Skolkovo Institute of Science and Technology

USING RNA EXPRESSION AS A QUANTITATIVE MOLECULAR PHENOTYPE TO STUDY HUMAN AND VERTEBRATE EVOLUTION

Doctoral Thesis

by

SONG GUO

DOCTORAL PROGRAM IN LIFE SCIENCES

Supervisor
Professor Philipp Khaitovich

Moscow - 2021

I hereby declare that the work presented in this thesis was carried out by myself at Skolkovo Institute of Science and Technology, Moscow, except where due acknowledgment is made, and has not been submitted for any other degree.

Song Guo

Prof. Dr. Philipp Khaitovich

Abstract

Different phenotypes bring diversity to the entire living world and are a prerequisite for evolution and natural selection. Why does each individual of a species with a very similar genome have such different appearances? Why do embryos have such similarities among species in a phylum? The mysterious molecular mechanism underlying phenotype evolution has been studied for many decades. A crucial question is what major molecular force drives the enormous changes in evolutionary history? Although genome-based studies have shown fruitful results, there are still several limitations. In addition to the genome, the expressions of protein-coding genes and regulatory transcripts, such as microRNAs, as informative endophenotypes, have been now used in studies of human and vertebrate evolution. In my Ph.D. work, I applied high-throughput RNA sequencing (RNA-seq) technology to study the transcriptome features in the evolution of embryonic development of vertebrate species and the emerging diversity of human populations, based on expression profiling of actual tissues. Two projects in my work are discussed according to the different evolutionary phenotypic changes.

Human populations display a number of phenotypic differences formed as a result of development. Analysis of lymphocyte cell lines has revealed significant differences in mRNA and microRNA (miRNA) expression between human populations. However, the extent of such population-associated differences in actual human tissues remains largely unexplored. Our previous study reported mRNA expression differences in 6.3% of expressed genes between placental samples from four populations: African Americans, European Americans, South Asian Americans, and East Asian Americans. In the study I report here, we further designed a work to exam the microRNA (miRNA) expression in the same placental samples and to check how different demographic variables affect it. The results of my analysis showed that among the twelve demographic variables assessed, two – population identity and sex of the newborn – significantly influenced the variation in miRNA expression in human placental samples. Based on further analyses, I observed that the miRNA effects associated with the sex of the newborn child further led to expression

inhibition of target genes in specific functional pathways. By contrast, analysis of target expression patterns indicated that population-associated miRNA differences might mainly represent neutral changes with minimal functional effects.

Embryonic development is very crucial for mature organisms. Any small aberrant change could lead to abnormal development. In contrast to the diversity of adult phenotypes between species, embryos in different animals have similar morphological characteristics in early development. Von Baer and Ernst Haeckel laid cornerstones for current evolutionary views on developmental programs by formulating two different paradigms of developmental stages' relationship among species. While von Baer did not accept the concept of species' evolution, his concepts were actively used in building evolutionary models of development, such as commonly used funnel and hourglass models. In modern biology, studies of the long debated relationship between embryonic development and evolution (Evo-devo) have shifted from the morphology of the embryos to using mRNA expression as the endophenotype. However, many such studies have been focusing on limited number of species. Our work aimed to extend this avenue of research by comparing developmental profiles of different species in a single taxonomic group. Our previous work has confirmed the applicability of the funnel model and indicated the coexistence of funnel and hourglass models at the level of gene expression in vertebrates. In my Ph.D. thesis, I aimed to extend our knowledge on the evolution of developmental processes to investigate possible functions of consistent and inconsistent embryonic developmental processes among vertebrates at the mRNA and miRNA levels. To this end, my work included two main sub-projects. In the first study, I compared developmental changes among nine evolutionarily distinct species: from oyster to mouse at mRNA expression levels across the bulk of whole embryos. I found that up to 34% of mRNA expression demonstrated near-linearly decreasing profiles across species' developmental programs, with the corresponding genes being enriched in spliceosome function and a group of housekeeping genes. By contrast, genes involved in the annotated developmental biological process did not show significantly co-expressed profiles between species. It

suggested that differences in tissue composition or differences in expression of regulators during development between species could be possible reasons.

Therefore, I extended my work to a study of small regulatory RNAs, miRNAs, and their regulatory relationship with organ development by comparing miRNA expression of eight vertebrates, ranging from human to chicken, at five developmental stages at the brain, heart, and liver. The results showed that less than half of miRNAs express in all three tissues, with the brain having more organ-specific expressed miRNAs. 35%~68% of miRNAs showed dynamic changes in at least one organ for each species. Further, miRNAs consistently expressed in the three organs were conserved between species and inhibited the target genes for cell cycle and protein localization to the membrane. By contrast, miRNAs expressed in three organs with different patterns were not conserved between species but showed regulatory function on the targets in organ development in the manner of a species-specific process.

In summary, these two projects, one focusing on the role of miRNA in driving recent evolutionary divergence of human populations, another including two studies, aimed at a better understanding of evolutionary conservation of mRNA developmental profiles among vertebrates and role of miRNAs in the regulation of this divergence and conservation, have expanded our understanding of transcriptome diversities in the evolution of embryonic development and the resulting consequences between populations. The results may provide clues to the links between molecular changes and complex traits, filling gaps in actual tissue studies between human populations and provoking another molecular layer to explore the relationship between development and evolution.

Keywords:
Phenotype; Gene expression; microRNA; Evolution; Evo-devo; Human population

Publications

1. **Guo, S**., S. Huang, X. Jiang, H. Hu, D. Han, C. S. Moreno, G. L. Fairbrother, D. A. Hughes, M. Stoneking and P. Khaitovich (2021). *Variation of microRNA expression in the human placenta driven by population identity and sex of the newborn.* **BMC Genomics** 22(1): 286.

2. Khrameeva, E., Kurochkin, I., Han, D., Guijarro, P., Kanton, S., Santel, M., . . **Guo S**, Khaitovich, P. (2020). *Single-cell-resolution transcriptome map of human, chimpanzee, bonobo, and macaque brains*. **Genome Res**, 30(5), 776-789. doi:10.1101/gr.256958.119

3. Mitina, A., Mazin, P., Vanyushkina, A., Anikanov, N., Mair, W., **Guo, S**., & Khaitovich, P. (2020). *Lipidome analysis of milk composition in humans, monkeys, bovids, and pigs.* **BMC Evol Biol**, 20(1), 70. doi:10.1186/s12862-020-01637-0

4. **Song Guo**, Haiyang Hu, Chuan Xu, Naoki Irie, Philipp Khaitovich
Comparative gene analysis in chordate embryonic development reveals co-existence of Von Baer's and Haeckel's postulates. **Communication biology (under review)**

Acknowledgments

I would like to express my special thanks to my supervisor, Prof. Dr. Philipp Khaitovich, for supervising my research and doctoral studies. I joined Prof. Philipp Khaitovich 's lab in 2007 when I was a Junior Research Assistant in his lab at CAS-MPG Partner Institute for Computational Biology, Shanghai Institute for Biological Sciences, CAS. Since 2017, I have been a Ph.D. student in his lab at Skolkovo Institute of Science and Technology. During this time, I have successfully developed into a computational biologist. I would also like to thank all the members of the dissertation committee. They sacrificed their precious time to review my dissertation, including Prof. Dr. Mikhail Gelfand and Prof. Dr. Dmitri Pervouchine.

I would like to express my sincere appreciation to all my former and current members at PICB for their genuine kindness, valuable discussions, and maintaining a lively atmosphere in which to do science. All of them are very valuable to me. Among them, I would like to express special thanks to Dr. Haiyang Hu and Dr. Chuan Xu for collaborating on my research studies and sharing their knowledge and scientific ideas with me, Xi Jiang, and Shuyun Huang for conducting experiments in my projects.

I am very grateful to Prof. Naoki Irie from the University of Tokyo for material support and scientific discussion on the project Evo-Devo. I would like to thank Dr. David A. Hughes and Prof. Mark Stoneking from Max Planck Institute for Evolutionary Anthropology for moral and material support on the human placenta project. I would like to thank Dr. Margarida Cardoso Moreira and Professor Henrik Kaessmann from the Center for Molecular Biology of Heidelberg University for their idea, supervision, and data support on the miRNA organ evolution project.

During my studies at Skoltech, I shared many wonderful times with all my colleagues and friends there. Thank you to Skoltech and PICB, whose financial support allowed me to complete my studies and research.

Finally, I am incredibly grateful for the support of my family. My father, a scientist, has always advised me to be patient in any scientific research. My sister, who advised me to study bioinformatics, has always encouraged me in my studies. My mother and niece have always made me happy.

Table of Contents

List of Figures

List of Tables

List of Supplementary Tables

List of Symbols, Abbreviations

miRNA– microRNA

BMI– Body mass index

SON– Sex of newborn

RPM– Reads per million

RNA-seq– RNA sequencing

ncRNA– non-coding RNA

ANOVA– analysis of variance.

MYA– Million years ago

C19MC– Chromosome 19 miRNA cluster

Evo-devo – Evolutionary developmental biology

TF – Transcriptional factor

NGS – Next generation sequencing

Chapter 1. Introduction

Phenotype describes the characteristics of an organism, such as morphology, development, or behavior. Each organism, individual, or cell type with a different spatiotemporal location has a distinct phenotype from a single cell to the entire organism. For example, monozygotic twins never have precisely the same phenotype but share the same genotype; they can have a significant difference in appearance if they live in different environments. The morphology of synapses in the brain varies across the lifespan, with the composition of synapses changing in all brain regions (Cizeron et al., 2020). Moreover, embryos show a high degree of morphological similarity across the animal kingdom, especially early in development, despite genetic and environmental differences. Such variation in phenotype gives rise to the overall diversity of the living world and is a prerequisite for evolution and natural selection. The mystery underlying development of the phenotype has been studied for many generations. Understanding the molecular basis of phenotypic divergence is essential and can help us elucidate evolution, physiology, and disease.

Many studies have shown that various genetic alterations can influence complex traits in animals and plants. For example, single nucleotide polymorphisms (SNP) associated with human body mass index (BMI) (Locke et al., 2015), body size (Wood et al., 2014) and blood lipid levels (Willer et al., 2013); copy number variations (CNVs) contributed to evolutionary processes underlying domestication, such as fruit size and disease resistance (Lye & Purugganan, 2019). However, only a small fraction of genetic contribution has been proven to associate with specific traits from hundreds of identified common genetic loci in genome-wide association studies (GWAS) (Visscher, Brown, McCarthy, & Yang, 2012). Most genetic changes in non-coding regulatory elements affecting gene expression changes have profound phenotypic implications (Levchenko, Kanapin, Samsonova, & Gainetdinov, 2018).

As a crucial molecular phenotype, gene expression converts molecular information from the genome into a functional gene product. It represents the result of genetic changes, epigenetic regulation, transcriptional and post-transcriptional regulation, and other molecular processes that might be strongly responsive to environmental factors. Gene expression describes the characteristic properties of cell types, tissues, individuals, and organisms in different environments. Therefore, gene expression is a direct endophenotype that correlates with phenotypes, although the evolution of gene expression in both primates and model organisms is likely under stabilizing selection (Gilad, Oshlack, & Rifkin, 2006). Nonetheless, an increasing number of comparative gene expression studies in animals and plants have led to many important insights into development, aging, disease, and evolution in recent decades.

One common type of essential post-transcriptional regulators is miRNA: short (~22 nucleotides) RNA molecules, derived from hairpin-shaped precursors, that binds to complementary sequences in the 3'-UTR of target mRNAs. It regulates gene expression by degrading target mRNAs or inhibiting their translation (Lewis, Burge, & Bartel, 2005; Lewis, Shih, Jones-Rhoades, Bartel, & Burge, 2003). miRNAs are involved in numerous normal and pathological cellular processes. Not surprisingly, it is now increasingly recognized that variations in miRNAs and their target genes' expression levels may also contribute to phenotypic variability. For instance, a study comparing miRNA expression in humans, chimpanzees, and rhesus macaques found that 13 human-specific changed miRNAs and their target genes were associated with neuronal function (H. Y. Hu et al., 2011). miRNA expression changes in their spatiotemporal pattern suggested that the changes may be an essential factor in the evolution of phenotypic diversity in vertebrates and invertebrates (Aboobaker, Tomancak, Patel, Rubin, & Lai, 2005; Krichevsky, King, Donahue, Khrapko, & Kosik, 2003; Leong, Abdullah, Ling, & Cheah, 2016; Rahmanian et al., 2019).

An obvious question is: What kind of evolutionary force may influence the phenotypic variation in human populations and vertebrate evolution? To answer this

question, we must first figure out the relationships between molecular phenotypes and morphological differences. Scientists have studied this at the genomic, transcriptomic, proteomic, and metabolic levels (Tkachev et al., 2019). However, the transcriptome-level approach has higher coverage, resolution, and lower cost than current mass spectrum quantification technologies and, therefore, is the first choice for us to study the evolution of coding and non-coding RNA expression patterns systematically and comprehensively. In my Ph.D. projects, I proposed to perform mRNA and microRNA expression comparisons on bulk tissues using NGS sequencing to improve our understanding of the molecular mechanisms of human and vertebrate evolution. Specifically, I first introduce my research on determining miRNA expression variation in placenta between human populations and the regulatory effects of population-associated miRNA expression in combination with mRNA expression data obtained from the same tissue in our previous study. And then, I introduce my second research, where I describe the relationship between embryonic development and evolution by examining mRNA and miRNA expression profiles on vertebrates and then attempting to explore the functionality of gene expression change from the perspective of miRNA regulation. In the following part, I will specifically present the research status of the human population and embryonic developmental evolution studies based on mRNA and miRNA.

1.1 Quantitative studies on the evolution of human populations

Phenotypic differences among humans can be attributed to the combined effects of genetic, epigenetic, and environmental factors. Human populations may differ in skin pigmentation, facial shape, disease susceptibility (Hemming et al., 2011; Wright Willis, Evanoff, Lian, Criswell, & Racette, 2010), environmental adaptability, microbiota composition (K. Li et al., 2016; L. Li & Zhao, 2015), brain anatomical organization (Bai et al., 2012; Tang et al., 2018), and so on. Any trait fixed in a population must undergo the long-term process of evolution to change gene frequencies within a gene pool.

The genetic basis for phenotypic variation in human populations has been extensively studied. Previous studies have identified many genetic variants associated with

human population-specific phenotypic traits or differential disease susceptibility, including differences in the frequency of single nucleotide polymorphisms (SNPs), copy number variations (CNVs), transposable elements (TEs), and DNA methylation (Armengol et al., 2009; de Magalhaes & Matsuda, 2012; Genomes Project et al., 2015; J. Guo et al., 2018; Husquin et al., 2018; McCarroll & Altshuler, 2007; Stranger et al., 2007; L. Wang, Rishishwar, Marino-Ramirez, & Jordan, 2017). However, some genetic variations may not affect fitness or phenotypic variation. 1000 genome project has found around 12,000 polymorphic variants in protein-coding region and up to 565,000 variants in the sites of known regulatory regions per genome. Only around 2,000 variants are associated with complex traits or rare diseases in European ancestry samples, but not in other populations (Genomes Project et al., 2015; O'Huallachain, Karczewski, Weissman, Urban, & Snyder, 2012).

In addition to genomic analyses, gene expression was used to assess the relative influence of genotype, environment, and non-neutral evolution on phenotypic variation. However, most of our current knowledge of expression variation among human populations comes from lymphoblastoid cell lines (LCLs) derived from individuals representing different populations, rather than native tissues (Daca-Roszak et al., 2018; Price et al., 2008; Spielman et al., 2007; Storey et al., 2007; W. Zhang et al., 2008). Gene expression profiles of cell lines, however, differ drastically from those of the tissues. Nonetheless, several such studies focused on mRNA expression in LCLs showed that 4.5-29% of the expressed genes were differentially expressed between human populations, including Europeans (CEU), Yoruba from sub- Saharan Africa (YRI), and two East Asian populations: Han Chinese (CHB) and Japanese (JPT). Specifically, 17% of expression variation was observed between CEU and YRI (Storey et al., 2007). A similar estimate of 12% expression variation was observed between CEU and YRI (Price et al., 2008). However, when non-genetic factors were considered, only 5% could be attributed to the population difference. Another similar result yielded a mean estimate of 0.7% of population divergence between CEU and YRI samples (Stranger et al., 2012). A recent study on five populations (89-95 individuals per populations), including CEPH (CEU),

Finns (FIN), British (GBR), Toscani (TSI), and Yoruba (YRI), reported 2% to 26% genes with differential expression between population pairs, requiring genes with FDR<0.05 and log2 fold change greater than two (Lappalainen et al., 2013). A multiple-omics study for three Southeast Asian populations on whole blood showed that 280 transcriptional probes, 107 lipids, and 5 miRNAs were identified to be differentially expressed between Chinese, Malays, and Indians (Saw et al., 2017).

Our previous study, focusing on placenta rather than blood-derived cell lines, reported mRNA expression differences in 7.8% of expressed genes between placental samples collected from the same hospital from four populations: African Americans, European Americans, South Asian Americans, and East Asian Americans (Hughes et al., 2015). Among these, most placental expression variation was consistent with a neutral model of evolution, however, still 35% of the differences were considered as potentially affected by positive selection. The significantly differentially expressed genes were enriched in biological pathways associated with immune response, cell signaling, and metabolism. So far, we have observed the variation of gene expression between populations, but we still do not know the potential molecular force that drives the population difference. As one of the essential post-transcriptional regulators, miRNA has been studied to a limited extent in human population evolution. To extend our study, we want to understand whether most miRNA expression variation is as neutral as genes or whether miRNA expression variation could be one of the driving forces for human population evolution.

Differences in miRNA expression between human populations were previously examined using LCLs derived from CEU and YRI individuals; this study showed population-associated expression differences for 33 out of 757 detected miRNAs, resulting in 55-88% down-regulation of their expressed target genes (Huang et al., 2011). Cancer studies examining circulating miRNA abundance also showed differences between African and non-African descent (Rawlings-Goss, Campbell, & Tishkoff, 2014; H. Zhao et al., 2010). In a study on miR-375 in plasma, differences in expression level and its CpG

methylation in the promoter region were found between Kazak and Han populations, contributing to the occurrence of type 2 diabetes (Chang et al., 2014).

However, few studies have been performed on native human tissues for both mRNA and miRNA due to limited sample collection. As a result, the extent of population-associated differences in natural human tissues remains largely unexplored. For example, what is the extent of population-associated differences in endophenotype in different tissues at different molecular levels? How significant is the contribution of trans-regulators to population diversity?

To attempt to address these questions and to understand the basic mechanisms of human variability and regulation in bulk tissues, in my Ph.D. project, I extended our previous mRNA study by analyzing miRNA expression in human placental samples from 36 individuals representing four genetically distinct human populations: African Americans, European Americans, South Asian Americans, and East Asian Americans, to explore the factors that determine miRNA expression variation and the regulatory effects induced by population-associated miRNA. The placenta is one of the few solid human tissues collected in substantial numbers in a controlled manner, allowing quantitative analysis of transient biomolecules such as RNA transcripts. The results are described in Chapter 3.

1.2 Quantitative studies in evolutionary developmental biology

Evolutionary developmental biology (EvoDevo), the study of the evolution of the developmental mechanisms underlying the morphological diversity of organisms, is a growing field of research. By understanding the evolution of an organism, we can see why it functions and has phenotypes as it does today. There are many interesting questions in this field. One long-debated topic is the general relationship between development and evolution.

The cornerstones of evolutionary developmental theory are the arguments formulated in von Baer's Laws of Embryology and Ernst Haeckel's Biogenetic Law.

Although von Baer did not accept the concept of evolution, his idea that the earlier stages of embryogenesis reflect common features of a broader taxonomic group (Von Baer, 1828) laid the foundation for current evolutionary views on the conservation of developmental programs. Ernst Haeckel, who advocated evolution, proposed in his Biogenetic Law that embryogenesis reflects the evolutionary past of the species. Thus, according to this concept, new developmental stages are added to the ancestral embryonic program to produce more recent species (Haeckel, 1866). After a long debate, the general concept of early embryogenesis conservation, reflecting a more ancestral stage, continued in a monotonic developmental conservation model known as the funnel model. More recently, a developmental conservation model called the developmental hourglass model was proposed, which placed the most significant stage conservation at a mid-embryonic part of embryogenesis, defined as the phylotypic period (Duboule, 1994).

Comparative gene expression studies have been widely used to explore the relationship between development and evolution. The applicability of the hourglass model to maintain developmental stages at mRNA expression levels has been supported in both animals (Domazet-Loso & Tautz, 2010; H. Hu et al., 2017; Irie & Kuratani, 2011; Irie & Sehara-Fujisawa, 2007; Kalinka et al., 2010) and plants (Cheng, Hui, Lee, Wan Law, & Kwan, 2015; Cridge, Dearden, & Brownfield, 2016; Quint et al., 2012). Recent studies demonstrated the applicability of the funnel model or the coexistence and validity of both funnel and hourglass models at different levels of trait evolution in vertebrates (Artieri, Haerty, & Singh, 2009; Bininda-Emonds, Jeffery, & Richardson, 2003; Comte, Roux, & Robinson-Rechavi, 2010; H. Hu et al., 2017; Levin et al., 2016; Piasecka, Lichocki, Moretti, Bergmann, & Robinson-Rechavi, 2013; Roux & Robinson-Rechavi, 2008; Uesaka, Kuratani, Takeda, & Irie, 2019). However, there have been many debates and questions.

One of the major barriers in testing the relationship between development and evolution, including the applicability of concepts proposed in Biogenetic law or von Baer's law of embryology, is the difficulty in identifying evolutionarily homologous developmental stages among distant species. While pioneering studies focused on the

investigation of embryonic morphology (Jeffery, Bininda-Emonds, Coates, & Richardson, 2002), recent works relied on developmental changes in mRNA expression profiles as an informative endophenotype (Bozinovic, Sit, Hinton, & Oleksiak, 2011; J. J. Li, Huang, Bickel, & Brenner, 2014; Piasecka et al., 2013; Yanai, Peshkin, Jorgensen, & Kirschner, 2011). The use of gene expression facilitates inter-species comparisons, as orthologous genes can be matched among species, and their expression profiles could be traced at all stages of development. Several studies investigated developmental gene expression patterns in mammals (R. A. Wagner, Tabibiazar, Liao, & Quertermous, 2005), vertebrates (Piasecka et al., 2013), chordates (Levin et al., 2016; Yanai et al., 2011), and fruit fly species (Tomancak et al., 2007). However, these studies either focused on gene expression patterns within a particular species or compared the gene expression within the concepts of the hourglass model paradigm. None of these studies have examined the role of gene expression change in large-scale quantitative comparisons. We still lack a direct overview of the relationship between morphology and temporal gene expression change with the scale of evolutionary history. We want to know whether previous results on single or few species could be extended to more distant species or not. Whether the embryonic developmental stage is comparable between species or not. Whether the embryonic developmental stage is functionally equivalent or not when anatomical structures become less similar with more distant species(Roux, Rosikiewicz, & Robinson-Rechavi, 2015). What is the molecular force contributing to the conservation and divergence of gene expression? How significant is the regulatory effect on final organ development?

It has been shown that miRNA is involved in embryonic development and tissue differentiation (Brennecke, Hipfner, Stark, Russell, & Cohen, 2003; Bushati, Stark, Brennecke, & Cohen, 2008). Several comparative analyze of miRNA sequences showed that the emergence of miRNA genes correlates with the morphological complexity of organisms and suggest that miRNAs play an important role in evolutionary changes of body structure (Heimberg, Sempere, Moy, Donoghue, & Peterson, 2008; Prochnik, Rokhsar, & Aboobaker, 2007; Sempere, Cole, McPeek, & Peterson, 2006). A study showed that ancient miRNA expression is conserved at multiple tissues between protostomes and

deuterostomes, including the central nervous system, sensory tissue, muscle, and gut (Christodoulou et al., 2010). It indicated that ancient miRNAs have the functions to establish tissue identity. We guess that most miRNAs might be substantially involved in maintaining cell differentiation rather than tissue identity. Moreover, many studies reported that miRNAs change in their spatiotemporal expression pattern in vertebrates and invertebrates (Aboobaker et al., 2005; Krichevsky et al., 2003; Leong et al., 2016; Rahmanian et al., 2019), suggesting that the changes may thus be a major factor in the evolution of phenotypic diversity.

In my Ph.D. project, I compared the dynamic gene expression profiles of eight chordate species (amphioxus, ciona, zebrafish, two frog species, turtle, chicken, and mouse) and one outgroup species (oyster). In contrast to the early conservation and hourglass model, we found a parallel relationship on embryonic developmental stages between two species dominating all comparisons. Further, the most conserved gene expression module in the developmental program involved genes highly expressed in the 2-cell/8-cell stage and was enriched in genes involved in spliceosome-related cellular processes. Different tissue composition or asynchronous developmental rates could account for some of the expression differences among species. To explore these questions, I further examined mRNA and miRNA gene expression covering the development of three major organs (brain, liver, and heart) from early organogenesis to adulthood in eight species (human, rhesus monkey, marmoset, mouse, rat, rabbit, opossum, and chicken) to investigate the primary characters underlying the evolution of each regulator and the regulatory relationship with its target gene. These results are described in Chapter 4.

1.3 The Aim of the thesis

The overall goal of my Ph.D. project is to use gene and miRNA expression as quantitative molecular traits to study human and vertebrate evolution. The broader intent is to understand transcriptome diversity in the evolution of embryonic development and population divergence and incorporate this information into a broader picture of molecular mechanisms underlying the evolution of complex traits.

Specific aims of my studies are:

1. Investigate the role of demographic factors that determine variation in miRNA expression in the placenta between human populations.
2. Assess the extent of human developmental regulation and variability at mRNA and miRNA levels.
3. Assess the relationship between development and phylogenetic divergence at evolutionary homologous developmental stages among distant species.
4. Understand the role of miRNA and its regulatory relationship on vertebrate organogenesis.
5. To explore the evolutionary mechanisms contributing to the maintenance and divergence of gene expression in development and their effect on the mature organ profile.

1.4 Structure of the thesis

My dissertation consists of five parts: a general introduction, a literature review, two main results, and a conclusion. My colleagues listed below handled all experiments involved in my dissertation. I will only focus on the description of the data analysis in the main results in chapters 3-4.

Chapter 2. Review of the literature

Chapter 2 provides an overview of applications to the comparative analysis of coding and non-coding RNA expression on development and evolution.

Chapter 3. The expression variation of microRNA in placental tissue among human populations and its functional implications.

In Chapter 3, I present results on variation in miRNA expression in the placenta in four human ethical groups. Prof. Dr. Mark Stoneking and Dr. David A. Hughes from Max Planck Institute for Evolutionary Anthropology, Germany, provided human placenta samples collected from Northside Hospital in Atlanta, Georgia, USA. Miss. Shuyun Huang

and Miss. Xi Jiang from Shanghai Institute of Nutrition and Health, CAS, China, extracted RNA and prepared small RNA libraries for sequencing. The manuscript was published at BMC genomics.

Chapter 4. Comparative transcriptome analysis quantifies molecular phenotypes in embryonic development and evolution.

In Chapter 4, I present the results of gene and miRNA expression analysis of vertebrate embryonic development. The RNA-seq dataset for gene expression was downloaded from our published study. The miRNA-seq dataset on three organs was generated in collaboration with Prof. Dr. Henrik Kaessmann and Dr. Margarida Cardoso Moreira from ZMBH - Center for Molecular Biology, Germany. Dr. Katharina Moessinger, Dr. Jean Halbert, and Dr. Maria Warnefors from ZMBH extracted small RNA and constructed the sequencing libraries. Data analysis was performed in collaboration with Dr. Haiyang Hu and Dr. Chuan Xu from Shanghai Institute of Nutrition and Health, CAS, China. Dr. Haiyang Hu identified one pairwise orthologous gene and one gene orthologous to nine species.

Chapter 2. Review of the Literature

This chapter gives a review on Next-generation sequencing (NGS) applications for comparative analysis of coding and non-coding RNA expression in development and evolution. High-throughput sequencing is a revolutionary technology for molecular biology research. High-throughput data and corresponding low cost of RNA-seq are promoting a rapidly growing number of applications in recent years. High coverage reads of RNA-seq can cover almost all expressed coding or non-coding RNA. Consequently, RNA-seq is sufficient to characterize the molecular properties of many processes. Particularly, RNA-seq allows us to discover novel coding transcript structures, long non-coding RNA, miRNA, or other hidden "junk DNA" combined with bioinformatics analysis at the cell, tissue, and organism levels. This section focuses on reviewing comparative studies of coding and non-coding RNA on development and evolution. For more details on comparative studies, see the reviews by (Gildor & Smadar, 2018; Liu, Yu, Liu, Li, & Li, 2015; Pantalacci & Semon, 2015; Romero, Ruvinsky, & Gilad, 2012; Roux et al., 2015)

2.1 Challenges in study design and RNA-seq data processing in Evo-devo

A well-designed study in evolution and development must consider at least three factors: coverage of multiple species, efficient developmental timing, and organs or cell types to be studied. The classic case-control study considers only one tissue type under different conditions, as in biomedical research or the adaptive evolution study. The species-specific molecular trait study usually compares only one organ or cell type under a unique time point in the evolutionary timeline. To interpret and identify the essential mechanisms behind phenotypes during development, we should integrate multi-dimensional data on different molecular levels, time points, and tissues. Here we will present some challenges on evo-devo.

Transcripts from multi-species RNA-seq data

Early studies in evo-devo used microarray techniques limited by the type of species available. In this decade, state-of-the-art RNA-seq technology opens many new

possibilities to work with more species, even without a corresponding genome. As described in the last section, complete coding transcript sequences can be obtained and quantified by de novo transcript assembly of RNA-seq reads when the genome or transcriptome is unavailable or not well annotated (Garber, Grabherr, Guttman, & Trapnell, 2011). However, the quality of assemblies can depend on read length, read redundancy, chimeric artifacts, and other factors. For example, in our previous comparative transcriptome study on eight chordates, we assembled the coding genes of African clawed frogs and amphioxus from our RNA-seq data (Table 1) (H. Hu et al., 2017), yielding almost twice as many genes as the western-clawed frog annotated by Ensembl, although these two frogs have very similar morphologies. However, we cannot assess the quality of the assembly with only short RNA-seq reads available. In the context of multiple species, methods that assemble transcripts for all species using the same procedure can reduce bias from different annotations (Perry et al., 2012), but it is time-consuming.

Species	**Annotated coding genes**
Chicken	15508
Turtle	18183
Mouse	22712
Western clawed frog	18442
African clawed frog	36341
Zebrafish	26194
Ascidian	15287
Amphioxus	25135

Table 1. The number of annotated coding genes for each species.
This table is adopted from the supplementary table of (H. Hu et al., 2017).

Further, it is crucial to identify the orthologous relationships of transcripts to compare gene expression between different species. It can be problematic when we compare multiple species, especially in distantly related species. The proportion of one-to-one ortholog genes decay very rapidly with increasing of the intergenomic evolutionary distance (Wolf & Koonin, 2012). Our previous study also showed that the number of one-to-one orthologs in eight chordates and oyster is about 10% of the pairwise one-to-one orthologs (Table 2). Genes that are most similar to each other in compared genomes identified by bidirectional best hits (BBHs) might not be orthologous (Tatusov, Koonin, &

Lipman, 1997). On the other hand, ortholog does not necessarily imply conservation of gene functions, and genes with equivalent functions are not necessarily orthologous vice versa. A study examining homologous proteins in Human and Mouse showed that paralogs often have a higher functional similarity than orthologs, even at lower sequence similarities (Gabaldon & Koonin, 2013; Nehrt, Clark, Radivojac, & Hahn, 2011). To preserve information as much as possible, we proposed an approach based on co-expressed modules to identify functional conservation or divergence between species in my project described in Chapter 4. However, not all species have well-annotated GO, and metabolic pathways. Some conserved long non-coding RNAs (LncRNA) could be overlooked, so further research is required to refine the concepts on the analysis of functional and evolutionary implications.

Type	**Gene number**
Eight chordates +Oyster 1:1 orthologs	1,704
Six vertebrates and Amphioxus 1:1 orthologs	2,393
Six vertebrates 1:1 orthologs	4,814
Tetrapods 1:1 orthologs	6,347
Amniotes 1:1 orthologs	10,718
Chicken-Turtle 1:1 orthologs	11,957
Frogs 1:1 orthologs	15,207

Table 2. The number of identified 1:1 orthologous gene.
The table is adopted from Hu et al. (H. Hu et al., 2017)

Comparing developmental time and rate of embryos between species

The development of an embryo or an organ is a dynamic process with the morphology change involving cellular differentiation and divergence, all reflected in RNA-seq time series data. Comparing the temporal change in trajectory profile between two species does not consider the absolute gene expression value at a single stage and could remove expression bias from sequencing batch or other effects. Somel et al. studied human brain-specific functions by comparing chimpanzee and rhesus macaque in the course of development after birth. The results showed that genes with consistent expression differences were enriched in mRNA processing and splicing-related processes. In contrast, the type of genes that underwent developmental remodeling was associated with brain-

specific functional processes (Somel et al., 2011). Comparing transcriptomes over time may provide new perspectives to assess molecular changes in heterochrony (Margolin, Khil, Kim, Bellani, & Camerini-Otero, 2014; Tan et al., 2013). However, we need to select the appropriate developmental time points in paired embryos or organs to compare the developmental trajectories of species. For closely related species this task is more straightforward, as researchers could determine the matched stages based on the embryo's developmental rate and morphology markers (Gildor & Ben-Tabou de-Leon, 2015; Levin, Hashimshony, Wagner, & Yanai, 2012; Yanai et al., 2011). However, the work is far from trivial. Recently, Li et al. proposed an approach to mapping developmental stages between two species based on the global alignment algorithm, which used genes highly expressed at a stage without prior knowledge of embryonic development (J. J. Li et al., 2014). In a sense, this is an effective method, but sometimes this algorithm could result in several alternative alignments. Therefore, this problem is still a big challenge. Efficient computational algorithms are needed.

Tissue complexity at different stages of development

Embryonic development begins with a zygote, passing through cleavage, gastrulation, organogenesis, then develops from larva to maturity. By the process of gastrulation, the cells of the embryo transform from a one-dimensional layer to three germ layers. Each germ layer corresponds to the development of specific tissue or organ during organogenesis. At the very early embryonic stage of development, cell compositions of different species undoubtedly have different proportions and localizations. Changes in gene expression level during the developmental process may correspond to developmental time, tissue composition changes, cell type compositions and other influences. Most previous evo-devo studies have used whole embryos or whole organs without considering tissue complexity for practical reasons, leading to bias in interspecies comparisons. New single-cell sequencing technology gives access to molecular information at a single cell level and uncovers different cell subtypes in a cell population (Yan et al., 2013; Yao et al., 2017). Some deconvolution methods could estimate the composition of cell types in bulk

tissues when marker genes are available (Margolin et al., 2014). However, new computational tools and algorithms are needed to match cell types in an unbiased manner or compare expression between different species given differences in transcript sequence that might influence expression level quantification (Konstantinides, Degabriel, & Desplan, 2018; X. Wang, Park, Susztak, Zhang, & Li, 2019).

2.2 Comparison of gene expression during development and evolution

Using RNA-seq technology, we are able to measure and compare gene expression levels between different species. This can achieve many applications in the field of development and evolution. A classical study is to understand the general relationship of evo-devo. The relationship between development and evolution has been discussed for many generations by biologists in evolutionary developmental embryology. Briefly, the concept began with von Baer's law, even though von Baer did not formulate it within an evolutionary framework but to describe variability of the existing life forms. This law postulated that embryos in different animals begin with similar morphological features and become increasingly different from each other as ontogeny progresses. Later, Ernst Haeckel proposed a Biogenetic Law of development most known for its slogan "ontogeny recapitulates phylogeny" which states that embryogenesis is a rapid repetition of the evolutionary history of a species. In modern biology, the "funnel-like" model (early conservation model) corresponds to von Baer's third law, which states that the earlier stages of embryogenesis reflect a more ancestral state than the later stages. In the 1990s, Denis Duboule proposed the "developmental hourglass model" (Duboule, 1994), which predicts that the middle period of embryogenesis is more conserved (termed the phylotypic stage) than the early and late stages among animal phyla. These two patterns of conservation have long been debated. New high-throughput sequencing technology makes it possible to include more species without genome sequence in large-scale studies to resolve this debate.

Quantifying transcriptome levels between different species within the phylum provided new insights into these two models from a molecular perspective. Most of the data supported the hourglass-like model over the funnel model. Kalinka et al. compared

expression profiles between six Drosophila species separated by 40 million years and proved an hourglass pattern in developmental evolution. A study among four vertebrate species (mouse, chicken, zebrafish, and frog) showed that the pharyngula stage has the most conserved gene expression, in agreement with the morphological phylotypic stage in the vertebrate phylum (Irie & Kuratani, 2011). The hourglass model of expression divergence has also been supported in nematode species (Levin et al., 2012), chordates (Domazet-Loso & Tautz, 2010; H. Hu et al., 2017), and even plants such as Arabidopsis thaliana (Quint et al., 2012) and fungi (Cheng et al., 2015). In addition to the hourglass model, some studies have shown the coexistence of the funnel and hourglass models in the evolution of various traits(Artieri et al., 2009; Bininda-Emonds et al., 2003; Comte et al., 2010; H. Hu et al., 2017; Levin et al., 2016; Malik, Gildor, Sher, Layous, & Ben-Tabou de-Leon, 2017; Piasecka et al., 2013; Roux & Robinson-Rechavi, 2008; Uesaka et al., 2019) or other models(Cordero, Sanchez-Villagra, & Werneburg, 2020). However, the results often introduce pattern bias when the species studied have distant phylogenetic distances. A recent large-scale comparative study of ten distantly related species representing ten phyla showed conservation of mid-development observed within phyla but not among different phyla (Levin et al., 2016). Our previous study of eight chordates split 550 million years ago also showed that the hourglass model is more applicable to vertebrates than chordates (H. Hu et al., 2017). Such methods relying on pairwise comparison between species have been criticized for failing to account for the phylogenetic relationships of species, especially for distant species (Dunn, Zapata, Munro, Siebert, & Hejnol, 2018). Further, more recently, substantial influence of the maternal transcriptome (Atallah & Lott, 2018), potentially affecting divergence at first developmental stages, came into consideration, complicating interpretation of previous observations. Overall, regardless of methodological inconsistencies, the relationship between ontogeny and phylogeny might be far more complicated than we initially appreciated.

In addition, comparative studies in closely related species revealed several insights into understanding molecular changes with morphology divergence under adaptation. Closely species or species in different populations have divergent morphology despite

having a highly conserved body plan. During embryonic development, the molecular mechanism of maintaining the conserved embryonic phenotype and, at the same time, providing adaptation to environmental changes is essential for understanding morphology divergence in adulthood. Many studies in animals demonstrated the relationship between embryogenesis and environmental changes, such as temperature (Kuntz & Eisen, 2014; Levine, Eckert, & Begun, 2011; Runcie et al., 2012), latitude (L. Zhao, Wit, Svetec, & Begun, 2015), acidity (Byrne et al., 2009; Evans, Chan, Menge, & Hofmann, 2013; Martin et al., 2011), and nutrition (Bennis-Taleb, Remacle, Hoet, & Reusens, 1999; Gluckman et al., 2007; Vickers, Breier, Cutfield, Hofman, & Gluckman, 2000). These responses to environmental challenges in early development will affect the adult organism, just as they do in humans. The human fetus and infant may alter their growth in response to imbalanced nutrition or other influences (Barker, Forsen, Uutela, Osmond, & Eriksson, 2001; Bateson et al., 2004; Gluckman, Hanson, Spencer, & Bateson, 2005).

In summary, the developmental process of any organism is essential to its long-term outcome. In response to any environmental changes during development, organisms have evolved several different strategies to adapt to or compensate for environmental perturbations. The mechanism of such developmental plasticity is not well understood. In the meantime, speciation could be considered as a long-term environmental adaptation until some traits fit in a population. Different lineages gain adaptations that improve their fitness over time, and some of these would be reflected as divergence. Comparative studies between multiple species could help us understand the molecular mechanism of conservation and divergence on evo-devo. Together with genome, transcriptome and epigenome level data, the multidimensional data on developmental time series and phylogenetic history could greatly improve our knowledge of evolution and development.

2.3 Comparison of expression of non-coding RNAs in evolution and development

In addition to protein-coding genes, several types of non-coding RNAs (ncRNA) exist in the genome; some are involved in the nuclear function of mRNA translation, such as ribosomal RNAs and tRNAs, or in splicing (small nuclear RNAs). One type of non-

coding RNA is the so-called “junk RNA”, which was first proposed by Britten and Davidson (Britten & Davidson, 1969, 1971) almost 50 years ago and was mostly found in the intergenic region of the genome. Recently, more and more types of “junk RNAs” have been detected, such as long non-coding RNAs (lncRNA) with lengths over 200nt; small non-coding RNAs with 18-30nt length, including miRNAs, Piwi-interacting RNAs (piRNAs), and endogenous small interfering RNAs (endo-siRNAs). In the last decade, advances in high-throughput sequencing have revealed a variety of non-coding RNAs and their functional characterization in both plants and animals. Although the elucidation of the functional role of ncRNAs is more delayed than that of protein-coding genes, the regulatory role on gene expression is widely accepted. piRNAs are predominantly expressed in the animal germline, where they silence transposable elements (Aravin et al., 2006; Grivna, Beyret, Wang, & Lin, 2006; Malone & Hannon, 2009), whereas endo-siRNAs are mainly expressed in oocytes and can regulate oocyte meiosis (Kaneda, Tang, O'Carroll, Lao, & Surani, 2009). miRNAs are expressed in both plants and animals and are essential for regulating gene expression by facilitating mRNA degradation or translational repression (Lewis et al., 2005; Lewis et al., 2003). Studies have shown that miRNAs are involved in development, cellular progression, signaling, morphogenesis, and organogenesis (Kloosterman & Plasterk, 2006). lncRNAs, another type of well-studied ncRNA, are present in the genome (Carninci et al., 2005; E. P. Consortium et al., 2007) and play essential roles in regulating critical cellular processes such as cell self-renewal and differentiation, DNA damage response, and immune surveillance (Atianand, Caffrey, & Fitzgerald, 2017; Sanchez & Huarte, 2013). Such regulatory functions of non-coding RNAs are important for understanding the mechanisms underlying phenotypic divergence within and between species. In the following, I focus only on the comparative studies of miRNAs gene expression towards the function of development and evolution.

The functional role of individual miRNA has been intensively studied (reviewed in (Kloosterman & Plasterk, 2006; Sayed & Abdellatif, 2011; Vidigal & Ventura, 2015)). For example, in animal development, the let-7 family, which is highly conserved between invertebrates and vertebrates, may regulate the embryonic developmental program from

early to adult transition in C. elegans (Abbott et al., 2005; Abrahante et al., 2003; M. Li, Jones-Rhoades, Lau, Bartel, & Rougvie, 2005); the miRNA family miR-1, which is conserved from worms to mammals, is involved in the progression of myogenesis and cardiogenesis (Aboobaker et al., 2005; Y. Zhao, Samal, & Srivastava, 2005); pancreatic islet-specific miR-375 can affect glucose-induced insulin secretion (Poy et al., 2004). Meanwhile, many studies suggest that the diverse changing miRNA expression patterns during embryonic development may imply potential functions for tissue genesis and phenotypic diversity in evolution (Aboobaker et al., 2005; Krichevsky et al., 2003; Leong et al., 2016; Rahmanian et al., 2019). The ENCODE project recently measured miRNA expression profiles in twelve organs during mouse embryonic development using high-throughput sequencing and, thus, provided a comprehensive atlas of miRNA expression alterations during development of the major organ systems. The study showed that 83% of tissue-specific miRNAs are preferentially expressed in neuronal and muscular tissues and the expression is well conserved between human and mouse (Rahmanian et al., 2019)

With the advances in next-generation sequencing technology and the increasing number of available genomes in recent years, miRNAs have been also widely studied on evolution. The spread of emerging miRNAs correlates with the morphological complexity of organisms (Heimberg et al., 2008; Prochnik et al., 2007; Sempere et al., 2006), as shown in previous studies. Recently, a study on miRNAs across 12 species of the genus Drosophila identified 649 completely new miRNA loci and found striking differences in birth and death rates across different miRNA classes (Mohammed et al., 2018). Another study for five organs in six mammals provided the first global overview of birth and expression of miRNAs and suggested the associated selective forces of evolution in mammals (Meunier et al., 2013). Similar comparative studies have been done in organs of domestic mammals (Penso-Dolfin, Moxon, Haerty, & Di Palma, 2018), *Micropterus salmoides* (W. Gong et al., 2018), embryonic development of *Grapholita molesta* (X. Wang et al., 2017) and *Chelicerates* (Leite et al., 2016). However, the relationship between miRNA evolution and development is not well studied. Only one study in two divergent fruit flies, *Drosophila melanogaster* and *D. virilis*, explored the phylotypic stage of

miRNAs during development and showed an hourglass pattern in miRNA expression, similar to that observed for protein-coding genes (Ninova, Ronshaugen, & Griffiths-Jones, 2014). However, understanding the constraints on miRNA expression in development needs to be expanded in the future.

In summary, gene and miRNA expression could be valid and informative as quantitative molecular phenotypes in development and evolution studies. Comparative studies in more species, even without available genome annotation, could further promote our understanding in the field of evo-devo with the help of next-generation sequencing technology.

Chapter 3. The expression variation of microRNA in placental tissue among human populations and its functional implications

Analysis of lymphocyte cell lines revealed substantial differences in the expression of mRNA and microRNA (miRNA) among human populations. The extent of such population-associated differences in actual human tissues remains largely unexplored. The placenta is one of the few solid human tissues that can be collected in substantial numbers in a controlled manner, enabling quantitative analysis of transient biomolecules such as RNA transcripts. Here, we analyzed microRNA (miRNA) expression in human placental samples derived from 36 individuals representing four genetically distinct human populations: African Americans, European Americans, South Asian Americans, and East Asian Americans. Sequence analysis of the miRNA fraction yielded 938 annotated and 70 novel miRNA transcripts expressed in the placenta. Of them, 82 (9%) of annotated and 11 (16%) of novel miRNAs displayed quantitative expression differences among populations, generally reflecting reported genetic and mRNA-expression-based distances. Several co-expressed miRNA clusters stood out from the rest of the population-associated differences in terms of miRNA evolutionary age, tissue-specificity, and disease-association characteristics. Among three non-environmental influenced demographic parameters, the second largest contributor to miRNA expression variation after population was the sex of the newborn, with 32 miRNAs (3% detected) exhibiting significant expression differences depending on whether the newborn was male or female. Male-associated miRNAs were evolutionarily younger and correlated inversely with the expression of target mRNA involved in neuron-related functions. In contrast, both male and female-associated miRNAs appeared to mediate different types of hormonal responses. Demographic factors further affected reported imprinted expression of 66 placental miRNAs: the imprinting strength correlated with the mother's weight, but not height.

3.1 Results

3.1.1 miRNA identification and characterization

We collected placenta samples from individuals representing four major human ethnic groups (further referred to as populations): African Americans, European Americans, South Asian Americans, and East Asian Americans (Figure 1A). All samples were collected at the same geographic location (Northside Hospital in Atlanta, Georgia) from residents of the area following a unified protocol to minimize potential biases that might influence the results. We sampled each placenta at five sites within the central villous parenchyma region and pooled the dissected samples before the mRNA and miRNA isolation (Hughes et al., 2015).

For each population, we sampled ten individuals (Table S1) and estimated miRNA expression levels using high-throughput transcriptome sequencing (RNA-seq) conducted on the Illumina HiSeq X10 platform. We pooled 14 samples with different index sequences, carried out 150-bp paired-end sequences, and then obtained an average of 31.4 million reads. Based on these data, we detected 938 miRNAs annotated in miRbase (v22) and 70 novel miRNAs (Figure 1B) with the total expression count among all samples greater than 100 reads. The novel miRNAs for each sample were detected using miRDeep2 algorithm with default parameters based on the human genome (hg38) and miRbase v22 as references (Friedlander, Mackowiak, Li, Chen, & Rajewsky, 2012). Predicted rRNA/tRNA reads and reads with the miRDeep2 score <=5, representing the miRNA hairpin properties matching, were removed from the following analyses. Sequencing reads mapped to overlapping miRNA genomic locations were merged across all samples. The detected miRNAs with no overlap with known mature miRNA genomic positions were considered as novel.

Based on a principal component analysis of all 42 samples, four samples were considered as outliers based on their dispersion and were removed from further analysis (Figure 2). In total, data from 8 African American, 10 European American, 8 South Asian American, and 10 East Asian American individuals were retained (Figure 1A).

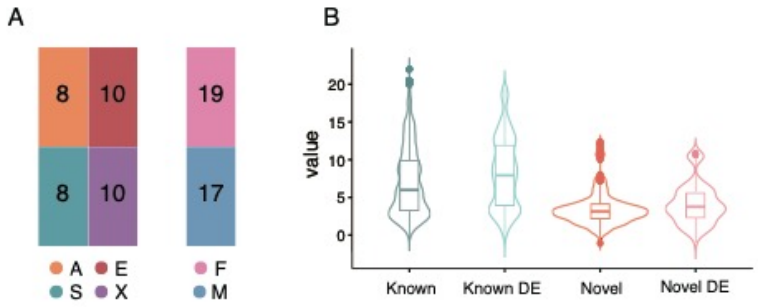

Figure 1. Sample information and miRNA expression distribution.

A. Schematic illustration of sample numbers according to population ID and newborn sex. The abbreviations here and in the text indicate: A – African Americans; E – European Americans; S – South Asian Americans; X – East Asian Americans; F - female newborn and M - male newborn. **B.** Violin plot showing miRNA expression distribution. Y-axis shows the quantile normalized log2-transformed miRNA read count values after removing the batch effect. X-axis labels indicate 938 annotated miRNAs (Known), 127 annotated miRNAs differentially expressed among populations (Known DE), 70 novel miRNAs (Novel), and 12 novel miRNAs differentially expressed among populations (Novel DE).

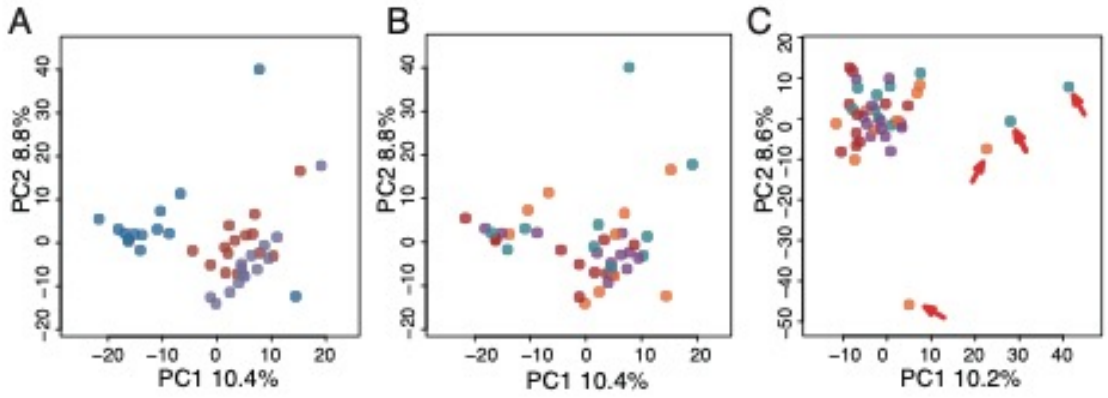

Figure 2. miRNA expression variation among 40 individuals.

A. Principal component analysis plots based on the miRNA expression of all 1,008 miRNAs in 40 individual placental samples before removing the batch effect. Colors indicate sequencing batch. Each dot represents a sample. **B.** Colors indicate human populations: orange – African Americans; red – European Americans; light blue – South Asian Americans; purple – East Asian Americans. Each dot represents a sample. **C.** Principal component analysis plots based on the miRNA expression of all 1,008 miRNAs in 40 individual placental samples after removing batch effect. Colors indicate human populations: orange – African Americans; red – European Americans; light blue – South Asian Americans; purple – East Asian Americans. Each dot represents a sample; Red arrows point to four outliers removed from further analysis.

3.1.2 miRNA expression variations in human placenta reflect human population divergence

To understand the contribution of genetic and environmental effects on human population differences and find the association between demographic variables and molecular diversities, we collected information for 26 demographic parameters (Table S1), including population identity. Among them, 12 variables had at least five observations for each level with sufficient variability to estimate their influence on miRNA expression levels. The 12 variables are including: birth delivery type (Cesarean or natural), sex of the newborn, maternal body mass index, number of pregnancies, has pregnancy infection or not, first pregnancy or not, mother's age, birth length of the newborn child, birth weight of the newborn child, number of children, drinker or not, and population identity. For these variables, we estimated miRNA expression variation using a multivariate Type I analysis of variance with the above sequential ordering, denoted as Model1. The variation explained by each factor was estimated by an eta-squared statistic using the sums of squares (ANOVA; aov() function in the R stats package; Table 3). The results showed that population identity was the major factor, with the other seven factors having a lesser effect, though still above 3% of the variance.

Demographic variable	Explained variance %
Population	10.00
Baby's sex	3.54
First pregnancy	3.49
Mother's age	3.45
Mother's bmi	3.38
Birth delivery type	3.36
Drinker	3.15
Pregnancy infection	3.09
Number of Kids	2.77
Pregnancy number	2.84
Baby Length	2.58
Baby Weight	2.37

Table 3. Average variance proportion of 12 demographic factors on all miRNAs across 1008 miRNAs

To assess the variables' effects more precisely, we recalculated the eta-squared statistics after subtracting the potential effects of four continuous demographic variables using linear regression as a second model (Model2). The four continuous variables modeled first were: maternal body mass index, birth weight of the newborn child, mother's age, and birth length of the child. The residuals for each miRNA were carried forward into subsequent linear regressions that included birth delivery type (Cesarean or natural), sex of the newborn and population identity. The estimations of variation agreed well between model1 and model 2. Besides individual differences, the most notable contributors to the miRNA expression variation were population identity and sex of the newborn (SON), explaining 11% and 4% of the total variation, respectively (Figure 3).

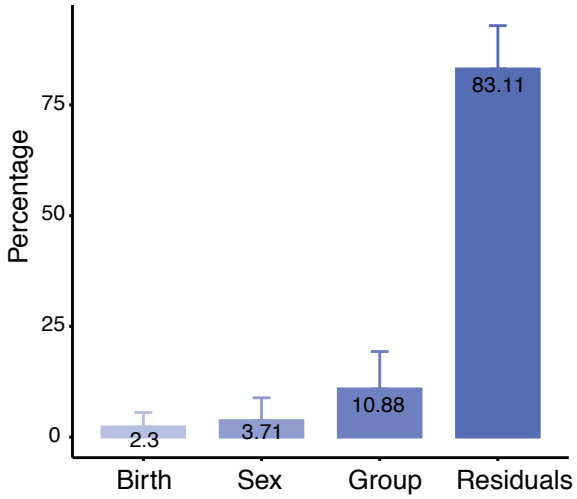

Figure 3. Percentage of total expression variance explained by newborn sex and population. Bars represents the mean of the variance proportion explained by the factors across 1008 miRNAs. Error bars represent the standard deviation of the variance.

3.1.3 miRNA expression in African Americans distant from the other populations

To identify miRNAs differentially expressed between populations, we next applied an ANOVA test with factors contributing the most to the general variation effects listed in the model in the following order: birth delivery type, sex of the newborn, and population. Accordingly, 139 miRNAs showed nominally significant expression differences among

populations, including 12 novel ones (ANOVA F-test, nominal $p<0.05$, FDR<36%; permutation $p<0.0001$; Figure 1B; Figure 4B).

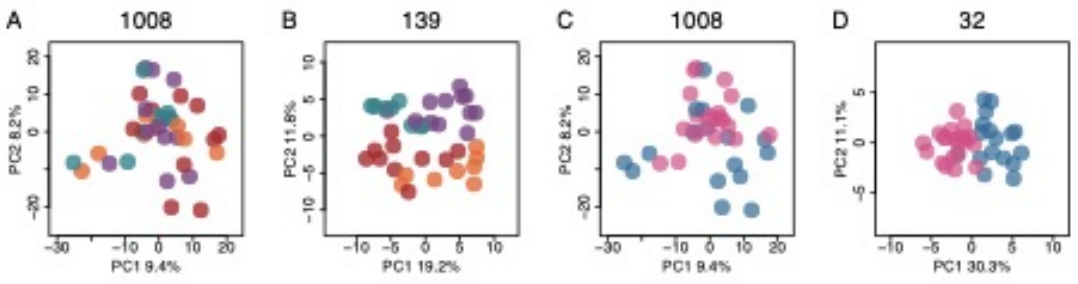

Figure 4. Principal component analysis on miRNA expression.
A, **C** Principal component analysis plots based on the miRNA expression of all 1,008 miRNAs (**A** colored according to population and **C** colored according to newborn sex), 139 miRNAs differentially expressed among populations (**B**), and 32 miRNAs differentially expressed depending on the sex of the newborn (**D**).

To further identify the miRNAs differentially expressed between any two human populations, we applied a two-sided pairwise t-test (pairwise.t.test in R multcomp package). Differentially expressed miRNAs were defined as those with a $p<0.05$ after Benjamini-Hochberg correction. Further analysis of the 139 miRNAs showing population-associated expression yielded 93 miRNAs with significant expression differences between at least one pair of populations (Student's t-test, Benjamini-Hochberg corrected $p<0.05$). Visualization of the distances among populations based on the expression of 93 or 139 population-associated miRNAs yielded dendrograms compatible with the genetic relationships among populations (Figure 5A, C, and D). The relative miRNA expression divergence among four populations investigated in the study is consistent with their genetic divergence as estimated using the 1000 Genomes data set (Two-sided Mantel permutation test, Spearman's correlation coefficients rho=0.771, p=0.08) (Genomes Project et al., 2015). Since genetic divergence is largely thought to reflect the accumulation of phenotypically neutral mutations (Kimura, 1983), it is therefore conceivable that miRNA variation among populations is similarly influenced by the phenotypically neutral changes. This notion

aligns with previous work suggesting that mRNA expression divergence includes a substantial proportion of functionally and phenotypically neutral changes (Hughes et al., 2015; Khaitovich et al., 2004).

Specifically, miRNA expression in African Americans was the most distant from the other populations, while the two Asian populations were most similar. Similarly, miRNA expression in African American population differed most from the other three based on analysis of 1,008 expressed miRNAs, but with lower expression divergence (Figure 5B).

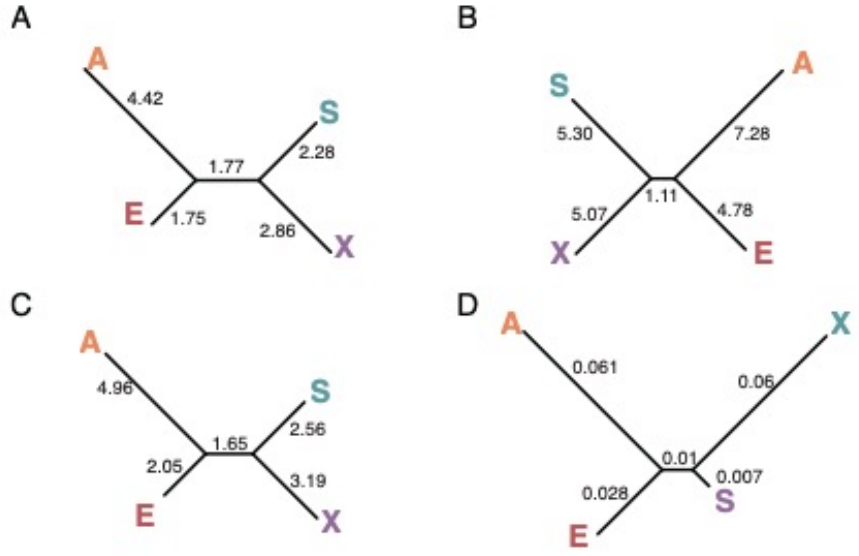

Figure 5. Characterization of miRNAs differentially expressed among human populations. **A.** Dendrogram based on expression levels of 93 population-associated miRNAs. **B.** Dendrogram based on expression of 1,008 detected miRNAs. **C.** Dendrogram based on expression levels of 139 population-associated miRNAs. The abbreviations here and in the text indicate: A – African Americans; E – European-Americans; S – South Asian Americans; X – East Asian Americans. Numbers indicate the branch length. **D.** Dendrogram of the genetic divergence among four human populations**.** Shown is a neighbor joining tree based on Fst values from the 1000 genomes project. Populations used in the analysis include African American, CEPH, Telugu, Han Chinese, Southern Han Chinese, and Kinh Vietnamese, to match the populations used in our study. The mean Fst value of the Han Chinese, Southern Han Chinese, and Kinh Vietnamese populations was considered as the value for East Asian population. The abbreviations indicate human populations: A – African Americans; E – European-Americans; S – South Asian Americans; X – East Asian Americans. Numbers indicate the branch length.

3.1.4 Molecular features and miRNA target function on six co-expressed miRNA clusters

We then asked the molecular characters on co-expressed miRNA clusters, which probably have similar functions and molecular features. To detect co-expression patterns among 93 population-associated miRNAs, we applied a hierarchical clustering method (hclust function in R) on z-transformed miRNA expression values with (1-rho) as the distance measure, where rho is the Spearman correlation coefficient(Figure 6; Figure 7).

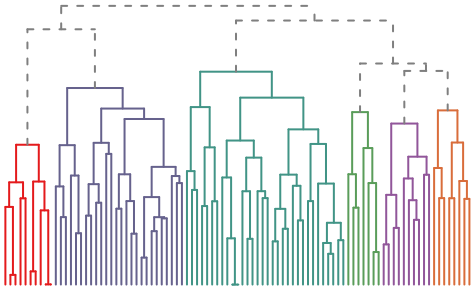

Figure 6. Differentially expressed among human populations.
Hierarchical clustering of 93 population-associated miRNAs based on correlation of their expression profiles. Colors represent six main clusters.

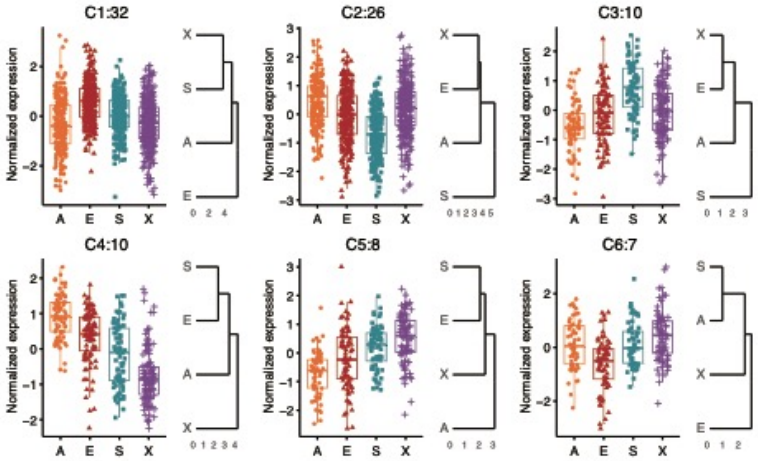

Figure 7. Clusters of miRNAs are differentially expressed among human populations.
Colors represent populations. Panel titles show the cluster name and the number of miRNAs in the cluster. Y-axis indicates Z-transformed miRNA expression values. The dendrograms on each panel's right represent the average normalized expression distances among populations based on the expression of cluster miRNAs.

Next, we characterized these clusters by showing significant feature enrichment on miRNA evolutionary age, expression tissue-specificity, and disease associations (two-sided Wilcoxon test, nominal $p<0.01$). Specifically, cluster 1 (C1), characterized by elevated expression in European American samples, contained significantly younger miRNAs than the bulk (two-sided Wilcoxon test, nominal $p<0.01$) (Figure 8A) and showed the highest miRNA expression tissue-specificity, restricted mainly to the placenta (Figure 8B). Further, cluster 5 (C5), characterized by low expression in African-Americans and elevated expression in Asian populations (Figure 7), showed the highest number of miRNA disease associations (Figure 8C; Figure 9; Table S3).

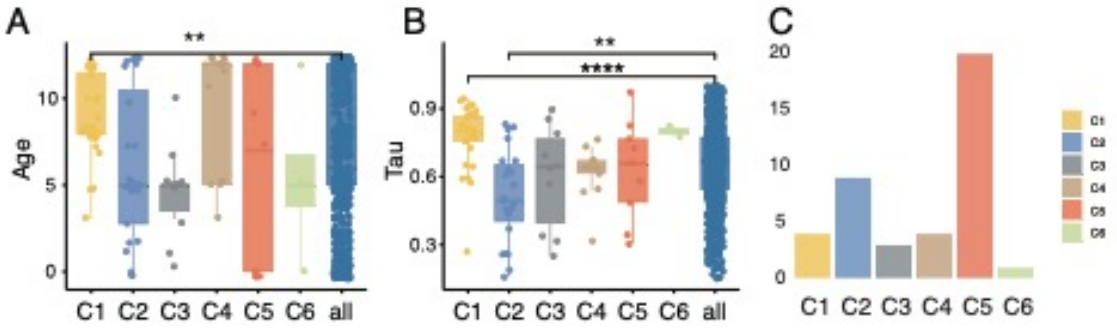

Figure 8. Characterization of miRNAs differentially expressed among human populations
A. Distribution of miRNA evolutionary age in the six clusters. The age scale extends from 433 Mya (age 0) to human-specific miRNA (age 12). Asterisks indicate the significance of the difference (two-sided Wilcoxon test, ** represents nominal $p<0.01$). **B.** Distribution of miRNA tissue expression index (Tau) in the six clusters. Large values represent greater expression tissue-specificity. Asterisks indicate the significance of the difference (two-sided Wilcoxon test, **** represents nominal $p<0.0001$). **C.** Number of miRNAs associated with disease in each cluster.

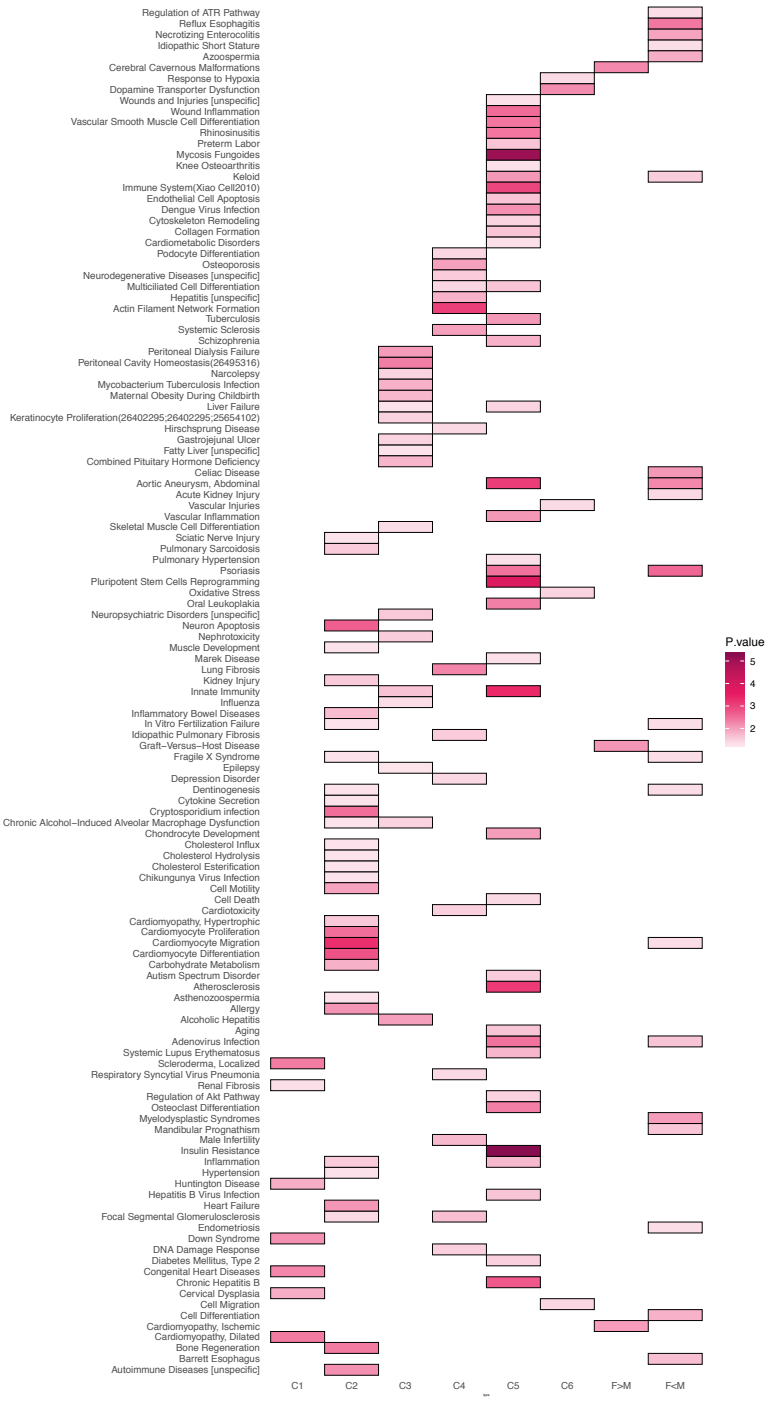

Figure 9. Enrichment for differentially expressed miRNAs in specific disease categories. Enrichment for miRNAs differentially expressed among human populations (clusters C1-6) and sex of newborn. Each term is associated with human diseases according to the TAM2.0 database with exclusion of cancer-related terms (J. Li et al., 2018). Color represents the value of -log(p-value). Only the cases with enrichment nominal $p<0.05$ are highlighted.

To further assess the potential effects of population-associated miRNAs on expression of their target genes, we examined the published mRNA expression dataset derived from a partially overlapping set of placental samples (Table S1) (Hughes et al., 2015). Only cluster 1 (C1) reveals significant downregulation of predicted targets of population-associated miRNAs (one-side Wilcoxon rank test, $p<0.05$, correlation $r<-0.5$). The potential targets of C1 miRNAs were enriched in the functional term associated with vasculogenesis and muscle organ development (Figure 10).

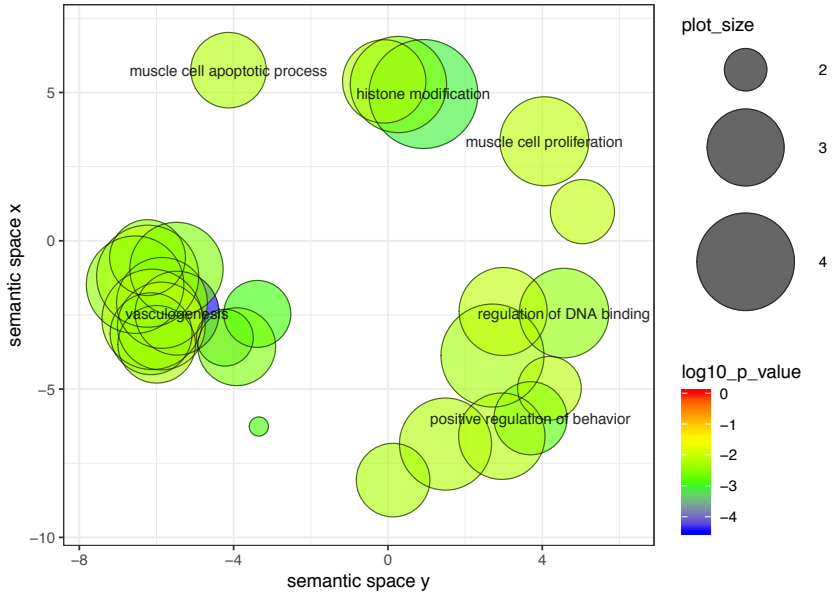

Figure 10. Visualization of enriched GO terms for miRNA targets of C1 miRNAs. Color represents the value of log10(p-value). The size of cycle represents the number of GO terms. The figure was visualized from REVIGO.

3.1.5 miRNA expression variation is associated with the sex of the newborn

In addition to population identity, we applied for an ANOVA test on the factor of sex of newborn (SON). 32 miRNAs showed expression differences (SON-associated miRNA, ANOVA F-test, nominal $p<0.01$, FDR<31%; permutation $p<0.05$; Figure 4D). Among it, 14 miRNAs were elevated in pregnancies with a male child (male-associated miRNA) and 18 in pregnancies with a female child (female-associated miRNA) (FDR<31%, permutation $p<0.05$; Figure 11; Figure 13A). All SON-associated miRNA

expression differences were reproduced in multiple populations, with 24 of the 32 miRNAs in all four populations (Exact binomial test, $p<0.01$; Figure 12). Notably, female-associated miRNAs were of significantly older evolutionary origin than most male-associated miRNAs (two-sided Wilcoxon test, nominal $p<0.05$; Figure 13B). Further, female-associated miRNAs were enriched in the imprinted mir-379 cluster (C14MC) implicated in the regulation of brain-specific functions (Winter, 2015) (hypergeometric test, Bonferroni corrected $p=4.58 \times 10^{-6}$). However, both female- and male-associated miRNA groups showed the same moderate tissue-specificity (Figure 13C).

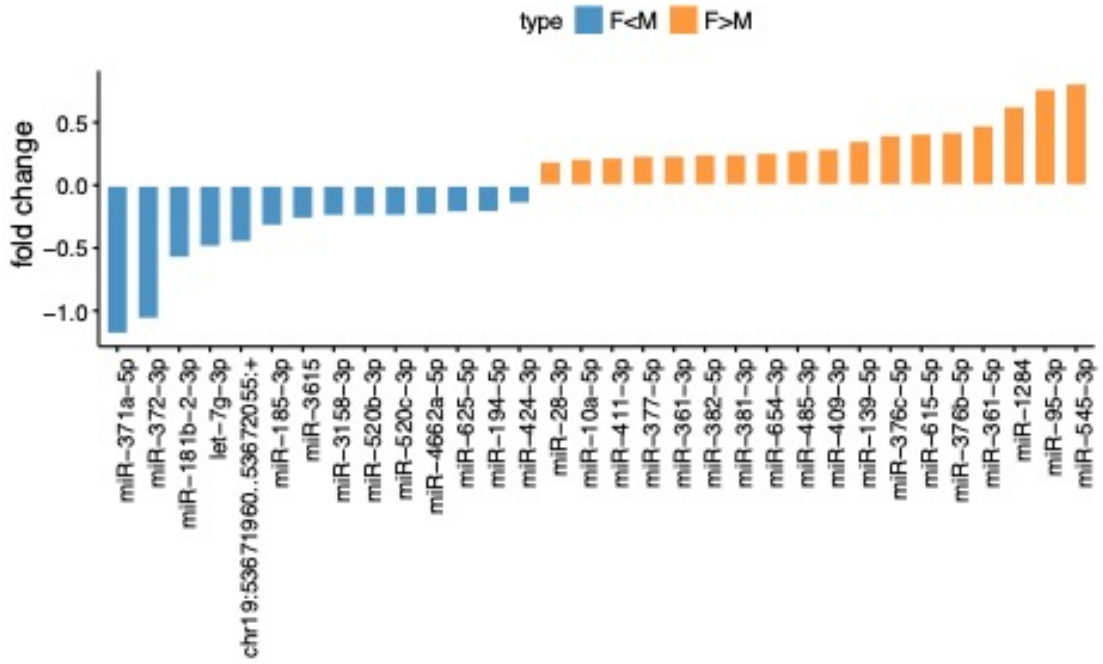

Figure 11. Characterization of miRNAs with newborn sex-associated expression.
Bar plot showing individual miRNA expression differences between placental samples from female vs. male newborn. Colors represent male-newborn-associated (F<M, blue) and female-newborn-associated (F>M, orange) miRNAs. Abbreviations: F – female newborn; M – male newborn.

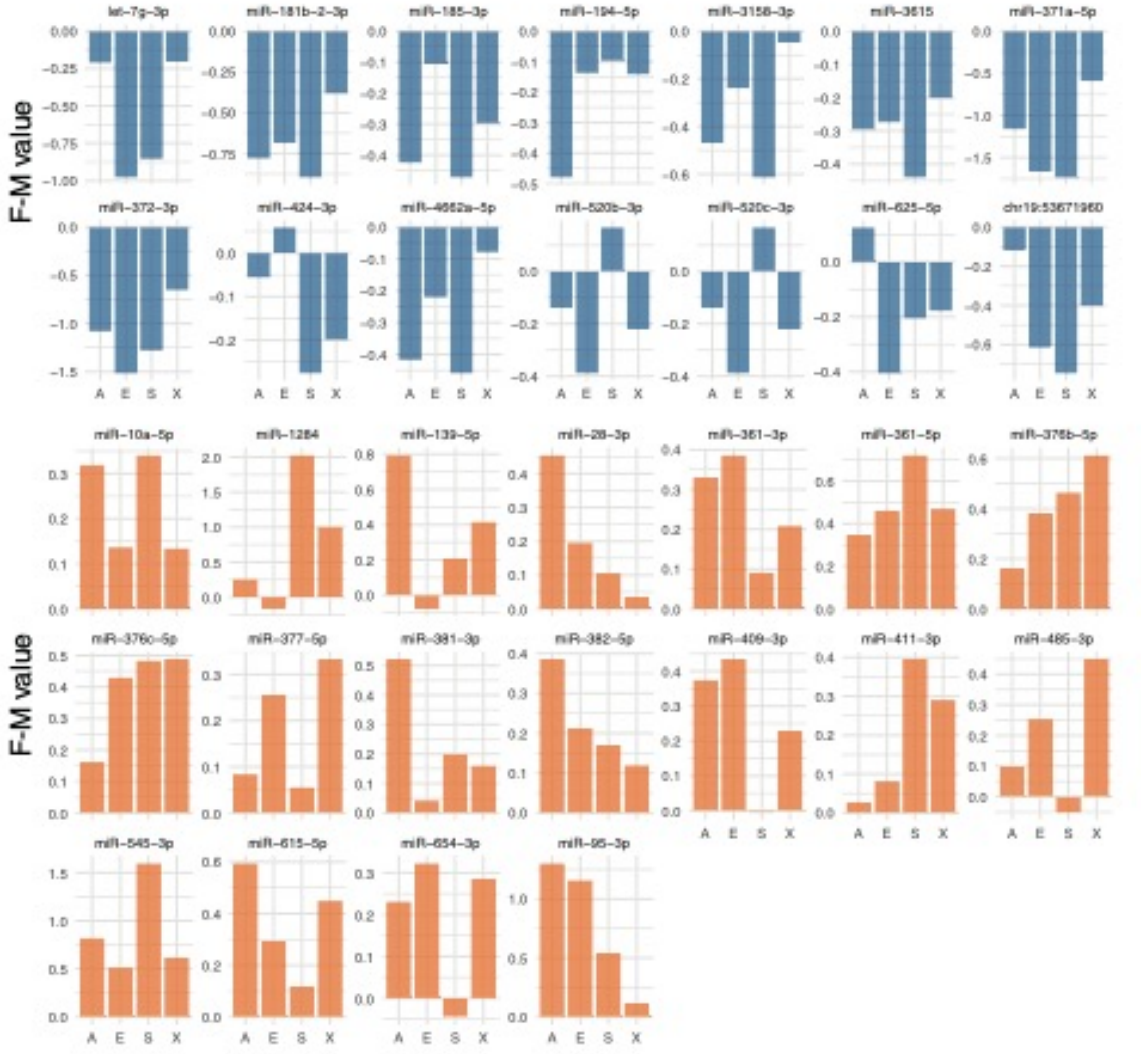

Figure 12. Expression of newborn-sex-associated miRNA in each human population. Each bar represents the expression difference between placental samples from female newborn vs. male newborn, displayed separately for each population. Colors represent male-newborn-associated (F<M, blue) and female-newborn-associated (F>M, orange) miRNAs. The abbreviations here and in the text indicate: A – African Americans; E – European Americans; S – South Asian Americans; X – East Asian Americans.

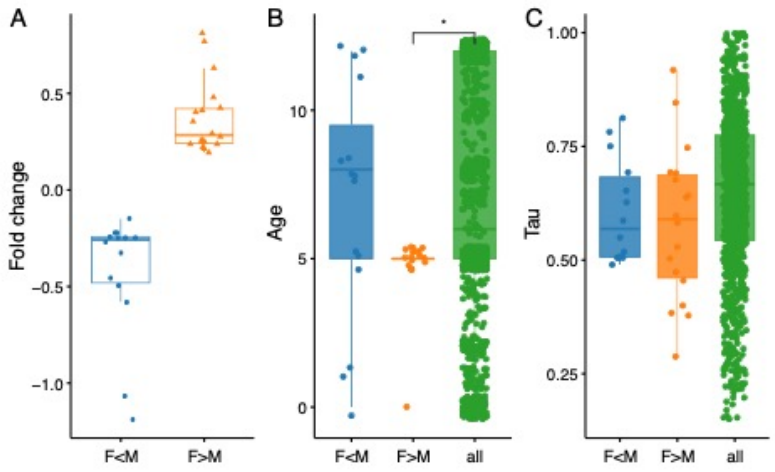

Figure 13. Characterization of miRNAs with newborn sex-associated expression.
A. Boxplot showing the distributions of miRNA expression fold-change for placental samples from female vs. male newborn infants. The blue and yellow boxes represent miRNAs with male-newborn-associated and female-newborn-associated expressions. Each dot represents one miRNA. **B.** Distribution of miRNA evolutionary age for male-newborn-associated (blue) and female-newborn-associated (orange) miRNA. The age scale extends from 433 Mya (age 0) to human-specific miRNA (age 12). Asterisks indicate the significance of the difference (two-sided Wilcoxon test, * represents nominal $p<0.05$). **C.** Distribution of miRNA tissue expression index (Tau) for male-newborn-associated (blue) and female-newborn-associated (orange) miRNA. Large values represent greater expression tissue-specificity.

Previous investigation of mRNA expression in human placenta reported 41 genes with sex-associated expression, 12 of them (30%) localized on sex chromosomes (Sood, Zehnder, Druzin, & Brown, 2006). The substantial prevalence for sex chromosome localization was not, however, the case for SON-associated differences in miRNA expression: of 32 miRNAs, four (13%) localize on sex chromosomes. Sex-associated differences in miRNA expression were similarly reported in human tissues other than the placenta. Specifically, miRNA analysis across postnatal brain development revealed 40 miRNAs with significant sex-biased expression differences in the prefrontal cortex regions, 93% of them female-biased (Ziats & Rennert, 2014). Further, investigation of four adult human tissues – brain, colorectal mucosa, peripheral blood, and cord blood – revealing 73 female-biased and 163 male-biased expressed miRNAs (Cui et al., 2018). Notably, two of 32 SON-associated miRNAs overlapped with miRNAs showing corresponding sex-biased

expression in the adult brain, and two overlapped with miRNAs showing such a bias in the peripheral blood.

3.1.6 Targets of sex of the newborn miRNAs were enriched in different functions

To assess the potential effects of SON-associated miRNA expression, we identified their potential targets in the published mRNA expression dataset derived from a partially overlapping set of placental samples (Hughes et al., 2015). In total, we classified 46 mRNAs as potential targets of male-associated miRNAs and 65 mRNAs as potential targets for female-associated miRNAs, using a combination of miRNA target predictions and the inverse relationship of miRNA and target expression profiles as selection criteria. Notably, the potential targets of male-associated miRNAs were enriched in functional terms associated with glutamate receptor signaling and endocrine processes (Figure 14A). By contrast, the potential targets of female-associated miRNAs were enriched in functions linked to steroid hormones, estradiol, glucocorticoid response, and cell differentiation and metabolic (Figure 14B), as previous studies singled out hormonal regulation was the main driving mechanism of miRNA sex-biased expression (Dai & Ahmed, 2014; Hao & Waxman, 2018).

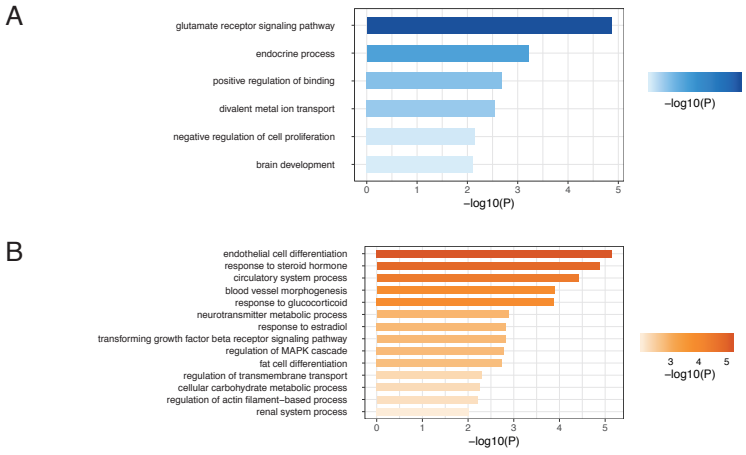

Figure 14. GO terms enriched in targets of newborn sex-associated miRNAs.
GO terms are enriched in targets of male-newborn-associated (**A** blue) and female-newborn-associated (**B** orange) miRNAs. X-axis and the number within circles indicate -log10-transformed p-values.

3.1.7 Imprinting miRNAs are associated with mother's BMI

One of the characteristic features of placental miRNA is the prevalence of imprinted expression, a term referring to complete or partial suppression of one of the parental alleles (Malnou, Umlauf, Mouysset, & Cavaille, 2018). To assess the extent of miRNA expression imprinting in our data, we focused on the largest characterized imprinted miRNA cluster, located on chromosome 19 (C19MC) and expressed almost exclusively in the placenta (Malnou et al., 2018; Noguer-Dance et al., 2010). This cluster locus contains 67 mature miRNAs (hg38: chr19:53,665,746-53,761,746), of which 66 were detected in our study. Expression analysis of these 66 miRNAs revealed a significant negative correlation with the mother's BMI (two-sided Wilcoxon test, $p=2.8x10^{-14}$) and weight ($p=1.7x10^{-10}$), but not height ($p=0.27$) (Figure 15A). This relationship was further apparent at the level of individual miRNAs (Spearman correlation, $p<0.05$; Figure 15B).

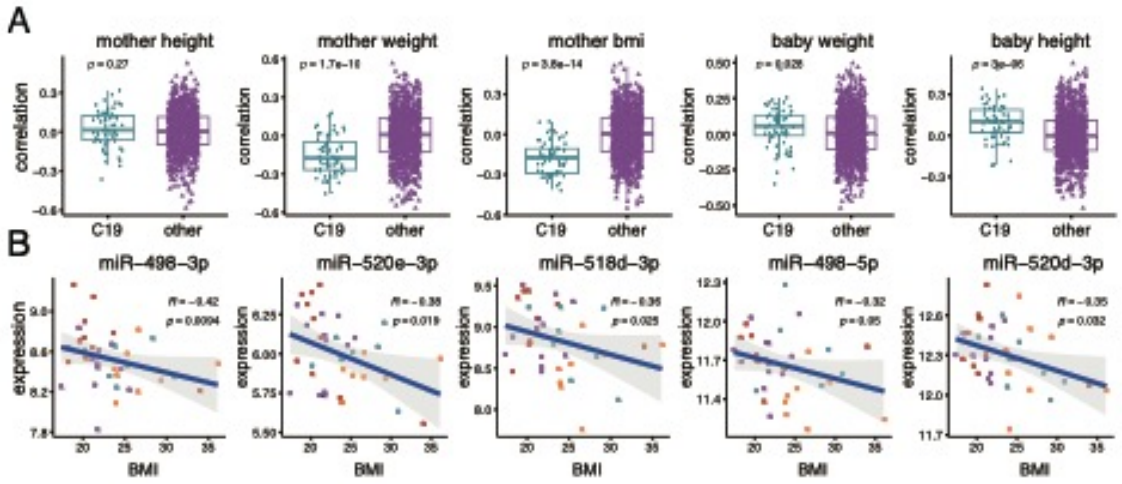

Figure 15. Expression of imprinted miRNAs in the C19MC cluster.

A. Correlation distribution between imprinted miRNAs located in the C19MC cluster and demographic variables. Panel titles indicate the demographic variable used in the comparison. P-values for a two-sided Wilcoxon test are shown within panels. **B.** Five miRNAs showing a significant expression correlation with mother's BMI. Each dot represents the expression level in a sample. Colors represent human populations as illustrated in Figure 1A. Shaded areas represent confidence intervals. Spearman's correlation coefficients rho (R) and p-values (p) are displayed in the top right corner of each scatter plot.

3.2 Methods

3.2.1 Sample collection and information

Placental samples were obtained from a previous study (Hughes et al., 2015), and included ten individuals from each ancestry (African American, European American, South Asian American (India) and East Asian American ancestry (Korea, China, Vietnam, and Taiwan)) from Northside Hospital in Atlanta, Georgia with the approval of the Northside Hospital Institutional Review Board (NSH #804), as described previously. In short, we dissected centrally located villus parenchyma tissue, avoiding the decidua, chorion, and amnion, at five non-adjacent locations of each placenta and pooled them into a single sample to homogenize the cell type composition among samples. The sample information with 26 demographic variables was provided under GEO (GSE66622).

The miRNA raw datasets generated during the current study are available in the Sequence Read Archive (SRA) with accession number PRJNA602793. Detailed sample information can be found both in the PRJNA602793 and in the Gene Expression Omnibus (GEO) under accession number GSE66622. The quantified gene expression data from the previous study (Hughes et al., 2015) were downloaded from GEO under accession number GSE66622.

3.2.2 Construction of indexed small RNA-Seq Libraries

Total RNA was isolated from human placenta using Trizol (Invitrogen, USA) according to the manufacturer's directions. RNA quality was determined using Agilent 6000 Nano chips on an Agilent Technologies 2100 Bioanalyzer. Samples with RIN>6 were selected for library construction. Sequencing libraries were constructed in a single batch according to the TruSeq SmallRNA Libary Preparation guide (Illumina) with no modification. In each sequencing lane on an Illumina HiSeq X10 platform, we pooled 14 samples with different index sequences and carried out 150-bp paired-end sequences. We randomly distributed 10 African American, 10 European American, 10 South Asian American, and nine East Asian American samples into three sequencing lanes. We further

included the tenth East Asian American sample (X14) into each of the three lanes to control for potential sequencing artifacts among the lanes (Table 1).

3.2.3 miRNA expression quantification

The trimmed raw sequences that were at least 17bp long were mapped allowing no mismatches to the sequences of known and novel mature miRNAs determined as described in the previous section, extended 8nt both up and downstream, using the Bowtie algorithm (Langmead, Trapnell, Pop, & Salzberg, 2009). To quantify the miRNA expression, only the reads from the R1 strand were considered. Following the protocol described in (H. Y. Hu et al., 2009), the expression value of each miRNA was calculated as the number of reads mapping to the reference sequence for the mature miRNA. All miRNA reads' count data were log2 transformed and quantile normalized. As sequences were obtained in three batches, the batch effect, identified by principal component analysis, was removed by the removeBatchEffect (limma) algorithm (Figure 2). Based on a principal component analysis of all 42 samples, four samples were considered as outliers based on their dispersion and were removed from further analysis (Figure 2C). In total, data from 8 African American, 10 European American, 8 South Asian American and 10 East Asian American individuals were retained.

3.2.4 Construction of population dendrograms

To illustrate the extent of the expression difference among human populations, we built unrooted neighbor joining trees (nj function in R package ape) using the Euclidean distances between the mean expressions of each miRNA in each population. The expression level of each miRNA was scaled by the mean expression of this miRNA across all samples. A single population tree was generated by estimating the mean of Euclidean distance across all miRNAs. A differentially expressed (DE) tree was generated by taking the mean Euclidean distance across 93 DE miRNAs or 139 DE miRNAs (Figure 5). The same approach was applied to human genetic data with distances based on the Fst values from the 1,000 genome project (Genomes Project et al., 2015). Populations used for genetic

data included African Americans, CEPH, Telugu, Han Chinese, Southern Han Chinese, and Kinh Vietnamese to match the populations used in our study. The mean Fst value of the Han Chinese, Southern Han Chinese, and Kinh Vietnamese populations was considered as the value for the East Asian population.

3.2.5 Dating of human miRNA evolutionary ages

To estimate miRNA evolutionary age, we first downloaded all hairpin sequences for the 23 species listed in Table S4 from miRBase v22. We then mapped the hairpin sequences to all the 23 genomes using blastn with parameter -evalue 1e-5. A positive hit was called when the sequence overlapped with another species' known miRNA coordinates using bedtools to intersect with default parameters. If a miRNA pair could be blasted reciprocally, it was considered to be a one-to-one miRNA hit. Next, according to Table S4, we assigned each miRNA to an age group from 0-12, where 0 is the oldest and 12 is the most recent evolutionary age group. We considered the evolutionary age of the pre-miRNA to correspond to the evolutionary age of the mature miRNA.

3.2.6 Quantification of miRNA tissue specificity

We downloaded 87 miRNA expression datasets derived from 12 healthy tissues from miRmine database (Panwar, Omenn, & Guan, 2017). Tissue specificity was measured by the tau index (Yanai et al., 2005) using quantile normalized log2-RPM expression data. A tau value can range between 0 for house-keeping genes and 1 for tissue-specific genes.

3.2.7 miRNA target prediction and functional analysis of miRNA targets

We downloaded miRNA predicted targets using miRNAtap (Simpson, 2019), requiring a predicted target to be identified by at least three of the five following methods: "pictar", "diana", "targetscan", "miranda" and "mirdb". To get potential targets for each cluster of population-associated miRNAs, we assessed if predicted target genes for cluster of miRNAs were negatively regulated compared to non-target genes (Spearman correlation coefficient rho<-0.5; one-sided Wilcoxon rank test $p< 0.05$). To identify potential targets

of SON-associated miRNAs, we required the absolute expression fold change of miRNA targets to be greater than 0.1 and the direction of change to be opposite to the direction of miRNA expression difference. The normalized expression of miRNA targets was obtained from (Hughes et al., 2015).

Gene ontology (GO) and pathway enrichment tests were processed with Metascape (Zhou et al., 2019) with all genes expressed in human placenta used as a background. The GO terms with $-\log_{10}(p) > 2$ were reported as significantly enriched.

3.2.8 miRNA disease association analysis

We used the database TAM2.0 (J. Li et al., 2018) to analyze the functional and disease associations of miRNAs with cancer-related terms masked. All expressed human placenta miRNAs were taken as a background. The terms with nominal $p<0.05$ were reported as associated diseases.

Chapter 4. Comparative transcriptome analysis quantifies molecular phenotypes in embryonic development and evolution

Comparative analyzes of embryonic development and evolution have been relatively well studied at the level of mRNA expression. However, most studies have focused either on gene expression patterns within a given species or compared the conservation of developmental stages within the concepts of the hourglass model. We still lack a direct overview of the relationship between morphology and temporal change in gene expression within a broader scale of evolutionary history. We want to understand the molecular forces contributing to the maintenance and divergence of the developmental expression profiles and estimate their association with morphological and functional organ evolution. Here, we analyzed the conservation and variation of the developmental mRNA expression profiles by comparing the developmental changes among nine evolutionarily distinct species: from oyster to mouse. To summarize our results in brief, we found that, among the nine species, the most conserved gene expression module contains genes highly expressed in 2-cell/8-cell stage, with gene functions associated with spliceosome-related cellular processes. Differences in tissue composition and asynchronous developmental rates of different organs could account for these expression differences among species. To explore these organ-dependent developmental effects, we further combined gene and miRNA expression datasets covering the development of three major organs (brain, liver, and heart) from early organogenesis to adulthood in eight species (human, rhesus monkey, marmoset, mouse, rat, rabbit, opossum, and chicken) to investigate the evolutionary features underlying the evolution of miRNA expression patterns and their regulatory relationship with the target genes within individual organs. The results showed that nearly half of the miRNAs are expressed in all three tissues, with the brain having more tissue-specifically expressed ones. Overall, 35%-68% of expressed miRNAs show dynamic developmental changes in at least one tissue for each species. About 50% of miRNAs displaying conserved expression profile among tissues and species could be associated with RNA processing and synaptic signaling functions. By contrast, miRNA showing

differential developmental expression trajectories among organs within species did not display functional conservation of their targets between human and mouse.

4.1 Results

4.1.1 Constructing gene expression atlas of embryonic development in nine species

We investigated the relationship between ontogeny and phylogeny of chordate species at the level of molecular phenotype using RNA-sequencing (RNA-seq) data collected over the entire course of embryonic development in eight chordate species of different organizational complexity (amphioxus, ciona, zebrafish, two species of frogs, turtle, chicken, and mouse) and an outgroup species (oyster) (H. Hu et al., 2017; G. Zhang et al., 2012). The data were collected at 11 to 20 developmental stages in each species and measured in duplicates (Figure 16).

Within species, differences among developmental stages explained approximately 90% of total expression variation (Figure 17; Figure 18). By contrast, other factors, such as biological replicates and sequence coverage, explained 1% and 0.6% of the variation, respectively (Figure 17). Similarly, based on principal component analysis of all samples for each species, a large variance (>50%) in gene expression can be explained by different developmental stages (Figure 18; the sum of the variation from PC1 and PC2).

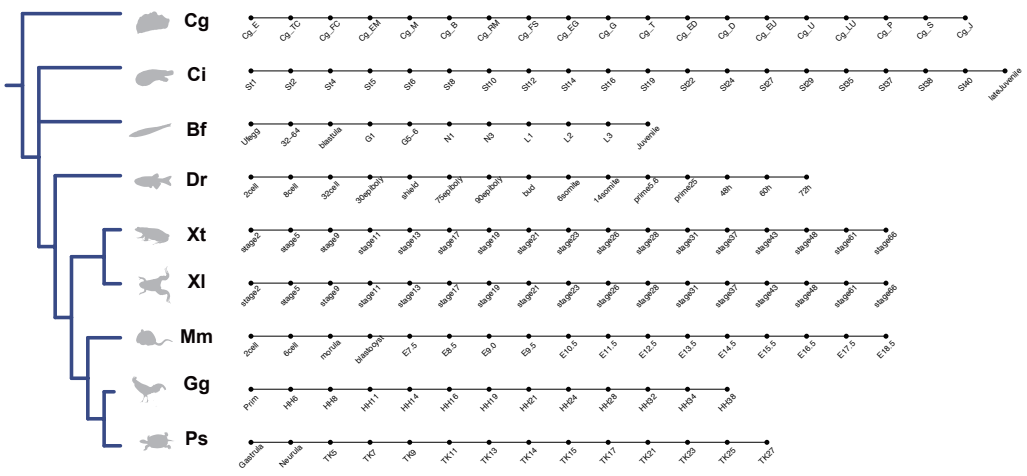

Figure 16. Species phylogeny and illustration of sampled developmental stages for each species.

Experimental scheme showing the phylogenetic relationship among species (dark blue dendrogram) and numbers of sampled developmental stages. The abbreviations here and throughout the figures indicate species: Cg – oyster (Crassostrea gigas), Ci – ciona (Ciona intestinalis), Bf – amphioxus (Branchiostoma floridae), Dr – zebrafish (Danio rerio), Xt – Western clawed frog (Xenopus tropicalis), Xl – African clawed frog (Xenopus laevis), Ps – turtle (Pelodiscus sinensis), Gg – chicken (Gallus gallus), Mm – mouse (Mus musculus).

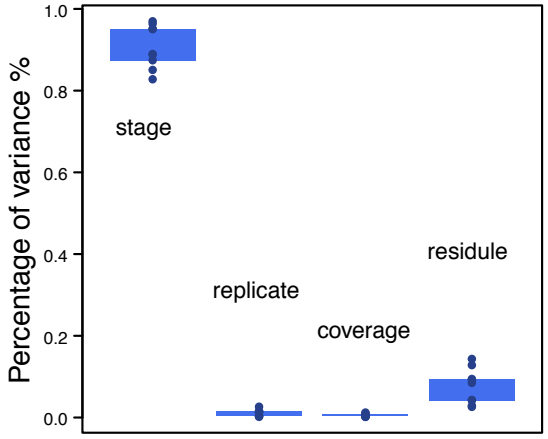

Figure 17. Variance explained by different factors based on ANOVA.

Boxes represent the interquartile values of the variance proportion explained by the factors marked near each box among nine species. Dots represent the mean explained variance proportion for each species.

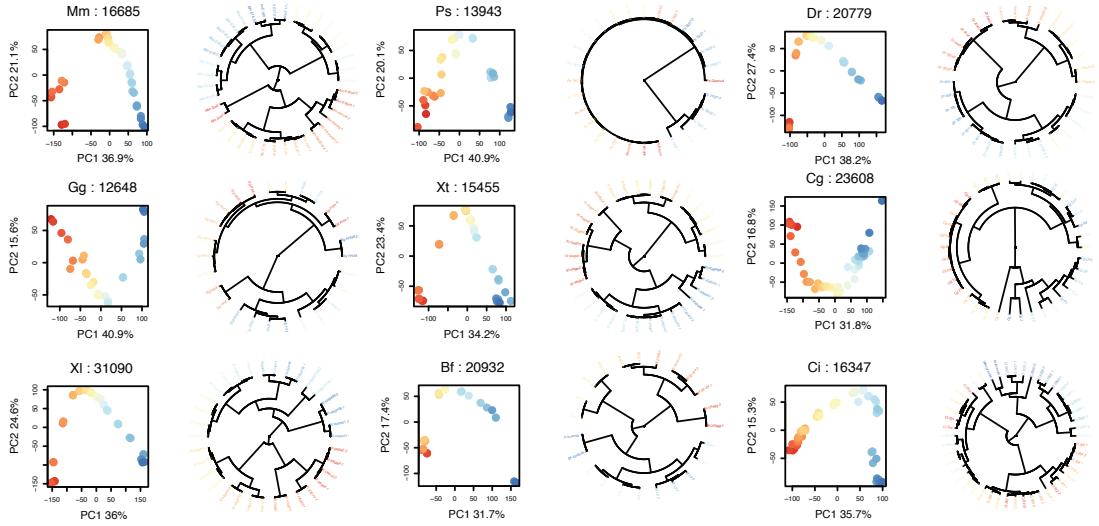

Figure 18. PCA plot and dendrogram based on expression variation within each species. Dots represent samples, color represents developmental stages (red – early, blue – late). Dendrograms were constructed using 1- Spearman correlation coefficient dissimilarity matrix. Panel titles show abbreviated species' identifiers and the number of expressed genes with FPKM > 1 in at least one developmental stage. The proportion of variance explained by each principal component is shown in the axis.

Further, we identified genes with expression levels dynamically changing with developmental stages, termed as dynamically changed genes, using polynomial regression models according to the method described in Somel et al. (Somel et al., 2009). On average, 85% of the detected genes showed significant expression differences between developmental stages in each species (dynamically changed genes) (polynomial test, Benjamini-Hochberg (BH) corrected p <0.05; Figure 19).

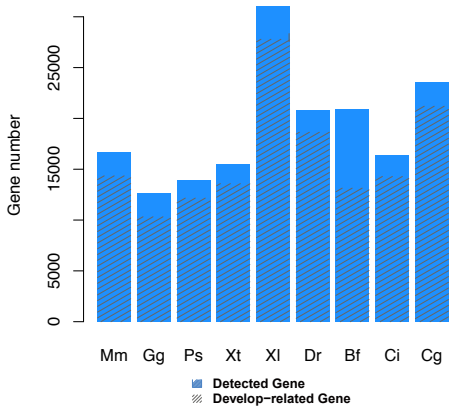

Figure 19. The number of detected and dynamically changed genes in each species.
The blue bar represents the number of detected protein coding genes; the shadow bar represents the number of dynamically changed genes

4.1.2 Aligning developmental gene expression patterns between species

To compare gene expression differences between species in the series of developmental stages, we first want to find the best temporal alignment between each species pair. We defined genes preferentially expressed at a particular developmental stage as stage-associated genes in each species following the method described in (H. Hu et al., 2017; J. J. Li et al., 2014). The stage-associated genes are potentially related to organ development, such as growth and reproduction in early embryonic stage. For each species, we required the gene on the stage with FPKM (fragments per kilobase of exon model per million reads mapped) > 2 and Z-score > 1.5. These criteria guaranteed that the genes are more highly expressed in one stage than in others and are expressed distinguishable from the background. On average, we identified 65%-90% of expressed genes as stage-associated genes for given species.

We then searched for the best temporal alignment between each species pair using the Needleman-Wunsch algorithm (J. J. Li et al., 2014) based on stage-associated genes. Surprisingly, despite substantial differences in organizational complexity, an approximately linear alignment of developmental stages was a dominant stage-matching pattern in each pair of species (Figure 20; Figure 21). These results were robust to the

selection of samples and stage-associated genes used in the alignment (bootstrap by randomly selected biological replicate for each developmental stage, Figure 21).

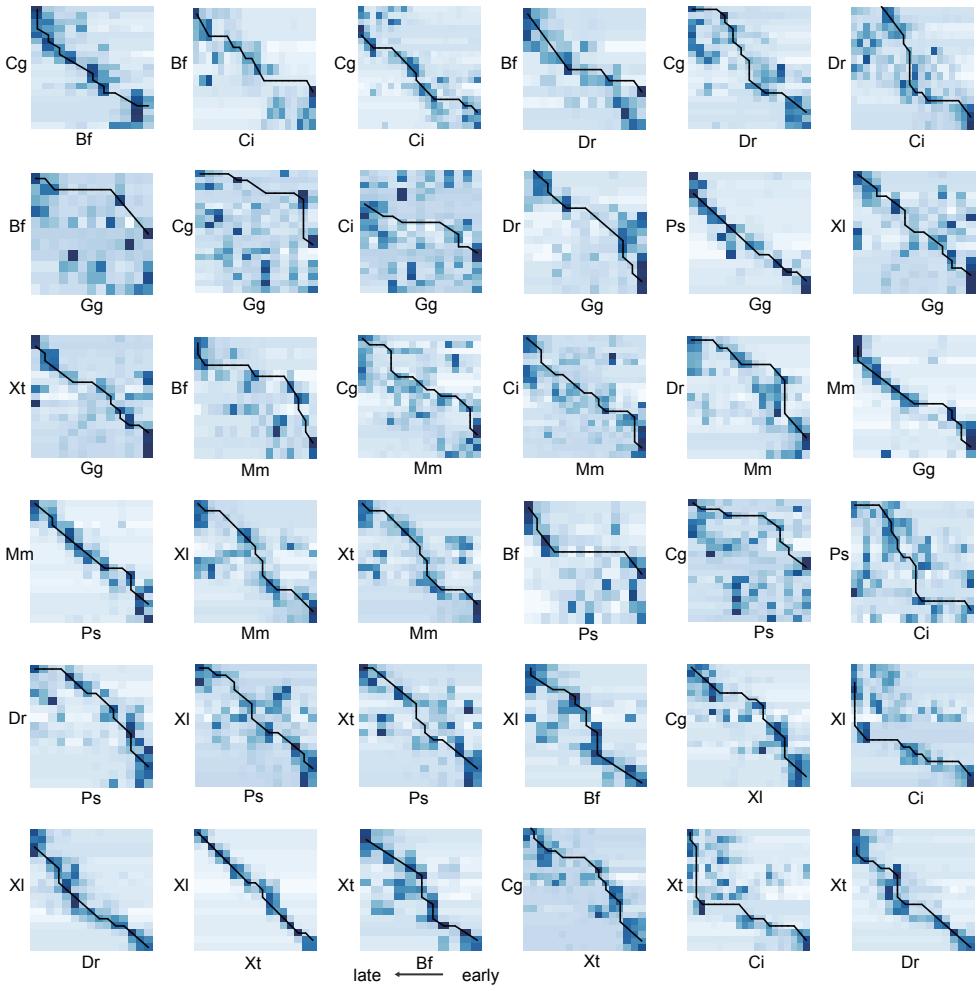

Figure 20. Heatmap of the pairing scores reflecting the overlap of stage-associate genes. Darker shade of blue represents greater overlap of stage-associate genes. Black line represents the optimal alignment path calculated using the Needleman-Wunsch algorithm.

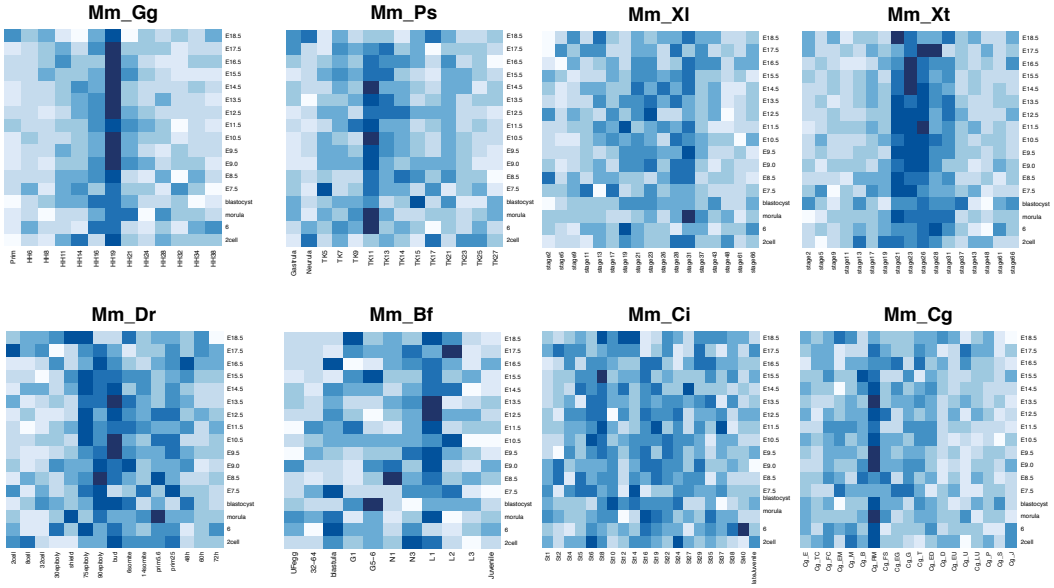

Figure 21. Heatmap of the average pairing scores calculated by randomly assigning stage-associated tags to expressed genes.

The heatmap shows pairing scores based on the average of 500 iterations of random assignments of stage-specific gene labels. Panel titles show abbreviated species identifiers. Darker shades of blue represent greater overlap between stages.

4.1.3 Comparative developmental dynamics of orthologous genes' expression among vertebrates and chordates

To further compare the gene expression changes across developmental stages among nine species, we first needed to align developmental gene expression profiles to the same time scale in all species. To do so, we utilized the mouse embryonic developmental stage as a baseline and aligned the other eight species' stages expression profiles to the mouse ones, as presented in Figure 22. Precisely, we mapped the postconceptional age of 33 stages cumulatively interpolated from eight species to the mouse age scale by interpolating them from the full mouse developmental curve fitted using a cubic smoothing spline with ten degrees of freedom. Then we compared the developmental gene expression curves among nine species based on z-transformed expression of each gene interpolated at these temporarily aligned 33 stage points.

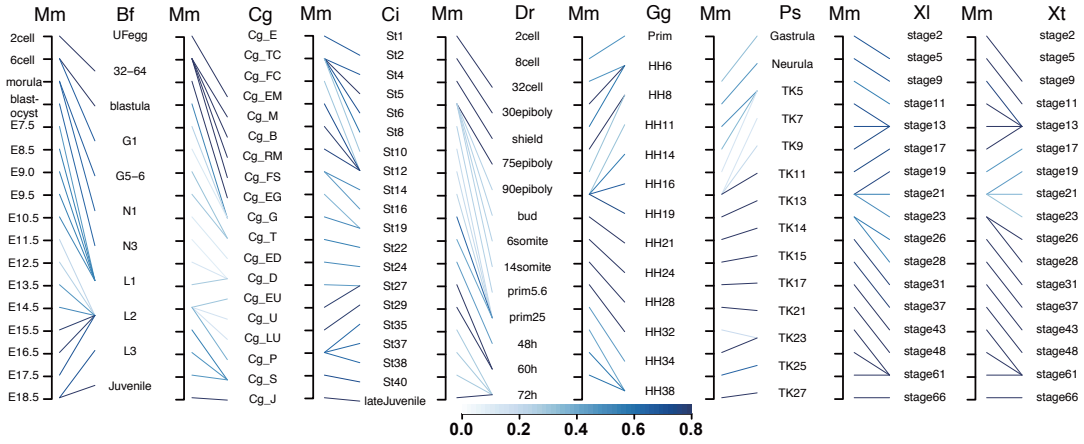

Figure 22. Pairwise alignment of developmental stages between mouse and the other species
Thick lines represent alignments supported by technical replicates. Thicker lines represent more stable alignments calculated by random subsampling of samples 500 times.

Next, we examined the developmental expression trajectories of the 244 orthologous genes classified as developmentally related in all nine species using the near-linear species alignment (Figure 22). These genes were grouped into six clusters based on their developmental profiles in unsupervised clustering analysis (Figure 23) using the complete linkage method (hclust function in R) with (1 – rho) as the distance measure, where rho is the Spearman correlation coefficient. Remarkably, 66% of the genes fell into a single cluster representing a declining expression pattern conserved across all nine species (CL9.1) (Figure 24A). By contrast, expression trajectories in other clusters differed more between species, with the magnitude of the differences directly proportional to the corresponding phylogenetic distances (Figure 25).

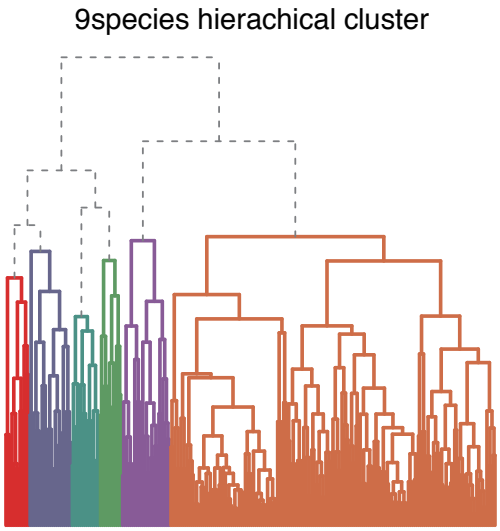

Figure 23. Clustering of dynamically changed gene expression trajectories in chordates. Hierarchical clustering of 244 nine-chordates orthologous genes based on the correlation of their developmental expression profiles. Colors represent six main clusters.

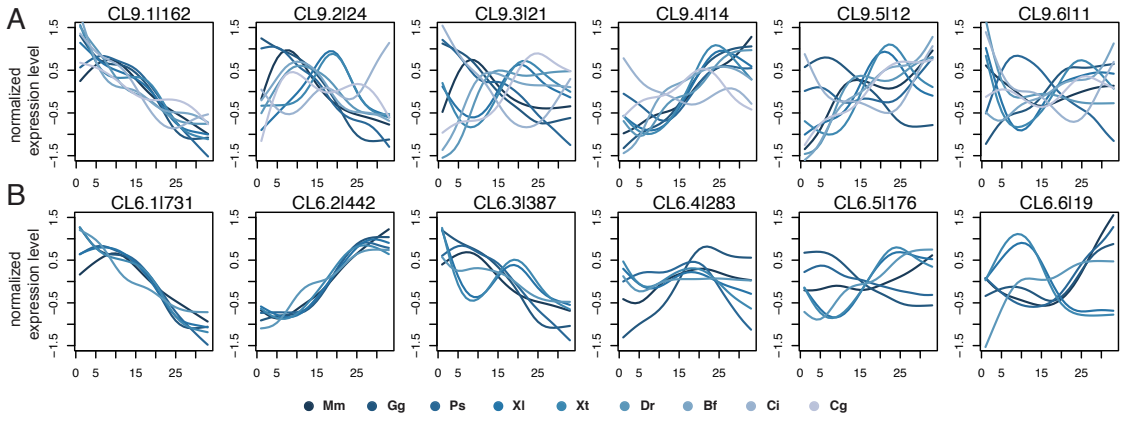

Figure 24. Clustering of gene expression profiles during embryonic development

A. Developmental gene expression patterns in each of six clusters based on nine-species gene orthologs (CL9). Colors represent species. Panel titles show cluster identifiers and number of contained genes. **B.** Developmental gene expression patterns in each of six clusters based on six vertebrate species gene orthologs (CL6). Panel titles show cluster identifiers and number of contained genes. Each line shows the average expression levels of all genes in a group for each species. X-axis indicates the temporal rank of developmental stages. Y-axis indicates the normalized gene expression levels (z-transformed).

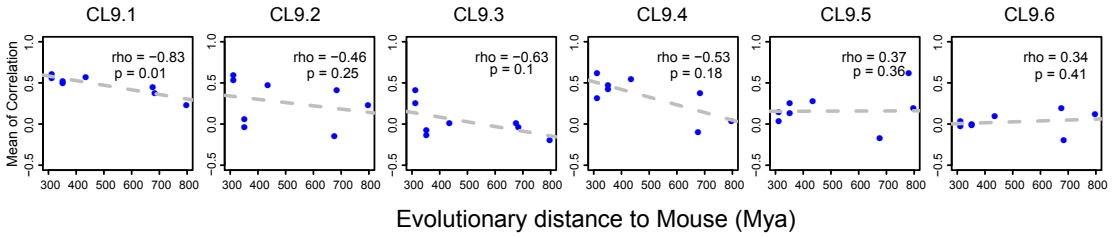

Figure 25. Relationship between the similarity of dynamically changed gene expression patterns and phylogenetic distances.

The relationship is shown for each of six clusters of dynamically changed genes orthologous among nine species. The expression similarity was calculated as the mean of Spearman correlation coefficients between mouse and non-mouse expression trajectories. The x-axis shows the phylogenetic distance (Mya) to the mouse. The Y-axis shows the average gene expression similarity.

Repetition of the same analysis, now based on expression profiles of 2,038 dynamically changed orthologs shared among six vertebrate species, revealed a CL9.1-like developmental pattern represented by a single cluster (CL6.1) (Figure 24B). Genes in clusters CL9.1 and CL6.1 overlapped significantly (73.5% of CL9.1) and were enriched in similar biological processes, including spliceosome and RNA processing (Figure 26; Table 4; Table 5). The second-largest cluster found in vertebrates (CL6.2) showed an opposite, ascending expression pattern conserved among species and contained genes enriched in signaling pathways, such as cGMP-PKG signaling pathway, oxytocin signaling pathway, and renin secretion. Notably, this cluster overlapped with the chordate cluster CL9.4, where the ascending pattern was conserved among the seven species, excluding ciona and oyster (Figure 24A). By contrast, genes in other clusters showed developmental profiles divergent among species, including patterns involving partial similarity, development shifts, and inversions.

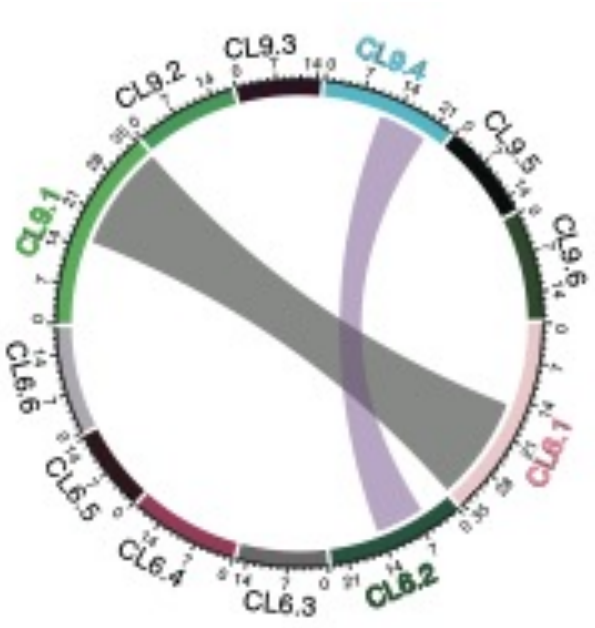

Figure 26. The circus plot shows the significance of gene overlapping between chordate (nine species) and vertebrate (six species) developmental expression clusters.

Fisher's exact test was applied to test the significance of pairwise overlapping between two clusters. The connections between entities display the value of -log10(p) after Bonferroni correction. Only the pairs of $p<0.00001$ are shown. The wider chords indicate greater and more significant overlap between genes in chordate (CL9) and vertebrate (CL6) developmental clusters.

Cluster	Pathway	P-value BF correction <0.05
1	Spliceosome	1.17E-16
1	Ribosome biogenesis in eukaryotes	1.38E-08
1	Pyrimidine metabolism	0.010666557
1	RNA transport	4.03E-05
1	RNA degradation	0.030735139
1	RNA polymerase	0.017144867
2	GABAergic synapse	0.012342312
2	Oxytocin signaling pathway	0.000928933
2	cGMP-PKG signaling pathway	0.025054815
2	Gap junction	0.044979311
2	Focal adhesion	0.000144876
2	Complement and coagulation cascades	0.002405745
2	Renin secretion	0.032170789
2	PI3K-Akt signaling pathway	0.007817167

2	ECM-receptor interaction	0.000728605
3	Ribosome	1.65E-07
3	Proteasome	0.022630144
5	Mineral absorption	0.044836792

Table 4. KEGG pathways enriched in genes contained in six developmental expression clusters based on six vertebrate species

The P-values of hypergeometric test were calculated using the *phyer* function in R package for KEGG pathway enrichment without controlling the false discovery rate and adjusted for multiple testing using Bonferroni (BF) correction. All expressed orthologous dynamically changed genes were considered as background genes. The pathways were downloaded from the KEGG website for mice.

Cluster	Pathway	P-value without correction <0.05
1	Spliceosome	0.023450918
2	Ribosome	0.047955162
2	Aminoacyl-tRNA biosynthesis	0.047955162
3	Thyroid hormone synthesis	0.045697618
3	Endocytosis	0.036820573
4	Oxytocin signaling pathway	0.014292659
4	cGMP-PKG signaling pathway	0.005042017
4	Renin secretion	0.005042017
4	Valine leucine and isoleucine degradation	0.005042017
5	Non-alcoholic fatty liver disease (NAFLD)	0.025744843
5	Alzheimer’s disease	0.047859084
5	Pentose phosphate pathway	0.006518905
5	Oxidative phosphorylation	0.025744843
5	Huntington’s disease	0.047859084
5	Parkinson’s disease	0.009243697
6	Hepatitis C	0.016806723
6	Measles	0.016806723
6	Non-alcoholic fatty liver disease (NAFLD)	0.049173554
6	Influenza A	0.016806723

Table 5. KEGG pathways enriched in genes contained in six developmental expression clusters based on nine chordate species

The P-values of hypergeometric test were calculated using the *phyer* function in R package for KEGG pathway enrichment without controlling the false discovery rate and adjusted for multiple testing using Bonferroni (BF) correction. All expressed orthologous dynamically changed genes were considered as background genes. The pathways were downloaded from the KEGG website for mice.

4.1.4 Comparative developmental expression dynamics of all mouse gene ortholog in chordates

Restricting our analysis to genes having annotated orthologs in all six or all nine examined species inevitable leads to strong bias towards the most conserved fraction of the protein coding genes. To include genes having no such annotated orthologs in each of the examined species into the comparative analysis of developmental expression profiles, we examined the expression trajectories of all 103,728 genes from nine species either directly present in the mouse genome or annotated to have a mouse ortholog and displaying significant developmental expression alteration in at least one of the species. The gene expression patterns of these genes were all aligned to the mouse development stages according to the matched stage relationships shown in Figure 22. An unsupervised clustering of the expression profiles of these genes using weighted correlation network analysis (WGCNA) based on z-transformed gene expressions yielded 20 clusters (Figure 27). Similar to the clustering based on the cross-species orthologous genes, monotonically decreasing and increasing expression patterns occupied large proportion of the developmental profiles accounted for 34% (CL20.2-20.3) and 29% (CL20.1) of all clustered genes, respectively. In further agreement with the result based on gene orthologs, spliceosome genes in all species were significantly enriched in clusters showing the descending expression pattern, whereas genes involved in ECM-receptor interaction and neuroactive ligand-receptor interaction were significantly enriched in clusters showing the ascending expression pattern (Figure 28).

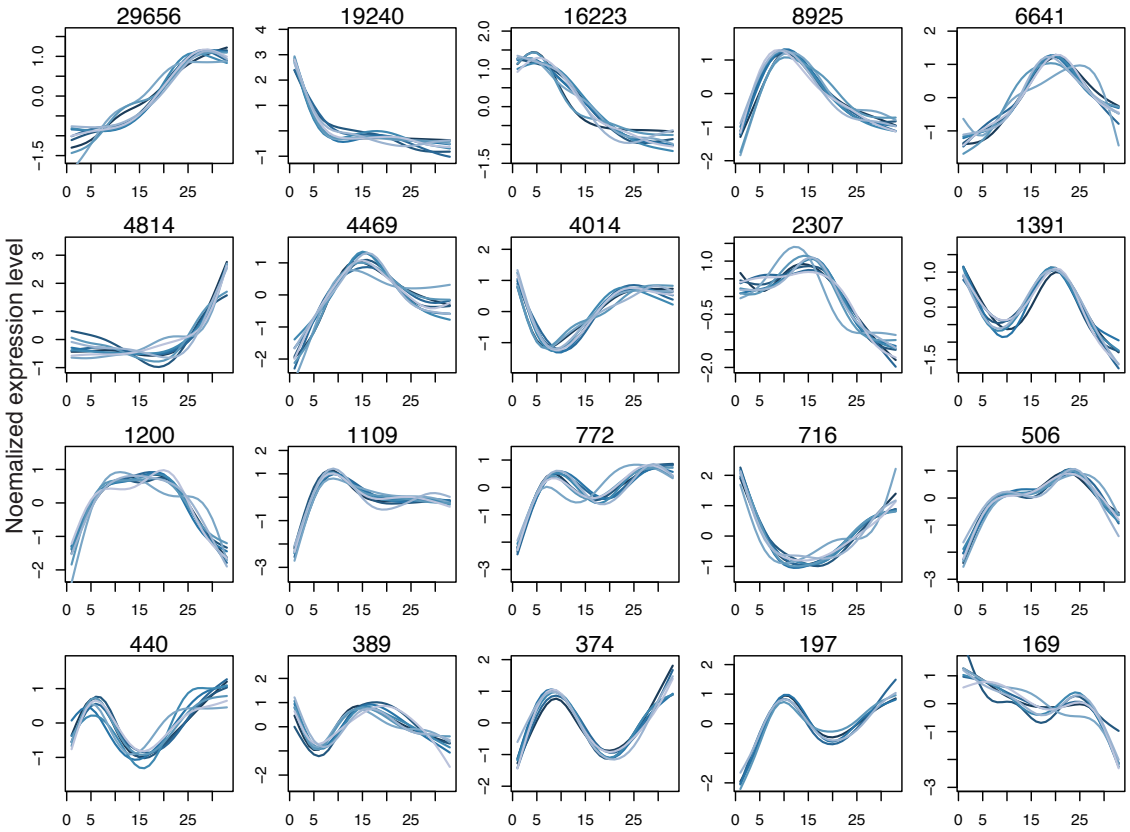

Figure 27. Clustering of developmental gene expression profiles based on all 103,728 genes with annotated mouse orthologs.

These 103,728 genes are the union of all dynamically changed mouse genes and mouse orthologs. The panels show twenty co-expression modules were grouped by weighted correlation network analysis (WGCNA) based on z-transformed gene expressions. Each line shows the average expression levels of all genes in a group for each species. The color from dark blue to light blue indicates species from Mouse to Oyster following the evolutionary tree.

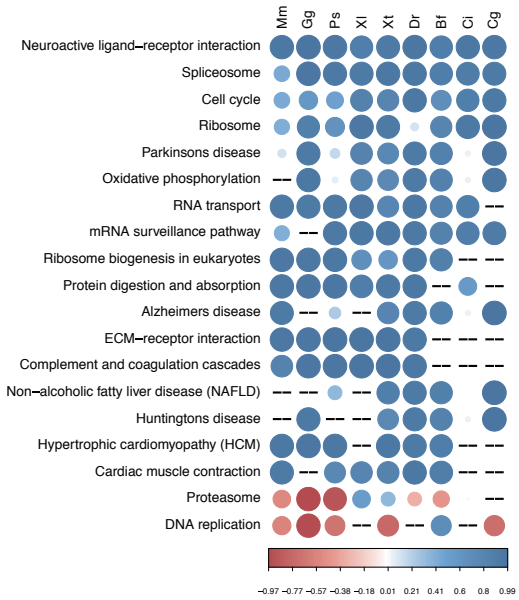

Figure 28. Expression trajectory correlation among species of genes within top 20 enriched KEGG pathways.

Hypergeometric test adjusted by Bonferroni correction was applied for each species in each gene group separately. The enriched KEGG pathway was selected with P<0.05 after correction and at least enriched in two species. Dot with the same color indicates that the mean of gene expression in enriched pathways for each species has high similarity. Blue represents the correlation between pathway genes and cluster 1 genes (which is increasing with developmental stage) is higher.

Moreover, the genes with descending expression pattern (CL20.2+CL20.3) were enriched on housekeeping genes and were consistent with previously reported results that evolutionarily older genes tend to be expressed at earlier developmental stages (Gao et al., 2018). Furthermore, a study examining embryonic gene expression reported the existence of such a pattern in isolated mouse tissues (Cardoso-Moreira et al., 2019), suggesting that descending expression reflects changes within embryo tissues rather than changes in embryo composition. Further, genes with ascending expression patterns (CL20.1) were enriched in tissue-specific expression genes (Figure 29 A-B, Fisher's exact test $p < 0.001$). It is consistent with the results based on gene orthologs (Figure 29 C-D, Fisher's exact test $p<0.001$).

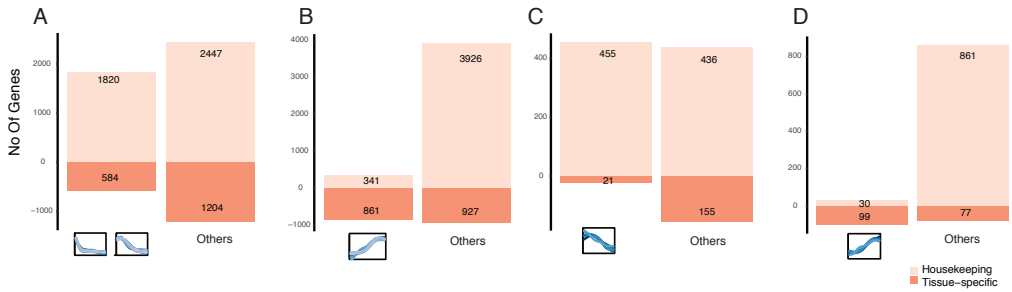

Figure 29. Tissue specificity of genes with ascending and descending developmental expression patterns.

Bars represent the number of tissue-specific and housekeeping genes for gene groups. Numbers within bars show the number of tissue-specific and housekeeping genes in each group. **A.** The genes with descending expression pattern (CL20.2+CL20.3) **B.** the genes with ascending expression pattern (CL20.1) **C**. The ortholog genes with descending expression pattern (CL6.1) **D.** The ortholog genes with ascending expression pattern (CL6.2)

It is worth noting that we found no conserved organ development related GO ontology modules between species (Figure 30). These variations could represent interspecific developmental heterochrony or differences in tissue composition.

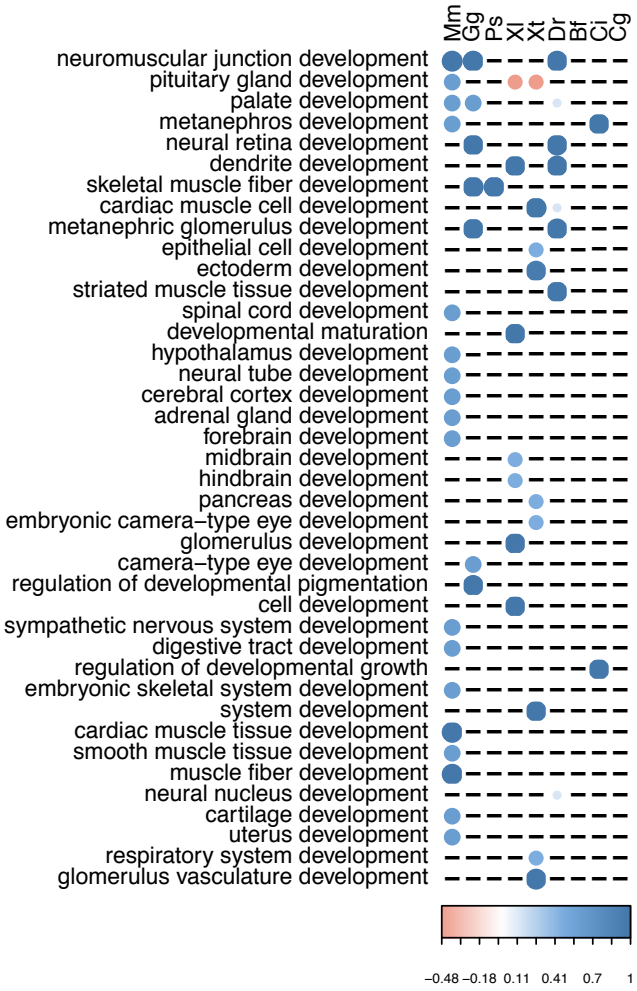

Figure 30. Expression tendency for genes enriched in "development" GO ontology term. Hypergeometric test adjusted by Bonferroni correction was applied for each species in each gene group separately. The enriched GO ontology was selected with P<0.05 after correction. Dot with the same color indicates that the mean of gene expression in enriched pathways for each species has high similarity. Blue color represents positive correlation between pathway genes and CL20.1 genes (which is increasing with developmental stage) is higher.

4.1.5 RNA processing functions are conserved among species and organs

To investigate the effects of tissue composition differences on developmental expression patterns, we took advantage of the published mouse dataset containing data

expression data for brain, liver, and heart development from early organogenesis to adulthood (Cardoso-Moreira et al., 2019). We based our analysis on expression of 10,130 genes reported in that study. Grouping these genes according to the direction of developmental gene expression changes in each tissue resulted in eight clusters. We then tested the relationship between these clusters and the ones obtained in our analysis of nine chordate species focusing on functional overlap between them. Consistent with our findings on descending expressed genes (CL20.2 in Figure 27), RNA processing-related metabolic pathways such as spliceosome, RNA transport, and mRNA surveillance pathway are significantly enriched in the group of descending genes in three organs (CL8.2 in Figure 31). This suggests that the developmental regulation of spliceosome pathway is conserved between species and organs. However, the neuroactive ligand-receptor interaction pathway, which we observed to be one of the most conserved pathways among all nine chordates with ascending expression curves, showed no enrichment in any group of genes in Cardoso-Moreira's dataset. Furthermore, descending genes overlapped substantially among datasets comprised of vertebrate genes in CL6.1 depicted in Figure 24B and the genes expressed in three mouse organs. By contrast, only a small proportion of genes overlapped between other cluster vertebrate genes and mouse organ developmental profiles (Figure 32). This lack of overlap might be due to the differences in tissue composition and the function of some regulators acting as the driving force for the diversity of gene expression in the different organs.

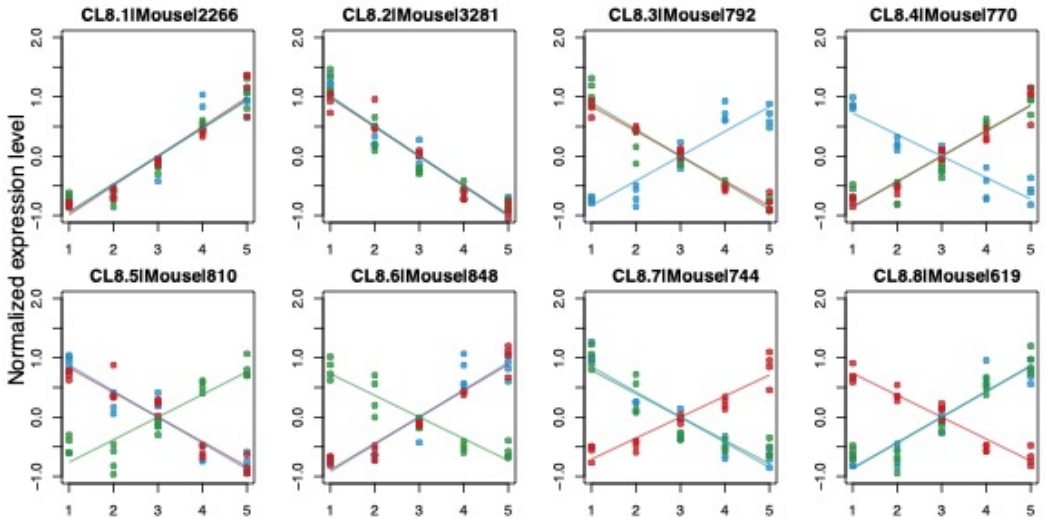

Figure 31. Clustering of mouse developmental mRNA expression trajectories from three organs.

Clustering of mRNAs based on the direction of developmental trajectories. Each dot represents the average expression levels of all genes in a group for each organ at each stage. The x-axis indices the ranked developmental stage, 1-5 represents the stage covering early prenatal, late prenatal, neonatal, juvenile, and adult. The y-axis shows the mean standardized expression level among mRNAs. The linear regression line represents the relationship between the average expression levels of all mRNAs in the cluster and assessed development stages. Blue color represents brain, green represents liver, red represents heart. The head indicates the number of genes.

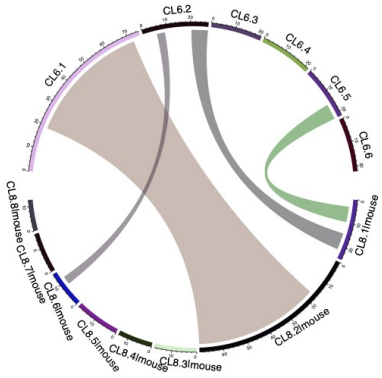

Figure 32. Circus plot depicting the significant overlapping between vertebrate and mouse developmental clusters.

Fisher's exact test was applied to test the significance of pairwise overlapping between two clusters. The connections between entities display the value of -log10(p) after Bonferroni correction. Only the pairs of $p<0.00001$ are shown. The wider chords indicate greater and more significant overlap between genes in six clusters of vertebrates (CL6) and genes from eight developmental clusters based on expression in three organs of mouse.

4.1.6 miRNAs development atlases on three vertebrate organs

Developmental expression profiles are driven by specific regulators. To investigate possible regulators of detected developmental alterations functioning as the driving force for the diversity of gene expression profiles in the different organs, we focused on miRNA, a key post-transcriptional regulator (Carroll, 2008). Role of miRNA regulation in early embryonic development, such as maternal to zygotic expression, is well established, but its role in driving developmental profiles in individual organs is not well assessed. To address this question, we generated miRNA expression developmental atlas in three organs – brain, heart, and liver – for seven mammalian species and one bird species as an outgroup. The mammals include three primates (human, rhesus macaque, and marmoset), three glires (mouse, rat, and rabbit), a marsupial (short-tailed opossum), and the bird is the red jungle fowl, henceforth referred to as 'chicken' (Figure 33A). We generated small RNA libraries for five developmental stages: early prenatal, late prenatal, neonatal, juvenile, and adult (Figure 33). Because marsupials give birth to very undeveloped young (Ferner, Schultz, & Zeller, 2017), all stages sampled in opossum are postnatal. The 401 small RNA libraries that comprise this cross-species miRNA developmental resource are described in Table S2.

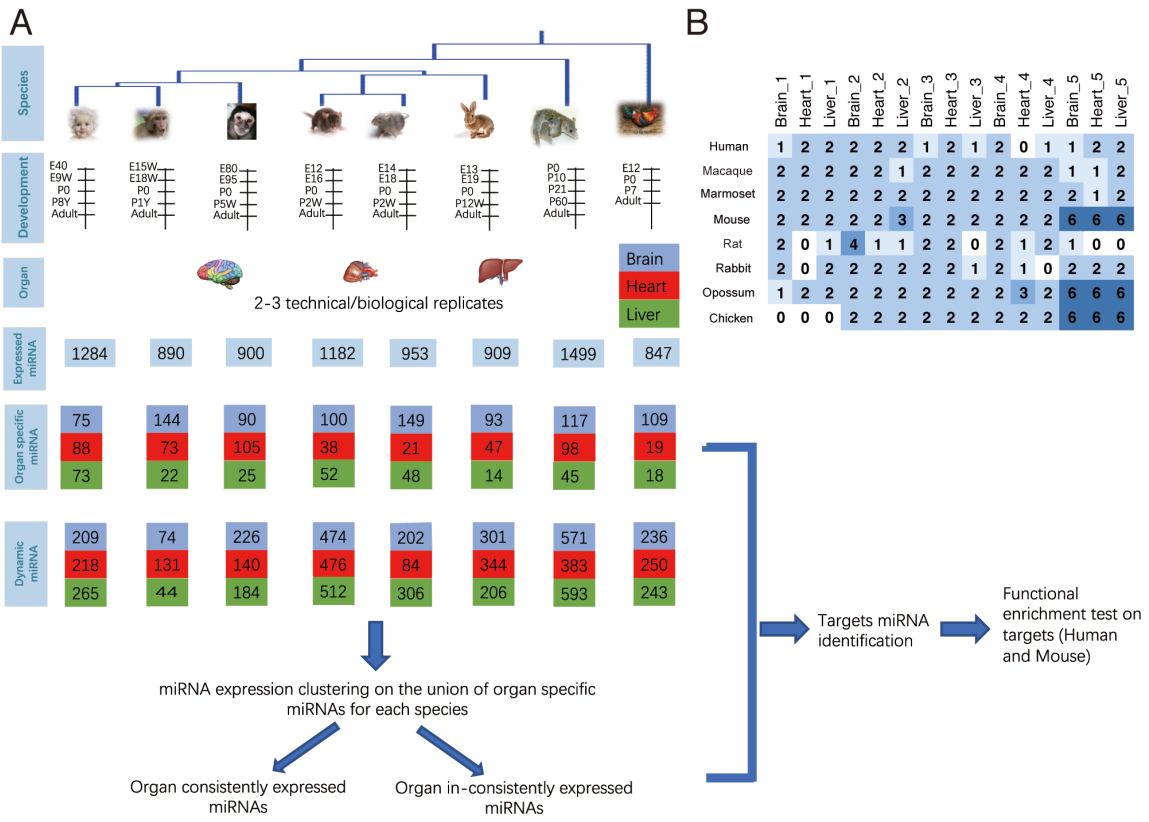

Figure 33 Species phylogeny tree and sample number for each stage

A. Workflow of the study with the illustration of species phylogeny tree, developmental stage. The number on each box represents the number of miRNAs in each identified category. **B.** Sample number on each stage for each tissue. The number of biological samples on each developmental stage of each organ. 1-5 represents the developmental stage from early embryonic development to adulthood.

The quality and depth of miRNA annotation differs significantly between mammalian species, so our first step was to annotate known and novel miRNAs contained in our expression data in each species using mirDeep2 (Friedlander et al., 2012). We identified between 900 (marmoset) and 1,499 (opossum) expressed miRNAs per species (i.e., covered by at least 100 sequencing reads across all samples within a species) (Figure 34). The high number of miRNAs in the opossum is consistent with a previous cross-species comparison of miRNAs in adult tissues (Meunier et al., 2013). Among the identified miRNAs, ~120-600 are novel mature miRNAs (Figure 34). As expected, the number of novel mature miRNAs follows the previously available resources, with human and mouse having lower numbers of novel miRNAs.

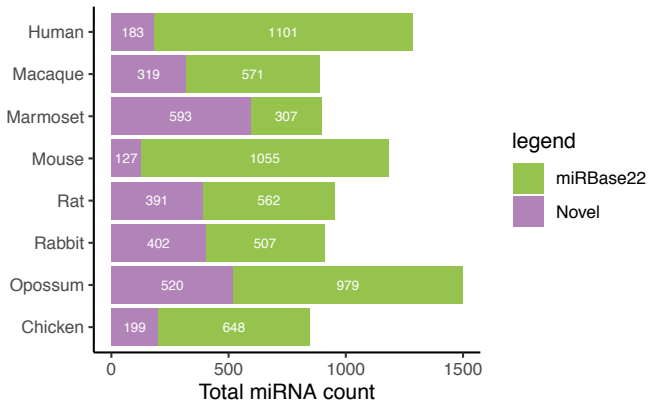

Figure 34. Expressed miRNA number in each species with total mapped reads counts >100. The expressed miRNAs were counted as the ones with a total mapped count of sequence reads >100. Green represents known miRNAs; purple indicates novel miRNAs.

Consistent with previous studies, we found that novel miRNAs are typically expressed at lower levels across species than known miRNAs (Figure 35) and in a smaller number of samples (Figure 36). Novel miRNAs are also evolutionarily young, with many being species-specific (Figure 37). Nevertheless, our set of novel miRNAs contains some that are conserved over long evolutionary periods. Even in humans, one of the best-annotated species, we identified 17 novel miRNAs that are also present in primates, and two that are also present in glirs (mouse, rat, and rabbit). Another distinctive feature is that novel miRNAs are more likely to overlap with repeat elements than known miRNAs (transposable elements and low complexity DNA sequences, Figure 38). Different types of repeat elements contribute differently to known and novel miRNAs in different species (Figure 39). Despite these differences, there are many common features between known and novel miRNAs, including the fact that for both groups, most (>85%) precursors are in intronic or intergenic regions, which has been widely observed (Olena & Patton, 2010). The result is also consistent across the eight species (Figure 40).

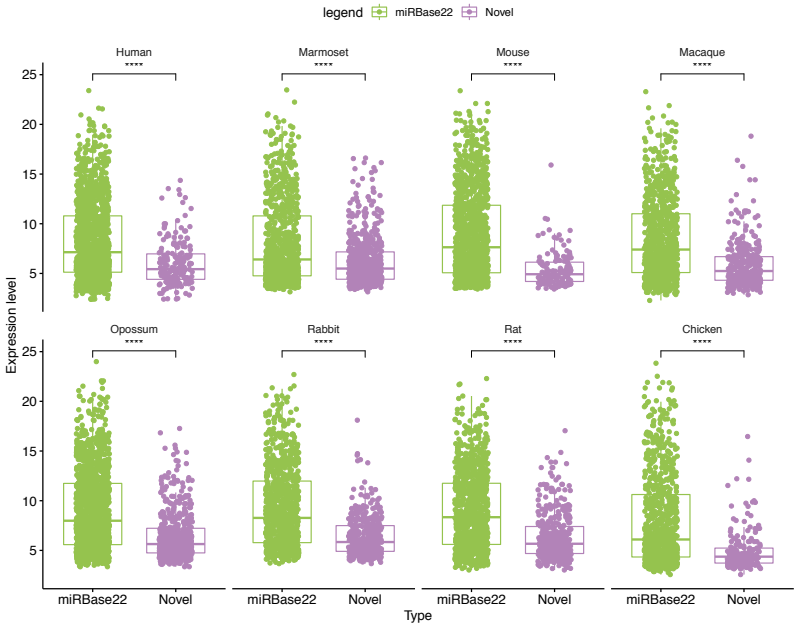

Figure 35. The distribution of miRNA expression values for known and novel miRNAs in each species.

The miRNA expression levels were estimated and normalized by TPM (transcript per million). Green represents known miRNAs; purple indicates novel miRNAs. Asterisks indicate the significance of the difference (two-sided Wilcoxon test, * represents nominal p<0.05).

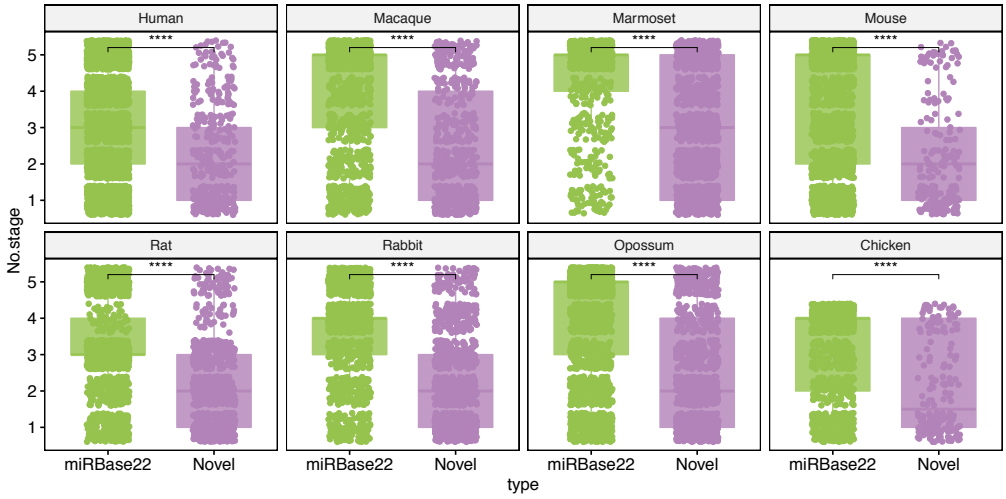

Figure 36. Coverage of expressed stages for each miRNA.

Green represents known miRNAs; purple indicates novel miRNAs. Asterisks indicate the significance of the difference (two-sided Wilcoxon test, * represents nominal p<0.05).

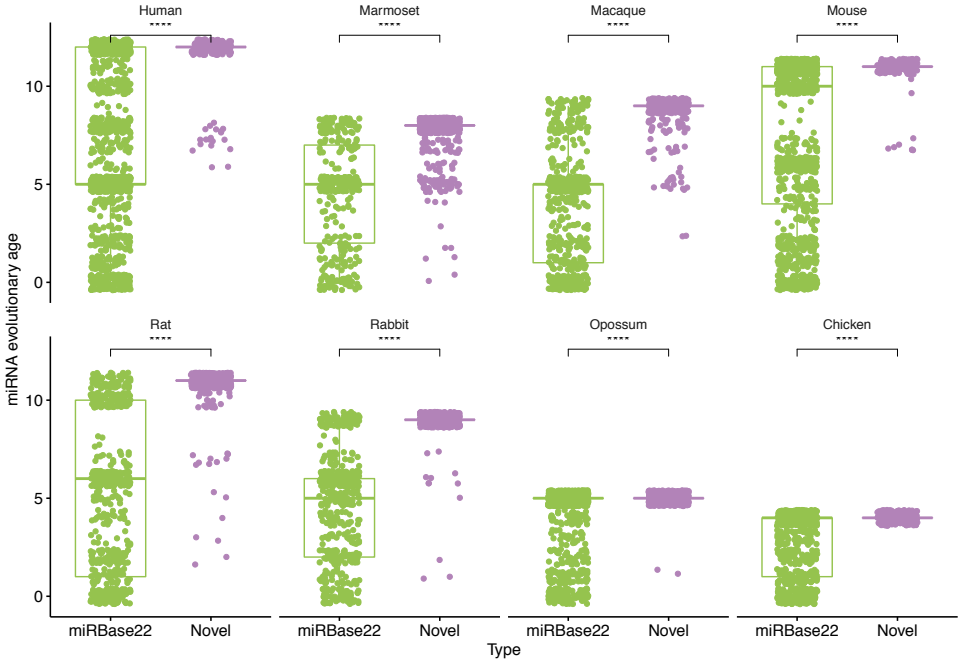

Figure 37. miRNA evolutionary age for known (green) and novel (purple) miRNA.
The age scale extends from the ancestry (age 0) to species-specific miRNA. Green represents known miRNAs; purple indicates novel miRNAs. Asterisks indicate the significance of the difference (two-sided Wilcoxon test, * represents nominal p<0.05).

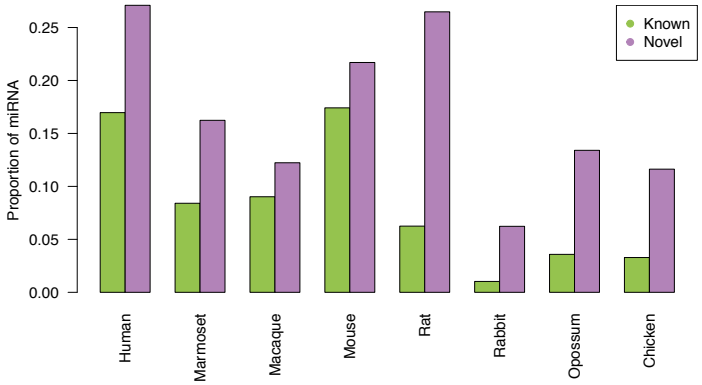

Figure 38. The overall proportion of miRNAs identified repeat elements by RepeatMasker.
Each color bar represents the proportion of miRNA identified as repeats. The purple color represents novel miRNA, the green color represents the known miRNAs.

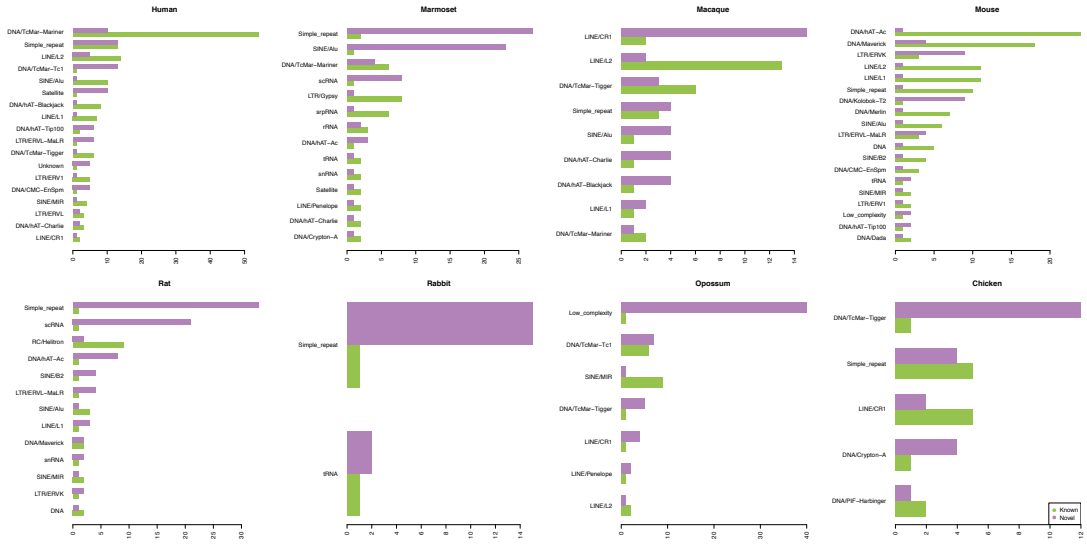

Figure 39. The proportion of miRNAs overlapping with repeat elements detected using RepeatMasker.

Each color bar represents the proportion of miRNA overlapped with repeat elements. The purple color represents the proportion for novel miRNA, while the green color describes the known miRNAs.

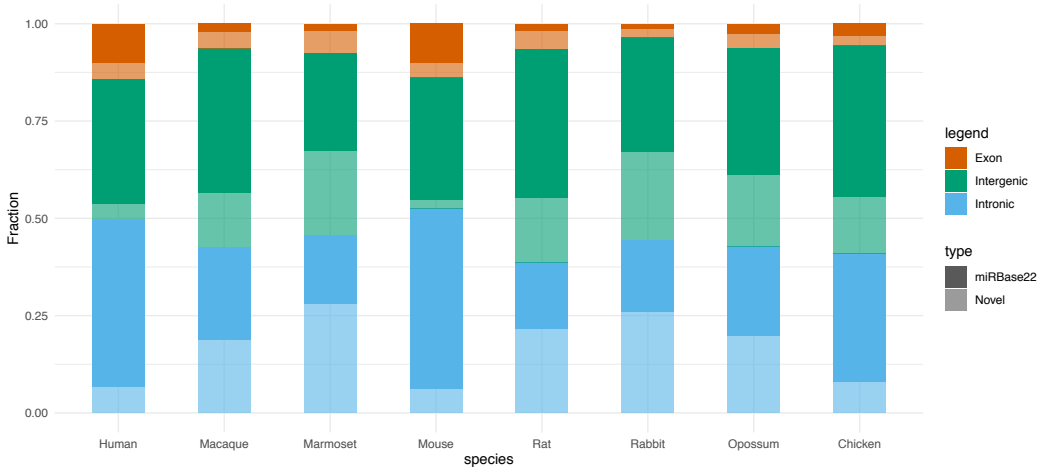

Figure 40. Genomic location of miRNA precursors.

Each color bar represents the proportion of miRNA precursors mapped to a specified type of genomic location. The light color represents the proportion for novel miRNA, while the dark color represents the known miRNAs.

4.1.7 Organ-specific miRNA

Previous studies suggested that ancient animal miRNAs play a conserved role in regulation of development timing, such as miR-100, miR-125 and let-7 which are expressed broadly in fly and vertebrate development (Christodoulou et al., 2010). Here we aimed to investigate how exactly organ-specific miRNAs evolve and their role in establishing tissue identity. The miRNA with expression greater than two (CPM>2) in at least four samples was considered as expressed miRNA of an organ. Only miRNAs expressed in a specific organ were defined as organ-specific miRNAs.

We found that novel miRNAs were not the dominant source of organ-specific miRNAs in each species (Figure 41). These organ-specific miRNAs tend to be expressed at low levels (Figure 42), evolutionarily young (Figure 43), and are at least 65% of organ-specific miRNAs are expressed in at least two developmental stages (Figure 44). Then, we only selected organ-specific miRNAs expressed in at least three developmental stages to test the role of tissue identity and function. We grouped these miRNAs by sequence, and a total of 30 miRNA families were explicitly expressed in one organ and at least three species (Table 6). Compared to liver and heart, brain has more organ-specific miRNAs (Figure 45; Table 6). The organ-specific miRNAs we identified contained some known examples, such as hsa-miR-129-1-3p and hsa-miR-124-5p specifically expressed in brain, and hsa-miR-208b-3p explicitly expressed in heart (Ludwig et al., 2016). In addition, miR-1 and miR-133 were previously found to be heart-specific and miR-137, miR-153 and miR-128 were found to be brain-specific (Z. Guo et al., 2014).

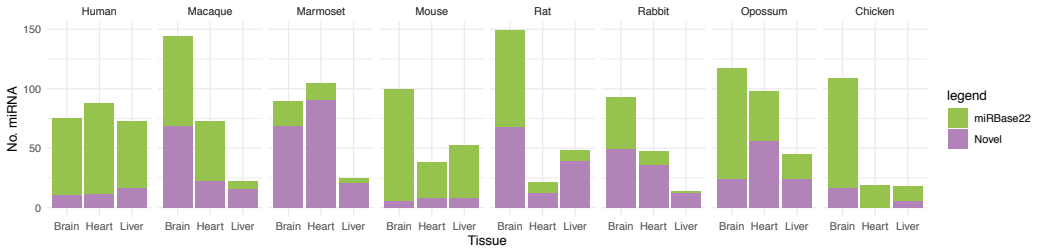

Figure 41. The number of organ-specific miRNAs in each species.

Each color bar represents the number of organ-specific miRNA. The purple color represents the number of novel miRNAs, while the green color represents the known miRNAs.

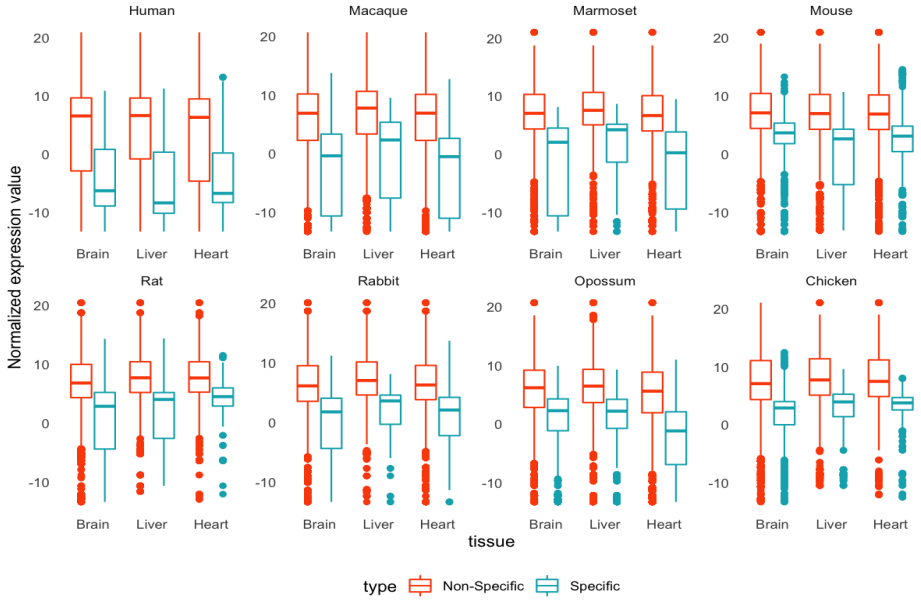

Figure 42. Expression levels of organ-specific miRNAs and non-organ specific miRNAs.

Y-axis represents the quantile normalized miRNA expression. miRNA expression for each organ was considered as the mean of expression on each development stage. Red represents non-organ-specific miRNA expression; Blue represents organ specific miRNA expression.

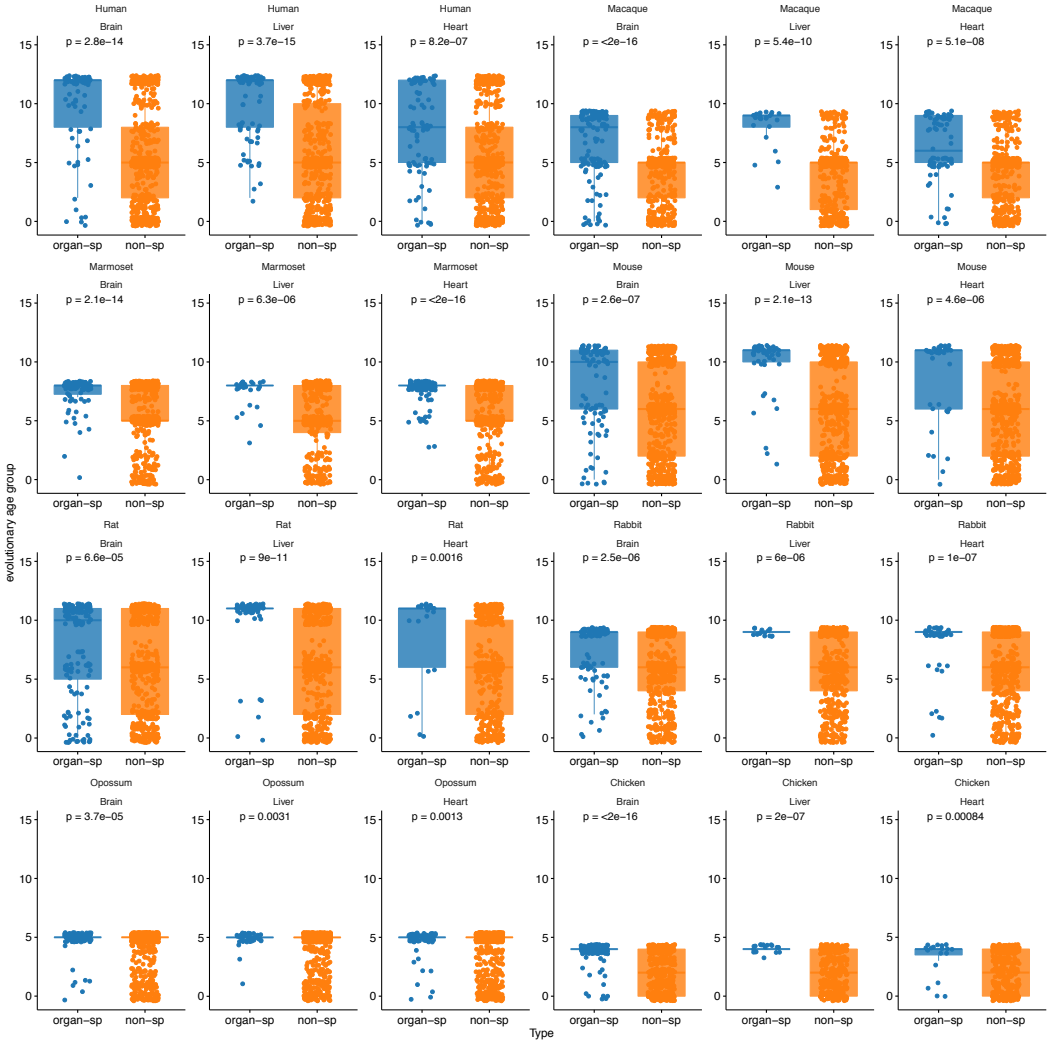

Figure 43. The distribution of miRNA age for organ specific and non-organ specific miRNAs.

Blue indicates the distribution of the evolutionary age group on organ-specific miRNAs; Orange indicates the distribution of the evolutionary age group on non-organ specific miRNAs; Y-axis represents the age group.

Figure 44. Developmental persistence of organ-specific miRNAs.

Colors indicate the number of stages each organ-specific miRNA was detected in each tissue of each species. 1-5 represents the number of stages each miRNA expressed. Each bar represents the percentage of organ-specific miRNAs.

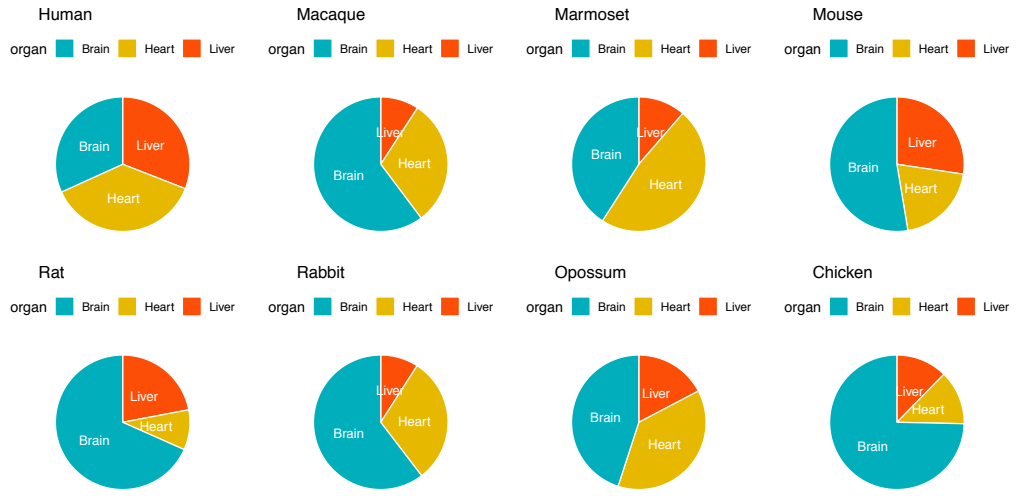

Figure 45. Proportion of organ-specific miRNAs in each species

Each slice of the pie chart represents the percentage of expressed organ-specific miRNA in each organ.

Organ specific miRNAs	Organ
gga-miR-551-5p; mmu-miR-551b-5p; mdo-miR-551b-5p; rno-miR-551b-5p; ocu-miR-551b-5p	Brain
mml-miR-217; rno-miR-217-5p; mdo-miR-449c-3p	Brain
hsa-miR-129-1-3p; gga-miR-129-3p; mmu-miR-129-1-3p; mml-miR-129-2-3p; mml-miR-129-1-3p; rno-miR-129-1-3p	Brain
hsa-miR-6794-5p; mml-miR-6794-5p; cja|chr22:12056225..12056288:+	Brain
gga-miR-383-3p; mmu-miR-383-3p; rno-miR-383-3p; ocu-miR-383-3p	Brain
mmu-miR-873a-5p; mml-miR-873-5p; rno-miR-873-5p; cja-miR-873	Brain
hsa-miR-137-5p; gga-miR-137-5p; gga|chrUn_NT_466449v1:12114..12173:+*; mmu-miR-137-5p; mml-miR-137-5p; rno-miR-137-5p; ocu-miR-137-5p	Brain
mmu-miR-325-5p; rno-miR-325-5p; ocu-miR-325-5p	Brain
mmu-miR-1251-5p; rno|chr7:32983166..32983226:-; ocu-miR-1251-5p; cja-miR-1251	Brain
gga-miR-449b-5p; gga-miR-449c-5p; rno|chr2:44897501..44897563:+; ocu-miR-449b-5p	Brain
gga-miR-184b-3p; mml|chr19:15693879..15693958:-; cja|chr10:131636604..131636677:-	Brain
hsa-miR-128-2-5p; gga-miR-128-2-5p; mmu-miR-128-2-5p; mml-miR-128b-5p; rno-miR-128-2-5p; ocu-miR-128b-5p	Brain
mmu-miR-1264-5p; rno|chrX:118131823..118131886:+; ocu-miR-1264-5p	Brain
hsa-miR-3085-5p; mml|chr9:93399332..93399394:-*; cja|chr12:84956783..84956844:-	Brain
mmu-miR-135a-2-3p; rno-miR-135a-3p; ocu-miR-135a-2-3p	Brain
hsa-miR-124-5p; rno-miR-124-5p; ocu-miR-124-5p	Brain
gga-miR-135a-1-3p; mml-miR-135a-1-3p; rno|chr8:114847395..114847455:+	Brain
mml-miR-153-3p; rno-miR-153-3p; cja-miR-153	Brain
mmu-miR-448-3p; mml-miR-448; rno-miR-448-3p; ocu-miR-448-3p;	Brain
mmu-miR-135b-3p; mdo-miR-135b-3p; mml-miR-135a-2-3p; rno-miR-135b-3p; cja|chr19:27527472..27527533:-*	Brain
mmu-miR-325-3p; rno-miR-325-3p; ocu-miR-325-3p	Brain
mmu-miR-3059-5p; mml-miR-3059-5p; rno|chr7:43324717..43324776:+; ocu-miR-3059-5p	Brain
hsa-miR-302d-3p; hsa-miR-302a-3p; hsa-miR-372-3p; ocu-miR-302c-3p; cja-miR-302a-3p; cja-miR-302d-3p	Heart
hsa-miR-3154; mml-miR-3154; cja|chr1:172180009..172180064:-	Heart
hsa-miR-208b-5p; hsa-miR-208a-5p; mmu-miR-208a-5p; mmu-miR-208b-5p; mml-miR-208b-5p; rno-miR-208a-5p; ocu-miR-208a-5p	Heart

hsa-miR-1-5p; gga-miR-1a-2-5p; gga-miR-1b-5p; gga-miR-1a-1-5p; mmu-miR-1a-1-5p; mmu-miR-1a-2-5p; mdo-miR-1-1-5p; mml-miR-1-1-5p	Heart
hsa-miR-133a-5p; mml-miR-133c-5p; rno-miR-133a-5p; ocu-miR-133a-5p;	Heart
hsa-miR-152-5p; mml-miR-152-5p; ocu-miR-152-5p;	Heart
hsa-miR-208b-3p; hsa-miR-208a-3p; mmu-miR-208a-3p; mmu-miR-208b-3p; mdo-miR-208a; mml-miR-208b-3p; mml-miR-208a-3p; rno-miR-208a-3p; ocu-miR-208b-3p; ocu-miR-208a-3p; cja-miR-208a; cja-miR-208b	Heart
mmu-miR-1247-3p; mml-miR-1247-3p; cja-miR-1247	Liver

Table 6. Identified organ-specific miRNAs expressed in three species and sharing the same seed region sequence.

To explore if these organ-specific miRNAs are biologically relevant to organ development, we applied Metascape to perform a functional enrichment test on the set of organ-specific miRNA targets. We used miRNAtap with at least three methods to predict human and mouse miRNA targets, which we then combined with the expression of target genes from Cardoso-Moreira's dataset (Cardoso-Moreira et al., 2019). However, compared to random miRNA-non target interactions, we only found weak repression on miRNA targets for human liver, heart, and mouse liver specific miRNAs (bootstrap the non-target genes 100 times for each miRNA; One-side Wilcoxon test; $p<0.05$; Figure 46). Further, we defined the genes with the strong repression (Spearman correlation rho< -0.5) as potential targets for organ-specific miRNAs. Overall, we found weak enrichment in GO ontology terms and KEGG pathways on strong down-regulated miRNA targets. However, we discovered that the targets downregulated by mouse liver-specific miRNAs are significantly enriched in biological functions such as dendrite development, brain development, axon guidance, and the Wnt signaling pathway, which makes sense given that those targets are highly expressed in the mouse brain and heart. To see if the lack of relevant function was due to a small number of targets, we included weakly repressed miRNA targets, defined as genes with a reverse correlation of less than 0.1. We would then discover more biological significance, but it would not be conserved between human and mouse, except for the targets of liver-specific miRNAs, which demonstrate biological relevance on brain development in both Humans and Mouse (Figure 47). Our observation shows the majority of targets predicted for each miRNA are only weakly repressed, as

reported in previous studies (Baek et al., 2008; Hausser & Zavolan, 2014; Selbach et al., 2008), and we can observe some functional miRNA regulatory effects on a phenotypic consequence from weakly repressed miRNA targets. However, we cannot evaluate the precise regulatory function of miRNAs because we found no significant repression effects on targets and didn't find the highly functional overlaps on targets between Human and Mouse. It could imply that organ-specific miRNAs are more likely to target multiple genes by chance, with possible deleterious consequences (Berezikov, 2011)

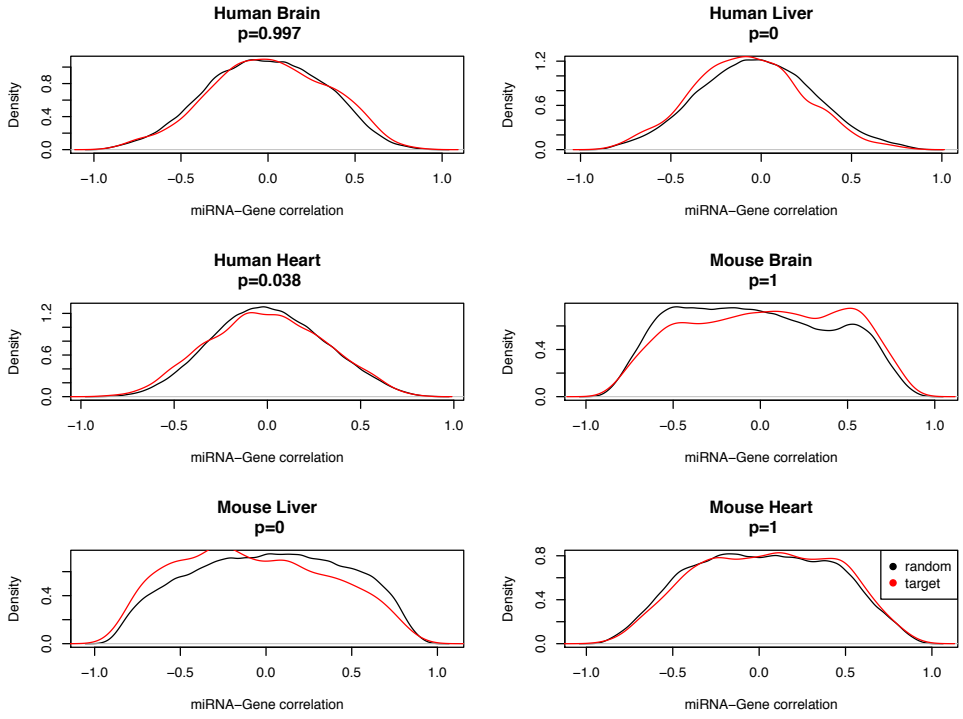

Figure 46. Distributions of miRNA-target correlations

The red line shows the distribution of Spearman correlations between miRNA-target pairs relative to miRNAs and non-target pairs. The black lines depict the random expectation estimated by the union of 100 bootstraps of miRNA-non-target pairs. The p-value based on the one-side Wilcoxon test is listed on the top of each panel. The X-axis indicates the miRNA-gene correlation.

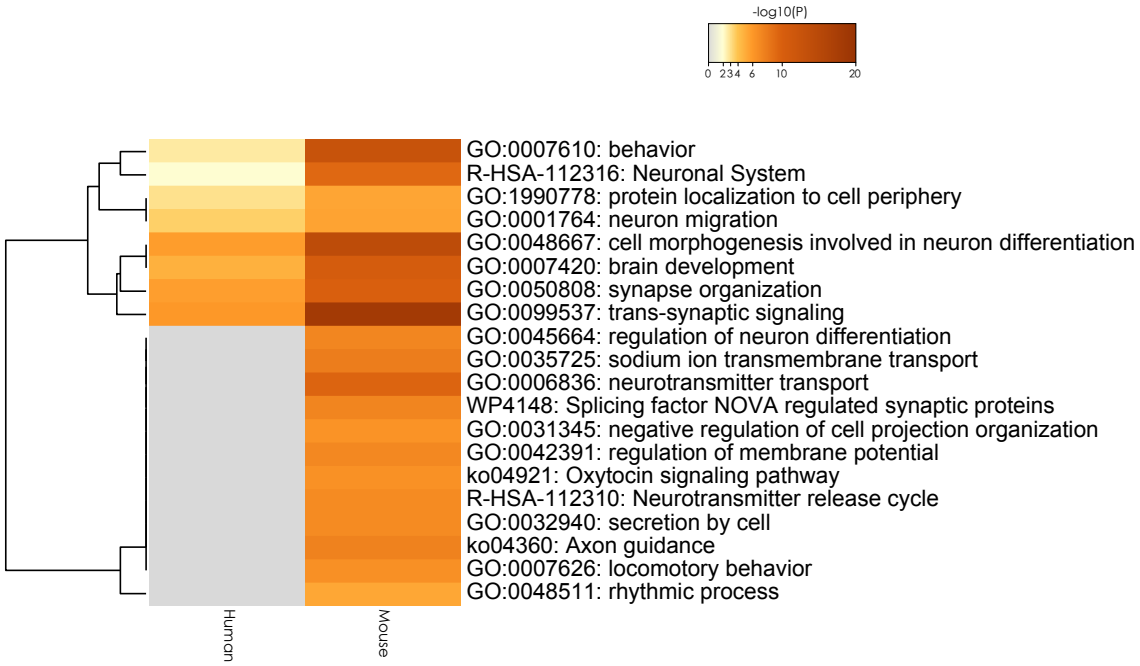

Figure 47. Gene Ontology terms enriched in the targets of liver-specific miRNAs.
Selected terms with the best p-value within each cluster, as the representative terms, are displayed in a dendrogram. The heatmap cells are colored by their enrichment p-values; light cells indicate the lack of enrichment for that term in the corresponding species. The significant terms are hierarchically clustered into a tree based on Kappa-statistical similarities among their gene memberships. 0.3 kappa score was applied as the threshold to cast the tree into term clusters.

4.1.8 Dynamics of miRNA expression during organ development

We used principal component analysis (PCA) to examine the global relationships among samples and identify the variables that explain most of the variance in gene expression among samples (Figure 48). Predictably, organ and developmental time are two main variables that contribute to expression variation. Samples are clustered by organ and are ordered from early to late development, suggesting miRNA profiles that diverge steadily during development and differ between organs. We further identified miRNAs with significant differential expression over time (termed “developmental dynamics”) using a linear regression approach with developmental stages ordered from 1 to 5 for each organ or species for each miRNA expressed at least four samples with TPM >2（$p < 0.01$; FDR <10% for species except opossum and rat liver, which is 11% and 15%, respectively;

permutation p <0.05). Except for macaque, 35%~68% of miRNAs show dynamic changes in at least one organ for each species. Only 22% of macaque miRNAs show dynamic changes (Figure 49). We also observed that there are fewer dynamic changed mRNAs than other species. But we did not find significant difference in RNA quality and sequencing reads coverage. We guess that may be due to improper sample collection. Moreover, among these dynamically changed miRNAs, only up to 23% are organ-specific miRNAs in rat identified in the previous section (Figure 50).

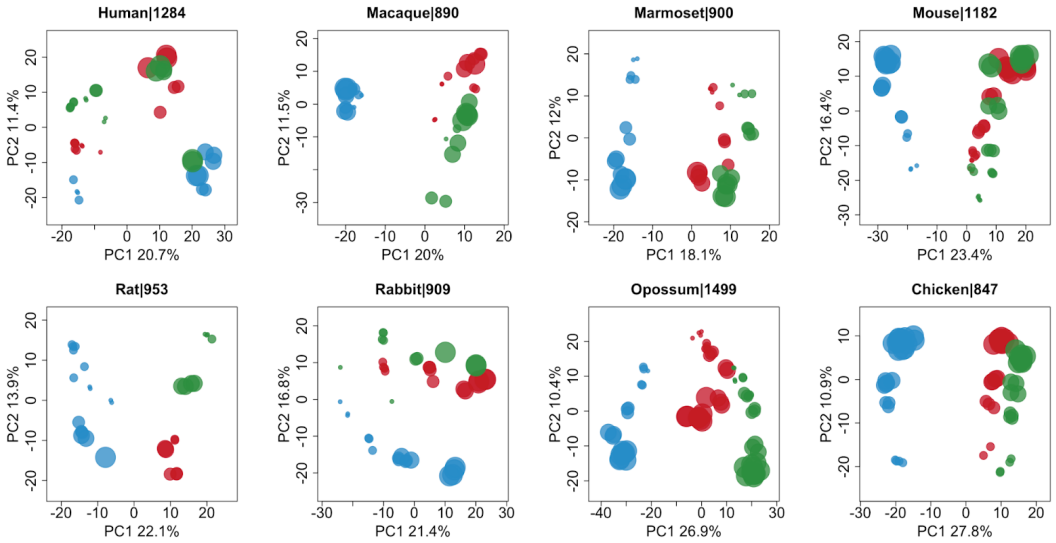

Figure 48. PCA plot for each species based on all expressed miRNA with total counts >100. Each dot represents a sample. The dot size dot represents the developmental stage from early prenatal, late prenatal, neonatal, juvenile, to adult. Blue represents the brain; green represents liver and red represents heart samples. The number separated by pipe symbol indicates the number of expressed miRNAs. The proportion of variance explained by each principal component is shown with the axis label.

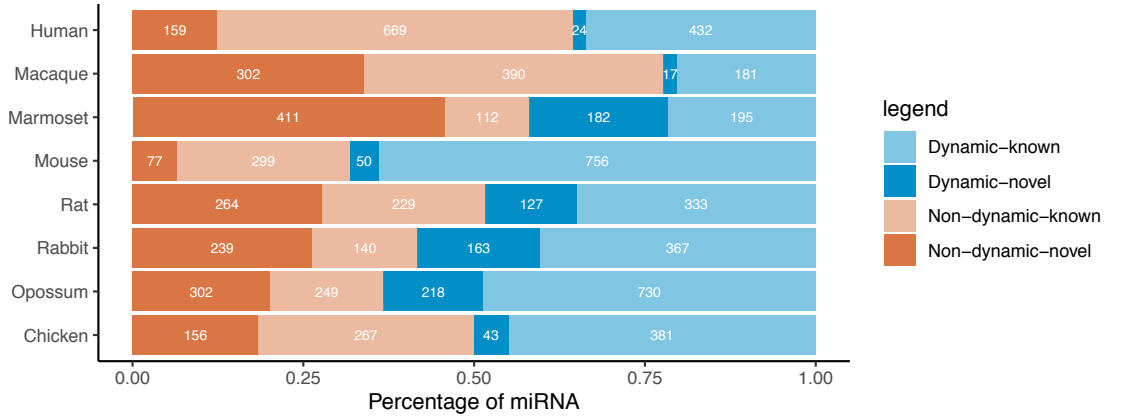

Figure 49. Proportion of total dynamically changed miRNAs.

Blue represents the proportion of dynamically changed miRNAs; Orange represents non-dynamically changed miRNAs; Dark color represents novel miRNAs; light color represents known miRNAs. The number of miRNAs is marked on each bar. The number is the union of dynamic changes at least on one organ.

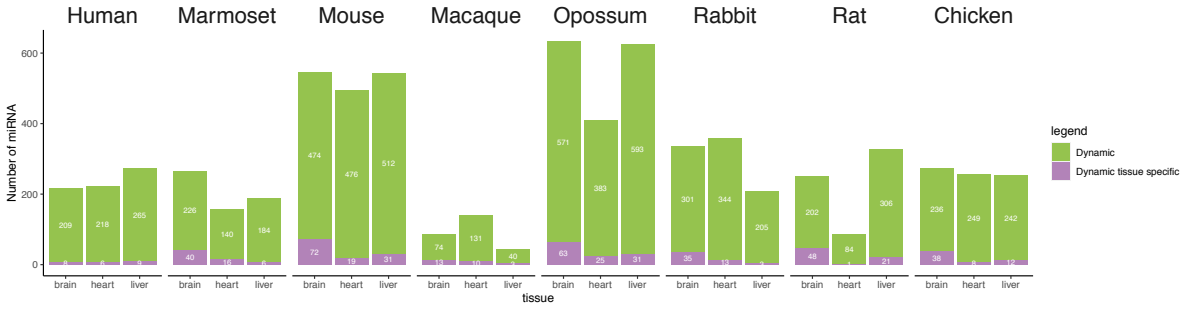

Figure 50. The number of dynamically changed miRNAs in each organ, including organ-specific ones.

Each bar with the number represents the number of miRNAs. Green represents all dynamically changed miRNAs. Purple represents dynamically changed organ-specific miRNAs.

4.1.9 Comparative dynamics of miRNA developmental expression profiles among organs

To assess the conservation and divergence of miRNA expression profiles among brain, liver, and heart between species, we used non-organ-specific miRNAs which dynamically changed at least one organ. Further, we classified them into eight groups according to the direction of expression changes along with each organ's developmental

stages (Figure 51 for Human; Figure 52 for Mouse). Two groups showed a consistent pattern of change among the three organs and were termed organ-consistent miRNAs. The rest of miRNAs displayed an anti-correlated expression profile between one organ and the other two organs, termed organ-inconsistent miRNAs. Overall, except for marmoset, rat, and opossum, more than 50% of the dynamically changed miRNAs showed consistent changes between brain, liver, and heart (Figure 53).

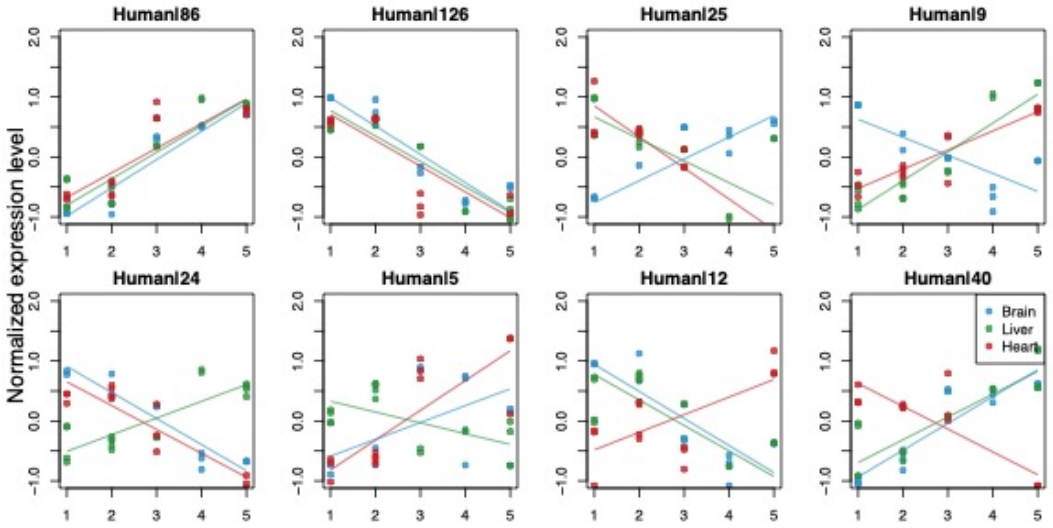

Figure 51. Clustering of human miRNAs' developmental expression profiles.
Clustering of miRNAs is based on the direction of the relationship between miRNA expression and development stages. Each dot represents the average expression levels of all miRNAs in a group for each organ at each stage. The linear regression line represents the relationship between the average expression levels of all miRNAs in a group and developmental stages. Blue represents brain, green represents liver, red represents heart. The x-axis indices the ranked developmental stage, 1-5 represents the stage covering early prenatal, late prenatal, neonatal, juvenile, and adult. The y-axis shows the mean standardized expression level among miRNAs. The number of miRNAs is indicated at the top of each panel.

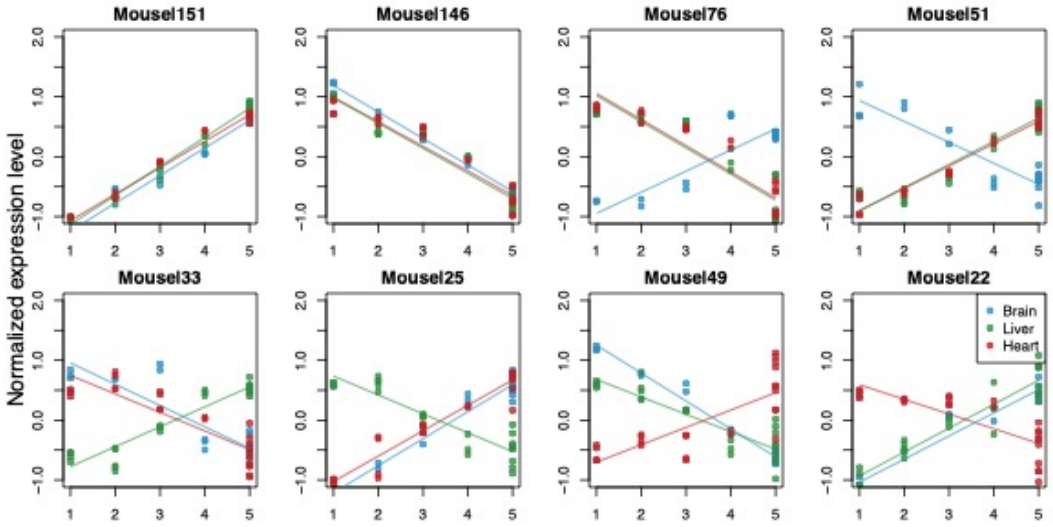

Figure 52. Clustering of mouse miRNAs' developmental expression profiles.

Clustering of miRNAs is based on the direction of the relationship between miRNA expression and development stages. Each dot represents the average expression levels of all miRNAs in a group for each organ at each stage. The linear regression line represents the relationship between the average expression levels of all miRNAs in a group and developmental stages. Blue represents brain, green represents liver, red represents heart. The x-axis indices the ranked developmental stage, 1-5 represents the stage covering early prenatal, late prenatal, neonatal, juvenile, and adult. The y-axis shows the mean standardized expression level among miRNAs. The number of miRNAs is indicated at the top of each panel.

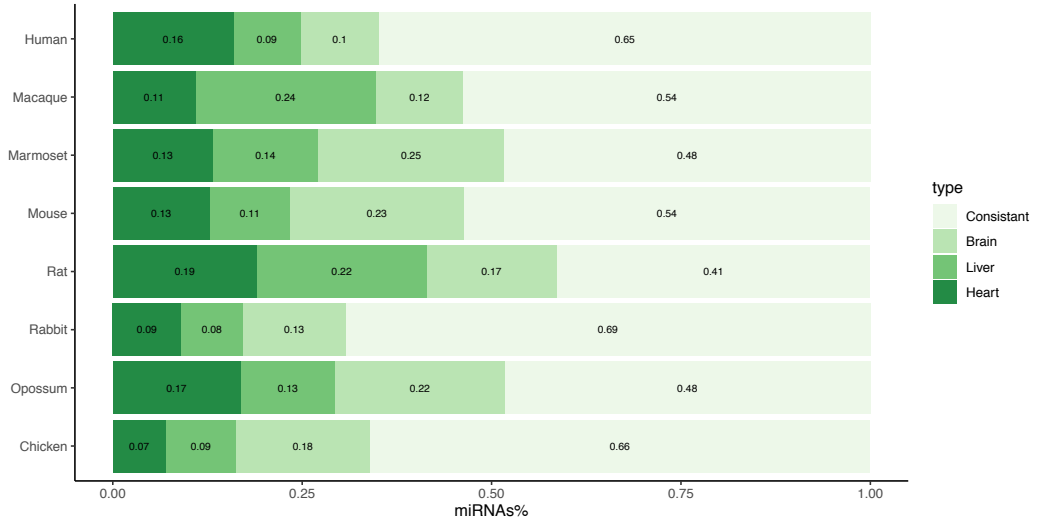

Figure 53. The percentage of miRNAs showing conserved and organ-specific expression patterns among three organs.

Light green represents dynamically changed miRNAs with same direction pattern among three organs; Darker green represents dynamically changed miRNAs with trajectories specific to one organ. The number on each bar indicates the percentage of miRNAs.

To assess whether organ-consistent miRNAs are more ancient with the importance of some origination mechanisms, we used a BLAST-based phylostratigraphic technique to calculate the evolutionary age of miRNA for each species. Each miRNA was allocated to a phylostratum, representing the miRNA's most ancient phylogenetic node (Table S4; Table S5). We used 0 as the most ancient phylogenetic node, which origins from Zebrafish and Fugu. However, our organ-consistent miRNAs did not tend to be evolutionarily older than organ-inconsistent miRNAs, although the age of organ-consistent miRNA is older than that of other expressed miRNAs (One side Wilcoxon test $p < 0.05$; Figure 54). The result indicated that the miRNAs expressed in multiple tissues are found to be old, even when they show different expression patterns across development in different organs. This observation aligns with the notion that miRNAs with divergent expression profiles between organs play roles in regulation of organogenesis.

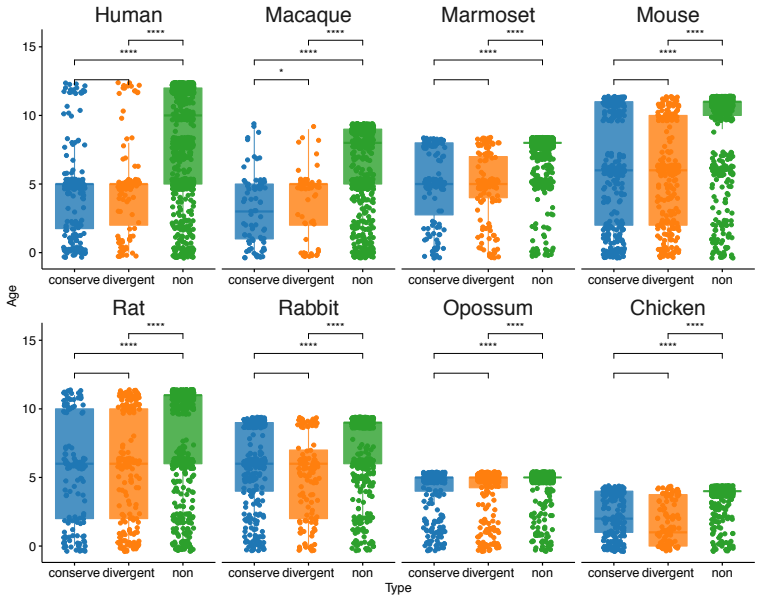

Figure 54. Age distributions of conserved and divergent miRNAs in three organs. Asterisks indicate the significance of the difference (one-sided Wilcoxon test, **represents nominal $p<0.01$; **** represents nominal $p<0.0001$). Blue represents the age of miRNAs consistently changing in development among three organs. Orange represents the age of miRNAs that show inconsistently changing among three organs, but all dynamically changes. Green represents the age of all other miRNAs.

4.1.10 Integrative analysis of miRNA and mRNA expression profiles during development identifies conserved function modules between human and mouse

Action of miRNA, as an essential posttranscriptional regulator, could be associated with the degradation or translational inhibition of mRNA by binding to the 3' UTR of target genes. This process is dominated by the seed (7-8nt) region of the miRNA. A common thought is miRNAs with similar seed target similar sets of genes and thus similar sets of biological functions or pathways. miRNA-mRNA regulatory networks occur within nearly all organisms and contribute to developmental, phenotypic, and speciation processes. To investigate whether the regulatory relationship is conserved during evolution, we sought to analyze them in two steps. First, we tested the conservation of the expression trajectories by checking whether miRNAs with the same seed, which has the same regulatory functions as usually assumed, fall into the same expression patterns. We found such significant clustering only for miRNAs with organ-consistent expression patterns between species (Figure 55A), but no significant expression profile conservation for miRNAs with organ-inconsistent developmental expression patterns (Figure 55B). Next, we investigated whether the targets of different miRNAs still perform similar functions, although there is no expression alignment on organ-inconsistent miRNAs.

Further, we focused on the human-mouse comparison to test whether miRNAs with different expression trajectories among organs are conserved in their target functions between these two species. To assess the potential effects of miRNAs on the expression of their target genes, we predicted human and mouse miRNA targets using miRNAtap with at least three algorithms and examined the published mRNA expression dataset derived from a partially overlapping set from our previous study (Cardoso-Moreira et al., 2019). The targets were further defined as anti-correlated with the miRNA expression (Spearman correlation, rho <-0.5). Functional enrichment test using Metascape on the group of targets with organ-consistent patterns displayed the conservation of genes and gene ontology terms between human and mouse (Figure 56). Specifically, the potential targets of organ-consistently increasing miRNAs were enriched in the functional term associated with cell

cycle and transcriptional activity in both human and mouse (Figure 57A). By contrast, synaptic signaling, protein localization to the membrane, and regulation of ion transmembrane transport are enriched in the targets of organ-consistently decreasing miRNAs (Figure 57B). However, there is less overlap with the targets of organ-inconsistent miRNAs between humans and mouse on both gene and GO ontology terms. It could be caused by the heterochrony of organogenesis or specific function during organ development of each species. It should be noted that we can still find some developmental-related functions enriched in organ-specific profiling genes, particularly in mouse (Figure 58). Specifically, targets of brain-specific changed miRNAs (CL4 in Figure 51 and Figure 52) are enriched in the neuron-related functions, such as dendrite development, synapse organization, and synaptic signaling, although there was slightly less significant functional enrichment for human targets (Figure 59).

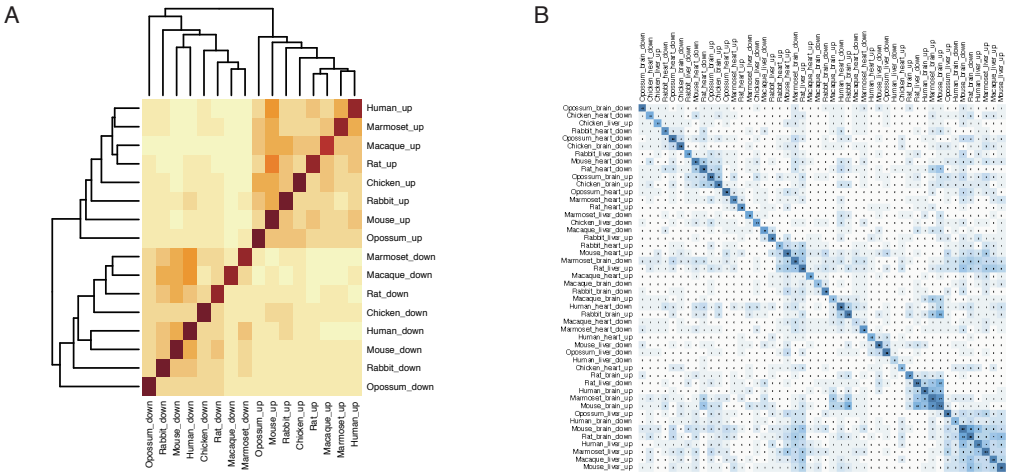

Figure 55. Heatmap visualizes the overlapped miRNA number between species in all pairwise comparisons

The heatmap cells are colored by the number of overlapped miRNA seeds. Darker color represents greater overlap. The hierarchical clustering was generated using the number of overlapped seeds based on Euclidean distance**. A**. Comparison between pairs of organ-consistent miRNA clusters for eight species. **B**. Comparison between pairs of organ-inconsistent miRNA clusters for eight species without displaying the dendrograms.

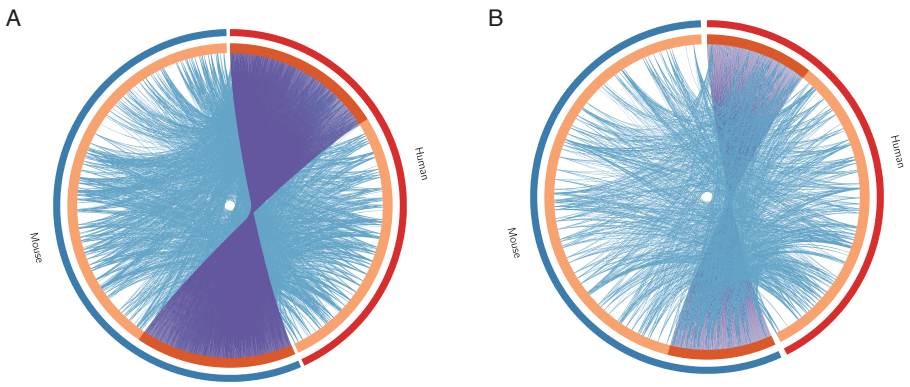

Figure 56. The circus plots showing miRNA target genes overlap between human and mouse by Metascape

On the outside, each arc represents the identity of each gene list. On the inside, each arc represents a gene list, where each gene has a spot on the arc. Dark orange represents the genes that appear in multiple lists, and light orange represents genes that are unique to that gene list. Purple lines link the same gene that are shared by multiple gene lists. Blue lines link the different genes where they fall into the same ontology term. The greater the number of purple links and the longer the dark orange arcs the greater is the overlap of miRNA target genes among human and mouse. Blue links indicate the amount of functional overlap among the miRNA target genes of human and mouse. **A**. miRNA target genes showing decreasing expression pattern among all three organs. **B**. miRNA target genes showing increasing expression pattern among all three organs.

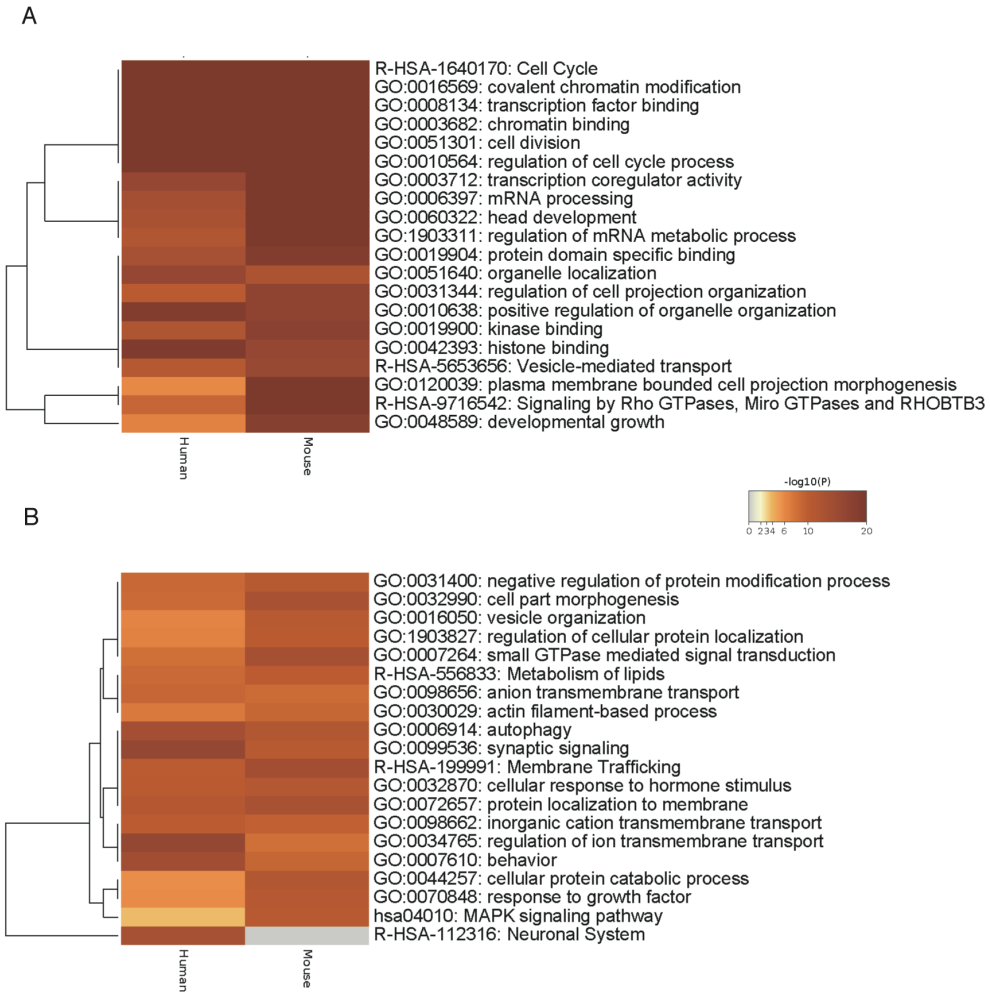

Figure 57. Gene Ontology terms enriched in miRNA target genes for organ-consistent miRNA developmental expression patterns

Selected terms with the best p-value within each cluster, as the representative terms, are displayed in a dendrogram. The heatmap cells are colored by their enrichment p-values, light cells indicate the lack of enrichment for that term in the corresponding species. The significant terms are hierarchically clustered into a tree based on Kappa-statistical similarities among their gene memberships. 0.3 kappa score was applied as the threshold to cast the tree into term clusters. **A**. Genes showing decreasing expression pattern among all three organs. **B**. Genes showing increasing expression pattern among all three organs.

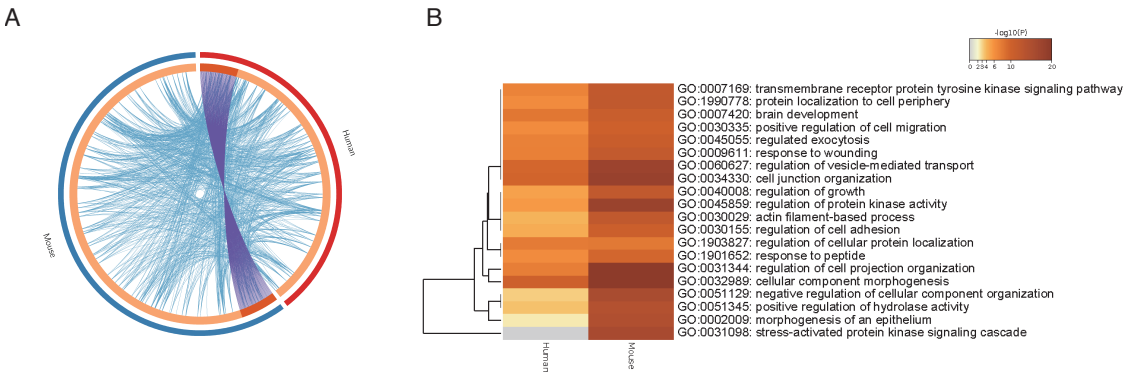

Figure 58. Gene Ontology terms enriched in miRNA target genes for all organ-inconsistent miRNA patterns

A. Circus plot shows how target genes overlap between human and mouse. On the outside, each arc represents the identity of each gene list. On the inside, each arc represents a gene list, where each gene has a spot on the arc. Dark orange represents the genes that appear in multiple, and light orange represents genes that are unique to that gene list. Purple lines link the same gene that are shared by multiple gene lists. Blue lines link the different genes where they fall into the same ontology term. **B**. Selected terms with the best p-value within each cluster, as the representative terms, are displayed in a dendrogram. The heatmap cells are colored by their enrichment p-values, light cells indicate the lack of enrichment for that term in the corresponding species. The significant terms are hierarchically clustered into a tree based on Kappa-statistical similarities among their gene memberships. 0.3 kappa score was applied as the threshold to cast the tree into term clusters.

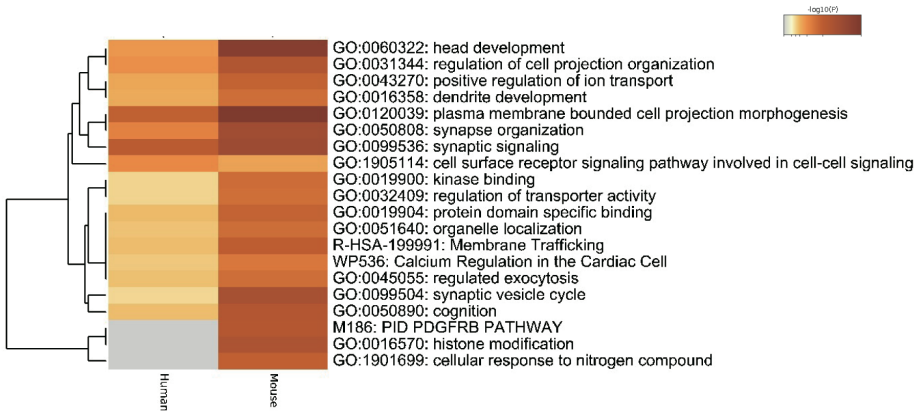

Figure 59. Gene Ontology terms enriched in target genes of brain-specific miRNAs.

Selected terms with the best p-value within each cluster, as the representative terms, are displayed in a dendrogram. The heatmap cells are colored by their enrichment p-values, light cells indicate the lack of enrichment for that term in the corresponding species. The significant terms are hierarchically clustered into a tree based on Kappa-statistical similarities among their gene memberships. 0.3 kappa score was applied as the threshold to cast the tree into term clusters.

4.2 Materials and Methods

4.2.1 Sample information and Datasets

We used three data sets to study the embryonic development of evolution on chordates. Briefly, we obtained the RNA-seq data and expression of *Crassostrea gigas* from the database GEO (accession number SRP014559) (G. Zhang et al., 2012), *Ciona intestinalis* from DDBJ (accession number DRA003460), and the rest of the species from DDBJ (accession number DRA000567).

To further understand the evolution of developmental processes of miRNA on different organ formation processes and regulatory relationships with potential targets, we obtained samples for brain, liver, and heart from eight species, human (Homo sapiens), macaque (Macaca mulatta), marmoset (Callithrix jacchus), mouse (Mus musculus), rat (Rattus norvegicus), rabbit (Oryctolagus cuniculus), opossum (Monodelphis domestica) and chicken (Gallus gallus) from the group of Prof. Dr. Henrik Kaessmann in the Center for Molecular Biology of Heidelberg University. Each sample covers 3-5 time points: early prenatal, late prenatal, neonatal, juvenile, and adult, but not very precisely. For opossums, the samples were all postnatal because they are born very early in development. The number of samples for each species was shown in Figure 33. The corresponding normalized gene expression data on the same dataset were downloaded from a previous study(Cardoso-Moreira et al., 2019).

4.2.2 Construction of indexed small RNA-Seq libraries

Due to difficulties in obtaining sufficient amounts of RNA, two RNA extraction methods were used in miRNA evo-devo study, as described in Warnefors et al. (2017) (Warnefors et al., 2017) as follows: Briefly, total RNA was extracted from 10 to 20 mg of tissue using the RNeasy Micro Kit and the miRNeasy Mini Kit according to the manufacturer's instructions. The small RNA libraries were sequenced on the Lausanne Genomic Technologies Facility on an Illumina HiSeq 2500 system. Adapters were removed from the raw sequencing reads, and only trimmed reads in size range of 15-28 nt

with a minimum quality score of 20 at all positions were retained. A technical replicate for sequencing was performed for each sample. All the above work was performed by the staff from Prof. Kaessmann's lab. Part of the data has been published on GEO (accession number: GSE102062). The miRNA quantification followed the procedure in section 3.2.3 "miRNA expression quantification".

4.2.3 RNA-seq data processing

To quantify gene expression of chordates, we mapped RNA-seq reads to the corresponding genome using Tophat (v2), allowing up to three mismatches and indels, except for ciona (Ci). In the case of ciona, we mapped reads with up to five mismatches and indels to the genome, as the ciona RNA-seq data and genome data are of slightly lower quality than the rest. Next, we filtered out very low expressed genes and identified the genes with maximum expression across all developmental stages exceeding 1 FPKM as expressed genes. For each stage, we calculated the expression as the mean of the expression of the replicated samples. More than 70% of coding genes annotated for a given species were reliably detected, except for amphioxus (41%).

4.2.4 Identification of dynamically changed mRNA/miRNA

We defined developmental changes in gene expression levels using polynomial regression models according to the method described in Somel et al.(Somel et al., 2009). For each gene, we selected the best regression model with 11-17 developmental stages (by rank) as the predictor and expression level as the response with Benjamini-Hochberg corrected $p < 0.05$. Genes that fit a significant regression model were designated as developmentally related genes.

We defined developmental changes in miRNA expression levels using linear regression models because there are only 3-5 developmental stages available for miRNA. We chose the linear model for each miRNA with 3-5 developmental stages (by rank) as predictor and expression level as responses. We randomly permuted the developmental stage 500 times and repeated the linear regression test for each miRNA. The miRNAs with

nominal p-value <0.01 (permutation p-value< 0.05) were designated as dynamically changed miRNAs.

4.2.5 Aligning of mouse developmental stage to developmental profiles of other species

To assess the transcriptome similarity in the course of developmental stages between mouse (Mm) and other species, we used stage-associated and 1:1 orthologous genes to align the developmental stages between mouse and other species following the method reported in (H. Hu et al., 2017; J. J. Li et al., 2014). The detailed description can be found in the supplementary materials of (H. Hu et al., 2017; J. J. Li et al., 2014). In brief, we calculated the pairing score using the hypergeometric test to evaluate the ratio of orthologs overlapping within the stage-associated genes pairwise. The pairing score can be written as:

$$\text{Paring score} = -\log_{10}(\text{Bonferroni corrected hypergeometric test P-values})$$

The paring score was used to quantify the significance of the overlap on each pairwise stage comparison between mouse and another species. From the paring score, we identified the relationship between mouse and other species (Figure 22). To check the stability of this relationship, we repeated the comparison 500 times with randomly assigned stage-associated genes for each stage of each species using the same procedure (Figure 21).

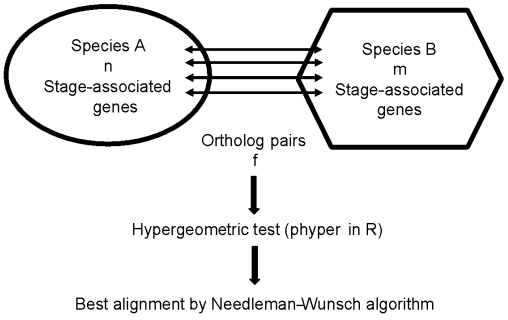

Figure 60. Workflow for stage alignment

According to the above scheme, we further assigned the corresponding stage alignment between mouse and other species using the Needleman-Wunsch algorithm with the gap penalty equal to one. To estimate the stability of alignment based on the mean of the replicates, we randomly chose one individual per stage, aligned two species 500 times, and calculated the frequency for each pair of alignments. The resulting frequency was presented by the thickness of the line in Figure 22.

4.2.6 Gene expression interpretation to mouse developmental stage

To compare the similarity of the expression profiles across developmental stages of nine species, we used the predicted developmental stage alignment presented in Figure 22 to create a unified alignment of eight species to the mouse developmental curves. To do so, we mapped 33 stages cumulatively interpolated from eight species to the full mouse developmental curve fitted using cubic smoothing spline with ten degrees of freedom. We then compared gene expression curves among nine species based on z-transformed expression of each gene interpolated at these 33 stage points.

4.2.7 Clustering of gene expression profiles in six vertebrate and nine chordate species

To investigate the expression pattern diversity in nine or six species, we used hierarchical clustering (hclust function in R) of z-transformed gene expression trajectories aligned among species with (1 - rho) as the distance measure, where rho is the Spearman correlation coefficient. We chose k equal six, as optimal, based on visual inspection of clusters obtained using different k values.

4.2.8 Clustering all co-expressed gene modules in nine distal species

To obtain complete transcriptomic information across all species, we combined 103,728 dynamically changed mouse-orthologous genes to identify the co-expressed modules by WGCNA with the parameter corType as “biocor” in 21 different modules. The gray module containing 176 genes, which cannot be clustered in any module, was removed for further analysis. Then, the GO and KEGG Pathway Enrichment Test were applied to

the expressed genes of each species in each module.

4.2.9 Functional annotation of developmental expression patterns

To verify the functions of genes in each module, we applied Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa & Goto, 2000) enrichment tests. GO annotation of mouse genes was downloaded from Bioconductor package "org.Mm.eg.db". Mouse pathway annotation was downloaded from http://rest.kegg.jp/list/pathway/ and http://rest.kegg.jp/link/genes/.

For the pathway enrichment test, the reference genes were the same as the GO-based analysis. We used a hypergeometric test (phyper in R) to assess enrichment in each KEGG pathway. Bonferroni correction was applied for genes in vertebrate clusters. Pathways with corrected $p<0.05$ were reported as significantly enriched. No correction was applied for genes from chordate clusters due to low statistical power, and the nominal p-value for pathway enrichment was set to $p <0.05$. This relaxed cutoff was used to assess the potential overlap between enriched functions found in vertebrate and chordate clusters.

For the GO enrichment test, the "elim" algorithm of topGO (Alexa & Rahnenfuhrer, 2016) was chosen to eliminate the hierarchical dependency of the GO terms. Fisher's exact test was applied to each GO term. The background set consisted of all dynamically changed genes in each group. For testing vertebrate gene clusters, the Benjamini-Hochberg correction was applied to "molecular function", "biological process", and "cellular component" independently. GO terms with BH-corrected $p<0.05$ were reported. We did not apply a multiple test correction for enrichment analysis on chordate gene clusters due to the low statistical power of the dataset and used a stricter cutoff of nominal $p <0.001$.

4.2.10 Dating of miRNA evolutionary age

To estimate miRNA evolutionary age for each species, we used the same procedure described in section 3.2.5 "Dating of human miRNA evolutionary ages". For each species,

we created an independent species tree to demonstrate the evolutionary age as described in Table S5.

4.2.11 Prediction of novel miRNAs

Novel miRNAs for each sample were detected using the miRDeep2 algorithm with default parameters based on the genome and miRbase v22 as references from Table 7. Predicted rRNA/tRNA reads and reads with the miRDeep2 score <=5 were removed in the following analyses. Sequencing reads mapped to overlapping miRNA genomic locations were merged across all samples. The detected miRNAs with no overlap with known mature miRNA genomic positions were considered novel.

Species	miRBase	UCSC Genome	Ensembl Gene annotation
Human	v22	hg38	Homo_sapiens.GRCh38.96.gtf.gz
Macaque	v22	rheMac8	Macaca_mulatta.Mmul_8.0.1.96.gtf.gz
Marmoset	v22	calJac3	Callithrix_jacchus.C_jacchus3.2.1.90.gtf.gz
Mouse	v22	mmu10_ucsc	Mus_musculus.GRCm38.96.gtf.gz
Rat	v22	rn6	Rattus_norvegicus.Rnor_6.0.96.gtf.gz
Rabbit	v22	oryCun2	Oryctolagus_cuniculus.OryCun2.0.96.gtf.gz
Opossum	v22	monDom5	Monodelphis_domestica.monDom5.96.gtf.gz
Chicken	v22	galgal5	Gallus_gallus.Gallus_gallus-5.0.86.gtf.gz

Table 7. Genome and miRNA annotation source for each species.

4.2.12 Sorting miRNA genomic location

Gene annotations were downloaded from Ensemble database matched to the genome and miRNA annotation (Table 7). The genome can be sorted as exon, intron, and intergenic region according to gene annotation. Only mature miRNA fully overlapped in one of the genomic regions was assigned.

4.2.13 miRNA target prediction and functional analysis of miRNA targets for Human and Mouse

We downloaded human and mouse miRNA predicted targets using miRNAtap (Simpson, 2019), where a predicted target had to be identified by at least three of the

following five methods: “pictar”, “diana”, “targetscan”, “miranda” and “mirdb”, as described in Section 3.2.7, “miRNA target prediction and functional analysis of miRNA targets”. The gene expression data of the same dataset were downloaded from a previous study (Cardoso-Moreira et al., 2019). The predicted targets negatively regulated with miRNA with Spearman correlation coefficient rho<-0.5 were considered potential targets for further functional enrichment tests. The predicted targets with Spearman correlation coefficient of rho<-0.1 were considered weakly repressed targets.

Gene ontology (GO) and pathway enrichment tests were processed by Metascape (Zhou et al., 2019) with all genes used as a background. The GO terms with $-\log_{10}(p) > 2$ were reported as significantly enriched.

4.2.14 Species tree construction

The species tree on nine chordates was generated using the web tool https://phylot.biobyte.de/ with NCBI taxonomy IDs (10090**,** 9031**,** 13735**,** 8364**,** 7955**,** 7719**,** 8355**,** 7739**,** 29159). The species separation time to the mouse was obtained from http://www.timetree.org/ (Kumar, Stecher, Suleski, & Hedges, 2017).

The species tree on eight vertebrates was generated under the web https://phylot.biobyte.de/ with NCBI taxonomy IDs (9606, 9544, 9483, 10090, 10116, 9986, 13616, 9031)

4.2.15 Statistical analysis and software

All statistical analyses and plots were performed in the R environment (http://www.r-project.org/), using packages preprocessCore, multcomp, limma, dendextend, ggpubr, ggsci, ggplot2, gridExtra, reshape2 and miRNAtap. Metascape was used for functional enrichment test on genes.

Chapter 5. Conclusions and Discussions

The molecular mechanisms underlying phenotypic diversity that emerged in human and vertebrate evolution has been investigated from different perspectives, such as genome complexity, transcript structure variation, gene expression differences, cis or trans-regulatory networks, epigenetic regulation, and lipid or metabolite composition divergence. However, biological technology and computational algorithm limitations delay our steps to find out the critical hidden genetic messages guiding developmental evolution, despite an accumulation of massive amounts of data at different omics levels. At present, we are still in the phase of collecting knowledge and getting more comprehensive data coverage. In my Ph.D. work, I quantified molecular phenotype by transcriptome expression, extending from protein-coding genes to microRNA to study human population divergence and vertebrate developmental evolution. Both studies added to our understanding of coding and non-coding RNAs to phenotypic divergence, development, and speciation.

5.1 Conclusions

The main goal of this thesis was to assess molecular alterations associated with human population divergence and vertebrate evolution of the embryonic developmental mechanisms using RNA-seq data for mRNA and miRNA expression as molecular phenotypes. This dissertation successfully demonstrates the main features of transcriptome diversity in these two types of evolutionary processes: evolution of vertebrate embryonic development and human population divergence. The key findings and contributions of this dissertation are summarized below.

1. In humans, when considering placental miRNA expression as an endophenotype, the population identity has a major effect on miRNA expression variation, and the miRNA expression divergence is partly consistent with the extent of the genetic divergence.

2. Another feature of placental miRNA expression in humans is that sex of a newborn is a demographic variable significantly associated with miRNA variation. Our

functional analysis showed that target genes regulated by these sex-of-the-newborn-associated miRNAs are related to hormonal processes. This observation aligns with previous studies reporting hormonal regulation as the primary driving mechanism of miRNA sex-biased expression.

3. In bulk, it appears that the effect of population-dependent miRNA expression differences in human placenta might mainly represent neutral evolutionary changes with minimal functional impacts, consistent with the overall concept of largely neutral effects of many or most gene expression differences detected among species.

4. In the embryonic developmental process, the genes with gradual descending expression profiles form the most conserved developmental expression profile both in the whole embryo analysis and in the study of isolated embryonic tissues.

5. Genes forming gradual descending pattern shared among chordates are involved in essential functions, such as RNA processing, matching observations reported in the study examining embryonic gene expression in isolated mouse tissues.

6. Genes forming gradual increasing expression at the late embryonic stages are substantially less conserved among vertebrate and chordate species and show no dominant involvement in particular biological functions.

7. In humans and mice, miRNA regulatory effects are more conserved for miRNAs displaying developmental expression trajectories shared among organs compared to miRNAs displaying organ-dependent developmental profiles.

8. Organ-specific miRNAs, which trend to be evolutionary young, do not show clear regulatory roles focused on specific functions in organ development, consistent with proposed gradual evolution of regulatory engagement of newly evolved miRNAs.

5.2 Discussion

In studying microRNA expression variation in placenta and its functional implications among the human populations, I based my work on a rare collection of well-characterized samples derived from a single geographic location comprising four human populations characterized by twelve demographic variables. My colleagues and I found that population affiliation and sex of a child had substantial influence on miRNA expression variation among human placental samples. Population has the most substantial influence explaining up to 11% of the total miRNA variance. The overall relative miRNA expression divergence among the four populations studied in the study is similar to their genetic divergence, implying that the presence of single nucleotide polymorphism (SNP) in different miRNA gene regions could be one cause of population difference, as we know that SNP in different miRNA gene regions can affect miRNA function or the efficiency of miRNA biogenesis (Duan, Pak, & Jin, 2007; Sun et al., 2009). We also know that CNV influences gene expression and demonstrates differences between human populations. Using the previous study's list of copy number variable miRNAs, we discovered four miRNAs in CNV regions among our 93 population-associated miRNAs (Marcinkowska, Szymanski, Krzyzosiak, & Kozlowski, 2011). Except for the genetic influence, environments, socioeconomic status, cultural upbringing, and other factors may correlate with population grouping, which we can investigate more precisely in the latter study. Concerning investigating the potential regulatory functions of miRNA on gene expression divergence among populations, we used mRNA expression data collected from the same samples in this study. However, we did not observe substantial regulatory effects of population-associated miRNAs on mRNA expression in our sample set: our analysis revealed a significant excess of expressional repression among predicted targets for one of the six miRNA clusters only in my study. This result appears to contrast the reported widespread population-specific downregulation of miRNA targets described in cell lines (Huang et al., 2011). While part of this discrepancy might be due to the limited statistical power of our study involving a total of approximately 40 samples, the rest could be caused by the unequal extent of the evolutionary constraint in tissues and cell lines. As other

regulators controlling multiple targets, miRNAs are under substantial evolutionary constraint from the expression level standpoint (Wu, Shen, & Tang, 2009; Xu et al., 2013). Assuming that most randomly arising population-specific miRNA expression differences are non-adaptive, the ones with large regulatory effects are likely to be detrimental and will not be observed in natural tissue, such as placenta. However, the artificial growth conditions of the cell lines could allow the manifestation of large-scale population-associated regulatory effects of miRNA variation. Furthermore, 32 female and male newborn-associated miRNAs were shown to have regulatory interactions with several biological processes. However, the targets of these two categories of miRNAs both have hormone-related functions, as evidenced by the findings of a recent study indicating sex hormones can influence miRNA expression. It's worth mentioning, however, that placental tissue can secrete steroid hormones, estrogen, and human placental lactogen, which could skew some of our findings. We could investigate the regulatory functions adjusting hormone-related miRNA targets in the future.

In the other part of my thesis work, I focused on comparative transcriptome analysis in embryonic development and evolution. Specifically, my colleagues and I explored the general relationship between ontogeny and phylogeny by comparing protein-coding gene and miRNA developmental expression trajectories among chordates and vertebrates, respectively. Unlike the traditional studies focusing on investigating the conserved phylotypic period among phylum, we compared dynamically changing developmental expression profiles and investigated biological functions of genes with shared patterns of developmental profile alterations. Our results show that the dominant developmental pattern conserved among species and containing the largest proportion of gene orthologs present in each of the assessed species represents expression trajectories that gradually descend along embryogenesis. This pattern dominated developmental trajectories both in chordates and vertebrates in the whole embryo analysis and was reproduced in the study of isolated mammalian and bird embryonic tissues (T. Gong et al., 2020). Agreement between whole-embryo-based and tissue-based observations indicates that the descending developmental expression reflects alterations within embryo tissues rather than changes in

embryo anatomical composition. By contrast, a group of genes with gradually increasing developmental expression was substantially less conserved among species and more enriched in tissue specific genes. One possible explanation is that genes with a conserved descending pattern have a higher proportion of older genes. It has been postulated that older gene promoters have DNase I hypersensitive sites (DHSs) at earlier developmental stages, allowing gene expression to be active at earlier stages (Gao et al., 2018). Another reason could be that those genes are subject to gene pleiotropy constraints. We found that 88% of mouse genes with descending expression profiles expressed at many stages, but only 48% of mouse genes with ascending expression profiles did. Other vertebrates have made similar observations.

To obtain a clearer view of developmental expression evolution, I further analyzed protein-coding and non-coding miRNA expression changes on three organs in vertebrate development. The results showed that approximately one-third of protein-coding genes and miRNA transcripts showed divergent developmental expressed profiles among organs, indicating miRNA's potential regulatory role in the evolution of organ-specific expression during embryogenesis. Regulatory effects of these organ-inconsistent miRNAs were less conserved between human and mouse compared to organ-consistent miRNAs. However, functional analysis revealed relevant biological terms, such as synapse development-related functions, as significantly enriched in the targets of miRNAs with brain-specific decreasing developmental profile, both in human and mouse. Other organ-specific miRNA targets are potentially involved in proliferation, growth, and development only in mouse, not human. What's more, we identified many newly emerged organ-specific miRNAs, whose expansion might be one of the driving forces behind the evolution of organ complexity, as suggested previously. While our current functional analysis of organ-specific miRNAs does not show related functions on tissue development, these miRNAs expressed at low levels and in specific organs are likely to target many genes by chance, with possible deleterious consequences (Berezikov, 2011). Even though I did not find significant enrichment of organ-specific miRNAs' targets in particular biological processes, I identified a total of 30 miRNA families that were explicitly expressed in one organ of at

least three species. These miRNA families partially overlapped with reported ones, showing consistent tissue specificity. Particularly, miR-129-1-3p, miR-124-5p, miR-137, miR-153, and miR-128 are specifically expressed in brain (Z. Guo et al., 2014); miR-208b-3p are heart-specific (Ludwig et al., 2016); miR-1 and miR-133 are muscle-specific (Z. Guo et al., 2014). My results further revealed that conserved and evolutionarily novel miRNAs differ in their contribution to embryonic development. The miRNA evolutionary conservation throughout all stages of embryogenesis was involved in RNA processing functions. By contrast, the brain-specific dynamically evolving miRNA-mRNA regulatory network contributed to synapse development in human and mouse.

One of the critical problems present in both parts of my thesis work is the determination of the mRNA-miRNA regulatory relationship. Both in population-level and species-level study, I often did not observe evident negative correlations between miRNA expression profiles and expression profiles of their targets, compared to the non-targets control. In both our datasets involved in miRNA-mRNA interaction analysis, we have few time points and relatively small sample sizes, potentially yielding insufficient statistical power to identify significantly anti-correlated miRNA targets. In addition, several technical factors might have further restricted our ability to detect miRNA-driven regulation of their predicted mRNA targets. Such factors include a mismatch between computational and experimentally verified miRNA target predictions, the known phenomenon that miRNA targeting leads to sequestering target mRNA out of the translational pool without degradation, and the complex and often a tissue-specific interplay between miRNAs and other regulators (Chou et al., 2016; Z. Guo et al., 2014). To overcome these issues, we could either collect more comprehensive sample sets or use additional methods to improve miRNA target predictions, such as AGO CLIP sequencing and CLASH, which could identify AGO binding and miRNA target sites directly. What's more, we could add transcriptional factor (TF) and TF-binding regulation to our current analysis, integrate this level of regulation with miRNA-driven one, and potentially reconstruct TF-miRNA-gene regulatory networks. Based on this reconstruction, we could then obtain further knowledge

of how gene products and regulatory networks evolve and how these evolutionary alterations might result in phenotypic change and speciation.

Additionally, our study did not consider the tissue and cell-type composition differences among species during embryogenesis. New cell types arise along with progenitors fading away during cell differentiation. However, bulk tissues analysis cannot provide information about cellular heterogeneity. Therefore, our comparison of bulk tissues among species could be affected by such cell type composition differences. Further research using high-resolution single-cell transcriptomes could elucidate a more precise relationship between cell types across species. In fact, single-cell cross-species analyses of human fetal and mouse embryos have already revealed the conserved and divergent transcriptome profiles during embryonic development (Cao et al., 2020; J. Cao et al., 2019). Large-scale single-cell transcriptome studies have provided useful resources for us to understand cell fate, organogenesis, tissue regeneration in different species, such as mouse, zebrafish, ciona, and *C.elegans* (C. Cao et al., 2019; Packer et al., 2019; Pijuan-Sala et al., 2019; D. E. Wagner et al., 2018). Such multi-species developmental resources provide an important foundation for future studies of developmental evolution in divergent species.

Furthermore, in my study, I oberseved that there are less overlap of tissue-speicfic miRNAs and its regulatory function among vetebrates. Similar observation were found in a study of teleost fish investigating gene expression variation among tissues, individuals, and populations. The results showed that half of the genes were differentially expressed among individuals within a population-tissue group and that only a small subset (31%) of tissue-specific differences were consistent across all three populations. It suggested that many tissue-specific differences in gene expression are unique to a single population and are therefore unlikely to contribute to fundamental distinctions between tissue types (Whitehead & Crawford, 2005). However, both of studies did not look at the effect of genetic diversity on gene and miRNA expression levels across tissues. An analysis of copy number variation (CNV) in multiple somatic tissues from six unrelated people revealed a significant amount of intraindividual genomic variation between tissues, demonstrating

that somatic tissues can be genetically diverse. 79% of these events have an impact on genes (O'Huallachain et al., 2012). The Genotype-Tissue Expression (GTEx) project that includes genotype, gene expression, histological and clinical data for 449 human individuals across 42 distinct tissues discovered that cis-acting genetic variants tend to affect either most tissues or a small number of tissues (G. T. Consortium et al., 2017). Accoordingly, to understand the causal of phenotype difference among tissues, individuals, or populations, we would integrate genetic information in the future studies.

In conclusion, my work provides a view of molecular phenotypes (mRNA and miRNA) variation and interplay in human and vertebrate evolution. I believe, first of all, my work highlights the value of quantitative endophenotypes as an essential supplement to morphology-based developmental studies. However, both parts of my thesis study could only be considered as a starting point for future research: we still lack complete molecular developmental alterations at other levels, and our data represents a mixed signature of cellular, organ, and organismal evolution. With the extension of other omics-data, such as genome, epigenome, single-cell resolution transcriptome, spatial transcriptome, proteome, and metabolome, we could understand the differential contributions of the various molecular mechanisms to the rise of phenotypic diversity. It must be noted that despite these limitations, the fact that our study reveals many novel miRNA and mRNA expression features indicates the importance of further studies investigating the role of gene-expression-level alterations in human and vertebrate evolution.

Bibliography

Abbott, A. L., Alvarez-Saavedra, E., Miska, E. A., Lau, N. C., Bartel, D. P., Horvitz, H. R., & Ambros, V. (2005). The let-7 MicroRNA family members mir-48, mir-84, and mir-241 function together to regulate developmental timing in Caenorhabditis elegans. *Dev Cell, 9*(3), 403-414. doi:10.1016/j.devcel.2005.07.009

Aboobaker, A. A., Tomancak, P., Patel, N., Rubin, G. M., & Lai, E. C. (2005). Drosophila microRNAs exhibit diverse spatial expression patterns during embryonic development. *Proc Natl Acad Sci U S A, 102*(50), 18017-18022. doi:10.1073/pnas.0508823102

Abrahante, J. E., Daul, A. L., Li, M., Volk, M. L., Tennessen, J. M., Miller, E. A., & Rougvie, A. E. (2003). The Caenorhabditis elegans hunchback-like gene lin-57/hbl-1 controls developmental time and is regulated by microRNAs. *Dev Cell, 4*(5), 625-637. doi:10.1016/s1534-5807(03)00127-8

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., . . . Tuschl, T. (2006). A novel class of small RNAs bind to MILI protein in mouse testes. *Nature, 442*(7099), 203-207. doi:10.1038/nature04916

Armengol, L., Villatoro, S., Gonzalez, J. R., Pantano, L., Garcia-Aragones, M., Rabionet, R., . . . Estivill, X. (2009). Identification of copy number variants defining genomic differences among major human groups. *PLoS One, 4*(9), e7230. doi:10.1371/journal.pone.0007230

Artieri, C. G., Haerty, W., & Singh, R. S. (2009). Ontogeny and phylogeny: molecular signatures of selection, constraint, and temporal pleiotropy in the development of Drosophila. *BMC Biol, 7*, 42. doi:10.1186/1741-7007-7-42

Atallah, J., & Lott, S. E. (2018). Evolution of maternal and zygotic mRNA complements in the early Drosophila embryo. *PLoS Genet, 14*(12), e1007838. doi:10.1371/journal.pgen.1007838

Atianand, M. K., Caffrey, D. R., & Fitzgerald, K. A. (2017). Immunobiology of Long Noncoding RNAs. *Annu Rev Immunol, 35*, 177-198. doi:10.1146/annurev-immunol-041015-055459

Baek, D., Villen, J., Shin, C., Camargo, F. D., Gygi, S. P., & Bartel, D. P. (2008). The impact of microRNAs on protein output. *Nature, 455*(7209), 64-71. doi:10.1038/nature07242

Bai, J., Abdul-Rahman, M. F., Rifkin-Graboi, A., Chong, Y. S., Kwek, K., Saw, S. M., . . . Qiu, A. (2012). Population differences in brain morphology and microstructure among Chinese, Malay, and Indian neonates. *PLoS One, 7*(10), e47816. doi:10.1371/journal.pone.0047816

Barker, D. J., Forsen, T., Uutela, A., Osmond, C., & Eriksson, J. G. (2001). Size at birth and resilience to effects of poor living conditions in adult life: longitudinal study. *BMJ, 323*(7324), 1273-1276. doi:10.1136/bmj.323.7324.1273

Bateson, P., Barker, D., Clutton-Brock, T., Deb, D., D'Udine, B., Foley, R. A., . . . Sultan, S. E. (2004). Developmental plasticity and human health. *Nature, 430*(6998), 419-421. doi:10.1038/nature02725

Bennis-Taleb, N., Remacle, C., Hoet, J. J., & Reusens, B. (1999). A low-protein isocaloric diet during gestation affects brain development and alters permanently cerebral cortex blood vessels in rat offspring. *J Nutr, 129*(8), 1613-1619. doi:10.1093/jn/129.8.1613

Berezikov, E. (2011). Evolution of microRNA diversity and regulation in animals. *Nat Rev Genet, 12*(12), 846-860. doi:10.1038/nrg3079

Bininda-Emonds, O. R., Jeffery, J. E., & Richardson, M. K. (2003). Inverting the hourglass: quantitative evidence against the phylotypic stage in vertebrate development. *Proc Biol Sci, 270*(1513), 341-346. doi:10.1098/rspb.2002.2242

Bozinovic, G., Sit, T. L., Hinton, D. E., & Oleksiak, M. F. (2011). Gene expression throughout a vertebrate's embryogenesis. *BMC Genomics, 12*, 132. doi:10.1186/1471-2164-12-132

Brennecke, J., Hipfner, D. R., Stark, A., Russell, R. B., & Cohen, S. M. (2003). bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila. *Cell, 113*(1), 25-36. doi:10.1016/s0092-8674(03)00231-9

Britten, R. J., & Davidson, E. H. (1969). Gene regulation for higher cells: a theory. *Science, 165*(3891), 349-357. doi:10.1126/science.165.3891.349

Britten, R. J., & Davidson, E. H. (1971). Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. *Q Rev Biol, 46*(2), 111-138. doi:10.1086/406830

Bushati, N., Stark, A., Brennecke, J., & Cohen, S. M. (2008). Temporal reciprocity of miRNAs and their targets during the maternal-to-zygotic transition in Drosophila. *Curr Biol, 18*(7), 501-506. doi:10.1016/j.cub.2008.02.081

Byrne, M., Ho, M., Selvakumaraswamy, P., Nguyen, H. D., Dworjanyn, S. A., & Davis, A. R. (2009). Temperature, but not pH, compromises sea urchin fertilization and early development under near-future climate change scenarios. *Proc Biol Sci, 276*(1663), 1883-1888. doi:10.1098/rspb.2008.1935

Cao, C., Lemaire, L. A., Wang, W., Yoon, P. H., Choi, Y. A., Parsons, L. R., . . . Chen, K. (2019). Comprehensive single-cell transcriptome lineages of a proto-vertebrate. *Nature, 571*(7765), 349-354. doi:10.1038/s41586-019-1385-y

Cao, J., O'Day, D. R., Pliner, H. A., Kingsley, P. D., Deng, M., Daza, R. M., . . . Shendure, J. (2020). A human cell atlas of fetal gene expression. *Science, 370*(6518). doi:10.1126/science.aba7721

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D. M., Hill, A. J., . . . Shendure, J. (2019). The single-cell transcriptional landscape of mammalian organogenesis. *Nature, 566*(7745), 496-502. doi:10.1038/s41586-019-0969-x

Cardoso-Moreira, M., Halbert, J., Valloton, D., Velten, B., Chen, C., Shao, Y., . . . Kaessmann, H. (2019). Gene expression across mammalian organ development. *Nature, 571*(7766), 505-509. doi:10.1038/s41586-019-1338-5

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M. C., Maeda, N., . . . Genome Science, Group. (2005). The transcriptional landscape of the mammalian genome. *Science, 309*(5740), 1559-1563. doi:10.1126/science.1112014

Carroll, S. B. (2008). Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell, 134*(1), 25-36. doi:10.1016/j.cell.2008.06.030

Chang, X., Li, S., Li, J., Yin, L., Zhou, T., Zhang, C., . . . Sun, K. (2014). Ethnic differences in microRNA-375 expression level and DNA methylation status in type 2 diabetes of Han and Kazak populations. *J Diabetes Res, 2014*, 761938. doi:10.1155/2014/761938

Cheng, X., Hui, J. H., Lee, Y. Y., Wan Law, P. T., & Kwan, H. S. (2015). A "developmental hourglass" in fungi. *Mol Biol Evol, 32*(6), 1556-1566. doi:10.1093/molbev/msv047

Chou, C. H., Chang, N. W., Shrestha, S., Hsu, S. D., Lin, Y. L., Lee, W. H., . . . Huang, H. D. (2016). miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database. *Nucleic Acids Res, 44*(D1), D239-247. doi:10.1093/nar/gkv1258

Christodoulou, F., Raible, F., Tomer, R., Simakov, O., Trachana, K., Klaus, S., . . . Arendt, D. (2010). Ancient animal microRNAs and the evolution of tissue identity. *Nature, 463*(7284), 1084-1088. doi:10.1038/nature08744

Cizeron, M., Qiu, Z., Koniaris, B., Gokhale, R., Komiyama, N. H., Fransen, E., & Grant, S. G. N. (2020). A brainwide atlas of synapses across the mouse life span. *Science, 369*(6501), 270-275. doi:10.1126/science.aba3163

Comte, A., Roux, J., & Robinson-Rechavi, M. (2010). Molecular signaling in zebrafish development and the vertebrate phylotypic period. *Evol Dev, 12*(2), 144-156. doi:10.1111/j.1525-142X.2010.00400.x

Consortium, Encode Project, Birney, E., Stamatoyannopoulos, J. A., Dutta, A., Guigo, R., Gingeras, T. R., . . . de Jong, P. J. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature, 447*(7146), 799-816. doi:10.1038/nature05874

Consortium, G. TEx, Laboratory, Data Analysis, Coordinating Center -Analysis Working, Group, Statistical Methods groups-Analysis Working, Group, Enhancing, GTEx groups, Fund, N. I. H. Common, . . . Montgomery, S. B. (2017). Genetic effects on gene expression across human tissues. *Nature, 550*(7675), 204-213. doi:10.1038/nature24277

Cordero, G. A., Sanchez-Villagra, M. R., & Werneburg, I. (2020). An irregular hourglass pattern describes the tempo of phenotypic development in placental mammal evolution. *Biol Lett, 16*(5), 20200087. doi:10.1098/rsbl.2020.0087

Cridge, A. G., Dearden, P. K., & Brownfield, L. R. (2016). Convergent occurrence of the developmental hourglass in plant and animal embryogenesis? *Ann Bot, 117*(5), 833-843. doi:10.1093/aob/mcw024

Cui, C., Yang, W., Shi, J., Zhou, Y., Yang, J., Cui, Q., & Zhou, Y. (2018). Identification and Analysis of Human Sex-biased MicroRNAs. *Genomics Proteomics Bioinformatics, 16*(3), 200-211. doi:10.1016/j.gpb.2018.03.004

Daca-Roszak, P., Swierniak, M., Jaksik, R., Tyszkiewicz, T., Oczko-Wojciechowska, M., Zebracka-Gala, J., . . . Zietkiewicz, E. (2018). Transcriptomic population markers for human population discrimination. *BMC Genet, 19*(1), 54. doi:10.1186/s12863-018-0663-2

Dai, R., & Ahmed, S. A. (2014). Sexual dimorphism of miRNA expression: a new perspective in understanding the sex bias of autoimmune diseases. *Ther Clin Risk Manag, 10*, 151-163. doi:10.2147/TCRM.S33517

de Magalhaes, J. P., & Matsuda, A. (2012). Genome-wide patterns of genetic distances reveal candidate Loci contributing to human population-specific traits. *Ann Hum Genet, 76*(2), 142-158. doi:10.1111/j.1469-1809.2011.00695.x

Domazet-Loso, T., & Tautz, D. (2010). A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns. *Nature, 468*(7325), 815-818. doi:10.1038/nature09632

Duan, R., Pak, C., & Jin, P. (2007). Single nucleotide polymorphism associated with mature miR-125a alters the processing of pri-miRNA. *Human Molecular Genetics, 16*(9), 1124-1131. doi:10.1093/hmg/ddm062

Duboule, D. (1994). Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. *Dev Suppl*, 135-142.

Dunn, C. W., Zapata, F., Munro, C., Siebert, S., & Hejnol, A. (2018). Pairwise comparisons across species are problematic when analyzing functional genomic data. *Proc Natl Acad Sci U S A, 115*(3), E409-E417. doi:10.1073/pnas.1707515115

Evans, T. G., Chan, F., Menge, B. A., & Hofmann, G. E. (2013). Transcriptomic responses to ocean acidification in larval sea urchins from a naturally variable pH environment. *Mol Ecol, 22*(6), 1609-1625. doi:10.1111/mec.12188

Ferner, K., Schultz, J. A., & Zeller, U. (2017). Comparative anatomy of neonates of the three major mammalian groups (monotremes, marsupials, placentals) and implications for the ancestral mammalian neonate morphotype. *J Anat, 231*(6), 798-822. doi:10.1111/joa.12689

Friedlander, M. R., Mackowiak, S. D., Li, N., Chen, W., & Rajewsky, N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res, 40*(1), 37-52. doi:10.1093/nar/gkr688

Gabaldon, T., & Koonin, E. V. (2013). Functional and evolutionary implications of gene orthology. *Nat Rev Genet, 14*(5), 360-366. doi:10.1038/nrg3456

Gao, L., Wu, K., Liu, Z., Yao, X., Yuan, S., Tao, W., . . . Liu, J. (2018). Chromatin Accessibility Landscape in Human Early Embryos and Its Association with Evolution. *Cell, 173*(1), 248-259 e215. doi:10.1016/j.cell.2018.02.028

Garber, M., Grabherr, M. G., Guttman, M., & Trapnell, C. (2011). Computational methods for transcriptome annotation and quantification using RNA-seq. *Nat Methods, 8*(6), 469-477. doi:10.1038/nmeth.1613

Genomes Project, Consortium, Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., . . . Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature, 526*(7571), 68-74. doi:10.1038/nature15393

Gilad, Y., Oshlack, A., & Rifkin, S. A. (2006). Natural selection on gene expression. *Trends Genet, 22*(8), 456-461. doi:10.1016/j.tig.2006.06.002

Gildor, T., & Ben-Tabou de-Leon, S. (2015). Comparative Study of Regulatory Circuits in Two Sea Urchin Species Reveals Tight Control of Timing and High Conservation

of Expression Dynamics. *PLoS Genet, 11*(7), e1005435. doi:10.1371/journal.pgen.1005435

Gildor, T., & Smadar, B. D. (2018). Comparative Studies of Gene Expression Kinetics: Methodologies and Insights on Development and Evolution. *Front Genet, 9*, 339. doi:10.3389/fgene.2018.00339

Gluckman, P. D., Hanson, M. A., Spencer, H. G., & Bateson, P. (2005). Environmental influences during development and their later consequences for health and disease: implications for the interpretation of empirical studies. *Proc Biol Sci, 272*(1564), 671-677. doi:10.1098/rspb.2004.3001

Gluckman, P. D., Lillycrop, K. A., Vickers, M. H., Pleasants, A. B., Phillips, E. S., Beedle, A. S., . . . Hanson, M. A. (2007). Metabolic plasticity during mammalian development is directionally dependent on early nutritional status. *Proc Natl Acad Sci U S A, 104*(31), 12796-12800. doi:10.1073/pnas.0705667104

Gong, T., Zhang, C., Ni, X., Li, X., Li, J., Liu, M., . . . Wang, Y. (2020). A time-resolved multi-omic atlas of the developing mouse liver. *Genome Res, 30*(2), 263-275. doi:10.1101/gr.253328.119

Gong, W., Huang, Y., Xie, J., Wang, G., Yu, D., Zhang, K., . . . Zhu, Y. (2018). Identification and comparative analysis of the miRNA expression profiles from four tissues of Micropterus salmoides using deep sequencing. *Genomics, 110*(6), 414-422. doi:10.1016/j.ygeno.2018.09.017

Grivna, S. T., Beyret, E., Wang, Z., & Lin, H. (2006). A novel class of small RNAs in mouse spermatogenic cells. *Genes Dev, 20*(13), 1709-1714. doi:10.1101/gad.1434406

Guo, J., Wu, Y., Zhu, Z., Zheng, Z., Trzaskowski, M., Zeng, J., . . . Yang, J. (2018). Global genetic differentiation of complex traits shaped by natural selection in humans. *Nat Commun, 9*(1), 1865. doi:10.1038/s41467-018-04191-y

Guo, Z., Maki, M., Ding, R., Yang, Y., Zhang, B., & Xiong, L. (2014). Genome-wide survey of tissue-specific microRNA and transcription factor regulatory networks in 12 tissues. *Sci Rep, 4*, 5150. doi:10.1038/srep05150

Haeckel, Ernst. (1866). Generelle Morphologie der Organismen. *Berlin: Georg Reimer*.

Hao, P., & Waxman, D. J. (2018). Functional Roles of Sex-Biased, Growth Hormone-Regulated MicroRNAs miR-1948 and miR-802 in Young Adult Mouse Liver. *Endocrinology, 159*(3), 1377-1392. doi:10.1210/en.2017-03109

Hausser, J., & Zavolan, M. (2014). Identification and consequences of miRNA-target interactions--beyond repression of gene expression. *Nat Rev Genet, 15*(9), 599-612. doi:10.1038/nrg3765

Heimberg, A. M., Sempere, L. F., Moy, V. N., Donoghue, P. C., & Peterson, K. J. (2008). MicroRNAs and the advent of vertebrate morphological complexity. *Proc Natl Acad Sci U S A, 105*(8), 2946-2950. doi:10.1073/pnas.0712259105

Hemming, J. P., Gruber-Baldini, A. L., Anderson, K. E., Fishman, P. S., Reich, S. G., Weiner, W. J., & Shulman, L. M. (2011). Racial and socioeconomic disparities in parkinsonism. *Arch Neurol, 68*(4), 498-503. doi:10.1001/archneurol.2010.326

Hu, H., Uesaka, M., Guo, S., Shimai, K., Lu, T. M., Li, F., . . . Consortium, Expande. (2017). Constrained vertebrate evolution by pleiotropic genes. *Nat Ecol Evol, 1*(11), 1722-1730. doi:10.1038/s41559-017-0318-0

Hu, H. Y., Guo, S., Xi, J., Yan, Z., Fu, N., Zhang, X., . . . Khaitovich, P. (2011). MicroRNA expression and regulation in human, chimpanzee, and macaque brains. *PLoS Genet, 7*(10), e1002327. doi:10.1371/journal.pgen.1002327

Hu, H. Y., Yan, Z., Xu, Y., Hu, H., Menzel, C., Zhou, Y. H., . . . Khaitovich, P. (2009). Sequence features associated with microRNA strand selection in humans and flies. *BMC Genomics, 10*, 413. doi:10.1186/1471-2164-10-413

Huang, R. S., Gamazon, E. R., Ziliak, D., Wen, Y., Im, H. K., Zhang, W., . . . Dolan, M. E. (2011). Population differences in microRNA expression and biological implications. *RNA Biol, 8*(4), 692-701. doi:10.4161/rna.8.4.16029

Hughes, D. A., Kircher, M., He, Z., Guo, S., Fairbrother, G. L., Moreno, C. S., . . . Stoneking, M. (2015). Evaluating intra- and inter-individual variation in the human placental transcriptome. *Genome Biol, 16*, 54. doi:10.1186/s13059-015-0627-z

Husquin, L. T., Rotival, M., Fagny, M., Quach, H., Zidane, N., McEwen, L. M., . . . Quintana-Murci, L. (2018). Exploring the genetic basis of human population differences in DNA methylation and their causal impact on immune gene regulation. *Genome Biol, 19*(1), 222. doi:10.1186/s13059-018-1601-3

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

Irie, N., & Sehara-Fujisawa, A. (2007). The vertebrate phylotypic stage and an early bilaterian-related stage in mouse embryogenesis defined by genomic information. *BMC Biol, 5*, 1. doi:10.1186/1741-7007-5-1

Jeffery, J. E., Bininda-Emonds, O. R., Coates, M. I., & Richardson, M. K. (2002). Analyzing evolutionary patterns in amniote embryonic development. *Evol Dev, 4*(4), 292-302. doi:10.1046/j.1525-142x.2002.02018.x

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

Kaneda, M., Tang, F., O'Carroll, D., Lao, K., & Surani, M. A. (2009). Essential role for Argonaute2 protein in mouse oogenesis. *Epigenetics Chromatin, 2*(1), 9. doi:10.1186/1756-8935-2-9

Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res, 28*(1), 27-30. doi:10.1093/nar/28.1.27

Kloosterman, W. P., & Plasterk, R. H. (2006). The diverse functions of microRNAs in animal development and disease. *Dev Cell, 11*(4), 441-450. doi:10.1016/j.devcel.2006.09.009

Konstantinides, N., Degabriel, S., & Desplan, C. (2018). Neuro-evo-devo in the single cell sequencing era. *Curr Opin Syst Biol, 11*, 32-40. doi:10.1016/j.coisb.2018.08.001

Krichevsky, A. M., King, K. S., Donahue, C. P., Khrapko, K., & Kosik, K. S. (2003). A microRNA array reveals extensive regulation of microRNAs during brain development. *RNA, 9*(10), 1274-1281. doi:10.1261/rna.5980303

Kumar, S., Stecher, G., Suleski, M., & Hedges, S. B. (2017). TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Molecular Biology and Evolution, 34*(7), 1812-1819. doi:10.1093/molbev/msx116

Kuntz, S. G., & Eisen, M. B. (2014). Drosophila embryogenesis scales uniformly across temperature in developmentally diverse species. *PLoS Genet, 10*(4), e1004293. doi:10.1371/journal.pgen.1004293

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol, 10*(3), R25. doi:10.1186/gb-2009-10-3-r25

Lappalainen, T., Sammeth, M., Friedlander, M. R., t Hoen, P. A., Monlong, J., Rivas, M. A., . . . Dermitzakis, E. T. (2013). Transcriptome and genome sequencing uncovers functional variation in humans. *Nature, 501*(7468), 506-511. doi:10.1038/nature12531

Leite, D. J., Ninova, M., Hilbrant, M., Arif, S., Griffiths-Jones, S., Ronshaugen, M., & McGregor, A. P. (2016). Pervasive microRNA Duplication in Chelicerates: Insights from the Embryonic microRNA Repertoire of the Spider Parasteatoda tepidariorum. *Genome Biol Evol, 8*(7), 2133-2144. doi:10.1093/gbe/evw143

Leong, J. W., Abdullah, S., Ling, K. H., & Cheah, P. S. (2016). Spatiotemporal Expression and Molecular Characterization of miR-344b and miR-344c in the Developing Mouse Brain. *Neural Plast, 2016*, 1951250. doi:10.1155/2016/1951250

Levchenko, A., Kanapin, A., Samsonova, A., & Gainetdinov, R. R. (2018). Human Accelerated Regions and Other Human-Specific Sequence Variations in the Context of Evolution and Their Relevance for Brain Development. *Genome Biol Evol, 10*(1), 166-188. doi:10.1093/gbe/evx240

Levin, M., Anavy, L., Cole, A. G., Winter, E., Mostov, N., Khair, S., . . . Yanai, I. (2016). The mid-developmental transition and the evolution of animal body plans. *Nature, 531*(7596), 637-641. doi:10.1038/nature16994

Levin, M., Hashimshony, T., Wagner, F., & Yanai, I. (2012). Developmental milestones punctuate gene expression in the Caenorhabditis embryo. *Dev Cell, 22*(5), 1101-1108. doi:10.1016/j.devcel.2012.04.004

Levine, M. T., Eckert, M. L., & Begun, D. J. (2011). Whole-genome expression plasticity across tropical and temperate Drosophila melanogaster populations from Eastern Australia. *Mol Biol Evol, 28*(1), 249-256. doi:10.1093/molbev/msq197

Lewis, B. P., Burge, C. B., & Bartel, D. P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell, 120*(1), 15-20. doi:10.1016/j.cell.2004.12.035

Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P., & Burge, C. B. (2003). Prediction of mammalian microRNA targets. *Cell, 115*(7), 787-798. doi:10.1016/s0092-8674(03)01018-3

Li, J., Han, X., Wan, Y., Zhang, S., Zhao, Y., Fan, R., . . . Zhou, Y. (2018). TAM 2.0: tool for MicroRNA set analysis. *Nucleic Acids Res, 46*(W1), W180-W185. doi:10.1093/nar/gky509

Li, J. J., Huang, H., Bickel, P. J., & Brenner, S. E. (2014). Comparison of D. melanogaster and C. elegans developmental stages, tissues, and cells by modENCODE RNA-seq data. *Genome Res, 24*(7), 1086-1101. doi:10.1101/gr.170100.113

Li, K., Dan, Z., Gesang, L., Wang, H., Zhou, Y., Du, Y., . . . Nie, Y. (2016). Comparative Analysis of Gut Microbiota of Native Tibetan and Han Populations Living at Different Altitudes. *PLoS One, 11*(5), e0155863. doi:10.1371/journal.pone.0155863

Li, L., & Zhao, X. (2015). Comparative analyses of fecal microbiota in Tibetan and Chinese Han living at low or high altitude by barcoded 454 pyrosequencing. *Sci Rep, 5*, 14682. doi:10.1038/srep14682

Li, M., Jones-Rhoades, M. W., Lau, N. C., Bartel, D. P., & Rougvie, A. E. (2005). Regulatory mutations of mir-48, a C. elegans let-7 family MicroRNA, cause developmental timing defects. *Dev Cell, 9*(3), 415-422. doi:10.1016/j.devcel.2005.08.002

Liu, T., Yu, L., Liu, L., Li, H., & Li, Y. (2015). Comparative Transcriptomes and EVO-DEVO Studies Depending on Next Generation Sequencing. *Comput Math Methods Med, 2015*, 896176. doi:10.1155/2015/896176

Locke, A. E., Kahali, B., Berndt, S. I., Justice, A. E., Pers, T. H., Day, F. R., . . . Speliotes, E. K. (2015). Genetic studies of body mass index yield new insights for obesity biology. *Nature, 518*(7538), 197-206. doi:10.1038/nature14177

Ludwig, N., Leidinger, P., Becker, K., Backes, C., Fehlmann, T., Pallasch, C., . . . Keller, A. (2016). Distribution of miRNA expression across human tissues. *Nucleic Acids Res, 44*(8), 3865-3877. doi:10.1093/nar/gkw116

Lye, Z. N., & Purugganan, M. D. (2019). Copy Number Variation in Domestication. *Trends Plant Sci, 24*(4), 352-365. doi:10.1016/j.tplants.2019.01.003

Malik, A., Gildor, T., Sher, N., Layous, M., & Ben-Tabou de-Leon, S. (2017). Parallel embryonic transcriptional programs evolve under distinct constraints and may enable morphological conservation amidst adaptation. *Dev Biol, 430*(1), 202-213. doi:10.1016/j.ydbio.2017.07.019

Malnou, E. C., Umlauf, D., Mouysset, M., & Cavaille, J. (2018). Imprinted MicroRNA Gene Clusters in the Evolution, Development, and Functions of Mammalian Placenta. *Front Genet, 9*, 706. doi:10.3389/fgene.2018.00706

Malone, C. D., & Hannon, G. J. (2009). Small RNAs as guardians of the genome. *Cell, 136*(4), 656-668. doi:10.1016/j.cell.2009.01.045

Marcinkowska, M., Szymanski, M., Krzyzosiak, W. J., & Kozlowski, P. (2011). Copy number variation of microRNA genes in the human genome. *BMC Genomics, 12*, 183. doi:10.1186/1471-2164-12-183

Margolin, G., Khil, P. P., Kim, J., Bellani, M. A., & Camerini-Otero, R. D. (2014). Integrated transcriptome analysis of mouse spermatogenesis. *BMC Genomics, 15*, 39. doi:10.1186/1471-2164-15-39

Martin, S., Richier, S., Pedrotti, M. L., Dupont, S., Castejon, C., Gerakis, Y., . . . Gattuso, J. P. (2011). Early development and molecular plasticity in the Mediterranean sea urchin Paracentrotus lividus exposed to CO2-driven acidification. *J Exp Biol, 214*(Pt 8), 1357-1368. doi:10.1242/jeb.051169

McCarroll, S. A., & Altshuler, D. M. (2007). Copy-number variation and association studies of human disease. *Nat Genet, 39*(7 Suppl), S37-42. doi:10.1038/ng2080

Meunier, J., Lemoine, F., Soumillon, M., Liechti, A., Weier, M., Guschanski, K., . . . Kaessmann, H. (2013). Birth and expression evolution of mammalian microRNA genes. *Genome Res, 23*(1), 34-45. doi:10.1101/gr.140269.112

Mohammed, J., Flynt, A. S., Panzarino, A. M., Mondal, M. M. H., DeCruz, M., Siepel, A., & Lai, E. C. (2018). Deep experimental profiling of microRNA diversity, deployment, and evolution across the Drosophila genus. *Genome Res, 28*(1), 52-65. doi:10.1101/gr.226068.117

Nehrt, N. L., Clark, W. T., Radivojac, P., & Hahn, M. W. (2011). Testing the ortholog conjecture with comparative functional genomic data from mammals. *PLoS Comput Biol, 7*(6), e1002073. doi:10.1371/journal.pcbi.1002073

Ninova, M., Ronshaugen, M., & Griffiths-Jones, S. (2014). Conserved temporal patterns of microRNA expression in Drosophila support a developmental hourglass model. *Genome Biol Evol, 6*(9), 2459-2467. doi:10.1093/gbe/evu183

Noguer-Dance, M., Abu-Amero, S., Al-Khtib, M., Lefevre, A., Coullin, P., Moore, G. E., & Cavaille, J. (2010). The primate-specific microRNA gene cluster (C19MC) is imprinted in the placenta. *Human Molecular Genetics, 19*(18), 3566-3582. doi:10.1093/hmg/ddq272

O'Huallachain, M., Karczewski, K. J., Weissman, S. M., Urban, A. E., & Snyder, M. P. (2012). Extensive genetic variation in somatic human tissues. *Proc Natl Acad Sci U S A, 109*(44), 18018-18023. doi:10.1073/pnas.1213736109

Olena, A. F., & Patton, J. G. (2010). Genomic organization of microRNAs. *J Cell Physiol, 222*(3), 540-545. doi:10.1002/jcp.21993

Packer, J. S., Zhu, Q., Huynh, C., Sivaramakrishnan, P., Preston, E., Dueck, H., . . . Murray, J. I. (2019). A lineage-resolved molecular atlas of C. elegans embryogenesis at single-cell resolution. *Science, 365*(6459). doi:10.1126/science.aax1971

Pantalacci, S., & Semon, M. (2015). Transcriptomics of developing embryos and organs: A raising tool for evo-devo. *J Exp Zool B Mol Dev Evol, 324*(4), 363-371. doi:10.1002/jez.b.22595

Panwar, B., Omenn, G. S., & Guan, Y. F. (2017). miRmine: a database of human miRNA expression profiles. *Bioinformatics, 33*(10), 1554-1560. doi:10.1093/bioinformatics/btx019

Penso-Dolfin, L., Moxon, S., Haerty, W., & Di Palma, F. (2018). The evolutionary dynamics of microRNAs in domestic mammals. *Sci Rep, 8*(1), 17050. doi:10.1038/s41598-018-34243-8

Perry, G. H., Melsted, P., Marioni, J. C., Wang, Y., Bainer, R., Pickrell, J. K., . . . Gilad, Y. (2012). Comparative RNA sequencing reveals substantial genetic variation in endangered primates. *Genome Res, 22*(4), 602-610. doi:10.1101/gr.130468.111

Piasecka, B., Lichocki, P., Moretti, S., Bergmann, S., & Robinson-Rechavi, M. (2013). The hourglass and the early conservation models--co-existing patterns of developmental constraints in vertebrates. *PLoS Genet, 9*(4), e1003476. doi:10.1371/journal.pgen.1003476

Pijuan-Sala, B., Griffiths, J. A., Guibentif, C., Hiscock, T. W., Jawaid, W., Calero-Nieto, F. J., . . . Gottgens, B. (2019). A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature, 566*(7745), 490-495. doi:10.1038/s41586-019-0933-9

Poy, M. N., Eliasson, L., Krutzfeldt, J., Kuwajima, S., Ma, X., Macdonald, P. E., . . . Stoffel, M. (2004). A pancreatic islet-specific microRNA regulates insulin secretion. *Nature, 432*(7014), 226-230. doi:10.1038/nature03076

Price, A. L., Patterson, N., Hancks, D. C., Myers, S., Reich, D., Cheung, V. G., & Spielman, R. S. (2008). Effects of cis and trans genetic ancestry on gene expression in African Americans. *PLoS Genet, 4*(12), e1000294. doi:10.1371/journal.pgen.1000294

Prochnik, S. E., Rokhsar, D. S., & Aboobaker, A. A. (2007). Evidence for a microRNA expansion in the bilaterian ancestor. *Dev Genes Evol, 217*(1), 73-77. doi:10.1007/s00427-006-0116-1

Quint, M., Drost, H. G., Gabel, A., Ullrich, K. K., Bonn, M., & Grosse, I. (2012). A transcriptomic hourglass in plant embryogenesis. *Nature, 490*(7418), 98-101. doi:10.1038/nature11394

Rahmanian, S., Murad, R., Breschi, A., Zeng, W., Mackiewicz, M., Williams, B., . . . Mortazavi, A. (2019). Dynamics of microRNA expression during mouse prenatal development. *Genome Res, 29*(11), 1900-1909. doi:10.1101/gr.248997.119

Rawlings-Goss, R. A., Campbell, M. C., & Tishkoff, S. A. (2014). Global population-specific variation in miRNA associated with cancer risk and clinical biomarkers. *BMC Med Genomics, 7*, 53. doi:10.1186/1755-8794-7-53

Romero, I. G., Ruvinsky, I., & Gilad, Y. (2012). Comparative studies of gene expression and the evolution of gene regulation. *Nat Rev Genet, 13*(7), 505-516. doi:10.1038/nrg3229

Roux, J., & Robinson-Rechavi, M. (2008). Developmental constraints on vertebrate genome evolution. *PLoS Genet, 4*(12), e1000311. doi:10.1371/journal.pgen.1000311

Roux, J., Rosikiewicz, M., & Robinson-Rechavi, M. (2015). What to compare and how: Comparative transcriptomics for Evo-Devo. *J Exp Zool B Mol Dev Evol, 324*(4), 372-382. doi:10.1002/jez.b.22618

Runcie, D. E., Garfield, D. A., Babbitt, C. C., Wygoda, J. A., Mukherjee, S., & Wray, G. A. (2012). Genetics of gene expression responses to temperature stress in a sea urchin gene network. *Mol Ecol, 21*(18), 4547-4562. doi:10.1111/j.1365-294X.2012.05717.x

Sanchez, Y., & Huarte, M. (2013). Long non-coding RNAs: challenges for diagnosis and therapies. *Nucleic Acid Ther, 23*(1), 15-20. doi:10.1089/nat.2012.0414

Saw, W. Y., Tantoso, E., Begum, H., Zhou, L., Zou, R., He, C., . . . Teo, Y. Y. (2017). Establishing multiple omics baselines for three Southeast Asian populations in the Singapore Integrative Omics Study. *Nat Commun, 8*(1), 653. doi:10.1038/s41467-017-00413-x

Sayed, D., & Abdellatif, M. (2011). MicroRNAs in development and disease. *Physiol Rev, 91*(3), 827-887. doi:10.1152/physrev.00006.2010

Selbach, M., Schwanhausser, B., Thierfelder, N., Fang, Z., Khanin, R., & Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. *Nature, 455*(7209), 58-63. doi:10.1038/nature07228

Sempere, L. F., Cole, C. N., McPeek, M. A., & Peterson, K. J. (2006). The phylogenetic distribution of metazoan microRNAs: insights into evolutionary complexity and constraint. *J Exp Zool B Mol Dev Evol, 306*(6), 575-588. doi:10.1002/jez.b.21118

Simpson, Maciej Pajak and T. Ian. (2019). miRNAtap: miRNAtap: microRNA Targets - Aggregated Predictions.

Somel, M., Franz, H., Yan, Z., Lorenc, A., Guo, S., Giger, T., . . . Khaitovich, P. (2009). Transcriptional neoteny in the human brain. *Proc Natl Acad Sci U S A, 106*(14), 5743-5748. doi:10.1073/pnas.0900544106

Somel, M., Liu, X., Tang, L., Yan, Z., Hu, H., Guo, S., . . . Khaitovich, P. (2011). MicroRNA-driven developmental remodeling in the brain distinguishes humans from other primates. *PLoS Biol, 9*(12), e1001214. doi:10.1371/journal.pbio.1001214

Sood, R., Zehnder, J. L., Druzin, M. L., & Brown, P. O. (2006). Gene expression patterns in human placenta. *Proc Natl Acad Sci U S A, 103*(14), 5478-5483. doi:10.1073/pnas.0508035103

Spielman, R. S., Bastone, L. A., Burdick, J. T., Morley, M., Ewens, W. J., & Cheung, V. G. (2007). Common genetic variants account for differences in gene expression among ethnic groups. *Nat Genet, 39*(2), 226-231. doi:10.1038/ng1955

Storey, J. D., Madeoy, J., Strout, J. L., Wurfel, M., Ronald, J., & Akey, J. M. (2007). Gene-expression variation within and among human populations. *Am J Hum Genet, 80*(3), 502-509. doi:10.1086/512017

Stranger, B. E., Forrest, M. S., Dunning, M., Ingle, C. E., Beazley, C., Thorne, N., . . . Dermitzakis, E. T. (2007). Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science, 315*(5813), 848-853. doi:10.1126/science.1136678

Stranger, B. E., Montgomery, S. B., Dimas, A. S., Parts, L., Stegle, O., Ingle, C. E., . . . Dermitzakis, E. T. (2012). Patterns of cis regulatory variation in diverse human populations. *PLoS Genet, 8*(4), e1002639. doi:10.1371/journal.pgen.1002639

Sun, G., Yan, J., Noltner, K., Feng, J., Li, H., Sarkis, D. A., . . . Rossi, J. J. (2009). SNPs in human miRNA genes affect biogenesis and function. *RNA, 15*(9), 1640-1651. doi:10.1261/rna.1560209

Tan, M. H., Au, K. F., Yablonovitch, A. L., Wills, A. E., Chuang, J., Baker, J. C., . . . Li, J. B. (2013). RNA sequencing reveals a diverse and dynamic repertoire of the Xenopus tropicalis transcriptome over development. *Genome Res, 23*(1), 201-216. doi:10.1101/gr.141424.112

Tang, Y., Zhao, L., Lou, Y., Shi, Y., Fang, R., Lin, X., . . . Toga, A. (2018). Brain structure differences between Chinese and Caucasian cohorts: A comprehensive morphometry study. *Hum Brain Mapp, 39*(5), 2147-2155. doi:10.1002/hbm.23994

Tatusov, R. L., Koonin, E. V., & Lipman, D. J. (1997). A genomic perspective on protein families. *Science, 278*(5338), 631-637. doi:10.1126/science.278.5338.631

Tkachev, A., Stepanova, V., Zhang, L., Khrameeva, E., Zubkov, D., Giavalisco, P., & Khaitovich, P. (2019). Differences in lipidome and metabolome organization of prefrontal cortex among human populations. *Sci Rep, 9*(1), 18348. doi:10.1038/s41598-019-53762-6

Tomancak, P., Berman, B. P., Beaton, A., Weiszmann, R., Kwan, E., Hartenstein, V., . . . Rubin, G. M. (2007). Global analysis of patterns of gene expression during Drosophila embryogenesis. *Genome Biol, 8*(7), R145. doi:10.1186/gb-2007-8-7-r145

Uesaka, M., Kuratani, S., Takeda, H., & Irie, N. (2019). Recapitulation-like developmental transitions of chromatin accessibility in vertebrates. *Zoological Lett, 5*, 33. doi:10.1186/s40851-019-0148-9

Vickers, M. H., Breier, B. H., Cutfield, W. S., Hofman, P. L., & Gluckman, P. D. (2000). Fetal origins of hyperphagia, obesity, and hypertension and postnatal amplification by hypercaloric nutrition. *Am J Physiol Endocrinol Metab, 279*(1), E83-87. doi:10.1152/ajpendo.2000.279.1.E83

Vidigal, J. A., & Ventura, A. (2015). The biological functions of miRNAs: lessons from in vivo studies. *Trends Cell Biol, 25*(3), 137-147. doi:10.1016/j.tcb.2014.11.004

Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five years of GWAS discovery. *Am J Hum Genet, 90*(1), 7-24. doi:10.1016/j.ajhg.2011.11.029

Von Baer, K. E. (1828). Uber Entwicklungsgeschichte Der Thiere.

Wagner, D. E., Weinreb, C., Collins, Z. M., Briggs, J. A., Megason, S. G., & Klein, A. M. (2018). Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science, 360*(6392), 981-987. doi:10.1126/science.aar4362

Wagner, R. A., Tabibiazar, R., Liao, A., & Quertermous, T. (2005). Genome-wide expression dynamics during mouse embryonic development reveal similarities to Drosophila development. *Dev Biol, 288*(2), 595-611. doi:10.1016/j.ydbio.2005.09.036

Wang, L., Rishishwar, L., Marino-Ramirez, L., & Jordan, I. K. (2017). Human population-specific gene expression and transcriptional network modification with polymorphic transposable elements. *Nucleic Acids Res, 45*(5), 2318-2328. doi:10.1093/nar/gkw1286

Wang, X., Li, Y., Zhang, J., Zhang, Q., Liu, X., & Li, Z. (2017). De novo characterization of microRNAs in oriental fruit moth Grapholita molesta and selection of reference genes for normalization of microRNA expression. *PLoS One, 12*(2), e0171120. doi:10.1371/journal.pone.0171120

Wang, X., Park, J., Susztak, K., Zhang, N. R., & Li, M. (2019). Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. *Nat Commun, 10*(1), 380. doi:10.1038/s41467-018-08023-x

Warnefors, M., Mossinger, K., Halbert, J., Studer, T., VandeBerg, J. L., Lindgren, I., . . . Kaessmann, H. (2017). Sex-biased microRNA expression in mammals and birds reveals underlying regulatory mechanisms and a role in dosage compensation. *Genome Res, 27*(12), 1961-1973. doi:10.1101/gr.225391.117

Whitehead, A., & Crawford, D. L. (2005). Variation in tissue-specific gene expression among natural populations. *Genome Biol, 6*(2), R13. doi:10.1186/gb-2005-6-2-r13

Willer, C. J., Schmidt, E. M., Sengupta, S., Peloso, G. M., Gustafsson, S., Kanoni, S., . . . Global Lipids Genetics, Consortium. (2013). Discovery and refinement of loci associated with lipid levels. *Nat Genet, 45*(11), 1274-1283. doi:10.1038/ng.2797

Winter, J. (2015). MicroRNAs of the miR379-410 cluster: New players in embryonic neurogenesis and regulators of neuronal function. *Neurogenesis (Austin), 2*(1), e1004970. doi:10.1080/23262133.2015.1004970

Wolf, Y. I., & Koonin, E. V. (2012). A tight link between orthologs and bidirectional best hits in bacterial and archaeal genomes. *Genome Biol Evol, 4*(12), 1286-1294. doi:10.1093/gbe/evs100

Wood, A. R., Esko, T., Yang, J., Vedantam, S., Pers, T. H., Gustafsson, S., . . . Frayling, T. M. (2014). Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet, 46*(11), 1173-1186. doi:10.1038/ng.3097

Wright Willis, A., Evanoff, B. A., Lian, M., Criswell, S. R., & Racette, B. A. (2010). Geographic and ethnic variation in Parkinson disease: a population-based study of US Medicare beneficiaries. *Neuroepidemiology, 34*(3), 143-151. doi:10.1159/000275491

Wu, C. I., Shen, Y., & Tang, T. (2009). Evolution under canalization and the dual roles of microRNAs: a hypothesis. *Genome Res, 19*(5), 734-743. doi:10.1101/gr.084640.108

Xu, J., Zhang, R., Shen, Y., Liu, G., Lu, X., & Wu, C. I. (2013). The evolution of evolvability in microRNA target sites in vertebrates. *Genome Res, 23*(11), 1810-1816. doi:10.1101/gr.148916.112

Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., . . . Tang, F. (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol, 20*(9), 1131-1139. doi:10.1038/nsmb.2660

Yanai, I., Benjamin, H., Shmoish, M., Chalifa-Caspi, V., Shklar, M., Ophir, R., . . . Shmueli, O. (2005). Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics, 21*(5), 650-659. doi:10.1093/bioinformatics/bti042

Yanai, I., Peshkin, L., Jorgensen, P., & Kirschner, M. W. (2011). Mapping gene expression in two Xenopus species: evolutionary constraints and developmental flexibility. *Dev Cell, 20*(4), 483-496. doi:10.1016/j.devcel.2011.03.015

Yao, Z., Mich, J. K., Ku, S., Menon, V., Krostag, A. R., Martinez, R. A., . . . Ramanathan, S. (2017). A Single-Cell Roadmap of Lineage Bifurcation in Human ESC Models of Embryonic Brain Development. *Cell Stem Cell, 20*(1), 120-134. doi:10.1016/j.stem.2016.09.011

Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., Xu, F., . . . Wang, J. (2012). The oyster genome reveals stress adaptation and complexity of shell formation. *Nature, 490*(7418), 49-54. doi:10.1038/nature11413

Zhang, W., Duan, S., Kistner, E. O., Bleibel, W. K., Huang, R. S., Clark, T. A., . . . Dolan, M. E. (2008). Evaluation of genetic variation contributing to differences in gene expression between populations. *Am J Hum Genet, 82*(3), 631-640. doi:10.1016/j.ajhg.2007.12.015

Zhao, H., Shen, J., Medico, L., Wang, D., Ambrosone, C. B., & Liu, S. (2010). A pilot study of circulating miRNAs as potential biomarkers of early stage breast cancer. *PLoS One, 5*(10), e13735. doi:10.1371/journal.pone.0013735

Zhao, L., Wit, J., Svetec, N., & Begun, D. J. (2015). Parallel Gene Expression Differences between Low and High Latitude Populations of Drosophila melanogaster and D. simulans. *PLoS Genet, 11*(5), e1005184. doi:10.1371/journal.pgen.1005184

Zhao, Y., Samal, E., & Srivastava, D. (2005). Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature, 436*(7048), 214-220. doi:10.1038/nature03817

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., . . . Chanda, S. K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun, 10*(1), 1523. doi:10.1038/s41467-019-09234-6

Ziats, M. N., & Rennert, O. M. (2014). Identification of differentially expressed microRNAs across the developing human brain. *Mol Psychiatry, 19*(7), 848-852. doi:10.1038/mp.2013.93

Sample	Population group	Baby gender	Baby length	Baby weight	birth	Mother age(year)	Mother height(cm)	Mother weight	First pregancy	Pregnancy number	children	drinker	pregancy_infection	meds	bmi	Sequence batch	smoker	pregancy_drinker
E15_a	E	M	52.07	3203.5	natural	30	170.18	125	YES	0	0	YES	NO	NO	19.577567	1	NO	NO
E16_a	E	F	53.98	3628.74	cesarean	32	154.94	180	NO	3	3	NO	YES	YES	34.010353	1	NO	NO
A7_a	A	M	50.8	3486.99	cesarean	32	160.02	150	YES	0	0	YES	NO	NO	26.571035	1	NO	YES
A1_a	A	M	54.61	3883.88	natural	38	157.48	120	NO	2	2	NO	YES	NO	21.948062	1	NO	NO
X1_a	X	F	45.72	2069.52	natural	36	157.48	95	NO	2	1	NO	NO	NO	17.375549	1	NO	NO
E11_c	E	F	50.8	3033.4	natural	34	165.1	142	NO	2	2	YES	YES	NO	23.629794	1	NO	NO
A13_a	A	F	50.8	3430.29	natural	35	160.02	135	NO	2	0	NO	NO	YES	23.913931	1	NO	NO
S5_a	S	F	53.34	3968.93	cesarean	31	162.56	170	YES	0	0	NO	NO	NO	29.180134	1	NO	NO
X15_a	X	M	50.8	2891.65	natural	28	152.4	120	YES	0	0	YES	NO	YES	23.435653	1	NO	NO
X14_a	X	F	52.07	3231.85	natural	31	154.94	115	NO	1	1	NO	YES	YES	21.728837	1	NO	NO
E2_a	E	M	55.88	3515.34	natural	27	165.1	127.6	YES	0	0	NO	NO	NO	21.233533	1	NO	NO
S9_c	S	F	50.17	3255	cesarean	28	149.86	125	NO	1	1	NO	NO	NO	25.246681	1	NO	NO
S10_c	S	M	50.8	3742.14	natural	29	154.94	150	NO	1	1	NO	NO	NO	28.341961	1	NO	NO
E12_a	E	M	52.07	3543.69	natural	29	162.56	110	YES	0	0	YES	NO	YES	18.881263	1	NO	YES
E9_c	E	M	53.34	3770.49	cesarean	28	157.48	110	NO	3	3	NO	NO	NO	20.119057	2	NO	NO
A6_a	A	M	50.8	3373.6	cesarean	29	167.64	149	YES	0	0	NO	YES	NO	24.048982	2	NO	NO
A16_a	A	M	47.63	3401.94	cesarean	26	152.4	150	NO	2	1	YES	NO	NO	29.294566	2	NO	NO
X12_a	X	M	53.34	3657.09	natural	31	175.26	145	NO	2	1	NO	NO	YES	21.412537	2	NO	NO
X5_a	X	M	49.53	3175.15	cesarean	23	152.4	112	NO	1	1	NO	NO	NO	21.873276	2	NO	NO
S2_a	S	M	54.61	3345.24	natural	30.11	162.56	140	NO	1	1	NO	NO	YES	24.030699	2	NO	NO
S3_a	S	F	47.63	2834.95	cesarean	32	162.56	135	NO	1	1	NO	NO	NO	23.172459	2	NO	NO
E7_a	E	M	49.53	3316.89	natural	29	167.64	126	NO	1	1	YES	NO	NO	20.336723	2	NO	YES

S6_a	S	F	48.26	3316.89	natural	32	157.48	130	NO	1	1	NO	NO	YES	23.777067	2	NO	NO
E3_c	E	M	46.99	2409.71	natural	31	149.86	100	YES	0	0	YES	NO	NO	20.197345	2	NO	NO
S11_c	S	F	49.53	3288.54	natural	27	167.64	125	YES	0	0	NO	YES	NO	20.175321	2	NO	NO
A8_a	A	F	52.1	3260.19	natural	29	157.48	130	NO	3	3	YES	NO	NO	23.777067	2	NO	YES
X14_a	X	F	52.07	3231.85	natural	31	154.94	115	NO	1	1	NO	YES	YES	21.728837	2	NO	NO
A12_a	A	M	52.07	3770.49	natural	18	168.91	135	YES	0	0	YES	NO	NO	21.462919	2	NO	NO
X8_a	X	F	54.61	3600.39	natural	33	154.94	135	NO	3	0	NO	NO	NO	25.507765	3	NO	NO
A17_a	A	F	53.34	3183.65	natural	19	152.4	185	YES	0	0	NO	NO	NO	36.129965	3	NO	NO
X13_a	X	M	46.99	2778.25	cesarean	34	154.94	130	NO	2	2	NO	YES	YES	24.563033	3	NO	NO
X10_a	X	F	45.72	2211.26	natural	34	160.02	110	YES	0	0	NO	NO	YES	19.485425	3	NO	NO
S4_a	S	M	47.63	2755.57	natural	35	162.56	180	NO	1	1	NO	NO	NO	30.896612	3	NO	NO
X14_a	X	F	52.07	3231.85	natural	31	154.94	115	NO	1	1	NO	YES	YES	21.728837	3	NO	NO
E10_a	E	F	46.99	3460.29	cesarean	35	172.72	120	YES	0	0	YES	NO	YES	18.24575	3	NO	NO
A14_a	A	F	48.26	2579.81	natural	19	165.1	160	YES	0	0	NO	NO	NO	26.62512	3	NO	NO
E20_a	E	F	52.07	3770.49	natural	35	170.18	135	NO	3	2	YES	NO	NO	21.143772	3	NO	NO
S8_a	S	M	48.9	2579.81	cesarean	28	157.48	128	YES	0	0	NO	NO	NO	23.411266	3	NO	NO
X3_a	X	M	52.07	3883.88	natural	30	170.18	135	NO	1	1	NO	NO	NO	21.143772	3	NO	NO
S7_c	S	F	50.8	3458.64	cesarean	29	159.7	135	YES	0	0	NO	NO	NO	24.009863	3	NO	NO
A2_a	A	M	52.07	2891.65	natural	29	149.86	125	NO	1	1	YES	NO	NO	25.246681	3	NO	NO
X7_a	X	F	54.61	3460.29	natural	34	157.48	100	YES	0	0	NO	NO	NO	18.290052	3	NO	NO

Table S1. Sample information of human placenta

Sample ID	Species	Tissue	Stage	Total.reads	Trimmed.reads	Sequencing.run	RNA.extraction
761Nf	Chicken	Brain	5. Adult	5366782	3531424	OBIWAN 170	TRIzol/Chloroform
761Nf_rseq	Chicken	Brain	5. Adult	28104194	14644080	WINDU 82	TRIzol/Chloroform
653Nf	Chicken	Brain	5. Adult	4311960	2897719	OBIWAN 170	TRIzol/Chloroform
653Nf_rseq	Chicken	Brain	5. Adult	12341569	6821122	WINDU 82	TRIzol/Chloroform
653Nf_rseq2	Chicken	Brain	5. Adult	37829193	23317500	OBIWAN 214	TRIzol/Chloroform
922Nf	Chicken	Brain	5. Adult	3894732	1953198	OBIWAN 170	TRIzol/Chloroform
922Nf_rseq	Chicken	Brain	5. Adult	12134941	5005042	WINDU 82	TRIzol/Chloroform
922Nf_rseq2	Chicken	Brain	5. Adult	48344094	21283910	OBIWAN 214	TRIzol/Chloroform
783Nf	Chicken	Brain	5. Adult	3519652	2991379	OBIWAN 170	TRIzol/Chloroform
783Nf_rseq	Chicken	Brain	5. Adult	13327333	9397446	WINDU 82	TRIzol/Chloroform
5740Nf	Chicken	Brain	5. Adult	28201958	12523981	OBIWAN 198	QIAgen miRNeasy
650Nf	Chicken	Heart	5. Adult	3991143	2492691	OBIWAN 170	[likely Trizol/chloroform]
650Nf_rseq	Chicken	Heart	5. Adult	20311436	9790022	WINDU 82	[likely Trizol/chloroform]
764Nf	Chicken	Heart	5. Adult	3576115	2901141	OBIWAN 170	TRIzol/Chloroform
764Nf_rseq	Chicken	Heart	5. Adult	19577438	12673884	WINDU 82	TRIzol/Chloroform
919Nf	Chicken	Heart	5. Adult	2993917	2272035	OBIWAN 170	TRIzol/Chloroform
919Nf_rseq	Chicken	Heart	5. Adult	16313659	9863642	WINDU 82	TRIzol/Chloroform
560Nf	Chicken	Heart	5. Adult	5122113	4008854	OBIWAN 170	TRIzol/Chloroform
560Nf_rseq	Chicken	Heart	5. Adult	23606030	14455623	WINDU 82	TRIzol/Chloroform
785Nf	Chicken	Heart	5. Adult	5579964	4251302	OBIWAN 170	TRIzol/Chloroform
785Nf_rseq	Chicken	Heart	5. Adult	28846461	17047705	WINDU 82	TRIzol/Chloroform
5736Nf	Chicken	Heart	5. Adult	13279472	9693476	OBIWAN 190	QIAgen miRNeasy
5736Nf_rseq	Chicken	Heart	5. Adult	11966261	8066709	OBIWAN 210	QIAgen miRNeasy
554Nf	Chicken	Liver	5. Adult	18282002	11878395	WINDU 67	TRIzol/Chloroform
5738Nf	Chicken	Liver	5. Adult	28361064	17828460	WINDU 67	QIAgen miRNeasy

917Nf	Chicken	Liver	5. Adult	14581174	8709557	WINDU 67	TRIzol/Chloroform
917Nf_rseq	Chicken	Liver	5. Adult	9586551	6371489	OBIWAN 210	TRIzol/Chloroform
5737Nf	Chicken	Liver	5. Adult	15801274	10374071	WINDU 67	QIAgen miRNeasy
5737Nf_rseq	Chicken	Liver	5. Adult	30336051	19728883	WINDU 168	QIAgen miRNeasy
558Nf	Chicken	Liver	5. Adult	4608549	3539788	OBIWAN 170	TRIzol/Chloroform
558Nf_rseq	Chicken	Liver	5. Adult	24791706	15684840	WINDU 82	TRIzol/Chloroform
786Nf	Chicken	Liver	5. Adult	4184150	3240921	OBIWAN 170	TRIzol/Chloroform
786Nf_rseq	Chicken	Liver	5. Adult	21900287	13897845	WINDU 82	TRIzol/Chloroform
6167Nf	Chicken	Brain	5. Adult	2159900	1395819	WINDU 142	QIAgen miRNeasy
6167Nf_rseq	Chicken	Brain	5. Adult	48495517	34547377	OBIWAN 263	QIAgen miRNeasy
1353Nf	Chicken	Brain	2. Late embryo	13742648	8897231	OBIWAN 192	QIAgen miRNeasy
1353Nf_rseq	Chicken	Brain	2. Late embryo	16196146	9431864	OBiWAN 210	QIAgen miRNeasy
1364Nf	Chicken	Brain	2. Late embryo	14830987	10614788	OBIWAN 192	QIAgen miRNeasy
1364Nf_rseq	Chicken	Brain	2. Late embryo	29533350	17773305	WINDU 168	QIAgen miRNeasy
1471Nf	Chicken	Heart	2. Late embryo	19324312	14373988	OBIWAN 190	QIAgen miRNeasy
1472Nf	Chicken	Heart	2. Late embryo	17637290	12105322	OBIWAN 190	QIAgen miRNeasy
1535Nf	Chicken	Liver	2. Late embryo	15261730	10627895	OBIWAN 190	QIAgen miRNeasy
1535Nf_rseq	Chicken	Liver	2. Late embryo	29055176	18423212	WINDU 168	QIAgen miRNeasy
1610Nf	Chicken	Liver	2. Late embryo	14085732	6614609	OBIWAN 222	QIAgen miRNeasy
1610Nf_rseq	Chicken	Liver	2. Late embryo	29240602	12001363	WINDU 114	QIAgen miRNeasy
1539Nf	Chicken	Brain	3. Neonate	16344400	10155143	OBIWAN 198	QIAgen miRNeasy
1539Nf_rseq	Chicken	Brain	3. Neonate	24176767	14708078	WINDU 168	QIAgen miRNeasy
1806Nf	Chicken	Brain	3. Neonate	96760185	65329399	OBIWAN 202	QIAgen miRNeasy
1541Nf	Chicken	Heart	3. Neonate	12658613	7261186	WINDU 67	QIAgen miRNeasy
1541Nf_rseq	Chicken	Heart	3. Neonate	6874587	4457851	OBIWAN 210	QIAgen miRNeasy
1659Nf	Chicken	Heart	3. Neonate	18868005	14263281	OBIWAN 192	QIAgen miRNeasy

1661Nf	Chicken	Liver	3. Neonate	70218276	10139502	OBIWAN 202	QIAgen miRNeasy
1661Nf_rseq	Chicken	Liver	3. Neonate	57784719	31850743	WINDU 168	QIAgen miRNeasy
1543Nf	Chicken	Liver	3. Neonate	23356239	10461306	OBIWAN 210	QIAgen miRNeasy
1543Nf_rseq	Chicken	Liver	3. Neonate	30846391	13810693	WINDU 168	QIAgen miRNeasy
1647Nf	Chicken	Brain	4. Juvenile	12868637	8321892	OBIWAN 198	QIAgen miRNeasy
1647Nf_rseq	Chicken	Brain	4. Juvenile	13819169	9426944	OBIWAN 210	QIAgen miRNeasy
1647Nf_rseq2	Chicken	Brain	4. Juvenile	15024907	10294246	OBIWAN 214	QIAgen miRNeasy
1654Nf	Chicken	Brain	4. Juvenile	9316188	5925988	OBIWAN 198	QIAgen miRNeasy
1654Nf_rseq	Chicken	Brain	4. Juvenile	29009013	18004488	OBIWAN 214	QIAgen miRNeasy
1649Nf	Chicken	Heart	4. Juvenile	12095458	7368158	WINDU 67	QIAgen miRNeasy
1649Nf_rseq	Chicken	Heart	4. Juvenile	11881540	7573017	OBIWAN 210	QIAgen miRNeasy
1649Nf_rseq2	Chicken	Heart	4. Juvenile	13133207	8475711	OBIWAN 214	QIAgen miRNeasy
1656Nf	Chicken	Heart	4. Juvenile	17617837	11379199	OBIWAN 192	QIAgen miRNeasy
1651Nf	Chicken	Liver	4. Juvenile	11156320	5777593	OBIWAN 198	QIAgen miRNeasy
1651Nf_rseq	Chicken	Liver	4. Juvenile	30906017	16926933	OBIWAN 214	QIAgen miRNeasy
1679Nf	Chicken	Liver	4. Juvenile	14436794	6020249	OBIWAN 210	QIAgen miRNeasy
1679Nf_rseq	Chicken	Liver	4. Juvenile	49044357	17577547	WINDU 114	QIAgen miRNeasy
5893Nf	Human	Liver	3. Neonate	50908029	24302304	OBIWAN 210	QIAgen RNeasy Micro modified
5893Nf_rseq	Human	Liver	3. Neonate	60470237	24429305	WINDU 168	QIAgen RNeasy Micro modified
5982Nf	Human	Heart	3. Neonate	1358363	701095	WINDU 142	QIAgen RNeasy Mini modified
5982Nf_rseq	Human	Heart	3. Neonate	119141370	70766344	OBIWAN 283	QIAgen RNeasy Mini modified
6042Nf	Human	Heart	5. Adult	11953770	6241347	OBIWAN 260	QIAgen miRNeasy
6042Nf_rseq	Human	Heart	5. Adult	127028761	74465552	OBIWAN 283	QIAgen miRNeasy
6048Nf	Human	Brain	5. Adult	39011081	8453223	WINDU 193	QIAgen RNeasy Mini modified
6048Nf_rseq	Human	Brain	5. Adult	52648295	12262850	WINDU 214	QIAgen RNeasy Mini modified
6295Nf	Human	Liver	5. Adult	35914484	3257751	WINDU 193	QIAgen RNeasy Mini modified

6295Nf_rseq	Human	Liver	5. Adult	52008581	4863982	WINDU 209	QIAgen RNeasy Mini modified
6147Nf	Human	Brain	3. Neonate	38672346	20288626	WINDU 142	QIAgen RNeasy Micro modified
6147Nf_rseq	Human	Brain	3. Neonate	70097401	36010160	WINDU 168	QIAgen RNeasy Micro modified
6043Nf	Human	Heart	5. Adult	96309702	33685666	WINDU 209	QIAgen miRNeasy
6106Nf	Human	Liver	5. Adult	31263266	11651810	OBIWAN 247	QIAgen RNeasy Mini modified
6106Nf_rseq	Human	Liver	5. Adult	40889573	15010937	WINDU 168	QIAgen RNeasy Mini modified
6362Nf	Human	Heart	3. Neonate	37153467	3590280	WINDU 214	QIAgen RNeasy Micro modified
5986Nf	Human	Liver	4. Juvenile	9603525	2682016	OBIWAN 222	QIAgen RNeasy Mini modified
5986Nf_rseq	Human	Liver	4. Juvenile	124584346	35166571	OBIWAN 283	QIAgen RNeasy Mini modified
5984Nf	Human	Brain	4. Juvenile	59463153	20080470	WINDU 209	QIAgen RNeasy Mini modified
5985Nf	Human	Brain	4. Juvenile	7810247	4889198	WINDU 190	QIAgen RNeasy Mini modified
5985Nf_rseq	Human	Brain	4. Juvenile	57618045	31193029	WINDU 209	QIAgen RNeasy Mini modified
3671Nf	Human	Heart	1. Early embryo	4962259	3341372	OBIWAN 210	QIAgen RNeasy Micro modified
2076Nf	Human	Liver	1. Early embryo	3380410	2500459	OBIWAN 192	QIAgen RNeasy Micro modified
2076Nf_rseq	Human	Liver	1. Early embryo	18053324	13280813	OBIWAN 214	QIAgen RNeasy Micro modified
2088Nf	Human	Heart	1. Early embryo	30534685	13196777	OBIWAN 210	QIAgen miRNeasy
2088Nf_rseq	Human	Heart	1. Early embryo	78119714	32216005	WINDU 168	QIAgen miRNeasy
3676Nf	Human	Liver	1. Early embryo	13594767	9877671	OBIWAN 192	QIAgen miRNeasy
3676Nf_rseq	Human	Liver	1. Early embryo	17993132	11760895	OBIWAN 210	QIAgen miRNeasy
2077Nf	Human	Brain	1. Early embryo	13885651	11006339	OBIWAN 192	QIAgen miRNeasy
2077Nf_rseq	Human	Brain	1. Early embryo	23525516	17377318	WINDU 168	QIAgen miRNeasy
2071Nf	Human	Heart	2. Late embryo	8687898	7027955	OBIWAN 192	QIAgen miRNeasy
2071Nf_rseq	Human	Heart	2. Late embryo	24205841	18062486	OBIWAN 210	QIAgen miRNeasy
2071Nf_rseq2	Human	Heart	2. Late embryo	26715603	19967793	OBIWAN 214	QIAgen miRNeasy
2039Nf	Human	Heart	2. Late embryo	8271136	5481289	OBIWAN 202	QIAgen miRNeasy
2039Nf_rseq	Human	Heart	2. Late embryo	24766706	18910311	OBIWAN 214	QIAgen miRNeasy

2022Nf	Human	Liver	2. Late embryo	9171587	7666323	OBIWAN 192	QIAgen miRNeasy
2022Nf_rseq	Human	Liver	2. Late embryo	16269992	12646316	OBIWAN 210	QIAgen miRNeasy
2022Nf_rseq2	Human	Liver	2. Late embryo	17936627	13927210	OBIWAN 214	QIAgen miRNeasy
2072Nf	Human	Liver	2. Late embryo	11993996	8427635	OBIWAN 192	QIAgen miRNeasy
2072Nf_rseq	Human	Liver	2. Late embryo	10749424	7300907	OBIWAN 210	QIAgen miRNeasy
2068Nf	Human	Brain	2. Late embryo	23205111	18478349	OBIWAN 202	QIAgen miRNeasy
2041Nf	Human	Brain	2. Late embryo	14335059	12151039	OBIWAN 210	QIAgen miRNeasy
6003Nf	Macaque	Brain	2. Late embryo	40197239	20282138	OBIWAN 247	QIAgen RNeasy Mini modified
6021Nf	Macaque	Heart	2. Late embryo	32198684	20020945	WINDU 142	QIAgen miRNeasy
6002Nf	Macaque	Brain	2. Late embryo	38038443	16205938	OBIWAN 247	QIAgen RNeasy Mini modified
6020Nf	Macaque	Heart	2. Late embryo	7607357	3630419	WINDU 190	QIAgen miRNeasy
6020Nf_rseq	Macaque	Heart	2. Late embryo	73235032	28291507	WINDU 209	QIAgen miRNeasy
6001Nf	Macaque	Liver	2. Late embryo	38277150	26712052	WINDU 142	QIAgen RNeasy Mini modified
6019Nf	Macaque	Heart	1. Early embryo	14169573	7115142	WINDU 142	QIAgen miRNeasy
6019Nf_rseq	Macaque	Heart	1. Early embryo	21748349	13930286	OBIWAN 280	QIAgen miRNeasy
5997Nf	Macaque	Liver	1. Early embryo	25286518	20483800	OBIWAN 260	QIAgen RNeasy Mini modified
5999Nf	Macaque	Brain	1. Early embryo	22949616	17833793	OBIWAN 260	QIAgen RNeasy Mini modified
6018Nf	Macaque	Heart	1. Early embryo	19865092	9607897	OBIWAN 260	QIAgen miRNeasy
6018Nf_rseq	Macaque	Heart	1. Early embryo	3683294	2043540	OBIWAN 280	QIAgen miRNeasy
5994Nf	Macaque	Liver	1. Early embryo	16751595	11755456	WINDU 142	QIAgen RNeasy Mini modified
5996Nf	Macaque	Brain	1. Early embryo	17092350	11400018	WINDU 142	QIAgen RNeasy Mini modified
6009Nf	Macaque	Brain	3. Neonate	20632816	15433115	OBIWAN 260	QIAgen RNeasy Mini modified
6023Nf	Macaque	Heart	3. Neonate	68583013	45804080	OBIWAN 283	QIAgen miRNeasy
6010Nf	Macaque	Liver	3. Neonate	67934915	47550802	OBIWAN 283	QIAgen RNeasy Mini modified
6015Nf	Macaque	Brain	4. Juvenile	18001078	7064706	OBIWAN 263	QIAgen RNeasy Mini modified
6015Nf_rseq	Macaque	Brain	4. Juvenile	17688441	8284887	OBIWAN 280	QIAgen RNeasy Mini modified

6025Nf	Macaque	Heart	4. Juvenile	33494492	12097589	OBIWAN 247	QIAgen miRNeasy
6311Nf	Macaque	Liver	4. Juvenile	57499681	14787807	WINDU 208	QIAgen RNeasy Mini modified
6024Nf	Macaque	Heart	4. Juvenile	33876945	16644944	WINDU 142	QIAgen miRNeasy
6013Nf	Macaque	Liver	4. Juvenile	14903781	10607251	OBIWAN 260	QIAgen RNeasy Mini modified
6013Nf_rseq	Macaque	Liver	4. Juvenile	24155891	17114585	WINDU 168	QIAgen RNeasy Mini modified
6364Nf	Macaque	Brain	4. Juvenile	44026670	1038639	WINDU 214	QIAgen RNeasy Mini modified
6006Nf	Macaque	Brain	3. Neonate	14776878	9321294	WINDU 142	QIAgen RNeasy Mini modified
6022Nf	Macaque	Heart	3. Neonate	93368181	45986639	WINDU 213	QIAgen miRNeasy
6022Nf_rseq	Macaque	Heart	3. Neonate	30948943	19113696	WINDU 214	QIAgen miRNeasy
6007Nf	Macaque	Liver	3. Neonate	9708128	5721977	OBIWAN 260	QIAgen RNeasy Mini modified
6007Nf_rseq	Macaque	Liver	3. Neonate	32787152	20780950	WINDU 168	QIAgen RNeasy Mini modified
5988Nf	Macaque	Brain	5. Adult	11178070	4255539	WINDU 108	QIAgen RNeasy Mini modified
5988Nf_rseq	Macaque	Brain	5. Adult	17514935	6661657	WINDU 119	QIAgen RNeasy Mini modified
5988Nf_rseq2	Macaque	Brain	5. Adult	21782095	9597092	WINDU 168	QIAgen RNeasy Mini modified
6026Nf	Macaque	Heart	5. Adult	54629202	22574923	WINDU 186	QIAgen miRNeasy
5989Nf	Macaque	Liver	5. Adult	11620004	6359924	WINDU 108	QIAgen RNeasy Mini modified
5989Nf_rseq	Macaque	Liver	5. Adult	20181094	10622767	WINDU 111	QIAgen RNeasy Mini modified
5993Nf	Macaque	Liver	5. Adult	12396418	7068171	WINDU 108	QIAgen RNeasy Mini modified
5993Nf_rseq	Macaque	Liver	5. Adult	22314326	12133393	WINDU 111	QIAgen RNeasy Mini modified
5814Nf	Marmoset	Heart	5. Adult	22064641	11810931	OBIWAN 210	QIAgen RNeasy Micro modified
5787Nf	Marmoset	Liver	5. Adult	16442899	12138539	OBIWAN 202	QIAgen RNeasy Micro modified
5798Nf	Marmoset	Liver	5. Adult	11657269	6170449	OBIWAN 210	QIAgen RNeasy Micro modified
5798Nf_rseq	Marmoset	Liver	5. Adult	19383914	9481912	WINDU 114	QIAgen RNeasy Micro modified
5800Nf	Marmoset	Brain	5. Adult	16664387	9393008	OBIWAN 202	QIAgen RNeasy Micro modified
5800Nf_rseq	Marmoset	Brain	5. Adult	19345589	11075784	OBIWAN 210	QIAgen RNeasy Micro modified
5801Nf	Marmoset	Brain	5. Adult	8516820	4735961	OBIWAN 202	QIAgen RNeasy Micro modified

5801Nf_rseq	Marmoset	Brain	5. Adult	25526956	15735208	OBIWAN 214	QIAgen RNeasy Micro modified
5313Nf	Marmoset	Brain	1. Early embryo	15583926	12142902	OBIWAN 202	QIAgen RNeasy Micro modified
5250Nf	Marmoset	Heart	1. Early embryo	5694589	3288106	OBIWAN 202	QIAgen RNeasy Micro modified
5250Nf_rseq	Marmoset	Heart	1. Early embryo	30065978	19917923	OBIWAN 214	QIAgen RNeasy Micro modified
5248Nf	Marmoset	Liver	1. Early embryo	41041491	27223566	OBIWAN 263	QIAgen Rneasy Micro modified
5315Nf	Marmoset	Brain	1. Early embryo	14964749	11803951	OBIWAN 202	QIAgen RNeasy Micro modified
5251Nf	Marmoset	Heart	1. Early embryo	25050858	14864597	OBIWAN 210	QIAgen RNeasy Micro modified
5241Nf	Marmoset	Liver	1. Early embryo	29974104	15271576	OBIWAN 210	QIAgen Rneasy Micro modified
5279Nf	Marmoset	Brain	2. Late embryo	22815552	16589990	OBIWAN 202	QIAgen RNeasy Micro modified
5282Nf	Marmoset	Heart	2. Late embryo	25999210	17465993	OBIWAN 192	QIAgen RNeasy Micro modified
5807Nf	Marmoset	Heart	2. Late embryo	42896166	16713328	WINDU 193	QIAgen RNeasy Micro modified
5284Nf	Marmoset	Liver	2. Late embryo	34403457	24289641	OBIWAN 210	QIAgen Rneasy Micro modified
5309Nf	Marmoset	Brain	2. Late embryo	15267940	9475519	WINDU 108	QIAgen RNeasy Micro modified
5309Nf_rseq	Marmoset	Brain	2. Late embryo	30046439	17948114	WINDU 111	QIAgen RNeasy Micro modified
6350Nf	Marmoset	Liver	2. Late embryo	45087691	13170849	WINDU 208	QIAgen RNeasy Micro modified
5815Nf	Marmoset	Heart	4. Juvenile	9137398	5465954	OBIWAN 210	QIAgen RNeasy Micro modified
5815Nf_rseq	Marmoset	Heart	4. Juvenile	11939950	6211182	WINDU 114	QIAgen RNeasy Micro modified
5812Nf	Marmoset	Heart	4. Juvenile	37316909	17562571	OBIWAN 263	QIAgen RNeasy Micro modified
5789Nf	Marmoset	Liver	4. Juvenile	8553564	5769136	OBIWAN 202	QIAgen RNeasy Micro modified
5789Nf_rseq	Marmoset	Liver	4. Juvenile	25117991	17877590	OBIWAN 214	QIAgen RNeasy Micro modified
5813Nf	Marmoset	Liver	4. Juvenile	3788601	2739281	OBIWAN 210	QIAgen RNeasy Micro modified
5813Nf_rseq	Marmoset	Liver	4. Juvenile	17072407	12497206	OBIWAN 240	QIAgen RNeasy Micro modified
5788Nf	Marmoset	Brain	4. Juvenile	12702974	9291754	OBIWAN 202	QIAgen RNeasy Micro modified
5788Nf_rseq	Marmoset	Brain	4. Juvenile	15310110	11287429	OBIWAN 210	QIAgen RNeasy Micro modified
5811Nf	Marmoset	Brain	4. Juvenile	50885761	27226848	OBIWAN 263	QIAgen RNeasy Micro modified
5269Nf	Marmoset	Brain	3. Neonate	21092052	15355321	OBIWAN 192	QIAgen RNeasy Micro modified

4330Nf	Marmoset	Brain	3. Neonate	15900887	12399533	OBIWAN 210	QIAgen RNeasy Micro modified
5809Nf	Marmoset	Heart	3. Neonate	23575944	12179995	OBIWAN 222	QIAgen RNeasy Micro modified
4332Nf	Marmoset	Liver	3. Neonate	30397806	23057573	OBIWAN 210	QIAgen RNeasy Micro modified
4332Nf_rseq	Marmoset	Liver	3. Neonate	33534532	25321916	OBIWAN 214	QIAgen RNeasy Micro modified
5810Nf	Marmoset	Liver	3. Neonate	7791780	4585717	OBIWAN 210	QIAgen RNeasy Micro modified
5810Nf_rseq	Marmoset	Liver	3. Neonate	27266818	13790078	WINDU 114	QIAgen RNeasy Micro modified
6181Nf	Marmoset	Heart	3. Neonate	11580645	5891082	WINDU 190	QIAgen RNeasy Micro modified
6181Nf_rseq	Marmoset	Heart	3. Neonate	58014708	24304011	WINDU 209	QIAgen RNeasy Micro modified
1880Nf	Mouse	Brain	4. Juvenile	13183371	6544168	OBIWAN 170	QIAgen miRNeasy
1880Nf_rseq	Mouse	Brain	4. Juvenile	18712236	9967132	OBIWAN 190	QIAgen miRNeasy
1876Nf	Mouse	Brain	4. Juvenile	23434428	13238158	OBIWAN 165	QIAgen miRNeasy
2643Nf	Mouse	Heart	4. Juvenile	34961824	22678400	OBIWAN 170	QIAgen miRNeasy
2645Nf	Mouse	Heart	4. Juvenile	34915184	21660913	OBIWAN 170	QIAgen miRNeasy
1878Nf	Mouse	Liver	4. Juvenile	17469163	12797934	OBIWAN 170	QIAgen miRNeasy
1882Nf	Mouse	Liver	4. Juvenile	21272917	7301972	WINDU 48	QIAgen miRNeasy
1882Nf_rseq	Mouse	Liver	4. Juvenile	19798517	7704344	WINDU 48	QIAgen miRNeasy
1944Nf	Mouse	Brain	5. Adult	19678998	13752405	WINDU 67	QIAgen miRNeasy
1948Nf	Mouse	Brain	5. Adult	10902927	7020495	OBIWAN 170	QIAgen miRNeasy
1948Nf_rseq	Mouse	Brain	5. Adult	17405298	11842422	OBIWAN 190	QIAgen miRNeasy
1960Nf	Mouse	Brain	5. Adult	27522613	19392888	OBIWAN 165	QIAgen miRNeasy
1956Nf	Mouse	Heart	5. Adult	14710731	7889861	OBIWAN 170	QIAgen miRNeasy
1956Nf_rseq	Mouse	Heart	5. Adult	17618707	7631811	WINDU 82	QIAgen miRNeasy
2663Nf	Mouse	Heart	5. Adult	7586786	4093603	OBIWAN 165	QIAgen miRNeasy
2663Nf_rseq	Mouse	Heart	5. Adult	22564280	11505051	OBIWAN 210	QIAgen miRNeasy
1950Nf	Mouse	Heart	5. Adult	9718687	6434132	OBIWAN 165	QIAgen miRNeasy
1950Nf_rseq	Mouse	Heart	5. Adult	36186713	22399353	OBIWAN 214	QIAgen miRNeasy

2669Nf	Mouse	Heart	5. Adult	12597867	8265720	OBIWAN 165	QIAgen miRNeasy
2669Nf_rseq	Mouse	Heart	5. Adult	9712442	6102795	OBIWAN 210	QIAgen miRNeasy
1946Nf	Mouse	Liver	5. Adult	23754848	18924169	OBIWAN 202	QIAgen miRNeasy
1958Nf	Mouse	Liver	5. Adult	10233655	8764644	OBIWAN 202	QIAgen miRNeasy
1958Nf_rseq	Mouse	Liver	5. Adult	10568327	8403983	OBIWAN 210	QIAgen miRNeasy
1962Nf	Mouse	Liver	5. Adult	10429309	7967863	OBIWAN 192	QIAgen miRNeasy
1962Nf_rseq	Mouse	Liver	5. Adult	13342382	9799649	OBIWAN 210	QIAgen miRNeasy
6170Nf	Mouse	Heart	5. Adult	38053404	23910424	OBIWAN 247	QIAgen miRNeasy
6139Nf	Mouse	Brain	5. Adult	8738962	5756367	WINDU 108	QIAgen miRNeasy
6139Nf_rseq	Mouse	Brain	5. Adult	25922941	20090510	OBIWAN 263	QIAgen miRNeasy
6140Nf	Mouse	Liver	5. Adult	55430085	24126130	WINDU 142	QIAgen miRNeasy
6142Nf	Mouse	Brain	5. Adult	35441845	15331453	WINDU 110	QIAgen miRNeasy
6143Nf	Mouse	Heart	5. Adult	30324367	15946840	WINDU 110	QIAgen miRNeasy
6144Nf	Mouse	Liver	5. Adult	34989722	16621383	WINDU 110	QIAgen miRNeasy
1954Nf	Mouse	Brain	5. Adult	14097566	9804432	OBIWAN 140	QIAgen miRNeasy
1954Nf_rseq	Mouse	Brain	5. Adult	12970167	9180978	OBIWAN 140	QIAgen miRNeasy
1952Nf	Mouse	Liver	5. Adult	17032443	10124271	OBIWAN 140	QIAgen miRNeasy
1952Nf_rseq	Mouse	Liver	5. Adult	17934701	10754448	OBIWAN 140	QIAgen miRNeasy
3036Nf	Mouse	Heart	1. Early embryo	7613908	3224245	OBIWAN 202	QIAgen miRneasy Micro
3036Nf_rseq	Mouse	Heart	1. Early embryo	24043139	9837062	OBIWAN 210	QIAgen miRneasy Micro
3037Nf	Mouse	Liver	1. Early embryo	16146450	9166741	OBIWAN 190	QIAgen miRneasy Micro
3037Nf_rseq	Mouse	Liver	1. Early embryo	14966596	7940481	OBIWAN 210	QIAgen miRneasy Micro
3042Nf	Mouse	Liver	1. Early embryo	14148529	5952954	OBIWAN 202	QIAgen miRneasy Micro
3042Nf_rseq	Mouse	Liver	1. Early embryo	45033454	21642000	OBIWAN 214	QIAgen miRneasy Micro
3040Nf	Mouse	Brain	1. Early embryo	14999954	9311517	OBIWAN 192	QIAgen miRneasy Micro
3040Nf_rseq	Mouse	Brain	1. Early embryo	12765868	7056021	OBIWAN 210	QIAgen miRneasy Micro

6354Nf	Mouse	Brain	1. Early embryo	48867289	8695462	WINDU 208	QIAgen miRNeasy Mini
6352Nf	Mouse	Heart	1. Early embryo	10915567	6582205	WINDU 190	QIAgen miRneasy Micro
6352Nf_rseq	Mouse	Heart	1. Early embryo	43024066	20961822	WINDU 209	QIAgen miRneasy Micro
2257Nf	Mouse	Brain	2. Late embryo	40083965	13741910	WINDU 63	QIAgen miRNeasy
2276Nf	Mouse	Brain	2. Late embryo	35129550	15313292	WINDU 63	QIAgen miRNeasy
2514Nf	Mouse	Heart	2. Late embryo	8750259	6211625	OBIWAN 192	QIAgen miRNeasy
2514Nf_rseq	Mouse	Heart	2. Late embryo	45773207	29749347	OBIWAN 214	QIAgen miRNeasy
2512Nf	Mouse	Heart	2. Late embryo	15308274	7340062	OBIWAN 222	QIAgen miRNeasy
2512Nf_rseq	Mouse	Heart	2. Late embryo	20407535	8197131	WINDU 114	QIAgen miRNeasy
2279Nf	Mouse	Liver	2. Late embryo	18893027	8578525	OBIWAN 190	QIAgen miRNeasy
2279Nf_rseq	Mouse	Liver	2. Late embryo	15929367	7354405	OBIWAN 210	QIAgen miRNeasy
6353Nf	Mouse	Liver	2. Late embryo	40979053	12893170	WINDU 208	QIAgen miRNeasy Mini
2272Nf	Mouse	Liver	2. Late embryo	18037151	6805013	OBIWAN 170	QIAgen miRNeasy
2272Nf_rseq	Mouse	Liver	2. Late embryo	14571011	5452774	OBIWAN 190	QIAgen miRNeasy
1906Nf	Mouse	Brain	3. Neonate	52842092	26099534	WINDU 63	QIAgen miRNeasy
1912Nf	Mouse	Brain	3. Neonate	45851372	16777229	WINDU 63	QIAgen miRNeasy
1908Nf	Mouse	Heart	3. Neonate	12382601	6562495	OBIWAN 170	QIAgen miRNeasy
1908Nf_rseq	Mouse	Heart	3. Neonate	13363988	6927375	OBIWAN 190	QIAgen miRNeasy
1914Nf	Mouse	Heart	3. Neonate	14817183	10184656	OBIWAN 222	QIAgen miRNeasy
1914Nf_rseq	Mouse	Heart	3. Neonate	25382420	17978104	WINDU 168	QIAgen miRNeasy
1916Nf	Mouse	Liver	3. Neonate	18551797	10146576	OBIWAN 170	QIAgen miRNeasy
1916Nf_rseq	Mouse	Liver	3. Neonate	24569368	13544062	WINDU 168	QIAgen miRNeasy
1910Nf	Mouse	Liver	3. Neonate	23113996	12541713	WINDU 142	QIAgen miRNeasy
4209Nf	Opossum	Heart	5. Adult	44183274	25990319	OBIWAN 247	QIAgen miRNeasy
6171Nf	Opossum	Brain	5. Adult	37897280	20335032	OBIWAN 247	QIAgen miRNeasy
4213Nf	Opossum	Heart	5. Adult	31573705	16417583	OBIWAN 247	QIAgen miRNeasy

3988Nf	Opossum	Brain	5. Adult	10584277	8352693	OBIWAN 170	QIAgen miRNeasy
3988Nf_rseq	Opossum	Brain	5. Adult	17308213	13385879	WINDU 92	QIAgen miRNeasy
3951Nf	Opossum	Brain	5. Adult	22650915	11678014	WINDU 63	QIAgen miRNeasy
5729Nf	Opossum	Brain	5. Adult	22831850	17791542	OBIWAN 165	QIAgen miRNeasy
3957Nf	Opossum	Brain	5. Adult	9658481	7794379	OBIWAN 202	QIAgen miRNeasy
3957Nf_rseq	Opossum	Brain	5. Adult	9410634	7529841	OBIWAN 210	QIAgen miRNeasy
3994Nf	Opossum	Brain	5. Adult	13511307	8735726	OBIWAN 202	QIAgen miRNeasy
3994Nf_rseq	Opossum	Brain	5. Adult	14211059	9250088	OBIWAN 210	QIAgen miRNeasy
3953Nf	Opossum	Heart	5. Adult	31121496	13033401	OBIWAN 170	QIAgen miRNeasy
3990Nf	Opossum	Heart	5. Adult	13350132	6511369	OBIWAN 170	QIAgen miRNeasy
3990Nf_rseq	Opossum	Heart	5. Adult	16654111	8095977	WINDU 92	QIAgen miRNeasy
3996Nf	Opossum	Heart	5. Adult	34703321	20257806	WINDU 67	QIAgen miRNeasy
5733Nf	Opossum	Liver	5. Adult	13294850	7896428	WINDU 63	QIAgen miRNeasy
5733Nf_rseq	Opossum	Liver	5. Adult	10334629	7742524	OBWIAN 170	QIAgen miRNeasy
5733Nf_rseq2	Opossum	Liver	5. Adult	15482914	11431920	WINDU 92	QIAgen miRNeasy
3955Nf	Opossum	Liver	5. Adult	26000778	21346938	OBIWAN 165	QIAgen miRNeasy
3991Nf	Opossum	Liver	5. Adult	27862175	22924132	OBIWAN 165	QIAgen miRNeasy
3997Nf	Opossum	Liver	5. Adult	4065321	2870757	OBIWAN 170	QIAgen miRNeasy
3997Nf_rseq	Opossum	Liver	5. Adult	11034811	7244883	OBIWAN 179	QIAgen miRNeasy
3997Nf_rseq2	Opossum	Liver	5. Adult	20506009	12218160	OBIWAN 283	QIAgen miRNeasy
5732Nf	Opossum	Liver	5. Adult	4747058	2653309	OBIWAN 170	QIAgen miRNeasy
5732Nf_rseq	Opossum	Liver	5. Adult	12921270	6732636	OBIWAN 179	QIAgen miRNeasy
5732Nf_rseq2	Opossum	Liver	5. Adult	41564508	21290348	OBIWAN 214	QIAgen miRNeasy
3961Nf	Opossum	Liver	5. Adult	31989274	19061667	WINDU 63	QIAgen miRNeasy
6146Nf	Opossum	Heart	5. Adult	15544268	9393472	OBIWAN 260	QIAgen miRNeasy
6146Nf_rseq	Opossum	Heart	5. Adult	34698680	24208044	OBIWAN 280	QIAgen miRNeasy

5139Nf	Opossum	Heart	1. Early embryo	9895845	6799996	WINDU 92	QIAgen miRNeasy
5139Nf_rseq	Opossum	Heart	1. Early embryo	18477672	12424284	OBIWAN 214	QIAgen miRNeasy
5039Nf	Opossum	Heart	1. Early embryo	5395258	2722301	OBIWAN 210	QIAgen miRNeasy
5039Nf_rseq	Opossum	Heart	1. Early embryo	25900889	13724838	OBIWAN 240	QIAgen miRNeasy
5041Nf	Opossum	Liver	1. Early embryo	9475807	5294196	WINDU 92	QIAgen miRNeasy
5041Nf_rseq	Opossum	Liver	1. Early embryo	19039503	10672269	OBIWAN 214	QIAgen miRNeasy
5040Nf	Opossum	Liver	1. Early embryo	11238655	7668371	OBIWAN 202	QIAgen miRNeasy
5040Nf_rseq	Opossum	Liver	1. Early embryo	13165335	9498793	OBIWAN 210	QIAgen miRNeasy
5293Nf	Opossum	Brain	1. Early embryo	60840347	21173950	WINDU 186	QIAgen RNeasy Mini modified
4809Nf	Opossum	Brain	2. Late embryo	10811477	5966717	OBIWAN 202	QIAgen RNeasy Mini modified
4809Nf_rseq	Opossum	Brain	2. Late embryo	86893104	49741604	OBIWAN 214	QIAgen RNeasy Mini modified
4813Nf	Opossum	Brain	2. Late embryo	16381861	9160312	OBIWAN 202	QIAgen RNeasy Mini modified
4813Nf_rseq	Opossum	Brain	2. Late embryo	29526516	18223104	OBIWAN 210	QIAgen RNeasy Mini modified
5042Nf	Opossum	Heart	2. Late embryo	9854194	7437943	OBIWAN 170	QIAgen miRNeasy
5042Nf_rseq	Opossum	Heart	2. Late embryo	10104872	6416935	WINDU 82	QIAgen miRNeasy
5045Nf	Opossum	Heart	2. Late embryo	16776399	7349863	WINDU 63	QIAgen miRNeasy
5045Nf_rseq	Opossum	Heart	2. Late embryo	11932800	6993576	WINDU 114	QIAgen miRNeasy
5044Nf	Opossum	Liver	2. Late embryo	8124312	5795918	OBIWAN 170	QIAgen miRNeasy
5044Nf_rseq	Opossum	Liver	2. Late embryo	8344740	5121159	WINDU 82	QIAgen miRNeasy
5044Nf_rseq2	Opossum	Liver	2. Late embryo	18207224	11136653	OBIWAN 283	QIAgen miRNeasy
5047Nf	Opossum	Liver	2. Late embryo	6587881	4773245	OBIWAN 170	QIAgen miRNeasy
5047Nf_rseq	Opossum	Liver	2. Late embryo	14474277	8817698	WINDU 82	QIAgen miRNeasy
4696Nf	Opossum	Brain	3. Neonate	14317670	6938322	WINDU 110	QIAgen RNeasy Micro modified
4696Nf_rseq	Opossum	Brain	3. Neonate	12815100	6301593	WINDU 114	QIAgen RNeasy Micro modified
4763Nf	Opossum	Heart	3. Neonate	5829454	4861029	OBIWAN 170	QIAgen miRNeasy
4763Nf_rseq	Opossum	Heart	3. Neonate	14548118	10027306	WINDU 82	QIAgen miRNeasy

4760Nf	Opossum	Heart	3. Neonate	53801900	24287202	WINDU 63	QIAgen miRNeasy
4765Nf	Opossum	Liver	3. Neonate	7223537	5578466	OBIWAN 170	QIAgen miRNeasy
4765Nf_rseq	Opossum	Liver	3. Neonate	20811277	13638902	WINDU 82	QIAgen miRNeasy
4762Nf	Opossum	Liver	3. Neonate	16810306	7183531	WINDU 63	QIAgen miRNeasy
4762Nf_rseq	Opossum	Liver	3. Neonate	9016015	6229420	OBIWAN 210	QIAgen miRNeasy
4762Nf_rseq2	Opossum	Liver	3. Neonate	9921235	6888994	OBIWAN 214	QIAgen miRNeasy
6297Nf	Opossum	Brain	3. Neonate	39723304	13526693	WINDU 186	QIAgen RNeasy Micro modified
4647Nf	Opossum	Brain	4. Juvenile	51889310	28669530	WINDU 67	QIAgen RNeasy Micro modified
4633Nf	Opossum	Brain	4. Juvenile	18493349	10512180	WINDU 63	QIAgen RNeasy Micro modified
4633Nf_rseq	Opossum	Brain	4. Juvenile	19888805	13007680	OBIWAN 283	QIAgen RNeasy Micro modified
4732Nf	Opossum	Heart	4. Juvenile	7481387	6439402	OBIWAN 170	QIAgen miRNeasy
4732Nf_rseq	Opossum	Heart	4. Juvenile	15160279	11284169	WINDU 82	QIAgen miRNeasy
4660Nf	Opossum	Liver	4. Juvenile	10945521	9392753	OBIWAN 170	QIAgen miRNeasy
4660Nf_rseq	Opossum	Liver	4. Juvenile	31617665	22836266	WINDU 82	QIAgen miRNeasy
6168Nf	Opossum	Heart	4. Juvenile	61308372	26130196	WINDU 186	QIAgen miRNeasy
6168Nf_rseq	Opossum	Heart	4. Juvenile	41825391	5746619	WINDU 214	QIAgen miRNeasy
4728Nf	Opossum	Heart	4. Juvenile	5783887	4297844	OBIWAN 170	QIAgen miRNeasy
4728Nf_rseq	Opossum	Heart	4. Juvenile	14624546	9244354	WINDU 82	QIAgen miRNeasy
6169Nf_rseq	Opossum	Liver	4. Juvenile	40560165	9597935	WINDU 214	QIAgen miRNeasy
4832Nf	Rabbit	Heart	4. Juvenile	5546393	4048080	OBIWAN 170	QIAgen miRNeasy
4832Nf_rseq	Rabbit	Heart	4. Juvenile	15215176	10337291	OBIWAN 179	QIAgen miRNeasy
6154Nf	Rabbit	Brain	4. Juvenile	51677246	24611095	WINDU 190	QIAgen RNeasy Micro modified
5968Nf	Rabbit	Brain	4. Juvenile	58922602	28029389	WINDU 190	QIAgen RNeasy Mini modified
3017Nf	Rabbit	Brain	5. Adult	6332426	4306206	OBIWAN 170	QIAgen miRNeasy
3017Nf_rseq	Rabbit	Brain	5. Adult	17442692	10982217	OBIWAN 179	QIAgen miRNeasy
3019Nf	Rabbit	Heart	5. Adult	5121240	3900139	OBIWAN 170	QIAgen miRNeasy

3019Nf_rseq	Rabbit	Heart	5. Adult	13783091	9728945	OBIWAN 179	QIAgen miRNeasy
3079Nf	Rabbit	Liver	5. Adult	27570091	18362510	OBIWAN 165	QIAgen RNeasy Micro modified
3025Nf	Rabbit	Brain	5. Adult	4656591	3692390	OBIWAN 170	QIAgen miRNeasy
3025Nf_rseq	Rabbit	Brain	5. Adult	12864838	9516160	OBIWAN 179	QIAgen miRNeasy
3027Nf	Rabbit	Heart	5. Adult	5787978	4317961	OBIWAN 170	QIAgen miRNeasy
3027Nf_rseq	Rabbit	Heart	5. Adult	15484681	10725910	OBIWAN 179	QIAgen miRNeasy
3029Nf	Rabbit	Liver	5. Adult	6719321	5044619	OBIWAN 165	QIAgen miRNeasy
3029Nf_rseq	Rabbit	Liver	5. Adult	11008428	8014716	OBIWAN 214	QIAgen miRNeasy
4828Nf	Rabbit	Liver	1. Early embryo	17752589	12241049	OBIWAN 170	QIAgen miRNeasy
4926Nf	Rabbit	Liver	1. Early embryo	19388383	6743448	WINDU 214	QIAgen RNeasy Micro modified
4924Nf	Rabbit	Brain	1. Early embryo	7375209	4198905	OBIWAN 165	QIAgen RNeasy Micro modified
4924Nf_rseq	Rabbit	Brain	1. Early embryo	15485460	7157886	WINDU 119	QIAgen RNeasy Micro modified
6151Nf	Rabbit	Brain	1. Early embryo	30701889	12061205	WINDU 193	QIAgen RNeasy Micro modified
4788Nf	Rabbit	Brain	2. Late embryo	14263350	11164407	OBWAN 170	QIAgen RNeasy Micro modified
4867Nf	Rabbit	Brain	2. Late embryo	15365662	10827706	OBIWAN 170	QIAgen RNeasy Micro modified
4867Nf_rseq	Rabbit	Brain	2. Late embryo	19392376	13077458	WINDU 168	QIAgen RNeasy Micro modified
4790Nf	Rabbit	Heart	2. Late embryo	15908078	9950197	OBIWAN 170	QIAgen RNeasy Micro modified
4790Nf_rseq	Rabbit	Heart	2. Late embryo	16137100	10119123	OBIWAN 190	QIAgen RNeasy Micro modified
4869Nf	Rabbit	Heart	2. Late embryo	12411150	8090262	OBIWAN 170	QIAgen RNeasy Micro modified
4869Nf_rseq	Rabbit	Heart	2. Late embryo	11691524	7703275	OBIWAN 190	QIAgen RNeasy Micro modified
4792Nf	Rabbit	Liver	2. Late embryo	14779992	8508105	OBIWAN 170	QIAgen RNeasy Micro modified
4792Nf_rseq	Rabbit	Liver	2. Late embryo	14156481	8256218	OBIWAN 190	QIAgen RNeasy Micro modified
4871Nf	Rabbit	Liver	2. Late embryo	16563022	10118727	OBIWAN 170	QIAgen RNeasy Micro modified
4871Nf_rseq	Rabbit	Liver	2. Late embryo	19287645	11373315	WINDU 168	QIAgen RNeasy Micro modified
3170Nf	Rabbit	Brain	3. Neonate	12685693	7951637	OBIWAN 170	QIAgen RNeasy Micro modified
3170Nf_rseq	Rabbit	Brain	3. Neonate	10960952	6991158	OBIWAN 190	QIAgen RNeasy Micro modified

3176Nf	Rabbit	Brain	3. Neonate	6663440	3865618	OBIWAN 192	QIAgen RNeasy Micro modified
3176Nf_rseq	Rabbit	Brain	3. Neonate	36147851	20040917	OBIWAN 214	QIAgen RNeasy Micro modified
3172Nf	Rabbit	Heart	3. Neonate	4465914	2268956	OBIWAN 170	QIAgen RNeasy Micro modified
3172Nf_rseq	Rabbit	Heart	3. Neonate	11117317	4566732	WINDU 82	QIAgen RNeasy Micro modified
3172Nf_rseq2	Rabbit	Heart	3. Neonate	51906102	25238335	OBIWAN 214	QIAgen RNeasy Micro modified
3178Nf	Rabbit	Heart	3. Neonate	53068497	3032490	WINDU 213	QIAgen RNeasy Micro modified
3174Nf	Rabbit	Liver	3. Neonate	16571203	8616940	OBIWAN 222	QIAgen RNeasy Micro modified
3174Nf_rseq	Rabbit	Liver	3. Neonate	25057775	11250389	WINDU 114	QIAgen RNeasy Micro modified
2527Nf	Rat	Liver	4. Juvenile	8833172	5741918	OBIWAN 170	QIAgen RNeasy Micro modified
2527Nf_rseq	Rat	Liver	4. Juvenile	16297171	10499727	OBIWAN 190	QIAgen RNeasy Micro modified
2759Nf	Rat	Liver	4. Juvenile	8600608	5073382	OBIWAN 202	QIAgen RNeasy Mini modified
2759Nf_rseq	Rat	Liver	4. Juvenile	38059521	19805013	OBIWAN 214	QIAgen RNeasy Mini modified
6173Nf	Rat	Brain	4. Juvenile	64424679	32461381	WINDU 186	QIAgen RNeasy Micro modified
6174Nf	Rat	Brain	4. Juvenile	61502013	30466096	WINDU 186	QIAgen RNeasy Micro modified
5806Nf	Rat	Heart	4. Juvenile	10311239	4594863	OBIWAN 202	QIAgen RNeasy Mini modified
5806Nf_rseq	Rat	Heart	4. Juvenile	15850809	6908783	WINDU 119	QIAgen RNeasy Mini modified
4092Nf	Rat	Brain	1. Early embryo	9492347	5427369	OBIWAN 190	QIAgen RNeasy Micro modified
4092Nf_rseq	Rat	Brain	1. Early embryo	34636393	18336028	OBIWAN 214	QIAgen RNeasy Micro modified
5966Nf	Rat	Liver	1. Early embryo	11101824	5023613	OBIWAN 222	QIAgen RNeasy Mini modified
5966Nf_rseq	Rat	Liver	1. Early embryo	13559492	5630599	WINDU119	QIAgen RNeasy Mini modified
5966Nf_rseq2	Rat	Liver	1. Early embryo	17033482	7437900	OBIWAN 283	QIAgen RNeasy Mini modified
4006Nf	Rat	Brain	1. Early embryo	22681084	10050930	OBIWAN 202	QIAgen RNeasy Micro modified
4006Nf_rseq	Rat	Brain	1. Early embryo	48182415	21452114	OBIWAN 283	QIAgen RNeasy Micro modified
3299Nf	Rat	Brain	2. Late embryo	4858093	3045265	OBIWAN 170	QIAgen miRNeasy
3299Nf_rseq	Rat	Brain	2. Late embryo	17251022	8932125	WINDU 82	QIAgen miRNeasy
3303Nf	Rat	Brain	2. Late embryo	3559836	2150574	OBIWAN 170	QIAgen miRNeasy

3303Nf_rseq	Rat	Brain	2. Late embryo	11436533	5756920	WINDU 82	QIAgen miRNeasy
5981Nf	Rat	Brain	2. Late embryo	7159223	1546033	WINDU 190	QIAgen RNeasy Mini modified
5980Nf	Rat	Brain	2. Late embryo	44921428	22472289	WINDU 213	QIAgen RNeasy Mini modified
2723Nf	Rat	Heart	2. Late embryo	14249774	5354783	OBIWAN 202	QIAgen RNeasy Micro modified
2723Nf_rseq	Rat	Heart	2. Late embryo	53023947	17372244	OBIWAN 214	QIAgen RNeasy Micro modified
2556Nf	Rat	Liver	2. Late embryo	65832046	13262020	WINDU 214	QIAgen RNeasy Micro modified
5728Nf	Rat	Brain	3. Neonate	15245638	8049028	OBIWAN 202	QIAgen RNeasy Micro modified
5728Nf_rseq	Rat	Brain	3. Neonate	37022005	16774403	WINDU 114	QIAgen RNeasy Micro modified
6176Nf	Rat	Brain	3. Neonate	79380770	29634762	WINDU 213	QIAgen RNeasy Micro modified
6177Nf	Rat	Heart	3. Neonate	78608711	29062942	WINDU 213	QIAgen RNeasy Micro modified
5967Nf	Rat	Heart	3. Neonate	17276945	8406273	OBIWAN 222	QIAgen RNeasy Mini modified
5967Nf_rseq	Rat	Heart	3. Neonate	24798522	10372942	WINDU 114	QIAgen RNeasy Mini modified

Table S2. Sample information of miRNA three organs

Category	Term	Count	Percent	Fold	P.value	Bonferroni	FDR	miRNA	type
Cluster	hsa-mir-512-1 cluster	14	0.28	7.2216	8.63E-11	9.06E-09	5.34E-08	hsa-mir-1323,hsa-mir-498,hsa-mir-520a,hsa-mir-519b,hsa-mir-520b,hsa-mir-520c,hsa-mir-524,hsa-mir-519d,hsa-mir-520g,hsa-mir-518e,hsa-mir-518a-1,hsa-mir-518d,hsa-mir-518a-2,hsa-mir-520h	C1
Cluster	hsa-mir-941-1 cluster	5	1	56.272	6.20E-10	5.21E-08	3.84E-07	hsa-mir-941-1,hsa-mir-941-2,hsa-mir-941-3,hsa-mir-941-4,hsa-mir-941-5	C4
Family	mir-941 family	5	1	56.272	6.20E-10	5.21E-08	3.84E-07	hsa-mir-941-1,hsa-mir-941-2,hsa-mir-941-3,hsa-mir-941-4,hsa-mir-941-5	C4
Family	mir-515 family	12	0.28571429	7.3690	2.87E-09	3.01E-07	8.87E-07	hsa-mir-518a-1,hsa-mir-518a-2,hsa-mir-518d,hsa-mir-518e,hsa-mir-519b,hsa-mir-519d,hsa-mir-520a,hsa-mir-520b,hsa-mir-520c,hsa-mir-520g,hsa-mir-520h,hsa-mir-524	C1
Cluster	hsa-mir-379 cluster	9	0.23076923	8.9278	5.09E-08	4.58E-06	3.15E-05	hsa-mir-411,hsa-mir-376c,hsa-mir-654,hsa-mir-376b,hsa-mir-381,hsa-mir-382,hsa-mir-485,hsa-mir-377,hsa-mir-409	F>M
Function	Insulin Resistance	4	0.28571429	25.265	5.56E-06	0.000828	0.00344	hsa-mir-223,hsa-mir-181b-1,hsa-mir-378c,hsa-mir-181b-2	C5
Disease	Mycosis Fungoides	3	0.6	53.057	8.81E-06	0.00131	0.00273	hsa-mir-181b-1,hsa-mir-223,hsa-mir-181b-2	C5
Function	Pluripotent Stem Cells Reprogramming	3	0.27272727	24.116	0.000141	0.021	0.0291	hsa-mir-29b-1,hsa-mir-181b-1,hsa-mir-181b-2	C5
Cluster	hsa-mir-34b cluster	2	1	56.272	0.000288	0.0242	0.0593	hsa-mir-34b,hsa-mir-34c	C4
Cluster	hsa-mir-3130-1 cluster	2	1	56.272	0.000288	0.0242	0.0593	hsa-mir-3130-1,hsa-mir-3130-2	C4
Family	mir-3130 family	2	1	56.272	0.000288	0.0242	0.0593	hsa-mir-3130-1,hsa-mir-3130-2	C4
Cluster	hsa-mir-3158-1 cluster	2	1	47.615	0.000408	0.044	0.2524	hsa-mir-3158-1,hsa-mir-3158-2	F<M
Family	mir-194 family	2	1	47.615	0.000408	0.044	0.2524	hsa-mir-194-1,hsa-mir-194-2	F<M
Family	mir-3158 family	2	1	47.615	0.000408	0.044	0.2524	hsa-mir-3158-1,hsa-mir-3158-2	F<M

Table S3. miRNA associated disease and function from TAM2.0

Number	Species	miRBase ID	Age group	Divergence time to Human (Mya)
1	Human	hsa	12	
2	Chimpanzee	ptr	11	6.4
3	Orangutan	ppy	10	15.2
4	Gibbon	nle	9	19.8
5	Rhesus	mml	8	28.81
6	Marmoset	caj	7	42.9
7	Mouse	mmu	6	89
8	Rat	rno		
9	Guinea Pig	cpo		
10	Rabbit	ocu		
11	Cow	bta	5	94
12	Horse	eca		
13	Dog	cfa		
15	Opossum	mdo	4	160
16	Platypus	oan	3	180
17	Lizard	aca	2	318
18	Chicken	gga		
19	Zebra finch	tgu		
20	Xenopus tropicalis	xtr	1	351.7
21	Tetraodon	tni	0	433
22	Fugu	fru		
23	Zebrafish	dre		

Table S4. Table of dating of human miRNA evolutionary age

number	species	Age group
1	Rabbit	9
2	Mouse	8
3	Rat	8
4	Human	7
5	Rhesus	
6	Marmoset	
7	Cow	6
8	Dog	
9	Armadillo	5
10	Opossum	4
11	Tasmanian devil	
12	Platypus	3
13	Chicken	2
14	Painted turtle	
15	Xenopus tropicalis	1
16	Fugu	0
17	Zebrafish	

number	species	Age group
1	Opossum	5
2	Human	4
3	Mouse	
4	Dog	
5	Platypus	3
6	Chicken	2
7	Xenopus tropicalis	1
8	Zebrafish	0

number	species	Age group
1	Chicken	4
2	Zebrafinch	3
3	Human	2
4	Opossum	
5	Platypus	
6	Xenopus tropicalis	1
7	Zebrafish	0

number	species	Age group
1	Rhesus	9
2	Chimpanzee	8
3	Orangutan	8
4	Gibbon	8
5	Human	8
6	Marmoset	7
7	Mouse	6
8	Rat	
9	Guinea Pig	
10	Rabbit	
11	Cow	5
12	Horse	
13	Dog	
15	Opossum	4
16	Platypus	3
17	Lizard	2
18	Chicken	
19	Zebra finch	
20	Xenopus tropicalis	1
21	Tetraodon	0
22	Fugu	
23	Zebrafish	

number	species	Age group
1	Marmoset	8
2	Chimpanzee	7
3	Orangutan	7
4	Gibbon	7
5	Human	7
6	Rhesus	7
7	Mouse	6
8	Rat	
9	Guinea Pig	
10	Rabbit	
11	Cow	5
12	Horse	
13	Dog	
15	Opossum	4
16	Platypus	3
17	Lizard	2
18	Chicken	
19	Zebra finch	
20	Xenopus tropicalis	1
21	Tetraodon	0
22	Fugu	
23	Zebrafish	

number	species	Age group
1	Mouse	11
2	Rat	10
3	Guinea Pig	9
4	Rabbit	8
5	Human	7
6	Rhesus	
7	Marmoset	
8	Cow	6
9	Dog	
10	Armadillo	5
11	Opossum	4
12	Tasmanian devil	
13	Platypus	3
14	Chicken	2
15	Painted turtle	
16	Xenopus tropicalis	1
17	Fugu	0
18	Zebrafish	

number	species	Age group
1	Rat	11
2	Mouse	10
3	Guinea Pig	9
4	Rabbit	8
5	Human	7
6	Rhesus	
7	Marmoset	
8	Cow	6
9	Dog	
10	Armadillo	5
11	Opossum	4
12	Tasmanian devil	
13	Platypus	3
14	Chicken	2
15	Painted turtle	
16	Xenopus tropicalis	1
17	Fugu	0
18	Zebrafish	

Table S5. Table of dating of miRNA evolutionary ages