

Sheffield
Chemoinformatics
Research
Group



The
University
Of
Sheffield.

Visualizaciones del espacio químico

Antonio de la Vega de León

Contenido

- Introducción
- Proyecciones de espacios multidimensionales
- Redes de compuestos químicos

Introducción

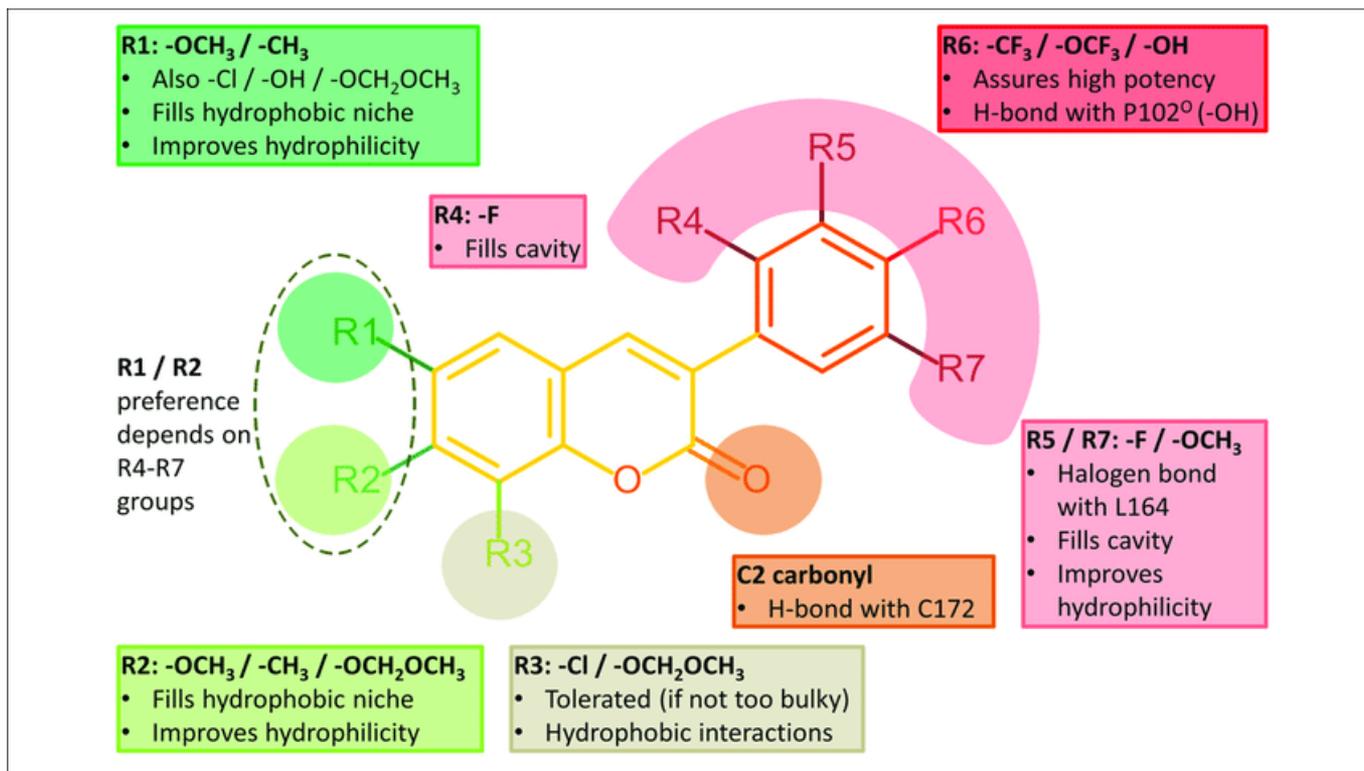
Espacio químico

Espacio compuesto por todos los compuestos químicos que es posible crear



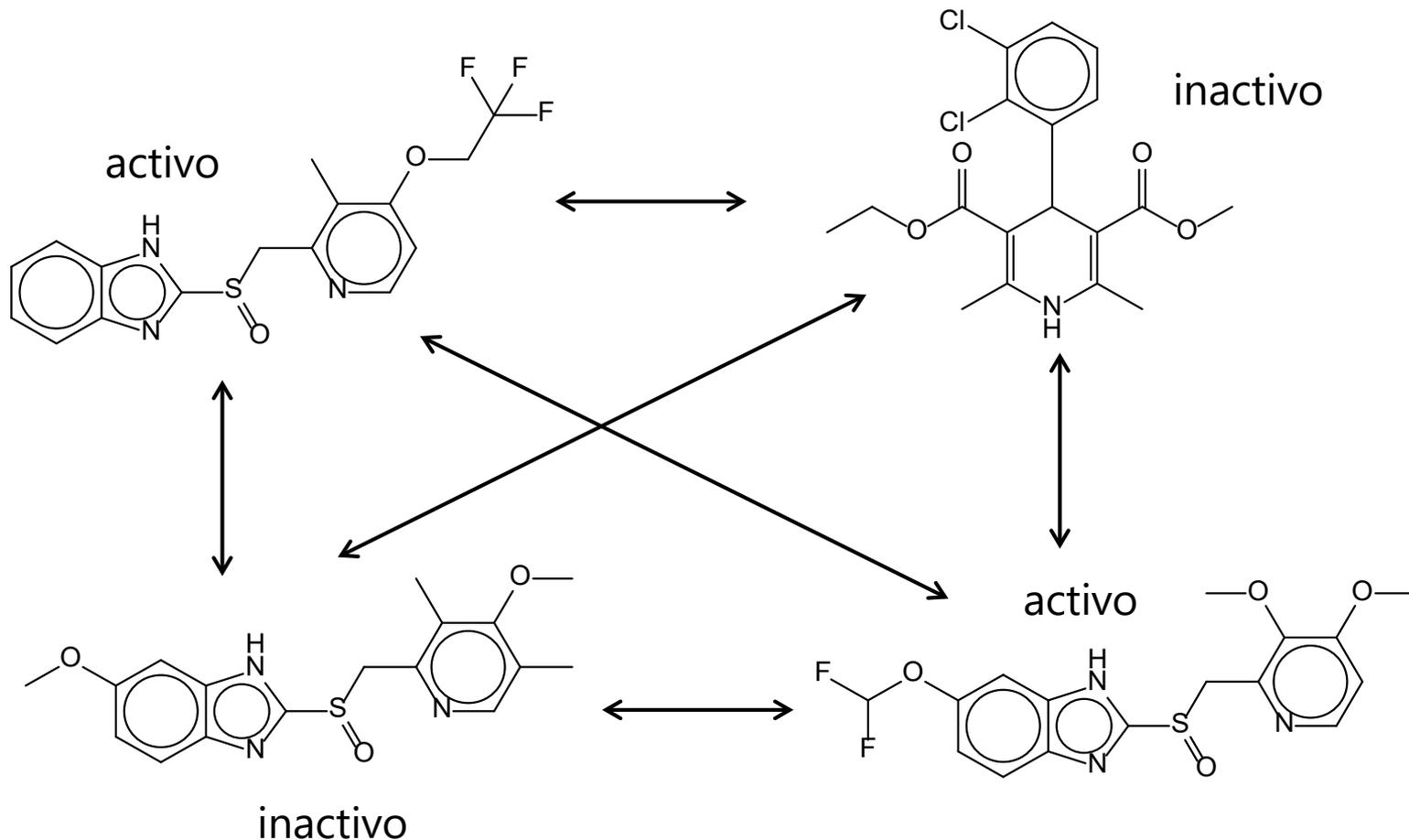
Espacio químico

- Análisis del espacio químico proporciona información sobre relaciones de estructura y propiedad



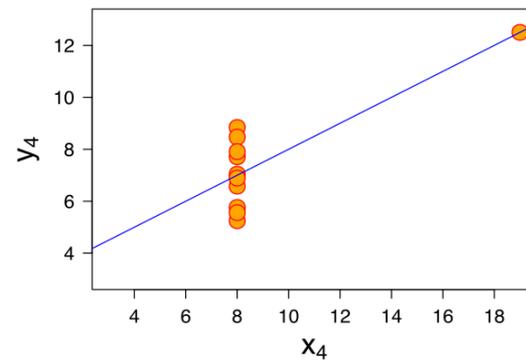
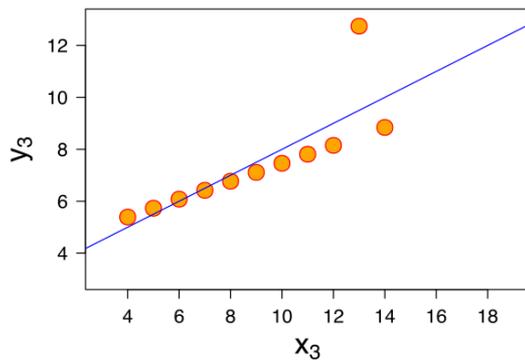
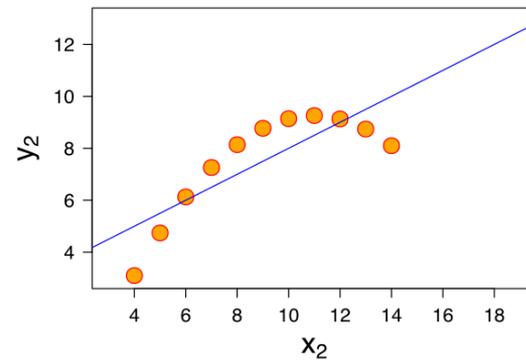
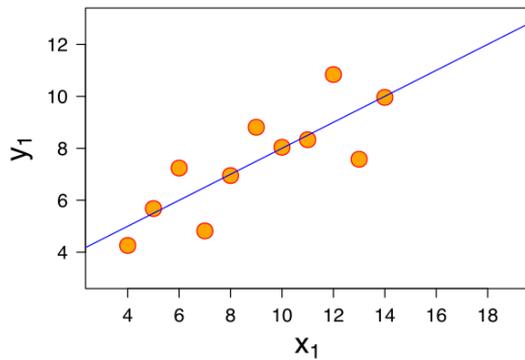
Visualizaciones

- Gráficos deben mostrar relaciones entre compuestos y las propiedades que son de interés



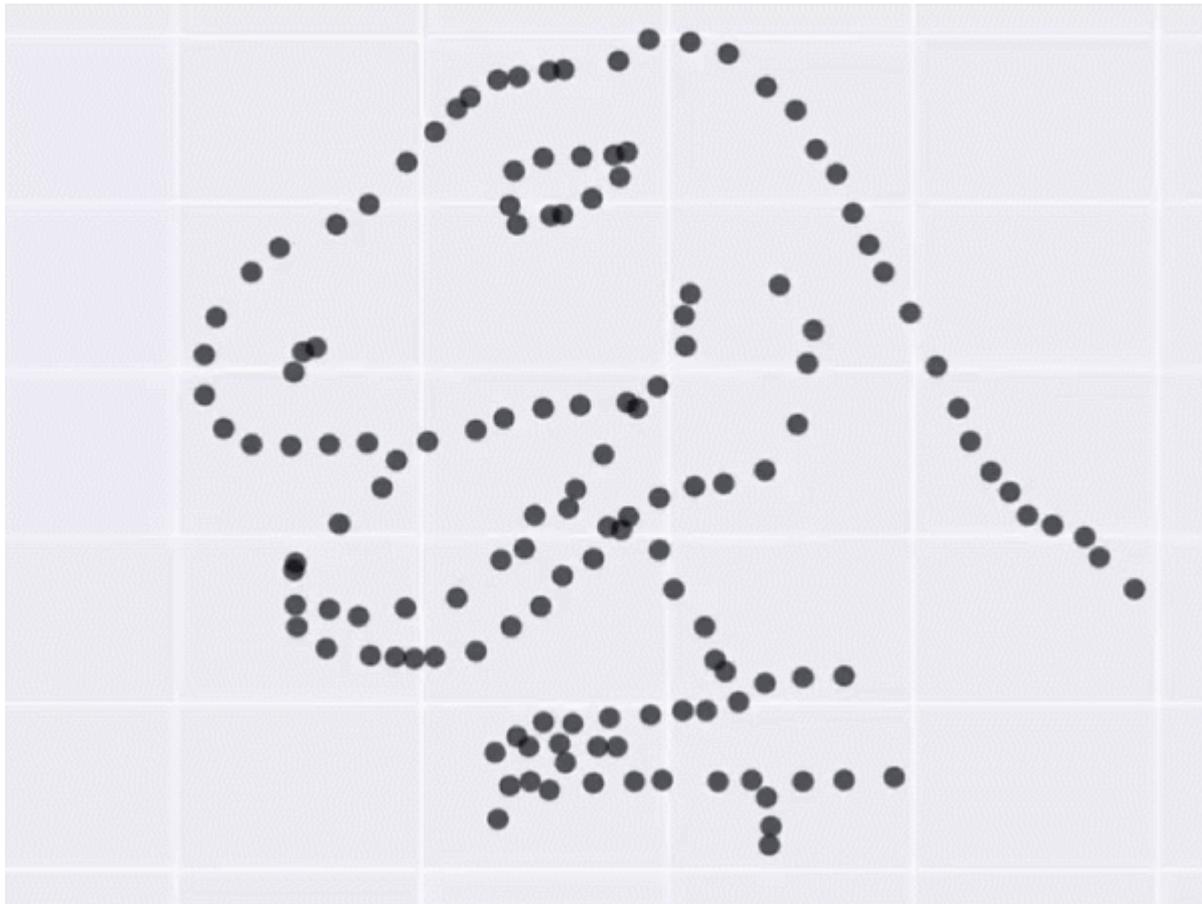
¿Porqué visualizar?

- Visualizaciones son capaces de mostrar relaciones difíciles de detectar con análisis estadísticos

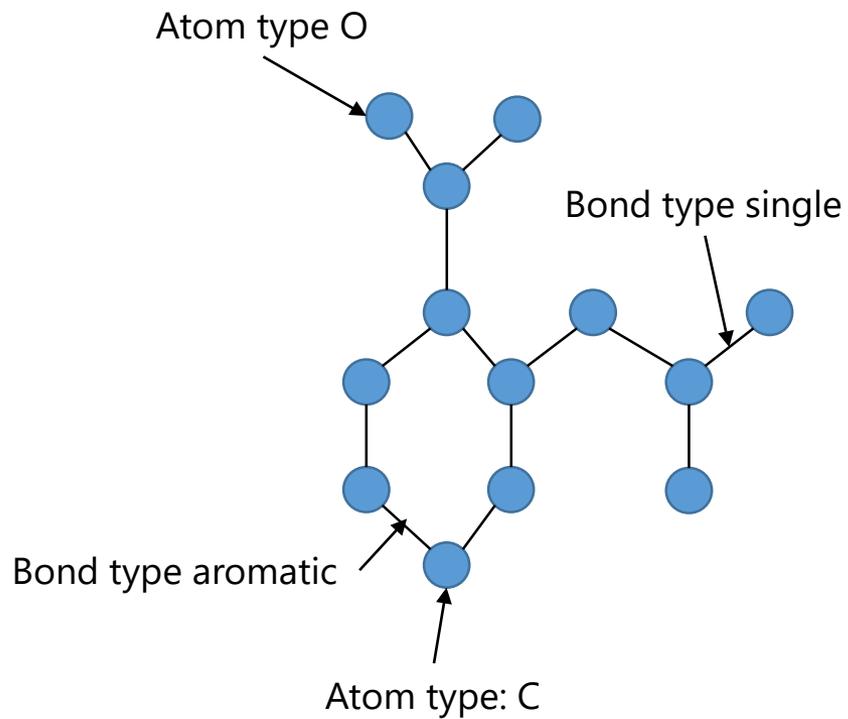
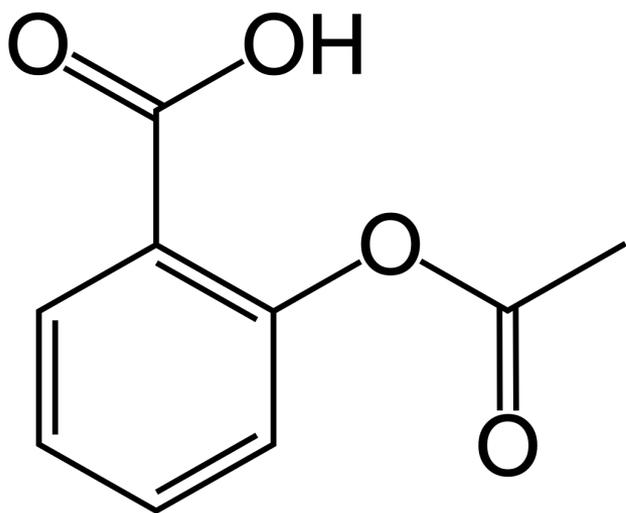


¿Porqué visualizar?

- Visualizaciones son capaces de mostrar relaciones difíciles de detectar con análisis estadísticos

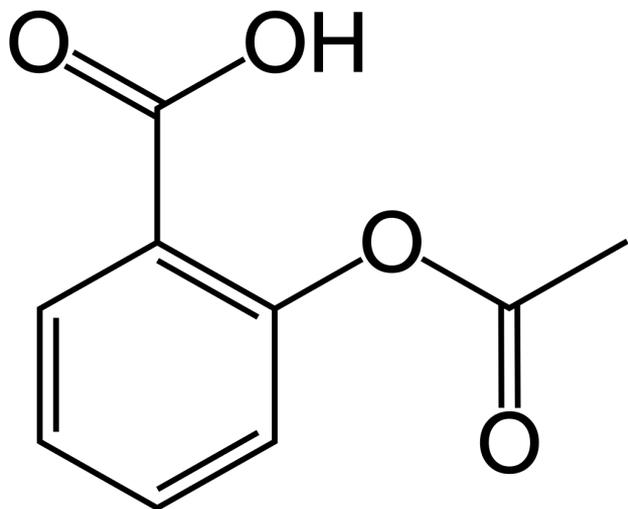


Representaciones de compuestos



Representaciones de compuestos

- Descriptores fisico-químicos



Tamaño

logP

Número de aceptores de hidrogeno

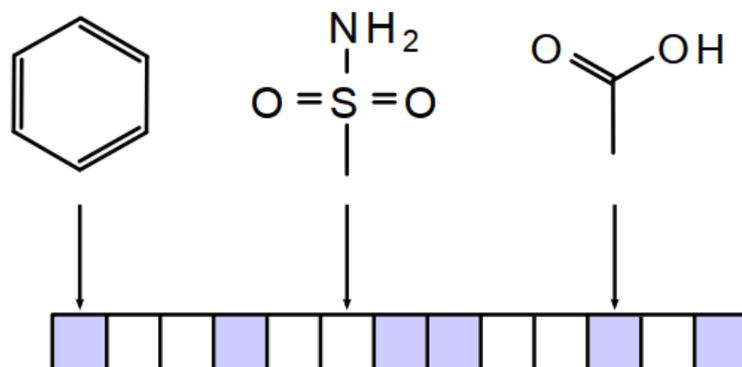
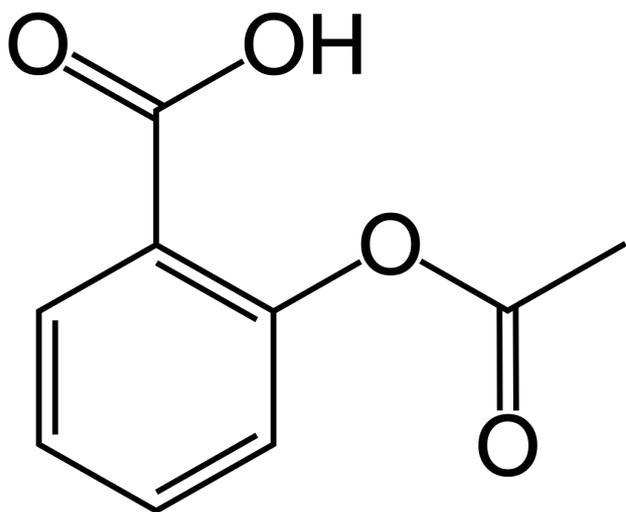
Porcentaje de átomos aromáticos

Suma de cargas

180	1.19	3	0.46	0
-----	------	---	------	---

Representaciones de compuestos

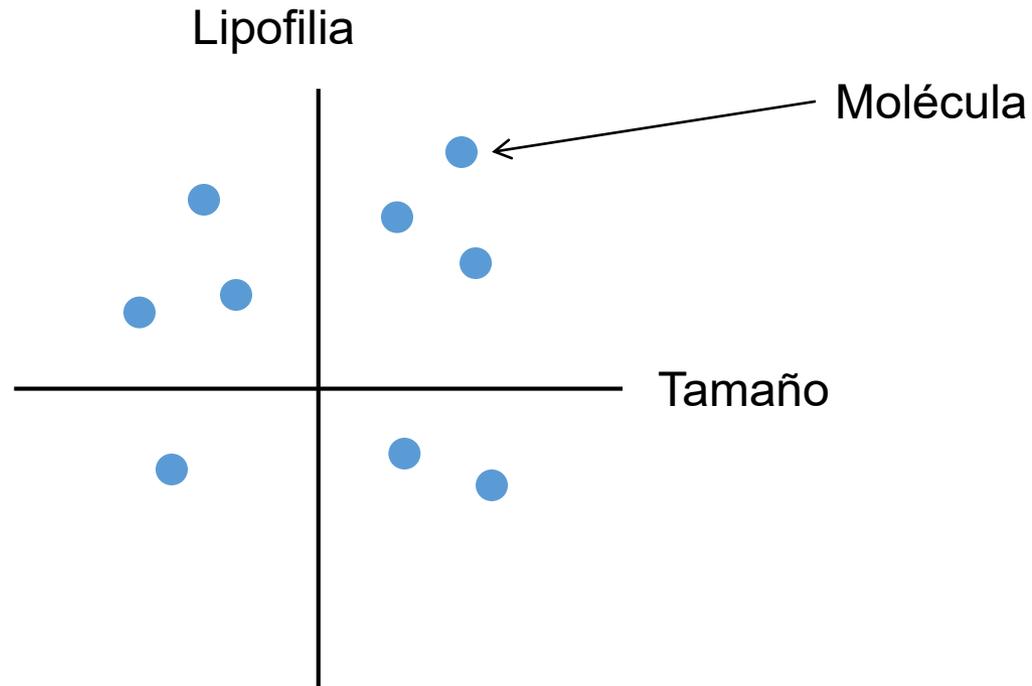
- Huellas digitales químicas



Proyecciones de espacios multidimensionales

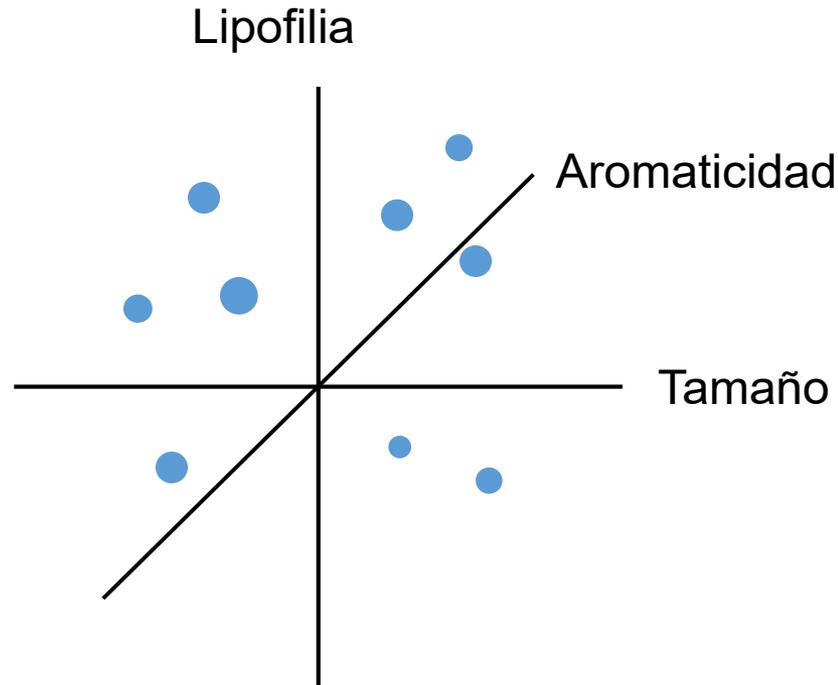
Mostrar espacios multidimensionales

- Problemas de mostrar más de dos dimensiones



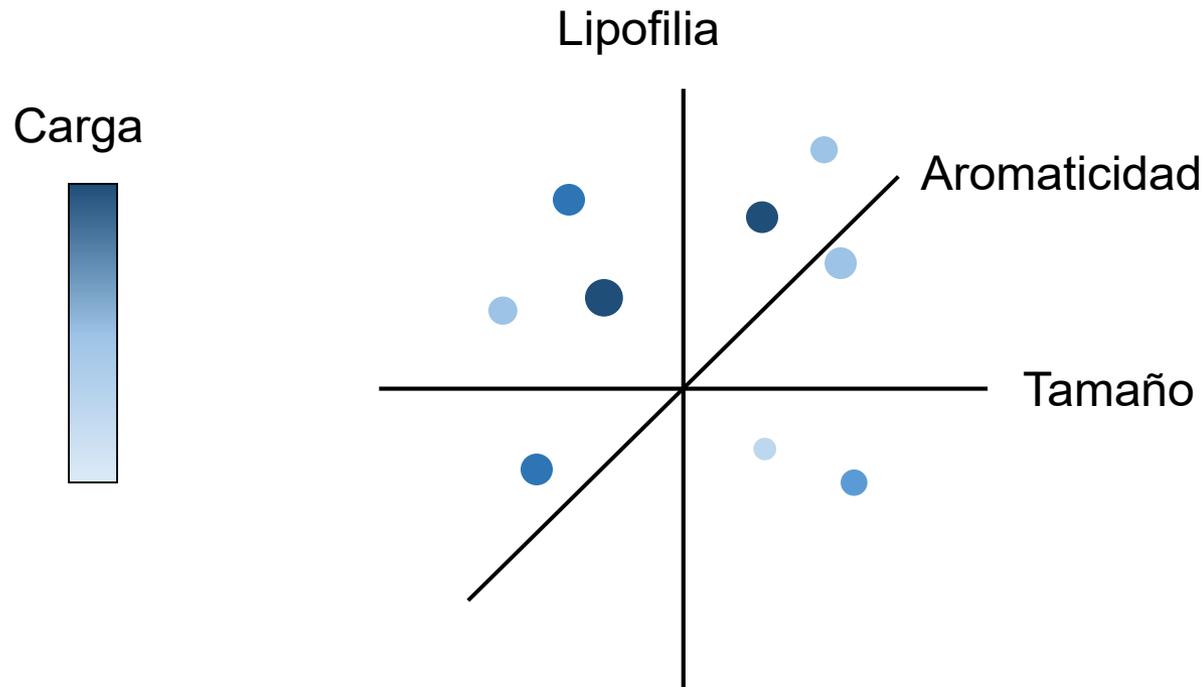
Mostrar espacios multidimensionales

- Problemas de mostrar más de dos dimensiones



Mostrar espacios multidimensionales

- Problemas de mostrar más de dos dimensiones

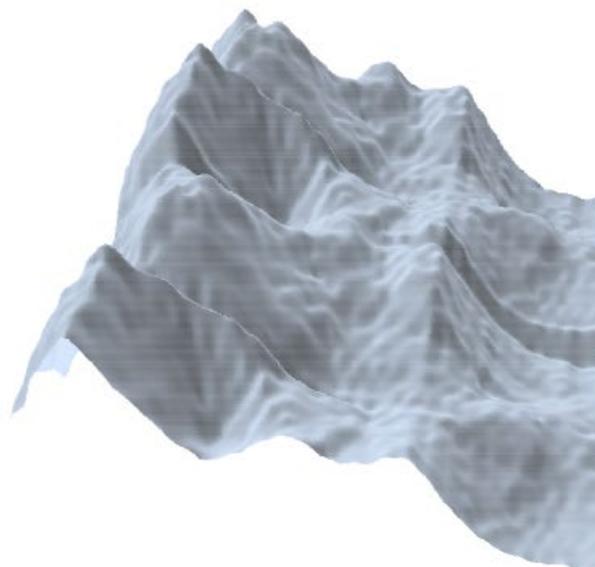
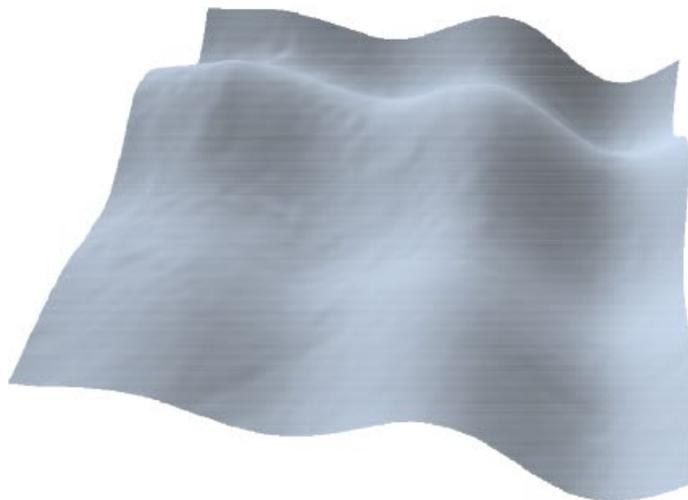


Reducción dimensional

- Distintos tipos de técnicas para reducir la dimensionalidad:
 - Análisis de componentes principales (PCA): obtienen una serie de ejes ortogonales que mantienen el máximo de variabilidad
 - Reducción multidimensional (MDS): proporciona coordenadas en dos dimensiones que mejor aproximan las distancias en el espacio químico

Paisajes de actividad

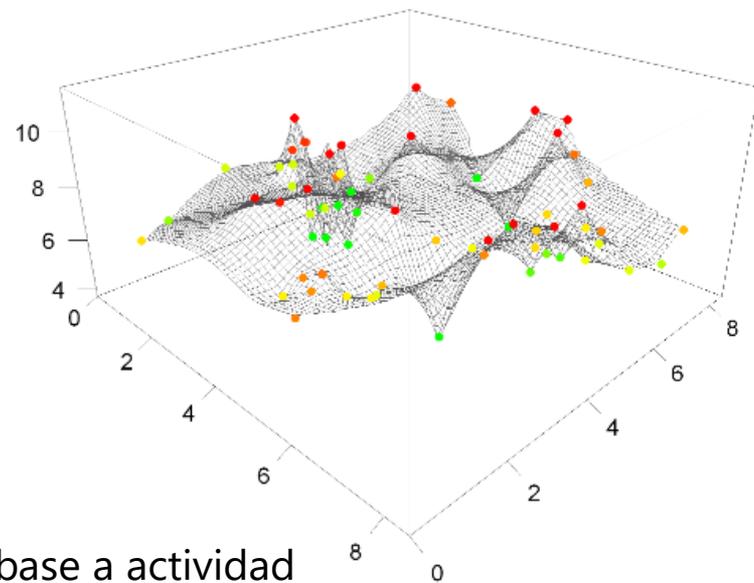
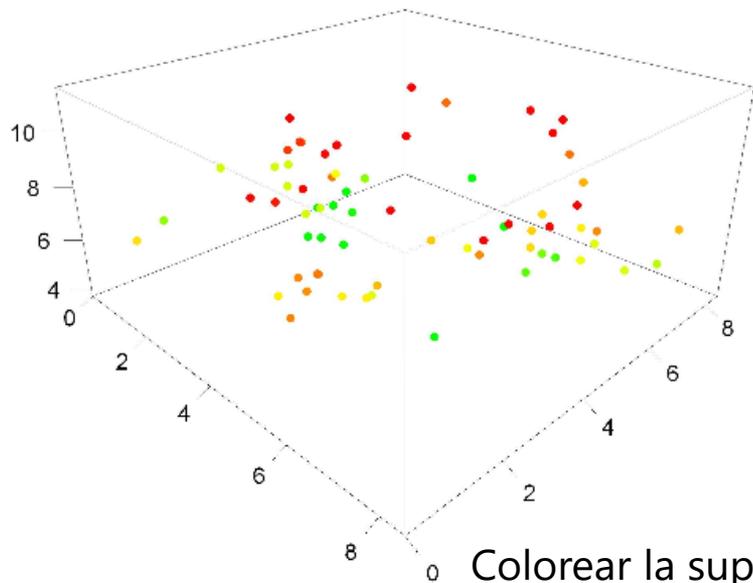
- Reducir la dimensionalidad y añadir la actividad en el eje Z



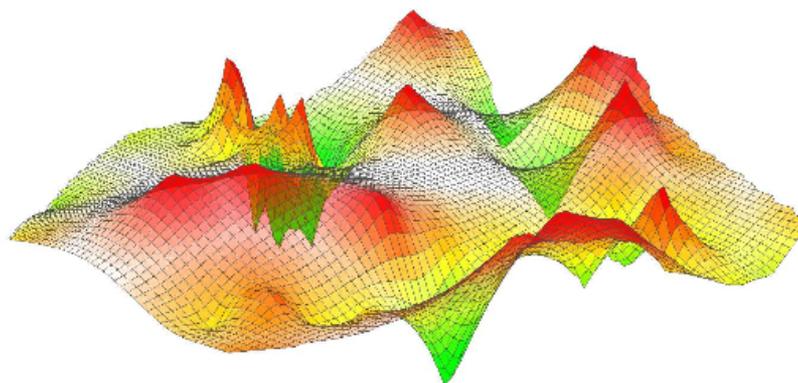
Paisajes de actividad

Reducir a 2D y usar actividad como 3^{ra} dimension

Interpolar una superficie entre los datos



Colorear la superficie en base a actividad

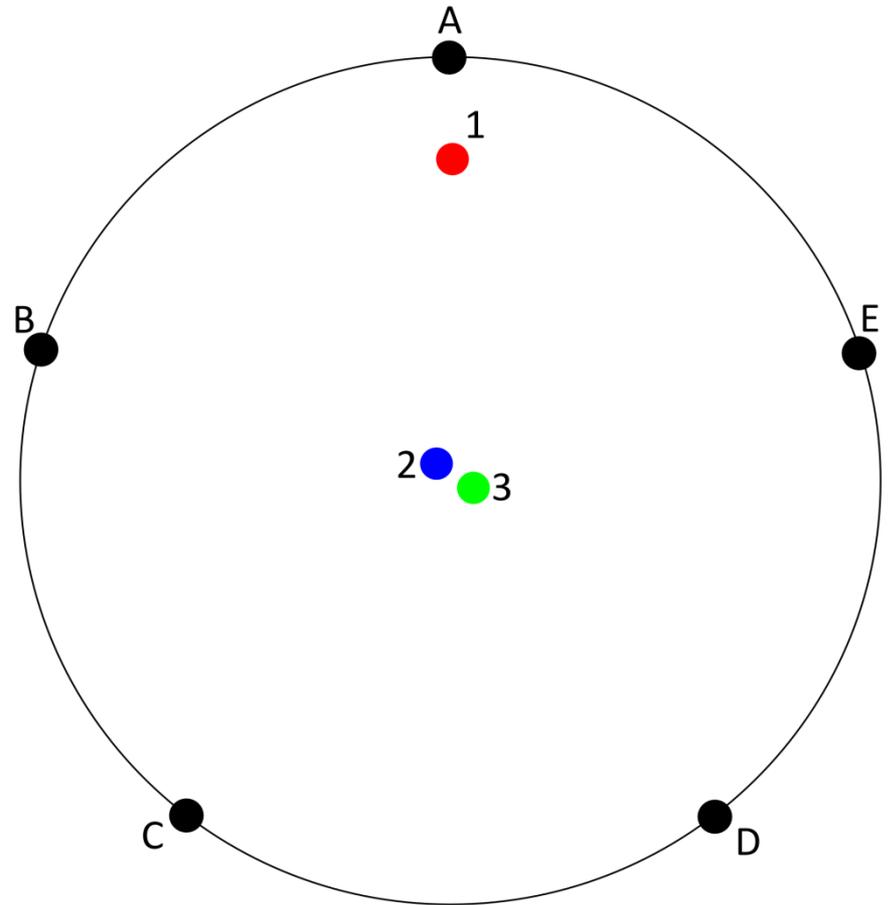


Expandir paisajes a multiples actividades

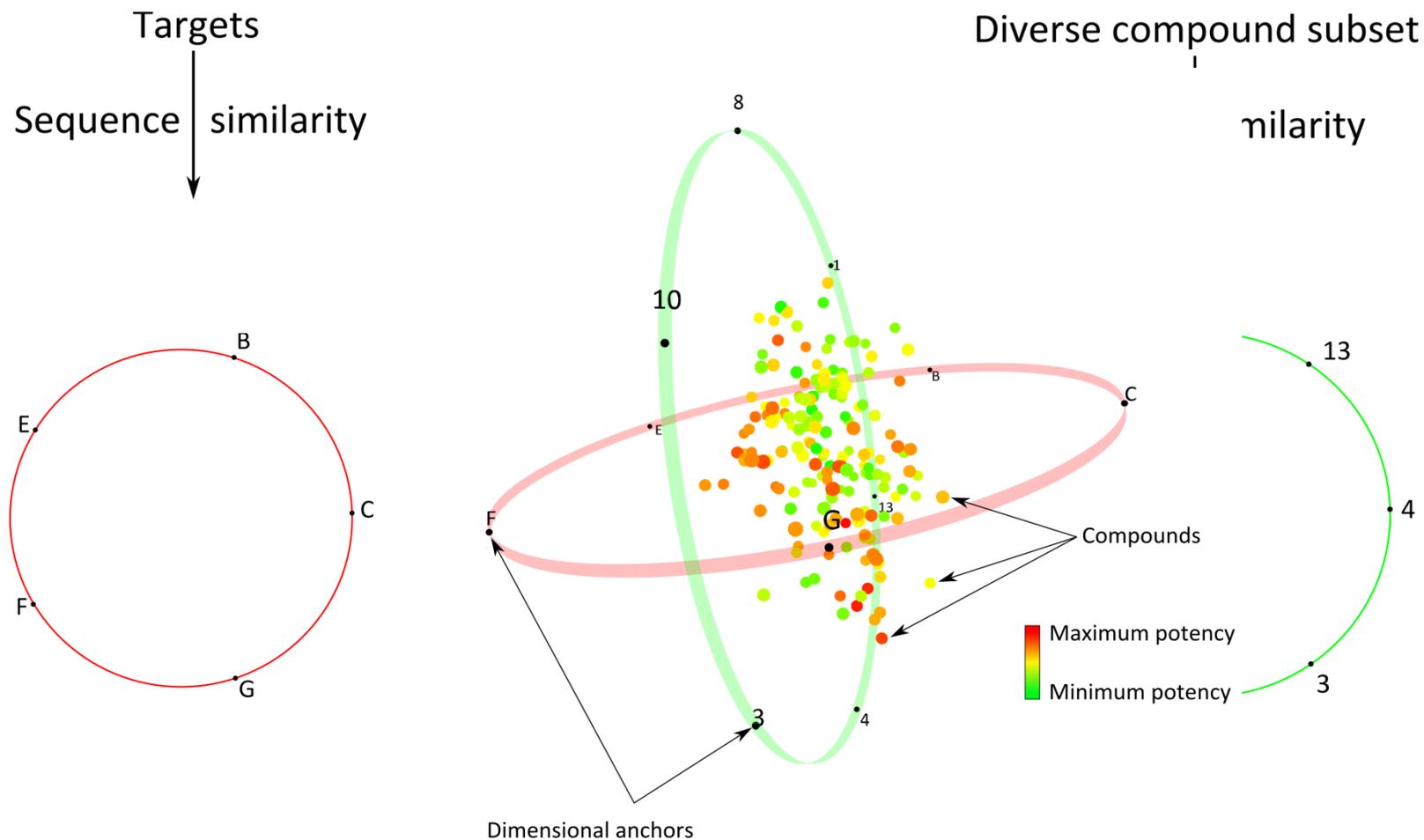
- Una limitación es que solo es capaz de mostrar una propiedad y no varias
- Sería interesante poder mostrar relaciones entre estructura y diferentes propiedades en un gráfico
- Por ejemplo, actividad contra un panel de quinasas
- Para ello, utilizamos técnicas de visualización multidimensional como RadViz

Paisajes de multiactividad 3D

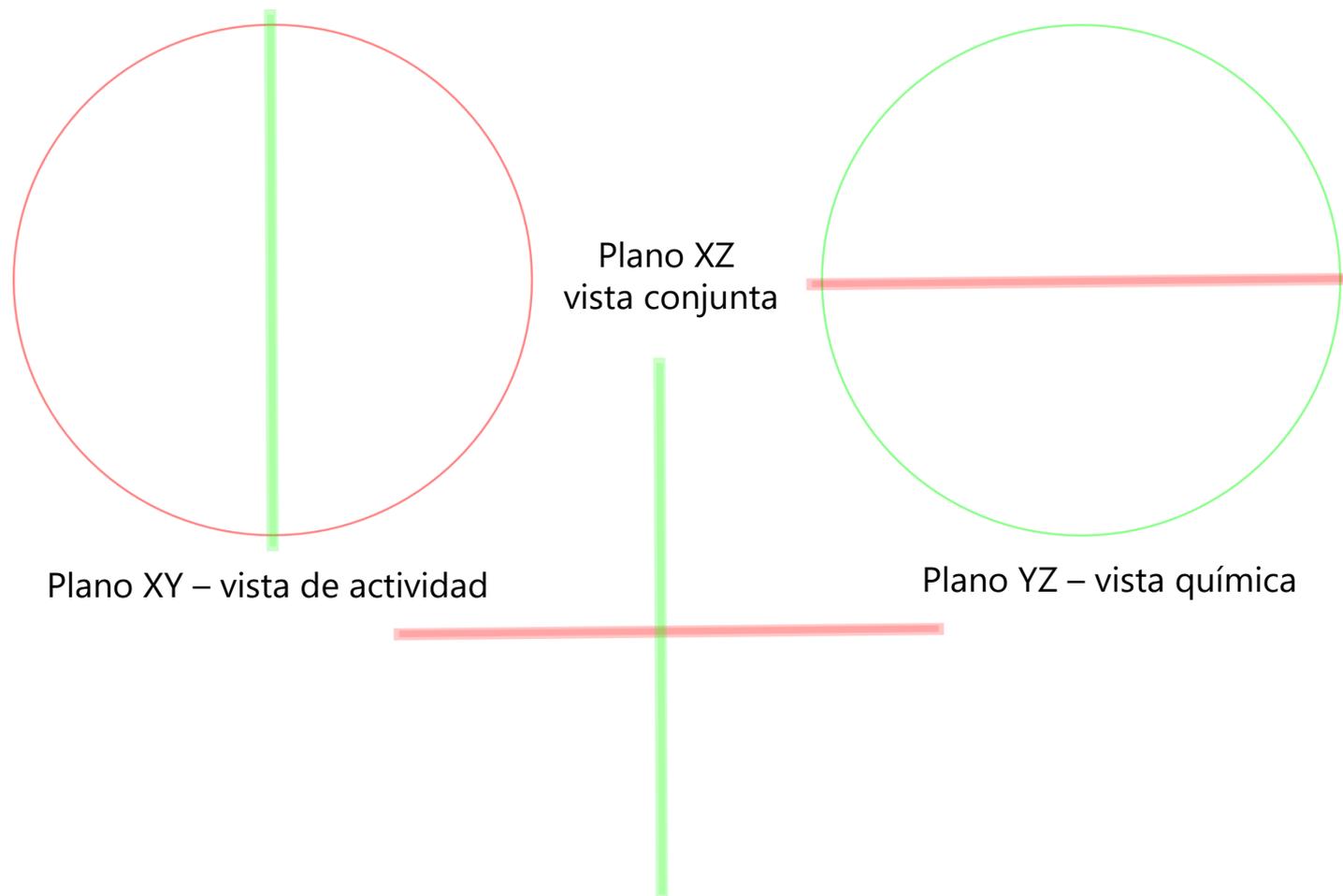
- Es una técnica de proyección
- Utiliza una serie de anclas dimensionales (A-E)
- La posición de los datos depende de los valores para cada ancla dimensional



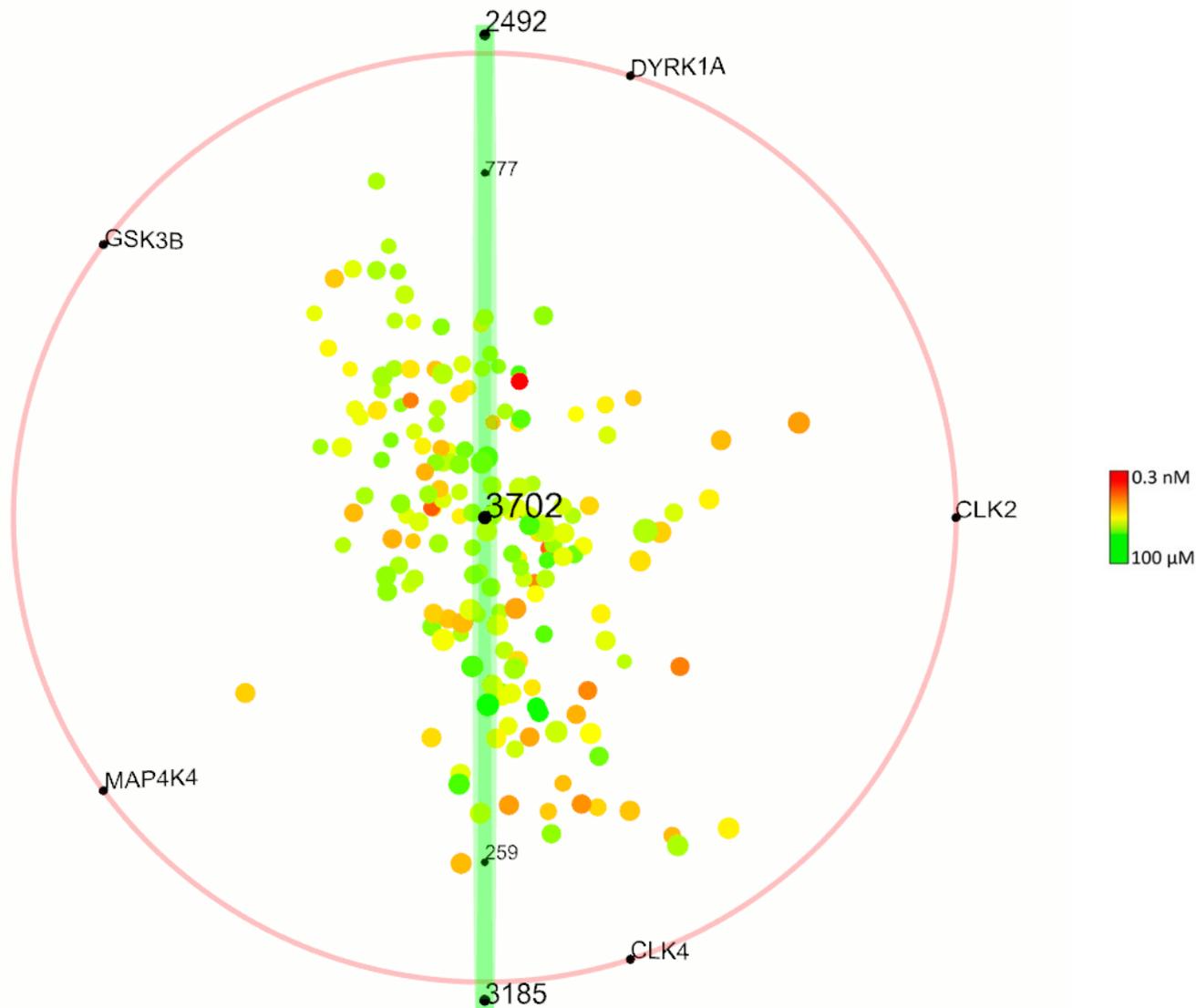
Paisajes de multiactividad 3D



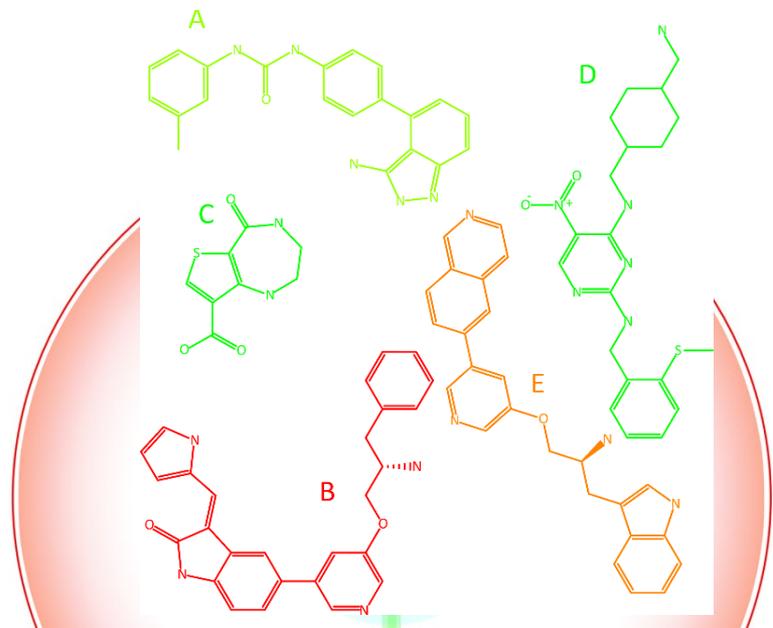
Paisajes de multiactividad 3D



Paisajes de multiactividad

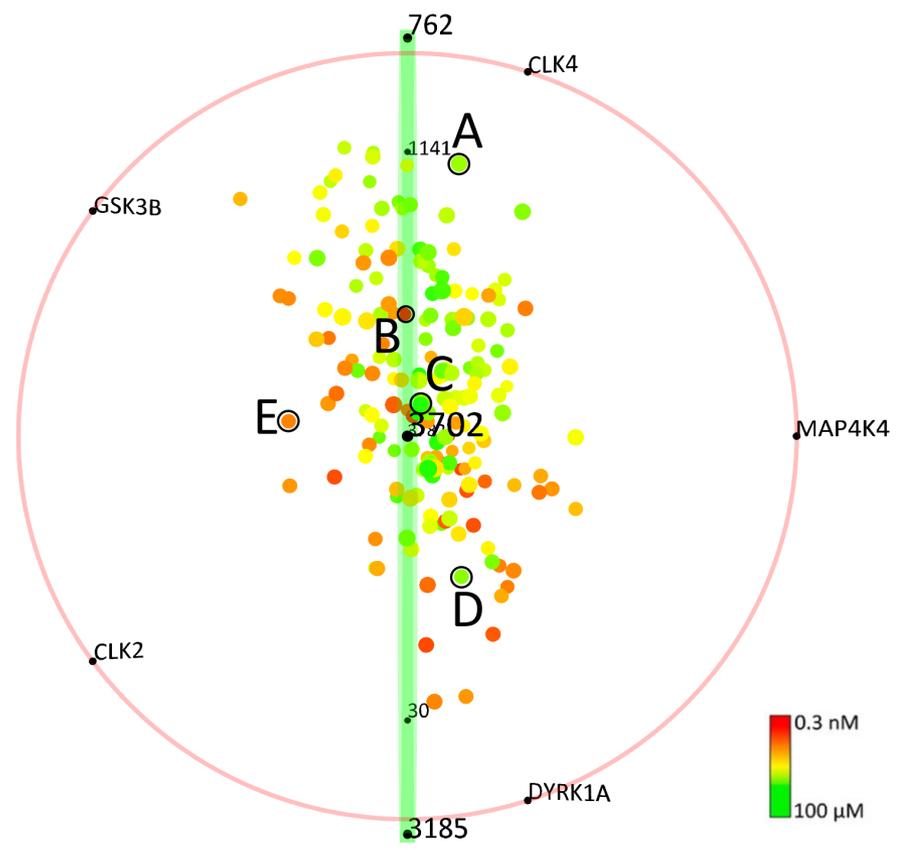


Paisajes de multiactividad 3D



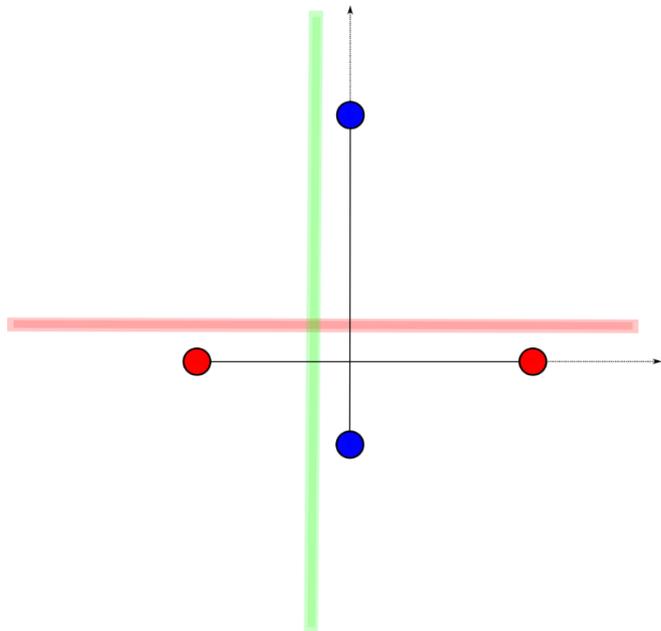
	A	B	C	D	E
MAP4K4	5.8	8.3	4.5	6.5	5.3
CLK4	7.1	9.5	5.4	5.5	9
GSK3B	5.4	8.6	4.3	5.7	6.8
CLK2	5.7	9.1	4.7	5.7	8.6
DYRK1A	5.3	9.4	5.2	5.5	8.5

 Selectivo
 Ambiguo
 No selectivo

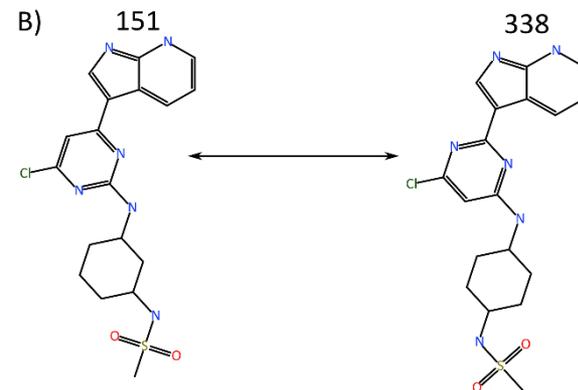
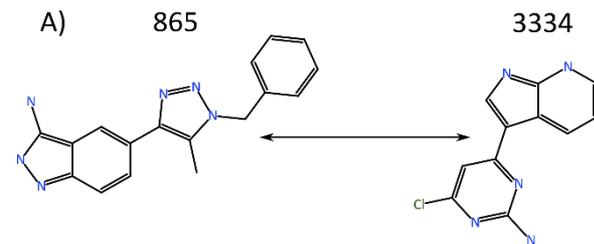


Paisajes de multiactividad 3D

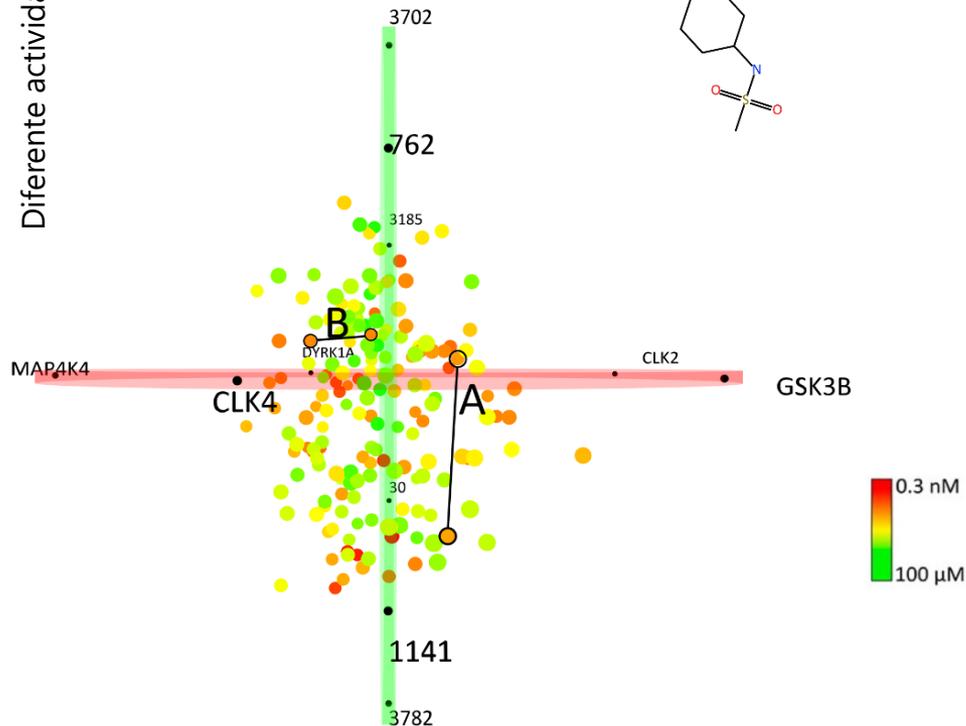
Actividad parecida, diferente estructura



Diferente actividad, estructura parecida



	A	A	B	B
MAP4K4	6.0	6.1	6.7	7.8
CLK4	8.0	7.7	8.5	7.2
GSK3B	7.2	6.9	6.1	7.8
CLK2	7.0	7.1	7.8	5.7
DYRK1A	7.0	6.9	7.3	7.3



Redes de compuestos químicos

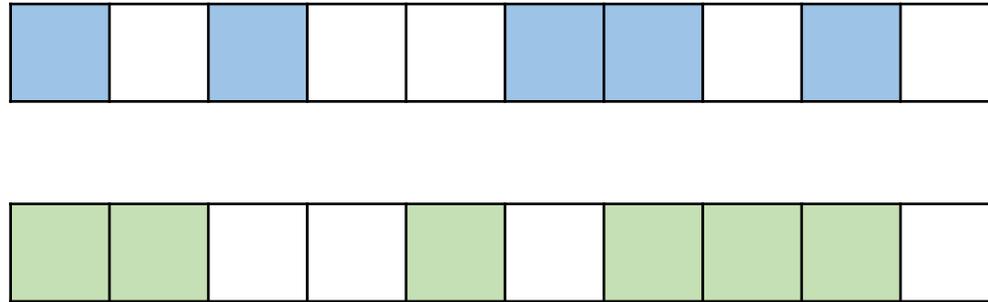
Redes de compuestos químicos

- Concepto similar al de redes sociales



Semejanza entre compuestos

- Compuestos se consideran similares si su semejanza es mayor que un valor específico



$$Tc(i, j) = \frac{c}{a + b - c}$$

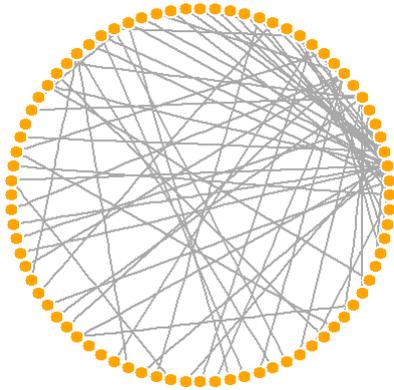
$$A = 5$$

$$B = 6$$

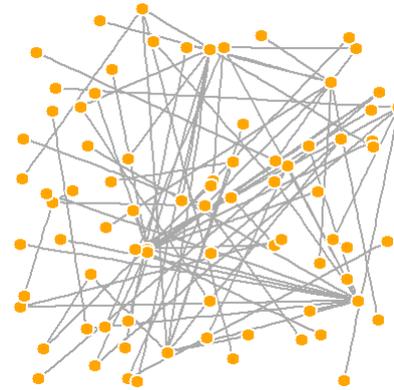
$$C = 3$$

$$Tc = 0.375$$

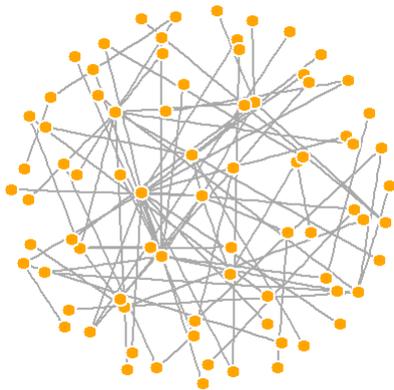
Disposición de redes



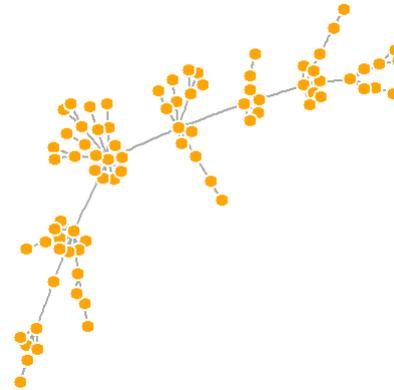
Círculo



Aleatorio



Esfera



Fruchterman-Reingold

Redes de compuestos

● — ● $T_c > x$

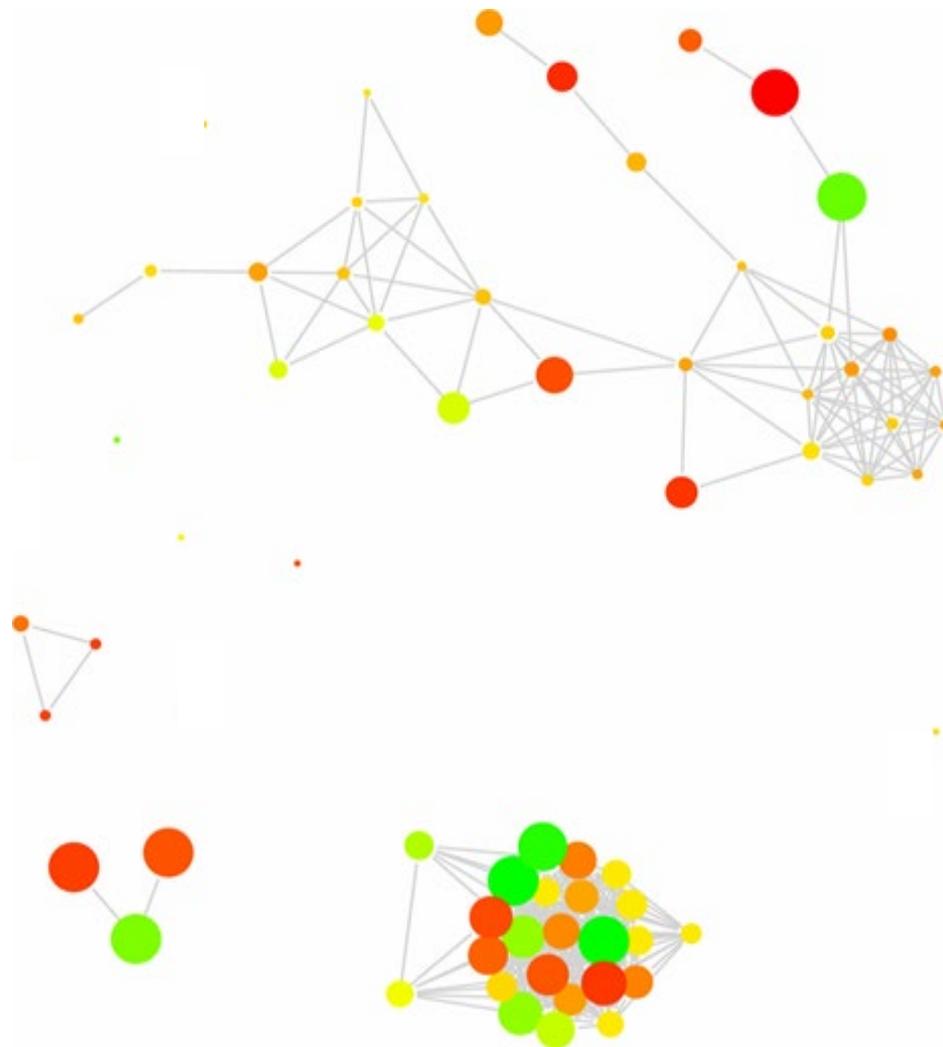
● ● $T_c < x$

 alta actividad

baja actividad

● alta discontinuidad REA

● baja discontinuidad REA

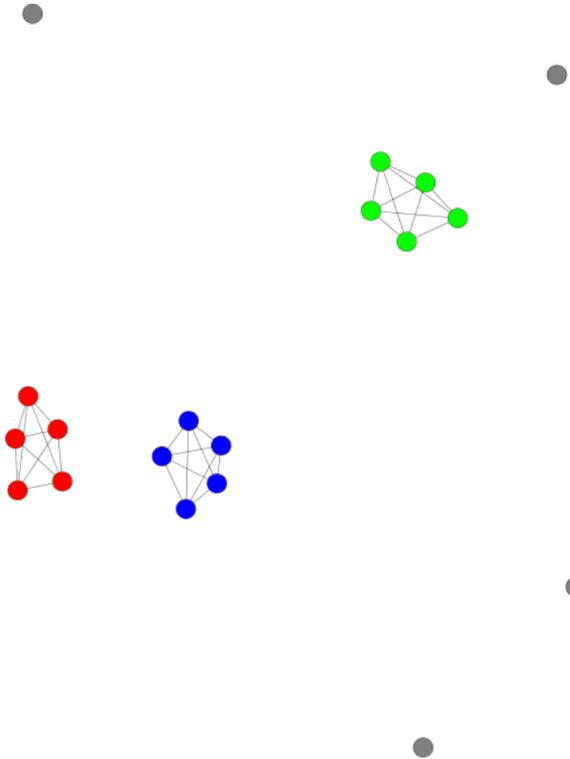


Limitación de estas redes

- Las distancias entre compuestos no tienen significado químico
- La red depende del valor límite que determina cuando dos compuestos son similares
- Cambiar el valor, cambia la disposición completamente
- Adaptamos una técnica de disposición que ayuda con estos problemas: Kamada-Kawai (KK)

Disposición Kamada-Kawai

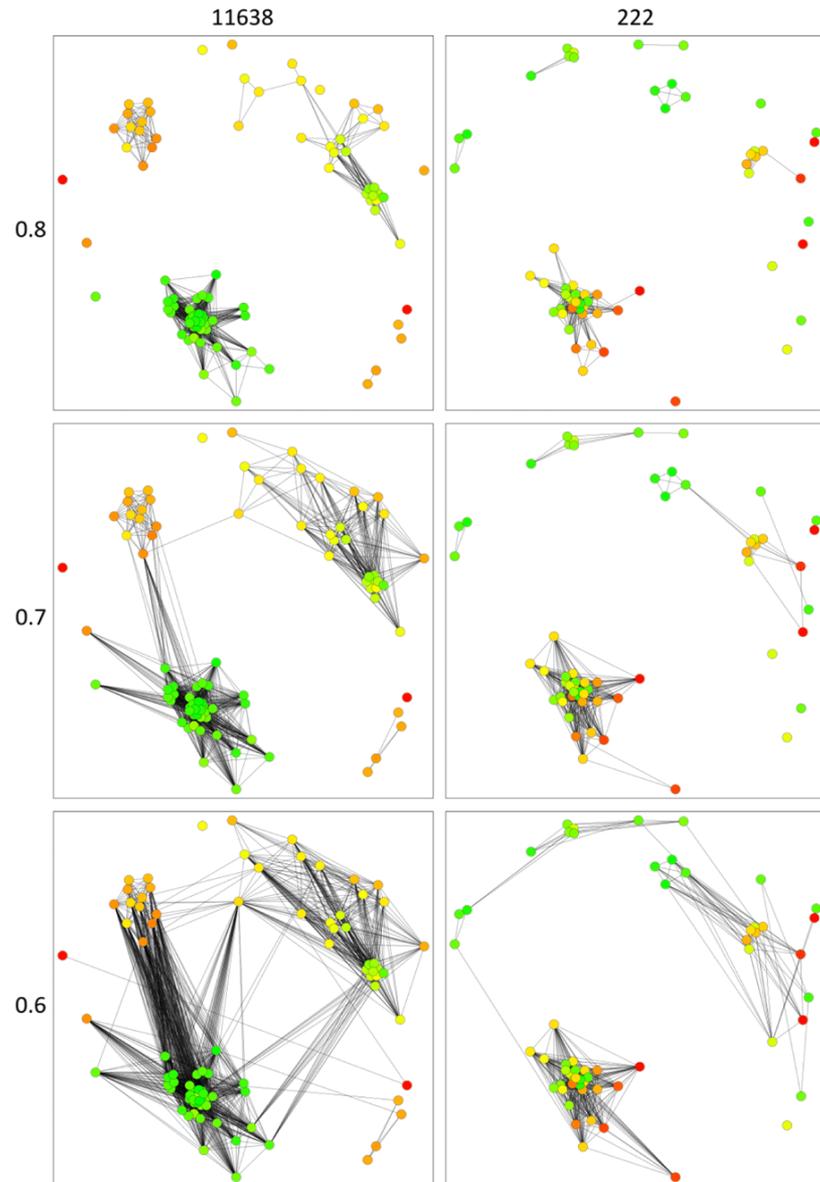
- Es capaz de mostrar la semejanza entre compuestos



	A	B	C	D
A	1.0-0.9	0.8-0.7	0.6-0.5	0.1-0.0
B	0.8-0.7	1.0-0.9	0.4-0.3	0.1-0.0
C	0.6-0.5	0.4-0.3	1.0-0.9	0.1-0.0
D	0.1-0.0	0.1-0.0	0.1-0.0	0.1-0.0

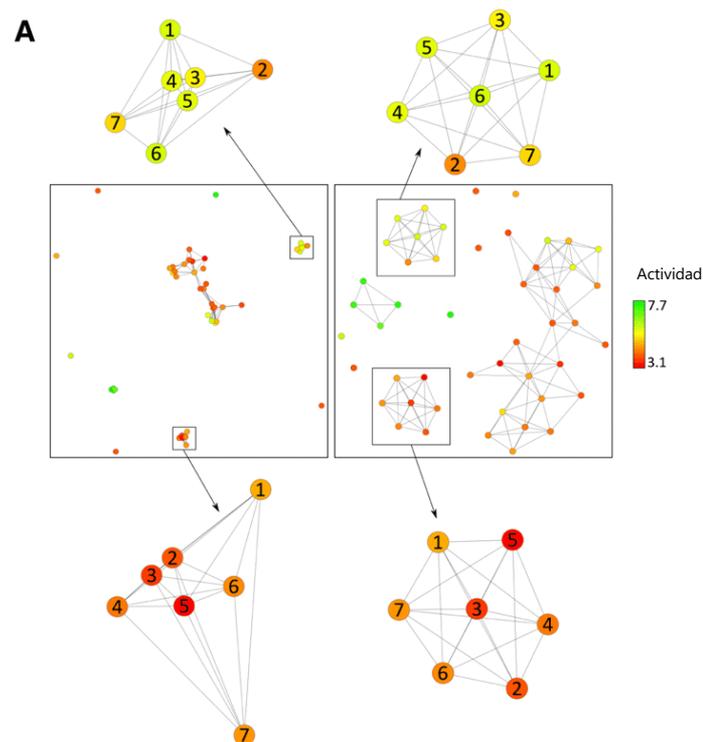
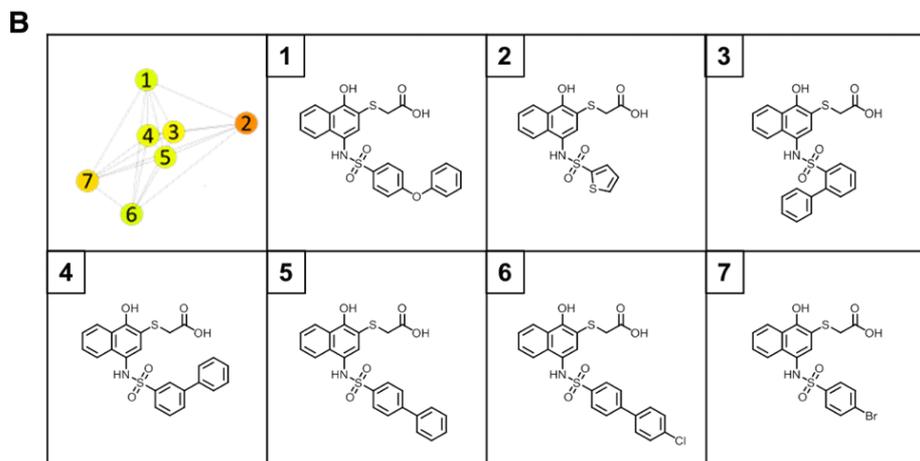
Disposición Kamada-Kawai

- Cuando se reduce el valor límite, el número de conexiones en un grupo se incrementa mucho pero no tanto el número de conexiones entre grupos

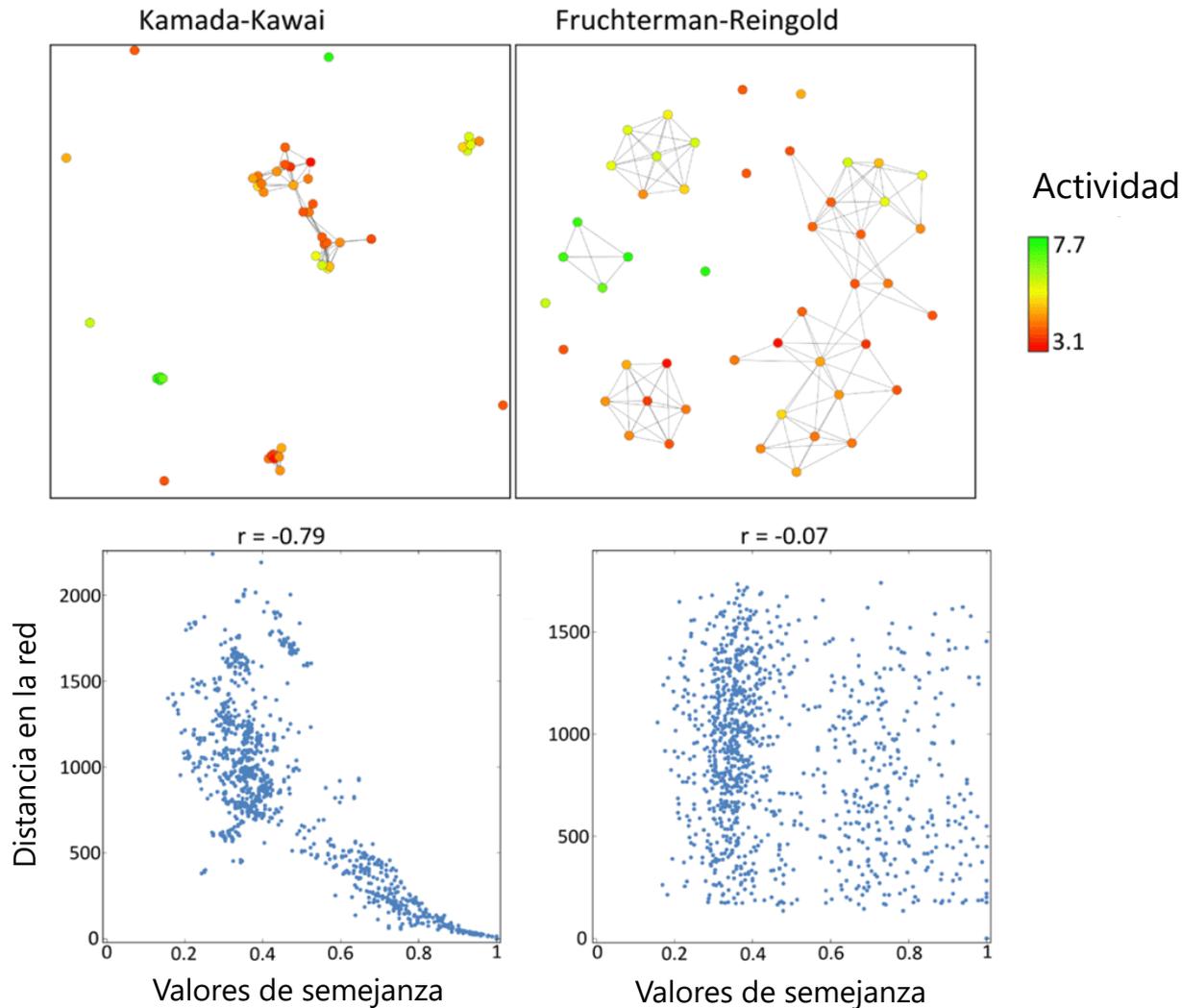


Disposición Kamada-Kawai

- La disposición Kamada-Kawai proporciona mucha más información sobre la relación entre compuestos



Disposición Kamada-Kawai



Conclusiones

- Visualización es un técnica muy útil para analizar grupos de moléculas e intentar extraer relaciones de estructura y actividad
- Las dos formas típicas de visualizar datos químicos son usar proyecciones multidimensionales o redes de compuestos

Muchas gracias por su
atención