第 7 卷◆第 1 期◆版本 1.0◆2025 年
文章类型：论文丨刊号（ISSN）：2705-1102(P) / 2705-1110(O)

生物信息学方法研究脂肪分化过程中的差异基因

肖昭晖[1]　韩桂艳[2*]
1 承德医学院研究生学院　2 承德医学院附属医院内分泌科
DOI:10.12238/bmtr.v7i1.11799

[摘　要] 目的：应用生物信息学方法鉴定脂肪分化过程中差异表达的基因。方法：利用NCBI中的GEO基因芯片公共数据库进行芯片数据搜索,最终选择数据集(GSE185481),使用软件R Studio的程序包识别差异基因。结果：通过对GSE185481进行分析,通过String和Cytoscape分析鉴定出7个枢纽基因。结论：脂肪分化过程中有差异基因的表达,本实验通过分析芯片数据(GSE185481)筛选出Hub基因：Mki67、Ccna2、Aurka、TOP2A、Cdc20、MMP9和Melk。
[关键词] 生物信息学；肥胖；差异基因
中图分类号：R723.14　文献标识码：A

Bioinformatics methods were used to study the differentially expressed genes during adipogenesis

Zhaohui Xiao[1]　Guiyan Han[2*]
1 Graduate School of Chengde Medical University
2 Department of Endocrinology, Affiliated Hospital of Chengde Medical University
[Abstract] Objective: To identify differentially expressed genes during adipocyte differentiation using bioinformatics methods.Methods: Chip data were searched in the GEO gene chip public database of NCBI, and the chip data (GSE185481) was ultimately selected as the analysis object. The differentially expressed genes were identified using the functions of R tools in the Bioconductor package and the Limma package.Results: Through the analysis of GSE185481,7 hub genes were identified by String and Cytoscape analysis.Conclusion: There are differentially expressed genes during adipocyte differentiation. In this experiment, hub genes were screened out by analyzing the chip data (GSE185481): Mki67、Ccna2、Aurka、TOP2A、Cdc20、MMP9 and Melk.
[Key words] Bioinformatics; Obesity; Differentially expressed genes

在过去的几十年里,肥胖已成为全球日益严重的公共卫生问题[1]。肥胖被定义为体脂(BF)的过度积累或异常分布,且肥胖会影响健康[2]。会导致多种其他疾病[3、4]。最近,基因表达谱分析与生物信息学相结合已成为研究与各种疾病相关的重要信号通路和基因的通用且有效的方法。因此,分析生物信息学数据库有助于预测、诊断和治疗多种疾病[5、6]。本研究用软件R Studio分析GEO数据库中的数据集(GSE185481),对该数据集进行挖掘,探讨脂肪分化过程中差异基因。

1 材料与方法

1.1材料。利用NCBI(National Center for Biotechnology Information)中的GEO(Gene Expression Omnibus)基因芯片公共数据库进行芯片数据搜索。登录网址为：https://www.ncbi.nlm.nih.gov/geo/,搜索关键词为小鼠(Mus musculus),肥胖(obesity)。筛选后,选择由数据集(GSE185481)作为分析对象,该数据集使用4组细胞：棕色前脂肪细胞(n=3)、分化的棕色脂肪细胞(n=3)、3T3-L1白色前脂肪细胞(n=3)、分化的3T3-L1白色脂肪细胞(n=3)。此数据集的平台是GPL1261,该平台是Affymetrix公司开发的[Mouse430_2]Affymetrix Mouse Genome 430 2.0 Array。

1.2方法。(1)差异基因的获得：使用软件R Studio中的程序包识别差异基因(Differently expressed genes,DEGs)。对棕色前脂肪细胞、分化的棕色脂肪细胞、3T3-L1 白色前脂肪细胞、分化的3T3-L1白色脂肪细胞的差异基因进行筛选时,选用错误发现率FDR(False Discovery Rate)小于0.05,以及log2FC的绝对值大于2,即差异表达倍数值Fold Change大于2为标准进行筛选。(2)蛋白互作网络：应用String在线数据库(STRING:functional protein association networks)获得差异基因的蛋白互作(protein-protein interaction,PPI)网络。最终使用Cytoscape中Cytoscape hub的不同算法选取出前10位hub基因。

2 结果

2.1差异基因的分析结果。使用软件R Studio的程序包对数据集GSE185481进行分析，获得了681个差异基因，其中上调基因有416个，下调基因有265个；得出了火山图(图1)和热图(图2)。

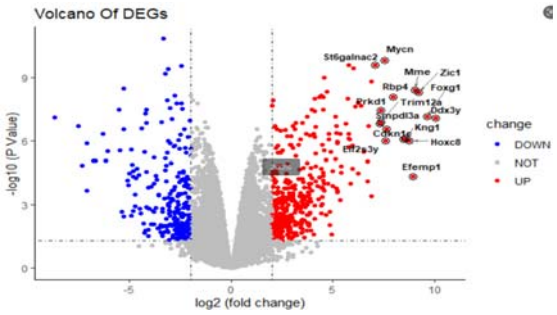

图 1

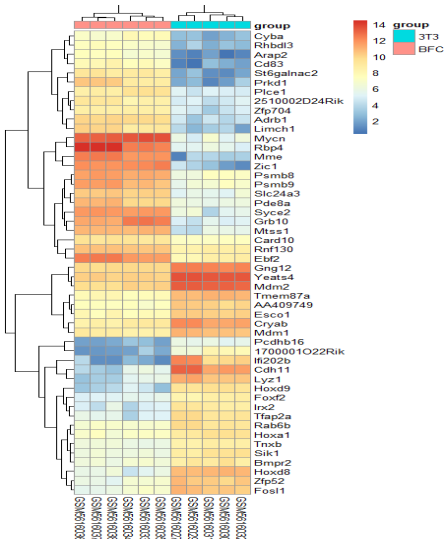

图 2

2.2差异表达基因的GO富集和KEGG通路富集分析结果

使用Bioconductor包中R工具的函数及ggplot程序包，R函数得出GO(图3)和KEGG图(图4)。

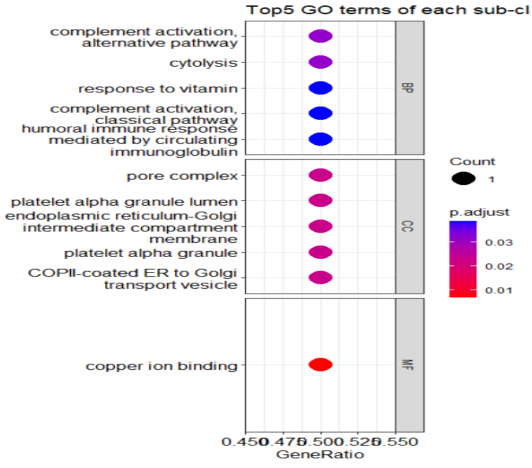

图 3

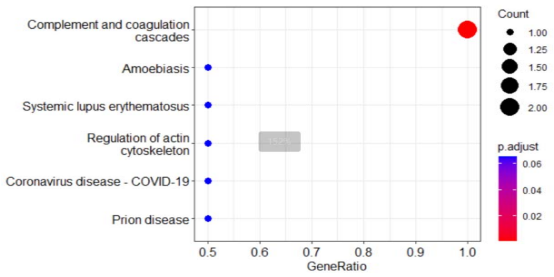

图 4

2.3差异表达基因的PPI网路及重要基因

使用STRING和CYTOSCAPE构建了PPI网络来理解ODEG的相互作用。然后，执行CYTONAC插件，获取复杂PPI网络中的集群功能模块。此外，为了鉴定该PPI网络中高度连接的基因，进行了cytoHubba插件，并在每种计算方法【MNC(图5)和EPC(图6)】中选择前10个基因，鉴定出7个枢纽基因。分别为Mki67、Ccna2、Aurka、TOP2A、Cdc20、MMP9和Melk。

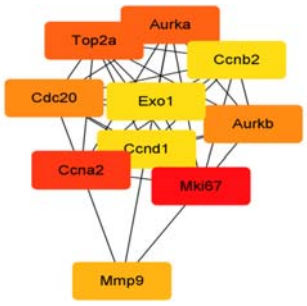

图 5

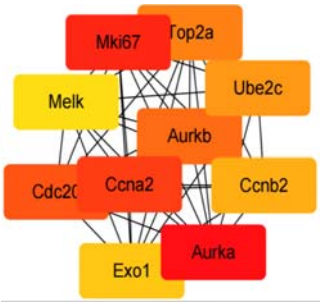

图 6

3 讨论

在本研究中，通过挖掘棕色前脂肪细胞、分化的棕色脂肪细胞、3T3-L1白色前脂肪细胞、分化的3T3-L1白色脂肪细胞数据集，进行了相关的生物信息学分析，研究脂肪分化过程中差异基因的表达。

本研究使用STRING和CYTOSCAPE构建了PPI网络。然后，执行CYTONAC插件，获取复杂PPI网络中的集群功能模块。并进行了cytoHubba插件，使用四种计算方法(closeness，degree，EPC和MNC)中选择前10个基因，鉴定出7个枢纽基因。分别为Mki67、Ccna2、Aurka、TOP2A、Cdc20、MMP9和Melk。

Mki67是本研究中发现的枢纽基因之一，是细胞增殖的标志物，Mki67基因通过核糖核酸的核糖体合成与细胞增殖相关。MKI67抗原存在于细胞周期的活性期(G1、S、G2 和有丝分裂)，而在另一个期(G0)不存在。Mki67是细胞增殖和肿瘤侵袭性的极好标志物，已发现它在肥胖期间响应脂肪组织扩张中具有更高的表达水平[7]。细胞周期蛋白2(Ccna2)是一种与细胞周期相关的基因，已被建议作为低级别胶质瘤的可能分子标记物[8]。有研究表明，支链氨基酸通过NADPH-FTO-m降低CDK2和Ccna2的表达来抑制肥胖和脂肪生成[9]。3T3-L1细胞中的FTO过表达和YTHDF2敲低显著提高了CDK2和Ccna2蛋白水平，并改善了表没食

子儿茶素没食子酸酯诱导的脂肪生成抑制。因此,m6FTO和YTHDF2介导依赖CDK2和Ccna2的表达,有助于表没食子儿茶素没食子酸酯诱导的脂肪生成抑制[10]。Aurora-A激酶是一种参与有丝分裂的丝氨酸/苏氨酸有丝分裂激酶,肠道上皮中的Aurka丢失导致肠道菌群失调和更高水平的短链脂肪酸,尤其是丙酸盐,导致AKT激活和脂质代谢调节基因表达,进而促进年龄诱导的肥胖[11]。据报道,TOP2A是活性增殖细胞的敏感和特异性标志物[12]。TOP2A可以在肥胖失调的转录组中发挥战略作用[13]。有研究表明Cdc20的表达在超重和肥胖女性的卵丘人颗粒卵丘细胞中上调[14]。MMP9是基质金属蛋白家族的一员,它在各种生理过程中起着至关重要的作用,包括胚胎发育、繁殖、血管形成、骨骼发育、伤口愈合、细胞迁移、学习和记忆[15]。MMP9可能在预防和治疗肥胖方面具有重要的生物学作用[16]。Melk是一种AMP活化蛋白激酶(AMPK)相关激酶,与肥胖相关的代谢缺陷有关。Smad蛋白通过差异调节饮食诱导的肥胖小鼠的MPK38活性,对肥胖相关代谢具有类别特异性影响[17]。

肥胖所引起的代谢型紊乱,会导致多种疾病,因此找到预防和治疗肥胖的靶点至关重要。通过对GSE185481进行分析,本研究证明Mki67、Ccna2、Aurka、TOP2A、Cdc20、MMP9和Melk为预防或治疗肥胖的潜在靶点。这些hub基因对肥胖的潜在作用,仍需要细胞或动物实验进一步确认。

[参考文献]

[1]Mayoral L P C, Andrade G M, Mayoral E P C, et al. Obesity subtypes, related biomarkers & heterogeneity[J].Indian Journal of Medical Research,2020,151(1):11-21.

[2]Bray G A. Evaluation of obesity: who are the obese?[J]. Postgraduate medicine,2003,114(6):19-38.

[3]Fuentes H A C.From basic mechanisms to clinical applications in heart protection, new players in cardiovascular diseases and cardiac theranostics: meeting report from the third international symposium on “New frontiers in cardiovascular research”[J].2016.

[4]Cabrera-Fuentes H A, Alba-Alba C, Aragones J, et al. Meeting report from the 2nd International Symposium on New Frontiers in Cardiovascular Research. Protecting the cardiovascular system from ischemia:: between bench and bedside[J]. Basic Research in Cardiology,2015,111.

[5]Sarafidis M,Lambrou G I,Zoumpourlis V,et al.An integrated bioinformatics analysis towards the identification of diagnostic, prognostic, and predictive key biomarkers for urinary bladder cancer[J].Cancers,2022,14(14):3358.

[6]Sai Swaroop R, Akhil P S, Sai Sanwid P, et al. Integrated multi-omic data analysis and validation with yeast model show oxidative phosphorylation modulates protein aggregation in amyotrophic lateral sclerosis[J].Journal of Biomolecular Structure and Dynamics,2023,41(12):5548-5567.

[7]Haim Y, Blüher M, Slutsky N, et al. Elevated autophagy gene expression in adipose tissue of obese humans: A potential non-cell-cycle-dependent function of E2F1[J].Autophagy,2015,11(11):2074-2088.

[8]Qi C, Lei L, Hu J, et al. Serine incorporator 2 (SERINC2) expression predicts an unfavorable prognosis of low-grade glioma(LGG):evidence from bioinformatics analysis[J]. Journal of Molecular Neuroscience,2020,70(10):1521-1532.

[9]Huang C,Luo Y,Zeng B, et al. Branched-chain amino acids prevent obesity by inhibiting the cell cycle in an NADPH-FTO-m6A coordinated manner[J]. The Journal of Nutritional Biochemistry,2023,122:109437.

[10]Wu R, Yao Y, Jiang Q, et al. Epigallocatechin gallate targets FTO and inhibits adipogenesis in an mRNA m6A-YTHDF2-dependent manner[J].International journal of obesity, 2018, 42(7):1378-1388.

[11]Sun N, Meng F, Zhao J, et al. Aurka deficiency in the intestinal epithelium promotes age-induced obesity via propionate-mediated AKT activation[J].International journal of biological sciences,2021,17(5):1302.

[12]DArcy N, Gabrielli B. Topoisomerase II inhibitors and poisons,and the influence of cell cycle checkpoints[J]. Current medicinal chemistry,2017,24(15):1504-1519.

[13]Nuncia-Cantarero M,Martinez-Canales S, Andrés-Pretel F, et al. Functional transcriptomic annotation and protein-protein interaction network analysis identify NEK2, BIRC5, and TOP2A as potential targets in obese patients with luminal A breast cancer [J].Breast cancer research and treatment,2018,168:613-623.

[14]Merhi Z, Polotsky A J, Bradford A P, et al. Adiposity alters genes important in inflammation and cell cycle division in human cumulus granulosa cell[J]. Reproductive Sciences, 2015,22(10):1220-1228.

[15]Mondal S,Adhikari N, Banerjee S,et al.Matrix metalloproteinase-9(MMP-9) and its inhibitors in cancer:A minireview [J].European journal of medicinal chemistry,2020,194:112260.

[16]Yuan K, Hu D, Mo X, et al. Uncovering the pathogenesis of obesity complicated with papillary thyroid carcinoma via bioinformatics and experimental validation[J]. Aging (Albany NY),2023,15(17):8729.

[17]Seong H A,Manoharan R,Ha H.Smad proteins differentially regulate obesity-induced glucose and lipid abnormalities and inflammation via class-specific control of AMPK-related kinase MPK38/MELK activity[J].Cell death & disease,2018,9(5):471.

作者简介：

肖昭晖(2000--),女,汉族,湖南长沙人,硕士,研究方向：内分泌。