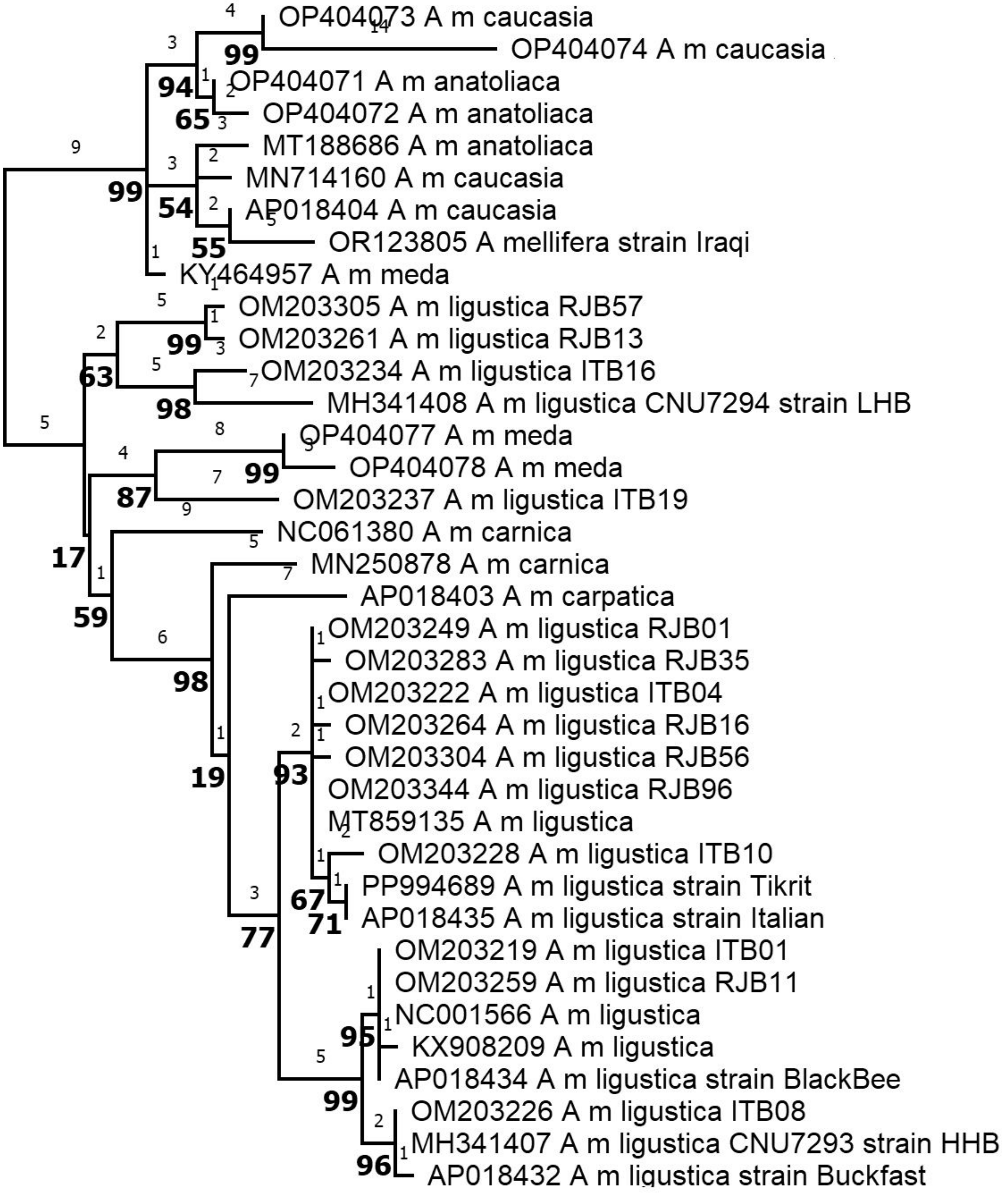
B



Supplementary Figure S2: **Patterns of short repeats across subspecies of *A. mellifera* and within *A. m. ligustica***. (**A**) **AAT** repeat region at 10.1 kb, and (**B**) **TA** repeat region at 14.6 kb. See distribution of counts for these and the **AT** repeat at 6.1 kb on Figure 1.

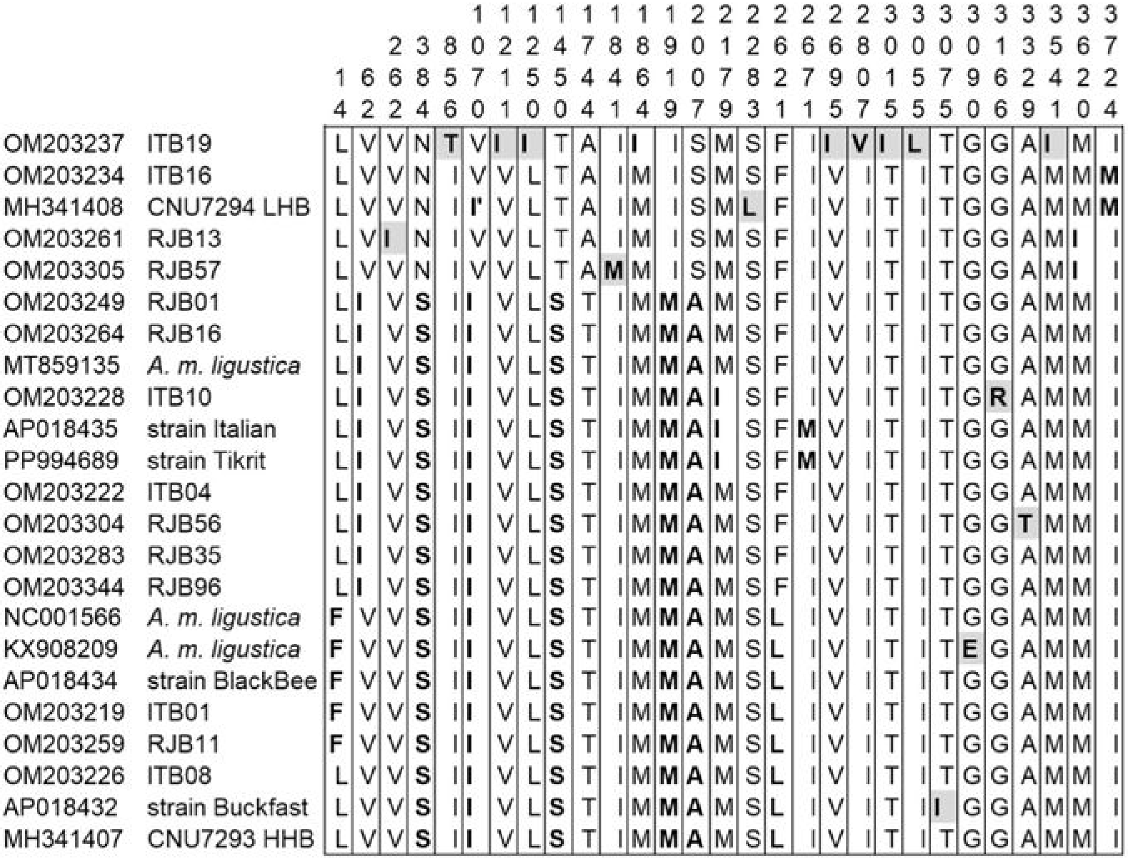


Supplementary Figure S3: Maximum Parsimony analysis of complete protein-coding nucleotide sequences, excluding *A. m. carnica* and *A. m. carpatica*. Support for the derived clade of *A. m. ligustica* sequences increases to ten SNPs (99% bootstrap support), and for the **HHB**-inclusive clade to five SNPs (99%). Support for the derived clade including **HHB** in the analogous amino acid-based tree increases to four substitutions (98%) (not shown).



Supplementary Figure S4: Names Matter - Maximum Parsimony analysis of complete mtDNA sequences from all GenBank accessions referred to taxa in the Southeast European

and Asia Minor clades of Carr (2023). Note polyphyletic distributions of sequences in the OP40407# series with respect to sequences referred to the same subspecific taxa. Note especially OP404077 & OP404078 referred to *A. m. meda* with respect to *A. m. ligustica* (**ITB19**). See main text for discussion of sequences referred to *A. m. carnica* and *A. m. carpatica.*



		14	62	262	384	856	1070	1211	1250	1450	1744	1841	1864	1949	2107	2179	2283	2621	2671	2695	2807	3015	3055	3075	3090	3166	3329	3541	3620	3724
OM203237	ITB19	L	V	V	N	**T**	V	**I**	**I**	T	A	I	**I**	I	S	M	S	F	I	**I**	**V**	**I**	**L**	T	G	G	A	**I**	M	I
OM203234	ITB16	L	V	V	N	I	V	V	L	T	A	I	M	I	S	M	S	F	I	V	I	T	I	T	G	G	A	M	M	**M**
MH341408	CNU7294 LHB	L	V	V	N	I	**I'**	V	L	T	A	I	M	I	S	M	**L**	F	I	V	I	T	I	T	G	G	A	M	M	**M**
OM203261	RJB13	L	V	**I**	N	I	V	V	L	T	A	I	M	I	S	M	S	F	I	V	I	T	I	T	G	G	A	M	**I**	I
OM203305	RJB57	L	V	V	N	I	V	V	L	T	A	**M**	M	I	S	M	S	F	I	V	I	T	I	T	G	G	A	M	**I**	I
OM203249	RJB01	L	**I**	V	**S**	I	**I**	V	L	**S**	T	I	M	**M**	**A**	M	S	F	I	V	I	T	I	T	G	G	A	M	M	I
OM203264	RJB16	L	**I**	V	**S**	I	**I**	V	L	**S**	T	I	M	**M**	**A**	M	S	F	I	V	I	T	I	T	G	G	A	M	M	I
MT859135	*A. m. ligustica*	L	**I**	V	**S**	I	**I**	V	L	**S**	T	I	M	**M**	**A**	M	S	F	I	V	I	T	I	T	G	G	A	M	M	I
OM203228	ITB10	L	**I**	V	**S**	I	**I**	V	L	**S**	T	I	M	**M**	**A**	**I**	S	F	I	V	I	T	I	T	G	**R**	A	M	M	I
AP018435	strain Italian	L	**I**	V	**S**	I	**I**	V	L	**S**	T	I	M	**M**	**A**	**I**	S	F	**M**	V	I	T	I	T	G	G	A	M	M	I
PP994689	strain Tikrit	L	**I**	V	**S**	I	**I**	V	L	**S**	T	I	M	**M**	**A**	**I**	S	F	**M**	V	I	T	I	T	G	G	A	M	M	I
OM203222	ITB04	L	**I**	V	**S**	I	**I**	V	L	**S**	T	I	M	**M**	**A**	M	S	F	I	V	I	T	I	T	G	G	A	M	M	I
OM203304	RJB56	L	**I**	V	**S**	I	**I**	V	L	**S**	T	I	M	**M**	**A**	M	S	F	I	V	I	T	I	T	G	G	**T**	M	M	I
OM203283	RJB35	L	**I**	V	**S**	I	**I**	V	L	**S**	T	I	M	**M**	**A**	M	S	F	I	V	I	T	I	T	G	G	A	M	M	I
OM203344	RJB96	L	**I**	V	**S**	I	**I**	V	L	**S**	T	I	M	**M**	**A**	M	S	F	I	V	I	T	I	T	G	G	A	M	M	I
NC001566	*A. m. ligustica*	**F**	V	V	**S**	I	**I**	V	L	**S**	T	I	M	**M**	**A**	M	S	**L**	I	V	I	T	I	T	G	G	A	M	M	I
KX908209	*A. m. ligustica*	**F**	V	V	**S**	I	**I**	V	L	**S**	T	I	M	**M**	**A**	M	S	**L**	I	V	I	T	I	T	**E**	G	A	M	M	I
AP018434	strain BlackBee	**F**	V	V	**S**	I	**I**	V	L	**S**	T	I	M	**M**	**A**	M	S	**L**	I	V	I	T	I	T	G	G	A	M	M	I
OM203219	ITB01	**F**	V	V	**S**	I	**I**	V	L	**S**	T	I	M	**M**	**A**	M	S	**L**	I	V	I	T	I	T	G	G	A	M	M	I
OM203259	RJB11	**F**	V	V	**S**	I	**I**	V	L	**S**	T	I	M	**M**	**A**	M	S	**L**	I	V	I	T	I	T	G	G	A	M	M	I
OM203226	ITB08	L	V	V	**S**	I	**I**	V	L	**S**	T	I	M	**M**	**A**	M	S	**L**	I	V	I	T	I	T	G	G	A	M	M	I
AP018432	strain Buckfast	L	V	V	**S**	I	**I**	V	L	**S**	T	I	M	**M**	**A**	M	S	**L**	I	V	I	T	I	**I**	G	G	A	M	M	I
MH341407	CNU7293 HHB	L	V	V	**S**	I	**I**	V	L	**S**	T	I	M	**M**	**A**	M	S	**L**	I	V	I	T	I	T	G	G	A	M	M	I

Supplementary Table T1: Amino acid substitutions in *A. m. ligustica*. Apomorphic states are inferred from Figures 1 & 2, and are offset to the left. There are 29 positions with amino acid substitutions, of which 15 are autapomorphic (bold, grey), including eight in **RGB19**, and 14 synapomorphic (Bold), including position #1070 with an inferred parallel change (**I'**) in **LHB**. See Table 1 for distribution of phylogenetically informative changes by position in the genome.