Suppl. Table S1. Sequencing results, number of reads kept after trimming and number of reads classified as possible bacteria contaminations.

Library	**Total read bases (Gb)**	**Number of reads**	**After trimming**	**Contaminant reads**
PE - 350bp	33.99	2 x 134,896,091	2 x 131,122,011	201,770
PE - 550bp	35.97	2 x 142,742,469	2 x 138,743,584	

Suppl. Table S2. Genome size and heterozygosity relative to different *k-mer* sizes.

***k-mer* size (bp)**	**Genome size (pb)**	**Heterozygosity (%)**
21	901,419,411	0.106906
24	909,233,236	0.0964907
27	915,858,453	0.0879122
30	921,364,602	0.0894091

Suppl. Table S3: Details about the number of sequences used for P450 gene family analyses. The final number of sequences refers to the total number of sequences used after filtering by Trimal, RogueNaRok and Treeshrink for the final phylogenetic inference.

Species	**Starting number of sequences**	**Final number of sequences**
Reticulitermes lucifugus	95	61
Reticulitermes speratus	106	71
Zootermopsis nevandesis	78	64
Macrotermes natalensis	91	71
Coptotermes formosanus	80	71
Cryptotermes secundus	99	92
Acromyrmex echinatior	66	52
Blattella germanica	145	102
Camponotus floridanus	113	52
Drosophila melanogaster	89	89
Locusta migratoria	132	94
Nasonia vitripennis	95	90
Apis mellifera	46	46
TOTAL	**1,235**	**955**

Suppl. Table S4: Details about the number of sequences used for the OR gene family analyses. The final number of sequences refers to the total number of sequences used after filtering by Trimal, RogueNaRok and Treeshrink for the final phylogenetic inference.

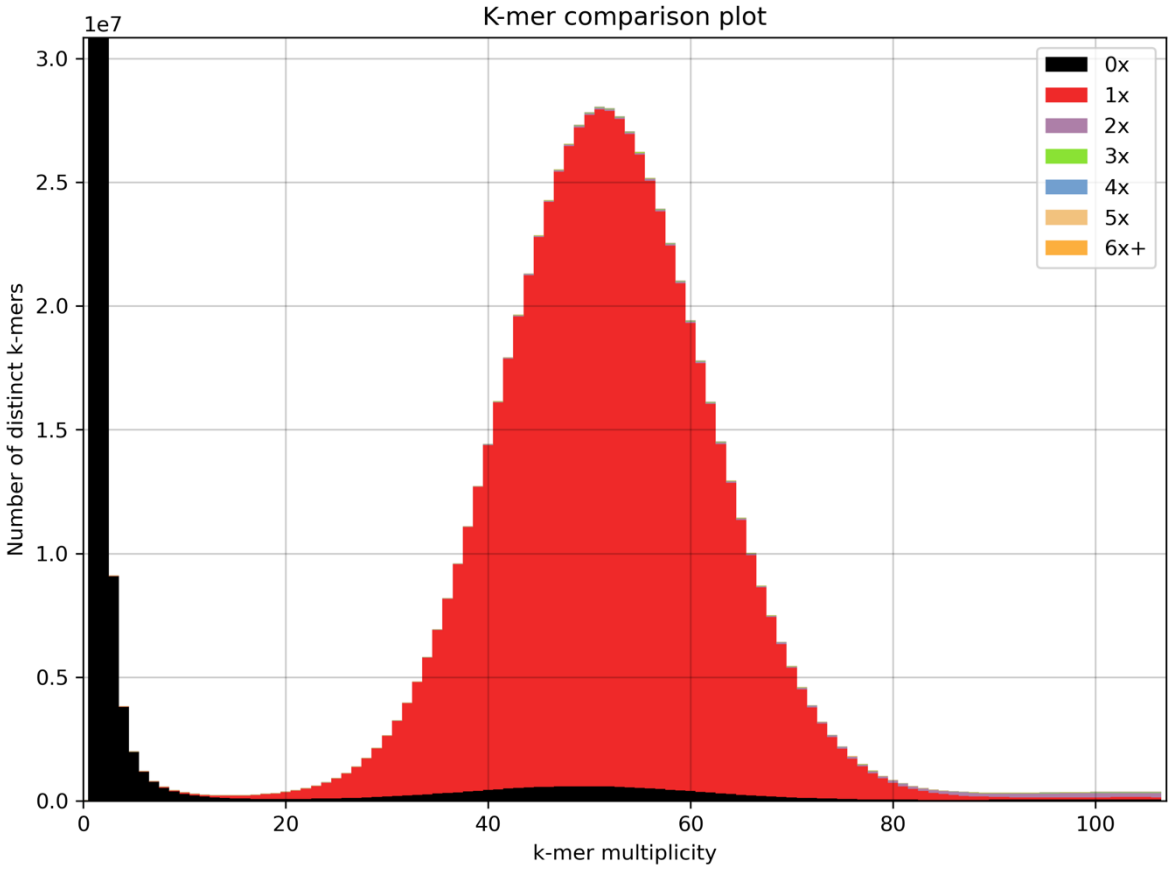
Specie	**Starting number of sequences**	**Final number of sequences**
Reticulitermes lucifugus	80	60
Reticulitermes speratus	22	5
Zootermopsis nevandesis	65	39
Macrotermes natalensis	14	12
Coptotermes formosanus	64	39
Cryptotermes secundus	63	37
Acromyrmex echinatior	82	61
Blattella germanica	128	114
Camponotus floridanus	162	57
Drosophila melanogaster	60	60
Locusta migratoria	75	47
Nasonia vitripennis	227	224
Apis mellifera	141	136
TOTAL	**1183**	**891**

Suppl. Table S5. Source and accession number of publicly available proteomes used in this study.

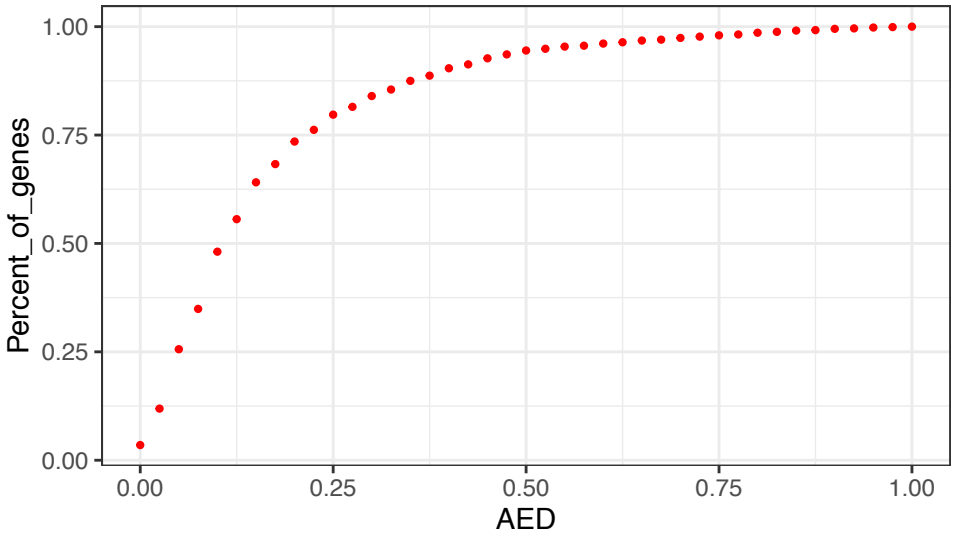
Specie	**Source**	**Accession number**
Coptotermes formosanus	NCBI	GCA_013340275.1
Macrotermes natalensis	GigaScience Database	http://dx.doi.org/10.5524/100057
Cryptotermes secundus	NCBI	GCF_002891415.2
Zootermopsis nevadensis	NCBI	GCF_002240175.1
Blattella germanica	i5K workspace	blager_OGS_v1.0
Locusta migratoria	i5K workspace	JAMg_OGSv1
Drosophila melanogaster	Uniprot reference proteomes	UP000000803
Camponotus floridanus	Uniprot reference proteomes	UP000000311
Acromyrmex echinatior	Uniprot reference proteomes	UP000007755
Nasonia vitripennis	Uniprot reference proteomes	UP000002358
Apis mellifera	Uniprot reference proteomes	UP000005203
Danaus plexippus	NCBI	GCF_009732855.1
Ephemera danica	i5K workspace	ephdan_OGSv1.0
Solenopsis invicta	NCBI	GCF_016802735.1
Reticulitermes speratus	FigShare	https://doi.org/10.6084/m9.figshare.c.5483235

Suppl. Table S6. Calibration points used for divergence time estimation.

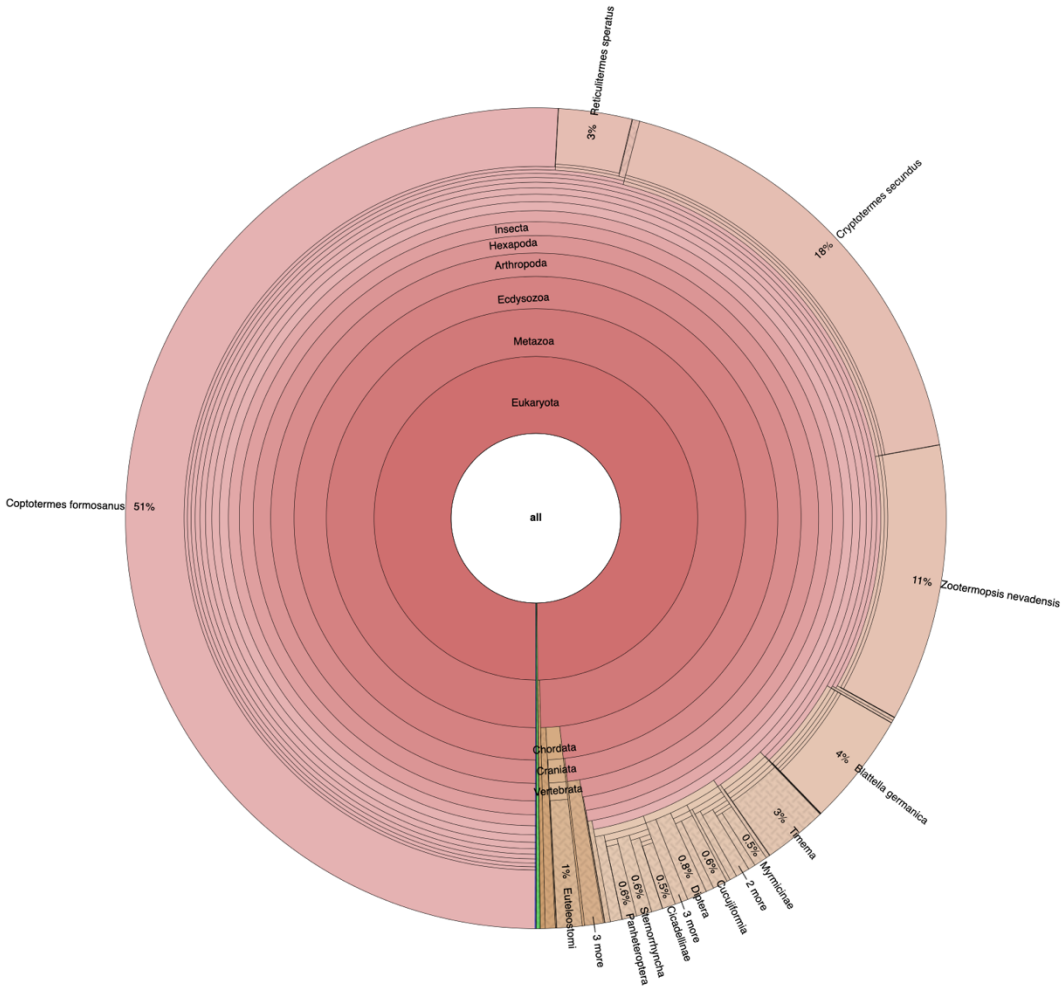
Node	**Estimate age of the split (Mya)**
R. lucifugus - *C. formosanus*	53
Z. nevadensis - other termites	132
L. migratoria - ingroup	295



Suppl. Fig. S1: Assembly spectra copy number plots; the plot shows nearly complete and homozygous assembly. K-mer present in the reads but missing from the assembly are represented in black while k-mer present in the reads and just once in the assembly are represented in red.



Suppl. Fig. S2: Cumulative distribution of AED values of annotated protein coding genes. All genes with an AED value < 0.5 (95 % of total genes) were kept in the final gene set.



Suppl. Fig. S3: Taxonomic annotation of *Reticulitermes lucifugus* proteome.