

## EntroPhyl: un programa para la entropía biológica en filogenia

SAID A. RETANA-SALAZAR<sup>1</sup>

<sup>1</sup>*RS consultores (Sección Bioinformática), Colonia Kennedy 53, San Sebastián, San José, Costa Rica, América Central*

---

**RESUMEN:** En este artículo se presenta el primer programa para calculadora disponible para el cálculo del nivel entrópico de los árboles obtenidos a partir de la matriz de datos para reconstrucción filogenética. El programa utiliza el algoritmo desarrollado y publicado por Brooks y Wiley (1986) para la evaluación de árboles filogenéticos.

**PALABRAS CLAVE:** entropía, calculo, entropía biológica, filogenia.

**ABSTRACT:** This paper introduces the first program for the calculation of the entropic level show by a phylogenetic the trees obtained from a given data matrix after a phylogenetic reconstruction. The program performs the algorithm developed and published by Brooks and Wiley for the evaluation of phylogenetic trees.

**KEYWORDS:** entropy, calculation, biological entropy, phylogeny

---

### INTRODUCCIÓN

Desde la década de los años 1920 se cita en la literatura la convergencia entre los sistemas entrópicos y los en evolución. Ambos mecanismos siguen patrones similares y en los últimos años se ha demostrado que ambos procesos se explican a través de estructuras matemáticas que involucran un principio descrito a través de un incremento unidireccional de una propiedad cuantificable y operacional (Demetrius 2000).

Todo esto ha sido teoría hasta que Brooks y Wiley (1986) sintetizan la forma en que las ecuaciones y medidas estadísticas de la entropía pueden ser aplicadas a sistemas en evolución. Desde el punto de vista práctico nada más importante que la aplicación de estos conceptos a la reconstrucción filogenética, con

lo que es posible evaluar entre muchos posibles arreglos filogenéticos obtenidos de una matriz de datos, todos con idénticos estadísticos y en consecuencia igualmente probables para la reconstrucción filogenética por parsimonia.

Las matrices de datos de grupos altamente homoplásicos suelen presentar desde decenas hasta miles de soluciones todas igualmente probables, por lo que es imposible saber si alguna corresponde al arreglo filogenético más probable.

Cuando una matriz es altamente homoplásica puede dar resultados equivocados al interpretar algunos caracteres como homólogos sin serlo y viceversa produciendo múltiples arreglos filogenéticos igualmente probables. Una matriz de datos en la cual la varianza de las homoplasias iguala o supera la de las

homologías va a dar resultados falsos sobre la matriz de datos utilizada. Por otra parte si se elige como mejor arreglo filogenético al más estable de los árboles, es decir al de menor entropía entonces tendremos un único árbol más probable (Brooks y Wiley 1986).

### EJEMPLO HIPOTÉTICO

Utilizando una matriz de datos hipotética con dos posibles resoluciones (cuadro 1). De esta matriz se obtienen dos árboles igualmente probables (Fig. 1)

Desde este punto de vista es imposible a través de la reconstrucción parsimoniosa determinar cual árbol es más probable y en consecuencia un árbol consenso estricto no resolverá los parentescos entre A,B,C sino que politomizará

estos tres grupos, siendo esta una resolución poco parsimoniosa.

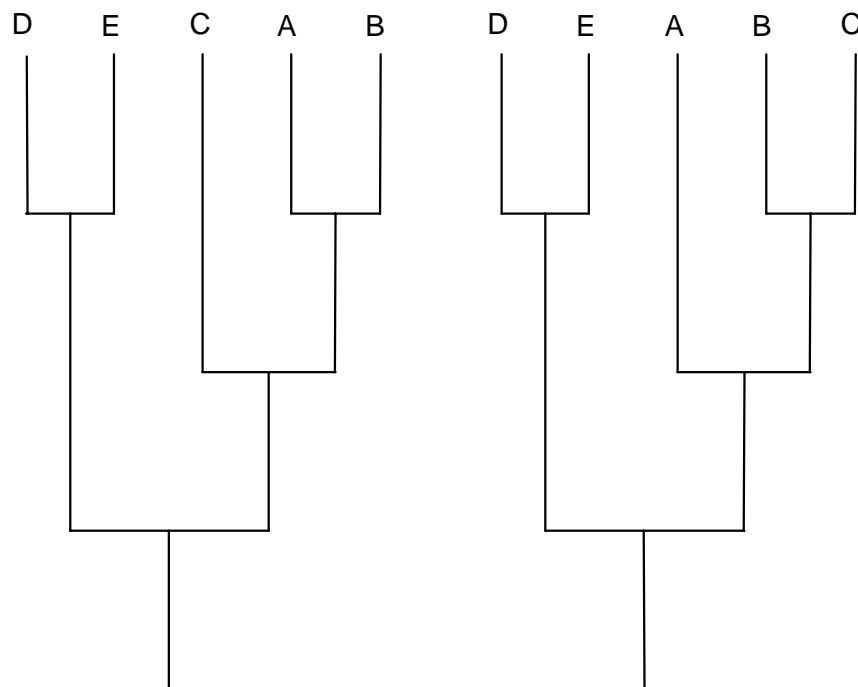
Al aplicar el concepto de entropía filogenética podemos calcular la entropía propia de cada árbol y el más probable será el menos entrópico.

### CÁLCULO DE LA ENTROPÍA

Para esto debemos conocer el número de caracteres en cada rama del árbol y crear un cálculo de probabilidades entre los caracteres y la topología según lo establecen Brooks y Wiley (1986).

**Cuadro 1:** Matriz de datos hipotética con dos posibles resoluciones igualmente parsimoniosas.

A	B	C	D	E
111100100000	111100111000	111000110100	100011000010	100011000001



**Figura 1:** Reconstrucciones posibles de la matriz del cuadro 1 ambas con L=13, CI=0,92, RI=0,90.

CÁLCULO DE ENTROPÍA  
UTILIZANDO EL ALGORITMO DE  
BROOKS Y WILEY

Entropía de la topología 1 (((AB)C))(DE))

Entropía de la topología 2 ((BC)A))(DE))

Aplicando la fórmula establecida por Brooks y Wiley (1986)  $H = -\sum p_i \ln p_i$ , se obtiene el valor de H o la entropía asociada a la topología, como es evidente  $H_1 < H_2$ , por lo que la topología 1 es más probable que la 2.

igual forma sirve para calcular la entropía por cada grupo independiente de la matriz, es decir las submatrices de cada clado.

PROGRAMA EntroPhyl EN SQL BASIC

*List1List2 → List3 ← (sum List3)*

*1List3) → List3 ←*

*(Log List3/Log 2)List3 → List3 ← sum List3*

**Cuadro 2:** Cálculo de entropía para una topología determinada

Grupos	Caracteres	P carac (Pc)	P topo (Pt)*	(Pc)(Pt)	P estándar
A	0	-	-	-	-
B	2	2/13	1/13	2/169xK	2/20
Ancestro 1	1	1/13	2/13	2/169xK	2/20
C	2	2/13	1/13	2/169xK	2/20
Ancestro 2	3	3/13	2/13	6/169xK	6/20
D	1	1/13	1/13	1/169xK	1/20
E	1	1/13	1/13	1/169xK	1/20
Ancestro 3	2	2/13	2/13	4/169xK	4/20
Ancestro Inic.	1	1/13	2/13	2/169xK	1/20
TOTAL	13			20/169	20/20=1
				K=169/20	H <sub>1</sub> =2,75

**Cuadro 3:** Cálculo de entropía para una topología determinada

Gr	Car	Pc	Pt*	Pc Pt	Pest.
B	2	2/13	1/13	2/169xK	2/20
C	1	1/13	1/13	1/169xK	1/20
Anc1	1	1/13	2/13	2/169xK	2/20
A	1	1/13	1/13	1/169xK	1/20
Anc2	3	3/13	2/13	6/169xK	6/20
D	1	1/13	1/13	1/169xK	1/20
E	1	1/13	1/13	1/169xK	1/20
Anc3	2	2/13	2/13	4/169xK	4/20
AncI	1	1/13	2/13	2/169xK	2/20
Tot	13			20/169	20/20=1
				K=169/20	H <sub>2</sub> =2,85

\* El cálculo de Pt se facilita al calcular el denominador con la fórmula  $3n-2$ , donde n=número de grupos.

Sin embargo, este cálculo es engorroso cuando la matriz de datos es muy extensa y aún más si las posibles soluciones son muchas. Con el programa EntroPhyl basta con alimentarlo con las columnas (Pc) y (Pt) para obtener el resultado, de ahí el valor de este programa, de

Este programa ha sido ampliamente probado durante 3 años en cálculos teóricos y reales con excelente resultado. Tiene la facilidad de correr rápidamente en cualquier calculadora de bolsillo programable que pueda incorporar este lenguaje.

## REFERENCIAS

**Brooks DR, Wiley EO. 1986.** Evolution as Entropy: Toward a Unified Theory of Biology, Chicago Univ. Press.

**Demetrius L. 2000.** Thermodynamics and Evolution (review). J. theor. Biol. 206: 1-16.

Recibido: 6 Diciembre 2005

Aceptado: 20 Enero 2006