Shared neural transcriptomic patterns underlie the repeated evolution of mutualistic cleaning behavior in *Labridae* wrasses

EcoEvoRxiv Preprint https://doi.org/10.32942/osf.io/4p638

Rebecca L. Young[1*], Chelsea A. Weitekamp[1#], Zegni Triki[2$], Yiheng Su[1], Abigail G. Beman[1], Redouan Bshary[2], and Hans A. Hofmann[1,3*]

Affiliations:

[1]Department of Integrative Biology, The University of Texas at Austin, Austin, TX, USA

[2]Institute of Biology, University of Neuchâtel, Neuchâtel, Switzerland

[3]Institute for Neuroscience, The University of Texas at Austin, Austin, TX, USA

*Correspondence to**:**
Dr. Hans A. Hofmann (hans@utexas.edu)
Dr. Rebecca L. Young (youngrl@utexas.edu)
Department of Integrative Biology
The University of Texas at Austin
2415 Speedway - C0930
Austin, TX 78745
USA

[#]Current address:
Dr. Chelsea A. Weitekamp
Center for Public Health & Environmental Assessment
U.S. Environmental Protection Agency
109 TW Alexander Dr.
Durham, NC 27709
USA

[$]Current address:
Dr. Zegni Triki
Institute of Ecology and Evolution
University of Bern
Wohlenstrasse 50a
CH-3032 Hinterkapplen
Switzerland

Keywords: cooperative behavior, comparative transcriptomics, amygdala, hippocampus, convergence, parallelism, phylogenetic comparative analysis

Abstract

Despite the remarkable diversity of life forms on earth, evolutionary biologists have discovered numerous instances where even distantly related species share astonishing similarities in how they behave, look, and function. Given the importance of happenstance in evolution (e.g., random mutations, genetic drift, environmental stochasticity), it is often assumed that the mechanisms underlying such convergent phenotypes are distinct. Nevertheless, recent discoveries that the same pathways can underlie convergently evolved phenotypes have reinvigorated questions about the predictability of evolution and whether broadly conserved genomic mechanisms facilitate phenotypic convergence. Here, we generated transcriptomes of the putative teleost homologs of the mammalian hippocampus and basolateral amygdala, broadly associated with spatial and social cognition, in six sympatric species of Labridae wrasses that vary in mutualistic cleaning behavior (including three non-cleaning, two facultative cleaning, and one obligate cleaning species) and combined differential gene expression, gene co-expression, and phylogenetic comparative analyses to test two hypotheses about convergent evolution and specialization of mutualistic cleaning behavior. We first identify genes and gene modules exhibiting parallel neurotranscriptomic patterns in the repeated evolution of facultative cleaning. We then examined whether expression and co-expression patterns associated with facultative cleaning are also shared in the obligate cleaner species in our dataset and found evidence for transcriptomic concordance, though no evidence for additional specialization. Taken together, our results provide insights into the convergent evolution and the neuromolecular basis of cooperative behavior and, more generally, illustrate the potential of phylogenetic comparative transcriptomics to unravel the mechanistic underpinnings of the repeated evolution of complex organismal phenotypes.

Introduction

For all the spectacular diversity generated in evolution, there are often remarkable similarities among species in how they behave, look, and function. Such similarity in phenotypes can reflect both shared evolutionary history – resulting from the multitude of molecular and developmental pathways shared by any given lineage of organisms – and/or convergent responses to similar ecological challenges (Gould, 1989; Brakefield, 2006; Losos, 2011; Stern, 2013; Blount et al., 2018). While there is now ample empirical evidence suggesting that convergence is much more common in nature than Gould (1991) predicted (Orgogozo, 2015; Blount et al., 2018), it is still unclear to which extent the prevalence of convergence is due to shared evolutionary history (Stern, 2013; Bolnick et al., 2018). First, studies of convergence often investigate replicate populations of the same or closely related species (Kocher et al., 1993; Foster and Baker, 2004; Mahler et al., 2013; Langerhans, 2017; Renn et al., 2018). In this case, the selection pressures of similar environments might eclipse the effects of historical contingency. At longer timescales, divergence, contingency, and stochasticity among species are expected to limit the potential for homoplasy (i.e., independently evolved similarity of a trait) (Gould, 1989; Wake et al., 2011; Orgogozo, 2015). Further, because many distinct genotypes, developmental pathways, and cellular origins can give rise to functionally equivalent and even homologous phenotypes (True and Haag, 2001; Young and Wagner, 2011), it has been suggested that the molecular and physiological processes underlying convergently evolved organismic traits (e.g., phenotypes or functions with independent evolutionary origins) are more likely to be nonparallel (i.e., associated with distinct underlying mechanisms) (Storz, 2016; Roda et al., 2017; Bolnick et al., 2018; Cheng et al., 2021). Nevertheless, there are now numerous examples where the same pathways or even genes appear to have been deployed repeatedly in the service of a convergently evolved phenotype (Castoe et

al., 2009; Zhen et al., 2012; Witt and Huerta-Sánchez, 2019), even at the level of the transcriptome (Chan et al., 2009; Gallant et al., 2014; Pankey et al., 2014; Pfenning et al., 2014; Gao et al., 2019; Young et al., 2019). In fact, recent progress resolving evolutionary relationships among animals indicates that such homoplasies are much more common than previously appreciated, even among distantly related taxa (Rokas and Carroll, 2008; Dunn et al., 2014; Lamichhaney et al., 2019).

A fascinating example of behavioral convergence is mutualistic cleaning. In this cooperative behavior, so called 'cleaner' species remove ectoparasites and dead tissue from their 'clients,' which are often larger species (Poulin and Grutter, 1996; Côté, 2000). Cleaning mutualisms have independently evolved in several marine vertebrates and invertebrates including shrimp, crabs, gobies, and wrasses as well as in fresh water and terrestrial systems (reviewed in: Poulin and Grutter, 1996; Côté, 2000; Vaughan et al., 2017). In fishes, repeated evolution of mutualistic cleaning consists not only of behavioral changes but is also associated with anatomical convergence in body elongation and musculoskeletal morphology and function of the feeding apparatus (Baliga and Mehta, 2014, 2019; Huie et al., 2020). The highest proportion and diversity of cleaner fishes are present in the *Labridae* wrasses with at least 58 species exhibiting mutualistic cleaning behavior during at least one life history stage (Côté, 2000; Baliga and Law, 2016; Vaughan et al., 2017). Resulting from an estimated 26 to 30 independent evolutionary transitions, mutualistic cleaning has emerged in wrasses over relatively recent evolutionary history (i.e., within the last 20 million years; Baliga and Law, 2016). Because behavior, like mutualistic cleaning, is closely tied to the neural transcriptome, repeated transitions to the complex mutualistic cleaning phenotype across wrasse species allows us to test the hypothesis that rapid and frequent evolution of complex behavioral phenotypes are facilitated by repeated deployment of parallel neurotranscriptomic mechanisms.

Within the *Labridae* wrasses, cleaning behavior varies in ontogenetic timing and behavioral and cognitive specialization, with some species relying on cleaning as a primary food source – obligate cleaners – and others cleaning only as juveniles or facultatively throughout ontogeny – facultative cleaners (Fig. 1; Côté, 2000; Barbu et al., 2011; Baliga and Law, 2016). Increased reliance on cleaning behavior (i.e., in obligate cleaners) is linked to increased behavioral and cognitive specialization, such as a greater diversity of client species, increased duration of cleaning bouts, opportunistic cheating, and cognitive performance during client interactions (Barbu et al., 2011; Gingins and Bshary, 2016). Further, behavioral specialization correlates with morphological specialization such that the highly specialized obligate cleaner species exhibit more limited morphological variation and increased body elongation (Baliga and Mehta, 2019). Interestingly, evolutionary transitions between juvenile and adult cleaners (obligate or facultative) are phylogenetically correlated suggesting that adult cleaning may have evolved from a juvenile cleaning state (Baliga and Law, 2016), perhaps by maintaining an early life history state via neoteny (or heterochronic changes in timing of developmental or ontogenetic events that maintain juvenile or early life states in the adult organism, "paedomorphosis" sensu Garstang, 1928). Though morphological and phylogenetic correlations suggest similarities in the underlying mechanistic bases and evolutionary trajectories of obligate and facultative cleaner types, whether the neurotranscriptomic underpinnings of cleaning behavior are shared by both cleaner types is unknown.

In general, the neural and molecular underpinnings of cooperative behavior like mutualistic cleaning (for review see: Soares et al., 2010; Soares, 2017) remain poorly understood (Weitekamp and Hofmann, 2014, 2017). However, studies by Soares and colleagues have begun to illuminate the neuroendocrine mechanisms of cleaning behavior, with a particular focus on the Indo-Pacific

bluestreak cleaner wrasse *Labroides dimidiatus*. For example, these authors showed that the nonapeptide arginine vasotocin (AVT, the non-mammalian homolog of arginine vasopressin) appears to inhibit cleaning behavior in *L. dimidiatus* (Cardoso et al., 2015), possibly via the V1a receptor subtype (Soares et al., 2012) and through modulation of their learning competence (Mendonça et al., 2013). While these studies have provided important insights into the regulation of mutualistic cleaning behavior, they have been limited to candidate neuroendocrine and neuromodulatory pathways and to one species that is most amenable to experimental manipulation. In fact, a systems-level understanding of cleaning behavior and its evolution based on genome-wide analyses of the gene co-expression networks is lacking. However, because RNA sequencing can be performed in principle on any tissue and species, neural transcriptomic comparisons across species can test hypotheses of convergent evolution and identify candidate brain regions and novel candidate genes associated with specific behavior phenotypes.

Even though the labrid brain has received little attention from neuroanatomists (but see Nugent et al., 2021; Lamm et al., 2022), there is an evolutionarily conserved Social Decision-Making Network (SDMN) that is critical for evaluating stimulus salience and regulating sexual, aggressive, and parental behavior across vertebrates (O'Connell and Hofmann, 2011b, 2012). Two SDMN nodes that have been well studied for their role in spatial and social cognition, respectively, deserve particular attention in the context of a phenotype as complex as cleaning behavior. First, the medial pallium (which forms the hippocampus in mammals) plays a critical role in spatial memory in mammals (O'Keefe and Nadel, 1978; Andersen et al., 2007; Humphries and Prescott, 2010) and has a functionally equivalent role in both avian and non-avian reptiles (Striedter, 2016; Butler, 2017) and teleost fish (Elliott et al., 2017; Trinh et al., 2019; Vinepinsky et al., 2020). And second, the lateral pallium (which mainly comprises the basolateral amygdala in mammals)

integrates multimodal sensory inputs and regulates affective and goal-directed behavior in mammals (Maeda and Mogenson, 1981; LeDoux, 2000; Moreno and González, 2007), similar to the situation in birds and reptiles (Martínez-García et al., 2002) as well as teleosts (Portavella et al., 2002). Together, the medial (area Dl in teleosts: lateral part of the dorsal telencephalon) and the lateral pallium (area Dm in teleosts: medial part of the dorsal telencephalon) are ideally suited for comparative analyses as they are complementary in function and reciprocally connected.

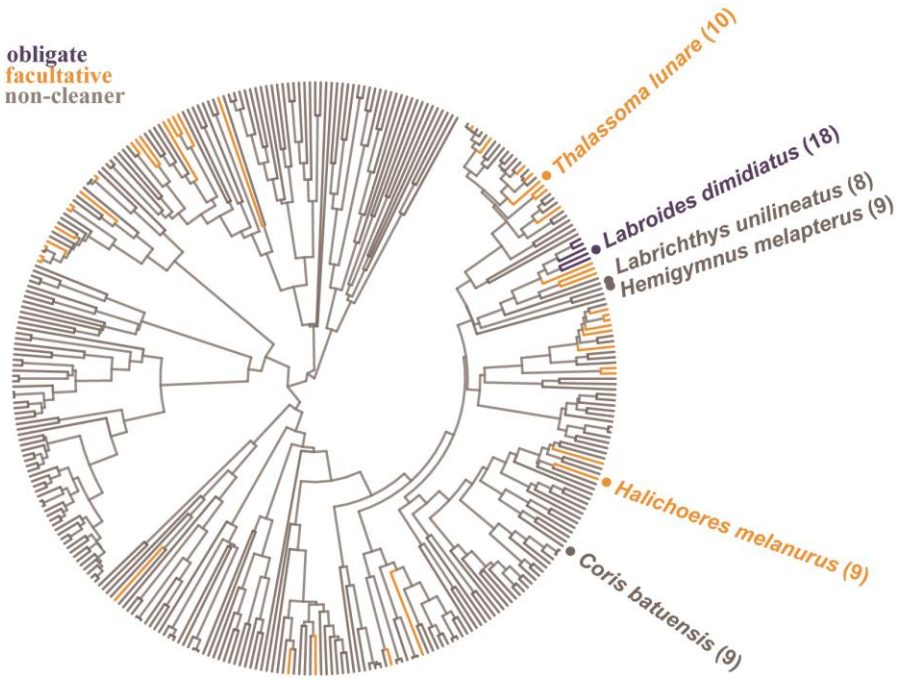


Figure 1. Evolutionary relationships and cleaner type of 343 wrasses (Labridae) (redrawn Baliga and Law, 2016). Branches are colored by species cleaning activity with obligate cleaners, facultative cleaners, and non-cleaners highlighted in purple, gold, and grey, respectively. Names and phylogenetic position of species of wrasses selected for this study are highlighted. The selected species represent differences in cleaning activity (obligate cleaners, facultative cleaners, and non-cleaners). The distribution of the selected species across the tree suggests three independent transitions to cleaning. The number of individuals sequenced for each species is shown in parentheses.

It is clear that differences in the neural transcriptome underlie behavioral variation (Zayed and Robinson, 2012). In fact, coordinated expression of neural genes associated with convergently evolved behavioral phenotypes can be conserved across even distantly related species (Pfenning

et al., 2014; Rittschof et al., 2014; Morandin et al., 2016; Renn et al., 2016, 2018; Young et al., 2019). However, linking transcriptomic and phenotypic variation across species requires an understanding of how gene expression evolves. First, transcriptomes are inherently noisy due to the stochastic nature of the biochemical reactions of transcription (Raser and O'Shea, 2005), and can be highly plastic in response to environmental and physiological fluctuations (reviewed in: de Jong et al., 2019), which can mask relevant evolutionary patterns depending on when and how samples are obtained and analyzed (Liang et al., 2018; Rittschof and Hughes, 2018; Fischer et al., 2021). Second, phenotypes can be altered by the changes in expression of individual genes or entire gene co-expression networks (Hartwell et al., 1999; Stuart et al., 2003; Barabasi et al., 2004; Carroll, 2008; Harrison et al., 2012; Mehta et al., 2021). Thus, evolutionary changes in temporal and spatial gene expression patterns (e.g., tissue or cell expression domains) over the course of ontogeny can result in the loss (addition) of genes from (into) existing gene co-expression networks, potentially influencing the phenotype (Hu et al., 2016; reviewed in: Halfon, 2017). Finally, because like other complex traits, the transcriptome is shaped by stabilizing selection (Gilad et al., 2006; Bedford and Hartl, 2009; Romero et al., 2012) and evolves via neutral or nearly neutral processes (e.g., drift; Khaitovich et al., 2004), testing hypotheses of convergent evolution of phenotypes using transcriptomics data requires examination of differential expression and gene co-expression combined with rigorous phylogenetic comparative analyses and tests of appropriate null hypotheses (Dunn et al., 2013; Young and Hofmann, 2019).

Here, we combine differential gene expression, gene co-expression, and phylogenetic comparative analyses of the neural transcriptome to test two main hypotheses about the evolution and specialization of mutualistic cleaning behavior in *Labridae* wrasses. Targeting brain regions associated with spatial and social cognition, we sequenced RNA extracted from a brain section

containing both putative teleost homologs of the mammalian hippocampus and basolateral amygdala (areas Dl and Dm, respectively; O'Connell and Hofmann, 2011b) from six sympatric wrasse species that vary in mutualistic cleaning behavior including three non-cleaning species, two facultative cleaning species, and one obligate cleaning species (Fig. 1). We then tested two hypotheses: first, we hypothesized that the repeated evolution of facultative cleaning is accompanied by parallel neurotranscriptomic patterns beyond what is expected by chance and after correcting for phylogenetic non-independence among species. Second, we hypothesized that gene expression and co-expression patterns associated with facultative cleaning are shared and specialized in obligate cleaners. An alternative hypothesis is that obligate cleaning evolves by maintaining an early life history state via neoteny. Testing this hypothesis would require comparisons between adult and juvenile sample across species and is beyond the scope of this study.

Results

Gene expression quantification across species

After pre-processing for quality control, 3'tagseq reads were mapped to the Nile tilapia (*Oreochromis niloticus*) coding sequences (Orenil1.0 Ensembl cDNA). Despite the evolutionary distance between the focal species and genomic reference (114 MYA) we obtained expression information for a large number of genes for each species (Supplementary Table S1). In interspecific analyses, especially when multiple species are aligned to non-species-specific reference genomes, interpretation of zero read counts is confounded by the possibility of an inability to align reads due to sequence divergence. As a result, we filtered the gene set such that

that each gene was expressed in at least one individual of each species. The resulting set of 10,411 genes was used for all downstream analyses.

Expression variation across species reveals neural transcriptomic signature of cleaning

We first asked how the neural transcriptome varies across species. We use Discriminant Analysis of Principal Components (DAPC) with the goal of summarizing and discriminating transcriptomic variation across the six species (as in Kenkel and Matz, 2016; Fig. 2A). We find a clear separation among species on the first two discriminant functions. Despite the fact that DAPC is blind to cleaner type, the species align by cleaner type on the first discriminant function. Second,

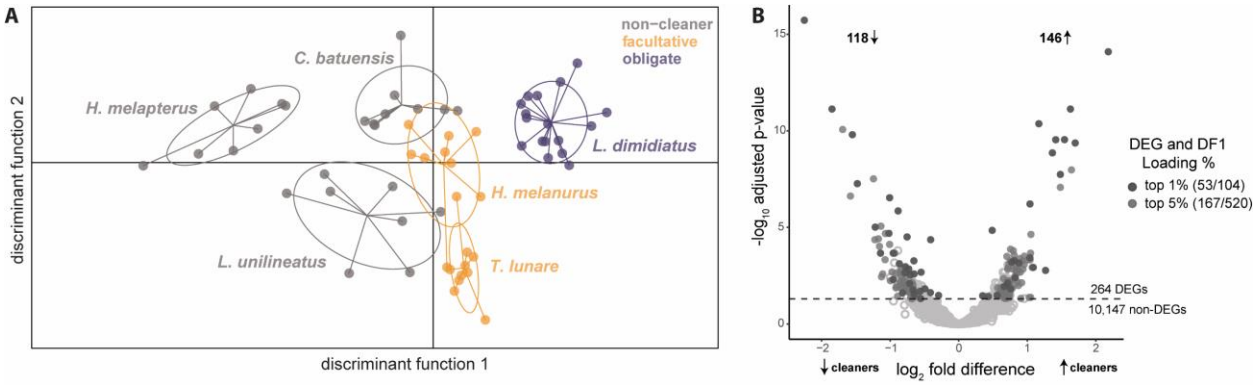


Figure 2. Variation in the neural transcriptome across species (A) and cleaner and non-cleaner species (B). (A) DAPC summarizes the main axis of expression variation across all 10,411 genes using PCA. Next, it minimizes variation within each species and maximizes variation across species using discriminate analysis. Though blind to cleaner affiliation, the species align by cleaner type affiliation (separated by colors) on the first discriminant function. The ellipsoid indicates inertia ellipses or a graphical summary of the point distribution. Vertical and horizontal lines are zero-scaled means of the discriminant functions. (B) 264 genes are differentially expressed between cleaners (both obligate and facultative) and non-cleaner species. Genes differentially expressed were more likely to contribute highly to the separation of species on discriminant function 1 (top 1% and top 5%, A).

we tested the hypothesis that species that engage in mutualistic cleaning at any life history stage (i.e., both obligate and facultative cleaners) exhibit distinct neural transcriptomic patterns from those species that do not clean (non-cleaner species). We identified 264 (2.5%) differentially expressed genes (DEGs) after eliminating species-specific variation and correcting for false discovery rate (adjusted p-value < 0.05), with 146 genes showing increased expression and 118

genes showing decreased expression in cleaners (both obligate and facultative) as compared to non-cleaners (Fig. 2B; Table 1). Genes differentially expressed between cleaner and non-cleaner species (Fig. 2B) were more likely to contribute to separation across species (Fig. 2). Genes with high loadings (top 5% 167/264 or 63.2% of DEGs, ecdf p = 1.3e-159; top 1%: 53/264 or 20% of DEGs, ecdf p = 1.2e-59) on the first discriminant function (Fig. 2A, x-axis) were more likely to differentially expressed between cleaners and non-cleaners (Fig. 2B, grayscale).

Table 1. Number of differentially expressed genes (DEGs, adjusted p-value < 0.05; out of 10,411 total genes in the analysis) in the cleaner type comparisons.

comparison	# of DEGs	# of non-DEGs
cleaners vs non-cleaners	264 (2.5%)	10,147 (97.5%)
facultative vs non-cleaner	342 (3.3%)	10,109 (96.7%)
obligate vs. non-cleaner	284 (2.7%)	10,127 (97.5%)

Facultative cleaner species exhibit parallel gene expression profiles

We tested the hypothesis that the repeated evolution of facultative cleaning is accompanied by parallel patterns of neural gene expression. First, we performed a differential expression analysis between the two facultative cleaner species (*H. melanurus* and *T. lunare*) and the three non-cleaner species (*C. batuensis, H. melapterus*, and *L. unilineatus*) with the aim of identifying genes consistently associated with facultative cleaning. We identified 342 (3.3%) differentially expressed genes (DEGs) after correcting for false discovery rate (adjusted *p*-value < 0.05), with 218 genes showing increased expression and 124 genes showing decreased expression in facultative cleaners as compared to non-cleaners (Fig. 3A). Second, to test explicitly whether the two facultative species (*H. melanurus* and *T. lunare*) are more similar to each other in gene expression profiles than to any other pair of species (including those that are more closely related), we performed a differential gene expression analysis of all pairwise species. We found that the

two facultative species had significantly fewer DEGs (624; 0/14 comparisons, ecdf $p = 0$) compared to all other species pairs (median = 1132.5; Fig 3B).

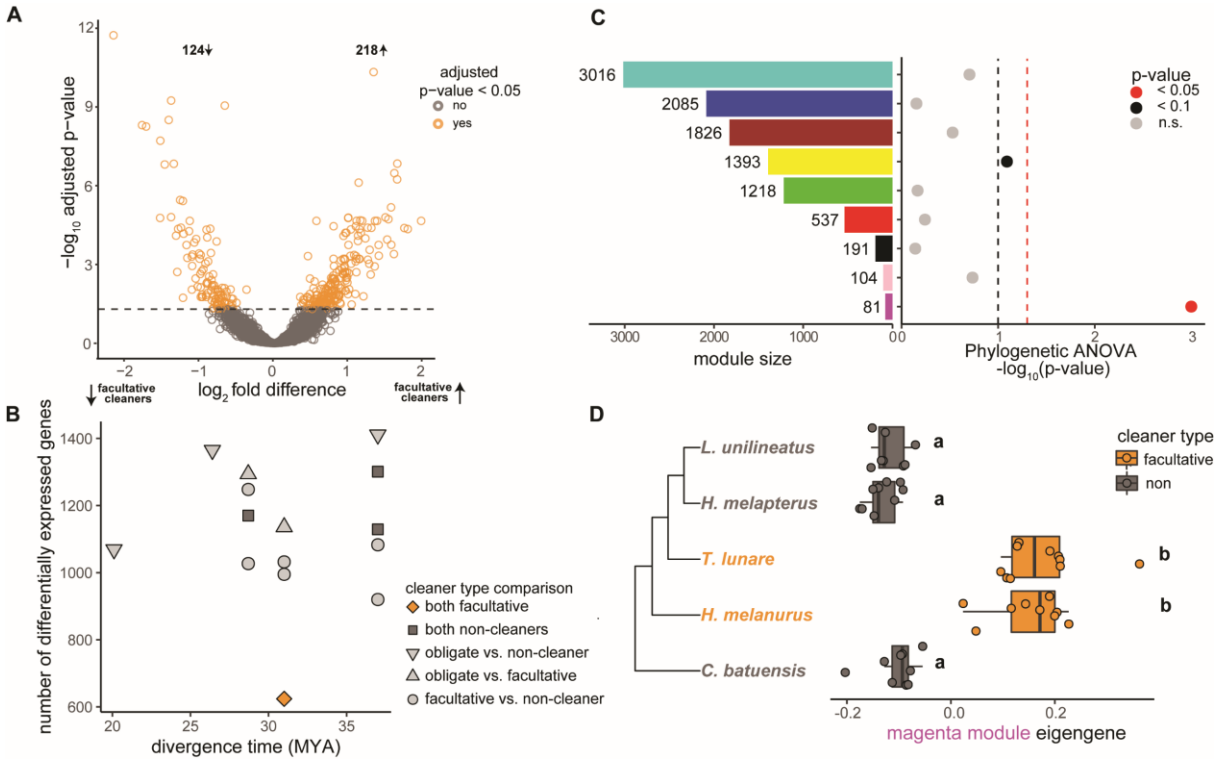


Figure 3. 342 genes are differentially expressed between facultative cleaner and non-cleaner species (A). Pairwise species comparisons between all species pairs found fewer DEGs between the facultative cleaner species as compared to all other species pairs (B: gold diamond). Estimated divergence times (millions of years ago) at each node were obtained from TimeTree (Hedges et al., 2006; Kumar et al., 2017). Combining gene co-expression analysis and phylogenetic comparative analyses finds one co-expression module (magenta) significantly associated with facultative cleaning (C). The number of genes contained in each module is indicated by module size (C). Dashed lines indicated statistical support for the phylogenetic ANOVA at $p < 0.1$ (black) and $p < 0.05$ (red) (C). Species-level co-expression comparisons reveal that the magenta module eigengene co-expression is similar within cleaner type (e.g., between the two facultative cleaner species, gold) and significantly different across cleaner type comparisons (e.g., between non-cleaner and facultative species, grey and gold, respectively) (D). Significance between pairwise comparisons is indicated by the associated letters (D).

Gene co-expression analysis identifies a gene module robustly associated with independent transitions to facultative cleaning

To further test the hypothesis that the repeated evolution of facultative cleaning is accompanied by parallel patterns of neural gene expression we used Weighted Gene Co-expression Network Analysis (WGCNA) of facultative and non-cleaner species. WGCNA of all 10451 genes

yielded nine modules varying in size from 81 to 3016 genes (Fig. 3C). Welch's t-test and ANOVA revealed a number of modules whose co-expression eigengene differs across facultative and non-cleaner types (Supplementary Table S2). We found seven modules that differ significantly across species (Supplementary Figure S1; Supplementary Table S3). Only one module (magenta) differed between facultative cleaners and non-cleaners after accounting for phylogenetic non-independence ($F_{(1,44)}$ = 274.6, p = $2.8e^{-20}$; phylogenetic ANOVA F = 326.7, p = 0.001; Fig. 3C and D; Supplementary Table S2). To assess the probability of identifying a module associated with cleaner type by chance, we used a permutation approach. We resampled genes within each species and performed WGCNA followed by our downstream module-trait association tests. 1000 iterations yielded a total of 7455 pseudo-modules. None of the pseudo-modules were considered significant in the phylogenetic ANOVA of cleaner type. Thus, the probability of discovering, by chance, a module associated with the facultative cleaner phenotype, such as the magenta module, is very low (p < 7.5e-3).

Integrative analysis uncovers candidate genes robustly associated with facultative cleaning

Finally, to identify gene expression signatures robustly associated with evolutionary transitions to facultative cleaning, we integrated analyses of evolutionary divergence (Expression Variance and Evolution model (EVE): Rohlfs and Nielsen, 2015), differential gene expression, and gene co-expression. First, we screened for genes with high intramodular connectivity in modules of interest. Identifying these so called hub genes provides a biologically motivated data reduction approach that has been shown to yield meaningful insight through identification of candidate genes and pathways (e.g. see: Gargalovic et al., 2006; Horvath et al., 2006; Liu et al., 2019). Second, we used EVE analysis (Rohlfs and Nielsen, 2015) to calculate evolutionary

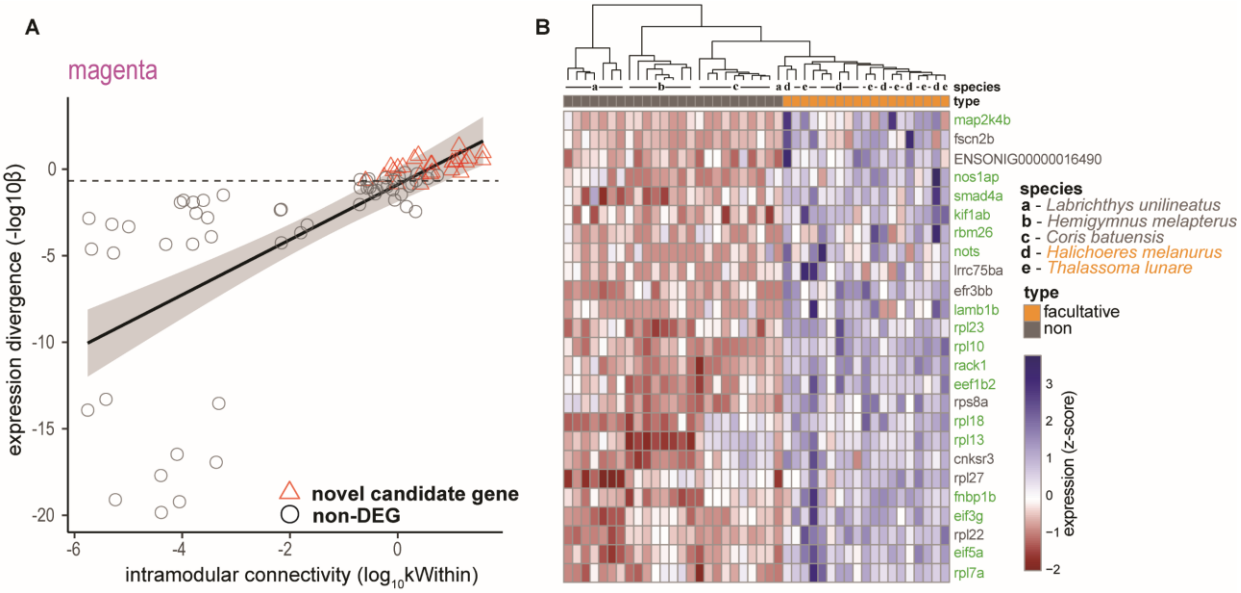


Figure 4. Gene connectivity in the facultative cleaner-associated magenta module is positively correlated with expression divergence across species ($F_{(1,79)}$ = 67.08, p = 3.7e-12) and fold difference in expression between facultative and non-cleaners ($F_{(1,79)}$ = 216.6, p = 2.2e-16) (A). 25 magenta module genes were significantly differentially expressed between facultative and non-cleaner species and had higher than average expression divergence (A: red triangles). Overall, the expression of these 25 candidate genes of facultative cleaning had higher expression in facultative cleaners as compared to non-cleaners (B). Cleaner type and species affiliation of each sample are shown using the color bar and letter annotation, respectively Candidate genes broadly associated with neural development and function, identified from the literature, are shown in green (B; Supplementary Table S6). Bootstrap support for sample clustering (B) is provided in Supplementary Figure S3. The relationship between differential expression (facultative cleaners compared to non-cleaners, *limma*: -$\log_{10}$ adjusted p-value) and evolutionary divergence (EVE: -$\log_{10}$B) is shown in Supplementary Figure S4.

divergence scores ($-\log_{10}\beta_i$) to identify genes that exhibited higher than average interspecific variability indicative of evolutionary divergence. Specifically, genes with $-\log_{10}\beta_i$ greater than the $-\log_{10}\beta$ shared across all 10451 genes ($-\log_{10}\beta_{shared}$ = -0.675) were considered to have high interspecific variability. We found that the more highly connected hub genes of the facultative cleaning-associated magenta module were also differentially expressed ($F_{(1,79)}$ = 238.8, p = 2.2e-16) and exhibited high evolutionary divergence scores ($F_{(1,79)}$ = 67.08, p = 3.7e-12; Fig. 4A). Specifically, we identified 25 differentially expressed magenta module genes with high interspecific variability as previously undescribed candidate genes associated with facultative cleaning (Fig. 4B; Supplementary Table S4). We were able to annotate 24 of these 25 genes with a molecular or cellular function. Of note, 17 (68%) of these genes have been broadly implicated in neural development and function (highlighted in green font Fig 4B; Supplementary Table S4). The relationship between evolutionary divergence ($-\log_{10}\beta$) and differential expression between facultative cleaners and non-cleaners ($-\log_{10}$ adjusted p-value) for all genes is shown in Supplementary Fig. S4).

Facultative and obligate cleaner species show concordant expression and co-expression patterns

Next, we asked whether the one obligate cleaner species (*L. dimidiatus*) in our analysis displayed gene (co-)expression patterns that were concordant with those of the facultative cleaners described above. To do this, we first compared the obligate cleaner (*L. dimidiatus*) and the non-cleaner species (*C. batuensis*, *H. melapterus*, and *L. unilineatus*) and found 284 (2.7%) DEGs with 205 genes showing increased expression and 79 genes showing decreased expression in the obligate cleaner species as compared to non-cleaners. We then compared differential gene expression between facultative cleaners and non-cleaners with differential expression in the

obligate cleaner and non-cleaners (Supplementary Figure S2A). Specifically, we examined the intersection of four gene sets (Tables 1 and 2): (1) genes that were differentially expressed and directionally concordant in both differential expression analyses (342 DEGs in facultative versus non-cleaners and 284 DEGs in obligate versus non-cleaners); (2) genes differentially expressed between facultative and non-cleaners (342 DEGs) and directionally concordant (but not significant) between obligate and non-cleaners (10,127 non-DEGs); (3) genes differentially expressed between facultative and non-cleaners (342 DEGs) and directionally discordant (but not significant) between obligate and non-cleaners (10,127 non-DEGs, as above in 3); and (4) differentially expressed in both analyses (342 DEGs in facultative versus non-cleaners and 284 DEGs in obligate versus non-cleaners, as above in 1), but directionally discordant. Comparing differential expression analyses, we found that 44 of the 342 DEGs (12.9%) between facultative cleaner and non-cleaner species were also DEGs in the concordant direction between the obligate and non-cleaners, which is significantly more than expected by chance (Fig. 5A and B; Tables 1 and 2: 44/284 DEGs in the obligate versus non-cleaner comparison; ecdf p = 5.1e-19). The expression of an additional 188 of these 342 DEGs (54.9%), while not significant, were concordant in the obligate versus non-cleaner species (Fig. 5A and B; Tables 1 and 2: 188/10,127 non-DEGs in the obligate versus non-cleaner comparison; ecdf p = 1). Finally, only 5 of these 342 DEGs (1.5%) were significantly differentially expressed, but directionally discordant in the obligate versus non-cleaner comparison (Fig. 5A and B; Tables 1 and 2: 5/284 DEGs in the obligate versus non-cleaner comparison; ecdf p =1) and 105 of these 342 DEGs (30.7%) were discordant between obligate versus non-cleaner species (Fig. 5A and B; Tables 1 and 2: 105/10,127 non-DEGs in the obligate versus non-cleaner comparison; ecdf p = .91). In other words, cleaner-level differential expression comparisons supported concordance of gene expression patterns in facultative and

obligate cleaners (Fig. 5A and B, dark blue). However, we did not find evidence for increased similarity in overall gene expression at the species level between facultative cleaners and the obligate cleaner species (Fig. 3B, upward triangles).

To further characterize gene co-expression changes associated with facultative cleaning or cleaning in general (i.e., obligate and facultative cleaners) we assessed differences in gene connectivity of genes from our nine focal co-expression modules in the obligate cleaner, the facultative cleaners, and the non-cleaner. Because the magenta module was uniquely associated with facultative cleaners (Fig. 3C and D), we focused on the connectivity of magenta module genes with all other genes in the transcriptome (kTotal) in the three cleaner types (all other modules are shown in Supplementary Figure S5 and Supplementary Table S7). We found that connectivity of magenta module genes was highest in the obligate cleaner and lowest in the non-cleaner species (Fig. 5D; $F_{(2,240)} = 68.1$; $p = 3.8E\text{-}24$) and differed significantly between all pairwise cleaner types (non-cleaners versus facultative cleaners: t ratio = -2.34, adjusted p-value = 1.92E-2; non-cleaners versus the obligate cleaner: t ratio = -8.48, adjusted p-value = 7.31E-17; facultative cleaners versus the obligate cleaner: t ratio = -6.13, adjusted p-value = 1.31E-9). Further, we found that gene connectivity in facultative cleaners was significantly correlated with connectivity in both obligate and non-cleaner species (Fig. 5E; non-cleaners versus facultative cleaners: Spearman's $\rho = 0.37$; adjusted $p = 0.001$; facultative cleaners versus the obligate cleaner: Spearman's $\rho = 0.44$; adjusted $p = 7.8E\text{-}5$). One other module, the brown module (Fig. 3C), exhibited a similar pattern of connectivity between cleaner types; however, effect sizes of connectivity differences and gene connectivity correlations between cleaner types were generally weaker, further statistical significance may reflect, in part, increased power associated with larger module size (i.e., 1826

genes in the brown module versus 81 genes in the magenta module; Supplementary Fig. S4; Supplementary Table S7).

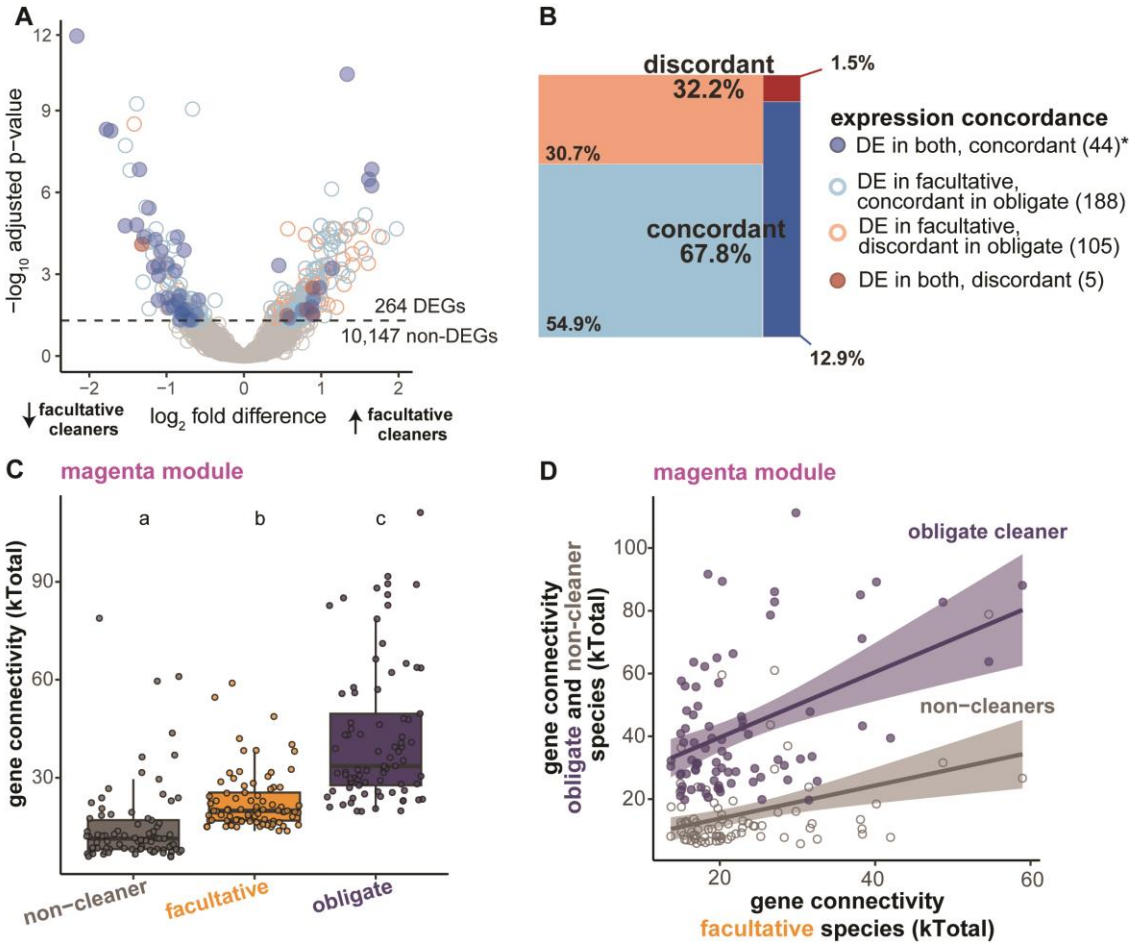


Figure 5. Concordance of expression (A and B) and co-expression patterns (C-D) in facultative and obligate cleaner species as compared to non-cleaner species. 67.8% of the 264 facultative versus non-cleaner DEGs, were expressed in the concordant direction in obligate versus non-cleaners (A and B, blues). We found more genes than expected by chance (12.9%) were differentially expressed between facultative cleaner and non-cleaner species were also differentially expressed and directionally concordant in expression between obligate and non-cleaner species (A, dark blue; 44 genes, p = 5.1e-19). Gene connectivity of the facultative cleaning associated magenta module (Fig. 3C and D) genes was highest in the obligate cleaner and lowest in non-cleaners (C; $F_{(2,240)}$ = 68.1; p = 3.8E-24) and differed across all pairwise combinations of cleaner types (Supplementary Table S7). Further, gene connectivity in facultative cleaners was significantly correlated with connectivity in both the obligate cleaner and non-cleaners (D; non-cleaners versus facultative cleaners: Spearman's ρ = 0.37; adjusted p = 0.001; facultative cleaners versus the obligate cleaner: Spearman's ρ = 0.44; adjusted p = 7.8E-5). Connectivity comparisons and associated statistics for all additional modules and non-cleaners versus the obligate cleaner are provided in Supplementary Figure S5 and Supplementary Table S7. The dashed line indicates adjusted p-value < 0.05 (A Significance between pairwise comparisons is indicated by the associated letters (C). Shading on the regression indicates the standard error (D).

Table 2. Directional concordance of gene expression in facultative and obligate cleaners as compared to non-cleaners. Of the 342 differentially expressed genes (DEGs) identified in facultative cleaners (compared to non-cleaners), those with a concordant differential expression patterns in obligate cleaners were significantly over-represented (highlighted by an asterisks).

	DEGs (both analyses)		Non-DEGs obligate vs. non-cleaner	
	Concordant	Discordant	Concordant	Discordant
Overlapping DEGs in obligate vs. non-cleaner comparison	**44***	5	188	105
Percentage of facultative vs. non-cleaner DEGs (n=342)	**12.9%**	1.5%	54.9%	30.7%
Fraction of obligate vs. non-cleaner DEGs (284)	15.5%	1.8%	-	-
Fraction of obligate vs. non-cleaner non-DEGs (10,127)	-	-	1.9%	1.0%

The obligate cleaner species L. dimidiatus shows no evidence of expression specialization compared to the facultative cleaner species

To test the hypothesis that obligate cleaning is a specialization of facultative cleaning, we compared differential gene expression between facultative cleaners and non-cleaners with differential expression in the obligate cleaner and non-cleaners and obligate cleaner and facultative cleaners. Expression differences (Supplementary Figure S6A and B) between obligate (*L. dimidiatus*) and facultative cleaners (*H. melanurus* and *T. lunare*) did not support the hypotheses that obligate cleaning is a specialization of facultative cleaning. Specifically, DEGs between facultative and non-cleaners are not specialized (i.e., more differentially expressed) in the obligate cleaner (Supplementary Figure S6B).

Discussion

In the present study, we took advantage of the repeated, independent evolution of mutualistic cleaning behavior in wrasses to test our main hypothesis that convergent evolution of complex behavioral phenotypes is facilitated – in part – by the similar modifications to conserved

neurotranscriptomic mechanisms. Using 3'tag-based RNA sequencing, we quantified and compared gene expression and co-expression in a brain section containing both the putative teleost homologs of the mammalian hippocampus and basolateral amygdala, broadly associated with spatial and social cognition (O'Connell and Hofmann, 2011a, 2012; Yang and Wang, 2017), in six species of *Labridae* wrasses that vary in mutualistic cleaning behavior (Fig. 1). By combining gene expression and co-expression analyses with phylogenetic comparative analyses and tests against appropriate null hypotheses, we first ask whether species with distinct cleaner phenotypes vary in transcriptome. Second, we ask whether neural repeated evolution of facultative cleaning is accompanied by parallel neurotranscriptomic patterns beyond what is expected by chance and after correcting for phylogenetic non-independence among species. Third, we investigate whether gene expression and co-expression patterns associated with facultative cleaning are shared and specialized in obligate cleaners. Comparisons of gene expression and co-expression across species and cleaner types indicate that neural gene expression varies between cleaner and non-cleaner species, provide strong support for shared neuromolecular basis of facultative cleaning, and show limited support for maintenance, but not specialization, of these neural transcriptomic patterns in the obligate cleaner species. Below we discuss these results and their implications for our understanding of mutualistic cleaning and cooperative behavior more generally as well as the mechanistic bases of convergent evolution of complex behavioral phenotypes.

Neural transcriptomes differ among cleaning and non-cleaning species

Characterizing the overall patterns of transcriptome variation across species revealed a strong signature of cleaner type. Using discriminant analysis of principal components (DAPC) we aimed to characterized transcriptomic variation that best highlighted species-level variation. While this

analysis was successful in distinguishing expression patterns among species, cleaner types also emerged. Blind to cleaner affiliation, we found species separated by cleaner type rather than evolutionary relatedness on the first discriminant function (Fig. 2A). Notably, the three cleaning species (i.e., one obligate cleaner and two facultative cleaners) each represent an independent evolutionary transition to cleaning. In addition, we identified 264 genes differentially expressed in cleaners (both obligate and facultative) compared to non-cleaners (Fig. 2B). Genes differentially expressed between cleaners and non-cleaners were more likely to contribute to species-level variation (Fig. 2B). Combined these analyses support the hypothesis that gene expression in the brain regions associated with spatial and social cognition differ between species that engage in mutualistic cleaning as compared with species that do not clean regardless of evolutionary history.

Parallel gene expression and co-expression patterns underlie independent the evolution of facultative cleaning

Our analysis of neural gene expression and co-expression patterns between species pairs and facultative cleaners and non-cleaners (Figs. 3 and 4) provides robust support for the hypothesis that shared neurotranscriptomic mechanisms underlie repeated evolution of facultative cleaning. First, we found that the facultative cleaning species exhibited significantly more similar neural transcriptomic patterns than any other pair of species regardless of evolutionary distance (Fig. 3B). Second, comparing combined facultative versus non-cleaners species, we found many genes differentially expressed, indicating high concordance of directional changes in expression in the facultative cleaner (342; Fig. 3A). Interestingly, the differential expression analysis revealed a strong directional bias between cleaner and non-cleaner comparisons. Two-thirds of the differentially expressed genes have increased expression in the facultative cleaner species (Fig.

3A). Similarly, 72% of the genes differentially expressed between obligate cleaners and non-cleaners exhibit increased expression in the obligate cleaner (Supplementary Fig. S2A). However, we did not find similar bias in gene activation when comparing facultative and obligate cleaner species (Supplementary Fig. S2B). Thus, our results suggest that the evolution, but not specialization, of complex mutualistic cleaning behavior is associated with increased gene activation in the putative teleost hippocampus and basolateral amygdala homologs targeted in our study. Given the role of stabilizing selection in transcriptome evolution (Khaitovich et al., 2004; Gilad et al., 2006; Bedford and Hartl, 2009; Romero et al., 2012) and the strong species-specific expression patterns (Supplementary Figure S1B and S7A), our finding of consistent directional bias in expression variation is notable. While any functional interpretation of this bias is necessarily speculative at this point, it is suggestive that many of the novel cleaner candidate genes we discovered are associated with synaptic function and plasticity, neuronal growth, and neurite elongation (Supplementary Table S5). Increased activity of these genes in mutualistic cleaning species could reflect enhanced spatial and social cognition in cleaners reflected compared to non-cleaners (Mehta, 2015). Alternatively, it may reflect the molecular processes associated with a decrease in central inhibitory control (*sensu* Roeder, 1935), which could enable cleaners to approach much larger and, often, predatory client fish.

Beyond similarity in expression of individual genes in the facultative cleaner species, using gene co-expression analysis (WGCNA), we found a gene co-expression module significantly associated with facultative cleaning after taking phylogenetic non-independence among species (Fig. 3C and D). While this 81 gene module is relatively small, compared to the other modules defined by WGCNA (Fig. 3C), our permutation analyses revealed that discovering a module of any size (greater than the minimum modules size of 50 genes) associated with facultative cleaning

after accounting for phylogenetic non-independence is highly unlikely. Moreover, expression of most co-expression modules (7 of the 9) diverge across species, underscoring the strong species signal present in the gene expression and co-expression patterns (Fig. 3C; Supplementary Figures 1 and 8) and further highlighting the unique transcriptomic pattern captured by the magenta module.

Integrative analysis uncovers candidate genes robustly associated with facultative cleaning

In general, the neural and molecular underpinnings of mutualistic cleaning behavior are poorly understood. However, some candidate neuroendocrine pathways have been implicated, with most studies focusing on females of the obligate cleaner species *L. dimidiatus* and finding evidence for a role of nonapeptides (arginine vasopressin, oxytocin), biogenic amines (serotonin, dopamine), and sex steroid hormones (estradiol, testosterone) in regulating different aspects of this behavior (for reviews see: Soares et al., 2010; Soares, 2017). It is important to note here that only the obligate cleaner, *L. dimidiatus*, engages in regular and frequent cleaning bouts throughout life, while the facultative cleaning species included in our analysis (*T. lunare and H. melanurus*) display cleaning behavior only as juveniles (Baliga and Law, 2016). Even though the candidate genes and pathways suggested by previous studies may well be critical in regulating frequency of cleaning or specific behavior during acute cleaning bouts at least in some cleaner species, we did not expect them to emerge in our analysis of the repeated evolution and specialization of mutualistic cleaning behavior, which indeed was the case. In fact, by sampling free swimming fish engaged in a variety of activities our study was designed to minimize any effects of ongoing behavior on our analysis, such that we could identify patterns of variation in *constitutive* expression of genes associated with the repeated evolution of facultative cleaning. Towards this goal, we

integrated independent analyses including differential gene expression and co-expression with gene connectivity and expression diversification. We found that the highly-connected hub genes from the facultative cleaner associated magenta module were significantly more likely to be differentially expressed between facultative cleaners and non-cleaners and exhibited high evolutionary divergence (Fig. 4A). Specifically, twenty-five facultative cleaning associated magenta module genes were both differentially expressed between facultative cleaners and non-cleaners and exhibited high evolutionary divergence relative to all other genes in the genome (Fig. 4A). All 25 genes exhibit increased expression in the facultative cleaner species, again indicating a bias towards increased gene expression in the evolution of mutualistic cleaning behavior. Two-thirds of these genes are broadly associated with neural development and function (Fig. 4B). While, to our knowledge, none of these genes have been previously associated the facultative cleaning, 80% (20/25) of these facultative cleaning associated genes were implicated in cleaning behavior in the obligate cleaner (*L. dimidiatus*; Kang et al., 2023). Importantly, detailed descriptions of ecology, life history, and morphology of Labridae species clearly indicate that there is not a single such attribute – other than cleaning behavior – that is more similar between the two facultative cleaner species, *T. lunare and H. melanurus*, than any other species in our analysis (Randall et al., 1988; Krattinger, 2016) (Supplementary Table S5). This observation suggests that the transcriptomic similarities we have discovered are indeed due to the shared behavioral strategy.

Evidence of shared, but not specialization, of expression and co-expression among cleaner types

Because species sampling was limited to one obligate cleaner species, *L. dimidiatus*, we were unable to integrate the obligate cleaner into our phylogenetic comparative analysis of cleaning-associated gene expression and co-expression patterns. However, we can ask whether

facultative and obligate species share similar gene expression and co-expression pattern. In fact, we found gene expression patterns in the facultative and obligate cleaners to be quite concordant as compared to the non-cleaner species. 68% of the genes differentially expressed in facultative cleaners were expressed in the concordant direction in the obligate cleaner (Fig. 5A and B). Further, we found significantly more genes than expected by chance to be differentially expressed in the concordant direction in both facultative and obligate species as compared to non-cleaners (Fig. 5B). These results suggest that the molecular underpinnings of cleaning behavior are shared, at least in part, between obligate and facultative cleaners.

To ask whether co-expression patterns were similarly shared, we examined co-expression of the genes contained in the magenta "cleaning" module (which we had previously found to be associated with facultative cleaning) and discovered that, compared to both facultative and non-cleaners, these genes were most highly connected with rest of the transcriptome in obligate cleaners (Fig. 5C and D). Given that only obligate cleaners regularly and frequently display cleaning as adults (the life stage at which we collected our samples off the reef), and given that any cleaning activity by facultative cleaners had ceased by the time they became adults (Baliga and Law, 2016), we speculate that the molecular and physiological processes supported by these genes might be more fully integrated with overall neural and behavioral activity in obligate cleaners.

Because increased dependency on cleaning behavior is linked to increased behavioral, cognitive, and morphological specialization in wrasses (Barbu et al., 2011; Gingins and Bshary, 2016; Baliga and Mehta, 2019), we asked whether a similar specialization will be evident in the neural transcriptome of the obligate cleaner species. We found no evidence for an increased specialization of cleaning-related gene expression. Though directionality of gene expression

differences (relative to non-cleaners) was generally concordant between facultative and obligate cleaners (Fig. 5A and B), genes differentially expressed in facultative cleaners were not more differentially expressed in the obligate cleaner (Supplementary Figure S6B). Alternatively, because evolutionary transitions between juvenile and adult cleaners (obligate or facultative) are phylogenetically correlated it has been hypothesized that adult cleaning evolved from a juvenile cleaning state (Baliga and Law, 2016), perhaps by maintaining an early life history state via neoteny. However, testing the hypothesis that obligate cleaning reflects a heterochronic shift (i.e., maintenance of an early life history stage in the obligate cleaner) requires gene expression comparisons across developmental time points, which is beyond the scope of the present study.

Conclusions

Overall, our analyses reveal a strong evolutionary signal in neuromolecular gene expression across the six species of wrasse, suggest that mutualistic cleaning, broadly, is associated with an increase in neural gene expression, and provides robust support for the hypothesis that independent evolution of facultative cleaning is associated with shared neurotranscriptomic mechanisms. Further, we find that gene expression and co-expression patterns are conserved in an obligate cleaner; however, the specialized cleaning behavior and correlated cognitive phenotypes cannot be explained by increased specialization in expression of genes evolutionarily associated with cleaning behavior.

Materials and Methods

Study species and their ecological and life history attributes

Six species of wrasses (Labridae) were selected for this study: *Labroides dimidiatus* (obligate cleaner), *Halichoeres melanurus* (facultative cleaner), *Thalassoma lunare* (facultative cleaner), *Hemigymnus melapterus* (non-cleaner), *Labrichthys unilineatus* (non-cleaner) and *Coris batuensis* (non-cleaner). Species were selected after Gingins and Bshary (2016) because they spread across the Labridae phylogenetic tree and represent differences in cleaning activity (Fig. 1; obligate cleaners, facultative cleaners, and non-cleaners), which in turn are not correlated with any other, potentially confounding, ecological, morphological, or life history attributes. In fact, detailed comparisons of these and other wrasse species (Randall et al., 1988; Krattinger, 2016) have demonstrated that the facultative and obligate cleaner species do not share any attributes that the other species in our study lack. Estimated divergence times were obtained from TimeTree (Hedges et al., 2006; Kumar et al., 2017). Phylogeny was redrawn from Baliga and Law (2016). Briefly, the newick formatted tree plotted using ggtree (Yu et al., 2017, 2018). Cleaner type assignments were based on Baliga and Law (2016) Table A.7. Baliga and Law (2016) distinguish juvenile and facultative cleaners. For our purposes we consider species that clean as juveniles or occasionally or opportunistically as adults both examples of facultative cleaning. We made one modification from Baliga and Law (2016), *H. melanurus* is not listed as a facultative cleaner; however, individuals of these species have been observed cleaning (reported in Gingins and Bshary, 2016).

Sample collection, tissue processing, RNA extraction, and 3'Tag sequencing

Samples were collected in August 2015 at Lizard Island Research Station, Great Barrier Reef, Australia. Employing hand and barrier nets (8 × 2 m; 1 cm mesh for large fish; and 2 m × 1 m, 5 mm mesh size for small fish) scuba divers collected fish in plastic zip bag with sufficient water. After each single capture, fish were handed to a researcher on the boat in order to process

the samples. Because these populations of the target species are female-biased and we did not want to decrease statistical power by having sex as a co-factor, no males were included in the analysis. The female fish were measured (Supplementary Table S5) and killed immediately on the boat by cervical transection and whole heads were transferred into 30 mL conical tubes with RNA*later* (ThermoFischer, Waltham, MA, USA) to preserve the integrity of the RNA as logistical obstacles did not allow the use of dry ice or liquid nitrogen for flash-freezing the samples on the boat. The elapsed time between capture and death ranged from 6 to 28 minutes. After 3 hours (on average) in RNAlater, the samples were transferred into O.C.T. Compound (Sakura, Torrance, CA, USA) and flash frozen in liquid nitrogen. At completion of sampling and preparations, all samples were shipped on dry ice to The University of Texas at Austin for further processing.

We discovered that the short time the samples were stored in RNAlater caused the tissue to become too fragile to reliably micro-dissect specific brain regions. Instead, frozen heads were sectioned on a cryostat microtome into 300µm thick slices. The second most rostral slice of the brain, containing pallial areas Dl and Dm (i.e., the putative teleost homologs of the mammalian hippocampus and basolateral amygdala, respectively (O'Connell and Hofmann, 2011), along with partial portions of septal and striatal territories, was collected and stored in 200ul of ice-cold homogenization buffer with 1-thioglycerol (Promega Corporation, Madison, WI, USA) and stored at -80C until further processing. Total RNA was extracted from each sample using Maxwell 16LEV simplyRNA tissue kit (Promega Corporation, Madison, WI, USA) following manufacturer instructions, including DNase treatment. RNA samples were then eluted into 40uL of nuclease-free water. RNA integrity was confirmed using an Agilent BioAnalyzer and sequencing libraries constructed using 3'tag sequencing approach (Lohman et al., 2016) for sequencing on an Illumina HiSeq 2000 instrument (Illumina, San Diego, CA, USA). RNA quality control, library

561 construction, and sequencing were performed by the University of Texas at Austin Genome
562 Sequencing and Analysis Facility.
563
564 *Read preprocessing and alignment and gene expression quantification*
565 3'tagseq raw reads were preprocessed prior to alignment using the following pipeline. Briefly,
566 custom perl scripts (after Meyer et al., 2011) using the FASTX-toolkit (Hannon, 2010) and
567 CUTADAPT v. 2.8 (Martin, 2011) were used to remove reads with a homo-polymer run of "A"≥8
568 bases, retain reads with minimum 20 bases, removal PCR duplicates – defined as sequences
569 sharing the same degenerate header and 20 bases of sequence – and filtered for quality (Phred
570 quality score > 20 for 90% of the nucleotides). Although a *L. dimidiatus* genome was recently
571 published (Kang et al., 2023), it would not be appropriate to use as a reference for read alignment
572 here due to the highly variable evolutionary distances (0 to ~37 million years, according to
573 http://timetree.org/) of the remaining five species in our analysis relative to *L. dimidiatus*, which
574 could distort any quantification of gene expression across species. Instead, we first aligned
575 preprocessed reads using the Burrows-Wheeler Aligner (bwa-mem; Li and Durbin, 2009) to a *de*
576 *novo* assembled brain transcriptome of the ocellated wrasse, *Symphodus ocellatus* (Nugent et al.,
577 2016), which is ~58 million years diverged from the species in our analysis (according to
578 http://timetree.org/). While mapping percentages averaged 19.4%, few reads mapped to contigs
579 annotated with known gene ids (average gene mapping percentage, 1.3%), precluding downstream
580 interpretation of any results. We therefore mapped reads to the more distant (~115 million years,
581 according to http://timetree.org/) but much better annotated *Oreochromis niloticus* genome
582 (Orenil1.0 Ensembl cdna). As expected, mapping percentages were relatively low, yet very
583 consistent (4.7% to 5.2%). All downstream analyses were thus carried out using gene counts from

584 *O. niloticus* read mapping. Reads mapped to the *O. niloticus* were converted to counts using
585 samtools (idxstats; Li et al., 2009) and TMM normalized (Robinson and Oshlack, 2010) using R
586 package NOISeq (Tarazona et al., 2011). TMM normalized expression values were used for all
587 downstream analyses. We filtered the gene set such that that each gene was expressed in at least
588 one individual of each species, which yielded a set of 10,411 genes. We also applied a more
589 stringent filter, such that only genes expressed in at least 4 individuals of each species were kept,
590 which yielded a set of 3,686 genes. Because the variance structure of both gene sets was very
591 similar (not shown), we used the more inclusive gene set for all downstream analyses.
592
593 *Neural transcriptomic variation across species*
594 To summarize overall variation in neural gene expression across species, we use
595 Discriminant Analysis of Principal Components (DAPC, using the R package *adegenet*: Jombart,
596 2008; Jombart and Ahmed, 2011; after Kenkel and Matz, 2016). DAPC combines unsupervised
597 (PCA) and supervised (DA) machine learning to classify samples. In DAPC, we first perform a
598 principal components analysis to summarize the major axes of variation in the neural transcriptome
599 across all samples. We then perform a discriminant analysis with the goal of discriminating among
600 the six species blind to cleaner type. Thus, the DA minimizes variation within each species and
601 maximized variation among species. We determined the number of PCs to retain using *a*-score
602 optimization, which measures the trade-off between power of discrimination and over-fitting. The
603 optimal *a*-score was achieved by retaining 3 PCs. Significance in overlap of DAPC loadings and
604 differentially expressed genes between facultative and non-cleaners was determined by a
605 hypergeometric test using the R function (*phyper*).
606

Phylogenetic comparative analysis of expression divergence (EVE model)

The Expression Variance and Evolution (EVE) model was used to characterize evolutionary divergence in expression of each gene (Rohlfs and Nielsen, 2015). The EVE model, parameterizes the ratio of intra- and interspecific variation of each gene across the phylogeny. Under stabilizing selection, this ratio (β_i) for any specific gene should be equivalent to the mean β_{shared} for all the genes across the transcriptome. $\beta_i > \beta_{shared}$ indicates high relative intraspecific variation associated with plasticity or diversifying selection within species. $\beta_i < \beta_{shared}$ indicates high divergence in expression across species associated with lineage-specific directional selection. Gene-specific β values were converted to "expression divergence scores" using a $-\log_{10}$ transformation. TMM-normalized counts of each gene shared across all six species and the phylogeny of the six wrasse species included in our study were used as input data for the EVE model. Gene-specific β_i and transcriptome β_{shared} were calculated for all six species and for the five non-cleaner and facultative cleaner species. To assess the effect of the obligate cleaner *L. dimidiatus* on gene divergence scores, we performed an EVE analysis on both all six species and on the five non-cleaner and facultative cleaner species. Spearman's rank correlation of gene-specific β_i values calculated with and without the obligate cleaner demonstrated that the results were highly concordant (Supplementary Figure S7; Spearman's $\rho = 0.87$, $p < 2.2e\text{-}16$).

To assess the efficacy of the EVE model to target genes exhibiting different patterns of variation and diversification across species, we examined expression co-variance patterns among individuals and species for different subsets of genes including all genes and the genes with the top and bottom 1% expression divergence scores (Supplementary Figure S8A). Not surprisingly, the genes with highest (top 1%) $-\log_{10}\beta_i$ values (i.e., genes whose expression is both variable and highly divergent between species) resulted in a robust hierarchical clustering of individuals by

630 species (Supplementary Figure S8B). Conversely, when we selected the genes with the lowest
631 (bottom 1%) $-\log_{10}\beta_i$ values (i.e., genes whose expression is variable but not divergent between
632 species), hierarchical clustering did not reveal any robust patterns either by species or any other
633 attribute (Supplementary Figure S8C). We calculated Pearson's correlations in gene expression
634 between all pairwise individuals. Individuals were clustered using 1-correlation and correlations
635 were plotted in a heatmap using R package pheatmap (Kolde, 2012). Cluster support values were
636 generated by using R package pvclust (1000 permuations; Suzuki and Shimodaira, 2006).
637
638 *Differential expression analysis of cleaner types and species*
639 Gene exhibiting differences in expression across cleaner types and species were identified
640 using the R package *limma* (linear models for microarray data: Smyth, 2005). We used several
641 different *limma* analyses. The broad cleaner (i.e., both obligate and facultative cleaner species)
642 versus non-cleaner comparisons were modeled as (obligate + facultative)/2 – non-cleaner while
643 blocking the random effect of species. For the facultative cleaner versus non-cleaner, obligate
644 cleaner versus non-cleaner, and obligate versus facultative cleaner comparisons, we compared
645 cleaner types while blocking the random effect of species. For species-level comparisons, we
646 calculated the number of DEGs for each pairwise combination of species. To test the hypothesis
647 that facultative species are more similar in gene expression than is expected by chance, we
648 compared the number of DEGs between the two facultative species to the distribution of
649 differentially expressed gene of all the remaining species pair combinations using *ecdf* in R as
650 above. For all analyses, differential expression was defined by an FDR adjusted p-value < 0.05
651 unless otherwise indicated. We used TMM-normalized counts of the 10,411 genes as the input
652 data set for all limma contrasts.

Phylogenetic Comparative Analysis of Gene Co-expression

To capture genes with correlated expression variation across individuals, species, and cleaner types, we performed a Weighted Gene Co-expression Network Analysis (WGCNA) (Langfelder and Horvath, 2008) including the genes shared across all species. WGCNA clusters genes by expression similarity and summarizes gene co-expression as module eigengenes (i.e., the first principal component of all the genes in each co-expression module). Each eigengene is the linear combination of gene expression values that explains the most variation in the expression of the genes contained in the module. For all analyses, WGCNA was performed with a minimum modules size of 50 and soft power thresholds were determined using WGCNA's softPower function.

To identify modules associated with facultative cleaning, we performed WGCNA including the two facultative cleaner species and the three non-cleaner species (Fig. 1). We assessed changes in gene co-expression across species and cleaner types using a Welch's t-test (due to unequal variances) or ANOVA depending on whether two or more groups were compared. For pairwise comparisons (e.g., among species) we used Tukey's honest signal difference post hoc test for ANOVAs with F-statistics with significance at $p < 0.05$. For cleaner type comparison between facultative and non-cleaner species, we confirmed associations between cleaner type and module eigengene taking phylogenetic non-independence into account using a phylogenetic ANOVA. Phylogenetic ANOVA was performed with the R package phytools (Revell, 2012). A phylogenetic ANOVA could not confirm whether any modules associate with obligate cleaning due to the inclusion of only one representative obligate cleaner species (*L. dimidiatus*). To determine the probability of identifying WGCNA modules associated with any subset of species

(e.g., facultative cleaners) by random chance we used a permutation approach. Specifically, to maintain phylogenetic structure but disrupt gene co-expression, TMM normalized gene counts were sampled without replacement within each species. For each sampling iteration (n = 1000) we performed WGCNA. The resultant modules were tested for difference among species and cleaner types using an ANOVA or Welch's t-test and phylogenetic ANOVA (as above). Significance of modules of interest was determined by comparing *p*-values from the observed WCGNA to the distribution of *p*-values generated in the permutation analysis.

To identify candidate (hub) genes of interest we quantified intramodular connectivity (kWithin: the connectivity of each gene with all other genes in a focal module) of all genes from modules of interest – i.e., significantly associated with cleaner type after accounting for phylogenetic non-independence – using WGCNA's *intramodularConnectivity* function.

Candidate genes and gene functions robustly associated with facultative cleaners

Genes robustly associated with facultative cleaners across analyses were identified as those differentially expressed between facultative and non-cleaner types (adjusted p-value < 0.05) and contained in modules of interest (i.e., significantly associated with cleaner type after accounting for phylogenetic non-independence). Using the primary literature, we highlight known, relevant functional associations of each candidate gene (Supplementary Table S4)

Concordance of gene expression and co-expression across facultative and obligate cleaner species

To determine whether gene expression patterns in the obligate cleaner species (*L. dimidiatus*) are concordant with patterns associated with facultative cleaning, we compared differential gene expression between facultative cleaners and non-cleaners with differential

699 expression in the obligate cleaner and non-cleaners. Significance in overlap of all sets was
700 determined using a hypergeometric test using the R function (*phyper*).

701 To ask whether gene connectivity persist across cleaner types, we performed WGCNA
702 analysis including only species from each cleaner type (i.e., the obligate cleaner *L. dimidiatus*, the
703 facultative cleaners *T. lunare* and *H. melanurus* only, or the non-cleaners *H. melapterus*, *L.*
704 *unilineatus*, and *C. batuensis*). We compared gene connectivity (kTotal) of genes from focal
705 modules (Fig. 3C and D) in WGCNA performed on non-cleaners, facultative cleaners, and the
706 obligate cleaner alone. kTotal was calculated using WGCNA's intramodularConnectivity function
707 and measures connectivity of each gene with all other genes in the transcriptome. We asked
708 whether gene connectivity of each module different among cleaner types using ANOVA followed
709 by estimated marginal mean and Cohen's D effect size *post hoc* pairwise comparisons between
710 cleaner types (Fig. 5D; Supplementary Figure S5; Supplementary Table S7) using the R packages
711 emmeans (Lenth, 2022) and rstatix (Kassambara, 2020), respectively. Spearman's rank correlation
712 was used to determine similar of gene connectivity between all pairwise cleaner types for genes in
713 each focal module (Fig. 3C and D) independently. Spearman's rank correlations were performed
714 using the rcorr function from R package Hmisc (Harrell, 2020) and p-values were adjusted for
715 multiple hypothesis testing using R stats function p.adjust and the Benjamini-Hochberg method.
716
717 *No evidence of expression specialization in the obligate cleaner species*
718 To test the hypothesis that obligate cleaning is a specialization of facultative cleaning, we
719 compared differential gene expression between facultative cleaners and non-cleaners with
720 differential expression in the obligate cleaner and facultative cleaners. We asked whether the 342
721 DEGs between facultative cleaner and non-cleaner species, the 44 DEGs shared and directionally

concordant in facultative and obligate cleaners versus non-cleaners, and differentially expressed magenta module genes exhibited additional specialization in expression (i.e., were more differentially expressed) in obligate versus facultative cleaners using a Spearman's Rank correlation.

Availability of Data and Materials

All sequence data in this publication will be deposited in the National Center for Biotechnology Information Gene Expression Omnibus. All metadata and scripts used to analyze data and generate figures will be publicly available on the Texas Data Repository at the time of publication and upon request prior to publication by an editor or reviewer.

Acknowledgments

We thank D. Arasappan and the University of Texas at Austin Texas Big Data in Biology Freshman Research Initiative stream (Howard Hughes Medical Institute grant #52008124) and J. Han, P. Bushal, J. Tran, and M. Truong, for help with data analysis. 3'tag sequencing was performed by the Genomic Sequencing and Analysis Facility at UT Austin, Center for Biomedical Research Support RRID# SCR_021713. We also thank Sandra Binning and Dominique Roche for fish catching, and the Lizard Island Research Station for their support. This work was supported by NSF Graduate Research Fellowship and NSF GROW Switzerland & Swiss Government Excellence Scholarship to CAW, by NSF grants IOS-1354942 and IOS-1501704 to HAH and IOS-1601734 to CAW and HAH, by the NSF BEACON Center for Science and Technology (DBI-0939454) to RLY and HAH, and by the Swiss National Science Foundation (grants number: 31003A_153067/1 and 310030B_173334/1 to R.B.).

745 Literature Cited

746 Andersen, P., Morris, R., Amaral, D., Bliss, T., and O'Keefe, J. 2007. The Hippocampus Book
747 (Oxford: Oxford University Press).

748 Baliga, V.B., and Law, C.J. 2016. Cleaners among wrasses: Phylogenetics and evolutionary
749 patterns of cleaning behavior within Labridae. *Mol. Phylogenet. Evol.* 94: 424–435.

750 Baliga, V.B., and Mehta, R.S. 2014. Scaling patterns inform ontogenetic transitions away from
751 cleaning in Thalassoma wrasses. *J. Exp. Biol.* 217: 3597–3606.

752 Baliga, V.B., and Mehta, R.S. 2019. Morphology, Ecology, and Biogeography of Independent
753 Origins of Cleaning Behavior Around the World. *Integr. Comp. Biol.* 59: 625–637.

754 Barabasi, A., Oltvai, Z.N., Barabási, A.-L., and Oltvai, Z.N. 2004. Network biology:
755 understanding the cell's functional organization. *Nat. Rev. Genet.* 5: 101–113.

756 Barbu, L., Guinand, C., Bergmüller, R., Alvarez, N., and Bshary, R. 2011. Cleaning wrasse
757 species vary with respect to dependency on the mutualism and behavioural adaptations in
758 interactions. *Anim. Behav.* 82: 1067–1074.

759 Bedford, T., and Hartl, D.L. 2009. Optimization of gene expression by natural selection. *Proc.*
760 *Natl. Acad. Sci.* 106: 1133 LP – 1138.

761 Blount, Z.D., Lenski, R.E., and Losos, J.B. 2018. Contingency and determinism in evolution:
762 Replaying life's tape. *Science (80-.).* 362: eaam5979.

763 Bolnick, D.I., Barrett, R.D.H., Oke, K.B., Rennison, D.J., and Stuart, Y.E. 2018. (Non)Parallel
764 Evolution. *Annu. Rev. Ecol. Evol. Syst.* 49: 303–330.

765 Brakefield, P.M. 2006. Evo-devo and constraints on selection. *Trends Ecol. Evol.* 21: 362–368.

766 Butler, A.B. 2017. Of Horse-Caterpillars and Homologies: Evolution of the Hippocampus and Its
767 Name. *Brain. Behav. Evol.* 90: 7–14.

768 Cardoso, S.C., Bshary, R., Mazzei, R., Paitio, J.R., Oliveira, R.F., and Soares, M.C. 2015.
769 Arginine vasotocin modulates associative learning in a mutualistic cleaner fish. *Behav. Ecol.*
770 *Sociobiol.* 69: 1173–1181.

771 Carroll, S.B. 2008. Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of
772 Morphological Evolution. *Cell* 134: 25–36.

773 Castoe, T.A., de Koning, A.P.J., Kim, H.-M., Gu, W., Noonan, B.P., Naylor, G., Jiang, Z.J.,
774 Parkinson, C.L., and Pollock, D.D. 2009. Evidence for an ancient adaptive episode of convergent
775 molecular evolution. *Proc. Natl. Acad. Sci.* 106: 8986 LP – 8991.

776 Chan, E.T., Quon, G.T., Chua, G., Babak, T., Trochesset, M., Zirngibl, R.A., Aubin, J., Ratcliffe,
777 M.J.H., Wilde, A., Brudno, M., et al. 2009. Conservation of core gene expression in vertebrate
778 tissues. *J. Biol.* 8: 33.

779 Cheng, Y., Miller, M.J., Zhang, D., Xiong, Y., Hao, Y., Jia, C., Cai, T., Li, S.-H., Johansson,
780 U.S., Liu, Y., et al. 2021. Parallel genomic responses to historical climate change and high
781 elevation in East Asian songbirds. *Proc. Natl. Acad. Sci.* 118: e2023918118.

Côté, I. 2000. Evolution and ecology of cleaning symbioses in the sea. *Oceanogr. Mar. Biol.* 38: 311–355.

Dunn, C.W., Giribet, G., Edgecombe, G.D., and Hejnol, A. 2014. Animal Phylogeny and Its Evolutionary Implications. *Annu. Rev. Ecol. Evol. Syst.* 45: 371–395.

Dunn, C.W., Luo, X., and Wu, Z. 2013. Phylogenetic Analysis of Gene Expression. *Integr. Comp. Biol.* .

Elliott, S.B., Harvey-Girard, E., Giassi, A.C.C., and Maler, L. 2017. Hippocampal-like circuitry in the pallium of an electric fish: Possible substrates for recursive pattern separation and completion. *J. Comp. Neurol.* 525: 8–46.

Fischer, E.K., Hauber, M.E., and Bell, A.M. 2021. Back to the basics? Transcriptomics offers integrative insights into the role of space, time and the environment for gene expression and behaviour. *Biol. Lett.* 17: 20210293.

Foster, S., and Baker, J. 2004. Evolution in parallel: new insights from a classic system. *Trends Ecol Evol* 19: 456–459.

Gallant, J.R., Traeger, L.L., Volkening, J.D., Moffett, H., Chen, P.-H., Novina, C.D., Phillips, G.N., Anand, R., Wells, G.B., Pinch, M., et al. 2014. Genomic basis for the convergent evolution of electric organs. *Sci.* 344: 1522–1525.

Gao, W., Sun, Y.-B., Zhou, W.-W., Xiong, Z.-J., Chen, L., Li, H., Fu, T.-T., Xu, K., Xu, W., Ma, L., et al. 2019. Genomic and transcriptomic investigations of the evolutionary transition from oviparity to viviparity. *Proc. Natl. Acad. Sci.* 116: 3646 LP – 3655.

Gargalovic, P.S., Imura, M., Zhang, B., Gharavi, N.M., Clark, M.J., Pagnon, J., Yang, W.-P., He, A., Truong, A., Patel, S., et al. 2006. Identification of inflammatory gene modules based on variations of human endothelial cell responses to oxidized lipids. *Proc. Natl. Acad. Sci.* 103: 12741 LP – 12746.

Garstang, W. 1928. Memoirs: The Morphology of the Tunicata, and its Bearings on the Phylogeny of the Chordata. *J. Cell Sci.* s2-72: 51–187.

Gilad, Y., Oshlack, A., and Rifkin, S.A. 2006. Natural selection on gene expression. *Trends Genet.* 22: 456–461.

Gingins, S., and Bshary, R. 2016. The cleaner wrasse outperforms other labrids in ecologically relevant contexts, but not in spatial discrimination. *Anim. Behav.* 115: 145–155.

Gould, S.J. 1989. A developmental constraint in Cerion, with comments on the definition and interpretation of constraint in evolution. *Evolution (N. Y).* 43: 516–539.

Gould, S.J. 1991. Wonderful life - the Burgess shale and the nature of history (New York, NY: WW. Norton & Company).

Halfon, M.S. 2017. Perspectives on Gene Regulatory Network Evolution. *Trends Genet.* 33: 436–447.

Hannon, G.J. 2010. FASTX-Toolkit.

Harrison, P.W., Wright, A.E., and Mank, J.E. 2012. The evolution of gene expression and the transcriptome-phenotype relationship. *Semin. Cell Dev. Biol.* 23: 222–229.

Hartwell, L.H., Hopfield, J.J., Leibler, S., and Murray, A.W. 1999. From molecular to modular cell biology. *Nature* 402: C47-52.

Hedges, S.B., Dudley, J., and Kumar, S. 2006. TimeTree: a public knowledge-base of divergence times among organisms. *Bioinforma.* 22: 2971–2972.

Horvath, S., Zhang, B., Carlson, M., Lu, K. V, Zhu, S., Felciano, R.M., Laurance, M.F., Zhao, W., Qi, S., Chen, Z., et al. 2006. Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target. *Proc. Natl. Acad. Sci.* 103: 17402 LP – 17407.

Hu, G., Hovav, R., Grover, C.E., Faigenboim-Doron, A., Kadmon, N., Page, J.T., Udall, J.A., and Wendel, J.F. 2016. Evolutionary Conservation and Divergence of Gene Coexpression Networks in Gossypium (Cotton) Seeds. *Genome Biol. Evol.* 8: 3765–3783.

Huie, J.M., Thacker, C.E., and Tornabene, L. 2020. Co-evolution of cleaning and feeding morphology in western Atlantic and eastern Pacific gobies. *Evolution (N. Y).* 74: 419–433.

Humphries, M.D., and Prescott, T.J. 2010. The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward. *Prog. Neurobiol.* 90: 385–417.

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24: 1403–1405.

Jombart, T., and Ahmed, I. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* 27: 3070–3071.

de Jong, T. V, Moshkin, Y.M., and Guryev, V. 2019. Gene expression variability: the other dimension in transcriptome analysis. *Physiol. Genomics* 51: 145–158.

Kang, J., Ramirez-Calero, S., Paula, J.R., Chen, Y., and Schunter, C. 2023. Gene losses, parallel evolution and heightened expression confer adaptations to dedicated cleaning behaviour. *BMC Biol.* 21: 180.

Kenkel, C.D., and Matz, M. V 2016. Gene expression plasticity as a mechanism of coral adaptation to a variable environment. *Nat. Ecol. Evol.* 1: 14.

Khaitovich, P., Weiss, G., Lachmann, M., Hellmann, I., Enard, W., Muetzel, B., Wirkner, U., Ansorge, W., and Pääbo, S. 2004. A Neutral Model of Transcriptome Evolution. *PLoS Biol.* 2: 682–689.

Kocher, T.D., Conroy, J.A., McKaye, K.R., and Stauffer, J.R. 1993. Similar Morphologies of Cichlid Fish in Lakes Tanganyika and Malawi Are Due to Convergence. *Mol. Phylogenet. Evol.* 2: 158–165.

Kolde, R. 2012. Pheatmap: pretty heatmaps. *R Packag. Version* 1: 726.

Krattinger, O. 2016. A phylogenetic comparison of cleaning behaviour and morphology in wrasses. University of Neuchatel.

Kumar, S., Stecher, G., Suleski, M., and Hedges, S.B. 2017. TimeTree: A Resource for

Timelines, Timetrees, and Divergence Times. *Mol. Biol. Evol.* 34: 1812–1819.

Lamichhaney, S., Card, D.C., Grayson, P., Tonini, J.F.R., Bravo, G.A., Näpflin, K., Termignoni-Garcia, F., Torres, C., Burbrink, F., Clarke, J.A., et al. 2019. Integrating natural history collections and comparative genomics to study the genetic architecture of convergent evolution. *Philos. Trans. R. Soc. B Biol. Sci.* 374: 20180248.

Lamm, M.S., Lamb, A.D., Klapheke, B.P., Tyler, W.A., and Godwin, J.R. 2022. Characterization and distribution of kisspeptins, kisspeptin receptors, GnIH, and GnRH1 in the brain of the protogynous bluehead wrasse (Thalassoma bifasciatum). *J. Chem. Neuroanat.* 121: 102087.

Langerhans, R.B. 2017. Predictability and Parallelism of Multitrait Adaptation. *J. Hered.* 109: 59–70.

Langfelder, P., and Horvath, S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9: 559.

LeDoux, J.E. 2000. Emotion Circuits in the Brain. *Annu. Rev. Neurosci.* 23: 155–184.

Li, H., and Durbin, R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinforma.* 25: 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Subgroup, G.P.D.P. 2009. The Sequence Alignment/Map (SAM) Format and SAMtools. *Bioinformatics* 25: 2078–2079.

Liang, C., Musser, J.M., Cloutier, A., Prum, R.O., and Wagner, G.P. 2018. Pervasive Correlated Evolution in Gene Expression Shapes Cell and Tissue Type Transcriptomes. *Genome Biol. Evol.* 10: 538–552.

Liu, Y., Gu, H.-Y., Zhu, J., Niu, Y.-M., Zhang, C., and Guo, G.-L. 2019. Identification of Hub Genes and Key Pathways Associated With Bipolar Disorder Based on Weighted Gene Co-expression Network Analysis . *Front. Physiol.* 10: 1081.

Lohman, B.K., Weber, J.N., and Bolnick, D.I. 2016. Evaluation of TagSeq, a reliable low-cost alternative for RNAseq. *Mol. Ecol. Resour.* 16: 1315–1321.

Losos, J.B. 2011. CONVERGENCE, ADAPTATION, AND CONSTRAINT. *Evolution (N. Y).* 65: 1827–1840.

Maeda, H., and Mogenson, G.J. 1981. Electrophysiological responses of neurons of the ventral tegmental area to electrical stimulation of amygdala and lateral septum. *Neuroscience* 6: 367–376.

Mahler, D.L., Ingram, T., Revell, L.J., and Losos, J.B. 2013. Exceptional Convergence on the Macroevolutionary Landscape in Island Lizard Radiations. *Science (80-.).* 341: 292 LP – 295.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17:.

Martínez-García, F., Martínez-Marcos, A., and Lanuza, E. 2002. The pallial amygdala of amniote vertebrates: evolution of the concept, evolution of the structure. *Brain Res. Bull.* 57:

463–469.

Mehta, M.R. 2015. From synaptic plasticity to spatial maps and sequence learning. *Hippocampus* 25: 756–762.

Mehta, T.K., Koch, C., Nash, W., Knaack, S.A., Sudhakar, P., Olbei, M., Bastkowski, S., Penso-Dolfin, L., Korcsmaros, T., Haerty, W., et al. 2021. Evolution of regulatory networks associated with traits under selection in cichlids. *Genome Biol.* 22: 25.

Mendonça, R., Soares, M.C., Bshary, R., and Oliveira, R.F. 2013. Arginine vasotocin neuronal phenotype and interspecific cooperative behaviour. *Brain. Behav. Evol.* 82: 166–176.

Meyer, E., Aglyamova, G. V, and Matz, M. V 2011. Profiling gene expression responses of coral larvae (Acropora millepora) to elevated temperature and settlement inducers using a novel RNA-Seq procedure. *Mol. Ecol.* 20: 3599–3616.

Moreno, N., and González, A. 2007. Evolution of the amygdaloid complex in vertebrates, with special reference to the anamnio-amniotic transition. *J. Anat.* 211: 151–163.

Nugent, B.M., Stiver, K.A., Alonzo, S.H., and Hofmann, H.A. 2016. Neuroendocrine profiles associated with discrete behavioral variation in Symphodus ocellatus, a species with male alternative reproductive tactics. *Mol. Ecol.*

Nugent, B.M., Stiver, K.A., Han, J., Kindsvater, H.K., Marsh-Rollo, S.E., Hofmann, H.A., and Alonzo, S.H. 2021. Morpholino-mediated knockdown of the brain mineralocorticoid receptor affects glucocorticoid signaling and neuroplasticity in wild ocellated wrasse (Symphodus ocellatus). *BioRxiv* 2021.11.17.468986.

O'Connell, L.A., and Hofmann, H.A. 2011a. Genes, hormones, and circuits: An integrative approach to study the evolution of social behavior. *Front. Neuroendocrinol.* 32: 320–335.

O'Connell, L.A., and Hofmann, H.A. 2011b. The Vertebrate mesolimbic reward system and social behavior network: A comparative synthesis. *J. Comp. Neurol.* 519: 3599–3639.

O'Connell, L.A., and Hofmann, H.A. 2012. Evolution of a Vertebrate Social Decision-Making Network. *Sci.* 336: 1154–1157.

O'Keefe, J., and Nadel, L. 1978. The hippocampus as a cognitive map (Oxford: Clarendon Press).

Orgogozo, V. 2015. Replaying the tape of life in the twenty-first century. *Interface Focus* 5:.

Pankey, M.S., Minin, V.N., Imholte, G.C., Suchard, M.A., and Oakley, T.H. 2014. Predictable transcriptome evolution in the convergent and complex bioluminescent organs of squid. *Proc. Natl. Acad. Sci.* 111: E4736 LP-E4742.

Pfenning, A.R., Hara, E., Whitney, O., Rivas, M. V, Wang, R., Roulhac, P.L., Howard, J.T., Wirthlin, M., Lovell, P. V, Ganapathy, G., et al. 2014. Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science (80-.).* 346: 1256846.

Portavella, M., Vargas, J.P., Torres, B., and Salas, C. 2002. The effects of telencephalic pallial lesions on spatial, temporal, and emotional learning in goldfish. *Brain Res. Bull.* 57: 397–399.

Poulin, R., and Grutter, A.S. 1996. Cleaning Symbioses: Proximate and Adaptive Explanations. *Bioscience* 46: 512–517.

Randall, J.E., Allen, G.R., and Steene, R.C. 1988. Fishes of the Great Barrier Reef & Coral Sea (Honolulu: University of Hawaii Press).

Raser, J.M., and O'Shea, E.K. 2005. Noise in gene expression: origins, consequences, and control. *Science* 309: 2010–2013.

Renn, S.C.P., Machado, H.E., Duftner, N., Sessa, A.K., Harris, R.M., and Hofmann, H.A. 2018. Gene expression signatures of mating system evolution. *Genome* 61: 287–297.

Revell, L.J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3: 217–223.

Rittschof, C.C., and Hughes, K.A. 2018. Advancing behavioural genomics by considering timescale. *Nat. Commun.* 9: 489.

Robinson, M.D., and Oshlack, A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11: R25.

Roda, F., Walter, G.M., Nipper, R., and Ortiz-Barrientos, D. 2017. Genomic clustering of adaptive loci during parallel evolution of an Australian wildflower. *Mol. Ecol.* 26: 3687–3699.

Roeder, K.D. 1935. An experimental analysis of the secual behavior of the praying mantis (Mantis Religiosa L.). *Biol. Bull.* 69: 203–220.

Rohlfs, R. V, and Nielsen, R. 2015. Phylogenetic ANOVA: The Expression Variance and Evolution Model for Quantitative Trait Evolution. *Syst. Biol.* 64: 695–708.

Rokas, A., and Carroll, S.B. 2008. Frequent and Widespread Parallel Evolution of Protein Sequences. *Mol. Biol. Evol.* 25: 1943–1953.

Romero, I.G., Ruvinsky, I., and Gilad, Y. 2012. Comparative studies of gene expression and the evolution of gene regulation. *Nat Rev Genet* 13: 505–516.

Safran, M., Rosen, N., Twik, M., BarShir, R., Iny Stein, T., Dahary, D., Fishilevich, S., and Lancet, D. 2022. The GeneCards Suite Chapter. In Practical Guide to Life Science Databases, pp. 27–56.

Smyth, G.K. 2005. Limma: linear models for microarray data. *Bioinforma. Comput. Biol. Solut. Using R Bioconductor* 397–420.

Soares, M.C. 2017. The Neurobiology of Mutualistic Behavior: The Cleanerfish Swims into the Spotlight. *Front. Behav. Neurosci.* 11: 191.

Soares, M.C., Bshary, R., Fusani, L., Goymann, W., Hau, M., Hirschenhauser, K., and Oliveira, R.F. 2010. Hormonal mechanisms of cooperative behaviour. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 365: 2737–2750.

Soares, M.C., Bshary, R., Mendonça, R., Grutter, A.S., and Oliveira, R.F. 2012. Arginine Vasotocin Regulation of Interspecific Cooperative Behaviour in a Cleaner Fish. *PLoS One* 7: e39583.

968 Stern, D.L. 2013. The genetic causes of convergent evolution. *Nat Rev Genet.* 14:.

969 Storz, J.F. 2016. Causes of molecular convergence and parallelism in protein evolution. *Nat. Rev.*
970 *Genet.* 17: 239–250.

971 Striedter, G.F. 2016. Evolution of the hippocampus in reptiles and birds. *J. Comp. Neurol.* 524:
972 496–517.

973 Stuart, J., Segal, E., Koller, D., and Kim, S. 2003. A gene-coexpression network for global
974 discovery of conserved genetic modules. *Science (80-.).* 302: 249–255.

975 Suzuki, R., and Shimodaira, H. 2006. Pvclust: an R package for assessing the uncertainty in
976 hierarchical clustering. *Bioinformatics* 22: 1540–1542.

977 Tarazona, S., García-Alcalde, F., Dopazo, J., Ferrer, A., and Conesa, A. 2011. Differential
978 expression in RNA-seq: a matter of depth. *Genome Res.* 21: 2213–2223.

979 Trinh, A.-T., Clarke, S.E., Harvey-Girard, E., and Maler, L. 2019. Cellular and Network
980 Mechanisms May Generate Sparse Coding of Sequential Object Encounters in Hippocampal-
981 Like Circuits. *Eneuro* 6: ENEURO.0108-19.2019.

982 True, J.R., and Haag, E.S. 2001. Developmental system drift and flexibility in evolutionary
983 trajectories. *Evol. Dev.* 3: 109–119.

984 Vaughan, D.B., Grutter, A.S., Costello, M.J., and Hutson, K.S. 2017. Cleaner fishes and shrimp
985 diversity and a re-evaluation of cleaning symbioses. *Fish Fish.* 18: 698–716.

986 Vinepinsky, E., Cohen, L., Perchik, S., Ben-Shahar, O., Donchin, O., and Segev, R. 2020.
987 Representation of edges, head direction, and swimming kinematics in the brain of freely-
988 navigating fish. *Sci. Rep.* 10: 14762.

989 Wake, D., Wake, M.H., and Specht, C. 2011. Homoplasy: From Detecting Pattern to
990 Determining Process and Mechanism of Evolution. *Science (80-.).* 331: 1032–1035.

991 Weitekamp, C.A., and Hofmann, H.A. 2014. Evolutionary themes in the neurobiology of social
992 cognition. *Curr. Opin. Neurobiol.* 28: 22–27.

993 Weitekamp, C.A., and Hofmann, H.A. 2017. Neuromolecular correlates of cooperation and
994 conflict during territory defense in a cichlid fish. *Horm. Behav.* 89: 145–156.

995 Witt, K.E., and Huerta-Sánchez, E. 2019. Convergent evolution in human and domesticate
996 adaptation to high-altitude environments. *Philos. Trans. R. Soc. B Biol. Sci.* 374: 20180235.

997 Yang, Y., and Wang, J.-Z. 2017. From Structure to Behavior in Basolateral Amygdala-
998 Hippocampus Circuits . *Front. Neural Circuits* 11:.

999 Young, R.L., Ferkin, M.H., Ockendon-Powell, N.F., Orr, V.N., Phelps, S.M., Pogány, Á.,
1000 Richards-Zawacki, C.L., Summers, K., Székely, T., Trainor, B.C., et al. 2019. Conserved
1001 transcriptomic profiles underpin monogamy across vertebrates. *Proc. Natl. Acad. Sci.* 116: 1331
1002 LP – 1336.

1003 Young, R.L., and Hofmann, H.A. 2019. Reply to Jiang and Zhang: Parallel transcriptomic
1004 signature of monogamy: What is the null hypothesis anyway? *Proc. Natl. Acad. Sci.* 116: 17629

LP – 17630.

Young, R.L., and Wagner, G.P. 2011. Why ontogenetic homology criteria can be misleading: lessons from digit identity transformations. *J. Exp. Zool. B. Mol. Dev. Evol.* 316B: 165–170.

Yu, G., Lam, T.T.-Y., Zhu, H., and Guan, Y. 2018. Two Methods for Mapping and Visualizing Associated Data on Phylogeny Using Ggtree. *Mol. Biol. Evol.* 35: 3041–3043.

Yu, G., Smith, D.K., Zhu, H., Guan, Y., and Lam, T.T.-Y. 2017. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* 8: 28–36.

Zayed, A., and Robinson, G.E. 2012. Understanding the Relationship Between Brain Gene Expression and Social Behavior: Lessons from the Honey Bee. *Annu. Rev. Genet.* 46: 591–615.

Zhen, Y., Aardema, M.L., Medina, E.M., Schumer, S., and Andolfatto, P. 2012. Parallel Molecular Evolution in an Herbivore Community. *Science (80-.).* 337: 1634–1637.

Supplementary Tables and Figures

All supplementary tables are provided in Youngetal_SupplementaryTables.xlsx

Supplementary Table S1. Number of genes and read mapping percentages by species and evolutionary distance and number of comparable genes across species.

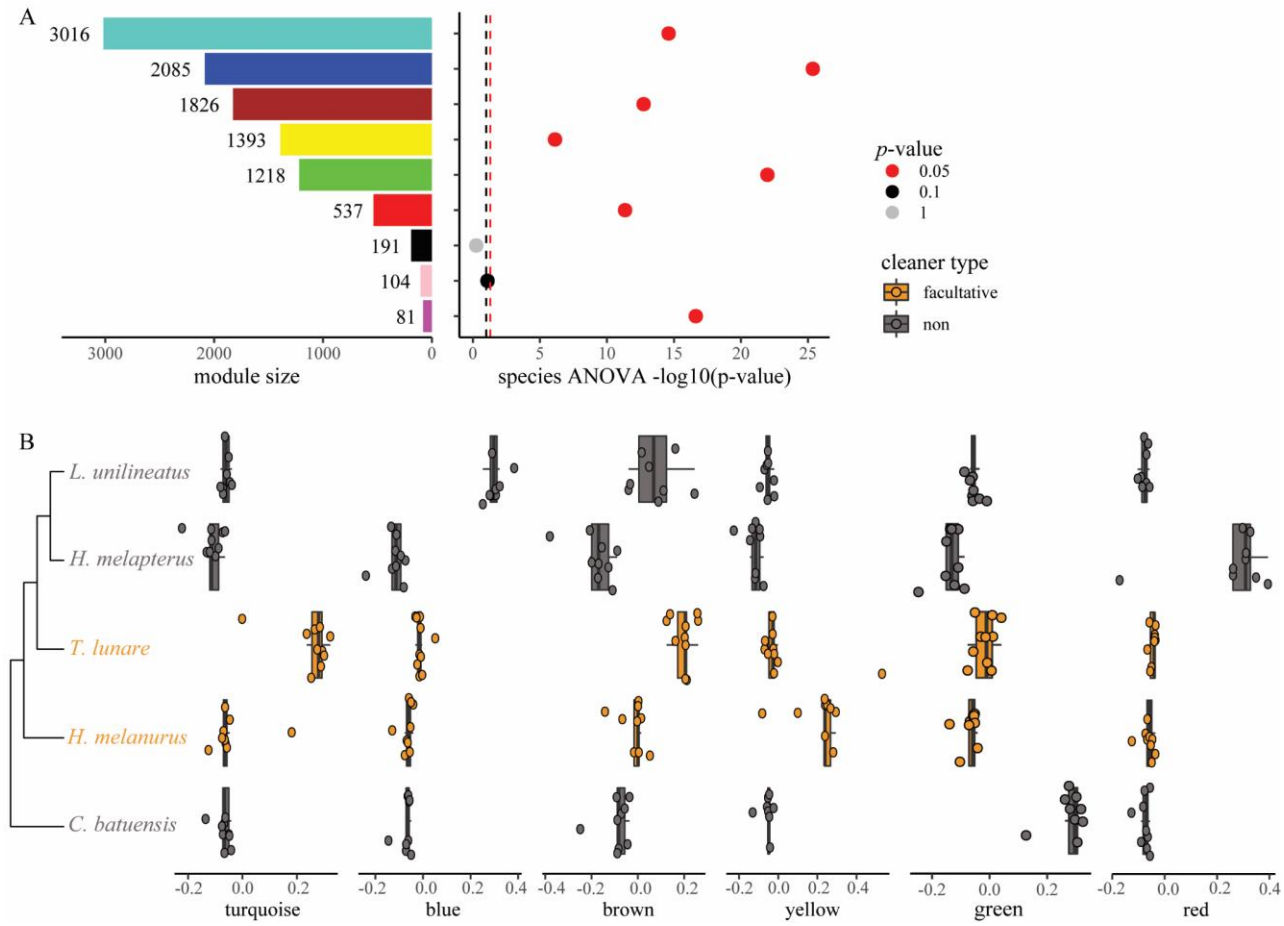
Supplementary Table S2. Variation in module eigengene expression between facultative cleaners and non-cleaners. Welch's t-statistics, phylogenetic ANOVA F-statistics, and *p*-values for all modules a provided. Modules are identified by color and size indicates the number of genes in each module. Significance at $p < 0.05$ after accounting for phylogenetic independence is indicated in bold.

Supplementary Table S3. Variation in module eigengene expression across species. ANOVA F-statistics and *p*-values are provided for all modules and Tukey's honest signal difference post hoc pairwise species q-statistics and *p*-values are provided for significant pairwise comparisons from modules with an ANOVA *p*-value < 0.05. Modules are identified by color and size indicates the number of genes in each module. Significance at $p < 0.05$ is indicated in bold.

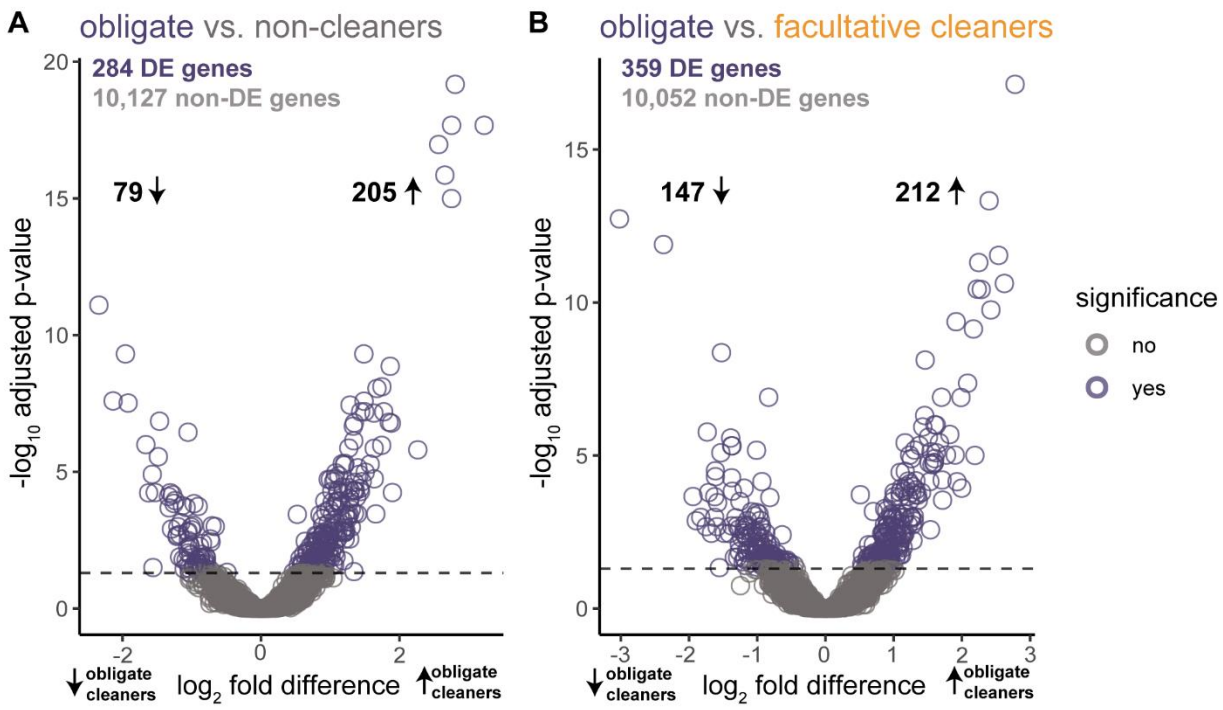
Supplementary Table S4. Candidate genes associated with independent transitions to facultative cleaning. ENSEMBL IDs, Gene IDs, functional annotations from GeneCards (Safran et al., 2022) and PubMed, $\log_2$ fold difference and associated adjusted p-value between facultative cleaners and non-cleaner, magenta module intramodular connectivity, and EVE model negative $\log_{10}$ beta values are provided for each gene.

Supplementary Table S5. Behavioral, ecological, life history, and morphological attributes of the six focal species included in this analysis. Lengths were measures from individuals sampled as part of this study. All other attributes were obtained from previous studies (Randall et al., 1988; Gingins and Bshary, 2016; Krattinger, 2016).

Supplementary Table S7. Results of an analysis of variance (ANOVA) comparing focal module gene connectivity across cleaner types, estimated marginal mean and Cohen's D effect size *post hoc* pairwise comparisons of connectivity (kTotal) between cleaner types, and correlation of gene connectivity between cleaner types. Genes from all nine focal modules differed significantly across cleaner types (indicated with an asterisks); however, modules differed in the directionality and effect size of overall connectivity differences (*t* ratio and Cohen's D; Supplementary Fig. S4) as well as correlation of gene connectivity scores (Spearman's rho).



Supplementary Figure S1. Seven gene co-expression modules differed across species (A). The number of gene contain in each module is indicated by module size (A). Dashed lines indicated statistical support for the ANOVA at $p < 0.1$ (black) and $p < 0.05$ (red) (A). Species-level co-expression comparisons illustrate differences in module eigengene expression across species. Species are colored by cleaner type (non-cleaner and facultative species, grey and gold, respectively). The magenta module is plotted in Figure 3B (B). Significance between pairwise comparisons is provided in Supplementary Table S5.

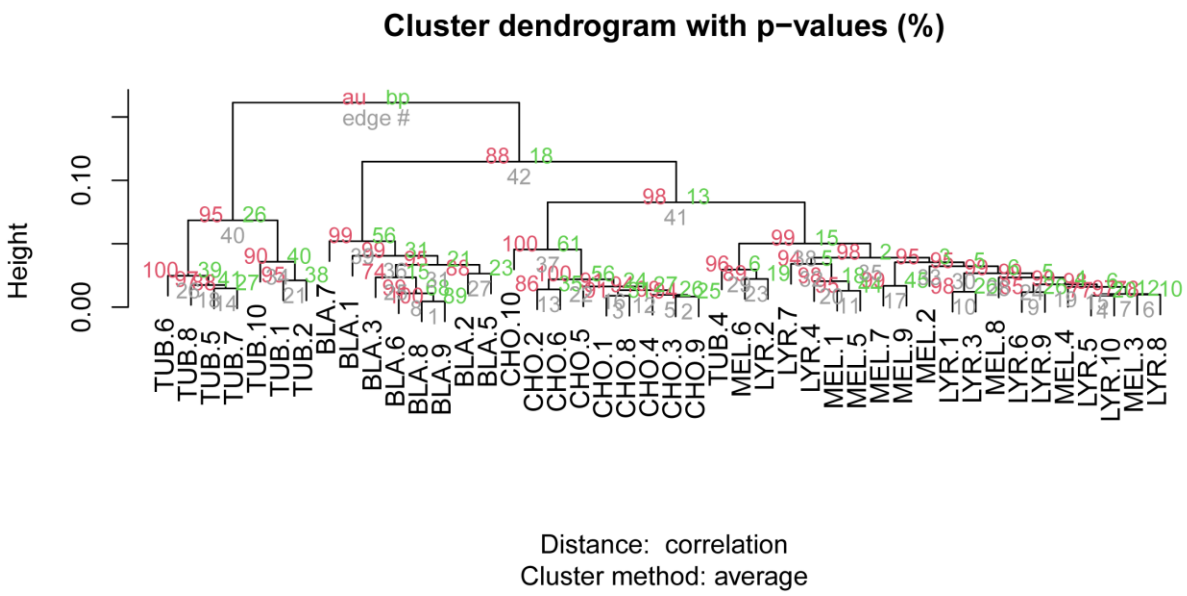


1056

1057 **Supplementary Figure S2**. Differential gene expression between obligate and non-cleaner

1058 species (A) and obligate and facultative cleaner species (B).

1059

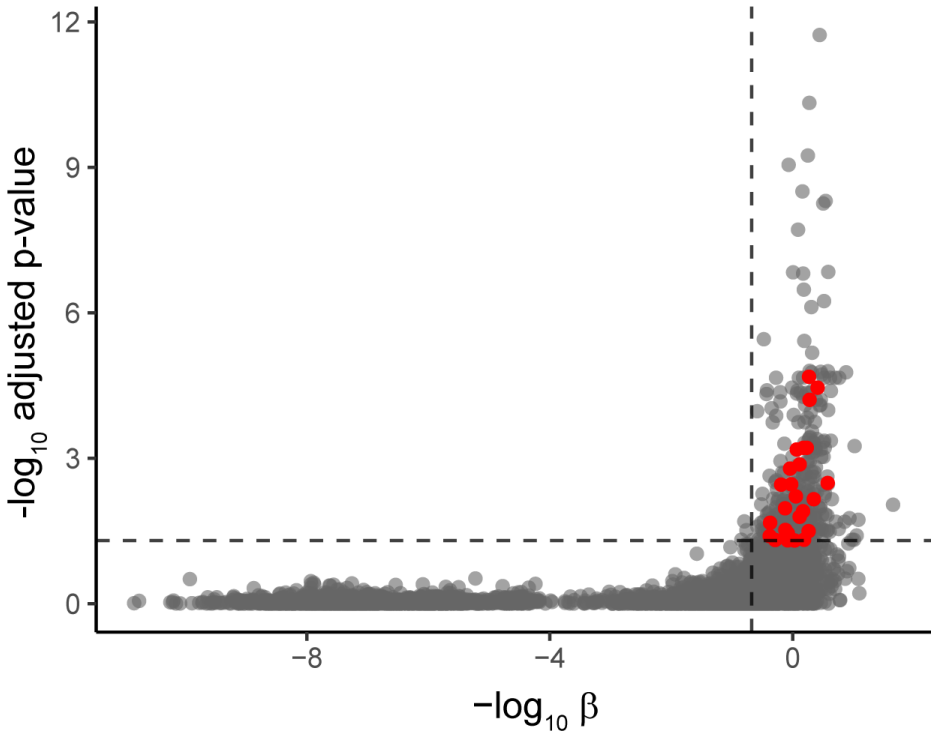


Supplementary Figure S3. Approximately unbiased p-values (red) and bootstrap probabilities (green) for sample clustering by facultative cleaning candidate gene expression Fig. 4B. Height indicates 1-correlation. While facultative cleaners *Halichoeres melanurus* (MEL) and *Thalassoma lunare* (LYR) cluster together, non-cleaners *Hemigymnus melapterus* (BLA), *Coris batuensis* (CHO), and *Labrichthys unilineatus* (TUB) cluster by species affiliation.

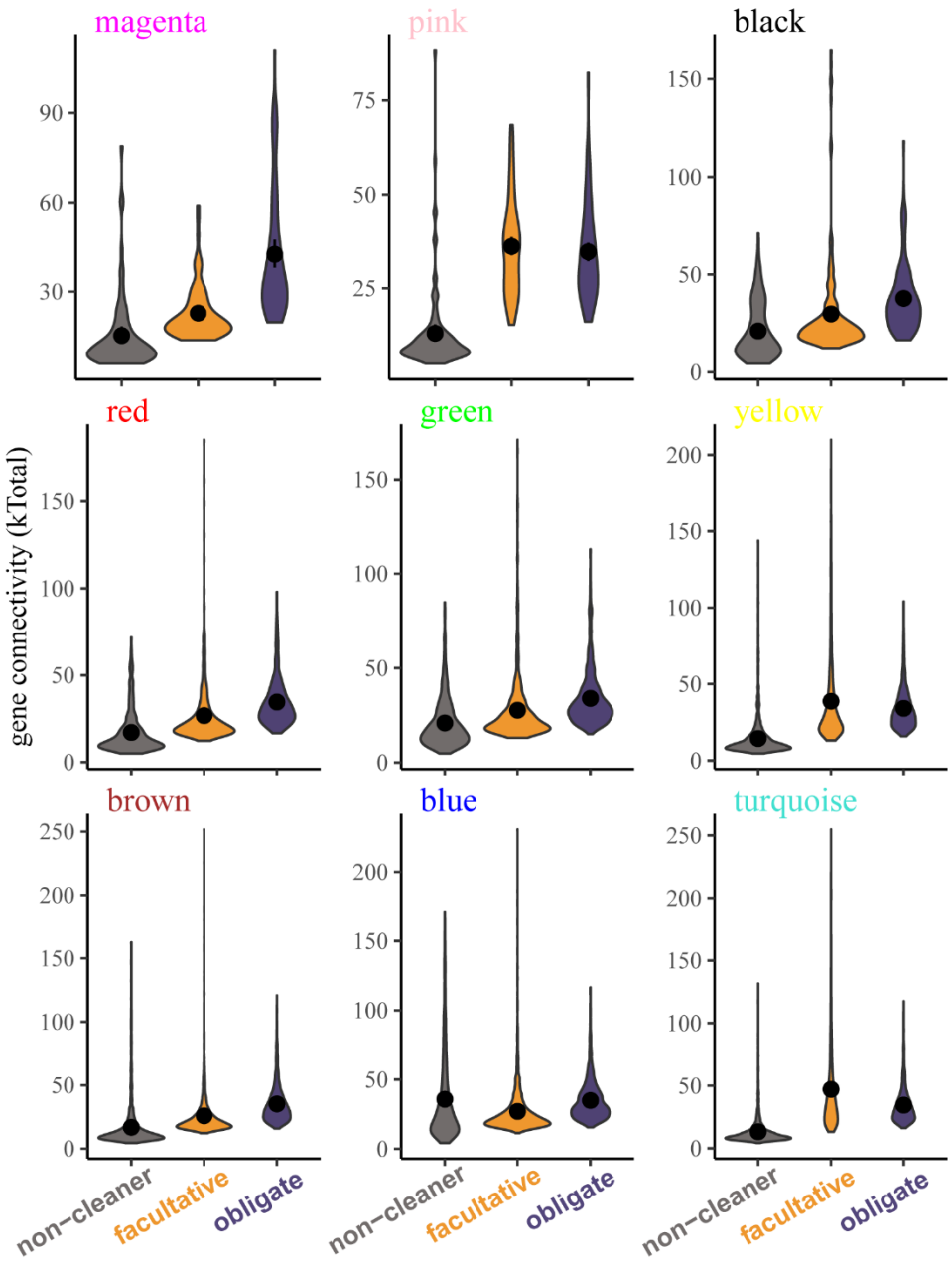
1060

1061

1062

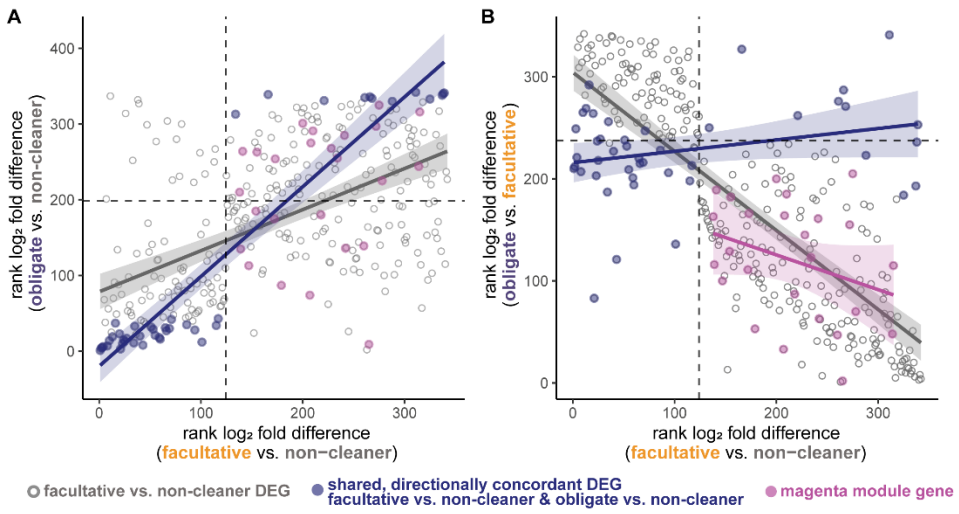


Supplementary Figure S4. Relationship between evolutionary divergence in expression across all species (EVE; -$\log_{10}$B) and differential expression between facultative cleaners and non-cleaners (*limma*; -$\log_{10}$ adjusted p-value). Genes with high evolutionary divergence in expression across species were more likely to be differentially expressed between facultative and non-cleaner species (Spearman's $\rho = 0.48$, $p < 2.2e\text{-}16$). The novel candidate genes identified in the magenta module (Fig. 4) are shown in red.



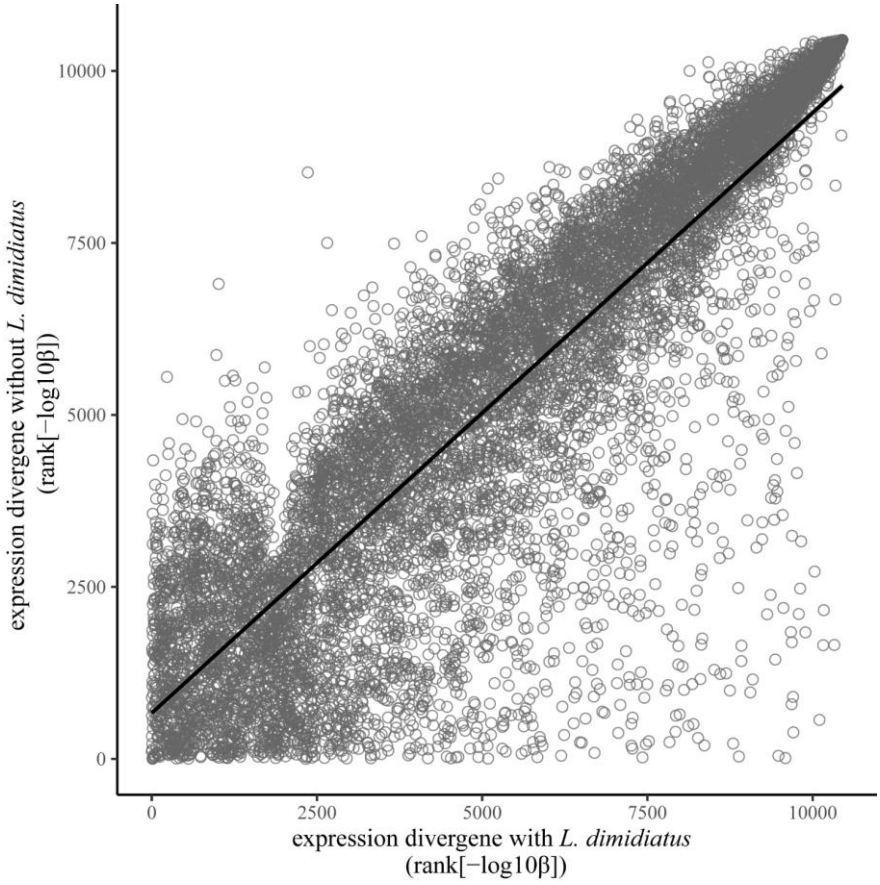
Supplementary Figure S5. Differences in gene connectivity across cleaner types. Gene connectivity (kTotal) was obtained from WGCNA on each cleaner type independently. Connectivity was compared for genes in each of our nine focal modules (Figure 3). We found that gene from two focal modules (magenta and brown) exhibited highest connectivity in the obligate cleaner and lowest connectivity in the non-cleaner species (Figure 5C; Supplementary Table S7).

Further, gene connectivity was correlated in all pairwise comparisons of cleaner types (Figure 5D; Supplementary Table S7). Means and standard deviations are shown as dots and whiskers within each violin plot. Modules are arranged in order of size (i.e., number of genes) with magenta being the smallest module and turquoise the largest (Figure 3).

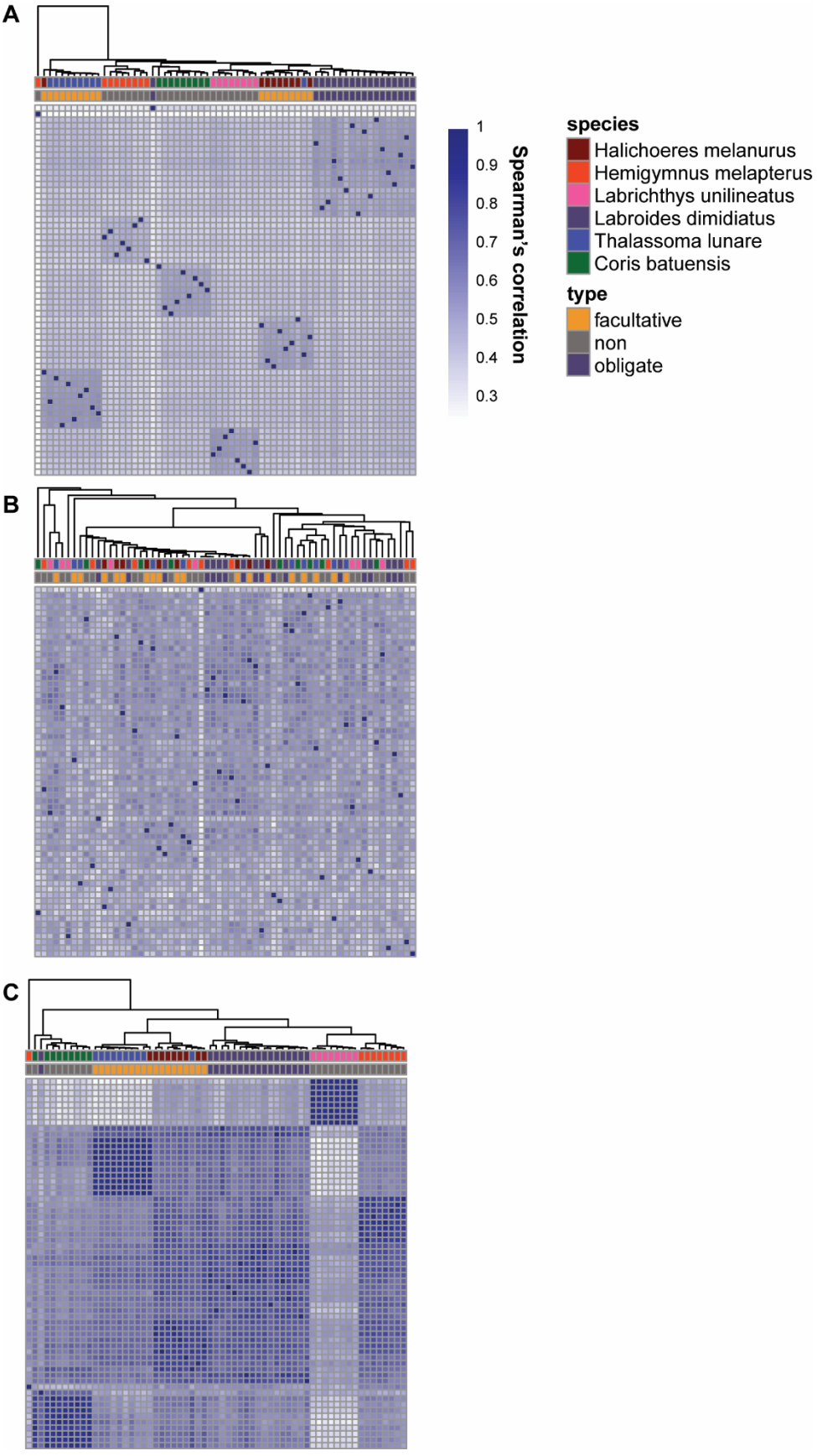


Supplementary Figure S6. Patterns of expression differences and variation between obligate (*L. dimidiatus*) and facultative cleaners (*H. melanurus* and *T. lunare*) with the non-cleaner species indicate significant correlations in expression of facultative cleaning related, differentially expressed genes (grey: Spearman's $\rho = 0.50$, $p < 2.2e\text{-}16$; blue: Spearman's $\rho = 0.93$, $p < 2.2e\text{-}16$), but not magenta module genes (magenta: Spearman's $\rho = 0.24$, $p = 0.12$) (A). Expression differences (B) between obligate (*L. dimidiatus*) and facultative cleaners (*H. melanurus* and *T. lunare*) does not support the hypotheses that obligate cleaning is a specialization of facultative cleaning (B). We found a negative correlation between gene expression differences in facultative and non-cleaner species versus obligate and facultative species for all differentially expressed genes (B, grey: Spearman's $\rho = -0.77$, $p < 2.2e\text{-}16$), genes differentially expressed in both facultative cleaners and the obligate cleaner as compared to the non-cleaner species (B, blue: Spearman's $\rho = 0.11$, $p = 0.48$), and genes differentially expressed and contained in the facultative cleaner-associated magenta module (B, magenta: Spearman's $\rho = -0.3$, $p = 0.14$). 342 DEGs between facultative and non-cleaner species are shown in grey, 44 DEGs shared and directionally

concordant in obligate vs. non-cleaners are blue, 81 magenta module genes are magenta, and shading indicates standard error.



Supplementary Figure S7. EVE analysis on all six species and on the five non-cleaner and facultative cleaner species were highly concordant. Spearman's rank correlation of gene-specific β_i values calculated with and without the obligate cleaner are significantly positively correlated (Spearmans $\rho = 0.87$, $p < 2.2e\text{-}16$).


A
B
C
Spearman's correlation
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
species
Halichoeres melanurus
Hemigymnus melapterus
Labrichthys unilineatus
Labroides dimidiatus
Thalassoma lunare
Coris batuensis
type
facultative
non
obligate

1094

Supplementary Figure S8. Sample clustering with all genes (A), the 105 genes with the lowest interspecific divergence (highest 1% beta values) (B), and the 105 genes with the highest interspecific divergence (lowest 1% beta values) (C). Clustering individual samples using genes identified as having low evolutionary divergence relative to other gene in the transcriptome results in loss of clustering by species affiliation (B). Clustering of samples by genes identified as having high evolutionary divergence relative to other genes in the transcriptome group by species and additionally group the facultative cleaning species (C). Correlation matrix of all samples is generated using Spearman's rank correlation. Correlation strength is indicated by intensity of color. Colored bars indicate species and cleaner type affiliations.