Sistema bioinformático de apoyo a la investigación en la unificación del patrimonio genealógico y genético

Researchers:

- Álvaro Rodríguez San Pedro. TFG en Grado en Ingeniería Informática en Ingeniería de Software de la UEx [1]. Codirigido por Félix Rodríguez Rodríguez y José Luis González Sánchez.

Idioma Indefinido

Descrição:

El desarrollo en el campo de la genética ha abierto nuevas puertas en el estudio de enfermedades y en el diagnóstico clínico. La medicina de precisión pretende aprovechar esas nuevas capacidades. Aunando el estudio de la variabilidad genética de los individuos, su entorno y sus hábitos de vida, pretende identificar estrategias de tratamiento y prevención de enfermedades para grupos concretos. En Extremadura, la aplicación de tales técnicas está supeditada a la disponibilidad de la información. Para ello, es necesaria una plataforma que permita la unificación del patrimonio genealógico y genético, conjuntamente con los datos clínicos.

Este TFG pretende responde a esta necesidad, desarrollando un sistema bioinformático que permita la mencionada integración y, además, que dé soporte a la investigación médica. El objetivo es estudiar los beneficios de la unificación de fuentes de información heterogéneas, como registros civiles, eclesiásticos, o historiales clínicos, al estudio de enfermedades hereditarias. El estudio se ha enfocado a un conjunto de personas de una población relevante desde el punto de vista genético, para secuenciar su genoma (concretamente los genes asociados a la enfermedad elegida) y unificarlo con su información genealógica. La zona a estudio es la del Valle del Jerte debido a las características que presenta en cuanto a su población y entorno.

Para gestionar y analizar la información, así como para extraer conocimiento de la misma, se plantea el uso de recursos de supercomputación en combinación con técnicas de Big Data y de inteligencia artificial.

El objetivo se materializa en la implantación de una infraestructura de apoyo a la investigación genética con las siguientes funcionalidades:

- Unificación de la información genealógica y personal e información genómica.
- Gestión de grandes volúmenes de datos.
- Facilitar el acceso y el análisis de la información.
- Extracción de información a partir la integración de todas las fuentes de datos.

Los resultados obtenidos muestran que se cuenta con 637 documentos almacenados en la base de datos, de los cuales 380 son de participantes y 257 de pedigríes. Esto representa aproximadamente un 4% del total de la población a estudio.

La diferencia de documentos pone de manifiesto que a pesar de ser pocas personas, un número elevado de las mismas están relacionadas. Analizando la relación de participantes de cada población, es posible sospechar que, retrotrayéndose lo suficiente, se pueden encontrar ancestros comunes para casi todos los participantes. Además, se cuenta con otros datos que proporcionan información útil de la población estudiada y que podrían incorporarse al estudio, proporcionando otro tipo de conclusiones. Éstos son la edad, la actividad profesional, etc. Hay que considerar que para cada gen hay información de hasta cinco bases de datos (aunque no siempre hay información de todas) y que los riesgos están organizados en grupos de cáncer, que a su vez se dividen en cánceres específicos.

Con todo, se ha conseguido una base de datos que aúna múltiples fuentes, aportando valor extra al proyecto. La última base de datos es la de variantes. En ella, hay información de 2 runs, un total de 559.374 variantes y 1.283.123 análisis de cobertura para transcritos, para un total de 1.842.499 documentos.

URL de origem:https://www.cenits.es/pt-pt/proyectos/sistema-bioinformatico-apoyo-investigacion-unificacion-patrimonio-genealogico-genetico

Ligações

[1] https://www.unex.es/