线粒体自噬与肺腺癌

# 1. 背景

线粒体自噬是自噬的一种，是细胞维持能量平衡和稳态的重要方式，这个过程主要由高度保守的一些基因调控，这些基因通称为线粒体自噬相关基因[[Wang et al., 2020](#ref-Wang2020)]。目前的研究显示线粒体自噬异常与多种肿瘤的发生相关[[Kitada et al., 1998](#ref-Kitada1998); [Liu et al., 2018](#ref-Liu2018); [Chang et al., 2017](#ref-Chang2017)]。研究发现线粒体自噬相关基因PINK1激活的线粒体自噬能够增强肺腺癌的耐药性，从而促进病情的发展[[Li et al., 2022](#ref-Li2022)]。但是其他线粒体自噬相关基因与肺腺癌发生与发展之间的关联目前还不清楚。本研究旨在通过多组学和机器学习的方法筛选与肺腺癌病情进展相关的线粒体自噬相关的基因。

# 2. 数据

## 2.1 线粒体自噬相关基因

使用关键词mitophagy在Gene Ontology，KEGG pathway和reactome pathway中检索线粒体自噬相关基因，并取并集。

## 2.2 肺腺癌测序数据

| Dataset ID | 样本类型 | 样本数 | 样本特征 | 数据类型 |
| --- | --- | --- | --- | --- |
| *GSE267108* | 肺腺癌 | 8 | 4个样本不表达PD-1，2个低表达，2个高表达 | single cell |
| GSE256091 | 肺腺癌 | 141 | 非吸烟患者的肺腺癌测序数据，包含EGFR，ALK不突变数据 | bulk RNAseq |
| *GSE253013* | 肺腺癌 | 89 | 包含肺腺癌及其对应的癌旁组织，以及富集的CD45+的免疫细胞 | single cell |
| GSE164789 | 肺腺癌 | 62 | 肺腺癌患者早期测序数据 | single cell |
| GSE165298 | 肺腺癌 | 6 | 3个肺腺癌和3个配对的癌旁组织 | bulk RNAseq |
| GSE10072 | 肺腺癌 | 107 | 包括吸烟，不吸烟和以前吸烟的肺腺癌样本 | bulk RNAseq |
| *GSE11969* | 肺腺癌 | 163 | 包括非小细胞肺癌和肺腺癌 | bulk RNAseq |
| GSE17475 | 肺腺癌 | 28 | 包括吸烟和不吸烟的肺腺癌样本 | bulk RNAseq |
| 200013947 | 肺腺癌 | 12 | 肺腺癌测序样本 | bulk RNAseq |
| GSE229705 | 肺腺癌 | 246 | 肺腺癌和对应的癌旁组织 | bulk RNAseq |
| GSE229253 | 肺腺癌 | 33 | 肺腺癌和对应的癌旁组织 | snRNAseq |
| GSE119267 | 肺腺癌 | 155 | 肺腺癌 | microarray |
| GSE90623 | 肺腺癌 | 211 | 手术后复发和非复发的StageI肺腺癌 | microarray |
| GSE46539 | 肺腺癌 | 230 | 115个肺腺癌和对应的癌旁组织 | microarray |
| GSE71181 | 肺腺癌 | 284 | 吸烟患者的肺腺癌和正常肺组织 | microarray |
| *GSE72094* | 肺腺癌 | 442 | 包括吸烟和不吸烟肺腺癌，以及KRAS, TP53突变信息 | microarray |
| *TCGA-LUAD* | 肺腺癌 | 519 | 肺腺癌 | bulk RNAseq |

# 3. 分析流程

*Figure1：* 分析流程图。

*Figure2：* 肺腺癌单细胞测序数据（GSE253013）质控、去批次、降维聚类以及细胞注释后的TSNE图。每种细胞在不同样本中的分布图。每种细胞marker基因表达气泡图和小提琴图。每种细胞前20个marker基因的热图。

*Figure3：* 取所有细胞的差异基因与线粒体自噬基因的交集(DE-MRGs)，并绘制韦恩图。绘制DE-MRGs表达相关性矩阵热图。绘制DE-MRGs在不同细胞中表达水平的TSNE图。

*Figure4：* 利用AUCell计算DE-MRGs在不同细胞中的活性值，绘制出所有细胞活性值的直方图，并选取一个临界值将所有的细胞划分成DE-MRGs高表活性和低活性两群，在TSNE图中标记出这两群细胞。分析这两群细胞的差异基因，并进行GO和KEGG富集分析。

*Figure5：* 将上皮细胞提取出来重新进行降维、聚类。按照已有的marker基因把上皮细胞划分成多个亚群，并绘制各个亚群marker基因的小提琴图和气泡图。以及绘制DE-MRGs在不同亚型上皮细胞中表达水平的TSNE图。对上皮细胞进行拟时序分析，并绘制各个亚型的上皮细胞沿着拟时序分布情况和各个细胞的拟时序值。

*Figure6：* 在上皮细胞中绘制DE-MRGs沿着拟时序表达水平变化的散点图。

*Figure7：* 对所有的大类细胞和各个亚群的上皮细胞进行细胞通讯分析。并绘制代表细胞间相互作用强弱（相互作用的个数，参与相互作用的ligand-receptor的表达水平之和）的热图。绘制出在肿瘤和癌旁组织中上皮细胞与其他细胞相互作用有差异的配体受体对，以及各个上皮细胞亚型之间在肿瘤和癌旁组织中有差异的受体配体对。

*Figure8：* 使用DE-MRGs对TCGA-LUAD进行一致性聚类分析。对不同亚群进行差异分析，绘制火山图，并检查DE-MRGs在各个亚群中属于高表达还是低表达。取所有亚群差异基因与所有单细胞差异基因的交集并绘制韦恩图，然后使用交集基因进行GO富集分析。

*Figure9：* 绘制出一致性聚类的结果图、DE-MRGs在不同亚群中的表达水平、不同亚群的PCA图、不同亚群免疫浸润水平的箱线图以及不同免疫细胞在不同亚群中的免疫浸润水平。

*Figure10：* 使用Cox和Lasso分析从TCGA-LUAD样本中筛选与生存相关的DE-MRGs。使用筛选到的DE-MRGs基因年龄和病情分级信息计算每个样本的风险值。利用风险值将合并后的样本划分成高风险组和低风险组。使用相同的方法计算出验证数据集GSE72094每个样本的风险值并进行生存分析，绘制出1、3、5年的生存曲线。对DE-MRGs进行生存分析，绘制DE-MRGs的风险比率图和列线图。

*Figure11：* 确认Cox与Lasso筛选到的DE-MRGs在肿瘤组织和癌旁组织中的表达差异（mRNA水平:RT-PCR, 蛋白水平：IHC）。并对高表达和低表达Cox与Lasso筛选筛选到的DE-MRG进行生存分析。计算Cox与Lasso筛选到的DE-MRGs的表达水平与各个免疫细胞浸润水平的相关系数以及不同免疫检查点基因的相关系数。计算不同免疫细胞在TCGA-LUAD高风险组与低风险组中的浸润水平。

*Figure12：* 下载TCGA-LUAD的突变数据，并画出高风险组和低风险组前20个高突变基因的突变信息。计算高风险组和低风险组中每个样本的肿瘤突变负荷，并绘制箱线图。对TCGA-LUAD样本中高突变组与低突变组进行生存分析（使用突变负荷中位数分组）。下载TCGA-LUAD拷贝数变异数据计算染色体每个位置的G-score，并绘制出高风险组和低风险组每个染色体位置的G-score。分别绘制出微卫星不稳定分数与突变负荷、 肿瘤免疫功能障碍与排除、肿瘤免疫功能障碍与微卫星不稳定的散点图。

# 4. 附表

1. 线粒体自噬相关基因。
2. 所有细胞的差异基因列表。
3. GP和KEGG富集分析结果。
4. 一致性聚类每个样本所属的cluster列表。
5. 免疫浸润分析结果。

**该设计方案参考这篇文章[**[**Wang et al., 2023**](#ref-Wang2023)**], IF=7.7**

# 参考文献

Chang JY, Yi H-S, Kim H-W, Shong M. 2017. [Dysregulation of mitophagy in carcinogenesis and tumor progression](https://doi.org/10.1016/j.bbabio.2016.12.008). Biochimica et Biophysica Acta (BBA) - Bioenergetics 1858: 633–640.

Kitada T, Asakawa S, Hattori N, Matsumine H, Yamamura Y, Minoshima S, Yokochi M, Mizuno Y, Shimizu N. 1998. [Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism](https://doi.org/10.1038/33416). Nature 392: 605–608.

Li Y, Chen H, Xie X, Yang B, Wang X, Zhang J, Qiao T, Guan J, Qiu Y, Huang Y-X, Tian D, Yao X, Lu D, Koeffler HP, Zhang Y, Yin D. 2022. [PINK1-mediated mitophagy promotes oxidative phosphorylation and redox homeostasis to induce drug-tolerant persister cancer cells](https://doi.org/10.1158/0008-5472.can-22-2370). Cancer Research 83: 398–413.

Liu Y, Yan J, Sun C, Li G, Li S, Zhang L, Di C, Gan L, Wang Y, Zhou R, Si J, Zhang H. 2018. [Ameliorating mitochondrial dysfunction restores carbon ion-induced cognitive deficits via co-activation of NRF2 and PINK1 signaling pathway](https://doi.org/10.1016/j.redox.2018.04.012). Redox Biology 17: 143–157.

Wang Y, Liu H-H, Cao Y-T, Zhang L-L, Huang F, Yi C. 2020. [The role of mitochondrial dynamics and mitophagy in carcinogenesis, metastasis and therapy](https://doi.org/10.3389/fcell.2020.00413). Frontiers in Cell and Developmental Biology 8.

Wang Z, Chen C, Ai J, Shu J, Ding Y, Wang W, Gao Y, Jia Y, Qin Y. 2023. [Identifying mitophagy-related genes as prognostic biomarkers and therapeutic targets of gastric carcinoma by integrated analysis of single-cell and bulk-RNA sequencing data](https://doi.org/10.1016/j.compbiomed.2023.107227). Computers in Biology and Medicine 163: 107227.