

gura 1) y automatizado para la identificación y cuantificación masiva de péptidos mediante marcaje isotópico estable con ^{18}O , ofreciendo unas prestaciones superiores en cuanto a identificación (mayor número de identificados con la misma tasa de error) y en cuanto a robustez en la detección de cambios de expresión estadísticamente significativos.

La biología computacional de la proteómica post-genómica

Segura Ruiz V.

Unidad de Proteómica, Genómica y Bioinformática, CIMA

Introducción

El estudio de la expresión de genes mediante técnicas de alto rendimiento, como los microarrays de ADN, obligó a un cambio de paradigma en la interpretación de los resultados obtenidos con respecto a la experimentación clásica empleada en biología. La bioinformática tiene entre sus principales objetivos dar respuesta a estas nuevas necesidades desde múltiples disciplinas, como son la estadística y la informática. En proteómica las necesidades computacionales han surgido principalmente de los experimentos de identificación de proteínas mediante geles 2D y espectrometría de masas, es decir, han estado orientadas a la gestión de grandes cantidades de información, el procesado de imagen y de espectros y la búsqueda en bases de datos. Sin embargo, la aplicación de métodos ya empleados en genómica a estos datos, o la aparición de nuevas tecnologías como los microarrays de proteínas o el SELDI-TOF, abren nuevas expectativas en el área de la proteómica, tanto en el ámbito científico como en el ámbito clínico.

Material y métodos

El análisis de los datos obtenidos en un experimento de microarrays de ADN requiere una serie de procesos independientemente de la plataforma o del objetivo del estudio (Hoheisel, 2006): cálculo de los niveles de expresión utilizando el método de normalización adecuado, acondicionamiento de la matriz de expresión, y la aplicación del método de selección

Bibliografía

- Martinez-Bartolome, S., Martín-Maroto, F., Navarro, P.J., López-Ferrer, D., et al., *Mol. Cell. Proteomics* (in review).
- Ramos-Fernandez, A., Lopez-Ferrer, D. & Vazquez, J. *Mol Cell Proteomics*, 2007, 6, 1274-1286.
- Elias, J.E. & Gygi, S.P. *Nat Methods*, 2007, 4, 207-214.

(foldchange, estadística t, modelos lineales, etc), con el fin de identificar los genes de interés. A partir de esta lista de genes existe una gran diversidad de opciones para la interpretación de los resultados: análisis de los perfiles de expresión, análisis funcional, estudio de factores de transcripción, o la integración de información mediante redes de genes con programas como Ingenuity (www.ingenuity.com).

En el ámbito de la proteómica, con el empleo de nuevas tecnologías como los microarrays de anticuerpos (Madoz-Gúrpide et al, 2007) o el SELDI-TOF (Vorderwülbecke, 2005), el análisis precisa de métodos similares a los empleados en genómica, ya sea para el proceso de normalización, o para la selección, mediante métodos estadísticos y herramientas bioinformáticas apropiadas, de las proteínas alteradas o de la expresión diferencial de péptidos.

Conclusiones

La bioinformática permite, con datos provenientes de la experimentación en proteómica, el desarrollo de nuevas aplicaciones, como son la identificación de biomarcadores o de nuevas dianas terapéuticas para el desarrollo de medicamentos.

Bibliografía

- JD., Hoheisel. *Nature Reviews Genetics* 2006, 7, 3, 200-210.

J., Madoz-Gúrpide, M., Cañamero, L., Sánchez, J., Solano, P., Alfonso, JI., Casal. *Molecular&Cellular Proteomics* 2007, 6, 12, 2150-2164.

S., Vorderwülbecke, S., Cleverley, SR., Weinberger, A., Wiesner. *Nature Methods*, 2005, 2, 393–395.

Desarrollo de una base de datos de proteínas identificadas en estudios de interacción patógeno-hospedador: PROTEOPATHOGEN

Fernández Vital V, Pascual-Montano A, Gil C.

Departamento Microbiología II, Unidad I. Facultad de Farmacia. UCM

Introducción

A medida que la tecnología proteómica mejora en sensibilidad y resolución se hace necesario el uso de herramientas adecuadas que permitan almacenar, analizar e interpretar los datos generados en los experimentos, tareas en las que la bioinformática juega un papel esencial. La combinación de técnicas como la electroforesis bidimensional y cromatografía, usadas para la separación de proteínas; y la espectrometría de masas, para su identificación, generan una enorme cantidad de datos muy complejos que han de ser convenientemente almacenados para que puedan ser consultados y analizados de manera eficiente. El objetivo de este trabajo es el desarrollo de una base de datos que resuelva esta necesidad existente en un laboratorio de proteómica.

Material y métodos

Datos biológicos. Los datos almacenados corresponden a proteínas procedentes de diversos experimentos enmarcados en estudios de patogenicidad microbiana e interacción patógeno-hospedador. La mayor parte de las identificaciones corresponden a la levadura *Candida albicans* (Fernández-Arenas et al., 2007) y a proteínas de macrófagos murinos (Martínez-Solano et al., 2006).

Implementación y estructura de la aplicación. La base de datos se presenta como un modelo servidor-cliente que permite múltiples conexiones y consultas a través de una interfaz web de acceso. El sistema gestor de bases de datos relacionales utilizado es PostgreSQL. Otros recursos usados en el desarrollo son php, Apache, Perl y R-cran. En todos los casos

se trata de software de libre acceso, ampliamente utilizado y de comprobada robustez y fiabilidad.

Resultados

PROTEOPATHOGEN es una base de datos en desarrollo que almacena y relaciona de forma eficiente información relativa a proteínas implicadas en interacción patógeno-hospedador. Además la creación de una interfaz de acceso web permite realizar consultas de manera sencilla e intuitiva. La unidad básica de consulta es la proteína, a través de la cual se puede recuperar información como el experimento y la especie de origen, los identificadores o la lista de péptidos correspondiente a la identificación.

Conclusiones

El desarrollo de una base de datos que integre y relacione la información generada en experimentos de proteómica no solo supone la forma adecuada de almacenar los datos sino que proporciona una herramienta de gran utilidad para el análisis e interpretación de los resultados de experimentos en proteómica.

Bibliografía

- Fernández-Arenas, E., Cabezón, V., Bermejo, C., Arroyo, J., Nombela, C., Díez-Orejas, R., Gil, C., 2007. *Molecular and Cellular Proteomics* 6(3), 460-78
- Martínez-Solano, L., Nombela, C., Molero, G., Gil, C., 2006. *Proteomics* 6:S133-S144