NRC Research Press

ARTICLE

Gene expression signatures of mating system evolution[1]

Suzy C.P. Renn, Heather E. Machado, Nina Duftner, Anna K. Sessa, Rayna M. Harris, and Hans A. Hofmann

Abstract: The diversity of mating systems among animals is astounding. Importantly, similar mating systems have evolved even across distantly related taxa. However, our understanding of the mechanisms underlying these convergently evolved phenotypes is limited. Here, we examine on a genomic scale the neuromolecular basis of social organization in cichlids of the tribe Ectodini from Lake Tanganyika. Using field-collected males and females of four closely related species representing two independent evolutionary transitions from polygyny to monogamy, we take a comparative transcriptomic approach to test the hypothesis that these independent transitions have recruited similar gene sets. Our results demonstrate that while lineage and species exert a strong influence on neural gene expression profiles, social phenotype can also drive gene expression evolution. Specifically, 331 genes (~6% of those assayed) were associated with monogamous mating systems independent of species or sex. Among these genes, we find a strong bias (4:1 ratio) toward genes with increased expression in monogamous individuals. A highly conserved nonapeptide system known to be involved in the regulation of social behavior across animals was not associated with mating system in our analysis. Overall, our findings suggest deep molecular homologies underlying the convergent or parallel evolution of monogamy in different cichlid lineages of Ectodini.

Key words: cichlid, gene expression, monogamy, polygyny, deep homology.

Résumé : La diversité des systèmes de reproduction parmi les animaux est étonnante. Fait important, des systèmes de reproduction semblables sont apparus chez des taxons très éloignés. Cependant, les mécanismes derrière cette convergence phénotypique demeurent méconnus. Dans ce travail, les auteurs examinent à l'échelle génomique les bases neuromoléculaires de l'organisation sociale chez des Ectodini, une tribu au sein des cichlidés retrouvés dans le lac Tanganyika. À l'aide de mâles et de femelles capturés sur le site et appartenant à quatre espèces très proches représentant deux transitions évolutives indépendantes de la polygynie à la monogamie, les auteurs utilisent une approche de transcriptomie comparée pour vérifier l'hypothèse voulant que ces transitions indépendantes se sont opérées à l'aide de jeux de gènes semblables. Les résultats montrent que, bien que la lignée et l'espèce exercent une grande influence sur les profils d'expression des gènes neuraux, le phénotype social peut également diriger l'évolution de l'expression génique. Spécifiquement, 331 gènes (~6 % des gènes étudiés) étaient associés à la monogameie indépendamment de l'espèce ou du sexe. Parmi ces gènes, les auteurs ont observé un fort biais (rapport de 4:1) en faveur de gènes plus fortement exprimés chez les individus monogames. Un système hautement conservé de nonapeptides, connu comme étant impliqué dans la régulation des comportements sociaux chez les animaux, a été associé au système de reproduction dans cette analyse. Globalement, ces résultats suggèrent que des homologies moléculaires profondément ancrées sous-tendent l'évolution convergente ou parallèle de la monogamie chez différentes lignées de cichlidés appartenant à la tribu des Ectodini. [Traduit par la Rédaction]

Mots-clés : cichlidés, expression génique, monogamie, polygynie, homologie profonde.

Introduction

Social behavior can vary greatly across species, and such diversity lends itself to comparative analyses of the underlying proximate mechanisms (Crews and Moore 1986). The evolution of mating systems is an excellent example of how modification of an animal's reproductive behavior in response to environmental conditions can maximize fitness (Emlen and Oring 1977). Throughout the animal kingdom, monogamy, along with biparental care, has evolved repeatedly in response to ecological factors such as predation pressure or the availability and distribution of mates and resources (Grant 1993). At the mechanistic level, the role of the nonapeptide arginine vasopressin (AVP) and its V1a receptor in pair bond formation has been well established in the prairie vole, *Microtus ochrogaster* (Ophir et al. 2008) as well as mice of the genus *Peromyscus* (Bendesky et al. 2017). Importantly, AVP, and its non-mammalian homolog arginine vasotocin (AVT) have been shown to be associated with mating system variation in a range of vertebrates (reviewed by Oldfield et al. 2013, 2015), sug-

Received 31 March 2017. Accepted 21 August 2017.

Corresponding Editor: Bradley Olson.

S.C.P. Renn. Department of Biology, Reed College.
H.E. Machado. Department of Biology, Reed College; Department of Biology, Stanford University.
N. Duftner and A.K. Sessa. Department of Integrative Biology, the University of Texas at Austin.
R.M. Harris. Department of Integrative Biology, the University of Texas at Austin; Institute for Cellular and Molecular Biology, the University of Texas at Austin.
H.A. Hofmann. Department of Integrative Biology, the University of Texas at Austin; Institute for Cellular and Molecular Biology, the University of Texas at Austin; Center for Computational Biology and Bioinformatics, Institute for Neuroscience, the University of Texas at Austin.
Corresponding authors: Hans A. Hofmann (email: hans@utexas.edu); Suzy C.P. Renn (email: renns@reed.edu).

[1]This paper is part of a Special Issue entitled Ecological Genomics.

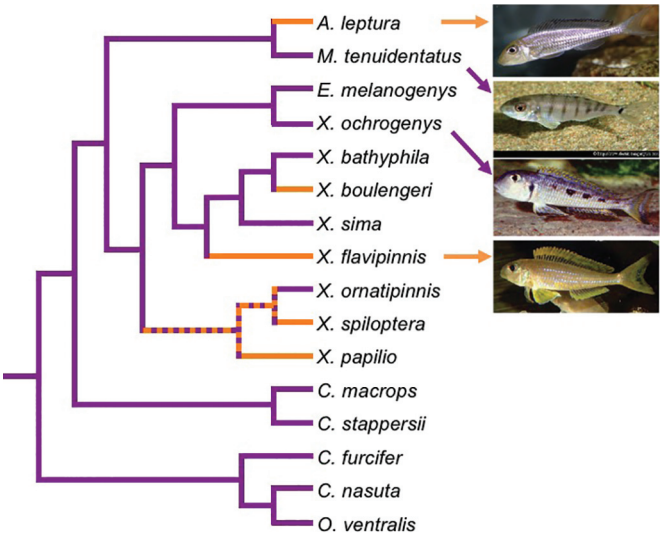
gesting deep molecular homologies underlying the independent evolution of sociality in different vertebrate clades. If we are to understand the neural and molecular building blocks of social behavior, and how these are shaped over developmental and evolutionary time, we need to examine closely related species with different mating strategies that differ in few ecological and morphological aspects. The extent to which patterns of deep homology more broadly explain repeated evolution of social phenotypes requires genome-wide studies to complement the analysis of candidate genes.

The cichlid family of fishes offers unique opportunities for genome-scale comparative studies examining the evolution of complex behavior patterns involved in behavioral decision making and motor output, as the rapid radiation of species with diverse social phenotypes allows for comparison across closely related species (Brawand et al. 2014; Hofmann 2003; Kocher 2004). Cichlids are a classic model system for studying the evolution of social behavior because of their extraordinary behavioral diversity (Kocher 2004; Pollen et al. 2007). The abundant literature on cichlid ecology, evolution, behavior, and genomics puts cichlids in an ideal position for studies on the molecular basis of ecological and evolutionary processes. The behavioral diversity and genomic resources of the cichlid system makes it uniquely suited to identify the molecular substrates underlying the evolution of mating systems across a range of carefully selected species. The cichlid lineages in Lake Tanganyika, in particular, display a wide range of mating strategies (e.g., monogamous and polygynous) and provide parental care in a maternal, biparental, or cooperative manner either via buccal incubation (mouth-brooding) or substrate guarding (see Barlow 2000 for review).

Within these lineages, the monophyletic cichlid tribe Ectodini exhibits striking differences in mating and parental care strategy, thus this clade is particularly well suited for a genomic exploration of the evolution of mating system. Habitat and social organization in this clade correlate with differences in gross morphology of the brain (Pollen et al. 2007) as well as visual behavior (Dobberfuhl et al. 2005). Importantly, Ectodini comprises species that display either monogamous (with biparental offspring care) or polygynous (with maternal only care) behavior and thus offers a unique opportunity to study the evolution of distinct mating systems within a phylogenetic framework. Kidd et al. (2012) reconstructed the evolutionary relationships among 16 species of this lineage and found that during 2.5–3 million years of evolution within this clade there have been 3–5 independent transitions from polygyny to monogamy (see also Koblmuller et al. 2004). These results suggest that the evolution of parental care strategies is not only remarkably fast, but also much more labile than previously expected.

Over the last decade, a multitude of genomic resources have been developed for cichlids (Aubin-Horth et al. 2007; Salzburger et al. 2008; Watanabe et al. 2004), including draft genomes and numerous transcriptomes for five African cichlid species (Brawand et al. 2014). Despite the amazing diversity in ecology and behavior, there is very little divergence in coding sequence (Loh et al. 2008; Salzburger et al. 2008; Watanabe et al. 2004). Importantly, a cDNA microarray constructed for Burton's Mouthbrooder cichlid, *Astatotilapia burtoni*, has been used to identify neural gene expression modules associated with complex social traits in this model system for social neuroscience (Duftner et al. 2008; O'Connell and Hofmann 2012; Renn et al. 2008; O'Rourke and Renn 2015). This array platform has also been applied successfully in a range of other teleost fishes for which little or no genomic information was available (Aubin-Horth et al. 2009; Cummings et al. 2008; Machado et al. 2009; Renn et al. 2004). Heterologous array hybridization (whereby a single array platform is used to study a group of closely related species) has been an effective tool for such comparative studies in several systems (e.g., Buckley 2007; Kassahn et al. 2007). This is particularly the case when competitive genomic

Fig. 1. Convergent evolution of mating strategies within the Ectodini and *Xenotilapia* clades from Lake Tanganyika. An ancestral character state reconstruction by maximum parsimony reveals multiple transitions from polygyny (purple) to monogamy (orange) adapted from Kidd et al. 2012. [Colour online.]



DNA hybridizations between species are used to provide a means for assessing the degree to which sequence divergence biases the results of expression studies. We previously introduced a masking technique that corrects interspecific gene expression differences with genomic DNA hybridization ratios for the same set of species and determines a threshold sequence divergence (Machado et al. 2009). This masking technique was first used to compare neural gene expression profiles between individual males and females from two closely related species of Ectodini, the polygynous *Enantiopus melanogenys* and the monogamous *Xenotilapia flavipinnis* (Machado et al. 2009). In that study, the authors asked whether sex-specific neural gene expression is more closely associated with mating system than with gonadal sex and found that the gene expression profiles are largely species specific, as relatively few genes showed conserved expression patterns associated with either sex (Machado et al. 2009). However, to test whether this pattern is generalizable to other independent (parallel) transitions from polygyny to monogamy in this system (Kidd et al. 2012) more species pairs have to be examined that represent independent transitions.

In the current study, we test further this hypothesis by examining the molecular basis of interspecific variation in social organization, specifically mating strategy, in four closely related cichlids of Ectodini. We aim to determine whether the independent evolutionary transitions from polygyny to monogamy have co-opted similar sets of genes, or whether the mechanisms that underlie seemingly similar mating strategies are wholly independent. We hypothesize that while some gene sets will be associated with lineage (phylogeny) and (or) gonadal sex, at least a subset of genes associated with mating system will be shared across the two contrasts. Such a correspondence in mating system specific gene sets across these parallel transitions from polygyny to monogamy would suggest broader patterns of deep homology.

Methods

Choice of species

We chose four closely related species from the monophyletic Ectodini clade differing in mating and parental care strategy (Fig. 1). From the primarily sand-dwelling genus *Xenotilapia* we selected *X. ochrogenys*, a polygynous species in which only the females provide maternal care, and the closely related *X. flavipinnis*,

a monogamous species with the male and female forming a pair bond and providing parental care. Similarly, the rock-dwelling *Asprotilapia leptura* is monogamous and biparental, but the closely related *Microdontochromis tenuidentata* (which lives in intermediate habitats) is a polygynous species in which only the females provide maternal care (Kidd et al. 2012; Pollen et al. 2007). In the following, we refer to *X. ochrogenys* and *X. flavipinnis* as the X-lineage and to *A. leptura* and *M. tenuidentata* as the nonX-lineage. Together, these two contrasts represent two independent transitions from the ancestral polygynous to a monogamous mating system (Kidd et al. 2012; Koblmuller et al. 2004).

Sample collection

Specimens were obtained by SCUBA (*X. flavipinnis*, *A. leptura*, *M. tenuidentata*) or netted by beach seine (*X. ochrogenys*) at field sites surrounding Kigoma (Tanzania) in July 2004 or Mpulungu (Zambia) in April 2005. Note that all species reproduce year-round, although during the wet season, adverse weather events are more likely and can disrupt breeding. The animals were collected at the following locations: *X. flavipinnis*: Katonga Beach (4°55′22.5″S, 29°40′20.5″E); *A. leptura*: Hilltop South Cliffs (4°53′24.50″S, 29°36′45.00″E); *X. ochrogenys*: TAFIRI beach (4°53′10.11″S, 29°37′12.69″E); *M. tenuidentata*: near Mpulungu (8°44′47.5″S, 31°07′30.5″E). The animals used in the present study ranged in standard length as follows. *A. leptura*: 9.3–10.0 cm (males) and 7.0–9.4 cm (females); *M. tenuidentata*: 6.63–7.4 cm (males) and 6.43–7.08 cm (females); *X. flavipinnis*: 5.8–6.5 cm (males) and 5.9–7.4 cm (females); and *X. ochrogenys*: 7.6–8.5 cm (males) and 6.5–7.4 cm (females). For each species, we collected sexually mature adult individuals (five of either sex). The males of all species possessed large gonads with distinct and mature sperm packages. The females of both monogamous species were gravid and most were still holding fry in their mouths after capture, as were several of the monogamous males. Four of the five *X. ochrogenys* females were also mouth-brooding. All *M. tenuidentata* females had large ovaries with eggs at various stages of maturation; however, by the time of capture none were holding fry anymore. Animals were euthanized as quickly as possible after capture and brains were rapidly removed and stored in RNAlater solution (Ambion) within 3 min of death. Total RNA was extracted from brains using the TRIzol protocol (Invitrogen) following homogenization of brain tissue. RNA quality and concentration was determined using the Bioanalyzer (Agilent).

Microarray platform

We used a second generation 19K *A. burtoni* microarray that contains features from a brain-specific library and a mixed-tissue library representing 1248 tentative contigs (TCs), 6572 singletons (sequences not belonging to a TC), and 6840 features without sequence information, for a total of 14 660 potential genes (Renn et al. 2004; Salzburger et al. 2008) (GEO platform GPL6416). After hybridization (see below for details) arrays were scanned (Axon 4000B: Axon Instruments) using Genepix 5.0 software (Axon Instruments).

Heterologous array-based comparative genomic hybridization

When heterologous hybridization is used to assess expression difference between two or more species, hybridization of genomic DNA from the two species onto the microarray can help identify genes with sequence divergence or other genomic characteristics that differ between the two species (Buckley 2007; Kassahn et al. 2007; Machado et al. 2009; Renn et al. 2010). Six heterologous array-based comparative genomic hybridizations (aCGH) incorporating dye-swaps and all pairwise comparisons were used to identify array features with genomic bias among the four test species (supplementary data, Fig. S1A[2]). We extracted genomic DNA from ethanol-stored fin clips of 5–10 individuals per species using a standard Proteinase K/Phenol protocol, pooled by species, and sheared to roughly 1.5 Kb (Hydroshear, Genome Solutions/Digilab). Each pooled DNA sample (3 μg) was fluorescently labeled with Alexa-Fluor conjugated dCTP by Klenow reaction (Invitrogen, BioPrime), and quantified (Nanodrop 3300) to match Cy3 and Cy5 samples for between species competitive genomic hybridizations that proceeded for ~16 h at 48 °C in Ambion Hyb Buffer 1 (Ambion) blocked by Cot-1DNA (Invitrogen) using the 19K *A. burtoni* cDNA microarray platform. Scanning, background correction, and normalization were conducted as described below for expression arrays.

Genomic hybridization masking procedure

Sequence divergence between the four species of Ectodini and the array platform species (*A. burtoni*) will influence the observed gene expression differences and thus confound our results (Cummings et al. 2008; Machado et al. 2009; Renn et al. 2010; Schumer et al. 2011). To determine which sequences have diverged enough to potentially bias hybridization efficacy, we examined the correlation between aCGH ratios and expression ratios (see below for hybridization protocol) for each pairwise species comparison. Both of these measures of relative hybridization to the array were calculated independently using the "lmFit" function in LIMMA and extracting each pairwise contrast between-species using a modified "contrasts.fit" function (Machado et al. 2009), which calculates the correct (rather than the approximate) standard errors even in the presence of missing data due to quality filters. For these masking steps, we treated features of similar sequence (belonging to the same tentative contigs, TCs) as independent.

For each pairwise comparison, we used an iterative, stepwise process (scripts available on GitHub: https://github.com/zrenn/Renn_etal_2017_Ectodini_expression-profiling). At each step the two additional array features that showed the greatest magnitude aCGH ratio were added to the mask (i.e., eliminated from further analysis). This process proceeded until a sufficient number of microarray features had been masked such that for an appropriately sized "test group" the correlation between aCGH ratio and gene expression ratio was no longer significant ($P > 0.05$). The test group size was set for each pairwise comparison as the number of microarray features (greater than 50 to avoid spurious correlation) with the greatest magnitude of aCGH ratio that resulted in the largest positive and statistically significant ($P < 0.05$) correlation between aCGH ratio and expression ratio, when only features with genomic bias in the same direction of expression bias were masked. This method produced six independent genomic masks, one for each pairwise species comparison, resulting in 399 features masked for *A. leptura* versus *M. tenuidentata*, 194 for *A. leptura* versus *X. flavipinnis*, 384 for *A. leptura* versus *X. ochrogenys*, 547 for *M. tenuidentata* versus *X. flavipinnis*, 460 for *M. tenuidentata* versus *X. ochrogenys*, and 36 for *X. flavipinnis* versus *X. ochrogenys*. The union of these six genomic masks, a total of 1319 array features, was masked out for all analyses of differential gene expression presented here.

Expression analysis using heterologous array hybridization

We used the nested loop design with dye-swap (Churchill 2002) to emphasize within-lineage comparison between species of different mating strategies, analyzing each sex independently (Fig. S1B[2]). As two-color microarray technology provides relative measures of gene expression for those samples that are compared directly or indirectly through competitive hybridization, the

[2]Supplementary data are available with the article through the journal Web site at http://nrcresearchpress.com/doi/suppl/10.1139/gen-2017-0075.

nested male loop and nested female loop produce independent gene lists that can then be compared. The nested female loops included nine X-lineage and nine nonX-lineage array comparisons between monogamous and polygynous species. The nested male loops included 10 X-lineage and nine nonX-lineage comparisons. For both sexes, eight between-lineage hybridizations allowed direct comparison between the X- and nonX-lineages, both within- and between-mating strategy. For each individual sample, 2 μg of total RNA from whole brain was reverse transcribed and labeled according to Renn et al. (2004). Briefly, amino-allyl dUTP (Sigma) was incorporated to the cDNA using oligo-dT(12-18) with SuperScript II (Invitrogen) according to the manufacturer's protocol, and dye-coupling with Cy3 or Cy5 (CyDye Post-Labeling Reactive Dye Pack, Amersham) was followed by RNA hydrolyzation and purification. The neutralized color reaction was purified and combined with the appropriate competitive sample in hybridization buffer containing SSC and HEPES buffer with poly(dA)-poly(dT) (Sigma) for blocking for overnight hybridization at 65 °C. Whole brain samples were used to allow for the use of wild-caught samples.

Analysis of differential expression

In addition to masking out features determined to exhibit hybridization bias due to sequence divergence (see above), raw expression data were filtered such that all faint (average intensity <2 standard deviations above background) and (or) small (diameter <60 pixels) features were removed prior to background correction and normalization as described above. Here, to test for differential regulation at the level of the gene, the intensities of features of similar sequence (i.e., members of tentative contigs, TCs) were averaged to a single value prior to fitting a linear model to the data (lmFit) and estimating expression differences (contrasts.fit) (Machado et al. 2009) and statistical significance (eBayes) (Smyth et al. 2005). Eight contrasts of interest were assessed to identify gene expression patterns associated with mating system (1 contrast), lineage (1 contrast), and species (6 pair-wise contrasts). Determination of significant differential gene expression bias was based on a threshold of $P < 0.005$.

Functional annotation and gene ontology analysis

Microarray features for which sequence information was available were annotated according to the available cichlid genome sequences (Brawand et al. 2014) by BLAST comparison to the available five species' tissue-specific transcriptomes. For features with no transcript alignments with bit score greater than 200 and for features with available sequence length greater than that of the available transcript, the initial sequence was retained. We then compared this set of sequences to the set of predicted genes for the five cichlid species. The best hit with a bit score greater than 200 was used to annotate the microarray feature by transferring available gene names.

Gene Ontology (GO) terms were used to give functional annotation to the array features and to test for significant over-representation of functional groups among the genes showing mating-system-biased and lineage-biased expression. We used the Generic GO slims (Mundodi and Ireland 2002), which were previously developed to provide a useful summary of GO annotation for comparison of genomes and transcriptomes when a broad overview of the ontology content is required. Of the 7820 unique *A. burtoni* sequences available for array features, 4824 have been annotated to GO Slim terms. Note that any single assembled sequence may be annotated in all three ontologies and according to multiple ontology terms. A total of 20 138 annotations have been applied: 8175 biological process annotations (including 66 GO Slim terms applied to 3591 array sequences), 7871 molecular function annotations (including 41 GO Slim terms applied to 4432 array sequences), and 4992 cellular component annotations (including 30 GO Slim terms applied to 3253 sequences).

Over-representation of biological process and molecular function GO terms for a regulated set of genes was determined in Cytoscape (Shannon et al. 2003) using the Biological Network Gene Ontology tool, BiNGO (Maere et al. 2005), which relies upon statistical significance of a hypergeometric distribution. Owing to the small number of genes for each ontology term and the relatively small number of genes that are regulated, there is little statistical power to identify significantly under-represented GO terms. Also, due to the highly non-independent nature of GO categories we use GO analysis as a hypothesis-generating tool and report only uncorrected hypergeometric P-values for over-representation.

Results and discussion

Our parallel nested loop designs consisted of 26 female and 28 male within-sex comparisons for analysis of gene expression patterns associated with lineage, species, and mating system using five males and females each from each species. The 1319 features that showed significant hybridization bias due to sequence variation were masked (i.e., removed) from the analysis (see methods), and quality filters were applied to all individual array features as described above prior to estimating expression coefficients. This resulted in 9888 genes available for the analysis of female gene expression, and only 5696 genes were available for the analysis of males (which contained a few arrays of poorer quality). Here, we report results for the intersection of 5232 genes that were available for analysis in both sexes (GEO: GSE 97082). We first confirm the effectiveness of this masking procedure before examining gene expression variation. We then report gene expression variation associated with lineage in males and in females. Next, we identify genes whose expression varies with mating system regardless of lineage and examine concordance between the sexes. Not surprisingly, we also identify genes that are regulated in a species-specific manner both within and across X- and nonX-lineages in both sexes. Finally, we use the Gene Ontology framework to provide further functional insights.

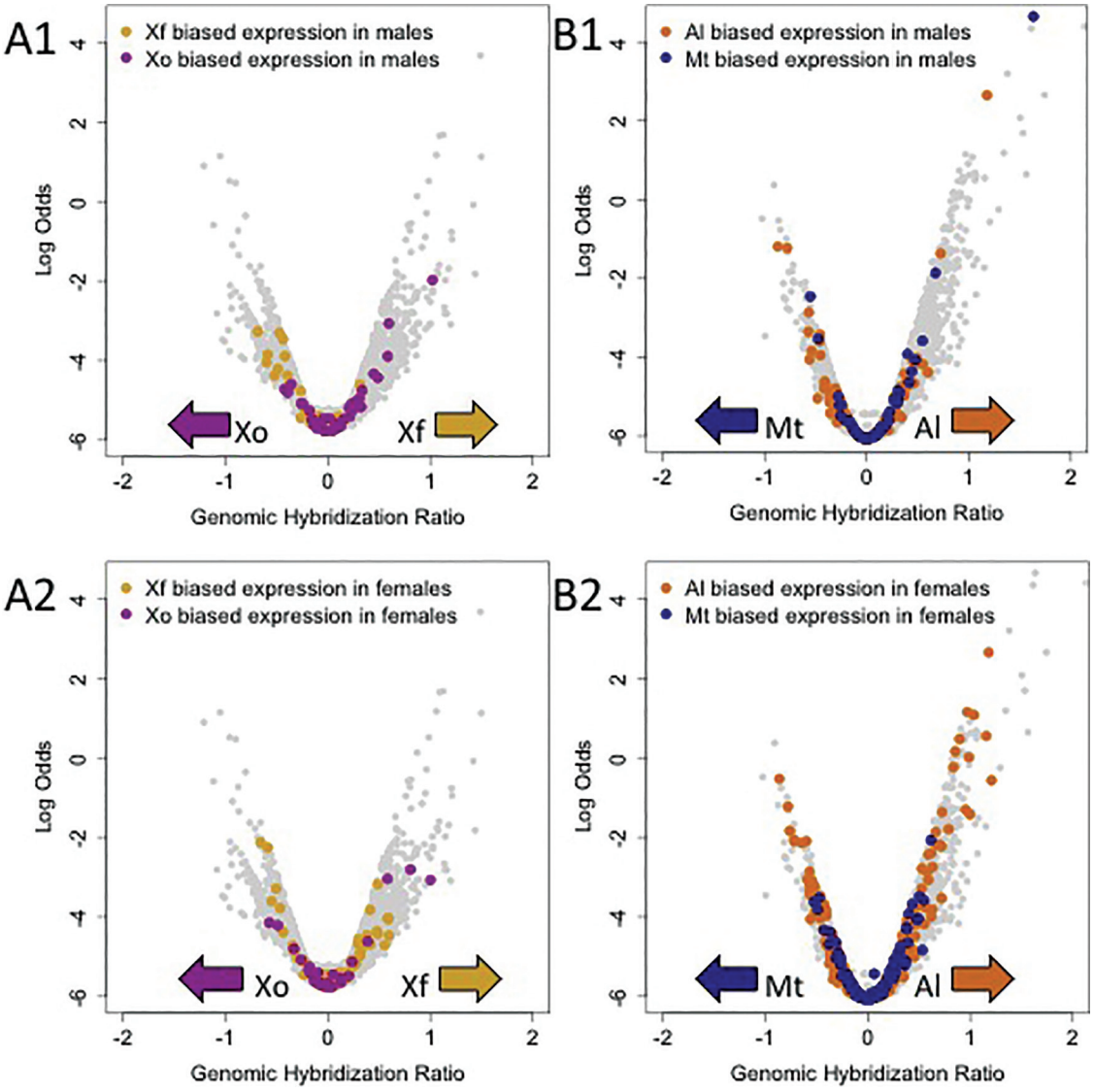
Ascertaining array features without sequence bias

To demonstrate the effectiveness of the masking procedure and confirm that these results represent true differences in gene expression levels, we asked whether genes identified as differentially expressed showed any (even non-significant) genomic hybridization bias. Importantly, for the 1284 genes that we found to be differentially expressed in any pair-wise species comparison, gene expression bias and genomic hybridization bias (aCGH ratio) were not correlated (Table S1[2]). Graphically, the array features that showed statistically significant expression variation in pair-wise species comparisons are randomly distributed with respect to aCGH ratios (Fig. 2). This was true for both lineages and both sexes. As previously demonstrated by Machado et al. (2009), these results confirm the suitability of heterologous array hybridization for the comparative analysis of gene expression patterns across species.

Gene expression varies according to phylogeny

To identify gene expression regulation related to lineage we directly compared gene expression patterns of the two closely related species from the X-lineage to those from the nonX-lineage, for each sex separately. In the analysis of males, we found 92 genes with increased expression in the X-lineage and 69 genes with increased expression in the nonX-lineage ($P < 0.005$). In the analysis of females, 274 genes were found to have increased expression in the X-lineage and 157 genes were found to have increased expression in the nonX-lineage ($P < 0.005$) (Table S2[2]). Our experimental design with a nested loop design for each sex did not allow calculation of direct contrasts between the sexes. Instead, we compared the two lineage-bias gene lists (derived separately for males and for females) to identify the set of genes that exhibit concordant

Fig. 2. Effectiveness of masking procedure in reducing error in gene expression analyses. Volcano plots showing statistical significance in expression variation (log odds) against log 2 genomic DNA hybridization ratios comparing (A) the X-lineage species *X. flavipinnis* (Xf) with *X. ochrogenys* (Xo) males (A1) and females (A2), and (B) the nonX-lineage species *A. leptura* (Al) with *M. tenuidentata* (Mt) males (B1) and females (B2). Colored arrows indicate direction of species bias for genomic hybridization ratios. Highlighted features indicate genes that showed statistically significant gene expression bias in the pair-wise species comparisons. The lack of color segregation for highlighted features demonstrates that the genomic masking procedure is effective in preventing bias by species-specific genomic DNA sequence characteristics in heterologous array hybridization experiments. [Colour online.]

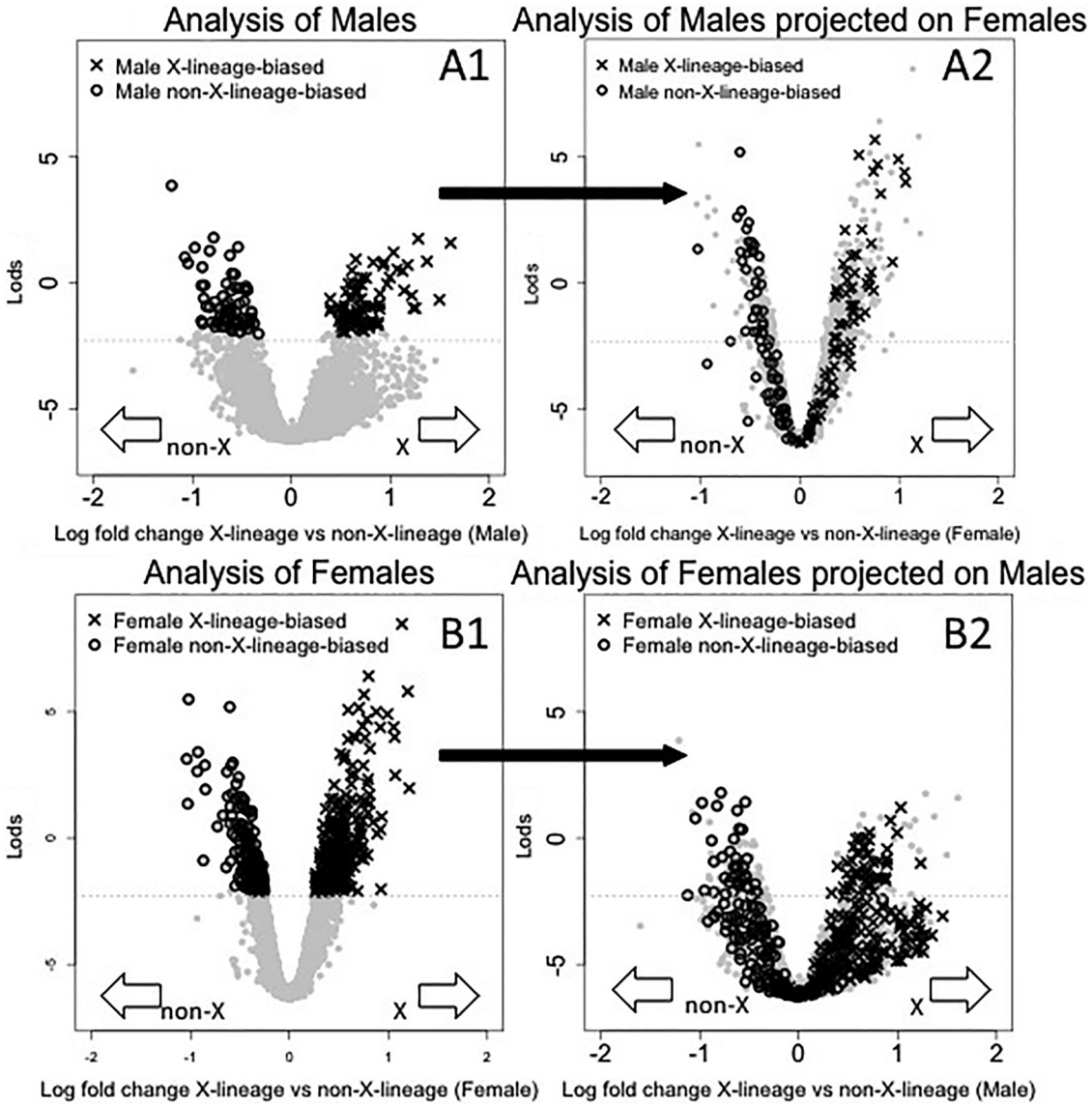


expression in the X-lineage compared to the nonX-lineage, independent of sex. We found 46 genes that were X-lineage biased and 15 genes that were nonX-lineage biased in both analyses. We found no genes that showed discordant lineage-biased expression between the male and female analyses.

We can visualize the role played by phylogenetic relatedness in regulating gene activity, independent of sex or mating system, by projecting the significant results from one sex on a volcano plot of the results from the analysis in the other sex (Fig. 3). Specifically, the majority of the genes whose mRNA levels we identified as lineage biased in males (Fig. 3A1) showed the same (though not always statistically significant) trend when projected onto the female analysis (Fig. 3A2). In the reverse projection, lineage-biased genes identified in the female analysis (Fig. 3B1) followed the same trend when projected onto the male analysis (Fig. 3B2). This high level of concordance demonstrates the strong influence of lineage.

Our study is one of only a handful that have examined lineage-specific gene expression on a genomic scale using more than one species pair or multiple subspecies from the same lineage. For example, Oldham et al. (2006) compared the brain transcriptomes of humans and chimpanzees and showed that gene expression profiles vary greatly across discrete brain regions, with those in subcortical brain regions more conserved across species than in the cerebral cortex. However, the inference was weakend by the limited number of species and lineages in the analysis. Other groups used comparative transcriptomics to determine whether a molecular correlate of a phylotypic stage can be identified during animal development (six species of *Drosophila* spec., Kalinka et al. 2010; four vertebrate species, Irie and Kuratani 2011). Maybe the most comprehensive study to date was carried out by Brawand et al. (2011), who examined the dynamics of mammalian transcriptome evolution by comparing six organs across 10 species of mammals (with a single representative from each lineage, plus an

Fig. 3. Graphical representation of gene expression variation according to lineage. Volcano plots showing expression variation differences between the X- and nonX-lineages (log 2 fold differences) plotted in grey against statistical significance (log odds) for male (A1 and B2) and female (B1 and A2) analyses. Genes whose expression pattern was identified as significantly lineage biased (highlighted in black) in the male analysis (A1) showed the same lineage bias (though not always significant) projected onto the female analysis (A2). Similarly, genes whose expression pattern was significantly lineage biased (highlighted in black) in the female analysis (B1) showed the same lineage bias (though not always significant) when projected onto the male analysis (B2).

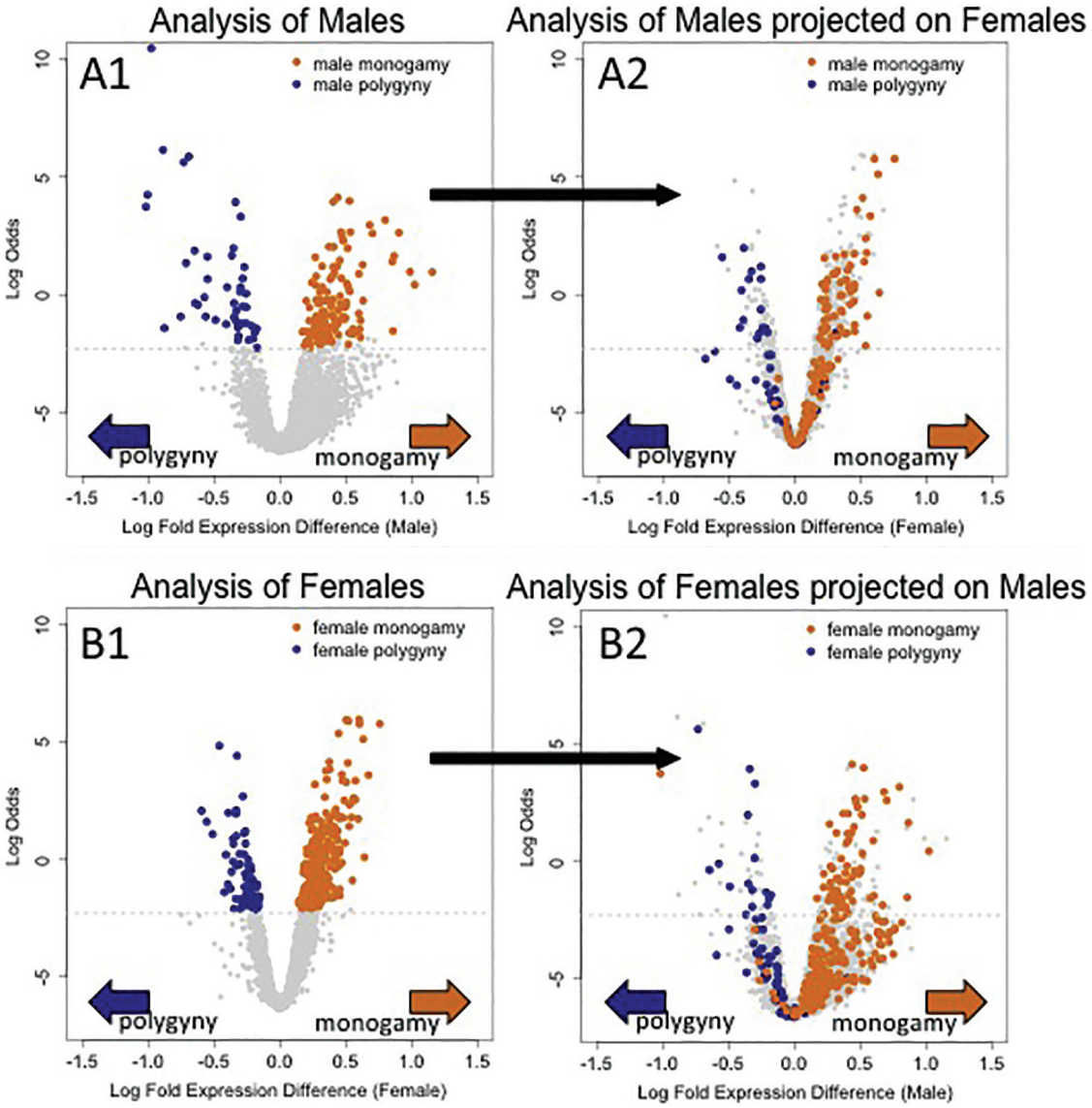


outgroup), demonstrating that the rate of gene expression evolution varies considerably among organs, lineages, and chromosomes. Together with the present study, these examples support the notion that comparative transcriptomics can provide novel and fundamental insights into biological problems in a way that is complementary to experimental studies.

Gene expression profiles are associated with mating system

Our experimental design (Fig. S1B[2]) emphasized direct comparisons of gene expression profiles of the two polygynous species (*X. ochrogenys* and *M. tenuidentata*) with the two monogamous species (*X. flavipinnis* and *A. leptura*). We chose this design as it provides considerably more power to discover genes whose expression patterns are associated with mating system compared with the species-centered analyses presented above. In the female analysis, 331 (~6%) of the genes analyzed here showed expression variation according to mating system, including 260 that were up-regulated in the monogamous species and 71 that were up-regulated in the polygynous species ($P < 0.005$). Similarly, in the male analysis a total of 247 genes showed mating-system-associated expression variation, including 131 up-regulated in the monogamous species and 46 up-regulated in the polygynous species (Table S3[2]). Importantly, not only were both the female and the male results biased (~4:1) toward genes more highly expressed in the monogamous species, but these gene lists were also highly concordant. Specifically, 81 genes were concordantly regulated according to mating system; the vast majority of those (66 vs. 15) were up-regulated in monogamous individuals (Fig. 4). Of note, only one gene showed discordant gene expression bias in this analysis. This strong pattern is also evident when we investigate trends regardless of statistical significance. Most of the genes identified as differentially expressed between monogamous and polygynous males (Fig. 4A1) showed the same direction of expression bias when projected onto the female analysis (Fig. 4A2) and vice versa (Figs. 4B1 and 4B2).

Fig. 4. Graphical representation of gene expression variation according to mating system. Volcano plots showing log 2 fold differences in gene expression according to mating system plotted against statistical significance (log odds). Genes identified as differentially expressed between monogamous (orange) and polygynous (purple) individuals in the analysis of males (A1) showed the same (though not always significant) trend when projected onto the analysis of females (A2). Conversely, mating system-associated genes identified in the analysis of females (B1) follow the same trend when projected onto the analysis of males (B2). [Colour online.]



The concordance between the evolution of neural gene expression in males and females during the transition from polygyny to monogamy is remarkable in light of the behavioral roles of the sexes, which have evolved in somewhat opposing directions in the monogamous species. Specifically, the transition to monogamy resulted in increased parental care by males and decreased or shared parental care by females. Our results suggest that changes in gene expression have been convergent or parallel, while changes at the behavioral level have been divergent (males evolved to increase care behavior while females reduced care). It should be noted, however, that gene expression studies aim at identifying differences. Statistical tests of equivalence are hampered by large data sets, the need to set arbitrary thresholds, and noisy data (but see Qiu and Cui 2010). Alternatively, the observed concordant evolution of gene expression may underlie parallel evolution of other behavior patterns such as pair bonding, which is common in monogamous but not in polygynous mating systems.

With regard to social evolution, research on hymenoptera has been important to our understanding of evolution of genome regulation on multiple timescales associated with social phenotypes. On fairly rapid timescales, the same genes appear to be associated with plasticity and adaptation for aggression levels (Alaux et al. 2009). Over longer timescales, genome regulation associated with provisioning and foraging behaviors appears to be largely conserved across bees (Sen Sarma et al. 2007) and even between wasps and bees, which are separated by ~100 million years of evolution (Toth et al. 2010). Similar attempts to address the evolution of gene expression associated with complex traits in animals have focused on physiological traits such as cold tolerance (Makinen et al. 2016), hypoxia tolerance (Kozak et al. 2014), salinity tolerance (Latta et al. 2012), predator avoidance (Fischer et al. 2014; Ghalambor et al. 2015), generation of a weakly electromagnetic field (Gallant et al. 2014), sulfur tolerance (Kelley et al. 2016), and eco-toxin resistance (Whitehead et al. 2012). One recent

study found similar gene expression modules in brain regions involved in vocal learning in songbirds, parakeet, hummingbirds, and humans (Pfenning et al. 2014). Similar to our identification of common gene expression patterns associated with mating strategy, the emerging picture from these studies suggests deep molecular homologies underlying the independent evolution of monogamy in different cichlid lineages of Ectodini.

Are nonapeptide systems associated with mating system variation?

Nonapeptide systems play important, albeit complex, roles in the regulation of social behavior across vertebrates (and beyond). The nonapeptide arginine vasopressin (AVP), and its non-mammalian homolog arginine vasotocin (AVT), have been examined especially in relation to pair bonding and social affiliation (e.g., prairie vole, *Microtus ochrogaster* (Winslow et al. 1993; Young et al. 2008); zebrafinch, *Taeniopygia guttata* (Goodson and Adkins-Regan 1999); convict cichlid, *Amatitlania nigrofasciata* (O'Connell et al. 2012; Oldfield et al. 2013; Oldfield and Hofmann 2011)). Recently, Oldfield et al. (2015) integrated the known diversity in AVP/AVT function in affiliative behavior with resource defence theory to explain variation in territory-based mating systems across vertebrates. By carefully reviewing the literature, these authors found that expression of AVP/AVT (and its V1a receptor) in one particular neural circuit involving the lateral septum of the forebrain is associated with territorial behavior in males of diverse species, likely due to effects of this system on social cognition. We therefore examined whether the expression levels of AVT was associated with mating system in either male or female Ectodini. (Note that probes for the nonapeptide oxytocin/mesotocin/isotocin, as well as any of the nonapeptide receptors, did not pass quality thresholds for inclusion in final analysis.) We found that expression did not vary significantly according to mating system (nor sex, nor species). It is, of course, possible that profiling gene expression in whole brains simply masked any variation in the AVT cell population that might be relevant to mating system. There are at least three functionally distinct AVT neuron groups in the teleost preoptic area alone (e.g., Greenwood et al. 2008), and likely also in the telencephalon (Rodriguez-Santiago et al. 2017), and these areas play distinct and sometimes even opposing roles in the regulation of social behavior (Greenwood et al. 2008). Alternatively, it may well be that, at least at this level of analysis, AVT is not involved in establishing the difference between monogamous and polygynous cichlids of Ectodini. Our finding that sequence variation in the AVT promoter across the entire Ectodini clade is not correlated with mating system (N. Duftner and H.A. Hofmann, unpublished results) supports this interpretation. Similarly, the number of preoptic AVT immunoreactive neurons (as a measure of AVT system activity) appears to be independent of mating system when examined in four species of Ectodini (including three of the four analyzed in the present study): *A. leptura*, *X. flavipinnis* (both monogamous); and *X. ochrogenys*, *Enantiopus melanogenys* (both polygynous) (C.A. Shumway and H.A. Hofmann, unpublished results). Clearly, nonapeptide systems (along with numerous other neuroendocrine and neuromodulatory pathways) are deeply homologous in that they are critically involved in the regulation of social behavior across animals (for review see Weitekamp and Hofmann 2016). It may, however, be naïve to expect their regulatory logic to be faithfully replicated in independent transitions to monogamy (or other complex social traits). Rather, a systems-level view facilitated by comparative transcriptomics may be much better suited to discover deeply homologous molecular substrates underlying convergent or parallel evolution of behavioral phenotypes and social systems.

Gene expression profiles associated with specific species

To identify gene expression differences that are likely to be involved in species-specific phenotypes (e.g., habitat, food preference, interspecific aggression, etc.), we next identified species-specific gene expression biases. Specifically, we identified genes that showed either significantly higher or significantly lower expression ($P < 0.005$) in one species contrasted against all three other species (see Table 1 for gene numbers and Table S4[2] for gene annotations). Interestingly, expression patterns of the polygynous species *M. tenuidentata* appeared to be the most divergent compared to the other three species, especially in females. Similarly, *X. ochrogenys* females, the other polygynous species, showed 10 times more genes with species-specific biased expression compared to the males of this species. We also found significant concordance (i.e., species specific in both sexes: Fisher's exact test $P < 0.05$) of gene sets (Table 1). This high level of concordance strongly demonstrates that species, regardless of sex, is a strong determinant for gene expression profile. Even though we cannot entirely rule out the possibility that, despite the masking procedure, species differences in genomic sequence might have affected these results, any such effects would likely be minor, such that the species-specific expression observed here primarily relates to genes sufficiently conserved to be detected as orthologs. Of course, taxon-specific genes, which comprise 10%–20% of protein-coding genes in many sequenced genomes, likely also underlie lineage-specific traits (Dai et al. 2008; Johnson and Tsutsui 2011; Khalturin et al. 2009), such as mating and parental care behavior.

Table 1. Species-specific gene expression variation in males and females of the four study species.

	Male (up/down)	Female (up/down)	Concordant (up/down)
Xenotilapia flavipinnis	230 (133/97)	218 (156/62)	50 (35/15)
Xenotilapia ochrogenys	41 (22/19)	339 (166/173)	17 (7/10)
Asprotilapia leptura	69 (38/31)	109 (67/42)	14 (11/3)
Microdontochromis tenuidentata	254 (88/166)	892 (327/565)	193 (66/127)

Note: Shown are numbers of genes significantly ($P < 0.005$) different in expression in any given species compared to the three other species for males and females as well as those that are shared (overlap) between the sexes. Numbers in parentheses indicate the number of genes significantly up- or down-regulated in the named species.

Table 2. Lineage-specific gene expression variation in males and females of the four study species.

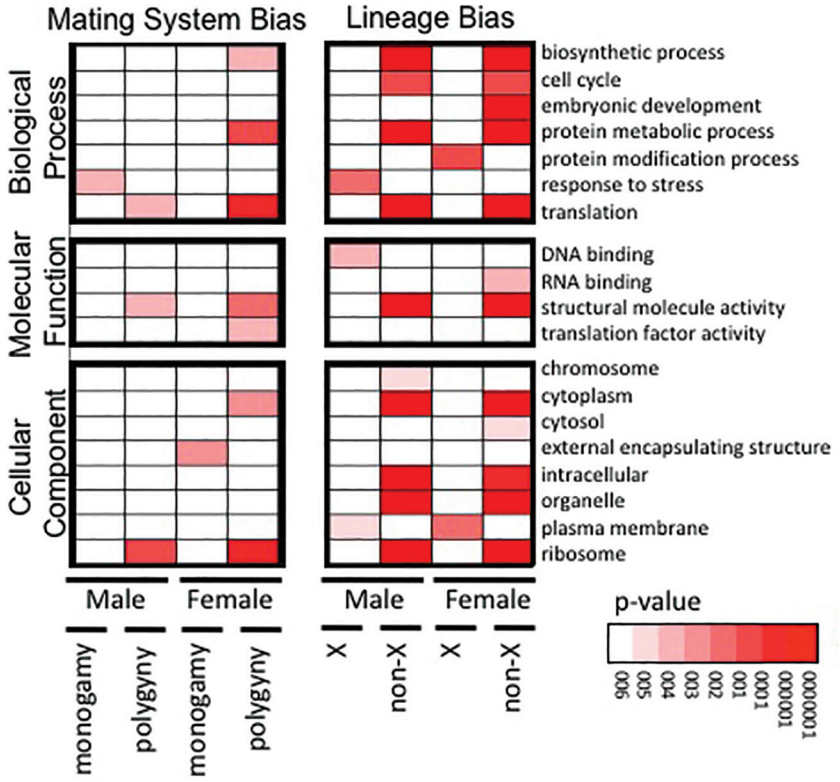
	Male (monogamy/polygyny)	Female (monogamy/polygyny)	Concordant	Discordant
X-lineage	91 (47/44)	103 (70/33)	8 (1/7)	8 (6/2)
NonX-lineage	207 (146/61)	593 (412/181)	109 (81/28)	0 (0/0)
Concordant	13 (7/6)	9 (7/2)	0 (0/0)	—
Discordant	10 (8/2)	15 (7/8)	—	—

Note: For each lineage the numbers of genes are shown that are significantly ($P < 0.005$) different in expression according to mating strategy as well as those that are shared (overlap) between the sexes. Numbers in parentheses indicate the number of genes significantly up-regulated in monogamous or polygynous species.

Species-specific gene expression profiles associated with mating system within and across lineages

We next conducted pairwise species comparisons to identify gene expression regulation associated with mating system that may be specific to a given species pair. First, we interrogated our dataset within lineage (X and nonX), comparing the expression profiles from each monogamous species to its closely related polygynous species (Table 2; gene annotations are provided in Table S5[2]). Notably for the analyses of both males and females, many more genes showed mating-system-dependent expression within the nonX-lineage than within the X-lineage, likely due to the very divergent *M. tenuidentata* expression profiles (see above). With regard to the direction of gene expression associated with

Fig. 5. Functional analysis using Gene Ontology. The generic GO-Slim was used to analyze gene function. Red shading indicates degree of statistical significance for over-represented categories. Gene expression biases were more prevalent across clades (X- vs. nonX-lineage). Translation, structural molecule activity, and ribosome are highly significant for mating system in polygynous individuals, but they were also significant in the nonX-lineage. Cell cycle, cytoplasm, and organelle categories are significant in nonX-individuals only. [Colour online.]



mating system, there were many more genes up-regulated in the monogamous species than were up-regulated in polygynous species (Table 2). This was true for both sexes in the nonX-lineage and for the females in the X-lineage.

Though we have less statistical power in this analysis, the genes identified reflect species-specific expression correlated with mating system. To focus on genes most likely related to mating system, we looked for concordance between the pairwise comparisons within sex and within lineage. We found that the mating system biases in gene expression in the X-lineage were highly concordant between males and females (109 concordant genes, 0 discordant genes), which was not the case in the nonX-lineage (8 concordant genes, 8 discordant genes) (Fig. S2[2]; Table 2). Next, we looked for concordance between the pairwise comparisons within sex but across lineage for gene expression regulation associated with mating system. Interestingly, very few genes appeared to be concordantly regulated at a statistically significant level, and similar numbers were discordantly regulated (Fig. S3[2]; Table 2). In summary, we found strong bias toward increased gene expression in monogamous individuals and this bias is concordant between sexes within the X-lineage but not concordant between sexes in the nonX-lineage nor between lineages regardless of sex. While we found no genes associated with mating system that showed concordant expression variation across both sexes and both lineages, there were genes that showed concordant regulation according to mating system in (*i*) both sexes for one lineage or, more importantly, (*ii*) in one sex or the other in both lineages. This suggests that the changes in gene expression regulation that have evolved in males and in females are not entirely parallel or may be obscured by evolution of other phenotypic traits.

When we analyzed the data at the species level, at reduced statistical power, we found a lack of overall concordance when considering both male and female comparisons between X- and nonX-lineages. This suggests that the most extreme and robust differences in brain gene expression between species (i.e., those that can be detected with smaller sample sizes) do not reflect any shared molecular mechanisms underlying similar mating system phenotypes. Instead, much of the species-specific differences in gene expression may be associated with species-specific phenotypes, or they may represent alternate mechanisms underlying similar phenotypes. Whatever the case, this result underscores the importance of including individuals from different lineages in the analysis, in addition to multiple comparisons of monogamy and polygyny. Because we see a strong phylogenetic signal, at the level of both lineage and species, the inclusion of two lineages and two independent instances of a transition to monogamy is critical for detecting gene expression variation associated with mating strategy.

Functional analysis using gene ontology

To infer functional information about the genes that show biased expression for mating strategy or lineage we applied generic GO-Slim annotations to the array features (Fig. 5; Table S6[2]). Among genes up-regulated in both males and females of polygynous species, the GO categories translation (biological process), structural molecule activity (molecular function), and ribosome (cellular component) were over-represented ($P < 0.05$). Identification of these GO categories was driven largely by numerous genes annotated to ribosomal proteins. Conversely, the GO categories that were significantly associated with monogamy in the analysis of the males (response to stress) or the females (external encapsulating structure) were unique to each sex. Interestingly, the GO categories translation (biological process), structural molecule activity (molecular function), and ribosome (cellular component)

were also significantly associated with nonX-lineage in both males and females, as were biosynthetic process and cell cycle (biological process), structural molecule activity (molecular function), and cytoplasm, intracellular and organelle (cellular component). These GO categories identified for the nonX-lineage were again largely driven by genes annotated to ribosomal proteins, while the GO term categories identified as enriched in the X-lineage (protein modification process and plasma membrane) included genes representing more neural-specific pathways. The biological interpretation of these results remains uncertain, as is often the case with GO analyses in transcriptome studies.

Conclusion

In the present study, we tested the hypothesis that independent evolutionary transitions from polygyny to monogamy in cichlids of Ectodini have co-opted similar brain gene expression profiles, independent of phylogeny and (or) sex. Male and female gene expression profiles were analyzed independently for lineage, species, and mating system. We identified signatures of gene expression that correspond to sex, species, lineage, and mating system. Importantly, we found high concordance between the male and female analyses for genes associated with either lineage or mating system, demonstrating that both phylogenetic relatedness and social phenotype can drive gene expression evolution. Further, among the genes associated with mating system independent of lineage, we find a strong bias (at a ratio of 4:1) toward genes that show increased expression in monogamous individuals, possibly due to the requirement for both males and females to change their affiliative and parental care behavior as monogamy evolved. A highly conserved nonapeptide system known to be critically involved in the regulation of social behavior across animals was not associated with mating system in our analysis. Our findings thus support the hypothesis that, on a genomic scale, independent evolutionary transitions from polygyny to monogamy were accompanied by similar changes in brain gene expression patterns, suggesting deep molecular homologies underlying the independent evolution of monogamy in different cichlid lineages of Ectodini. The extent to which these results hold more broadly across vertebrates is currently unknown, motivating a fascinating area of future research.

Acknowledgements

We thank George and Wilbroad Kazumbe, Sarah Bahan, and Alex Pollen for assistance in the field, and Lin Winton and Celeste Kidd for laboratory assistance. We also thank Rebecca Young Brim for reading earlier versions of this manuscript. We are grateful to the Tanzania Fisheries Research Institute (TAFIRI), the Tanzania Commission on Science and Technology (COSTECH), and Alfeo Nikundiwe (University of Dar es Salaam) for their kind support of our research. Finally, we extend sincere thanks to Andy Cohen and Ellinor Michel from the Nyanza Project, John Fitzpatrick, Saskia Marijnissen, and the Vaitha brothers for providing materials and support to our fieldwork. Sara Natale and Andrew Winterman, among other NSF-REU funded undergraduate students, contributed improvements to the LIMMA code, under advice of Albyn Jones (Reed College). This research was supported by an Erwin-Schrödinger postdoctoral fellowship provided by the FWF Austrian Science Fund (N.D.); NIH National Research Service Award MH070180-02, the Jordan Award of the American Cichlid Association, a Murdock Life Trust Foundation award, and National Science Foundation (NSF) grants IOS 0818957 and DEB 1021582 (S.C.P.R.); and by NSF grants IOS 0217915 and IOS 0843712, the Alfred P. Sloan Foundation, and a Dwight W. and Blanche Faye Reeder Centennial Fellowship in Systematic and Evolutionary Biology (H.A.H.).

References

Alaux, C., Sinha, S., Hasadsri, L., Hunt, G.J., Guzman-Novoa, E., DeGrandi-Hoffman, G., et al. 2009. Honey bee aggression supports a link between gene regulation and behavioral evolution. Proc. Natl. Acad. Sci. U.S.A. **106**(36): 15400–15405. doi:10.1073/pnas.0907043106. PMID:19706434.

Aubin-Horth, N., Desjardins, J.K., Martei, Y.M., Balshine, S., and Hofmann, H.A. 2007. Masculinized dominant females in a cooperatively breeding species. Mol. Ecol. **16**(7): 1349–1358. doi:10.1111/j.1365-294X.2007.03249.x. PMID:17391260.

Aubin-Horth, N., Letcher, B.H., and Hofmann, H.A. 2009. Gene-expression signatures of Atlantic salmon's plastic life cycle. Gen. Comp. Endocrinol. **163**(3): 278–284. doi:10.1016/j.ygcen.2009.04.021. PMID:19401203.

Barlow, G.W. 2000. Cichlid fishes: Nature's grand experiment in evolution. Harper Collins Publishers.

Bendesky, A., Kwon, Y.M., Lassance, J.M., Lewarch, C.L., Yao, S.Q., Peterson, B.K., et al. 2017. The genetic basis of parental care evolution in monogamous mice. Nature, **544**(7651): 434–439. doi:10.1038/nature22074. PMID:28424518.

Brawand, D., Soumillon, M., Necsulea, A., Julien, P., Csárdi, G., Harrigan, P., et al. 2011. The evolution of gene expression levels in mammalian organs. Nature, **478**(7369): 343–348. doi:10.1038/nature10532.

Brawand, D., Wagner, C.E., Li, Y.I., Malinsky, M., Keller, I., Fan, S., et al. 2014. The genomic substrate for adaptive radiation in African cichlid fish. Nature, **513**(7518): 375–381. doi:10.1038/nature13726. PMID:25186727.

Buckley, B.A. 2007. Comparative environmental genomics in non-model species: using heterologous hybridization to DNA-based microarrays. Journal of Exp. Biol. **210**(9): 1602–1606. doi:10.1242/jeb.002402.

Churchill, G.A. 2002. Fundamentals of experimental design for cDNA microarrays. Nat. Genet. **32**(Suppl.): 490–495. PMID:12454643.

Crews, D., and Moore, M.C. 1986. Evolution of mechanisms controlling mating behavior. Science, **231**(4734): 121–125. doi:10.1126/science.3941893. PMID:3941893.

Cummings, M.E., Larkins-Ford, J., Reilly, C.R.L., Wong, R.Y., Ramsey, M., and Hofmann, H.A. 2008. Sexual and social stimuli elicit rapid and contrasting genomic responses. Proc. R. Soc. B Biol. Sci. **275**(1633): 393–402. doi:10.1098/rspb.2007.1454.

Dai, H.Z., Chen, Y., Chen, S.D., Mao, Q.Y., Kennedy, D., Landback, P., et al. 2008. The evolution of courtship behaviors through the origination of a new gene in *Drosophila*. Proc. Natl. Acad. Sci. U.S.A. **105**(21): 7478–7483. doi:10.1073/pnas.0800693105. PMID:18508971.

Dobberfuhl, A.P., Ullmann, J.F.P., and Shumway, C.A. 2005. Visual acuity, environmental complexity, and social organization in African cichlid fishes. Behav. Neurosci. **119**(6): 1648–1655. doi:10.1037/0735-7044.119.6.1648. PMID:16420167.

Duftner, N., Larkins-Ford, J., Legendre, M., and Hofmann, H.A. 2008. Efficacy of RNA amplification is dependent on sequence characteristics: implications for gene expression profiling using a cDNA microarray. Genomics, **91**(1): 108–117. doi:10.1016/j.ygeno.2007.09.004. PMID:18006269.

Emlen, S.T., and Oring, L.W. 1977. Ecology, sexual selection, and the evolution of mating systems. Science, **197**(4300): 215–223. doi:10.1126/science.327542. PMID:327542.

Fischer, E.K., Harris, R.M., Hofmann, H.A., and Hoke, K.L. 2014. Predator exposure alters stress physiology in guppies across timescales. Horm. Behav. **65**(2): 165–172. doi:10.1016/j.yhbeh.2013.12.010. PMID:24370688.

Gallant, J.R., Traeger, L.L., Volkening, J.D., Moffett, H., Chen, P.H., Novina, C.D., et al. 2014. Genomic basis for the convergent evolution of electric organs. Science, **344**(6191): 1522–1525. doi:10.1126/science.1254432. PMID:24970089.

Ghalambor, C.K., Hoke, K.L., Ruell, E.W., Fischer, E.K., Reznick, D.N., and Hughes, K.A. 2015. Non-adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. Nature, **525**(7569): 372–375. doi:10.1038/nature15256. PMID:26331546.

Goodson, J.L., and Adkins-Regan, E. 1999. Effect of intraseptal vasotocin and vasoactive intestinal polypeptide infusions on courtship song and aggression in the male zebra finch (*Taeniopygia guttata*). J. Neuroendocrinol. **11**(1): 19–25. PMID:9918225.

Grant, J.W.A. 1993. Whether or not to defend? The influence of resource distribution. Mar. Behav. Physiol. **23**: 137–153. doi:10.1080/10236249309378862.

Greenwood, A.K., Wark, A.R., Fernald, R.D., and Hofmann, H.A. 2008. Expression of arginine vasotocin in distinct preoptic regions is associated with dominant and subordinate behaviour in an African cichlid fish. Proc. R. Soc. B Biol. Sci. **275**(1649): 2393–2402. doi:10.1098/rspb.2008.0622.

Hofmann, H.A. 2003. Functional genomics of neural and behavioral plasticity. J. Neurobiol. **54**(1): 272–282. doi:10.1002/neu.10172. PMID:12486709.

Irie, N., and Kuratani, S. 2011. Comparative transcriptome analysis reveals vertebrate phylotypic period during organogenesis. Nat. Comm. **2**: 248. doi:10.1038/ncomms1248.

Johnson, B.R., and Tsutsui, N.D. 2011. Taxonomically restricted genes are associated with the evolution of sociality in the honey bee. BMC Genomics, **12**: 164. doi:10.1186/1471-2164-12-164. PMID:21447185.

Kalinka, A.T., Varga, K.M., Gerrard, D.T., Preibisch, S., Corcoran, D.L., Jarrells, J., et al. 2010. Gene expression divergence recapitulates the developmental hourglass model. Nature, **468**(7325): 811–814. doi:10.1038/nature09634. PMID:21150996.

Kassahn, K.S., Caley, M.J., Ward, A.C., Connolly, A.R., Stone, G., and Crozier, R.H. 2007. Heterologous microarray experiments used to identify the early gene response to heat stress in a coral reef fish. Mol. Ecol. **16**(8): 1749–1763. doi:10.1111/j.1365-294X.2006.03178.x. PMID:17402988.

Kelley, J.L., Arias-Rodriguez, L., Martin, D.P., Yee, M.C., Bustamante, C.D., and Tobler, M. 2016. Mechanisms underlying adaptation to life in hydrogen sulfide-rich environments. Mol. Biol. Evol. **33**(6): 1419–1434. doi:10.1093/molbev/msw020. PMID:26861137.

Khalturin, K., Hemmrich, G., Fraune, S., Augustin, R., and Bosch, T.C.G. 2009. More than just orphans: are taxonomically-restricted genes important in evolution? Trends Genet. **25**(9): 404–413. doi:10.1016/j.tig.2009.07.006. PMID:19716618.

Kidd, M.R., Duftner, N., Koblmuller, S., Sturmbauer, C., and Hofmann, H.A. 2012. Repeated parallel evolution of parental care strategies within *Xenotilapia*, a genus of cichlid fishes from Lake Tanganyika. PLoS ONE, **7**(2): e31236. doi:10.1371/journal.pone.0031236. PMID:22347454.

Koblmuller, S., Salzburger, W., and Sturmbauer, C. 2004. Evolutionary relationships in the sand-dwelling cichlid lineage of Lake Tanganyika suggest multiple colonization of rocky habitats and convergent origin of biparental mouthbrooding. J. Mol. Evol. **58**(1): 79–96. doi:10.1007/s00239-003-2527-1. PMID:14743316.

Kocher, T.D. 2004. Adaptive evolution and explosive speciation: the cichlid fish model. Nat. Rev. Genet. **5**(4): 288–298. doi:10.1038/nrg1316. PMID:15131652.

Kozak, G.M., Brennan, R.S., Berdan, E.L., Fuller, R.C., and Whitehead, A. 2014. Functional and population genomic divergence within and between two species of killifish adapted to different osmotic niches. Evolution, **68**(1): 63–80. doi:10.1111/evo.12265. PMID:24134703.

Latta, L.C., Weider, L.J., Colbourne, J.K., and Pfrender, M.E. 2012. The evolution of salinity tolerance in *Daphnia*: a functional genomics approach. Ecol. Lett. **15**(8): 794–802. doi:10.1111/j.1461-0248.2012.01799.x. PMID:22583985.

Loh, Y.-H.E., Katz, L.S., Mims, M.C., Kocher, T.D., Yi, S.V., and Streelman, J.T. 2008. Comparative analysis reveals signatures of differentiation amid genomic polymorphism in Lake Malawi cichlids. Genome Biol. **9**(7): R113. doi:10.1186/gb-2008-9-7-r113. PMID:18616806.

Machado, H.E., Pollen, A.A., Hofmann, H.A., and Renn, S.C. 2009. Interspecific profiling of gene expression informed by comparative genomic hybridization: a review and a novel approach in African cichlid fishes. Integr. Comp. Biol. **49**(6): 644–659. doi:10.1093/icb/icp080. PMID:21665847.

Maere, S., Heymans, K., and Kuiper, M. 2005. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics, **21**(16): 3448–3449. doi:10.1093/bioinformatics/bti551.

Makinen, H., Papakostas, S., Vollestad, L.A., Leder, E.H., and Primmer, C.R. 2016. Plastic and evolutionary gene expression responses are correlated in European grayling (*Thymallus thymallus*) subpopulations adapted to different thermal environments. J. Hered. **107**(1): 82–89. doi:10.1093/jhered/esv069. PMID:26297731.

Mundodi, S., and Ireland, A. 2002. Generic GO Slim. Available from http://www.geneontology.org/GO.slims.shtml.

O'Connell, L.A., and Hofmann, H.A. 2012. Social status predicts how sex steroid receptors regulate complex behavior across levels of biological organization. Endocrinology, **153**(3): 1341–1351. doi:10.1210/en.2011-1663. PMID:22166981.

O'Connell, L.A., Matthews, B.J., and Hofmann, H.A. 2012. Isotocin regulates paternal care in a monogamous cichlid fish. Horm. Behav. **61**(5): 725–733. doi:10.1016/j.yhbeh.2012.03.009. PMID:22498693.

Oldfield, R.G., and Hofmann, H.A. 2011. Neuropeptide regulation of social behavior in a monogamous cichlid fish. Physiol. Behav. **102**(3–4): 296–303. doi:10.1016/j.physbeh.2010.11.022. PMID:21112347.

Oldfield, R.G., Harris, R.M., Hendrickson, D.A., and Hofmann, H.A. 2013. Arginine vasotocin and androgen pathways are associated with mating system variation in North American cichlid fishes. Horm. Behav. **64**(1): 44–52. doi:10.1016/j.yhbeh.2013.04.006. PMID:23644171.

Oldfield, R.G., Harris, R.M., and Hofmann, H.A. 2015. Integrating resource defence theory with a neural nonapeptide pathway to explain territory-based mating systems. Front. Zool. **12**: S16. doi:10.1186/1742-9994-12-S1-S16. PMID:26813803.

Oldham, M.C., Horvath, S., and Geschwind, D.H. 2006. Conservation and evolution of gene coexpression networks in human and chimpanzee brains. Proc. Natl. Acad. Sci. U.S.A. **103**(47): 17973–17978. doi:10.1073/pnas.0605938103. PMID:17101986.

Ophir, A.G., Wolff, J.O., and Phelps, S.M. 2008. Variation in neural V1aR predicts sexual fidelity and space use among male prairie voles in semi-natural settings. Proc. Natl. Acad. Sci. U.S.A. **105**(4): 1249–1254. doi:10.1073/pnas.0709116105. PMID:18212120.

O'Rourke, C., and Renn, S.C.P. 2015. Integrating adaptive trade-offs between parental care and feeding regulation. Curr. Opin. Behav. Sci. **6**: 160–167. doi:10.1016/j.cobeha.2015.11.010.

Pfenning, A.R., Hara, E., Whitney, O., Rivas, M.V., Wang, R., Roulhac, P.L., et al. 2014. Convergent transcriptional specializations in the brains of humans and song-learning birds. Science, **346**(6215): 1256846. doi:10.1126/science.1256846. PMID:25504733.

Pollen, A.A., Dobberfuhl, A.P., Scace, J., Igulu, M.M., Renn, S.C.P., Shumway, C.A., and Hofmann, H.A. 2007. Environmental complexity and social organization sculpt the brain in Lake Tanganyikan cichlid fish. Brain Behav. Evol. **70**(1): 21–39. doi:10.1159/000101067. PMID:17389793.

Qiu, J., and Cui, X.Q. 2010. Evaluation of a statistical equivalence test applied to microarray data. J. Biopharm. Stat. **20**(2): 240–266. doi:10.1080/10543400903572738. PMID:20309757.

Renn, S.C.P., Aubin-Horth, N., and Hofmann, H.A. 2004. Biologically meaningful expression profiling across species using heterologous hybridization to a cDNA microarray. BMC Genomics, **5**: 42. doi:10.1186/1471-2164-5-42. PMID:15238158.

Renn, S.C.P., Aubin-Horth, N., and Hofmann, H.A. 2008. Fish and chips: functional genomics of social plasticity in an African cichlid fish. J. Exp. Biol. **211**(18): 3041–3056. doi:10.1242/jeb.018242. PMID:18775941.

Renn, S.C.P., Machado, H.E., Jones, A., Soneji, K., Kulathinal, R.J., and Hofmann, H.A. 2010. Using comparative genomic hybridization to survey genomic sequence divergence across species: a proof-of-concept from *Drosophila*. BMC Genomics, **11**: 271. doi:10.1186/1471-2164-11-271. PMID:20429934.

Rodriguez-Santiago, M., Nguyen, J., Winton, L.S., Weitekamp, C.A., and Hofmann, H.A. 2017. Arginine vasotocin is expressed in surprising regions of the teleost forebrain. Front. Endocrinol. **8**: 195. doi:10.3389/fendo.2017.00195.

Salzburger, W., Renn, S.C.P., Steinke, D., Braasch, I., Hofmann, H.A., and Meyer, A. 2008. Annotation of expressed sequence tags for the East African cichlid fish *Astatotilapia burtoni* and evolutionary analyses of cichlid ORFs. BMC Genomics, **9**: 96. doi:10.1186/1471-2164-9-96. PMID:18298844.

Schumer, M., Krishnakant, K., and Renn, S.C.P. 2011. Comparative gene expression profiles for highly similar aggressive phenotypes in male and female cichlid fishes (*Julidochromis*). J. Exp. Biol. **214**(19): 3269–3278. doi:10.1242/jeb.055467. PMID:21900474.

Sen Sarma, M., Whitfield, C.W., and Robinson, G.E. 2007. Species differences in brain gene expression profiles associated with adult behavioral maturation in honey bees. BMC Genomics, **8**: 202. doi:10.1186/1471-2164-8-202. PMID:17603883.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., et al. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. **13**(11): 2498–2504. doi:10.1101/gr.1239303. PMID:14597658.

Smyth, G.K., Michaud, J., and Scott, H.S. 2005. Use of within-array replicate spots for assessing differential expression in microarray experiments. Bioinformatics, **21**(9): 2067–2075. doi:10.1093/bioinformatics/bti270. PMID:15657102.

Toth, A.L., Varala, K., Henshaw, M.T, Rodriguez-Zas, S.L., Hudson, M.E., and Robinson, G.E. 2010. Brain transcriptomic analysis in paper wasps identifies genes associated with behaviour across social insect lineages. Proc. R. Soc. B Biol. Sci. **277**(1691): 2139–2148. doi:10.1098/rspb.2010.0090.

Watanabe, M., Kobayashi, N., Shin-i, T., Horiike, T., Tateno, Y., Kohara, Y., and Okada, N. 2004. Extensive analysis of ORF sequences from two different cichlid species in Lake Victoria provides molecular evidence for a recent radiation event of the Victoria species flock — Identity of EST sequences between *Haplochromis chilotes* and *Haplochromis* sp. "Redtailsheller". Gene, **343**(2): 263–269. doi:10.1016/j.gene.2004.09.013. PMID:15588581.

Weitekamp, C.A., and Hofmann, H.A. 2016. Brain systems underlying social behavior. *In* Evolution of nervous systems. *Edited by* G.F. Striedter and J.H. Kaas. pp. 327–334.

Whitehead, A., Pilcher, W., Champlin, D., and Nacci, D. 2012. Common mechanism underlies repeated evolution of extreme pollution tolerance. Proc. R. Soc. B Biol. Sci. **279**(1728): 427–433. doi:10.1098/rspb.2011.0847.

Winslow, J.T., Hastings, N., Carter, C.S., Harbaugh, C.R., and Insel, T.R. 1993. A role for central vasopressin in pair bonding in monogamous prairie voles. Nature, **365**(6446): 545–548. doi:10.1038/365545a0. PMID:8413608.

Young, K.A., Liu, Y., and Wang, Z.X. 2008. The neurobiology of social attachment: a comparative approach to behavioral, neuroanatomical, and neurochemical studies. Comp. Biochem. Physiol. C Toxicol. Pharmacol. **148**(4): 401–410. doi:10.1016/j.cbpc.2008.02.004. PMID:18417423.