



La llegada de la Bioinformática al laboratorio clínico

 18 min.



La revolución de los conocimientos conseguida gracias a los avances científicos producidos en las áreas de biología molecular, genética, genómica y la aparición de la bioinformática buscan mejorar la salud humana, perfeccionando diagnósticos, pronósticos y tratamientos. ManLab ha incorporado en su laboratorio el área de Bioinformática, ésta permitirá profundizar la transferencia biotecnológica, asociando el mundo molecular con el mundo clínico. Estas nuevas herramientas permitirán avanzar hacia una era de la medicina genómica personalizada que posibilitará adecuar la terapia a las características genéticas individuales de cada paciente. En un futuro próximo el diagnóstico, el pronóstico, el tratamiento y la prevención de enfermedades se adaptará a la información personal genética y fenotípica de cada individuo.



Patricio Yankilevich ¹
María Silvia Pérez ²

1. Jefe del Grupo de Bioinformática del Instituto de Investigación en Biomedicina de Buenos Aires (IBiBA) - CONICET - Partner Institute of the Max Planck Society. Godoy Cruz 2390. 1425, CABA, Argentina.
2. Departamento de Medicina Genómica y Biología Molecular – ManLab



Correspondencia:
pyankilevich@ibioba-mpsp-conicet.gov.ar,
maria.perez@manlab.com.ar



Casi una década después de la finalización del primer borrador del genoma humano, la comunidad biomédica está en condiciones de marcar el comienzo de una nueva era de la investigación científica que vincula conocimientos biológicos fundamentales con los conocimientos clínicos. A través de nuevos enfoques holísticos hoy es posible desarrollar y evaluar hipótesis que incorporan el conocimiento genotípico, fenotípico y medio ambiental para el estudio de enfermedades complejas.

Esta traslación de los descubrimientos científicos e innovaciones tecnológicas hacia la salud de los pacientes, mejorando diagnósticos, pronósticos y tratamientos, se está acelerando gracias a una revolución de conocimiento producida en las áreas de biología molecular, genética, genómica y la aparición de la bioinformática como un nuevo campo de investigación científica. Actualmente es posible analizar secuencias de ADN (inclusive genomas humanos completos), analizar secuencias de ARN y cuantificar como se expresan nuestros genes, analizar secuencias de proteínas, su estructura, expresión y modificaciones y estudiar la presencia, estructura y cantidad de otras pequeñas moléculas que participan en nuestro metabolismo, con el objetivo final de

mejorar la salud humana.

Los grandes proyectos de investigación como el proyecto genoma humano no tendrían sentido sin un análisis bioinformático de acompañamiento e interpretación. Sin lugar a dudas la bioinformática tendrá un impacto significativo en la investigación médica y en el laboratorio de diagnóstico clínico. La interpretación de la información genómica es utilizada cada vez más en la práctica de la medicina. El rápido progreso de las tecnologías de laboratorio y el desarrollo de métodos computacionales asociados han permitido incorporar la secuenciación masiva del ADN en la identificación y el descubrimiento de biomarcadores en mezclas biológicas complejas tales como sangre, orina o tejido. A partir de la bioinformática, también denominada biología computacional, podemos hoy identificar biomarcadores clínicamente relevantes que apoyan el diagnóstico, pronóstico y detección de enfermedades.

Medicina personalizada

En estos momentos nos encontramos en los albores de la era de la medicina genómica personalizada, que es la posibilidad de adecuar la terapia a las características genéticas individuales de cada paciente. Al igual que las tecnologías de la información transformaron varios aspectos de nuestras vidas, así también la interpretación del genoma humano está transformando la medicina.

El desarrollo en los últimos años

de las denominadas tecnologías de secuenciación masiva permite actualmente obtener millones de secuencias de ADN a una velocidad y costo cada vez más reducido. En un futuro próximo será más rentable secuenciar el genoma completo de un paciente y utilizar métodos bioinformáticos de análisis de secuencias para implementar los test genéticos que realizar múltiples tests independientes para genes específicos. En un futuro próximo el diagnóstico, el pronóstico, el tratamiento y lo más importante, la prevención de enfermedades se adaptará a la información personal genética y fenotípica de cada individuo (ver Figura 1).

Estas tecnologías están permitiendo alcanzar logros científicos trascendentales, con la resolución de las bases genéticas de enfermedades mendelianas a la cabeza. Su potencial ha permitido el desarrollo de nuevas aplicaciones y pruebas biológicas que van a revolucionar por ejemplo el

diagnóstico de cáncer y el diagnóstico prenatal no invasivo de enfermedades genéticas. La información generada por estas tecnologías es utilizada en:

- Medicina → permitiendo a los médicos realizar una evaluación del riesgo genético probabilístico con el fin de realizar mejores informes preventivos, diagnósticos y decisiones de tratamiento. En casos de detección temprana de riesgos será posible sugerir modificaciones sobre los hábitos y comportamientos del paciente hacia una vida más saludable.

- Ensayos clínicos → permitiendo acelerar los tiempos y reducir los riesgos de los estudios clínicos, al reclutar participantes en base a su perfil genético e historia familiar. Al combinar la genética con resultados del estudio será posible informar el desarrollo terapéutico o identificar las causas genéticas en la respuesta a fármacos y efectos secundarios.

- Investigación → permitiendo una mejor comprensión de la biología de las enfermedades a partir de la exploración genética para dilucidar vías de la enfermedad e identificar nuevas dianas farmacológicas. Así mismo, utilizando grandes poblaciones de controles sanos se podrán identificar las variantes o mutaciones genéticas de protección.



Figura 1. Escenario genómico de la medicina personalizada



Bioinformática y secuenciación masiva

Si bien la secuenciación masiva

Mucho más que resultados.

IACA

LABORATORIOS

- Genética Molecular
- Filiación
- Estudios Forenses
- Citometría de Flujo
- Enf. Metabólicas
- Screening Neonatal
- Toxicología Laboral
- Enf. Infecciosas
- Histocompatibilidad

www.iaca.com.ar

Sede Central Bahía Blanca: San Martín 68 | (B8000BIF) | Tel: +54 0291 459-9999 | laboratorios@iaca.com.ar

Sede Buenos Aires C.A.B.A: Tel: +54 011 43710046 | Móvil: 011 15 513 22214 | buenosaires@iaca.com.ar

Sede Mar del Plata: Móvil: 0223 15 424 9300 | mardelplata@iaca.com.ar

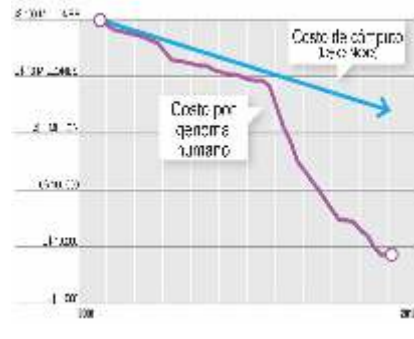
surgió inicialmente con fines de investigación, en la actualidad está siendo utilizada en centros de salud y laboratorios clínicos. Los principales hospitales de investigación y laboratorios clínicos del mundo están activamente implementando prácticas de trabajo para evaluar cómo y cuándo comenzar a ofrecer análisis de genoma completo a sus pacientes. La bioinformática tiene aquí un rol clave, teniendo como desafío lograr que el análisis de secuencias cumpla los requerimientos de eficiencia y reproducibilidad requeridos y que la integración de los datos genómicos personales con otras fuentes de información permita estimar las probabilidades de eventos clínicos como diagnóstico, respuesta al tratamiento o la prognosis de los pacientes. Los datos genómicos de la población general nos proporcionan probabilidades a priori que pueden ser combinadas con datos individuales para calcular las estimaciones de eventos clínicos en cada persona. Por lo tanto, la integración de datos poblacionales con las mediciones genómicas individuales permitirá la práctica de la medicina personalizada. En la actualidad, la secuenciación masiva se está utilizando en los laboratorios clínicos mayoritariamente para hacer diagnóstico a través de paneles, donde un conjunto específico de genes son examinados simultáneamente por una condición o enfermedad determinada.

La era de la información digital ha sido posible gracias a avances fundamentales en la tecnología de semiconductores, hecho que conduce a una disminución del costo computacional. Desde entonces las velocidades de procesamiento se han duplicado continuamente cada dos años y el costo de cómputo ha disminuido en la misma proporción (eso se conoce como la Ley de Moore); además las aplicaciones de software hace mucho que han ganado en cantidad y calidad. En la tecnología de secuenciación del ADN, los avances tecnológicos están ocurriendo incluso más rápido, sobrepasando el crecimiento en poder de cómputo (ver Figura 2). La capacidad de secuenciación, que se duplica cada seis meses, junto con la disminución de los costos y velocidades cada vez mayores, han facilitado en gran medida el descu-

brimiento de asociaciones entre variaciones genéticas en la secuencia del genoma y la enfermedad humana. Hoy es posible secuenciar un genoma humano completo en menos de 24 horas.



Figura 2. Evolución del costo de secuenciación por genoma humano.



Los instrumentos de secuenciación de hoy no sólo realizan la química de la secuenciación, sino que además llevan a cabo la conversión inicial de las secuencias a datos computacionales y realizan el control de calidad de las mismas, un proceso comúnmente conocido como análisis primario. El costo de secuenciación (sin considerar el análisis e interpretación de los datos) se ha reducido drásticamente a menos de 5.000 dólares por genoma completo y cerca de 1.000 dólares por un exoma completo, es decir, solo las partes codificantes de los genes, lo que corresponde aproximadamente al 2% de la secuencia total del genoma humano. Por mucho menos dinero, es posible realizar una resecuenciación dirigida solo a la porción del genoma donde se encuentran las variantes o mutaciones genéticas hoy conocidas, brindando la flexibilidad de elegir que marcadores de enfermedad se desean secuenciar. Los fragmentos asociados a todas las variantes conocidas, corresponden aproximadamente a un 0,03% de la secuencia del genoma humano. La secuenciación ha evolucionado al punto que ya no es algo extraordinario, sino que ahora es accesible, rápida y económica, y ha dejado de ser el foco de la innovación. Como resultado, la secuenciación pronto será accesible para muchos laboratorios clínicos y hospitales, sin embargo la cantidad de datos genómicos está a punto de multiplicarse por

un orden de magnitud.

Aplicaciones médicas de estas tecnologías se están dando en varios países, por ejemplo, los mayores prestadores médicos de los Estados Unidos ya aprueban reembolsos por los tests de detección genética prenatal no invasiva. En estos tests se analizan secuencias de ADN fetales circulantes en la sangre de la madre embarazada. Otro ejemplo ocurre en el Reino Unido, donde este mismo año han puesto en marcha un programa de investigación con 4 millones de dólares de financiación de la Wellcome Trust que permitirá a todos los pacientes de cáncer acceder a los tests genéticos por secuenciación masiva. El resultado deseado del programa es generar los laboratorios, la analítica, la interpretación y las capacidades clínicas del National Health Service (NHS) necesarios para que la información de los tests de predisposición genética al cáncer estén regularmente accesibles en la clínica. La salud pública está comenzando a conocer y utilizar los beneficios de la secuenciación. Otras áreas donde se están adoptando estas tecnologías, además de diagnóstico prenatal y oncología, incluyen la detección de defectos congénitos en recién nacidos, enfermedades infecciosas e inmunología.

Por otro lado, en el campo de la investigación se han obtenido varios resultados de análisis de datos de secuenciación masiva muy relevantes que sugieren otras potenciales aplicaciones en la clínica. Por ejemplo, en una anotación clínica de una secuencia completa del genoma humano de un paciente se analizó riesgo de enfermedades comunes, enfermedades raras y se predijo la respuesta a cientos de fármacos [1]. En otro trabajo, los investigadores evaluaron el riesgo genético para cientos de enfermedades comunes y enfermedades raras identificando variantes asociadas a los diferentes desordenes.

En un análisis similar de una familia (madre, padre, hija, hijo) se consiguió una mayor precisión en la interpretación de variantes genéticas debido a la utilización de un genoma específico de esa etnia familiar en el análisis de las variantes encontradas. [2]. Otros estudios han descubierto la



STAMBOULIAN
LABORATORIO



PORQUE UN DIAGNÓSTICO PRECISO NECESITA RESULTADOS CONFIABLES.

Nuestro laboratorio integral está al servicio del profesional, brindando resultados confiables y asesoramiento en su interpretación, facilitando información precisa para colaborar en el diagnóstico, seguimiento y prevención de las enfermedades.

Nuestro compromiso: brindar un servicio personalizado a través de un equipo de especialistas, cumplir con los más exigentes estándares de calidad, y garantizar confiabilidad y exactitud en los resultados.

/ Biología Molecular
/ Citometría de Flujo
/ Hematología y Hemostasia
/ Inmunoserología

/ Microbiología
/ Química Clínica
/ Endocrinología
/ Virología

PLANTA DE LABORATORIO
Av. Scalabrini Ortiz 676

DPTO. COMERCIAL LABORATORIO
4858-7061 al 63

laboratorio@stamboulían.com.ar

Centro de Atención Telefónica
5411 4515-3000

www.stamboulían.com.ar

STAMBOULIAN
PRIMERO, LA SALUD

genética de enfermedades raras mediante la secuenciación del genoma, [3, 4, 5]. También se ha avanzado en la comprensión de las bacterias que colonizan el intestino y que constituyen el microbioma, conocimiento que sin duda es importante para comprender la respuesta a los fármacos, así como la respuesta inmune. Estos estudios permitirán una mejor comprensión de cómo la dieta y el estado nutricional influyen en la composición y dinámica de las comunidades microbianas del intestino, y la respuesta innata y adaptativa de nuestro sistema inmunológico [6]. En términos de la integración con otras modalidades, un individuo fue seguido por más de 14 meses con un análisis del genoma, transcriptoma, proteoma (incluyendo paneles de auto-anticuerpos) y metaboloma [7]. En el contexto de una infección viral sufrida por este individuo los investigadores observaron el inicio putativo de diabetes mellitus tipo 2. Esto lo lograron integrando millones de mediciones, lo que sugiere que algún día podremos ver la evolución de las respuestas moleculares y rápidamente intervenir para mejorar la salud.

Análisis e interpretación del genoma humano

Si bien el costo de secuenciación de genomas ha disminuido considerablemente, el costo del análisis posterior de los datos no ha seguido el mismo ritmo. El desafío en este caso es triple: ya que el análisis bioinformático de estos datos genómicos debe ser rápido, barato y lo más importante, preciso. Una vez que la muestra del paciente ha sido secuenciada, se obtienen millones de pequeñas secuencias las cuales deben ser alineadas contra un genoma de referencia y ensambladas para formar un genoma completo, obteniendo así una visión completa de las variaciones genéticas y de este modo arribando al genotipo del paciente. Este paso es comúnmente conocido como análisis secundario, debemos imaginarlo como el armado de un enorme rompecabezas de miles de millones de piezas. Dado que este paso requiere de una potente infraestructura computacional y de bioinformáticos capacitados, los costos por realizar un

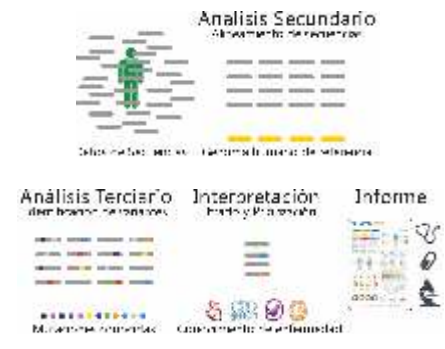
análisis secundario son significativos. Aunque la secuenciación de un único genoma crea terabytes de datos brutos, los datos del genoma ya ensamblado (genotipo) recorta la cantidad de datos a almacenar en mil veces, de terabytes a unos pocos gigabytes.

Una vez concluido el ensamblado del genoma y su genotipo identificado, este debe ser interpretado por relevancia biológica. Este paso se denomina comúnmente análisis terciario y es donde las variantes o mutaciones genéticas son identificadas (ver Figura 3). Aunque existen muchos tipos de variantes genéticas, las mutaciones que afectan exactamente un único nucleótido (Single Nucleotide Variants - SNVs) son las más prevalentes [8]. En nuestro genoma de 6 mil millones de bases se estima que cada uno de nosotros posee unas 10 millones de SNVs, las cuales pueden ser identificadas por comparación contra un genoma humano de referencia. La última versión del genoma humano de referencia es de febrero de 2009 y es de acceso público a través de distintos sitios como: University of California, Santa Cruz (UCSC), the National Center for Biotechnology Information (NCBI), the Ensembl project (ENSEMBL) y Genome Reference Consortium (GRC). Los polimorfismos de un único nucleótido (Single Nucleotide Polymorphism - SNPs) son variantes conocidas donde la sustitución de un nucleótido por otro ocurre en por lo menos el 1% de la población. Aunque los SNPs han sido muy estudiados para encontrar asociaciones con enfermedad y se encuentran anotados en bases de datos, hasta ahora solo unos pocos miles de SNPs han sido asociados a rasgos fenotípicos específicos en forma estadísticamente significativa [9]. Otras variaciones que también pueden ser identificadas son las variaciones estructurales, como las variaciones en el número de copias de segmentos de ADN (Copy Number Variations – CNVs), las inserciones y deleciones (InDels) o translocaciones de secciones enteras de ADN. El paso siguiente en el flujo de análisis del genoma de un individuo es la interpretación de las variaciones identificadas. En este paso una cierta variante o mutación genética se asocia con una enfermedad determinada o

rasgo físico (fenotipo).



Figura 3. Análisis e interpretación de secuencias del genoma humano



El reto de la interpretación está en encontrar aquellas variaciones que tienen un impacto en la función biológica, que puedan ser perjudiciales, para luego diagnosticar o prescribir un tratamiento adecuado. De alguna manera esto es como buscar una aguja en un pajar, aunque por suerte contamos con las computadoras y algoritmos para hacerlo. Si bien aún no existe un proceso estandarizado de cómo realizar esta interpretación, existen varias estrategias, métodos y algoritmos que, conjuntamente con un buen poder de procesamiento, se pueden utilizar para lograr una interpretación del análisis terciario en forma rápida y eficaz (ver Figura 4). Estos resultados permitirán a los médicos y científicos extraer el significado biológico de los complejos datos genómicos, conectando mutaciones específicas con marcadores de riesgo conocidos de manera sistemática y reproducible.

Ahora es posible secuenciar el genotipo del paciente y obtener la lista de todas las variaciones candidatas, el siguiente paso es combinar este conocimiento con otros datos médicos - historia clínica, el medio ambiente, el entorno familiar, el microbioma, la dieta, los hábitos, etc. Para mejorar la atención al paciente y desarrollar terapias personalizadas es necesario poder integrar y aprovechar estos datos moleculares y clínicos complejos provenientes de una variedad de fuentes internas, externas y

públicas. Sin embargo, existe una preocupación cada vez mayor porque las variaciones genéticas comunes (SNPs) no estarían explicando toda la variabilidad genética observada en fenotipos relevantes en salud y enfermedad; por lo tanto, la interpretación de variantes raras (en particular en regiones reguladoras no codificantes) tiene también gran relevancia en este contexto.

Los estudios genéticos pueden tener distintas finalidades, como por ejemplo:

- Identificación de mutaciones en genes de enfermedades de transmisión hereditaria, donde se buscan mutaciones constitucionales heredadas de los padres (por ejemplo Fibrosis quística o Hemocromatosis).
- Identificación de mutaciones en genes candidatos de causar enfermedades complejas, para luego realizar estudios funcio-

nales de los mismos (por ejemplo Parkinson o las enfermedades autoinmunes).

- Identificación de mutaciones somáticas en genes asociados a enfermedades como cáncer y diabetes, entre otras.



Figura 4. Interpretación. Filtrado y priorización de variantes en genomas humanos [Adaptado de 10]



Conclusión

Hace más de un siglo, una nueva tecnología provocó una nueva disciplina médica: la radiología. Al ir madurando la radiología, la placa de rayos X se convirtió en un estándar. En poco tiempo, los datos de radiología se habían convertido en parte de la historia clínica del paciente. Lo mismo sucederá con los datos genómicos. Al igual que hoy en los hospitales muchos pacientes son derivados a la unidad de rayos, en un futuro próximo habrá también una unidad de genómica, donde inmediatamente se secuenciarán sus genes y/o genomas para detectar las mutaciones presentes para ser analizadas e integradas con el resto de su historia clínica. Pronto pasaremos de investigar el genoma y las enfermedades a desarrollar diagnósticos y transformar el conocimiento en nuevas y mejores terapias.

Se ha estado hablando de la promesa del genoma desde hace tiempo. Si bien el costo y la complejidad asociada al proceso

**Antes de solicitar un
análisis clínico,
lo primero que hay que
es el**

**analizar
laboratorio.**

Más de 30 años de trayectoria dedicados al cuidado de los pacientes.

OAA ✓

Organismo
Argentina de
Acreditación

Laboratorio Clínico

LM 013
Consulte siempre acreditación
en: www.oaa.org.ar

La excelencia se construye desde
la calidad, la tecnología y el profesionalismo.
Resultados confiables para efectuar
diagnósticos certeros.

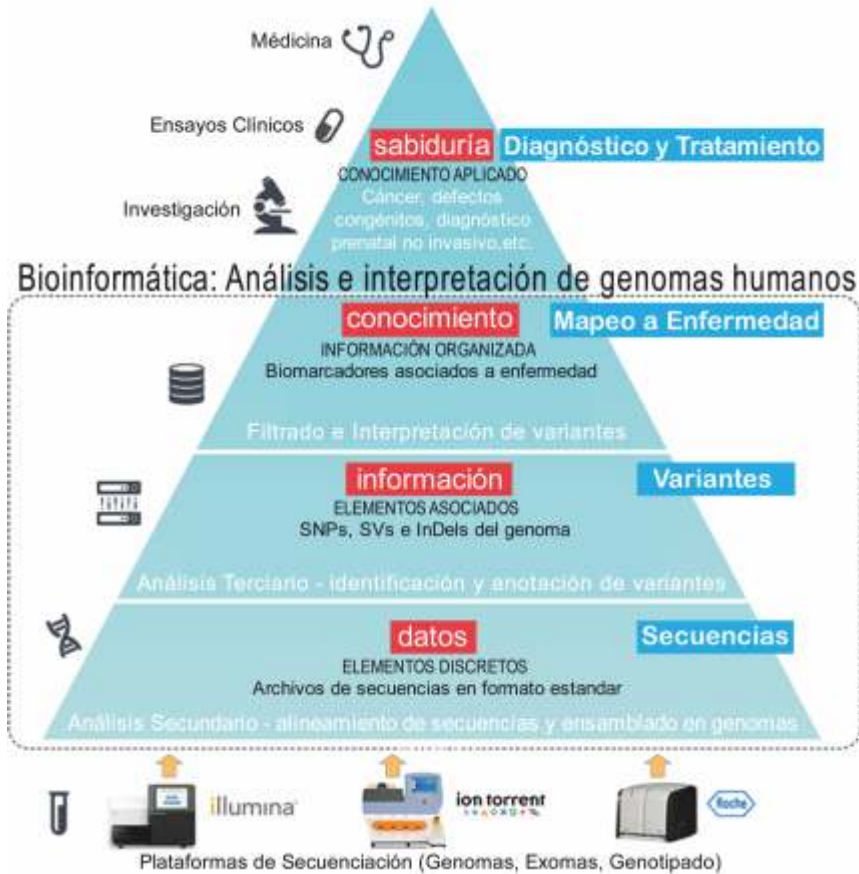


LABORATORIO DE MEDICINA
ANÁLISIS CLÍNICOS | Dr. Raul Gutman

www.labmedicina.com - Tel: 011 4514 9370 - Acreditado Norma ISO 15189



Figura 5. Esquema de plataforma bioinformática de análisis e interpretación de genomas humanos.



de secuenciación del genoma humano han sido hasta el momento los mayores cuellos de botella, ahora estamos muy cerca de poder aprovechar el poder de los diagnósticos y las terapias basadas en el ADN. En nuestro país contamos ya con el conocimiento y los recursos biotecnológicos y computacionales para lograr implantar este tipo de prácticas.

Al entrar en la segunda década del siglo XXI, las inversiones en biología molecular, bioinformática, en gestión de la enfermedad y en los esfuerzos de interpretación del genoma humano están finalmente dando sus frutos. La medicina personalizada promete revolucionar la práctica de la medicina, transformar la industria de la salud y en última instancia conducirnos hacia una vida más larga y saludable. En la actualidad los médicos, ge-

netistas e investigadores están pidiendo utilizar estas tecnologías, ya que por ejemplo una identificación de genes defectuosos podría mostrar cómo tratar un tumor o ayudar a diagnosticar las causas subyacentes de algunas enfermedades.

El Instituto de Investigación en Biomedicina de Buenos Aires (IBioBA) se encuentra prestando una asesoría a ManLab para el análisis e interpretación de datos provenientes de tecnologías de secuenciación masiva. Este tipo de emprendimientos seguramente puedan contribuir a la profundización de la transferencia biotecnológica, asociando aún más el mundo molecular con el mundo clínico, lo que permitirá en el futuro tomar decisiones asociadas al diagnóstico, tratamiento, pronóstico y prevención de enfermedades en forma rápida y mejor informada.

Referencias

- [1] Ashley E. A., Butte A. J., Wheeler MT, et al. Clinical assessment incorporating a personal genome. *Lancet*; 375: 1525-35.
- [2] Dewey, F. E, et al. Phased Whole-Genome genetic risk in a family quartet using a major allele reference sequence. *2011 PLoS Genet* 7
- [3] Worthey, E. A., et al. Making a definitive diagnosis: successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. *2011. Genetics in medicine* 13: 255-262.
- [4] Mardis, E. R. Applying next-generation sequencing to pancreatic cancer treatment. *2012. Nat Rev Gastroenterol Hepatol* 9, 477-486.
- [5] Chen, R., et al. Personal omics profiling reveals dynamic molecular and medical phenotypes. *2012. Cell* 148:1293-1307.
- [6] Kau, L., et al. Human nutrition, the gut microbiome and the immune system. *Nature* 474 No. 7351 (2011), pp. 327-336.
- [7] Cronin, M., et al. Analytical validation of the oncoType DX genomic diagnostic test for recurrence prognosis and therapeutic response prediction in Node-Negative, estrogen Receptor-Positive breast cancer. *2007. Clin Chem* 53:1084-1091.
- [8] The 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature*, Vol. 467, No. 7319. (2010), pp. 1061-1073.
- [9] Manolio T. A., et al. Finding the missing heritability of complex diseases. *Nature In Nature*, Vol. 461, No. 7265. (2009), pp. 747-753
- [10] Hayden E. C. Gene sequencing leaves the laboratory. *Nature*, Vol. 494, No. 7437. (2013), pp. 290-291.



Los bioquímicos de **La Rioja**
cuentan con nosotros ...
Somos Socios Complementarios



Parque Nacional Talampaya

Informes: (5411) 4508 2091 - www.manlab.com.ar

MANLAB[®]
Diagnóstico Bioquímico y Genómico