

CAMBioMit, Desarrollo de modelos predictivos de **Cáncer Colorrectal** basados en la **interacción** entre **factores ambientales y genéticos** de la biogénesis mitocondrial

Objetivo del proyecto

Desarrollar modelos predictivos para la identificación de mutaciones en la biogénesis mitocondrial y su interacción con factores ambientales en la incidencia del cáncer colorrectal.

Financiación del proyecto

Junta de Castilla y León. (Orden EDU/986/2017, de 8 de noviembre, "B. O. C. y L. de 10 de noviembre).

Participantes del proyecto

Universidad de León, www.unileon.es

SCAYLE, Supercomputación Castilla y León,
www.scayle.es

Periodo de ejecución

Octubre del año **2017** a noviembre del **2019**.



Vicente Martín, investigador principal, y dos miembros del proyecto, Carmen Gonzalez y María Garcia.

Justificación del proyecto

El desarrollo de modelos basados en datos genéticos y ambientales del proyecto supondrán una mejora en la capacidad para la predicción del Cáncer Colorrectal (CCR). La aplicación directa de estos modelos facilitará una elección más personalizada en las estrategias de prevención primaria y secundaria del CCR e indirectamente, una mejora e innovación en el desarrollo de metodologías que, incluyendo información ambiental, genética y/o de biomarcadores, podrán ser aplicadas a otras patologías.

Este proyecto está enmarcado en la línea de trabajo sobre el papel de la biogénesis mitocondrial que esta UIC está liderando dentro del Consorcio de Genética y Epidemiología del Cáncer Colorrectal (GECCO). Este consorcio es el más relevante a nivel internacional en su campo, por lo que es de esperar un impacto altamente relevante en la capacidad de liderazgo internacional de nuestra UIC, en una prometedora línea de investigación en oncología y epidemiología del cáncer.

En cuanto a la contribución en el marco de las tecnologías facilitadoras esenciales, nuestra propuesta incide en el ámbito de la biotecnología industrial, toda vez que trata de buscar modelos que permitan una selección más adecuada de biomarcadores con fines predictivos. Proponer vías de desarrollo y mejora de los modelos predictivos establecidos hasta el momento, dará como resultado una mejor valoración del riesgo de padecer CCR y por ello una aplicación socioeconómica clara, al reducir costes directos e indirectos de las estrategias actuales de prevención primaria y sobre todo secundaria.

Asimismo, el proyecto plantea una importante acción interdisciplinar con la colaboración de áreas de investigación que incluyen la epidemiología ambiental y genética en la que trabaja la UIC, y también el campo de la ingeniería informática y de datos, que es el marco de interés de las empresas colaboradoras en el proyecto. Los modelos predictivos planteados se basan en el desarrollo de sistemas inteligentes de gestión de la información, basados en la extracción avanzada de datos y el aprendizaje automático, que son considerados como algunas de las TIC con más proyección y recorrido. Son estas, por otro lado, unas tecnologías facilitadoras transversales cuya evolución y desarrollo van a tener un efecto más allá del campo de acción de este proyecto.

Es reseñable también, dentro del interés del proyecto, su vinculación con las líneas de trabajo planteadas en el Horizonte 2020, dentro de sus

retos en Salud, cambio demográfico y bienestar. Así, el objetivo de este proyecto se enmarca dentro de la búsqueda de soluciones y el planteamiento de formas innovadoras e interdisciplinares para la comprensión de los factores determinantes de la salud y de la enfermedad. En este sentido, el proyecto responde a la necesidad de integración de datos de comportamiento, medioambientales y provenientes de las ciencias "ómicas", buscando mejorar la capacidad de identificación de personas y poblaciones con alto riesgo de enfermedad con el objetivo de desarrollar estrategias personalizadas, estratificadas y colectivas tendentes a una prevención eficaz y eficiente de las enfermedades.

Los modelos predictivos generados como resultado del proyecto, gracias a la colaboración entre la Universidad de León, a la que pertenece la UIC y las empresas que participan en su desarrollo, se espera que den lugar a la generación y explotación de un modelo de utilidad que haga más eficientes y adecuadas las intervenciones en prevención primaria y secundaria del CCR, considerándose una obligación ética adecuar las intervenciones a las personas.

Por otro lado, es esperable que la difusión del conocimiento generado tanto sobre el desarrollo de este tipo de modelos predictivos, basados en el machine learning y los algoritmos genéticos, como sobre el papel del pathway de la biogénesis mitocondrial y de los SNPs implicados en el mismo, puedan ser de interés para el posterior desarrollo de nuevos modelos predictivos tanto para el CCR como para otras patologías o el estudio de otras rutas genéticas.

Funciones de SCAYLE

El grupo no contaba con la potencia de cálculo necesaria para llevar a cabo los análisis para ese fin, por ello se solicitaron recursos de SCAYLE, así como el asesoramiento ante requerimientos informáticos que superaban la capacidad de los sistemas informáticos de la UIC-109.

Se requirió la utilización del software R para llevar a cabo los análisis estadísticos y gráficos necesarios.

Líder del proyecto

G.I. INTERACCIONES GEN- AMBIENTE-SALUD, de la Universidad de León. El investigador principal del proyecto y director del grupo es Vicente Martín.

Este grupo de investigación desarrolla estudios sobre los hábitos de vida y su interacción con factores genéticos.