Multi-Omics Data In The Prediction Of Kidney Cancer Subgroups

EC6070 - Computer Engineering Research Project 1

By:

Maduranga W.P.N. (2018/E/073)

Rodrigo S.M. (2018/E/102)

Supervisor: Dr. Pratheeba J.

Co-Supervisor: Dr. Thuseethan S.

Annotated Bibliography

Submitted By: Maduranga W.P.N. (2018/E/073)

3rd Aug 2022

Table of Contents

Annotated Bibliography 3
Annotation 01 3
Annotation 02 4
Annotation 03 5
Annotation 04 6
Annotation 05 7
Appendix 8

Annotated Bibliography

Annotation 01

(1) Muhamed Ali, Ali, Hanqi Zhuang, Ali Ibrahim, Oneeb Rehman, Michelle Huang, and Andrew Wu. 2018. "**A Machine Learning Approach for the Classification of Kidney Cancer Subtypes Using miRNA Genome Data**" Applied Sciences 8, no. 12: 2422. https://doi.org/10.3390/app8122422

(2) In this paper, they proposed a machine learning approach for the classification of kidney cancer subtypes using miRNA genome data. (3) The authors downloaded data from TCGA, found the most discriminative miRNAs and then used a machine learning tool to group kidney cancer five subtypes. (4) Their research focused on kidney cancer sub-type detection and classification in an effort to assist researchers in medicine to address the key points of kidney subtypes and their characteristics. (5) The article is useful to our research topic, as authors suggested that the identified miRNAs in this study can be used as biomarker candidates for kidney cancer subtype classification. (6) The main limitation of this article is that the effectiveness of these selected miRNAs were not validated, (7) thus the authors indicated that further, the effectiveness must be validated by wet-lab experiments or further clinic studies. (8) The article is forming the basis of our research; since authors did the same research as ours and only difference is using mRNAs genome data instead of multi-omics data.

Annotation 02

(1) Lovino, M., Bontempo, G., Cirrincione, G., Ficarra, E. (2020). "**Multi-omics Classification on Kidney Samples Exploiting Uncertainty-Aware Models**". In: Huang, DS., Jo, KH. (eds) Intelligent Computing Theories and Application. ICIC 2020. Lecture Notes in Computer Science (), vol 12464. Springer, Cham. https://doi.org/10.1007/978-3-030-60802-6_4

(2) In this paper, they proposed a method consisting of a tree-based multi-layer perceptron (MLP), which estimates the class-membership probabilities for classification. (3) The authors downloaded data from Genomic Data Commons (GDC) database for kidney tumor subtypes, selected samples for which all mRNA, miRNA and methylation data are available. Some stomach samples (do not belong to the kidney classes) were obtained from GDC for test the model. They proposed a tree MLP classifier and it was compared with support vector machine (SVM), random forest (RF) classifier, standard MLP and Bayesian neural network (BNN) classifiers to give relevance to all the omics and also to label as Unknown those samples for which the classifier is uncertain in its prediction. (4) Their research focused on creating automatic tools to integrate different omics information, which may favor clinical practice. (5) The article is useful to our research topic, as authors used SVM and RF for the same problem and those showed high classification rates. (6) The main limitation of this article is the standard consensus is given by the absence of a measure to check the relevance of each individual omics in the classification. (7) Authors indicated that tree MLP architecture is reliable for classification on the individual omics exploiting uncertainty-aware models. (8) This research is on the same domain and gave an idea about Unknown sample's affect in multi omics.

Annotation 03

(1) Hu F, Zeng W, Liu X. **A Gene Signature of Survival Prediction for Kidney Renal Cell Carcinoma by Multi-Omic Data Analysis**. International Journal of Molecular Sciences. 2019; 20(22):5720. https://doi.org/10.3390/ijms20225720

(2) In this paper, they performed a multi-omics analysis to build a multi-gene prognosis signature for Kidney renal cell carcinoma (KIRC). (3) The authors downloaded multiplatform genomics datasets from The Cancer Genome Atlas (TCGA), identified 863 differentially expressed genes (DEGs) with an altered DNA methylation status, found 189 methylated differentially expressed genes (MDEGs) as prognosis-related genes, selected 7 and generate a risk score to predict prognosis based on the expression of the seven genes on KIRC patient. (4) Their research focused on methylation and expression profiles to provide a reliable prognostic model for KIRC patients. (5) The article is useful to our research topic, as authors suggested seven-MDEG signature for predicting survival in KIRC patients. (6) The main limitation of this paper is that the signature was only validated on the TCGA cohort. Thus authors indicate that this signature needs to be further investigated in multiple datasets with different populations. (7) Their results showed that their signature is an independent prognostic factor in KIRC patients and can more accurately predict overall survival in KIRC patients than a tumor stage system. (8) This study is on one of the kidney cancer subgroups, hence gives an idea on kidney cancer related studies such as DEGs and MDEs.

Annotation 04

(1) He, Z., Liu, H., Moch, H. et al. **Machine learning with autophagy-related proteins for discriminating renal cell carcinoma subtypes**. Sci Rep 10, 720 (2020). https://doi.org/10.1038/s41598-020-57670-y

(2) In this paper, they tested the possibility of using numeric data acquired from software-based quantification of certain marker proteins, key autophagy proteins (ATGs), obtained from immunohistochemical (IHC) images of renal cell carcinomas (RCC). (3) The authors found a Tissue microarray (TMA) containing 237 RCCs from untreated patients and 18 normal kidney tissues from healthy donors belongs to Department of Pathology and Molecular Pathology, University and University Hospital Zurich and analyzed data using different data cleaning and machine learning techniques to classify the subtypes of RCCs. (4) Their research focused on indicating the potential for bioinformatics approaches in tumor classification based on the expression levels of certain ATGs in RCC. (5) The article is useful to our research topic, as authors suggested that histone methyltransferases and microRNA-145 may have diagnostic value for discrimination of certain subtypes of RCC. (6) The main limitation of this article is that most of the tested proteins could not differentiate papillary renal cell carcinomas (pRCC) from normal tissue, (7) thus the authors suggested that evaluation of autophagy and/or ATGs are less sufficient for pRCC prediction as compared with other subtypes. (8) Here authors did the same research as ours and only difference is using autophagy-related proteins instead of multi-omics data.

Annotation 05

(1) Eliana Marostica, Rebecca Barber, Thomas Denize, Isaac S. Kohane, Sabina Signoretti, Jeffrey A. Golden, Kun-Hsing Yu; **Development of a Histopathology Informatics Pipeline for Classification and Prediction of Clinical Outcomes in Subtypes of Renal Cell Carcinoma.** Clin Cancer Res 15 May 2021; 27 (10): 2868–2878. https://doi.org/10.1158/1078-0432.CCR-20-4119

(2) In this study, they developed informatics pipelines to connect Renal cell carcinoma (RCC) histopathology images with genomic information, clinical profiles, and biomarkers of response to immune checkpoint blockade. (3) The authors obtained whole-slide histopathology images and demographic, genomic and clinical data from The Cancer Genome Atlas (TCGA), the Clinical Proteomic Tumor Analysis Consortium, Brigham and Women's Hospital (Boston, MA), and developed fully automated convolutional neural networks to diagnose renal cancers and connect quantitative pathology patterns with patients' genomic profiles and prognoses. (4) Their research focused on predict the subtypes, prognoses, and genomic aberrations of patients with RCC using histopathology images to guide clinical decision making, improve patients' outcomes, and reduce the cost of cancer management. (5) The article is useful to our research topic, as authors suggested that integrated information from multiple modalities, including multi-omics are extensible to the histopathology evaluation of other complex diseases. (6) The main limitation of this article is that even they employed a large dataset in this study, may not capture the full spectrum of morphology heterogeneity in RCC, (7) authors indicated that further the performance needs to be evaluated in tumors with atypical histology manifestations. (8) This research used omics data of kidney cancer from TCGA data like our research for even subtyping a particular kidney cancer subgroup, where we are subgrouping the kidney cancer.

Appendix

[1] Citation

[2] Introduction

[3] Aims and Research methods

[4] Scope

[5] Usefulness

[6] Limitations

[7] Conclusions

[8] Reflection