Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Nature. Author manuscript; available in PMC 2016 January 14.

Published in final edited form as:

Nature. 2015 February 5; 518(7537): 98–101. doi:10.1038/nature13965.

Transferred interbacterial antagonism genes augment eukaryotic innate immune function

Seemay Chou[1,*], **Matthew D. Daugherty**[2,3,*], **S. Brook Peterson**[1], **Jacob Biboy**[4], **Youyun Yang**[5], **Brandon L. Jutras**[6,7], **Lillian K. Fritz-Laylin**[8], **Michael A. Ferrin**[1], **Brittany N. Harding**[1], **Christine Jacobs-Wagner**[6,7,9,10], **X. Frank Yang**[5], **Waldemar Vollmer**[4], **Harmit S. Malik**[2,3], and **Joseph D. Mougous**[1]

Joseph D. Mougous: mougous@u.washington.edu

[1]Department of Microbiology, University of Washington School of Medicine, Seattle, WA 98195, USA

[2]Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA

[3]Howard Hughes Medical Institute, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA

[4]Centre for Bacterial Cell Biology, Institute for Cell and Molecular Biosciences, Newcastle University, Newcastle upon Tyne, NE2 4AX, UK

[5]Department of Microbiology and Immunology, Indiana University School of Medicine, Indianapolis, IN 46202, USA

[6]Microbial Diversity Institute, Yale University, New Haven, CT 06516, USA

[7]Howard Hughes Medical Institute, Yale University, New Haven, CT 06516, USA

[8]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA 94158, USA

[9]Department of Microbial Pathogenesis, Yale University, New Haven, CT 06516, USA

[10]Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, CT 06516, USA

Abstract

Horizontal gene transfer (HGT) allows organisms to rapidly acquire adaptive traits[1]. Though documented instances of HGT from bacteria to eukaryotes remain rare, bacteria represent a rich source of new functions potentially available for co-option[2]. One benefit that genes of bacterial origin could provide to eukaryotes is the capacity to produce anti-bacterials, which have evolved

Correspondence to: Joseph D. Mougous, mougous@u.washington.edu.

*These authors contributed equally to this work.

Author Contributions

S.C., M.D.D., H.S.M. and J.D.M. designed the study. S.C., M.D.D., S.B.P., J.B., Y.Y., B.L.J., L.K.F.-L., M.A.F., B.N.H., C.J.-W., X.F.Y., W.V., H.S.M., and J.D.M. performed experiments, analyzed data, and provided intellectual input into aspects of this study. S.C., M.D.D., S.B.P., H.S.M., and J.D.M. wrote the manuscript; all authors contributed to its editing.

The authors declare no competing financial interests.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

in prokaryotes as the result of eons of interbacterial competition. The type VI secretion amidase effector (Tae) proteins are potent bacteriocidal enzymes that degrade the cell wall when delivered into competing bacterial cells by the type VI secretion system (T6SS)[3]. Here we show that *tae* genes have been transferred to eukaryotes on at least six occasions, and that the resulting domesticated amidase effector (*dae*) genes have been preserved for hundreds of millions of years via purifying selection. We show that the *dae* genes acquired eukaryotic secretion signals, are expressed within recipient organisms, and encode active antibacterial toxins that possess substrate specificity matching extant Tae proteins of the same lineage. Finally, we show that a *dae* gene in the deer tick *Ixodes scapularis* limits proliferation of *Borrelia burgdorferi*, the etiologic agent of Lyme disease. Our work demonstrates that a family of horizontally acquired toxins honed to mediate interbacterial antagonism confers previously undescribed antibacterial capacity to eukaryotes. We speculate that the selective pressure imposed by competition between bacteria has produced a reservoir of genes encoding diverse antimicrobial functions that are tailored for facile co-option by eukaryotic innate immune systems.

Eukaryotes can acquire new functions through the exchange of genetic material with other domains of life[1]. Indeed, Bacteria-to-Eukarya HGT underlies the adaptation and diversification of many microbial eukaryotes, such as algae, choanoflagellates, and protozoa[4,5]. The acquisition of bacterial genes by metazoans is rare. Among the transferred genes, many are not expressed and have no known function[6], while others have roles in endosymbiont maintenance[7,8]. Relatively few reports provide evidence of transferred elements conferring traits directly beneficial strictly to their metazoan recipients[2]. One recent example is the discovery that phytophagous mites and Lepidoptera species exploit a horizontally acquired bacterial cysteine synthase in order to feed on plants producing cyanogenic defense compounds[9].

Genes that can independently provide new functionality to a recipient organism are strong candidates for domestication following HGT[6,10]. The Tae proteins are small, single domain enzymes that can rapidly digest the bacterial cell wall[11]. These proteins comprise four phylogenetically distinct families (Tae1–4) that share no overall sequence homology and display unique specificities against peptidoglycan (PG)[3,12]. In the course of probing *tae* distribution, we made the serendipitous observation that homologs are found in distantly related eukaryotic genomic and expression datasets ranging from unicellular protozoa to multicellular metazoans (Fig. 1a). The genes did not appear to derive from contaminating bacterial DNA; most contain introns and are located in genomic regions flanked by eukaryotic genes (Extended Data Fig. 1)[13]. We therefore refer to these eukaryotic loci as domesticated amidase effector (*dae*) genes, and hypothesized they encode antibacterial toxins horizontally acquired from bacteria. Maximum likelihood and Bayesian phylogenetic analyses revealed that trees of bacterial *tae2–4* families each contained two distinct monophyletic clades of eukaryotic *dae* genes (Fig. 1b; Extended Data Figs. 2–4). Thus, we conclude that three of the four known *tae* gene families have been acquired by eukaryotes from diverse bacteria in at least six HGT events (Fig. 1a). Our survey is biased by the status of genome sequencing efforts; therefore, these six instances are likely an underestimate of eukaryotic *tae* acquisitions.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Three of the *dae* genes we found are limited to individual or closely related eukaryotes (light green, light blue and dark blue; Fig. 1a). These could represent recent HGT events, or reflect limited genomic and transcriptomic sampling of related species. The remaining three *dae* genes appear to be the result of ancient HGT events. For instance, we found *dae2* in 10 species of ticks and mites (Fig. 1b). This dense sampling, a shared intron between the *dae2* genome sequence of *I. scapularis* and *Metaseiulus occidentalis*, and the fact that the tick and mite *dae2* gene phylogeny closely resembles the established phylogeny of these organisms, lead us to conclude that vertical transmission followed a single HGT event of a bacterial *tae2* gene to the common ancestor of ticks and mites approximately 400 million years ago (MYA) (Figs. 1b–d; Extended Data Fig. 1, Fig. 5a and 5b)[14]. The complete genome sequence of the Acariform mite *Tetranychus urticae* does not possess *dae2*, indicating loss of the gene has also occurred. Partial *dae2* sequences in the genomes of two scorpions species and the horseshoe crab share an intron position with *dae2* from ticks and mites, suggesting that *dae2* introduction into arthropods may have occurred as early as 550 MYA (Extended Data Fig. 5c). Similarly, dense sampling of *dae4* genes in gastropod and bivalve mollusks, as well as a shared *dae4* intron position across all sampled mollusks and an annelid, dates the origin of *dae4* in these animals to at least 400 MYA (light red; Fig. 1a and Extended Data Fig. 1 and Fig. 4)[15]. Finally, a second *dae4* present in a species of choanoflagellates, sea anemones, acorn worms and lancelets is most parsimoniously explained by a single HGT event followed by vertical inheritance and loss in multiple lineages, dating this *dae4* acquisition prior to the base of the metazoan lineage (>800 MYA) (dark red; Fig. 1a and Extended Data Fig. 4). However, owing to sparse sampling and lack of evidence of shared synteny, we cannot rule out more recent HGT to and between these eukaryotic lineages[4]. In summary, we find compelling evidence that at least two animal lineages have retained a bacterially derived antibacterial gene for hundreds of millions of years.

Several lines of evidence led us to hypothesize that *dae* genes provide an adaptive function to their eukaryotic hosts. We found strong signatures of purifying selection acting on *dae2* and *dae4* genes (Extended data Table 1). Additionally, eukaryotic Sec signals were identified in the majority of Dae proteins, including representatives from each of the predicted HGT events (Extended data Fig. 6). Secretion of bacterial Tae proteins occurs through the Sec-independent T6SS; thus, acquisition of a Sec signal is indicative of functional specialization involving export from eukaryotic cells. Lastly, the majority of Dae proteins possess the cysteine–histidine catalytic dyad and flanking motifs of their corresponding Tae families, consistent with retention of enzymatic activity (Extended data Fig. 6).

We next sought evidence of expression for eukaryotic *dae* homologs belonging to each of the transferred bacterial *tae* families. We found *dae2* expression during both the unfed nymphal and unfed adult life stages of the hard tick *I. scapularis*, with levels significantly elevated in adults (Fig. 2a). In the amoeba *Naegleria gruberi*, we observed a basal level of expression of each of the three *dae3* homologs in trophozoite (amoeba) cells, which increased during differentiation into flagellates (Fig. 2b). A published expression profile of the lancelet *Branchiostoma floridae* indicates that expression of *dae4* is enriched at the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

neurula stage of development[16]. Together, these data strongly support the hypothesis that *dae* genes have been functionally integrat2 ed into recipient physiology.

The Tae families display unique specificities against PG. Within PG typified by Gram-negative Proteobacteria, enzymes from families 1 and 4 cleave at the γ-D-glutamyl-*meso*-diaminopimelic acid (*m*DAP) bond, whereas those from 2 and 3 cleave the *m*DAP-D-alanine bond crosslinking the peptide stems (Fig. 2c)[3,12,17]. To test whether Dae proteins can hydrolyze PG, we incubated purified Dae2–4 representatives from *I. scapularis*, *N. gruberi*, and *B. floridae*, respectively, with isolated *E. coli* PG sacculi. HPLC analysis of reaction products demonstrated that each of the enzymes hydrolyzes PG (Fig. 2d and Fig. 2e). Remarkably, Dae2, Dae3, and Dae4 display substrate specificity matching that of the characterized extant Tae homologs within corresponding families (Fig. 2c). These data support the hypothesis that *dae* homologs, derived from three *tae* families, have been retained in eukaryotic genomes due to their PG amidase activity. We did not find evidence supporting the transfer of housekeeping bacterial amidases, leading us to speculate that genes encoding T6S effectors – enzymes that intoxicate recipient cells at exceedingly low concentrations – might be especially amenable to preservation following HGT[18].

Within eukaryotes, enzymes with PG-degrading activity might play immuno-regulatory roles, or act directly as antibacterial factors like the Tae toxins[19]. To explore the functional significance of a domesticated *tae*, we focused on *dae2* from the deer tick *I. scapularis*, an important vector for numerous diseases, including Lyme borreliosis and anaplasmosis[20]. Western blot analysis of adult *I. scapularis* demonstrated that Dae2 is present in the salivary glands and midgut (Fig. 3a). *I. scapularis* is an ectoparasite that requires a blood meal for life stage transitions; pathogens are typically acquired during feeding and transmitted to a new host at the next blood meal. Accordingly, the midgut and salivary glands interface with bacterial pathogens and influence their transmission[21].

To understand how Dae2 could contribute to innate bacterial defense within *I. scapularis*, we tested its capacity to cleave diverse PG structures representative of bacteria the organism encounters in the environment[22]. Consistent with its ability to degrade *E. coli* PG, we found that Dae2 degrades a related form of the cell wall present in Firmicutes belonging to the class Bacilli (Extended data Fig. 7a)[17]. We did not detect cleavage of the lysine-type PG found in *Streptococcus pneumoniae*, which represents the second major PG-type found in Firmicutes (Extended data Fig. 7b). Though the ultrastructure of the *B. burgdoferi* sacculus is not well defined, its amino acid composition appears to differ from that of well-characterized bacterial cell walls[23]. Incubation of *B. burgdorferi* sacculi with Dae2 led to the accumulation of specific enzymatic degradation products, indicating that the cell wall of this organism is also a substrate of the amidase (Extended data Fig. 8).

The Dae proteins are reminiscent of an evolutionarily conserved group of bacteriophage-related eukaryotic innate immune amidases, the PG recognition proteins (PGRPs)[19]. Some PGRPs are directly bacteriocidal and act by hydrolyzing PG, whereas others exert antibacterial activity through alternative mechanisms[24]. We found that exogenous Dae2 is not toxic to intact *E. coli* cells. In contrast, Dae2, but not a catalytically-inactive variant of the enzyme (C43A), administered to outer membrane (OM)-permeabilized *E. coli* or

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

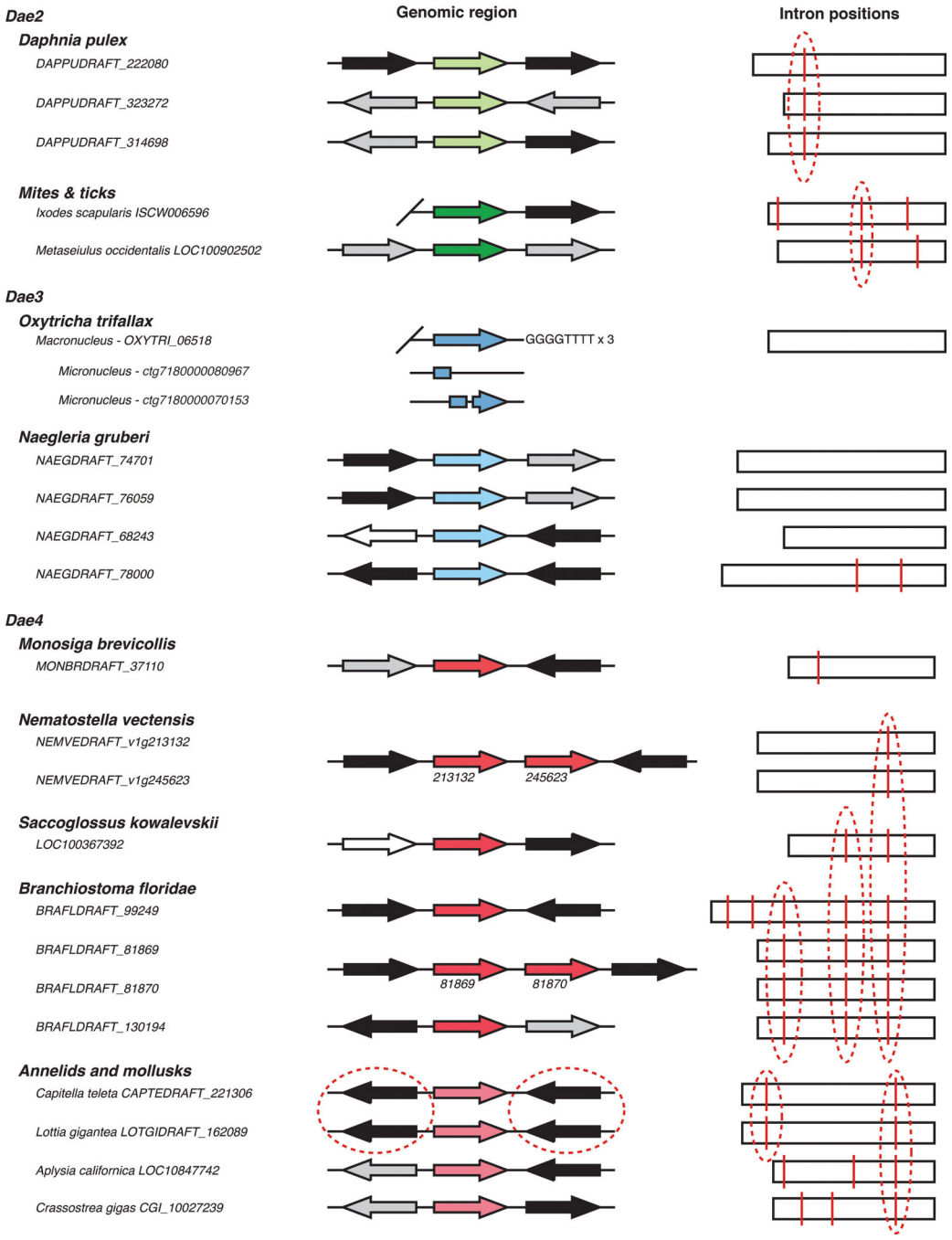
targeted to the periplasm via the Sec pathway is highly lytic (Figs 3b–d). Moreover, exogenous Dae2 is bacteriocidal against *B. subtilis*, which has cell surface exposed PG (Fig. 3d). Together, these results strongly suggest Dae2-dependent antibiosis is solely the result of its amidase activity and that the enzyme would require OM permeabilizing agents such as antimicrobial peptides to act *in vivo*.

B. burgdorferi is the causative agent of Lyme disease, the most prevalent vector-borne illness in the United States[25]. Given the antibacterial activity of Dae2 (Figs 3b–d), its ability to cleave *B. burgdorferi* PG *in vitro* (Extended Data Fig. 8), and its localization to sites that interface with bacteria (Fig. 3a), we hypothesized that Dae2 could play a role in regulating *B. burgdorferi* populations in *I. scapularis*. We tested this possibility using RNAi-mediated knockdown of *dae2* (Fig. 3e). RNAi-treated nymphal ticks were fed to repletion on *B. burgdorferi*-infected mice, and spirochaete load was assessed at engorgement and again after 2 weeks. At repletion, we observed no detectable difference in *B. burgdorferi* levels in control and experimental RNAi-treated ticks, indicating Dae2 activity does not limit initial acquisition of the bacterium (Extended Data Fig. 9a). In contrast, at 2 weeks post-engorgement, *B. burgdorferi* levels were significantly elevated in the *dae2* knockdown group (Fig. 3f). The effect of Dae2 disruption on *B. burgdorferi* levels is unlikely to be due to variations in tick feeding or general fitness, as we observed no difference between the groups in engorgement weights at either time point (Extended Data Fig. 9b). Furthermore, overall bacterial load was similar between groups, suggesting that the increase in *B. burgdorferi* did not result from gross changes in populations of tick-associated microbes (Extended Data Fig. 9c). The ability of Dae2 to act on a wide range of bacterial cell walls leaves open the possibility that compositional changes to the tick microbiome may contribute to the effect of the knockdown on *B. burgdorferi*[26]. We conclude based on these findings that Dae2 contributes to the innate ability of *I. scapularis* to control *B. burgdorferi* levels following acquisition. This has potential ramifications for Lyme disease transmission, as spirochaete load in the tick can influence transmission efficiency[27].

Here, we demonstrate that bacterial genes encoding antibacterial effectors of the T6SS have been horizontally transferred to diverse eukaryotes. The recurrent and independent transfer of *tae* genes to distinct eukaryotic lineages suggests that these toxins can confer immediate fitness benefits by supplying new function to the innate immune system[10]. Recent studies have revealed that the number and diversity of factors mediating interbacterial antagonism is greater than once appreciated. Thus, we speculate that competition between bacteria generates a reservoir of genes – beyond the *tae* superfamily – with the potential to confer antimicrobial capacity to eukaryotes upon acquisition.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Extended Data

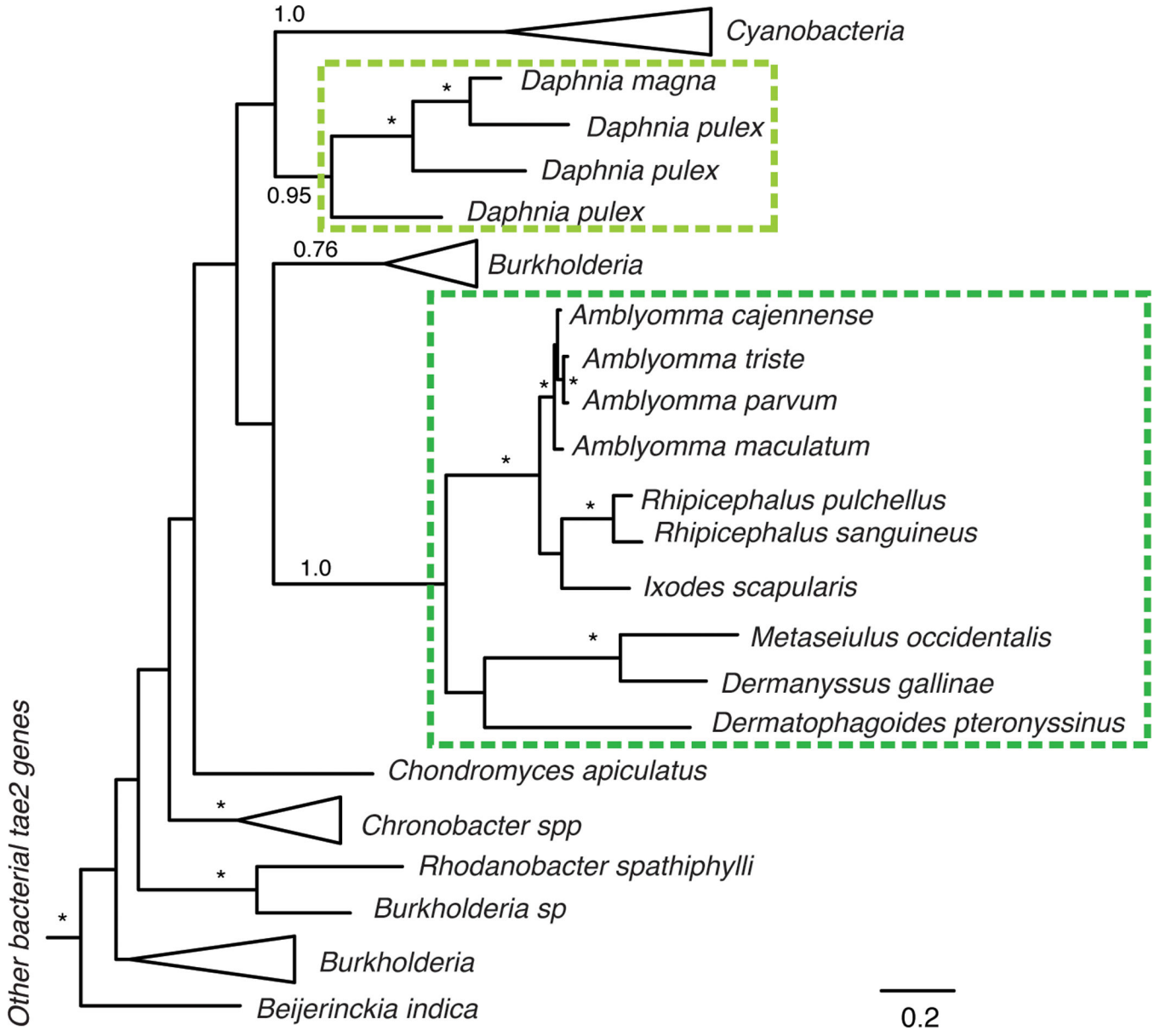


Extended Data Figure 1. Genomic evidence for validated eukaryotic *dae* genes

Eukaryotic *dae* genes from the indicated organisms are listed adjacent to schematic representations of available predicted open reading frames (color-coded according to family as in Fig. 1a) and corresponding genomic context of *dae* genes. Flanking genes are colored-coded according to organisms that homologs of these genes are found in (broadly in eukaryotes, black; only closely related eukaryotic species, grey; both bacteria and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

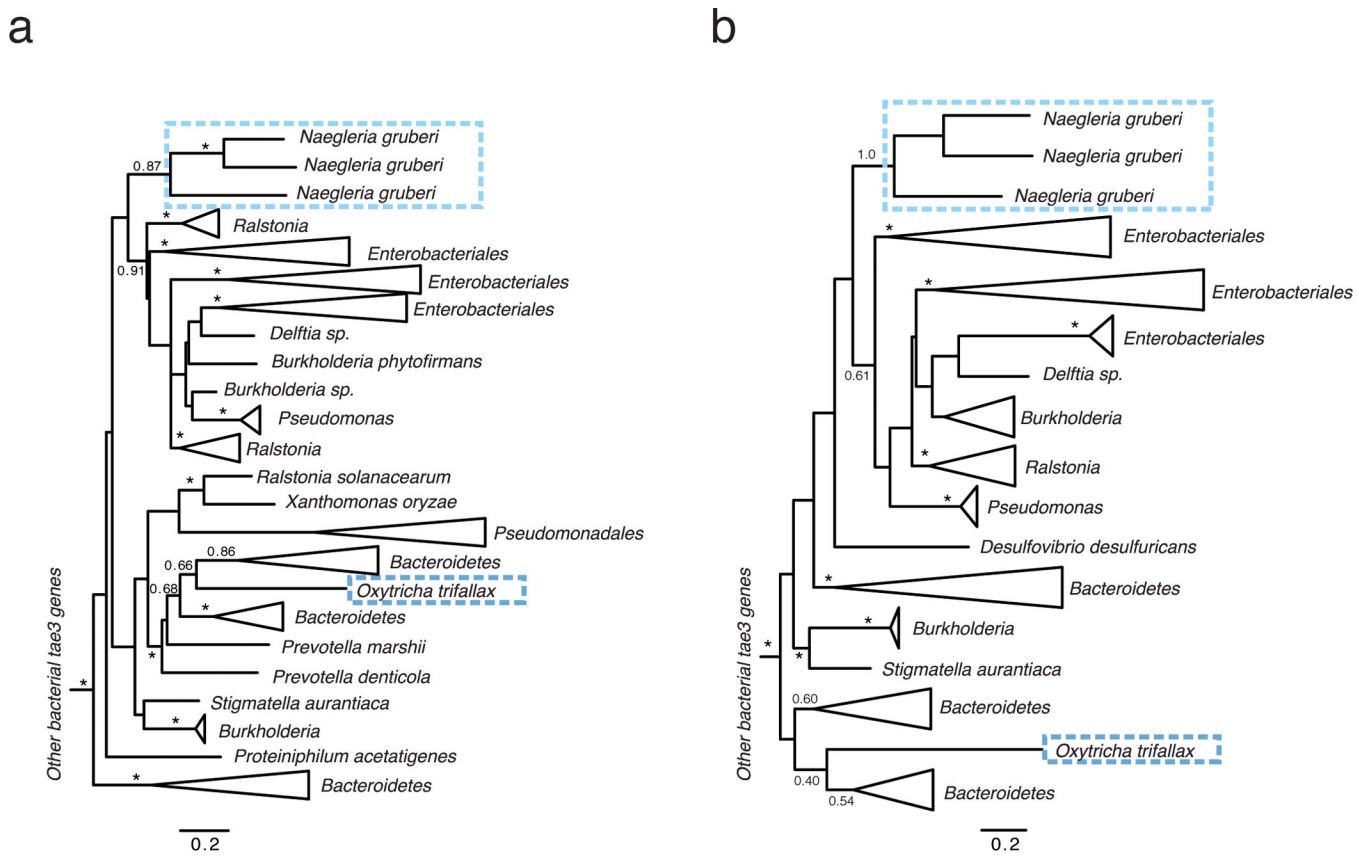
eukaryotes, white). Diagonal lines denote ends of genomic contigs. In the right column, splice sites (red vertical lines) and conserved intron positions (red dashed circles) are shown. In *Oxytricha trifallax*, the somatic nucleus (macronucleus) contains ~16,000 chromosomes and is a rearranged form of the germline nucleus (micronucleus)[13]. The complete *dae3* gene in *Oxytricha* is found in the macronucleus on a chromosome with three characteristic GGGGTTTT telomere sequences. Three fragments comprising the *dae3* gene are found in the micronuclear genome (http://oxy.ciliate.org/). In *Nematostella vectensis* and *Branchiostoma floridae*, lineage-specific duplication events have resulted in two adjacent *dae4* paralogs with gene names labeled (numbers). In *Capitella teleta* and *Lottia gigantea*, shared synteny on both sides of the *dae4* gene is indicated (red dashed circles).



Extended Data Figure 2. Phylogenetic tree of bacterial *tae2* and eukaryotic *dae2* genes

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

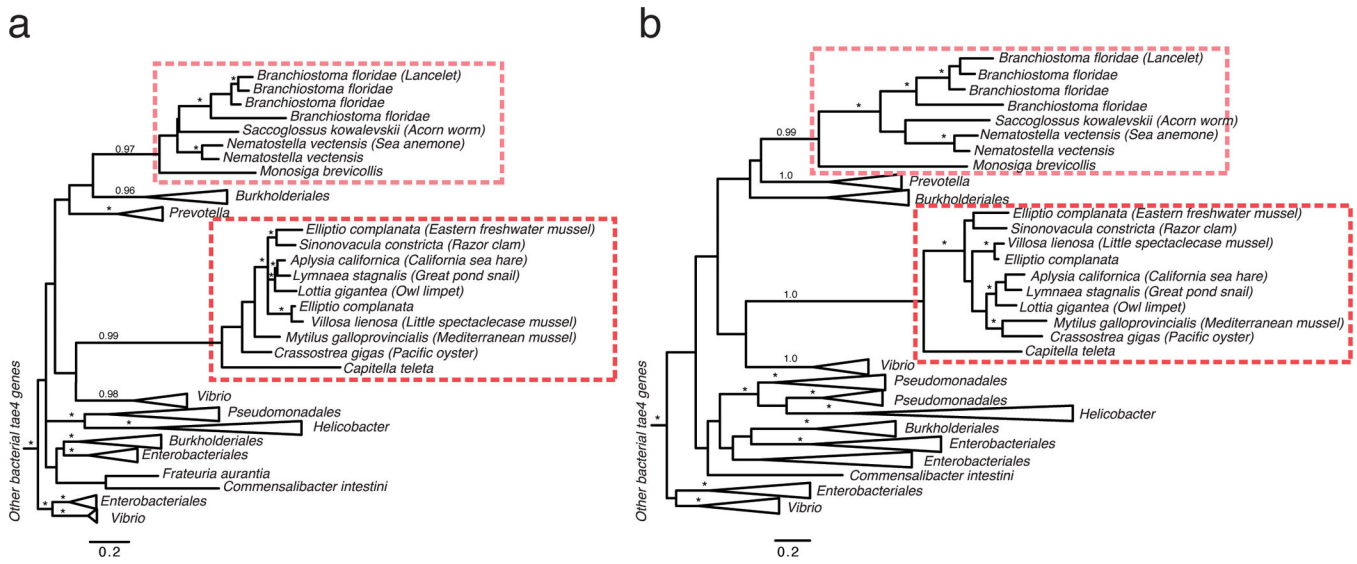
A phylogenetic tree was constructed using Bayesian methods in MrBayes[34] to compare to the maximum likelihood tree shown in Fig. 1b. Branch support > 0.7 is indicated by asterisks or by numbers. The scale bar shows estimated divergence in amino acid changes per residue. Eukaryotic *dae2* genes are indicated by dashed boxes, which highlight two separate HGT events. In both phylogenetic trees, the two eukaryotic *dae2* clades are well supported as monophyletic clades, supporting our conclusion of two HGT events. Likewise, many major bacterial groups are well supported in both trees. Differences in the overall topology of the trees, mostly owing to changes in deep branches that are not well supported in either phylogenetic tree, reflect uncertainty in the ancient history of these genes and should therefore be treated with caution.



Extended Data Figure 3. Phylogenetic tree of bacterial *tae3* and eukaryotic *dae3* genes

Phylogenetic trees were constructed using either maximum likelihood methods (**a**) or Bayesian methods (**b**). Branch support > 0.7 is indicated by asterisks or by numbers. The scale bar shows estimated divergence in amino acid changes per residue. Eukaryotic *dae3* genes are indicated by dashed boxes, which highlight two separate HGT events. In both trees, the two eukaryotic *dae3* clades are well supported as monophyletic clades, supporting our conclusion of two separate HGT events. Likewise, many major bacterial groups are well supported in both trees. Differences in the overall topology of the trees, mostly owing to changes in deep branches that are not well supported in either phylogenetic tree, reflect uncertainty in the ancient history of these genes and should therefore be treated with caution.

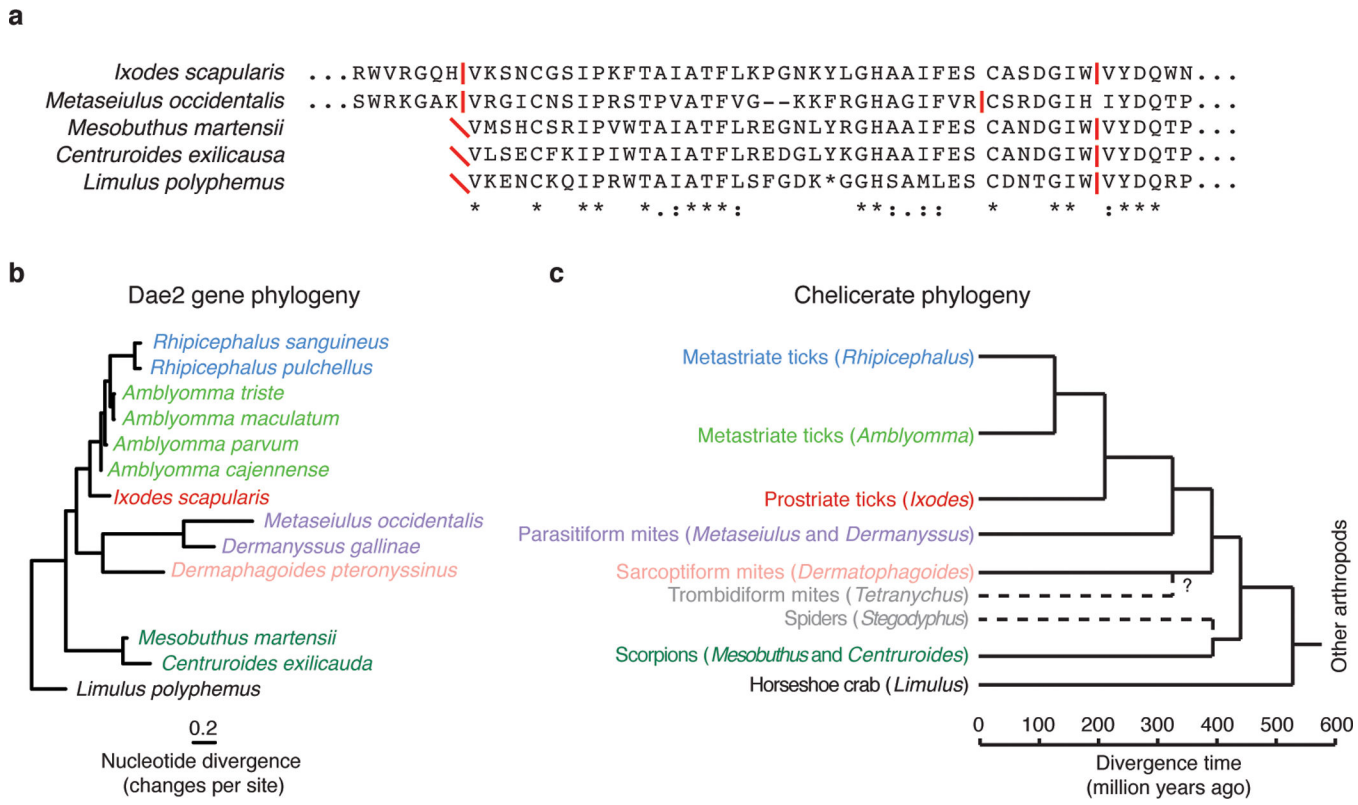
Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Extended Data Figure 4. Phylogenetic tree of bacterial *tae4* and eukaryotic *dae4* genes

Phylogenetic trees were constructed using either maximum likelihood methods (**a**) or Bayesian methods (**b**). Branch support > 0.7 is indicated by asterisks or by numbers. The scale bar shows estimated divergence in amino acid changes per residue. Eukaryotic *dae4* genes are indicated by dashed boxes, which highlight two separate HGT events. In both trees, the two eukaryotic *dae4* clades are well supported as monophyletic clades, supporting our conclusion of two separate HGT events. Likewise, many major bacterial groups are well supported in both trees. Differences in the overall topology of the trees, mostly owing to changes in deep branches that are not well supported in either phylogenetic tree, reflect uncertainty in the ancient history of these genes and should therefore be treated with caution.

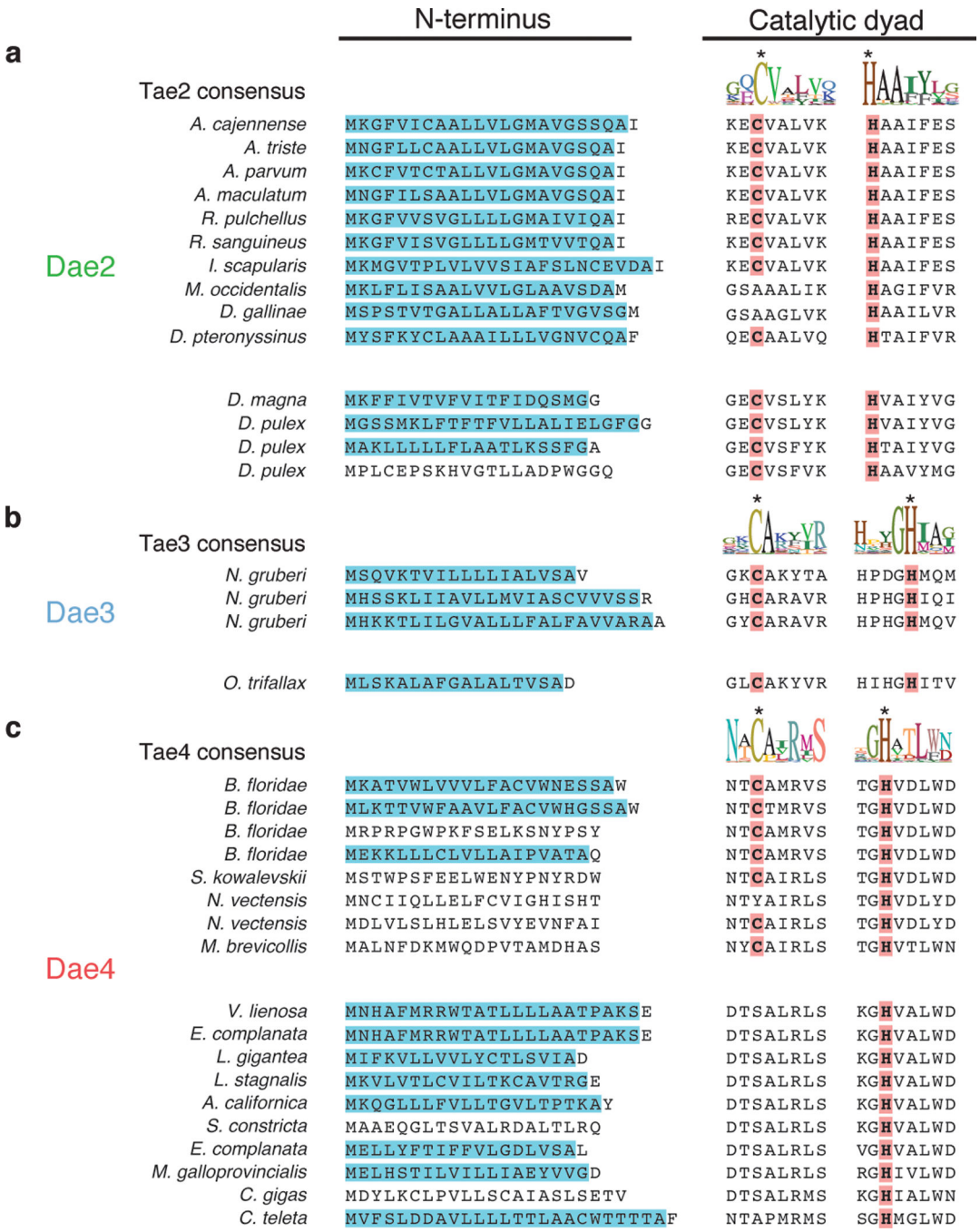
Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Extended Data Figure 5. Evidence for *dae2* in other chelicerates

a, Phylogenetic tree based on partial nucleotide sequences of *dae2* from the indicated chelicerate species. Scale bar shows estimated divergence, in substitutions per nucleotide. **b,** Chelicerate phylogeny with approximate dates of divergence[14]. The unknown divergence time of sarcoptiform and trombidiform mites is indicated by a question mark. We find no evidence for *dae2* in the complete genome of the trombidiform mite *Tetranychus urticae* nor in the partial (several species) or complete genome (*Stegodyphus mimosarum*) of any spider. Putative *dae2* gene loss events in trombidiform mites and spiders are denoted (dashed lines). **c,** Alignment of Dae2 from ticks and mites (*I. scapularis* and *M. occidentalis*) with Dae2 sequences from partially assembled genomes of two scorpions (*Mesobuthus martensii* and *Centruroides exilicauda*) and the horseshoe crab (*Limulus polyphemus*). Splice junctions are denoted (horizontal red lines). All three alignable partial sequences start (red diagonal slashes) in the same position as the shared splice site in tick and mite *dae2* genes, suggesting this is likely the beginning of the exons in all *dae* genes shown. A second intron position is shared between the tick, scorpion and horseshoe crab *dae* genes and is nearby the mite intron position.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

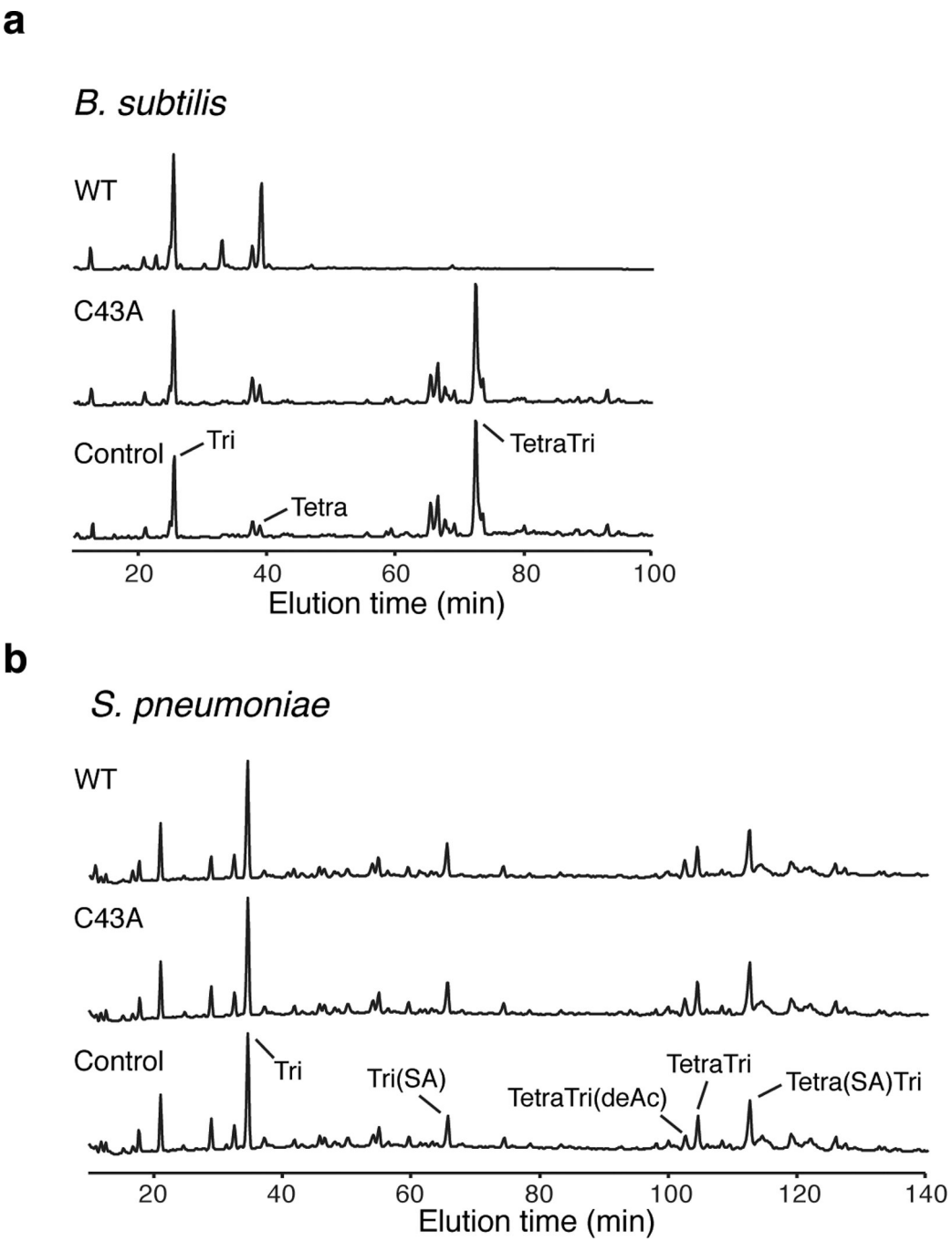


Extended Data Figure 6. Evidence for retention of important catalytic motifs and recurrent eukaryotic-specific addition of secretion signals

a-c, Alignments for the predicted Dae N-terminal signal sequences (shaded blue) and catalytic motifs (right) are shown for each of the families. The consensus sequence logo of residues surrounding the cysteine and histidine positions of catalytic dyads from extant Tae enzymes are shown above alignments from each family. Below are aligned eukaryotic Dae proteins in these same regions. Representatives derived from distinct HGT events are separated by a space. Predicted N-terminal secretion signals (blue) and predicted catalytic

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

residues (red) are colored. Lowering the cutoff value in SignalP[36] from the default value of 0.45 to the 'sensitive' value of 0.34 predicted a signal peptide in residues 1-21 of *C. gigas* Dae4.



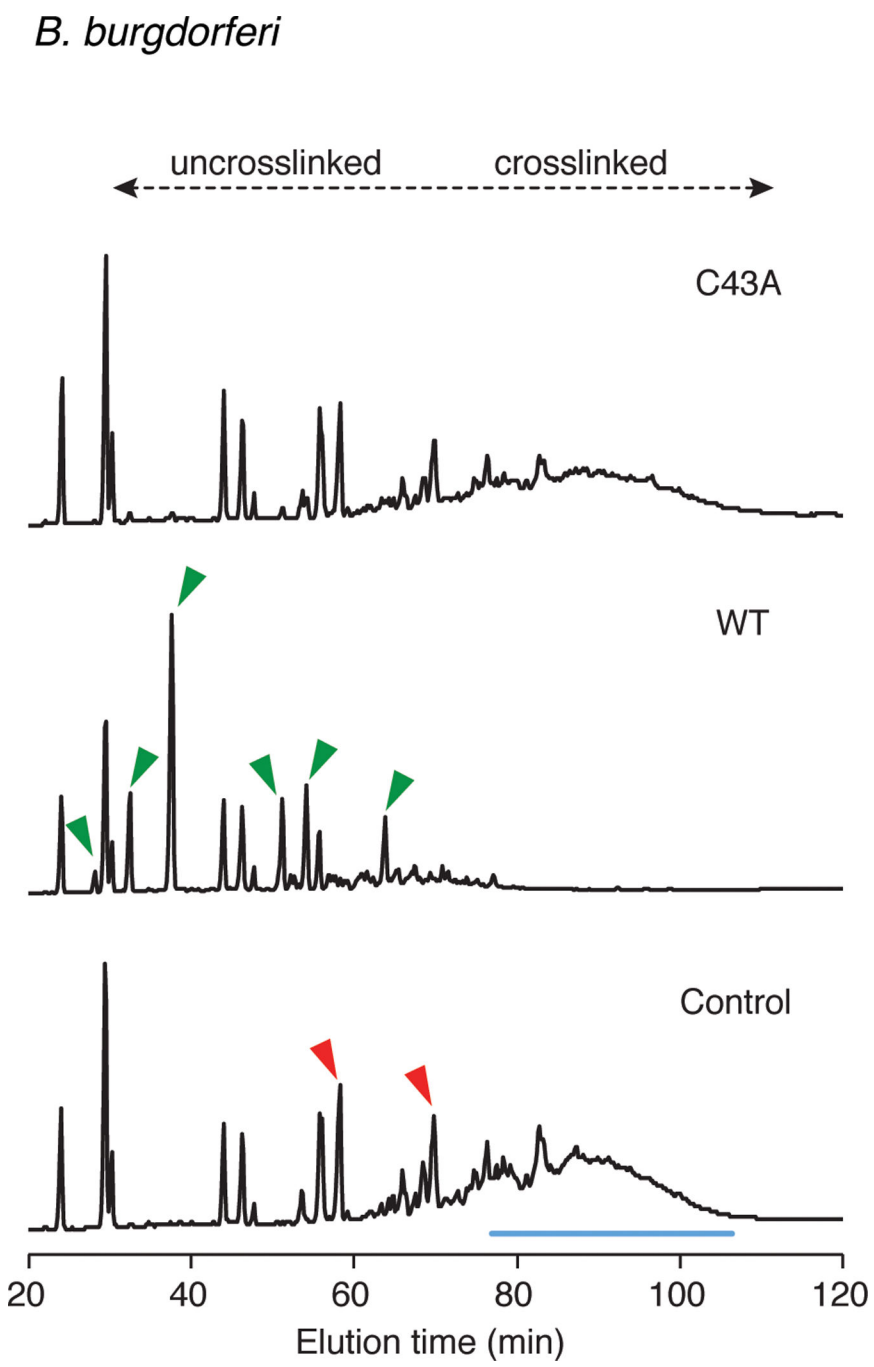
Extended Data Figure 7. Dae2 degrades *m*DAP- but not Lys-type PG

a,b, Partial HPLC chromatograms of sodium borohydride-reduced soluble PG fragments (muropeptides) from *Bacillus subtilis* (**a**) or *Streptococcus pneumoniae* (**b**). PG sacculi products resulting from incubation with buffer (Control) or the indicated Dae2 proteins (WT

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

or C43A), followed by cellosyl digestion are shown. Major peaks are labeled. **a,** Muropeptides from *B. subtilis* include Tri (Glc-*N*Ac–Mur*N*Ac(reduced (r))–L-Ala–D-γ-Glu–*m*DAP(amidated (NH_2))), Tetra (Glc*N*Ac–Mur*N*Ac(r)–L-Ala–D-γ-Glu–*m*DAP(NH_2)–D-Ala), and TetraTri (Glc*N*Ac–Mur*N*Ac–L-Ala–D-γ-Glu–*m*DAP(NH_2)–*D*-Ala–*m*DAP(NH_2)–D-γ-Glu–L-Ala–Mur*N*Ac(r)–Glc*N*Ac). **b,** Muropeptides from *S. pneumoniae* include Tri (Glc*N*Ac–Mur*N*Ac(r) –L-Ala–D-γ-Gln–L-Lys) and TetraTri (Glc*N*Ac–Mur*N*Ac–L-Ala–D-γ-Gln–L-Lys–D-Ala–L-Lys–D-γ-Gln–L-Ala–Mur*N*Ac(r) – Glc*N*Ac). L-Ser–L-Ala branch is indicated by (SA) and deacetylation by (deAc).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

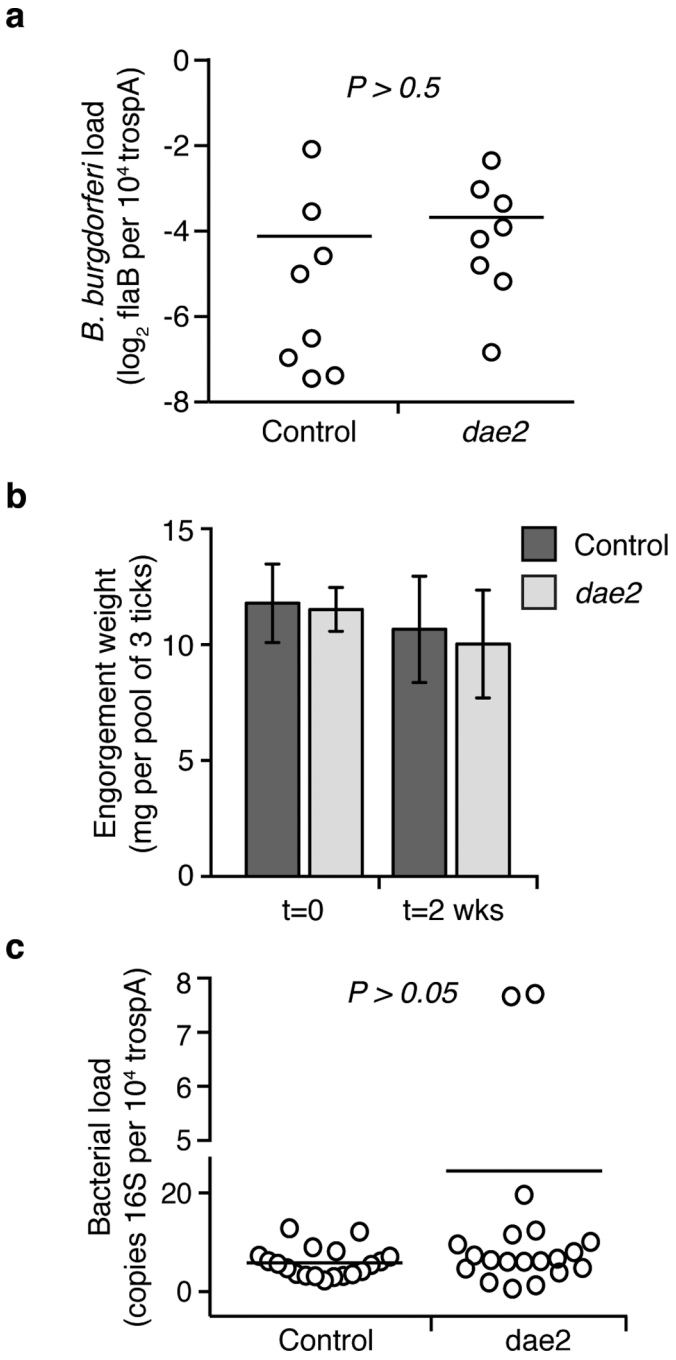


Extended Data Figure 8. Dae2 is active against *B. burgdorferi* PG

HPLC elution profiles of *B. burgdorferi* sacculi incubated with buffer (Control) or the indicated Dae2 proteins (WT or C43A), followed by cellosyl digestion are shown. Discrete peaks lost (red) or produced (green) upon digestion by Dae2 are denoted with arrowheads in Control and WT chromatograms, respectively. Unresolved peaks, likely corresponding to a complex mixture of multi-crosslinked species cleaved by Dae2, are also highlighted (blue line). *B. burgdorferi* PG composition is complex and not yet resolved, thus approximate

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

elution times of uncrosslinked versus crosslinked species are based on *E. coli* muropeptides in the same solvent system.



Extended Data Figure 9. Disruption of *dae2* expression does not significantly alter tick physiology at repletion

a, Knockdown of *dae2* does not increase the *B. burgdorferi* burden in infected nymphs at engorgement. Loads were quantified by qPCR analysis of *flaB*, a *B. burgdorferi*-specific gene, and normalized to *trospa*, a tick-specific gene. n=20. For this and subsequent panels, each data point represents a pool of 3 nymphs, and horizontal bars represent mean values,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

which were not significantly different in a two-tailed nonparametric Mann-Whitney test ($p>0.5$). **b,** Disruption of *dae2* expression did not affect engorgement weights of nymphal ticks fed on *B. burgdorferi*-infected mice. Tick weights were measured at repletion and 2 weeks post-repletion. Error bars +/− s.d., n=8. **c,** Overall bacterial load was not affected by knockdown of *dae2*. Bacterial load was assessed by qPCR analysis of the 16S rRNA normalized against tick-specific gene, *trospA*. Load is represented on both a linear (below) and $\log_2$ (above) scale, which is denoted by a gap on y-axis.

Evolutionary analyses of *dae* and *tae* gene families

Summary of results from maximum likelihood tests of aligned *dae* or *tae* sequences from the indicated species, using SLAC in the HyPhy software package[35]. The overall gene dN/dS ratio (ratio of non-synonymous changes to synonymous changes) is shown, indicating an overall signature of purifying selection. Individual codons with a statistically significant signature of purifying selection ($p<0.05$) were also calculated and are expressed as a percentage of the total number of codons used in the analysis. In the same analyses, no codons were found with a statistically significant signature of positive selection.

	Gene family	Species group	Number of species	dN/dS	Codons evolving under purifying selection
Eukaryotic amidases	*dae2*	Ticks & mites	10	0.20	21% (25 of 120)
	dae4	Mollusks	9	0.18	40% (59 of 149)
Prokaryotic amidases	*tae2*	*Cronobacter*	10	0.12	23% (30 of 129)
	tae3	*Acinetobacter*	10	0.08	30% (45 of 150)
	tae4	*Pseudomonas*	12	0.15	46% (75 of 163)

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank L. Holland for assistance with transcriptome analysis, D. Vollmer and C. Aldridge for PG preparation, J. Parrish for microinjection assistance, J. Young for assistance with phylogenetic analyses, H. Merrikh for sharing equipment, and T. Alber, C. Fuqua, K. Clay, E. Rynkiewicz, U. Pal, C. Grundner, G. Nester, P. Singh and members of the Malik and Mougous laboratories for helpful discussions. This work was funded by the NIH (AI080609 to J.D.M), the Defense Threat Reduction Agency (HDTRA-1-13-014 to J.D.M.), and the BBSRC (BB/I020012/1 to W.V.). S.C. was supported by an HHMI-sponsored Life Sciences Research Foundation fellowship, M.A.F. by the ASM Undergraduate Research Fellowship, and M.D.D. by an Irvington Institute Fellowship from the Cancer Research Institute. C.J.-W. and H.S.M. are investigators of the Howard Hughes Medical Institute. J.D.M. holds an Investigator in the Pathogenesis of Infectious Disease Award from the Burroughs Wellcome Fund.

References

1. Boto L. Horizontal gene transfer in the acquisition of novel traits by metazoans. Proc Biol Sci. 2014; 281:2013–2450.
2. Dunning Hotopp JC. Horizontal gene transfer between bacteria and animals. Trends Genet. 2011; 27:157–163. [PubMed: 21334091]
3. Russell AB, et al. A widespread bacterial type VI secretion effector superfamily identified using a heuristic approach. Cell host & microbe. 2012; 11:538–549. [PubMed: 22607806]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

4. Andersson JO. Gene transfer and diversification of microbial eukaryotes. Annual review of microbiology. 2009; 63:177–193.
5. Schonknecht G, et al. Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. Science. 2013; 339:1207–1210. [PubMed: 23471408]
6. Keeling PJ, Palmer JD. Horizontal gene transfer in eukaryotic evolution. Nature reviews. Genetics. 2008; 9:605–618.
7. Husnik F, et al. Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. Cell. 2013; 153:1567–1578. [PubMed: 23791183]
8. Nikoh N, Nakabachi A. Aphids acquired symbiotic genes via lateral gene transfer. BMC Biol. 2009; 7:12. [PubMed: 19284544]
9. Wybouw N, et al. A gene horizontally transferred from bacteria protects arthropods from host plant cyanide poisoning. Elife. 2014; 3:e02365. [PubMed: 24843024]
10. Moran Y, Fredman D, Szczesny P, Grynberg M, Technau U. Recurrent horizontal transfer of bacterial toxin genes to eukaryotes. Molecular biology and evolution. 2012; 29:2223–2230. [PubMed: 22411854]
11. Russell AB, et al. Type VI secretion delivers bacteriolytic effectors to target cells. Nature. 2011; 475:343–347. [PubMed: 21776080]
12. Chou S, et al. Structure of a Peptidoglycan Amidase Effector Targeted to Gram-Negative Bacteria by the Type VI Secretion System. Cell Rep. 2012; 1:656–664. doi:S2211-1247(12)00141-6 [pii] 10.1016/j.celrep.2012.05.016. [PubMed: 22813741]
13. Swart EC, et al. The Oxytricha trifallax macronuclear genome: a complex eukaryotic genome with 16,000 tiny chromosomes. PLoS Biol. 2013; 11:e1001473. [PubMed: 23382650]
14. Jeyaprakash A, Hoy MA. First divergence time estimate of spiders, scorpions, mites and ticks (subphylum: Chelicerata) inferred from mitochondrial phylogeny. Exp Appl Acarol. 2009; 47:1–18. [PubMed: 18931924]
15. Warnke KM, Meyer A, Ebner B, Lieb B. Assessing divergence time of Spirulida and Sepiida (Cephalopoda) based on hemocyanin sequences. Mol Phylogenet Evol. 2011; 58:390–394. [PubMed: 21145402]
16. Yu JK, et al. A cDNA resource for the cephalochordate amphioxus Branchiostoma floridae. Dev Genes Evol. 2008; 218:723–727. [PubMed: 18773220]
17. Vollmer W, Blanot D, de Pedro MA. Peptidoglycan structure and architecture. FEMS Microbiol Rev. 2008; 32:149–167. doi:FMR094 [pii] 10.1111/j.1574-6976.2007.00094.x. [PubMed: 18194336]
18. Russell AB, Peterson SB, Mougous JD. Type VI secretion system effectors: poisons with a purpose. Nature reviews. Microbiology. 2014; 12:137–148. [PubMed: 24384601]
19. Dziarski R, Gupta D. The peptidoglycan recognition proteins (PGRPs). Genome biology. 2006; 7:232. [PubMed: 16930467]
20. Sonenshine, DE.; Roe, RM. Biology of ticks. 2nd edn. Oxford University Press; 2013.
21. Hajdusek O, et al. Interaction of the tick immune system with transmitted pathogens. Front Cell Infect Microbiol. 2013; 3:26. [PubMed: 23875177]
22. Hawlena H, et al. The arthropod, but not the vertebrate host or its environment, dictates bacterial community composition of fleas and ticks. The ISME journal. 2013; 7:221–223. [PubMed: 22739493]
23. Beck G, Benach JL, Habicht GS. Isolation, preliminary chemical characterization, and biological activity of Borrelia burgdorferi peptidoglycan. Biochem Bioph Res Co. 1990; 167:89–95.
24. Kashyap DR, et al. Peptidoglycan recognition proteins kill bacteria by activating protein-sensing two-component systems. Nat Med. 2011; 17:676–683. [PubMed: 21602801]
25. Radolf JD, Caimano MJ, Stevenson B, Hu LT. Of ticks, mice and men: understanding the dual-host lifestyle of Lyme disease spirochaetes. Nature reviews. Microbiology. 2012; 10:87–99. [PubMed: 22230951]
26. Narasimhan S, et al. Gut microbiota of the tick vector Ixodes scapularis modulate colonization of the lyme disease spirochete. Cell host & microbe. 2014; 15:58–71. [PubMed: 24439898]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

27. Zhang L, et al. Molecular interactions that enable movement of the Lyme disease agent from the tick gut into the hemolymph. PLoS pathogens. 2011; 7:e1002079. [PubMed: 21695244]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

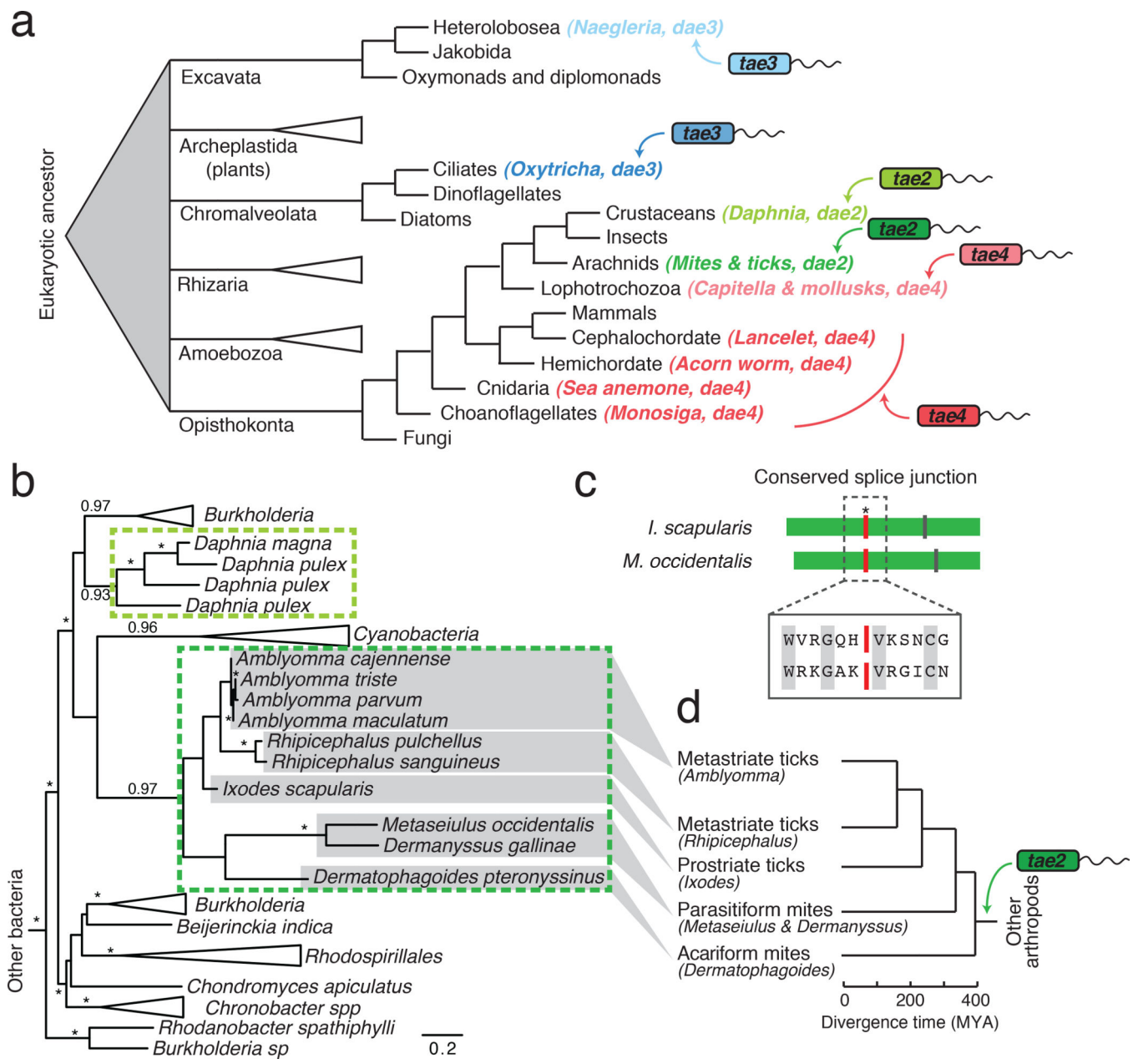


Figure 1. Recurrent horizontal gene transfer of type VI amidase effector (*tae*) genes into diverse eukaryotic lineages

a, Schematized phylogenetic tree of basal eukaryotic lineages[28] showing instances of *tae* transfer (arrows) from bacteria to eukaryotes, coded by color (*tae* family) and shading (acquisition events). **b,** Maximum likelihood phylogenetic tree of *tae2* and *dae2* genes. Representatives are boxed and color-coded according to Fig. 1a. Branch support > 0.7 indicated by asterisks or numbers. Scale bar shows estimated divergence in amino acid changes per residue. Dashed lines highlight separate HGT events. **c,** Schematic alignment of tick (*I. scapularis*) and mite (*M. occidentalis*) *dae2* genes with shared (red line, asterisk) and unique (vertical lines) intron positions denoted. Aligned residues surrounding the splice site

are shown (boxed) with conserved amino acids indicated (grey). **d,** Tick and mite phylogeny with approximate dates of divergence based on concordance with the *dae2* gene tree (**c**)[14].

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

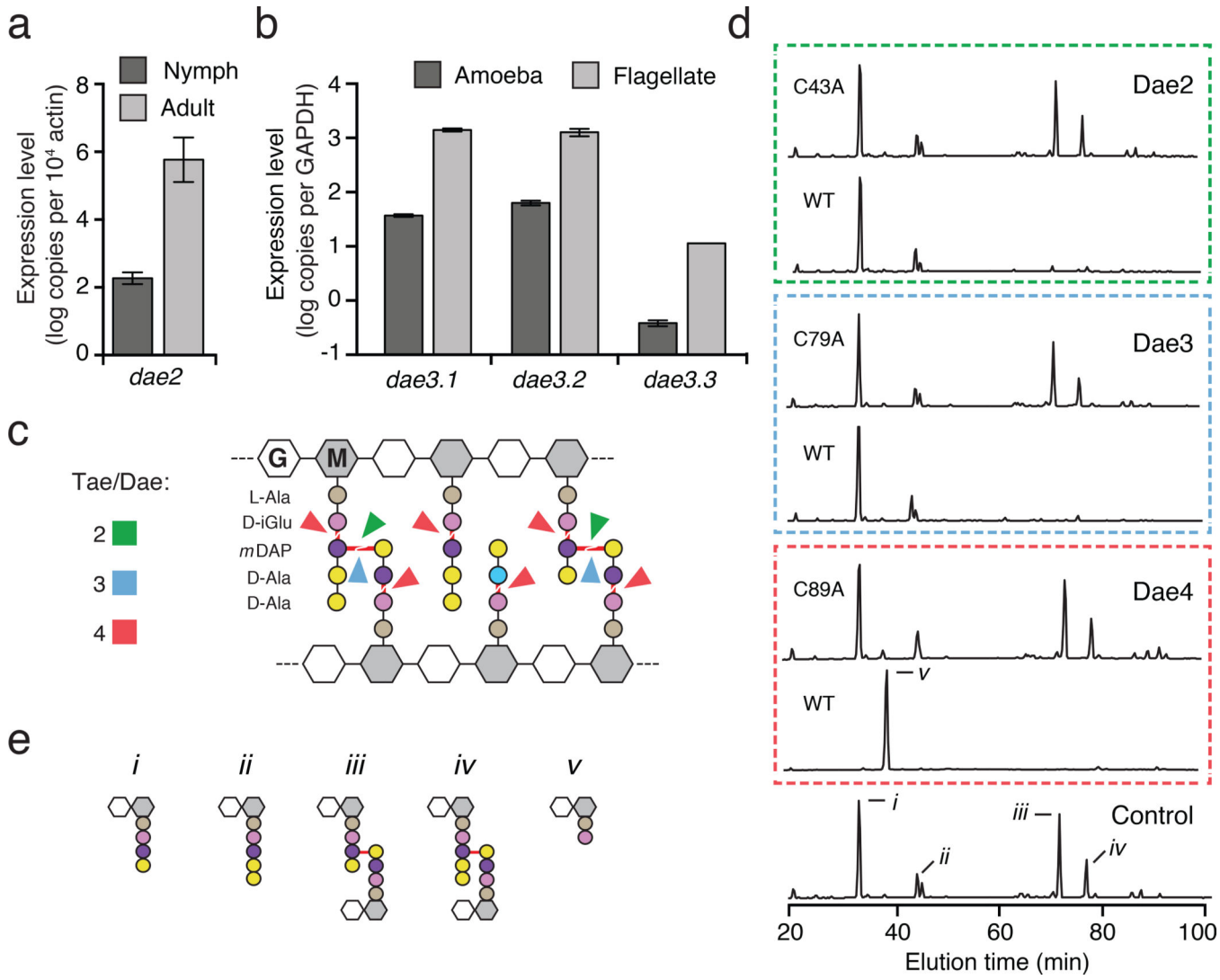


Figure 2. Eukaryotic *dae* genes encode differentially-expressed PG amidases with conserved specificity

a, b, Expression profile of *I. scapularis* and *N. gruberi dae* genes at the indicated life stages as measured by qRT-PCR. Levels of each *N. gruberi dae3* genes (Dae3.1–3.3) were determined. Error bars +/− s.d., n=3. **c,** Schematic representation of typical Gram-negative PG showing cleavage sites (red lines) for Tae and Dae families 2–4 (colors correspond to Fig. 1a). **d,** Partial HPLC chromatograms of *E. coli* PG sacculi products resulting from incubation with buffer (control), native and catalytically inactive (C43A, C79A, C89A) Dae enzymes and cellosyl. **e,** Major HPLC peaks assigned previously by mass spectrometry correspond to disaccharide-linked tetrapeptide (*i*), pentapeptide (*ii*), tetrapeptide–tetrapeptide (*iii*), pentapeptide–tetrapeptide (*iv*), and dipeptide (*v*)[3].

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

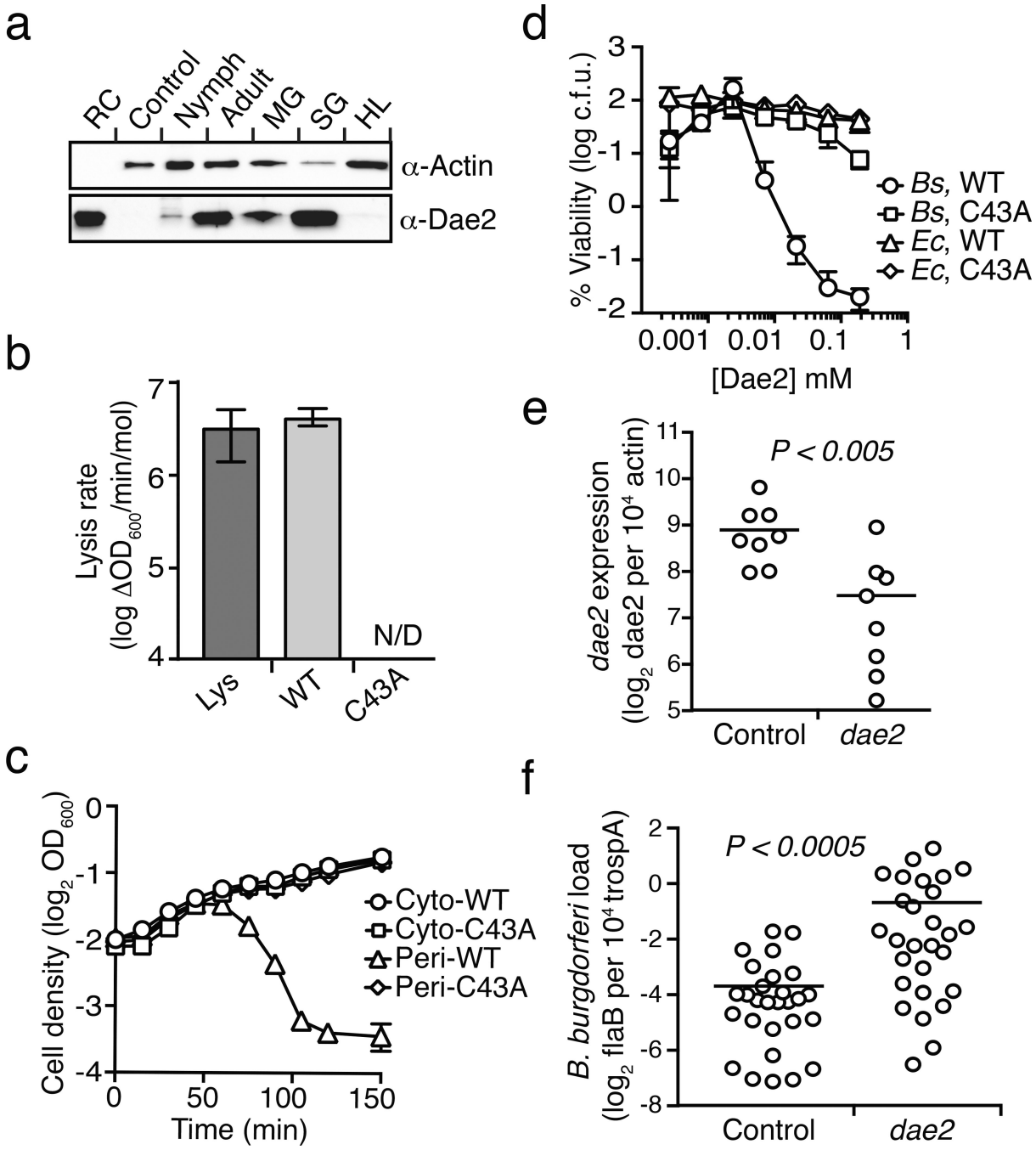


Figure 3. Dae2 is a bacteriolytic toxin that restricts the proliferation of *B. burgdorferi* in the tick *I. scapularis*

a, Western blot analysis of Dae2 in unfed adult and nymphal total tissue (total), midgut (MG), salivary gland (SG), and hemolymph (HL) extracts from *I. scapularis*. Recombinant Dae2 protein (RC) and tissue from a closely related species, *Dermacentor variablis* (Control), were included. Actin levels were examined as a loading control. **b,** Lytic activity of lysozyme (Lys) and Dae3 (WT, C43A) proteins against permeabilized *E. coli*. Error bars +/− s.d., n=3. **c,** Growth of *E. coli* expressing native (cyto-) or periplasm-targeted (peri-)

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Dae2 proteins. Error bars +/− s.d., n=3. **d,** Bacterial killing activity of indicated proteins against *B. subtilis* (*Bs*) and *E. coli* (*Ec*) cells. Error bars +/− s.d., n=3. **e,** *Dae2* transcript levels quantified by qRT-PCR in RNAi-treated engorged ticks. **f,** At 2-weeks post-engorgment, spirochaete levels were quantified in ticks that had received the indicated RNAi treatments, using qPCR analysis of *flaB*, a *B. burgdorferi*-specific gene, and normalized to *trospa*, a tick-specific gene. n=20. Each data point in Figs 3e and 3f represents 3 nymphs. Horizontal bars represent mean values, which were significantly different in a two-tailed nonparametric Mann-Whitney test ($p<0.05$).