

TALLER DE BIOCOMPUTACIÓN

ARN y grafos

M.C. Hugo Armando Guillén Ramírez
Grupo de Biocomputación, Ciencias de la Computación, CICESE.

 hugoagr@gmail.com / hguillen@cicese.edu.mx

 <https://hugoguillen.github.io>

 [/r/procrastinando](https://www.reddit.com/r/procrastinando)

INTRODUCCIÓN |

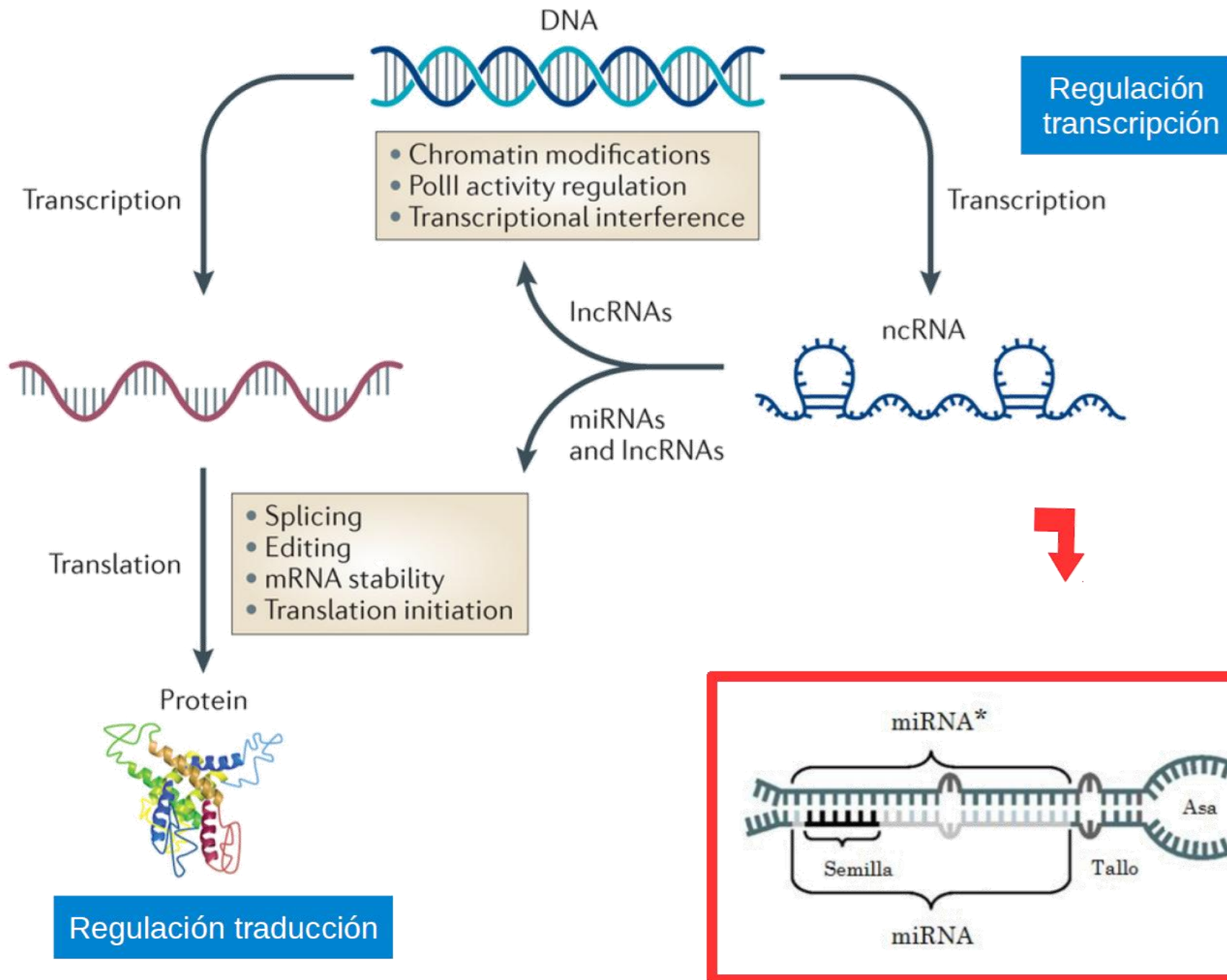


Imagen: ARN no codificante

CLASIFICACIÓN

Etiquetar ejemplos **desconocidos** a partir de ejemplos **conocidos** (ya etiquetados).



Miniature Pinscher



Miniature Schnauzer



Norfolk Terrier



Poodle
(Toy / Miniature)



?



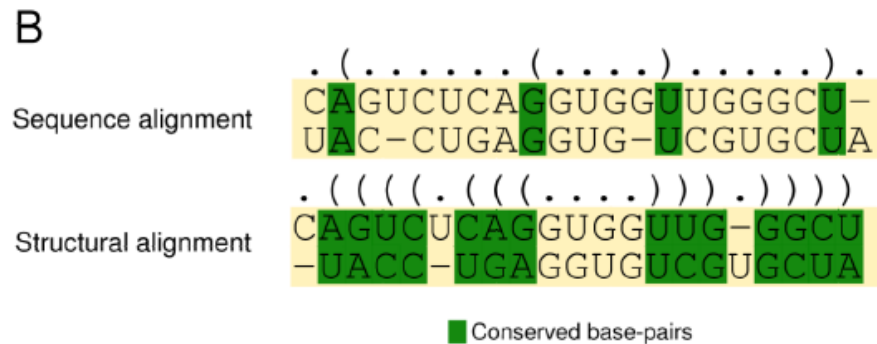
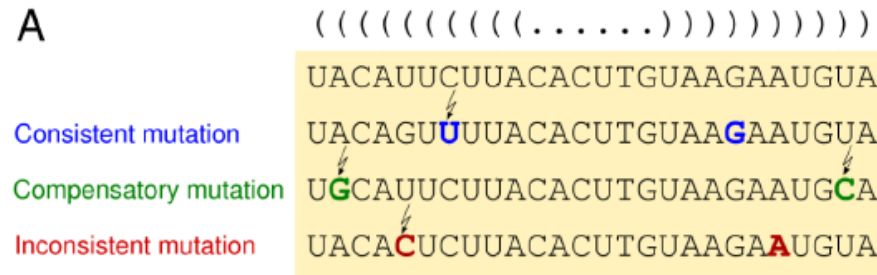
Schipperkee



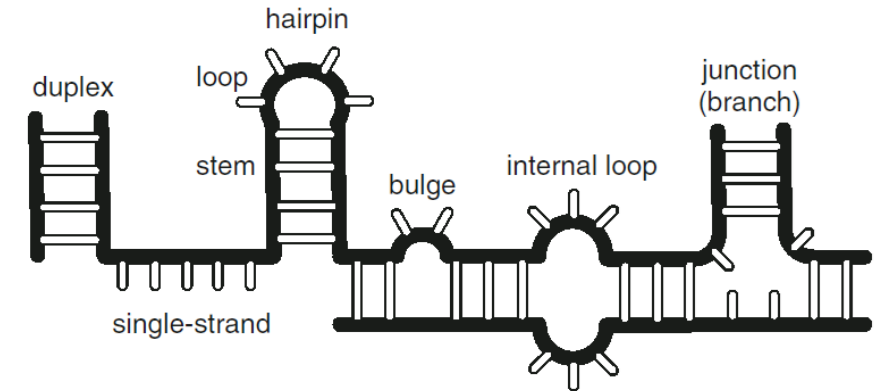
IN CS, IT CAN BE HARD TO EXPLAIN
THE DIFFERENCE BETWEEN THE EASY
AND THE VIRTUALLY IMPOSSIBLE.

Características

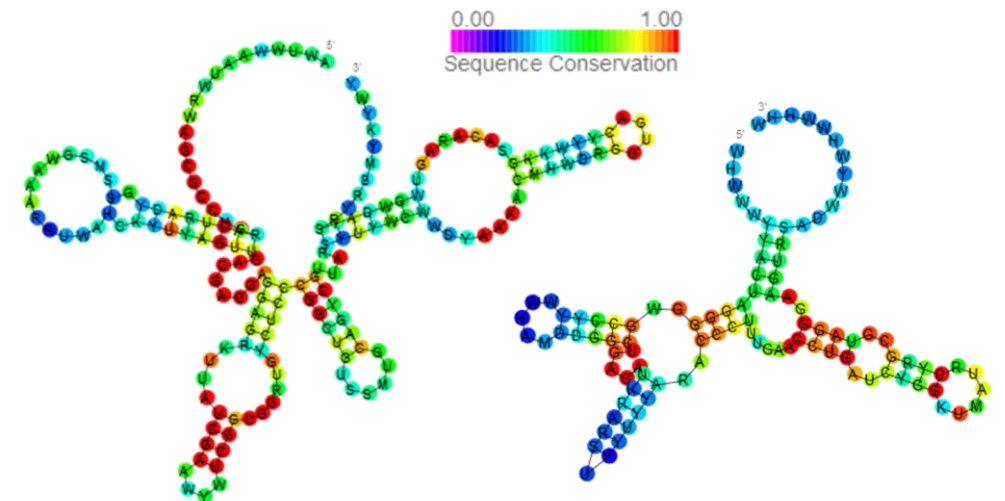
PLEGAMIENTO, ESTRUCTURA SECUNDARIA Y ALINEAMIENTOS



- A) 3 secuencias que se pliegan en un hairpin alineadas.
- B) Alineamientos a nivel de secuencia vs estructural.

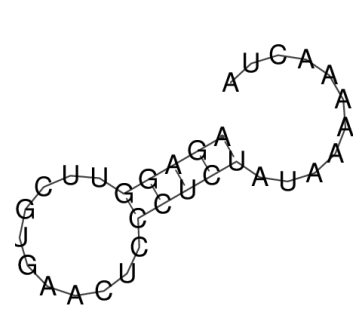


Nomenclatura de subestructuras de ARN

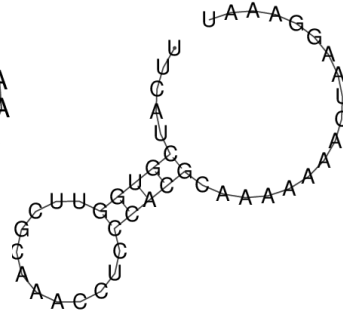


glmS (RF00234) [275 seqs] **TPP (RF00059) [3616 seqs]**

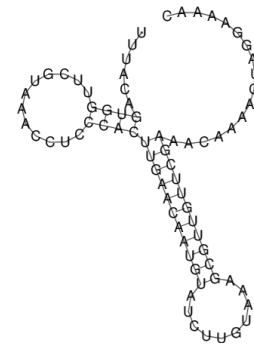
RFAM's consensus secondary structure of glmS and TPP.



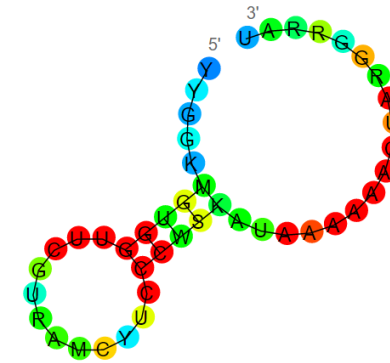
CP001793.1



CP000746.1

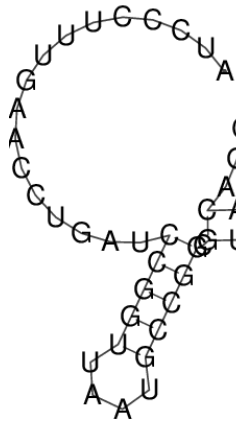


ACZR01000018.1

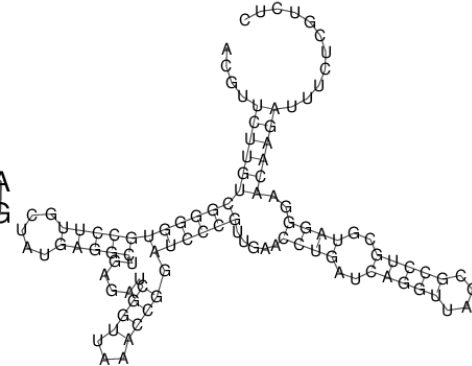


PreQ1

Estructura secundaria para secuencias de
tamaño mínimo, medio, máximo, y estructura
consenso RFAM de PreQ1



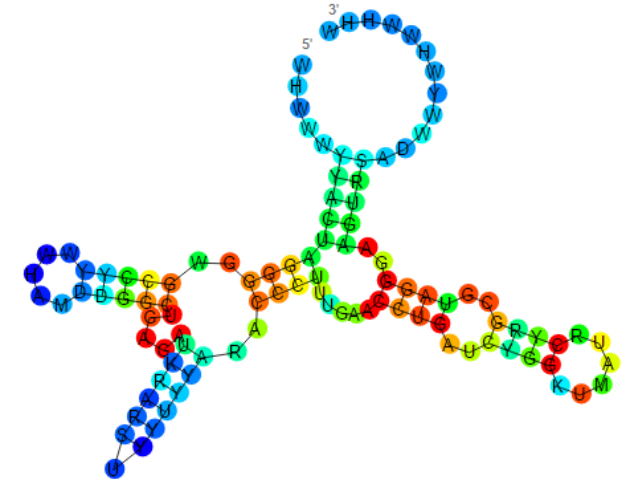
AACY024024797.1



AEAO01000314.1



FN869568.1

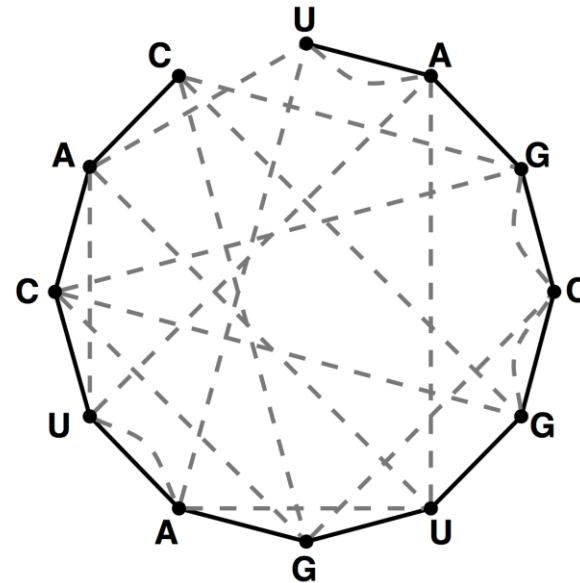
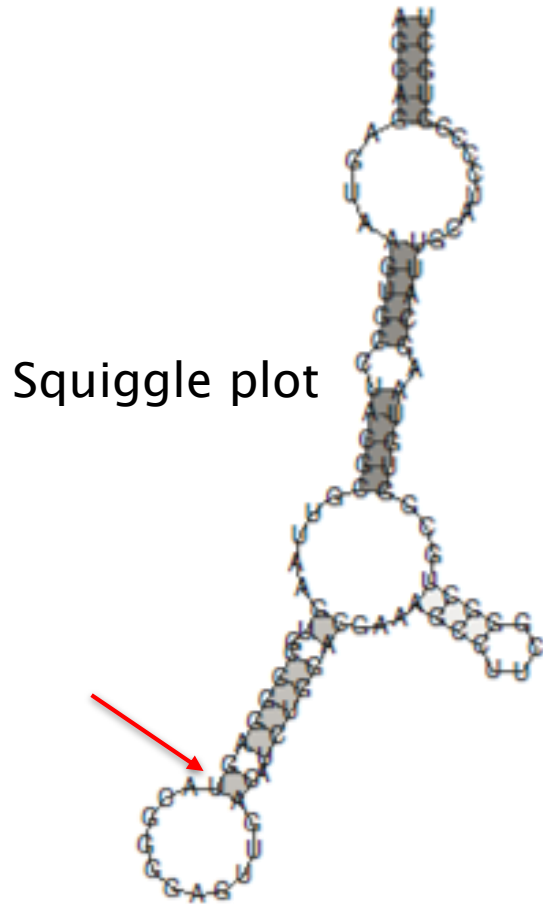


TPP

Estructura secundaria para secuencias de
tamaño mínimo, medio, máximo, y estructura
consenso RFAM de TPP

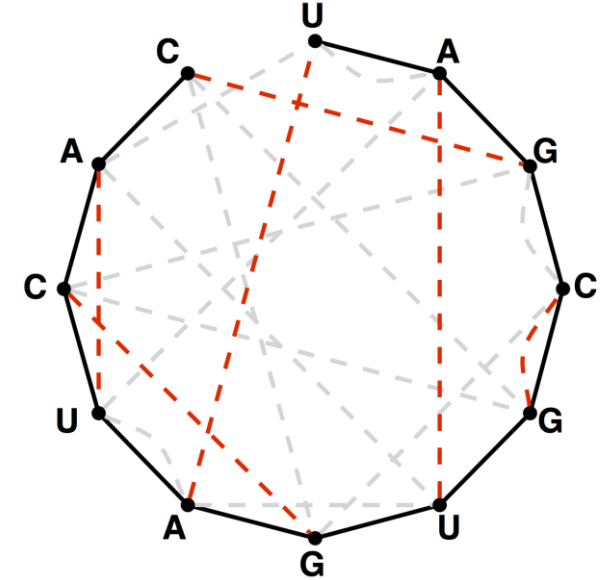
PLEGAMIENTO DE ARN

PLEGAMIENTO



Bonding graph para
UAGCGUGAUCAC.

Matching perfecto que
representa una posible
estructura secundaria.



¿Cuántos matching
perfectos hay para un
bonding graph dado?

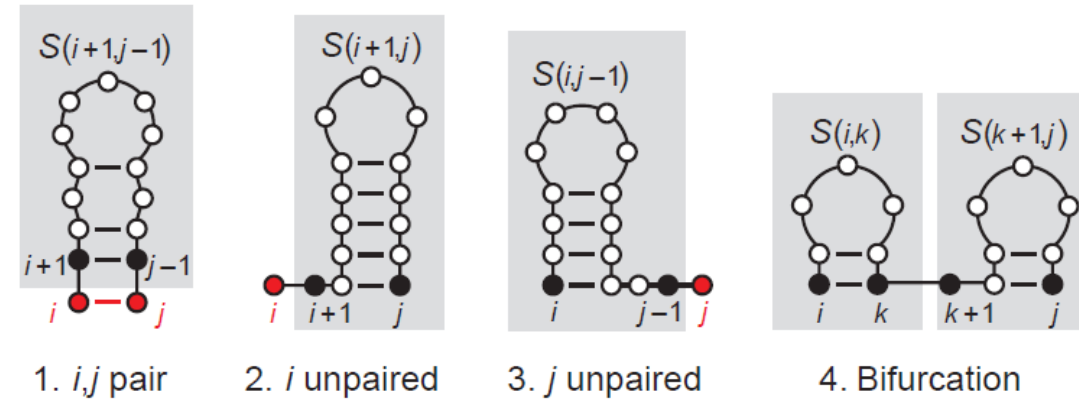
AGCAGAGUAAGUGCCUACGCGUUAAGUGCCGGAGUACGGGGAGUUGACAUCUGGACG;
((((.....((((.....((((.....(((.....)))))))).

Notación dot-
bracket

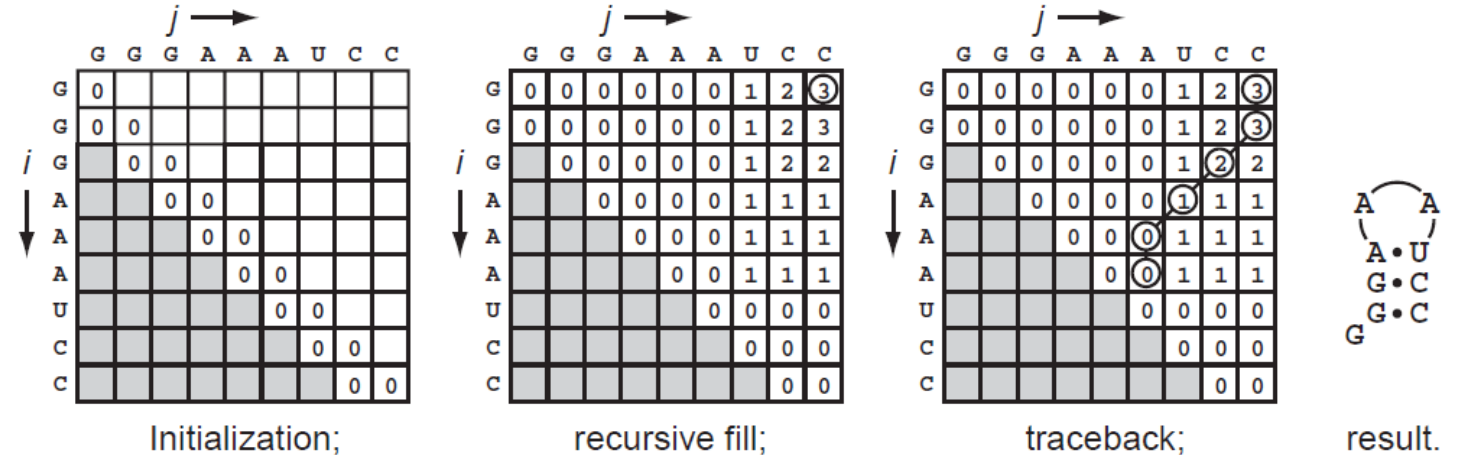
$$S(i,j) = \max \begin{cases} S(i+1,j-1) + 1 & \text{[if } i,j \text{ base pair]} \\ S(i+1,j) \\ S(i,j-1) \\ \max_{i < k < j} S(i,k) + S(k+1,j) \end{cases}$$

Algoritmo de Nussinov y Jakobson (1980)

a Recursive definition of the best score for a sub-sequence i,j looks at four possibilities:



b Dynamic programming algorithm for all sub-sequences i,j , from smallest to largest:

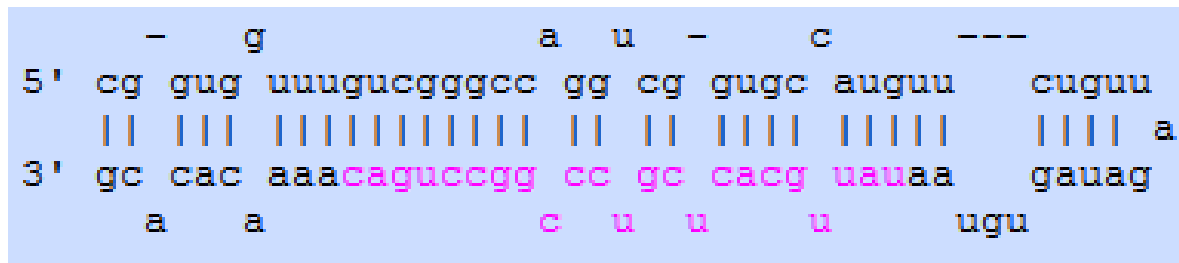


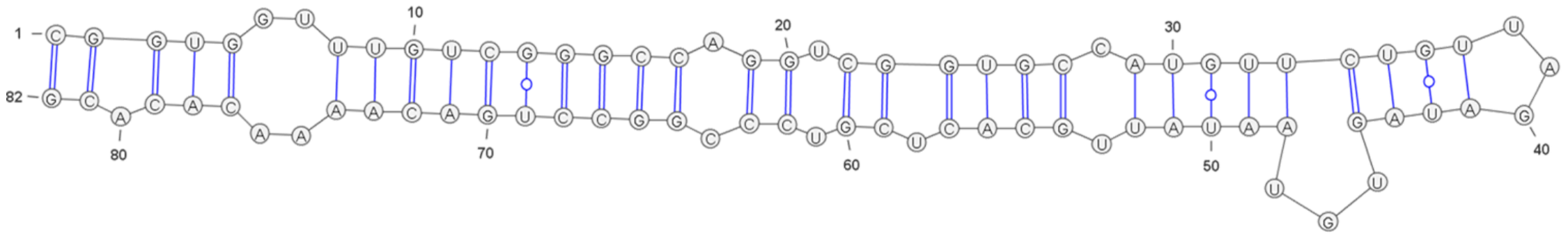
Xenoturbella bocki miR-92a stem-loop

- La búsqueda de herramientas no invasivas de diagnóstico y manejo de cancer es extremadamente importante.
- Los micro ARNs (miRNAs) muestran potencial como biomarcadores y pueden encontrarse circulando en suero. Algunos son específicos en pacientes de tumores.
- miR-92 está presente en individuos sanos en el suero pero en niveles variables y presenta cambios en respuesta a ciertos cánceres.

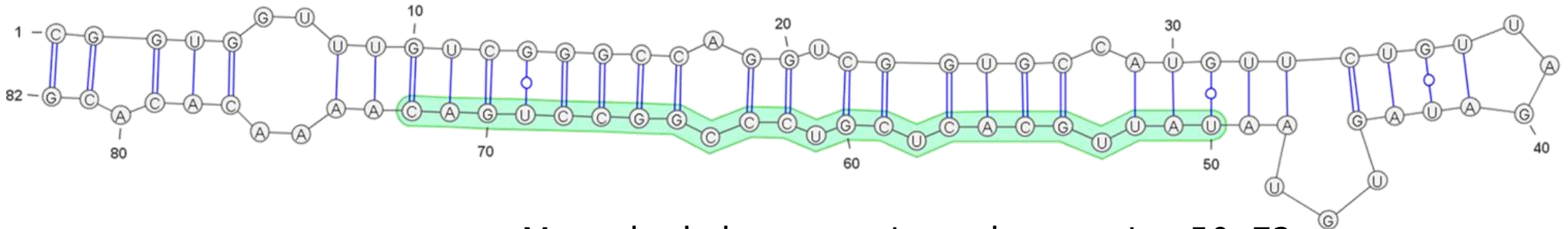
>xbo-mir-92a MI0017684

CGGUGGUUUUGUCGGGCCAGGUCGGUGCCAUGUUCUGUUAGAUAGUGUAAUUAUUGCACUCGUCGCCGGCCUGACAAAACACACG





Visualización de xbo-mir-92a MI0017684 en VARNA usando el algoritmo NAVIEW.



Marcado de la secuencia madura, region 50-72.

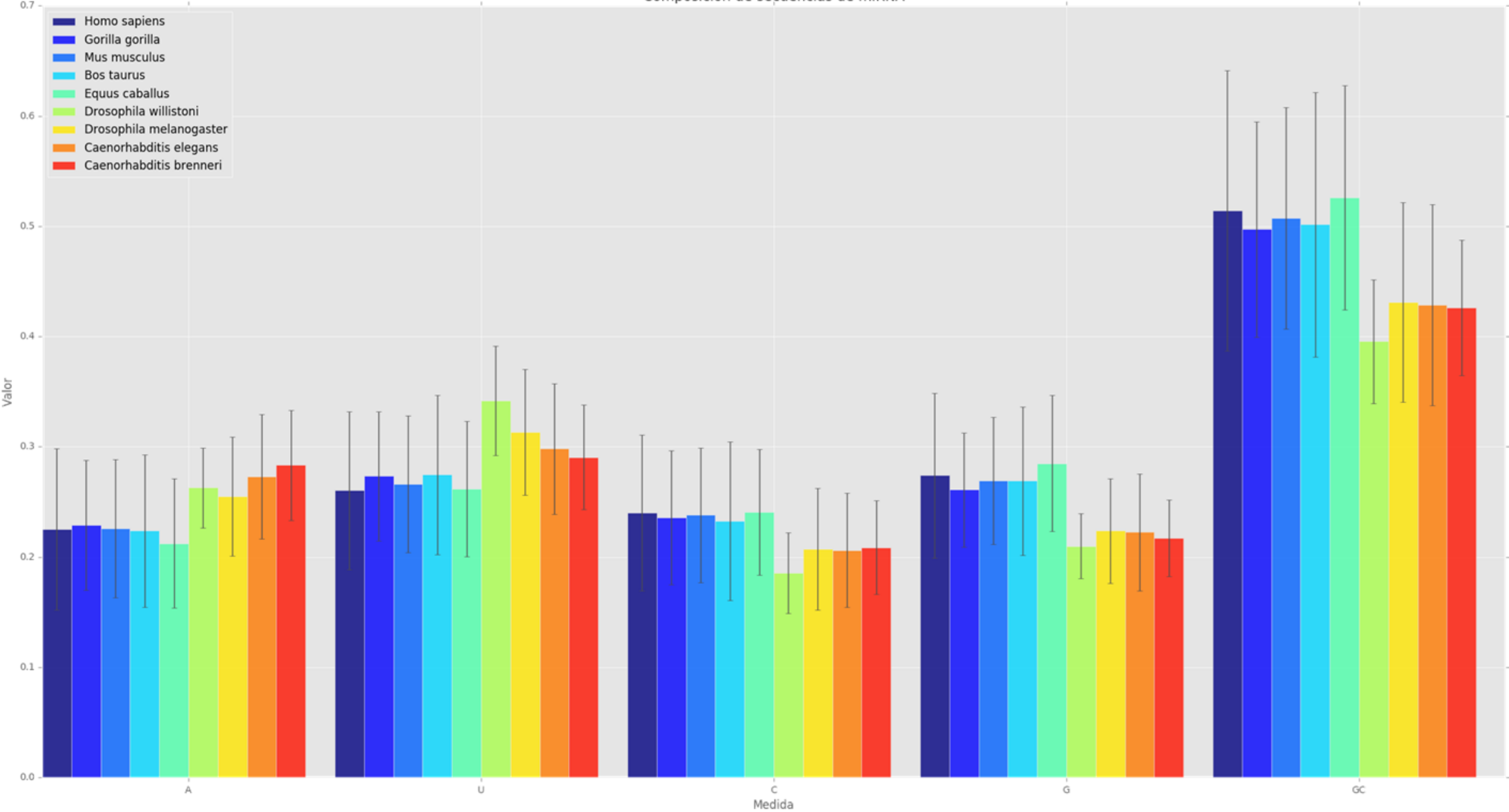
```

      -   g             a   u   -   c       ---
5'  cg gug uuugucggggcc gg cg gugc auguu   cuguu
    || ||| ||||| ||||| || || |||| ||||| ||||| a
3'  gc cac aaacaguccgg cc gc cacg uauaa   gauag
      a   a             c   u   u       u       ugu

```

Representación de la estructura del miRNA en miRBase.

Composición de secuencias de miRNA



¡GRACIAS POR SU
ATENCIÓN!