Nuevo método para el análisis acelerado de datos masivos generados por los barridos genómicos completos

19/04/2010

Nature Genetics - 19 de abril de 2010

Los estudios genéticos representan la nueva frontera del descubrimiento en Biología y Medicina y la implementación de técnicas moleculares cada vez más potentes y complejas están permitiendo revolucionar la investigación médica hoy y la práctica de la medicina mañana.

En los próximos 10 años es muy posible que la secuencia de nuestro genoma sea parte de nuestro historial médico. Los mayores problemas no serán probablemente ni la tecnología ni el costo para obtenerlo, sino la capacidad informática para analizar, interpretar y almacenar las ingentes cantidades de información generadas.

Actualmente, el tratamiento de las cantidades masivas de datos que se manejan en estudios genéticos supone un tiempo de procesado de ordenador que puede llegar en algunos casos a ser prohibitivo. Sin embargo, esto no es más que un atisbo de las necesidades futuras ya que las bases de datos contienen miles de sujetos y millones de datos genéticos por sujeto.

El problema será varios órdenes de magnitud mayor cuando estas bases de datos contengan los miles de millones de pares de bases del genoma de cada individuo en millones de personas. Por lo tanto es esencial el desarrollar métodos de cálculo estadístico para copar con las nuevas necesidades y así evitar que el análisis de los datos se convierta en el cuello de botella del progreso científico y medico.

Podemos pensar en un paralelismo con algo que se ha hecho común en nuestras vidas en la última década.

Sería imposible transmitir la cantidad de fotografías que se mueven por internet hoy en día con la velocidad disponible en los albores de internet. Pero llega un momento que incluso aumentar la velocidad no es suficiente y hay que desarrollar algoritmos que reduzcan el numero de pixeles sin afectar calidad y resolución (por ejemplo el popular formato JPG).

El trabajo que estos investigadores, entre los que se encuentra el Dr. José María Ordovás, ha llevado a cabo realiza algo similar con los datos genéticos, al comprimir la información que se analiza sin pérdida alguna de resolución y así conseguir reducciones espectaculares en el tiempo de procesado necesario para llevar a cabo estudios genéticos masivos de días a minutos, esta aceleración hace factible estudios cada vez más complejos y que el progreso en este área de la biomedicina pueda seguir adelante sin esta barrera que teníamos hasta ahora.

Mixed linear model approach adapted for genome-wide association studies.

Nature Genetics, 2010 Apr;42(4):355-60 Zhang Z, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu J, Arnett DK, Ordovas JM, Buckler ES. Artículo en PubMed.

Ver noticia

URL de origen:https://www.cnic.es/es/noticias/nuevo-metodo-para-analisis-acelerado-datos-masivos-generados-por-barridos-genomicos