

IPCARF

IPCARF: improving lncRNA-disease association prediction using incremental principal component analysis feature selection and a random forest classifier

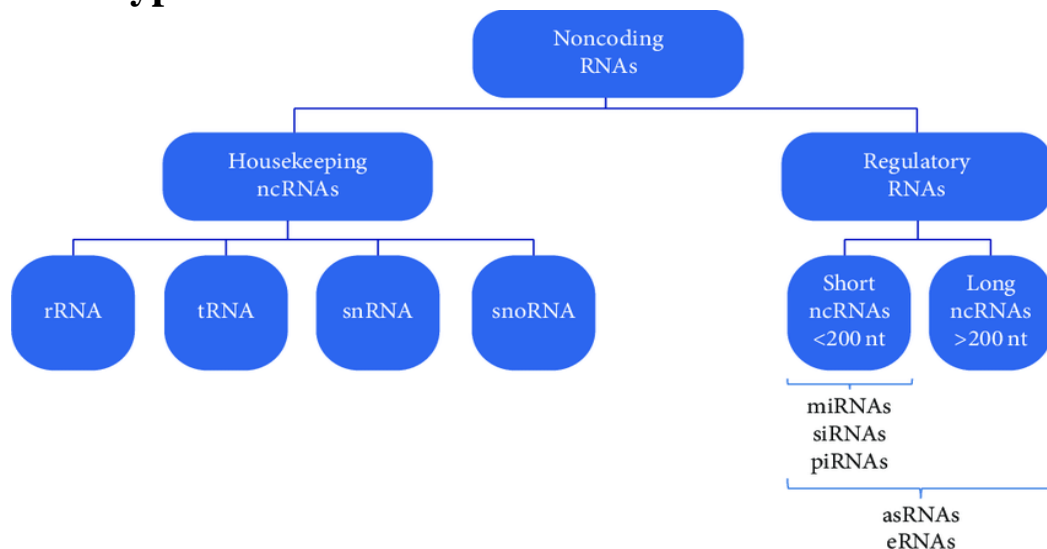
Abstract

Background

- **RNA:**

RNA is a vital molecule found in your cells, and it is necessary for life. Pieces of RNA are used to construct proteins inside of your body so that new cell growth may take place.

- **RNA types :**



- **RNA in diseases :**

There is important connections have been discovered between RNA and human disease. For example, as described previously, some miRNAs are capable of regulating cancer-associated genes in ways that facilitate tumor development. In addition, the dysregulation of miRNA metabolism has been linked to various neurodegenerative diseases, including Alzheimer disease.

- **lncRNA-disease :**

Studies discovered that lncRNAs play critical roles in a broad range of biological processes and are associated with a number of diseases, i.e. cancer, cardiovascular disease, and neurodegeneration disease, but

also speeds up the identification of potential biomarkers for disease diagnoses, treatments, and drug response predictions.

Result:

Using a combination of incremental principal component analysis (IPCA) and random forest (RF) algorithms and by integrating multiple similarity matrices, we propose a new algorithm (IPCARF) based on integrated machine learning technology for predicting lncRNA-disease associations.

Conclusions:

IPCARF which has shown excellent performance in predicting lncRNA-disease associations. Although the IPCARF method has achieved good prediction results, it still has some limitations that should be improved in future studies

Introduction:

However, as the amount of archived biological data continues to grow, it has become increasingly difficult to detect potential human lncRNA-disease associations from these enormous biological datasets using traditional biological experimental methods. Consequently, developing new and effective computational methods to predict potential human lncRNA diseases is essential. So, several publicly available databases have emerged that have introduced greater opportunities for predicting potential lncRNA-disease associations through computational methods. Using computational methods to predict potential lncRNA-disease associations is important, because such associations can effectively improve our understanding of disease pathogenesis and guide treatments. In this study, we proposed a novel model called IPCARF to predict lncRNA-disease associations and compared it with the existing LRLSLDA, LRLSLDA-LNCSIM, TPGLDA, NPCMF, and ncPred prediction methods using 10CV. These methods have achieved excellent performances for predicting lncRNA-disease associations. The comparison results show that the prediction results of the IPCARF method are better than those of the compared methods.