Title: Genome-wide, genic, and site-specific changes associated with the convergent evolution of sociality in spiders

Authors: Chao Tong[a,*], Leticia Avilés[b], Linda S. Rayor[c], Alexander S. Mikheyev[d,e], Timothy A. Linksvayer[a,f,*]

Affiliations:

[a] Department of Biology, University of Pennsylvania, Philadelphia, PA, 19104, USA

[b] Department of Zoology, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada

[c] Department of Entomology, Cornell University, Ithaca, New York 14853, USA

[d] Ecology and Evolution Unit, Okinawa Institute of Science and Technology, Okinawa, 904-0495, Japan

[e] Evolutionary genomics group, Research School of Biology, Australian National University, Canberra, 0200, Australia

[f] Department of Biological Sciences, Texas Tech University, Lubbock, TX, 79409, USA

[*] Correspondence: Chao Tong, Timothy A. Linksvayer

Email: tongchao1990@gmail.com (C.T.), tlinksvayer@gmail.com (T.A.L.)

C.T. 0000-0001-5202-5507; L.A. 0000-0002-9609-2952; L.S.R. 0000-0002-7132-6281; A.S.M. 0000-0003-4369-1019; T.A.L. 0000-0001-7034-1546

Classification

Major – Biological Sciences; Minor – Evolution

Keywords

comparative genomics; convergent evolution; sociality; spiders

Author Contributions

CT and TAL conceived and designed the study. LA and LSR provided spider samples and expertise. ASM provided reagents and sequencing support. CT generated libraries for RNA sequencing. CT and TAL performed all transcriptomic and comparative genomic analyses and created all figures. CT and TAL wrote the paper, with input from all authors.

This PDF file includes:

Main Text

Figures 1 to 4

Abstract

Sociality is a striking phenotypic innovation that independently evolved dozens of times across animals. Sociogenomic approaches have begun to elucidate the molecular underpinnings of social life in social insects and vertebrates, but the degree to which the convergent evolution of sociality involves convergent molecular evolution remains controversial and largely unknown. Spiders are a powerful system for identifying the genomic causes and consequences of social life because sociality is estimated to have independently evolved 15 times, and each origin likely occurred recently, within the past few million years. To determine if there are statistically supported genomic signatures of protein-coding sequence evolution associated with the convergent evolution of sociality in spiders, we compared the genomes or transcriptomes of 24 spider species that vary in social organization and represent at least seven independent origins of sociality. We identified hundreds of genes that experienced shifts in patterns of molecular evolution during the convergent evolution of sociality, and these genes were enriched for several annotated functions, including neural function, neurogenesis, and behavior, as well as immune function, growth, and metabolism. We also found evidence that directional selection for specific substitutions repeatedly occurred in social species for several genes, in particular the calcium channel gene *TPC1*. Finally, supporting previous results, we found elevated genome-wide rates of molecular evolution in social species, resulting mainly from relaxation of selection. Altogether, we identify genome-wide, genic, and site-specific changes that repeatedly occurred during the evolution of sociality in spiders.

Significance Statement

The transition from solitary to social life is a major transition in evolution that repeatedly occurred, but the genetic underpinnings are largely unknown. To identify genomic changes associated with sociality, we compared the genomes of 24 spider species, representing seven recent and independent origins of sociality. We identified hundreds of genes and many functional classes of genes that tended to experience shifts in molecular evolution, including genes that have previously been implicated in animal social behavior and human behavioral disorders. Our study shows that while the precise genetic details vary, the repeated evolution of sociality in spiders predictably leaves genome-wide signatures and involves sets of genes with conserved functions that may play general roles in the evolution of social behavior.

Main Text

Introduction

A longstanding goal of biology is to identify the genetic causes of major phenotypic innovations such as multicellularity, sociality, and flight (1). When individuals within a population vary for the trait of interest, identifying associations between genetic variants and traits of interest is relatively straightforward with experimental mapping populations or genome-wide association studies (2). However, major phenotypic innovations often only differ between divergent lineages. This makes identifying their genetic underpinnings more challenging because of the difficulty in distinguishing the relevant genetic differences from those unrelated to the phenotype of interest (3,4).

Increasingly, comparative genomic approaches have been used to elucidate the genomic underpinnings of phenotypic innovations, in particular when phenotypes of interest repeatedly evolve in different evolutionary lineages, so-called convergent phenotypic evolution (5-9). If specific genetic changes and genomic signatures are common across independent origins of a phenotype of interest, comparative genomic approaches can provide a powerful means to identify genetic changes that are repeatedly involved, and thus likely to be especially important in phenotypic innovation. At the same time, these studies can begin to answer a second complementary longstanding question: to what degree are the genetic changes underlying independent phenotypic innovation the same, and more generally, how repeatable is the evolutionary process (10,11)?

Sociality is an exemplar phenotypic innovation that convergently evolved repeatedly and sporadically across vertebrates, insects, spiders, and crustaceans (12). Sociality and social behavior more generally are hypothesized to evolve via changes in a set of deeply conserved genes (13-17). However, the degree to which convergent social evolution in diverse lineages consistently involves evolutionary changes in the expression or protein-coding sequence of the same or similar genes remains largely unclear. Many transcriptomic studies have emphasized significantly overlapping sets of genes or molecular pathways underlying the expression of key social phenotypes in different lineages (18-21), but other transcriptomic studies emphasize the importance of lineage-specific genes, or genes with lineage-specific expression (22-24). Social insect studies searching for patterns of molecular evolution associated with the convergent evolution of sociality have also found differing results. For example, a study comparing population genomic signatures of selection in two bee and one wasp species, altogether representing two origins of sociality, found significant overlap in genes experiencing positive selection among all three species (25). In contrast, the largest comparative genomic study of sociality to date, a comparative genomic study of ten bee species representing two independent origins and two independent elaborations of eusociality, found low and insignificant overlap for rapidly evolving genes associated with

social evolution in each lineage (26). The authors concluded that the evolution of sociality in each lineage involved distinct evolutionary routes (26).

One difficulty with comparative genomic studies of sociality is that in most lineages there are relatively few origins of sociality, which typically occurred many millions of years ago, making the genomic signatures of the evolution of sociality difficult to detect (4). For example, social hymenopteran insects (wasps, bees, and ants), and in particular the bees, are often described as ideal study systems because they include the full range of social complexity and multiple origins of sociality (26, 27). However, there are only an estimated total 2-3 independent origins of sociality within the bees, and 7-8 across all social hymenopterans, and these origins occurred between 20 and 100 million years ago (4, 26, 27). Other groups with many independent origins of sociality may complement studies in social insects and help to further identify common genomic signatures of sociality.

Like the social insects, spiders exhibit the full range of social complexity and multiple independent origins of sociality (28-30). Sociality evolved independently in spiders an estimated 15-16 times (31, 32) and each independent origin is thought to be relatively recent, at most a few million years ago (33, 34). Spiders are classified into several categories of social organization, including solitary, subsocial, prolonged subsocial, and social: solitary species live in individual nests, having dispersed from the egg sac soon after hatching (28, 32); subsocial spiders have nests that contain a single mother and up to a few dozen offspring, which may remain together for several instars before dispersing to initiate their own nests (28, 29); prolonged subsocial spiders form colonies containing a single mother and multiple cohorts of offspring that remain in their natal nest until late adolescence or sexual maturity before dispersing independently to found their own colony (29); and social spiders (i.e. non-territorial permanent-social species) form colonies that contain multiple adult females and offspring that remain in the natal nest through maturity, mating with each other to produce new generations that reoccupy the natal nest (28). In social and subsocial species, individuals cooperate in building and maintaining the communal nest, capture prey cooperatively, and share their food. In social species, individuals additionally exhibit communal brood care. Social spider species have additional distinct features, including colonies that grow over multiple generations and in some species can reach colony sizes of tens of thousands of individuals, female-biased sex ratios, high reproductive skew, and high rates of inbreeding (30, 35).

Spider sociality has a twiggy phylogenetic distribution, where the closest relatives of most social species are not social. Several previous studies have compared pairs of social species and their closest nonsocial (i.e. subsocial) relatives within a genus in order to identify genetic consequences of spider sociality and associated shifts in breeding system from outbreeding to inbreeding (33, 36-39). These studies, together

with a preliminary comparative genomic study including two social species from one genus and several solitary species (40), identified putative genomic signatures of the evolution of sociality in spiders, including elevated genome-wide rates of molecular evolution likely caused by relaxed purifying selection (36-40).

Here, we use a comparative genomic approach to determine whether there are common genomic signatures across repeated origins of spider sociality that can provide insight into the genomic causes and consequences of social life. We used 24 total spider species in four lineages that range in social organization and represent an estimated seven independent origins of sociality (34, 41-43). We focused our comparisons on testing whether the convergent evolution of sociality in spiders is associated with the following signatures of convergent molecular evolution: (i) convergent genome-wide patterns of molecular evolution; (ii) convergent shifts in evolutionary rates for specific genes; and (iii) convergent amino acid substitutions in specific genes.

Results

Spider samples

We constructed transcriptomes in parallel from 16 ethanol-stored specimens from species in the genera *Anelosimus*, *Theridion* (Theriidiade), *Stegodyphus* (Eresidae), and members of the family Sparassidae (SI Appendix, Table S1). We complemented these new transcriptomes with an additional two available genomes and six transcriptomes from another eight related species (38, 44-48) (SI Appendix, Table S2). Altogether, our study included 24 total spider species across the range of spider sociality: eight social species, one prolonged subsocial species, five subsocial species, and ten solitary species. Since we were interested in identifying genomic changes associated with sociality, we compared the eight social species (hereafter labeled as “social”) with the 16 species in the remaining categories (hereafter all grouped together as “nonsocial”) (Fig. 1A).

Transcriptome assembly, ortholog identification, and species phylogeny

Following *de novo* assembly and annotation, each of 22 spider transcriptomes had an average of 33,870 transcripts representing the complete or partial protein-coding regions of genes (SI Appendix, Table S2). Of the 1,066 core arthropod orthologous genes in the BUSCO database (49), 85.2% to 92.1% were present in the 22 spider transcriptome assemblies (SI Appendix, Table S2), indicating that all transcriptomes were relatively complete. We also included three available spider genomes from two social *Stegodyphus* species and a solitary outgroup, *Acanthoscurria geniculata* (SI Appendix, Table S2).

We used a reciprocal hit search strategy with Orthofinder (50) and OrthoDB (https://www.orthodb.org/), and identified a total of 7,580 orthologous gene groups (OGGs) shared by all 25 spider species, and 1,618 core single-copy orthologs across all species (SI Appendix, Table S2). Finally, we used coalescent single-copy orthologs to reconstruct a species phylogeny (Fig. 1A, SI Appendix, Fig. S1) that was strongly supported (bootstrap value = 100) for all nodes, and was mostly consistent with previously-published phylogenies (34, 41–43, 51) (Fig. S1).

Social spiders exhibit convergent genome-wide patterns of accelerated molecular evolution that is caused mainly by the relaxation of selection

As a first test of signatures of convergent molecular evolution, we asked whether social branches had different rates of genome-wide molecular evolution (i.e. dN/dS) compared to the nonsocial branches. We estimated the genome-wide ratio of nonsynonymous to synonymous substitution (dN/dS) across the spider phylogeny, using CodeML in PAML 4.7a (52) with concatenated 1,618 single-copy ortholog data. Social branches had higher genome-wide dN/dS that estimated by HyPhy 2.5 (53) than nonsocial branches (likelihood ratio test (LRT), $P < 0.00001$) (Fig. 1B).

Elevated dN/dS can be caused by increased positive selection, relaxed purifying selection, or a combination of both. RELAX (54) quantifies the degree to which shifts in the distribution of dN/dS across individual genes or whole genomes are caused by overall *relaxation* of selection (i.e. weakening of both purifying selection and positive selection, towards neutrality) versus overall *intensification* of selection (i.e. strengthening of both purifying selection and positive selection, away from neutrality). Specifically, RELAX models the distribution of three categories of dN/dS (i.e. positive selection, neutral evolution, purifying selection) across a phylogeny and compares the distributions for foreground (i.e. social) branches to background (i.e. nonsocial) branches and estimates a parameter K that indicates overall relaxation ($K < 1$) or intensification ($K > 1$). We found that the shifts in genome-wide dN/dS between social and nonsocial branches were caused mainly by relaxation of both purifying and positive selection ($K = 0.80$, LRT, $P < 0.001$) (Fig. 1C).

Many genes experience convergent patterns of accelerated or decelerated evolution in social spiders

Next we asked if there was evidence for convergent molecular evolution at the gene level. First we used RELAX (54) to identify genes that experienced relaxed or intensified selection in all social branches relative to all nonsocial branches. Consistent with the genome-wide RELAX results emphasizing relaxation of selection, out of 5,243 single-copy orthologs included in the RELAX analysis, more genes

showed evidence of relaxation of selection (598 genes, K < 1, LRT *P* < 0.05, Fig. 2A and 2B, SI Appendix, Table S3) compared to intensification of selection (275 genes, K > 1, LRT *P* < 0.05, Fig. 2A & 2C, SI Appendix, Table S3). Genes experiencing significant relaxation of selection in social branches were enriched for 57 gene ontology (GO) terms, including many related to neural function, neurogenesis, and cell growth (Fig. 2D, SI Appendix, Table S4), such as neuron projection extension (GO:1990138), regulation of nervous system process (GO:0031644), regulation of neuron differentiation (GO:0045664), regulation of extent of cell growth (GO:0061387), and regulation of cell morphogenesis (GO:0022604). In addition, genes experiencing intensification of selection in social branches were enriched for 71 GO terms, including terms relating to energy metabolic, immune response, nervous system development, reproduction and chemosensory behavior (Fig. 2E, SI Appendix, Table S5), such as NADH dehydrogenase complex assembly (GO:0010257), regulation of protein catabolic process (GO:0042176), defense response to virus (GO:0051607), synapse organization (GO:0050808), peripheral nervous system development (GO:0007422), gonad development (GO:0008406), development of primary sexual characteristics (GO:0045137), chemosensory behavior (GO:0007635) and olfactory behavior (GO:0042048). The pattern of relaxation of selection in social relative to nonsocial branches can be caused by relaxation of purifying selection, relaxation of positive selection, or both, and similarly, intensification of selection can be caused by intensification of purifying selection, intensification of positive selection, or both. Indeed, inspection of the RELAX results for individual genes showed that genes experienced relaxation of selection because of either mainly relaxed purifying selection (Fig. 2D), mainly relaxed positive selection (Fig. 2B), or both (Fig. 2B). Similarly, genes experienced intensification of selection either because of intensified purifying selection (Fig. 2C), intensified positive selection (Fig. 2C), or both (Fig. 2C).

Further, we applied the relative evolutionary rate (RER) test implemented in RERconverge (55) that is specifically designed to test for signatures of convergent molecular evolution underlying convergent phenotypic evolution. This test examines the rate of protein sequence evolution on every branch across the phylogeny for each gene, standardized to the distribution of rates across all genes. Genes for which these standardized rates (RER) are consistently either higher or lower in foreground branches (defined here as the tips for each social species as well as the branch leading to the common ancestor of *A. eximus* and *A. domingo*, which may have also been social (42, 51); see Figure 1A) compared to background branches (i.e. all remaining branches) as experiencing accelerated molecular evolution in social species or in nonsocial species. We used a phylogenetically restricted permutation strategy, dubbed permulations, to assess the statistical significance of genes and enriched GO terms (9, 56). This permulation approach combines phylogenetic simulations and permutations to construct null expectations for *P* values, and has been shown to be unbiased and more conservative than other approaches such as permutations alone (9, 56).

We ran 10,000 permulations and identified genes and enriched GO terms (Fig. 3A) that experienced significant convergent shifts in RERs in social compared to nonsocial branches. Specifically, out of 7,345 single-copy orthologs included in REconverge analysis, we identified 170 genes experiencing significantly accelerated molecular evolution in social branches, namely rapidly evolving genes in social spiders (permulation $P < 0.05$) (Fig. 3A and 3B, SI Appendix, Table S6) and we identified 39 genes experiencing significantly decelerated molecular evolution in social branches, namely rapidly evolving genes in nonsocial spiders (permulation $P < 0.05$) (Fig. 3A & 3X, SI Appendix, Table S6). We found 46 GO terms significantly associated with social acceleration (permulation $P < 0.05$), mainly related to nervous system development and neural function (Fig. 3D), such as cerebral cortex development (GO:0021987), neurotransmitter secretion (GO:0007269), memory (GO:0007613) and visual learning (GO:0008542) (SI Appendix, Table S7). We found 22 GO terms significantly associated with social deceleration (permulation $P < 0.05$) (Fig. 3E), mainly associated with metabolism and DNA repair processes, such as carbohydrate metabolic process (GO:0005975), ATP metabolic process (GO:0046034), DNA repair (GO:0006281) and DNA replication checkpoint (GO:0000076) (SI Appendix, Table S8).

To clarify further what patterns of selection contribute to the acceleration and deceleration of protein evolution in social branches, we examined the overlap of lists of genes identified by RERconverge and RELAX. The lists of social-accelerated genes and genes experiencing relaxed selection in social branches overlapped more than expected by chance (permutation $P = 0.00017$; Fig. S2), indicating that elevated rates of protein evolution for some genes in social branches is caused primarily by relaxation of purifying selection. Similarly, lists of social-decelerated genes and genes experiencing intensified selection in social branches were significantly overlapping (permutation $P = 0.015$; SI Appendix, Fig. S2), indicating that reduced rates of protein evolution for some genes in social branches is caused primarily by intensified purifying selection.

Social spiders harbor convergent amino acid substitutions

We asked if there were any amino acid substitutions consistently associated with the evolution of sociality in spiders. We used FADE in HyPhy 2.5 (53) to identify sites experiencing directional selection towards specific amino acids in social compared to nonsocial branches. After filtering sites that were also identified by FADE as experiencing directional selection in nonsocial relative to social branches, we found 907 sites in 169 genes with evidence (Bayes Factor > 100) for directional selection in social relative to background branches. Since we were most interested in identifying sites that showed a repeated pattern of substitution across the social branches, we further focused on sites that were inferred by FADE to have experienced the same substitution at least four separate times in social branches, resulting in a list of 90

sites in 19 genes (SI Appendix, Table S9). We inspected protein alignments for these 19 genes and found that there were no substitutions that occurred in all social branches and no nonsocial branches. However, site 64 in the ion channel gene, two pore calcium channel protein 1 (*TPC1*) had seven identical inferred substitutions from valine to glycine in six social spider species as well as a single nonsocial spider, the prolonged subsocial *Delena cancerides*, and another two substitutions from valine to alanine in two social *Anelosimus* species, *A. oritoyacu* and *A. guacamayos*, which are thought to have independently evolved sociality (Fig. 4A). Besides *D. cancerides*, all other nonsocial species had the presumed ancestral valine.

As two substitutions at site 64 repeatedly occurred in *TPC1* in social spider branches (Fig. 4A), we asked if these amino acid changes are predicted to have a functional effect. We generated a three-dimensional protein structure homology model of TPC1, consisting of 12 transmembrane helices connected by loop regions, forming a transporter protein-like structure (Fig. 4B). The substrate transportation cavity is surrounded by four transmembrane helices: TM2, TM3, TM4, TM6, TM7 and TM9 (Fig. 4C and 4D). Site 64 is located on the TM4 domain facing the ion channel and close to the periplasm, indicating that change from valine to glycine or alanine at this site (Fig. 4E) might affect the activity of ion transport.

Discussion

We used a comparative genomic approach with 24 spider species, representing seven independent and likely recent origins of sociality, in order to answer a longstanding question: what genetic changes underpin the convergent evolution of sociality? We found the following genomic signatures of convergent phenotypic evolution: (i) Supporting previous results in spiders and insects, we found elevated genome-wide rates of molecular evolution in social species compared to nonsocial species, resulting mainly from relaxation of selection, which is an expected genomic consequence of living in inbred social groups with reduced effective population size. (ii) We identified hundreds of genes that experienced shifts in patterns of molecular evolution during the convergent evolution of sociality. These genes were enriched for Gene Ontology (GO) terms related to neural function, neurogenesis, and behavior, as well as immune function, growth, and metabolism, indicating that changes in these functions are likely important during or following the transition to social life. (iii) We found evidence for specific substitutions that tended to occur in social species, with substitutions at one site in the calcium channel gene *TPC1* found almost exclusively in social species. Overall, our results suggest that patterns of natural selection that lead to, and result from the repeated evolution of sociality in spiders, tend to act on sets of genes with similar functions and also repeatedly lead to genome-wide signatures.

Consistent with several previous studies in social spiders (33, 36-39) and social insects (26, 57, 58), we found that social branches tend to experience elevated genome-wide rates of molecular evolution (i.e. dN/dS) compared to nonsocial branches. Such elevated genome-wide rates of molecular evolution are hypothesized to be caused by reduced effective population size in social species, which results in relaxation of purifying selection and positive selection experienced by genes (33, 36-39). In social spiders, factors causing this reduced effective population size include inbreeding, female-biased sex ratios, and reproductive skew. Supporting this hypothesis, we found that the genome-wide pattern was primarily driven by the relaxation of both purifying selection and positive selection (Fig. 1C), so that genes in social branches tend to experience more relaxed, neutral evolution when compared to genes in nonsocial branches. This genome-wide pattern is likely to be a longer-term *consequence* of the switch from outbreeding to inbreeding, and to female-biased sex ratios, which likely occurred after the origin of sociality. There was also variation for genome-wide dN/dS that did not depend on sociality (Fig. 1A), which might be explained by lineage-specific differences in effective population size or other factors. Similarly, in a previous study in hymenopteran insects, social branches experienced elevated genome-wide dN/dS compared to nonsocial branches, but most variation existed between lineages, with bees -- regardless social organization -- having the highest dN/dS (58).

We identified approximately 200 genes that tended to experience convergent shifts in evolutionary rates during the transition to social life (Fig. 2A). Such gene-wide shifts could be caused by changes in the patterns of purifying selection, positive selection or both (Fig. 2D). We separately identified hundreds of genes that experienced an average relaxation or intensification of selection during the transition to social life (Fig. 2C). The list of genes experiencing *elevated* rates of protein evolution significantly overlapped with the list of genes experiencing *relaxation* of selection (Fig. S2), indicating that the pattern of gene-wide elevated protein evolution that we detected is often caused by the relaxation of purifying selection during the transition to sociality. Similarly, the list of genes experiencing *decreased* rates of protein evolution significantly overlapped with the list of genes experiencing *intensification* of selection (Fig. S2), indicating that gene-wide decreased protein evolution is largely caused by intensified purifying selection during the transition to sociality.

The genes that we identified as experiencing convergent patterns of molecular evolution were not randomly distributed with respect to annotated gene function, as would be expected if the gene-wide patterns were driven only by genome-wide relaxation of selection associated with reduced effective population size. Instead, these genes were enriched for many specific annotated functional categories, notably those related to neural function, neurogenesis, and behavior, as well as immune function, metabolism, development, and reproduction (Fig. 2 & 3). We can infer that genes and molecular pathways with these functions likely played important roles in the evolution of sociality, either in initial

genetic changes directly associated with the evolution of social traits and the transition to social life, or genetic changes that were later consequences of social life or the elaboration of social phenotypes.

Genes that experienced shifts in patterns of molecular evolution during the transition to sociality were consistently enriched for functions relating to neural function, neurogenesis, and behavior (SI Appendix, Table S4, S5, S6), indicating that changes in these functions were likely particularly important. Genes experiencing relaxation of selection in social branches were enriched for many GO terms associated with neural function or neural development, including neurotransmitter receptor transport and regulation of neuron differentiation (Fig. 2B). Similarly, genes tending to experience elevated evolutionary rates in social branches were enriched with functional annotations such as cerebral development, neurotransmitter secretion, visual learning, and memory (Fig. 3B). Genes experiencing intensified selection were also enriched for synapse organization and peripheral nervous system development, as well as behavior-related terms such as olfactory behavior and negative regulation of locomotion (Fig. 2C). A fundamental phenotypic difference between social and nonsocial species is the extent of social interactions with conspecifics, which would have required increased behavioral tolerance towards conspecifics as an important preliminary change necessary for the evolution of cooperative group living in social spiders (35, 59, 60) and other animals (61). Furthermore, shifts in behavioral development and the timing of social behaviors, reproductive behavior, and dispersal behaviors, are also hypothesized to be critical in the evolution of sociality in spiders (30, 35) and insects (15). Finally, changes in the size of certain brain regions and sensory organs are hypothesized to be important for detecting and processing social signals and social information (62). The changes that we detected in genes associated with neural function, neurogenesis, and behavior may underlie changes in these important behavioral and neural phenotypes.

Genes involved in neurotransmitter secretion and neurotransmitter receptor transport stand out as interesting candidates influencing the evolution of spider social behavior because many genes annotated with these functions experienced significant shifts in evolutionary rates in social spiders and have also previously been implicated in the regulation of social behaviors in other animals (SI Appendix, Table S6). For example, the gene *adenosine A_{2A} receptor* (*adora2a*), which experienced relaxed selection in social spiders, regulates glutamate and dopamine release and influences motivation for social interactions in mice (63, 64), and variation in *adora2a* has been associated with mood and anxiety disorders in humans (65, 66). The gene *dystrophin, isoforms A/C/F/G/H* (*dys*), which experienced accelerated evolution in social spiders, affects social behavior, communication, and synaptic plasticity in mice (67, 68), and mutations in *dys* are associated with autism and intellectual disability in humans (68). The gene *osmotic avoidance abnormal protein 3* (*osm-3*), which experienced relaxed selection in social spiders, regulates social feeding and aggregation behavior in nematodes (69). The gene *calsyntenin-1* (*clstn1*), which

experienced relaxed selection and accelerated evolution in social spiders, regulates synapse formation and synapse plasticity in mice (70), is essential for chemotaxis learning in nematodes (71), and is a candidate for Fragile X syndrome and autism in humans (70, 72). Finally, the genes *synaptosomal-associated protein 25kDa* (*snap-25*), *synaptobrevin*, and *syntaxin-1A* (*stx1a*) interact to regulate the secretion of neurotransmitters and neuromodulators (73). Each of these has been implicated in the expression of social behavior in various animals (74-76), and also implicated in abnormal behavioral phenotypes in humans, including autism and schizophrenia (77, 78). Each of these experienced significant relaxed or accelerated rates of molecular evolution in social spiders, except for *stx1a* (RERconverge permulation $P = 0.06$). Further studies will be necessary to validate the function and mechanism of these candidates influencing spider social behavior.

Besides critical changes in behavioral and neural phenotypes, transitions to social living are also associated with several other phenotypic changes. Social living is hypothesized to be associated with increased disease pressure, although researchers have found mixed support for genetic changes in immune-related genes in social insects compared to solitary insects (79-81). Genes experiencing intensified selection in social spiders were enriched for several immune-related GO terms such as defense response to virus and defense response to other organisms, which is consistent with social spiders experiencing increased disease pressure. Transitions to social living in spiders are also thought to be associated with shifts in development, including increasing paedomorphosis, and a trend towards smaller brood sizes (35) and smaller body size (82). Genes experiencing relaxed selection in social spiders were enriched for GO terms including regulation of cell size and developmental growth, which could underlie these phenotypic changes. Finally, another phenotypic consequence of transitions to sociality in spiders, which may occur over longer time scales, is the change of breeding system from outbreeding to inbreeding and the evolution of female-biased sex ratios (30, 35). Genes experiencing intensified selection were enriched for GO terms including gonad development and male mating behavior, which could be associated with these shifts in the breeding system. Metabolism-related genes and gene functions have also frequently been implicated in the expression of various types of social behavior in many animals (20, 83), and comparative genomic studies in social insects have identified metabolism-associated genes as experiencing accelerated evolution in eusocial bees (84) and ants (85). Consistent with these previous studies, we found that genes experiencing accelerated evolution and intensified selection in spiders were also enriched for metabolism-related functions (e.g., the GO terms ATP metabolic process and regulation of cellular protein catabolic process).

In addition to genome-wide and gene-wide signatures, we also identified some signatures of site-specific convergent molecular evolution. Specifically, 90 sites in 19 genes showed evidence for directional selection for specific amino acids in social relative to nonsocial species, and genes with such signatures

of specific substitutions were enriched for several GO terms related to nervous system development and synapse organization. However, inspection of the amino acid alignments for these genes showed that amino acid substitutions that occurred repeatedly in social branches also tended to be found in one or more nonsocial branches. An amino acid substitution at position 64 in the calcium channel gene *two pore calcium channel protein 1* (*TPC1*) was found in all eight social species and just one nonsocial species. Interestingly, the only nonsocial species that also had a substitution at this position was *Delena cancerides*, which stands as being relatively more socially complex: *D. cancerides* is classified as prolonged subsocial because individuals spend most of their lives living in social groups with their mother and siblings, while the other nonsocial species in our study are classified as solitary or subsocial. Homology modeling of the protein structure of *TPC1* indicated that position 64 is located in a transmembrane domain facing the ion channel and close to the periplasm and may affect ion transport activity. Ion channel genes are widely expressed in the nervous system and influence a variety of behavioral phenotypes (86), including social behaviors such as depression-related behavior (87), acquired fear-related behavior (88), and courtship behavior (89). Our finding suggests that the amino acid substitution at position 64 in TPC1 across social species might contribute to a key social trait, and this gene and specific position is an interesting candidate for future functional tests.

We used phylogenetic comparative approaches to identify genome-wide, gene-wide, and site-specific signatures of convergent molecular evolution. We identified hundreds of genes that tended to experience different rates of molecular evolution in the replicate social branches compared to background nonsocial branches, and these genes were enriched for certain functions. However, we did not identify genes or sites within genes that *always* experienced evolutionary shifts in each of the seven branches that independently evolved sociality. For example, inspection of patterns of gene-wide relative evolutionary rates (e.g., Fig. 3D, 3E) shows that genes with a relatively strong association between social status and evolutionary rate still showed a lot of variation within each category. This means that we identified certain genes and functions that tended to experience shifts in evolutionary rates, but the exact details differed somewhat between the seven independent origins of sociality. Thus, specific substitutions or changes in specific genes are not required for the evolution of spider sociality, but changes do tend to occur in certain sets of genes or molecular pathways. Put another way, the strong statistical signatures across seven independent origins of sociality indicate that social spiders tend to travel similar evolutionary paths in terms of the functional genetic changes and genome-wide consequences, but the precise genetic details do vary. In practice, these results also emphasize that for complex phenotypes such as sociality, datasets representing many independent origins may be necessary to have the statistical power to detect these signatures. When samples representing multiple independent origins are not available, lineage-specific processes such as lineage-specific adaptations or historical contingency can be confounded with genetic changes that are actually associated with the phenotype of interest (3, 4, 40, 90).

Our study adds to the growing list of studies that use comparative genomic approaches to elucidate the molecular underpinnings of convergent evolution and the degree to which similar phenotypic outcomes result from the same or similar genetic changes (5-9). Collectively, these studies show that some convergent phenotypes, especially those involving relatively simple physiological adaptation, such as the evolution of antibiotic resistance, pesticide resistance, or resistance to neurotoxins (91), often predictably involve the same substitutions or changes in one family of genes. Other phenotypic changes that are relatively more complex, such as the evolution of sociality, still have identifiable genomic signatures across independent origins, but these signatures are not as highly predictable in the sense that they always involve the same substitution or changes in the same genes, but instead involve changes in genes with similar functions.

Finally, our study emphasizes the potential for ethanol-preserved specimens to be used in transcriptome-based comparative genomics research. Increasingly, ethanol-preserved specimens from collections are being used for DNA-based phylogenomic and comparative genomic studies (92), but RNA-based studies have usually been restricted to samples collected in liquid nitrogen or RNA later, and stored at -80 C (93). Our study builds on previous studies that successfully extracted RNA from ethanol-stored samples and subsequently conducted transcriptome-based phylogenomics (94, 95). We extracted RNA from spider specimens that had been preserved in ethanol at room temperature for up to 11 years (SI Appendix, Table S1), and generated relatively high quality RNA sequence data for subsequent comparative molecular evolution analyses (SI Appendix, Table S2). Collectively, our study highlights the heretofore unappreciated potential of transcriptomics using ethanol-preserved specimens for comparative genomics and phylogenomics analyses, especially in species with large genomes, such as spiders, where sequencing and assembling whole genomes remains challenging.

Materials and Methods

Sample collection

We collected spider specimens of 16 species that range in social organization, including social, prolonged subsocial, subsocial and solitary, and stored the samples in 95% ethanol (Decon Labs, USA). Specifically, this set of spider specimens includes five social and three subsocial species from the genus *Anelosimus*, one social and four solitary species from the genus *Theridion*, and one prolonged subsocial and two solitary species from the family Sparassidae (Fig.1A, SI Appendix, Table S1). Full details of this as well as other materials and methods are provided in the SI Appendix.

Acknowledgements

We would like to thank Drs. Sarah Rose, Marc Milne, and Jonathan Pruitt for kindly providing ethanol-preserved spider specimens. This work was funded by the National Institutes of Health Grant GM115509 to TAL.

References

1. A. Wagner, The molecular origins of evolutionary innovations. *Trends Genet.* 27, 397–410 (2011).
2. T. F. C. Mackay, E. A. Stone, J. F. Ayroles, The genetics of quantitative traits: challenges and prospects. *Nat. Rev. Genet*. 10, 565–577 (2009).
3. W. Enard, The Molecular Basis of Human Brain Evolution. *Curr. Biol.* 26, R1109–R1117 (2016).
4. T. A. Linksvayer, B. R. Johnson, Re-thinking the social ladder approach for elucidating the evolution and molecular basis of insect societies. *Curr Opin Insect Sci* 34, 123–129 (2019).
5. Y. Zhen, M. L. Aardema, E. M. Medina, M. Schumer, P. Andolfatto, Parallel molecular evolution in an herbivore community. *Science* 337, 1634–1637 (2012).
6. J. Parker, et al., Genome-wide signatures of convergent evolution in echolocating mammals. *Nature* 502, 228–231 (2013).
7. R. Partha, et al., Subterranean mammals show convergent regression in ocular genes and enhancers, along with adaptation to tunneling. *Elife* 6 (2017).
8. C. R. Feldman, E. D. Brodie Jr, E. D. Brodie 3rd, M. E. Pfrender, Constraint shapes convergence in tetrodotoxin-resistant sodium channels of snakes. *Proc. Natl. Acad. Sci. U. S. A.* 109, 4556–4561 (2012).
9. A. Kowalczyk, R. Partha, N. L. Clark, M. Chikina, Pan-mammalian analysis of molecular constraints underlying extended lifespan. *Elife* 9 (2020).
10. E. B. Rosenblum, C. E. Parent, E. E. Brandt, The molecular basis of phenotypic convergence. *Annu. Rev. Ecol. Evol. Syst.* 45, 203-226 (2014).
11. D. I. Bolnick, R. D. H. Barrett, K. B. Oke, D. J. Rennison, Y. E. Stuart, (Non) parallel evolution. *Annu. Rev. Ecol. Evol. Syst.* 49, 303–330 (2018).
12. D. R. Rubenstein, P. Abbot, Comparative Social Evolution (Cambridge University Press, 2017).
13. A. L. Toth, G. E. Robinson, Evo-devo and the evolution of social behavior. *Trends Genet.* 23, 334–341 (2007).
14. G. V. Amdam, A. Csondes, M. K. Fondrk, R. E. Page Jr, Complex social behaviour derived from maternal reproductive traits. *Nature* 439, 76–78 (2006).
15. T. A. Linksvayer, M. J. Wade, The evolutionary origin and elaboration of sociality in the aculeate Hymenoptera: maternal effects, sib-social effects, and heterochrony. *Q. Rev. Biol.* 80, 317–336 (2005).
16. L. A. O'Connell, H. A. Hofmann, The vertebrate mesolimbic reward system and social behavior network: a comparative synthesis*. J. Comp. Neurol.* 519, 3599–3639 (2011).
17. G. E. Robinson, C. M. Grozinger, C. W. Whitfield, Sociogenomics: social life in molecular terms. *Nat. Rev. Genet.* 6, 257–270 (2005).

18. A. J. Berens, J. H. Hunt, A. L. Toth, Comparative transcriptomics of convergent evolution: different genes but conserved pathways underlie caste phenotypes across lineages of eusocial insects. *Mol. Biol. Evol.* 32, 690–703 (2015).

19. R. L. Young, et al., Conserved transcriptomic profiles underpin monogamy across vertebrates. *Proc. Natl. Acad. Sci. U. S. A.* 116, 1331–1336 (2019).

20. C. C. Rittschof, et al., Neuromolecular responses to social challenge: common mechanisms across mouse, stickleback fish, and honey bee. *Proc. Natl. Acad. Sci. U. S. A.* 111, 17929–17934 (2014).

21. M. R. Warner, L. Qiu, M. J. Holmes, A. S. Mikheyev, T. A. Linksvayer, Convergent eusocial evolution is based on a shared reproductive groundplan plus lineage-specific plastic genes. *Nat. Commun.* 10, 2651 (2019).

22. B. R. Johnson, N. D. Tsutsui, Taxonomically restricted genes are associated with the evolution of sociality in the honey bee. *BMC Genomics* 12, 164 (2011).

23. W. C. Jasper, et al., Large-scale coding sequence change underlies the evolution of postdevelopmental novelty in honey bees. *Mol. Biol. Evol.* 33, 1379 (2016).

24. S. Sumner, The importance of genomic novelty in social evolution. *Mol. Ecol.* 23, 26–28 (2014).

25. K. A. Dogantzis, et al., Insects with similar social complexity show convergent patterns of adaptive molecular evolution. *Sci. Rep.* 8, 10388 (2018).

26. K. M. Kapheim, et al., Genomic signatures of evolutionary transitions from solitary to group living. *Science* 348, 1139–1143 (2015).

27. S. D. Kocher, R. J. Paxton, Comparative methods offer powerful insights into social evolution in bees. *Apidologie* 45, 289–305 (2014).

28. L. Avilés, J. Guevara, “Sociality in Spiders” in Comparative Social Evolution, R. Rubenstein, P. Abbot, Eds. (Cambridge University Press, 2017), pp. 188–223.

29. E. C. Yip, L. S. Rayor, Maternal care and subsocial behaviour in spiders. Biol. Rev. Camb. Philos. Soc. 89, 427–449 (2014).

30. Y. Lubin, T. Bilde, “The Evolution of Sociality in Spiders” in Advances in the Study of Behavior, (Academic Press, 2007), pp. 83–145.

31. I. Agnarsson, L. Avilés, J. A. Coddington, W. P. Maddison, Sociality in theridiid spiders: repeated origins of an evolutionary dead end. *Evolution* 60, 2342–2351 (2006).

32. L. Avilés, “Social Spiders” in Encyclopedia of Social Insects, C. K. Starr, Ed. (Springer, 2020), pp. 1–10.

33. I. Agnarsson, L. Avilés, W. P. Maddison, Loss of genetic variability in social spiders: genetic and phylogenetic consequences of population subdivision and inbreeding. *J. Evol. Biol.* 26, 27–37 (2013).

34. J. Johannesen, Y. Lubin, D. R. Smith, T. Bilde, J. M. Schneider, The age and evolution of sociality in Stegodyphus spiders: a molecular phylogenetic perspective. *Proc. Biol. Sci.* 274, 231–237 (2007).

35. L. Avilés, 23• Causes and consequences of cooperation and permanent-sociality in spiders. The evolution of social behavior in insects and arachnids, 476–498 (1997).

36. V. Settepani, et al., Evolution of sociality in spiders leads to depleted genomic diversity at both population and species levels. *Mol. Ecol.* 26, 4197–4210 (2017).

37. V. Settepani, J. Bechsgaard, T. Bilde, Phylogenetic analysis suggests that sociality is associated with reduced effectiveness of selection. *Ecol. Evol.* 6, 469–477 (2016).

38. J. Bechsgaard, et al., Evidence for faster X chromosome evolution in spiders. *Mol. Biol. Evol.* 36, 1281–1293 (2019).

39. T. M. Mattila, J. S. Bechsgaard, T. T. Hansen, M. H. Schierup, T. Bilde, Orthologous genes identified by transcriptome sequencing in the spider genus Stegodyphus. *BMC Genomics* 13, 70 (2012).

40. C. Tong, G. M. Najm, N. Pinter-Wollman, J. N. Pruitt, T. A. Linksvayer, Comparative Genomics Identifies Putative Signatures of Sociality in Spiders. *Genome Biol. Evol.* 12, 122–133 (2020).

41. M. A. Arnedo, I. Agnarsson, R. G. Gillespie, Molecular insights into the phylogenetic structure of the spider genus Theridion (Araneae, Theridiidae) and the origin of the Hawaiian Theridion-like fauna. *Zool. Scr.* 36, 337–352 (2007).

42. Y. Luo, et al., Global Diversification of Anelosimus Spiders Driven by Long Distance Overwater Dispersal and Neogene Climate Oscillations. *Syst. Biol.* 69 (6), 1122-1136 (2020).

43. I. Agnarsson, L. S. Rayor, A molecular phylogeny of the Australian huntsman spiders (Sparassidae, Deleninae): implications for taxonomy and social behaviour. *Mol. Phylogenet. Evol.* 69, 895–905 (2013).

44. K. W. Sanggaard, et al., Spider genomes provide insight into composition and evolution of venom and silk. *Nat. Commun.* 5, 3765 (2014).

45. P. J. P. Croucher, M. S. Brewer, C. J. Winchell, G. S. Oxford, R. G. Gillespie, De novo characterization of the gene-rich transcriptomes of two color-polymorphic spiders, Theridion grallator and T. californicum (Araneae: Theridiidae), with special reference to pigment genes. *BMC Genomics* 14, 862 (2013).

46. L. Shao, S. Li, Early Cretaceous greenhouse pumped higher taxa diversification in spiders. *Mol. Phylogenet. Evol.* 127, 146–155 (2018).

47. R. Fernández, et al., Phylogenomics, Diversification Dynamics, and Comparative Transcriptomics across the Spider Tree of Life. *Curr. Biol.* 28, 2190–2193 (2018).

48. S. Liu, A. Aageaard, J. Bechsgaard, T. Bilde, DNA Methylation Patterns in the Social Spider, Stegodyphus dumicola. *Genes* 10 (2019).

49. F. A. Simão, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, E. M. Zdobnov, BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212 (2015).

50. D. M. Emms, S. Kelly, OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 16, 157 (2015).

51. I. Agnarsson, W. P. Maddison, L. Avilés, The phylogeny of the social Anelosimus spiders (Araneae: Theridiidae) inferred from six molecular loci and morphology. *Mol. Phylogenet. Evol.* 43, 833–851 (2007).

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

53. S. L. Kosakovsky Pond, et al., HyPhy 2.5—A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. *Mol. Biol. Evol.* 37, 295–299 (2020).

54. J. O. Wertheim, B. Murrell, M. D. Smith, S. L. Kosakovsky Pond, K. Scheffler, RELAX: detecting relaxed selection in a phylogenetic framework. *Mol. Biol. Evol.* 32, 820–832 (2015).

55. A. Kowalczyk, et al., RERconverge: an R package for associating evolutionary rates with convergent traits. *Bioinformatics* 35, 4815–4817 (2019).

56. E. Saputra, A. Kowalczyk, L. Cusick, N. Clark, M. Chikina, Phylogenetic Permulations: a statistically rigorous approach to measure confidence in associations between phenotypes and genetic elements in a phylogenetic context. Cold Spring Harbor Laboratory, 2020.10.14.338608 (2020).

57. J. Romiguier, et al., Population genomics of eusocial insects: the costs of a vertebrate-like effective population size. *J. Evol. Biol.* 27, 593–603 (2014).

58. A. Weyna, J. Romiguier, Relaxation of purifying selection suggests low effective population size in eusocial Hymenoptera and pollinating bees. *bioRxiv,* 2020.04.14.038893 (2020).

59. C. Viera, I. Agnarsson, "Parental Care and Sociality" in Behaviour and Ecology of Spiders: Contributions from the Neotropical Region, C. Viera, M. O. Gonzaga, Eds. (Springer International Publishing, 2017), pp. 351–381.

60. E. J. Kullmann, Evolution of social behavior in spiders (Araneae; Eresidae and Theridiidae). Am. Zool. 12, 419–426 (1972).

61. N. Lin, C. D. Michener, Evolution of Sociality in Insects. *Q. Rev. Biol.* 47, 131–159 (1972).

62. B. Wittwer, et al., Solitary bees reduce investment in communication compared with their social relatives. *Proc. Natl. Acad. Sci. U. S. A.* 114, 6569–6574 (2017).

63. L. López-Cruz, et al., Adenosine A2A receptor deletion affects social behaviors and anxiety in mice: Involvement of anterior cingulate cortex and amygdala. *Behav. Brain Res.* 321, 8–17 (2017).

64. L. López-Cruz, et al., Ethanol and Caffeine Effects on Social Interaction and Recognition in Mice: Involvement of Adenosine A2A and A1 Receptors. *Front. Behav. Neurosci.* 10, 206 (2016).

65. S. Oliveira, et al., Impact of genetic variations in ADORA2A gene on depression and symptoms: a cross-sectional population-based study. *Purinergic Signal.* 15, 37–44 (2019).

66. D. van Calker, K. Biber, K. Domschke, T. Serchov, The role of adenosine receptors in mood and anxiety disorders. *J. Neurochem.* 151, 11–27 (2019).

67. R. Miranda, et al., Altered social behavior and ultrasonic communication in the dystrophin-deficient mdx mouse model of Duchenne muscular dystrophy. *Mol. Autism* 6, 60 (2015).

68. F. Daoud, et al., Role of mental retardation-associated dystrophin-gene product Dp71 in excitatory synapse organization, synaptic plasticity and behavioral functions. *PLoS One* 4, e6574 (2008).

69. M. de Bono, D. M. Tobin, M. W. Davis, L. Avery, C. I. Bargmann, Social feeding in Caenorhabditis elegans is induced by neurons that detect aversive stimuli. *Nature* 419, 899–903 (2002).

70. K. Cheng, et al., Calsyntenin-1 Negatively Regulates ICAM5 Accumulation in Postsynaptic Membrane and Influences Dendritic Spine Maturation in a Mouse Model of Fragile X Syndrome. *Front. Neurosci.* 13, 1098 (2019).

71. D. D. Ikeda, et al., CASY-1, an ortholog of calsyntenins/alcadeins, is essential for learning in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. U. S. A.* 105, 5260–5265 (2008).

72. S. V. Ranneva, K. S. Pavlov, A. V. Gromova, T. G. Amstislavskaya, T. V. Lipina, Features of emotional and social behavioral phenotypes of calsyntenin2 knockout mice. *Behav. Brain Res.* 332, 343–354 (2017).

73. N. A. Ramakrishnan, M. J. Drescher, D. G. Drescher, Direct interaction of otoferlin with syntaxin 1A, SNAP-25, and the L-type voltage-gated calcium channel Cav1.3. *J. Biol. Chem.* 284, 1364–1372 (2009).

74. I. Corradini, et al., Epileptiform activity and cognitive deficits in SNAP-25+/- mice are normalized by antiepileptic drugs. *Cereb. Cortex* 24, 364–376 (2014).

75. T. Fujiwara, M. Sanada, T. Kofuji, K. Akagawa, Unusual social behavior in HPC-1/syntaxin1A knockout mice is caused by disruption of the oxytocinergic neural system. *J. Neurochem.* 138, 117–123 (2016).

76. S. D. Kocher, et al., The genetic basis of a social polymorphism in halictid bees. *Nat. Commun*. 9, 4338 (2018).

77. D. Braida, et al., Association between SNAP-25 gene polymorphisms and cognition in autism: functional consequences and potential therapeutic strategies. *Transl. Psychiatry* 5, e500 (2015).

78. I. Corradini, C. Verderio, M. Sala, M. C. Wilson, M. Matteoli, SNAP-25 in neuropsychiatric disorders. *Ann. N. Y. Acad. Sci.* 1152, 93–99 (2009).

79. K. Meusemann, J. Korb, M. Schughart, F. Staubach, No evidence for single-copy immune-gene specific signals of selection in termites. *Front. Ecol. Evol.* 8 (2020).

80. L. Viljakainen, et al., Rapid evolution of immune proteins in social insects. *Mol. Biol. Evol.* 26, 1791–1801 (2009).

81. S. Otani, N. Bos, S. H. Yek, Transitional Complexity of Social Insect Immunity. *Frontiers in Ecology and Evolution* 4, 69 (2016).

82. L. Grinsted, et al., Prey to predator body size ratio in the evolution of cooperative hunting—a social spider test case. *Dev. Genes Evol.* 230, 173–184 (2020).

83. C. C. Rittschof, G. E. Robinson, Behavioral Genetic Toolkits: Toward the Evolutionary Origins of Complex Phenotypes. *Curr. Top. Dev. Biol.* 119, 157–204 (2016).

84. S. H. Woodard, et al., Genes involved in convergent evolution of eusociality in bees. *Proc. Natl. Acad. Sci. U. S. A.* 108, 7472–7477 (2011).

85. J. Roux, et al., Patterns of positive selection in seven ant genomes. *Mol. Biol. Evol.* 31, 1661–1685 (2014).

86. N. S. Atkinson, et al., Molecular separation of two behavioral phenotypes by a mutation affecting the promoters of a Ca-activated K channel. *J. Neurosci.* 20, 2988–2993 (2000).

87. M. W. Coryell, et al., Acid-sensing ion channel-1a in the amygdala, a novel therapeutic target in depression-related behavior. *J. Neurosci.* 29, 5381–5388 (2009).

88. J. A. Wemmie, et al., Overexpression of acid-sensing ion channel 1a in transgenic mice increases acquired fear-related behavior. *Proc. Natl. Acad. Sci. U. S. A.* 101, 3621–3626 (2004).

89. Y. Ding, A. Berrocal, T. Morita, K. D. Longden, D. L. Stern, Natural courtship song variation caused by an intronic retroelement in an ion channel gene. *Nature* 536, 329–332 (2016).

90. F. Delsuc, M.-K. Tilak, Naked but not Hairless: the pitfalls of analyses of molecular adaptation based on few genome sequence comparisons. *Genome Biol. Evol.* 7, 768–774 (2015).

91. J. W. McGlothlin, et al., Historical Contingency in a Multigene Family Facilitates Adaptive Evolution of Toxin Resistance. *Curr. Biol.* 26, 1616–1621 (2016).

92. A. E. Z. Short, T. Dikow, C. S. Moreau, Entomological Collections in the Age of Big Data. *Annu. Rev. Entomol.* 63, 513–530 (2018).

93. Z. Wang, M. Gerstein, M. Snyder, RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57–63 (2009).

94. H. M. Gough, J. M. Allen, E. F. A. Toussaint, C. G. Storer, A. Y. Kawahara, Transcriptomics illuminate the phylogenetic backbone of tiger beetles. *Biol. J. Linn. Soc. Lond.* 129, 740–751 (2020).

95. A. L. Bazinet, M. P. Cummings, K. T. Mitter, C. W. Mitter, Can RNA-Seq resolve the rapid radiation of advanced moths and butterflies (Hexapoda: Lepidoptera: Apoditrysia)? An exploratory study. *PLoS One* 8, e82615 (2013).

Figures

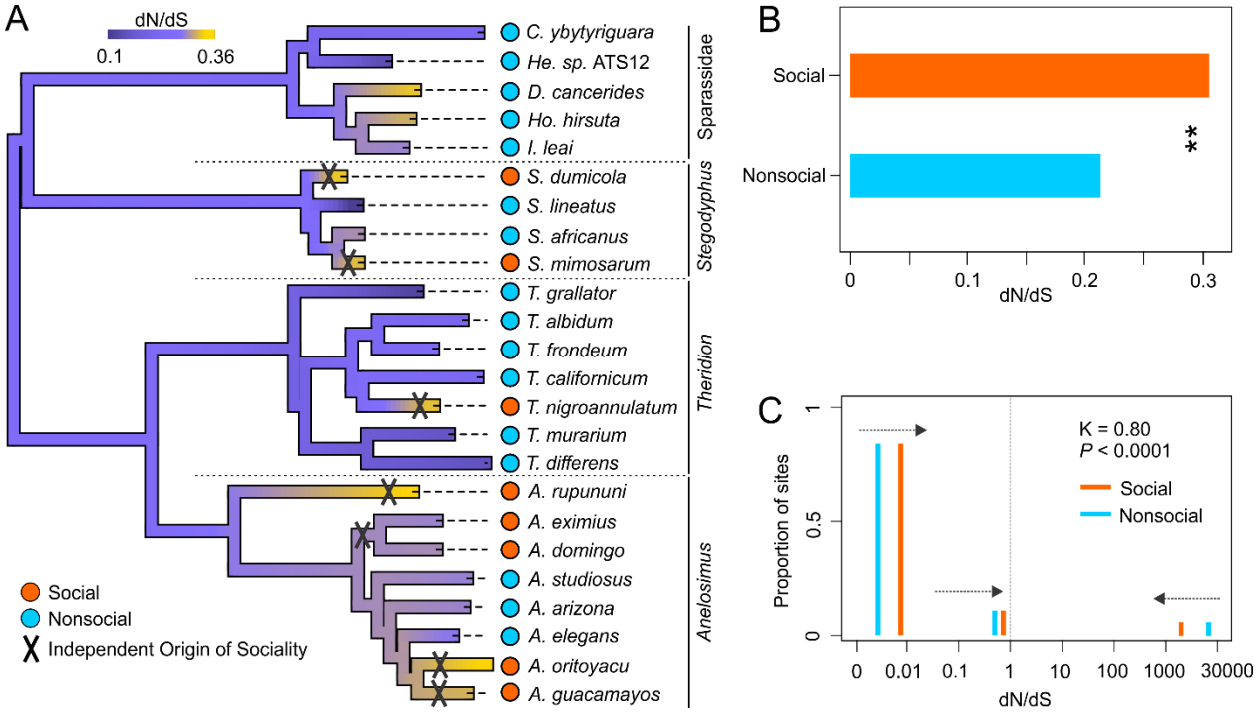
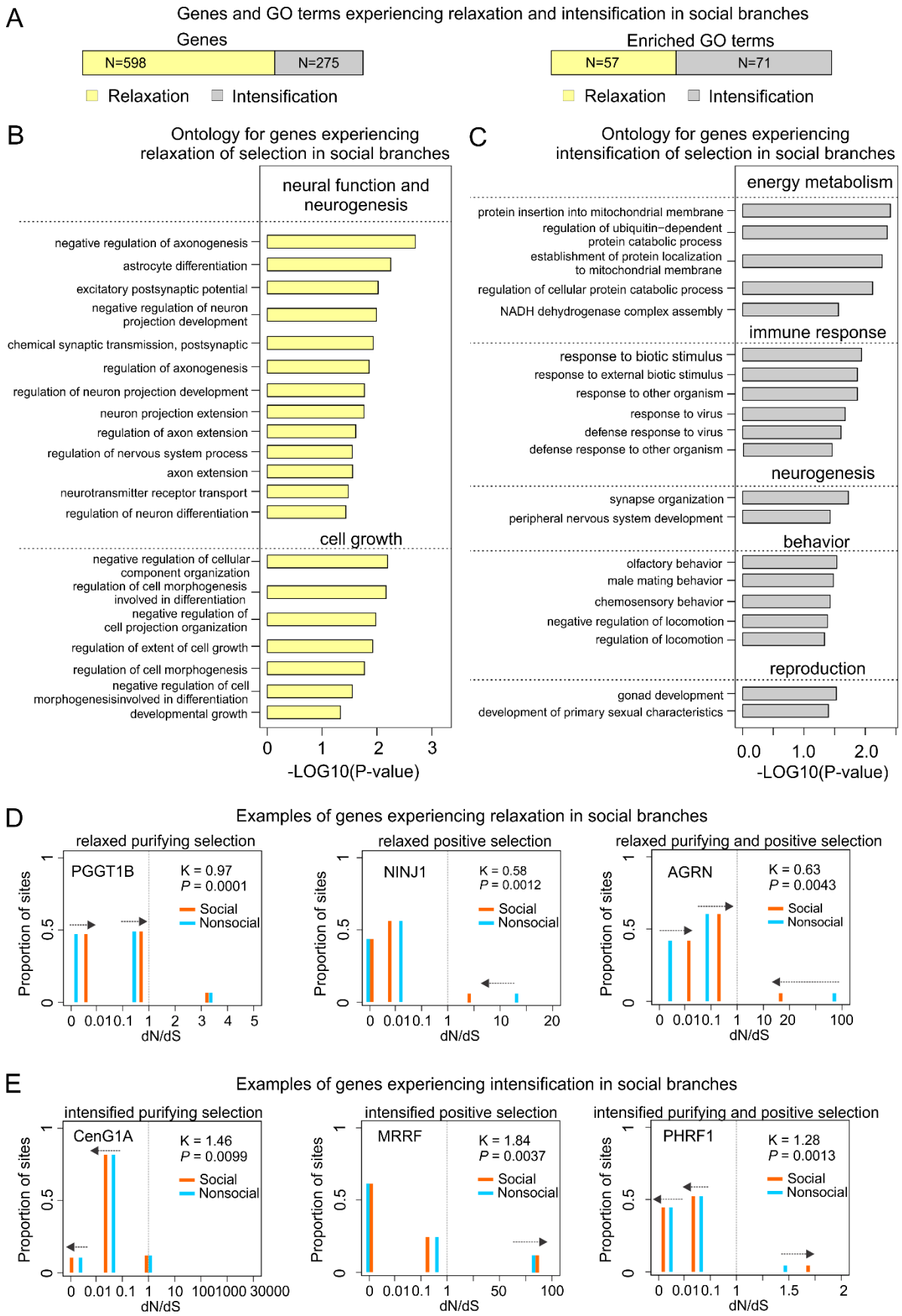


Figure 1. Spiders experienced elevated genome-wide molecular evolution during the transition to sociality, driven by the relaxation of purifying and positive selection. (A) phylogram showing the phylogenetic relationships among the 24 species included in the study. Branch color shows estimated genome-wide rate of molecular evolution (dN/dS) and branch length is proportional to sequence divergence. Social and nonsocial spider taxa are indicated by orange or sky blue dots, respectively. X's indicate independent origins of sociality as inferred in previous studies (31, 32). (B) barplot depicting the genome-wide dN/dS ratios for social and nonsocial branches, estimated by HyPhy. ** $P < 0.001$ (likelihood ratio test, LRT). (C) plot showing the pattern of selection experienced across sites in the genomes of social and nonsocial spiders, estimated with the Partitioned Descriptive model in RELAX (54). The distribution of dN/dS across sites in the genomes are illustrated by three categories of dN/dS for social (orange) and nonsocial (sky blue) branches. The vertical dashed line at dN/dS = 1 represents neutral evolution, bars at dN/dS > 1 represent sites experiencing positive selection, and bars at dN/dS <1 represent sites experiencing purifying selection. The arrows show the direction of change in dN/dS between nonsocial and social branches, demonstrating relaxation of purifying and positive selection associated with the transition to sociality. K < 1 indicates genome-wide relaxation of selection.


A
Genes and GO terms experiencing relaxation and intensification in social branches
Genes
N=598
N=275
Relaxation
Intensification
Enriched GO terms
N=57
N=71
Relaxation
Intensification
B
Ontology for genes experiencing relaxation of selection in social branches
neural function and neurogenesis
negative regulation of axonogenesis
astrocyte differentiation
excitatory postsynaptic potential
negative regulation of neuron projection development
chemical synaptic transmission, postsynaptic
regulation of axonogenesis
regulation of neuron projection development
neuron projection extension
regulation of axon extension
regulation of nervous system process
axon extension
neurotransmitter receptor transport
regulation of neuron differentiation
cell growth
negative regulation of cellular component organization
regulation of cell morphogenesis involved in differentiation
negative regulation of cell projection organization
regulation of extent of cell growth
regulation of cell morphogenesis
negative regulation of cell morphogenesisinvolved in differentiation
developmental growth
0
1
2
3
-LOG10(P-value)
C
Ontology for genes experiencing intensification of selection in social branches
energy metabolism
protein insertion into mitochondrial membrane
regulation of ubiquitin−dependent protein catabolic process
establishment of protein localization to mitochondrial membrane
regulation of cellular protein catabolic process
NADH dehydrogenase complex assembly
immune response
response to biotic stimulus
response to external biotic stimulus
response to other organism
response to virus
defense response to virus
defense response to other organism
neurogenesis
synapse organization
peripheral nervous system development
behavior
olfactory behavior
male mating behavior
chemosensory behavior
negative regulation of locomotion
regulation of locomotion
reproduction
gonad development
development of primary sexual characteristics
0.0
1.0
2.0
-LOG10(P-value)
D
Examples of genes experiencing relaxation in social branches
relaxed purifying selection
PGGT1B
K = 0.97
P = 0.0001
Social
Nonsocial
Proportion of sites
dN/dS
relaxed positive selection
NINJ1
K = 0.58
P = 0.0012
Social
Nonsocial
Proportion of sites
dN/dS
relaxed purifying and positive selection
AGRN
K = 0.63
P = 0.0043
Social
Nonsocial
Proportion of sites
dN/dS
E
Examples of genes experiencing intensification in social branches
intensified purifying selection
CenG1A
K = 1.46
P = 0.0099
Social
Nonsocial
Proportion of sites
dN/dS
intensified positive selection
MRRF
K = 1.84
P = 0.0037
Social
Nonsocial
Proportion of sites
dN/dS
intensified purifying and positive selection
PHRF1
K = 1.28
P = 0.0013
Social
Nonsocial
Proportion of sites
dN/dS

718

Figure 2. RELAX identifies a substantial set of genes that tend to experience relaxation or intensification of selection during the transition to sociality. (A) the number of protein-coding genes and enriched GO terms experiencing relaxation or intensification of selection. Genes or enriched GO terms are depicted in chalk yellow and light grey, respectively. (B) many enriched GO terms for genes experiencing relaxation of selection were associated with neural function, neurogenesis, and cell growth. 20 out of a total of 57 significantly ($P < 0.05$, Fisher's exact test) enriched GO terms (biological process, BP) are shown. (C) enriched GO terms (BP) for genes experiencing intensification of selection were associated with energy metabolism, immune response, nervous system development, reproduction and behavior. 20 out of a total of 71 significantly enriched GO terms are shown. (D) examples of genes that experienced relaxation of selection in social branches caused by relaxed purifying selection, relaxed positive selection, or both. (E) examples of genes that experienced intensification of selection in selection caused by intensified purifying selection, intensified positive selection, or both.

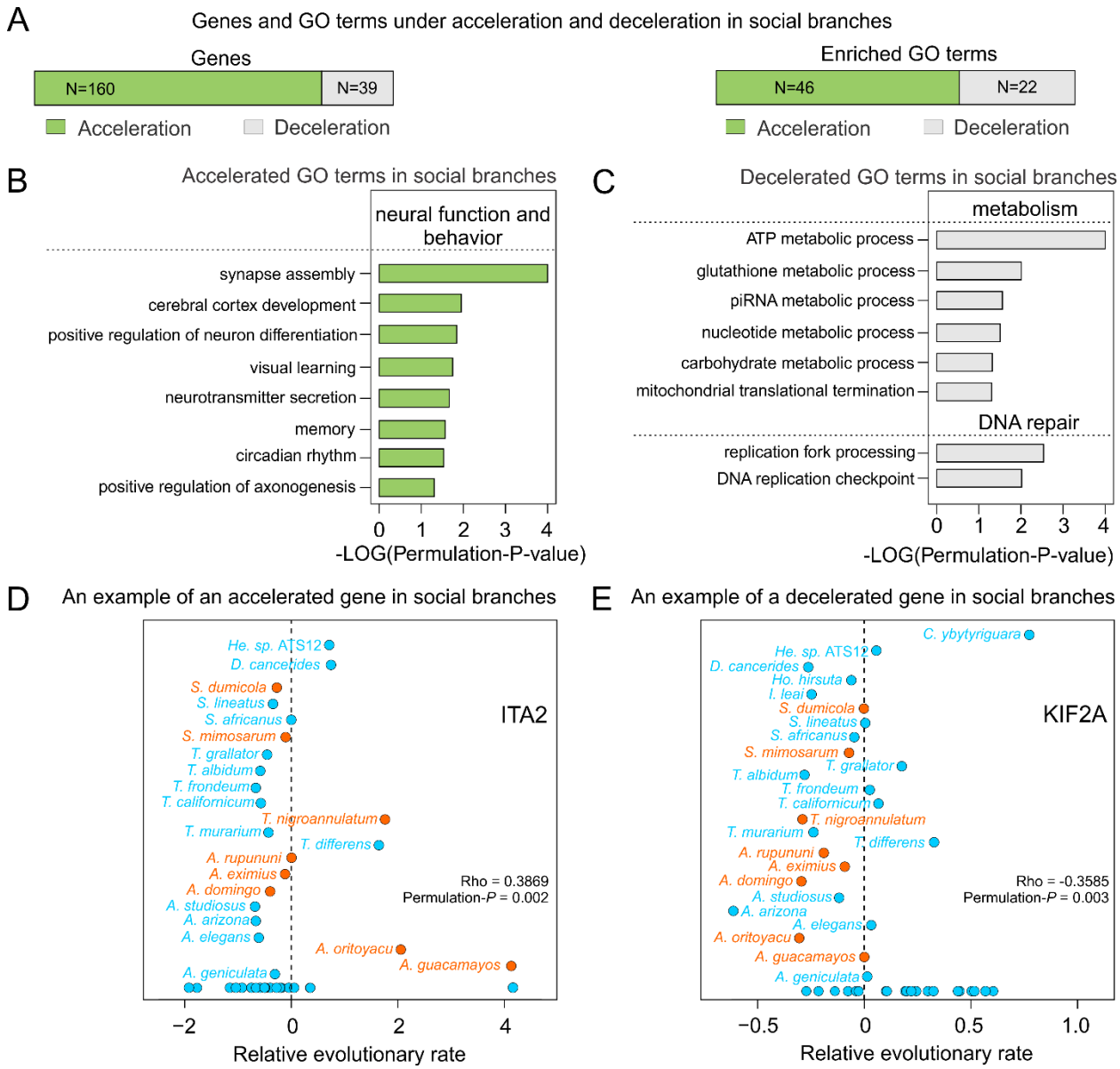


Figure 3. RERconverge identifies over 200 genes that tended to experience shifts in gene-wide rates of molecular evolution during the transition to sociality. (A) the number of genes and GO terms experiencing acceleration or deceleration. Genes or enriched GO terms are depicted in green and light grey, respectively. (B) 8 out of a total of 46 significantly (permulation *P* value < 0.05) enriched GO terms (biological process, BP) in social branches, such as main categories of neural function and behavior. (C) 8 out of a total of 22 significantly decelerated GO terms (BP) in social branches, including main categories of metabolism and DNA repair functions. (D) an example of an accelerated gene in social branches, the nervous system development related gene, Integrin alpha-PS2 (ITA2). Social spider taxa are depicted in orange, nonsocial taxa are in sky blue. (E) an example of a decelerated gene in social branches, Kinesin-like protein KIF2A (KIF2A), a microtubule-associated motor protein with effects on neural development and function.

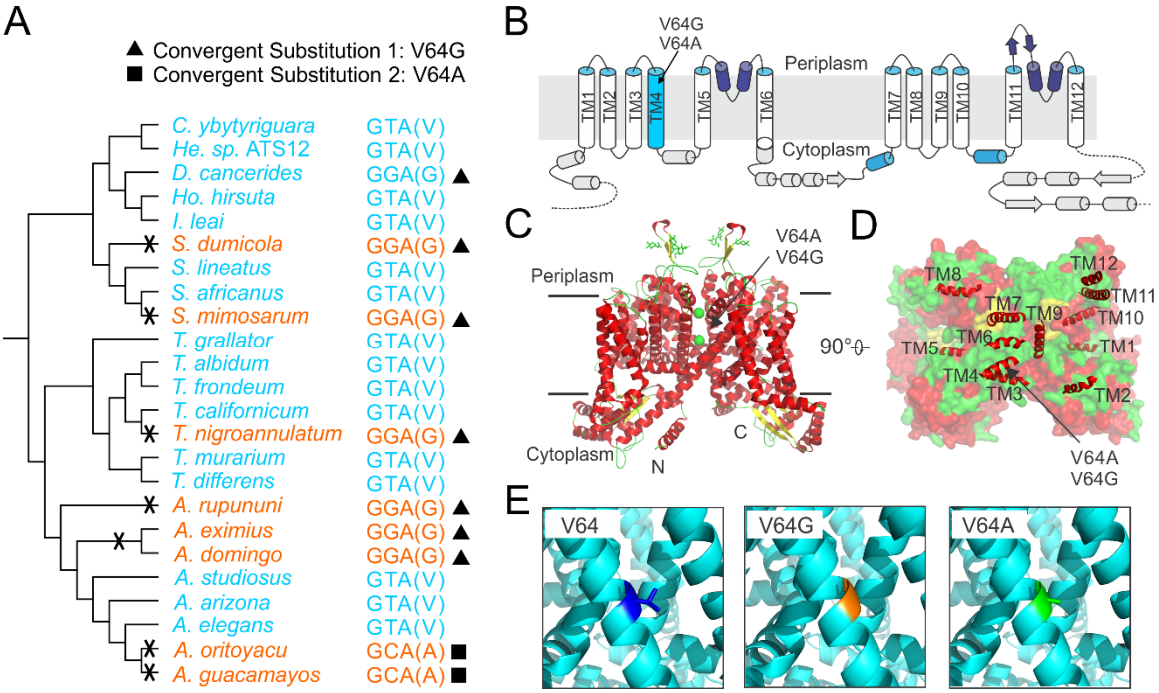


Figure 4. Amino acid substitutions in an ion channel gene occur almost exclusively in social branches and are predicted to cause protein structural changes. (A) Amino acid substitutions at position 64 of the ion channel gene Two pore calcium channel protein 1 (TPC1) were found in all social species and only a single nonsocial species, the prolonged subsocial *D.* cancerides. Triangles indicate a substitution from valine to glycine and squares represent a substitution of valine to alanine. Social spider species are depicted in orange, and nonsocial spider taxa in sky blue. X's on the phylogenetic tree represents the independent origins of sociality. (B) Schematic diagram depicting the topology and domain arrangement in TPC1. Periplasmic or cytoplasmic regions (dark line between boxes) and the 12 transmembrane (TM) domains (boxes) of TPC1 are shown. The substitutions V64G and V64A are located in the TM4 domain. (C) Ribbon representation of the protein structure of spider TPC1 viewed from within the membrane. The amino acid residue V64 is located at the center of TPC1, close to the periplasm side. (D) Periplasm (top) view of the homology structure of TPC1, showing the location of the substitutions at position 64. The 12 transmembrane alpha helices are shown as ribbons. (E) Residue V64 (blue color) and substitutions V64G (orange) and V64A (green) are shown as stick representations.