**Effects of candidate genes on milk fat synthesis in ruminants: A meta-analysis**

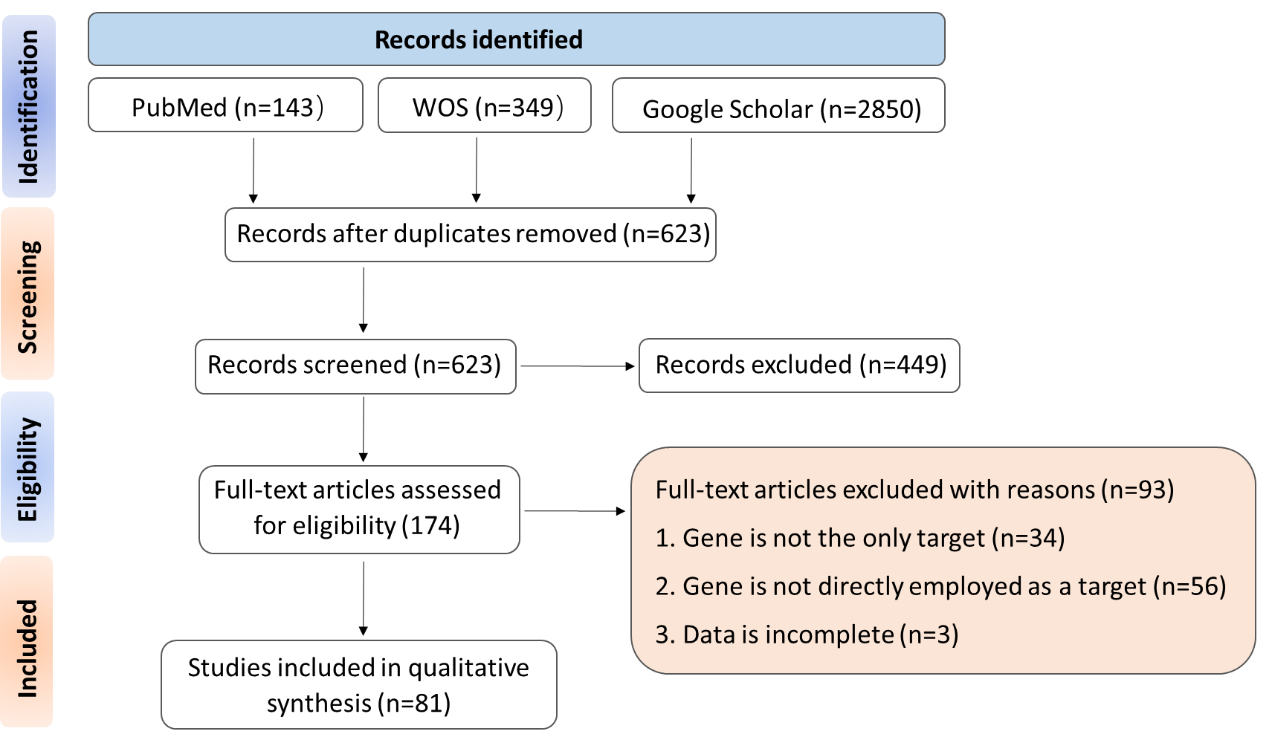
**Lily Liu12\*, Wenquan Yan1, Min Yan1, Songlin Liu1, Qin Zhang3, Emily Clark2\*, Jinhai Wang2\***

**Tables**

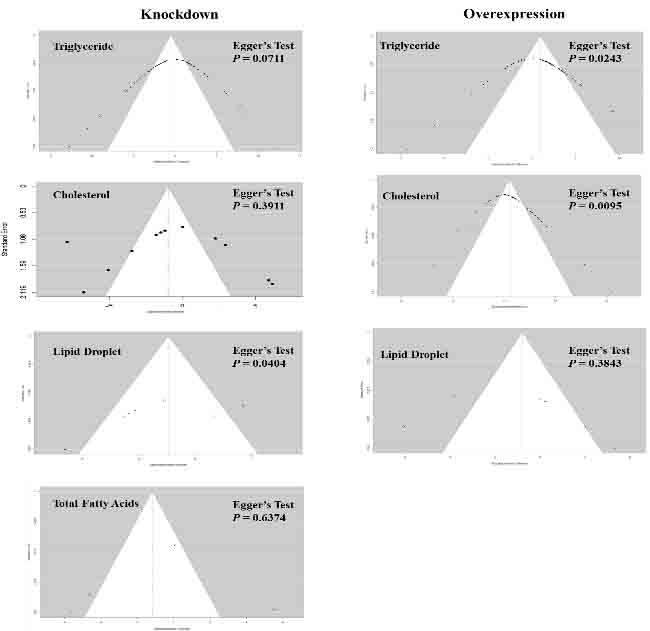
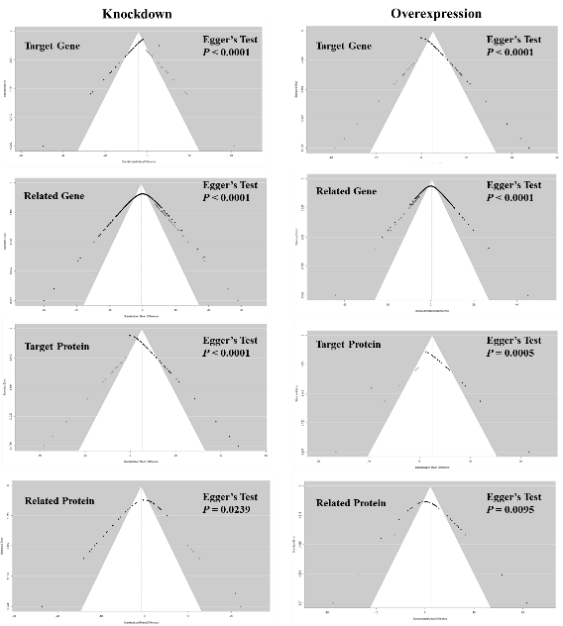
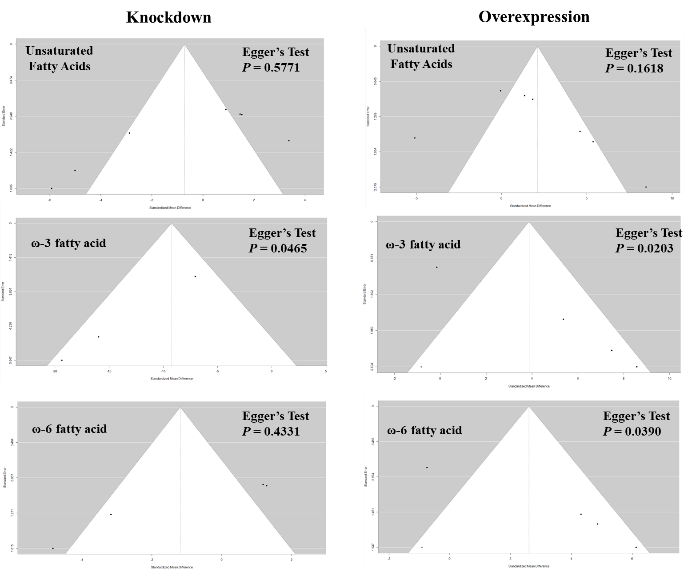
**Table S1.** Overall effect size calculation outcomes for gene/protein expression, triglyceride (TG), lipid droplet (LD), cholesterol (CHO), unsaturated fatty acids (UFAs). SMD, standardized mean difference; k, the number of effect sizes for each level of moderators; CI, confidence interval; PI, predicted interval.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Group** | **Parameter** | **k** | **SMD** | ***I*2 %** | **95% CI** | | **P-value** | **95% PI** | |
| **Lower** | **Upper** | **Lower** | **Upper** |
| **Knock-down** | Target Gene | 55 | -5.7 | 66.7 | -6.57 | -4.83 | **<0.0001** | -10.59 | -0.81 |
| Related Gene | 497 | -1.29 | 85.3 | -1.83 | -0.75 | **<0.0001** | -6.27 | 3.7 |
| Target Protein | 23 | 3.18 | 0.0 | 2.65 | 3.7 | **<0.0001** | 2.65 | 3.7 |
| Related Protein | 43 | -1.82 | 93.2 | -4.56 | 0.93 | 0.194 | -11.2 | 7.56 |
| TG | 50 | -0.82 | 87.6 | -0.71 | 0.7 | 0.0711 | -6.64 | 5 |
| LD | 6 | -1.67 | 71.6 | -3.27 | -0.07 | **0.0404** | -5.32 | 1.98 |
| CHO | 12 | -1.23 | 92.9 | -3.85 | 1.39 | 0.3568 | -9.91 | 7.44 |
| UFAs | 7 | -0.63 | 89.0 | -3.24 | 1.99 | 0.6382 | -7.44 | 6.19 |
| **Over-expression** | Target Gene | 48 | 7.97 | 86.5 | 6.37 | 9.57 | **<0.0001** | -1.05 | 16.99 |
| Related Gene | 444 | 1.21 | 85.3 | 0.48 | 1.93 | **0.0011** | -5 | 7.42 |
| Target Protein | 22 | 3.52 | 64.6 | 2.47 | 4.57 | **<0.0001** | -0.03 | 7.33 |
| Related Protein | 36 | 0.96 | 88.9 | -0.66 | 2.58 | 0.2467 | -5.37 | 7.28 |
| TG | 48 | 0.99 | 85.7 | 0.13 | 1.85 | **<0.0001** | -4.47 | 6.45 |
| LD | 6 | 1.22 | 87.6 | -1.53 | 3.97 | 0.3843 | -5.6 | 8.04 |
| CHO | 8 | -0.83 | 82.6 | -2.71 | 1.06 | 0.3903 | -5.73 | 4.08 |
| UFAs | 7 | 0.85 | 88.7 | -2.33 | 4.02 | 0.6017 | -6.62 | 8.32 |

**Figures**



**Figure.S1** Flow chart for search strategies and selection details based on the statement of Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).



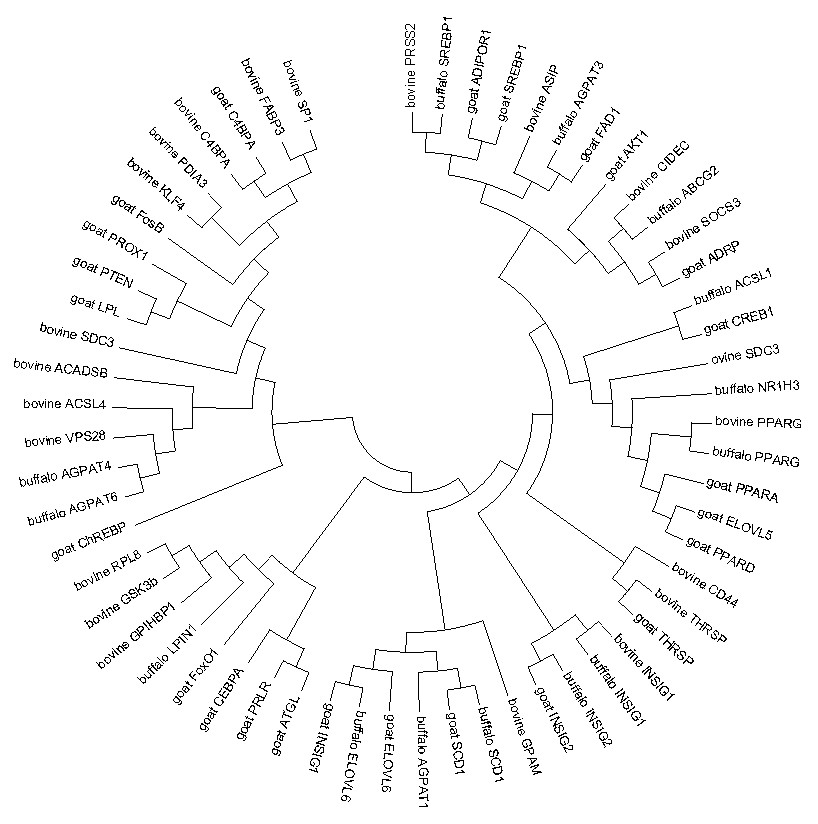
**Figure.S2** Evidence of publication bias. Funnel plots of standardized mean difference of means for Target Gene, Related Gene, Target Protein, Related Protein, Triglyceride, Cholesterol, Lipid Droplet, Unsaturated Fatty Acid, in knockdown group and overexpression group. Egger’s test and corresponding P-values were illustrated. *P* < 0.05 indicated the funnel plots exhibited significant asymmetry, indicating potential publication bias. Therefore, “trim and fill” was applied to test if there were any additional studies to be implanted to eliminate publication bias. Theoretically, the updated funnel plots show the additional missing studies imputed in white by “trim and fill” if extra studies should be added (The funnel plots were the same as original ones after “trill and fill”). This suggests that this asymmetry is not due to publication bias. Please refer to the Discussion section of the main text for more details.

A screenshot of a graph

Description automatically generatedA screenshot of a graph

Description automatically generated

**Figure S3.** Sensitivity analysis based on influence method. There are no potential effect-size outliers for all subgroups, except Target Gene in knockdown group. In this case, the leave-one-out method was employed to detect whether these outliers had opposing effects on the results. Please refer to the Results section of the main text for more details. Note: red dots indicate the outliers.



**Figure.S4** Molecular phylogenetic analysis by Maximum Likelihood method

**Notes**

**Note 1:** Screening process

In details, to ensure repeatability of the screening process, two analysts (Lily Liu and Jinhai Wang) screened titles/abstracts and full-texts independently and checked for agreement on included vs. excluded studies. Analyst 1 screened out all titles/abstracts to check the suitability of each study (N = 3342) while analyst 2 checked for repeatability. Analysts 1 and 2 agreed on the inclusion or exclusion (i.e. suitability) of over 90% of the studies when screening abstracts. Analyst 1 then screened all full texts that passed the titles/abstracts screening stage (N = 623 studies) to determine whether a study was suitable to be included in the meta-dataset, while analyst 2 screened ~ 30% of full texts to test for repeatability (N = 180). Analysts 1 and 2 agreed on the inclusion or exclusion of over 95% of the studies when screening full texts.

**Note 2:** Data extraction

Similarly, to assure repeatability of data extraction, three analysts checked data obtained from ~ 10% of papers (i.e. 55 studies and 55 effect sizes in Target gene expression of Gene knockdown dataset; OR 53 studies and 514 effect sizes in Target gene expression of Gene overexpression dataset; OR 53 studies and 498 effect sizes in Related gene expression of Gene knockdown dataset; OR 43 studies and 445 effect sizes in Related gene expression of Gene overexpression dataset). To do this, 4 analysts first independently extracted data and calculated effect sizes (standardized mean difference, SMD) for each of the 349 rows of the meta-dataset. The difference of calculated SMDs from two analysts were not significant (P > 0.05), indicating strong repeatability.

**Note 3:** Calculation of effect size

*Hedges’g* served as the effect size because it corrects for differences in sampling effort among studies and adjusts for small sample sizes and was calculated using the following formula:

A math equations with numbers

Description automatically generated with medium confidence

where mean*T* is the mean of the transgenic/gene-edited group, mean*C* is the mean of WT group, n1 and n2 are two sample sizes from these groups, and *SD*12 and *SD*22 are the estimated population variance of both groups.

Based on the formula above mentioned, we collected data on means, standard deviation/standard error (SD/SE), and sample sizes of each parameter in both groups from: text in the Methods/Results section, figures using ImageJ.JS (<https://ij.imjoy.io/>), the supplementary information or raw data provided with the paper, or a data repository, in that order. SD were determined using the formula SD = SE\*(n)1/2 if only SE and sample size (n) are supplied in a single publication. Additionally, when medians and inter-quartile ranges were reported, we converted these to SD.