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

Rebecca L. Young[1*], Chelsea A. Weitekamp[1#], Zegni Triki[3$], Yiheng Su[1], Redouan Bshary[3], 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 Cellular and Molecular Biology, 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
Department of Zoology
Stockholm University
Universitetsvägen 10 A
SE-106 91 Stockholm
Sweden

Keywords: cooperative behavior, RNA-seq, 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 (1–5). While there is now ample empirical evidence suggesting that convergence is much more common in nature than Gould (6) predicted (3, 7), it is still unclear to which extent the prevalence of convergence is due to shared evolutionary history (2, 8). First, studies of convergence often investigate replicate populations of the same or closely related species (9–13). 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) (1, 7, 14). Further, because many distinct genotypes, developmental pathways, and cellular origins can give rise to functionally equivalent and even homologous phenotypes (15, 16), 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) (8, 17–19). 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 (20–22), even at the level of the transcriptome (23–28). 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 (29–31).

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 (32, 33). 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 (32, reviewed in: , 33, 34). 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 (35–37). 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 (32, 34, 38). 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; , 38). 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 (32, 38, 39). 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 (39, 40). Further, behavioral specialization

correlates with morphological specialization such that the highly specialized obligate cleaner species exhibit more limited morphological variation and increased body elongation (36). 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 (38), 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 41). 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: , 42, 43) remain poorly understood (44, 45). 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* (46), possibly via the V1a receptor subtype (47) and through modulation of their learning competence (48). 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

130 species can test hypotheses of convergent evolution and identify candidate brain regions and novel
131 candidate genes associated with specific behavior phenotypes.

132 Even though the labrid brain has received little attention from neuroanatomists (but see 49,
133 50), there is an evolutionarily conserved Social Decision-Making Network (SDMN) that is critical
134 for evaluating stimulus salience and regulating sexual, aggressive, and parental behavior across
135 vertebrates (51, 52). Two SDMN nodes that have been well studied for their role in spatial and
136 social cognition, respectively, deserve particular attention in the context of a phenotype as complex
137 as cleaning behavior. First, the medial pallium (which forms the hippocampus in mammals) plays
138 a critical role in spatial memory in mammals (53–55) and has a functionally equivalent role in both
139 avian and non-avian reptiles (56, 57) and teleost fish (58–60). And second, the lateral pallium
140 (which mainly comprises the basolateral amygdala in mammals) integrates multimodal sensory
141 inputs and regulates affective and goal-directed behavior in mammals (61–63), similar to the
142 situation in birds and reptiles (64) as well as teleosts (65). Together, the medial (area Dl in teleosts:
143 lateral part of the dorsal telencephalon) and the lateral pallium (area Dm in teleosts: medial part of
144 the dorsal telencephalon) are ideally suited for comparative analyses as they are complementary
145 in function and reciprocally connected.

146 It is clear that differences in the neural transcriptome underlie behavioral variation (66). In
147 fact, coordinated expression of neural genes associated with convergently evolved behavioral
148 phenotypes can be conserved across even distantly related species (Pfenning et al., 2014; Rittschof
149 et al., 2014; Morandin et al., 2016; Renn et al., 2016, 2018; Young et al., 2019). However, linking
150 transcriptomic and phenotypic variation across species requires an understanding of how gene
151 expression evolves. First, transcriptomes are inherently noisy due to the stochastic nature of the
152 biochemical reactions of transcription (67), and can be highly plastic in response to environmental

and physiological fluctuations (reviewed in: , 68), which can mask relevant evolutionary patterns depending on when and how samples are obtained and analyzed (69–71). Second, phenotypes can be altered by the changes in expression of individual genes or entire gene co-expression networks (72–77). 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 (78, reviewed in: , 79). Finally, because like other complex traits, the transcriptome is shaped by stabilizing selection (80–82) and evolves via neutral or nearly neutral processes (e.g., drift; , 83), 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 (84, 85).

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 the brain regions associated with spatial and social cognition, we sequenced RNA extracted from the putative teleost homologs of the mammalian hippocampus and basolateral amygdala (areas Dl and Dm, respectively; , 51)

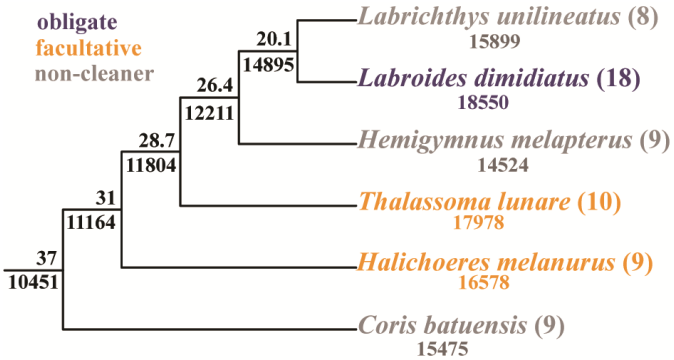


Figure 1. The six species of wrasses (Labridae) selected for this study represent differences in cleaning activity (obligate cleaners, facultative cleaners, and non-cleaners). The number of individuals sequenced for each species is shown in parentheses. The number of genes expressed in each species and shared across species at each node is shown. Estimated divergence times (millions of years ago) at each node were obtained from TimeTree (106, 107).

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.

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 (Fig. 1; 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 10,451 genes was used for all downstream analyses.

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. We first 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 973 (9.3%) differentially expressed genes (DEGs) after correcting for false discovery rate (adjusted *p*-value < 0.05), with 670 genes showing increased expression and 303 genes showing decreased expression in facultative cleaners as compared to non-cleaners (Fig. 2A). To quantify the probability that this difference in expression between facultative cleaners and non-cleaners is greater than expected by chance, we performed differential expression analysis for all 2 vs. 3 combinations of facultative and non-cleaner species. The number of DEGs identified in these nine additional comparisons ranged from 455 to 981 genes (median = 641), suggesting that identifying 973 DEGs is somewhat unlikely (i.e., more DEGs than 9/10 comparisons; empirical cumulative distribution function, ecdf, $p = 0.07$; Fig. 2A; Supplementary Table S2). Of note, we found 981 DEGs when the two most closely related non-cleaner species (*L. unilineatus* and *H. melapterus*, both non-cleaners; Fig. 1) are compared to the remaining three species (*C. batuensis*, *H. melanurus*, and *T. lunare*)

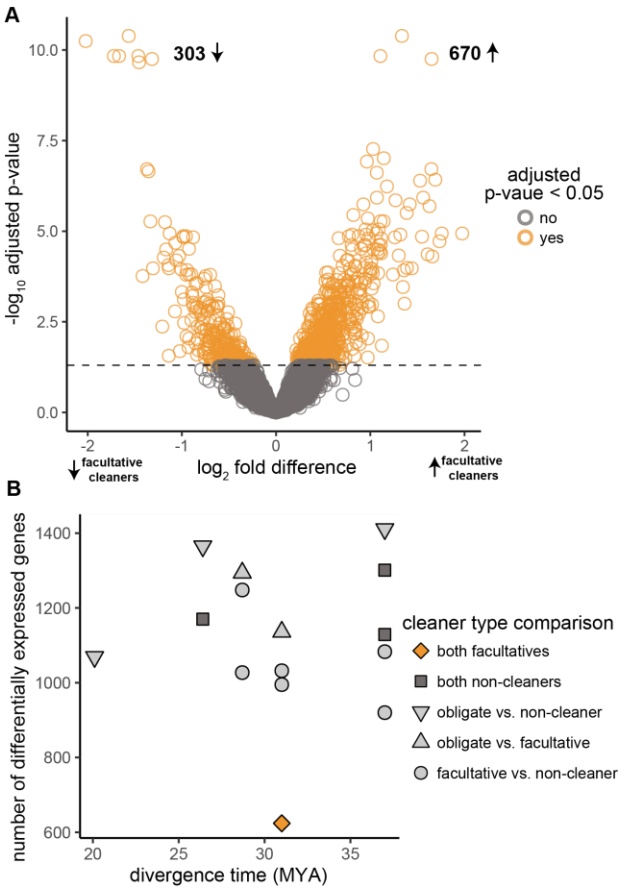


Figure 2. 973 genes are differentially expressed between facultative cleaner and non-cleaner species (A). Pairwise species comparisons between all species pairs found fewer DEGs between the facultative cleaner species as compared to all other species pairs (B: gold diamond).

(Supplementary Table S2). 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 2B).

To identify high-level functions (i.e., KEGG pathways; , 86, 87) enriched in the 973 gene differentially expressed between facultative and non-cleaner species, we performed pathway analysis using pathfindR (88). 47 KEGG pathways were enriched in our differentially expressed genes including several broadly associated with synaptic function and plasticity, neuronal growth, and neurite elongation such as ribosomal biogenesis (89), ubiquitin-mediated proteolysis (90), extracellular matrix receptors (91), O-glycan biosynthesis (92), and mTOR signaling (93) (Supplementary Table S3).

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. 3A). Welch's t-test and ANOVA revealed a number of modules whose co-expression eigengene differs across facultative and non-cleaner types (Supplementary Table S4). We found seven modules that differ significantly across species (Supplementary Figure S1; Supplementary Table S5). 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. 3A and B; Supplementary Table S4). 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).

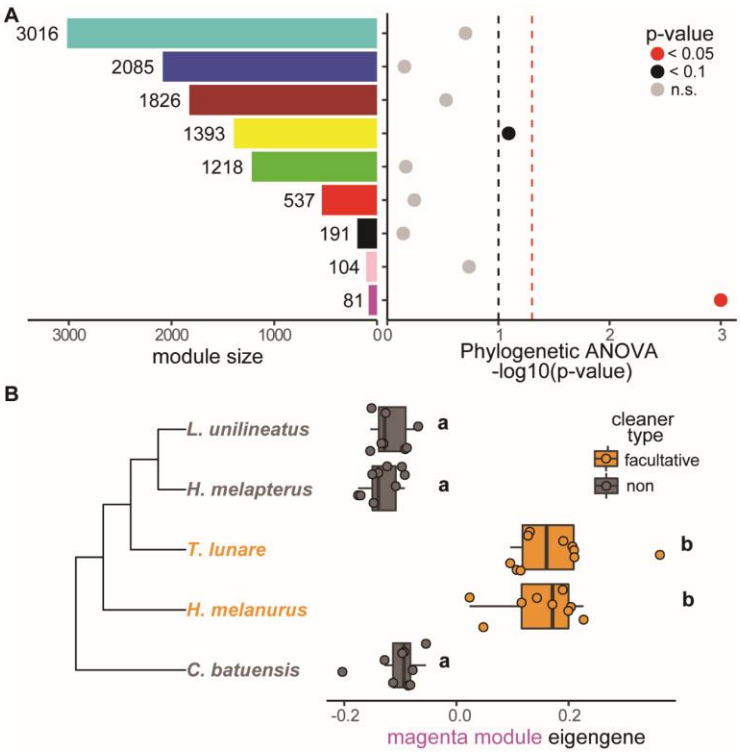


Figure 3. Combining gene co-expression analysis and phylogenetic comparative analyses finds one co-expression module (magenta) significantly associated with facultative cleaning (A). The number of genes contained in each module is indicated by module size (A). Dashed lines indicated statistical support for the phylogenetic ANOVA at p < 0.1 (black) and p < 0.05 (red) (A). 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) (B). Significance between pairwise comparisons is indicated by the associated letters (B).

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): , 94), 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 (95, 96, e.g. see: , 97). Second, we used EVE analysis (94) to calculate evolutionary 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 ($F_{(1,79)} = 216.6$, $p = 2.2e\text{-}16$) and exhibited high evolutionary divergence scores ($F_{(1,79)} = 67.08$, $p = 3.7e\text{-}12$; Fig. 4A). Specifically, we identified 41 differentially expressed magenta module genes with high interspecific variability as previously undescibed candidate genes associated with facultative cleaning (Fig. 4B; Supplementary Table S6). We were able to annotate 39 of these 41 genes with a molecular or cellular function. Of note, 25 (61%) of these genes have been broadly implicated in neural development and function (highlighted in green font Fig 4B; Supplementary Table S6).

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-

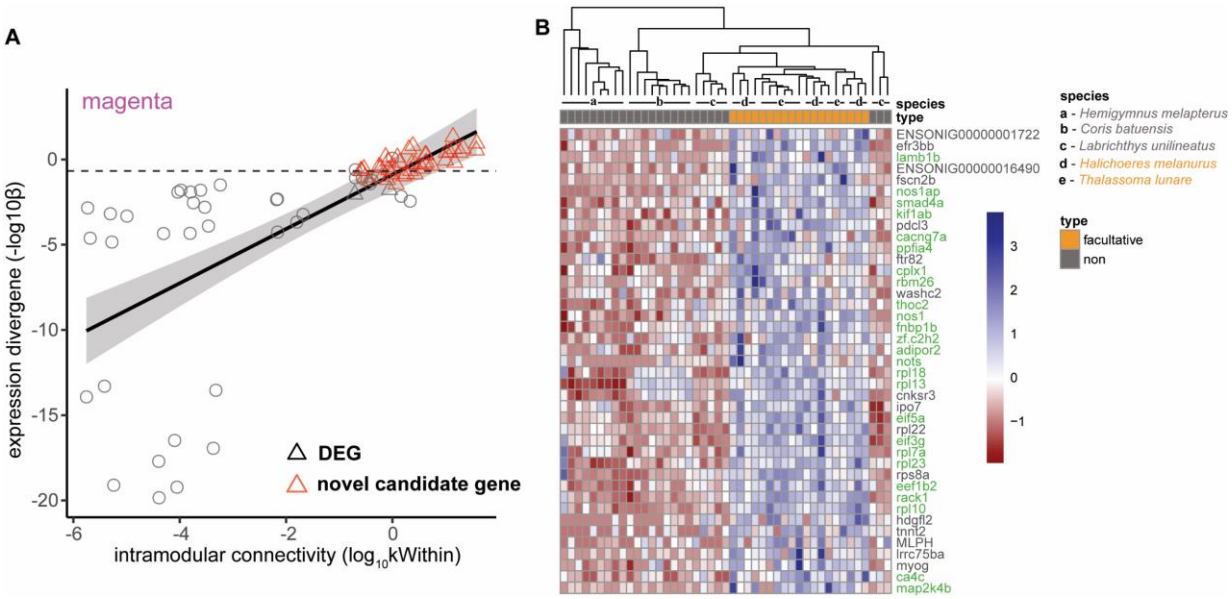


Figure 4. Gene connectivity in the facultative cleaner associated magenta module is positively correlated with expression divergence across species ($F_{(1,79)}$ = 67.08, p = 3.7e-12) and fold difference in expression between facultative and non-cleaners ($F_{(1,79)}$ = 216.6, p = 2.2e-16) (A). 41 magenta module genes were significantly differentially expressed between facultative and non-cleaner species and had higher than average expression divergence (A: red triangles). Overall, expression of these 41 candidate genes of facultative cleaner 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.

cleaner species (*C. batuensis*, *H. melapterus*, and *L. unilineatus*) and found 1091 (10.4%) DEGs with 696 genes showing increased expression and 395 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: (1) genes that were differentially expressed and directionally concordant in both differential expression analyses (973 DEGs in facultative versus non-cleaners and 1091 DEGs in obligate versus non-cleaners); (2) genes differentially expressed between facultative and non-cleaners (973 DEGs) and directionally concordant (but not significant)

between obligate and non-cleaners (9360 non-DEGs); (3) genes differentially expressed between facultative and non-cleaners (973 DEGs) and directionally discordant (but not significant) between obligate and non-cleaners (9360 non-DEGs, as above in 3); and (4) differentially expressed in both analyses (973 DEGs in facultative versus non-cleaners and 1091 DEGs in obligate versus non-cleaners, as above in 1), but directionally discordant. Comparing differential expression analyses, we found that 197 of the 973 DEGs (20.2%) between facultative cleaner and non-cleaner species were also significant DEGs in the concordant direction between the obligate and non-cleaners, which is significantly more than expected by chance (Fig. 5A and B: 197/1091 DEGs in the obligate versus non-cleaner comparison; ecdf p = 3.1e-22). The expression of an additional 494 of these 973 DEGs (50.8%), while not significant, were concordant in the obligate versus non-cleaner species (Fig. 5A and B: 494/9360 non-DEGs in the obligate versus non-cleaner comparison; ecdf p = 1). Finally, only 18 of these 973 DEGs (1.8%) were significantly differentially expressed, but directionally discordant in the obligate versus non-cleaner comparison (Fig. 5A and B: 18/1091 DEGs in the obligate versus non-cleaner comparison; ecdf p = 1) and 264 of these 973 DEGs (27.1%) were discordant between obligate versus non-cleaner species (Fig. 5A and B: 264/9360 non-DEGs in the obligate versus non-cleaner comparison; ecdf p = 1). In other words, cleaner-level differential expression comparisons supported concordance of gene expression patterns in facultative and obligate cleaners (Fig5 A 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. 2B, upward triangles).

To further characterize gene co-expression changes associated with facultative cleaning or cleaning in general (i.e., obligate and facultative cleaners) we characterized maintenance of

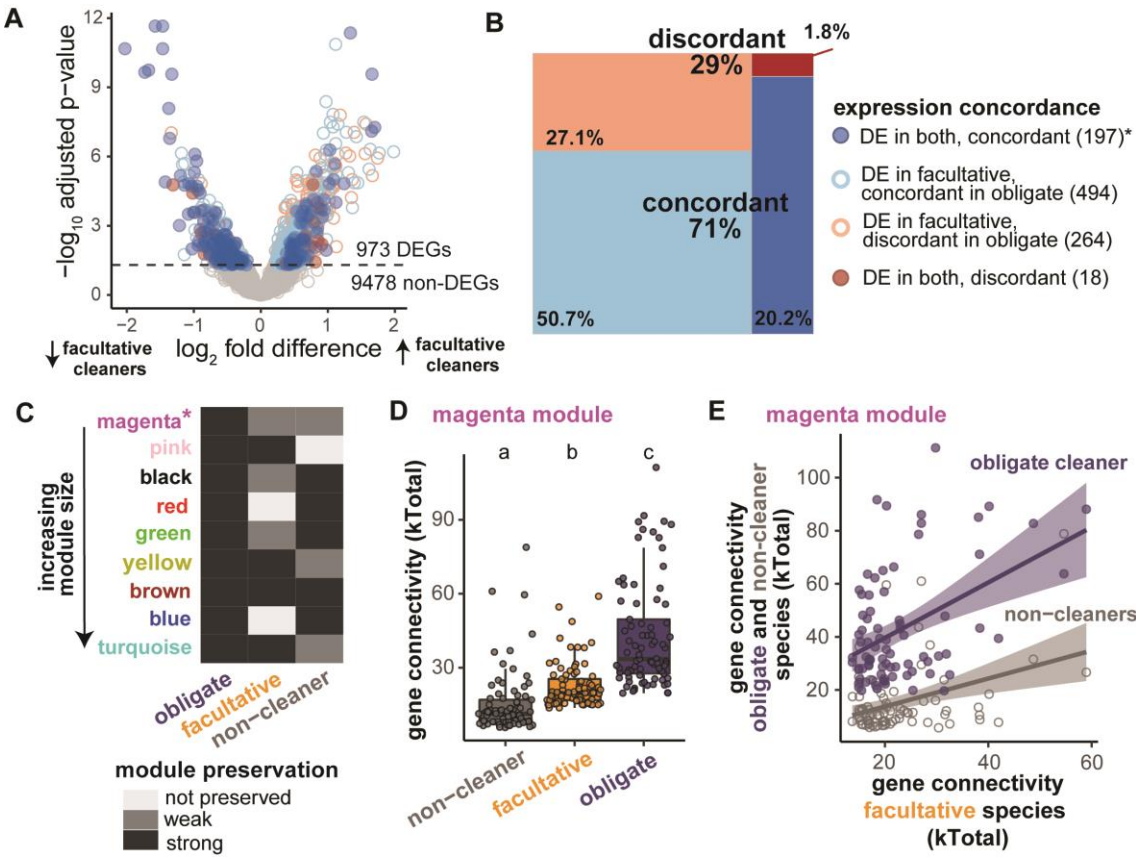


Figure 5. Concordance of expression (A and B) and co-expression patterns (C-E) in facultative and obligate cleaner species as compared to non-cleaner species. 71% of the 973 facultative versus non-cleaner DEGs, were expressed in the concordant direction in obligate versus non-cleaners (A and B, blues). We found more genes than expected by chance (20.2%) 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; 197 genes, p = 3.1e-22). We found overall strong module preservation across cleaner types of our focal WGCNA modules (C; Fig. 3). Gene connectivity of the facultative cleaning associated magenta module genes was highest in the obligate cleaner and lowest in non-cleaners (D; $F_{(2,240)} = 68.1$; p = 3.8E-24) and differed across all pairwise combinations of cleaner types (Supplementary Table S8). Further, gene connectivity in facultative cleaners was significantly correlated with connectivity in both the obligate cleaner and non-cleaners (E; 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 S4 and Supplementary Table S8. Dashed line indicates adjusted p-value < 0.05 (A). Strong preservation (black) indicates a $Z_{summary} > 10$, weak preservation (dark gray) indicates $2 < Z_{summary} > 10$, and not preserved (light gray) indicates $Z_{summary} < 2$ (C). Significance between pairwise comparisons is indicated by the associated letters (D). Shading on the regression indicates the standard error (E).

gene co-expression patterns defined in the facultative cleaner and non-cleaner WGNCA (focal co-expression modules; Fig. 3) in WGCNA analyses including the obligate cleaner only (*L.*

323 *dimidiatus*), the facultative cleaners only (*H. melanurus* and *T. lunare*), and the non-cleaners only
324 (*C. batuensis*, *H. melapterus*, and *L. unilineatus*). In general, the nine modules defined in the
325 facultative and non-cleaner WGCNA (focal co-expression modules; Fig. 3) were preserved in the
326 additional analyses (Fig. 5C). Of the 27 module comparisons across WGCNA analyses, we found
327 18 module comparisons with strong preservation ($Z_{summary} > 10$), 6 module comparisons with weak
328 preservation ($2 < Z_{summary} > 10$), and 3 module comparisons showing no evidence of preservation
329 ($Z_{summary} < 2$) (after 98). Finally, we assessed differences in gene connectivity of genes from our
330 nine focal co-expression modules in the obligate cleaner, the facultative cleaners, and the non-
331 cleaner. Because the magenta module was uniquely associated with facultative cleaners (Fig. 3),
332 we focused on the connectivity of magenta module genes with all other genes in the transcriptome
333 (kTotal) in the three cleaner types (all other modules are shown in Supplementary Figure S4 and
334 Supplementary Table S8). We found that connectivity of magenta module genes was highest in
335 the obligate cleaner and lowest in the non-cleaner species (Fig. 5D; $F_{(2,240)} = 68.1$; $p = 3.8E\text{-}24$)
336 and differed significantly between all pairwise cleaner types (non-cleaners versus facultative
337 cleaners: *t* ratio = -2.34, adjusted *p*-value = 1.92E-2; non-cleaners versus the obligate cleaner: *t*
338 ratio = -8.48, adjusted *p*-value = 7.31E-17; facultative cleaners versus the obligate cleaner: *t* ratio
339 = -6.13, adjusted *p*-value = 1.31E-9). Further, we found that gene connectivity in facultative
340 cleaners was significantly correlated with connectivity in both obligate and non-cleaner species
341 (Fig. 5E; non-cleaners versus facultative cleaners: Spearman's $\rho = 0.37$; adjusted $p = 0.001$;
342 facultative cleaners versus the obligate cleaner: Spearman's $\rho = 0.44$; adjusted $p = 7.8E\text{-}5$). One
343 other module, the brown module (Fig. 3), exhibited a similar pattern of connectivity between
344 cleaner types; however, effect sizes of connectivity differences and gene connectivity correlations
345 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 S8);

The obligate cleaner species L. dimidiatus shows no evidence of expression specialization or canalization 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. To test the hypothesis that life history transitions from a juvenile cleaner state to non-cleaner adult state in facultative cleaners is associated with increased variation in the neuromolecular bases of cleaning, we compared variation in gene expression between facultative and non-cleaner species (Fig. 2) in each facultative cleaning species (*H. melanurus* and *T. lunare*) compared to the obligate cleaner (*L. dimidiatus*). Neither expression differences (Supplementary Figure S5A and B) nor variation (Supplementary Figure S5C and D) between obligate (*L. dimidiatus*) and facultative cleaners (*H. melanurus* and *T. lunare*) support the hypotheses that obligate cleaning is a specialization or a canalized life history state of facultative cleaning. Specifically, DEGs between facultative and non-cleaners are not specialized (i.e., more differentially expressed) in the obligate cleaner (Supplementary Figure S5B). In addition, genes from the facultative cleaner associated magenta module are more (Supplementary Figure S5C) or equally variable (Supplementary Figure S5D) in the obligate cleaner (*L. dimidiatus*) as compared to two the facultative cleaning species.

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 the putative teleost homologs of the mammalian hippocampus and basolateral amygdala, broadly associated with spatial and social cognition (52, 99, 100), 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 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 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 provide strong support for shared neuromolecular basis of facultative cleaning and 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.

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. 2-4) 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. 2B). 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 (973; Fig. 2A). In fact, only one other combination of species exhibited more differentially expressed genes (Supplementary Table S2). Comparing the two most closely related non-cleaner species to the remaining three species yield 981 differentially expressed genes, highlighting the role of evolutionary history in shaping gene expression patterns (13, 101, 102). Interestingly, the differential expression analysis revealed a strong directional bias between cleaner and non-cleaner comparisons. The number of genes exhibiting increased expression in the facultative cleaner species was more than double the number exhibiting decreased expression (Fig. 2A). Similarly, nearly twice as many genes exhibiting increased expression in the obligate cleaner as compared to the non-cleaners (Supplementary Fig. 2A). Biased directional expression differences did not result from the unbalance design of the differential expression analysis (i.e., comparing two facultative to three non-cleaner species), as comparisons across all combinations of species revealed that the number of differentially expressed genes was not associated with the number of species included (Supplementary Table S2). Finally, we did not find similar bias in gene activation when comparing facultative and obligate cleaner species (Supplementary Fig. 2B). 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 (80–83) and the strong species-specific expression patterns (Supplementary Figure S1B and S7A), our finding of consistent directional bias in expression variation is notable; however, why a bias towards gene activation would accompany the evolution of mutualistic cleaning is unclear. A possible clue comes from the pathway analysis of genes differentially expressed between facultative cleaners and non-cleaners, which uncovered several KEGG pathways broadly associated with synaptic function and plasticity, neuronal growth, and neurite elongation (Supplementary Table S3). Thus, increased activity of these pathways in mutualistic cleaning species could reflect differences in learning-dependent synaptic plasticity (103) across species, amenable for future testing.

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. 3). While this 81 gene module is relatively small, compared to the other modules defined by WGCNA (Fig. 3A), 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. 3; 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 (38). 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. 4). Specifically, forty-one 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. 4). All 41 genes exhibiting

increased expression in the facultative cleaner species, again indicating a bias towards increased gene expression in the evolution of mutualistic cleaning behavior. Half of these genes are broadly associated with neural development and function (Fig. 4). To our knowledge, none of these candidate genes have been previously associated the mutualistic cleaning, facultative or obligate. 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 (104, 105) (Supplementary Table S7). 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. 71% 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. While gene co-expression patterns (as determined by WGCNA) did show considerable module preservation across all three behavioral

types (Fig. 5C), when we then examined the genes contained in the magenta "cleaning" module (which we had previously found to be associated with facultative cleaning) more closely, we discovered that, compared to both facultative and non-cleaners, these genes were most highly connected with rest of the transcriptome in obligate cleaners (Fig. 5D,E). 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 (38), 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 S5B). 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 (38), perhaps by maintaining an early life history state via neoteny. Life history transitions from a juvenile cleaner state to non-cleaner adult may result in decreased canalization (i.e., a reduction in the tendency for similar phenotypes to emerge regardless of internal and external context) and a resultant increased variation in the neuromolecular bases of cleaning

behavior. While we find no evidence for increased variation in expression in the facultative species (Supplementary Figure S5C and D), 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.

Methods

Study species and their ecological and life history attributes

Six species of wrasses (Labridae) were selected for this study (Fig. 1): *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 (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 (104, 105) 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 (106, 107).

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 (Fig. 1) 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. Fish were measured (Supplementary Table S3) and killed immediately on the boat by cervical transection and whole heads were transferred into 30 mL conical tubes with RNALater (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 (108) for sequencing on an Illumina HiSeq 2000 instrument (Illumina, San Diego, CA, USA). RNA quality control, library construction, and sequencing were performed by the University of Texas at Austin Genome Sequencing and Analysis Facility.

Read preprocessing and alignment and gene expression quantification

3'tagseq raw reads were preprocessed prior to alignment using the following pipeline. Briefly, custom perl scripts (after 109) using the FASTX-toolkit (110) and CUTADAPT v. 2.8 (111) were as used to remove reads with a homo-polymer run of "A"≥8 bases, retain reads with minimum 20 bases, removal PCR duplicates – defined as sequences sharing the same degenerate header and 20 bases of sequence – and filtered for quality (Phred quality score > 20 for 90% of the nucleotides). Preprocessed reads were aligned to the *Oreochromis niloticus* coding sequences (Orenil1.0 Ensembl cds) using the Burrows-Wheeler Aligner (bwa-mem; 112) resulting in an average mapping percentage of 4.9%. In addition to the *O. niloticus* reference, reads were mapped to a *de novo* assembled brain transcriptome from the more closely related (58 million years diverged)

ocellated wrasse, *Symphodus ocellatus* (113). Mapping reads to the *S. ocellatus de novo* assembled contigs resulted in an increased overall mapping percentage (average contig mapping percentage 19.4%) as compared to mapping to the *O. niloticus* reference; however, few reads mapped to contigs annotated with known gene ids (average gene mapping percentage, 1.3%). To facilitate downstream interpretation, the remaining analyses were done using gene counts from *O. niloticus* read mapping. Reads mapped to the *O. niloticus* were converted to counts using samtools (idxstats; 114) and TMM normalized (115) using R package NOISeq (116). TMM normalized expression values were used for all downstream analyses. We filtered the gene set such that that each gene was expressed in at least one individual of each species. The resulting set of 10,451 genes was used for all downstream analyses.

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 (94). 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 (Fig. 1) 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 S6; 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 S7). 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 species (Supplementary Figure S7B). Conversely, when we selected the genes with the lowest (bottom 1%) $-\log_{10}\beta_i$ values (i.e., genes whose expression is variable but not divergent between species), hierarchical clustering did not reveal any robust patterns either by species or any other attribute (Supplementary Figure S7C). We calculated Pearson's correlations in gene expression between all pairwise individuals. Individuals were clustered using 1-correlation and correlations were plotted in a heatmap using R package pheatmap (117). Cluster support values were generated by using R package pvclust (1000 permuations; , 118).

Differential expression analysis of cleaner types and species

Gene exhibiting differences in expression across cleaner types and species were identified using the R package *limma* (linear models for microarray data: , 119). For cleaner type comparisons, all individuals of a particular cleaner type were included as biological replicates even

if they were different species. To determine whether there were more DEGs between the facultative cleaners (*H. melanurus* and *T. lunare*) and non-cleaners (*C. batuensis, H. melapterus,* and *L. unilineatus*) we performed all possible comparisons of those five species (nine comparison). We then quantified the probability of discovering the observed number of DEGs between the facultative cleaners and non-cleaners using the empirical cumulative distribution function (*ecdf*) in R. For species-level comparisons, we calculated the number of DEGs for each pairwise combination of species. To test the hypothesis that facultative species are more similar in gene expression than is expected by chance, we compared the number of DEGs between the two facultative species to the distribution of differentially expressed gene of all the remaining species pair combinations using *ecdf* in R as above. For all analyses, differential expression was defined by an FDR adjusted p-value < 0.05 unless otherwise indicated. We used TMM-normalized counts of the 10,451 genes as the input data set for all limma contrasts.

KEGG pathway analysis

To functionally characterize genes associated with facultative cleaning we performed pathway analysis using pathfindeR (88). Using the input gene set, pathfindR first identifies gene that are interconnected in protein-protein interaction network and second identifies pathways enriched in the interconnected sets. As recommended for analysis of non-human data, we used the StringDB functional protein association network for *Oreochromis niloticus* (120) to identify interconnected genes and identified KEGG pathways (86) enriched in those interconnected gene sets. Genes differentially expressed (at p-value < 0.05; 2254 genes) were included in the analysis. Because StringDB used gene symbols only those 1523 genes that could be converted from *Oreochromis niloticus* ENSEMBL ids could be include in the analysis.

642

643 *Phylogenetic Comparative Analysis of Gene Co-expression*

644 To capture genes with correlated expression variation across individuals, species, and
645 cleaner types, we performed a Weighted Gene Co-expression Network Analysis (WGCNA) (121)
646 including the genes shared across all species. WGCNA clusters genes by expression similarity and
647 summarizes gene co-expression as module eigengenes (i.e., the first principal component of all the
648 genes in each co-expression module). Each eigengene is the linear combination of gene expression
649 values that explains the most variation in the expression of the genes contained in the module. For
650 all analyses, WGCNA was performed with a minimum modules size of 50 and soft power
651 thresholds were determined using WGCNA's softPower function.

652 To identify modules associated with facultative cleaning, we performed WGCNA
653 including the two facultative cleaner species and the three non-cleaner species (Fig. 1). We
654 assessed changes in gene co-expression across species and cleaner types using a Welch's t-test
655 (due to unequal variances) or ANOVA depending on whether two or more groups were compared.
656 For pairwise comparisons (e.g., among species) we used Tukey's honest signal difference post hoc
657 test for ANOVAs with F-statistics with significance at $p < 0.05$. For cleaner type comparison
658 between facultative and non-cleaner species, we confirmed associations between cleaner type and
659 module eigengene taking phylogenetic non-independence into account using a phylogenetic
660 ANOVA. Phylogenetic ANOVA was performed with the R package phytools (122). A
661 phylogenetic ANOVA could not confirm whether any modules associate with obligate cleaning
662 due to the inclusion of only one representative obligate cleaner species (*L. dimitiatus*). To
663 determine the probability of identifying WGCNA modules associated with any subset of species
664 (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 S6)

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

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

To ask whether gene co-expression network structure persist across cleaner types, we quantified preservation of the focal modules – defined in the WGCNA analysis including facultative and non-cleaner species only – in WGCNA analysis including only species from each cleaner type (i.e., the obligate cleaner *L. dimidiatus*, the facultative cleaners *T. lunare* and *H. melanurus* only, or the non-cleaners *H. melapterus*, *L. unilineatus*, and *C. batuensis*). Module preservation is quantified as the $Z_{summary}$ using the *modulePreservation* in WGCNA. $Z_{summary}$ integrates multiple preservation statistics into a single overall measure and can be used to determine conservation of gene connectivity of modules of interest across distinct WGCNA analyses (121). The higher the value of a $Z_{summary}$ the more preserved the gene connectivity of the module across analyses, with values greater than ten considered strongly preserved, values between two and ten considered weakly preserved, and values less than two considered not preserved (121). To further characterize preservation of gene connectivity in cleaner species, we compared gene connectivity (kTotal) of genes from focal modules (Fig. 3) in WGCNA performed on non-cleaners, facultative cleaners, and the obligate cleaner alone. kTotal was calculated using WGCNA's *intramodularConnectivity* function and measures connectivity of each gene with all other genes in the transcriptome. We asked whether gene connectivity of each module different among cleaner types using ANOVA followed by estimated marginal mean and Cohen's D effect size *posthoc* pairwise comparisons between cleaner types (Fig. 5D; Supplementary Figure S4; Supplementary Table S8) using the R packages *emmeans* (123) and *rstatix* (124), respectively. Spearman's rank correlation was used to determine similar of gene connectivity between all pairwise cleaner types for genes in each focal module (Fig. 3) independently. Spearman's rank

correlations were performed using the *rcorr* function from R package *Hmisc* (125) and p-values were adjusted for multiple hypothesis testing using R stats function *p.adjust* and the Benjamini-Hochberg method.

No evidence of expression specialization or canalization in the obligate 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 facultative cleaners. We asked whether the 973 DEGs between facultative cleaner and non-cleaner species, the 197 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. To test the hypothesis that life history transitions from a juvenile cleaner state to non-cleaner adult state in facultative cleaners is associated with increased variation in the neuromolecular bases of cleaning, we tested for equivalence between species in expression variation for all 973 DEGs, 197 DEGs shared and directionally concordant in facultative and obligate cleaners versus non-cleaners, and magenta module genes using the TOSTtwo function of the TOSTER package in R (126).

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.).

749 Literature Cited

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

752 2. D. L. Stern, The genetic causes of convergent evolution. *Nat Rev Genet.* **14** (2013).

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

755 4. J. B. Losos, CONVERGENCE, ADAPTATION, AND CONSTRAINT. *Evolution (N. Y).*
756 **65**, 1827–1840 (2011).

757 5. P. M. Brakefield, Evo-devo and constraints on selection. *Trends Ecol. Evol.* **21**, 362–368
758 (2006).

759 6. S. J. Gould, *Wonderful life - the Burgess shale and the nature of history* (WW. Norton &
760 Company, 1991).

761 7. V. Orgogozo, Replaying the tape of life in the twenty-first century. *Interface Focus* **5**
762 (2015).

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

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

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

770 11. D. L. Mahler, T. Ingram, L. J. Revell, J. B. Losos, Exceptional Convergence on the
771 Macroevolutionary Landscape in Island Lizard Radiations. *Science (80-.).* **341**, 292 LP –
772 295 (2013).

773 12. R. B. Langerhans, Predictability and Parallelism of Multitrait Adaptation. *J. Hered.* **109**,
774 59–70 (2017).

775 13. S. C. P. Renn, *et al.*, Gene expression signatures of mating system evolution. *Genome* **61**,
776 287–297 (2018).

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

779 15. R. L. Young, G. P. Wagner, Why ontogenetic homology criteria can be misleading:
780 lessons from digit identity transformations. *J. Exp. Zool. B. Mol. Dev. Evol.* **316B**, 165–70
781 (2011).

782 16. J. R. True, E. S. Haag, Developmental system drift and flexibility in evolutionary
783 trajectories. *Evol. Dev.* **3**, 109–119 (2001).

784 17. J. F. Storz, Causes of molecular convergence and parallelism in protein evolution. *Nat.*
785 *Rev. Genet.* **17**, 239–250 (2016).

18. Y. Cheng, *et al.*, Parallel genomic responses to historical climate change and high elevation in East Asian songbirds. *Proc. Natl. Acad. Sci.* **118**, e2023918118 (2021).

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

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

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

22. T. A. Castoe, *et al.*, Evidence for an ancient adaptive episode of convergent molecular evolution. *Proc. Natl. Acad. Sci.* **106**, 8986 LP – 8991 (2009).

23. J. R. Gallant, *et al.*, Genomic basis for the convergent evolution of electric organs. *Sci.* **344**, 1522–1525 (2014).

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

25. A. R. Pfenning, *et al.*, Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science (80-.).* **346**, 1256846 (2014).

26. E. T. Chan, *et al.*, Conservation of core gene expression in vertebrate tissues. *J. Biol.* **8**, 33 (2009).

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

28. W. Gao, *et al.*, Genomic and transcriptomic investigations of the evolutionary transition from oviparity to viviparity. *Proc. Natl. Acad. Sci.* **116**, 3646 LP – 3655 (2019).

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

30. S. Lamichhaney, *et al.*, Integrating natural history collections and comparative genomics to study the genetic architecture of convergent evolution. *Philos. Trans. R. Soc. B Biol. Sci.* **374**, 20180248 (2019).

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

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

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

34. D. B. Vaughan, A. S. Grutter, M. J. Costello, K. S. Hutson, Cleaner fishes and shrimp

diversity and a re-evaluation of cleaning symbioses. *Fish Fish.* **18**, 698–716 (2017).

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

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

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

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

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

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

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

42. M. C. Soares, *et al.*, Hormonal mechanisms of cooperative behaviour. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **365**, 2737–2750 (2010).

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

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

45. C. A. Weitekamp, H. A. Hofmann, Evolutionary themes in the neurobiology of social cognition. *Curr. Opin. Neurobiol.* **28**, 22–27 (2014).

46. S. C. Cardoso, *et al.*, Arginine vasotocin modulates associative learning in a mutualistic cleaner fish. *Behav. Ecol. Sociobiol.* **69**, 1173–1181 (2015).

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

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

49. B. M. Nugent, *et al.*, Morpholino-mediated knockdown of the brain mineralocorticoid receptor affects glucocorticoid signaling and neuroplasticity in wild ocellated wrasse (Symphodus ocellatus). *bioRxiv*, 2021.11.17.468986 (2021).

50. M. S. Lamm, A. D. Lamb, B. P. Klapheke, W. A. Tyler, J. R. Godwin, 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 (2022).

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

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

53. J. O'Keefe, L. Nadel, *The hippocampus as a cognitive map* (Clarendon Press, 1978).

54. P. Andersen, R. Morris, D. Amaral, T. Bliss, J. O'Keefe, *The Hippocampus Book* (Oxford University Press, 2007).

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

56. G. F. Striedter, Evolution of the hippocampus in reptiles and birds. *J. Comp. Neurol.* **524**, 496–517 (2016).

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

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

59. A.-T. Trinh, S. E. Clarke, E. Harvey-Girard, L. Maler, Cellular and Network Mechanisms May Generate Sparse Coding of Sequential Object Encounters in Hippocampal-Like Circuits. *eneuro* **6**, ENEURO.0108-19.2019 (2019).

60. E. Vinepinsky, *et al.*, Representation of edges, head direction, and swimming kinematics in the brain of freely-navigating fish. *Sci. Rep.* **10**, 14762 (2020).

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

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

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

64. F. Martínez-García, A. Martínez-Marcos, E. Lanuza, The pallial amygdala of amniote vertebrates: evolution of the concept, evolution of the structure. *Brain Res. Bull.* **57**, 463–469 (2002).

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

66. A. Zayed, G. E. Robinson, Understanding the Relationship Between Brain Gene

Expression and Social Behavior: Lessons from the Honey Bee. *Annu. Rev. Genet.* **46**, 591–615 (2012).

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

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

69. C. C. Rittschof, K. A. Hughes, Advancing behavioural genomics by considering timescale. *Nat. Commun.* **9**, 489 (2018).

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

71. E. K. Fischer, M. E. Hauber, A. M. Bell, 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 (2021).

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

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

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

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

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

77. T. K. Mehta, *et al.*, Evolution of regulatory networks associated with traits under selection in cichlids. *Genome Biol.* **22**, 25 (2021).

78. G. Hu, *et al.*, Evolutionary Conservation and Divergence of Gene Coexpression Networks in Gossypium (Cotton) Seeds. *Genome Biol. Evol.* **8**, 3765–3783 (2016).

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

80. T. Bedford, D. L. Hartl, Optimization of gene expression by natural selection. *Proc. Natl. Acad. Sci.* **106**, 1133 LP – 1138 (2009).

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

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

83. P. Khaitovich, *et al.*, A Neutral Model of Transcriptome Evolution. *PLoS Biol.* **2**, 682–689

(2004).

84. R. L. Young, H. A. Hofmann, Reply to Jiang and Zhang: Parallel transcriptomic signature of monogamy: What is the null hypothesis anyway? *Proc. Natl. Acad. Sci.* **116**, 17629 LP – 17630 (2019).

85. C. W. Dunn, X. Luo, Z. Wu, Phylogenetic Analysis of Gene Expression. *Integr. Comp. Biol.* (2013) https:/doi.org/10.1093/icb/ict068.

86. H. Ogata, *et al.*, KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* **27**, 29–34 (1999).

87. M. Kanehisa, Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951 (2019).

88. E. Ulgen, O. Ozisik, O. U. Sezerman, pathfindR: An R Package for Comprehensive Identification of Enriched Pathways in Omics Data Through Active Subnetworks . *Front. Genet.* **10** (2019).

89. M. Hetman, L. P. Slomnicki, Ribosomal biogenesis as an emerging target of neurodevelopmental pathologies. *J. Neurochem.* **148**, 325–347 (2019).

90. A. N. Hegde, The ubiquitin-proteasome pathway and synaptic plasticity. *Learn. Mem.* **17**, 314–327 (2010).

91. M. E. Kerrisk, L. A. Cingolani, A. J. Koleske, ECM receptors in neuronal structure, synaptic plasticity, and behavior. *Prog. Brain Res.* **214**, 101–131 (2014).

92. K. Itoh, S. Nishihara, Mucin-Type O-Glycosylation in the Drosophila Nervous System . *Front. Neuroanat.* **15** (2021).

93. N. Takei, H. Nawa, mTOR signaling and its roles in normal and abnormal brain development . *Front. Mol. Neurosci.* **7** (2014).

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

95. S. Horvath, *et al.*, Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target. *Proc. Natl. Acad. Sci.* **103**, 17402 LP – 17407 (2006).

96. Y. Liu, *et al.*, Identification of Hub Genes and Key Pathways Associated With Bipolar Disorder Based on Weighted Gene Co-expression Network Analysis . *Front. Physiol.* **10**, 1081 (2019).

97. P. S. Gargalovic, *et al.*, Identification of inflammatory gene modules based on variations of human endothelial cell responses to oxidized lipids. *Proc. Natl. Acad. Sci.* **103**, 12741 LP – 12746 (2006).

98. P. Langfelder, S. Horvath, Eigengene networks for studying the relationships between co-expression modules. *BMC Syst. Biol.* **1**, 54 (2007).

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

100. Y. Yang, J.-Z. Wang, From Structure to Behavior in Basolateral Amygdala-Hippocampus Circuits . *Front. Neural Circuits* **11** (2017).

101. D. Jiang, J. Zhang, Parallel transcriptomic changes in the origins of divergent monogamous vertebrates? *Proc. Natl. Acad. Sci.* **116**, 17627 LP – 17628 (2019).

102. D. Brawand, *et al.*, The evolution of gene expression levels in mammalian organs. *Nature* **478**, 343–348 (2011).

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

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

105. J. E. Randall, G. R. Allen, R. C. Steene, *Fishes of the Great Barrier Reef & Coral Sea*, 2nd Ed. (University of Hawaii Press, 1988).

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

107. S. Kumar, G. Stecher, M. Suleski, S. B. Hedges, TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Mol. Biol. Evol.* **34**, 1812–1819 (2017).

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

109. E. Meyer, G. V Aglyamova, M. V Matz, 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 (2011).

110. G. J. Hannon, FASTX-Toolkit (2010).

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

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

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

114. H. Li, *et al.*, The Sequence Alignment/Map (SAM) Format and SAMtools. *Bioinformatics* **25**, 2078–9 (2009).

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

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

117. R. Kolde, Pheatmap: pretty heatmaps. *R Packag. version* **1**, 726 (2012).

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

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

120. D. Szklarczyk, *et al.*, STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, D447–D452 (2015).

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

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

123. R. V. Lenth, emmeans: Estimated Marginal Means, aka Least-Squares Means (2022).

124. A. Kassambara, rstatix: Pipe-Friendly Framework for Basic Statistical Tests (2020).

125. F. E. J. Harrell, Hmisc: Harrell Miscellaneous (2020).

126. D. Lakens, Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. *Soc. Psychol. Personal. Sci.* **8**, 355–362 (2017).

127. M. Safran, *et al.*, "The GeneCards Suite Chapter" in *Practical Guide to Life Science Databases*, (2022), pp. 27–56.

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. Number of comparable genes across species shown in Figure 1.

Supplementary Table S2. Number of DEGs in all two by three species comparisons. The comparison of facultative and non-cleaner species shown in Figure 2A is bolded.

Supplementary Table S3. 47 KEGG pathways enriched in genes differentially expressed (*p*-value < 0.05) between the facultative cleaner and non-cleaner species.

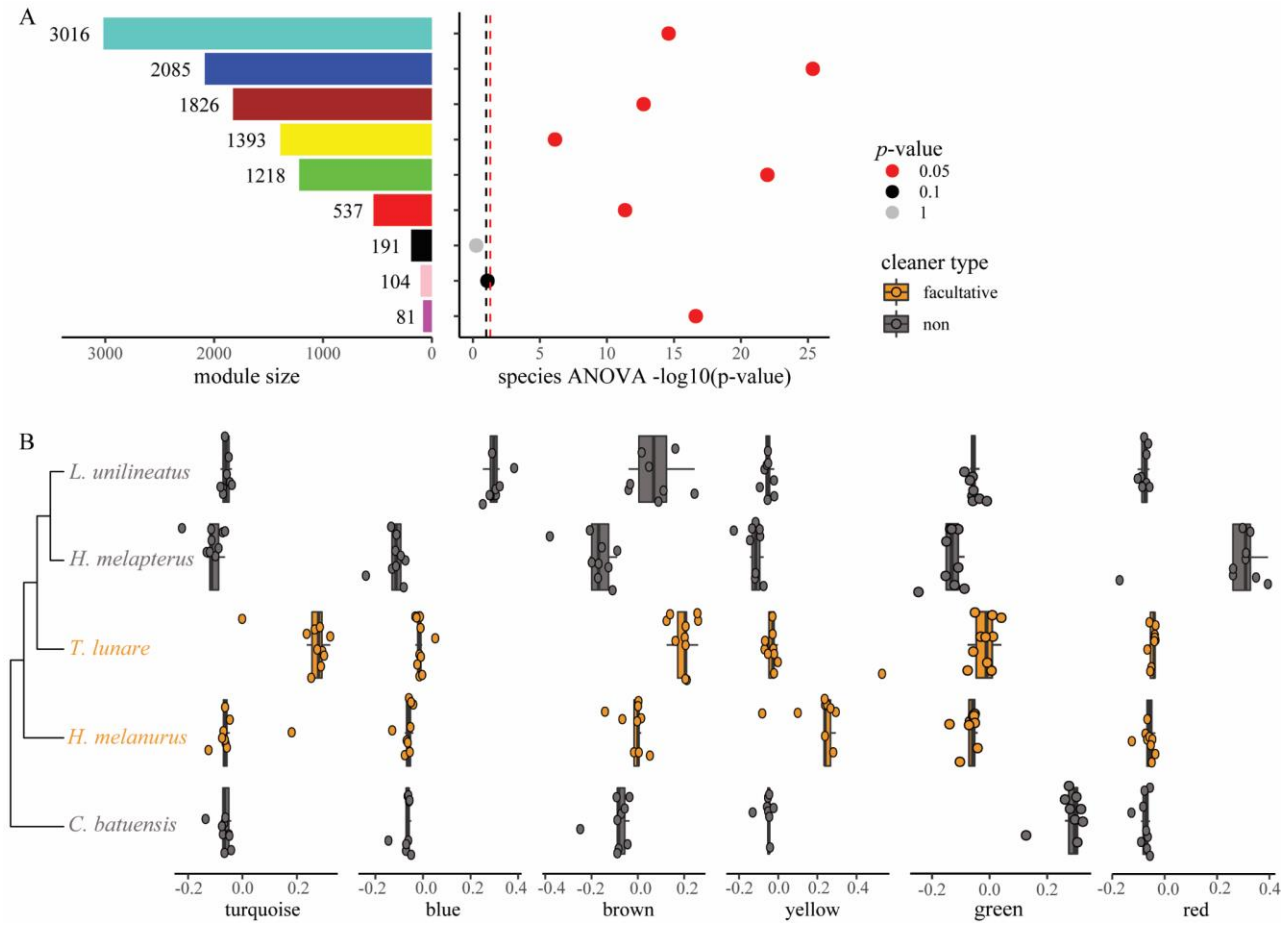
Supplementary Table S4. Variation in module eigengene expression between facultative cleaners and non-cleaners. Welch's t-statistics, phylogenetic ANOVA F-statistics, and *p*-values for all modules a provided. Modules are identified by color and size indicates the number of genes in each module. Significance at $p < 0.05$ after accounting for phylogenetic independence is indicated in bold.

Supplementary Table S5. 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.

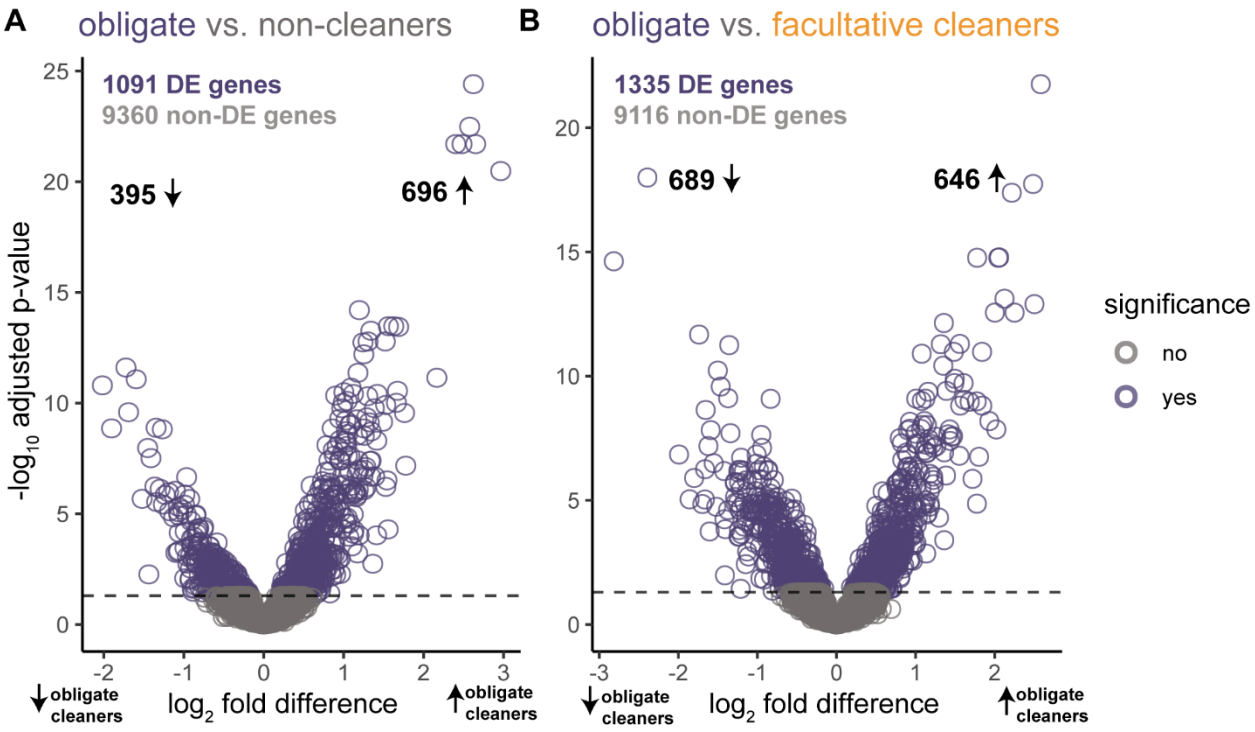
Supplementary Table S6. Candidate genes associated with independent transitions to facultative cleaning. ENSEMBL IDs, Gene IDs, functional annotations from GeneCards (127) and PubMed, $\log_2$ fold difference and associated adjusted p-value between facultative cleaners and non-cleaner, magenta module intramodular connectivity, and EVE model negative $\log_{10}$ beta values are provided for each gene.

Supplementary Table S7. 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 (40, 104, 105).

Supplementary Table S8. Results of an analysis of variance (ANOVA) comparing focal module gene connectivity across cleaner types, estimated marginal mean and Cohen's D effect size *posthoc* 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).

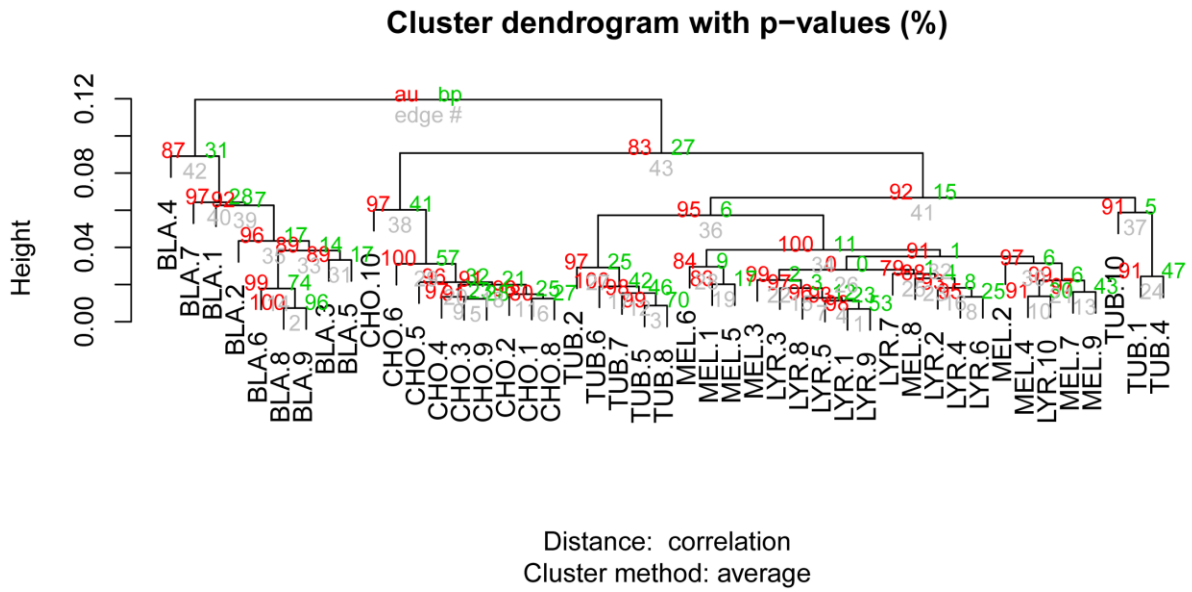


Supplementary Figure S1. Seven gene co-expression modules differed across species (A). The number of gene contain in each module is indicated by module size (A). Dashed lines indicated statistical support for the ANOVA at $p < 0.1$ (black) and $p < 0.05$ (red) (A). Species-level co-expression comparisons 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.



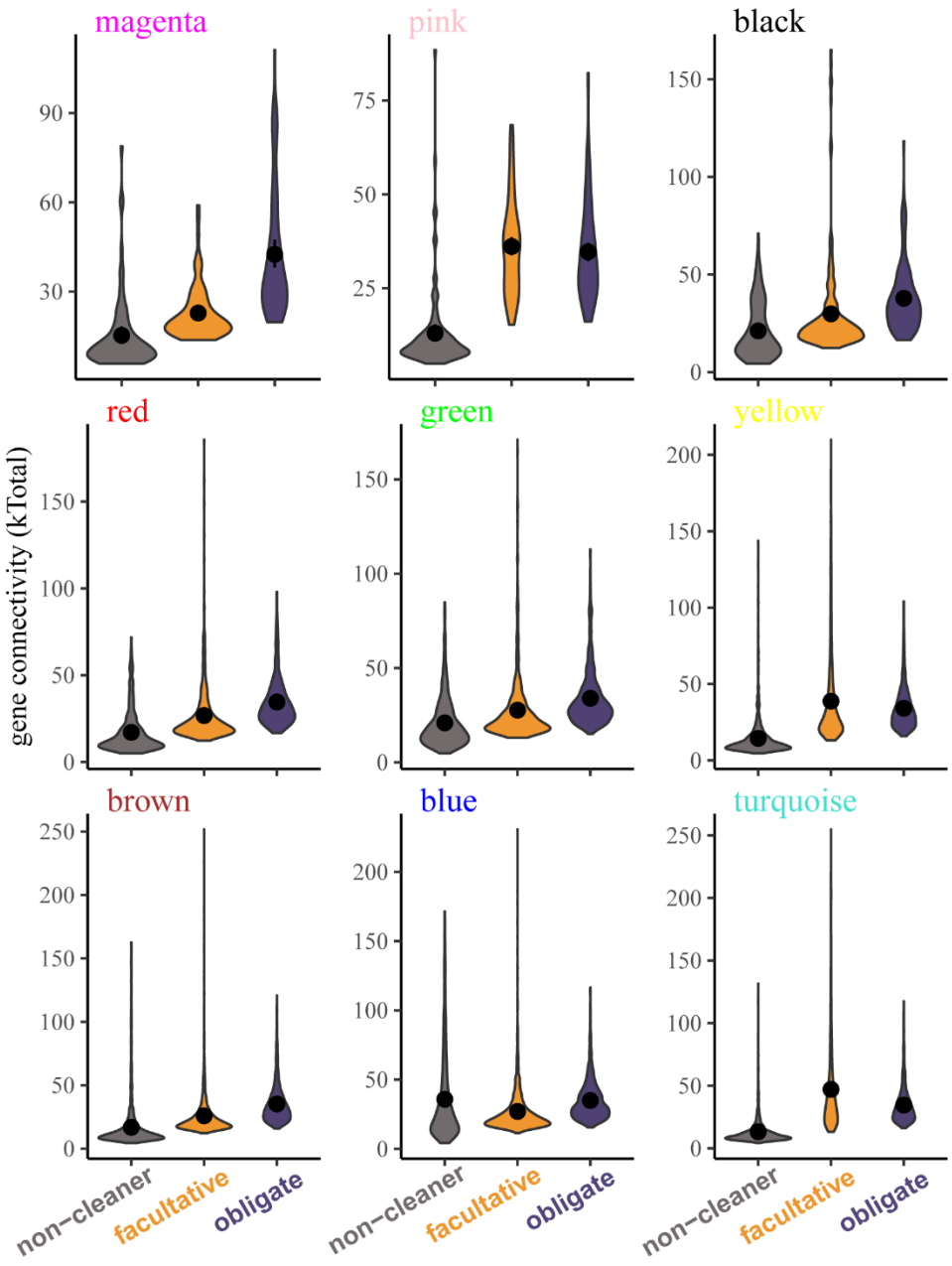
Supplementary Figure S2. Differential gene expression between obligate and non-cleaner species (A) and obligate and facultative cleaner species (B).

1072



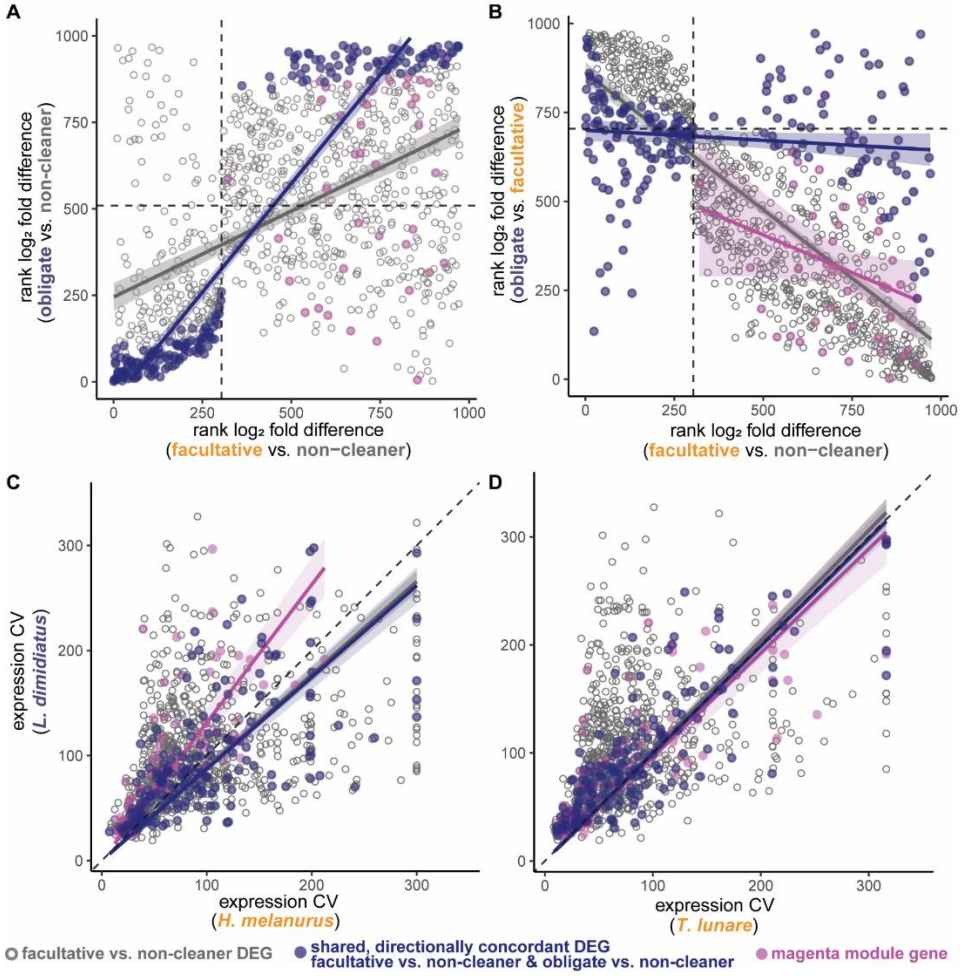
Supplementary Figure S3. Approximately unbiased p-values (red) and bootstrap probabilities (green) for sample clustering by facultative cleaning candidate gene expression Fig. 4B. Height indicates 1-correlation. While facultative cleaners *Halichoeres melanurus* (MEL) and *Thalassoma lunare* (LYR) cluster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

1073



Supplementary Figure S4. Differences in gene connectivity across cleaner types. Gene connectivity (kTotal) was obtained from WGCNA on each cleaner type independently Connectiviy 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 5D; Supplementary Table S8). Further, gene

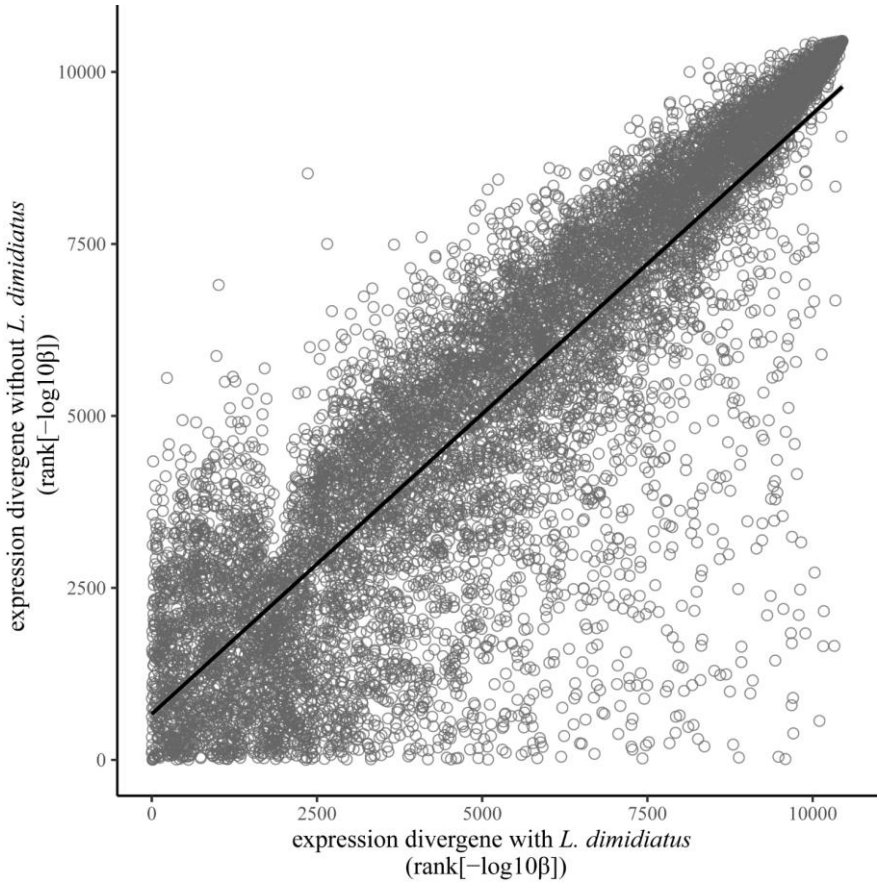
connectivity was correlated in all pairwise comparisons of cleaner types (Figure 5E; Supplementary Table S8). Means and standard deviations are shown as dots and whiskers within each violin plot. Modules are arranged in order of size (i.e., number of genes) with magenta being the smallest module and turquoise the largest (Figure 3).



Supplementary Figure S5. 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). Neither expression differences (B) nor variation (C and D) between obligate (*L. dimidiatus*) and facultative cleaners (*H. melanurus* and *T. lunare*) support the hypotheses that obligate cleaning is a specialization (B) or a canalized life history state (C and D) of facultative cleaning. We found a negative correlation between gene expression differences in facultative and non-cleaner species versus obligate and

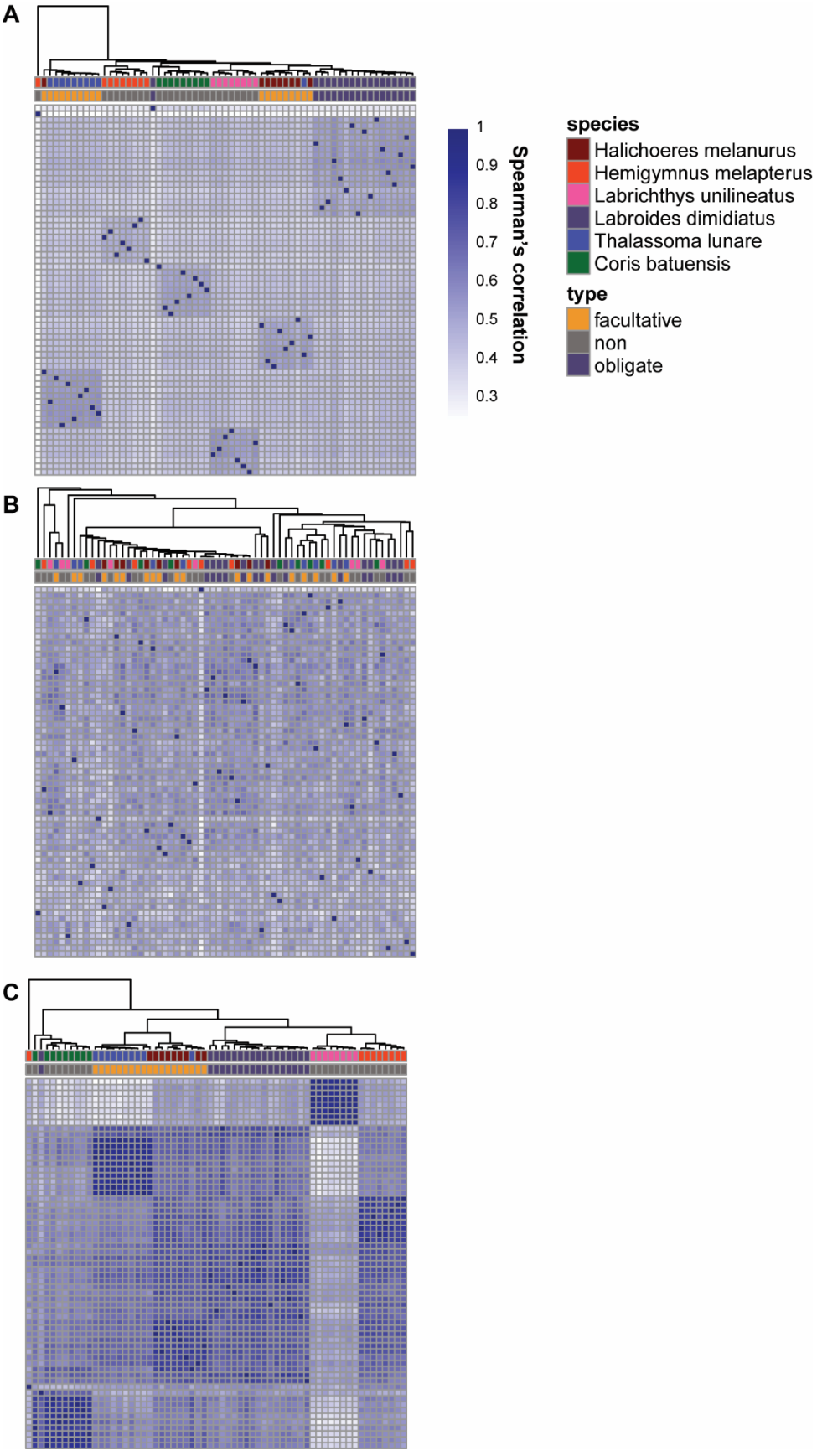
facultative species for all differentially expressed genes (B, grey: Spearman's $\rho = -0.77$, $p < 2.2e-16$), genes differentially expressed in both facultative cleaners and the obligate cleaner as compared to the non-cleaner species (B, blue: Spearman's $\rho = -0.18$, $p = 0.011$), and genes differentially expressed and contained in the facultative cleaner-associated magenta module (B, magenta: Spearman's $\rho = -0.31$, $p = 0.046$). Genes from the facultative cleaner associated magenta module are more (B) or equally variable (C) in the obligate cleaner (*L. dimidiatus*) as compared to two the facultative cleaning species. The 973 DEGs between facultative and non-cleaner species are slightly less variable (C) or equally variable (D) in the obligate cleaner (*L. dimidiatus*) as compared to the facultative cleaning species. For all DEGs and for the magenta module genes, variation between the facultative cleaners (*H. melanurus* and *T. lunare*) and the obligate cleaner (*L. dimidiatus)* differed significantly from zero (C, *H. melanurus* – DEGs: $t_{(2045.5)} = -4.1$, *p*-value = 5.0e-5 and magenta genes: $t_{(112.7)} = -5.5$, *p*-value = 2.1e-7; D, *T. lunare* – DE genes: $t_{(2061.9)} = -9.6$, *p*-value = 1.7 e-21 and magenta genes: $t_{(116.3)} = -6.0$, *p*-value = 2.2e-8) and did not exhibit equivalent variance (C, *H. melanurus* – DE genes: $t_{(2045.5)} = -2.9$, *p*-value = 0.998 and magenta genes: $t_{(112.7)} = -5.5$, *p*-value = 1.0; D, *T. lunare* – DE genes: $t_{(2061.9)} = -8.5$, *p*-value = 1.0 and magenta genes: $t_{(116.3)} = -5.7$, *p*-value = 1.0). In comparison with *H. melanurus*, there was a small reduction in variation in *L. dimidiatus* when all DEGs were compared and in comparison with *T. lunare* there was a very small reduction in variation in *L. dimidiatus* when the magenta DEGs were compared, providing very limited support for expression canalization (i.e., a reduction in variability). For the 197 DEGs shared and directionally concordant in both the facultative and obligate cleaner versus non-cleaner comparison variation between the facultative cleaners (*H. melanurus* and *T. lunare*) and the obligate cleaner (*L. dimidiatus)* did not differ significantly from zero (C, *H. melanurus* – shared, directionally concordant DEGs – $t_{(391.5)} = 0.44$, *p*-value = 0.33; D,

T. lunare – shared, directionally concordant DEGs – $t_{(390.3)} = -1.4$, p-value = 0.06).Variation in expression is quantified as the coefficient of variation (C and D). Dashed lines indicate the expression rank at which DEGs change from decreased to increased expression (A and B) or indicate equivalent CVs on the x- and y-axes (C and D). For all plots, 973 DEGs between facultative and non-cleaner species are shown in grey, DEGs shared and directionally concordant in obligate vs. non-cleaners are blue, magenta module genes are magenta, and shading indicates standard error.



Supplementary Figure S6. 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 $\rho = 0.87$, $p < 2.2e\text{-}16$).

1127


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

1128

Supplementary Figure S7. 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.