

ORIGINAL RESEARCH

Molecular evolution of bumble bee vitellogenin and vitellogenin-like genes

Fang Zhao¹ | Claire Morandin²  | Kai Jiang¹ | Tianjuan Su¹ | Bo He¹ | Gonghua Lin¹  | Zuhao Huang¹¹School of Life Sciences, Jinggangshan University, Ji'an, China²Department of Ecology and Evolution, Biophore, University of Lausanne, Lausanne, Switzerland**Correspondence**

Gonghua Lin and Zuhao Huang, School of Life Sciences, Jinggangshan University, Ji'an, China.

Email: lingonghua@163.com; hzhow@163.com

Funding information

Science and Technology Foundation of Jiangxi Provincial Department of Education, Grant/Award Number: GJJ201007; National Natural Science Foundation of China, Grant/Award Number: 31860105

Abstract

Vitellogenin (Vg), a storage protein, has been significantly studied for its egg yolk precursor role in oviparous animals. Recent studies found that vitellogenin and its Vg-like homologs were fundamentally involved in many other biological processes in social insects such as female caste differences and oxidative stress resilience. In this study, we conducted the first large-scale molecular evolutionary analyses of vitellogenin coding genes (*Vg*) and *Vg-like* genes of bumble bees, a primitively eusocial insect belonging to the genus *Bombus*. We obtained sequences for each of the four genes (*Vg*, *Vg-like-A*, *Vg-like-B*, and *Vg-like-C*) from 27 bumble bee genomes (nine were newly sequenced in this study), and sequences from the two closest clades of *Bombus*, including five *Apis* species and five *Tetragonula* species. Our molecular evolutionary analyses show that in bumble bee, the conventional *Vg* experienced strong positive selection, while the *Vg-like* genes showed overall relaxation of purifying selection. In *Apis* and *Tetragonula*; however, all four genes were found under purifying selection. Furthermore, the conventional *Vg* showed signs of strong positive selection in most subgenera in *Bombus*, apart from the obligate parasitic subgenus *Psithyrus* which has no caste differentiation. Together, these results indicate that the conventional *Vg*, a key pleiotropic gene in social insects, is the most rapidly evolving copy, potentially due to its multiple known social functions for both worker and queen castes. This study shows that concerted evolution and purifying selection shaped the evolution of the *Vg* gene family following their ancient gene duplication and may be the leading forces behind the evolution of new potential protein function enabling functional social pleiotropy.

KEYWORDSvitellogenin, *Bombus*, positive selection, functional pleiotropy

Fang Zhao and Claire Morandin are contributed equally.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

Vitellogenin (Vg) is a phospholipoglycoprotein, and the precursor protein of vitellin, a protein required for egg yolk formation by most oviparous species (Spieth et al., 1991). In insects, vitellogenin is synthesized in the fat body, released into the hemolymph, and finally taken up by developing oocytes to be consumed throughout embryogenesis (Hagedorn & Kunkel, 1979; Pan et al., 1969; Raikhel & Dhadialla, 1992; Tufail & Takeda, 2008). Despite its main egg yolk function, Vg is not female-specific and can also be found in males of some species although in smaller amount (Piulachs et al., 2003; Trenczek & Engels, 1986; Tufail & Takeda, 2008). Moreover, Vg has been extensively studied for its multifunctional effects in social insect life histories. In honey bee (*Apis mellifera*), in addition to its central involvement in the division of labor between queens and workers (Tufail & Takeda, 2008; Weil et al., 2009), Vg is known to be involved in regulation of nonreproductive features of colonies, such as aging and queen longevity (Corona et al., 2007; Excels, 1974), temporal worker division of labor (Bloch & Grozinger, 2011; Guidugli et al., 2005; Münch & Amdam, 2010; Nelson et al., 2007), and royal jelly production (Amdam et al., 2003).

In honey bee, only one conventional Vg gene can be found and has been extensively studied for its multiple phenotypic effects on both queen and worker traits (Amdam et al., 2003). Pleiotropic genes are expected to be evolutionarily constrained since mutations that increase fitness for one trait might decrease overall fitness via antagonistic effects on other traits (Otto, 2004). Interestingly however, many previous studies have not found support for a negative relationship between the pleiotropy of a given gene and its selection pressure, measured as the ratio of nonsynonymous to synonymous substitutions (dN/dS ; e.g., Razeto-Barry et al., 2011; Twyman et al., 2018; Vedanayagam & Garrigan, 2015). Vg displayed high rates of adaptive evolution, and positive selection signs of this gene were repeatedly detected in eusocial hymenopteran species (Kent et al., 2011; Morandin et al., 2014; Salmela et al., 2016).

Multiple roles for a single protein are also projected to lead to a gene duplication event and to favor the multiple acquired roles. Besides the conventional Vg, three homologs called Vg-like proteins were recently discovered in ants (Morandin et al., 2014). These Vg homologs have arisen from an ancient gene duplication event. Two of these homologs, Vg-like-A and Vg-like-B, can be found in all insect species studied, while Vg-like-C was so far only found in Hymenoptera (Kohlmeier et al., 2018; Morandin et al., 2014). These homologs exhibit differences in their conserved protein domains and have undergone rapid evolution after duplications (Morandin et al., 2014). Their role is currently unknown, but their structural variation suggests variable functions. In honey bee, Vg-like-A displays the closest structural and functional similarities to Vg and responded strongly to inflammatory and oxidative conditions, thus is likely associated with the aging process (Salmela et al., 2016). Vg-like-A also showed a strong temporal expression variation and may be involved in wintering worker longevity (Ricigliano et al., 2018). Furthermore, Vg-like-A is linked to the regulation of nursing behaviors in the

ant *Temnothorax longispinosus* (Kohlmeier et al., 2018). During duplication, Vg-like-B lost several Vg structural elements, which may suggest that Vg-like-B may perform only few of the Vg original functions, such as coping with oxidative stress (Morandin et al., 2014). Four protein domains (N-sheet, α -helical, vWFD, and polyserine linker) were found in Vg, while only the N-sheet was detected in Vg-like-C, potentially implying specialization, and could possibly be involved in neurobiological functions (Salmela et al., 2016).

Bumble bees are a group of insects belonging to the genus *Bombus* (Hymenoptera: Apidae). Bumble bees, honey bees, and stingless bees (*Tetragonula*) are phylogenetically close relatives (Peters et al., 2017). There are about 250 known bumble bee species belonging to 15 subgenera, mainly distributed in the northern hemisphere (Cameron et al., 2007; Williams et al., 2008). Bumble bees are often described as primitively eusocial because their social organization is simpler than that of the honeybee. Unlike honey bees or stingless bees, most bumble bee species have an annual cycle, with queens single-handedly founding nests (Goulson, 2010). Bumble bees pass through several distinct phases during their annual life cycle, including solitary and eusocial phases. At the final stage of their colony cycle, termed the competition phase, the queen and workers will compete intensely over the production of males (Amsalem et al., 2015).

The fascinating life history and high levels of biological and ecological heterogeneity make bumble bees an outstanding model system for the study of molecular evolution. First of all, the biological and ecological characteristics of bumble bees can largely differ among the different subgenera (Williams et al., 2008). For instance, the tongue lengths of bumble bees are very diversified among different subgenera. Some subgenera such as *Orientalibombus*, *Subterraneobombus*, and *Sibiricobombus* which favor deep flowers have very long tongues, while others have relatively short tongues (Williams et al., 2008). Bumble bees are also extremely diversified in their habitats: for example, *Mendacibombus* and *Alpinobombus* species prefer alpine/arctic, while *Orientalibombus* generally use forest habitats. Most strikingly, there are obligate parasitic species in the subgenus *Psithyrus* which enslaves the species from other subgenera (Amsalem et al., 2015). To our knowledge, such extreme biological and ecological diversifications within a single genus have not been previously reported in honey bees or stingless bees. Also, honey bees and stingless bees are primarily tropical insects, with relatively stable environments, while bumble bees mainly occur in cool climates in general with more variable environments. Furthermore, bumble bees biological and ecological characteristics may also deviate among different species within a single subgenus. For example, in at least seven subgenera, the distribution elevation significantly varies among different species (An et al., 2014), leading to distinct genomic evolution rates among species (Lin et al., 2019). Consequently, we hypothesize that Vg and Vg-like genes should be under distinct selection forces in bumble bees compared to honey bees or stingless bees.

To date, for most multi-species Vg analysis, bumble bee Vg sequences were compared with species from other genera (Du

et al., 2019; Li et al., 2010; Salmela et al., 2016), probably due to the lack of sequence data from multiple bumble bee species. In this study, based on 27 bumble bee genomes, and combined with genomes of their phylogenetically close relatives (honey bees and stingless bees), we conducted the first large-scale molecular evolutionary analyses of bumble bee *Vg* and *Vg-like* genes to understand the selective patterns of *Vg* gene family in bumble bees.

2 | MATERIALS AND METHODS

2.1 | DNA extraction and libraries construction

Adult workers from nine species of *Bombus* (*B. lantschouensis*, *B. sichelii*, *B. impetuosus*, *B. laesus*, *B. skorikovi*, *B. bohemicus*, *B. trifasciatus*, *B. waltoni*, and *B. convexus*) were live-trapped using sweep nets from the field. Samples of *B. trifasciatus* were collected in Ji'an, Jiangxi Province, China, while the other eight species were collected in Qinghai Province, China, between 2017 and 2019. After collection, samples were immediately stored at -20°C .

DNA extractions, libraries construction, and sequencing were done by the company Novogene. In brief, DNA was extracted from

whole-body of each individual (one worker per species) using a DNeasy Blood and Tissue Kit (Qiagen). Concentration and quality of extracted DNA was examined with an Agilent 2100 Bioanalyzer (Agilent Technologies). Libraries were sequenced paired-end for 150 cycles (1×150 bp) on an Illumina HiSeq 2000 system. Read quality was inspected with FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). Clean reads were used for de novo assembly using IDBA-UD (Peng et al., 2012) with default settings, and contigs shorter than 200 bp were discarded.

2.2 | Discovery of *Vg* and *Vg-like* sequences

In addition to our nine newly sequenced *Bombus* species, previously published bumble bee genomes were added to our dataset (see Jackson et al., 2020; Kent et al., 2018; Lin et al., 2019; Sadd et al., 2015; Tian et al., 2019). As a result, a total of 27 bumble bee species belonging to 10 subgenera (2 to 5 species for each genus) were used for this study (Figure 1; Table S1). The nucleotide sequences of *Vg*, *Vg-like-A*, *Vg-like-B*, and *Vg-like-C* of *B. impatiens* were downloaded from GenBank (*Vg*, XP_003492277.1; *Vg-like-A*, XP_003494500.3; *Vg-like-B*, XP_012245829.1; and *Vg-like-C*, XP_003489777.1). We

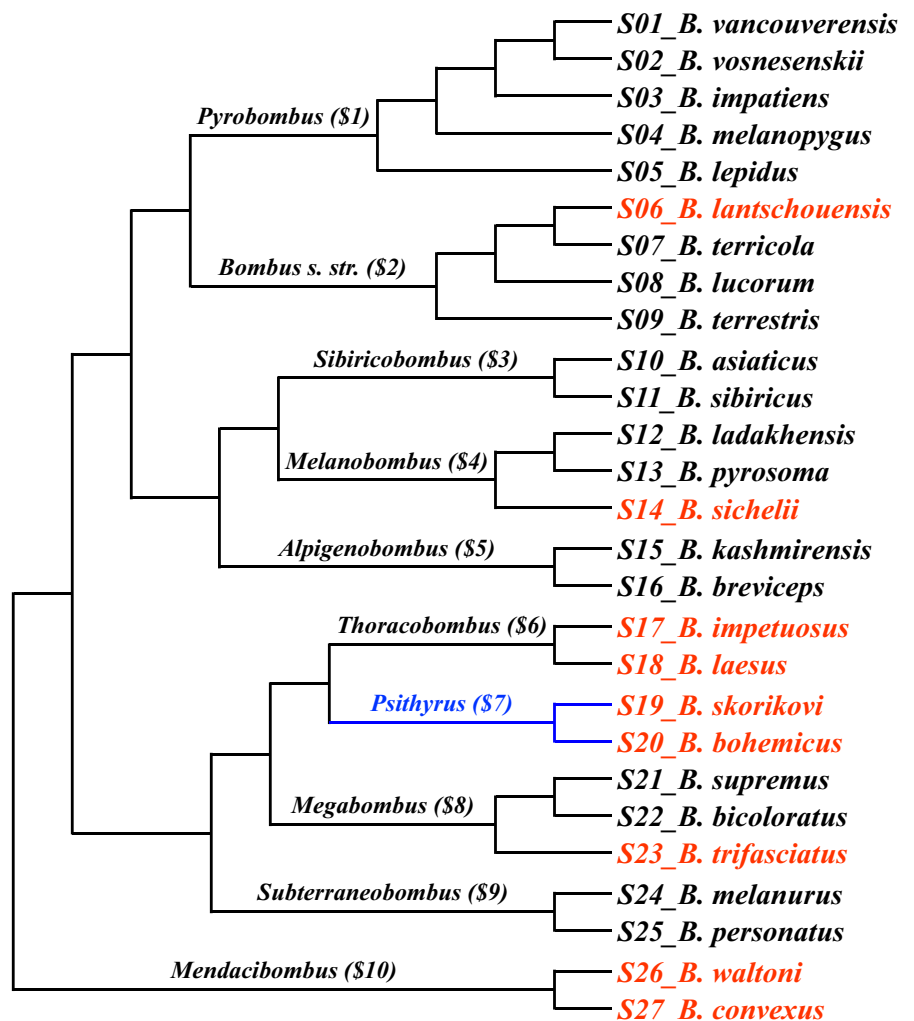


FIGURE 1 Phylogenetic relationships of the 27 bumble bee species involved in this study (red color, species whose genomes were sequenced in this study; blue color, obligate parasitic subgenus)

retrieved the corresponding coding sequences from the 26 other bumble bee genomes using the software Exonerate v2.4.0 (Slater & Birney, 2005), with default settings and with *B. impatiens* Vg and Vg-like protein sequences as queries.

Additionally, we aimed to retrieve Vg and Vg-like sequences from stingless bees (*Tetragonula* spp.) and honey bees (*Apis* spp.). These two genera are phylogenetically close relatives of *Bombus* (Peters et al., 2017) and five species of each genus (*Tetragonula carbonaria*, *T. clypearis*, *T. davenporti*, *T. hockingsi*, and *T. mellipes*; *Apis cerana*, *A. dorsata*, *A. florea*, *A. laboriosa*, and *A. mellifera*) were available with genomic resources in GenBank (Table S2). To do so, we obtained Vg and Vg-like sequences from GenBank for *Apis mellifera* (Vg, NP_001011578; Vg-like-A, XP_001121939; Vg-like-B, XP_395423; Vg-like-C, XP_001122505) and used them as queries to extract the other *Apis* species sequences using Exonerate with default settings. In the same manner, the translated protein sequences of the four genes from *B. impatiens* and *A. mellifera* were used as queries to extract the corresponding sequences from the five *Tetragonula* species.

The Vg and Vg-like sequences from each species were aligned for each genus separately using ClustalW (Codons) program embedded in the software MEGA v10.1.7 (Kumar et al., 2018) with default settings and verified by visual inspection. The DNA sequence variations were calculated using DnaSP v 6.12.03 (Rozas et al., 2017).

The reference species tree of *Bombus* (Figure 1) was drawn by TreeGraph2 v2.15.0 (Stöver & Müller, 2010) according to previous studies (Cameron et al., 2007; Williams et al., 2008, 2012). The reference species tree of *Tetragonula* (Figure S1) was drawn according to Rasmussen and Cameron (2010) and Hereward et al. (2020). The reference tree of *Apis* was drawn according to Raffiudin and Crozier (2007; Figure S1).

2.3 | Molecular evolutionary analyses among different genera

We used the CODEML program in the PAML package v4.9j (Yang, 2007) to study the selection pressures affecting the different genes and to test for patterns of molecular evolution. Our aim was to estimate synonymous and nonsynonymous substitutions using the branch tests and site tests (Bielawski & Yang, 2005).

We firstly tested whether the overall selection of Vg and Vg-like genes of bumble bees deviated from honey bees and stingless bees. The M0 (one-ratio) was used to estimate the overall selection (dN/dS , ratio of nonsynonymous / synonymous substitution rates) across all sites, and the alignment of each gene for each genus (*Bombus*, *Apis*, and *Tetragonula*) was separately used to calculate the dN/dS ratio of each genus. Next, we calculated the pairwise dN/dS ratios (Yang & Nielsen, 2000) across the species within each genus using the YN00 program from the PAML package. The values of pairwise dN/dS ratios of each gene were compared between *Bombus* and each of the other two genera using a Mann-Whitney rank test in SPSS v25.0. The relationships between dN and dS and between dN/dS and dS were tested using a linear regression analysis in SPSS.

2.4 | Molecular evolutionary analyses within *Bombus*

We first tested whether differences of selection pressures exist among the ten bumble bee subgenera. The “several dN/dS ratios” branch model (BM) was used to estimate the dN/dS ratios of each of the ten subgenera independently. The ten external branches corresponding to the ten subgenera were viewed as different foregrounds, whereas all the internal branches were viewed as a common background (Figure 1). The likelihood ratio tests (LRTs) between M0 (null model) and BM models were conducted by comparing twice the difference in log-likelihood values ($2\Delta lnL$) against a chi-square distribution ($df = 2$). The obtained ten dN/dS ratios of the bumble bee subgenera were compared across the four different genes using one-way ANOVA and paired samples t tests in SPSS. Moreover, we also used the branch-site model called Model A to test for positive selections in each subgenera. The null model for Model A is Model A1, which is a modify on Model A, but with $\omega_2 = 1$ fixed (Yang et al., 2005; Zhang et al., 2005). In each run, one target subgenus was marked as the foreground branch while the remaining nine subgenera were viewed as backgrounds. Again, $2\Delta lnL$ values between Model A and Model A1 were used to conduct LRTs for robustness with χ^2 test ($df = 1$).

Lastly, we studied the extent of selection for each Vg and Vg-like gene set of *Bombus* individually by dividing the data into four data sets (one for each orthologous gene) and comparing the neutral model (M1a) with a model allowing positive selection (M2a). The $2\Delta lnL$ values between the M1a and M2a models were used to test for robustness using LRTs with χ^2 test ($df = 2$), and positively selected sites were identified with the Bayes Empirical Bayes (BEB; Yang et al., 2005).

It should be mentioned that, in order to reduce false discovery rate, a Benjamini–Hochberg correction in R program (“p.adjust” command) was used where necessary (see below).

3 | RESULTS

3.1 | Overall genetic variation of Vg and Vg-like genes

A total of 2,095 million DNA reads were obtained from the nine bumble bee species (*B. lantschouensis*, *B. sichelii*, *B. impetuosus*, *B. laesus*, *B. skorikovi*, *B. bohemicus*, *B. trifasciatus*, *B. waltoni*, and *B. convexus*), with a total size of 314 Gb (SRA accession No. PRJNA667279). Coding sequences of the four Vg genes were successfully obtained for the 27 bumble bee species (File S1). The lengths of the aligned sequences (stop codons not considered) were 5,337, 4,569, 4,260, and 960 base pairs, respectively. Vg was the most variable in terms of nucleotide sequence identity among the four genes, with 2,389 (42.98%) variable nucleotide sites, including 39 indel codons. In contrast, Vg-like-B was the most conservative, with 11.56% variable nucleotide sites, and no indel was detected. Vg-like-A and Vg-like-C

TABLE 1 Genetic variations of *Vg* and *Vg-like* genes of the bumble bees (*Bombus*), stingless bees (*Tetragonula*), and honey bees (*Apis*; stop codons were not considered)

Sequence	Index	<i>Vg</i>	<i>Vg-like-A</i>	<i>Vg-like-B</i>	<i>Vg-like-C</i>
<i>Bombus</i> (N = 27)	Total sites	5,337	4,569	4,260	960
	Variable sites	2,294	903	476	193
	Variable percent	42.98%	19.76%	11.17%	20.10%
<i>Tetragonula</i> (N = 5)	Total sites	5,331	4,503	4,245	948
	Variable sites	140	73	54	39
	Variable percent	2.63%	1.62%	1.27%	4.11%
<i>Apis</i> (N = 5)	Total sites	5,319	4,524	4,260	978
	Variable sites	839	514	194	213
	Variable percent	15.77%	11.36%	4.55%	21.78%

TABLE 2 The overall *dN/dS* ratio of *Vg* and *Vg-like* genes based on MO model

Genus	<i>Vg</i>	<i>Vg-like-A</i>	<i>Vg-like-B</i>	<i>Vg-like-C</i>
<i>Bombus</i> (N = 27)	1.311	0.349	0.077	0.196
<i>Tetragonula</i> (N = 5)	0.302	0.295	0.072	0.107
<i>Apis</i> (N = 5)	0.260	0.211	0.076	0.301

showed similar levels of genetic variation and *Vg-like-C* sequence included three indel codons (Table 1).

The sequences of the four genes from five stingless bee species and five honey bee species are also provided in File S1. Unlike bumble bees, stingless bees and honey bees *Vg-like-C* sequence was the most variable in terms of nucleotide sequence identity among the four genes, follows by *Vg* and *Vg-like-A* (Table 1). Consistent with the bumble bee sequences, *Vg-like-B* was the most conserved gene for both stingless and honey bees (Table 1). The overall sequence variation information can be found in Table 1.

3.2 | Molecular evolution between *Bombus* and the other two genera

We first characterized the extent of positive and purifying selection in each of the orthologous copies separately (*Vg*, *Vg-like-A*, *Vg-like-B*, and *Vg-like-C*). The *dN/dS* ratios based on the MO model showed that *Bombus* conventional *Vg* was under strong positive selection (*dN/dS* = 1.311), whereas the *Vg-like* genes were under purifying selection (*dN/dS* = 0.349, 0.077 and 0.196, respectively, Table 2). In *Apis* and *Tetragonula*, all four genes were found under purifying selection. Moreover, in *Tetragonula*, *Vg-like-A*, *B*, and *C* are the most conserved copies (*dN/dS* = 0.295, 0.072 and 0.107, respectively, Table 2), whereas the conventional *Vg* evolve more rapidly (*dN/dS* = 0.302). In *Apis*, however, *Vg-like-C* was evolving more rapidly (*dN/dS* = 0.301) than the conventional *Vg* (*dN/dS* = 0.260), *Vg-like-A* (*dN/dS* = 0.211), and *Vg-like-B* (*dN/dS* = 0.076). In all three genera, *Vg-like-B* was the most conserved gene (Table 2).

Pairwise analyses of *Vg* selection showed that, for any of the three genera, *dN* significantly increased with increasing *dS* (Figure 2). Linear regression and Benjamini–Hochberg correction ($n = 12$) showed that the relationship of *dN* and *dS* of *Bombus* species follow a function of $dN = 0.022 + 0.924 * dS$, with a strong correlation ($R^2 = 0.882$, $p < .001$). Similar patterns were found among *Tetragonula* species ($dN = -6.26E-4 + 0.385 * dS$, $R^2 = 0.985$, $p < .001$) and *Apis* species ($dN = 0.016 + 0.212 * dS$, $R^2 = 0.788$, $p < .001$). For all *Vg* orthologous copies, *dN* and *dS* were also positively correlated, at the exception of *Vg-like-C* sequence in *Apis* ($R^2 = 0.281$, $p = .115$, Table 3). The linear regression information is shown in Table 3 and Figure S2.

Mann–Whitney rank test and Benjamini–Hochberg correction ($n = 8$) showed that pairwise *dN/dS* values of bumble bees *Vg* were significantly larger than those of stingless bees and honey bees ($Z = -5.393$, $p < .001$). Pairwise *dN/dS* values of bumble bees *Vg-like-A* were significantly larger than those of honey bees ($Z = -5.128$, $p < .001$), but not significantly deviated from stingless bees ($Z = -1.023$, $p = .350$). No significant differences of pairwise *dN/dS* values for *Vg-like-B* could be detected between bumblebees and stingless bees ($Z = -1.540$, $p = .165$) nor honey bees ($Z = -0.255$, $p = .799$). In contrast, pairwise *dN/dS* values of bumble bees *Vg-like-C* were significantly smaller than those of stingless bees ($Z = -3.306$, $p = .002$) and honey bees ($Z = -4.491$, $p < .001$).

3.3 | Molecular evolution within *Bombus*

Similar as the MO model, the branch model showed that *Bombus* conventional *Vg* was under positive selection with $dN/dS = 1.331 \pm 0.302$ (mean \pm SD), whereas the *Vg-like* genes were under purifying selection ($dN/dS = 0.344 \pm 0.064$, 0.074 ± 0.024 and 0.159 ± 0.080 , respectively [Table 4; Figure 3]). One-Way ANOVA based on the branch model results showed that the *dN/dS* values significantly deviated among the four genes ($F = 131.588$, $df = 3$, $p < .001$). Paired sample t test and Benjamini–Hochberg correction ($n = 6$) confirmed this finding ($df = 9$, $p < .006$) and showed a rank order of $Vg > Vg-like-A > Vg-like-C > Vg-like-B$ (Figure 3). According to the likelihood ratio test, the branch model significantly deviated from the MO ($df = 10$, $p < .001$) for *Vg*, but not

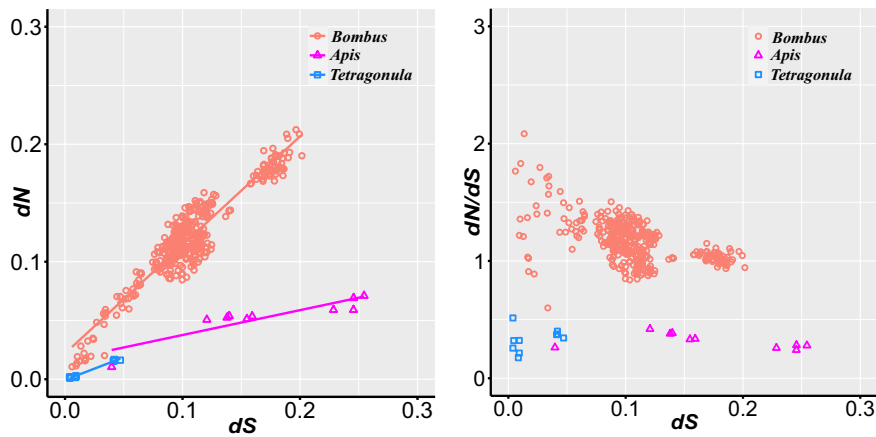


FIGURE 2 Scatter plot of dN versus dS (left) and dN/dS versus dS (right). The dN , dS , and dN/dS values were calculated under pairwise models among different species within genus *Bombus*, *Tetragonula*, and *Apis*, respectively

Gene	Genus	Equation	Statistical index
Vg	<i>Bombus</i>	$dN = 0.924*dS + 0.022$	$R^2 = .882, p < .001$
	<i>Tetragonula</i>	$dN = 0.385*dS - 6.26E-4$	$R^2 = .985, p < .001$
	<i>Apis</i>	$dN = 0.212*dS + 0.016$	$R^2 = .788, p < .001$
Vg-like-A	<i>Bombus</i>	$dN = 0.434*dS - 8.51E-4$	$R^2 = .914, p < .001$
	<i>Tetragonula</i>	$dN = 0.208*dS + 1.41E-3$	$R^2 = .979, p < .001$
	<i>Apis</i>	$dN = 0.212*dS + 1.74E-3$	$R^2 = .666, p = .004$
Vg-like-B	<i>Bombus</i>	$dN = 0.090*dS - 4.90E-4$	$R^2 = .709, p < .001$
	<i>Tetragonula</i>	$dN = 0.063*dS + 4.72E-4$	$R^2 = .970, p < .001$
	<i>Apis</i>	$dN = 0.078*dS + 1.13E-4$	$R^2 = .760, p = .001$
Vg-like-C	<i>Bombus</i>	$dN = 0.271*dS - 6.48E-3$	$R^2 = .856, p < .001$
	<i>Tetragonula</i>	$dN = 0.251*dS + 8.03E-4$	$R^2 = .977, p < .001$
	<i>Apis</i>	$dN = 0.129*dS + 0.057$	$R^2 = .281, p = .115$

TABLE 3 Linear regression between pairwise dN and dS of among bumble bee species

TABLE 4 dN/dS ratios for 10 bumblebee subgenera based on branch model and LRTs between branch model and M0 (null model)

Subgenus	Vg	Vg-like-A	Vg-like-B	Vg-like-C
<i>Pyrobombus</i>	1.514	0.379	0.091	0.221
<i>Bombus s. str.</i>	1.118	0.413	0.086	0.112
<i>Sibiricobombus</i>	1.712	0.411	0.065	0.185
<i>Melanobombus</i>	1.628	0.312	0.062	0.046
<i>Alpigenobombus</i>	1.344	0.286	0.038	0.142
<i>Thoracobombus</i>	1.589	0.362	0.112	0.138
<i>Psithyrus</i>	0.713	0.300	0.068	0.262
<i>Megabombus</i>	1.368	0.315	0.054	0.187
<i>Subterraneobombus</i>	1.182	0.234	0.058	0.034
<i>Mendacibombus</i>	1.137	0.425	0.106	0.259
$2\Delta\ln L$	33.518	9.330	11.658	14.514
p ($df = 10$)	<.001	.501	.309	.151

for the Vg-like genes ($df = 10, p \geq .151$; Table 4). For *Bombus* Vg, nine out of ten subgenera showed a dN/dS ratios over 1 (average = 1.40, range: 1.12–1.71, Table 4), with the exception of the subgenus *Psithyrus* (Figure 3) for which the dN/dS ratios was only 0.713.

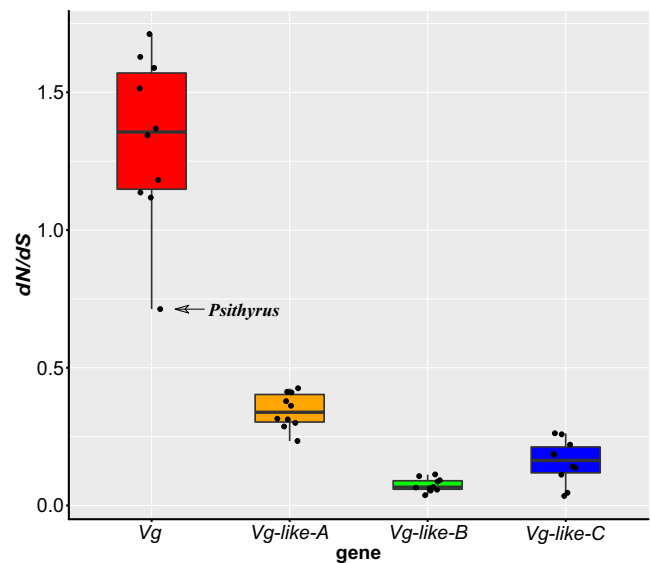


FIGURE 3 The dN/dS ratios of Vg and Vg-like genes of bumble bees based on branch model

The M2a model showed that 32.1% of the *Bombus* Vg codons were under purifying selection ($dN/dS < 1$), 46.2% under neutral selection ($dN/dS = 1$), and 21.6% under positive selection ($\omega = 4.111$).

TABLE 5 Statistical results of site models in *Vg* and *Vg-like* genes of 27 bumble bee species

Gene	M1a (np = 54)	M2a (np = 56)	LRT (df = 2)
<i>Vg</i>	$p0 = 0.436, \omega0 = 0.083; p1 = 0.564, \omega1 = 1.000; \lnL = -31,891.635$	$p0 = 0.321, \omega0 = 0.070; p1 = 0.462, \omega1 = 1.000; p2 = 0.216, \omega2 = 4.111; \lnL = -31,416.198$	$2\Delta\lnL = 950.874, p < .001$
<i>Vg-like-A</i>	$p0 = 0.716, \omega0 = 0.064; p1 = 0.284, \omega1 = 1.000; \lnL = -13,541.057$	$p0 = 0.746, \omega0 = 0.091; p1 = 0.231, \omega1 = 1.000; p2 = 0.023, \omega2 = 3.917; \lnL = -13,526.160$	$2\Delta\lnL = 29.794, p < .001$
<i>Vg-like-B</i>	$p0 = 0.968, \omega0 = 0.050; p1 = 0.032, \omega1 = 1.000; \lnL = -9,062.035$	$p0 = 0.971, \omega0 = 0.052; p1 = 0.028, \omega1 = 1.000; p2 = 0.001, \omega2 = 6.668; \lnL = -9,060.276$	$2\Delta\lnL = 3.581, p = .167$
<i>Vg-like-C</i>	$p0 = 0.877, \omega0 = 0.075; p1 = 0.123, \omega1 = 1.000; \lnL = -2,892.983$	$p0 = 0.886, \omega0 = 0.078; p1 = 0.000, \omega1 = 1.000; p2 = 0.114, \omega2 = 1.071; \lnL = -2,892.958$	$2\Delta\lnL = 0.050, p = .975$

TABLE 6 Branch-site model results for each of the 10 bumblebee subgenera ($N_{0.5}$, number of sites which were positively selected with posterior probability >0.5 ; $N_{0.95}$, number of sites which were positively selected with posterior probability >0.95)

Foreground subgenus	LRT (df = 1)	$N_{0.5}$	$N_{0.95}$
<i>Pyrobombus</i>	$2\Delta\lnL = 94.471, p < .001$	52	10
<i>Bombus s. str.</i>	$2\Delta\lnL = 6.709, p = .012$	3	0
<i>Sibiricobombus</i>	$2\Delta\lnL = 12.667, p < .001$	26	1
<i>Melanobombus</i>	$2\Delta\lnL = 57.192, p < .001$	80	5
<i>Alpigenobombus</i>	$2\Delta\lnL = 19.542, p < .001$	41	1
<i>Thoracobombus</i>	$2\Delta\lnL = 27.831, p < .001$	107	0
<i>Psithyrus</i>	$2\Delta\lnL = 0.000, p = 1.000$	1	0
<i>Megabombus</i>	$2\Delta\lnL = 10.140, p = .002$	15	0
<i>Subterraneobombus</i>	$2\Delta\lnL = 6.037, p = .016$	6	0
<i>Mendacibombus</i>	$2\Delta\lnL = 41.450, p < .001$	44	0

For the *Vg-like* genes, most codons were under either purifying selection or neutral selection (Table 5). In *Bombus*, Bayes Empirical Bayes analysis showed that 150 codons (8.43%) in the conventional *Vg* and 3 codons (0.20%) in *Vg-like-A* were detected as positively selected sites (probability $>.95$). Yet, we could not detect any positively selected sites for *Vg-like-B* and *Vg-like-C*. The M2a models were robust for *Vg* and *Vg-like-A*, but not for *Vg-like-B* or *Vg-like-C* (Table S3).

The branch-site models as well as Benjamini–Hochberg correction ($n = 10$) showed that for nine of the ten *Bombus* subgenera (except for *B. Psithyrus*), three or more sites were under positive selection with posterior probability >0.5 . And in four subgenera (*Pyrobombus*, *Sibiricobombus*, *Melanobombus*, and *Alpigenobombus*), at least one site was under positive selection with posterior probability $>.95$. For the same nine subgenera, the log-likelihood values of Model A significantly deviated from its

counterpart Model A1 ($p < .016$), indicating the likelihood of this result. For *Psithyrus*, only one site showed signs of positive selection. However, because there were not any deviations between Model A and Model A1 ($2\Delta\lnL = 0.000, p = 1.000$), this result could not be supported. The details of the branch-site model results are listed in Table 6.

4 | DISCUSSION

Vitellogenin is a multifunctional hemolymph protein that is characterized by its central role in social insect's division of labor and affects multiple aspects of social insect life histories (Amdam et al., 2003). Bumble bees are primitively eusocial insects which represents an intermediate stage in the evolution between solitary and eusociality (Goulson, 2010). Unlike the “true” eusocial bees such as honey bees or stingless bees, most bumble bee species pass through both solitary and eusocial phases during their life cycle. Moreover, bumble bees are extremely diversified in their biological and ecological characteristics among different subgenera. Here we show that in the bumble bees, *Vg* has experienced elevated rate of evolution and is under strong positive selection. Its homologous genes, *Vg-like* genes that have diverged both from *Vg* and from each other, do not display such patterns. Additionally, signs of positive selection were also absent from its sister clades, *Tetragonula* and *Apis*'s *Vg*.

A unique advantage of bumble bees is the fact that sequences from the two closest clades are publicly available, so we can investigate the forces driving vitellogenin evolution across the phylogeny tree and ask whether *Vg* and *Vg-like* genes in distinctive genus differ in their molecular evolutionary rates. Our study determines that the conventional *Vg* is the most rapidly evolving copy in bumblebees. Interestingly, we observed that unlike in bumble bees, no sign of positive selection could be found for the conventional *Vg* of both *Tetragonula* and *Apis*. The pairwise analyses showed that all the three genera had increased dN with increasing dS . However, the slope of the linear regression of *Bombus* was much higher than that of *Tetragonula* and *Apis*, indicating a higher overall dN/dS ratio in *Bombus*. Although there were considerable overlaps of dS between *Bombus* and the other two genera, the pairwise dN/dS values among *Bombus* genera were larger than any of the counterparts of *Tetragonula* and *Apis* (Figure 2). In *Tetragonula* and *Apis*, the purifying

selection pressure on *Vg* might be slightly lower because *Vg*-related traits underlying adaptive evolution may differ between the genera. Thus, our results suggest that duplication, positive evolution, and purifying selection may be major evolutionary forces driving *Vg* gene evolution across divergent taxa.

Vg is best known for its primary role in the formation of egg yolk in egg-laying animals (Tufail & Takeda, 2008); however, in social insects, *Vg* has probably acquired additional functions (Guidugli et al., 2005) and fulfills roles related to behavior and survival (Havukainen et al., 2011; Kent et al., 2011; Nelson et al., 2007; Seehuus et al., 2006). Functional pleiotropy plays an important role in molecular evolution (Paaby & Rockman, 2013). Although increased purifying selection against pleiotropic mutations (McGuigan et al., 2014) and slow rates of evolution of pleiotropic genes (Salathe et al., 2006) have been observed in some cases, more recent studies have demonstrated that gene pleiotropy may increase evolutionary rate (Razeto-Barry et al., 2011; Twyman et al., 2018; Vedanayagam & Garrigan, 2015). While viewed as a primitively eusocial genus, the single conventional *Vg* protein in bumble bees showed inclines of pleiotropy in previous studies. For instance, *Vg* was found to be expressed in *B. hypocrita* in several castes including queen, worker, and even drone (Li et al., 2010). Additionally, *Vg* mRNA was detected in various tissues including flight muscles in *B. terrestris* as well as *B. lantschouensis* (Jedlicka et al., 2016; Zhen et al., 2018). This pattern is consistent with the fact that the positive selection that we detected may drive the evolution of novel protein function and thus enables functional pleiotropy of bumble bee's *Vg* (Razeto-Barry et al., 2011). Similar patterns have been observed in the *Vg* genes of other eusocial hymenopteran species (Kent et al., 2011; Morandin et al., 2014; Salmela et al., 2016). Thus, our results suggest that *Vg* functional pleiotropy may have arisen due to strong positive selection acting on it, and may further indicate the appearance of a novel, unknown *Vg* function in bumble bees.

Functional pleiotropy of a gene is also predicted to lead to a duplication event, such as the duplication of the ancestral gene leading to *Vg* and *Vg-like* genes (Morandin et al., 2014). Gene duplication is an important source of new genetic material for selection to act upon (Force et al., 1999; Lynch & Force, 2000; Ohno, 1970; Zhang, 2003). After duplication, the duplicated gene copy can acquire functions different from those of the ancestral gene (Gu et al., 2002; Khaladkar & Hannehalli, 2012; Morandin et al., 2014; Wagner, 2000). Unlike *Vg*, the three *Vg-like* genes showed signs of purifying selection in all three genera. Multiple factors could be affecting the rate of sequence evolution such as the number of pleiotropic interactions, the gene expression levels, or their tissue-specific expression patterns, that could not be detected in this study. Further studies on evolutionary patterns of *Vg-like* genes across social insect species and on their functions are needed to fully understand their roles in social insects and the selection pressures they experience.

An exception to the otherwise overall positive selection on vitellogenin in our set of bumble bee species is the subgenus *Psithyrus*. We found that the subgenus-based branch model analyses showed

that almost all subgenera had a *dN/dS* ratio over 1. Incredibly however, the *dN/dS* ratio of the subgenus *Psithyrus* was much lower than the other subgenera (only 0.713). Moreover, the subgenus-based branch-site model analyses indicated that all subgenera but *Psithyrus* had significant deviations between Model A (alternative hypothesis) and Model A1 (null hypothesis). These results repeatedly indicate the distinctive evolution of *Psithyrus Vg*. Bumble bees in the subgenus *Psithyrus* have annual life cycles similar to those of typical bumble bee species, except that instead of founding their own nest and rearing workers, they steal a nest from "true" bumble bees (Goulson, 2010). Initially, these bumble bees were formerly described as a separate genus *Psithyrus*, but it is now widely accepted that they belong to a subgenus within *Bombus*, with the subgenus *Thoracobombus* as the sister group (Cameron et al., 2007; Williams et al., 2008). These bumblebees also exhibit social parasitism, with the absence of a worker caste, or the need to forage for nectar and pollen to provision developing larvae (Lhomme & Hines, 2019). Although it is not clear whether the lower positive selection level of *Psithyrus* was a secondary evolutionary event, it is probable that their fundamentally different life history has influenced the evolution of *Vg* even within just one genus. In fact, due to the absence of caste differentiation in *Psithyrus*, it is conceivable that its simplified lifestyle reduced the functional pleiotropy demand on *Vg* and thus caused a lower *dN/dS* ratio.

ACKNOWLEDGMENTS

The research was supported by the Science and Technology Foundation of Jiangxi Provincial Department of Education (GJJ201007) and the National Natural Science Foundation of China (No. 31860105).

CONFLICT OF INTEREST

The authors declare no conflict of interest.


AUTHOR CONTRIBUTION

Fang Zhao: Investigation (equal); Methodology (equal). **Claire Morandin:** Investigation (equal); Methodology (equal). **Kai Jiang:** Methodology (supporting). **Tianjuan Su:** Methodology (supporting). **Bo He:** Methodology (supporting). **Gonghua Lin:** Conceptualization (lead); Funding acquisition (equal); Investigation (lead); Methodology (equal). **Zuhao Huang:** Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (lead).

DATA AVAILABILITY STATEMENT

The genome data (BioProject ID: PRJNA667279) are available from the Sequence Read Archive of the National Centre for Biotechnology Information (NCBI). The other data generated or analyzed during this study are included in this published article and its supplementary files.

ORCID

Claire Morandin  <https://orcid.org/0000-0002-0811-901X>
Gonghua Lin  <https://orcid.org/0000-0001-7670-7932>

REFERENCES

- Amdam, G. V., Norberg, K., Hagen, A., & Omholt, S. W. (2003). Social exploitation of vitellogenin. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 1799–1802. <https://doi.org/10.1073/pnas.0333979100>
- Amsalem, E., Grozinger, C. M., Padilla, M., & Hefetz, A. (2015). The physiological and genomic bases of bumble bee social behaviour. *Advances in Insect Physiology*, *48*, 37–93.
- An, J., Huang, J., Shao, Y., Zhang, S., Wang, B., Liu, X., Wu, J., & Williams, P. H. (2014). The bumble bees of North China (Apidae, *Bombus* Latreille). *Zootaxa*, *3830*(1), 1–89.
- Bielawski, J. P., & Yang, Z. (2005). Maximum likelihood methods for detecting adaptive protein evolution. In R. Nielsen (Ed.), *Statistical methods in molecular evolution* (pp. 103–124). Springer.
- Bloch, G., & Grozinger, C. M. (2011). Social molecular pathways and the evolution of bee societies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*, 2155–2170. <https://doi.org/10.1098/rstb.2010.0346>
- Cameron, S. A., Hines, H. M., & Williams, P. H. (2007). A comprehensive phylogeny of the bumble bees (*Bombus*). *Biological Journal of the Linnean Society*, *91*, 161–188. <https://doi.org/10.1111/j.1095-8312.2007.00784.x>
- Corona, M., Velarde, R. A., Remolina, S., Moran-Lauter, A., Wang, Y., Hughes, K. A., & Robinson, G. E. (2007). Vitellogenin, juvenile hormone, insulin signalling, and queen honey bee longevity. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(17), 7128–7133.
- Du, L., Wang, M., Li, J., He, S., Huang, J., & Wu, J. (2019). Characterization of a vitellogenin receptor in the bumblebee, *Bombus lantschouensis* (Hymenoptera, Apidae). *Insects*, *10*, 445. <https://doi.org/10.3390/insects10120445>
- Excels, W. (1974). Occurrence and significance of vitellogenins in female castes of social Hymenoptera. *American Zoologist*, *14*, 1229–1237. <https://doi.org/10.1093/icb/14.4.1229>
- Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y., & Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, *151*(4), 1531–1545. <https://doi.org/10.1093/genetics/151.4.1531>
- Goulson, D. (2010). *Bumblebees: Behaviour, ecology, and conservation*. Oxford University Press.
- Gu, Z., Nicolae, D., Lu, H. H. S., & Li, W. H. (2002). Rapid divergence in expression between duplicate genes inferred from microarray data. *Trends in Genetics*, *18*, 609–613. [https://doi.org/10.1016/S0168-9525\(02\)02837-8](https://doi.org/10.1016/S0168-9525(02)02837-8)
- Guidugli, K. R., Nascimento, A. M., Amdam, G. V., Barchuk, A. R., Omholt, S., Simões, Z. L. P., & Hartfelder, K. (2005). Vitellogenin regulates hormonal dynamics in the worker caste of a eusocial insect. *FEBS Letters*, *579*, 4961–4965. <https://doi.org/10.1016/j.febslet.2005.07.085>
- Hagedorn, H. H., & Kunkel, J. G. (1979). Vitellogenin and vitellin in insects. *Annual Review of Entomology*, *24*, 475–505. <https://doi.org/10.1146/annurev.en.24.010179.002355>
- Havukainen, H., Halskau, Ø., & Amdam, G. V. (2011). Social pleiotropy and the molecular evolution of honey bee vitellogenin. *Molecular Ecology*, *20*, 5111–5113. <https://doi.org/10.1111/j.1365-294X.2011.05351.x>
- Hereward, J. P., Smith, T. J., Brookes, D. R., Gloag, R., & Walter, G. H. (2020). Tests of hybridisation in *Tetragonula* stingless bees using multiple genetic markers. *bioRxiv*. <https://doi.org/10.1101/2020.03.08.982546>
- Jackson, J. M., Pimsler, M. L., Oyen, K. J., Strange, J. P., Dillon, M. E., & Lozier, J. D. (2020). Local adaptation across a complex bioclimatic landscape in two montane bumble bee species. *Molecular Ecology*, *29*, 920–939. <https://doi.org/10.1111/mec.15376>
- Jedlicka, P., Ernst, U. R., Votavova, A., Hanus, R., & Valterova, I. (2016). Gene expression dynamics in major endocrine regulatory pathways along the transition from solitary to social life in a bumblebee. *Bombus Terrestris*. *Frontiers in Physiology*, *7*, 574. <https://doi.org/10.3389/fphys.2016.00574>
- Kent, C. F., Dey, A., Patel, H., Tsvetkov, N., Tiwari, T., MacPhail, V. J., Gobeil, Y., Harpur, B. A., Gurtowski, J., Schatz, M. C., Colla, S. R., & Zayed, A. (2018). Conservation genomics of the declining North American Bumblebee *Bombus terricola* reveals inbreeding and selection on immune genes. *Frontiers in Genetics*, *9*, 316. <https://doi.org/10.3389/fgene.2018.00316>
- Kent, C. F., Issa, A., Bunting, A. C., & Zayed, A. (2011). Adaptive evolution of a key gene affecting queen and worker traits in the honey bee, *Apis mellifera*. *Molecular Ecology*, *20*, 5226–5235. <https://doi.org/10.1111/j.1365-294X.2011.05299.x>
- Khaladkar, M., & Hannenhalli, S. (2012). Functional divergence of gene duplicates—A domain-centric view. *BMC Evolutionary Biology*, *12*, 126. <https://doi.org/10.1186/1471-2148-12-126>
- Kohlmeier, P., Feldmeyer, B., & Foitzik, S. (2018). Vitellogenin-like A—Associated shifts in social cue responsiveness regulate behavioral task specialization in an ant. *PLoS Biology*, *16*(6), e2005747. <https://doi.org/10.1371/journal.pbio.2005747>
- Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*, 1547–1549. <https://doi.org/10.1093/molbev/msy096>
- Lhomme, P., & Hines, H. M. (2019). Ecology and evolution of cuckoo bumble bees. *Annals of the Entomological Society of America*, *112*(3), 122–140. <https://doi.org/10.1093/aesa/say031>
- Li, J., Huang, J., Cai, W., Zhao, Z., Peng, W., & Wu, J. (2010). The vitellogenin of the bumblebee, *Bombus hypocrita*: Studies on structural analysis of the cDNA and expression of the mRNA. *Journal of Comparative Physiology B*, *180*, 161–170. <https://doi.org/10.1007/s00360-009-0434-5>
- Lin, G., Huang, Z., Wang, L., Chen, Z., Zhang, T., Gillman, L. N., & Zhao, F. (2019). Evolutionary rates of bumblebee genomes are faster at lower elevations. *Molecular Biology and Evolution*, *36*(6), 1215–1219. <https://doi.org/10.1093/molbev/msz057>
- Lynch, M., & Force, A. (2000). The probability of duplicate gene preservation by subfunctionalization. *Genetics*, *154*, 459–473. <https://doi.org/10.1093/genetics/154.1.459>
- McGuigan, K., Collet, J. M., Allen, S. L., Chenoweth, S. F., & Blows, M. W. (2014). Pleiotropic mutations are subject to strong stabilizing selection. *Genetics*, *197*(3), 1051–1062. <https://doi.org/10.1534/genetics.114.165720>
- Morandin, C., Havukainen, H., Kulmuni, J., Dhaygude, K., Trontti, K., & Helanterä, H. (2014). Not only for egg yolk—Functional and evolutionary insights from expression, selection, and structural analyses of *Formica* ant vitellogenins. *Molecular Biology and Evolution*, *31*(8), 2181–2193. <https://doi.org/10.1093/molbev/msu171>
- Münch, D., & Amdam, G. V. (2010). The curious case of aging plasticity in honey bees. *FEBS Letters*, *584*, 2496–2503. <https://doi.org/10.1016/j.febslet.2010.04.007>
- Nelson, C. M., Ihle, K., Amdam, G. V., Fondrk, M. K., & Page, R. E. (2007). The gene vitellogenin has multiple coordinating effects on social organization. *PLoS Biology*, *5*, 673–677. <https://doi.org/10.1371/journal.pbio.0050062>
- Ohno, S. (1970). *Evolution by gene duplication*. George Allen & Unwin Ltd.
- Otto, S. P. (2004). Two steps forward, one step back: The pleiotropic effects of favoured alleles. *Proceedings of the Royal Society B: Biological Sciences*, *271*, 705–714. <https://doi.org/10.1098/rspb.2003.2635>
- Paaby, A. B., & Rockman, M. V. (2013). The many faces of pleiotropy. *Trends in Genetics*, *29*(2), 66–73. <https://doi.org/10.1016/j.tig.2012.10.010>
- Pan, M. L., Bell, W. J., & Tefler, W. H. (1969). Vitellogenic blood protein synthesis by insect fat body. *Science*, *165*, 393–394. <https://doi.org/10.1126/science.165.3891.393>
- Peng, Y., Leung, H. C. M., Yiu, S. M., & Chin, F. Y. L. (2012). IDBA-UD: A de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*, *28*, 1420–1428. <https://doi.org/10.1093/bioinformatics/bts174>
- Peters, R. S., Krogmann, L., Mayer, C., Donath, A., Gunkel, S., Meusemann, K., Kozlov, A., Podsiadlowski, L., Petersen, M., Lanfear, R., Diez, P. A.,

- Heraty, J., Kjer, K. M., Klopstein, S., Meier, R., Polidori, C., Schmitt, T., Liu, S., Zhou, X., ... Niehuis, O. (2017). Evolutionary history of the Hymenoptera. *Current Biology*, 27(7), 1013–1018. <https://doi.org/10.1016/j.cub.2017.01.027>
- Piulachs, M. D., Guidugli, K. R., Barchuk, A. R., Cruz, J., Simões, Z. L. P., & Bellés, X. (2003). The vitellogenin of the honey bee, *Apis mellifera*: Structural analysis of the cDNA and expression studies. *Insect Biochemistry and Molecular Biology*, 33(4), 459–465. [https://doi.org/10.1016/S0965-1748\(03\)00021-3](https://doi.org/10.1016/S0965-1748(03)00021-3)
- Raffiudin, R., & Crozier, R. H. (2007). Phylogenetic analysis of honey bee behavioral evolution. *Molecular Phylogenetics and Evolution*, 43, 543–552. <https://doi.org/10.1016/j.ympev.2006.10.013>
- Raikhel, A. S., & Dhadialla, T. (1992). Accumulation of yolk proteins in insect oocytes. *Annual Review of Entomology*, 37, 217–251. <https://doi.org/10.1146/annurev.en.37.010192.001245>
- Rasmussen, C., & Cameron, S. A. (2010). Global stingless bee phylogeny supports ancient divergence, vicariance, and long distance dispersal. *Biological Journal of the Linnean Society*, 99, 206–232. <https://doi.org/10.1111/j.1095-8312.2009.01341.x>
- Razeto-Barry, P., Diaz, J., Cotoras, D., & Vasquez, R. A. (2011). Molecular evolution, mutation size and gene pleiotropy: A geometric reexamination. *Genetics*, 187, 877–885. <https://doi.org/10.1534/genetics.110.125195>
- Ricigliano, V. A., Mott, B. M., Floyd, A. S., Copeland, D. C., Carroll, M. J., & Anderson, K. E. (2018). Honey bees overwintering in a southern climate: Longitudinal effects of nutrition and queen age on colony-level molecular physiology and performance. *Scientific Reports*, 8, 1–11. <https://doi.org/10.1038/s41598-018-28732-z>
- Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large datasets. *Molecular Biology and Evolution*, 34, 3299–3302. <https://doi.org/10.1093/molbev/msx248>
- Sadd, B. M., Barribeau, S. M., Bloch, G., de Graaf, D. C., Dearden, P., Elsiik, C. G., Gadau, J., Grimmelikhuijzen, C. J. P., Hasselmann, M., Lozier, J. D., Robertson, H. M., Smaghe, G., Stolle, E., Van Vaerenbergh, M., Waterhouse, R. M., Bornberg-Bauer, E., Klasberg, S., Bennett, A. K., Câmara, F., ... Worley, K. C. (2015). The genomes of two key bumblebee species with primitive eusocial organization. *Genome Biology*, 16, 76. <https://doi.org/10.1186/s13059-015-0623-3>
- Salathe, M., Ackermann, M., & Bonhoeffer, S. (2006). The effect of multifunctionality on the rate of evolution in yeast. *Molecular Biology and Evolution*, 23, 721–722. <https://doi.org/10.1093/molbev/msj086>
- Salmela, H., Stark, T., Stucki, D., Fuchs, S., Freitak, D., Dey, A., Kent, C. F., Zayed, A., Dhaygude, K., Hokkanen, H., & Sundström, L. (2016). Ancient duplications have led to functional divergence of Vitellogenin-like genes potentially involved in inflammation and oxidative stress in honey bees. *Genome Biology and Evolution*, 8, 495–506. <https://doi.org/10.1093/gbe/evw014>
- Seehuus, S. C., Norberg, K., Gimsa, U., Krekling, T., & Amdam, G. V. (2006). Reproductive protein protects sterile honey bee workers from oxidative stress. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 962–967.
- Slater, G. S. C., & Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics*, 6, 31.
- Spieth, J., Nettleton, M., Zucker-Aprison, E., Lea, K., & Blumenthal, T. (1991). Vitellogenin motifs conserved in nematodes and vertebrates. *Journal of Molecular Evolution*, 32(5), 429–438. <https://doi.org/10.1007/BF02101283>
- Stöver, B. C., & Müller, K. F. (2010). TreeGraph 2: Combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics*, 11, 7. <https://doi.org/10.1186/1471-2105-11-7>
- Tian, L., Rahman, S. R., Ezray, B. D., Franzini, L., Strange, J. P., Lhomme, P., & Hines, H. M. (2019). A homeotic shift late in development drives mimetic color variation in a bumble bee. *Proceedings of the National Academy of Sciences of the United States of America*, 116(24), 11857–11865. <https://doi.org/10.1073/pnas.1900365116>
- Trenczek, T., & Engels, W. (1986). Occurrence of vitellogenin in drone honeybees (*Apis mellifica*). *International Journal of Reproduction and Development*, 10(3), 307–311.
- Tufail, M., & Takeda, M. (2008). Molecular characteristics of insect vitellogenins. *Journal of Insect Physiology*, 54, 1447–1458. <https://doi.org/10.1016/j.jinsphys.2008.08.007>
- Twyman, H., Andersson, S., & Mundy, N. I. (2018). Evolution of CYP2J19, a gene involved in colour vision and red coloration in birds: Positive selection in the face of conservation and pleiotropy. *BMC Evolutionary Biology*, 18(1), 22. <https://doi.org/10.1186/s12862-018-1136-y>
- Vedanayagam, J. P., & Garrigan, D. (2015). The effects of natural selection across molecular pathways in *Drosophila melanogaster*. *BMC Evolutionary Biology*, 15(1), 203. <https://doi.org/10.1186/s12862-015-0472-4>
- Wagner, A. (2000). Decoupled evolution of coding region and mRNA expression patterns after gene duplication: Implications for the neutralist-selectionist debate. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 6579–6584. <https://doi.org/10.1073/pnas.110147097>
- Weil, T., Korb, J., & Rehli, M. (2009). Comparison of queen-specific gene expression in related lower termite species. *Molecular Biology and Evolution*, 26, 1841–1850. <https://doi.org/10.1093/molbev/msp095>
- Williams, H., Brown, M. J. F., Carolan, J. C., An, J., Goulson, D., Aytakin, A. M., Best, L. R., Byvaltsev, A. M., Cederberg, B., Dawson, R., Huang, J., Ito, M., Monfared, A., Raina, R. H., Schmid-Hempel, P., Sheffield, C. S., Šima, P., & Xie, Z. (2012). Unveiling cryptic species of the bumblebee subgenus *Bombus* s. str. worldwide with COI barcodes (Hymenoptera: Apidae). *Systematics and Biodiversity*, 10(1), 21–56.
- Williams, P. H., Cameron, S. A., Hines, H. M., Cederberg, B., & Rasmont, P. (2008). A simplified subgeneric classification of the bumblebees (genus *Bombus*). *Apidologie*, 39(1), 46–74.
- Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, 24, 1586–1591. <https://doi.org/10.1093/molbev/msm088>
- Yang, Z., & Nielsen, R. (2000). Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Molecular Biology and Evolution*, 17, 32–43. <https://doi.org/10.1093/oxfordjournals.molbev.a026236>
- Yang, Z., Wong, W. S. W., & Nielsen, R. (2005). Bayes empirical Bayes inference of amino acid sites under positive selection. *Molecular Biology and Evolution*, 22, 1107–1118. <https://doi.org/10.1093/molbev/msi097>
- Zhang, J. (2003). Evolution by gene duplication: An update. *Trends in Ecology & Evolution*, 18, 292–298. [https://doi.org/10.1016/S0169-5347\(03\)00033-8](https://doi.org/10.1016/S0169-5347(03)00033-8)
- Zhang, J., Nielsen, R., & Yang, Z. (2005). Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Molecular Biology and Evolution*, 22, 2472–2479. <https://doi.org/10.1093/molbev/msi237>
- Zhen, C., Yang, H., Luo, S., Huang, J., & Wu, J. (2018). Broad-complex Z3 contributes to the ecdysone-mediated transcriptional regulation of the vitellogenin gene in *Bombus lantschouensis*. *PLoS One*, 13(11), e0207275. <https://doi.org/10.1371/journal.pone.0207275>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Zhao F, Morandin C, Jiang K, et al. Molecular evolution of bumble bee vitellogenin and vitellogenin-like genes. *Ecol Evol*. 2021;11:8983–8992. <https://doi.org/10.1002/ece3.7736>