

# Identification of Key Genes Affecting Porcine Fat Deposition Based on Co-Expression Network Analysis of Weighted Genes

**Kai Xing**

Beijing University of Agriculture

**Huatao Liu**

China Agricultural University

**Fengxia Zhang**

China Agricultural University

**Yibing Liu**

China Agricultural University

**Yong Shi**

China Agricultural University

**Xiangdong Ding**

China Agricultural University

**Chuduan Wang** (✉ [cdwang@cau.edu.cn](mailto:cdwang@cau.edu.cn))

China Agricultural University

---

## Research

**Keywords:** pigs, fat deposition, RNA-seq, WGCNA

**Posted Date:** January 19th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-148726/v1>

**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 at Journal of Animal Science and Biotechnology on August 20th, 2021. See the published version at <https://doi.org/10.1186/s40104-021-00616-9>.

# Abstract

**Background:** Fat deposition is an important economic consideration for pig production. The amount of fat deposition in pigs seriously affects production efficiency, quality, and reproductive performance, while also affecting consumers' choice of pork. Weighted gene co-expression network analysis (WGCNA) has been shown to be effective in pig genetic studies. Therefore, this study aimed to identify modules that co-express genes associated with fat deposition in pigs (Songliao black and Landrace breeds) with extreme levels of backfat (high and low), and to identify the central genes in each of these modules.

**Results:** We used RNA-seq of different pig tissues to construct a gene expression matrix consisting of 12 862 genes from 36 samples. Eleven co-expression modules were identified using WGCNA; the number of genes in these modules ranged from 39 to 3363. We found four co-expression modules were significantly correlated with backfat thickness. A total of 14 genes ( RAD9A , IGF2R , SCAP , TCAP , DGAT1 , GPS2 , IGF1 , MAPK8 , FABP , FABP5 , LEPR , UCP3 , APOF , and FASN ) were found to be related to fat deposition.

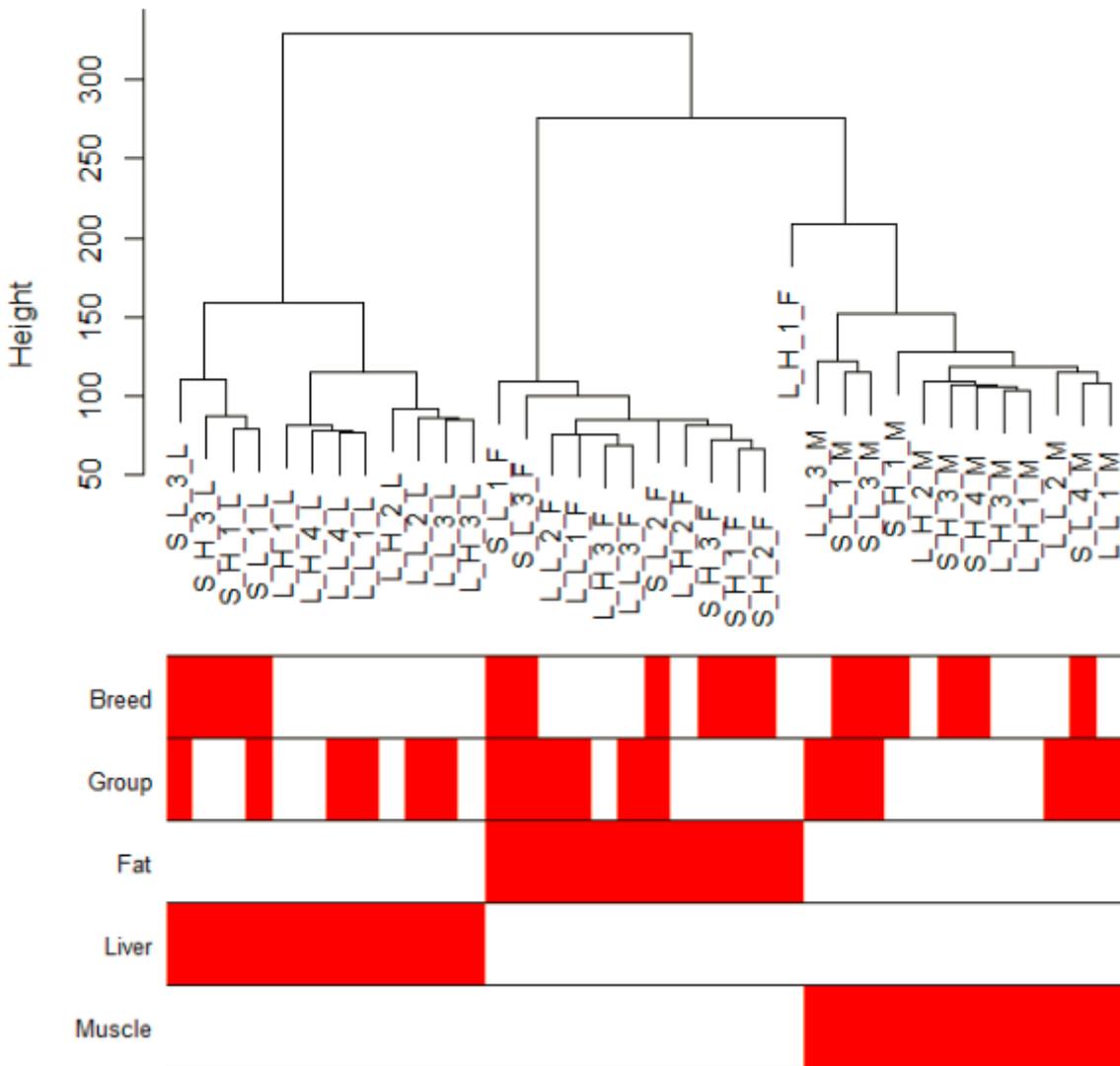
**Conclusions:** RAD9A , TCAP , GPS2 , and APOF were found to be the key genes in the four modules according to the degree of gene connectivity. Combining the results of differential gene analysis, APOF was proposed as a strong candidate gene for body size traits. This study explores the key genes that regulate porcine fat deposition and lays the foundation for further research into the molecular regulatory mechanisms behind porcine fat deposition.

## Full Text

Due to technical limitations, full-text HTML conversion of this manuscript could not be completed. However, the latest manuscript can be downloaded and [accessed as a PDF](#).

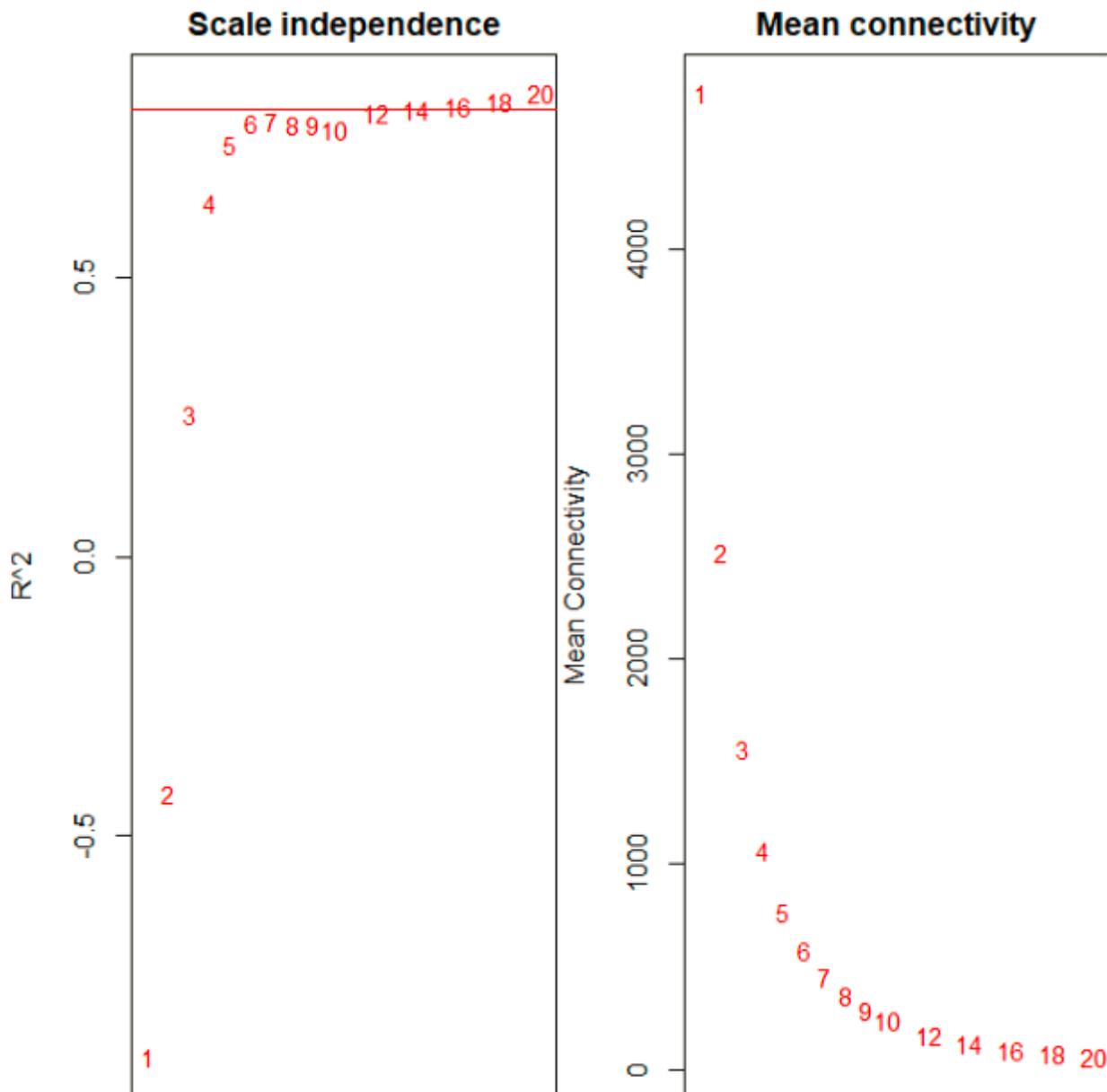
## Figures

### Sample dendrogram and trait heatmap



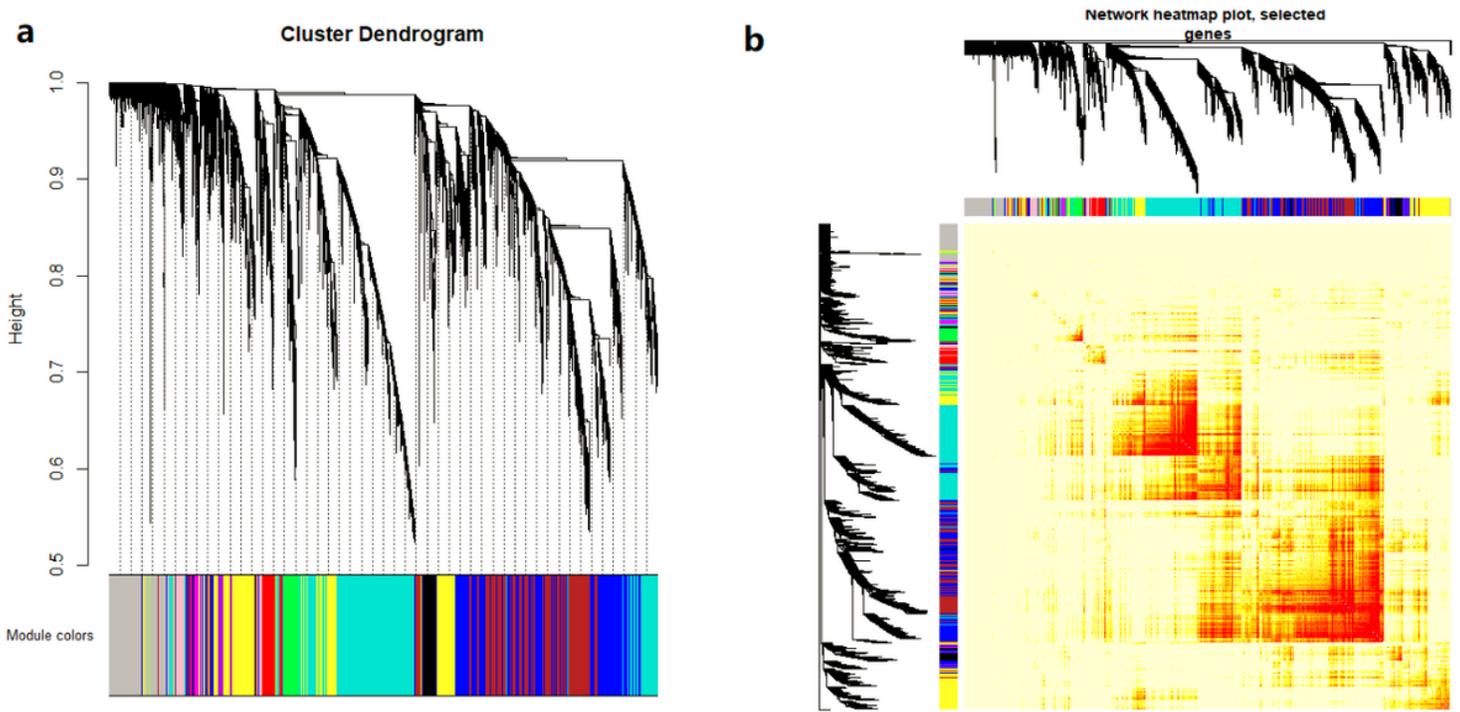
**Figure 1**

Sample clustering Clustering was carried out according to the samples to eliminate low-quality samples. From the clustering tree, there were no outliers



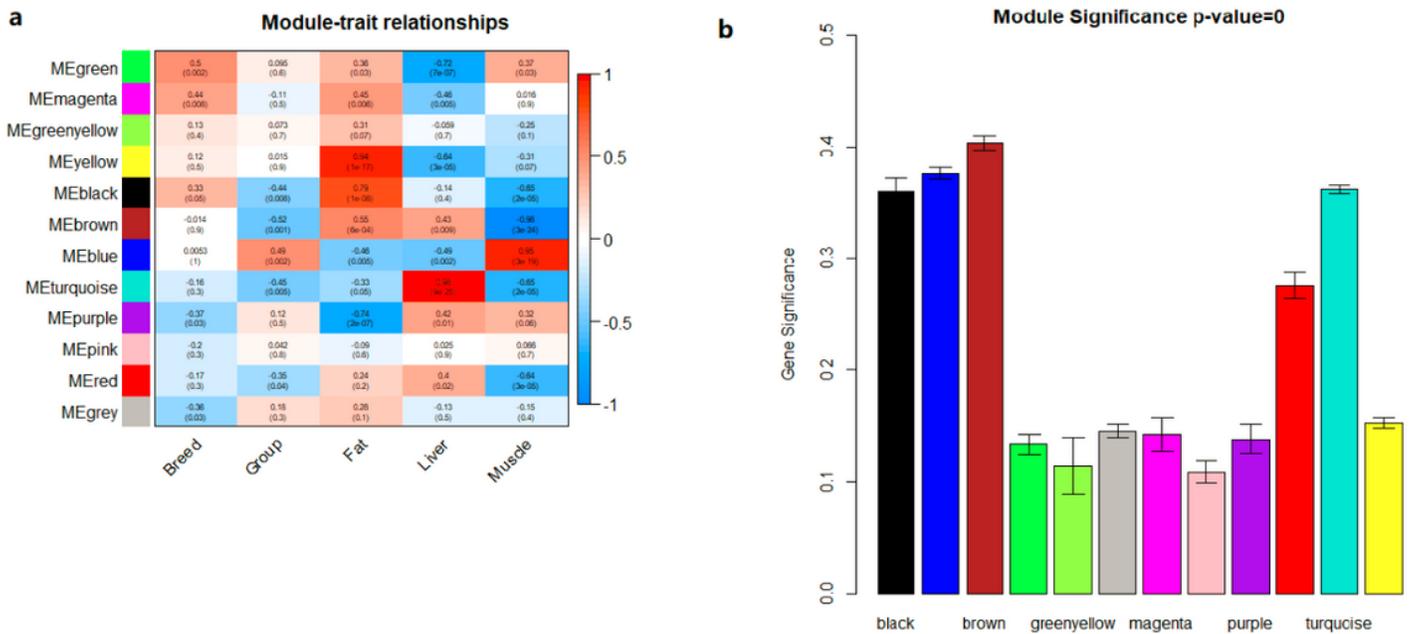
**Figure 2**

Analysis of network topology for various soft-thresholding powers According to the definition of soft threshold, try to choose a large  $R^2$  value. According to the suggestion of WGCNA package, choose an  $R^2$  value greater than 0.8, that is, the value above the red line in the left figure. The figure on the right shows the average connectivity of the constructed network. The larger the network is, the closer the gene is, and the more conducive it is to screen out hub genes.



**Figure 3**

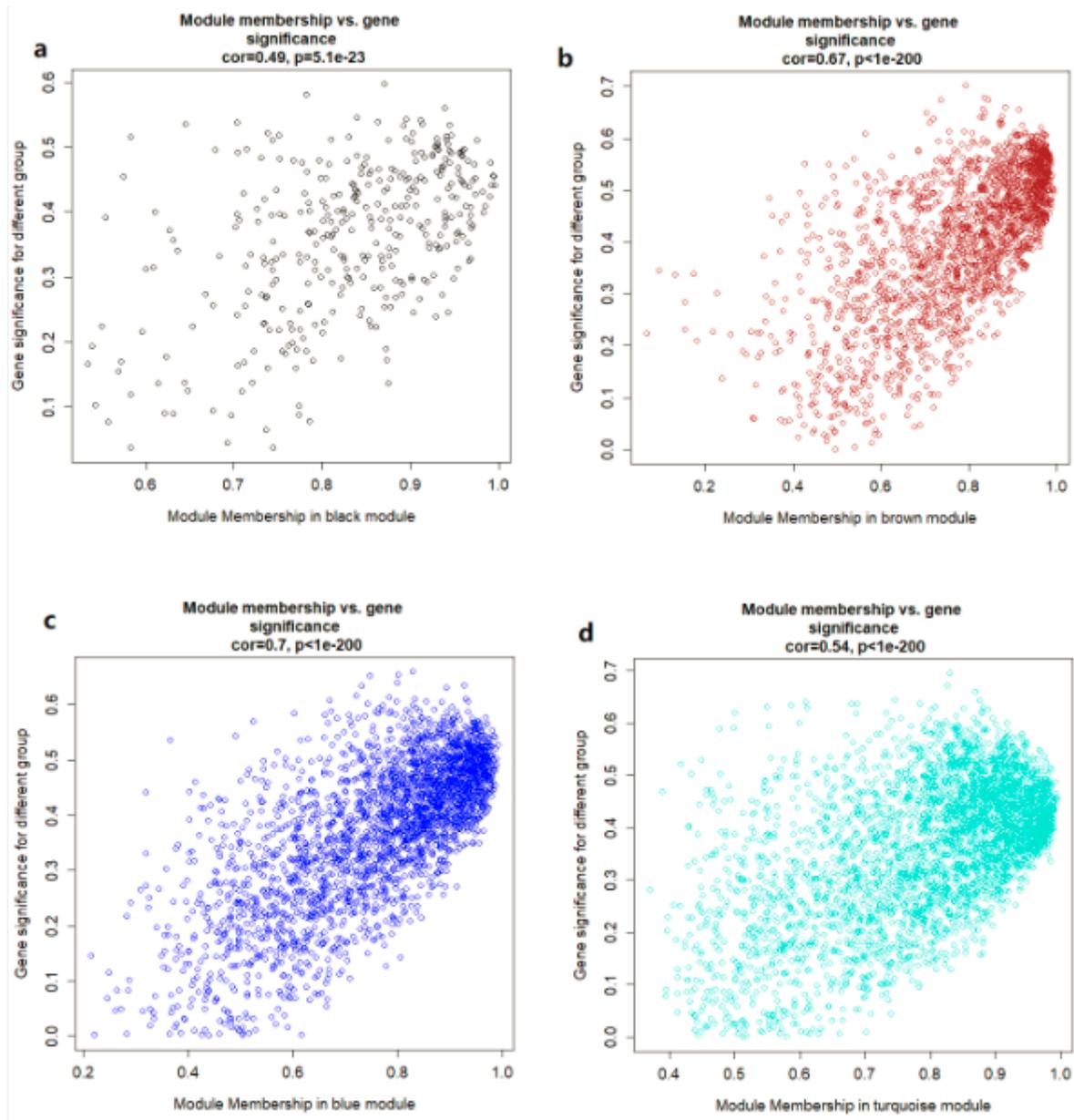
Division of gene modules Both a and b show the clustering of genes, and the division of gene modules is based on this result. Branches of the same color are divided into the same module of genes.



**Figure 4**

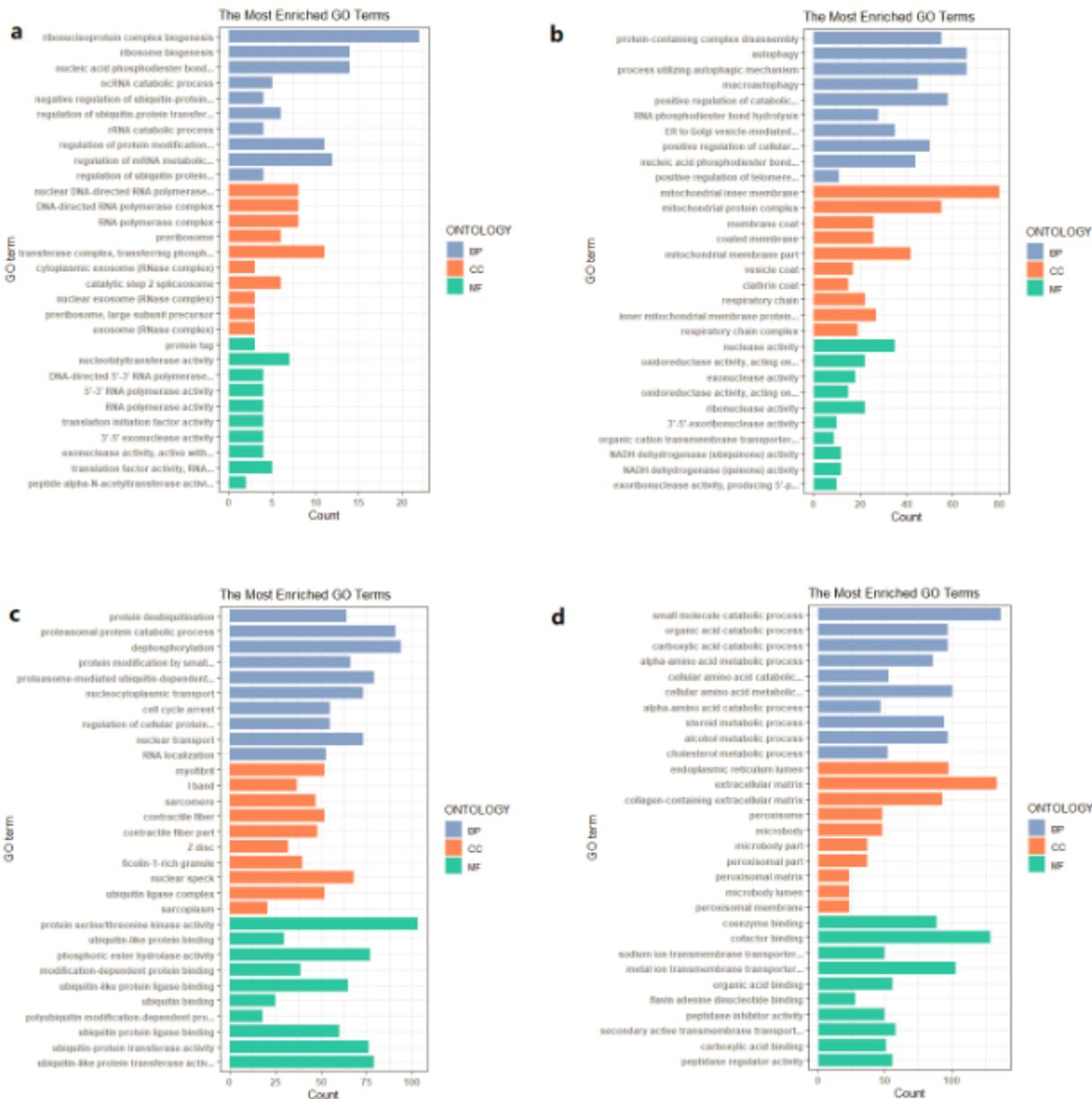
Module-trait associations In a, the correlation between the gene module and the sample information is shown, where the x-axis represents the sample information, including varieties, grouping of backfat thickness and sample tissue type. The y-axis represents each gene module. In the figure, the darker the color is, the higher the correlation is, in which red is the positive correlation and blue is the negative

correlation. Significance is in brackets and is represented by P value. Figure b shows the absolute correlation between genes in each module for backfat thickness.



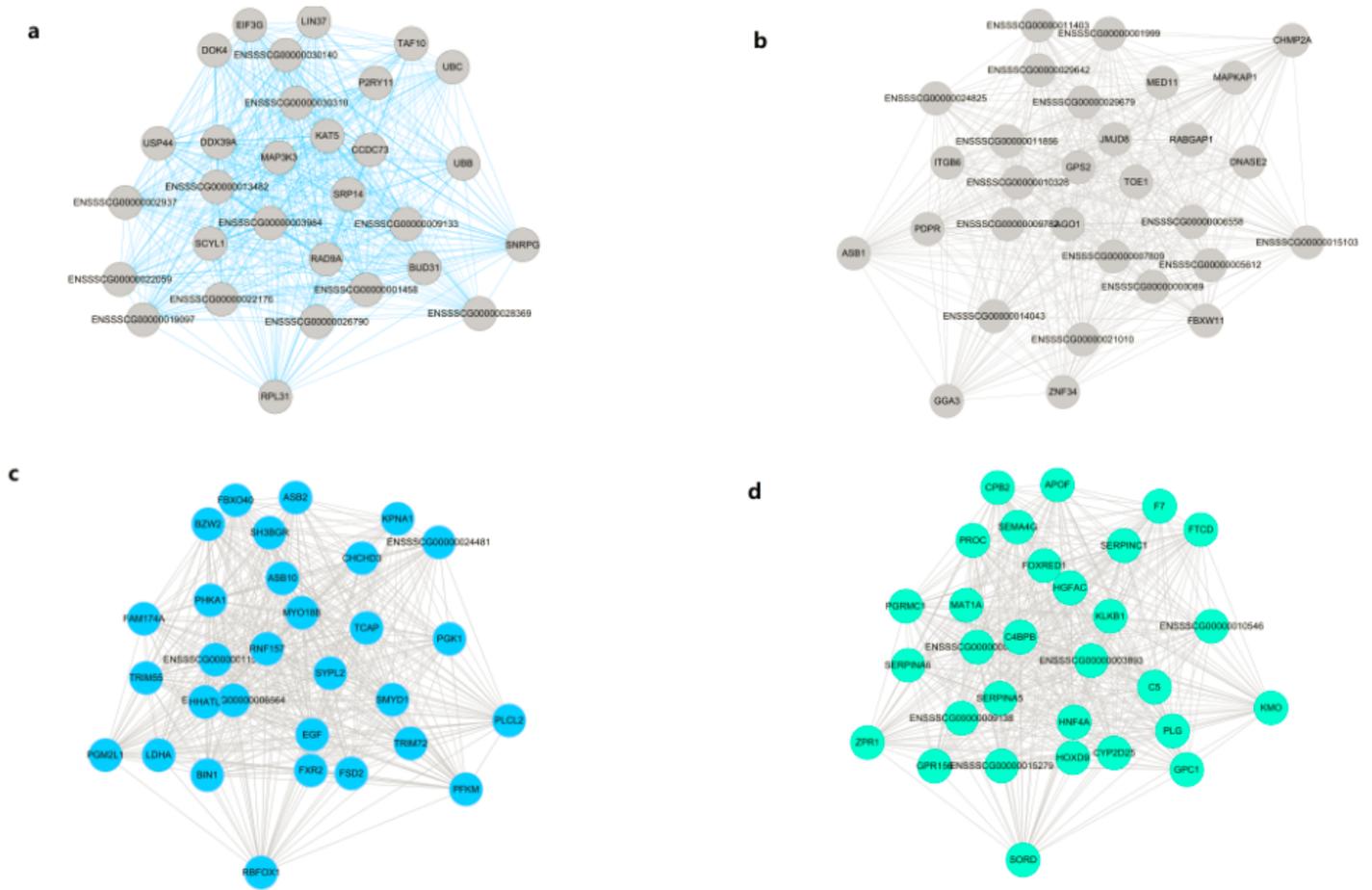
**Figure 5**

Scatterplot of Gene Significance (GS) for backfat vs. Module Membership (MM) in candidate modules. The figure shows the significance of the genes in the four modules. The x-axis represents the value of Module Membership in each module, and the y-axis represents Gene Significance for backfat; a, b, c, and d represent the genes in the black, brown, blue, and turquoise modules respectively. The gene in the upper right corner of each graph is the hub gene we need to look for. These genes are highly correlated with phenotypes and have a high MM, which is a good representation of the gene module.



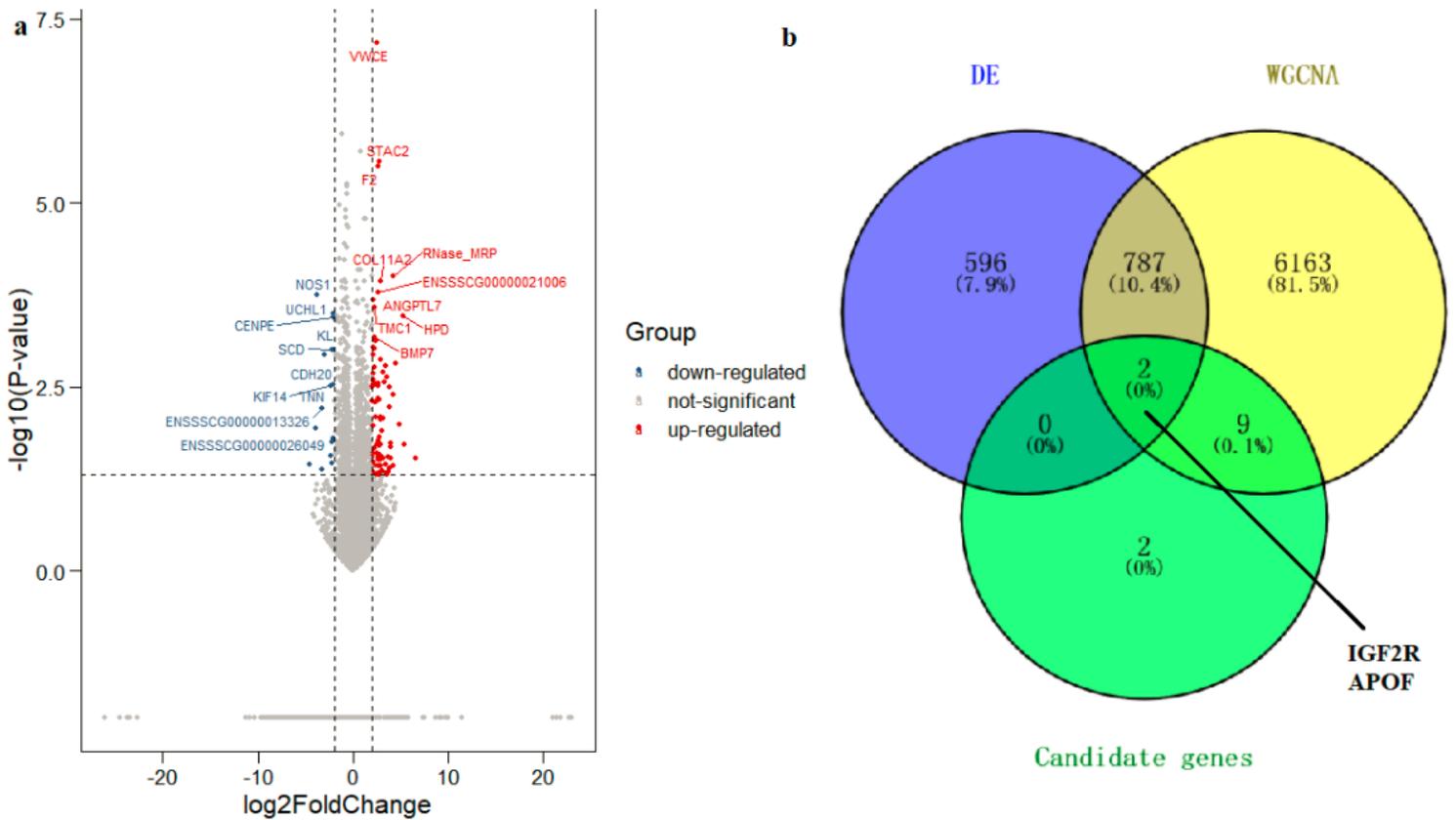
**Figure 6**

GO enrichment analysis of genes in co-expression modules. In the figure, a, b, c and d respectively show the enrichment of the genes in black, blue, brown and sapphire modules from three different biological categories. The y-axis represents the gene enriched entries, and the x-axis represents the number of genes enriched to the same entry. In each figure, different biological categories are represented with different colors, in which CC stands for cellular Components, MF stands for molecular functions, and BP stands for biological processes.



**Figure 7**

Visualization of the top 30 genes in a given module: (a) black, (b) brown, (c) blue, and (d) turquoise. The figure shows the network relationship diagram of hub genes in different modules, in which the thicker the line between genes indicates the higher the correlation of genes, and the more the number of lines between the gene and other genes indicates the more important the gene is.



**Figure 8**

Volcanogram (a) and Venn map (b) of differentially expressed genes (DEGs) Figure a shows the situation of the differential genes. The x-axis represents the multiple of difference, which is denoted by  $\log_2(\text{FoldChange})$ . The larger the absolute value is, the larger the multiple of difference is. The y-axis represents the significance of the difference, which is denoted by  $-\log_{10}(\text{P-value})$ . The larger the value is, the more significant the difference is. The figure shows the names of the top 20 genes with the most significant differences. Figure b shows the Venn Map of various analysis methods, among which intersection is the most important candidate gene.

## Supplementary Files

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

- [SupplementaryTable.docx](#)