



UNIVERSIDAD NACIONAL
AUTONOMA DE MEXICO

FACULTAD DE ESTUDIOS
SUPERIORES IZTACALA

T E S I S

**“Modelo computacional de procesos
evolutivos y posibles consecuencias sobre
el gradualismo filogenético”**

Que para obtener el título de

Biólogo

Presenta

José Arturo Alcántara Rodríguez

Director de tesis:
Fís. Fernando Ángeles Uribe



2005



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos

A la Universidad Nacional Autónoma de México, por todo lo que me ha dado y porque estoy orgulloso de ser parte de ella.

A Fernando Angeles Uribe, porque este trabajo también es tuyo, por tu apoyo y por los conocimientos que has compartido conmigo.

A mis profesores, quienes han contribuido grandemente a la culminación de esta meta.

En memoria de Roberto Rico Montiel, por tu agradable manera de enseñar, tus clases siempre fueron muy interesantes y divertidas, por haber creído en este proyecto que me hubiera gustado vieras concluido. Te recuerdo con cariño.

A todos los que han contribuido en mi formación profesional y personal.

Dedicatorias

...A mi madre, por enseñarme a vivir en libertad, por los valores que me has inculcado desde niño, porque sin tu amor y apoyo incondicional no hubiera alcanzado esta meta. Por el inmenso amor que te tengo.

...A Brenda, por su amor y apoyo, por toda la felicidad que hemos aprendido a compartir juntos y porque esta es la culminación de un paso mas en nuestro proyecto de vida y el inicio de otro. Te amo.

...A mis abuelos, por su enseñanza de vida y fortaleza, por su amor y apoyo de todos los días. Son muy importantes en mi vida.

...A Yola, por ese cariño, por creer en mí y alentarme siempre a conseguir mis metas. Te quiero mucho.

...A Micky, por ser alguien tan importante en la formación de mi personalidad, por subirme a la azotea cuando tenía cinco años a ver el cometa Halley y despertar en mí el deseo de conocer más. Sabes que eres una persona muy especial para mí.

...A Mauri, por los buenos momentos, por aquellas veces que me cuidabas cuando era niño y nos divertíamos juntos.

...A Mari, por siempre preocuparte por mi y apoyarme, porque has sido a lo largo de mi vida una persona muy importante.

...A Nico, por involucrarme en el trabajo familiar, enseñándome a usar las herramientas y por la confianza que depositaste en mi para hacerlo bien.

...A mis amigos, porque sin el valor de su amistad mi vida sería muy distinta.

Agradecimientos	4
Dedicatorias	5
Introducción	6
Antecedentes	13
Justificación	16
Objetivos	16
Materiales y Métodos	17
El Universo	19
La Comida.....	19
Los Organismos.....	20
Movimiento.....	28
Reproducción.....	30
Bicho Tipo	30
Gráficas de Disimilitud y Disimilitud Acumulada.....	32
Gradualismo y discontinuidad	34
Gráficas de Promedio de Células.....	36
Resultados	38
Modelo 1.....	39
Modelo 2.....	45
Modelo 3.....	50
Modelo 4.....	56
Modelo 5.....	62
Modelo 6.....	68
Comparativo entre los diferentes modelos.....	73
Análisis de resultados	75
Conclusiones	79
Anexo 1	83
Apéndice A	85
Referencias	97

Introducción

La uniformidad que presenta el código genético en todos los organismos sugiere que la vida tuvo un origen único y que todas las formas actuales se derivan de un arreglo primitivo de compuestos orgánicos capaz de autoreplicarse. La existencia de la gran diversidad biológica así como la relación presente entre la morfofisiología de los organismos y el ambiente en el que habitan han llevado a la biología a postular teorías evolutivas que tratan de explicar este fenómeno, entre las que se destacan: La Teoría Sintética de la Evolución y la Teoría de los Equilibrios Puntuados. [3]

En las décadas de los años 1930 y 1940 los trabajos de Dobzhansky, Mayr, Simpson y Huxley sentaron las bases para la formulación de la Teoría Sintética de la Evolución, tomando como punto de partida el principio de selección natural propuesto por Darwin y Wallace en la segunda mitad del siglo XIX, la TSE desechó por completo la idea de la herencia de caracteres adquiridos e incorporó los principios de la genética Mendeliana. En lo que se refiere al proceso de especiación, la TSE propone que el surgimiento de nuevas especies es un proceso gradual y continuo, resultado de la acumulación de cambios en las frecuencias alélicas de una población y que dichos cambios están dirigidos por la selección natural de los individuos en una población, la aparición de variaciones entre los individuos es explicada mediante las mutaciones en el material genético. [19], [3]

En 1972 Gould y Eldredge postularon la Teoría de los Equilibrios Puntuados con la finalidad de explicar las discontinuidades morfológicas encontradas por los paleontólogos en el registro fósil. La TEP sugiere que el cambio evolutivo se lleva a cabo por medio de la selección natural actuando sobre las especies (no sobre los individuos como propone la TSE) y tiene una naturaleza episódica, se consideró que la ausencia de formas intermedias entre una y otra especie fósil se debe a que la evolución no es un proceso gradual sino que ocurre en un corto periodo de tiempo que consideran, debe ser de alrededor de cien mil años, después de esto las especies permanecen sin cambios durante millones de años en periodos de estasis. [6], [11]

El gradualismo filogenético es un tema de fuerte debate entre las diferentes corrientes evolutivas. En la actualidad, no existen argumentos contundentes que

confirmen alguna de ellas y esto puede deberse a que generalmente se excluyen de las investigaciones los mecanismos bioquímicos de expresión y regulación de los genes, así como los efectos sobre el tipo y número de fenotipos posibles a partir de un estado evolutivo. Una visión holística debería incluir no solo el efecto de los mecanismos de selección sobre el fenotipo, además debe contemplar la manera en que los mecanismos moleculares intervienen en la definición de las rutas evolutivas. Aunque esta puede parecer una hipótesis demasiado fuerte debido a la naturaleza abierta del problema, no está de más reconsiderar los modelos actuales para incluir estos efectos y estudiar las posibles consecuencias sobre las poblaciones, tema central de este trabajo.

Es un hecho que el proceso evolutivo solo es posible gracias a la transmisión de información genética de una generación a otra de seres vivos, para Maynard Smith la vida debe ser definida por la posesión de las propiedades necesarias para asegurar la evolución por selección natural, esto es multiplicación, variación y herencia. [29]

Ya que los seres vivos utilizan la información para reproducirse, el concepto de información es importante para caracterizarlos, la palabra información proviene del Latín *informatio* que quiere decir iluminación, exposición o “cuerpo de conocimientos”. Shannon y Weaver desarrollaron a finales de los años cuarenta una teoría conocida como “Teoría Matemática de la Información”, ellos definen la información transmitida a través de un canal como “La selección de una instancia de un conjunto de posibles mensajes x ” esto es, el emisor comunica por medio de un canal cuál de todos los mensajes x posibles debe ser escogido por el receptor, de tal manera que la cantidad de información en el mensaje está dada por el número de posibles elementos a seleccionar. Por conveniencia la decisión entre dos elementos en un conjunto ha sido tomada como una unidad de información, el *bit*. [28]

La cantidad de información en *bits* puede ser calculada de la siguiente forma:

$$H = \log_2 N$$

Donde N es el número de estados posibles que puede adoptar un símbolo.

La cantidad de información contenida en el genotipo de un organismo puede ser cuantificada a partir del número de nucleótidos que contiene su genoma, para cada nucleótido en la secuencia de su ADN son necesarios dos *bits* de

información debido a que dicho nucleótido puede tener uno de cuatro valores posibles (Adenina, Timina, Guanina o Citosina). De tal manera que si:

$$H = \log_2(4)$$

$$H = 2 \text{ bits por base}$$

De este modo se puede calcular la cantidad de información contenida en el genotipo de cualquier organismo; si cada nucleótido necesita dos *bits* para ser representado, entonces el genoma entero contiene:

$$H = 2G$$

Donde *G* es el número de nucleótidos del genoma.

Por lo tanto, si el genoma de *E. Coli* consta aproximadamente de 4.2×10^6 pares de bases.[18]

La cantidad de información es:

$$H = 2 (4.2 \times 10^6)$$

$$H = 8.4 \times 10^6 \text{ bits/célula}$$

Quastler utilizó *B. Pycnoticus* como modelo para calcular la cantidad de información contenida en el fenotipo de una célula bacteriana, utilizó el siguiente método basado en la entropía termodinámica para calcular dicho valor: (Ver Apéndice A)

$$H = - \frac{\Delta s / \text{cell}}{k \ln_e 2}$$

El valor que obtuvo fue de:

$$H = 10^{13} \text{ bits/célula.}$$

La diferencia en la cantidad de información contenida en el fenotipo de las bacterias con respecto a su genotipo implica una aparente ganancia de

información. Es un hecho que a partir del material genético son construidos los fenotipos de los organismos; sin embargo, tomando en consideración las estimaciones del contenido informático en bacterias se puede afirmar que la cantidad de información en el genotipo de los organismos es menor a la necesaria para construir sus fenotipos, esto sugiere que durante la ontogenia de los seres vivos existe un comportamiento emergente¹ que genera las estructuras no descritas en el material genético [27].

Debido a la diferencia en la cantidad de información contenida en el genotipo con respecto al fenotipo se puede esperar que un gen determine más de una característica y que por lo tanto, algunas características estén determinadas por conjuntos de genes, de no ser así, no bastarían los genes para producir todas las características presentes en los fenotipos, esto no quiere decir que una secuencia de nucleótidos pueda generar proteínas con diferentes secuencias, sino que la interacción de los genes con el ambiente o entre ellos mismos durante el desarrollo o morfogénesis produce un fenómeno de emergencia, es decir, el sistema resultante (fenotipo) no puede ser explicado a partir de la suma de sus partes (genotipo).

El número de genes requerido para producir un determinado fenotipo suele ser mucho menor que el número de partes que lo componen, por ejemplo, en el cerebro de los humanos hay alrededor de un billón de conexiones neuronales pero sólo existen unos treinta mil genes activos en el genoma humano. Esto implica que el fenotipo está de alguna manera comprimido en el genotipo y que durante el desarrollo se lleva a cabo la descompresión de la información mediante el uso de una serie de algoritmos determinados por las propiedades fisicoquímicas de la materia que interviene en el proceso.[30]

Un ejemplo de esto es el plegamiento de las proteínas, una secuencia de nucleótidos determina la secuencia en la que serán arreglados los aminoácidos, pero no establece los ángulos que guardarán, entre sí ni la funcionalidad de la proteína; son las propiedades fisicoquímicas de las moléculas las que determinan las estructuras secundarias y terciarias de las proteínas, de una secuencia lineal de símbolos se pasa a una estructura tridimensional de átomos con una aparente ganancia de información debido a que en dicha secuencia

¹ La emergencia de un sistema indica la posesión de cualidades y atributos que no se sustentan en las partes aisladas y que, por otro lado, los elementos o partes de un sistema actualizan propiedades y cualidades que sólo son posibles en el contexto de un sistema dado.

precursora no está especificada la geometría de la molécula así como la distribución de las cargas eléctricas que son fundamentales para el funcionamiento de las proteínas [26].

Este fenómeno de emergencia se puede observar desde la formación de micelas a partir de lípidos, la formación de membranas celulares, la síntesis de proteínas con estructuras altamente complejas, el arreglo de las proteínas con otras moléculas para integrar organelos celulares, etc. ninguna de estas tareas está determinada de forma explícita en el genoma de los seres vivos, no es posible encontrar una secuencia que determine el arreglo de una red neuronal o cualquier otro tipo de estructura más allá de la secuencia de aminoácidos de una proteína, ni siquiera las estructuras secundarias de estas, las instrucciones para sintetizar las diferentes estructuras que conforman a los organismos se determinan durante el desarrollo.

La sugerencia de que existe una mayor cantidad de información en el fenotipo que en el genotipo implica que se produzcan alteraciones importantes en el fenotipo cuando el genotipo sufre mutaciones, debido a que la morfogénesis puede ser considerada como un proceso caótico, un cambio en un símbolo del genotipo puede alterar toda una secuencia de acontecimientos, por ejemplo, al modificar la secuencia de una proteína se pueden alterar sus estructuras secundaria y terciaria teniendo como consecuencia un cambio en el funcionamiento de una célula, y a su vez esto modificar la estructura de un tejido, un órgano y así sucesivamente afectar a niveles más altos.

Como se mencionó anteriormente la alteración de un símbolo en el genotipo de un organismo podría resultar en una desviación significativa del fenotipo producido, esto podría tener consecuencias sobre el modelo de gradualismo filogenético debido a la posibilidad de obtener mutantes con variaciones fenotípicas importantes a partir de mutaciones puntuales, teniendo como posible efecto la aparición repentina de especies.[11]

En un intento interdisciplinario por comprender los procesos inherentes de la vida, la disciplina de la Vida Artificial se encarga de estudiar los sistemas artificiales que exhiben comportamientos característicos de los sistemas vivos

naturales, trata de explicar con modelos² los fenómenos propios de la vida, sin restringirse a los ejemplos basados en la química del carbono encontrados en la tierra, con este fin se realizan experimentos en biología y química así como simulaciones por computadora y razonamientos teóricos de los procesos que ocurren a nivel molecular, social y evolutivo. Actualmente existen dos corrientes en materia de Vida Artificial, la postura “suave” pretende extraer la lógica elemental de los seres vivos mediante el uso de herramientas como las matemáticas, la física informática, la genética entre otras, se implementan programas de computadora que simulan el comportamiento de sistemas vivos con la finalidad de entender procesos como el origen de la vida, la inteligencia, la evolución, dinámicas ecológicas y sociales entre otras. Por otra parte, los seguidores de la postura “dura” afirman que los seres vivos pueden ser reproducidos artificialmente en un laboratorio o una computadora a partir de elementos inertes, debido a que aseguran que sólo se trata de sistemas complejos³ que pueden ser implementados en un medio artificial, llevan a cabo experimentos para “crear vida”[17]

La mayoría de las investigaciones en el campo de la Vida Artificial se centran en la síntesis y simulación de sistemas vivos, la Evolución Digital es una rama de estudio de la Vida Artificial que ocupa formas de vida artificiales representadas en forma de códigos autoreplicantes de computadora que habitan un área especialmente preparada en un sistema computacional, todo esto con la finalidad de aprender acerca de los aspectos fundamentales del proceso evolutivo. [1]

El presente trabajo aborda el problema de la evolución bajo la perspectiva de la Evolución Digital, mediante el uso de una computadora se realizó la simulación de una población de organismos autoreplicantes con un mecanismo de expresión genética basada en un algoritmo modificado de *IFS* (Anexo1), esto con el fin de emular la diferencia en la cantidad de información entre el genotipo y el fenotipo observada en los organismos reales. Las poblaciones artificiales están sujetas a la presión selectiva ejercida por la cantidad y

² Modelo. Sustitución de un objeto original con uno más simple, ej. ecuaciones matemáticas aplicadas a la realidad, descripciones de eventos en un lenguaje ordinario, simulaciones computacionales de fenómenos reales.

³ Sistemas complejos. Aquellos sistemas formados por una gran cantidad de elementos que interaccionan localmente y dan lugar a un comportamiento global que no puede ser explicado a partir de las propiedades de un solo elemento sino como el resultado de las interacciones de un colectivo, ej. cerebro, seres vivos, sistemas sociales o económicos. [7]

contenido energético del alimento presente en el ambiente, de esta manera, existen individuos con determinadas características morfológicas que tienen mayor probabilidad de reproducirse, dando como resultado un proceso de selección natural darwiniana de “fin abierto”, es decir, la probabilidad diferencial de reproducción entre los individuos que conforman la población está en función de las condiciones ambientales y no fue implementado un programa que asigne dicha probabilidad basado en ningún atributo morfológico o genético de los individuos.

Antecedentes

Maynard Smith [29] sugiere que la vida debe ser definida por la posesión de las propiedades necesarias para asegurar la evolución por selección natural, esto es, “entidades con las propiedades de multiplicación, variación y herencia están vivas, mientras que aquellas que carecen de una o más de estas propiedades no lo están.”

Ernest Mayr [21] propone que la posesión de un programa genético es la diferencia absoluta entre los organismos y la materia inanimada, asegura que los programas de computadora hechos por el hombre son la única cosa en el mundo inanimado que se puede comparar con la vida.

Para Christopher Langton [17] la vida debe ser definida como una propiedad de organización de la materia más que como una propiedad de la materia así organizada, afirma que la transferencia y retención de información es una característica esencial de la vida “Una buena parte de lo que hacen los seres vivos está basado en el procesamiento de información, no sólo materia ni energía, también información. Los organismos vivos usan información para reconstruirse, para localizar comida, para mantener su estructura interna... la estructura misma es información. Se puede concluir que en los seres vivos la manipulación de información ha tomado el control, dominando la manipulación de energía”.

Von Neumann[23] se refiere a las máquinas capaces de operar autónomamente con el término “autómata”, un autómata es una maquina con un comportamiento definido matemáticamente que procesa información y procede lógicamente. Sugiere además, que los seres vivos pueden ser vistos como autómatas naturales capaces de autoreplicarse y plantea la posibilidad de que un autómata artificial pueda hacer copias de sí mismo, sentando así las bases de la Vida Artificial.

John Conway [17] desarrolla *Life*, un autómata celular bidimensional en el que el estado de cada células depende del estado de sus ocho vecinas inmediatas, dando como resultado un comportamiento global emergente.

John Holland [17] utiliza la selección natural darwiniana para desarrollar un algoritmo de computadora capaz de resolver problemas sin indicarle cómo hacerlo, este método de optimización conocido como *Algoritmo Genético* genera una población de soluciones aleatoriamente y le asigna una probabilidad de reproducirse a cada individuo en función de su aptitud para resolver el problema en cuestión, además se establece una tasa de mutación que genera variaciones de las soluciones, optimizando así el proceso por medio de la selección natural.

Richard Dawkins [4] desarrolla un programa de computadora usando un algoritmo recursivo que genera poblaciones de *bioformas* (dibujos a base de píxeles), cada *bioforma* es el resultado de la expresión de genes que son mutados aleatoriamente en cada generación. El usuario selecciona una *bioforma* y se produce una nueva generación con variaciones aleatorias de ésta, dando como resultado una evolución dirigida por la presión selectiva ejercida por el usuario.

Thomas Ray [25] Desarrolla el *Software* “Tierra”, se trata de un ambiente artificial dentro de la memoria de una computadora en el que existen programas capaces de reproducirse y mutar su genoma aleatoriamente, además establece una serie de reglas para que las mutaciones generen códigos entendibles para la computadora. En este universo virtual la memoria de la computadora juega el papel del espacio, debido a que alberga a los programas, por otra parte, el procesador provee a los individuos de la energía necesaria para reproducirse, ya que es el que realiza las operaciones lógicas que permiten la copia y mutación de los códigos. La simulación comienza con un individuo cuyo genoma le permite reproducirse, cuando aparece un código capaz de autoreplicarse con una menor cantidad de instrucciones éste se ve favorecido por la selección natural ya que ocupa menos tiempo de procesador (energía) y menos memoria (espacio), teniendo así, mayores probabilidades de reproducirse.

Setlow y Pollard [27] calculan la cantidad de información contenida en el fenotipo de una bacteria por diversos medios y obtienen un valor de entre 10^{12} y 10^{13} bits, y en apariencia es muy superior a la contenida en el ADN, esto ha dado lugar a diversas teorías que disminuyen la importancia de los genes durante el desarrollo y achacan la emergencia de estructura a otros factores.

Goodwin [10] Propone que la morfogénesis de los seres vivos se lleva a cabo en un sistema complejo a través del cual los genes contribuyen al desarrollo del

fenotipo, sin embargo las propiedades macroscópicas son producto de fenómenos emergentes que determinan el tipo de variabilidad fenotípica puede existir.

Stanley y Miikkulainen [30] Definen la embriología artificial como una subdisciplina de la computación evolutiva en la cual los fenotipos están codificados en los genotipos y son decodificados mediante un proceso de desarrollo inspirado en los sistemas biológicos, esto implica un aumento en la complejidad del fenotipo con respecto al genotipo producido mediante la interacción de la información genética con el ambiente.

Hogeweg [13] Desarrolla un autómata celular que simula un proceso de morfogénesis, la simulación comienza con siete tipos de células diferentes e integra un genoma con las reglas a seguir para llevar a cabo la división y diferenciación celular. Establece poblaciones que constan de veinte agentes capaces de autoreplicarse (el genoma puede sufrir mutaciones) y asigna a cada individuo la probabilidad de reproducción en función su grado de diferenciación celular. El modelo describe un comportamiento gradual en lo que respecta al genotipo (a manera de reloj molecular), en cuanto al fenotipo se observan discontinuidades en las morfologías a través de la historia, lo cual explica como un fenómeno de sensibilidad a las mutaciones.

Fontana y Schuster [8] Encontraron discontinuidades en la evolución del fenotipo con respecto al genotipo, usaron un modelo basado en una población replicante de ARN con la finalidad de generar historias evolutivas para entender mejor la noción de discontinuidad fenotípica y como influye esta en el proceso de adaptación. Consideraron como genotipo a la secuencia de las unidades replicantes y como fenotipo a la estructura secundaria producto de sus secuencias.

Justificación

Debido a las dificultades técnicas que existen para realizar estudios evolutivos en campo o laboratorio y a que aún no existe una postura en materia de evolución que trate de incorporar el mecanismo de la expresión genética, se utilizará un modelo computacional que simule el comportamiento de una población sometida a presión selectiva, además de un algoritmo de expresión de la información basado en un autómata celular, como una alternativa para dar un enfoque diferente a algunos aspectos relacionados con la teoría evolutiva. Para dicho fin, se tomarán en cuenta las siguientes consideraciones:

El código genético es evidentemente digital.

El Dogma Central de la Biología Molecular tiene un carácter fuertemente computacional y determinístico.

Los seres vivos utilizan la información para reproducirse.

La existencia de fenómenos de emergencia en sistemas caóticos tales como los seres vivos.

Objetivos

Desarrollar un modelo computacional que simule el comportamiento de una población de agentes capaces de reproducirse y mutar, en un ambiente que ejerce presión selectiva.

Simular la expresión de los genotipos a través de un mecanismo de expansión de información inspirado en una abstracción de la manera en que ocurre en la naturaleza.

Realizar un análisis estadístico de los fenotipos obtenidos a diferentes tiempos en la simulación.

Comparar el comportamiento evolutivo de un modelo con expansión de información contra uno sin expansión.

Materiales y Métodos

Se desarrolló el programa *Bichos* mediante una serie de algoritmos en una computadora compatible con el *IBM PC* usando el lenguaje de programación *Borland Delphi®* versión 7 basado en *Object Pascal* bajo el sistema operativo Windows XP que simula una población de organismos evolucionando en un entorno gráfico, alimentándose, reproduciéndose, mutando y muriendo. La programación está apegada en la medida de lo posible al modelo de objetos, lo cual facilitó el manejo de las propiedades de cada individuo.

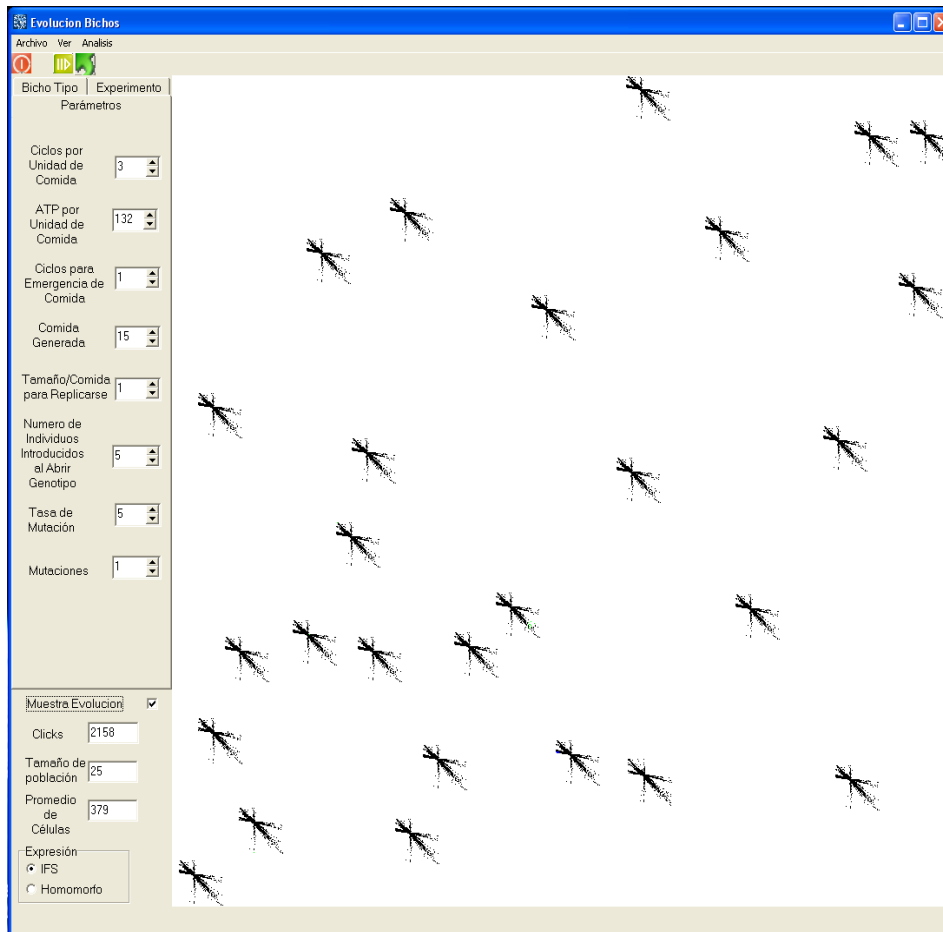


Figura 1. Ventana del programa “*Bichos*”.

Los individuos son organismos pluricelulares de reproducción asexual cuya morfología está determinada por el arreglo de sus células, las cuales no están diferenciadas y en su totalidad se especializan en la captación de alimento. Son heterótrofos, no pueden producir la energía necesaria para vivir y reproducirse, debido a esto, dependen de la existencia de productores primarios para su supervivencia.

Los organismos comparten el ambiente con productores primarios unicelulares que carecen de la capacidad de movimiento y que se reproducen a una tasa que puede ser regulada por el usuario, de igual manera es posible ajustar la cantidad de energía aportada por dichos productores a los consumidores, dando como resultado el control por parte del usuario sobre la presión selectiva del ambiente al disminuir la tasa de reproducción y la cantidad de energía de los productores.

Los consumidores transitan en el medio describiendo una trayectoria browniana modificada consumiendo una cantidad de energía directamente proporcional al número de células con las que cuenta, cuando en su recorrido una de sus células hace contacto con un productor lo ingieren y acumulan la energía aportada por éste.

Los individuos son capaces de producir copias de sí mismos, cuentan con un programa genético que incluye un genoma representado por una cadena de símbolos binarios que puede ser alterada por el cambio de valor de un elemento en caso de ocurrir una mutación, además de un mecanismo de expresión del material genético que determina el arreglo de sus células. Cuando alcanza un determinado estado energético que está relacionado con su número de células, (a mayor cantidad de células, mayor cantidad de energía necesaria para reproducirse) el individuo produce una copia de sí mismo y pierde la mitad de su energía, la cual es pasada al “hijo” junto con la información genética para hacer la copia. El desarrollo o morfogénesis se lleva a cabo en un instante y los individuos nacen con todas sus funciones y forma definitivas.

Dentro de la población existen individuos con variaciones en la morfología; el arreglo y número de células se ve alterado cuando ocurre una mutación en el genoma, dichas diferencias en la morfología tienen como consecuencia una capacidad reproductiva diferencial entre los individuos, permitiendo de esta manera la optimización de las características fenotípicas de la población para determinados ambientes.

El Universo

El Universo es finito y limitado (no toroidal como en algunos Autómatas celulares), se trata de un rectángulo de 800×800 píxeles de color negro que corresponden a cada localidad dentro del mismo. Por conveniencia, el eje coordenado “y” está invertido siguiendo los estándares de manejo gráfico en los sistemas de cómputo. Los eventos suceden en intervalos de tiempo discretos que denominaremos con el nombre de *Ticks*, cuya duración en tiempo real (segundos, minutos, etc.) depende de las capacidades del Hardware en el que se ejecute la simulación y de la cantidad de individuos en el universo, de tal manera que la velocidad del procesador y la cantidad de individuos son factores determinantes de la duración de los *Ticks*, pero no intervienen en el resultado final de una simulación. En un *Tick* cada individuo realiza un desplazamiento en una dirección determinada por un algoritmo modificado de movimiento browniano y consume una cantidad de energía determinada en función de su número de células, puede “reproducirse” (generar un nuevo individuo a partir de su propio genoma y energía), “mutar” (alterar su genoma) o “morir” (desaparecer del Universo) además, se genera un número de partículas de comida elegible por el usuario.

Cada individuo está confinado al Universo, por lo que la localidad central (x, y) de cada individuo cumple con la siguiente desigualdad:

$$25 \leq x \leq 800 \text{ y } 25 \leq y \leq 800$$

La Comida.

Al inicio, la comida está representada por 5000 píxeles de color amarillo que se generan y colocan aleatoriamente en el universo, y se regeneran a un ritmo ajustable por el usuario,

Si $Ticks \equiv C_a \Rightarrow$ Aparece C_g partículas de comida en localidades aleatorias

Ticks es el número de ciclos del universo (contador absoluto).

C_a es el número de *Ticks* que deben pasar antes de que aparezca la comida.

C_g es el número de partículas de comida que aparecen cada C_a *Ticks*.

De igual forma, la cantidad de energía aportada por cada unidad de alimento puede ser ajustada en la GUI¹ para que soporte un determinado número de células en los organismos por *Tick* del universo. El consumo de cada individuo está dado por

$$\Delta ATP = ATP - \frac{A}{C_c} + 1$$

ATP es la cantidad de comida almacenada por el individuo.

A es el área en número de píxeles que lo conforman.

C_c es el número de células soportadas por unidad de comida.

Los Organismos.

El genotipo de cada individuo está almacenado en un arreglo de 48 *Bytes*².

Gen : Array[0..GLen-1] of Byte;

En total 48x8=384 *bits*, cuyo estado heredan a sus descendientes con una probabilidad de mutación ajustable por el usuario desde un *bit* por nacimiento.

Se manejaron dos mecanismos de expresión de información genética para crear el fenotipo de cada organismo, uno denominado *simple* y el otro llamado *IFS*.

Mecanismo Simple

En el mecanismo simple, los bits son interpretados por pares en una de cuatro “bases”, U, R, D y L correspondientes a Up, Right, Down y Left respectivamente, siguiendo la equivalencia que se presenta a continuación:

¹ GUI = *Graphic User Interface* o Interfase Gráfica de Usuario.

² Byte. Arreglo de ocho bits.

	0	1	<i>1^{er} bit</i>
0	U	R	
1	D	L	

2^o bit

Cada letra especifica un desplazamiento unitario en la dirección arriba (U), derecha (R), abajo (D) e izquierda (L). La coordenada de inicio es la (0,0) y es a partir de este punto donde se inicia el crecimiento en función de la secuencia de bases.

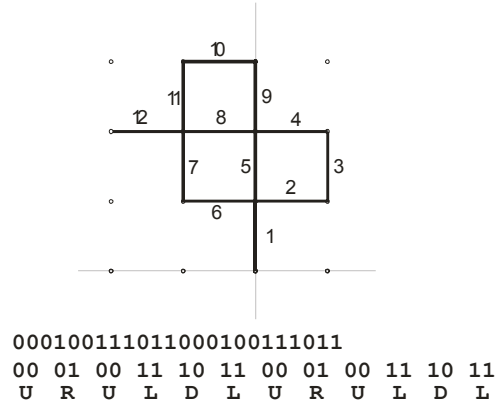


Figura 2. En la figura anterior se muestra un trozo de código, cómo es agrupado por pares de bits y su equivalencia en bases. Una vez obtenida la equivalencia, se inicia el trazado en el origen de coordenadas siguiendo el desplazamiento como se indica.

Cuando se ha calculado el fenotipo en su totalidad, se escala para cubrir una matriz de 50×50 píxeles, esto debido a que la longitud del genoma es constante y para limitar la extensión de fenotipos largos, como en el caso del genotipo (RRRR...RRR) que tornarían complicado el manejo en la computadora de los organismos. Dado que el genoma consta de 384 bits, el número de bases (y por lo tanto la longitud máxima) es de 192. Aun cuando el número de píxeles del fenotipo puede ser mayor que el número de bits del genotipo, no hay un aumento sustancial en la cantidad de información, ya que es relativamente fácil determinar el genotipo a partir de un fenotipo dado.

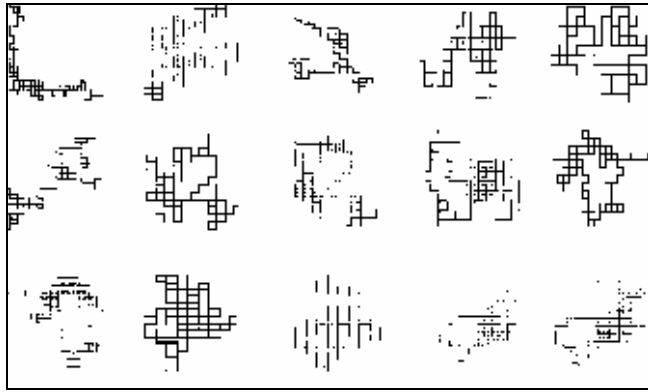


Figura 3. Algunos de los fenotipos obtenidos cuando se someten diferentes genotipos al algoritmo de expresión *IFS*.

Mecanismo *IFS*

El algoritmo *IFS* (*Iterated Function System*) modificado es usado en la simulación para la generación de formas complejas a partir de una cantidad reducida de información. En este caso, se decidió usar 384 *bits* (48 Bytes) en el genotipo y convertirlos en 24 números reales en un rango de -1 a 1 que sirven para alimentar cuatro transformaciones del algoritmo *IFS* (Anexo 1) dando como resultado una morfología autoescalada de 50×50 píxeles con una cantidad de información de 2500 *bits*, lo cual equivale a 6.5 veces la cantidad del genotipo.

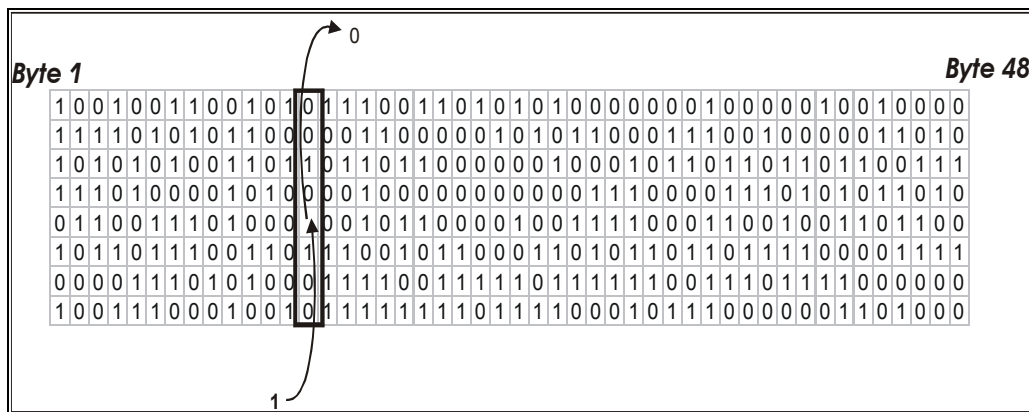


Figura 4. Representación de los 384 bits que componen el genoma de un individuo, las mutaciones ocurren en un bit al azar durante el nacimiento de un individuo

Debido a la naturaleza del algoritmo, no es fácil recuperar un genotipo a partir del fenotipo, a menos, por supuesto, que se conozcan las transformaciones.

Las células de los individuos son idénticas, no hay diferenciación ni funcionalidad específica, sólo sirven para alimentarse.

Dado que cada transformación está compuesta de seis números reales y cada *IFS* está compuesta de cuatro transformaciones, son necesarios 24 números reales para definir un *Morfo*. Esto da $384/24$ o sea 16 *bits* (dos *bytes*) por número. Cada elemento de la transformación está compuesto por cuatro dígitos los cuales son obtenidos a partir de cada *nibble*³ de esos dos *bytes*, lo que da un total de cuatro dígitos por elemento. La paridad del primer *nibble* determina el signo del elemento (positivo si es par y negativo si es impar). Después de esto, se verifica que el dígito generado por cada *nibble* subsiguiente (segundo a cuarto) no exceda el valor de nueve. Para esto, se compara el valor contra ocho y si lo excede o es igual, el *nibble* es enmascarado a través de la operación lógica *AND* con el número nueve (1001 en binario)

³ Nibble. Arreglo de 4 bits.

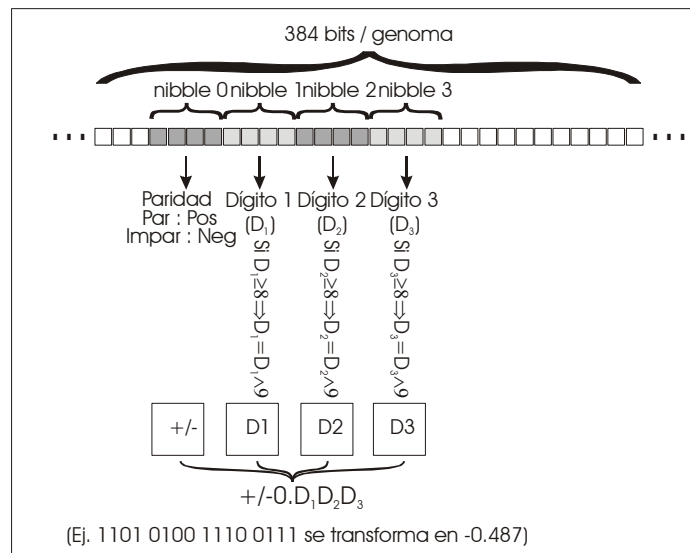


Figura 5. De esta manera se obtienen 24 números reales entre -0.999 y 0.999. Es importante notar que este intervalo fue elegido para evitar que la transformación modifique la escala de manera drástica. El factor de amplificación está dado por el determinante de A_i y si cada $a_{ij} < 1$ es fácil asegurar que $|A_i| < 2$ al menos.

La expresión del genotipo se realizó generando imágenes de 50×50 píxeles a partir de una matriz de elementos binarios que indican la presencia o ausencia de una célula en una localidad determinada, dicha matriz puede ser expresada de la siguiente manera.

$$\begin{array}{ccc}
 x_1, y_1 & \cdots & x_{50}, y_1 \\
 \vdots & & \vdots \\
 x_1, y_{50} & \cdots & x_{50}, y_{50}
 \end{array}$$

El valor de cada elemento de la matriz puede ser 1 o 0, si es 1, la locación (x, y) de una contraparte gráfica contará con un píxel de un color distinto al del universo el cual será capaz de capturar alimento, un valor de 0 indica la ausencia de célula en una locación.

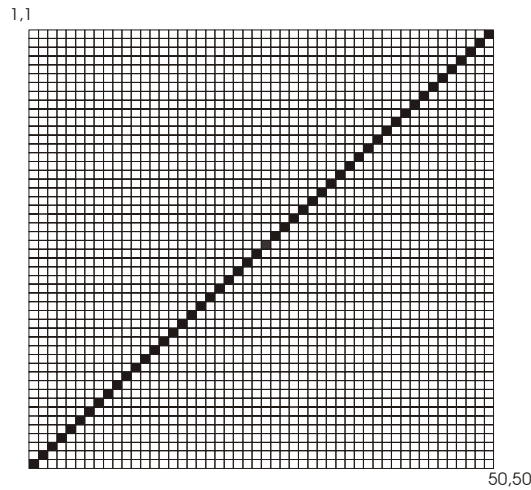


Figura 6. Fenotipo de 2500 bits (cada cuadro es un bit debido a que puede adoptar sólo dos valores 1 o 0), construido en un entramado de elementos capaces de alojar células, los cuadros en negro indican la presencia de células.

Para determinar que localidades deberán alojar una célula se obtienen cuatro transformaciones $A_1 \dots A_4$ a partir del genotipo. Cada transformación está compuesta por seis números reales en un rango de -0.999 a 0.999 , arreglados en una matriz de 2×2 ($a_{11} \dots a_{22}$ la transformación) y un vector de 1×2 (b_1 y b_2 desplazamiento).

$$A_1 = \left(\begin{array}{cc|c} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \end{array} \right)$$

$$\vdots$$

$$A_4 = \left(\begin{array}{cc|c} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \end{array} \right)$$

En cada iteración⁴ es utilizada una de las cuatro transformaciones antes descritas, la selección de dicha transformación es aleatoria y la probabilidad de elegir una en particular está dada por el valor porcentual del determinante de cada matriz.

⁴ Repetición de un proceso.

$$\begin{aligned}
|A_1| &= a_{11}^1 \cdot a_{22}^1 - a_{12}^1 \cdot a_{21}^1 \\
&\vdots \\
|A_4| &= a_{11}^4 \cdot a_{22}^4 - a_{12}^4 \cdot a_{21}^4
\end{aligned}$$

En esta fórmula, el superíndice indica el número de transformación. Debido a que la sumatoria será diferente de 1 en la mayoría de los conjuntos de transformaciones, es necesario normalizar las probabilidades para que el programa pueda escoger una de las cuatro para realizar la transformación.

$$\begin{aligned}
P^1 &= \frac{|A_1|}{\sum_i |A_i|} \\
&\vdots \\
P^4 &= \frac{|A_4|}{\sum_i |A_i|}
\end{aligned}$$

Para cada iteración de la función se genera un número aleatorio Z entre 0 y 1 y se procede a seleccionar la transformación a ser utilizada de la siguiente manera.

$$\begin{aligned}
\text{Si } Z > P^3 &\Rightarrow K = 4 \\
\text{Si } Z > P^2 &\Rightarrow K = 3 \\
\text{Si } Z > P^1 &\Rightarrow K = 2 \\
\text{Si } Z < P^1 &\Rightarrow K = 1
\end{aligned}$$

Donde K es el número de transformación A_k a ser utilizada. El superíndice indica el no. de transformación.

Al inicio de la *ontogenia* se generan valores aleatorios para la primera coordenada (x, y) , posteriormente estos son obtenidos mediante la iteración de la siguiente función.

$$A_k \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

Al inicio, se hace un estimado del dominio calculando los límites (x_{min}, x_{max}) y (y_{min}, y_{max}) para 1000 iteraciones. Durante el desarrollo, cada valor de (x, y) es forzado a permanecer dentro de los límites de 50×50 píxeles mediante la siguiente fórmula:

$$i = Int \left((x - X_{min}) \frac{50}{X_{Max} - X_{Min}} \right)$$

$$j = Int \left((y - Y_{min}) \frac{50}{Y_{Max} - Y_{Min}} \right)$$

Se itera 4000 veces la función y se obtiene una morfología de píxeles, cada conjunto de transformaciones tiene su propio atractor y genera siempre la misma forma.

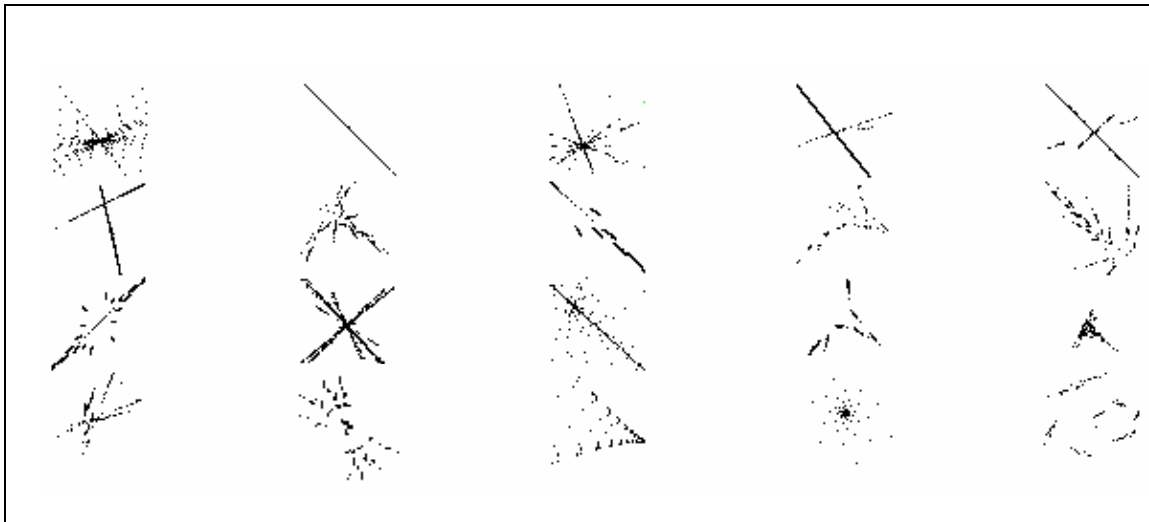


Figura 7. Algunos de los fenotipos obtenidos cuando se someten diferentes genotipos al algoritmo de expresión *IFS*.

La morfología es almacenada en una matriz para su posterior manipulación y al mismo tiempo se cuentan el número de píxeles encendidos (“células”).

Computacionalmente, los individuos son generados a partir de un objeto llamado *Tmorph* con las siguientes propiedades:

```

Class(TObject)
  X,Y      : Single;           // Posición calculada
  P        : Array[1..NTrans] of Single; // Probabilidad de c/transformación
  Pk       : Array[1..NTrans] of Single; // Probabilidad acumulada
  MaxX,MaxY,MinX,MinY : Double; // Límites del morfo

  Constructor Create;         // Crea el espacio para el morfo
  Procedure GetGen(Var Gen:TGenoma); // Extrae las transformaciones del genoma
  Procedure Step;            // Iteración sencilla
  Function Det(Var A:TMat):Single; // Función que entrega el determinante de una matriz
  Procedure Onotgeny(Var Feno:TFeno); // Genera el morfo.
End;

```

Movimiento

Cada individuo transita por el universo describiendo una trayectoria Browniana modificada donde se asigna una dirección nueva igual a la dirección de

movimiento anterior con una variación de ángulo limitada a $\pi/4$ a cada lado ($\pi/2$ en total).

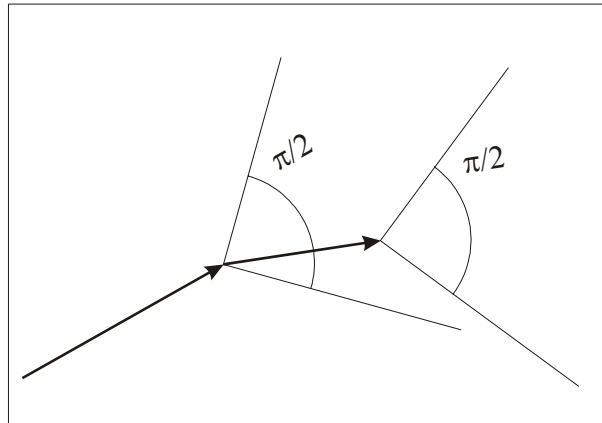


Figura 8. Trayectoria browniana seguida por el individuo, a diferencia de un movimiento completamente aleatorio donde el desplazamiento promedio es casi nulo (parece estar vibrando en un solo lugar), la trayectoria browniana sí presenta desplazamiento.

y su centro no puede cruzar los límites, esto es, la posición resultante de su trayectoria se haya acotada por las coordenadas limítrofes del Universo. La orientación de los individuos no cambia durante toda su vida. No cuentan con sensores ni inteligencia que les permitan identificar y/o dirigirse a los lugares con mayor densidad de alimento ni capturar un píxel que se encuentre inmediatamente fuera de su cobertura, de tal manera que la capacidad para conseguir alimento depende únicamente de su forma, entre mayor sea el área con comida que barre el individuo en su desplazamiento, mayor es su capacidad para alimentarse. Con cada comida el individuo obtiene una cierta cantidad de energía, misma que puede ser modificada por el usuario y que es consumida a un ritmo determinado por el número de píxeles que compone cada individuo, los que tienen un mayor número de píxeles tienen un mayor gasto energético que los que tienen pocos. Esto ayuda a establecer un equilibrio entre tamaño y consumo que permite obtener morfologías no triviales de tal manera que un individuo con el máximo número de células (2500) formaría un cuadrado de 50x50 células y tendría una gran eficiencia en la captación de alimento, pero lo consumiría a un ritmo muy acelerado lo cual representa una desventaja.

Reproducción

Cuando un individuo acumula un número de comida proporcional al número de células con las que cuenta, éste se reproduce, generando un nuevo individuo con el mismo genoma (a menos que ocurra una mutación) y dividiendo la cantidad de alimento en partes iguales entre el padre y el hijo. El nuevo individuo aparece en la misma posición que el padre, pero debido al carácter azaroso del movimiento, siguen trayectorias distintas inmediatamente.

$$\text{Si } ATP > \left(\frac{A}{ATP} \right) \Rightarrow \text{Re produce}$$

ATP es la cantidad de comida almacenada por el individuo.

A es el área del individuo en número de píxeles.

ATP_{rep} es el inverso de la cantidad mínima de alimento por celda para reproducirse. Esto es, a mayores valores de *ATP_{rep}* menor es la cantidad de alimento necesaria. El valor inicial es 1 (tanto alimento como área del individuo).

Bicho Tipo

Con el fin de comparar los patrones evolutivos de las poblaciones artificiales dotadas con uno u otro de los mecanismos de expresión genética propuestos se desarrollaron instrumentos informáticos que permiten el acceso al “registro fósil” el cual, cabe señalar, es completo tanto en tiempo como en individuos, es decir, se trabajó con el total de las características fenotípicas de cada individuo de la población a intervalos de tiempo muy pequeños para realizar el análisis filogenético.

El *Bicho tipo* se construyó con la finalidad de tener una referencia representativa de todos los fenotipos de la población en un momento determinado, de esta manera se puede construir un *Registro Fósil* con individuos “promedio” de la población en diferentes tiempos.

Debido a que la morfología de los individuos está constituida por la presencia o ausencia de células en las diferentes localidades de un arreglo matricial bidimensional y a que la naturaleza cerrada del universo en el que se lleva a cabo la simulación permite encontrar a todos los individuos de una población

en cualquier momento, es posible calcular con precisión la probabilidad de encontrar una célula en una determinada localidad del arreglo matricial que conforma a los organismos artificiales, esto se logra mediante el uso de un algoritmo que recorre todas y cada una de las localidades del entramado celular de cada uno de los individuos existentes en un momento determinado de la simulación, la probabilidad de encontrar una célula determinada en la población está dada por el número de individuos que la comparten entre el número total de individuos.

$$Pc_{x,y} = \frac{n}{N}$$

Pc es la probabilidad de encontrar la célula $[x, y]$.

n es el número de individuos que comparten dicha célula.

N es el numero total de individuos de la población.

La construcción del *Bicho Tipo* se realizó a intervalos de *Ticks* establecidos por el usuario mediante el uso de un umbral de probabilidad preestablecido en 50%, pero que puede ser modificado en la GUI, si la probabilidad de encontrar una determinada célula es mayor al umbral establecido, entonces dicha célula existirá en el *Bicho Tipo* de lo contrario no.

$$\text{Si } Pc_{x,y} > \text{Umbral} \Rightarrow \text{Célula}_{x,y} \text{ Existe}$$

Esto se expresó Computacionalmente de la siguiente manera:

```

Procedure TFormMain.CalcEvo;

Var   I,Sum,
      X,Y,
      Max,Posi:Integer;

begin
  For Y:=-BSize2 to BSize2 Do //Recorre todas las celulas
    For X:=-BSize2 to BSize2 Do
      Begin
        Sum:=0;
        For I:=1 to NBichos Do //Recorre todos los bichos
          If Pueblo[I].Feno[X,Y]<>0 Then //Si la celula x,y existe
            Inc(Sum); //incrementa el numero de bichos que la poseen
          Evo[X,Y]:=Sum/NBichos; // Calcula la probabilidad de aparecer para la
        End; //celula x,y
        SetLength(Tipos,High(Tipos)+2); //Guarda registro fosil de Bichos tipo
        Max:=FormEvo.Bar.Max; //prepara el umbral para generar bicho tipo
        Posi:=FormEvo.Bar.Max-FormEvo.Bar.Position;
        For Y:=-BSize2 to BSize2 Do //recorre todas las celulas para generar bicho tipo
          For X:=-BSize2 to BSize2 Do
            Begin
              //escribe matriz de probabilidades
              FormEvo.SG.Cells[X+BSize2+1,Y+BSize2+1]:=FloatToStr(Evo[X,Y]);
              If Max*Evo[X,Y]>=Posi Then //Si la probabilidad de la celula x,y > umbral
                Tipos[High(Tipos)][X,Y]:=1 //pinta dicha celula en bicho tipo
            Else

```



```

Tipos[High(Tipos)][X,Y]:=0;//si no, no se pinta
End;
FormEvo.PBTipo.OnPaint(Self);
PBTipo.OnPaint(Self);
end;

```

Para las simulaciones hechas en este trabajo el intervalo para la creación de *Bichos Tipo* fue de 200 *Ticks* y el umbral de probabilidad de 50%.

Gráficas de Disimilitud y Disimilitud Acumulada.

Con la finalidad de visualizar la acumulación de cambios fenotípicos a través del tiempo y establecer así patrones evolutivos se generaron gráficas de *Disimilitud* y *Disimilitud Acumulada* entre los *Bichos Tipo* en intervalos de tiempo de 200 *Ticks*.

Los datos utilizados para graficar la *Disimilitud* se obtuvieron del programa *Bichos* mediante un algoritmo que evalúa las diferencias fenotípicas (arreglo de células) entre *Bichos Tipo* temporalmente consecutivos.

$$D \text{ de Tipos}_{i,j} = 1 - \frac{N_c}{AT_i + AT_j / 2}$$

D Es la proporción de similitud entre los *Bichos Tipo* i y j

N_c Es el número de células que comparten los *Bichos Tipo* i y j

AT_i, AT_j Son el número de células de los *Bichos Tipo* i y j respectivamente.

La *Disimilitud Acumulada* para cada *Tiempo* se obtuvo mediante un algoritmo que suma las *Disimilitudes* de los *Bichos Tipo* consecutivos y las acumula con respecto al tiempo.

$$Da_t = Da_{t-1} + D_t$$

Da Es la *Disimilitud Acumulada*.

D Es la *Disimilitud*.

t Es el tiempo entre la generación de un *Bichos Tipo* y otro.

Los valores de *Disimilitud*, *Disimilitud Acumulada* y *Tiempo* fueron organizados en una matriz de datos de la siguiente forma:

$$\begin{array}{ccc}
 t_1 & D_1 & Da_1 \\
 \vdots & \vdots & \vdots \\
 t_n & D_n & Da_n
 \end{array}$$

Para la generación de las graficas los datos fueron exportados en una archivo de formato *.csv (Coma Separated Values) para ser procesados en Microsoft© Excel 2000.

Tiempo	Disimilitud	Disimilitud Acumulada
200	0.01	0.1
400	0.015	0.115
600	0.024	0.139
800	0.015	0.154
1000	0.1	0.254
1200	0.12	0.374
1400	0.23	0.604
1600	0.19	0.794
1800	0.01	0.804
2000	0.013	0.817
2200	0.038	0.855
2400	0.45	1.305

Tabla 1. Ejemplo de las tablas de Excel obtenidas con el Programa *Bichos*.

Para graficar los datos se elaboro un formato personalizado de gráficos de Excel en el que se incluyen dos ejes “y”, el izquierdo para *Disimilitud* y el derecho para *Disimilitud Acumulada*, mientras que el tiempo es expresado en el eje “x”. A continuación se muestra un ejemplo de una gráfica de *Disimilitud* y *Disimilitud Acumulada* con los datos de la *Tabla 1*.

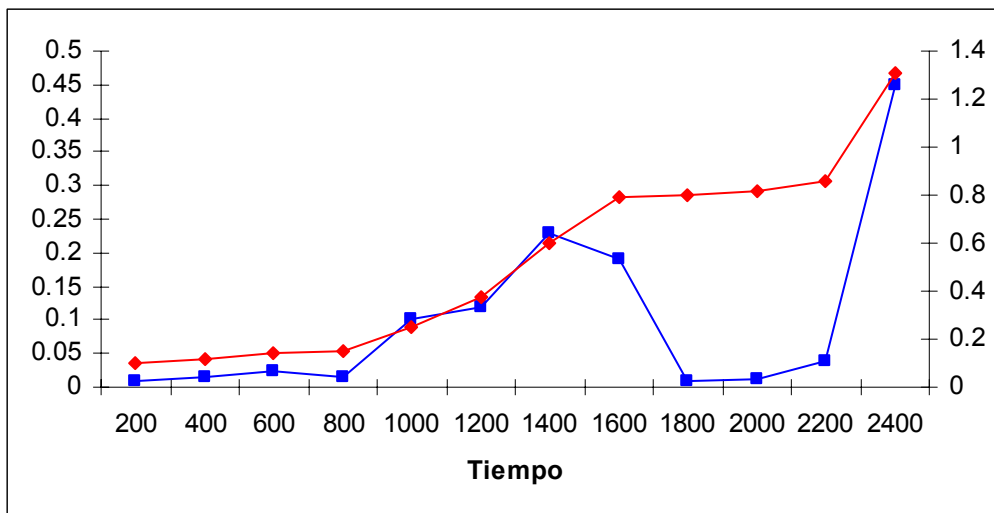


Figura 9. Gráfica de *Disimilitud* y *Disimilitud Acumulada*, el eje “y” de la izquierda representa la *Disimilitud* entre dos *Bichos Tipo* consecutivos temporalmente, mientras el de la derecha incorpora los valores de *Disimilitud Acumulada*, el eje “x” representa el tiempo. Es posible observar la aceleración de la acumulación de cambios fenotípicos en los tiempos en los que la *Disimilitud* alcanza picos máximos, en estos tiempos la pendiente de la *Disimilitud Acumulada* aumenta como resultado de una mayor aportación a la acumulación de cambios.

Gradualismo y discontinuidad

Las gráficas de *Disimilitud Acumulada* permiten establecer patrones evolutivos debido a que en ellas se expresa el cambio fenotípico de una población a través del tiempo, en ellas es posible observar la presencia o ausencia de cambios fenotípicos así como su magnitud en un determinado intervalo.

Patrón gradualista.

Un patrón gradualista se puede definir como un comportamiento evolutivo en el cual la *Disimilitud Acumulada* crece por medio de pequeños y continuos cambios en el registro fósil a través del tiempo. A continuación se muestra una tabla con valores de *Disimilitud Acumulada* que generan un comportamiento gradual, así como la grafica obtenida.

Disimilitud	Acumulada
0.0013	0.0013
0.0016	0.0023
0.0015	0.0032
0.0014	0.004
0.001	0.0044
0.0015	0.0053
0.0015	0.0062
0.0016	0.0072
0.0015	0.0081
0.0012	0.0087

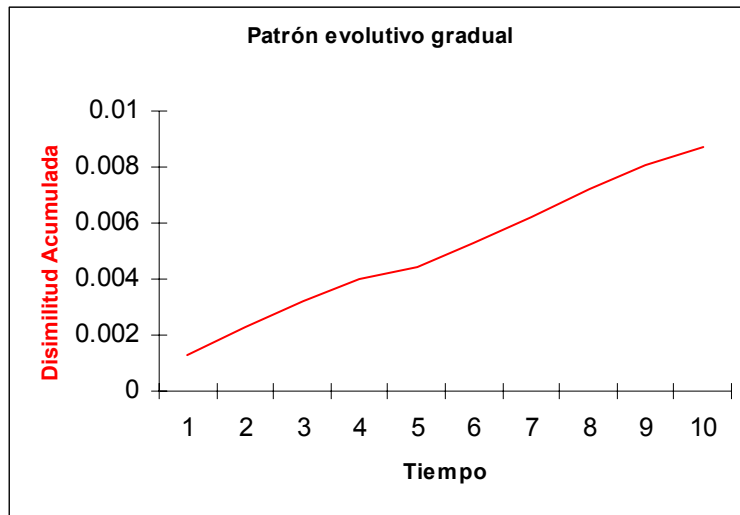


Figura 10 Patrón evolutivo gradual, se puede observar que la *Disimilitud* no difiere mucho entre dos *Bichos tipo* consecutivos y que la *Disimilitud Acumulada* graficada con respecto al tiempo genera una pendiente casi continua.

Este tipo de gráficas indican un cambio gradual de las poblaciones con respecto al tiempo ya que no presentan repentinos aumentos de la *Disimilitud* que se vean reflejados en la grafica de *Disimilitud Acumulada*, este tipo de cambio evolutivo se da de una manera similar a la propuesta por la TSE, ya que los cambios se van acumulando lentamente a través del tiempo. [3]

Patrón Discontinuo.

Los patrones discontinuos se caracterizan por la aparición de series con considerables aumentos de la *Disimilitud* entre *Bichos Tipo* consecutivos que se ven reflejados en la grafica de *Disimilitud Acumulada*, la continuidad de las pendientes en este tipo de gráficas se ve interrumpida por aumentos abruptos de la *Disimilitud*.

Disimilitud	Acumulada
0.0012	0.001
0.001	0.0022
0.009	0.0032
0.01	0.0122
0.02	0.0222
0.019	0.0422
0.018	0.0612
0.009	0.0792
0.001	0.0882
0.001	0.0892
0.0013	0.0902

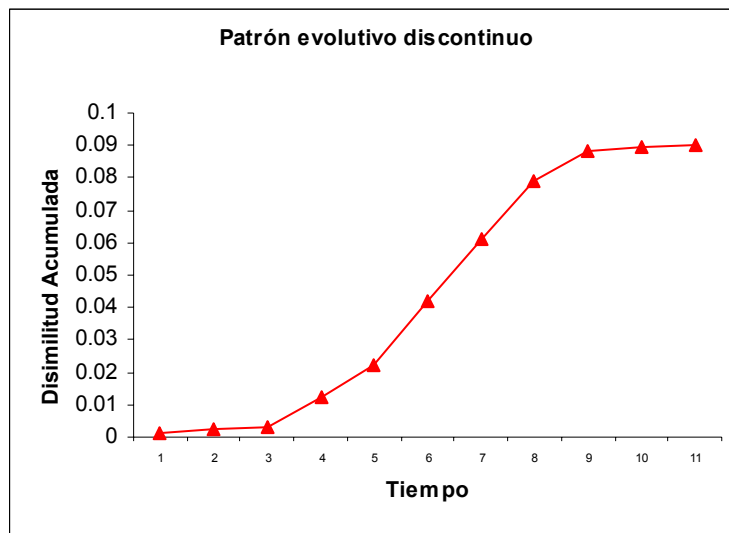


Figura 11. Patrón evolutivo discontinuo, se observa un aumento repentino de la *Disimilitud* (Celdas marcadas con negrita en la tabla de arriba) lo cual produce una discontinuidad en la pendiente de la gráfica de *Disimilitud Acumulada*.

La aparición de este tipo de gráficas indica un comportamiento evolutivo a “Saltos”, parecido al propuesto por la TEP.[6]

Gráficas de Promedio de Células.

Con la finalidad de observar la respuesta adaptativa de las poblaciones se obtuvo una tabla en formato *.csv del programa *Bichos* en la que se incluyen el número promedio de células en la población y la cantidad energética del alimento a diferentes tiempos.

Los datos obtenidos fueron procesados en Microsoft© Excel 2000 para la generación de las gráficas, estas están constituidas por dos ejes “y” en los que se grafican los valores de cantidad de energía por unidad de alimento y el número promedio de células en la población, el tiempo está representado en el eje “x”.

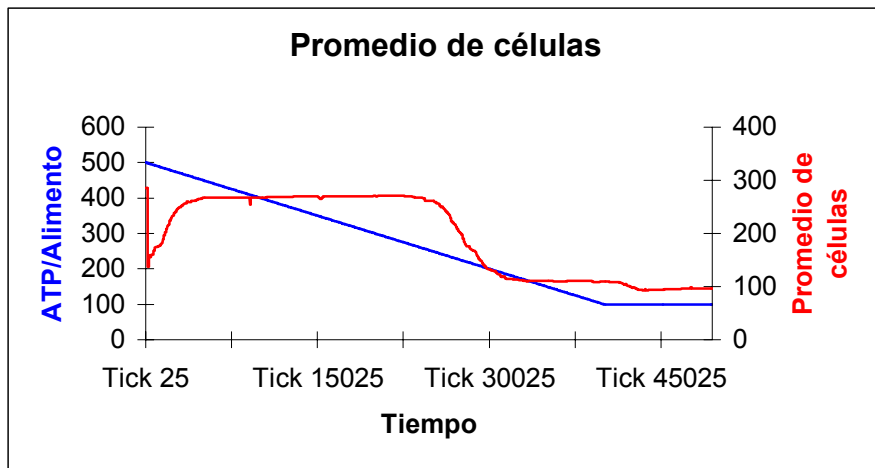


Figura 12. En esta gráfica se observa una disminución gradual de la energía contenida en el alimento (línea azul) con respecto al tiempo y la respuesta adaptativa en el número promedio de células (línea roja). En este ejemplo el número promedio de células cae como consecuencia de que la cantidad de energía disponible en el universo disminuye ya que el consumo de energía por parte de los organismos está determinado en función de su número de células.

Resultados

Se llevó a cabo una serie de 20 simulaciones en el Programa “Bichos” en una computadora IBM compatible con un procesador Intel Celeron® de 2.2 GHz con una memoria RAM de 256 MB bajo el sistema operativo Windows XP Profesional versión 2002. El diseño experimental constó de cuatro modelos de poblaciones artificiales diferenciados por el tipo de expresión genética y la variación en la presión ejercida por el ambiente a través del tiempo. Se realizaron cinco repeticiones para cada modelo poblacional y se observaron los patrones evolutivos de cada una mediante la graficación de la *Disimilitud Acumulada* a través del tiempo.

Modelo Poblacional	Expresión genética	Presión ambiental	Número de Repeticiones
1	<i>Simple</i>	No	5
2	<i>IFS</i>	No	5
3	<i>Simple</i>	Gradual	5
4	<i>IFS</i>	Gradual	5
5	<i>Simple</i>	Drástico	5
6	<i>IFS</i>	Drástico	5

Tabla 1 Modelos poblacionales para el diseño experimental.

La presión ambiental fue aumentada en cuatro de los modelos mediante la disminución del contenido energético en el alimento, reduciendo el número de células soportadas por unidad de alimento y aumentando el número de partículas de comida necesarias para reproducirse. Al inicio de las simulaciones en las que se ejerció presión ambiental a las poblaciones se comenzó en el *Tick 0* con una cantidad de energética de $500ATP/Partícula$ de Alimento y se concluyó en el *Tick 50,000* con un valor de $100ATP/Partícula$ de Alimento.

Las condiciones iniciales del sistema fueron las mismas para todos los modelos, se eligieron de tal manera que los ambientes creados por éstas no representaran presión selectiva para la mayoría de los fenotipos generados mediante los dos sistemas de expresión genética propuestos.

El número de ciclos por unidad de comida, el *ATP* por unidad de comida, el número de ciclos para la emergencia de comida y la comida generada determinan la cantidad de energía disponible en el sistema. El *tamaño/comida para replicarse* determina la cantidad de energía necesaria para que un individuo se reproduzca en función de su número de células. La tasa de mutación y el número de mutaciones determinan la frecuencia de mutaciones en la población.

Ciclos por Unidad de Comida	5
<i>ATP</i> por Unidad de Comida	500
Ciclos para la Emergencia de Comida	1
Comida Generada	30
Tamaño/Comida para Replicarse	1
Tasa de Mutación	5
Numero de mutaciones	1

Tabla 2. Parámetros iniciales preestablecidos para las simulaciones.

Modelo 1.

Poblaciones con expresión de información genética simple no sometidas a presión ambiental.

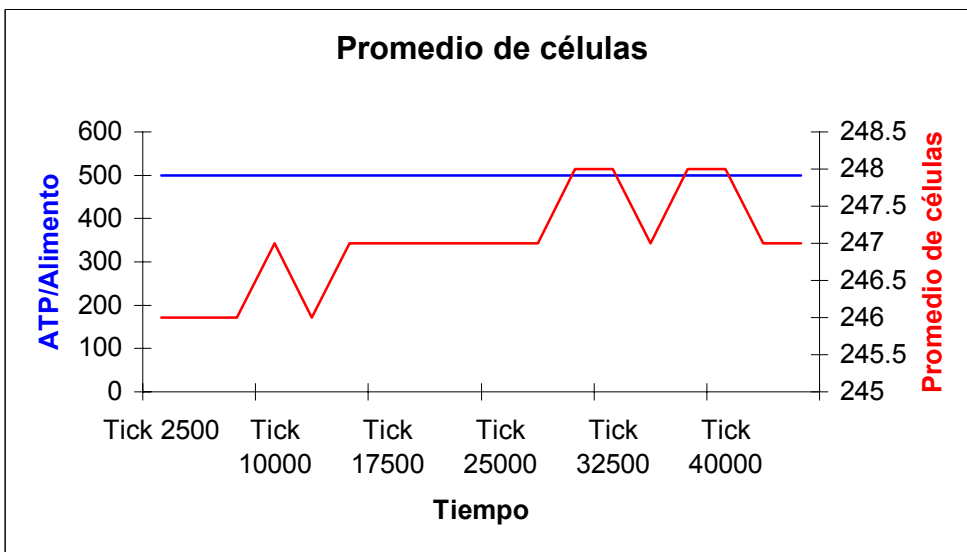
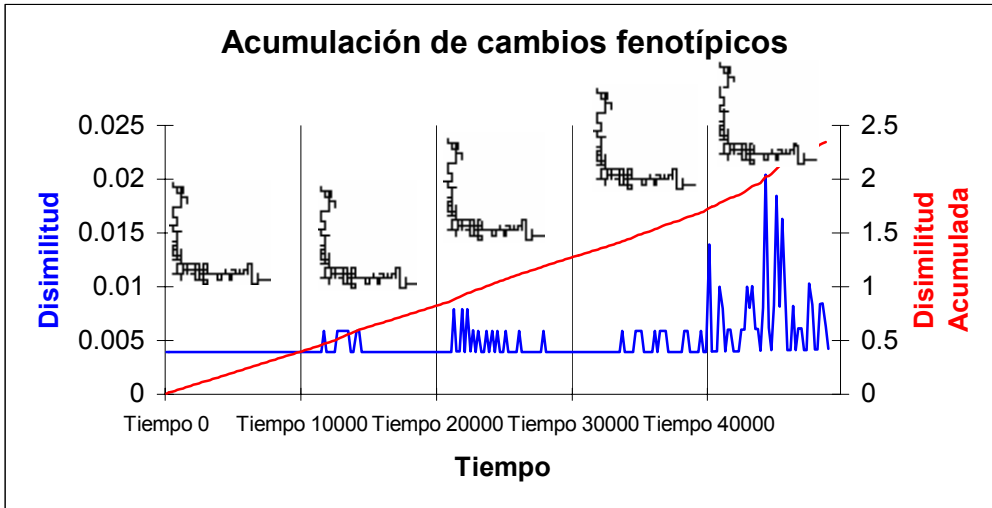
Se realizaron cinco repeticiones con este modelo de población, cada simulación se inicio con la generación de un individuo de genoma aleatorio y se dejó correr durante 50,000 *Ticks*, las condiciones ambientales en este modelo no fueron alteradas durante todo el tiempo que duró la simulación siendo ajustadas al inicio de la simulación con los parámetros preestablecidos (Tabla 2).

Se obtuvieron los siguientes resultados:

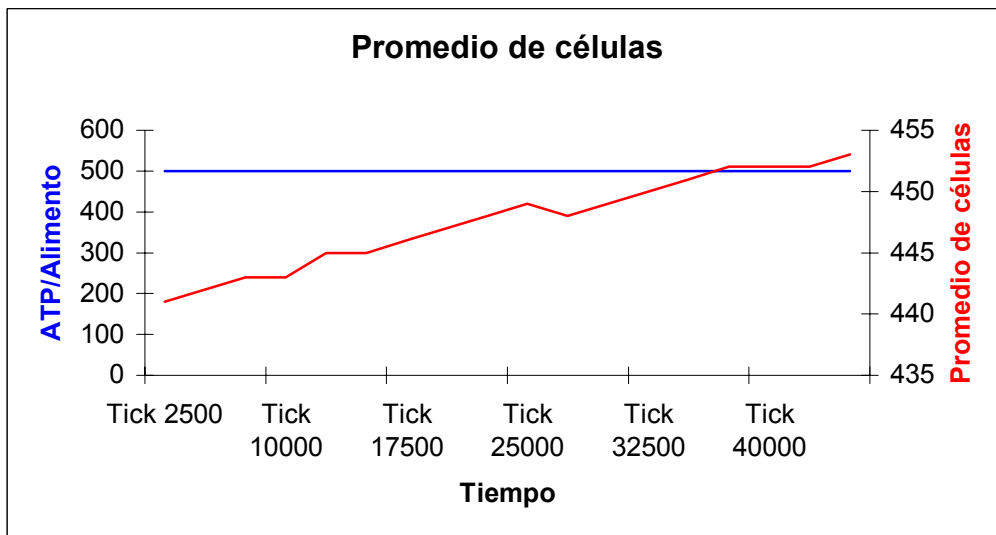
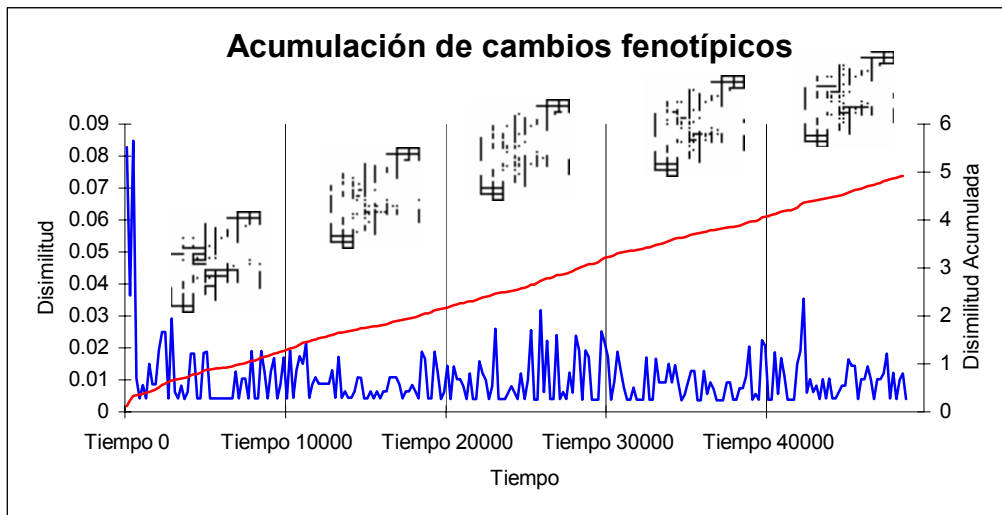
	<i>Disimilitud Máxima</i>	<i>Disimilitud Acumulada Máxima</i>	<i>Tasa de Evolución</i>
Repeticón 1	2.04	2.346	0.002258356
Repeticón 2	8.481	4.916	0.011538607
Repeticón 3	4.014	5.424	0.024972977
Repeticón 4	3.03	2.53	0.005946198
Repeticón 5	3.114	1.7	0.001271294

Tabla 3. Resultados de del modelo 1.

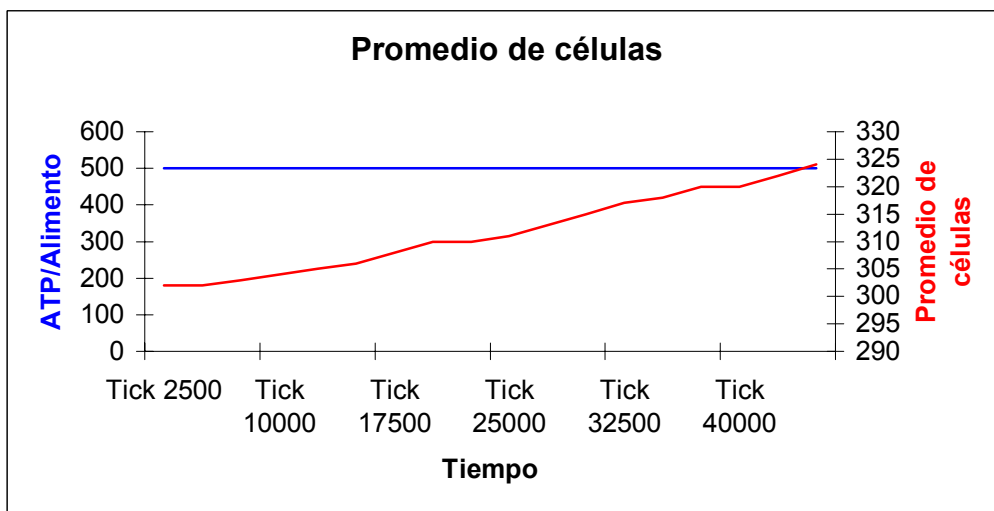
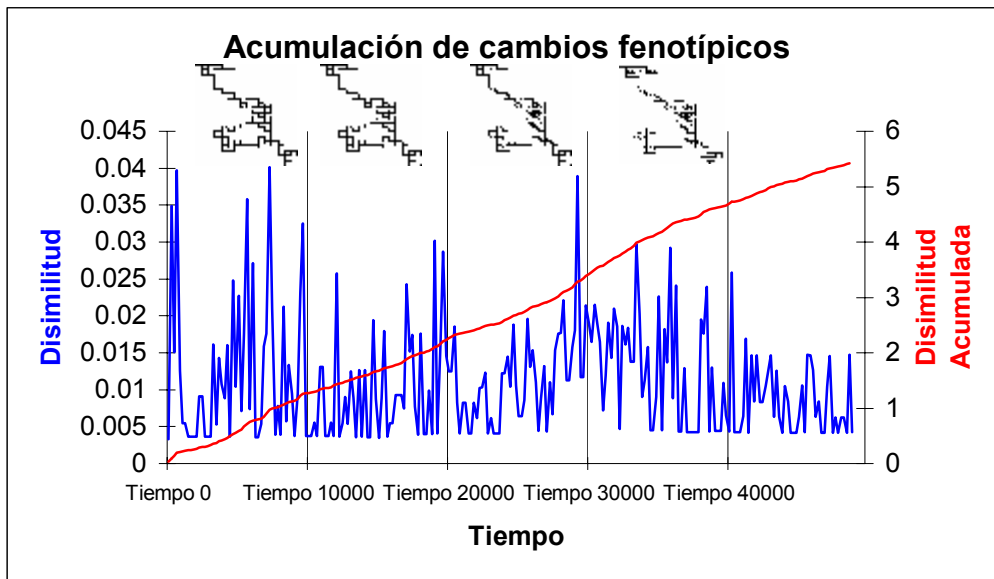
Repetición 1.



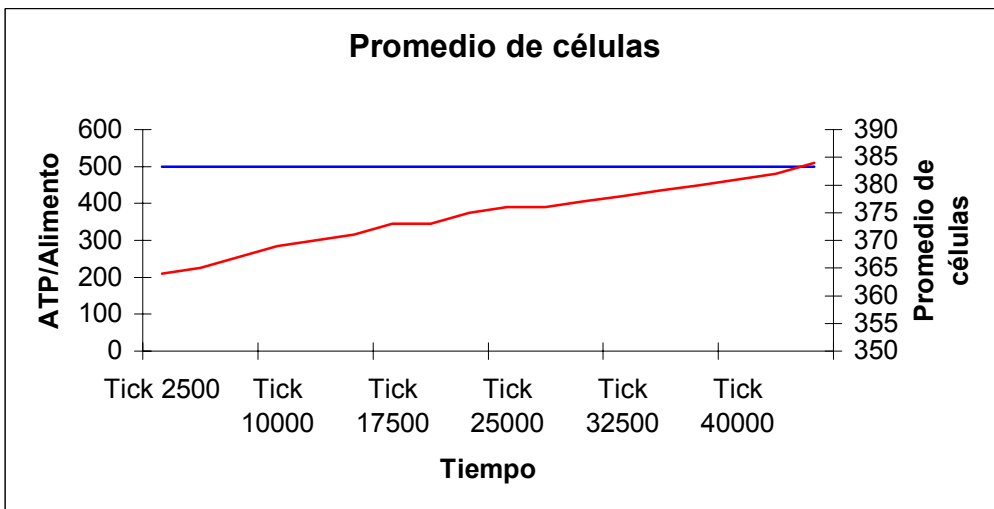
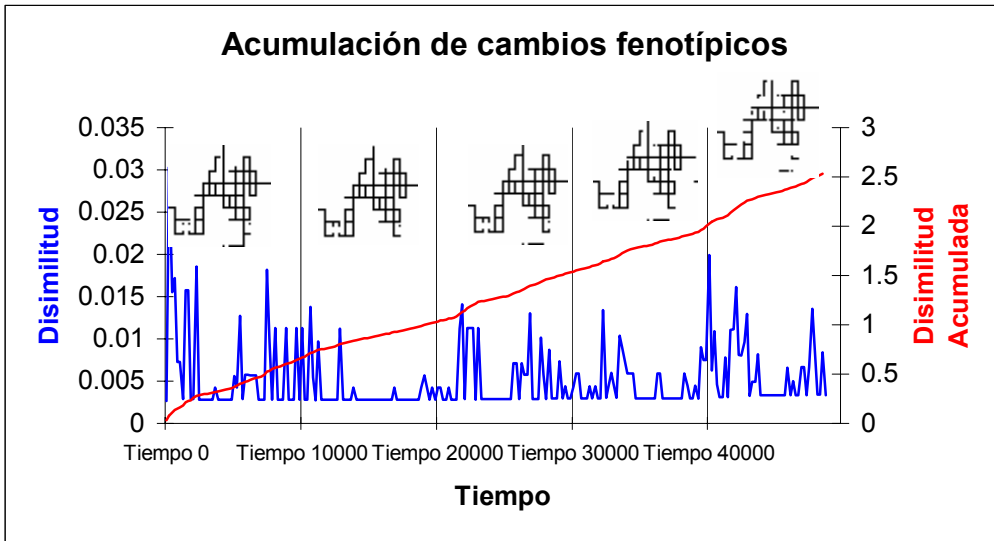
Repetición 2.



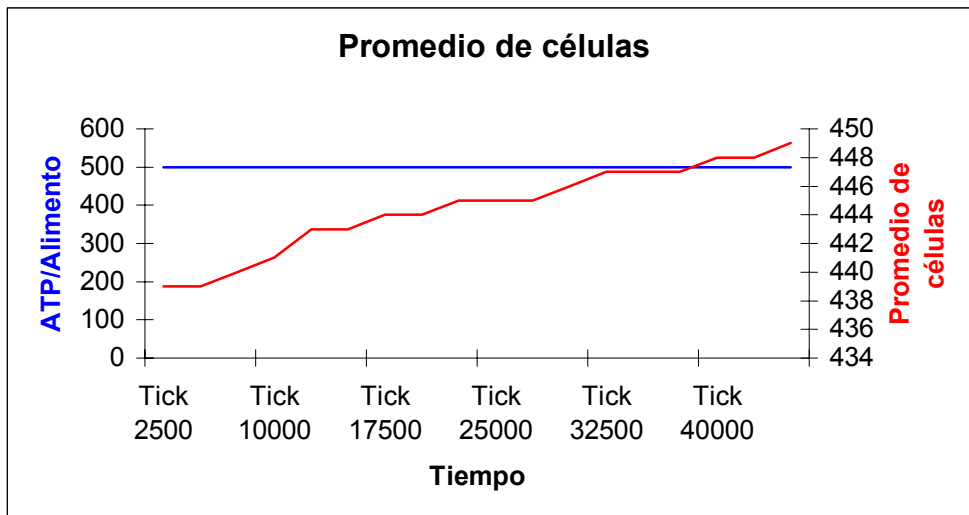
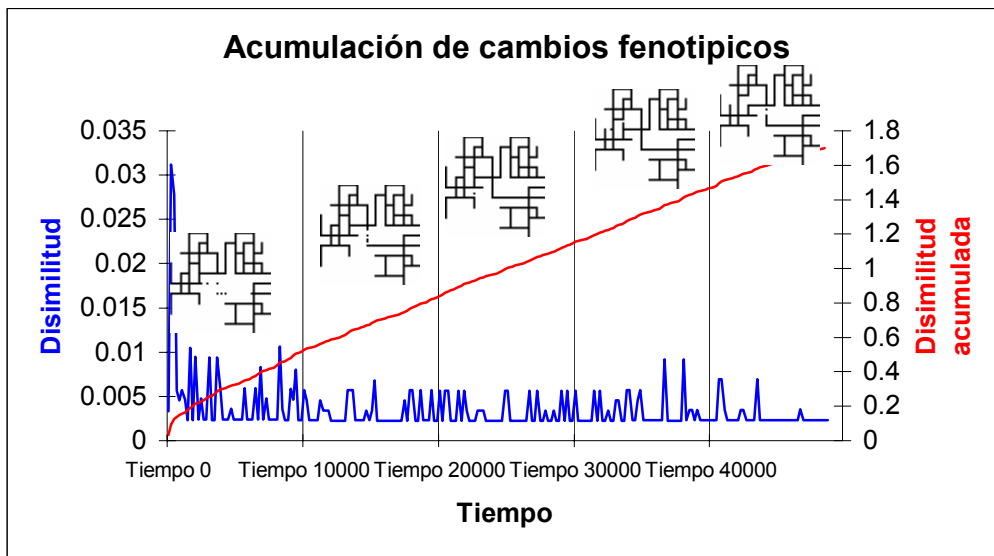
Repetición 3.



Repetición 4.



Repetición 5.

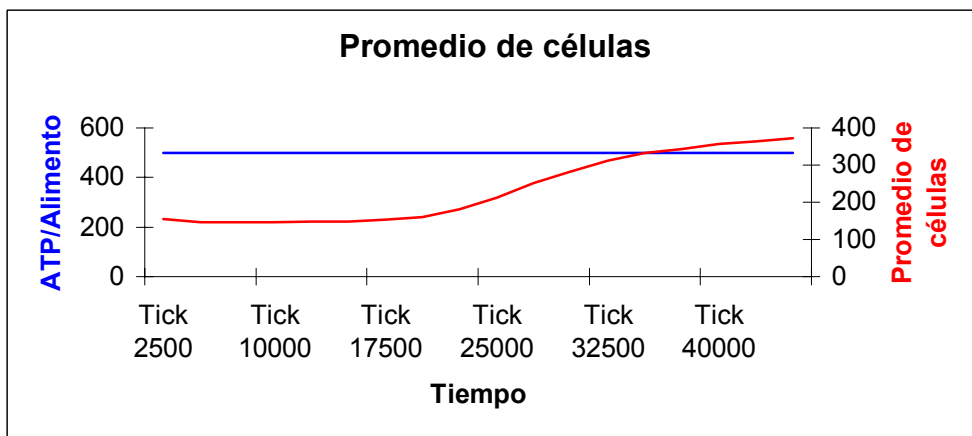
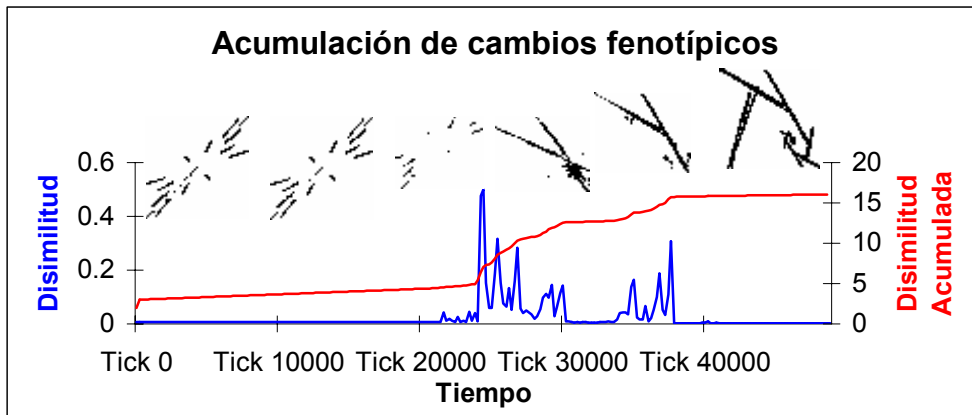


Modelo 2.

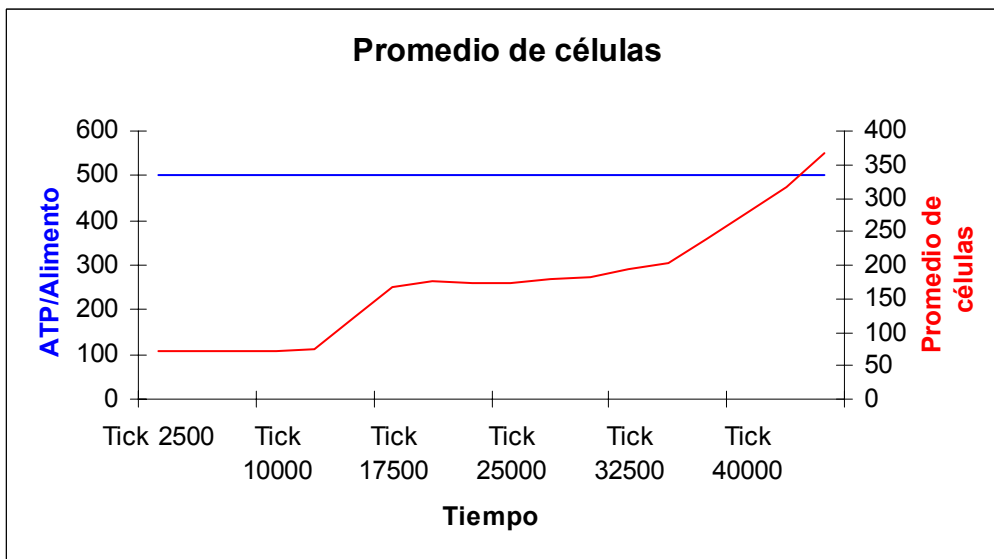
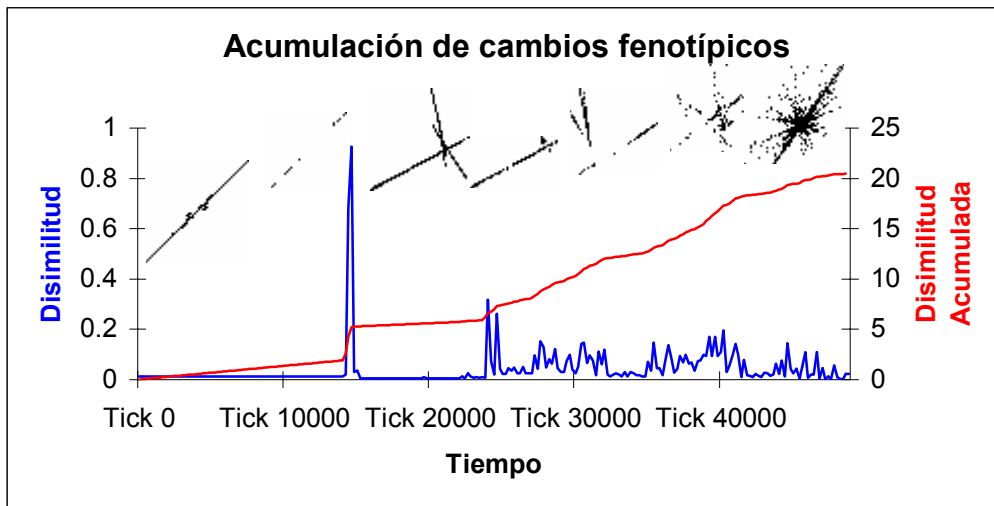
Poblaciones con expresión de información genética *IFS* no sometidas a presión ambiental.

	<i>Disimilitud Máxima</i>	<i>Disimilitud Acumulada Máxima</i>	<i>Tasa de Evolución</i>
Repetición 1	50	16.056	0.04638219
Repetición 2	92.592	20.49	0.045057315
Repetición 3	2.883	3.143	0.022982635
Repetición 4	52.941	12.642	0.047439249
Repetición 5	8.635	6.429	0.022899806

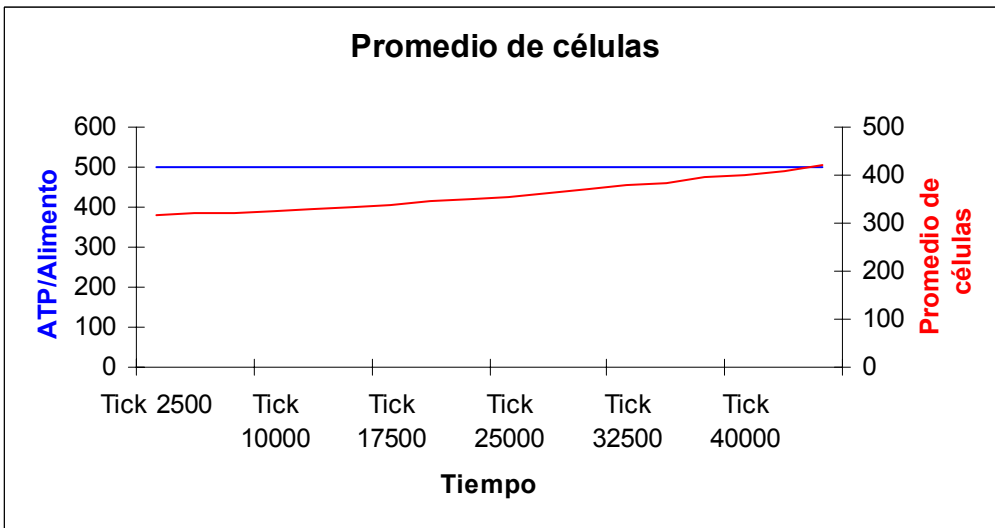
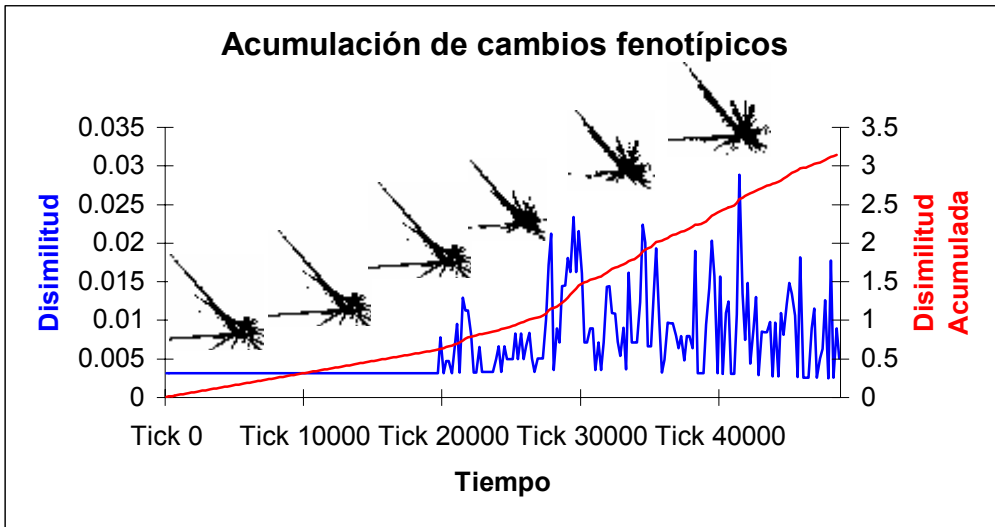
Repetición 1.



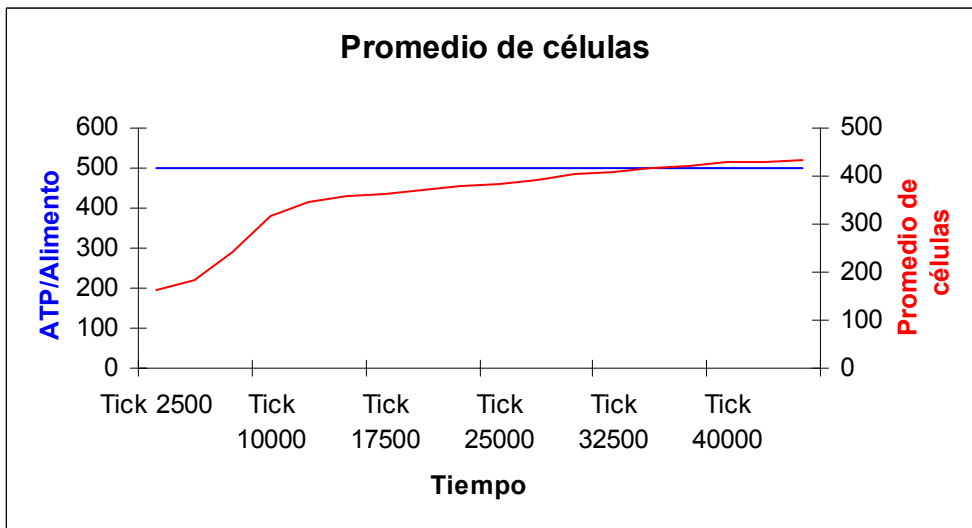
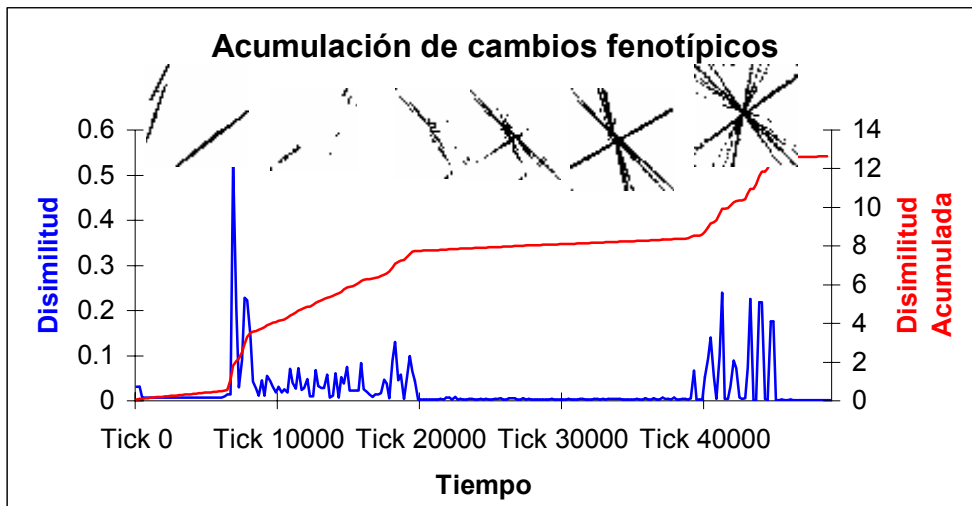
Repetición 2.



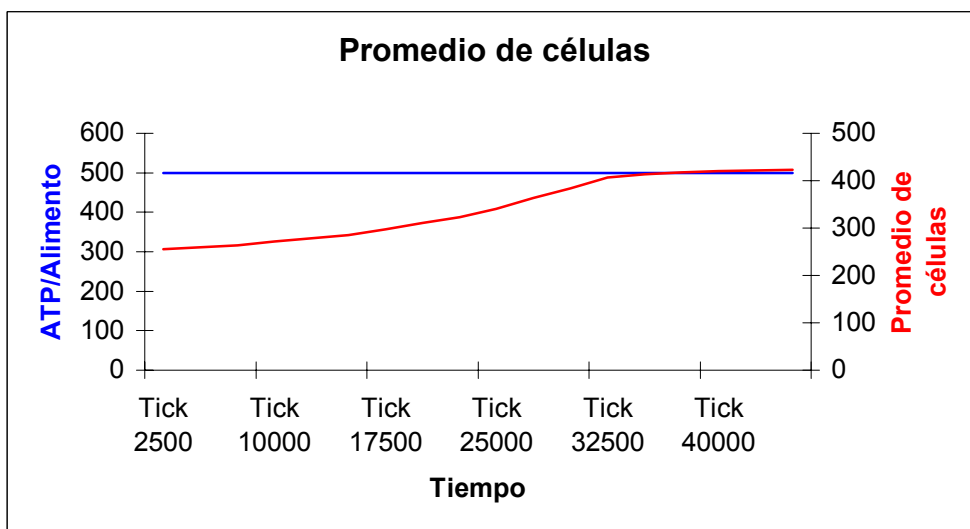
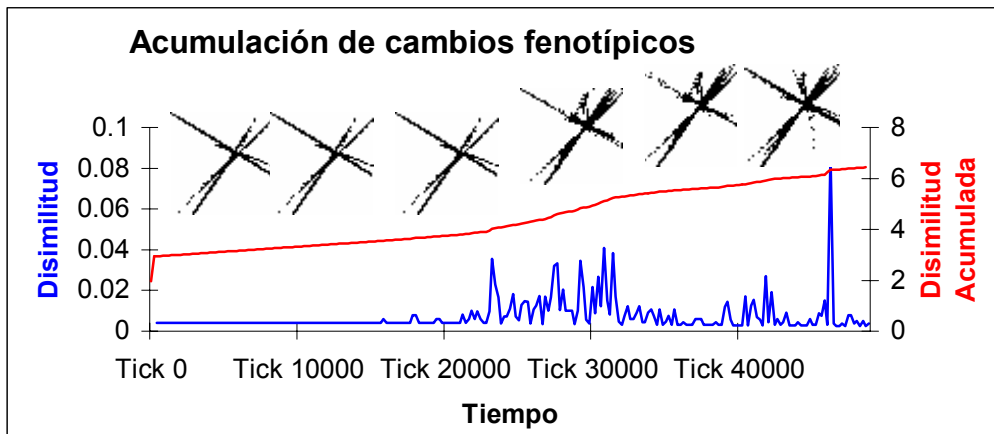
Repetición 3.



Repetición 4.



Repetición 5.

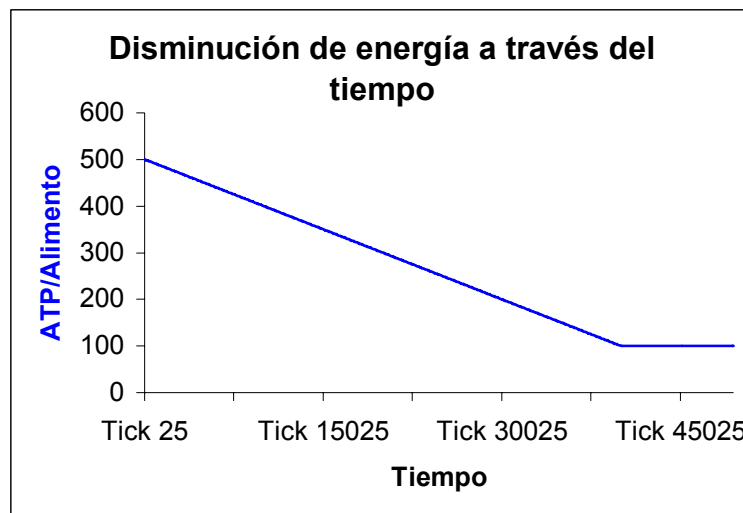


Modelo 3.

Poblaciones con expresión de información genética simple sometidas a un aumento gradual de la presión ambiental.

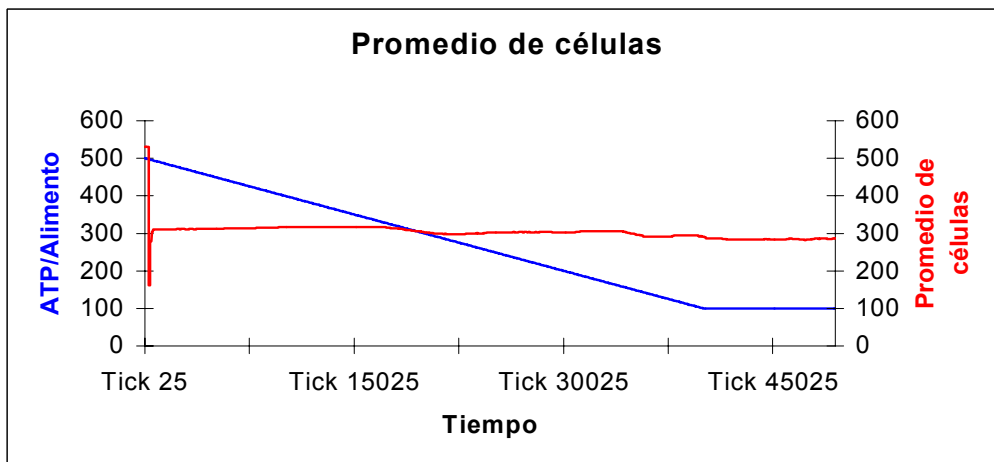
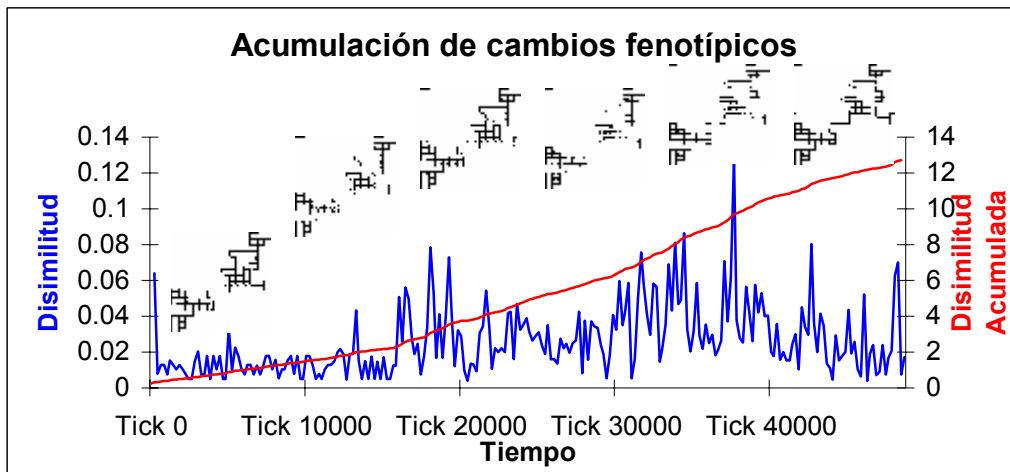
Se realizaron cinco repeticiones con este modelo poblacional, cada simulación se inició con la generación de un individuo de genoma aleatorio y se dejó correr durante 50,000 *Ticks*, las condiciones ambientales en este modelo fueron ajustadas al inicio como lo indica la Tabla 2 y durante el tiempo de la simulación se disminuyó gradualmente la cantidad de *ATP* contenida en el alimento a una tasa de $1ATP/100\ Ticks$, aumentando así gradualmente el número de comidas necesarias para no morir y reproducirse.

La disminución de la cantidad de energía se llevó a cabo gradualmente en este modelo, se comenzó con un valor de $500ATP/Alimento$ y se disminuyó una unidad cada 100 *Ticks* de la simulación hasta llegar al *Tick* 40,000 con un contenido energético de $100ATP/Alimento$. A partir de este punto no se disminuyó más la energía contenida en el alimento, debido a que la mayoría de los fenotipos no son capaces de sobrevivir con menos de esta cantidad de energía, la siguiente gráfica muestra el cambio en la cantidad de energía a través de la historia de la simulación.

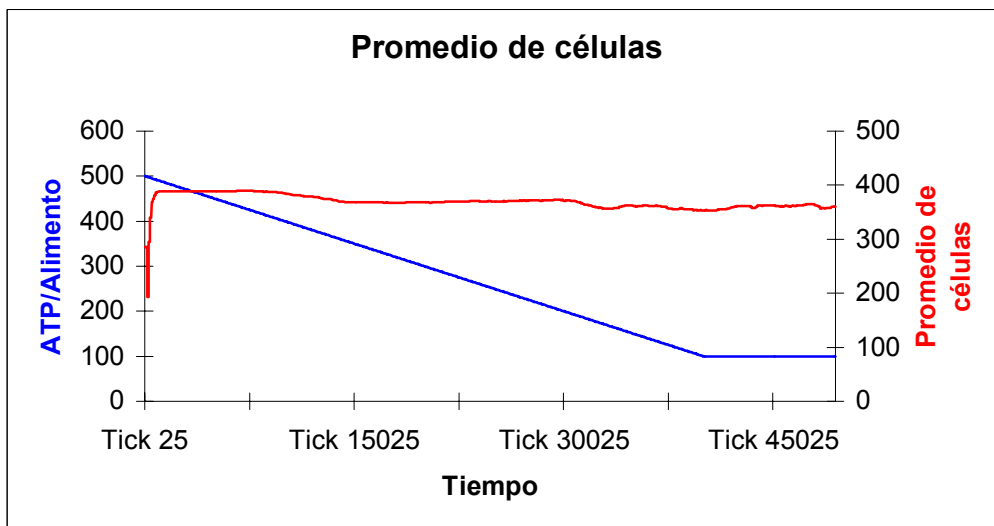
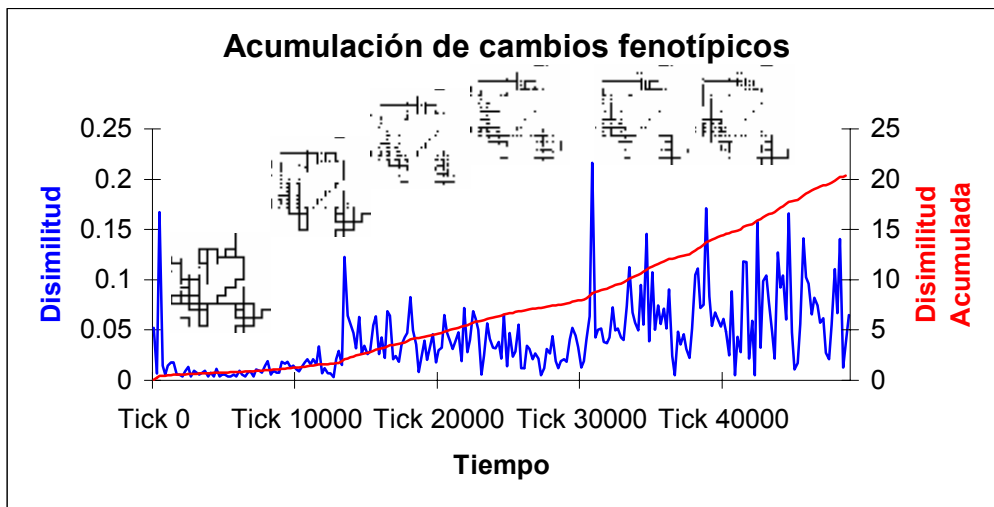


	<i>Disimilitud Máxima</i>	<i>Disimilitud Acumulada Máxima</i>	<i>Tasa de Evolución</i>
Repetición 1	12.834	12.708	0.0287
Repetición 2	21.6	20.352	0.0366
Repetición 3	22.869	6.341	0.0299
Repetición 4	14.884	9.138	0.0247
Repetición 5	24.352	24.336	0.0344

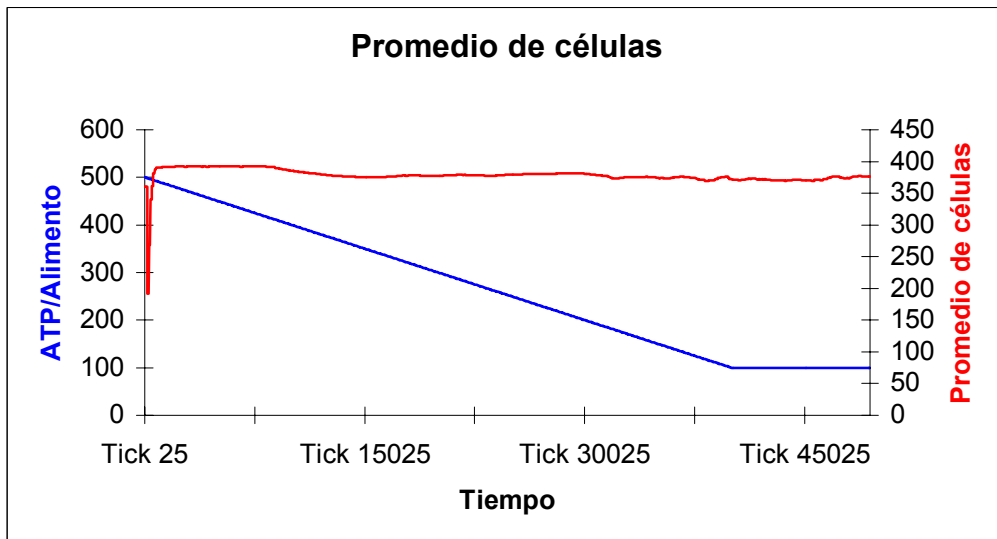
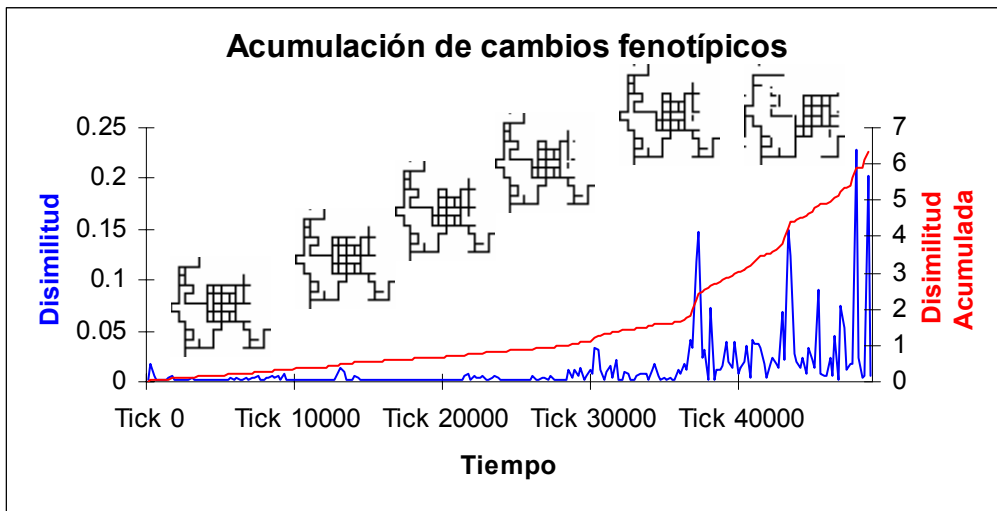
Repetición 1.



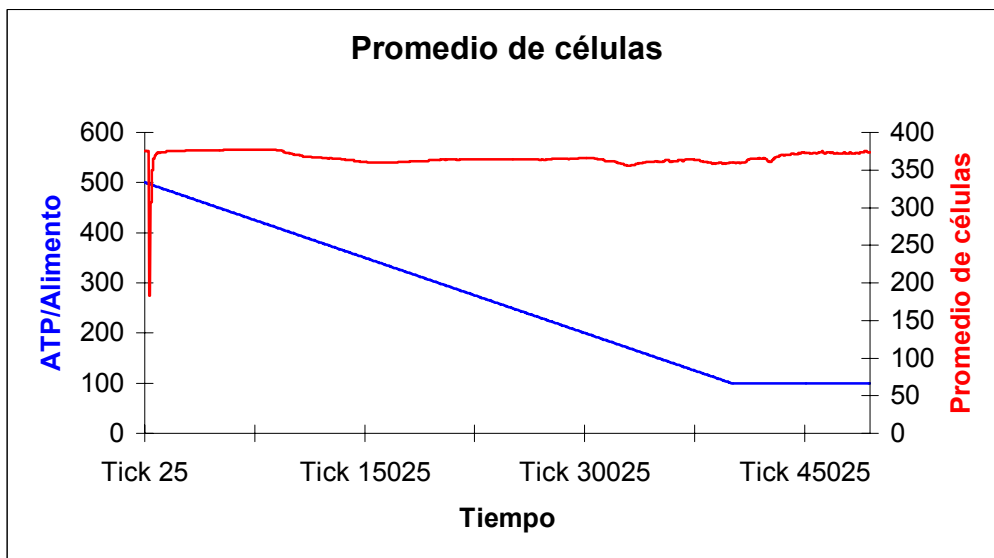
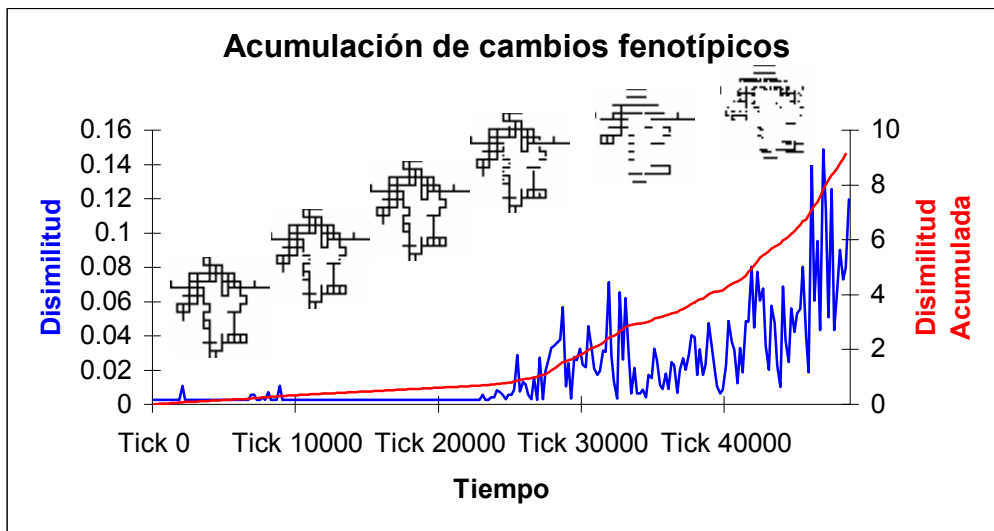
Repetición 2.



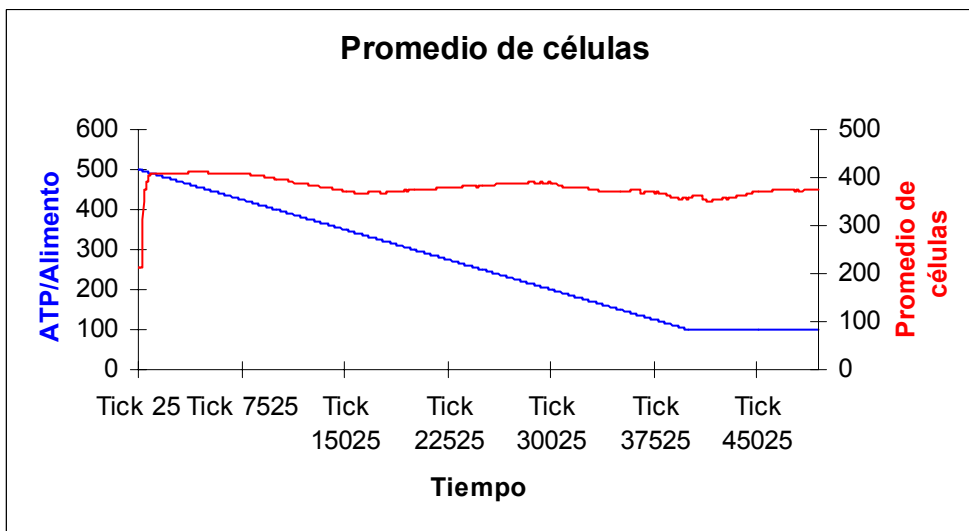
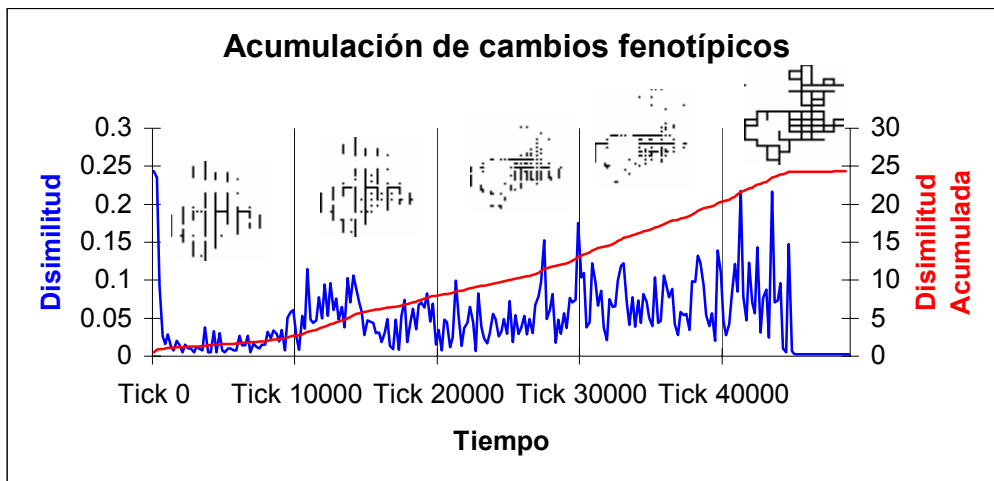
Repetición 3.



Repetición 4.



Repetición 5.



Modelo 4.

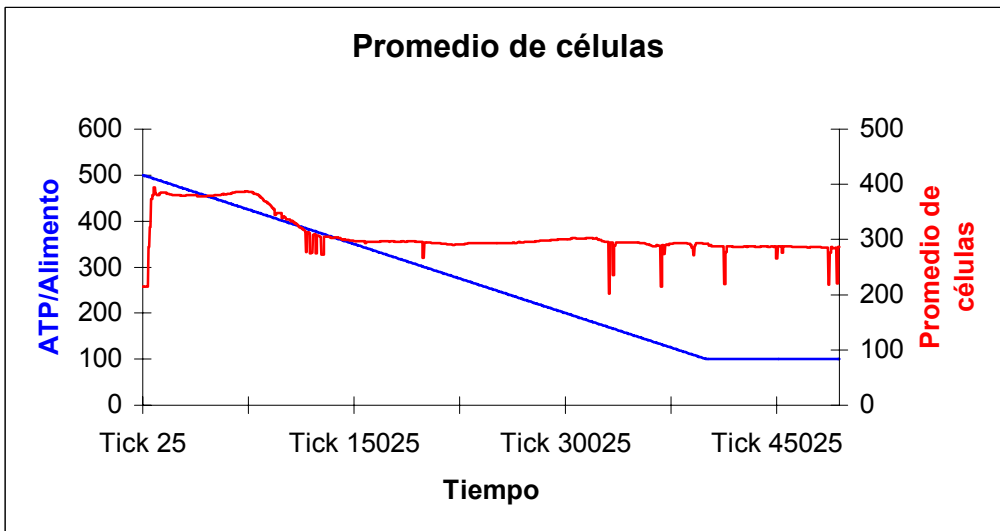
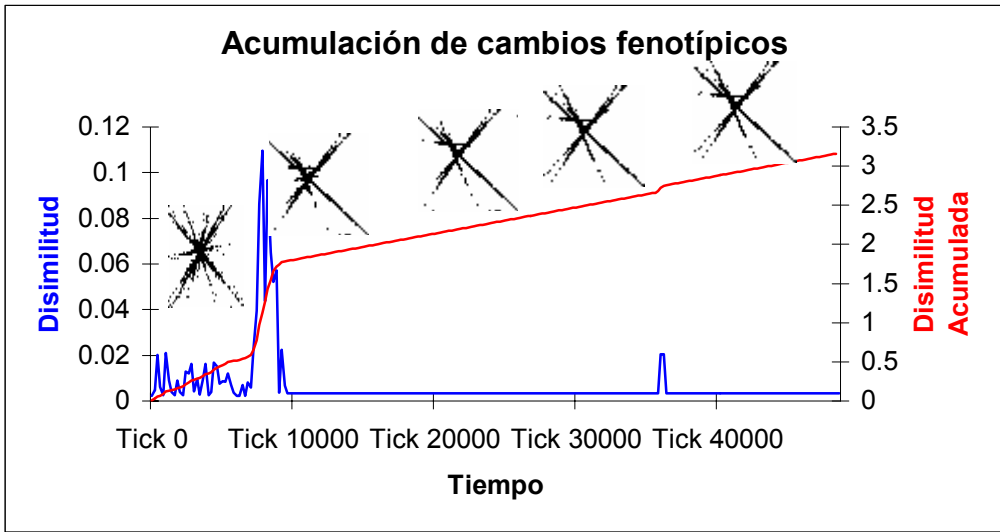
Poblaciones con expresión de información genética *IFS* sometidas a un aumento gradual de la presión ambiental.

Se realizaron cinco repeticiones con este modelo poblacional, cada simulación se inició con la generación de un individuo de genoma aleatorio y se dejó correr durante 50,000 *Ticks*, las condiciones ambientales en este modelo fueron ajustadas al inicio como lo indica la Tabla 2 y durante el tiempo de la simulación se disminuyó gradualmente la cantidad de *ATP* contenida en el alimento a una tasa de 1*ATP*/100 *Ticks*.

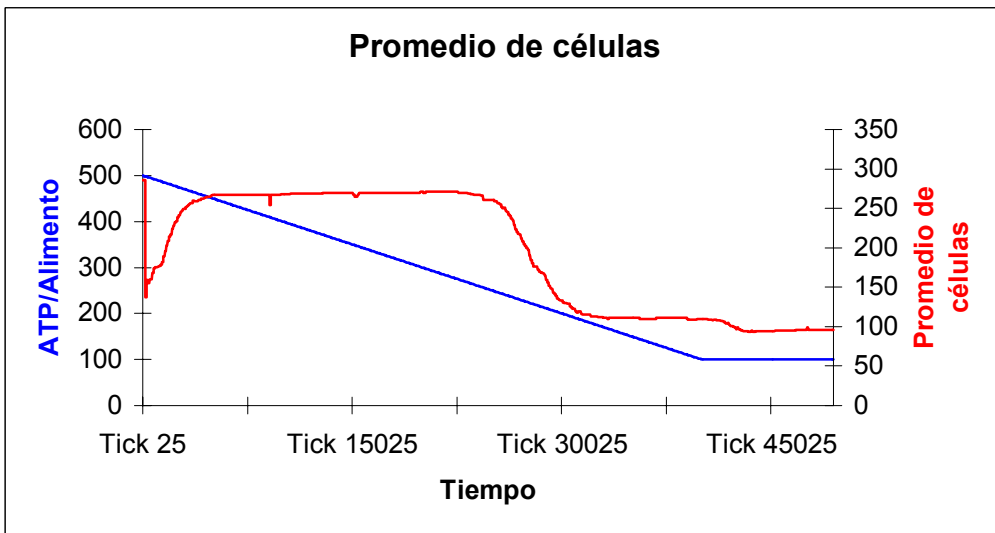
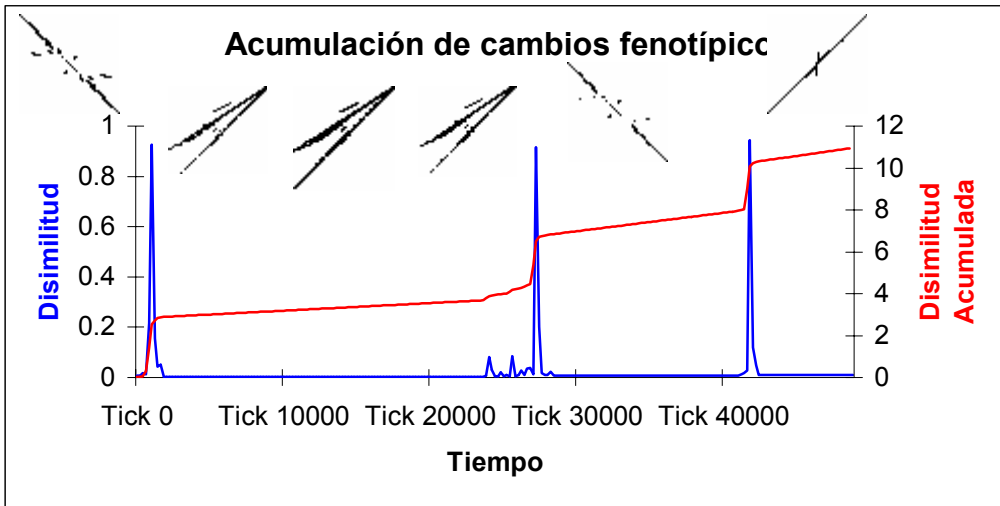
La disminución de la cantidad de energía se llevó a cabo gradualmente en este modelo, se comenzó con un valor de 500*ATP*/Alimento y se disminuyó una unidad cada 100 *Ticks* de la simulación hasta llegar al *Tick* 40,000 con un contenido energético de 100*ATP*/Alimento.

	<i>Disimilitud Máxima</i>	<i>Disimilitud Acumulada Máxima</i>	<i>Tasa de Evolución</i>
Repeticón 1	10.965	3.159	0.0243
Repeticón 2	94.475	10.943	0.0481
Repeticón 3	86.046	15.56	0.0464
Repeticón 4	25.263	5.883	0.0206
Repeticón 5	91.071	3.6853	0.04576

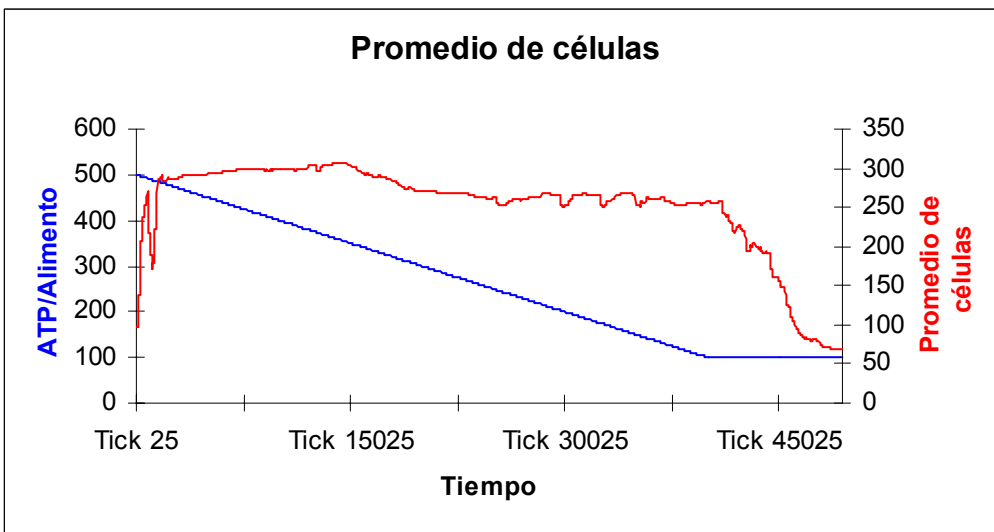
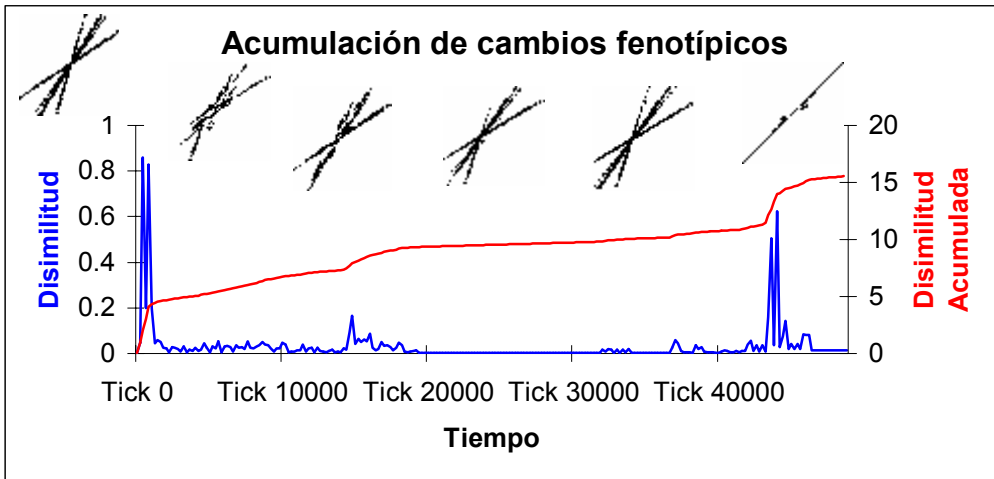
Repetición 1.



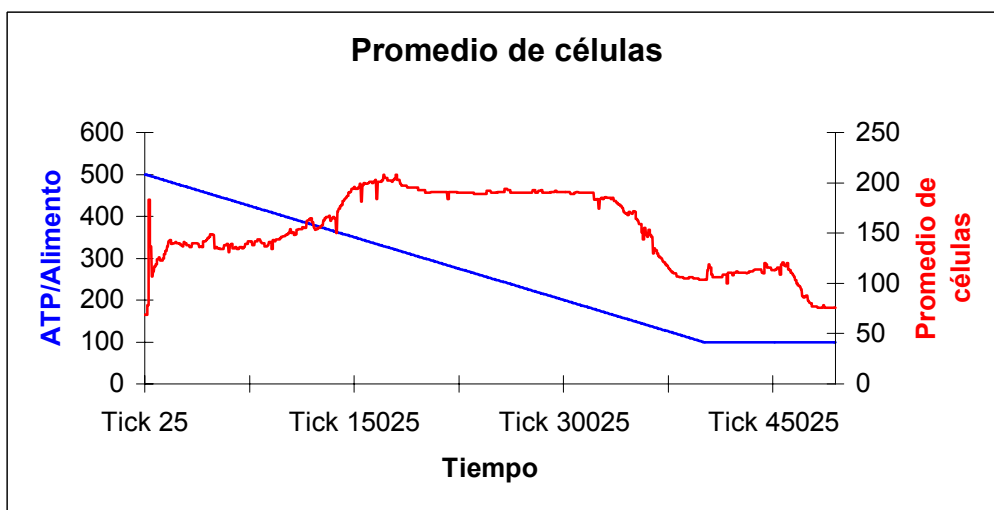
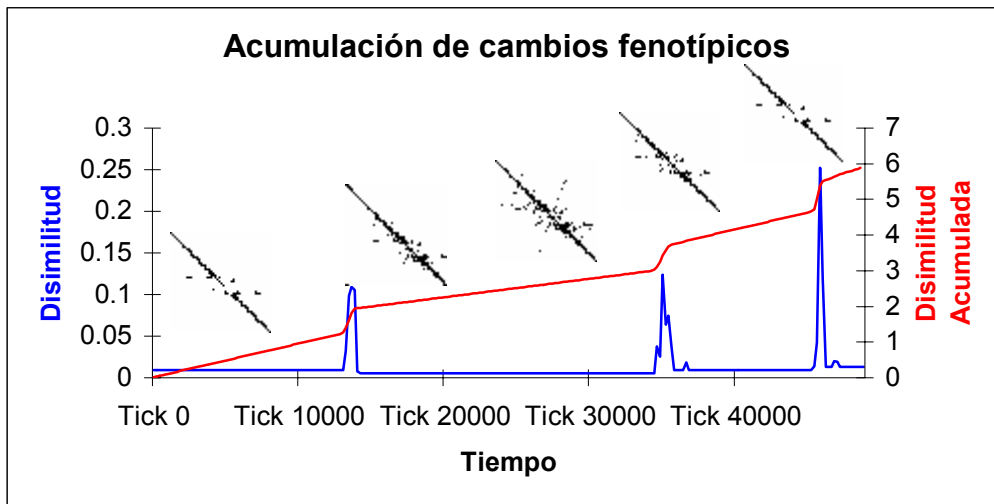
Repetición 2.



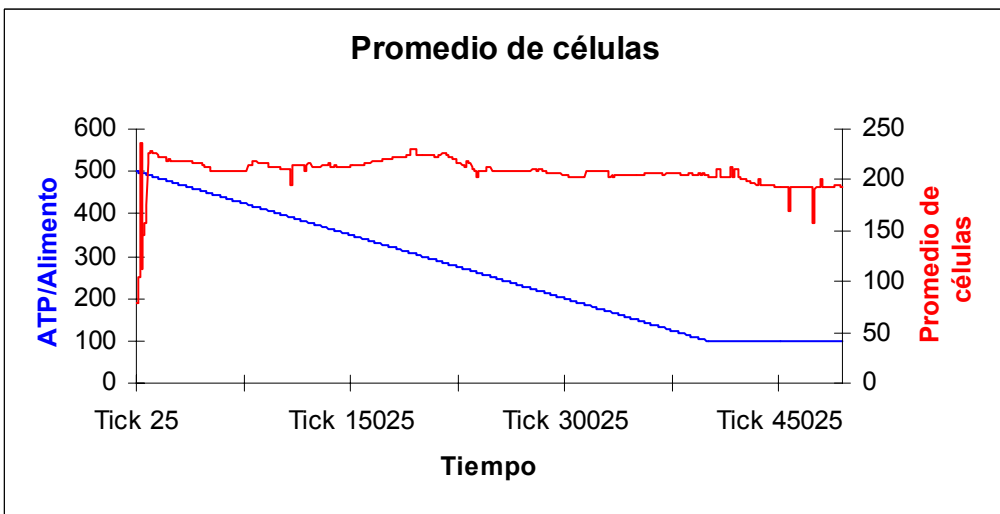
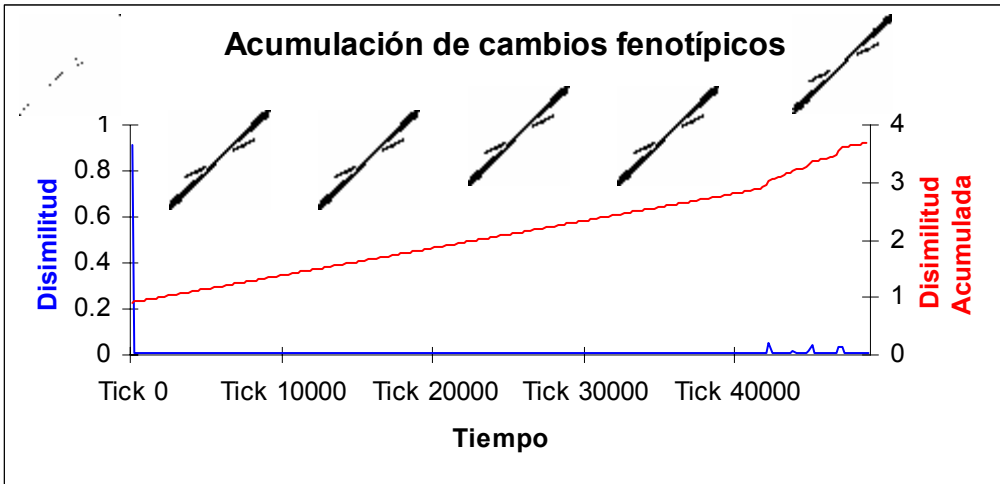
Repetición 3.



Repetición 4.



Repetición 5.



Modelo 5.

Poblaciones con expresión de información genética simple sometidas a aumentos drásticos de presión ambiental.

Se realizaron cinco repeticiones con este modelo poblacional, cada simulación se inició con la generación de un individuo de genoma aleatorio y se dejó correr durante 50,000 *Ticks*, las condiciones ambientales en este modelo fueron ajustadas al inicio como lo indica la Tabla 2 y posteriormente se disminuyó gradualmente la cantidad de *ATP* contenida en el alimento a una tasa de 100*ATP*/10,000 *Ticks*, aumentando así gradualmente el número de comidas necesarias para no morir y reproducirse.

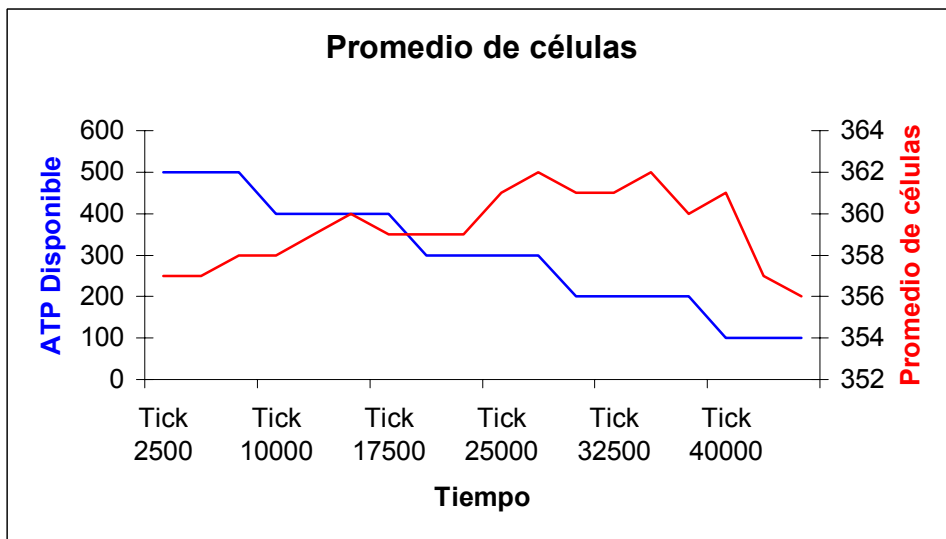
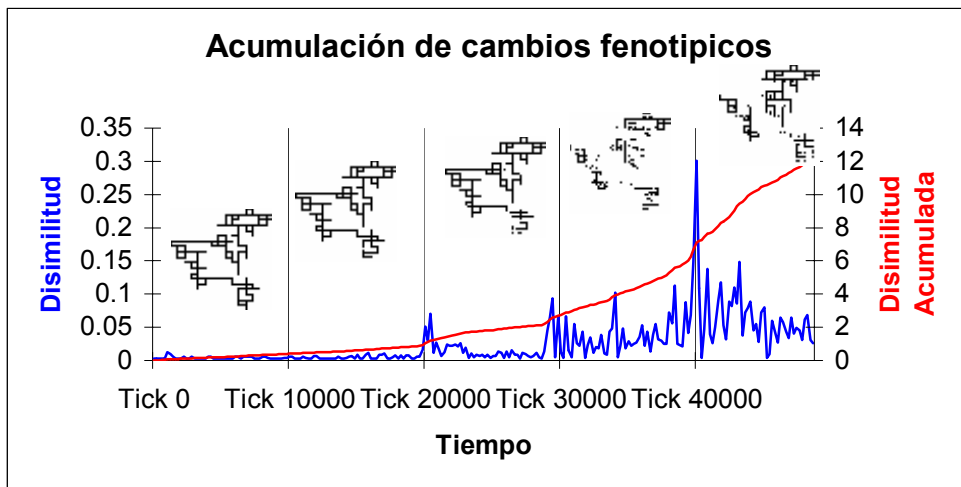
La disminución de la cantidad de energía se llevó acabo mediante largos periodos sin cambios seguidos de una disminución más o menos drástica de la cantidad de *ATP* contenida en las partículas de alimento, la siguiente tabla muestra el cambio en la cantidad de energía a través de la historia de la simulación.

Tiempo	<i>ATP</i>
<i>Tick</i> 0	500
<i>Tick</i> 10,000	400
<i>Tick</i> 20,000	300
<i>Tick</i> 30,000	200
<i>Tick</i> 40,000	100

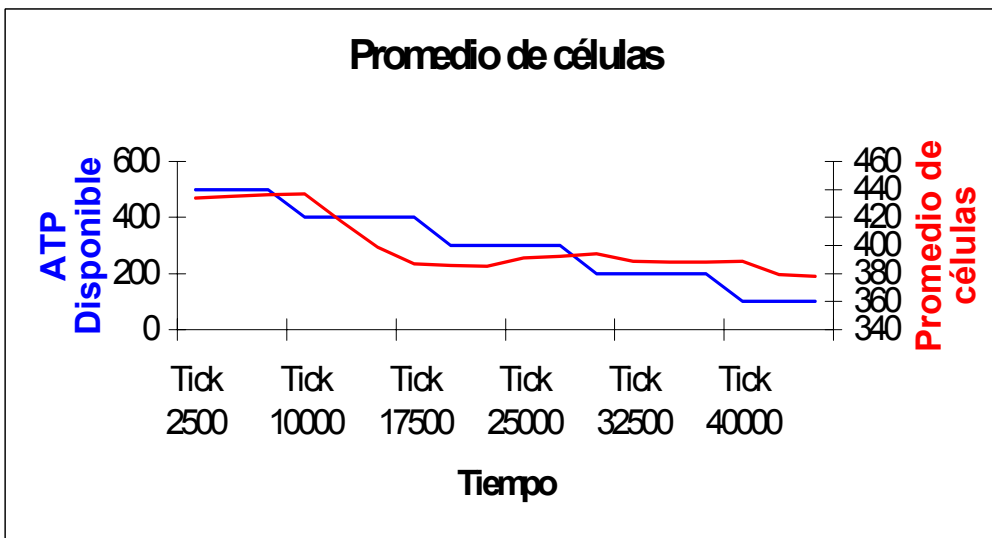
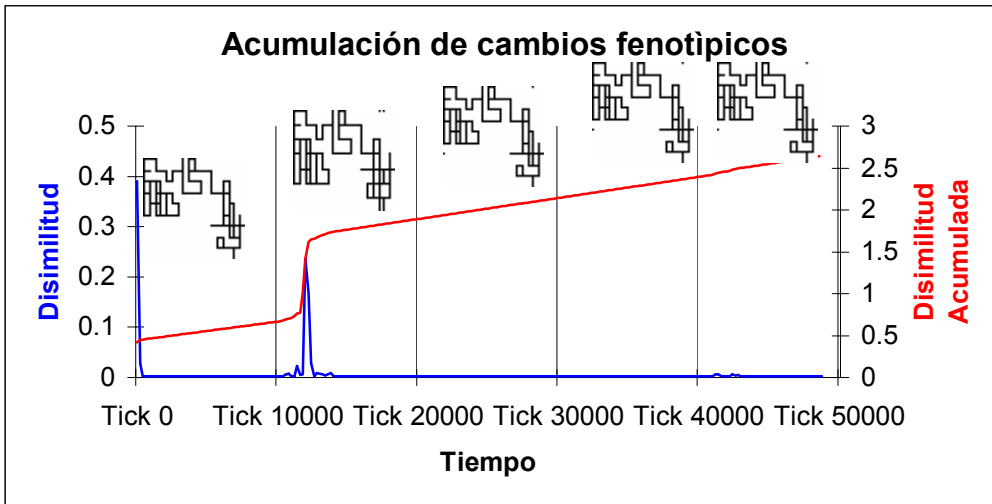
Tabla 4. Disminución de la cantidad de *ATP* en el alimento.

	<i>Disimilitud Máxima</i>	<i>Disimilitud Acumulada Máxima</i>	Tasa de Evolución
Repetición 1	30.097	12.11	0.033593161
Repetición 2	39.12	2.64	0.000126897
Repetición 3	40.932	15.567	0.038200184
Repetición 4	46.044	12.23	0.036768787
Repetición 5	15.83	11.57	0.031945471

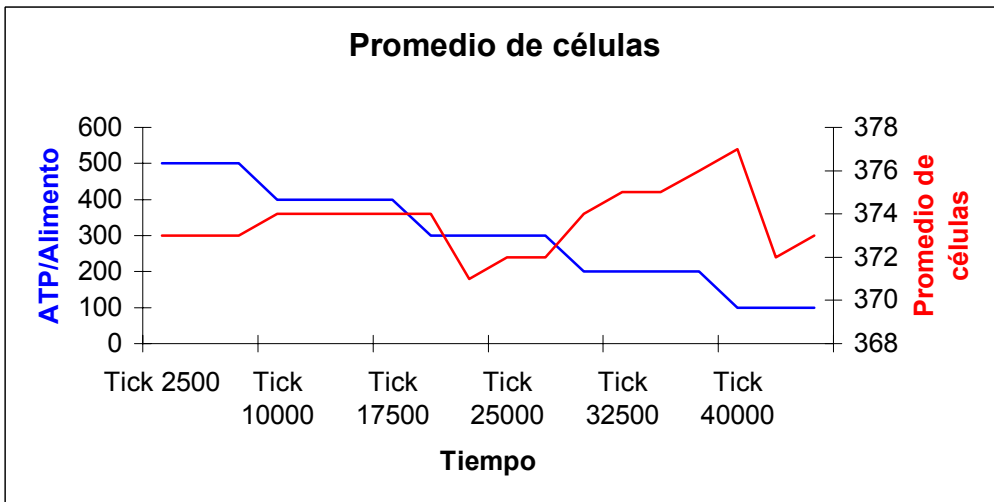
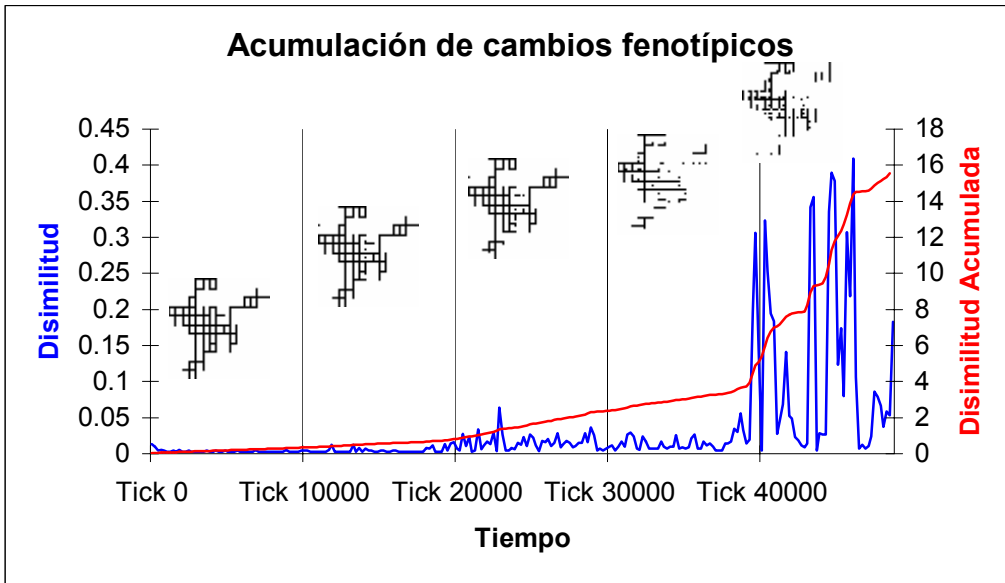
Repetición 1.



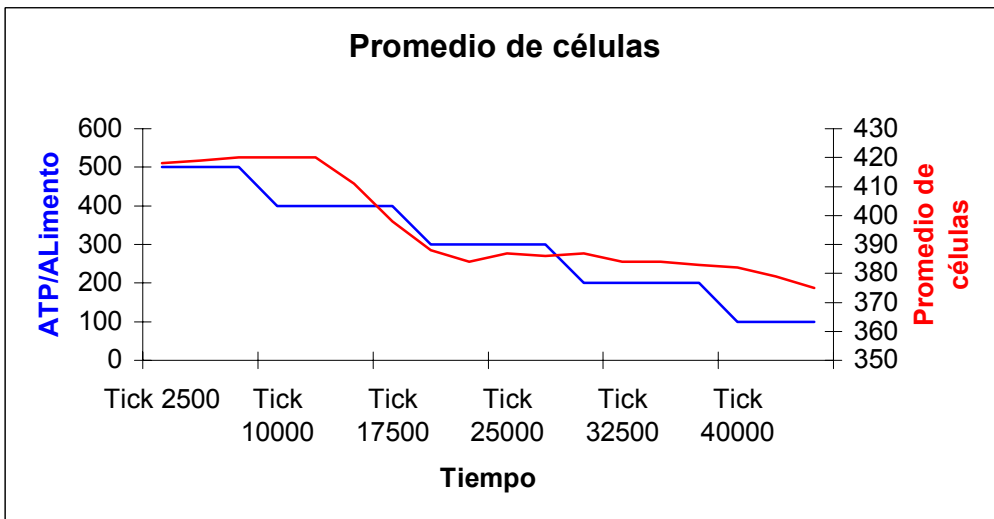
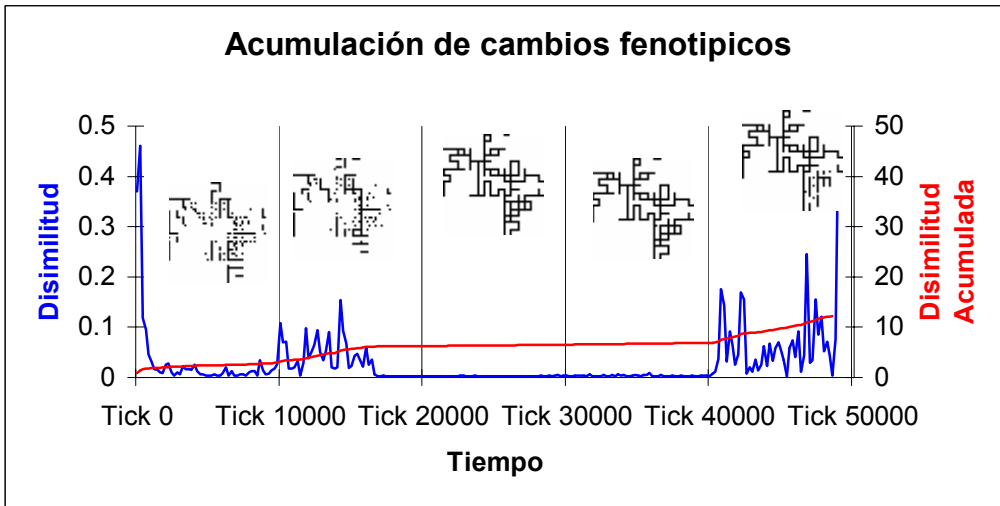
Repetición 2.



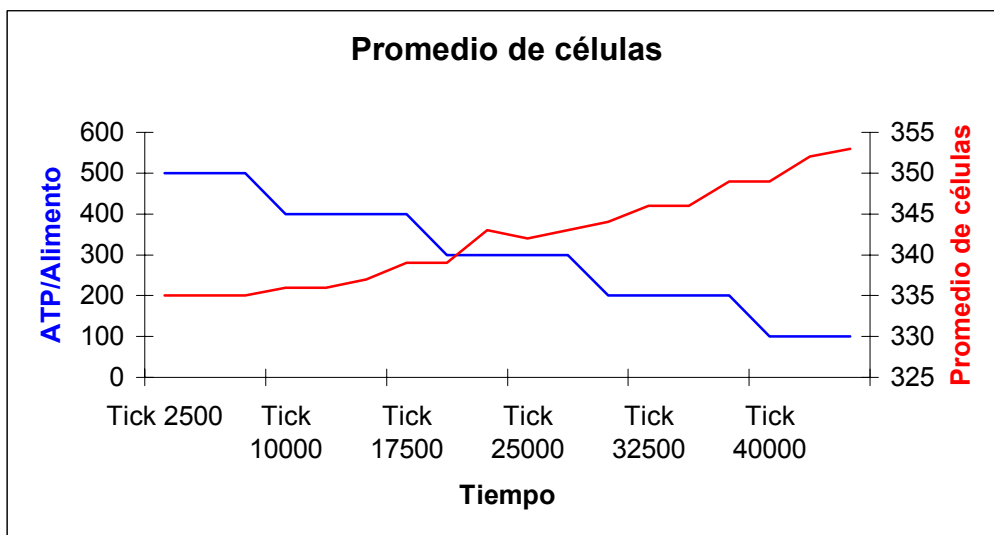
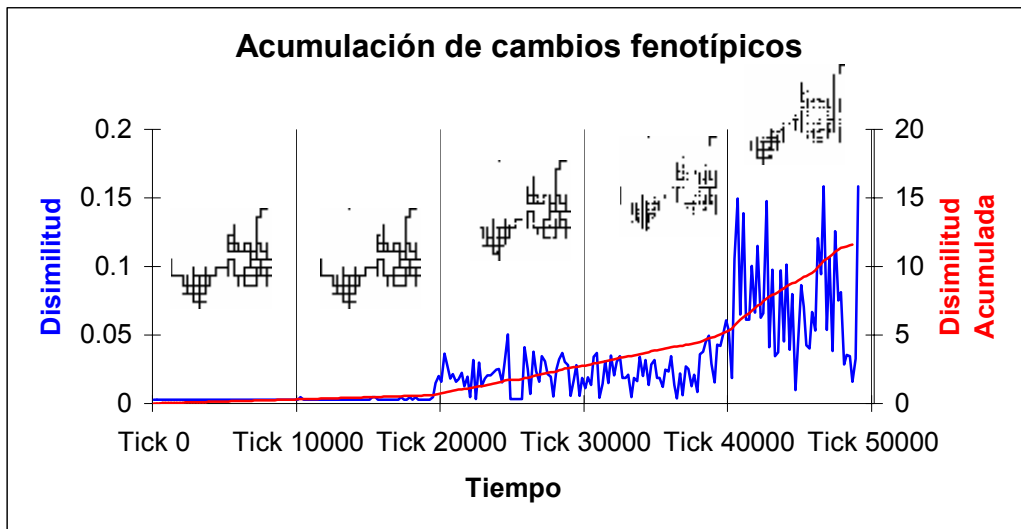
Repetición 3.



Repetición 4.



Repetición 5.

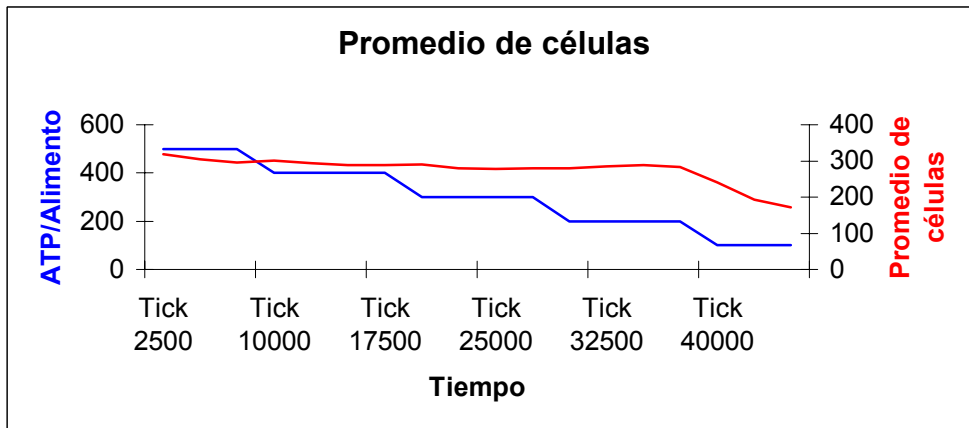
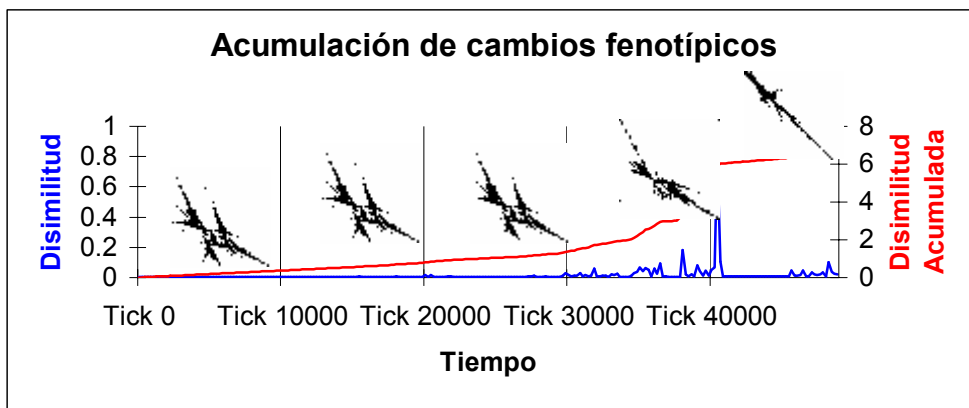


Modelo 6.

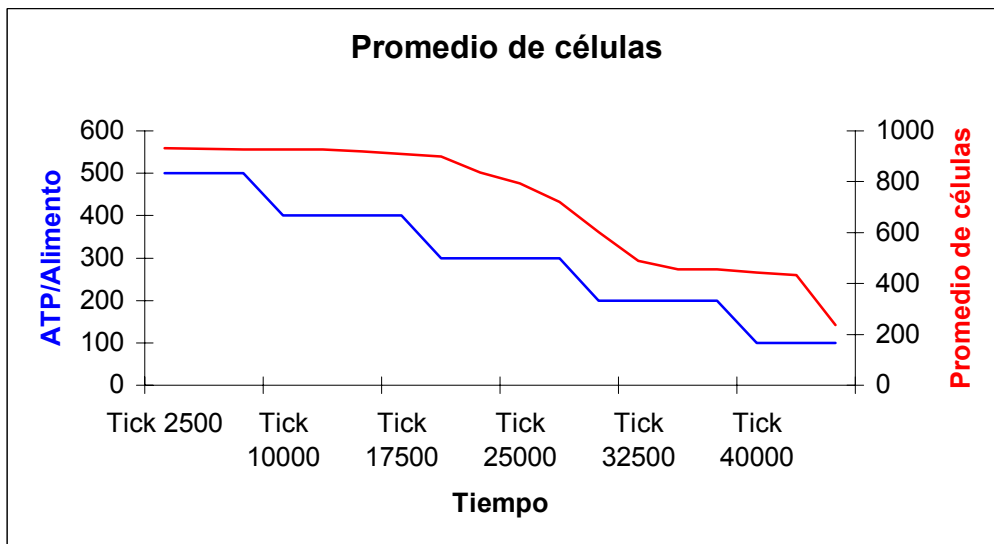
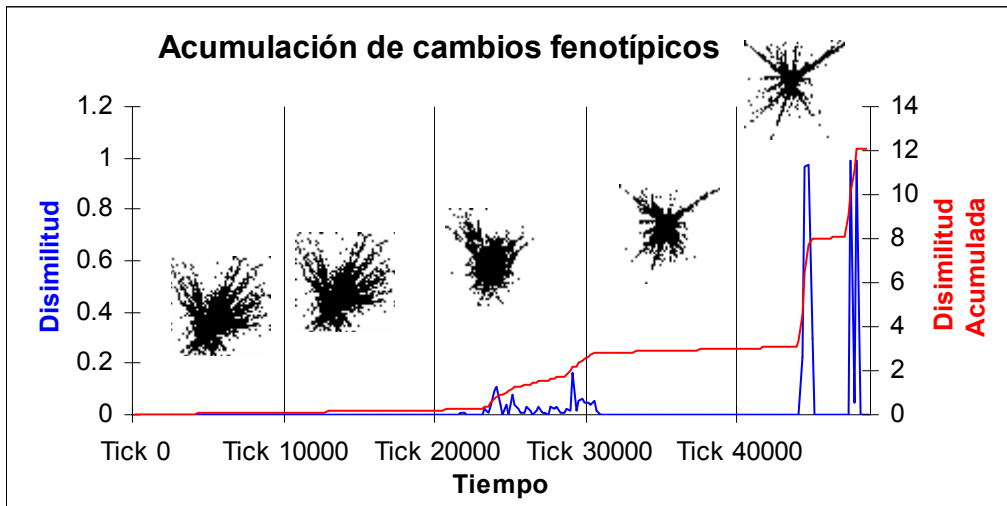
Poblaciones con expresión de información genética *IFS* sometidas a aumentos drásticos de presión ambiental.

	<i>Disimilitud Máxima</i>	<i>Disimilitud Acumulada Máxima</i>	<i>Tasa de Evolución</i>
Repetición 1	82.75	7.252	0.048448865
Repetición 2	99.17	12.136	0.032797815
Repetición 3	81.81	18.116	0.050505051
Repetición 4	15.33	4.307	0.033670034
Repetición 5	26.712	11.84	0.031631191

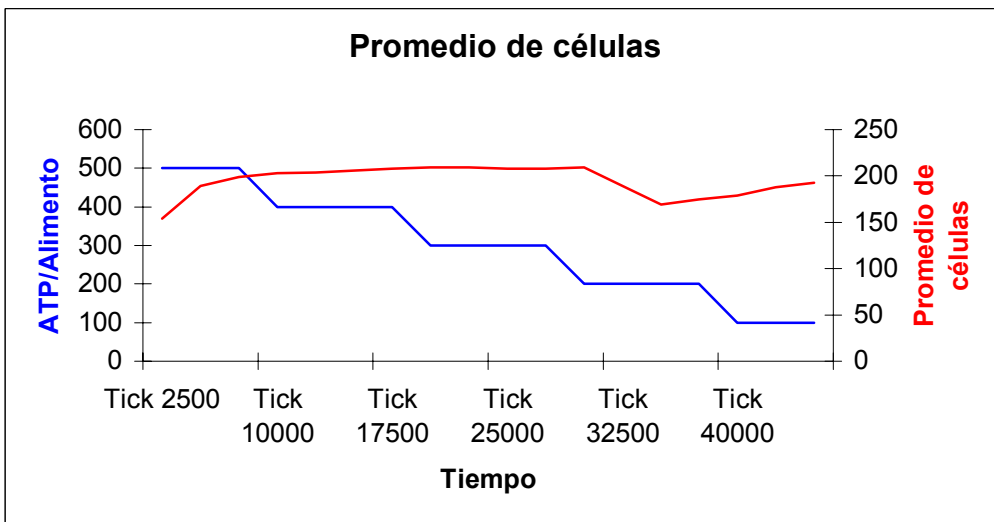
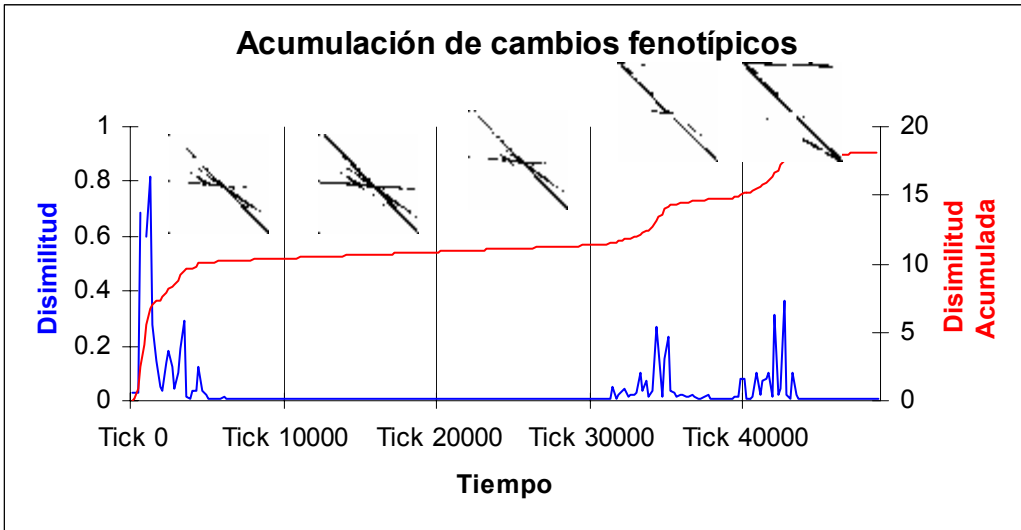
Repetición 1.



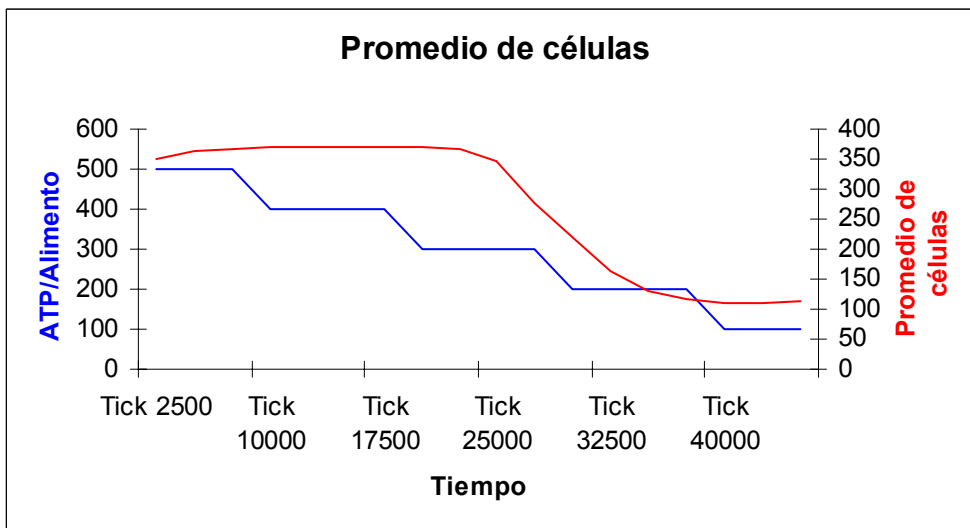
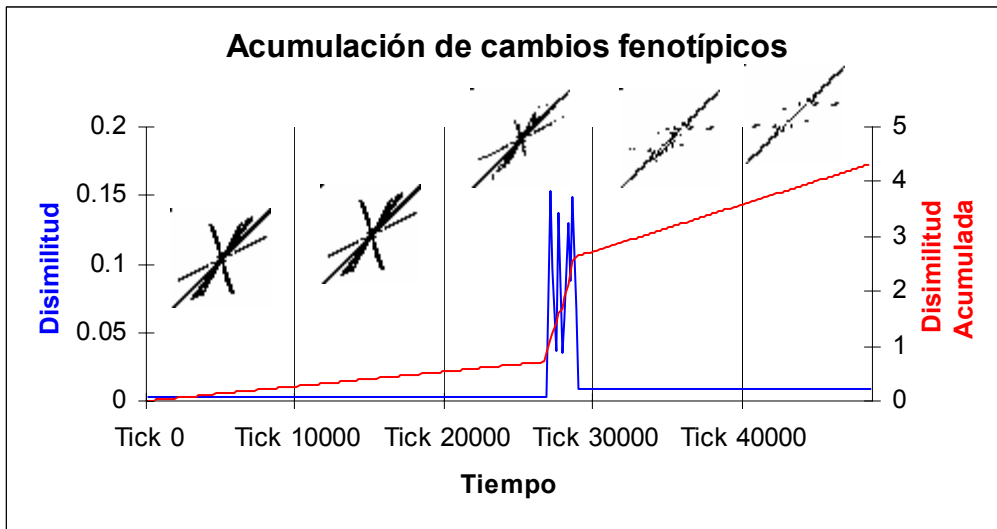
Repetición 2.



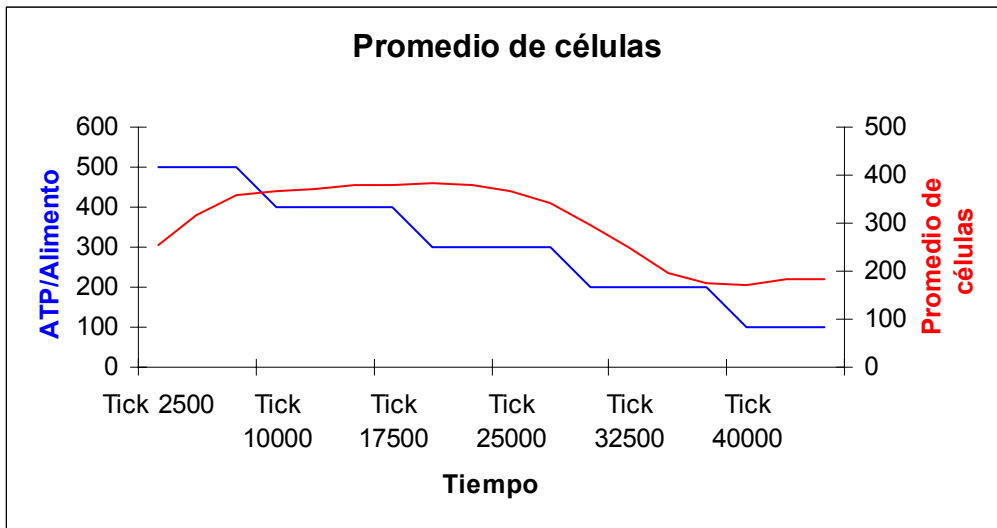
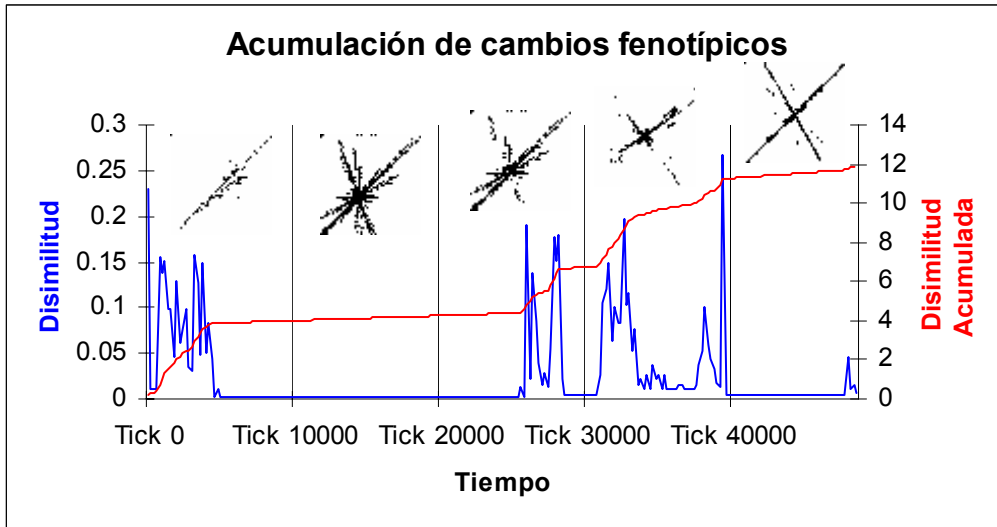
Repetición 3.



Repetición 4.



Repetición 5.



Comparativo entre los diferentes modelos

Se encontró una tendencia general en los modelos que incluyen una expresión genética *IFS* a tener valores altos de *Disimilitud Máxima*, mientras que los modelos de expresión simple no presentaron grandes saltos.

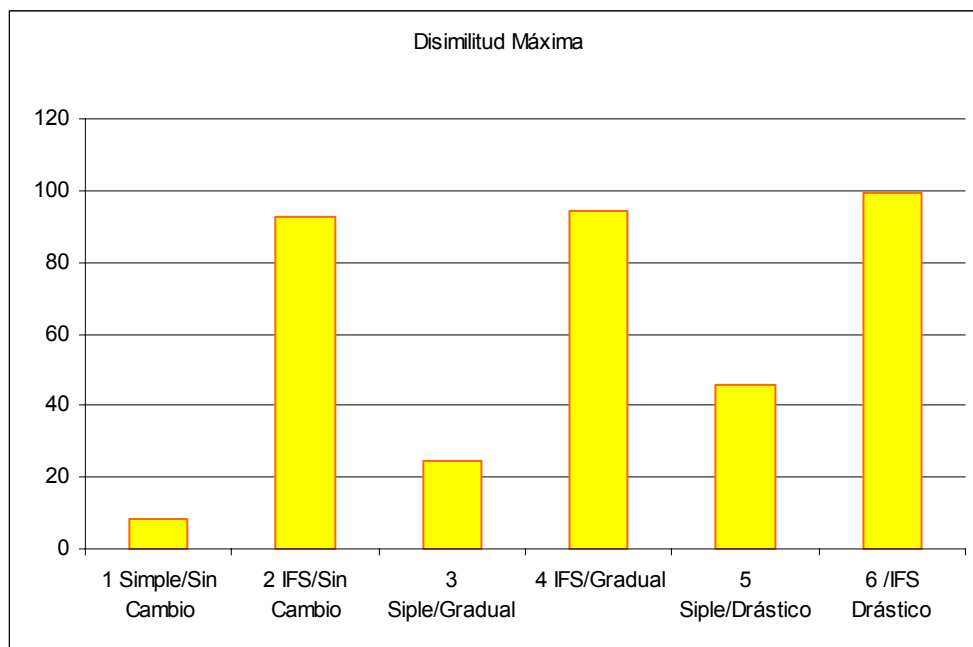


Figura 1. *Disimilitud Máxima* en los diferentes modelos.

Por otra parte, la *Disimilitud Acumulada Máxima* alcanzo valores ligeramente mas altos en los modelos dos y tres, siendo el modelo uno, el que acumulo menos cambios.

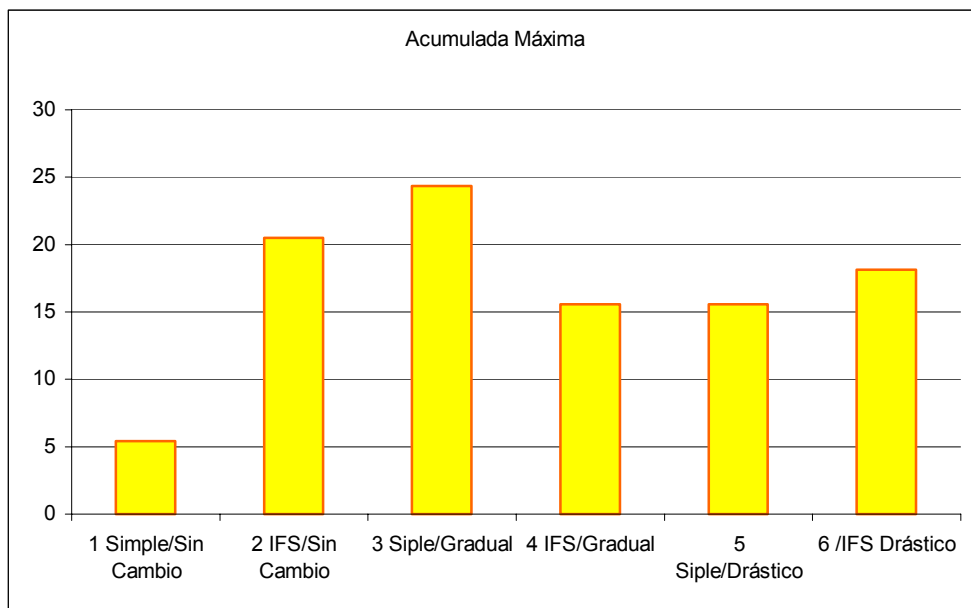


Figura 2. *Disimilitud Acumulada* máxima en los diferentes modelos.

Por ultimo, la *Tasa Evolutiva* alcanzo sus mayores valores en los modelos con expresión genética *IFS*

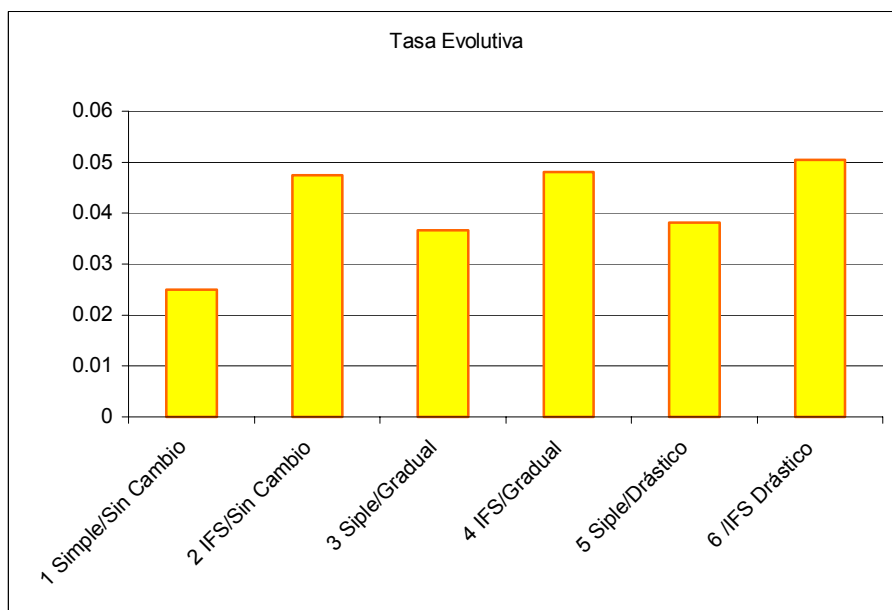


Figura 3. *Tasa Evolutiva* máxima en los diferentes modelos.

Análisis de resultados

Modelo 1. Poblaciones con expresión de información genética simple no sometidas a presión ambiental.

En este modelo la *Disimilitud* máxima encontrada entre dos *Tipos* históricamente consecutivos fue de 8.481%, siendo éste el menor valor en esta categoría en comparación con las otras poblaciones, mientras que la *Disimilitud Acumulada* máxima alcanzó un total de 5.424 siendo este modelo el que menos variación fenotípica presentó. La tasa evolutiva máxima encontrada en las cinco repeticiones fue de 0.0249 *Cambios/Tick*, la cual es la más baja con respecto a los otros modelos.

Los patrones evolutivos en este modelo se caracterizan por poblaciones que varían muy poco a través del tiempo y acumulan gradualmente cambios con la tendencia general a aumentar el número promedio de células, las gráficas de *Disimilitud Acumulada* muestran que los cambios fenotípicos se acumulan de manera gradual durante la historia de la simulación sin presentar discontinuidades drásticas, es decir, el cambio evolutivo se da mediante la acumulación de pequeñas variaciones y en apariencia está dirigido por la selección natural, ya que hay una tendencia general en todas las repeticiones realizadas a que las poblaciones aumenten su número promedio de células, lo cual sugiere que la aparición de una mutación que genere un individuo con un número ligeramente mayor de células y que aumente el área de captura de alimento de este, le dará mayores probabilidades de reproducirse en un ambiente como este.

Modelo 2. Poblaciones con expresión de información genética *IFS* no sometidas a presión ambiental.

La mayor *Disimilitud* entre *Tipos* consecutivos en las repeticiones realizadas fue de 92.59%, este valor representa un cambio en el que sólo permaneció el 7.4% de coincidencias fenotípicas entre dos *Tipos* separados por un lapso de 200 *Ticks*, lo cual indica una discontinuidad en el registro fósil, la máxima *Disimilitud Acumulada* registrada en este modelo fue de 20.49, siendo uno de los modelos con mayor acumulación de cambios. La mayor tasa de evolución

calculada para este modelo fue de 0.0474 *Cambio/Tick*, siendo una de las mayores encontradas entre los diferentes modelos.

Las gráficas de *Disimilitud Acumulada* muestran saltos abruptos en los que el fenotipo promedio cambia de manera drástica en un lapso reducido de tiempo, se puede observar en las gráficas que estas discontinuidades son provocadas por cambios bruscos que en este modelo están acompañados, por lo general, de un aumento repentino del promedio de células en las poblaciones. Este comportamiento indica la aparición de mutantes fenotípicamente muy diferentes al organismo que les dio origen, lo cual es atribuible al mecanismo de expresión genética *IFS*. Cuando aparece un individuo con un mayor número de células distribuidas de tal manera que le permiten capturar más alimento en un ambiente como este (alto contenido de energía en el alimento), tiene una mayor probabilidad de reproducirse y eventualmente puede llegar a ser más abundante y sustituir a otros fenotipos, siendo de esta manera que este tipo de poblaciones evolucionan mediante selección natural.

Modelo 3. Poblaciones con expresión de información genética simple sometidas a un aumento gradual de la presión ambiental.

La *Disimilitud* máxima encontrada entre dos *Tipos* históricamente consecutivos fue de 24.352%, valor que representa un cambio moderado en comparación con los otros modelos, por otra parte, la *Disimilitud Acumulada* máxima fue de 24.336, siendo este el valor más alto encontrado en todas las simulaciones realizadas, lo cual representa que este modelo acumula la mayor cantidad de cambios, aunque en su mayoría son pequeños. La tasa evolutiva más alta encontrada en este modelo fue de 0.0366 *Cambios/Tick*, valor muy similar al del Modelo 5, en el cual se utiliza el mismo tipo de expresión genética y hay aumento de presión ambiental.

Las gráficas de *Disimilitud Acumulada* en este modelo muestran cambios en el ángulo de la pendiente, lo cual indica una aceleración de la acumulación de cambios fenotípicos del *Tipo*, esto parece indicar la sustitución gradual de un *Tipo* por otro en función de la selección natural, el número de células promedio de las poblaciones en este modelo no parece seguir ningún patrón de aumento o disminución.

Modelo 4. Poblaciones con expresión de información genética *IFS* sometidas a un aumento gradual de la presión ambiental.

La mayor *Disimilitud* entre *Tipos* históricamente consecutivos encontrada para este modelo fue de 94.475%, este valor representa una gran discontinuidad en el registro fósil ya que indica un cambio drástico en el fenotipo promedio de la población en un lapso muy pequeño en comparación con el tiempo que dura la simulación. La *Disimilitud Acumulada* máxima alcanzó un total de 15.56, este valor indica una acumulación moderada de cambios en comparación con otros modelos, sobre todo con aquellos de expresión genética simple, los cuales tienden a acumular cambios muy pequeños pero de manera constante. La mayor tasa de evolución calculada para este modelo fue de 0.0481 *Cambios/Tick*, siendo la segunda más alta encontrada en los diferentes modelos.

El comportamiento evolutivo de este modelo se caracteriza por poblaciones que permanecen estables con un fenotipo durante largos periodos de tiempo después de los cuales se encuentran picos en el valor de *Disimilitud* que pueden llegar a ser de más de 90%, es decir, en un lapso pequeño de tiempo es posible que el fenotipo promedio de la población cambie en una proporción en la que sólo conserve 10% de sus características. El promedio de células en la población tiende a caer después de un pico en la *Disimilitud*, esto indica que un fenotipo poco frecuente en la población, con una menor cantidad de células que el promedio, repentinamente se vuelve más abundante como resultado de la selección natural debido a que en este modelo la cantidad de energía va disminuyendo continuamente aquellos fenotipos con un menor requerimiento energético (menos células) pueden tener más oportunidades de reproducirse y eventualmente sustituir por completo a otro fenotipo.

Modelo 5. Poblaciones con expresión de información genética simple sometidas a aumentos drásticos de presión ambiental.

Para este modelo la máxima *Disimilitud* encontrada entre dos *Tipos* fue de 46.044, el mayor valor para una población con expresión genética simple y representa una discontinuidad moderada en el registro fósil en un intervalo de 200 *Ticks*. La acumulación de cambios en el fenotipo promedio fue representada con un valor de *Disimilitud Acumulada* máxima de 15.567, valor moderado en comparación con el modelo 3, en el cual se utiliza el mismo mecanismo de expresión genética pero con un aumento gradual de la presión

ambiental. La tasa evolutiva máxima encontrada en las cinco repeticiones fue de 0.0382 *Cambios/Tick*, la cual es muy similar a la del modelo 3.

En este modelo se encuentran pendientes irregulares en las gráficas de *Disimilitud Acumulada*, incluso se encuentran algunos escalones que indican una aceleración en el cambio del fenotipo promedio, estos escalones o cambios de pendiente indican la sustitución de un fenotipo por otro y se observan frecuentemente después de un cambio drástico en la presión ambiental

Modelo 6. Poblaciones con expresión de información genética *IFS* sometidas a aumentos drásticos de presión ambiental.

En este modelo se encontró una *Disimilitud* máxima de 99.17, la mayor de todos los modelos propuestos, este valor indica que un fenotipo fue sustituido por otro con el que sólo comparte menos del 1% de sus características en un lapso de 200 *Ticks*, este fenómeno de sustitución drástica de fenotipos se puede explicar por la muerte repentina de la mayoría de los individuos con el fenotipo dominante como resultado del aumento de la presión ambiental y el aumento de la frecuencia del nuevo fenotipo debido a la disminución de la competencia por alimento. Aun que la *Disimilitud* máxima se encontró en este modelo, la *Disimilitud Acumulada* máxima no fue la mayor, ésta obtuvo un valor de 18.116, este valor en combinación con el de *Disimilitud* indican la presencia de pendientes muy suaves seguidas de cambios drásticos. La tasa de evolución fue la mayor encontrada en todas las simulaciones realizadas, alcanzando un valor de 0.0505.

Los patrones evolutivos en este modelo siguen tendencias parecidas a las del modelo 4, periodos de estabilidad seguidos de cambios repentinos de las frecuencias fenotípicas como consecuencia de la muerte o disminución de la tasa de reproducción de un fenotipo y el aumento de la frecuencia de otro que posteriormente lo sustituirá, siendo la magnitud del cambio de pendiente en la gráfica de *Disimilitud Acumulada* el indicador de la rapidez con la que esto ocurre.

Conclusiones

Se compararon las tendencias evolutivas de los modelos de población incluidos en este trabajo tomando en cuenta los patrones de cambio fenotípico a través del tiempo presentes en los diferentes modelos, estos patrones se observaron en las gráficas de *Disimilitud Acumulada* y expresan la magnitud del cambio fenotípico con respecto al tiempo, entre mayor sea la magnitud del cambio evolutivo en un intervalo determinado de tiempo, mayor será la pendiente de la función en el segmento correspondiente a dicho intervalo. Al comparar las diferencias en las formas generales de las gráficas de *Disimilitud Acumulada* de las diferentes poblaciones fue posible establecer patrones evolutivos graduales o discontinuos para los diferentes modelos.

En los modelos con expresión de información genética simple se observa un cambio gradual y muy pequeño durante toda la simulación, se caracterizan por tener pocos y generalmente pequeños aumentos del ritmo de evolución, en ocasiones se ve reflejada la acción de la selección natural en el número promedio de células por individuo, dependiendo de la morfología del primer individuo y el tipo de presión ambiental al que se someta la población, es posible observar pequeños ascensos o descensos del promedio de células como una respuesta adaptativa, este tipo de evolución gradual concuerda con lo propuesto por la TSE, la cual establece que el cambio evolutivo se produce de manera gradual por la acumulación de pequeños cambios en las frecuencias fenotípicas de las poblaciones dirigidos por la selección natural.

Por otra parte, los modelos con mecanismos de expresión genética *IFS* muestran periodos con pendientes poco pronunciadas que indican la ausencia de cambios fenotípicos, seguidos de pendientes que pueden llegar casi hasta 90° que indican un cambio drástico del fenotipo promedio, como resultado de discontinuidades en el registro fósil. Se observa también que el cambio evolutivo está dirigido por la selección natural, ya que el promedio de células de las poblaciones cambia marcadamente en función de la presión ambiental, es decir, en ambientes ricos en alimento el fenotipo tiende a ser más robusto, mientras que en ambientes pobres el número de células tiende a disminuir. La ausencia de cambios fenotípicos durante largos periodos de tiempo seguidos de cambios drásticos indican una tendencia evolutiva comparable con el modelo de macroevolución propuesto en la TEP.[6]

Se estableció que la magnitud con la que actúa la selección natural sobre una población depende de la presión ambiental ejercida sobre esta, se encontraron diferencias marcadas en la magnitud del cambio evolutivo entre los modelos sometidos a presión ambiental y los que no, siendo mayor este cambio en aquellos modelos en los que se incluyeron aumentos de la presión ambiental a través del tiempo. En un ejemplo biológico este fenómeno se ha observado en casos como el melanismo industrial de *Lepidópteros* en el que la presión ambiental aumenta de manera drástica al cambiar el color de los árboles y hacer más evidentes estos insectos para las aves que se alimentan de ellos, dando como resultado un cambio rápido de la población. [19]

Se observó la sustitución del fenotipo dominante, este fenómeno ocurre cuando existe un fenotipo predominante en la población y otro en una proporción menor, si el fenotipo que se encuentra con menor frecuencia en la población repentinamente es más favorecido por el ambiente, comenzara a reproducirse más rápidamente y aumenta su número consumiendo el alimento lo que provoca la eventual extinción del otro fenotipo, este fenómeno se observó hasta tres veces en una misma repetición en las poblaciones con expresión genética *IFS*.

En comparación con los modelos de expresión genética simple en los que cada gen esta “mapeado” directamente con una sola característica fenotípica, en el modelo de expresión *IFS* existe una expansión de información y no se puede hablar de un mapeo uno a uno entre el genotipo y el fenotipo, esta relación no lineal se encuentra presente en los sistemas biológicos debido a que las secuencias de nucleótidos en el material genético producen secuencias de aminoácidos, pero no contienen las instrucciones para arregarlos tridimensionalmente para conformar las proteínas y mucho menos para organizarlas y generar sistemas vivos, esta aparente ganancia de información se da durante el desarrollo mediante la aplicación de leyes fisicoquímicas. El mecanismo de expresión genética *IFS* fue implementado con la finalidad de simular de una manera generalizada el efecto de la morfogénesis sobre el proceso evolutivo.

El efecto que este modelo de desarrollo ontogénico expansor de información tuvo sobre las poblaciones en las que se implementó fue la aparición de discontinuidades en el gradualismo del proceso evolutivo, esto se debe a que el mecanismo que regula la generación del fenotipo a partir del genotipo es

susceptible a entregar variaciones fenotípicas importantes a partir de errores en la copia del material genético como consecuencia de la expansión de información ocurrida durante la ontogenia, es decir, si durante la reproducción de un individuo ocurre una mutación en el genoma del descendiente, este puede ser morfológicamente muy diferente al organismo que le dio origen. La aparición de este tipo de mutantes tiene como consecuencia una mayor probabilidad de encontrar individuos con derivaciones morfológicas en la población y esta diversidad acentúa el efecto de la selección natural siendo posible observar en un intervalo pequeño de tiempo el fenómeno de sustitución del fenotipo dominante.

Las observaciones realizadas en este trabajo son exclusivas para sistemas con reproducción asexual, sugieren que la posesión de un mecanismo de desarrollo tiene como consecuencia un tipo de evolución similar al propuesto en la TEP por Gould y Eldredge, al menos en los sistemas haploides. Por otra parte, se podría esperar que en los sistemas no haploides el efecto de las mutaciones se vea atenuado debido a la existencia de los alelos, que pueden ser vistos como “copias de seguridad” de un gen (no necesariamente son copias exactas pero sí son genes que ya existen en la población), impidiendo de esta manera una derivación morfológica drástica del fenotipo cuando ocurre una mutación en el genotipo como consecuencia de que la mutación estará presente solo en un alelo mientras que el otro se puede expresar de manera normal, disminuyendo así el efecto de las mutaciones, de ser cierto esto se puede explicar que los organismos de reproducción sexual evolucionen de manera gradual, debido a que la selección natural actúa sobre el fenotipo, el cual al ser derivado de un sistema que posee información genética por duplicado tendría menor probabilidad de expresar fenotípicamente las mutaciones ocurridas en el genoma y generar individuos con diferencias morfológicas importantes.

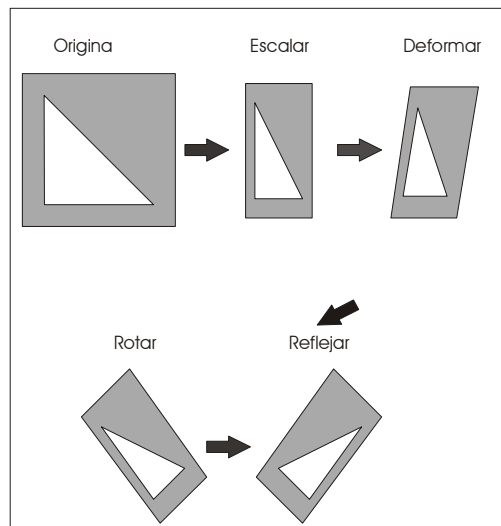
Algunas teorías de especiación sugieren que el aislamiento reproductivo de las poblaciones es un factor determinante para la generación de nuevas especies, como consecuencia de este aislamiento se puede generar un alto nivel de endogamia en la población, teniendo como resultado que la aparición de una mutación en un gen se pueda difundir en la población y eventualmente debido al alto grado de endogamia se pueden generar individuos con su par de alelos mutados y al no tener la “copia de seguridad” del gen original, el nuevo gen será expresado en su totalidad y puede generar una derivación importante del fenotipo original, lo cual podría resultar en una malformación o una enfermedad severa, pero también podría tener una importante ventaja reproductiva, la

selección natural se encargaría de eliminar de la población a aquellos individuos enfermos y de aumentar la frecuencia de aquellos con alguna ventaja reproductiva. Esta aparición de variantes fenotípicas drásticas se ha observado en el aislamiento reproductivo inducido por el hombre en especies domésticas, e incluso en él mismo, tal es el caso de la hemofilia en las mujeres de la realeza y otros tipos de enfermedades producidas por la cruce entre familiares.

Anexo 1.

El *IFS* (Iterated Function System) es un lenguaje de geometría fractal cuyos elementos son transformaciones matemáticas y algoritmos computacionales, en informática se han utilizado para codificar imágenes de un alto grado de complejidad con una mínima cantidad de información.

La idea central del *IFS* es tomar un punto P y copiarlo en otra localidad determinada por una transformación lineal del plano (se fija un nuevo sistema de coordenadas relativo al primero) y repetir la operación con el nuevo punto una y otra vez para generar una imagen fractal. A diferencia de otros tipos de fractales como los conjuntos de Mandelbrot o Julia en los que no es posible deducir a partir del atractor (imagen) el conjunto que los genera, el uso de *IFS* permite mediante el uso de varias transformaciones del plano generar imágenes como helechos, árboles y hasta rostros humanos.



Diferentes tipos de transformación lineal del plano.

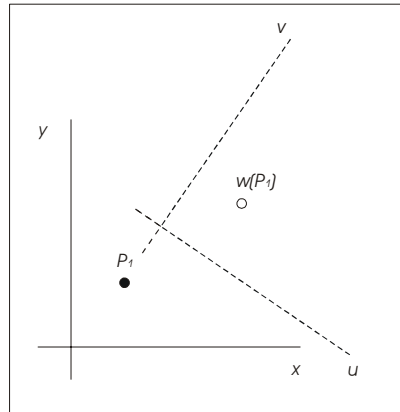
Para cualquier punto P , una transformación lineal F puede ser representada con respecto a un determinado sistema de coordenadas por una matriz de 2×2 .

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Donde, si $P = (x, y)$ y $F(P) = (u, v)$, entonces

$$u = ax + by,$$

$$v = cx + dy.$$



Transformación lineal de un plano, se fija un nuevo sistema de coordenadas a partir de uno previo mediante el uso de los coeficientes de una matriz de 2×2 .

A la matriz de transformación lineal se suma un vector de desplazamiento, de tal manera que la ubicación de un punto está determinada por seis números.

$$\left(\begin{array}{cc|c} a & b & e \\ c & d & f \end{array} \right)$$

Por lo tanto si $P = (x, y)$ y $w(P) = (u, v)$ entonces

$$u = ax + by + e,$$

$$v = cx + dy + f.$$

Apéndice A

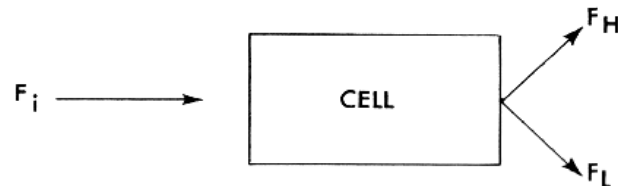
THE INFORMATION CONTENT OF A BACTERIAL CELL

HENRY LINSCHITZ, Department of Chemistry, Syracuse University

In this paper we present a brief, approximate calculation of an upper limit of the "degree of organization" of a bacterial cell. Expressed differently, we calculate a limit of the entropy change of the raw material which is converted to cell substance. This limit is obtained by estimating the entropy increase due to the metabolism represented by the cell waste per generation and equating this to the entropy decrease in assembling the cell. In order to compare this estimate with previous values (Dancoff and Quastler, this volume), the entropy is then expressed in units of "bits of information" per cell. For certain autotrophic bacteria, this number, representing an upper limit to the structural information content, is about 10^{13} bits. The significance of this result is discussed in terms of the structural and functional organization of the cell.

THERMODYNAMIC BASIS

The relation of the second law of thermodynamics to living systems has been discussed by many authors; e.g., Brillouin ('49), Schrödinger ('45), and Prigogine and Wiame ('46). For the purpose of our discussion, consider the following arrangement, in which a quantity of food and raw material at an average free energy level F_i is ingested by the cell per unit time:



This raw material is then metabolized and converted to products at two other average levels of free energy respectively higher (F_H) and lower (F_L) than F_i . The material produced at the high level may or may not be excreted from the cell. If it is retained within the cell then the rate of change of free energy due to the transformation of food material into protoplasm is $\Delta \cdot F_{internal} = F_H - \alpha F_i$ in which α is, roughly, the "fraction" of raw material converted to a high level product per unit time. More precisely, α is a stoichiometric factor relating the amount of cell substance produced per unit time to an equivalent amount of food-stuff. If the food intake may be split into two channels, one leading toward high and the other toward low level products, α is readily found. In the case of many autotrophs with simple nutrient requirements this is a convenient procedure. Similarly, the free energy change per unit time associated with the waste material is $\Delta \cdot F_{external} = G_L - \beta F_i$, in which β is a stoichiometric factor analogous to α .

The second law now requires that

$$\Delta \cdot F_{internal} + \Delta \cdot F_{external} \leq 0. \quad (1)$$

This inequality represents the familiar fact that the cell may produce within itself material having a *higher* free energy than an equivalent amount of food, provided that it is able to couple cell synthesis with other metabolic processes leading to a *net decrease* in free energy. This increase in free energy associated with the conversion of raw material into protoplasm may involve a heat term, an entropy term, or both. For example, if the heat of combustion of the cell is greater than that of an equivalent amount of food, the increase in free energy of the cell involves a positive ΔH term. This is most clearly the case in green plants. However, for cells whose carbon source is already at the carbohydrate level, growth will not be endothermic to any appreciable extent. The increase in free energy, if any, of the cell substance, is then due to a decrease in entropy.

The thermodynamic condition on the entropy changes associated with cell metabolism (Prigogine and Wiame, '46) is

$$\Delta \cdot S_{int} + \Delta \cdot S_{ext} \geq 0. \quad (2)$$

Once again, it is clear that the cell may *decrease* the entropy of *part* of the universe (i.e., its own substance) provided that it is

able to couple this process with other entropy-increasing processes involving either its own structure or interchanges (thermal and material) with the environment. A cell feeding on low entropy material thus has available a potentially large entropy increase which it can use to maintain its own unstable state of low entropy, or to decrease further the entropy of part of its foodstuffs in the course of its growth. In Schrödinger's ('45) vivid phrase, "The organism feeds on negative entropy."

If the cell were concerned solely with protoplasm synthesis, and if the anabolic and catabolic reactions involved in this process were perfectly matched and perfectly coupled, then no useful work would be lost to the environment and, from Eq. (1)

$$\Delta^{\circ}F_{int} = - \Delta^{\circ}F_{ext} \quad (3)$$

Actually, the cell is constantly doing very many things besides growing and a large part of the free energy supplied by the conversion of food to waste is thus diverted and not fixed in cellular structures. Moreover, since all these processes, including synthesis, must occur at reasonably high rates, finite gradients of various types, particularly of chemical potential, exist within the cell. The above equality is undoubtedly far from true. However, our purpose is to obtain *limits* on the entropy of the cell, and for that purpose, let us *assume* that the equality does hold, i.e., that that cell is perfectly efficient in protoplasm synthesis, and does nothing else. Then,

$$\Delta^{\circ}H_{int} - T\Delta^{\circ}S_{int} = - (\Delta^{\circ}H_{ext} - T\Delta^{\circ}S_{ext}) \quad (4)$$

or,

$$\Delta^{\circ}S_{int} = \frac{1}{T} [\Delta^{\circ}F_{ext} + \Delta^{\circ}H_{int}]. \quad (5)$$

In order to obtain $\Delta^{\circ}S_{int}$ we must now determine experimentally $\Delta^{\circ}F_{ext}$, the free energy change of the foodstuff ingested per unit time and discarded as waste, and add to this $\Delta^{\circ}H_{int}$, the difference in heat content between cell substance added per unit time and a stoichiometrically equivalent amount of food material. In any actual experiment, $\Delta^{\circ}F_{ext}$ includes the free energy which the cell must supply to carry on all its activities other than growth, such as osmotic work, internal viscous flow, locomotion, electrical dissipation, repairs of existing unstable structures, etc. These may

all be summed under the general heading of "work of maintenance" (Needham, '31). In addition, $\Delta \dot{F}_{ext}$ includes the losses due to inefficiency in the coupling of reactions involved in cell synthesis. $\Delta \dot{F}_{ext}$ will therefore be more strongly negative than what would correspond to ideally efficient growth and the calculation thus overestimates the "degree of organization" of the cell, i.e., gives too negative a value for the entropy change due to growth alone. To minimize the maintenance error, one must clearly take metabolic quotient data for cells at their optimum rate of growth. $\Delta \dot{S}_{int}$ then is the *limit* of the entropy change per unit time of the food material incorporated into the cell. If unit time is taken to be one cell generation, $\Delta \dot{S}_{int}$ will then be the limiting value of the entropy of the cell itself, referred to the entropy level of the food material.

Some comments on Eq. (5) may be helpful. The quantity $-\Delta F/T$ is the maximum increase of entropy of the universe due to an isothermal chemical process in which ΔF is the change in free energy of the reacting substances. Thus, if $\Delta F = \Delta H - T\Delta S$, where ΔH is the heat absorbed by the reactants from the environment (negative, in an exothermic process), then $-\Delta H/T$ is the corresponding entropy change of the *environment*. By definition, ΔS is the entropy change of the *reactants*. Hence, the *total* entropy change in the process is

$$-\frac{\Delta H}{T} + \Delta S = -\frac{\Delta F}{T}.$$

This assumes that all the energy liberated in the reaction appears, as heat, at temperature T . However, if the reaction is energy-coupled to another process, some of this energy may be stored again as heat of formation of the new products. For example, if a cell is allowed to grow in a calorimeter, the heat absorbed from the environment per unit time is $\Delta H_{ext} + \Delta H_{int}$, not ΔH_{ext} . If ΔH_{int} is positive, it represents energy which is fixed in newly formed cell substance, which never appears as heat and which therefore cannot contribute toward purchasing negative entropy for the cell.

Since the energy-releasing metabolic reaction must be spontaneous, $\Delta \dot{F}_{ext}$ will be negative. The sign of $\Delta \dot{S}_{int}$ thus depends on the value of $\Delta \dot{H}_{int}$ relative to $\Delta \dot{F}_{ext}$. Experiments show that the heat of combustion of protoplasm (dry weight) is not very far from that of an equivalent weight of carbohydrate. For the combustion of a carbohydrate $\Delta F \approx \Delta H$. Hence, for cells whose carbon source is car-

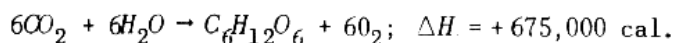
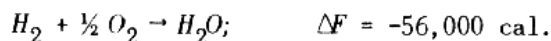
bohydrate, the relative magnitudes of ΔF_{ext} and ΔH_{int} are governed essentially by the energy efficiency of growth, i.e., the ratio

$$\frac{\text{heat content of unit weight of cells}}{(\text{heat content of food received to synthesize unit weight of cells}) - (\text{heat of maintenance})}$$

Since this ratio seems to be less than one (Rahn, '40) (with considerable uncertainty because of the energy required for maintenance), ΔS_{int} should be negative and the entropy of the foodstuff is expected to decrease upon conversion into cells. One interesting result of this viewpoint is the following: a low energy efficiency of an organism may represent either a real inefficiency or may merely be the consequence of a large negative entropy required for growth whose positive equivalent must be supplied in some manner; in this case it is simply by the discharge of more heat into the environment. The true measure of "thermodynamic efficiency" is evidently the ratio, $\Delta F_{int}/\Delta F_{ext}$, which requires the difficult evaluation of the entropy of the cell. In this paper the procedure is inverted, the efficiency assumed to be unity and thus a limiting value of the entropy is obtained.

NUMERICAL CALCULATION

A convenient organism for which sufficient data for a calculation are available is the autotrophic bacterium *B. pycnoticus*. This organism grows on a simple, completely inorganic diet, using the oxidation of gaseous hydrogen as an energy supply and bicarbonate as a source of carbon (Stephenson, '30). Measurements of the relative volumes of H_2 and O_2 absorbed during growth and the reasonable assumption that CO_2 reduction is carried to the carbohydrate level lead to the following reactions representing the overall assimilative process:



The free energy and heat data quoted here are taken from the review by Baas-Becking and Parks ('27).

Under the most favorable growth conditions, Ruhland found that eight volumes of H_2 were required for the absorption of one volume

of CO_2 . Thus, in order to reduce one gram, or 1/12 mole of carbon from the level of CO_2 to that of carbohydrate, 8/12 moles of H_2 are combusted, or $\Delta F_{ext} = (8/12)(-56,000 \text{ cal.}) = -37,400 \text{ cal.}$ per gram of carbon assimilated.

The change in heat content of one gram of carbon upon conversion from CO_2 to cell substance (taken as carbohydrate) is just the negative of the heat of combustion of 1/72 mole of glucose. Thus

$$\Delta H_{int} = \frac{675,000}{72} = +9,400 \text{ cal. per gram of carbon assimilated.}$$

Within the limits of error of this estimate, we may neglect the change in heat content of the raw materials other than CO_2 which are converted to cell substance.

Analysis of bacterial cells shows that they contain approximately 10% carbon (Knaysi, '51). One gram of carbon therefore represents about 10 grams of cells. The bacteria dealt with here have a volume of roughly $(1\mu)^3 = 10^{-12} \text{ cc.} = 10^{-12} \text{ gram.}$ Hence, one gram of carbon represents 10^{13} cells.

The change in entropy of the *assimilated* material upon conversion from food to cells, per gram of cellular carbon, fed as CO_2 is then given by Eq. (5) as

$$\Delta S_{int} = \frac{1}{300} [-37,400 + 9,400] \text{ cal./deg.} \sim -90 \text{ cal./deg.}$$

This corresponds to about $-10 \text{ cal./deg./gm.}$ of cells, $-10^{-11} \text{ cal./deg. per cell}$, or to about $-5 \times 10^{12} \text{ cal./deg. per "mole" of cells}$, regarding the entire cell as a single molecule. In order to obtain the number of "bits of information" per cell, we use the relation (Linschitz, Information and Physical Entropy, this volume)

$$H_{bits} = -\frac{\Delta S/\text{mole}}{R \ln_e 2} = -\frac{\Delta S/\text{cell}}{k \ln_e 2}$$

$$(R \ln_e 2 = 1.377)$$

and thus obtain the limiting value of $10^{13} \text{ bits/cell.}$

DISCUSSION

The entropy value obtained above represents, as pointed out previously, a lower limit. In reality, ΔS_{cell} cannot be so strongly

negative as indicated. However, the figure of -10 cal./deg/gram of cells does not seem unreasonable, when compared with that corresponding to simple, physical ordering processes. Thus, the entropy of condensation of water at 100°C is -1.45 cal./deg./gm. The order of magnitude also would not be seriously affected if data on other autotrophs were used. The free energy efficiency (with respect to carbon assimilation) among the organisms studied varies from the highest figure of 25% (for the hydrogen bacteria used in our calculation) to about 5% for other forms (Baas-Becking and Parks, '27). This still yields about 10^{13} bits per cell. The uncertainty in ΔH_{int} is also not serious, provided that ΔH_{int} is considerably less than ΔF_{ext} .

One may attempt an estimate of the error involved in this calculation. The energy dissipated in maintenance and inefficient balancing of negative entropy and positive entropy production by the cell will both lead to too low (negative) values. However, with rapidly growing cells (generation time ~ 1 hour) the relative effect of maintenance is decreased. Also, there is some indication that rapid growth is not necessarily related to inefficient entropy metabolism. Thus, rapidly reproducing molds and bacteria show about the same energy efficiency of growth as larger forms (Rahn, '40). If we estimate that the cell even at optimum growth wastes, at worst, 99% of its entropy supply in maintenance and dissipative activity, then the information content per cell is reduced to $10^{10} - 10^{11}$ bits. This range of values is similar to that estimated by Dancoff and Quastler (this volume) on the basis of quite different arguments, involving structural models and genetic characteristics of cells.

On the basis of this calculation, one may also estimate limiting values of the entropy of formation of proteins from simple carbon, nitrogen and oxygen sources. In the case treated above, taking the cell to be 10% protein and considering *all* of the entropy change to reside in the formation of these highly specific and improbable molecules, we obtain -10 cal./deg./gram of cells or -100 cal./deg./gram of protein. Assuming the average molecular weight of protein to be 10^5 , this yields an estimate of -10^7 cal./deg./mole of protein or 10^7 bits/molecule. It must be pointed out that if the configurational entropy of the protein due to the permutations of residues along the peptide chain turns out to be small, then this calculation is sensitive to the particular forms in which nitrogen is supplied in the nutrient material.

Correlations may also be looked for between the energy-efficiency of growth and the nutritive requirements of the cell. Suppose that an organism is available which has very simple requirements and suppose also that mutations of this organism can be obtained which require progressively more and more growth factors. These factors may be fed in such small quantities that they do not seriously affect the calorific or free energy value of ingested nutrients, although they make growth possible. The cell with special growth requirements presumably is deficient in many of the enzyme systems which the autotrophic form possesses. If this means that the enzymes, or similar protein molecules are lacking completely, then the specialized cell should have a slightly less negative entropy than the autotrophs. Therefore, the specialized cell should not have to produce so much positive entropy (by thermal losses) in order to obtain the negative entropy required for growth, and one might expect that the growth requirements become more specific and numerous, i.e., the more "parasitic" the organism, the higher should the energy-efficiency be. It may be significant that some highly cultivated strains of brewer's yeast show extremely high energy-efficiencies in growth (Rahn, '40).

STRUCTURAL ENTROPY AND FUNCTIONAL ORGANIZATION

Although the gross thermodynamic calculation presented here seems straightforward, the question arises regarding the biological significance of the "entropy" we have obtained, and of the conversion of this entropy to "information," using a factor derived from the thermodynamics of inert systems. Does this procedure lead us to a number which, in some way, represents the *living* character of the system? Various problems immediately appear in this connection, all of which seem to be concerned with the difficulty of defining the entropy of a system composed of *functionally interdependent* units, and of relating this to the information contained therein.

To begin with, we state explicitly that for the purpose of definition of physical entropy, the living cell is considered to be a mixture of various chemical species, at definite activities, pressures and temperatures, and that the entropy is assumed to be expressible as an explicit function of these thermodynamic variables.

From a microscopic viewpoint, one could, in principle, knowing all the species, structural forces and resultant quantum states, write down the entropy of any arbitrary mixture of n components in p phases joined at s surfaces. However, this physical entropy does *not* seem to provide the proper measure in which to express the *functional* information contained within the organism. Consider, for example, the problem of measuring the entropy of an iron crystal weighing say, 10 grams at 0°C. The third law of thermodynamics provides a definite procedure for doing this, and the value thus obtained has a precisely defined physical meaning, in terms of the average number of quantum states available to the crystal when it is immersed in a thermostat at the given temperature. If our piece of iron is now shaped into the form of a gear wheel, the change in its physical entropy will be negligible. If anything, the entropy will increase, due to the formation of imperfections at the crystal surface. However, from the standpoint of a mechanism into which the gear is to be inserted, it is *just* the shaping operation and not at all the microscopic configuration which gives the gear its value and which embodies the information required for the gear to carry out its function. Increasing the physical entropy of the gear by dissolving some nickel in the iron will have no effect on its articulation into the machine. On the other hand, if in shaping the gear, one tooth was not properly cut, no appreciable change in its physical entropy would result although the effect, as far as the operation of the machine were concerned, would be to destroy its function completely. It is evident that the degree of organization, or information content, of the gear wheel must be judged *relative to its function*, and that, at least for macroscopic units, this functional organization bears little relation to the physical (or microscopic) entropy of the unit. In the case of a cell containing numerous protein molecules, of highly specific function, the microscopic, or structural, entropy may be closely related to the functional information content if the relation of the various units is maintained by definite chemical linkages, so that the organized entity is truly molecular. However, if the interaction occurs in the cell simply by proximity, by such agencies as diffusion of substrates, inhibitors or organizers, or by hormonal or even electrical communication, in short, primarily by virtue of the existence of a common enclosing membrane, then one would expect that the physical entropy does not measure the essential organization of the functioning complex.

This conclusion is also reached by another argument, which may possibly clarify the situation. The statistical mechanical definition of entropy as $S = k \ln P$ must be used with caution in this context because P , the number of *available* thermodynamic complexions, does *not* correspond to *all* the possible (even non-lethal) complexions that may conceivably exist, involving the multitude of compounds and structures within the organism. On the contrary, the available number of complexions is severely restricted by the course of differentiation built into the zygote and by the requirements of survival of the organism. As Schrödinger (1945) clearly points out, in biology, order arises out of order, whereas in physics, order arises out of disorder. The entropy function is the expression of the physical, not the biological mechanism. In physics, entropy is a measure simply of the *structure* of a system, whereas in biology, what is really being studied is a *process*, structure-dependent, obviously, but still a *functioning* structure. One should expect that the functioning of a cell, through which it is constantly producing positive and negative physical entropy must be incorporated somehow into a measure of its biological information content. Thus, a smoothly functioning or live cell must be considered to have a higher degree of organization than a poorly functioning or dead cell, simply by virtue of its organized activity, which is to say, its continued existence.

This concept may possibly be made a little more precise by means of the following definition of a functional "entropy": out of all conceivable thermodynamic complexions to be considered in counting P , in the formula $S = k \ln P$, let us admit only those whose realizations would not affect the functioning of the system. The choice of process with respect to which the "function-state" is defined specifies the nature of the corresponding functional "entropy." Thus, in a dead organism (or disabled machine) a very large number of complexions are possible, in all of which the machine continues to be disabled, and the entropy is correspondingly large. However, in a living organism (or working machine) only a small number of complexions are available, *consistent with the maintenance of function*, and the entropy is low. Moreover, the change in P is discontinuous, according as the change in "function-state" is discontinuous. Imagine again, for example, the machine in which a single small, critical gear may be inserted or removed. The insertion of the gear and the resumption of function immediately effects the

coupling and interaction of all the parts and leads to a discontinuous change in the number of complexions which are functionally permissible. Something like a Gibbs' paradox seems to be involved here, sharply distinguishing the organization of the functioning from the non-functioning machine, although, once again, we note that the physical entropy of the two systems will be practically identical. It is interesting that attempts have been made to determine real calorimetric effects accompanying death or cessation of function in cells. Inconclusive results so far show a slight evolution of heat (Meyerhoff, '12; Lepeschkin, '37). The sudden availability of new thermodynamic complexions following death and the relaxation of homeostatic control might, under adiabatic conditions, result in a very slight drop in temperature. At the same time, the release of various gradients in chemical, osmotic, and electrical potential could well lead to sufficient energy dissipation and heat evolution to mask the entropy effect entirely.

The inadequacy of physical entropy alone to measure the biological information content is also seen in the extensive nature of entropy, by which the entropy of two identical cells is twice that of each cell. However, the biologist and the communication engineer might both argue that little more information is present in two identical cells than is already present in one. From the point of view of information theory the situation corresponds to a message in which a very high degree of redundancy exists between successive units, or in this case, generations.

In the light of the above discussion, some comment is needed on the validity of the conversion of physical entropy to bits of information, using the factor $R \ln 2$. This factor is derived from a model in which the lack of orientation coupling between the molecules at each site in the lattice is the expression of the lack of interdependence between successive binary choices. In a functioning cell, on the other hand, the correctness of every binary choice involved in assembling the cell is determined by the effect of this choice on the function of the whole organism. In other words, in this case a very high degree of functional coupling exists which is not considered in the purely structural information represented by the $R \ln 2$ conversion factor. Nevertheless, if the cell is regarded as a single "molecule" and the "molar entropy" computed on this basis, then all the interactions are kept within the molecule and

the maximum number of binary choices leading to a *structure* consistent with cell function should still be given by this calculation.

The resulting value obtained in the "numerical calculation" section thus represents the *physical entropy* of the cell, which is to say its *microscopic, structural organization*. This is the amount of negative physical entropy that the cell must produce in order to make itself. However, the calculation tells us nothing of the particular way in which this entropy is channeled to produce a functioning cell, nor of the macroscopic functional organization established thereby. This latter type of organization is the result of non-structural coupling between functioning units constituting the organism. Some other measure of this organization, which embodies the biological information residing in the organism, must be sought. The evolutionary principle would indicate that, at this level, one will find a tendency in nature to increase order, rather than to decrease it.

REFERENCES

- Baas-Becking, L. G. M., and Parks, G. S., *Physiol. Rev.* 7:85 (1927).
Brillouin, L., *Am. Scientist* 37:554 (1949).
Knaysi, G. A., *Elements of Bacterial Cytology*, Comstock Pub. Co., Inc., Ithaca, 1951.
Lepeschkin, W. W., *Protoplasma* 28:175-199 (1937).
Meyerhof, O., *Arch.f.ges. Physiol.* 146:159-184 (1912).
Needham, J., *Chemical Embryology*, Vol. II, University Press, Cambridge, England, 1931.
Prigogine, I., and Wiame, J. M., *Experientia* 2:451-453 (1946).
Rahn, O., *Growth* 4:77-80 (1940).
Schrödinger, E., *What is Life?*, University Press, Cambridge, England, 1945.
Stephenson, M., *Bacterial Metabolism*, 1st ed., Longmans, Green & Co., Ltd. London, 1930.

Referencias

1. Adami, C., Wilke, O., (2004). *Experiments in Digital Evolution (Editor's Introduction to the Special Issue)*. *Artificial Life*. 10:117-122.
2. Bentley, P., Corne, D., (2000). *An introduction to creative evolutionary systems.*, Great Britain.
3. Blanc, M. (1982). *Las teorías de la evolución hoy*. *Mundo científico* 12(2): 288- 303.
4. Dawkins, R., (1990). *The Blind Watchmaker*, Penguin Books, 1990.
5. Eguiarte, L., (1986). *Una guía para principiantes a la genética de poblaciones*. *Ciencias*, 3: 30-38
6. Eldredge, N., Gould, S. J. (1972). *Punctuated equilibria: an alternative to phyletic gradualism*. T. Schopf (ed) ;*Models in Paleobiology*. Freeman. Cooper and Co. , San Francisco,. Pp 82-115
7. Evans, M W. & Marciniak, J (1987). *Software Quality Assurance and Management*. New York, NY. John Wiley & Sons, Inc.
8. Fontana, W., & Schuster, P. (1998). *Continuity in evolution: On the nature of transitions*. *Science*, 280, 1451–1455.
9. Gerhart, J., (1997). *Cells, Embryos, And Evolution.*, Blakwell Science., USA.
10. Goodwin, B.C., (1994). *How The Leopard Changed Its Spots*. Weidenfeld and Nicolson, London.
11. Gould, S. J., (1980). *El pulgar del panda*. Ed. Grijalbo-Mondador. Barcelona 153-158.
12. Herrán, M. (1998). *Computación evolutiva*. Solo programadores, 43
13. Hogeweg, P., (2000). *Shapes in the Shadow: Evolutionary Dynamics of Morphogenesis*. *Artificial Life* 6: 85–10.
14. Huynen, M. A., & Hogeweg, P. (1994). *Pattern generation in molecular evolution: Exploitation of the variation in RNA landscapes*. *Journal of Molecular Evolution*, 39, 71–79.

15. Kampis, G. (1991). *Self-Modifying Systems in Biology and Cognitive Science. A New Framework for Dynamic, Information and Complexity*. Pergamon Press. Great Britain.
16. Kimura. M. (1982). *The neutral theory of molecular evolution*
17. Levy, S. (1992). *Artificial Life, A Report From The Frontier Where Computers Meet Biology*. Vintage Books, USA.
18. Lewin, B., (2000). *Genes VII*. Oxford University Press., UK. Pp. 75.
19. Majerus, M. E. N., (1999). *Evolución y mantenimiento del melanismo industrial en los Lepidoptera*. Bol. SEA, 26 : 637-649 [fecha de acceso 11 de enero de 2003]; URL disponible en <http://entomologia.rediris.es/sea/bol/vol26/s5/articulo/>.
20. Martínez, J., (1986). *¿Que es la evolución molecular?* Ciencias, 3: 24-29
21. Mayr, E., (1982). *The Growth of Biological Thought*. Cambridge, Great Britain.
22. Merelo, G., (1999). *Informática evolutiva: Algoritmos genéticos*. [fecha de acceso 12 de Marzo de 2003]; URL disponible en <http://kal-el.ugr.es/~jmerelo/>
23. Neumann, V. (1948). *General and Logical Theory of Automata*. Lectura en el *Hixon Symposium* tomada de Taub, A. H., (1963)., *John von Neumann: collected Works*, Pergamon, Elmsford, N.Y. vol. 5, pp. 288-328.
24. Olea-Franco. A. (1986). *La teoría del equilibrio puntuado. Una alternativa al Neodarwinismo*. Ciencias, 3: 46-58.
25. Ray, T. S. (1994). *An evolutionary approach to synthetic biology, Zen and the art of creating life*. Artificial Life 1(1/2): 195-226. MIT Press.
26. Richards, F, M., (1991). *The Protein Folding Problem*.
27. Setlow, R. Pollard, E. (1964). *Molecular Biophysics*. Addison Wesley. EUA 1964, 2ª ed.
28. Shannon, C. E., (1948). *A Matematical Theory of Comunication*. Reprinted with corrections from The Bell System Technical Journal, Vol. 27, pp. 379-423, 623-656.
29. Smith, J. M., (1958). *The Theory OF Evolution*. Cambridge university press, Great Britain.

30. Stanley, K. O., Miikkulainen, R. A., (2003). *Taxonomy for Artificial Embryogeny*. *Artificial Life* 9: 93–130.
31. Turing, A.M., (1937). *On Computable Numbers, With an Application to the Entscheidungsproblem*, *Proceedings of the London Mathematical Society*, (2) 42, pp 230-265; correction *ibid.* 43, Pp 544-54.