BITS :: Call for Abstracts 2019 - Oral communication

Type	Oral communication
Session	Algorithms in Bioinformatics
Title	Identification of active Mutational Signatures on transcript factor binding profiles
All Authors	Stamoulakatou E(1), Pinoli P(1)
Affiliation	

(1) Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milano

Motivation

All cancers carry somatic mutations which are caused by multiple mutational processes. Development of novel mathematical approaches for pan-cancer data analysis has identified mutational signatures i.e., patterns of somatic mutations caused by different mutational processes. Knowledge of mutational signatures is valuable for cancer treatment and cancer prevention. There are two conceptualizations of mutational signatures, with the most known model to be described by Alexandrov et al. [1] which separately counts all possible nucleotide triplets whose central base is mutated. Thus, each “Alexandrov signature” consists of 96 mutation probabilities that indicate which of the changes are occurring most frequently due to the mutational process it describes. Up to now, several bioinformatic packages have been developed to address the decomposition of a cancer genome’s mutation catalog into mutational signatures in order to shed more light about cancer etiology [2,3]. However, those methods cannot be used for the decomposition of specific local regions such as occurrences of a transcription factor motif (which disruption by mutations has an important impact on oncogenesis and tumor progression), since the background distribution of nucleotides in such regions may significantly differ from the one of the whole genome.

Methods

In this study, we extended the R package "decompTumor2Sig"[2] in order to identify the mutational signatures that dysregulate the transcription factor binding sites and get rid of the bias that the transcription factor motif context creates. Given a matrix S each of whose rows represents a signature and a vector g of the frequencies of mutations observed in an individual tumor, the "decompTumor2Sig" package uses quadratic programming to calculate a vector w that minimizes the error of Sw=g. Thus, w represents the contribution of each mutational signature to the individual tumor. Our contribution is to determine a correction factor to be applied to the individual frequency of mutations g before performing the minimization of the quadratic error. The correction factor is calculated for each transcription factor’s motif by dividing the frequencies of the possible triplets in the whole genome by the frequencies of the triplets induced by the position weight matrices (PWM) of a given motif.

Results

Here we show the application of our study to real data using mutations for skin cancer (MELA-AU) from ICGC portal [4] and the 19bp long Jaspar [5] MA0139.1 CTCF motif and illustrate its effectiveness to identify the mutational signatures which affect this specific motif. The figure shows the contribution of each signature considering mutations that fall in CTCF sites from 40 patients before and after applying the correction factor. As we can notice, without the correction factor the main signature is the 23 which is not associated with skin cancers but has very high probability because of the sequence context of the CTCF motif. Applying the correction factor we clearly see that signatures 7 and 11 are the main contributors and according to literature these signatures have been found predominantly in skin cancers.

References

[1] Alexandrov et al: Signatures of mutational processes in human cancer. Nature 500, 415-421 (2013).

[2]Krüger S, Piro RM decompTumor2Sig: Identification of mutational signatures active in individual tumors. BMC Bioinformatics (2019).

[3]Rosenthal R. et al. Deconstructsigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. Genome Biology (2016)

[4] Zhang J et al. International Cancer Genome Consortium Data Portal--a one-stop shop for cancer genomics data. Database (2011)

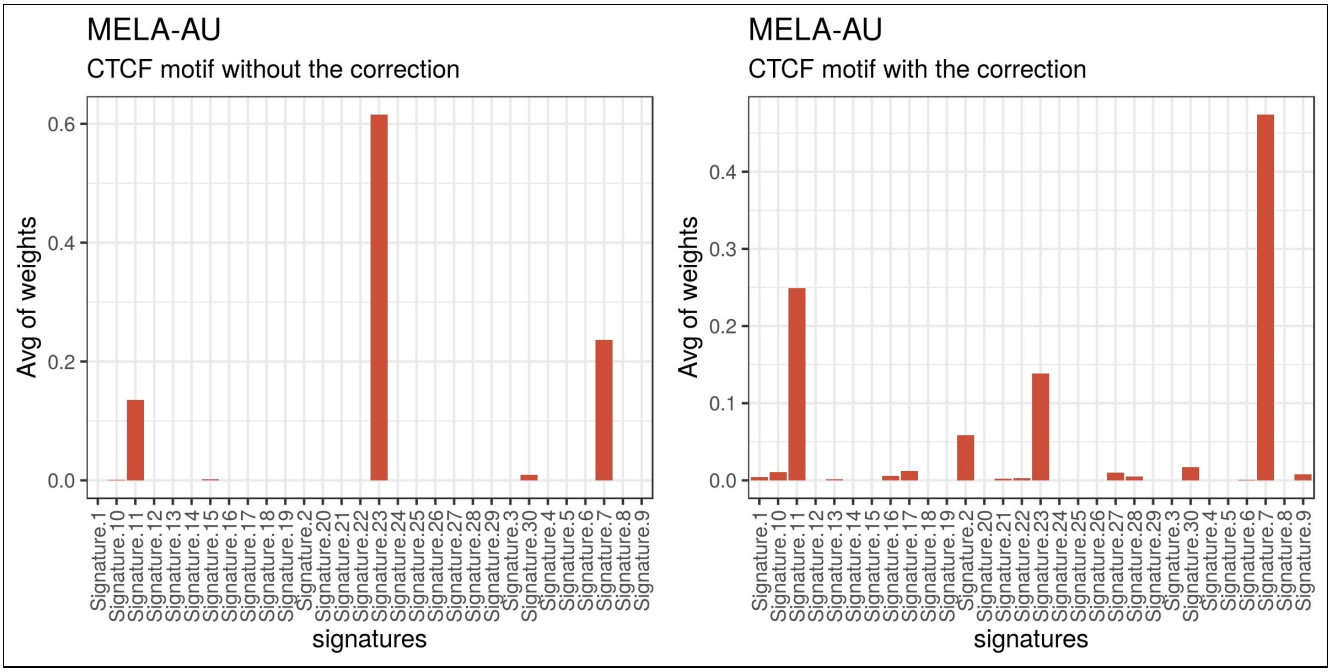
[5] A Khan et al: JASPAR 2018: update of the open-access database of transcription factor binding profiles

and its web framework, Nucleic Acids Research(2018)

Info

-

Figure



Availability	-

Corresponding Author

Name, Surname	Eirini, Stamoulakatou
Email	eirini.stamoulakatou@polimi.it
Submitted on	17.04.2019