



Review

Role of bioinformatics databases and tools in radiation biology

Yizhe Gao, Qingjie Liu *

China CDC Key Laboratory of Radiological Protection and Nuclear Emergency, National Institute for Radiological Protection, Chinese Center for Disease Control and Prevention, Beijing, 100088, China

ARTICLE INFO

Keywords:

Radiobiology
Bioinformatics
Radioresistance
Immune cell enrichment

ABSTRACT

Bioinformatics has become increasingly integral to radiation biology, also known as radiobiology, providing substantial support through data storage, conversion, visualization, and sharing. This review aims to deepen understanding of bioinformatics application in radiobiology by introducing key databases and analytical tools in radiobiology, including general bioinformatics databases, radiobiology-specific databases, data processing tools, and statistical analysis tools for differentially expressed genes (DEGs) and LC/MS analysis. This review also discusses bioinformatics applications in radiobiological fields, such as radioresistance and immune cell enrichment. Despite these advances, challenges such as data interoperability remain. Methods and projects to address these issues, such as GeCo and GMQL, are also examined.

1. Introduction

Radiobiology is the branch of biology concerned with the effect of ionizing radiation on organisms. It is also a field of clinical and basic medical science that involves the study of the health effects of radiation, and the application of biology in radiological techniques and procedures for diagnostics and treatment.¹ The recent advances in biotechnology and computation methods facilitate the development in the omics approach, which plays a pivotal role in radiobiological research with identification of radiation-induced biomarkers and pathway analysis. This review summarizes prevalent bioinformatics databases and tools in radiobiology, providing an overview of their application scenarios and the statistical methods behind them. We delve into applications of bioinformatics in radiobiology, and propose methods and ongoing projects that possibly resolve these problems. To ensure conciseness, statistical principles are outlined and only few application examples are included.

2. Bioinformatics databases utilized in radiobiology

Bioinformatics databases are supportive to radiation biological research in that it connects the data collected in radiation biology research to the genetic and radiotherapy data.² Even though some databases are not available for public, researchers with permissions can access to these various data to broaden understanding of ionizing radiation effects on biological systems and further influence radiation protection and medicine research in the world. This part will provide an

overview of major bioinformatics databases utilized in radiobiology research.

2.1. General public databases

2.1.1. The Cancer Genome Atlas (TCGA)

The Cancer Genome Atlas (TCGA) is a widely utilized comprehensive repository for cancer genomics.³ It offers clinical, transcriptomic, and genomic information pertaining to a wide variety of human malignancies. The TCGA contains information on 33 forms of cancer with details of tumor by categories including that of RNAseq, miRNAseq, DNA methylation, CNV, SNP, and so forth. TCGA data can be applied for identification of biomarkers associated with radiosensitivity or radioresistance. Genome-wide data is a valuable resource for studying radiation-induced mutation and clinical outcomes with radiological treatment in TCGA allow researchers to study correlation between genetic profiles with treatment outcomes, optimizing treatment plans for patients with the tumors.

The effective use of TCGA requires careful consideration of its complexity, access restrictions, and the demographic scope of its data. As TCGA data predominantly originate from U.S. populations, their applicability may be limited in global or ethnically diverse studies. This geographic and demographic concentration could affect the generalization of research findings to other populations, particularly in areas with genetic, environmental, or lifestyle differences.

* Corresponding author.

E-mail address: liuqingjie@nirp.chinacdc.cn (Q. Liu).