

# Didelphis albiventris: an overview of an unprecedented transcriptome sequencing of the white-eared opossum

**Íria Gabriela Dias dos Santos**

Universidade Federal de Minas Gerais Instituto de Ciencias Biologicas

**Tiago Antônio de Oliveira Mendes**

Universidade Federal de Vicosa

**Gerluza Aparecida Borges Silva**

Universidade Federal de Minas Gerais Instituto de Ciencias Biologicas

**Amanda Maria Sena Reis**

Universidade Federal de Minas Gerais Instituto de Ciencias Biologicas

**Cláudia Barros Monteiro Vitorello**

Universidade de Sao Paulo Escola Superior de Agricultura Luiz de Queiroz

**Patrícia Dayane Carvalho Schaker**

Universidade de Sao Paulo Escola Superior de Agricultura Luiz de Queiroz

**Roberto Hirochi Herai**

Pontificia Universidade Catolica do Parana

**André Brait Carneiro Fabotti**

Universidade Federal de Minas Gerais

**Luiz Lehmann Coutinho**

Universidade de Sao Paulo Escola Superior de Agricultura Luiz de Queiroz

**Erika C Jorge** (✉ [erika.cris.jorge@gmail.com](mailto:erika.cris.jorge@gmail.com))

Universidade Federal de Minas Gerais Instituto de Ciencias Biologicas <https://orcid.org/0000-0001-5390-5266>

---

## Research article

**Keywords:** Didelphis albiventris; white-eared opossum; RNA-seq; transcriptome; newborn; postnatal

**Posted Date:** November 1st, 2019

**DOI:** <https://doi.org/10.21203/rs.2.12871/v3>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

**Version of Record:** A version of this preprint was published on November 15th, 2019. See the published version at <https://doi.org/10.1186/s12864-019-6240-x>.

## Abstract

**Background**The white-eared opossum (*Didelphis albiventris*) is widely distributed throughout Brazil and South America. It has been used as an animal model for studying different scientific questions ranging from the restoration of degraded green areas to medical aspects of Chagas disease, leishmaniasis and resistance against snake venom. As a marsupial, *D. albiventris* can also contribute to the understanding of the molecular mechanisms that govern the different stages of organogenesis. Opossum joeys are born after only 13 days, and the final stages of organogenesis occur when the neonates are inside the pouch, depending on lactation. As neither the genome of this opossum species nor its transcriptome has been completely sequenced, the use of *D. albiventris* as an animal model is limited. In this work, we sequenced the *D. albiventris* transcriptome by RNA-seq to obtain the first catalogue of differentially expressed (DE) genes and gene ontology (GO) annotations during the neonatal stages of marsupial development. **Results**The *D. albiventris* transcriptome was obtained from whole neonates harvested at birth (P0), at five days of age (P5) and at ten days of age (P10). The de novo assembly of these transcripts generated 85,338 transcripts. Approximately 30% of these transcripts could be mapped against the amino acid sequences of *M. domestica*, the evolutionarily closest relative of *D. albiventris* to be sequenced thus far. Among the expressed transcripts, 2,077 were found to be DE between P0 and P5, 13,780 between P0 and P10, and 1,453 between P5 and P10. The enriched GO terms were mainly related to the immune system, blood tissue development and differentiation, vision, hearing, digestion, the CNS and limb development. **Conclusions**The elucidation of opossum transcriptomes provides an out-group for better understanding the distinct characteristics associated with the evolution of mammalian species. This study provides the first transcriptome sequences and catalogue of genes for a marsupial species at different neonatal stages, allowing the study of the mechanisms involved in organogenesis.

## Background

Marsupials are one of the three large modern groups of mammals and are considered the closest external group to eutherian mammals (placentals). They comprise a group of 350 extant species found in the Americas, mostly in South America and Australasia [1].

The most intriguing characteristic of marsupials is related to the way they develop: after a very short period of intrauterine development, they are born with only a few early developed structures (including the external nostrils, the mouth and the forelimbs), while all other structures are still in early stages of development. In general, lactation plays a major role in the development of marsupial joeys [2], whereas the placenta is short-lived, and most marsupial placentas are either non-invasive, as in the tammar wallaby [3, 4], or invasive in only the last few days of pregnancy, as in the South American grey short-tailed opossum [3, 5]. Immediately after birth, neonates climb into the marsupium pouch, where each attaches to one of the mother's teats, and a complex system of lactation is established while the main organogenesis stages of development take place [1-12]. In the pouch, marsupial neonates are in an altricial developmental stage equivalent to E10-12 of mouse embryos or 10 weeks of human development [6].

Marsupials can be considered an excellent model for understanding the mechanisms that govern cellular differentiation during organogenesis due their close phylogenetic relationship to eutherians and the availability of neonates in the pouch (overcoming the difficulty of access when studying in utero development). For instance, different aspects of neocortex development, expansion and evolution have recently been revealed using embryos and neonates of the grey short-tailed opossum (*Monodelphis domestica*) and the tammar wallaby (*Macropus eugenii*) as models [13, 14]. Studies using grey short-tailed opossum neonates revealed new insights into bronchioalveolar [15] and chondrocranial development in mammals [16]. Morphological and morphometric study of the skin of marsupial neonates suggested the participation of this organ in gaseous exchange on the basis of the investigation of species such as the eastern quoll (*Dasyurus viverrinus*), grey short-tailed opossum (*Monodelphis domestica*), southern brown bandicoot (*Isodon obesulus*), long-nosed bandicoot (*Perameles nasuta*), brush-tail possum (*Trichosurus vulpecula*), koala (*Phascolarctos cinereus*), long-nosed potoroo (*Potorous tridactylus*), brush-tailed rock wallaby (*Petrogale penicillata*), red-necked pademelon (*Thylogale thetis*) and black-striped wallaby (*Macropus dorsalis*) [17]. Marsupials, especially didelphids, have been employed as models for the study of mammalian evolution [18-20].

Such important biological characteristics can be discovered in marsupials with strong support from genomic and/or transcriptomic sequences available in public databanks. The present work adds to a growing number of new marsupial genome and transcriptome datasets that have been released in the last few years, which have mainly been developed to the study of the marsupial placenta, lactation and immune system, in species such as the grey short-tailed opossum (*Monodelphis domestica*) [21, 22], the Tasmanian devil (*Sarcophilus harrisii*) [23, 24], the tammar wallaby (*Macropus eugenii*) [12, 25, 26], the koala (*Phascolarctos cinereus*) [27-29], the long-nosed bandicoot (*Perameles nasuta*) [30, 31], the fat-tailed dunnart (*Sminthopsis crassicaudata*) [32], the monito del monte (*Dromiciops gliroides*) [33], and the Tasmanian tiger (*Thylacinus cynocephalus*) [34].

Based on these previous findings and genomic data availability limitations, the goal of this work was to sequence the transcriptome of the white-eared opossum (*Didelphis albiventris*) during the first days of its postnatal development. This is the first transcriptome obtained from a South American marsupial species during distinct stages of neonatal development. *D. albiventris* undergoes intra-uterine development for 13 days, in which the last ~3 days are dependent on a non-invasive placenta [35]; after birth, the neonates remain with the mother for ~100 days depending on lactation [36, 37]. For this work, we used neonates harvested at birth and at 5 and 10 days old.

These marsupials are widely distributed in South America (including Cerrado, Caatinga and Pantanal areas) [38-42], and they have been used as model organisms to understand human infections such as Chagas disease [43, 44] and leishmaniasis [45]. Cáceres [46] suggested that *D. albiventris* is an important agent as a seed spreader. The seeds of many plant species (including pioneers) may aid in the restoration of degraded environments after they are eaten by opossums, as they remain viable after passing through the intestine [46]. *D. albiventris* and other didelphids are resistant to the

venom of snakes such as *Bothrops* spp., *Crotalus durissus* and *Lachesis muta* [47-49]. They are also resistant to intoxication by millipedes [50], which are toxic to many vertebrates. Our group has been trying to establish *D. albiventris* as a model for understanding odontogenesis stages during development. This opossum exhibits complete heterodont dentition that is closer to that of humans than is that of rodents, the typical model for these studies [8, 51-55]. We can characterize the morphological stages of early tooth development (dental lamina, bud, cap, and bell stages) in this species [8, 51-55]. However, the use of sequences from *M. domestica* (the closest evolutionary relative) to develop molecular approaches to study *D. albiventris* has not been successful.

Analysis of the *D. albiventris* transcriptome has generated genetic information for this promising species as a new model organism for studies on the regulatory molecular mechanisms of organogenesis, providing a better understanding of marsupial species contributing to their preservation and supporting evolutionary developmental biology research [56].

## Results

### RNA-seq and transcriptome assembly of *D. albiventris*

This work generated transcriptomes (RNA-seq) of the white-eared opossum during the first days of its postnatal development. The analysis of all RNA-seq samples was performed; on average, over 47.5 million raw reads per library/biological replicate were obtained for each postnatal stage (**Table 1**) with a 100 bp paired-end fragment length. Using the Trinity tool, 85,338 transcripts were assembled, considering a minimum fragment overlap of 35 bp and contigs with a length of at least 300 bp.

**Table 1. RNA-seq data.**

Samples type	Total number of reads	Sample description
P0	50,523,432	Biological replicate 1 for P0
	51,490,816	Biological replicate 2 for P0
	47,805,980	Biological replicate 3 for P0
P5	50,433,754	Biological replicate 1 for P5
	47,708,526	Biological replicate 2 for P5
	47,014,722	Biological replicate 3 for P5
P10	46,287,998	Biological replicate 1 for P10
	52,374,156	Biological replicate 2 for P10
	33,968,928	Biological replicate 3 for P10

### Sequence composition among biological replicates

Principal component analysis (PCA) showed that the biological replicates for each developmental stage could be grouped together, forming 3 distinct clusters corresponding to the different sets of biological conditions (**Figure 1A**).

The colourmap matrix constructed to show the sample-to-sample distances confirmed that the three biological replicates for each developmental stage were very similar to each other, except for one P0 sample that showed similarity to the P5 samples (see "P0 B" in **Figure 1B**). The expression levels of each transcript throughout the three developmental stages were represented as a colourmap matrix, allowing the observation that the expression of several transcripts did not change considerably across samples P0, P5 and P10. It was also possible to observe that some transcripts varied between the samples, showing either increasing or decreasing expression across the sample types (**Figure 1C**).

### Differentially expressed genes between opossum postnatal stages

We identified 14,300 differentially expressed (DE) transcripts between the P0, P5 and P10 biological stages. Among the total DE transcripts, 2,077 transcripts were DE between stages P0 and P5, 13,780 between stages P0 and P10, and 1,453 between stages P5 and P10 (**Figure 2A**). Additionally, 213 of these transcripts were DE exclusively between stages P0 and P5 and 200 between P5 and P10 (**Figure 2A**). The most striking difference was observed between P0 and P10, as 11,008 DE transcripts were identified between these stages (**Figure 2A**). Additionally, 131 transcripts were found to be DE between all analysed developmental stages (**Figure 2A**).

Among the 14,300 DE transcripts, 579 were found to be downregulated (**Figure 2B**) and 1,492 to be upregulated (**Figure 2C**) at P5 in relation to P0; 6,438 were downregulated (**Figure 2B**), and 7,207 upregulated (**Figure 2C**) at P10 in relation to P0; and 846 were downregulated (**Figure 2B**), and 607 were upregulated (**Figure 2C**) at P10 in relation to P5.

We also identified a few transcripts that were exclusively expressed in each of the analysed stages: 109 transcripts were exclusively expressed at P0, 149 at P5 and 237 at P10 (**Additional File 3**).

### Amino acid sequence similarities between *D. albiventris* and *M. domestica*

The BLASTX analysis of the *D. albiventris* translated nucleotide sequences against the amino acid sequences from *M. domestica* showed that the sequences that exhibited 30% coverage were more likely to present higher identity values and less likely to present lower identity values (**Additional File 1**).

### ***D. albiventris* nucleotide sequence similarities to other marsupial transcriptomes**

BLASTN analysis of the *D. albiventris* assembled nucleotide sequences against available marsupial transcriptomes showed higher identity values between the *D. albiventris* and *M. domestica* (90%) or *S. harrisii* (90%) sequences compared to those of *P. cinereus* (20%) and *M. eugenii* (20%) (**Additional File 2**).

### **Transcriptome assembly validation**

#### *Validation by RT-qPCR*

We validated four DE transcripts from our RNA-seq analysis by RT-qPCR. The selected transcripts were associated with genes involved in enriched biological processes related to haematopoietic cells (HMGB3) [71, 72], platelet agglutination (APOH) [73, 74], erythrocytes (HBZ) [75], and the eye lens (CRYGB) [76, 77].

Although there were a limited number of genes to validate these data, we observed total concordance in the direction of expression but not in the magnitude of the expression change between the RNA-seq and RT-qPCR results (**Figure 3**). In the RNA-seq data (**Additional File 4**), HMGB3 was found to be downregulated at P5 compared to its expression at P0, while it was upregulated at P10 compared to P5. CRYGB was upregulated at P5 relative to P0 and at P10 relative to P0; APOH was downregulated at P10 relative to P5; and HBZ was downregulated at P10 relative to P0.

#### *Validation by in situ hybridization (ISH)*

A ~400 bp fragment of opossum titin mRNA could also be amplified and cloned. Sections of P0, P5 and P10 tissues were selected to show the presence of titin mRNA in skeletal muscles by using the ISH method (**Additional File 5**). The RNA-seq data showed higher expression levels of titin transcripts at later postnatal stages.

Our results revealed the expected muscle-specific staining for titin in all postnatal stages. At P0 (**Additional File 5B and C**) and P10 (**Additional File 5I**), the tongue muscle fibres were strongly stained by ISH. At P5, the ventral skeletal muscle of the body wall (**Additional File 5D**) and longitudinal and transversal sections of limb skeletal muscle (**Additional File 5F**) were strongly stained by ISH. The diaphragm (**Additional File 5A, E and H**) showed moderate staining of the skeletal muscle. At P10, the heart showed moderate staining (**Additional File 5G**), and the extrinsic muscles of the eye showed strong staining (**Additional File 5J**).

### **GO pathway enrichment analysis**

The enriched GO terms for the DE transcripts were generally related to the cell cycle, metabolism and tissue survival as well as the development of systems such as the (i) immune system, (ii) blood tissue, (iii) hearing, (iv) vision, (v) digestion, (vi) the excretory system, (vii) the nervous system, and (viii) the limbs (**Tables 2 and 3**).

The most significant GO terms are shown in **Figures 4 to 8** and **Additional File 6** for all 85,338 transcripts expressed among the three postnatal stages (**Figure 4**); for the 14,300 DE transcripts expressed among the three postnatal stages (**Figure 5**); for the 2,077 DE transcripts between P0 and P5 (**Figure 6A-B**); for the 579 downregulated DE transcripts (**Figure 7A-B**) and 1,492 upregulated DE transcripts (**Figure 7C-D**) at P5 relative to P0; for the 1,453 DE transcripts between P5 and P10 (**Figure 6C-D**); and for the 846 downregulated DE transcripts (**Figure 8A-B**) and 607 upregulated DE transcripts (**Figure 8C-D**) at P10 relative to P5.

## **Discussion**

In this work, we sequenced the transcriptomes of whole neonates of the white-eared opossum (*Didelphis albiventris*) in three postnatal developmental stages to reveal the mechanisms of development in this South American marsupial species. To the best of our knowledge, this is the first RNA-seq transcriptome assembly available for this species. Although some studies have investigated the transcriptomes of certain tissues of marsupial neonates [15, 78, 79], this is the first differential expression analysis of whole marsupial individuals in distinct neonatal development stages. The generated database, composed of 85,338 transcripts, could be used to reveal the mechanisms of late tissue specification and differentiation in a marsupial model, contributing to the understanding of molecular phylogenetic aspects of mammalian evolution.

### **Transcriptome resources**

Our transcriptome database was obtained from three *Didelphis albiventris* siblings per developmental stage (P0, P5 and P10), which is considered a sufficient number of biological replicates for a reliable quantitative inferential analysis [80]. On average, 47.5 million reads were obtained per library/biological replicate of each postnatal stage. Although the general homogeneity of the sequenced biological triplicates is supported by the PCA clustering results, Euclidian distance-based analysis, represented by a colourmap matrix, suggested that a single P0 replicate slightly deviated from the

other samples at the same stage. Discordance in gene expression patterns between siblings may occur due to epigenetic factors [81, 82] such as marsupium environment conditions and lactation.

The coverage of just 30% between the amino acid sequences of *D. albiventris* and its closest evolutionary relative sequenced to date, *M. domestica*, revealed the importance of the opossum transcriptome for understanding the biology of this species and the evolution of the marsupial group. Many reads from unannotated transcripts may be lost, and various multi-mapping reads are observed because of reads falling within exons that are shared by different transcript isoforms of the same gene [80].

Considering the similarity of mapped nucleotide sequences (BLASTN), the assembled *D. albiventris* transcripts showed more similarity at the nucleotide level with sequences of *M. domestica* (grey short-tailed opossum) and, surprisingly, *S. harrisi* (Tasmanian devil) than with *P. cinereus* (koala) and *M. eugenii* (tammar wallaby). Taken together, these results corroborate the importance of generating molecular information for a new South American marsupial species, as its phylogenetically closest sequenced relative shows low levels of sequence similarity. New marsupial phylogenetic relationships can now be revealed from these sequences.

RNA-seq validation is an important step for establishing and verifying biological interpretations that may arise from database analysis [83]. The need for this type of validation has been discussed since the correlation between RNA-seq and RT-qPCR data has been consistently confirmed. Indeed, the present study supported these observations: the quantification of upregulated and downregulated transcripts by RT-qPCR revealed a correlation with the levels and per-stage differences of transcript expression determined by RNA-seq. However, the RT-qPCR assays were performed on the same biological samples used to perform RNA-seq, which supports the validity of bioinformatics DE analysis for the discovery of up- and downregulated molecules between different stages. However, the RT-qPCR assays still present the limitation of not having the power to represent the biological expression of a gene in a population. Other independent biological samples need to be tested to confirm the detected gene expression levels [83].

Transcriptome assembly usually tends to be more complex than genome assembly and to be more challenging in eukaryote organisms than in prokaryotes, especially when it requires *de novo* assembly [84, 85]. Thus, we decided to validate the assembled transcriptome by ISH as well. Since the young opossum climbs and suckles from the day of its birth [6], we evaluated skeletal muscle development at the studied stages. The strong staining for titin hybridization reflected the presence of mature skeletal muscle fibres in the tongue at P0 and P10, in the body wall at P0, in the limbs at P5, and in the extrinsic muscles of the eye at P10. Additionally, weaker staining for titin hybridization showed a lower degree of maturation in the diaphragm at the three stages and in the heart at P10. The moderate hybridization of titin at P10 in the heart is consistent with reports that most ventricular cardiac muscle cells are fully differentiated by the end of the fifteenth postnatal week [86]. Thus, our data confirmed the specificity of the synthesized probe. The observation of gene expression in a semi-embryonic state model will allow further studies on molecular pathways during mammalian development. Moreover, the muscles of marsupials and other mammals have been recently studied for their role in non-shivering thermogenesis (NST), which is important for endothermic regulation [87]. The elucidation of the *D. albiventris* transcriptome will allow these molecular studies to be extensively performed on opossums.

### **The main tissues and organs in the development of the white-eared opossum**

The opossum transcriptome obtained during different developmental stages allowed us to establish the first gene expression patterns for different tissues from this marsupial model. As our first dive into this dataset, we decided to only analyse differentially expressed transcripts between each neonatal stage, focusing on the development of tissues that are known to be of particular importance for marsupial functions. Consequently, we could obtain some idea of the most important genic events controlling each stage.

Gene ontology terms associated with the development of the immune system were found to be enriched in our analysis of the *D. albiventris* DE transcripts. In marsupials, the thymus, bone marrow and lymph nodes are still immature at birth [88]. Until the adaptive immune system develops, immune protection of the marsupial newborn is achieved by prenatal transfer of immunoglobulins, antimicrobial compounds within the pouch and immune compounds within the milk [89]. We found a number of “immune”-associated GO terms during the *D. albiventris* developmental stages. The enriched terms grouping the P0 and P5 DE transcripts included “chemokine-mediated signalling pathway”, “intestinal immune network for IgA production”, “leukocyte migration”, “lymphocyte co-stimulation”, “negative regulation of tumour necrosis factor production”, “neutrophil chemotaxis”, “positive regulation of inflammatory response” and “thymic T cell selection”. The enriched terms grouping the P5 and P10 transcripts included “activated T cell proliferation”, “acute inflammatory response”, “cytokine biosynthetic process”, “cytokine secretion”, “defence response to bacterium”, “innate immune response in mucosa”, “neutrophil migration”, “positive regulation of chemokine production” and “positive regulation of interleukin-6 and -8 production”. These data show that *D. albiventris* starts to establish its adaptive immune system at birth, and the main pathways are activated in P5 and P10. Additionally, several genes linked to coagulation, fibrinolysis and complement system components were found to be enriched in our database, similar to what is observed during *M. domestica* early postnatal stages [15]. These components are all interconnected: the complement protein network, an essential component of the innate and humoral immune system [90], participates in several non-inflammatory processes, such as coagulation [91], along with components of the coagulation and fibrinolysis systems [90]. Genes related to erythrocyte differentiation were also observed, such as the erythroid transcription factor KLF1 (also known as EKLF), which is expressed by erythro-myeloid progenitors (EMPs) [92], erythroid cells [93], and megakaryocyte-erythrocyte progenitors (MEPs) [93].

Regarding the use of the primary jaw joint, the marsupial neonate approximates the embryonic condition of mammalian ancestors [94]. Therefore, marsupial neonates have recently come to be considered an excellent resource for studies on the transition from primary to novel jaw articulation

through the analysis of inner ear development [95, 96]. In the present study, we found enriched upregulated genes at P0 relative to P5 associated with cochlea development including HES1 (Hairy and Enhancer of Split 1), PTK7 (Tyrosine-Protein Kinase-Like 7) and SOBP (Sine Oculis Binding Protein Homologue), which were categorized under the GO terms “cellular response to retinoic acid” and “cochlea development”. Retinoic acid is indispensable for inner ear development and plays a role in the molecular cascade of hair cell regeneration [97-100]. The differentiation of the inner ear hair cells, located in the cochlea, is negatively regulated by HES1 [101], and the stereocilia orientation in these cells is guided by PTK7 [102, 103]. The SOBP gene is expressed mainly in the cochlear duct, regulating the cellular fate and standardization of the organ of Corti [104, 105]. These data suggest that our gene expression profile of *D. albiventris* at the P0 neonatal stage might include candidate genes that could reveal the molecular mechanisms that determine the development of the cochlea and other inner ear structures.

Eutherian neonates [106] and chicken embryos [107] have been used in studies on eye lens development due the great importance of studying eye-related pathologies that affect humans and can be identified soon after birth, such as congenital cataracts and retinoblastoma [108]. Regarding the importance of understanding eye lens development, our data showed enriched genes that were upregulated at P5 relative to P0 such as CRYBA1, CRYBA4, CRYBB1, and CRYBB2, which were grouped under the following GO terms: “visual perception” and “sensory perception of light stimulus”. The beta-crystalline gene cluster, composed of CRYBA1, CRYBA4, CRYBB1 and CRYBB2, is expressed in the vertebrate eye lens from embryonic to adult life [109]. Additionally, other genes such as AIPL1 (expressed during rod and cone development [110]) and GJC1 (expressed in retina bipolar cells and retinal ganglion cells during development [111]), which were found to be related to enriched GO terms in our results, could be targets for understanding the marsupial visual system.

The gastrointestinal system of postnatal marsupials has been explored, and the results contribute to telling the story of the evolution of this system in vertebrates [112-115]. Studies on marsupial pouch joeys have shown that every phase of milk contains specific factors that trigger phenotypic changes in the neonate stomach [116] and that the milk regulates the microbiota of the neonate intestine [117]. Furthermore, studies indicate that the milk plays a role in accelerating the development and maturation of the stomach [116]. However, the molecular mechanisms involved in these changes during the development of the marsupial stomach are not yet known [116]. Hence, the enriched genes linked to the digestive system identified in this study, such as SLC9A4 and MOGAT3, could help to understand the development of the mammalian gastrointestinal system. These genes are mainly related to stomach, intestine and pancreas functions. SLC9A4 (also known as NHE4) is a Na<sup>+</sup>/H<sup>+</sup> exchanger that is highly expressed in the basolateral membrane of gastric parietal cells, playing a role in the maintenance of normal levels of gastric acid and gastric epithelial cell differentiation [118]. MOGAT3 (or MGAT3) is expressed in the pancreas and intestines [119, 120] and may play a role in triglyceride absorption [119].

The development of the mammal kidneys is another topic that is still under investigation [121]. Several recent studies have aimed to trace the early stages of renal development to understand the mechanisms involved and implement strategies for the *in vitro* engineering of kidney tissue replacement (in foetuses and neonates, for example) to develop more precocial, more efficient, and lower-cost therapeutic approaches [122-124]. Among the enriched GO terms identified in the three neonatal stages, the gene IRX1 is correlated with the early stages of kidney development. This gene is required for the formation of intermediate segments of the pronephros [125] and is necessary to maintain the identity of the pronephric territory and define its size [126]. IRX1 is expressed in the proximal tubular segment and the intermediate tubule segments of the pronephros [125, 126] and exhibits an important function in the differentiation of a proximo-medial S-shaped body (SSB) subdomain [127]. Thus, IRX1 could provide valuable information about mammalian kidney development and evolution.

Finally, the rate of brain development in marsupials varies more than in eutherians, and this characteristic might be associated with their longer lactation period [128]. Because the forebrain of marsupials primarily develops postnatally, it is possible to study the brain development of the young in the pouch in stages that are equivalent to human mid-embryogenesis [129]. A study in which in-pouch electroporation was performed using newborns of the Australian marsupial fat-tailed dunnart (*Sminthopsis crassicaudata*) demonstrated that marsupials can be an excellent *in vivo* model for studying forebrain development and evolution [129]. In the present study, several genes were found to be related to enriched GO pathways involved in nervous tissue development and differentiation, including MSX1 and GPR37L1. While MSX1 is expressed in the spinal cord [130] and is related to the development of the midline structure of the forebrain [131], GPR37L1 is practically restricted to the brain, where it is expressed in neurons and glia [132, 133].

While cranial and adult structures have been receiving attention, there is a gap in the knowledge of the evolution of mammalian post-cranial structures, such as the limbs [134]. A recent morphological study suggests that the marsupial forelimb may place a constraint on evolution due the functional requirements of the crawl to the teat, which would limit morphological variation before and at the time of birth [134]. This peculiar mode of reproduction in marsupials could be responsible for their lower diversity in limb morphology [134]. The authors also emphasize that “adult integration studies may not be indicative of developmental integration” [134], highlighting the importance of studies on embryos and neonates. The enriched genes related to embryonic hindlimb morphogenesis found among opossum developmental stages include FGF4 and MSX1. FGF4 is expressed in the thoracic and pelvic limbs [135] and in the apical ectodermal ridge of the limb bud [136], playing a role in the distal growth of the developing limb bud [135, 137]. MSX1 encodes transcription factors that are crucial for limb development [138].

## Conclusion

In this study, the transcriptomes of whole young animals were sequenced, assembled and compared at three different ages, generating the first comprehensive catalogue of *D. albiventris* transcripts, enabling detailed analysis of the unique attributes of this native Brazilian marsupial. This

transcriptomic dataset is a useful resource for molecular genetic studies of the opossum and could be used as an animal model to investigate distinct organ formation and development. As an important model for mammals, the generated dataset provides novel and important dynamic modifications at the molecular level for better understanding of the evolutionary changes in marsupials and other vertebrates.

## Methods

### Macroscopic and microscopic procedures

Postnatal (P) *Didelphis albiventris* (Didelphidae, Marsupialia) individuals at 0, 5 and 10 days of age were harvested from the pouches of pregnant females captured during the breeding season (from June to January) at the Ecologic Station of the Federal University of Minas Gerais (19° 52' S, 43° 58' W, Belo Horizonte, MG, Brazil). Six specimens were harvested for each stage and euthanized by decapitation; three of the specimens were randomly selected for RNA-seq procedures and three for histological procedures. Neonates from the same developmental stage were harvested from the same mother and came from the same litter (siblings). All animal procedures were conducted following the instructions of the ethical committee (CEUA/UFMG, ethical approval 220/2008), the Brazilian Institute of Environment and Renewable Natural Resources (IBAMA, SISBIO 27354-2), and the Brazilian laws for the use of animals in scientific experiments.

According to the committee's instructions, for the removal of the postnatal individuals attached to the mammary papillae, the adult female was immobilized by restraining her cervical region, pelvic limbs and tail. Rotation and traction movements of the embryos were performed to remove the young from the female mammary papillae. These procedures did not require the use of sedation or anaesthetics. Decapitation using a scalpel or steel blade is an immediate, quick procedure, especially in postnatal individuals, without the need for anaesthetics.

After harvesting, the young animals were measured and decapitated, and their bodies were divided into halves along the median plane. All sections of this report adhere to the ARRIVE Guidelines for reporting animal research [57]. A completed ARRIVE guidelines checklist is included in **Checklist S1**.

### In situ hybridization (ISH)

For ISH, oligonucleotides (ATGGCAGCAGGCGGCAGC and TCAGCATCTTAGCAGACATGGGG) were designed using the *D. albiventris* sequence of the titin transcript assembled in this work as a template to allow the amplification of a fragment of ~400 bp of its mRNA. The amplified fragment was cloned into the *pGEM-T Easy Vector* (Promega). The antisense RNA probe was synthesized from the SP6 promoter of this clone in the presence of digoxigenin-UTP (Roche).

*D. albiventris* half trunks were fixed in 4% paraformaldehyde (PFA) for 2 days (P0), 4 days (P5), or 7 days (P10) at room temperature (RT) and embedded in paraffin with the anterior surface (for head samples) or the median surface (for body samples) facing the cut plane. Serial sections of 5 µm were obtained. The tissues were pre-treated with 10 µg/ml proteinase K (Sigma-Aldrich), and pre-hybridization was performed in the hybridization buffer (50% formamide, 5x SSC pH 4.5, 1% SDS, 500 µg/ml tRNA, 50 µg/ml heparin) in a water bath at 55°C. Incubation with the antisense RNA probe was performed at 55°C overnight, followed by incubation with 2% Boehringer blocking reagent (BBR)/MABT (0.2 M maleic acid, 0.3 M NaCl, 0.4 M NaOH, 0.1% Tween 20, DEPC-treated water), followed by another incubation with an anti-DIG-AP (Roche) antibody at 1:2,000 in 2% BBR/MABT. Probe detection was performed after washing the samples in MABT, NTMT (5 M NaCl, Tris pH 9.5, 2 M MgCl<sub>2</sub>, 0.1% Tween 20, 1 M levamisole, DEPC-treated water), and NTMT with 20 µl/ml NBT/Bcipy (nitro-blue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate; Roche). The slides were mounted with 80% glycerol in PBS. The hybridized sections were photographed under a BX-41 Olympus microscope using a Q-Color 3/Olympus capture system.

### Total RNA extraction and RNA-seq sequencing

After euthanasia, the animal samples were macerated and homogenized in *TRizol Reagent*® (Invitrogen) using Dispensing Tools S10N-10G (IKA WERKE). *TRizol Reagent*® (Invitrogen) was used at 1 ml of *TRizol* per 100 mg of tissue. The quality of the isolated total RNA was evaluated in a Bioanalyzer (Agilent). RNA integrity numbers (RINs) ranging from 8 to 10 were considered suitable for RNA-seq.

For cDNA library construction, 2 µg of total RNA was treated with 1 U of *DNase I, Amplification Grade* (Invitrogen) and purified according to the Illumina protocol using magnetic microspheres for messenger RNA separation. The purified mRNA was fragmented in Illumina buffer. *SuperScript II* (Invitrogen) and oligo(dT) were used for reverse transcription of the first cDNA strand. The second cDNA strand was synthesized using *RNase H* and *DNA Polymerase I* (Illumina). The ends of the molecules were treated with *T4 DNA Polymerase* and *Klenow DNA Polymerase* (Illumina). The 3' end of the synthesized cDNA were phosphorylated with *T4 PNK* (Illumina) and adenylated with *Klenow exo* (Illumina). Adaptors were added to the cDNA ends, and the samples were purified and selected by size (200 bp ±25 bp) through agarose gel electrophoresis (*QIAquick Gel Extraction Kit*, QIAGEN). The purified cDNA was quantified by RT-qPCR using adaptor-specific oligonucleotides (Illumina).

RNA-seq sequencing was performed using the *HiSeq 2500* platform (Illumina) according to the recommendations of the manufacturer and the paired-end reads protocol. Each sample was sequenced until reaching approximately 5 million reads/library.

### Transcriptome assembly and annotation

The *D. albiventris* transcriptome was assembled using a *de novo* approach. The raw data were trimmed to remove adaptors, sequencing artefacts, and low-quality fragments using the *NGS QC Toolkit* [58] and a Phred quality score of 15. High-quality reads were subjected to *in silico* normalization prior to *de novo* assembly. For normalization, decreasing sequencing coverage of highly represented regions with a fragment density greater than 30x sequencing coverage was considered. The normalized data were subjected to *de novo* transcriptome assembly using the *Trinity* tool [59]. In the assembly process, a minimum fragment overlap of 35 bp was considered. Only contigs longer than 300 bp were included within the assembled transcriptome for further analysis, which involved the following approaches: (i) differential gene expression analysis; (ii) gene ontology (GO) sequence annotation using *Blast2GO* [60]; and *KAAS* [61, 62].

### Differential gene expression analysis

For gene expression analysis, the assembled transcriptome was used as a gene model reference. All the raw RNA-seq data from the libraries were used in this step following data quality filtering and read alignment.

The libraries for each postnatal stage were separately aligned to the assembled reference transcriptome using the default parameters of *Bowtie* software [63] and then labelled together as biological replicates for statistical analysis.

A count matrix was constructed to represent the absolute number of aligned reads for each assembled transcript, with rows representing transcripts and columns representing fragment counts for a specific sample. The data distribution was fit according to a negative binomial distribution method, and a false discovery rate (FDR) correction [64] of 5% was then applied to control false-positive significance of transcript expression variation.

Statistical analysis was performed using the *DESeq2* package of *R/Bioconductor* [65]. Differential expression between developmental stages was detected for those transcripts with statistical significance ( $p < 0.05$ ). Differentially expressed (DE) genes were analysed with *Blast2GO* for GO annotation [60]. Quantitative differences in differentially expressed transcripts between the developmental stages were represented in a Venn diagram to identify common or exclusive transcripts among stages. A transcript was considered “exclusive” if its expression was detected in only one stage of development among those analysed in this work. The upregulated [ $\log_2(\text{fold change}) > 1$ ] and downregulated [ $\log_2(\text{fold change}) < -1$ ] genes between each pair of stages were also represented in Venn diagrams. The Venn diagram images were generated using Venny 2.1 [65].

The count matrix was used to calculate and generate a Euclidean distance matrix for hierarchical sample clustering, grouping samples according to the most similar transcriptome profiles. This method was used to generate a colourmap matrix representing how the sequenced samples (biological replicates) were correlated. The normalized count matrix was also employed to perform PCA (principal component analysis) using the *R* library *ggplots* from *DESeq/Bioconductor*. Additionally, a single linkage method was used to generate a dendrogram and a colourmap matrix to correlate all sample expression profiles according to colour, ranging from green (most different) to red (identical profiles).

A table of transcripts that were simultaneously expressed at the three stages was used to calculate and generate a matrix containing the expression levels of each transcript in each developmental stage. Additionally, a single linkage method was used to generate a dendrogram and a colourmap matrix to correlate the expression levels according to colour, ranging from green (lower expression values) to red (higher expression values). The colourmap matrix was obtained using the *d3heatmap* package of RStudio v3.4.1.

### Coverage between marsupial amino acid sequences

The Basic Local Alignment Search Tool (BLAST) X (NCBI BLAST+, available at <ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/>) with default parameters was used to compare the nucleotide sequences of the *D. albiventris* transcripts to the amino acid sequences of *M. domestica* (available at [ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/002/295/GCF\\_000002295.2\\_MonDom5/GCF\\_000002295.2\\_MonDom5\\_protein.faa.gz](ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/002/295/GCF_000002295.2_MonDom5/GCF_000002295.2_MonDom5_protein.faa.gz)). The *D. albiventris* nucleotide sequences were translated into amino acid sequences to verify how many and which proteins of *M. domestica* correspond to those of *D. albiventris*. Sequences of *D. albiventris* showing 20% to 50% coverage identity to the *M. domestica* sequences were evaluated. The probability density graphics of the coverage identity were generated in the *gplots* and *ggplot2* packages of RStudio software.

### Coverage between marsupial nucleotide sequences

The coverage between the nucleotide sequences of *D. albiventris* and other marsupial species (*M. domestica*, *S. harrisi*, *P. cinereus*, and *M. eugenii*) was evaluated with the Basic Local Alignment Search Tool (BLAST) N (NCBI BLAST+, available at <ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/>). This analysis was performed between the assembled skunk transcriptome and the assembled transcriptome available in the *Monodelphis domestica* NCBI database ([ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/002/295/GCF\\_000002295.2\\_MonDom5/GCF\\_000002295.2\\_MonDom5\\_ma.fna.gz](ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/002/295/GCF_000002295.2_MonDom5/GCF_000002295.2_MonDom5_ma.fna.gz)) and those of *Sarcophilus harrisi* ([ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/189/315/GCF\\_000189315.1\\_Devil\\_ref\\_v7.0/GCF\\_000189315.1\\_Devil\\_ref\\_v7.0\\_rna.fna.gz](ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/189/315/GCF_000189315.1_Devil_ref_v7.0/GCF_000189315.1_Devil_ref_v7.0_rna.fna.gz)) and *Phascolarctos cinereus* ([ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/002/099/425/GCF\\_002099425.1\\_phaCin\\_unsw\\_v4.1/GCF\\_002099425.1\\_phaCin\\_unsw\\_v4.1\\_rna.fna.gz](ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/002/099/425/GCF_002099425.1_phaCin_unsw_v4.1/GCF_002099425.1_phaCin_unsw_v4.1_rna.fna.gz)). For *Macropus eugenii*, we used the genome sequence ([ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/004/035/GCA\\_000004035.1\\_Meug\\_1.1/GCA\\_000004035.1\\_Meug\\_1.1\\_genomic.fna.gz](ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/004/035/GCA_000004035.1_Meug_1.1/GCA_000004035.1_Meug_1.1_genomic.fna.gz)), as its

transcriptome is not available. The probability density graphics of the coverage identity were generated in the *gplots* and *ggplot2* packages of RStudio software.

## RT-qPCR

cDNA was synthesized using 1 µg of total RNA following the recommendations of the *RevertAid™ H Minus First Strand cDNA Synthesis* kit (Fermentas). Diluted cDNAs (1:10 v/v) and *SYBR® Green Master Mix* (Bio-Rad) were used for the quantification of target genes via the quantitative reverse transcription polymerase chain reaction (RT-qPCR) approach in a *Rotor-Gene 3000* system (Corbett Research). Fluorescence was measured at the end of extension, providing the cycle threshold ( $C_T$ ) values. All reactions, including those with a template and omitting the template (negative controls), were run in biological triplicates (three biological samples for each developmental stage – P0, P5 and P10). The four evaluated genes of interest (GOI) followed by their forward (F) and reverse (R) primers in the 5' to 3' direction on the sense strand are as follows: HMGB3 (TCCATTATTACACAAACCAAGCA; GGAGGAAGAGGAGGAAGAGG), CRYGB (CATGGGCCATCAGTATTACC; GGAGCGTATTTTCATTCATGTG), APOH (ATGAGCCAGGGGAACAAA; TTCACAGTATTAGGGTAGTCAAAGG) and HBZ (GCCCTCGTCCTCACCATCT; TGACAGCATCTCCAATAGCAC). The UBC gene was used as a reference gene (CACTTGCTGCTGCTCTTC; TGCCTCTTATTTGACCTTCTTC). The oligonucleotides were designed to amplify an ~200 bp amplicon from within different exons following the Minimum Information for Quantitative RT-PCR Experiments (MIQE) guidelines [66]. The applied cycling parameters were as follows: 50°C for 2 min, 95°C for 2 min, followed by 40 cycles of 95°C for 15 sec, 60°C for 30 sec and 72°C for 20 sec.

The analysis of differential gene expression was performed using REST 2009 (*Relative Expression Software Tool*, V.2.0.13) software via randomization tests (*Pair Wise Fixed Reallocation Randomization Test*<sup>®</sup>) [67] with 95% significance.

## DIAMOND and BLASTP analyses

DIAMOND v0.9.19 (double index alignment of next-generation sequencing data) [68] was used to align the assembled transcripts against the NCBI-nr database. The following parameter settings were used for the program: -e 0.00001 -k 1 -masking 0 -matrix blosum62 -gapopen 11 -gapextend 1 -more-sensitive -salltitles -sallseqid. The results returned only the hits with the best score for each opossum transcript.

All DIAMOND subject IDs were listed without redundancies and used as the protein database input in the NCBI tool Batch Entrez (<https://www.ncbi.nlm.nih.gov/sites/batchentrez>). The output was used as the input for BLASTP (Basic Local Alignment Search Tool) analysis, using the NCBI BLAST+ tool (<ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/>) together with the *M. domestica* amino acid sequences ([ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/002/295/GCF\\_000002295.2\\_MonDom5/GCF\\_000002295.2\\_MonDom5\\_protein.faa.gz](ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/002/295/GCF_000002295.2_MonDom5/GCF_000002295.2_MonDom5_protein.faa.gz)). The following parameter settings were used for the program: -evalue 1e-5 -max\_target\_seqs 1. Thus, we obtained UniProt Entry IDs (using Retrieve ID tool; <http://www.UniProt.org/uploadlists/>) for further ClueGO analysis (**Additional File 3**). The grey short-tailed opossum was selected for further analysis using GO enrichment networks, as data for this species are available in ClueGO, and the DIAMOND results showed the greatest number of hits for *M. domestica*.

## GO pathway enrichment analysis

The plugin ClueGO (v.2.3.5) [69] for Cytoscape (v3.4.0) [70] was used for GO enrichment analysis. The node colours represent the functional groups. The node size represents the term enrichment significance. The edges connecting the nodes were based on the Kappa statistic (Kappa score threshold of 0.4). Only the most significant term in the group was labelled. The right-sided hypergeometric test was used to identify overrepresented GO terms, and the Benjamini-Hochberg method was used for the correction of p-values ( $p < 0.05$ ).

## List Of Abbreviations

AIPL1: Aryl Hydrocarbon Receptor Interacting Protein Like 1

Anti-DIG-AP: anti-digoxigenin antibody, conjugated with alkaline phosphatase

APOH: Apolipoprotein H

BBR: Boehringer Blocking Reagent

Bcip: 5-bromo-4-chloro-3'-indolyphosphate

BLAST: Basic Local Alignment Search Tool

BLASTN: search nucleotide databases using a nucleotide query

BLASTP: search protein databases using a protein query

BLASTX: search protein databases using a translated nucleotide query

bp: base pair

cDNA: complementary DNA

CNS: central nervous system

CRYBA1: Crystallin Beta A1

CRYBA4: Crystallin Beta A4

CRYBB1: Crystallin Beta B

CRYBB2: Crystallin Beta B2

CRYGB: Crystallin Gamma B

C<sub>T</sub>: cycle threshold

DE: differentially expressed

DEPC: diethyl pyrocarbonate

DIAMOND: double index alignment of next-generation sequencing data

EKLF: Erythroid Kruppel-Like Factor

EMP: erythro-myeloid progenitor

F: forward primer

FDR: False Discovery Rate

FGF4: Fibroblast Growth Factor 4

GJC1: Gap Junction Protein Gamma 1

GO: gene ontology

GOI: genes of interest

GPR37L1: G Protein-Coupled Receptor 37 Like 1

HBZ: Hemoglobin subunit Zeta

HES1: Hairy and Enhancer of Split 1

HMGB3: High Mobility Group Box 3

ID: identifier

IgA: immunoglobulin A

IRX1: Iroquois transcription factor

ISH: *in situ* hybridization

KLF1: Kruppel Like Factor 1

MABT: maleic acid buffer containing Tween 20

MEP: megakaryocyte-erythrocyte progenitor

MGAT3: Mannosyl (Beta-1,4)-Glycoprotein Beta-1,4-N-Acetylglucosaminyltransferase

MIQE: Minimum Information for Quantitative RT-PCR Experiments

MOGAT3: Monoacylglycerol O-Acyltransferase 3

mRNA: messenger RNA

MSX1: Msh Homeobox 1

NBT: nitro-blue tetrazolium

NCBI: National Center for Biotechnology Information

NGS: Next Generation Sequencing

NHE4: Na<sup>+</sup>/H<sup>+</sup> Exchanger 4

nr: non-redundant

NST: non-shivering thermogenesis

NTMT: NaCl, Tris-HCl, MgCl<sub>2</sub>, Tween 20

oligo(dT): deoxythymine nucleotides

p, or p-value: probability value, or significance value

P: postnatal stage

PBS: phosphate-buffered saline

PCA: Principal Component Analysis

PFA: paraformaldehyde

PTK7: Tyrosine-Protein Kinase-Like 7

R: reverse primer

REST: Relative Expression Software Tool

RIN: RNA integrity number

RNA: ribonucleic acid

RNA-seq: RNA sequencing

RT: room temperature

RT-qPCR: quantitative reverse transcription polymerase chain reaction

SD: standard deviation

SDS: sodium dodecyl sulfate

SLC9A4: Solute Carrier Family 9 Member A4

SOBP: Sine Oculis Binding Protein Homolog

SSB: S-shaped body

SSC: saline-sodium citrate

tRNA: transfer RNA

UBC: Ubiquitin C

UTP: uridine-5'-triphosphate

v/v: volume/volume

## **Declarations**

**Ethics approval and consent to participate**

All animal procedures were conducted following the instructions of the ethical committee (CEUA/UFMG, ethical approval 220/2008), the Brazilian Institute of Environment and Renewable Natural Resources (IBAMA, SISBIO 27354-2), and the Brazilian laws for the use of animals in scientific experiments. No experiment was conducted on humans or with human cells or tissues in this study.

### Consent for publication

Not applicable.

### Availability of data and materials

The raw transcriptome sequencing data (RNA-seq) from the biological replicates and developmental stages of *Didelphis albiventris* are available under the NCBI-BioProject submission code PRJNA544055.

The datasets supporting the conclusions of this article are included within the article and its additional files.

### Competing interests

The authors declare that they have no competing interests.

### Funding

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. The present study was also supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) and Pró-Reitoria de Pesquisa da UFMG (PRPq/UFMG). Erika C Jorge, Luiz L Coutinho and Cláudia BM Vitorello received a scholarship from CNPq.

### Authors' Contributions

The study was conceived by ECJ and GABS; the experiments were designed by ECJ, IGDS, LLC, CBMV, TAOM and RHH and performed by IGDS, RHH, PDCS and ABCF. Illustrations were drafted by IGDS and RHH. Data analysis was performed by IGDS, TAOM, RHH, ABCF, AMSR and ECJ. IGDS, TAOM and ECJ wrote the paper. All authors have read and approved the manuscript.

### Acknowledgements

We wish to thank Professors Germán A. B. Mahecha, Anderson J. Ferreira, Marcos A. de Sá, José C. Nogueira and Aristóteles G. Neto; Fabiano Pais, PhD, and Rodrigo Kato, PhD; and the laboratory technicians Francisco A. P. Cotrim and Horácio Montenegro for their contributions.

## References

- [1] May-Collado LJ, Kilpatrick CW, Agnarsson I. Mammals from 'down under': a multi-gene species-level phylogeny of marsupial mammals (Mammalia, Metatheria). *PeerJ*. 2015; doi:10.7717/peerj.805.
- [2] Renfree MB. Marsupial reproduction: the choice between placentation and lactation. *Oxford Reviews of Reproductive Biology*. 1983;5:1-26.
- [3] Freyer C, Zeller U, Renfree MB. Ultrastructure of the placenta of the tammar wallaby, *Macropus eugenii*: comparison with the grey short-tailed opossum, *Monodelphis domestica*. *J Anat*. 2002; doi:10.1046/j.1469-7580.2002.00084.x.
- [4] Freyer C, Renfree MB. The mammalian yolk sac placenta. *J Exp Zool B Mol Dev Evol*. 2009; doi: 10.1002/jez.b.21239.
- [5] Tyndale-Biscoe H, Renfree M. *Reproductive Physiology of Marsupials*. Cambridge: Cambridge University Press; 1987.
- [6] Smith KK. Craniofacial development in marsupial mammals: developmental origins of evolutionary change. *Dev Dyn*. 2006;235:1181-93.
- [7] Aplin KP, Archer M. Recent advances in marsupial systematics with a new syncretic classification. In: Archer M, editor. *Possums and opossums: Studies in evolution*. Sydney: Surrey Beatty; 1987. p. XV-LXXII.
- [8] Fonseca CT, Alves JB. Dental development of *Didelphis albiventris* (Marsupialia): I-incisors and canines. *Braz J Biol*. 2006;66:53-60.
- [9] Deakin JE. Marsupial genome sequences: providing insight into evolution and disease. *Scientifica (Cairo)*. 2012; doi:10.6064/2012/543176.
- [10] Pharo EA. Marsupial milk: a fluid source of nutrition and immune factors for the developing pouch young. *Reprod Fertil Dev*. 2019; doi:10.1071/RD18197.
- [11] Sharp JA, Wanyonyi S, Modepalli V, Watt A, Kuruppath S, Hinds LA, Kumar A, Abud HE, Lefevre C, Nicholas KR. The tammar wallaby: A marsupial model to examine the timed delivery and role of bioactives in milk. *Gen Comp Endocrinol*. 2017; doi:10.1016/j.ygcen.2016.08.007.

- [12] Guernsey MW, Chuong EB, Cornelis G, Renfree MB, Baker JC. Molecular conservation of marsupial and eutherian placentation and lactation. *Elife*. 2017; doi:10.7554/eLife.27450.
- [13] Sauerland C, Menzies BR, Glatzle M, Seeger J, Renfree MB, Fietz SA. The Basal Radial Glia Occurs in Marsupials and Underlies the Evolution of an Expanded Neocortex in Therian Mammals. *Cereb Cortex*. 2018; doi:10.1093/cercor/bhw360.
- [14] Ramamurthy DL, Krubitzer LA. Neural Coding of Whisker-Mediated Touch in Primary Somatosensory Cortex Is Altered Following Early Blindness. *J Neurosci*. 2018; doi:10.1523/JNEUROSCI.0066-18.2018.
- [15] Modepalli V, Kumar A, Sharp JA, Saunders NR, Nicholas KR, Lefèvre C. Gene expression profiling of postnatal lung development in the marsupial gray short-tailed opossum (*Monodelphis domestica*) highlights conserved developmental pathways and specific characteristics during lung organogenesis. *BMC Genomics*. 2018; doi:10.1186/s12864-018-5102-2.
- [16] Sánchez-Villagra MR, Forasiepi AM. On the development of the chondrocranium and the histological anatomy of the head in perinatal stages of marsupial mammals. *Zoological Lett*. 2017; doi:10.1186/s40851-017-0062-y.
- [17] Ferner K. Skin structure in newborn marsupials with focus on cutaneous gas exchange. *J Anat*. 2018; doi:10.1111/joa.12843.
- [18] Oliveira ÉV, Zimicz N, Goin FJ. Taxonomy, affinities, and paleobiology of the tiny metatherian mammal *Minusculodelphis*, from the early Eocene of South America. *Naturwissenschaften*. 2016; doi:10.1007/s00114-015-1331-2.
- [19] Jansa SA, Barker FK, Voss RS. The early diversification history of didelphid marsupials: a window into South America's "Splendid Isolation". *Evolution*. 2014; doi:10.1111/evo.12290.
- [20] Flores DA, Giannini N, Abdala F. Evolution of post-weaning skull ontogeny in New World opossums (Didelphidae). *Organisms Diversity & Evolution*. 2018; doi:10.1007/s13127-018-0369-3.
- [21] Mikkelsen TS, Wakefield MJ, Aken B, Amemiya CT, Chang JL, Duke S, Garber M, Gentles AJ, Goodstadt L, Heger A, Jurka J, Kamal M, Mauceli E, Searle SM, Sharpe T, Baker ML, Batzer MA, Benos PV, Belov K, Clamp M, Cook A, Cuff J, Das R, Davidow L, Deakin JE, Fazzari MJ, Glass JL, Grabherr M, Greally JM, Gu W, Hore TA, Huttley GA, Kleber M, Jirtle RL, Koina E, Lee JT, Mahony S, Marra MA, Miller RD, Nicholls RD, Oda M, Papenfuss AT, Parra ZE, Pollock DD, Ray DA, Schein JE, Speed TP, Thompson K, VandeBerg JL, Wade CM, Walker JA, Waters PD, Webber C, Weidman JR, Xie X, Zody MC; Broad Institute Genome Sequencing Platform.; Broad Institute Whole Genome Assembly Team., Graves JA, Ponting CP, Breen M, Samollow PB, Lander ES, Lindblad-Toh K. Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. *Nature*. 2007; doi:10.1038/nature05805.
- [22] Hansen VL, Schilkey FD, Miller RD. Transcriptomic Changes Associated with Pregnancy in a Marsupial, the Gray Short-Tailed Opossum *Monodelphis domestica*. *PLoS One*. 2016; doi:10.1371/journal.pone.0161608.
- [23] Murchison EP, Schulz-Trieglaff OB, Ning Z, Alexandrov LB, Bauer MJ, Fu B, Hims M, Ding Z, Ivakhno S, Stewart C, Ng BL, Wong W, Aken B, White S, Alsop A, Becq J, Bignell GR, Cheetham RK, Cheng W, Connor TR, Cox AJ, Feng ZP, Gu Y, Grocock RJ, Harris SR, Khrebtukova I, Kingsbury Z, Kowarsky M, Kreiss A, Luo S, Marshall J, McBride DJ, Murray L, Pearse AM, Raine K, Rasolonjatovo I, Shaw R, Tedder P, Tregidgo C, Vilella AJ, Wedge DC, Woods GM, Gormley N, Humphray S, Schroth G, Smith G, Hall K, Searle SM, Carter NP, Papenfuss AT, Futreal PA, Campbell PJ, Yang F, Bentley DR, Evers DJ, Stratton MR. Genome sequencing and analysis of the Tasmanian devil and its transmissible cancer. *Cell*. 2012; doi:10.1016/j.cell.2011.11.065.
- [24] Hewavisenti RV, Morris KM, O'Meally D, Cheng Y, Papenfuss AT, Belov K. The identification of immune genes in the milk transcriptome of the Tasmanian devil (*Sarcophilus harrisii*). *PeerJ*. 2016; doi:10.7717/peerj.1569.
- [25] Renfree MB, Papenfuss AT, Deakin JE, Lindsay J, Heider T, Belov K, Rens W, Waters PD, Pharo EA, Shaw G, Wong ESW, Lefevre CM, Nicholas KR, Kuroki Y, Wakefield MJ, Zenger KR, Wang C, Ferguson-Smith M, Nicholas FW, Hickford D, Yu H, Short KR, Siddle HV, Frankenberg SR, Chew KY, Menzies BR, Stringer JM, Suzuki S, Hore TA, Delbridge ML, et al: Genome sequence of an Australian kangaroo, *Macropus eugenii*, provides insight into the evolution of mammalian reproduction and development. *Genome Biol*. 2011; doi:10.1186/gb-2011-12-8-r81.
- [26] Lefèvre CM, Digby MR, Whitley JC, Strahm Y, Nicholas KR. Lactation transcriptomics in the Australian marsupial, *Macropus eugenii*: transcript sequencing and quantification. *BMC Genomics*. 2007; doi:10.1186/1471-2164-8-417.
- [27] Hobbs M, Pavasovic A, King AG, Prentis PJ, Eldridge MD, Chen Z, Colgan DJ, Polkinghorne A, Wilkins MR, Flanagan C, Gillett A, Hanger J, Johnson RN, Timms P. A transcriptome resource for the koala (*Phascolarctos cinereus*): insights into koala retrovirus transcription and sequence diversity. *BMC Genomics*. 2014; doi:10.1186/1471-2164-15-786.
- [28] Johnson RN, O'Meally D, Chen Z, Etherington GJ, Ho SYW, Nash WJ, Grueber CE, Cheng Y, Whittington CM, Dennison S, Peel E, Haerty W, O'Neill RJ, Colgan D, Russell TL, Alquezar-Planas DE, Attenbrow V, Bragg JG, Brandies PA, Chong AY, Deakin JE, Di Palma F, Duda Z, Eldridge MDB, Ewart KM, Hogg CJ, Frankham GJ, Georges A, Gillett AK, Govendir M, Greenwood AD, Hayakawa T, Helgen KM, Hobbs M, Holleley CE, Heider TN, Jones EA, King A,

- Madden D, Graves JAM, Morris KM, Neaves LE, Patel HR, Polkinghorne A, Renfree MB, Robin C, Salinas R, Tsangaras K, Waters PD, Waters SA, Wright B, Wilkins MR, Timms P, Belov K. Adaptation and conservation insights from the koala genome. *Nat Genet.* 2018; doi:10.1038/s41588-018-0153-5.
- [29] Morris KM, O'Meally D, Zaw T, Song X, Gillett A, Molloy MP, Polkinghorne A, Belov K. Characterisation of the immune compounds in koala milk using a combined transcriptomic and proteomic approach. *Sci Rep.* 2016; doi:10.1038/srep35011.
- [30] Udy DB, Voorhies M, Chan PP, Lowe TM, Dumont S. Draft De Novo Transcriptome of the Rat Kangaroo *Potorous tridactylus* as a Tool for Cell Biology. *PLoS One.* 2015; doi: 10.1371/journal.pone.0134738.
- [31] Morris KM, Weaver HJ, O'Meally D, Desclozeaux M, Gillett A, Polkinghorne A. Transcriptome sequencing of the long-nosed bandicoot (*Perameles nasuta*) reveals conservation and innovation of immune genes in the marsupial order Peramelemorphia. *Immunogenetics.* 2018; doi:10.1007/s00251-017-1043-1.
- [32] Whittington CM, O'Meally D, Laird MK, Belov K, Thompson MB, McAllan BM. Transcriptomic changes in the pre-implantation uterus highlight histotrophic nutrition of the developing marsupial embryo. *Sci Rep.* 2018; doi:10.1038/s41598-018-20744-z.
- [33] Nespolo RF, Gaitan-Espitia JD, Quintero-Galvis JF, Fernandez FV, Silva AX, Molina C, Storey KB, Bozinovic F. A functional transcriptomic analysis in the relict marsupial *Dromiciops gliroides* reveals adaptive regulation of protective functions during hibernation. *Mol Ecol.* 2018; doi: 10.1111/mec.14876.
- [34] Feigin CY, Newton AH, Doronina L, Schmitz J, Hipsley CA, Mitchell KJ, Gower G, Llamas B, Soubrier J, Heider TN, Menzies BR, Cooper A, O'Neill RJ, Pask AJ. Genome of the Tasmanian tiger provides insights into the evolution and demography of an extinct marsupial carnivore. *Nat Ecol Evol.* 2018; doi:10.1038/s41559-017-0417-y.
- [35] Krause WJ, Cutts H. Placentation in the Opossum, *Didelphis virginiana*. *Cells Tissues Organs.* 1985; doi:10.1159/000146058.
- [36] Streilein DE. The ecology of small mammals in the semi-arid Brazilian Caatinga. 3. Reproduction biology and population ecology. *Ann Carnegie Mus.* 1982; 51:251-269.
- [37] Talice RV, Lagomarsino JC. Comportamiento sexual y nacimientos en cautividad de la 'comadreja overa': *Didelphis azarae*. *Congreso Sudamericano Zool.* 1959; 5:81-96.
- [38] Bertassoli BM, Santos AC, Oliveira FD, Oliveira DM, Assis-Neto AC, Carvalho AF. Morfologia da laringe e traqueia de gambás (*Didelphis* sp.). *Cienc Anim Bras.* 2013; doi:10.5216/cab.v14i2.17044.
- [39] Kirsch JA. The comparative serology of Marsupialia, and a classification of marsupials. *Aust J Zool Suppl Ser.* 1977; doi:10.1071/AJZS052.
- [40] Gonçalves NJN, Mançanares CAF, Miglino MA, Samoto VY, Martins DS, Ambrosio CE, Ferraz RH, Carvalho AF. Morphological aspects of genital female of the opossum (*Didelphis* sp.). *Braz J Vet Res Anim Sci.* 2009;46:332-8.
- [41] Paglia, A.P., Fonseca, G.A.B., Rylands, A.B., Herrmann, G., Aguiar, L.M.S., Chiarello AG, Leite YLR, Costa LP, Siciliano S, Kierulff MCM, Mendes SL, Tavares VC, Mittermeier RA, Patton JL. Annotated Checklist of Brazilian Mammals. Arlington: Conservation International; 2012. p. 1-76.
- [42] Cáceres NC, de Moraes Weber M, Melo GL, Meloro C, Sponchiado J, Carvalho Rdos S, Bubadué Jde M. Which Factors Determine Spatial Segregation in the South American Opossums (*Didelphis aurita* and *D. albiventris*)? An Ecological Niche Modelling and Geometric Morphometrics Approach. *PLoS One.* 2016; doi:10.1371/journal.pone.0157723.
- [43] Bezerra CM, Cavalcanti LP, Souza Rde C, Barbosa SE, Xavier SC, Jansen AM, Ramalho RD, Diotaiuti L. Domestic, peridomestic and wild hosts in the transmission of *Trypanosoma cruzi* in the Caatinga area colonised by *Triatoma brasiliensis*. *Mem Inst Oswaldo Cruz.* 2014;109:887-98.
- [44] Orozco MM, Enriquez GF, Alvarado-Otegui JA, Cardinal MV, Schijman AG, Kitron U, Gürtler RE. New sylvatic hosts of *Trypanosoma cruzi* and their reservoir competence in the humid Chaco of Argentina: a longitudinal study. *Am J Trop Med Hyg.* 2013; doi:10.4269/ajtmh.12-0519.
- [45] Paiz LM, Donalísio MR, Richini-Pereira VB, Motoie G, Castagna CL, Tolezano JE. Infection by *Leishmania* spp. in Free-Ranging Opossums (*Didelphis albiventris*) in an Environmentally Protected Area Inhabited by Humans in Southeastern Brazil. *Vector Borne Zoonotic Dis.* 2016;16:728-30.
- [46] Cáceres NC. Food habits and seed dispersal by the white-eared opossum, *Didelphis albiventris*, in the southern Brazil Stud Neotrop Fauna Environ. 2002; doi:10.1590/S1676-06032010000200004.
- [47] Vellard J. Resistencia de los *Didelphis* (*Zarigueya*) a los venenos ofídicos; nota prévia. *Rev Bras Biol.* 1945;5:463-7.
- [48] Vellard J. Investigaciones sobre inmunidad natural contra los venenos de serpientes. *J Pub Mus Hist Nat Javier Prado.* 1949;4:72-96.

- [49] Farah MFL, One M, Novello JC, Toyama MH, Perales J, Moussatché H, Domont GB, Oliveira B, Marangoni S. Isolation of protein factors from opossum (*Didelphis albiventris*) serum, which protect against *Bothrops jararaca* venom. *Toxicon*. 1996;34:1067-71.
- [50] Santori RT. Discrimination of millipedes by the opossum *Didelphis albiventris* (Marsupialia, Didelphidae). *J Adv Zool*. 1998;19:118-9.
- [51] Dos Santos ÍG, Jorge EC, Copola AG, Bertassoli BM, Goes AM, Silva GA. FGF2, FGF3 and FGF4 expression pattern during molars odontogenesis in *Didelphis albiventris*. *Acta Histochem*. 2017; doi:10.1016/j.acthis.2016.12.001.
- [52] Santos IGD. Análise da expressão dos Fatores de Crescimento Fibroblásticos 2, 3 e 4 durante a odontogênese de molares do gambá *Didelphis albiventris*, um modelo promissor para estudos em Biologia do Desenvolvimento. In: Biblioteca Digital. Universidade Federal de Minas Gerais. 2014. <http://www.bibliotecadigital.ufmg.br/dspace/handle/1843/BUBD-AEPPD6>. Accessed 16 Jul 2017.
- [53] Torres CB, Alves JB, Silva GA, Goes VS, Nakao LY, Goes AM. Role of BMP-4 during tooth development in a model with complete dentition. *Arch Oral Biol*. 2008; doi:10.1016/j.archoralbio.2007.07.005.
- [54] Torres CB, Goes VS, Goes AM, Pacífico LG, Silva GA, Junior NL, Alves JB. Fibroblast growth factor 9: cloning and immunolocalisation during tooth development in *Didelphis albiventris*. *Arch Oral Biol*. 2006; doi:10.1016/j.archoralbio.2005.08.003.
- [55] Fonseca, CT. Estudo histológico do desenvolvimento dentário do *Didelphis albiventris* (Lund, 1841) – Didelphidae Marsupialia [Master's Thesis]. Belo Horizonte (MG): Universidade Federal de Minas Gerais; 1996.
- [56] Gilbert SF, Bosch TC, Ledón-Rettig C. Eco-Evo-Devo: developmental symbiosis and developmental plasticity as evolutionary agents. *Nat Rev Genet*. 2015; doi:10.1038/nrg3982.
- [57] Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol*. 2010; doi:10.1371/journal.pbio.1000412.
- [58] Patel RK, Jain M. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One*. 2012; doi:10.1371/journal.pone.0030619.
- [59] Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. Trinity: reconstructing a full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat Biotechnol*. 2011; doi:10.1038/nbt.1883.
- [60] Conesa A, Götz S. Blast2GO: A comprehensive suite for functional analysis in plant genomics. *Int J Plant Genomics*. 2008; doi:10.1155/2008/619832
- [61] Tanabe M, Kanehisa M. Using the KEGG database resource. *Curr Protoc Bioinformatics*. 2012; doi:10.1002/0471250953.bi0112s38.
- [62] Kanehisa M. The KEGG database. *Novartis Found Symp*. 2002;247:91-101.
- [63] Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012; doi:10.1038/nmeth.1923.
- [64] Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol*. 1995;57:289-300.
- [65] Oliveros JC. Venny. An interactive tool for comparing lists with Venn's diagrams. 2007-2015. <http://bioinfogp.cnb.csic.es/tools/venny/index.html>. Accessed 12 March 2018.
- [66] Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem*. 2009;55:611-22.
- [67] Pfaffl MW, Horgan GW, Dempfle L. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res*. 2002;30:e36.
- [68] Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*. 2015; doi:10.1038/nmeth.3176.
- [69] Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pagès F, Trajanoski Z, Galon J. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*. 2009;25:1091-3.
- [70] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003; doi:10.1101/gr.1239303.
- [71] Nemeth MJ, Kirby MR, Bodine DM. Hmgb3 regulates the balance between hematopoietic stem cell self-renewal and differentiation. *Proc Natl Acad Sci U S A*. 2006;103:13783-8.

- [72] Nemeth MJ, Curtis DJ, Kirby MR, Garrett-Beal LJ, Seidel NE, Cline AP, Bodine DM. Hmgb3: an HMG-box family member expressed in primitive hematopoietic cells that inhibits myeloid and B-cell differentiation. *Blood*. 2003;102:1298-306.
- [73] Tang L, Zeng W, Lu X, Wang QY, Liu H, Cheng ZP, Wu YY, Hu B, Jian XR, Guo T, Wang HF, Hu Y. Identification of APOH polymorphisms as common genetic risk factors for venous thrombosis in the Chinese population. *J Thromb Haemost*. 2014; doi:10.1111/jth.12679.
- [74] Berger JS, Rockman CB, Guyer KE, Lopez LR. Proatherogenic oxidized low-density lipoprotein/ $\beta$ 2-glycoprotein I complexes in arterial and venous disease. *J Immunol Res*. 2014; doi:10.1155/2014/234316.
- [75] Chui DH, Mentzer WC, Patterson M, Iarocci TA, Embury SH, Perrine SP, Mibashan RS, Higgs DR. Human embryonic zeta-globin chains in fetal and newborn blood. *Blood*. 1989;74:1409-14.
- [76] Bloemendal H, de Jong WW. Lens proteins and their genes. *Prog Nucleic Acid Res Mol Biol*. 1991;41:259-81.
- [77] Bloemendal H, de Jong W, Jaenicke R, Lubsen NH, Slingsby C, Tardieu A. Ageing and vision: structure, stability and function of lens crystallins. *Prog Biophys Mol Biol*. 2004;86:407-85.
- [78] Royall AH, Frankenberg S, Pask AJ, Holland PWH. Of eyes and embryos: subfunctionalization of the CRX homeobox gene in mammalian evolution. *Proc Biol Sci*. 2019; doi:10.1098/rspb.2019.0830.
- [79] Saunders NR, Noor NM, Dziegielewska KM, Wheaton BJ, Liddelow SA, Steer DL, Ek CJ, Habgood MD, Wakefield MJ, Lindsay H, Truettner J, Miller RD, Smith AI, Dietrich WD. Age-dependent transcriptome and proteome following transection of neonatal spinal cord of *Monodelphis domestica* (South American grey short-tailed opossum). *PLoS One*. 2014; doi:10.1371/journal.pone.0099080.
- [80] Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, Szczesniak MW, Gaffney DJ, Elo LL, Zhang X, Mortazavi A. A survey of best practices for RNA-seq data analysis. *Genome Biol*. 2016; doi:10.1186/s13059-016-0881-8.
- [81] Plomin R. Commentary: Why are children in the same family so different? Non-shared environment three decades later. *Int J Epidemiol*. 2011; doi:10.1093/ije/dyq144.
- [82] Poulsen P, Esteller M, Vaag A, Fraga MF. The epigenetic basis of twin discordance in age-related diseases. *Pediatr Res*. 2007;61:38R-42R.
- [83] Fang Z, Cui X. Design and validation issues in RNA-seq experiments. *Brief Bioinform*. 2011;12:280-7.
- [84] Moreton J, Izquierdo A, Emes RD. Assembly, Assessment, and Availability of De novo Generated Eukaryotic Transcriptomes. *Front Genet*. 2016;6:361.
- [85] Hölzer M, Marz M. De novo transcriptome assembly: A comprehensive cross-species comparison of short-read RNA-Seq assemblers. *Gigascience*. 2019; doi:10.1093/gigascience/giz039.
- [86] Krause WJ. *An Atlas of Opossum Organogenesis: Opossum Development*. Boca Raton: Universal Publishers; 2008.
- [87] Nowack J, Giroud S, Arnold W, Ruf T. Muscle Non-shivering Thermogenesis and Its Role in the Evolution of Endothermy. *Front Physiol*. 2017; doi:10.3389/fphys.2017.00889.
- [88] Borthwick CR, Young LJ, Old JM. The development of the immune tissues in marsupial pouch young. *J Morphol*. 2014; doi:10.1002/jmor.20250.
- [89] Edwards MJ, Hinds LA, Deane EM, Deakin JE. A review of complementary mechanisms which protect the developing marsupial pouch young. *Dev Comp Immunol*. 2012; doi:10.1016/j.dci.2012.03.013.
- [90] Oikonomopoulou K, Ricklin D, Ward PA, Lambris JD. Interactions between coagulation and complement—their role in inflammation. *Seminars in Immunopathology*. 2012;34:151-65.
- [91] Ricklin D, Hajishengallis G, Yang K, Lambris JD. Complement: a key system for immune surveillance and homeostasis. *Nat Immunol*. 2010;11:785-97.
- [92] Xue L, Galdass M, Gnanapragasam MN, Manwani D, Bieker JJ. Extrinsic and intrinsic control by EKLf (KLF1) within a specialized erythroid niche. *Development*. 2014;141:2245-54.
- [93] Tallack MR, Magor GW, Dartigues B, Sun L, Huang S, Fittock JM, Fry SV, Glazov EA, Bailey TL, Perkins AC. Novel roles for KLF1 in erythropoiesis revealed by mRNA-seq. *Genome Res*. 2012; doi:10.1101/gr.135707.111.
- [94] Sánchez-Villagra MR, Gemballa S, Nummela S, Smith KK, Maier W. Ontogenetic and phylogenetic transformations of the ear ossicles in marsupial mammals. *J Morphol*. 2002; doi:10.1002/jmor.1085.

- [95] Anthwal N, Joshi L, Tucker AS. Evolution of the mammalian middle ear and jaw: adaptations and novel structures. *J Anat.* 2013; doi:10.1111/j.1469-7580.2012.01526.x.
- [96] Tucker AS. Major evolutionary transitions and innovations: the tympanic middle ear. *Philos Trans R Soc Lond B Biol Sci.* 2017;372:20150483.
- [97] Frenz DA, Liu W, Cvekl A, Xie Q, Wassef L, Quadro L, Niederreither K, Maconochie M, Shanske A. Retinoid signaling in inner ear development: A "Goldilocks" phenomenon. *Am J Med Genet A.* 2010; doi:10.1002/ajmg.a.33670.
- [98] Romand R, Dollé P, Hashino E. Retinoid signaling in inner ear development. *J Neurobiol.* 2006;66:687-704.
- [99] Raz Y, Kelley MW. Retinoic acid signaling is necessary for the development of the organ of Corti. *Dev Biol.* 1999;213:180-93.
- [100] Rubbini D, Robert-Moreno À, Hoijman E, Alsina B. Retinoic Acid Signaling Mediates Hair Cell Regeneration by Repressing p27kip and sox2 in Supporting Cells. *J Neurosci.* 2015; doi:10.1523/JNEUROSCI.1099-15.2015.
- [101] Zheng JL, Shou J, Guillemot F, Kageyama R, GaoWQ. Hes1 is a negative regulator of inner ear hair cell differentiation. *Development.* 2000;127:4551-4560.
- [102] Lu X, Borchers AG, Jolicoeur C, Rayburn H, Baker JC, Tessier-Lavigne M. PTK7/CCK-4 is a novel regulator of planar cell polarity in vertebrates. *Nature.* 2004;430:93-8.
- [103] Montcouquiol M, Rachel RA, Lanford PJ, Copeland NG, Jenkins NA, Kelley MW. Identification of Vangl2 and Scrb1 as planar polarity genes in mammals. *Nature* 2003; 423:173-7. In: Lhoumeau AC, Puppo F, Prébet T, Kodjabachian L, Borg JP. PTK7: a cell polarity receptor with multiple facets. *Cell Cycle.* 2011;10:1233-6.
- [104] Kelley MW, Driver EC, Puligilla C. Regulation of cell fate and patterning in the developing mammalian cochlea. *Curr Opin Otolaryngol Head Neck Surg* 2009;17:381-7.
- [105] Puligilla C, Kelley MW. Building the world's best hearing aid; regulation of cell fate in the cochlea. *Curr Opin Genet Dev.* 2009;19:368-73.
- [106] Morishita H, Eguchi S, Kimura H, Sasaki J, Sakamaki Y, Robinson ML, Sasaki T, Mizushima N. Deletion of autophagy-related 5 (Atg5) and Pik3c3 genes in the lens causes cataract independent of programmed organelle degradation. *J Biol Chem.* 2013; doi:10.1074/jbc.M112.437103.
- [107] Basu S, Rajakaruna S, Reyes B, Van Bockstaele E, Menko AS. Suppression of MAPK/JNK-MTORC1 signaling leads to premature loss of organelles and nuclei by autophagy during terminal differentiation of lens fiber cells. *Autophagy.* 2014; doi:10.4161/auto.28768.
- [108] Mansoor N, Mansoor T, Ahmed M. Eye pathologies in neonates. *Int J Ophthalmol.* 2016; doi:10.18240/ijo.2016.12.22.
- [109] Wistow G. The human crystallin gene families. *Hum Genomics.* 2012; doi:10.1186/1479-7364-6-26.
- [110] Kirshman LT, Kolandaivelu S, Frederick JM. The Leber congenital amaurosis protein, AIPL1, is needed for the viability and functioning of cone photoreceptor cells. *Hum Mol Genet.* 2010;19:1076-87.
- [111] Blankenship AG, Hamby AM, Firl A, Vyas S, Maxeiner S, Willecke K, Feller MB. The role of neuronal connexins 36 and 45 in shaping spontaneous firing patterns in the developing retina. *J Neurosci.* 2011;31:9998-10008.
- [112] Chhour KL, Hinds LA, Jacques NA, Deane EM. An observational study of the microbiome of the maternal pouch and saliva of the tammar wallaby, *Macropus eugenii*, and of the gastrointestinal tract of the pouch young. *Microbiology.* 2010; doi:10.1099/mic.0.031997-0.
- [113] Kwek J, De longh R, Nicholas K, Familiari M. Molecular insights into evolution of the vertebrate gut: focus on stomach and parietal cells in the marsupial, *Macropus eugenii*. *J Exp Zool B Mol Dev Evol.* 2009; doi:10.1002/jez.b.21227.
- [114] Menzies BR, Shaw G, Fletcher TP, Renfree MB. Early onset of ghrelin production in a marsupial. *Mol Cell Endocrinol.* 2009; doi:10.1016/j.mce.2008.10.033.
- [115] Sharp JA, Modèpalli V, Enjapoori AK, Bisana S, Abud HE, Lefevre C, Nicholas KR. Bioactive Functions of Milk Proteins: a Comparative Genomics Approach. *J Mammary Gland Biol Neoplasia.* 2014; doi:10.1007/s10911-015-9331-6.
- [116] Kwek JH, longh RD, Digby MR, Renfree MB, Nicholas KR, Familiari M. Cross-fostering of the tammar wallaby (*Macropus eugenii*) pouch young accelerates fore-stomach maturation. *Mech Dev.* 2009; doi:10.1016/j.mod.2009.01.003.
- [117] Schanbacher FL, Talhouk RS, Murray FA, Gherman LI, Willett LB. Milk-borne bioactive peptides. *Int Dairy J.* 1998;8:393-403.
- [118] Gawenis LR, Greeb JM, Prasad V, Grisham C, Sanford LP, Doetschman T, Andringa A, Miller ML, Shull GE. Impaired gastric acid secretion in mice with a targeted disruption of the NHE4 Na<sup>+</sup>/H<sup>+</sup> exchanger. *J Biol Chem.* 2005;280:12781-9.

- [119] Yue YG, Chen YQ, Zhang Y, Wang H, Qian YW, Arnold JS, Calley JN, Li SD, Perry WL 3rd, Zhang HY, Konrad RJ, Cao G. The acyl coenzymeA:monoacylglycerol acyltransferase 3 (MGAT3) gene is a pseudogene in mice but encodes a functional enzyme in rats. *Lipids*. 2011; doi:10.1007/s11745-011-3537-1.
- [120] Cheng D, Nelson TC, Chen J, Walker SG, Wardwell-Swanson J, Meegalla R, Taub R, Billheimer JT, Ramaker M, Feder JN. Identification of acyl coenzyme A:monoacylglycerol acyltransferase 3, an intestinal specific enzyme implicated in dietary fat absorption. *J Biol Chem*. 2003;278:13611-4.
- [121] Lefevre JG, Short KM, Lamberton TO, Michos O, Graf D, Smyth IM, Hamilton NA. Branching morphogenesis in the developing kidney is governed by rules that pattern the ureteric tree. *Development*. 2017; doi:10.1242/dev.153874.
- [122] Brophy P. The developing kidney: issues and opportunities. *Semin Fetal Neonatal Med*. 2017; doi:10.1016/j.siny.2017.01.003.
- [123] McMahon AP. Development of the Mammalian Kidney. *Curr Top Dev Biol*. 2016; doi:10.1016/bs.ctdb.2015.10.010.
- [124] Oxburgh L, Carroll TJ, Cleaver O, Gossett DR, Hoshizaki DK, Hubbell JA, Humphreys BD, Jain S, Jensen J, Kaplan DL, Kesselman C, Ketchum CJ, Little MH, McMahon AP, Shankland SJ, Spence JR, Valerius MT, Wertheim JA, Wessely O, Zheng Y, Drummond IA. (Re)Building a Kidney. *J Am Soc Nephrol*. 2017; doi:10.1681/ASN.2016101077.
- [125] Reggiani L, Raciti D, Airik R, Kispert A, Brändli AW. The prepattern transcription factor *Irx3* directs nephron segment identity. *Genes Dev*. 2007;21:2358-70.
- [126] Alarcón P, Rodríguez-Seguel E, Fernández-González A, Rubio R, Gómez-Skarmeta JL. A dual requirement for Iroquois genes during *Xenopus* kidney development. *Development*. 2008; doi:10.1242/dev.023697.
- [127] Heliot C, Desgrange A, Buisson I, Prunskaitė-Hyyryläinen R, Shan J, Vainio S, Umbhauer M, Cereghini S. HNF1B controls proximal-intermediate nephron segment identity in vertebrates by regulating Notch signalling components and *Irx1/2*. *Development*. 2013; doi:10.1242/dev.086538.
- [128] Darlington RB, Dunlop SA, Finlay BL. Neural development in metatherian and eutherian mammals: variation and constraint. *J Comp Neurol*. 1999;411:359-68.
- [129] Paolino A, Fenlon LR, Kozulin P, Richards LJ, Suárez R. Multiple events of gene manipulation via in pouch electroporation in a marsupial model of mammalian forebrain development. *J Neurosci Methods*. 2018; doi:10.1016/j.jneumeth.2017.09.004.
- [130] Liu Y, Helms AW, Johnson JE. Distinct activities of *Msx1* and *Msx3* in dorsal neural tube development. *Development*. 2004;131:1017-28.
- [131] Bach A, Lallemand Y, Nicola MA, Ramos C, Mathis L, Maufras M, Robert B. *Msx1* is required for dorsal diencephalon patterning. *Development*. 2003;130:4025-36.
- [132] Meyer RC, Giddens MM, Schaefer SA, Hall RA. GPR37 and GPR37L1 are receptors for the neuroprotective and glioprotective factors prosaptide and prosaposin. *Proc Natl Acad Sci U S A*. 2013;110:9529-34.
- [133] Smith NJ. Drug Discovery Opportunities at the Endothelin B Receptor-Related Orphan G Protein-Coupled Receptors, GPR37 and GPR37L1. *Front Pharmacol*. 2015; doi:10.3389/fphar.2015.00275.
- [134] Kelly EM, Marcot JD, Selwood L, Sears KE. The development of integration in marsupial and placental limbs. *Integrative Organismal Biology*. 2019; doi:10.1093/iob/oby013.
- [135] Kosaka N, Sakamoto H, Terada M, Ochiya T. Pleiotropic function of FGF-4: its role in development and stem cells. *Dev Dyn*. 2009;238:265-76.
- [136] Suzuki HR, Sakamoto H, Yoshida T, Sugimura T, Terada M, Solursh M. Localization of *Hst1* transcripts to the apical ectodermal ridge in the mouse embryo. *Dev Biol*. 1992;150:219-22.
- [137] Niswander L, Martin GR. FGF-4 and BMP-2 have opposite effects on limb growth. *Nature*. 1993;361:68-71.
- [138] Bensoussan-Trigano V, Lallemand Y, Saint Clément C, Robert B. *Msx1* and *Msx2* in limb mesenchyme modulate digit number and identity. *Dev Dyn*. 2011;240:1190-202.

## Tables

**Table 2. Enriched DE transcripts between P0 and P5.**

Tissue, system or organ	GO term	Gene name	Encoded protein - expression and function
<b>Immune system</b>	Antigen processing and presentation	CD8A	Mediation of efficient cell-cell interactions within the immune system
	Mucosal immune response	PIGR	Facilitation of the transcytosis of soluble polymeric isoforms of immunoglobulin A and immune complexes
	Thymic T cell selection; T cell selection	CD74	Association with MHC class II and regulation of antigen presentation for the immune response
<b>Blood tissue</b>	Erythrocyte differentiation	AHSP	Molecular chaperone that binds specifically to free alpha-globin and is involved in haemoglobin assembly
		DYRK3	Acts as a negative regulator of EPO-dependent erythropoiesis
		KLF1	Transcription regulator of erythrocyte development that probably serves as a general switch factor during erythropoiesis
<b>Hearing</b>	Cochlea development	HES1	Expressed in the organ of Corti during its terminal differentiation
		PTK7	Guides stereocilia orientation
		SOBP	Regulates the cellular fate and standardization of the organ of Corti
<b>Vision</b>	Visual perception	AIPL1	Expressed during rod and cone development
		CRYBA1, CRYBA4, CRYBB1, CRYBB2	Crystallins are required for the development and maintenance of the transparency of the eye lens
		GJC1	Expressed in retina bipolar cells and retinal ganglion cells during development
<b>Digestion</b>	Gastric acid secretion	SLC9A4	Required for normal levels of gastric acid secretion, secretory membrane development, parietal cell maturation and/or differentiation and, at least secondarily, for chief cell differentiation
		GAST	Gastrin is a hormone that stimulates the secretion of hydrochloric acid by the gastric mucosa, which results in gastrin formation inhibition. Also acts as a mitogenic factor for gastrointestinal epithelial cells
<b>Excretory system</b>	Metanephric nephron development	IRX1	Expressed and required during different stages of pronephros development. Additionally, it is crucial in the regionalization and patterning of tissues and organs during metazoan development
<b>Nervous system</b>	Midbrain development; regulation of astrocyte differentiation; regulation of glial cell differentiation	GPR37L1	Expressed in neurons and glia
	Midbrain development	MSX1	Related to the development of the midline structure of the forebrain, expressed in the spinal cord
<b>Limbs</b>	Embryonic hindlimb morphogenesis	FGF4	Expressed in the thoracic and pelvic limbs and the apical ectodermal ridge of the limb bud
		MSX1	Encodes transcription factors that are crucial for limb development

Table 3. Enriched DE transcripts between P5 and P10.

Tissue, system or organ	GO term	Gene name	Encoded protein - expression and function
Immune system	Complement activation	C1S	Serine protease that is a major constituent of human complement subcomponent C1
		C4BPA, C4BPB	Controls the classical pathway of complement activation
		C6	Component of the complement cascade
	Mucosal immune response	PLA2G1B	Secreted member of the phospholipase A2 (PLA2) class of enzymes produced by the pancreatic acinar cells
RAB17		The small Rab GTPases are key regulators of intracellular membrane trafficking from the formation of transport vesicles to their fusion with membranes	
Blood tissue	Blood coagulation	APOH	Binds to cardiolipin
		FGB, FGG	Plays a role in blood clotting and platelet aggregation
		HRG	Binds to numerous ligands and modulates immunity, vascularization and coagulation processes
		PDGFA	Regulation of embryonic development, cell proliferation, cell migration, survival and chemotaxis
		PROC, PROZ	Regulation of blood coagulation
	Fibrin clot formation	APOH	Binds to cardiolipin
		FGB, FGG	Important for blood clotting and platelet aggregation
	Complement and coagulation cascade	F7	Initiates the extrinsic pathway of blood coagulation
		F9	Participates in the intrinsic pathway of blood coagulation
		PLG	Plasminogen is activated by proteolysis and converted to plasmin and angiostatin
	Erythrocyte differentiation	AHSP	Prevents the harmful aggregation of alpha-haemoglobin during normal erythroid cell development
		KLF1	Transcription regulator of erythrocyte development that probably serves as a general switch factor during erythropoiesis
	Digestion	Fat digestion and absorption; glycerolipid metabolism	MOGAT3
Excretory system	Regulation of kidney development	FAT4	Plays a role in the orientation of cell divisions and tubule elongation during kidney development

## Supplemental File Legends

File name: Additional file 1

File format: .docx

**Additional File 1. Probability density values for the coverage identity of the *D. albiventris* vs *M. domestica* sequence alignment.** A coverage of 30% is more likely to be associated with higher identity values and a less likely to be associated with lower identity values. Min: Minimum. 1st Qu: first quantile. 3rd Qu: third quantile. Max: maximum.

File name: Additional file 2

File format: .docx

**Additional File 2. Probability density values for BLASTN analysis between *D. albiventris* and other marsupial species.** Probability density values between *D. albiventris* and other marsupial species (*M. domestica*, *S. harrisi*, *P. cinereus* and *M. eugenii*) determined with BLASTN. According to the analysis of the median and mean values, the coverage percentages for which there is a greater probability of presenting higher identity values and a lower tendency to present lower identity values (bold text) are as follows: 90% for *M. domestica* and *S. harrisi*; 20% for *P. cinereus* and *M. eugenii*. Min: Minimum. 1st Qu: first quantile. 3rd Qu: third quantile. Max: maximum.

File name: Additional file 3

File format: .xlsx

**Additional File 3. DIAMOND results for the assembled opossum transcripts.** The data are presented in seven sheets: all results for the three stages (“total”), all results for each pair of stages (“P0xP5”, “P0xP10” and “P5xP10”), differentially expressed results for each pair of stages (“P0xP5\_DE”, “P0xP10\_DE” and “P5xP10\_DE”), and differentially expressed genes exclusively expressed in each stage (“P0\_DE\_exclusive”; “P5\_DE\_exclusive”; “P10\_DE\_exclusive”).

File name: Additional file 4

File format: .tif

**Additional File 4. RT-qPCR validation of expression profiles for selected genes.** The RNA-seq expression results for four selected genes (HMGB3, CRYGB, APOH and HBZ) were confirmed by RT-qPCR. Relative expression was evaluated at the later stage in relation to the early stage (red line): expression at P5 in relation to P0 (A), expression at P10 in relation to P0 (B), and expression at P10 in relation to P5 (C). The expression values (C<sub>T</sub>) were normalized against the reference gene UBC. The boxed area in a whisker-box plot encompasses 50% of all observations, the dotted line represents the sample median, and the whiskers represent the outer 50% of observations. \*p<0.05. The results were analysed using REST 2009 software.

File name: Additional file 5

File format: .tif

**Additional File 5. P0, P5 and P10 hybridized tissues for titin.** Sagittal view of *D. albiventris* hybridized tissues for titin mRNA at P0 (A to D), P5 (E and F), and P10 (G to J). Diaphragm (A, E and H). Tongue (B and C). Ventral skeletal muscle (D). Longitudinal and transverse views of limb skeletal muscle (F). Sagittal view of the heart (G). Frontal view of the tongue and mandible skeletal muscle (I). Frontal view of the eye (J). ClueGO networks for overrepresented GO functional groups related to muscle based on the DE transcripts between P0 and P5 (K). *Asterisk*, Meckel’s cartilage; *at*, atrium; *ca*, cartilage; *d*, diaphragm; *em*, extrinsic muscle of eye; *it*, small intestine; *li*, liver; *ln*, eye lens; *lsm*, longitudinal view of skeletal muscle; *lu*, lungs; *mm*, mandible skeletal muscle; *pd*, periderm; *se*, heart septum; *sm*, skeletal muscle; *st*, stomach; *to*, tongue; *tsm*, transversal view of skeletal muscle; *ve*, ventricle. *Bar scale*, 50 µm (C, D); 150 µm (A, F, I); 300 µm (B, E, G, J); 650 µm (H).

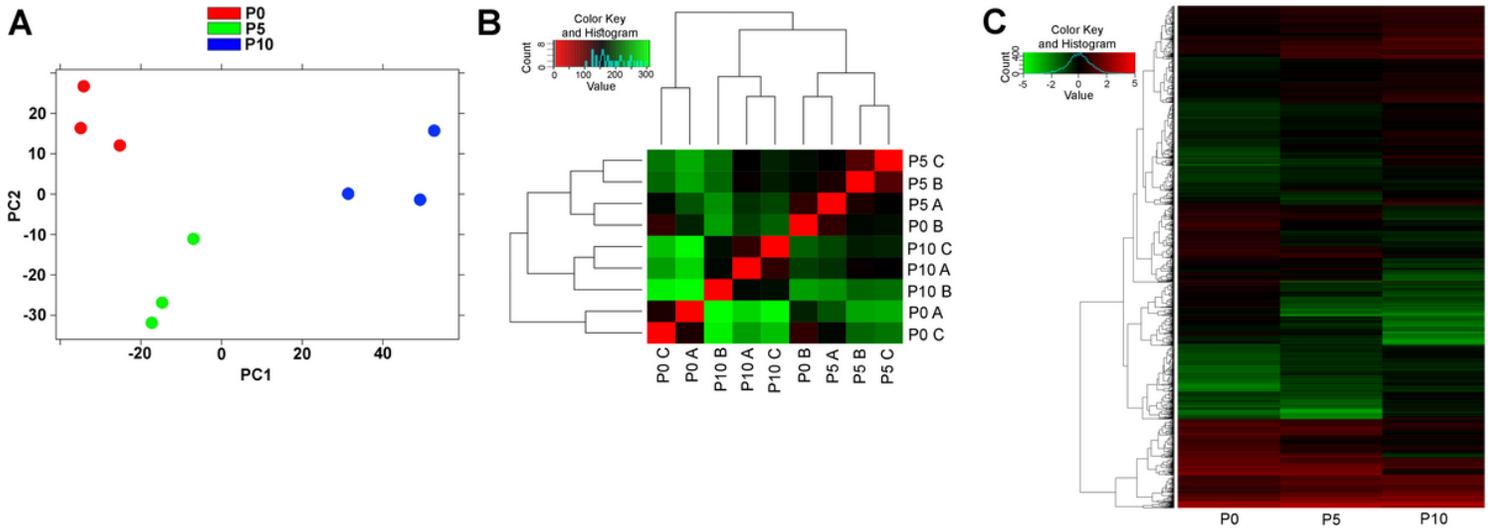
File name: Additional file 6

File format: .xlsx

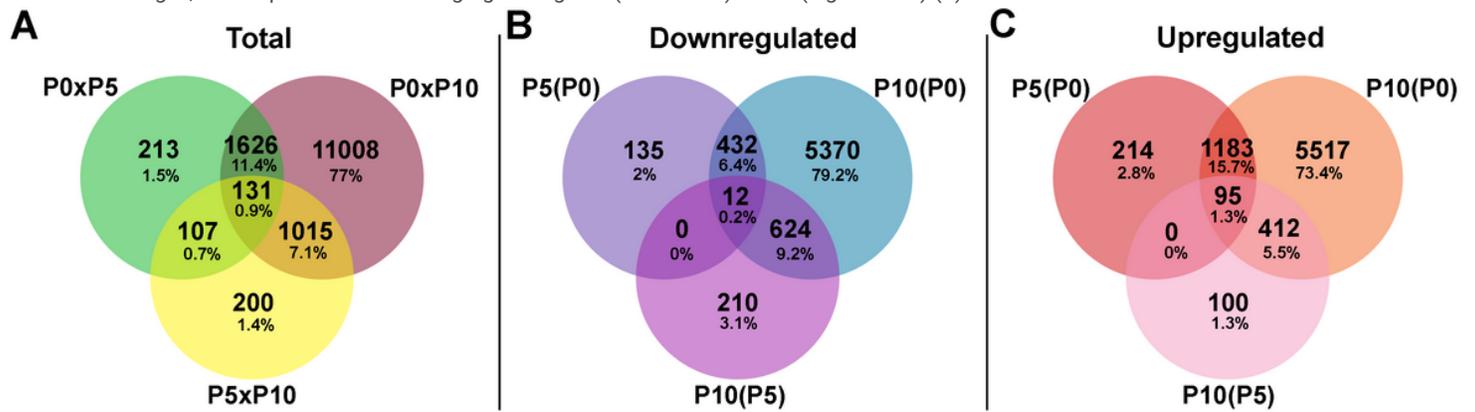
**Additional File 6. Enriched GO terms obtained through ClueGO analysis.** Gene names and GO terms of biological processes for enriched transcripts analysed with ClueGO.

The data are presented in eight sheets: all results for the three stages (“total”), differentially expressed results for the three stages (“total DE”), differentially expressed results for P0xP5 (“P0xP5 DE”), differentially expressed results for downregulated transcripts at P5 in relation to P0 [“P5(P0) DE down”], differentially expressed results for upregulated transcripts at P5 in relation to P0 [“P5(P0) DE up”], differentially expressed results for P5xP10 (“P5xP10 DE”), differentially expressed results for downregulated transcripts at P10 in relation to P5 [“P5xP10 DE down”], differentially expressed results for upregulated transcripts at P10 in relation to P5 [“P5xP10 DE up”].

## Figures



**Figure 1**  
 Bioinformatic analysis of RNA-seq data. Homogeneity of the biological triplicates determined by PCA (A) and the colourmap matrix (B) based on transcript expression values between P0, P5 and P10. In the colourmap matrix, the differences between the samples are indicated as follows: green dots represent the most different samples, and the red dots represent the least different samples (B). Colourmap matrix based on transcripts expressed at the three stages, with expression levels ranging from green (low values) to red (high values) (C).



**Figure 2**  
 Venn diagrams. Venn diagram for differentially expressed ( $p < 0.05$ ) transcripts between P0 and P5, P5 and P10, and P0 and P10 (A). Venn diagram for differentially expressed ( $p < 0.05$ ) transcripts showing downregulation [ $\log_2(\text{fold change}) < -1$ ; B] or upregulation [ $\log_2(\text{fold change}) > 1$ ; C]. The transcripts are described as exhibiting an increase or decrease in the expression level at the older stage compared to the younger stage. Images were generated using Venny 2.1.

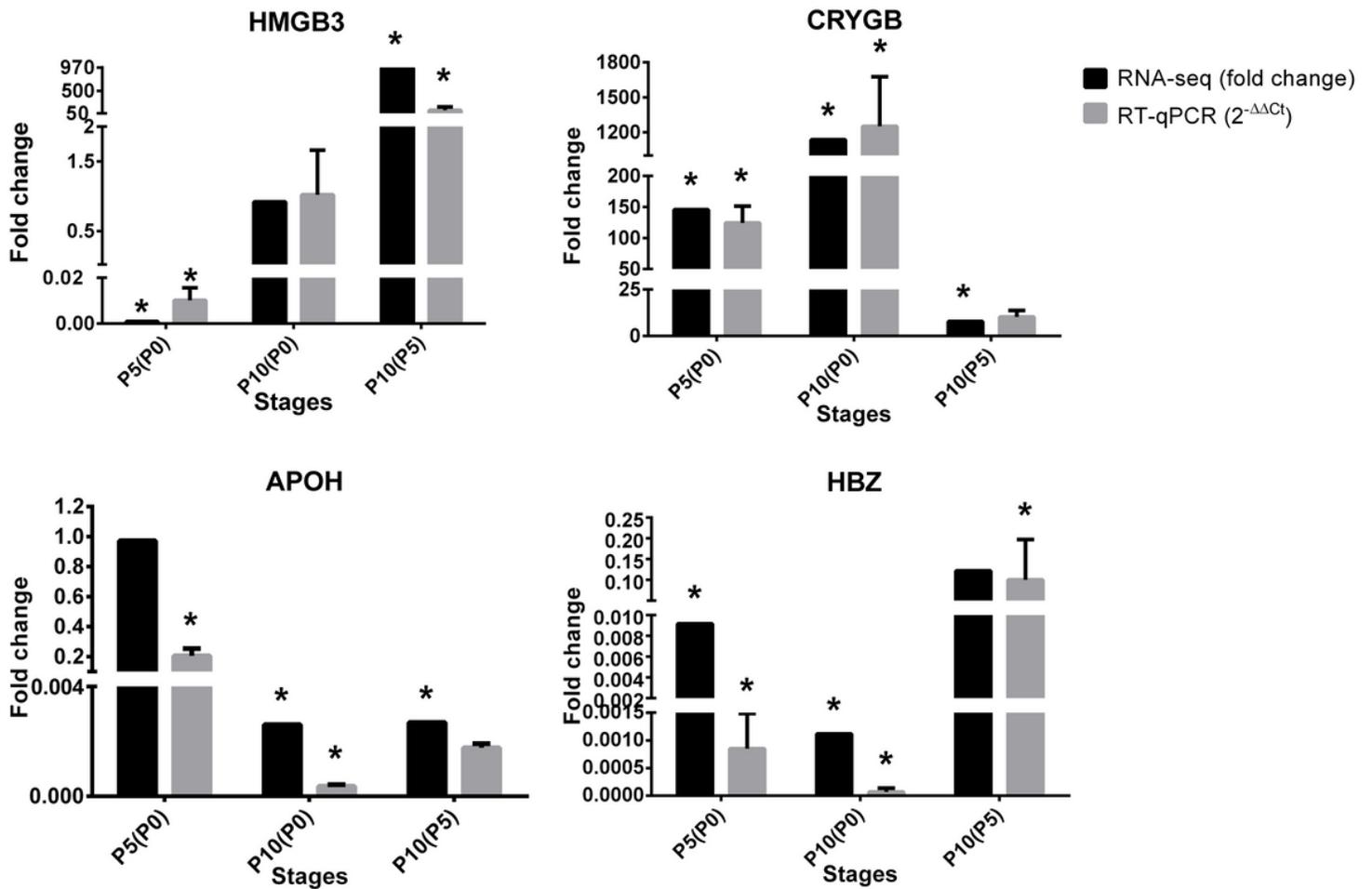
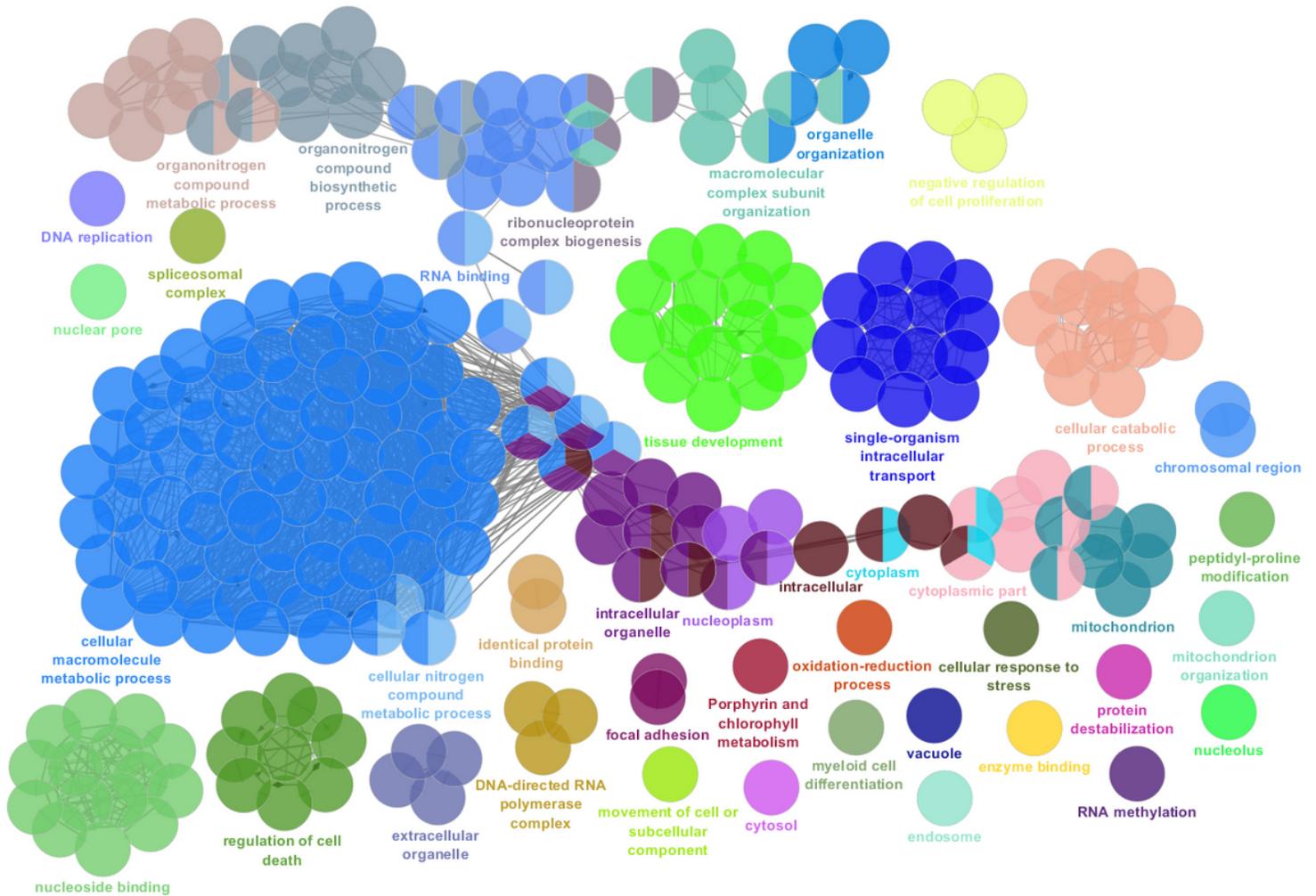


Figure 3

Validation of the RNA-seq expression profiles by RT-qPCR. To evaluate the validity of the Illumina sequencing results and confirm the observed patterns of expression, a subset of four differentially expressed genes among the most upregulated and downregulated genes between the stages was selected and evaluated by RT-qPCR analysis. The expression patterns indicate agreement between the two methods. Relative expression was evaluated between the developmental stages, and P0 was used as a calibrator. Asterisk,  $p < 0.05$ ; X-axis, developmental stage comparisons (P0 vs P5, P0 vs P10, P5 vs P10); Y-axis, fold change ( $2^{-\Delta\Delta Ct}$  values for RT-qPCR and fold change values for RNA-seq) in mRNA levels; bars, mean; error bars, standard deviation (SD).



**Figure 4**

Analysis of enriched GO functional groups of expressed genes between the developmental stages. Functionally grouped networks of enriched categories for expressed genes between the developmental stages, annotated for the cellular component, biological process, molecular function, and immune system process KEGG and GO terms. All genes were considered in this analysis, including differentially expressed and non-differentially expressed genes. Only the most significant term in the group is labelled. GO terms are represented as nodes. The edges connecting the nodes are based on the Kappa statistic (Kappa score threshold of 0.4), which measures the overlap of shared genes between terms. The right-sided hypergeometric test was used for statistical inference, and the Benjamini-Hochberg method was applied in p value correlation ( $p < 0.05$ ).

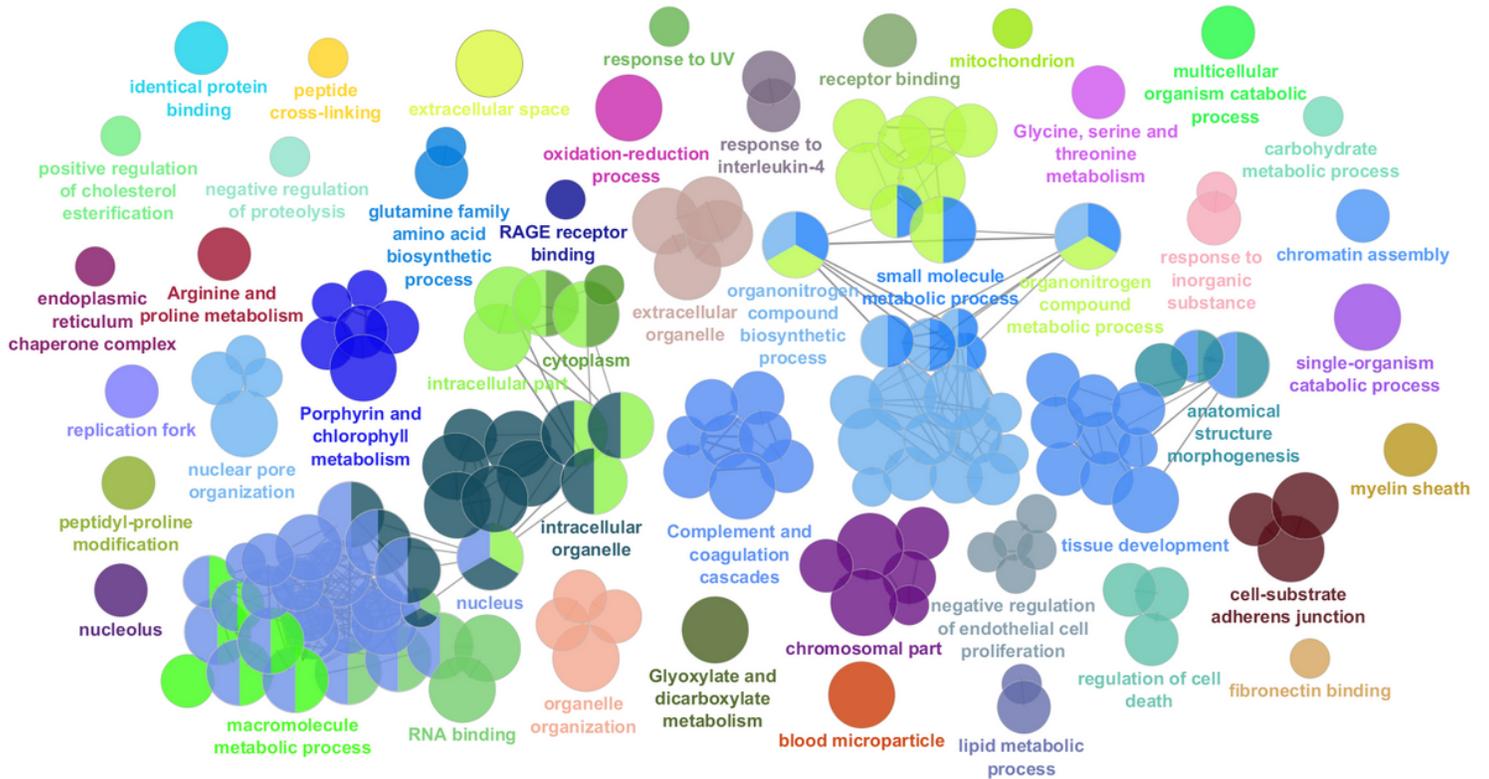


Figure 5

Analysis of enriched GO functional groups of differentially expressed genes between the developmental stages. Functionally grouped networks of enriched categories for differentially expressed genes between the developmental stages, annotated for the cellular component, biological process, molecular function, and immune system process KEGG and GO terms. Only the most significant term in the group is labelled. GO terms are represented as nodes, and the node size represents the term enrichment significance. The edges connecting the nodes are based on the Kappa statistic (Kappa score threshold of 0.4), which measures the overlap of shared genes between terms. The right-sided hypergeometric test was used for statistical inference, and the Benjamini-Hochberg method was applied for p value correlation ( $p < 0.05$ ).

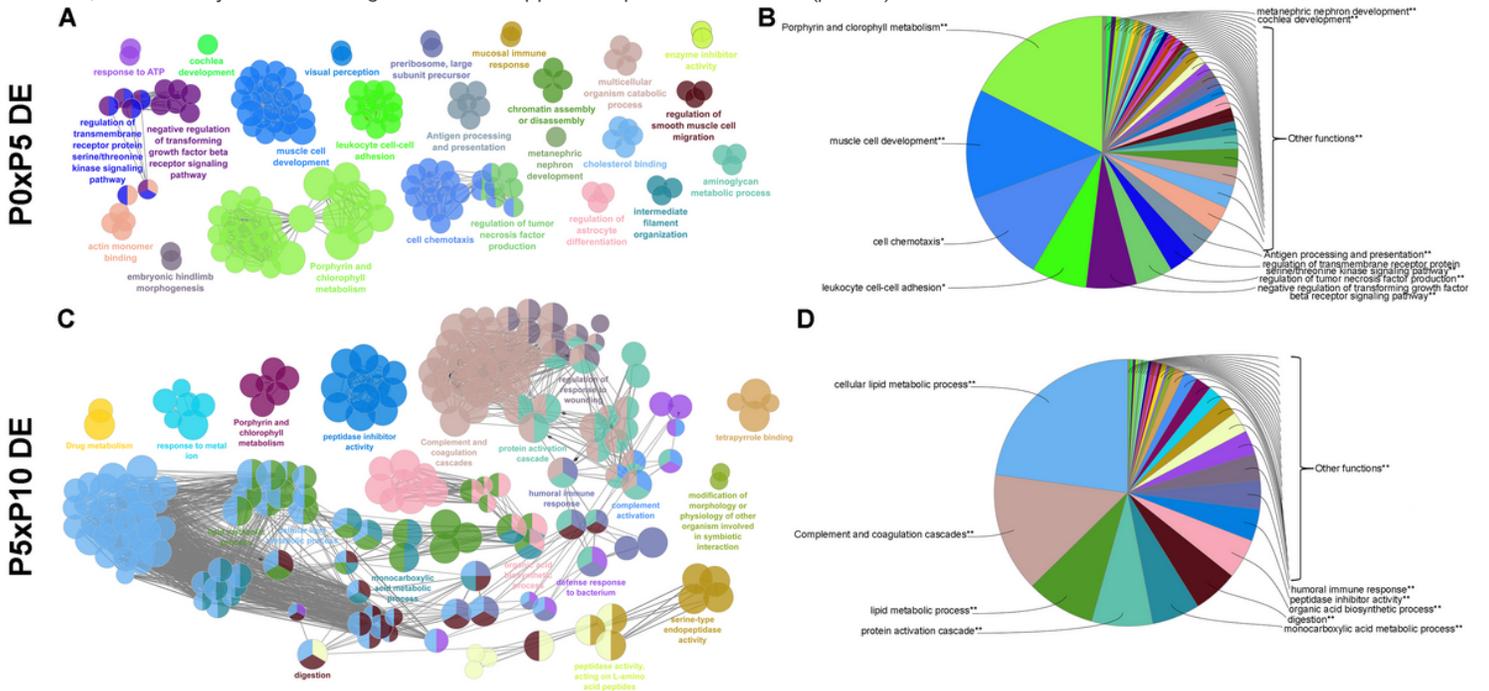
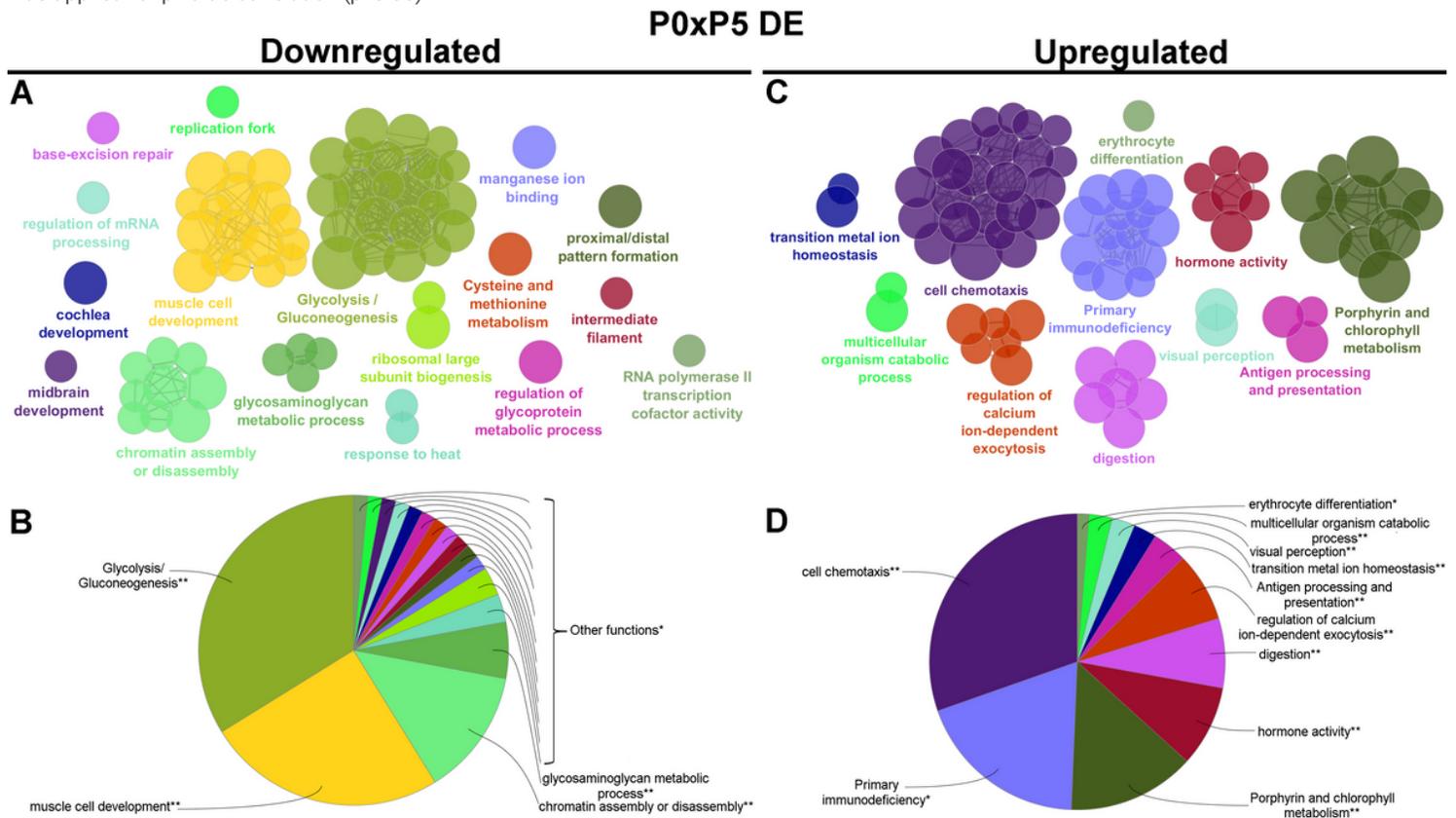


Figure 6

Analysis of enriched GO functional groups of DE genes at P0xP5 and P5xP10. Functionally grouped networks of enriched categories for DE genes between P0 and P5 (A) and between P5 and P10 (B), annotated for the cellular component, biological process, molecular function, and immune system

process KEGG and GO terms. Only the most significant term in the group is labelled. GO terms are represented as nodes, and the node size represents the term enrichment significance. The edges connecting the nodes are based on the Kappa statistic (Kappa score threshold of 0.4), which measures the overlap of shared genes between terms. The right-sided hypergeometric test was used for statistical inference, and the Benjamini-Hochberg method was applied for p value correlation ( $p < 0.05$ ).

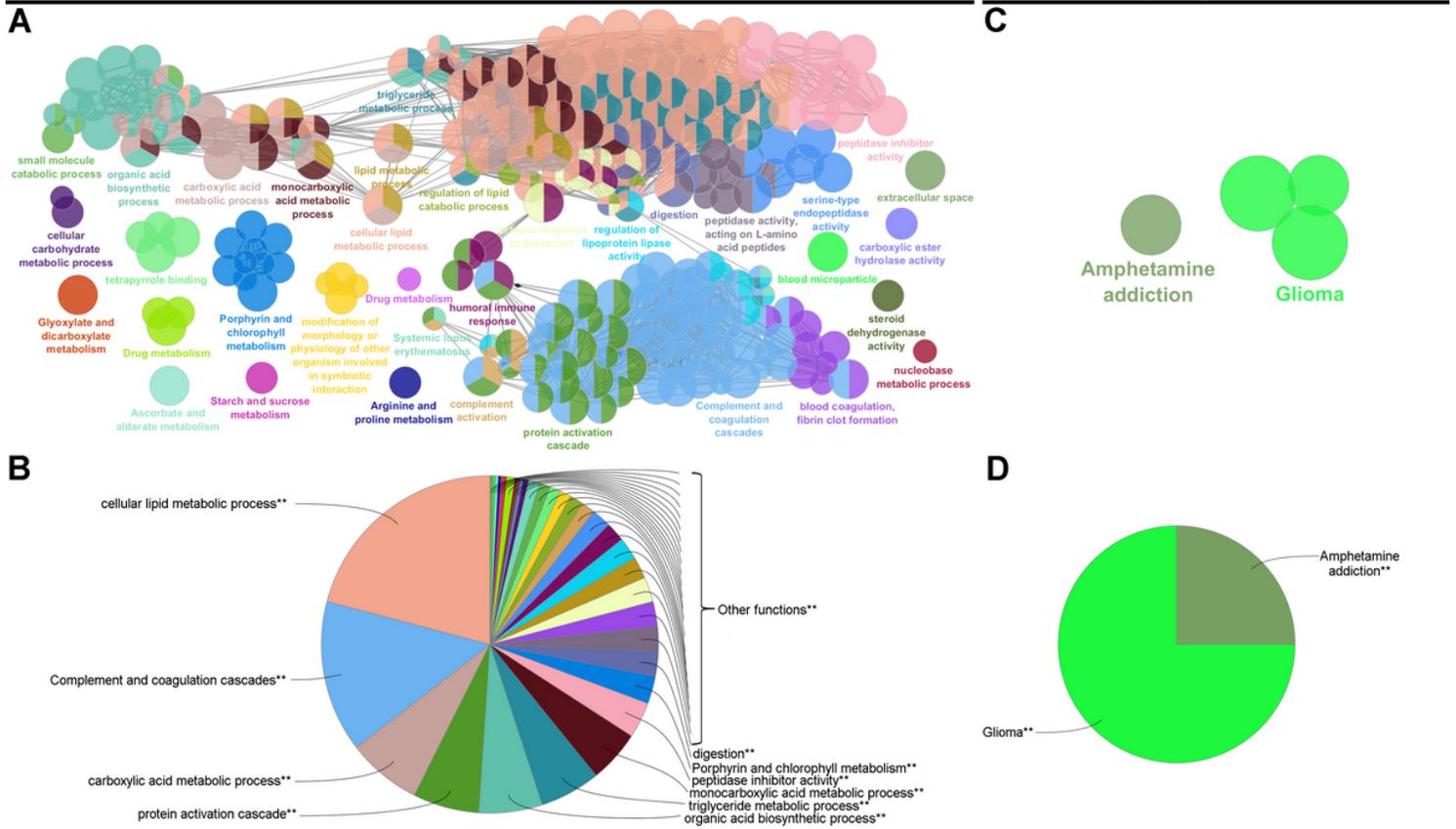


**Figure 7**  
 Analysis of enriched GO functional groups of downregulated and upregulated DE genes at P0xP5. A functionally grouped network of enriched categories was generated for the downregulated (A and B) and upregulated (C and D) differentially expressed genes at P5 in relation to P0. The most significant term in each functional group is represented. The cellular component, biological process, molecular function, and immune system process KEGG and GO terms were selected for this analysis. The right-sided hypergeometric test was used for statistical inference, and the Benjamini-Hochberg method was applied for p value correlation ( $p < 0.05$ ).

## P5xP10 DE

### Downregulated

### Upregulated



**Figure 8**  
 Analysis of enriched GO functional groups of downregulated and upregulated DE genes at P5xP10. A functionally grouped network of enriched categories was generated for the downregulated (A and B) and upregulated (C and D) differentially expressed genes at P10 in relation to P5. The most significant term in each functional group is represented. The cellular component, biological process, molecular function, and immune system process KEGG and GO terms were selected for this analysis. The right-sided hypergeometric test was used for statistical inference, and the Benjamini-Hochberg method was applied for p value correlation ( $p < 0.05$ ).

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile2.docx](#)
- [Additionalfile1.docx](#)
- [Additionalfile6.xlsx](#)
- [Additionalfile4.tif](#)
- [Additionalfile3.xlsx](#)
- [Additionalfile5.tif](#)
- [ChecklistS12ndsubmission.pdf](#)