

Genómica funcional y análisis de microarrays

M. Carme Ruíz de Villa
Alex Sánchez-Pla

PID_00192741

Material docente de la UOC

M. Carme Ruíz de Villa

Alex Sánchez-Pla

El encargo y la creación de este material docente han sido coordinados por el profesor: Daniel Riera Terrén (2013)

Primera edición: febrero 2013
© M. Carme Ruíz de Villa, Alex Sánchez-Pla
Todos los derechos reservados
© de esta edición, FUOC, 2013
Av. Tibidabo, 39-43, 08035 Barcelona
Diseño: Manel Andreu
Realización editorial: Eureka Media, SL
Depósito legal: B-1.722-2013

Ninguna parte de esta publicación, incluido el diseño general y la cubierta, puede ser copiada, reproducida, almacenada o transmitida de ninguna forma, ni por ningún medio, sea éste eléctrico, químico, mecánico, óptico, grabación, fotocopia, o cualquier otro, sin la previa autorización escrita de los titulares del copyright.

Introducción

Esta asignatura es una introducción al análisis de datos de alto rendimiento, particularmente los microarrays de expresión, tal como se está llevando a cabo por la comunidad bioinformática a principios de la segunda década del siglo XXI.

El análisis de microarrays ha revolucionado la biología de principios de este siglo en cuanto que ha conllevado un cambio de paradigma: se ha pasado de poder estudiar los genes de uno en uno a poder hacerlo todos a la vez. Esto, sin embargo no ha sido sin coste, dado que los datos que se han ido generando presentan tanto problemas técnicos como una considerable cantidad de ruido. La eliminación de estas fuentes de error puede llevarse a cabo usando métodos estadísticos, que también se han revelado de gran importancia a la hora de analizar los datos, ya sea para seleccionar genes relevantes, para buscar perfiles de variación común entre los genes o para construir clasificadores a nivel molecular.

Por todo lo anterior el análisis de datos de microarrays es una disciplina que se encuentra a caballo entre la estadística y la bioinformática y su dominio requiere la comprensión de los procesos biotecnológicos que generan los datos, el manejo de las herramientas bioinformáticas necesarias para manipular la información y la comprensión de los métodos estadísticos subyacentes a los análisis.

Con estos requisitos un manual para el análisis de microarrays tiene que ser necesariamente extenso en tanto que tendrá que tratar mínimamente los tres aspectos considerados.

El texto que se presenta a continuación es el fruto de la experiencia en el aula –presencial y virtual– de más de 5 años de docencia en este tema y pretende buscar un equilibrio entre la formación y la información para lo cual se han seleccionado cuidadosamente los temas y puntos a tratar, de manera que el estudiante pueda adquirir un sólido conocimiento sin que los detalles le hagan perder la perspectiva, en la que, probablemente, sea su primera incursión en este tema.

Aunque no parezca lo usual en este tipo de textos, en este caso se hace necesaria una pequeña sección de agradecimientos. La versión final de este documento no habría llegado a buen puerto sin la desinteresada colaboración de muchas personas. Ferran Briansó, Xavier de Pedro, Eudald Illa, Ricardo Gonzalo y Josep Gregori han contribuido con sus comentarios a mejorar muchas partes del texto y con su ayuda a detectar un sinnúmero de pequeños problemas. Marta Casals, Oriol Rua y Silvia Cardona han invertido un buen número de horas

en revisar y corregir imágenes texto y código. Sin la ayuda de todos ellos, este documento no habría visto la luz. Los errores no detectados corresponden naturalmente a los autores.

Finalmente, un texto de estas características está plagado de programas que –esperamos– podéis querer probar por vosotros mismos, para lo que podéis además necesitar los datos de los ejemplos que deseáis reproducir. Con el fin de facilitar esta forma de aprendizaje, todos los códigos del texto así como los datos utilizados y otros materiales de interés los podéis encontrar en el aula.

Objetivos

Los objetivos de esta asignatura son los siguientes:

- 1.** Comprender la tecnología de los microarrays y conocer sus diversas variantes.
- 2.** Adquirir conocimientos básicos de R y estadística que permitan llevar a cabo de forma comprensiva análisis básicos de microarrays.
- 3.** Conocer los principales problemas que pueden resolverse con microarrays.
- 4.** Saber llevar a cabo análisis básicos con datos reales.
- 5.** Conocer las limitaciones de los microarrays y su relación con las nuevas tecnologías de alto rendimiento como la ultrasecuenciación.

Contenidos

Módulo didáctico 1

Preliminares

M. Carme Ruíz de Villa y Alex Sánchez-Pla

1. Introducción a los microarrays
2. El lenguaje estadístico R
3. Fundamentos de estadística

Módulo didáctico 2

Análisis de datos de microarrays

M. Carme Ruíz de Villa y Alex Sánchez-Pla

1. El proceso de análisis de datos de microarrays
2. Diseño de experimentos de microarrays
3. Exploración de los datos, control de calidad y preprocesado
4. Selección de genes diferencialmente expresados
5. Después de la selección: análisis de listas de genes
6. Caso resuelto: Selección de genes diferencialmente expresados
7. Métodos avanzados: Búsqueda de patrones y predicción
8. Caso resuelto: Descubrimiento de clases