

## **CaMoDi: a new method for cancer module discovery**

Alexandros Manolakos 1, Idoia Ochoa 1, Kartik Venkat 1, Andrea J. Goldsmith 1 and

Olivier Gevaert 2

1 Department of Electrical Engineering, Stanford University, USA

2 Department of Radiology, Stanford University, USA

### **Abstract**

#### **Background**

Identification of genomic patterns in tumors is an important problem, which would enable the community to understand and extend effective therapies across the current tissue-based tumor boundaries. With this in mind, in this work we develop a robust and fast algorithm to discover cancer driver genes using an unsupervised clustering of similarly expressed genes across cancer patients. Specifically, we introduce CaMoDi, a new method for module discovery which demonstrates superior performance across a number of computational and statistical metrics.

#### **Results**

The proposed algorithm CaMoDi demonstrates effective statistical performance compared to the state of the art, and is algorithmically simple and scalable - which makes it suitable for tissue-independent genomic characterization of individual tumors as well as groups of tumors. We perform an extensive comparative study between CaMoDi and two previously developed methods (CONEXIC and AMARETTO), across 11 individual tumors and 8 combinations of tumors from The Cancer Genome Atlas. We demonstrate that CaMoDi is able to discover modules with better average consistency and homogeneity, with similar or better adjusted  $R^2$  performance compared to CONEXIC and AMARETTO.

#### **Conclusions**

We present a novel method for Cancer Module Discovery, CaMoDi, and demonstrate through extensive simulations on the TCGA Pan-Cancer dataset that it achieves comparable or better performance than that of CONEXIC and AMARETTO, while achieving an order-of-magnitude improvement in computational run time compared to the other methods.