

ENSAYO

## Bioinformática en Colombia: presente y futuro de la investigación biocomputacional

Alfonso Benítez-Páez, Sonia Cárdenas-Brito

Grupo de Análisis Bioinformático, GABi, Centro de Investigación y Desarrollo en Biotecnología, CIDBIO, Bogotá, D.C., Colombia

La bioinformática, como la conocemos hoy en día, es una de las ciencias con mayor proyección en la adquisición de conocimiento científico. Contrario a lo que se piensa, esta ciencia tuvo sus inicios en los años 50, pero no fue sino hasta hace un par de décadas cuando tuvo su verdadero auge con la creación de las primeras bases de datos y el desarrollo de algoritmos computacionales diseñados para el análisis de secuencia.

El desarrollo científico y tecnológico alcanzado a nivel mundial, constantemente nos lleva a evaluar las posibilidades de transferencia de esos logros a nuestra sociedad científica. En este punto, la bioinformática se ha sumado a la larga lista de retrasos científicos de los cuales padece constantemente nuestra sociedad.

No obstante, las nuevas proyecciones y políticas de investigación y desarrollo logradas recientemente en nuestro país, han abierto definitivamente el camino para la aplicación y desarrollo de la bioinformática, la cual durante años ha sido mantenida tímidamente como objeto de estudio por pocos grupos de investigación de nuestro país. Una vez fueron propuestos los actuales modelos de sostenimiento y desarrollo con base biotecnológica, se ha dado un salto generacional en la investigación computacional en Colombia y nuestros objetivos científicos, aunque a largo plazo, están trazados al mismo nivel de países como Chile, Argentina, Brasil y México, que son los referentes inmediatos para la región.

En este breve ensayo queremos resaltar la situación actual en la investigación bioinformática llevada a cabo en nuestro país, así como también, plantear las perspectivas que se esperan para esta ciencia de gran impacto a nivel mundial.

**Palabras clave:** biología computacional, biología molecular, biotecnología, genómica, Colombia.

### Bioinformatics in Colombia: state of the art and perspectives

Bioinformatics emerged about 50 years ago, but it was developed greatly during the early 1980's by robust databases such as GenBank, EMBL, and DNA Database of Japan (DDBJ). Bioinformatic routines were rapidly adapted once the main algorithms for sequence analysis became available worldwide. As in other science fields, bioinformatics had minimal impact in low-income countries of Latin America until the last decade.

We revised the bioinformatics state of art in Colombia and found a few bioinformatics groups carrying out basic computational biology research. Nowadays, bioinformatics in Colombia has a hopeful scenario thanks to recent science policies adopted by the Colombian Government. Such policies have been adopted in order to establish a new model of sustainable scientific research.

In this brief report we revise the bioinformatics state of the art in Colombia. Finally, we conclude with some considerations for the proposed science model and we describe different perspectives of interest for the Colombian scientific community.

**Key words:** Computational biology, molecular biology, biotechnology, genomics, Colombia.

### Inicios de la bioinformática en el mundo

Contrariamente a la concepción que tiene la mayoría de investigadores vinculados a los campos de acción de la biología molecular, la genética, la bioquímica

y demás ciencias afines, la bioinformática tuvo sus inicios hace, aproximadamente, 50 años.

Durante la última mitad de la década de los 50, se inició la carrera por la secuenciación de pequeñas moléculas biológicas. El constante impulso del hombre por saber de qué estamos hechos comenzó en 1956 con la secuenciación de la primera proteína. De esta forma, Margaret O. Dayhoff describió la secuencia de la insulina bovina, un pequeño péptido de 51 aminoácidos.

Correspondencia:

Alfonso Benítez-Páez, Centro de Investigación y Desarrollo en Biotecnología, CIDBIO, Bogotá, D.C.

Telefax: (+571) 470 4923

abenitez@cidbio.org

Recibido: 21/04/09; aceptado: 19/11/09

Otras secuenciaciones de péptidos de similares características fueron realizadas en los años posteriores, pero no fue sino hasta 1964 cuando se obtuvo la primera secuencia de un ácido nucleico. Dicha secuencia pertenecía a un ARNt que descodifica un codón de alanina en *Saccharomyces cerevisiae*.

Con la constante producción de información biológica, la cual crecía a un ritmo lento comparado con el actual volumen de generación de datos, se creó la necesidad de recopilar y organizar toda la información generada a partir de dichos proyectos de secuenciación. En 1965, la misma Margaret Dayhoff creó la primera base de datos de secuencias biológicas, en la cual almacenó y puso a disposición de la comunidad científica todas las secuencias de ADN y proteínas descritas hasta la fecha. Ocho años más tarde, en 1973, se anunció y creó la base de datos más antigua que se conoce y la cual sigue vigente, el *Protein Data Bank* (PDB) (1). Esta base de datos surgió inicialmente con el propósito de almacenar las estructuras proteicas determinadas por cristalografía de rayos X.

Hoy, 27 años después de su lanzamiento, además de servir como base de datos de estructuras de proteínas, también lo hace como reservorio de estructuras de toda clase de macromoléculas conocidas: ADN, ARN (ARNm, ARNr y ARNt) y grandes complejos proteicos asociados con todo tipo de biomoléculas. En la actualidad, se pueden encontrar dichas estructuras dilucidadas por otras metodologías diferentes a la cristalografía, siendo la resonancia magnética (RM) y la criomicroscopía electrónica alternativas para el estudio de estructuras macromoleculares.

Durante 1978, de nuevo Margaret Dayhoff fue la encargada de generar la primera matriz de sustitución de aminoácidos, denominada PAM (*Point Accepted Mutation*) (2). Tal avance en la interpretación de patrones de secuencia, obtenidos a partir de información biológica, abrió el camino a los estudios sobre evolución molecular que actualmente aportan una visión más aproximada a las verdaderas relaciones filogenéticas entre especies.

A pesar de la relevancia de los acontecimientos descritos hasta ahora, los mayores avances en la bioinformática acontecieron en años posteriores, con la generación de algoritmos matemáticos para el análisis de secuencias, así como también, la creación de las bases de datos y herramientas para el análisis de secuencias con un mayor impacto a nivel global.

## Era dorada de la bioinformática

Entrados los años 80, la bioinformática había cobrado ya un nuevo sentido en la investigación científica y, conscientes de ello, varios grupos de investigación de prestigio, como el *Theoretical Biology and Biophysics Group* adscrito al instituto norteamericano *The Alamos National Laboratory*, junto con *Stanford University*, dieron origen a la base de datos más conocida en el mundo, el *GenBank*. Dicho proyecto fue financiado por los *National Institutes of Health* (NIH) de los Estados Unidos y otras importantes instituciones gubernamentales, como el *United States Department of Energy* y el *United States Department of Defense*.

Casi paralelamente, en 1981, Temple Smith y Michael Waterman revisaron extensamente los algoritmos matemáticos propuestos por Needleman y Wunsch en 1970 para la comparación de secuencias biológicas. Como resultado de sus análisis, generaron el conocido algoritmo de alineamiento local que permitió optimizar la comparación de secuencias biológicas a lo largo de un alineamiento. Este método de análisis se convirtió en la contribución más importante para la comparación directa de secuencias y en la piedra angular del alineamiento por pares de secuencias (3).

Consolidados los grupos de investigación con base bioinformática alrededor del mundo, se inició, entonces, la era de la creación de las grandes bases de datos. Pocos años después de la creación del *GenBank*, se generaron sus versiones europea y asiática, conocidas como la base de datos EMBL (*European Molecular Biology Laboratory*) y DDBJ (*DNA Data Bank of Japan*) en 1981 y 1984, respectivamente. Casi inmediatamente después de la creación de los bancos de secuencias, se empezó a cuestionar el significado de la información almacenada en ellos. Como consecuencia, en 1985 se reportó el algoritmo FASTA (*FAST-All*) de comparación de secuencias, el cual directamente operaba como motor de búsqueda de secuencias similares dentro de la base de datos *GenBank* (4).

Durante los años 1987 a 1990, se dio impulso a las bases de datos para secuencias de proteínas que dio como resultado la creación de *SwissProt* y *PIR* (*Protein Information Resource*). En esa misma época, y más exactamente en 1990, se originó otro de los hitos más importantes de la bioinformática. La implementación del algoritmo BLAST (*Basic Local Alignment Tool*) revolucionó completamente la exploración y búsqueda de secuencias biológicas en bases de datos (5). El estudio publicado por

Stephen Altschul y colaboradores se considera como el método computacional de mayor uso y la referencia literaria de mayor citación en la historia de la ciencia.

De este modo, un nuevo panorama en la investigación de la biología computacional se abrió gracias a:

- la infraestructura computacional lograda para almacenar y administrar datos;
- los modelos matemáticos diseñados para analizar e interpretar patrones biológicos, y
- los avances tecnológicos que desde finales de los años 90 incrementaron sustancialmente el volumen de procesamiento de las muestras biológicas.

Con la creación del consorcio para la ejecución del proyecto de secuenciación del genoma humano (*Human Genome*, HUGO), en 1993, se inició la era genómica. Siendo la decodificación del genoma humano un proyecto tan ambicioso, otros consorcios se hicieron con el logro de secuenciar los primeros genomas no virales. En 1996, *S. cerevisiae* fue el primer genoma eucariota secuenciado. En 1995 y 1997, respectivamente, los genomas de *Haemophilus influenzae* (6) y *Escherichia coli* (7) fueron los primeros genomas bacterianos publicados. Entre los años 1998 y 2000, se entregaron los genomas secuenciados de *Caenorhabditis elegans*, *Pseudomonas aeruginosa*, *Drosophila melanogaster* y *Arabidopsis thaliana*. En el año 2003, se finalizó la secuencia definitiva del genoma humano y, aunque al día de hoy no se sabe con certeza cuántos genes poseemos, este hecho se logró gracias a la proyección y explotación del potencial de la industria con base biotecnológica.

Como consecuencia, todos los avances tecnológicos destinados al campo de la genómica lograron proyectar a la comunidad científica hacia la secuenciación del genoma de cualquier organismo de interés científico o económico. Hoy en día, existen más de 1.200 genomas de bacterias completamente secuenciados y más de 4.000 proyectos en curso (GOLD Database, <http://www.genomesonline.org>, actualizado a abril de 2010). La base de datos de genomas eucariotas más completa, ENSEMBL (<http://www.ensembl.org>), cuenta con más de 50 genomas completamente secuenciados y anotados (versión 58, mayo de 2010).

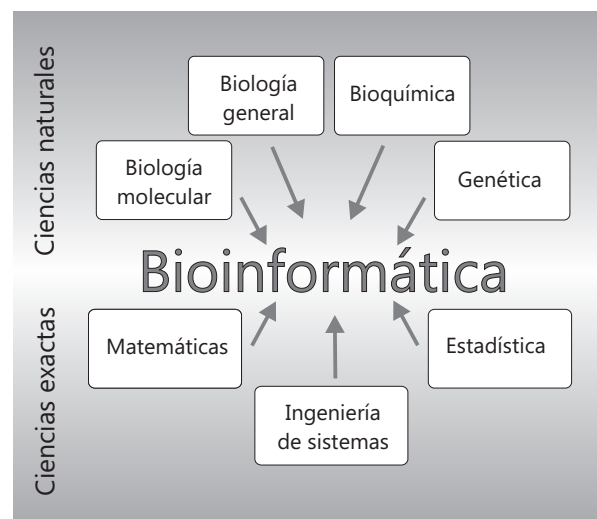
Esta gran cantidad de información generada a partir de los múltiples proyectos de secuenciación hace que las bases de datos crezcan a un ritmo acelerado y casi insostenible. Al mismo tiempo, el enorme volumen de datos conlleva a que el número de herramientas de análisis y algoritmos computacionales surjan al mismo ritmo para poder extraer un conocimiento tangible que permita dar un significado a la diversidad biológica de nuestro ecosistema.

### Bioinformática en Colombia

Lejos de los avances antes mencionados, y que han contribuido enormemente a la progresión de la bioinformática a nivel global, debemos posicionarnos objetivamente como una sociedad de muy baja producción de conocimiento bioinformático.

Dejando de lado las dificultades económicas que nos impiden tener un mayor progreso científico-tecnológico a cualquier nivel, el pobre desarrollo de la bioinformática en Colombia tiene factores adicionales de fondo. Dichos factores radican esencialmente en el déficit académico en cuanto a la enseñanza de la bioinformática. Teniendo en cuenta que la academia es el principal gestor de la investigación científica, tanto en el contexto de la educación pública como de la privada, la carencia de adecuados programas de formación produce un crecimiento nulo y, por ende, una pobre oferta de investigadores en este campo.

Aunque la bioinformática como tal se presenta como una ciencia multidisciplinaria (figura 1), en la cual



**Figura 1.** Componente disciplinario de la bioinformática. Muestra los principales campos de las ciencias naturales y exactas que convergen en la bioinformática.

es necesario poseer una adecuada formación en diversas áreas de las ciencias naturales y ciencias exactas, ya se han logrado establecer programas académicos competitivos en Europa y Estados Unidos, y algunos en Latinoamérica. La mayoría de los programas de formación de investigadores se basan en títulos propios de maestría y doctorado, los cuales van en aumento a medida que se revisan los programas académicos superiores.

En Colombia, en la actualidad no se ha desarrollado ningún programa sólido para la formación integral de bioinformáticos competentes. No obstante, algunos programas de maestría y doctorado de instituciones como la Universidad Nacional de Colombia y la Universidad de los Andes, han incorporado módulos semestrales de bioinformática que, a su vez, están lejos de tener el poder educativo requerido para generar verdaderos profesionales en este campo. A cambio, dichas cátedras sólo funcionan como herramienta difusora, más no orientadora, de la existencia de la bioinformática como tema de investigación.

Debido al componente multidisciplinario de la bioinformática, resultaría difícil imaginarse un profesional con vastos conocimientos en los campos de la biología molecular, la bioquímica, la genética, las matemáticas, la estadística y la ingeniería de sistemas; sin embargo, los conocimientos de cada una de estas áreas deben ser adaptados de una forma armónica para suplir las necesidades básicas, tanto del profesional como de la actualidad científica que lo circunscriba.

Otro punto que afecta a la poca progresión de la bioinformática en Colombia, se deriva directamente de lo arriba expuesto. **En nuestro país, la financiación de proyectos de investigación por parte de las principales entidades públicas y privadas destinadas al apoyo científico, es insuficiente.** Ese hecho sería consecuencia directa de dos eventualidades:

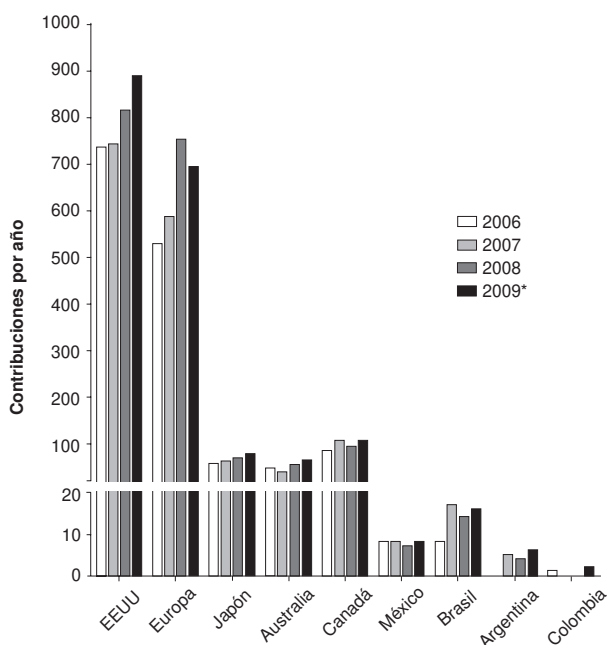
que, gracias a la falta de programas de formación académica en bioinformática, no exista promoción de proyectos de investigación dirigidos a la resolución de problemas propios y que involucren un componente bioinformático directo y bien planificado, y

que existan iniciativas de proyectos de investigación en el área, pero que no sean tomados como prioridad por falta de adecuados comités evaluadores que objetivamente valoren el impacto científico de dichas propuestas.

En cualquiera de los dos casos, la falta de profesionales con trayectoria y adecuado soporte científico, tanto para promover líneas de investigación como para la evaluación de proyectos en bioinformática, sería una consecuencia directa de la falta de formación de científicos especializados.

Dichas falencias pueden parecer generales para muchas de las áreas o ramas emergentes de la investigación científica en nuestro país, pero cabe destacar que su diagnóstico ya fue elaborado para la biotecnología y que fue recopilado en un completo estudio publicado por Colciencias, la Universidad Nacional de Colombia y CORPOGEN (8). Dicho estudio hace referencia al desarrollo de la biotecnología en Colombia y dentro de las líneas por considerar en el nuevo esquema del Plan Nacional de Biotecnología está el impulso de la investigación en bioinformática.

Al someter a una valoración objetiva nuestra contribución en la investigación bioinformática,



**Figura 2.** Contribuciones científicas en el área de la bioinformática a nivel mundial. Comparación de los artículos publicados por grupos de investigación pertenecientes a países industrializados, Brasil, México, Argentina y Colombia. La búsqueda de publicaciones científicas fue hecha en la base de datos *PubMed* a través del buscador *Entrez*. Los datos de Europa fueron generados por la contribución de los grupos de investigación de España, Italia, Alemania, Reino Unido, Holanda y Francia.

\* Los datos de 2009 son una proyección a partir de la tendencia observada entre los meses de enero y junio. Escala inferior, de 0 a 20 publicaciones por año; escala superior, de 21 a 1.000 publicaciones por año.



podemos ver que la producción científica en Colombia es casi nula en comparación con la contribución que tradicionalmente tienen los grupos de investigación de Estados Unidos, Europa o Canadá (figura 2). Paralelamente, también podemos diagnosticar que nuestra producción científica-bioinformática es deficiente, en número, comparada con nuestros referentes regionales inmediatos, Brasil, México y Argentina. Tales datos fueron obtenidos mediante una búsqueda de publicaciones en las más relevantes revistas científicas de este campo, como lo son: *Bioinformatics*, *PLoS Computacional Biology*, *BMC Bioinformatics*, *BMC Genomics* y *Journal of Computational Biology*. Dichas revistas gozan de los más altos índices de impacto en este campo de acuerdo con el *Journal Citation Reports 2007* (Thompson Reuters).

A pesar del panorama negativo que presenta el desarrollo de la bioinformática en nuestro país, cabe destacar la labor de algunos grupos de investigación que, a pesar de no contar con grandes presupuestos destinados a financiar sus investigaciones, han podido generar resultados que deben ser tomados como base o ejemplo de futuros modelos de investigación por desarrollarse en un mediano o corto plazo en Colombia.

### **Situación actual de la investigación en bioinformática en Colombia**

Haciendo una búsqueda por los grupos y centros de investigación que realizan estudios de componente bioinformático, podemos encontrar que son pocos los grupos dedicados a ello y aún menos los que han podido proyectar su trabajo a nivel internacional con publicaciones de mediano impacto (9-22).

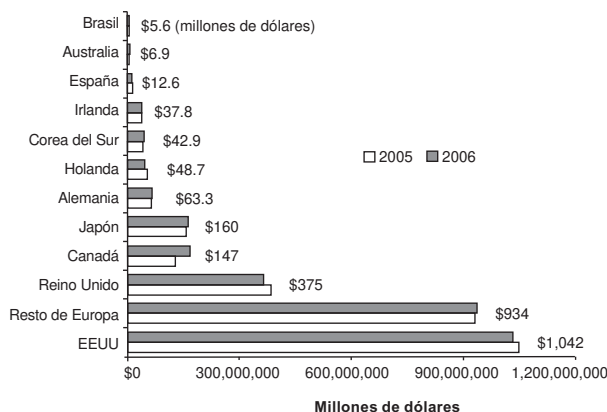
Entre estos últimos, podemos mencionar al Grupo de Parasitología Molecular (GEPAMOL) de la Universidad del Quindío, el Centro de Bioinformática del Instituto de Biotecnología de la Universidad Nacional de Colombia, el Grupo de Investigación en Bioquímica Computacional de la Pontificia Universidad Javeriana y el Grupo de Análisis Bioinformático (GABi) del Centro de Investigación y Desarrollo en Biotecnología. Esta muestra, aunque pequeña, no deja de ser vital y demuestra el potencial investigador en el campo de la bioinformática de los grupos colombianos. Ella debe ser, ante todo, un punto de partida concreto para el apoyo de la investigación en bioinformática que se genere a partir de las nuevas políticas para desarrollo biotecnológico en el país.

Además de la latente producción científica de nuestros grupos de investigación, cabe destacar la acción de impacto mundial que desarrollan desde hace varios años CENICAFE y su proyecto Genoma del Café (<http://bioinformatics.cenicafe.org>), el cual tiene como objeto principal un extenso análisis de genómica funcional y estructural del café colombiano.

Finalmente, debemos mencionar y destacar la misión de algunos grupos de investigación que constantemente promueven y difunden el conocimiento en bioinformática. Desde comienzos de la actual década, han organizado diversos seminarios, simposios y cursos de entrenamiento en herramientas bioinformáticas. Dichos entrenamientos están dirigidos a complementar la formación científica de los investigadores colombianos, así como a proyectar sus estudios a un nuevo campo de actuación donde se contemple el análisis computacional como piedra angular de las investigaciones contemporáneas. Algunos de los grupos promotores de entrenamientos son el Grupo BIMAC de la Universidad del Cauca, el Centro de Bioinformática del Instituto de Biotecnología de la Universidad Nacional de Colombia y el Grupo de Análisis Bioinformático (GABi). Estos dos últimos, miembros partícipes activos de las reuniones programadas en el marco de la Red Iberoamericana de Bioinformática (recientemente reagrupada como Sociedad Iberoamericana de Bioinformática, SolBio) (RIB, <http://chirimoyo.ac.uma.es/rib>), red que en el 2005 tuvo como lugar de encuentro la ciudad de Cartagena de Indias y que reúne anualmente a grupos de investigación bioinformática de habla hispana.

En una perspectiva general podemos concluir que, aunque pobre en infraestructura, mas no en calidad, existe una verdadera actividad de investigación en el campo de la bioinformática en Colombia. Lamentablemente, dada la actual coyuntura económica del país, esta rama de la ciencia corre serio peligro de ser uno más de los muchos campos de actuación científica que queda a la deriva por falta de apoyo financiero en nuestro país.

Sin embargo, debemos esperar que la nueva ley de ciencia, tecnología e innovación, aprobada por el Congreso de la República de Colombia a finales de 2008, mejore el desarrollo científico y tecnológico del país, y proporcione un mayor apoyo a la biotecnología, y que con esta última pueda destinar recursos que promuevan definitivamente la investigación bioinformática en Colombia.



**Figura 3.** Presupuesto destinado a la financiación de investigaciones en el campo de la genómica. Cifras obtenidas del estudio de Pohlhaus y Cook-Deegan 2006 (23). Los países están ordenados de acuerdo con la financiación concedida en el año 2006. Las cifras presentadas son el promedio de los años 2005 y 2006. Los datos para "Resto de Europa" comprenden los recursos destinados para investigación en Austria, Dinamarca, Finlandia, Islandia, Estonia, Francia, Hungría, Italia, Polonia y Suecia.

A este respecto, es interesante observar la destinación de recursos económicos a la investigación en bioinformática a nivel global. Una de las ramas más influyentes de la bioinformática es la genómica. A partir de los datos publicados por Pohlhaus y Cook-Deegan (23), se pudo hacer una comparación directa y, al mismo tiempo, una estimación del impacto que tiene la genómica en los países industrializados (figura 3). El resultado de esta comparación se correlaciona directamente con el patrón observado en la figura 2, con lo cual se puede discernir que, a mayor inversión en bioinformática, mayor será la contribución científica en el campo. Entre nuestros referentes inmediatos, sólo existen datos de la inversión de Brasil en genómica, la cual se presenta con un promedio anual de US\$ 5,6 millones entre 1997 y 2007 (23).

Dado el bajo impacto que hasta ahora ha tenido la bioinformática en Colombia y, más aún, la nula dedicación a la especializada rama de la genómica, no existen datos comparativos al respecto para esos años. No obstante, mediante la convocatoria 392 de Colciencias (marzo de 2007) se inició el proceso de financiación de la investigación en bioinformática y de la genómica. A partir de esa convocatoria, se destinó una partida de US\$ 1,2 millones para la ejecución del proyecto GeBIX (**Genómica y Bioinformática de Ambientes Extremos**) durante dos años, con posibilidad de extender la financiación a otros tres años por el mismo monto global.

De esta manera, Colombia se sumó a la larga lista de países interesados en la investigación genómica para el aprovechamiento de recursos bióticos propios. Y, aunque lejos de la inversión presupuestada para nuestro referente más cercano, Brasil (con 10 veces más inversión en genómica en años anteriores), resulta satisfactorio saber que este tipo de investigación es una realidad en nuestro país, la cual debe ser aprovechada al máximo con estudios coherentemente planeados, adecuada formación profesional y apoyo experto internacional.

### Perspectivas para el futuro mediano

Presentados el estado actual de la investigación bioinformática en Colombia, junto con las posibilidades y tendencias de desarrollo científico nacionales, y si a ello agregamos la situación actual a nivel mundial, no cabe más opción que planear estrategias de aplicación a corto plazo, con el fin de afianzar la bioinformática como una importante vía para el desarrollo sostenible y el aprovechamiento de nuestros recursos naturales de acuerdo con lo planteado en el modelo "La biotecnología, motor de desarrollo para la Colombia de 2015" (8).

Previo al mencionado modelo publicado por Colciencias en 2008, que es el resultado de dos años de seguimiento de la actividad científica y biotecnológica en el país, se ha dado un primer paso en la recuperación del modelo institucional para la gestión y producción de conocimiento científico como el que durante varios años sostuvo al Instituto Nacional de Salud. En su afán por crear una institución de componente biotecnológico y que estuviera en sintonía con las actuales tendencias de investigación sostenible, Colciencias y el Servicio Nacional de Aprendizaje (SENA) adjudicaron, en 2007, la creación de un Centro de Excelencia en Metagenómica y Bioinformática al consorcio GeBIX, conformado por siete prestigiosas instituciones que garantizarán una adecuada gestión de los recursos otorgados.

Recientemente, el consorcio dio a conocer el comienzo de sus operaciones a través de su portal de internet ([www.gebix.org.co](http://www.gebix.org.co)); sus objetivos principales son el caracterizar ecosistemas de ambientes extremos y el aprovechamiento de recursos genéticos de nuestro país. Con la creación de este centro de excelencia, el principal órgano gestor e impulsor de la ciencia y la tecnología en Colombia, Colciencias, no sólo establece las pautas de los futuros proyectos de investigación por financiar, sino que da un salto generacional y

promueve la creación de un centro de investigación con tendencia y líneas de investigación con carácter posgenómico. De igual forma, da inicio a la financiación de proyectos con componente bioinformático, con lo cual, imperativamente, crea la necesidad de formar personal en bioinformática, de los cuales carece nuestra sociedad científica.

Siguiendo con lo antes expuesto, se ha empezado a cumplir con el segundo de los requisitos para una nueva política de ciencia y tecnología planteada en la reforma al Programa Nacional de Biotecnología (PNB) de 1991, por parte del mismo Colciencias (8). Sin embargo, el primer requisito, el cual consta de la implementación de un adecuado programa de formación de capital humano, debería haberse promovido antes o, al menos, simultáneamente al apoyo a la investigación.

En este punto se hace realmente imprescindible que en el país se generen programas dirigidos a la formación específica de investigadores en bioinformática y, dado el contexto actual de la educación superior en el país, dichos programas:

- deberían ir dirigidos a la obtención de títulos propios de maestría y doctorado en bioinformática, y
- al menos, los programas de doctorado deberían ser financiados en su totalidad por el organismo gestor de la ciencia y tecnología de turno.

Si asumimos en nuestro futuro programa de ciencia y tecnología tales consideraciones, lograremos un salto tanto cualitativo como cuantitativo en la producción científica, no sólo en el área que tratamos en este ensayo, sino en cualquier área de la investigación científica.

Tales consideraciones son las que se han manejado tradicionalmente en los países europeos, Estados Unidos y, recientemente, en México, Argentina y Brasil, nuestros referentes regionales. Desde la aplicación de tales políticas en países latinoamericanos, ha sido evidente el surgimiento del potencial tecnológico de la región. Actualmente, Brasil y Argentina son líderes mundiales en cultivo y producción de alimentos transgénicos tendientes a suplir las necesidades nutricionales en sus respectivas naciones e, inclusive, a nivel mundial. Por lo tanto, si nos proponemos emerger tecnológicamente en la región, debemos tomar, considerar e implementar, de acuerdo con nuestras necesidades, los programas de ciencia y tecnología vigentes para nuestros referentes regionales y mundiales más próximos.

En cuanto a los modelos europeos, cabe destacar que la financiación de investigaciones junto con la retribución laboral de los estudiantes de doctorado, hacen que el sector científico sea competitivo, productivo y, a su vez, sea considerado como eje importante de desarrollo. Por lo tanto, ante los estímulos nulos para que los estudiantes se incorporen a programas de doctorado nacionales, la aplicación de las consideraciones antes expuestas generará un aumento sustancial del nivel educativo nacional, por el cual, más personas tendrán acceso a programas de posgrado y se contrarrestaría la fuga de cerebros que durante décadas ha dejado al país sin capacidad innovadora.

Además del estímulo de financiación de investigaciones en el campo de la bioinformática y a la creación y financiación de adecuados estudios de maestría y doctorado en el país, surge la necesidad de crear grandes centros de ciencia y tecnología que alberguen y concentren la producción científica en un ambiente académico y multidisciplinario. Ya se ha dado el primer paso con la creación del consorcio GeBIX para el estudio de la metagenómica y bioinformática de ambientes extremos. En un intento más aproximado a lo requerido, a finales del año 2008 Colciencias abrió la convocatoria para la creación de un “modelo de gestión organizacional” (sic.) para un futuro centro de bioinformática y biología computacional en Colombia (convocatoria 484).

La plausible creación de un centro de tal envergadura en nuestro país resultaría en el soporte definitivo para el establecimiento de la bioinformática como campo de estudio para la consecución de la ciencia y tecnología sostenibles que deseamos. Sin duda alguna, dicho centro de investigación y otros posibles que concentren grandes disciplinas deben ser de administración nacional, a cuya gestión puedan acoplarse satisfactoriamente los programas de formación profesional, de donde se obtendrían los futuros doctores que sean capaces de generar conocimiento y lineamientos científicos compatibles con las necesidades de nuestra sociedad.

El cambio generacional en las políticas de ciencia en Colombia se hace cada vez más latente. Durante el desarrollo de este ensayo, el gobierno de Colombia ha promulgado de manera oficial la Ley 1289 de 2009, Ley de Ciencia, Tecnología e Innovación. Con ella se otorgan nuevas facultades administrativas a Colciencias, el cual, a partir de febrero de 2009, empieza a ejercer como Departamento Administrativo de Ciencia, Tecnología

e Innovación. Como consecuencia, la Ley 1289 le confiere a Colciencias una mayor autonomía para administrar los recursos destinados a la ciencia para el beneficio y desarrollo de nuestra sociedad.

Finalmente, y por ello no menos importante, se hace un llamado especial a la integración y asociación de grupos y científicos para generar una red nacional de bioinformática. Este tipo de ejercicios es el único capaz de dar a conocer tanto a nuestros grupos como nuestros campos de acción científica. Simultáneamente, sirven como motor en la retroalimentación de los grupos de investigación para generar nuevo conocimiento a través de la integración con sus pares. Ya somos muchos los grupos que trabajamos en el campo de la biología computacional. Sin embargo, hasta la fecha no se tienen propuestas para la creación de una red en la cual podamos compartir espacios de conocimiento e integración. Reiteramos el llamado y a través de este reporte nos hacemos partícipes directos para evitar que la bioinformática se convierta en una ciencia en la cual lo que prime sea la competencia. Tales prácticas sólo van en detrimento del saber científico y de eso ya hemos padecido bastante.

### Conflicto de intereses

Los autores declaran que no tuvieron ni poseen conflictos de interés para la elaboración del presente escrito. De igual forma, declaran que no han recibido ningún tipo de retribución económica de fuentes industriales, farmacéuticas o gubernamentales para expresar las opiniones presentadas en este ensayo.

### Financiación

Este trabajo se llevó a cabo con fondos provenientes del Centro de Investigación y Desarrollo en Biotecnología, CIDBIO.

### Referencias

- Bernstein FC, Koetzle TF, Williams GJ, Meyer EF, Jr., Brice MD, Rodgers JR, *et al.* The Protein Data Bank: a computer-based archival file for macromolecular structures. *J Mol Biol.* 1977;112:535-42.
- Dayhoff M, Schwartz R, Orcutt B. A model of evolutionary change in proteins. En: Atlas of Protein Sequence and Structure. Dayhoff MO, editor. Washington D.C.: National Biomedical Research Foundation-NBR; 1978. p. 345-52.
- Smith TF, Waterman MS. Identification of common molecular subsequences. *J Mol Biol.* 1981;147:195-7.
- Lipman DJ, Pearson WR. Rapid and sensitive protein similarity searches. *Science.* 1985;227:1435-41.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;215:403-10.
- Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, *et al.* Whole-genome random sequencing and assembly of *Haemophilus influenzae*. *Science.* 1995;269:496-512.
- Blattner FR, Plunkett G 3rd, Bloch CA, Perna NT, Burland V, Riley M, *et al.* The complete genome sequence of *Escherichia coli* K-12. *Science.* 1997;277:1453-74.
- Peña MD, Castellanos O, Carrizosa S, Jiménez C, Del Portillo P. La biotecnología, motor de desarrollo para la Colombia de 2015. Bogotá D.C.: Colciencias; 2008.
- Gutiérrez AJ, Arenas AF, Gómez-Marín JE. Molecular evolution of serine/arginine splicing factors family (SR) by positive selection. *In Silico Biol.* 2006;6:32.
- Matute DR, Barreto-Hernández E, Falquet L. Hunting for insect-specific protein domains. *In Silico Biol.* 2006;6:35-42.
- Arenas AF, Gutiérrez AJ, Gómez-Marín JE. Evolutionary origin of the protozoan parasites histone-like proteins (HU). *In Silico Biol.* 2007;8:2.
- Benítez-Páez A. Sequence analysis of the receptor activity-modifying proteins family, new putative peptides and structural conformation inference. *In Silico Biol.* 2006;6:467-83.
- Benítez-Páez A, Cárdenas-Brito S. Dissection of functional residues in receptor activity-modifying proteins through phylogenetic and statistical analyses. *Evol Bioinform Online.* 2008;4:153-67.
- Garnica DP, Pinzón AM, Quesada-Ocampo LM, Bernal AJ, Barreto E, Grunwald NJ, *et al.* Survey and analysis of microsatellites from transcript sequences in *Phytophthora* species: frequency, distribution, and potential as markers for the genus. *BMC Genomics.* 2006;7:245.
- Acevedo OE, Lareo LR. Amino acid propensities revisited. *OMICS.* 2005;9:391-9.
- Mejía-Guerra MK, Lareo LR. In silico identification of regulatory elements of GRIN1 genes. *OMICS.* 2005;9:106-15.
- Sáenz H, Lareo L, Poutou RA, Sosa AC, Barrera LA. Computational prediction of the tertiary structure of the human iduronate 2-sulfate sulfatase. *Biomédica.* 2007;27:7-20.
- Ulloa JC, Matiz A, Lareo L, Gutiérrez MF. Molecular analysis of a 348 base-pair segment of open reading frame 2 of human astrovirus. A characterization of Colombian isolates. *In Silico Biol.* 2005;5:537-46.
- Narváez G, Lareo L, Rincón J. Mathematical models to correlate molecular topology with substrate affinity of the glycine antagonist in glutamate receptors. *Biomédica.* 2007;27:116-32.
- Benítez-Páez A, Cárdenas-Brito S. Dissection of functional residues in receptor activity-modifying proteins through phylogenetic and statistical analyses. *Evol Bioinform Online.* 2008;4:153-69.
- Restrepo-Montoya D, Vizcaino C, Niño LF, Ocampo M, Patarroyo ME, Patarroyo MA. Validating subcellular localization prediction tools with mycobacterial proteins. *BMC Bioinformatics.* 2009;10:134.
- Benítez-Páez A. Considerations to improve functional annotations in biological databases. *OMICS* 2009;13:527-32.
- Pohlhaus JR, Cook-Deegan RM. Genomics research: world survey of public funding. *BMC Genomics.* 2008;9:472.