

UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

PRINCIPIOS DE LA INFORMACIÓN BIOLÓGICA

T E S I S

QUE PARA OBTENER EL TÍTULO DE

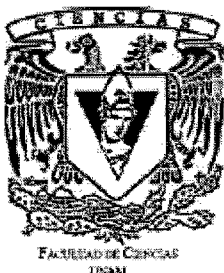
BIÓLOGO

P R E S E N T A:

JOSÉ AGUSTÍN MERCADO REYES

DIRECTOR DE TESIS:
DR. PABLO PADILLA LONGORIA

2009





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.

A peine exprimons-nous quelque chose qu'étrangement nous le dévaluons. Nous pensons avoir plongé au plus profond des abîmes, et quand nous revenons à la surface, la goutte d'eau ramenée à la pointe pâle de nos doigts ne ressemble plus à la mer dont elle provient. Nous nous figurons avoir découvert une mine de trésors inestimables, et la lumière du jour ne nous montre plus que des pierres fausses et des tessons de verre; et le trésor, inaltéré, n'en continue pas moins à briller dans l'obscur.

Maurice Maeterlinck, *La Morale Mystique*

Agradecimientos

En primer lugar, y sobre todo, a mis padres: Gracias, Mamá, gracias, Pop. No sólo me han apoyado en todos sentidos a lo largo de mi vida: han tolerado mis neurosis y han platicado conmigo desde metodología científica y arte hasta mis inquietudes más fúnebres. Hicieron de mí lo que soy y no podría pedir, o siquiera imaginar, a mejores padres.

A Pablo, mi asesor; por su invaluable guía a través de este trabajo, por haber sido el primero que creyó en estas ideas (¡y por seguirlo haciendo!), por su increíble y polifacético conocimiento; pero también por su amistad que ha permitido pasar horas platicando de intereses comunes, ideas nuevas y proyectos futuros.

A mi jurado: Ana Barahona, Alicia Negrón, Víctor Valdés y Luis Delaye. Su participación en proceso enriqueció enormemente mi tesis; pero además me ayudaron en imprevistos surgidos y me aconsejaron en el trayecto (¡incluso intercontinentalmente!). El respeto que les tengo, personal, académica y científicamente, me hace esperar y proponerme que esta tesis no sea la última colaboración que tengo con ustedes.

Al Doctísimo Roberto Ruiz Guadalajara, probablemente mi mayor influencia intelectual.

A Adriana, que siempre ha pensado que soy listo, que me ha aguantado desde que me acuerdo, que me conoce como soy. Y también (¿por qué no?) al perro Draco: casi todo el texto de este trabajo fue escrito escuchando junto a mí su respiración o sus ronquidos.

A quienes me ayudaron a soportar los primeros meses del 2008. Marco, Delfo: sin ustedes este trabajo hubiera llegado a un fin abrupto e irremediable.

A Iván, que aunque no presencié la escritura de esta tesis me acompañó y toleró en el largo proceso de los últimos trámites: companions in misfortune in oblivion's popped slope.

Álvaro: eres quien más sabe de mí, quien mejor me conoce, quien más me ha ayudado. Gracias por aguantar cada pregunta, retórica o no.

Al inevitable Christian Wittgensteiner. Te encuentras en este trabajo.

A mi gente de la facultad, en particular a esas tres reinas con quienes ya no existen los silencios incómodos – la Morales, por esas pláticas eternas y esos momentos camaleónicos, desde la Casa del Lago de Chapultepec hasta las comidas de las fiestas del barrio; a la Nay, probablemente la persona con la que más me he reído en mi vida; y a la Elis Froster: *todos* los que me conocen te conocen, y saben lo mucho que te quiero.

A Elito: gracias por todo: lo creas o no, una sola de esa frase tuya dio origen a las primeras sesenta páginas de este trabajo. Eres *genial*, en todos los sentidos de la palabra.

A Eda y a Patty, los únicos destellos de coherencia en ese edificio de mosaicos color mostaza. Ambas, a su manera, me enseñaron la belleza del cambio y la permanencia. LL, lo creas o no, diste luz verde a este texto aquella tarde en el cuarto de cultivo. LL, extraño nuestras horas en la barda de piedra volcánica. A ambas las conocí el mismo día; ambas me daban un poco de miedo.

A los amigos que desde hace más de diez años me acompañan, probablemente por compartir mi humor en ocasiones corrosivo: Toño, Sandra, doña Bernalg.

Por último -but not least- quiero agradecer a Carlos Rodríguez, mi chiquitín especial, mi amigo más cercano y más antiguo. Además de tener una fe inexplicable en mi trabajo, tomó parte activa en éste. Sus excelentes figuras (figs. 3, 17, 24) ilustran y aclaran los argumentos más importantes del trabajo.

To all you.

Índice

Prólogo	11
Sumario.....	12
Parte I: Definición de Vida	
1. 1 Introducción	
El lenguaje.....	13
Necesidad de una definición.....	14
Problemas característicos.....	17
Problemas actuales.....	21
1. 2 Las definiciones actuales de vida	
Elementos comunes.....	24
Características de una definición.....	25
Reproducción.....	28
Darwinismo e identidad.....	31
1. 3 Complejidad física e informacional	
Consideraciones termodinámicas.....	38
Información:	
<i>como mensaje</i>	42
<i>física</i>	44
<i>contextual</i>	44
<i>como patrones</i>	45
<i>semántica</i>	45
Complejidad Algorítmica.....	50
1. 4 Semiótica	
Símbolos.....	55
Códigos naturales.....	57
Definición de vida y consideraciones finales.....	63

Parte II: El origen de la información

2.1 La paradoja del huevo y la gallina

El Origen de la vida.....	66
El Mundo del RNA: “Replicación primero”.....	66
Problemas químicos del mundo del RNA.....	68
<i>Síntesis prebiótica de nucleótidos</i>	69
<i>Polimerización específica del RNA</i>	70
<i>Algunas soluciones propuestas</i>	71
Autoorganización: “Metabolismo primero”.....	72
<i>Las teorías de Wächtershäuser y Morowitz</i>	73
<i>Metabolismos abstractos</i>	77
Problemas conceptuales de las teorías de la Replicación Primero.....	79
<i>Replicación específica</i>	79
<i>Catástrofes</i>	80
<i>Surgimiento de genes</i>	80
<i>Impedimentos estéricos</i>	81
<i>Continuidad</i>	82
Justificación de la propuesta: La necesidad de teorías alternativas.....	83

2.2 El Código genético

Regularidades y patrones

<i>La tercera letra del codón: Redundancia</i>	86
<i>La segunda letra del codón: Hidrofilia</i>	86
<i>La primera letra del codón: Biosíntesis</i>	88
<i>El número de codones de cada aminoácido. El caso de la arginina</i>	90
<i>Reducción de errores</i>	91
Elementos del código genético.....	92
Algunas teorías acerca del origen del código genético	
<i>El accidente congelado / Coevolución</i>	93
<i>Estereoquímica</i>	95

2.3	Una teoría nueva	
	Preliminares.....	96
	Mecanismos de reacción de Copley.....	97
	Propuesta:	
	1. <i>Principios de la información biológica</i>	101
	2. <i>Un código genético primitivo</i>	105
	3. <i>RNA catalítico: rRNA</i>	109
	4. <i>Orígenes del tRNA</i>	111
	5. <i>Aminoacil-tRNA sintetasas (aaRSs)</i>	115
	6. <i>Otras funciones posibles del RNA</i>	116
	Propuestas a largo plazo	
	<i>Experimentos in vitro</i>	117
	<i>Experimentos in silico</i>	118
	<i>Umbrales</i>	119
	<i>Metabolismo no específico</i>	120
	<i>Un experimento multidisciplinario</i>	121
2.4	Conclusión	
	Preguntas abiertas y argumentos principales.....	122
	El Origen de la vida (2).....	123
	Epílogo.....	127
	Anexo: Aminoácidos esenciales.....	129
	Bibliografía citada.....	131

Prólogo

El tiempo, la vida y la muerte, la mente e incluso el arte son fenómenos que nos representan; incluso podemos decir que son fenómenos que *somos*. En ocasiones, cuando se consideran las propiedades de éstos uno tiende a olvidar, entre datos experimentales, la noción y la relevancia del objeto del estudio. También ocurre el caso opuesto: existe el riesgo de perderse en ramificaciones metafísicas que tal vez son interesantes pero en gran parte irrelevantes para lograr los objetivos concretos que se deben de proponer al inicio de toda investigación.

A lo largo de este escrito, que es en última instancia un ensayo acerca de la naturaleza de la vida, traté de no perder de vista estas dos posiciones aparentemente contradictorias. Por un lado, traté de mantener el mayor rigor científico que me fue posible; por otro, intenté conservar ese asombro, no siempre agradable, de describirme a mí mismo y a todos los demás seres vivos que conozco.

El objetivo último y único de este trabajo es proponer un punto de investigación experimental del origen de la vida, y justificar la elección de ese punto determinado y no otro. Por supuesto, y como lo demostró el fracaso de la geometría euclidiana, en el mundo “real”, el que existe fuera de los modelos simples pocas veces hay “puntos” y son mucho más usuales los gradientes, superposiciones, ramificaciones, etcétera. La primera parte del trabajo intenta delimitar, pues, el momento a investigar; la segunda parte es un acercamiento experimental hacia la búsqueda de soluciones.

Platón afirmaba que la opinión no es conocimiento, y por tanto no es ciencia. Yo sé, desde hace tiempo, que todo lo que pueda decir no son más que opiniones. Por esto, la ambición más grande de este trabajo no es que sirva para asentar verdades esenciales, sino como punto de partida para nuevas ideas, sean éstas derivadas de la continuación de lo que aquí se propone o de la crítica que reciba. Generar nuevas ideas es en sí una función envidiable.

Sumario

El presente trabajo se divide en dos partes, cada una de las cuales es una unidad en sí misma. Sin embargo, ambas están íntimamente relacionadas: los argumentos centrales de una apoyan a la otra.

Dado que este escrito es esencialmente una investigación del origen de la vida, la primera parte (“**Definición de vida**”) se puede considerar como una larga y, creo, necesaria introducción: es un intento de formular una descripción, meramente funcional, del proceso *vida*. Se comienza por señalar la dificultad o incluso la imposibilidad de definir aun los conceptos más simples de nuestro lenguaje, fenómeno ya analizado por Ludwig Wittgenstein en sus *Investigaciones Filosóficas*. Bajo tal nota precautoria, se procede a analizar las más importantes características utilizadas en los diversos acercamientos a las definiciones de vida. Finalmente, se hace énfasis para construir una descripción, en dos características de los organismos vivos conocidos: el manejo semiótico de la información llevado a cabo por el código genético y el contexto fisicoquímico en el que tal proceso ocurre. Dicha descripción está evidentemente influenciada por los trabajos de Erwin Schrödinger, en particular su libro *What is Life?*.

La segunda mitad del trabajo (“**El origen de la información**”) es una propuesta experimental encaminada a dilucidar el origen del código genético. Para iniciar tal discusión, se contrasta los dos principales grupos de teorías del origen de la vida (los llamados de “metabolismo primero” y de “información primero”), y se enumera algunos problemas conceptuales importantes del segundo grupo, que incluye al llamado “Mundo del RNA”. Tomando, pues, como base las teorías de Metabolismo Primero (y en particular los metabolismos de superficie propuestos por Wächtershäuser), se analiza las distintas teorías del surgimiento del código genético y se propone una original, con base en una serie de mecanismos de reacción publicados recientemente por Shelley Copley. Para concluir, se propone un programa experimental a futuro y una secuencia alternativa de los eventos que pudieron haber dado origen a la vida terrestre.

PARTE I: DEFINICIÓN DE VIDA

*Sag nicht »Es gibt keine ›letzte‹ Erklärung«.
Das ist gerade so, als wolltest du sagen:
»Es gibt kein letztes Haus in dieser Strasse;
man kann immer noch eines dazubauen.«*

Ludwig Wittgenstein, *Philosophische Untersuchungen*

1. Introducción

Lenguaje

La experiencia humana es inimaginable sin el lenguaje. La capacidad de comunicarnos tal y como nos comunicamos, ya sea con otros o con nosotros mismos, reside en el fundamento mismo de nuestra existencia. Sin embargo, cualquier intento de acercarse racionalmente al lenguaje resulta inevitablemente en un fracaso. No me refiero únicamente a la gran imperfección lógica que Schrödinger (Schrödinger 1998) observa en cada idioma: las palabras son finitas y las definiciones son necesariamente circulares. Ludwig Wittgenstein (Wittgenstein 1958) va un paso más allá con sus llamados “juegos lingüísticos”, experimentos mentales que buscaban deconstruir a la comunicación en sus partes. Sus preguntas (¿Cómo aprendimos nuestro lenguaje? ¿Cómo utilizamos nuestro lenguaje?) exploran la capacidad de la mente humana de funcionar sobre cimientos inexistentes. Por ejemplo, intenta definir la palabra “juego”, y descarta una a una las definiciones tentativas que él mismo propone. Sin embargo, cualquier persona utiliza la palabra “juego” de manera correcta, e incluso es capaz de discernir qué es un juego y qué no lo es cuando se le pregunta. El lenguaje, concluye, está asociado íntimamente a nuestra mente y a nuestra realidad ordinaria. Al transportarlo a un medio metafísico (por ejemplo, buscar definiciones esencialistas) comienzan a surgir dificultades filosóficas.

Aún los lenguajes formales, creados como construcciones puramente sintácticas, resultan falibles. El caso más famoso fue la ardua tarea que Bertrand Russell y Alfred N. Whitehead emprendieron al tratar de construir un sistema formal que fuera completo y consistente en sí mismo. Tal tarea, aunque monumental, sólo cubría los principios básicos de las matemáticas: los números ordinales, cardinales y reales, extendiéndose al campo de la teoría de conjuntos. Durante un tiempo se consideró a esta obra (*Principia*

mathematica) la base indiscutible de la teoría matemática; y sin embargo, unos veinte años después, Kurt Gödel postuló sus teoremas que no sólo demostraban que un trabajo tan exhaustivo y básico tenía una falla intrínseca, sino que todos los sistemas lingüísticos, lenguajes formales incluidos, eran inherentemente incongruentes.

Tal vez sea por estas fallas que en el caso del lenguaje científico ocurre un fenómeno extraño: los objetos principales de estudio de cada rama permanecen sin definición. La física no ha logrado dar una respuesta satisfactoria cuando se pregunta qué es el tiempo; la filosofía matemática aún debate si los números son entidades reales o abstracciones humanas sin manifestación fenomenológica. El caso de la biología no es distinto: entre las preguntas fundamentales de cada una de sus áreas (¿Qué es una especie? ¿Qué es un individuo?) se encuentra la pregunta fundamental: ¿Qué es la vida?. Tal persistencia no debe de ser una sorpresa, a la luz del fracaso constante del lenguaje por sostenerse a sí mismo.

Necesidad de una definición

Karl Popper, sentenciosamente, escribe: “Las preguntas como ‘¿Qué es la vida?’ o ‘¿Qué es la gravedad?’ no desempeñan ningún papel en la ciencia” (Popper 1995). Para Popper sólo existen dos clases de definiciones, que se encuentran los polos de una dicotomización evidentemente artificial. Popper coloca al *definiendum* del lado izquierdo y al *definiens* del lado derecho. La primera clase de definición es la aristotélica o *esencialista* (hago notar que los términos utilizados están definidos de manera estipulativa). Ésta es la definición que va de izquierda a derecha: la palabra o término suscitan una pregunta que es contestada con la fórmula definitoria. Así, la esencia de la palabra tiene que ser paralela a la esencia de la cosa a la que se refiere: la descripción de la esencia *surge* de la palabra. Popper niega toda utilidad a la definición esencialista con distintos argumentos. El principal dice que el método esencialista sustituye problemas factuales por problemas meramente verbales. Es decir, este método no analiza al fenómeno sino al lenguaje.

Tales argumentos son comprensibles. En la elaboración de un trabajo como éste, concerniente a las definiciones, es necesario detenerse en cada paso para comprobar que no se ha convertido en un ensayo puramente lingüístico o etimológico. Éste, pues, es un trabajo para aclarar un término del lenguaje y al hacerlo, utilizar los conceptos que

surgen de la desambiguación como apoyo para proponer nuevas perspectivas. No es un trabajo que se ocupe del lenguaje *en sí*, y de ninguna manera toma al lenguaje ni al signo lingüístico [vida] para desarrollar una propuesta.

La opción que Popper da, pues, es la alternativa en la dicotomía, una definición *nominalista*, que va de derecha a izquierda. Es decir, las definiciones sólo sirven como abreviaturas de conceptos más largos, inventadas *ad hoc* para hacer que el discurso científico sea más fluido. Así, Popper únicamente acepta, en el lenguaje científico, una nomenclatura con base en descubrimientos o invenciones; sólo considera “definibles” casos como una nueva cepa de bacterias, en la que la definición sólo es una manera corta (un nombre) de decir una frase larga (los métodos de tinción a los que responde, los genes característicos, la especie y variedad, etc). Primero vienen los conceptos; después la palabra que los nombrará.

Resultan aparentes las fallas inmediatas del acercamiento popperiano. Su propuesta no es válida como metodología porque toda palabra es una definición nominalista: cualquier palabra refleja un concepto que se expresa con un grupo mayor de palabras, con mayor o menor precisión. Además, el enfoque nominalista puede caer en prácticas de abuso de neologismos; Ernst Mayr solía redefinir o crear palabras para evitar resonancias históricas que él consideraba anticientíficas. Por ejemplo, diferencia “teleonómico” de “teleológico” para cargar al primer término de un significado genético (Mayr, 1992). Para algunos autores, esta práctica no es más que una creación de eufemismos (ver, por ejemplo, Sheldrake 1995)

Señalar las fallas en la propuesta de Popper es menos pertinente, para fines de este trabajo, que analizar su crítica. Él descalifica todo intento de articular una definición o siquiera una descripción de un fenómeno visible y conocido previamente, como el movimiento, la gravedad, el tiempo o la vida. Tales programas de investigación no se acercan a un problema de palabras, sino de hechos complejos y trascendentes que ocurren constantemente frente a nuestros ojos. El intento de entenderlos ha dado frutos constantes a la ciencia, y se ha ramificado a otras áreas como la filosofía e incluso el arte. Al acercarse al problema del origen de la vida, se hace inmediatamente presente la necesidad de tener una definición válida, independientemente de las imposibilidades y tautologías inherentes al lenguaje, y de las voces, como la de Popper, que dicen que esa práctica es estéril y confusa. No puedo insistir lo suficiente que este esfuerzo no es ni debe ser una práctica lingüística, basada en el estudio del uso de las palabras, ni se trata

de conciliar en un solo término el abanico de implicaciones que tiene el vocablo "vida". El esencialismo en una definición científica debe de ser un esfuerzo por condensar una teoría acerca de la fenomenología de un concepto: el esqueleto de un argumento que puede ser apoyado o contradicho, ya sea por cualidades lógicas de la argumentación misma o por la constante actividad de elucidación epistemológica o acumulación de información.

Tomando a la definición desde este punto de vista, la actividad de definir en este caso tiene ramificaciones profundas. Además de estructurar las propuestas de descripción de un fenómeno, es claro que sirve de punto de partida para el estudio del origen de la vida, en sus versiones experimentales o históricas (es decir, en la búsqueda de vida artificial o en las diferentes teorías del surgimiento de la vida como la conocemos actualmente). Antes de tratar de fabricar o estudiar cualquier cosa necesitamos una base conceptual sólida para saber cuándo hemos tenido éxito, o cuál es el punto de inicio en el que debemos basar nuestro estudio.

Además, en el caso particular de la vida existe otro factor: la búsqueda de vida en otros lugares. Considerado de cierta manera, es la razón menos práctica y más lejana a nuestras posibilidades. Resulta más fácil la búsqueda de vida inteligente productora de tecnología: las perturbaciones que causarían en el espacio (pensemos en cualquier tipo de ondas electromagnéticas con patrones repetitivos) se propagan mucho más rápido y a mayor distancia que los efectos fisicoquímicos de la vida en sí. Una definición como las propuestas hasta ahora, incluyendo el presente trabajo, implican una observación detallada del objeto; en algunos casos, como se verá, esta observación necesita prolongarse indeterminadamente. Sin embargo, es importante considerarlo; para mí, más que una posibilidad realizable a corto plazo, resulta un recurso para realizar experimentos mentales y enfocar, desde perspectivas inusuales, la solidez de los conceptos que se proponen. Descartar características que pueden ser únicas a la vida en nuestro planeta hasta ir dejando solamente las de inclusión inevitable nos acercan a la definición esencialista que se quiere lograr.

Problemas característicos

Hay pocas cosas tan difíciles de definir como el concepto “vida”. De cierta forma, antes de los problemas de *realizar* una definición, existen obstáculos previos que se añaden a los obstáculos implícitos del lenguaje. Un buen ejemplo para entender la problemática del lenguaje en este caso particular es el origen de la formalización llevada a cabo por David Hilbert en 1920. Este caso es presentado en el libro *Life Itself* (Rosen 1991) para ilustrar la dualidad de la ciencia en su acercamiento cualitativo o cuantitativo, pero resulta análogo a problemas más generales y usuales del lenguaje. Con el creciente número de paradojas en los sistemas tradicionales matemáticos, Hilbert afirmó que el problema estaba en la semántica del lenguaje, que irrumpía en un mundo que (según su ideal) debería ser completamente sintáctico. En la geometría, por ejemplo, existen palabras que se habían utilizado en los axiomas que tenían referentes fuera del sistema de la geometría. “Triángulo” o “círculo”, además de ser una secuencia de letras con relaciones entre sí, denotan objetos en un mundo externo al de la proposición. Esto quiere decir que si utilizamos la palabra “círculo” en un axioma, éste no puede contener a su referente. Por lo tanto, no importa cuán complejo sea un sistema formal: si contiene referentes externos nunca podrá ser completo.

Este ejemplo trata de ilustrar los problemas que tiene el lenguaje a muchos niveles. De hecho presenta suficientes problemas aparentemente irresolubles si se lo maneja como un sistema puramente sintáctico; al elevar la perspectiva y considerarlo como una construcción social el panorama se vuelve terriblemente complejo. El problema de Hilbert no se originaba al tratar de definir “triángulo”, o tratar de fundamentar una serie de reglas para la descripción y uso de este fragmento del lenguaje; tenía raíz en el lenguaje mismo, y el problema del lenguaje se proyectaba a través de esa palabra, añadiéndose a los problemas de la palabra misma. De la misma manera, todos los problemas que presenta la palabra “vida” no se limitan a nuestra ignorancia de su funcionamiento, origen y estructura. Algunos anteceden a este enfoque, y son problemas que están relacionados con el uso que hacemos de la palabra, las percepciones e intuiciones que tenemos del fenómeno y de los referentes que son utilizados. Antes, pues, de empezar a analizar diversas definiciones de vida y sus características principales (y sobre todo antes de intentar proponer una definición nueva), tenemos que considerar los problemas del lenguaje en este caso particular. Si la

palabra “triángulo” presenta problemas debido a sus referentes externos, la palabra “vida” es un cúmulo gigante de implicaciones e ideas preformadas, muchas veces contradictorias, cuyo origen se pierde en el pasado.

Tratar de describir la vida es una empresa particularmente delicada, considerando que es uno de los pocos conceptos que se refieren a nosotros sin dar importancia a las características que inconscientemente consideramos como propias. El ser humano tiene que sufrir un proceso de generalización para ser agrupado con el resto de los seres vivos, sin importar (en un primer acercamiento) nuestra capacidad para razonar o abstraer, nuestras sociedades complejas o las prácticas que de una u otra manera nos ponen aparte. Esta es, creo, una de las razones por las que casi nunca nos pensamos como referentes de la palabra vida. La otra razón, de naturaleza distinta pero probablemente relacionada, es el tipo de personas a quienes le interesa una definición de vida: la mayoría de la investigación hecha en este sentido es realizada por la comunidad que se dedica a tratar de dilucidar el origen de la vida. Por lógica (y por restricciones termodinámicas y químicas, entre otras), uno no puede proponer que la vida se cristalizó espontáneamente en un metazoario o en una gimnosperma. Aunque es posible obtener un orden complejo a partir de un sistema masivamente desordenado, lo más probable es que el camino hacia la vida compleja como la conocemos hoy comenzó en un sistema comparativamente simple. El concepto de vida mínima es popular en la investigación de origen de la vida: idealmente, provee simultáneamente de un punto de apoyo y de un objetivo hacia el cual dirigirse.

Además, la palabra vida, utilizada más o menos coloquialmente, implica conceptos que nada tienen que ver con el uso que se le da en este trabajo. De manera relativamente relacionada, “vida” se refiere a nuestra experiencia particular y subjetiva. En términos generales podemos decir que nuestra presencia en el mundo y nuestra percepción de él comienzan y acaban abruptamente. Es así como la palabra vida se relaciona con la existencia cotidiana e interior. Partiendo de este punto, se generan expresiones como vida espiritual, vida media y muchas otras, que se refieren a una serie de actividades o periodos de tiempo. Es importante distinguir este tipo de expresiones como metáforas.

Biológicamente, y en el sentido más general, la palabra vida implica una condición que separa a la materia orgánica de la inorgánica. Esta especie de definición es tautológica (vida = materia orgánica; materia orgánica = materia viva) y con ella sólo

pretendo delimitar el campo de trabajo en cuanto a lenguaje; esta definición, si fue utilizada alguna vez, está ya en desuso. Aunque parece ser lo suficientemente clara, al menos con respecto a la manera de utilizar la palabra, la descripción da pie a ambigüedades que es importante resolver. En principio, la dificultad más importante se relaciona con la individualidad que implica el concepto. Vida, en el sentido biológico, puede entenderse como 1) un organismo individual, lo cual a su vez acarrea una serie de objeciones y ambigüedades que se tratarán más adelante; 2) El conjunto de dichos organismos, siendo así la vida entendida como los efectos y consecuencias de las interacciones; o 3) El conjunto de dichos organismos en un contexto temporal, visión que generalmente se enfoca en la reproducción (y consecuentemente, a la evolución). Las definiciones de vida, inevitablemente, le dan más importancia a uno de los tres sentidos. Más que un sesgo en la manera de explicar, cada uno representa una manera distinta de entender el fenómeno. Al confundir o mezclar voluntariamente estas dos posturas surgen proposiciones sin fundamentos, como las teorías de Gaia en su forma más radical.

La diferencia reside en la dicotomía de concebir a la vida como un proceso (tanto espacial como temporal), que no tiene representación en el mundo material sino que existe únicamente como una serie de eventos en el tiempo, o concebirla como el sistema que lleva a cabo tal proceso. Aunque tales posturas son válidas, este trabajo se enfoca en el entendimiento de los seres vivos como entes físicos: me parece que es la postura más útil para lograr dos objetivos, i.e. diagnosticar materia viva con parámetros evaluables en un punto y momento determinado, y fijar *estudio*¹ del origen de la vida en un punto del tiempo mediante la aparición de las características propias de la vida. La vida como proceso es simultáneamente causa y consecuencia de varios eventos a distintos niveles, desde molecular hasta mundial. Sin embargo, el objetivo de este trabajo no es una definición a nivel histórico. La evolución es una serie de eventos fundamental para el surgimiento de la vida y su devenir hasta la forma en que la percibimos hoy, pero la definición aquí trabajada no necesitará incluir cada ancestro para ser funcional. De la misma manera, un organismo no puede sobrevivir sin la interacción que lleva a cabo con innumerables elementos en su medio, pero la definición no se basará en las interacciones. En todo caso, tal vez sea necesario moverse en

¹ Enfatizo “estudio” pues señalar a un punto cualquiera en el tiempo como “el origen de la vida” probablemente es artificial. Como se verá, aún dentro de los parámetros menos ambiguos puede existir un gradiente de posibilidades.

parámetros de complejidad, en vez de hacerlo en escalas temporales y espaciales. Es decir, el objetivo de este trabajo es presentar argumentos que indiquen que tanto una célula de un organismo como el organismo en sí están vivos; existen en contextos muy distintos, y sin embargo logran establecer cierto número de propiedades comunes y características.

Todas estas consideraciones se refieren a un mismo problema: el tratar de definir un término que no sólo tiene una implicación diferente en el contexto social en el que se use, sino que incluso dentro de un grupo, en este caso la comunidad científica, la palabra parece tener un significado distinto para cada grupo de investigación. Por supuesto, esto es sólo un reflejo del problema al que apunta Ferdinand de Saussure en su Curso de Lingüística (Saussure 1993): la lingüística, hasta ese momento, se había engañado en cierta forma al pretender estudiar la prescripción (es decir, la estipulación del conjunto de reglas que determinan la estructura de un lenguaje) y no la descripción (que se refiere al uso “real” de las palabras, en el contexto de la praxis). Como es obvio, esta reevaluación de los objetivos de la investigación lingüística no creó la diferencia entre el “ideal” y el gran grupo de variaciones en la práctica: simplemente la hizo evidente.

La innovación de Saussure es a la vez analogía y causa (una de las causas, al menos) del problema que nos ocupa. Cada miembro de la comunidad científica ofrece una definición distinta de vida porque ésta está sesgada a su práctica habitual. Es inevitable que Robert Shapiro, crítico incansable del mundo del RNA, prefiera definiciones que den importancia a cuestiones metabólicas y energéticas de la célula (punto de vista expuesto extensamente en Shapiro, 2007), mientras que Stuart Kauffman haga énfasis en el surgimiento de estructuras complejas². Por su parte, Francis Crick (Crick 1981) y Leslie Orgel³ prefieren enfatizar la replicación (concepto íntimamente relacionado con la estructura del DNA) y la selección natural. Emmeche (*Life is a computational phenomenon*) y Kull (*Life is a self-reading text*), investigadores de biosemiótica, dan definiciones sesgadas al manejo de información de la vida. No importa que las definiciones sirvan a intereses personales; de hecho es una característica fundamental del discurso individual. Sin embargo, vemos que la preferencia de cada

² “Life is an expected, collectively self organized property of catalytic polymers” (Kauffman 1993). Otros investigadores que apoyan esta corriente son Freeman Dyson y Christian De Duve.

³ “Living things are CITROENS: Complex information transforming reproducing objects that evolve by natural selection” (las definiciones de Orgel y Emmeche, junto con muchas otras, son citadas por Barbieri 2003)

científico está marcada por sus opiniones acerca del origen de la vida. Posteriormente se desarrollarán las ramificaciones e implicaciones de esta correlación, por lo demás un poco obvia.

Problemas actuales.

“Vida es un sistema capaz de evolución por selección natural”: tal es la definición casi minimalista que da Carl Sagan en 1970 (citado en Ruiz-Mirazo 2004). Es el ejemplo más claro del sesgo que existe entre los intentos de definición. La teoría de la selección natural, junto con sus transformaciones y actualizaciones (como las distintas etapas de la Síntesis Moderna), y la elucidación de la estructura y funcionamiento del DNA parecen indicar propiedades esenciales de la vida. De todos los objetos que conocemos, los únicos que presentan ambas características (evolución darwiniana e información codificada en DNA) son los seres vivos. Sin embargo, ¿es ésta una razón suficiente para afirmar que es la única opción para definir la vida?

Ciertamente no. Aunque la mayoría de las definiciones propuestas hasta ahora (y prácticamente la totalidad de las definiciones tomadas realmente en serio) siguen esos criterios, hay dos puntos importantes que indican que los resultados están lejos de ser completamente satisfactorios. Primero, ninguna de las definiciones ha podido evitar inconvenientes contradicciones y excepciones bastante escandalosas. Éste no es el lugar para abundar en los problemas wittgensteinianos del lenguaje: es posible iniciar un cuestionamiento *ad infinitum* a cada definición y encontrar fallas, cada vez más sutiles pero inevitables. Segundo, aún sin tomar en cuenta las definiciones que entienden a la palabra “vida” como el estado de ánimo, la conciencia, la planeación y elaboración de objetivos personales o una mezcla de este tipo de conceptos, las definiciones publicadas en el medio científico rara vez son compatibles y en ocasiones parecen hablar de cosas distintas; sin embargo, nadie puede descalificar o validar un solo punto de vista.

Es vital analizar el segundo punto. Creo que no existe un consenso acerca de la naturaleza del objeto “vida”; es decir, en la literatura se encuentran definiciones que abarcan un espectro enorme del universo físico. Las definiciones que ven a la vida como una estructura material se enfocan sobre todo en las propiedades de intercambio de energía y de organización física, a veces haciendo mención a propiedades termodinámicas; también existen definiciones que tratan a la vida como una propiedad

de la materia, lo cual es suficientemente ambiguo como para que sea aplicable a cualquier nivel de la jerarquía biológica, desde los cuestionables individuos hasta el planeta entero. Existen definiciones que se basan en la vida como un proceso, lo cual tiene implícito un acercamiento temporal muy claro. Muy aparte se encuentran las definiciones de vida que se basan en propiedades más bien metafísicas; entre las más importantes de éstas están las contribuciones del campo de la biosemántica, que se acerca al problema de la vida desde el análisis de las propiedades informacionales de la información genética. Por supuesto, muchas definiciones propuestas utilizan más de un punto de vista de los citados. Aunque es válido hacerlo, el mezclar distintos tipos de características puede llevar a callejones sin salida, conceptualmente. Por ejemplo, las definiciones que intentan integrar la vida como proceso (en general desde el punto de vista evolutivo) y la vida como estructura muchas veces se encuentran en problemas para llevar a buen término la unión entre la descripción física y la descripción temporal. Generalmente uno puede notarlo cuando se justifica el salto conceptual haciendo alusión a un sistema físico que “en potencia” es capaz de realizar un número de actividades. El argumento va más o menos así: aunque en su tiempo de existencia un organismo no lleve a cabo ciertas funciones necesarias para definirlo como ser vivo, es suficiente que nosotros sepamos que las *pudo haber hecho* para considerarlo como viviente. Este tipo de contradicciones de abordarán con más detalle más adelante; baste decir, por ahora, que la potencialidad no puede ser tomada como un parámetro para una definición congruente, y mucho menos para una definición funcional, que es uno de los objetivos de este trabajo.

Es imposible decir cuál acercamiento es correcto. Todos han surgido y han sido defendidos principalmente por ser aparentes. Por ejemplo: observamos que los seres vivos se reproducen de una manera particular. Prácticamente cualquier cosa en el universo es reproducible, más o menos fielmente. Existen muchas cosas que pueden autorreproducirse, como el fuego (una llama puede dividirse por sus propios medios en dos prácticamente idénticas). Sin embargo, la reproducción de los seres vivos conlleva una serie de eventos que poco tienen que ver con la aparente duplicación de una llama. En general, la reproducción de la vida es tan especial porque se trata de un proceso dirigido a organizar una serie de elementos producidos *de novo*, anabólica e internamente, hasta lograr un duplicado, con posibles variaciones, del sistema complejo original. La reproducción de los seres vivos es un proceso exclusivo de éstos, y es

reconocible en todas sus variantes (sean asexuales o sexuales); por tanto, parece lógico utilizar este criterio como uno de los puntos que definen a la vida. Y es así con cada parámetro propuesto para la definición: la evolución darwiniana, la información genética, el metabolismo, incluso la muerte. Sin embargo, ¿esta manera de pensar es realmente válida, o es una trampa lógica, un *non sequitur*? La reproducción llevada a cabo de esa manera es una característica particular de toda la vida que conocemos, en un nivel o en otro. Dejemos a un lado, por ahora, las contradicciones que implica la frase “todos los seres vivos se reproducen”, pues sabemos por experiencia que no es así. Sin tomar en cuenta los problemas conceptuales y posibles tautologías que yacen detrás de esa afirmación aparentemente sencilla, es importante preguntarse si realmente una característica presente en todos los miembros del conjunto es necesariamente un parámetro de definición de éste.

Tomemos a la definición como una serie de características necesarias y suficientes para, en primer lugar, delimitar un conjunto. Los elementos del conjunto, pues, constituyen el *definiendum* y la descripción del conjunto es el *definiens*. No es difícil pensar en un conjunto en el que todos los elementos tengan cosas en común que no estén estipuladas en la definición. Consideremos, por ejemplo, a los libros de una colección editorial cualquiera. Todos tienen un formato específico de colores; sin embargo, ni el diseño ni la distribución de colores son parte del *definiens* del conjunto, aunque sean una característica común. Más aún, la editorial puede cambiar de formato de colores y los nuevos libros seguirán siendo parte de la colección. El formato de colores común responde a una razón (la apariencia unificada de los productos de una misma colección) y al mismo tiempo hace aparente no es suficiente que una característica sea común para ser necesaria como parte de la definición.

Otro ejemplo, un poco distinto; pensemos en el conjunto de los números que se pueden formar con los signos [1], [4], [7], un conjunto claramente infinito. Todos los números incluidos en el conjunto se representarán, al menos en el sistema arábigo de escritura numérica, con signos contruidos únicamente con líneas rectas. Una vez más, los elementos del conjunto presentan una característica común; sin embargo, esta característica no obedece a ninguna razón: está sujeta a la arbitrariedad del signo numérico (Saussure 1993). Los valores de los números expresados y las posibles relaciones entre ellos no están conectados, ni siquiera indirectamente, con las características gráficas de los signos que los representan.

Este tipo de problemas se han considerado pocas veces, sea por descuido o por falta de información: es sabido que los nuevos descubrimientos y teorías abren camino a nuevas perspectivas en el campo de las definiciones. Las definiciones de vida en general analizan cada parámetro evaluando si éste se aplica a todos los casos de “vida” conocidos; es necesario, desde ahora, no sólo ver si es una característica común, sino si es una característica necesaria para una definición coherente.

2. Las definiciones actuales de vida

Elementos comunes

Las definiciones de vida suelen tener varios elementos comunes. Esto, aunque difícilmente signifique un consenso, indica claramente ciertas propiedades evidentemente características. Por supuesto, ninguna definición es imparcial: cualquier esfuerzo está modificado *a priori* por los intereses de quien lo realiza. No es raro encontrar una correlación entre la definición y el área de investigación de quien la propone, o al menos su escuela de pensamiento, si es que es posible hablar de ellas. Esto, a su vez, refleja la posición que toma cada científico en su programa de búsqueda de origen de la vida. Las divergencias que existen en la manera de acercarse al problema de la vida (i.e. su origen y naturaleza) son analizadas en la literatura, llegando al punto de dividir a la comunidad científica en dos bandos, dependiendo del parámetro utilizado (Pereto, 2005, Anet 2004). Sin embargo, es importante aclarar que este fenómeno es inevitable pero de ninguna manera negativo; la diversidad de puntos de vista enriquece la discusión aportando información que una visión unilateral nunca podría utilizar.

Los dos elementos más recurrentes son el metabolismo y la replicación: lo que Ruiz-Mirazo llama “los dos aspectos fundamentales de la fenomenología biológica. (Ruiz-Mirazo, 2004). Curiosamente, cada uno representa un extremo del espectro de las definiciones de los seres vivos: la replicación implica una vida comunal, en la que un individuo sólo está vivo en función de su capacidad reproductiva y hereditaria; el metabolismo implica un sistema cerrado sobre sí mismo, en un punto espaciotemporal determinado⁴. Esta polarización tiene, hasta cierto punto, las mismas caras de la

⁴ Una tercera característica recurrente es la presencia de barreras (“boundaries”) que delimitan el sistema vivo de su medio; sin embargo creo pertinente incluir esta característica dentro de la visión metabólica al describir integralmente a un sistema desde el punto de vista energético. El sistema debe de estar

dicotomización de Weismann (Weismann 1893) cuando propone las dos unidades fundamentales de la vida, el plasma germinal y el somatoplasma. Las caracterizaciones que hago de cada posición, claro, son radicales, pero es evidente que representan a los acercamientos a la vida de la biología molecular y la bioquímica, respectivamente.

Existe un tercer grupo de características, completamente heterogéneo, que aparecen con relativa frecuencia en las definiciones, sobre todo cuando los autores proponen una lista de características que tienen que ser evaluadas una a una (típicamente una *checklist*). Éstas resultan ser descripciones poco prácticas de alguna característica de la vida, y muchas veces tienden a ser redundantes o a incluir elementos que se pueden agrupar bajo un concepto más general; sus problemas serán analizados y ejemplificados más adelante. Sin embargo, en éste grupo se incluyen los puntos de vista relativamente recientes derivados de la biosemiótica (el campo del conocimiento que estudia las propiedades y orígenes de los símbolos en los seres vivos), que tienen gran potencial de innovación y experimentación. No es el único ejemplo: las recientes propuestas acerca del surgimiento espontáneo de orden en sistemas al límite del caos, y de niveles de complejidad, entre otras, esclarecen ciertas propiedades fundamentales de los organismos. Una mirada detallada a ambos grupos revelará que están íntimamente relacionados con los tradicionales de replicación y metabolismo, respectivamente, dándoles sin embargo un enfoque nuevo que ha reportado un gran potencial explicativo y predictivo, como se verá más adelante.

Características de una definición

Para construir una definición viable, no es suficiente mencionar al metabolismo y a la reproducción de los seres vivos; incluso puede decirse que no es *necesario*. Esto depende, por supuesto, del punto de vista que se tome al acercarse al problema. Por eso ahora es pertinente establecer ciertos principios básicos de este trabajo: qué parámetros se toman para construir una definición; qué características debe de tener ésta; qué acercamiento se le dará a la palabra Vida, y cuáles no se manejarán en el escrito.

Ruiz-Mirazo (Ruiz Mirazo 2004) y sus colaboradores establecen 5 puntos básicos que la definición de vida debe de cumplir; ésta debe de:

1. Ser coherente con los conocimientos actuales de física, química y biología.

delimitado y sus barreras deben de permitir el paso de materia y energía para lograr el tipo de estabilidad lejos del equilibrio que los sistemas vivos presentan.

2. Ser consistente consigo misma y no redundante
3. Poseer poder explicatorio (i.e. debe de arrojar luz sobre la naturaleza de la vida y los procesos que la originaron) y elegancia conceptual.
4. Ser universal
5. Ser suficiente y mínima, pero específica.

La mayoría de estos puntos son necesarios para cualquier definición, y la lista es una guía invaluable de la solidez de una definición. Sin embargo, en muchas ocasiones esta lista de requisitos puede servir para descalificar una definición viable. La “elegancia conceptual”, por ejemplo, es un parámetro bastante subjetivo. Es importante señalar que el poder de explicación de la definición es, precisamente, lo que permite establecer alguna manera (directa o indirecta) de comprobación experimental, y lo que permite relacionar directamente una definición funcional de la vida con un modelo para explicar su origen.

Ya se han mencionado a las listas de características que funcionan como definiciones. Se presentan algunos problemas en estas definiciones descriptivas, generalmente redundancias, inclusión de conceptos no necesarios, problemas de excesiva generalidad o especificidad (o ambas simultáneamente, por paradójico que sea).⁵ Este tipo de definiciones son las analizadas por Ludwig Wittgenstein, con su famoso ejemplo del término “Juego”, que ya se ha mencionado. Es frecuente que las construcciones abstractas de la mente humana necesiten definiciones así; en el uso frecuente del lenguaje juzgamos como correcto o incorrecto el uso de un término aún sin saber definirlo con exactitud. Lo que se realiza en estos casos es un tipo particular de definición ostensiva, la llamada definición politética. Se piensan o señalan ejemplos del término a definir (en el caso de Wittgenstein, “juegos”) y se encuentran los “parecidos familiares”. (Wittgenstein, 1958, párrafo 67). Lingüísticamente pueden ser aceptables, pero en una definición de vida los problemas son demasiados.

Por ejemplo, Koshland (Koshland, 2002) propone siete puntos (“pilares”) para definir a la vida. Incluye dos, Improvisación y Adaptabilidad, y la frontera entre ambas (si es que la hay) es bastante difusa. En efecto, el autor da una dudosa explicación:

“Improvisation is a form of adaptability, but it is too slow for many of the environmental hazards (...) Adaptability could arguably include improvisation (pillar number 2), but improvisation is a mechanism to change the fundamental

⁵ Considérese, por ejemplo, la definición, “un escritorio es un mueble sobre el cual se trabaja”.

program, whereas adaptability (pillar number 6) is a behavioral response that is part of the program.”

Inmediatamente se hacen aparentes varias incógnitas. ¿Acaso los mecanismos de cambio no están implícitos en el “programa”? ¿Cómo se define una respuesta de comportamiento? ¿No es el comportamiento en muchas instancias (si no es que siempre) un punto de partida para crear una diferencia de supervivencia y reproducción (es decir, un mecanismo de “cambio de programa”)? Por otro lado, “Programa” (que dicho sea de paso aparece también como el pilar número uno) parecería referirse a la información genética que posee cada organismo; en efecto, el concepto es definido como el plan organizado de los ingredientes y cinéticas de los procesos implementados por el DNA. Esta definición mezcla los conceptos de metabolismo y genoma, pero parece dar prioridad a este último. Sin embargo, bajo el apartado de Improvisación, Koshland escribe:

“If, for example, a warm period changes to an ice age so that the program is less effective, the system will need to change its program to survive.”

Tomada en conjunto, esta serie de razonamientos ciertamente parece lamarkiana, por la simple razón de que el autor mezcla conceptos que se refieren a la vida como una comunidad con otros que la tratan como un individuo discreto en un tiempo y espacio determinado (como la Compartimentalización, el pilar número cuatro). Como se puede ver, el uso de conceptos ambiguos da lugar a repeticiones, contradicciones y lagunas conceptuales que hacen que la definición sea inutilizable.

Existen otros “pilares” en su escrito que presentan problemas de distinta índole. Sin embargo, este trabajo no es el lugar para discutir las diferentes contradicciones, contraejemplos e implicaciones que contiene este acercamiento a la descripción de vida. Baste con que sirva para ejemplificar lo peligrosas que pueden ser las *checklist*, y por qué no parecen tener la elegancia de las definiciones formuladas tradicionalmente. Las características de la vida se presentan en cada aspecto de su existencia, conformándola como un fenómeno muy particular, pero, como ya se ha dicho, la simple enumeración de dichas características puede incluir alguna repetitiva o no fundamental.

La definición buscada aquí, pues, no puede ser ni politética ni descriptiva (en el sentido de las *checklist*); lo que se intenta lograr en este trabajo es una definición esencialista. Por supuesto, esto no quiere decir que se trate de enunciar lo que Aristóteles llamaba *to ti en einai*, la sustancia que constituía la naturaleza de la cosa separada de los accidentes (Aristóteles, 2004); en tal sentido, este trabajo no trata de encontrar la llamada *vis vitalis*. Esencialista, en este contexto, quiere decir intentar llegar a la comprensión de la teoría que yace detrás de la definición: una teoría de lo que la cosa necesita ser para ser. Un excelente ejemplo es la definición de Maturana y Varela (Maturana, 1994). “Los sistemas vivos son máquinas autopoieticas” representa una teoría condensada en una frase. Incluso la palabra “autopoiesis” es un neologismo acuñado por Maturana para agrupar una serie de características en un fenómeno común. Además, implica una visión mecanicista de la vida (i.e., los seres vivos son máquinas). En una definición de este tipo no es suficiente enunciar; para que tenga validez se tiene que explicar la teoría que subyace a la manera de explicar el término.

Reproducción

Una revisión rápida de la literatura acerca de la naturaleza de la vida revela una tendencia fundamental: considerar a la reproducción como una de las características esenciales de la vida. En efecto, como ya se mencionó arriba, es uno de dos aspectos fundamentales del estudio biológico (el otro siendo el metabolismo). La vida como la conocemos es inconcebible sin esa autopropagación que se presenta, sin excepción aparente, en todos los grupos. No es inusual encontrar una identificación de la reproducción (o replicación, pues en ocasiones se usa de manera indistinta) como parámetro único que define a la vida. Addy Pross (Pross, 2003) afirma que en un sentido general y químico, “*life is a manifestation of replication*”.

Antes de considerar las consecuencias de tal criterio, es pertinente considerar el objeto y naturaleza del fenómeno reproductivo, que puede ser llevado a cabo en varios niveles biológicos. El concepto de reproducción no es simplemente una multiplicación numérica de algún elemento: está ligado íntimamente al concepto de *autopropagación*. Esto quiere decir que un objeto cualquiera da lugar a copias de sí mismo. En ese sentido, encontramos este proceso en diversas manifestaciones biológicas:

1. El sentido más mínimo y el más utilizado en el estudio del origen de la vida es el molecular. En este caso, el concepto se refiere casi exclusivamente a la multiplicación de ácidos nucleicos, en parte porque son las únicas moléculas que conocemos que son capaces de copiarse a sí mismas.⁶
2. Por otra parte, está la reproducción a nivel celular; en este sentido, el término se relaciona meramente con la progresión de la célula (la unidad fundamental de la vida) a través del llamado ciclo celular. A grandes rasgos, éste comienza con la producción de elementos necesarios para la duplicación y culmina con el evento de la división celular⁷. También es una característica presente en todas las células vivas, sin importar su tipo.
3. Finalmente, la reproducción se considera en ocasiones como un evento a nivel de organismo. En el caso de los organismos unicelulares, la reproducción del individuo completo se puede igualar a la celular. Sin embargo, si tomamos en cuenta organismos multicelulares, el proceso (sea asexual o sexual) es mucho más complejo, y, estrictamente, sólo una porción del individuo será la parte replicante. En este nivel reproductivo, se hace evidente la dicotomía entre el germoplasma y el somatoplasma, propuesta por Weissman. Esto se analizará más adelante, pues tal idea ha tenido y sigue teniendo repercusiones profundas.

A primera vista, la reproducción parece ser un evento inevitable de la vida. Sin embargo, la inclusión de este criterio simple trae consigo una serie de problemas. El más evidente es que no resulta suficiente para definir inequívocamente a los seres vivos; la reproducción se lleva a cabo en distintas modalidades en todo el universo físico. El fuego y los cristales, como ya han considerado muchos autores, parecen reproducirse: una llama puede dar lugar a otras llamas independientes; un fragmento de cristal puede provocar la aparición de cristales en una solución saturada. Tales ejemplos

Intuitivamente consideramos que la reproducción del fuego es distinta a la de los seres vivos. Tales ejemplos no sólo sirven para mostrar cómo es que objetos claramente inertes cumplen con alguna característica propuesta sino que dichas características son compartidas por sistemas que nada tienen que ver con los sistemas vivos. Es decir: un

⁶ En teoría solamente. Aparentemente, es difícil que una molécula de RNA sirva como un templado o molde de manera espontánea, es decir, sin ayuda de una enzima protéica de actividad específica.

⁷ El caso de la meiosis, una división celular especializada de surgimiento posterior, es un evento relacionado con la reproducción sexual, considerada en el siguiente punto.

crystal y un ser vivo sólo comparten sus características superficialmente; en niveles más profundos (p.e., termodinámicamente) no tienen ninguna similaridad. La formación de un cristal aumenta el orden en una zona de un sistema; sin embargo el sistema aumenta su entropía como un todo. Por otro lado, los seres vivos consumen entropía negativa, como lo ha postulado Schroedinger, elevando en la entropía de su entorno; esto implica, como se verá más adelante, que los procesos termodinámicos de la vida se alejan constantemente del equilibrio. Un cristal es un objeto en un equilibrio inerte. Así pues, las características que parecen compartir son en realidad derivadas de dos *fenómenos causales distintos*.

Por otro lado, los virus o las partículas subvirales (tales como los viroides, los virus satélites) presentan un tipo de replicación idéntica o análoga a la replicación del DNA celular. Los virus, por ejemplo, contienen la información genética para cada una de sus partes e indudablemente se reproducen. Sin embargo, su inclusión dentro de los sistemas vivos es controversial pues no pueden reproducirse con medios propios: necesitan del aparato de replicación y traducción de la célula hospedera. Esta única pero fundamental diferencia es la característica que excluye a los virus del conjunto de la vida en la visión de la mayoría de los autores que han tocado este punto. Se puede observar asimismo que un organismo de reproducción exclusivamente sexual no puede reproducirse por sus propios medios: siempre necesita a otro individuo para llevar a cabo una fecundación y por tanto el proceso de reproducción⁸.

Un problema más ante este criterio es el paso de la abstracción conceptual hacia la fenomenología individual: muchos seres vivos no se reproducen nunca. Por ejemplo, un ser vivo completamente fértil puede morir sin dejar descendencia debido a una serie de contingencias (incluso por volición propia, en el caso de los seres humanos). ¿Debemos de considerarlo como un objeto no viviente? De la misma manera, no es inusual que un individuo sea estéril, ya sea desde su nacimiento o por entrar en una fase no reproductiva de su ciclo de vida. ¿Quiere esto decir que las mujeres en edad postmenopáusica o los híbridos infértiles no están vivos? También es interesante considerar el caso de la parte estéril de cada individuo vivo, i.e. las células somáticas. Tales células pueden multiplicarse, y sin embargo no son capaces de reproducir al individuo

⁸ Como me ha señalado el Dr. Víctor Valdés, los priones representan una excepción en este párrafo. Parecen reproducirse pero lo hacen modificando proteínas "normales" ya existentes; mi imposibilidad para decidir si tal evento es una replicación "real" tal vez se derive del sesgo implícito que tenemos los seres humanos, pues no conocemos otros sistemas indudablemente vivos además de los terrestres, y todos ellos llevan a cabo una replicación de información a partir de la contenida en los ácidos nucleicos.

completo, cosa que sólo pueden llevar a cabo las células germinales. Es casi imposible encontrar una posición que sea conciliadora y concreta.

Estos ejemplos son sólo eso: ejemplos de una serie de problemas presentados por el concepto de reproducción. No pretendo debatir cada uno, y proponer una serie de soluciones lógicas. Sólo trato de señalar la ambigüedad y arbitrariedad del parámetro. Algo con tal carga de subjetividad difícilmente puede ser tomado en cuenta como un criterio válido. Las dificultades no son superficiales, ni son causadas por excepciones despreciables: tienen necesariamente una raíz más profunda. Tal vez sea por utilizar un término demasiado laxo, o no delimitado hasta sus últimas consecuencias (pues la reproducción se lleva a cabo en muchos niveles, desde molecular hasta orgánico, y el término en sí no aclara específicamente a qué se refiere). Posiblemente se simplifica demasiado un evento inherentemente complejo, la vida. Incluso puede ser que, como afirman ciertos detractores del estudio de la vida en sí, aún no tenemos las armas apropiadas, sea en el campo conceptual o en el tecnológico. Sin embargo se puede decir que el uso del término reproducción es inadecuado. También se rechaza el método de corrección del término, agregando condiciones que permitan la exclusión o inclusión de excepciones (por ejemplo, afirmar que la reproducción “debe de ser por los propios medios” para excluir a los virus, o que “no es necesario reproducirse sino haber tenido la potencia para hacerlo”, para incluir a los individuos en edad avanzada). Este añadir cláusulas a nuestra definición no sólo es insuficiente y carece de poder explicativo o predictivo: es una labor potencialmente infinita.

Darwinismo e Individualidad

La reproducción de los seres vivos está íntimamente ligada a un concepto fundamental en el estudio de la biología, el de la evolución darwiniana; para algunos, de hecho, el más importante. Una de las frases más famosas del gran genetista ruso Theodosius Dobzhansky es “Nada en biología tiene sentido si no es observado a la luz de la evolución”. Ya se ha citado anteriormente la definición de Sagan, que ejemplifica ese grupo de definiciones que otorgan el mayor peso conceptual al proceso evolutivo. Este tipo de pensamiento es, probablemente, uno de los más usuales en el intento de describir la naturaleza de la vida. En efecto, la definición operacional de la NASA dice que la vida es “*a self-sustained chemical system capable of undergoing Darwinian*

evolution” (Luisi, 1998). Antes de analizar este tipo de acercamiento, creo que es necesario considerar qué se entiende por “evolución darwiniana”.

Según Stephen Jay Gould (2002), la teoría de la selección natural, como fue planteada por Charles Darwin, implica tres conceptos fundamentales e innegables: las dinámicas de los seres vivos implican superfertilidad, heredabilidad y variación. Tomando estos tres hechos, se deduce el proceso: ciertos individuos con variaciones tendrán una probabilidad más alta de reproducción y supervivencia, y a su vez tales variaciones serán transmitidas a la descendencia. Después de la llamada Síntesis Moderna se comenzó a entender la selección natural a la luz de la genética, con trabajos tan importantes como los de T. Dobzhansky (1975). Al cambiar un genotipo, cambia a su vez el fenotipo, sobre los que actuará la selección. Los fenotipos más adecuados a su medio probablemente verán aumentado su número con el paso de las generaciones (debido a la reproducción y supervivencia mayores). Este resumen de la teoría darwiniana no es más que una simplificación extrema de la complejidad de los fenómenos evolutivos⁹.

Al observar el esquema básico de la selección natural, resulta difícil entender la crítica que la reduce a una tautología: “la teoría de la selección natural afirma que los más adecuados sobreviven porque son más adecuados”. El supuesto argumento circular es que la adecuación representa simultáneamente la causa y el efecto de la supervivencia (y, claro, viceversa). Sin embargo, la supuesta tautología no es más que una trampa del lenguaje. El origen de la tautología reside en el concepto de adecuación, que simultáneamente refuta esta crítica pero da lugar a otras de resolución más difícil. En cierto sentido, la adecuación no es más que un elemento de un modelo para explicar las características y la diversificación de organismos; no tiene una representación real y unidimensional en el universo físico. No es una característica que pueda ser medida de la misma manera que, por ejemplo, la altura de un árbol o la temperatura de un líquido; surge *a posteriori*, al cuantificar la fertilidad y la supervivencia de un organismo en un contexto determinado; es una magnitud relativa y no absoluta.¹⁰ Es necesario notar que desde este punto de vista, la adecuación es un parámetro con función meramente epistemológica; es un concepto de segundo orden, derivado de los conceptos

⁹ Han surgido interpretaciones “ultragenéticas” como las propuestas por Richard Dawkins (Dawkins 1989), en el que apartentemente se puede reducir la esencia de la vida a los genes, y el “somatoplasma”, como le llamaría Weismann, es sólo una “máquina de supervivencia” de éstos.

¹⁰ Futuyma (Futuyma 2005) define tal parámetro (*fitness*) como el éxito reproductivo de una entidad; por lo tanto, es el promedio de la contribución de un alelo o un genotipo a las siguientes generaciones.

cuantificables de reproducción y supervivencia. Su carácter de abstracción invalida la crítica de la tautología: la adecuación no causa ni hace surgir a la supervivencia. Estos argumentos en contra de la selección, surgidos de la falibilidad normal del lenguaje, no serán tomados en cuenta en este trabajo. La adecuación surge como un parámetro cuantificable visto desde el punto de restricciones físicas y ecológicas, sobre todo en la ecología poblacional, trascendiendo la simple tautología de la sobrevivencia del sobreviviente.

Es innegable la realidad del proceso; las consecuencias de éste son visibles fácilmente e incluso susceptibles a la experimentación, como fue patente desde la publicación del libro *El Origen de las Especies* (Darwin 1968). La selección artificial es uno de los ejemplos más claros; su mecanismo puede observarse de manera acelerada gracias a la manipulación antropogénica de la reproducción: se escogen los caracteres de interés y los individuos que los presentan son los que se utilizan para las siguientes cruces. Lo que se cuestiona aquí es que esta particularidad podría resultar inadecuada para diagnosticar a la vida. Surgen varias inconsistencias al tratar de tomar a la evolución (darwiniana o no) como un parámetro fundamental de la vida; éstas se agravan, por supuesto, al tomarla como parámetro único.

El *caveat* más evidente reside en tomar la evolución como sinónimo de selección natural. La evolución es, en su forma más mínima, la aparición de un cambio heredable; de la manera más extensa, es el cambio en las características de los organismos a través de las generaciones. La selección natural, junto con otros procesos, se mueve en un mesocosmos: se enfoca en individuos y, particularmente, la producción de descendencia de cada uno, aunque sus resultados a largo plazo sólo sean visibles en poblaciones. Es difícil decir que la evolución es el resultado únicamente de la selección natural; es un proceso que ocurre en muchos niveles. Consideremos, por ejemplo, una bacteria, que representa una sola copia del genoma; en cuanto ocurre una mutación existe un cambio irreversible, cuyos efectos pueden ser visibles aún antes de que tal bacteria se reproduzca; es decir, la selección natural no tomó parte alguna en este caso de evolución. Además, varios autores han señalado la importancia de las restricciones físicas, independientemente de su postura ante el darwinismo (Gould y Lewontin, 1979; Goodwin, 1998). En trabajos recientes la importancia de las restricciones físicas ha ido en aumento: Smit, por ejemplo, reporta una serie de estructuras secundarias de el RNA ribosomal que se conservan independientemente de cambios evolutivos extensos entre

clados distantes, por lo cual los autores sugieren que dichas secuencias no fueron escogidas *per se* mediante un proceso de selección, sino que están correlacionadas a la estructura a gran escala del rRNA; secuencias aleatorias presentadas por los autores tienen el mismo sesgo hacia dichos motivos estructurales, a pesar de que no tienen ninguna historia evolutiva (Smit 2006)

Sin embargo existe un problema mayor: aquel de la individualidad de los organismos. Consideremos el caso de las células HeLa, una línea celular ampliamente usada en el laboratorio que es prácticamente inmortal y de reproducción acelerada. Fueron cultivadas de los tumores cervicales de Henrietta Lacks, que murió de cáncer el 4 de octubre de 1951. El problema de la individualidad es particularmente fuerte aquí; las células HeLa provinieron de Henrietta Lacks, pero es indudable que la sobrevivieron y que su reproducción ha sido mayor. Es difícil, entonces, evaluar si algún mecanismo de selección actuó en este caso. Es probable que el volumen de las células HeLa en la actualidad supere por mucho el volumen total de Henrietta cuando ella estaba viva. Además, aún cuando estaban en el cuerpo de Lacks, las células eran genéticamente distintas a Henrietta. Actualmente se ha propuesto (sin mucho éxito, pero con argumentos sólidos basados en la ambigüedad del concepto de individuo) que las células HeLa sean consideradas como otra especie, *Helacyton gartleri*. Tibor Gánti (Gánti 2003) argumenta que la unidad de un sistema es un criterio absoluto de vida¹¹, y puede ser diagnosticada por el surgimiento de un nuevo conjunto de propiedades, comparado con el conjunto de propiedades de sus partes aisladas. Esto no es necesariamente cierto: la idea de emergencia se relaciona más con la complejidad física que con la capacidad de discernir una unidad individual.

Existen otros tantos ejemplos de la dificultad de determinar exactamente en dónde reside la selección natural, qué parámetros se deben utilizar, cómo definir ciertos conceptos básicos para su entendimiento como adaptación, adecuación, etc. Sin embargo, su identificación con una característica necesaria y exclusiva de la vida (un lugar común de la biología) es inviable después de los argumentos derivados de los experimentos clásicos de Spiegelman.

En 1967 Sol Spiegelman (Spiegelman 1967, Kacian 1972) propuso una línea de investigación experimental de la teoría darwiniana. Durante su serie de trabajos con la

¹¹ Los criterios absolutos de Ganti son individualidad, capacidad metabólica, estabilidad inherente, regulación de procesos y manejo de información; los criterios potenciales son crecimiento y multiplicación, cambio hereditario, y mortalidad (es decir, capacidad de morir).

replicasa de RNA del fago Q β diseñó un experimento en el que agregaba el genoma ribonucleico del fago a una solución con la replicasa, nucleótidos libres y sales. Después de cierto tiempo, en el que la replicasa polimerizaba nuevas cadenas de RNA, Spiegelman tomó un poco del RNA producido y lo colocó nuevamente en una solución de replicasa. Eventualmente, las cadenas que se polimerizaban eran más y más cortas, hasta finalmente obtener una cadena de 218 nucleótidos que ahora es conocida como el “monstruo de Spiegelman”, cuya velocidad de replicación es alta en comparación con el genoma completo inicial, de aproximadamente 4500 bases. Usualmente se dice que tal experimento es el primer intento de lograr una secuencia de RNA autorreplicante, lo cual es erróneo. La única participación del ácido nucleico es como templado o molde para la replicación; el “monstruo de Spiegelman” no codifica para una proteína funcional. De hecho Oehlschläger y Eigen (Oehlschläger 1997) demostraron que se puede reducir aún más la parte replicada, llegando hasta 48 bases, que contenían únicamente el sitio de unión de la replicasa a la cadena de RNA.

El experimento de Spiegelman cobra gran importancia en este trabajo pues demuestra que el mecanismo de la selección natural no es exclusivo de los seres vivos. Las moléculas de RNA llevan a cabo un cambio gradual con base en la rapidez de la replicación; sin embargo, una parte importante del fenómeno de evolución de esta secuencia es que su semántica no toma ninguna parte en el proceso. Es decir, las instrucciones que codifica no tienen ninguna relevancia para la optimización del proceso de replicación. Esto es fundamentalmente distinto al comportamiento biológico de este tipo de moléculas, y es la razón por la que esta evolución no tiene el comportamiento dinámico que parecen tener los seres vivos: en palabras de S. J. Gould, los seres vivos presentan un “muro a la izquierda” en su gráfica de complejidad, que les impone una complejidad mínima; los RNA de Spiegelman parecen ser su opuesto y tener un “muro a la derecha” (Gould, 1996). Sin embargo, no es posible afirmar que este experimento se base en otro principio además de la “selección natural”: en este caso la producción de copias (o descendencia) diferencial de cada una de las moléculas.

Parece incongruente que un comportamiento considerado tradicionalmente exclusivo de los seres vivos pueda A) presentarse en un sistema completamente inerte y B) tener un resultado diametralmente distinto al comportamiento observado en los seres

vivos.¹² La razón es bastante simple: como afirma Goodwin (Goodwin 1998), “no hay nada especialmente biológico en la selección natural: no es más que un término usado por los biólogos para describir el modo en que una forma es reemplazada por otra como resultado de sus diferentes propiedades dinámicas”.

Esto es observable no sólo en fenómenos tan claramente conectados con la biología como la replicación de RNA. La estabilización dinámica (la forma generalizada de referirse a la “selección natural” es visible fácilmente en cualquier clase de fenómeno oscilatorio químico. En la conocida reacción de Belussov-Zhabotinsky se forman distintos patrones con las ondulaciones causadas por la reacción. Existen, por ejemplo, patrones de círculos concéntricos y patrones de espirales, surgidos por una ligera perturbación. Cuando aparecen los espirales, poco a poco se extienden y superan en número a los círculos concéntricos, pues la longitud de onda es más corta en los primeros (lo cual quiere decir que el tiempo entre onda y onda es más corto, y por lo tanto su formación en general es más rápida).

Todos los argumentos anteriores apuntan a una conclusión general: la selección natural no es una ley biológica; es simplemente un comportamiento estadístico, que puede presentarse en varias modalidades, en varios tipos de sistemas (no necesariamente vivos). Es verdad que todos los seres vivos lo comparten; sin embargo, también es probable que se deba a algún proceso, más fundamental, que subyace en el funcionamiento de la vida. Se ha discutido poco la posibilidad de que la selección natural, e incluso la evolución, sean sólo consecuencias inevitables de otras características, las cuales se tratan de dilucidar en este trabajo.¹³

Si negamos que la evolución por selección natural es una característica necesaria y exclusiva de los seres vivos, es inevitable preguntarse el papel de la historia en los sistemas biológicos; esta es sólo una manera de plantear la siguiente pregunta: ¿Tiene el tiempo cabida en la descripción necesaria de los seres vivos?

La experiencia humana universal es que el tiempo es una parte fundamental de la vida. Nuestra propia existencia parece moverse inevitablemente hacia adelante en la

¹² De hecho, se podría proponer la hipótesis que en un sistema biológico el equilibrio se alcance en un punto medio: una población heterogénea de moléculas de tamaño medio, que se repliquen rápidamente pero no lo suficientemente como para agotar los recursos. Esto hace eco en los estudios de Eigen y Schuster con los hiperciclos, la preocupación de que un RNA “egoísta” pueda desequilibrar el mundo del RNA conduciendo inevitablemente a una catástrofe de error.

¹³ De hecho, las teorías de Eigen y Schuster (Eigen y Schuster, 1977, 1978 ab), fundamentales para el estudio poblacional de ciclos moleculares autocatalíticos como modelos matemáticos, *requieren* que la vida no sea una condición necesaria para la selección.

línea que percibimos como “tiempo”. Sin embargo, esto no quiere decir necesariamente que la inclusión del factor tiempo en una descripción de la vida sea unilateral. Por ejemplo, la visión histórica (evolucionista) de la definición de vida es un intento por explicar la temporalidad inherente en los seres vivos. Sin embargo, es inútil definir a un ser vivo con base en su evolución; como todos los fenómenos del universo, proviene de una cadena de causalidad que en último término es irrelevante para su descripción: no es necesario recapitular la historia del origen del objeto definido para lograr una definición competente. Por supuesto, toda comprensión extensiva de los organismos requiere un enfoque multifacético, y la evolución juega un papel fundamental para entender los procesos que rigen a los organismos; el argumento que aquí se propone es que la temporalidad de la vida se debe de buscar en otra escala; análogamente, uno puede definir un sistema termodinámico sin considerar los eventos y procesos que dieron lugar a su estado inicial. Cuando se define “sistema vivo” de manera histórica, en realidad se está definiendo “vida” como un flujo de sistemas vivos que forman parte de un todo (y como se ve, esta definición no dice mucho). En efecto, es imposible diagnosticar la capacidad evolutiva o la historia de un ser vivo con información que ese ser vivo nos proporciona: es necesario evaluar a sus descendientes a lo largo de un número indefinido de generaciones. ¿Qué pasaría, por ejemplo, si ese sistema vivo sólo se reproduce una vez en un número enorme de años? Incluso en nuestro planeta, existen especies de plantas cuyo ciclo reproductivo rebasa el tiempo de una vida humana. Así, encontramos dos fallas fundamentales en el acercamiento histórico de la definición de la vida:

- Necesitaríamos poseer más información de la que nos proporciona el sistema vivo. Es decir, la información acerca de la definición tendría que provenir de *afuera* del sistema definido. Esto es inaceptable para nuestros fines.
- No es posible lograr una definición *hic et nunc*¹⁴ de un ser vivo: se tendría que esperar a que se diera un proceso evolutivo, observando al objeto en cuestión durante un lapso de tiempo indeterminado. Esto, más que una falla conceptual, es un problema práctico. Por supuesto, no es posible sesgar una definición por los alcances tecnológicos para analizar a los objetos definidos; sin embargo, es conveniente tener siempre en cuenta que una de las funciones de la definición de vida es diagnosticar si un sistema extraño está vivo o no.

¹⁴ Lat. “Aquí y ahora”

Por tales razones, considero que se debe evitar un acercamiento histórico (evolutivo) de la definición de vida es irrelevante. Sin embargo, es innegable que la vida es un sistema *temporal*. La razón más fundamental de esto es que los sistemas vivos funcionan con base en procesos termodinámicos. Considero que éste es el acercamiento pertinente a la naturaleza temporal de los seres vivos. Como en el estudio de cualquier proceso, se necesitan dos momentos *cualquiera* en el tiempo: un estado inicial del sistema, y un estado final que presentará cambios con respecto al primero. Esta pareja de momentos es la muestra mínima representativa para conocer la naturaleza y el comportamiento de cualquier sistema dinámico y lo dota de una descripción *temporal pero no histórica*. Por lo tanto, se debe de encuadrar a la descripción de los seres vivos dentro de un marco fisicoquímico. Esto, por supuesto, no es una afirmación original; fue escrita explícitamente dentro del libro *What is life?*, de Schrödinger.

3. Complejidad física e informacional

Consideraciones termodinámicas

Probablemente el trabajo que más ha influido en la investigación de la naturaleza de la vida como fenómeno fisicoquímico ha sido el trabajo de Erwin Schrödinger, *What is life?* (Schrödinger 2006¹⁵). En él, por primera vez se presentan los conceptos base en la comprensión física de los sistemas vivos. Dedujo un mecanismo por el cual la información puede ser transmitida por medio de la herencia (lo cual es aún más sorprendente si se considera que no se había dilucidado la estructura del DNA), y argumentó que una de las características principales de los seres vivos es la manera en que crean “orden a partir del desorden” del medio. Tal observación asombró a los fisicoquímicos desde tiempos de Boltzmann, pues parecía romper la segunda ley de la termodinámica. Sólo fue hasta el trabajo de Schrödinger que se concretó un esbozo del mecanismo utilizado.

¹⁵ Traducción al español en ed. Tusquets. La primera edición aparece en 1944, basada en las conferencias impartidas en Dublín en 1943.

Es necesario discutir someramente las tres posibles situaciones que conciernen a la termodinámica en cuanto a estudio de procesos (Nicolis 1989). La primera situación es un sistema *termodinamicamente en equilibrio*, en el que el cambio de los parámetros es cero. Tal proceso es expresado por la ecuación

$$(1) \quad J_{ieq} = \delta X_i / \delta t = 0$$

que dice que los flujos de energía o materia a través del sistema (dependientes del cambio de propiedades en una diferencia de tiempo) son nulos. Hago notar que esto sigue siendo la descripción de un proceso aunque un punto y otro del proceso sean iguales en propiedades. Es decir, la reversión del tiempo no representa ninguna diferencia.

La segunda situación es un sistema fuera de equilibrio debido a ciertas restricciones en los parámetros (gradientes, restricciones mecánicas, etc). Al quitar estas restricciones, el sistema dividido por diferencias en las propiedades intercambiará energía y/o materia hasta que alcance un estado de equilibrio termodinámico. Este comportamiento está predicho por la llamada ley cero de la termodinámica.

La tercera situación se presenta cuando un sistema es mantenido fuera del equilibrio, generalmente por un gradiente, y alcanza un estado estable. Este tipo de fenómenos son los llamados por Ilya Prigogine “sistemas disipativos”, pues en el sistema aparecen características emergentes cuya función es disipar el gradiente energético que constantemente se aplica al sistema. Los sistemas alejados del equilibrio en un estado estable presentan dos particularidades:

1. Las funciones de estado no se limitan a las funciones X_i de la ecuación anterior; también juegan un papel fundamental las llamadas funciones λ , que son fuerzas impuestas por el exterior para conservar el gradiente energético y por tanto el alejamiento del equilibrio.
2. Las propiedades emergentes generalmente se manifiestan en un comportamiento ordenado cuya función es aumentar la eficiencia de intercambio de entropía con el ambiente. De esta manera, el proceso termodinámico se manifiesta en una organización elevada que no va en contra de la segunda ley de la termodinámica, a veces expresada como la tendencia inherente al desorden; la organización mantiene un intercambio entrópico con el exterior que hacen posible la estabilidad del sistema.

Boltzmann fue el primero en notar la extraña tendencia de los seres vivos de mantenerse organizados, y realizando una explicación darwinista afirma que los seres vivos no luchan por recursos ni por energía (los cuales existen en exceso en el planeta y en sus propios sistemas), sino que “compiten por entropía”. Schrödinger expandió este concepto afirmando que los seres vivos parecen comportarse mecánicamente y no termodinámicamente, pues su tendencia de evadir el desorden y la caída inevitable al equilibrio se parece más al comportamiento estático de la materia cuando se acerca al cero absoluto y la entropía del proceso es considerada como nula. Esto es así sólo en apariencia: los procesos vivos son procesos termodinámicos mantenidos por el consumo de *entropía negativa o negentropía*, como le llama Schrödinger. Se sabe ahora que esto es en realidad la combinación de dos fenómenos energéticos de los seres vivos: la importación orden del medio en forma de moléculas altamente energéticas que se degradarán por medio de reacciones liberadoras de entropía, y la producción de calor que se liberará al medio (von Stockar 1999). En última instancia, la entropía negativa proviene del gradiente de energía solar utilizado por los seres fotosintéticos.

Así pues, tanto en el acercamiento de Schrödinger como en el de Prigogine la interacción con el medio es fundamental para determinar las funciones del sistema. Prigogine expande el concepto de Schrödinger introduciendo las propiedades emergentes causadas por los gradientes y la afirmación de que el comportamiento de los procesos disipativos es no lineal. Esto implica, entre otras cosas, que los sistemas disipativos tienen una respuesta particular *desproporcionada* con respecto al estímulo aplicado.

Massimo Rizotti (Rizotti 1996) afirma que los seres vivos no son sistemas disipativos, pues “los seres vivos no son pequeñas porciones de una gran solución acuosa, pues pueden existir lejos del agua”, y los sistemas disipativos “no pueden ser aislados del líquido en el que surgen, ni pueden ser delimitados por una frontera definida”. Rizotti parece estar hablando exclusivamente de uno de los ejemplos más sencillos y predilectos de Prigogine, las celdas de Bénard que surgen por un fenómeno de convección en un gradiente de temperatura. Como ya se mencionó, los sistemas disipativos no se definen por las características citadas por Rizotti, sino por su comportamiento termodinámico estable pero lejano al equilibrio, y la influencia de las funciones externas en su desarrollo. Los sistemas disipativos existen en gran variedad: desde los comportamientos de las eras glaciales o los fenómenos atmosféricos, hasta

láser y seres vivos. En efecto, para Scheider y Kay (Schneider 1994, 1995) los distintos niveles de organización biológica (desde células hasta ecosistemas) comprenden una serie de propiedades emergentes, cuya función termodinámica es disipar de manera eficiente el gradiente energético creado en la tierra por el Sol.

Es importante mencionar en este punto un fenómeno que, aunque presente en la discusión, no ha sido nombrado: la complejidad del sistema. Como señala Lazcano (Lazcano 2005), existe una tendencia a proponer generación espontánea de orden basándose en “principios de auto-organización” o “propiedades emergentes”, sin definir ni unos ni otras. Esto es demasiado cierto: tales argumentos (o falsos argumentos) frecuentemente aparecen sin realizar ninguna aclaración de su significado, volviéndolos expresiones vacías. El termino complejidad está ocupado aquí en el sentido en que lo utiliza Prigogine, cuando describe esta propiedad de los procesos fisicoquímicos disipativos, como las células de Bénard o las oscilaciones químicas. Algo complejo se caracteriza por: 1) Multiplicidad de soluciones posibles, es decir, no linealidad (como se muestra en la figura); 2) Interacción de un gran número de elementos, sean homogéneos o heterogéneos; 3) Surgimiento de propiedades emergentes, que derivan a la larga en un comportamiento concertado entre las partes. La unión de estos tres factores causan que el comportamiento del sistema y sus elementos pueda ser considerado caótico. Sin embargo, el caos surgido de tales comportamientos está inmerso en un sistema completamente determinista, por lo que es fundamentalmente distinto al caos surgido por “ruido” aleatorio.

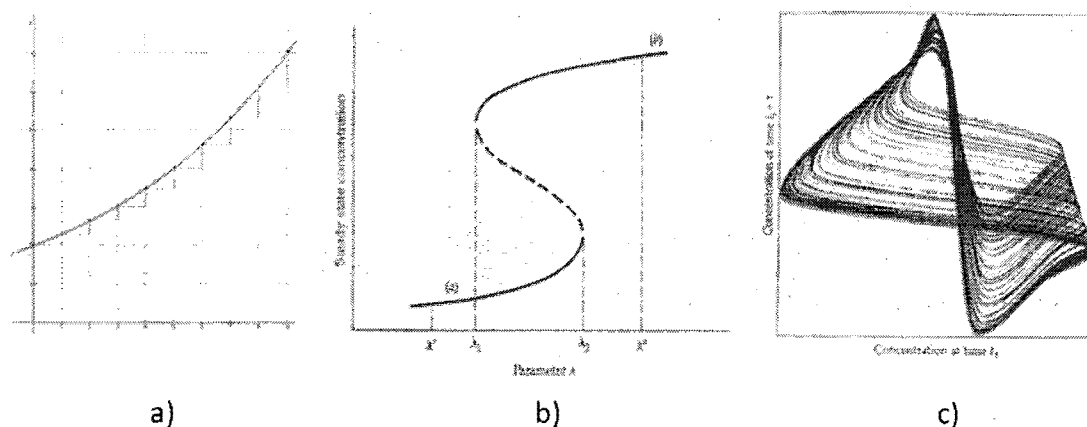


Fig. 1 Tres gráficas: a) Gráfica lineal; b) Gráfica no lineal: en la sección comprendida entre λ_1 y λ_2 cada punto posee múltiples soluciones, por lo que un cambio mínimo puede tener repercusiones impredecibles; c) Gráfica no lineal que muestra un atractor caótico (la gráfica representa las concentraciones de los reactivos en la reacción BZ). [b) y c) están tomadas de Nicolis, 1989]

Información

Las características termodinámicas anteriormente discutidas, aunque son parte funcional de la definición, son el *género próximo*; es decir, un conjunto en el que se inscribe el subconjunto que se quiere definir. La *diferencia específica*, la característica que delimita tal subconjunto, tiene que ver con las propiedades informacionales del sistema, y por tanto es necesario dejar un poco de lado la discusión fisicoquímica para enfocarnos en otras ramas aparentemente distintas en su totalidad, como teoría de la información e incluso semiótica. Más adelante, sin embargo, se verá que la información y la termodinámica están conectadas de manera profunda, sobre todo en los puntos de entropía informacional y complejidad algorítmica.

El concepto de información está cargado de subjetividad, tal vez más que muchos otros conceptos analizados aquí. Es un poco distinto al caso de la vida: no es difícil reconocer, intuitivamente, un sistema vivo. Los problemas en tal caso surgen cuando se quiere definir inequívocamente la esencia del sistema. La información, por otro lado, no *existe* como tal, sino que al parecer su aparición se realiza en cuanto es percibida por un observador. Sin embargo, creo que los seres vivos conocidos son únicos no sólo en el tipo de información que utilizan, sino en la manera en que la manejan; en última instancia, como concluiré más adelante, la singularidad de los seres vivos reside en el uso de códigos. Así, es necesario al menos mencionar los distintos significados que puede acarrear la palabra “información”, y cuál será de utilidad para nuestro análisis posterior.

Se ha analizado la naturaleza de la información de distintas maneras; los enfoques utilizados van desde secuencias lineales de símbolos hasta representaciones fenomenológicas de la totalidad del universo.¹⁶ De esta manera, es posible contrastar los diferentes acercamientos y hacer una lista de diversos sentidos válidos en los que se puede hablar de información. Por supuesto, dicha lista no es exhaustiva. Algunas de estas visiones de la información, además, han derivado una serie de teorías para el

¹⁶ Griffiths (2001) propone una división de la información entre causal (una simple relación entre los estados de la fuente y el receptor) e intencional (conexiones simbólicas con la posibilidad de ser falsas). Estas categorías corresponden, respectivamente y de manera bastante laxa, a las informaciones sintáctica y semántica. Espero que el análisis presentado aquí haga ver que el concepto de información puede ser mucho más complejo. Por ejemplo, la falsedad de una información puede ser derivada del ruido en el canal *sensu* Shannon. También podemos considerar información completamente simbólica, como lo es el código genético; es difícil imaginar una información “falsa” dada en este sistema, puesto que el código genético es arbitrario.

estudio e interpretación del fenómeno informativo; tal vez la más importante y utilizada es *la teoría de la información* propuesta por Shannon.

Información como mensaje

En 1948 Claude Shannon publica su artículo clásico, *A Mathematical Theory of Communication*. (Shannon, 1948). En él expone un modelo de transferencia de información completamente subordinado los intereses de la ingeniería. De hecho, históricamente, su desarrollo se debe al interés de optimizar la comunicación telegráfica; el interés de la teoría, en un principio, se centró en el tiempo de transferencia de un mensaje, la manera de optimizar la fidelidad de la transmisión y de reducir los errores, y las características de distintos tipos de codificación del mismo. Con el paso de los años, fue aparente que la teoría de Shannon tenía implicaciones mucho más profundas, dando lugar a la rama de las matemáticas llamada teoría de la información.

Se tienen que hacer ciertas consideraciones para entender cabalmente el papel de los mensajes en esta teoría:

- a) Los mensajes están compuestos con los signos de un alfabeto finito. En un sistema binario, por ejemplo, el alfabeto disponible consiste en los signos gráficos 0 y 1.
- b) Los mensajes son concatenaciones finitas de dichos símbolos. Es decir, un observador siempre puede medir la longitud de un mensaje determinado.
- c) Un mensaje de una longitud X es sólo un elemento del conjunto de mensajes de longitud X que pueden ser construidos con el alfabeto dado.
- d) Los mensajes frecuentemente tienen un significado. Sin embargo, para este acercamiento, cualquier consideración semántica es irrelevante. El mensaje es seleccionado del conjunto de mensajes posibles, y todos ellos tienen la misma importancia.

Shannon, partiendo de estos principios, considera a la información como *la reducción de la incertidumbre del sistema*. Consideremos, por ejemplo, un mensaje de longitud $x=4$, en un sistema binario. Los mensajes posibles, en este sistema, son $2^{(4)}=16$, por lo que cualquier mensaje dado tiene una probabilidad de ser escogido de 0.0625 (si todos los mensajes tienen la misma probabilidad). Cuando la longitud del mensaje

aumenta a, por ejemplo, $x=10$, el número de mensajes posibles aumenta (1024) y las probabilidades de cada uno se reducen (aprox. 0.001). Con esto, la incertidumbre crece; y cuando el mensaje finalmente se conoce, la información que contiene es mayor, que la que puede tener un mensaje de longitud $x=4$. Lo anterior puede ser dicho con la siguiente ecuación (Ash 1965),

$$(2) \quad I(X|Y) = H(X) - H(X|Y)$$

es decir: la información que un mensaje (Y) nos da acerca de una variable aleatoria (X), es la reducción de la incertidumbre de tal variable, $H(X)$, que causa el evento. La incertidumbre $H(X)$ también es llamada entropía en la teoría de la comunicación.

Información física

El universo físico, considerado de cierta manera, conlleva en su existencia una cantidad de información. Cada objeto tiene propiedades intrínsecas que pueden ser medidas para producir datos que hablen de él, más o menos precisamente. Esto no significa que las mediciones y percepciones que tenemos los humanos acerca del universo sean absolutas; simplemente que reflejan propiedades que contienen inherentemente un valor informativo. No es relevante, en este momento, discutir el inevitable grado de subjetividad con que se percibe el universo, o la validez de tal o cual manejo de los datos obtenidos. Una variante cercana de considerar al universo como información, es considerar que los hechos poseen información; es decir, las situaciones generan siempre información al hacer que cierto estado se vuelva factual. (Israel, 1990)

Información potencial (información contextual)

Sin embargo, resulta imposible decir si todo el universo está compuesto de información representada fenomenológicamente. Tal vez una manera menos extrema de considerar tal visión sea la llamada información contextual (Popa 2004). Este tipo de información se encuentra en el contexto de los eventos, modificando las probabilidades de ciertas ocurrencias fisicoquímicas. Las diferencias en concentración de un compuesto que favorezca cierta reacción química ocasionan un cambio en el desarrollo del sistema,

de una gran variedad de maneras. De cierta manera es una expresión de la potencialidad de un elemento para realizar ciertas actividades; así, podemos relacionarlo con el comportamiento enzimático. La actividad catalítica de cierta enzima no está codificada más que de una manera indirecta en el genoma; depende en gran medida de su ambiente, de su conformación particular (que puede ser una entre muchas posibles para la misma proteína), de la presencia y concentración del sustrato, etc.

Este tipo de información, al menos en sistemas químicos, puede crear a su vez eventos informativos, como cambios en la concentración de gradiente. Esto tiene una relación íntima con la modificación de las propiedades entrópicas del sistema, entre otras cosas. Así, podemos visualizar una relación entre la entropía informacional de Shannon y la entropía fisicoquímica.

Información como patrones

La información puede ser expresada como patrones ordenados; es lo que Popa (Popa 2004) llama información explícita. Los patrones informativos son generalmente estructurales, y pueden ser copiados sin ningún tipo de instrucción. Esto quiere decir que la información se utiliza como patrón, se extrae una representación del objeto; químicamente, generalmente la información utilizada en este sentido es una repetición de una configuración tridimensional, o una complementariedad, como en el caso de las cadenas de ácidos nucleicos.

Información semántica

Los acercamientos propuestos por Shannon, a primera vista, resultan un tanto contradictorios a lo que esperaríamos en un análisis de la información. ¿Cómo es posible que exista una medida de la información de un mensaje, pero no de su significado, validez o precisión? Aunque tal análisis aporta datos valiosos acerca de la información de un mensaje, hay ciertas instancias (como la información de la literatura, por ejemplo, o la bioinformación) en las que simplemente es insuficiente. Consideremos, por ejemplo, un texto escrito. La información de Shannon puede decir

las características matemáticas de tal información; incluso puede se pueden modificar los parámetros de análisis mediante la inclusión de las propiedades probabilísticas del idioma en que está escrito, por ejemplo. Sin embargo, ese tipo de texto contiene otros niveles de información además de las propiedades de una secuencia de símbolos. Por ejemplo, es pertinente preguntarse:

- ¿Las palabras que componen el texto son válidas en algún idioma? Es decir, ¿conlleven una información conceptual?
- ¿El ordenamiento de las palabras produce ideas coherentes?
- Si es así, ¿Qué tipo de información es la dada?. La información de un texto escrito puede ser de naturaleza instructiva, literaria, etc.

El caso es similar, aunque no idéntico, a la bioinformación. Uno puede preguntarse si la información genética puede codificar para una proteína; si tal proteína tiene algún tipo de actividad, considerando el contexto; si la actividad que realiza es estructural, catalítica o de señalización; etc. Para el análisis de Shannon, como ya se mencionó, todas estas consideraciones resultan completamente irrelevantes.

Es claro, desde un primer acercamiento, que esta lista no es exhaustiva. La información puede clasificarse de muchas otras maneras, como el grado y la manera de codificación que se utilice, las propiedades semánticas del mensaje, etc. Además, los tipos de información mencionados aquí no son mutuamente excluyentes. Tomemos, por ejemplo, una secuencia de un gen funcional. Tal información está codificada mediante un alfabeto predefinido, en una secuencia concatenada y finita de símbolos. Estas características recuerdan, por supuesto, a la teoría de la información de Shannon; de hecho, el sistema y sus componentes resultan muy similares a los presentados en el esquema

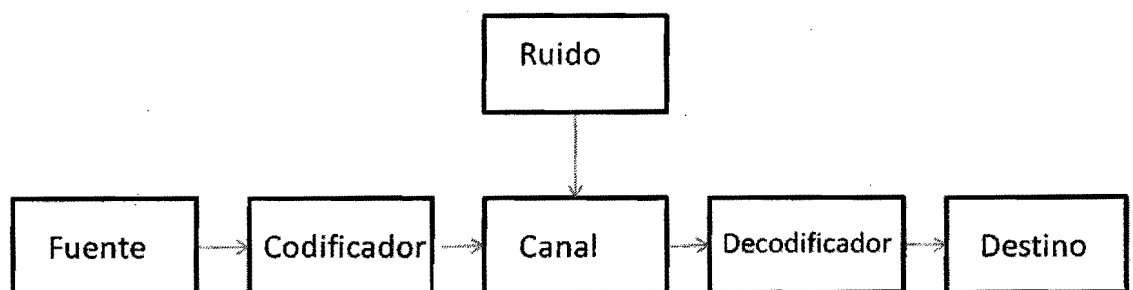


Fig. 2 Sistema de comunicación según la teoría de Shannon

Sin embargo, esta manera de considerar a un mensaje resulta incompleta, pues no considera el sentido o significado de dicho mensaje; la información que contiene sólo importa en función a la reducción de la incertidumbre general del sistema, como ya se ha visto en la ecuación (2). El gen existe dentro de un contexto integral, ya sea celular o pluricelular (que resulta aún más complejo); debido a esto, las consecuencias de su mensaje adquieren una multidimensionalidad imposible de predecir o analizar con las propuestas de Shannon. En efecto: en el caso del DNA, el significado es fundamental.

Dentro de un sistema celular (para tomar un sistema mínimo indiscutiblemente vivo), los diferentes tipos de información se entrelazan y sobreponen; sin embargo, podemos ver que se pueden trazar límites entre un proceso y otro, y definirlos lo suficiente como para articularlos. Las diferencias entre las distintas funciones y percepciones del fenómeno informativo, en ese momento, se vuelven visibles.

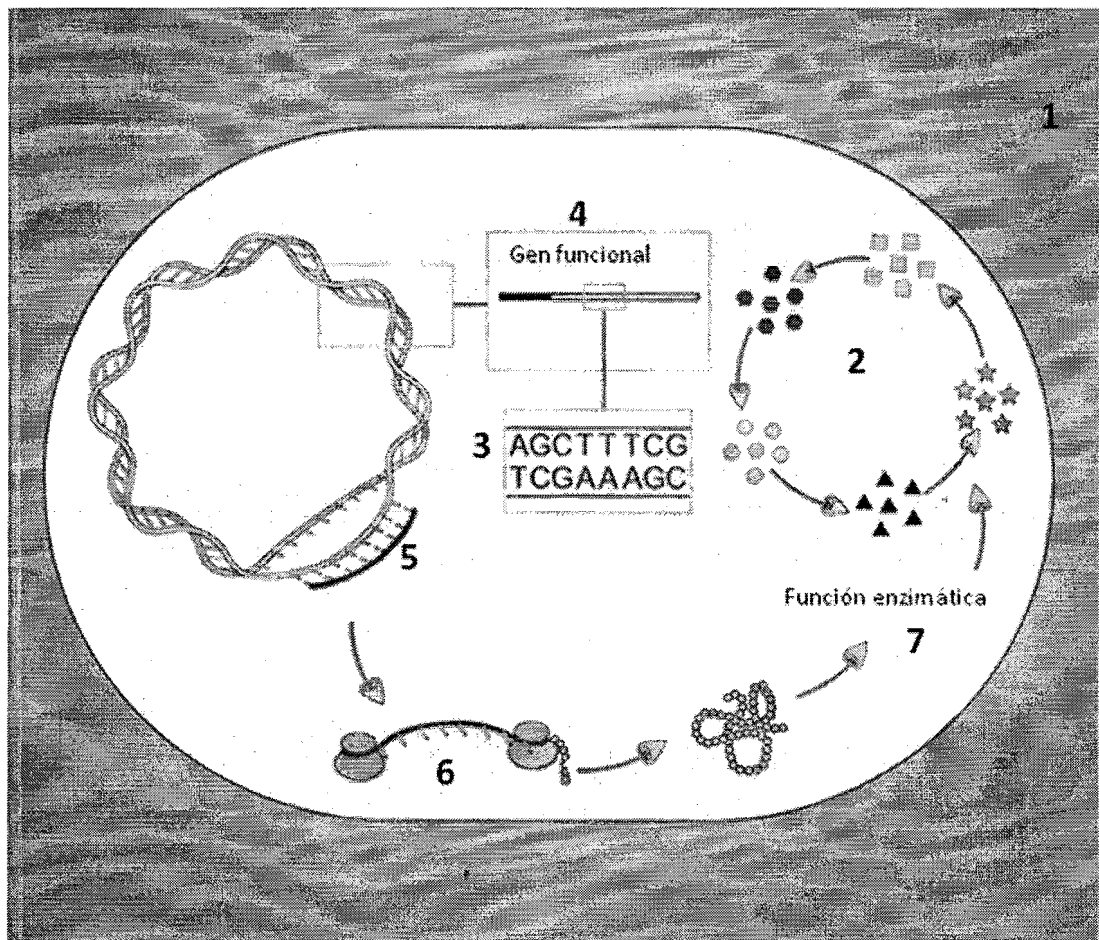


Fig. 3 Tipos de información celular, según el texto: 1, información física; 2, reacciones químicas; 3, secuencia lineal; 4, información semántica; 5, información como patrones; 6, códigos (en este caso, traducción); 7, información teleológica.

Consideremos, pues, a la célula como un sistema en el que conviven distintas entidades informativas; es pertinente enumerar algunas de las más importantes (esquemáticas de manera simple en la figura 3):

1. Como todo ser físico, la simple existencia de la célula conlleva una serie de datos informativos. Ya se ha mencionado que este tipo de información es potencialmente infinita; puede abarcar desde medidas simples y unidimensionales como las medidas de superficie o volumen hasta datos complejos como tipo celular, además de listas aparentemente infinitas (composición molecular o atómica, tipos de enzimas, genes, etc.). Por sí sola, tal cantidad de información no tiene ningún sentido explícito; una serie de relaciones tiene que ser dilucidada para que ésta adquiera cierta coherencia para nosotros. Sin embargo, parámetros invaluablemente pueden ser deducidos de la cuidadosa observación del sistema como un todo, como sus propiedades termodinámicas o bioquímicas.
2. La célula viva, como la conocemos, es un sistema inherentemente químico; todos sus procesos, independientemente de su naturaleza estructural, energética o de comunicación, son llevados a cabo mediante reacciones restringidas a un tipo de química muy específico. La información contextual o potencial se hace aparente aquí. Aunque la información física descrita anteriormente describe teóricamente (aunque, por supuesto, nunca práctica o empíricamente) la totalidad de la existencia de la célula, la información potencial se enfoca en procesos sutilmente distintos. Cada molécula de la célula, hasta cierto punto, puede estar involucrada en un número de reacciones, dependiendo del contexto. En este sentido, este tipo de información está estrechamente relacionada a la termodinámica del sistema; ciertas condiciones químicas favorecen cambios en el equilibrio de ciertas reacciones, lo cual a su vez cambia de manera retroactiva las condiciones, en una relación constante de *feedback*. Este tipo de intercambio informativo, pues, no sólo toma en cuenta la existencia física de cada molécula de la célula, sino también las relaciones entre ellas, así como los cambios a gran escala en el sistema a lo largo del tiempo. Celularmente, esto implica condiciones que pueden ser utilizadas para distintos procesos: como ya se ha dicho, esa “reducción en la aleatoriedad” puede dar lugar a gradientes químicos, transiciones de fase, e incluso alguna especie de compartimentalización.

3. Cuando se habla de información celular, lo primero que viene a la mente es el DNA. La molécula informacional por antonomasia puede ser vista como una secuencia de letras concatenadas, realizada con un alfabeto finito (de hecho, un alfabeto de sólo cuatro letras). Al mirarlo de esta manera, puede analizarse perfectamente (si bien de manera incompleta y hasta cierto punto insatisfactoria) con la teoría de información iniciada por Shannon. En efecto, la información en el DNA puede analizarse con base en sus propiedades como entropía (hablando en sentido informacional, claro), incertidumbre, etc. Por supuesto, pronto se harían aparentes ciertas propiedades de la concatenación de letras: su ordenamiento dista mucho de ser completamente aleatorio, y su entropía es muy baja. Dichos estudios pueden esclarecer ciertas propiedades del lenguaje utilizado por la célula; sin embargo, resultan incompletos: este nivel informacional ignora por completo el posible significado del mensaje analizado, y el significado resulta fundamental en los procesos celulares.
4. En el DNA, entonces, el significado es fundamental. En este nivel, la información que contiene no sólo es importante como una instrucción precisa para construir una proteína (el más común e inmediato de los procesos que dirige). También es importante considerar una serie de elementos, de cierta manera gramaticales, que aparecen en el funcionamiento del lenguaje genético. Éstos incluyen el ordenamiento de distintos elementos, puntuaciones, reguladores, activadores, etc. Esto, en su conjunto, produce una serie de significados que se pueden observar directamente como procesos. La analogía salta a la vista: la teoría de Shannon se ocupa de la sintaxis (la forma del lenguaje), pero la semántica, que se ocupa del significado, es necesaria en un estudio coherente y completo.
5. Dentro de la célula, y específicamente en los ácidos nucleicos, podemos encontrar la transferencia de información como patrón (Popa 2004). Este tipo de información, como ya se ha dicho, es copiada simplemente siguiendo un modelo físico, sin ningún tipo de instrucción externa. En este sentido, encontramos varias modalidades de transferencias de patrones: la transferencia $A \rightarrow A$, en la que una molécula se copia de manera idéntica; el tipo $A^+ \rightarrow A^-$, en la que una molécula produce su complementario, como en el caso de la replicación de

DNA; y finalmente, el tipo $A \rightarrow B$, en el que una molécula produce otra de naturaleza distinta, como en la transcripción del DNA a RNA.

6. Otro tipo de información se presenta al analizar el cambio de sistema de expresión de ésta. La información del DNA se tiene que traducir a una secuencia de aminoácidos, utilizando una equivalencia específica. Es decir: la información de ácidos nucleicos se encuentra codificada. Sin embargo, la información del código no se encuentra almacenada en ningún lugar físico; sólo se hace presente al momento de poner en funcionamiento la maquinaria de traducción. Es el resultado de un número de interacciones moleculares complejas y de una cadena de causalidad que determina la equivalencia entre una “palabra” del ácido nucleico (un codón) y un aminoácido. Éste es el tipo de información fundamental para el presente trabajo, que se dedicará a analizarla y a proponer un posible origen.
7. Finalmente, todos los procesos de información que realiza la célula pueden tener una repercusión en su contexto que podría parecer teleológica (destinada a un fin particular). El ejemplo más impactante de éste tipo de procesos se presenta si consideramos a la célula no como un elemento aislado, sino como parte de un sistema pluricelular. Prácticamente todos los tipos celulares, en este contexto, externan señales para otras células, ya sea parácrina o endócrinamente. Estas señales tienen como función afectar al sistema en su totalidad, pero no tienen ninguna repercusión directa en la célula en sí. La información que activa y controla estos procesos no está explícitamente codificada en ninguna parte: es el resultado de una sinergia compleja entre la información genética, el contexto integral de la célula y las señales que otras células del sistema manden.

Complejidad Algorítmica

En este punto, creo necesario retomar un concepto ya analizado y observarlo desde otra perspectiva: el término “complejidad”, que informacionalmente toma un significado distinto. Hasta ahora, se habían considerado cierto número de características para describir a la complejidad de una manera principalmente fisicoquímica. Sin embargo, a la luz de la teoría de Shannon y de la manera en que ésta analiza secuencias de signos, Kolmogorov (Gammerman y Vok, 1999) propone el

concepto de “complejidad algorítmica”. Tal tipo de complejidad se define como la longitud mínima de un algoritmo computacional para obtener una secuencia dada. Cuando la complejidad algorítmica es máxima, su valor iguala a la longitud de la secuencia.

Este concepto es expandido por Abel y Trevors (Abel, 2005), que crean tres subconjuntos informacionales, utilizando a la complejidad y a la funcionalidad como los dos parámetros que definen cada uno. Por ejemplo: pensemos en los dos extremos del espectro de mensajes posibles, como son vistos por la teoría de la información, tomando arbitrariamente una longitud de mensaje $l=50$, en un mensaje genético (es decir, con cuatro dígitos o letras). En un extremo, la secuencia de ordenamiento máximo: una repetición de 50 letras iguales. Este mensaje tiene al mismo tiempo el mayor orden y la menor complejidad. Una secuencia de una sola letra repetida es la más comprimible en un algoritmo sencillo: “repetir A 50 veces”. Este tipo de secuencias compone el grupo llamado OSC (Ordered Sequence Complexity). En términos fisicoquímicos, están determinadas por un sesgo probabilístico o por leyes físicas. Por ejemplo, la adsorción de nucleótidos en montmorillonita tiende a producir secuencias de poliadeninas (Ertem 2004). Del otro lado del espectro, se encuentra la secuencia menos ordenada que podemos lograr: una secuencia en la que aparecen, de manera aleatoria y equiprobable, cada una de las cuatro letras posibles. Esta secuencia no se puede comprimir en un algoritmo sencillo pues no tiene ningún patrón definido; para transmitirla, se tiene que enunciar letra por letra. Es, por lo tanto, la más compleja. El sesgo de probabilidades, en este grupo, es mínimo o nulo; las letras aparecen (o se polimerizan, en el caso del DNA) de manera estocástica. Las secuencias que poseen estas características componen el grupo RSC (Random Sequence Complexity). Es importante hacer notar aquí que esta noción de complejidad no es compatible con la complejidad de los sistemas físicos: ésta presupone, como ya se ha visto, un grado alto de auto-organización, propiedades emergentes, etcétera.

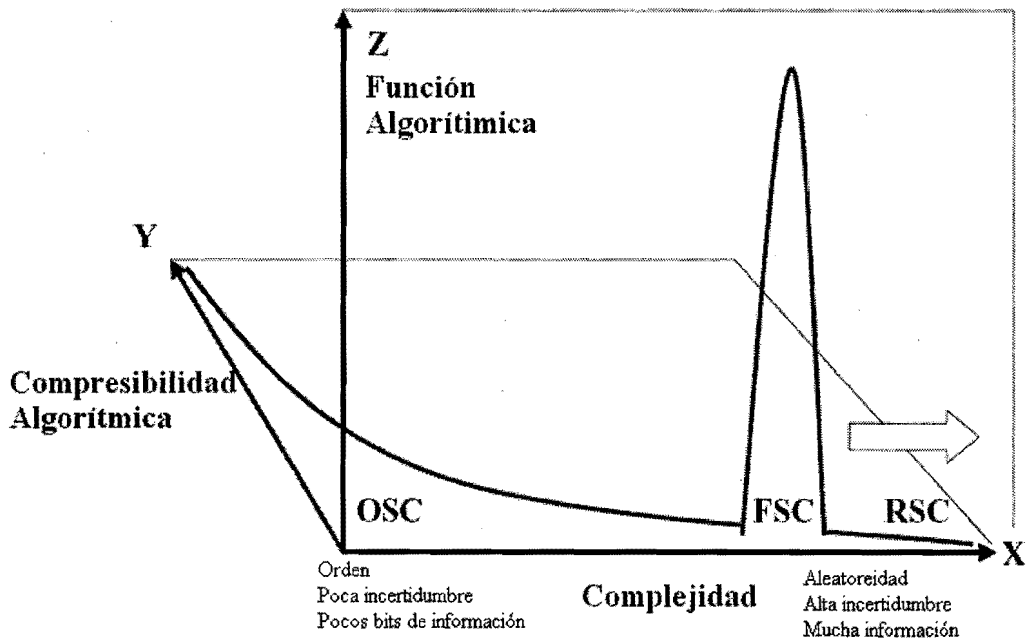


Fig. 4 Tres subconjuntos de complejidad algorítmica

Ni las secuencias aleatorias ni las ordenadas tienen la posibilidad de escribir ningún algoritmo, sin importar que tan simple sea éste. Tanto los patrones inmutables del orden máximo como la inestabilidad de la secuencia aleatoria son incapaces de albergar ningún contenido *instructivo*, independientemente de las diferencias entre la cantidad de información según Shannon. Abel y Trevors proponen una región en el gradiente de complejidad de los mensajes en los que se encuentran características importantes para la vida. Las secuencias encontradas en dicha región son llamadas por los autores “secuencias con complejidad funcional”. Si son analizados con los métodos de Shannon, la FSC (“Functional Sequence Complexity”) se encuentra muy cercana a la RSC¹⁷. Sin embargo, a diferencia de una aleatoriedad completa, las secuencias de FSC tienen función algorítmica; esto quiere decir que hacen *algo* en un contexto operacional. Este hecho único es lo que hace incompleto el análisis de Shannon.

Reiteraré que en el segundo párrafo de su primer artículo importante, Shannon se muestra perfectamente consciente de la irrelevancia del significado en los mensajes analizados. En el desarrollo del presente estudio, como ya se ha mencionado, se persigue otro fin: tratar de enfatizar la importancia de las relaciones simbólicas entre varios componentes de los sistemas biológicos. Posteriormente, trata de explorar ciertas

¹⁷ Comparar con el “orden a partir del desorden” de Schrödinger, y con el concepto de “orden al borde del caos” de Kauffman

posibilidades de surgimiento de dichos símbolos. Tal fenómeno se encuentra en el centro del funcionamiento biológico. Cada evento dentro de una célula innegablemente viva está dirigido por el significado que toman los genes, simples secuencias de cuatro signos; al final de la cadena de la comunicación, el metabolismo, uno de los destinos posibles del mensaje, se pone en funcionamiento, con una serie de características particulares: qué sustrato se va a tomar, y en qué se va a convertir; dónde, cómo, cuándo y qué tan rápido se realizará la reacción química especificada; qué producto es el objetivo final; etc. Acaso tal tejido algorítmico sea lo que causa la sensación teleológica de los seres vivos, esa percepción de que todos están orientados a un fin específico. La teleología no es una característica ontológica; es simplemente una ilusión que provoca el entramado autoconsistente de algoritmos programados en el genoma.¹⁸

En este punto nos enfrentamos con un problema. Toda secuencia con FSC es parte de un sistema integral. El mensaje biológico sólo es tal cuando ha pasado de la fuente (el DNA) hasta el receptor (la cadena de aminoácidos), pasando por la maquinaria de traducción que lo decodifica. Además, sólo tiene sentido en un contexto específico espacio-temporal. Un solo gen de un operón tiene sentido únicamente si se toma en cuenta como respuesta a un estímulo, y como parte de un conjunto mayor de información que se expresa bajo las mismas circunstancias para un fin determinado. Esto es lo que hace tan complicado de definir el grado de función; porque no es una mera magnitud. Las funciones biológicas difieren cuantitativa y cualitativamente, y se entrelazan para formar una construcción completa: cada parte sirve a un objetivo, independientemente de los artefactos o contingencias que se puedan presentar en el tiempo. Por supuesto, esto no implica ninguna clase de volición celular; simplemente quiero expresar la inconsistencia de ver a la información biológica como a) una simple cadena de letras, ignorando conscientemente sus repercusiones físicas, o b) un mensaje traducible, ignorando el contexto en que se produce. Es por eso que Abel y Trevors afirman que Shannon no analizó más que lo que está dentro del círculo en la siguiente figura.

¹⁸ No quiere decir que absolutamente todo esté programado en el genoma; simplemente que el funcionamiento del genoma es integral. Los eventos puramente físicos, por ejemplo, son igualmente importantes.

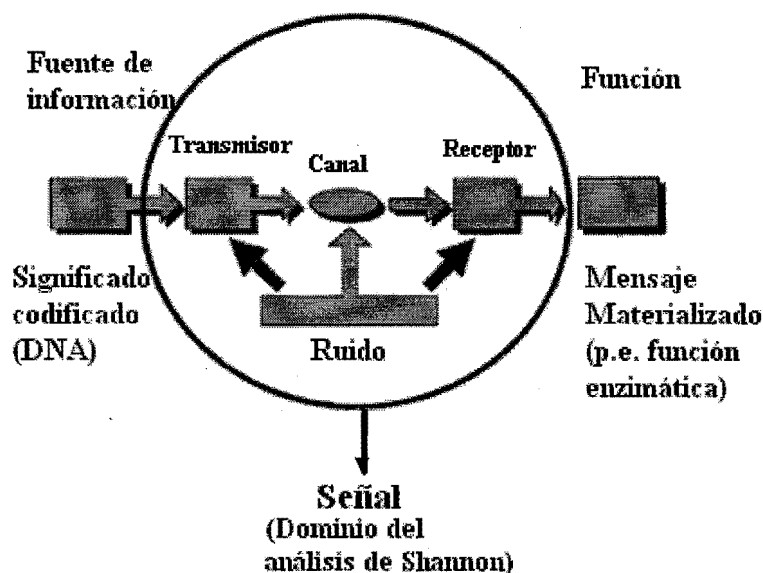


Fig. 5 Ruta de la información biológica; lo que está fuera del círculo es imposible de analizar por el método de Shannon

Ricard (Ricard 2003) afirma que los parámetros que deja fuera el análisis de Shannon (en particular, la discontinuidad de las variables y la no uniformidad de la red informacional biológica) es, precisamente, lo que hace que el sistema se comporte como una unidad integral y compleja.

¿Cuál es la consecuencia tomar en cuenta tales características? Creo que la más importante es ésta: el funcionamiento biológico de un mensaje celular (es decir, de una comunicación, a cualquier nivel), sólo puede ser apreciado *a posteriori*. Es decir que podemos predecir la secuencia de la proteína, e incluso eventualmente se logrará predecir su conformación, pero la mayoría de la información que es relevante en estos sistemas se aprende de manera empírica, después del hecho¹⁹. Nuestro conocimiento acerca de cualquier sistema biológico trasciende el conocimiento que aporta un análisis de Shannon, pues éste último es simplemente un modo parcial e incompleto de conocer. Por otra parte, la cualidad *a posteriori* del sistema genético tiene otra consecuencia. No sólo dificulta enormemente el análisis del fenómeno, sino que significa que no se pueden tomar ni siquiera como modelos de origen de la vida. Es un caso similar a la selección: la selección natural actúa sobre sistemas definitivamente vivos; la selección

¹⁹ Hablo aquí de un observador para poder esclarecer el punto. Sin embargo, el observador tiene tanta cabida aquí como en el sistema de Shannon; es prácticamente irrelevante. Quiero decir que con observador o sin él, el gen se *realiza* después de llevar a cabo su función

no puede ser ni la causa ni la explicación de la vida en sí. (Afirmar lo contrario significaría negar que su dominio es exclusivamente los sistemas vivos, lo cual crearía un argumento tautológico). De la misma manera, los algoritmos no pueden ser la causa de sí mismos; su surgimiento no puede ser espontáneo, ni fragmentario.

Un artículo reciente (Durston 2007) trata de desarrollar las ideas de Abel y Trevors. Sin embargo, toma simplemente las partes conservadas de las proteínas, dándole así valores de complejidad. Es decir, un aminoácido importante para la función conlleva un valor de complejidad alto. Este acercamiento es distinto a los dos tipos de complejidad analizados en este trabajo; sólo se mencionará que no tiene una relación relevante con las características algorítmicas propuestas por el primer artículo.

4. Semiótica

Símbolos

Independientemente de las características de la información manejada por los sistemas vivos, la propuesta de este trabajo se enfoca en el modo de manejarla; los procesos informativos de la vida no sólo son asombrosos por las características ya mencionadas en las secciones anteriores, sino porque se hace uso de un sistema semiótico sofisticado.

Los seres humanos manejamos sistemas semióticos de manera automática: prácticamente todas las actividades que realizamos contienen códigos preestablecidos. La propia mente humana, y probablemente muchas de las percepciones fenomenológicas de otros animales, funcionan en su totalidad por la creación, uso y modificación de *relaciones simbólicas*. Se puede afirmar que todo lenguaje es una colección de símbolos, es decir, un sistema semiótico²⁰.

Según Charles Peirce (1991), existen 3 tipos de signos. El primero, el icono, es un objeto que posee una cualidad que lo hace apto para ser un signo. El icono puede referirse o no a un objeto determinado; puede ser incluso que su objeto no tiene existencia. Es un tipo de signo que está determinado en sí mismo en su posibilidad de hacer referencia a algo. Una línea trazada en un papel es un icono de la abstracción de una línea recta, aunque no haya nadie para interpretarlo y aunque no haya sido hecho

²⁰ Queiroz y el-hani (Queiroz 2006) definen sistema simbólico como aquellos que producen, comunican, reciben, computan e interpretan diversas clases de símbolos.

con el propósito expreso de representarla. Podemos extender la definición de Peirce: la información, tal como la maneja Shannon, es un signo icónico, pues estos existen sin ninguna carga semántica. Una secuencia de caracteres existe como signo en sí mismo dentro de la teoría de la información. Los iconos sólo se componen de un primer elemento, su propia existencia.

El segundo tipo de signo es el índice o sema. Está determinado por una relación fenomenológica entre dos elementos; es decir, depende de la existencia de ambos. Una veleta indica la dirección del viento gracias a la relación física de ambos. Una fotografía indica las condiciones físicas en las que fue tomada la imagen; la relación física entre ella y los objetos que representa en un momento determinado del tiempo es innegable y no necesita intérprete. Como vemos, los índices se componen de dos partes: el signo y el objeto al que hace referencia. No necesitan un tercer elemento que los conecte, pues están conectados intrínsecamente. Este tipo de signo conecta directamente con los tipos de información y con lo que he llamado “información como patrones”. Dentro de una cadena de DNA, la timina presupone la adenina, y la guanina presupone la citosina, pues son físicamente complementarias.

Consideremos, para explicar el tercer tipo de signo, la relación entre el grafo A y el sonido con el que se relaciona, [a]. Entre el signo y el sonido no existe ninguna relación física, ni están conectados por la naturaleza de ambos. Las marcas en una superficie que componen el grafo y las vibraciones del aire que provocan el sonido necesitan una cierta abstracción entre sonido y visión, realizada por una mente que los interpreta. Sin la percepción humana tal relación no existiría. Por lo tanto, el símbolo es el tipo de signo más complejo; necesita de tres elementos (representamen, objeto e intérprete) para cobrar sentido. Más aún, el símbolo depende de acuerdos o *convenciones* para funcionar; su carácter de representamen es precisamente que él mismo es una regla que indica su significado. *El símbolo es una ley adquirida que funcionará en el futuro indefinido.* Por supuesto, como todas las convenciones, está sujeta a los cambios de una evolución constante.

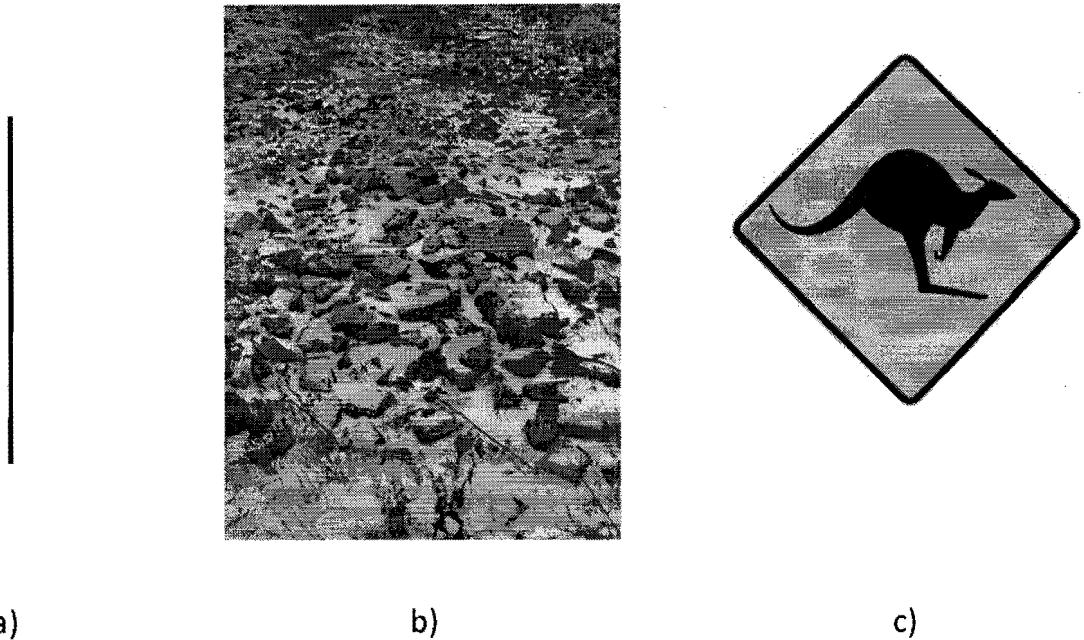


Fig. 6 – Tipos de signos. a) El icono es signo y objeto del signo simultáneamente; b) El índice depende físicamente del objeto; c) El símbolo está conectado conceptualmente con su objeto, mediante una abstracción humana.

El ejemplo más cotidiano de un sistema simbólico es el lenguaje. Aunque nos hemos acostumbrado a usarlo sin ningún problema, es una red extremadamente compleja de símbolos que funcionan en varios niveles. Cualquier palabra genera una serie de asociaciones. Por ejemplo, la palabra “libro”, que finalmente es una serie de 5 sonidos (o de 5 grafos, dependiendo el medio de representación), se conecta directamente con el concepto abstracto de “libro” como un bloque anónimo de hojas de papel. A su vez, ambos se conectan con signos más individuales como libros específicos. Y en una asociación más, pueden referirse a objetos físicos determinados, independientes de toda interpretación subjetiva; es decir, objetos concretos que existen fuera del observador. Tales afirmaciones dan lugar a la vieja disputa acerca de la validez de términos como “objetivo”, “subjetivo”, “observador”, etc., cuyo análisis está fuera de los fines de este trabajo.

Códigos naturales

Cualquier lenguaje es un código, y cualquier código es una serie de relaciones simbólicas. Esto no presenta ningún problema lógico si consideramos siempre a una mente animal como intérprete. Sin embargo, y como se menciona al principio de éste

apartado, una de las características más sorprendentes de todos los seres vivos es la presencia de códigos reales sin necesidad de una mente que realice las convenciones sobre las cuales funcionan. Varios autores han analizado a este punto (principalmente Barbieri, 2003) y resulta evidente que los sistemas simbólicos funcionan dentro de los seres vivos a varios niveles. Las abstracciones de la mente humana parecen ser sólo un ejemplo de un sistema semiótico natural. Recordemos que la mente humana, a final de cuentas, reside sobre una base física: la fisiología cerebral. Esto conlleva una serie de discusiones filosóficas (principalmente en las cuestiones de la naturaleza física o intangible de la mente, y acerca de la artificialidad o naturalidad de los objetos); de nuevo, tales cuestiones no conciernen a este trabajo.

Marcello Barbieri ha analizado ya distintos tipos de códigos o “convenciones naturales”, desde la equivalencia aminoácido-codón hasta el significado semántico de la expresión protéica durante el desarrollo de un organismo. Sin embargo, el más fundamental y definitivamente el más difundido de los códigos en los seres vivos es el código genético, es decir, la serie de reglas que permiten a cualquier célula traducir la información contenida en ácidos nucleicos a información protéica. El código determina una correlación no ambigua entre cada triplete o codón posible de nucleótidos, y un aminoácido específico. Justamente aquí radica la diferencia entre este trabajo y el de Barbieri: en el considerar al código genético como el más fundamental. Barbieri lo considera como equivalente, pero este es el único código sin una carga semántica (es decir, sólo es representado como una correlación sin significado).

		Segunda Base				
		U	C	A	G	
Primera Base	U	Phe	Ser	Tyr	Cys	U
		Phe	Ser	Tyr	Cys	C
		Leu	Ser	STOP	STOP	A
		Leu	Ser	STOP	Trp	G
	C	Leu	Pro	His	Arg	U
		Leu	Pro	His	Arg	C
		Leu	Pro	Gln	Arg	A
		Leu	Pro	Gln	Arg	G
	A	Ile	Thr	Asn	Ser	U
		Ile	Thr	Asn	Ser	C
		Ile	Thr	Lys	Arg	A
		Met	Thr	Lys	Arg	G
	G	Val	Ala	Asp	Gly	U
		Val	Ala	Asp	Gly	C
		Val	Ala	Glu	Gly	A
		Val	Ala	Glu	Gly	G

Fig. 7 – El código genético estándar

Dado que existen $4^3=64$ combinaciones posibles de tres nucleótidos, el código genético consta de 64 reglas de correlación. El triplete de nucleótidos AAA, por ejemplo, codifica para lisina en el código estándar.

El código genético mostrado en la figura no es el único, pero es el que traduce la gran mayoría de los genes conocidos; en consecuencia es llamado el código genético estándar. Existen, sin embargo, pequeñas variaciones en organismos específicos, principalmente en procariontes, arqueobacterias y en los aparatos de traducción de los organelos eucariontes que poseen genoma propio (como mitocondrias y plástidos). Esto no quiere decir que el código es ambiguo. Dentro de los organismos en los que se han encontrado cambios en las reglas de codificación, éstos son mínimos y siempre los mismos dentro de cada taxón.

(a) Nuclear variants

	U	C	A	G	
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	Diplomonads Acetabularia Some ciliates -Other ciliates
	UUC Phe	UCC Ser	UAC Tyr	UGC Cys	
	UUA Leu	UCA Ser	UAA TYP	UGA TER	
C	CUU Leu	CCU Pro	CAU His	CGU Arg	Cys Trp Euplozia Mycoplasma Spiroplasma
	CUC Leu	CCC Pro	CAC His	CGC Arg	
	CUA Leu	CCA Pro	CAA Gln	CGA Arg	
A	CUG Leu	CCG Pro	CAG Gln	CGG Arg	Nonsense Mycoplasma Spiroplasma
	AUU Ile	ACU Thr	AAU Asn	AGU Ser	
	AUC Ile	ACC Thr	AAC Asn	AGC Ser	
G	AUA Ile	ACA Thr	AAA Lys	AGA Arg	Nonsense Micrococcus
	AUG Met	ACG Thr	AAG Lys	AGG Arg	
	GUU Val	GCU Ala	GAU Asp	GGU Gly	
G	GUC Val	GCC Ala	GAC Asp	GGC Gly	
	GUA Val	GCA Ala	GAA Glu	GGA Gly	
	GUG Val	GCG Ala	GAG Glu	GGG Gly	

(b) Mitochondrial variants

	U	C	A	G	
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	Various Some chlorophytes (UAG = Leu) Some chlorophytes (UAG = Ala) Platyhelminths (UAA = Tyr)
	UUC Phe	UCC Ser	UAC Tyr	UGC Cys	
	UUA Leu	UCA Ser	UAA TER	UGA TER	
C	CUU Leu	CCU Pro	CAU His	CGU Arg	Ancestral mitochondrion -Diatyostellum -Plants Chondrus crispus Some prymnesophytes
	CUC Leu	CCC Pro	CAC His	CGC Arg	
	CUA Leu	CCA Pro	CAA Gln	CGA Arg	
A	CUG Leu	CCG Pro	CAG Gln	CGG Arg	Trp Nonsense Yeast
	AUU Ile	ACU Thr	AAU Asn	AGU Ser	
	AUC Ile	ACC Thr	AAC Asn	AGC Ser	
G	AUA Ile	ACA Thr	AAA Lys	AGA Arg	Nonsense Candida Prototheca (algae)
	AUG Met	ACG Thr	AAG Lys	AGG Arg	
	GUU Val	GCU Ala	GAU Asp	GGU Gly	
G	GUC Val	GCC Ala	GAC Asp	GGC Gly	Various Bacteria (Ser) -Drosophila (nonsense) -Vertebrates (Gly) -Tunicates (TER)
	GUA Val	GCA Ala	GAA Glu	GGA Gly	
	GUG Val	GCG Ala	GAG Glu	GGG Gly	

Fig. 8 - Algunas variantes del código genético (Knight 1999)

En la segunda mitad del trabajo se considerarán las implicaciones de las variaciones del código, así como otras propiedades particulares. En este momento, el punto de interés es analizar la relación aminoácido-codón como un signo, que se comporta fundamentalmente de la misma manera que la relación entre un grafo y un sonido de una letra. El codón no tiene ninguna relación física intrínseca con el

aminoácido que codifica; aunque el sistema haya surgido por factores deterministas, sin el conjunto de reglas es prácticamente imposible deducirlo (así como es imposible deducir el sonido [a] a partir del grafo A). La cadena de DNA y la proteína que codifica son mundos independientes unidos por el código, como afirma Marcello Barbieri. Sin embargo, existe una diferencia entre ambos sistemas: la naturaleza del intérprete. En el caso de las letras escritas y pronunciadas, la mente humana es el punto de unión de los dos elementos; los símbolos normalmente presuponen la existencia de la mente, es decir, un ente metafísico, para que surja una interpretación abstracta que los una. En cambio el código genético antecede a la metafísica. En él, todos los elementos (signo, objeto, intérprete y sistema de reglas) son de naturaleza puramente física. El argumento central es el siguiente: *los seres vivos cuentan con el único sistema simbólico funcional generado fuera de la mente humana*, al menos dentro del universo físico conocido.

En la teoría semiótica tradicional no son considerados casos de este tipo. Durante todo el siglo XX los signos se han estudiado principalmente mediante dos acercamientos opuestos: la visión lingüística estructuralista de Saussure (Saussure 1993), que está emparentado más bien con la antropología del lenguaje; y la teoría abstracta de Peirce, tomando al mundo y a sus partes como signos, pero siempre dando por sentado que el intérprete es el observador humano. En ambas la atención se centra en el sujeto que interpreta. Bajo la visión tradicionalista, es difícil decidir en dónde reside el “sujeto” que hace funcionar las reglas; puede ser el RNA de transferencia (el adaptador que une ambos tipos de moléculas), el ribosoma, o la célula como un sistema completo. Por otro lado, cualquier intento para localizar la cualidad interpretativa de un sistema físico sería hecho en función de un interés, bastante difundido en las ciencias biológicas, de antropomorfizar los objetos de estudio. Por lo tanto, no señalaré el lugar de residencia de las convenciones naturales, ni jerarquizaré los diversos componentes para dar a cada uno mayor o menor importancia semiótica. Creo que lo más que se puede decir al respecto es que la interpretación es una propiedad emergente de un sistema complejo.

Es necesario insistir en que la relación simbólica existente en los sistemas vivos no es así llamada por una semejanza meramente superficial. Por ejemplo, se podría metaforizar una reacción química como si funcionara semióticamente, diciendo que el reactivo es el signo, el producto es el objeto, y tal vez el catalizador podría considerarse como intérprete. Esto no es más que una metáfora que podría aplicarse a cualquier

sistema de tres componentes. En efecto, cualquier proceso tiene una serie de condiciones iniciales, condiciones finales y un factor causal; sin embargo, éstas no son relaciones semióticas verdaderas. El sustrato de una reacción y su producto están conectados necesariamente no sólo por una continuidad física, sino por un lapso temporal: no pueden existir ambos de manera simultánea. Las relaciones semióticas son, en este sentido, no temporales y no espaciales. Dado a esto, pueden expresarse de manera abstracta mediante la tabla de codones de la figura. En efecto, tienen una base física de muchos componentes: las aminoacil-sintetasas, los RNA de transferencia, el aparato de traducción. Sin embargo, las reglas como tales son una *abstracción natural*. No necesitan de una molécula física para realizarse, y por esto somos capaces de decodificar cualquier secuencia de DNA, aunque ésta no codifique para ninguna proteína funcional, incluso aunque tal molécula nunca haya tenido una existencia física. Esta propiedad de las reglas de codificación también se puede apreciar en las distintas fases de lectura que posee cada secuencia genética; al estar codificadas en tripletes, es posible obtener 3 conjuntos distintos de aminoácidos de una sola cadena de nucleótidos. Es decir: se puede empezar a “leer” la secuencia desde el nucleótido n , desde el nucleótido $n+1$ o desde el nucleótido $n+2$. (El nucleótido $n+3$ nos sitúa de nuevo en el primer marco de lectura, el que comienza en el nucleótido n).

Una de las consecuencias más importantes de la existencia de un código es la creación de dos formatos independientes de información; esto, a su vez, hace posible la inevitable separación de dos sistemas: fenotipo y genotipo. Esta es una propiedad necesaria en los sistemas biológicos conocidos, y probablemente sea una característica difundida en todas las formas de vida. Los dos cambios más evidentes gracias a esta separación son la posibilidad de un almacenaje independiente de la información (y su replicación con fidelidad suficiente para no ocasionar una catástrofe de error), y la naturaleza reemplazable de las proteínas, las cuales se pueden sintetizar indefinidamente *de novo* mientras exista la copia de la cadena de RNA que las codifica. La disociación de la parte puramente informacional (nucléica) y los componentes funcionales (protéicos) son necesarios para una serie de actividades sin las cuales la vida se disolvería en su medio. Esto es, en parte, uno de los elementos que la hace mucho más resistente a los cambios que los sistemas disipativos de Prigogine, analizados anteriormente.

Argumentos similares han sido expuestos por varios autores, sobre todo en los escritos del reciente campo de la biosemiótica, surgido por los trabajos pioneros de H.H. Pattee. Sin embargo, existe una diferencia fundamental: Pattee reconoce dos “modos” de la vida, uno dinámico (los procesos bioquímicos en particular) y otro simbólico (las instrucciones genéticas) (Pattee 2001, Hoffmeyer 2001). Éstos, perpetuándose a sí mismos mutuamente en lo que se llama *semantic closure* (Pattee 1982 citado en Umerez 2001), representan un modelo epistemológico más que una descripción física del sistema; es una expresión semiótica del problema del huevo y la gallina. La relación simbólica es la que hay entre el almacenaje de información (los genes) y la función que éstos llevan a cabo (Etxeberria 2001). Como se ve, esta propiedad es la capacidad de la descendencia materializar las funciones de sus antecesores. Pattee afirma que su visión simbólica se refiere al control de expresión para mejor adecuación, que presupone una manipulación compleja de información. La diferencia fundamental, pues, entre la propuesta simbólica de este trabajo (más acorde con la visión de Barbieri) y la escuela tradicional de biosemiótica está expresada por la convicción que la unidad biosemiótica mínima está representada por la célula (Kull 2002, Emmeche 2000).

Definición de vida. Consideraciones finales.

*quid est ergo tempus? si nemo ex me quaerat, scio;
si quaerenti explicare velim, nescio.*

San Agustín, *Confesiones*

Se ha dicho que la existencia de códigos simbólicos es un fenómeno exclusivo de los seres vivos; incluso se propone tomar como característica para construir una definición. Sin embargo, el lenguaje humano posee este tipo de relaciones simbólicas; las extensiones formales de éste (los lenguajes matemáticos y los programas de computadoras) también funcionan así. ¿Quiere esto decir, como ya se ha afirmado antes, que el lenguaje y la cultura humana, o los virus informáticos están vivos? Esta clase de objetos evolucionan darwinianamente, y tienen un lenguaje simbólico; bajo la mirada de varias definiciones, nuestra cultura está literalmente viva.

Las características fisicoquímicas que se han discutido sirven como un contexto para evitar este tipo de discusiones ligeramente ociosas. El rechazo a igualar a la cultura humana, al lenguaje animal o a los programas de computadora con los seres vivos que conocemos es casi instintivo. ¿Cómo hablar del nacimiento o de la muerte de un virus

informático? Aún cuando el programa falle y deje de funcionar, tal interrupción es fenomenológicamente inexistente. Además, los seres vivos parecen funcionar sobre sí mismos. Cambiar el contexto en el que se manifiestan probablemente interrumpa el funcionamiento termodinámico, pero no hace que sus sistemas simbólicos se vuelvan incoherentes. Cambiar el contexto de un programa informático, sin embargo, hace que los algoritmos programados pierdan validez.

De tal manera, la primera parte de este trabajo descansa sobre las nociones de *genus proximum* y *differentia specifica*; es decir, parte de un conjunto del universo conocido (el conjunto de los sistemas termodinámicos) y dentro de él, trata de definir un subconjunto con características particulares (el uso de lenguajes simbólicos para sustentar y perpetuar su orden). Tal método tiene dos implicaciones principales. Por un lado, es un intento de conciliar la visión puramente física de la vida con la escuela biológica que afirma que la vida es un fenómeno que no se puede reducir a leyes físicas. (Mayr 2004), discusión que en mi punto de vista trasciende el simple rechazo a las visiones epistemológicamente reduccionistas. La vida está innegablemente circunscrita en el universo físico; por tanto, las leyes físicas deben de poder hacer una descripción de sus formas y procesos. Es ingenuo pensar, sin embargo, que se puede describir de manera satisfactoria los mecanismos intrincados de información con un enfoque puramente termodinámico, así como es ingenuo pensar que se pueden deducir las características emergentes de un sistema únicamente con el estudio de sus partículas. Por otra parte, la definición basada en género y diferencia es útil para descartar algunos sistemas que se pueden considerar como seres vivos utilizando otros criterios. Una definición que se limita a enumerar características de los seres vivos sin delimitar el conjunto en el que se encuentran ocasiona ambigüedades sin solución. ¿Cuántas características son necesarias y suficientes? ¿Qué ocurre si se presentan todas menos una? La clasificación taxonómica de Linnaeus se basa precisamente en este sistema, previendo las posibles confusiones que pueden surgir de otro modo.

La búsqueda este escrito se interrumpe en este punto: en la propuesta de un género y una diferencia específica. No se propone ninguna definición. Nuestro conocimiento de la vida es limitado, no sólo por la muy reciente elucidación de hechos sino porque sólo la conocemos en su encarnación terrestre. Además, la filosofía del siglo XX ha creado un punto de vista mucho más pesimista que en épocas pasadas. Una de las observaciones más conocidas (y más superficiales) de las investigaciones

filosóficas de Wittgenstein presenta de manera contundente la falibilidad del lenguaje. No se puede mirar alrededor nuestro sin encontrarnos un sin número de objetos o eventos cuyos nombres son ambiguos. Una silla, por ejemplo, es tan difícil de definir como un sistema vivo. Sin embargo, espero que la inclusión de un elemento físico y uno aparentemente metafísico establezca o siga una dirección fértil en la visión de los seres vivos. Puede ser que el debate constante alrededor de la posibilidad de reducir la vida a la física se deba a la aparente libertad de los sistemas simbólicos, mencionada por Hoffmeyer y Emmeche 1991: "*No natural law restricts the possibility space of a written (or spoken) text*". Tal es la libertad que la arbitrariedad proporciona a los símbolos.

El presente trabajo se fundamenta en la irresoluble paradoja de la naturaleza de las definiciones: son imposibles y extremadamente útiles. Las definiciones se comportan como el Demonio de Maxwell: son construcciones imposibles que sin embargo replantean y cuestionan lo que se cree conocer acerca de cualquier evento. Las características propuestas aquí servirán, idealmente, para marcar caminos de investigación y las críticas hechas a lo dicho aquí continuarán el avance sobre ellos. Por ejemplo, alguien más ingenioso o sagaz encontrará algún ejemplo de un sistema simbólico no vivo, materializado en el universo físico, y tal vez el estudio de los orígenes de este hipotético sistema servirá para comprender mejor el origen o la naturaleza de la vida.

PARTE II: EL ORIGEN DE LA INFORMACIÓN

1. La paradoja del huevo y la gallina

El origen de la vida

Los procesos fundamentales de los seres vivos están dictados por un sistema cerrado sobre sí mismo: la información del DNA codifica para proteínas, las cuales en última instancia servirán para mantener el metabolismo que permitirá la replicación del DNA. El origen de este sistema, pues, es una de las grandes incógnitas del estudio del origen de la vida. Usualmente se llama a este problema *la paradoja del huevo y la gallina*, porque aparentemente no puede existir uno sin la existencia anterior del otro. Esto ha causado que los autores que trabajan en este campo se hayan separado en dos grandes grupos, antagónicos y casi mutuamente exclusivos: los que se centran en el surgimiento inicial del metabolismo, y los que prefieren estudiar la replicación como primer el primer evento que ocurrió en el camino del origen de la vida. Este trabajo es una propuesta acerca del surgimiento del código genético; el establecimiento de este sistema es probablemente posterior a la organización de la química prebiótica, y por tanto la indagación de las posibles rutas de la evolución química no es uno de sus objetivos. Sin embargo, ambos escenarios son fundamentalmente distintos y crean condiciones características.

El mundo de RNA: “Replicación primero”

Como afirman Gerald Joyce y Leslie Orgel (Joyce 1999), existen varias versiones de la teoría del mundo del RNA; sin embargo, la característica común es que todas postulan la existencia de una fase en el surgimiento de la vida en la que proteínas codificadas genéticamente no tomaban parte en la catálisis de reacciones bioquímicas. El argumento central de estas teorías es la capacidad que tienen los ácidos nucleicos para cumplir dos funciones: vehículo de información y molécula catalizadora.

La versión más aceptada de la teoría inicia con un modelo similar a la sopa prebiótica de Oparin, en la que se encontrarían los “bloques de construcción” necesarios. Dichos bloques, en el caso del RNA, pueden ser nucleótidos activados,

nucleósidos o los constituyentes por separado (fosfatos, azúcares y bases nitrogenadas). A partir de esta mezcla y por medio de evolución química o prebiótica, se propone la generación espontánea de una cadena de RNA, probablemente de un número reducido de bases, que tenga propiedades autorreplicantes; es decir, que actúe sobre sí misma o sobre copias de sí misma para producir cadenas complementarias, para después actuar sobre estas últimas para producir más copias de sí misma, generando una población de RNA autorreplicante. A partir de aquí, prácticamente todos los procesos que llevan al origen de la vida dependen de la selección natural darwiniana. Primero, aumenta la fidelidad del replicador para que la población no degenere en copias erróneas y no funcionales; a este evento se llama “catástrofe del error”. Algunos autores proponen el surgimiento de ribozimas (es decir, moléculas de RNA con actividad catalítica) con diversas funciones: síntesis de las partes esenciales de los nucleótidos, membranas lipídicas, y demás componentes necesarios para el metabolismo básico. En un último paso, surgen las características que encontramos en las células actuales: traducción, enzimas protéicas, metabolismo complejo y la transcripción de la información genética a DNA, ácido nucléico de mayor estabilidad que el RNA.

El atractivo principal del mundo del RNA es que parece presentar una solución viable a la paradoja del “huevo y la gallina”; como el RNA puede funcionar como catalizador y almacén de información, con lo cual se rompe el círculo de requisitos. En teoría, no es necesario un sistema complejo de traducción surgido espontáneamente para establecer la función dual de los sistemas vivos, funciones enzimáticas y almacenamiento de información. En ocasiones, y en vista de los problemas presentados por la síntesis abiótica de las cantidades suficientes de los elementos que componen el RNA, se han propuesto varias moléculas que pueden almacenar información, precursoras del RNA, como se verá más adelante.

Una característica importante de la teoría del mundo del RNA es la necesidad lógica de continuidad entre el primer RNA autorreplicante y todas las formas de vida a las que dio origen. Dado que la resolución de la paradoja del huevo y la gallina planteada por estas teorías se basa en el almacenamiento de información de la propia molécula catalítica, si dicha información se pierde la resolución deja de tener sentido. Esto no quiere decir que podemos encontrar a la secuencia autorreplicadora en el genoma de todos los seres vivos; quiere decir que hay una continuidad de linaje entre las poblaciones de moléculas del mundo del RNA y los seres actuales.

Problemas químicos del mundo del RNA

Las críticas al mundo del RNA se basan principalmente en la dificultad que existe para sintetizar de manera prebiótica los componentes necesarios para producir ácidos nucleicos, y la especificidad de la polimerización de los ácidos nucleicos. Para comprender cabalmente las dificultades, es preciso analizar un poco la estructura de los ácidos nucleicos que actualmente utilizan los seres vivos.

Los ácidos nucleicos son polímeros de nucleótidos. Las tres unidades básicas de éstos son una base nitrogenada, una pentosa y un grupo fosfato, como se muestra a continuación (Alberts 2002)

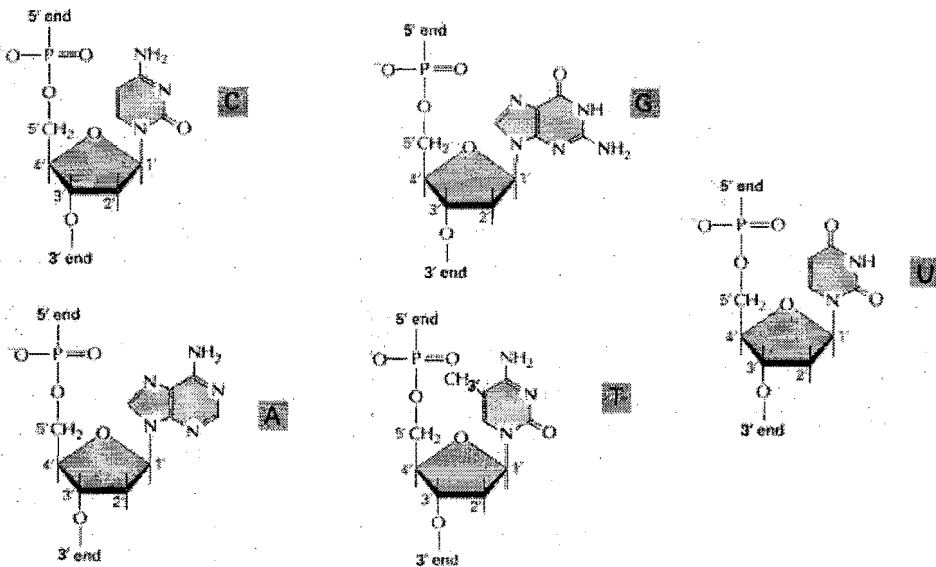


Fig. 9 - Los 5 nucleótidos más utilizados por los seres vivos (modificado de Alberts, 2002)

Las pentosas de los ácidos nucleicos utilizadas por prácticamente todos los seres vivos actuales son la ribosa (en el RNA) y la desoxirribose (en el DNA). En 4 de los 5 carbonos de dichas pentosas se pueden señalar características importantes:

- C1: Unión a una de las 5 bases nitrogenadas que utilizan los seres vivos (A, T, G, U, C)
- C2: Determina si la pentosa es ribosa, si está unido a un grupo hidroxilo, o desoxirribose, si está unido a un átomo de hidrógeno.
- C3: Se une con un enlace fosfodiéster al siguiente nucleótido.
- C5: En este carbono se encuentra la unión al grupo fosfato del nucleótido.

Los polímeros de DNA utilizan 4 tipos de bases nitrogenadas: dos purinas (adenina y guanina) y dos pirimidinas (citosina y timina). Estas bases forman los llamados enlaces Watson-Crick, que son puentes de hidrógeno entre las bases A-T y entre las bases G-C.

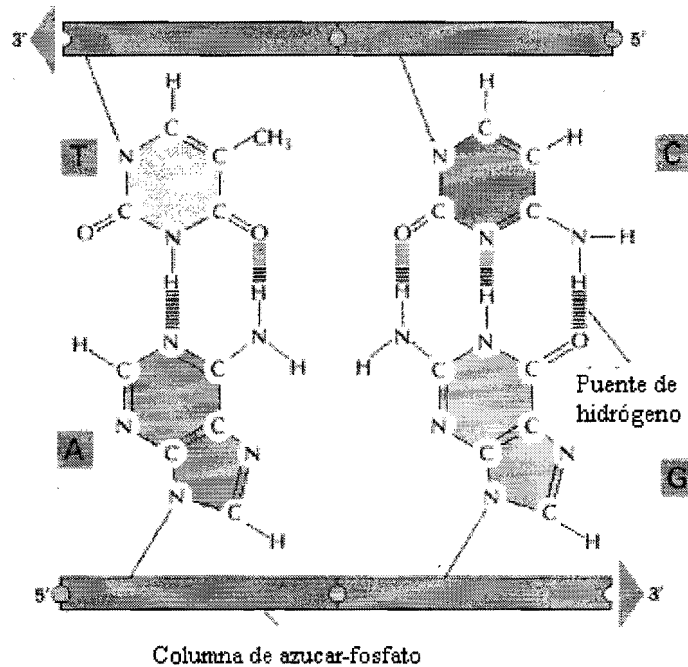


Fig. 10 - Apareamiento tradicional entre nucleótidos del DNA (modificado de Alberts 2002)

Por otro lado, los nucleótidos monoméricos tienen que ser “activados”; es decir, tienen que estar unidos en el carbono 5 a un grupo trifosfato. En cuanto se polimeriza, este enlace altamente energético se hidroliza, permitiendo la formación de un enlace covalente y liberando un pirofosfato. A continuación se mencionan de manera somera los principales problemas de la síntesis prebiótica de RNA y algunas de las soluciones propuestas;

Síntesis prebiótica de nucleótidos

- Formación prebiótica de azúcares: El azúcar que forma parte de la columna del RNA (la ribosa) es particularmente inestable. Su vida media es reducida y tiene un grado alto de reactividad debido al grupo OH que se encuentra en el carbono 2'; esto, por supuesto, es una desventaja para su permanencia en el medio pero es justamente la característica que la mayoría de las veces le permite al RNA funcionar como enzima.

Es posible obtener ribosa con la llamada síntesis de Butlerow (Orgel 2004), pero dicha reacción produce una mezcla compleja de azúcares, de los cuales la ribosa sólo es una pequeña parte. Además, al ser químicamente tan similares, dichos azúcares no pueden ser separados más que mediante métodos químicos de alta eficiencia (Shapiro 1988). Aunque se han reportado diversas modificaciones al sistema, Schwarz (Schwarz 1999, Schwarz y de Graaf 1993) afirma que no producen una selectividad aumentada del azúcar de interés.

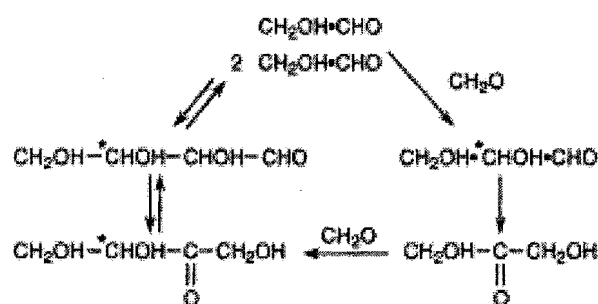


Fig. 11 - El ciclo hipotético y autocatalítico de la formosa (Orgel 2004)

- Formación prebiótica de bases: se han reportado diversos tipos de síntesis prebióticas para las cuatro bases nitrogenadas del RNA. La más mencionada es la síntesis de adenina mediante la polimerización de HCN (ácido cianhídrico). Tal síntesis es de importancia extrema, pues el HCN es el compuesto más producido por los experimentos tipo Miller (descargas eléctricas a través de una atmósfera reductora de metano y nitrógeno). Sin embargo, dichas reacciones se llevan a cabo en condiciones extremas y en el mejor de los casos la cantidad producida no rebasa el 0.1% (Schwarz 1998). De la misma manera, se han modelado mecanismos de reacción para las otras bases, pero generalmente utilizan condiciones ambientales y concentraciones tan altas de sustrato que difícilmente son representativas del ambiente prebiótico.

Polimerización específica del RNA

El RNA no sólo es problemático en cuanto a su complejidad monomérica; su polimerización implica una serie de problemas que no son fáciles de resolver. De hecho, cualquier polímero informativo tiene que ser homogéneo en varios sentidos: en las partes de la "columna" de polimerización, en la quiralidad de sus elementos y, en caso de tener varias alternativas, que los enlaces se formen en el mismo átomo. En el caso de

los ácidos nucleicos, el enlace debe de formarse siempre en los carbonos 3' y 5'; además, los azúcares tienen que ser D-ribosas (o D-desoxirribosas). En caso de incorporar un elemento de quiralidad distinta o en un átomo incorrecto (en el sentido 5'-2', o en 5'-5', por ejemplo, cosa que ocurre frecuentemente en la polimerización no específica del RNA), la adición de nucleótidos a la cadena es interrumpida. Por supuesto, los ácidos nucleicos no son los únicos que presentan este problema: en el caso de incorporar un D-aminoácido a una cadena peptídica, ésta queda bloqueada y no puede seguir polimerizándose.

Asimismo, hay moléculas que pueden incorporarse a la cadena e interrumpir la polimerización, aunque sean de la quiralidad correcta. Tal es el caso de ácidos carboxílicos en los péptidos, o cualquier compuesto que contenga un grupo hidroxilo reactivo en el caso de los ácidos nucleicos.

Algunas soluciones propuestas

Debido a tales dificultades, se ha propuesto un mundo del pre-RNA, que considera que un ancestro del RNA, de estructura química distinta. Una cantidad importante de publicaciones propone un ancestro que utiliza las mismas bases nitrogenadas, pero cuya "columna" (la parte de la molécula que se polimeriza y que actualmente se forma con ribosa y fosfatos) es de naturaleza química menos compleja. Algunos de los pre-RNA más investigados son más simples químicamente; los ácidos nucleicos peptídicos (PNA) tienen una columna simple similar a las de las proteínas, con una alternancia de grupos amino y carboxilo, lo cual significa que no necesitan fosfatos para polimerizarse. Además, los PNA son aquirales por lo que se simplifica el mecanismo de polimerización. También se han investigado los TNA, ácidos nucleicos con treosas, azúcares que aunque tienen sólo cinco átomos de carbono pueden formar heterohélices con cadenas de RNA. (Orgel 2004). Otros polímeros propuestos se basan en aminoaldehidos, ditioles, hidroxiácidos, etc. (Joyce y Orgel, 1992)

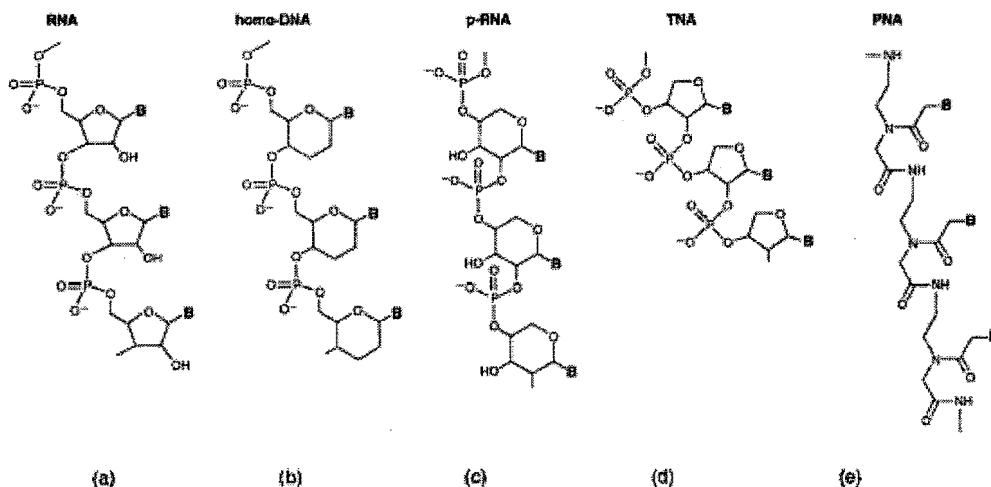


Fig. 12 - El RNA y distintas alternativas prebióticas propuestas (Orgel 2004)

Por supuesto, todas las dificultades para la síntesis de los ácidos nucleicos (en particular el ribonucleico) no son exclusivas de las teorías del mundo del RNA. Tanto el RNA como el DNA son las bases del funcionamiento de los seres vivos actuales, y su surgimiento debe de ser explicado cabalmente en cualquier modelo completo de origen de la vida, sin importar si éste se basa en el surgimiento inicial del genoma, del metabolismo o algún otro sistema. De nuevo, la síntesis prebiótica de los distintos componentes no es el punto central de esta tesis, por lo que sólo se ha presentado un análisis somero de las dificultades.

Autoorganización: “Metabolismo primero”

El otro grupo importante de teorías que tratan de explicar el origen de la vida son las que proponen el surgimiento de un sistema metabólico²¹. Las teorías que proponen el surgimiento de un metabolismo rudimentario son mucho más variadas que sus contrapartes enfocadas en la replicación. La propuesta de la replicación es lineal: una molécula autocatalítica que produce copias de sí misma. Las propuestas metabólicas tienen inherentemente un campo más amplio de acción: el metabolismo es mucho más variado químicamente que las reacciones de la biología molecular. Se puede proponer el surgimiento del metabolismo a través del surgimiento prebiótico de aminoácidos; es decir, analizando las propiedades de aminoácidos y oligopéptidos para llevar a su vez

²¹ Tal división (replicación primero vs. Metabolismo primero) es natural y lógica. Los dos grupos de teoría representan las dos características más impactantes de la vida: el manejo de información y de energía.

reacciones cada vez más complejas. Se puede, también, realizar un acercamiento a través de la bioenergética, y analizar la posibilidad de formación de moléculas que puedan funcionar como una moneda energética prebiótica. Otro camino posible es la formación prebiótica de lípidos, para formar membranas y compartimentos que permitirán elevar la concentración de algún metabolito para llevar a cabo reacciones que en una solución poco concentrada no serían posibles. Finalmente (aunque probablemente esta lista no sea exhaustiva), se pueden tomar las rutas metabólicas modernas más importantes (por ejemplo, el ciclo de Krebs, la fotosíntesis, la biosíntesis de nucleótidos y aminoácidos) y proponer distintos mecanismos para llevarlas a cabo sin las enzimas actuales, esos catalizadores de alta eficiencia. Una teoría plausible es la propuesta de Wächtershäuser, que, aunque ha sido criticada (Orgel, 2000; de Duve, 1991), se sostiene como una solución relativamente sencilla al problema del origen del metabolismo. Además, se pueden realizar propuestas con un alto grado de abstracción, en la que se pueden representar a los metabolitos con signos y enfocarse en las relaciones de grupos abstractos de moléculas, y no tanto en las propiedades físicas de moléculas “reales”. Dichas teorías se discuten más adelante, en la sección de “metabolismos abstractos”.

Teorías de Wächtershäuser y de Morowitz

Una de las teorías que proponen el metabolismo primero es la de Günther Wächtershäuser (Wächtershäuser 1988, 1990, 1994). Su tesis central es que la vida inició como un sistema químico unido aniónicamente a una superficie de carga positiva, por ejemplo pirita. Las moléculas están constantemente sometidas a una presión selectiva que favorece a las que tienen más grupos con cargas negativas. Los polímeros que cuenten con muchos grupos fosfato, por ejemplo, tendrán más estabilidad que moléculas pequeñas con cargas débiles. Esta teoría contrasta con las teorías de Oparin y Haldane: para Wächtershäuser, la concentración de “bloques de construcción biológicos” en solución es despreciable, por lo que una molécula que se separa de la superficie se pierde irremediablemente. Para el autor, esta es la razón por la cual una gran parte de los metabolitos importantes de los seres vivos actuales tienen cargas fuertes negativas (por ejemplo, grupos hidroxilo, fosfatos, sulfuros, etc). También explica la carga negativa en la cara interna de la membrana bilipídica.

Una de las partes más importantes de la teoría del metabolismo de superficie es la propuesta para explicar origen del ciclo del ácido cítrico, la cual incluye la fuente de energía que se utiliza en estos procesos. El ciclo del ácido cítrico, también llamado ciclo de Krebs o de los ácidos tricarbóxicos, es probablemente el punto central del metabolismo de los seres vivos actuales. No sólo es el mecanismo bioquímico por el cual se genera poder reductor (en forma de $\text{NADH}+\text{H}^+$ o FADH_2) para obtener ATP en el proceso de respiración celular, sino que muchas de las moléculas que intervienen en él son utilizadas para la biosíntesis de diversos compuestos orgánicos, como bases nitrogenadas, aminoácidos, lípidos, etc.

Wächtershäuser propone una versión reductora (es decir, en reversa) del ciclo de Krebs en un ambiente termófilo, como una ventila hidrotermal, y con una bioquímica de superficie. En vez de producir energía y liberar átomos de carbono en forma de CO_2 , la dinámica del ciclo de ácidos tricarbóxicos en reversa tomaría energía de la oxidación de la superficie para llevar a cabo las reacciones, y su función principal sería fijar carbono para sintetizar ácidos tricarbóxicos más complejos y con enlaces energéticos. Una evidencia importante para esta teoría es que la fijación del carbono en organismos de ecosistemas hidrotermales se lleva a cabo gracias al funcionamiento de un ciclo de ácidos tricarbóxicos en reversa (rTCA por sus siglas en inglés), utilizando para un ciclo completo 4 o 5 moléculas de ATP para solventar el gasto energético (Williams, 2006). Por supuesto, el rTCA actual depende completamente de catalizadores enzimáticos, y el modelo prebiótico de Wächtershäuser debe prescindir de ellos.

El modelo de Wächtershäuser es un eco de las preocupaciones que expresa el autor acerca de los posibles problemas de la llamada “sopa primigenia”. Si consideramos la difusión de los compuestos en solución, la sopa primigenia tiene que estar restringida a volúmenes relativamente reducidos y altas concentraciones, por lo que el autor responde creando un modelo de superficie en el que al romperse el enlace iónico que une al ácido con la superficie mineral de carga positiva, tal molécula se considera perdida en tanto no exista una compartimentalización primitiva. Por otra parte, Wächtershäuser señala su desconfianza hacia la inductividad de la teoría de la sopa: “tan pronto como se encontraban trazas de un compuesto nuevo, se afirmaba que éste estaba también en el caldo”, aunque “las condiciones químicas de algunas reacciones prebióticas eran incompatibles con las condiciones de otras” (Wächtershäuser 1997). Aunque se han demostrado experimentalmente los distintos

pasos de las reacciones primarias de esta teoría por separado, nunca se ha logrado reproducir un ciclo de reacciones. Esto no es inusual: es difícil predecir las condiciones exactas de la organización de un sistema complejo. Además, el ambiente que necesitan los experimentos (anaerobiosis y condiciones elevadas de presión y temperatura) hacen costoso y complicado llevar a cabo investigaciones experimentales.

Por su parte, la propuesta que Harold Morowitz (Morowitz 2000, Smith y Morowitz 2005) ha realizado a lo largo de sus escritos no es un modelo químico en sí; tal vez sea más pertinente verlo como un complemento a las propuestas de metabolismo primero como las de Wächtershäuser. Morowitz señala que el centro del metabolismo, el ciclo de Krebs, es universal a todas las especies. Sin embargo, actualmente tal ciclo funciona como una especie de transducción de energía, tomando compuestos energéticos y degradándolos, produciendo así especies químicas más simples, energía en forma de compuestos trifosfatados y especies reactivas ($\text{NADH}+\text{H}^+$ y FADH_2), que serán utilizadas en la respiración celular. Morowitz y su grupo también insisten en los problemas que plantea el modelo de Oparin-Haldane (difusión en solución acuosa, concentración de reactivos irreal en condiciones prebióticas); además, agregan que los modelos de sopa prebiótica aceptan que la vida pudo haber surgido en un ambiente enriquecido en moléculas orgánicas *por reacciones fundamentalmente distintas a las del metabolismo actual*. Tal suposición no sólo aumenta la complejidad irreducible de los modelos (se tienen que buscar rutas que adapten esas químicas distintas a la bioquímica actual), sino que no tienen ningún poder explicativo porque no restringen la química prebiótica.

La propuesta de Morowitz, pues, parte de tres hechos fundamentales: i) Los once componentes del ciclo de Krebs son un núcleo único del metabolismo terrestre, pues de sus elementos surgen prácticamente todas las moléculas de la bioquímica actual; ii) la secuencia de las reacciones se da de manera oxidativa en los organismos fotoautótrofos y heterótrofos, pero se presenta en reversa (reductivamente) en varios quimiolitautótrofos (organismos que derivan su energía de reacciones químicas inorgánicas); y iii) existe evidencia geofísica que sugiere que la Tierra primitiva, antes del surgimiento de la fotosíntesis, tenía ambiente reductor al menos en ciertas zonas. Los autores afirman que en las condiciones propuestas no se favorece la degradación de los reactivos del ciclo que son relativamente complejos, siendo el acetato y el succinato los dos puntos de menor energía libre de formación (Smith y Morowitz 2004). En el

mismo artículo, afirman que el ciclo reductivo de ácidos tricarboxílicos (rTCA) tiene características que favorecen su aparición y organización espontánea en ambientes prebióticos: es autocatalítico mediante una serie de pasos relativamente corta (lo cual permite que aparezca en ambientes de baja complejidad relativa); es redundante en algunos puntos (como por ejemplo, la síntesis de oxaloacetato por medio de piruvato o citrato); y todas las reacciones son cinéticamente de primer orden, con lo cual se evita la necesidad de barreras de difusión para favorecerlas.

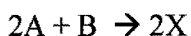
Tanto el metabolismo de Wächtershäuser como la propuesta de Morowitz se basan en la posibilidad de llevar a cabo el ciclo de ácidos tricarboxílicos de manera reductiva, lo cual no sólo es posible sino que se puede encontrar en varios organismos, ya sea de manera constitutiva o facultativa (Williams 2006). Además de los géneros en los que se ha encontrado (*Proteobacteria*, *Chlorobium* y *Desulfobacter*, por mencionar algunos), se han encontrado diversos genes del ciclo en bacterias de vida libre de los ambientes hidrotermales (Campbell 2004).

El acercamiento experimental es complejo: las propuestas de metabolismo primero generalmente requieren condiciones extremas de temperatura y presión, y por no existe un cuerpo de estudios de laboratorio tan amplio como el que tiene la teoría de Oparin-Haldane. Además, mientras que los estudios de la sopa prebiótica se basan en la presencia de los compuestos, la preocupación principal de los autores de los ciclos químicos prebióticos es justamente la organización de un proceso completo. Sin embargo, varios puntos de ésta han sido demostrados en condiciones similares a las de las ventilas hidrotermales²². Es innegable que la escuela de experimentos de síntesis prebióticas ha aportado información invaluable acerca de las posibilidades químicas de ciertos sistemas; sin embargo, los autores mencionados en este apartado señalan problemas clave que persisten. Los modelos que presentan tanto Morowitz y Wächtershäuser son acercamientos alternativos a un problema antiguo, y probablemente aporten buena parte de las herramientas para su resolución; al menos, son útiles para re-evaluar varios puntos de vista que se consideran implícitos cuando se discute el origen de la vida.

²² Ver, por ejemplo, los trabajos de Cody (2000, 2001, 2004 y 2005); son de particular interés porque, además de demostrar la plausibilidad de las reacciones del ciclo rTCA en condiciones de presión y temperatura elevadas, conectan las propuestas de Morowitz con las de Wächtershäuser mostrando que los sulfuros de metales de transición (en particular, FeS) pueden funcionar como catalizadores minerales de dichas reacciones.

Metabolismos abstractos

Como ya se ha dicho, es posible proponer metabolismos sin especificar las moléculas y las reacciones de las que se habla. Esto podría parecer contradictorio o inútil, pero un ejemplo cotidiano muestra su utilidad *como modelo*. Consideremos la ecuación de una reacción química cualquiera:



Tal expresión, sin hacer evidente la naturaleza de los reactivos, proporciona información de un tipo posible de reacción (hago notar que tal fórmula funciona como una abstracción de la reacción de formación de agua, entre otras reacciones: $2H_2 + O_2 = 2H_2O$). Los mecanismos de metabolismo abstracto funcionan de manera similar: exponen cierto tipo de reacciones y permiten evaluar la plausibilidad de la realización del modelo, dejando abierta la posibilidad de utilizar distintos metabolitos.

Es importante mencionar que estos modelos son simplificaciones para evaluar el comportamiento dinámico de una población de partículas. Nuestro conocimiento de la química de partículas elementales y el papel exacto que toman en las reacciones es limitado. Además, nuestra capacidad de computación hace que llevar a cabo iteraciones múltiples con un número tan grande de elementos sea poco factible. En este sentido, los modelos abstractos y el diseño tradicional de experimentos son similares. En ambos casos, se trata de reducir las variables para evaluar el comportamiento de un fenómeno cualquiera y establecer lo más certeramente posible una relación de causalidad entre dos eventos. Los experimentos de laboratorio no son nunca una copia fiel del comportamiento del objeto estudiado en condiciones no controladas, de lo que ocurre en la "Naturaleza".

A continuación se presentan de manera somera dos modelos de metabolismos abstractos de particular interés.

- "La Cristalización de la vida"

Según Stuart Kaufmann (Kaufmann 1993), el surgimiento de la vida es una característica esperada de la materia, materializándose como una propiedad de un colectivo molecular heterogéneo; lejos de ser un evento único y aislado, como se ha afirmado muchas veces, es incluso probable.

La lógica de su modelo implica un conjunto de moléculas con dos propiedades catalíticas aleatorias: unión y corte de otras moléculas. Kaufmann demostró que al

incrementar la complejidad de las moléculas (dada en el número de monómeros que las forman), las conexiones aumentan y la red completa se vuelve autocatalítica.

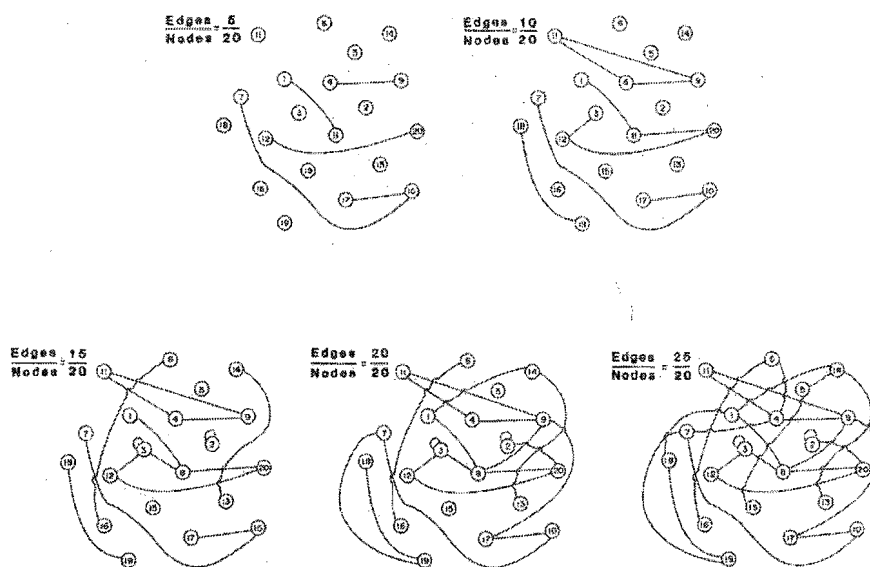


Fig. 13 - Surgimiento de ciclos en grafos aleatorios (Kauffman 1993)

En la figura se observan grafos aleatorios demostrando el principio funcional de los ciclos autocatalíticos de Kauffman; sin importar el orden de las conexiones, cuando la relación aristas/nodos es mayor a uno, comienzan a surgir ciclos de tamaños diversos. Kauffman relaciona los nodos con los posibles polímeros compuestos por, en el caso más simple, dos tipos de monómeros, y las aristas con las reacciones de polimerización que dan lugar a tales secuencias. El resultado es el mismo: después de que el número de conexiones rebasa un umbral crítico, espontáneamente aparecen ciclos autocatalíticos como los que se han propuesto químicamente.

- AIChem (Fontana, 1991)

El modelo de Walter Fontana, llamado AIChem se basa en estructuras puramente gramaticales en forma de funciones matemáticas, y se compone de varias partes:

- 1) Un universo definido por el cálculo lambda, en la que las funciones matemáticas representan objetos reales;
- 2) Interacciones entre estos objetos; es decir, una función matemática puede representar un objeto o una instrucción;
- 3) Reglas de colisión, que comprenden principalmente lo que ocurre con los objetos después de que interactúan (si desaparecen o perduran); lo que ocurre con el producto

de la interacción (es decir, su función y su actividad); y el límite computacional del modelo;

4) Un sistema que determine el comportamiento de las “partículas” matemáticas en su conjunto; Fontana define un mapa de iteraciones definidas o alternativamente un “gas de Turing”, que es un proceso estocástico que crea dinámicas adicionales.

Después de que Fontana pone a funcionar su modelo, éste crea redes de interacciones de cooperación, que eventualmente se cierra. Por lo tanto, el autor argumenta que existen ciertas propiedades emergentes inevitables de los sistemas en los que interactúan partículas múltiples: el surgimiento de hiperciclos auto-replicadores, comportamientos de auto-organización y un aumento de orden y complejidad. En cierta manera, el modelo de Fontana es un reflejo puramente matemático del modelo químico propuesto por Kauffman.

Aunque los resultados de Fontana predicen un sistema autoreplicador, no quiere decir que éste sea compuesto por ácidos nucleicos. En todo caso, abogaría por lo contrario: que aún sin la propiedad replicativa un sistema complejo se puede organizar en una red o un ciclo autocatalítico sin necesidad de ningún depósito informativo.

Problemas conceptuales de las teorías de Replicación Primero

Ambos tipos de teorías, sean de replicación o de metabolismo primero, presentan problemas graves. Sin embargo, el mundo del RNA presenta fallas aparentemente irresolubles que no son de naturaleza química. Tales fallas se hacen evidentes cuando se analiza la continuidad lógica necesaria; la aparente resolución de la paradoja del huevo y la gallina pierde gran parte de su atractivo cuando éstas se observan cuidadosamente.

Replicación específica.- El mundo del RNA siempre comienza, teóricamente, con una molécula que se puede copiar a sí misma. En esto reside todo su atractivo: una molécula catalizadora que puede guardar información. La catálisis es necesariamente la autorreplicación: una molécula de RNA que catalice, por ejemplo, la síntesis de glicina no tiene ningún potencial para ser seleccionada. Sin embargo, ¿qué clase de mecanismo permitiría una replicación *específica* de las moléculas autorreplicadoras? Es decir, si una molécula funciona como una polimerasa, tenderá a copiar todas las moléculas que estén a su alrededor si no tiene una dirección compleja como la que proporciona el

ambiente celular actual (y más aún si tal polimerasa es primitiva y no específica). Es improbable que una polimerasa surgida del azar cuente con mecanismos como un sitio de iniciación de replicación y un sitio activo que reconozca tal sitio de iniciación. Además, los problemas son mayores si consideramos el mecanismo de replicación de los ácidos nucleicos, basado en la complementaridad de bases. Si se utiliza un ácido nucleico como templado o molde para producir otro, se obtiene, además de la molécula original, una complementaria pero *de ninguna manera similar*, con excepción de la longitud de ambas. Las propiedades físicas de ambas moléculas serán radicalmente distintas. Tales diferencias se hacen más evidentes aún si se considera que el apareamiento tradicional de los ácidos nucleicos se lleva a cabo de manera antiparalela. Una molécula que sea “autocatalítica” tiene que reconocer no una, sino dos entidades físicas completamente distintas para llevar a cabo un ciclo de replicación completo.

Catástrofes.- Existen dos puntos críticos en la reproducción de moléculas autocatalíticas. El primero, el punto crítico de fidelidad de la polimerización para que ésta produzca moléculas funcionales con la suficiente frecuencia para evitar que el sistema completo degenere en caos y desorden. El segundo punto crítico es el surgimiento de moléculas parasíticas, moléculas que por alguna razón u otra son replicadas por el sistema prebiótico y sin embargo no toman parte en ninguna de las reacciones necesarias de mantenimiento. Tales formas parasíticas pueden funcionar de diversas maneras: acortando el ciclo reduciendo su complejidad y por tanto su repertorio de reacciones, o simplemente produciendo moléculas no catalíticas que sólo acaparan recursos.

Surgimiento de genes.- Sin embargo, aceptemos por un momento el escenario típico del mundo del RNA. Tendríamos un volumen determinado (de la magnitud y características que sean) lleno de una solución de autorreplicadores (tanto la molécula catalítica como su molde complementario). ¿Qué evento sería el siguiente en la secuencia de evolución molecular? ¿Cómo surgen los nuevos genes, por ejemplo los necesarios para la síntesis de pentosas o bases nitrogenadas? Como señalan Trevors y Abel, (Trevors 2004), las secuencias *per se* no tienen ningún valor de selección. Además de aumentar el repertorio de reacciones útiles, las nuevas moléculas de RNA deberán (al azar) tener

algún inicio de replicación, algún módulo que las haga propensas a ser replicadas más frecuentemente que las demás secuencias, igualmente azarosas pero no catalíticas.

Impedimentos estéricos .- Consideremos por un momento la replicación moderna del DNA. Existe una maquinaria compleja de enzimas, con particular énfasis en la helicasa, que sirven para “estirar” el DNA. En efecto, para que una enzima utilice como molde una cadena de DNA ésta debe de haber perdido toda estructura secundaria, asumiendo una forma prácticamente lineal.²³ El problema fundamental aquí no es precisamente la ausencia de complejos protéicos que preparen al ácido nucléico prebiótico para su replicación (aunque es un problema secundario importante), sino es la refutación física de la resolución de la paradoja del huevo y la gallina. Una molécula de RNA puede servir *teóricamente* como un almacén de información y como una molécula enzimática. Sin embargo, estas dos funciones requieren conformaciones diametralmente distintas. Una molécula de RNA catalítico depende de una conformación secundaria, que a su vez es causada por el apareamiento de bases nitrogenadas al interior de la molécula. Dicho sea de paso, tal conformación tridimensional es, generalmente, la conformación que espontáneamente se generará por ser la que se encuentra en un mínimo energético. Para que una molécula sea un almacén funcional de energía, sin embargo, necesita estar completamente desnaturalizada; es imposible acceder completamente a la información de una molécula de RNA plegada. Aunque existen condiciones físicas que hacen posible la desnaturalización efectiva del RNA (como por ejemplo, temperaturas elevadas o pH extremo), es necesario tener en cuenta que para que el modelo del mundo del RNA se lleve a cabo las dos formas tienen que coexistir en un mismo tiempo y espacio (Benner 1999).

Tal problema se relaciona en cierta manera con el problema que tiene una molécula para ser realmente autorreplicante. En ocasiones se asume que la molécula originaria del mundo del RNA tiene la capacidad de copiarse a sí misma. Esto, en sí, es una imposibilidad física, pues esta molécula no sólo tendría que estar simultáneamente en una conformación tridimensional (catalítica) y lineal (informativa), sino que el sitio activo tendría que actuar sobre sí mismo. Tal cosa representa un problema

²³ De hecho, en los casos en que existe algún tipo de estructura secundaria, se utiliza como una manera de regulación genética, como por ejemplo el caso del operón del triptófano y etc, en la que la formación de loops y hairpins en el ácido nucléico interfieren con el funcionamiento de las enzimas que se unen a él, o incluso causan su disociación interrumpiendo la replicación, traducción o transcripción (según sea el caso).

topológico irresoluble (al menos en las tres dimensiones en las que existimos), pues, para ponerlo en los términos del modelo clásico de enzima y sustrato, el sitio activo tendría que tener forma de llave y cerradura simultáneamente. Sería como tratar de meter una caja dentro de sí misma. Las soluciones que se han buscado a este punto son aún más implausibles que el surgimiento azaroso de polimerasas de RNA. Por ejemplo, se ha propuesto que por azar o por complementariedad (von Kiedrowsky 1986) esta primera ribozima catalice la formación de copias de sí misma sin necesitar un molde; tal serie de pasos se limita a moléculas estrictamente palíndromas, y su actividad catalítica se limita a ligar dos fragmentos complementarios que se unan a ella con puentes de hidrógeno.

Continuidad. – El más grave problema, a mi parecer, es el de la ruptura de continuidad como única salida del mundo del RNA, sin importar la versión de esta teoría que se esté manejando. Surge cuando se considera que las moléculas catalizadoras son fundamentalmente distintas en su naturaleza. Probablemente exista la posibilidad de recrear una célula completamente funcional dependiendo únicamente de ribozimas (los RNAs catalizadores parecen ser asombrosamente versátiles). Sin embargo, es un hecho innegable que los seres vivos actuales no funcionan así. Toda teoría del mundo del RNA implica que de una manera u otra, tiene que haber una etapa de transición hacia el mundo de DNA-proteína. Hasta ahora, no existe ningún método que permita, como en ocasiones se ha sugerido, una transición lenta de la información lineal del RNA a proteína. Se ha propuesto, tomando en cuenta la llamada “degeneración” (o, como le llamaré preferentemente en adelante, la redundancia) del código genético, que éste comenzó intercalando una base extra por cada base de la secuencia original, dando lugar a un código de codones de dos nucleótidos. Después, según la propuesta, se intercalará una base más entre cada diplete, dando lugar al lenguaje de tripletes que existe hoy en día. Como se puede apreciar, tal mecanismo no sólo es extremadamente complejo, sino que no tiene ninguna base (ni actual ni como una posibilidad química) para presuponerlo.

Sin embargo el problema de una transcripción de un RNA funcional a un mRNA de tripletes no es la improbabilidad de encontrar un mecanismo plausible para hacerlo, sino que *tal transcripción sería completamente inútil*. La funcionalidad de la cadena de RNA está determinada por su propia secuencia, y por la estructura secundaria causada por ésta. Una ribozima que catalice, por ejemplo, la síntesis de la ribosa, no codificará,

utilizada como mRNA, para una proteína funcional. La probabilidad de que una enzima de RNA codifique para una enzima proteica cualquiera es despreciable, y estas dos características presentes en una cadena de RNA (función como ribozima y como mRNA) no tendrán ninguna conexión causal. La mayor implicación de esta observación a primera vista obvia es la pérdida completa de la información metabólica del mundo del RNA al surgir el mundo de proteínas, información supuestamente surgida al azar por selección natural. Al surgir el mundo de proteínas todas las secuencias de RNA que existían hasta el momento se harán inmediatamente obsoletas, y se tendrá que empezar nuevamente a construir, en *tabula rasa*, los algoritmos metabólicos de los ácidos nucleicos, esta vez en el formato de tripletes. Es justo en este punto en el que la solución de la paradoja del huevo y la gallina pierde su atractivo: no resolvió un problema de una hipótesis; simplemente lo trasladó de un punto (el surgimiento de las primeras moléculas informativas) a otro posterior (el surgimiento del formato actual de información). Esta idea no es nueva: fue mencionada por Crick (1968) cuando descartó el surgimiento de códigos genéticos basados en nucleótidos individuales (es decir, cuatro "codones"), luego el paso a dinucleótidos (16 "codones") y finalmente trinucleótidos (los 64 codones del código conocido), argumentando que tal paso rompería el Principio de Continuidad necesario para la transmisión de información.

Justificación de la propuesta: La necesidad de teorías alternativas

Como se ha visto, el mundo del RNA parece resolver ciertos problemas conceptuales pero plantea otros igualmente complicados. En cierta manera, esta teoría resuelve simultáneamente el surgimiento de la catálisis y la información, pero crea dos conjuntos de interrogantes, flanqueándola cronológicamente: el primero, el conjunto de problemas *pre-RNA*: si el metabolismo surgió con una serie de ribozimas, ¿cómo se fabrican los compuestos necesarios para polimerizar dichas moléculas? El estudio del surgimiento de nucleótidos por medio de síntesis prebiótica ha sido poco fructífero, y los resultados que se han logrado requieren condiciones controladas y específicas, que difícilmente se encontrarían en el mundo prebiótico (véase Orgel 2004 para una reseña). El segundo conjunto es el reseñado en la sección anterior: es difícil imaginar un mecanismo de paso de un sistema complejo únicamente compuesto por reacciones ribozimáticas a uno con similitudes superficiales. Ni la maquinaria de traducción de

los seres vivos ni los escenarios *in vitro* de los experimentos modernos logran traducir la función, sino simplemente crear un sistema simbólico entre dos elementos. Una traducción simbólica de un carácter de un sistema a otro que además resulte en una traducción de función es altamente improbable.²⁴

¿En qué reside el problema del huevo y la gallina? El fondo de la cuestión es que el sistema de traducción es extremadamente complejo, y además cada una de sus partes parece estar hecha específica y expresamente para la función que desempeña. Los tRNA, las aminoacil-transferasas, los ribosomas e incluso el lenguaje de tripletes forman en conjunto una maquinaria precisa, y la ausencia de una de los engranes conlleva la inutilización instantánea de todos los demás. Parece complicado explicar el surgimiento gradual y secuencial de los elementos, pero es intuitivamente imposible que todos los elementos hayan surgido *de novo*, de un momento a otro, con sus funciones predeterminadas, o que hayan surgido independientemente y que por algún capricho probabilístico se engarzen de manera perfecta en una maquinaria que, además, es fundamental para todas las formas de vida conocidas.

Trevors (2004) afirma que los seres vivos presentan tres problemas principales, cuando se trata de estudiar su origen. Los problemas que presentan la naturaleza de las primeras moléculas protéicas, informacionales, el surgimiento del metabolismo, las características ambientales en que surgió, la dicotomía superficie-solución, etc, son a mi manera de ver problemas fundamentales. Sin embargo, los tres problemas que señala Trevors son, por mucho, los más desconcertantes y enigmáticos, e incluso herméticos. Los tres problemas forman una especie de paradoja del huevo y la gallina tripartita, pues cada uno debió de surgir independientemente de los demás, y sin embargo ninguno tiene sentido ni uso sin los otros dos. De hecho, la relación entre los tres problemas es la causa profunda de la paradoja del huevo y la gallina.

²⁴ Tal afirmación se sostiene en cualquier sistema simbólico, como la traducción de un lenguaje a otro. Nabokov provee un ejemplo literario en los comentarios del poema *Pale Fire*: un periódico ruso publica un recuento de la coronación del Zar, y la palabra *korona* (corona) aparece con un error tipográfico: *vorona* (cuervo). Al día siguiente, en una fe de erratas, de nuevo se comete un error, esta vez escribiendo *korova* (vaca). Nabokov compara la serie rusa (*korona-vorona-korova*) con su traducción inglesa: *crown-crow-cow*. Es evidente que además de conservar su significado, la serie de palabras conserva su relación tipográfica, separadas unas de otras por sólo un carácter. “La probabilidad de esta doble coincidencia”, dice Nabokov, “desafia al cómputo”.

El primer problema, según Trevors, es el surgimiento de las *instrucciones celulares*. Es decir, el surgimiento del “programa celular”, expresado en una secuencia de ácido nucléico, que da lugar a un cúmulo de información con significado semántico. El segundo es el establecimiento de un “sistema operativo” completo, cuya característica más fundamental es la dualidad DNA-proteína. Como se puede observar, tal dualidad se basa tanto en la existencia del aparato de traducción, como en la capacidad de replicación del DNA. El tercero es el surgimiento de un código como tal, que relaciona dos entidades completamente distintas de manera simbólica; tal código es la relación abstracta entre tripletes y aminoácidos, especificada por el código genético.

Aceptemos por un momento, pues, que las instrucciones celulares (los complejos y entrelazados algoritmos del DNA) surgieron poco a poco, gracias a la polimerización casual, los errores de copia y la selección (ya sea ésta llamada “natural” o “química”). Los dos problemas restantes (sin que esto signifique que el primero ha sido resuelto por esa ambigua concesión) tienen un punto en común: la relación triplete-aminoácido. Un programa de investigación que trate de esclarecer el origen físico de tal sistema probablemente ayudará a entender en gran medida el surgimiento del sistema de traducción y el establecimiento de tales sistemas de almacenamiento informativo. Ésta es la principal razón por la que este trabajo, dedicado a investigar el origen de la vida, se enfoque específicamente en el código genético. Es necesario buscar alternativas para resolver las complicaciones intrínsecas presentadas por las teorías tradicionales.

No es una coincidencia que en la sección anterior una de las características fundamentales de los seres vivientes residiera en las relaciones organizadas por el código genético, y que en esta segunda parte se considere el origen del código como uno de los puntos del origen de la vida. Lejos de ser una tautología, ambas consideraciones se derivan la una de la otra, y presuponen a la otra; en palabras de Stuart Kaufmann, este argumento cíclico no es un círculo vicioso, sino virtuoso.

2. El código genético

Regularidades y patrones

La tercera letra del codón: Redundancia

Para poder comprender el código genético y proponer una teoría acerca de su surgimiento, es necesario conocer ciertas propiedades que éste posee. Prácticamente desde el momento de su elucidación se comenzaron a visualizar ciertas regularidades que hacían evidente que su organización no era arbitraria. La primera de ellas, por supuesto, es la llamada "degeneración" del código (a la que en este trabajo se le llamará redundancia). Como se puede comprobar con sólo echando un vistazo a la tabla de codones, muchos de éstos se pueden agrupar en familias por sus dos primeras letras, siendo la tercera indiferente para la traducción. Por ejemplo, todos los codones que comienzan con CC codificarán para el aminoácido Prolina, sin importar la última letra. En otros casos, las agrupaciones se dan por parejas: lo que determina el aminoácido a codificar son las dos primeras letras del codón y la identidad química del último nucleótido (es decir, si éste es una pirimidina o una purina). Por ejemplo: los codones AAR codificarán para Lisina, y los AAY codificarán para Asparagina. Como ya se ha mencionado, esta característica es la propiedad más evidente que indica un orden en el código genético. Un acomodo aleatorio, incluso si fuera redundante, no produciría estas agrupaciones por familias. Francis Crick propuso poco tiempo después de la elucidación del código la hipótesis del bamboleo (Wobble), que sugiere que las primeras dos bases del codón son las más importantes para la codificación, y que el apareamiento de la última base es relativamente débil, por lo que permite una ambigüedad de la relación codón-anticodón. Tal hipótesis se comprobó posteriormente.

La segunda letra del codón: Hidrofilia

Una relación menos obvia es la que existe entre la segunda letra del codón y el aminoácido codificado. Por alguna razón, esta base se encuentra asociada a las propiedades fisicoquímicas de dicho aminoácido, en particular a su polaridad y en última instancia a su hidrofobicidad o hidrofilia.

Como se puede observar en la tabla 1, la segunda letra de los codones de los aminoácidos más hidrofóbicos es U; en el caso de los aminoácidos más hidrofílicos (es

decir, los aminoácidos polares con carga) la segunda letra es A. En el caso de los otros dos nucleótidos (G y C) la relación es un poco más ambigua, pues codifican para aminoácidos con hidrofobicidad intermedia. Existen dos grandes excepciones: la cisteína (hidrofóbica) y la arginina (altamente hidrofílica).

Tabla 1 – Índice de hidrofilia de los aminoácidos esenciales

Aminoácido	Índice de Hidrofobicidad	Codones usados
Isoleucina	4.5	AU*
Valina	4.2	GU*
Leucina	3.8	CU* GUR
Fenilalanina	2.8	UUY
Cisteína	2.5	UGY
Metionina	1.9	AUG
Alanina	1.8	GC*
Prolina	1.6	CC*
Glicina	-0.4	GG*
Treonina	-0.7	AC*
Serina	-0.8	UC* AGY
Triptofano	-0.9	UGG
Tirosina	-1.3	UAY
Histidina	-3.2	CAY
Aspartato	-3.5	GAY
Glutamato	-3.5	GAR
Asparagina	-3.5	AAR
Glutamina	-3.5	CAR
Lisina	-3.9	AAAY
Arginina	-4.5	CG* AGR

Es interesante notar que cuando aunque la guanina tiende a codificar aminoácidos de hidrofobicidad intermedia en la segunda posición, parece ser una especie de comodín. Algunos autores consideran que la guanina, en el “código dentro

del código”, es un nucleótido informacionalmente neutro. G en la segunda posición no tiene una identidad fisicoquímica definida, pues tiene aminoácidos representativos de todo el espectro de hidrofilia; además cabe mencionar que en los tres aminoácidos con redundancia de 6 codones, dos son consistentes en la segunda letra (Leucina = CUN/GUR; Arginina = CGN/AGR). En el tercero (Serina = UCN/AGY), la pareja “externa” (AGY) tiene G en la segunda posición.

La primera letra del codón: Biosíntesis

Wong (1975) probablemente fue el primer autor que publicó un comentario acerca de la relación biosintética entre los codones similares, con los análisis de contigüidad que realizó al exponer su teoría de la coevolución. Tradicionalmente se atribuye la primera publicación explícita de tal relación a Taylor y Coates (1989). Dichos autores lo resumen de esta manera: los codones iniciados con A son la familia del aspartato, que incluye a la asparagina, treonina, isoleucina, metionina y lisina; el aspartato no está incluido dentro de esta familia sino que es el precursor de cinco aminoácidos mencionados. Los iniciados con U son casi todos aromáticos, y se agrupan en la familia del shikimato compuesto derivado a su vez del fosfoenolpiruvato (PEP) y la eritrosa 4 fosfato (E4P); estos son triptófano, fenilalanina y tirosina. Otros aminoácidos codificados por los codones que inician en U (cisteína y serina) se relacionan directamente con el triptófano mediante su ruta biosintética. Los codones que comienzan en C son la familia del glutamato, que incluye arginina, glutamina y prolina; de nuevo, el glutamato no está incluido dentro de su familia, sino que es el precursor de los tres anteriores. La G, nuevamente, es un nucleótido más ambiguo que los demás pues no tiene ninguna ruta biosintética preferencial ni algún precursor común. Sin embargo, existe una cierta regularidad: todos sus aminoácidos (glicina, alanina, valina, aspartato y glutamato) son relativamente sencillos, son los aminoácidos que se encuentran en más cantidad en los experimentos de química abiótica, lo cual quiere decir que su síntesis es fácil de realizar. Además, los aminoácidos cuyos codones comienzan en G se sintetizan con una sola reacción química de sus precursores (ya sean éstos parte de la glucólisis o del ciclo de ácidos tricarbóxicos). Esta característica se retomará después pues es fundamental para la propuesta de código primitiva de codones que realiza Copley (Copley 2005)

El número de codones de cada aminoácido. El caso de la arginina.

Massimo di Giulio ha dedicado buena parte de su investigación reciente a encontrar relaciones significativas entre el número de codones asignados a cada aminoácido y distintas propiedades de éstos. La relación más evidente es la que se aprecia al comparar el número de codones con la complejidad química de cada aminoácido (expresada en peso molecular). Esta relación había sido mencionada en Taylor y Coates (1985); entre mayor sea el peso molecular de un aminoácido, menor el número de codones asignados.

A simple vista se puede apreciar la misma excepción de los casos anteriores: la arginina, con un peso molecular de 174, tiene asignados seis codones. Es pertinente analizar en este momento las posibles razones (o, al menos, las razones reportadas en literatura) para explicar la consistencia de este aminoácido en presentarse como una excepción a la luz de las regularidades generales del código genético.

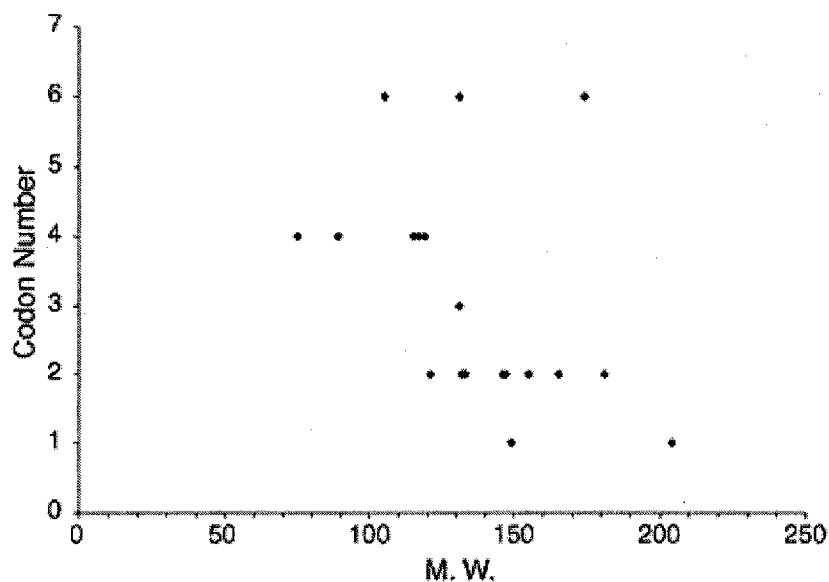


Fig. 15 - Correlación entre peso molecular y número de codones. El punto superior izquierdo corresponde a la Arginina. (Di Giulio, 2005)

Los trabajos más importantes son los llevados a cabo por el grupo que apoya la teoría estereoquímica (Knight 2000, Yarus 2000, 2005). En sus artículos reportan constantemente la selección de ribozimas que se unen a residuos de arginina, *cuyos sitios activos están constituidos precisamente por secuencias de los codones asignados*

actualmente a la arginina. Esto apunta a que, al menos en este caso, es posible que exista un factor estereoquímico en la asignación de codones. (Como se verá más adelante, no existen evidencias similares que relacionen de esta manera a otros codones con sus aminoácidos).

Por otro lado, Di Giulio ha realizado un estudio similar a la comparación de peso molecular con el número de codones, tomando otras propiedades (como el índice de termofilia y de barofilia (Di Giulio 2005) de los aminoácidos). De esta manera, ha encontrado una correlación directamente proporcional de las propiedades extremófilas de cada aminoácido (comparando los aminoácidos que utilizan preferentemente los organismos extremófilos y aquellos que usan los no extremófilos) y el número de codones que presenta cada uno. La arginina tiene un alto índice de termofilia y barofilia, aunque en los organismos mesófilos dicho aminoácido esté sujeto a fuertes restricciones selectivas. Estos resultados llevan a Di Giulio a proponer que el origen del código (y extrapolándolo con la propuesta del presente trabajo, el origen de la vida) se llevó a cabo en ambientes de presión y temperatura elevadas.

Reducción de errores

Las propiedades anteriormente mencionadas tienen un efecto combinado: causan que el código genético tenga propiedades de reducción de errores. Durante los últimos cuarenta años, ha existido una discusión centrada en la cadena de causalidad de estos eventos. ¿Es el código genético resistente a errores por un hecho contingente, o esta característica ha surgido precisamente gracias a un proceso de selección natural?

En un artículo clásico, Freeland (Freeland 1998) investiga la posibilidad de que azarosamente se haya logrado una configuración con tal cantidad de adaptaciones aparentes, y llega a la conclusión de que ésta es mínima. Desde un punto de vista ingenieril, pueden fabricarse códigos genéticos que sean mucho más eficientes que el utilizado por los seres vivos terrestres actuales en sus propiedades de disminución de errores. (Knight 1999, Di Giulio 1998). Es importante tener en cuenta que no se pretende que la configuración actual es la óptima. Probablemente después de un periodo de establecimiento, las relaciones codón-aminoácido se fijaron, mediante selección, en un pico cercano del paisaje adaptativo. Los resultados de Freeland indican que de un millón de códigos generados aleatoriamente, sólo uno resulta “mejor” que el código natural. Por supuesto, como él mismo lo discute, tales resultados son informativos pero

de ninguna manera definitivos: debido a que su técnica asigna pesos a los distintos tipos de mutación, sus parámetros resultan algo aleatorios. Existen muchas maneras de asignar esos pesos; Freeland, por ejemplo, decide ignorar la regularidad fisicoquímica de la segunda base, por lo cual en sus cálculos la eficiencia de cualquier código aleatorio es igual que el código natural, si se considera únicamente la base intermedia. Al considerar la estructura y función de cualquier proteína hipotética, vemos que esto no es así: un cambio de un aminoácido polar por otro polar tendría un efecto menos grave en el plegado proteico o en las regiones transmembranales que cambiar un aminoácido polar por uno alifático no polar. En cualquier caso, el código genético natural resulta significativamente resistente a errores.

Sin embargo, la resistencia a errores no es un argumento definitivo a favor de la construcción del código genético mediante la selección natural. Éste no se consolidó de manera aleatoria, pero las regularidades pueden ser explicadas mediante una serie de restricciones físicas, como se propone más adelante. En efecto, es posible que la cualidad reductora de errores del código sea simplemente un subproducto de tales restricciones. Tan sólo la redundancia de la última letra del codón da cuenta de gran parte de su eficiencia; por ejemplo, en los casos en que la familia de un aminoácido es de cuatro miembros tres mutaciones de nueve posibles resultan silenciosas. En el caso del codón CUA (Lys), cuatro mutaciones de nueve son silenciosas. Si tomamos en cuenta las otras regularidades ya discutidas *y además asumimos que éstas fueron determinantes para la asignación de codones y no al revés*, la corrección y reducción de errores se puede considerar como un producto afortunado pero completamente probable del surgimiento del código por otros mecanismos.

Elementos del código genético

Antes de comenzar a comparar teorías acerca del surgimiento del código genético, es necesario mencionar los elementos que lo constituyen. En la primera parte de este trabajo se ha mencionado que el código es una relación simbólica y por lo tanto abstracta; sin embargo, necesariamente tiene que haber ciertos elementos físicos que permitan su realización material. Tales elementos componen un sistema que es, posiblemente, el más complejo dentro de los elementos celulares; en él intervienen al menos tres clases distintas de moléculas de RNA, además de enzimas proteicas y

aminoácidos monoméricos. Es, además, uno de los procesos que están extendidos a lo largo de todo el espectro de la vida conocida, por lo que se puede asumir que es por lo menos tan antiguo como el último ancestro común (LUCA, por sus siglas en inglés). En resumen, se tienen cuatro elementos principales. El mensaje, codificado en el RNA mensajero (**mRNA**); el **ribosoma**, que cataliza la lectura del mensaje y la polimerización del péptido; el adaptador, una molécula de RNA unida covalentemente a un aminoácido específico, llamada RNA de transferencia (**tRNA**); y por último el péptido resultante, producto análogo al mRNA. A grandes rasgos, el mRNA, transcrito directamente del DNA de la célula, se asocia a un ribosoma (un aparato riboprotéico compuesto por dos grandes subunidades). Éste permite la asociación de tRNAs específicos a cada triplete, por medio de una asociación Watson-Crick entre el mRNA y el tRNA; después, cataliza la polimerización de la cadena de aminoácidos, uno por cada triplete del mRNA, hasta llegar a una señal de alto (llamado codón de STOP), momento en el cual las dos subunidades se disocian y la proteína queda libre.

Los detalles de este proceso varían, sobre todo entre eucariontes y procariontes. No sólo por el detalle más evidente, la existencia de la compartimentalización nuclear en los eucariontes; también existen diferencias entre las estructuras de los ribosomas de ambos grupos, los factores de liberación que se unen a los codones de STOP, el grado de fidelidad con que se unen el tRNA y el mRNA, y las modificaciones post-transcripcionales que se llevan a cabo en el mRNA, por mencionar sólo unos cuantos ejemplos. En este trabajo no se abundará sobre las distintas versiones de la maquinaria de traducción (mas que en los casos en que sea necesario para la argumentación); la atención estará enfocada a los rasgos comunes, pues se asume que las modificaciones, probablemente ocurridas a lo largo de la evolución de los eucariontes, son poco importantes para comprender el origen del sistema y por tanto quedan fuera de los límites de la investigación.

Principales teorías del origen del código genético

El accidente congelado / Coevolución

Una de las primeras teorías acerca del origen del código genético fue propuesta por Francis Crick (Crick 1968). Comúnmente se piensa que tal teoría, llamada “el accidente congelado”, propone que el ordenamiento de los codones no es más que una

configuración que se estableció aleatoriamente en el origen de la vida. Es decir, que la asignación de codones no es racionalizable y por tanto no admite ningún tipo de explicación causal más que el azar. Tal visión incompleta de la teoría lleva a un rechazo prematuro y fácil, pues, como ya se ha visto, en el código genético hay una serie de propiedades cuyo surgimiento aleatorio sería improbable (en particular la llamada “resistencia a los errores” del código, que tiene sus raíces en las propiedades fisicoquímicas de los aminoácidos codificados). Esta manera de ver la hipótesis de Crick cumple tal vez con el criterio del rasero de Occam (“la explicación más sencilla es probablemente la correcta”), pero tiene muchos puntos en contra. Al contrario de lo que se puede suponer al echar un primer vistazo al código genético, éste es todo menos arbitrario. Las propiedades que ya se han mencionado en un apartado anterior (las redundancias del código, las características fisicoquímicas uniformes que se asocian con la primera y segunda letra del codón, etc) difícilmente se podrían haber configurado de tal manera gracias a un acomodo completamente aleatorio. Todas estas propiedades han hecho que se considere al código genético como un *código minimizador de errores*.

En realidad, la hipótesis de Francis Crick es frecuentemente malentendida, probablemente por la manera en que éste decidió nombrarla. El congelamiento del código genético es la última etapa de una secuencia de pasos. El primer paso es el establecimiento de un código primitivo en un solo individuo (o en una población endogámica de individuos). Cuando tal población ganara complejidad, tanto metabólica como informacional, se agregarían más aminoácidos al código hasta llegar a un umbral en el que más modificaciones ocasionarían una catástrofe. Dada la cantidad de aminoácidos y la complejidad de las proteínas codificadas en este punto (una complejidad seguramente menor a la de las proteínas actuales), cualquier cambio haría que las enzimas primitivas se volvieran inservibles. En este punto es en el que finalmente ocurre el congelamiento del código. El término “accidente” se refiere a que, en algún momento o en otro, algunos de los aminoácidos fueron asignados por eventos azarosos. “En su versión extrema”, dice Crick, “la teoría implica que la asignación de aminoácidos a los codones fue dada, en este punto, solamente por el azar”.

Como se ve, si se habla de tal hipótesis sin conocer más que el nombre, es fácilmente refutable. El código no está “congelado”, pues ahora se conocen numerosos códigos alternativos, principalmente en organelos y en bacterias. Tampoco resulta ser un “accidente”, pues si lo fuera resultaría difícil explicar estadísticamente las numerosas

propiedades de resistencia a errores que tiene el código. Sin embargo, si se considera el arco completo de la teoría, se verá que el congelamiento simplemente es una manera de llamar a la clara restricción que existe en el código actual, aunque existen algunos mecanismos para reasignación de codones²⁵. En realidad, vista holísticamente, la teoría del accidente congelado está estrechamente relacionada con una de las teorías más aceptadas del surgimiento del código, la teoría de la co-evolución.

Wong (Wong 1975) propuso por vez primera la teoría de la co-evolución, basándose principalmente en las regularidades biosintéticas que presenta el código. En su primera elaboración, presenta parejas de aminoácidos separados por un solo paso enzimático, y los compara con la tabla del código genético. En efecto, la mayoría de las parejas establecidas por Wong tienen dominios de codones contiguos, con la excepción de las parejas Glu-Pro, Glu-Arg, Asp-Thr y Asp-Lys.. Tal relación es significativa pero no definitiva; tan sólo hay que mencionar que los dominios de codones tienen 6 dominios contiguos de 15 posibles, y que las parejas de aminoácidos no son exhaustivas. Además, ciertos aminoácidos tienen rutas múltiples de biosíntesis.

Estereoquímica

Otra de las hipótesis pioneras, ésta propuesta por Woese (Woese 1967), es un posible mecanismo puramente físico de asignación de codones. En ella, se considera que los codones y los aminoácidos se asociaban físicamente en el origen del código genético, probablemente por una compatibilidad estereoquímica. Aunque es una de las teorías menos investigadas por falta de datos experimentales sólidos que la apoyen, con el advenimiento de la tecnología de diseño de aptámeros se descubrió una relación inusual: existe una afinidad entre la arginina y sus codones presentes en la secuencia de los aptámeros realizados *in vitro*. Se han encontrado otras relaciones, menos específicas, entre otros aminoácidos y sus codones (isoleucina y tirosina, sobre todo) Yarus 2000. El poder explicativo de esta teoría es limitado: no provee información acerca de las posibles interacciones de otros aminoácidos, sino sólo se enfoca en encontrar las interacciones de los 20 esenciales; tampoco puede proponer una razón válida de por qué

²⁵ Es conveniente considerar un argumento no inusual en este debate: “Si el código genético está sujeto a cambios, debe de estar sujeto, pues, a selección natural; después de casi cuatro mil millones de años, debe de haber mejorado e incluso tal vez esté aún en proceso de refinamiento”. La existencia de códigos alternos es un evento que necesita de condiciones muy específicas, como lo ha propuesto Osawa (Evolution of the genetic code, 1995. Además, el código estándar se encuentra en la raíz del árbol de la vida: es compartido por la mayoría de los seres vivos de los tres dominios de Woese, además de los virus

se asignaron tales aminoácidos y no otros, fisicoquímica o estereoquímicamente similares. Además, aún en caso de que se encontraran tales relaciones en todos y cada uno de los aminoácidos con sus respectivos codones (y ese no es el caso), no explica las regularidades del código ya mencionadas.

3. Una teoría nueva

Preliminares

Existen al menos dos observaciones que es conveniente hacer antes de comenzar a hablar de una propuesta para el surgimiento del código genético. La primera es analizar el papel del RNA a la luz de los problemas explicados en la sección TAL. Es innegable que el RNA necesita estar presente en las primeras etapas de la vida; así lo podemos ver con su participación en los sistemas obligatorios de la célula en sus diferentes funciones (principalmente mRNA, tRNA y rRNA). También es innegable su valor catalítico; incluso el ribosoma, un aparato ribonucleoprotéico, en realidad funciona como una ribozima.

También se tiene que hacer notar la universalidad del código genético. Es cierto, como argumentan los detractores de la teoría del accidente congelado (o más bien: los detractores del nombre de la teoría), que no existe un solo código genético. Sin embargo esto no quiere decir que existan códigos alternativos, sino que las variantes *son claramente alteraciones excepcionales sobre el código estándar*. Como ya se ha mencionado, tales cambios se dan en condiciones poco probables en las que la mutación de una característica tan fundamental como la relación de un codón con un aminoácido no resulta fatal. Para explicar que exista un solo código genético (el llamado estándar), creo necesario que ocurra una o más de las siguientes eventualidades: 1) en el surgimiento del código genético ya existía una compartimentalización, por lo cual las interacciones codón-aminoácido se transmitieron desde un ancestro común delimitado; 2) el surgimiento del código genético se llevó a cabo en un sistema sin compartimentalización pero delimitado, ya sea un cuerpo de agua pequeño y concentrado (en el escenario de la sopa prebiótica) o en una superficie limitada; 3) el código genético surgió como muchas variaciones sincrónicas y sintópicas, y el código actual fue seleccionado por alguna razón, ya sea como un código “puro” entre otros con

menos “adecuación” (sea lo que sea que esto implique), o como una fusión de características de más de un código; 4) el código genético es *una necesidad*, en el sentido de Monod (Monod 1970); es decir que las condiciones en las que surgió hacen probable la configuración que adoptó.

Mecanismos de reacción de Copley

Shelley Copley (2005), frente a la imposibilidad de explicar mecanísticamente las regularidades ya analizadas del código genético, propone un mecanismo de asociación puramente químico entre el aminoácido y su codón. Existen ciertas suposiciones iniciales:

1. En todo nucleótido de RNA, existen varias posibilidades químicas. El oxígeno del C2, por ejemplo, es altamente reactivo. También es conocido el uso de grupos fosfato, iones (en particular Mg $2+$) y cofactores para llevar a cabo distintas reacciones. Esto es prácticamente un hecho comprobado: en la mayoría de las ribozimas, por ejemplo, tal reactividad está detrás de la actividad catalítica del RNA en cuestión.
2. Existe una fuente constante de elementos de las rutas bioquímicas centrales, en particular de los componentes del ciclo de los ácidos tricarboxílicos. Se asume que tal serie de reacciones no es simplemente un producto arbitrario llegado hasta nuestros días por el linaje de la vida terrestre, sino que es uno de los atractores fisicoquímicos que han sido impuestos a la vida: es una de las soluciones energéticas que restringen a la vida. Tal suposición representa una línea de investigación en sí misma; ya se ha reseñado en la sección dedicada a la teoría de Wächtershäuser. El argumento central es que el metabolismo del ciclo de Krebs es una pieza necesaria para entender el origen de la vida que conocemos.

Siendo las regularidades biosintéticas un factor central en la teoría de Copley, ella propone una unión covalente de un precursor de aminoácido (derivado del ciclo de ácidos tricarboxílicos) a un dinucleótido. Tal unión resulta ventajosa para explicar las reacciones que transformarán al precursor en el aminoácido correspondiente,

proveyendo, en particular, cierto ambiente químico, características de orientación y facilitando energéticamente el cambio.

Se proponen tres tipos de reacciones: 1) una aminación reductiva del precursor; 2) fosforilación del precursor; y 3) decarboxilación del piruvato.

La reacción 1) es simplemente una aminación reductiva, reacción metabólica que genera los aminoácidos cuyos codones empiezan con G. Son los menos complejos, y en el caso de la glicina y la alanina, no se llevan a cabo más reacciones pues tales aminoácidos no poseen posiciones reactivas. El aspartato y el glutamato pueden sufrir una fosforilación, por lo que necesitan cierta protección ante su reactividad. Se propone que la formación de puentes de hidrógeno entre la amina exocíclica de G en primera posición y A en segunda pueda proteger el grupo carboxilo del aminoácido, haciendo que subsecuentes modificaciones sean improbables. Es importante mencionar que la Guanina en primera posición es el único nucleótido que tiene tal amina exocíclica que alcance al precursor unido en 2' de la primera base.

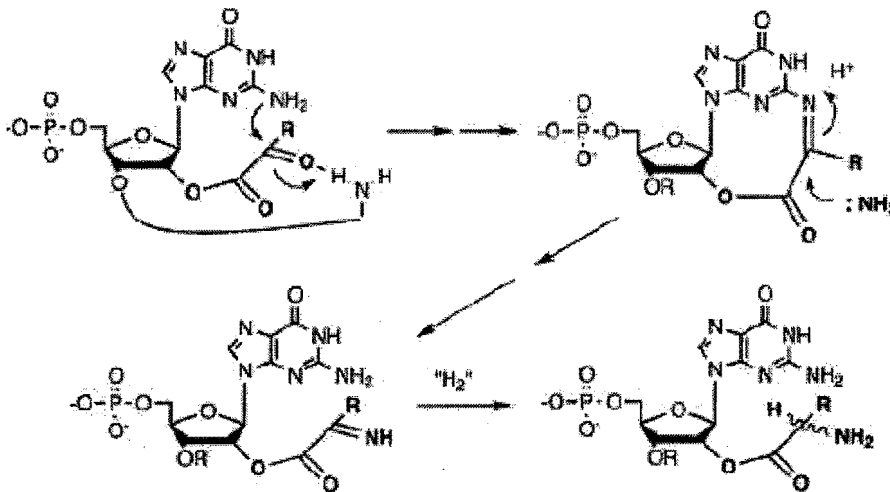


Fig. 16 - Aminación reductiva catalizada por Guanina en primera posición (tomada de Copley 2005). El grupo amino de la primera imagen pertenece a la segunda base del dinucleótido, simplificado después como R en 3'.

La reacción 2), la fosforilación del precursor, puede ser realizada por distintas bases aunque cada una tiene posibilidades reactivas distintas, dependiendo del α -ceto ácido involucrado. Se propone que el ataque sobre los diversos precursores es apoyado por los grupos exocíclicos de la primera base del dinucleótido, lo cual plantea una explicación mecánica de la relación directa que existe entre precursores específicos y las bases A, U y C en primera posición. (La base G en primera posición, como se detalla en la reacción

uno, determina el tipo de reacción y no el precursor específico). La fosforilación en el caso del α -cetoglutarato, por ejemplo, se ve facilitada por el puente de hidrógeno que puede formar la amina exocíclica de la citosina con el oxígeno del fosfato necesario; al no tener amina exocíclica, el uracilo tendría una reactividad reducida en este caso. Los análisis realizados por los autores indican que en el caso de este precursor, la reactividad esperada es $C > A > U$.

Algo similar ocurre en el caso del oxaloacetato cuya fosforilación también es favorecida por una amina exocíclica. Sin embargo, siendo de cadena más corta, los autores achacan razones estéricas a que el orden de reactividad se distinto al del α -cetoglutarato: en este caso, es $A > C > U$. Por último, la fosforilación del piruvato para formar fosfoenolpiruvato puede ser catalizada con ayuda de un ion metálico como Mg^{2+} ; éste es un cofactor utilizado actualmente por las enolasas para realizar este tipo de reacción. Es factible que tal ion pueda ser sustituido por otro ion metálico, (como Fe^{2+} , lo cual estaría en conformidad con la tesis de Wäctershäuser), para ser sustituido en enzimas protéicas por uno de valencia similar en etapas posteriores: sin embargo, el artículo de Copley no hace mención de esta posibilidad. El artículo reporta un impedimento de cargas tanto en los nucleótidos UA y UU, lo cual explicaría que éstos estén asociados a codones de STOP y a otros mecanismos biosintéticos, respectivamente.

La reacción 3) es, según los autores, la que comprende el mecanismo más complejo, la decarboxilación de una molécula de piruvato seguida de un ataque nucleofílico en una segunda molécula (de piruvato en caso de Val y Leu; de α -cetobutirato en caso de Ile). Actualmente, tal reacción se lleva a cabo con la asistencia de un cofactor, pirofosfato de tiamina. Copley y sus colaboradores sugieren que el uracilo en segunda posición puede funcionar como un cofactor primitivo, explicando la agrupación de estos aminoácidos (hidrofóbicos y ramificados) en los codones cuya segunda letra es U. Resulta de particular interés que en la valina, codificada por GUN, se reúnen las reacciones 1 y 3 (primero una decarboxilación de piruvato seguida de una aminación reductiva).

Como se puede observar en la figura, tal repertorio de reacciones limita el número de aminoácidos. Por lo tanto, y como se ha propuesto desde los primeros estudios del origen del código, se debe de considerar la posibilidad de un código primitivo. La teoría del bamboleo y el escaso valor informativo de la tercera letra del codón en muchos de

los casos (al menos en lo que se refiere a identidad química del aminoácido) sugieren que es probable que el código haya surgido en un sistema en el que se leían²⁶ sólo las dos primeras bases de cada triplete, ignorando la tercera.

Un punto que es pertinente notar es que los dinucleótidos contienen en sí mismos cierta información acerca del código genético. De hecho, la identidad química de tales dinucleótidos *es lo único que es necesario para materializar la relación semiótica del código* en esas primeras etapas. Ésta se presenta como una asociación puramente física, y no representa una relación simbólica como tal, bajo los parámetros establecidos en la Parte I de este trabajo; sin embargo, la relación simbólica que existe en los seres vivos actuales es compleja, y necesita de una serie de elementos cuya función parece ser exclusivamente realizar tal relación. Como en muchos problemas en evolución, resulta improbable que tales elementos surgieran de manera independiente, y espontáneamente encajaran en el sistema complejo del código. En las primeras etapas, pues, es probable que las convenciones naturales surjan de relaciones dictadas enteramente por leyes físicas. Este punto de vista no es reduccionista: tales leyes son insuficientes (al menos para el entendimiento humano) de explicar la complejidad de las propiedades emergentes que surgen al cruzar cierto umbral.

Otro punto de importancia capital, en un sentido más mecanístico, es el siguiente: Copley se muestra intrigada ante su propio modelo químico pues los elementos relacionados son el aminoácido y el codón. Desde cierto punto de vista, esto tiene menos sentido que encontrar un mecanismo que relacione al aminoácido con el anticodón, ya que ambos se encuentran necesariamente relacionados en los tRNA actuales. Los autores opinan que se tiene que postular un mecanismo para una transesterificación del aminoácido hacia una molécula de RNA que se asocie al codón, obteniendo así una especie de tRNA primitivo.

²⁶ El término "Lectura" está aplicado con mucha libertad en este punto de la discusión. La "lectura" en el escenario propuesto por Copley y su grupo es simplemente afinidad y ambiente químico causado por el dinucleótido.)

Propuesta

*We accept explanations ultimately
because they are aesthetically pleasing*

H. H. Pattee

1. Principios de la información biológica

El mecanismo de reacción propuesto por Copley trasciende el sistema del mundo del RNA, aún cuando fue concebido en este marco teórico. Ya se ha hablado de lo sorprendente que le resulta a la autora que los aminoácidos se asocien a sus codones y no a sus anticodones, lo cual produciría un tRNA primitivo para traducir un mensaje prefabricado (que, a su vez, proveería un mecanismo para pasar del mundo del RNA a los sistemas riboprotéicos).

Este artículo, sin embargo, es medular para este trabajo porque provee un mecanismo puramente químico de asociación. Además este mecanismo es completamente novedoso; nunca antes se había vislumbrado la capacidad de los nucleótidos para proporcionar un ambiente químico apropiado para la transformación de los precursores (α -ceto-ácidos del ciclo de Krebs) en aminoácidos. Tal mecanismo cae fuera de las teorías usuales del surgimiento del código genético; es, tal vez, un punto medio entre la teoría de coevolución y la teoría estereoquímica.

El presente trabajo propone un mecanismo teórico del surgimiento del código genético, basado en la propuesta de Copley; se analizarán las consecuencias de tal modelo y algunos datos que concuerdan con él. Es difícil presentar datos y afirmar inequívocamente que demuestran o prueban la teoría: simplemente parecen no contradecirla. También se considerarán algunos problemas e interrogantes de dicha teoría, y finalmente se propondrán algunos experimentos posibles a realizar, tanto informáticamente como en laboratorio. Tales experimentos serán un punto de partida para continuar esta investigación a largo plazo.

Se ha mencionado el problema de continuidad, ya analizado por Crick. Debido a éste, creo necesario tomar un número de supuestos: que en el surgimiento del código, la información del mensaje surge en una cadena de RNA o una análoga (pues el surgimiento en una cadena de RNA evita la necesidad de proponer un mecanismo de traducción de la molécula análoga al RNA); que el lenguaje surge como tripletes y no aumentando gradualmente; y finalmente, que toda la información no involucrada

directamente en el surgimiento del código, aún cuando haya estado en forma de RNA, se pierde por completo, y por tanto resulta irrelevante para el modelo del surgimiento. También creo necesario hacer explícito que se utiliza el llamado “sueño del biólogo molecular”, como lo llama Leslie Orgel: se asume que existe una fuente de moléculas con la quiralidad que los seres vivos utilizan (D-pentosas y L-Aminoácidos). Aunque el modelo de Copley presenta una solución para la quiralidad de los aminoácidos, y se han descrito ciertas condiciones en las que se producen preferentemente D-carbohidratos (Bielsky 2007), la quiralidad sigue siendo un problema importante para los químicos prebióticos, y queda fuera de los objetivos de este trabajo.

Comenzaré por decir que los tripletes de RNA, como complejos moleculares, son construcciones relativamente estables; en la polimerización dependiente de un templado o molde en realidad se oligomerizan dímeros o trímeros, en un paso sin catalizador. Después, se unen por medio de enlaces de hidrógeno al molde y se lleva a cabo una ligación (Schwartz 1998). Además, se ha publicado, en los experimentos de autorreplicación, que la ligación de oligómeros (en particular trímeros) es más eficiente que cuando se utilizan monómeros (Burmeister 1998). La razones por las que la ligación química de trímeros (von Kiedrowski 1986) es más factible que la policondensación de monómeros (Orgel 1995) parecen ser la variedad de catalizadores que permiten tal ligación, la probabilidad menor de formar enlaces no característicos del RNA y las restricciones reducidas de utilizar purinas o pirimidinas como molde. No resulta improbable, pues, que los tripletes con los que funciona el código genético estén relacionados con estas particularidades del RNA.

Consideremos un complejo de un triplete de RNA unido covalentemente a un aminoácido. El modelo de Copley presenta una unión de aminoácidos específicos con codones específicos, en el carbono 2' de la primera base²⁷. Las características informativas del RNA surgen, en la presente propuesta, no como una molécula a la que es posible leer, sino como una molécula que provee de ciertas condiciones fisicoquímicas. Es decir: la información biológica y su codificación surge independientemente de la complementariedad de las bases nitrogenadas, aunque tal característica, presente siempre en potencia, es determinante para la propagación del sistema. El proceso de lectura es un proceso complejo: es justamente el punto en el que radica la relación simbólica más básica

²⁷ Aunque Bosenberger afirma que es más probable que ocurran espontáneamente tripletes con grupos funcionales extraños en el carbono 2' de la tercera base

del lenguaje celular. No es necesaria una simple complementaridad “ciega” (un reconocimiento no simbólico de un patrón por otro patrón), sino un reconocimiento por un adaptador, prefabricado por enzimas específicas para una función determinada, en un proceso ayudado por una maquinaria que además realiza reacciones de polimerización peptídica.

Teniendo estos complejos (tripletes de RNA aminoacilados específicamente), es posible afirmar que se tienen los fundamentos de la traducción; también se pueden proponer varios escenarios en los que, paso a paso, se establecen sistemas que almacenan información, valiéndose (sólo hasta este punto) de las propiedades de complementariedad de los ácidos nucleicos. A continuación presento el que me parece más sencillo.

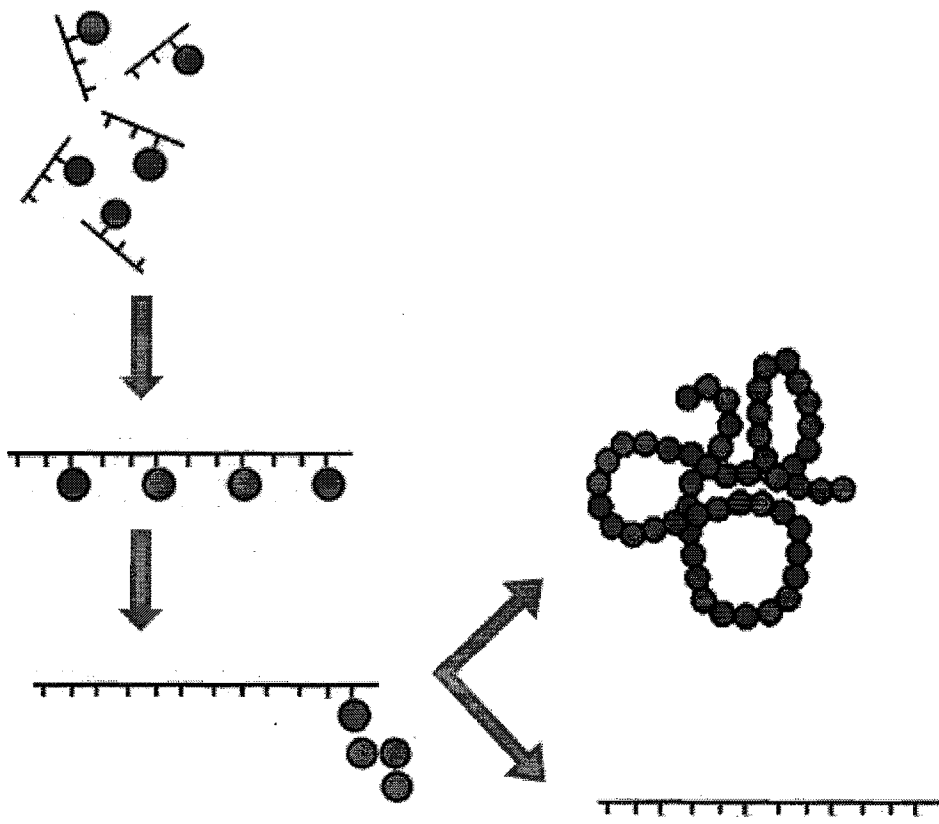


Fig. 17 - Mecanismo propuesto del origen de la codificación protéica

Los complejos aminoacil-RNA se pueden polimerizar en cadenas mayores, lo cual produciría cadenas de RNA con un aminoácido unido covalentemente cada tercera base. Hago notar que esta distancia es similar (si no idéntica) a la distancia existente en la traducción actual: los tRNA se unen a cada triplete en un espacio estéricamente limitado. Algunos autores opinan que los adaptadores están hechos de ácidos nucleicos

en vez de proteínas porque, además de representar un menor costo, los aminoácidos necesitarían de una estructura mayor para lograr una fidelidad aceptable, lo cual no sería permitido por las restricciones espaciales entre codón y codón.

Es posible, entonces, imaginar un sistema en el que utilizando esta molécula, se lleve a cabo una reacción relativamente sencilla de polimerización. Esta reacción probablemente fue llevada a cabo inicialmente por algún catalizador y no de manera espontánea, como propone Bosenberger: se ha demostrado que el rRNA lleva a cabo esa unión mediante posicionamiento de los componentes y no como una reacción de catálisis ácido-base. Esto permite ampliar las perspectivas de un escenario de surgimiento, pues tal proceso parece ser independiente de pH (Bieling 2006), a diferencia de lo que antes se había propuesto: un mecanismo de ataques nucleofílicos dependientes del carácter de zwitterion de cada aminoácido (Rodnina 2003).

Como señala Copley, la ventaja de la formación de aminoácidos mediante el ambiente químico proporcionado por los dinucleótidos provee de un catalizador reusable: el aminoácido puede ser separado por hidrólisis y el dinucleótido queda libre para ser unido nuevamente a un precursor simple, derivador del ciclo de los ácidos tricarbónicos. Sin embargo, no es descabellado imaginar el papel de la molécula resultante de RNA: es el primer depósito de información surgido, mediante el cual será posible reproducir la proteína formada. Complejos de trinucleótidos aminoacilados podrán unirse, por las propiedades complementarias de los ácidos nucleicos, a templados no catalíticos, los cuales tendrán un tiempo de vida más largo si no toman parte en ninguna reacción.

Esta propuesta plantea el origen del RNA mensajero de manera simultánea al origen de la codificación primitiva. Las rutas que se pueden seguir desde el origen de ambos elementos son variadas. Las enzimas codificadas pueden tener muchas funciones catalíticas que aumenten la eficiencia del sistema; los procesos más importantes, sin embargo, son los relacionados con la polimerización de péptidos y ácidos nucleicos. Tal es la única función catalítica del RNA que tendrá una continuidad histórica, y no como replicadores o autorreplicadores, como se verá más adelante. El presente modelo utiliza tripletes de RNA con el aminoácido unido covalentemente en la primera base; sin embargo, es pertinente, para futuros análisis, considerar que la química prebiótica no era específica y el aminoácido se puede unir en cualquiera de las bases, lo que ocasionaría que este modelo se volviera más complejo. Este factor ya ha sido analizado, desde otro

punto de vista, por Wu (2005), que propone la unión de dos códigos de dímeros (el de las dos primeras letras de cada codón y el de las dos últimas) para producir el código actual de tripletes, basándose sobre todo en la distribución de las redundancias de los aminoácidos.

El uso de complejos aminoacil-RNA ya ha sido analizado por el grupo de Bosenberger (Bosenberger 2004), los cuales se adhieren a la teoría estereoquímica del surgimiento del código genético. Como ya se ha mencionado antes, tal teoría tiene poca evidencia experimental que la respalde; sin embargo, modelos como el de Bosenberger y el presentado en este trabajo son, en sí mismos y en mi opinión, una alternativa a las teorías tradicionales.

2. *Un código genético primitivo*

Es probable que el código genético haya sufrido modificaciones desde sus orígenes, sean estos los propuestos aquí o no. Ya se ha mencionado que no está exento de ser modificado, lo cual es posible comprobar con sólo ver las variaciones de los códigos genéticos de procariontes y organelos. No es descabellado considerar la realización un código genético primitivo, como ya lo han hecho muchos autores para facilitar el estudio de su posible origen y de sus características generales. Ya se ha mostrado el propuesto por Copley (2005); a continuación se encuentra el propuesto por Bosenberger (2004), excluyendo los codones codificados por un solo aminoácido y los que se codifican con los codones UNN y NAN, pues según los autores son en los grupos en los que se concentran los aminoácidos más complejos. El código resultante sólo cuenta con 9 familias funcionales de codones, y 8 ó 9 aminoácidos. Las casillas vacías representan codones *nonsense*.

		SECOND BASE					
		G	U	C			
FIRST BASE	G	Gly Gly 2 Gly Gly	Val Val 1 Val Val	Ala Ala 2 Ala Ala	THIRD BASE	U	C A G
	A	Ser 2 Or Arg 1	Ile Ile 1 Ile Ile	Thr Thr 2 Thr Thr		U	
	C	Arg Arg 1 Arg Arg	Leu Leu 1 Leu Leu	Pro Pro 2 Pro Pro		C	
						A	

Fig. 18 - Tabla primitiva de codones según Bosenberger (2004)

Para proponer un código primitivo se pueden considerar ciertas modificaciones posibles:

- En un principio, los codones de STOP no tenían factores complejos para la traducción; lo más probable es que eran codones *nonsense* sin sentido protéico que interrumpían la traducción. Es posible que desde el inicio sólo estuvieran presentes los que existen actualmente, pero la presencia de codones de STOP es flexible. Es más probable, sin embargo, que el número de codones *nonsense* haya disminuido que aumentado. Tal fenómeno se puede observar en las instancias en las que un codón de STOP tiene un aminoácido asignado en códigos alternativos.
- Se pueden excluir del código genético primitivo los aminoácidos que sólo son codificados por un codón (Metionina y Triptofano), que además son los dos aminoácidos más complejos bioquímicamente. Según el modelo de Osawa (Osawa 1992), el uso restringido de un codón, sobre todo por sesgos en la cantidad de bases AU, puede causar una modificación en el código sin un costo adaptativo alto.

- Asimismo, pueden excluirse con más reservas, los aminoácidos codificados por los codones de la familia de STOP: Tirosina y Cisteína. Se basa sobre todo en el supuesto que los codones de STOP (que comienzan con los diptetes UA y UG) comprendían inicialmente estos ocho codones. Sin embargo, la Cisteína se sintetiza con cierta facilidad de manera prebiótica si se incluyen compuestos con azufre, en particular H₂S, el cual es uno de los principales reactivos del modelo de Wächterhäuser.
- Además de los aminoácidos ya mencionados, la Lisina y la Histidina requieren rutas de biosíntesis complejas, por lo que se propone (Bosenberger 2004) que fueron incluidos en etapas posteriores.
- Se pueden considerar codones ambiguos. Es decir: una familia de codones codifica para un grupo de aminoácidos de propiedades similares pero sin establecer la relación unívoca del código actual. Por ejemplo, Leucina e Isoleucina, isómeros estructurales con propiedades similares de hidrofobicidad pueden ser codificados ambiguamente por los codones cuya segunda letra es U.
- El experimento de Miller y sus resultados, independientemente de la posición que se tome con respecto a la existencia de la sopa prebiótica, son una fuente de datos extremadamente valiosa. Los aminoácidos reportados son moléculas con una alta probabilidad de síntesis espontánea, por lo cual es posible que hubieran estado presentes en las primeras etapas de la formación del código. Por supuesto, al modificar distintos parámetros del medio (concentración de reactivos, temperatura, etc) se modifica la cantidad de los aminoácidos producidos, pero los resultados de Miller pueden ser tomados como un punto de partida: Gly, Ala, Val, Leu, Ile, Pro, Asp, Glu, Ser, Thr (Miller 1987)

La figura resume estas características. Tomándolas en conjunto, es posible proponer un catálogo de aminoácidos reducido que constituyan las primeras etapas del código genético. También se puede considerar la manera de codificación inicial de cada uno. (**Ala, Gly, Pro, Ser, Thr** como más importantes por facilidad biosintética y por tener 4 codones; **Asn, Asp, Glu, Gln / Leu, Ile, Val, Phe** como grupos ambiguos por polaridad; y **Arg** como estereoquímica). Dados estos datos, se puede construir un código primitivo que presente tanto influencias biosintéticas como fisicoquímicas.

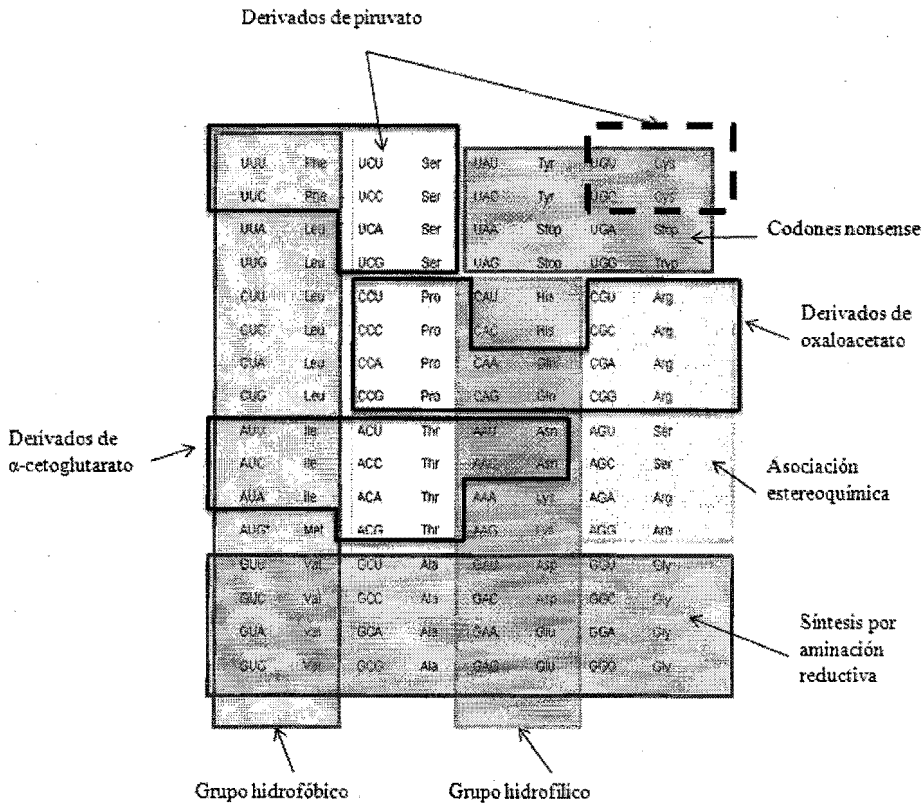


Fig. 19 – Regularidades del código genético utilizadas para proponer una tabla primitiva de codones.

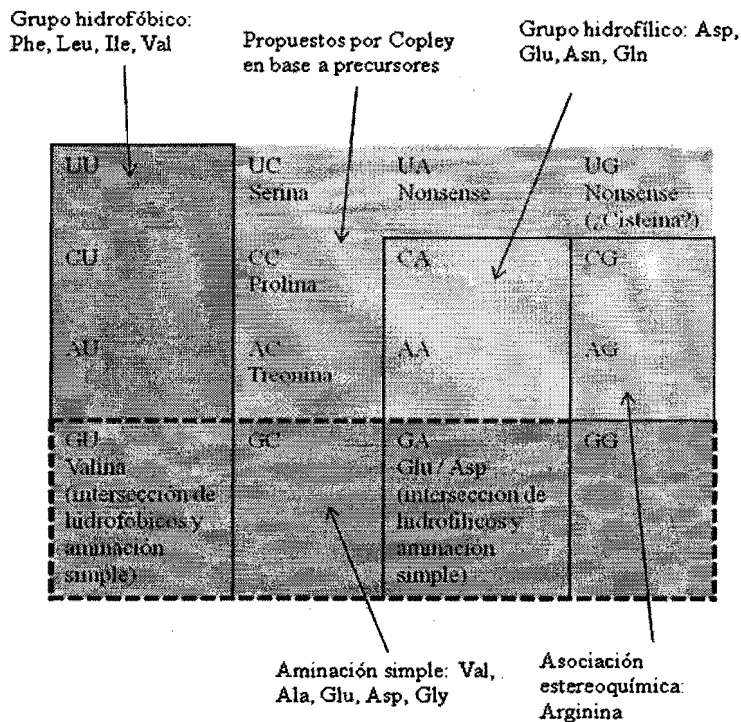


Fig. 20 - Código primitivo propuesto. Los grupos hidrofílico e hidrofóbico son de codificación ambigua.

3. *RNA catalítico: rRNA*

La participación del RNA como agente catalítico es relevante para el proceso de continuidad histórica sólo en el surgimiento del elemento que se convertirá en el RNA ribosomal. La complejidad de tal molécula, incluso sin tomar en cuenta el andamiaje protéico con el que se asocia, hace que analizar su surgimiento sea un trabajo extensivo. En el presente escrito es necesario limitarse a señalar ciertas características que pueden apuntar hacia un escenario de su origen.

Se debe de tomar en cuenta que aún siendo el centro de la síntesis peptídica, la función y estructura del ribosoma no se ha comprendido plenamente. Una reacción unidimensional como la aminólisis del enlace estérico por el aminoacil-tRNA en turno puede ocurrir por varios mecanismos. En esto interviene, por ejemplo, la naturaleza switteriónica de los aminoácidos, ya que dependiendo de las condiciones del medio los mecanismos de movimiento de protones pueden variar. Es cierto que en soluciones acuosas los aminoácidos generalmente se encuentran en su forma protonada (es decir, su grupo amino está en la forma NH_3^+), pero el ambiente creado por la maquinaria del ribosoma puede variar estas condiciones. Además, es todavía una incógnita si el rRNA toma parte de la reacción

Hasta este momento, parecen vislumbrarse tres datos importantes:

- 1) La catálisis de polimerización de proteínas se lleva a cabo en un sitio activo compuesto por RNA (el rRNA de la subunidad más grande), por lo que el ribosoma se puede considerar como una ribozima modificada.
- 2) La reacción ribosómica de polimerización es mucho menos eficiente que las reacciones enzimáticas tradicionales; tal cosa probablemente se deriva del carácter ribozomático de la reacción
- 3) La reacción no es una reacción de hidrólisis o aminólisis ácido-base, como sugeriría la bioquímica tradicional basada en soluciones acuosas; parece ser que se trata de un proceso en el que tienen papeles importantes fenómenos que lo hacen fisicoquímicamente necesario, como la exclusión de moléculas de agua, posicionamiento y el consecuente cambio de la entropía de la reacción.

La reacción de polimerización peptídica primitiva puede situarse en varios ambientes posibles, dado a que el ambiente inmediato lo da el catalizador. Lo fundamental en cuanto al rRNA es que posiblemente sea el único fósil verdaderamente

catalítico del surgimiento de la vida. Ya se mencionó que los detalles de la hipótesis del mundo del RNA (es decir, la posibilidad catalítica de los diferentes RNA, su plegamiento y su posible origen) quedan fuera de los límites de este trabajo. Sin embargo, siguiendo la lógica de esta propuesta, los RNA catalíticos que se pudieron generar aleatoriamente sólo habrían servido para fabricar los “bloques de construcción” moleculares para un sistema riboprotéico, pero no como depósito de la información que a la larga habrá de producir la información biológica actual. Por supuesto, esa acción catalítica, aunque no heredada, puede ser fundamental para la estabilidad y mantenimiento del sistema en las primeras etapas de establecimiento de sistemas complejos metabólicos.²⁸

La reacción del rRNA, por otra parte, no es una especialmente complicada. Illangasekare (Illangasekare 1999) ha demostrado que una molécula bastante pequeña de RNA (29 bases) es capaz de realizar (no específicamente) tanto aminoacilación como transferencia de péptidos a moléculas de RNA. Es probable que esta función haya surgido como una molécula catalítica de RNA al azar, y después se haya elaborado sobre la estructura pre-existente, sin modificar en centro activo, su parte más importante, que siempre ha sido constituida por RNA. Tomando en cuenta que es una reacción fundamental para la célula y aún así no es excepcionalmente eficiente, se podría sugerir que si en efecto el mundo del RNA existió y tuvo una importancia fundamental para la vida actual, existirían funciones metabólicas compuestas de ribozimas o riboproteínas. La ausencia de tales características parece indicar que la historia fundamental de la vida comienza con las proteínas, de alguna o de otra manera. Es altamente probable que el origen del sistema ribonucleoprotéico de traducción se localice en etapas muy tempranas de la vida; análisis recientes de la identidad del último ancestro común (LCA) muestran que las proteínas implicadas en la traducción, con dominios de unión a RNA, están altamente conservadas y son de las más antiguas (Delaye 2005, Becerra 2007)

²⁸ Otras funciones conocidas y difundidas del RNA catalítico son el intrón auto-catalítico (esta vez, auto-catalítico quiere decir que realiza una actividad de edición sobre sí mismo, no auto-replicante) de *Tetrahymena* (la primera actividad catalítica del RNA que fue descubierta) y la ribozima RNasaP, que degrada el RNA. Comparadas con las funciones ribosomales, estas ribozimas parecen ser complementos y no partes esenciales del sistema vivo.

4. Orígenes del tRNA

El tRNA en la célula funciona como el adaptador que conecta la información del RNA mensajero con su traducción protéica. Es una molécula que la mayoría de las veces lleva a cabo su función sin ser traducida.²⁹

La estructura secundaria del tRNA es cruciforme y comprende cuatro brazos y tres asas. El aminoácido será unido covalentemente al extremo 3' del aceptor, y el brazo del anticodón contiene la secuencia complementaria al codón específico. Los otros dos brazos parecen tener función únicamente estructural. Se distinguen por los nucleótidos modificados que presentan regularmente: el brazo T ψ C contiene la secuencia (ribo)timidina – pseudouridina – citosina, y el brazo D contiene dos o más nucleótidos de dihidrouridina.

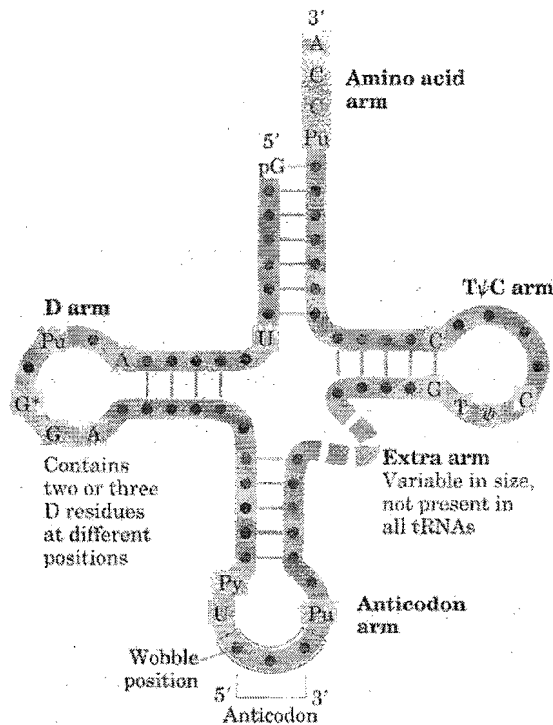


Fig. 21 - Estructura general de los tRNA (Lehninger 2000)

La estructura secundaria del tRNA es simétrica; Eigen y Schuster mencionan que esta característica es derivada de la antigüedad de la molécula. Como hacen notar en su artículo, la hebra + y la hebra - son extremadamente similares (es decir, son

²⁹ En ocasiones, como se discute más adelante, el tRNA puede tener una función de mensajero.

complementariamente simétricas), por lo que tomar un tRNA como modelo para replicación produce 2 moléculas de tRNA. Las hebras complementarias no son idénticas, por lo que la replicación de tRNAs, además de los errores cometidos por la polimerasa darían lugar a una población de moléculas similares que se comportarían como una solo entidad evolutiva en el modelo de hiperciclo.

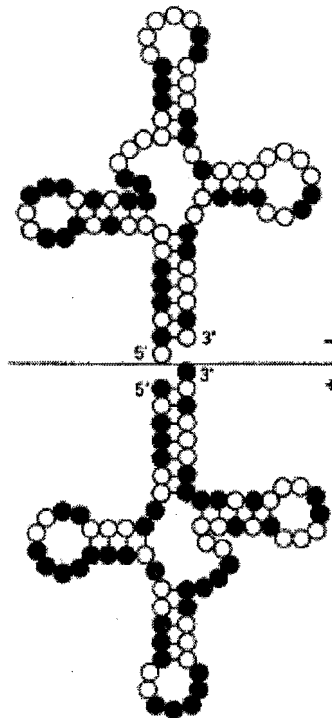


Fig. 22 - Simetría de una molécula de tRNA y su complementaria

Massimo di Giulio, en una serie de artículos, va más allá. Él propone que la molécula misma del RNA, que es simétrica internamente, proviene de la unión de dos “horquillas” de RNA unidas, ambas con la identidad del tRNA marcada por la presencia de un anticodón. La unión de estas horquillas provocaría que un anticodón fuera conservado para funcionar en la maquinaria ribosomal, y el otro se convirtiera en lo que hoy es la región que permite la identificación del tRNA por las aminoacil-transferasas. Di Giulio (2006) se basa en gran parte en algunos organismos, como *Nanoarchaeum equitans*, cuyo tRNA se transcribe en dos mitades, que se encuentran en lugares diferentes del genoma; también en la simetría ya mencionada de la molécula del tRNA. Partes fundamentales de las teorías de Di Giulio han sido corroboradas por Nagaswamy y Fox (Nagaswamy 2005), cuyos experimentos y análisis “demuestran inequívocamente

que la ligación de minihélices de RNA pueden resultar en RNAs complejos cuya estructura es muy similar a los tRNA modernos”.

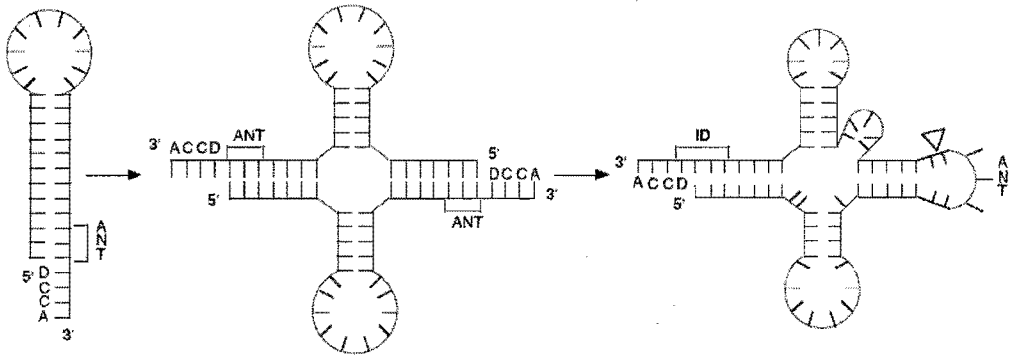


Fig. 23 – Modelo de Di Giulio (2006) para la evolución del tRNA, basado en la duplicación de minihélices. ACCD = brazo aceptor. ANT = anticodón.

Ambas observaciones son compatibles con la propuesta de este trabajo. Aunque este trabajo se ocupe estrictamente de los orígenes primitivos del la asociación aminoácido-RNA y el surgimiento de la codificación, se puede vislumbrar un origen del sistema de traducción con adaptadores a partir del modelo propuesto aquí. Es posible aprovechar la simetría de muchos grados del tRNA. Considérese nuevamente un conjunto de trímeros de RNA unidos selectivamente a aminoácidos. Ya se ha mencionado que es posible polimerizar dichos oligómeros con mayor eficiencia que utilizando monómeros aislados. Así, mediante una serie de pasos de polimerización de oligómeros y elongación utilizando moldes intramoleculares es posible lograr una molécula simétrica de RNA, como se muestra en la ilustración. Orgel (1995) ya ha mencionado la facilidad con la que la tercera estructura (en forma de J) puede utilizarse a sí misma como molde intramolecular.

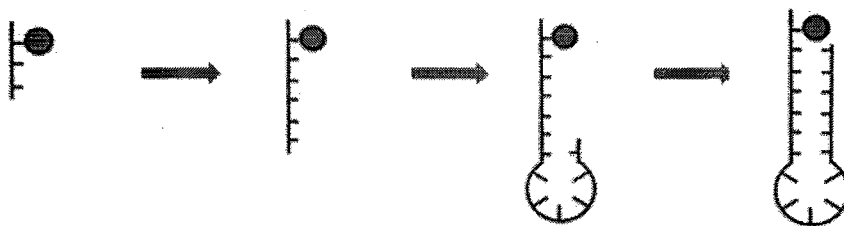


Fig. 24 – Formación de una minihélice de tRNA a partir de un complejo RNA-aminoácido

Existen evidencias de que ha ocurrido este mecanismo:

- Bajo ciertas condiciones, algunos RNA de transferencia pueden funcionar como mensajeros. Éstos son los llamados tmRNA, y tienen funciones relativamente simples. No codifican para ninguna proteína funcional, sino que al ser utilizados como mensajeros agregan una serie de aminoácidos que marcan a las proteínas nacientes como defectuosas, y éstas serán degradadas. También funcionan para “rescatar” ribosomas que por una razón u otra han quedado unidos a la proteína que sintetizaban. (Gillet 2001)
- El segmento en donde se une el aminoácido es de aparición tardía. En algunos organismos, primero se sintetiza el tRNA completo y después se agrega, post-transcripcionalmente, la secuencia RCCA. (Altman, 1999)
- Aun considerando la variedad y la antigüedad de las moléculas de RNA de transferencia, se esperaría encontrar alguna reminiscencia del mecanismo; la única posible es la presencia de codones en el tRNA, específicamente en el brazo aceptor.
- En el tRNA se concentra una gran cantidad de nucleótidos extraños, que como la inosina, pseudouridina, dihidrouridina y otras modificaciones de los nucleótidos normales. Se ha publicado aproximadamente 80 modificaciones en el tRNA citoplásmico (Sprinzl 2005) y se ha demostrado que tales modificaciones son fundamentales para el funcionamiento correcto de la molécula, tanto en el reconocimiento del codón por el anticodón como en el reconocimiento de la molécula por el ribosoma o las aminoacil-tRNA-sintetasas (Yarian 2002). Tales modificaciones sugieren que el tRNA tuvo sus orígenes en un contexto en el que la síntesis de nucleótidos era indiscriminada, es decir, no mediada por ninguna enzima específica, sea ésta de naturaleza protéica o nucléica.
- Una regularidad inexplicable de los tRNA es su periodicidad, de tres en tres, de su contenido de purinas y pirimidinas; tal cosa llevó a Crick a proponer su teoría de RNY, en la que los mensajes alternaban preferentemente ese tipo de nucleótidos (es decir, purina-nucleótido-pirimidina). Si los trímeros del inicio de la información biológica tenían ese sesgo estructural, puede explicar tal regularidad del tRNA.
- La simetría de las moléculas ya ha sido notada por Eigen y Schuster, y fue llevada más allá por Di Giulio. También es necesario mencionar que Orgel

reportó que la polimerización intramolecular tiene eficiencia alta, utilizando a la misma secuencia de RNA como molde y como primer. (Orgel 1995)

5. *Aminoacil-tRNA Sintetasas (aaRS)*

Dado a que existen evidencias de que las aaRSs tuvieron su origen después del surgimiento de las moléculas de tRNA (Ribas de Pouplana 1998), lo que concuerda con el modelo presentado aquí, es posible afirmar que el origen del código genético no fue dependiente de las aaRSs, ni que podremos encontrar propiedades reveladoras en ellas para proponer nuevos caminos de evolución temprana. Además, siendo construcciones protéicas, cualquier modificación que causen a la teoría del origen del código genético tiene el riesgo de causar una nueva paradoja del huevo y la gallina. Sin embargo, su surgimiento ocurrió eventualmente, por lo cual podremos asumir que fue un paso necesario para la evolución de una codificación eficiente; es pertinente, pues, mencionar unas cuantas de sus características.

Existen dos tipos de aminoacil-tRNA-sintetasas; cada una codifica para diez aminoácidos distintos. El único aminoácido cuyos tRNA son sintetizados por los dos tipos de enzimas es la Lisina. La tabla siguiente enlista los aminoácidos dependientes de cada tipo de aa-tRNA-sintetasa, y el mecanismo de reacción de cada una:

Como se puede observar, existe una tendencia de los aminoácidos hidrofóbicos a tener aa-RS de clase I. Sin embargo, si estas enzimas surgieron en etapas posteriores al establecimiento del código genético, las posibles regularidades han sido derivadas de las propias regularidades del código genético. En algunos casos, como las parejas Ile-Leu, Glu-Gln y Asn-Asp pueden representar una ambigüedad primitiva del código en el momento del surgimiento de dichas enzimas.

Tabla 2 – Clasificación de los aminoácidos esenciales según su tipo de aa-RNA sintetasa

Clase I	Clase II
Arginina	Alanina
Cisteína	Asparagina
Glutamina	Aspartato
Glutamato	Glicina
Isoleucina	Histidina
Leucina	Lisina
Metionina	Fenilalanina
Triptofano	Prolina
Tirosina	Serina
Valina	Treonina

6. Otras funciones posibles del RNA

Como ya se ha mencionado, proponer un modelo en el que la codificación por tripletes se encuentre en el origen de la vida no excluye el importante papel jugado por el RNA. Simplemente se indica que cualquier avance bioquímico logrado por un sistema de ribozimas se tendría que “reinventar” en el cambio hacia el sistema actual. Sin embargo, debido a la capacidad potencialmente catalítica del RNA y sobre todo su presencia obligada en algún punto de la evolución temprana de la vida, hace que sea coherente asumir que algunas reacciones se llevaban a cabo por moléculas contingentes de RNA. Sin embargo, y a diferencia de las teorías del mundo del RNA, se propone aquí el supuesto de que dichas moléculas, en caso de existir, no funcionaban como depósitos de información ni se necesitaban heredar, sino simplemente tomaban parte del sistema al asistir en procesos metabólicos. La lectura de una molécula de ácido nucléico es un proceso más complejo que el simple apareamiento lineal de bases nitrogenadas, por lo que un metabolismo completo de ribozimas es posible en teoría pero altamente improbable como un evento a largo plazo.

En la tabla se muestran algunas de las funciones en las que podría haber asistido el RNA (tomada de Joyce 1999, en donde se citan los trabajos que reportan dicha actividad). Sin embargo, una de las funciones de importancia fundamental es la función del RNA es como factor de unión de moléculas. Ya se habló del ribosoma como una ribozima, probablemente la única función catalítica del RNA extendida en todos los

dominios de la vida; también podemos considerar una ribozima no heredada cuya función sea ligar fragmentos de RNA.

Tabla 3 - Algunas funciones catalíticas de los ácidos nucleicos

Transferencia de fosfodiéster
Ruptura de enlace fosfodiéster
Ligación de polinucleótidos
Fosforilación de polinucleótidos
Polimerización de mononucleótidos
Aminoacilación de polinucleótidos
Transferencia de aminoácidos
Formación de enlaces peptídicos

Propuestas a largo plazo

Con base en la propuesta anterior es posible visualizar una serie de investigaciones a largo plazo, para validar los puntos fundamentales de la teoría. Como en la mayoría de las hipótesis científicas, existen dos acercamientos posibles: la realización de modelos matemáticos y la experimentación directa en el laboratorio.

Experimentos in vitro

En mi opinión, los puntos fundamentales a probar son los que conciernen la viabilidad de polimerización de distintas moléculas: la polimerización de los complejos de trímeros de RNA unidos a aminoácidos específicos, y posteriormente la polimerización de los aminoácidos en péptidos. Ya se ha demostrado que moléculas relativamente sencillas pueden funcionar como activadores de aminoácidos o aminoacil-RNA primitivos (Illangasekare 1999, entre otros). Los experimentos que serían necesarios a continuación estarían encaminados a verificar si no existe ningún impedimento estérico para que se apareen dos bases nitrogenadas del RNA si una de ellas tiene un aminoácido unido covalentemente.

Además de verificar la viabilidad estérica de tal construcción, se pueden hacer experimentos para evaluar la probabilidad de surgimiento de polímeros. Por ejemplo, se puede poner a una cadena de tripletes unidos a aminoácidos en presencia de la ribozima

ribosomal, para demostrar si es posible que un ancestro de este ácido nucleico haya tomado parte en este origen de la vida. La fidelidad de la producción enzimática es un punto menos relevante: ya se ha comentado que probablemente existía un grado alto de ambigüedad en el código genético. Incluso existen autores (Trifonov 2000) que han propuesto que originalmente las proteínas se seleccionaban exclusivamente por su conformación física globular, y por tanto el código genético no necesitaba traducir más que dos tipos de aminoácidos: hidrofóbicos e hidrofílicos. Tal mecanismo podría producir poblaciones de proteínas cuya similitud, por más burda que hubiera sido, era suficiente para otorgar la misma función.

Tal programa puede ser expandido hacia el origen del tRNA moderno. Ya se ha demostrado que para que los tRNAs sean cargados correctamente únicamente necesitan de la mini- hélice del brazo aceptor. ¿Qué tan eficientemente se polimerizaría una proteína relativamente corta, en la presencia de una ribozima ribosomal y con la ayuda de estos proto-tRNAs? ¿Qué tanta diversidad de tRNAs específicos se pueden lograr con un conjunto de mini-hélices complementarias, es decir, con el mecanismo propuesto por Di Giulio? El programa experimental que derivaría de estas preguntas sería sin duda costoso, pero con la creciente diversidad de técnicas de manejo de RNA es de realización factible; definitivamente la dificultad de realizarlo sería menor a la que existe en los experimentos de condiciones hipertermófilas e hiperbarófilas.

Experimentos in silico

También se debe de llevar a cabo un modelo matemático para describir la dinámica de los sistemas aquí propuestos. Ciertamente, los experimentos in vitro aportan datos irrefutables e incluso en ocasiones inesperados; pero un modelo matemático posee un poder de predicción al cual se puede apelar al realizar el diseño experimental "húmedo". Los modelos de Kauffman y de Fontana demuestran que se pueden asociar parámetros relativamente simples con comportamientos complejos.

Tales son los trabajos que se pueden utilizar como base: uno de los proyectos que se propone aquí es incorporar a tales modelos un sistema semiótico, justo como el código genético pero sin estar sesgado al sistema de traducción. Es decir, se puede construir un sistema en el que la señal y la respuesta estén ligadas por un sistema de símbolos arbitrarios y observar su desarrollo a lo largo del tiempo. Hago notar que tal

sistema no está necesariamente restringido a un funcionamiento intracelular: al poseer la facultad de modelar libremente, el signo y su objeto pueden pertenecer a dos sistemas distintos y ser relacionados arbitrariamente por un factor que pueda ingresar a ambos. Considérese, por ejemplo, a los sistemas análogos ya existentes, como por ejemplo las señales endócrinas o el reconocimiento inmunológico de los animales complejos.

Es importante señalar la similitud entre el funcionamiento del modelo aquí propuesto (en sus primeras etapas, al menos) y el funcionamiento viral: ambos pueden ser vistos como individuos con poca información biológica codificada que dependen de un contexto metabólico robusto (el metabolismo prebiótico, en el caso del origen de la vida; el del hospedero, en caso de los virus). La evolución de ambos sistemas puede ser modelada matemáticamente y comparada, para establecer los umbrales necesarios para mantener una estabilidad mínima.

Umbrales

Durante la elaboración de este trabajo llamó mi atención una cantidad importante de umbrales (*thresholds*) que existen a diversos niveles, y que marcan la aparición de las características de la vida. El primero y el más evidente es el umbral de energía mínima que la vida necesita para llevar a cabo sus funciones. Sin embargo, existen otros tantos que es necesario analizar en investigaciones futuras.

- Prigogine menciona umbrales de complejidad molecular (tanto en número de partículas como en distancia del equilibrio termodinámico) en el que el comportamiento repentinamente se vuelve altamente organizado.
- Kaufmann hace mención específica a umbrales de complejidad en sus redes moleculares y en el comportamiento de sistemas dinámicos. Hace notar que el orden surgido al borde del caos es una propiedad emergente no deducible del comportamiento o de las leyes que rigen a los elementos individuales, dependiente de la interacción constante del sistema con su medio.
- Trevors y Abel mencionan un umbral de complejidad algorítmica mínima para que una secuencia de información tenga un sentido “funcional” o semántico. Este umbral se relaciona significativamente con los anteriores porque los tres autores relacionan el surgimiento de un comportamiento organizado con la cercanía de un régimen aleatorio. (Ver fig. 4)

- Del modelo del hiperciclo de Eigen y Schuster se desprende otro conjunto de umbrales, cuya presencia determina la aparición de errores fatales o catástrofes. Tales errores tienden a desviar recursos del hiperciclo hacia un solo componente, disminuyendo gradualmente la complejidad y difundiéndose en última instancia al sistema. Tales errores son los siguientes:
 - Catástrofe del error: Las moléculas se deben de replicar con un margen mínimo de fidelidad; de otra manera, se producirían grandes números de moléculas inservibles para el hiperciclo, utilizando los recursos disponibles
 - Catástrofe de los parásitos: Pueden existir, ya sea por mutaciones de las moléculas del ciclo o por síntesis externa, moléculas que acaparen los recursos del hiperciclo dirigiéndolas hacia su propia replicación sin aportar al siguiente paso del proceso, causando así que el ciclo se detenga en ese punto.
 - Catástrofe del corto circuito: El hiperciclo requiere, para su mantenimiento a lo largo del tiempo, un número mínimo de elementos. Si, por mutación o síntesis externa, surge una molécula que cree un bypass y disminuya la longitud del ciclo a un número de elementos menor al necesario, el hiperciclo perderá complejidad gradualmente y nuevamente el resultado será su difusión en el medio.

En trabajos posteriores se deberán implementar estos umbrales o restricciones en modelos matemáticos para evaluar el significado que tienen en el origen y el mantenimiento de la complejidad biológica.

Metabolismo no específico

Las primeras etapas del código genético vieron la luz en un mundo metabólico de especificidad reducida; las enzimas protéicas no estaban disponibles por razones obvias y las reacciones que se llevaban a cabo eran poco finas, sin importar qué tipo de metabolismo original tomemos en cuenta. El presente trabajo sufre de una simplificación conceptual extrema: toma en cuenta únicamente los compuestos que se conocen actualmente en los seres vivos. La falla reside, sobre todo, en el conjunto de cuatro nucleótidos del RNA y 20 aminoácidos esenciales. Es lógico suponer que el código genético se estableció como un subconjunto de interacciones que incluían una

diversidad mucho mayor: basta ver la variedad de bases que se incorporan al tRNA actual.

Es imposible saber cuál era la diversidad molecular disponible en el inicio de la vida. Se ha tratado de incluir en las propuestas de este trabajo extendiendo, por ejemplo, la relación codón-aminoácido a una relación más profunda: codón-aminoácido-precursor de aminoácido. También se han mencionado, como en el trabajo de Copley, posibles asignaciones primitivas a otros aminoácidos, como ornitina y 2,4-diaminobutirato (posibles precursores de Arg) u homoserina (posible precursor de Thr). Sin embargo, sólo se ha trabajado con las cuatro bases del RNA por simplicidad; un trabajo más completo exploraría las posibles interacciones de otras bases y su relevancia para el establecimiento del código genético. Tal vez el código genético sea una depuración de un código más grande potencial, que involucra un número mayor de nucleótidos.

Un experimento multidisciplinario

Es posible proponer un experimento que utilice recursos tanto informáticos como de laboratorio para comprobar la viabilidad de un código ambiguo. En dicho experimento se parte de un código de alta ambigüedad, como el propuesto en este trabajo, y se utiliza para traducir, *in silico*, la secuencia de mRNA de una proteína moderna, cuya actividad sea fácilmente cuantificable. Debido a la ambigüedad, diversas iteraciones de la traducción producirán proteínas con diferencias de aminoácidos (por ejemplo: los codones que comienzan con GA podrán codificar para cualquier aminoácido hidrofílico). Después de obtener un número considerable de proteínas similares pero no idénticas, se tomará una muestra significativa, cuya similitud con la proteína original sea variable. Mediante clonación o síntesis artificial tales proteínas se podrán aislar físicamente, y su actividad podrá ser evaluada.

La hipótesis de este experimento es que la ambigüedad del código afecta cuantitativamente (acaso también cualitativamente) la actividad enzimática, pero no la anula, por lo cual no es necesario un código unívoco primitivo para comenzar un metabolismo llevado a cabo por proteínas codificadas por RNA.

4. Conclusión

Preguntas abiertas y argumentos principales

Antes de concluir, es necesario hacer notar que aunque este trabajo es un acercamiento a un posible origen de la vida, hay varios puntos que se evitan a pesar de su gran importancia. En general, tales puntos no fueron tratados por ser problemas de gran complejidad cuyo alcance queda fuera de esta discusión, cuyo punto exclusivo es proponer el origen de una codificación primitiva pero cualitativamente igual a la actual (esto es, tripletes de bases de RNA que codifican para aminoácidos). Algunos de estos problemas son:

- El origen de la quiralidad. Los compuestos biológicos tienen una quiralidad determinada, cuya causa desafía las teorías actuales. Se ha mencionado este problema en la sección de problemas de la síntesis prebiótica del RNA, pero únicamente de manera superficial.
- El ambiente prebiótico en el que surgió el código y la fuente de “bloques de construcción” necesarios para llevar a cabo el modelo. Si asumimos, como ya se ha mencionado, un escenario como el propuesto por Wächtershäuser, el medio en el que se lleva a cabo todo el proceso es similar a las ventilas hidrotermales mediante un metabolismo inicial en superficies. Sin embargo, las ventajas y desventajas del modelo, como el cambio de la vida media de los compuestos por las condiciones, no ha sido discutido.
- El surgimiento de la compartimentalización, que es una característica fundamental de la vida en el punto de vista de Tibor Gánti (más aún que la reproducción o la adaptación). Existen varios modelos para el surgimiento de membranas, algunos incluso propuestos por el mismo Wächtershäuser para dar continuidad a su teoría, pero el metabolismo de lípidos, el origen de micelas, sus propiedades fisicoquímicas y demás mecanismos implicados quedan fuera de las perspectivas de este trabajo.
- La identidad del Último Ancestro Común (LUCA, por sus siglas en inglés); este punto, sin embargo, parece encontrarse en una etapa muy posterior al surgimiento de la vida: antes de la separación evolutiva de los dominios de

Woese pero después de obtener un metabolismo complejo y utilizar el DNA como almacén informativo (Becerra 2007).

- El surgimiento del DNA como un almacén de información de más estabilidad que el RNA.

Como cualquier hipótesis, ésta está basada sobre supuestos para simplificar la argumentación. Es imposible, en un trabajo finito, abarcar temas tan amplios como la geofísica terrestre, la resonancia de los anillos heterocíclicos de los nucleótidos o la probabilidad de surgimiento de una secuencia polimerizadora. En resumen, los argumentos más importantes de este trabajo son los siguientes:

1. Las características físicas de la vida pueden ser descritas por leyes físicas. La vida, como proceso, posee características termodinámicas particulares que hacen posible su inclusión en un conjunto de sistemas complejos, i.e., sistemas disipativos como los descritos por Prigogine.
2. La codificación (es decir, el manejo de un lenguaje simbólico utilizando medios puramente físicos) es una característica de la vida. Es el factor que permite la disyunción entre el genotipo y el fenotipo, y todas las implicaciones que trae consigo esta separación.
3. Existen problemas, sobre todo lógicos, que dificultan el paso del llamado mundo del RNA a los sistemas vivos actuales. Por lo tanto, es necesario explorar la posibilidad de que el origen de la codificación sea el mismo que el de la información biológica.
4. Toda teoría evolutiva debe de basarse en el Principio de Continuidad (Crick 1968). Por lo tanto, es necesario explorar la posibilidad de que la codificación se dio desde sus inicios como la actual: en forma de codones. Se propone en el presente trabajo un mecanismo plausible.

El Origen de la Vida (2)

Aunque el modelo presentado aquí es distinto al llamado mundo del RNA en puntos cruciales, no lo es en su totalidad. En ambos existe, por ejemplo, el problema de la síntesis prebiótica de nucleótidos, cuya respuesta elude a los esfuerzos de los investigadores aún hoy. También hay puntos en común: el modelo de Wächtershauser, aunque no ha sido demostrado experimentalmente en su totalidad, puede considerarse

como una fuente de “bloques de construcción” de las macromoléculas, sean éstas péptidos o ácidos nucleicos. No es posible insistir lo suficiente en un argumento central: independientemente de la existencia de un mundo de RNA como se ha propuesto o de otra alternativa prebiótica, el RNA es un elemento necesario en la vida actual. Se le puede encontrar a lo largo del espectro del árbol de la vida, realizando en todos los organismos las mismas funciones catalíticas e informacionales. Esto indica su importancia y la necesidad de considerar su presencia desde las primeras etapas de la vida. Sin embargo, hay entre ambas propuestas diferencias fundamentales, surgidas únicamente por tratar de llevar a último término los razonamientos que se derivan del mecanismo del surgimiento del código genético. En caso de que se demuestre la imposibilidad de algún paso de tal mecanismo, espero que los caminos señalados aquí sirvan para seguir considerando alternativas viables.

En primer lugar, este modelo no considera a una molécula de RNA autorreplicadora. Aunque la replicación como tal es importante, una molécula informativa cuyas instrucciones se basan únicamente en su perpetuación aporta poco al sistema; su aparición es posible, incluso probable, pero para su funcionamiento tal aparición debe surgir en una red de mantenimiento y abastecimiento de materia prima. Ya se han mencionado otros problemas en una sección anterior.

En segundo lugar, se considera como accesorios a los RNA catalíticos que realicen funciones metabólicas. Esto no les resta importancia: posiblemente sin la posibilidad catalítica de las moléculas de RNA la vida nunca se hubiera materializado. Con el término “accesorios” me refiero a que tales moléculas no fueron heredadas a los seres vivos que conocemos actualmente, y por tanto la línea genealógica de la vida debe de ser buscada en otro lugar. Una excepción de importancia capital es el rRNA: su actividad catalítica parece ser un requisito ineludible de la polimerización codificada de péptidos. Como el modelo del mundo del RNA, en este modelo también cabe considerar la aparición espontánea de oligopéptidos temporales y no heredados, con actividad catalítica poco eficiente y poco específica pero valiosa en las primeras etapas de la vida.

Tales argumentos llevan a la tercera diferencia, en la que he insistido varias veces a lo largo del trabajo. En el mundo del RNA, los primeros genomas se constituirían de RNA autorreplicante y posteriormente de dicho RNA y RNA que codifica para ribozimas. En esta propuesta, se considera que los elementos actuales estaban

prefigurados de manera primitiva en las etapas tempranas de la vida: código, mensaje, adaptadores (tRNA) y polimerizadores (rRNA) tienen su origen en el principio de la información biológica. Tal vez sus funciones se superponían, como es el caso ya mencionado del tRNA mensajero, y con el paso del tiempo las identidades de cada una se fueron afinando y tomando la forma que se conoce actualmente. En las figuras se esquematizan las diferencias entre ambas propuestas.

Ambos esquemas son una gran simplificación y su formato es distinto, pero muestran las diferencias entre una concepción y otra. Mientras que el modelo general del mundo del RNA propone un metabolismo controlado casi exclusivamente por ribozimas y un surgimiento gradual de las proteínas codificadas, en el presente trabajo se propone un origen común de los péptidos codificados y los RNA mensajeros y de transferencia. El rRNA, por otra parte, se propone como uno de los únicos elementos heredados del grupo de enzimas (sean de naturaleza protéica o nucléica) que pudieron haber apoyado el surgimiento de la vida; todos los demás miembros de este grupo de tamaño desconocido se han perdido. Como se indica en la nota de la figura, se pueden incluir en este grupo otras ribozimas actuales de importancia, como la RNAsa P³⁰; sin embargo, son moléculas excepcionales, ya que como se ha mencionado la lectura, como es conocida actualmente, adquirió su importancia en un momento posterior.

³⁰ Además, recientemente se ha descubierto una RNAsa P compuesta exclusivamente de proteína (Holzmann 2008)

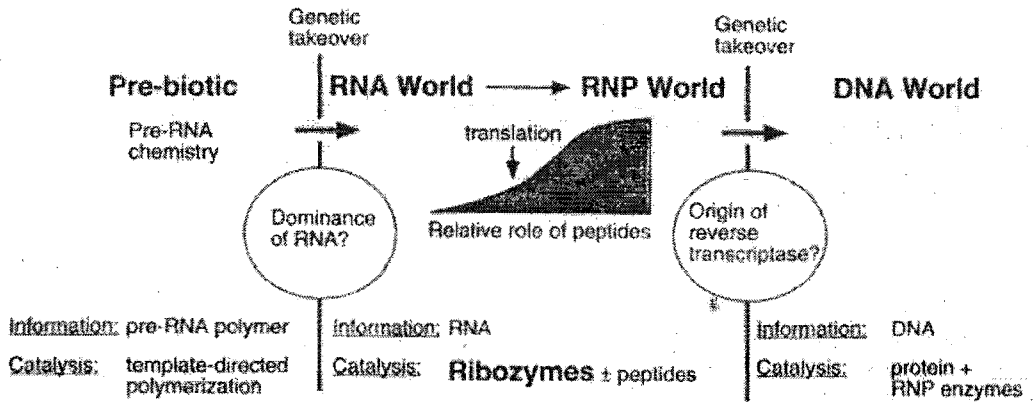


Fig. 25 - Modelo general del Mundo del RNA (tomado de Cech y Golden, 1999)

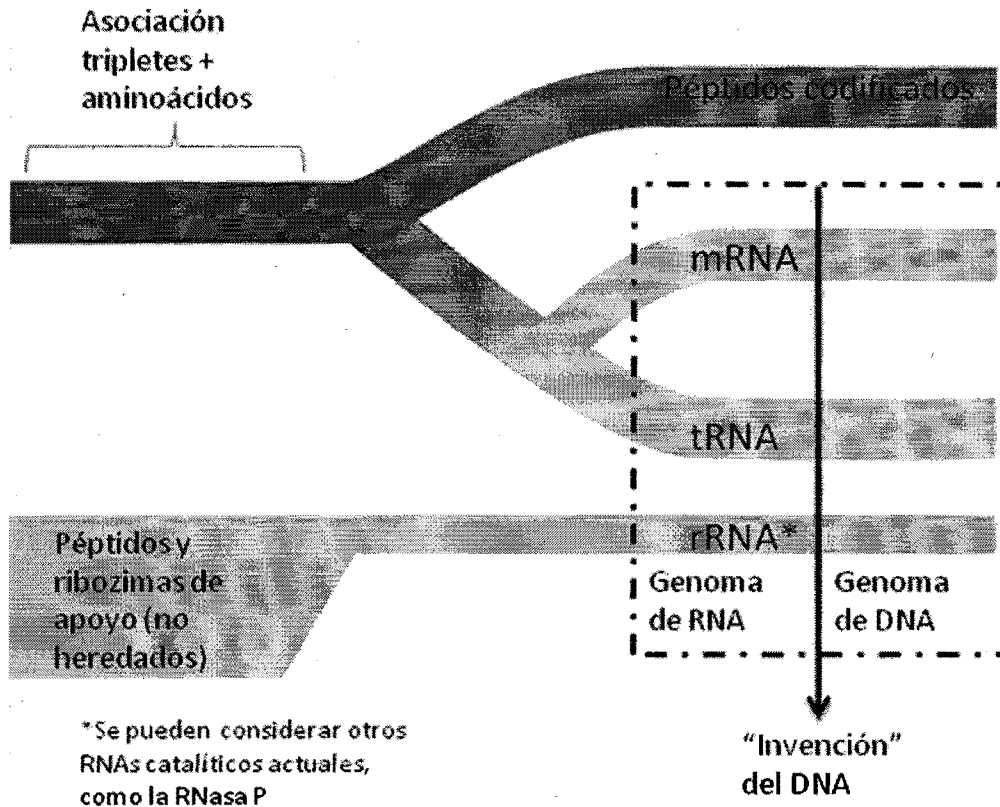


Fig. 26 - Historia propuesta de las moléculas actuales.

Epílogo

El 24 de septiembre de 1920 Henri Bergson presentó una conferencia en Oxford en la que hizo una afirmación sorprendente: el futuro puede afectar al presente y el presente puede afectar al pasado. Más aún, no hablaba en un tono relativista; por ejemplo, no se refería a que un evento que hoy nos parece maravilloso puede parecer terrible el día de mañana. Bergson se refería a que el futuro puede cambiar las posibilidades del presente de manera retroactiva. (Bergson, 2002)

Menciono el mismo ejemplo que él, para aclarar su afirmación. El texto de *Hamlet* ha sido posible, como una concatenación de símbolos, incluso como una secuencia de eventos e ideas, desde el principio de los tiempos, en el sentido en que no existe ninguna imposibilidad física irresoluble. Sin embargo, sin el advenimiento de la mente de Shakespeare, la tragedia del príncipe de Dinamarca es imposible. *Hamlet* no estaba almacenada bajo llave o en la biblioteca de Babel de Borges, a la espera de que Shakespeare la encontrara. De la misma manera, la obra maestra literaria del futuro es posible (en el sentido de que es no-imposible), pero tal posibilidad sólo se materializa en el presente cuando la obra haya sido creada en el futuro. En sus palabras, la obra maestra futura *va a haber sido posible, pero todavía no lo es*.

La situación del arte es, como en muchos otros aspectos, paralela a la de la ciencia. No tiene mucho sentido calcular la probabilidad del surgimiento espontáneo de *Hamlet* como una concatenación de letras y afirmar que ese número nos da la probabilidad del surgimiento de la obra en una civilización cualquiera. El cálculo real es mucho más complejo. Es muy probable que si se corriera la “cinta” de la historia del universo, no surgiría la obra de Shakespeare. Las pequeñas perturbaciones, el *clinamen* de Epicúreo, probablemente ocasionarían un desarrollo universal completamente distinto. Aún más: si el fenómeno artístico surgiera en el universo, lo más probable es que lo haría en formas distintas e inimaginables para nosotros.

De la misma manera es válido pero poco informativo realizar cálculos para intentar analizar la probabilidad del surgimiento de un genoma cualquiera. La complejidad metabólica, ecológica, genética, etc, de los seres vivos actuales era imposible, en el sentido bergsoniano, hace cuatro mil millones de años. El desarrollo que siguió nuestro planeta hizo que el día de hoy la vida como la conocemos parezca

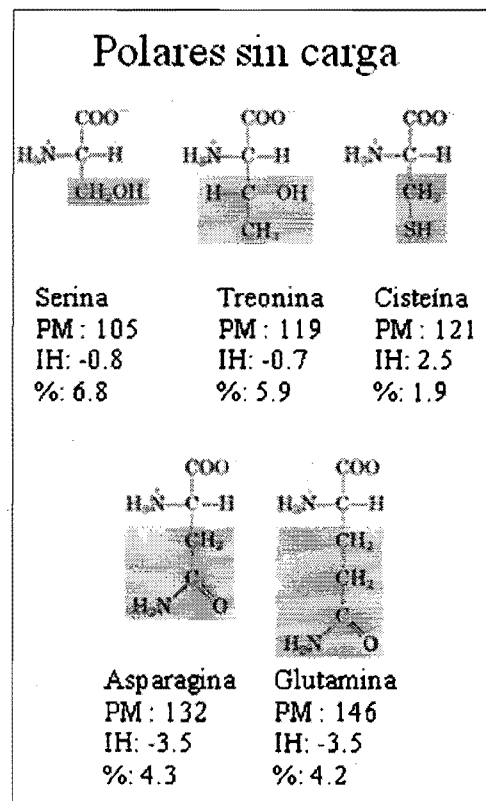
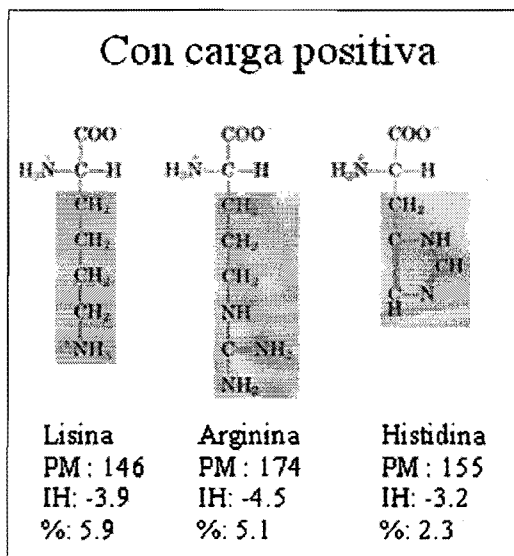
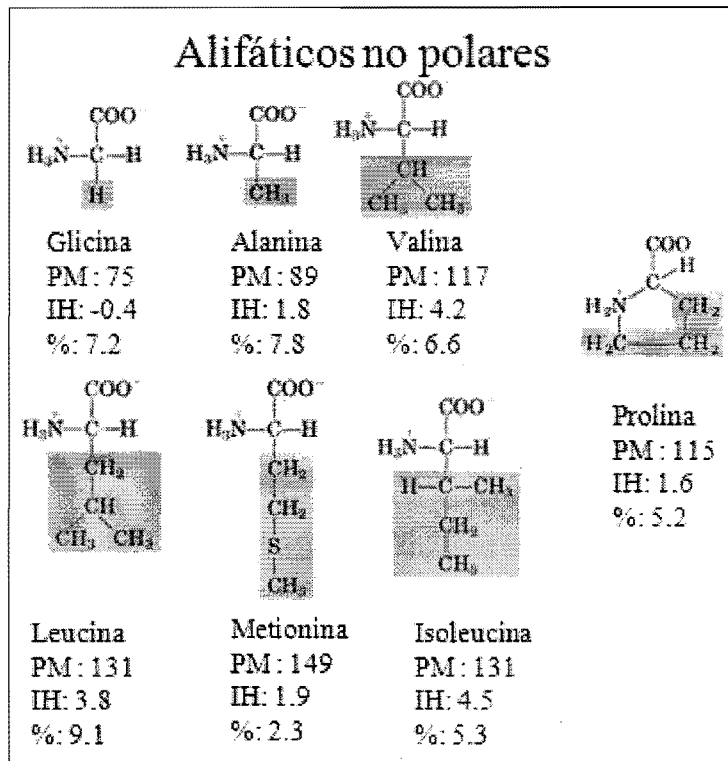
una necesidad; lo afirmo una vez más: si corriéramos la cinta de la historia del universo, la vida que acaso surgiría sería completamente distinta a la que nosotros conocemos.

Dado a este paradójico pero ilustrativo ejemplo, la deducción paso por paso del surgimiento de la vida como la conocemos es una tarea que posiblemente sobrepase las capacidades humanas. Como afirma Eschenmoscher (citado en Popa 2004), “el origen de la vida no puede ser descubierto, tiene que ser inventado”. El presente trabajo trata de tener siempre en cuenta tal principio. Se ha presentado aquí una propuesta del posible mecanismo de origen del código genético; aún cuando se comprueben sus supuestos fundamentales, incluso si se logra llevar a cabo en sistemas físicos, no existe ninguna garantía de que fue el mecanismo que se tuvo lugar en los albores de la vida terrestre.

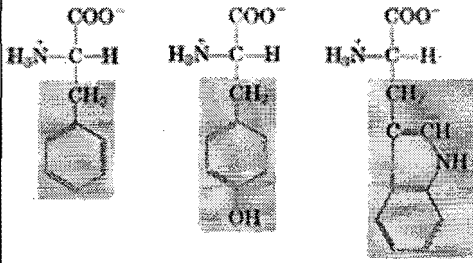
Así, creo que es posible afirmar que se llevaron a cabo los objetivos de este trabajo: “proponer un punto de investigación experimental del origen de la vida, y justificar la elección de ese punto determinado y no otro”. Es factible que los resultados obtenidos por este trabajo sean aplicados para construir sistemas simbólicos naturales, o que sean usados para buscar rastros de vida en otros lugares. Sin embargo, es prácticamente imposible que sean una narración fiel del nuestro pasado, por definición incognoscible.

Anexo 1: Estructura química de los aminoácidos esenciales (Lehninger 2000)

(PM = Peso Molecular; IH=Índice de Hidrofilia; %= Porcentaje promedio de aparición en proteínas)

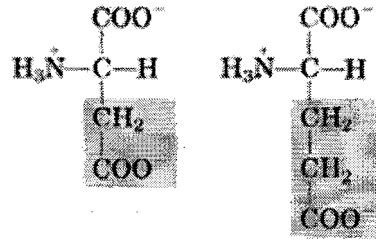


Aromáticos



Fenilalanina	Tirosina	Triptofano
PM: 165	PM: 181	PM: 204
IH: 2.8	IH: -1.3	IH: -0.9
‰: 3.9	‰: 3.2	‰: 1.4

Con carga negativa



Aspartato	Glutamato
PM: 133	PM: 147
IH: -3.5	IH: -3.5
‰: 5.3	‰: 6.3

Bibliografía citada

1. Abel, D.L. and J.T. Trevors, *Three subsets of sequence complexity and their relevance to biopolymeric information*. *Theor Biol Med Model*, 2005. **2**: p. 29.
2. Altman, S. and L. Kirsebom, *Ribonuclease P*, in *The RNA World*, R. Gesteland, J. Atkins, and T. Cech, Editors. 1999, Cold Spring Harbor Laboratory Press: New York.
3. Anet, F.A., *The place of metabolism in the origin of life*. *Curr Opin Chem Biol*, 2004. **8**(6): p. 654-9.
4. Aristóteles, *Metafísica*. Sepan Cuantos. 2004, México: Editorial Porrúa.
5. Ash, R.B., *Information theory*. 1965, New York: Dover Publications.
6. Barbieri, M., *The Organic Codes*. 2003, Cambridge: Cambridge University Press.
7. Becerra, A., et al., *The very early stages of biological evolution and the nature of the last common ancestor of the three major cell domains*. *Annu Rev Ecol Evol Syst*, 2007. **38**: p. 361-79.
8. Benner, S.A., et al., *Did the RNA world exploit an expanded genetic alphabet?* in *The RNA World*, R. Gesteland, T. Cech, and J. Atkins, Editors. 1999, Cold Spring Harbor Laboratory Press: New York.
9. Bergson, H., *The Possible and the Real*, in *Key Writings*. 2002, Continuum International Publishing Group.
10. Bieling, P., et al., *Peptide bond formation does not involve acid-base catalysis by ribosomal residues*. *Nat Struct Mol Biol*, 2006. **13**(5): p. 423-8.
11. Bielski, R. and M. Tencer, *A possible path to the RNA world: enantioselective and diastereoselective purification of ribose*. *Orig Life Evol Biosph*, 2007. **37**(2): p. 167-75.
12. Borsenberger, V., et al., *Exploratory studies to investigate a linked prebiotic origin of RNA and coded peptides*. *Chem Biodivers*, 2004. **1**(2): p. 203-46.
13. Bruce, A., et al., *Molecular Biology of the Cell*. 4th ed. 2002, New York: Garland Science.
14. Campbell, B.J. and S.C. Cary, *Abundance of reverse tricarboxylic acid cycle genes in free-living microorganisms at deep-sea hydrothermal vents*. *Appl Environ Microbiol*, 2004. **70**(10): p. 6282-9.

15. Cech, T. and B.L. Golden, *Mechanisms of RNA Catalysts*, in *The RNA World*, R. Gesteland, J. Atkins, and T. Cech, Editors. 1999, Cold Spring Harbor Laboratory Press: New York.
16. Cody, G.D., *Geochemical connections to primitive metabolism*. *Elements*, 2005. **1**(3): p. 139-43.
17. Cody, G.D., et al., *Assaying the catalytic potential of transition metal sulfides for abiotic carbon fixation*. *Geochim Cosmochim Acta*, 2004. **68**(10): p. 2185-96.
18. Cody, G.D., et al., *Primordial carbonylated iron-sulfur compounds and the synthesis of pyruvate*. *Science*, 2000. **289**(5483): p. 1337-40.
19. Cody, G.D., et al., *Geochemical roots of autotrophic carbon fixation: hydrothermal experiments in the system citric acid, H₂O-(±FeS)-(±NiS)*. *Geochim Cosmochim Acta*, 2001. **65**(20): p. 3557-76.
20. Copley, S.D., E. Smith, and H.J. Morowitz, *A mechanism for the association of amino acids with their codons and the origin of the genetic code*. *Proc Natl Acad Sci U S A*, 2005. **102**(12): p. 4442-7.
21. Crick, F.H., *The origin of the genetic code*. *J Mol Biol*, 1968. **38**(3): p. 367-79.
22. Crick, F.H., *Life Itself: Its origin and nature*. 1981, New York: Simon and Schuster.
23. Darwin, C., *The Origin of Species by means of Natural Selection*. 1968, London: Penguin Books.
24. Dawkins, R., *The Selfish Gene*. 1989, Oxford: Oxford University Press.
25. de Duve, C. and S.L. Miller, *Two-dimensional life?* *Proc Natl Acad Sci U S A*, 1991. **88**: p. 10014-7.
26. Delaye, L., A. Becerra, and A. Lazcano, *The last common ancestor: what's in a name?* *Orig Life Evol Biosph*, 2005. **35**(6): p. 537-54.
27. Di Giulio, M., *The ocean abysses witnessed the origin of the genetic code*. *Gene*, 2005. **346**: p. 7-12.
28. Di Giulio, M., *The non-monophyletic origin of the tRNA molecule and the origin of genes only after the evolutionary stage of the last universal common ancestor (LUCA)*. *J Theor Biol*, 2006. **240**(3): p. 343-52.
29. Di Giulio, M. and M. Medugno, *The historical factor: the biosynthetic relationships between amino acids and their physicochemical properties in the origin of the genetic code*. *J Mol Evol*, 1998. **46**(6): p. 615-21.
30. Dobzhansky, T., *Nothing in biology makes sense except in the light of evolution*. *The American Biology Teacher*, 1973. **35**: p. 125-9.

31. Dobzhansky, T., *Genética del proceso evolutivo*. 1975, México: Editorial Extemporáneos.
32. Durston, K.K., et al., *Measuring the functional sequence complexity of proteins*. *Theor Biol Med Model*, 2007. **4**: p. 47.
33. Eigen, M. and P. Schuster, *The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle*. *Naturwissenschaften*, 1977. **64**(11): p. 541-65.
34. Eigen, M. and P. Schuster, *The hypercycle. A principle of natural self-organization. Part B: The Abstract Hypercycle*. *Naturwissenschaften*, 1978. **65**(1): p. 7-41.
35. Eigen, M. and P. Schuster, *The hypercycle. A principle of natural self-organization. Part C: The Realistic Hypercycle*. *Naturwissenschaften*, 1978. **65**(7): p. 341-69.
36. Emmeche, C., *Closure, function, emergence, semiosis, and life: the same idea? Reflections on the concrete and the abstract in theoretical biology*. *Ann N Y Acad Sci*, 2000. **901**: p. 187-97.
37. Ertem, G., *Montmorillonite, oligonucleotides, RNA and origin of life*. *Orig Life Evol Biosph*, 2004. **34**(6): p. 549-70.
38. Etxeberria, A. and A. Moreno, *From complexity to simplicity: nature and symbols*. *Biosystems*, 2001. **60**(1-3): p. 149-57.
39. Fontana, W., *Algorithmic Chemistry*, in *Artificial Life II*, C. Langton, et al., Editors. 1991, Addison-Wesley: Santa Fe.
40. Freeland, S.J. and L.D. Hurst, *The genetic code is one in a million*. *J Mol Evol*, 1998. **47**(3): p. 238-48.
41. Futuyma, D.J., *Evolution*. 2005, Boston: Sinauer.
42. Gammerman, A. and V. Vovk, *Kolmogorov Complexity: sources, theory and applications*. *The Computer Journal*, 1999. **42**(4): p. 252-5.
43. Gánti, T., *The Principles of Life*. 2003: Oxford University Press.
44. Gillet, R. and B. Felden, *Emerging views on tmRNA-mediated protein tagging and ribosome rescue*. *Mol Microbiol*, 2001. **42**(4): p. 879-85.
45. Goodwin, B., *Las Manchas del Leopardo*. 1998, Barcelona: Tusquets Editores.
46. Gould, S.J., *Full House: The spread of excellence from Plato to Darwin*. 1997, New York: Three Rivers Press.
47. Gould, S.J., *The structure of evolutionary theory*. 2002, Boston: Harvard University Press.

48. Gould, S.J. and R.C. Lewontin, *The spandrels of San Marco and the Panglossian Paradigm: A critique of the adaptationist programme*. Proc Roy Soc Lon B, 1979. **205**(1161): p. 581-98.
49. Griffiths, P.E., *Genetic Information: A metaphor in search of a theory*. Philos Sci, 2001. **68**(3): p. 394-412.
50. Hoffmeyer, J., *Life and reference*. Biosystems, 2001. **60**(1-3): p. 123-30.
51. Holzmann, J., et al., *RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme*. Cell, 2008. **135**(3): p. 462-74.
52. Illangasekare, M. and M. Yarus, *A tiny RNA that catalyzes both aminoacyl-RNA and peptidyl-RNA synthesis*. Rna, 1999. **5**(11): p. 1482-9.
53. Israel, D. and J. Perry, *What is information*, in *Information, Language and Cognition*, P. Hanson, Editor. 1990, British Columbia Press: Vancouver.
54. Joyce, G.F., *Reactions catalyzed by RNA and DNA enzymes*, in *The RNA World*, R. Gesteland, J. Atkins, and T. Cech, Editors. 1999, Cold Spring Harbor Laboratory Press: New York.
55. Joyce, G.F. and L.E. Orgel, *Prospects for understanding the origin of the RNA world*, in *The RNA World*, R. Gesteland, J. Atkins, and T. Cech, Editors. 1999, Cold Spring Harbor Laboratory Press: New York.
56. Kacian, D.L., et al., *A replicating RNA molecule suitable for a detailed analysis of extracellular evolution and replication*. Proc Natl Acad Sci U S A, 1972. **69**(10): p. 3038-42.
57. Kaufmann, S., *The Origins of Order*. 1993: Oxford University Press.
58. Knight, R.D., S.J. Freeland, and L.F. Landweber, *Selection, history and chemistry: the three faces of the genetic code*. Trends Biochem Sci, 1999. **24**(6): p. 241-7.
59. Knight, R.D. and L.F. Landweber, *Guilt by association: the arginine case revisited*. Rna, 2000. **6**(4): p. 499-510.
60. Koshland, D.E., Jr., *Special essay. The seven pillars of life*. Science, 2002. **295**(5563): p. 2215-6.
61. Kull, K., *A sign is not alive - a text is*. Sign Systems Studies, 2002. **30**(1): p. 327-36.
62. Lazcano, A., *Prebiotic evolution and the origin of life: Is a system-level understanding feasible?* in *Systems Biology*, I. Rigoutsos and G. Stephanopoulos, Editors. 2005, Oxford University Press.

63. Lehninger, A., D. Nelson, and M. Cox, *Principles of Biochemistry*. 3rd ed. 2000, New York: W. H. Freeman & co.
64. Luisi, P.L., *About various definitions of life*. *Orig Life Evol Biosph*, 1998. **28**(4-6): p. 613-22.
65. Maturana, H. and F. Varela, *De máquinas y seres vivos: Autopoiesis, la organización de lo vivo*. 6th ed. 1994, Buenos Aires: Grupo Editorial Lumen.
66. Mayr, E., *The idea of teleology*. *Journal of the History of Ideas*, 1992(53): p. 117-35.
67. Mayr, E., *What makes biology unique? Considerations on the autonomy of a scientific discipline*. 2004, Cambridge: Cambridge University Press.
68. Miller, S.L., *Which organic compounds could have occurred on the prebiotic earth?* *Cold Spring Harb Symp Quant Biol*, 1987. **52**: p. 17-27.
69. Monod, J., *Le Hasard et la Nécessité*. 1970, Paris: Éditions du Seuil.
70. Morowitz, H.J., et al., *The origin of intermediary metabolism*. *Proc Natl Acad Sci U S A*, 2000. **97**(14): p. 7704-8.
71. Nagaswamy, U. and G.E. Fox, *RNA ligation and the origin of tRNA*. *Orig Life Evol Biosph*, 2003. **33**(2): p. 199-209.
72. Nicolis, G. and I. Prigogine, *Exploring Complexity: An introduction*. 1989, New York: W. H. Freeman and Company.
73. Oehlenschläger, F. and M. Eigen, *30 years later--a new approach to Sol Spiegelman's and Leslie Orgel's in vitro evolutionary studies. Dedicated to Leslie Orgel on the occasion of his 70th birthday*. *Orig Life Evol Biosph*, 1997. **27**(5-6): p. 437-57.
74. Orgel, L.E., *Unnatural selection in chemical systems*. *Acc Chem Res*, 1995. **28**: p. 109-18.
75. Orgel, L.E., *Self-organizing biochemical cycles*. *Proc Natl Acad Sci U S A*, 2000. **97**(23): p. 12503-7.
76. Orgel, L.E., *Prebiotic chemistry and the origin of the RNA world*. *Crit Rev Biochem Mol Biol*, 2004. **39**(2): p. 99-123.
77. Osawa, S., et al., *Recent evidence for evolution of the genetic code*. *Microbiol Rev*, 1992. **56**(1): p. 229-64.
78. Pattee, H.H., *The physics of symbols: bridging the epistemic cut*. *Biosystems*, 2001. **60**(1-3): p. 5-21.
79. Peirce, C.S., *Peirce on Signs: Writings on Semiotic By Charles Sanders Peirce*. 1991: The University of North Carolina Press.

80. Pereto, J., *Controversies on the origin of life*. Int Microbiol, 2005. **8**(1): p. 23-31.
81. Popa, R., *Between Necessity and Probability: Searching for the Definition and Origin of Life*. 2004, Los Angeles: Springer Verlag.
82. Popper, K., *Escritos Selectos*, ed. D. Miller. 1995, México: Fondo de Cultura Económica.
83. Pross, A., *The driving force for life's emergence: kinetic and thermodynamic considerations*. J Theor Biol, 2003. **220**(3): p. 393-406.
84. Queiroz, J. and C.N. El-Hani, *Towards a multi-level approach to the emergence of meaning processes in living systems*. Acta Biotheor, 2006. **54**(3): p. 179-206.
85. Ribas de Pouplana, L., et al., *Genetic code origins: tRNAs older than their synthetases?* Proc Natl Acad Sci U S A, 1998. **95**(19): p. 11295-300.
86. Ricard, J., *What do we mean by biological complexity?* C R Biol, 2003. **326**(2): p. 133-40.
87. Rizotti, M., *Defining life: the central problem in theoretical biology*. 1996, Padova: University of Padova.
88. Rodnina, M.V. and W. Wintermeyer, *Peptide bond formation on the ribosome: structure and mechanism*. Curr Opin Struct Biol, 2003. **13**(3): p. 334-40.
89. Rosen, R., *Life Itself: A comprehensive inquiry into the nature, origin and fabrication of Life*. 1991, New York: Columbia University Press.
90. Ruiz-Mirazo, K., J. Pereto, and A. Moreno, *A universal definition of life: autonomy and open-ended evolution*. Orig Life Evol Biosph, 2004. **34**(3): p. 323-46.
91. Saussure, F., *Curso de Lingüística General*. 1993, México: Editorial Planeta.
92. Schneider, E. and J. Kay, *Life as a manifestation of the second law of thermodynamics*. Mathematical and Computer Modelling, 1994. **19**(6-8): p. 25-48.
93. Schneider, E. and J. Kay, *Order from disorder: the thermodynamics of complexity in biology*, in *What is life? The next fifty years*, M. Murphy and L. O'Neill, Editors. 1995, Cambridge University Press: Cambridge.
94. Schrödinger, E., *Mi Concepción del Mundo*. 1998, Barcelona: Tusquets Editores.
95. Schrödinger, E., *¿Qué es la vida?* 2006, Barcelona: Tusquets Editores.

96. Shannon, C.E., *A Mathematical Theory of Communication*. The Bell System Technical Journal, 1948. **27**: p. 379-423, 623-656.
97. Shapiro, R., *Prebiotic ribose synthesis: a critical analysis*. Orig Life Evol Biosph, 1988. **18**: p. 71-85.
98. Shapiro, R., *A replicator was not involved in the origin of life*. IUBMB Life, 2000. **49**(3): p. 173-6.
99. Shapiro, R., *A simpler origin for life*. Sci Am, 2007. **296**(6): p. 46-53.
100. Sheldrake, R., *The presence of the past: Morphic Resonance*. 1995: Park Street Press.
101. Smit, S., M. Yarus, and R. Knight, *Natural selection is not required to explain universal compositional patterns in rRNA secondary structure categories*. Rna, 2006. **12**(1): p. 1-14.
102. Smith, E. and H.J. Morowitz, *Universality in intermediary metabolism*. Proc Natl Acad Sci U S A, 2004. **101**(36): p. 13168-73.
103. Spiegelman, S., *An in vitro analysis of a replicating molecule*. Am Sci, 1967. **55**(3): p. 221-64.
104. Sprinzl, M. and K.S. Vassilenko, *Compilation of tRNA sequences and sequences of tRNA genes*. Nucleic Acids Res, 2005. **33**(Database issue): p. D139-40.
105. Taylor, F.J. and D. Coates, *The code within the codons*. Biosystems, 1989. **22**(3): p. 177-87.
106. Trevors, J.T. and D.L. Abel, *Chance and necessity do not explain the origin of life*. Cell Biol Int, 2004. **28**(11): p. 729-39.
107. Trifonov, E.N., *Consensus temporal order of amino acids and evolution of the triplet code*. Gene, 2000. **261**(1): p. 139-51.
108. Umerez, J., *Howard Pattee's theoretical biology--a radical epistemological stance to approach life, evolution and complexity*. Biosystems, 2001. **60**(1-3): p. 159-77.
109. von Kiedrowski, G., *A self-replicating hexadeoxynucleotide*. 1986. **25**: p. 932-35.
110. von Stockar, U. and J.-S. Liu, *Does microbial life always feeds on negative entropy? Thermodynamic analysis of microbial growth*. Biochimica et Biophysica Acta, 1999. **1412**: p. 191-211.
111. Wachtershauser, G., *Before enzymes and templates: theory of surface metabolism*. Microbiol Rev, 1988. **52**(4): p. 452-84.

112. Wachtershauser, G., *Evolution of the first metabolic cycles*. Proc Natl Acad Sci U S A, 1990. **87**(1): p. 200-4.
113. Wachtershauser, G., *Life in a ligand sphere*. Proc Natl Acad Sci U S A, 1994. **91**(10): p. 4283-7.
114. Wachtershauser, G., *The origin of life and its methodological challenge*. J Theor Biol, 1997. **187**(4): p. 483-94.
115. Weismann, A., *The Germ-Plasm: A Theory of Heredity*. 1893, New York: Charles Schreiber's Sons.
116. Williams, T.J., et al., *Evidence for autotrophy via the reverse tricarboxylic acid cycle in the marine magnetotactic coccus strain MC-1*. Appl Environ Microbiol, 2006. **72**(2): p. 1322-9.
117. Wittgenstein, L., *Investigaciones Filosóficas*. 3ra ed. 1958, México: Instituto de Investigaciones Filosóficas UNAM / Crítica.
118. Woese, C.R., *The genetic code: the molecular basis for genetic expression*. 1967, New York: Harper & Row.
119. Wong, J.T., *A co-evolution theory of the genetic code*. Proc Natl Acad Sci U S A, 1975. **72**(5): p. 1909-12.
120. Wu, H.L., S. Bagby, and J.M. van den Elsen, *Evolution of the genetic triplet code via two types of doublet codons*. J Mol Evol, 2005. **61**(1): p. 54-64.
121. Yarus, M., *RNA-ligand chemistry: a testable source for the genetic code*. Rna, 2000. **6**(4): p. 475-84.
122. Yarus, M., J.G. Caporaso, and R. Knight, *Origins of the genetic code: the escaped triplet theory*. Annu Rev Biochem, 2005. **74**: p. 179-98.