



Estudio personalizado de datos ómicos en paragangliomas y feocromocitomas

Silvia Noemí Vázquez Cuesta

Máster universitario en Bioinformática y bioestadística UOC-UB
Área 3, subárea 5.

Laia Bassanganyas Bars
Ferrán Prados Carrasco

12 de Junio de 2019



Esta obra está sujeta a una licencia de Reconocimiento-NoComercial-SinObraDerivada [3.0 España de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/3.0/es/)

B) GNU Free Documentation License (GNU FDL)

Copyright © 2019 SILVIA NOEMÍ VÁZQUEZ CUESTA.

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.

A copy of the license is included in the section entitled "GNU Free Documentation License".

C) Copyright

© (Silvia Noemí Vázquez Cuesta)

Reservados todos los derechos. Está prohibido la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidos la impresión, la reprografía, el microfilme, el tratamiento informático o cualquier otro sistema, así como la distribución de ejemplares mediante alquiler y préstamo, sin la autorización escrita del autor o de los límites que autorice la Ley de Propiedad Intelectual.

FICHA DEL TRABAJO FINAL

Título del trabajo:	<i>Estudio personalizado de datos ómicos en paragangliomas y feocromocitomas</i>
Nombre del autor:	<i>Silvia Noemí Vázquez Cuesta</i>
Nombre del consultor/a:	<i>Laia Bassanganyas Bars</i>
Nombre del PRA:	<i>Ferrán Prados Carrasco</i>
Fecha de entrega (mm/aaaa):	06/2019
Titulación:	<i>Máster universitario en Bioinformática y bioestadística UOC-UB</i>
Área del Trabajo Final:	<i>M0.179 TFM-Estadística y Bioinformática 3</i>
Idioma del trabajo:	<i>Español</i>
Palabras clave	<i>Paragangliomas, feocromocitomas, ómica.</i>
<p>Resumen del Trabajo (máximo 250 palabras): <i>Con la finalidad, contexto de aplicación, metodología, resultados y conclusiones del trabajo.</i></p>	
<p>Introducción: Los feocromocitomas y los paragangliomas son tumores neuroendocrinos que se consideran fuertemente influenciados por la genética, ya que aproximadamente el 40% de los pacientes sufre alguna mutación en la línea germinal.</p> <p>Además de las mutaciones en la línea germinal, se han encontrado mutaciones somáticas, patrones de expresión génica y de metilación que han permitido su clasificación en dos grandes grupos.</p> <p>Aun así, estos estudios se han llevado a cabo en cohortes y era necesaria una aproximación más personalizada para cada tipo de paciente.</p> <p>Materiales y métodos: Para este estudio se han utilizado datos del proyecto “<i>The Cancer Genome Atlas</i>” con 179 casos entre feocromocitomas y paragangliomas. Se han analizado datos de alteraciones en el número de copia, de expresión génica, de mutaciones y de metilación. Se han utilizado aplicaciones web como CNApp para el análisis de datos de CNAs y el resto de análisis se han realizado principalmente en R con la librería TCGAbiolinks.</p>	

Resultados: Integrando los distintos tipos de datos se han obtenido resultados interesantes en cada grupo tumoral, algunos datos aparecen comunes en todos los grupos, como alteraciones en funciones de adhesión celular o diferenciación celular, pero en algunos grupos aparecen alteraciones en funciones que pueden estar más relacionadas con la agresividad de dicho tumor en concreto, como procesos de angiogénesis y señalización celular.

Conclusiones: En este trabajo se ha visto que, aunque estos tumores pueden ser muy parecidos en algunos aspectos, en otros difieren y esos datos son importantes tanto para la prognosis como para posibles tratamientos.

Abstract (in English, 250 words or less):

Introduction: Pheochromocytomas and paragangliomas are neuroendocrine tumors that are considered strongly influenced by genetics, since approximately 40% of patients suffer some mutation in the germ line.

In addition to mutations in the germ line, somatic mutations, gene expression and methylation patterns have been found, which have allowed their classification into two large groups.

Even so, these studies have been carried out in cohorts and a more personalized approach was needed for each type of patient.

Materials and methods: The data analyzed in this study were from "The Cancer Genome Atlas" project. There were data from 179 cases between pheochromocytomas and paragangliomas. We analyzed data of copy number alterations, gene expression, mutations and methylation. Web applications such as CNApp have been used for the analysis of CNA data and the rest of the data were analyzed with R and the TCGAbiolinks package.

Results: With the integration of the different types of data, results have been obtained in each tumor group, some data appear common in all groups, such as functions of alterations in cell adhesion or cell differentiation, but in some groups appear alterations in functions that may be more related with the aggressiveness of one tumor in particular, as angiogenesis processes and cell signaling.

Conclusions: In this work we have seen that, although these tumors can be very similar in some ways, in others they differ, and these data are important both for the prognosis and for future treatments.

Índice

1. Introducción	1
1.1 Contexto y justificación del Trabajo	1
1.2 Objetivos del Trabajo	3
1.3 Enfoque y método seguido	3
1.4 Planificación del Trabajo	5
1.5 Breve resumen de productos obtenidos	9
1.6 Breve descripción de los otros capítulos de la memoria	9
2. Materiales y métodos:.....	11
2.1 Datos utilizados:	11
2.2 Herramientas para el análisis de datos:	12
3. Resultados:.....	13
3.1. Análisis de datos de CNAs:	15
3.2. Datos de metilación:	22
3.3. Datos de mutaciones:	25
3.4. Datos de expresión génica:	32
5. Discusión:	38
3. Conclusiones	41
4. Glosario	43
5. Bibliografía.....	44
6. Anexos.....	49

Lista de figuras

Figura 1. Imagen de la localización de los feocromocitomas y los paragangliomas.[3]	1
Figura 2. Diagrama de Gantt con la planificación mensual de las tareas.	7
Figura 3. Diagrama de Gantt con la planificación semanal de las tareas.	8
Figura 4. Frecuencia de CNAs de todas las muestras juntas.	16
Figura 5. Frecuencia de CNAs de las muestras agrupadas en (de izquierda a derecha) Paragangliomas extra-adrenales, Paragangliomas y feocromocitomas.	16
Figura 6. Comparación de CNAs por regiones entre los distintos grupos diagnósticos.	17
Figura 7. Puntuaciones de los distintos CNAs según el grupo diagnóstico.	18
Figura 8. Gráficos de barras con los procesos biológicos más representados en PhM y PhNOS según el tipo de CNA.	19
Figura 9. Gráficos de barras con los procesos biológicos más representados en PaM y PaNOS según el tipo de CNA.	20
Figura 10. Gráficos de barras con los procesos biológicos más representados en EAPaM y EAPaNOS según el tipo de CNA.	21
Figura 11. Distribución de las medias de metilación entre los distintos grupos diagnósticos.	22
Figura 12. Distribución de las medias de metilación en PhM y PhNOS según los distintos tipos de CNAs.	24
Figura 13. Gráficos volcano con diferencias de metilación entre el tercer y el primer cuartil de los distintos tipos de CNAs en PhNOS.	24
Figura 14. Resumen de datos de mutaciones de todas las muestras juntas.	26
Figura 15. Rutas oncogénicas afectadas por mutaciones en sus genes de todo el conjunto de muestras del estudio.	27
Figura 16. Datos de mutaciones de PhM en base a los cuartiles de los distintos CNAs.	29
Figura 17. Datos de mutaciones de PhNOS en base a los cuartiles de los distintos CNAs.	30
Figura 18. Gráfico que integra todos los datos de las muestras con los genes que se han encontrado más frecuentemente mutados.	31
Figura 19. Relación de las puntuaciones de MATH y de BCS en PhM.	32

Lista de tablas

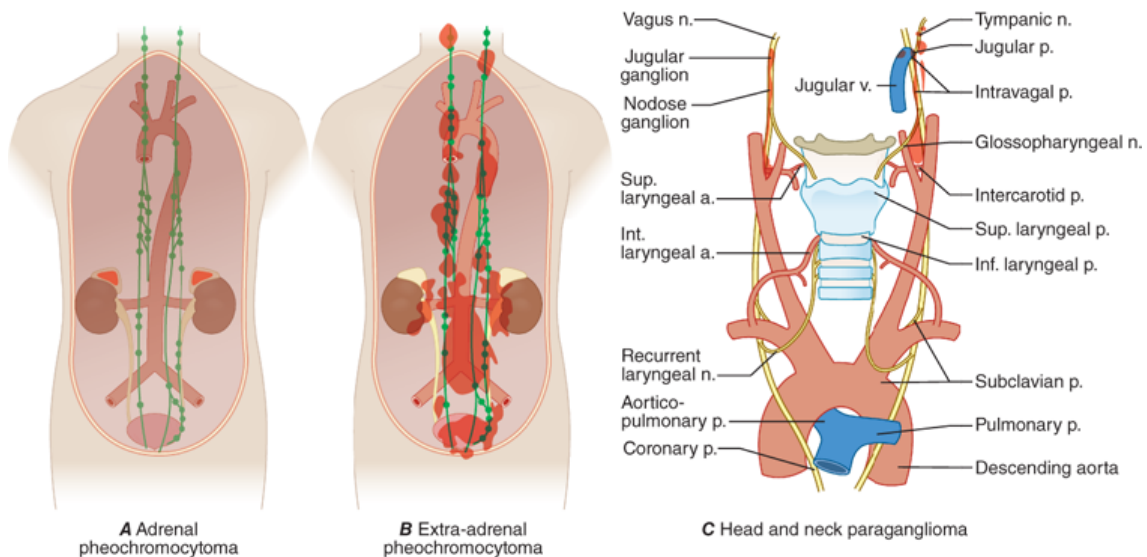
Tabla 1. Descripción de la planificación del trabajo según las tareas a realizar.	6
Tabla 2. Datos demográficos de los pacientes incluidos en el estudio.	14
Tabla 3. Resumen de datos de las muestras que se van a analizar.	15
Tabla 4. Resultados de análisis de procesos biológicos en PhM.	33
Tabla 5. Resultados de análisis de procesos biológicos en PhNOS.	34
Tabla 6. Resultados de análisis de procesos biológicos en PaM.	35
Tabla 7. Resultados de análisis de procesos biológicos en EAPaM.	36
Tabla 8. Resultados de análisis de procesos biológicos en EAPaNOS.	37

1. Introducción

1.1 Contexto y justificación del Trabajo

Los feocromocitomas y los paragangliomas (PPGL) son tumores neuroendocrinos que se originan a partir de células cromafines de la médula adrenal y de progenitores de la cresta neural de paraganglios extra-adrenales respectivamente [1]. El 90% de este tipo de tumores son feocromocitomas que se localizan en la médula adrenal. De los paragangliomas, el 85% se localizan en el abdomen, el 12% se encuentran en el tórax, y en cabeza y cuello se localizan solamente el 3% [2]. (Figura 1) Estos últimos son los únicos que no producen catecolaminas.

Muchos de estos tumores son benignos y se curan mediante cirugía, pero hay otros casos en los que la morbilidad y la mortalidad son más altas debido a la agresividad de determinados tumores. [1]



Source: D. L. Kasper, A. S. Fauci, S. L. Hauser, D. L. Longo, J. L. Jameson, J. Loscalzo: Harrison's Principles of Internal Medicine, 19th Edition. www.accessmedicine.com
Copyright © McGraw-Hill Education. All rights reserved.

Figura 1. Imagen de la localización de los feocromocitomas y los paragangliomas.[3]

Los PPGL se consideran fuertemente influenciados por la genética, ya que aproximadamente el 40% de los pacientes sufre alguna mutación en la línea germinal [4].

Por ello las guías internacionales recomiendan el genotipado de genes susceptibles de sufrir mutaciones en estos pacientes, ya que los resultados pueden dar información sobre la prognosis de la enfermedad, el tratamiento, etcétera [5-7].

Además de mutaciones, en los PPGL se han encontrado distintos patrones de metilación y de expresión génica que han permitido clasificarlos en dos grandes grupos [8].

Se ha descrito que los tumores pertenecientes al grupo de hipometilación cuentan un número mayor de alteraciones en el número de copia, sin embargo los clasificados dentro del grupo de hipermetilación cuentan con más inestabilidad cromosómica. [9, 10]

Al mismo tiempo, los perfiles de metilación también han sido de utilidad para estimar el pronóstico de la enfermedad, permitiendo clasificar a los pacientes según el riesgo de desarrollar metástasis. [11]

Otro dato importante a tener en cuenta en el estudio de tumores son las alteraciones en el número de copia (CNAs de sus siglas del inglés copy number alterations).

Se han descrito unos nuevos marcadores de CNAs que los clasifican en CNAs focales, para alteraciones que afectan a menos de un 50% del brazo de un cromosoma, y CNAs extensas o anchas (broad) cuando afectan a un cromosoma entero o a un brazo de un cromosoma. [12]

Según Teresa Davoli et al. “Los distintos tipos de CNAs pueden predecir firmas de expresión génica que están asociadas con dos características distintivas del cáncer: la proliferación celular y la huida inmunológica.”

Según sus datos existe una correlación positiva entre puntuaciones altas de CNAs y la expresión de genes relacionados con regulación del ciclo celular y proliferación.[12]

Los CNAs focales estarían más relacionados con genes que intervienen en el ciclo celular y la proliferación celular, mientras que los CNAs anchos estarían relacionados con procesos de la respuesta inmune.[12]

Por todos los datos anteriormente descritos sobre la presencia de alteraciones importantes de material genético, se pensó que sería importante realizar un estudio más personalizado que se centrara en cada tipo de tumor en particular, por si se encontrasen diferencias relevantes entre ellos y esto pudiese ofrecer información de cara tanto al tratamiento como para el pronóstico de la enfermedad.

La mayoría de los estudios acerca de estos tumores son estudios de cohortes que han servido para hacer distintas clasificaciones de los tumores, pero con este trabajo se pretende abordar una aproximación más personalizada de este tipo de análisis, ya que parece que no hay demasiados trabajos enfocados desde este punto de vista.

1.2 Objetivos del Trabajo

Los objetivos principales de este trabajo con sus objetivos secundarios han sido:

1. El primer objetivo principal ha sido la exploración de datos ómicos en los paragangliomas y los feocromocitomas. Sus objetivos secundarios han sido:
 - 1.1 Explorar datos de CNAs
 - 1.2 Explorar datos de metilación
 - 1.3 Explorar datos de expresión génica
 - 1.4 Explorar datos de mutaciones

2. El segundo objetivo principal del trabajo ha sido el estudio de la correlación e interpretación de los datos ómicos.
 - 2.1 Analizar de manera integrada los datos ómicos
 - 2.2 Correlacionar los datos ómicos con los datos clínicos
 - 2.3 Interpretar los datos ómicos

1.3 Enfoque y método seguido

Para la realización de este estudio se han utilizado datos públicos de paragangliomas y feocromocitomas pertenecientes al programa “The Cancer Genome Atlas”

Los datos analizados son los siguientes:

- Datos CNAs de microarray Affimetrix SNP6.0:
[gdac.broadinstitute.org PCPG.Merge_snp_genome_wide_snp_6_broad.mit.edu_Level_3_segmented_scna_minus_germline_cnv_hg18_seg.Level_3.2016012800.0.0.tar.gz](https://gdac.broadinstitute.org/PCPG/Merge_snp_genome_wide_snp_6_broad.mit.edu_Level_3_segmented_scna_minus_germline_cnv_hg18_seg.Level_3.2016012800.0.0.tar.gz)
- Datos de Metilación de Illumina Infinium HM450:
[gdac.broadinstitute.org PCPG.Merge_methylation_humanmethylation450_jhu usc.edu_Level_3_within_bioassay_data_set_function_data.Level_3.2016012800.0.0.tar.gz](https://gdac.broadinstitute.org/PCPG/Merge_methylation_humanmethylation450_jhu usc.edu_Level_3_within_bioassay_data_set_function_data.Level_3.2016012800.0.0.tar.gz)
- Datos de expresión de RNAseq de Illumina:
[gdac.broadinstitute.org PCPG.Merge_rnaseqv2_illuminahiseg_rnaseqv2_unc.edu_Level_3_RSEM_genes_normalized_data.Level_3.2016012800.0.0.tar.gz](https://gdac.broadinstitute.org/PCPG/Merge_rnaseqv2_illuminahiseg_rnaseqv2_unc.edu_Level_3_RSEM_genes_normalized_data.Level_3.2016012800.0.0.tar.gz)
- Datos de mutaciones de SureSelect v2 de Illumina:
[gdac.broadinstitute.org PCPG.Mutation_Packager_Oncotated_Calls.Level_3.2016012800.0.0.tar.gz](https://gdac.broadinstitute.org/PCPG/Mutation_Packager_Oncotated_Calls.Level_3.2016012800.0.0.tar.gz)
- Datos clínicos:
[gdac.broadinstitute.org PCPG.Clinical_Pick_Tier1.Level_4.2016012800.0.0.tar.gz](https://gdac.broadinstitute.org/PCPG/Clinical_Pick_Tier1.Level_4.2016012800.0.0.tar.gz)

La metodología a seguir fue la siguiente:

- Primero realizar un análisis descriptivo de las muestras del estudio para tener una visión general de sus datos.
- Clasificarlas en pequeños grupos basados en características clínicas.
- De cada grupo analizar por separado los datos de CNAs, metilación, mutaciones y expresión génica. Primero se realizaron los estudios de CNAs en todos los grupos, luego los de metilación, luego los de mutaciones y luego los de expresión génica.
- Integrar los resultados de todos esos datos en cada grupo y buscar su significado biológico.
- Otra manera de hacerlo sería ir haciendo los diferentes análisis en un grupo de muestras, luego hacer todos los análisis en otro grupo de muestras y así sucesivamente.

- Se pensó que la primera estrategia podía mejorar la eficiencia del proceso de análisis.
- Para el análisis de datos de CNAs se utilizó la aplicación web desarrollada en R CNApp tool. [13]
- Para el análisis de datos de metilación se planteó utilizar los paquetes methyAnalysis [14] y methylKit [15]
- Para los análisis de mutaciones y expresión génica se utilizarían distintos paquetes de Bioconductor. [16]
- La decisión final fue analizar los datos de CNAs con CNApp tool, utilizar el paquete TCGAbiolink para los datos de metilación y de expresión génica por estar enfocado directamente a este formato de datos. Para los datos de mutaciones se utilizó el paquete maftools. [17]

1.4 Planificación del Trabajo

Para la realización de este trabajo se realizó la planificación que queda plasmada en la tabla 1.

Se realizó una clasificación de tareas en la que el objetivo general 1 corresponde al grupo de tareas 1.3 Análisis de datos. Fase 1, con sus objetivos específicos englobados dentro de ella.

El objetivo general 2 corresponde al grupo de tareas 1.4 Análisis de datos. Fase 2, con sus objetivos específicos englobados en este grupo.

1	TFM	Tipo	Inicio	Fin
		proyecto	20/2/19	20/6/19
1.1	Estudio preliminar	grupo	20/2/19	4/3/19
1.1.1	Estudio de la bibliografía relacionada con los tipos de tumores	tarea	20/2/19	23/2/19
1.1.2	Estudio de la bibliografía relacionada con el análisis de datos de CNAs	tarea	24/2/19	27/2/19
1.1.3	Estudio de la bibliografía relacionada con análisis de datos de metilación.	tarea	26/2/19	1/3/19
1.1.4	Elaboración de la PEC 0	tarea	1/3/19	3/3/19
1.1.5	Entrega de la PEC 0	hito	4/3/19	4/3/19
1.2	PEC 1	grupo	5/3/19	18/3/19
1.2.1	Repaso de la bibliografía y estudio de nueva bibliografía sobre metodología	tarea	5/3/19	8/3/19
1.2.2	Definición específica de objetivos	tarea	9/3/19	10/3/19
1.2.3	Elaboración del diagrama de Gantt	tarea	11/3/19	13/3/19
1.2.4	Elaboración de la memoria de la PEC 1	tarea	14/3/19	17/3/19
1.2.5	Entrega de la PEC 1	hito	18/3/19	18/3/19
1.3	Análisis de datos. Fase 1	grupo	19/3/19	24/4/19
1.3.1	Obtención de los datos	tarea	19/3/19	20/3/19
1.3.2	Análisis de datos de los CNAs	tarea	20/3/19	24/3/19
1.3.3	Análisis de datos de metilación	tarea	25/3/19	30/3/19
1.3.4	Análisis de datos de expresión génica	tarea	31/3/19	4/4/19
1.3.5	Análisis de datos de mutaciones	tarea	5/4/19	11/4/19
1.3.6	Elaboración de la PEC 2	tarea	12/4/19	23/4/19
1.3.7	Entrega de la PEC 2	hito	24/4/19	24/4/19
1.4	Análisis de datos. Fase 2	grupo	26/4/19	20/5/19
1.4.1	Estudio relacional de los datos obtenidos	tarea	26/4/19	4/5/19
1.4.2	Estudio de la bibliografía sobre implicaciones de los datos obtenidos	tarea	5/5/19	10/5/19
1.4.3	Elaboración de la PEC 3	tarea	11/5/19	19/5/19
1.4.4	Entrega de la PEC 3	hito	20/5/19	20/5/19
1.5	Cierre de la memoria	grupo	21/5/19	5/6/19
1.5.1	Finalización de la memoria	tarea	21/5/19	4/6/19
1.5.2	Entrega de la memoria. PEC 4	hito	5/6/19	5/6/19
1.6	Elaboración de la presentación	grupo	6/6/19	13/6/19
1.6.1	Elaboración de la presentación	tarea	6/6/19	12/6/19
1.6.2	Entrega de la presentación. PEC 5a	hito	13/6/19	13/6/19
1.7	Defensa de la memoria	grupo	20/6/19	20/6/19
1.7.1	Defensa de la memoria. PEC 5b	hito	20/6/19	20/6/19

Tabla 1. Descripción de la planificación del trabajo según las tareas a realizar.

Se realizó un diagrama de Gantt con la planificación de las tareas, se generaron dos vistas de dicho diagrama, una más general con la planificación en vista mensual (Figura 2) y otro con más detalle de las semanas (Figura 3). En esta planificación no se tuvieron en cuenta los días no laborables para excluirlos de la planificación porque eran días en los que si se iba a seguir con el trabajo. El Diagrama de Gantt se realizó con la aplicación web <https://app.teamgantt.com> [18]

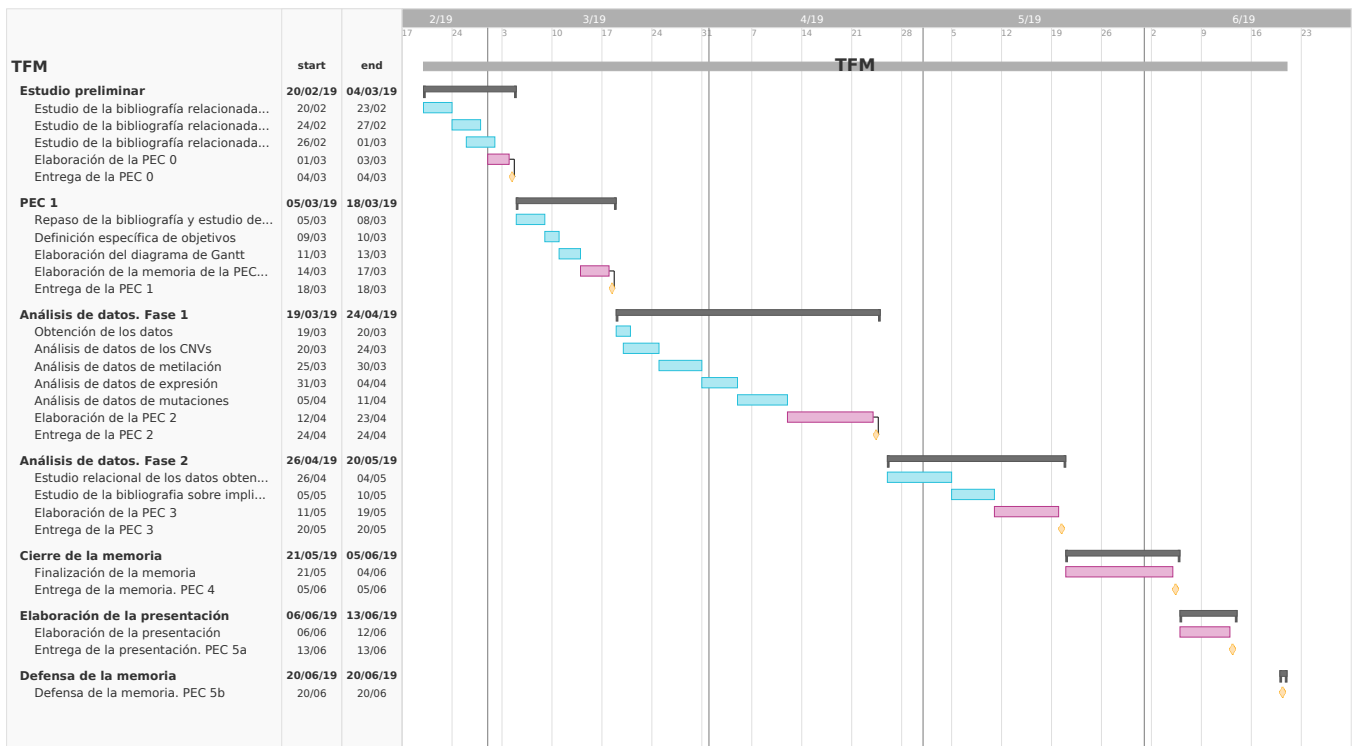


Figura 2. Diagrama de Gantt con la planificación mensual de las tareas.



Figura 3. Diagrama de Gantt con la planificación semanal de las tareas.

Los hitos que han sido marcados para este proyecto son las entregas de las pruebas de evaluación (PEC), puesto que marcan etapas del proyecto que deben haberse realizado y sin su cumplimiento sería muy difícil sacar adelante el proyecto. Estas son:

- Entrega de la PEC 0, con el resumen del proyecto
- Entrega de la PEC 1, con la planificación.
- Entrega de la PEC 2, con el objetivo 1 cumplido
- Entrega de la PEC 3, con el objetivo 2 cumplido

- Entrega de la PEC 4, cierre de la memoria del proyecto
- Entrega de la PEC 5, con la entrega de la presentación y con la defensa pública.

Los recursos necesarios para realizar este trabajo son un ordenador con unos requisitos mínimos de memoria de 8GB a 1600MHz DDR3 y procesador de 1,6 GHz Intel Core i5. A parte de esto se necesita tener conexión a internet.

1.5 Breve resumen de productos obtenidos

Como producto de este proyecto se esperaba obtener la siguiente información acerca de los paragangliomas y los feocromocitomas:

- Datos de CNAs
- Datos de metilación
- Datos de expresión génica
- Datos de mutaciones
- La integración de los datos ómicos mencionados anteriormente
- La correlación de los datos ómicos con los datos clínicos
- Significado biológico de los anteriores resultados obtenidos

1.6 Breve descripción de los otros capítulos de la memoria

Los siguientes capítulos del trabajo son:

Materiales y métodos: donde se explica con detalle la naturaleza de los datos utilizados para la realización de este trabajo. También se detallan las herramientas utilizadas para la realización de todos los análisis y los test estadísticos que se hayan aplicado.

Resultados: en este capítulo se exponen los resultados obtenidos de cada tipo de análisis y de la integración entre ellos.

Discusión: donde se discute y compara los resultados obtenidos con lo publicado previamente en la bibliografía.

Conclusiones: en este capítulo se detallan las conclusiones que se han podido obtener de este estudio.

Glosario: en esta sección se encuentran las definiciones de términos técnicos utilizados en el trabajo.

Bibliografía: donde encontramos las referencias bibliográficas utilizadas en este trabajo.

Anexo: donde encontramos información adicional que no se haya explicado en detalle en los demás capítulos para no dificultar la lectura del trabajo.

2. Materiales y métodos:

2.1 Datos utilizados:

Para este trabajo se han utilizado datos públicos de paragangliomas y feocromocitomas pertenecientes al programa “The Cancer Genome Atlas” (TCGA) (<https://www.cancer.gov/tcga>). La referencia de los datos de este proyecto es TCGA-PCPG.

Esta colección cuenta con 179 casos de los que se obtuvieron muestras de tejido tumoral, muestras de tejido normal y muestras de sangre.

El estudio pasó los comités de ética de los distintos centros participantes y las muestras fueron obtenidas tras la firma del consentimiento informado de los pacientes.

Los métodos seguidos para el procesamiento de las muestras y la obtención de los datos ómicos son los descritos previamente por Fishbein et al. [19]

Los datos ómicos con los que se ha trabajado han sido:

Datos de CNAs obtenidos a partir de microarray Affimetrix SNP6.0. Los datos de CNAs descargados son los de procesamiento de nivel 3 que en este caso corresponden a datos de segmentación basados en sondas del genoma de referencia hg19.

Datos de metilación obtenidos del array de Illumina Infinium HM450. También se descargaron los datos de nivel 3, con los valores beta ya obtenidos.

Datos de expresión génica obtenidos a partir de RNAseq de Illumina ya normalizados mediante RSEM.

Datos de mutaciones de SureSelect v2 de Illumina. En este caso es un archivo MAF con datos de mutaciones somáticas.

Datos clínicos de las muestras y de los pacientes participantes en el estudio.

Como el objeto de este estudio era la medicina personalizada se realizó una clasificación de los datos en seis grupos según el diagnóstico que aparecía en los datos clínicos. Una vez hecho esto se realizaron análisis

generales de todas las muestras y luego divididas en estos seis grupos para ver si existían diferencias entre ellos.

2.2 Herramientas para el análisis de datos:

Para realizar los distintos tipos de análisis de este trabajo se han utilizado diversos recursos como son:

La página del GDC (<https://portal.gdc.cancer.gov/exploration>)[16] para realizar la primera exploración de los datos.

La versión local de la aplicación web CNApp para los análisis de CNAs. [13]

Se utilizó el software estadístico R versión 3.5.3 (2019-03-11) y el entorno RStudio versión 1.1.463 para la mayoría de análisis de datos de expresión génica, mutaciones y datos de metilación. [20, 21]

Los paquetes que se utilizaron en R fueron: maftools [17] para los análisis de mutaciones, TCGABiolinks y TCGAWorkflow tanto para la descarga de los datos desde TCGA como para los análisis de expresión génica y metilación. [22, 23]

También se utilizaron varios paquetes para el resumen de datos, creación y manejo de tablas y gráficos, como son summarytools, SummarizedExperiment, dplyr, ggpubr, gplots, ggplot2, RColorBrewer, sqldf, knitr, data.table, Hmisc, DT, magrittr, qwraps2, GenomicRanges. [24-38]

Para realizar comparaciones centradas en algunos genes de interés y poder ver su comportamiento en los distintos diagnósticos, se utilizó la herramienta web cBioportal versión 2.2.1.[39, 40]

Además se utilizó la herramienta web Gene Ontology [41-43] para corroborar los datos obtenidos con la función de Enrichment Analysis de la librería TCGABiolinks.

3. Resultados:

En este estudio se ha realizado un análisis de datos ómicos a partir de datos de 179 casos de paragangliomas y feocromocitomas procedentes del TCGA.

Lo primero que se hizo fue un análisis descriptivo de los datos, tanto clínicos (Tabla 2), como de los datos de las muestras (Tabla 3), subdividiendo este tipo de tumores en seis grupos según el diagnóstico primario.

Estos seis grupos diagnósticos son:

- Feocromocitoma maligno (PhM)
- Feocromocitoma no especificado de otra manera (NOS, de las siglas en inglés “Not otherwise specified”), (PhNOS)
- Paraganglioma maligno (PaM)
- Paraganglioma NOS (PaNOS)
- Paraganglioma extra-adrenal maligno (EAPaM)
- Paraganglioma extra-adrenal NOS (EAPaNOS).

Como se puede observar en la tabla 2 el diagnóstico que cuenta con más muestras es el PhNOS con 110 seguido del PhM con 39. De los demás tipos de tumores hay muy pocos casos siendo el que más tiene el EAPaNOS con 18 casos y el que menos el PaM con sólo 2 casos.

La edad de los pacientes comprende desde un mínimo de 20 años hasta un máximo de 84 años, aunque la media suele estar alrededor de los 40 años.

Salvo en el caso del PaNOS en la mayoría de los grupos prácticamente la mitad de los casos son hombres y la mayoría de raza blanca.

En general no parece que haya mucha mortalidad y los grupos en los que ningún paciente ha fallecido son EAPaM y PaNOS.

	Extra-adrenal paraganglioma, maligno (N = 6)	Extra-adrenal paraganglioma, NOS (N = 18)	Paraganglioma, maligno (N = 2)	Paraganglioma, NOS (N = 4)	Feocromocitoma, maligno (N = 39)	Feocromocitoma, NOS (N = 110)
Edad en el diagnóstico						
media (sd)	39.17 ± 9.91	50.83 ± 17.05	45.50 ± 13.44	49.25 ± 7.89	51.00 ± 15.98	46.87 ± 14.96
min	21	25	36	40	20	20
max	48	76	55	58	78	84
Género n(%)						
Masculino	3 (50)	9 (50)	1 (50)	1 (25)	19 (49)	45 (41)
Raza n(%)						
Blanco	4 (67)	16 (89)	2 (100)	2 (50)	32 (82)	92 (84)
Afro-Americano	2 (33)	2 (11)	0 (0)	2 (50)	6 (15)	8 (7)
Estado n(%)						
Vivo	6 (100)	17 (94)	1 (50)	4 (100)	36 (92)	109 (99)

Tabla 2. Datos demográficos de los pacientes incluidos en el estudio. En el resumen descriptivo de las muestras (Tabla 3) se puede observar que en todos los grupos hay más o menos el mismo número de muestras tumorales que muestras de sangre y además encontramos muestras de metástasis y de adicional-nuevo primario en PaM y en PhM.

En el caso de EAPaM encontramos que la mayoría de las muestras provienen del retroperitoneo (45%), seguido de cabeza, cara o cuello (36%).

Las muestras de EAPaNOS se aislaron principalmente de retroperitoneo (66%).

En el caso de los paragangliomas encontramos que en PaM un 67% de las muestras se aíslan de sistema nervioso, mientras que en PaNOS sólo existen datos de dos de las muestras que son aisladas retroperitoneo.

Los feocromocitomas son los tumores de los que más muestras hay y la mayoría en ambos diagnósticos se aislaron de las glándulas adrenales.

	Extra-adrenal paraganglioma, maligno (N = 11)	Extra-adrenal paraganglioma, NOS (N = 32)	Paraganglioma, maligno (N = 6)	Paraganglioma, NOS (N = 8)	Feocromocitoma, maligno (N = 79)	Feocromocitoma, NOS (N = 210)
Tejido/Órgano de origen n(%)						
Glándula Adrenal	0 (0)	0 (0)	2 (33)	0 (0)	75 (95)	206 (98)
Médula de Glándula Adrenal	0 (0)	0 (0)	0 (0)	0 (0)	2 (3)	0 (0)
Córtex de Glándula Adrenal	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	2 (1)
Sistema Nervioso	0 (0)	0 (0)	4 (67)	0 (0)	0 (0)	0 (0)
Retroperitoneo	5 (45)	21 (66)	0 (0)	2 (25)	0 (0)	2 (1)
Tejidos blandos	0 (0)	2 (6)	0 (0)	0 (0)	0 (0)	0 (0)
Corazón	0 (0)	1 (3)	0 (0)	0 (0)	0 (0)	0 (0)
Tórax	0 (0)	3 (9)	0 (0)	0 (0)	0 (0)	0 (0)
Cabeza, cara o cuello	4 (36)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Cuerpo aórtico y otros paraganglios	0 (0)	0 (0)	0 (0)	0 (0)	2 (3)	0 (0)
Mediastino	2 (18)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Tipo de muestra n(%)						
Tumor primario	5 (45)	15 (47)	2 (33)	4 (50)	38 (48)	98 (47)
Sangre Normal	6 (55)	17 (53)	2 (33)	4 (50)	39 (49)	107 (51)
Tejido Sólido Normal	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	5 (2)
Metástasis	0 (0)	0 (0)	1 (17)	0 (0)	1 (1)	0 (0)
Adicional - Nuevo Primario	0 (0)	0 (0)	1 (17)	0 (0)	1 (1)	0 (0)

Tabla 3. Resumen de datos de las muestras que se van a analizar.

En este trabajo se va a realizar el análisis de datos de CNAs en primer lugar y el resto de los análisis se van a centrar en los resultados obtenidos de los análisis de CNAs, haciendo una clasificación de las muestras de los distintos tumores según las puntuaciones que tengan en CNAs focales (FCS), amplios (BCS) y generales (GCS).

3.1. Análisis de datos de CNAs:

Para tener una visión general de los datos, lo primero que se realizó fue un gráfico de frecuencias de CNAs en todas las muestras. (Figura 4)

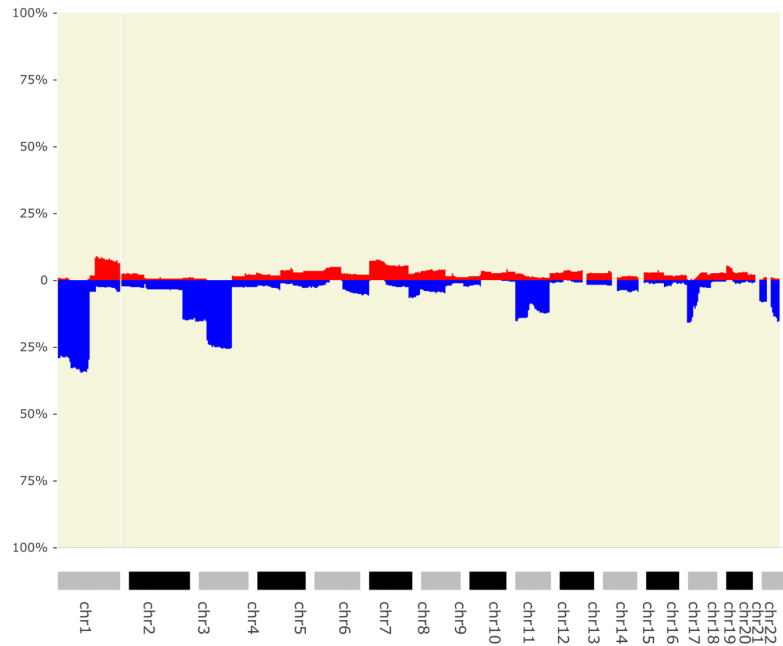


Figura 4. Frecuencia de CNAs de todas las muestras juntas.

En rasgos generales, se observa que hay más pérdida de material genético que ganancia. Las pérdidas más llamativas aparecen en los cromosomas 1, 3 y 11.

Los siguientes gráficos de frecuencias de CNAs que se realizaron fueron comparando los diagnósticos maligno y NOS en los tres grupos tumorales, extra-adrenal paraganglioma, paraganglioma y feocromocitoma. (Figura 5).

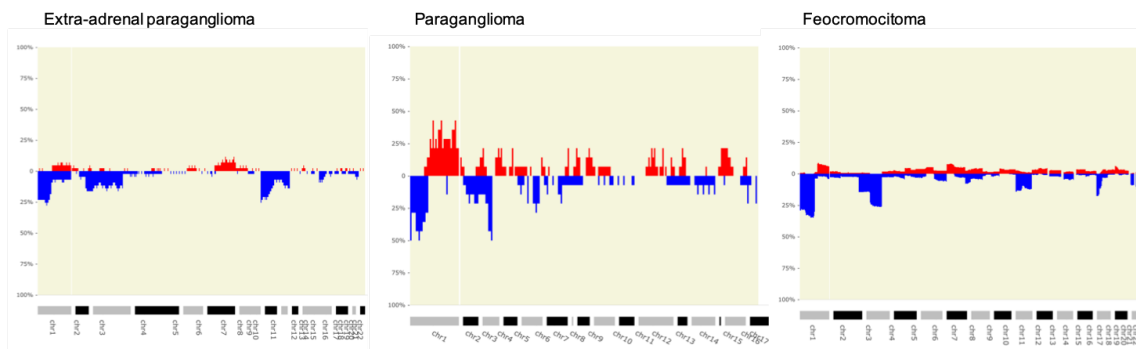


Figura 5. Frecuencia de CNAs de las muestras agrupadas en (de izquierda a derecha) Paragangliomas extra-adrenales, Paragangliomas y feocromocitomas.

Al observar los gráficos, si parece que entre los dos diagnósticos se encuentren diferencias de CNAs y también se puede apreciar que los tres tipos de tumores mantienen más o menos el mismo patrón de CNAs.

Cuando realizamos un estudio de comparaciones de todos los CNAs (Figura 6) vemos que se obtienen más regiones con diferencias significativas entre PaM-EAPaM, PaM-EAPaNOS, PaNOS-PaM, PhM-PaM y PhNOS-PaM.

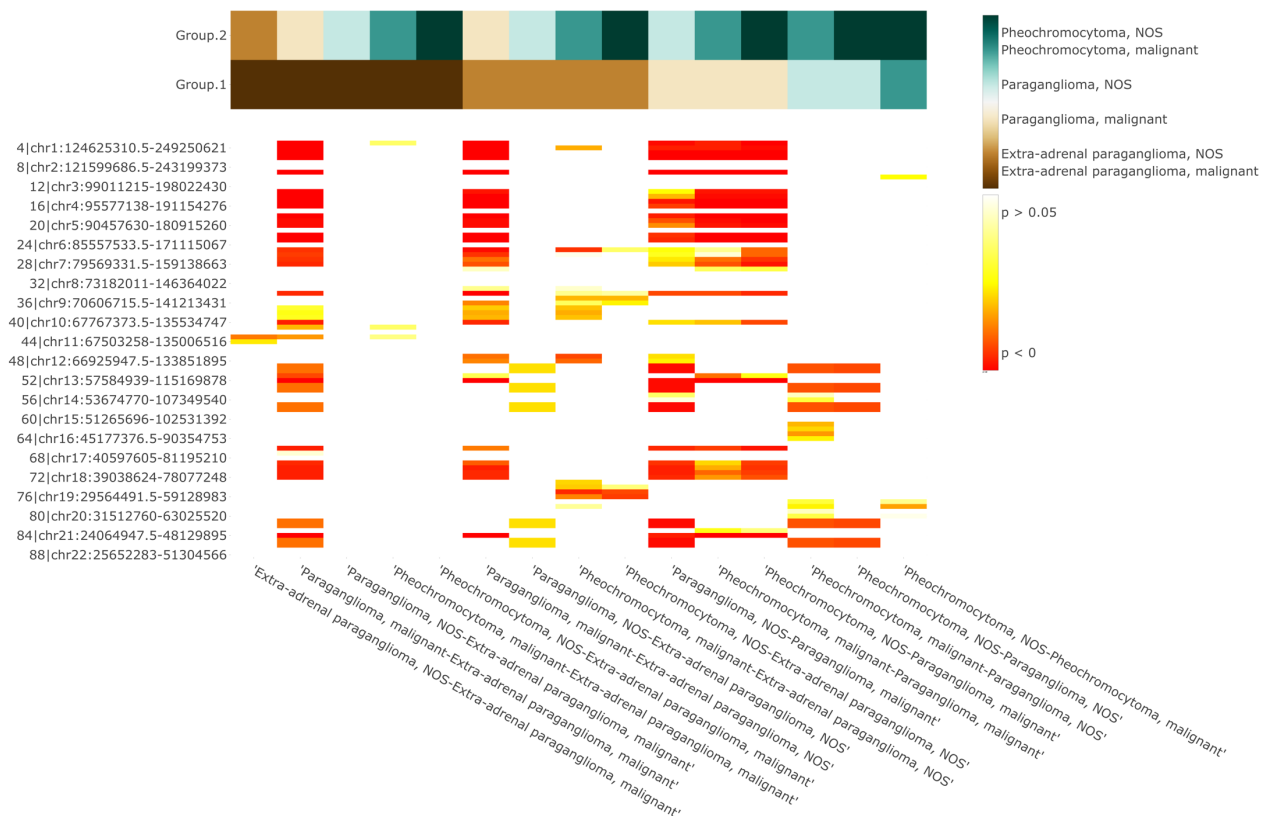


Figura 6. Comparación de CNAs por regiones entre los distintos grupos diagnósticos.

Posteriormente se estudiaron las diferencias de puntuaciones de los CNAs focales, los CNAs anchos y los CNAs generales en los seis grupos diagnósticos. (Figura 7).

Se observa que, en todos los casos, según el test de Kruskal Wallis, las diferencias entre los diagnósticos son significativas. Si embargo, si nos fijamos sólo en las diferencias entre maligno y NOS dentro de cada diagnóstico, sólo se observan diferencias significativas en el caso del

paraganglioma. Además el PaM es el tipo de tumor con los valores más altos de los tres tipos de CNAs.

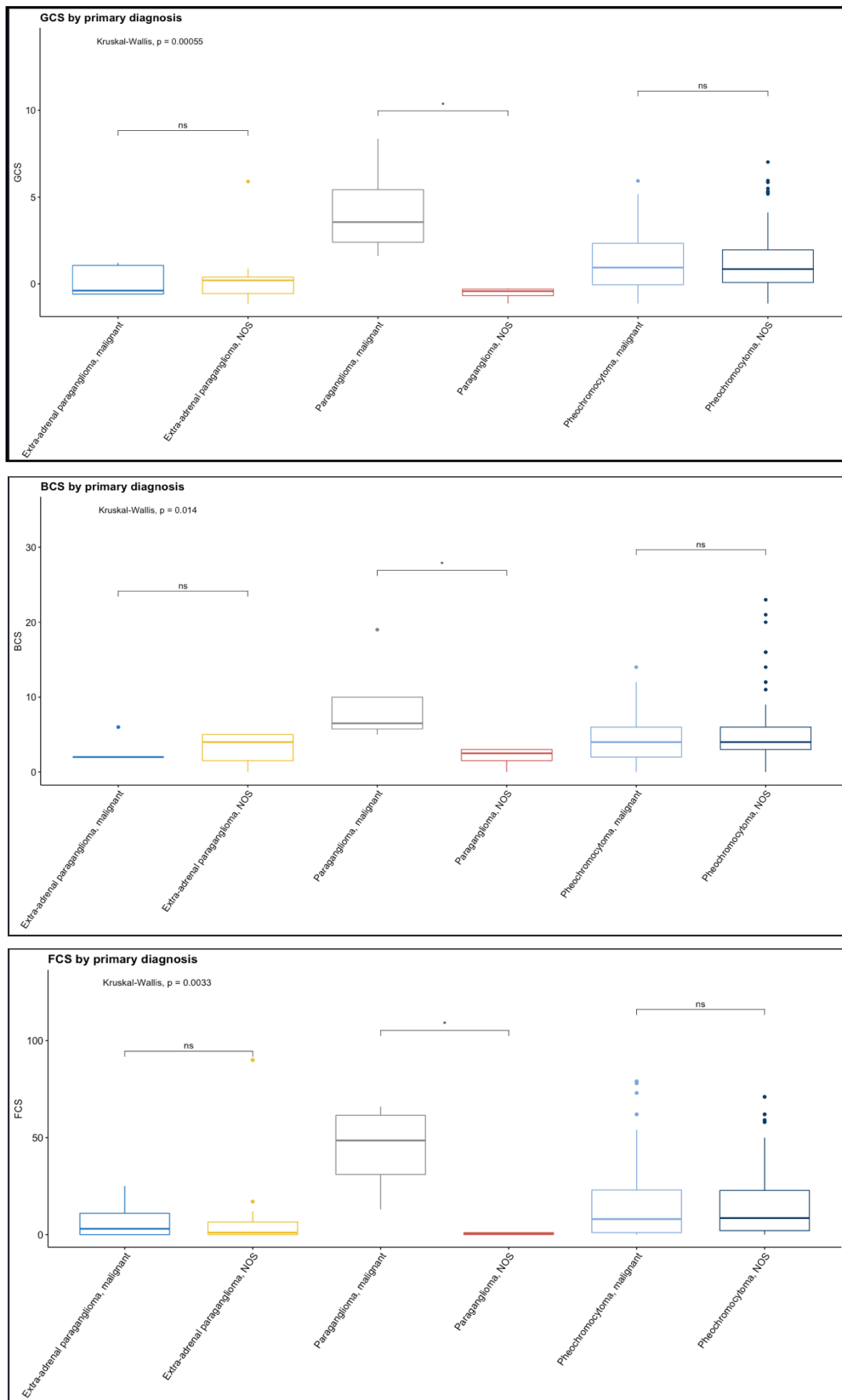


Figura 7. Puntuaciones de los distintos CNAs según el grupo diagnóstico.

Para realizar una selección de genes en cada grupo diagnóstico, se realizaron comparaciones dentro de cada grupo entre las muestras tumorales y las muestras de sangre en base a los distintos tipos de CNAs. Se exportaron los datos de puntuaciones de los distintos tipos de CNAs de cada grupo diagnóstico, se anotaron y se realizó una selección de aquellos genes que aparecían con pérdida o ganancia de material genético y, en los casos en los que se pudo, que tuviesen un p-valor menor de 0,01.

Con estos datos se realizaron análisis de enriquecimiento funcional de genes centrados en el proceso biológico. (Figuras 5, 6 y 7)

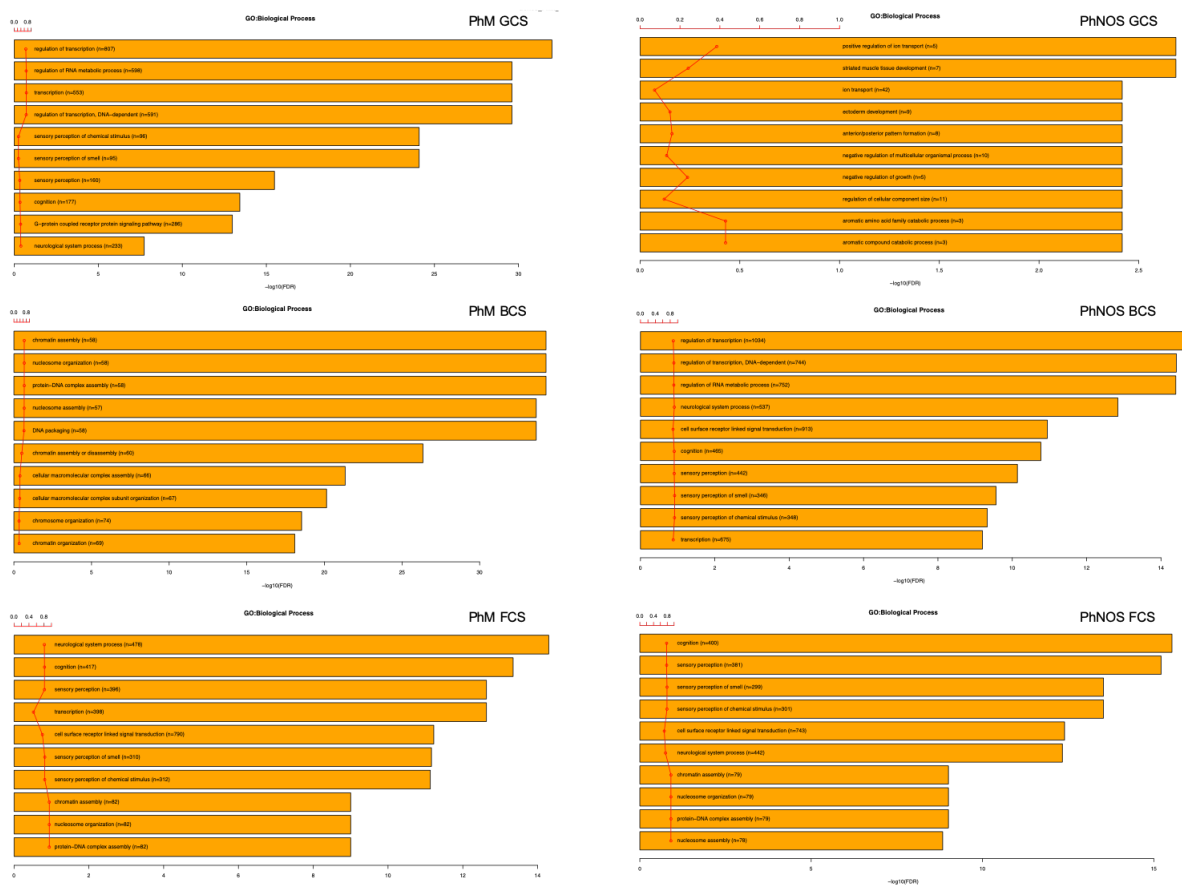


Figura 8. Gráficos de barras con los procesos biológicos más representados en PhM y PhNOS según el tipo de CNA.

Como se observa en la figura 8, en los tumores PhM las funciones biológicas que aparecen en los GCS están relacionadas la mayoría con la transcripción del ácido ribonucleico (ARN) y también con la percepción de

estímulos químicos y cascadas de señalización de receptores acoplados a proteínas G.

Con respecto a los BCS, se observa que principalmente las funciones afectadas están relacionadas con organización del material genético como organización y ensamblado del nucleosoma, ensamblado y organización de la cromatina, empaquetamiento del ácido desoxirribonucleico (ADN), etc.

Sin embargo, en el caso de los FCS están relacionados con genes implicados en percepción sensorial, transcripción, organización del nucleosoma o ensamblado de la cromatina.

Cuando observamos el grupo PhNOS vemos que hay bastantes cambios en las funciones biológicas representadas, en primer lugar, en GCS vemos funciones relacionadas con el desarrollo celular, el transporte celular y alguna función de catabolismo. En este tipo tumoral se observa que el número de genes disminuye bastante, ya que no salieron apenas diferencias significativas.

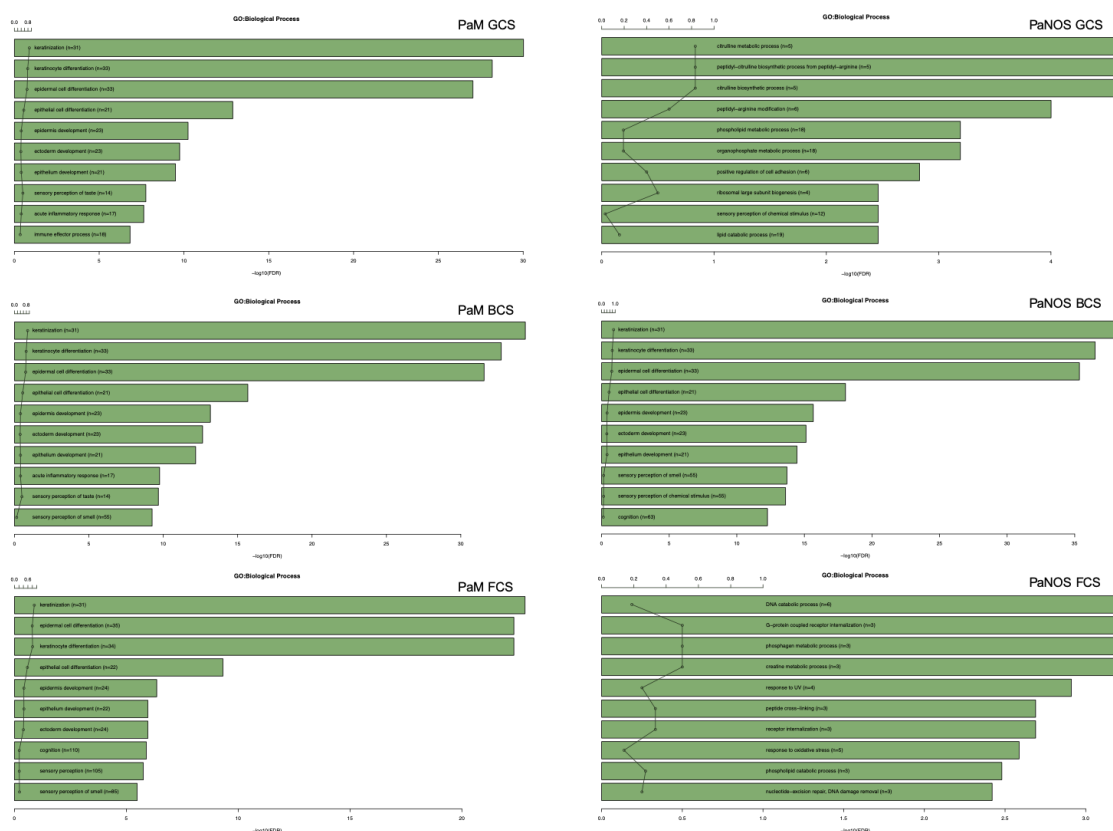


Figura 9. Gráficos de barras con los procesos biológicos más representados en PaM y PaNOS según el tipo de CNA.

En la figura 9 se pueden observar los datos de GCS y BCS en PaM y en ambos casos las funciones más representadas tienen relación con diferenciación celular, desarrollo celular y respuesta inmune. Con respecto a los FCS la única diferencia que se aprecia es que no aparecen genes relacionados con respuesta inmune.

En cambio en PaNOS, las funciones más representadas en GCS están relacionadas con metabolismo, catabolismo y regulación de la adhesión celular.

Cuando se observan los datos de BCS las funciones que aparecen están relacionadas con desarrollo y diferenciación celular, y sin embargo los FCS están relacionados con genes que intervienen en funciones de catabolismo del ADN, internalización del receptor acoplado a proteínas G, metabolismo, respuestas a estrés oxidativo y a eliminación de ADN dañado.

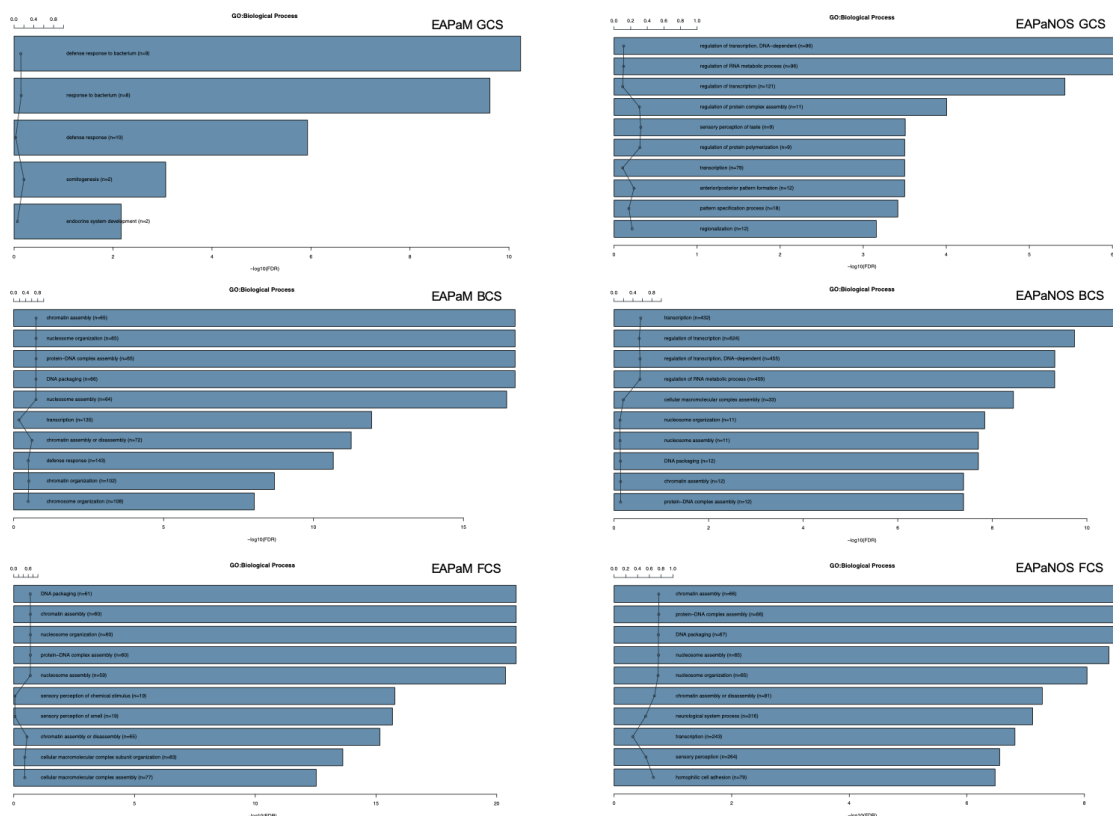


Figura 10. Gráficos de barras con los procesos biológicos más representados en EAPaM y EAPaNOS según el tipo de CNA.

Siguiendo con estos análisis se observan las funciones de los genes de los distintos tipos de CNAs dentro del grupo de EAPaM, (Figura 10) se aprecia que las funciones de los genes asociados a GCS están relacionados con respuestas de defensa y sistema endocrino.

Las funciones en el caso de los BCS están relacionadas con organización de la cromatina, el ADN y el nucleosoma, hay genes relacionados con transcripción y otros relacionados con respuestas de defensa. En este tipo de tumores las funciones de genes relacionadas con FCS son muy parecidas a las de BCS.

Por último, al estudiar el grupo de EAPaNOS, se encuentran funciones relacionadas con regulación de la transcripción y organización del material genético en los tres tipos de CNAs.

3.2. Datos de metilación:

Los datos de metilación que se estudiaron fueron los de muestras tumorales solamente. El primer análisis que se llevó a cabo fue un estudio de los valores medios de metilación, en base a los valores beta de cada grupo tumoral, y se pudo observar que no había grandes diferencias de metilación entre ellos. (Figura 11)

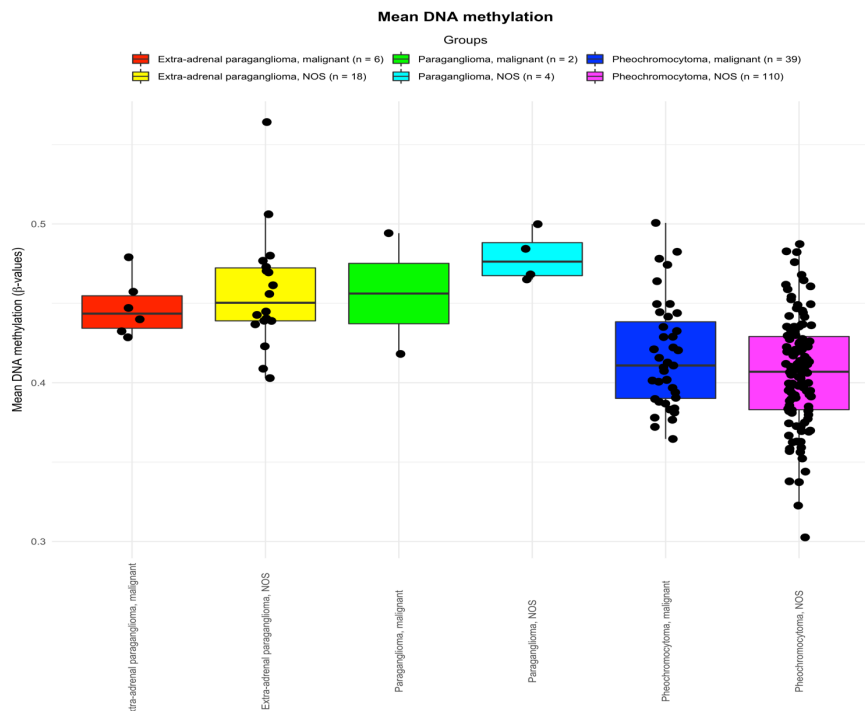


Figura 11. Distribución de las medias de metilación entre los distintos grupos diagnósticos.

Un dato que se puede observar en este primer análisis, es que parece que los dos grupos de feocromocitomas tienden a un menor grado de metilación con respecto a los otros grupos.

También se realizaron gráficos volcano, para ver si se encontraban genes con diferencias en el estado de metilación que fuesen estadísticamente significativas, comparando los distintos grupos tumorales, pero no se encontró ninguna diferencia significativa.

El siguiente paso fue estudiar la diferencia de metilación por grupo diagnóstico, subdividiendo las muestras en grupos según su puntuación en los distintos tipos de CNAs.

Se clasificaron las muestras de cada grupo diagnóstico por cuartiles, según la puntuación que tuviesen de cada tipo de CNA. Con esta clasificación se estudió si había diferencias de metilación, principalmente entre el tercer y primer cuartil de cada tipo de CNA en cada grupo diagnóstico. Este tipo de análisis no se pudo llevar a cabo en algunos grupos diagnósticos como PaM por el número de muestras tan reducido. En otros casos no se pudo obtener resultados porque dentro de alguno de los cuartiles no había muestras suficientes para poder realizar la comparación.

En ningún caso se encontraron diferencias significativas, aunque en algunos casos como PhM y PhNOS en todos los tipos de CNAs se podía apreciar una ligera tendencia a que siempre hubiese menor metilación en el tercer cuartil (es decir mayor número de CNAs) que en el primero. (Figura 12)

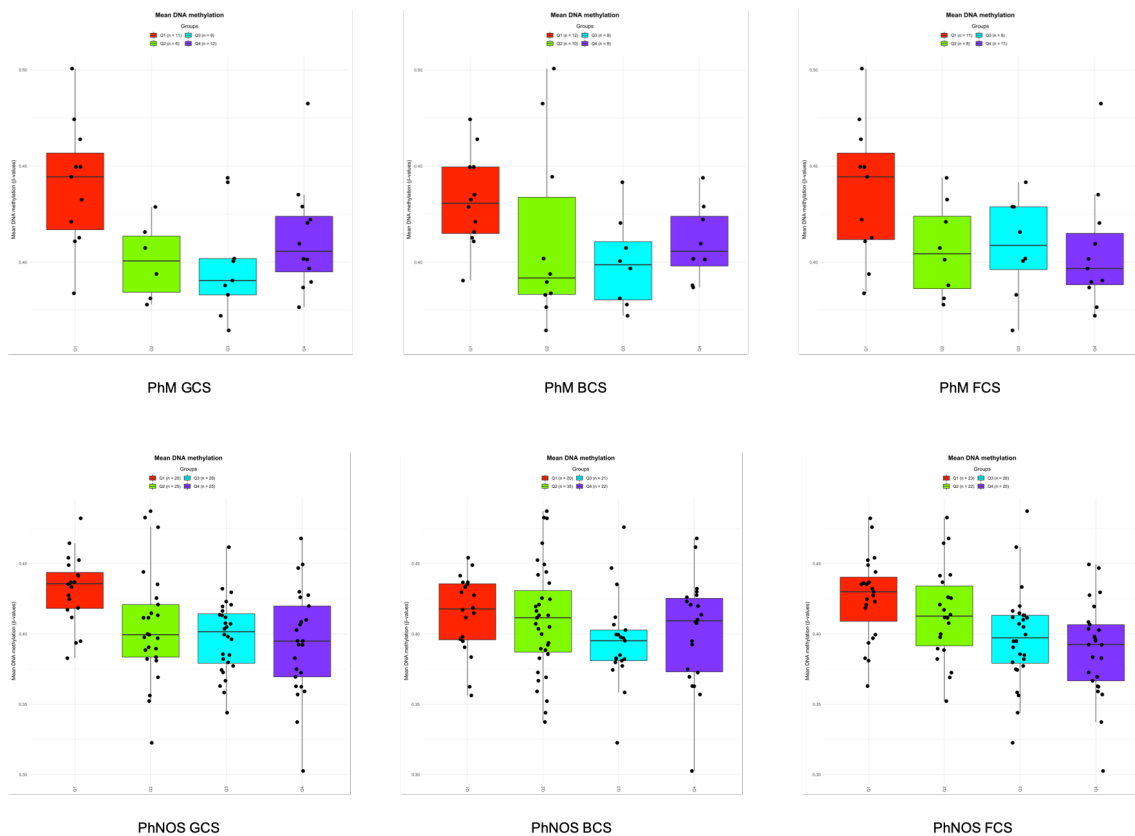


Figura 12. Distribución de las medias de metilación en PhM y PhNOS según los distintos tipos de CNAs.

También se estudió qué genes podían tener diferencias de metilación entre las muestras que se encontraban en el tercer cuartil de los distintos CNAs frente a las que se encontraban en el primer cuartil.

En este caso salieron resultados significativos en el grupo de PhNOS (Figura 13). En los tres tipos de CNAs los genes que salían con un estado de metilación significativamente distinto fueron, prácticamente todos, hipometilados en el tercer cuartil con respecto al primer cuartil.

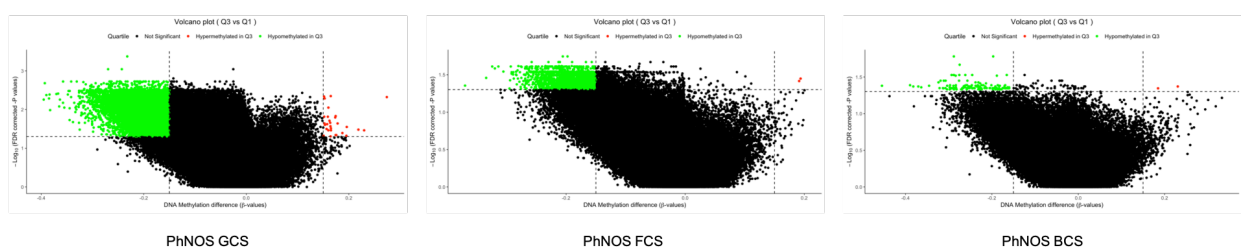


Figura 13. Gráficos volcano con diferencias de metilación entre el tercer y el primer cuartil de los distintos tipos de CNAs en PhNOS.

Al analizar las funciones biológicas de estos genes se encontró que en GCS, donde aparece un mayor número de genes hipometilados de manera estadísticamente significativa, aparecen un gran número de procesos distintos, entre ellos desarrollo y morfología neuronal, adhesión celular, quimiotaxis, organización de la matriz extracelular, procesos de transducción de señales, procesos de migración, procesos metabólicos y respuesta inmune.

Sin embargo, en el caso de FCS ya encontramos menos genes y menos procesos biológicos, entre ellos desarrollo, morfología celular y diferenciación celular.

Los pocos genes que se encuentran en BCS están implicados en procesos de regulación de transducción de señales, comunicación y señalización celular.

3.3. Datos de mutaciones:

En el caso de las mutaciones se realizó un primer análisis general de todas las muestras, en el que se vio que la mayoría de las mutaciones en estos tumores son mutaciones con cambio de sentido, el tipo de mutación más frecuente es de poliformismos de un solo nucleótido (SNPs del inglés "Single nucleotide polymorphism"). La mediana de mutaciones por muestra es de 8 y los 10 genes que se encuentran más frecuente mutados son NF1, HRAS, EPAS1, RET, MUC16, ATRX, ABCA13, TTN, CSDE1 y AHNAK.(Figura 11) También se aprecia que los genes HRAS, RET, ABCA13, RYR1 y VHL sólo aparecen mutados en muestras de feocromocitomas. (Figura 14)

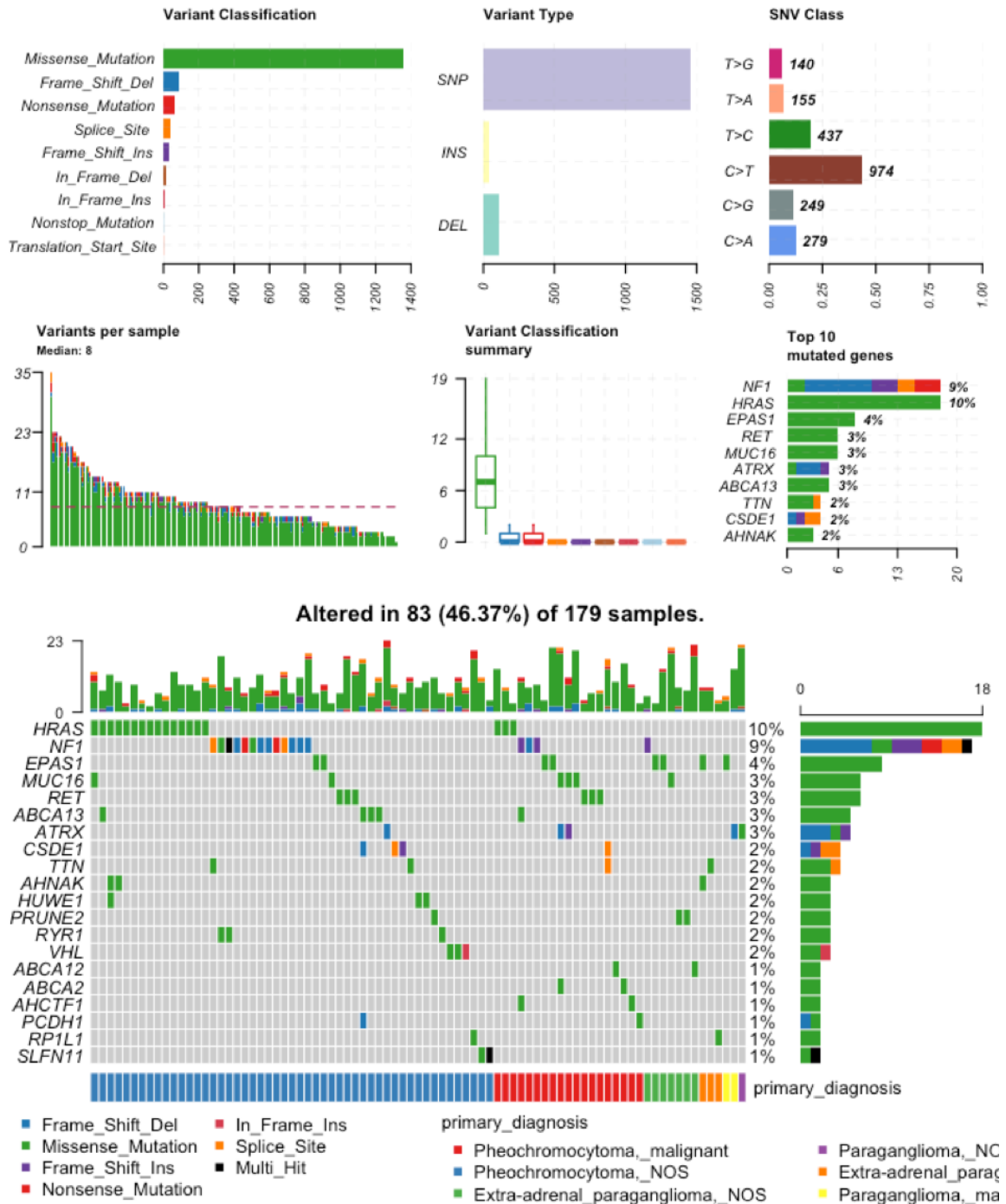


Figura 14. Resumen de datos de mutaciones de todas las muestras juntas.

También se realizó un análisis del número de genes mutados pertenecientes a rutas oncogénicas y se vio que las rutas con más genes afectados eran RTK-Ras, NOTCH, WNT e Hippo. (Figura 15)

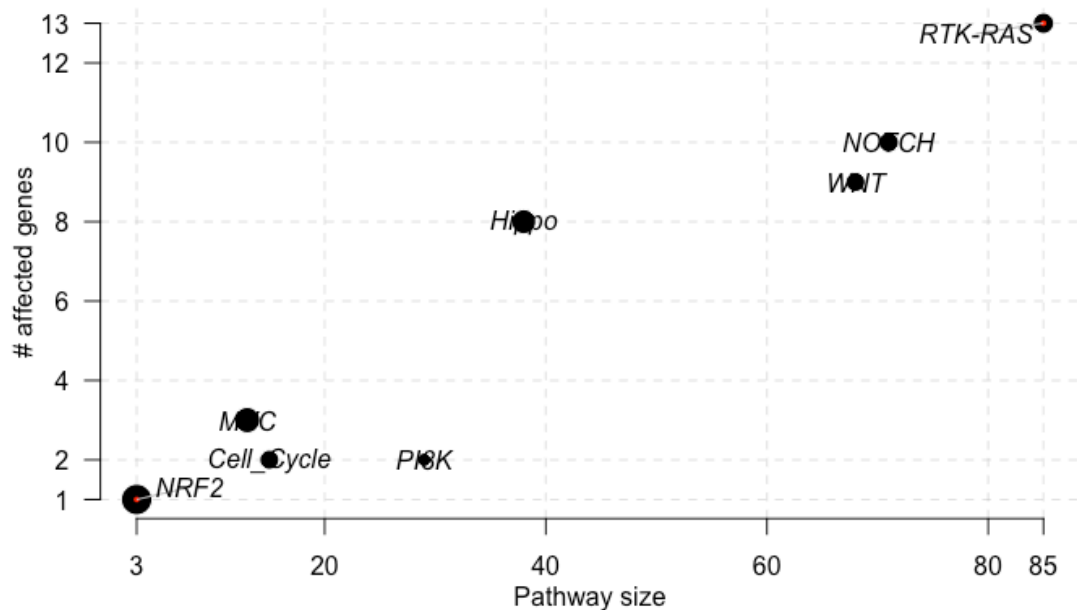


Figura 15. Rutas oncogénicas afectadas por mutaciones en sus genes de todo el conjunto de muestras del estudio.

Al realizar este tipo de análisis ya separando por grupos diagnósticos, en el caso de PhM encontramos, entre los diez genes más frecuentemente mutados, PCDH1, MAST4 y AHCTF1, que no se encontraban en el análisis de todas las muestras juntas. En el resto de los resultados no hay diferencias.

Al realizar este mismo análisis en las muestras pertenecientes al grupo de PhNOS, lo primero que llama la atención es que hay varias diferencias en los genes que se encuentran más mutados, tanto con respecto al conjunto total de muestras, como con respecto a PhM.

También se encuentran grandes diferencias con respecto a los grupos anteriormente nombrados en PaM, sólo coinciden los genes EPAS1 y ATRX con alguno de los grupos anteriores, aunque este grupo ha quedado reducido en este análisis a dos muestras y esto puede hacer que los datos sean tan variables.

El grupo diagnóstico PaNOS también cuenta con muy pocas muestras y los resultados son muy diferentes, sólo coincide ATRX con los demás grupos, los demás genes son todos distintos.

Tanto EAPaM como EAPaNOS tienen genes muy diferentes con mutaciones, sólo coinciden con los demás grupos AHNAK y EPAS1 respectivamente.

Como se contaba con los datos de puntuaciones de los distintos tipos de CNAs, se quiso analizar si había diferencias en los genes que aparecían mutados en las muestras según se encontrasen en distintos cuartiles de puntuación de CNAs. Puesto que los únicos grupos con un número considerable de muestras eran PhM y PhNOS, sólo se pudo realizar este análisis en dichos grupos. En la figura 16 se encuentran los datos de PhM y se observa que en BCS el gen HRAS sólo aparece mutado en las muestras del tercer cuartil. En cuanto a FCS los genes MUC16 y ATRX sólo se encuentran mutados en el rango más alto de puntuación de este tipo de CNAs y en cuanto a GCS pasa lo mismo con estos genes y además con EPAS1.

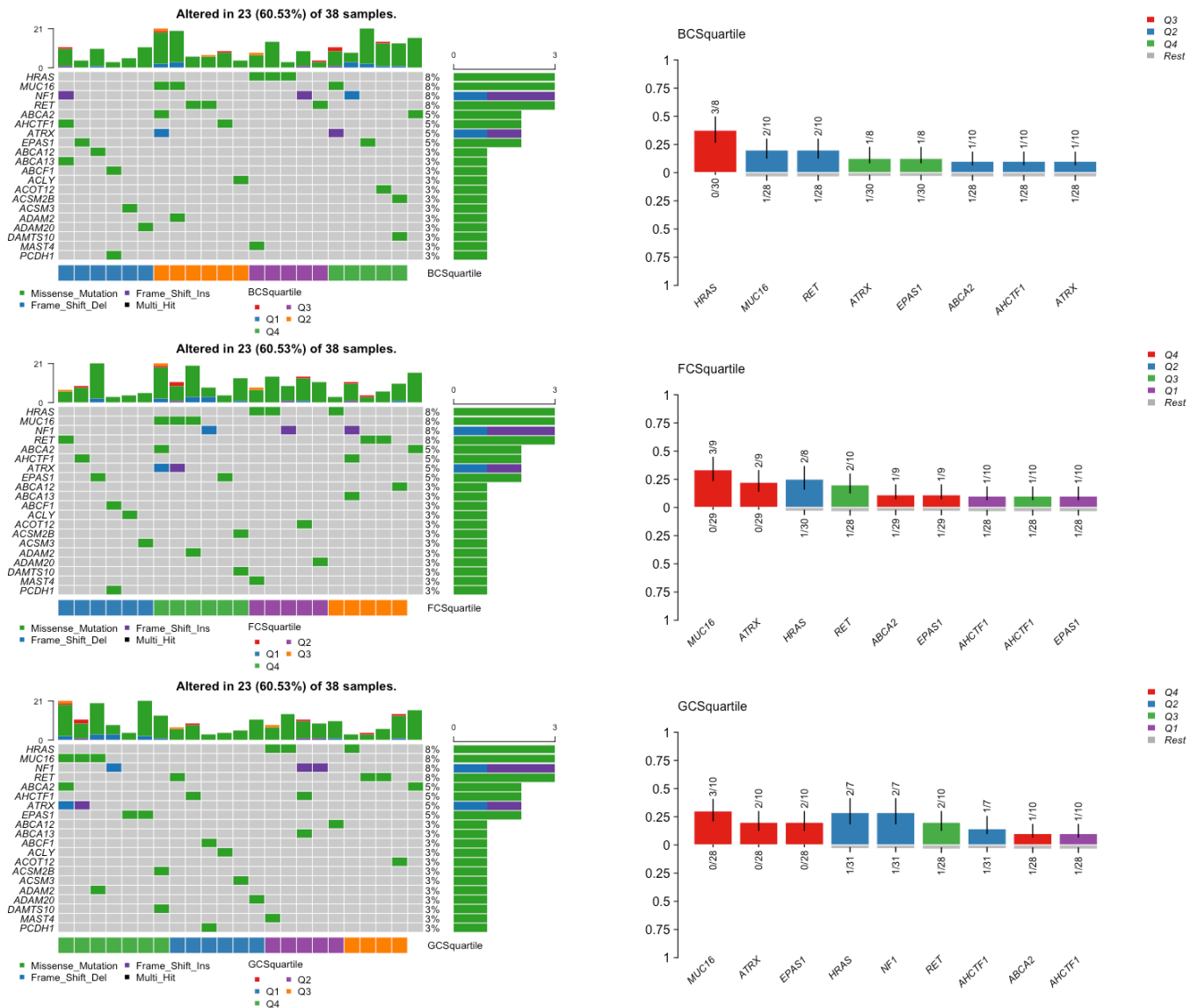


Figura 16. Datos de mutaciones de PhM en base a los cuartiles de los distintos CNAs.

Al realizar estos mismos análisis en las muestras de PhNOS, se encuentra que no hay ningún gen que esté sólo mutado en alguno de los cuartiles de puntuación y no en los demás, salvo el caso de VHL en el primer cuartil de FCS. (Figura 17)

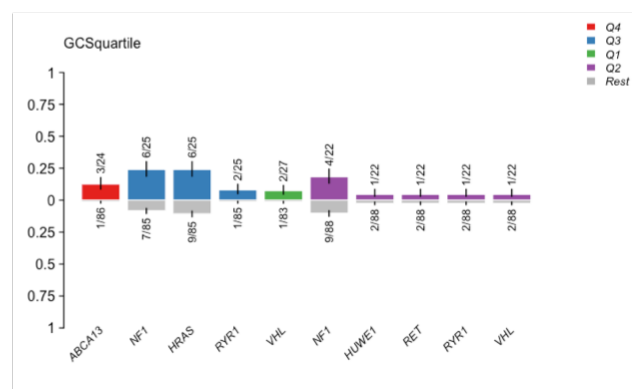
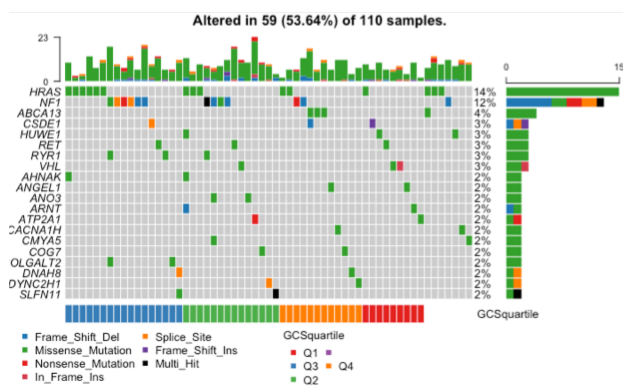
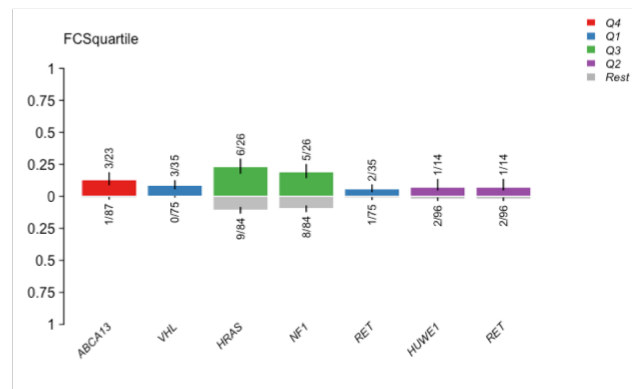
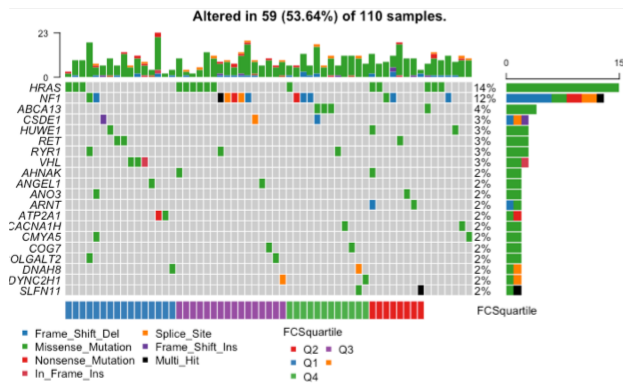
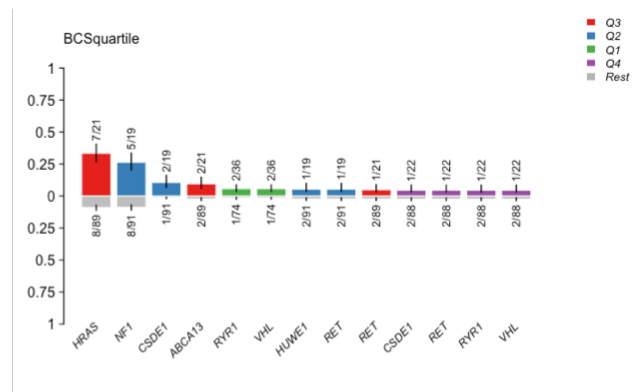
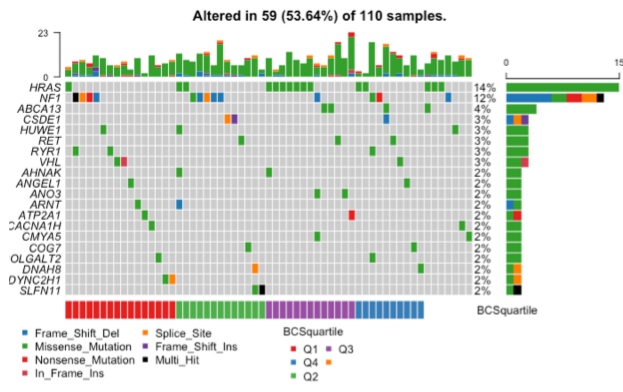


Figura 17. Datos de mutaciones de PhNOS en base a los cuartiles de los distintos CNAs.

Al estudiar algunos de los genes que salen mutados en más grupos diagnósticos, y algunos de ellos con relación a mayor puntuación en los cuartiles de CNAs, se podían observar diferencias en los grados de metilación y de expresión. (Figura 18) En algunos genes, como NF1 se puede observar que las muestras en las que aparece mutado tienen una expresión menor que en el resto de muestras. Otros como HRAS y RET aparecen siempre hipermetilados en muestras de paragangliomas independientemente de si tienen mutación o no. EPAS en las muestras en la que aparece con una mutación está hipometilado.

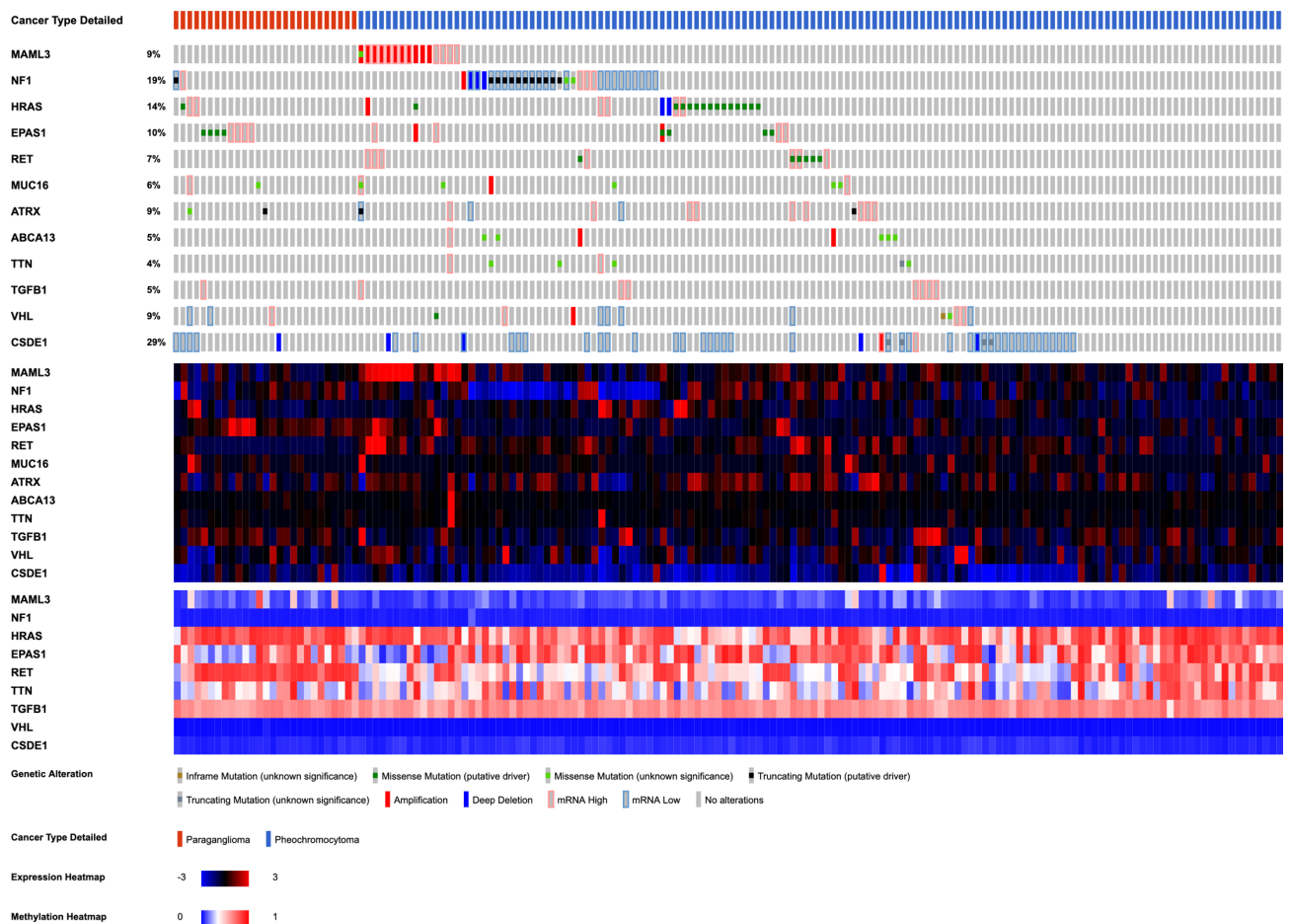


Figura 18. Gráfico que integra todos los datos de las muestras con los genes que se han encontrado más frecuentemente mutados.

Aparte de esto, se ha descrito un marcador llamado MATH que sirve para determinar el grado de heterogeneidad genética intratumoral, y se ha visto que hay una relación entre este marcador y la tasa de supervivencia, a mayor índice de MATH menor tasa de supervivencia. Por esto parecía interesante estudiar si podía haber relación entre este marcador y los distintos tipos de CNAs. Al realizar este estudio en todos los grupos tumorales únicamente en el grupo PhM se vio una correlación positiva entre MATH y BCS. (Figura 19)

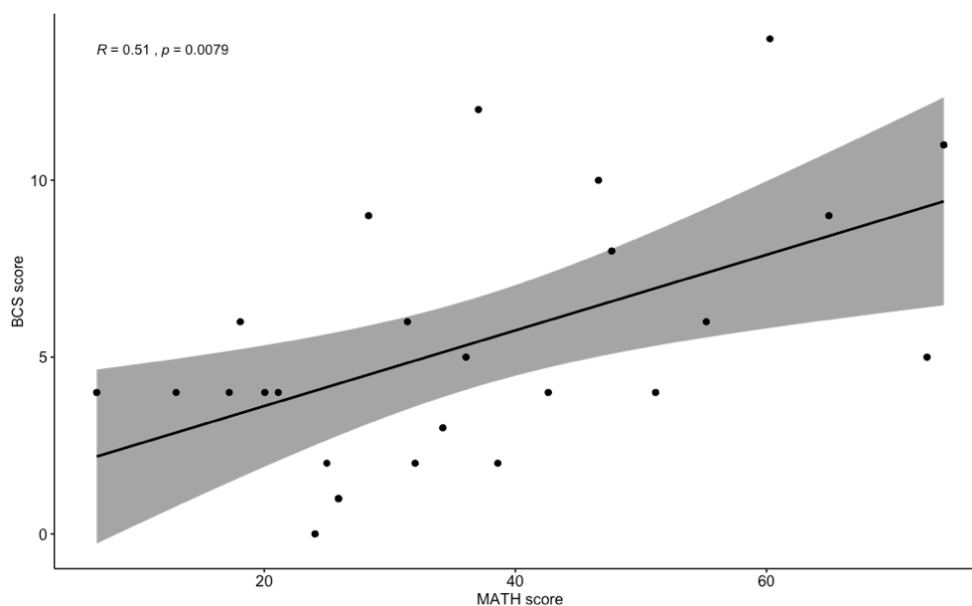


Figura 19. Relación de las puntuaciones de MATH y de BCS en PhM.

3.4. Datos de expresión génica:

Con los datos de expresión génica se realizaron los siguientes análisis, tomando los datos de los cuartiles de los distintos CNAs, se estudiaron diferencias de expresión entre las muestras que se encontraban en el tercer cuartil y las que se encontraban en el primer cuartil de cada tipo de CNA. Estos análisis se repitieron en cada grupo diagnóstico.

Posteriormente se realizaron análisis de enriquecimiento funcional para saber qué procesos podían estar modificados en estos tumores.

PhM_BCS_Q1 genes Vs BCS_Q3	FDR	genes totales	genes en común
GO:0048812~neuron projection morphogenesis	0.00353	71	8
GO:0006936~muscle contraction	0.00353	29	5
GO:0042273~ribosomal large subunit biogenesis	0.00353	8	3
GO:0031175~neuron projection development	0.00353	125	10
GO:0003012~muscle system process	0.00353	31	5
GO:0048666~neuron development	0.00593	163	11
GO:0050768~negative regulation of neurogenesis	0.00593	11	3
GO:0030182~neuron differentiation	0.00593	145	10
GO:0006801~superoxide metabolic process	0.00593	12	3
GO:0034101~erythrocyte homeostasis	0.00593	12	3
PhM_FCS_Q1 genes Vs FCS_Q3	FDR	genes totales	genes en común
GO:0008202~steroid metabolic process	2.33e-08	139	13

<i>GO:0006694~steroid biosynthetic process</i>	1.61e-06	40	7
<i>GO:0042446~hormone biosynthetic process</i>	2.41e-05	11	4
<i>GO:0006704~glucocorticoid biosynthetic process</i>	3.89e-05	5	3
<i>GO:0008207~C21-steroid hormone metabolic process</i>	7.73e-05	6	3
<i>GO:0034754~cellular hormone metabolic process</i>	1.03e-03	29	4
<i>GO:0010817~regulation of hormone levels</i>	1.33e-03	32	4
<i>GO:0045765~regulation of angiogenesis</i>	2.05e-03	16	3
<i>GO:0010575~positive regulation vascular endothelial growth factor production</i>	2.52e-03	5	2
<i>GO:0034382~chylomicron remnant clearance</i>	2.52e-03	5	2
PhM_GCS_Q1 genes Vs GCS_Q3	FDR	genes totales	genes en común
<i>GO:0008202~steroid metabolic process</i>	3.32e-07	139	13
<i>GO:0006694~steroid biosynthetic process</i>	3.32e-07	40	8
<i>GO:0042446~hormone biosynthetic process</i>	5.59e-05	11	4
<i>GO:0006704~glucocorticoid biosynthetic process</i>	7.20e-05	5	3
<i>GO:0008207~C21-steroid hormone metabolic process</i>	1.43e-04	6	3
<i>GO:0007155~cell adhesion</i>	6.75e-04	318	13
<i>GO:0022610~biological adhesion</i>	6.75e-04	318	13
<i>GO:0034754~cellular hormone metabolic process</i>	1.73e-03	29	4
<i>GO:0010817~regulation of hormone levels</i>	2.30e-03	32	4
<i>GO:0034382~chylomicron remnant clearance</i>	3.34e-03	5	2

Tabla 4. Resultados de análisis de procesos biológicos en PhM.

En PhM tanto en GCS como en FCS aparecían genes relacionados con procesos de regulación, metabolismo y biosíntesis de hormonas esteroideas y glucocorticoides y también procesos relacionados con la adhesión celular. (Tabla 4)

Sin embargo, los procesos encontrados en BCS no se parecen en nada, se encuentran procesos como desarrollo neuronal o formación de subunidad grande ribosomal. (Tabla 4)

Con respecto a PhNOS en el caso de GCS nos encontramos funciones de angiogénesis y proliferación celular, mientras que en BCS nos encontramos con procesos relacionados con la fuerza de contracción del corazón relacionada con señales químicas, adhesión celular, regulación de la transcripción, procesos vasculares y señalización por segundos mensajeros. Cuando estudiamos FCS vemos que los procesos también difieren con los anteriores y están relacionados con la ruta de señalización

Wnt, respuestas a estímulos de hormonas esteroideas y regulación de componentes de biogénesis celular entre otros. (Tabla 5)

PhNOS_BCS_Q1 genes Vs BCS_Q3	FDR	genes totales	genes en común
<i>GO:000122~negative regulation of transcription from RNA polymerase II promoter</i>	0.000668	137	7
<i>GO:0035150~regulation of tube size</i>	0.000668	30	4
<i>GO:0050880~regulation of blood vessel size</i>	0.000668	30	4
<i>GO:0003018~vascular process in circulatory system</i>	0.000748	33	4
<i>GO:0003057~regulation of the force of heart contraction by chemical signal</i>	0.001335	4	2
<i>GO:0007155~cell adhesion</i>	0.001335	318	9
<i>GO:0022610~biological adhesion</i>	0.001335	318	9
<i>GO:0019932~second-messenger-mediated signaling</i>	0.001362	84	5
<i>GO:0044057~regulation of system process</i>	0.001374	140	6
<i>GO:0050678~regulation of epithelial cell proliferation</i>	0.001374	20	3
PhNOS_FCS_Q1 genes Vs FCS_Q3	FDR	genes totales	genes en común
<i>GO:0006820~anion transport</i>	0.00819	93	4
<i>GO:0010769~regulation of cell morphogenesis involved in differentiation</i>	0.00819	53	3
<i>GO:0007223~Wnt receptor signaling pathway, calcium modulating pathway</i>	0.00819	20	2
<i>GO:0007588~excretion</i>	0.00819	20	2
<i>GO:0048545~response to steroid hormone stimulus</i>	0.00819	66	3
<i>GO:0051260~protein homooligomerization</i>	0.00819	23	2
<i>GO:0044087~regulation of cellular component biogenesis</i>	0.00819	73	3
<i>GO:0030182~neuron differentiation</i>	0.00819	145	4
<i>GO:0050770~regulation of axonogenesis</i>	0.00975	29	2
PhNOS_GCS_Q1 genes Vs GCS_Q3	FDR	genes totales	genes en común
<i>GO:0042127~regulation of cell proliferation</i>	5.73e-06	363	24
<i>GO:0001525~angiogenesis</i>	1.54e-05	51	9
<i>GO:0060191~regulation of lipase activity</i>	1.66e-04	38	7
<i>GO:0044057~regulation of system process</i>	2.49e-04	140	12
<i>GO:0048754~branching morphogenesis of a tube</i>	2.57e-04	18	5
<i>GO:0003018~vascular process in circulatory system</i>	3.85e-04	33	6
<i>GO:0008284~positive regulation of cell proliferation</i>	3.85e-04	205	14
<i>GO:0001763~morphogenesis of a branching structure</i>	3.85e-04	21	5
<i>GO:0043085~positive regulation of catalytic activity</i>	6.15e-04	339	18
<i>GO:0045639~positive regulation of myeloid cell differentiation</i>	6.15e-04	13	4

Tabla 5. Resultados de análisis de procesos biológicos en PhNOS.

En la tabla 6 aparece el resultado de PaM, en todas las comparaciones han salido genes relacionados con desarrollo, morfogénesis y señalización celular.

Sólo había genes que se expresasen de manera diferencial en muestras de BCS dentro del grupo de PaNOS. Los procesos en los que estaban implicados eran adhesión y señalización celular principalmente. (Tabla 6)

PaM_FCS_Q1 genes Vs FCS_Q3	FDR	genes totales	genes en común
<i>GO:0003002~regionalization</i>	5.83e-08	55	14
<i>GO:0048568~embryonic organ development</i>	5.98e-08	68	15
<i>GO:0009952~anterior/posterior pattern formation</i>	5.98e-08	50	13
<i>GO:0048562~embryonic organ morphogenesis</i>	1.29e-07	54	13
<i>GO:0007389~pattern specification process</i>	1.81e-07	101	17
<i>GO:0048705~skeletal system morphogenesis</i>	5.31e-07	26	9
<i>GO:0048706~embryonic skeletal system development</i>	6.77e-07	27	9
<i>GO:0048704~embryonic skeletal system morphogenesis</i>	9.23e-07	21	8
<i>GO:0007267~cell-cell signaling</i>	5.39e-06	317	28
<i>GO:0009792~embryonic development ending in birth or egg hatching</i>	1.11e-05	110	15
PaM_GCS_Q1 genes Vs GCS_Q3	FDR	genes totales	genes en común
<i>GO:0003002~regionalization</i>	5.83e-08	55	14
<i>GO:0048568~embryonic organ development</i>	5.98e-08	68	15
<i>GO:0009952~anterior/posterior pattern formation</i>	5.98e-08	50	13
<i>GO:0048562~embryonic organ morphogenesis</i>	1.29e-07	54	13
<i>GO:0007389~pattern specification process</i>	1.81e-07	101	17
<i>GO:0048705~skeletal system morphogenesis</i>	5.31e-07	26	9
<i>GO:0048706~embryonic skeletal system development</i>	6.77e-07	27	9
<i>GO:0048704~embryonic skeletal system morphogenesis</i>	9.23e-07	21	8
<i>GO:0007267~cell-cell signaling</i>	5.39e-06	317	28
<i>GO:0009792~embryonic development ending in birth or egg hatching</i>	1.11e-05	110	15

Tabla 6. Resultados de análisis de procesos biológicos en PaM.

El siguiente análisis fue en EAPaM en el que sólo salieron datos de GCS y FCS y en el caso de GCS los procesos fueron adhesión celular al sustrato, y metabolismo de fósforo y fosfato, mientras que en el caso de FCS los procesos estaban relacionados con señalización celular, secreción y transporte de hormonas y regulación negativa de la activación de leucocitos y linfocitos. (Tabla 7)

EAPaM_FCS_Q1 genes Vs FCS_Q3	FDR	genes totales	genes en común
<i>GO:0007610~behavior</i>	0.00413	193	6
<i>GO:0007267~cell-cell signaling</i>	0.00456	317	7
<i>GO:0006939~smooth muscle contraction</i>	0.00456	11	2
<i>GO:0002790~peptide secretion</i>	0.00456	15	2
<i>GO:0015833~peptide transport</i>	0.00456	15	2
<i>GO:0030072~peptide hormone secretion</i>	0.00456	15	2
<i>GO:0051250~negative regulation of lymphocyte activation</i>	0.00456	15	2
<i>GO:0009914~hormone transport</i>	0.00456	16	2
<i>GO:0046879~hormone secretion</i>	0.00456	16	2
<i>GO:0002695~negative regulation of leukocyte activation</i>	0.00456	17	2
EAPaM_GCS_Q1 genes Vs GCS_Q3	FDR	genes totales	genes en común
<i>GO:0045449~regulation of transcription</i>	0.00535	1172	1
<i>GO:0007218~neuropeptide signaling pathway</i>	0.00535	56	4
<i>GO:0010811~positive regulation of cell-substrate adhesion</i>	0.00535	9	2
<i>GO:0006793~phosphorus metabolic process</i>	0.00535	621	13
<i>GO:0006796~phosphate metabolic process</i>	0.00535	621	13
<i>GO:0006833~water transport</i>	0.00555	12	2
<i>GO:0010810~regulation of cell-substrate adhesion</i>	0.00555	12	2
<i>GO:0042044~fluid transport</i>	0.00555	12	2
<i>GO:0045785~positive regulation of cell adhesion</i>	0.00785	15	2
<i>GO:0009991~response to extracellular stimulus</i>	0.00785	49	3

Tabla 7. Resultados de análisis de procesos biológicos en EAPaM.

Cuando estudiamos el grupo EAPaM encontramos procesos de regulación negativa de biosíntesis y metabolismo de ácidos nucleicos y de Adenosín monofosfato cíclico (cAMP), desarrollo de la glándula adrenal y respuestas de defensa y miedo tanto en el grupo de GCS como en el de BCS, aunque en todos los casos sólo se encontró un gen implicado. (Tabla 8). El grupo de FCS muestra unos procesos distintos relacionados con adhesión celular y metabolismo de lípidos.

EAPaNOS_BCS_Q1 genes Vs BCS_Q3	FDR	genes totales	genes en común
<i>GO:0001661~conditioned taste aversion</i>	0.00206	5	1
<i>GO:0030800~negative regulation of cyclic nucleotide metabolic process</i>	0.00206	6	1
<i>GO:0030803~negative regulation of cyclic nucleotide biosynthetic process</i>	0.00206	6	1
<i>GO:0030809~negative regulation of nucleotide biosynthetic process</i>	0.00206	6	1
<i>GO:0030815~negative regulation of cAMP metabolic process</i>	0.00206	6	1
<i>GO:0030818~negative regulation of cAMP biosynthetic process</i>	0.00206	6	1
<i>GO:0045980~negative regulation of nucleotide metabolic process</i>	0.00206	6	1
<i>GO:0030325~adrenal gland development</i>	0.00206	7	1
<i>GO:0001662~behavioral fear response</i>	0.00206	9	1
<i>GO:0002209~behavioral defense response</i>	0.00206	9	1
EAPaNOS_FCS_Q1 genes Vs FCS_Q3	FDR	genes totales	genes en común
<i>GO:0006690~icosanoid metabolic process</i>	0.00191	26	2
<i>GO:0033559~unsaturated fatty acid metabolic process</i>	0.00191	26	2
<i>GO:0007155~cell adhesion</i>	0.00191	318	4
<i>GO:0022610~biological adhesion</i>	0.00191	318	4
<i>GO:0016337~cell-cell adhesion</i>	0.00420	188	3
<i>GO:0010543~regulation of platelet activation</i>	0.00954	5	1
<i>GO:0006631~fatty acid metabolic process</i>	0.00954	98	2
<i>GO:0005513~detection of calcium ion</i>	0.00954	6	1
<i>GO:0015810~aspartate transport</i>	0.00954	8	1
<i>GO:0031018~endocrine pancreas development</i>	0.00954	8	1
EAPaNOS_GCS_Q1 genes Vs GCS_Q3	FDR	genes totales	genes en común
<i>GO:0001661~conditioned taste aversion</i>	0.00419	5	1
<i>GO:0030800~negative regulation of cyclic nucleotide metabolic process</i>	0.00419	6	1
<i>GO:0030803~negative regulation of cyclic nucleotide biosynthetic process</i>	0.00419	6	1
<i>GO:0030809~negative regulation of nucleotide biosynthetic process</i>	0.00419	6	1
<i>GO:0030815~negative regulation of cAMP metabolic process</i>	0.00419	6	1
<i>GO:0030818~negative regulation of cAMP biosynthetic process</i>	0.00419	6	1
<i>GO:0045980~negative regulation of nucleotide metabolic process</i>	0.00419	6	1
<i>GO:0030325~adrenal gland development</i>	0.00419	7	1
<i>GO:0001662~behavioral fear response</i>	0.00419	9	1
<i>GO:0002209~behavioral defense response</i>	0.00419	9	1

Tabla 8. Resultados de análisis de procesos biológicos en EAPaNOS.

5. Discusión:

Con este trabajo se pretendía realizar un estudio de datos ómicos de feocromocitomas y paragangliomas de manera “personalizada”. La mayoría de los estudios que existen en la literatura son de cohortes, por ello el primer paso de este estudio fue separar los datos disponibles en seis grupos según el diagnóstico principal que se había anotado para cada caso.

Con las muestras separadas en esos seis grupos se realizaron los estudios de CNAs, poniendo de manifiesto las diferencias que había entre los distintos tipos de tumores.

En el caso de los feocromocitomas se vio que los genes que aparecían más frecuentemente mutados coincidían bastante en los dos diagnósticos, pero no completamente. Además de esto, algunas de estas mutaciones sólo aparecía en las muestras con puntuaciones más altas en alguno de los tipos de CNAs. Estos genes son RET, NF1, HRAS, EPAS1 y ATRX, estos genes ya fueron previamente descritos en este tipo de tumores por Lawrence et al. en 2013.[44] En cambio las muestras de paragangliomas eran mucho más heterogéneas y sólo se encontraba dentro de algún grupo diagnóstico algún gen de los anteriores nombrados, sobre todo EPAS1, relacionándose, como en el caso de los feocromocitomas, con altas puntuaciones de CNAs.

Estos datos de mutaciones podrían ayudar a predecir el pronóstico de la enfermedad, puesto que ya se han relacionado algunos de ellos con distintos pronósticos, por ejemplo, mutaciones somáticas de EPAS1 sin mosaicismo se relacionan con buen pronóstico de la enfermedad [45], mientras que mutaciones en ATRX se relacionan con un comportamiento más agresivo del tumor [46]. También se ha descrito que los genes relacionados con la ruta WNT sólo se encuentran mutados en las células somáticas y por lo tanto servirían de marcador para tumores esporádicos y no familiares [19].

En el caso de los análisis de datos de expresión génica se encontraron datos variables entre los distintos diagnósticos, viéndose en algunos casos relación de las funciones asociadas sólo a un tipo tumoral, por

ejemplo, en el caso de PhM en el que los genes que aparecían diferencialmente expresados estaban relacionados con procesos de regulación, metabolismo y biosíntesis de hormonas esteroideas y glucocorticoides. Y en el grupo de PhNOS se encontraba alguna función de respuesta a estímulos de hormonas esteroideas o la fuerza de contracción del corazón. Estas funciones no aparecen en otros grupos de tumores y esto podría ser porque los feocromocitomas se encuentran en la médula adrenal rodeadas de la corteza adrenal que es donde se producen corticoides y otras hormonas esteroideas.

En el caso de PhNOS, también se encuentran funciones relacionadas con la angiogénesis que pueden estar relacionados con la presencia en estas muestras de mutaciones en los genes EPAS1 y VHL [47]. En general las funciones más representadas en el resto de tumores tenían relación con procesos de metabolismo, adhesión celular y señalización celular, pero con pequeñas diferencias o funciones añadidas en cada grupo.

Con respecto a las diferencias encontradas en metilación del ADN, sólo se encontraron diferencias significativas en el grupo PhNOS, donde se encontraron genes hipometilados del tercer cuartil de los distintos grupos de CNAs comparados con el primer cuartil. Estos genes que aparecen hipometilados en las muestras que cuentan con mayor número de CNAs están relacionados con procesos de diferenciación y adhesión celular, quimiotaxis, organización de la matriz extracelular, procesos de transducción de señales, procesos de migración, procesos metabólicos y respuesta inmune. Está descrito que estos procesos están relacionados con la ruta RTK-Ras, en los análisis de mutaciones de este grupo diagnóstico esta ruta era la que contaba con más genes con mutaciones. [48].

Sin embargo, en este trabajo se ha visto que al subclasificar las muestras hay grupos que quedan un número muy reducido de ellas y varios análisis no se han podido realizar en estos grupos precisamente por este motivo. Además, datos tan sensibles igual podrían cambiar si se aumenta la n de estos pequeños grupos, así que habría que esperar a realizar estos estudios con un número mayor de muestras en algunos grupos para realmente poder sacar unos datos más fiables.

3. Conclusiones

En este trabajo se han llevado a cabo análisis de datos ómicos provenientes de muestras de feocromocitomas y paragangliomas y se ha llegado a las siguientes conclusiones:

1. Se ven diferencias en los resultados de las muestras de feocromocitomas y de paragangliomas como primera agrupación de muestras.
2. También se observan diferencias en los resultados obtenidos según la clasificación utilizada por grupo diagnóstico, por lo que, además de las clasificaciones moleculares descritas previamente, sería conveniente también tener en cuenta la clasificación según el grupo diagnóstico, a la hora de buscar modificaciones genéticas o incluso a la hora del tratamiento de la enfermedad.
3. Es importante tener en cuenta los resultados de varios tipos de datos ómicos, ya que todos pueden aportar información que puede interconectarse e influir finalmente en procesos biológicos que den información del pronóstico de la enfermedad, ya sea porque denotan más agresividad o para guiar sobre la mejor diana del tratamiento.

Se han logrado todos los objetivos planteados para este trabajo, y se ha mantenido la planificación inicial y los plazos marcados, aunque se han quedado un poco más ajustados de lo ideal. Con respecto a la metodología, se ha seguido la planteada en un inicio y los resultados han sido buenos para poder relacionar los distintos datos ómicos.

Aún así hay que admitir, que la idea inicial del trabajo era ahondar más en este tipo de datos, pero no se tuvo bien en cuenta que realizar un análisis de esta envergadura partiendo de un nivel de conocimiento casi nulo en este campo y de manera totalmente independiente, requeriría mucho más tiempo del que se da para realizar este trabajo. Por lo tanto, sería muy interesante poder realizar más análisis de una manera mucho más fina y dirigida a determinadas regiones del genoma o incluso sobre

determinados genes o promotores. Todo ello sirviéndose de más herramientas bioinformáticas y de más información bibliográfica.

4. Glosario

- **ADN:** Ácido desoxirribonucleico.
- **ARN:** Ácido ribonucleico
- **BCS:** Puntuación de alteraciones en el número de copia amplios.
- **CNA:** Alteraciones en el número de copia.
- **EAPaM:** Paraganglioma extra adrenal maligno.
- **EAPaNOS:** Paraganglioma extra adrenal no especificado de otra manera.
- **FCS:** Puntuación de alteraciones en el número de copia focales.
- **Feocromocitoma:** tumor neuroendocrino de células cromafines que se encuentra en la médula de las glándulas adrenales.
- **GCS:** Puntuación de alteraciones en el número de copia generales.
- **MATH:** medida de la heterogeneidad genética intratumoral.
- **Mutaciones con cambio de sentido:** mutación en la que el cambio de un nucleótido resulta en la codificación de un aminoácido distinto en la proteína.
- **PaM:** Paraganglioma maligno.
- **PaNOS:** Paraganglioma no especificado de otra manera.
- **Paraganglioma:** Tumor neuroendocrino de células cromafines que se sitúan fuera de la glándula adrenal, en los paraganglios.
- **PhM:** Feocromocitoma maligno.
- **PhNOS:** Feocromocitoma no especificado de otra manera.
- **PPGL:** Feocromocitoma y paraganglioma.
- **SNP:** Polimorfismo de un único nucleótido.
- **TCGA:** Proyecto “The Cancer Genome Atlas”
- **TCGA-PCPG:** código del proyecto “The Cancer Genome Atlas” dedicado al estudio de los feocromocitomas y los paragangliomas.

5. Bibliografía

- [1] A. Cascón, L. Remacha, B. Calsina, and M. Robledo, "Pheochromocytomas and Paragangliomas: Bypassing Cellular Respiration," (in eng), *Cancers (Basel)*, vol. 11, no. 5, May 2019, doi: 10.3390/cancers11050683.
- [2] M. A. Sevilla García *et al.*, "Paragangliomas de cabeza y cuello: revisión de 89 casos en 73 pacientes," vol. 58, ed. Acta Otorrinolaringológica Española: Elsevier, 2007, pp. 94-100.
- [3] D. L. Kasper, A. S. Fauci, S. L. Auser, D. L. Longo, J. L. Jameson, and J. Loscalzo, "Harrison's Principles of internal medicine," 19th ed: McGraw-Hill Education.
- [4] J. Favier, L. Amar, and A. P. Gimenez-Roqueplo, "Paraganglioma and phaeochromocytoma: from genetics to personalized medicine," (in eng), *Nat Rev Endocrinol*, vol. 11, no. 2, pp. 101-11, Feb 2015, doi: 10.1038/nrendo.2014.188 nrendo.2014.188 [pii].
- [5] J. W. Lenders *et al.*, "Pheochromocytoma and paraganglioma: an endocrine society clinical practice guideline," (in eng), *J Clin Endocrinol Metab*, vol. 99, no. 6, pp. 1915-42, Jun 2014, doi: 10.1210/jc.2014-1498.
- [6] P. F. Plouin *et al.*, "European Society of Endocrinology Clinical Practice Guideline for long-term follow-up of patients operated on for a phaeochromocytoma or a paraganglioma," (in eng), *Eur J Endocrinol*, vol. 174, no. 5, pp. G1-G10, May 2016, doi: 10.1530/eje-16-0033.
- [7] L. Fishbein, S. Merrill, D. L. Fraker, D. L. Cohen, and K. L. Nathanson, "Inherited mutations in pheochromocytoma and paraganglioma: why all patients should be offered genetic testing," (in eng), *Ann Surg Oncol*, vol. 20, no. 5, pp. 1444-50, May 2013, doi: 10.1245/s10434-013-2942-5.
- [8] P. L. Dahia *et al.*, "A HIF1alpha regulatory loop links hypoxia and mitochondrial signals in pheochromocytomas," (in eng), *PLoS Genet*, vol. 1, no. 1, pp. 72-80, Jul 2005, doi: 10.1371/journal.pgen.0010008.
- [9] E. Letouze *et al.*, "SDH mutations establish a hypermethylator phenotype in paraganglioma," (in eng), *Cancer Cell*, vol. 23, no. 6, pp. 739-52, Jun 10 2013, doi: 10.1016/j.ccr.2013.04.018.

- [10] S. Backman *et al.*, "Global DNA Methylation Analysis Identifies Two Discrete clusters of Pheochromocytoma with Distinct Genomic and Genetic Alterations," (in eng), *Sci Rep*, vol. 7, p. 44943, Mar 22 2017, doi: 10.1038/srep44943 srep44943 [pii].
- [11] A. A. de Cubas *et al.*, "DNA Methylation Profiling in Pheochromocytoma and Paraganglioma Reveals Diagnostic and Prognostic Markers," (in eng), *Clin Cancer Res*, vol. 21, no. 13, pp. 3020-30, Jul 1 2015, doi: 10.1158/1078-0432.CCR-14-2804 1078-0432.CCR-14-2804 [pii].
- [12] T. Davoli, H. Uno, E. C. Wooten, and S. J. Elledge, "Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy," (in eng), *Science*, vol. 355, no. 6322, Jan 20 2017, doi: 10.1126/science.aaf8399.
- [13] S. Franch-Expósito *et al.*, "CNApp: a web-based tool for integrative analysis of genomic copy number alterations in cancer," *bioRxiv*, p. 479667, 2018, doi: 10.1101/479667.
- [14] B. R. Du P, "*methyAnalysis: DNA methylation data analysis and visualization* .," ed: R package version 1.26.0., 2019.
- [15] A. Akalin *et al.*, "methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles," (in eng), *Genome Biol*, vol. 13, no. 10, p. R87, Oct 3 2012, doi: 10.1186/gb-2012-13-10-r87.
- [16] <https://portal.gdc.cancer.gov/exploration> (accessed Marzo 2019).
- [17] A. Mayakonda, D. C. Lin, Y. Assenov, C. Plass, and H. P. Koeffler, "Maftools: efficient and comprehensive analysis of somatic variants in cancer," (in eng), *Genome Res*, vol. 28, no. 11, pp. 1747-1756, Nov 2018, doi: 10.1101/gr.239244.118.
- [18] <https://app.teamgantt.com> (accessed 20/FEB/2019).
- [19] L. Fishbein *et al.*, "Comprehensive Molecular Characterization of Pheochromocytoma and Paraganglioma," (in eng), *Cancer Cell*, vol. 31, no. 2, pp. 181-193, Feb 13 2017, doi: 10.1016/j.ccell.2017.01.001.
- [20] R. C. Team, "R: A Language and Environment for Statistical Computing," ed: R Foundation for Statistical Computing, 2019.
- [21] R. Team, "RStudio: Integrated Development for R. ," ed. RStudio, Inc., Boston, MA, 2015.

- [22] A. Colaprico *et al.*, "TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data," (in eng), *Nucleic Acids Res*, vol. 44, no. 8, p. e71, May 5 2016, doi: 10.1093/nar/gkv1507.
- [23] T. C. Silva *et al.*, "TCGA Workflow: Analyze cancer genomics and epigenomics data using Bioconductor packages," (in eng), *F1000Res*, vol. 5, p. 1542, 2016, doi: 10.12688/f1000research.8923.2.
- [24] D. Comtois, "summarytools: Tools to Quickly and Neatly Summarize Data. R package version 0.9.2.," ed, 2019.
- [25] M. Morgan, V. Obenchain, J. i. Hester, and H. Pagès, "SummarizedExperiment: SummarizedExperiment container. R package version 1.12.0.," ed, 2018.
- [26] R. F. Hadley Wickham, Lionel Henry and Kirill Müller, "dplyr: A Grammar of Data Manipulation. R package version 0.8.0.1. ," ed, 2019.
- [27] A. Kassambara, "ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.2.," ed, 2018.
- [28] G. R. Warnes *et al.*, "gplots: Various R Programming Tools for Plotting Data. R package version 3.0.1.1.," ed, 2019.
- [29] H. Wickham, "ggplot2: Elegant Graphics for Data Analysis.," ed. Springer-Verlag New York, 2016.
- [30] E. Neuwirth, "RColorBrewer: ColorBrewer Palettes. R package version 1.1-2.," ed, 2014.
- [31] G. Grothendieck, "sqldf: Manipulate R Data Frames Using SQL. R package version 0.4-11.," ed, 2017.
- [32] Y. Xie, "knitr: A General-Purpose Package for Dynamic Report Generation in R. R package version 1.22.," ed, 2019.
- [33] M. Dowle and A. Srinivasan, "data.table: Extension of `data.frame`. R package version 1.12.2.," ed, 2019.
- [34] F. E. Harrell Jr, "Hmisc: Harrell Miscellaneous. R package version 4.2-0.," ed, 2019.
- [35] Y. Xie, J. Cheng, and X. Tan, "DT: A Wrapper of the JavaScript Library 'DataTables'. R package version 0.5.," ed, 2018.
- [36] S. Milton Bache and H. Wickham, "magrittr: A Forward-Pipe Operator for R. R package version 1.5.," ed, 2014.
- [37] P. DeWitt, "qwraps2: Quick Wraps 2. R package version 0.4.1.," ed, 2019.

- [38] M. Lawrence *et al.*, "Software for computing and annotating genomic ranges," (in eng), *PLoS Comput Biol*, vol. 9, no. 8, p. e1003118, 2013, doi: 10.1371/journal.pcbi.1003118.
- [39] E. Cerami *et al.*, "The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data," (in eng), *Cancer Discov*, vol. 2, no. 5, pp. 401-4, May 2012, doi: 10.1158/2159-8290.cd-12-0095.
- [40] J. Gao *et al.*, "Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal," (in eng), *Sci Signal*, vol. 6, no. 269, p. p11, Apr 2 2013, doi: 10.1126/scisignal.2004088.
- [41] M. Ashburner *et al.*, "Gene ontology: tool for the unification of biology. The Gene Ontology Consortium," (in eng), *Nat Genet*, vol. 25, no. 1, pp. 25-9, May 2000, doi: 10.1038/75556.
- [42] H. Mi *et al.*, "PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements," (in eng), *Nucleic Acids Res*, vol. 45, no. D1, pp. D183-d189, Jan 4 2017, doi: 10.1093/nar/gkw1138.
- [43] The Gene Ontology Consortium, "The Gene Ontology Resource: 20 years and still GOing strong," *Nucleic Acids Research*, vol. 47, no. D1, pp. D330-D338, 2018, doi: 10.1093/nar/gky1055.
- [44] M. S. Lawrence *et al.*, "Mutational heterogeneity in cancer and the search for new cancer-associated genes," (in eng), *Nature*, vol. 499, no. 7457, pp. 214-218, Jul 11 2013, doi: 10.1038/nature12213.
- [45] A. Buffet *et al.*, "Mosaicism in HIF2A-related polycythemia-paraganglioma syndrome," (in eng), *J Clin Endocrinol Metab*, vol. 99, no. 2, pp. E369-73, Feb 2014, doi: 10.1210/jc.2013-2600.
- [46] L. Fishbein *et al.*, "Whole-exome sequencing identifies somatic ATRX mutations in pheochromocytomas and paragangliomas," (in eng), *Nat Commun*, vol. 6, p. 6140, Jan 21 2015, doi: 10.1038/ncomms7140.
- [47] M. Ema, S. Taya, N. Yokotani, K. Sogawa, Y. Matsuda, and Y. Fujii-Kuriyama, "A novel bHLH-PAS factor with close sequence similarity to hypoxia-inducible factor 1alpha regulates the VEGF expression and is potentially involved in lung and vascular development," (in eng), *Proc Natl*

Acad Sci U S A, vol. 94, no. 9, pp. 4273-8, Apr 29 1997, doi: 10.1073/pnas.94.9.4273.

- [48] Y. Pylayeva-Gupta, E. Grabocka, and D. Bar-Sagi, "RAS oncogenes: weaving a tumorigenic web," (in eng), *Nat Rev Cancer*, vol. 11, no. 11, pp. 761-74, 2013, doi: 10.1038/nrc3106.

6. Anexos

1. Documento de R con el código de los análisis realizados.

https://drive.google.com/open?id=1NcAfDFPtDDJGc1de44o6smZ_PGpKjxLy

2. Documento HTML con el código utilizado en el documento de R.

https://drive.google.com/open?id=1NcAfDFPtDDJGc1de44o6smZ_PGpKjxLy