



BIOINFORMÁTICA: TECNOLOGÍAS DE LA INFORMACIÓN AL SERVICIO DE LA BIOLOGÍA Y OTRAS CIENCIAS

Bioinformatics: Information Technology in the Service of Biology and other Sciences

Miguel Angel Alcalde-Alvites¹

Universidad Nacional Mayor de San Marcos, Perú

Recibido: 4-10-2014

Aceptado: 19-12-2014

RESUMEN

El presente artículo presenta una revisión de la bioinformática, desde sus inicios y su evolución en el tiempo, las definiciones por diversos autores, las aplicaciones que tiene, los software de acceso libre para el almacenamiento y procesamiento de datos y de qué manera estos aportan a la bioinformática y a las ciencias que están inmersas en las mismas. Concluyendo que el uso de las tecnologías es una alternativa de solución para esta inmensidad de datos para campos de la biología actuales y revolucionarias como la genómica, la proteómica, la transcriptómica y la metabólica.

Palabras Clave: *Bioinformática, Tecnologías, Biología, ADN, Ciencias computacionales.*

ABSTRACT

This article presents an overview of bioinformatics, since its beginnings and how it has been evolved over time, the definitions from different authors, the application it has and the free access software storage and processing of data and how these contribute to bioinformatics and sciences that are part of them. Concluding that the use of this technology is an alternative solution for this vast area of data for current and revolutionary biology like a genomics, proteomics, transcriptomics and metabolomics.

Keywords: *Bioinformatics, Technology, Biology, DNA, Computational sciences.*

INTRODUCCIÓN

El avance y auge en los progresos científicos en biología y otras ciencias han generado grandes cantidades de datos genéticos, más aun con el descubrimiento del genoma humano, así como la información provenientes de genomas completos de diferentes organismos vivos se han desarrollado nuevas herramientas informáticas o computacionales para albergar dicha información, no solo como un archivo inmenso, sino también que pueda ser utilizada a nivel mundial por especialistas que lo requieran con fines diagnósticos y terapéuticos o simplemente para contrastar sus hallazgos o almacenar nueva información; Castro, Gutiérrez y Cristancho, (2010) refiere que en las últimas dos décadas el almacenamiento de información biológica en bases de datos públicos se ha hecho común y crecido de una manera exponencial, indicando que para agosto del 2009 la base de datos GenBank contenía 108'431,692 y que este

¹Facultad de Ciencias Biológicas, Escuela Académica Profesional de Genética y Biotecnología. miguelalcalde.unmsm.edu.pe@gmail.com



crecimiento continua, doblando su tamaño aproximadamente cada 35 meses, en especial los genomas completos representan un área con una alta tasa de crecimiento, esto está dando una enorme cantidad de información haciendo que las computadoras sean indispensables para analizar datos biológicos, haciendo surgir una nueva ciencia denominada como la Bioinformática.

MÉTODO

El presente artículo tiene un fundamento teórico y conceptual basado en la lectura, análisis, revisión e interpretación de libros, tesis, artículos científicos de bioinformática, para lo cual se ha realizado una búsqueda en base de datos de SciElo, Redalyc, Dialnet y DOAJ, así como en Google book y Google Scholar, empleándose descriptores: Bioinformática, base de datos, uso de tecnologías para biología, historia de la bioinformática. Para el análisis y revisión de la literatura se ha considerado el periodo desde el año 2004 al 2014, tiempo en el que tiene una mayor incidencia la aplicación de la bioinformática en biología y otras ciencias. Después de realizar este procedimiento se organizó el artículo estructurándolo desde su historia, definiciones, aplicaciones y la descripción de los software de libre acceso.

1. Historia de la Bioinformática

Castro et al (2010) menciona que desde la década de los cincuenta donde se inicia los estudios de la estructura del ADN con Watson y Crick empieza la investigación en Bioinformática, en 1988 se crea el Centro Nacional de Informática y Biotecnología, así como el proyecto de secuenciación del genoma humano, el cual acelera el desarrollo de varias técnicas y tecnologías, dando paso a otras ramas como las proteómica, las transcriptómica y la metabolómica, produciendo una gran cantidad de información, lo cual da impulso al campo de la bioinformática. Cañedo y Arencibia (2004) refieren que el nacimiento de la bioinformática se debe al desarrollo de la genética sobre todo en la aplicación de las leyes de Mendel y el descubrimiento de la estructura del ADN, ya que se descubrieron un volumen enorme de información que era necesario guardar y analizar. Al proporcionarles la

tecnología las herramientas prácticas para que se pueda explorar las proteínas y ADN, considerando a las proteínas como moléculas portadoras de información basada en tres aspectos; primero, el código genético demuestra como una sucesión de nucleótidos puede transformarse en una sucesión de aminoácidos, segundo, la información molecular en una sucesión de aminoácidos determina la estructura espacial tridimensional de las proteínas y tercero la información en el ADN que también determina la función de las proteínas y observar el encadenamiento de residuos más pequeños llamados aminoácidos y nucleótidos, todo lo cual llevó a unir la teoría de la información, teorías matemáticas y genéticas, permitiéndole a Claude Shannon padre de la teoría de información a desarrollar una metáfora algebraica para modelar el código genético. A diferencia de los autores anteriores Lahoz-Beltrá, (2004) refiere que el nacimiento de la bioinformática data desde los estudios realizados por Margaret Oakley Dayhoff, al utilizar los ordenadores en la medicina y biología para crear bases de datos de almacenamiento para la información de las secuencias del ADN y las proteínas, a través de tarjetas perforadas realizando cálculos en un problema biológico, siendo esta base de datos utilizadas posteriormente en estudios de evolución molecular, lo cual permitió que a mediados de los años 80 se conozca como bioinformática, desde ese momento la utilización de los ordenadores en Biología estaría ligada al desarrollo y evolución de la informática, permitiendo la clasificación de especies con la introducción de los métodos de Taxonomía numérica, el estudio y modelización de procesos genéticos con apoyo de las tecnologías a través de los Algoritmos Genéticos. Medina et al (2012) la bioinformática se inicia con la primera secuencia de proteína en 1958, seguida de la primera secuencia de ADN en 1975, con el desarrollo de la técnica molecular del PCR (reacción en cadena de la polimerasa) en 1986 y con el inicio de la era Genómica. Sea como haya surgido esta disciplina lo cierto es que está permitiendo conocer nuevos campos para el bienestar del ser humano y de otras especies. Bustos, Moreno y Duque, (2011) refiere que en 1866 Mendel descubre los genes, en 1871 los ácidos nucleicos, en 1953 se descubre la estructura del ADN, entre 1975 y 1979 se aísla el primer

gen humano y se consigue la secuenciación completa del genoma humano, lo cual provoca una explosión de información, en vista de esta necesidad los especialistas en tecnologías, se sirvan de estas para crear la bioinformática.

Barraza, (s/f) refiere que la bioinformática ha tenido diversas fases, iniciándose con la genética con Mendel y sus leyes en 1865 a 1930, con la búsqueda del gen y la sustancia hereditaria entre 1865 y 1952, una segunda fase con la biología molecular, con el descubrimiento de la doble hélice y el código genético en 1953 a 1966, con la aparición de la ingeniería genética y la reacción en cadena de la polimerasa en 1967 a 1989, su tercera fase con la genómica con el proyecto genoma y la era post-genómica en 1990 al 2001 y en la actualidad con las nuevas tendencias en proteómica, metabolómica, biología sistémica, unido esto a los primeros modelos computacionales con Alan Turing en 1930 a 1950, luego con la evolución de las arquitecturas de computadora realizadas por Von Neumann en 1950 a 1970, con la formalización de las bases de datos en 1970 a 1980, con la era de la PC en 1980 a 1990, con el internet y la globalización en 1990, en la actualidad con Grid Computing, computación ubicua y nano computación, las cuales han permitido esta nueva ciencia.

2. Bioinformática

El definir a la bioinformática desde las perspectivas de diversos autores, es un indicativo de la importancia que tiene. Castro et al (2010) la define como la aplicación de las técnicas computacionales para entender, organizar y analizar la información asociada a las macromoléculas. Ramírez et al (2014) indican que es la ciencia que examina la estructura y función de genes y proteínas a través del uso de análisis computacionales, estadística y patrones de reconocimiento.

El Diario Gestión (marzo, 17 del 2014) enfatiza que la bioinformática es el uso de la tecnología de la información, estadística y algoritmos para almacenar, recuperar y analizar datos biológicos generados durante la investigación en ciencias de la vida. Cañedo y Arencibia (2004) menciona que es una disciplina emergente que utiliza las tecno-

logías de la información, y se encuentra entre las ciencias de la vida y de la información proporcionando herramientas y recursos para favorecer la investigación biomédica, desarrollando sistemas que sirvan para comprender el flujo de información desde los genes a las estructuras moleculares, su función bioquímica, conducta biológica para conocer su influencia en las enfermedades y la salud con la finalidad de utilizar esta información para desarrollar nuevas formas de tratar, curar o prevenir la diversidad de enfermedades que afectan a la humanidad. Asimismo esta disciplina está sirviendo para unir diversas áreas de la ciencia como la enzimología, la genética, la biología estructural, la medicina, la morfología y la ecología entre otros campos. Para Medina et al (2012) la bioinformática es el uso de técnicas computacionales, matemáticas y estadísticas para el análisis, interpretación y generación de datos biológicos, a través del estudio de datos del ADN (Ácido Desoxirribonucleico), proteínas, genómicas, mutación/polimorfismo, almacenando los datos de estos genes y observando que hacen los mismos. Brown, 2000 citado por Medina et al 2012 la definió como el uso de computadoras para la adquisición, manejo y análisis de la información biológica. Coltell, (2004) indica que es una disciplina científica y tecnológica en la que interaccionan en armonía la biología genética y molecular con los enfoques metodológicos y tecnológicos de la ciencia de la computación y la ingeniería informática, para la obtención y gestión del conocimiento biológico genómico y proteómico. Bustos et al (2011) define a la bioinformática como el uso de herramientas computacionales que permiten analizar, depurar y agilizar el manejo de grandes cantidades de datos de la biología en términos fisicoquímicos y permitir comprender y organizar la información asociada. Asimismo cita a la National Institute of Health de Estados Unidos (NIH), quien la define como la investigación, desarrollo o aplicación de herramientas computacionales y propuestas científicas para extender y facilitar el uso de datos biológicos, médicos o sanitarios, incluyendo la adquisición, almacenamiento, organización, análisis y visualización de los mismos.

Trelles, (s/f c) por su parte hace referencia que la bioinformática es el campo científico interesa-

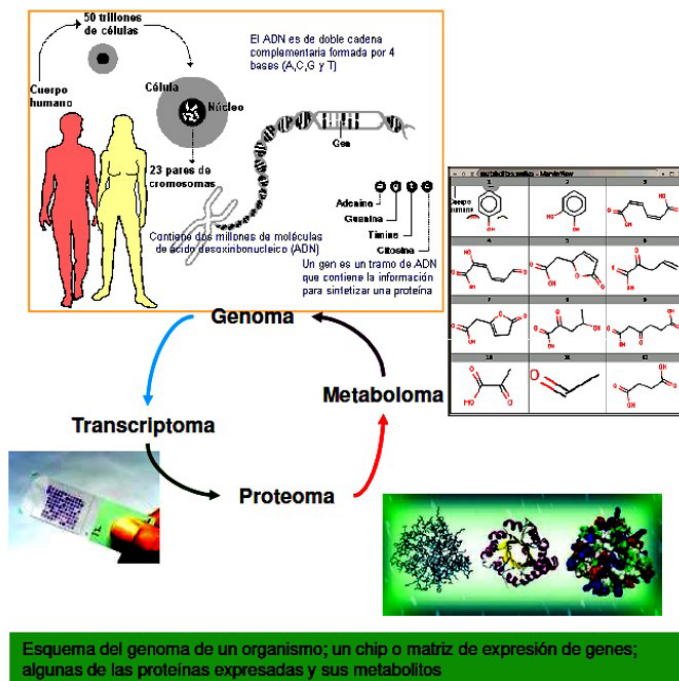
do en el manejo computacional de todo tipo de información biológica tanto si se trata de genes y sus productos como de organismos completos o incluso sistemas ecológicos, a su vez es el encargado de la gestión, tratamiento y análisis de información biológica mediante el desarrollo y uso de procedimientos y sistemas basados en las tecnologías.

3. Aplicaciones de Bioinformática

Rodríguez et al 2006 citado por Bustos et al 2011, indica que la aplicación de la bioinformática se da desde tres niveles, el primero a partir de la célula, el segundo es el análisis de datos de expresión genética ARN, que son transcritos por un conjunto de genes y como tercer nivel es el árbol de la vida. Salcedo (2014) menciona que está siendo aplicada en problemas que implican evaluar y entender la dispersión y la variación de marcadores genéticos, modelaje molecular, genómica, proteómica y minería de datos biológicos, también se puede aplicar en medicina forense, antropología, manejo y control de plagas, conservación y desarrollo de vacunas y drogas, mejoramiento genético de animales y plantas, entre otros. Bustos et al (2011) Refiere que actualmente se está desarrollando aplicaciones en el campo de descubrimiento de nuevas drogas, portales de información genética, medicina forense, mejoramiento agrícola y ecológico. Barraza, (s/f) por su parte menciona que la bioinformática se puede aplicar en el campo de la medicina, para el descubrimiento de drogas, para la medicina personalizada, para terapias genéticas y para prevención, en la agricultura para el mejoramiento de cultivos, tratamiento de plagas y enfermedades, en la ecología y medio ambiente, en energía, antropología e investigación forense. Trelles, (s/f a) menciona que los campos de la bioinformática comprende el análisis al genoma que es el material genético almacenado en las células de un organismo particular, a la genómica considerada como el estudio integral del origen, evolución y funcionamiento del genoma, así como la estructura espacial de las moléculas, proteínas y ADN y las funciones asociadas a ella, el proteoma que es el conjunto de proteínas expresadas por el organismo en determinadas células, la transcriptoma, conjunto de

todas la moléculas de ARN, la cual refleja la actividad de los genes y las concentraciones de proteínas y el último campo en que se aplica la bioinformática es el de metaboloma que es la colección de todos los datos metabólicos de un organismo como producto final de la expresión génica, que incluye la huella genética y química que acompaña a un proceso celular específico.

Figura 1. Campos relacionados a la Bioinformatica. Fuente: Trelles, (s/f e)



4. Software de Bioinformática

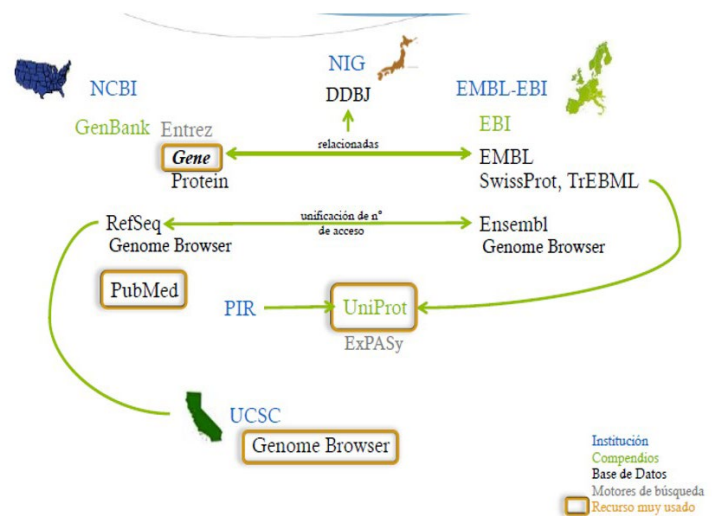
Trelles, (s/f e) refiere que una base de datos es un software (programa) para organizar datos, el cual incluye secuencias, estructuras, expresión génica, etc., además considera que existen tres tipos de bases de datos: primarias son las que contienen información original de los objetos biológicos, entre ellos están SwissProt, EMBL, GenBank entre otros; secundarias, son datos obtenidos a partir de una base de datos primarios, destacan Prosite, Pfam, scop, cath, etc. Y base datos compuestas, son las que integran una variedad de fuentes de datos primarios, sirviendo para evitar búsquedas múltiples en diferentes fuentes. Bustos et al (2011) indica que desde 1980 la bases de datos del Laboratorio de Biología Molecular Europeo (EMBL), el European Molecular Biology Laboratory (NCBI) National Center Biotechnology

Information de Estados Unidos y el Laboratorio Japonés DDBJ recopilan las secuencias nucleóticas de manera colaborativa, de manera que cada nueva entrada es automáticamente intercambiada con las otras dos restantes.

Santa María (s/f) menciona que la base de datos de GenBank almacena una variedad de tipos de secuencias de ADN, ARN, aminoácidos, secuencia de transcripto, gen, cromosoma, genoma, secuencia de mutación SNP comparte de manera constante sus datos, los cuales son de acceso libre en la página del NCBI: <http://www.ncbi.nlm.nih.gov/genbank/>, además presenta información de artículos, publicaciones científicas, distintos software relacionados con el análisis de secuencias de simple acceso; otra de estas base de datos importante es Entrez Gene que brinda a partir del nombre del gen información sobre la localización cromosómica, transcriptos asociados (nucleótidos), productos génicos (proteínas). Otra base de datos es el EBI que está compuesta por genoma, proteína y nucleótido, <http://www.ebi.ac.uk/>. UniProt es la base de datos que brinda información de proteínas, a través de tres bases de datos Swiss-Prot, Translated EMBL y Protein Sequence Database, disponible para todos <http://www.uniprot.org/>. El Expert Protein Analysis System (ExPASy) es un compendio de herramientas de análisis en proteómica, permite buscar/recuperar datos, en diversas áreas de las ciencias biológicas, incluida la proteómica, la genómica, la filogenia, la biología de sistemas, genética de poblaciones, transcriptómica, <http://www.expasy.org/>. Navegador de genoma, es una base de datos con una interfaz gráfica para representar secuencias y otros datos en función de su posición en los cromosomas, a través de sus tres navegadores principales el Genome Browser NCBI, Genome Browser de la universidad de california, santa Cruz UCSC, <http://genome.ucsc.edu/>.

Ramírez et al (2014) indican que para utilizar la base de datos de secuencias nucleótidos, es necesario tener en cuenta los conocimientos básicos y procedimientos de biología molecular para que se obtengan las secuencias de ADN. Trelles, (s/f d) Refiere que para reconstruir la secuencia completa de un ADN original se utilizan programas de ordenador, los cuales realizan los encajes de frag-

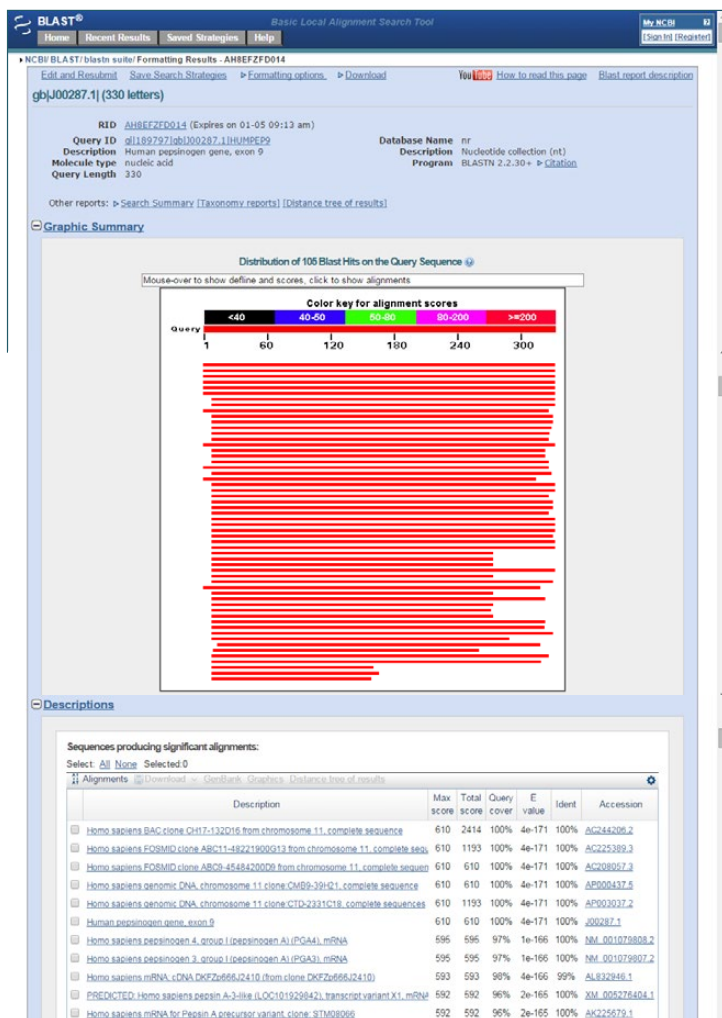
Figura 2. Secuencia de Base de datos. Fuente: Santa María (s/f)



mentos basado en el solapamiento de los mismos, para lo cual se puede utilizar Phred/phrad/consed que es un paquete de software que lee cromatogramas (trace files), asigna valores de calidad a las bases individuales de una secuencia, identifica y enmascara secuencias correspondientes a vectores (plásmido), ensamblan secuencias individuales en contigs, hacen sequence finishing, y está disponible en www.phrap.org

Ramírez et al (2014) hace alusión que además de las bases de datos, existe software que utiliza esta información para obtener datos importantes como la similitud de secuencias por medio de alineamiento en regiones de estas, gracias a programas como BLAST (basic local alignment search tool). La comparación se realiza por medio de secuencias nucleóticas o de aminoácidos, esta herramienta se encuentra en el GenBank <http://blast.ncbi.nlm.nih.gov/Blast.cgi> en la que muestra muchas variantes. pBLAST para el caso de proteínas (aminoácidos), nBLAST para el caso de nucleótidos y otros como el BlastX o tBlastx para nucleótidos traducidos. Esta comparación se realiza con una secuencia que se desea contrastar respecto a otras que se encuentran en la base de datos, donde aparece una pantalla con las secuencias más parecidas, el tamaño que presentan estas secuencias y a que organismo pertenecen.

Figura 3. Uso del software BLAST para la enzima del pepsinogeno en *Homo sapiens*.

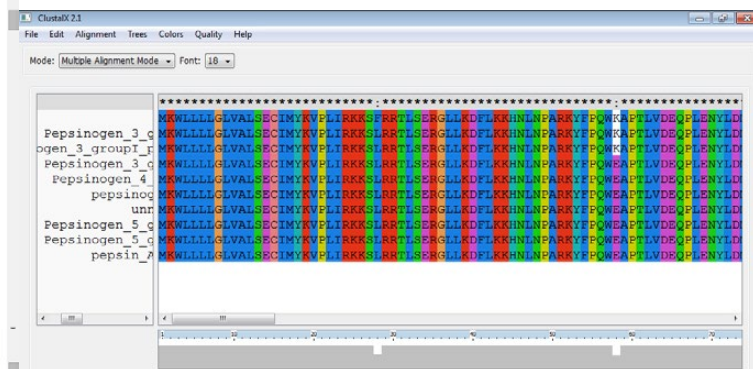


Además del BLAST, la página del NCBI brinda el programa para la obtención de ORFs <http://www.ncbi.nlm.nih.gov/gorf/gorf.html> (marcos de lectura abierta), ingresando alguna secuencia. Este software es muy necesario, permitiendo analizar en una secuencia la posible existencia de ADN codificante (proteínas) por medio de la observación de tripletes de inicio y tripletes de terminación en dicha secuencia, además permite el análisis BLAST para relacionar la posible proteína con alguna ya existente, que se encuentra en la base de datos del GenBank.

Otro programa especializado, que lo menciona Herrera (s.f), para el alineamiento de secuencias es el Clustal W/X <http://www.clustal.org/cluster2/> es un software indicado para el alineamiento múltiple que permite escoger las secuencias que se desea alinear, previamente descargadas de las

bases de datos en el formato FASTA, formato universal para muchos de estos programas. Luego de la alineación de las secuencias muestra por medio de distintos símbolos como asterisco, punto o dos puntos, sí en esa posición estas son idénticas, presentan sustituciones menos conservadas o si presentan sustituciones muy conservadas, respectivamente.

Figura 4. Uso del software Clustal para el alineamiento de secuencias de distintos organismos para la enzima del pepsinogeno.

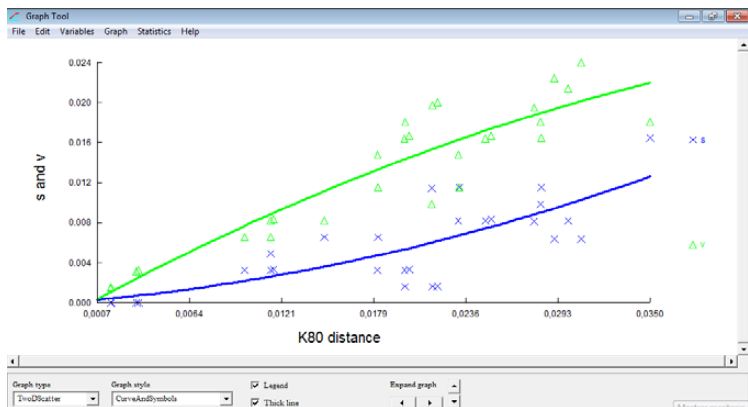


A partir de la base de datos de secuencias estas se pueden editar, por medio de cromatogramas (graficas obtenidas por medio de la cromatografía que demuestra el orden y la composición de los nucleótidos en una muestra extraída de algún organismo) que se observa con ayuda del programa chromas lite http://technelysium.com.au/?page_id=13 disponible de forma gratuita para su descarga. Para realizar trabajos en este software se requiere de conocimiento acerca de la obtención de primers, que permitirán al investigador obtener un contig o secuencias consenso de ADN por medio de otro programa como el CAP3 assembly <http://doua.prabi.fr/software/cap3> este programa accede la posibilidad de estudiar un segmento del genoma completo, especialmente cuando se buscan genes con ciertas características y cuando se desea establecer la secuencia de grandes fragmentos de un cromosoma.

Ramírez et al (2014) muestra también que la evaluación de secuencias de nucleótidos es importante para estudios de evolución que implican análisis filogenéticos, en este caso se puede realizar con el programa DAMBE <http://dambe.bio.uottawa.ca/dambe.asp>, que es de acceso libre para su descarga. Este software permite por medio de secuen-

cias el análisis de la frecuencias de los nucleótidos que se encuentran entre los organismos analizados, muestra graficas de transiciones (mutaciones puntuales de la secuencias del ADN de purinas a purina o de pirimidina a pirimidina) y transversiones (mutaciones puntuales de la secuencias del ADN de purina a pirimidina o viceversa) versus la divergencia genética que ha sufrido estos organismos (las secuencias de ese marcador) por medio de varios modelos de sustituciones nucleótídicas que se elija para luego realizar el análisis respectivo de este.

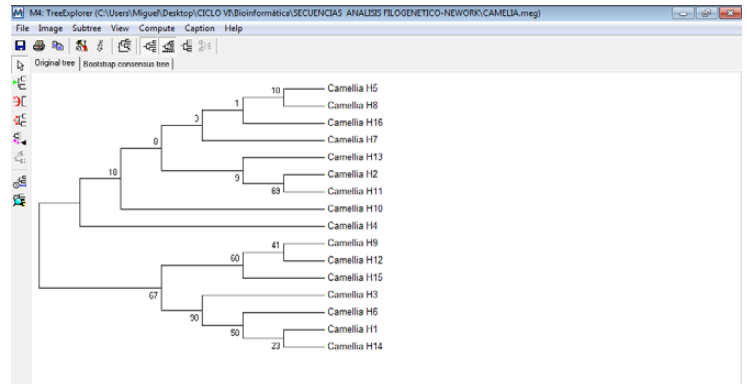
Figura 5. Uso del software DAMBE para el analisis de transiciones y transversiones vs diversidad genetica para la enzima glutamina sintasa en el género *Passiflora*.



Ramirez et al (2014) menciona que existen otros software como MEGA (molecular evolutionary genetic analysis) <http://www.megasoftware.net>, que permite realizar análisis filogenéticos, como la realización de árboles filogenéticos por medio de comparaciones con bases de datos, estimar tasas de evolución molecular y proveer hipótesis evolutivas, además de presentar un uso sencillo y la descarga es gratuita. El programa brinda información importante sobre las secuencias anotadas para el análisis; por ejemplo, este mostrará sitios conservados, sitios variables, sitios informativos o sinapomorfias, que son los principales caracteres para la elaboración de los árboles filogenéticos, y de los sitios autopomórficos. Además presenta opciones para hallar distancias genéticas entre las taxas (organismos ingresados en el software), la inferencia filogenética o la elaboración de los arboles filogenéticos a partir de las distancias genéticas con un análisis bootstrap, que es el soporte estadístico para conocer el nivel de confianza con

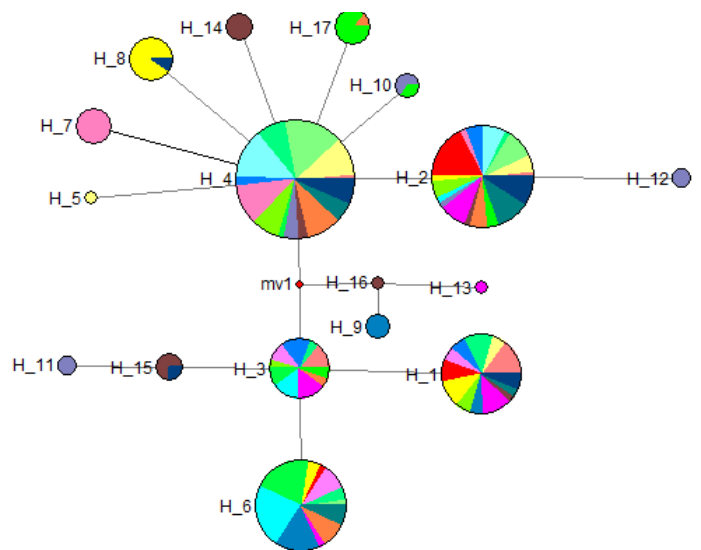
el que se está elaborando el árbol filogenético.

Figura 6. Uso del software MEGA para la elaboracion del arbol filogenetico del gen PAL para el genero *Camelia*.



El NETWORK <http://www.megasoftware.net> y DNAsp <http://www.ub.edu/dnasp> se utiliza para el análisis intraespecífico y poblacional, en análisis relacionados a la filogeografía y la genética de poblaciones, es de acceso libre, la descarga es gratuita. Estos programas permiten la elaboración de redes de haplotipos (combinación de alelos de diferentes loci de un cromosoma que son transmitidos juntos) por medio de un conjunto de secuencias que son puestas en tablas, que muestran los sitios informativos que se toman en cuenta para la elaboración del network.

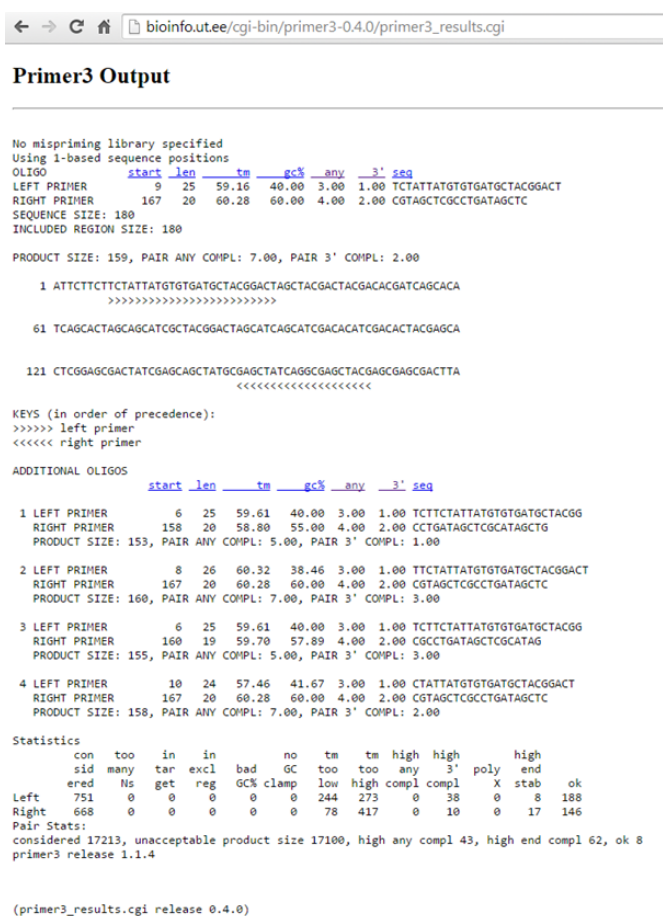
Figura 7. Uso del software NETWORK para elaborar la red de estos 17 haplotipos para el genero *Camelia*.



Según Salcedo (2014), con el avance de las ómicas: el inicio revolucionario de la genómica, y

otras que se añadieron por su importancia como, la transcriptómica que evalúa el nivel de expresión de los genes, se dio inicio a la creación de bases de datos y software que permiten la predicción de estructuras del RNA como es el caso del tRNA Database <http://gtrnadb.ucsc.edu/> . Esta base de datos incluye un software, basado en el apareamiento de las bases de la secuencia, permitiendo observar la formación de las estructuras secundarias de los tRNA de distintos organismos ya estudiados; no es el único método de evaluación para estos tipos de RNA, existe otro programa como el mfold <http://mfold.rna.albany.edu/?q=mfold>, que presenta un software para la predicción de estructuras del RNA basado en la energía libre que tiene cada estructura interna, formada por la unión de los enlaces de la misma secuencia, creando un nivel de estabilidad para su conformación conocido como entalpia.

Figura 8. Uso del software PRIMER 3

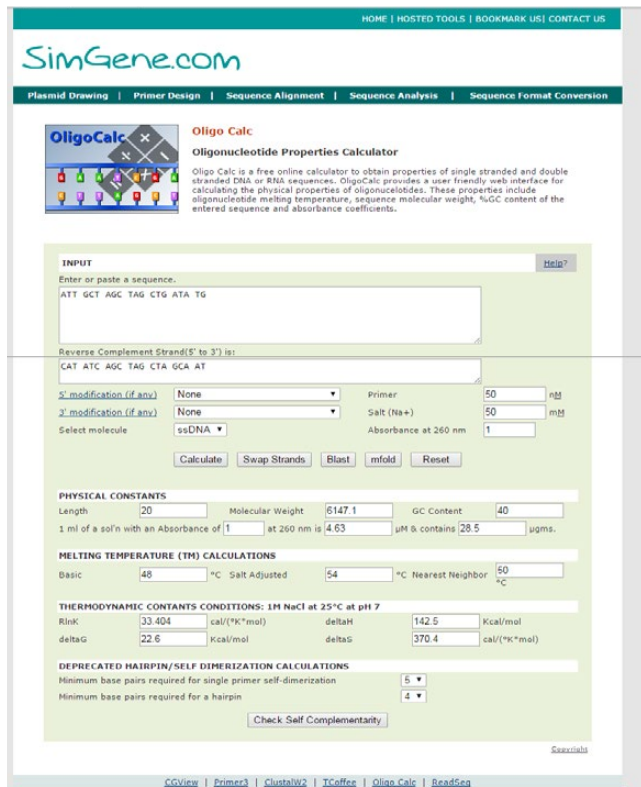


Ramírez et al (2014) pone énfasis al diseño de primers como un tema muy importante debido

a que estos son los iniciadores de la reacción en cadena de la polimerasa (PCR) que permite amplificar muchas veces una secuencia para su análisis. Los primers se unen al DNA molde de forma exacta y precisa de modo que amplifique lo que se requiere. Los primers presentan características importantes como un tamaño de 18 a 24 nucleótidos, presentan cierto porcentaje de GC (guanina-citosina) de aproximadamente 45 al 55%, una temperatura de alineamiento entre 45-55°C y que estos no deben formar heterodímeros ni estructuras secundarias que impidan la hibridación con el DNA.

Las características señaladas son evaluadas por distintos programas como en la página de oligocalc <http://www.basic.northwestern.edu/biotools/oligo.html> , la cual expresa de forma clara cada una de las características anteriores, para que el investigador analice si es adecuado usar el primer que se crea con los conocimientos previos. Mientras que la página de primer 3: <http://bioinfo.ut.ee/primer3-0.4.0/> , brinda de forma directa los posibles primers (forward y reverse) que se pueden utilizar, ingresando la secuencia que se desea amplificar.

Figura 9. Uso del software Oligocalc



DISCUSIÓN Y CONCLUSIONES

El avance del campo molecular de la biología, la extensa información que aumenta cada día de muchos organismos y el fenómeno de la globalización con un progreso de las tecnologías informáticas; han llevado a que se tenga que crear ciertos programas que contengan toda esta información, de modo que cualquier persona en el mundo lo pueda observar, extraer, comprender y analizar. Esto ha sido puesto en práctica gracias a la bioinformática.

Desde los inicios de la bioinformática, en donde lo principal era almacenar todo tipo de información como secuencias de una proteína, aminoácidos de las proteínas, etc.; es donde se dio inicio la creación de bases de datos, que con el tiempo se hicieron más especializadas y con una mejor clasificación, de modo que existe, base de datos solo para genes, solo para proteínas, solo para transcritos; aunque también existen las que presentan a todas estas juntas de modo que las hace más completas, además de contener también artículos de investigación en los que usan esta información o envían está a la base de datos para poder compartirla con otros investigadores. Un punto interesante de estas bases de datos es que no están exenta a correcciones de algún tipo sea por parte del autor o de otro investigador, que serán informadas y luego corregidas, si es necesario por los revisores de este programa.

Además de las bases de datos, que fue el punto de inicio de la bioinformática, se empezó a observar que toda esta información, sobre todo el de la secuencias de ADN, podía ser usada de otra forma, como la observación de estructuras que pueden formar, características de las proteínas, comparación de secuencias, relaciones evolutivas, predicciones para la formación de ciertas estructuras, etc.; todo esto con ayuda de software especiales como BLAST, Clustalw/x, MEGA, NETWORK, DAMBE, etc. La eficiencia, sencillez de su uso, rapidez, fácil acceso de estos programas, además de que muchos de ellos no presentan algún costo para su descarga o utilización, hacen que la bioinformática sea un campo que tenga una gran aceptación y produzca mayor beneficios que otras.

Finalizaremos mencionando que cada año se está

duplicando y hasta triplicando la información en las distintas bases de datos, siendo este un punto a favor de la investigación; además que se están creando otros programas con la finalidad de obtener ya genomas, proteomas, transcriptomas completos de distintos organismos, que están dando el auge de las ciencias ómicas. Todos estos avances en la bioinformática al final se observarán en la mejora de la terapia para muchas enfermedades, para las que tienen cura y las que no, en el mejoramiento genético de muchas especies de plantas y animales, las relaciones evolutivas de muchas taxas que aún no han sido clasificadas o las recién descubiertas y una nueva solución a problemas ecológicos.

REFERENCIAS BIBLIOGRÁFICAS

- Barraza, F. (s.f) Introducción a la Bioinformática, power point, http://cic.javerianacali.edu.co/wiki/lib/exe/fetch.php?media=materias:bioinfo_sesion1.pdf
- Barreto, E. (2008) Bioinformática: una oportunidad y un desafío. *Revista Colombiana Biotecnol.* 10 (1) 132-138
- Benitez-Paez, A. y Cardenas-Brito, S. (2010) Bioinformática en Colombia: presente y futuro de la investigación biocomputacional. *Biomédica*, 30, 170-177
- Bustos, L., Moreno, R. y Duque, N. (2011) Modelo de una bodega de datos para el soporte a la investigación bioinformática. *Scientia et Technica*, 16 (49) 145-152
- Castro, L., Gutiérrez, R. y Cristancho, M. (2010) Necesidad de un Centro Nacional de Bioinformática y Biología Computacional para Colombia, *Revista Colomb Biotecnol*, 12 (1) 179-189
- Cañedo, R. y Arencibia, R. (2004) Bioinformática: En busca de los secretos moleculares de la vida. *Acimed*, 12 (6) recuperado de http://www.bvs.sld.cu/revistas/aci/vol12_6_04/aci02604.htm
- Coltell, O. (2004) Integración de la Bioinformática en la Investigación Genómica Cardiovascular: Aplicaciones en el Framingham Heart Study. Tesis doctoral, Universitat Jaume I, Castellon-España. Recuperado de <http://www.tdx.cat/bitstream/handle/10803/10476/coltell.pdf;jsessionid=D117E-2BE123DA1833F645C30D2EAAEA5.tdx2?sequence=1>
- Cuadros, M., Cano, C. y Blanco, A. (2010) Recurso Web para el aprendizaje de aplicaciones bioinformáticas: Analisis de microarrays. *Actas de las I Jornadas sobre Innovación docente y adaptación*, 53-56
- Chicano, E. (2010) BlasStorP Herramienta NCBI_Blast a nivel local, Proyecto fin de master Universidad Internacional de Andalucía. Recuperado de http://dspace.unia.es/bitstream/handle/10334/401/0102_Chicano.pdf;sequence=1
- De la Puente, R. (2012) Mapa Genético y análisis de se-

cuencias nucleotídicas de interés en Lens, Tesis Doctoral, Universidad de León

Diario Gestión (marzo 17 del 2014) Mercado Peruano de bioinformática crece a mayor velocidad en la región. Sección empresas, Lima Perú

Franco, M., Cediél, J. y Payán, C. (2008) Breve historia de la bioinformática. Revista Colombia Médica, 39 (1) 117-120

González, D. (2010) Optimización del método SCAR (Sequence Characterized Amplified Region) que favorece el aislamiento de loci polimórficos para estudios filogenéticos en taxa cercanamente relacionados.

Herrera, J. (s.f) El software libre en bioinformática. Recuperado de http://www.arareko.net/bioinformatics/free_software/index.pdf.es.pdf

Martínez, J. (2007) La bioinformática como herramienta para la investigación en salud humana, Salud Publica de México, edición especial, 49 64-66

Medina, J., Garzón, F., Tafurth, P. y Barbosa, J. (2012) Recopilación Bioinformática. Universidad Distrital Francisco José de Caldas

Meneses, C. Rozo y L., Franco, J. (2011) Tecnologías bioinformáticas para el análisis de secuencias de ADN. Scienti et Technica, 16 (49) 116-121

Lahoz-Beltrá, R. (2004) Bioinformática, Simulación, vida artificial e inteligencia artificial. Ediciones Díaz de Santos S.A. Madrid

Leyva, M. y Febles, J. (s.f) Métodos ágiles en el desarrollo de software para la bioinformática. Recuperado de <http://publicaciones.uci.cu/index.php/SC/article/viewFile/25/26>

Ouzounis, C. (2012) Rise and Demise of Bioinformatics? Promise and Progress, PLoS Computational Biology, 8 (4) 1-5 Recuperado de [http://www.ploscompbiol.org/article/](http://www.ploscompbiol.org/article/fetchObject.action?uri=info%3Adoi%2F10.1371%2Fjournal.pcbi.1002487&representation=PDF)

Ramírez, R., Ramírez, P., Romero, P., Congrains, C. y Ramírez, J. (2014) Guía Práctica de Bioinformática, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Biológicas, EAP Genética y Biotecnología.

Rodríguez, E. (2013) Introducción a la Bioinformática, CINVESTAV, power point 16 de mayo del 2013.

Salcedo, F. (2014) Ciencias Ómicas: Bioinformática, genómica, proteómica, Transcriptoma y Metagenómica, power point, recuperado de http://es.slideshare.net/ef_salcedo/clase-04-ciencias-micas-bioinformtica-genmica-protemica-trascriptmica-y-metagenmica

Santa María, R. (s/f) Blast Universidad de Salamanca-España. Recuperado de http://vis.usal.es/rodrigo/documentos/bioinfo/temas/4_BLAST.pdf

Trelles, O. (s/f a) Bioinformática Básica, Tema 3: Introducción a la Bioinformática, Proyecto Open Course Ware, power point recuperado de http://ocw.unia.es/ciencias-de-la-vida/bioinformatica-basica/materiales/bloques-2-y-3/03-IntroBioInfo-1-v2_El%20dominio%20de%20aplicacion%20y%20areas%20de%20interes.pdf

Trelles, O. (s/f b) Bioinformática y biomedicina, Nuevos retos de la supercomputación. UNIA, 40-43 Recuperado de <http://riuma.uma.es/xmlui/bitstream/handle/10630/5523/40-43.pdf?sequence=1>

Trelles, O. (s/f c) Presentación y programación del curso Bioinformática, Proyecto Open Course Ware-UNIA, recuperado de <http://ocw.unia.es/ciencias-de-la-vida/bioinformatica-basica/materiales/bloques-0-y-1/00-PresenyProgram.pdf>