



# Búsqueda mediante técnicas computacionales de nuevas dianas terapéuticas para la artritis reumatoide

**Jéssica Ferreira Vicente**

Máster en bioinformática y bioestadística

Descubrimiento de fármacos asistido por Computados (CADD)

**Miguel Romero Cuevas**

**María Jesús Marco Galindo**

05/06/2018



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

Copyright © 2018 Jessica Ferreira Vicente.

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

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

## FICHA DEL TRABAJO FINAL

<b>Título del trabajo:</b>	<i>Búsqueda mediante técnicas computacionales de nuevas dianas terapéuticas para la artritis reumatoide</i>
<b>Nombre del autor:</b>	<i>Jéssica Ferreira Vicente</i>
<b>Nombre del consultor/a:</b>	<i>Miguel Romero Cuevas</i>
<b>Nombre del PRA:</b>	<i>María Jesús Marco Galindo</i>
<b>Fecha de entrega (mm/aaaa):</b>	06/2018
<b>Titulación:</b>	<i>Máster en Bioinformática y Bioestadística</i>
<b>Área del Trabajo Final:</b>	<i>Descubrimiento de Fármacos Asistido por Computador (CADD)</i>
<b>Idioma del trabajo:</b>	<i>Castellano</i>
<b>Palabras clave</b>	<i>Rheumatoid Arthritis, therapeutic targets</i>
<b>Resumen del Trabajo (máximo 250 palabras):</b> <i>Con la finalidad, contexto de aplicación, metodología, resultados i conclusiones del trabajo.</i>	
<p>La artritis reumatoide (AR) es una enfermedad inflamatoria crónica de naturaleza autoinmune que afecta entre el 0,3 y el 1% de la población mundial. Los afectados por AR sufren fuertes dolores además de una pérdida de movimientos causada por la inflamación y el desgaste de las articulaciones, provocando un importante decrecimiento en la calidad de vida del paciente. Este trastorno es más frecuente en mujeres que en hombres y más habitual a edad avanzada que en sujetos jóvenes.</p> <p>Existen fármacos cuyo objetivo es reducir los síntomas y mejorar la calidad de vida de los afectados por la AR, éstos pueden ser controladores de los síntomas o modificadores de la enfermedad, en el segundo caso se logra ralentizar el avance del deterioro de las articulaciones, así como reducir la inflamación.</p> <p>Este trabajo tiene como objetivo buscar mediante técnicas computacionales nuevas dianas terapéuticas que puedan tener una mejor asociación con la AR y puedan ser nuevas vías de investigación para el desarrollo de nuevos fármacos que permitan mejorar el tratamiento de la enfermedad.</p>	
<b>Abstract (in English, 250 words or less):</b>	

Rheumatoid Arthritis (RA) is an autoimmune chronic inflammatory disease which affects between the 0,3 and the 1% of the world's population. RA patients tend to suffer pain and loss of the movements caused by joint inflammation. This disease is three times more frequent in women than in men and the probability of affection increases with the age.

There are drugs that reduce the symptoms and improve the life quality of the RA affected, these can control de symptoms or modify the disease. The second type's objective is to slow down the progress of the joint deterioration, as well as reduce the inflammation.

The main objective of this project is to discover using computational techniques and tools new therapeutic targets whose may improve the association with the RA and become new investigational ways for the development of new drugs.

## Tabla de contenido

1. Introducción .....	9
1.1 Contexto y justificación del Trabajo.....	9
1.2 Objetivos del Trabajo.....	10
1.3 Enfoque y método seguido.....	10
1.4 Planificación del Trabajo.....	11
1.5 Breve resumen de productos obtenidos .....	12
1.6 Breve descripción de los otros capítulos de la memoria .....	12
2. Metodología .....	13
2.1 Búsqueda bibliográfica .....	13
2.2 Bases de datos .....	13
2.3 Integración de los datos.....	14
2.4 Selección de los nodos centrales .....	14
2.5 Validación de los genes seleccionados .....	14
2.6 Test de viabilidad de las dianas terapéuticas .....	15
3. Resultados .....	15
3.1 Búsqueda bibliográfica .....	15
Diagnóstico de la AR .....	15
Tratamiento de la AR .....	17
3.2 Bases de datos .....	20
Disgenet .....	20
MalaCard.....	22
Kyoto Encyclopedia of Genes and Genomes (KEGG) .....	24
Comparative Toxicogenomics Database (CTD) .....	25
3.3 Integración de los datos.....	27
Análisis .....	27
Integración en Cytoscape.....	28

3.4 Selección de nodos centrales .....	30
3.5 Validación de los genes.....	30
3.6 Test de viabilidad de las dianas terapéuticas .....	32
3. Conclusiones.....	41
4. Glosario .....	42
5. Bibliografía .....	43
6. Anexos.....	45
Anexo I.....	45
Anexo II.....	49

## Lista de figuras

Figura 1. Diagrama de Gantt .....	12
Figura 2. Criterios de clasificación de la AR [15].....	16
Figura 3. Árbol de decisión para el diagnóstico de la artritis reumatoide en pacientes susceptibles a padecerla. ....	17
Figura 4. Buscador de DisGeNET. ....	20
Figura 5. Ficha de la artritis reumatoide en DisGeNET.....	20
Figura 6. Página principal de MalaCards. ....	22
Figura 7. Ficha de la AR en MalaCards.....	22
Figura 8. Página principal de KEGG. ....	24
Figura 9. Ficha de la enfermedad en KEGG. ....	24
Figura 10. Mapa del pathway de la AR.....	25
Figura 11. Página principal de cdt. ....	25
Figura 12. Ficha de la AR en cdt.....	26
Figura 13. Página principal de GeneMANIA .....	28
Figura 14. Red biológica generada por GeneMANIA.....	29
Figura 15. Visualización de la red en Cytoscape.....	29
Figura 16. Selección de nodos centrales en Cytoscape.....	30
Figura 17. Página de inicio de DrugBank. ....	30
Figura 18. Ejemplo de ficha de un gen en DrugBank.....	31
Figura 19. Pantalla de inicio de la herramienta DrugEBillity .....	33
Figura 20. Resultado de los dominios relacionados con el gen TNFAIP3 .....	33
Figura 21. Ficha del dominio seleccionado.....	34
Figura 22. Página de inicio de Swiss-model.....	34
Figura 23. Resultados de Swiss-model .....	35
Figura 24. Resultados de los templates de Swiss-model.....	35
Figura 25. Estructura del dominio seleccionado .....	36
Figura 26. Caja que engloba el ligando seleccionado para el virtual screening .....	36

Figura 27. Introducción de los datos e MTIOpenScreen para el Virtual Screening.....	37
Figura 28. Introducción de los datos de la caja de Autodock en MTIOpenScreen.....	37
Figura 29. Captura de la tabla con los mejores 1000 ligandos.....	38
Figura 30. Interacciones de los ligandos con energías de interacción más favorable.....	39
Figura 31. Superposición de los primeros 2 ligandos [1] y de los 10 primeros [2].....	40



# 1. Introducción

## 1.1 Contexto y justificación del Trabajo

La Artritis Reumatoide (AR) es una enfermedad crónica multiorgánica que generalmente afecta las articulaciones, aunque también se manifiesta en otros tejidos. También se la define como una poliartritis simétrica deformante. La AR es de naturaleza autoinmune y se da cuando el sistema inmunológico ataca las articulaciones provocando inflamación en el tejido sinovial y causando dolores e hinchazón. Si no se trata debidamente, la AR puede provocar erosión en los huesos, deformación y pérdida de funcionalidad [1].

Puede manifestarse a cualquier edad, aunque suele hacerlo entre los 30 y 60 años, con una incidencia máxima entre los 40 y 60. Las mujeres son tres veces más propensas a desarrollar AR, sobre todo en el periodo de postparto. Aunque no se conozca la causa exacta, se han reconocido algunos factores de riesgo como el historial familiar, el tabaco y la obesidad. La AR afecta entre el 0,3% y el 1,2% de la población mundial, con una prevalencia en España de entre el 0,3 y el 1,6% [2].

Al desconocer las causas es difícil, tanto diagnosticar este tipo de artritis como de hallar una cura para ella. Por lo que los tratamientos actuales se centran en aliviar el dolor, reducir la inflamación, detener o retrasar el daño articular y mejorar las funciones y bienestar del paciente. Para ello es de suma importancia la monitorización del paciente para evaluar el avance de la enfermedad y ajustar el tratamiento [3], en algunos casos se llega a recomendar al paciente una prótesis de la articulación en cuestión [4]. La tabla 1 muestra los distintos tipos de fármaco que se están usando actualmente para tratar la AR.

Tabla 1. Resumen de los tipos de fármacos para el tratamiento de la Artritis Reumatoide

<b>Tipo de fármaco</b>	<b>Objetivo</b>	<b>Ejemplos</b>
<i>Controladores de los síntomas</i>	Aliviar el dolor y reducir la inflamación	Aspirina, ibuprofeno, corticoides
<i>Modificadores de la enfermedad reumática (DMARD)</i>	Aliviar el dolor, reducir la inflamación y retrasar el daño articular	Dentro de esta categoría se hallan distintos tipos de fármaco (csDMARDs, bDMARDs, tsDMARDs)

Por lo expuesto, es necesario el desarrollo de nuevos tratamientos que puedan suponer una mejora terapéutica o una reducción de los inconvenientes de los fármacos actuales. En este sentido, el descubrimiento de fármacos es un proceso complejo con distintas etapas entre las que destacan la búsqueda de nuevas dianas biológicas contra la enfermedad, la viabilidad de dichas dianas, la búsqueda de posibles fármacos que actúen contra éstas, la validación de los mismos y ciclos iterativos de diseño y mejora. Dada toda la casuística y medios necesarios surge el Descubrimiento de Fármacos Asistido por Computadoras (CADD), que se puede considerar una disciplina que utiliza distintos métodos bioinformáticos para el descubrimiento de fármacos contra enfermedades, reduciendo de esta manera el coste, y, sobre todo, el tiempo de desarrollo [5].

Como se ha indicado, uno de los puntos importantes dentro del descubrimiento de nuevos fármacos es la búsqueda de nuevas dianas biológicas contra una enfermedad, en esta etapa está siendo de gran importancia la gran cantidad de información genómica, siendo esta además de dominio público a través de bases de datos disponibles en Internet, y las herramientas computacionales para extraer e integrar información que permiten realizar este tipo de estudio.

La finalidad del presente trabajo es la aplicación de herramientas y métodos computacionales para la búsqueda de dianas biológicas, así como su validación y propuesta como posibles dianas terapéuticas de la AR.

## **1.2 Objetivos del Trabajo**

El objetivo principal de este trabajo es buscar mediante técnicas computacionales nuevas dianas terapéuticas que puedan tener una mejor asociación con la AR y puedan ser nuevas vías de investigación para el desarrollo de nuevos fármacos que permitan mejorar el tratamiento de la enfermedad.

Para ello, se propone los siguientes objetivos específicos:

1. Análisis terapéutico de la AR.
2. Búsqueda de dianas terapéuticas en distintas fuentes de información.
3. Integración de los datos obtenidos.
4. Priorización y selección de las dianas candidatas.
5. Testar la viabilidad (druggability) de las moléculas y propuesta final.

## **1.3 Enfoque y método seguido**

Para la realización de este proyecto vamos a utilizar la información genómica, proteómica y de rutas metabólicas disponible en bases de datos públicas. Cada día se generan cantidades ingentes de información, y uno de los grandes retos de la actualidad es analizarla e integrarla para sacar conocimiento útil.

En este proyecto extraemos información relevante a la AR de las bases de datos públicas, la integramos gracias a un software con sus interrelaciones, a partir de estas relaciones escogemos nodulos (genes) centrales basados en algoritmos internos del programa relacionado con la teoría de grafos. A partir de estos genes, cargamos toda la red proteínas relacionadas, y a partir del mismo principio seleccionaremos las proteínas centrales que serán nuestras candidatas probables.

A continuación, vemos cuál es su capacidad para ser usada en terapéutica o “druggability” y si existen posibles moléculas para interaccionar con ella, y podrían ser posibles candidatos a fármacos. Se trata de una metodología muy utilizada dentro del campo del Diseño de Fármacos Asistido por Computador (CADD, por sus siglas en inglés) [6].

#### 1.4 Planificación del Trabajo

El trabajo se ha realizado en distintas partes, una primera de definición y planificación, una segunda de desarrollo y por último de cierre y presentación.

La primera parte del trabajo es la **definición de contenidos**, donde se estableció la temática y objetivos del proyecto, así como el **plan de trabajo a seguir**.

El desarrollo de trabajo se ha dividido en dos fases, siendo el objetivo de la primera analizar la información actual sobre la artritis reumatoide, tanto a nivel de diagnóstico-terapéutico como a estudios desarrollados de identificación de genes y biomoléculas relacionadas con la enfermedad. Los puntos principales son:

- **Revisión bibliográfica de libros y artículos científicos** sobre la artritis reumatoide, su diagnóstico y tratamientos actuales.
- **Identificación de genes o biomoléculas relevantes en la literatura**

La segunda fase del desarrollo es mucho más extensa que la primera, empezando por la **búsqueda en bases de datos** de la AR para hallar los genes con mayor relación, la **integración** de éstos, así como la comparación de los resultados obtenidos por los distintos métodos. La información resultante se ha preparado para la introducción en Cytoscape, que **selecciona las dianas centrales**. Finalmente, se ha estudiado la viabilidad de dichas moléculas como fármaco para la artritis reumatoide. Los puntos principales son:

- **Búsqueda en bases de datos** como, por ejemplo:
  - a. Disgenet
  - b. MalaCards Database
  - c. Kyoto Encyclopedia of Genes and Genomes
  - d. Comparative Toxicogenomics Database
  - e. Drug bank
- **Integración de la información obtenida.**
- **Selección de dianas centrales con la ayuda de Cytoscape.**
- **Validación de las dianas seleccionadas en Drugbank.**
- **Test de la viabilidad de las dianas identificadas en DrugEBility.**

Por último, se ha redactado la memoria y preparado la presentación y cierre del proyecto.

La temporización de este plan de trabajo se muestra a continuación, en la figura 1:

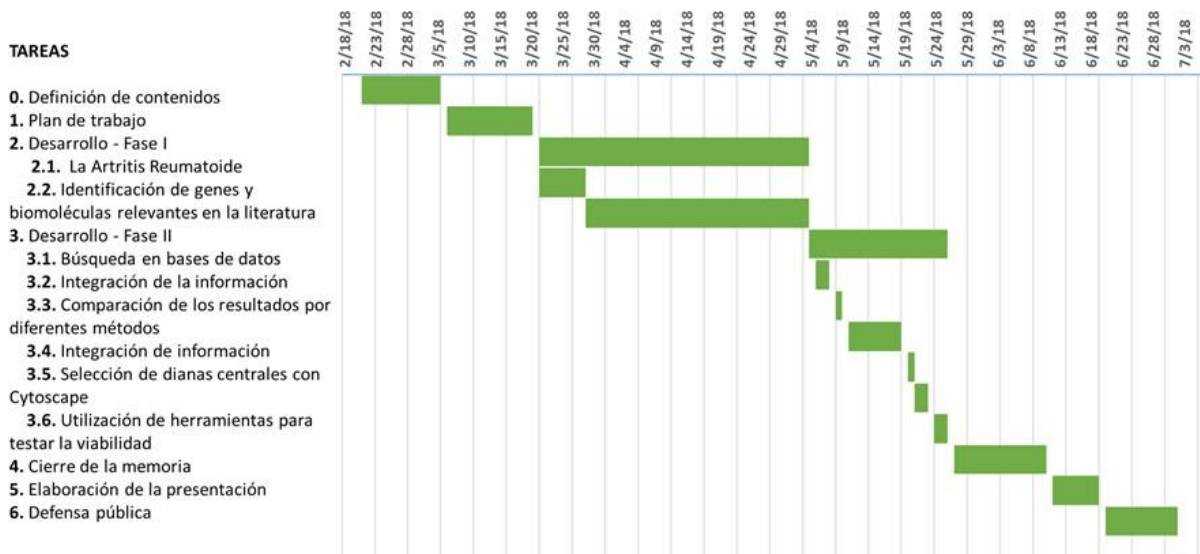


Figura 1. Diagrama de Gantt.

### 1.5 Breve resumen de productos obtenidos

Se ha propuesto un listado de genes y proteínas que podrían ser susceptibles de actuar como dianas terapéuticas contra la AR.

Del listado anterior, se han obtenido dos dianas terapéuticas no relacionadas en las bases de datos con la Artritis Reumatoide, una de ellas sin definir ni siquiera como diana.

Con éstas se han hallado 4 moléculas que se podrían proponer para el tratamiento de la AR y un dominio con dos puntos de unión y con un buen resultado como base a un posible futuro fármaco para la AR.

### 1.6 Breve descripción de los otros capítulos de la memoria

La memoria se compone por un capítulo de introducción, un capítulo donde se explica la metodología seguida, uno con los resultados obtenidos en cada paso del proceso y las conclusiones.

## 2. Metodología

### 2.1 Búsqueda bibliográfica

Para la búsqueda de recursos bibliográficos se han empleado bases de datos como, PubMed, *Google Scholar* y *Science Direct*. El objetivo de esta búsqueda ha sido entrar en contexto, conociendo la enfermedad (causas, síntomas y diagnóstico) y los tratamientos actuales.

*Google Scholar* forma parte del conjunto de servicios ofrecidos por la empresa Google, en este caso para búsquedas concretas de literatura escolar.

*PubMed* comprende más de 28 millones de citas de literatura biomédica de MEDLINE, revistas científicas y libros electrónicos, éstas pueden incluir enlaces al texto completo y a la página de la publicación.

*Science Direct* es un portal que da acceso al usuario a una extensa base de datos de publicaciones y revistas científicas. Permite buscar de una forma cómoda e intuitiva, usando filtros para aumentar la precisión de la búsqueda. Si bien las dos primeras bases de datos son de tipo referencial, la tipología de esta es de acceso al documento, permitiendo llegar de forma directa a los mismos.

### 2.2 Bases de datos

Tras finalizar el análisis terapéutico, se realiza una búsqueda en bases de datos genómicos para hallar los genes con mayor relación con la AR. Estas bases de datos son las siguientes:

- DisGeNET [7]: es una plataforma gratuita que integra información de enfermedades humanas asociada con genes. Se trata de una de las mayores colecciones de genes y variantes involucradas en enfermedades. Las entradas de estudios de asociación genética se ordenan según un sistema de puntuación propio que refleja la recurrencia de una asociación en todas las fuentes de datos. Esta puntuación se actualiza cada vez que se incluyen nuevas fuentes. Además, para facilitar el ranking de los genes asociados a la enfermedad, se usan dos métricas:
  - Índice de especificidad de la enfermedad (DSI): es inversamente proporcional al número de enfermedades asociadas a un gen particular.
  - Índice de "pleiotropía" de la enfermedad (DPI): es proporcional al número de diferentes clases de enfermedades a las que se asocia un gen.
- Malacards [8]: es una base de datos que integra información sobre las enfermedades y sus anotaciones, usando la arquitectura y estrategia de la base de datos GeneCards de genes humanos. Utiliza 44 fuentes de datos. Para cada enfermedad se extraen anotaciones específicas y conexiones entre enfermedades. Los resultados se muestran por secciones en las que aparece toda la información recolectada acerca de dicha enfermedad. Los genes relacionados con una enfermedad determinada se ordenan según su Score, un valor que tiene en cuenta la frecuencia de aparición del gen y la enfermedad.

- Kyoto Encyclopedia of Genes and Genomes (KEGG) [9]: es una base de datos que integra información genómica, química y sistémica-funcional. Por una parte, contiene la información funcional en la base de datos "PATHWAY" en forma de representaciones gráficas de procesos celulares. Los datos de las enfermedades se hallan en la base de datos "DISEASE", en la cual se pueden encontrar datos como la identificación, descripción, enlace al Pathway y fármacos relacionados.
- Comparative Toxicogenomics Database (CTD) [10]: proporciona información sobre las interacciones entre productos químicos y génicos así como su relación con las enfermedades. Incluye más de 30 millones de conexiones tóxico-genómicas.

### 2.3 Integración de los datos

Análisis de los resultados extraídos a través de las distintas metodologías seguidas e integración de los datos en el programa Cytoscape. Dicho programa se trata de un software libre para visualización de redes biológicas y análisis [11]. Esta aplicación permite la instalación de diferentes plugins externos, lo que le aporta mayores funcionalidades. Un ejemplo es el plugin de Disgenet, que facilita la interpretación y análisis de su base de datos o el plugin de Reactome.

Para la integración de los datos se ha planteado la siguiente metodología:

1. Unir los 15 genes más relevantes de cada base de datos.
2. Generar la red biológica mediante GeneMANIA.
3. Introducir y analizar esta red en Cytoscape para seleccionar los genes con más interacciones.

GeneMANIA es una aplicación online para generar hipótesis sobre funciones de los genes, analizar listas de genes y priorizar para ensayos funcionales. Para ello es necesario introducir una lista con el id de los genes [12], también existe un plugin para Cytoscape.

### 2.4 Selección de los nodos centrales

Se seleccionan los nodos centrales aplicando un filtro en Cytoscape. El objetivo es localizar los genes con mayor número de interacciones.

### 2.5 Validación de los genes seleccionados

Tras seleccionar los genes, se ha realizado la búsqueda de éstos en la base de datos de DrugBank para:

1. Conocer si el gen consta como diana terapéutica para la artritis reumatoide
2. Saber si hay algún fármaco asociado al gen

DrugBank (<https://www.drugbank.ca>) es una base de datos con acceso online que contiene información molecular integrada sobre fármacos, sus mecanismos, interacciones y dianas [13]. En este punto se seleccionarán las dianas ya descritas y sus fármacos. Se seleccionan las moléculas no descritas como diana para realizar un estudio de viabilidad así como la propuesta

de nuevos fármacos. A través de esta base de datos se tiene acceso a posibles fármacos descritos contra la diana.

## 2.6 Test de viabilidad de las dianas terapéuticas

Para ello se usa la herramienta DrugEBllity (<https://www.ebi.ac.uk/chembl/drugability/blast/search>) del European Bioinformatics Institute (EBI). Se trata de un buscador de “drugability” basado en la estructura molecular. Esta herramienta permite encontrar y clasificar dominios que podrán usarse para el tratamiento de la diana seleccionada.

Dentro de la lista de dominios posibles, se selecciona el que tenga mejor puntuación y que, a la vez, sea “tractable” y “druggable”.

## 2.7. Búsqueda de moléculas que podrían interactuar con las dianas.

Para aquellas dianas no descritas como tales, se realizará un estudio de búsqueda de moléculas que podrían interactuar con ellas y ser candidatas a fármacos. Para ellos, con la secuencia del dominio procedemos a la propuesta de un modelo 3D para poder realizar un screening virtual de posibles moléculas que pudieran unirse al mismo. El modelo se creará mediante Swiss Model.

Swiss-model es un servidor que permite realizar modelos comparativos de estructuras tridimensionales de proteínas de forma automática. Éste permite diseñar experimentos de manera efectiva como mutagénesis dirigida, estudios de mutaciones relacionadas con enfermedades o el diseño de inhibidores específicos basado en estructura.

A partir de esta estructura, se busca primero mediante Autodock la zona que se desea analizar y con la herramienta MTIOpenScreen se hallan posibles ligandos para el dominio identificado inicialmente por Swiss-model. MTIOpenScreen es un servidor web para virtual screening basado en la estructura, facilita la identificación de nuevos compuesto bioactivos para la innovación y mejora de la efectividad de los procesos bioquímicos y de descubrimiento de nuevos fármacos [14].

# 3. Resultados

## 3.1 Búsqueda bibliográfica

El objetivo principal de la búsqueda bibliográfica era conocer la enfermedad, pero principalmente los tratamientos actuales y dianas terapéuticas ya descubiertas.

Diagnóstico de la AR

En 2010 se establecieron unos criterios consensuados entre el *American College of Rheumatology* (ACR) y la *European League Against Rheumatism* (EULAR) siguiendo un estudio de seguimiento de pacientes además de datos externos. Los criterios de diagnosis de la artritis reumatoide pasan por tres puntos [15]:

1. Presencia confirmada por análisis de sangre y radiografías de sinovitis en por lo menos una articulación.
2. Ausencia de un diagnóstico alternativo que explique mejor la sinovitis.

3. Alcanzar un mínimo de 6 puntos de 10 posibles de un interrogatorio (Figura 2) dividido en 4 dominios:
  - a. Número y localización de las articulaciones afectadas (0 – 5 puntos)
  - b. Anormalidad serológica (0 – 3 puntos)
  - c. Elevada respuesta en la fase aguda (0 – 1 punto)
  - d. Duración de los síntomas (2 niveles, 0 – 1 punto)

	Score
Target population (Who should be tested?): Patients who	
1) have at least 1 joint with definite clinical synovitis (swelling)*	
2) with the synovitis not better explained by another disease†	
Classification criteria for RA (score-based algorithm: add score of categories A–D; a score of $\geq 6/10$ is needed for classification of a patient as having definite RA)‡	
A. Joint involvement§	
1 large joint¶	0
2–10 large joints	1
1–3 small joints (with or without involvement of large joints)#	2
4–10 small joints (with or without involvement of large joints)	3
>10 joints (at least 1 small joint)**	5
B. Serology (at least 1 test result is needed for classification)††	
Negative RF <i>and</i> negative ACPA	0
Low-positive RF <i>or</i> low-positive ACPA	2
High-positive RF <i>or</i> high-positive ACPA	3
C. Acute-phase reactants (at least 1 test result is needed for classification)‡‡	
Normal CRP <i>and</i> normal ESR	0
Abnormal CRP <i>or</i> abnormal ESR	1
D. Duration of symptoms§§	
<6 weeks	0
$\geq 6$ weeks	1

Figura 2. Criterios de clasificación de la AR [15]

Finalmente, para diagnosticar la AR se utiliza un algoritmo del tipo árbol de decisión que responde a las anteriores preguntas de un modo práctico y sencillo. Este algoritmo se puede observar en la figura 3.



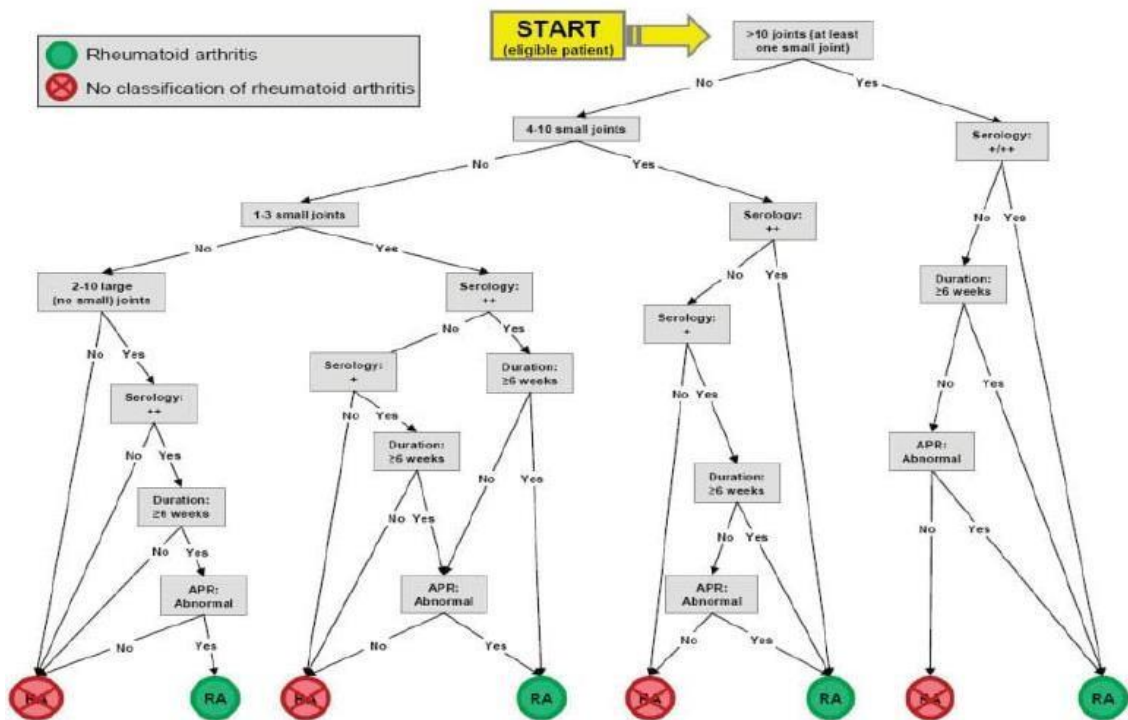


Figura 3. Árbol de decisión para el diagnóstico de la artritis reumatoide en pacientes susceptibles a padecerla.

### Tratamiento de la AR

Es importante tratar la artritis reumatoide desde etapas tempranas para reducir el riesgo a que las articulaciones se erosionen, deformen o incluso pierdan su funcionalidad. Los tratamientos administrados, se centran en reducir los síntomas de la enfermedad retrasando el daño articular y aliviando el dolor. Los fármacos pueden ser de tres tipos, los fármacos que controlan los síntomas o los modificadores de la enfermedad (DMARD).

- Medicamentos para controlar los síntomas [16]:
  - AINE o antiinflamatorios no esteroides: Se usan para aliviar el dolor y combatir la inflamación. Por ejemplo: aspirina, ibuprofeno, naproxeno, etc.
  - Analgésicos: Se usan para aliviar el dolor, pero no tienen efecto sobre la inflamación producida por el daño articular. Por ejemplo: acetaminofeno, tramadol, propoxifeno, etc.
  - Corticosteroides: Alivian los síntomas de la AR, se pueden tomar en bajas dosis junto con los demás tipos de medicamentos. Tiene muchos beneficios y también efectos secundarios, el objetivo es tomar una dosis efectiva y baja que evite los efectos secundarios.
- Medicamentos modificadores de enfermedades reumáticas (DMARD Disease-Modifying Anti-Rheumatic Drugs): tienen la capacidad de reducir y hasta revertir síntomas, progresión de los daños articulares e inhabilidad para trabajar, todo ello interfiriendo en el proceso de la enfermedad. Los hay de dos clases mayoritarias componentes químicos sintéticos (sDMARDs) y agentes biológicos (bDMARDs).

Recientemente se ha propuesto un cambio en la nomenclatura de los sDMARDs, clasificándolos entre convencionales (csDMARDs) y dirigidos (tsDMARDs) [4].

El resumen de fármacos clasificados como modificadores de enfermedad reumática y las terapias biológicas se muestran en la siguiente tabla:

Tabla 2. Fármacos actuales para el tratamiento de la AR.

<b>Tipo fármaco</b>	<b>Nombre</b>	<b>Acción primaria</b>	<b>Diana</b>	<b>Síntomas</b>	<b>Efectos adversos</b>
csDMARD	Azathioprine	Inmuno supresor. Bloquea el metabolismo de la purina y la síntesis del ADN	HPRT1 y RAC1	Reducir la inflamación	Mayor riesgo de contraer infecciones o, en casos poco frecuentes, linfoma
csDMARD	Certolizumab Pegol	Anticuerpo humano anti TNF	TNF	Reducir la inflamación	
csDMARD	Hydroxychloroquine	Desconocido	TLR7 y TLR9	Alivia la inflamación, la hinchazón y el dolor de las articulaciones.	Daños en la retina, cefalea, náuseas.
csDMARD	Leflunomide	Inhibe la síntesis de la pirimidina	DHODH, AHR y PTK2B	Reducir la inflamación	Anormalidades en la función del hígado, fallo hepático, aumento de nivel de toxicidad hepática, pérdida de peso y diarrea.
csDMARD	Methotrexate	Es un antimetabólito antineoplástico con propiedades inmuno supresoras	DHFR	Reducir la inflamación	Anormalidades en la función del hígado y de la médula ósea, daño hepático, malestar gástrico, cefalea, náuseas y diarrea.
bDMARD	Abatacept	Inhibe la coestimulación de las células T	CD80 y CD86	Reducir la inflamación	Cefalea, infección de las vías respiratorias altas, nasofaringitis y náuseas.
bDMARD	Adalimumab	Anticuerpo humano anti TNF alpha	TNF	Reducir la inflamación	No disponible
bDMARD	Anakinra	Antagonista del receptor de Interleukina-1	IL1	Reducir la inflamación	Irritación de la piel en la zona de la inyección, riesgo de infección bacteriana.

<i>bDMARD</i>	Etanercept	Inhibidor TNF alpha y beta	TNF	Reducir la inflamación	Infecciones, cancer, vasculitis, enfermedades hepáticas, anomalías hepáticas como anemia aplásica y linfoma, alergia severa y meningitis aséptica
<i>bDMARD</i>	Golimumab	Anticuerpo monoclonal del TNF	TNF	Reducir la inflamación	Infecciones graves, sepsis bacteriana, tuberculosis. En niños y adolescentes también linfoma
<i>bDMARD</i>	Infliximab	Anticuerpo quimérico anti TNF alpha	TNF	Reducir la inflamación	Todavía no se han establecido los efectos adversos en humanos.
<i>bDMARD</i>	Rituximab	Anti célula B	MS4A1, FCGR3B, C1R, C1Q	Reducir la inflamación	No disponible
<i>bDMARD</i>	Tocilizumab	Anticuerpo monoclonal bloqueador del receptor (IL-6R)	IL6R	Reducir la inflamación	Infecciones de las vías respiratorias altas, nasofaringitis, cefalea e hipertensión.
<i>tsDMARD</i>	Baricitinib	Inhibidor selectivo y reversible de quinasas Janus	JAK1, JAK2, PTK2B, JAK3	Reducir la inflamación	Todavía no se han establecido los efectos adversos en humanos.
<i>tsDMARD</i>	Tofacitinib	Inhibidor de quinasas Janus	JAK1, JAK2, JAK3, TYK2, CYP3A4, CYP2C19 y ALB	Reducir la inflamación	Todavía no se han establecido los efectos adversos en humanos.

El primer fármaco que se recomienda a un paciente con AR son los csDMARDs, si estos no funcionan, se pueden realizar terapias biológicas [17].

Las terapias biológicas (bDMARDs) actuales tienen una alta eficacia, sin embargo, entre el 30 y el 50% de los pacientes no responden debidamente a dicho tratamiento. Uno de los principales motivos es que la AR provoca, en algunos casos, una compleja desregulación del sistema inmunológico, por lo que la estrategia de las terapias biológicas no es siempre exitosa, ya su diana son citoquinas pro-inflamatorias específicas y sus receptores. Otros motivos son los efectos secundarios que afectan aproximadamente un 14% de los pacientes, el modo de

aplicación (inyecciones) y el elevado coste de este tipo de tratamiento en comparación con fármacos tradicionales, si un bDMARD no funciona, no se puede probar otro fármaco de la misma clase [17].

### 3.2 Bases de datos

Se han consultado todas las bases mencionadas en el apartado de metodología: “2.2 Bases de datos”. Los resultados se muestran a continuación:

Disgenet

Para la búsqueda en disgenet es necesario acceder a su página web (<http://www.disgenet.org/>) y seleccionar la pestaña “Search” (Figura 4).

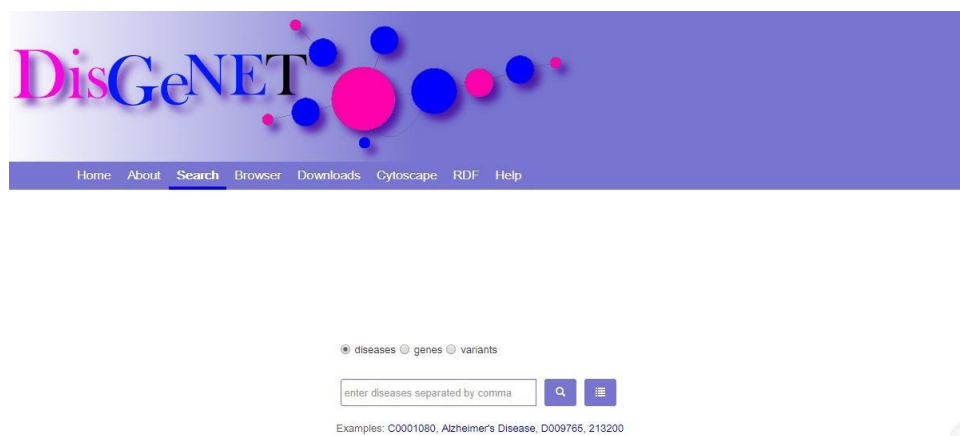


Figura 4. Buscador de DisGeNET.

Hay tres posibles opciones: enfermedades, genes y variantes. Seleccionamos enfermedades e introducimos Rheumatoid Arthritis, lo que nos lleva a la ficha de la enfermedad (Figura 5).

**Rheumatoid Arthritis**

**UMLS CUI** C0003873

**Type** disease

**MeSH Class** Immune System Diseases ; Musculoskeletal Diseases ; Skin and Connective Tissue Diseases

**MeSH** D001172

**OMIM** 607218,180300

**Semantic Type** Disease or Syndrome

**Phenotypes** Abnormality of the skeletal system

**Disease Ontology** disease of anatomical entity

[Top 10 gene associations for this disease](#)

[Top 10 diseases that share genes with this disease](#)

[Top 10 SNPs for this disease](#)

[All evidences for this disease](#)

[All diseases that share genes with this disease](#)

Figura 5. Ficha de la artritis reumatoide en DisGeNET.

Los resultados obtenidos en la ficha son, primero, una pequeña descripción de la enfermedad junto con distintos códigos de identificación según algunas fuentes, seguido por los TOP 10 de los genes asociados, de las enfermedades que comparten genes y los SNPs relacionados con la enfermedad y, por último, todas las evidencias y enfermedades que comparten genes con la AR.

En este caso seleccionamos la primera opción “*Top 10 gene associations for this disease*” y una vez dentro, se descarga una tabla de Excel con los resultados obtenidos, en este caso 1848 genes. Para facilitar este análisis, seleccionamos los 15 primeros (Tabla 3):

Tabla 3. Top 15 de genes relacionados con la AR obtenidos a través de DisGeNET.

<b>Gene</b>	<b>Gene Full Name</b>	<b>DPI</b>	<b>DSI</b>	<b>Score</b>
<i>PTPN22</i>	protein tyrosine phosphatase, non-receptor type 22	0,821	0,458	0,567
<i>TNF</i>	tumor necrosis factor	0,964	0,251	0,506
<i>HLA-DRB1</i>	major histocompatibility complex, class II, DR beta 1	0,964	0,317	0,440
<i>IL1B</i>	interleukin 1 beta	0,964	0,3	0,323
<i>CRP</i>	C-reactive protein	0,929	0,381	0,315
<i>CTLA4</i>	cytotoxic T-lymphocyte associated protein 4	0,929	0,395	0,299
<i>PADI4</i>	peptidyl arginine deiminase 4	0,643	0,557	0,298
<i>IL6</i>	interleukin 6	0,964	0,275	0,297
<i>MTHFR</i>	methylenetetrahydrofolate reductase	0,964	0,325	0,287
<i>IL10</i>	interleukin 10	0,964	0,297	0,286
<i>IL6ST</i>	interleukin 6 signal transducer	0,786	0,523	0,284
<i>STAT4</i>	signal transducer and activator of transcription 4	0,786	0,489	0,277
<i>TRAF1</i>	TNF receptor associated factor 1	0,643	0,565	0,271
<i>IL1RN</i>	interleukin 1 receptor antagonist	0,964	0,359	0,259
<i>TNFAIP3</i>	TNF alpha induced protein 3	0,75	0,545	0,243

Los resultados obtenidos demuestran la baja homogeneidad en los datos para la AR en las distintas fuentes de datos, con un Score máximo de 0,57, además los genes muestran un bajo índice de especificidad y un elevado índice de pleiotropía, lo que indica que no son específicos de la AR y que están relacionados con diversas clases de enfermedad.

## MalaCard

Se accede a través del siguiente enlace: <http://www.malacards.org/>. Cuya página principal se muestra en la Figura 6.

The screenshot shows the MalaCards website homepage. At the top, there is a navigation bar with the MalaCards logo and the text 'HUMAN DISEASE DATABASE'. To the right, there are logos for 'WEIZMANN INSTITUTE OF SCIENCE' and 'LifeMap SCIENCES'. A search bar is located in the center of the top bar. Below the navigation bar, there are several tabs: 'Home', 'User Guide', 'Analysis Tools', 'News and Views', 'Disease Lists/Categories', and 'About'. On the right side, there are 'Log In' and 'Sign Up' buttons. The main content area features the title 'MalaCards: The human disease database' and a brief description of the database. A network diagram titled 'Disease network for Acute Sphenoidal Sinusitis' is displayed, showing connections between various conditions like ophthalmoplegia, blindness, cavernous sinus thrombosis, etc. On the right side, there are sections for 'NGS Analysis Tools' (TGex, VarElect) and 'Affiliated Databases' (GeneCards, LifeMap, PathCards, GeneLoc).

Figura 6. Página principal de MalaCards.

En el buscador se introduce la enfermedad, en este caso “Rheumatoid Arthritis”. Tras seleccionar la entrada de la AR, se abre la ficha de la enfermedad (Figura 7).

The screenshot shows the MalaCards disease page for Rheumatoid Arthritis. The page header includes the MalaCards logo and the text 'HUMAN DISEASE DATABASE'. The title 'Rheumatoid Arthritis' is prominently displayed, along with its MCID (RHM011) and MIFTS (91). Below the title, there are categories: 'Rare diseases, Bone diseases, Respiratory diseases, Genetic diseases'. A navigation bar at the top of the page lists various sections: 'Genes', 'Variations', 'Tissues', 'Related diseases', 'Publications', 'Pathways', 'Symptoms & Phenotypes', and 'Drugs'. A 'Download' button and 'Expand all tables' option are also present. The main content area is titled 'Aliases & Classifications for Rheumatoid Arthritis' and shows a list of aliases and classifications, including 'Arthritis or Polyarthritis, Rheumatic' and 'Atrophic Arthritis'. A sidebar on the left allows users to jump to different sections of the page.

Figura 7. Ficha de la AR en MalaCards.

Los resultados se muestran por secciones, siendo la primera un resumen de las fuentes de información en la que aparece la clasificación y descripción de la AR según cada una. Seguido de los resultados de fármacos, expresión y genes.

Para exportar los resultados en formato de hoja de cálculo, es necesario registrarse e iniciar sesión.

Se obtiene un total de 114 genes, de los cuales seleccionamos los 15 primeros, que se muestran en la Tabla 4.

Tabla 4. Top 15 de genes relacionados con la AR obtenidos a través de MalaCards.

<b>Symbol</b>	<b>Description</b>	<b>Score</b>
<i>PTPN22 (Elite Gene)</i>	Protein Tyrosine Phosphatase, Non-Receptor Type 22	492.18
<i>SLC22A4 (Elite Gene)</i>	Solute Carrier Family 22 Member 4	154.64
<i>CIITA (Elite Gene)</i>	Class II Major Histocompatibility Complex Transactivator	152.78
<i>PADI4 (Elite Gene)</i>	Peptidyl Arginine Deiminase 4	147.79
<i>NFKBIL1 (Elite Gene)</i>	NFKB Inhibitor Like 1	147.51
<i>HLA-DRB1 (Elite Gene)</i>	Major Histocompatibility Complex, Class II, DR Beta 1	147.35
<i>IL10 (Elite Gene)</i>	Interleukin 10	145.81
<i>CD244 (Elite Gene)</i>	CD244 Molecule	137.57
<i>IL6</i>	Interleukin 6	70.06
<i>CCL2</i>	C-C Motif Chemokine Ligand 2	63.90
<i>TNF</i>	Tumor Necrosis Factor	61.86
<i>IL17A</i>	Interleukin 17A	57.89
<i>IL1A</i>	Interleukin 1 Alpha	53.58
<i>FCRL3</i>	Fc Receptor Like 3	53.34
<i>STAT4</i>	Signal Transducer And Activator Of Transcription 4	52.88

Los genes clasificados como “elite genes” son los que podrían estar más asociados con causar la enfermedad. En este caso, 8 de los 15 genes se clasifican como tal y se corresponden con los que reportan un mayor valor de Score. A mayor Score, mayor relevancia tiene el gen con respecto a la enfermedad.

## Kyoto Encyclopedia of Genes and Genomes (KEGG)

Se accede a través de <http://www.kegg.jp/kegg/kegg2.html> donde se selecciona la opción DISEASE para obtener los resultados de la enfermedad. La pantalla de inicio se muestra en la Figura 8.



### KEGG - Table of Contents

Menu **PATHWAY** BRITE MODULE KO GENES LIGAND NETWORK DISEASE DRUG DBGET

Search **DISEASE** for

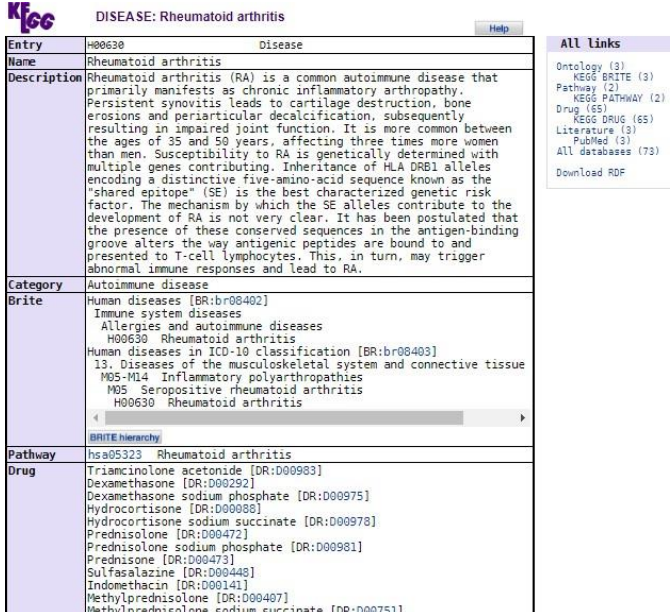
#### Data-oriented entry points

Category	Entry Point	Content	DBGET Search
Systems information	<b>KEGG PATHWAY</b>	KEGG pathway maps	PATHWAY
	<b>KEGG BRITE</b>	BRITE hierarchies and tables	BRITE
	<b>KEGG MODULE</b>	KEGG modules	MODULE
Genomic information	<b>KO (KEGG Orthology)</b>	Functional orthologs	ORTHOLOGY
	<b>KEGG GENOME</b>	KEGG organisms (complete genomes)	GENOME
	<b>KEGG GENES</b> <b>KEGG SSDB</b>	Genes and proteins GENES sequence similarity	GENES
Chemical information (KEGG LIGAND)	<b>KEGG COMPOUND</b>	Small molecules	COMPOUND
	<b>KEGG GLYCAN</b>	Glycans	GLYCAN
	<b>KEGG REACTION</b>	Reactions and reaction classes	REACTION RCLASS
	<b>KEGG ENZYME</b>	Enzyme nomenclature	ENZYME
Health information	<b>KEGG NETWORK</b>	Disease-related network elements	NETWORK VARIANT
	<b>KEGG DISEASE</b>	Human diseases	DISEASE
	<b>KEGG DRUG</b>	Drugs and drug groups	DRUG DGROUP
	<b>KEGG ENVIRON</b>	Health related substances	ENVIRON
	<b>KEGG MEDICUS</b>	Japanese drug labels (JAPIC) FDA drug labels (DailyMed)	

KEGG is developed by Kanehisa Laboratories. See Kanehisa et al. (2017) for updates of KEGG.

Figura 8. Página principal de KEGG.

En este caso la información obtenida tras realizar dicha búsqueda es una ficha de la enfermedad en la que aparece la descripción de la AR, categoría, pathway y fármacos usados. Tal como se aprecia en la Figura 9.



**KEGG** DISEASE: Rheumatoid arthritis Help

<b>Entry</b>	H00630 Disease
<b>Name</b>	Rheumatoid arthritis
<b>Description</b>	Rheumatoid arthritis (RA) is a common autoimmune disease that primarily manifests as chronic inflammatory arthropathy. Persistent synovitis leads to cartilage destruction, bone erosions and periarticular decalcification, subsequently resulting in impaired joint function. It is more common between the ages of 35 and 50 years, affecting three times more women than men. Susceptibility to RA is genetically determined with multiple genes contributing. Inheritance of HLA DRB1 alleles encoding a distinctive five-amino-acid sequence known as the "shared epitope" (SE) is the best characterized genetic risk factor. The mechanism by which the SE alleles contribute to the development of RA is not very clear. It has been postulated that the presence of these conserved sequences in the antigen-binding groove alters the way antigenic peptides are bound to and presented to T-cell lymphocytes. This, in turn, may trigger abnormal immune responses and lead to RA.
<b>Category</b>	Autoimmune disease
<b>Brite</b>	Human diseases [BR:br08402] Immune system diseases Allergies and autoimmune diseases H00630 Rheumatoid arthritis Human diseases in ICD-10 classification [BR:br08403] 13. Diseases of the musculoskeletal system and connective tissue M05-M14 Inflammatory polyarthropathies M05 Seropositive rheumatoid arthritis H00630 Rheumatoid arthritis
<b>Pathway</b>	hsa05323 Rheumatoid arthritis
<b>Drug</b>	Triamcinolone acetonide [DR:D009883] Dexamethasone [DR:D002921] Dexamethasone sodium phosphate [DR:D00975] Hydrocortisone [DR:D00088] Hydrocortisone sodium succinate [DR:D00978] Prednisolone [DR:D00472] Prednisolone sodium phosphate [DR:D00981] Prednisone [DR:D00473] Sulfasalazine [DR:D00448] Indomethacin [DR:D00141] Methylprednisolone [DR:D00407] Methylprednisolone sodium succinate [DR:D00751]

**All links**

- Ontology (3)
- KEGG BRITE (3)
- Pathway (2)
- KEGG PATHWAY (2)
- Drug (65)
- KEGG DRUG (65)
- Literature (3)
- PubMed (3)
- All databases (73)
- Download RDF

Figura 9. Ficha de la enfermedad en KEGG.



En este caso no aparecen los genes asociados con la enfermedad. Como dato interesante sí que aparece el Pathway (Figura 10), en el que se aprecian los genes que intervienen y los momentos en los que lo hace.

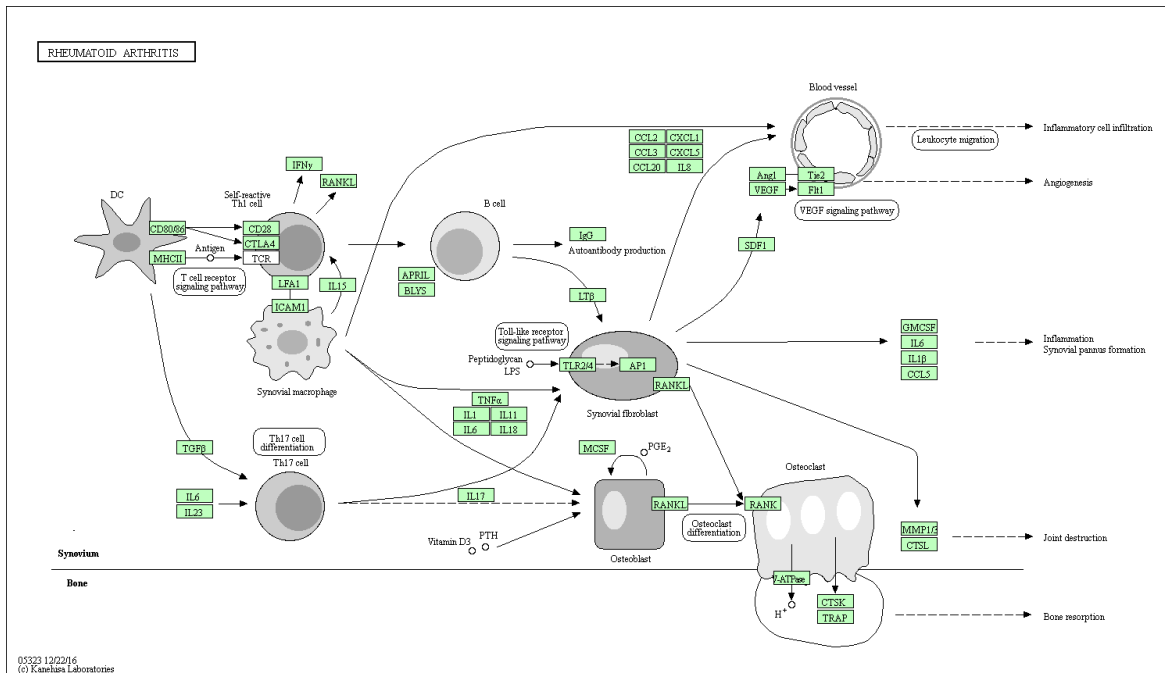


Figura 10. Mapa del pathway de la AR.

La ficha del Pathway (Figura 11) muestra los datos de identificación y descripción de la enfermedad, la miniatura del mapa, las enfermedades que comparten este proceso, los fármacos y la lista de los genes que forman parte de éste, en este caso son 90.

### Comparative Toxicogenomics Database (CTD)

El acceso a esta base de datos se encuentra en el siguiente enlace: <http://ctdbase.org/>. A la derecha de la página principal (Figura 11) se halla el buscador, en el que seleccionaremos la opción "Diseases".

La imagen muestra la interfaz de usuario de la Comparative Toxicogenomics Database (CTD). En la parte superior, se encuentra el logo de CTD con el lema "Illuminating how chemicals affect human health." y el menú de navegación: Home, Search, Analyze, Download, Help. A la derecha, hay un enlace para "YOUR QUERIES | CONTACT US".

El contenido principal incluye:

- Connect. Compare.** Descripción de CTD como una base de datos robusta que busca avanzar el entendimiento de cómo las exposiciones ambientales afectan la salud humana.
- Discover.** Una lista de preguntas orientadas a la investigación, como "¿Qué enfermedades humanas están asociadas con un gen/proteína?" y "¿Qué genes/proteínas interactúan con un químico?".
- 37,781,030 TOXICOGENOMIC RELATIONSHIPS!** Un recuento de las relaciones almacenadas en la base de datos.
- Keyword Search:** Una sección para buscar por enfermedades, con un campo de entrada para el ID y un botón de búsqueda.
- Updated Chemicals:** Una lista de sustancias químicas recientemente agregadas, incluyendo Acetaminophen, Methyl Methanesulfonate, y otros.

Figura 11. Página principal de cdt.

Una vez abierta la ficha de la enfermedad (Figura 12), la información se dispone en forma de pestañas. Cada una contiene un tipo de información.

**Arthritis, Rheumatoid**

**Name** ? Arthritis, Rheumatoid

**Synonyms** ? RA | Rheumatoid Arthritis | RHEUMATOID ARTHRITIS, SUSCEPTIBILITY TO

**Definition** ? A chronic systemic disease, primarily of the joints, marked by inflammatory changes in the synovial membranes and articular structures, widespread fibrinoid degeneration of the collagen fibers in mesenchymal tissues, and by atrophy and rarefaction of bony structures. Etiology is unknown, but autoimmune mechanisms have been implicated.

**Categories** ? Connective tissue disease | Immune system disease | Musculoskeletal disease

**MeSH<sup>®</sup> ID** ? **D001172**

**OMIM<sup>®</sup> ID** ? 180300

**External Links** DO ?  
DOID:7148

**Ancestors** ? Top ↑

1. Diseases ← Musculoskeletal Diseases ⚠ ⚡ ← Joint Diseases ⚠ ⚡ ← Arthritis ⚠ ⚡ ⚡  
← Arthritis, Rheumatoid ⚠ ⚡ ⚡
2. Diseases ← Musculoskeletal Diseases ⚠ ⚡ ⚡ ← Rheumatic Diseases ⚠ ⚡ ⚡ ←

Figura 12. Ficha de la AR en cdt.

La lista de genes relacionados se puede extraer en formato de hoja de cálculo de Excel, en total se dispone de 28425 entradas. La tabla 5 muestra el top 15 de genes asociados a la AR.

Tabla 5. Top 15 de genes asociados a la AR obtenidos a través de cdtbase.

Gene Symbol	Gene ID	Inference Score	Reference Count
IL1B	3553	98,88	455
IFNG	3458	98,56	436
IL6	3569	93,82	464
PTGS2	5743	87,82	306
IL10	3586	85,71	412
CXCR4	7852	68,93	355
SOD2	6648	68,65	239
CXCL8	3576	67,79	274
TLR2	7097	65,53	206
VEGFA	7422	65,08	220
MPO	4353	63,35	392
PTGS1	5742	60,61	259
IL23A	51561	56,82	350
CAT	847	56,54	407
STAT1	6772	55,22	213

En este caso, los genes se disponen según su “Inference Score”, valor que refleja el grado de similitud entre la red químico-gen-enfermedad y una red aleatoria. Cuán mayor sea este valor, más probable es que la red de inferencia tenga una conectividad atípica [18].

### 3.3 Integración de los datos

#### Análisis

Se analizan los resultados obtenidos a partir de las bases de datos DisGeNET, MalaCards y cdt. Para las tres se dispone de una hoja de cálculo con los resultados obtenidos tras la búsqueda de la AR.

Al unir los genes obtenidos mediante las 3 bases de datos, se obtiene un total de 30387, de los cuales, únicos quedan 21899. Para facilitar el análisis, éste se llevará a cabo juntando los más relevantes de cada base de datos, en este caso 45 genes totales. Puesto que cada base de datos se rige por unos parámetros de puntuación, es de especial interés incluir en el estudio los genes más relevantes de cada fuente. Una vez generada esta tabla, se eliminan los genes duplicados.

*Tabla 6. Genes asociados con la AR según las distintas bases de datos con los valores de referencia que usa cada una de ellas.*

<b>Gene</b>	<b>Disgenet</b>	<b>Malacards</b>	<b>cdt</b>
PTPN22	0,56707601	492,18	
TNF	0,5055899	61,86	
HLA-DRB1	0,44	147,35	
IL1B	0,32268422		98,88
CRP	0,31475301		
CTLA4	0,29911346		
PADI4	0,29772256	147,79	
IL6	0,29682814	70,06	93,82
MTHFR	0,28663576		
IL10	0,28570636	145,81	85,71
IL6ST	0,28405539		
STAT4	0,27679844	52,88	
TRAF1	0,27059639		
IL1RN	0,25934863		
TNFAIP3	0,2429736		
SLC22A4		154,64	
CIITA		152,78	
NFKBIL1		147,51	
CD244		137,57	
CCL2		63,9	
IL17A		57,89	
IL1A		53,58	
FCRL3		53,34	
IFNG			98,56
PTGS2			87,82
CXCR4			68,93
SOD2			68,65
CXCL8			67,79

TLR2			65,53
VEGFA			65,08
MPO			63,35
PTGS1			60,61
IL23A			56,82
CAT			56,54
STAT1			55,22

En total se obtiene una lista de 35 genes únicos, de los cuales únicamente 2 de ellos se repiten en las tres bases de datos (IL6 e IL10). El número de genes comunes entre las distintas bases de datos se muestra en la siguiente tabla:

Tabla 7. Número de genes que se repiten en las distintas bases de datos.

DisGeNET + MalaCards + cdt	DisGeNET + MalaCards	DisGeNET + cdt	MalaCards + cdt
2	7	3	2

Para las bases de DisGeNET y MalaCards se obtienen 7 genes repetidos de los 30 totales. La base de datos con resultados más dispares es cdt.

### Integración en Cytoscape

Los 35 genes únicos se introducen en la GeneMANIA accediendo a través del enlace siguiente: <https://genemania.org/>, cuya página principal se muestra en la figura 13.

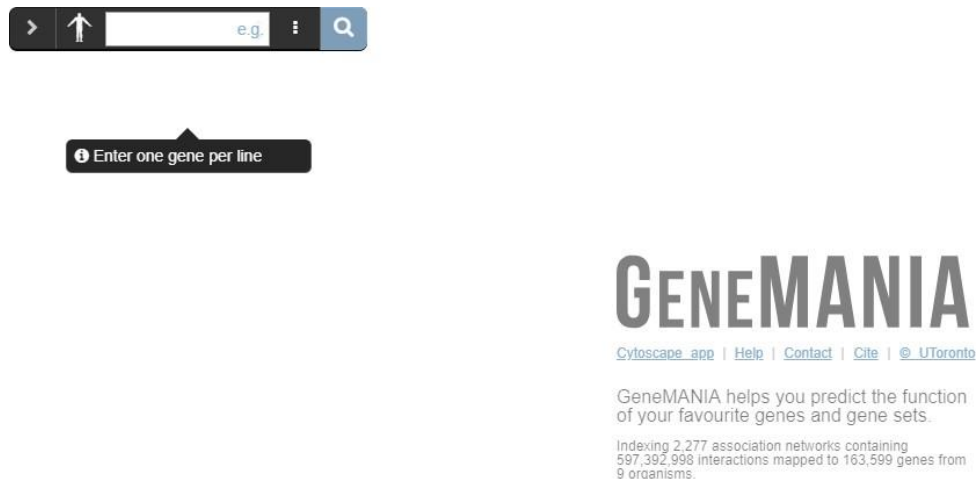


Figura 13. Página principal de GeneMANIA.

Tras introducir la lista de los genes en la aplicación, se genera una red biológica (Figura 14) con los genes introducidos y unos genes añadidos por la propia aplicación al mostrar interacciones con los datos introducidos.

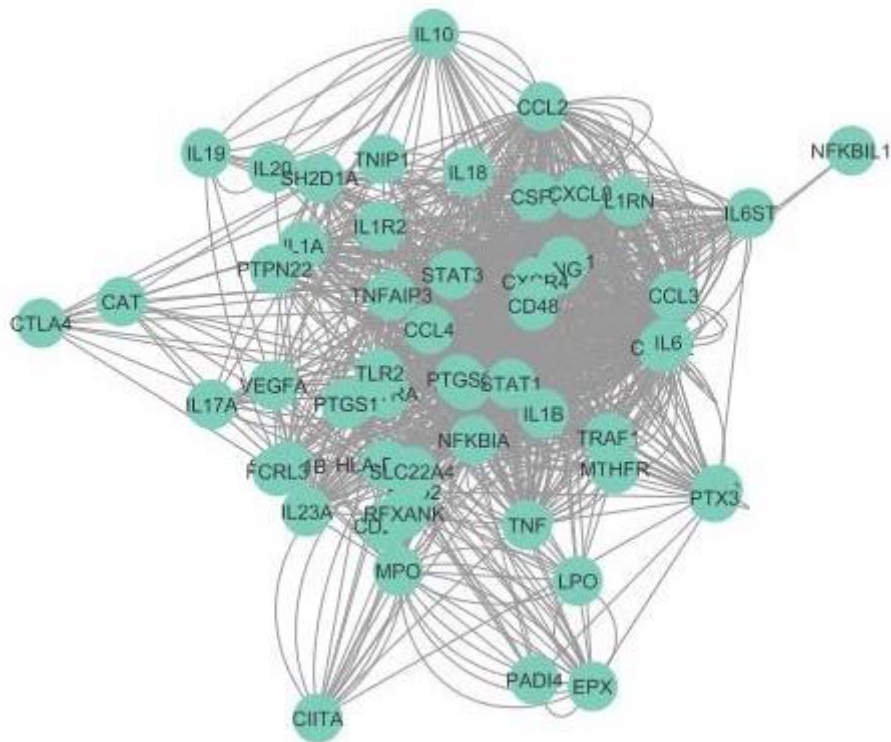


Figura 14. Red biológica generada por GeneMANIA.

Al introducir los datos a Cytoscape, éstos pueden ser analizados de una forma más cómoda. Se organizan los nodos de modo que los que tengan más interacciones se sitúen en el centro de la red (Figura 15).

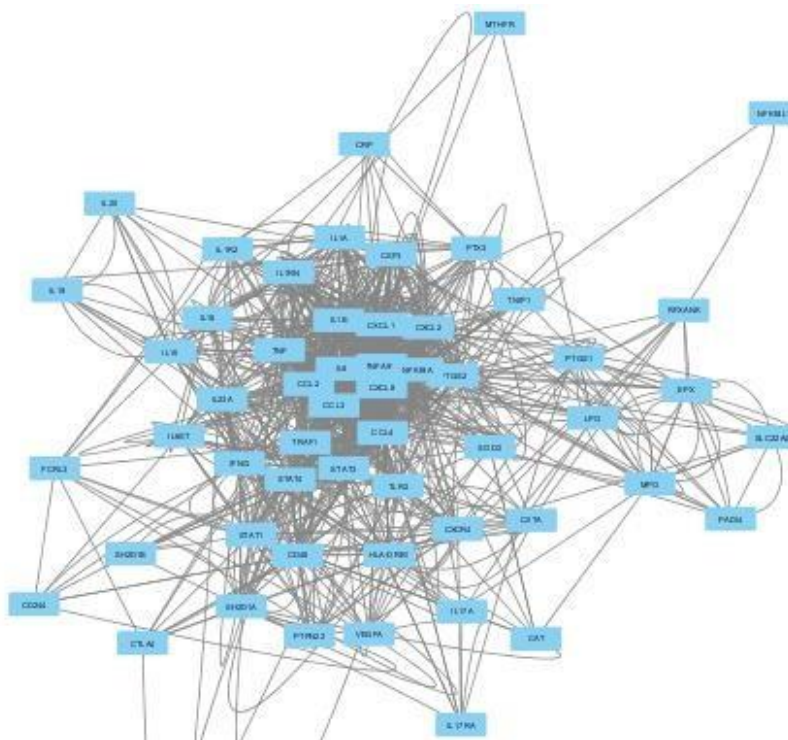


Figura 15. Visualización de la red en Cytoscape.

### 3.4 Selección de nodos centrales

Se han seleccionado los 5 nodos con mayor número de interacciones (Figura 16) aplicando un filtro a la red.

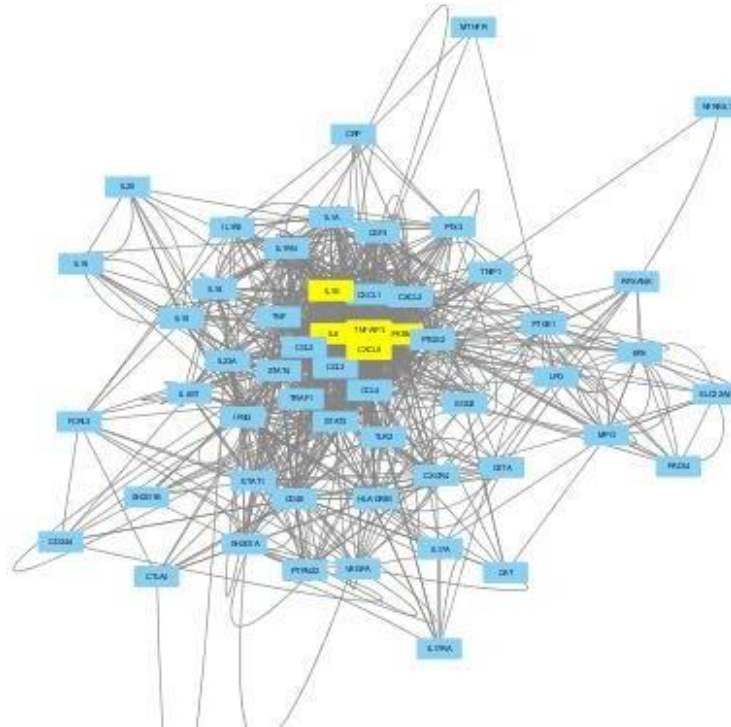


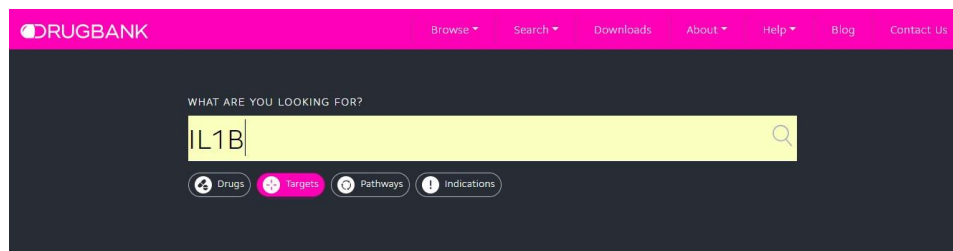
Figura 16. Selección de nodos centrales en Cytoscape.

La selección de genes resultante es:

**IL1B, IL6, TNFAIP3, CXCL8 y NFKBIA.**

### 3.5 Validación de los genes

Para verificar que los genes están descritos como dianas terapéuticas, se realiza una búsqueda en la base de datos de DrugBank (Figura 15) del gen como "Target".



The DrugBank database is a unique bioinformatics and cheminformatics resource that combines detailed drug data with comprehensive drug target information.

The latest release of DrugBank (version 5.1.0, released 2018-04-02) contains 11,141 drug entries including 2,555 approved small molecule drugs, 965 approved biotech (protein/peptide) drugs, 121 nutraceuticals

Figura 17. Página de inicio de DrugBank.

Tras seleccionar la entrada del gen, se abre una ficha de éste. Los puntos más importantes para este estudio son, por un lado, los medicamentos asociados (entrando en la ficha de cada uno se puede consultar si este se está indicando para la AR) y, por otro lado, la secuencia del gen o de la proteína para el posterior análisis de viabilidad de tratamiento.

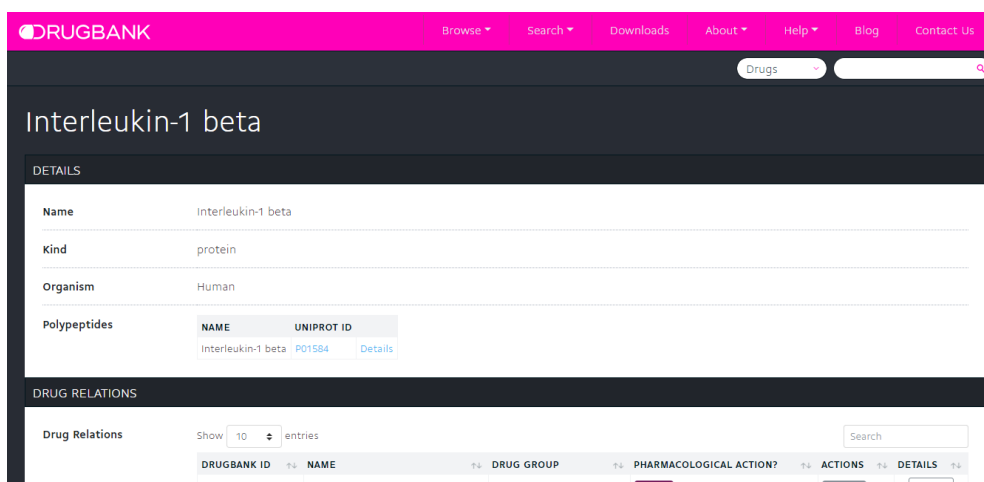


Figura 18. Ejemplo de ficha de un gen en DrugBank.

Se verifica la relación de estos genes con la AR, así como la existencia de fármacos asociados.

Tabla 8. Relación entre los genes y la AR según DrugBank

Gen	Nombre completo Gen	Diana	AR	Medicamentos	Medicamentos AR
<i>IL1B</i>	Interleukin-1 beta	SI	SI	14	3
<i>IL6</i>	Interleukin-6	SI	SI	13	4
<b><i>TNFAIP3</i></b>	<b>TNF alpha induced protein 3</b>	<b>NO</b>	<b>NO</b>	<b>0</b>	<b>0</b>
<b><i>CXCL8</i></b>	<b>Interleukin-8</b>	<b>SI</b>	<b>NO</b>	<b>4</b>	<b>0</b>
<i>NFKBIA</i>	NF-kappa-B inhibitor Alpha	SI	SI	3	1

Los genes ***TNFAIP3*** y ***CXCL8*** no están descritos como diana terapéutica de la AR, de hecho, el *TNFAIP3* no aparece descrito como diana terapéutica en DrugBank. En la tabla 9 se visualizan todos los medicamentos relacionados con el *CXCL8* y que se proponen como prueba para dicha enfermedad.

Tabla 9. Medicamentos relacionados con el gen *CXCL8*.

NAME	DRUG GROUP	PHARMACOLOGICAL ACTION?	ACTIONS	INDICATIONS
<b><i>Rivanicline</i></b>	investigational	Unknown	antagonist	Colitis ulcerativa
<b><i>MDX-018</i></b>	investigational	Unknown		Tratamiento de enfermedades autoinmunes y desórdenes inflamatorios

<b>ABT-510</b>	investigational	Unknown		Tratamiento de linfoma, melanoma y tumores sólidos
<b>Tapinarof</b>	investigational	Unknown		Tratamiento para psoriasis y desórdenes psoriáticos

La lista completa de los medicamentos, así como sus indicaciones y dianas terapéuticas, se encuentra en el [Anexo I](#). Algunos de ellos tratan incluso dos de las dianas seleccionadas y varias dianas más halladas para la artritis reumatoide y por sus características podría ser interesante realizar pruebas con pacientes de la AR.

A continuación, se describen algunos medicamentos que, dadas sus características, podrían ser de interés en el tratamiento de la AR son:

- Binimetinib se está investigando para el tratamiento de cáncer de pulmón, pero podría ser interesante para la AR debido a que su mecanismo es inhibir la encima MEK que regula la biosíntesis de las citoquinas inflamatorias. Además, inhibe tres dianas que están definidas como dianas de la AR, para algunas de ellas se ha hallado información sobre su efecto en pacientes con artritis reumatoide durante la búsqueda bibliográfica.
- Andrographolide es un producto botánico que se extrae de una hierba autóctona de China y sus propiedades también podrían ser interesantes para mejorar la calidad de vida de los afectados por la AR, debido a sus efectos antiinflamatorios.

El gen TFNAIP3, que no aparece en la base de datos de DrugBank, y por lo tanto no puede ser validado por esta metodología.

### 3.6 Test de viabilidad de las dianas terapéuticas

Tras realizar la búsqueda de cada gen y de todos los fármacos o ligandos de cada uno, se obtiene un gen que no aparece descrito como diana terapéutica. Para continuar, se ha obtenido la secuencia FASTA del gen a través de Uniprot (<http://www.uniprot.org/>):

```
>sp|P21580|TNAP3_HUMAN Tumor necrosis factor alpha-induced protein 3
OS=Homo sapiens OX=9606 GN=TNFAIP3 PE=1 SV=1
MAEQVLPQALYLSNMRKAVKIRERTPEDIFKPTNGIIHHFKTMHRYTLEMFRTCQFCPQF
REIIHKALIDRNIQATLESQKLNWCREVRKLVALKTNKDGNCIMHATSQYMWGVQDSDL
VLRKALFSTLKETDTRNFKFRWQLESLSKSEFVETGLCYDTRNWNDEWDNLIKMASTDP
MARSGLQYNSLEEIHIFVLCNILLRPIIVISDKMLRSLESNSFAPLKVGGIYLPWHWA
QECYRPIVVLGYDSHHFVPLVTLKDSGPEIRAVPLVNRDRGRFEDLKVHFLTDPENEMKE
KLLKEYLMVIEIPVQGDHGTTHLINAAKLDEANLPKEINLVDDYFELVQHEYKKWQENS
EQGRREGHAQNPMEPSVPLSLMDVKCETPNCPPFMSVNTQPLCHECSERRQKNQNKLPK
LNSKPGPEGLPGMALGASRGEAYEPLAWNPEESTGGPHSAPPTAPSPFLFSETTAMKCRS
PGCFPTLNQHNCFERCHNARQLHASHAPDHTRHLDPGKCQACLQDVTRTFNGICSTCF
KRTTAEASSSLSTSLPPSCHQRSKSDPSRLVRSPPHRSCHRAGNDAPAGCLSQAARTPGD
RTGTSKCRKAGCVYFGTPEKNGFCTLCFIEYRENKHFAAASGKVSPTASRFQNTIPCLGR
ECGLTGSTMFEGYCQKCFIEAQNQRFFHEAKRTEEQLRSSQRRDVPRTTQSTSRPKCARAS
CKNILACRSEELCMECQHPNQRMGPGAHRGEPAPEDPPKQRCRAPACDHFNGAKCNGYCN
ECFQFKQMYG
```



Esta secuencia se introduce en la herramienta DrugEBIility de la “European Bioinformatics Institute” (EBI), cuya página principal se observa en la figura 19.

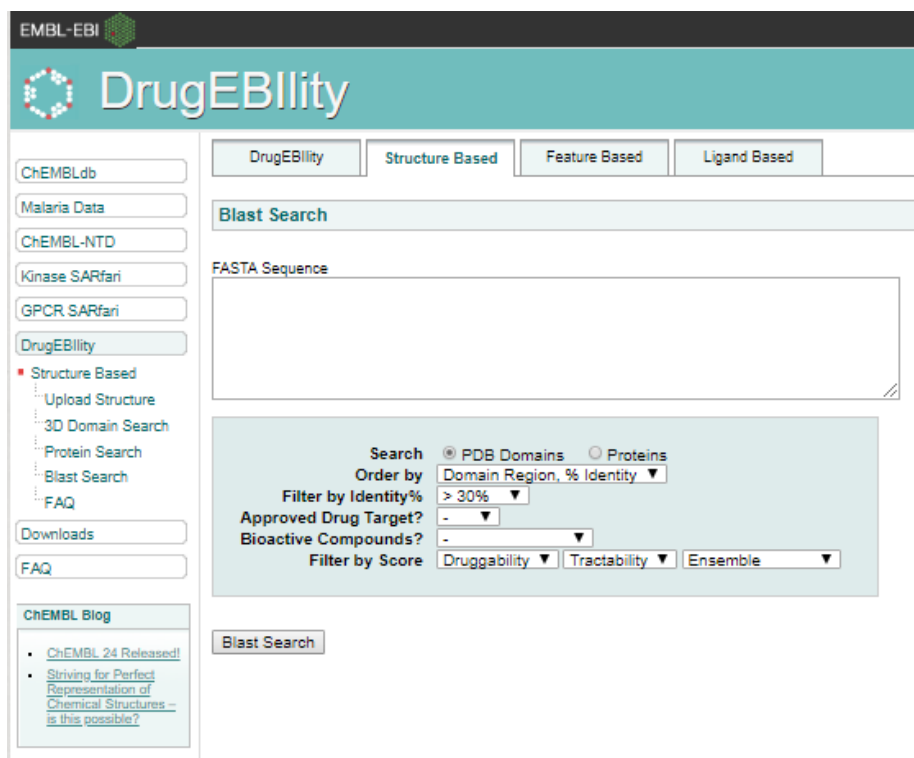


Figura 19. Pantalla de inicio de la herramienta DrugEBIility.

Se ha introducido la secuencia dejando los demás parámetros por defecto. El resultado se muestra en la Figura 20.

Blast cutoff has an E value of 0.0001. Green domains are druggable or tractable, red domains are neither druggable nor tractable.

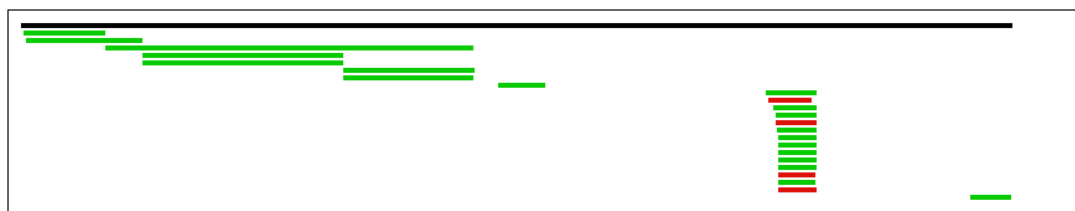


Figura 20. Resultado de los dominios relacionados con el gen TNFAIP3.

Se obtiene un total de 23 dominios, de los cuales 19 son tratables o “druggables” y solamente dos de ellos son ambos. La tabla completa se encuentra en el Anexo II.

Para continuar con el estudio se ha seleccionado de entre los dos dominios mencionados el que mejor valor de “Ensemble” obtiene. En este caso es el resultado 19, con un ensemble “tractable” + “druggable” de 0.70. Para acceder a la ficha de éste (Figura 21), es preciso presionar sobre el ID del dominio.

**Domain Family Details:**

Domain ID	4161239 [ PDBe PISA Biological Unit <sup>2</sup> ]
PDB	3QJ4
Fold	UNMATCHED
Superfamily	UNMATCHED
Family	UNMATCHED 3QJ4 HETERO
Gene	
Other PDB(s)	

**Site Druggability Details:**

Reset	Site 1	Site 2	Site 3	Site 4
Druggable <sup>2</sup>	1.00	0.00	0.00	0.00
Confidence	0.67	0.96	0.96	0.96
Tractable <sup>2</sup>	1.00	0.00	0.00	0.00
Confidence	0.92	0.86	0.86	0.86
Ensemble <sup>2</sup>	0.70	-0.99	-0.99	-0.99
Volume [Å <sup>3</sup> ]	1779.05	1494.7	1117.55	932.77
Buried Surface [%]	77.6	58.54	67.06	65.96
Show Site	●	○	○	○
Show Residues	●	○	○	○
Ligand	○			

Legend: Green = Druggable, Yellow = Tractable, Pink = Undruggable

**Family Druggability Average:**

Name	UNMATCHED 3QJ4 HETERO
Avg Druggability	Tractable: , Druggable: , Ensemble: <a href="#">Details</a>

Figura 21. Ficha del dominio seleccionado.

Por otro lado, clicando sobre Align en la tabla del Anexo II se obtiene la secuencia de la proteína.

>3650499|3QJ4

Length = 31

Score = 69.3 bits (168), Expect = 3e-11, Method: Composition-based stats.

Identities = 31/31 (100%), Positives = 31/31 (100%)

Query: 605 SKCRKAGCVYFGTPENKGFCTLCFIEYRENK 635

SKCRKAGCVYFGTPENKGFCTLCFIEYRENK

Sbjct: 1 SKCRKAGCVYFGTPENKGFCTLCFIEYRENK 31

Con la secuencia del dominio procedemos a la propuesta de un modelo 3D para poder realizar un screening virtual de posibles moléculas que pudieran unirse al mismo. El modelo se creará mediante Swiss Model (Figura 22).

**Start a New Modelling Project**

Target Sequence:  (Format must be FASTA. Guide: plain string, or a valid UniProt(B AC))

Project Title:

Email:

Supported Inputs:

- Sequence(s)
- Target-Template Alignment
- User Template
- DeepView Project

By using the SWISS-MODEL server, you agree to comply with the following terms of use and to cite the corresponding articles.

You are currently not logged in - to take advantage of the workspace, please log in or create an account. (There is no requirement to create an account to use any part of SWISS-MODEL, however you will gain the benefit of seeing a list of your previous modelling projects here.)

Figura 22. Página de inicio de Swiss-model.

El resultado de la búsqueda de la secuencia puede observarse en la Figura 23. Se obtienen “templates” basados en la estructura 3D de la propia proteína, código de acceso de la Protein Data Bank (PDB) 3oj3.

Template	Sequence Identity	Blounit Oligo State	Description
3oj3.1.B	100.00	hetero-dimer	Tumor necrosis factor alpha-induced protein 3
3oj3.3.B	100.00	hetero-dimer	Tumor necrosis factor alpha-induced protein 3
3oj4.1.C	100.00	hetero-trimer	Tumor necrosis factor alpha-induced protein 3
2eqe.1.A	100.00	monomer	Tumor necrosis factor, alpha-induced protein 3
2c7n.1.A	44.00	hetero-dimer	RAB GUANINE NUCLEOTIDE EXCHANGE FACTOR 1

Figura 23. Resultados de Swiss-model

Los resultados de la búsqueda de “templates” se aprecian en la Figura 24.

Sort	Name	Title	Coverage	GMQE	QSQE	Identity	Method	Oligo State	Ligands
<input checked="" type="checkbox"/>	3oj3.1.B	Tumor necrosis factor alpha-induced protein 3	<div style="width: 100%;"></div>	0.99	-	100.00	X-ray, 2.5Å	hetero-dimer Δ	1 x ZN <sup>ES</sup> ▼
<input type="checkbox"/>	3oj3.3.B	Tumor necrosis factor alpha-induced protein 3	<div style="width: 100%;"></div>	0.99	-	100.00	X-ray, 2.5Å	hetero-dimer Δ	1 x ZN <sup>ES</sup> ▼
<input type="checkbox"/>	3oj4.1.C	Tumor necrosis factor alpha-induced protein 3	<div style="width: 100%;"></div>	0.99	-	100.00	X-ray, 3.4Å	hetero-trimer Δ	1 x ZN <sup>ES</sup> ▼
<input type="checkbox"/>	2eqe.1.A	Tumor necrosis factor, alpha-induced protein 3	<div style="width: 100%;"></div>	0.96	-	100.00	NMR	monomer ✓	1 x ZN <sup>ES</sup> ▼
<input type="checkbox"/>	2c7n.1.A	RAB GUANINE NUCLEOTIDE EXCHANGE FACTOR 1	<div style="width: 71%;"></div>	0.71	-	44.00	X-ray, 2.1Å	hetero-dimer Δ	1 x ZN <sup>ES</sup> ▼
<input type="checkbox"/>	2c7m.1.A	RAB GUANINE NUCLEOTIDE EXCHANGE FACTOR 1	<div style="width: 71%;"></div>	0.71	-	44.00	X-ray, 2.4Å	hetero-dimer Δ	1 x ZN <sup>ES</sup> ▼
<input type="checkbox"/>	2c7n.3.A	RAB GUANINE NUCLEOTIDE EXCHANGE FACTOR 1	<div style="width: 71%;"></div>	0.71	-	44.00	X-ray, 2.1Å	hetero-dimer Δ	1 x ZN <sup>ES</sup> ▼
<input type="checkbox"/>	2c7n.5.A	RAB GUANINE NUCLEOTIDE EXCHANGE FACTOR 1	<div style="width: 71%;"></div>	0.71	-	44.00	X-ray, 2.1Å	hetero-dimer Δ	1 x ZN <sup>ES</sup> ▼
<input type="checkbox"/>	2fif.1.B	Rab5 GDP/GTP exchange factor	<div style="width: 71%;"></div>	0.71	-	44.00	X-ray, 2.5Å	hetero-dimer Δ	1 x ZN <sup>ES</sup> ▼
<input type="checkbox"/>	2fif.2.B	Rab5 GDP/GTP exchange factor	<div style="width: 71%;"></div>	0.71	-	44.00	X-ray, 2.5Å	hetero-dimer Δ	1 x ZN <sup>ES</sup> ▼
<input type="checkbox"/>	2fif.3.B	Rab5 GDP/GTP exchange factor	<div style="width: 71%;"></div>	0.71	-	44.00	X-ray, 2.5Å	hetero-dimer Δ	1 x ZN <sup>ES</sup> ▼
<input type="checkbox"/>	2fid.1.C	Rab5 GDP/GTP exchange factor	<div style="width: 71%;"></div>	0.71	-	44.00	X-ray, 2.8Å	hetero-trimer Δ	1 x ZN <sup>ES</sup> ▼

Figura 24. Resultados de los templates de Swiss-model

El dominio “druggable” corresponde al dominio tipo dedo de Zinc de Tumor necrosis factor alpha-induce protein 3, con una identidad de 100. Su ficha y estructura se observan en la Figura 25.

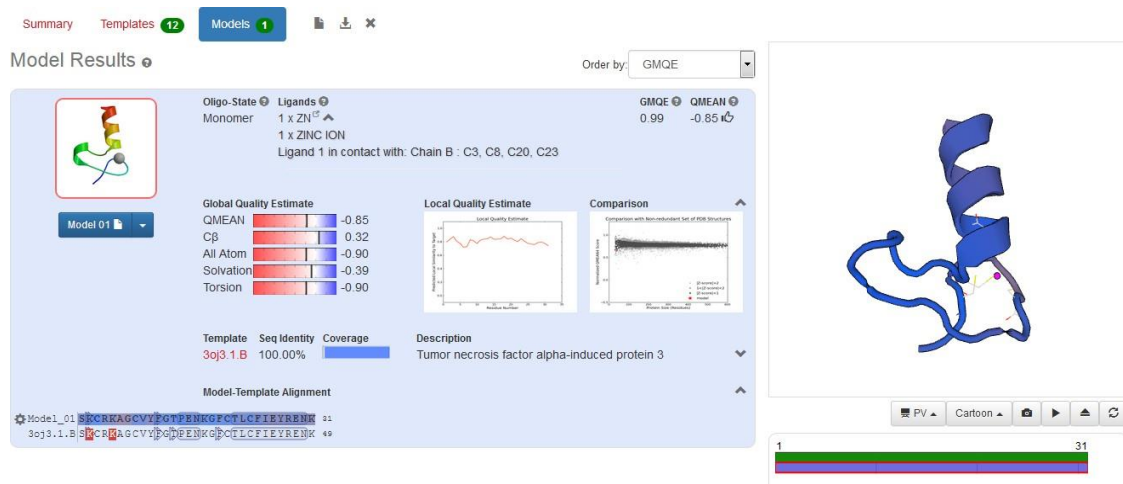


Figura 25. Estructura del dominio seleccionado

Desde la web se puede descargar el archivo de la estructura 3D del modelo en formato pdb.

Para usar este archivo como entrada para realizar virtual screening se usará el software local Autodocktools para definir las características del cubo de búsqueda (Figura 26), donde se definirán las coordenadas del centro del cubo y tamaño de lados para recoger todo el dominio sobre el que realizar el virtual screening de los ligandos.

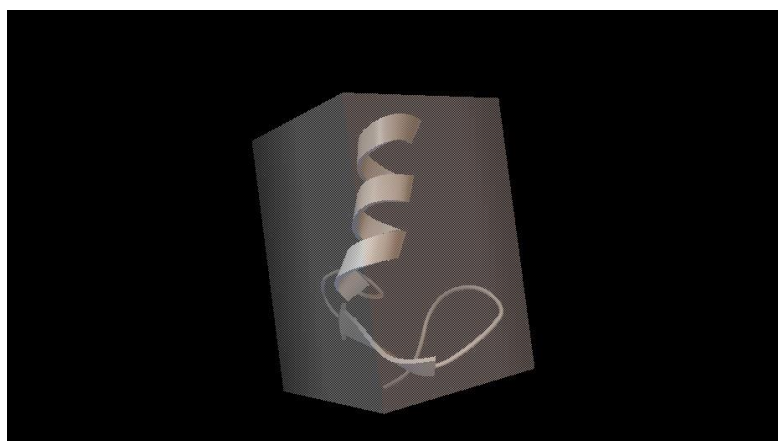


Figura 26. Caja que engloba el ligando seleccionado para el virtual screening

Las coordenadas de búsqueda vendrán definidas por:

Centro de la caja de búsqueda (grid box):

X= -8.535

Y= -6.254

Z= 17.939

Tamaño de la caja de búsqueda (puntos):

X= 44 puntos

Y= 62 puntos

Z=42 puntos

A continuación, se realiza el Virtual screening con MTIOpenScreen, accediendo al portal a través del siguiente enlace: <http://bioserv.rpbs.univ-paris-diderot.fr/services/MTiOpenScreen/>, como protein receptor se introduce la estructura del ligando en formato PDB tal como se observa en la figura 27.

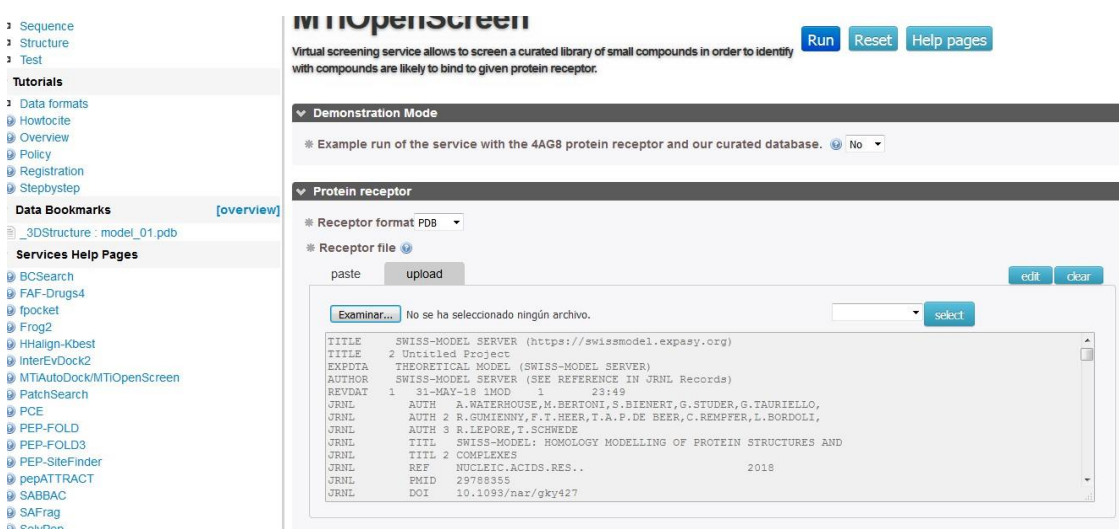


Figura 27. Introducción de los datos e MTIOpenScreen para el Virtual Screening.

Los parámetros a seleccionar son la quimioteca de compuestos que usará la herramienta en el Virtual Screening (se elige diverse-lib debido a su alta variabilidad en ligandos provenientes de distintas bases de datos [14]) y por último el filtro "Is lead-like?", que omite los demás parámetros a introducir.

Finalmente, se introducen los datos de la caja de búsqueda calculados en el punto anterior tal como se indica en la siguiente figura:

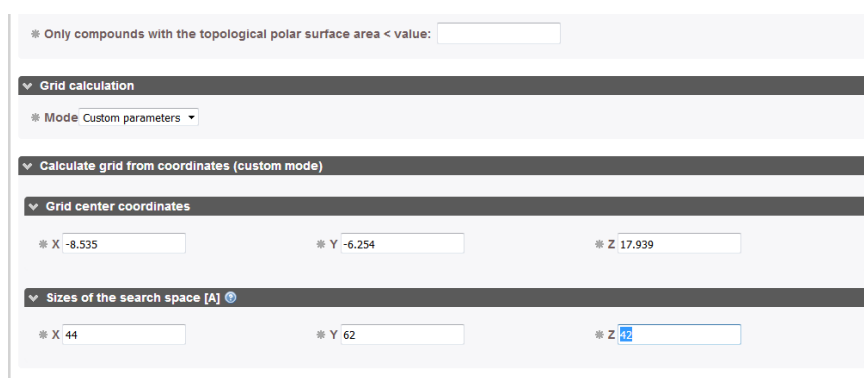


Figura 28. Introducción de los datos de la caja de Autodock en MTIOpenScreen.

Una vez introducidos estos datos, se inicia el proceso, que resulta en una búsqueda de los 100 mejores ligandos.

En la pestaña “Job” se puede visualizar el informe del trabajo realizado y los resultados del virtual screening, tabla con los distintos ligandos (Figura 29) ordenados de menor energía de interacción a mayor. Los softwares de docking y virtual screening tienen una función de cálculo de la energía de interacción entre receptor y ligando buscando las conformaciones más accesibles energéticamente, es decir las de menor energía son las más estables.

Compound	Model ID	Energy	nRot	Library	isLeadLike	HBA	HBD	LogP	MW	TP
860538_Intermediate	1	-7.9	4	diverse-lib	Y	7	1	3.99	381.3868	91.6
17443648_Accepted	1	-7.7	1	diverse-lib	Y	3	1	3.59	273.28544	46.1
99358243_Intermediate	1	-7.5	1	diverse-lib	Y	5	0	2.1	351.39574	65.3
49644536_Intermediate	1	-7.4	1	diverse-lib	Y	3	0	3.97	307.7305	47.0
24296088_Intermediate	1	-7.3	2	diverse-lib	Y	5	2	3.77	356.37404	85.6
24270812_Intermediate	1	-7.3	2	diverse-lib	Y	5	2	2.42	299.2533032	82.3
24837121_Accepted	1	-7.3	4	diverse-lib	Y	5	0	3.68	406.49738	66.0
3716217_Intermediate	1	-7.3	3	diverse-lib	Y	6	0	3.18	372.4198	62.4

Figura 29. Captura de la tabla con los mejores 1000 ligandos.

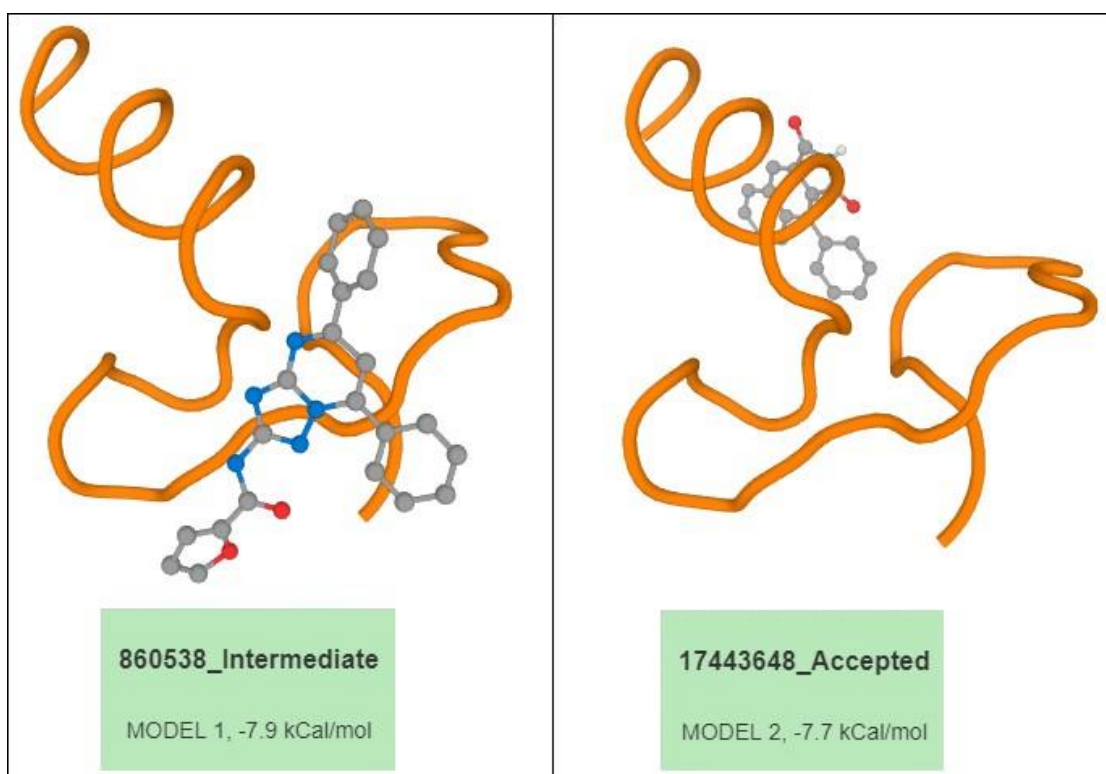
Además del dato de energía de interacción se ofrece información sobre si son moléculas “cabeza de serie” (leadlike), por tanto, sólo se obtendrán moléculas que tienen potencialidad para funcionar como cabezas de serie de una nueva serie sintética.

Además, se da el número de enlaces rotables, átomos aceptores de enlaces de hidrógeno, átomos dadores de enlace de hidrógeno, logP (variable que mide la hidrofobicidad), peso molecular. Todas ellas variables importantes para calcular la regla del 5 de Lipinski<sup>1</sup>.

<sup>1</sup> La regla de Lipinski da información sobre las características como posible fármaco de una molécula. Esta regla se usa para determinar si un compuesto químico tiene las características que lo hagan adecuado para convertirse en un medicamento de vía oral para humanos. Éstas se basan en la observación de que muchos fármacos administrados de forma oral son relativamente pequeñas y lipófilas. La regla de Lipinski establece que un fármaco oral activo no incumple más de uno de los siguientes criterios [19]:

- No más de 5 enlaces dadores de hidrógeno.
- No más de 10 enlaces aceptores de hidrógeno.
- Masa molecular inferior a 500 daltons.
- La variable logP no sea superior a 5

Las mejores interacciones se observan en la figura 30:



*Figura 30. Interacciones de los ligandos con energías de interacción más favorable.*

Según se aprecia en las dos figuras anteriores se observan 2 modos de unión, el del ligando 860538\_Intermediate que se une en la zona de loop y la lámina Beta y la del ligando 17443648\_Accepted en la hélice alfa (normalmente la de reconocimiento para unión a ADN), lo que podría suponer una estrategia para el diseño de nuevos ligandos. Para facilitar la visualización de estos ligandos y de los 10 primeros, véanse la Figura 31.

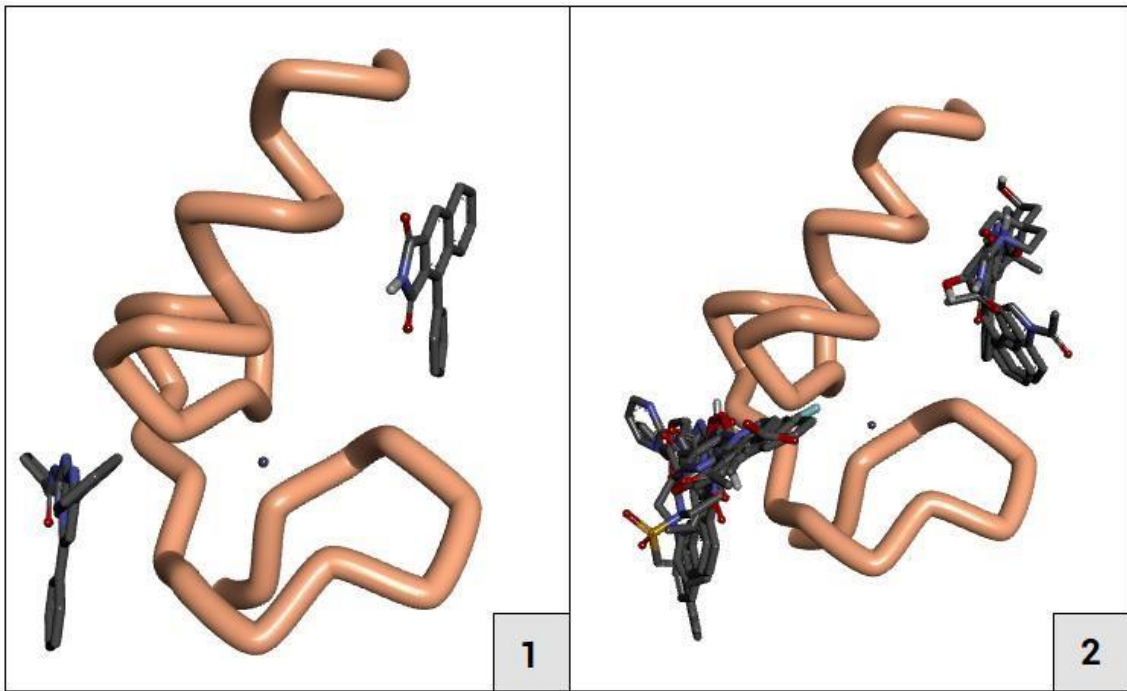


Figura 31. Superposición de los primeros 2 ligandos [1] y de los 10 primeros [2]

La figura anterior muestra los puntos de unión al dominio preferidos por los ligandos, en este caso, son dos puntos de unión claramente marcados.



### 3. Conclusiones

Gracias a este proyecto he sido consciente de la importancia de las herramientas digitales y de la bioinformática para incrementar la eficiencia en procesos que sin estas posibilidades serían muy lentos.

En cuanto a los objetivos inicialmente planteados, se han podido cumplir todos los que se habían definido, de modo que:

1. Con la revisión bibliográfica he podido comprender la complejidad de la artritis reumatoide (AR) tanto por el desconocimiento de las causas por las que se manifiesta, como por las dificultades a la hora de diagnosticarla y la baja eficiencia de los fármacos usados hasta el momento.
2. Se han encontrado más de 20.000 genes relacionados con la AR mediante distintas bases de datos. De éstos, se ha hecho una selección de 45, siendo el criterio seguido, que se trataran de los 15 genes más relevantes según cada base de datos.
3. Gracias a la integración de estos datos y el uso de herramientas que en este caso han sido GeneMania y Cytoscape, se ha hecho una selección de los 5 genes con mayor número de interacciones.
4. Se han buscado las dianas candidatas y se han clasificado según si están descritas como diana terapéutica y más concretamente, si lo están para la AR. De los 5 genes seleccionados, uno de ellos ha resultado no estar descrito como diana de la AR y otro no aparece descrito como diana terapéutica. Se ha priorizado el estudio en hallar los medicamentos del gen no descrito como diana de la artritis reumatoide con el objetivo de proponerlos como posible tratamiento y en estudiar la viabilidad o druggability de la molécula no descrita como diana.
5. Se ha testeado la viabilidad del gen TNFAIP3, encontrando un dominio del tipo dedo de zinc y una lista de posibles ligandos que se unen a éste por dos puntos de unión diferentes.

Con respecto a la planificación del trabajo, para cumplir con los plazos, se decidió reducir la selección de genes a estudiar de los casi 30.000 a 45, seleccionando los 15 más relevantes de cada base de datos. La principal razón de esta decisión fue el tiempo de procesado por parte de GeneMania con todos los genes. Inicialmente se intentó realizar este proceso, pero tras muchas horas sin resultados y con una apariencia de que no se estaba realizando ningún cálculo, se optó por reducir dicha selección.

## 4. Glosario

**CADD:** Conjunto de técnicas computacionales que permiten, mediante distintas herramientas, la obtención de distintos fármacos con el objetivo de distintas enfermedades.

**Docking:** Método que predice la conformación de una molécula, al estar unida a otra, para formar un complejo estable.

**Dominio:** Partes de una proteína que cumplen una función determinada en la misma.

**Druggable:** Dicho de la habilidad de un gen a ser considerado como diana a ser tratada por un fármaco.

**Druggability:** Capacidad que tienen determinadas moléculas de ser susceptibles de unirse a un fármaco (o futuro fármaco) de forma resistente. Se evalúa mediante un estudio de si es necesaria su secuencia proteica, y mediante el SCORE, se puede saber cuáles de sus dominios son más “druggables”.

**Drug-like:** Se dice de un compuesto que presenta propiedades similares a un fármaco.

**Virtual screening:** Cribado virtual. Proceso para buscar compuestos que se puedan unir a las proteínas diana propuestas, presentando una energía libre de enlace favorable.

## 5. Bibliografía

- [1] J. Richard D. Brasington, "Clinical features of rheumatoid arthritis," *Rheumatology*, Elsevier Ltd, pp. 704–711, 2015.
- [2] L. Carmona, "Epidemiología de la artritis reumatoide," vol. 29, no. 3, pp. 86–89, 2013.
- [3] M. L. González, J. Rueda, H. González, E. Cantor, and A. Martínez, "Artritis reumatoide temprana: Resultados clínicos y funcionales de una cohorte en un centro de alta complejidad, Cali-Colombia," *Rev. Colomb. Reumatol.*, vol. 23, no. 3, pp. 148–154, 2016.
- [4] J. S. Smolen *et al.*, "EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs : 2013 update," pp. 1–18, 2013.
- [5] I. M. Kapetanovic, "Computer-aided drug discovery and development (CADD): In silico-chemico-biological approach," *Chem. Biol. Interact.*, vol. 171, no. 2, pp. 165–176, 2008.
- [6] G. Sliwoski, S. Kothiwale, J. Meiler, and E. W. Lowe, "Computational methods in drug discovery.," *Pharmacol. Rev.*, vol. 66, no. 1, pp. 334–95, 2014.
- [7] J. Pi, J. Deu-pons, E. Centeno, J. Garc, F. Sanz, and L. I. Furlong, "DisGeNET : a comprehensive platform integrating information on human disease-associated genes and," vol. 45, no. October 2016, pp. 833–839, 2017.
- [8] N. Rappaport *et al.*, "Original article MalaCards : an integrated compendium for diseases and their annotation," vol. 2013, pp. 1–14, 2013.
- [9] M. Kanehisa, S. Goto, Y. Sato, M. Furumichi, and M. Tanabe, "KEGG for integration and interpretation of large-scale molecular data sets," vol. 40, no. November 2011, pp. 109–114, 2012.
- [10] A. P. Davis *et al.*, "The Comparative Toxicogenomics Database : update 2017," vol. 45, no. September 2016, pp. 972–978, 2017.
- [11] S. Lotia, J. Montojo, Y. Dong, G. D. Bader, and A. R. Pico, "BIOINFORMATICS APPLICATIONS NOTE Systems biology Cytoscape App Store," vol. 29, no. 10, pp. 1350–1351, 2013.
- [12] K. Zuberi *et al.*, "GeneMANIA Prediction Server 2013 Update," vol. 41, pp. 115–122, 2013.
- [13] D. S. Wishart *et al.*, "DrugBank 5 . 0 : a major update to the DrugBank database for 2018," vol. 46, no. November 2017, pp. 1074–1082, 2018.
- [14] J. Rey *et al.*, "MTiOpenScreen : a web server for structure-based," vol. 43, no. April, pp. 448–454, 2015.
- [15] D. Aletaha *et al.*, "2010 Rheumatoid arthritis classification criteria: An American College of Rheumatology/European League Against Rheumatism collaborative initiative," *Arthritis Rheum.*, vol. 62, no. 9, pp. 2569–2581, 2010.

- [16] “Artritis reumatoide: Síntomas, diagnóstico y tratamientos,” *Arthritis Foundation National Office*. [Online]. Available: <http://espanol.arthritis.org/espanol/disease-center/imprimia-un-folleto/dc-ar/>.
- [17] O. Foot and A. Malaviya, “Treatment of rheumatoid arthritis: current and future,” *Futur. Prescr.*, vol. 14, no. 2, pp. 8–12, 2013.
- [18] B. L. King, A. P. Davis, M. C. Rosenstein, T. C. Wieggers, and C. J. Mattingly, “Ranking Transitive Chemical-Disease Inferences Using Local Network Topology in the Comparative Toxicogenomics Database,” vol. 7, no. 11, 2012.
- [19] C. A. Lipinski, “Rule of five in 2015 and beyond: Target and ligand structural limitations, ligand chemistry structure and drug discovery project decisions,” *Adv. Drug Deliv. Rev.*, 2016.

## 6. Anexos

### Anexo I

Tabla con todos los resultados de la base de datos DrugBank resumidos. Contiene información sobre cada fármaco o ligando, sus indicaciones y dianas terapéuticas.

<b>NAME</b>	<b>DRUG GROUP</b>	<b>PHARMACOLOGICAL ACTION?</b>	<b>ACTIONS</b>	<b>INDICATIONS</b>	<b>TARGETS</b>
<b><i>Minocycline</i></b>	approved, investigational	unknown	modulator	Tratamiento de infecciones	rpsI, rpsD, <b>IL1B</b> , ALOX5, MMP9, VEGFA, CASP1, CASP3, CYCS
<b><i>Talmapimod</i></b>	investigational	unknown		Tratamiento del dolor agudo o crónico y <b>artritis reumatoide</b>	MAPK14, TNF, <b>IL1B</b> , MT-CO2
<b><i>Etiprednol dicloacetate</i></b>	investigational	unknown		Tratamiento del asma y enfermedad de Crohn	<b>IL1B</b>
<b><i>VP025</i></b>	investigational	unknown		Tratamiento del alzheimer y desórdenes inflamatorios	<b>IL1B</b>
<b><i>VX-702</i></b>	investigational	unknown		Tratamiento de enfermedad de arteria coronaria, desórdenes inflamatorios y <b>artritis reumatoide</b>	MAPK14, TNF, <b>IL1B</b> , <b>IL6</b>

<b>Andrographolide</b>	investigational	unknown		Tratamiento de la colitis ulcerativa e infecciones de las vías respiratorias altas	TNF, <b>IL1B</b> , <b>IL6</b> , NFKB2, NFKB1
<b>VX-765</b>	investigational	unknown		Tratamiento de desórdenes inflamatorios, psoriasis y desórdenes psoriático	CASP1, <b>IL1B</b>
<b>Gevokizumab</b>	investigational	unknown		Tratamiento de acné, osteoartritis y uveítis de Behcet	<b>IL1B</b>
<b>Binimetinib</b>	investigational	unknown		Cáncer y carcinoma de pulmón	TNF, <b>IL1B</b> , <b>IL6</b>
<b>Canakinumab</b>	approved, investigational	yes	binder	Tratamiento de síndromes periódicos asociados a la criopirina (CAPS) y la artritis sistémica idiopática juvenil	<b>IL1B</b>
<b>Gallium nitrate</b>	approved, investigational	yes	antagonist	Tratamiento de la hipercalcemia y linfoma non-hodgkin	RRM2, ATP6V1B2, BGLAP, RPI, <b>IL1B</b>
<b>Riloncept</b>	approved, investigational	unknown	binder	Tratamiento de síndromes periódicos asociados a la criopirina (CAPS)	IL1RN, <b>IL1B</b> , IL1A
<b>Dilmapimod</b>	investigational	unknown		Tratamiento de inflamación, dolor, <b>artritis reumatoide</b> y enfermedad de la arteria coronaria	TNF, <b>IL1B</b> , <b>IL6</b>

<b>Foreskin (neonatal)</b>	<b>4</b>	approved	yes	agonist	Tratamiento de úlceras en las piernas y pie diabético	FGF1, EGFR, CSF2RA, <b>IL1B</b> , <b>IL6</b> , IFNG, FGFR2, PDGFRA, TGFBR2, TGFB1, TNF, VEGFA
<b>Ginseng</b>		approved, investigational, nutraceutical	unknown	antagonist	No disponible	<b>IL6</b> , PTGS2, AHR
<b>YSIL6</b>		investigational	unknown		Tratamiento en la enfermedad de Crohn, psoriasis y desórdenes psoriáticos y <b>artritis reumatoide</b>	TNF, <b>IL6</b>
<b>CRx-139</b>		investigational	unknown		Tratamiento de la <b>artritis reumatoide</b>	TNF, <b>IL6</b> , CRP, IL10
<b>Tapinarof</b>		investigational	unknown		Tratamiento para psoriasis y desórdenes psoriáticos	IL12B, <b>IL6</b> , IL2, <b>CXCL8</b>
<b>Atiprimod</b>		investigational	unknown		Tratamiento para cáncer/tumores, mieloma múltiple y <b>artritis reumatoide</b>	TNF, <b>IL6</b>
<b>Siltuximab</b>		approved, investigational	yes	antagonist antibody	Tratamiento de pacientes con la enfermedad multicéntrica de Castleman	<b>IL6</b>
<b>Polaprezinc</b>		approved, investigational	unknown	inhibitor	Enfermedad de úlcera péptica y dispepsia	TNF, <b>IL6</b> , IL3, FLT1, NGF, PDGFRB, HSP90AA1,

					HSP90AB1
<b>Foreskin fibroblast (neonatal)</b>	approved	unknown	agonist	Tratamiento de úlceras de pies diabéticos	<b>IL6</b> , FLT1, IL1R1, TGFBR2, FGF1, IFNG, FGFR2, CSF2RA, PDGFRB, TGFB1, TNF
<b>ABT-510</b>	investigational	unknown		Tratamiento de linfoma, melanoma y tumores sólidos	FGF2, HGF, VEGFA, <b>CXCL8</b>
<b>MDX-018</b>	investigational	unknown		Tratamiento de enfermedades autoinmunes y desórdenes inflamatorios	<b>CXCL8</b>
<b>Rivanicline</b>	investigational	unknown	antagonist	Colitis ulcerativa	CHRN2, <b>CXCL8</b>
<b>Bardoxolone methyl</b>	investigational	unknown		Tratamiento de linfoma, melanoma y tumores sólidos	<b>NFKBIA</b>
<b>Astaxanthin</b>	investigational	unknown		Desórdenes e infecciones oculares, cancer/tumores y asma	<b>NFKBIA</b>
<b>Acetylsalicylic acid</b>	approved, vet_approved	unknown	inhibitor	Alivio temporal de dolor e inflamación incluyendo <b>artritis reumatoide</b>	PTGS1, PTGS2, AKR1C1, PRKAA1, EDNRA, IKBKB, TP53, HSPA5, RPS6KA3, <b>NFKBIA</b> , NFKB2



## Anexo II

Resultado de la búsqueda de la secuencia FASTA del gen TNFAIP3 en DrugEBility.

#	Domain	Tractable	Druggable	Ensemble	PDB	Family	Parent Protein	InChEMBL	Approved Drug?	Query Start	Query End	% Ident	E-Value	Align
1	<a href="#">3360507</a>	1	0	-0.82	3dkb:A	SirA-like	<a href="#">P21580</a>	No	No	3	68	100	1e-32	=
2	<a href="#">3220045</a>	1	0	-0.86	2vfj:A	UNMATCHED 2VFJ	<a href="#">P21580</a>	No	No	5	97	100	5e-49	≡
3	<a href="#">3360508</a>	1	0	-0.78	3dkb:A	UNMATCHED 3DKB	<a href="#">P21580</a>	No	No	69	362	97	1e-173	=
4	<a href="#">3360510</a>	1	0	-0.79	3dkb:B	PFAM OTU	<a href="#">P21580</a>	No	No	98	257	95	8e-89	≡
5	<a href="#">3220046</a>	1	0	-0.96	2vfj:A	PFAM OTU	<a href="#">P21580</a>	No	No	98	257	83	9e-74	=
6	<a href="#">3360511</a>	1	0	-0.83	3dkb:B	Methyltetrahydrofolate-utilizing methyltransferases	<a href="#">P21580</a>	No	No	258	362	100	1e-58	≡
7	<a href="#">3220047</a>	1	0	-0.86	2vfj:A	Methyltetrahydrofolate-utilizing methyltransferases	<a href="#">P21580</a>	No	No	258	361	100	4e-58	=
8	<a href="#">3080599</a>	1	0	-0.98	2eqg:A	PFAM zf-A20	<a href="#">P21580</a>	No	No	382	419	94	5e-13	≡
9	<a href="#">4078415</a>	1	1	0.69	3oj4:EDF	UNMATCHED 3OJ4 HETERO		No	No	595	635	80	7e-13	=
10	<a href="#">3080597</a>	0	0	-0.98	2eqe:A	PFAM zf-A20	<a href="#">P21580</a>	No	No	597	631	100	3e-14	≡
11	<a href="#">4207877</a>	1	1	-0.66	3oj3:DICA	UNMATCHED 3OJ3 HETERO		No	No	601	635	97	4e-15	=

12	<a href="#">4091763</a>	1	0	-0.97	3oj3:IA	UNMATCHED 3OJ3 HETERO		No	No	603	635	100	5e-14	≡
13	<a href="#">3650496</a>	0	0	-0.98	3oj3:l	PFAM zf-A20	<a href="#">P21580</a>	No	No	603	635	100	1e-12	≡
14	<a href="#">4174387</a>	1	0	-0.98	3oj3:GO	UNMATCHED 3OJ3 HETERO		No	No	604	635	100	3e-13	≡
15	<a href="#">4188075</a>	1	0	-0.97	3oj3:KC	UNMATCHED 3OJ3 HETERO		No	No	605	635	100	9e-13	≡
16	<a href="#">4144960</a>	1	0	-0.97	3oj3:DL	UNMATCHED 3OJ3 HETERO		No	No	605	635	100	9e-13	≡
17	<a href="#">4167652</a>	1	0	-0.98	3oj3:EM	UNMATCHED 3OJ3 HETERO		No	No	605	635	100	9e-13	≡
18	<a href="#">4193669</a>	1	0	-0.83	3oj3:HP	UNMATCHED 3OJ3 HETERO		No	No	605	635	100	9e-13	≡
19	<a href="#">4161239</a>	1	1	0.70	3oj4:BCA	UNMATCHED 3OJ4 HETERO		No	No	605	635	100	1e-12	≡
20	<a href="#">4024396</a>	0	0	-0.97	3oj3:Nf	UNMATCHED 3OJ3 HETERO		No	No	605	634	100	3e-12	≡
21	<a href="#">4048439</a>	1	0	-0.95	3oj3:JB	UNMATCHED 3OJ3 HETERO		No	No	605	634	100	3e-12	≡
22	<a href="#">3650499</a>	0	0	-0.98	3oj4:C	PFAM zf-A20	<a href="#">P21580</a>	No	No	605	635	100	3e-11	≡
23	<a href="#">3080598</a>	1	0	-0.97	2eqf:A	PFAM zf-A20	<a href="#">P21580</a>	No	No	758	790	100	1e-13	≡