A network-based algorithm for identifying drug repurposing opportunities for complex diseases

G. Fiscon[1,2], F. Conte[2], and P. Paci[1,2,3]
[1] *Department of Computer, Control and Management Engineering "A. Ruberti", Sapienza University of Rome;*
[2] *Institute for Systems Analysis and Computer Science "A. Ruberti", National Research Council of Rome;*
[3] *Karolinska Institutet, Stockholm, Sweden*

***Abstract*— SAveRUNNER is a recently developed network-based tool to efficiently identify novel medical indications for currently approved drugs (known as drug repurposing strategy). Up to now, SAveRUNNER has been gainfully applied to unveil repurposable solutions for several diseases, including viral infection (e.g., SARS, COVID-19, HIV, Malaria, Ebola), breast cancer, progressive disorders of central nervous system (e.g., Amyotrophic Lateral Sclerosis, Multiple Sclerosis), and other neurodegenerative diseases (e.g., Alzheimer's Disease). Here, SAveRUNNER algorithm and its main applications are described.**

***Keywords*— Network Theory, Network Medicine, Drug Repurposing.**

I. Introduction

Drug repurposing represents an effective drug discovery strategy to find new uses from existing drugs out of the scope of their original medical indication, which can shorten the time, reduce the cost, and increase the success rate of molecular entities compared to the traditional de novo drug discovery [1].

In this context, we recently developed a network-based algorithm, called SAveRUNNER (Searching off-lAbel dRUg aNd NEtwoRk) [2], [3], which predicts new uses for already approved drugs by exploiting concepts from the emerging field of Network Medicine [4]. According to the Network Medicine paradigm, the human interactome network (i.e., the cellular network of all physical molecular interactions) [5] can be considered as map, where the molecular determinants of a given disease (*disease genes*) tend to co-localize and agglomerate in specific regions (*disease modules*), whose perturbation can contribute to the manifestation of the pathophenotype [4]. In this perspective, even the action of a drug can be considered as a local perturbation of the human interactome and thus, for a drug to be effective against a specific disease, its target proteins should be in the immediate vicinity of the corresponding disease module [6]. Following these principles, SAveRUNNER predicts drug-disease associations by quantifying the interplay between the drug targets and the disease-associated genes in the human interactome through the computation of a novel network-based similarity measure, which prioritizes associations between drugs and diseases located in the same network neighbourhoods.

In the last two years, SAveRUNNER has been acknowledged as promising tool to find drugs repurposable opportunities for several diseases (Fig. 1). In particular, the pioneer application was for predicting repurposable drugs against COVID-19 [3], [7]. In the same study, we also analyzed other related diseases linked to the COVID-19 infection as genetic similarity (i.e., Severe Acute Respiratory Syndrome - SARS), as risk factors (e.g., cardiovascular diseases, diabetes mellitus), or other viral infections (i.e., malaria, HIV and Ebola), and immune disorders (i.e., rheumatoid arthritis) [3]. More recently, it has been successfully applied to identify candidate repurposable drugs also for Amyotrophic Lateral Sclerosis (ALS) [8], Multiple Sclerosis (MS) [9], breast invasive carcinoma subtypes [10], and neurodegenerative diseases, such as Alzheimer's Disease (AD) [11] (Fig. 1). In all these studies, SAveRUNNER turned out as a novel promising tool for an efficient screening of the repurposable opportunities, outlining where looking at for further ad-hoc experimental validations.

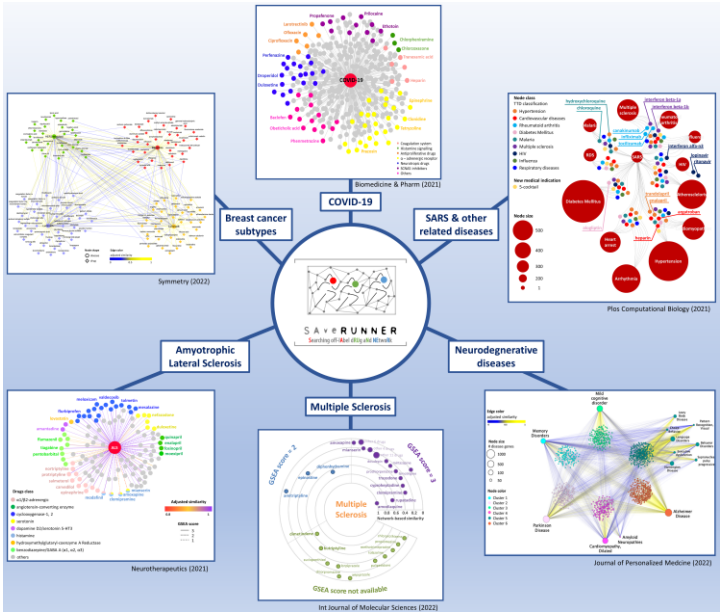

Fig. 1. Overview of SAveRUNNER applications. Each panel corresponds to a different study, where the drug-disease network computed by SAveRUNNER was shown. In these networks, nodes refer to drugs (smaller nodes) and diseases of interest (bigger nodes), and a link occurs between a drug and a disease if their corresponding drug-target proteins and disease-associated genes are proximal in the human interactome more than by chance. The link is weighted according to the novel network-based similarity measure implemented by SAveRUNNER [3].

II. Methods

A. Input data

1) Human interactome network

The human interactome used in all the SAveRUNNER applications was downloaded from [6], including 217,160 interactions connecting 15,970 unique proteins.

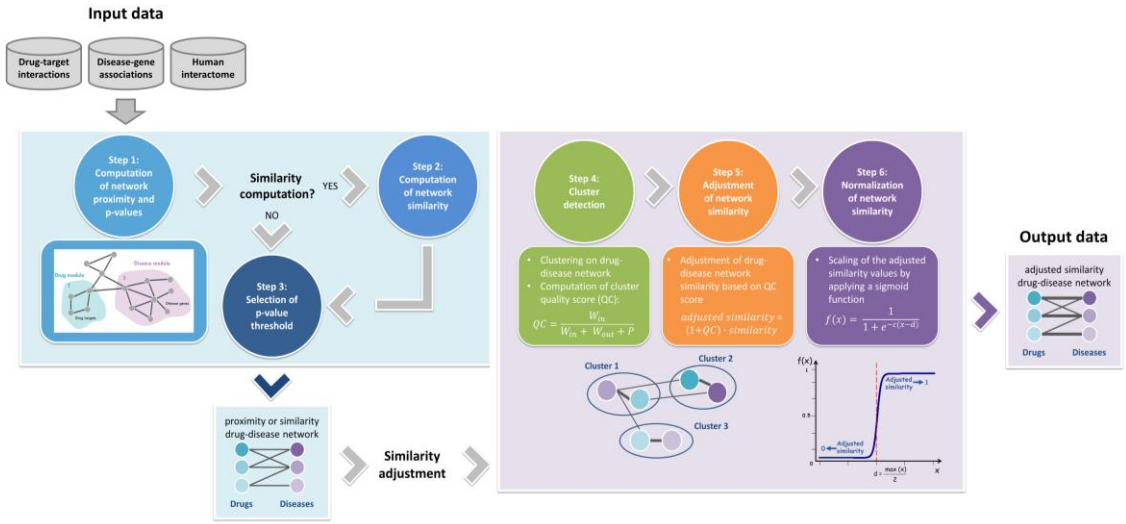

Fig. 2. Overview of SAveRUNNER algorithm [2], [3]**.** Taking as input a list of disease genes, drug targets, and the human interactome, SAveRUNNER releases a drug-disease network via six steps. In steps 1-3, it builds a proximity-based bipartite drug-disease network, where nodes are both drugs and diseases and edges are the statistically significant drug-disease associations. In steps 4-7, it builds a similarity-based bipartite drug-disease network, where the weights represent the adjusted similarity measure computed to prioritize the predicted drug-disease associations by rewarding the associations between drugs and diseases belonging to the same network cluster.

2) Disease gene associations

Disease-associated genes were retrieved from Phenopedia (last release April 2020) [12] and DisGeNET (last release October 2021) [13], two of the most popular and largest publicly available collections of disease genes obtained by integrating and homogeneously annotating data from expert-curated repositories and GWAS catalogues. For one of SAveRUNNER application to COVID-19 [7], transcriptomic data from whole blood cells of patients with COVID-19 was retrieved from GEO repository (GSE152641) and differentially expressed genes between infected cells and heathy controls were used as *disease genes*. Yet, for SAveRUNNER application to breast cancer subtypes [10], due to a poor knowledge of the genes associated with each subtype, transcriptomic data from patients affected by breast invasive carcinoma was retrieved from The Cancer Genome Atlas (TCGA) repository and stratified according to PAM50 classification [14]. Then, a list of genes (known as *switch genes*) previously identified by SWIM software [15] as likely associated to each subtype were used as *disease genes* [14].

3) Drug-target interactions

Drug-target interactions were retrieved from DrugBank [16] (5.1.6 release of April 2020), the most acknowledged database about drugs information and drug-target interactions. The targets' Uniprot IDs provided by DrugBank were mapped to Entrez gene IDs by using the BioMart – Ensembl tool (https://www.ensembl.org/), obtaining a total of 2,165 genes interacting with 1,873 drugs.

B. Algorithm

SAveRUNNER algorithm takes as input the human interactome, a list of disease-gene associations related to disease(s) of interest, and the list of drug-target interactions and, via the implementation of several steps depicted in Fig. 2 and detailed in [2], [3], gives as output a weighted drug-disease bipartite network.

In the network released by SAveRUNNER, nodes are drugs and diseases, and a drug and a disease are connected if the corresponding drug targets and disease genes are closer in the interactome than expected by chance (p-value ≤ 0.05). The weight of their association is based on the implementation of a novel network-based similarity measure, known as *adjusted similarity* (AS), defined in Eq. (1).

$$AS(p) = \frac{1}{1+e^{-c\left[\frac{(1+QC)(m-p)}{m} - d\right]}} \quad (1)$$

In Eq (1), p refers to the network proximity measure described in Eq. (2) [6].

$$p(T,S) = \frac{1}{\|T\|}\sum_{t\epsilon T} \min_{s\epsilon S} d(t,s) \quad (2)$$

where $p(T,S)$ represents the average shortest path length between drug target t in the drug module T and the nearest disease gene s in the disease module S; QC is a quality cluster score that rewards associations between drugs and diseases located in the same network cluster; m is max(p); c and d are the steepness and the midpoint of AS(p), respectively. A comprehensive description of SAveRUNNER methodology can be found in [2], [3].

C. GSEA

To test whether the repurposable drugs predicted by SAveRUNNER could counteract the gene expression perturbations caused by the disease of interest (i.e., up-regulate genes down-regulated by the disease or vice versa), the so-called gene set enrichment analysis (GSEA) was downstream performed [3] in most of the studies. This analysis takes as input: (i) differentially expressed genes of the understudied disease (generally obtained by GEO or TCGA repository) to use as *disease signature*; (ii) differentially expressed genes of drug-treated human cell lines (obtained by Connectivity Map (CMap) database [17]) to use as *drug signature*. For each drug both included in the CMap database and predicted by SAveRUNNER, an enrichment score (ES) was computed to evaluate whether the effect of the drug could counteract the effect of the disease (ES < 0, drug and disease signature negatively correlated) or

not (ES > 0, drug and disease signature positively correlated) [17]. For each tested drug, a GSEA score equal to 1 was assigned if that drug was found to have ES < 0 with respect to a given disease, or equal to 0, otherwise. By repeating this procedure for several independent disease signatures, the final GSEA score ranges from 1 to *n* (with *n* is the total number of disease signatures tested): the greater the GSEA score, the greater the number of disease datasets satisfying this criterion.

III. Results

A. COVID-19, SARS, and other related diseases

SAveRUNNER was first applied to COVID-19 and a panel of 14 diseases COVID-related for genetic similarity, comorbidity, or for their association to drugs with ongoing clinical trials for treating COVID-19 patients [3]. In particular, we tested: SARS, as it is caused by the coronavirus with the highest sequence identity with SARS-CoV-2 [18] and there existed a well-established knowledge of its associated disease genes; diabetes, cardiovascular diseases, and hypertension, whose comorbidity in COVID-19 patients was observed; and finally other viral infections (i.e., malaria, HIV and Ebola) and immune disorders (i.e., rheumatoid arthritis), since drugs approved for their treatment were investigated for their potential effect to fight coronavirus disease [19]. In this study, SAveRUNNER prioritized 282 potential anti-SARS repurposable drugs and 98 anti-COVID-19, including ACE-inhibitors, monoclonal antibodies, and thrombin inhibitors that were further in-silico corroborated by GSEA, considering drug-treated cell lines from CMap and coronaviruses-induced transcriptomics data from GEO repository.

SAveRUNNER was also used to study COVID-19 starting from differentially expressed genes obtained by exploiting transcriptomic data from COVID-19 infected patients and healthy control samples [7]. In this study, SAveRUNNER identified 399 repurposable drugs for COVID-19, whose top-ranked ones were molecules involved in the modulation of the coagulation system (e.g., *heparin* and *tranexamic acid*), antihistaminic drugs, mast cell stabilizers (e.g., *chlorzoxazone* and *chlorpheniramine*), anti-proliferative drugs including tyrosine kinase (TRK) inhibitors and antibiotics (e.g., *larotrectinib* and *ciprofloxacin*), alpha-adrenergic receptor agents (e.g., *clonidine* and *prazosin*); drugs affecting the central nervous system (e.g., *perfenazine* and *droperidol*), and inhibitors of the sodium voltage-gated channel alpha subunit 5 (SCN5A) involved in cardiac rhythm control (e.g., *propafenone* and *prilocaine*).

B. Amyotrophic Lateral Sclerosis and Multiple Sclerosis

SAveRUNNER was applied to study ALS [8] and MS [9]. In both cases, by prioritizing the network-predicted drugs according to the decreasing value of their network similarity with ALS and MS, SAveRUNNER highlighted some interesting drug modulators of histamine receptors. The expression of histamine receptors were shown to increase in cortical tissues of patients affected by MS [9], thus suggesting that histamine-related molecules can be potential therapeutic strategies for the disease. Moreover, it has been shown that the histaminergic modulation might participate to the mechanisms of ALS insurgence, progression, and therapy as well, since other studies found that histamine-related genes could be modifiers in ALS, and thus supporting their role as candidate biomarkers and therapeutic targets [20]. The histaminergic compounds predicted by SAveRUNNER were also in-silico-validated by GSEA analysis, considering drug-treated cell lines from CMap as drug signatures, and three different datasets of transcriptomics data of cortex tissues from patients affected by ALS [8] and MS [9] collected from GEO repository as disease signatures for ALS and MS, respectively.

C. Breast cancer subtypes

SAveRUNNER was applied to study four subtypes of breast invasive carcinoma stratified based on PAM50 classification (i.e., luminal A, luminal B, Her2-enriched, and basal-like) [10]. In this study, SAveRUNNER predicted: (i) 74 repurposable drugs for luminal A, including as top-ranked *teniposide* that has been studied as suppressor of the growth of subcutaneous cell lines of luminal A (MCF-7) [21]; (ii) 54 repurposable drugs for luminal B, including as top-ranked *ivabradineI*, recently repurposed as a novel therapy for breast cancer [22]; (iii) 79 repurposable drugs for HER2-enriched, including as top-ranked several compounds for the treatment of bacterial infections; and (iv) 77 repurposable drugs for basal-like, whose top-ranked one was *metergoline*, acting as antagonist of certain subtypes of 5-HT receptor, which has been showed to induce cell proliferation of triple-negative breast cancer via 5-HT7 receptor signaling [23]. Yet, the 20% of the total predicted drugs have also an indication originally related to cancer treatment, including *dexrazoxane*, *ixabepilone*, *capecitabine* already approved for fighting breast cancers. Most of the predicted repurposable drugs were also in-silico-validated by GSEA, considering breast cancer drug-treated cell lines from CMap as drug signatures and transcriptomics data of breast cancer patients from TCGA as diseases signature.

D. Neurodegenerative diseases

SAveRUNNER was applied to study AD and other 13 AD-related disorders, including Mild Cognitive Impairment (MCI), which is the stage between the expected cognitive decline of normal aging and the more serious decline of dementia, other neurological diseases (i.e., Parkinson's disease, Huntington disease, supranuclear palsy progressive, and Lewy body disease), and other diseases that can be symptoms, risk factors, or consequences of AD (i.e., language disorders, memory disorders, visual pattern recognition, executive dysfunction, choice behavior, cardiomyopathy, amyloid neuropathies). SAveRUNNER predicted some interesting compound belonging to the family of: (i) beta-blockers, including *betaxolol*, originally used for hypertension treatment and already linked to a lower risk of developing AD [24]; (ii) multi-kinase inhibitors, including *regorafenib*, whose beneficial effects were recently investigated on neuroinflammation and AD pathology in vitro and in vivo [25]; (iii) anti-platelet inhibitors, including *clopidogrel*, whose therapy has been associated with lower risk of developing AD and whose anti-inflammatory effect has been recently studied in in-vivo model of AD [26]; and mTOR inhibitors, including *sirolimus*, an immunosuppressant drug used in transplantation medicine to prevent organ rejection, whose modulation could have

potential application to delay memory loss and age-related neurodegenerative diseases [27]. All these compounds were also corroborated by GSEA, considering neural drug-treated cell lines available from CMap as drug signatures and three different datasets of transcriptomic data from brain tissues of patients affected by AD from GEO as disease signature.

IV. Conclusion

We presented an overview of SAveRUNNER, a network-based approach that offers a promising framework to efficiently detect putative novel uses for marketed drugs against diseases of interest. We reviewed its main applications, including but not limited to viral infections, breast cancer, and neurodegenerative progressive disorders.

Acknowledgments

This work was partially funded by PON Ricerca e Innovazione 2014-2020 – DM 1062/202, BiBiNet project (grant n: H35F21000430002) within the POR-Lazio FESR 2014-2020, by Finalizzata Giovani Ricercatori 2021 (grant n.: GR-2021-12372614, CUP: B83C22007560001), by Progetto di Ricerca di Ateneo 2021 (grant n: RM12117A34663A2C) and 2022 (grant n: RP12218139131053) of Sapienza University of Rome.

References

[1] S. Pushpakom *et al.*, «Drug repurposing: progress, challenges and recommendations», *Nat. Rev. Drug Discov.*, vol. 18, fasc. 1, pp. 41–58, gen. 2019, doi: 10.1038/nrd.2018.168.

[2] G. Fiscon e P. Paci, «SAveRUNNER: an R-based tool for drug repurposing», *BMC Bioinformatics*, vol. 22, fasc. 1, p. 150, mar. 2021, doi: 10.1186/s12859-021-04076-w.

[3] G. Fiscon, F. Conte, L. Farina, e P. Paci, «SAveRUNNER: A network-based algorithm for drug repurposing and its application to COVID-19», *PLOS Comput. Biol.*, vol. 17, fasc. 2, p. e1008686, feb. 2021, doi: 10.1371/journal.pcbi.1008686.

[4] A.-L. Barabási, N. Gulbahce, e J. Loscalzo, «Network medicine: a network-based approach to human disease», *Nat. Rev. Genet.*, vol. 12, fasc. 1, pp. 56–68, gen. 2011, doi: 10.1038/nrg2918.

[5] M. Caldera, P. Buphamalai, F. Müller, e J. Menche, «Interactome-based approaches to human disease», *Curr. Opin. Syst. Biol.*, vol. 3, pp. 88–94, giu. 2017, doi: 10.1016/j.coisb.2017.04.015.

[6] F. Cheng *et al.*, «Network-based approach to prediction and population-based validation of in silico drug repurposing», *Nat. Commun.*, vol. 9, fasc. 1, p. 2691, 12 2018, doi: 10.1038/s41467-018-05116-5.

[7] P. Sibilio *et al.*, «In silico drug repurposing in COVID-19: A network-based analysis», *Biomed. Pharmacother.*, vol. 142, p. 111954, ott. 2021, doi: 10.1016/j.biopha.2021.111954.

[8] G. Fiscon, F. Conte, S. Amadio, C. Volonté, e P. Paci, «Drug Repurposing: A Network-based Approach to Amyotrophic Lateral Sclerosis», *Neurotherapeutics*, mag. 2021, doi: 10.1007/s13311-021-01064-z.

[9] S. Amadio, F. Conte, G. Esposito, G. Fiscon, P. Paci, e C. Volonté, «Repurposing Histaminergic Drugs in Multiple Sclerosis», *Int. J. Mol. Sci.*, vol. 23, fasc. 11, Art. fasc. 11, gen. 2022, doi: 10.3390/ijms23116347.

[10] F. Conte, P. Sibilio, G. Fiscon, e P. Paci, «A Transcriptome- and Interactome-Based Analysis Identifies Repurposable Drugs for Human Breast Cancer Subtypes», *Symmetry*, vol. 14, fasc. 11, Art. fasc. 11, nov. 2022, doi: 10.3390/sym14112230.

[11] G. Fiscon, P. Sibilio, A. Funari, F. Conte, e P. Paci, «Identification of Potential Repurposable Drugs in Alzheimer's Disease Exploiting a Bioinformatics Analysis», *J. Pers. Med.*, vol. 12, fasc. 10, Art. fasc. 10, ott. 2022, doi: 10.3390/jpm12101731.

[12] W. Yu, M. Clyne, M. J. Khoury, e M. Gwinn, «Phenopedia and Genopedia: disease-centered and gene-centered views of the evolving knowledge of human genetic associations», *Bioinformatics*, vol. 26, fasc. 1, pp. 145–146, gen. 2010, doi: 10.1093/bioinformatics/btp618.

[13] J. Piñero *et al.*, «DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants», *Nucleic Acids Res.*, vol. 45, fasc. D1, pp. D833–D839, gen. 2017, doi: 10.1093/nar/gkw943.

[14] A. M. Grimaldi *et al.*, «The New Paradigm of Network Medicine to Analyze Breast Cancer Phenotypes», *Int. J. Mol. Sci.*, vol. 21, fasc. 18, Art. fasc. 18, gen. 2020, doi: 10.3390/ijms21186690.

[15] P. Paci e G. Fiscon, «SWIMmeR: an R-based software to unveiling crucial nodes in complex biological networks», *Bioinformatics*, vol. 38, fasc. 2, pp. 586–588, gen. 2022, doi: 10.1093/bioinformatics/btab657.

[16] D. S. Wishart *et al.*, «DrugBank 5.0: a major update to the DrugBank database for 2018», *Nucleic Acids Res.*, vol. 46, fasc. Database issue, p. D1074, gen. 2018, doi: 10.1093/nar/gkx1037.

[17] J. Lamb *et al.*, «The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease», *Science*, vol. 313, fasc. 5795, pp. 1929–1935, set. 2006, doi: 10.1126/science.1132939.

[18] Y. Zhou, Y. Hou, J. Shen, Y. Huang, W. Martin, e F. Cheng, «Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2», *Cell Discov.*, vol. 6, fasc. 1, pp. 1–18, mar. 2020, doi: 10.1038/s41421-020-0153-3.

[19] A. R. Sahin *et al.*, «2019 Novel Coronavirus (COVID-19) Outbreak: A Review of the Current Literature», *EURASIAN J. Med. Oncol.*, vol. 4, fasc. 1, pp. 1–7, 2020.

[20] C. Volonté, S. Apolloni, e S. Amadio, «The Histamine and Multiple Sclerosis Alliance: Pleiotropic Actions and Functional Validation», *Curr. Top. Behav. Neurosci.*, ago. 2021, doi: 10.1007/7854_2021_240.

[21] B. Chu *et al.*, «Preparation and evaluation of teniposide-loaded polymeric micelles for breast cancer therapy», *Int. J. Pharm.*, vol. 513, fasc. 1, pp. 118–129, nov. 2016, doi: 10.1016/j.ijpharm.2016.09.005.

[22] K.-C. Mok *et al.*, «Repurposing hyperpolarization-activated cyclic nucleotide-gated channels as a novel therapy for breast cancer», *Clin. Transl. Med.*, vol. 11, fasc. 11, p. e578, nov. 2021, doi: 10.1002/ctm2.578.

[23] V. Cınar, Z. Hamurcu, A. Guler, N. Nurdinov, e B. Ozpolat, «Serotonin 5-HT7 receptor is a biomarker poor prognostic factor and induces proliferation of triple-negative breast cancer cells through FOXM1», *Breast Cancer Tokyo Jpn.*, ago. 2022, doi: 10.1007/s12282-022-01391-9.

[24] E. E. Beaman *et al.*, «Blood–brain barrier permeable β-blockers linked to lower risk of Alzheimer's disease in hypertension», *Brain*, p. awac076, feb. 2022, doi: 10.1093/brain/awac076.

[25] K.-M. Han *et al.*, «Regorafenib Regulates AD Pathology, Neuroinflammation, and Dendritic Spinogenesis in Cells and a Mouse Model of AD», *Cells*, vol. 9, fasc. 7, p. 1655, lug. 2020, doi: 10.3390/cells9071655.

[26] D. Pan *et al.*, «Anti-platelet Therapy Is Associated With Lower Risk of Dementia in Patients With Cerebral Small Vessel Disease», *Front. Aging Neurosci.*, vol. 14, 2022.

[27] R. X. Santos, S. C. Correia, S. Cardoso, C. Carvalho, M. S. Santos, e P. I. Moreira, «Effects of rapamycin and TOR on aging and memory: implications for Alzheimer's disease», *J. Neurochem.*, vol. 117, fasc. 6, pp. 927–936, giu. 2011, doi: 10.1111/j.1471-4159.2011.07262.x.