



UNIVERSIDAD SIMÓN BOLÍVAR
DECANATO DE ESTUDIOS PROFESIONALES
COORDINACIÓN DE INGENIERÍA DE LA COMPUTACIÓN

**USO DE DATOS GENÓMICOS DE ALTO RENDIMIENTO EN EL
MODELADO DE LA RESPUESTA DE LÍNEAS CELULARES DE CÁNCER
A DROGAS ANTICÁNCER.**

Por:
Valeria Pestana Gianvittorio

Realizado con la asesoría de:
Prof. Ivette Carolina Martínez

PROYECTO DE GRADO DE INTERCAMBIO ESTUDIANTIL
Presentado ante la Ilustre Universidad Simón Bolívar
como requisito parcial para optar al título de
Ingeniero de Computación

Sartenejas, marzo de 2015



UNIVERSIDAD SIMÓN BOLÍVAR
DECANATO DE ESTUDIOS PROFESIONALES
COORDINACIÓN DE INGENIERÍA DE LA COMPUTACIÓN

**USO DE DATOS GENÓMICOS DE ALTO RENDIMIENTO EN EL
MODELADO DE LA RESPUESTA DE LÍNEAS CELULARES DE CÁNCER
A DROGAS ANTICÁNCER.**

Por:
Valeria Pestana Gianvittorio

Realizado con la asesoría de:
Prof. Ivette Carolina Martínez

PROYECTO DE GRADO DE INTERCAMBIO ESTUDIANTIL
Presentado ante la Ilustre Universidad Simón Bolívar
como requisito parcial para optar al título de
Ingeniero de Computación

Sartenejas, marzo de 2015



UNIVERSIDAD SIMÓN BOLÍVAR
VICERRECTORADO ACADÉMICO
DECANATO DE ESTUDIOS PROFESIONALES
Coordinación de Ingeniería de la Computación

ACTA DE EVALUACIÓN DE
PROYECTO DE GRADO DE INTERCAMBIO ESTUDIANTIL

CÓDIGO DE LA ASIGNATURA: EP3422

FECHA: 24/03/2015

TÍTULO DEL TRABAJO: Uso de DATOS GENÓMICOS DE ALTO RENDIMIENTO EN EL
MODELADO DE LA RESPUESTA DE LÍNEAS CELULARES DE CÁNCER A DROGAS ANTICÁNCER.

NOMBRE DEL ESTUDIANTE: Valeria Pestana Gianvittorio

CARNÉ: 0910655

TUTOR: Prof. Ivette Martínez

JURADO: Profes. Marlene Goncalves y Angela Di Serio

APROBADO:

REPROBADO:

OBSERVACIONES: _____

El jurado considera por unanimidad que el trabajo es EXCEPCIONALMENTE BUENO:

SI: NO: En caso positivo, justificar razonadamente: _____

Marlene Goncalves
Prof. Marlene Goncalves

Angela Di Serio
Prof. Angela Di Serio

Ivette Martinez
Prof. Ivette Martínez

Notas: Colocar los sellos de los respectivos Departamentos Académicos. Para jurados externos usar el sello de la Coordinación. Este documento debe ir sin enmiendas.

Resumen

El propósito de este estudio es evaluar opciones para la posterior construcción de un marco estadístico y computacional confiable que, provisto de una cantidad sustancial de datos *omics* (sobre sistemas complejos como lo son el genoma y el transcriptoma), permita predecir respuestas biológicamente relevantes a terapias anticáncer vía modelos de regresión lineal. Para entrenar los modelos se emplearon conjuntos de datos de gran escala provenientes de publicaciones recientes donde se han reportado respuestas de ciertas líneas celulares de cáncer a un cierto número de drogas; adicionalmente, se utilizaron amplias colecciones de datos de alto rendimiento que permitieron caracterizar a las líneas celulares de cáncer vía perfiles genómicos. Con el fin de neutralizar discrepancias entre los conjuntos de datos y obtener rasgos de mayor nivel, los perfiles ómicos originales fueron transformados vía análisis de enriquecimiento a puntajes cumulativos de ciclos o genes. Se evaluó el desempeño de modelos univariantes y multivariantes ajustados utilizando variables crudas de perfiles y puntajes de enriquecimiento. Se encontró que los puntajes de enriquecimiento superan a los datos crudos en situaciones donde el contenido de información en conjuntos de datos moleculares es pobre, demostrando entonces que pueden hacerse grandes mejoras al momento de usar análisis de enriquecimiento para obtener modelos de predicción de sensibilidad que sean transferibles y biológicamente relevantes.



Be who you are
and say what you feel because
those who mind don't matter
and those who matter don't mind

~ Dr. Seuss

*To my parents Rita and Eduardo,
my sisters Astrid and Martina,
and my partner Benoît.*

AGRADECIMIENTOS

A mi tutor en *Science for Life Laboratories* (SciLifeLab) en Estocolmo, Andrey Alexeyenko, por haberme introducido a esta nueva área de estudios y por haberme dado la oportunidad de llevar a término este proyecto, a la vez proporcionando constante atención y apoyo.

Al profesor Jens Lagergren por haber aceptado ser mi supervisor en KTH y por su amabilidad.

A la profesora Carolina Martínez, por habernos tomado, a mí y a este trabajo, bajo su ala ante la USB.

Al programa de intercambio entre la USB y KTH, por haberme dado la oportunidad de experimentar una maravillosa cultura y su excelente sistema académico.

A mis seres queridos, por proporcionar apoyo incondicional en todo momento.

Índice General

Índice General	vii
Índice de Tablas	ix
Índice de Figuras	xi
Lista de Abreviaturas	xiii
Introducción	1
1. Marco Teórico	5
1.1. Conceptos de Biología	5
1.1.1. <i>Omics</i>	5
1.1.2. Biología de Sistemas	6
1.1.3. Líneas Celulares de Cáncer	7
1.1.4. Ciclos Biológicos	9
1.2. Modelos de Regresión Lineal	10
1.2.1. Mínimos Cuadrados	12
1.2.2. Regresión Penalizada	13
1.2.2.1. Elastic Net	13
1.3. Análisis de Enriquecimiento	14
1.3.1. NEA: <i>Network Enrichment Analysis</i>	15
2. Datos	19
2.1. Sensibilidad a Drogas	20
2.2. Perfiles Crudos	22
2.3. Network Enrichment Analysis	23

2.3.1. Red	23
2.3.2. FGSSs	23
3. Construcción de los Modelos de Predicción	24
3.1. Derivación de AGSSs y Ejecución de NEA y GNEA	25
3.1.1. Ejecución de NEA y GNEA	29
3.2. Modelos Univariantes	32
3.3. Modelos Multivariantes	33
3.3.1. Modelos de Complejidad Limitada	34
4. Experimentos y Resultados	36
4.1. Correlatos Uni-Característica	37
4.2. Exploración de Diferentes Niveles de Complejidad	40
4.3. Modelos para Drogas Compartidas	45
5. Conclusiones	51
Bibliografía	54
A. Organización de los Datos de Trabajo	59
B. Correlatos Uni-Característica	61
C. Exploración de Diferentes Niveles de Complejidad	63
D. Modelos para Drogas Compartidas	74
D.1. Lista de Drogas Comunes	78

Índice de Tablas

1.1. Bases de datos de anotación de genomas.	17
2.1. Autores principales y proyectos asociados a los estudios farmacogenómicos. . .	19
3.1. Convención de nombres, y dimensiones, para datos crudos, NEA y GNEA. .	30
4.1. Fracciones de correlatos uni-característica.	39
B.1. Fracciones de correlatos uni-característica.	61
C.1. Leyenda de figuras para modelos con distintos niveles de complejidad.	64
D.1. Leyenda de tablas para comparación de modelos entre estudios.	76
D.2. Leyenda de figuras para comparación de modelos entre estudios.	76
D.3. Drogas compartidas entre diferentes estudios farmacogenómicos.	78
D.4. Modelos para drogas compartidas, P: Basu, T: Barretina, r_s	79
D.5. Modelos para drogas compartidas, P: Barretina, T: Basu, r_s	79
D.6. Modelos para drogas compartidas, P: Basu, T: Barretina, p -value	80
D.7. Modelos para drogas compartidas, P: Barretina, T: Basu, p -value	80
D.8. Modelos para drogas compartidas, P: Basu, T: Barretina, MSE	81
D.9. Modelos para drogas compartidas, P: Barretina, T: Basu, MSE	81
D.10. Modelos para drogas compartidas, P: Basu, T: Garnett, r_s (1/2)	82
D.11. Modelos para drogas compartidas, P: Basu, T: Garnett, r_s (2/2)	83
D.12. Modelos para drogas compartidas, P: Garnett, T: Basu, r_s (1/2)	84
D.13. Modelos para drogas compartidas, P: Garnett, T: Basu, r_s (2/2)	85
D.14. Modelos para drogas compartidas, P: Basu, T: Garnett, p -value (1/2)	86
D.15. Modelos para drogas compartidas, P: Basu, T: Garnett, p -value (2/2)	87
D.16. Modelos para drogas compartidas, P: Garnett, T: Basu, p -value (1/2)	88

D.17. Modelos para drogas compartidas, P: Garnett, T: Basu, p -value (2/2)	89
D.18. Modelos para drogas compartidas, P: Basu, T: Garnett, MSE (1/2)	90
D.19. Modelos para drogas compartidas, P: Basu, T: Garnett, MSE (2/2)	91
D.20. Modelos para drogas compartidas, P: Garnett, T: Basu, MSE (1/2)	92
D.21. Modelos para drogas compartidas, P: Garnett, T: Basu, MSE (2/2)	93
D.22. Modelos para drogas compartidas, P: Barretina, T: Garnett, r_s	94
D.23. Modelos para drogas compartidas, P: Garnett, T: Barretina, r_s	95
D.24. Modelos para drogas compartidas, P: Barretina, T: Garnett, p -value	96
D.25. Modelos para drogas compartidas, P: Garnett, T: Barretina, p -value	97
D.26. Modelos para drogas compartidas, P: Barretina, T: Garnett, MSE	98
D.27. Modelos para drogas compartidas, P: Garnett, T: Barretina, MSE	99

Índice de Figuras

1.1.	Representación esquemática de un perfil genómico asociado a alguna CCL.	8
1.2.	Ciclo biológico asociado al proceso de apoptosis.	10
1.3. (a)	Esquema sencillo de análisis de enriquecimiento en redes.	16
1.3. (b)	Ejemplo sencillo de una red de interacción de genes.	16
2.1. (a)	Matrices de datos de respuesta (sensibilidad) a drogas.	20
2.1. (b)	Matrices de datos crudos (perfils genómicos).	20
2.2. (a)	Cantidad de drogas compartidas entre los estudios farmacogenómicos.	21
2.2. (b)	Cantidad de CCLs compartidas entre los estudios farmacogenómicos.	21
3.1.	Derivación de AGSs y ejecución de NEA y GNEA.	28
3.2.	Representación esquemática de los datos de entrada para la regresión lineal.	31
4.1.	Diagrama de dispersión de p -valores para la droga Atra.	37
4.2.	Leyenda para interpretar los diagramas de cajas.	39
4.3.	Modelos de diversa complejidad, Basu, r_s	41
4.4.	Modelos de diversa complejidad, Barretina, r_s	42
4.5.	Modelos de diversa complejidad, Garnett, r_s	43
4.6.	Modelos para drogas compartidas, P: Basu, T: Barretina, r_s	48
4.7.	Modelos para drogas compartidas, P: Barretina, T: Basu, r_s	48
4.8.	Modelos para drogas compartidas, P: Basu, T: Garnett, r_s	49
4.9.	Modelos para drogas compartidas, P: Garnett, T: Basu, r_s	49
4.10.	Modelos para drogas compartidas, P: Barretina, T: Garnett, r_s	50
4.11.	Modelos para drogas compartidas, P: Garnett, T: Barretina, r_s	50
A.1.	Esquema de la estructura utilizada para organizar los datos de trabajo.	60

C.1. Modelos de diversa complejidad, <i>Basu</i> , r_s	65
C.2. Modelos de diversa complejidad, <i>Basu</i> , p -value	66
C.3. Modelos de diversa complejidad, <i>Basu</i> , <i>MSE</i>	67
C.4. Modelos de diversa complejidad, <i>Barretina</i> , r_s	68
C.5. Modelos de diversa complejidad, <i>Barretina</i> , p -value	69
C.6. Modelos de diversa complejidad, <i>Barretina</i> , <i>MSE</i>	70
C.7. Modelos de diversa complejidad, <i>Garnett</i> , r_s	71
C.8. Modelos de diversa complejidad, <i>Garnett</i> , p -value	72
C.9. Modelos de diversa complejidad, <i>Garnett</i> , <i>MSE</i>	73
D.1. Modelos para drogas compartidas, P: <i>Basu</i> , T: <i>Barretina</i> , r_s	100
D.2. Modelos para drogas compartidas, P: <i>Barretina</i> , T: <i>Basu</i> , r_s	100
D.3. Modelos para drogas compartidas, P: <i>Basu</i> , T: <i>Barretina</i> , p -value	101
D.4. Modelos para drogas compartidas, P: <i>Barretina</i> , T: <i>Basu</i> , p -value	101
D.5. Modelos para drogas compartidas, P: <i>Basu</i> , T: <i>Barretina</i> , <i>MSE</i>	102
D.6. Modelos para drogas compartidas, P: <i>Barretina</i> , T: <i>Basu</i> , <i>MSE</i>	102
D.7. Modelos para drogas compartidas, P: <i>Basu</i> , T: <i>Garnett</i> , r_s	103
D.8. Modelos para drogas compartidas, P: <i>Garnett</i> , T: <i>Basu</i> , r_s	103
D.9. Modelos para drogas compartidas, P: <i>Basu</i> , T: <i>Garnett</i> , p -value	104
D.10. Modelos para drogas compartidas, P: <i>Garnett</i> , T: <i>Basu</i> , p -value	104
D.11. Modelos para drogas compartidas, P: <i>Basu</i> , T: <i>Garnett</i> , <i>MSE</i>	105
D.12. Modelos para drogas compartidas, P: <i>Garnett</i> , T: <i>Basu</i> , <i>MSE</i>	105
D.13. Modelos para drogas compartidas, P: <i>Barretina</i> , T: <i>Garnett</i> , r_s	106
D.14. Modelos para drogas compartidas, P: <i>Garnett</i> , T: <i>Barretina</i> , r_s	106
D.15. Modelos para drogas compartidas, P: <i>Barretina</i> , T: <i>Garnett</i> , p -value	107
D.16. Modelos para drogas compartidas, P: <i>Garnett</i> , T: <i>Barretina</i> , p -value	107
D.17. Modelos para drogas compartidas, P: <i>Barretina</i> , T: <i>Garnett</i> , <i>MSE</i>	108
D.18. Modelos para drogas compartidas, P: <i>Garnett</i> , T: <i>Barretina</i> , <i>MSE</i>	108

Listas de Abreviaturas

ADN Abreviatura, en español, de ácido desoxirribonucleico.

AGS Abreviatura en inglés de conjunto de genes alterados (*altered gene set*).

ARN Abreviatura, en español, de ácido ribonucleico.

CCL Abreviatura en inglés de línea celular de cáncer (*cancer cell line*).

CCLE Siglas en inglés del proyecto enciclopedia de líneas celulares de cáncer (*Cancer Cell Line Encyclopedia*).

CGP Siglas en inglés del proyecto del genoma del cáncer (*Cancer Genome Project*).

CNA Abreviatura en inglés de alteración en número de copias (*copy number alteration*).

CTRP Siglas en inglés del proyecto portal de respuestas a terapias para el cáncer (*Cancer Therapeutics Response Portal*).

FGS Abreviatura en inglés de grupo funcional de genes (*functional gene set*).

GE Abreviatura en inglés de expresión de gen(es) (*gene expression*).

GO Abreviatura en inglés de ontología de genes (*gene ontology*).

GSEA Abreviatura en inglés de análisis de enriquecimiento de grupos de genes (*gene set enrichment analysis*).

HTS Abreviatura en inglés de secuenciación de alto rendimiento (*high-throughput sequencing*).

KEGG Siglas en inglés de la enciclopedia de genes y genomas de Kyoto (*Kyoto Encyclopedia of Genes and Genomes*).

NEA Abreviatura en inglés de análisis de enriquecimiento en redes (*network enrichment analysis*).

GNEA Abreviatura en inglés de análisis de enriquecimiento en redes aplicado a genes (*gene-wise network enrichment analysis*).

NGS Abreviatura en inglés de secuenciación de próxima generación (*next generation sequencing*).

RNA Abreviatura en inglés de ácido ribonucleico (*ribonucleic acid*).

mRNA Abreviaturas en inglés de ARN mensajero (*messenger RNA*).

ncRNA Abreviaturas en inglés de ARN no codificante (*non coding RNA*).

rRNA Abreviaturas en inglés de ARN ribosomal (*ribosomal RNA*).

tRNA Abreviaturas en inglés de ARN de transporte (*transport RNA*).

TCGA Siglas en inglés del proyecto atlas del genoma del cáncer (*The Cancer Genome Atlas*).

Introducción

Desde sus comienzos en los años 30 la biología molecular, rama de la biología que estudia los fenómenos biológicos a nivel molecular [1], ha triunfado en la interrelación de las áreas de biología, física y química. Descubrimientos importantes, incluyendo la demostración de la relación existente entre genes y proteínas en 1940 [2] y el descubrimiento de la estructura de doble hélice de la molécula de ADN en 1953 [3], tomaron lugar sobre un período de un cuarto de siglo. El surgimiento de tecnologías sofisticadas en las últimas décadas ha llevado a que la biología molecular prospere en cooperación con otras disciplinas como la estadística y la computación.

Actualmente, gracias a los avances en tecnología de secuenciación, está disponible una cantidad masiva de información de alta calidad sobre el genoma, transcriptoma y proteoma. La alta demanda de tecnologías de secuenciación de bajo costo ha impulsado el desarrollo de métodos de alto rendimiento o *high-throughput sequencing* (HTS)¹ logrando la automatización y el incremento en la escala de los experimentos² por medio de la paralelización del proceso de secuenciación.

La repentina baja en los costos de la tecnología, aunado al aumento en la rapidez de procesamiento, ha llevado a un incremento casi exponencial en la cantidad de datos de secuenciación disponibles. En contraste, los costos de cómputo no han bajado tan drásticamente³. El potential de conocimiento contenido en la nueva abundancia de datos está retenido por el cuello de botella que representa el análisis de datos [5].

¹También conocido como *next generation sequencing* (NGS).

²Produciendo hasta miles de secuencia de manera concurrente [4].

³Sobre el mismo período de cuatro años (2007 al 2011), los costos de las tecnologías de secuenciación decrementaron por un factor de 800, mientras que los de cómputo habrán bajado por un factor de quizás cuatro [5].

La terminación de varios proyectos de genoma, el gran incremento en datos *omics*⁴ y los avances en los experimentos de *high-throughput* y en bioinformática permitieron que la disciplina de la biología de sistemas emergiera como un movimiento propio. La biología de sistemas usa un enfoque holístico a la investigación biológica y biomédica, reemplazando la más tradicional perspectiva reduccionista.

Una aplicación contemporánea para la biología de sistemas, en combinación con análisis panómico⁵, es la medicina de precisión, cuya intención es analizar a nivel molecular la causa de la enfermedad de un paciente para proporcionar tratamientos dirigidos a abordar la afección. La medicina de precisión es un nuevo modelo de asistencia médica que parece ser prometedor en el tratamiento de pacientes con cáncer [7] (el cáncer permanece hasta hoy como la principal causa de muerte en el mundo, reclamando 8.2 millones de vidas en el 2012 [8]).

Existen incontables proyectos, iniciativas e investigaciones que representan un esfuerzo en la guerra contra el cáncer⁶; algunos se enfocan únicamente en cerrar la brecha entre los volúmenes colosales de datos y la habilidad de aplicar esta información para mejorar la respuesta de los pacientes, otros buscan producir los mencionados datos o se ocupan de profundizar y mejorar la comprensión de los mecanismos moleculares del cáncer. A la larga, todos están orientados hacia mejorar la respuesta clínica de los pacientes con cáncer.

Este trabajo enfoca su atención en tres artículos recientes. Cada publicación testeó un panel de líneas celulares de cáncer (que reflejan la diversidad genómica de cánceres en humanos, y para las que se contaba con una detallada descripción genética) acopladas con perfiles farmacológicos para diversas drogas anticáncer (algunas que están bajo investigación preclínica⁷) con el objetivo de identificar predictores de sensibilidad a las drogas basándose

⁴Datos sobre sistemas biológicos complejos como el genoma, el transcriptoma, el proteoma, etc.

⁵*Panomics* se refiere al rango de tecnologías de biología molecular, incluyendo *genomics*, *proteomics*, *metabolomics*, *transcriptomics*, etc [6].

⁶Entre ellos: *The Cancer Genome Atlas* (TCGA), *Cancer Genome Project* (CGP), *Cancer Target Discovery and Development* (CTD²).

⁷La etapa de investigación que precede a los ensayos clínicos (pruebas en humanos).

en genética, linaje (del cáncer) y expresión de genes de las líneas de cáncer celular. El uso sistemático de perfiles en estudios farmacogenómicos se espera que acelere el surgimiento de regímenes terapeúticos personalizados mediante el descubrimiento, en la fase preclínica, de predictores genéticos (*biomarkers*) para la respuesta ante las drogas que luego serían incorporados al diseño de ensayos clínicos para el cáncer.

Estudios farmacogenómicos recientes [9–11] han presentado vastos y exhaustivos conjuntos de datos describiendo la respuesta de líneas celulares de cáncer a ciertas drogas anticáncer.

Amplias colecciones de datos *high-throughput* permitieron a los autores explicar parcialmente las respuestas diferenciadas de líneas celulares de cáncer a ciertas drogas, atribuyéndolas a mutaciones, alteraciones en el número de copias de genes y a cambios a nivel de transcripción. A pesar de ser bien recibidos por la comunidad de investigadores, estos trabajos recibieron fuertes críticas debido a las discrepancias de sus resultados [12].

Los datos genómicos entre los estudios [10] y [11] están bien correlacionados; sin embargo, se observa una gran discordancia en los datos de sensibilidad (respuesta) a las drogas. Esto representa un obstáculo si se desea utilizar los recursos asociados para construir o validar modelos de predicción de respuesta a drogas que sean confiables. La fuente de las discrepancias, que aún no se ha determinado, tiene potenciales implicaciones en la posibilidad de usar los resultados para evaluar asociaciones gen-cáncer al momento de seleccionar drogas anticáncer [12].

Este trabajo tiene por objetivo evaluar opciones para la posterior construcción de un marco estadístico y computacional confiable que, provisto de una cantidad sustancial de datos *omics*, permita explorar y predecir la respuesta a drogas (o terapias) anticáncer.

Con la intención de neutralizar las discrepancias en los datos provenientes de diferentes estudios y también de mejorar la explicación y predicción, los perfiles ómicos originales fueron transformados en puntajes cumulativos de ciclos vía *network enrichment analysis* (NEA [13]).

El marco NEA se espera que sea más robusto y sensible debido al uso de conocimiento biológico incrustado en la red global de interacción genes-proteínas. Para implementar esto, las listas de genes alterados son generalizadas al nivel de ciclos por medio de puntajes ciclo-específicos de enriquecimiento, así se obtienen rasgos agregados de mayor nivel; luego, los valores obtenidos son utilizados en lugar de las variables crudas de perfiles genómicos.

El uso de NEA también ha de ayudar al reducir la notoria alta dimensionalidad de características de datos *omics* de unos 20000 genes a unos cientos de ciclos biológicos (circuitos de genes). Métodos de regresión por encogimiento son luego empleados para reducir aún más el número de predictores.

Este trabajo está estructurado en 5 capítulos. En el capítulo 1 se introducen conceptos teóricos que incluyen términos de biología pertinentes al trabajo, definiciones matemáticas y otros métodos utilizados para resolver el problema planteado. El capítulo 2 comprende una descripción de los datos utilizados. En el capítulo 3 se detalla la construcción de los modelos de predicción. En el capítulo 4 se describen los experimentos llevados a cabo con los modelos de predicción y se discuten los resultados obtenidos. Finalmente, en el capítulo 5 se presentan conclusiones y sugerencias para trabajos futuros.

Capítulo 1

Marco Teórico

1.1. Conceptos de Biología

A continuación se presentan una serie de conceptos biológicos que permitirán al lector ubicarse en el contexto del problema y comprender el significado biológico de los datos que se manejan en los capítulos posteriores.

1.1.1. *Omics*

El sufijo ‘-omics’ es utilizado para formar sustantivos que dan nombre a áreas de estudio de la biología, a los objetos de estudio de dichas áreas suele referírsele con el sufijo ‘-oma’⁸. Ambos términos son empleados en el contexto de sistemas grandes y complejos como lo son el genoma (conjunto completo de ADN en un organismo o una célula, incluyendo genes y secuencias no codificantes), el transcriptoma (el conjunto completo de transcriptos de mRNA en un organismo o el conjunto de todas las moléculas de ARN producidas en la célula, incluyendo mRNA, rRNA, tRNA y ncRNA), el proteoma (el conjunto completo de proteínas

⁸Para denotar totalidad de algún tipo (el conjunto entero de una clase de sustancias para una especie o individuo).

expresadas por un organismo), entre otros. Existen diversos tipos de estudios de *omics*, los más relevantes para este trabajo son: *genomics* (estudio del genoma), *transcriptomics* (estudio del transcriptoma) y *proteomics* (estudio del proteoma)⁹.

Las tecnologías *omics* y otros neologismos asociados son más que un simple juego de palabras. Han transformado la escala y el diseño de los experimentos por medio de la generación de copiosas cantidades de datos y de la alteración del proceso fundamental de investigación. Tradicionalmente los experimentos y la recolección de datos obedecen al intento de probar una hipótesis declarada previamente, sin embargo, con el enfoque de *omics* los datos pueden ser recolectados sin haber hecho una pregunta previa; esta inversión en el paradigma de investigación promete el descubrimiento de nuevos mecanismos de respuesta y toxicidad ante las drogas y a la larga mejorará el enfoque simplista y reduccionista de modelos experimentales que hasta ahora sólo ofrecían una fotografía de lo que en realidad es un sistema biológico complejo y dinámico [15].

1.1.2. Biología de Sistemas

La biología de sistemas tiene por objetivo entender las complejas interrelaciones entre entidades biológicas [19]. Esto se logra mediante la integración de datos provenientes de todos los niveles de complejidad biológica (genoma, transcriptoma, proteoma, metaboloma, etc) y el uso de métodos de modelado cuantitativo para modelar matemática y computacionalmente sistemas biológicos complejos. Usualmente la entrada o la salida de estos métodos se presenta en la forma de grafos complejos (redes).

La llegada de la biología de sistemas viene de la mano del desarrollo de tecnologías de secuenciación de alto rendimiento que permiten recolectar las masivas cantidades de datos requeridas por este enfoque; la biología de sistemas también se apoya en el campo de la

⁹Ver en la Bibliografía: [14–18].

bioinformática no sólo para manejar y almacenar los inmensos conjuntos de datos pero por sobre todo para analizarlos (extrayendo principios y conocimientos de los datos).

La biología de sistemas puede entenderse como: *i)* un novedoso paradigma de percepción de procesos que busca entenderlos dentro de un contexto mucho más amplio que antes [19]; en lugar de enfocarse en genes individuales (como se hacía previamente), este enfoque intenta entender cómo los genes y sus productos funcionan juntos en una red (a nivel celular por ejemplo). *ii)* Un campo de estudio que emplea un enfoque holístico a la investigación biológica y biomédica.

1.1.3. Líneas Celulares de Cáncer

El alcance actual de las investigaciones sobre el cáncer demanda de material de tumores animales y humanos. Los modelos *in vitro* son ampliamente utilizados hoy en día en estudios experimentales [20].

Las líneas celulares de cáncer (CCLs, del inglés *cancer cell lines*) son células originarias de tumores de cáncer que crecen en un cultivo de tejido por varias generaciones [21]. Se dividen en el tiempo bajo ciertas condiciones *in vitro* [22], siendo por décadas un pilar de la biología de tumores y del descubrimiento de drogas [10].

La probabilidad de que el cáncer de un paciente responda a un tratamiento puede determinarse a partir de las alteraciones en el genoma somático del cáncer. Por ello, uno de los principales intereses de la investigación en torno al cáncer es la asociación de drogas anticáncer a alteraciones genómicas específicas que podrían potencialmente ser utilizadas como indicadores de efectividad terapéutica (*biomarkers*) [11].

La caracterización de CCLs con perfiles genómicos se ha utilizado para revelar patrones de sensibilidad de CCLs a compuestos químicos. Por conveniencia y simplicidad, en este trabajo se define un perfil genómico como un vector fila ‘línea celular-específico’ de valores asociados

g_1	g_2	g_3	\dots								\dots	g_p
v_1	v_2	v_3	\dots								\dots	v_p

Figura 1.1: Representación esquemática de un perfil genómico asociado a alguna CCL. Un **valor** v_i es la medida de alguna característica (expresión, número de copias, mutación) del **gen** g_i en la CCL.

a genes, el valor en cuestión es una característica medida del gen respectivo en dicha línea celular; la figura 1.1 muestra una representación esquemática de un perfil genómico, de alguna característica particular, que podría caracterizar a una CCL. Las características con las que se trataron en este trabajo son:

- Expresión de Genes (GE¹⁰): es el proceso de utilización de la información codificada en un gen para sintetizar un producto funcional de un gen (una proteína o una molécula funcional de ARN) [23]. La alteración, dentro de una célula, de la expresión de algunos genes podría desencadenar una serie de efectos secundarios sobre la expresión de otros genes; que una célula se vuelva maligna se atribuye a la agregación de todos estos cambios [24].
- Alteración en el Número de Copias (CNA¹¹): se define como una alteración en el número de copias de un gen en una región genómica, comparado con la mayoría de los individuos en la población [25]; suele ser resultado de mutaciones somáticas que ocurren durante el desarrollo clonal del tumor [26].
- Mutación: es un cambio permanente en la secuencia de nucleótidos del genoma de una célula; las mutaciones pueden ser el resultado de errores durante la división celular o de daños al ADN no reparados (causados por algún agente en el ambiente) [27].

¹⁰Del inglés *gene expression*.

¹¹Del inglés *copy number alteration*.

1.1.4. Ciclos Biológicos

Un ciclo biológico es un circuito de relaciones entre moléculas en una célula que necesariamente incluye macromoléculas (genes, proteínas) y permite una cierta producción metabólica o la regulación de un estado celular. Un ciclo puede desencadenar el ensamblado de nuevas moléculas, activar o desactivar genes o disparar un proceso global en la célula (p. ej. la muerte celular o apoptosis, un diagrama del ciclo asociado puede observarse en la figura 1.2). Esta variedad de funciones da pie a múltiples formas de clasificar ciclos moleculares [28].

La farmacogenómica surgió en los años 50 como una ciencia interesada en estudiar e investigar las implicaciones farmacológicas de mutaciones individuales de genes. La observación de que la mayoría de las respuestas a las drogas parecen ser multifactoriales despertó el interés de los científicos por la farmacogenómica [29].

Hasta hace poco tiempo, muchos investigadores esperaban que los cánceres fueran impulsados por errores genéticos únicos y recurrentes, de modo que podrían ser tratados con drogas diseñadas para atacar estos eventos específicos. Desafortunadamente, este enfoque ha fallado para muchos tipos de cáncer. Estudios recientes que descifraron genomas de células de cáncer encontraron un rango de diferentes mutaciones que podrían provocar el mismo tipo de cáncer en distintos pacientes (datos generados por [30]). Sin considerar perfiles genómicos de tumores de cada individuo, sólo se puede aspirar, a lo sumo, a una droga o combinación de drogas que funcione para la mayoría de los pacientes (pero no para todos).

Luchar contra muchos enemigos genéticos es una tarea abrumadora. Se espera que el estudio de este complejo panorama molecular pueda facilitarse al generalizar las mutaciones genéticas perturbadoras a ciclos biológicos importantes. En lugar de diseñar drogas que ataquen cada mutación genética específica, los desarrolladores de drogas podrían enfocar su atención en sólo unos pocos ciclos biológicos. Los pacientes posteriormente tomarían las drogas que más probablemente funcionen para ellos, dado el patrón individual de ciclos.

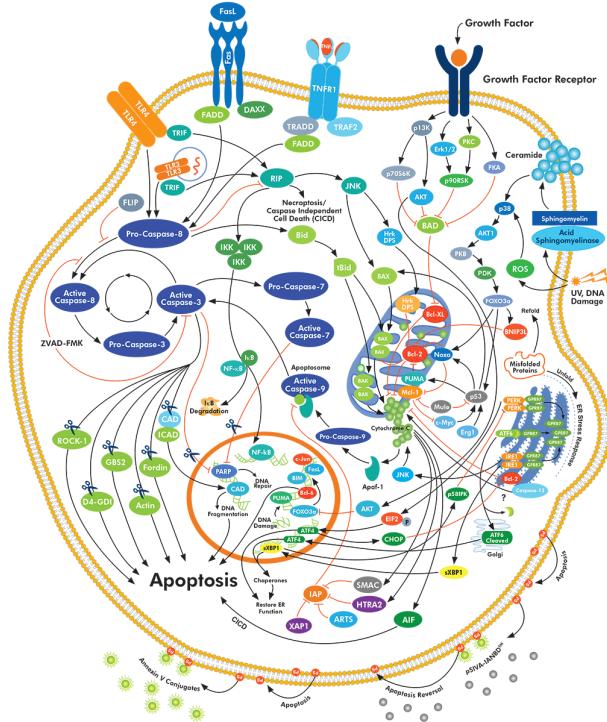


Figura 1.2: Ciclo biológico asociado al proceso de apoptosis (muerte celular). En el cáncer se evidencia una proliferación celular excesiva que puede resultar de la inhibición del proceso de apoptosis a causa de genes participantes en el ciclo (p. ej. p53, ras, c-myc) que se encuentran mutados o desactivados.

1.2. Modelos de Regresión Lineal

La regresión lineal es un método utilizado para modelar la relación entre una variable dependiente y y una o más variables explicativas X . Dado un conjunto de n observaciones estadísticas y p predictores, $\{y_i, x_{i1}, x_{i2}, \dots, x_{ip}\}_{i=1}^n$, un modelo lineal (que asume una relación lineal entre y_i y los x_{ij} con $j = 1, \dots, p$) toma la siguiente forma:

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \epsilon_i \quad i = 1, \dots, n \quad (1.1)$$

el conjunto de n ecuaciones suele expresarse como:

$$y = X\beta + \epsilon \quad (1.2)$$

donde

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} X_1^T \\ X_2^T \\ \vdots \\ X_n^T \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}, \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix} \quad (1.3)$$

- y_i es la variable de respuesta (o variable dependiente).
- $x_{i1}, x_{i2}, \dots, x_{ip}$ son variables explicativas (regresoras, predictoras o independientes). En la matriz de diseño, X , las filas representan observaciones y las columnas predictores. Se suele incluir una constante, por ello $x_{i0} = 1$ para $i = 1, \dots, n$.
- β es el vector parámetro, sus elementos se denominan coeficientes de regresión.
- ϵ_i es el término de error (o perturbación), una variable aleatoria que añade ruido y ha de capturar otros factores, aparte de x_{ij} , $j = 1, \dots, p$, que influencian a la variable dependiente y_i .

La regresión lineal busca la curva que mejor se ajusta a un conjunto de mediciones, i. e. una línea en una regresión simple (una sola variable explicativa), una superficie (dos variables explicativas), etc.

1.2.1. Mínimos Cuadrados

Los modelos de regresión lineal suelen ajustarse usando el método de los mínimos cuadrados, que busca minimizar la suma del cuadrado de los residuos¹². Para un conjunto de n observaciones (incluyendo medida del resultado) y p mediciones de entrada, $\{y_i, x_{i1}, x_{i2}, \dots, x_{ip}\}_{i=1}^n$, mínimos cuadrados ajusta la función modelo $f(X, \beta) = X\beta$ (con p parámetros ajustables que vendrían siendo los elementos de β). El mejor modelo (óptimo) es el que minimiza la norma Euclídea (norma ℓ^2) del vector de residuos $r = f(X, \beta) - y$.

$$\min_{\beta} \|r\|_2 = \min_{\beta} \|X\beta - y\|_2 = \min_{\beta} \left\| \begin{pmatrix} \left(\sum_{j=0}^p x_{1j} \beta_j \right) - y_1 \\ \left(\sum_{j=0}^p x_{2j} \beta_j \right) - y_2 \\ \vdots \\ \left(\sum_{j=0}^p x_{nj} \beta_j \right) - y_n \end{pmatrix} \right\|_2 = \min_{\beta} \sqrt{\sum_{i=1}^n \left[\left(\sum_{j=0}^p x_{ij} \beta_j \right) - y_i \right]^2} \quad (1.4)$$

como al minimizar una función también se minimiza su raíz cuadrada, basta con minimizar la cantidad dentro de la raíz; es equivalente decir entonces que el mejor modelo es aquel que minimiza la suma S de los residuos r_i al cuadrado.

$$r_i = f(X_i^T, \beta) - y_i = \left(\sum_{j=0}^p x_{ij} \beta_j \right) - y_i \quad \text{para } i = 1, \dots, n \quad (1.5)$$

$$S = \sum_{i=1}^n r_i^2 = \sum_{i=1}^n \left[\left(\sum_{j=0}^p x_{ij} \beta_j \right) - y_i \right]^2 \quad (1.6)$$

La solución a un problema de mínimos cuadrados puede encontrarse resolviendo el sistema de ecuaciones $y = X\beta$, de modo que la solución analítica (recuérdese que se desea hallar β) sería $\beta = (X^T X)^{-1} X^T y$.

¹²El residuo es la diferencia entre el valor observado y el estimado.

1.2.2. Regresión Penalizada

En ocasiones el método de mínimos cuadrados no produce resultados satisfactorios debido a [31]: *i*) Cuando $p \gg n$, o cuando las columnas de X presentan alta correlación, la varianza de β es grande. *ii*) Si p es muy grande, puede que se desee un conjunto reducido de predictores que permita dilucidar las relaciones más relevantes entre y y X .

Algunas alternativas para solventar el problema incluyen [31]: *i*) Escoger un subconjunto de predictores importantes basado en métodos *ad-hoc* de ensayo y error (p. ej. guiarse por el p -valor para decidir si una variable ha de pertenecer o no al modelo); el problema de este enfoque es que no puede ser replicado. *ii*) Un posible enfoque sistemático consiste en revisar exhaustivamente todos los posibles subconjuntos de regresores y emplear algún criterio para elegir el subconjunto óptimo; el problema es que el número de subconjuntos crece de forma exponencial conforme crece p y cambios infinitesimalmente pequeños en los datos pueden resultar en estimados completamente diferentes. *iii*) Usar regresión penalizada, esta permite lograr el objetivo de la selección de subconjuntos de manera más estable y eficiente.

La regresión penalizada permite simultáneamente seleccionar variables y estimar parámetros, esto la hace una herramienta atractiva para análisis estadístico. Los métodos de regresión penalizada presentan ventajas en selección de variables y predicción de respuesta cuando se trabaja con datos de alta dimensión (como es el caso de este estudio donde $p \gg n$). A continuación se describe el método de regresión penalizada utilizado en el presente trabajo:

1.2.2.1. Elastic Net

Elastic net (red elástica) es un método de regresión penalizada que minimiza la suma de residuos cuadrados sujeta a una cota sobre una penalización que es una combinación lineal de las normas ℓ^1 y ℓ^2 del vector de coeficientes β [32].

$$\min_{\beta} \|y - X\beta\|_2 \quad \text{s.t.} \quad (1 - \alpha)\|\beta\|_1 + \alpha\|\beta\|_2 \leq t \quad \text{para algún } t \quad (1.7)$$

El parámetro α controla la combinación convexa de las penalidades ℓ^1 y ℓ^2 . *Elastic net* realiza de manera simultánea selección de variables y encogimiento continuo y además es capaz de seleccionar grupos de variables correlacionadas gracias a la penalidad ℓ^2 [32, 33]. *Elastic net* es como una malla de pescar elástica que retiene todos los “peces grandes” [33].

- La parte ℓ^1 de la penalidad permite la generación de soluciones dispersas y la superación de la dificultad $p \gg n$.
- La parte ℓ^2 de la penalidad elimina la limitación sobre el número de variables seleccionadas a la vez que promueve el efecto de agrupamiento¹³.

1.3. Análisis de Enriquecimiento

El análisis de enriquecimiento permite la caracterización biológica de conjuntos experimentales de genes mediante la búsqueda de categorías funcionales conocidas, p. ej. ciclos biológicos, que se encuentran sobre o bajo representadas en el conjunto derivado experimentalmente; es probable obtener un mejor entendimiento biológico al estudiar grupos de genes en lugar de analizar genes de manera individual.

Típicamente, un conjunto de genes alterados (AGS, del inglés *altered gene set*) derivado experimentalmente suele ser una lista de genes, de un perfil, expresados diferentemente y ordenados por algún criterio (ej. ver los genes ordenados del perfil en la figura 1.3 (a)); la alteración generalmente está asociada a un estado biológico particular, p. ej. un tumor.

El análisis de enriquecimiento identifica cuáles grupos funcionales de genes (FGS, del inglés *functional gene sets*) previamente conocidos (ej. el conjunto de **genes anaranjados** en la figura 1.3 (a)), están sobre o bajo representados en un AGS; se dice que hay enriquecimiento cuando el FGS está sobre representado en el AGS; si está bajo representado se dice que hay

¹³Evidenciado en modelos lineales donde los coeficientes de variables altamente correlacionadas tienden a ser iguales (hasta un cambio de signo si la correlación es negativa).

agotamiento. La valoración de enriquecimiento o agotamiento se basa en alguna medida estadística utilizada por el método particular de enriquecimiento, típicamente suele utilizarse un z -valor (de modo que un z -valor mayor a un valor umbral indica enriquecimiento, si es menor a dicho umbral indica que hay agotamiento); ej. z -valor basado en la cantidad de genes solapados entre el FGS y el AGS o en la cantidad de enlaces entre genes del FGS y genes del AGS (ej. enlaces morados en la figura 1.3 (a)). La tabla 1.1 lista algunas bases de datos que contienen FGSs utilizados en el análisis de enriquecimiento; incluyen categorías biológicas informativas (p. ej. ciclos biológicos, conjuntos de genes, términos de vocabulario de ontología de genes y productos de genes) que han sido estudiadas, descritas y caracterizadas a fondo.

Existen varias herramientas y métodos para realizar análisis de enriquecimiento, cada una emplea algún test de estadística para evaluar enriquecimiento (y agotamiento).

1.3.1. NEA: *Network Enrichment Analysis*

Network enrichment analysis (NEA) puede verse como una extensión de *gene set enrichment analysis* (GSEA)¹⁴ en el contexto de una red. NEA extiende la estadística de superposición (o solapado) utilizada por GSEA a enlaces de la red entre genes en los AGSs y genes en los FGSs. NEA cuantifica la sobre o bajo representación de FGSs en los vecinos de red del AGS, en lugar de en el AGS en si.

Este enfoque integra dos tipos de información biológica: *i*) información funcional conocida disponible para genes, i. e. FGSs y *ii*) conectividad de la red para casi todos los genes que codifican proteínas. Incorporar información topológica de la red de interacción de genes aumenta el número de enriquecimientos potencialmente detectables entre grupos de genes, haciendo el poder estadístico de NEA superior al de GSEA. El enfoque NEA permite explotar el valioso y creciente conocimiento sobre las interacciones biológicas a diferentes niveles¹⁵.

¹⁴GSEA es una estrategia común para realizar análisis de enriquecimiento.

¹⁵Incluyendo ADN, ARN y proteínas.

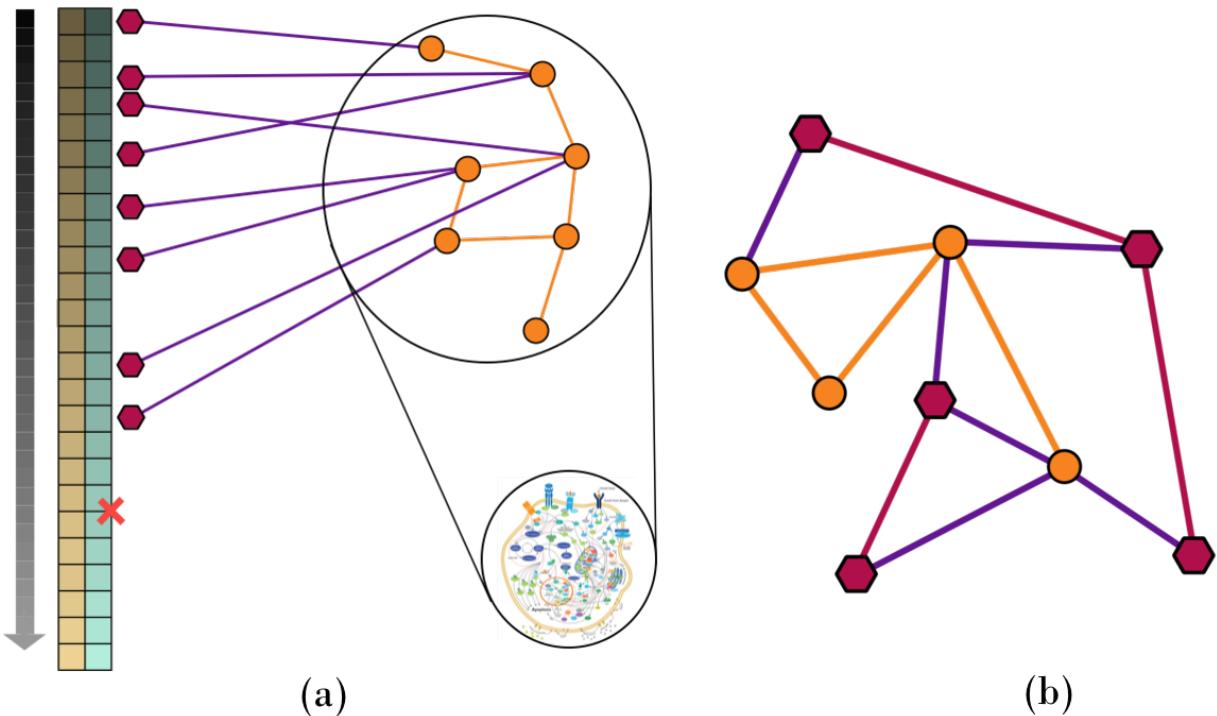


Figura 1.3: (a) Esquema sencillo de análisis de enriquecimiento en redes. Los **genes** del perfil (a la izquierda) están ordenados por alguna métrica, forman parte del AGS aquellos con un **valor** superior al valor umbral indicado con una **x**. Los **círculos anaranjados** son genes de un FGS (asociado al ciclo biológico de la apoptosis), los **hexágonos rosados** son genes del AGS asociados a genes del FGS y los **enlaces morados** son enlaces entre genes del AGS y genes del FGS.

Figura 1.3: (b) Ejemplo sencillo de una red de interacción de genes.

- Los **hexágonos rosados** son genes del AGS y los **arcos rosados** son enlaces internos dentro del AGS (cuya cantidad es igual a N_{AGS} , ver la ecuación 1.9).
 - Los **círculos anaranjados** son genes del FGS y los **arcos anaranjados** son enlaces internos dentro del FGS (cuya cantidad es igual a N_{FGS} , ver la ecuación 1.9).
 - Los **enlaces morados** son enlaces entre genes del AGS y genes del FGS (cuya cantidad es igual a $n_{AGS-FGS}$, ver la ecuación 1.8).

Ejemplo de cálculo del puntaje de enriquecimiento z para el AGS y el FGS representados:

- Se sustituyen en la ecuación 1.9 $N_{AGS} = 3$, $N_{FGS} = 4$ y $N = 13$.

$$\hat{n}_{AGS-FGS} = \frac{4 \times 3}{13} = \frac{6}{13}$$

- Se sustituyen en la ecuación 1.8 $n_{AGS-FGS} = 6$, $\hat{n}_{AGS-FGS} = 6/13$ y $\sigma_{AGS-FGS} = 3\sqrt{10}/13$ (calculado a partir de N_{global} , N_{AGS} y N_{FGS}).

$$z = \frac{6 - \frac{6}{\frac{3\sqrt{10}}{13}}}{\frac{3\sqrt{10}}{13}} = \frac{12}{5}\sqrt{10} \approx 7,589$$

Base de Datos	Descripción
KEGG [34, 35] <i>Kyoto Encyclopedia of Genes and Genomes</i>	<ul style="list-style-type: none"> • Ciclos Metabólicos.
<i>Gene Ontology</i> (GO) [36]	<ul style="list-style-type: none"> • Vocabulario controlado para genes y productos de genes.
MSigDB [37] <i>Molecular Signatures Database</i>	<ul style="list-style-type: none"> • Colección de grupos de genes anotados, para usarse con software de GSEA.

Tabla 1.1: Bases de datos de anotación de genomas. Estas bases de datos contienen FGSs utilizados para realizar análisis de enriquecimiento.

El puntaje de enriquecimiento en NEA es una estadística z que transmite el excedente de conectividad en la red entre un AGS y un FGS por encima de lo que se esperaría por azar. Para inferencia estadística NEA usa el modelo de muestreo hipergeométrico para obtener la distribución de la conectividad de red bajo la hipótesis nula de no asociación entre los AGSs y los FGSs.

Para estimar probabilísticamente relaciones funcionales aparentes entre conjuntos de genes con NEA, puede utilizarse el z -valor estándar para la conectividad de la red biológica entre la lista nueva (AGS) y el FGS (comúnmente un ciclo o término de GO) [13]. Este z -valor caracteriza probabilísticamente el número de enlaces (arcos) en la red global que conectan genes del AGS con genes del FGS (enlaces **morados** en la figura 1.3 (b)), según [13,38], puede computarse a partir del número de enlaces observados¹⁶ ($n_{AGS-FGS}$) y esperados ($\hat{n}_{AGS-FGS}$) y su desviación estándar ($\sigma_{AGS-FGS}$):

$$z = \frac{n_{AGS-FGS} - \hat{n}_{AGS-FGS}}{\sigma_{AGS-FGS}} \quad (1.8)$$

Un ejemplo del cálculo de este valor puede observarse en la descripción de la figura 1.3 (b).

¹⁶Enlaces **morados** en la figura 1.3 (b).

El uso del z -valor se justifica mientras se asuma una distribución normal, para la cual la transformación a p -valores es trivial, sin embargo, si el AGS o el FGS contiene muy pocos genes (p. ej. menos de 10) o los genes tienen una conectividad muy baja en general, entonces el supuesto de normalidad podría verse violado. Para estimar el número de enlaces, entre un AGS y un FGS, que se esperaría observar por azar, en lugar de aleatorizar la red, se hace uso de una fórmula binomial más simple (ver ejemplo en la descripción de la figura 1.3 (b)):

$$\hat{n}_{AGS-FGS} = \frac{N_{AGS}N_{FGS}}{2N_{global}} \quad (1.9)$$

que emplea valores cumulativos de conectividad (número total de enlaces entre todos los genes de la red completa; esto sería en la figura 1.3 (b) los enlaces **rosados**, internos del AGS, los enlaces **anaranjados**, internos del FGS, y los enlaces **morados**, de la interacción entre genes en el AGS y genes en el FGS). La desviación de este valor con respecto al observado ($n_{AGS-FGS}$) sigue una distribución chi cuadrado que no se ve afectada por N s pequeños. Además, es veloz y no requiere de múltiples rondas de aleatorización de la red como propone [39]. Es importante acotar que esta simplificación sólo es aplicable al momento de cuantificar enlaces directos; evaluar otros patrones de red más complicados requeriría aleatorizar la red. Los valores χ^2 fueron transformados en p -valores y luego transformados de vuelta a z -valores; como los últimos siguen una distribución normal bajo la hipótesis nula, el procedimiento permite utilizar las matrices resultantes de puntajes z como entrada para una regresión lineal.

Además de relaciones AGS-FGS, NEA puede analizar relaciones entre un gen individual y un grupo de genes, cosa que no es posible con *Gene Set Enrichment Analysis* (GSEA) por ejemplo. En esta modalidad de NEA, que se denota por GNEA (del inglés *gene-wise* NEA), el FGS se presenta como un único gen, y los cálculos son idénticos a los de las ecuaciones 1.8 y 1.9. Esto permite identificar genes específicos, incluyendo aquellos que aún no han sido asignados a alguna categoría funcional (ciclo biológico, término de ontología, etc).

Capítulo 2

Datos

Se detallan en este capítulo los datos utilizados en los procedimientos descritos en el capítulo 3. Los datos de sensibilidad a drogas y de perfiles genómicos (i. e. expresión de genes, alteración en el número de copias y mutaciones), asociados a los estudios publicados por Basu [9], Barretina [10] y Garnett [11] fueron descargados de CTRP (<http://www.broadinstitute.org/ctrp>), CCLE (<http://www.broadinstitute.org/ccle/>) y CGP (<http://www.cancerrxgene.org>) respectivamente. La red y la colección de FGSs para NEA [13] fueron descargadas de http://research.scilifelab.se/andrey_alexeyenko/downloads.html. La tabla 2.1 lista el autor principal y el proyecto asociado a cada uno de los estudios farmacogenómicos considerados en el presente trabajo.

Publicación	Siglas del Proyecto	Nombre del Proyecto
Basu et al. [9]	CTRP	<i>Cancer Therapeutics Response Portal</i> ¹⁷
Barretina et al. [10]	CCLE	<i>Cancer Cell Line Encyclopedia</i> ¹⁸
Garnett et al. [11]	CGP	<i>Cancer Genome Project</i> ¹⁹

Tabla 2.1: Autores principales y proyectos asociados a los estudios farmacogenómicos.

¹⁷Portal de Respuestas a Terapias de Cáncer, <http://www.broadinstitute.org/ctrp/>

¹⁸Enciclopedia de Líneas Celulares de Cáncer, <http://www.broadinstitute.org/ccle/home>

¹⁹Proyecto del Genoma del Cáncer, <https://www.sanger.ac.uk/research/projects/cancergenome/>

2.1. Sensibilidad a Drogas

Basu [9], Barretina [10] y Garnett [11] condujeron estudios cuantitativos donde midieron la respuesta de CCLs (caracterizadas genómicamente) al ser acopladas con un grupo de drogas anticáncer seleccionadas²⁰ (i. e. midieron el efecto de un grupo seleccionado de drogas anticáncer sobre CCLs caracterizadas genómicamente).

Para cada estudio se compiló una matriz con los datos de respuesta a drogas, cada fila en la matriz corresponde a una observación o muestra (una CCL) y cada columna corresponde a un compuesto o droga; ver figura 2.1 (a). El elemento a_{ij} es la medida de sensibilidad o respuesta de la i -ésima CCL a la j -ésima terapia o droga (i. e. el efecto de la j -ésima droga sobre la i -ésima CCL). Una columna en la matriz entonces provee los valores de respuesta de todas las CCLs ante una droga específica, esto significa que una columna de la matriz puede emplearse como variable dependiente o de respuesta (vector y) al ejecutar regresión lineal.

Se denotan estas matrices de datos de respuesta por Basu, Barretina y Garnett.

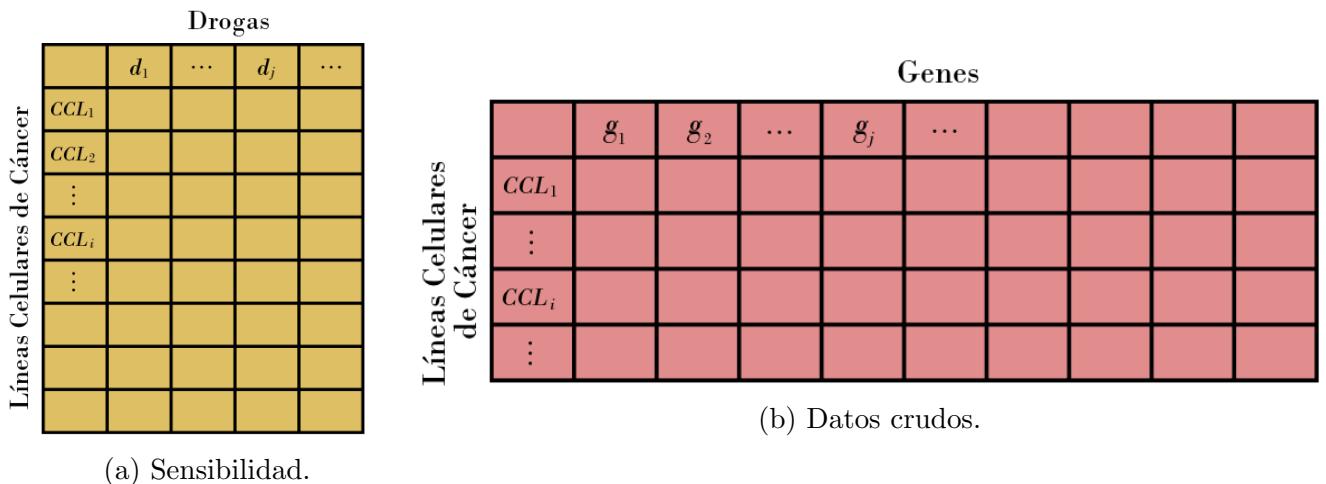


Figura 2.1: (a) Matrices de datos de respuesta (sensibilidad) a drogas.

Figura 2.1: (b) Matrices de datos crudos (perfiles genómicos).

²⁰Ver los artículos publicados por cada estudio para obtener detalles sobre las mediciones y el ajuste de curvas para los datos de respuesta a drogas.

Las dimensiones, por estudio, de cada **matriz de sensibilidad** (figura 2.1 (a)) son:

- Basu: 242 CCLs \times 203 drogas
- Barretina: 504 CCLs \times 24 drogas
- Garnett: 712 CCLs \times 138 drogas

Los nombres de drogas y CCLs en **Basu**, **Barretina** y **Garnett** fueron curados manualmente con la finalidad de hacer una correspondencia entre compuestos y CCLs idénticos. Las figuras 2.2 (a) y (b) muestran el número de drogas y CCLs compartidas²¹ entre los tres estudios.

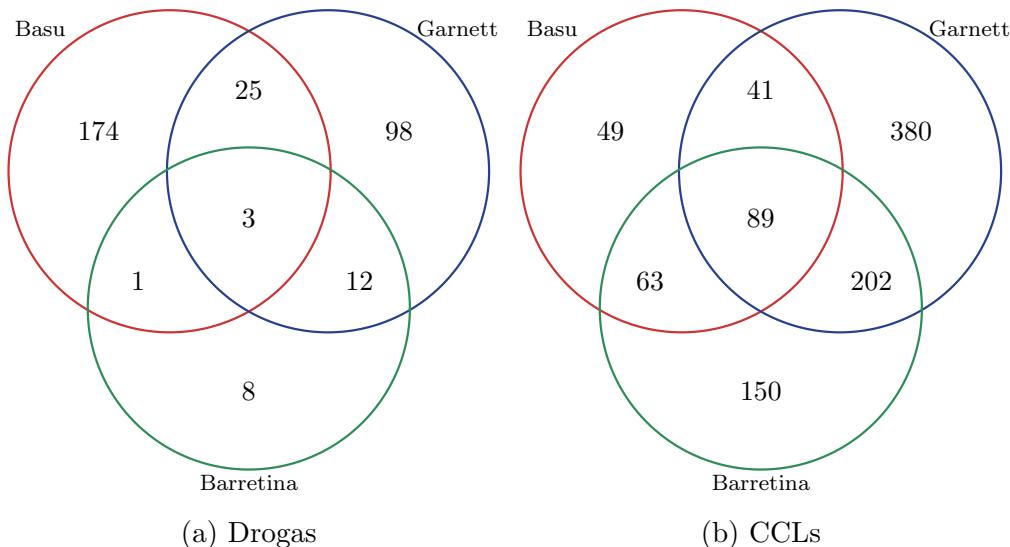


Figura 2.2: (a) Cantidad de drogas compartidas entre los estudios farmacogenómicos.

Figura 2.2: (b) Cantidad de CCLs compartidas entre los estudios farmacogenómicos.

²¹Ver la tabla D.3 en el apéndice D.1 para mayor detalle con respecto a las drogas comunes.

2.2. Perfiles Crudos

En este trabajo se usa el término datos crudos para referirse a datos de perfiles genómicos que no han sido sometidos a algún tipo de análisis de enriquecimiento. Este estudio emplea perfiles de genes (para expresión de genes, alteraciones en el número de copias y mutaciones) tomados todos de la CCLE. Para cada tipo de perfil, los datos fueron organizados en una matriz en donde las filas son perfiles genómicos. Por conveniencia, se utiliza `affymetrix1`, `snp6` y `mgs` para hacer referencia a las matrices de perfiles de GE, CNA y mutaciones.

Para cada tipo de perfil se compiló una matriz con las características medidas. Cada fila en la matriz corresponde a una observación o muestra (una CCL) y cada columna corresponde a un gen; ver figura 2.1 (b). El elemento a_{ij} es el valor de la característica (GE, CNA o mutación) del j -ésimo gen en la i -ésima CCL.

Las dimensiones, por tipo de perfil, de cada **matriz de datos crudos** (figura 2.1 (b)):

- Expresión de Genes (`affymetrix1`): 1034 CCLs \times 18900 genes
- Alteraciones en el Número de Copias (`snp6`): 994 CCLs \times 21217 genes
- Mutaciones (`mgs`): 904 CCLs \times 1667 genes

En aras de la simplicidad, este trabajo sólo considera datos binarios sobre mutaciones, es decir, la presencia o ausencia de una mutación en un gen. Por dicha razón la matriz de perfiles de mutaciones (`mgs`) es booleana (i. e. los datos son de carácter binario, cada entrada de la matriz toma un valor en el conjunto $\{0, 1\}$); el elemento a_{ij} indica si el j -ésimo gen en la i -ésima CCL está mutado ($a_{ij} = 1$) o no ($a_{ij} = 0$).

2.3. Network Enrichment Analysis

La entrada requerida por NEA incluye una lista de AGSs derivada de los datos de perfiles (en el capítulo 3.1 se describe como fueron derivados los AGSs para este estudio), una lista de FGSS para los cuales se desea evaluar su enriquecimiento o agotamiento en los AGSs y por último, una red de interacción de genes y productos de genes.

2.3.1. Red

La red utilizada en este trabajo resulta de la combinación de cuatro redes (incluyendo la red de datos de interacción de *functional coupling* [40], ciclos curados de KEGG, complejos de proteínas conocidos y anotaciones sopesadas con términos de GO) [13]. La red final contiene 1445027 enlaces funcionales entre más de 16000 genes distintos [13].

2.3.2. FGSS

Se utilizó una lista de 330 FGSS para caracterizar los AGSs. Dicha lista incluye ciclos de la base de datos de KEGG, términos de GO relacionados con los sellos distintivos del cáncer (*Hallmarks of Cancer*²²) [41], ciclos biológicos relacionados con el cáncer, otros FGSS tomados de estudios que reportan sobre proyectos de gran escala acerca del genoma del cáncer, un ciclo responsable del desarrollo metastásico (EMT, *epithelial-mesenchymal transition*) y un ciclo responsable del cambio en ph de tumores específicos, que da a un tumor la habilidad de crecer en un ambiente hipoxíco²³, (interruptor acídico, en inglés *acidic switch*).

²²Seis características (sellos, *hallmarks*), comunes a todos los cánceres, que gobiernan la transformación de una célula normal a un tumor maligno [41].

²³Bajas condiciones de oxígeno.

Capítulo 3

Construcción de los Modelos de Predicción

Los procedimientos necesarios para la construcción de modelos de predicción fueron ejecutados en una computadora con un procesador Intel Core 2 Duo 2.53 GHz (64-bit) y 8 GB de memoria RAM, bajo el sistema operativo Mac OS X 10.6.8.

Se utilizó el lenguaje de programación R para realizar la manipulación y análisis de datos que se llevó a cabo bajo el ambiente de desarrollo integrado RStudio. El análisis de enriquecimiento se llevó a cabo utilizando **NEArender**, un nuevo paquete de R (aún bajo desarrollo), y los modelos de *elastic net* se obtuvieron utilizando la librería **glmnet** [42] que implementa un algoritmo rápido para estimar modelos de regresión lineal generalizados con penalidades convexas vía descenso de coordenadas computado a lo largo del camino de regularización [43].

En este capítulo se detalla cómo a partir de las matrices de datos crudos de perfiles genómicos se obtienen, vía análisis de enriquecimiento, matrices de datos enriquecidos (puntajes de enriquecimiento z). Posteriormente se describe el proceso utilizado para construir modelos lineales (univariantes y multivariantes) de predicción de sensibilidad.

3.1. Derivación de AGSs y Ejecución de NEA y GNEA

La intención de derivar un conjunto de genes alterados (AGS) es determinar cuáles genes hacen a una particular línea celular de cáncer (CCL) diferente de las demás. Para cada CCL, en cada una de las matrices de perfiles genómicos crudos, se deriva un AGS (ver figura 3.1 (a)) utilizando diferentes estrategias que en esta sección se describen. Las estrategias para derivar AGSs para datos de mutaciones difieren de aquellas aplicables a datos de expresión de genes (GE) y de alteraciones en el número de copias (CNA). Se describen a continuación dos enfoques para derivar AGSs para GE y CNA, y dos enfoques para datos de mutaciones.

GE & CNA

Sea F una matriz $n \times p$ de datos crudos de perfiles de GE o CNA (figura 2.1 (b)), sea f_{ij} un elemento de dicha matriz. Se derivan AGSs para cada perfil en F (i. e. cada fila, asociada a una CCL, de F) utilizando los siguientes métodos.

Top L

Los genes de un perfil son ordenados por una medida de ordenamiento definida, el AGS incluye los *top L* genes (primeros L genes) del *ranking*. Aquí un valor de corte viene dado al imponer un límite, L , sobre el tamaño de los AGSs.

Se sigue el siguiente procedimiento:

- Se calcula la media del valor medido de un gen g sobre todos los perfiles, i. e. muestras de CCLs (ver en la figura 3.1 (a) la columna resaltada en **morado**):

$$\bar{F}_j = \frac{1}{n} \sum_{i=1}^n f_{ij} \quad j = 1, \dots, p \quad (3.1)$$

- Para un perfil $k \in \{1, \dots, n\}$, se define la medida de ordenamiento para los genes en el perfil de la siguiente manera:

$$r_{kj} = |\bar{F}_j - f_{kj}| \quad j = 1, \dots, p \quad (3.2)$$

- Se denota por g los nombres de genes y se define entonces el siguiente conjunto:

$$S = \{(g_j, r_{kj}) \mid j = 1, \dots, p\} \quad (3.3)$$

y el orden total $(S, \mathcal{R}_{\geq}) = \{(\tilde{g}_j, \tilde{r}_{kj}) \mid j = 1, \dots, p\}$ (3.4)

Se define la relación \mathcal{R}_{\geq} como:

$$\left(\forall (a, b), (c, d) : (a, b), (c, d) \in S : (a, b) \geq (c, d) \Leftrightarrow b \geq d \right) \quad (3.5)$$

nótese que \mathcal{R}_{\geq} es reflexiva, antisimétrica, transitiva y total (comparable), por ende (S, \mathcal{R}_{\geq}) es un orden total.

- Se define el AGS para el perfil genómico k de la siguiente manera:

$$\text{AGS}_k = \{\tilde{g}_l \mid l = 1, \dots, L\} \quad \text{para algún } 0 \leq L \leq p \quad (3.6)$$

Significant

Se realiza una prueba unilateral z para los genes en un perfil y pasan a formar parte del AGS aquellos que se desvén de manera significativa, por ende el tamaño de los AGSs no es fijo y puede variar de una CCL a otra.

- Para un perfil genómico $k \in \{1, \dots, n\}$ se calcula el valor unilateral z para los genes en el perfil de la siguiente manera:

$$z_{kj} = \frac{\bar{F}_j - f_{kj}}{\sqrt{\frac{\bar{F}_j(1-\bar{F}_j)}{n}}} \quad \text{donde} \quad \bar{F}_j = \frac{1}{n} \sum_{i=1}^n f_{ij} \quad j = 1, \dots, p \quad (3.7)$$

- El AGS comprenderá los genes más desviados de la media (aquellos cuyo z -valor es mayor a algún valor fijado Z).

$$\text{AGS}_k = \{g_j \mid j = 1, \dots, p \wedge |z_{kj}| > Z\} \quad (3.8)$$

Mutaciones

Sea F una matriz $n \times p$ de datos crudos de perfiles de mutaciones. Se deriva un AGS para cada perfil en F utilizando los siguientes métodos.

Total

- Para un perfil genómico $k \in \{1, \dots, n\}$

$$\text{AGS}_k = \{g_j \mid j = 1, \dots, p \wedge f_{kj} = 1\} \quad (3.9)$$

El AGS asociado a la línea celular incluye todos los genes mutados del perfil.

Reference

Mediante un procedimiento especial de NEA (disponible como una función separada en el paquete **NEArender**), se evalúa el enriquecimiento de genes presentes en un AGS derivado con la estrategia *total* con respecto a un ciclo biológico de referencia; cada gen del AGS se estudia de manera individual para asignarle un valor de enriquecimiento, aquellos genes cuyo puntaje z de enriquecimiento sea mayor a un Z dado pasarán a formar parte del nuevo AGS.

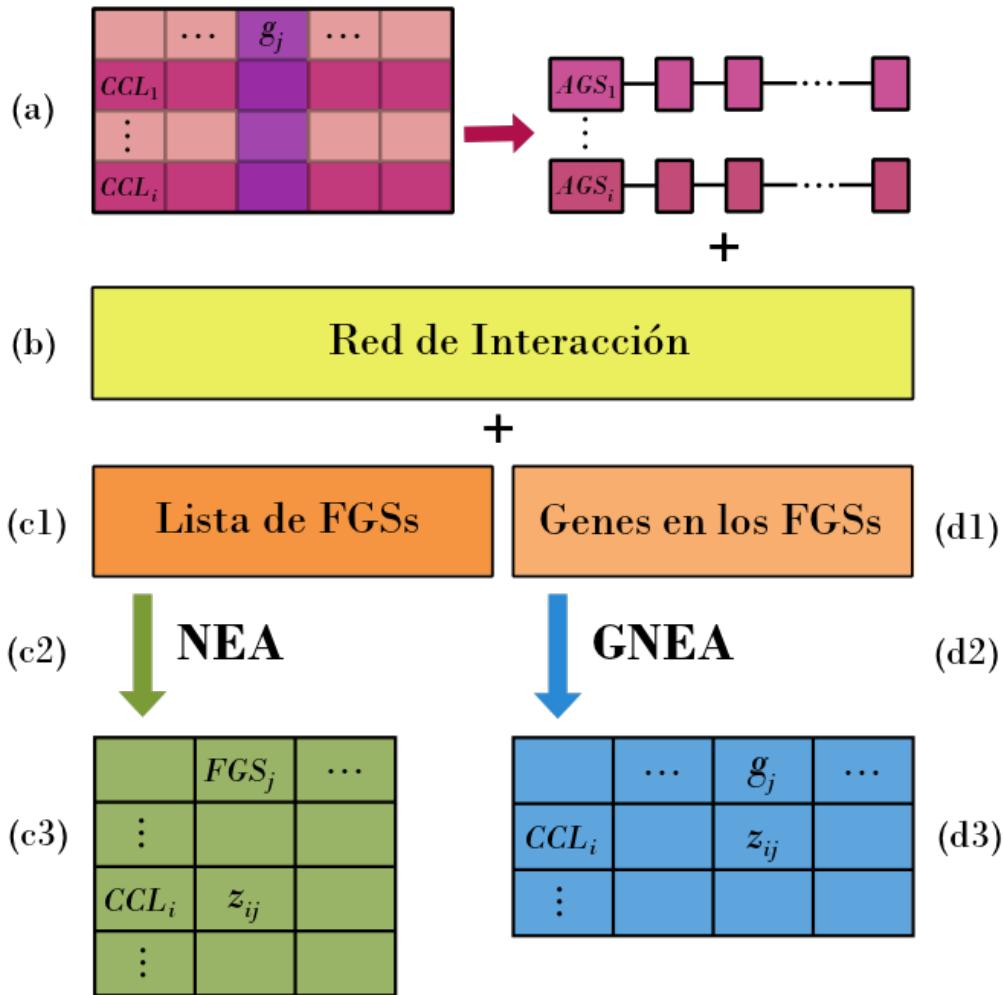


Figura 3.1: Derivación de AGSs y ejecución de NEA y GNEA.

- Para cada fila (asociada a una CCL) de una **matriz de datos crudos** de perfiles genómicos (figura 2.1 (b)) se deriva un **AGS** utilizando alguno de los métodos descritos en la sección 3.1.
- Es la **red de interacción de genes y proteínas** descrita en la sección 2.3.1.
- Es la **lista de FGSs** descrita en la sección 2.3.2.
- Es una lista que comprende cada uno de los genes en (c1).
- Se ejecuta NEA utilizando como entrada (a), (b) y (c1), i. e. la **lista de AGSs**, la **red** y la **lista de FGSs**. Los puntajes z resultantes se almacenan en (c3).
- La **matriz de datos NEA**. El elemento z_{ij} indica el z -valor (puntaje de enriquecimiento) del j -ésimo FGS en el AGS asociado a la i -ésima CCL.
- Se ejecuta GNEA utilizando como entrada (a), (b) y (d1), i. e. la **lista de AGSs**, la **red** y la **lista de genes en los FGSs**. Los puntajes z resultantes se almacenan en (d3).
- La **matriz de datos GNEA**. El elemento z_{ij} indica el z -valor (puntaje de enriquecimiento) del j -ésimo gen en el AGS asociado a la i -ésima CCL.

3.1.1. Ejecución de NEA y GNEA

Para las CCLs presentes en datos de perfiles de GE y CNA se derivaron cuatro conjuntos de listas de AGSs: *top* 200 genes de la clasificación, *top* 400 genes de la clasificación, *significant* (genes significativamente desviados) tomando $Z = 3,09$ que equivale a un p -valor < 0,0001 y una versión filtrada (vía NEA) de los genes más desviados (*significant*). El AGS filtrado se genera reduciendo la lista original de AGSs, usualmente demasiado grande, en un $\sim 70\%$ de manera que sólo los genes relacionados con ciclos importantes participan en el enriquecimiento (de manera similar a como se deriva un AGS *reference* a partir de un AGS *total* cuando se emplean datos de mutaciones).

Para las CCLs presentes en los datos de perfiles de mutaciones se derivaron dos conjuntos de AGSs: el primero utiliza el enfoque *total*; el segundo evalúa enriquecimiento de los genes presentes en los AGSs obtenidos de *total* con respecto a un ciclo grande involucrado en el cáncer (que contiene más de 250 genes), “super-” pathway KEGG_05200, fijando $Z = 3,09$.

Para cada conjunto de AGSs (figura 3.1 (a)), se ejecutó NEA (figura 3.1 (c2)) utilizando la red (figura 3.1 (b)) y la lista de FGSs (figura 3.1 (c1)) descritos en las secciones 2.3.1 y 2.3.2 respectivamente. Para GNEA (figura 3.1 (d2)) se utilizó la misma red (figura 3.1 (b)), considerando cada uno de los genes individuales en la lista de FGSs un FGS (ciclo) en si (figura 3.1 (d1)). Los puntajes de enriquecimiento resultantes fueron organizados en matrices ((c3) y (d3) en la figura 3.1) donde el elemento a_{ij} de una matriz da el z -valor de enriquecimiento del j -ésimo FGS/gen (en el caso de NEA/GNEA) en el AGS asociado a la i -ésima CCL.

Tipo de Perfil			
Expresión de Genes (GE)		Alteraciones en el Número de Copias (CNA)	Mutaciones
RAW	affymetrix1 1034 × 18900	snp6 994 × 21217	mgs 904 × 1667
NEA	pwnea.top.200.affymetrix1 pwnea.top.400.affymetrix1 pwnea.significant.affymetrix1 pwnea.significant.filtered.maxi.affymetrix1 1034 × 330	pwnea.top.200.snp6 pwnea.top.400.snp6 pwnea.significant.snp6 pwnea.significant.filtered.maxi.snp6 994 × 330	pwnea.total.mgs pwnea.reference.mgs 904 × 330
GNEA	gneatop.200.affymetrix1 gneatop.400.affymetrix1 gneasignificant.affymetrix1 gneasignificant.filtered.maxi.affymetrix1 1034 × 19027	gneatop.200.snp6 gneatop.400.snp6 gneasignificant.snp6 gneasignificant.filtered.maxi.snp6 994 × 19027	gneatotal.mgs gneareference.mgs 904 × 19027

Tabla 3.1: Convención de nombres, y dimensiones, para datos crudos, NEA y GNEA. Las filas indican tipo de dato y las columnas indican tipo de perfil genómico. La última fila de cada bloque en la tabla indica las dimensiones de las matrices, dadas como N° CCLs × N° Genes o N° CCLs × N° FGSS (para la fila NEA).

En la tabla 3.1 se listan los nombres convenidos de las matrices de datos que se utilizarán como entrada para la construcción de modelos de regresión lineal:

- RAW: matrices de datos crudos de perfiles genómicos (figura 2.1 (b)).
- NEA: matrices de puntajes NEA obtenidas utilizando diferentes métodos de generación de AGSSs (figura 3.1 (c3)).
- GNEA: matrices de puntajes GNEA obtenidas utilizando diferentes métodos de generación de AGSSs (figura 3.1 (d3)).

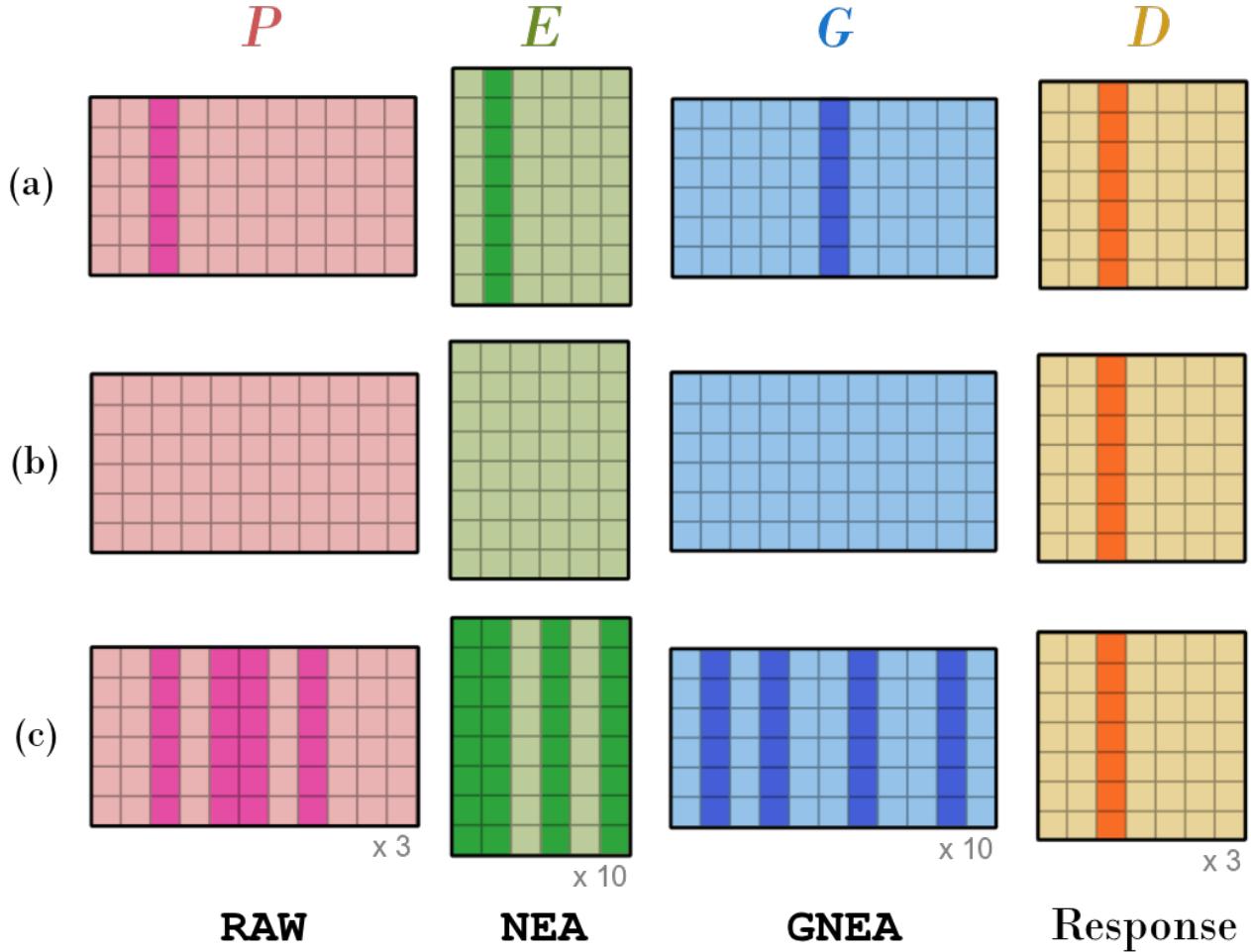


Figura 3.2: Representación esquemática de los datos de entrada para la regresión lineal.

P: matriz genérica de datos crudos de perfiles (fila **RAW** de la tabla 3.1, figura 2.1 (b)).

E: matriz genérica de datos NEA (fila **NEA** de la tabla 3.1, figura 3.1 (c3)).

G: matriz genérica de datos GNEA (fila **GNEA** de la tabla 3.1, figura 3.1 (d3)).

D: matriz genérica de datos de sensibilidad de CCLs a drogas (figura 2.1 (a)).

En la regresión lineal, $X \subseteq_{\text{cols}} F$ donde $F \in \{P, E, G\}$ y $y = d$ donde $d \in_{\text{cols}} D$.

$X \subseteq_{\text{cols}} F$ indica que X es igual a un subconjunto de las columnas de F .

$d \in_{\text{cols}} D$ indica que d es igual a una de las columnas de D .

Cada fila representa la entrada a uno de los tipos de modelos construidos:

- Modelos univariantes (sección 3.2).
- Modelos multivariantes de complejidad ilimitada (sección 3.3).
- Modelos multivariantes de complejidad limitada (sección 3.3.1).

3.2. Modelos Univariantes

Son modelos que sólo manejan una variable independiente i. e. una única característica (gen o FGS) por modelo. Los modelos univariantes de la forma $y = X\beta$ fueron ajustados utilizando la función `lm()` de R, con $y = \textcolor{red}{d}$ y $X = \textcolor{violet}{f}$ donde:

- $\textcolor{red}{d}$ es un vector columna ‘droga-específico’ tomado de una de las matrices de datos de respuesta $\textcolor{brown}{D}$.
- $\textcolor{violet}{f}$ es un vector columna ‘característica-específico’ para:
 - Un **gen** de una de las matrices de perfiles genómicos $\textcolor{red}{P}$.
 - Un **FGS** de una de las matrices de datos NEA $\textcolor{brown}{E}$.
 - Un **gen** de una de las matrices de datos GNEA $\textcolor{blue}{G}$.

La representación esquemática para $\textcolor{red}{d}$ y para $\textcolor{violet}{f}$ puede apreciarse en la figura 3.2 (a). Es importante notar que para realizar la regresión sólo deben considerarse las filas que representan CCLs comunes tanto a $\textcolor{red}{d}$ como a $\textcolor{violet}{f}$ y para las cuales se encuentran todos los datos disponibles (deben eliminarse las filas/CCLs con perfiles incompletos, i. e. medidas faltantes para algunas características, o para las cuales se carece del valor de sensibilidad/respuesta a la droga en cuestión).

Se ajustaron modelos lineales univariantes para todas las **drogas** posibles en cada una de las matrices de datos de sensibilidad $\textcolor{brown}{D}$ (3 en total) acopladas a todos las posibles **características**:

- **genes** de cada una de las matrices de perfiles genómicos ($\textcolor{red}{P}$) listadas en la fila **RAW** de la tabla 3.1 (3 en total).
- **FGSs/genes** de cada una de las matrices de datos NEA/GNEA ($\textcolor{brown}{E}/\textcolor{blue}{G}$) listadas en la fila **NEA/GNEA** de la tabla 3.1 (10 en total).

3.3. Modelos Multivariantes

Los modelos multivariantes de la forma $y = X\beta$ fueron ajustados vía el método *elastic net* haciendo uso de la función `cv.glmnet()` de R, con $y = \mathbf{d}$ y $X = \mathbf{F}$ donde:

- \mathbf{d} es un vector columna ‘droga-específico’ tomado de una de las matrices de datos de respuesta \mathcal{D} .
- $\mathbf{F} \in \{\mathcal{P}, \mathcal{E}, \mathcal{G}\}$

La representación esquemática para \mathbf{d} y para \mathbf{F} puede apreciarse en la figura 3.2 (b). Es importante notar que para realizar la regresión sólo deben considerarse las filas que representan CCLs comunes tanto a \mathbf{d} como a \mathbf{F} y para las cuales se encuentran todos los datos disponibles (deben eliminarse las filas/CCLs con perfiles incompletos, i. e. medidas faltantes para algunas características, o para las cuales se carece del valor de sensibilidad/respuesta a la droga en cuestión).

La función `cv.glmnet()` del paquete `glmnet` de R fue utilizada para ajustar los modelos lineales de *elastic net* realizando validación cruzada de 10 iteraciones²⁴ (el error es acumulado y se calcula su promedio sobre las iteraciones). Se asignó igual peso a las penalidades ℓ^1 y ℓ^2 colocando $\alpha = 0,5$.

Se ajustaron modelos lineales multivariantes de *elastic net* para todas las **drogas** posibles en cada una de las matrices de datos de sensibilidad \mathcal{D} acopladas con todas las **matrices** posibles (cada \mathcal{P} , \mathcal{E} y \mathcal{G} listadas respectivamente en las filas **RAW**, **NEA** y **GNEA** de la tabla 3.1).

Como no se restringe explícitamente el número de predictores (variables) seleccionados²⁵, se hace referencia a estos modelos como modelos de complejidad ilimitada.

²⁴En la validación cruzada de k iteraciones, se realiza una partición del conjunto de datos observados en k sub-muestras de igual tamaño, una se retiene para validación y las restantes $k - 1$ se utilizan para entrenar al modelo. El proceso se repite k veces (cada sub-muestra es utilizada exactamente una vez para validar).

²⁵Esto puede hacerse bien sea ajustando algunos parámetros de la función `cv.glmnet()` o preseleccionando características como se describe en la sección 3.3.1.

3.3.1. Modelos de Complejidad Limitada

No limitar de manera explícita el número de características suministradas a `cv.glmnet()` da al algoritmo demasiada libertad al momento de realizar selección de variables (muy frecuentemente reteniendo demasiadas variables en el modelo), resultando en modelos complejos que aparentan desempeñar bien debido a su tendencia a sobreajustar. Para evitar tal situación, se produjeron modelos de complejidad limitada vía preselección de un subconjunto de K características que han de usarse como entrada para *elastic net*.

Los modelos multivariantes *elastic net* de complejidad limitada, con $K \in \{10, 25\}$, se ajustaron de igual manera a como se ajustaron los modelos multivariantes *elastic net* de complejidad ilimitada (ver sección 3.3); en este caso en lugar de usar $X = \textcolor{violet}{F}$ se utiliza $X \subseteq_{\text{cols}} \textcolor{violet}{F}$ ²⁶ (note los **genes**, **FGSs** y **genes** resaltados respectivamente en las matrices $\textcolor{red}{P}$, $\textcolor{brown}{E}$ y $\textcolor{blue}{G}$ en la figura 3.2 (c); estos corresponden a las K características preseleccionadas).

Se ajustaron modelos multivariantes *elastic net* de complejidad limitada ($K \in \{10, 25\}$) para todas las **drogas** posibles en cada una de las matrices de datos de sensibilidad $\textcolor{brown}{D}$ acopladas con las K características preseleccionadas para todas las **matrices** posibles (cada $\textcolor{red}{P}$, $\textcolor{brown}{E}$ y $\textcolor{blue}{G}$ listadas respectivamente en las filas RAW, NEA y GNEA de la tabla 3.1).

Preselección de Características vía *Ranking*

Para cada par $(\textcolor{violet}{F}, \textcolor{red}{d})$ con $\textcolor{violet}{F} \in \{\textcolor{red}{P}, \textcolor{brown}{E}, \textcolor{blue}{G}\}$ (alguna de las matrices definidas en la tabla 3.1) y $\textcolor{red}{d}$ una columna ‘droga-específica’ de $\textcolor{brown}{D}$:

1. Se computa el coeficiente de correlación de Spearman (r_s) entre $\textcolor{red}{d}$ y cada columna ‘característica-específica’ (\mathbf{f}_j) de $\textcolor{violet}{F}$.

$$r_s(\textcolor{red}{d}, \mathbf{f}_j) \quad \text{para } j = 1, \dots, p \text{ (todas las características en } \textcolor{violet}{F})$$

²⁶Ver definición de \subseteq_{cols} en la descripción de la figura 3.2.

2. Las características se ordenan decrecientemente por el valor de $|r_s(\mathbf{d}, \mathbf{f}_j)|$.

El subconjunto seleccionado comprenderá las primeras K características del *ranking*.

La matriz `mgs` con datos binarios discretos es la única excepción al enfoque descrito. Para perfiles de mutaciones se utilizaron p -valores de ANOVA²⁷ para seleccionar los subconjuntos de genes.

Para cada una de las 365 drogas (203 en `Basu`, 24 en `Barretina` y 138 en `Garnett`) se ajustaron:

- 235354 modelos univariantes (18900 por cada gen en `affymetrix1`, 21217 por cada gen en `snp6`, 1667 por cada gen en `mgs`, 330×10 por cada FGS en cada matriz de datos NEA y 19027×10 por cada gen en cada matriz de datos GNEA). En total se ajustaron 85904210 (365×235354) modelos univariantes.
- 23 modelos multivariantes para cada nivel de complejidad (ilimitada, limitada a 25 características y limitada a 10 características), uno por cada una de las 3 matrices de perfiles genómicos crudos, uno por cada una de las 10 matrices de datos NEA y uno por cada una de las 10 matrices de datos GNEA. En total se ajustaron 8395 (365×23) modelos multivariantes para cada nivel de complejidad (ilimitada, limitada a 25 características y limitada a 10 características).

²⁷Prueba F de un modelo lineal donde $y = \mathbf{d}$ y $X = \mathbf{f}_j$.

Capítulo 4

Experimentos y Resultados

Los modelos ajustados serán resumidos en esta sección por estudio farmacogenómico (o datos de respuesta, i. e. Basu, Barretina o Garnett), tipo de perfil (GE, CNA o mutaciones, o respectivamente `affymetrix1`, `snp6` y `mgs`), tipos de datos (crudos/`RAW`, NEA o GNEA/NEA o GNEA; los últimos dos incluyendo todas las alternativas para derivar listas de AGSs), complejidad de modelos y por último, utilizamos drogas analizadas por más de un estudio (“drogas comunes”) para comparar las opciones listadas al momento de transferir modelos multivariantes a conjuntos independientes de datos.

Entre otras metas, se quería probar si los datos NEA y GNEA presentarían un mejor desempeño que los datos crudos. Se compararon las combinaciones de tipos de perfiles por tipos de datos de acuerdo a cuántos modelos buenos (entrenados con estos datos) se observan. Además de la medida estándar de error cuadrático medio (MSE, del inglés *mean squared error*), se podría juzgar bien sea el coeficiente de correlación o el *p*-valor de la respuesta predicha por un ajuste de regresión versus la respuesta verdadera. Es de destacar que el *p*-valor es también una función del número disponible de CCLs para cada modelo, dado que este número es rara vez menor a 100, cualquier coeficiente de correlación mayor a p. ej. 0,25 sería significativo a $p < 0,01$ de modo que comparar fracciones de coeficientes mayores a un

certo umbral sería una medida informativa de desempeño relativo. Por esta razón, se emplean en este reporte coeficientes de correlación de Spearman (r_s), p -valores y errores cuadráticos medios (MSE).

4.1. Correlatos Uni-Característica

Se ejecutó este experimento con la intención de evaluar la concordancia entre los datos de sensibilidad (respuesta) de los diferentes estudios farmacogenómicos y con la esperanza de descubrir coincidencias ocultas que no hayan sido reportadas previamente.

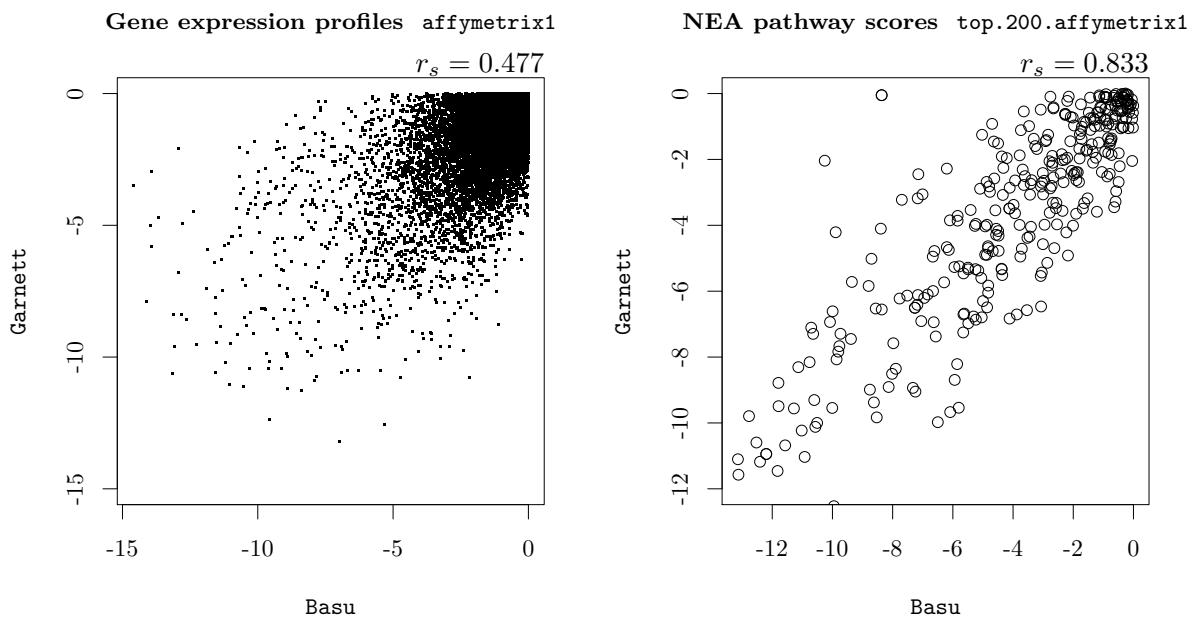


Figura 4.1: Diagrama de dispersión de p -valores para la droga Atra.

Se llevó a cabo un experimento piloto realizando un análisis uni-característica sistemático. Los resultados indicaron que el uso de NEA y GNEA ha de ser ventajoso en contraste al uso de datos crudos para construir modelos multivariantes. La idea de este análisis es tomar cada droga común y observar cómo las correlaciones individuales o p -valores, para cada tipo de perfil y de dato, se preservan entre diferentes estudios.

Para dos estudios A y B (p. ej. Garnett y Basu en la figura 4.1), para una droga determinada (Atra en la figura 4.1) y para todos los tipos de dato de un tipo particular de perfil²⁸ (en la figura 4.1 se observan perfiles de expresión de genes, `affymetrix1`, a la izquierda y `pwnea.top.200.affymetrix1` a la derecha) se calcula la significación de modelos univariantes para cada característica disponible (en la figura 4.1 serían 18900 características de expresión de genes en `affymetrix1` y 330 FGSs en `pwnea.top.200.affymetrix1`). Los p -valores obtenidos, y los respectivos coeficientes de correlación para los dos estudios A y B son graficados para visualizar concordancia entre correlaciones ‘gen vs. sensibilidad a drogas’ entre estudios. La figura 4.1 demuestra que los puntajes de FGSs de NEA fueron mucho más consistentes entre los estudios que los valores originales (crudos) de expresión de genes.

Se aplicó este procedimiento de forma sistemática a cada droga compartida entre cada par de estudios (el número de drogas compartidas entre cada par de estudios es de 4, 28 y 15 para Basu y Barretina, Basu y Garnett, y Barretina y Garnett, respectivamente). Los resultados fueron luego resumidos por tipo de perfil y dato como fracciones de buenos coeficientes de correlación/ p -valores mayores/menores que un cierto valor umbral. La tabla 4.1 presenta los resultados resumidos para valores de correlación; las filas indican conjuntos de datos (tipo de perfil y de dato) y las columnas dan el valor umbral de correlación rc , de modo que la entrada a_{ij} de la tabla indica la fracción de correlaciones mayores a rc_j usando datos del conjunto asociado a la fila i .

Se observan mayores correlaciones entre diferentes estudios cuando se usan datos de enriquecimiento NEA y en especial GNEA. Este resultado permite concluir que: *i*) contrario a las conclusiones establecidas por [12], si existe cierto nivel de concordancia y acuerdo entre los datos de sensibilidad reportados por los estudios [10] y [11] y *ii*) el uso de NEA y GNEA ha de ser ventajoso en contraste al uso de datos crudos para construir modelos multivariantes.

²⁸Es decir, la correspondiente matriz de perfiles crudos \mathbf{P} y todas las matrices NEA (\mathbf{E}) y GNEA (\mathbf{G}) derivadas de \mathbf{P} .

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
snp6	0.333	0.118	0.02	0	0	0	0	0	0
pwnea.top.200.snp6	0.186	0.039	0	0	0	0	0	0	0
pwnea.top.400.snp6	0.167	0.059	0.02	0	0	0	0	0	0
pwnea.significant.snp6	0.137	0	0	0	0	0	0	0	0
mgs	0.038	0	0	0	0	0	0	0	0
pwnea.total.mgs	0.275	0.069	0.01	0	0	0	0	0	0
pwnea.reference.mgs	0.284	0.088	0.01	0	0	0	0	0	0
pwnea.significant.filtered.maxi.snp6	0.245	0.127	0.059	0.02	0	0	0	0	0
gneatop.200.snp6	0.167	0.098	0.069	0.029	0.01	0.01	0	0	0
gneatop.400.snp6	0.235	0.147	0.069	0.029	0.02	0	0	0	0
gneatop.400.affymetrix1	0.343	0.127	0.059	0.029	0.01	0	0	0	0
gneatotal.mgs	0.314	0.127	0.039	0.039	0.01	0	0	0	0
gneareference.mgs	0.363	0.147	0.078	0.039	0.02	0	0	0	0
affymetrix1	0.402	0.196	0.088	0.069	0.01	0	0	0	0
gneasignificant.snp6	0.255	0.186	0.127	0.069	0.039	0.029	0	0	0
pwneasignificant.affymetrix1	0.471	0.363	0.196	0.137	0.049	0.029	0.02	0	0
pwneasignificant.filtered.maxi.affymetrix1	0.49	0.353	0.225	0.147	0.098	0.059	0.029	0.01	0
pwneatop.200.affymetrix1	0.471	0.382	0.275	0.157	0.108	0.069	0.029	0.01	0
gneatop.400.affymetrix1	0.559	0.373	0.265	0.167	0.127	0.078	0.049	0.029	0
gneasignificant.affymetrix1	0.539	0.353	0.265	0.196	0.108	0.088	0.039	0	0
pwneatop.400.affymetrix1	0.559	0.461	0.343	0.216	0.108	0.069	0.039	0.02	0
gneatop.200.affymetrix1	0.608	0.48	0.353	0.225	0.157	0.098	0.059	0.029	0.01
gneasignificant.filtered.maxi.affymetrix1	0.578	0.471	0.304	0.235	0.147	0.078	0.01	0	0

Tabla 4.1: Fracciones de correlatos uni-característica.

La siguiente leyenda ha de utilizarse para interpretar los diagramas de caja en el presente capítulo (figuras 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 4.10 y 4.11) y en los apéndices C y D.

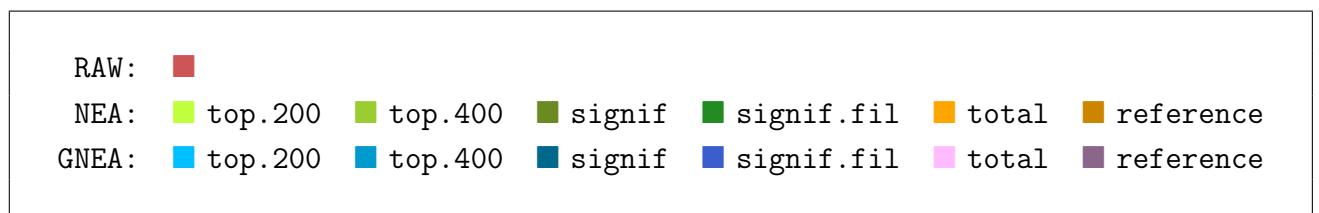


Figura 4.2: Leyenda para interpretar los diagramas de cajas.

4.2. Exploración de Diferentes Niveles de Complejidad

Este experimento se ejecuta con la intención de comparar el desempeño de los modelos multivariantes *elastic net* de acuerdo al tipo de perfil, el tipo de dato y el nivel de complejidad de los modelos.

Luego de ajustar los modelos lineales multivariantes *elastic net* usando `cv.glmnet()`, los modelos fueron asociados a una cantidad de coeficientes de correlación r_s (determinada por el número de drogas presentes en el conjunto de datos de respuesta utilizado para producir los modelos). Las distribuciones de estos coeficientes de correlación se comparan por tipo de perfil, tipo de dato y nivel de complejidad de los modelos en una serie de diagramas de cajas que pueden ser vistos en las figuras 4.3, 4.4 y 4.5 (ver leyenda asociada en la figura 4.2) correspondientes a los estudios de Basu, Barretina y Garnett respectivamente. A continuación se resumen algunas de las observaciones más relevantes:

- Una simple inspección de las medianas de cada grupo de modelos en los diagramas de cajas revela que los valores de r_s no presentan desviaciones perceptibles con respecto a la normalidad.
- Para todos los estudios, los modelos de complejidad ilimitada ajustados con datos crudos parecen superar a NEA para todos los tipos de perfiles. Sin embargo, los datos GNEA no presentaron un peor desempeño que los datos crudos para mutaciones (`mgs`).
- Se observa que para los datos NEA, la elección del método para derivar AGSs parece no tener un efecto significativo sobre los valores de correlación. Esta observación es válida para todos los estudios, niveles de complejidad y tipos de perfil. Por otra parte, los arreglos GNEA derivados usando los métodos `top.200` y `top.400` (para generar las listas de AGSs) basados en `affymetrix1` y `snp6` parecen desempeñar mejor que otros métodos.

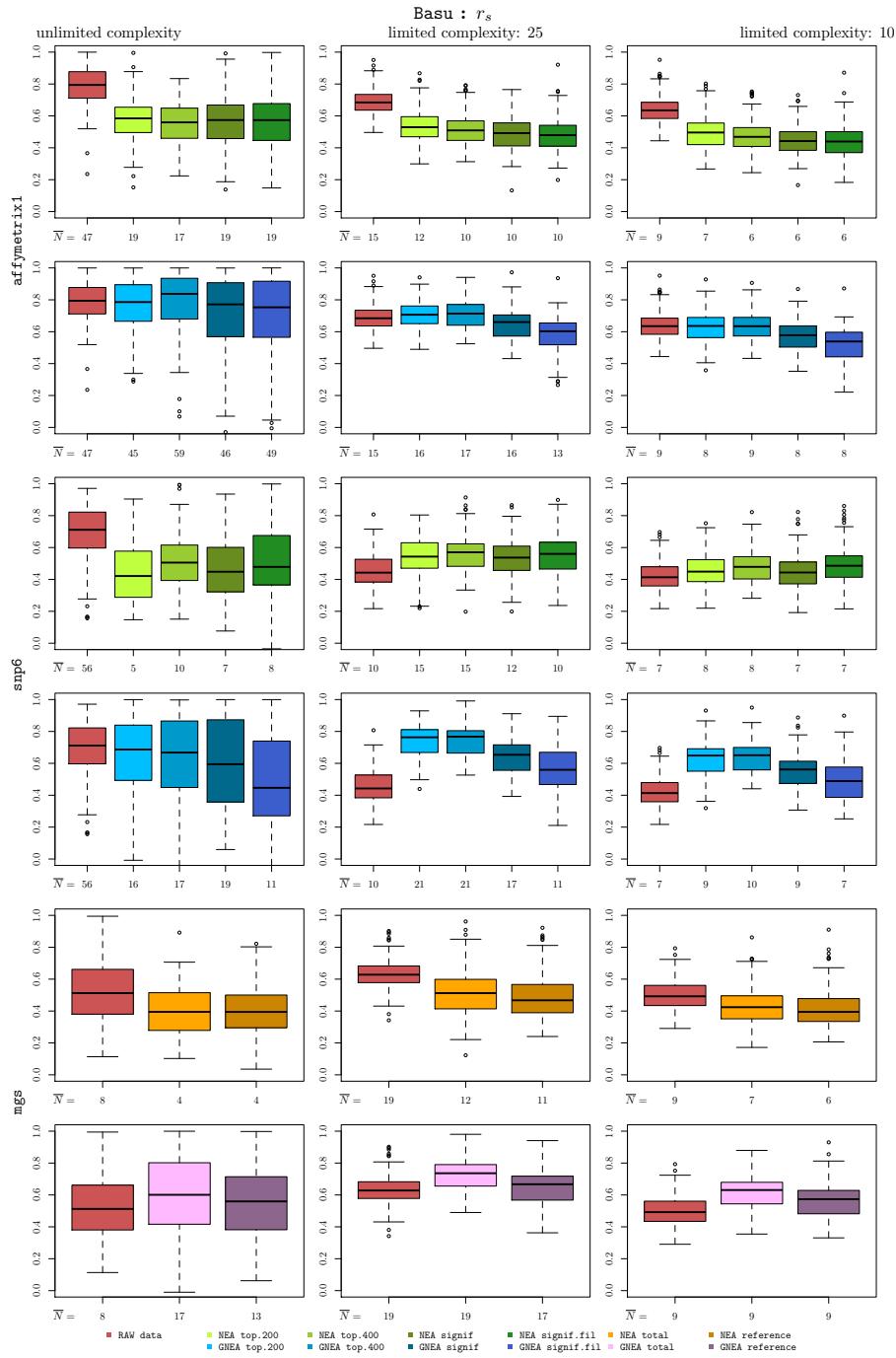


Figura 4.3: Modelos de diversa complejidad, Basu, r_s

- Para todos los estudios, NEA y especialmente GNEA superan a los datos crudos en los modelos de complejidad limitada que usan perfiles de CNA (snp6).
- En los modelos que utilizan información sobre mutaciones puntuales (mgs), para todos los estudios y niveles de complejidad GNEA parece superar ligeramente a su contraparte de datos crudos.

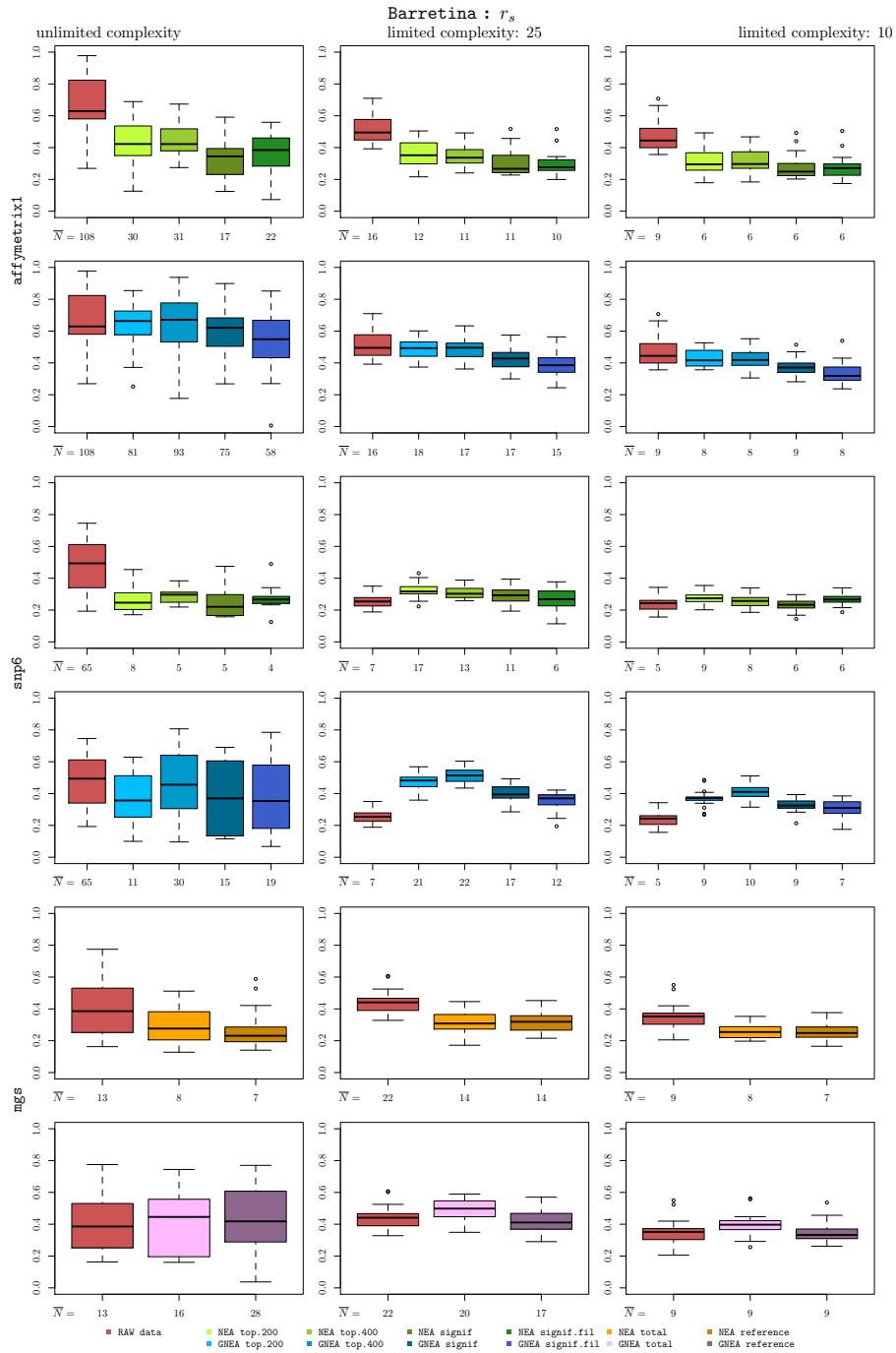


Figura 4.4: Modelos de diversa complejidad, Barretina, r_s

La observación más impactante es que los datos crudos superan a NEA (ver las filas 1, 3 y 5 de la primera columna de las figuras 4.3, 4.4 y 4.5). Luego de inspeccionar cuidadosamente las complejidades de los modelos, notamos que el desempeño está casi perfectamente determinado por el número de características (predictores) en el modelo (véase \bar{N} en la leyenda). Sin embargo, es sabido que los modelos más complejos tienden a sobreajustarse a los

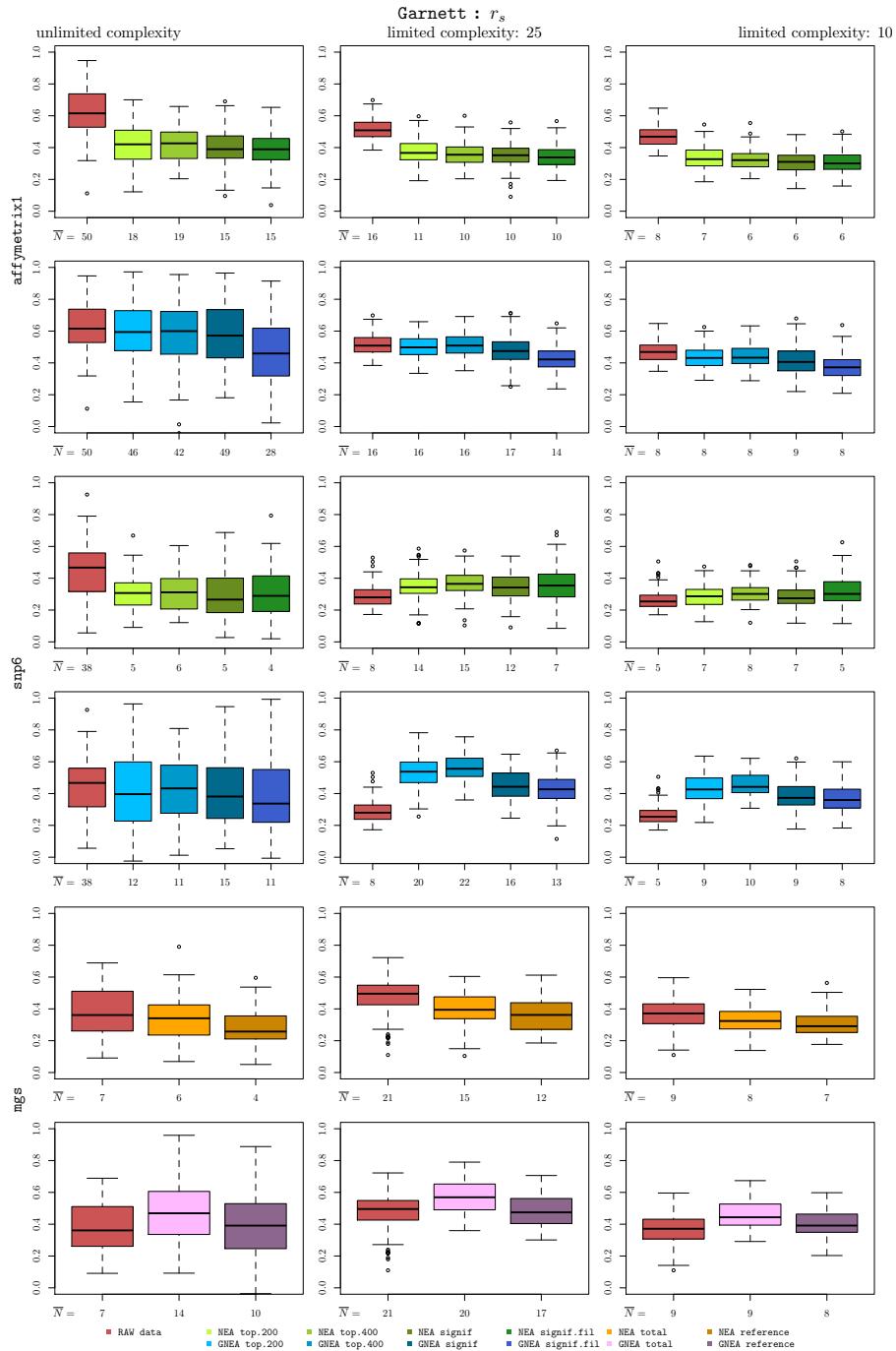


Figura 4.5: Modelos de diversa complejidad, Garnett, r_s

datos. Para contrarrestar esto, se siguen los procedimientos descritos en la sección 3.3.1 para limitar explícitamente el número de características utilizadas; los resultados para modelos de complejidad limitada se presentan en la segunda y tercera columna de las figuras 4.3, 4.4 y 4.5.

La diferencia entre datos crudos, NEA y GNEA usando perfiles de expresión de genes y de mutaciones probablemente se explica por la menor dimensionalidad, ya que es entre 30 y 50 veces menor en las matrices NEA (tan solo 330 características). Por otra parte, la dimensionalidad de las matrices GNEA es del mismo orden (19027 características) que las de datos crudos. El hecho que GNEA ligeramente supera a los datos crudos `mgs` puede explicarse por la alta fracción de mutaciones pasajeras irrelevantes (en los perfiles) cuyo efecto se trata de eliminar usando las listas de AGSs derivadas con el método *reference*.

En los modelos de complejidad limitada que usan perfiles de CNA (`snp6`), NEA y particularmente GNEA superan a sus contrapartes crudas. Esto puede explicarse por el hecho de que mientras los modelos limitados con datos crudos sólo cuentan con acceso a los primeros 25 o 10 genes (y estos puede que provengan de la misma región cromosómica, es decir, transmiten la misma información), cada una de las 25 o 10 características seleccionadas por NEA/GNEA colecciona información de un número mucho mayor de genes con alteraciones en número de copias, provisto que estos estaban conectados a FGSSs/genes en la red; también, cuando se ordenan las características por correlación se están seleccionando lo que en teoría son las mejores características (si bien es menos eficiente que la selección realizada por `glmnet`, no puede evadirse la estructura de los datos).

Estos resultados permiten concluir que el uso de NEA, y especialmente el uso de GNEA, ha de ser de alto interés en situaciones donde: *i*) se desean construir modelos de predicción con un número limitado de predictores y *ii*) el contenido de información en conjuntos de datos moleculares es pobre, p. ej. al usar perfiles de alteraciones en el número de copias o perfiles de mutaciones (más aún, se considera beneficioso dado que estas características genómicas tradicionalmente son mejor valoradas, porque son consideradas como más casuales, que valores de expresión de genes).

4.3. Modelos para Drogas Compartidas

Este experimento se llevó a cabo con la intención de explorar el desempeño de modelos construídos a partir de características transferidas. El objetivo es evaluar la viabilidad de utilizar las características de un modelo ajustado utilizando un conjunto de datos de sensibilidad para construir otro modelo haciendo uso de un conjunto distinto de datos de sensibilidad; la comparación se hace para la misma droga (común a al menos un par de estudios, i. e. conjuntos de datos de sensibilidad/respuesta) y se consideran únicamente modelos multivariantes *elastic net* de complejidad ilimitada.

La más común, y en general mejor aceptada, manera de demostrar que un enfoque funciona es mostrando que se preserva el desempeño en un conjunto diferente de datos. Se tuvo la oportunidad de probar diferentes enfoques y los modelos obtenidos por validación independiente y a la larga juzgar su desempeño en estos escenarios. Fueron comparados modelos multivariantes de complejidad ilimitada para drogas compartidas entre pares de estudios, se ajustaron modelos nuevos empleando las características de un modelo en un estudio y utilizándolas como entrada a una regresión lineal para producir un modelo con los datos de sensibilidad de otro estudio.

Para este experimento, se utilizaron diferentes datos de sensibilidad a drogas (variable y) mientras que el resto de los datos para la regresión (matriz X , i. e. los datos crudos, NEA y GNEA para los tipos de perfiles `affymetrix1`, `snp6` y `mgs`) se mantuvieron constantes. Es poco probable que un modelo de K características ajustado con datos de un estudio, p. ej. Basu, funcione para otro estudio, p. ej. Barretina, usando los mismos coeficientes. Más aún, los *rankings* difieren en gran medida de un estudio a otro, de modo que el enfoque de ‘complejidad limitada’ sería imposible de aplicar en este contexto. Por estas razones, el siguiente procedimiento fue utilizado para comparar modelos de drogas compartidas entre pares de estudios:

- Sea d una droga evaluada por dos estudios farmacogenómicos, A y B .
- Se tomaron las características seleccionadas (coeficientes distintos de cero) de un modelo multivariante de complejidad ilimitada para la droga d , ajustado utilizando datos de sensibilidad del estudio A como respuesta, se denota a este modelo por m_A , dichas características se usan como entrada para producir un modelo multivariante de complejidad ilimitada para la misma droga d empleando esta vez datos del estudio B como respuesta, se denota a este modelo por m_{BA} ; nótese que los datos de respuesta (variable y) varían mientras que las observaciones (matriz X) se mantienen. En paralelo se realiza el proceso opuesto, es decir, características seleccionadas del modelo m_B (ajustado originalmente utilizando datos de sensibilidad del estudio B) se utilizan como entrada para producir un modelo m_{AB} (usando datos de sensibilidad del estudio A).

Se introduce a continuación la notación Primary (P) y Transferred (T). En el escenario descrito, si m_A es el modelo primario P, entonces m_{AB} es el modelo transferido T (creado utilizando como entrada características transferidas del modelo m_B), de modo similar, si m_B es el modelo primario P entonces m_{BA} es el modelo transferido T²⁹.

Las tablas y diagramas de cajas en el apéndice D resumen las medidas estadísticas para estos modelos. Para una lista de todas las drogas compartidas entre pares de estudios referirse a la tabla D.3 en el apéndice D.1.

Las distribuciones de coeficientes de correlación r_s fueron comparadas por tipo de dato y de perfil para pares de estudios P y T, en los diagramas de cajas mostrados en las figuras 4.6 & 4.7 (Basu y Barretina), 4.8 & 4.9 (Basu y Garnett) y 4.10 & 4.11 (Barretina y Garnett); ver leyenda en la figura 4.2. Se resumen algunas observaciones a continuación:

²⁹Otra observación con respecto a la notación: si m_A y m_{AB} son los modelos primario y transferido respectivamente, entonces al estudio A se le denota por P (estudio primario) y al estudio B se le denota por T (estudio transferido).

- Por sobre todo, es importante recordar que el número de drogas compartidas entre cualquier par de estudios es bastante bajo (ver figura 2.2 (a) en la sección 2.1), de modo que no se puede estrictamente evaluar la significación de estos hallazgos. No obstante, aún es posible adquirir información acerca del desempeño de los modelos analizando las figuras.
- Para la mayoría de los pares de estudios, los modelos transferidos parecen no desempeñar tan bien como sus contrapartes primarios. Esta observación es aplicable a todos los tipos de datos y perfiles.
- Luego de transferir a un nuevo estudio, la complejidad de los modelos *elastic net* de facto disminuyó 2, 3 y hasta 4 veces y puede ser tan baja como $N < 2$ (ver valores de \bar{N} en las leyendas). Entonces, a pesar de que su desempeño bajo tales condiciones limitadas y poco favorables no fue satisfactorio, los resultados obtenidos pueden utilizarse como punto de partida para futuras mejoras al método.

No se esperaba necesariamente obtener buenos resultados inmediatamente luego de transferir características seleccionadas, de hecho, la probabilidad de que esto suceda es bastante baja; en el futuro sería interesante probar ajustar los parámetros de la función `cv.glmnet()` para explorar en mayor detalle el posible valor agregado que puedan aportar NEA y GNEA, el cual podría estarse obviando en este escenario debido a la metodología seguida. Otra opción más técnica a explorar es emplear únicamente las características que estén bien representadas (i. e. pocos valores faltantes) en ambos estudios.

Considerando el buen desempeño de los modelos univariantes, se puede concluir a partir de estos resultados que existen ciertos inconvenientes con la metodología empleada y que han de considerarse otras alternativas para realizar transferencia de características de modelos multivariantes entre diferentes conjuntos de datos de respuesta a drogas.

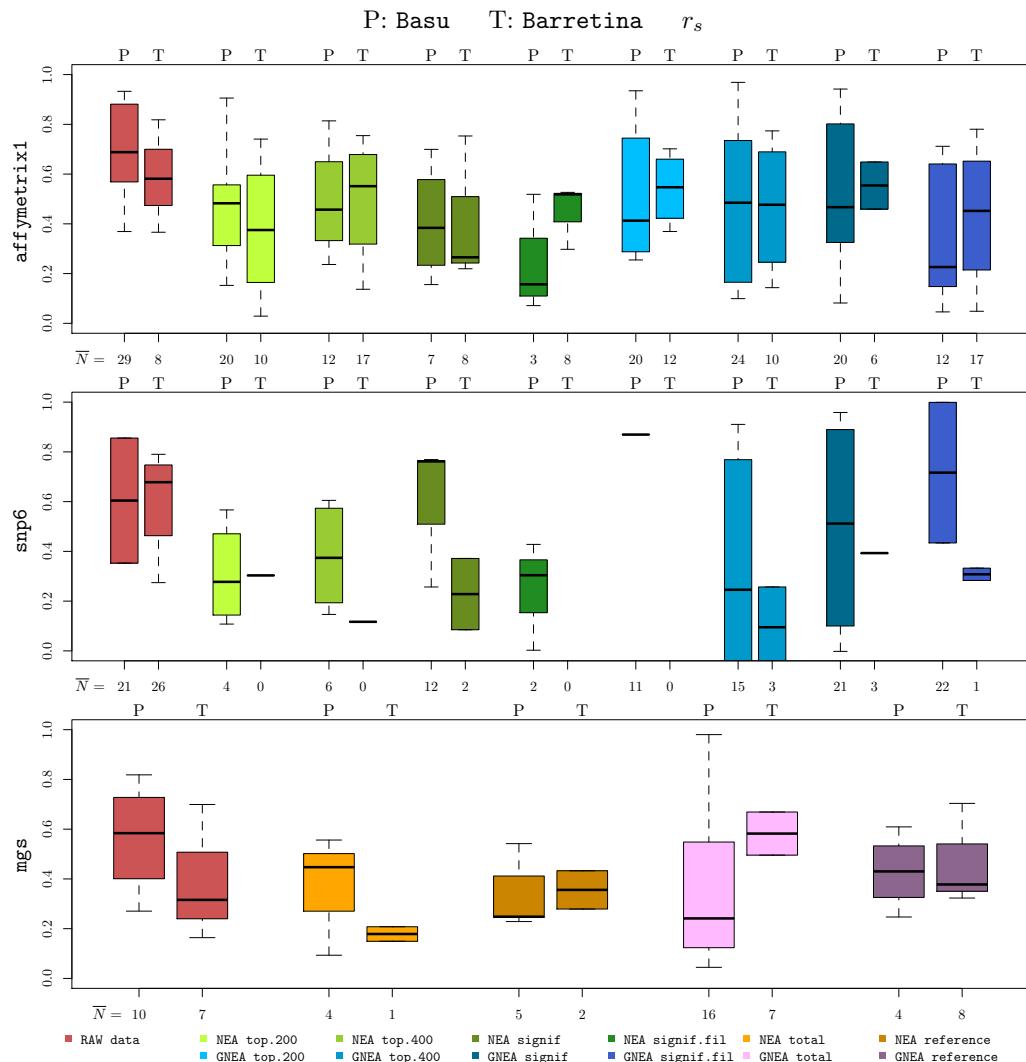


Figura 4.6: Modelos para drogas compartidas,
P: Basu, T: Barretina, r_s

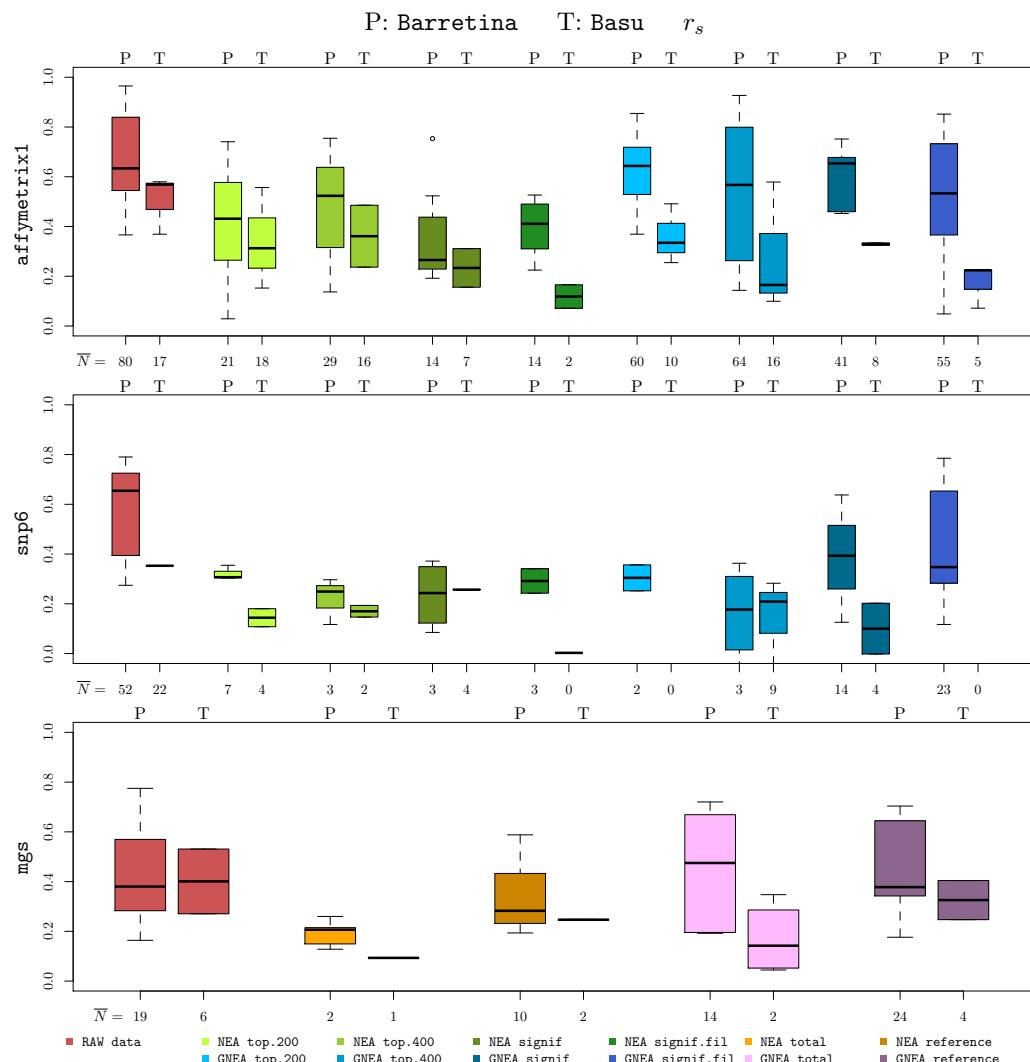


Figura 4.7: Modelos para drogas compartidas,
P: Barretina, T: Basu, r_s

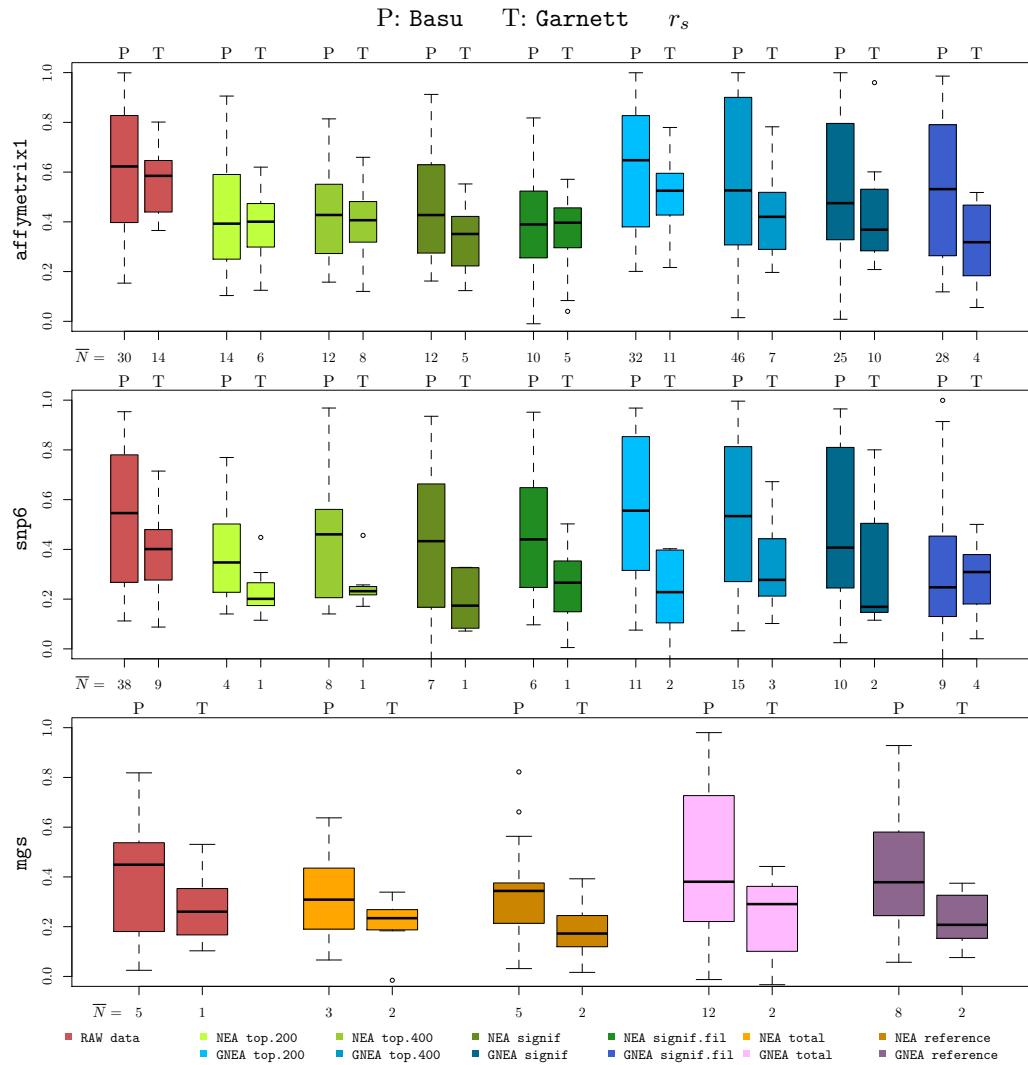


Figura 4.8: Modelos para drogas compartidas,
P: Basu, T: Garnett, r_s

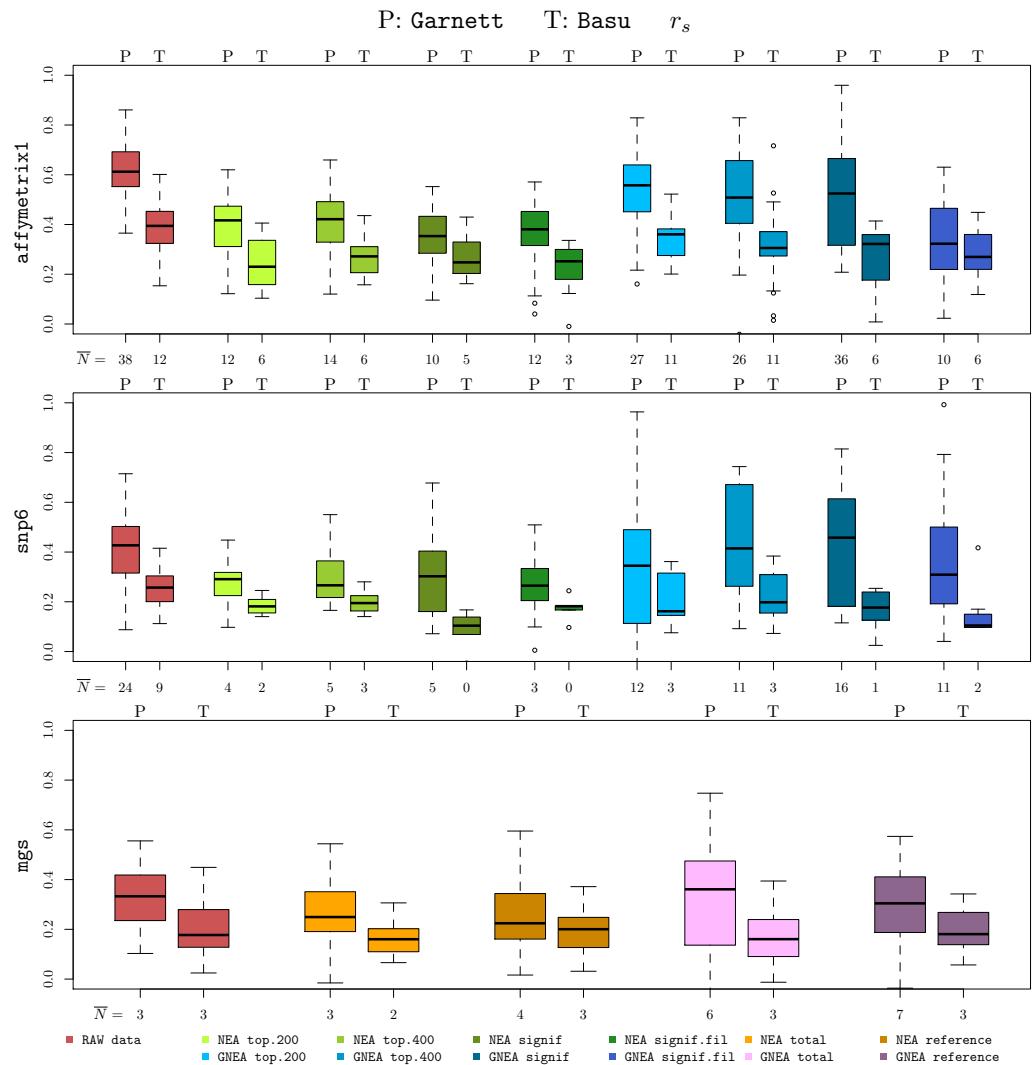


Figura 4.9: Modelos para drogas compartidas,
P: Garnett, T: Basu, r_s

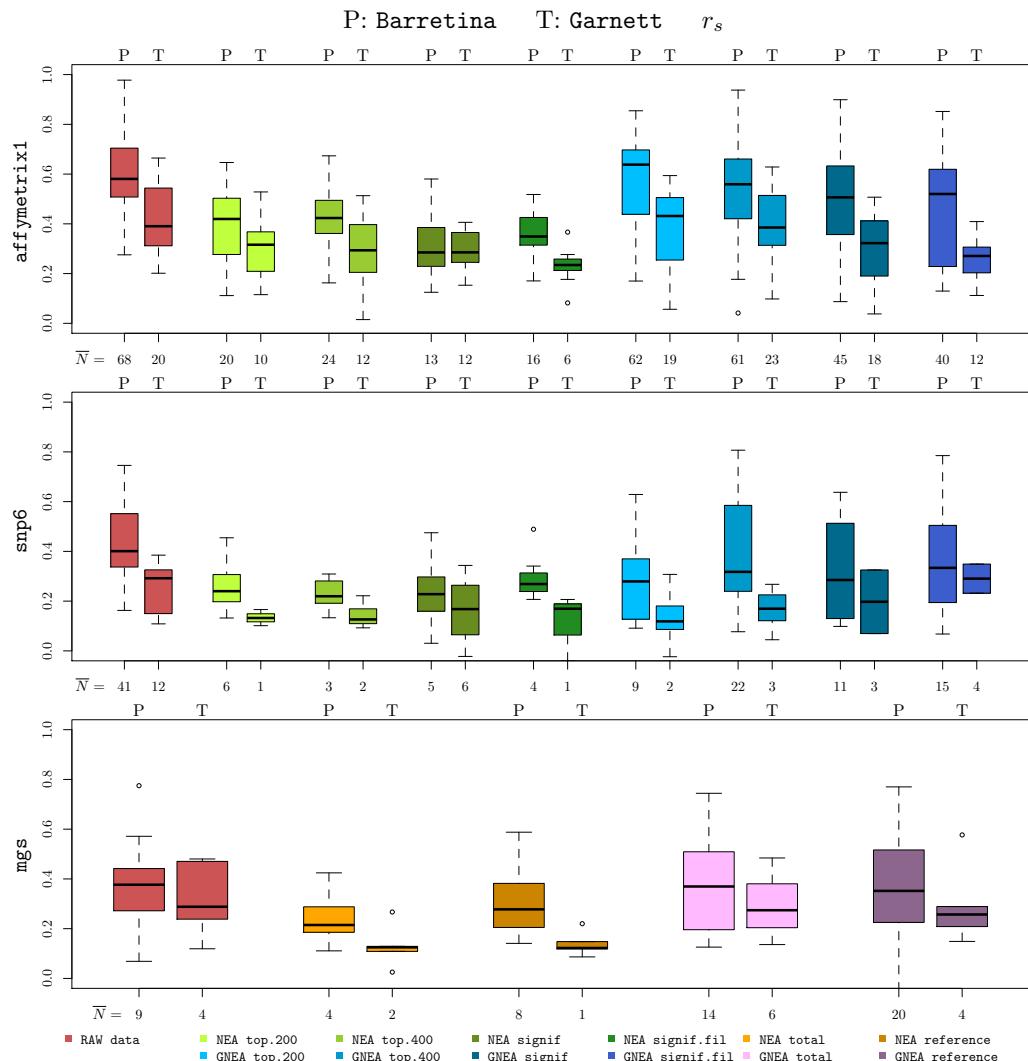


Figura 4.10: Modelos para drogas compartidas,
P: Barretina, T: Garnett, r_s

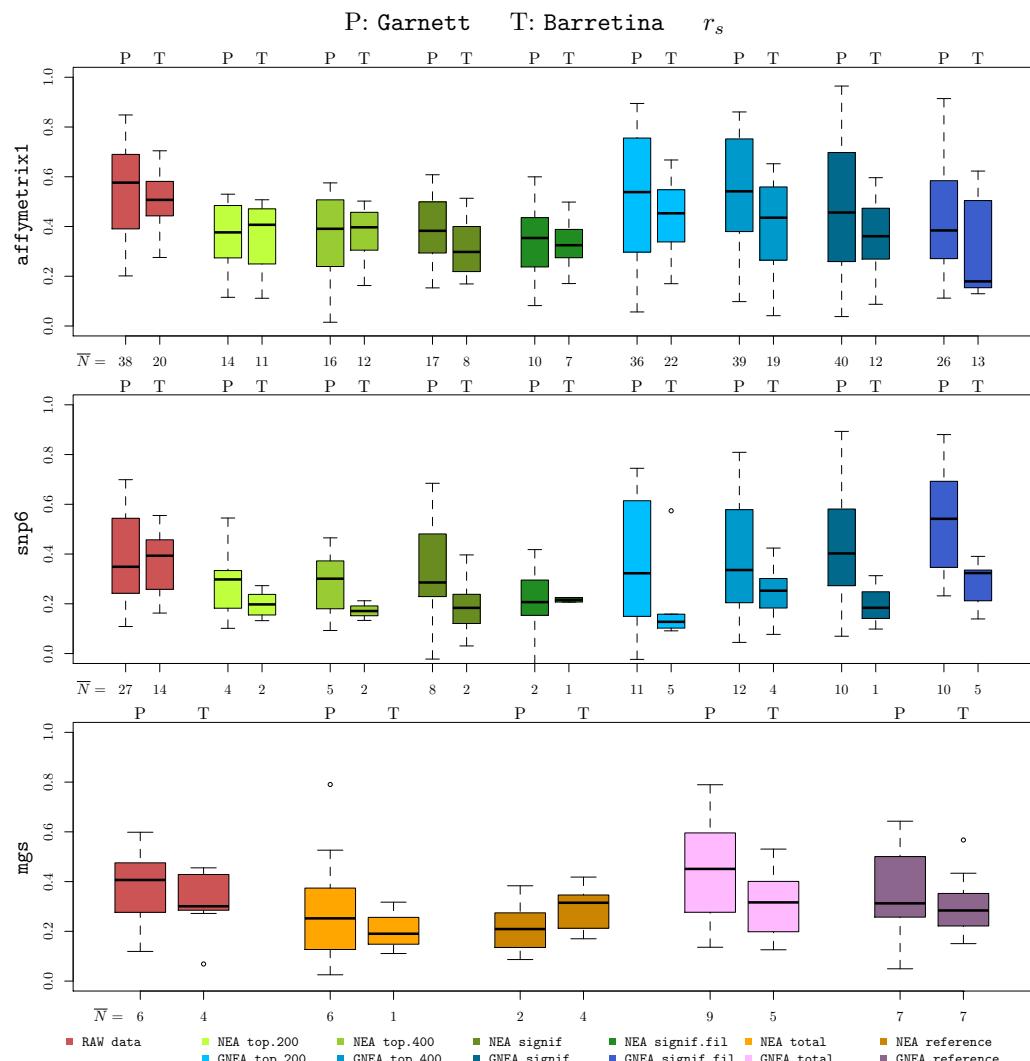


Figura 4.11: Modelos para drogas compartidas,
P: Garnett, T: Barretina, r_s

Capítulo 5

Conclusiones

Este estudio no pretendía alcanzar predicciones viables perfectas de actividad de drogas, en cambio, la intención era encontrar mejores opciones y características para investigaciones futuras (dado que las del estado del arte fallan tan frecuentemente).

El estudio piloto de correlatos uni-características mostró que los puntajes de NEA, y especialmente GNEA, fueron mucho más consistentes entre estudios que los valores crudos, frecuentemente alcanzando mayores valores de correlación. La mayoría de los estudios no logran encontrar conservación de modelos entre conjuntos de datos recolectados de manera independiente; sin embargo, los modelos lineales univariantes (una única característica) ajustados en este trabajo muestran que existe una concordancia sustancial entre diferentes estudios farmacogenómicos, demostrando así el alto contenido de información transferible y biológicamente relevante en NEA y GNEA.

La superioridad de los modelos *elastic net* no fue tan evidente al momento de comparar drogas compartidas entre estudios farmacogenómicos. Esto ciertamente es un inconveniente de la metodología empleada; simplemente hay demasiados enfoque que probar y parámetros

que ajustar, de modo que un estudio más detallado ha de considerarse como trabajo a futuro. Se tuvo una oportunidad única de probar enfoques y los modelos obtenidos por validación independiente y a la larga juzgar su desempeño en este escenario; por esta razón no se invirtió recursos en explorar ajustes alternativos para procedimientos de validación cruzada, elección de λ y α , etc.

Se conocía a priori que el contenido de información era más alto en los datos de perfiles de expresión de genes; como ejemplo, el número de correlatos de genes individualmente significativos era mayor a 1000; por ello no se esperaba que otros tipos de perfiles (alteraciones en el número de copias de genes y mutaciones puntuales) compitieran con valores de expresión de genes. Tampoco se estaba esperando que las características NEA y GNEA derivadas de `affymetrix1` llegasen a superar a los datos crudos. Sin embargo, en situaciones donde el contenido de información en conjuntos de datos moleculares es pobre, utilizar NEA y especialmente GNEA ha de ser de alto interés. A pesar de la muy baja complejidad de los modelos transferidos, estos en muchos casos compiten (no eran significativamente peores) con los respectivos modelos de datos crudos cuando se emplea información de alteraciones en el número de copias (`snp6`) y de mutaciones puntuales (`mgs`); esto es beneficioso dado que estas características genómicas tradicionalmente son mejor valoradas (pensadas como más casuales) que los valores de expresión.

Poniendo de lado la perspectiva puramente estadística, se considera que las características de NEA son más interpretables en el sentido biológico dado que comprendían los perfiles crudos con pobre sobreposición en una manera que los biólogos de cáncer pueden entender. Por ejemplo, se sabe que los patrones de mutaciones puntuales pueden presentar una sobreposición pobre entre pacientes con un mismo tipo de cáncer, sin embargo, es sabido que hay

ciclos comunes que se ven interrumpidos por dichas mutaciones. En NEA y GNEA se pueden identificar bien sea puntos directos o sus *proxies* (representantes) y estratificar a los pacientes por la medida en que se ven afectados ciertos ciclos.

El enfoque propuesto para la predicción de sensibilidad basada en modelos multivariantes *elastic net* usando datos crudos, de NEA y de GNEA está lejos de ser impecable y tan sólo proporciona un punto de partida para explorar otras alternativas. Algunas direcciones posibles que deberían considerar trabajos a futuro serían:

- Investigar maneras de identificar los ciclos más adecuados para análisis de enriquecimiento (en lugar de utilizar existentes, muy frecuentemente irrelevantes, ciclos de terceros).
- Únicamente utilizar características robustas (aquellas que cuentan con perfiles más o menos enteros en todos los estudios) al momento de analizar transferencia de modelos.
- Validar los modelos utilizando perfiles de líneas celulares de cáncer generados de forma independiente.
- Encontrar datos de drogas más recientes o considerar utilizar datos de tratamiento vs. supervivencia del TCGA (*The Cancer Genome Atlas*, el Atlas del Genoma del Cáncer).
- Al momento de construir los conjuntos de genes alterados, comparar los valores en los perfiles genómicos de líneas celulares de cáncer con valores en perfiles genómicos de tejidos sanos, de modo que la magnitud de la diferencia entre los valores en el tumor y el tejido sano se utilice como medida de ordenamiento para establecer el *ranking* de los genes, seleccionando entonces los genes del tumor que más difieren con respecto al respectivo tejido sano.

Bibliografía

- [1] “molecular biology (noun) [Concise Encyclopedia].” Merriam-Webster’s online dictionary.
- [2] G. W. Beadle and E. L. Tatum, “Genetic Control of Biochemical Reactions in Neurospora,” *Proceedings of the National Academy of Sciences of the United States of America, PNAS*, vol. 27, pp. 499–506, November 1941.
- [3] J. D. Watson and F. H. C. Crick, “Molecular Structure of Nucleic Acids,” *Nature*, vol. 171, pp. 737–738, April 1953.
- [4] G. M. Church, “Genomes for All,” *Scientific American*, vol. 294, pp. 46–54, January 2006.
- [5] A. Pollack, “DNA Sequencing Caught in Deluge of Data,” *The New York Times*, November 2011.
- [6] A. S. of Clinical Oncology, “The Molecular Landscape of Cancer: Using Panomics to Drive Change,” *ASCO Vision Statement*, December 1969.
- [7] T. C. G. Atlas, “Impact of Cancer Genomics on Precision Medicine for the Treatment of Cancer,” *About Cancer Genomics*.

- [8] I. A. for Research on Cancer (IARC) and W. H. O. (WHO), “Cancer Fact Sheets (N°297),” *GLOBOCAN 2012; Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012*, 2012.
- [9] A. Basu, N. E. Bodycombe, *et al.*, “An Interactive Resource to Identify Cancer Genetic and Lineage Dependencies Targeted by Small Molecules,” *Cell*, vol. 154, pp. 1151–1161, August 2013.
- [10] J. Barretina, G. Caponigro, *et al.*, “The Cancer Cell Line Encyclopedia enables predictive modeling of anticancer drug sensitivity,” *Nature*, vol. 483, pp. 603–607, March 2012.
- [11] M. J. Garnett, E. J. Edelman, *et al.*, “Systematic identification of genomic markers of drug sensitivity in cancer cells,” *Nature*, vol. 483, pp. 570–575, March 2012.
- [12] B. Haibe-Kains, N. El-Hachem, *et al.*, “Inconsistency in large pharmacogenomic studies,” *Nature*, vol. 504, pp. 389–393, December 2013.
- [13] A. Alexeyenko, W. Lee, *et al.*, “Network enrichment analysis: extension of gene-set enrichment analysis to gene networks,” *BMC Bioinformatics*, vol. 13, September 2012.
- [14] “Emerging Technologies: Omics, Bioinformatics, Computational Biology. Overview,” *AltTox.org. Non-animal Methods for Toxicity Testing*, July 2014.
- [15] “What is “Omics” Science?,” *OMICS-ETHICS Research Group*.
- [16] “Definition: genome,” *Scitable by nature EDUCATION*.
- [17] “Definition: transcriptome,” *Scitable by nature EDUCATION*.
- [18] “Definition: proteome,” *Scitable by nature EDUCATION*.
- [19] G. Potters, “Systems Biology of the Cell,” *Nature Education*, vol. 3, no. 9, p. 33, 2010.

- [20] A. F. Gazdar, L. Girad, *et al.*, “Lung Cancer Cell Lines as Tools for Biomedical Discovery and Research,” *Journal of the National Cancer Institute*, vol. 102, pp. 1310–1321, September 8 2010.
- [21] “cell line. (n.d.).” The American Heritage®Medical Dictionary, 2007.
- [22] “cancer cell line.” NCI Dictionary of Cancer Terms.
- [23] D. A. M. Mandal, “What is gene expression?.” News Medical.
- [24] “Cancer Genetics and Genomics.” Yale Cancer Center (Research Programs).
- [25] P. Collins, Francis S. M.D., “Copy Number Variation (CNV).” National Human Genome Research Institute’s Talking Glossary of Genetic Terms.
- [26] N. R. Zhang, “DNA Copy Number Profiling in Normal and Tumor Genomes,” in *Frontiers in Computational and Systems Biology* (J. Feng, W. Fu, *et al.*, eds.), London: Springer, 2010.
- [27] “mutation.” NCI Dictionary of Cancer Terms.
- [28] N. H. G. R. Institute, “Biological Pathways,” *genome.gov Education Fact Sheets*, June 2012.
- [29] W. Kalow, “Pharmacogenetics and pharmacogenomics: origin, status and the hope for personalized medicine,” *The Pharmacogenomics Journal*, vol. 6, pp. 162–165, January 2006.
- [30] “*The Cancer Genome Atlas (tcga) research network.*”
- [31] P. Breheny, “Penalized regression: Introduction.” University of Kentucky BST764 (Applied Statistical Modeling for Medicine and Public Health) Course Notes, August 30 2011.

- [32] H. Zou and T. Hastie, “Regularization and Variable Selection via the Elastic Net.” Stanford University, Department of Statistics, April 21 2005.
- [33] H. Zou and T. Hastie, “Regularization and variable selection via the elastic net,” *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 67, no. 2, pp. 301–320, 2005.
- [34] M. Kanehisa and S. Goto *Nucleic Acids Research*, vol. 28, no. 1, pp. 27–30, 2000.
- [35] M. Kanehisa, S. Goto, *et al.* *Nucleic Acids Research*, vol. 42, pp. 199–205, November 2014.
- [36] T. G. O. Consortium, “Gene Ontology: tool for the unification of biology,” *Nature Genetics*, vol. 25, pp. 25–29, May 2000.
- [37] A. Subramanian, P. Tamayo, *et al.*, “Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, pp. 15545–15550, October 25 2005.
- [38] J. E. McCormack, M. G. Harvey, *et al.*, “A Phylogeny of Birds Based on Over 1,500 Loci Collected by Target Enrichment and High-Throughput Sequencing,” *PLOS ONE*, vol. 8, January 2013.
- [39] K. A. Eriksen, I. Simonsen, *et al.*, “Modularity and Extreme Edges of the Internet,” *Physical Review Letters*, vol. 90, pp. 289–300, April 2003.
- [40] A. Alexeyenko and E. L. L. Sonnhammer, “Global networks of functional coupling in eukaryotes from comprehensive data integration,” *Genome Research*, vol. 19, pp. 1107–1116, 2009.

- [41] D. Hanahan and R. A. Weinberg, “Hallmarks of Cancer: The Next Generation,” *Cell*, vol. 144, pp. 646–674, March 4 2011.
- [42] J. Friedman, T. Hastie, *et al.*, “Lasso and elastic-net regularized generalized linear models.” Package ‘glmnet’, May 24 2014.
- [43] J. Friedman, T. Hastie, *et al.*, “Regularization Path for Generalized Linear Models via Coordinate Descent,” *Journal of Statistical Software*, vol. 33, pp. 1–22, January 2010.

Apéndice A

Organización de los Datos de Trabajo

En la figura A.1 se muestra una representación esquemática de la estructura de R utilizada para organizar los conjuntos de datos descritos en el capítulo 2. Se trata de una lista de listas en R, donde cada nodo final es una matriz. Los nodos destacados en color **ocre** corresponden a matrices de datos de respuesta a drogas para cada estudio farmacogenómico (figura 2.1 (a)); los nodos en color **rojo** corresponden a matrices de datos crudos de perfiles ómicos (GE, CNA y mutaciones), figura 2.1 (b)); los nodos resaltados en colores **verde** y **anaranjado** corresponden a matrices de datos NEA (GE o CNA y mutaciones respectivamente), figura 3.1 (c3); y los nodos en colores **azul** y **morado** corresponden a matrices de datos GNEA (GE o CNA y mutaciones respectivamente), figura 3.1 (d3)³⁰.

Esta arquitectura facilitó el manejo y manipulación de datos al mantener todo en la misma estructura, permitiendo acceso a matrices específicas de forma cómoda y sencilla. Más aún, la anatomía actual permite la incorporación de nuevos conjuntos de datos (bien sea de las categorías ya existentes o de nuevas categorías que pueden añadirse fácilmente), e.g. los nodos destacados en color **marrón** corresponden a matrices de datos adicionales que podrían ser de interés para futuras investigaciones.

³⁰Los nombres de las matrices están indicados en la tabla 3.1.

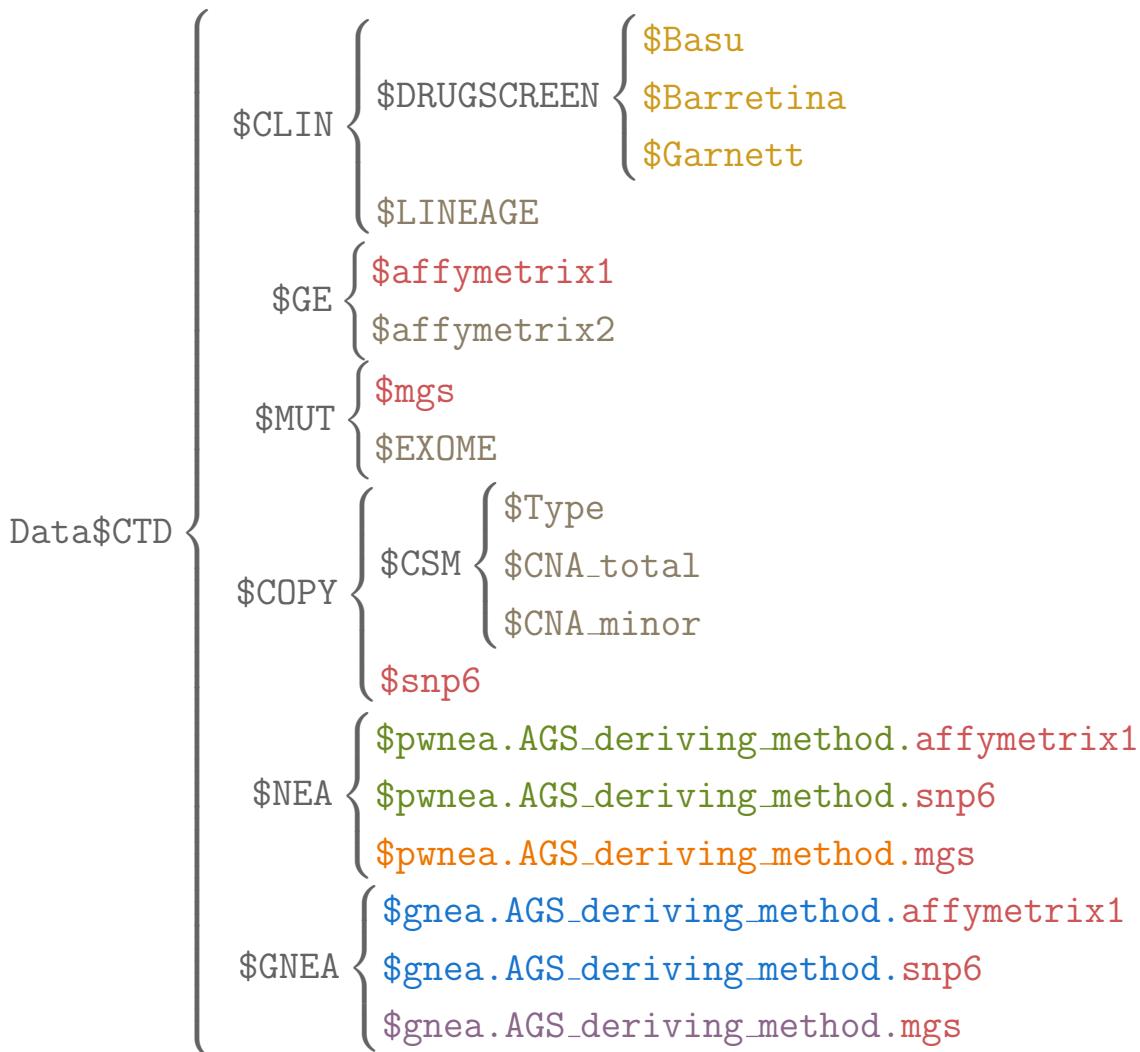


Figura A.1: Esquema de la estructura utilizada para organizar los datos de trabajo.

Apéndice B

Correlatos Uni-Característica

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
snp6	0.333	0.118	0.02	0	0	0	0	0	0
pwnea.top.200.snp6	0.186	0.039	0	0	0	0	0	0	0
pwnea.top.400.snp6	0.167	0.059	0.02	0	0	0	0	0	0
pwnea.significant.snp6	0.137	0	0	0	0	0	0	0	0
mgs	0.038	0	0	0	0	0	0	0	0
pwnea.total.mgs	0.275	0.069	0.01	0	0	0	0	0	0
pwnea.reference.mgs	0.284	0.088	0.01	0	0	0	0	0	0
pwnea.significant.filtered.maxi.snp6	0.245	0.127	0.059	0.02	0	0	0	0	0
gneea.top.200.snp6	0.167	0.098	0.069	0.029	0.01	0.01	0	0	0
gneea.top.400.snp6	0.235	0.147	0.069	0.029	0.02	0	0	0	0
gneea.significant.filtered.maxi.snp6	0.343	0.127	0.059	0.029	0.01	0	0	0	0
gneea.total.mgs	0.314	0.127	0.039	0.039	0.01	0	0	0	0
gneea.reference.mgs	0.363	0.147	0.078	0.039	0.02	0	0	0	0
affymetrix1	0.402	0.196	0.088	0.069	0.01	0	0	0	0
gneea.significant.snp6	0.255	0.186	0.127	0.069	0.039	0.029	0	0	0
pwnea.significant.affymetrix1	0.471	0.363	0.196	0.137	0.049	0.029	0.02	0	0
pwnea.significant.filtered.maxi.affymetrix1	0.49	0.353	0.225	0.147	0.098	0.059	0.029	0.01	0
pwnea.top.200.affymetrix1	0.471	0.382	0.275	0.157	0.108	0.069	0.029	0.01	0
gneea.top.400.affymetrix1	0.559	0.373	0.265	0.167	0.127	0.078	0.049	0.029	0
gneea.significant.affymetrix1	0.539	0.353	0.265	0.196	0.108	0.088	0.039	0	0
pwnea.top.400.affymetrix1	0.559	0.461	0.343	0.216	0.108	0.069	0.039	0.02	0
gneea.top.200.affymetrix1	0.608	0.48	0.353	0.225	0.157	0.098	0.059	0.029	0.01
gneea.significant.filtered.maxi.affymetrix1	0.578	0.471	0.304	0.235	0.147	0.078	0.01	0	0

Tabla B.1: Fracciones de correlatos uni-característica.

Cada celda en la tabla B.1 indica la fracción de correlatos, para el tipo de dato y perfil asociado a la fila, cuyo valor de correlación r_s excede el valor indicado al tope de la columna. Se asocian colores a cada par de valores de correlación consecutivos (ver columnas de la tabla) y el color asociado a cada fila representa el valor más alto de correlación alcanzado para el tipo de dato y perfil representado por dicha fila. Entonces, por ejemplo, el color de la sexta fila, de abajo hacia arriba, `pwnea.top.400.affymetrix1` comunica que algunos correlatos asociados a `pwnea.top.400.affymetrix1` exceden el valor **0.8**.

Apéndice C

Exploración de Diferentes Niveles de Complejidad

Las estadísticas de r_s , p -valor y MSE para los modelos multivariantes *elastic net* ajustados según los procedimientos descritos en la sección 3.3 se resumen en diagramas de caja en las páginas que siguen.

Para cada combinación de estudio farmacogenómico y estadística, se presenta una matriz de diagramas de cajas:

- Las columnas de esta matriz indican el nivel de complejidad de los modelos, que decrece de izquierda a derecha (complejidad ilimitada, complejidad limitada a 25 características y complejidad limitada a 10 características).
- Cada par de filas indica un tipo de perfil utilizado: expresión de genes (GE)/**affymetrix1**, alteraciones en el número de copias (CNA)/**snp6** o mutaciones puntuales/**mgs**. La primera fila del par reporta resultados para NEA, la segunda reporta resultados para GNEA.

- \bar{N} (debajo de cada caja) es el número promedio de características seleccionadas por el conjunto de modelos representados por la caja.
- Se presentan resultados para diversos conjuntos de datos crudos, NEA y GNEA (ver las matrices listadas en la tabla 3.1). Una leyenda de colores para identificar los distintos conjuntos de datos crudos y de enriquecimiento se provee al final de cada matriz de diagramas de cajas y también puede visualizarse en la figura 4.2.

Estudio \ Estadística	r_s	p-value	MSE
Basu	Figura C.1	Figura C.2	Figura C.3
Barretina	Figura C.4	Figura C.5	Figura C.6
Garnett	Figura C.7	Figura C.8	Figura C.9

Tabla C.1: Leyenda de figuras para modelos con distintos niveles de complejidad.

La tabla C.1 provee una leyenda para facilitar la búsqueda de figuras de acuerdo al estudio farmacogenómico y la estadística medida (tenga en cuenta que las figuras en el documento aparecen en orden de número ascendente). Por ejemplo, la figura C.1 reporta r_s para los modelos ajustados utilizando como respuesta datos de sensibilidad de **Basu**.

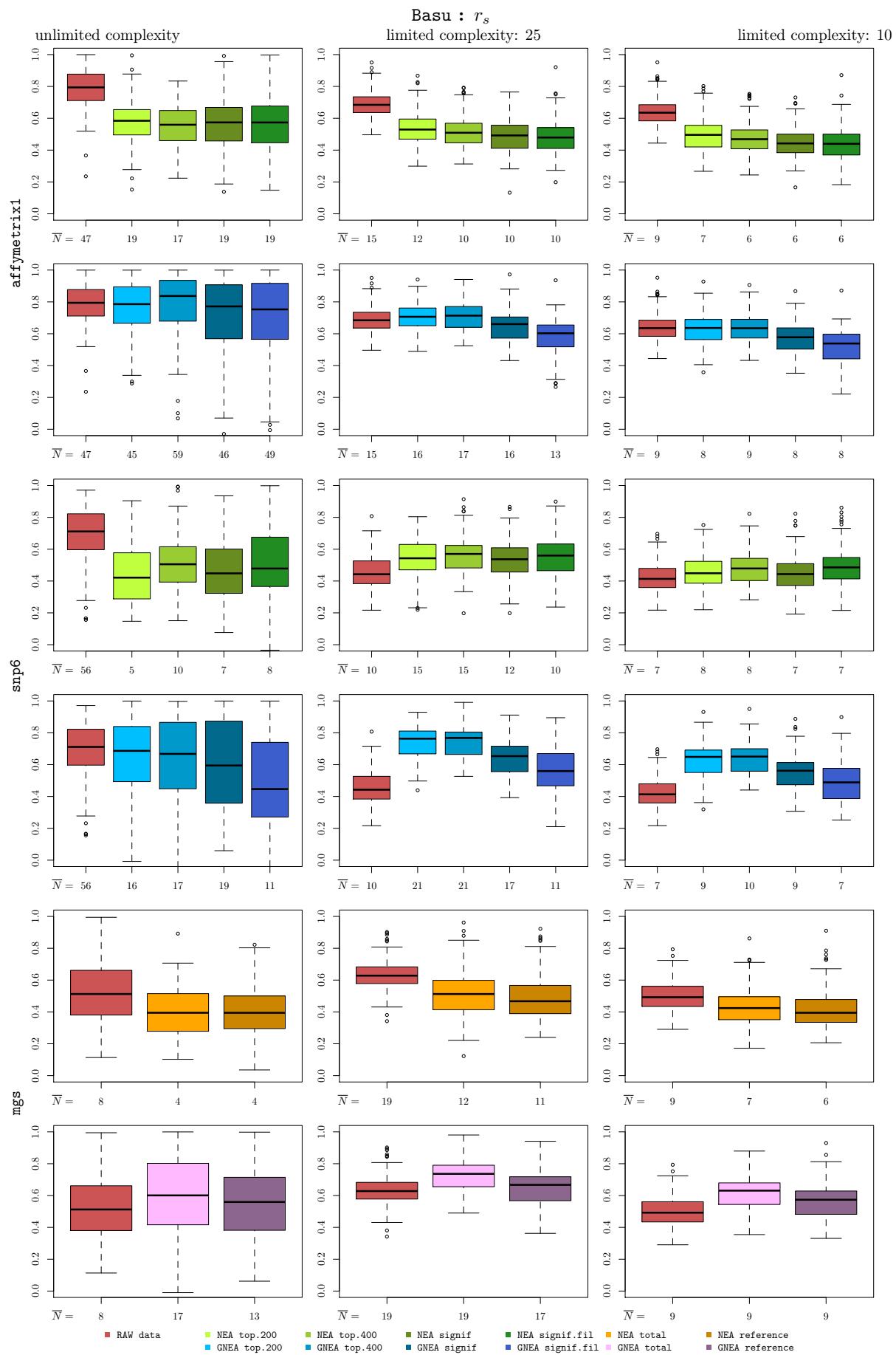
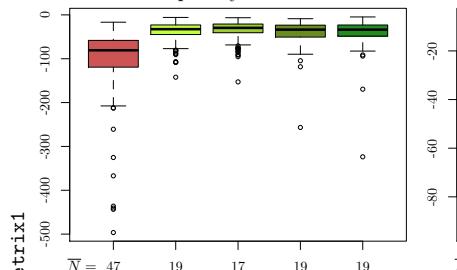
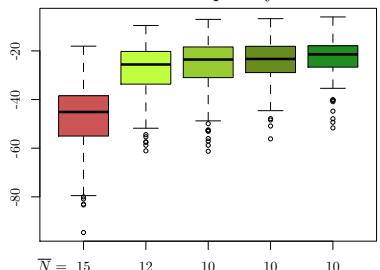


Figura C.1: Modelos de diversa complejidad, Basu, r_s

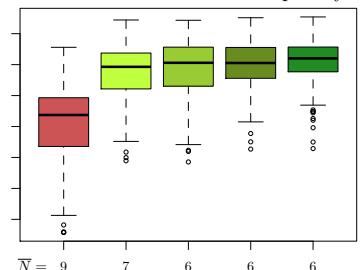
Basu : p -value
unlimited complexity



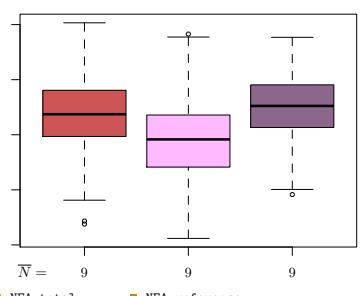
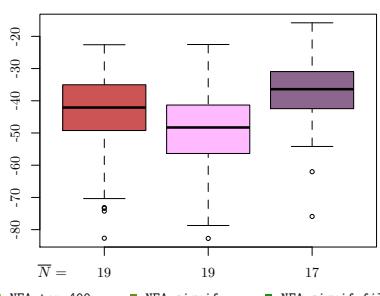
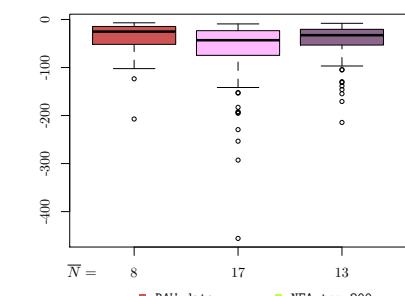
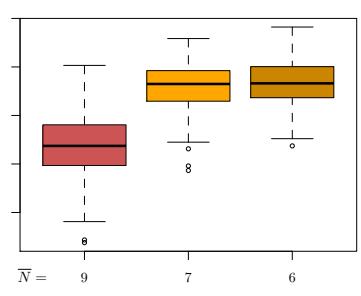
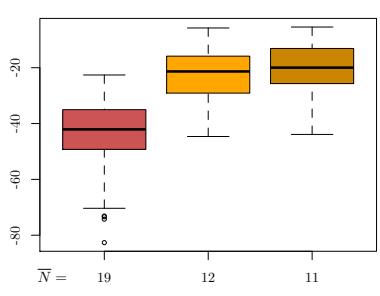
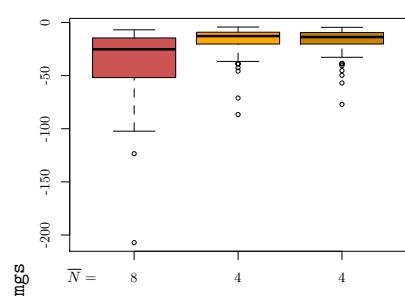
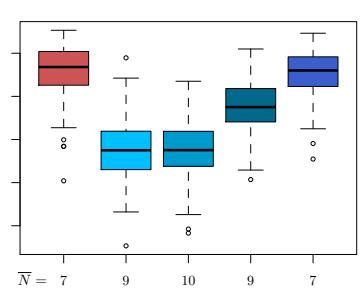
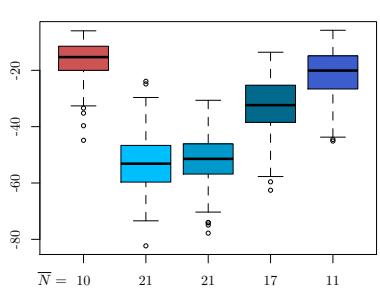
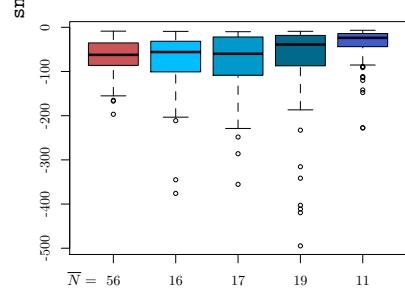
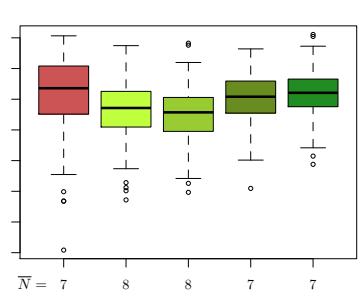
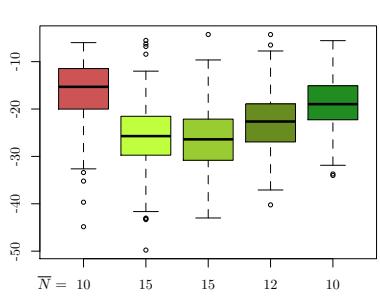
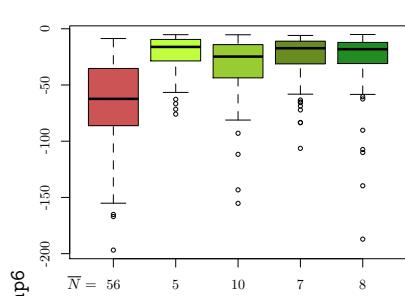
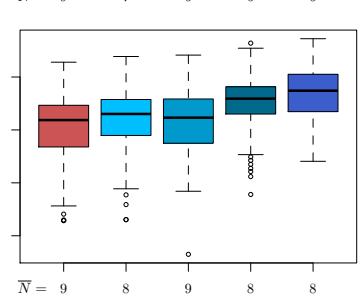
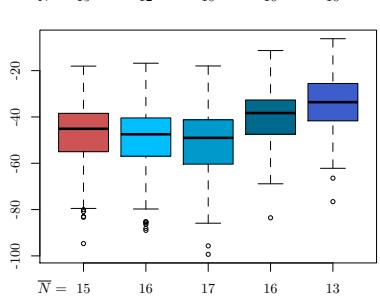
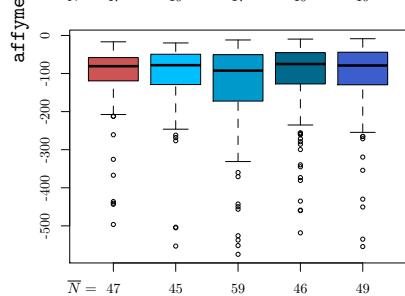
limited complexity: 25



limited complexity: 10



66



■ RAW data ■ NEA top.200 ■ GNEA top.200 ■ NEA top.400 ■ GNEA top.400 ■ NEA signif. fil ■ GNEA signif. fil ■ NEA total ■ GNEA total ■ NEA reference ■ GNEA reference

Figura C.2: Modelos de diversa complejidad, Basu, p -value

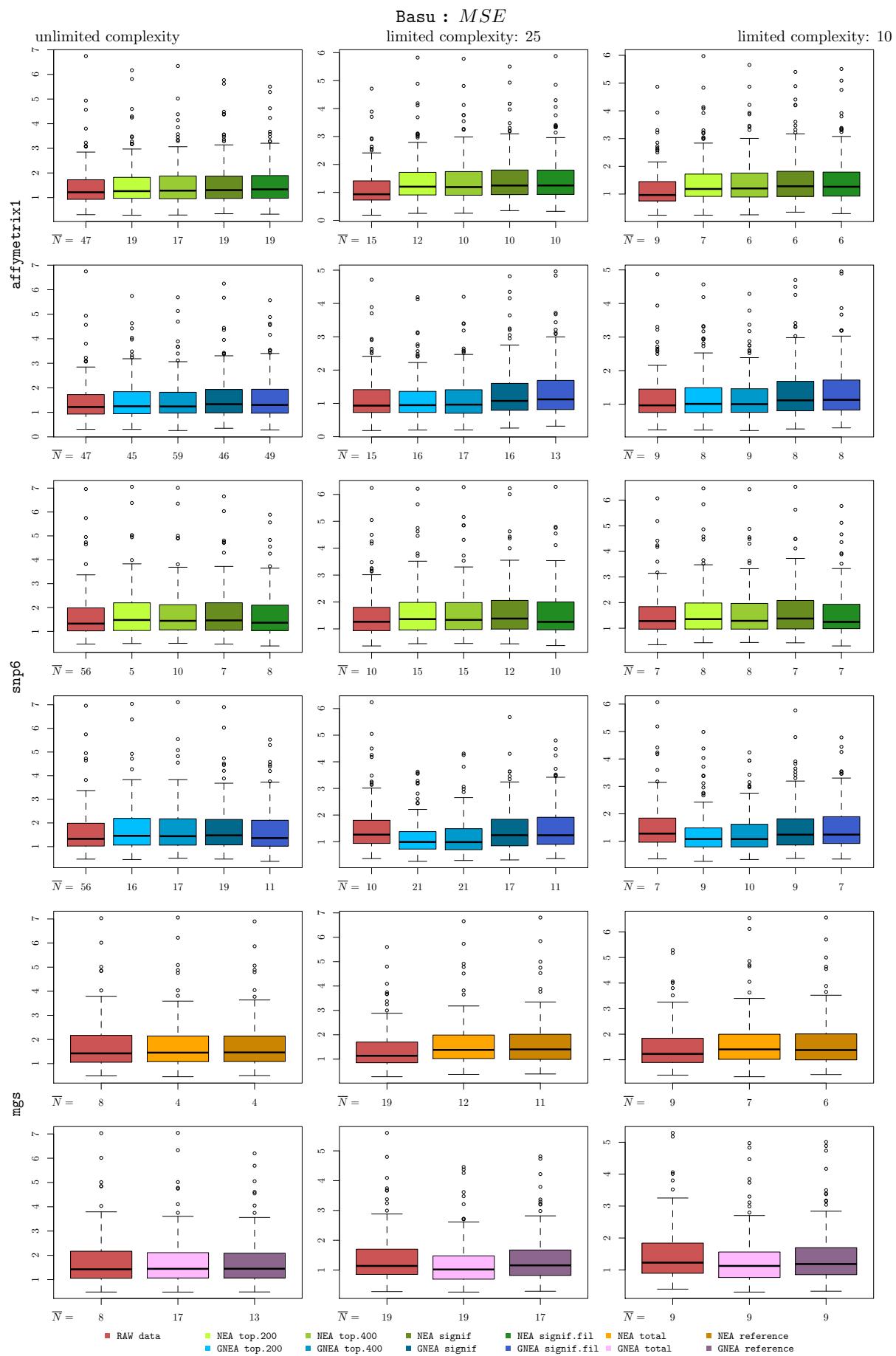
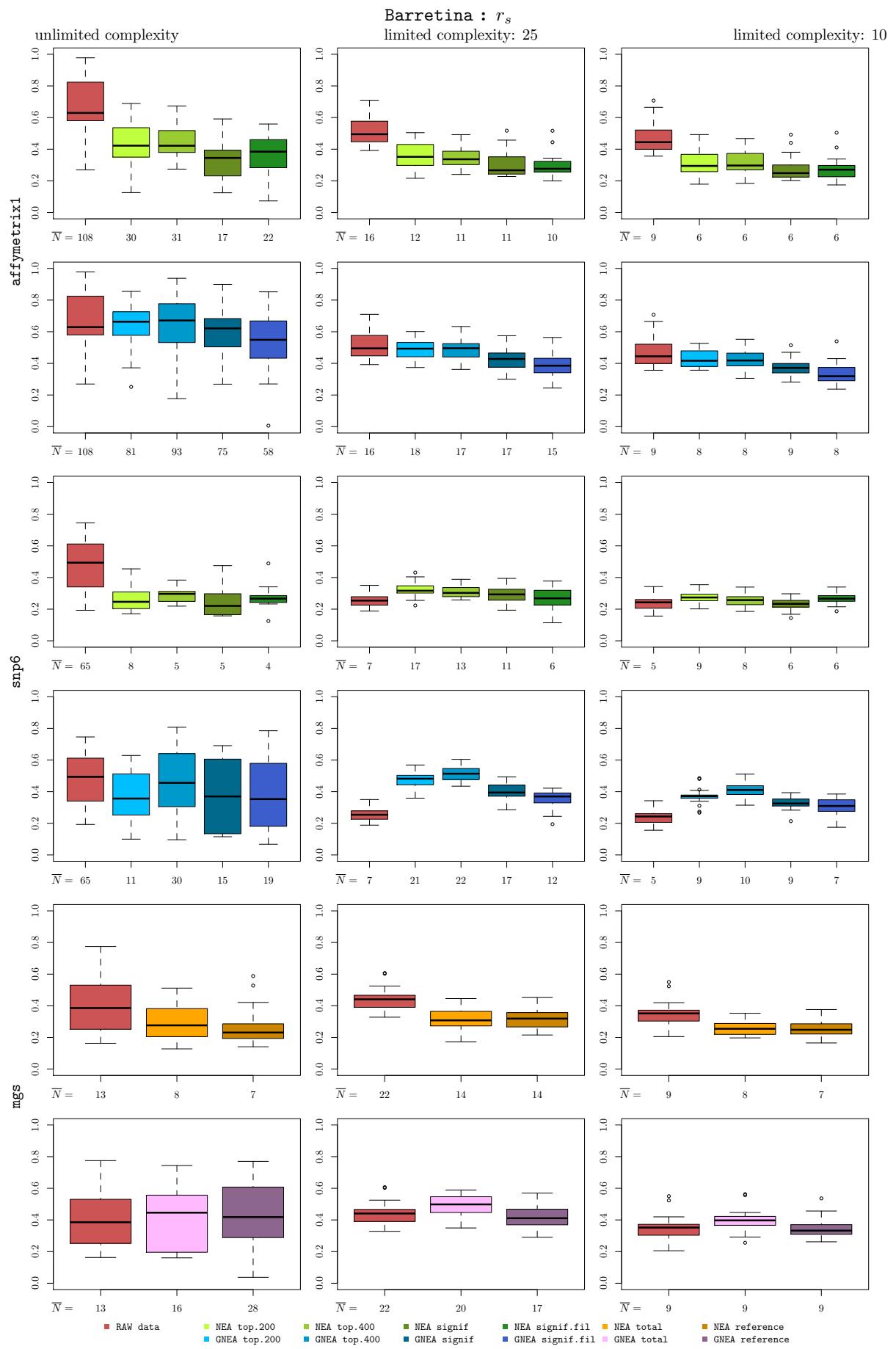


Figura C.3: Modelos de diversa complejidad, Basu, MSE

Figura C.4: Modelos de diversa complejidad, Barretina, r_s

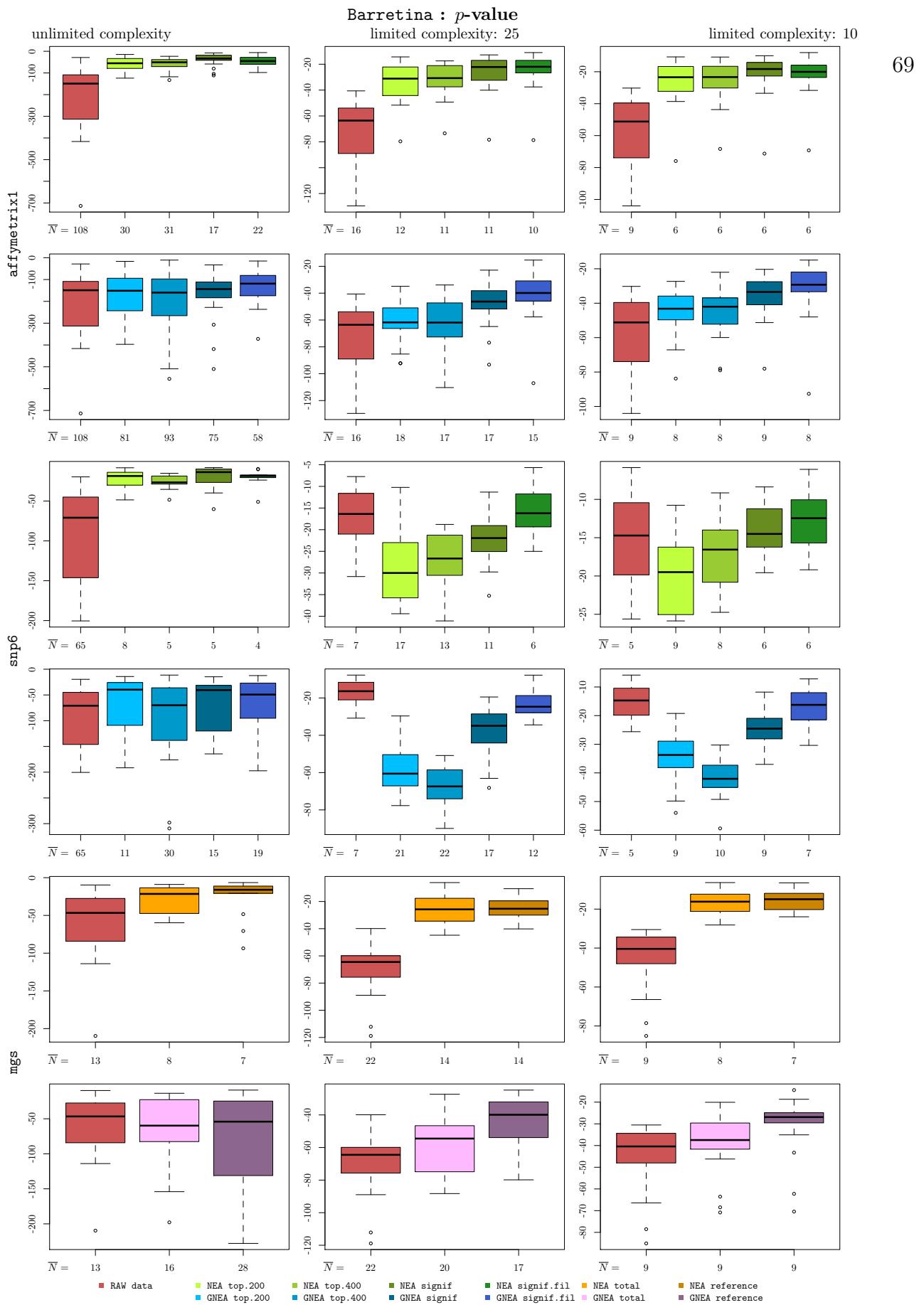


Figura C.5: Modelos de diversa complejidad, Barretina, p -value

Barretina : MSE

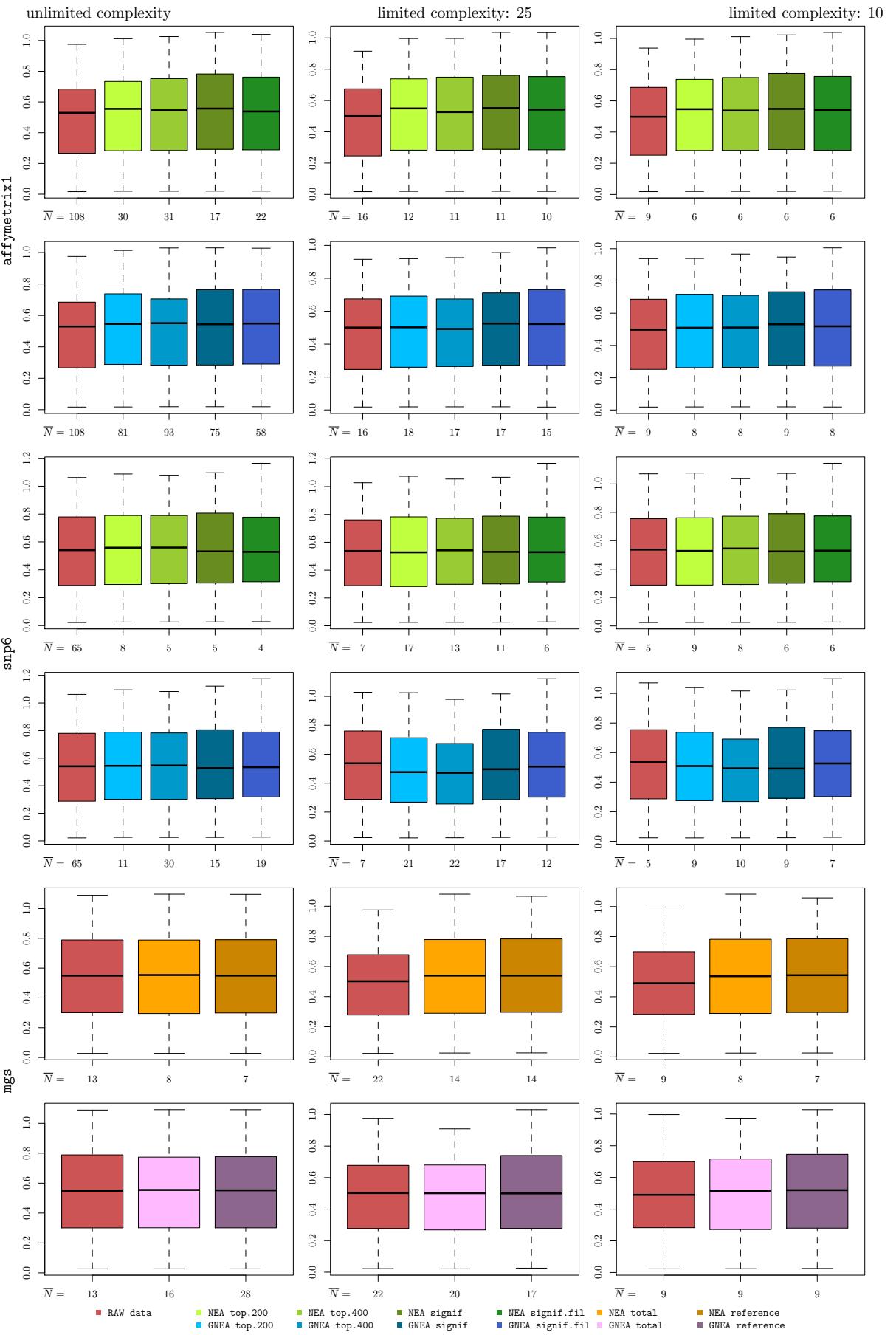
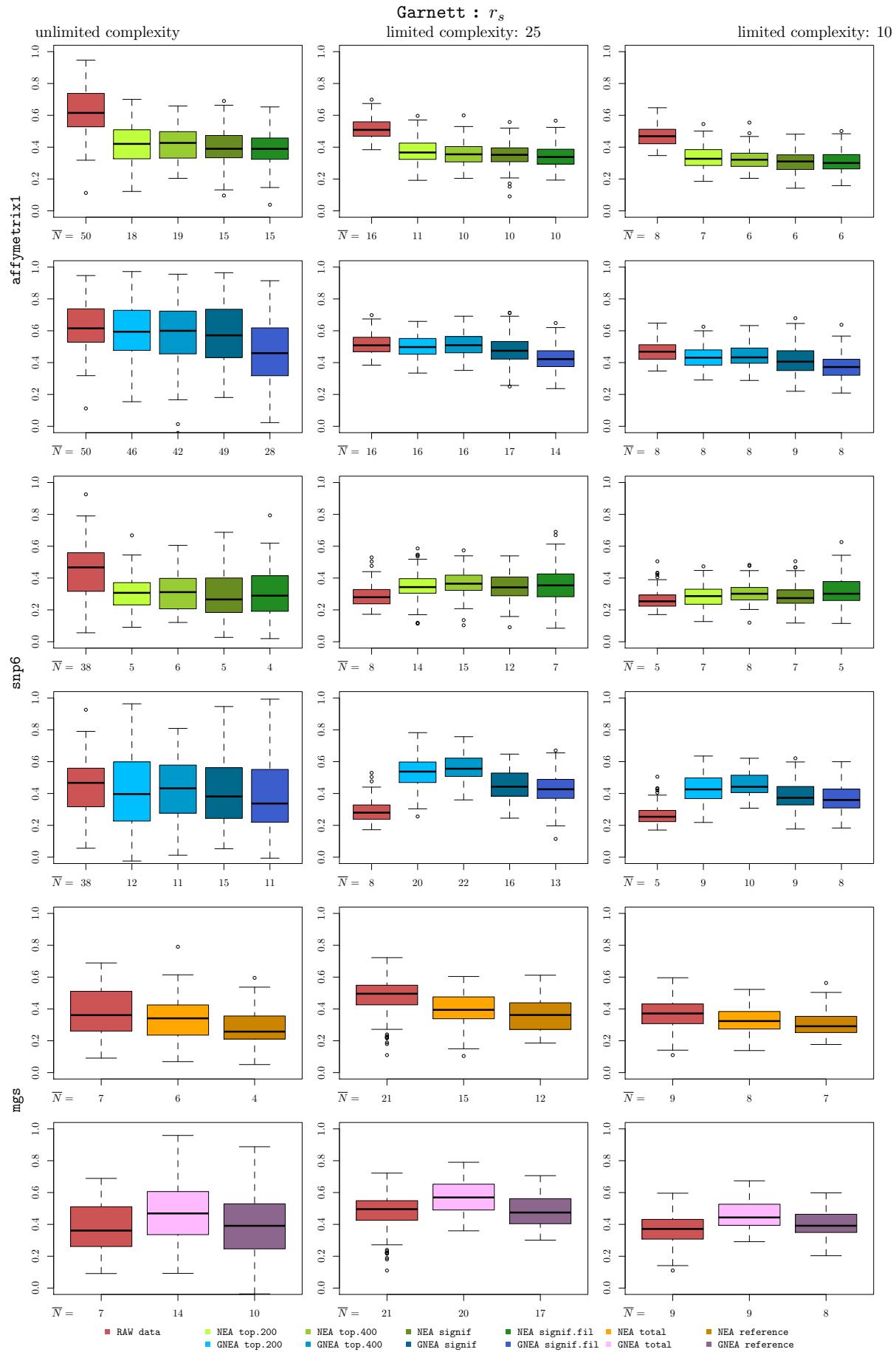
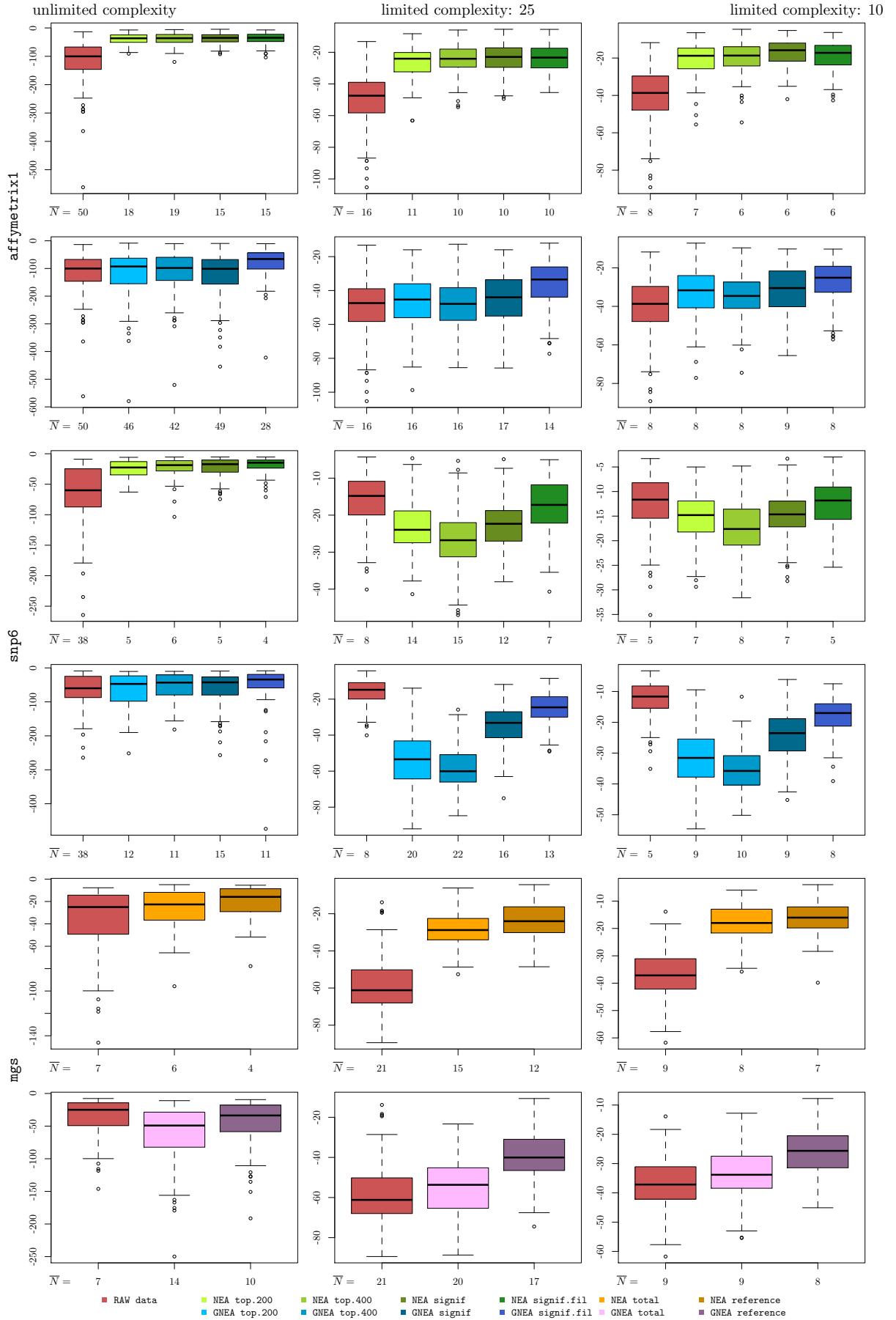


Figura C.6: Modelos de diversa complejidad, Barretina, *MSE*


 Figura C.7: Modelos de diversa complejidad, Garnett, r_s

Garnett : p -value



72

Figura C.8: Modelos de diversa complejidad, Garnett, p -value

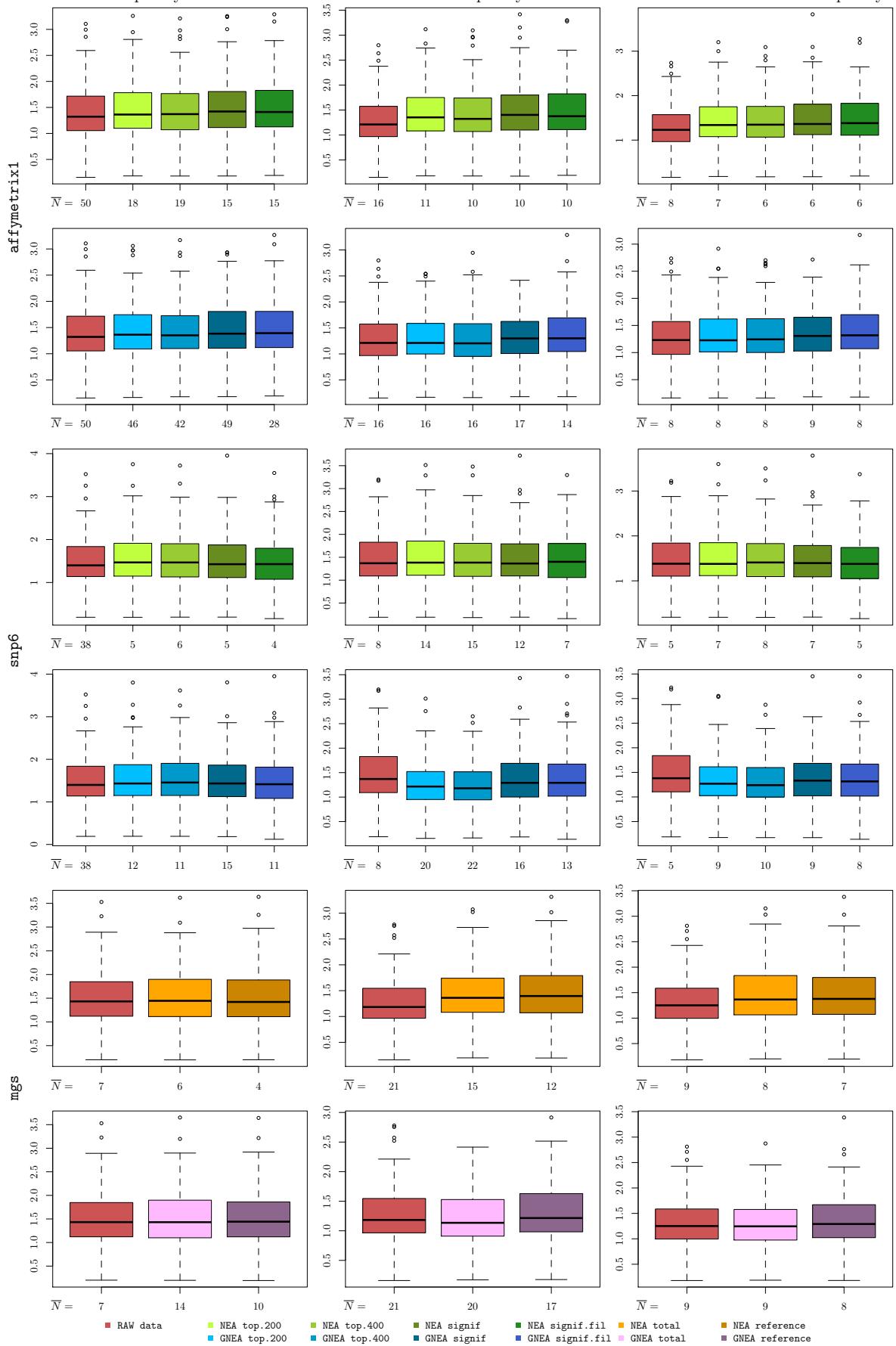


Figura C.9: Modelos de diversa complejidad, Garnett, MSE

Apéndice D

Modelos para Drogas Compartidas

Este apéndice presenta un compendio de tablas y diagramas de cajas que resumen las estadísticas de r_s , p -valor y MSE asociadas a modelos primarios (P) y transferidos (T) ajustados utilizando el enfoque descrito en la sección 4.3.

Para cada par de estudios farmacogenómicos (A, B) (con $A, B \in \{\text{Basu}, \text{Barretina}, \text{Garnett}\}$ y $A \neq B$), las estadísticas mencionadas, asociadas a los modelos multivariantes *elastic net* (de complejidad ilimitada) para drogas comunes a A y B , se presentan en tablas y diagramas de caja, discerniendo por tipo de dato, tipo de perfil como también por P y T (reportamos tanto el caso $P = A$ y $T = B$ como el caso $P = B$ y $T = A$).

Para cada combinación posible de P y T asociada a un par de estudios farmacogenómicos reportamos:

1. Una tabla:

- Las filas de la tabla corresponden a drogas evaluadas por ambos estudios. Aplica aquí el mismo esquema de codificación de colores utilizado en la tabla D.3.

- Cada fila comprende tres sub-filas (una para cada tipo de perfil).
- Las columnas corresponden a tipos de datos (crudos, NEA y GNEA; los últimos dos refinados por método de derivación de AGSs).
- Cada columna está compuesta por dos sub-columnas (una reporta el valor de la estadística para el modelo P y la otra reporta el mismo valor estadístico para el modelo T).
- Un asterisco azul, *, al lado del valor de una celda indica que no se pudo ajustar un modelo de *elastic net* para esa droga particular usando los datos de respuesta, i. e. estudio farmacogenómico, y entrada, i. e. tipo de perfil y dato, en cuestión (respectivos a la fila y columna).
- Un asterisco rojo, *, al lado del valor de una celda indica que el modelo asociado en cuestión (considerando la combinación particular de droga, estudio farmacogenómico, tipo de perfil y tipo de dato) seleccionó cero predictores (variables) al momento de realizar la regresión, en otras palabras, el único coeficiente distinto de cero es el intercepto β_0 .

2. Una columna de diagramas de cajas:

- Las filas indican el tipo de perfil empleado: expresión de genes (GE)/**affymetrix1**, alteraciones en el número de copias (CNA)/**snp6** o mutaciones/**mgs**.
- Se presentan resultados para cada conjunto de datos crudos, NEA y GNEA (ver las matrices listadas en la tabla 3.1) como pares de cajas, la de la izquierda para modelos primarios (P) y la de la derecha para modelos transferidos (T). Una leyenda de colores permite la identificación correcta de cada tipo de dato (incluyendo el refinamiento por método de derivación de AGSs en los casos NEA y GNEA), la leyenda puede apreciarse con mayor claridad en la figura 4.2.

- \bar{N} (debajo de cada caja) es el número promedio de características seleccionadas por el grupo de modelos asociados a la caja.

Estudio		Estadística		
P	T	r_s	p-valor	MSE
Basu	Barertina	D.4	D.6	D.8
Barretina	Basu	D.5	D.7	D.9
Basu	Garnett	D.10 & D.11	D.14 & D.15	D.18 & D.19
Garnett	Basu	D.12 & D.13	D.16 & D.17	D.20 & D.21
Barretina	Garnett	D.22	D.24	D.26
Garnett	Barretina	D.23	D.25	D.27

Tabla D.1: Leyenda de tablas para comparación de modelos entre estudios.

La tabla D.1 provee una leyenda para facilitar la consulta de tablas de acuerdo a la combinación estudio farmacogenómico P - T y estadística medida (téngase en cuenta que las tablas aparecen en orden de número ascendente). Por ejemplo, la tabla D.4 reporta, para cada droga estudiada por tanto Basu como Barretina, los valores r_s para modelos ajustados utilizando los posibles datos de entrada (ver tabla 3.1) con Basu siendo el estudio farmacogenómico P y a Barretina siendo el estudio farmacogenómico T.

Estudio		Estadística		
P	T	r_s	p-valor	MSE
Basu	Barertina	D.1	D.3	D.5
Barretina	Basu	D.2	D.4	D.6
Basu	Garnett	D.7	D.9	D.11
Garnett	Basu	D.8	D.10	D.12
Barretina	Garnett	D.13	D.15	D.17
Garnett	Barretina	D.14	D.16	D.18

Tabla D.2: Leyenda de figuras para comparación de modelos entre estudios.

La tabla D.2 provee una leyenda para facilitar la consulta de figuras (con diagramas de cajas) de acuerdo a la combinación estudio farmacogenómico P - T y estadística medida (téngase en cuenta que las tablas aparecen en orden de número ascendente). Por ejemplo, la figura D.4 contiene diagramas de cajas para distribuciones de valores de r_s para los modelos ajustados teniendo a Basu como el estudio farmacogenómico P y a Barretina como el estudio farmacogenómico T (los modelos están agrupados por tipo de perfil y de dato, refinando en este nivel por método de derivación de AGSs en el caso de datos NEA y GNEA).

D.1. Lista de Drogas Comunes

Basu - Barretina (CTRP - CCLE)	Basu - Garnett (CTRP - CGP)	Barretina - Garnett (CCLE - CGP)
nutlin selumetinib tanespimycin 1685458	nutlin selumetinib tanespimycin atra bez235 bi2536 bms536924 bmx754807 bortezomib bx795 chir99021 etoposide gemcitabine gw843682x linsitinib mk2206 navitoclax nvpauy922 olaparib pac1 pathenolide qs11 sirolimus sn38 tipifarnibp1 tozasertib tw37 vorinostat	nutlin selumetinib tanespimycin azd0530 erlotinib lapatinib nilotinib paclitaxel pd0325901 pd0332991 crizotinib pha665752 plx4720 sorafenib nvptae684

Tabla D.3: Drogas compartidas entre diferentes estudios farmacogenómicos.

En la tabla D.3, las drogas en color rojo son comunes todos los (tres) estudios farmacogenómicos.

Primary (P): Basu, Transferred (T): Barretina

		r_s																											
		NEA														GNEA													
		RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference			
1685458	affy1	NA*	NA*	NA*	0.3	NA*	0.14	NA*	0.27	NA*	NA*					0.29	0.37	NA*	0.35	0.08	NA*	0.05	0.05						
	snp6	NA*	0.27	NA*	NA*	0.37	0.12	0.76	NA*	0.3	NA*					NA*	NA*	-0.05	-0.07	NA*	NA*	NA*	NA*						
	mgs	NA*	0.16									NA*	NA*	0.23	NA*									0.19	NA*	NA*	NA*		
nutlin	affy1	0.93	0.58	0.41	0.03	0.43	0.5	0.46	0.22	0.52	0.3					NA*	0.48	0.74	0.14	0.6	0.46	0.71	0.38						
	snp6	0.86	0.65	0.37	0.3	0.61	NA*	NA*	NA*	NA*	NA*					0.87	NA*	0.77	NA*	0.82	NA*	0.43	0.28						
	mgs	0.64	0.32									0.45	NA*	0.54	0.28									0.98	0.5	0.46	0.38		
selumetinib	affy1	0.88	0.82	0.91	0.74	0.81	0.75	0.7	0.75	0.15	0.52					0.93	0.7	0.97	0.77	0.8	0.65	0.66	0.52						
	snp6	NA*	0.79	0.57	NA*	0.57	NA*	0.77	0.37	0.43	NA*					NA*	NA*	0.91	NA*	0.96	0.39	1	0.33						
	mgs	0.82	0.7									0.56	0.15	0.41	0.43									0.75	0.67	0.61	0.7		
tanespimycin	affy1	0.8	0.37	0.55	0.45	NA*	0.6	NA*	NA*	NA*	0.53					0.74	0.62	0.39	0.61	0.94	NA*	0.62	0.78						
	snp6	NA*	0.7	NA*	NA*	NA*	NA*	NA*	NA*	NA*	NA*					NA*	NA*	NA*	0.26	NA*	NA*	NA*	NA*						
	mgs	NA*	NA*									NA*	0.21	NA*	NA*									0.26	NA*	NA*	0.32		

Tabla D.4: Modelos para drogas compartidas, P: Basu, T: Barretina, r_s

		Primary (P): Barretina, Transferred (T): Basu														r_s														
		NEA														GNEA														
		RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference				
1685458	affy1	0.63	NA*	0.41	NA*	0.36	NA*	0.35	NA*	0.32	NA*					0.58	0.25	0.53	NA*	0.68	NA*	0.54	NA*							
	snp6	0.41	NA*	NA*	NA*	0.25	0.19	NA*	NA*	0.24	NA*					NA*	NA*	0.1	-0.05	NA*	NA*	0.36	NA*							
	mgs	0.38	NA*									0.26	NA*	0.19	NA*									0.19	0.04	0.18	NA*			
nutlin	affy1	0.51	0.37	0.23	0.15	0.27	0.24	0.19	0.16	0.22	0.17					0.67	NA*	0.18	0.16	0.45	NA*	0.35	0.23							
	snp6	0.38	0.35	0.35	0.11	NA*	NA*	NA*	NA*	0.34	NA*					0.36	NA*	NA*	0.21	NA*	0	0.79	NA*				0.45	0.22	0.36	0.25
	mgs	0.44	0.27									0.21	0.09	0.23	0.24															
selumetinib	affy1	0.86	0.57	0.62	0.56	0.67	0.49	0.52	0.31	0.46	0.07					0.85	0.49	0.93	0.58	0.75	0.33	0.68	0.07							
	snp6	0.75	NA*	NA*	0.18	NA*	0.15	0.33	0.26	NA*	0					0.25	NA*	NA*	0.28	0.64	0.2	0.65	NA*				0.72	0.35	0.68	0.4
	mgs	0.77	0.53									0.21	NA*	0.59	0.25															
tanespimycin	affy1	0.97	0.58	0.53	0.31	0.55	NA*	0.24	NA*	0.41	NA*					0.74	0.33	0.82	0.1	0.66	0.33	0.85	0.22							
	snp6	0.66	NA*	0.31	NA*	0.3	NA*	0.16	NA*	NA*	NA*					NA*	NA*	0.36	NA*	0.13	NA*	0.12	NA*				0.2	0.06	0.61	NA*
	mgs	0.25	NA*									0.13	NA*	0.29	NA*															

Tabla D.5: Modelos para drogas compartidas, P: Barretina, T: Basu, r_s

Primary (P): Basu, Transferred (T): Barretina
p-value

	NEA														GNEA												
	RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference		
	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	
1685458	affy1	NA*	NA*	NA*	-7.11	NA*	-6.11	NA*	-7.26	NA*	NA*					-32.75	-9.47	NA*	-17.98	-24.53	NA*	-24.66	-15.96				
	snp6	NA*	-9.3	NA*	NA*	-20.41	-5.6	-83.32	NA*	-20.16	NA*					NA*	NA*	-25.13	-0.03	NA*	NA*	NA*	NA*				
	mgs	NA*	-4.33							NA*	NA*	-6.52	NA*											-24.77	NA*	NA*	NA*
nutlin	affy1	-147.28	-24.28	-16.61	-3.67	-16.5	-17.74	-13.79	-5.7	-21.19	-8.44					NA*	-19.05	-49.36	-0.58	-44.04	-20.77	-62.44	-16.65				
	snp6	-80.84	-26.59	-11.52	-5.15	-29.3	NA*	NA*	NA*	NA*	NA*					-109.17	NA*	-64.24	NA*	-75.38	NA*	-23.33	-6.44				
	mgs	-38.38	-10.92							-18.07	NA*	-29.45	-5.92											-195.2	-16.01	-23.32	-14.22
selumetinib	affy1	-87.15	-57.98	-90.53	-40.86	-57.37	-40.35	-38.6	-46.92	-6.7	-18.46					-119.68	-37.76	-172.62	-53.6	-57.17	-36.56	-39.93	-24.51				
	snp6	NA*	-48.81	-26.04	NA*	-24.75	NA*	-43.69	-13.07	-12.63	NA*					NA*	NA*	-97.57	NA*	-136.79	-11.92	-227.83	-8.64				
	mgs	-56.79	-31.6							-18.57	-2.38	-9.56	-10.76											-45.7	-31.34	-24.63	-38.69
tanespimycin	affy1	-70.79	-11.66	-28.76	-19.85	NA*	-37.36	NA*	NA*	NA*	-23.6					-66.15	-32.56	-12.43	-31.88	-168.16	NA*	-36.4	-63.93				
	snp6	NA*	-46.47	NA*	NA*	NA*	NA*	NA*	-0.24	NA*	NA*					NA*	NA*	NA*	-9.3	NA*	NA*	NA*	NA*				
	mgs	NA*	NA*							NA*	-6.02	NA*	NA*											-10.88	NA*	NA*	-7.62

Tabla D.6: Modelos para drogas compartidas, P: Basu, T: Barretina, p-value

Primary (P): Barretina, Transferred (T): Basu
p-value

	NEA														GNEA												
	RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference		
	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	
1685458	affy1	-131.54	NA*	-42.05	NA*	-35.35	NA*	-33.29	NA*	-33.64	NA*					-95.17	-12.83	-76.2	NA*	-151.91	NA*	-93.98	NA*				
	snp6	-45.46	NA*	NA*	NA*	-16.93	-10.88	NA*	NA*	-9.7	NA*					NA*	NA*	-11.42	-1.55	NA*	NA*	-31.41	NA*				
	mgs	-29.15	NA*							-14.66	NA*	-10.24	NA*											-15.22	-2.68	-15.44	NA*
nutlin	affy1	-80.38	-40.07	-16.85	-6.52	-23.11	-14.37	-11.77	-8.83	-15.62	-8.95					-178.61	NA*	-10.9	-10.43	-73.86	NA*	-52.43	-15.46				
	snp6	-49.81	34.81	-40.45	-3.9	NA*	NA*	NA*	NA*	-23.7	NA*					43.81	NA*	NA*	-12.96	NA*	-4.89	-197.31	NA*				
	mgs	-65.11	-21.3							-15.7	-4.63	-20.71	-21.85											-66.61	-7.05	-43.09	-15.73
selumetinib	affy1	-328.79	-85.27	-109.86	-85.21	-132.74	-59.41	-79.82	-27.85	-62.14	-3.86					-340.71	-75.08	-509.23	-104.08	-227.73	-32.43	-162.02	-10.57				
	snp6	-200.45	NA*	NA*	-11.47	NA*	-7.43	-27.69	-17.13	NA*	-1.84					-25.89	NA*	NA*	-28.42	-140.86	-12.22	-117.29	NA*				
	mgs	-209.67	-64.21							-10.33	NA*	-93.65	-16.24											-154.04	-26.97	-131.1	-35.17
tanespimycin	affy1	-713.63	-101.9	-87.91	-30.88	-86.11	NA*	-20.79	NA*	-57.98	NA*					-223.91	-29.66	-323.22	-2.67	-176.71	-38.61	-372.14	-16.08				
	snp6	-189.05	NA*	-30.01	NA*	-27.16	NA*	-8.55	NA*	NA*	NA*					NA*	NA*	-55.92	NA*	-21.29	NA*	-15.65	NA*				
	mgs	-27.49	NA*							-13.4	NA*	-19.4	NA*											-20.05	-2.28	-150.06	NA*

Tabla D.7: Modelos para drogas compartidas, P: Barretina, T: Basu, p-value

Primary (P): Basu, Transferred (T): Barretina
MSE

		NEA												GNEA													
		RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference	
		P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T
1685458	affy1	0.55*	0.56*	0.56*	0.55	0.56*	0.56	0.57*	0.57	0.58*	0.58*					0.58	0.55	0.58*	0.52	0.56	0.56*	0.56	0.55				
	snp6	0.56*	0.55	0.57*	NA*	0.56	0.57	0.58	NA*	0.61	0.63*					0.57*	NA*	0.57	NA	0.61*	NA*	0.62*	0.63*				
	mgs	0.59*	0.57									0.58*	0.58*	0.58	0.58*										0.58	0.58*	0.58*
nutlin	affy1	1.89	1.99	2.4	2.43	2.28	2.25	2.38	2.33	2.42	2.29					2.47*	2.28	2.41	NA	2.49	2.16	2.36	2.26				
	snp6	2.43	2.09	2.46	2.49	2.39	NA*	2.31*	NA*	2.32*	2.26*					2.49	2.57*	2.19	NA*	2.24	NA*	2.28	2.26				
	mgs	2.47	2.41									2.56	2.54*	2.4	2.4										2.29	2.35	2.54
selumetinib	affy1	2.12	1.95	2.54	2.33	2.46	2.61	2.75	2.46	2.96	2.89					2.33	2.26	2.14	1.99	2.75	2.45	2.87	2.62				
	snp6	3.01*	2.88	2.93	NA*	2.86	NA*	2.64	2.68	2.81	NA*					2.99*	3.12*	2.95	NA*	2.88	2.95	2.63	2.85				
	mgs	2.78	2.8									2.82	3.02	2.96	2.91										2.96	2.69	2.84
tanespimycin	affy1	1.68	1.73	1.84	1.71	1.84*	1.82	1.86*	1.84*	1.83*	1.83					1.86	1.77	1.85	1.65	1.72	1.85*	1.71	1.79				
	snp6	1.78*	1.72	1.77*	1.75*	1.76*	1.74*	1.73*	NA	1.82*	NA*					1.75*	NA*	1.76*	1.69	1.72*	1.74*	1.83*	1.85*				
	mgs	1.91*	1.87*									1.92*	1.81	1.91*	1.9*										1.87	1.89*	1.91*

Tabla D.8: Modelos para drogas compartidas, P: Basu, T: Barretina, *MSE*

Primary (P): Barretina, Transferred (T): Basu
MSE

		NEA												GNEA															
		RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference			
		P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T		
1685458	affy1	0.98	NA*	1.01	NA*	1.03	NA*	1.05	NA*	1.04	NA*					1.01	1.05	1.03	NA*	1.03	1.48*	1.03	1.09*						
	snp6	1.06	NA*	1.09*	NA*	1.08	1.05	1.07*	1.07*	1.1	1.11*					1.08*	NA*	1.07	1.07	1.07*	NA*	1.1	NA*			1.09	1.09	1.09	NA*
	mgs	1.09	NA*									1.1	NA*	1.1	1.1*														
nutlin	affy1	0.5	0.48	0.52	0.53	0.52	0.51	0.52	0.52	0.52	0.52					0.53	NA*	0.53	0.53	0.52	0.53*	0.52	0.52						
	snp6	0.52	0.51	0.53	0.53	0.53*	0.53*	0.49*	NA*	0.5	NA*					0.51	0.52*	0.52*	0.51	0.48*	0.48	0.49	0.51*						
	mgs	0.51	0.5									0.52	0.53	0.5	0.5										0.52	0.53	0.51	0.51	
selumetinib	affy1	0.68	0.74	0.82	0.8	0.87	0.84	0.94	0.93	0.93	NA					0.77	0.84	0.71	0.78	0.9	0.95	0.9	0.97						
	snp6	0.86	NA*	0.98*	0.96	0.98*	0.97	0.94	0.93	0.98*	0.98					0.98	NA*	0.98*	0.97	0.94	0.95	0.96	0.97*						
	mgs	0.8	0.79									0.96	0.96*	0.9	0.92										0.8	0.9	0.79	0.86	
tanespimycin	affy1	0.07	0.08	0.1	0.1	0.1	NA*	0.1	NA*	0.1	NA*					0.1	0.1	0.1	NA	0.1	0.11	0.11	0.11						
	snp6	0.11	NA*	0.11	NA*	0.11	NA*	0.11	NA*	0.11*	NA*					0.11*	NA*	0.11	NA*	0.11	NA*	0.11	NA*						
	mgs	0.11	NA*									0.11	NA*	0.11	NA*										0.11	0.11	0.11	NA*	

Tabla D.9: Modelos para drogas compartidas, P: Barretina, T: Basu, *MSE*

		NEA														GNEA															
		RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference					
		P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T						
atra	affy1	0.81	0.58	0.61	0.53	0.46	0.44	0.65	0.39	0.45	0.49					0.95	0.55	0.67	0.52	0.83	0.6	0.95	0.49								
	snp6	0.85	0.43	NA*	NA*	NA*	NA*	0.23	0.43	0.07	0.82	NA*				NA*	NA*	NA*	NA*	NA*	NA*	0.27	NA*			0.91	NA*	0.62	0.12		
	mgs	0.51	NA*									NA*	0.27	NA*	0.15																
bez235	affy1	0.78	0.37	0.43	0.47	0.5	0.34	0.7	0.45	0.82	0.45					0.46	0.53	0.9	0.27	0.82	0.21	0.99	0.33								
	snp6	0.61	NA*	0.77	NA*	0.45	0.26	0.94	NA*	0.66	0.35					0.59	0.18	NA*	NA*	0.81	0.17	0.45	0.31			NA*	NA*	0.66	0.17		
	mgs	0.32	NA*									0.42	NA*	0.35	0.04																
bi2536	affy1	0.89	NA*	0.67	0.35	0.5	0.66	0.44	NA*	0.48	0.04					0.75	NA*	0.94	NA*	0.61	NA*	0.91	NA*								
	snp6	0.86	NA*	0.53	NA*	0.63	NA*	0.78	0.33	0.78	0.26					0.56	-0.13	0.87	0.16	0.6	NA*	0.85	NA*			NA*	NA*	0.68	NA*		
	mgs	NA*	NA*									0.25	NA*	NA*	NA*																
bms536924	affy1	0.72	NA*	0.39	0.2	0.4	NA*	0.45	0.13	0.3	0.08					0.79	NA*	0.96	0.29	0.79	0.41	0.76	NA*								
	snp6	0.67	NA*	0.26	NA*	0.55	NA*	0.42	NA*	NA*	0.29					NA*	NA*	NA*	NA*	0.46	NA*	0.13	NA*								
	mgs	NA*	NA*									0.42	0.21	0.3	0.39										0.65	0.13	0.38	NA*			
bms754807	affy1	0.69	0.59	0.54	NA*	NA*	0.45	NA*	0.2	0.33	0.3					0.99	0.6	0.67	0.57	NA*	0.29	NA*	NA*								
	snp6	0.53	NA*	NA*	NA*	0.44	NA*	0.26	NA*	NA*	NA*					NA*	NA*	0.87	NA*	0.45	0.16	NA*	NA*			NA*	NA*	NA*	0.15		
	mgs	0.72	NA*									NA*	NA*	NA*	0.22																
bortezomib	affy1	0.94	0.42	0.57	0.48	0.6	0.33	0.84	0.33	0.51	0.5					0.83	NA*	0.99	0.32	0.92	0.31	0.92	NA*								
	snp6	0.6	0.28	NA*	NA*	0.47	0.22	NA*	NA*	NA*	0.15					NA*	NA*	0.48	0.35	0.42	NA*	0.23	NA*								
	mgs	0.52	NA*									0.31	NA*	NA*	NA*										0.72	NA*	0.55	0.08			
bx795	affy1	0.79	0.55	0.62	0.51	0.63	0.48	0.44	0.41	0.57	0.57					0.8	NA*	0.83	0.5	NA*	0.37	0.75	0.28								
	snp6	0.61	0.55	NA*	NA*	NA*	NA*	NA*	NA*	0.52	NA*					NA*	NA*	NA*	NA*	0.25	NA*	0.22	NA*				0.84	NA*	0.93	0.2	
	mgs	NA*	NA*									0.25	NA*	NA*	0.02																
chir99021	affy1	0.75	NA*	0.63	0.31	0.51	0.53	0.56	0.16	0.44	0.3					0.65	0.22	0.76	0.7	0.48	0.6	NA*	0.15								
	snp6	0.56	0.35	NA*	NA*	NA*	NA*	NA*	NA*	0.08	NA*	0.27				NA*	0.1	NA*	NA*	0.82	NA*	NA*	NA*				0.61	NA*	NA*	0.25	
	mgs	NA*	NA*									NA*	NA*	NA*	NA*																
etoposide	affy1	0.86	0.55	0.62	0.45	0.61	NA*	0.74	NA*	0.7	0.45					0.96	NA*	0.98	NA*	1	0.33	0.96	0.32								
	snp6	0.95	NA*	0.33	NA*	NA*	0.17	0.42	NA*	NA*	0.38					0.85	NA*	NA*	NA*	NA*	NA*	0.17	NA*				0.74	NA*	NA*	NA*	
	mgs	NA*	NA*									0.44	0.26	0.66	0.19																
gemcitabine	affy1	0.76	0.62	0.53	0.22	0.65	NA*	0.36	0.35	0.52	0.46					0.93	NA*	0.98	0.37	0.51	0.21	0.79	NA*								
	snp6	0.8	0.51	0.48	NA*	0.78	NA*	NA*	NA*	0.44	NA*					0.89	0.28	0.75	NA*	0.41	NA*	NA*	NA*				NA*	0.11	NA*	NA*	
	mgs	0.36	NA*									NA*	NA*	0.48	NA*																
gw843682x	affy1	0.86	NA*	0.8	0.46	0.61	NA*	0.46	NA*	0.64	NA*					0.84	NA*	0.93	0.4	0.95	NA*	0.96	NA*								
	snp6	0.8	0.12	NA*	NA*	0.48	0.21	NA*	NA*	0.85	0.1					0.52	NA*	0.86	NA*	0.93	0.17	NA*	NA*				0.48	NA*	0.46	NA*	
	mgs	NA*	NA*									0.45	NA*	0.36	NA*																
linsitinib	affy1	0.78	0.59	0.83	0.16	0.42	0.14	0.63	0.12	0.51	0.32					1	0.45	0.6	0.2	0.48	NA*	0.25	0.06								
	snp6	0.57	0.23	NA*	NA*	0.49	NA*	NA*	NA*	NA*	NA*					0.27	NA*	0.3	NA*	NA*	NA*	0.05	NA*					0.71	0.32	0.55	0.37
	mgs	0.54	0.35									0.64	NA*	0.21	0.28																
mk2206	affy1	0.93	0.66	0.81	0.46	0.79	0.41	0.91	NA*	0.57	0.25					0.83	0.58	1	0.47	0.94	0.96	0.81	NA*								
	snp6	0.9	NA*	NA*	0.45	NA*	NA*	0.73	NA*	0.65	NA*					NA*	NA*	NA*	NA*	0.87	0.8	0.44	NA*				0.85	0.34	NA*	NA*	
	mgs	NA*	NA*									NA*	-0.02	0.34	0.14																
navitoclax	affy1	0.92	0.44	0.39	0.31	0.38	0.24	0.64	0.26	0.5	0.38					0.75	0.43	0.96	0.54	0.63	0.35	0.85	0.31								
	snp6	0.77	0.48	0.54	0.19	NA*	NA*	0.28	0.11	NA*	NA*					NA*	NA*	0.37	NA*	0.11	0.12	0.29	0.04				0.48	NA*	0.39	NA*	
	mgs	0.48	NA*									0.19	NA*	0.35	NA*																

Tabla D.10: Modelos para drogas compartidas, P: Basu, T: Garnett, r_s (1/2)

		NEA														GNEA												
		RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference		
		P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	
nutlin	affy1	0.93	0.62	0.41	0.29	0.43	0.31	0.46	0.41	0.52	0.38					NA*	-0.09	0.74	0.44	0.6	0.49	0.71	0.24					
	snp6	0.86	0.59	0.37	0.31	0.61	0.46	NA*	0.33	NA*	NA*					0.87	0.4	0.77	0.67	0.82	0.44	0.43	NA*					
	mgs	0.64	0.27									0.45	0.27	0.54	0.27										0.98	0.4	0.46	0.33
nvpauy922	affy1	1	0.65	0.53	0.35	0.59	0.4	0.52	0.43	0.59	0.42					0.65	0.67	0.79	NA*	0.77	0.29	0.77	NA*					
	snp6	0.78	0.4	0.52	NA*	NA*	NA*	NA*	NA*	NA*	NA*					0.79	NA*	0.59	NA*	NA*	0.57	NA*	NA*					
	mgs	NA*	NA*									0.38	0.18	NA*	NA*										0.29	NA*	NA*	NA*
olaparib	affy1	0.64	0.61	0.62	0.46	0.44	0.44	0.43	0.36	0.58	0.41					0.63	0.43	0.47	0.4	0.82	0.52	0.61	NA*					
	snp6	0.87	0.44	NA*	NA*	NA*	NA*	NA*	0.24	0.52	NA*					NA*	0.4	0.73	0.54	0.29	NA*	NA*	0.38					
	mgs	0.43	NA*									0.51	0.19	NA*	NA*										NA*	0.1	NA*	NA*
pac1	affy1	0.78	0.51	0.41	0.41	0.55	0.44	0.31	0.33	0.4	NA*					NA*	NA*	0.82	NA*	0.38	0.26	0.35	NA*					
	snp6	NA*	0.24	0.27	0.12	0.53	NA*	NA*	NA*	0.25	NA*					NA*	NA*	0.36	NA*	NA*	0.3	0.18						
	mgs	NA*	NA*									NA*	NA*	0.38	NA*										0.33	NA*	NA*	NA*
parthenolide	affy1	0.84	0.37	0.8	NA*	0.73	NA*	0.85	NA*	0.76	0.44					0.93	0.64	0.76	NA*	0.95	0.56	0.77	0.47					
	snp6	0.86	0.29	NA*	0.22	0.97	0.24	0.65	NA*	NA*	NA*					NA*	NA*	NA*	NA*	0.96	NA*	0.91	NA*					
	mgs	0.67	0.25									NA*	NA*	0.56	NA*										0.83	0.29	0.75	NA*
qs11	affy1	NA*	NA*	0.29	NA*	0.45	0.12	0.2	NA*	NA*	NA*					0.47	0.47	0.47	NA*	0.25	NA*	0.17	NA*					
	snp6	NA*	NA*	NA*	NA*	NA*	NA*	0.48	NA*	0.43	0.5					0.97	NA*	NA*	0.26	NA*	NA*	0.46	NA*					
	mgs	NA*	NA*									NA*	0.34	0.82	NA*										0.33	0.38	0.33	0.22
selumetinib	affy1	0.88	0.8	0.91	0.62	0.81	0.64	0.7	0.55	0.15	NA*					0.93	0.78	0.97	0.78	0.8	0.51	0.66	0.51					
	snp6	NA*	NA*	0.57	NA*	0.57	NA*	0.77	NA*	0.43	NA*					NA*	0.11	0.91	0.29	0.96	NA*	1	NA*					
	mgs	0.82	0.53									0.56	NA*	0.41	NA*										0.75	NA*	0.61	NA*
sirolimus	affy1	0.65	NA*	0.38	0.39	0.49	NA*	0.53	NA*	0.38	NA*					0.69	NA*	0.68	NA*	0.6	NA*	0.62	0.1					
	snp6	0.48	NA*	0.36	NA*	NA*	NA*	0.66	NA*	0.34	NA*					0.57	NA*	0.7	NA*	NA*	NA*	0.82	NA*					
	mgs	NA*	NA*									NA*	NA*	0.36	0.1										0.66	NA*	0.25	NA*
sn38	affy1	0.88	0.7	0.65	0.54	0.59	0.62	0.63	0.47	0.51	0.34					0.83	0.64	0.99	NA*	0.56	0.43	0.93	0.32					
	snp6	0.81	0.71	0.46	NA*	0.47	NA*	0.59	NA*	0.95	0.22					0.91	NA*	0.97	NA*	NA*	0.14	NA*	NA*					
	mgs	NA*	NA*									NA*	NA*	0.34	NA*										NA*	0.09	0.87	NA*
tanespimycin	affy1	0.8	0.68	0.55	NA*	NA*	NA*	NA*	NA*	NA*	NA*					0.74	0.52	0.39	NA*	0.94	0.27	0.62	NA*					
	snp6	NA*	NA*	NA*	0.16	NA*	NA*	NA*	NA*	NA*	NA*					NA*	0.39	NA*	0.1	NA*	NA*	NA*	NA*					
	mgs	NA*	0.17									NA*	NA*	NA*	NA*										0.26	NA*	NA*	NA*
tipifarnibp1	affy1	0.58	0.42	0.57	0.12	0.55	0.26	0.31	NA*	0.37	0.24					0.7	0.31	0.78	0.28	0.5	0.27	0.3	NA*					
	snp6	0.65	0.34	NA*	NA*	NA*	NA*	0.43	NA*	NA*	NA*					NA*	NA*	NA*	NA*	0.41	NA*	0.06	NA*					
	mgs	0.53	NA*									NA*	NA*	NA*	NA*										0.36	NA*	0.34	NA*
tozaseribt	affy1	0.56	0.4	0.39	0.36	0.41	0.35	0.42	0.19	0.4	NA*					0.79	NA*	1	0.46	0.42	0.54	0.16	NA*					
	snp6	0.5	0.09	NA*	NA*	NA*	NA*	NA*	NA*	0.44	NA*					NA*	NA*	0.63	NA*	NA*	NA*	0.2	NA*					
	mgs	NA*	NA*									NA*	NA*	NA*	NA*										0.78	NA*	0.52	NA*
tw37	affy1	NA*	0.47	0.4	0.21	0.43	0.32	0.65	0.24	0.77	0.55					0.88	0.57	1	0.24	NA*	0.28	0.66	0.18					
	snp6	0.59	0.41	0.39	0.2	0.77	NA*	0.85	NA*	0.54	NA*					NA*	NA*	1	0.26	0.36	NA*	0.77	0.5					
	mgs	0.72	NA*									NA*	NA*	NA*	NA*										0.81	0.44	0.79	NA*
vorinostat	affy1	0.87	0.66	0.67	0.59	0.71	0.62	0.71	0.48	0.6	0.52					0.82	0.5	0.96	NA*	0.84	0.56	0.87	0.52					
	snp6	0.78	NA*	0.32	NA*	NA*	NA*	0.48	NA*	0.38	NA*					0.5	NA*	NA*	0.63	NA*	NA*	NA*	NA*					
	mgs	NA*	0.1									0.31	NA*	0.27	0.17									0.38	-0.03	0.51	0.35	

Tabla D.11: Modelos para drogas compartidas, P: Basu, T: Garnett, r_s (2/2)

		NEA														GNEA													
		RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference			
		P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T		
atra	affy1	0.54	0.39	0.42	0.38	0.4	0.36	0.37	0.33	0.4	0.33					0.56	0.44	0.59	0.35	0.76	0.4	0.48	0.45						
	snp6	0.35	0.28	0.44	NA*	0.4	NA*	0.1	0.14	0.26	NA*					0.56	NA*	0.67	NA*	0.48	NA*	NA*	NA*						
	mgs	NA*	NA*									0.35	NA*	0.19	NA*										0.51	NA*	0.22	0.24	
bez235	affy1	0.51	0.4	0.38	0.14	0.37	0.2	0.26	0.23	0.33	0.25					0.6	0.25	0.45	0.14	0.26	0.2	0.33							
	snp6	NA*	0.29	NA*	NA*	0.27	NA*	NA*	NA*	0.17	0.18					0.29	NA*	0.43	NA*	0.29	0.12	0.19	NA*			NA*	NA*	0.15	0.06
	mgs	NA*	0.18								NA*	NA*	0.05	NA*															
bi2536	affy1	NA*	NA*	0.54	0.1	0.5	NA*	NA*	NA*	0.19	NA*					NA*	NA*	NA*	NA*	NA*	NA*	NA*	NA*						
	snp6	0.34	NA*	0.35	0.14	0.33	0.28	0.4	NA*	0.25	0.24					0.1	NA*	0.3	0.38	0.61	0.18	NA*	NA*						
	mgs	0.33	NA*								NA*	0.07	NA*	NA*										NA*	NA*	NA*	0.17		
bms536924	affy1	NA*	0.45	0.48	0.31	0.48	NA*	0.31	0.3	0.49	0.21					NA*	NA*	0.83	0.42	0.94	0.34	0.46	0.37						
	snp6	0.23	0.25	0.29	NA*	0.25	0.19	NA*	0.1	0.47	NA*					NA*	NA*	NA*	NA*	0.81	NA*	NA*	0.1						
	mgs	NA*	NA*									0.36	0.31	0.47	NA*										0.48	0.16	NA*	0.15	
bms754807	affy1	0.78	0.26	0.23	0.14	0.42	NA*	0.28	NA*	0.32	0.17					0.67	0.44	0.73	0.2	0.73	NA*	0.32	NA*						
	snp6	0.56	0.21	0.13	NA*	NA*	0.14	NA*	-0.05	0.2	NA*					0.07	NA*	NA*	NA*	0.14	NA*	NA*	NA*						
	mgs	NA*	0.17									0.22	NA*	0.22	NA*										NA*	NA*	0.19	NA*	
bortezomib	affy1	0.69	NA*	0.53	0.16	0.53	0.32	0.4	0.39	0.56	0.31					0.66	0.2	0.63	0.72	0.52	0.41	0.54	0.38						
	snp6	0.5	0.11	NA*	NA*	0.38	0.16	NA*	NA*	0.25	NA*					NA*	NA*	0.74	0.19	0.3	NA*	NA*	NA*						
	mgs	0.48	NA*									0.39	NA*	NA*	NA*										0.6	NA*	-0.04	NA*	
bx795	affy1	0.76	0.45	0.55	0.35	0.51	0.44	0.42	0.26	0.57	0.23					0.66	0.38	0.69	0.37	0.44	NA*	0.54	0.36						
	snp6	0.65	0.33	NA*	NA*	0.55	NA*	0.51	NA*	0.33	NA*					0.69	NA*	0.67	NA*	0.21	0.02	NA*	-0.1						
	mgs	NA*	NA*								NA*	NA*	0.14	NA*											0.75	0.35	0.39	0.34	
chir99021	affy1	0.57	0.34	0.32	0.37	0.44	0.3	0.3	NA*	0.33	0.26					0.47	0.38	0.61	0.4	0.63	NA*	0.37	NA*						
	snp6	0.48	0.2	0.29	NA*	NA*	NA*	0.16	NA*	0.33	NA*					0.68	NA*	NA*	NA*	0.47	NA*	NA*	NA*						
	mgs	0.41	NA*								NA*	NA*	0.33	NA*											0.26	0.16	0.31	NA*	
etoposide	affy1	0.59	0.36	0.23	NA*	NA*	0.23	NA*	NA*	0.42	0.18					0.24	0.36	NA*	0.31	0.3	0.18	0.27	NA*						
	snp6	NA*	0.3	NA*	NA*	0.18	NA*	NA*	NA*	0.28	NA*					NA*	NA*	NA*	NA*	NA*	NA*	NA*	NA*						
	mgs	NA*	NA*									0.23	0.11	0.36	0.24										NA*	0.24	0.41	NA*	
gemcitabine	affy1	0.63	0.32	0.12	0.12	NA*	NA*	0.28	NA*	0.5	0.18					0.48	0.28	0.43	0.34	0.23	0.25	NA*	0.26						
	snp6	0.45	0.26	NA*	0.18	0.27	NA*	0.18	NA*	NA*	NA*					0.22	0.15	NA*	0.07	0.26	NA*	0.25	NA*						
	mgs	NA*	0.02								NA*	NA*	NA*	0.26											0.21	NA*	0.52	NA*	
gw843682x	affy1	NA*	0.32	0.47	NA*	NA*	NA*	NA*	NA*	NA*	NA*					NA*	NA*	-0.04	0.03	NA*	NA*	NA*	NA*						
	snp6	0.29	NA*	0.29	NA*	0.28	0.15	NA*	NA*	0.13	NA*					NA*	NA*	NA*	NA*	0.81	0.23	NA*	NA*						
	mgs	0.34	NA*								NA*	NA*	NA*	NA*											NA*	NA*	NA*	0.14	
linsitinib	affy1	0.86	0.32	0.15	NA*	0.28	0.16	0.1	0.23	0.36	NA*					0.54	NA*	0.46	NA*	0.82	0.14	0.13	0.12						
	snp6	0.49	NA*	NA*	NA*	NA*	NA*	NA*	NA*	NA*	NA*					NA*	0.08	0.68	NA*	NA*	NA*	NA*	0.1						
	mgs	0.4	0.18								NA*	0.19	0.24	0.15										0.47	0.16	0.28	0.19		
mk2206	affy1	0.75	0.15	0.27	0.19	0.36	0.21	0.38	0.22	0.15	NA*					0.6	0.24	0.51	0.3	0.86	0.29	NA*	0.15						
	snp6	0.4	NA*	0.27	NA*	0.2	NA*	NA*	0.1	NA*	NA*					0.36	NA*	0.69	NA*	0.78	0.24	NA*	0.13						
	mgs	NA*	NA*									0.19	NA*	0.25	0.03										0.61	NA*	0.39	NA*	
navitoclax	affy1	0.48	0.5	0.35	0.32	0.33	0.21	0.43	0.34	0.35	0.29					0.55	0.38	0.69	0.49	0.56	0.37	0.41	0.36						
	snp6	0.69	0.27	0.43	0.25	0.17	NA*	0.3	NA*	0.21	NA*					0.65	NA*	0.25	NA*	0.46	NA*	0.08	0.1						
	mgs	NA*	NA*								NA*	NA*	NA*	0.17											0.45	NA*	0.3	NA*	

Tabla D.12: Modelos para drogas compartidas, P: Garnett, T: Basu, r_s (1/2)

		NEA														GNEA											
		RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference	
		P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T
nutlin	affy1	0.69	0.56	0.38	0.18	0.32	0.21	0.32	0.17	0.39	NA*					0.16	NA*	0.57	0.28	0.89	NA*	0.32	0.14				
	snp6	0.7	0.41	0.3	0.21	0.37	0.24	0.5	NA*	NA*	NA*					0.32	0.36	0.7	0.24	0.55	0.25	NA*	NA*				
	mgs	0.43	0.45									0.43	0.11	0.38	0.37										0.6	0.39	0.5
nvpauy922	affy1	0.63	0.39	0.35	NA*	0.43	0.27	0.5	NA*	0.45	0.34					0.44	NA*	0.5	0.33	0.47	0.33	0.44	0.28				
	snp6	0.39	0.2	0.1	NA*	0.31	NA*	NA*	NA*	0.2	NA*					NA*	NA*	0.09	NA*	0.56	NA*	NA*	NA*				
	mgs	NA*	NA*									0.1	NA*	NA*	NA*										NA*	0.02	NA*
olaparib	affy1	0.61	0.43	0.48	0.4	0.6	0.26	0.49	0.29	0.45	0.3					0.64	0.38	0.73	0.28	0.74	0.36	0.26	0.22				
	snp6	0.63	0.35	0.37	NA*	NA*	NA*	0.31	NA*	NA*	0.1					0.79	NA*	0.56	0.18	0.62	NA*	0.51	NA*				
	mgs	0.24	NA*									0.15	0.21	0.5	NA*										0.09	NA*	NA*
pac1	affy1	0.85	0.37	0.45	0.21	0.45	0.27	0.36	0.18	0.24	NA*					0.57	NA*	0.62	0.12	0.59	NA*	0.33	0.25				
	snp6	0.43	NA*	0.33	NA*	NA*	0.21	0.22	NA*	0.39	0.17					0.49	NA*	NA*	0.13	0.79	NA*	0.79	NA*				
	mgs	0.23	NA*									NA*	NA*	0.41	0.09										0.27	NA*	0.45
parthenolide	affy1	0.61	0.2	NA*	0.15	NA*	0.2	NA*	0.43	0.4	NA*					0.77	0.26	0.79	0.3	0.85	0.17	0.63	NA*				
	snp6	0.34	NA*	0.3	NA*	0.36	0.22	0.61	NA*	NA*	NA*					0.96	NA*	NA*	NA*	0.48	NA*	0.38	NA*				
	mgs	0.49	0.14									0.54	NA*	0.6	0.26										0.47	0.2	NA*
qs11	affy1	0.56	NA*	0.34	0.11	0.21	NA*	0.33	NA*	0.3	NA*					0.53	NA*	0.41	0.13	NA*	NA*	NA*	NA*				
	snp6	NA*	NA*	NA*	NA*	NA*	NA*	NA*	NA*	0.17	0.39	NA*				0.33	NA*	0.41	NA*	0.38	NA*	NA*	NA*				
	mgs	NA*	NA*									0.39	NA*	NA*	NA*										0.48	0.04	0.57
selumetinib	affy1	0.62	0.45	0.47	0.39	0.54	0.36	0.52	0.35	0.34	-0.01					0.81	0.52	0.81	0.43	0.63	0.32	0.42	NA*				
	snp6	NA*	NA*	0.24	NA*	0.38	0.18	NA*	0.07	NA*	NA*					0.11	NA*	0.33	NA*	NA*	NA*	NA*	0.42				
	mgs	0.56	0.38									NA*	NA*	NA*	0.1										NA*	0.2	NA*
sirolimus	affy1	NA*	0.35	0.51	0.27	NA*	NA*	0.29	NA*	NA*	NA*					NA*	0.28	NA*	0.29	NA*	NA*	0.02	NA*				
	snp6	NA*	0.18	0.23	NA*	NA*	NA*	NA*	NA*	0.27	NA*					NA*	NA*	NA*	0.2	NA*	NA*	NA*	NA*				
	mgs	0.25	NA*									0.2	NA*	0.22	0.22										NA*	-0.01	NA*
sn38	affy1	0.72	0.6	0.42	0.3	0.46	0.27	0.49	0.16	0.31	0.12					0.62	0.3	0.53	0.53	0.56	0.01	0.08	0.32				
	snp6	0.48	0.19	NA*	0.18	NA*	NA*	NA*	NA*	0.29	0.18					NA*	0.32	NA*	NA*	0.19	NA*	NA*	NA*				
	mgs	NA*	NA*									NA*	NA*	NA*	NA*										0.14	NA*	NA*
tanespimycin	affy1	0.81	0.5	0.46	0.26	0.53	NA*	0.5	NA*	0.42	NA*					0.72	0.33	0.75	0.01	0.73	0.35	0.53	0.22				
	snp6	0.18	NA*	0.32	NA*	0.19	NA*	NA*	NA*	0.14	NA*					0.47	NA*	0.19	NA*	NA*	NA*	NA*	NA*				
	mgs	0.2	NA*									0.24	NA*	0.23	NA*										NA*	0.09	NA*
tipifarnibp1	affy1	0.55	NA*	0.21	0.22	0.32	0.29	NA*	0.2	0.36	0.25					0.44	0.27	0.52	0.31	0.35	0.32	0.31	NA*				
	snp6	0.47	NA*	NA*	NA*	NA*	NA*	NA*	NA*	NA*	NA*					NA*	NA*	NA*	NA*	0.13	0.13	NA*	NA*				
	mgs	NA*	0.11									NA*	NA*	NA*	NA*										NA*	NA*	NA*
tozasertib	affy1	0.82	NA*	0.49	0.24	0.4	0.23	0.34	0.21	NA*	NA*					0.83	NA*	0.75	0.31	0.95	0.01	NA*	NA*				
	snp6	0.22	0.15	NA*	NA*	NA*	NA*	0.68	NA*	0.51	NA*					NA*	NA*	NA*	NA*	0.78	NA*	NA*	0.17				
	mgs	NA*	NA*									0.29	NA*	NA*	NA*										NA*	NA*	NA*
tw37	affy1	0.71	NA*	0.47	0.2	0.35	0.3	0.41	0.2	0.46	NA*					0.62	0.38	0.5	NA*	0.62	NA*	0.09	NA*				
	snp6	0.57	0.2	0.28	0.16	NA*	0.2	NA*	NA*	NA*	NA*					0.41	NA*	0.44	0.38	0.69	NA*	0.99	NA*				
	mgs	NA*	NA*									NA*	NA*	0.23	NA*										0.38	NA*	0.42
vorinostat	affy1	0.65	0.53	0.44	0.41	0.45	0.43	0.32	0.33	0.35	0.3					0.29	0.43	0.4	0.27	0.7	0.38	0.32	0.23				
	snp6	NA*	0.42	0.31	NA*	NA*	NA*	0.38	NA*	0.23	NA*					-0.02	0.16	NA*	NA*	0.12	NA*	NA*	NA*				
	mgs	0.13	NA*									NA*	0.16	0.22	0.2										-0.14	0.25	0.53

Tabla D.13: Modelos para drogas compartidas, P: Garnett, T: Basu, r_s (2/2)

Primary (P): Basu, Transferred (T): Garnett
 p -value

		NEA														GNEA																								
		RAW		top.200				top.400				signif		signif.filt		total		reference		top.200				top.400				signif		signif.filt		total		reference						
		P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T									
atra	affy1	-88.22	-47.51	-56.95	-40.95	-35.81	-33.36	-65.79	-41.34	-41.92	-41.89									-167.98	-41.57	-69.48	-44.88	-106.09	-50.82	-168.65	-51.59													
	snp6	-112.48	-12.06	NA*	NA*	NA*	NA*	-6.68	-10.86	-0.72	-53.11	NA*								NA*	NA*	-16.14	NA*																	
	mgs	-17.68	NA*									NA*	-3.95	NA*	-2.39														-141.35	NA*	-51.95	-0.77								
bez235	affy1	-57.09	-15.68	-17.04	-16.41	-20.55	-9.62	-42.76	-15.97	-73.48	-14.82									-25.64	-29.35	-96.3	-9.54	-73.64	-8.84	-220.76	-9.41													
	snp6	-32.28	NA*	-55.13	NA*	-15.6	-6.96	-106.36	NA*	-23.78	-7.26									-30.6	-6.56	NA*	NA*	-45.83	-2.13	-16	-6.94				NA*	NA*	-36.11	-7.23						
	mgs	-7.63	NA*										-13.45	NA*	-9.9	-0.07																								
bi2536	affy1	-116.99	NA*	-41.84	-12.04	-25.87	-37.5	-21.33	NA*	-23.43	-3.37									-60.46	NA*	-137.63	NA*	-43.69	NA*	-119.32	NA*													
	snp6	-100.16	NA*	-21.91	NA*	-34.53	NA*	-57.53	-7.85	-60.42	-3.48									-29.69	-0.5	-108.47	-1.93	-35.23	NA*	-81.61	NA*													
	mgs	NA*	NA*										-5.83	NA*	NA*	NA*																								
bms536924	affy1	-93.46	NA*	-23.5	-5.3	-22.91	NA*	-31.64	-3.66	-14.74	-5.37									-118.48	NA*	-329.75	-18.4	-146.57	-26.55	-131.59	NA*													
	snp6	-72.81	NA*	-8.17	NA*	-48.87	NA*	-31.18	NA*	NA*	-5.37									NA*	NA*	NA*	NA*	-38.93	NA*	-8.9	NA*													
	mgs	NA*	NA*										-27.41	-7.35	-14.47	-17.37																								
bms754807	affy1	-52.75	-31.62	-28.04	NA*	NA*	-17.93	NA*	-4.06	-12.82	-4.84									-268.13	-36.72	-47.41	-29.05	NA*	-14.39	NA*	NA*													
	snp6	-24.23	NA*	NA*	NA*	-17.27	NA*	-7.26	NA*	NA*	NA*									NA*	NA*	-96.42	NA*	-17.36	-3.34	NA*	NA*													
	mgs	-54.77	NA*											NA*	NA*	NA*	-3.59																							
bortezomib	affy1	-142.31	-13.72	-22.51	-18.86	-25.05	-8.45	-75.59	-8.36	-19.26	-15.6								-76.01	NA*	-260.78	-3.29	-114.32	-7.17	-113.24	NA*														
	snp6	-31.11	-11.28	NA*	NA*	-15.37	-5.19	NA*	NA*	NA*	-2.11								NA*	NA*	-25.65	-9.68	-18.25	NA*	-10.32	NA*														
	mgs	-25.76	NA*											-9.18	NA*	NA*	NA*																							
bx795	affy1	-66.44	-29.44	-38.78	-26.47	-39.2	-20.27	-21.13	-19.2	-28.52	-25.57								-76.96	NA*	-81.66	-22.73	NA*	-15.99	-70.16	-14.35														
	snp6	-32.1	-26.64	NA*	NA*	NA*	NA*	NA*	NA*	-13.47	NA*								NA*	NA*	NA*	NA*	-9.06	NA*	-7.45	NA*														
	mgs	NA*	NA*										-7.96	NA*	NA*	-0.01																								
chir99021	affy1	-63.21	NA*	-36.05	-10.52	-19.85	-20.11	-27.51	-4.38	-19.43	-6.53								-47.32	-5.39	-60.65	-46.1	-32.1	-38.7	NA*	-5.75														
	snp6	-24.47	-7.64	NA*	NA*	NA*	NA*	NA*	-1.76	NA*	-6.85								NA*	-5.58	NA*	NA*	-63.62	NA*	NA*	NA*														
	mgs	NA*	NA*											NA*	NA*	NA*	NA*																							
etoposide	affy1	-99.56	-34.83	-43.06	-21.75	-42.02	NA*	-59.3	NA*	-55.25	-27.12								-186.87	NA*	-260.05	NA*	-518.28	-13.86	-185.23	-10.16														
	snp6	-167	NA*	-6.59	NA*	NA*	-2.94	-14.64	NA*	NA*	-8.15								-108.99	NA*	NA*	NA*	NA*	-11.06	NA*															
	mgs	NA*	NA*										-12.75	-4.49	-45.24	-5.02																								
gemcitabine	affy1	-58.12	-37.12	-27.37	-4.2	-40.71	NA*	-17.48	-12.21	-23.74	-21.3								-128.93	NA*	-198.65	-15.02	-35.13	-3.22	-78.88	NA*														
	snp6	-73.53	-21.37	-16.22	NA*	-53.24	NA*	NA*	NA*	-12.95	NA*								-105.64	-5.98	-63.35	NA*	-9.55	NA*	NA*	NA*														
	mgs	-7.95	NA*											NA*	NA*	-23.35	NA*																							
gw843682x	affy1	-135.8	NA*	-106.75	-25.69	-54.24	NA*	-38.7	NA*	-61.64	NA*								-125.96	NA*	-196.71	-0.9	-235.15	NA*	-254.48	NA*														
	snp6	-105.64	-1.94	NA*	NA*	-24.27	-5.87	NA*	NA*	-90.25	-1.06								-33.7	NA*	-133.99	NA*	-186.81	-6.61	NA*	NA*														
	mgs	NA*	NA*										-20.24	NA*	-12.48	NA*																								
linsitinib	affy1	-69.37	-40.87	-89.05	-5.32	-17.64	-5.68	-40.13	-6.73	-29.84	-12.79								-505.08	-22.32	-48.71	-7.05	-35.4	NA*	-21.92	-2.34														
	snp6	-39.58	-3.87	NA*	NA*	-17.9	NA*	NA*	NA*	NA*	NA*								-13.48	NA*	-13.86	NA*	NA*	NA*	-8.63	NA*														
	mgs	-25.22	-9.75											-38.88	NA*	-6.79	-10.03																							
mk2206	affy1	-65.62	-32.25	-27.75	-15.13	-38.88	-14.9	-52.67	NA*	-20.99	-6.39								-42.11	-23.04	-202.27	-16.71	-76.59	-86.76	-45.62	NA*														
	snp6	-62.4	NA*	NA*	-11.81	NA*	NA*	-27.01	NA*	-14.51	NA*								NA*	NA*	NA*	NA*	-37.02	-34.06	-12.51	NA*														
	mgs	NA*	NA*											NA*	-0.08	-6.59	-2.92																							
navitoclax	affy1	-260.76	-41.43	-25.89	-16.05	-27.42	-10.21	-64.65	-14.68	-40.53	-25.65								-113.77	-35.05	-317.25	-53.28	-87.57	-34.41	-176.35	-26.87														
	snp6	-120.16	-48.17	-51.33	-8.2	NA*	NA*	-11.92	-5.13	NA*	NA*								NA*	NA*	-22.7	NA*	-10.18	-5.68	-18.06	-0.57														
	mgs	-28.37	NA*											-8.1	NA*	-21.33	NA*																							

Tabla D.14: Modelos para drogas compartidas, P: Basu, T: Garnett, p -value (1/2)

Primary (P): Basu, Transferred (T): Garnett
 p -value

		NEA												GNEA															
		RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference			
		P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T				
nutlin	affy1	-147.28	-26.54	-16.61	-8.14	-16.5	-9.11	-13.79	-12.3	-21.19	-13.65					NA*	-0.39	-49.36	-13.84	-44.04	-20.46	-62.44	-6.07						
	snp6	-80.84	-24.44	-11.52	-7.87	-29.3	-16.27	NA*	-9.29	NA*	NA*					-109.17	-12.18	-64.24	-30.86	-75.38	-12.53	-23.33	NA*						
	mgs	-38.38	-6.95									-18.07	-7.25	-29.45	-8.08									-195.2	-13.71	-23.32	-11.48		
nvpauy922	affy1	-443.37	-41.38	-21.96	-9.94	-27.82	-14.59	-23.32	-16.67	-36.75	-22.37							-42.88	-34.44	-84.59	NA*	-66.84	-12.48	-83.42	NA*				
	snp6	-65.11	-11.49	-19.76	NA*	NA*	NA*	NA*	NA*	NA*	NA*							-76.06	NA*	-51.5	NA*	NA*	-34.11	NA*	NA*				
	mgs	NA*	NA*									-15.49	-1.38	NA*	NA*										-9.87	NA*	NA*	NA*	
olaparib	affy1	-54.03	-41.25	-35.36	-19.93	-16.05	-15.36	-12.71	-13.92	-36.39	-21.68							-63.24	-19.47	-34.32	-16.78	-104.74	-25.27	-65.38	NA*				
	snp6	-97.38	-13.91	NA*	NA*	NA*	NA*	NA*	-5.75	-22.51	NA*							NA*	-17.36	-71.09	-24.1	-16.47	NA*	NA*	-13.23				
	mgs	-26.24	NA*									-25.11	-3.45	NA*	NA*										NA*	-2.06	NA*	NA*	
pac1	affy1	-127.48	-40.19	-23.87	-24.66	-43.04	-24.53	-13.88	-14.27	-20.51	NA*							NA*	NA*	-137.77	NA*	-27.01	-9.74	-26.55	NA*				
	snp6	NA*	-7.7	-14.44	-3.44	-43.67	NA*	NA*	NA*	-9.45	NA*							NA*	NA*	-20.88	NA*	NA*	NA*	-22.61	-12.03		-15.86		
	mgs	NA*	NA*																										
parthenolide	affy1	-61.97	-9.98	-48.1	NA*	-37.81	NA*	-64.08	NA*	-45.25	-12.13							-94.34	-28.27	-40	NA*	-113.83	-22.72	-48.71	-16.14				
	snp6	-63.72	-4.17	NA*	-2.95	-143.29	-4.39	-18.76	NA*	NA*	NA*							NA*	NA*	NA*	NA*	-125.36	NA*	-61.12	NA*				
	mgs	-25.55	-6.3																						-60.68	-5.58	-35.74	NA*	
qs11	affy1	NA*	NA*	-13.05	NA*	-20.89	-5.16	-11.03	NA*	NA*	NA*							-44.47	-23.18	-29.26	NA*	-27.97	NA*	-24.7	NA*				
	snp6	NA*	NA*	NA*	NA*	NA*	NA*	-19.25	NA*	-13.56	-12.19							-190.95	NA*	NA*	-5.07	NA*	NA*	-27.71	NA*				
	mgs	NA*	NA*																						-9.92	-15.25	-10.56	-5.43	
selumetinib	affy1	-87.15	-59.31	-90.53	-27.51	-57.37	-25.16	-38.6	-21.36	-6.7	NA*							-119.68	-52.74	-172.62	-59.95	-57.17	-15.87	-39.93	-20.14				
	snp6	NA*	NA*	-26.04	NA*	-24.75	NA*	-43.69	NA*	-12.63	NA*							NA*	-0.14	-97.57	-4.78	-136.79	NA*	-227.83	NA*				
	mgs	-56.79	-17.73									-18.57	NA*	-9.56	NA*										-45.7	NA*	-24.63	NA*	
sirolimus	affy1	-79.4	NA*	-36.67	-34.51	-49.55	NA*	-59.29	NA*	-34.01	NA*							-92.22	NA*	-82.71	NA*	-71.06	NA*	-82.03	-0.22				
	snp6	-45.89	NA*	-22.86	NA*	NA*	NA*	-68.72	NA*	-12.66	NA*							69.4	NA*	-104.54	NA*	NA*	NA*	-120.12	NA*				
	mgs	NA*	NA*																						-76.03	NA*	-18.62	NA*	
sn38	affy1	-100.61	-52.94	-44.32	-31.23	-36	-36.58	-34.37	-20.95	-28.82	-15.5							-75.44	-39.66	-269.01	NA*	-32.21	-21.4	-130.5	-8.75				
	snp6	-72.4	-47.54	-15.15	NA*	-17.43	NA*	-26.27	NA*	-107.37	-4.58							-121.15	NA*	-188.94	NA*	-3.96	NA*	NA*					
	mgs	NA*	NA*																						NA*	-0.44	-95.65	NA*	
tanespimycin	affy1	-70.79	-38.8	-28.76	NA*	NA*	NA*	NA*	NA*	NA*	-2.92							-66.15	-25.41	-12.43	NA*	-168.16	-10.94	-36.4	NA*				
	snp6	NA*	NA*	-3.18	NA*	NA*	NA*	NA*	NA*	-1.87							NA*	-10.83	NA*	-0.82	NA*	NA*	NA*			-10.88	NA*	NA*	NA*
	mgs	NA*	-2.91																										
tipifarnibp1	affy1	-77.79	-33.05	-68.26	-3.3	-63.18	-12.81	-29.27	NA*	-33.22	-13.33							-119.56	-17.8	-155.43	-17.18	-78.5	-19.03	-55.21	NA*				
	snp6	-94.75	-22.75	NA*	NA*	NA*	NA*	-38	NA*	NA*	NA*							NA*	NA*	NA*	NA*	-55.14	NA*	-21.23	NA*				
	mgs	-55.53	NA*																						-29.55	NA*	-20.45	NA*	
tozasertib	affy1	-56.34	-24.5	-26.1	-19.46	-26.59	-20.97	-27.82	-6.74	-25.72	NA*							-134.01	NA*	-Inf	-33.57	-35.99	-48.31	-19.29	NA*				
	snp6	-37.91	-2.1	NA*	NA*	NA*	NA*	NA*	NA*	-16.13	NA*							NA*	NA*	-78	NA*	NA*	-13.12	NA*					
	mgs	NA*	NA*																						-116.54	NA*	-39.28	NA*	
tw37	affy1	NA*	-15.37	-13.14	-6.06	-13.95	-8.72	-29.05	-5.43	-46.22	-19							-92.7	-25.79	-331.18	-5.43	NA*	-7.26	-38.48	-3.33				
	snp6	-25.41	-14.11	-12.01	-3.8	-48.65	NA*	-65.76	NA*	-19.11	NA*							NA*	NA*	-286.1	-4.89	-17.48	NA*	-41.96	-17.37				
	mgs	-33.21	NA*																						-60.36	-11.9	-53.59	NA*	
vorinostat	affy1	-164.96	-72.35	-80.9	-58.25	-85.42	-63.09	-84.36	-43.34	-59.66	-48.89							-135.34	-44.18	-308.65	NA*	-140.93	-58.55	-156.68	-40.8				
	snp6	-116.28	NA*	-12.77	NA*	NA*	NA*	-31.45	NA*	-15.08	NA*							-30.95	NA*	NA*	NA*	-56.21	NA*	NA*	NA*				
	mgs	NA*	-1.92									-10.94	NA*	-8.9	-4.31										-18.92	-1.24	-35.92	-14.65	

Tabla D.15: Modelos para drogas compartidas, P: Basu, T: Garnett, p -value (2/2)

Primary (P): Garnett, Transferred (T): Basu
p-value

RAW		NEA												GNEA												
		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference		
		P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	
atra	affy1	-73.09	-45.25	-50.73	-43.93	-45.17	-39.22	-39.41	-30.12	-40.93	-34.17					-81.74	-54.44	-87.84	-40.3	-163.29	-43.96	-53.32	-50.75			
	snp6	-28.56	-17.44	-50.53	NA*	-40.87	NA*	-5.84	-4.06	-12.62	NA*					-67.82	NA*	-106.9	NA*	-56.62	NA*	NA*	NA*			
	mgs	NA*	NA*									-25.55	NA*	-8.41	NA*									-55.73	NA*	-10.73
bez235	affy1	-143.04	-33.84	-39.93	-10.07	-32.9	-12.88	-28.23	-25.14	-37.87	-20.71					-232.92	-11.01	-145.34	-17.49	-82.86	-31.62	-47.46	-47.86			
	snp6	NA*	-17.76	NA*	NA*	-16.73	NA*	NA*	NA*	-9.96	-8.61					-108.55	NA*	-102.62	NA*	-52.94	-4.69	-47.23	NA*			
	mgs	NA*	-13.85									NA*	NA*	-8.11	NA*									NA*	NA*	-18.14
bi2536	affy1	NA*	NA*	-37.69	-6.27	-50	NA*	NA*	NA*	-7.37	NA*					NA*	NA*	NA*	NA*	NA*	NA*	NA*	NA*			
	snp6	-19.91	NA*	-23.72	-2.25	-10.8	-9.66	-30.85	NA*	-17.68	-3.08					-14.72	NA*	-13.53	-18.02	-85.62	-6.11	NA*	NA*			
	mgs	-36.93	NA*									NA*	-2.37	NA*	NA*									NA*	NA*	-5.88
bms536924	affy1	NA*	-20.77	-27.71	-12.35	-22.39	NA*	-8.96	-3.76	-30.52	-6.17					NA*	NA*	-139.36	-29.35	-255.84	-18.21	-35.05	-18.62			
	snp6	-9.3	-6.11	-8.46	NA*	-7.77	-5.93	NA*	-4.57	-23.67	NA*					NA*	NA*	NA*	NA*	-150.9	NA*	NA*	-3.19			
	mgs	NA*	NA*									-16.72	-10.49	-34.27	NA*									-27.78	-6.79	NA*
bms754807	affy1	-295.49	-30.75	-30.42	-7.69	-61.4	NA*	-30.19	NA*	-36.11	-13.8					-226.16	-66.77	-258.99	-14.78	-265.44	NA*	-58.46	NA*			
	snp6	-117.56	-16.61	-9.43	NA*	NA*	-8.51	NA*	-0.58	-7.4	NA*					-15.53	NA*	NA*	NA*	-19.44	NA*	NA*	NA*			
	mgs	NA*	-8.42									-17.18	NA*	-17.71	NA*									NA*	NA*	-13.68
bortezomib	affy1	-65.4	NA*	-29.94	-4.18	-29.89	-12.44	-16.48	-15.23	-35.25	-4.94					-61.51	-8.31	-55.6	-57.43	-42.32	-14.67	-47.13	-8.35			
	snp6	-42.07	-2.39	NA*	NA*	-19.96	-3.59	NA*	NA*	-6.71	NA*					NA*	NA*	-101.41	-4.73	-27.87	NA*	NA*	NA*			
	mgs	-31	NA*									-20.49	NA*	NA*	NA*									-50.55	NA*	-12.89
bx795	affy1	-185.73	-45.29	-81.84	-28.75	-60.78	-45.34	-39.38	-20.2	-90.35	-17.59					-138.16	-37.18	-151.84	-25.97	-69.48	NA*	-103.5	-29.37			
	snp6	-121.08	-27.77	NA*	NA*	-78.41	NA*	-57.52	NA*	-14.76	NA*					-139.86	NA*	-156.21	NA*	-18.06	-1.32	NA*	-0.17			
	mgs	NA*	NA*									NA*	NA*	-7.43	NA*									-179.35	-32.71	-34.53
chir99021	affy1	-106.4	-28.94	-31.55	-35.18	-52.76	-24.91	-23.57	NA*	-32.29	-19.12					-64.84	-39.45	-115.65	-43.5	-123.55	NA*	-49.6	NA*			
	snp6	-70.36	-14.66	-22.69	NA*	NA*	NA*	-7.34	NA*	-22.83	NA*					-146.12	NA*	NA*	NA*	-68.1	NA*	NA*	NA*			
	mgs	-42.08	NA*									NA*	NA*	-29.07	NA*									-14.56	-5.47	-24.18
etoposide	affy1	-133.82	-24.78	-16.94	NA*	NA*	-13.92	NA*	NA*	-52.06	-11.97					-59.67	-35.49	NA*	-31.89	-52.51	-9.06	-55.05	NA*			
	snp6	NA*	-32.46	NA*	NA*	-12.67	NA*	NA*	NA*	-20.42	NA*					NA*	NA*	NA*	NA*	NA*	NA*	NA*				
	mgs	NA*	NA*									-11.91	-4.13	-29.36	-15.03									NA*	-16.68	-54.6
gemcitabine	affy1	-145.86	-33.02	-7.25	-5.33	NA*	NA*	-26.31	NA*	-80.04	-6.5					-152.19	-24.72	-86.84	-40.55	-26.6	-15.04	NA*	-20.09			
	snp6	-90.74	-25.37	NA*	-10.96	-21.21	NA*	-20.3	NA*	NA*	NA*					-43.95	-13.53	NA*	-3.27	-45.53	NA*	-55.25	NA*			
	mgs	NA*	-0.83									NA*	NA*	-14.18										-10.99	NA*	-135.34
gw843682x	affy1	NA*	-8.44	-28.49	NA*	NA*	NA*	NA*	NA*	NA*	NA*					NA*	NA*	-13.09	-7.19	NA*	NA*	NA*	NA*			
	snp6	-13.47	NA*	-20.92	NA*	-14.38	-5.11	NA*	NA*	-6.32	NA*					NA*	NA*	NA*	NA*	-171.82	-15.37	NA*	NA*			
	mgs	-34.22	NA*									NA*	NA*	NA*	NA*									NA*	NA*	NA*
linsitinib	affy1	-295.43	-24.39	-12.12	NA*	-24.88	-8.69	-11.04	-18.04	-39.69	NA*					-92.5	NA*	-64.83	NA*	-264.64	-5.55	-14.2	-6.3			
	snp6	-63.59	NA*	NA*	NA*	NA*	NA*	NA*	NA*	NA*	NA*					NA*	-0.39	-142.62	NA*	NA*	NA*	NA*	-1.53			
	mgs	-33.86	-8.82									NA*	-9.58	-14.85	-6.71									-49.33	-7.09	-24.71
mk2206	affy1	-165.9	-7.38	-20.18	-11.31	-29.44	-12.15	-31.87	-14.07	-7.84	NA*					-92.48	-14.14	-59.55	-20.75	-282.01	-21.1	NA*	-6.15			
	snp6	-42.67	NA*	-14.68	NA*	-13.67	NA*	NA*	-4.27	NA*	NA*					-31.28	NA*	-123.93	NA*	-169.29	-13.02	NA*	-3.18			
	mgs	NA*	NA*									-7.5	NA*	-14.72	-0.19									-88.9	NA*	-33.99
navitoclax	affy1	-65.95	-71.7	-37.45	-27.46	-34.55	-15.75	-37.84	-27.15	-35.86	-24.36					-70.9	-46.35	-118.94	-64.43	-89.56	-40.7	-63.43	-43.3			
	snp6	-124.71	-18.59	-36.25	-14.62	-8.72	NA*	-22.37	NA*	-8.49	NA*					-107.81	NA*	-12.7	NA*	-48.59	NA*	-8.58	-3.56			
	mgs	NA*	NA*									NA*	NA*	NA*	-8.05									-40.31	NA*	-18.41

Tabla D.16: Modelos para drogas compartidas, P: Garnett, T: Basu, p-value (1/2)

Primary (P): Garnett, Transferred (T): Basu
 p -value

		NEA												GNEA														
		RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference		
		P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T			
nutlin	affy1	-146.18	-91.54	-32.47	-10.63	-25.04	-9.73	-24.93	-9.99	-34.06	NA*					-13.58	NA*	-72.27	-18.15	-288.58	NA*	-29.28	-8.17					
	snp6	-126.6	-35.49	-23.51	-10.12	-29.79	-12.97	-54.49	NA*	NA*	NA*					-22.22	-25.49	-127.31	-13.58	-66.92	-12.23	NA*	NA*					
	mgs	-42.67	-47.3							-43.61	-4.17	-38.91	-36.89										-85.48	-32.46	-57.22	-21.98		
nvpauy922	affy1	-175.2	-44.74	-36.56	NA*	-49.27	-12.84	-81.81	NA*	-61.78	-31.34					-115.21	NA*	-129.96	-21.26	-90.83	-12.28	-94.73	-10.46					
	snp6	-62.06	-10.57	-8.74	NA*	-33.62	NA*	NA*	NA*	-15.58	NA*					NA*	NA*	-21.52	NA*	-156.76	NA*	NA*	NA*					
	mgs	NA*	NA*							-7.44	NA*	NA*	NA*									NA*	-1.55	NA*	NA*			
olaparib	affy1	-98.72	-46.43	-53.41	-41.24	-85.12	-21.06	-59.22	-24.15	-51.92	-29.25					-100.35	-43.97	-143.32	-21.92	-165.38	-37.18	-44.86	-24.86					
	snp6	-102.03	-31.19	-38.47	NA*	NA*	NA*	-18.23	NA*	NA*	-3.5					-179.55	NA*	-81.43	-10.57	-88.36	NA*	-39.16	NA*					
	mgs	-13.09	NA*							-6.25	-11.51	-49.1	NA*										-11.2	NA*	NA*	NA*		
pac1	affy1	-287.08	-39.66	-53.57	-13.4	-53.28	-20.55	-32.21	-8.42	-16.28	NA*					-85.84	NA*	-111.14	-10.11	-93.21	NA*	-38.85	-15.14					
	snp6	-54.23	NA*	-26.57	NA*	NA*	-13.19	-14.17	NA*	-22.1	-3.45					-60.24	NA*	NA*	-5.37	-187.13	NA*	-128.09	NA*					
	mgs	-12.98	NA*							NA*	NA*	-43.67	-3.78										-19.46	NA*	-52.1	NA*		
parthenolide	affy1	-45.84	-6.17	NA*	-4.98	NA*	-5.75	NA*	-19.81	-18.58	NA*					-88.87	-10.05	-83.06	-10.04	-121.17	-6.43	-50.08	NA*					
	snp6	-15.01	NA*	-11.23	NA*	-15.76	-7.72	-41.64	NA*	NA*	NA*					-251.57	NA*	NA*	NA*	-27.42	NA*	-15.21	NA*					
	mgs	-23.47	-3.55							-33.83	NA*	-41.65	-9.03										-26.59	-5.53	NA*	NA*		
qs11	affy1	-94.02	NA*	-38.25	-4.68	-17.14	NA*	-32.86	NA*	-30.17	NA*					-93.29	NA*	-46.45	-4.04	NA*	NA*	NA*	NA*					
	snp6	NA*	NA*	NA*	NA*	NA*	NA*	NA*	-4.42	-24.07	NA*					-28.99	NA*	-42.95	NA*	-48.93	NA*	NA*	NA*					
	mgs	NA*	NA*							-38.01	NA*	NA*	NA*										-63.28	-0.27	-84.37	-2.61		
selumetinib	affy1	-81.02	-41.18	-40.44	-28.71	-54.19	-24.37	-58.15	-18.59	-22.8	-0.57					-188.11	-50.27	-194.37	-30.03	-92.12	-20.48	-46.51	NA*					
	snp6	NA*	NA*	-13.5	NA*	-28.27	-8.16	NA*	-4.66	NA*	NA*					-10.4	NA*	-25.76	NA*	NA*	NA*	NA*	-31.14		NA*	-8.31	NA*	-11.02
	mgs	-63.26	-20.2							NA*	NA*	NA*	-3.56															
sirolimus	affy1	NA*	-11.64	-44.01	-4.84	NA*	NA*	-13.98	NA*	NA*	NA*					NA*	-11.37	NA*	-5.28	NA*	NA*	-13.24	NA*					
	snp6	NA*	-3.47	-5.87	NA*	NA*	NA*	NA*	NA*	-13.93	NA*					NA*	NA*	NA*	-8.8	NA*	NA*	NA*	NA*					
	mgs	-14.08	NA*							-9.08	NA*	-5.63	-5.63										NA*	-8.49	NA*	NA*		
sn38	affy1	-146.07	-58.43	-34.37	-13.89	-43.29	-12.45	-59.4	-7.43	-21.41	-4.39					-166.28	-24.95	-97.8	-64.02	-145.99	-4.69	-18.99	-20.11					
	snp6	-89.95	-8.09	NA*	-6.32	NA*	NA*	NA*	NA*	-22.66	-6.13					NA*	-25.26	NA*	NA*	-33.49	NA*	NA*	NA*					
	mgs	NA*	NA*							NA*	NA*	NA*	NA*										-15.98	NA*	NA*	NA*		
tanespimycin	affy1	-272.61	-74.76	-52.1	-17.54	-64.46	NA*	-60.35	NA*	-36.95	NA*					-188.42	-23.72	-236.15	-0.42	-201.93	-26.05	-114.42	-9.42					
	snp6	-8.89	NA*	-28.65	NA*	-16.04	NA*	NA*	NA*	-6.81	NA*					-84.32	NA*	-17.29	NA*	NA*	NA*	NA*	NA*					
	mgs	-18.37	NA*							-13.63	NA*	-16.73	NA*										NA*	-6.95	NA*	NA*		
tipifarnibp1	affy1	-144.06	NA*	-17.52	-19.61	-27.37	-23.72	NA*	-12.88	-39.59	-14.48					-125.71	-24.26	-164.35	-27.09	-80.01	-40.04	-42.56	NA*					
	snp6	-117.91	NA*	NA*	NA*	NA*	NA*	NA*	NA*	NA*	NA*					NA*	NA*	NA*	NA*	-24.02	-8.78	NA*	NA*					
	mgs	NA*	-8.65							NA*	NA*	NA*	NA*										NA*	NA*	NA*	-6.25		
tozasertib	affy1	-132.67	NA*	-31.31	-5.76	-21.96	-6.25	-12.43	-4.02	NA*	NA*					-148.82	NA*	-98.3	-12.46	-245.26	-3.11	NA*	NA*					
	snp6	-9.03	-4.44	NA*	NA*	NA*	NA*	-74.42	NA*	-26.41	NA*					NA*	NA*	NA*	NA*	-112.27	NA*	NA*	-4.57					
	mgs	NA*	NA*							-8.83	NA*	NA*	NA*										NA*	NA*	NA*	NA*		
tw37	affy1	-154.07	NA*	-51.68	-7.09	-31.39	-13.15	-42.17	-9.42	-62.59	NA*					-122.15	-36.18	-75.29	NA*	-134.07	NA*	-29.57	NA*					
	snp6	-83.53	-9.26	-29.28	-3.36	NA*	-14.2	NA*	NA*	NA*	NA*					-54.39	NA*	-68.42	-39.48	-158.27	NA*	-473.93	NA*					
	mgs	NA*	NA*							NA*	NA*	-7.74	NA*										-34.73	NA*	-51.23	NA*		
vorinostat	affy1	-115.32	-51.23	-31.68	-27.77	-48.17	-28.67	-29.14	-25.73	-30.93	-15.83					-31.91	-44.77	-69.24	-15.07	-189.63	-29.65	-56.17	-27.34					
	snp6	NA*	-37.62	-22.43	NA*	NA*	NA*	-43.91	NA*	-11.93	NA*					-33.72	-8.82	NA*	NA*	-38.25	NA*	NA*	NA*					
	mgs	-22.01	NA*							NA*	-3.34	-6.04	-4.32										-16.28	-6.59	-76.66	-10.71		

Tabla D.17: Modelos para drogas compartidas, P: Garnett, T: Basu, p -value (2/2)

	NEA														GNEA											
	RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference	
	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T
atra	affy1	0.98	0.83	1.06	1	1.1	1.03	0.92	0.98	1.01	0.93					1.08	1.01	1.09	0.97	0.84	0.98	0.98	0.99			
	snp6	1.53	1.4	1.56*	1.55*	1.6*	1.51	1.64	NA	1.15	1.22*					1.55*	1.55*	1.58*	1.59*	1.63*	1.63*	1.2	NA*			
	mgs	1.46	NA*								1.48*	1.53	1.49*	1.47										1.49	1.5*	1.49
bez235	affy1	1.31	1.3	1.25	1.28	1.32	1.32	1.28	1.34	1.24	1.18					1.32	1.23	1.26	1.34	1.32	1.3	1.13	1.3			
	snp6	1.19	NA*	1.24	NA*	1.31	1.37	1.13	NA*	0.91	0.86					1.34	1.34	1.36*	1.43*	1.29	1.34	0.97	0.94			
	mgs	1.26	NA*								1.22	NA*	1.25	NA										1.28*	NA*	1.21
bi2536	affy1	1.5	NA*	1.82	2.11	1.85	1.97	1.92	NA*	1.86	NA					1.83	NA*	1.86	NA*	2.06	NA*	1.82	NA*			
	snp6	1.91	2.3*	2.2	2.3*	2.21	2.3*	2.35	2.28	2.43	2.34					2.21	NA	2.16	NA	2.31	2.45*	2.37	NA*			
	mgs	2.29*	2.26*								2.25	NA*	2.26*	NA*										2.26*	NA*	2.25
bms536924	affy1	1.5	NA*	1.72	1.69	1.67	1.71*	1.67	1.68	1.71	1.73					1.59	NA*	1.42	1.69	1.62	1.7	1.66	1.74*			
	snp6	1.59	1.74*	1.73	1.73*	1.69	1.73*	1.64	NA*	1.51*	1.47					1.76*	NA*	1.74*	NA*	1.67	1.69*	1.5	NA*			
	mgs	1.76*	NA*								1.75	1.74	1.76	1.72										1.69	1.75	1.75
bms754807	affy1	1.79	1.71	1.8	2*	2.01*	1.95	2.01*	2.02	2.04	2.05					1.82	1.96	1.91	1.93	2.02*	2	2.03*	2.09*			
	snp6	1.95	2.02*	2.02*	2.02*	1.97	NA*	2.07	NA*	1.92*	1.87*					2*	2.05*	1.97	NA*	2.02	2.06	1.87*	NA*			
	mgs	2.07	NA*								2.05*	2.15*	2.1*	2.02										2.09*	NA*	2.09*
bortezomib	affy1	1.79	1.98	1.77	1.8	1.92	2	1.79	1.93	1.9	1.88					1.81	2.01*	1.48	2.01	1.69	2	1.57	2.04*			
	snp6	2.15	2.15	2.15*	NA*	2.08	2.04	2.1*	NA*	2.17*	2.16					2.16*	NA*	2.02	2.12	2.08	2.08*	2.15	NA*			
	mgs	2.03	2.05*								1.97	2.03*	2.08*	NA*										1.86	2.03*	1.99
bx795	affy1	1	0.97	0.87	0.97	1	1.03	1.05	0.94	0.95	0.93					0.97	1.11*	0.87	1.02	1.12*	1.07	1.09	1.07			
	snp6	1.04	0.97	1.09*	NA*	1.07*	1.08*	1.07*	1.05*	0.99	1.09*					1.08*	1.06*	1.06*	1.08*	1.06	1.07*	1.06	NA*			
	mgs	1.11*	NA*								1.1	NA*	1.1*	NA										1.06	1.1*	1.07
chir99021	affy1	1.61	1.87*	1.65	1.86	1.74	1.8	1.68	1.86	1.88	1.86					1.65	1.89	1.63	1.36	1.83	1.61	1.93*	1.89			
	snp6	1.91	1.83	1.96*	1.92*	1.94*	NA*	1.94*	NA	1.93*	1.97					1.91*	1.89	1.93*	NA*	1.81	1.97*	1.97*	NA*			
	mgs	1.95*	1.98*								1.95*	NA*	1.96*	1.96*										1.91	1.95*	1.96*
etoposide	affy1	2.29	2.56	2.32	2.62	2.38	NA*	2.21	NA*	2.36	2.66					1.88	3.38*	1.86	NA*	2.44	3.15	2.07	3.05			
	snp6	2.1	NA*	3.12	NA*	3.23*	3.12	3.24	NA*	2.92*	2.73					2.94	NA*	3.16*	NA*	3.21*	NA*	2.88	2.92*			
	mgs	3.17*	NA*								3.13	3.1	2.9	3.08										3.06	NA*	3.11*
gemcitabine	affy1	2.8	2.7	2.97	NA	2.72	NA*	3.08	3.25	3.1	3.12					2.65	3.44*	2.47	3.41	3.11	3.39	2.88	NA*			
	snp6	3.26	3.21	3.44	NA*	3.27	3.62*	3.61*	3.57*	3.6	NA*					3.2	3.41	3.54	NA*	3.62	3.63*	3.73*	3.65*			
	mgs	3.36	NA*								3.39*	NA*	3.23	NA*										3.38*	NA	3.41*
gw843682x	affy1	1.82	NA*	2.55	2.48	2.14	NA*	2.41	NA*	2.29	NA*					2.07	NA*	1.75	NA	2.36	NA*	2.24	NA*			
	snp6	2.05	2.84	2.85*	2.87*	2.69	2.84	2.99*	NA*	2.74	NA					2.86	NA*	2.84	NA*	2.77	2.96	2.99*	NA*			
	mgs	2.86*	2.85*								2.79	NA*	2.79	NA*										2.82	NA*	2.76
linsitinib	affy1	0.85	1.01	1	0.98	0.99	0.99	0.87	NA	0.97	0.96					0.78	0.95	0.94	1.01	0.95	1.02*	0.96	1			
	snp6	1.05	1.05	1.07*	NA*	1.06	NA*	1.1*	NA*	1.02*	NA*					1.07	NA*	1.04	1.06*	1.08*	NA*	0.99	NA*			
	mgs	1.01	0.99								0.94	NA*	1.04	0.94										1.02	1.03	1.02
mk2206	affy1	1.05	1.29	1.18	1.32	1.08	1.29	1.15	1.92*	1.41	1.61					1.17	1.35	0.98	1.87	1.05	0.98	1.14	NA*			
	snp6	1.73	1.87*	1.84*	1.37	1.94*	1.87*	1.59	NA*	1.55	NA*					1.85*	1.83*	1.79*	1.89*	1.7	1.39	1.54	NA*			
	mgs	1.86*	NA*								1.87*	NA	1.8	1.8										1.84	1.83	1.91*
navitoclax	affy1	2.07	2.42	2.8	2.75	2.82	2.89	2.86	2.87	2.85	2.86					2.48	2.67	2.54	2.5	2.59	2.65	2.69	2.77			
	snp6	2.63	2.6	2.91	2.89	2.93*	2.92*	2.85	2.9	2.97*	2.95*					2.93*	2.95*	2.91	2.96*	2.88	2.93	2.95	NA			
	mgs	2.96	NA*								2.97	NA*	2.98	NA*										2.96	2.99*	2.96

Tabla D.18: Modelos para drogas compartidas, P: Basu, T: Garnett, MSE (1/2)

Primary (P): Basu, Transferred (T): Garnett
MSE

	NEA														GNEA												
	RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference		
	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	
nutlin	affy1	1.89	2.21	2.4	2.42	2.28	2.46	2.38	2.34	2.42	2.28					2.47*	NA	2.41	2.36	2.49	2.34	2.36	2.47				
	snp6	2.43	2.25	2.46	2.39	2.39	2.37	2.31*	2.29	2.32*	NA*					2.49	2.2	2.19	2.31	2.24	2.26	2.28	NA*				
	mgs	2.47	2.35									2.56	2.49	2.4	2.48									2.29	2.33	2.54	2.42
nvpauy922	affy1	0.8	0.89	0.87	0.94	0.89	0.9	0.88	1.01	0.91	0.88					0.82	0.97	0.93	1.02*	0.94	1.01	0.7	1.03*				
	snp6	0.88	0.95	1.03	1.08*	1.07*	1.07*	1.15*	NA*	0.96*	0.97*					0.94	NA*	1.03	1.07*	1.1*	1.05	0.95*	NA*				
	mgs	1*	NA*									0.94	NA	0.99*	NA*									0.97	NA*	0.98*	NA*
olaparib	affy1	1.45	1.39	1.32	1.33	1.58	1.43	1.58	1.51	1.4	1.46					1.4	1.46	1.48	1.47	1.33	1.53	1.39	1.6*				
	snp6	1.26	1.52	1.56*	1.56*	1.62*	NA*	1.53*	1.41	1.21	NA*					1.55*	1.54	1.59	1.36	1.49	1.49*	1.28*	1.2				
	mgs	1.54	1.58*									1.56	NA	1.56*	1.59*									1.61*	NA	1.62*	NA*
pac1	affy1	0.82	0.86	0.88	0.85	0.85	0.85	0.89	0.9	0.88	0.91*					0.9*	0.91*	0.88	0.91*	0.9	0.9	0.89	0.92*				
	snp6	0.88*	0.88	0.89	0.88	0.85	NA*	0.87*	0.87*	0.91	0.93*					0.88*	0.9*	0.88	NA*	0.87*	0.87*	0.92	0.92				
	mgs	0.93*	0.93*									0.93*	NA*	0.91	0.92*									0.93	0.92*	0.93*	0.93*
parthenolide	affy1	1.13	1.83	1.38	NA*	1.45	NA*	1.34	NA*	1.37	1.66					1.14	1.46	1.49	2.18*	1.4	1.48	1.44	1.52				
	snp6	1.6	1.9	1.9*	1.89	1.73	1.93	1.8	2*	2.17*	NA*					1.97*	1.91*	1.92*	NA*	1.81	1.94*	2.06	2.2*				
	mgs	1.91	1.93									1.95*	2.06*	1.9	2.01*									1.72	1.85	1.86	NA*
qs11	affy1	1.08*	1.08*	1.05	1.09*	1.03	1.09	1.09	1.09*	1.1*	1.09*					1.05	1.47	1.08	1.1*	1.04	NA*	1.01	NA*				
	snp6	1.08*	NA*	1.11*	NA*	1.11*	NA*	1.14	NA*	1.04	1.04					1.08	1.1*	1.07*	1.06	1.14*	1.13*	0.99	NA*				
	mgs	1.1*	NA*									1.09*	1.06	1.09	NA*									1.08	1.03	1.08	1.1
selumetinib	affy1	2.12	2.03	2.54	2.35	2.46	2.71	2.75	2.91	2.96	3.01*					2.33	2.42	2.14	2.42	2.75	2.71	2.87	2.69				
	snp6	3.01*	NA*	2.93	3.02*	2.86	2.99*	2.64	NA*	2.81	NA*					2.99*	NA	2.95	2.91	2.88	NA*	2.63	NA*				
	mgs	2.78	2.6									2.82	NA*	2.96	NA*									2.96	NA*	2.84	NA*
sirolimus	affy1	1.25	NA*	1.25	1.28	1.2	NA*	1.27	1.49*	1.35	NA*					1.19	NA*	1.24	NA*	1.28	NA*	1.3	NA				
	snp6	1.35	NA*	1.41	1.44*	1.43*	NA*	1.32	NA*	1.11	1.12*					1.38	NA*	1.38	NA*	1.38*	NA*	1.02	NA*				
	mgs	1.46*	1.46*									1.46*	1.46*	1.45	NA									1.42	NA*	1.45	NA*
sn38	affy1	2.33	2.32	2.5	2.57	2.71	2.75	2.75	2.74	2.73	3.02					2.65	2.66	2.41	3.34*	3.04	2.94	2.68	3.11				
	snp6	2.77	2.89	3.24	NA*	3.34	NA*	3.59	NA*	3.65	3.53					3.05	NA*	3.21	NA*	3.68*	3.61	3.76*	NA*				
	mgs	3.32*	NA*									3.31*	NA*	3.05	NA*									3.33*	NA	3.1	NA*
tanespimycin	affy1	1.68	1.68	1.84	1.87*	1.84*	1.86*	1.86*	1.86*	1.83*	1.81					1.86	1.84	1.85	1.85*	1.72	1.81	1.71	1.88*				
	snp6	1.78*	1.74*	1.77*	1.74	1.76*	1.73*	1.73*	NA*	1.82*	NA					1.75*	1.74	1.76*	NA	1.72*	NA*	1.83*	NA*				
	mgs	1.91*	1.87									1.92*	1.9*	1.91*	1.89*									1.87	NA*	1.91*	NA*
tipifarnibp1	affy1	0.63	0.67	0.6	0.74	0.61	0.72	0.67	NA*	0.67	0.72					0.58	0.73	0.61	0.74	0.65	0.69	0.62	0.76*				
	snp6	0.68	0.74	0.76*	NA*	0.75*	NA*	0.71	NA*	0.67*	NA*					0.75*	NA*	0.76*	NA*	0.74	0.74*	0.67	NA*				
	mgs	0.75	NA*									0.77*	NA*	0.77*	NA*									0.75	NA*	0.75	NA*
tozasertib	affy1	2.21	2.46	2.36	2.48	2.4	2.4	2.38	2.5	2.48	NA*					2.27	2.61*	2.08	2.45	2.51	2.37	2.54	NA*				
	snp6	2.44	2.56	2.57*	NA*	2.58*	NA*	2.53*	2.51*	2.23	2.33*					2.53*	NA*	2.42	NA*	2.48*	2.45*	2.27	NA*				
	mgs	2.56*	NA*									2.57*	2.57*	2.6*	NA*									2.56	NA*	2.54	NA*
tw37	affy1	2.13*	2.06	1.99	2.05	2.05	1.93	1.85	2.09	1.78	1.99					1.78	1.92	1.87	2.04	2.09*	2.05	1.86	2.02				
	snp6	2.1	2.13	2.06	2.14	1.81	NA*	1.92	NA*	1.57	NA*					2.19*	2.15*	1.69	2.09	2.14	2.12*	1.54	1.65				
	mgs	2.03	NA*									2.06*	NA*	2.08*	2.08*									1.96	1.96	2.01	2.05*
vorinostat	affy1	0.96	1.03	1.14	1.12	1.23	1.12	1.21	1.27	1.17	1.16					1.13	1.22	1	1.63*	1.24	1.21	1.16	1.29				
	snp6	1.36	NA*	1.61	1.63*	1.67*	NA*	1.68	1.71*	1.78	1.79*					1.6	1.64*	1.64*	NA*	1.68	1.72*	1.75*	NA*				
	mgs	1.6*	NA									1.59	NA*	1.58	NA									1.56	NA	1.59	1.54

Tabla D.19: Modelos para drogas compartidas, P: Basu, T: Garnett, MSE (2/2)

Primary (P): Garnett, Transferred (T): Basu
MSE

RAW		NEA										GNEA													
		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference	
		P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T		
atra	affy1	1.4	1.37	1.4	1.38	1.38	1.36	1.47	1.46	1.44	1.38					1.41	1.44	1.38	1.41	1.46	1.43	1.53	1.42		
	snp6	1.55	1.57	1.58	NA*	1.57	NA*	1.55	1.56	1.52	1.54*					1.58	NA*	1.56	NA*	1.56	NA*	1.56*	1.55*		
	mgs	1.57*	1.57*							1.51	NA*	1.56	NA*									1.55	1.56*	1.56	1.56
bez235	affy1	0.54	0.56	0.56	0.58	0.56	0.57	0.58	0.59	0.58	0.58					0.55	0.57	0.56	0.58	0.58	0.56	0.59	0.55		
	snp6	0.59*	0.58	0.59*	0.59*	0.59	0.6*	0.62*	0.62*	0.57	0.55					0.59	0.6*	0.58	NA*	0.61	0.63	0.56	0.57*		
	mgs	0.61*	NA							0.61*	0.61*	0.6	0.61*								0.61*	NA*	0.6	0.59	
bi2536	affy1	2.14*	2.17*	2	2.14	2.12	2.16*	2.16*	2.15*	2.14	2.15*					2.13*	2.15*	2.15*	2.15*	2.16*	2.14*	2.17*	2.13*		
	snp6	2.09	2.14*	2.13	2.14	2.14	2.12	2.14	2.15*	1.81	1.82					2.14	2.15*	2.14	2.14	2.02	2.09	1.84*	1.83*		
	mgs	1.83	NA*							1.96*	NA	1.94*	NA*									1.94*	NA*	1.96*	1.95
bms536924	affy1	2.4*	2.31	2.28	2.38	2.34	2.4*	2.39	2.38	2.29	2.35					2.4*	2.4*	2.19	2.28	1.93	2.26	2.34	2.29		
	snp6	2.38	2.35	2.38	2.4*	2.39	2.38	2.31*	2.24	2.2	NA*					2.4*	NA*	2.43*	NA*	2.26	2.31*	2.27*	NA		
	mgs	2.31*	NA*							2.27	2.23	2.27	2.33*									2.29	2.3	2.32*	2.27
bms754807	affy1	1.39	1.62	1.61	1.69	1.53	NA*	1.63	NA*	1.59	1.64					1.54	1.58	1.43	1.68	1.41	NA*	1.61	NA*		
	snp6	1.57	1.64	1.68	NA*	1.71*	1.71	1.69*	NA	1.69	NA*					1.72	NA*	1.73*	1.74*	1.7	1.72*	1.71*	NA*		
	mgs	1.72*	1.71							1.7	NA*	1.68	NA*									1.71*	NA*	1.68	NA*
bortezomib	affy1	1.14	1.2*	1.1	1.2	1.06	1.15	1.16	1.15	1.14	1.19					1.13	1.2	1.12	1.2	1.15	1.14	1.2	1.17		
	snp6	1.17	1.21	1.2*	NA*	1.2	1.2	1.29*	NA*	1.29	NA*					1.2*	NA*	1.19	1.18	1.2	1.21*	1.32*	1.29*		
	mgs	1.21	1.25*							1.23	1.25*	1.24*	NA*									1.21	1.24*	1.24	1.23*
bx795	affy1	0.9	1.02	0.95	1.03	0.98	0.99	1.1	1.06	1.14	1.11					0.92	1.03	0.96	1.09	1.07	NA*	1.06	1.14		
	snp6	1.02	1.04	1.14*	NA*	1.1	NA*	1.11	NA*	1.12	1.14*					1.15	NA*	1.15	NA*	1.13	NA	1.15*	NA		
	mgs	1.14*	NA*							1.14*	1.15*	1.15	NA*									1.12	1.12	1.13	1.12
chir99021	affy1	1.78	1.92	1.86	1.82	1.87	1.87	1.94	1.99*	1.93	1.9					1.83	1.82	1.81	1.84	1.89	1.98*	1.88	NA*		
	snp6	1.88	1.89	1.96	NA*	1.99*	NA*	1.96	NA*	1.82	NA*					1.93	NA*	1.98*	NA*	1.96	1.97*	1.85*	NA*		
	mgs	1.93	NA*							1.95*	NA*	1.95	NA*									1.94	1.96	1.92	NA*
etoposide	affy1	1.5	1.53	1.59	1.61*	1.6*	1.6	1.61*	1.61*	1.58	1.62					1.59	1.57	1.6*	1.58	1.6	1.59	1.6	1.63*		
	snp6	1.61*	1.62	1.61*	1.61*	1.6	NA*	1.63*	1.63*	1.56	NA*					1.63*	1.62*	1.62*	NA*	1.65*	NA*	1.57	1.57*		
	mgs	1.5*	NA*							1.49	1.49	1.49	1.48									1.5*	1.48	1.49	NA*
gemcitabine	affy1	1.46	1.49	1.56	1.55	1.57*	1.58*	1.56	1.57*	1.47	1.57					1.53	1.56	1.54	1.5	1.56	1.56	1.59*	1.55		
	snp6	1.54	1.53	1.50*	1.52	1.55	1.56*	1.56	NA*	1.52*	1.52*					1.57	1.52	1.58*	1.57	1.58	1.59*	1.48	NA*		
	mgs	1.49*	NA							1.49*	NA*	1.49*	1.46									1.48	NA*	1.47	NA*
gw843682x	affy1	2.16*	2.07	2.07	2.13*	2.15*	2.14*	2.15*	2.16*	2.15*	2.15*					2.15*	2.14*	2.11	2.12	2.14*	2.16*	2.14*	2.14*		
	snp6	2.13	2.15*	2.12	NA*	2.12	2.13	2.16*	NA*	1.66	1.67*					2.14*	2.14*	2.16*	2.14*	2.07	2.12	1.7*	NA*		
	mgs	1.8	NA*							1.89*	1.9*	1.89*	1.91*									1.9*	1.89*	1.9*	1.89
linsitinib	affy1	2.13	2.61	2.58	2.6*	2.55	2.57	2.58	2.59	2.57	2.59*					2.51	2.62*	2.58	2.61*	2.51	2.6	2.59	2.6		
	snp6	2.55	2.6*	2.59*	NA*	2.6*	2.6*	2.59*	NA*	2.6*	NA*					2.65*	NA	2.61	2.65*	2.64*	NA*	2.6*	NA		
	mgs	2.59	2.55							2.59*	2.57	2.57	NA									2.58	2.58	2.59	2.57
mk2206	affy1	1.44	1.48	1.48	1.49	1.47	1.46	1.49	1.47	1.48	1.49*					1.44	1.47	1.5	1.47	1.44	1.43	1.49*	1.48		
	snp6	1.49	1.51*	1.48	NA*	1.49	NA*	1.52*	1.51	1.55*	1.54*					1.51	NA*	1.5	NA*	1.54	1.53	1.54*	1.53		
	mgs	1.49*	NA*							1.49	NA*	1.48	NA									1.43	1.5*	1.48	NA*
navitoclax	affy1	1.76	1.71	1.75	1.77	1.77	1.87	1.81	1.8	1.79	1.78					1.77	1.76	1.7	1.93	1.73	1.72	1.73	1.82		
	snp6	1.77	1.92	1.89	1.92	1.91	NA*	1.91	1.94*	1.88	NA*					1.89	NA*	1.91	1.93*	1.92	1.94*	1.89	1.87		
	mgs	1.94*	1.93*							1.93*	1.92*	1.92*	1.92									1.91	1.93*	1.91	1.94*

Tabla D.20: Modelos para drogas compartidas, P: Garnett, T: Basu, MSE (1/2)

		NEA														GNEA													
		RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference			
		P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T		
nutlin	affy1	1.41	1.44	1.91	1.93	1.97	1.94	1.91	1.94	1.93	2*					1.98	NA*	1.93	1.91	1.84	2*	1.97	1.99						
	snp6	1.78	1.89	1.92	1.93	1.94	1.96	1.92	NA*	1.75*	NA*					1.93	1.92	1.89	1.9	1.9	1.94	1.75*	1.75*						
	mgs	1.65	1.58									1.77	1.96	1.69	1.68										1.62	1.84	1.7	1.77	
nvpauy922	affy1	1.05	1.1	1.14	1.17*	1.07	1.14	1.12	1.17*	1.13	1.14					1.11	1.17*	1.11	1.16	1.11	1.15	1.13	1.18						
	snp6	1.14	1.15	1.17	1.18*	1.17	NA*	1.13*	NA*	0.82	NA*					1.19*	1.18*	1.17	1.18*	1.12	NA*	0.82*	NA*						
	mgs	1.06*	NA*									1.05	1.06*	1.06*	NA*										1.06*	NA	1.05*	NA*	
olaparib	affy1	1.27	1.26	1.25	1.27	1.26	1.36	1.32	1.35	1.35	1.37					1.29	1.26	1.2	1.42	1.3	1.44	1.43	1.43						
	snp6	1.32	1.42	1.5	NA*	1.48*	NA*	1.38	NA*	1.34*	1.34					1.43	NA*	1.45	1.46	1.37	1.38*	1.33	NA*						
	mgs	1.47	1.47*									1.46	1.43	1.39	NA*										1.45	NA*	1.46*	NA*	
pac1	affy1	1.69	1.73	1.67	1.76	1.72	1.73	1.74	1.76	1.77	1.78*					1.73	NA*	1.69	1.79	1.76	1.8*	1.74	1.75						
	snp6	1.76	NA*	1.78	1.79*	1.79*	1.77	1.74	NA*	1.72	1.73					1.75	NA*	1.77*	1.76	1.72	NA*	1.66	1.74*						
	mgs	1.81	NA*									1.82*	NA*	1.79	1.82										1.8	1.82*	1.78	NA*	
parthenolide	affy1	2	2.02	2.06*	2.04	2.04*	2.02	2.08*	2.03	2.07	2.05*					1.84	2.01	1.91	1.93	1.99	2.04	1.97	2.07*						
	snp6	2.04	2.05*	2.04	NA*	2.06	2.04	1.92	2.07*	2*	NA*					1.9	NA*	2*	NA*	1.96	2.03*	1.99	2*						
	mgs	2.11	2.08									2.08	NA*	1.98	2.1										2.1	2.09	2.13*	2.11*	
qs11	affy1	2.21	NA*	2.23	2.25	2.22	2.26*	2.2	2.26*	2.21	NA*					2.25	2.27*	2.28	2.26	2.27*	2.26*	2.25*	2.25*						
	snp6	2.26*	NA*	2.26*	NA*	2.27*	NA*	2.21*	2.18	2.23	2.26*					2.25	2.25*	2.22	NA*	2.14	NA*	2.3*	2.28*						
	mgs	2.25*	NA*									2.21	NA*	2.25*	2.24*										2.23	NA	2.21	NA	
selumetinib	affy1	1.8	1.75	1.87	1.87	1.86	1.82	1.86	1.9	1.93	NA					1.72	1.91	1.76	1.89	1.9	1.91	1.88	1.96*						
	snp6	1.96*	NA*	1.97	1.96*	1.95	1.96	1.96*	1.96	1.96*	1.96*					1.99	NA*	1.97	2.01*	2.02*	2.01*	1.94*	1.92						
	mgs	1.92	1.89									1.96*	1.95*	1.96*	1.94										1.96*	1.93	1.96*	1.89	
sirolimus	affy1	1.98*	1.97	1.88	1.93	1.97*	1.96*	1.94	1.94*	1.97*	1.97*					1.96*	1.89	1.94*	1.93	1.97*	1.96*	1.95	1.97*						
	snp6	1.95*	1.97	1.95	1.95*	1.96*	NA*	1.93*	1.94*	1.73	1.78*					1.96*	1.97*	1.99*	1.95	1.95*	NA*	1.77*	1.77*						
	mgs	2.02	NA*									2.03	NA*	2	2.01											2.02*	2.01	2.01*	2*
sn38	affy1	0.72	0.73	0.84	0.87	0.83	0.87	0.87	0.86	0.88	0.88					0.84	0.85	0.84	0.82	0.87	0.87	0.89	0.88						
	snp6	0.87	0.87	0.88*	0.88	0.88*	0.88*	0.89*	0.88*	0.88	0.9					0.89*	0.88	0.89*	0.89*	0.89	NA*	0.91*	NA*						
	mgs	0.87*	NA*									0.88*	NA*	0.87*	0.87*											0.86	NA*	0.87*	0.87*
tanespimycin	affy1	0.9	0.91	1.15	1.19	1.16	NA*	1.11	NA*	1.18	NA*					1.07	1.19	1.16	NA	1.11	1.22	1.15	1.26						
	snp6	1.25	NA*	1.24	NA*	1.24	NA*	1.17*	NA*	1.1	NA*					1.23	NA*	1.23	NA*	1.2*	NA*	1.1*	NA*						
	mgs	1.2	NA*									1.2	NA*	1.21	NA*											1.21*	1.19	1.21*	NA*
tipifarnibp1	affy1	1.18	1.23*	1.2	1.21	1.18	1.16	1.25*	1.22	1.2	1.23					1.22	1.22	1.18	1.19	1.18	1.15	1.23	1.26*						
	snp6	1.21	1.23*	1.23*	NA*	1.24*	NA*	1.23*	1.23*	1.12*	NA*					1.25*	NA*	1.24*	NA*	1.24	1.21	1.11*	1.11*						
	mgs	1.09*	1.08									1.09*	NA*	1.08*	NA*												1.08*	1.08*	1.08*
tozasertib	affy1	2.99	3.31*	3.26	3.23	3.21	3.19	3.24	3.25	3.29*	3.37*					3.06	3.3*	3.17	3.28	2.93	3.27	3.27*	3.3*						
	snp6	3.25	3.23	3.25*	NA*	3.3*	NA*	2.78	NA*	2.93	3.11*					3.28*	NA*	3.26*	3.27*	2.65	NA*	3.09*	3.1						
	mgs	3.23*	NA*									3.09	NA*	3.26*	NA*											3.2*	3.19*	3.22*	3.2*
tw37	affy1	0.16	NA*	0.18	0.19	0.18	0.19	0.18	0.19	0.2	0.2*					0.16	0.19	0.18	0.2*	0.18	NA*	0.2	0.2*						
	snp6	0.19	0.2	0.19	0.2	0.2*	0.2	0.19*	0.19*	0.16*	0.16*					0.19	NA*	0.19	0.2	0.18	0.2*	0.12	0.16*						
	mgs	0.2*	0.2*									0.2*	NA*	0.2	NA*											0.2	0.2*	0.19	0.2*
vorinostat	affy1	0.37	0.36	0.39	0.39	0.39	0.4	0.41	0.4	0.41	0.4					0.41	0.37	0.41	0.41	0.42	0.39	0.42	0.42						
	snp6	0.42*	0.4	0.41	0.42*	0.42*	NA*	0.4	0.42*	0.34	0.34*					0.42	0.42	0.43*	NA*	0.42	0.42*	0.35*	NA*						
	mgs	0.4	NA*									0.4*	0.4	0.4	NA											0.4	0.39	0.38	0.39

Tabla D.21: Modelos para drogas compartidas, P: Garnett, T: Basu, MSE (2/2)

		NEA														GNEA											
		RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference	
		P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T
tanespimycin	affy1	0.97	0.66	0.53	0.35	0.55	0.45	0.24	0.3	0.41	0.24					0.74	0.56	0.82	0.57	0.66	0.48	0.85	0.26				
	snp6	0.66	NA*	0.31	0.17	0.3	NA*	0.16	NA*	NA*	-0.04					NA*	0.18	0.36	0.04	0.13	NA*	0.12	NA*				
	mgs	0.25	NA*									0.13	0.11	0.29	0.12										0.2	NA*	0.61
azd0530	affy1	0.58	0.26	0.4	0.18	0.43	0.21	0.12	NA*	0.35	0.22					0.77	0.22	0.53	0.24	0.66	0.21	0.62	0.14				
	snp6	0.58	0.33	NA*	NA*	NA*	NA*	0.47	NA*	0.49	NA*					0.1	NA*	0.81	NA*	0.61	NA*	NA*	NA*				
	mgs	NA*	NA*									0.29	0.12	NA*	NA*										0.57	NA*	0.52
selumetinib	affy1	0.86	0.54	0.62	0.5	0.67	0.5	0.52	0.37	0.46	0.18					0.85	0.52	0.93	0.59	0.75	0.41	0.68	0.28				
	snp6	0.75	NA*	NA*	NA*	NA*	NA*	0.33	NA*	NA*	NA*					0.25	-0.02	NA*	NA*	0.64	NA*	0.65	NA*				
	mgs	0.77	0.47									0.21	NA*	0.59	NA*										0.72	NA*	0.68
erlotinib	affy1	0.58	0.45	0.51	0.39	0.6	0.38	0.36	0.27	0.39	0.23					0.68	0.5	0.7	0.42	0.58	0.5	0.62	0.27				
	snp6	0.67	0.18	0.34	0.13	NA*	0.09	0.23	0.18	0.29	NA*					0.55	NA*	0.61	NA*	0.61	NA*	0.63	0.35				
	mgs	NA*	NA*									0.27	NA*	0.28	NA*										0.36	NA*	0.39
lapatinib	affy1	0.62	0.49	0.48	0.37	0.47	0.4	0.34	0.27	0.32	0.25					0.7	0.46	0.66	0.51	0.64	0.41	0.74	0.41				
	snp6	0.58	0.31	0.2	NA*	NA*	NA*	0.16	NA*	0.27	0.21					0.37	NA*	0.32	0.12	0.37	NA*	0.52	NA*				
	mgs	0.16	NA*									NA*	NA*	0.19	NA*										NA*	NA*	0.77
nilotinib	affy1	0.55	0.31	0.33	0.21	0.37	0.23	0.24	0.25	0.31	0.27					0.7	0.31	0.8	0.21	0.27	0.39	0.27	0.27				
	snp6	0.34	0.12	0.24	0.1	0.31	NA*	0.16	0.15	NA*	NA*					NA*	0.09	0.64	0.27	NA*	NA*	NA*	NA*				
	mgs	NA*	NA*									NA*	NA*	NA*	NA*										NA*	NA*	NA*
nutlin	affy1	0.51	0.36	0.23	0.21	0.27	0.25	0.19	0.22	0.22	0.22					0.67	0.06	0.18	0.33	0.45	0.19	0.35	0.15				
	snp6	0.38	0.32	0.35	NA*	NA*	0.22	NA*	NA*	0.34	NA*					0.36	NA*	NA*	0.13	NA*	NA*	0.79	NA*				
	mgs	0.44	0.26									0.21	0.27	0.23	0.22									0.45	0.28	0.36	0.29
paclitaxel	affy1	0.8	0.39	0.53	0.36	0.5	0.02	0.38	NA*	0.46	NA*					0.78	NA*	0.64	NA*	0.9	NA*	0.6	NA*				
	snp6	0.5	0.15	0.18	NA*	0.22	0.11	0.35	NA*	NA*	NA*					NA*	NA*	0.14	NA*	0.26	NA*	NA*	NA*				
	mgs	0.2	0.12									0.36	NA*	0.2	NA*										NA*	NA*	0.2
pd0325901	affy1	0.8	0.65	0.65	0.53	0.56	0.51	0.58	0.38	0.44	0.28					0.85	0.59	0.81	0.63	0.62	0.51	0.67	0.3				
	snp6	0.68	0.11	NA*	NA*	NA*	NA*	0.2	0.34	NA*	0.17					NA*	NA*	0.12	0.23	0.13	0.33	0.37	0.23				
	mgs	0.57	0.48									0.28	NA*	0.53	NA*										0.62	0.48	0.72
pd0332991	affy1	0.62	0.56	0.54	0.26	0.49	0.37	0.46	0.41	0.49	0.37					0.59	NA*	0.59	0.39	0.69	0.27	0.66	0.36				
	snp6	0.54	0.38	0.45	NA*	0.27	NA*	NA*	NA*	NA*	NA*					NA*	NA*	0.34	NA*	0.42	NA*	0.34	NA*				
	mgs	0.35	0.24									0.42	0.13	0.42	0.15										0.16	0.14	0.3
crizotinib	affy1	0.75	0.38	0.3	NA*	0.49	0.2	0.2	NA*	0.43	NA*					0.66	0.43	0.63	0.46	0.61	0.37	0.54	0.27				
	snp6	0.34	NA*	0.31	NA*	0.22	NA*	0.23	NA*	0.27	NA*					NA*	NA*	0.66	NA*	NA*	NA*	0.5	NA*				
	mgs	0.39	NA*									NA*	NA*	NA*	0.12										0.17	NA*	NA*
pha665752	affy1	0.58	0.2	0.4	0.16	0.37	0.19	0.26	NA*	0.52	0.21					0.25	NA*	0.42	0.1	0.62	0.11	0.55	0.11				
	snp6	0.34	NA*	0.23	NA*	NA*	NA*	NA*	-0.02	NA*	NA*					0.35	NA*	0.31	NA*	NA*	NA*	0.11	NA*				
	mgs	NA*	NA*									0.17	NA*	NA*	NA*										0.74	NA*	0.12
plx4720	affy1	0.55	0.39	0.43	0.32	0.39	0.34	0.22	0.32	0.26	0.08					0.53	0.23	0.94	0.37	0.44	0.25	0.45	0.32				
	snp6	0.37	0.28	NA*	NA*	NA*	0.13	0.27	NA*	0.25	0.17					0.28	0.31	NA*	0.2	NA*	NA*	0.22	NA*				
	mgs	0.46	0.31									NA*	NA*	0.14	NA*										0.49	0.27	0.48
sorafenib	affy1	0.75	NA*	0.38	NA*	0.39	NA*	0.27	NA*	0.32	NA*					0.71	NA*	0.49	NA*	0.43	0.19	0.43	NA*				
	snp6	0.47	NA*	0.3	NA*	NA*	0.17	NA*	NA*	NA*	NA*					NA*	0.12	0.56	NA*	NA*	NA*	NA*	NA*				
	mgs	0.38	NA*									NA*	0.03	0.23	NA*										0.44	NA*	0.33
nvptae684	affy1	0.98	0.24	0.13	0.12	0.4	0.02	0.32	0.15	NA*	NA*					0.72	0.28	0.68	0.31	0.77	0.04	NA*	NA*				
	snp6	0.29	0.37	0.24	NA*	NA*	NA*	NA*	NA*	NA*	NA*					0.63	NA*	0.78	NA*	0.12	0.07	NA*	NA*				
	mgs	NA*	NA*									NA*	NA*	NA*	0.09										0.2	NA*	NA*

Tabla D.22: Modelos para drogas compartidas, P: Barretina, T: Garnett, r_s

		NEA																GNEA									
		RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference	
		P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T
tanespimycin	affy1	0.81	0.7	0.46	0.51	0.53	0.45	0.5	0.17	0.42	0.39					0.72	0.67	0.75	0.65	0.73	0.41	0.53	0.62				
	snp6	0.18	0.4	0.32	NA*	0.19	NA*	NA*		0.12	0.14	NA*				0.47	NA*	0.19	0.25	NA*	NA*	NA*	NA*				
	mgs	0.2	NA*									0.24	0.11	0.23	NA*									NA*	NA*	NA*	0.43
azd0530	affy1	0.68	0.5	0.47	0.41	0.58	0.47	NA*	0.25	0.53	0.33					0.65	0.37	0.45	0.39	0.67	0.47	0.83	NA*				
	snp6	0.66	0.55	NA*	NA*	NA*	NA*	NA*	0.4	NA*	NA*					NA*	NA*	0.81	NA*	0.4	0.1	NA*	NA*				
	mgs	NA*	NA*									0.3	NA*	NA*	NA*									NA*	0.53	0.31	NA*
selumetinib	affy1	0.62	0.58	0.47	0.47	0.54	0.5	0.52	0.48	0.34	0.35					0.81	0.53	0.81	0.64	0.63	0.54	0.42	0.45				
	snp6	NA*	0.55	0.24	NA*	0.38	NA*	NA*	NA*	NA*	NA*					0.11	0.13	0.33	NA*	NA*	NA*	NA*	0.32				
	mgs	0.56	0.42									NA*	NA*	NA*	0.31									NA*	0.38	NA*	0.35
erlotinib	affy1	0.69	0.55	0.47	0.43	0.49	0.47	0.61	0.32	0.6	0.5					0.89	0.55	0.65	0.52	0.96	0.37	0.67	NA*				
	snp6	0.48	0.4	0.54	0.27	0.35	NA*	0.46	NA*	0.42	NA*					0.66	NA*	NA*	0.32	NA*	NA*	0.82	NA*				
	mgs	NA*	NA*									NA*	NA*	NA*	NA*									0.79	NA*	NA*	NA*
lapatinib	affy1	0.85	0.48	0.53	0.5	0.52	0.4	0.5	0.39	0.54	0.41					0.8	0.55	0.83	0.52	0.85	0.59	0.78	0.6				
	snp6	0.68	0.44	0.54	0.14	0.4	NA*	NA*	NA*	0.41	0.23					NA*	0.1	0.34	0.14	0.89	NA*	NA*	0.39				
	mgs	0.41	0.07								0.53	NA*	NA*	0.21										0.52	NA*	0.5	0.32
nilotinib	affy1	0.5	0.28	0.29	0.2	0.32	0.32	0.29	0.18	0.46	0.17					0.41	0.31	0.38	0.3	0.84	0.09	0.63	0.13				
	snp6	0.6	NA*	0.23	0.13	0.33	0.13	0.27	0.03	0.11	NA*					0.58	NA*	0.58	NA*	0.43	NA*	NA*	NA*				
	mgs	NA*	NA*								NA*	NA*	0.22	NA*										NA*	NA*	NA*	NA*
nutlin	affy1	0.69	0.5	0.38	0.16	0.32	0.21	0.32	0.23	0.39	0.23					0.16	0.44	0.57	0.04	0.89	0.27	0.32	0.17				
	snp6	0.7	0.32	0.3	0.21	0.37	NA*	0.5	NA*	NA*	NA*					0.32	NA*	0.7	NA*	0.55	NA*	NA*	0.35				
	mgs	0.43	0.44								0.43	0.32	0.38	0.35										0.6	0.36	0.5	0.35
paclitaxel	affy1	0.59	0.44	0.53	NA*	0.2	NA*	NA*	NA*	NA*	0.32					NA*	0.2	NA*	0.23	NA*	0.36	NA*	NA*				
	snp6	0.36	0.16	0.3	0.2	0.3	0.21	0.29	0.24	NA*	NA*					0.19	NA*	0.46	NA*	NA*	NA*	NA*	NA*				
	mgs	0.35	NA*								NA*	NA*	NA*	NA*										NA*	NA*	0.05	0.15
pd0325901	affy1	0.73	0.58	0.49	0.5	0.54	0.46	0.47	0.51	0.43	0.39					0.68	0.64	0.81	0.65	0.68	0.6	0.53	0.5				
	snp6	0.24	0.48	NA*	NA*	NA*	NA*	0.56	NA*	0.3	NA*					0.74	NA*	0.49	0.08	0.61	0.18	0.88	0.32				
	mgs	0.59	0.46								NA*	0.19	NA*	0.42										0.69	0.43	0.64	0.57
pd032991	affy1	0.8	0.51	0.38	0.26	0.43	0.29	0.55	0.29	0.44	0.31					0.64	0.34	0.7	NA*	0.46	NA*	0.45	0.15				
	snp6	0.54	0.22	0.33	0.17	NA*	0.17	NA*	NA*	NA*	NA*					0.42	NA*	NA*	0.2	0.31	NA*	0.19					
	mgs	0.29	0.27								0.32	0.2	0.2	0.17										0.35	0.13	0.55	0.22
crizotinib	affy1	0.84	0.63	0.37	0.25	0.44	0.3	0.46	0.2	0.38	NA*					0.85	0.47	0.86	0.43	0.76	NA*	0.91	0.13				
	snp6	0.34	0.22	NA*	NA*	NA*	NA*	0.68	NA*	0.28	NA*					NA*	NA*	NA*	0.42	0.66	NA*	0.34	0.14				
	mgs	NA*	0.3								NA*	NA*	0.32	NA*										NA*	NA*	NA*	NA*
pha665752	affy1	0.55	0.39	0.29	0.38	0.4	NA*	NA*	NA*	NA*	0.44	0.32				0.25	NA*	0.38	0.26	0.45	0.33	0.41	NA*				
	snp6	NA*	NA*	0.42	NA*	0.29	NA*	0.28	NA*	NA*	NA*					NA*	0.09	0.8	0.29	NA*	NA*	0.54	NA*				
	mgs	NA*	NA*								NA*	NA*	NA*	NA*										NA*	0.21	0.37	-0.05
plx4720	affy1	0.78	0.42	0.52	0.45	0.52	0.35	0.51	0.31	0.37	0.2					0.79	0.42	0.77	0.56	0.72	0.22	0.74	0.18				
	snp6	0.53	0.29	0.18	NA*	0.47	NA*	NA*	0.18	0.29	0.21					0.72	0.16	0.55	NA*	0.29	NA*	NA*	0.23				
	mgs	0.6	0.3								NA*	NA*	NA*	NA*										0.42	0.27	0.31	0.23
sorafenib	affy1	NA*	0.51	0.25	0.11	NA*	0.16	NA*	NA*	NA*	NA*					NA*	0.17	NA*	0.23	0.18	0.19	NA*	0.17				
	snp6	NA*	NA*	NA*	0.27	0.4	NA*	NA*	NA*	NA*	NA*					0.65	NA*	NA*	0.23	NA*	NA*	0.56	NA*				
	mgs	0.46	NA*								0.21	NA*	0.35	NA*										NA*	0.19	0.42	0.24
nvptae684	affy1	0.63	NA*	0.51	NA*	0.3	0.42	0.42	0.41	0.29	NA*					0.83	0.64	0.8	0.44	0.45	0.35	NA*	NA*				
	snp6	0.68	0.39	NA*	NA*	NA*	NA*	NA*	NA*	NA*	NA*					NA*	0.57	0.63	NA*	0.26	NA*	NA*	NA*				
	mgs	NA*	NA*								0.79	NA*	0.16	NA*									NA*	NA*	NA*	NA*	

Tabla D.23: Modelos para drogas compartidas, P: Garnett, T: Barretina, r_s

Primary (P): Barretina, Transferred (T): Garnett
p-value

		NEA												GNEA																		
		RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference						
		P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T							
tanespimycin	affy1	-713.63	-159.33	-87.91	-36.28	-86.11	-57.25	-20.79	-28.87	-57.98	-17.94					-223.91	-93.49	-323.22	-101.79	-176.71	-74.83	-372.14	-32.97									
	snp6	-189.05	NA*	-30.01	-8.66	-27.16	NA*	-8.55	NA*	NA*	-0.9					NA*	-13.94	-55.92	-0.01	-21.29	NA*	-15.65	NA*									
	mgs	-27.49	NA*							-13.4	-3.25	-19.4	-4.55										-20.05	NA*	-150.06	NA*						
azd0530	affy1	-123.6	-23.29	-47.59	-12.86	-58.32	-18.8	-10.14	NA*	-49.62	-12.75					-278.56	-19.17	-108.68	-19.4	-208.33	-15.43	-188.79	-10.57									
	snp6	-121.01	-32.53	NA*	NA*	NA*	NA*	-60.12	NA*	-50.98	NA*					-20.16	NA*	-309.27	NA*	-139.62	NA*	NA*	NA*	-99.57	NA*	-90.58	-7.09					
	mgs	NA*	NA*							-25.81	-2.8	NA*	NA*											-154.04	NA*	-131.1	NA*					
selumetinib	affy1	-328.79	-81.29	-109.86	-71.44	-132.74	-67.92	-79.82	-41.82	-62.14	-8.91					-340.71	-79.64	-509.23	-92.82	-227.73	-43.19	-162.02	-29.59									
	snp6	-200.45	NA*	NA*	NA*	NA*	NA*	-27.69	NA*	NA*	NA*					-25.89	-0.22	NA*	NA*	-140.86	NA*	-117.29	NA*									
	mgs	-209.67	-56.5							-10.33	NA*	-93.65	NA*											-154.04	NA*	-131.1	NA*					
erlotinib	affy1	-100.77	-56.85	-77.06	-36.72	-106	-35.72	-34.37	-19.82	-47.57	-16.74					-153.48	-65	-165.1	-49.46	-125.03	-84.8	-152.86	-21.64									
	snp6	-149.01	-11.28	-44.52	-10.47	NA*	-4.91	-13.31	-11.11	-20.58	NA*					-109.01	NA*	-122.1	NA*	-120.03	NA*	-97.4	-28.06									
	mgs	NA*	NA*							-23.44	NA*	-19.26	NA*											-37.03	NA*	-46.18	NA*					
lapatinib	affy1	-126.96	-71.39	-65.51	-37.67	-58.57	-39.48	-33.5	-26.61	-30.27	-15.98					-182.43	-69.98	-141.42	-81.74	-146.79	-69.74	-236.61	-61.38									
	snp6	-135.13	-31.75	-14.51	NA*	NA*	NA*	-9.45	NA*	-18.15	-9.59					-37.06	NA*	-39.02	-6.15	-48.23	NA*	-65.52	NA*									
	mgs	-9.71	NA*							NA*	NA*	-12.54	NA*											NA*	NA*	-227.92	-5.81					
nilotinib	affy1	-90.81	-40.65	-32.84	-17.44	-35.16	-18.61	-17.61	-17.74	-25.68	-18.55					-181.19	-30.91	-263.78	-20.37	-33.08	-54.55	-34.71	-29.64									
	snp6	-39.27	-6.26	-15.39	-2.77	-26.55	NA*	-7.99	-9.77	NA*	NA*					NA*	-6.08	-139.32	-21.24	NA*	NA*	NA*	NA*			NA*	NA*	NA*	NA*			
	mgs	NA*	NA*							NA*	NA*	NA*	NA*												-66.61	-20.68	-43.09	-23.89				
nutlin	affy1	-80.38	-36.52	-16.85	-15.94	-23.11	-20.57	-11.77	-13.99	-15.62	-14.74					-178.61	-2.39	-10.9	-31.02	-73.86	-16.34	-52.43	-11.65									
	snp6	-49.81	-26.03	-40.45	NA*	NA*	-10.71	NA*	NA*	-23.7	NA*					-43.81	NA*	NA*	-7.42	NA*	NA*	-197.31	NA*			-66.61	-20.68	-43.09	-23.89			
	mgs	-65.11	-19.06							-15.7	-19.67	-20.71	-19.68																			
paclitaxel	affy1	-297.71	-47.78	-80.9	-29.19	-64.09	-0.81	-28.86	NA*	-61.97	NA*					-294.64	NA*	-169.32	NA*	-510.06	NA*	-164.04	NA*									
	snp6	-99.18	-5.73	-11.33	NA*	-18.58	-5.26	-39.85	NA*	NA*	NA*					NA*	NA*	-24.36	NA*	-32.99	NA*	NA*	NA*			NA*	NA*	-25	NA*			
	mgs	-12.77	-3.57							-46.36	NA*	-11.15	NA*																			
pd0325901	affy1	-267.88	-124.88	-118.49	-77.74	-74.46	-63.26	-111.32	-46.07	-66.07	-25.04					-396.97	-119.12	-302.35	-127.76	-190.57	-96.65	-215.1	-45.25									
	snp6	-188.89	-4.62	NA*	NA*	NA*	NA*	-10.12	-35.02	NA*	-5.1					NA*	NA*	-14.63	-24.66	-14.71	-35.17	-61.65	-25.36			-96.5	-54.33	-157.45	-69.55			
	mgs	-84.27	-50.99							-19.6	NA*	-70.66	NA*																			
pd0332991	affy1	-94.36	-71.78	-67.88	-20.02	-50.86	-28.37	-44.44	-34.82	-54.38	-26.31					-84.52	NA*	-85.99	-30.82	-126.39	-14.59	-118.85	-28.79									
	snp6	-74.24	-35.33	-48.02	NA*	-22.28	NA*	NA*	NA*	NA*	NA*					NA*	NA*	-36.11	NA*	-59.17	NA*	-41.05	NA*									
	mgs	-28.62	-11.03							-50.12	-6.56	-48.35	-6.3														-13.61	-10.69	-24.75	-17.86		
crizotinib	affy1	-216.72	-37.79	-23.37	NA*	-78.92	-10.27	-11.02	NA*	-45.18	NA*					-148.69	-50.21	-138.99	-58.16	-133.2	-34.24	-86.16	-24.24									
	snp6	-44.5	NA*	-29.76	NA*	-15.13	NA*	-14.55	NA*	-18.34	NA*					NA*	NA*	-176.16	NA*	NA*	NA*	-93.07	NA*						-18.11	NA*	NA*	NA*
	mgs	-58.73	NA*							NA*	NA*	NA*	-3.49																			
pha665752	affy1	-110.28	-13.27	-42.01	-4.07	-33.97	-8.3	-15.5	NA*	-72.03	-12.81					-16.72	NA*	-54.35	-3.42	-141.82	-5	-94.74	-8.63									
	snp6	-37.66	NA*	-16.56	NA*	NA*	NA*	-0.13	NA*	NA*	-9.55	NA*	NA*	NA*		-39.77	NA*	-25.54	NA*	NA*	NA*	-12.35	NA*			-197.82	NA*	-9.08	NA*			
	mgs	NA*	NA*																													
plx4720	affy1	-108.69	-60.54	-72.05	-41.36	-58.86	-44.33	-25.02	-39.61	-30.87	-14.03					-102.98	-35.52	-55.27	-55.77	-64.18	-36.87	-81.14	-47.2									
	snp6	-58.59	-36.57	NA*	NA*	NA*	-6.08	-25.58	NA*	-17.31	-5.85					-28.11	-28.45	NA*	-15.98	NA*	NA*	-26.01	NA*			-68.72	-41.07	-83.36	-47.56			
	mgs	-60.56	-46.18							NA*	NA*	-6.51	NA*																			
sorafenib	affy1	-206.35	NA*	-33.99	NA*	-41.19	NA*	-27.46	NA*	-32.08	NA*					-187.59	NA*	-72.28	NA*	-52.95	-2.32	-57.21	NA*									
	snp6	-67.56	NA*	-29.1	NA*	NA*	-10.47	NA*	NA*	NA*	NA*					NA*	-6.1	-107.49	NA*	NA*	NA*	NA*	NA*			-62.49	NA*	-37.67	NA*			
	mgs	-34.66	NA*							NA*	-0.54	-16.05	NA*																			
nvptae684	affy1	-Inf	-26.68	-16.16	-5.04	-66.76	-1.6	-33.14	-10.77	NA*	NA*					-277.88	-37.32	-248.06	-44.73	-307.12	-12.21	NA*	NA*									
	snp6	-56.79	-52.25	-13.12	NA*	NA*	NA*	NA*	NA*	NA*	NA*					-191.53	NA*	-297.97	NA*	-26.41	-14.27	-20.87	NA*			-25.42	NA*	NA*	NA*			
	mgs	NA*	NA*							NA*	NA*	NA*	-2.28																			

Tabla D.24: Modelos para drogas compartidas, P: Barretina, T: Garnett, p-value

Primary (P): Garnett, Transferred (T): Barretina
 p -value

		NEA														GNEA															
		RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference					
		P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T				
tanespimycin	affy1	-272.61	-161.64	-52.1	-66.31	-64.46	-51.01	-60.35	-8.93	-36.95	-33.51					-188.42	-116.58	-236.15	-97.87	-201.93	-37.47	-114.42	-110.43								
	snp6	-8.89	-52.54	-28.65	NA*	-16.04	NA*	NA*	-1.73	-6.81	NA*					-84.32	NA*	-17.29	-14.94	NA*	NA*	NA*	NA*								
	mgs	-18.37	NA*									-13.63	-4.08	-16.73	NA*									NA*	NA*	NA*	-38.36				
azd0530	affy1	-57.2	-29.95	-24.75	-19.16	-41.11	-25.54	NA*	-5.09	-32.85	-11.52					-53.22	-13.8	-24.32	-21.24	-58.02	-26.97	-107.44	NA*								
	snp6	-55.79	-34.55	NA*	NA*	NA*	NA*	NA*	-19.48	NA*	NA*					NA*	NA*	-101.24	NA*	-18.49	-4.85	NA*	NA*								
	mgs	NA*	NA*									-10.58	NA*	NA*	NA*									NA*	-30	-13.5	NA*				
selumetinib	affy1	-81.02	-66.79	-40.44	-42.81	-54.19	-48.36	-58.15	-44.63	-22.8	-21.85					-188.11	-50.33	-194.37	-86.73	-92.12	-62.55	-46.51	-50.79								
	snp6	NA*	-58.04	-13.5	NA*	-28.27	NA*	NA*	NA*	NA*	NA*					-10.4	-4.35	-25.76	NA*	NA*	NA*	NA*	-18.33				NA*	-23.88	NA*	-19.88	
	mgs	-63.26	-33.49									NA*	NA*	NA*	-18.98																
erlotinib	affy1	-63.86	-32.49	-26.93	-20.6	-27.5	-28.24	-42.95	-11.72	-37.73	-26.1					-137.12	-37.14	-55.05	-28.5	-238.39	-12.51	-50.84	NA*								
	snp6	-30.45	-19.86	-29.71	-7.59	-14.81	NA*	-24.32	NA*	-15.26	NA*					-47.22	NA*	NA*	-15.32	NA*	NA*	-63.18	NA*				-77.78	NA*	NA*	NA*	
	mgs	NA*	NA*									NA*	NA*	NA*	NA*																
lapatinib	affy1	-127.53	-31.74	-34.71	-31	-34.33	-21.22	-31.37	-18.99	-33.7	-17.47					-105.23	-42.81	-119.08	-40.06	-124.14	-45.71	-93.33	-49.22								
	snp6	-63.14	-27.03	-35	-4	-20.33	NA*	NA*	NA*	-13.43	-4.54					NA*	-7.37	-19.11	-10.73	-142.88	NA*	NA*	-16.1					-32.76	NA*	-27.18	-9.07
	mgs	-13.81	-1.06									-28.64	NA*	NA*	-4.63																
nilotinib	affy1	-93.01	-47.93	-32.57	-21.74	-37.58	-36.35	-22.86	-14	-56.98	-15.87					-58.81	-48.08	-43.94	-38.77	-265.07	-27.3	-133.6	-31.89								
	snp6	-103.77	NA*	-14.78	-3.69	-26.68	-5.94	-23.14	-0.19	-11.28	NA*					-98.08	NA*	-109.18	NA*	-43.85	NA*	NA*	NA*					NA*	NA*	NA*	NA*
	mgs	NA*	NA*									NA*	NA*	-12.91	NA*																
nutlin	affy1	-146.18	-75.99	-32.47	-5.71	-25.04	-13.09	-24.93	-13.23	-34.06	-13.45					-13.58	-49.09	-72.27	-1.17	-288.58	-19.94	-29.28	-11.37								
	snp6	-126.6	-23.61	-23.51	-12.24	-29.79	NA*	-54.49	NA*	NA*	NA*					-22.22	NA*	-127.31	NA*	-66.92	NA*	NA*	-20.6					-85.48	-36.34	-57.22	-36.76
	mgs	-42.67	-42.86									-43.61	-23.92	-38.91	-31.95																
paclitaxel	affy1	-71.47	-19.57	-39.64	NA*	-5.93	NA*	NA*	NA*	NA*	-13.69					NA*	-5.87	NA*	-5.22	NA*	-14.94	NA*	NA*								
	snp6	-18.41	-4.95	-13.62	-3.73	-12.93	-5.49	-16.54	-4.81	NA*	NA*					-31.75	NA*	-58.5	NA*	NA*	NA*	NA*	NA*					NA*	NA*	-13.72	-4.07
	mgs	-35.62	NA*									NA*	NA*	NA*	NA*																
pd0325901	affy1	-153.12	-86.82	-62.85	-62.48	-77.85	-49.69	-63.91	-70.93	-50.3	-40.92					-135.14	-108.76	-228.34	-116.83	-154.02	-89.6	-91.46	-74.34								
	snp6	-17.59	-61.8	NA*	NA*	NA*	NA*	-61.61	NA*	-16.2	NA*					-190.27	NA*	-53.89	-2.34	-95.41	-4.4	-216.18	-23.3					-117.97	-32.29	-87.66	-60.13
	mgs	-80.95	-34.97									NA*	-7.82	NA*	-32.93																
pd0332991	affy1	-245.34	-69.83	-24.99	-11.05	-34.98	-16.56	-77.48	-19.55	-50.04	-24.07					-137.55	-21.8	-150.28	NA*	-87.61	NA*	-87.78	-15.2								
	snp6	-106.14	-12.19	-25.92	-5.34	NA*	-10.28	NA*	NA*	NA*	NA*					-45.63	NA*	NA*	NA*	-18.14	-18.92	NA*	-10.48					-33.77	-18.23	-78.61	-14.39
	mgs	-22.96	-19.36									-26.45	-10.89	-10.62	-6.32																
crizotinib	affy1	-130.32	-55.98	-15.17	-7.03	-19.67	-10.63	-26.29	-4.38	-19.97	NA*					-136.03	-31.32	-147.27	-23.72	-95.7	NA*	-174.99	-11.08								
	snp6	-12.9	-6.93	NA*	NA*	NA*	NA*	-64.54	NA*	-10.8	NA*					NA*	NA*	NA*	-23.08	-53.5	NA*	-22.92	-7.47					NA*	NA*	NA*	NA*
	mgs	NA*	-10.46									NA*	NA*	-11.49	NA*																
pha665752	affy1	-48.61	-20.25	-9.44	-19.65	-19.71	NA*	NA*	NA*	-25.91	-11.3					-16.91	NA*	-23.85	-7.26	-25.7	-11.98	-17.78	NA*								
	snp6	NA*	NA*	-19.87	NA*	-10.19	NA*	-7.59	NA*	NA*	NA*					NA*	-2.55	-113.87	-9.14	NA*	NA*	-37.73	NA*					NA*	-5.12	-16.97	-0.18
	mgs	NA*	NA*									NA*	NA*	NA*	NA*																
plx4720	affy1	-175.85	-53.46	-68.44	-52.41	-68.52	-26.75	-64.05	-24.18	-34	-14.42					-190.87	-51.73	-167.37	-80.43	-137.34	-29.71	-138.81	-21.63								
	snp6	-74.79	-27.85	-8.74	NA*	-52.24	NA*	NA*	-7.36	-13.62	-6.9					-125.32	-7.96	-63.08	NA*	-18.61	NA*	NA*	-7.68					-39.31	-26.38	-32.16	-26.93
	mgs	-73.17	-23.07									NA*	NA*	NA*	NA*																
sorafenib	affy1	NA*	-34.5	-10.73	-3.52	NA*	-4.06	NA*	NA*	NA*	NA*					NA*	-9.61	NA*	-9.66	-9.56	-6.8	NA*	-6.28					NA*	-5.58	-21.85	-8.13
	snp6	NA*	NA*	NA*	NA*	-8.79	-19.73	NA*	NA*	NA*	NA*					-64.24	NA*	NA*	-9.41	NA*	NA*	-26.22	NA*					NA*			
	mgs	-23.07	NA*									-7.56	NA*	-14.99	NA*																
nvptae684	affy1	-70.92	NA*	-39.34	NA*	-10.09	-20.69	-24.01	-16.9	-10.26	NA*					-158.17	-59.53	-122.85	-20.99	-52.43	-16.55	NA*	NA*								
	snp6	-79.07	-17.51	NA*	NA*	NA*	NA*	NA*	NA*	NA*	NA*					NA*	-41.96	-77.86	NA*	-21.07	NA*	NA*	NA*					NA*	NA*	NA*	NA*
	mgs	NA*	NA*									-95.75	NA*	-5.44	NA*																

Tabla D.25: Modelos para drogas compartidas, P: Garnett, T: Barretina, p -value

Primary (P): Barretina, Transferred (T): Garnett
MSE

		NEA												GNEA													
		RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference	
		P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T		
tanespimycin	affy1	0.07	0.08	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.11					0.1	0.09	0.1	0.1	0.1	0.1	0.11	0.1				
	snp6	0.11	0.11*	0.11	0.11	0.11	0.11*	0.11	NA*	0.11*	NA					0.11*	0.11	0.11	NA	0.11	0.11	NA*					
	mgs	0.11	0.11*									0.11	0.11	0.11	0.11									0.11	NA*	0.11	NA*
azd0530	affy1	0.73	0.76	0.76	0.78	0.76	0.78	0.79	NA*	0.76	0.79					0.76	0.76	0.71	0.76	0.77	0.79	0.76	0.79				
	snp6	0.78	0.75	0.8*	NA*	0.79*	NA*	0.82	NA*	0.87	NA*					0.8	NA*	0.78	0.8*	0.83	0.83*	0.88*	NA*				
	mgs	0.79*	NA*									0.79	0.79	0.79*	NA*									0.77	NA*	0.77	0.77
selumetinib	affy1	0.68	0.77	0.82	0.82	0.87	0.83	0.94	0.93	0.93	0.96					0.77	0.86	0.71	0.81	0.9	0.89	0.9	0.92				
	snp6	0.86	NA*	0.98*	0.98*	0.98*	0.98*	0.94	NA*	0.98*	NA*					0.98	NA	0.98*	0.98*	0.94	NA*	0.96	NA*				
	mgs	0.8	0.8									0.96	NA*	0.9	NA*									0.8	NA*	0.79	NA*
erlotinib	affy1	0.67	0.71	0.74	0.75	0.73	0.76	0.81	0.81	0.8	0.82					0.71	0.76	0.69	0.73	0.77	0.75	0.79	0.83				
	snp6	0.78	0.83	0.84	0.84	0.85*	0.85	0.85	0.86	0.78	0.81*					0.83	0.86*	0.85	NA*	0.84	NA*	0.79	0.77				
	mgs	0.86*	NA*									0.86	NA*	0.85	NA*									0.86	0.86*	0.85	NA*
lapatinib	affy1	0.56	0.61	0.64	0.65	0.65	0.65	0.7	0.69	0.71	0.72					0.57	0.62	0.58	0.6	0.66	0.65	0.61	0.64				
	snp6	0.61	0.69	0.73	0.74*	0.73*	0.73*	0.75	NA*	0.71	0.7					0.73	NA*	0.73	0.73	0.73	0.76*	0.67	NA*				
	mgs	0.71	0.72*									0.72*	0.72*	0.71	NA*									0.72*	0.72*	0.7	0.71
nilotinib	affy1	0.57	0.56	0.6	0.6	0.6	0.61	0.62	0.59	0.6	0.6					0.57	0.61	0.58	0.6	0.6	0.59	0.59	0.58				
	snp6	0.61	0.62	0.62	0.62	0.6	0.62*	0.6	0.59	0.55*	0.55*					0.63*	0.62	0.6	0.62	0.6*	0.6*	0.56*	NA*				
	mgs	0.59*	NA*									0.59*	NA*	0.6*	0.59*									0.59*	NA*	0.59*	NA*
nutlin	affy1	0.5	0.48	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.51					0.53	NA	0.53	0.51	0.52	0.53	0.52	0.51				
	snp6	0.52	0.52	0.53	0.53*	0.53*	0.53	0.49*	0.49*	0.5	NA*					0.51	0.52*	0.52*	0.51	0.48*	0.48*	0.49	NA*				
	mgs	0.51	0.49									0.52	0.51	0.5	0.5									0.52	0.5	0.51	0.49
paclitaxel	affy1	0.06	0.08	0.08	0.08	0.08	NA	0.08	NA*	0.08	NA*					0.07	NA*	0.08	NA*	0.07	NA*	0.08	NA*				
	snp6	0.08	0.08	0.08	0.08*	0.08	0.08	0.08	0.09*	0.09*	NA*					0.09*	0.08*	0.08	0.08*	0.09	NA*	0.09*	NA*				
	mgs	0.08	0.08									0.08	NA*	0.08	NA*									0.08*	NA*	0.08	0.08*
pd0325901	affy1	0.78	0.77	0.91	0.9	0.94	0.94	0.96	1	1.01	1.05					0.9	0.85	0.83	0.85	0.95	0.9	1	1.01				
	snp6	1.04	1.07	1.07*	NA*	1.08*	NA*	1.1	1.09	1.16*	1.17					1.1*	1.09*	1.08	1.04	1.12	1.08	1.18	1.16				
	mgs	0.99	0.95									1.09	NA*	1.03	NA*									0.96	0.98	0.92	0.88
pd0332991	affy1	0.82	0.84	0.83	0.87	0.8	0.84	0.87	0.87	0.89	0.88					0.85	0.92*	0.84	0.86	0.85	0.9	0.87	0.87				
	snp6	0.88	0.88	0.89	0.92*	0.92	NA*	0.91*	NA*	0.92*	NA*					0.92*	0.91*	0.9	NA*	0.89	0.89*	0.89	NA*				
	mgs	0.92	0.89									0.93	0.92	0.93	0.92									0.93	0.91	0.91	0.91
crizotinib	affy1	0.42	0.44	0.45	0.46*	0.43	0.45	0.46	0.47*	0.46	0.46*					0.43	0.44	0.45	0.42	0.44	0.46	0.45	0.45				
	snp6	0.46	0.46*	0.46	NA*	0.46	NA*	0.46	0.46*	0.47	0.47*					0.47*	NA*	0.46	NA*	0.47*	0.47*	0.47	0.48*				
	mgs	0.44	NA*									0.46*	NA*	0.46*	0.46									0.46	NA*	0.46*	NA*
pha665752	affy1	0.75	0.78	0.73	0.78	0.75	0.78	0.77	NA*	0.76	0.78					0.78	0.79*	0.77	0.78	0.76	0.78	0.77	0.78				
	snp6	0.78	NA*	0.78	0.79*	0.79*	0.79*	0.79*	NA	0.73*	NA*					0.78	NA*	0.78	0.79*	0.78*	NA*	0.72	0.73*				
	mgs	0.79*	NA*									0.78	NA*	0.78*	NA*									0.78	NA*	0.78	0.79*
plx4720	affy1	0.69	0.68	0.72	0.72	0.75	0.72	0.77	0.76	0.76	0.77					0.71	0.72	0.66	0.7	0.73	0.71	0.73	0.75				
	snp6	0.77	0.74	0.78*	0.79*	0.78*	0.78	0.77	NA*	0.77	0.78					0.78	0.75	0.78*	0.77	0.78*	0.79*	0.79	NA*				
	mgs	0.69	0.68									0.79*	NA*	0.79	NA*									0.71	0.69	0.66	0.66
sorafenib	affy1	0.67	NA*	0.71	0.73*	0.67	NA*	0.71	NA*	0.72	NA*					0.69	NA*	0.7	NA*	0.69	NA	0.69	NA*				
	snp6	0.69	NA*	0.71	NA*	0.72*	0.7	0.71*	NA*	0.7*	NA*					0.72*	0.72	0.72	NA*	0.72*	NA*	0.7*	0.7*				
	mgs	0.67	0.69*									0.7*	NA	0.68	0.69*									0.68	NA*	0.69	0.7*
nvptae684	affy1	0.43	0.52	0.53	0.53	0.5	0.53	0.52	0.53	0.5*	0.5*					0.5	0.52	0.5	0.51	0.52	0.53	0.5*	NA*				
	snp6	0.5	0.48	0.5	NA*	0.51*	NA*	0.49*	NA*	0.51*	NA*					0.46	NA*	0.5	0.52*	0.49	0.48	0.51	NA*				
	mgs	0.5*	NA*									0.5*	0.5*	0.5*	NA									0.5	NA*	0.5*	NA*

Tabla D.26: Modelos para drogas compartidas, P: Barretina, T: Garnett, *MSE*

Primary (P): Garnett, Transferred (T): Barretina
MSE

		NEA														GNEA														
		RAW		top.200		top.400		signif		signif.filt		total		reference		top.200		top.400		signif		signif.filt		total		reference				
		P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T			
tanespimycin	affy1	0.9	0.86	1.15	1.07	1.16	1.12	1.11	1.22	1.18	1.17					1.07	0.98	1.16	1.07	1.11	1.2	1.15	1.07							
	snp6	1.25	1.22	1.24	1.24*	1.24	1.25*	1.17*	NA	1.1	NA*					1.23	NA*	1.23	1.23	1.2*	1.19*	1.1*	1.09*							
	mgs	1.2	1.21*									1.2	1.21	1.21	1.21*										1.21*	1.21*	1.21*	1.17		
azd0530	affy1	2.59	2.67	2.81	2.83	2.82	2.92	3*	NA	2.65	2.87					2.88	2.79	2.87	2.79	2.93	2.88	2.76	3*							
	snp6	2.67	2.6	2.95*	NA*	2.94*	NA*	2.95*	2.9	2.85*	2.83*					2.99*	2.99*	2.88	2.97*	2.86	2.94	2.89*	NA*							
	mgs	2.83*	NA*									2.79	2.85*	2.91*	NA*										2.83*	2.41	2.87	2.85*		
selumetinib	affy1	1.8	1.75	1.87	1.83	1.86	1.78	1.86	1.77	1.93	1.92					1.72	1.84	1.76	1.88	1.9	1.81	1.88	1.8							
	snp6	1.96*	1.84	1.97	NA*	1.95	NA*	1.96*	1.98*	1.96*	NA*					1.99	1.98	1.97	NA*	2.02*	2.01*	1.94*	1.9							
	mgs	1.92	1.84									1.96*	1.97*	1.96*	1.93										1.96*	1.94	1.96*	1.89		
erlotinib	affy1	2.29	2.34	2.53	2.55	2.55	2.55	2.65	2.7	2.49	2.36					2.35	2.67	2.44	2.45	2.44	2.69	2.56	3.05*							
	snp6	2.65	2.65	2.72	2.75	2.72	NA*	2.82	2.85*	3	3.03*					2.76	2.78*	2.82*	2.73	2.86*	2.86*	2.98	2.98*							
	mgs	2.78*	NA*									2.82*	2.84*	2.85*	2.81*										2.7	2.84*	2.82*	2.87*		
lapatinib	affy1	1.94	2.2	2.19	2.23	2.24	2.38	2.33	2.37	2.25	2.28					2.1	2.21	2.21	2.06	2.27	2.08	2.31	2.19							
	snp6	2.42	2.39	2.55	2.56	2.47	NA*	2.61*	2.64*	2.84	2.83					2.62*	2.58	2.58	2.54	2.54	2.65*	2.85*	2.79			2.5	NA*	2.46	2.5	
	mgs	2.5	NA									2.49	NA*	2.54*	2.49															
nilotinib	affy1	1.16	1.17	1.27	1.27	1.26	1.24	1.35	1.32	1.33	1.34					1.29	1.29	1.32	1.29	1.2	1.27	1.26	1.24							
	snp6	1.29	1.36*	1.34	1.36	1.28	1.34	1.28	NA	1.23	NA*					1.29	NA*	1.37	1.38*	1.29	NA*	1.25*	NA*							
	mgs	1.32*	NA*									1.33*	NA*	1.3	NA*										1.33*	NA*	1.32*	NA*		
nutlin	affy1	1.41	1.47	1.91	1.96	1.97	1.95	1.91	1.91	1.93	1.92					1.98	1.85	1.93	NA	1.84	1.9	1.97	2.01							
	snp6	1.78	1.9	1.92	1.91	1.94	NA*	1.92	NA*	1.75*	1.77*					1.93	2*	1.89	NA*	1.9	NA*	1.75*	1.75							
	mgs	1.65	1.59									1.77	1.74	1.69	1.66										1.62	1.71	1.7	1.64		
paclitaxel	affy1	2.26	2.14	2.19	2.28*	2.28	2.3*	2.31*	2.31*	2.3*	2.23					2.34*	2.28	2.3*	2.27	2.3*	2.24	2.3*	2.3*							
	snp6	2.25	2.3	2.23	2.29	2.27	2.26	2.08	2.09	1.79*	NA*					2.3	NA*	2.29	2.3*	2.11*	2.13*	1.82*	NA*							
	mgs	2.02	2.14*									2.15*	2.16*	2.16*	2.12*										2.13*	NA*	2.15	2.16		
pd0325901	affy1	1.37	1.3	1.55	1.5	1.48	1.48	1.57	1.56	1.52	1.49					1.45	1.34	1.45	1.36	1.5	1.34	1.56	1.52							
	snp6	1.71	1.61	1.73*	NA*	1.73*	NA*	1.65	1.71*	1.69	NA*					1.63	NA*	1.72	NA	1.66	NA	1.63	1.57				1.69	1.6	1.69	1.51
	mgs	1.65	1.61									1.76*	1.73	1.74*	1.71															
pd0332991	affy1	1.16	1.25	1.54	1.52	1.48	1.51	1.43	1.48	1.5	1.56					1.46	1.51	1.4	1.56*	1.49	1.55*	1.47	1.54							
	snp6	1.38	1.55	1.57	1.54	1.55*	1.55	1.6*	NA*	1.58*	NA*					1.54	NA*	1.54*	1.54*	1.55	1.5	1.57*	1.52							
	mgs	1.49	1.44									1.55	1.56	1.55	1.55										1.52	1.44	1.47	1.47		
crizotinib	affy1	1.75	1.69	1.85	1.9	1.81	1.86	1.71	1.86	1.76	1.92*					1.66	1.65	1.69	1.75	1.77	1.92*	1.8	1.87							
	snp6	1.87	1.86	1.94*	1.93*	1.9*	1.91*	1.74	1.85*	1.87	1.89*					1.91*	NA*	1.89*	1.8	1.82	NA*	1.83	1.87							
	mgs	1.95*	1.92									1.98*	NA*	1.95	NA*											1.94*	1.95*	1.95*	NA*	
pha665752	affy1	1.67	1.66	1.66	1.63	1.63	1.69*	1.7*	1.7*	1.6	1.64					1.65	1.66*	1.64	1.62	1.63	1.65	1.63	1.67*							
	snp6	1.68*	1.66*	1.65	1.66*	1.66	NA*	1.56	NA*	1.55*	NA*					1.67*	1.65	1.65	1.61	1.6*	NA*	1.52	1.53*							
	mgs	1.57*	NA*									1.56*	1.58*	1.56*	NA*										1.56*	1.54	1.53	NA		
plx4720	affy1	1.58	1.56	1.62	1.58	1.62	1.76	1.75	1.69	1.73	1.8					1.52	1.56	1.63	1.64	1.6	1.66	1.62	1.74							
	snp6	1.75	1.72	1.82	NA*	1.72	NA*	1.83*	1.81	1.69	1.67					1.77	1.79	1.8	NA*	1.79	NA*	1.7*	1.67				1.71	1.65	1.64	1.61
	mgs	1.65	1.63									1.81*	NA*	1.8*	1.79*															
sorafenib	affy1	1.79*	1.76	1.78	1.79	1.8*	1.79	1.8*	1.83*	1.81*	1.81*					1.81*	1.72	1.81*	1.75	1.79	1.76	1.81*	1.78							
	snp6	1.83*	1.8*	1.8*	1.78	1.78	NA*	1.75*	NA*	1.57*	NA*					1.78	NA*	1.79*	1.75	1.76*	NA*	1.57	NA*							
	mgs	1.75	1.74*									1.74	NA*	1.67	1.78*											1.75*	1.72	1.75	1.73	
nvptae684	affy1	1.97	2.06*	1.93	2.04*	2.02	2.03	2.05	1.92	2.04	NA*					1.87	2.01	1.93	1.99	1.98	2	2.03*	NA*							
	snp6	1.97	1.95	2.03*	2.04*	2.07*	NA*	1.88*	NA*	2.3*	NA*					2.05*	2	2.03	2.07*	1.86	1.9*	2.3*	2.32*							
	mgs	2.1*	NA*									2.09	NA*	2.1	NA*											2.14*	2.13*	2.15*	NA*	

Tabla D.27: Modelos para drogas compartidas, P: Garnett, T: Barretina, MSE

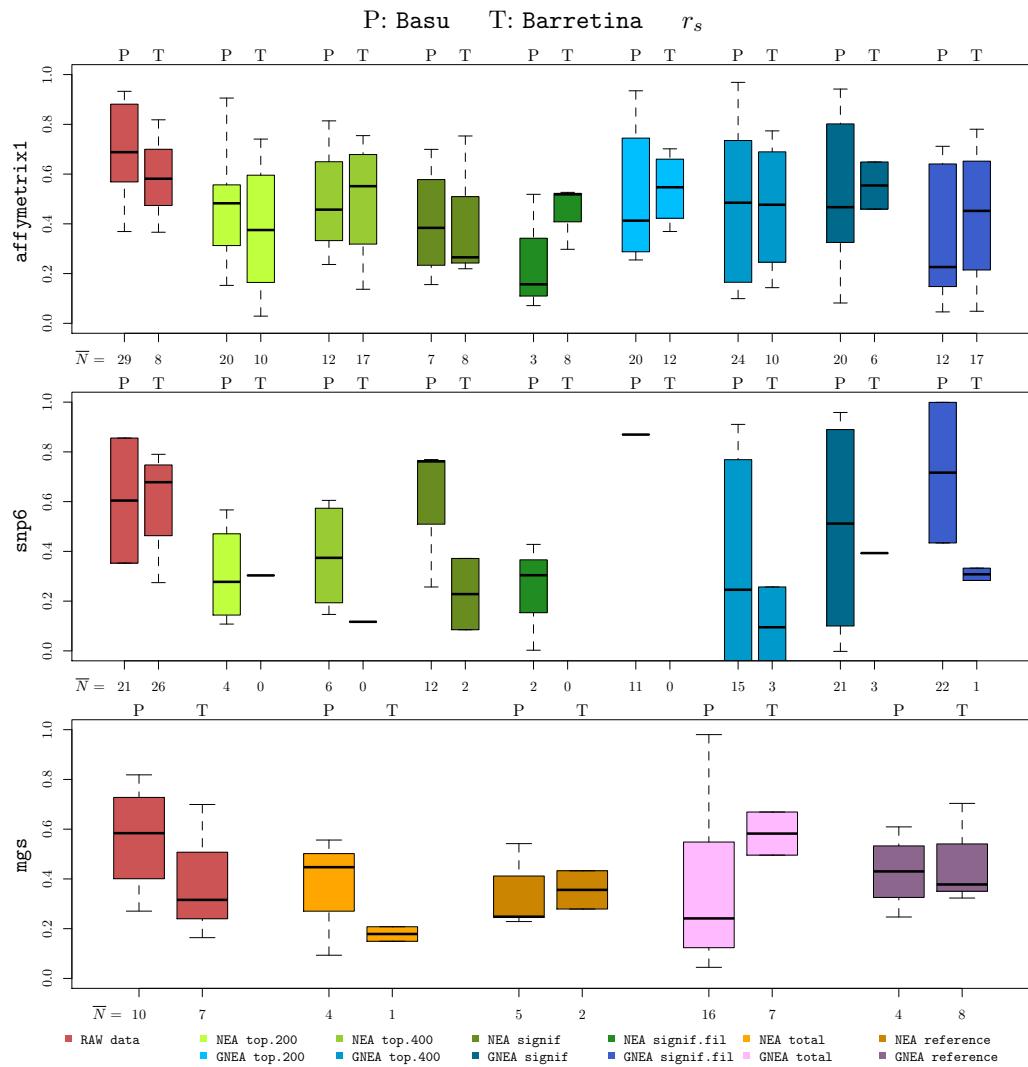


Figura D.1: Modelos para drogas compartidas,
P: Basu, T: Barretina, r_s

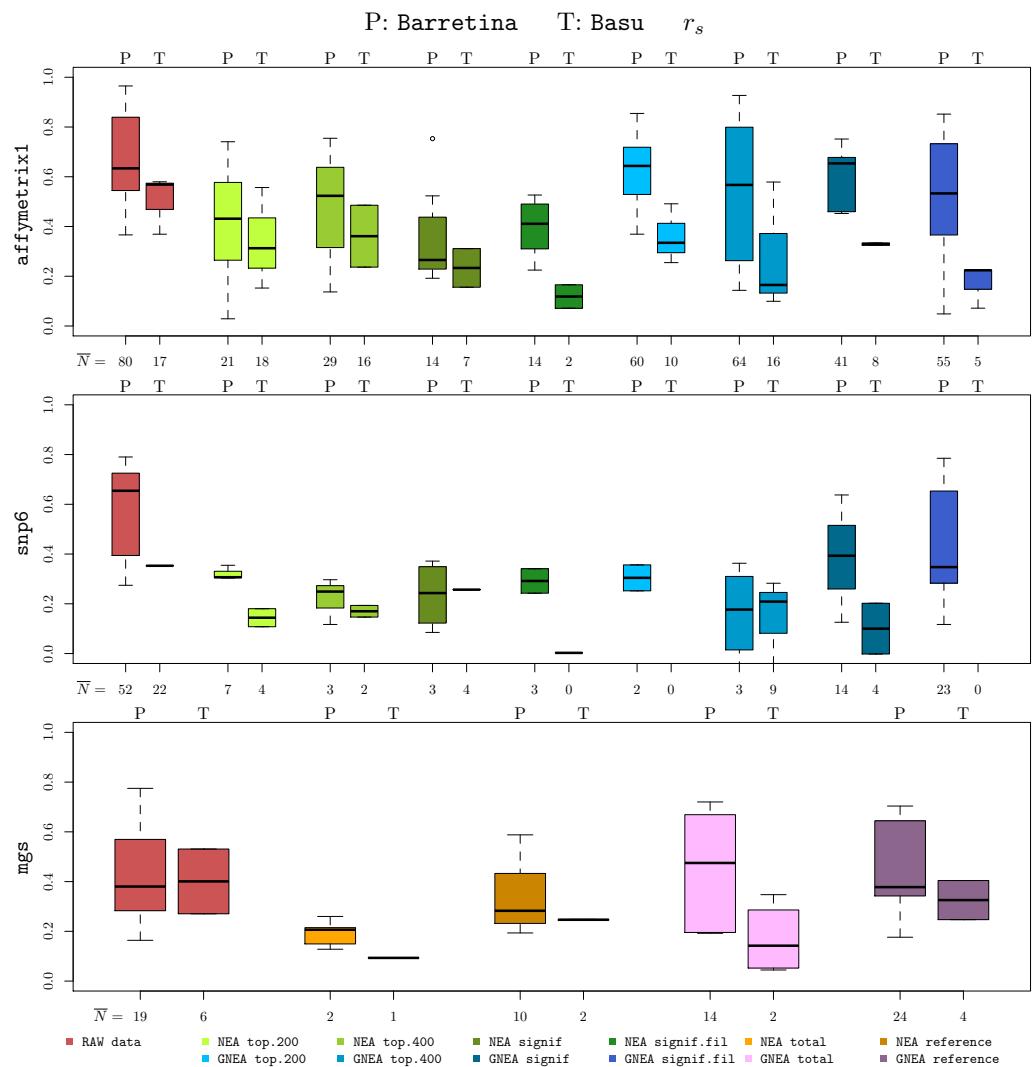


Figura D.2: Modelos para drogas compartidas,
P: Barretina, T: Basu, r_s

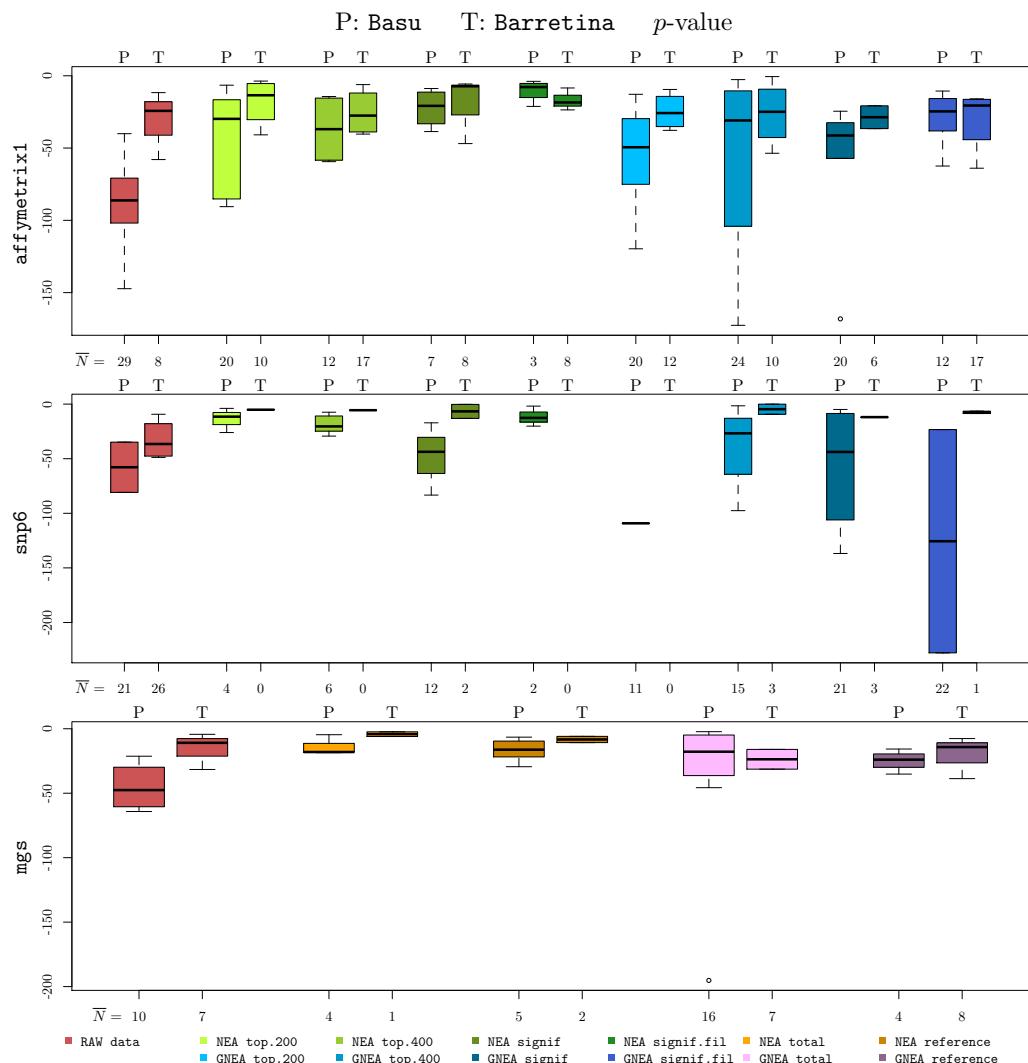


Figura D.3: Modelos para drogas compartidas,
P: Basu, T: Barretina, *p*-value

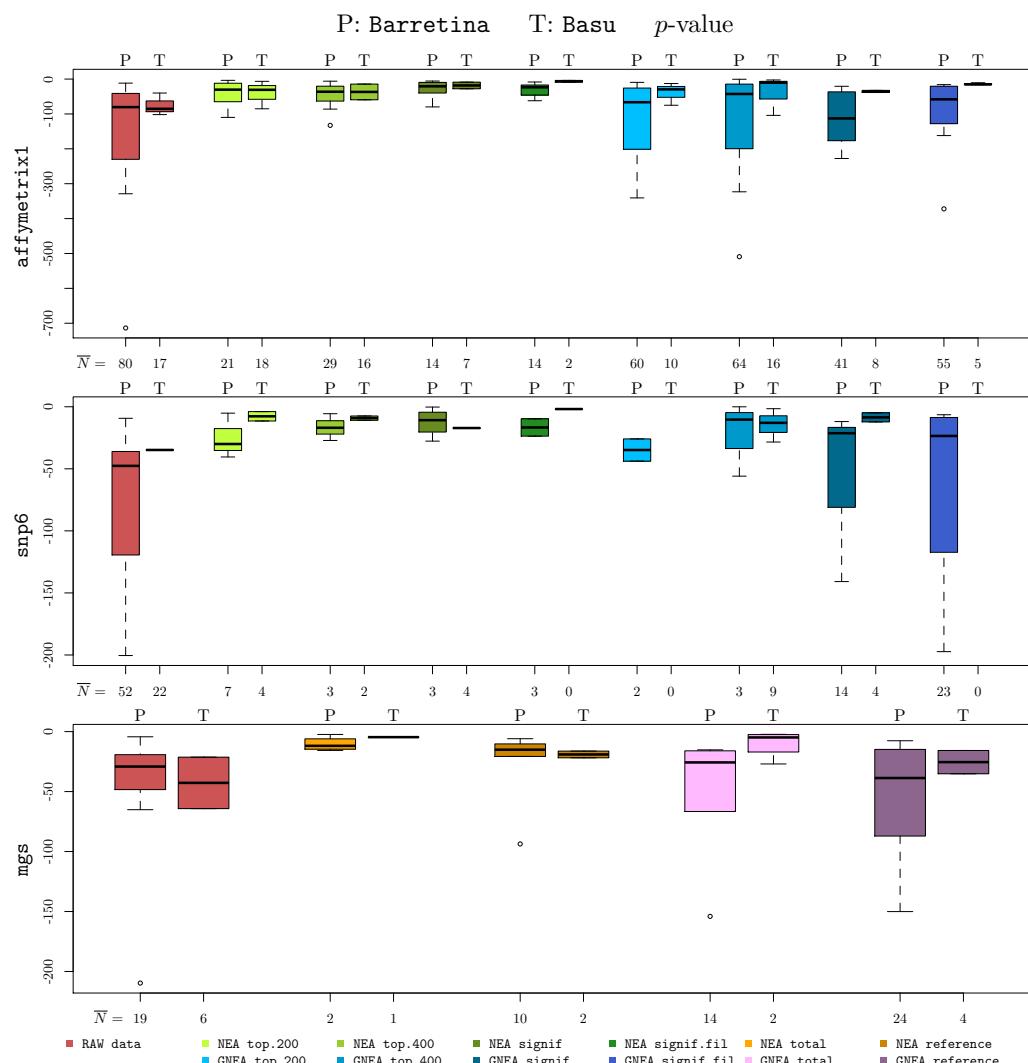


Figura D.4: Modelos para drogas compartidas,
P: Barretina, T: Basu, *p*-value

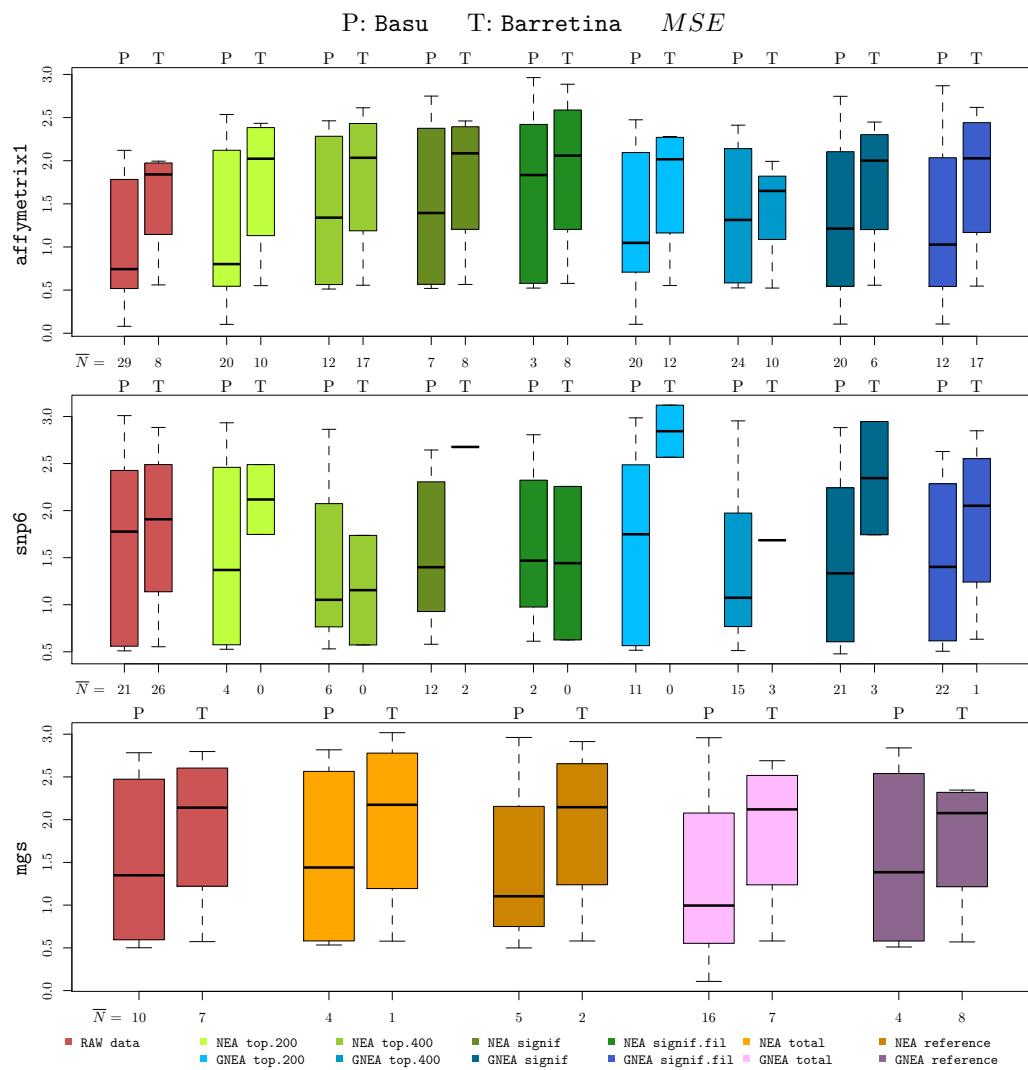


Figura D.5: Modelos para drogas compartidas,
P: Basu, T: Barretina, *MSE*

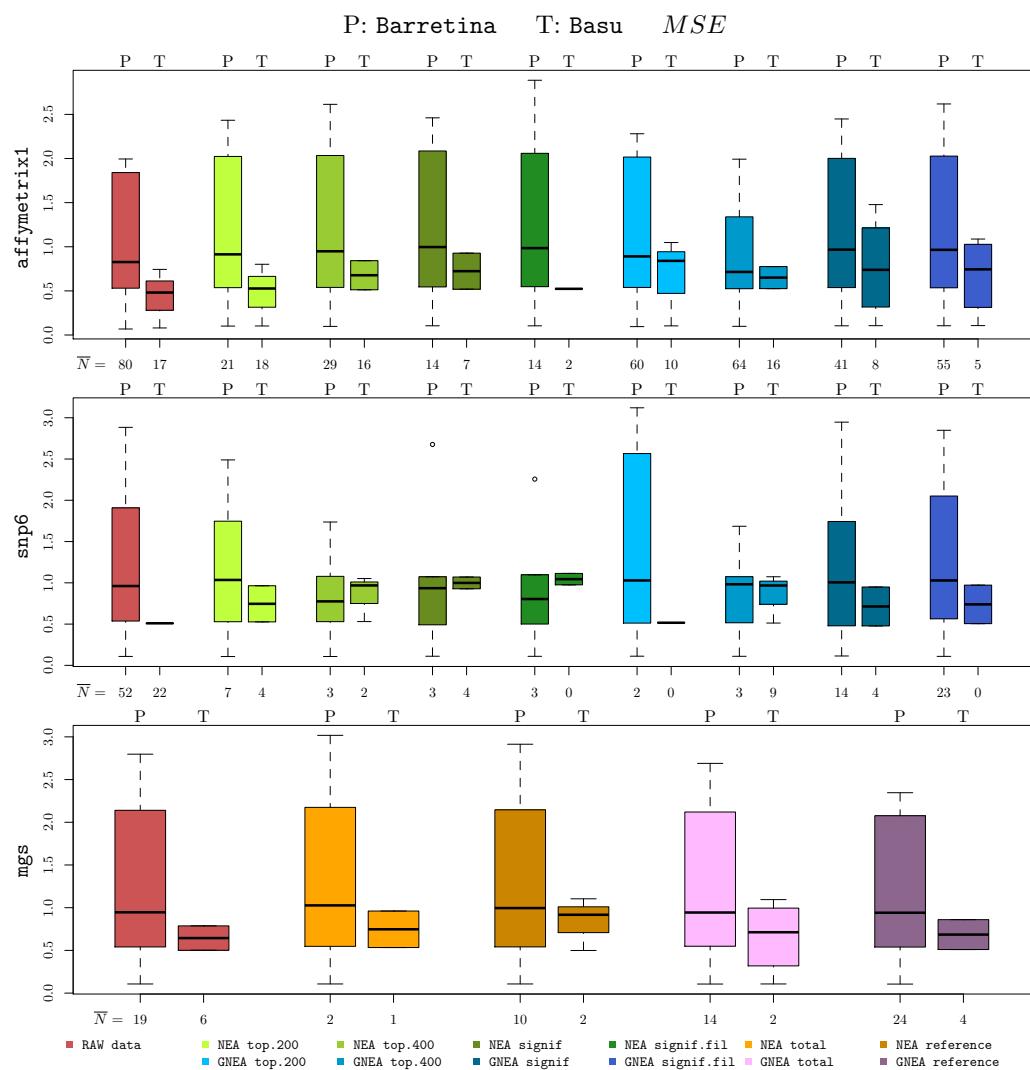


Figura D.6: Modelos para drogas compartidas,
P: Barretina, T: Basu, *MSE*

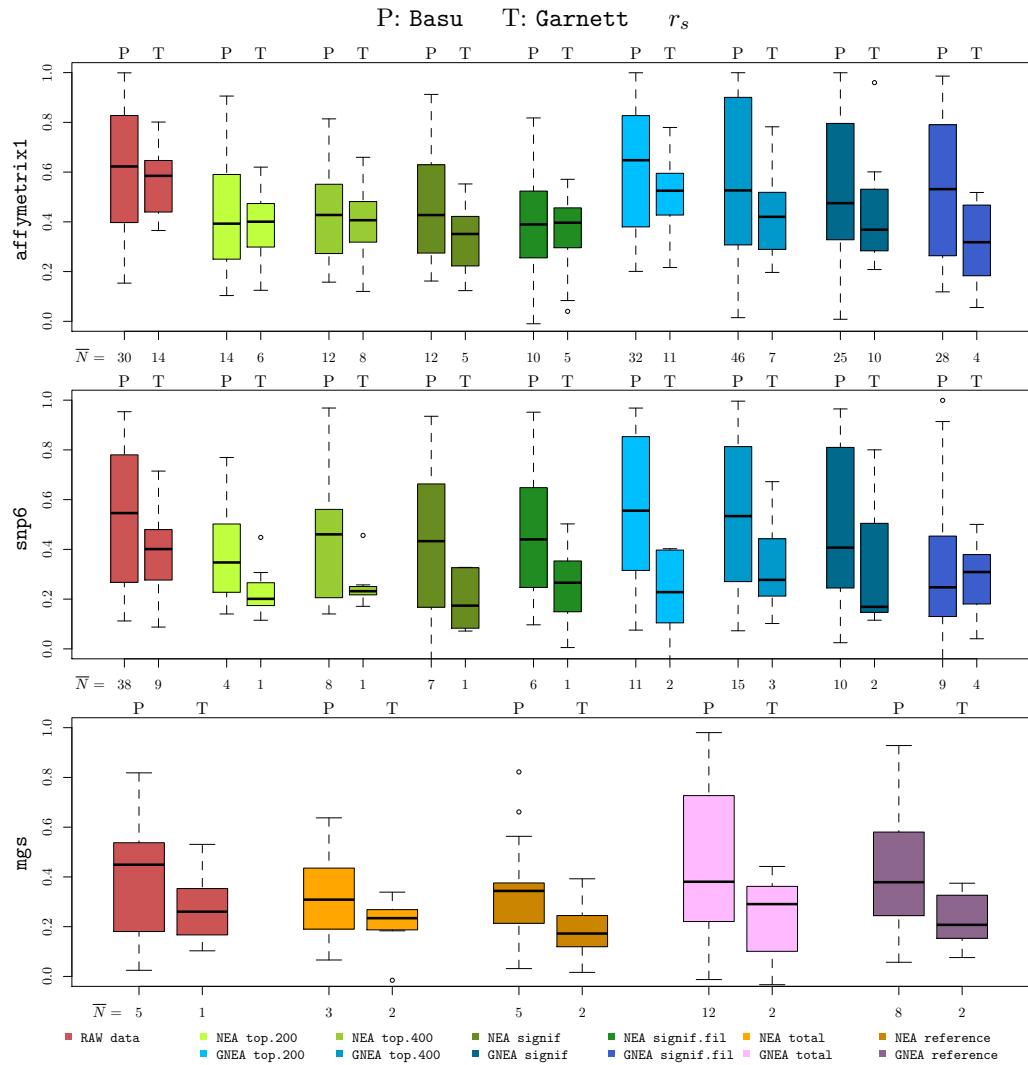


Figura D.7: Modelos para drogas compartidas,
P: Basu, T: Garnett, r_s

103

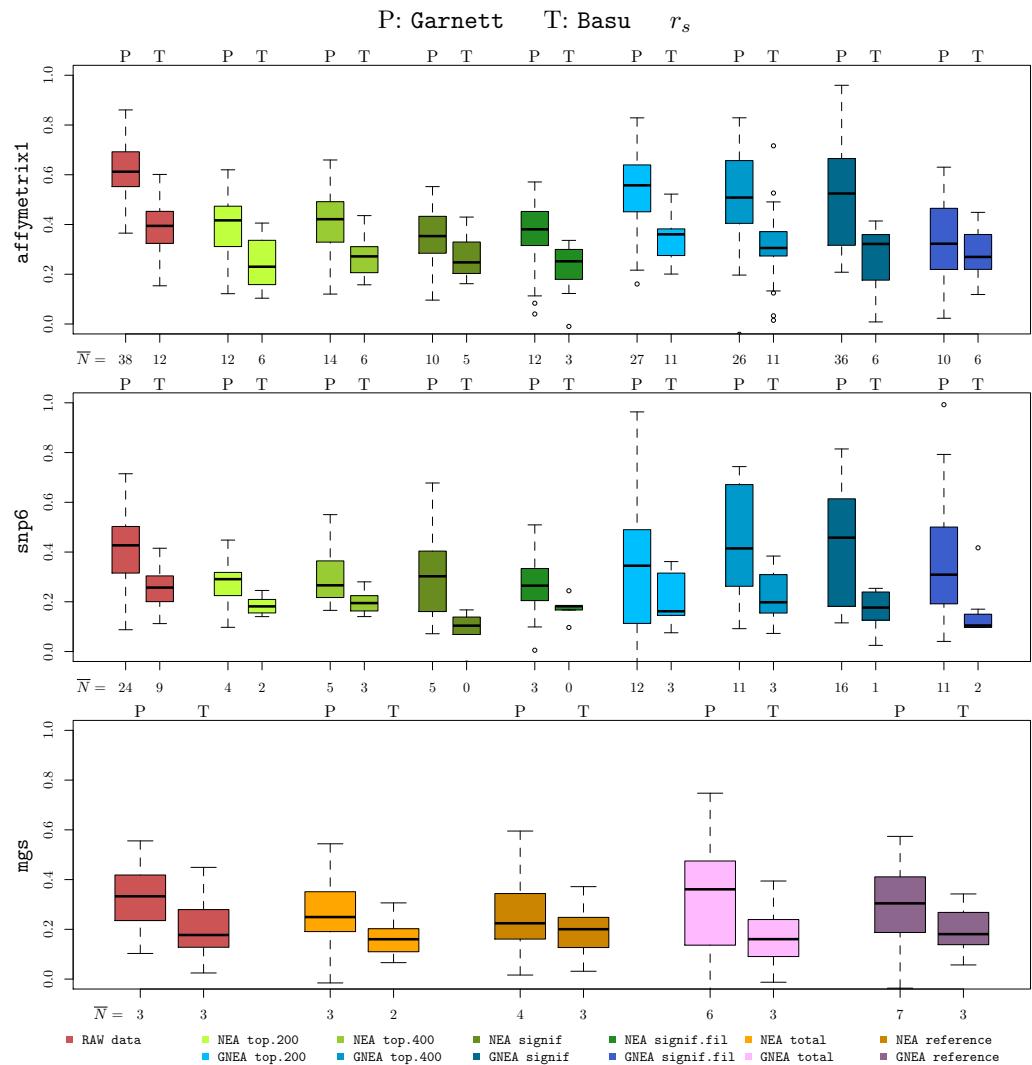


Figura D.8: Modelos para drogas compartidas,
P: Garnett, T: Basu, r_s

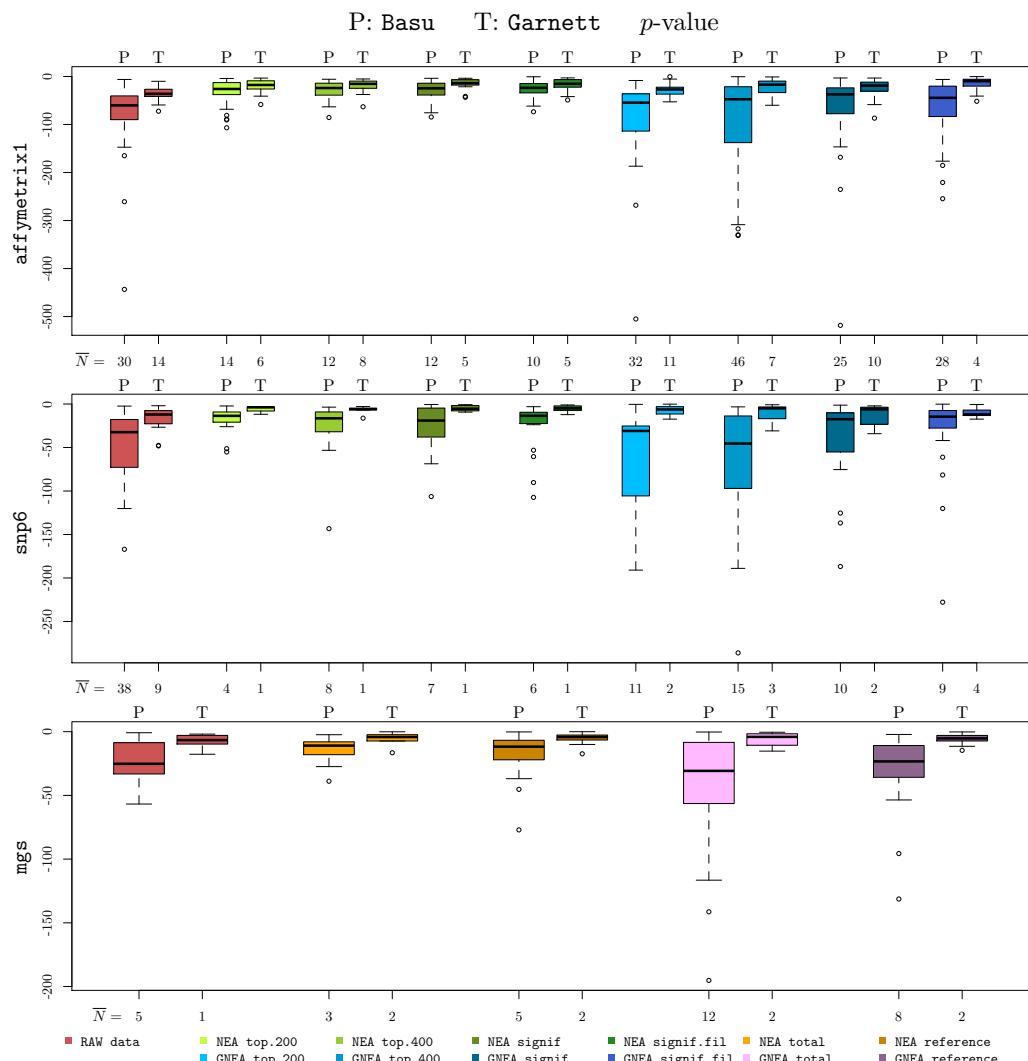


Figura D.9: Modelos para drogas compartidas,
P: Basu, T: Garnett, p -value

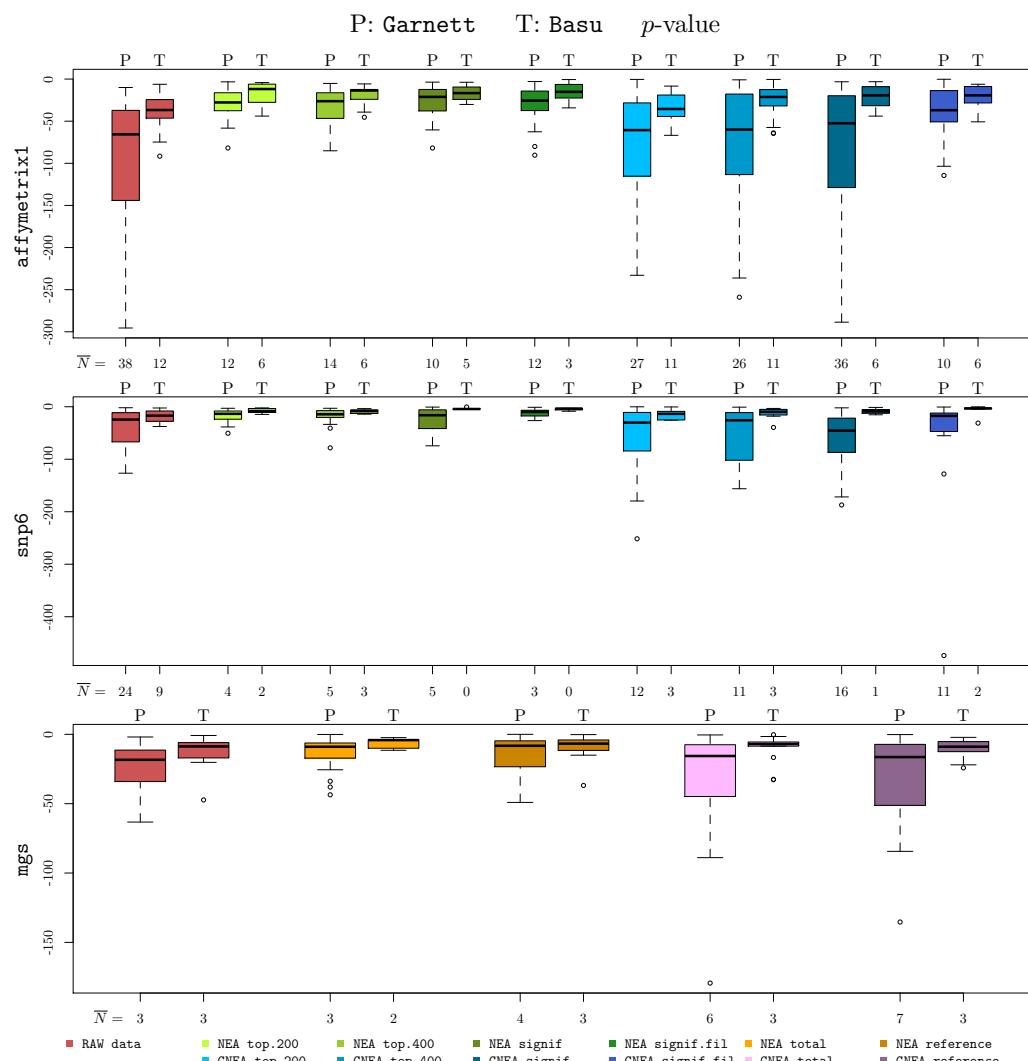


Figura D.10: Modelos para drogas compartidas,
P: Garnett, T: Basu, p -value

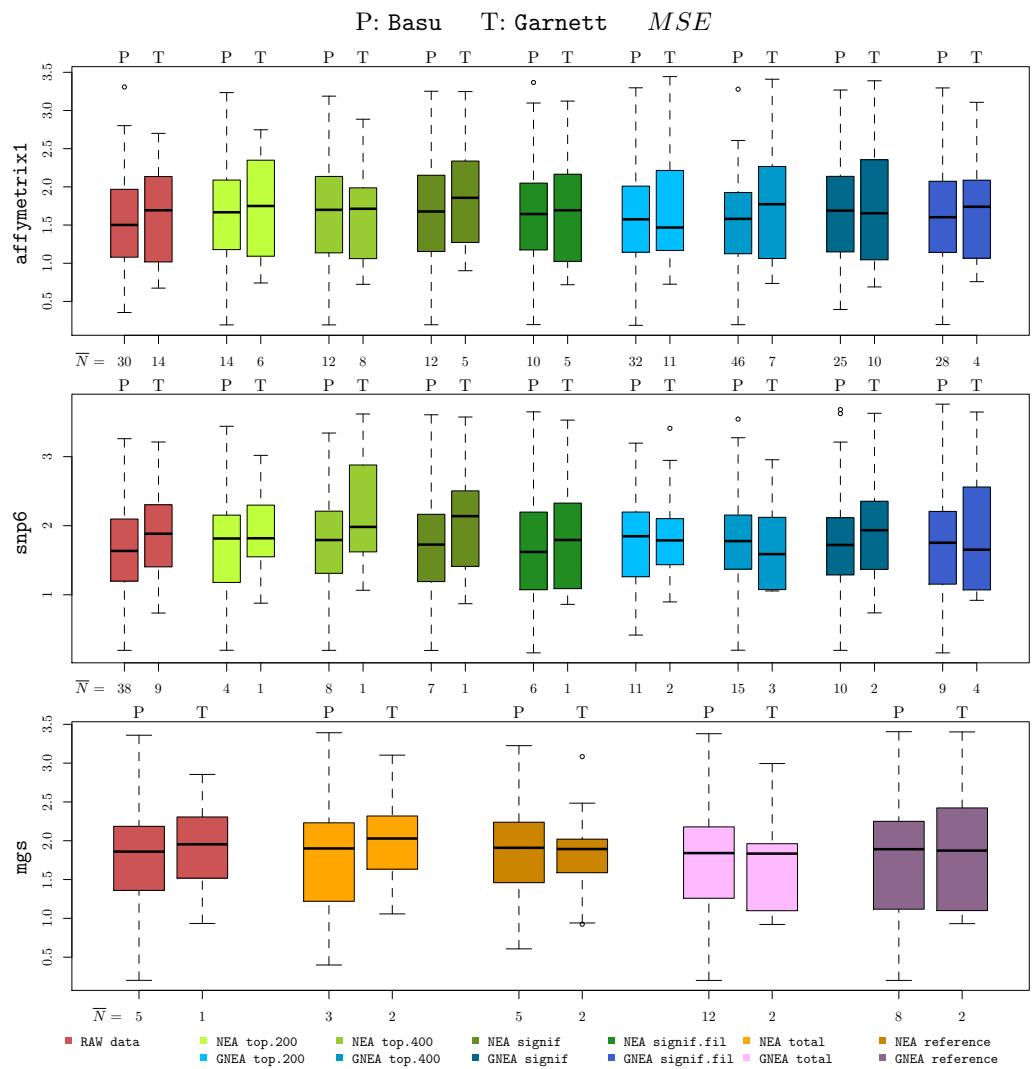


Figura D.11: Modelos para drogas compartidas,
P: Basu, T: Garnett, *MSE*

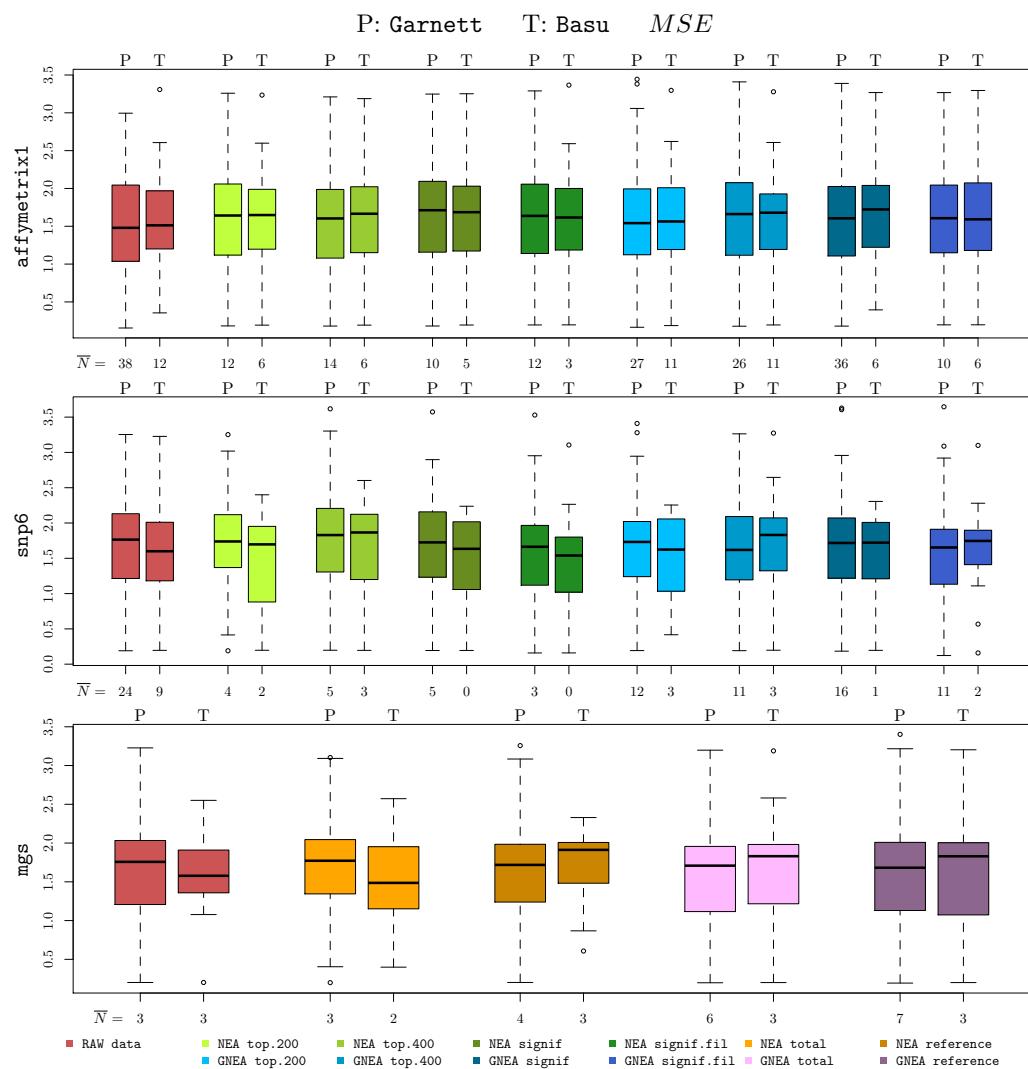


Figura D.12: Modelos para drogas compartidas,
P: Garnett, T: Basu, *MSE*

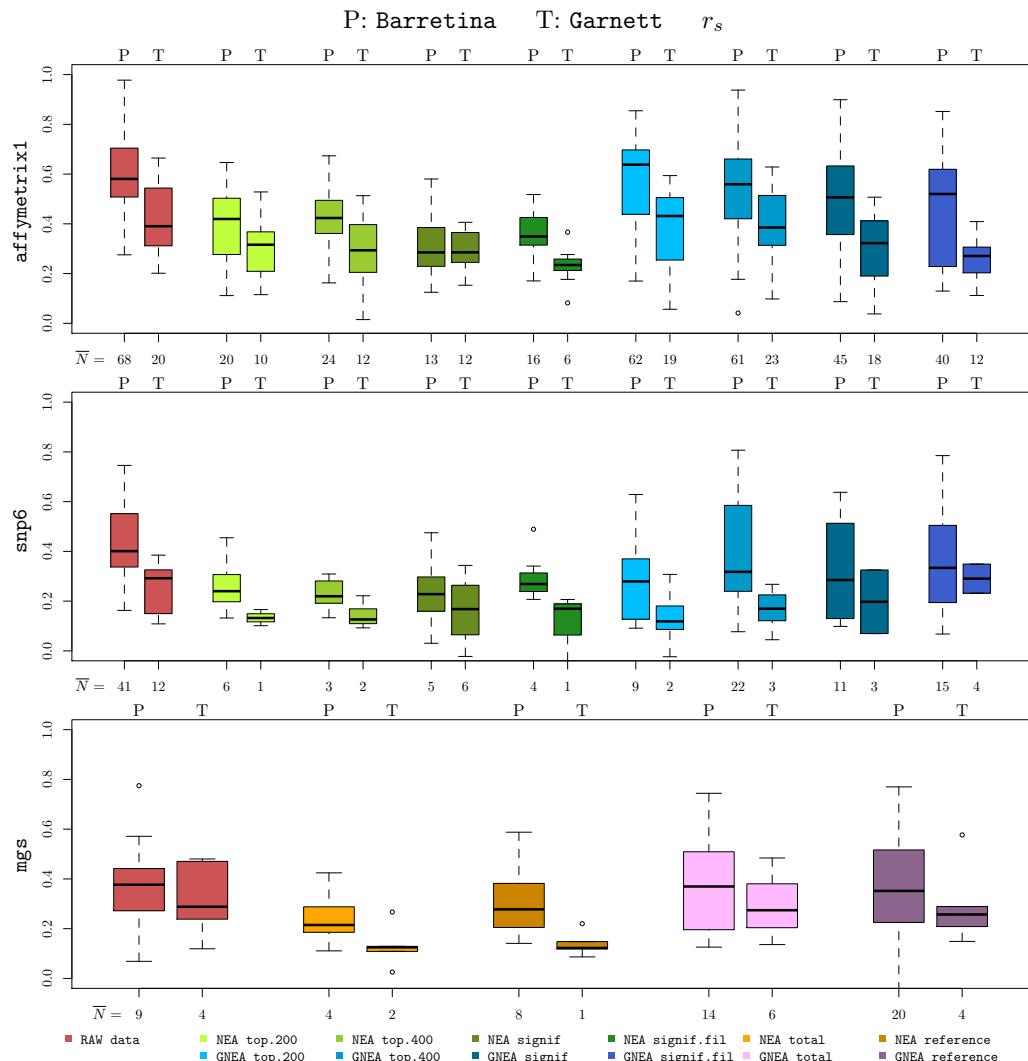


Figura D.13: Modelos para drogas compartidas,
P: Barretina, T: Garnett, r_s

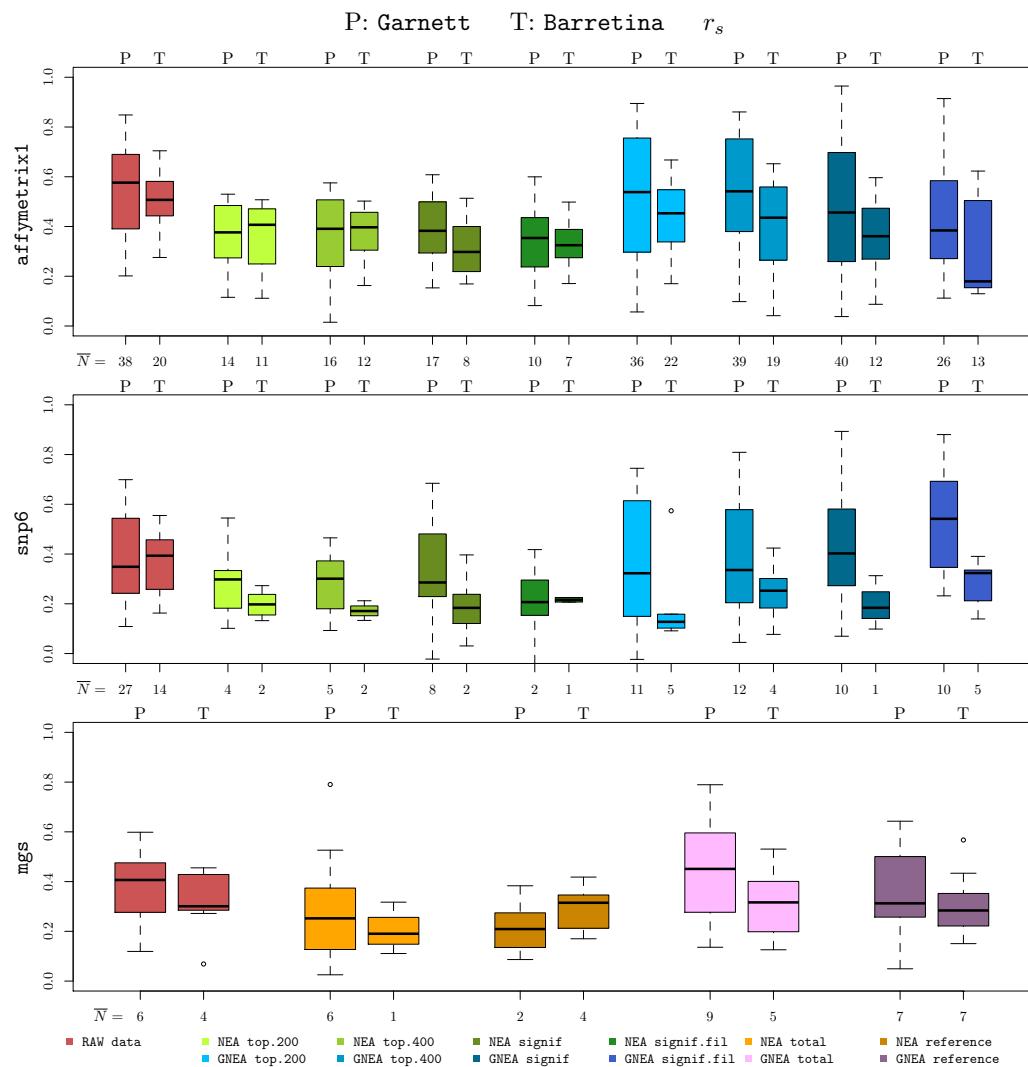


Figura D.14: Modelos para drogas compartidas,
P: Garnett, T: Barretina, r_s

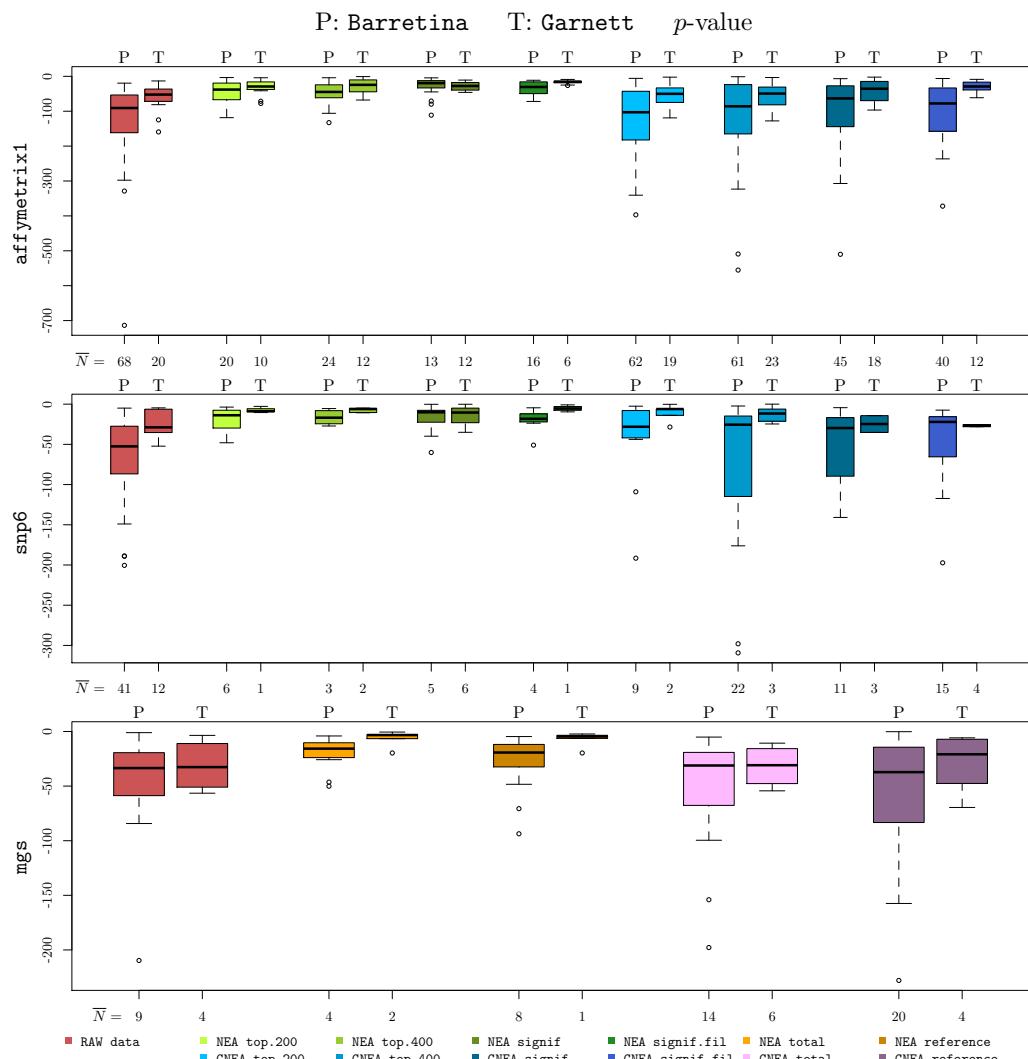


Figura D.15: Modelos para drogas compartidas,
P: Barretina, T: Garnett, *p*-value

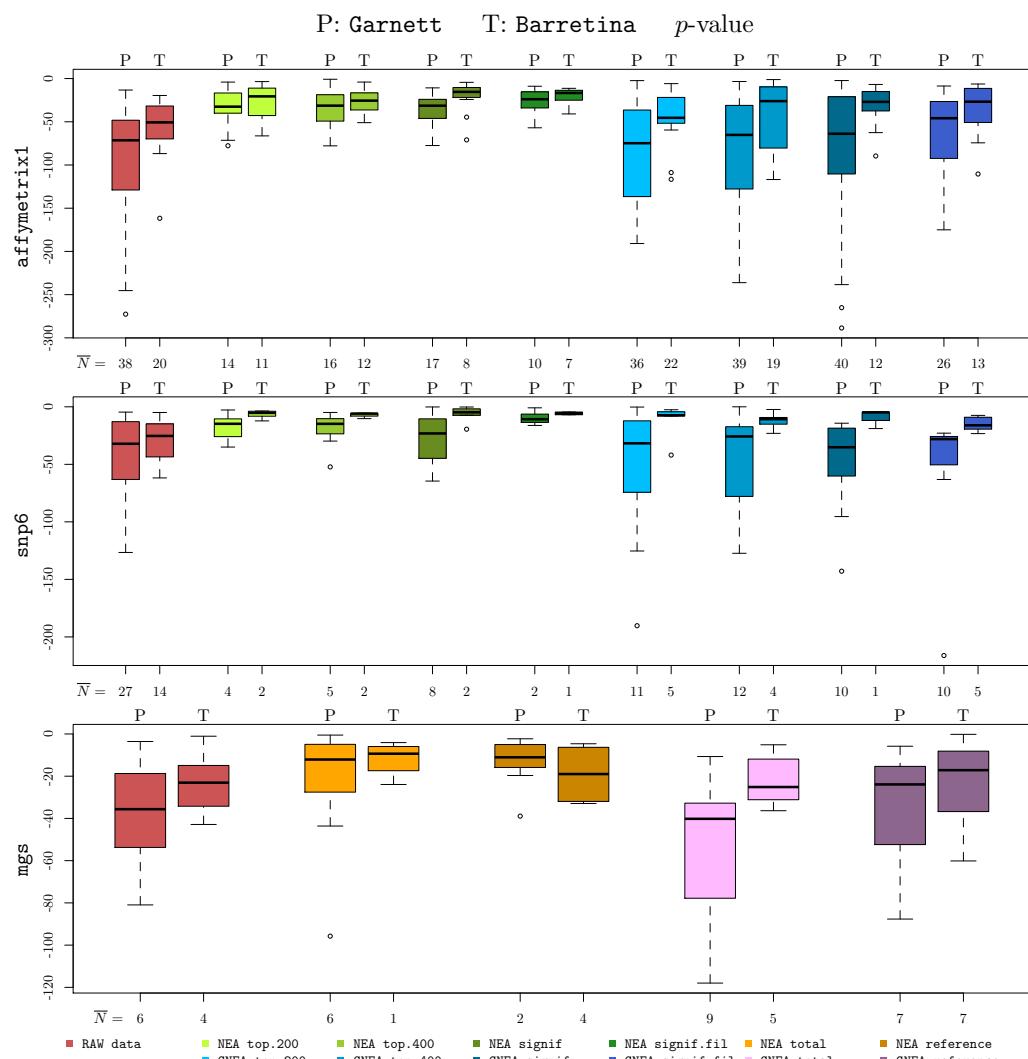


Figura D.16: Modelos para drogas compartidas,
P: Garnett, T: Barretina, *p*-value

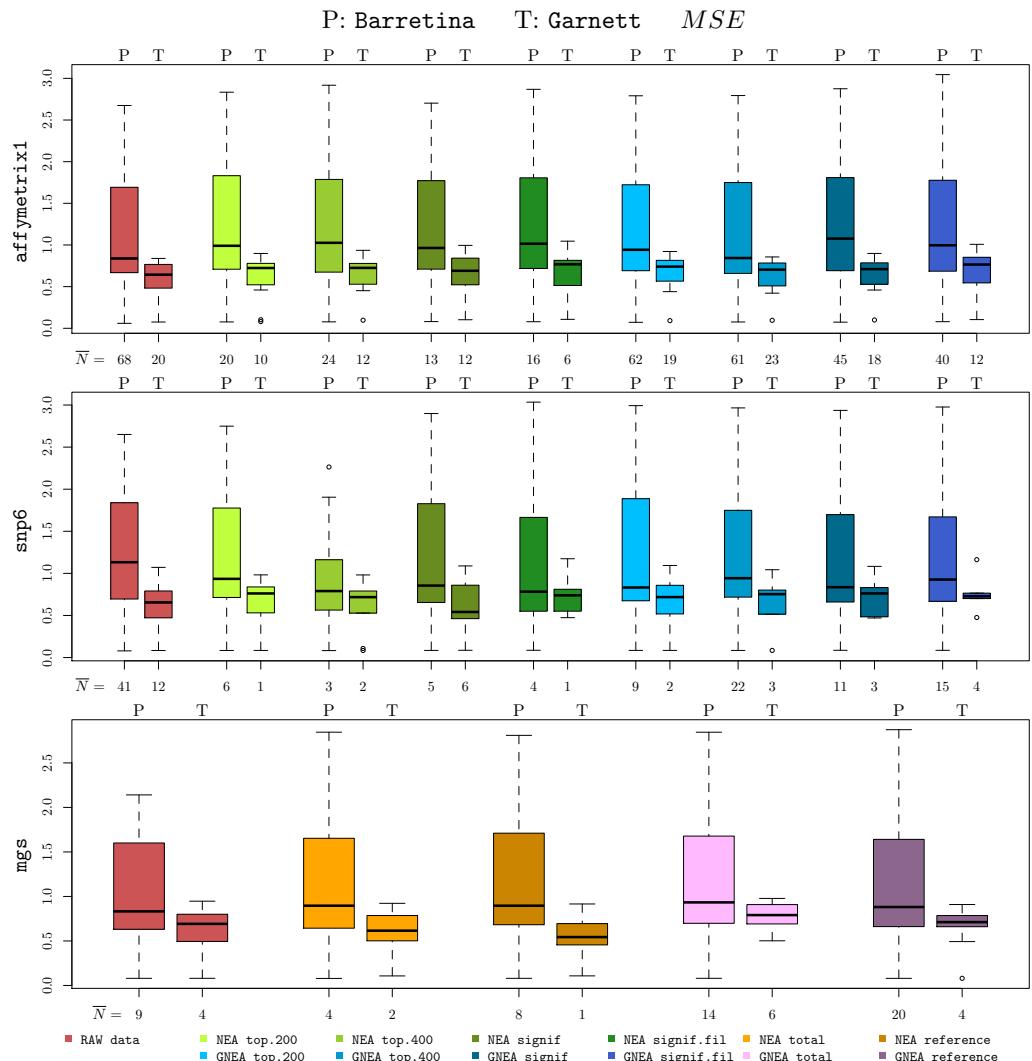


Figura D.17: Modelos para drogas compartidas,
P: Barretina, T: Garnett, *MSE*

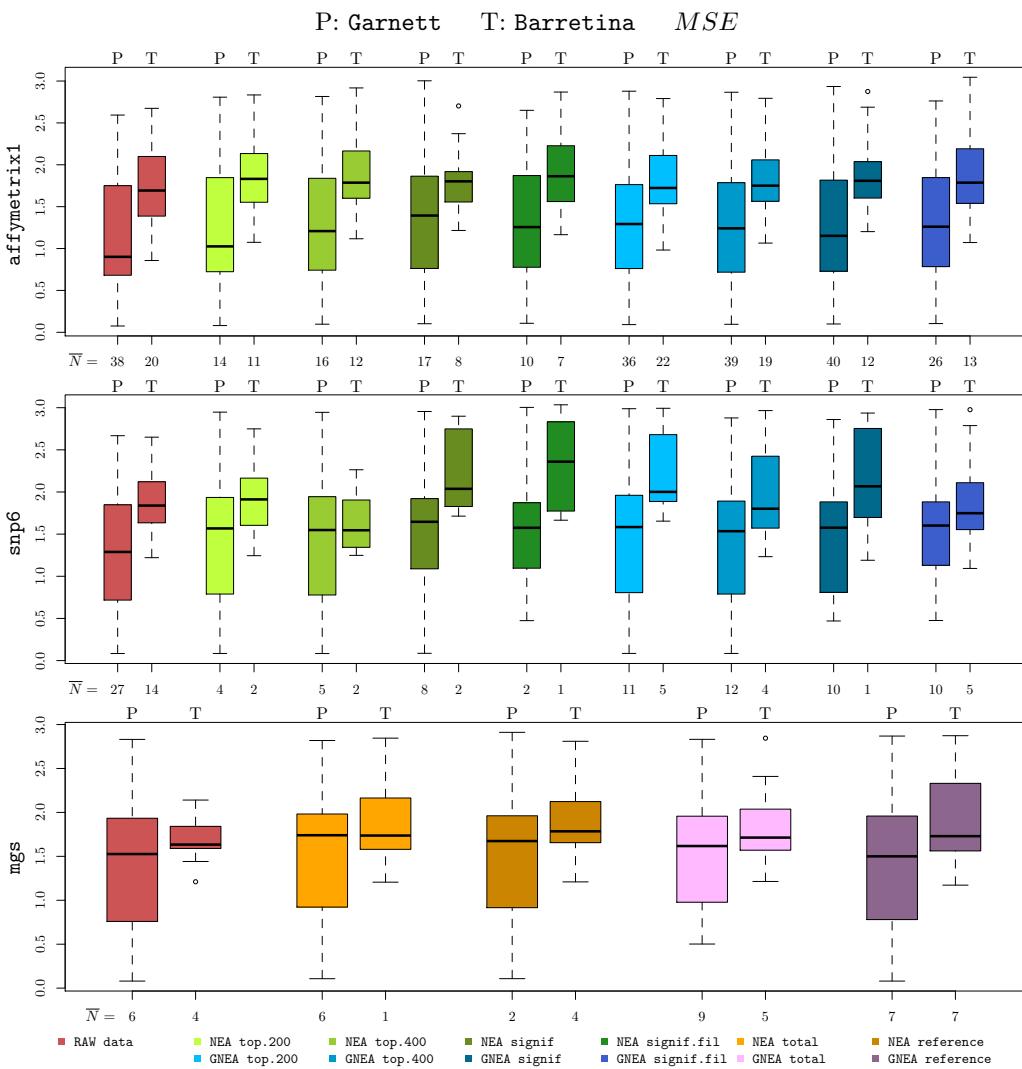


Figura D.18: Modelos para drogas compartidas,
P: Garnett, T: Barretina, *MSE*