

# Análise da transcrição de material genético

Suzylaine da Silva Lima  
Orientador: Alexandre Ferreira Ramos

## 1 RESUMO

O estudo tem como objetivo analisar a dinâmica de transcrição de material genético. Tal mecanismo quer compreender como a oscilação ocorre ao longo do tempo e do espaço, e assim estruturar o mecanismo que coordena a construção de material genético.

## 2 INTRODUÇÃO

Com a evolução da tecnologia, novas ferramentas científicas surgiram, tal progresso propiciou o avanço em diversas frentes de estudos que demandam complexas e sofisticadas ferramentas para extrair e analisar informações como, no caso, dados de material genético.

Atualmente, há diversas frentes de estudos na área em questão, dentre elas a análise da dinâmica de produção de material genético. Buscando contribuir com o tema formou-se uma parceria com o laboratório Mounia Lagha Lab e Department of Applied Mathematics and Statistics, and Center for Developmental Genetics, na qual os laboratórios compartilham dados com a produção de material genético.

A base compartilhada contém a intensidade de fluorescência identificada durante o processo de transcrição de material genético de embriões de *Drosophila*. O inseto em questão foi selecionado, pois possui um ciclo de vida curto, fato que propicia gerar dados robustos em curto espaço de tempo.

Para gerar os insumos necessários a produção de material genético utilizou fluorescência para inferir a sua intensidade ao longo do tempo, pois tal técnica permite identificar os momentos em que houve a ativação da transcrição do material genético. Os dados compartilhados entre os laboratórios representam

o mesmo estudo, de forma a identificar dois pontos de fluorescência no núcleo, pontos esses denominados por *spots*. Porém, a técnica para identificar os *spots* são diferentes, isso porque uma das bases conta com a produção da fluorescência a partir da inserção de proteínas no núcleo, tal proteína sofre fragmentação quando dá-se início ao processo de produção de material genético, fator que gera luminosidade ao longo da transcrição [3]. Já outra base de dados gera a fluorescência a partir de separação de camadas, sendo a reação com MS2 e PP7.

## 3 MÉTODOS

Para compreender a dinâmica da fluorescência ao longo do tempo foram observados alguns núcleos, para assim analisar se há algum tipo de padrão.

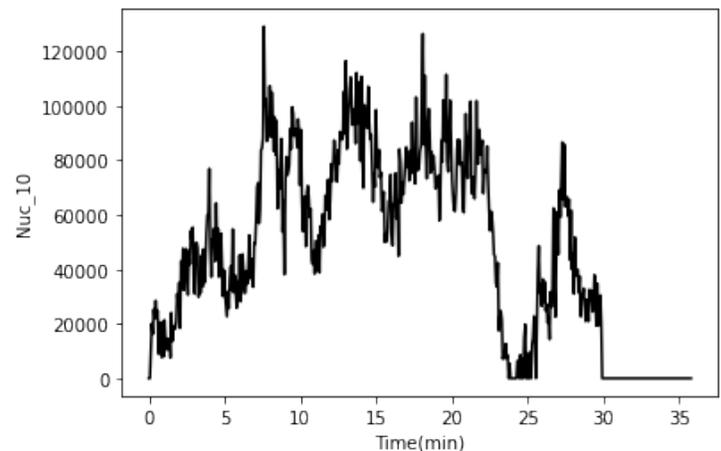


Figura 1. Gráfico que apresenta a intensidade da fluorescência (AU) ao longo do tempo (min) em um dos *spots* do núcleo. Dados da base que faz a identificação através da quebra por proteína.

A base de dados observada contém 4 embriões diferentes, sendo 3 gerados a partir de MS2 e PP7 e 1 gerado a partir da ruptura por proteína.

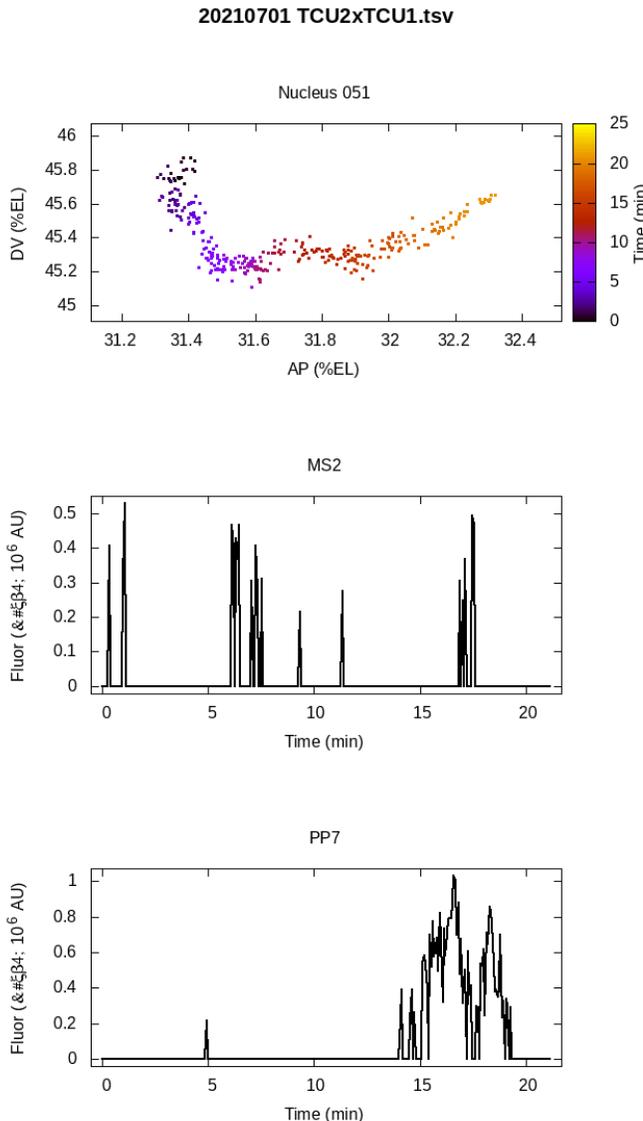


Figura 2. Gráficos que contém a movimentação dos *spots* ao longo do espaço; Intensidade da fluorescência(AU) do *spot* 01 ao longo do tempo(Min); Intensidade da fluorescência(AU) do *spot* 02 ao longo do tempo(Min).

Com observações preliminares nos núcleos foi possível observar a importância da exploração da intensidade de fluorescência, por conta

disso o estudo utilizou a técnica *K-means* de clusterização para entender a formação de grupos da fluorescência, inicialmente a técnica foi aplicada na base de dados gerada a partir da ruptura por proteína. A técnica de clusterização tem como objetivo montar grupos que maximizem a similaridade entre os seus elementos, assim minimizando o contato entre grupos que possuem características com pouca similaridade. O algoritmo *K-means* conta com uma técnica iterativa para fragmentar um grupo de dados. Tal técnica busca minimizar a distância dos registros de um conjunto de dados com  $k$  centros de forma iterativa.

## 4 RESULTADOS PRELIMINARES

Seguindo para análise de grupos da fluorescência a técnica de *K-means* contou com o  $k=3$ , valor adotado a partir método Elbow. Na Figura 3 é possível compreender que a técnica conseguiu separar bem os *spots*, de forma que a intensidade da fluorescência está vinculada ao momento que o material genético começou a ser transcrito, no caso, no primeiro ou no segundo momento.

Os resultados, a partir da base analisada, permitem concluir que de acordo com o momento em que a transcrição se deu é possível inferir qual a capacidade de produção de material genérico.

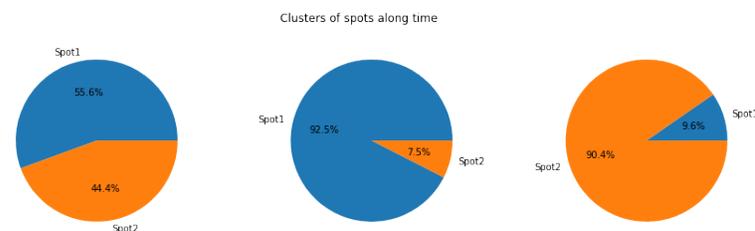


Figura 3. Gráficos que contém Distribuição dos grupos por *spots*

Como próximo passo está sendo analisado a distribuição espacial, pois assim será possível observar se é possível estabelecer tendência de fluorescência de acordo com a região e, assim, avançar no entendimento do fator de oscilação da fluorescência.

## 5 REFERENCIAS BIBLIOGRÁFICAS

[1] Pimmett, Virginia; Dejean, Matthieu; Fernandez, Carola; Trullo, Antonio; Bertrand, Edouard: Quantitative imaging of transcription in living *Drosophila* embryos reveals the impact of core promoter motifs on promoter state dynamics, 2021

[2] Guindalini, Camila; Tufik, Sergio: Use of microarrays in the search of gene expression patterns: application to the study of complex phenotypes, 2007

[3] Sloutskin, A. et al. ElemeNT: a computational tool for detecting core promoter elements. *Transcription* 6, 41–50 (2015)

[4] Luann Farias Palma, AGRUPAMENTO DE DADOS: K- MÉDIAS, 2018

[5] Bussab, WO; Morettin, PA. *Estatística Básica*. São Paulo: Editora Saraiva, 2006 (5ª Edição).