TestBio: a tool for estimating the bioactivity of phytocomplexes tutorial

TestBio is a web-based tool designed to estimate the bioactivity of phytocomplexes, complex mixtures of compounds found in plants. It addresses the challenge of predicting the potential therapeutic effects of plant extracts by comparing their constituent molecules to a library of known bioactive compounds. TestBio is intended for researchers, scientists, and anyone interested in exploring the potential health benefits of plants.

In this tutorial we will show how to use testbio

- Identify the species whose bioactivity you want to predict, in this case rosemary (*Rosmarinus officinalis*)
- Look for a study in which the plant is analyzed to identify its composition, see reference 1 for rosemary,
- Identify the compounds and use the SMILES notation on the PubChem website (see reference 3) in the case of rosemary:

rosemary extract

Acido carnosico

CC(C)C1=C(C(=C2C(=C1)CC[C@@H]3[C@@]2(CCCC3(C)C)C(=O)O)O)O

Carnosol

CC(C)C1=C(C(=C2C(=C1)[C@@H]3C[C@@H]4[C@@]2(CCCC4(C)C)C(=O)O3)O)O

Ursolic acid

C[C@@H]1CC[C@@]2(CC[C@@]3(C(=CC[C@H]4[C@]3(CC[C@@H]5[C@@]4(CC[C@@H](C5(C)C)O)C)C)[C@@H]2[C@H]1C)C)C(=O)O

Rosmarinic acid

C1=CC(=C(C=C1C[C@H](C(=O)O)OC(=O)/C=C/C2=CC(=C(C=C2)O)O)O)O

Genkwanin

COC1=CC(=C2C(=C1)OC(=CC2=O)C3=CC=C(C=C3)O)O

Luteolin 7 glucoside

C1=CC(=C(C=C1C2=CC(=O)C3=C(C=C(C=C3O2)O[C@H]4[C@@H]([C@H]([C@@H]([C@H](O4)CO)O)O)O)O)O)O

Omoplantaginin

COC1=C(C=C2C(=C1O)C(=O)C=C(O2)C3=CC=C(C=C3)O)O[C@H]4[C@@H]([C@H]([C@@H]([C@H](O4)CO)O)O)O

Epigallocatechin

C1[C@H]([C@H](OC2=CC(=CC(=C21)O)O)C3=CC(=C(C(=C3)O)O)O)O

Neoesperidin

C[C@H]1[C@@H]([C@H]([C@H]([C@@H](O1)O[C@@H]2[C@H]([C@@H]([C@H](O[C@H]2OC3=CC(=C4C(=O)C[C@H](OC4=C3)C5=CC(=C(C=C5)OC)O)O)CO)O)O)O)O)O

Quinic acid

C1[C@H](C([C@@H](CC1(C(=O)O)O)O)O)O

Syringic acid

COC1=CC(=CC(=C1O)OC)C(=O)O

Ferulic acid

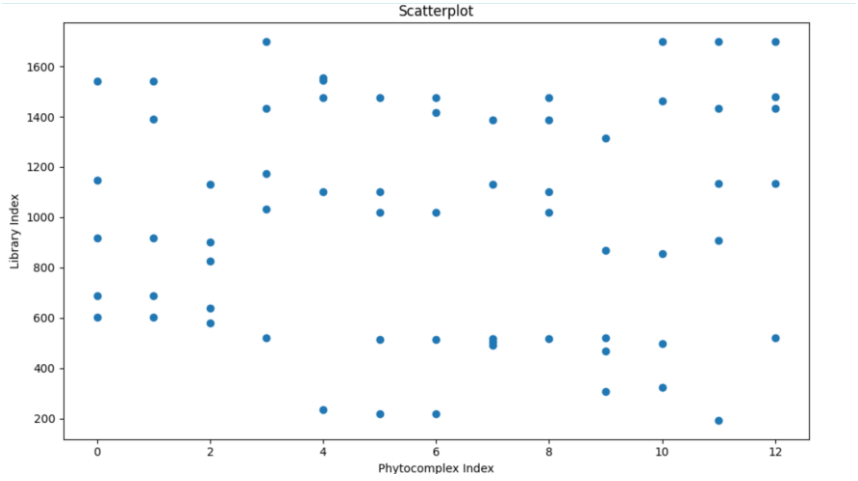
COC1=C(C=CC(=C1)/C=C/C(=O)O)O

Caffeic acid

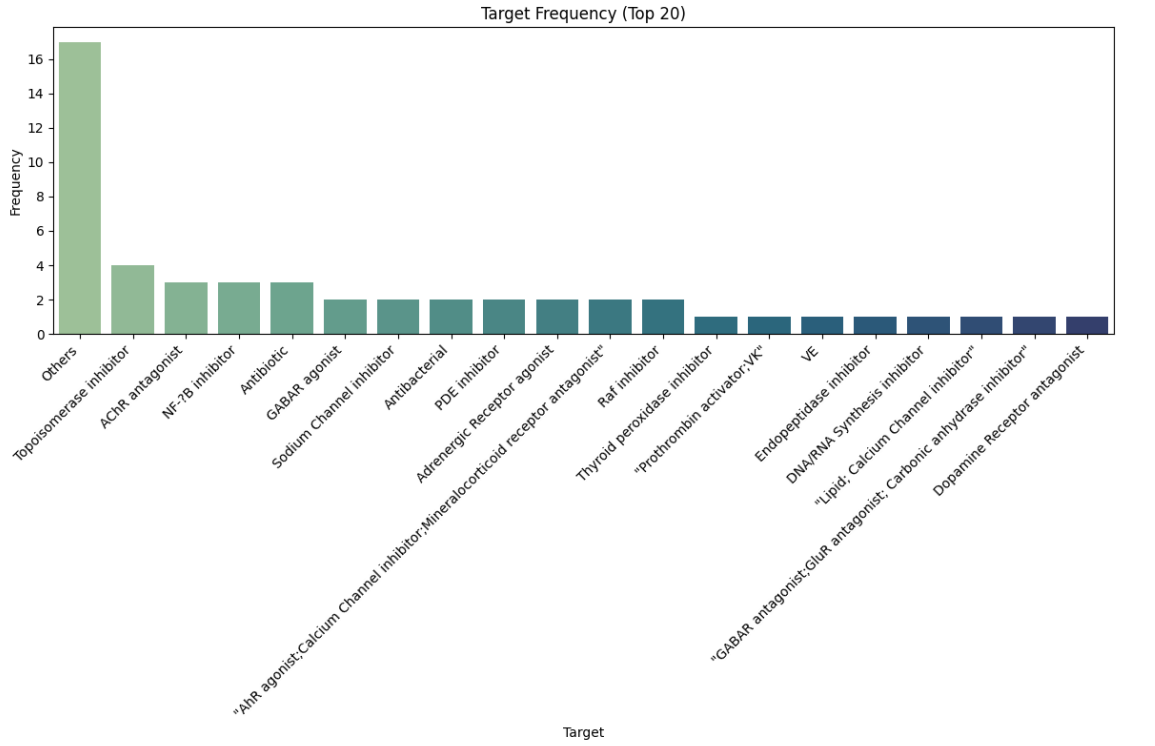
C1=CC(=C(C=C1/C=C/C(=O)O)O)O

- Load the compounds in smiley format (one per line), you can help yourself with a text editor and copy paste all the molecules at once
- Press the "processes molecules" button

Result: you will get 4 graphs and a table

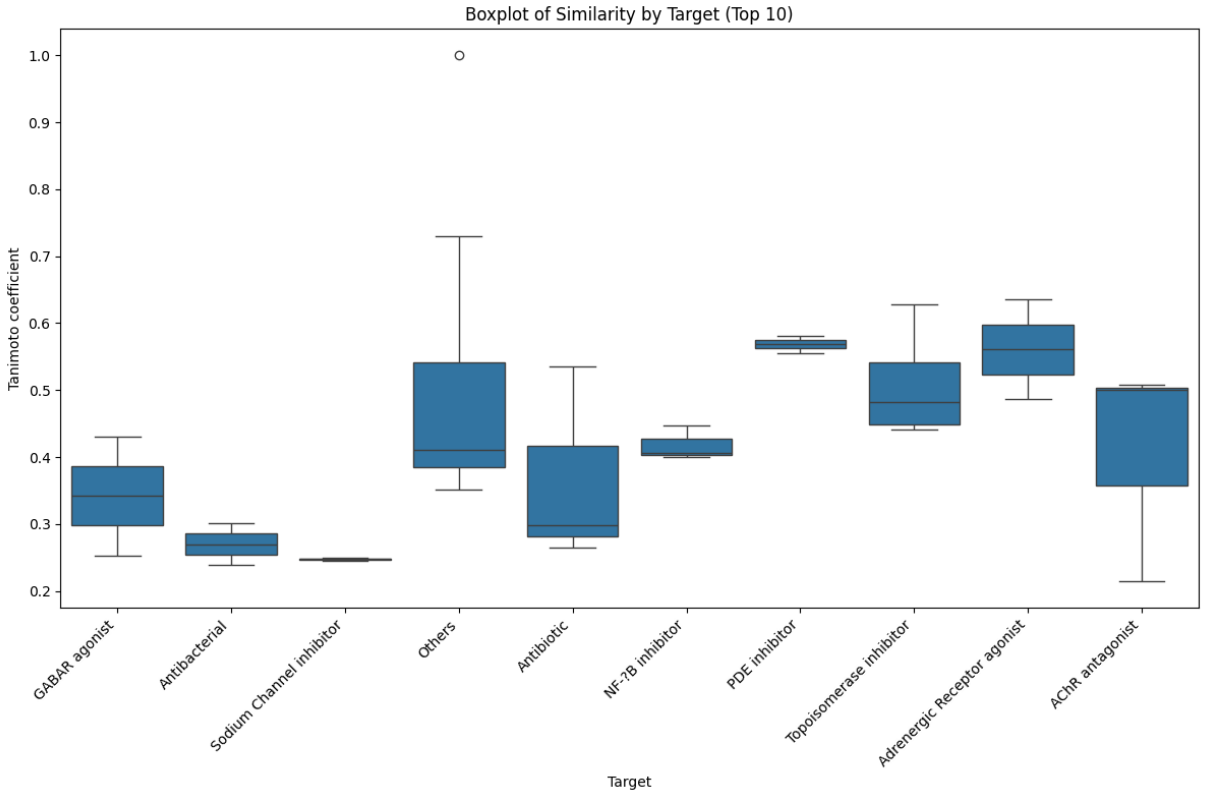


This scatter plot displays the comparison of fingerprint similarities between compounds extracted from rosemary and a reference library. Each point represents a compound, with its position determined by its similarity score on the x-axis (Phytocomplex Index) and the corresponding library compound's score on the y-axis (Library Index). This visualization allows for the assessment of how closely the rosemary compounds match those in the library, potentially indicating similar bioactivity.



The bar chart illustrates the frequency of different protein targets associated with rosemary compounds based on Tanimoto similarity scores. The most frequent target is classified as "Others," indicating that the library contains known compounds with unknown specific protein targets. This category may include compounds like limonene, which are recognized for their bioactivity but lack a definitive target in the current library. Among the identified targets, **topoisomerase inhibitors** are notable due to their emerging interest in therapies against infectious and cancerous cells. Topoisomerases are enzymes crucial for DNA replication and repair, making them attractive targets for therapeutic intervention. The presence of topoisomerase inhibitors in rosemary suggests potential applications in these areas, warranting further investigation. Other frequent targets, such as AChR antagonists, NF-kB inhibitors, and antibacterial compounds, also point to diverse potential bioactivities of rosemary extracts.

Boxplot



This boxplot displays the distribution of Tanimoto similarity scores between rosemary compounds and known bioactive compounds in a reference library, categorized by their target protein.

How to read the boxplot:

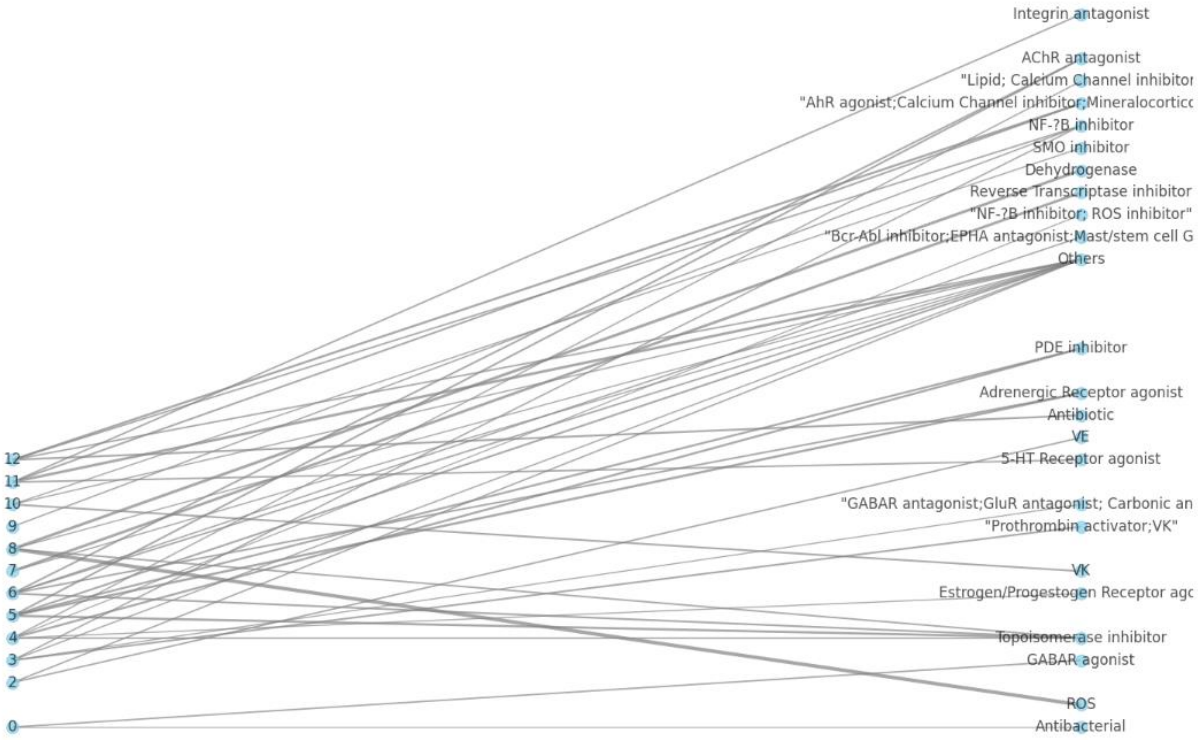
- **The box:** Represents the interquartile range (IQR), containing the middle 50% of the data. The line inside the box represents the median [1] similarity score.
- **The whiskers:** Extend from the box to the lowest and highest values within 1.5 times the IQR. Points outside the whiskers are considered outliers.

Results for rosemary compounds:

- **GABAR agonist and Antibacterial:** These targets show a wide range of similarity scores, suggesting that rosemary contains compounds with varying degrees of similarity to known GABAR agonists and antibacterials.
- **Sodium Channel inhibitor:** The data is tightly packed with low similarity scores suggesting that rosemary compounds do not exhibit strong similarity to known Sodium Channel inhibitors.
- **Others:** This category, representing compounds with unknown targets, also shows a wide distribution of similarity scores.
- **Topoisomerase inhibitor:** The boxplot indicates that rosemary contains compounds with moderate similarity to known topoisomerase inhibitors, with a cluster of higher similarity scores suggesting potential interactions with this important target.
- **Adrenergic Receptor agonist and AChR antagonist:** Similar to topoisomerase inhibitors, rosemary exhibits moderate similarity to these targets as well.

Overall, the boxplot highlights the diverse potential bioactivities of rosemary compounds, with notable similarity observed for several targets, including topoisomerase inhibitors, GABAR agonists, antibacterials, adrenergic receptor agonists, and AChR antagonists. The wide range of similarity scores within some targets suggests that rosemary may contain a variety of compounds with varying degrees of potency or efficacy.

Network graph



This network graph represents the interactions between compounds extracted from rosemary and a library of known bioactive molecules, visualized as a bipartite network.

Nodes on the left represent rosemary compounds, while **nodes on the right** represent target proteins or biological activities. **Edges connecting these nodes** indicate a predicted interaction or similarity between a rosemary compound and a specific target, based on fingerprint analysis.

The graph highlights the diverse potential bioactivities of rosemary, as its compounds show interactions with a variety of targets, including:

- **Enzymes:** Topoisomerase, PDE inhibitor, Dehydrogenase
- **Receptors:** Adrenergic Receptor agonist, AChR antagonist, 5-HT-Receptor agonist, GABAR agonist, Estrogen/Progestogen Receptor agonist
- **Other targets:** Antibacterial, Antibiotic, Integrin antagonist, NF-kB inhibitor, and others.

The **thickness of the edges** could indicate the strength of the interaction or the degree of similarity between the compounds and targets. A higher number of connections for a rosemary compound suggests it might have multiple bioactivities. The "Others" category on the right indicates compounds in the library with known bioactivity but unspecified targets, highlighting areas where further research is needed. Overall, this network visualization provides a comprehensive overview of the potential molecular interactions of rosemary compounds, suggesting its diverse pharmacological potential.

Comparison Results (Tanimoto coefficient):

Query Index	Library Index	Tanimoto Similarity	Target	Bioactivity
0	918	0.431	GABAR agonist	The immediate precursor in the biosynthesis of SEROTONIN from tryptophan. It is used as an antiepileptic and antidepressant.
0	604	0.302	Antibacterial	Aliskiren hemifumarate is a direct renin inhibitor
0	688	0.262	Raf inhibitor	Telmisartan is an angiotensin II receptor antagonist (ARB) used in the management of hypertension. Generally, angiotensin II receptor blockers (ARBs) such as telmisartan bind to the angiotensin II type 1 (AT1) receptors with high affinity,
0	1148	0.256	Endopeptidase inhibitor	Polyphenolic compounds with molecular weights of around 500-3000 daltons and containing enough hydroxyl groups (1-2 per 100 MW) for effective cross linking of other compounds (ASTRINGENTS). The two main types are HYDROLYZABLE TANNINS and CO
0	1541	0.250	Sodium Channel inhibitor	Terpin hydrate is an expectorant, commonly used to loosen mucus in patients presenting with acute or chronic bronchitis, and related conditions. It is derived from sources such as oil of turpentine, oregano, thyme and eucalyptus. Though it
1	918	0.253	GABAR agonist	The immediate precursor in the biosynthesis of SEROTONIN from tryptophan. It is used as an antiepileptic and antidepressant.
1	1541	0.245	Sodium Channel inhibitor	Terpin hydrate is an expectorant, commonly used to loosen mucus in patients presenting with acute or chronic bronchitis, and related conditions. It is derived from sources such as oil of turpentine, oregano, thyme and eucalyptus. Though it
1	604	0.239	Antibacterial	Aliskiren hemifumarate is a direct renin inhibitor

This table presents the initial results of a Tanimoto coefficient-based comparison between compounds extracted from rosemary (Query Index) and a library of known compounds (Library Index).

How to Read the Table:

- **Query Index:** A numerical identifier for each compound extracted from rosemary.
- **Library Index:** A numerical identifier for each compound in the reference library.
- **Tanimoto Similarity:** A score between 0 and 1 representing the similarity between the rosemary compound and the library compound. Higher scores indicate greater similarity.
- **Target:** The known protein target or biological activity associated with the library compound.
- **Bioactivity:** A brief description of the known bioactivity of the library compound.

Initial Results Summary:

The table shows that rosemary compound 0 exhibits the highest similarity (0.431) to a compound in the library known to be a GABAR agonist. This suggests that rosemary compound 0 may also interact with GABAR, a neurotransmitter receptor in the brain. Additionally, rosemary compound 0 shows some similarity (0.302) to a compound with antibacterial activity.

Rosemary compound 1 also shows the highest similarity to the same GABAR agonist (0.253), though with a lower score than compound 0. It also exhibits similarity to an antibacterial compound (0.239).

The table provides a glimpse into the potential bioactivities of rosemary compounds, with the highest similarity scores observed for compounds associated with GABAR agonism and antibacterial activity. Further analysis of the complete dataset is needed to draw comprehensive conclusions about the full range of potential bioactivities in rosemary.

References

1. JOSE´ DEL CAMPO, MARIE-JOSE'PHE AMIOT, AND CHRISTOPHE NGUYEN-THE
   Antimicrobial Effect of Rosemary Extracts
   Journal of Food Protection, Vol. 63, No. 10, 2000, Pages 1359–1368
2. PubChem
   https://pubchem.ncbi.nlm.nih.gov/
3. Approveddrugslibrary. Thanks to Dr. William Bains for publishing the library on Researchgate
   https://www.researchgate.net/post/FDA_approved_drugs_in_SDF_format

Acknowledgment

Website design and content creation supported by Gemini 2.0 Flash.

Disclaimer

This tool is for research purposes only and not for medical advice.