



Informe final ID2017_037

Herramientas bioinformáticas web públicas para el estudio de alteraciones genéticas en tumores cerebrales



Ester Gangoso Rodríguez
Scottish Center for Regenerative Medicine
Edinburgh University, Scotland, U.K.
Unidad de Investigación Consolidada
JCyL013 USAL



Arantxa Taberero Urbieta
Instituto de Neurociencias de Castilla y León
Dpto. de Bioquímica y Biología Molecular
Unidad de Investigación Consolidada
JCyL013 USAL

Resumen:











Es mucha y muy variada la información publicada sobre las alteraciones genéticas que dan lugar a la aparición de los Glioblastomas (los tumores cerebrales más frecuentes y fatales). Sin embargo, está muy dispersa y resulta de difícil acceso para no especialistas en el tema. Por ello, la utilización de un sistema sencillo y accesible facilitaría la consulta de toda esa información tanto a estudiantes interesados en Neuro-Oncología, como posiblemente a facultativos en formación, implicados en el tratamiento de los tumores cerebrales. Hasta la ejecución de este proyecto no se disponía de nada similar en la USAL, o en otras Universidades de nuestro entorno, por lo que la puesta en marcha de este proyecto significará un salto cualitativo.

El uso de plataformas web es algo cada vez más común, pero en este proyecto, vamos a descubrir al alumno nuevos recursos más específicos de la Neuro-Oncología. El estudiante, una vez familiarizado con el uso eficiente de los recursos que le proponemos, será capaz en el futuro de extrapolarlo a otros casos, trabajando de forma autónoma.

A lo largo de éste curso se han recopilado todos los materiales y la información necesaria y se han elaborado los recursos propuestos. Todo ello se ha incorporado a un nuevo curso en Studium, titulado: “Herramientas bioinformáticas web públicas para el estudio de alteraciones genéticas en tumores cerebrales” que se hará accesible a los estudiantes el próximo curso, en el primer cuatrimestre, que es cuando se imparte la asignatura de Master para la que se propuso su diseño.

Elaboración del Cuestionario Inicial de Evaluación:

Para calibrar el grado de conocimiento de los estudiantes al comienzo de la actividad hemos elaborado un cuestionario muy simple que no les lleva más de 5 min contestar y nos da una idea de cuál es el nivel que debe darse a las explicaciones para hacerlas más comprensibles y efectivas.

<p>Pregunta 1 Sin contestar Puntúa como 1</p> <p> Marcar pregunta</p> <p> Editar pregunta</p>	<p>¿Sabes si existen bases de datos con información de muestras de pacientes?</p> <p>Seleccione una:</p> <ul style="list-style-type: none"><input type="radio"/> a. No, debido a la protección de datos<input type="radio"/> b. Si, algunas de ellas abiertas a todo tipo de público<input type="radio"/> c. Si, pero solo para uso de personal de hospital e investigadores<input type="radio"/> d. Si, pero todas son de pago
<p>Pregunta 2 Sin contestar Puntúa como 1</p> <p> Marcar pregunta</p> <p> Editar pregunta</p>	<p>¿Los glioblastomas tienen alteraciones genéticas?</p> <p>Seleccione una:</p> <ul style="list-style-type: none"><input type="radio"/> a. No, las células del cerebro no pueden sufrir mutaciones debido a su importancia.<input type="radio"/> b. Si, es una característica de estos tumores<input type="radio"/> c. No, todavía no se conocen alteraciones genéticas.<input type="radio"/> d. Si, pero solo las células que no son neuronas
<p>Pregunta 3 Sin contestar Puntúa como 1</p> <p> Marcar pregunta</p> <p> Editar pregunta</p>	<p>¿Alguno de estos genes puede estar alterado en los glioblastomas?</p> <p>Seleccione una:</p> <ul style="list-style-type: none"><input type="radio"/> a. ARTX<input type="radio"/> b. NF1<input type="radio"/> c. PTEN<input type="radio"/> d. PTEN y NF1
<p>Pregunta 4 Sin contestar Puntúa como 1</p> <p> Marcar pregunta</p> <p> Editar pregunta</p>	<p>¿Qué % representan los glioblastomas con alteraciones en el gen PTEN ?</p> <p>Seleccione una:</p> <ul style="list-style-type: none"><input type="radio"/> a. < 1%<input type="radio"/> b. Entre 1 y 5 %<input type="radio"/> c. Entre 10-15 %<input type="radio"/> d. Entre 20-35 %
<p>Pregunta 5 Sin contestar Puntúa como 1</p> <p> Marcar pregunta</p> <p> Editar pregunta</p>	<p>¿Dónde se encuentran las mutaciones dentro del gen PTEN?</p> <p>Seleccione una:</p> <ul style="list-style-type: none"><input type="radio"/> a. Fuera de la región codificante<input type="radio"/> b. Pueden localizarse en cualquier región<input type="radio"/> c. Solo en la región de unión al DNA<input type="radio"/> d. En el codón de inicio, para la interrupción de la traducción

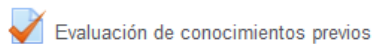
Diseño del Curso en Studium:

Se ha tratado de estructurar un curso de studium de forma muy sencilla y asequible, de manera que se va a profundizar, paulatinamente, en el tema sin que suponga un esfuerzo excesivo para el estudiante.

Herramientas bioinformáticas web públicas para el estudio de alteraciones genéticas en tumores cerebrales



¿Qué sabes sobre la genética de los tumores cerebrales?



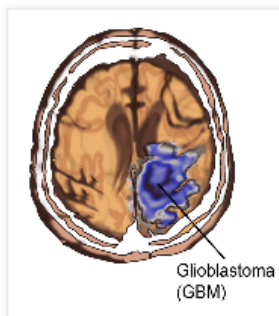
Evaluación de conocimientos previos

Para evaluar vuestros conocimientos previos al comienzo de la actividad, os rogamos que intentéis responder el cuestionario que se plantea. Por favor no busquéis información en ningún sitio para responder. No se aplicará en ningún momento la calificación obtenida, sólo se trata de saber qué información tenéis al respecto.

Debéis seleccionar la respuesta que consideréis más adecuada

Una vez cerrado el cuestionario se harán accesibles todos los recursos. Muchas gracias por vuestra colaboración

Objetivos del curso



Herramientas bioinformáticas web públicas para el estudio de alteraciones genéticas en tumores cerebrales

Ester Gangoso (1) y Arantxa Tabernero (2,3)

(1) Scottish Center for Regenerative Medicine. Edinburgh University. Edinburgh. UK (2) Instituto de Neurociencias de Castilla y León. Salamanca. Spain (3) Dpto Bioquímica y Biología molecular. USAL Salamanca. Spain

La existencia de alteraciones genéticas determinantes de enfermedades, concretamente de tumores del sistema nervioso central (SNC), es un hecho relevante y clave futura de los tratamientos personalizados. De igual manera, la realización de estudios genéticos detallados es fundamental para una prevención eficaz. Existen muchos tipos de tumores que afectan al SNC, pero nos centraremos en este curso en los más frecuentes y graves, los Glioblastomas.

Esperamos que con este recurso los estudiantes consigan manejar eficazmente las bases de datos de tumores de pacientes. Queremos que sean capaces de descubrir el potencial que tienen y cuánta información se puede sacar de ellas. Para ello, nos hemos propuesto los siguientes objetivos:

- 1) Familiarizarles con el uso de los portales web públicos más útiles que contienen bases de datos reales obtenidos de casos clínicos perfectamente referenciados.
- 2) Enseñarles a buscar y extraer, a partir de estos portales, toda la información referente tanto a los diferentes tipos de alteraciones presentes en los genes implicados en la aparición de glioblastomas como a los niveles de expresión de los mismos en las diferentes situaciones clínicas.
- 3) Entrenarles en el empleo de las herramientas bioinformáticas públicas más útiles para llevar a cabo lo expuesto en el objetivo 2.
- 4) Demostrarles que lo aprendido puede extrapolarse al estudio de cualquier otro tipo de tumor y cómo pueden llevarlo a cabo de forma autónoma en el futuro.

Introducción a la genética de los Glioblastomas

Los Glioblastomas, son los tumores cerebrales más comunes y con un peor pronóstico. En los últimos años se ha tratado de identificar cuáles son las alteraciones genéticas presentes en este tipo de tumores. En la figura 1 se recogen las más comunes. Los genes más alterados son EGFR, PDGFR, PTEN, TP53, CDKN2A y NF1. Pero no son los únicos y no todos están presentes en el mismo tumor. Por ello, la heterogeneidad de estos tumores tanto en el tipo de células que los componen como en las alteraciones genéticas presentes, hace que sea difícil encontrar nuevos tratamientos.

Tener acceso a información genética de tumores de pacientes, hace que sea muy relevante a la hora de diseñar un estudio o una estrategia terapéutica. Por ello, en este ejercicio queremos que aprendáis a usar, de manera básica, la base de datos pública, cBioportal (perteneciente, entre otros, al Memorial Sloan Kettering Cancer Center, NY, USA) En ella se podrá obtener información sobre cómo varían las alteraciones genéticas analizadas en cada paciente individual. En este portal, podéis encontrar además bases de datos de muestras de pacientes de todos los tipos de cáncer.

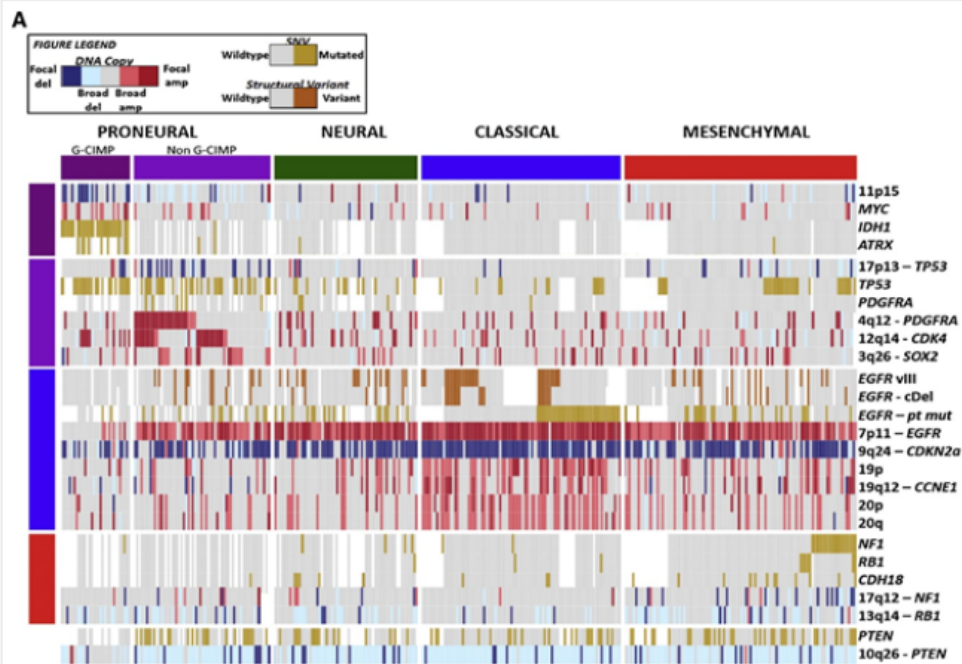


Figura 1: Alteraciones moleculares más frecuentes en Glioblastomas (Brennan et al, 2013, Cell)

EGFR (Epidermal Growth Factor Receptor)

PDGFR (Platelet Derived Growth Factor Receptor)

PTEN (Phosphatase and tensin homolog, a tumour suppressor)

TP53 (Tumour suppressor)

CDKN2A (Cyclin-dependent kinase Inhibitor 2A, a tumour suppressor)

NF1 (Neurofibromin, a tumour suppressor)

Tras la Introducción General y la enumeración de objetivos se procede a explicar la forma de adquirir la información y procesarla con la ayuda de los recursos WEB públicos. Para ello se ha elaborado un sencillo tutorial explicativo que les guiará paso a paso.

¿Cómo y en que WEB busco la información sobre el tumor que quiero estudiar?

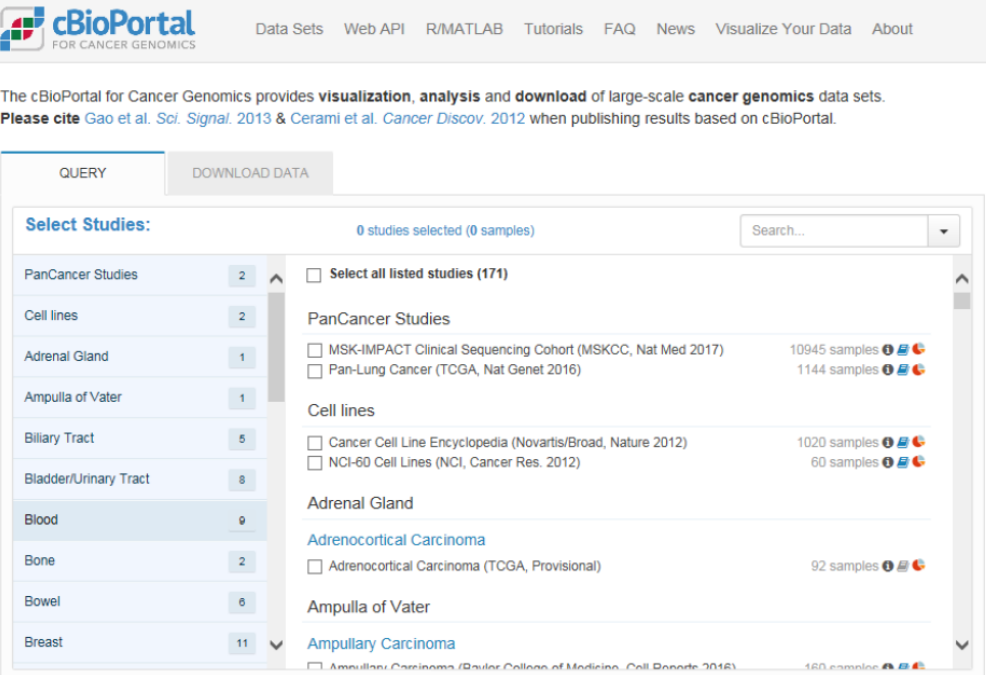


Tutorial sobre el uso de la WEB cBioportal 708.4KB documento PDF

Instrucciones para saber el tipo de información que puedo encontrar y como hacerlo.

Tutorial sobre el uso de la WEB cBioPortal

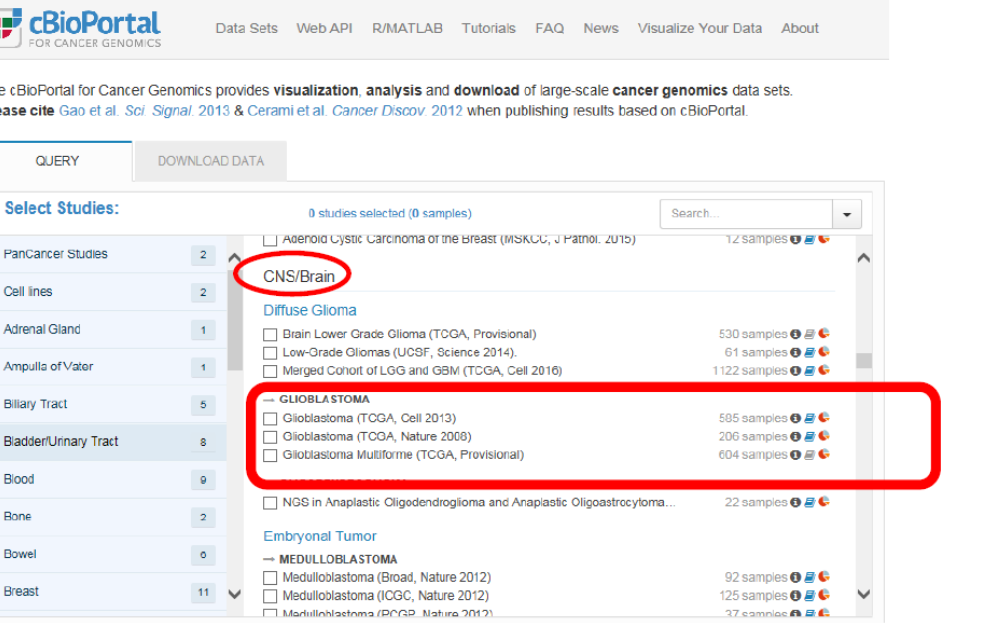
1- Buscar bases de datos confeccionadas a partir de muestras de pacientes, para ello ir por ejemplo a cBioportal, portal especializado en cáncer. Entrar en la web: <http://www.cbioportal.org/>



The screenshot shows the cBioPortal interface. At the top, there is a navigation bar with the cBioPortal logo and links for Data Sets, Web API, R/MATLAB, Tutorials, FAQ, News, Visualize Your Data, and About. Below the navigation bar, a message states: "The cBioPortal for Cancer Genomics provides visualization, analysis and download of large-scale cancer genomics data sets. Please cite Gao et al. *Sci. Signal.* 2013 & Cerami et al. *Cancer Discov.* 2012 when publishing results based on cBioPortal."

The main content area is divided into two tabs: "QUERY" and "DOWNLOAD DATA". Under the "QUERY" tab, there is a "Select Studies:" section. It shows a list of studies on the left and a detailed view of selected studies on the right. The left sidebar lists various cancer types with their respective sample counts: PanCancer Studies (2), Cell lines (2), Adrenal Gland (1), Ampulla of Vater (1), Biliary Tract (5), Bladder/Urinary Tract (8), Blood (9), Bone (2), Bowel (6), and Breast (11). The right pane shows a list of studies under the heading "0 studies selected (0 samples)". The studies listed include: MSK-IMPACT Clinical Sequencing Cohort (MSKCC, Nat Med 2017) with 10945 samples; Pan-Lung Cancer (TCGA, Nat. Genet 2016) with 1144 samples; Cancer Cell Line Encyclopedia (Novartis/Broad, Nature 2012) with 1020 samples; NCI-60 Cell Lines (NCI, Cancer Res. 2012) with 60 samples; Adrenocortical Carcinoma (TCGA, Provisional) with 92 samples; and Ampullary Carcinoma (Baylor College of Medicine, Cell Reports 2016) with 160 samples.

2- Desplazarse con el cursor, hasta encontrar CNS/Brain. En este apartado se encuentran diferentes cánceres del sistema nervioso central. Nosotros nos centraremos en los glioblastomas. Como podéis observar hay tres bases de datos referentes a este tumor.



The screenshot shows the cBioPortal interface. At the top, there is a navigation bar with the cBioPortal logo and links for Data Sets, Web API, R/MATLAB, Tutorials, FAQ, News, Visualize Your Data, and About. Below the navigation bar, a message states: "The cBioPortal for Cancer Genomics provides visualization, analysis and download of large-scale cancer genomics data sets. Please cite Gao et al. *Sci. Signal.* 2013 & Cerami et al. *Cancer Discov.* 2012 when publishing results based on cBioPortal."

The main content area is divided into two tabs: "QUERY" and "DOWNLOAD DATA". Under the "QUERY" tab, there is a "Select Studies:" section. It shows a list of studies on the left and a detailed view of selected studies on the right. The left sidebar lists various cancer types with their respective sample counts: PanCancer Studies (2), Cell lines (2), Adrenal Gland (1), Ampulla of Vater (1), Biliary Tract (5), Bladder/Urinary Tract (8), Blood (9), Bone (2), Bowel (6), and Breast (11). The right pane shows a list of studies under the heading "0 studies selected (0 samples)". The studies listed include: Adenoid Cystic Carcinoma of the breast (MSKCC, J. Pathol. 2015) with 12 samples; Diffuse Glioma; Brain Lower Grade Glioma (TCGA, Provisional) with 530 samples; Low-Grade Gliomas (UCSF, Science 2014) with 61 samples; Merged Cohort of LGG and GBM (TCGA, Cell 2016) with 1122 samples; **GLIOBLASTOMA** (highlighted with a red rectangle): Glioblastoma (TCGA, Cell 2013) with 585 samples; Glioblastoma (TCGA, Nature 2008) with 206 samples; Glioblastoma Multiforme (TCGA, Provisional) with 604 samples; NGS in Anaplastic Oligodendroglioma and Anaplastic Oligoastrocytoma... with 22 samples; Embryonal Tumor; and **MEDULLOBLASTOMA** (highlighted with a red rectangle): Medulloblastoma (Broad, Nature 2012) with 92 samples; Medulloblastoma (ICGC, Nature 2012) with 125 samples; Medulloblastoma (BCSP, Nature 2012) with 37 samples.

3- Seleccionar la primera opción Glioblastomas (TCGA, Cell 2013). Dejar las funciones que selecciona el programa por defecto en cuanto al perfil genómico. Ahora vamos a introducir uno de los genes más alterados en los glioblastomas, *Pten*. Hay que introducirlo en mayúsculas. Después seleccionar Submit Query.

Se pueden buscar varios genes a la vez, para tener más datos. Pero en este caso, para hacerlo más sencillo, lo haremos sólo con uno.

The screenshot shows the cBioPortal search interface. The browser address bar displays www.cbioportal.org/index.do. The search results are filtered to 'CNS/Brain' and 'Diffuse Glioma'. Under 'Diffuse Glioma', the following options are listed:

- Brain Lower Grade Glioma (TCGA, Provisional) - 530 samples
- Low-Grade Gliomas (UCSF, Science 2014) - 61 samples
- Merged Cohort of LGG and GBM (TCGA, Cell 2016) - 1122 samples
- GLIOBLASTOMA**
- Glioblastoma (TCGA, Cell 2013) - 585 samples
- Glioblastoma (TCGA, Nature 2008) - 206 samples
- Glioblastoma Multiforme (TCGA, Provisional) - 604 samples
- OLIGODENDROGLIOMA
- NGS in Anaplastic Oligodendroglioma and Anaplastic Oligoastrocytomas tumors (MSK, Neu...) - 22 samples
- Embryonal Tumor
- MEDULLOBLASTOMA

 Below the search results, the 'Select Genomic Profiles' section has 'Mutations' and 'Putative copy-number alterations from GISTIC' checked. The 'Select Patient/Case Set' dropdown is set to 'Tumor Samples with sequencing and CNA data (281)'. The 'Enter Gene Set' dropdown is set to 'User-defined List' and contains the text 'PTEN'. A green message below the gene set says 'All gene symbols are valid.' The 'Submit Query' button is circled in red.

4- En la pestaña **OncoPrint**, podéis obtener información del tipo de alteraciones genéticas en las muestras de esta base de datos. *Pten* se encuentra alterado en el 31 % de las muestras (86/281). Cada rectángulo, representa una muestra. En la leyenda de colores, podéis ver qué tipo de alteraciones están presentes.

The screenshot shows the cBioPortal OncoPrint view for 'Glioblastoma (TCGA, Cell 2013)'. The page title is 'Glioblastoma (TCGA, Cell 2013)' and the subtitle is 'Tumor Samples with sequencing and CNA data (281 samples) / 1 Genes'. The 'Gene Set / Pathway is altered in 86 (30.6%) of queried samples' is displayed. The 'OncoPrint' tab is selected, and the 'Case Set' is 'Tumor Samples with sequencing and CNA data (281 patients / 281 samples)'. The 'PTEN' gene is highlighted, showing that it is altered in 86 (31%) of 281 sequenced cases/patients (281 total). The legend for 'Genetic Alteration' includes:

- Deep Deletion (blue)
- Inframe Mutation (unknown significance) (orange)
- Missense Mutation (putative driver) (green)
- Missense Mutation (unknown significance) (light green)
- No alterations (grey)
- Truncating Mutation (putative driver) (black)

 The heatmap shows 281 samples, with 86 samples having alterations in the PTEN gene.

5- En la pestaña **Mutations**, podéis observar donde aparecen las mutaciones más frecuente. Como se puede comprobar, las mutaciones se encuentran distribuidas a lo largo del gen. A la derecha, resaltado en rojo, se describe el tipo de mutaciones presente: Missense¹, truncating², inframe³ or other.



1.-Mutación puntual no sinónima en la cual se produce un cambio en un único nucleótido, provocando la aparición de un codón que codifica para un aminoácido diferente o 2.-por un codón de terminación. Esto puede provocar un producto proteico incapaz de cumplir su función. Mutaciones de este tipo son responsables por ejemplo de enfermedades tales como la anemia de células falciformes, la epidermólisis bullosa o la esclerosis lateral amiotrófica. 3.- A veces las mutaciones implican el cambio en la fase de lectura (+1)

Evaluación de la actividad mediante Cuestionario Final

De acuerdo con las observaciones realizadas a lo largo del desarrollo de las diferentes actividades, se elaborará un cuestionario final más complejo con el que se evaluarán las competencias adquiridas. Posteriormente se pasará a los alumnos una encuesta configurable a través de Studium para que emitan su valoración del recurso y propongan mejoras.