Supplemental methods and materials: The genomic consequences and persistence of sociality in spiders

Jilong Ma[1,2*], Jesper Bechsgaard[2], Anne Aagaard[2], Palle Villesen[1]. Trine Bilde[2*$], Mikkel Heide Schierup[2*$]

1. Bioinformatics Research Centre, Aarhus University, DK-8000 Aarhus C, Denmark
2. Department of Biology, Aarhus University, DK-8000 Aarhus C, Denmark

* Corresponding authors: aujilongm@birc.au.dk, trine.bilde@bio.au.dk, mheide@birc.au.dk
$ shared last authorship

Section 1 - Sampling and Sequencing

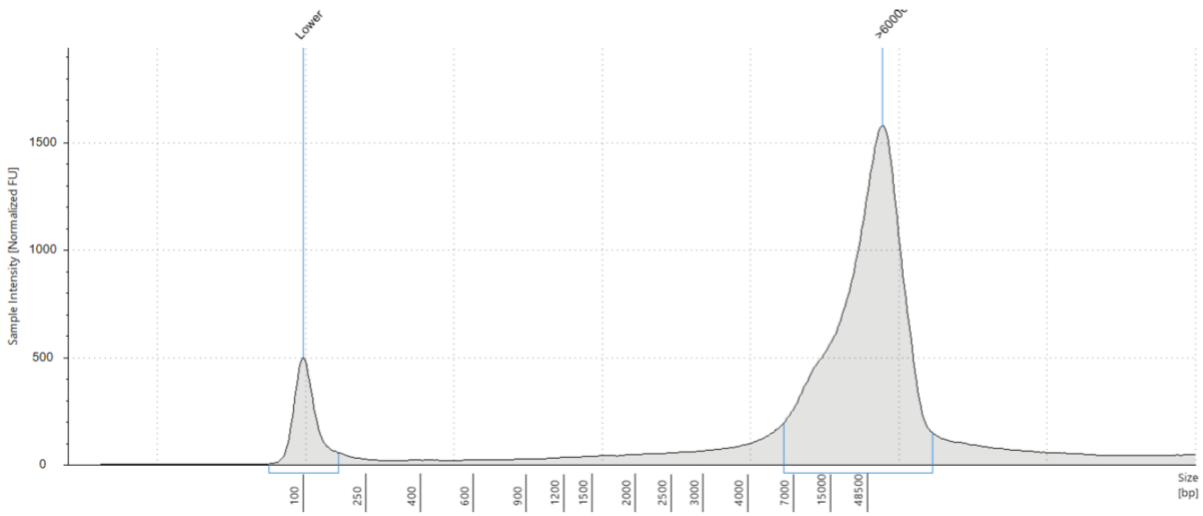
DNA - long reads

A single individual from *S. dumicola* (Namibia - Otavi), *S. tentoriicola* (South Africa - Tierberg), *S. mimosarum* (South Africa - Weenen), *S. sarasinorum* (India - Unknown), *S. bicolor* (Namibia - Betta) and *S. lineatus* (Israel - Negev Desert) were sampled, and DNA was extracted using the MagAttract HWM DNA kit from Qiagen (Hilden, Germany). The DNA was fragmented to 15-20 kb fragments using Megaruptor 3, and libraries were prepared using Pacific Biosciences protocol for HiFi library prep using SMRTbell® ExpressTemplate Prep Kit 3.0. Final library was size selected using BluePippin with a 10kb cut-off. Each library was sequenced on three 8M SMRT cells on Sequel II instrument using Sequel II Binding kit 2.2 and Sequencing chemistry v2.0. Loading was performed by adaptive loading, movie time: 30 hours, pre-extension: 2 hours. Between 66 and 99 gb data was obtained (see Supplementary Table S1).

Tapestation profiles of DNA extractions for long read sequencing:

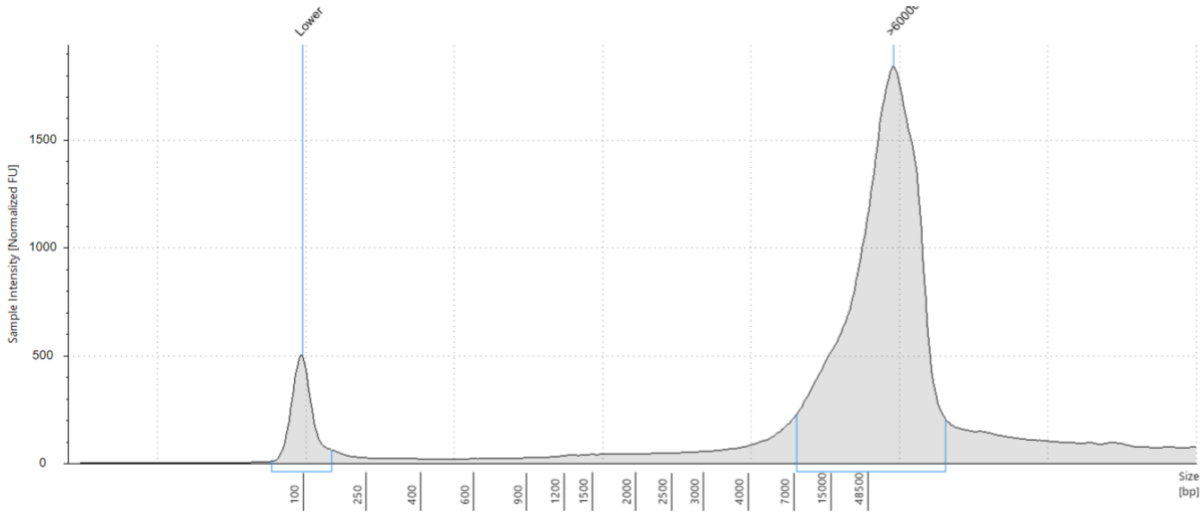
Electropherogram of DNA used for PacBio HiFi sequencing on an Agilent TapeStation for each of the six species. The peak around 100bp serves as a standard for both fragment size and intensity.

Stegodyphus dumicola:



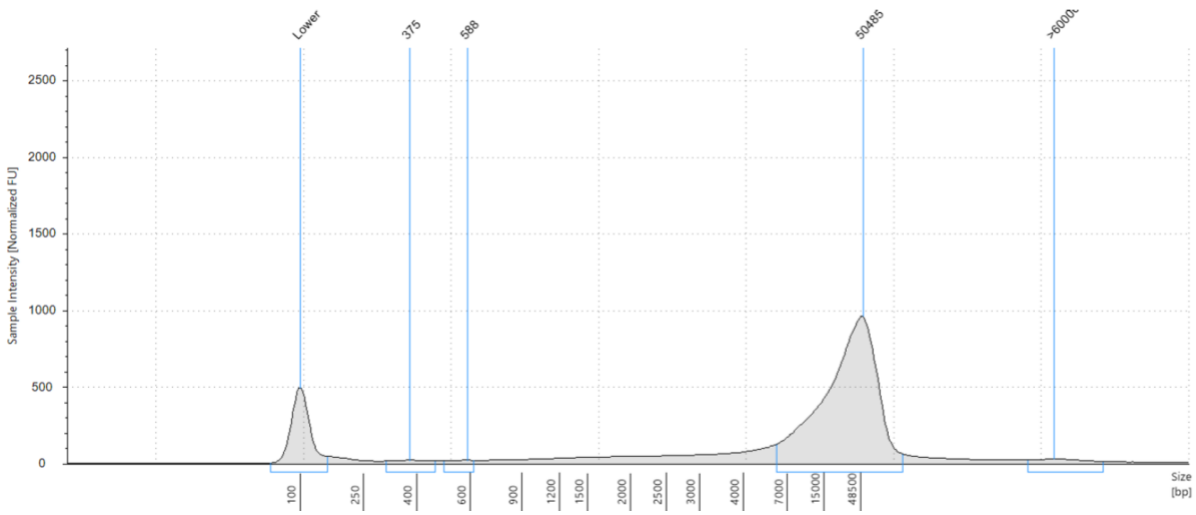
1

2 *Stegodyphus tentoriicola*:

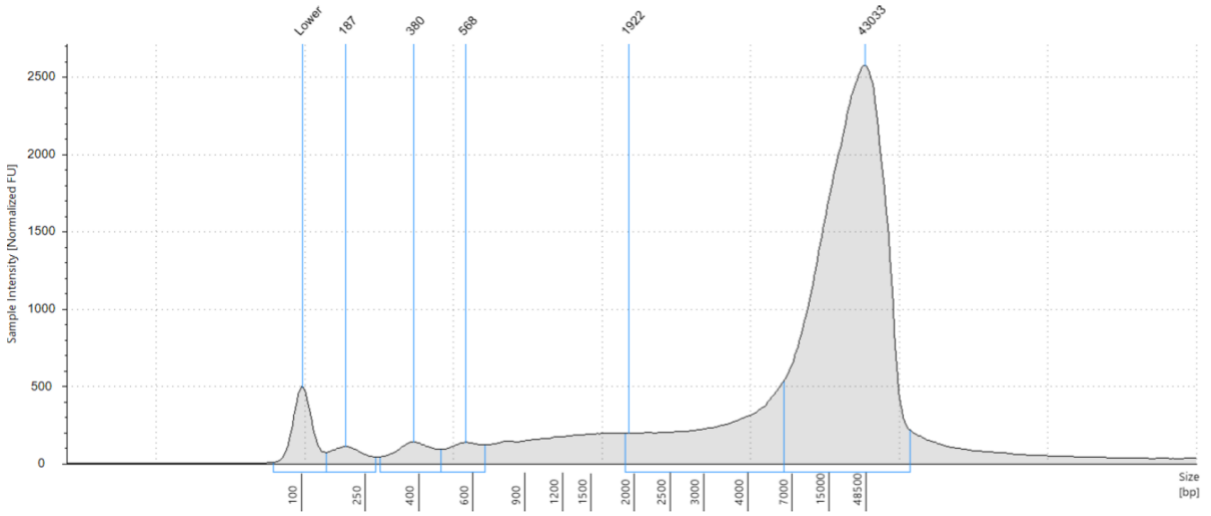


3

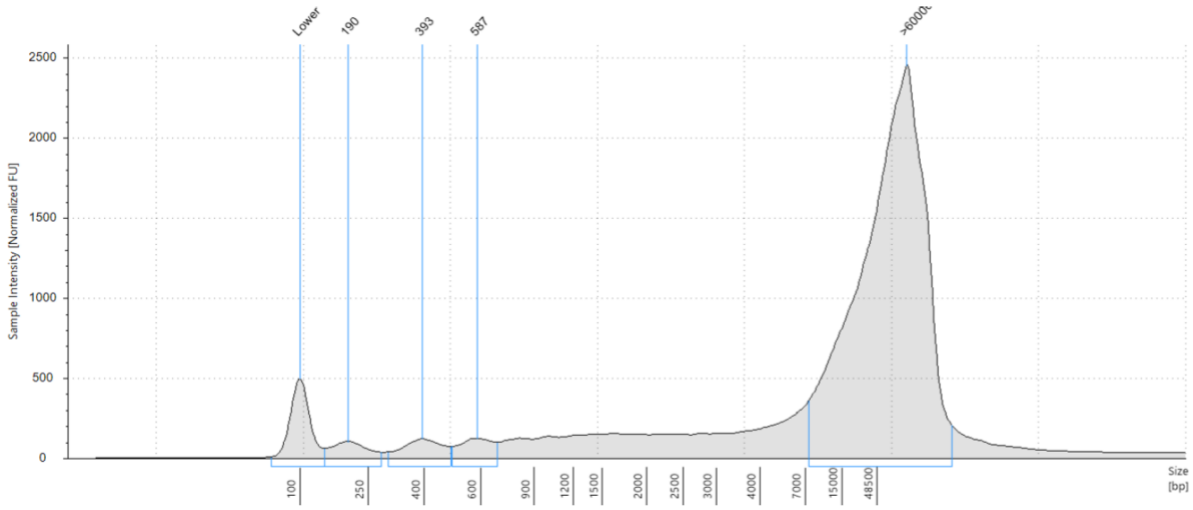
4 *Stegodyphus mimosarum*:



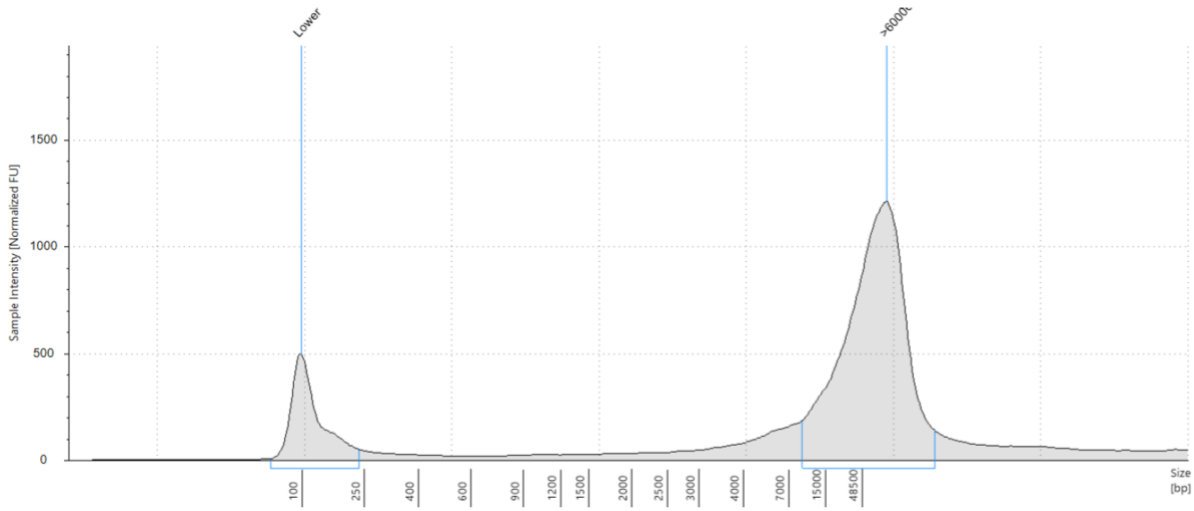
Stegodyphus sarasinorum:



Stegodyphus bicolor:



Stegodyphus lineatus:



DNA - Hi-C

A single individual from each species from the same populations that were sampled for PacBio sequencing were sampled for hi-C sequencing. The social individuals were taken from the same nests. Libraries were prepared from 6 legs per species using the Dovetail® Omni-C® Kit, and each library was sequenced using DNBSEQ-G400 to obtain between 139 and 153 gb 150PE data per sample (see Supplementary Table S1).

Resequencing

From each of the social species two individuals from isolated genetic lineages were sampled for resequencing to use for estimating 'social' d_N/d_S ratios (pi_N/pi_S in practice). From *S. mimosarum*, individuals were sampled in South Africa (Weenen) and Madagascar (Antananarivo); from *S. dumicola*, individuals were sampled from two populations in Namibia (Otavi and Betta); from *S. sarasinorum*, individuals were sampled from Himalaya and Sri Lanka(Settepani et al. 2014). Five single individuals from the subsocial *S. pacificus* were sampled in India for resequencing to be used for generating a reference genome by aligning reads to chromosome-level assembly of *S. sarasinorum*. DNA from all individuals were extracted using the Qiagen Blood and Tissue kit, Qiagen (Hilden, Germany), and the DNA was sequenced using DNBSEQ-G400 to obtain at between 40 and 146 gb 150PE data per sample (see Supplementary Table S1). DNBSEQ is a high-throughput DNA sequence technology developed in BGI. Its quality score per nucleotide per read would be lower than HiFi. Nevertheless, our individuals have a decent coverage depth per genome to call and filter for solid variants. Hence, the final variant quality from short read sequencing (piN/piS) and HiFi genomes (d_N/d_S) would be comparable. We are confident that technology differences would not raise noise that heavily biased the comparison of results.

RNA resequencing

To guide the annotation of protein-coding genes, we sequenced the transcriptomes of several individuals from each species. Three to five families were established by controlled crosses, and three offspring from each family were sampled as spiderlings (n=3), subadults (n=3) and adults (n=3). All parental individuals came from the same populations as used for long read and Hi-C sequencing. Families were produced as follows: 1) Social species were mated among individuals from the same nest. Initially we created several groups (15+) of 6 subadults to be sure they were virgins until we could determine the sex. When they became adults, we kept only groups with 1 male and 5 females. A

few weeks after they were all sexually mature, we froze down the male, and placed the females in individual boxes. The females that laid an egg sac were kept. After the eggs hatched and before matriphagy, one female per group was frozen down and so were six of her offspring. 2) Subsocial individuals were raised from hatching egg sacs produced by females collected in the wild. When sexually mature, females and males with separate mothers were mated (15+), and the male was frozen down. The females that laid an egg sac were kept. After the eggs hatched and before matriphagy, five females were frozen down and so were six of their offspring.

We extracted RNA all individuals using the Qiagen RNeasy Mini kit, and sequencing libraries were constructed using NEBNext Ultra II Directional RNA Library Prep Kit that were sequenced on Illumina NovaSeq 6000 to obtain ~6GB 150PE data per individual. (see Supplementary Table S2)

Section 2 - *De novo* assemblies and annotations

De novo assembly

We generated chromosome-level assemblies for all six species with PacBio HiFi long reads and Hi-C sequencing. We started with using Hifiasm(Cheng et al. 2021), a haplotype-resolved *de novo* assembler for PacBio HiFi reads, to assemble contigs for each species using the default settings. Both PacBio HiFi reads and Hi-C reads were used in this process. We then selected the haplotype with the longer phased assembly graph to retrieve the fasta sequence for contig scaffolding.

Next, we used the Juicer software to align the Hi-C reads to the long read contigs generated in the previous step. We subsequently employed the 3D-DNA (Dudchenko et al. 2017) run-asm-pipeline.sh script to order and orient the contigs based on the aligned Hi-C reads. We customize the settings with "-r 0 --editor-repeat-coverage 30 --editor-coarse-stringency 20" to omit mis-join correction rounds in the 3D-DNA scaffolding pipeline. This decision was made because the mis-join correction tends to

break long contigs joined from PacBio HiFi long reads in our practice, which would introduce more mis-joining. This customization resulted in a single "mega-scaffold" containing ordered and oriented contigs, with 500 bp Ns introduced at the joints.

Finally, we examined the mega-scaffold Hi-C contact map to identify contigs belonging to the same chromosome, as they exhibited distinct intra- and inter-chromosomal Hi-C contact patterns. We then manually reviewed, edited, and split the mega-scaffold into chromosome-level scaffolds for each species using Juicebox(Durand et al. 2016), ultimately generating the final chromosome-level fasta file. The final Hi-C contact maps for all six species are shown in Figure S1. The Hi-C contact map indicates that the scaffolding of chromosomes are visually clean with very few minor misjoined contigs between chromosomes. Further manual curation on the HiC-scaffolding could have been possible to "solve" the few seemingly mis-joints but we should claim clearly that there is no robust way in confirming the precision of the manual curation result. The inherent characteristics of larger contigs derived from long-read HiFi sequencing ensure that the credibility of subsequent analyses focused on local variants and genes remains intact. This is due to the fact that potential local misjoins, typically occurring at a broader spatial scale, do not significantly impact these analyses.

Genome Annotations

RepeatModeler2 (Flynn et al. 2020) was initially applied to construct repeat databases that are specific to each species. Following this, we employed RepeatMasker (Tarailo-Graovac and Chen 2009) to soft-mask the genome assembly for each species, by integrating each species-specific repeat database and the Repbase Arthropoda repeat database(Bao et al. 2015) to form the repeats library.

We used STAR(Dobin et al. 2013) to do spliced alignments for the RNA sequence from every individual sample. The aligned RNA bam files from all the individual samples for a species were then combined using Samtools. This collective data served as transcriptome hints for predicting genes.

We expedited the annotation process by running the BRAKER2(Brůna et al. 2021) ETP mode pipeline independently on each repeat-masked chromosome for each species in parallel. This process involved the use of the aligned species RNA bam file and the NCBI *S. dumicola* protein sequence as the transcriptome evidence and protein homology evidence respectively. The annotations derived from each chromosome were then assembled and merged into a single annotation file in the gff3 format for each species. Lastly, we used BUSCO(Simão et al. 2015) to gauge the completeness of the genome annotations, employing the Arthropoda ortholog database for this purpose.

The above annotation pipeline was completed for the majority of the chromosomes in all species, with two exceptions where the BRAKER2 pipeline using ETP mode failed to annotate the local part of the genome. The HiC_scaffold_11 (dum_8) of *S. dumicola* is annotated with the BRAKER2 with only RNA transcriptome data as the evidence. For the ending half of HiC_scaffold_16 (mim_6) of *S. mimosarum*, we used blat to search for *S. bicolor* mRNA sequence against the part of the genome sequence missing annotations. The hits of the blat search were further parsed as the hints for AUGUSTUS gene prediction. The results from AUGUSTUS gene prediction were combined with the results from BRAKER2 ETP mode.

Incorporation of species without chromosome-level assembly

We do not have the genome assembled from *S. africanus*. To include *S. africanus* in the analyses, we downloaded the transcriptome data of *S. africanus* from Bechsgaard et.al 2019 (Bechsgaard et al. 2019) and used Trinity (Haas et al. 2013) to assemble transcripts. We used DIAMOND(Buchfink et al. 2021) blastx mode to align the transcripts to the database of all the translated amino acid sequences from the 10065 single-copy orthologous groups of all six species. 5590 out of 10065 single-copy orthologous groups find one or more hits from the Trinity transcriptome assemblies. We further check the aligned percentage of the hitted orthologous genes and filtered for a percentage of 75% to consider a transcript from *S. africanus* being a valid match to a certain single-copy ortholog groups identified, which ended up with 2649 single-copy orthologous groups.

We also do not have a genome assembled from *S. pacificus*. To include *S. pacificus* genes into the 2649 single-copy orthologous groups identified, we first used BWA-MEM2 (Vasimuddin et al. 2019) to aligned short-read DNA sequence from a *S. pacificus* individual to the *S.sarasinorum* reference genome, which is the closest sister species of it. Then we used SAMtools (Danecek et al. 2021) consensus to call the consensus sequence as the genome of *S. pacificus* based on the BAM file. The sequences of 2649 ortholog genes in *S. pacifcus* are then retrieved using the genome annotation file of *S. sarasinorum*. This process can be challenging when there are indels found in *S. pacificus* compared to *S. sarasinorum* reference when we construct the consensus sequence of *S. pacificus*. The sequence of *S. pacificus* and *S. sarasinorum* will not be in alignment as they have different total lengths with different genome coordinate systems. In practice, we fill gaps for deletions and remove insertions in *S. pacificus* based on the reference of *S. sarasinorum* when we construct the consensus sequences. This practice is achieved by setting parameters for SAMtools being "consensus --show-del yes --show-ins no". By ignoring the indels in such a way, we maintain the same coordinate systems between the *S. pacificus* and *S. sarasinorum*. Orthologous sequences in S. pacificus can be then retrieved directly using the genome annotation file of *S. sarasinorum* to build multiple sequence alignments across all 8 species for d_N/d_S estimations.

Sex Chromosome Identification

We used bwa-mem2 to align reads from a single male individual to the reference genome for each species with a genome assembled. The read depth at each position, covered by at least one read, was obtained using samtools depth. Subsequently, the depth distribution across each scaffolded chromosome was visualized. Chromosomes exhibiting a relative depth in mean and median of half compared to others were identified as X Chromosomes, as detailed in Figure S8.

Genome Quality Control

To assess genome quality, we employ Mercury, which measures k-mer completeness and base pair accuracy. In the case of subsocial species, where we successfully reconstruct both haplotypes, Merqury's metrics are provided for the diploid genome assembly. Conversely, for social species

characterized by extensive inbreeding, leading to the acquisition of a single haplotype, Merqury's metrics are presented exclusively for the haploid assembly.

Section 3 - d_N/d_S ratio estimations

Estimating the social transition time using intensity of selection

The selection efficiency is expected to be relaxed after social transitions as multiple traits (reproductive skew, female biased sex ratio, inbreeding) acts to reduce Ne and elevate effects of drift. Phylogenetic methods (Such as PAML in this study) provide d_N/d_S estimation as a single value per lineage (shown in Figure 2.D). The single d_N/d_S value for the social lineages thus becomes a weighted mean of varying selection efficiency through time. We simplify the scenario of selection efficiency change as a single instant of social transition with a lower d_N/d_S before and a higher d_N/d_S after (Figure M1). Once knowing the weighted d_N/d_S across whole lienages and the two d_N/d_S values before and after the transition, the time fraction of the social period can be derived.

As it is challenging to estimate the d_N/d_S before and after the social transition in social lineages directly, we use approximations. We use the d_N/d_S from the subsocial sister species of each social species as the approximation for d_N/d_S before the social transition and pi_N/pi_S between isolated populations of social species as the approximation for d_N/d_S after the social transition ("social d_N/d_S").

Benchmarking for substitution of "social d_N/d_S" with pi_N/pi_S across social species populations

The general application of d_N/d_S aims to test the selection strength by investigating the fixed non-synonymous and synonymous substitutions between species. We used pi_N/pi_Sbetween divergent populations of the same species as an approximation of "social d_N/d_S". This implication has a risk of overestimating the d_N/d_S. As d_N/d_S compares the fixed nucleotide difference between species under certain selection strength, while pi_N/pi_S includes the nucleotide differences that are not fixed by selection yet. Thus, the sites included in the pi_N/pi_S would likely contain more deleterious non-synonymous mutations that are not removed by selection, which leads to an overestimated d_N/d_S. The

overestimation could happen if the selection strength on the non-synonymous sites are much higher than the selection strength on synonymous sites, which result in a higher observed pi_N/pi_S than the actual d_N/d_S after fixation.

First, we benchmarked the reliability of this substitution strategy by comparing the site frequency spectrum of non-synonymous and synonymous sites in an *S. dumicola* population, which has the shortest population divergent time with potentially more polymorphisms unfixed between populations. We called genotypes for 9 individuals from *S. dumicola* using the GATK pipeline and filtered for bi-allelic single nucleotide variants where the genotype quality is over 30. We used snpEff (version 5.2) (Cingolani et al. 2012) to build a database with our de novo assemblies and annotations and classify variants as either missense variants or synonymous variants. After the classification, the site frequency spectrum was built for missense variants and synonymous variants respectively (Figure S7). The high similarity of N-sites spectrum and S-sites spectrum suggest the ongoing selection acting similarly on N-sites and S-sites, countering the concern that selection is more effective in removing polymorphism in N-sites than S-sites.

Second, we evaluate whether there are substantial amounts of deleterious polymorphisms segregating in the social species leading to over estimation of d_N/d_S by using pi_N/pi_S. We assume that deleterious mutations that can be removed by selection are less likely to segregate within the population. We then quantify the fraction of common variants out of the polymorphisms we identified between the two individuals selected from separate populations. We have 20 individuals (40 alleles per site, 10 individuals from each population) from *S. dumicola* and 15 individuals (30 alleles per site, 8 individuals from Madagascar and 7 individuals from mainland Africa) from *S. mimosarum*. We assume different minimum thresholds based on minor allele count (2 to 10) we found in each species to filter for “common variants” (See result in Table S5). With a conservative filter that the minor allele count has to be greater or equal than 10, common variants compose 66.43% of polymorphisms used for pi_N/pi_S estimation of *S. dumicola* are 91.52% of polymorphisms used for pi_N/pi_S estimation of *S. mimosarum*. The high fraction of common alleles of the used polymorphisms for pi_N/pi_S estimation

also support that there is only a small proportion of segregating deleterious mutations that can lead to overestimation of "social d_N/d_S".

In the end, we select the polymorphisms that are fixed differently between the two isolated populations used for cross-population pi_N/pi_S estimations. We find that 17.60% of polymorphisms are fixed in *S. dumicola* and 74.34% of polymorphisms are fixed in *S. mimosarum* (Table S6). We then estimate "social d_N/d_S" using only fixed polymorphism across populations, which would be expected to reflect the true selection efficiency free from unremoved deleterious segregating sites. We find the point estimation of "social d_N/d_S" for *S. mimosarum* is 0.2990 using fixed polymorphisms compared to previous estimation of 0.3043 using polymorphic sites between two individuals (Figure S12). And the new "social d_N/d_S" estimation for *S. dumicola* is 0.3558 compared to previous estimation of 0.3366. The high similarity between the two estimations quantifies the potential bias to overestimate "social d_N/d_S" is arguably neglectable in *S. mimosarum* and *S. dumicola*.

We do not have more individuals to create population sets for Himalyas population and Sri Lanka population of *S. sarasinorum.* Thus, no empirical benchmarking has been done for S. sarasinorum. However, population divergence time is longer for *S. sarasinorum* (110 kya) compared to *S. dumicola* (20kya), indicating the concern for overestimation would not be larger than for *S. dumicola*.

Coding gene sequence alignment and filtering for d_N/d_S estimation

For getting reliable alignments for d_N/d_S estimation, we made a strict and conserved filtering for ortholog groups. We ended up analyzing 2302 autosomal genes and 347 single-copy ortholog groups across the 8 species. After retrieving the nucleotide sequence from all species of each ortholog group, we did the alignment using MACSE alignSequences to account for potential frameshift since we are aligning coding sequences(Ranwez et al. 2018). With MACSE, we ensured that the alignment of each single-copy ortholog is always a multiple of 3 in length.

We do resampling estimation of d_N/d_S for X Chromosomes and autosomes separately. We randomly sample 500 or 100 ortholog groups out of the 2302 autosomal genes or 347 X Chromosome genes respectively. The random sample was repeated 500 times for autosomal genes and 100 times for X Chromosome genes to get standard error of the mean estimation. For each sampled set of genes, we concatenated the alignment using GoAlign(Lemoine and Gascuel 2021). We then checked for the concatenated alignment for every codon which contains gaps in any of the species, any codon with gaps in alignment will be marked as removed. The codon marked as removed further divided the whole alignment into continuous alignment blocks in different sizes. An overall distribution of the polymorphic site fraction and alignment block size (Figure S2) suggest that a small size of alignment block with high polymorphic fraction represents local mis-alignment in general. Hence, we filter for alignment blocks that are at least 300 continuous nucleotides long and the fraction of polymorphic sites in a local alignment block should be less than 15%. The above filter was conserved and might lead to fewer sites in final d_N/d_S estimation, which can be compensated with the concatenation process, but should be more robust to be free from mis-alignment.

Resampling strategy for quantifying confidence interval of d_N/d_S estimations

The choice of "sampling 500 times of 500 genes out of 2302 autosomal genes without replacement" is relatively arbitrary and aiming for a reasonable estimation for getting variation estimates for our genome-wide selection intensity (d_N/d_S) estimations. If we choose too many genes, for example 2000, then each run of the resampling without replacement will be similar, which may not reflect the true variance in the genome-wide selection intensity estimation, because data between resampling runs are expected to be highly overlapping. Meanwhile, we aim for a "genome-wide" selection intensity estimate, which requires a substantial amount of genes. Otherwise, gene-specific selection intensity contributes to higher noise. That is why we arbitrarily choose 500 genes, which should not result in a high overlapping rate between resampled data sets but still be able to reflect a genome-wide signal. We show that 500 rounds of resampling already reveal a normal distribution of estimated genome-wide selection intensity (Figure S10). It is also worth noting that the resampling here requires

substantial computation resources as each round requires a resampling of the raw sequence data, thus we limited the number of rounds to what we regard a sufficient amount.

To get a confidence interval of estimated social transition time, we performed 10000 rounds of calculation for social transition times. The calculation of social transition time need four values, which are the d_N/d_S on the social lineage, the d_N/d_S on the subsocial sister species lineage, and the pi_N/pi_S from the isolated populations of social species, and the dS branch length of the social lineage. All the four values have a confidence interval estimated from the previous resampling process (500*500). To account for the influence of uncertainty of selection intensity and species divergence time on our estimation of social transition, we sampled four values 10000 times from each their distribution (built based on the 500*500 resamples), which allow us to explore the confidence interval of social transition times. This final resampling process is less computational intensive since we only need to sample values from a known distribution, which allows us to do it as many times as needed.

Testing for relaxed selection

We use RELAX from HyPhy (Wertheim et al. 2015) to test for the observed higher d_N/d_S in social species due to intensification of positive selection or relaxed purifying selection. We used the concatenated alignment of all autosomes and X Chromosomes separately and 4 different grouping strategies of testing branches and reference branches. For each social species, we used the social species branch as the "Test" and the subsocial sister branches and common ancestor branches as the "Reference". We also have an extra contrasting group where all the social species branches are used as the "Test" and rest branches in the phylogeny are used as the "Reference".

RELAX classify the sites of a alignment into three categories, thus each category are assigned with a fraction:

a. Sites with d_N/d_S closer to 1. (o2 in Figure S3)
b. Sites with d_N/d_S much larger than 1.(o1 in Figure S3)
c. Sites with d_N/d_S much smaller than 1. (o3 in Figure S3)

Then RELAX estimates a d_N/d_S value (shown on x-axis of Figure S3) of sites in each category for reference branches and test branches separately. When relaxed selection dominates the observation of higher d_N/d_S in test branches than reference branches, the d_N/d_S value estimated for test branches in category a and c will be both shifted towards $d_N/d_S = 1$ compared to results of reference branches.

Supplementary Figures

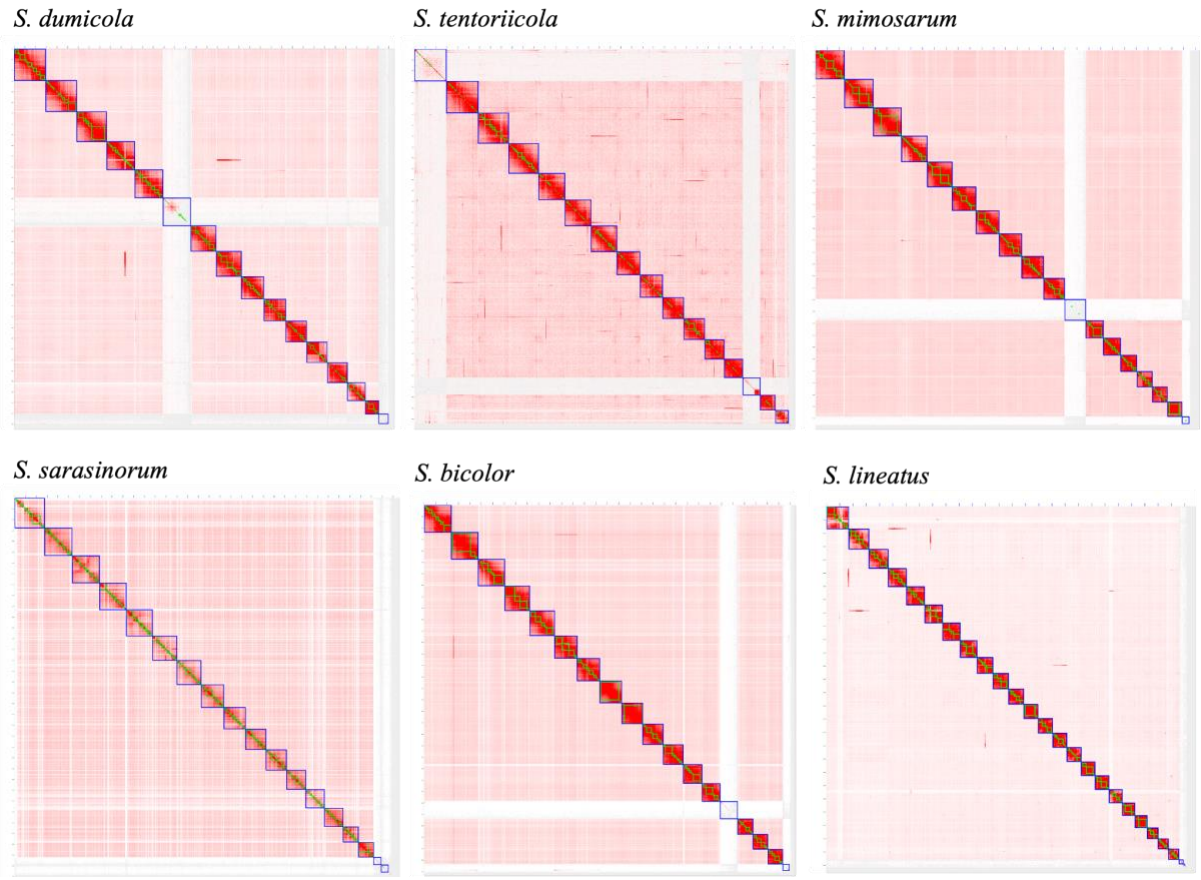


Figure S1. HiC contact map of genome assemblies. The density of red color denotes the hic contact density between regions in the assembly. The green box denotes initial contigs assembled from Pacbio HiFi long reads using hifiasm. The blue box denotes the candidate chromosome-level scaffolds. Each blue box supported by higher intra-scaffold density of HiC contact pattern compared to inter-scaffold HiC contact is identified as a chromosome. The blue box without high intra-scaffold density of HiC contact remains scaffolded.

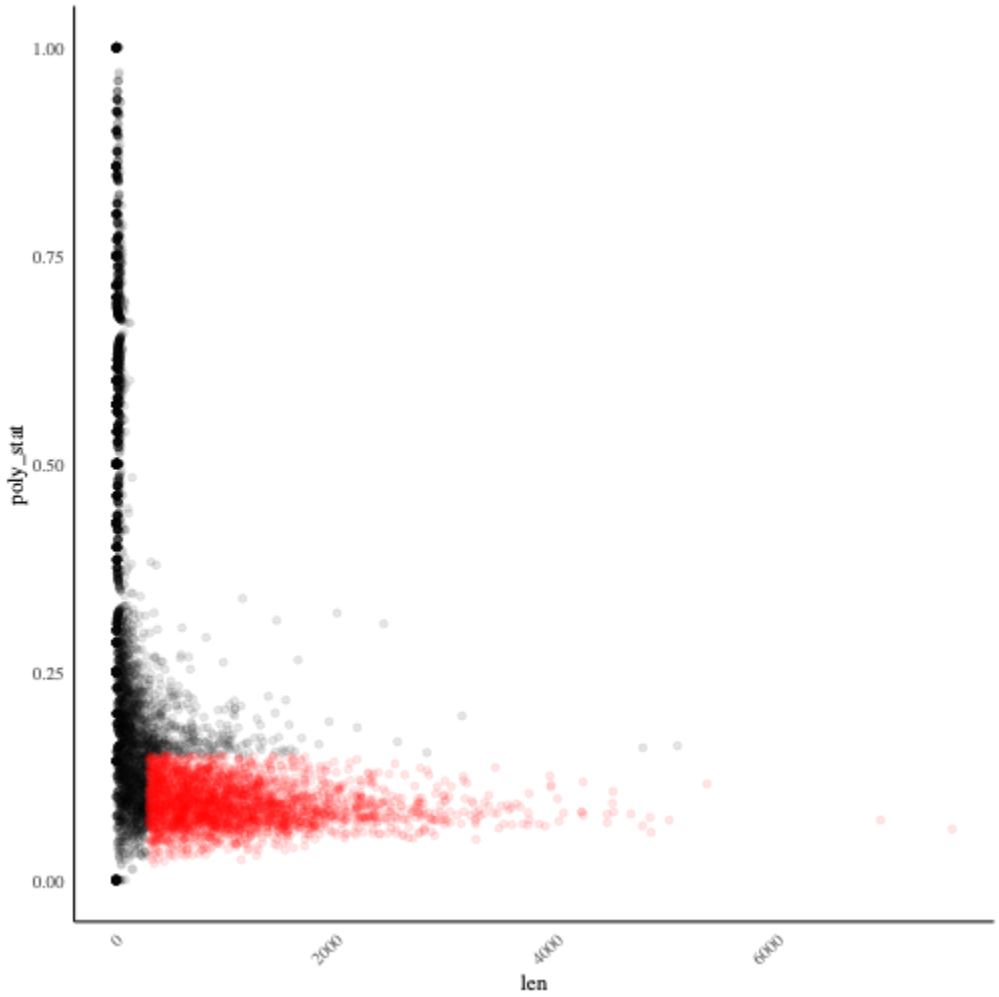
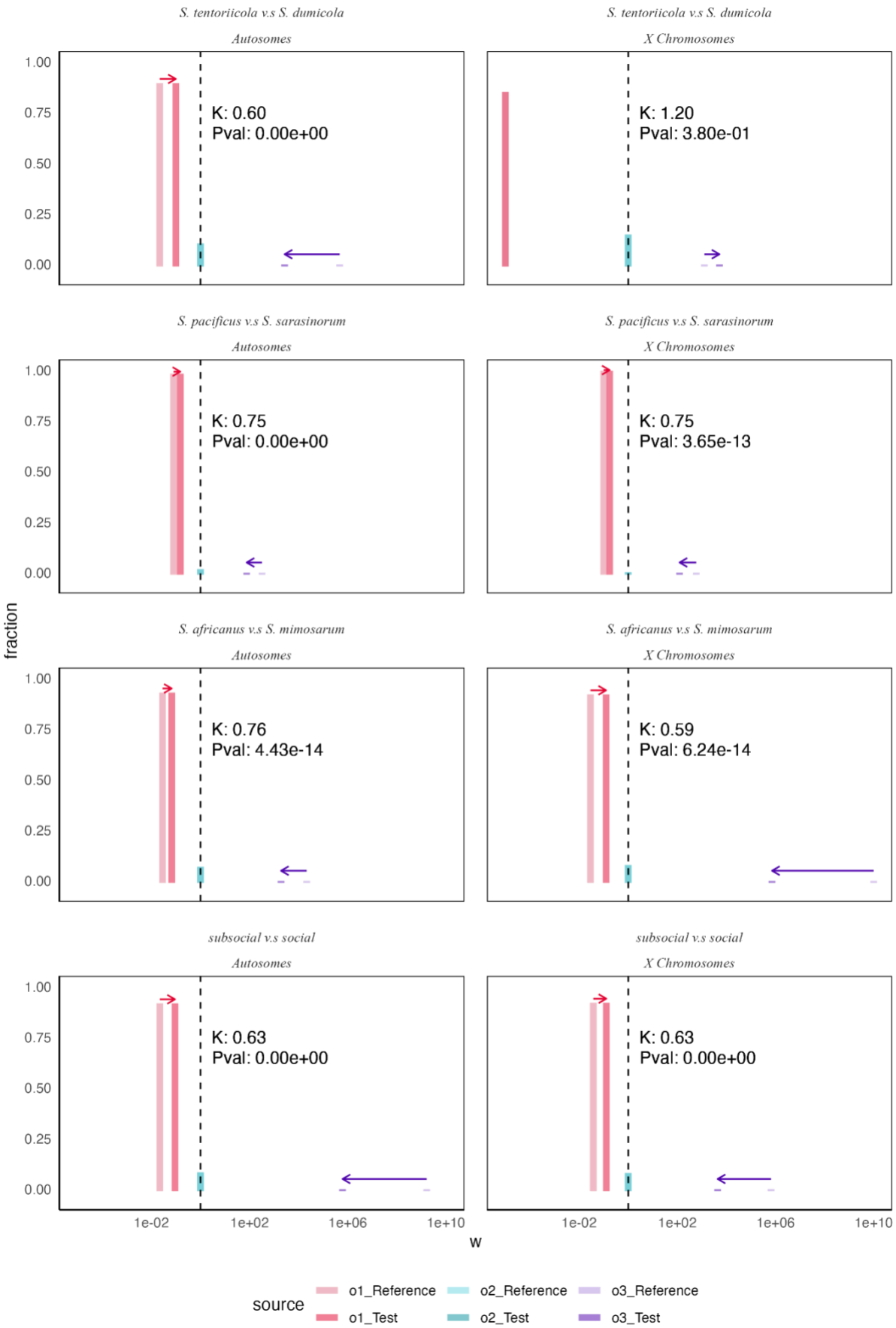


Figure S2 The size of alignment blocks without gaps versus the fraction of polymorphic sites for all alignment blocks of concatenated 2302 autosome gene alignments. The alignment blocks colored in red are further used for estimating d_N/d_S in PAML(Yang 2007).


S. tentoriicola v.s S. dumicola
Autosomes
K: 0.60
Pval: 0.00e+00
S. tentoriicola v.s S. dumicola
X Chromosomes
K: 1.20
Pval: 3.80e-01
S. pacificus v.s S. sarasinorum
Autosomes
K: 0.75
Pval: 0.00e+00
S. pacificus v.s S. sarasinorum
X Chromosomes
K: 0.75
Pval: 3.65e-13
S. africanus v.s S. mimosarum
Autosomes
K: 0.76
Pval: 4.43e-14
S. africanus v.s S. mimosarum
X Chromosomes
K: 0.59
Pval: 6.24e-14
subsocial v.s social
Autosomes
K: 0.63
Pval: 0.00e+00
subsocial v.s social
X Chromosomes
K: 0.63
Pval: 0.00e+00
fraction
1.00
0.75
0.50
0.25
0.00
1e-02
1e+02
1e+06
1e+10
w
source
o1_Reference
o1_Test
o2_Reference
o2_Test
o3_Reference
o3_Test

1

Figure S3. RELAX results for different contrasting pairs using genes from autosomes and X Chromosomes separately. The social lineage(s) are always used as the test set and their corresponding sister subsocial lineage(s) are used as the reference set. K > 1 implies intensification of positive selection and K < 1 implies relaxation of purifying selection. o1 represents the sites that are under strong purifying selection (w << 1), o2 represents the sites that are close to neutral (w ≈ 1), o3 represents the sites that are under positive selection (w >> 1). The x axis shows estimated w (d_N/d_S) for each category (o1, o2, o3) and the y axis shows the fraction of sites in the genome of the corresponding w categories (o1, o2, o3) for reference branches and test branches respectively. The arrow indicates the shifting direction from reference branches to test branches in the corresponding w category (o1, o2, o3). Relaxed selection in the test branches will be reflected as the shifting direction of arrows points to w = 1 (black dashed line) on both sides of the black dashed line.

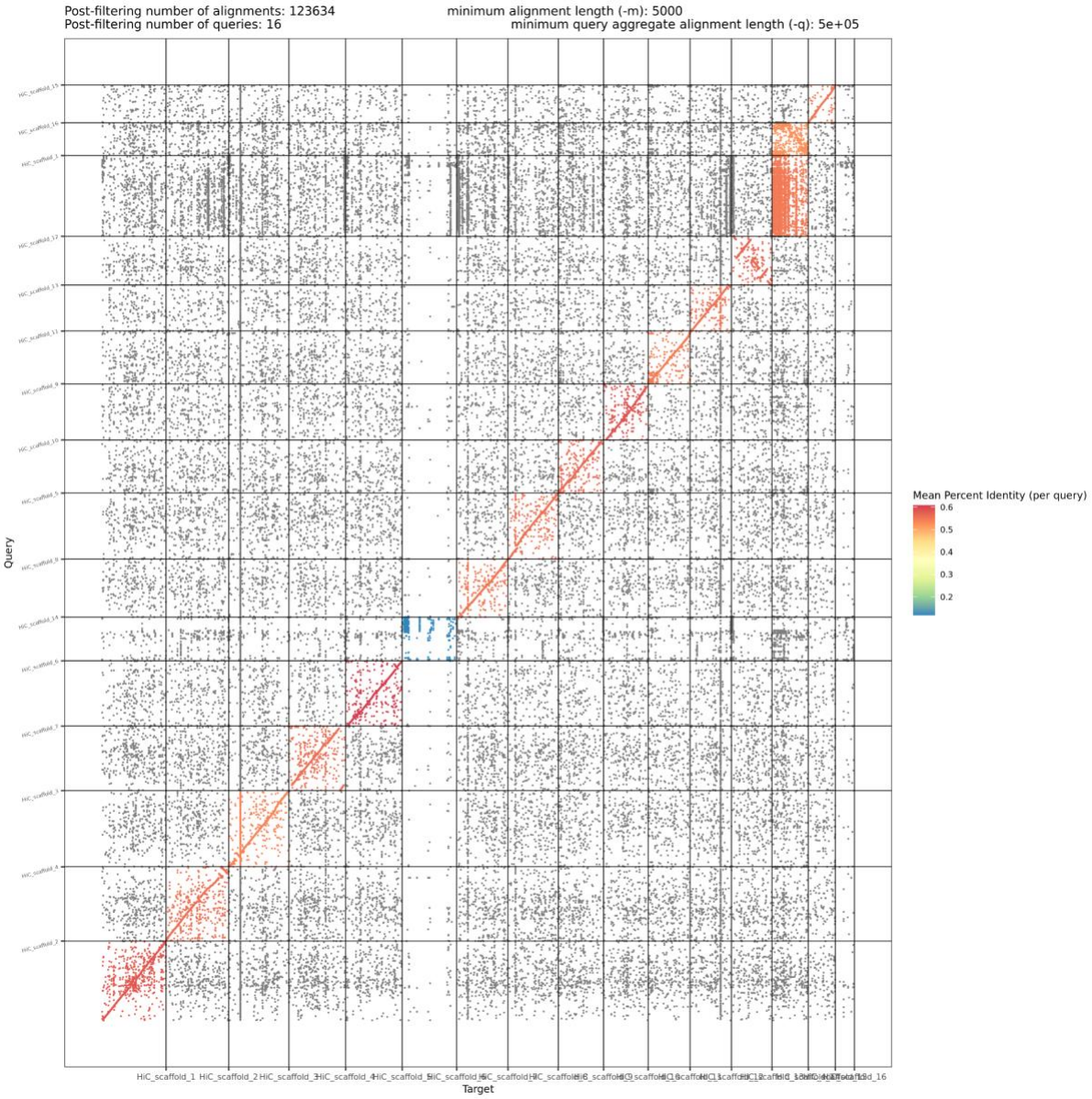


Figure S4. The dot plot from dotPlotly between genome assemblies of *S. dumicola* (x-axis) and *S. tentoriicola* (y-axis). Each point denotes an alignment length of minimum 5000 base pairs. Percentage of identity is shown for each comparison group of chromosomes.

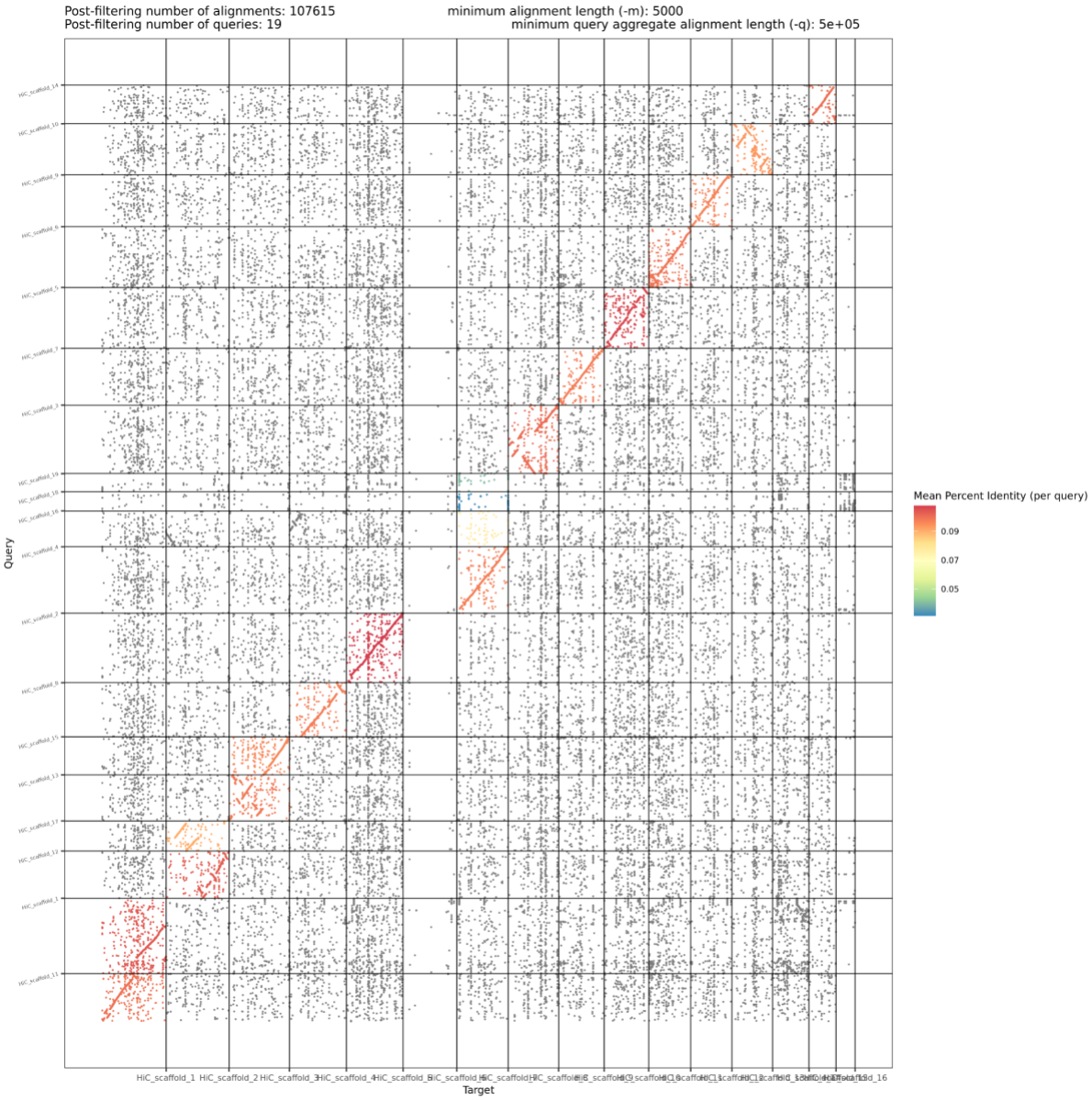


Figure S5. The dot plot from dotPlotly between genome assemblies of *S. dumicola* (x-axis) and *S. sarasinorum* (y-axis). Each point denotes an alignment length of minimum 5000 base pairs. Percentage of identity is shown for each comparison group of chromosomes.

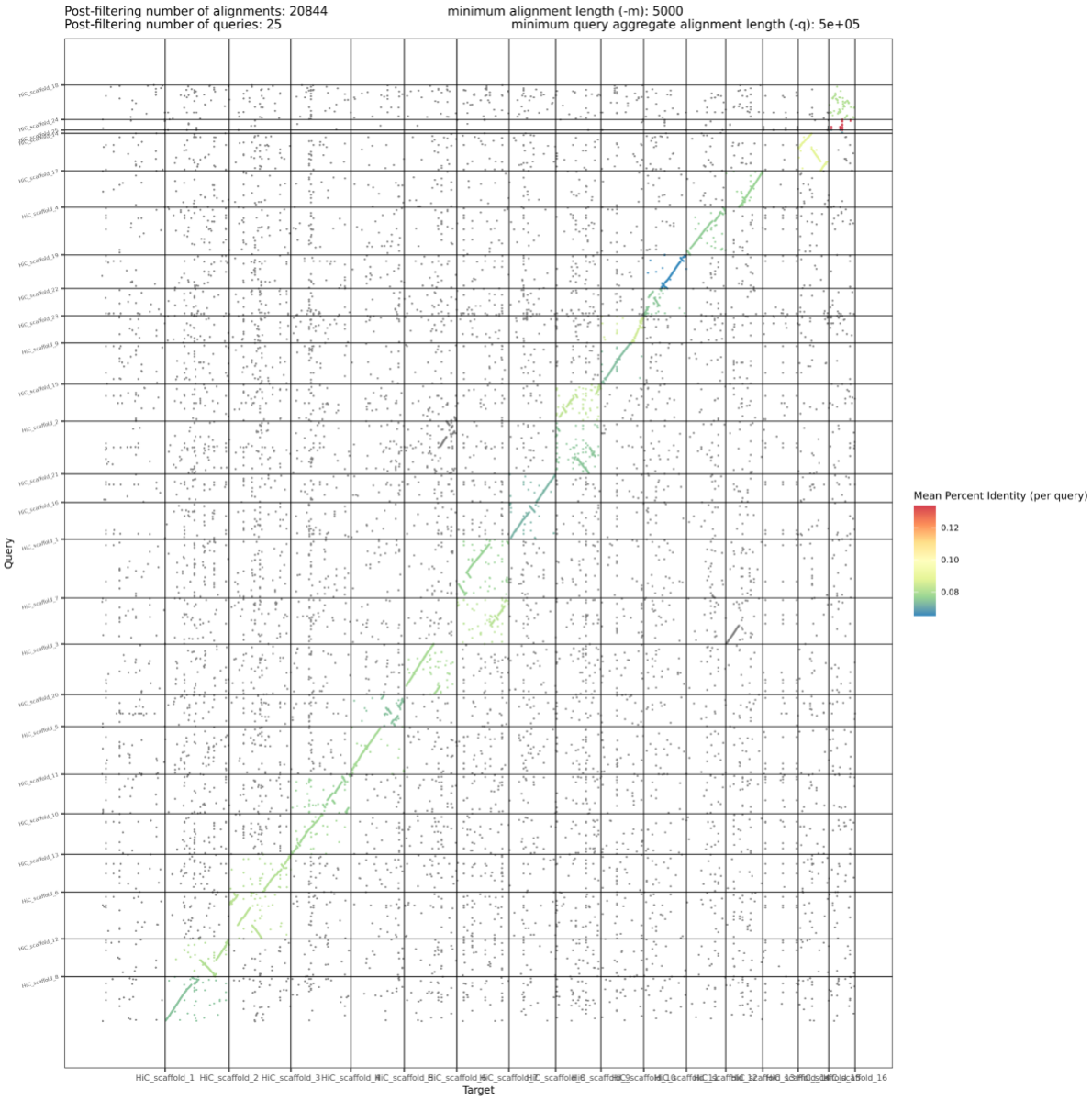


Figure S6. The dot plot from dotPlotly between genome assemblies of *S. tentoriicola* (x-axis) and *S. lineatus* (y-axis). Each point denotes an alignment length of minimum 5000 base pairs. Percentage of identity is shown for each comparison group of chromosomes.

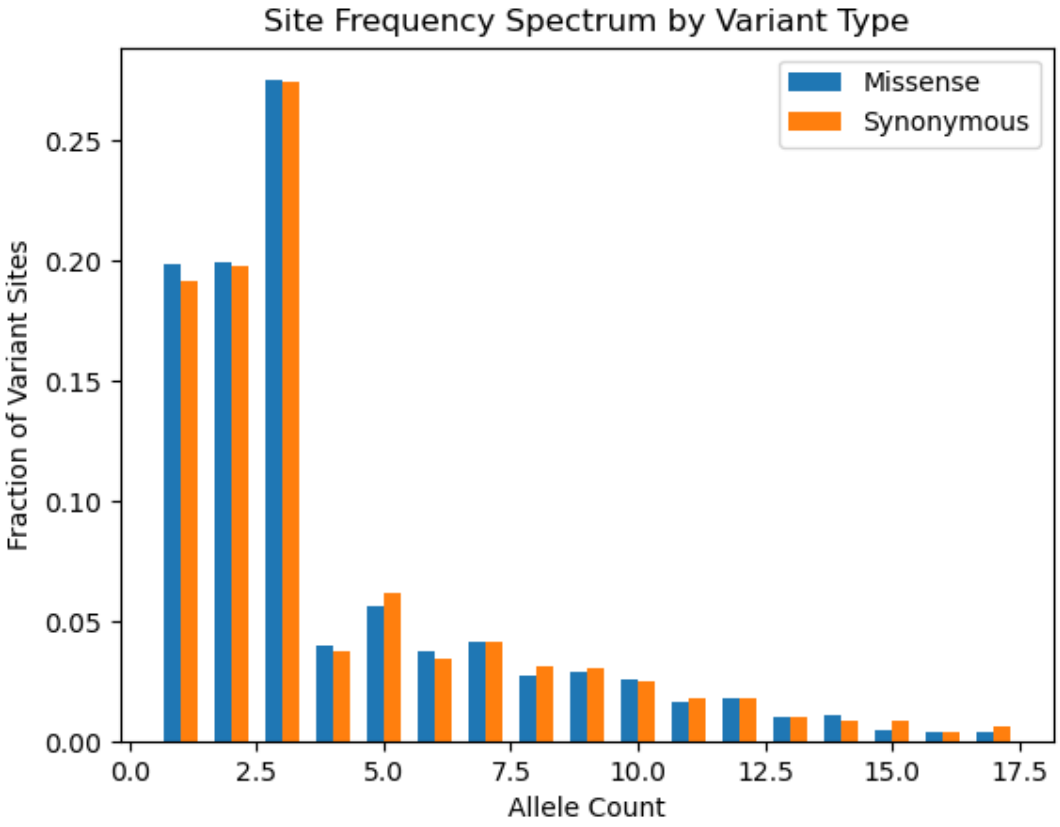


Figure S7. Site frequency spectrums for missense variants and synonymous variants separately based on 9 individuals from *S. dumicola.*

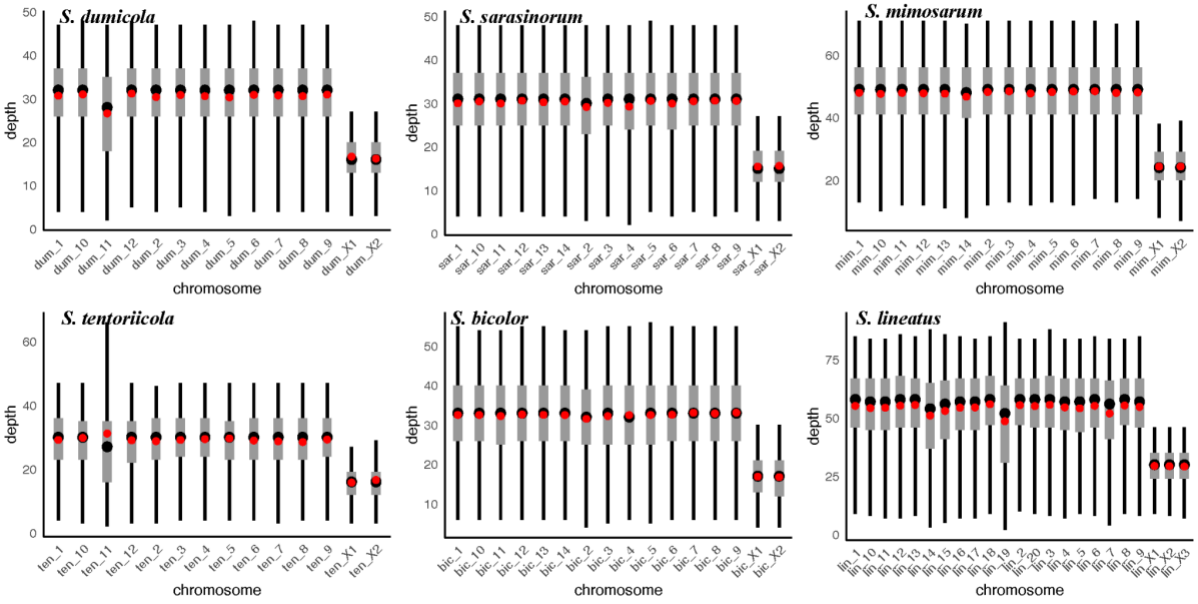


Figure S8. A boxplot of chromosome-level depth distribution from a single male individual for each species. The black lines mark from quantile 2.5% to quantile 97.5%. The grey boxes mark from

quantile 25% to quantile 75%. The black points and red points mark the median depth and mean depth respectively.

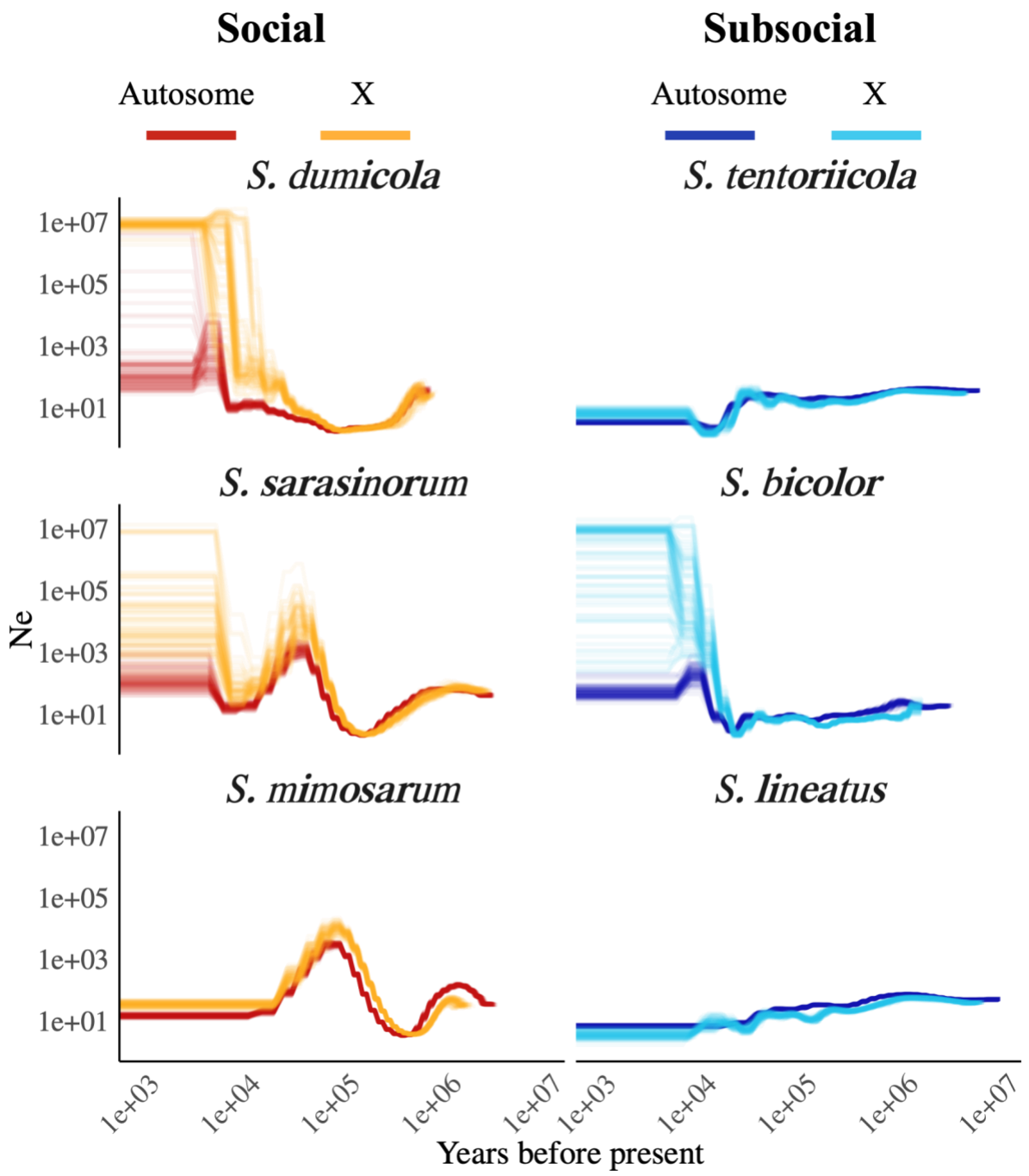


Figure S9. The historical effective population size inferred from the Pairwise Sequentially Markovian Coalescent (PSMC) model with 100 rounds of bootstrapping, setting segment size of 100000bp in resampling process, for different *Stegodyphus* species with chromosome-level assembly. Results from autosomes and X Chromosomes are shown separately for each species

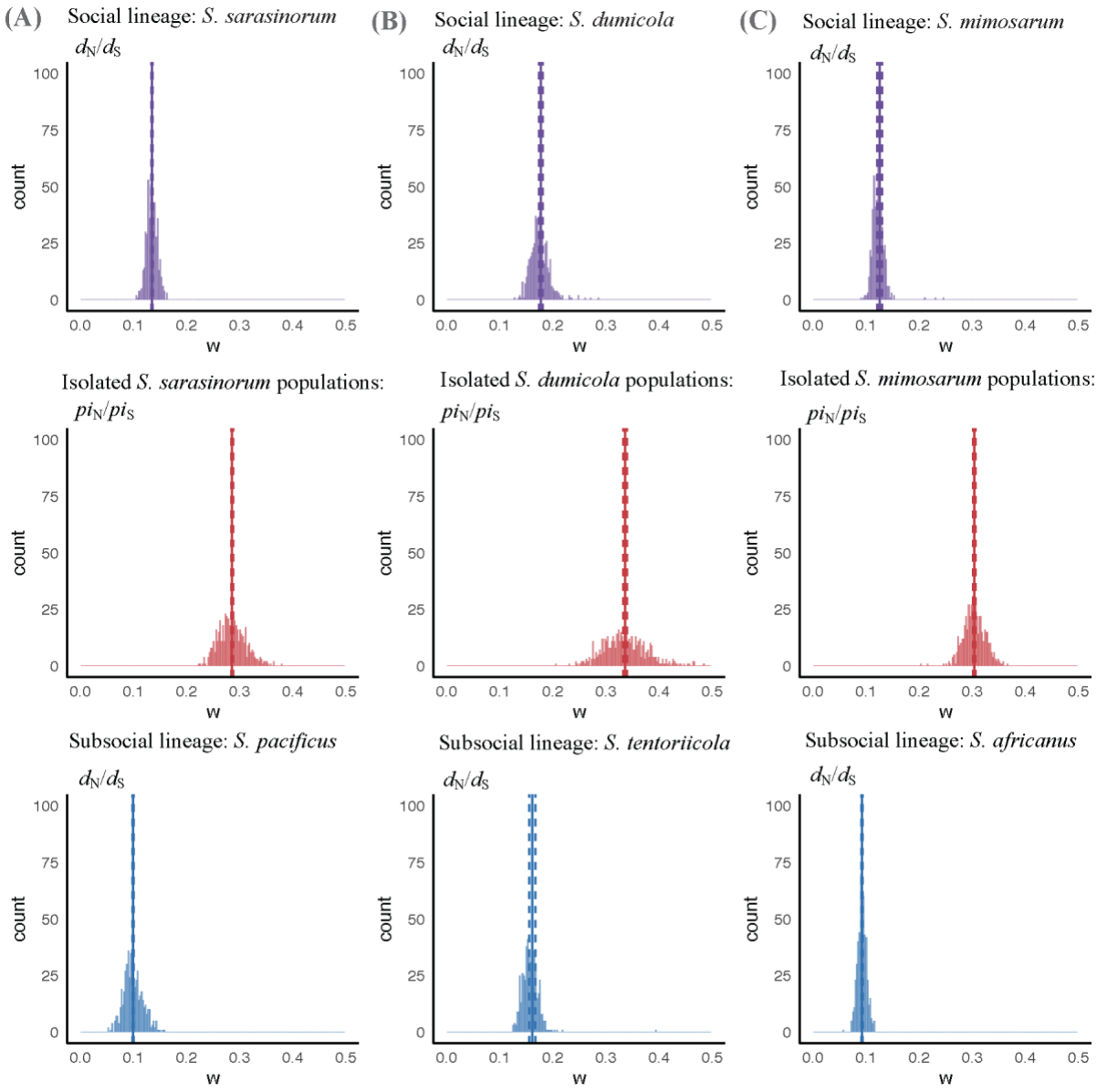


Figure S10. The distributions of d_N/d_S and pi_N/pi_S in each species, based on 500 independent resampling runs, where each run samples random 500 genes out of 2303 autosomal genes without replacement. The point estimation from the mean and 95% confidence interval of the mean are denoted by vertical lines in solid lines and dash lines respectively.

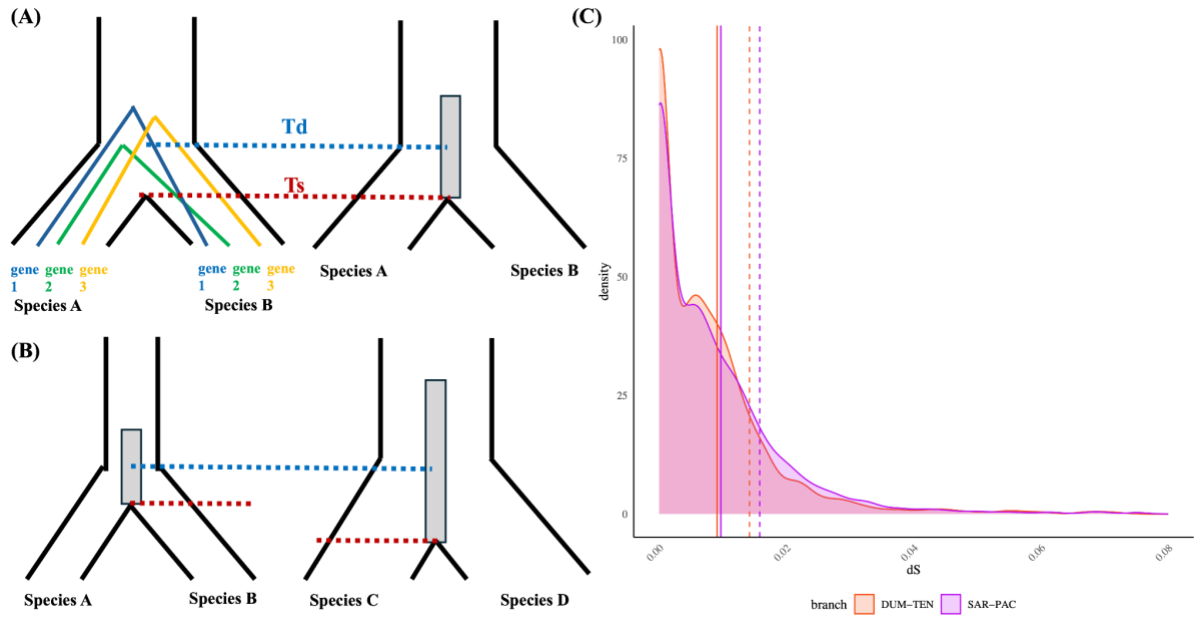


Figure S11. Coalescence expectations of distance between species divergence time (Td) and speciation time (Ts) for *S. dumicola* and *S. sarasinorum*. (A) Different ortholog genes can coalesce at different time points between a pair of species. This provides a time interval on average of 2Ne generations for a gene to coalesce in the ancestral populations, which is the distance between species divergence time (Td) and speciation time (Ts) (from the speciation to the middle point of the grey box). (B) Difference in expected coalescence time between Td and Ts for different ancestral population size. The species pair of species A and species B has a smaller ancestral Ne than the pair of species C and species D. (C) Distribution of dS from each autosomal ortholog gene (2303 in total) for *S. dumicola* (red) and *S. sarasinorum* (pink). The solid vertical lines are the median of the dS distributions, which is expected to correlate with Td (blue dashed line) in (A) and (B). The dashed vertical lines are the 75% quantile of the dS distributions. The distance between the solid vertical line and dashed vertical line in each species is expected to correlate with the width of the distribution, which represents the height of the grey box illustrated in (B).

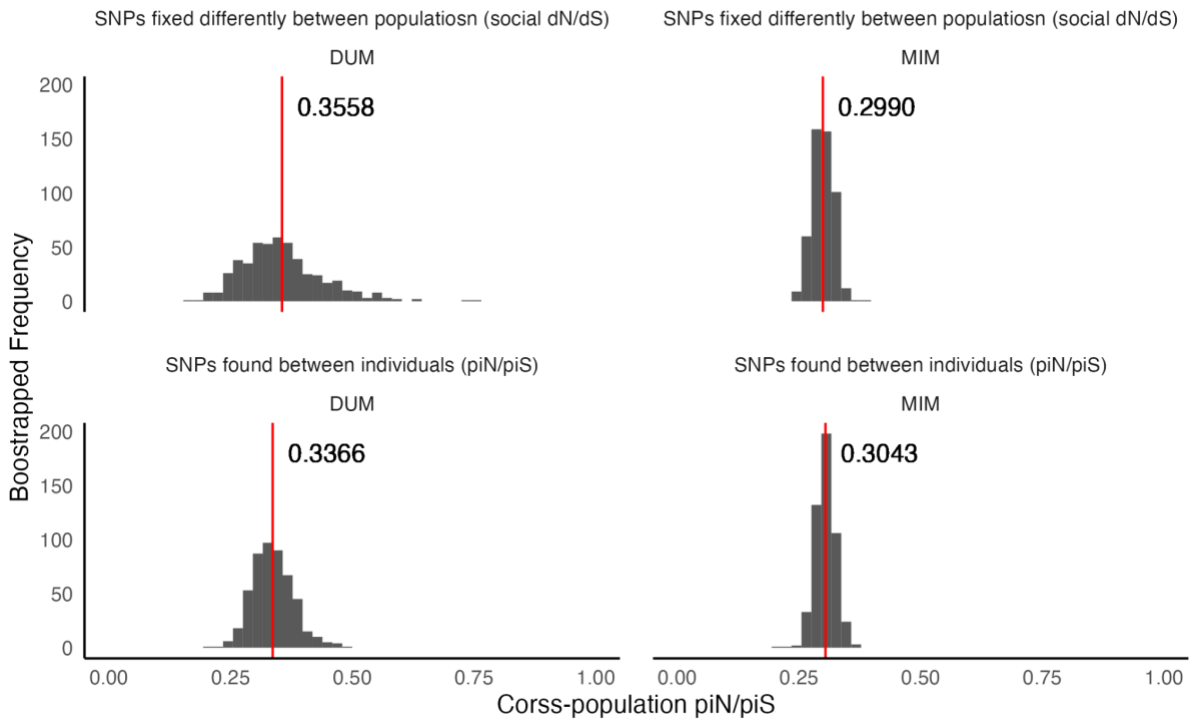


Figure S12. Benchmarking for using pi_N/pi_S as an approximation of "social d_N/d_S" in *S. dumicola* and *S. mimosarum*. The distribution shows the bootstrap pi_N/pi_S estimations of two different subsets of SNPs. 1.We filtered SNPs to analyse only polymorphic sites between an individual from two isolated populations (bottom graphs, pi_N/pi_S shown in Figure 4 and Figure 5) and 2. Using population data, we further filtered SNPs to keep only polymorphic sites fixed differently between two populations (upper graph, "social d_N/d_S"). The point estimation is shown as the vertical red lines and marked with text.

Supplementary Reference

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

Bechsgaard J, Schou MF, Vanthournout B, Hendrickx F, Knudsen B, Settepani V, Schierup MH, Bilde T. 2019. Evidence for Faster X Chromosome Evolution in Spiders. *Mol Biol Evol* **36**: 1281–1293.

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**: lqaa108.

Buchfink B, Reuter K, Drost H-G. 2021. Sensitive protein alignments at tree-of-life scale using

DIAMOND. *Nat Methods* **18**: 366–368.

Cheng H, Concepcion GT, Feng X, Zhang H, Li H. 2021. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat Methods* **18**: 170–175.

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* **6**: 80–92.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. 2021. Twelve years of SAMtools and BCFtools. *Gigascience* **10**. http://dx.doi.org/10.1093/gigascience/giab008.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.

Dudchenko O, Batra SS, Omer AD, Nyquist SK, Hoeger M, Durand NC, Shamim MS, Machol I, Lander ES, Aiden AP, et al. 2017. De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. *Science* **356**: 92–95.

Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, Aiden EL. 2016. Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Syst* **3**: 99–101.

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**: 9451–9457.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* **8**: 1494–1512.

Lemoine F, Gascuel O. 2021. Gotree/Goalign: toolkit and Go API to facilitate the development of phylogenetic workflows. *NAR Genom Bioinform* **3**: lqab075.

Ranwez V, Douzery EJP, Cambon C, Chantret N, Delsuc F. 2018. MACSE v2: Toolkit for the Alignment of Coding Sequences Accounting for Frameshifts and Stop Codons. *Mol Biol Evol* **35**: 2582–2584.

Settepani V, Bechsgaard J, Bilde T. 2014. Low genetic diversity and strong but shallow population differentiation suggests genetic homogenization by metapopulation dynamics in a social spider. *J Evol Biol* **27**: 2850–2855.

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

Tarailo-Graovac M, Chen N. 2009. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr Protoc Bioinformatics* **Chapter 4**: 4.10.1–4.10.14.

Vasimuddin M, Misra S, Li H, Aluru S. 2019. Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. In *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, pp. 314–324.

Wertheim JO, Murrell B, Smith MD, Kosakovsky Pond SL, Scheffler K. 2015. RELAX: detecting

relaxed selection in a phylogenetic framework. *Mol Biol Evol* **32**: 820–832.

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* **24**: 1586–1591.