Comparative transcriptome analysis revealed omnivorous adaptation of the small intestine of Melinae

Lidong wang
Qufu Normal University
Honghai zhang (zhanghonghai67@126.com)
Qufu Normal University

Research Article

Keywords: RNA-seq, Mink, Tyrosinase, Fur, Coat color

Posted Date: April 14th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-598218/v1

Abstract

Background

Mink fur is an important economic product, but the molecular mechanism of its color formation and fur development is still unclear. We used RNA-seq to analyze the skin transcriptome of young and adult minks with two different hair colors.

Result

The minks were made up of black adults (AB), white adults (AW), black juveniles (TB), and white juveniles (TW) (three each). We found that thirteen KRTAP genes and five signaling pathways (JAK-STAT signaling Pathway (cfa04630), Signaling pathways regulating pluripotency of stem cells (cfa04550), ECM-receptor interaction (cfa04512), Focal adhesion (cfa04510) and Ras signaling pathway (cfa04014)) were related to mink fur development through pair comparison and cross-screening among different subgroups. We also found a tyrosinase family (TYR, TYRP1, TYRP2) involved in mink hair color formation. The expression level of TYR in black young minks was higher than that in white young minks, but this phenomenon was not observed in adult minks.

Conclusion

Our study found significant differences in adult and juvenile mink skin transcriptome, which may shed light on the mechanism of mink fur development. At the same time, the skin transcriptome of black and white minks was also different, and the comparison was different by age, suggesting that the genes regulating hair color function in early childhood rather than adulthood. This study can provide molecular support for breeding mink coat color and improving fur quality.

Introduction

American mink (*Neovison vison*) is one of the major farmed species of Mustelidae and one of the world's major suppliers of quality fur[1]. Minks are semi-aquatic, so their coats are thick and waterproof to protect them from the cold and move better in the water[2]. In the early 19th century, wild mink populations declined rapidly due to the economic value of their pelts. So during this time, the first mink farms were established by T.D. Phillips and W.Woodcock[3]. Farmed minks were originally standard dark brown, but with genetic mutations and advances in farming techniques, more than 100 color types are available. Short-haired black mink and red-eyed white mink are the most common breeds, as black can be worn with almost any color of clothing and white can be dyed in any color.

The characters displayed by species are divided into qualitative character and quantitative character, and coat color belongs to the qualitative character. Qualitative character refers to the non-continuous quantitative change between different phenotypes of the same trait, which is dominated by a few decisive

genes[4]. Seven mechanisms for determining mink color have been revealed in the current study, including Aleutian[5], Albino[6], Palomino[7], Silverblue, and Hedlund White[1], Moyle Brown[8], Himalayan[9]. The genes identified are LYST (Lysosomal trafficking regulator), TYR (Tyrosinase, Albino, and Himalayan), TYRP1 (Tyrosinase-related protein 1), MLPH (Melanophilin), and MITF (Microphthalmia-associated transcription factor), RAB38 (Ras-related protein-38). However, these studies only focus on the changes in a gene sequence, and little attention has been paid to the changes in gene expression.

In addition, the economic value of mink fur depends not only on the color but also on the overall quality (length, density) of the fur. These characteristics associated with hair follicle development can be called quantitative characters. The development and growth of mink hair cannot be studied independently of the hair follicle[10]. Hair follicles undergo differentiation, proliferation, and apoptosis periodically and are regulated by a variety of signaling factors[11]. Most of these key signaling molecules belong to Wnt (Wingless-related), Shh (Sonic Hedgehog), BMP (Bone morphogenetic protein), FGFs (Fibroblast growth factors), TGF(Transforming growth factor), and Notch signaling pathway[12]. By comparing and analyzing the skin transcriptome of minks during development and maturity, we can reveal the key genes and pathways during the skin development of minks.

Therefore, in this study, to reveal the mechanisms of fur development and color determination in short-haired black minks and red-eyed white minks, we conducted transcriptome sequencing of skin tissues from adult and juvenile individuals of these two coat colors. The PCA analysis, differential expression analysis, and enrichment analysis were used in this study. Through the comparison between different groups, the comparison between the adult group and the juvenile group focused on discovering the genes and pathways related to the development of mink fur, and the comparison between the black group and the white group explored the determining mechanism of coat color. This study provides a more comprehensive overview of key genes and pathways that influence the economic value of mink fur. It provides the foundation for mink breeding and new color type cultivation.

Materials And Methods

Sample collection and preparation

Twelve healthy young (ten-day-old) white minks, adult white minks, young (ten-day-old) black minks, and adult black minks (3 of each type) were selected as samples from a mink farm in Shandong province. All sample procedures and experimental methods were approved by the Qufu Normal University Institutional Animal Care and Use Committee (No. 2022033), Qufu, China. Two pieces of skin (1.0 cm in diameter) from the back were collected via a punch skin biopsy under local anesthesia and were immediately placed in liquid nitrogen. Young white skin samples were numbered TW1, TW2, and TW3. Adult white skin samples were numbered AW1, AW2, and AW3. The numbers of young black skin samples were TB1, TB2, and TB3. Adult black skin samples were numbered AB1, AB2, and AB3.

RNA quantification and qualification

Total RNA was extracted from the 12 skin samples using the RNeasy Mini Kit (QIAGEN, USA). Total amounts and integrity of RNA were assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA).

Library preparation for Transcriptome sequencing

Total RNA was used as input material for the RNA sample preparations. In simple terms, the mRNA is purified from the total RNA using poly-T oligo-attached. Divalent cations are fractured at elevated temperatures in a first-chain synthetic reaction buffer (5X). The first-strand cDNA was synthesized using random hexamer primers and M-Mulv reverse transcriptase, and the RNA was degraded using RNaseH. The second strand of cDNA was then synthesized using DNA polymerase I and dNTP. The rest of the overhang is converted to blunt ends by exonuclease/polymerase activity. After adenylation of the 3 '-end of the DNA fragment, the Adaptor with hairpin ring structure is attached and prepared for hybridization. AMPure XP system of Beckman Coulter, Beverly, USA was used to purify the library fragments and selected the 370–420 bp cDNA fragments. PCR amplification was performed, the PCR product was purified with AMPure XP microbeads, and finally, the library was obtained.

To ensure the quality of the library, it is necessary to test the library. After the library was constructed, the library was quantified by a Qubit2.0 fluorescence analyzer, then diluted to 1.5ng/ul, and the library insertion size was detected by Agilent 2100 bioanalyzer. After the size of the insert reached the expected size, qRT-PCR was used to accurately quantify the effective concentration of the library (the effective concentration of the library was higher than 2nM) to ensure the quality of the library.

Clustering and sequencing

After the libraries were qualified, the different libraries were collected according to the effective concentration and the target data volume output by the machine, and then sequenced by Illumina NovaSeq 6000. Generates an end reading of the 150bp pair. The basic principle of sorting is that composition and sort occur simultaneously (composition sort). Four fluorescent-labeled dNTP, DNA polymerase, and splicing primers were added to the sequenced flow cells for amplification. When the sequence cluster extends the complimentary chain, each fluorescently labeled dNTP can release corresponding fluorescence. The sequencer captures the fluorescence signal and converts the optical signal into a sequencing peak through computer software to obtain the sequence information of the fragment to be measured.

Quality control

Image data measured by a high-throughput sequencer is converted into sequence data (read) by CASAVA base recognition. Raw data in FastQ format (raw reads) is first processed by an internal Perl script. In this step, clean reads are obtained by removing reads containing adapters, reads containing N bases, and reads of low quality from the raw data. Meanwhile, Q20, Q30, and GC contents of clean data were calculated. All downstream analyses are based on high-quality clean data.

Reads mapping to the reference genome

The reference genome comes from unpublished data from the Kunming Institute of Zoology, Chinese Academy of Sciences. The reference genome index was constructed using Hisat2 (V2.0.5), and the paired clean reads were compared with the reference genome using Hisat2 (V2.0.5)[13]. We choose Hisat2 as the mapping tool because Hisat2 can generate a splicing connection database from a gene model annotation file, its mapping results are superior to other non-splicing mapping tools.

Novel transcripts prediction

The mapped reads of each sample were assembled by StringTie[14](v1.3.3b) in a reference-based approach. StringTie uses a novel network flow algorithm with optional ab initio assembly steps to assemble and quantify full-length transcripts representing multiple spliced variants of each gene loci.

Quantification of gene expression level

The featureCounts v1.5.0-p3 was used to count the read numbers mapped to each gene[15]. The FPKM value of each gene was then determined based on the length of the gene and the read count mapped to the gene[16]. FPKM, the expected number of fragments per thousand base pairs of transcriptional sequence, is the most commonly used method to estimate gene expression level, taking into account the effects of sequencing depth and gene length on the read count.

Differential expression analysis

Differential expression analysis of two groups was performed using the DESeq2 R package (1.20.0)[17]. DESeq2 provides arithmetical procedures for determining differential expression in digital gene expression data using a model based on the negative binomial distribution. To control the error detection rate, we use Benjamini and Hochberg's method to adjust the p-value obtained. Padj < = 0.05 and |log2(foldchange)| >= 1 were set as the threshold for significantly differential expression.

GO and KEGG enrichment analysis of differentially expressed genes

The gene ontology (GO) enrichment analysis of differentially expressed genes was performed by clustering map R package (3.8.1), which corrected the gene length bias. We believe that GO items with a p-value less than 0.05 are significantly enriched in differentially expressed genes. KEGG is a database resource for understanding the advanced functions and uses of biological systems, such as cells, organisms, and ecosystems, from large-scale molecular datasets generated from molecular-level information, particularly genome sequencing and other high-throughput experimental techniques (http://www.genome.jp/kegg/)[18]. We used Cluster Profile R Package (3.8.1) to detect statistical enrichment of differentially expressed genes in the KEGG pathway.

Result

Data quality control and novel gene prediction

After filtering out low-quality reads and removing the adaptor sequences, we finally obtained 80.71Gb of clean data (Table 1). The effective average GC ratio was 51.9%. Quality value Q30 means that the base recognition accuracy rate is 99.9%, which is above 93.57%, and that the ratio of each sample aligned to the reference genome was 94.69–96.58% (Table S1). Additionally, most reads were mapped to the EXON region (Table S2), which overall indicates good sequencing data and high utilization. We predicted and assembled 1,943 new genes using software and annotated them functionally using Pfam (Table S3). These data were used for the analysis of gene expression.

Table 1
Sequencing data statistics

Sample	Library	raw_reads	clean_reads	clean_bases	error_rate	Q20	Q30	GC_pct
AB1	AB1	45947094	45190978	6.78G	0.03	97.7	93.57	52.55
AB2	AB2	43685582	42895824	6.43G	0.03	97.83	93.81	52.43
AB3	AB3	46277030	45357284	6.8G	0.03	97.82	93.83	52.28
AW1	AW1	44938430	44155634	6.62G	0.03	97.73	93.69	52.94
AW2	AW2	48298078	47408254	7.11G	0.03	97.88	93.94	51.17
AW3	AW3	46758560	46088186	6.91G	0.03	97.84	93.87	51.52
TB1	TB1	43634474	42882980	6.43G	0.03	97.76	93.71	52.22
TB2	TB2	47393846	46084054	6.91G	0.02	98.24	94.89	51.14
TB3	TB3	46278052	44685668	6.7G	0.02	98.18	94.64	49.18
TW1	TW1	44342806	43591538	6.54G	0.03	97.68	93.56	52.31
TW2	TW2	44828808	43975908	6.6G	0.03	97.71	93.62	52.77
TW3	TW3	46740826	45858686	6.88G	0.03	97.75	93.72	52.91

Q20, Q30:Represent the base recognition error rate of 1% and 0.1% in the sequencing process.

Quantitative analysis of gene expression

The gene expression level of each sample was quantitatively analyzed and then combined to obtain the expression matrix of all samples (Table S4). After the expression values of all genes in each sample (FPKM) were calculated, the distribution of gene expression levels in different samples was displayed by a box graph (Fig. 1A). The correlation of gene expression levels between samples is an important index to test whether the experiment is reliable and the sample selection is reasonable. The heat map of correlation between samples shows that age factors divide them into two groups and Pearson correlation coefficient R^2 within the group is greater than 0.8(Fig. 1B). PCA analysis of gene expression values (FPKM) in all samples was performed. The results showed that they were divided into adult and juvenile groups at the PC1 level, and further grouped by coat color at the PC2 level (Fig. 1C). This shows the high reliability of our sample collection and grouping.

Differentially expressed genes (DEGs) in American mink skin

To reveal the genes differentially expressed in the fur development group(AB vs TB, AW vs TW) and coat color group(AB vs AW, TB vs TW) of American mink, four combinations were compared and analyzed(Table 2, Fig. 2). In terms of fur development, there were 6453 and 5689 differentially expressed genes in white and black groups, respectively. In the coat color difference combination, there were 116 differentially expressed genes in the adult group and 257 differentially expressed genes in the juvenile group. The differential genes in all comparison groups were combined as differential gene sets, cluster analysis was carried out on the differential gene sets, and genes with similar expression patterns were clustered together (Fig. 3, Fig S1). In addition, we plotted differential gene Venn plots from pairwise combinations of the coat color group and the fur development group (Fig. 4a,b). There were only four differentially expressed genes shared by adults with different coat colors and juveniles with different coat colors. Unfortunately, the two non-coding RNAs and SCD5 (Stearoyl-CoA Desaturase 5) and RSPH4A (Radial Spoke Head Component 4A) were not associated with the coat color difference. There were 4172 differentially expressed genes in the white group and black group of different ages. These differentially expressed genes may be related to the fur development of minks.

Table 2
Statistics of differentially expressed genes in different combinations

Group	ALL	UP	Down
AWvsTW	6453	2979	3474
ABvsTB	5689	2794	2895
AWvsAB	116	18	98
TWvsTB	257	132	125

Analysis of GO and KEGG pathways

To further investigate the biological relevance of all DEGs, we performed a GO analysis of DEGs to identify the enrichment of biological processes in each group. To make the enrichment results more accurate, we took the intersecting part of the two significant results of the fur development group (AW vs TW and AB vs TB). Significantly enriched to obtain 73 terms, including 18 biological processes (BP), 17 cellular components (CC), and 38 molecular functions (MF) (Table S5). Interestingly, the terms of keratin filaments (GO:0045095) related to fur development are also included. Finally, 13 genes including KRTAP1-3, KRTAP2-3, KRTAP3-3, KRTAP4-2, KRTAP4-5, KRTAP4-11, KRTAP4-12, KRTAP9-2, KRTAP10-7, KRTAP10-8, KRTAP10-12, KRTAP12-1, and KRTAP29-1 were identified, and all of them were down-regulated genes. Similarly, the enrichment results of the coat color difference group (AW vs AB and TW vs TB) were intersected. There are only three GO terms: structural molecule activity (GO:0005198), sulfotransferase activity (GO:0008146), and transferase activity (GO:0016782).

KEGG annotation analysis results of differentially expressed genes between groups are shown in Fig. 5. Similar to the above analysis, we obtained the metabolic pathways of fur development and coat color difference respectively. There were 35 cross-enriched related metabolic pathways in different age combinations of the same coat color, among which the JAK-STAT signaling Pathway (cfa04630), Signaling pathways regulating pluripotency of stem cells (cfa04550), ECM-receptor interaction (cfa04512), Focal adhesion (cfa04510) and Ras signaling pathway (cfa04014) were related to fur development (Table S6). Interestingly, only one metabolic pathway was found to be Tyrosine metabolism (cfa00350) by the intersection of enrichment results of different coat colors at the same age. There are five differentially expressed genes in this pathway, including DCT, TYR, AOX4, GSTZ1, and ADH5.

Discussion

Mink is one of the most important fur animals in the world. Previous studies only focused on the diversity and formation mechanism of fur color, but in this study, we not only explored the formation mechanism of fur color but also focused on key genes and pathways in the process of fur development. Three young and three adults of the short-haired black mink breed and three young and three adults of red-eyed white mink were selected as our experimental samples. A comparison of skin transcriptome between different groups was performed to achieve our research objectives. Animal coat color is a qualitative character, which is determined by the content of eumelanin and pheomelanin in the hair[19]. In IFPCS (International Federation of Pigment Cell Societies), 661 genes related to coat color have been published so far[20]. The research on fur development mainly focuses on the molecular mechanism of hair follicle development, including keratin[21, 22], keratin-associated proteins[23], signal pathway[24], gene family[25], and growth factor[26, 27].

Transcriptome sequencing of 12 mink skin samples yielded a total of 538 million clean reads. Correlation analysis between the samples showed that the samples could be divided into two components by age, indicating that the difference in gene expression caused by different coat colors was very small. Further principal component analysis (PCA) on the expression levels of samples can find that there are differences in samples of a different colors. This shows that our experimental design is reasonable. Differential gene expression analysis showed that the differentially expressed genes shared by adult black and white minks (AB vs AW) and juvenile black and white minks (TB vs TW) were not related to coat color formation. At the same time, adult black and juvenile black minks (AB vs TB) shared too many differentially expressed genes (4172) with adult white and juvenile white minks (AW vs TW). The genes related to fur development could not be identified.

Therefore, we searched for genes related to hair color formation and fur development through GO enrichment and KEGG enrichment analysis. The analysis showed that differentially expressed genes were enriched in 73 GO terms, and the most significant term we looked at about fur development was Keratin Filaments (GO:0045095). The keratin filaments are intervening filaments made of keratin that are found in various epithelial cells[28]. These GO terms identified thirteen keratin-associated protein genes (KRTAP), all of which were down-regulated in adult minks. The KRT is the main protein that makes up the outer layers of hair and skin[29, 30]. Mutations in the KRT gene have been found to cause changes in hair shape in

dogs[31], cats[32], and mice[33], resulting in curly hair. It has also been found that mutations in the KRT gene cause hereditary skin diseases in humans, such as ichthyosis, congenital tachycardia, and palmoplantar epidermolytic keratosis[34, 35]. KRTAP can be divided into homocysteine (HS-KAPS) and high glycine tyrosine (HGT-KAPS) according to the composition of amino acids[36, 37]. The thirteen KRTAP genomes we identified were HS-KAPS. It has been reported that the KRTAP gene is related to wool weight, strength, diameter, elongation, and other traits[38–41]. In addition, different numbers of cysteine-containing repeats in human KRTAP1 and KRTAP2 genes lead to length polymorphism, which changes the interaction between KRT and KRTAP and results in differences in hair traits among individuals[42]. From this, we can infer that these thirteen KRTAP genes play an important role in the development of mink fur.

KEGG enrichment analysis also found five developmental-related signaling pathways. In mammals, the JAK-STAT pathway is the main signal transduction mechanism for a variety of cytokines and growth factors[43]. Pluripotent stem cells (PSCs) are basic cells with an indefinite self-renewal capacity, which is associated with the TGF[44], and BMP[45] signaling pathways. Extracellular matrix (ECM) is a mixture of macromolecules with complex structure and function, which plays an important role in the morphogenesis of tissues and organs and the maintenance of cell and tissue structure and function[46]. Adhesion molecules, while catalytically inactive, can do this by binding to growth factor receptors, which can initiate, integrate, or feedback adhesion-based signals[47]. Ras proteins are GTPases, which act as molecular switches in signaling pathways that regulate cell proliferation, survival, growth, migration, differentiation, or cytoskeletal activity[48]. In conclusion, we speculate that these signaling pathways may be related to mink fur development.

In addition, the Tyrosine metabolism pathway in KEGG enrichment was related to hair color. TYR and DCT in genes enriched by this pathway are related to melanogenesis. The tyrosinase gene family that regulates melanin production mainly consists of TYR (Tyrosinase), TYRP1 (Tyrosinase Related Protein 1), and TYRP2 (Dopachrome Tautomerase, DCT)[49]. Tyrosinase is a rate-limiting enzyme in the process of melanin synthesis. The expression level and activity of TYR directly affect the expression of eumelanin and melanin and then affect the animal hair color[50]. TYRP1 and TYRP2 play a catalytic role in the last steps that control the type of melanin melanocytes produce[51]. In our results, the expression level of TYR in black young minks was higher than that in white young minks, but this phenomenon was not observed in adult minks. Interestingly, TYRP1 and TYRP2 were more expressed in white young minks than in black minks. In adult minks, only TYRP2 was highly expressed in black minks, while there was no significant difference in the expression of TYR and TYRP1. Therefore, it can be inferred that the regulatory genes of mink hair color have been playing a role in early childhood, but the difference in gene expression level in adults is not significant. In conclusion, the tyrosinase gene family plays an important role in the regulation of mink hair color.

Conclusions

In summary, we have identified differences in skin transcriptome between juvenile and adult minks and between black and white minks. In this study, skin transcriptomes of four groups of minks (AB, AW, TB, TW) were sequenced, and it was found that thirteen KRTAP genes and five signaling pathways were involved in

the fur development of minks. Meanwhile, TYR, TYRP1, and TYRP2 are closely related to the formation of mink fur color. The results of this study can provide a reference for studying the internal molecular mechanism of mink fur development and color formation.

Declarations

Acknowledgments

We thank the members of our lab and staff on the farm for their assistance in the study.

Author contributions

G.L and S.Z. carried out the experimental work and the data collection and interpretation. T.L. and L.S. participated in the design and coordination of experimental work, and the acquisition of data. Y.D and H.Z. participated in the study design, data collection, analysis of data, and preparation of the manuscript. L.W carried out the study design, the analysis, and the interpretation of data and drafted the manuscript. All authors read and approved the final manuscript.

Funding

Funding for the project was from the National Natural Science Fund of China (Grant numbers 31872242, 32070405, 31900311).

Availability of data and materials

The datasets generated during the current study are available in the NCBI Sequence Read Archive (SRA): PRJNA825111

Competing interests

The authors declare no conflict of interest.

Ethics approval and consent to participate

All sample procedures and experimental methods were approved by the Qufu Normal University Institutional Animal Care and Use Committee (No. 2022033), Qufu, China.

Consent for publication.

Not applicable.

Competing interests

The authors declare no conflict of interest.

References

1. Manakhov AD, Andreeva TV, Trapezov OV, Kolchanov NA, Rogaev EI: **Genome analysis identifies the mutant genes for common industrial Silverblue and Hedlund white coat colours in American mink**. Sci Rep 2019, **9**(1):4581.
2. Winfield IJ: **Vertebrates: Fish, Amphibians, Reptiles, Birds, Mammals**. *Elsevier* 2009.
3. Joergensen G: **Mink production**: Mink production.; 1985.
4. Falconer DS, Mackay TF: **Introduction to quantitative genetics**: Longman; 1960.
5. Anistoroaei R, Krogh AK, Christensen K: **A frameshift mutation in the LYST gene is responsible for the Aleutian color and the associated Chediak-Higashi syndrome in American mink**. Anim Genet 2013, **44**(2):178–183.
6. Anistoroaei R, Fredholm M, Christensen K, Leeb T: **Albinism in the American mink (Neovison vison) is associated with a tyrosinase nonsense mutation**. Anim Genet 2008, **39**(6):645–648.
7. Cirera S, Markakis MN, Kristiansen T, Vissenberg K, Fredholm M, Christensen K, Anistoroaei R: **A large insertion in intron 2 of the TYRP1 gene associated with American Palomino phenotype in American mink**. Mamm Genome 2016, **27**(3–4):135–143.
8. Manakhov AD, Mintseva MY, Andreev IA, Uralsky LI, Andreeva TV, Trapezov OV, Rogaev EI: **Genome analysis of American minks reveals link of mutations in Ras-related protein-38 gene to Moyle brown coat phenotype**. Sci Rep 2020, **10**(1):15876.
9. Benkel BF, Rouvinen-Watt K, Farid H, Anistoroaei R: **Molecular characterization of the Himalayan mink**. Mamm Genome 2009, **20**(4):256–259.
10. Li YC, He DQ, Ma YH, Ma Q, Ding W, Chen YH, Zhang M, Luo F, Chen LY, Wang JK *et al*: **Skin transcriptome analysis identifies the key genes underlying fur development in Chinese Tan sheep in the birth and Er-mao periods**. Gene 2022, **820**:146257.
11. Polakis P: **Wnt signaling and cancer**. Genes Dev 2000, **14**(15):1837–1851.
12. Mikkola ML: **Genetic basis of skin appendage development**. Semin Cell Dev Biol 2007, **18**(2):225–236.
13. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B: **Mapping and quantifying mammalian transcriptomes by RNA-Seq**. Nat Methods 2008, **5**(7):621–628.
14. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL: **StringTie enables improved reconstruction of a transcriptome from RNA-seq reads**. Nat Biotechnol 2015, **33**(3):290–295.
15. Yang L, Smyth GK, Wei S: **featureCounts: an efficient general purpose program for assigning sequence reads to genomic features**. Bioinformatics 2014(7):923–930.
16. Bray NL, Pimentel H, Melsted P, Pachter L: **Near-optimal RNA-Seq quantification**. Computer Science 2015.
17. Love MI, Huber W, Anders S: **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2**. Genome Biology 2014, **15**(12):550.
18. Ogata H, Goto S, Sato K, Fujibuchi W, Kanehisa M: **KEGG: kyoto Encyclopedia of Genes and Genomes**. Nucleic Acids Research 1999, **27**(1):29–34.
19. Lamoreux ML, Wakamatsu K, Ito S: **Interaction of Major Coat Color Gene Functions in Mice as Studied by Chemical Analysis of Eumelanin and Pheomelanin**. Pigment Cell Research 2001, **14**(1).

20. Laura, Baxter, Dawn, Watkins-Chow, William, Pavan, Stacie: **A curated gene list for expanding the horizons of pigmentation biology**. Pigment Cell & Melanoma Research 2018.
21. Gu LH, Coulombe PA: **Keratin expression provides novel insight into the morphogenesis and function of the companion layer in hair follicles**. Journal of Investigative Dermatology 2007, **127**(5):1061–1073.
22. Parry AL, Nixon AJ, Craven AJ, Pearson AJ: **The Microanatomy, Cell Replication, and Keratin Gene Expression of Hair Follicles during a Photoperiod-Induced Growth Cycle in Sheep**. Acta Anatomica 1995.
23. Yu Z, Gordon SW, Nixon AJ, Bawden CS, Rogers MA, Wildermoth JE, Maqbool NJ, Pearson AJ: **Expression patterns of keratin intermediate filament and keratin associated protein genes in wool follicles**. Differentiation 2009, **77**(3):307–316.
24. Millar SE, Willert K, Salinas PC, Roelink H, Nusse R, Sussman DJ, Barsh GS: **WNT signaling in the control of hair growth and structure**. Developmental Biology 1999, **207**(1):133–149.
25. Jung HS, Francis-West PH, Widelitz RB, Jiang TX, Ting-Berreth S, Tickle C, Wolpert L, Chuong CM: **Local inhibitory action of BMPs and their relationships with activators in feather formation: implications for periodic patterning**. Developmental Biology 1998, **196**(1):11–23.
26. Rao RK, Philipps AF, Williams CS, McCracken DM, Koldovsky O: **Luminal stability of insulin-like growth factors I and II in developing rat gastrointestinal tract**. Journal of pediatric gastroenterology and nutrition 1998, **26**(2):179–185.
27. Weger N, Schlake T: **Igf-I signalling controls the hair growth cycle and the differentiation of hair shafts**. The Journal of investigative dermatology 2005, **125**(5):873–882.
28. A PACE, B MBO: **'Hard' and 'soft' principles defining the structure, function and regulation of keratin intermediate filaments**. Current Opinion in Cell Biology 2002, **14**(1):110–122.
29. Steinert PM, Torchia DR, Mack JW: **Structural Features of Keratin Intermediate Filaments**: The Biology of Wool and Hair; 1988.
30. Moll R, Divo M, Langbein L: **The human keratins: biology and pathology**. Histochemistry and Cell Biology 2008, **129**(6):705–733.
31. Cadieu E, Neff MW, Quignon P, Walsh K, Chase K, Parker HG, Vonholdt BM, Rhue A, Boyko A, Byers A *et al*: **Coat variation in the domestic dog is governed by variants in three genes**. Science 2009, **326**(5949):150–153.
32. Gandolfi B, Outerbridge CA, Beresford LG, Myers JA, Pimentel M, Alhaddad H, Grahn JC, Grahn RA, Lyons LA: **The naked truth: Sphynx and Devon Rex cat breed mutations in KRT71**. Mammalian Genome 2010, **21**(9):509–515.
33. Runkel F, Klaften M, Koch K, B?Hnert V, Büssow H, Fuchs H, Franz T, Angelis M: **Morphologic and molecular characterization of two novel Krt71 (Krt2-6g) mutations: Krt71rco12 and Krt71rco13**. *Mammalian Genome* 2006, **17**(12):1172–1182.
34. Lemke JR, Kernland-Lang K, H?Rtnagel K, Itin P: **Monogenic Human Skin Disorders**. Dermatology 2014, **229**(2):55–64.

35. Schweizer J, Langbein L, Rogers MA, Winter H: **Hair follicle-specific keratins and their diseases**. Experimental Cell Research 2007, **313**(10):2010–2020.
36. Wu DD, Irwin DM, Zhang YP: **Molecular evolution of the keratin associated protein gene family in mammals, role in the evolution of mammalian hair** (**vol 8, pg 241**, 2008). *Bmc Evolutionary Biology* 2009, **9**.
37. Khan, Imran, Maldonado, Emanuel, Vasconcelos, Vítor, O'Brien, Stephen J, Johnson, Warren E: **Mammalian keratin associated proteins (KRTAPs) subgenomes: disentangling hair diversity and adaptation to terrestrial and aquatic environments**. BMC Genomics 2014.
38. Zhou H, Gong H, Li S, Luo Y, Hickford J: **A 57-bp deletion in the ovine KAP6-1 gene affects wool fibre diameter**. Journal of Animal Breeding and Genetics 2015, **132**(4).
39. Shaobin, Li, Huitong, Zhou, Hua, Gong, Fangfang, Zhao, Jiqing, Wang: **Identification of the Ovine Keratin-Associated Protein 22 – 1 (KAP22-1) Gene and Its Effect on Wool Traits**. *Genes* 2017.
40. Hua G, Zhou H, Dyer JM, Hickford J: **Identification of the ovine KAP11-1 gene (KRTAP11-1) and genetic variation in its coding sequence**. Molecular Biology Reports 2011, **38**(8):5429–5433.
41. Gong H, Zhou H, Hickford J: **Polymorphism of the ovine keratin-associated protein 1–4 gene (KRTAP1-4)**. *Molecular Biology Reports* 2010, **37**(7):3377–3380.
42. Rogers MA, Langbein L, Praetzel-Wunder S, Winter H, Schweizer J: **Human hair keratin-associated proteins (KAPs)**. International Review of Cytology 2006, **251**:209.
43. Rawlings JS, Rosler KM, Harrison DA: **The JAK/STAT signaling pathway**. J Cell Sci 2004, **117**(Pt 8):1281–1283.
44. Itoh F, Watabe T, Miyazono K: **Roles of TGF-beta family signals in the fate determination of pluripotent stem cells**. Semin Cell Dev Biol 2014, **32**:98–106.
45. Dimitriou R, Giannoudis PV: **Discovery and development of BMPs**. Injury-international Journal of the Care of the Injured 2005, **36**(3-supp-S):S28-S33.
46. Rosso F, Giordano A, Barbarisi M, Barbarisi A: **From cell-ECM interactions to tissue engineering**. Journal of Cellular Physiology 2010, **199**(2):174–180.
47. Comoglio PM, Boccaccio C, Trusolino L: **Interactions between growth factor receptors and adhesion molecules: breaking the rules**. Current opinion in cell biology 2003, **15**(5):565.
48. Parmeggiani A: **GTPases – molecular switches of cellular signaling pathways**. Trends in Biochemical Sciences 2001.
49. Jimbow K, Park JS, Kato F, Hirosaki K, Toyofuku K, Chen H, Yamashita T: **Assembly, Target-Signaling and Intracellular Transport of Tyrosinase Gene Family Proteins in the Initial Stage of Melanosome Biogenesis**. Pigment Cell Research 2010, **13**(4):222–229.
50. Hirobe T, Wakamatsu K, Ito S: **Excess Tyrosine Stimulates Eumelanin and Pheomelanin Synthesis in Cultured Slaty Melanocytes from Neonatal Mouse Epidermis**. Zoologicalence 2007, **24**(3):209–217.
51. Makpol S, Jam FA, Rahim NA, Khor SC, Ismail Z, Yusof Y, Ngah W: **Comparable down-regulation of TYR, TYRP1 and TYRP2 genes and inhibition of melanogenesis by tyrostat, tocotrienol-rich fraction and**

tocopherol in human skin melanocytes improves skin pigmentation. La Clinica Terapeutica 2014, **165**(1):e39.

Figures

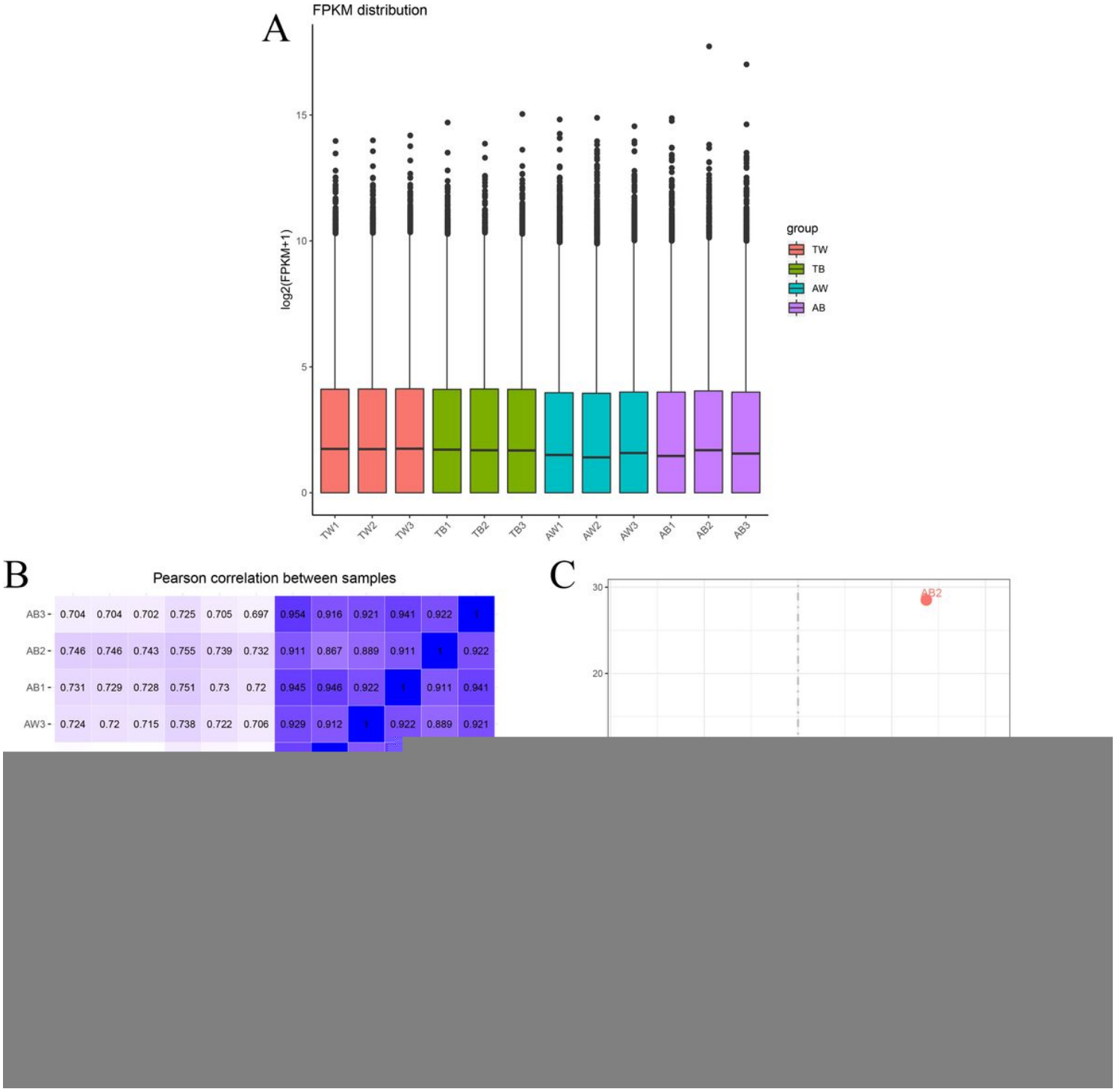

Figure 1

Fig 1 A Comparison of the distribution of gene expression levels in different samples. **B** Sample correlation heat map. C Principal component analysis (PCA).

Figure 2

Fig 2 Volcano map of differentially expressed genes in different comparison combinations

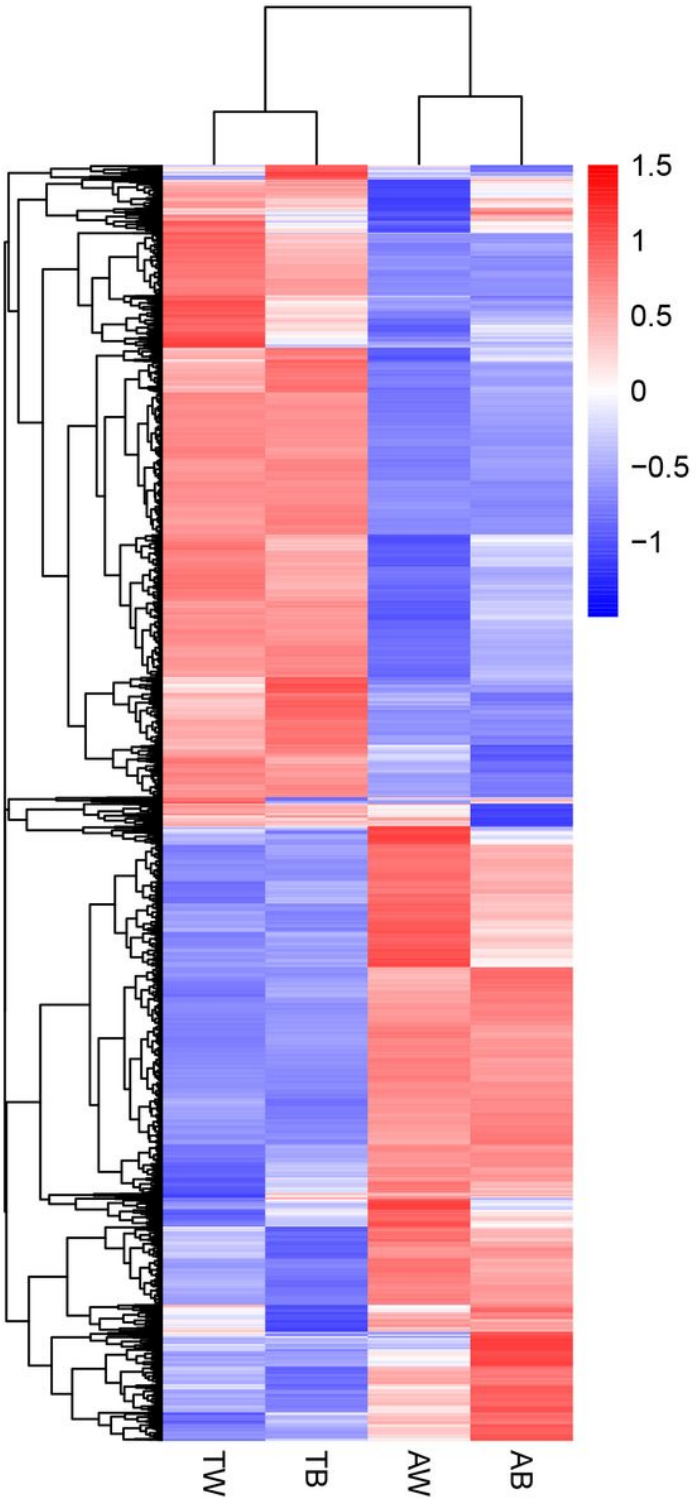

Figure 3

Fig 3 Cluster heat map of differentially expressed genes

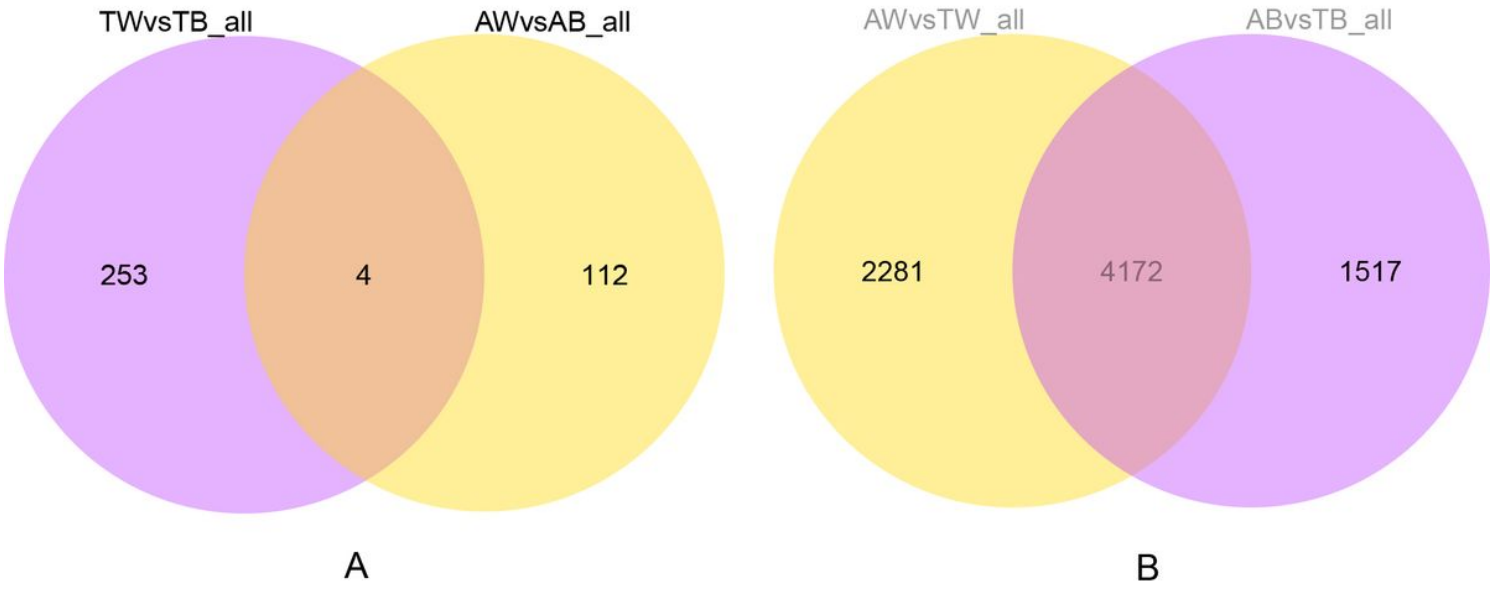

Figure 4

Fig 4 A Venn diagram of differentially expressed genes in different hair colors of two ages. **B** Venn diagram of differentially expressed genes at different ages of two coat colors.

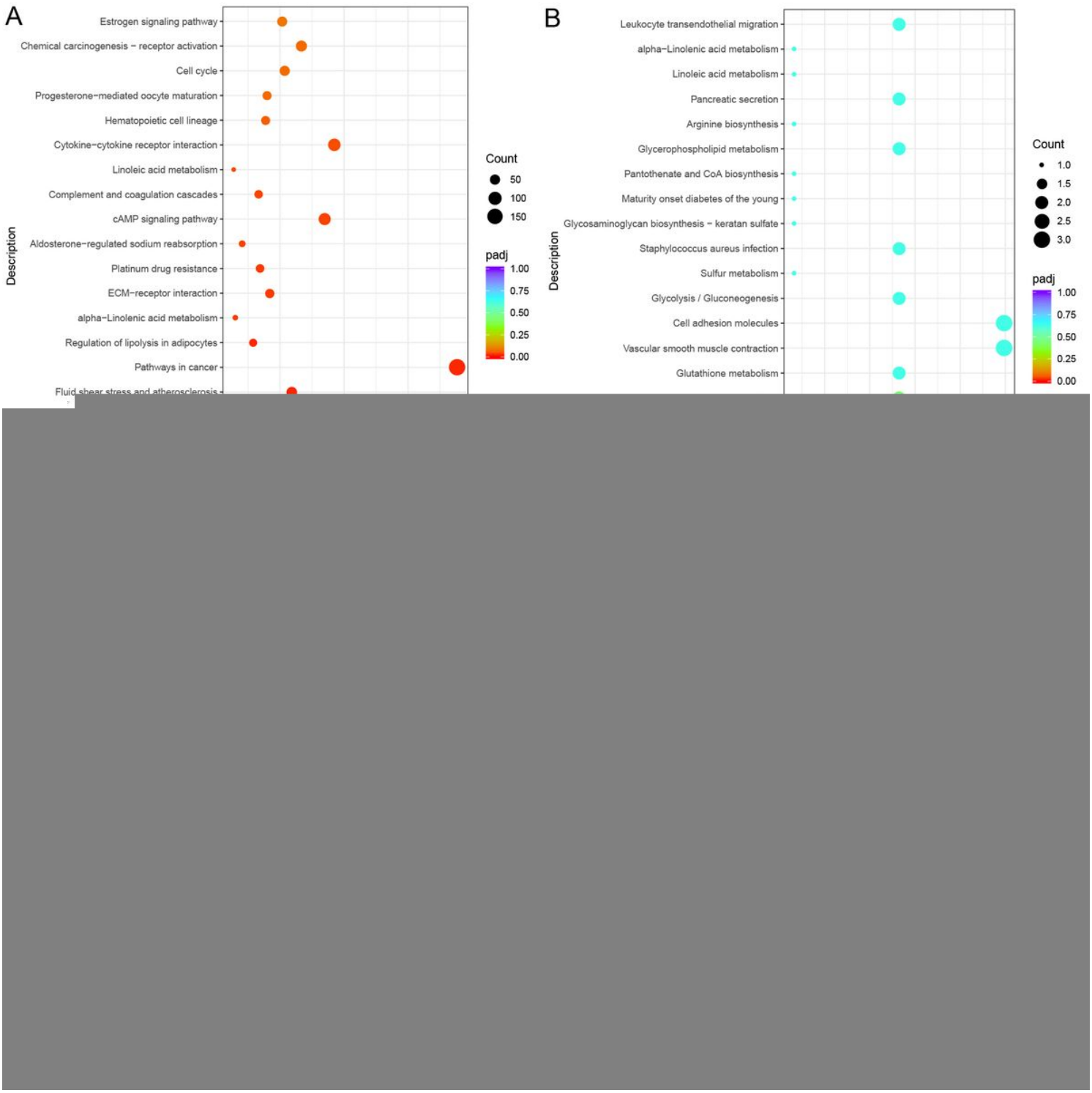

Figure 5

Fig 5 Scatter plot of KEGG enrichment analysis for differentially expressed genes in different combinations (**A**: AB vs TB, **B**: AW vs AB, **C**: AW vs TW, **D**: TW vs TB).

Supplementary Files

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

- FigS1.pdf
- TableS1.xlsx
- TableS2.xlsx
- TableS3.xlsx
- TableS4.xls
- TableS5.xlsx
- TableS6.xlsx