Longevity and transposon defense, the case of termite reproductives

Daniel Elsner[a], Karen Meusemann[a], and Judith Korb[a,1]

[a]Evolutionary Biology & Ecology, University of Freiburg, Freiburg D-79104, Germany

Edited by Gene E. Robinson, University of Illinois at Urbana–Champaign, Urbana, IL, and approved March 30, 2018 (received for review March 7, 2018)

Social insects are promising new models in aging research. Within single colonies, longevity differences of several magnitudes exist that can be found elsewhere only between different species. Reproducing queens (and, in termites, also kings) can live for several decades, whereas sterile workers often have a lifespan of a few weeks only. We studied aging in the wild in a highly social insect, the termite *Macrotermes bellicosus*, which has one of the most pronounced longevity differences between reproductives and workers. We show that gene-expression patterns differed little between young and old reproductives, implying negligible aging. By contrast, old major workers had many genes up-regulated that are related to transposable elements (TEs), which can cause aging. Strikingly, genes from the PIWI-interacting RNA (piRNA) pathway, which are generally known to silence TEs in the germline of multicellular animals, were down-regulated only in old major workers but not in reproductives. Continued up-regulation of the piRNA defense commonly found in the germline of animals can explain the long life of termite reproductives, implying somatic cooption of germline defense during social evolution. This presents a striking germline/soma analogy as envisioned by the superorganism concept: the reproductives and workers of a colony reflect the germline and soma of multicellular animals, respectively. Our results provide support for the disposable soma theory of aging.

transposable elements | aging | social insects | termites | disposable soma

All animals age. However, some have a lifespan of a few weeks only (e.g., the fruit fly *Drosophila melanogaster*, the nematode *Caenorhabditis elegans*, the water flea *Daphnia longispina*), whereas others live for centuries (e.g., the clam *Arctica islandica*) (1, 2). Social insects like honey bees, ants, and termites have lifespan differences within single colonies that can differ by two orders of magnitude (3–5). Reproducing queens (and, in termites, also kings) can live for 20 y whereas the nonreproducing workers live only a few weeks to months. This makes social insects especially promising new models, as the within-species variation in rate of senescence only exists at the range of between-species variation outside social insects. Additionally, as social insect colonies are generally composed of families, all colony members share the same genetic background and differences in longevity are caused by differences in gene expression. Studying the molecular underpinning of aging in social insects is an emerging field with groundbreaking research in the honey bee and some ants (6–14). For termites, which evolved sociality independently from social Hymenoptera (wasps, bees, and ants), we are aware of only a single study (15).

From the nematode *C. elegans* to humans, aging in solitary organisms has been linked to an increased activity of transposable elements (TEs) (16–21). Active TEs can "jump" within the genome, thereby disturbing the regulation and expression of other genes, for instance, by transposing into another gene or regulatory region. The germline is protected against uncontrolled TE activity by the PIWI-interacting RNA (piRNA) pathway (22–24). Proteins from the PIWI group (Aubergine, Argonaute 3, Piwi), which are part of the Argonaute superfamily, interact with piRNAs and a core of other proteins to prevent activity of TEs (24).

To gain insights into the mechanisms underlying caste-specific aging in social insects, and especially in termites, we compared gene-expression profiles by using de novo-sequenced transcriptomes of heads from old and young individuals from different castes of the fungus-growing termite *Macrotermes bellicosus* (Smeathman, 1781; Fig. 1*A* and Dataset S1). This species has one of the highest longevity skews (80–120 fold) of all social insects, with queens and kings living as long as 20 y whereas workers have a lifespan of only a few months. Whereas workers are completely sterile, *Macrotermes* queens are specialized egg-layers that produce ca. 20,000 eggs per day (25). Given their long lifespan, they are probably the most fecund terrestrial animals. *M. bellicosus* has two worker castes, majors and minors, which differ in behavior, sex, and development. Minor workers are mainly involved in "indoor" tasks, such as brood care and constructing mounds; they are females and can develop into soldiers, thereby extending their lifespan (26, 27). Major workers have a higher random extrinsic mortality already early in life, as they generally do the more dangerous tasks like foraging for food; they are males and present developmental endpoints (26, 27).

To uncover molecular underpinnings of longevity differences, we analyzed transcriptomes of old and young queens, kings, and major and minor workers sampled from field colonies of a population that we have studied for more than 20 y. This gives us the unique opportunity not only to obtain and age long-lived reproductives from the wild but also to determine life expectancy and sample young and old individuals accordingly. Thus, it was possible to compare reproductives and nonreproductives from the wild that were at the maximum range of their natural lifespan (*Methods*). Although *Macrotermes* reproductives can live for as long as 20 y in the laboratory, in the study population, the median lifetime expectancy after reaching the epigeal mound phase is 6 y (Fig. S1). This is because of the common occurrence of army ants that kill complete

Significance

Social insects such as honey bees or termites are promising new models for aging research. In contrast to short-lived models like the fruit fly or mouse, the reproductives of an insect colony have exceptionally long lifespans. This offers important new avenues for gerontology, especially as mechanisms underlying aging are highly conserved among animals. We studied aging in a termite from the wild. Our results suggest that aging in this species, as in other animals, is related to the activity of transposable elements (TEs; also known as "jumping genes"). Yet reproductives seem to be protected by a process that normally silences TEs in the germline of animals. This suggests that natural selection used a mechanism from the germline to protect whole animals.

Author contributions: J.K. designed research; D.E., K.M., and J.K. performed research; D.E., K.M., and J.K. analyzed data; and D.E., K.M., and J.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The raw sequence data used in this publication have been deposited in the ArrayExpress database (https://www.ebi.ac.uk/arrayexpress/), accession no. E-MTAB-6767. Additional archives are deposited at the Dryad Digital Repository (dx.doi.org/10.5061/dryad.7n3df16).

See Commentary on page 5317.

[1]To whom correspondence should be addressed. Email: judith.korb@biologie.uni-freiburg.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1804046115/-/DCSupplemental.

Published online May 7, 2018.

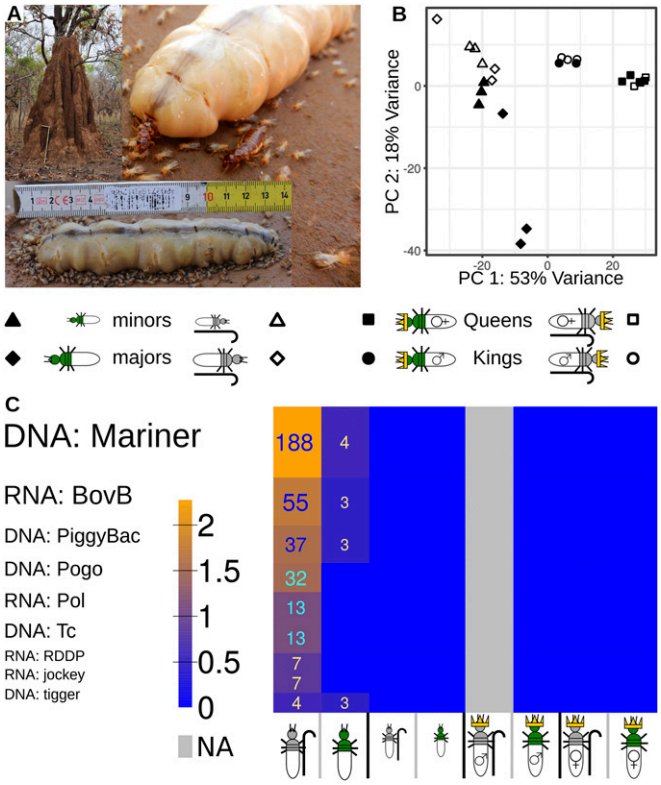

Fig. 1. *M. bellicosus* and its gene-expression pattern between castes and age classes. (*A*) An *M. bellicosus* mound with royal chamber harboring the queen (large whitish individual), the king (dark brown individual), and minor and major workers (smaller individuals). (*B*) Results of a PCA including all transcriptomes. Castes are separated along the first component (PC1), which explains 53% of the variance. Age classes of workers are separated along the second component (PC2), which accounts for 18% of the variance. By contrast, age classes of kings and queens cluster together. (*C*) Heat map of the DEGs for identified TEs. The height of the rows reflects the number of different TEs (shown as numbers). The color legend reflects the log10 of respective numbers (gray in the old kings, as they showed no significantly up-regulated genes). Termite symbols with a crown indicate reproductives: ♀, queens; ♂, kings; small and large termite symbols without a crown are minor and major workers, respectively. Old individuals are shown with a walking stick, in gray, with open symbols. Young individuals are shown in green with closed symbols.

colonies and function as the major cause of random extrinsic mortality after colonies are established (28). Accordingly, old reproductives in this study had an age of at least 6 y, and young reproductives of 1 y, after mound emergence. The age differences between young and old major and minor workers were in the range of months (*Methods*).

Results

Age-Related Changes in Gene Expression in Different Castes. By using a total of 24 transcriptomes, a principal component analysis (PCA) showed that gene expression clustered by caste and age (*Methods* and Fig. 1*B*). The first component separated castes, with kings being placed close to queens. Both worker castes formed a large cluster. The second component separated young and old workers, whereas all reproductives grouped together. No axis separated age classes of reproductives, implying that there are few age-related differences in gene expression in kings and queens.

This was supported when analyzing significantly differentially expressed genes (DEGs) at the single-gene level. We focused on the differential gene expression between young and old age classes within each caste (detailed in *Methods*). Of 13,959 expressed genes analyzed (Datasets S2, S3, and S11), only 26 genes were differentially expressed between old and young queens (18 up-regulated in old and 8 up-regulated in young). Kings had only two DEGs, which were up-regulated in young kings. This contrasts with the high number of DEGs in workers: 67 in minors (17 up-regulated in old and 50 up-regulated in young) and more than 5,000 in majors (3,125 up-regulated in old and 2,491 up-regulated in young; Dataset S3). These differences of aging in terms of DEGs between reproductives and workers are especially striking given that old and young workers differed in age by a few months only, whereas the age differences in reproductives was at least 5 y. Independent of age, queens had the most distinctive gene-expression pattern, with 717 genes being queen-specifically overexpressed compared with all other castes, whereas kings and minor and major workers had 50, 88, and 179, respectively (Fig. S2).

Age-Related Changes in Gene Expression and TE Activity. Analyzing the function of DEGs by using *D. melanogaster* homologs and orthologs in Database for Annotation, Visualization and Integrated Discovery (DAVID) did not reveal any significant enrichments in old or young individuals for queens, kings, and minor workers (complete functional annotation is provided in Dataset S4). However, young major workers were characterized by an overexpression of genes related to metabolism compared with old major workers (Fig. S3, Table S1, and Dataset S5).

In old major workers, a striking 356 of 2,387 annotated (with *Macrotermes natalensis*) up-regulated genes (14.9%) were TEs. This significantly differed compared with young major workers, for which only 13 of 2,096 annotated up-regulated genes (0.6%) were TEs (contingency analysis, $\chi^2 = 301.87$, $P < 0.001$). Note that the different number of genes up-regulated in old or young major workers compared with the DEGs mentioned earlier is because, here, we considered only annotated genes. Up-regulated TEs comprised DNA transposons as well as retrotransposons (Fig. 1*C* and Dataset S6). We identified active TEs of the types Line/L1, Mariner Mos 1 (>50%), BovB (15%), PiggyBac (10%), Pogo (8%), and others (Fig. 1*C*). Not all of the active TEs contained functional protein domains. Hence, some are likely nonautonomous TEs (29).

High TE expression is specific for major workers, and it differed significantly between major workers compared with the other castes (contingency analysis, majors vs. all other castes combined, $\chi^2 = 6.85$, $P = 0.008$). We found no TEs to be differentially expressed between age classes in queens and minor workers (Fig. 1*C*). Between old and young kings, only in the young kings were genes up-regulated, and none of them were TEs.

These differential-expression (DE) analyses were done with DESeq2. A corresponding pattern was obtained when using edgeR (*SI Methods*, *SI Results*, and Dataset S7) (30, 31). Furthermore, a quantitative real-time PCR (qRT-PCR) experiment for two TEs, Mariner Mos 1 and BovB, with additional samples from several colonies, confirmed the results based on transcriptome data (Datasets S8 and S11). Both TEs were significantly up-regulated in old major workers compared with young major workers based on samples from a total of seven colonies (Mariner Mos 1, old, 0.55 ± 0.06 SEM, $n = 6$; young, 0.37 ± 0.03 SEM, $n = 5$; t test for unequal variances, $t = -2.61$, $P = 0.028$; BovB, old, 0.62 ± 0.09 SEM, $n = 6$; young, 0.31 ± 0.01 SEM, $n = 5$; t test for equal variances, $t = -3.42$, $P = 0.018$; note, we did not find old and young major workers for each colony; one sample was analyzed per colony). Hence, our results show that old major workers were characterized by a high activity of TEs. As high TE activity has been linked to senescence in *D. melanogaster* (19–21, 32–34) and other animals (17, 18), the short intrinsic lifespan of major workers may be a result of their high TE activity.

This leads to the following questions. Which mechanisms might prevent TE activity in reproductives and minor workers, and which mechanisms are switched on in young but off in old major workers?

Expression of TE-Silencing Pathways in Different Castes. In *D. melanogaster*, TEs are silenced by piRNAs, which function in the germline (22–24, 35). To test whether similar TE-silencing mechanisms occur in *M. bellicosus*, and to what extent they are

differentially expressed, we identified orthologs of key genes known to be involved in piRNA-mediated TE silencing (22, 24, 35), checked protein domains (Dataset S11), and analyzed their expression pattern (Fig. 2).

piRNA pathway. piRNAs are a distinct class of small noncoding RNAs that form the piRNA-induced silencing complex (piRISC) in the germline of many animals (18–24, 35). The piRISC protects the integrity of the genome by silencing TEs. In contrast to siRNA or miRNA processing, it is Dicer-independent. The piRNA pathway can be separated into two parts: de novo biogenesis of piRNAs and the "ping-pong" amplification cycle. During biogenesis in somatic cells, primary piRNAs are produced from piRNA clusters (i.e., sites in the genome with high piRNA coverage) that serve as recognition and complementary binding sites for TEs in the ping-pong amplification cycle. The latter couples piRNA biogenesis to target silencing and amplifies primary piRNA in fly germ cells (Fig. 2*A*). There is considerable overlap of genes between both parts of the piRNA pathway (Table S2). Based on the recent review by Czech and Hannon (24), we unambiguously identified nine genes from the piRNA pathway (Table S3). This includes three piRNA genes (*ago3*, *aub1*, and *aub2*) that are characterized by a Piwi and PAZ domain, typical for *D. melanogaster* PIWIs (Fig. S4; *Methods*; rationale of naming these genes is described in *SI Results*).

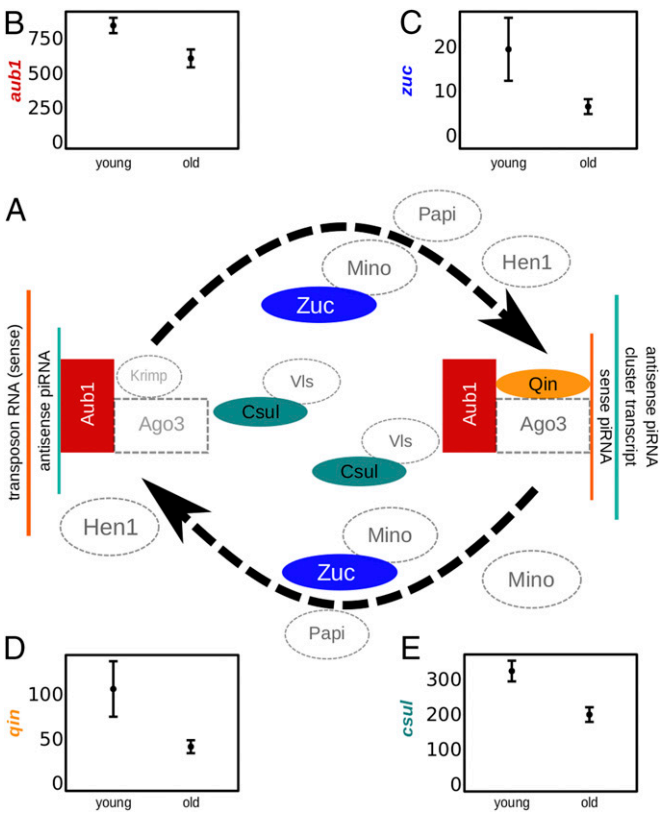

Fig. 2. DEGs that are involved in TE silencing according to results for *D. melanogaster*. (*A*) Shown is a simplified version of the ping-pong amplification cycle [after Czech and Hannon (24)] of the piRNA pathway, which is involved in silencing TEs in the germline of animals. Ago3 with its sense piRNA recognizes and cleaves piRNA cluster transcripts, which are loaded into Aub1, probably assisted by a Tudor protein and Qin, which prevents Aub1/Aub1 pairings. Maturation of piRNAs involves several proteins, including Mino, Zuc, Papi, and Hen1, and results in mature antisense piRNA–Aub1 complexes. Methylation by Capsuleen (Csul) is essential for stable Ago3 and Aub proteins (36). These complexes can detect and slice transposon RNAs. The 3′ cleaved product is then loaded in Ago3 and, after modification, the cycle can restart. The expression of the following genes involved in the ping-pong cycle was down-regulated in old major workers compared with young ones (mean ± SEM): (*B*) *aub1* (*C*) *zuc*, (*D*) *qin*, and (*E*) *csul*.

Ping-pong amplification cycle. Several genes involved in the ping-pong amplification cycle were significantly down-regulated in old compared with young major workers (Table S2). They include the key genes *aub1* as well as *capsuleen*, *qin*, and *zucchini* (Fig. 2). These are genes representing major processes of the ping-pong cycle in *D. melanogaster*, such as piRNA processing or linking Aub and Ago3 to transfer transcripts (Fig. 2). In the edgeR analysis, *aub1* was no longer significantly differentially expressed, but the remaining DEGs of the ping-pong cycle were significantly down-regulated in old major workers (Dataset S7). By contrast, none of the identified ping-pong cycle genes were differentially expressed between age classes in queens, kings, or minor workers (Table S2). This suggests that, specifically in old major workers, the ping-pong amplification cycle is less active, in line with an increased TE activity.

De novo biogenesis. In contrast to the differential expression of genes from the ping-pong cycle, none of the genes specifically involved in the somatic de novo biogenesis of piRNAs known from *D. melanogaster* was differentially expressed, such as *armi* or putative *gasz* (Table S2). *SoYb*, which was significantly down-regulated in the old major workers, is part of the de novo biogenesis but is germline-specific (37). One shared gene, *zucchini*, belongs to the ping-pong cycle and the de novo biogenesis. The expression of these genes was confirmed in DESeq2 and edgeR. These results imply that the somatic, de novo biogenesis part of the piRNA pathway is not associated with TE activity in old major workers, but that the germline-specific ping-pong amplification cycle is.

Discussion

Our results show that several genes from the ping-pong cycle that, in other organisms, protect the germline against TE activity (and hence aging) are specifically down-regulated in old compared with young major workers. Their expression pattern is diametrically opposite to the activity of TEs. This supports the hypothesis that a defense mechanism against TEs known from the germline of flies and other animals functions in termites to protect the soma of reproductives and minor workers. Note that we analyzed head transcriptomes (Dataset S1); this introduces selectivity (*SI Methods*) but now allows us to exclude any germline influence. We do not rule out that other processes such as oxidative stress may play a role in aging, but we focused on the striking pattern of differentially expressed TEs in this study (the complete list is provided in Datasets S2 and S3).

Comparing Major and Minor Workers. We found major differences between minor and major workers in age-related gene expression. Only major workers showed an increased TEs activity and a down-regulation of piRNA genes with age, whereas minor workers were more similar to the reproductives. Two factors, which may be interlinked, can account for this when applying evolutionary life history reasoning and regarding colony fitness as the relevant fitness level. The latter is reasonable because all workers in this species are sterile and cannot gain any direct fitness; colony fitness increases their indirect fitness. Hence, the workers of this species can be regarded as analogous to the soma of a multicellular organism, and we propose that aging theory (38, 39), such as the disposable soma theory (40, 41), applies equivalently. As a first factor, age-specific random extrinsic mortality differs between major and minor workers in *M. bellicosus*. Major workers do mainly outdoor tasks (i.e., foraging) for which random extrinsic mortality is extremely high in this species (42). By contrast, minor workers do indoor tasks when young and later do outdoor tasks (43). As extrinsic mortality is much higher outdoors than indoors, life history theory should analogously predict that selection against intrinsic aging (e.g., investment into physiological mechanisms that increase longevity such as TE defenses) should be lower in major compared with minor workers as the reproductive value of the major workers is lower. The reproductive value is measured/realized in these sterile castes via colony fitness. Second, major workers are developmental endpoints whereas minor workers can develop into soldiers. This additionally increases the reproductive

value of minors compared with majors and hence selects more strongly against (intrinsic) aging in minors. Both factors give minor workers a higher reproductive value than major workers. This may explain the benefit of a higher continued investment in TE defense in minor workers. Life history models specifically tailored to social insects are required to test this explanation.

TE Activity, Aging, and the Germline. There is substantial evidence, from *C. elegans* to humans, that senescence correlates with increasing TE activity (16–24, 44), and several studies suggest a causal link for *D. melanogaster* (32–34, 21). Uncontrolled TE activity is a threat to genome integrity as well as proper functioning of gene expression. Silencing TEs is especially important in the germline, as it "needs" more protection than disposable soma cells to sustain itself through generations. Accordingly, the piRNA pathway is mainly active in the germline, although recent evidence suggests that it can also be active in the brain and maybe some other tissues (44, 45). We can only speculate why this defense mechanism is largely restricted to the germline. Most likely, it is (energetically) costly, and a disposable soma is not worthwhile to be protected under all circumstances (40). Protection against TEs may present some of the germline-maintenance costs, which should receive more attention in aging research (41).

The Germline Analogy in Social Insects. In highly social insects with sterile workers, the queens and kings can be regarded as equivalent to the germline of a colony, whereas the workers represent the disposable soma. Similar to organs in a multicellular organism, virtually all reproduction is channeled through long-lived queens (and, in termites, also kings), whereas the workers specialize on tasks like food provisioning or health care. This organism-like analogy is reflected in the superorganism concept. Our results provide support for this analogy and the concept of the disposable soma. Causal evidence has to be shown in further studies. The reproductives, and to some extent the "more valuable" minor workers, seem to be protected against aging by a mechanism that is mainly active in the germline of solitary animals.

Methods

Sampling, Age Determination, Laboratory Work, and Sequencing. The *M. bellicosus* samples were collected in Kakpin, next to the Comoé National Park in Côte d'Ivoire (coordinates 8°39′N 3°46′W), where we have been studying this species for more than 20 y. We collected individuals of four castes (two age classes each): major and minor workers, queens and kings. The castes were identified by using the description in ref. 46. Of each of the worker castes, six individuals (three old and three young) were sampled. Sample size was constrained by the availability of colonies of known age and a low number of major workers of clearly distinct age (as detailed later). Of the sexual castes, five kings (two old, three young) and seven queens (three old, four young) were collected (Datasets S1 and S8).

We compared queens and kings from colonies that just had started to build a termite mound the previous year (i.e., young) with those that had mounds continuously inhabited for at least 6 y (i.e., old). Queens and kings are known to live as long as 20 y in the laboratory. The median longevity after emergence of an epigeal mound in our study site was estimated with Kaplan–Meier survival analysis for 36 mounds (Fig. S1) with the use SPSS 23 software (47).

For the workers, all individuals were collected in and around the queen chamber to control for a confounding effect of task on gene expression. Age was determined by checking mandible wear, a marker for age in termite workers (48, 49). Age differences between young and old reproductives were at least 5 y, whereas those for workers were only a few months.

Transcriptomes were generated from termite heads to avoid bacterial contamination from the gut. Additionally, by limiting ourselves to a single body part, we tried to avoid confounding factors/dilution of the signal. Therefore, the head was separated from all individuals. Altogether, 24 samples (one individual per sample; Dataset S1) were stored in Eppendorf tubes with RNAlater before the samples were transported to the laboratory for RNA extraction. Heads of all individuals were flushed with RNAlater (Qiagen) to ensure complete permeation of the head tissue. The samples were transported on ice to the field station and kept overnight at 4 °C to ensure full permeation. Then samples were kept frozen in a −20 °C refrigerator for no more than 2 wk. During transport, the samples were kept in an insulated container, cooled on ice, until arrival in the laboratory in Freiburg, Germany. There, samples were kept in a −80 °C freezer until further processing. Eleven additional samples were treated in the same manner for the qRT-PCR analysis; eight of these were collected in the Pendjari National Park (coordinates 11°19′ N 1°35′E; Dataset S8).

Total RNA was extracted from the head of single individuals according to a modified version of the peqGOLD TriFast protocol. Modifications were an incubation time of 10 min instead of 5 min at room temperature after adding 900 μL peqGOLD TriFast and inverting the tube. After adding chloroform and after incubation, centrifugation was modified to 10 min at 4 °C, 12,388 × *g*. After transfer of the aqueous phase into a new Eppendorf tube, we added another centrifugation step for 5 min at 4 °C (12,388 × *g*). The supernatant was then transferred into a new Eppendorf tube and 2 μL nuclease-free glycogen (5 mg/mL) was added per sample. After adding isopropanol, the mixture was incubated for 10 min at room temperature and not on ice. After washing the RNA pellet with 75% ethanol and vortexing, the sample was centrifuged for 5 min at 4 °C, 7,674 × *g*. This step was repeated three times.

After dissolving the pellet in nuclease-free water, samples were kept for at least 1 h at 4 °C. DNA digestion was performed by using DNase I Amplification Grade Kit (Sigma-Aldrich). The concentration and quality of the isolated RNA was checked with an Agilent Bioanalyzer (Agilent RNA 6000 Nano Kit) and sent to BGI in Hong Kong on dry ice.

Library preparation was done by BGI. Libraries for 22 samples showing slight degradation of RNA were constructed with the Ribo-Zero rRNA Removal Kit (human/mouse/rat; Illumina). Additionally, for 5 of these 22 samples, RnaseH treatment was performed. For two samples, the TruSeq RNA Kit was applied following the BGI sample preparation protocol. Details for each sample are given in Dataset S1.

Amplified libraries were sequenced on an Illumina HiSeq4000 platform with strategies of 100-bp paired-end reads, generating ∼4 Gb of raw data for each sample. After sequencing, index sequences from the machine reads were demultiplexed (sorted and removed) by a proprietary in-house tool at BGI.

Preprocessing of RNAseq Raw Reads. Raw sequence reads from all samples provided by BGI were used for all further steps. After quality control with FastQC v. 0.11.4 (50), raw sequence reads were checked for a minimum length of 70 bp, and adapters (including two BGI in-house adapters) were trimmed using Trimmomatic v. 0.33 (51). Trimming and other subsequent steps were facilitated by GNU parallel v. 20141022 (52).

Gene Expression and Mapping Against the Genome of *M. natalensis*. Because there is no sequenced *M. bellicosus* genome, we used the genome of *M. natalensis* (Haviland, 1898) v. 1.0 as a mapping backbone in our study (53). This species is closely related to our study species *M. bellicosus*. To test the suitability of *M. natalensis* as backbone (54), we downloaded mitochondrial cytochrome oxidase I (COI) sequences and sequence fragments for both *Macrotermes* species from GenBank (two sequence fragments of *M. natalensis* and six sequence fragments of *M. bellicosus*; Dataset S11). After aligning the sequences with the program Seaview (55) with the implemented clustal-omega (-clustalo) algorithm, we calculated with distmat (EMBOSS package v. 6.6.0.0) the distances between all COI sequences (all base positions considered, Kimura correction method) (56). The short evolutionary distances (<11) between COI sequences of *M. natalensis* and *M. bellicosus* suggest a short divergence time. This is in line with Brandl et al. (57), who determined the divergence time between the major clades of *Macrotermes* to be ∼8 My. Hence, *M. natalensis* can be used as a mapping backbone.

Trimmed and cleaned raw reads were mapped against the *M. natalensis* genome by using the program Tophat2 v. 2.1.0 with default settings (58). We created bowtie2-indices (59) based on the genome assembly (v. 1.0) and the corresponding GFF file as reference on which Tophat2 relies. The resulting BAM files were sorted by gene name by using SAMtools (v. 0.1.19–96b5f2294a) (60). HTSeq v. 0.6.0 was then used to count the mapped raw reads per gene (stranded = no, type = gene, mode = union, order = name) for each sample (61). A plain text table summarizing all samples was compiled (Dataset S11). Based on all valid counts identified by HTSeq, a background file with all expressed and mapped genes in *M. bellicosus* was built for later enrichment and pathway analyses.

Identification and Annotation of Clusters of Orthologous Groups. We inferred a set of clusters of orthologous sequence groups (COGs) from the official gene sets, i.e., all protein-coding sequences, at amino-acid level of the fruit fly *D. melanogaster* (v. r6.11) and the termite *M. natalensis* (v. 1.2; Dataset S11). For both species, we kept only the longest isoform per gene and removed all sequences with internal stop codons and/or selenocysteines (U) by custom-made perl scripts. The latter may cause problems in downstream analyses. For inferring COGs, we used OrthoFinder v. 0.4.0, a graph-based approach that relies on an all-vs.-all BLAST and MCL clustering and corrects for gene length bias (62). We considered only 1:1 orthologs.

All transcripts of the *M. natalensis* genome assembly that had a positive match with the *M. bellicosus* raw reads after HTSeq were matched to all pathways, Gene Ontology (GO) terms, and protein domains in the Interpro database (63), including the databases Pfam, Gene3D, and PANTHER, with InterProScan (5.16–55.0, December 2015) (64) using default settings.

Gene-Expression Analysis. DESeq2 v. 1.10.1 (30) was used in R v. 3.3.1 (65) to analyze differential gene expression between young and old individuals separately for each caste and between castes regardless of age. In addition, the obtained normalized read counts (Dataset S11) were used internally in DESeq2 to determine differential expression of all genes.

The count data were variance-transformed to perform a PCA and create a distance-clustering heat map. This allowed us to visualize the data and compare the effect of age and caste on gene expression. To test for pairwise expression differences, we used the generalized linear model with negative binominal distribution as implemented in DESeq2. We compared young and old age classes within each caste by combining caste and age factors into a group factor, as described in the DESeq2 vignette. Additionally, we also compared differential gene expression between castes (Dataset S11). False discovery rate (FDR)-adjusted *P* values were used to correct for multiple testing (66). Venn diagrams were produced to visualize the number of shared and unique genes between castes by using the online tool Venny 2 (67). We used the adjusted *P* values to determine DEGs of the piRNA pathway (Table S2 and Dataset S9). To corroborate our results, we additionally applied the algorithm implemented in edgeR to our data (31) (*SI Methods*, *SI Results*, Fig. S5, and Dataset S7).

Functional Annotation and Enrichment. A GO enrichment analysis (GO category, Biological Process) was done with the DAVID 6.8 Web tool (68). We supplemented our 1:1 ortholog set with homologs obtained by a BLASTP search between *M. natalensis* and *D. melanogaster* (BLAST v. 2.2.31+; National Center for Biotechnology Information) (69) with default settings and an e-value cutoff of $1e^{-5}$ of *M. natalensis* vs. *D. melanogaster*. We generated a tag cloud for significant results ($P < 0.05$ after FDR correction). Tag clouds were created in R 3.3.1 using the package tagcloud v. 0.6 (https://cran.r-project.org/package=tagcloud).

Transposable Elements. We chose the annotation of the *M. natalensis* genome provided by Haofu Hu (Dataset S2) to identify TEs (51). Additionally, we checked the Pfam database (70) (Pfam A, release 30), the Pfam annotation from the InterPro (61), and the Dfam database (71) for the presence of TEs in our DEGs (Dataset S11).

We compared the number of differentially expressed TEs between young and old individuals within each caste by using contingency analyses. This was possible only for major workers, as differentially expressed TEs were missing between age classes in the other castes. Additionally, we tested whether the frequency of differentially expressed TEs differed between major workers and the other castes by using contingency analyses. All tests were two-tailed, and all analyses were performed with SPSS 23 (47).

RNAi Pathway Genes Involved in TE Silencing. We identified genes related to the piRNA pathway and their relatives (Table S3 and Dataset S10): we obtained gene symbols and FlyBase IDs of *D. melanogaster* (e.g., "AGO1," "FBgn0262739"; Dataset S10) from FlyBase (72). We used the FlyBase IDs to search in OrthoDB v. 9.1 (73) for COGs of the following species: *D. melanogaster* (DMEL), *Apis mellifera* (AMEL), *Tribolium castaneum* (TCAS), *Zootermopsis nevadensis* (ZNEV), and *Blattella germanica* (BGER). We downloaded amino-acid sequences for each COG of the aforementioned species. We aligned the sequences separately for each COG with MAFFT v7.294b, choosing the G-INS-i vsm algorithm to avoid overalignment of sequences [(option allowshift–unalignlevel-leavegappyregion (74)]. We used hmmbuild implemented in HMMER v3.1b2 to construct a hidden Markov model (HMM) from each multiple sequence alignment (75). The HMM was used to search with hmmsearch (HMMER) against the official gene set v. 1.2 of *M. natalensis* to identify candidate sequences for each COG. We searched all candidate sequences with hmmscan (HMMER) against the Pfam database (Pfam A, release 30; inclusion threshold $1e^{-2}$). The candidate sequences were searched reciprocally against the official gene set of *D. melanogaster* (v. r6.11) with BLASTP and an e-value cutoff of $1e^{-10}$. In cases for which the original sequence of *D. melanogaster* was not identical with the one after the reciprocal BLASTP search, we considered results as ambiguous.

For selected genes related to the piRNA pathway (Dataset S10), we inferred phylogenetic trees of gene groups. We included sequences of all species mentioned earlier (DMEL, AMEL, TCAS, BGER, ZNEV, *Macrotermes*) plus amino acid sequences from the recently sequenced genome of the termite *Cryptotermes secundus* (data provided by J.K.). In particular, we created sets of sequences that we previously had assigned to more than one gene with the HMM search and reciprocal BLAST search against *D. melanogaster*: the AGO1-3/Piwi/Aub set (AGO1, AGO2, AGO3, Piwi/Aubergine; Dataset S10). For the set of sequences, the amino acid selenocysteine (U) was replaced by "X." We aligned each set of sequences with the G-INS-i vsm algorithm (option allowshift–unalignlevel-leavegappyregion) and subsequently checked each multiple sequence alignment for ambiguously aligned sequence regions with Aliscore v. 2 (76, 77), allowing the maximal number of pairwise comparisons, the e-option for gappy alignments, and otherwise default settings. We removed all positions suggested by Aliscore with Alicut v. 2.3 (https://www.zfmk.de/en/research/research-centres-and-groups/utilities; Dataset S11). We conducted 10 maximum-likelihood tree searches per sequence set with IQ-Tree v. 1.5.3 (78). We used completely random start trees and the automated estimation of the best model considering implemented nuclear models (with +G, +I, or +G+I, four rate categories, and the median approximation for +G site rates) plus the free rate model LG4X (79) with four rate categories. Statistical support was inferred from 1,000 nonparametric bootstrap replicates. We plotted bootstrap support on the maximum-likelihood tree with the best log-likelihood value. We visualized the unrooted trees with bootstrap support by using Seaview v. 4.5.4 (55) and graphically processed the trees with Inkscape v0.91 (www.inkscape.org/).

The sequence of Mnat_13309 in the genome of *M. natalensis* was short (*SI Results* and Dataset S11). To obtain further information, we assembled the original raw reads from *M. bellicosus* into contigs with Trinity v2.0.6 (80) and confirmed the identity of the best matching contig to Mnat_13309 as *ago3* using a reciprocal BLAST against *D. melanogaster* and a domain search against Pfam release 30. The contig was most similar to *D. melanogaster* Ago3 and contained Piwi and PAZ domains.

ACKNOWLEDGMENTS. We thank K. Eduard Linsenmair, N'Golo A. Kone, and the Comoé Research Station for logistic support; David Kouassi Kouame for assistance in the field; Florentine Schaub for assistance in the laboratory; Haofu Hu, Ecology and Evolution, Department of Biology, University of Copenhagen, Copenhagen, Denmark for providing the gene annotation of *M. natalensis*; the Csec-consortium for providing access to the *C. secundus* genome; the i5K consortium, Coby Schal, Xavier Belles, and Stephen Richards for providing access to the *B. germanica* genome; Ondrej Hlinka for help with the Cluster Commonwealth Scientific and Industrial Research Organization in Australia; Robert Waterhouse for information on data preparation; Panos Provataris for bioinformatic scripts; Manuel Monroy Kuhn for bioinformatic discussions; Volker Nehring for English editing; Fabian Staubach for advice regarding the qRT-PCR controls; and three anonymous reviewers for helpful comments. This research was supported by Deutsche Forschungsgemeinschaft Grant KO1895/19-1 (to J.K.) within the Research Unit FOR2281 and by the state of Baden-Württemberg through Baden-Württemberg High Performance Computing. The Office Ivoirien des Parcs et Réserves provided sampling and export permits. The study was conducted according to the Nagoya protocol.

1. Jones OR, et al. (2014) Diversity of ageing across the tree of life. *Nature* 505:169–173.
2. Butler PG, Wanamaker AD, Jr, Scourse JD, Richardson CA, Reynolds DJ (2013) Variability of marine climate on the North Icelandic Shelf in a 1357-year proxy archive based on growth increments in the bivalve *Arctica islandica*. *Palaeogeogr Palaeoclimatol Palaeocol* 373: 141–151.
3. Keller L, Genoud M (1997) Extraordinary lifespans in ants: A test of evolutionary theories of ageing. *Nature* 389:958–960.
4. Keller L (1998) Queen lifespan and colony characteristics in ants and termites. *Insectes Soc* 45:235–246.
5. Korb J, Thorne B (2017) Sociality in termites. *Comparative Social Evolution*, eds Rubenstein DR, Abbot P (Cambridge Univ Press, Cambridge, UK), pp 84–123.
6. Rueppell O, Christine S, Mulcrone C, Groves L (2007) Aging without functional senescence in honey bee workers. *Curr Biol* 17:R274–R275.
7. Seehuus S-C, Norberg K, Gimsa U, Krekling T, Amdam GV (2006) Reproductive protein protects functionally sterile honey bee workers from oxidative stress. *Proc Natl Acad Sci USA* 103:962–967.
8. Jemielity S, et al. (2007) Short telomeres in short-lived males: What are the molecular and evolutionary causes? *Aging Cell* 6:225–233.
9. Rascón B, Hubbard BP, Sinclair DA, Amdam GV (2012) The lifespan extension effects of resveratrol are conserved in the honey bee and may be driven by a mechanism related to caloric restriction. *Aging (Albany NY)* 4:499–508.
10. Münch D, Kreibich CD, Amdam GV (2013) Aging and its modulation in a long-lived worker caste of the honey bee. *J Exp Biol* 216:1638–1649.
11. Seehuus S-C, Taylor S, Petersen K, Aamodt RM (2013) Somatic maintenance resources in the honeybee worker fat body are distributed to withstand the most life-threatening challenges at each life stage. *PLoS One* 8:e69870.

12. Lucas ER, Keller L (2014) Ageing and somatic maintenance in social insects. *Curr Opin Insect Sci* 5:31–36.
13. von Wyschetzki K, Rueppell O, Oettler J, Heinze J (2015) Transcriptomic signatures mirror the lack of the fecundity/longevity trade-off in ant queens. *Mol Biol Evol* 32: 3173–3185.
14. Lucas ER, Privman E, Keller L (2016) Higher expression of somatic repair genes in long-lived ant queens than workers. *Aging (Albany NY)* 8:1940–1951.
15. Tasaki E, Kobayashi K, Matsuura K, Iuchi Y (2017) An efficient antioxidant system in a long-lived termite queen. *PLoS One* 12:e0167412.
16. Maxwell PH, Burhans WC, Curcio MJ (2011) Retrotransposition is associated with genome instability during chronological aging. *Proc Natl Acad Sci USA* 108: 20376–20381.
17. Dennis S, Sheth U, Feldman JL, English KA, Priess JR (2012) *C. elegans* germ cells show temperature and age-dependent expression of Cer1, a Gypsy/Ty3-related retrotransposon. *PLoS Pathog* 8:e1002591.
18. De Cecco M, et al. (2013) Transposable elements become active and mobile in the genomes of aging mammalian somatic tissues. *Aging (Albany NY)* 5:867–883.
19. Li X, et al. (2013) A resurrected mammalian hAT transposable element and a closely related insect element are highly active in human cell culture. *Proc Natl Acad Sci USA* 110:E478–E487.
20. Chen H, Zheng X, Xiao D, Zheng Y (2016) Age-associated de-repression of retrotransposons in the *Drosophila* fat body, its potential cause and consequence. *Aging Cell* 15:542–552.
21. Wood JG, et al. (2016) Chromatin-modifying genetic interventions suppress age-associated transposable element activation and extend life span in *Drosophila*. *Proc Natl Acad Sci USA* 113:11277–11282.
22. Siomi MC, Sato K, Pezic D, Aravin AA (2011) PIWI-interacting small RNAs: The vanguard of genome defence. *Nat Rev Mol Cell Biol* 12:246–258.
23. Rozhkov NV, Hammell M, Hannon GJ (2013) Multiple roles for Piwi in silencing *Drosophila* transposons. *Genes Dev* 27:400–412.
24. Czech B, Hannon GJ (2016) One loop to rule them all: The ping-pong cycle and piRNA-guided silencing. *Trends Biochem Sci* 41:324–337.
25. Kaib M, Hacker M, Brandl R (2001) Egg-laying in monogynous and polygynous colonies of the termite *Macrotermes michaelseni* (Isoptera, Macrotermitidae). *Insectes Soc* 48:231–237.
26. Noirot C (1969) Formation of castes in the higher termites. *Biology of Termites*, eds Krishna K, Weesner FM (Academic, New York), Vol 1, pp 311–350.
27. Gerber C, Badertscher S, Leuthold R (1988) Polyethism in *Macrotermes bellicosus* (Isoptera). *Insectes Soc* 35:226–240.
28. Korb J, Linsenmair KE (2001) The causes of spatial patterning of mounds of a fungus-cultivating termite: Results from nearest-neighbour analysis and ecological studies. *Oecologia* 127:324–333.
29. Wicker T, et al. (2007) A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* 8:973–982.
30. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550.
31. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26: 139–140.
32. Driver CJ, McKechnie SW (1992) Transposable elements as a factor in the aging of Drosophila melanogaster. *Ann N Y Acad Sci* 673:83–91.
33. Woodruff RC (1992) Transposable DNA elements and life history traits. I. Transposition of P DNA elements in somatic cells reduces the lifespan of Drosophila melanogaster. *Genetica* 86:143–154.
34. Nikitin AG, Woodruff RC (1995) Somatic movement of the mariner transposable element and lifespan of *Drosophila* species. *Mutat Res* 338:43–49.
35. Yang F, Xi R (2017) Silencing transposable elements in the *Drosophila* germline. *Cell Mol Life Sci* 74:435–448.
36. Kirino Y, et al. (2009) Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability. *Nat Cell Biol* 11:652–658.
37. Handler D, et al. (2011) A systematic analysis of *Drosophila* TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors. *EMBO J* 30:3977–3993.
38. Hamilton WD (1966) The moulding of senescence by natural selection. *J Theor Biol* 12: 12–45.
39. Charlesworth B (1980) *Evolution in Age-Structured Populations* (Cambridge Univ Press, Cambridge, UK).
40. Kirkwood TBL (1977) Evolution of ageing. *Nature* 270:301–304.
41. Maklakov AA, Immler S (2016) The expensive germline and the evolution of ageing. *Curr Biol* 26:R577–R586.
42. Korb J, Linsenmair KE (2002) Evaluation of predation risk in the collectively foraging termite *Macrotermes bellicosus*. *Insectes Soc* 49:264–269.
43. Hinze B, Leuthold RH (1999) Age related polyethism and activity rhythms in the nest of the termite *Macrotermes bellicosus* (Isoptera, Termitidae). *Insectes Soc* 46:392–397.
44. Perrat PN, et al. (2013) Transposition-driven genomic heterogeneity in the *Drosophila* brain. *Science* 340:91–95.
45. Zuo L, Wang Z, Tan Y, Chen X, Luo X (2016) piRNAs and their functions in the brain. *Int J Hum Genet* 16:53–60.
46. Ruelle JE (1970) Revision of the termites of the genus *Macrotermes* from the Ethiopian region (Isoptera: Termitidae). *Bull Br Mus (Nat Hist)* 24:365–444.
47. IBM Corp (2013) IBM SPSS Statistics for Windows. Version 23 (IBM, Armonk, NY).
48. Skaife SH (1955) *Dwellers in Darkness: An Introduction to the Study of Termites* (Longmans Green, London).
49. Buchli HHR (1958) L'origine des castes et les potentialités ontogéniques des Termites Européens du genre *Reticulitermes* Holmgren, ed Masson et Cie. PhD dissertation (Libraires de l' Académie de Médecine, Saint-Germain, Paris).
50. Anders S (2010) FastQC: A quality control tool for high throughput sequence data, version 0.11.4 (Babraham Institute, Cambridge, UK) Available at https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
51. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120.
52. Tange O (2011) GNU parallel–The command-line power tool. *Login USENIX Mag* 36: 42–47.
53. Poulsen M, et al. (2014) Complementary symbiont contributions to plant decomposition in a fungus-farming termite. *Proc Natl Acad Sci USA* 111:14500–14505.
54. Hornett EA, Wheat CW (2012) Quantitative RNA-seq analysis in non-model species: Assessing transcriptome assemblies as a scaffold and the utility of evolutionary divergent genomic reference species. *BMC Genomics* 13:361.
55. Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* 27:221–224.
56. Rice P, Longden I, Bleasby A (2000) EMBOSS: The European molecular biology open software suite. *Trends Genet* 16:276–277.
57. Brandl R, et al. (2007) Divergence times in the termite genus *Macrotermes* (Isoptera: Termitidae). *Mol Phylogenet Evol* 45:239–250.
58. Kim D, et al. (2013) TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14:R36.
59. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359.
60. Li H, et al.; 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079.
61. Anders S, Pyl PT, Huber W (2015) HTSeq–A Python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166–169.
62. Emms DM, Kelly S (2015) OrthoFinder: Solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* 16: 157.
63. Mitchell A, et al. (2015) The InterPro protein families database: The classification resource after 15 years. *Nucleic Acids Res* 43:D213–D221.
64. Jones P, et al. (2014) InterProScan 5: Genome-scale protein function classification. *Bioinformatics* 30:1236–1240.
65. R Core Team (2016) R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna). Available at https://www.R-project.org/.
66. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc B* 57:289–300.
67. Oliveros JC (2007) Venny. An interactive tool for comparing lists with Venn's diagrams (Spanish National Biotechnology Centre–CSIC, Madrid). Available at bioinfogp.cnb.csic.es/tools/venny/index.html.
68. Huang W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4:44–57.
69. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410.
70. Finn RD, et al. (2016) The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res* 44:D279–D285.
71. Hubley R, et al. (2016) The Dfam database of repetitive DNA families. *Nucleic Acids Res* 44:D81–D89.
72. Gramates LS, et al.; The Fly Base Consortium (2017) FlyBase at 25: Looking to the future. *Nucleic Acids Res* 45:D663–D671.
73. Zdobnov EM, et al. (2017) OrthoDB v9.1: Cataloging evolutionary and functional annotations for animal, fungal, plant, archaeal, bacterial and viral orthologs. *Nucleic Acids Res* 45:D744–D749.
74. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol Biol Evol* 30:772–780.
75. Eddy SR (2011) Accelerated profile HMM searches. *PLoS Comput Biol* 7:e1002195.
76. Misof B, Misof K (2009) A Monte Carlo approach successfully identifies randomness in multiple sequence alignments: A more objective means of data exclusion. *Syst Biol* 58: 21–34.
77. Kück P, et al. (2010) Parametric and non-parametric masking of randomness in sequence alignments can be improved and leads to better resolved trees. *Front Zool* 7: 10.
78. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268–274.
79. Le SQ, Dang CC, Gascuel O (2012) Modeling protein evolution with several amino acid replacement matrices depending on site rates. *Mol Biol Evol* 29:2921–2936.
80. Grabherr MG, et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29:644–652.
81. Untergasser A, et al. (2012) Primer3–New capabilities and interfaces. *Nucleic Acids Res* 40:e115.
82. Weil T, Rehli M, Korb J (2007) Molecular basis for the reproductive division of labour in a lower termite. *BMC Genomics* 8:198.
83. Iwashita S, et al. (2006) A tandem gene duplication followed by recruitment of a retrotransposon created the paralogous *bucentaur* gene (bcntp97) in the ancestral ruminant. *Mol Biol Evol* 23:798–806.
84. Dowling D, et al. (2016) Phylogenetic origin and diversification of RNAi pathway genes in insects. *Genome Biol Evol* 8:3784–3793.