分析报告

# M1 and M2

为了分析C，Y和W组的表型跟倾向于M1还是M2，我们到Gene Expression Omnibus（GEO）数据库上寻找M1和M2巨噬细胞的测序数据，我们发现Schmidt等人的测序数据符合我们的要求（GSE36952）1。因此，我们将其下载下来，然后使用下面的公式分别将我们自己的测序数据和Schmidt等人的测序数据标准化。

其中表示比对到基因g的reads数，D表示测序深度，表示基因g的长度（bp）。

数据标准化后分别计算C，Y和W与M1和M2的spearman相关系数。相关系数矩阵使用ward.D2进行聚类。结果显示C，Y和W的表型都更倾向于M2，但是C组的表型更接近于M2表型，W组的表型与M2的相似度最小。

In order to analyze whether the phenotype of C, Y, and W groups tends towards M1 or M2, we searched for sequencing data of M1 and M2 macrophages in the Gene Expression Omnibus (GEO) database. We found that the sequencing data of Schmidt et al. met our requirements1. Therefore, we downloaded it and standardized our own sequencing data and Schmidt et al.’s sequencing data using the following formula.

Among them, represents the number of reads aligned to gene g, D represents sequencing depth, and represents the length (bp) of gene g.

Calculate the Spearman correlation coefficients of C, Y, and W with M1 and M2 after data standardization. The results showed that the phenotypes of C, Y, and W were more inclined towards M2, but the phenotypes of Group C were closer to M2, while the phenotypes of Group W had the smallest similarity to M2.

# fatty acid biosysnthesis

为了研究C，Y和W三组哪个的脂肪合成代谢最为旺盛，我们首先从KEGG数据库中下载了fatty acid biosysnthesis和Biosynthesis of unsaturated fatty acids两个与脂肪合成相关代谢通路中的基因，然后分析了这些基因在C，Y和W三组中的表达水平。最后我们使用ward.D2聚类方法将不同的样本进行聚类，聚类结果显示不同处理组的表型相似度大于组间的相似度，说明实验条件处理的较为成功。分析结果显示W组的脂肪合成代谢最为旺盛，尤其是ACSL4基因，在W组的表达水平非常高，已有的研究显示其主要参与长链脂肪酸的合成。

In order to investigate which group C, Y, and W had the most vigorous fat synthesis metabolism, we first downloaded the genes related to fat acid biosynthesis and Biosynthesis of Unsaturated fat acids from the KEGG database, and then analyzed the expression levels of these genes in the C, Y, and W groups. The results showed that the fat synthesis metabolism in group W was the most vigorous, especially the ACSL4 gene, which had a very high expression level in group W and was mainly involved in the synthesis of long-chain fatty acids.

# 图说明

* fatty\_acid\_biosysnthesis.pdf 每一列是一个样本，每一行是一个参与脂肪合成的基因。将行和列分别用ward.D2聚类算法急性聚类，表型像是的样本会聚集到一起，表达水平相近的基因也会聚到一起。
* similarity\_score.pdf 计算出来的不同样本与M1和M2表型的spearman相关系数矩阵，然后用ward.D2对相关系数矩阵进行聚类。

# 参考文献

1. Schmidt, S. V. *et al.* The transcriptional regulator network of human inflammatory macrophages is defined by open chromatin. *Cell Res.* **26**, 151–170 (2016).