Mammal placental phenotypes are predictable from microRNA repertoires.

Jonathan Fenn[1], Jessica C. Edge[2], Vladimir Ovchinnikov[1,3], Olga Amelkina[2,4,5], Meléa Sinclair[1], Tiago H. C. De Bem[6], Alessandra Bridi[6], Irene Malo-Estepa[2], Angela Gonella-Diaza[7], Felipe Perecin[6], Juliano C. da Silveira[6], Guilherme Pugliesi[8], Flávio Vieira Meirelles[6], Zenab Butt[2], Haidee Tinning[2], Bastian Fromm[9], James O. McInerney[10], Niamh Forde[2*], Mary J. O'Connell[1,11*]

[1] School of Life Sciences, Faculty of Medicine and Health Sciences, University of Nottingham, Nottingham, United Kingdom

[2] Discovery and Translational Sciences Department, Leeds Institute of Cardiovascular and Metabolic Medicine, School of Medicine, University of Leeds, Leeds, United Kingdom

[3] Human Genetics, Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, CB10 1SA, United Kingdom

[4] Department of Reproduction Biology, Leibniz Institute for Zoo and Wildlife Research, Berlin, Germany

[5]Smithsonian's National Zoo and Conservation Biology Institute, Washington DC, USA

[6] Departamento de Medicina Veterinária, Faculdade de Zootecnia e Engenharia de Alimentos, Universidade de São Paulo, Pirassununga, Brazil

[7] North Florida Research and Education Center, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, Florida, USA

[8] Department of Animal Reproduction, School of Veterinary Medicine and Animal Science, University of São Paulo, Pirassununga, Brazil

[9] The Arctic University Museum of Norway, UiT - The Arctic University of Norway, Lars Thørings veg 10, 9006 Tromsø, Norway

[10] Department of Ecology and Evolution, Institute of Infection, Veterinary & Ecological Sciences, Faculty of Health and Life Sciences, University of Liverpool, Liverpool, United Kingdom

[11]The Natural History Museum, London, United Kingdom

*corresponding authors : n.forde@leeds.ac.uk and mary.o'connell@nottingham.ac.uk

Keywords: Mammals, reproduction, convergent evolution, microRNA, placenta

SUMMARY

Despite >100 million years of mammal diversification, similar placental morphologies have independently evolved multiple times, presenting a long-standing evolutionary puzzle: what genetic mechanisms lead to convergent forms? MicroRNAs (miRNAs), despite their diversity in mammals and known roles as developmental regulators, remain under-examined as drivers of morphological evolution. We identified presence-absence patterns for 429 miRNA gene families across 300 eutherian genomes and discovered that placental phenotype is highly predictable from genomic miRNA repertoires (classification accuracy >94-97%). We identified 30 miRNA gene families associated with placentation type, whose gene targets are enriched for developmental processes. Notably, convergent placental morphologies consistently involve identical miRNA families, revealing that evolution is constrained to predictable genetic pathways. We demonstrate evidence of mechanism through MIR-11968, associated with cotyledonary placentation and displaying tissue-restricted expression in cows. MiRNA-mediated regulation therefore constrains placental morphological diversification into reproducible programs, offering insights into how genetic architecture shapes the predictability of convergent evolution.

INTRODUCTION

Extant Eutheria share key defining traits, including implantation of the embryo and placental presence throughout embryonic development. However, these traits, while universally eutherian, manifest high levels of inter-species diversity, with well documented differences in embryonic development, implantation into the maternal endometrium, development of the placental shape and venous system, and the degree of invasiveness of the maternal blood supply at the maternal-fetal interface [1,2]. These innovations in placental morphology are cited as a potential mechanism for increasing reproductive isolation and fuelling speciation [1], making them a vital consideration when investigating diversification, adaptation and speciation within Eutheria.

Placental divergence has been studied intensely in terms of protein coding contributors, *e.g.* gene families disproportionately expressed in placenta [3], lineage specific variation in gene expansions in growth hormone and prolactin gene family and pregnancy specific glycoprotein gene family size [4], lineage-specific gene families such as IFNT in ruminants (which also displays variation across

ruminants) [5,6], and the independent integration of endogenous retro-viral elements such as synctins [7,8] in various lineages and their subsequent co-option to similar functions. Missing from our current understanding is what regulatory molecules, with the capacity to affect thousands of genes in dynamic, flexible and tissue specific ways, are coordinating these paths to placental phenotypic variation over short evolutionary timescales.

The eutherian-specific miRNA family expansion is the largest known miRNA expansion[9] , and has been linked to the evolution of the mammal placenta, embryonic implantation and other unique aspects of mammal reproduction and development [10–12]. Expression of many of these eutherian-unique miRNAs is required for successful embryonic implantation and subsequent placental development. Our previous research has demonstrated the essential role of miRNA repertoires in the origin of the placenta [13], specifically a cohort of 13 miRNA gene families evolved at the origin of placental mammals and have been retained in all descendent lineages. For example, MIR-378 [14] facilitates trophoblast invasion through their targeting of the Nodal pathway. Conversely, miRNA families that pre-date the origin of mammals such as MIR-155 [15], MIR-210 [16] and MIR-29 [17], which are conserved across vertebrates (*i.e.* MIR-155 [18]), and Bilateria (*i.e.* MIRs -210 and -29 [19]), have been shown to inhibit trophoblast invasion, and can be pathogenic if overexpressed [16,17]. These associations suggest that characteristic elements of eutherian reproduction may indeed be connected to the rapid expansion and diversification of miRNA repertoires in eutherian genomes since their divergence from marsupials. There is also a potential distinction between newer eutherian-specific miRNAs that promote placentation, and embryonic invasion, and older miRNAs that play a more inhibitory role.

Despite the central role of placenta in the evolution of novel reproductive strategies and life history traits in animals, it is still not clear which miRNAs are essential for placental development across all mammals (the 'conserved' or 'core' component) and which miRNAs are responsible for the diverse lineage-specific phenotypes. Previously, we showed that expression of a conserved 'core toolkit' is induced in a species-specific manner in response to conserved early pregnancy markers including the key hormone of pregnancy - progesterone (P4) - in bovine and human cells, resulting in downstream regulatory variation associated with different, species-specific implantation strategies [13]. These miRNAs did not respond to species-specific pregnancy signals, suggesting they mediate conserved mammal signalling rather than contribute to pregnancy strategy diversity [13]. Today, the rapid growth of sequenced genomes and transcriptomes for mammals can be used to understand which miRNA gene families are indispensable for eutherian reproduction, versus those that underpin phenotype diversity [20–22](Figure S1).

Reconstruction of the placental phenotype of the last common ancestor (LCA) of Eutheria from extant taxa has inferred a discoid shape, with an invasive, hemochorial placental membrane and a labyrinthine venous system [23,24]. We also know that derived placental phenotypes can arise convergently across Eutheria - for example villous placental venous systems have evolved in Xenarthra (sloths and armadillos), Perissodactyla (horses, rhinos and tapirs), Artiodactyla (even-toed ungulates), and primates. There is, however, generally strong phylogenetic structure to their distribution, with some traits being highly clade-specific (e.g. the cotyledonary placental phenotype in Pecoran ruminants). Broad associations between mammal placental phenotype and life history have been drawn. For example, larger longer-lived species with smaller litters are associated with diffuse placentas and epitheliochorial membranes, while smaller shorter-lived species with larger litters are associated with discoid placentas, labyrinthine venous systems and hemochorial membranes [24]. However, these patterns do not universally hold true across Eutheria. In fact, convergent placental phenotypes are observed in species inhabiting quite different ecological niches, *e.g.*, the independent evolution of villous venous systems in armadillos, primates and artiodactyls, and endotheliochorial membranes in rodents, elephants and manatees [1]. These observations highlight that the complex and dynamic evolutionary history of placental traits is shaped by both phylogenetic constraints and convergent adaptations to diverse ecological and reproductive pressures. The selective pressures and evolutionary dynamics underlying the evolution of reproductive traits are clearly complex, but may be better understood by analysing corresponding evolutionary changes in the regulatory molecular networks governing their development.

We address two fundamental questions in evolutionary biology: what molecular mechanisms drive the evolution of placental diversity? And, when similar placental complex traits emerge independently, do they use the same mechanisms, or do they evolve distinct solutions? MiRNAs, as post-transcriptional regulators of gene expression that are commonly implicated in reproductive outcomes, are an important layer of developmental control that can help answer these questions. The conventional expectation is that similar placental structures appearing in different clades would be controlled by distinct genetic mechanisms rather than shared molecular pathways. However, these morphologies are quite complex and their evolutionary emergence would involve coordinated changes in genome content, or expression patterns of large numbers of genes. Therefore, we hypothesise that despite the remarkable diversity of placental phenotypes across Eutheria, convergent evolution of specific placental traits is indeed underpinned by consistent molecular mechanisms, specifically involvement of the same miRNA gene families (*e.g.* independent loss of the same miRNA family in different lineages). If this hypothesis were correct, it would also imply that

placental morphology might be predictable by taking a census of the miRNA repertoire. Our hypothesis challenges the traditional view of convergent evolution as representing multiple independent 'solutions' to the emergence of homoplastic traits. Instead, we propose that the evolutionary pathways available for placental innovation may be more constrained and predictable than previously thought.

Using complete-genome sequence data from 326 mammal species, we investigated whether the miRNA profiles specifically associated with convergent placental phenotypes evolved independently and differently, or if similar regulatory mechanisms have repeatedly emerged to control similar morphological outcomes. We find that miRNA repertoire is indeed highly predictive of placental phenotype, and we uncover two broad patterns of miRNA-linked phenotypic innovation. Convergent traits are commonly associated with loss of specific miRNA families, mimicking experimental gene deletion, whilst non-convergent, single-origin traits are associated with the acquisition of completely new miRNA genes, as is most apparent in species with cotyledonary placentas. This demonstrates that lineage-specific phenotypes evolve via novel gene acquisition, mimicking natural gene insertion experiments. Furthermore, *in vitro* and *in vivo* analysis of the function of miRNA gene family MIR-11968, identified in our study as having very strong and clade-specific associations with the cotyledonary placental phenotype, revealed tissue- and species-specific expression patterns in bovine endometrial and placental tissue. This highlights the impact that expression of a single miRNA can have on crucial stages of mammal development. This work provides insight into the molecular underpinnings of mammal pregnancy success, identifies regulatory drivers of placental diversification, and highlights miRNA families as candidates for investigating reproductive pathologies in clinical and agricultural settings.

MATERIALS & METHODS

Genome Acquisition and Selection

In total 420 species genome assemblies were sourced from: DNA Zoo [20], the Zoonomia project [21] and the Vertebrate Genomes Project (VGP) [22] (Table S1 and Figure S2). Genomes were included for downstream analysis based on the completeness of 'core' vertebrate miRNA families, reported in every vertebrate species in MirGeneDB - any assemblies with fewer than 52 of the 53 expected miRNA families present were excluded. This was to allow for a rare, but possible loss of an otherwise

vertebrate-wide miRNA family. As detailed in the miRNA identification section, a MirMachine score per genome was also calculated [25].

Of 525 mammal genome assemblies, 396 recovered at least 52 of 53 core vertebrate miRNA gene families and had a MirMachine completeness score >=0.8. These 396 assemblies were filtered so that a single representative genome from each species was retained, favouring genomes with higher contiguity, as measured by contig:chromosome ratio [26]. This reduced the set to 326 unique species: 300 eutherians; 24 marsupials; and 2 monotremes. There were no significant differences between genome repositories in the metrics used to assess the quality of genomes for miRNA detection (Figure S3).

miRNA Identification

[*NOTE: All miRNA nomenclature in this study is provided by MirGeneDB (Table S2 contains corresponding miRBase nomenclature)*].

Using BLASTn [27] the database of mammal-exclusive miRNA gene families provided by MirGeneDB [19] was queried against all 300 eutherian genomes (e-value set to e^{-5}) and the results were converted to a miRNA gene family-level presence-absence matrix in R [28]. In parallel, MirMachine [25] was used to detect miRNA sequences expected to be found within mammals, by specifying the node as 'Mammalia' and the covariance model as 'Combined'. The bitscore-filtered GFF output files provided by MirMachine were used for all further analysis. MirMachine also estimated genome completeness with regards to a taxon's full miRNA gene family repertoire, and this genome quality score is calculated from the proportion of observed vs expected miRNA gene families present for a taxon within a given clade (in this case Mammalia). Genomes were deemed of acceptable quality if they contained at least ~80% of the total miRNA gene families expected from MirMachine Given the increased risk of false negative sequence detection relative to false positives, a miRNA gene family was counted as present if detected by MirMachine, BLAST, or both. While both approaches agreed on the presence-absence status for most miRNA gene families (on average 94.1% per genome), some were only detected by one method (*e.g.* MIR-1277 only detected by MirMachine) (discrepancies listed in Table S3). To determine the extent to which miRNA-phenotype associations were overrepresented in mammal-exclusive miRNA gene families, a presence-absence matrix of miRNA gene families which arose prior to divergence of mammals was also compiled, using the same approach *i.e.* BLASTn [27] searches,

querying all MirGeneDB derived pri-miRNA sequences of families that arose prior to emergence of mammals [19] against our 300 genomes.

To mitigate false classification of miRNA family absences, genomes used in the study were of highest possible quality and contiguity. For species represented by multiple assemblies, a preferred assembly was chosen based on completeness of the core 53 vertebrate miRNA gene families followed by the total number of miRNAs captured. In cases where these values were equal, genome contiguity (Contig:Chromosome ratio [26]) was used to select an assembly for analysis. To generate an updated putative 'core toolkit' of eutherian miRNAs, the finalised presence-absence matrix was queried for miRNA gene families which were present in every eutherian species.

Throughout the study we applied the consensus mammal phylogeny [21], with support for the Atlantogenata rooting [29,30]. The phylogenetic placement of species that are in our dataset but not in these phylogenomic studies was achieved using TimeTree [31].

miRNA Presence-Absence Structure

To ascertain levels of co-occurrence between miRNA gene families, and the presence of any defined clusters (of miRNA families themselves, and species according to their miRNA repertoires), the miRNA gene family presence-absence matrix was used to compute Jaccard dissimilarities between miRNA gene families across all genomes, and between genomes. Hierarchical clusters were generated from the resultant dissimilarity matrices, using the *hclust* function of the *stats* package [28]. Clustering methods were compared by bootstrapping for optimal stability using the *clusterboot* function of the *fpc* package [32], using the Ward D2 clustering method. The MetaMDS function of the *vegan* package [33] was used to visualise these clusters and extract the weightings of the distinguishing miRNA gene families.

Phenotype Data

Phenotype data from 334 mammal species was collected from the literature, as originally compiled by Wildman et al [1], and where appropriate incorporating additional data from the literature. Of these 334 species, 97 species were in our genome dataset, *i.e.* had high-quality genomes. Phenotype data includes placental membrane type (hemochorial, epitheliochorial or endotheliochorial), venous

system (villous, labyrinthine or trabecular) and placenta shape (cotyledonary, diffuse, zonary or discoid/bidiscoid). To determine whether phenotypes tended to co-occur (*i.e.* to assess whether a species with a given placenta shape is likely to have a specific venous pattern and/or membrane type) these three phenotypes were tested for non-independence with a log linear model in R and their clustering was assessed using Multiple Correspondence Analysis (MCA). A Mantel test using the *mantel* function of the *vegan* package [33] was used to determine whether similarity in phenotype is associated with similarity in miRNA gene family repertoire. Inter-species distances for the Mantel test were extracted from the phenotype MCA scores and from the presence-absence miRNA gene family matrix.

Ancestral Reconstruction

Ancestral phenotype reconstruction was performed using the R package *phytools* [34]. Reconstruction of each phenotypic state in the eutherian LCA was performed by taking summary results from 100 stochastic tree reconstructions, using the MCMC method of the *make.simmap* function of *phytools* [34]. The presence or absence of each miRNA gene family at the tips of the tree were used to reconstruct the ancestral state of the miRNA gene family with Dollo parsimony, which allows only one instance of gene gain, using the *Claddis* package in R [35]. Dollo reconstruction methods were chosen, as according to the MirGeneDB definitions used in this study, each miRNA family is by definition a family of single evolutionary origin [19]. A total of 1000 stochastic tree reconstructions were performed per miRNA gene family, and the most commonly-supported point of gain and point(s) of losses (if any losses occurred) were recorded (Table S17 and Figure S10). To determine whether mammal-specific miRNA gene families existed prior to the emergence of Eutheria, reconstruction was also performed for the mammal LCA, by including 26 non-eutherian species: 2 monotremes - echidna & platypus - and 24 marsupial species (Table S4).

Random Forest

To test the ability of miRNA gene families to predict placental phenotype, a random forest was built in R using the *ranger* package [36] for each of the three key placental phenotype categories. The training/test sets were made from a random 50:50 split of the data, and 500 trees were used in each forest, chosen as the point where out-of-bag error rates stopped decreasing for all forests. To assess

the importance of mammal-specific miRNA gene families in predicting phenotype, relative to background miRNA gene family repertoire, random forests were repeated using 174 miRNA gene families present in other non-mammal clades (*i.e.* pre-dating the origin of mammals) [19] (Table S5). The reconstructed miRNA gene family repertoire of the eutherian LCA was tested using each random forest to predict ancestral phenotypes.

For each phenotype, a control random forest was implemented which retained the same prevalence as the true presence-absence data for each miRNA gene family but with randomisation across species. MiRNA gene families were included as potentially having significant association with phenotype if they met the criteria of having: (i) a variable importance value of higher than the mean value plus one standard deviation, and (ii), at least a twofold increase in variable importance relative to the corresponding randomly simulated data (Table S6).

Evolink

Phylogeny-informed associations between phenotype and miRNA gene family presence were assessed using Evolink [37]. This approach parses binary phenotype data, necessitating phenotype states to be analysed separately to determine if there are associations with miRNA gene family presence-absence. For example, for the membrane phenotype, separate analyses were performed to test associations with each phenotype state; hemochorial, endotheliochorial and epitheliochorial membranes. For each phenotype, a list of miRNA gene families displaying significant association ($p<0.05$) - either positive or negative - with each phenotype state was produced (Table S7).

Phenotype-Associated miRNA gene Families

MiRNA gene families most likely to be implicated in determining placental phenotype, hereafter termed 'phenotype-associated miRNAs', were defined as those: (i) present in the random forest high-importance shortlist, and (ii), determined to be statistically-significant in the Evolink analysis (full output of miRNA-phenotype associations in Table S15). The ancestral reconstruction of the 30 phenotype-associated miRNAs was compared with the phenotypic reconstructions to determine whether miRNA gene families were gained or lost prior to-, coincident with-, or after-, the emergence of derived phenotypes.

miRNA Expression Data from Literature

For each of the 9 miRNA gene families in the updated core toolkit, and each of the 30 phenotype-associated miRNAs, any associations with developmental or reproductive biology described in clinical and laboratory studies were extracted from the literature and collated (Table S8). Endometrial miRNA expression data was compiled from published datasets for 7 eutherian species as of April-May 2024 (Table S9), and the presence of phenotype-associated miRNA expression across these species was determined. It was not possible to analyse the expression of the miRNA gene families in placental tissue due to reduced availability of cross-species datasets.

Target Prediction and Gene Ontology

To perform target prediction analyses on our miRNA gene families of interest, the 3' untranslated regions (UTRs) of each of the 300 eutherian genomes which passed quality filtering, were extracted using the *threeUTRsByTranscript* function of the *GenomicRanges* package in R[38]. The corresponding sequences were extracted from genome files using *BEDTools* [39]. If a gene had multiple alternate transcripts of the corresponding 3' UTR in a genome, the longest transcript was chosen for further analysis, with the goal of increasing the likelihood of detecting any potential seed-sequence matches. In order to assess the quality of UTR annotation within each genome, a 'UTR quality score' was calculated by performing a principal component analysis (PCA) on 3 metrics: the average UTR sequence length, the average proportion of ambiguous or uncalled bases per UTR sequence, and the proportion of annotated UTRs relative to genes in the genome. PC1 explained 54.8% of variance in these metrics, with higher PC1 scores corresponding to genomes with longer UTRs, UTRs with fewer ambiguous or uncalled nucleotides and higher proportions of UTRs per genes (Table S10). A quality cut-off threshold of the bottom 5th percentile value (UTR score -2.24) was used for inclusion in downstream analyses, leaving a total of 257 genomes (Figure S4). For the included genomes, where a species was represented more than once, the genome with the highest UTR score was selected, leaving a UTR database of 230 species. Some genomes had unresolvable gene naming issues in their annotations which prevented their processing via TargetScan [40], leaving a final set of 145 unique species for further analysis.

TargetScan was used to predict targets for each miRNA homolog in the mRNA 3'UTRs present in each genome [40] (excluding any UTRs corresponding to genes which were present in fewer than 5 genomes). Mature miRNA sequences for each genome were extracted from mirGeneDB by matching genomic miRNA sequences identified in the original BLAST search with the corresponding mature sequences taken from mirGeneDB . For each genome, the output of TargetScan provided a matrix of putative interactions between mature miRNA seed sequences and corresponding regions of mRNA 3' UTRs (*i.e.* 6-, 7- or 8-mers). To refine the target predictions for each miRNA, a machine-learning approach, TEC-miTarget [41], was used to validate seed sequence-UTR matches, with each match being scored for validity according to the provided model (trained on the miRaw dataset from TEC-miTarget). Any targets with prediction scores of 0.5 or higher were brought forward for further analysis.

To determine whether we see common biological pathways targeted by miRNA gene families across species, a gene ontology (GO) overrepresentation analysis was performed on the miRNA targets of each miRNA gene family, using the *enrichGO* function of the *clusterProfiler* R package [42], using human genome annotations ('org.Hs.eg.db'). For the background gene universe, all targets of all miRNAs from that species were used.

In order to assess whether GO terms conserved in the targets of core and phenotype-associated miRNA families were at higher prevalence compared to other miRNAs present only in mammals, we examined the top 100 most conserved GO terms for each miRNA which contained the terms 'development' or 'morphogenesis'. Differences between the three miRNA family categories ('core', 'phenotype-associated' and 'other') were identified using a Kruskal Wallis rank sum test, and between-group differences were examined using a pairwise Wilcox test.

Across all 9 core-toolkit miRNAs, terms were summarised as reduced-terms clusters using the *calculateSimMatrix* and *reduceSimMatrix* functions of the R package *rrvgo* [43], with default similarity threshold of 0.7 and term reduction by uniqueness. Phenotype-associated miRNA GO terms were divided into three subsets according to which miRNA gene families were associated with the three phenotypic states (membrane type, placental shape, and placental venous pattern). Within each phenotype, the 100 most conserved GO terms were summarised using *rrvgo.*

MIR-11968 Expression Analysis & Target Prediction

Phenotype-associated miRNA gene family MIR-11968, uniquely associated with cotyledonary placenta shape, was taken for further functional analysis. The expression levels of three MIR-11968 paralogs were experimentally assessed in *Bos taurus* (a model cotyledonary species). We used *in vivo* bovine endometrial samples collected on Day 16 of pregnancy (the day of pregnancy recognition) following embryo transfer of *in vivo* (*n=6*), *in vitro* (*n=8*), or cloned (*n=9*) embryos on day 7 of a synchronised cycle as described in De Bem *et al* [44]. Samples were collected from caruncular (aglandular regions and the site of cotyledonary formation) and intercaruncular regions from each animal (*n=23*).

Bovine endometrial luminal epithelial (*n=3*) and stromal (*n=3*) cells were isolated from reproductive tracts collected from an abattoir and grown *in vitro* as described in Tinning *et al* [45]. Stromal and luminal epithelial cells were separately treated with vehicle control (PBS) or 1µg/ml recombinant ovine IFNT (roIFNT) for 24 hours. Bovine endometrial epithelial glandular organoids (*n=3*) were derived from reproductive tracts collected from an abattoir [46] and treated with vehicle control (PBS) or 1µg/ml recombinant roIFNT for 24 hours on day 8 of growth post isolation. To act as a control, human immortalised endometrial epithelial cells (Ishikawa cells; n=3) were grown to confluency in DMEM/F12 with 10% FBS (fetal bovine serum) and 1% GSP (glutamine, streptomycin and penicillin). Untreated cells were plated (approximately 200,000 cells per well in a 6 well plate) for 24 hours.

All RNA was extracted using the miRNeasy Mini Kit as per manufacturers' instructions (Qiagen N.V., Hilden, Germany). For all samples, RNA concentration was measured using the DeNovix DS-11 FX+ spectrophotometer (DeNovix Inc., Delaware, USA) and diluted to 5ng/µl using DNase/RNase free water. The miRCURY LNA RT Kit (Qiagen N.V.) was used to reverse transcribe cDNA from RNA. MiRCURY LNA miRNA PCR Assays for bta-miR-11986 (miR-11968 paralog described in miRBase, but not MirGeneDB - potentially not a real miRNA, as evidenced by read count data on miRBase[47]), bta-Mir-11968-P1 and bta-Mir-11968-P2 were obtained from Qiagen. cDNA was diluted 1:60 for use in the assays and 3µl added to each well along with 5µl SYBR green, 1µl PCR assay and 1µl water. Real time qPCR was performed on the LightCycler 96 (Roche, Basel, Switzerland) and LightCycler 96 software used to calculate absolute quantification of results. Data was then exported to Microsoft Excel for further analysis and statistical tests and graphing was performed using Prism Graphpad. A one-way ANOVA with multiple comparisons test was performed to compare bta-Mir-11968-P1 expression in in vitro cell types with and without roIFNT treatment. A paired t-test was used to compare caruncular with intercaruncular expression of bta-miR-11986, bta-Mir-11968-P1 and bta-Mir-11968-P2.

For the two MIR-11968 paralogs described by MirGeneDB (P1 & P2) we tested whether there were specific functional pathways overrepresented within their pools of targets (as predicted from

TargetScan and TEC-miTarget analyses). This was done using the cow-specific gene ontology database (Org.Bt.eg.db)[48].

RESULTS

miRNAs detected across multiple genomes show low levels of phylogenetic clustering

Across the 326 mammal genomes, 429 miRNA gene families present only in mammals were identified: 299 were exclusive to Eutheria (*i.e.* not in monotremes or marsupials), 107 were unique to marsupials and monotremes, and a total of 23 miRNA gene families arose at the LCA of either Theria or Mammalia. After quality filtering for genome completeness, 61 miRNA gene families were found in >90% of eutherian genomes (Table S11 and Table S12). To understand the similarities of miRNA repertoires within Eutheria, hierarchical clustering was performed on the miRNA presence-absence matrix. Eutherians fell into three well-supported clusters: simian primates (Infraorder: *Simiiformes*), rodents (Order: *Rodentia*), and all other species. In addition, miRNA families also formed three stable clusters: one corresponding to a core group of somewhat uniform distribution across the eutherian tree, and two smaller clusters, with increased prevalence in *Rodentia* (18 species), and *Simiiformes* (44 species), respectively (Figure 1). Concordance between the dendrogram derived from hierarchical clustering and the consensus species phylogeny was very low (normalised Robinson-Foulds distance = 0.894) indicating that across Eutheria, miRNA repertoires generally have low agreement with phylogenies derived from standard molecular phylogenetics. The clade-specific clusters in rodents and simian primates represent distinct exceptions to this pattern, as each clade contains well-defined clusters of miRNA gene families uniquely associated with their respective taxonomic groups.

Ancestral reconstruction placed 121 miRNA gene families in the eutherian LCA, with only 9 inherited from the LCA of all mammals (including monotremes) and 15 from the therian LCA, indicating a burst of eutherian-specific miRNA diversification. 17 miRNA families present in the LCA of Eutheria belong to the cluster associated with the Rodentia, and 7 to the cluster associated with Simian primates (Figure 1), indicating retention of these miRNA families in their respective clades. The other 97 families belong to the broad cluster with no strong phylogenetic signal (note - these miRNA families show some discrepancy with those placed at the eutherian ancestral node by MirGeneDB). The widely-shared

miRNA gene family cluster is, therefore, mostly derived from a miRNA expansion arising between the stem mammal and stem eutherian lineages (Table S17).

Analysis of our substantially expanded dataset revealed nine miRNA gene families to be truly synapomorphic across Eutheria: MIR-28, MIR-331, MIR-423 (as previously reported in Taylor *et al* [13]), along with MIR-1251, MIR-506, MIR-154, MIR-3059, MIR-328 and MIR-376. The remaining 10 miRNA gene families previously reported to be synapomorphic [13], while widely distributed, show discrete losses across Eutheria when examined across this comprehensive sampling (Figure 1). Many of these losses are convergent: MIR-542 was absent from 7 rodent species, as a result of 3 separate loss events; MIR-433 was lost from both species of *Rousettus* fruit bats, as well as both representative species of *Suidae*; MIR-340 is absent in both representative hyrax species (*Procaviidae*), and finally, MIR-185 is absent in 3 members of *Pilosa*, a result of two independent losses. While functional redundancy may explain these losses through other miRNA families assuming similar roles, no lost core miRNA family shares identical mature sequences with other families (Table S13). However, paralogs from two miRNA families show potential seed sequence compensation: MIR-340 potentially compensated by MIR-452 (present in all three species lacking MIR-340), and MIR-505 potentially compensated by MIR-7181 (with the exception of the Madagascan fruit bat, *Eidolon dupreanum,* where both MIR-505 and MIR-7181 are absent). Despite these sporadic losses, core miRNA families show 95.6-99.6% prevalence across species, indicating significant roles in eutherian evolution.

miRNA repertoire is strongly associated with phenotype

Combinations of placental phenotypes are non-random (loglinear analysis p~0), *i.e.* variations in the membrane, venous system and shape tend to fall into specific groupings (Figure 2). Specifically, all placentas with epitheliochorial membranes have either cotyledonary or diffuse placental shape, and villous venous system, while those with hemochorial membranes almost all have discoid shape and labyrinthine venous system. The ordination pattern in the MCA plot reveals distinct clusters of associated traits (Figure S5). Distances extracted from MCA scores are significantly associated with distances in miRNA gene family presence-absence (Mantel $r = 0.193$, $p = 0.0009$), suggesting that miRNA repertoire is strongly associated with phenotype state. While some of the observed association may stem from shared phylogenetic signals in both miRNA repertoire and placental phenotype, our analysis shows that the non-phylogenetic distribution of miRNA repertoires across the study species is itself significant. This strengthens the observed association between miRNA and phenotype.

Uniquely Derived phenotypes are associated with miRNA gain

Three phenotype-associated miRNA families (MIR-11968, MIR-10182 and MIR-3431) show no secondary losses and are exclusively linked to cotyledonary placentation. MIR-11968 emerged in the pecoran ruminant LCA, while MIR-10182 and MIR-3431 arose earlier in the Ruminantia LCA, which includes both pecoran ruminants and mouse deer (*Tragulidae*) (Figure 3). To investigate the functional significance of the association between MIR-11968 and the cotyledonary phenotype, we examined MIR-11968 expression across bovine endometrial tissues and cell types, providing the first insights into how clade-specific miRNAs may drive specialised placental phenotypes.

As expected (given the clade specific nature of MIR-11968), we found no detectable expression of bta-miR-11986, Mir-11968-P1 or Mir-11968-P2 in human Ishikawa (endometrial epithelial cell line) cells. *In vitro* analysis of expression levels in bovine endometrial luminal epithelial, stromal and organoid cells showed bta-miR-11986 and Mir-11968-P2 are not detected. However, Mir-11968-P1 was expressed in bovine luminal epithelial cells and organoids, but not in stromal cells. Bta-miR-11986 was significantly ($p<0.0001$, one-way ANOVA with Šidák's test for multiple comparisons) more highly expressed in bovine endometrial organoids than luminal epithelial cells. Treatment with the pregnancy recognition signal roIFNT did not significantly alter expression of bta-Mir-11968-P1 in organoids or luminal epithelial cells compared to their respective controls. Bta-Mir-11968-P1 was more highly expressed ($p<0.0001$, one-way ANOVA with Šidák's test for multiple comparisons) in organoids treated with roIFNT than luminal epithelial cells treated with roIFNT (Figure 4A). To determine whether MIR-11968 shows tissue-specific expression patterns relevant to cotyledonary placentation, we next examined its expression in caruncular versus intercaruncular endometrial tissues, which represent the specialized versus non-specialized regions of the bovine endometrium.

In bovine species, the caruncular tissue forms the primary site of fetal-maternal interaction, where caruncles interdigitate with cotyledons of the placenta, whereas the surrounding intercaruncular regions contain glandular tissue. Using caruncular and intercaruncular endometrial samples collected *in vivo*, we find expression of bta-miR-11986, bta-Mir-11968-P1 and bta-Mir-11968-P2 in both tissues. The expression of bta-miR-11986 ($p=0.009$), bta-mir-11968-P1 ($p=0.004$) and bta-Mir-11968-P2 was significantly ($p=0.0009$) higher in caruncular tissue compared to intercaruncular (Figure 4C-E). Finally, embryo developmental competency had no significant effect ($p>0.05$) on bta-miR-11986, bta-Mir-11968-P1 or bta-Mir-11968-P2 expression *in vivo*.

After filtering for TEC score, 5,117 targets of Mir-11968-P1 were identified with only one GO term showing significant overrepresentation: GO:0007264 (small GTPase-mediated signal transduction). GTPases are involved in many stages of mammal development [49,50] , and their expression levels have been associated with pre-eclampsia in humans [51]. These results indicate they have a potential role in facilitating the development of novel placental structures. Contrastingly, after filtering for TEC score there were no targets predicted for the MIR-11968 paralog - Mir-11968-P2, *i.e.* these two family members display highly divergent patterns of expression and targeting.

Unlike the unique and unreversed property of MIR-11968, convergent evolution characterises most non-ancestral phenotypes (with the exception of cotyledonary shape and trabecular venous pattern). Indeed, some cases show alignment between miRNA family evolution and phenotypic convergence through preferential retention of ancestral miRNAs. For example, MIR-1842 has been lost 22 times across Eutheria but only once in a diffuse placental species (*Tragulus javanicus*). In Carnivora, where zonary placenta evolved secondarily from diffuse placenta, MIR-1842 has been lost 4 times, suggesting reduced selective pressure for its retention following the loss of the diffuse phenotype. Due to correlated evolution of chorial membrane and placental shape, MIR-1842 shows similar retention patterns in epitheliochorial membrane clades, again with only *T. javanicus* showing loss. Three eutherian-ancestral families (MIR-3187, MIR-197, MIR-1307) show higher retention rates in clades with a villous venous pattern: 72.2%, 94.4%, and 96.3% respectively, versus their retention rates of 35.3%, 70.2%, and 71.3% in other species. Given that only two venous phenotypes exist (an exception being the *Callithrix jacchus* with a trabecular pattern), these results could alternatively reflect increased propensity to lose the ancestral labyrinthine phenotype when these miRNAs are lost.

Multiple miRNA families show coordinated losses across lineages with similar placental features. Zonary placentation, for example, has only ever arisen in organisms that have lost MIR-872, MIR-3187, and MIR-6119, and MIR-598 is absent from all but three zonary species. Likewise, endotheliochorial membranes, which arose independently at least six times, consistently associate with MIR-3187 loss (through three ancestral loss events) and MIR-598 loss in all species except aardvark and large tree shrew (four independent losses). These patterns indicate that specific miRNA losses seem to be either necessary, or they strongly potentiate the evolution of particular placental morphologies, and this is a consequence of the same regulatory changes recurring across distantly related lineages (Figure 3; complete phylogenetic reconstructions in Figures S11-S13, Table S17).

These miRNA-phenotype associations reveal two quite different evolutionary patterns. Convergent traits show repeated miRNA loss events, suggesting that certain developmental constraints make

specific regulatory losses advantageous when evolving particular phenotypes. For example, all endotheliochorial membrane lineages independently lost MIR-3187, indicating strong selective pressure against this miRNA in that developmental context. Non-convergent traits, in contrast, are associated with miRNA gains, as seen with cotyledonary placenta and MIR-11968 acquisition in ruminants. This suggests that novel regulatory gains can enable unique morphological innovations but depend on rare, lineage-specific evolutionary opportunities. The balance between convergence driven by loss, and innovation driven by gain, are likely to reflect the developmental complexity of each trait and the genomic flexibility available for its evolution.

MiRNA gene family repertoire accurately predicts phenotype

Using our new and expanded dataset, the reconstructed phenotype of the eutherian LCA was inferred to consist of hemochorial membrane, labyrinthine venous system and discoid shape (Figure S6), in concordance with previous studies [1,23] . We also identified several phenotypes that have arisen independently in Eutheria. Epitheliochorial membranes evolved independently in the Lemuroidea, and in the common ancestor of Artiodactyla and Perissodactyla, while endotheliochorial membranes later emerged as a secondary innovation in the Carnivora. Diffuse membranes arose separately in Lemuroidea and in the LCA of Perissodactyla and Artiodactyla. Zonary placental structure evolved independently in three lineages: Carnivora; Paenungulata (hyraxes & elephants), and aardvarks. Similarly, villous membranes emerged independently in four distinct groups: Perissodactyla; Artiodactyla; Cingulata (armadillos), and in Primates. Together, these patterns highlight extensive convergent evolution of placental traits in mammals, implying that similar selective pressures have repeatedly driven the emergence of similar placental phenotypes across diverse lineages, or at the very least that there are fundamental constraints on the combinations of phenotypes that can evolve.

As determined by the random forest analysis, the placental phenotypes of extant species were well predicted by miRNA repertoire (out-of-bag error rates: 0.03-0.06). However, the reconstructed eutherian LCA miRNA repertoire incorrectly predicted epitheliochorial membrane, diffuse shape, and villous venous pattern instead of the inferred ancestral hemochorial-discoid-labyrinthine combination. Naturally, more recently evolved miRNAs are not available to use in the predictions for the eutherian LCA. This prediction failure suggests extensive regulatory network rewiring throughout eutherian evolutionary history, where ancestral miRNAs performed different developmental roles than their modern counterparts. This indicates that the placental miRNA toolkit is evolutionarily

plastic, with strong contemporaneous phenotype-miRNA associations arising through derived gains and losses rather than from tinkering with otherwise broadly conserved ancestral regulatory programs.

The failure to predict the eutherian LCA phenotype from miRNA repertoire reconstruction is clarified by additional analyses showing that mammal-specific miRNA families drive phenotypic prediction accuracy. When we included non-mammal miRNA families as controls, out-of-bag error rates increased significantly, indicating that miRNA evolution subsequent to the origin of mammals is more important for placental phenotyping than ancient miRNA families (Table S14). This explains why the eutherian LCA repertoire - dominated by ancient miRNAs and devoid of recent, derived miRNAs - cannot predict ancestral placental phenotype, while extant genomes with their complement of derived mammal miRNAs show strong predictive power. According to our criteria, 25 miRNA families show high importance for placental shape prediction, 28 for membrane type, and 21 for venous pattern (Table S6), representing prime candidates for functional investigation of placental development. Evolink analysis, which accounts for phylogeny, identified 30 miRNA gene families significantly associated with placental phenotypes, with MIR-3187 notably associated with 6 different phenotype states (Tables S7, S16). Associations with ancestral phenotypes were typically negative, reflecting the rarity of ancestral ‘core toolkit’ miRNA loss events.

Targets of core and phenotype-associated miRNAs display a range of biological functions including developmental processes

Gene targets of core miRNAs show enrichment for developmental GO terms among their 100 most conserved associations, including 'stem-cell differentiation' and 'regulation of developmental growth' (Figure 5), confirming developmental roles for miRNA families present across all eutherian mammals. Similarly, phenotype-associated miRNAs showed strong conservation of development-related GO term associations, with 'stem cell differentiation', 'morphogenesis of branching epithelium', and 'developmental growth involved in morphogenesis' evident across all three placental phenotypes (Figure S9, Table S16). Interestingly, the set of core miRNA families showed significantly higher prevalence of 'development' and 'morphogenesis' terms compared to the phenotype-associated miRNAs and other mammal-unique families (Kruskal-Wallis $p=2.12\times10^{-3}$, Figure S14), suggesting core miRNAs have broader developmental regulatory roles while phenotype-associated miRNAs show more specialized functions.

Transcriptomic analysis of endometrial tissue across seven species (Figure 6) revealed expression of all nine core eutherian-synapomorphic miRNA families in at least one species, with six (MIR-154, MIR-328, MIR-376, MIR-28, MIR-331, MIR-423) expressed across all species. Among the ten families with rare losses identified by Taylor et al. [13], five showed universal endometrial expression while MIR-505 was undetected in any species. Overall, 58 of the 104 phenotype-associated miRNA families showed endometrial expression when present in a species' genome, ranging from universal expression in rat to only 6 of 24 families in goat, though experimental design differences may influence these patterns.

DISCUSSION

Placental evolution shows surprising genomic predictability: miRNA gene family repertoires accurately predict placental morphology across mammal species, with convergent phenotypes consistently involving losses of the same miRNA families. The ability to accurately predict placental characteristics from miRNA gene family presence-absence patterns indicates that developmental constraints channel placental innovation through limited genetic pathways rather than allowing unlimited mechanistic diversity, and that there is a tight link between evolutionary patterns for these particular elements and the resulting phenotypes. The independent evolution of these tight genotype-phenotype pairs suggests that there are only a limited number of evolutionary avenues available for placental diversity. The mammal placenta is a complex organ, its evolution having required exaptation of maternal and foetal tissues, as well as co-option of existing networks of gene regulation, in a manner which may be evolutionarily constrained, as suggested by the convergent evolution of viviparity in mammals and squamates [52–55]. It stands to reason that similar patterns of evolutionary constraint may influence the diversification of placental form and function, such that shared physiological and genomic characteristics in distantly-related taxa may lead to the adaptation of similar placental phenotypes, given a similar background of selective pressures.

MIR-3187, MIR-197, and MIR-1307, while having patchy distributions across the phylogeny, are strongly associated with zonary placenta absence. Our reconstructions suggest multiple independent losses, even though this conflicts with the documented rarity of miRNA gene loss in evolution [56]. Two alternative explanations merit consideration: firstly, rapid sequence divergence may cause some miRNA families to become undetectable rather than truly lost. Alternatively, our single-origin assumption for miRNA families may be incorrect, and multiple independent gains are not accounted

for by the taxonomy of miRNA gene families used in this study [19]. Future analyses that account for similarity and convergence between mature miRNA sequences from different miRNA families will provide further insight. Irrespective of the underlying evolutionary mechanism, the genomic presence-absence patterns for recognisably homologous miRNA families, remain strongly predictive of placental phenotype. This suggests that whether through loss, gain, or sequence divergence, the detectability of these miRNA families correlates strongly with specific developmental programs underlying placental morphology.

Ancestral reconstruction of miRNA gene families showed one miRNA family which arose at the ancestor of Pecora (MIR-11968), a clade with a characteristic cotyledonary placenta, and two (MIR-2331 and MIR-3957) which arose at the ancestor of Ruminantia, which includes infraorders Pecora and Tragulina (chevrotains). Of these, only MIR-11968 showed presence in every taxon tested within Pecora and no others, and as such it is the strongest candidate within our phenotype-associated miRNA gene families of a novel miRNA gene family directly involved in a marked shift in placental phenotype, and being necessary to maintain that phenotype. The detection of MIR-11968 in bovine glandular endometrial epithelial organoids but not in human Ishikawa cells, as well as its presence in the caruncular tissue *in vivo* provides strong evidence that this miRNA gene family is involved in the peri-implantation period of pregnancy. Our *in vitro* and *in vivo* analyses of MIR-11968 demonstrated persistent expression in relevant reproductive tissues and supported a causal effect on cotyledonary placental shape, showing that this miRNA-phenotype relationship is not simply correlational. The opposing patterns of expression of MIR-11968 in glandular and aglandular cell types in pregnant vs non-pregnant animals, suggests that the presence of the embryo influences upregulation of MIR-11968 in non-glandular cells of the endometrium via conceptus-derived molecules or maternal-conceptus interactions.

The Core Ancestral Eutherian Toolkit Revisited

Dense genomic sampling revealed that many previously identified 'core' eutherian miRNA families show sporadic losses, contradicting assumptions about their universal requirement for mammal pregnancy. While miRNA loss is typically rare, these specific losses show no significant phenotypic associations and mostly occur in taxa retaining ancestral placental states. The losses may affect developmental processes not examined here, such as implantation strategy, rather than the three placental phenotypes analysed in this study.

miRNA Targets and Biological Significance

The strong evolutionary associations between miRNA repertoire and phenotype, did not translate into clear mechanistic insights from target prediction analysis - revealing an important aspect of placental regulatory evolution. While core miRNA families showed developmental pathway enrichment, the low conservation of specific predicted pathways across species - combined with placenta-associated miRNAs showing similar developmental associations to other miRNAs - suggests that placental evolution operates through distributed regulatory networks rather than discrete pathway switches. This pattern indicates that miRNAs quite likely influence placental phenotypes through coordinated fine-tuning of multiple regulatory circuits. Evolving such a mechanism may be more robust than relying on single developmental pathways, or the focussed activity of a single protein-coding gene. The extensive gene target repertoires, typically numbering in the thousands, are characteristic of miRNAs and make them ideally suited for such distributed regulatory control, potentially explaining both their evolutionary plasticity but also the predictable phenotypic outcomes we observe.

Conclusion

We demonstrate that placental phenotypes are predictable from miRNA gene repertoires across 300 mammal species, with convergent morphologies consistently involving the same miRNA families. This genomic predictability reveals developmental constraints that channel placental evolution down a small number of avenues, *via* both miRNA gene gain and loss events. Experimental validation of MIR-11968 in bovine caruncular tissue provides direct evidence linking clade-specific miRNA evolution to morphological innovation.

Where we have previously shown that miRNA evolution was important for the origin of placental mammals [13], the current findings establish miRNAs as key molecular architects of subsequent mammal reproductive diversity and provide a framework for understanding how genomic constraints shape complex trait evolution. Importantly, placental phenotypic diversity has not emerged through a single genetic master switch, but through this complex, multilayered regulatory landscape. It has been coordinated by a network of miRNAs - some ancient and persistent fixtures in mammal genomes since the origin of Eutheria, others displaying more variable patterns of presence and absence.

The evolution of placenta through miRNA regulation stands in contrast to the way that, say, Homeobox genes operate in development. Homeobox genes act as master switches, defining broad body plans with discrete, often dramatic effects [57,58], like positioning major body segments or

specifying cell lineages [59,60]. In contrast, miRNAs simultaneously fine-tune the expression of thousands of genes across complex regulatory networks. Where homeobox gene mutations can directly affect radical developmental disruptions, changes in miRNA repertoire clearly produce subtle modifications that manifest as alternative placental phenotypes. This means that miRNAs can generate nuanced phenotypic diversity through distributed, coordinated genetic adjustments across entire developmental landscapes. The outcome of miRNA repertoire evolution has been a sophisticated fine-tuning of placental development across thousands of species, both extant and extinct.

By looking across the breadth of extant mammal genomes we have been able to identify what works in terms of regulation of placental phenotype - something that wouldn't be possible by merely studying individual species etc. Evolution has tried and tested many more combinations of miRNAs etc than we observe but the ones we observe are the ones that were most successful … …

Acknowledgements

This work was supported by the Biotechnology and Biological Sciences Research Council (award numbers BB/X007332/1, BB/X007367/1 and BB/R017522/1) and FAPESP (2016/50433-9, 2016/22790-1, 2017/50438-3 and 2018/14137-1). MJO'C would like to thank the Leverhulme Trust for her personal fellowship to complete this work (RF-2024-492). We also thank the Vertebrate Genome Project, DNAzoo and Zoonomia for open access to their range of mammal genomes, and the high-performance

computing service at the University of Nottingham for access to computing facilities. BF acknowledges funding through the Tromsø forskningsstiftelse grant (TFS) [20_SG_BF 'MIRevolution'].

Author Contributions

MJO'C and NF conceived the project. JF designed and carried out all computational and statistical analyses with input from MJO'C and JOMcI. JCE, HT and ZB designed and performed *in vitro* analyses with input from NF. NF, FVM, JdS, GP, and TDB designed the *in vivo* experiment. HT, TDB, AB, IME, AGD, JdS, GP, FVM and NF performed the *in vivo* experiments. HT, IME, and JCE performed the *in vivo* sample analyses. OA collated endometrial miRNA expression datasets, VO prepared placental morphology data, MS contributed to generating the consensus tree for phylogeny-informed analyses, VO and BF provided technical assistance with MirGeneDB. MJO'C supervised all computational & evolutionary aspects of the work and NF all *in vitro* and *in vivo* aspects of the work. All authors contributed to writing the manuscript.

References

1. Wildman, D.E., Chen, C., Erez, O., Grossman, L.I., Goodman, M., and Romero, R. (2006). Evolution of the mammalian placenta revealed by phylogenetic analysis. Proc. Natl. Acad. Sci. U. S. A. *103*, 3203–3208. https://doi.org/10.1073/pnas.0511344103.

2. Enders, A.C., and Carter, A.M. (2012). The evolving placenta: Convergent evolution of variations in the endotheliochorial relationship. Placenta *33*, 319–326. https://doi.org/10.1016/j.placenta.2012.02.008.

3. Cross, J.C., Baczyk, D., Dobric, N., Hemberger, M., Hughes, M., Simmons, D.G., Yamamoto, H.,

and Kingdom, J.C.P. (2003). Genes, development and evolution of the placenta. Placenta *24*, 123–130. https://doi.org/10.1053/plac.2002.0887.

4. Haig, D. (2008). Placental Growth Hormone-Related Proteins and Prolactin-Related Proteins. Placenta *29*, 36–41. https://doi.org/10.1016/j.placenta.2007.09.010.

5. Bazer, F.W., Spencer, T.E., and Ott, T.L. (1997). Interferon tau: a novel pregnancy recognition signal. Am. J. Reprod. Immunol. N. Y. N 1989 *37*, 412–420. https://doi.org/10.1111/j.1600-0897.1997.tb00253.x.

6. Bazer, F.W., Spencer, T.E., Johnson, G.A., and Burghardt, R.C. (2011). Uterine receptivity to implantation of blastocysts in mammals. Front. Biosci.-Sch. *3*, 745–767. https://doi.org/10.2741/S184.

7. Denner, J. (2016). Expression and function of endogenous retroviruses in the placenta. APMIS Acta Pathol. Microbiol. Immunol. Scand. *124*, 31–43. https://doi.org/10.1111/apm.12474.

8. Lokossou, A.G., Toudic, C., and Barbeau, B. (2014). Implication of Human Endogenous Retrovirus Envelope Proteins in Placental Functions. Viruses *6*, 4609–4627. https://doi.org/10.3390/v6114609.

9. Clarke, A.W., Høye, E., Hembrom, A.A., Paynter, V.M., Vinther, J., Wyrożemski, Ł., Biryukova, I., Formaggioni, A., Ovchinnikov, V., Herlyn, H., et al. (2025). MirGeneDB 3.0: improved taxonomic sampling, uniform nomenclature of novel conserved microRNA families and updated covariance models. Nucleic Acids Res. *53*, D116–D128. https://doi.org/10.1093/nar/gkae1094.

10. Glazov, E.A., McWilliam, S., Barris, W.C., and Dalrymple, B.P. (2008). Origin, Evolution, and Biological Role of miRNA Cluster in DLK-DIO3 Genomic Region in Placental Mammals. Mol. Biol. Evol. *25*, 939–948. https://doi.org/10.1093/molbev/msn045.

11. Martins, M., Galfrè, S., Terrigno, M., Pandolfini, L., Appolloni, I., Dunville, K., Marranci, A., Rizzo, M., Mercatanti, A., Poliseno, L., et al. (2021). A eutherian-specific microRNA controls the translation of Satb2 in a model of cortical differentiation. Stem Cell Rep. *16*, 1496–1509. https://doi.org/10.1016/j.stemcr.2021.04.020.

12. Paikari, A., D. Belair, C., Saw, D., and Blelloch, R. (2017). The eutheria-specific miR-290 cluster modulates placental growth and maternal-fetal transport. Development *144*, 3731–3743. https://doi.org/10.1242/dev.151654.

13. Taylor, A.S., Tinning, H., Ovchinnikov, V., Edge, J., Smith, W., Pullinger, A.L., Sutton, R.A., Constantinides, B., Wang, D., Forbes, K., et al. (2023). A burst of genomic innovation at the origin of placental mammals mediated embryo implantation. Commun. Biol. *6*, 1–8. https://doi.org/10.1038/s42003-023-04809-y.

14. Luo, L., Ye, G., Nadeem, L., Fu, G., Yang, B.B., Honarparvar, E., Dunk, C., Lye, S., and Peng, C. (2012). MicroRNA-378a-5p promotes trophoblast cell survival, migration and invasion by targeting Nodal. J. Cell Sci. *125*, 3124–3132. https://doi.org/10.1242/jcs.096412.

15. Dai, Y., Qiu, Z., Diao, Z., Shen, L., Xue, P., Sun, H., and Hu, Y. (2012). MicroRNA-155 inhibits proliferation and migration of human extravillous trophoblast derived HTR-8/SVneo cells via down-regulating cyclin D1. Placenta *33*, 824–829. https://doi.org/10.1016/j.placenta.2012.07.012.

16. Anton, L., Olarerin-George, A.O., Schwartz, N., Srinivas, S., Bastek, J., Hogenesch, J.B., and Elovitz, M.A. (2013). miR-210 Inhibits Trophoblast Invasion and Is a Serum Biomarker for Preeclampsia. Am. J. Pathol. *183*, 1437–1445. https://doi.org/10.1016/j.ajpath.2013.07.021.

17. Li, P., Guo, W., Du, L., Zhao, J., Wang, Y., Liu, L., Hu, Y., and Hou, Y. (2012). microRNA-29b contributes to pre-eclampsia through its effects on apoptosis, invasion and angiogenesis of trophoblast cells. Clin. Sci. *124*, 27–40. https://doi.org/10.1042/CS20120121.

18. Vigorito, E., Kohlhaas, S., Lu, D., and Leyland, R. (2013). miR-155: an ancient regulator of the immune system. Immunol. Rev. *253*, 146–157. https://doi.org/10.1111/imr.12057.

19. Fromm, B., Høye, E., Domanska, D., Zhong, X., Aparicio-Puerta, E., Ovchinnikov, V., Umu, S.U., Chabot, P.J., Kang, W., Aslanzadeh, M., et al. (2022). MirGeneDB 2.1: toward a complete sampling of all major animal phyla. Nucleic Acids Res. *50*, D204–D210. https://doi.org/10.1093/NAR/GKAB1101.

20. Dudchenko, O., Batra, S.S., Omer, A.D., Nyquist, S.K., Hoeger, M., Durand, N.C., Shamim, M.S., Machol, I., Lander, E.S., Aiden, A.P., et al. (2017). De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. Science *356*, 92–95. https://doi.org/10.1126/science.aal3327.

21. Genereux, D.P., Serres, A., Armstrong, J., Johnson, J., Marinescu, V.D., Murén, E., Juan, D., Bejerano, G., Casewell, N.R., Chemnick, L.G., et al. (2020). A comparative genomics multitool for scientific discovery and conservation. Nature *587*, 240–245. https://doi.org/10.1038/s41586-020-2876-6.

22. Rhie, A., McCarthy, S.A., Fedrigo, O., Damas, J., Formenti, G., Koren, S., Uliano-Silva, M., Chow, W., Fungtammasan, A., Kim, J., et al. (2021). Towards complete and error-free genome assemblies of all vertebrate species. Nature *592*, 737–746. https://doi.org/10.1038/s41586-021-03451-0.

23. Armstrong, D.L., McGowen, M.R., Weckle, A., Pantham, P., Caravas, J., Agnew, D., Benirschke, K., Savage-Rumbaugh, S., Nevo, E., Kim, C.J., et al. (2017). The core transcriptome of mammalian placentas and the divergence of expression with placental shape. Placenta *57*, 71–78. https://doi.org/10.1016/j.placenta.2017.04.015.

24. Lewitus, E., and Soligo, C. (2011). Life-History Correlates of Placental Structure in Eutherian Evolution. Evol. Biol. *38*, 287–305. https://doi.org/10.1007/s11692-011-9115-x.

25. Umu, S.U., Paynter, V.M., Trondsen, H., Buschmann, T., Rounge, T.B., Peterson, K.J., and Fromm, B. (2023). Accurate microRNA annotation of animal genomes using trained covariance models of curated microRNA complements in MirMachine. Cell Genomics *3*, 100348. https://doi.org/10.1016/j.xgen.2023.100348.

26. Wang, P., and Wang, F. (2023). A proposed metric set for evaluation of genome assembly quality. Trends Genet. *39*, 175–186. https://doi.org/10.1016/j.tig.2022.10.005.

27. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

28. Team, R.C. (2018). R: A language and environment for statistical computing.

29. Tarver, J.E., dos Reis, M., Mirarab, S., Moran, R.J., Parker, S., O'Reilly, J.E., King, B.L., O'Connell, M.J., Asher, R.J., Warnow, T., et al. (2016). The Interrelationships of Placental Mammals and the Limits of Phylogenetic Inference. Genome Biol. Evol. *8*, 330–344. https://doi.org/10.1093/gbe/evv261.

30. Morgan, C.C., Foster, P.G., Webb, A.E., Pisani, D., McInerney, J.O., and O'Connell, M.J. (2013). Heterogeneous Models Place the Root of the Placental Mammal Phylogeny. Mol. Biol. Evol. *30*, 2145–2156. https://doi.org/10.1093/molbev/mst117.

31. Kumar, S., Suleski, M., Craig, J.M., Kasprowicz, A.E., Sanderford, M., Li, M., Stecher, G., and Hedges, S.B. (2022). TimeTree 5: An Expanded Resource for Species Divergence Times. Mol. Biol. Evol. *39*, msac174. https://doi.org/10.1093/molbev/msac174.

32. Hennig, C. (2024). fpc: Flexible Procedures for Clustering. R Package Version 22-13.

33. Oksanen, J., Simpson, G.L., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Solymos, P., Stevens, M.H.H., Szoecs, E., et al. (2024). vegan: Community Ecology Package.

34. Revell, L.J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. *3*, 217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x.

35. Lloyd, G.T. (2016). Estimating morphological diversity and tempo with discrete character-taxon matrices: implementation, challenges, progress, and future directions. Biol. J. Linn. Soc. *118*, 131–151. https://doi.org/10.1111/bij.12746.

36. Wright, M.N., and Ziegler, A. (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. J. Stat. Softw. *77*, 1–17. https://doi.org/10.18637/jss.v077.i01.

37. Yang, Y., and Jiang, X. (2023). Evolink: a phylogenetic approach for rapid identification of genotype–phenotype associations in large-scale microbial multispecies data. Bioinformatics *39*, btad215. https://doi.org/10.1093/bioinformatics/btad215.

38. Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for Computing and Annotating Genomic Ranges. PLOS Comput. Biol. *9*, e1003118. https://doi.org/10.1371/journal.pcbi.1003118.

39. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

40. Agarwal, V., Bell, G.W., Nam, J.-W., and Bartel, D.P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. eLife *4*, e05005. https://doi.org/10.7554/eLife.05005.

41. Yang, T., Wang, Y., and He, Y. (2024). TEC-miTarget: enhancing microRNA target prediction based on deep learning of ribonucleic acid sequences. BMC Bioinformatics *25*, 159. https://doi.org/10.1186/s12859-024-05780-z.

42. Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS J. Integr. Biol. *16*, 284–287. https://doi.org/10.1089/omi.2011.0118.

43. Sayols, S. (2023). rrvgo: a Bioconductor package for interpreting lists of Gene Ontology terms. MicroPublication Biol. https://doi.org/10.17912/micropub.biology.000811.

44. De Bem, T.H.C., Bridi, A., Tinning, H., Sampaio, R.V., Malo-Estepa, I., Wang, D., Vasconcelos, E.J.R., Nociti, R.P., Ávila, A.C.F.C.M. de, Rodrigues Sangalli, J., et al. (2024). Biosensor capability of the endometrium is mediated in part, by altered miRNA cargo from conceptus-derived extracellular vesicles. FASEB J. *38*, e23639. https://doi.org/10.1096/fj.202302423RR.

45. Tinning, H., Taylor, A., Wang, D., Constantinides, B., Sutton, R., Oikonomou, G., Velazquez, M.A., Thompson, P., Treumann, A., O'Connell, M.J., et al. (2020). The role of CAPG in molecular communication between the embryo and the uterine endometrium: Is its function conserved in species with different implantation strategies? FASEB J. *34*, 11015–11029. https://doi.org/10.1096/fj.202000882RR.

46. Edge, J.C., Amelkina, O., Tinning, H., Giovanardi, G., Mancinelli, E., Gardner, S., Vasconcelos, E.J., Pensabene, V., Forbes, K., O'Connell, M.J., et al. (2024). Bovine endometrial organoids: A new tool to study conceptus-maternal interactions in mammals. Preprint at bioRxiv, https://doi.org/10.1101/2024.12.13.628419 https://doi.org/10.1101/2024.12.13.628419.

47. miRBase entry: bta-mir-11986 https://mirbase.org/hairpin/MI0038209.

48. Carlson, M. (2023). org.Bt.eg.db: Genome wide annotation for Bovine.

49. Duquette, P.M., and Lamarche-Vane, N. (2014). Rho GTPases in embryonic development. Small GTPases *5*, e972857. https://doi.org/10.4161/sgtp.29716.

50. Zhao, Z., and Rivkees, S.A. (2000). Tissue-specific expression of GTPas RalA and RalB during embryogenesis and regulation by epithelial-mesenchymal interaction. Mech. Dev. *97*, 201–204. https://doi.org/10.1016/S0925-4773(00)00427-5.

51. Orozco-Hernández, E., Clemente-Guerrero, M., Saavedra-Molina, A., and Manzo-Avalos, S. (2014). Activity state of the RhoA GTPase and expression of effector molecules in placentas of women with preeclampsia (802.28). FASEB J. *28*, 802.28. https://doi.org/10.1096/fasebj.28.1_supplement.802.28.

52. Blackburn, D.G. (2015). Evolution of vertebrate viviparity and specializations for fetal nutrition: A quantitative and qualitative analysis. J. Morphol. *276*, 961–990. https://doi.org/10.1002/jmor.20272.

53. Griffith, O.W., and Wagner, G.P. (2017). The placenta as a model for understanding the origin and evolution of vertebrate organs. Nat. Ecol. Evol. *1*, 1–10. https://doi.org/10.1038/s41559-017-0072.

54. Roberts, R.M., Green, J.A., and Schulz, L.C. (2016). The Evolution of the Placenta. Reprod. Camb. Engl. *152*, R179–R189. https://doi.org/10.1530/REP-16-0325.

55. Van Dyke, J.U., Brandley, M.C., and Thompson, M.B. (2014). The evolution of viviparity: Molecular and genomic data from squamate reptiles advance understanding of live birth in amniotes. Reproduction *147*, R15–R26. https://doi.org/10.1530/REP-13-0309.

56. Tarver, J.E., Taylor, R.S., Puttick, M.N., Lloyd, G.T., Pett, W., Fromm, B., Schirrmeister, B.E., Pisani, D., Peterson, K.J., and Donoghue, P.C.J. (2018). Well-Annotated microRNAomes Do Not Evidence Pervasive miRNA Loss. Genome Biol. Evol. *10*, 1457–1470. https://doi.org/10.1093/gbe/evy096.

57. Gehring, W.J. (1993). Exploring the homeobox. Gene *135*, 215–221. https://doi.org/10.1016/0378-1119(93)90068-E.

58. Gehring, W.J., and Hiromi, Y. (1986). Homeotic genes and the homeobox. Annu. Rev. Genet. *20*, 147–173. https://doi.org/10.1146/annurev.ge.20.120186.001051.

59. Kongsuwan, K., Webb, E., Housiaux, P., and Adams, J.M. (1988). Expression of multiple homeobox genes within diverse mammalian haemopoietic lineages. EMBO J. *7*, 2131–2138. https://doi.org/10.1002/j.1460-2075.1988.tb03052.x.

60. McGinnis, W., and Krumlauf, R. (1992). Homeobox genes and axial patterning. Cell *68*, 283–302. https://doi.org/10.1016/0092-8674(92)90471-N.

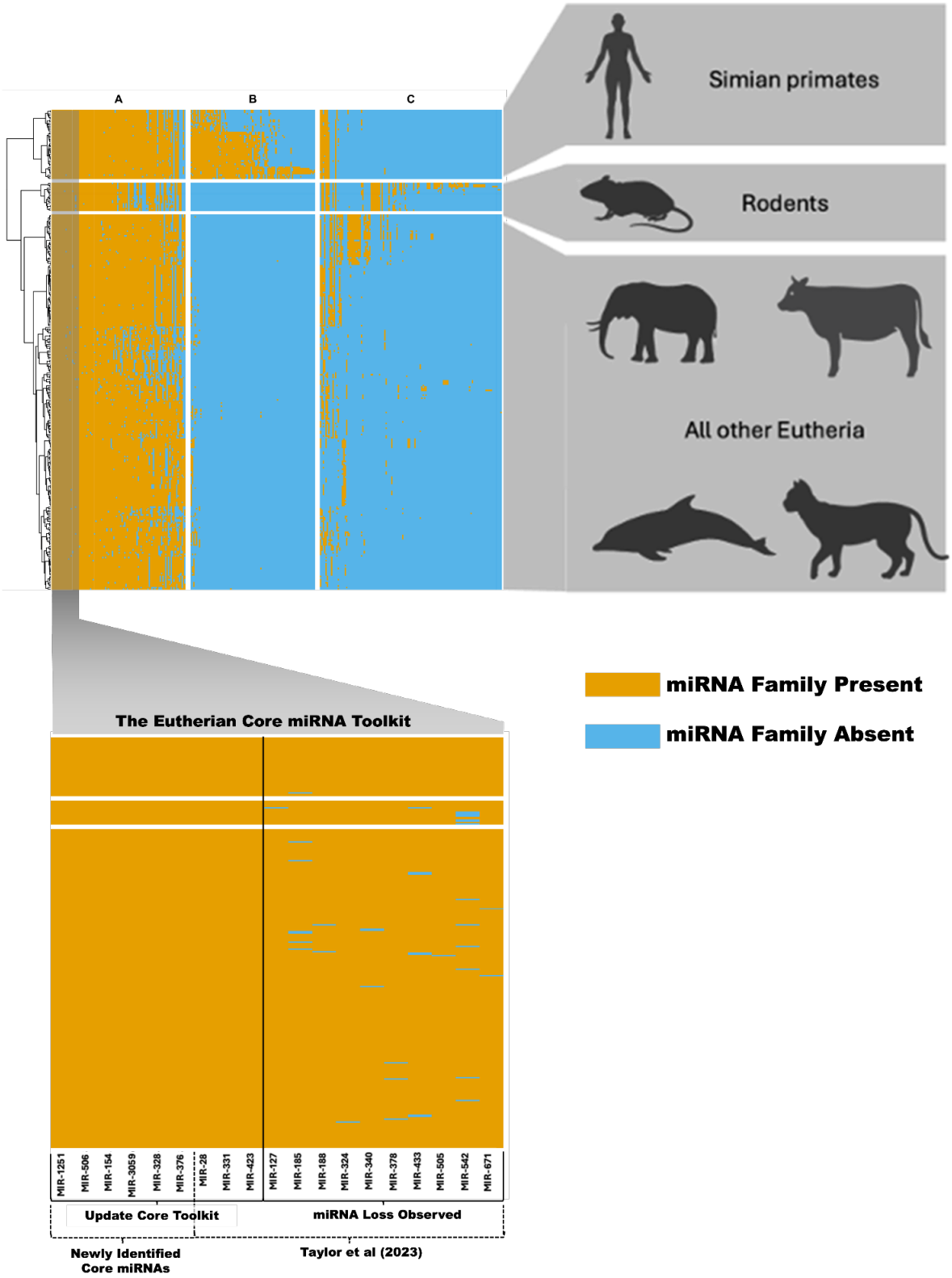

Figure 1. The structure of miRNA family presence and absence in Eutherian mammals is defined by distinct clusters and a core toolkit of shared miRNA families. TOP - Heatmap of miRNA family presence (yellow) and absence (blue) across the Eutherian tree, with rows representing individual

species (n=300) and columns representing miRNA families (n=322) with miRNA presence being detected with BLAST searches from mirGeneDB databases and MirMachine analysis. Species formed stable clusters (n=3) according to Jaccard dissimilarity of miRNA repertoires, with discrete clusters comprising simian primates (top cluster) and rodents (middle cluster). The bottom custer comprises all other Eutherian mammals. miRNA families also formed 3 discrete clusters, with cluster A showing no strong taxonomic signal, but clusters B and C being overrepresented in simian primates and rodents, respectively. BOTTOM - Subset of the miRNA presence/absence heatmap showing updates to the Core Eutherian miRNA toolkit as described in Taylor et al (2023). The 9 leftmost columns show miRNA families found in all Eutherian species that were checked, with 3 of these (Mir-28, -331 & -423) being carried over from Taylor et al. The 10 rightmost columns show miRNA families previously purported to be present in every Eutherian species, but in which we have found rare losses.

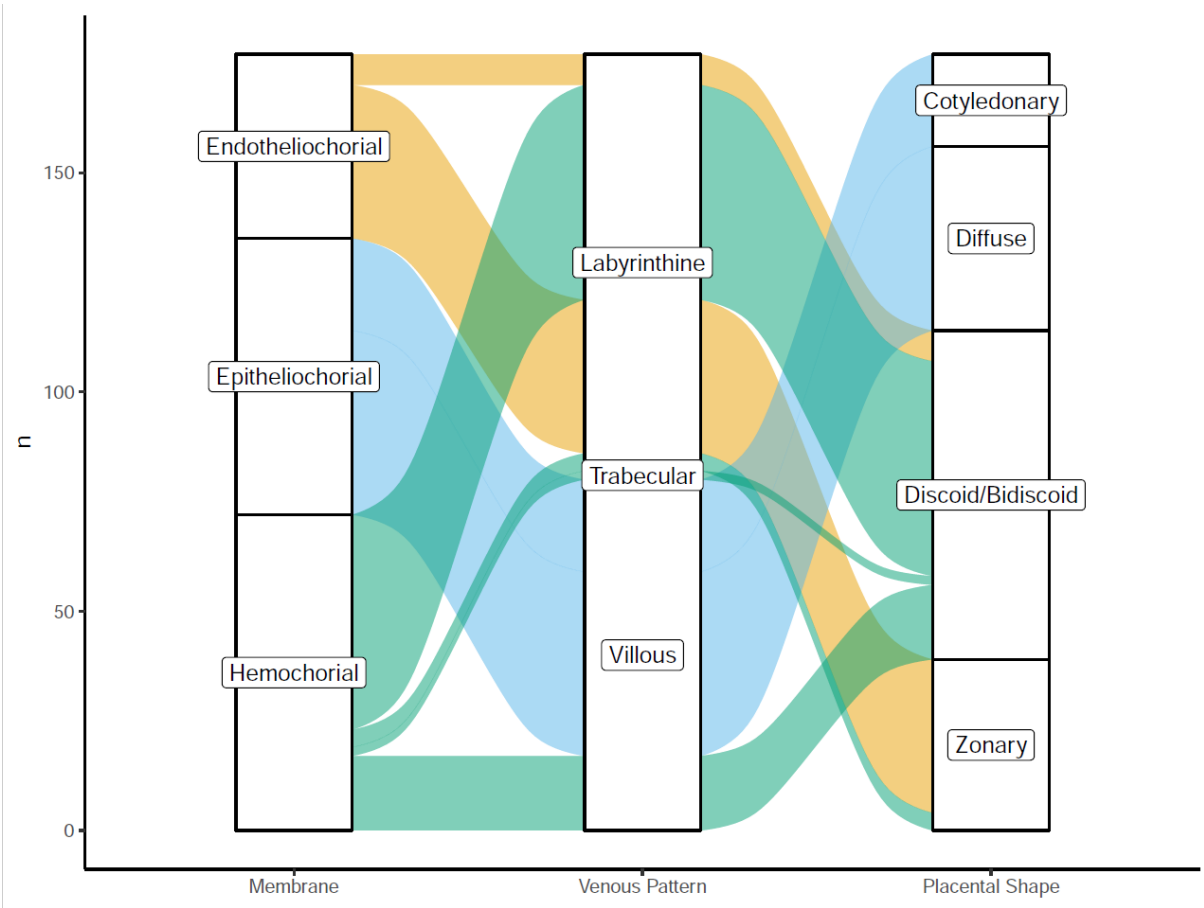

Figure 2. Associations between placental phenotypes are non-random. Alluvial plot showing non-random association between placental phenotype states (loglinear p~0). Endotheliochorial

membranes (yellow) co-occur with labyrinthine venous systems, and zonary & discoid placenta shape. Epitheliochorial membranes (blue) co-occur with villous venous systems, and diffuse & cotyledonary placenta shapes. Hemochorial membranes (green) co-occur with all venous system, and discoid & zonary placenta shapes.

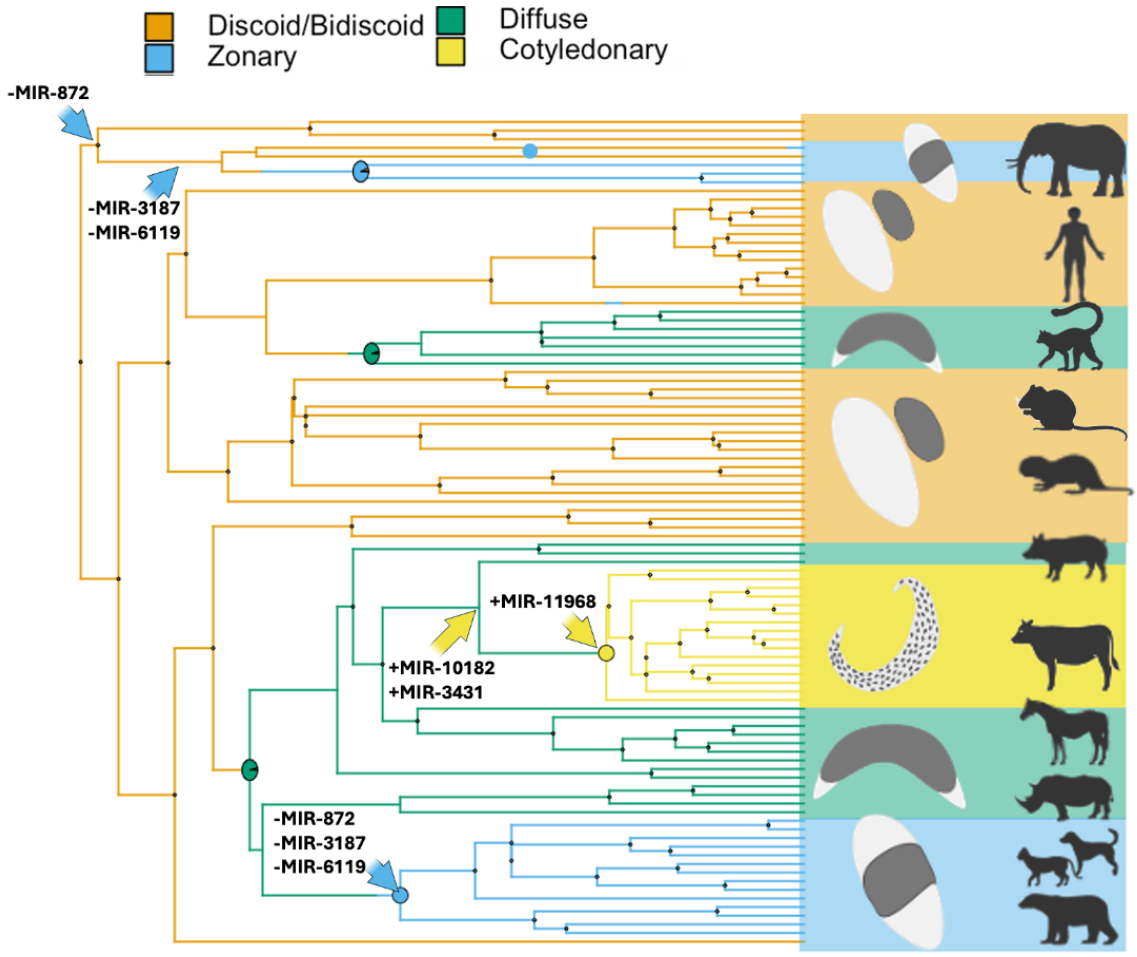

Figure 3. miRNA gene family gain and loss associated with innovation in placental shape across Eutheria. Ancestral placental shape phenotypes, inferred from 100 MCMC reconstructions, are plotted on an underlying mammalian phylogeny (left). Images on right of the tree show placental shapes in each clade, and silhouette of representative species (taken from BioRender [1]). Gain and loss of strongly phenotype-associated miRNA families is shown (denoted with '+' and '-' respectively), with arrows pointing to the branch upon which they emerged, and arrow colours corresponding to the phenotype state with which they are associated. MIRs-11968, -10182 and -3431 are strongly

associated by their gain to the cotyledonary placenta shape, present only in ruminants of infraorder *Pecora*. MIRs-872, -3187 and -6119 are present in the eutherian ancestor, but are lost in both ancestors of clades which developed zonary placentas (note that these families show additional losses in other clades, not shown here).

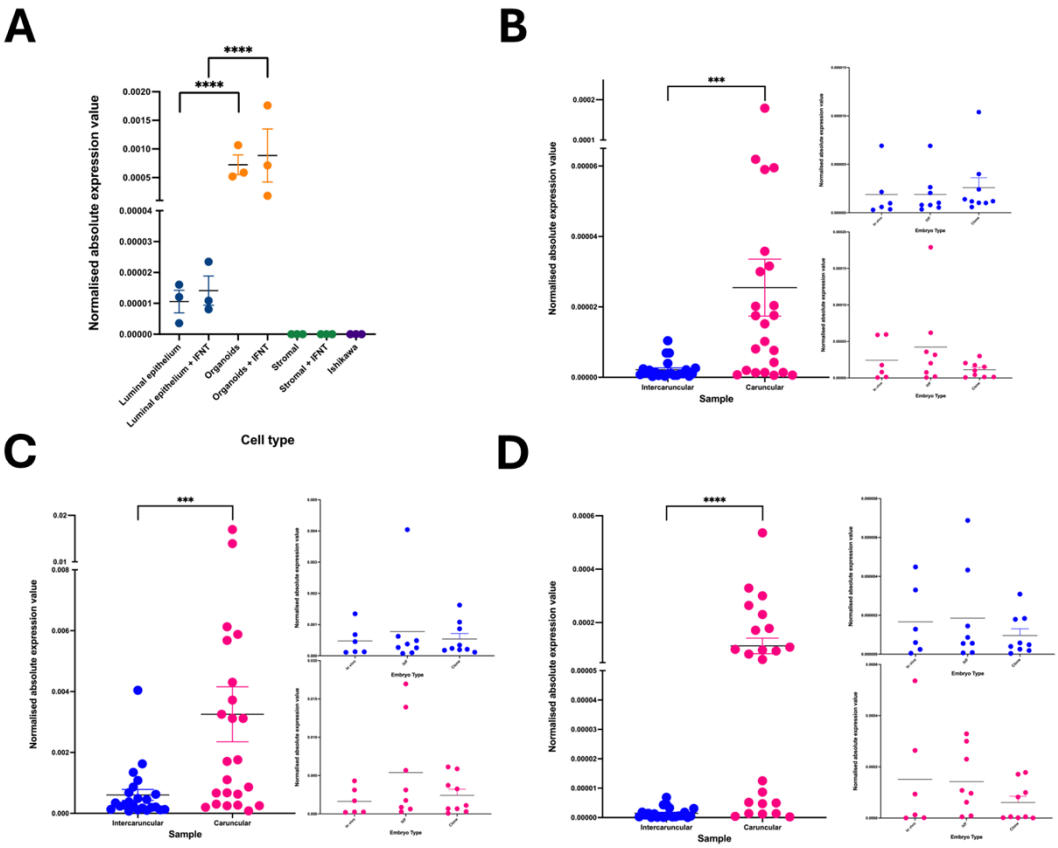

Figure 4. MIR-11968, showing complete evolutionary association with cotyledonary placentas, shows species- and tissue-specific expression in bovine endometrium. Function of the cotyledonary placenta-associated MIR-11968 paralogues were assessed through *in vitro* and in *vivo* analysis. (A) Mir-11968-P1 expression in different in vitro models of the endometrium. Expression of Mir-11968-P1 was measured by qRT-PCR in primary bovine luminal epithelial cells (n=3), bovine glandular epithelial organoids (n=3) and bovine stromal cells (n=3) and in each cell type treated with 1000ng/µl recombinant ovine IFNT for 24 hours (n=3). Mir-11968-P1 expression was also measured in human endometrial epithelial (Ishikawa) cells (n=3) as a negative control. Significant differences where

p<0.0001 are denoted by **** (B-D) MiR-11986, Mir-11968-P1 and Mir-11968-P2 expression in bovine caruncular or intercaruncular endometrial tissue. Expression of miR-11986 (B), Mir-11968-P1 (C) and Mir-11968-P2 (D) was measured by qRT-PCR in bovine caruncular (n=25) and intercaruncular (n=25) endometrial samples on Day 16 of pregnancy. Summary lines in all boxplots represent mean values, error bars represent standard error of the mean. Significant differences in expression where p<0.01 are denoted by ***, and where p<0.001 are denoted by ****.

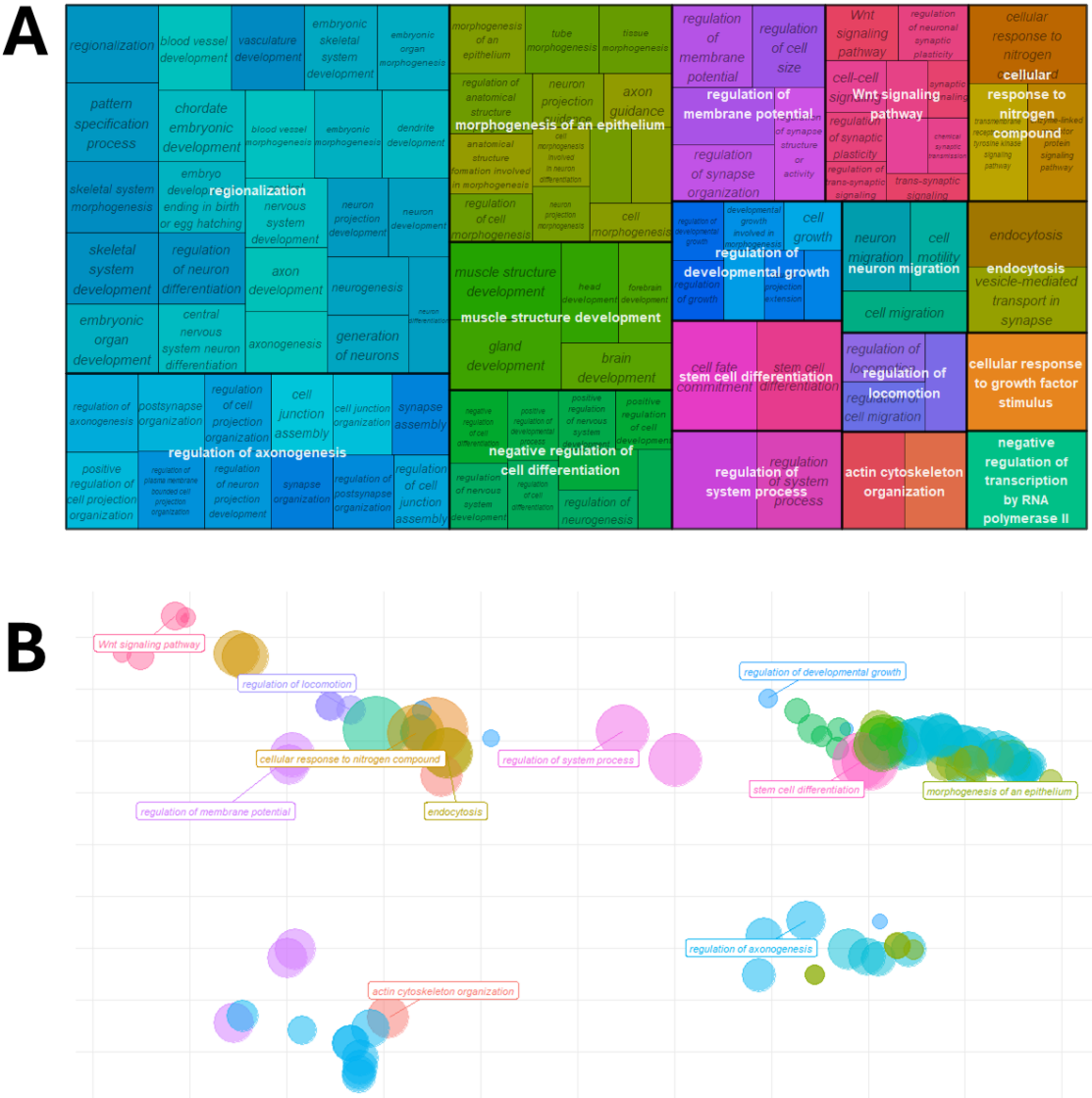

Figure 5. Core Eutherian miRNA families have a diversity of gene targets conserved across species, including development processes. A) Clustered treemap plot and **B)** scatter map plot from rrvgo summary of the top 100 most-conserved biological-process level gene ontology terms, clustered by term similarity. GO terms were derived from UTR gene targets of synapomorphic 'core' Eutherian miRNA gene families, as predicted from Targetscan and TEC-miTarget. Most genes consistently targeted by core eutherian miRNA families across species are implicated in developmental, morphogenic and regionalisation processes, suggesting a crucial role for conserved miRNA regulatory networks in eutherian development.

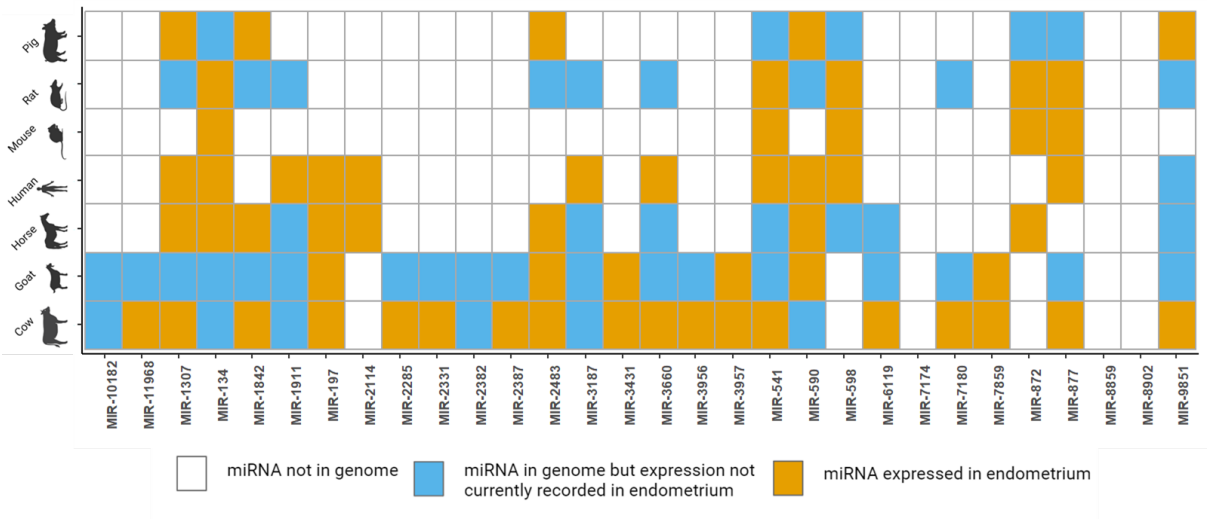

Figure 6. miRNA families associated with phenotype are commonly expressed in the endometrium across Eutheria. Heatmap of 7 eutherian species, showing where phenotype-associated miRNAs are recorded as being expressed in the endometrium, when present in the genome, as compiled from published databases (Table S9).