

LIVRO DE RESUMOS



DÉCIMA PRIMEIRA SEMANA DA
GRADUAÇÃO E PÓS-GRADUAÇÃO DO
INSTITUTO DE FÍSICA DE SÃO CARLOS - USP

2021



Universidade de São Paulo
Instituto de Física de São Carlos

XI Semana Integrada do Instituto de
Física de São Carlos

Livro de Resumos

São Carlos
2021

Semana Integrada do Instituto de Física de São Carlos

SIFSC 11

Coordenadores

Prof. Dr. Vanderlei Salvador Bagnato

Diretor do Instituto de Física de São Carlos – Universidade de São Paulo

Prof. Dr. Luiz Vitor de Souza Filho

Presidente da Comissão de Pós Graduação do Instituto de Física de São Carlos – Universidade de São Paulo

Prof. Dr. Luís Gustavo Marcassa

Presidente da Comissão de Graduação do Instituto de Física de São Carlos – Universidade de São Paulo

Comissão Organizadora

Arthur Deponte Zutião

Artur Barbedo

Beatriz Kimie de Souza Ito

Beatriz Souza Castro

Carolina Salgado do Nascimento

Edgard Macena Cabral

Fernando Camargo Soares

Gabriel dos Reis Trindade

Gabriel dos Santos Araujo Pinto

Gabriel Henrique Armando Jorge

Giovanna Costa Villefort

Inara Yasmin Donda Acosta

Humberto Ribeiro de Souza

João Hiroyuki de Melo Inagaki

Kelly Naomi Matsui

Leonardo da Cruz Rea

Letícia Cerqueira Vasconcelos

Natália Carvalho Santos

Nickolas Pietro Donato Cerioni

Vinícius Pereira Pinto

Normalização e revisão – SBI/IFSC

Ana Mara Marques da Cunha Prado

Maria Cristina Cavarette Dziabas

Maria Neusa de Aguiar Azevedo

Sabrina di Salvo Mastrantonio

Ficha catalográfica elaborada pelo Serviço de Informação do IFSC

Semana Integrada do Instituto de Física de São Carlos
(11: 06 set. - 10 set. : 2021: São Carlos, SP.)
Livro de resumos da XI Semana Integrada do Instituto de
Física de São Carlos/ Organizado por João H. Melo Inagaki [et al.].
São Carlos: IFSC, 2021.

412 p.

Texto em português.

1. Física. I. Inagaki, João H. de Melo, org. II. Título

ISBN 978-65-993449-3-0

CDD 530

PG204

Predição de interações miRNA-mRNA com Predictive Bi-Clustering Trees

ILIDIO, P.¹; CERRI, R.²; THIEMANN, O.¹; SANTOS, R. A. C.¹; FRAJACOMO, H. C.²

ilidio@alumni.usp.br

¹Instituto de Física de São Carlos - USP²Universidade Federal de São Carlos - UFSCar

O projeto consiste em implementar e aperfeiçoar o algoritmo Predictive Bi-Clustering Trees (PBCT), proposto por Pliakos *et al.* em 2015 (1), e aplicá-lo à predição da existência de interação entre microRNAs (miRNAs) e RNAs mensageiros (MTI, *mirna-target interaction*), a partir de características extraídas computacionalmente das sequências biológicas. Os miRNAs são pequenas sequências (22 bp) de RNAs não codificantes capazes de desencadear a degradação dirigida de mRNAs, por meio do recrutamento de proteínas Argonautas e subsequente formação de complexos proteicos de silenciamento. (2) Trata-se de um importante mecanismo de regulação pós-transcricional da expressão gênica, com presença quase unânime nos processos biológicos que observamos. Nos humanos, estima-se computacionalmente que mais da metade dos genes codificadores de proteínas são alvos da regulação por miRNAs (2) e diversos fenótipos patológicos já foram associados a um funcionamento irregular desses processos, incluindo múltiplos tipos de câncer (2), de forma a sugerir o uso de miRNAs como biomarcadores e possíveis alvos terapêuticos. Atualmente, o principal parâmetro para identificação de alvos de miRNAs é o pareamento de suas bases 2 a 9, que configuram a região conhecida como *seed*, com o mRNA. Contudo, nota-se que não se trata de um padrão absoluto (3), não sendo necessária a completude do pareamento em muitos casos e observando-se participação de regiões mais próximas à ponta 3' em outros, especialmente para sítios-alvo em CDS, dado que a maioria das interações se faz presente na UTR 3' dos transcritos codificantes. (3) Softwares como o TargetScan levam em conta principalmente esses fatores de alinhamento, combinados a outras características genéricas das sequências, como o conteúdo GC, para prever pares miRNA-mRNA interagentes. O presente projeto objetiva a introdução de outras características mais complexas dessas sequências no processo preditivo, observando se a interação ainda se faz clara mesmo na ausência de alinhamentos, o que poderia trazer novos *insights* sobre o funcionamento das MTIs. Para este trabalho, os pares conhecidos de interação serão providos pela base de dados MiRTarBase, enquanto sequências dos microRNAs podem ser obtidas da miRBase. Como algoritmo preditivo, faz-se uso das PBCTs. Elas consistem em árvores de decisão regressivas que operam sobre uma matriz de interações, na qual cada miRNA é representado por uma coluna e cada transcrito alvo, por uma linha. (1) A presença ou ausência de interação são representadas na matriz por bits 1 e 0, respectivamente, nas posições da matriz correspondentes a cada par miRNA-mRNA em questão. O método se destaca pela eficiência no uso de memória e processamento ao utilizar o formato dos dados em duas dimensões, não precisando guardar os atributos de entrada uma vez para cada par miRNA-mRNA. Além disso, o modelo carrega toda a explicabilidade das árvores de decisão, de suma utilidade na tarefa de desvendar os processos biológicos.

Palavras-chave: microRNA. Predictive Bi-clustering tree. Interação miRNA-alvo. Extração de características de sequências biológicas.

Referências:

- 1 PLIAKOS, K.; GEURTS, P. VENS, C.. Global multi-output decision trees for interaction prediction. **Machine Learning**, v. 107, n. 8, p. 1257-1281, 2018.
- 2 JONAS, S.; IZAURRALDE, E.. Towards a molecular understanding of microRNA-mediated gene silencing. **Nature Reviews Genetics**, v. 16, n. 7, p. 421-433, 2015.
- 3 CHIPMAN, L. B.; PASQUINELLI, A. E. miRNA targeting: growing beyond the seed. **Trends in Genetics**, v. 35, n. 3, p. 215-222, 2019.