Comparative Genomics Sheds Light on the Convergent Evolution of Miniaturized Wasps

Hongxing Xu,[†,1] Xinhai Ye [ID],[†,2,3,4] Yajun Yang,[1] Yi Yang,[2] Yu H. Sun,[5] Yang Mei,[2] Shijiao Xiong,[2] Kang He [ID],[2] Le Xu,[2] Qi Fang,[2] Fei Li,*,[2] Gongyin Ye,*,[2] and Zhongxian Lu*,[1]

[1]State Key Laboratory for Managing Biotic and Chemical Treats to the Quality and Safety of Agroproducts, Institute of Plant Protection and Microbiology, Zhejiang Academy of Agricultural Sciences, Hangzhou, China

[2]State Key Laboratory of Rice Biology & Ministry of Agricultural and Rural Affairs Key Laboratory of Molecular Biology of Crop Pathogens and Insects, Institute of Insect Sciences, Zhejiang University, Hangzhou, China

[3]Shanghai Institute for Advanced Study, Zhejiang University, Shanghai, China

[4]Institute of Artificial Intelligence, College of Computer Science and Technology, Zhejiang University, Hangzhou, China

[5]Department of Biology, University of Rochester, Rochester, NY, USA

[†]These authors contributed equally to this work.

*Corresponding authors: E-mails: luzxmh@163.com; chu@zju.edu.cn; lifei18@zju.edu.cn.

Associate editor: Guang Yang

Abstract

Miniaturization has occurred in many animal lineages, including insects and vertebrates, as a widespread trend during animal evolution. Among Hymenoptera, miniaturization has taken place in some parasitoid wasp lineages independently, and may have contributed to the diversity of species. However, the genomic basis of miniaturization is little understood. Diverged approximately 200 Ma, _Telenomus_ wasps (Platygastroidea) and _Trichogramma_ wasps (Chalcidoidea) have both evolved to a highly reduced body size independently, representing a paradigmatic example of convergent evolution. Here, we report a high-quality chromosomal genome of _Telenomus remus_, a promising candidate for controlling _Spodoptera frugiperda_, a notorious pest that has recently caused severe crop damage. The _T. remus_ genome (129 Mb) is characterized by a low density of repetitive sequence and a reduction of intron length, resulting in the shrinkage of genome size. We show that hundreds of genes evolved faster in two miniaturized parasitoids _Trichogramma pretiosum_ and _T. remus_. Among them, 38 genes exhibit extremely accelerated evolutionary rates in these miniaturized wasps, possessing diverse functions in eye and wing development as well as cell size control. These genes also highlight potential roles in body size regulation. In sum, our analyses uncover a set of genes with accelerated evolutionary rates in _Tri. pretiosum_ and _T. remus_, which might be responsible for their convergent adaptations to miniaturization, and thus expand our understanding on the evolutionary basis of miniaturization. Additionally, the genome of _T. remus_ represents the first genome resource of superfamily Platygastroidea, and will facilitate future studies of Hymenoptera evolution and pest control.

Key words: miniaturization, convergent evolution, chromosome-level genome, parasitoid wasp, _Telenomus remus_, Platygastroidea.

Introduction

Miniaturization, that is, body size reduction, is an important evolutionary process in animal evolution and one of the principal directions of evolution in insects (Hanken and Wake 1993; Polilov 2015). However, the genomic basis of this process is poorly understood. In Hymenoptera evolution (mainly in parasitoid wasps), miniaturization is thought to be one of the key triggers to species radiation, which may optimize parasitoid lifestyle and allow for successfully attacking a variety of new hosts (Peters et al. 2017). Moreover, many extremely small wasps are found in different superfamilies (mainly in Chalcidoidea and Platygastroidea), suggesting that multiple transition events are likely contributing to the body size reduction (fig. 1A). For example, some species of the early branches of superfamily Chalcidoidea are

in extremely small size, including family Mymaridae and Trichogrammatidae (most of them are smaller than 1 mm), implying the chalcid ancestor is likely to be a miniaturized parasitoid wasp (Lindsey et al. 2018; Peters et al. 2018). In addition, some extremely small-sized wasps are also found in families Eulophidae, Encyrtidae, and Aphelinidae (fig. 1A). However, the phylogeny of Hymenoptera reveals that the closest relatives of Chalcidoidea are superfamily Diaprioidea, with an average body size between 2 and 18 mm (medium-sized in parasitoid wasps) (Peters et al. 2017, 2018). This piece of evidence further supports that the miniaturization event has taken place in the ancestor of chalcid, following divergence from Diaprioidea. However, due to the lack of high-resolution phylogeny in Chalcidoidea, we cannot dismiss the possibility of the independent miniaturizations in each
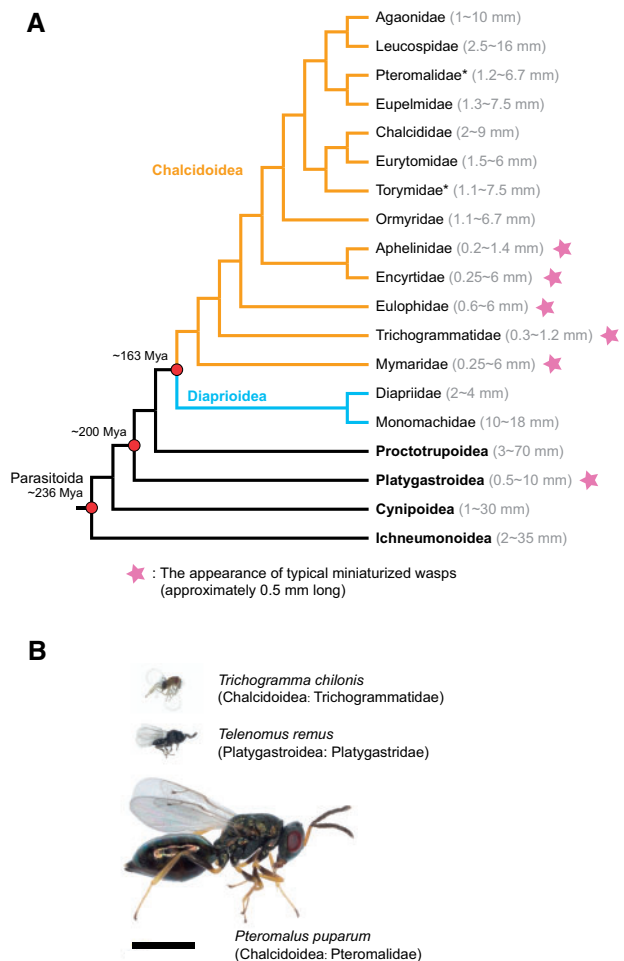
Article

Open Access

Fig. 1. The convergent evolution of miniaturized wasps. (*A*) Independent evolution of miniaturizations on the Chalcidoidea lineage and also on the Platygastroidea lineage. The phylogeny of Hymenoptera was obtained from Peters et al. (2017) and Zhang et al. (2020). The asterisks indicate that Pteromalidae and Torymidae are polyphyletic, and this phylogeny represents the main topology of Hymenoptera evolution. The general body size of each superfamily or family was obtained from Goulet and Huber (1993) and He (2004). Pink stars indicate the presence of miniaturized parasitoid wasps (~0.5 mm) in this group. (*B*) Comparison of body size among three parasitoid wasps, including *Trichogramma chilonis* (Chalcidoidea: Trichogrammatidae), *Telenomus remus* (Platygastroidea: Platygastridae), and *Pteromalus puparum* (Chalcidoidea: Pteromalidae). The photos of parasitoid wasps were taken by the digital microscope VHX-7100 (KEYENCE). Scale bars: 1,000 μm.

species branch. Another dramatic body size reduction event has occurred in superfamily Platygastroidea, with *Telenomus* wasps (family Platygastridae) as representative species (smaller than 1 mm) (Johnson 1984, 1992). Therefore, the miniaturizations in Chalcidoidea and Platygastridae are likely independent, and their lineages represent the classical model of convergent evolution (fig. 1*A*).

In different insect taxa, the decreased body size is accompanied by several concurrent morphological changes, which also support the convergent evolution of miniaturization (Polilov 2015). Indeed, in addition to the reduction of almost all organs and tissues, miniaturized parasitoid wasps share

several distinctive morphological characteristics, including the decreased number of wing veins, the relative volume increase of the central nervous system especially the brain, and a considerable increase of chromatin compaction in the neuron nuclei (van der Woude et al. 2013; Polilov 2015; Diakova et al. 2018; Lindsey et al. 2018). A previous genomic study on *Trichogramma* revealed rapid evolution of its genome and proposed that these changes may contribute to the adaptation to miniaturization (Lindsey et al. 2018).

In this study, we leverage the genomic data of two representative extremely small-sized wasps (~0.5 mm in length) from these two superfamilies and other normal-sized wasps (at least 2 mm long) to investigate the convergent evolution of miniaturization in parasitoid wasps. As for the two tiny wasps, *Trichogramma pretiosum* is a well-known small insect from family Trichogrammatidae (superfamily Chalcidoidea), with adults measuring only about 0.3 mm in length (Lindsey et al. 2018). *Telenomus remus* was chosen to represent the extremely small-sized wasps in superfamily Platygastroidea, with a typical length of about 0.5 mm (fig. 1*B*). As the high-quality reference genome of *Tri. pretiosum* is now available, here, we first report a high-quality chromosome-level genome assembly of *T. remus*, representing the first genome sequence of the superfamily Platygastroidea. The relatively small genome size of *T. remus* is likely due to the reduction of repetitive sequences (including transposable elements [TEs]) and total intron length. Comparative genomics and evolutionary analyses suggest hundreds of genes evolved convergently faster in two tiny parasitoids *Tri. pretiosum* and *T. remus*, which may be relevant to their adaptations of miniaturized body size. Further comparisons using strict cutoffs highlight 38 genes, with extremely accelerated evolutionary rates in *Tri. pretiosum* and *T. remus* specifically, and several of them have been implicated in body/organ size control. Our study significantly advances the understanding of genetic mechanisms underlying convergent miniaturization during parasitoid wasp evolution. Additionally, our high-quality genome assembly presented herein serves as a new genomic resource for the super-diverse Hymenoptera and will greatly facilitate pest control using parasitoid wasps.

Results

Chromosome-Level Assembly and Genomic Characteristics of the Egg Parasitoid Wasp *T. remus*

The genome size of *T. remus* was estimated to be 104.4 Mb based on the flow cytometry method (supplementary fig. 1, Supplementary Material online). Using single-molecule real-time PacBio reads (167.7 Gb), we first generated a contig-level assembly with contig N50 length of 5.0 Mb (table 1 and supplementary table 1, Supplementary Material online). Then, we used chromatin conformation capture (Hi-C) data to improve this assembly into a chromosomal assembly (ZJU_Trem_1.0) comprising ten contiguous chromosomes, consistent with the haploid number (fig. 2*A* and supplementary tables 2 and 3, Supplementary Material online) (Dreyfus and Breuer 1944; Gokhman 2009). Our chromosomal assembly of *T. remus* spans 129.0 Mb with scaffold N50 length of

Table 1. Assembly Statistics for *Telenomus remus* Genome and Other Five Chromosomal Hymenopteran Genomes.

|	*Telenomus remus*	*Pteromalus puparum*	*Nasonia vitripennis*	*Belonocnema treatae*	*Aphidius gifuensis*	*Apis mellifera*
Assembly version	ZJU_Trem_1.0	ZJU_Ppup_chr_1.0	Nvit_psr_1.1	B_treatae_v1	ASM1490517v1	Amel_HAv3.1
Assembly size (Mb)	129.0	338.1	297.3	1,538.7	156.9	225.3
Number of assembled chromosomes	10	5	5	10	6	16
Contig N50 (Mb)	5.0	0.038	7.2	0.02	3.9	5.4
Scaffold N50 (Mb)	11.9	65.8	24.7	151.0	27.5	13.6
Protein-coding genes	15,082	17,656	24,388	14,488	10,443	9,935
Repeats (%)	10.7	44.4	20.0	81.7	28.9	7.8
GC (%)	36.5	40.6	41.7	31.2	26.5	32.5
Complete BUSCO (%)	98.4	98.0	96.5	97.1	99.0	98.1
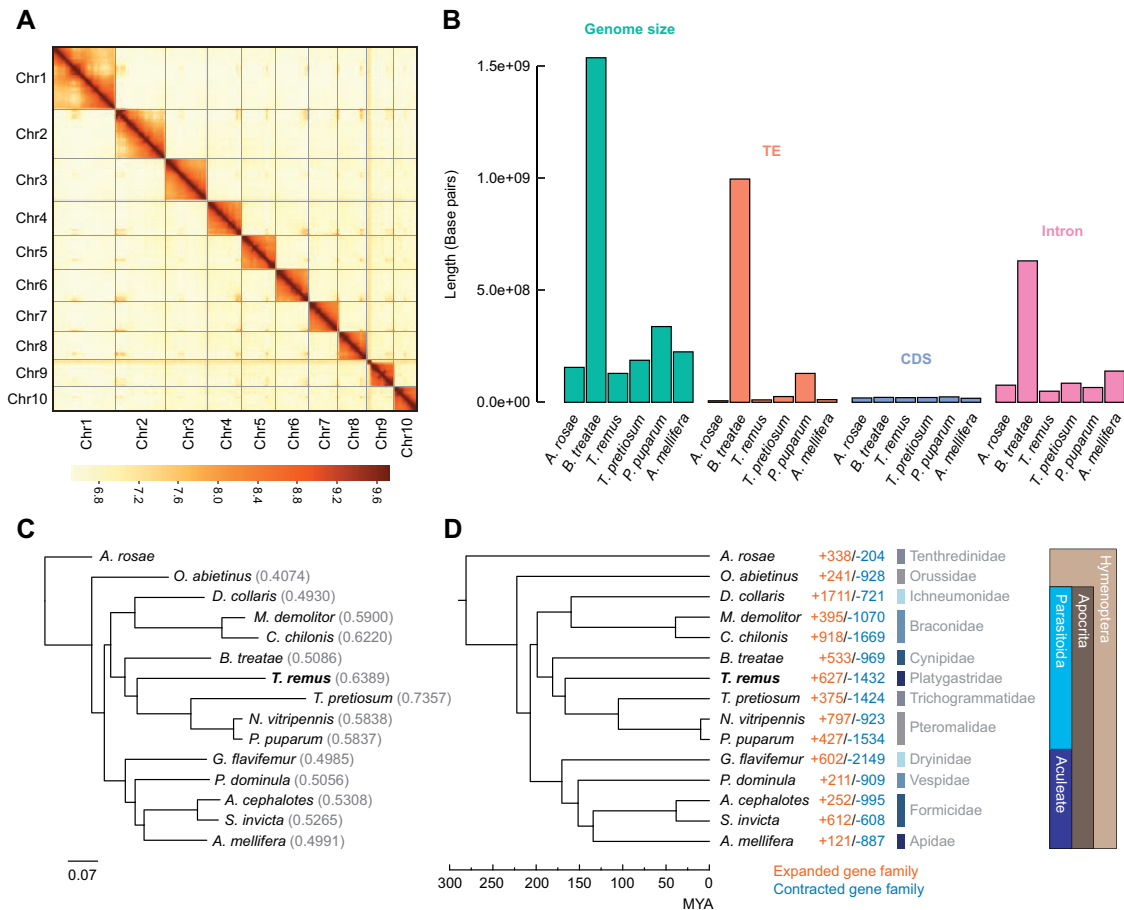
Fig. 2. Chromosome-level genome assembly of *Telenomus remus* and genome evolution of Hymenoptera. (*A*) Heatmap of Hi–C interactions among all chromosomes of *T. remus*. (*B*) The bar plots comparing the assembled genome size, total TE length, total CDS length, and total intron length among six hymenopteran insects. (*C*) Phylogenetic tree constructed using concatenated protein sequences of 2,189 one-to-one orthologous genes. The values in parentheses are the distances (number of substitutions per site) between each of the hymenopterans and the sawfly *Athalia rosae* (outgroup). Bootstrap values based on 1,000 replicates are equal to 100 for each node. (*D*) The Phylogenomic tree with estimated species divergence time. The numbers of inferred gene family expansion (in orange) and contraction (in blue). Totally 9,441 gene families were used for gene family expansion and contraction analysis.

11.9 Mb and over 98% of the assembled sequences have been anchored onto chromosomes (supplementary tables 2 and 3, Supplementary Material online). The assembly is of high integrity and accurate as over 99% of Illumina reads could be successfully mapped to the assembly (supplementary table 4, Supplementary Material online). Moreover, 98.4% complete BUSCO genes have been found in the assembly, which is

comparable to, or better than, those of other published high-quality hymenopteran genomes (table 1).

The genome size of *T. remus* (129 Mb) is smaller than most of the sequenced hymenopteran species (80% are between 180 and 340 Mb) (Branstetter et al. 2018). Compared with the other sequenced hymenopteran genomes, we found that the small size of the *T. remus* genome is likely due to the

reduction of repetitive sequences (including TEs) and total intron length (fig. 2B and supplementary table 5, Supplementary Material online). In total, our assembly reveals that the repeat sequences comprise 10.7% of the genome and most of the repeats were annotated as TEs (9.3% of the genome) (supplementary table 6, Supplementary Material online). The *T. remus* genome has a lower TE proportion compared with many hymenopterans, such as *Pteromalus puparum* (42.1%), *Microplitis demolitor* (35.3%), *Belonocnema treatae* (80.6%), and *Solenopsis invicta* (37.5%). In contrast, we also noticed some hymenopterans with low TE content, including the honeybee *Apis mellifera* (5.6%), the sawfly *Athalia rosae* (4.6%), and *Tri. pretiosum* (14.1%), but their genome sizes are larger than that of *T. remus* (fig. 2B). To investigate the TEs reduction in the *T. remus* genome, we annotated the TEs using known TE libraries in the 15 hymenopteran genomes and performed a series of comparative analyses. In the *T. remus* genome, we identified 40 TE superfamilies from four major TE classes, including DNA transposons (DNA, 2.2%), long-terminal repeats (LTR, 2.0%), long-interspersed nuclear elements (LINE, 0.3%), and short-interspersed nuclear element (SINE, 0.02%) (supplementary table 6, Supplementary Material online). The two most abundant TE superfamilies in the *T. remus* genome are the Gypsy (1.4%) from LTR retrotransposons and the Helitron (0.8%) from DNA transposons, which were also found in other 14 hymenopteran genomes, suggesting these TEs are highly conserved in Hymenoptera. Some of the TE sequences from 64 superfamilies, presented in at least one hymenopteran genome, were lost in the *T. remus* genome. In particular, the *T. remus* genome loses several TEs that are highly conserved across over 80% of the other 14 hymenopteran genomes we examined. For example, DNA/hAT-hATm was lost in the *T. remus* genome, whereas it is abundant (total mean length = 1,278 kb) in 12 out of the 15 hymenopteran genomes tested, including the basal hymenopteran species *Orussus abietinus* and *Ath. rosae*. Other conserved TE losses were found in DNA/TcMar-Mariner, DNA/TcMar-Tigger, and LINE/R1 (supplementary table 7, Supplementary Material online). However, no *T. remus*-specific TE losses were detected.

Genome annotation revealed that 16.0% (20.7 Mb) of the genomic sequences code for proteins in *T. remus*. Although *T. remus* has a larger proportion of coding region than other hymenopterans, the total lengths of the coding sequences in their genome are similar (between 17 and 24 Mb) (fig. 1B). A total of 15,082 protein-coding genes were annotated in the *T. remus* genome. This number is similar to most hymenopterans, but smaller than *Nasonia vitripennis*, which has 24,388 predicted protein-coding genes (table 1). These findings imply that gene loss and coding region reduction may not drive the relatively small genome size of *T. remus*. Additionally, we found that the reduction in intron length also contributed to the reduced genome size of *T. remus* (fig. 2B).

Whole-genome alignment analyses revealed a low level of synteny between the chromosome-level genomes of *T. remus* and other three hymenopterans (*P. puparum*, *B. treatae*, and *A. mellifera*), suggesting that massive genome rearrangements have occurred in the evolutionary history of Hymenoptera (supplementary fig. 2, Supplementary Material online). These results are supported by the previous study of the chromosome synteny between *P. puparum* and *N. vitripennis* based on a CDS pairwise synteny search method (Ye et al. 2020).

Phylogenetic Relationship

We identified 2,189 one-to-one orthologous genes by a mixed phylogeny-network approach within *T. remus* and 14 other hymenopteran species. All the proteins of these 2,189 genes were aligned and concatenated for generating the maximum likelihood (ML) tree. Our ML tree showed that *T. remus* (from family Platygastridae) was placed as a sister to the chalcidoid clade (superfamily Chalcidoidea), which includes three wasps (fig. 2C and D). The oak gall wasp *B. treatae* from the family Cynipidae was identified as the sister lineage of *T. remus* and three chalcidoid wasps. This reconstructed topology is consistent with the topology revealed by a previous study mainly based on transcriptome data (Peters et al. 2017). In addition, we also estimated the divergence time of *T. remus* and other hymenopteran species and found that *T. remus* diverged from other wasps approximately 166.5 Ma, during the Jurassic period (fig. 2D).

Gene Content Comparison and Gene Family Evolution

A total of 12,118 orthogroups (OGs) were identified by a mixed phylogeny-network method (Broccoli) among *T. remus* and the 14 other hymenopteran insects. Broccoli firstly performs phylogenetic analyses on most proteins and builds a network of orthologous relationships. Then, a parameter-free machine learning algorithm is implemented to identify orthologous groups from the orthology network (Derelle et al. 2020). In *T. remus*, 11,465 genes (78.4%) were detected in OGs and 3,617 were unassigned. These 3,617 unassigned genes might be species-specific in *T. remus*, and they were enriched in gene ontology (GO) terms related to digestion, temperature compensation of the circadian clock, and transmission of nerve impulse (FDR-adjusted $P < 0.05$) (supplementary table 8, Supplementary Material online). Comparison between *T. remus* and four other hymenopterans identified 6,629 core OGs and 296 *T. remus*-specific OGs (supplementary fig. 3, Supplementary Material online).

Gene family evolution analysis performed using Cafe revealed 627 expanded gene families and 1,432 contracted gene families on the *T. remus* terminal branch since it separated from other wasps (fig. 2D). A series of gene families related to chemosensory were found to be expanded in the *T. remus* genome, including odorant receptor, gustatory receptor, and G-protein coupled receptor. These may contribute to the host location of the parasitoid wasp. Additionally, we found that the folate transporter gene family was expanded in the *T. remus* genome, which may help *T. remus* transport and utilize folate (also known as vitamin B9) more efficiently.

High Evolutionary Rates of Two Tiny Parasitoid Wasps, *T. remus* and *Tri. pretiosum*

In our phylogenetic analysis, we found that the branch lengths of two tiny parasitoid wasps, *T. remus* and *Tri.*

pretiosum, are longer than that of other hymenopterans, suggesting their higher protein evolutionary rates compared with other hymenopterans analyzed in this study (fig. 2C). Furthermore, *Tri. pretiosum* shows a higher evolutionary rate than *T. remus* (fig. 2C). These results were statistically significant by Tajima's relative rate test (supplementary table 9, Supplementary Material online). To bolster the evidence for faster protein evolution in tiny parasitoid wasps, we performed additional analyses by adding another *Trichogramma* species (*Tri. brassicae*) to the current phylogeny and replacing *Tri. pretiosum* with *Tri. brassicae*. Consistently, both analyses supported the finding that the proteins evolved significantly faster in *Trichogramma* wasps and *T. remus* than in other hymenopterans we tested (supplementary fig. 4 and tables 10 and 11, Supplementary Material online). This result is to be expected, given the close relationship between *Tri. pretiosum* and *Tri. brassicae* within the same genus. It should also be noted that *Tri. brassicae* is the only tiny wasp with currently available genomic data; however, the gene annotation was poor, with only 66.7% complete BUSCO genes (912 genes) can be found (supplementary table 12, Supplementary Material online). Considering that adding in this low-quality genome might introduce bias into the subsequent analyses, we thus decided to exclude this genome from the formal analysis on rapidly evolving protein identification.

To further identify the rapidly evolving proteins in *Tri. pretiosum and T. remus*, we firstly reconstructed the phylogeny of each of the 2,189 one-to-one orthologous proteins obtained from Broccoli and extracted the branch lengths for each protein from each species to the hymenopteran ancestor (HA). Next, we ranked the branch lengths within each OG (table 2). The first place represented the longest branch, suggesting the most amino acid substitutions after divergence from the HA. As expected, we observed a significant tendency of a protein to have the longest branch length in *Tri. pretiosum* or *T. remus* ($P < 0.00001$, χ^2 test). *Trichogramma pretiosum* has 921 first place proteins, and *T. remus* has 503, the next highest number of first place proteins.

The wasp with the third highest number of first place proteins is *Cotesia chilonis*, however, with only 285 first place proteins. The analyses of pairwise branch length density distribution additionally showed that *T. remus* has much longer branch lengths than other hymenopterans but is slightly shorter than *Tri. pretiosum* (fig. 3A–F). These comparisons further support the idea that the proteins in *Tri. pretiosum and T. remus* have evolved under the higher evolutionary rates among hymenopterans.

Proteins with first place assigned to *T. remus* were significantly enriched for six GO terms including intracellular transport, cellular localization, and organelle organization (supplementary table 13, Supplementary Material online). And proteins with first place assigned to *T. pretisosum* were significant enriched in chromatin organization, developmental process, and transcription (supplementary table 14, Supplementary Material online). In contrast, no significantly enriched GO terms were found when we analyzed the *C. chilonis* orthologs, which had the longest branch within each OG (supplementary table 15, Supplementary Material online).

The causes of the rapid protein evolution in *Tri. pretiosum and T. remus* remain unclear. A previous study has suggested that the rapid evolution of proteins in *Tri. pretiosum* might be related to small body size and egg parasitism (Lindsey et al. 2018). Although *Tri. pretiosum and T. remus* belong to different superfamilies in Hymenoptera (Chalcidoidea and Platygastroidea, the divergence time between them was estimated to be about 166.5 Ma), they resemble each other closely and share many important characteristics. For example, they are both egg parasitoid wasps and extremely miniaturized (<0.5 mm). Thus, we hypothesize that the convergently rapid protein evolution in *Tri. pretiosum and T. remus* might be associated with the adaptations to miniaturization, and their specialized egg parasitoid life. The alternative interpretation for rapid protein evolution is that *Tri. pretiosum and T. remus* both have relatively short generation times (within 10 days per generation in the laboratory), which may contribute to their high mutation rates.

Table 2. Numbers of Hymenopteran Proteins with Different Rank (from 1 to 14) in Each Species.

Rank	Nvit	Ppup	Tpre	Trem	Btre	Mdem	Cchi	Dcol	Gfla	Pdom	Acep	Sinv	Amel	Oabi
1	25	38	921	503	54	48	285	43	73	51	52	20	67	14
2	173	134	453	441	52	200	347	46	59	65	59	62	84	16
3	257	267	242	245	80	277	348	51	82	77	83	91	71	21
4	225	226	188	302	101	301	272	92	92	87	119	92	73	22
5	293	250	116	163	145	255	288	95	116	106	141	109	84	31
6	253	273	77	173	114	300	207	94	115	141	160	143	96	43
7	199	206	56	125	226	209	136	149	157	154	200	188	134	51
8	150	186	53	84	220	181	90	186	192	191	230	220	152	58
9	165	138	26	47	215	123	75	196	221	225	228	284	169	72
10	138	140	22	40	206	105	43	213	198	244	268	283	209	81
11	119	132	19	31	192	81	44	200	264	261	250	249	231	115
12	96	110	8	21	229	55	32	286	230	232	213	220	318	136
13	73	68	6	10	236	36	14	368	248	236	129	170	319	275
14	23	21	2	4	119	18	8	170	142	119	57	58	182	1,254

NOTE.—Rank 1, that is, first place, represents a protein having the longest branch length within an orthologous group. Rank 14, also as 14th place, means a protein having the shortest branch length within an orthologous group. Branch lengths were extracted for each protein from each species to the hymenopteran ancestor, indicating the amino acid substitutions after divergence from the hymenopteran ancestor. In total, 2,189 one-to-one orthologous proteins were used for this analysis.

Convergently Accelerated Proteins of *Tri. pretiosum* and *T. remus* May Reflect Adaptations to Body Size Miniaturization

In particular, we noticed a high proportion of rapidly evolving proteins shared between *T. remus* and *Tri. pretiosum*. Among 921 OGs with *Tri. pretiosum* protein as the longest branch (first place proteins), 34.3% (316/921) OGs had *T. remus* protein as second place protein, representing a significantly higher proportion than expected (the expected proportion is 7.7%, $P < 0.00001$, χ^2 test) (supplementary table 16, Supplementary Material online). Similarly, a significantly higher proportion of *Tri. pretiosum* proteins as second place proteins (51.1%, 257/503) was found when the *T. remus*

proteins were assigned as first place proteins (the expected proportion is 7.7%, $P < 0.00001$, χ^2 test) (supplementary table 17, Supplementary Material online). Therefore, we concluded a nonrandom overlap of proteins experienced rate accelerations in both *Tri. pretiosum* and *T. remus* (i.e., convergent rate accelerations). Totally, 573 proteins have the longest and the second longest branches leading to Tpre (*Tri. pretiosum*) and Trem (*T. remus*), that is, Tpre–Trem fast proteins (fig. 3G). GO enrichment analysis revealed that the Tpre–Trem fast proteins were enriched in biological regulation, regulation of cellar process, and regulation of intracellular signal transduction (supplementary table 18, Supplementary Material online).
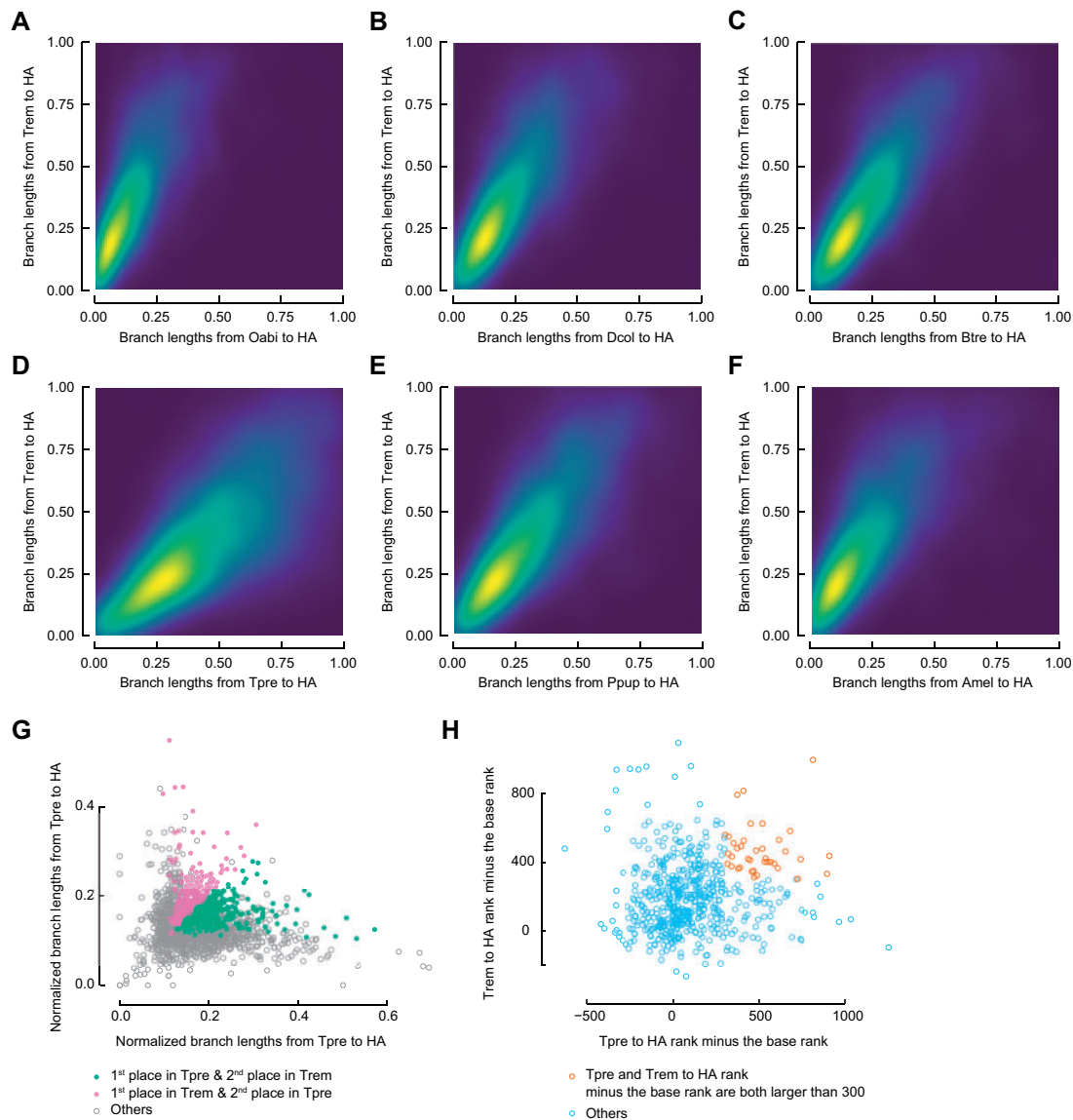

FIG. 3. Accelerated protein evolution of *Telenomus remus* and *Trichogramma pretiosum*. (*A–F*) Two-dimensional density distribution of branch lengths of proteins from species to the hymenopteran ancestor (HA). The *T. remus* (Trem) has much greater branch lengths than other five hymenopteran insects, except *Tri. pretiosum* (Tpre). (*G*) Branch lengths of proteins from Trem and Tpre to the HA, relative to the total branch lengths highlighting proteins with the longest and the second longest branches leading to Tpre and Trem as compared with other hymenopteran insects (first place in Tpre and second place in Trem or first place in Trem and second place in Tpre, i.e., the Tpre–Trem rapid proteins). (*H*) The differences in the Trem and Tpre to HA ranks compared with the base ranks (the ranks of total branch length) of the Tpre–Trem rapid proteins. We highlighted the proteins with the rank difference (The species to HA rank minus the base rank) greater than 300 in both Tpre and Trem.

Next, we developed a rank-based branch length comparison method to further identify the proteins which evolved specifically faster in *Tri. pretiosum* and *T. remus*. In our procedure, we also took the potential influence of baseline evolutionary rates, because protein evolutionary rates vary but are significantly correlated between genomes (Tourasse and Li 2000; Pal et al. 2006). In brief, we firstly assigned ranks to proteins based on their total branch length (from 1 to 2,189), that is, the base rank, representing the baseline rate of each protein. Second, we ranked these proteins by their distance from the *Tri. pretiosum* and *T. remus* to the HA, respectively, that is, the Tpre to HA rank and the Trem to HA rank. Then, we compared the Tpre to HA rank and Trem to HA rank to the base rank of each protein, respectively, and the rank difference implied the evolutionary rate difference between *Tri. pretiosum* or *T. remus* protein and the baseline rate. To determine which proteins largely increased both their Tpre to HA rank and Trem to HA rank when compared with their base rank, we selected the proteins with the rank differences (the species to HA rank minus the base rank) greater than 300 in both *Tri. pretiosum and T. remus*, and obtained totally 43 proteins (supplementary table 19, Supplementary Material online). Compared with the 573 Tpre–Trem fast-evolving proteins, we previously identified, our analysis pinpointed 38 overlapped proteins, likely evolved faster in both *Tri. pretiosum and T. remus* specifically, that is, Tpre–Trem accelerated proteins (fig. 3H and table 3).

To answer which biological functions were enriched among Tpre–Trem accelerated proteins, we first performed GO enrichment analysis, but the result did not show clear GO term enrichments (FDR-adjusted $P > 0.05$), which may due to the relatively small gene set. To investigate the functions of these Tpre–Trem accelerated proteins, we also searched databases of genes/pathways and their functional information (e.g., Uniprot database). We observed a number of proteins related to development including SPRED2, MKP3, STK3, ATX2L, MARK3, PAXBP1, TSSK2, NEURL2, and FGFR1OP2. Interestingly, among them, two proteins (MARK3 and FGFR1OP2) were reported as negative regulators for hippo signaling, which is an evolutionarily conserved signaling pathway that controls organ size during development from flies to humans (Saucedo and Edgar 2007; Staley and Irvine 2012; Kwan et al. 2016; Zheng et al. 2017). Moreover, two proteins (SPRED2 and MKP3) were related to the negative regulations of Ras signaling pathways, which are important in controlling several functions including cell proliferation, growth, and migration (Kim et al. 2004; King et al. 2005). The cell growth and proliferation may also affect cell features and eventually contribute to the body size miniaturization of *Tri. pretiosum and T. remus* (Amodeo and Skotheim 2016). Another two proteins (STK36 and SUFU) belong to the hedgehog signaling pathway, which plays an essential role during growth and patterning in many organs (Kogerman et al. 1999; Maloveryan et al. 2007; Raducu et al. 2016; Han et al. 2019). Mutations of hedgehog signaling pathway genes were found in the tallest terrestrial animal, giraffes, and these changes in hedgehog signaling pathway genes were thought to be related to the body size variation in ruminants (Chen et al. 2019). In

addition, we also found five proteins in the Tpre–Trem accelerated protein set were involved in organ or tissue development including PAXBP1 (muscle organ development), TSSK2 (multicellular organism development), INHBB (neuronal morphogenesis and optic lobe development), ATX2L (chaeta development and compound eye development), and BUN (peripheral nervous system morphogenesis, eye development, and oogenesis). Particularly, we discovered an EIF2B1 protein experienced strong rate acceleration in both *Tri. pretiosum and T. remus* (table 3). EIF2B1 is an essential regulator for protein synthesis, and its homologous protein in mammals regulates eIF2 signaling and may interact with EIF2S1 and EIF2AK3 which resulted in body size variation (Zhang et al. 2002; Gupta et al. 2010). Other Tpre–Trem accelerated proteins were involved in a variety of biological processes, including chromatin organization, transcription, translation, cell–cell adhesion, cell cycle, and protein ubiquitination, which may also contribute to body size miniaturization in these two parasitoid wasps (table 3).

According to our Hymenoptera phylogeny in figure 2C, the rate acceleration related to miniaturizations may occur in three branches independently, including the terminal branches of *Tri. pretiosum* and *T. remus*, and the Chalcidoidea ancestral branch. Therefore, three models of miniaturization related rate acceleration could have happened during the convergent miniaturizations of *Tri. pretiosum* and *T. remus*, including: 1) rate accelerations in the terminal branches of *Tri. pretiosum* and *T. remus*, and the Chalcidoidea ancestral branch; 2) rate accelerations in the terminal branches of *Tri. pretiosum* and *T. remus*; 3) rate accelerations in the terminal branch of *T. remus*, and in the Chalcidoidea ancestral branch. To assess the protein evolutionary rate accelerations in *Tri. pretiosum*, *T. remus*, and the Chalcidoidea ancestor, we applied an additional approach to test if their branches convergently shifted to higher rates than the average rate in each protein. Briefly, we first estimated relative evolutionary rates for all branches of each orthologous protein group by normalizing their branch lengths across all trees to the master branch lengths. Branch lengths are corrected for the heteroskedastic relationship between average branch length and variance by weighted regression. Then, the Wilcoxon rank-sum test and correlation analysis were implemented to test if the branches of interest share higher or lower rates when compared with other branches (Chikina et al. 2016). First, we used this relative evolutionary rate method to identify those proteins that convergently accelerated their evolutionary rates relative to their overall rates of evolution throughout the trees, specifically in the branches leading to the two tiny parasitoid wasps *Tri. pretiosum and T. remus*, that is, their terminal branches, and the Chalcidoidea ancestral branch. In total, 71 proteins showed statistically significant convergently accelerated rates in all three branches ($P < 0.05$, Wilcoxon rank-sum test, supplementary table 20, Supplementary Material online). Nineteen of 71 were previous identified as Tpre–Trem accelerated proteins, including USP46 (a role in behavior, possibly by regulating GABA action), MKP3 (negative regulations of Ras signaling pathways), UBE4B (protein ubiquitination), and ATX2L (chaeta

Table 3. The 38 Accelerated Proteins Specifically in *Trichogramma pretiosum* and *Telenomus remus*.

OG ID	Gene	Annotation	Function	Evolutionary Mode	Strictly Convergent Amino Acid Changes
OG_5183[a,b,c]	USP46	Ubiquitin carboxyl-terminal hydrolase 46	Chromatin organization	RS in both Tpre and Trem	S344L
OG_10449[a,c]	UBE4B	Ubiquitin conjugation factor E4 B	Protein ubiquitination	RS in both Tpre and Trem	L485S, V644A, N828Y
OG_3437[a,c]	ATX2L	Ataxin-2-like protein	Development	PS in Tpre and RS in Trem	P681N
OG_3601[a]	PAXBP1	PAX3- and PAX7-binding protein 1	Development	RS in both Tpre and Trem	K734N
OG_9629[a,b]	Mtf2	Metal-response element-binding transcription factor 2	Transcription	RS in both Tpre and Trem	V82F
OG_977[a,b]	RSRC2	Arginine/serine-rich coiled-coil protein 2	RNA binding	—	Not detected
OG_9799[a,b]	Cdh23	Cadherin-23	Cell–cell adhesion	—	A148V, A318V, T1719V
OG_4013[a,c]		Putative uncharacterized protein	—	—	L1377N
OG_141[a,c]	PSMD10	26S proteasome non-ATPase regulatory subunit 10	—	—	L159A
OG_2860[a]	MFAP1	Microfibrillar-associated protein 1	Cell cycle	—	M29S
OG_3196[a,b]	GTF2E2	General transcription factor IIE subunit 2	Transcription	RS in both Tpre and Trem	Not detected
OG_3642[a]	GDL	Gonadal protein gdl	—	RS in both Tpre and Trem	Not detected
OG_6550[a,c]	TOM1L2	TOM1-like protein 2	Protein transport	RS in both Tpre and Trem	T150S
OG_684[a]	SKT36	Serine/threonine-protein kinase 36	Development		Not detected
OG_8291[a]	POLDIP2	Polymerase delta-interacting protein 3	Translation	RS in both Tpre and Trem	Not detected
OG_10958[a]	Rab3	Ras-related protein Rab-3	Nervous system	—	Q301T
OG_9926[a]	TSSK2	Testis-specific serine/threonine-protein kinase 2	Development	RS in both Tpre and Trem	Not detected
OG_6103[a]	MARK3	MAP/microtubule affinity-regulating kinase 3	Microtubule	—	Not detected
OG_8789[a,b]	Mkp3	Dual specificity protein phosphatase Mpk3	Development	PS in Tpre and RS in Trem	Not detected
OG_4394[b]		Putative uncharacterized protein	—	PS in Tpre and RS in Trem	I302S
OG_9269	MARCHF6	E3 ubiquitin-protein ligase MARCH6	Protein ubiquitination	RS in both Tpre and Trem	Not detected
OG_286[c]	Fgfr1op2	FGFR1 oncogene partner 2 homolog	Development	RS in both Tpre and Trem	Not detected
OG_7012[c]	Slmap	Sarcolemmal membrane-associated protein	Development	—	Not detected
OG_6467[b]	Asator	Tau-tubulin kinase homolog Asator	Microtubule	RS in both Tpre and Trem	Not detected
OG_4467	SS18L1	Calcium-responsive transactivator	Chromatin organization	RS in both Tpre and Trem	Not detected
OG_5607	UBL7	Ubiquitin-like protein 7	Protein ubiquitination	—	Not detected
OG_618	NEURL4	Neuralized-like protein 2	Development	RS in both Tpre and Trem	A2987V
OG_9066[b]	INHBB	Inhibin beta B chain	Nervous system	—	Not detected
OG_6080	cdc14ab	Dual specificity protein phosphatase CDC14AB	Cell cycle	RS in both Tpre and Trem	S432N, V563A
OG_905		Putative uncharacterized protein	—	—	Not detected
OG_3022[b]	SSFA2	Sperm-specific antigen 2	—	—	R528V
OG_7791	SLC2A1	Solute carrier family 2, facilitated glucose transporter member 1	Glucose transporter	RS in both Tpre and Trem	I292F, D519V, N611S
OG_8816[b]	EIF2B1	Translation initiation factor eIF-2B subunit alpha	Translation	RS in both Tpre and Trem	H323E
OG_5384[c]	RABEP1	Rab GTPase-binding effector protein 1	Membrane	RS in both Tpre and Trem	Not detected

(continued)

Table 3. Continued

OG ID	Gene	Annotation	Function	Evolutionary Mode	Strictly Convergent Amino Acid Changes
OG_3344[b]	SUFU	Suppressor of fused homolog	Transcription	RS in both Tpre and Trem	Not detected
OG_10954[b]	BTG3	Protein BTG3	Cell cycle	RS in both Tpre and Trem	Not detected
OG_8575[b]	Spred2	Sprouty-related, EVH1 domain-containing protein 2	Development	PS in Tpre	Not detected
OG_3792	BUN	Protein bunched, class 2/F/G isoform	Nervous system	PS in Tpre and RS in Trem	Not detected

NOTE.—RS, relaxed selection; PS, positive selection.
[a]A protein shows significantly rate accelerations on the terminal branches of *Tri. pretiosum* and *T. remus*, and the Chalcidoidea ancestral branch.
[b]A protein shows significantly rate accelerations on the terminal branches of *Tri. pretiosum* and *T. remus*.
[c]A protein shows significantly rate accelerations on the terminal branches of *T. remus* and the Chalcidoidea ancestral branch ($P < 0.05$, Wilcoxon-rank sum test).

development and compound eye development) (table 3). Three representative cases of the accelerated proteins in the terminal branches of *T. remus* and *Tri. pretiosum* and the Chalcidoidea ancestral branch were shown in figure 4A–C. Additionally, we also identified the proteins that convergently accelerated their evolutionary rates, specifically in the terminal branches of *Tri. pretiosum* and *T. remus*, and obtained totally 39 candidate proteins ($P < 0.05$, Wilcoxon rank-sum test, supplementary table 21, Supplementary Material online). Among them, 14 proteins were identified as Tpre–Trem accelerated proteins, including SUFU (a negative regulator in the hedgehog/smoothened signaling pathway), EIF2B1 (an essential regulator for protein synthesis), INHBB (neuronal morphogenesis and optic lobe development), and SPRED2 (a negative regulator of Ras signaling pathway) (table 3). These proteins showed significant rate accelerations specifically in the *Tri. pretiosum* and *T. remus* lineages, but not in the Chalcidoidea ancestor (fig. 4D–F). Lastly, we detected the proteins with significant rate accelerations specifically in the *T. remus* branch and the Chalcidoidea ancestor, and finally found 48 candidate proteins ($P < 0.05$, Wilcoxon rank-sum test, supplementary table 22, Supplementary Material online). Nine of them were in the Tpre–Trem accelerated protein set, including FGFR1OP2 (a negative regulator for hippo signaling) (fig. 4G–I). Overall, our relative evolutionary rate analyses based on three different convergent evolutionary models therefore explained the rate acceleration models of most of the Tpre–Trem accelerated proteins (30/38), indicating they may experience different convergent evolutionary models during the miniaturizations of *Tri. pretiosum* and *T. remus*.

Additionally, we applied this relative evolutionary rate analysis to two new genomics data sets: 1) adding *Tri. brassicae* to the original data set; and 2) replacing *Tri. pretiosum* with *Tri. brassicae* in the original data set. In the first data set, 25 of 38 previously identified Tpre–Trem accelerated proteins were still found as a single copy in all tested species. And among them, 16 proteins showed significant rate acceleration in relevant branches (supplementary table 23, Supplementary Material online). In the second data set, 24 of 38 identified Tpre–Trem accelerated proteins were single-copy proteins, and half of them were supported by relative evolutionary

rate analysis (supplementary table 24, Supplementary Material online). In total, 19 of 38 previously identified Tpre–Trem accelerated proteins were verified in our additional tests, despite a relatively incomplete genome of *Tri. brassicae* (only keeping 66.7% complete BUSCO genes).

Taken together, our scanning and functional annotation provide evidence that hundreds of proteins evolved faster in two tiny parasitoids *Tri. pretiosum* and *T. remus* than other hymenopterans we tested. Several proteins in these two parasitoid wasps showed significant rate accelerations with different convergent evolutionary models, which might be associated with the adaptations to body size miniaturization and might be relevant to the astonishing radiation of parasitoid wasp species (Peters et al. 2017).

Relaxed Constraint May Mainly Drive the Rates Accelerations in *Tri. pretiosum* and *T. remus*

The Tpre–Trem accelerated proteins/genes we identified above could have resulted from either positive selection or relaxed constraint. To figure out the contribution of positive selection in these Tpre–Trem accelerated genes, we applied an adaptive branch-site random effects likelihood model (aBSREL) to test whether a proportion of sites have evolved under positive selection in the *Tri. pretiosum* and *T. remus* terminal branches. Among 38 genes, we detected positive selected signals in five *Tri. pretiosum* genes but no *T. remus* gene showed evidence of positive selection. This result suggested that the evolutionary models of these five genes may be different between these two parasitoid wasps, and the contribution of positive selection to the rate accelerations was relatively small. Next, we scanned these Tpre–Trem accelerated genes using RELAX, to determine whether the strength of natural selection has been relaxed or intensified along the *Tri. pretiosum* and *T. remus* lineages. A total of 20 (52.6%) genes showed relaxed constraint in both *Tri. pretiosum* and *T. remus* lineages (table 3). This is significantly higher than the proportion of the genes under relaxed selection (35.9%) in all 2,189 single-copy genes ($P < 0.05$, χ^2 test). Additionally, we also found that the genes under relaxed selection were significantly likely to be assigned as the previously identified Tpre–Trem fast genes (249/573), indicating the important contribution of
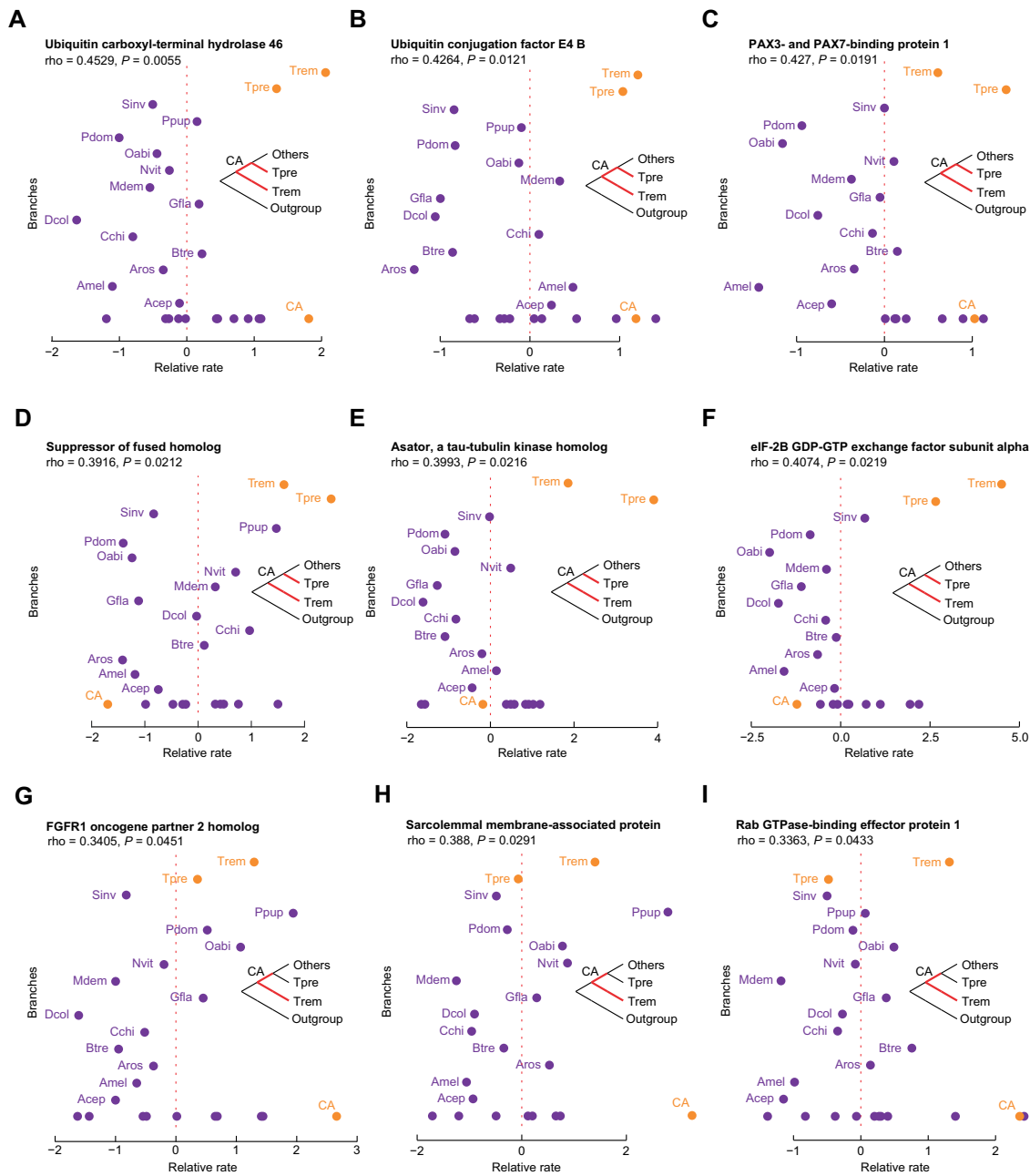
FIG. 4. Nine representative cases of the accelerated proteins in *Telenomus remus* and *Trichogramma pretiosum* based on three different convergent evolutionary models. RERconverge package was used to identify proteins with convergent rates of evolution in *T. remus* (Trem) and *Tri. pretiosum* (Tpre), and the Chalcidoidea ancestor (CA) ($P < 0.05$, Wilcoxon-rank sum test). (*A–C*) These three cases indicate that Trem branch, Tpre branch, and CA branch showed strong evolutionary rate acceleration compared with other branches (purple points), suggesting that rate accelerations occurred in the terminal branches of *Tri. pretiosum* and *T. remus*, and the Chalcidoidea ancestral branch, convergently. (*D–F*) These three cases indicate that Trem branch and Tpre branch showed strong evolutionary rate acceleration compared with other branches (purple points and CA branch), implying that rate accelerations occurred in the terminal branches of *Tri. pretiosum* and *T. remus*, convergently. (*G–I*) These three cases indicate that Trem branch and CA branch showed strong evolutionary rate acceleration compared with other branches (purple points and Tpre branch), suggesting that rate accelerations occurred in the terminal branches of *T. remus* and the Chalcidoidea ancestral branch, convergently. Bolded red branches in each phylogeny indicate the evolutionary positions of rate accelerations of each protein. Aros, *Ath. rosae*; Oabi, *O. abietinus*; Dcol, *D. collaris*; Mdem, *M. demolitor*; Cchi, *C. chilonis*; Nvit, *N. vitripennis*; Ppup, *P. puparum*; Btre, *B. treatae*; Gfla, *G. flavifemur*; Pdom, *P. dominula*; Sinv, *S. invicta*; Acep, *Atta cephalotes*; Amel, *A. mellifera*.

relaxed constraint in the rates accelerations in *Tri. pretiosum* and *T. remus* ($P < 0.01$, χ^2 test). In summary, our selection analyses implied that relaxed constraint might mainly drive the rate accelerations in *Tri. pretiosum* and *T. remus*.

No Evidence of the Convergent Amino Acid Changes Contributed to the Rate Accelerations

Site-specific convergent amino acid changes have been reported in convergent phenotypic evolutionary cases, and

they might contribute to convergent rate shifts (Chikina et al. 2016). Therefore, we investigated the convergent amino acid changes in the evolution of the 38 Tpre–Trem accelerated proteins. We identified all strictly convergent amino acid changes using a PCOC detection pipeline (posterior probability > 0.9, PCOC model) by integrating both protein alignments and the phylogeny of Hymenoptera (Rey et al. 2018). Strictly convergent amino acid sites were defined as those that converged to the exact same amino acid in both *Tri. pretiosum* and *T. remus*. Totally, we found strictly convergent amino acid changes in 17 proteins (table 3). We therefore speculate that the convergent rate accelerations in those proteins may not mainly result from the convergence amino acid substitutions. However, these convergence amino acid substitutions may also play essential roles in adaptations to miniaturization in *Tri. pretiosum* and *T. remus*, during their convergent evolution.

Discussion

This study reported a chromosome-level genome assembly for the parasitoid wasp *T. remus*, representing the first genome sequence of the family Platygastridae and the superfamily Platygastroidea. Our genome sequencing, annotation, and downstream analyses advance the knowledge of *T. remus* genomics and genetics, which will guide future research within the Hymenoptera and parasitoid wasp community.

The relatively small genome size of *T. remus* has resulted from the reduction of repeats (including TEs) and intron size. The reduction of TE may also lead to the shortened intron size in *T. remus*, as intron size is correlated to TE number in *Drosophila melanogaster* (Cridland et al. 2013). Although we also speculated a possible relationship between genome size and body size, our data failed to support this hypothesis, because the three sequenced small-sized *Trichogramma* wasps have larger genome sizes with moderate repeat contents (~200 Mb) (Lindsey et al. 2018; Ferguson et al. 2020).

Miniaturization is regarded as one of the principal directions in parasitoid wasp evolution, which may allow parasitoids to attack a variety of new hosts successfully and trigger species radiation during Hymenoptera evolution (Polilov 2015; Peters et al. 2017). By integrating our survey of body size and the phylogeny of parasitoid wasps (Peters et al. 2017), we found multiple probable significant miniaturization events are more likely located in two superfamilies during parasitoid wasp evolution: one is superfamily Platygastroidea (e.g., *T. remus* which we sequenced in this study), the other is superfamily Chalcidoidea (e.g., *Trichogramma* wasps) (fig. 1). However, we note that the miniaturization may not take place at the platygastroidean ancestor, since the more basal branches of Platygastroidea species remain a medium body size, for example, *Inostemma* sp. (~2 mm length). Previous phylogenic studies suggested that the chalcid ancestor is likely a miniaturized wasp, implying miniaturization may have occurred at the chalcid ancestor (Peters et al. 2018; Zhang et al. 2020). The more basal species of Chalcidoidea are in small sizes, such as family Mymaridae (known as fairyfly) and Trichogrammatidae (most of them are smaller than 1 mm),

whereas several species further reverted to a larger body size during Chalcidoidea evolution, for example, some species of family Leucospidae are larger than 10 mm (Ye et al. 2017). According to their evolutionary positions in the Hymenoptera tree (Peters et al. 2017), it is likely that the miniaturizations in Platygastroidea and Chalcidoidea occurred independently. Therefore, in our comparative genomics and evolutionary analyses, the miniaturizations in *T. remus* and *Tri. pretiosum* were considered as two independent evolutionary events, representing an example of convergent evolution.

Interestingly, as an advance of the previous study on the rapid evolution of *Tri. pretiosum* (Lindsey et al. 2018), our analyses revealed that about a fourth of all single-copy coding sequences (573/2,189) of the two extremely small wasps (*T. remus* and *Tri. pretiosum*, in ~0.5 mm) show a rapid evolution signature compared with other hymenopterans examined. We thus hypothesize that their convergently rapid evolution may contribute to the adaptation of their miniaturizations. Next, a strict rank-based branch length comparison method was used to further identify the proteins that evolved specifically faster in *Tri. pretiosum* and *T. remus*, highlighting 38 Tpre–Trem accelerated proteins. Following relative evolutionary rate analyses provided evidence that these Tpre–Trem accelerated proteins may experience different convergent evolutionary models during the miniaturizations of *Tri. pretiosum* and *T. remus*, including: 1) rate accelerations in the terminal branches of *Tri. pretiosum* and *T. remus*, and the Chalcidoidea ancestral branch, convergently; 2) rate accelerations in the terminal branches of *Tri. pretiosum* and *T. remus*; 3) rate accelerations in the terminal branch of *T. remus*, and in the Chalcidoidea ancestral branch, convergently. Our findings also reveal the complexity of miniaturization evolution in the parasitoid wasps.

Miniaturization in animals is an evolutionary process frequently followed by structural simplification and size reduction of organs, tissues, and cells (Kogerman et al. 1999; Polilov 2015; Diakova et al. 2018). Among 38 Tpre–Trem accelerated proteins, we found a number of proteins with functional categories in development, cell size, chromatin organization, transcription, and translation, which may be associated with the independent miniaturizations in these two tiny parasitoid wasps. For example, two convergently accelerated proteins (MARK3 and FGFR1OP2) were reported as negative regulators for Hippo signaling pathway, which controls organ size in animals through the regulation of cell proliferation and apoptosis (Saucedo and Edgar 2007; Staley and Irvine 2012; Kwan et al. 2016; Zheng et al. 2017). MARK3 is an important protein in the regulation of microtubule cytoskeleton organization, axis specification, and cell polarity, and knockdown of the orthologous protein of MARK3 in *Drosophila* (par-1, flybase ID: FBgn0260934) leads to a significant reduction of eye size (Ansar et al. 2018). FGFR1OP2 belongs to the SIKE family and localizes to the Striatin-interacting phosphatase and kinase (STRIPAK) complex, which is involved in numerous cellular and developmental processes in eukaryotic organisms (Kuck et al. 2019). In addition, SPRED2 (orthologous to Spred in *Drosophila*, flybase ID: FBgn0020767), a sprouty-related

suppressor of Ras signaling, is essential for the differentiation of neuronal cells and myocytes (Wakioka et al. 2001). MKP3 is a well-known negative regulator in the Ras signaling pathway responsible for cell fate determination and proliferation during development, including photoreceptor cell differentiation and wing vein formation (Kim et al. 2004). In *Drosophila*, down-regulation of the Ras signaling pathway inhibits wing vein formation (Guichard et al. 1999). Consistently, miniaturizations in *Tri. pretiosum* and *T. remus* have both resulted in a highly reduced wing venation. We therefore speculate the rate accelerations of the *MKP3* orthologs in *Tri. pretiosum* and *T. remus* may contribute to the wing vein simplification during their independent miniaturizations. Moreover, EIF2B1 is an essential regulator for translational initiation, and may participate in the human body size variation by regulating eIF2 signaling (Zhang et al. 2002; Gupta et al. 2010). BUN is an essential protein in cell proliferation and its homolog in mammals is known as the mammalian tumor suppressor TSC-22, which promotes cellular growth (Gluderer et al. 2008). The convergently accelerated proteins discussed above may therefore explain the miniaturization in *Tri. pretiosum* and *T. remus* from different aspects, such as eye development, wing development, and nervous system. We envision further functional investigations of these proteins we highlighted in regulating the body size of parasitoid wasp, especially for the proteins that have not been documented to regulate body size so far. In this work, we analyzed the evolution of coding sequences across Hymenoptera and successfully identified convergently accelerated proteins in *Tri. pretiosum* and *T. remus*, which might be related to their miniaturizations. However, in animals, the changes in regulatory sequences also play critical roles in morphological changes (Carroll 2008). Thus, we suggest future work to assess the rate shifts in noncoding regions, for example, *cis*-regulatory sequences. Additionally, more genome sequencing data of tiny parasitoid wasps (e.g., from families Mymaridae, Aphelinidae Encyrtidae, and Eulophidae in fig. 1A) will help to provide more robust evidence to support our findings about convergent miniaturization in parasitoid wasp evolution.

Our comprehensive analysis revealed that more than half of the Tpre–Trem accelerated genes were under a significantly relaxed constraint on both *Tri. pretiosum* and *T. remus* branches ($P < 0.05$, RELAX), suggesting a contribution of relaxed selection to rate accelerations of these genes. It is worth mentioning that some genes under relaxed selection may also have experienced short episodes of positive selection (Anisimova et al. 2001). Relaxed selection may cause the accumulation of deleterious genetic variants in genes and lead to loss of function. However, relaxed selection may also not change the functions of genes (Chikina et al. 2016; Cui et al. 2019). Further functional studies should be carried out to test the functions of these accelerated genes.

The parasitoid wasps are among the most important natural enemies of agricultural and forest pests and have been used as biological control agents for a long time (Chen et al. 2014). Understanding the genetic basis of the parasitoid wasp miniaturization might also benefit pest control. It can, for instance, be used to guide gene-editing techniques to create small-size parasitoid wasps that are more accessible to pests

(i.e., can enter some cracks to find the host) and consume less energy than larger parasitoids. The parasitoid wasp we sequenced (*T. remus*) in this study is a promising biological control agent against *Spodoptera* pests, including the fall armyworm *Spodoptera frugiperda*, a notorious lepidopteran pest that has rapidly spread over the world in the past few years (Cock et al. 1985; Kenis et al. 2019; Liao et al. 2019).

Overall, our findings bridge the gap between evolutionary rate accelerations and the convergent evolution of miniaturization in parasitoid wasps, which is one of the major trends in the evolution of parasitoid wasps. In parallel, we also screen out a number of accelerated candidate genes that may be associated with body size reduction. This study provides a valuable genomic resource for *T. remus* as well as insights into the genome evolution of parasitoid wasps. Lastly, our high-quality genome underpins further research of *T. remus* and greatly facilitates pest control.

Materials and Methods

Genome Size Estimation, Genome Sequencing, and Transcriptome Sequencing

The genome size of *T. remus* was estimated based on the flow cytometry method described in (He et al. 2016). More than 500 adults of *T. remus* were used for genomic DNA extraction by the sodium dodecyl sulfate extraction method (Zhou et al. 1996). Both PacBio Sequel and Illumina Hiseq Xten platforms were applied to sequence the genome of *T. remus*. For Illumina sequencing, a 350-bp insert Illumina TruSeq fragment was constructed from the qualified genomic DNA using a Truseq Nano DNA HT Sample preparation Kit, and then sequenced on the Hiseq Xten platform. Raw reads with low-quality bases, adapter sequences, and reads containing poly-N were removed using the Fastp v.0.20.0 (Chen et al. 2018), and the clean reads were used for subsequent analysis. For long-read sequencing, a SMRT bell library was constructed using the PacBio SMRTbell Express Template Prep Kit 2.0, and then sequenced on one SMRT cell using the PacBio Sequel sequencer. We also generated a paired-end RNA-seq library from pooled adults, and sequenced using the Illumina HiSeq Xten platform.

Genome Assembly and Hi-C Scaffolding

To obtain the primary genome assembly, SMARTdenovo (Liu et al. 2021) was used to assemble the genome of *T. remus* with the PacBio long reads with default parameters. Then, we polished the assembly with the Illumina short reads using Pilon v1.24 (Walker et al. 2014) with default parameters.

To further improve the primary genome assembly to the chromosomal level, we prepared the Hi-C library following the standard protocol modified for application to whole insects (Xiao et al. 2020; Ye et al. 2020). Then, Hi-C libraries were quantified and sequenced using the Illumina HiSeq platform (insert size is 350 bp). Next, we used HiC-Pro v2.8.0 (Servant et al. 2015) for quality control with default parameters. After filtration, clean sequencing reads were mapped to the draft genome by Bowtie2 v2.3.2 (Langmead and Salzberg 2012). The unique paired-end reads mapping close to the

restriction sites were retained. Finally, LACHESIS was applied to produce the chromosome-level assembly for *T. remus*. BUSCO v5 (Simao et al. 2015) was used to assess the genome assembly completeness with the insect (odb10) protein set.

Repeat Sequence Annotation

To identify the TEs in the *T. remus* and other 15 hymenopteran genomes, we first constructed species special de novo repeat libraries for each species using RepeatModeler2 with the LTR structural discovery pipeline (Flynn et al. 2020). Then, we screened for TEs and low-complexity DNA sequences in each genome using RepeatMasker v4.0.7 (Tempel 2012) against both the de novo repeat library generated by RepeatModeler2 and the RepBase v26.03 library (Bao et al. 2015).

Chromosomal Synteny

Chromosomal synteny analysis was performed by D-GENIES online tool (http://dgenies.toulouse.inra.fr/) (Cabanettes and Klopp 2018) based on repeat-masked chromosomal genomes with default parameters. In this study, we selected four species for chromosomal synteny analysis, including *T. remus* (this study), *P. puparum* (Ye et al. 2020), *B. treatae* (NCBI: GCF_010883055.1), and *A. mellifera* (Wallberg et al. 2019).

Genome Annotation

We predicted protein-coding gene models from the *T. remus* genome by integrating three approaches including ab initio gene prediction, gene expression-based gene prediction, and homology-based gene prediction. Ab initio gene prediction was performed by BRAKER v2.1.5 (Brůna et al. 2021), a combination of self-training GeneMark-ET ver 4.65_lic (Ter-Hovhannisyan et al. 2008) and AUGUSTUS v3.1 (Stanke et al. 2004). For gene expression-based gene prediction, The RNA-seq reads were first aligned to the reference genome using HISAT2 v2.2.1 (Kim et al. 2015). Then, StringTie v2.1.0 (Pertea et al. 2015) was used to assemble the potential transcripts. Next, we used TransDecoder v5.4.0 (https://github.com/TransDecoder/TransDecoder) to identify the candidate coding regions within each transcript sequence. For homology-based gene prediction, Arthropoda proteins were firstly obtained from OrthoDB (https://v100.orthodb.org/) and aligned to the masked genome by GenomeThreader v1.7.1 (Gremme et al. 2005). All of the predicted gene models above were then combined to create a consensus gene set using EVidenceModeler v1.1.1 (Haas et al. 2008). Protein sequences encoded by gene models against the SwissProt and Trembl databases using BLASTP (e value <1e-5) for functional annotation. Protein domains of gene models were identified by InterproScan-5 (Jones et al. 2014). GO annotation analysis was performed using Blast2Go v5.2 (Conesa et al. 2005).

Comparative Genomics

We used Broccoli v1.2 (Derelle et al. 2020), a mixed phylogeny-network approach, to identify high-precision orthologous groups among the 15 hymenopteran genomes including *Ath. rosae* (RefSeq assembly accession number GCF_000344095.2), *O. abietinus* (RefSeq assembly accession number GCF_000612105.2), *Diadromus collaris* (GenBank assembly accession number GCA_009394715.1), *M. demolitor* (RefSeq assembly accession number GCF_000572035.2), *C. chilonis* (http://www.insect-genome.com), *B. treatae* (RefSeq assembly accession number GCF_010883055.1), *T. remus* (this study), *Tri. pretiosum* (RefSeq assembly accession number GCF_000599845.2), *N. vitripennis* (RefSeq assembly accession number GCF_009193385.2), *P. puparum* (Ye et al. 2020), *Gonatopus flavifemur* (Yang et al. 2021), *Polistes dominula* (RefSeq assembly accession number GCF_001465965.1), *Atta cephalotes* (RefSeq assembly accession number GCF_000143395.1), *S. invicta* (RefSeq assembly accession number GCF_016802725.1), *A. mellifera* (RefSeq assembly accession number GCF_003254395.2). The basal hymenopteran *Ath. rosae* was used as an outgroup. The comparison of the assembled genome size, total TE length, total CDS length, and total intron length among hymenopteran insects was conducted by in-house python scripts. We used Cafe v4.2.1 (De Bie et al. 2006) to analyze the evolution of the gene family sizes, with the results from Broccoli and the phylogenetic tree with divergence times as inputs.

Phylogenetic Analysis

The phylogenetic relationships between *T. remus* and other hymenopterans were determined using a genome-wide set of 2,189 one-to-one orthologous genes which were obtained from Broccoli v1.2 (Derelle et al. 2020). The protein sequences of one-to-one orthologous genes were aligned using MAFFT v.7.305b with the L-INS-i algorithm (Katoh and Standley 2013). Poorly aligned regions in each multiple sequence were eliminated by trimAl v1.2 (Capella-Gutierrez et al. 2009). The phylogenetic tree was reconstructed using the ML criterion with IQ-TREE v2.1.2 based on concatenated protein sequences (Nguyen et al. 2015). The best-fitting model was inferred by ModelFinder implemented in IQ-TREE v2.1.2 (Kalyaanamoorthy et al. 2017). Statistical support for the phylogenetic tree was assessed by Ultrafast bootstrap analysis using 1,000 replicates. Divergence times were originally inferred using r8s v1.81 (Sanderson 2003). Six calibration time points were based on the previous study: Orussoidea + Apocrita: 211–289 Ma, Apocrita: 203–276 Ma, Ichneumonoidea: 151–218 Ma, Chalcidoidea: 105–159 Ma, Aculeata: 160–224 Ma, Cynipoidea: 181–246 Ma (Peters et al. 2017).

Protein Evolutionary Rate Comparison

First, we performed Tajima's relative rate test for the pairwise comparison of the protein evolutionary rates (pairwise distances to the HA) between *T. remus* and other 14 hymenopteran species using the R package pegas (Paradis 2010) based on concatenated protein sequences. Moreover, we also performed Tajima's relative rate test in two additional genomic data sets: one with *Tri. brassicae* (GenBank assembly accession number GCA_902806795.1) added to the original data and the other one replacing *Tri. pretiosum* with *Tri. brassicae* in the original data. To construct the phylogenies of these two data sets, the "Broccoli-MAFFT-trimAl-IQ-TREE" pipeline we

described above was used. In total, 1,574 one-to-one orthologous genes were obtained for the first data set, and 1,618 one-to-one orthologous genes were obtained for the second data set. However, due to the low score of gene annotation of *Tri. brassicae* (with only 66.7% complete BUSCO genes), we therefore excluded this genome from the following analysis on rapidly evolving protein identification. To identify proteins with higher evolutionary rates in *T. remus* and *Tri. pretiosum*, we reconstructed the phylogeny of each of the 2,189 one-to-one orthologous proteins using IQ-TREE v2.1.2 (Nguyen et al. 2015) with the option "-m MFP," and extracted the branch lengths and calculated the distances from each species to the HA in each OG. Then, we assigned species from 1 to 14 within each OG (rank 1, i.e., first place, indicates a protein has the longest branch length within an orthologous group, whereas rank 14, that is, 14th place, means a protein has the shortest branch length within an orthologous group). Proteins with the longest or the second longest branches leading to *Tri. pretiosum* and *T. remus* were identified as Tpre–Trem fast proteins. To further identify the proteins which evolved specifically faster in *Tri. pretiosum and T. remus*, we employed a rank-based branch length comparison method describe in Lindsey et al. (2018). Briefly, we assigned each OG from 1 to 2,189 within all the single copy orthologous groups based on its total branch length (base rank, i.e., overall rank) and the distance from *Tri. pretiosum and T. remus* to the HA, respectively. Rank difference between species to HA and base rank represents the rate accelerated rank or rate decelerated rank. We thus selected the proteins with the rank difference greater than 300 in both *Tri. pretiosum and T. remus* as Tpre–Trem accelerated proteins. In additional, RERconverge v0.1.0 (Chikina et al. 2016; Partha et al. 2017) was used to estimates the correlation between relative evolutionary rates of genes and the evolution of a convergent trait across our hymenopteran phylogeny.

Convergent Amino Acid Substitutions

Convergent amino acid substitutions were detected using the PCOC detection pipeline with option "-f 0.9" for Tpre–Trem accelerated proteins (Rey et al. 2018). The gene tree and amino acid alignment of each Tpre–Trem accelerated protein generated in the previous analysis were used as inputs. Our strict criterion only identified the strictly convergent amino acid sites, which were defined as those that converged to the exact same amino acid in both *Tri. pretiosum and T. remus*.

Selection Analysis

All single copy genes were subjected to phylogenetic models of codon evolution in order to detect signatures of positive selection and relaxation of constraint over the *T. pretisoum and T. remus* branches. For positive selection, we used the adaptive branch-site random effects model (aBSREL) implemented in the HyPhy software package to fit the full adaptive model and run the likelihood ratio test to compare the full model to a null model where the branches are not allowed to have rate classes of ω (dN/dS) >1 (Smith et al. 2015). For relaxed selection, the likelihood ratio test in RELAX was utilized to comparing the model fixing $k = 1$ and the model

allows k to be estimated (Wertheim et al. 2015). The *Tri. pretiosum* and *T. remus* branches were alternatively set as the foreground branches for the test. A significant result of $k < 1$ indicates that selection strength has been relaxed along the test branches.

Functional Enrichment Analysis

All GO enrichment analyses were conducted by GOATOOLS v1.0.6 (Klopfenstein et al. 2018).

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Author Contributions

Z.L., G.Y., and F.L. conceived and supervised the project. H.X. and Yajun Yang collected the samples. K.H. performed flow cytometry experiments. X.Y. led the bioinformatic analyses and Yi Yang and Y.M. participated in the analyses. H.X., X.Y., and Yi Yang drafted the manuscript. X.Y., Yi Yang, H.X., Z.L., G.Y., F.L., L.X., Y.H.S., S.X., and Q.F. revised the manuscript. All authors have read and approved the final manuscript.

Data Availability

All the data generated in this study are available at the National Center for Biotechnology Information (NCBI), under BioProject number PRJNA723667. The genome assembly has been deposited at GenBank under accession number JAGTPF000000000.

References

Amodeo AA, Skotheim JM. 2016. Cell-size control. *Cold Spring Harb Perspect Biol*. 8(4):a019083.

Anisimova M, Bielawski JP, Yang ZH. 2001. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. *Mol Biol Evol*. 18(8):1585–1592.

Ansar M, Chung H, Waryah YM, Makrythanasis P, Falconnet E, Rao AR, Guipponi M, Narsani AK, Fingerhut R, Santoni FA, et al. 2018. Visual impairment and progressive phthisis bulbi caused by recessive pathogenic variant in MARK3. *Hum Mol Genet*. 27(15):2703–2711.

Bao W, Kojima KK, Kohany O. 2015. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob DNA*. 6:11.

Branstetter MG, Childers AK, Cox-Foster D, Hopper KR, Kapheim KM, Toth AL, Worley KC. 2018. Genomes of the Hymenoptera. *Curr Opin Insect Sci*. 25:65–75.

Brůna T, Hoff KJ, Lomsadze A, Stanke M, Borodovsky M. 2021. BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genom Bioinform*. 3(1):lqaa108.

Cabanettes F, Klopp C. 2018. D-GENIES: dot plot large genomes in an interactive, efficient and simple way. *PeerJ*. 6:e4958.

Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25(15):1972–1973.

Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134(1):25–36.

Chen L, Qiu Q, Jiang Y, Wang K, Lin Z, Li Z, Bibi F, Yang Y, Wang J, Nie W, et al. 2019. Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits. *Science* 364(6446):eaav6202.

Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34(17):i884–i890.

Chen XX, Tang P, Zeng J, van Achterberg C, He JH. 2014. Taxonomy of parasitoid wasps in China: an overview. *Biol Control*. 68:57–72.

Chikina M, Robinson JD, Clark NL. 2016. Hundreds of genes experienced convergent shifts in selective pressure in marine mammals. *Mol Biol Evol*. 33(9):2182–2192.

Cock MJW, Bennett FD, Hughes IW, Simmonds FJ, Yaseen M. 1985. A review of biological control of pests in the Commonwealth Caribbean and Bermuda up to 1982. Farnham Royal (United Kingdom): Commonwealth Agricultural Bureaux.

Consea A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21(18):3674–3676.

Cridland JM, Macdonald SJ, Long AD, Thornton KR. 2013. Abundance and distribution of transposable elements in two Drosophila QTL mapping resources. *Mol Biol Evol*. 30(10):2311–2327.

Cui R, Medeiros T, Willemsen D, Iasi LNM, Collier GE, Graef M, Reichard M, Valenzano DR. 2019. Relaxed selection limits lifespan by increasing mutation load. *Cell* 178(2):385–399.e320.

De Bie T, Cristianini N, Demuth JP, Hahn MW. 2006. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* 22(10):1269–1271.

Derelle R, Philippe H, Colbourne JK. 2020. Broccoli: combining phylogenetic and network analyses for orthology assignment. *Mol Biol Evol*. 37(11):3389–3396.

Diakova AV, Makarova AA, Polilov AA. 2018. Between extreme simplification and ideal optimization: antennal sensilla morphology of miniaturized *Megaphragma* wasps (Hymenoptera: Trichogrammatidae). *PeerJ* 6:e6005.

Dreyfus A, Breuer ME. 1944. Chromosomes and sex determination in the parasitic Hymenopteron *Telenomus fariai* (Lima). *Genetics* 29(1):75–82.

Ferguson KB, Kursch-Metz T, Verhulst EC, Pannebakker BA. 2020. Hybrid genome assembly and evidence-based annotation of the egg parasitoid and biological control agent *Trichogramma brassicae*. *G3 (Bethesda)* 10(10):3533–3540.

Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. 2020. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc Natl Acad Sci U S A.* 117(17):9451–9457.

Gluderer S, Oldham S, Rintelen F, Sulzer A, Schutt C, Wu XD, Raftery LA, Hafen E, Stocker H. 2008. Bunched, the Drosophila homolog of the mammalian tumor suppressor TSC-22, promotes cellular growth. *BMC Dev Biol*. 8:10.

Gokhman VE. 2009. Karyotypes of parasitic hymenoptera. Dordrecht: Springer.

Goulet H, Huber JT. 1993. Hymenoptera of the world: an identification guide to families. Ottawa (ON): Research Branch, Agriculture Canada.

Gremme G, Brendel V, Sparks ME, Kurtz S. 2005. Engineering a software tool for gene structure prediction in higher organisms. *Inform Software Tech*. 47(15):965–978.

Guichard A, Biehs B, Sturtevant MA, Wickline L, Chacko J, Howard K, Bier E. 1999. rhomboid and Star interact synergistically to promote EGFR MAPK signaling during *Drosophila* wing vein development. *Development* 126(12):2663–2676.

Gupta S, McGrath B, Cavener DR. 2010. PERK (EIF2AK3) regulates proinsulin trafficking and quality control in the secretory pathway. *Diabetes* 59(8):1937–1947.

Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. 2008. Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to assemble spliced alignments. *Genome Biol*. 9(1):R7.

Han Y, Wang B, Cho YS, Zhu J, Wu J, Chen Y, Jiang J. 2019. Phosphorylation of Ci/Gli by fused family kinases promotes hedgehog signaling. *Dev Cell*. 50(5):610–626.e614.

Hanken J, Wake DB. 1993. Miniaturization of body-size – organismal consequences and evolutionary significance. *Annu Rev Ecol Syst*. 24(1):501–519.

He JH. 2004. Hymenopteran insect fauna of Zhejiang. Beijing (China): Science Press.

He K, Lin K, Wang G, Li F. 2016. Genome sizes of nine insect species determined by flow cytometry and k-mer analysis. *Front Physiol*. 7:569.

Johnson NF. 1984. Systematics of Nearctic *Telenomus*: classification and revisions of the podisi and phymatae species groups (Hymenoptera: Scelionidae). *Bull Ohio Biol. Surv*. 6:1–113.

Johnson NF. 1992. Catalog of world species of Proctotrupoidea, exclusive of Platygastridae (Hymenoptera). *Mem Am Entomol Inst*. 51:1–825.

Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, et al. 2014. InterProScan 5: genomescale protein function classification. *Bioinformatics* 30(9):1236–1240.

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. 14(6):587–589.

Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. *Mol Biol Evol*. 30(4):772–780.

Kenis M, du Plessis H, Van den Berg J, Ba MN, Goergen G, Kwadjo KE, Baoua I, Tefera T, Buddie A, Cafa G, et al. 2019. *Telenomus remus*, a candidate parasitoid for the biological control of *Spodoptera frugiperda* in Africa, is already present on the continent. *Insects* 10(4):92.

Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 12(4):357–U121.

Kim M, Cha GH, Kim S, Lee JH, Park J, Koh H, Choi KY, Chung JK. 2004. MKP-3 has essential roles as a negative regulator of the Ras/mitogenactivated protein kinase pathway during *Drosophila* development. *Mol Cell Biol*. 24(2):573–583.

King JAJ, Straffon AFL, D'Abaco GM, Poon CLC, I STT, Smith CM, Buchert M, Corcoran NM, Hall NE, Callus BA, et al. 2005. Distinct requirements for the Sprouty domain for functional activity of Spred proteins. *Biochem J*. 388(Pt 2):445–454,.

Klopfenstein DV, Zhang LS, Pedersen BS, Ramirez F, Vesztrocy AW, Naldi A, Mungall CJ, Yunes JM, Botvinnik O, Weigel M, et al. 2018. GOATOOLS: a Python library for Gene Ontology analyses. *Sci Rep*. 8(1):10872.

Kogerman P, Grimm T, Kogerman L, Krause D, Unden AB, Sandstedt B, Toftgard R, Zaphiropoulos PG. 1999. Mammalian suppressor-offused modulates nuclear-cytoplasmic shuttling of Gli-1. *Nat Cell Biol*. 1(5):312–319.

Kuck U, Radchenko D, Teichert I. 2019. STRIPAK, a highly conserved signaling complex, controls multiple eukaryotic cellular and developmental processes and is linked with human diseases. *Biol Chem*. 400:1005–1022.

Kwan J, Sczaniecka A, Heidary Arash E, Nguyen L, Chen CC, Ratkovic S, Klezovitch O, Attisano L, McNeill H, Emili A, et al. 2016. DLG5 connects cell polarity and Hippo signaling protein networks by linking PAR-1 with MST1/2. *Genes Dev*. 30(24):2696–2709.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 9(4):357–359.

Liao YL, Yang B, Xu MF, Lin W, Wang DS, Chen KW, Chen HY. 2019. First report of *Telenomus remus* parasitizing *Spodoptera frugiperda* and its field parasitism in southern China. *J Hymenopt Res*. 73:95–102.

Lindsey ARI, Kelkar YD, Wu X, Sun D, Martinson EO, Yan Z, Rugman-Jones PF, Hughes DST, Murali SC, Qu J, et al. 2018. Comparative genomics of the miniature wasp and pest control agent *Trichogramma pretiosum*. *BMC Biol*. 16(1):54.

Liu H, Wu S, Li A, Ruan J. 2021. SMARTdenovo: a de novo assembler using long noisy reads. *Gigabyte* 1–9. doi: 10.46471/gigabyte.15.

Maloveryan A, Finta C, Osterlund T, Kogerman P. 2007. A possible role of mouse Fused (STK36) in Hedgehog signaling and Gli transcription factor regulation. *J Cell Commun Signal.* 1(3–4):165–173.

Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 32(1):268–274.

Pal C, Papp B, Lercher MJ. 2006. An integrated view of protein evolution. *Nat Rev Genet.* 7(5):337–348.

Paradis E. 2010. pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* 26(3):419–420.

Partha R, Chauhan BK, Ferreira Z, Robinson JD, Lathrop K, Nischal KK, Chikina M, Clark NL. 2017. Subterranean mammals show convergent regression in ocular genes and enhancers, along with adaptation to tunneling. *Elife* 6:e25884.

Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol.* 33(3):290–295.

Peters RS, Krogmann L, Mayer C, Donath A, Gunkel S, Meusemann K, Kozlov A, Podsiadlowski L, Petersen M, Lanfear R, et al. 2017. Evolutionary history of the Hymenoptera. *Curr Biol.* 27(7):1013–1018.

Peters RS, Niehuis O, Gunkel S, Blaser M, Mayer C, Podsiadlowski L, Kozlov A, Donath A, van Noort S, Liu S, et al. 2018. Transcriptome sequence-based phylogeny of chalcidoid wasps (Hymenoptera: Chalcidoidea) reveals a history of rapid radiations, convergence, and evolutionary success. *Mol Phylogenet Evol.* 120:286–296.

Polilov AA. 2015. Small is beautiful: features of the smallest insects and limits to miniaturization. *Annu Rev Entomol.* 60:103–121.

Raducu M, Fung E, Serres S, Infante P, Barberis A, Fischer R, Bristow C, Thezenas ML, Finta C, Christianson JC, et al. 2016. SCF (Fbxl17) ubiquitylation of Sufu regulates Hedgehog signaling and medulloblastoma development. *EMBO J.* 35(13):1400–1416.

Rey C, Gueguen L, Semon M, Boussau B. 2018. Accurate detection of convergent amino-acid evolution with PCOC. *Mol Biol Evol.* 35(9):2296–2306.

Sanderson MJ. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* 19(2):301–302.

Saucedo LJ, Edgar BA. 2007. Filling out the Hippo pathway. *Nat Rev Mol Cell Biol.* 8(8):613–621.

Servant N, Varoquaux N, Lajoie BR, Viara E, Chen CJ, Vert JP, Heard E, Dekker J, Barillot E. 2015. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* 16:259.

Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31(19):3210–3212.

Smith MD, Wertheim JO, Weaver S, Murrell B, Scheffler K, Kosakovsky Pond SL. 2015. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol Biol Evol.* 32(5):1342–1353.

Staley BK, Irvine KD. 2012. Hippo signaling in *Drosophila*: recent advances and insights. *Dev Dyn.* 241(1):3–15.

Stanke M, Steinkamp R, Waack S, Morgenstern B. 2004. AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Res.* 32(Web Server issue):W309–W312.

Tempel S. 2012. Using and understanding RepeatMasker. *Methods Mol Biol.* 859:29–51.

Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M. 2008. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Res.* 18(12):1979–1990.

Tourasse NJ, Li WH. 2000. Selective constraints, amino acid composition, and the rate of protein evolution. *Mol Biol Evol.* 17(4):656–664.

van der Woude E, Smid HM, Chittka L, Huigens ME. 2013. Breaking Haller's rule: brain-body size isometry in a minute parasitic wasp. *Brain Behav Evol.* 81(2):86–92.

Wakioka T, Sasaki A, Kato R, Shouda T, Matsumoto A, Miyoshi K, Tsuneoka M, Komiya S, Baron R, Yoshimura A. 2001. Spred is a Sprouty-related suppressor of Ras signalling. *Nature* 412(6847):647–651.

Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, et al. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9(11):e112963.

Wallberg A, Bunikis I, Pettersson OV, Mosbech MB, Childers AK, Evans JD, Mikheyev AS, Robertson HM, Robinson GE, Webster MT. 2019. A hybrid de novo genome assembly of the honeybee, *Apis mellifera*, with chromosome-length scaffolds. *BMC Genomics* 20(1):275.

Wertheim JO, Murrell B, Smith MD, Kosakovsky Pond SL, Scheffler K. 2015. RELAX: detecting relaxed selection in a phylogenetic framework. *Mol Biol Evol.* 32(3):820–832.

Xiao H, Ye X, Xu H, Mei Y, Yang Y, Chen X, Yang Y, Liu T, Yu Y, Yang W, et al. 2020. The genetic adaptations of fall armyworm *Spodoptera frugiperda* facilitated its rapid global dispersal and invasion. *Mol Ecol Resour.* 20(4):1050–1068.

Yang Y, Ye X, Dang C, Cao Y, Hong R, Sun YH, Xiao S, Mei Y, Xu L, Fang Q, et al. 2021. Genome of the pincer wasp *Gonatopus flavifemur* reveals unique venom evolution and a dual adaptation to parasitism and predation. *BMC Biol.* 19(1):145.

Ye X, Yan Z, Yang Y, Xiao S, Chen L, Wang J, Wang F, Xiong S, Mei Y, Wang F, et al. 2020. A chromosome-level genome assembly of the parasitoid wasp *Pteromalus puparum*. *Mol Ecol Resour.* 20(5):1384–1402.

Ye XH, van Achterberg C, Yue Q, Xu ZF. 2017. Review of the Chinese Leucospidae (Hymenoptera, Chalcidoidea). *Zookeys* (651):107–157.

Zhang JX, Lindsey ARI, Peters RS, Heraty JM, Hopper KR, Werren JH, Martinson EO, Woolley JB, Yoder MJ, Krogmann L. 2020. Conflicting signal in transcriptomic markers leads to a poorly resolved backbone phylogeny of chalcidoid wasps. *Syst Entomol.* 45(4):783–802.

Zhang PC, McGrath B, Li SA, Frank A, Zambito F, Reinert J, Gannon M, Ma K, McNaughton K, Cavener DR. 2002. The PERK eukaryotic initiation factor 2 alpha kinase is required for the development of the skeletal system, postnatal growth, and the function and viability of the pancreas. *Mol Cell Biol.* 22(11):3864–3874.

Zheng Y, Liu B, Wang L, Lei H, Pulgar Prieto KD, Pan D. 2017. Homeostatic control of Hpo/MST kinase activity through autophosphorylation-dependent recruitment of the STRIPAK PP2A phosphatase complex. *Cell Rep.* 21(12):3612–3623.

Zhou J, Bruns MA, Tiedje JM. 1996. DNA recovery from soils of diverse composition. *Appl Environ Microbiol.* 62(2):316–322.