

Herramientas Computacionales

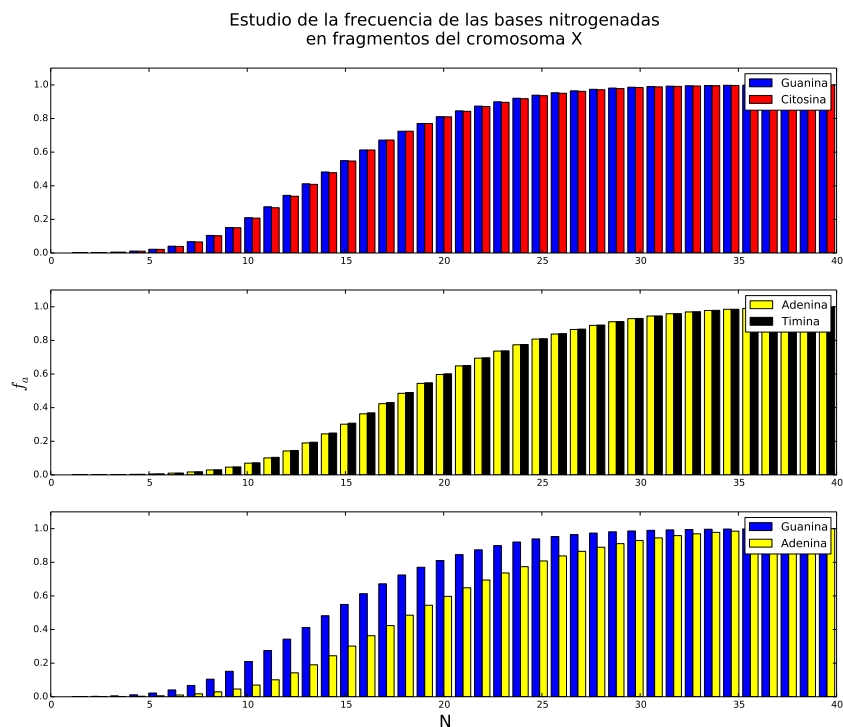
Taller 9 - PYTHON: herramientas estadísticas

Octubre de 2014

1. 100 pt El cromosoma X

El archivo `cromox.dat`¹ es un archivo de texto que contiene una versión simplificada de la secuencia genética del cromosoma X. Cada línea tiene 71 caracteres, 70 letras del genoma y el carácter nueva línea `\n`.

- 40 pt Hacer un ciclo que analice todo el archivo, que en cada iteración lea una línea completa (71 caracteres), que cuente usando la función `count` el número de veces que aparece en esa línea cada base nitrogenada (A, G, T, C), y que añada esas cantidades a los arreglos: `countA`, `countG`, `countT` y `countC`.
- 25 pt Calcular el promedio, la desviación estándar y la mediana de los cuatro arreglos.
- 35 pt Haga tres histogramas cumulativos: uno que incluya en los mismo ejes los histogramas de `countG` y `countC`, los de `countA` y `countT`, los de `countG` y `countA`. Todos los histogramas deben tener leyendas apropiadas, y deben estar normalizados. Los resultados deben ser similares a lo mostrado en la figura.



¹Datos obtenidos de NCBI.