Conserved transcriptomic profiles underpin monogamy across vertebrates

Rebecca L. Young[a,b,1], Michael H. Ferkin[c], Nina F. Ockendon-Powell[d], Veronica N. Orr[e], Steven M. Phelps[a,f,g], Ákos Pogány[h], Corinne L. Richards-Zawacki[i], Kyle Summers[j], Tamás Székely[d,k,l], Brian C. Trainor[e], Araxi O. Urrutia[k,m], Gergely Zachar[n], Lauren A. O'Connell[o], and Hans A. Hofmann[a,b,f,g,1]

[a]Department of Integrative Biology, The University of Texas at Austin, Austin, TX 78712; [b]Center for Computational Biology and Bioinformatics, The University of Texas at Austin, Austin, TX 78712; [c]Department of Biological Sciences, The University of Memphis, Memphis, TN 38111; [d]Life Sciences, University of Bristol, Bristol BS8 1TQ, United Kingdom; [e]Department of Psychology, University of California, Davis, CA 95616; [f]Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin, TX 78712; [g]Institute for Neuroscience, The University of Texas at Austin, Austin, TX 78712; [h]Department of Ethology, Eötvös Loránd University, Budapest H-1117, Hungary; [i]Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260; [j]Department of Biology, East Carolina University, Greenville, NC 27858; [k]Milner Centre for Evolution, University of Bath, Bath BA2 7AY, United Kingdom; [l]Centre for Networks and Collective Behaviour, University of Bath, Bath BA2 7AY, United Kingdom; [m]Instituto de Ecología, Universidad Nacional Autónoma de México, Ciudad de México 04510, Mexico; [n]Department of Anatomy, Histology, and Embryology, Semmelweis University, Budapest H-1094, Hungary; and [o]Department of Biology, Stanford University, Stanford, CA 94305

Edited by Gene E. Robinson, University of Illinois at Urbana–Champaign, Urbana, IL, and approved November 26, 2018 (received for review August 14, 2018)

Social monogamy, typically characterized by the formation of a pair bond, increased territorial defense, and often biparental care, has independently evolved multiple times in animals. Despite the independent evolutionary origins of monogamous mating systems, several homologous brain regions and neuropeptides and their receptors have been shown to play a conserved role in regulating social affiliation and parental care, but little is known about the neuromolecular mechanisms underlying monogamy on a genomic scale. Here, we compare neural transcriptomes of reproductive males in monogamous and nonmonogamous species pairs of *Peromyscus* mice, *Microtus* voles, parid songbirds, dendrobatid frogs, and *Xenotilapia* species of cichlid fishes. We find that, while evolutionary divergence time between species or clades did not explain gene expression similarity, characteristics of the mating system correlated with neural gene expression patterns, and neural gene expression varied concordantly across vertebrates when species transition to monogamy. Our study provides evidence of a universal transcriptomic mechanism underlying the evolution of monogamy in vertebrates.

evolution | social behavior | gene expression | deep homology | mating systems

The diversity of animal social behavior has motivated a wealth of studies that explore variation in behavioral repertoires, sensory and cognitive specializations, and the ecological contexts in which they have evolved. Despite this extensive variation, the action of hormones, specifically sex steroids and neuropeptides (1), and other candidate pathways appears to be remarkably conserved in the regulation of social behavior (e.g., refs. 2 and 3). Moreover, recent studies support the intriguing hypothesis that coordinated activity of conserved gene sets underlies independent evolutionary transitions to similar behavioral phenotypes (4–8). It should thus not be a surprise that behavioral phenotypes may share molecular mechanisms regardless of their evolutionary history. Like extant animals, the most recent common ancestor had to meet challenges imposed by fluctuating internal and external conditions. The mechanisms used by these ancestral organisms to maintain homeostasis serve as the building blocks for the evolution of more derived behavioral responses as evidenced by the conserved role of homologous brain regions in processing social signals (9–11). At the molecular level, a "toolkit" of molecular pathways and gene networks can be preserved for hundreds of millions of years (12), and phenotypic novelty often can be attributed to new uses of such conserved gene sets (13, 14). The pervasiveness of conserved gene modules is highlighted by phenologs—functionally and physiologically unrelated phenotypes in different species with a statistical overrepresentation of shared sets of underlying orthologous genes (15). Finally, recent progress resolving evolutionary relationships among metazoan animals indicates that homoplasy is much more common than previously appreciated (16), even among phenotypes with overlapping molecular mechanisms [e.g., the nervous systems (17, 18)]. These discoveries have transformed our thinking about the origins and evolution of morphological and developmental phenotypes, but are rarely applied to investigations of the evolution of behavior.

Uncovering universal mechanisms of similar phenotypes requires a broadly comparative approach (19–21). Here, we ask to what extent similar neural transcriptomic profiles are associated with variation in social behavior across vertebrates, using mating system evolution as an example, and discuss the importance of

Significance

Social monogamy, typically characterized by the formation of a pair bond, increased territorial defense, and often biparental care, has evolved numerous times in animals. Despite the independent evolutionary origins of monogamous mating systems, several homologous brain regions and neuroendocrine pathways play conserved roles in regulating social affiliation and parental care, but little is known about the evolution of the neuromolecular mechanisms underlying monogamy. Here, we show that shared transcriptomic profiles are associated with monogamy across vertebrates and discuss the importance of our discovery for understanding the origins of behavioral diversity. We compare neural transcriptomes of reproductive males in monogamous and nonmonogamous species pairs of mice, voles, parid songbirds, frogs, and cichlid fishes. Our results provide evidence of a universal transcriptomic code underlying monogamy in vertebrates.

Author contributions: R.L.Y., L.A.O., and H.A.H. designed research; R.L.Y., M.H.F., N.F.O.-P., V.N.O., S.M.P., Á.P., C.L.R.-Z., K.S., T.S., B.C.T., A.O.U., G.Z., L.A.O., and H.A.H. performed research; R.L.Y., L.A.O., and H.A.H. analyzed data; and R.L.Y. and H.A.H. wrote the paper.

The authors declare no conflict of interest

This article is a PNAS Direct Submission.

Data deposition: All sequence data in this publication have been deposited in National Center for Biotechnology Information Gene Expression Omnibus, https://www.ncbi.nlm.nih.gov/geo (SuperSeries GSE123301, accession nos. GSM3499527–GSM3499536). All metadata and protocols/scripts are available on the Texas Data Repository (https://dataverse.tdl.org/dataverse/monogamy).

[1] To whom correspondence may be addressed. Email: youngrl@utexas.edu or hans@utexas.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1813775116/-/DCSupplemental.

our discovery for understanding the origins of behavioral diversity. While a more narrow focus within a clade may reveal more candidate genes with similar expression (6, 22), these results cannot be generalized across clades, limiting their broader implications. Monogamous species with nonmonogamous close relatives can be found in at least four major vertebrate clades (teleosts, amphibians, birds, mammals), providing an unparalleled opportunity to examine whether repeated transitions to a particular mating system evolve via shared transcriptomic mechanisms. We compare neural transcriptomes of reproductive males in monogamous and nonmonogamous species pairs of *Peromyscus* mice, *Microtus* voles, passeroid songbirds, dendrobatid frogs, and ectodine cichlid fishes. We characterize similarity in neural gene expression patterns among species in relation to similarity in mating system, ecological attributes, and evolutionary divergence time. While neither similarity in ecology nor in divergence time between clades explained gene expression similarity, neural gene expression varied concordantly across vertebrates between males of monogamous and nonmonogamous species. Genes with highly increased or decreased expression in monogamous species of one clade are likely to also have highly increased or decreased expression in the monogamous species of another clade. Our study provides evidence of a universal transcriptomic mechanism underlying monogamy in vertebrates.

Animal mating systems can be characterized as a suite of reproductive, parental, and agonistic phenotypes that can be highly variable among closely related (and even within) species depending on sex ratio as well as ecological factors such as predation risk, resource distribution, and extent of competition (23–25). Despite this potential diversity, similar mating systems have been described in numerous distantly related species. Social monogamy, for example, is typically characterized by the formation of a pair bond, increased territorial defense, and often biparental care; this suite of social behaviors has evolved independently numerous times (23, 26). Studies of pair bonding and parental care in mammals, birds, and fishes reveal a conserved, albeit complex, role for arginine vasopressin (AVP) and oxytocin (OT) as well as their receptors in regulating social affiliation (27–31). Recent studies have advanced our understanding of variation and the evolution of such pathways, for example, by illustrating how life history trade-offs underlie molecular, transcriptomic, and epigenetic variation among individuals (32). Less attention has been given to characterizing the complexities of neuromolecular mechanisms underlying monogamy on a genomic scale, including identifying novel candidate genes and pathways.

Results and Discussion

Shared and Unique Patterns Across Clades. Within each clade we compared expression of orthologous gene groups (OGGs) between species pairs (*SI Appendix*, Fig. S3). We found that the mean difference in expression between species pairs was near zero across all 1,979 OGGs (*SI Appendix*, Fig. S3*A*). Intriguingly, gene expression was least variable between the two bird species (*SI Appendix*, Fig. S3*B*), which may be explained by the fact that the mating systems of these species pairs are considerably more similar than those of the other clades (Fig. 2). Specifically, both *Anthus spinoletta* and *Prunella modularis* can form pair bonds (although less frequently in *P. modularis*) and exhibit direct paternal care (Fig. 2 and *SI Appendix*, Fig. S1 and Table S1). Despite these differences in OGG expression variation across clades, clade-specific Gene Ontology (GO) analysis revealed high conservation of GO term enrichment highlighting cell communication, signaling receptor activity, and membrane proteins as consistently associated with monogamy-related expression across clades (*SI Appendix*, Fig. S4).

Across Vertebrates, Gene Expression Varies Concordantly Between Males of Monogamous and Nonmonogamous Species. To assess concordance of OGG expression variation between monogamous and nonmonogamous species across vertebrate clades, we used the differential expression analysis software package DESeq2 (33). We assessed differential OGG expression multiple times with distinct evolutionary groupings from mammals to all clades where monogamous species of distinct clades were entered as interspecific replicates of monogamy (Fig. 3). We find that differences in OGG expression are generally concordant, particularly among OGGs that exhibit larger fold-differences (Fig. 3). In addition, most OGGs maintain directional concordance. For example, many of the OGGs that show increased expression (positive $\log_2$ fold-difference) in one evolutionary group (e.g., mammals) show increased expression in other evolutionary groups (Fig. 3). These results indicate that across vertebrates monogamous species recruit a common set of OGGs despite evolutionarily independent transitions to similar mating systems. We find that, as the evolutionary frame of reference is expanded and more distantly related clades are added to the analysis, fewer OGGs retain significance. In particular, we find a large decrease in above-threshold OGGs when amniote species pairs (i.e., mammalian and avian species pairs) are included compared with only the mammalian species pairs (Fig. 3). This finding likely reflects the decreased expression variation notable in our bird species pair (*SI Appendix*, Fig. S3). While the observed effect of the mating system for some OGGs may be smaller at broader taxonomic scales, adding species pairs increases the statistical power. Thus, with the exception of the comparison between mammals and birds described above, the decrease in number of OGGs meeting our threshold cutoff is quite small (Fig. 3).

When we included all species pairs, we identified 123 OGGs (6.2%) associated with monogamy across vertebrates (*SI Appendix*, Fig. S5). We find a number of OGGs significant at one level of analysis failing to meet the significance threshold at another (Fig. 3). Many differential expression analysis approaches, including DESeq2, rely on expression variation among biological replicates to determine differential expression. This approach is limiting when biological replicates are highly variable (34) as is the case here, where species of different clades are included as interspecific replicates of monogamy. In addition to evolutionary distance, a number of biological and technical features likely generate noise in our analysis (e.g., ecological differences among species, course-grained tissue sampling, and technical variation during sequencing). To identify shared transcriptomic patterns across monogamous species and assess whether the degree of overlap in expression between monogamous species is statistically significant, we utilize the Rank-Rank Hypergeometric Overlap (RRHO) approach (34). Comparing ranked fold-differences enables discovery of OGGs that share patterns of expression among monogamous species without requiring that expression values be similar across evolutionarily distant clades. Additionally, rather than adhering to discrete thresholds to identify candidates with similar expression, RRHO identifies candidates with coordinated directional shifts in expression using a sliding threshold (34) (Fig. 4).

Using RRHO including all 1,979 OGGs, we find an enrichment of OGG overlap in the on-diagonal extremes (i.e., at high $\log_2$ fold-differences in expression between monogamous and nonmonogamous species in both clades being compared; Fig. 4*D*). Most notably, we find enriched overlap of OGGs exhibiting decreased expression in monogamous species of all clades. In all comparisons, the concordant down-regulated (i.e., bottom left) quadrant of the RRHO plots has the greatest overall significance (i.e., most significant quadrant mean $-\log_{10} P$ value), strongly indicative of a universal signature of expression among monogamous vertebrates (Fig. 4 *C* and *D*). We generally do not find enrichment at small $\log_2$ fold-differences with a few exceptions (e.g., frogs and voles). This observation is likely due to the fact that the majority of OGGs exhibit small fold-differences (although they may be important for monogamy-related behavior). Subtle variation in expression of genes involved in clade-specific or basic cellular functions could mask potentially important similarities at this level. A consequence of this limitation of the RRHO approach is that, for OGGs with small fold-differences, a large overlap in OGGs is required to reach significant enrichment. Interestingly, clade comparisons including frogs exhibit the most enrichment in the off-diagonal (i.e., discordant) regions. Specifically, many orthologous gene groups that show decreased expression in other groups increase in the monogamous frogs, a finding consistent with the general trend of increased gene expression in the monogamous frogs (*SI Appendix*, Fig. S3*A*). This pattern is clear in all pairwise comparisons involving frogs, suggesting that the monogamy-related expression patterns in frogs are the least similar to the other clades. Comparisons including birds show the least enrichment in OGG overlap, possibly because the mating systems of the bird species in our study are relatively similar to each other, as discussed above. Importantly, however, birds still share the same overall pattern of OGG overlap in the on-diagonal extremes and show strong similarities in expression patterns in OGGs with high differential expression (e.g., *Bottom Left* and *Top Right* quadrants of birds vs. mice and birds vs. fish, Fig. 4*D*). Variation among clades in overall and specific patterns of OGG enrichment may reflect a difference among clades in species-specific features of monogamy. For example, male parental care is ancestral in the poison frog clade (35). Thus, biparental care associated with monogamy in our focal species evolved from male-biased care in our frog clade rather than from female-biased care as in the other clades examined here. Future analyses—including species exhibiting some but not all reproductive, parental, and agonistic phenotypes of monogamy as well as a more anatomically focused analysis—may shed light on this variation, enabling decoupling and association of gene expression patterns with more specific behavioral phenotypes.

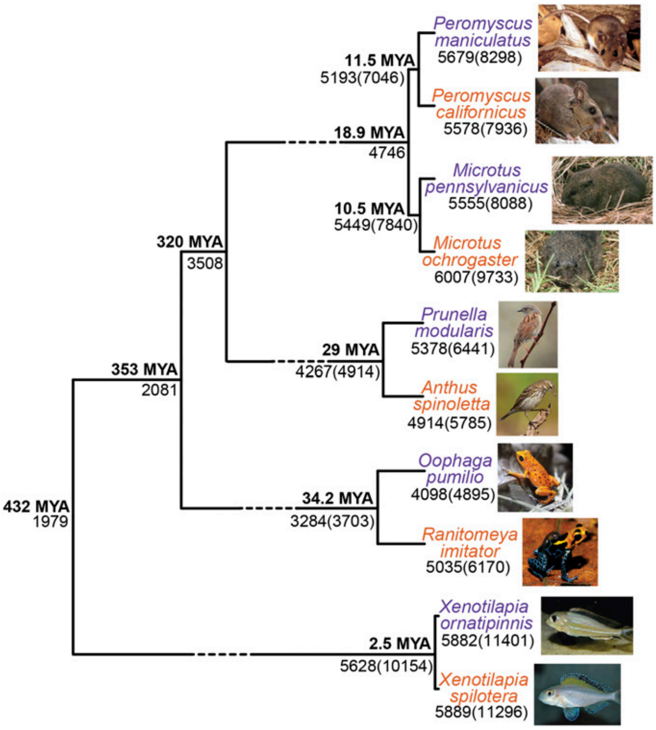

Fig. 1. Representative species exhibiting monogamous mating systems (including forming pair bonds, engaging in biparental care, and defending territory) were included in the analysis. To reduce clade-specific gene expression patterns and identify neural expression patterns associated with a monogamous mating system, monogamous species were paired with a closely related nonmonogamous species. Monogamous species for each clade are shown in orange, and nonmonogamous are shown in purple. Genes were grouped into OGGs using orthoMCL (44). Of the 6,125 OGGs identified across clades, the number of OGGs and genes in parentheses with expression data are shown at each tip (*SI Appendix*, Table S3). The number of shared OGGs and genes in parentheses (for species pairs) at each node are shown. Median divergence estimate was obtained from TimeTree (*SI Appendix*, Table S5) (46). Dashed branch lengths are not drawn to scale.

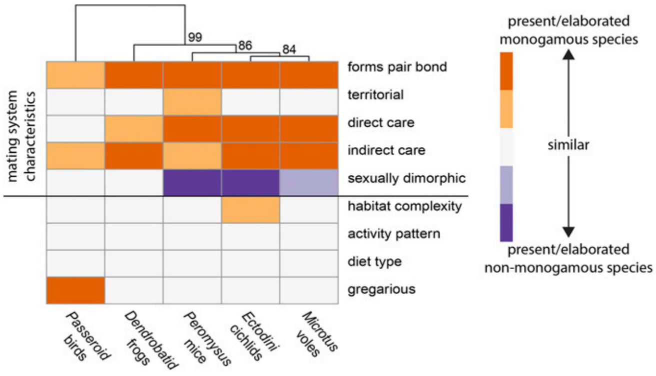

Fig. 2. Mating system and ecological distances between focal monogamous and nonmonogamous species pairs. Divergence between species pairs is indicated by color intensity. Orange indicates character states that are expressed (or more elaborated) in the monogamous species. Purple characters are expressed (or more elaborated) in the nonmonogamous species. White indicates that the character state is similar in the species pair. Profiles for all species in this study are provided in *SI Appendix*, Fig. S1 and Tables S1 and S2.

Discovery of Candidates Underlying Monogamy. We identified 24 candidate genes that are robustly (i.e., independently in both differential expression and RRHO analyses) associated with monogamous mating systems across vertebrates (Fig. 5, *SI Appendix*, Table S7 and Dataset S1). Interestingly, across lineages, genes involved in neural development, cell–cell signaling, synaptic activity (e.g., GRM6), learning and memory (e.g., DSCAM), and cognitive function are up-regulated in monogamous males. Conversely, genes involved in gene transcription (e.g., genes associated with RNA Polymerase II) and AMPA receptor regulation are down-regulated. These differences might indicate increased neural plasticity in the face of tighter transcriptional regulation in monogamous males. A more detailed discussion of these candidate genes, including a description of neurolocalization, is provided in *SI Appendix*, Table S7 and Dataset S1. While the majority of the candidate OGGs identified by both approaches to differential expression analysis (DESeq2 and RRHO) show increased expression in monogamous males, the greatest overlap enrichment occurs in OGGs with decreased expression in monogamous males (Fig. 4). Many of these OGGs do not survive threshold-dependent DESeq2 analysis and thus are eliminated from our list of robust candidates. Traditional candidates associated with monogamy-related behavior (e.g., AVP and OT as well as their receptors) were not identified among those consistently expressed in monogamous males, and in some species were not identified in the RNA-seq analysis (*SI Appendix*, Fig. S6). This finding does not indicate that these traditional candidates are unimportant in regulating monogamy-related behavior across these clades, but rather that their roles (if they are conserved across monogamous species) are not reflected in a consistent expression pattern at the coarse neuroanatomical scale that we sampled.

Phylogenetic and Ecological Correlates. The combined discovery of monogamy-related OGG expression by differential expression analysis (Fig. 3) and enrichment of gene expression rank correlation among clades (Fig. 4) supports the hypothesis that shared

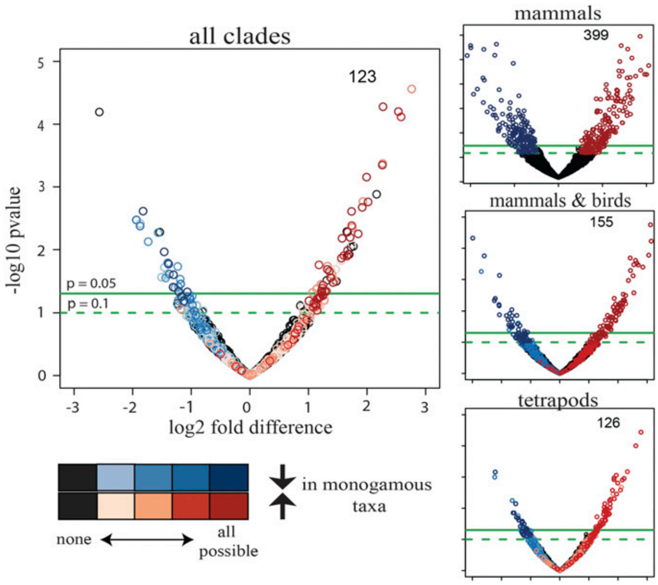

Fig. 3. Volcano plots indicating which of the 1,979 OGGs identified across all clades are differentially expressed at different taxonomic levels (mammals only, mammals and birds, tetrapods, all clades). Differential expression analysis was performed using DESeq2, where the monogamous (nonmonogamous) species of each clade were included as interspecific replicates of monogamy (nonmonogamy). Black circles show no differential expression at any taxonomic level. Differential expression analysis was performed on distinct evolutionary subgroups: mammals only (399 OGGs concordantly regulated), mammals and birds (155 OGGs), tetrapods (i.e., including frogs; 126 OGGs), and all four clades of vertebrates (i.e., including fishes; 123 OGGs). OGGs with a $-\log_{10}$ *P* value > 1 and a $\log_2$ fold-difference less than −1 (blue) or greater than 1 (red), respectively, are highlighted. The darker each circle, the more concordant across clades is the expression of the OGG that it represents. As more lineages are added to the analysis, more OGGs that are significant in one analysis fall below the significance threshold in another; however, adding species pairs increases the statistical power due to the increased number of interspecific replicates; thus, with the exception of evolutionary-subgroup mammals only versus mammals and birds, the decrease in number of OGGs meeting our threshold cutoff is small.

gene expression patterns underlie the behavioral expression of monogamous mating systems across vertebrate clades. However, it is possible that historical and ecological features shared between clades influence similarity in neural gene expression patterns. By design, species pairs were quite similar in their ecological attributes, differing primarily in specific characteristics of their mating systems (Fig. 2 and *SI Appendix*, Fig. S1 and Tables S1 and S2). Thus, the selection of species pairs by clade should minimize confounding ecological factors; however, several other factors may play a role in gene expression similarity. First, divergence times between species pairs vary between ∼2.5 and 34 million years (36). Historical contingency can bias the path of evolution such that more closely related species may be more similar due to shared evolutionary history. Second, elaboration of mating systems varies among species such that comparisons between monogamous and nonmonogamous mating systems are not equivalent across the clades. For example, the bird species included in this study share a number of mating system characteristics (Fig. 2). To assess the role of evolutionary history and mating system on gene expression divergence, we compared evolutionary and mating system distances to OGG expression distance for all species pairs. Neither evolutionary distance nor mating system distance correlated with OGG expression divergence between species pairs (Fig. 6 *A* and *B*). Notably, however, we find that the birds and the frogs are the most similar in their respective transcriptomes and are also the most similar in characteristics of mating system, while at the same time being the most distantly related of all species pairs (birds: 29 Mya; frogs: 34.2 Mya; Fig. 1 and *SI Appendix*, Fig. S3). When expression and mating system variation attributable to phylogeny is removed (using phylogenetic independent contrasts), we find a significant relationship between neural gene expression and mating system (Fig. 6*C*). Even though phylogenetic relatedness and ecological attributes affect neural transcriptome similarity across species in complex ways, together these observations indicate a critical role for mating system in driving gene expression similarity in the brain.

Conclusions. Using a comparative transcriptomics approach, we asked whether independent transitions to a monogamous mating system across four major clades of vertebrates are associated with shared neural gene expression patterns. A shared mechanistic basis of social behavior across distantly related clades has been documented at the level of neural circuitry where brain-region-specific expression of neurochemical genes is remarkably conserved in the Social Decision Making Network of the vertebrate fore- and midbrain (37). Further neural gene expression comparisons of aggressive behavior in bees, stickleback, and mice (4) provide

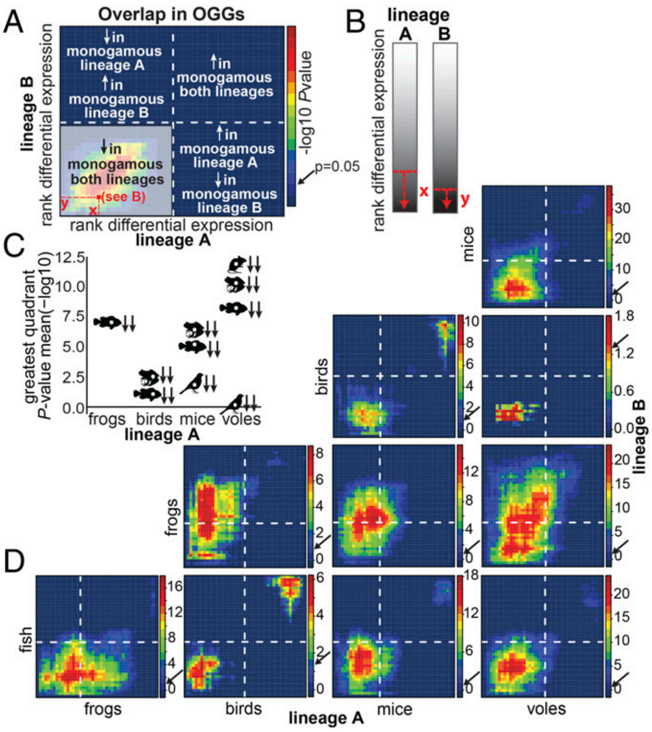

Fig. 4. RRHO of monogamy-related $\log_2$ fold-differences in gene expression for the 1,979 OGGs identified across all clades. Ranked $\log_2$ fold-differences in monogamous vs. nonmonogamous mRNA levels are binned into 44 sets of 45 OGGs from the most down-regulated to the most up-regulated in the monogamous species of each clade. OGG set overlap is compared in four quadrants defined by the transition between down- and up-regulation in each clade (*A*, dashed lines). The color of each pixel of the matrix (*A*, red square) indicates the enrichment in OGG set overlap at and above that differential expression threshold (*B*) and is expressed as the negative $\log_{10}$ of the Benjamini–Yekutieli-corrected *P* value. Significance of the enrichment is indicated by the pixel color with warm colors indicating increased enrichment. For each pairwise comparison of clades, the strength of OGG set overlap is summarized as the most significant quadrant mean negative $\log_{10}$ of the BY-corrected (*C*). Mean, median, and maximum *P* values for each quadrant are provided in *SI Appendix*, Table S6. Arrows next to the silhouettes indicate the directionality in lineage A (first) and lineage B (second) (*C*). RRHO analyses are shown for each pairwise clade comparison (*D*). Negative $\log_{10}$ of the BY-corrected *P* value color scale varies across plots. Dashed lines indicate the position of the switch point from down- to up-regulation in the monogamous species of each clade. Arrows on the color scale indicate the color at *P* value = 0.05.

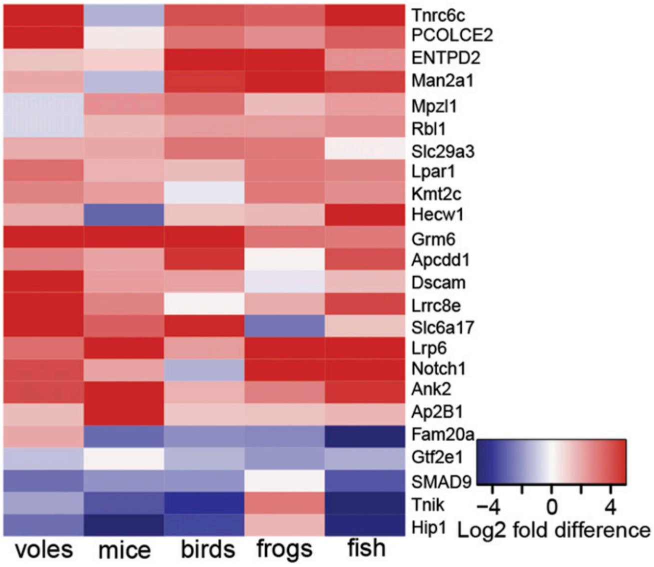

Fig. 5. Discovery of monogamy candidate genes (OGGs). To ensure a rigorous and conservative approach toward identifying OGGs robustly associated with monogamy across clades, any given gene/OGG had to fulfill two selection criteria. First, an OGG had to show a ±1 $\log_2$ fold expression difference between the monogamous and nonmonogamous species in at least four clades. Second, an OGG had to be among the most up- or the most down-regulated in 6 of the 10 RRHO analyses (as shown in Fig. 4); this again required an OGG to be concordantly expressed in at least four of the five lineages. Which clade (if any) is nonconcordant can vary for each OGG. Relative expression levels of these candidates are illustrated in the heatmap. Note that no clade is more often discordant than any other (*SI Appendix*, Table S7).

support for the hypothesis that broad similarities in social behavior across species may result from independent recruitment of ancient gene modules. However, inference from this analysis was limited by the complexity of homology inference among brains in bees and brain regions sampled in mice and stickleback. We expand on this previous comparative neural gene expression study of social behavior by comparing equivalent brain dissections and providing a phylogenetically informed assessment of transcriptome similarity in monogamous species of distinct and distantly related clades. Furthermore, by carefully scoring mating system characteristics and ecological attributes and by accounting for phylogenetic nonindependence, we test alternative hypotheses underlying the origins of transcriptomic similarities that originated in independent evolutionary transitions to monogamy across vertebrates. While our sampling of whole fore- and midbrains provides a coarse-grained view of the neuromolecular correlates of mating system, this approach circumvents potential problems with the sometimes tenuous homology inference of brain regions, ensuring consistent sampling across clades.

Our analysis reveals that monogamous species of distantly related vertebrate clades share more gene expression similarities than would be expected by chance, particularly in OGGs that are highly differentially expressed in monogamous species. Shared OGG expression across these clades provides evidence of deeply homologous mechanisms underlying the suite of reproductive, parental, and agonistic behavioral phenotypes that accompany monogamous mating systems. Our discovery adds to the growing body of research highlighting shared mechanistic bases of independently evolved morphological and physiological phenotypes (12). The present work illustrates that the independent evolution of complex behavioral phenotypes, like monogamy, is better conceptualized as the product of both parallel and convergent processes where many components of the underlying mechanism arise via parallel recruitment of deeply shared gene networks and some mechanisms evolve de novo (38–40). We argue that a conceptual shift is critical for achieving progress toward testing mechanistic hypotheses of the evolution of complex social behavior using comparative ‘omic approaches (as has been suggested for morphological traits; cf. ref. 41). Because the preponderance of evidence supports ancient, deeply homologous molecular and developmental pathways in evolutionary diversification, similar phenotypes should be expected to have mechanistic similarities regardless of evolutionary origins. Under this model, independent evolutionary transitions to similar behavioral phenotypes are expected to share similar mechanisms as a consequence of (*i*) integration of new traits with shared, ancestral physiological processes and their associated developmental and genetic pathways; and (*ii*) similar challenges to basic physiology that organisms experience in the evolution of these similar phenotypes. Whereas previous accounts of transcriptomic similarities among evolutionarily distinct complex behavior have focused on gene expression similarities, this conceptual approach provides a context to leverage both similarities and differences in gene expression to understand the mechanistic basis of complex trait evolution. Overall, our results indicate that independent evolutionary transitions to a monogamous mating system across vertebrate clades are accompanied by similar changes in gene expression in the male fore- and midbrain. In combination with the discovery of shared transcriptomic mechanisms associated with aggression (4), learned vocalizations in birds and humans

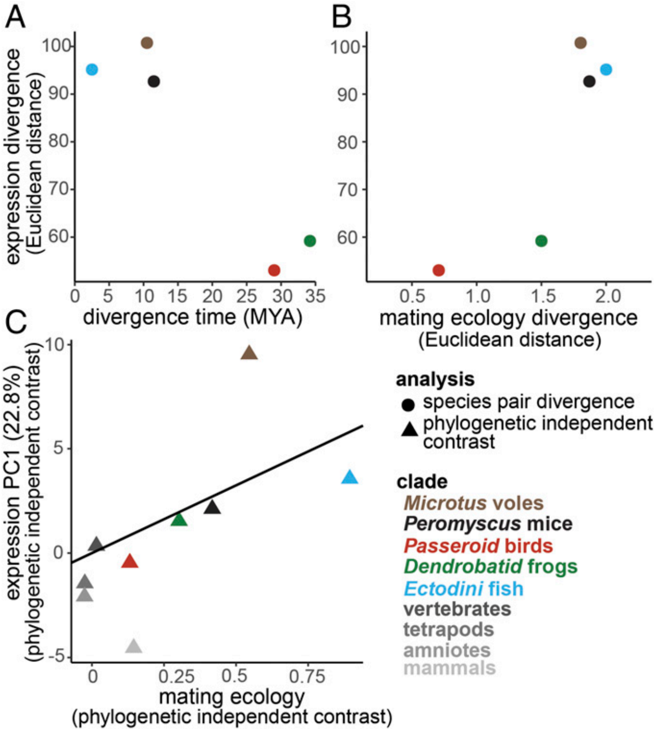

Fig. 6. Evolutionary divergence (*A*) and divergence in mating system characteristics (*B*) within clades does not correlate with gene expression distance (Spearman $\rho = -0.8$, $P = 0.95$; $\rho = 0.7$, $P = 0.12$, respectively). However, when expression variation due to phylogeny is removed (using phylogenetically independent contrasts) (*C*), we find a significant relationship between mating system and neural gene expression across clades (linear regression $r^2 = 0.6$, $t = 3.8$, $P = 0.005$). Mating system score was calculated as the sum of the mating system characteristics where higher values indicate more elaborated monogamy (i.e., males form pair bonds, provide direct and indirect parental care, exhibit high levels of territoriality, and are less sexually dimorphic; Fig. 2 and *SI Appendix*, Fig. S1 and Tables S1 and S2). A principal component analysis was performed on the differentially expressed (±1 $\log_2$ fold difference in at least one clade) and variable (*Upper* quartile of variance across all species) OGGs. This subset of 401 OGGs was also used for OGG expression divergence in *A* and *B*. See *SI Appendix*.

(5), and caste differentiation in hymenoptera (7), the results presented here considerably expand our understanding of how behavioral diversity evolves.

Materials and Methods

All animal care and use practices were approved by University of Texas at Austin; University of Memphis; University of California, Davis; University of Bath; East Carolina University; and Tulane University. Using an unbiased approach to identify neural gene expression patterns associated with a monogamous mating system and to limit clade-specific patterns in our cross-clade analysis, we sequenced and compared neural transcriptomic profiles from reproductive males of closely related monogamous and nonmonogamous species from four major classes of vertebrates ($n = 3$ pooled individuals per species): Mammalia (*Microtus ochrogaster* versus *Microtus pennsylvanicus* and *Peromyscus californicus* versus *Peromyscus maniculatus*); Reptilia–Aves (*A. spinoletta* versus *P. modularis*); Amphibia (*Ranitomeya imitator* versus *Oophaga pumilio*); and Actinopterygii (*Xenotilapia spilotera* versus *Xenotilapia ornatipinnis*) (Fig. 1). All sequence data in this publication have been deposited in National Center for Biotechnology Information Gene Expression Omnibus (42). Procedures for sample collection are detailed in *SI Appendix*. These selected species pairs differ in mating system characteristics, but are similar in other ecological attributes (Fig. 2; *SI Appendix*, Fig. S1 and Tables S1 and S2; and ref. 43).

One challenge associated with comparative analysis of gene expression patterns across distantly related species is identifying homologous tissues and comparable orthologous genes. To limit the requirement of brain region homology inference across distantly related clades, we extracted RNA from the combined fore- and midbrain tissues after hindbrain removal. To improve comparability in the transcriptomic analysis, we focus on expression of OGGs rather than individual genes. Across our 10 species we identified 1,979 OGGs using the sequenced-based ortholog-calling software package OrthoMCL (44). Our focus was on identifying monogamy-related expression patterns. Thus, when an OGG contained more than one gene (*SI Appendix*, Table S4; voles: 588, 30.0%; mice: 536, 27%; birds: 320, 16%; frogs: 228, 12%; fishes: 747, 38%), the gene with the highest $\log_2$ fold-difference between the monogamous and nonmonogamous species pairs was used for the remainder of the analysis (as in ref. 4). Genes in the same OGGs were generally concordant in directionality of expression difference (*SI Appendix*, Fig. S2 and Table S4). Thus, the selection of the most differentially expressed paralog did not obscure the overall similarity in expression pattern and allowed for downstream analysis of candidate genes. Thus, for each OGG and each clade the gene with the largest expression difference between the monogamous and nonmonogamous species was selected as the representative gene (Dataset S2 and ref. 45). For brevity, we refer to expression of this representative gene as OGG expression.

ACKNOWLEDGMENTS. We thank A. Ball, R. Harris, and R. Kar for assistance with the research; A. Battenhouse, B. Goetz, and E. Ortego (Center for Computational Biology and Bioinformatics, University of Texas at Austin), C. Jordan, and the Texas Advanced Computing Center (University of Texas Austin) for technical support; and D. Crews and members of the H.A.H. laboratory for discussion and helpful comments on earlier versions of this manuscript. This work was supported by the Alfred P. Sloan Foundation (BR-4900); NSF Grants IOS-1354942, IOS-1501704, and IOS-1601734 (to H.A.H.); NIH Grant R01 MH85069-S2 (to B.C.T.); and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences (to Á.P.).

1. O'Connell LA, Hofmann HA (2011) The vertebrate mesolimbic reward system and social behavior network: A comparative synthesis. *J Comp Neurol* 519:3599–3639.
2. White SA, Fisher SE, Geschwind DH, Scharff C, Holy TE (2006) Singing mice, songbirds, and more: Models for FOXP2 function and dysfunction in human speech and language. *J Neurosci* 26:10376–10379.
3. Campbell P, Reep RL, Stoll ML, Ophir AG, Phelps SM (2009) Conservation and diversity of Foxp2 expression in muroid rodents: Functional implications. *J Comp Neurol* 512:84–100.
4. Rittschof CC, et al. (2014) Neuromolecular responses to social challenge: Common mechanisms across mouse, stickleback fish, and honey bee. *Proc Natl Acad Sci USA* 111:17929–17934.
5. Pfenning AR, et al. (2014) Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science* 346:1256846.
6. Renn SCP, et al. (2018) Gene expression signatures of mating system evolution. *Genome* 61:287–297.
7. Morandin C, et al. (2016) Comparative transcriptomics reveals the conserved building blocks involved in parallel evolution of diverse phenotypic traits in ants. *Genome Biol* 17:43.
8. Shpigler HY, et al. (2017) Deep evolutionary conservation of autism-related genes. *Proc Natl Acad Sci USA* 114:9653–9658.
9. Porges SW (2001) The polyvagal theory: Phylogenetic substrates of a social nervous system. *Int J Psychophysiol* 42:123–146.
10. Crews D (2005) Evolution of neuroendocrine mechanisms that regulate sexual behavior. *Trends Endocrinol Metab* 16:354–361.
11. Goodson JL (2005) The vertebrate social behavior network: Evolutionary themes and variations. *Horm Behav* 48:11–22.
12. Shubin N, Tabin C, Carroll S (2009) Deep homology and the origins of evolutionary novelty. *Nature* 457:818–823.
13. Stergachis AB, et al. (2014) Conservation of trans-acting circuitry during mammalian regulatory evolution. *Nature* 515:365–370.
14. Carroll SB, Grenier JK, Weatherbee SD (2005) *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design* (Blackwell Scientific, Malden, MA), 2nd Ed.
15. McGary KL, et al. (2010) Systematic discovery of nonobvious human disease models through orthologous phenotypes. *Proc Natl Acad Sci USA* 107:6544–6549.
16. Dunn CW, Giribet G, Edgecombe GD, Hejnol A (2014) Animal phylogeny and its evolutionary implications. *Annu Rev Ecol Evol Syst* 45:371–395.
17. Liebeskind BJ, Hillis DM, Zakon HH, Hofmann HA (2016) Complex homology and the evolution of nervous systems. *Trends Ecol Evol* 31:127–135.
18. Holland LZ, et al. (2013) Evolution of bilaterian central nervous systems: A single origin? *Evodevo* 4:27.
19. O'Connell LA, Hofmann HA (2011) Genes, hormones, and circuits: An integrative approach to study the evolution of social behavior. *Front Neuroendocrinol* 32:320–335.
20. Toth AL, Robinson GE (2007) Evo-devo and the evolution of social behavior. *Trends Genet* 23:334–341.
21. Thomas AL, Davis SM, Dierick HA (2015) Of fighting flies, mice, and men: Are some of the molecular and neuronal mechanisms of aggression universal in the animal kingdom? *PLoS Genet* 11:e1005416.
22. Reinius B, et al. (2008) An evolutionarily conserved sexual signature in the primate brain. *PLoS Genet* 4:e1000100.
23. Emlen S, Oring L (1977) Ecology, sexual selection, and the evolution of mating systems. *Science* 197:215–223.
24. Kokko H, Rankin DJ (2006) Lonely hearts or sex in the city? Density-dependent effects in mating systems. *Philos Trans R Soc Lond B Biol Sci* 361:319–334.
25. Trivers R (1972) Parental investment and sexual selection. *Sexual Selection and the Descent of Man*, ed Campbell B (Aldine, Chicago), pp 136–179.
26. Shuster SM, Wade MJ (2003) *Mating Systems and Strategies*, eds Krebs JR, Clutton-Brock TH (Princeton Univ Press, Princeton, NJ).
27. Winslow JT, Hastings N, Carter CS, Harbaugh CR, Insel TR (1993) A role for central vasopressin in pair bonding in monogamous prairie voles. *Nature* 365:545–548.
28. Young KA, Liu Y, Wang Z (2008) The neurobiology of social attachment: A comparative approach to behavioral, neuroanatomical, and neurochemical studies. *Comp Biochem Physiol C Toxicol Pharmacol* 148:401–410.
29. O'Connell LA, Matthews BJ, Hofmann HA (2012) Isotocin regulates paternal care in a monogamous cichlid fish. *Horm Behav* 61:725–733.
30. Goodson JL, 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* 11:19–25.
31. Ophir AG, Gessel A, Zheng D-J, Phelps SM (2012) Oxytocin receptor density is associated with male mating tactics and social monogamy. *Horm Behav* 61:445–453.
32. Okhovat M, Berrio A, Wallace G, Ophir AG, Phelps SM (2015) Sexual fidelity trade-offs promote regulatory variation in the prairie vole brain. *Science* 350:1371–1374.
33. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550.
34. Plaisier SB, Taschereau R, Wong JA, Graeber TG (2010) Rank-rank hypergeometric overlap: Identification of statistically significant overlap between gene-expression signatures. *Nucleic Acids Res* 38:e169.
35. Brown JL, Morales V, Summers K (2010) A key ecological trait drove the evolution of biparental care and monogamy in an amphibian. *Am Nat* 175:436–446.
36. Hedges SB, Marin J, Suleski M, Paymer M, Kumar S (2015) Tree of life reveals clock-like speciation and diversification. *Mol Biol Evol* 32:835–845.
37. O'Connell LA, Hofmann HA (2012) Evolution of a vertebrate social decision-making network. *Science* 336:1154–1157.
38. Stern DL (2013) The genetic causes of convergent evolution. *Nat Rev Genet* 14:751–764.
39. Powell R, Mariscal C (2015) Convergent evolution as natural experiment: The tape of life reconsidered. *Interface Focus* 5:20150040.
40. Orgogozo V (2015) Replaying the tape of life in the twenty-first century. *Interface Focus* 5:20150057.
41. Mabee PM (2006) Integrating evolution and development: The need for bioinformatics in evo-devo. *Bioscience* 56:301–309.
42. Young RL, et al. (2018) Data from "Conserved transcriptomic profiles underpin monogamy across vertebrates." Gene Expression Omnibus. Available at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE123301. Deposited December 3, 2018.
43. Young RL (2018) Data from "Species ecology and mating characteristics." Texas Data Repository. Available at https://dataverse.tdl.org/dataset.xhtml?persistentId=doi:10.18738/T8/IWHEKH. Deposited December 11, 2018.
44. Li L, Stoeckert CJ, Jr, Roos DS (2003) OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res* 13:2178–2189.
45. Young RL (2018) Data from "Orthologous gene groups (OGGs) and their expression values." Texas Data Repository. Available at https://dataverse.tdl.org/dataset.xhtml?persistentId=doi:10.18738/T8/CLHOFF. Deposited December 11, 2018.
46. Kumar S, Stecher G, Suleski M, Hedges SB (2017) TimeTree: A resource for timelines, timetrees, and divergence times. *Mol Biol Evol* 34:1812–1819.