

INTRODUCTION

The prediction of targets for bioactive compounds identified in medicinal plants traditionally used in the world and their potential correlations with the diseases reported in this use or prospected by *in silico* target prediction research help in the evolution of studies involving these diseases and accelerate the discovery of drugs. They also reveal the potential for new effects of these compounds. Among the countless medicinal species, we have *Syzygium cumini* (L.) Skeels, a tropical tree (Myrtaceae) popularly known as “Jambolao” (Brazilian Portuguese).

OBJECTIVE

Our aim was to provide more information about *Syzygium cumini*, such as the pharmacological effects of its compounds and a better understanding of related diseases.

RESULTS AND DISCUSSION

SwissTargetPrediction

Virtual Screening

Source	Target	Common name	UniProt ID	CHEMBL ID	Target Class
1,2,6-Trimethylglucose	Squalene monooxygenase (SQUAL)		Q14534	CHEMBL3592	Enzyme
1-galloylglucose	Aldose reductase	AKR1B1	P15221	CHEMBL3900	Enzyme
Beta-Sitosterol	Niemann-Pick C1-like protein NPC1L1	Q9UHC9	CHEMBL2027		Other membrane protein
Betulinic acid	SUMO-activating enzyme sub-SAE1	Q9UBED	CHEMBL2095174		Enzyme
Betulinic acid	SUMO-activating enzyme sub-UBA2	Q9UBI2	CHEMBL2095174		Enzyme
Betulinic acid	DNA polymerase beta (By the POLB)	P09746	CHEMBL2392		Enzyme
Betulinic acid	Aldo-keto reductase family 1 AKR1B10	Q60218	CHEMBL5583		Enzyme
Caffeic Acid	Carbonic anhydrase II	CA2	P00918	CHEMBL205	Lyase
Caffeic Acid	Arachidonate 5-lipoxygenase ALOX5	P09917	CHEMBL215		Oxidoreductase
Caffeic Acid	Carbonic anhydrase VII	CA7	P43166	CHEMBL2326	Lyase
Caffeic Acid	Carbonic anhydrase I	CA1	P09915	CHEMBL261	Lyase
Caffeic Acid	Carbonic anhydrase VI	CA6	P23280	CHEMBL3025	Lyase
Caffeic Acid	Matrix metalloproteinase 9	MMP9	P14780	CHEMBL321	Protease
Caffeic Acid	Carbonic anhydrase XII	CA12	Q43570	CHEMBL3242	Lyase
Caffeic Acid	Matrix metalloproteinase 1	MMP1	P03956	CHEMBL332	Protease
Caffeic Acid	Matrix metalloproteinase 2	MMP2	P08253	CHEMBL333	Protease
Caffeic Acid	Protein-tyrosine phosphatase PTPN1	P18031	CHEMBL335		Phosphatase
Caffeic Acid	Carbonic anhydrase XIV	CA14	Q5ULV7	CHEMBL3510	Lyase
Caffeic Acid	Carbonic anhydrase IX	CA9	Q18790	CHEMBL3594	Lyase
Caffeic Acid	Carbonic anhydrase VB	CASB	Q9Y2D0	CHEMBL3969	Lyase
Caffeic Acid	Carbonic anhydrase VA	CASA	P35218	CHEMBL4789	Lyase
Alfloxacin	Aldehyde reductase	AKR1B1	D15131	CHEMBL1900	Enzyme

Figure 2. Target list of *S. cumini* compounds resulting from target prediction.

METHODS

Pubchem database

S. cumini compounds

VIDA 4.4.0²

22 compounds

Structure preparatio

SwissTargetPredic₃

Target Prediction

125 targets

STRING 11.04 DAVID 6.8^{5,6,7}

Functional integration

Cytoscape 3.8.0⁸

6 diseases

Network

This number represents the 125 different targets found by predicting targets for a probability equal to or greater than 80%.

STRING 11.0 Functional integration

The opacity of the links between the targets refers to the level of trust that is related to the reliability of the analysis.

This complex network reveals 602 connections between the targets.

Figure 3. Network of interaction between the targets of the identified compounds of *S. cumini*.

INTRODUCTION

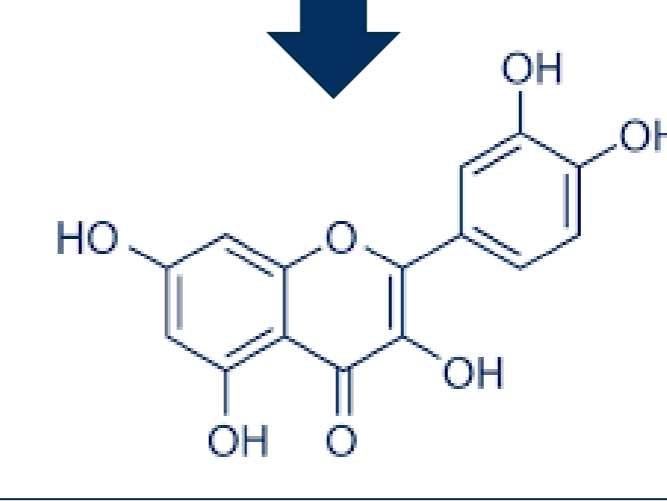
Prospection of bioactive compounds

Bioactive compounds

Target prediction



Syzygium Cumini



Target, disease and gene correlation

Figure 1. Prospection of *S. cumini* bioactive compounds, virtual screening and correlation of genes corresponding to diseases.

RESULTS AND DISCUSSION

Cytoscape 3.8.0

Network

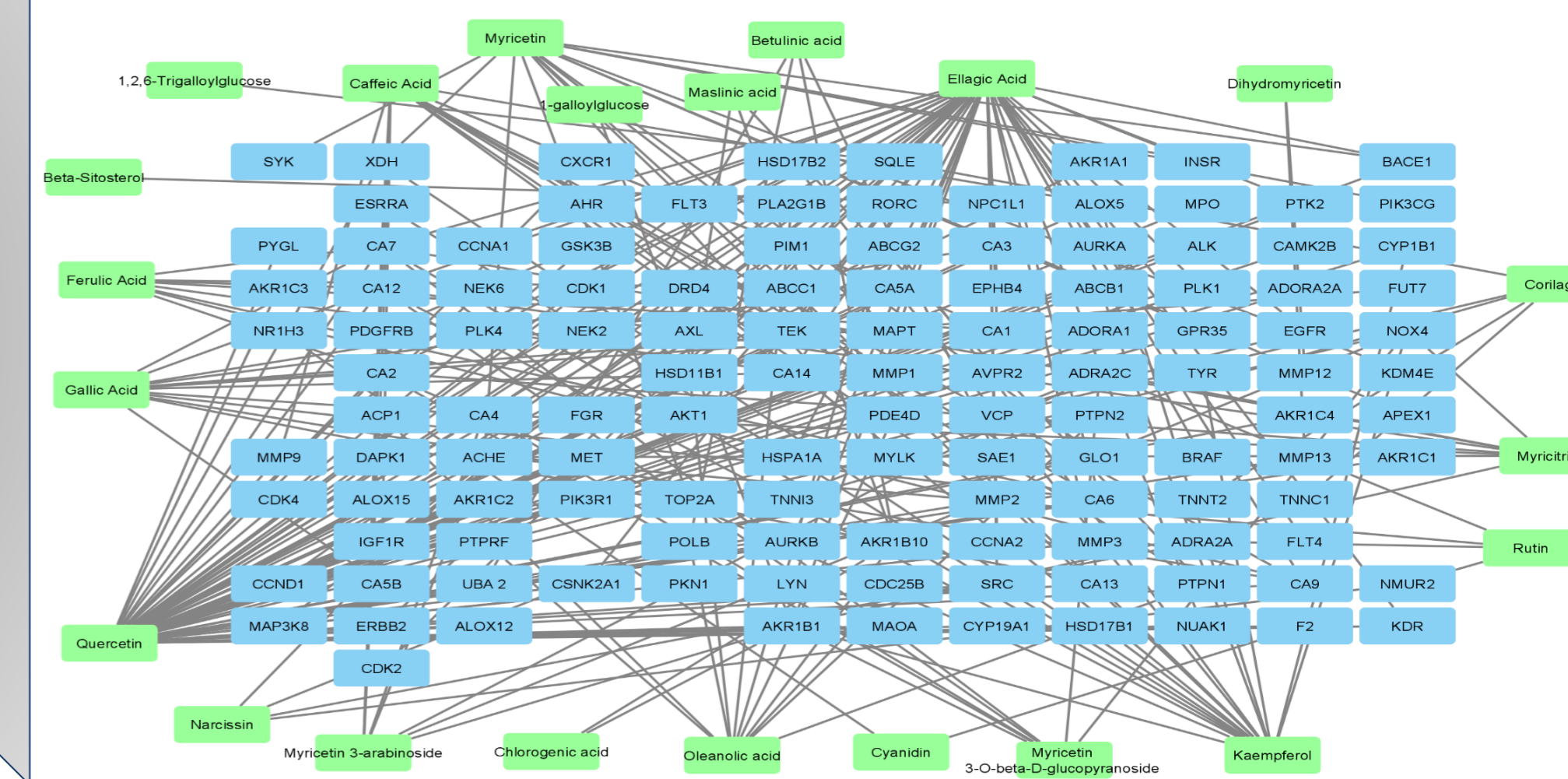


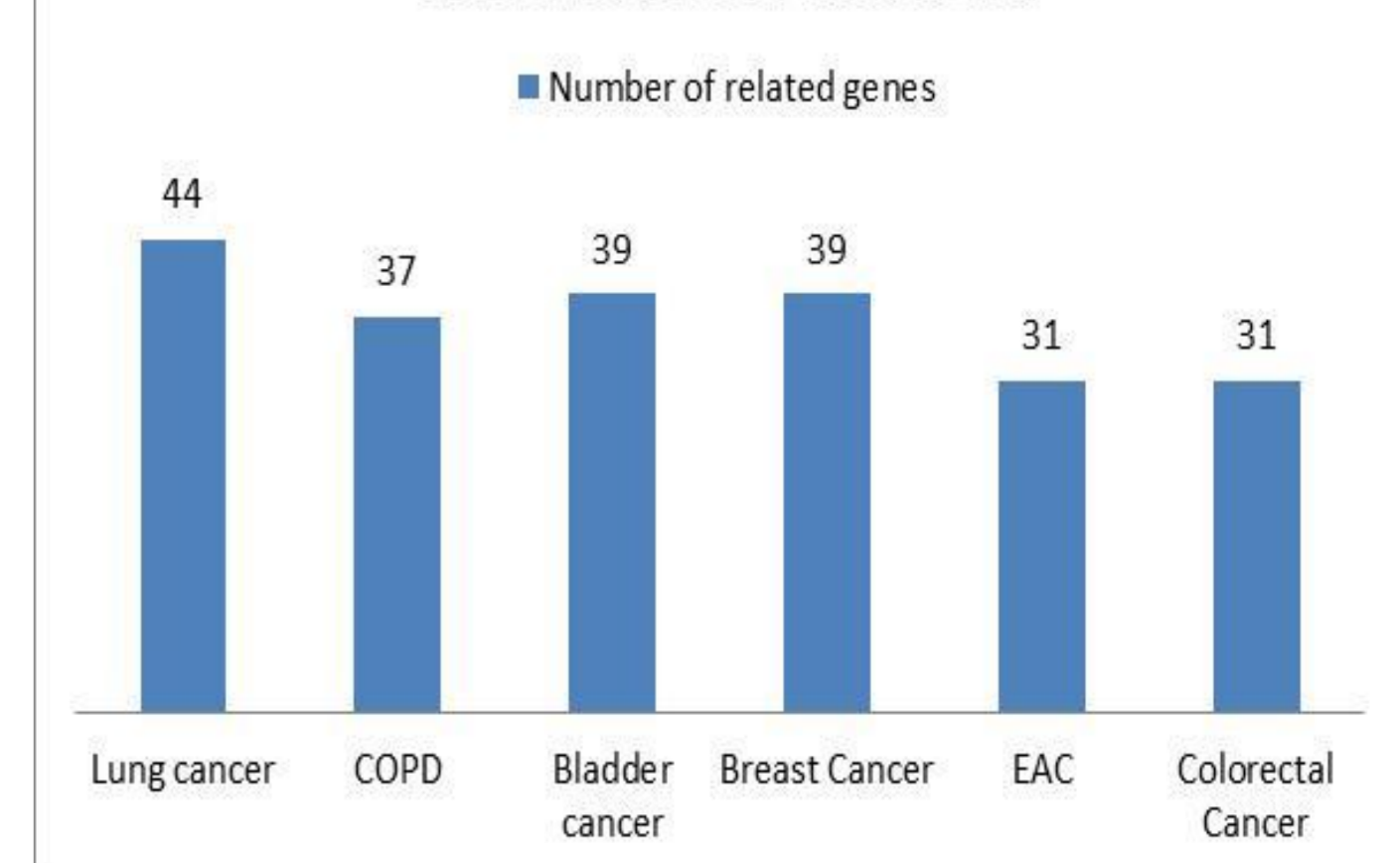
Figure 4. Network of interactions between compounds (green) and targets (blue).

This complex network reveals 242 connections between compounds and targets.

DAVID 6.8

Functional integration

Gene-related diseases



Graphic 1. Possible diseases related to targets quantified in descending order of reliability.

This graph reveals the influence of *S. cumini* compounds, through the targets found in virtual screening, with diseases such as: lung cancer, chronic obstructive pulmonary disease (COPD), bladder cancer, breast cancer, adenocarcinoma esophagus (EAC) and colorectal cancer.

CONCLUSION

Our results suggest a wide application of the identified compounds of *S. cumini*. The analyzes carried out revealed as targets, some types of cancer and COPD. Therefore, the bioactive compounds of *S. cumini* possibly have pharmacological and clinical interest effects.

REFERENCES

1. Franco RR, Ribeiro Zabisky LF, Pires de Lima Júnior J, Mota Alves VH, Justino AB, Saraiva AL, Goulart LR, Espindola FS. Antidiabetic effects of *Syzygium cumini* leaves: A non-hemolytic plant with potential against process of oxidation, glycation, inflammation and digestive enzymes catalysis. *J Ethnopharmacol.* 2020 Oct 28;261:113132. doi: 10.1016/j.jep.2020.113132.
2. The software VIDA 4.4.0, OpenEye Scientific Software, Santa Fe, NM.
3. Antoine Daina, Olivier Michielin, Vincent Zoete, SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules, *Nucleic Acids Research*, Volume 47, Issue W1, 02 July 2019, Pages W357–W364, <https://doi.org/10.1093/nar/gkz382>.
4. B. Snel, G. Lehmann, P. Bork, M. A. Huynen, STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene, *Nucleic Acids Research*, Volume 28, Issue 18, 15 September 2000, Pages 3442–3444, <https://doi.org/10.1093/nar/28.18.3442>.
5. Da Wei Huang, Brad T. Sherman, Qina Tan, Joseph Kir, David Liu, David Bryant, Yongjian Guo, Robert Stephens, Michael W. Baseler, H. Clifford Lane, Richard A. Lempicki, DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists, *Nucleic Acids Research*, Volume 35, Issue suppl_2, 1 July 2007, Pages W169–W175, <https://doi.org/10.1093/nar/gkm415>.
6. Da Wei Huang, Brad T. Sherman, Richard A. Lempicki, Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists, *Nucleic Acids Research*, Volume 37, Issue 1, 1 January 2009, Pages 1–13, <https://doi.org/10.1093/nar/gkn923>.
7. Huang, D., Sherman, B. & Lempicki, R. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4, 44–57 (2009). <https://doi.org/10.1038/nprot.2008.211>.
8. P. Shannon, "Cytoscape: A software environment for integrated models of biomolecular interaction networks", *Genome Res.*, vol. 13, no. 22, pp. 2498-2504, 1971.