



Anchor Based Sequence Clustering Algorithm for Efficient & Accurate Motif Discovery

Tech ID: 27614 / UC Case 2017-608-0

BRIEF DESCRIPTION

A new strategy for speeding up motif discovery, an anchor based sequence clustering algorithm (ASC) that works significantly faster and with greater accuracy than current motif finding algorithms.

BACKGROUND

Finding sequence patterns from a set of protein sequences, referred to as motif discovery, is a critical component in the biological and medical sectors. As a result, motif discovery has been an important area of research. There has been a rise in instrumental techniques that make it possible to generate thousands of sequences at once. However while this accelerated method can work on a larger scale it sacrifices accuracy for speed depreciating the value of the data.

DESCRIPTION

Researchers at the University of California, Santa Barbara have created a new strategy for speeding up motif discovery, an anchor based sequence clustering algorithm (ASC) that works significantly faster and with greater accuracy than current motif finding algorithms. This new strategy groups sequences containing the same motif together, creating a subset of sequences. By limiting the input to subsets, traditionally time consuming algorithms can now be completed in a more efficient time span. The ASC can process millions of protein sequences within one day while currently available motif finding algorithms take multiple years. ASC is easily adaptable and can be used on any existing motif finding algorithm, including MEME.

ADVANTAGES

- ▶ Operates with improved accuracy
- ▶ Operates significantly faster
- ▶ Can be used in any existing motif finding algorithm

APPLICATIONS

- ▶ Motif Discovery
- ▶ Biological and medical sectors

CONTACT

Pasquale S. Ferrari
ferrari@tia.ucsb.edu
tel: .

INVENTORS

- ▶ Liu, Honglei
- ▶ Yan, Xifeng

OTHER INFORMATION

KEYWORDS

Motif discovery, algorithm,

anchor based sequencing,

indsoftw

CATEGORIZED AS

- ▶ **Computer**
- ▶ **Software**

RELATED CASES

2017-608-0

