UNIVERSIDAD EAFIT

INGENIERÍA DE SISTEMAS

STO0263 TÓPICOS ESPECIALES EN TELEMÁTICA, 2020-1

GRUPOS 031 Y 032

PROYECTO 4 – HPC

**Transcripción ADN a ARN**

**Autor:** Juan Diego Saldarriaga – jsalda23@eafit.edu.co

**Repositorio:** <https://github.com/jsalda23/Proyecto-4-telematica>

**Google Drive:** https://drive.google.com/drive/folders/1-H0NcGYouu1cw6j-Xs-oeUgeZRQfHF2E?usp=sharing

**1. Problema o caso de estudio a resolver**

Durante le ejecución de este proyecto vamos a resolver el problema descrito en la página: <http://rosalind.info/problems/rna/>

Este problema consiste en realizar la transcripción de ADN a ARN. Está transcripción es el primer paso de la expresión genética (Proceso por el cual se codifican los ácidos nucleicos en proteínas). Este proceso consta de 3 etapas:

* **Iniciación:** La ARN polimerasa se une a una secuencia de ADN que se encuentra al inicio del un gen. Una vez unida, la ARN polimerasa separa las cadenas de ADN para proporcionar el molde de cadena sencilla necesario para la transcripción. [2]
* **Elongación:** El transcrito de ARN tiene la misma información que la cadena de ADN, pero contraria al molde en el gen. Sin embargo, a diferencia de la cadena ADN contiene la base uracilo (U) en lugar de timina (T). [2]
* **Terminación:** Una vez transcritas, estas secuencias provocan que el transcrito sea liberado de la ARN polimerasa. [2]

Una cadena ADN está compuesta por adenina (A), citosina (C), guanina (G) y timina (T), por lo tanto, para su fácil manejo y comprensión, se trabaja con sus siglas A, C, G y T. La cadena ARN a diferencia de la ADN está compuesta por A, C, G y uracilo (U). Sin embargo, cuando se pasa de ADN a ARN la T cambia por A, la A cambia por U, la C cambia por G, y la G cambia por C, así como se evidencia en la siguiente imagen:

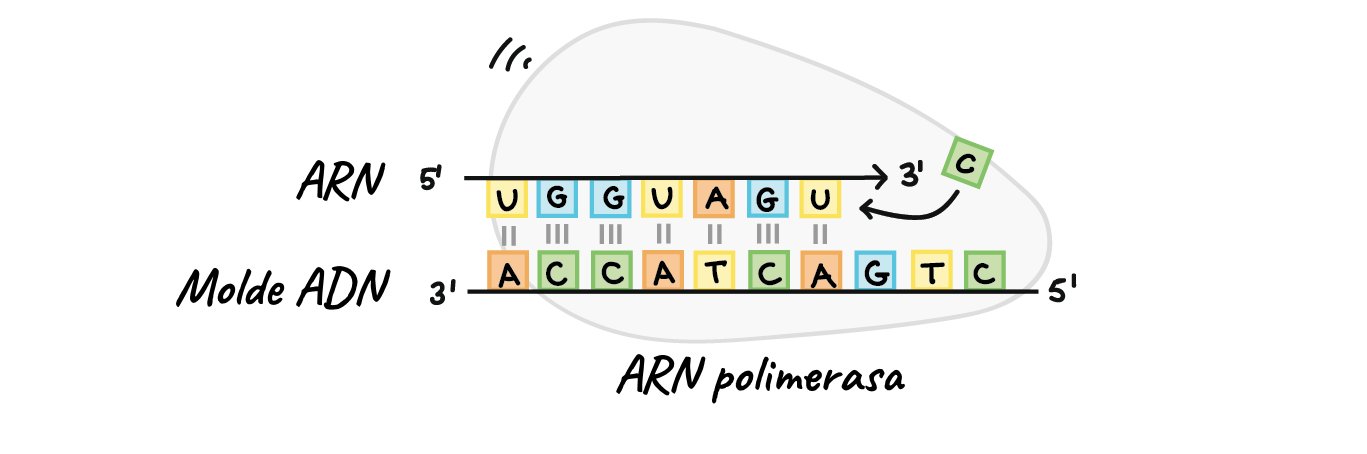


Imagen extraída de: https://es.khanacademy.org/science/biology/gene-expression-central-dogma/transcription-of-dna-into-rna/a/overview-of-transcription

**2. Objetivos y alcance**

**Alcance**

La solución será dividida en dos partes:

1. Programa secuencial, el cual nos servirá para establecer la línea base de la solución.
2. Diseño paralelo, el cual deberá estar en la capacidad de al recibir más de un archivo con código ADN y realizar simultáneamente su conversión a ARN.

**Objetivo Principal**

El objetivo principal del proyecto es desarrollar una solución para el problema descrito anteriormente, usando algoritmos paralelos en memoria compartida y MPI, bajo la metodología PCAM.

**Objetivos secundarios**

* Hallar los tiempos de procesamiento en paralelo.
* Hallar el *SpeedUp* y Eficiencia para varios escenarios de recursos.
* Realizar el análisis de rendimiento con varios números de procesadores.

**3.Requerimientos técnicos**

* Se hará el desarrollo en Python (esto porque su tiempo de aprendizaje es menor a c++).
* Se usará la librería mpi4py en Python.
* La implementación del algoritmo se hará el clúster del MOOC de Intel.
* Se usará GITHUB como repositorio, y deben quedar documentados todos los avances en cuanto a desarrollo y documentación.
* Se debe aplicar la técnica de diseño de algoritmo PCAM (particionamiento, comunicación, aglomeración, mapeo)
* Para realizar el análisis de rendimiento en paralelo, se deberá hacer con 2, 3 y 4 procesadores.

**4. Plan de trabajo**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Formulación del proyecto | Diseñar e implementar algoritmo secuencial | Diseñar e implementar algoritmo paralelo | Implementar en clúster | Análisis de rendimiento |
| Mayo 24 |  |  |  |  |  |
| Mayo 25 |  |  |  |  |  |
| Mayo 26 |  |  |  |  |  |
| Mayo 27 |  |  |  |  |  |
| Mayo 28 |  |  |  |  |  |
| Mayo 29 |  |  |  |  |  |
| Mayo 30 |  |  |  |  |  |
| Mayo 31 |  |  |  |  |  |
| Junio 1 |  |  |  |  |  |

**5. Referencias**

1. <http://rosalind.info/problems/rna/>
2. <https://es.khanacademy.org/science/biology/gene-expression-central-dogma/transcription-of-dna-into-rna/a/overview-of-transcription>
3. <https://www.genome.gov/es/genetics-glossary/ACGT>
4. <https://mpi4py.readthedocs.io/en/stable/>