Association Rule Mining and In silico Analysis of Ethiopian Traditional Medicine Prescription

Lemessa Etana Bultum

leme3388@gmail.com

Dong-A University

Gwangmin Kim

Korea Advanced Institute of Science and Technology (KAIST)

Seon-Woo Lee

Dong-A University

Doheon Lee

Korea Advanced Institute of Science and Technology (KAIST)

Research Article

Keywords: Natural product-based medicines, Drug discovery, Data mining, Network analysis, Thera-peutic targets

Posted Date: July 25th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-4652485/v1

Additional Declarations: No competing interests reported.

Abstract

Multicomponent traditional medicine prescriptions are commonly used for disease treatment in Ethiopia. However, the lack of consistency in these prescriptions across practitioners, cultures, and locations has impeded the development of reliable therapeutic medicines. Therefore, a systematic analysis of traditional medicine information is essential to identify consistent and dependable medicinal materials, fostering harmony in these practices. In this study, we compiled and analyzed a dataset comprising 505 prescriptions, encompassing 567 different medicinal materials used to treat 106 diseases. Utilizing association rule mining, we uncovered significant associations between human diseases and medicinal materials. Additionally, in silico analysis was conducted to provide support for these associations and validate traditional medicinal uses. The results of the study revealed significant associations between diseases and medicinal materials. Notably, wound healing showed a strong association with Rumex abyssinicus Jacq, prompting further molecular-level investigation to confirm this association. In silico analysis of the phytochemicals of Rumex abyssinicus Jacq identified 756 therapeutic targets enriched in various KEGG pathways and biological processes. Using the random-walk with restart algorithm in the CODA PPI network, we identified disease associated with these targets, including cancer, inflammation, congenital malformations, and diseases of metabolic, immune, respiratory, and neurological systems. Furthermore, numerous hub target genes in the PPI network were directly associated with the wound healing process, supporting the findings of the association rule mining and the traditional use of Rumex abyssinicus Jacq for treating wounds. In conclusion, this research revealed important associations between diseases and medicinal materials, emphasizing the therapeutic potential of Rumex abyssinicus Jacq. It provides a foundational understanding for further exploration of the efficacy of traditional and natural-products based medicines.

1. Introduction

Over time, numerous cultures worldwide have developed their own traditional healing systems. In recent years, there has been a significant increase in the popularity and acceptance of traditional, complementary, and alternative medicine (TCAM) in both developing and developed countries [1]. Traditional medicine (TM) is widely used and practiced in countries like China, Korea, India, and Africa [2]. In China, traditional herbal medicine accounts for 30–50% of the total medicinal consumption [3]. In African countries such as Ghana, Mali, Nigeria, and Zambia, herbal medicines are the primary treatment for approximately 60% of children with high fever caused by malaria [4]. This trend reflects the growing interest in natural product therapy and researches on a global scale. This led to an increase in the availability of academic articles and electronic databases focusing on traditional medicine. An example of such a resource is the Compound Combination-Oriented Natural Product Database with Unified Terminology (COCONUT), which consolidates 21,456 oriental medicine prescriptions from Korean, Chinese, and Japanese herbal medicine sources [5]. Additionally, the integrated Ethiopian traditional herbal medicine and phytochemicals database (ETM-DB) encompasses details on 573 multi-herb prescriptions derived from combinations of 265 herbs found in Ethiopian natural products [6]. The ETM-DB serves as a valuable resource to augment the knowledge and utilization of traditional medicines in Ethiopia, primarily focusing on their application in natural product-based drug discovery research. As traditional medicines continue to gain acceptance globally, the accumulation of knowledge in such databases becomes increasingly invaluable.

Ethiopia is a home to a rich variety of cultures, comprising over 80 ethnic groups, each with its own unique traditions and knowledge in the realm of traditional healthcare. The country's history is deeply intertwined with the use of traditional medicine, which encompasses a wide range of practices, including preventive, curative, and even surgical methods, showcasing the diverse and vibrant Ethiopian cultures [7]. According to the New Partnership for Africa's

Development, a significant portion of the Ethiopian population, estimated between 60% and 79%, still relies on indigenous traditional medicine for their healthcare needs [8]. This continuing reliance on traditional medicine can be attributed to various factors. As a developing nation, access to modern healthcare remains limited in many parts of the country. Moreover, the expenses associated with modern medical services make them unaffordable for many people. In contrast, traditional practices are more familiar and culturally accepted, leading to their continued usage. Beyond the cultural significance, Ethiopia's natural resource abundance contributes to the popularity of traditional medicine practices. The country is blessed with a remarkable diversity of flora, including 6,500 to 7,000 floral species, of which 10–12% are exclusive to Ethiopia. Remarkably, more than 1,000 plant species, accounting for 15% of the total flora, are utilized for medicinal purposes [9], [10], [11]. In this way, Ethiopia's heritage, cultural acceptance, and abundant natural resources all converge to sustain the utilization of traditional medicine, bridging the gap in healthcare access and meeting the needs of its people.

In Ethiopia, knowledge about the usage of traditional medicine is derived from both written and unwritten sources [11]. However, this valuable information is scattered across various conventional and modern references. The indigenous traditional medicine practitioners, who have inherited this wisdom from their ancestors, have preserved ancient records and oral traditions, forming the foundation of traditional literature. On the other hand, contemporary literature, mostly from academic and research institutions, delves into topics such as Ethiopian ethnobotany, ethnomedicine, phytochemical studies, and the biological evaluation of medicinal plants [6], [12]. These diverse sources of knowledge contribute to a comprehensive understanding of the rich traditions and practices of traditional medicine in Ethiopia.

The utilization of multicomponent prescriptions for disease treatment in Ethiopia exhibit variations based on the healer, cultural practices within society, and geographic regions. As a result, the data collected on traditional medicine usage leads to large datasets with overlapping and duplicated entries, presenting difficulties in establishing a consistent and reliable selection of therapeutic medicinal materials for prevalent diseases. To address this issue and promote harmonized practices, a systematic organization and analysis of prescription data are essential. In this study, we utilized data mining and network analysis techniques to analyze prescription data, aiming to identify significant associations between common diseases and medicinal materials. Through this effort, our goal is to establish a reliable and consistent collection of medicinal materials that practitioners can readily adopt. Furthermore, the results obtained from this analysis serve as a fundamental basis for conducting additional investigations into the therapeutic effects of these medicinal materials on the associated diseases.

In order to support the findings of the association and affirm the claimed traditional therapeutic uses of the associated medicinal materials, we conducted in silico analysis. Through this process, we identified molecular-level associations between diseases and their associated medicinal materials. For instance, one such association was found between wound healing and Rumex abyssinicus Jacq (Polygonaceae). Subsequently, we performed an in silico analysis of the phytochemicals within R. abyssinicus Jacq to uncover related therapeutic targets, associated pathways, biological processes, and diseases.

This study involved compiling traditional medicine prescription data from multiple sources and using data mining and in silico analysis techniques to reveal valuable insights concealed within traditional medicine documents. To the best of our knowledge, no prior comprehensive research has explored Ethiopian traditional medicine to substantiate traditional knowledge and conduct molecular-level in silico analysis of medicinal materials. The findings could provide valuable knowledge for traditional medicine practitioners and researchers in the field of natural product-based drug development.

2. Material and methods

2.1. Data source and data pre-processing

Like other traditional medicine systems such as oriental medicine, Ethiopian traditional medicines predominantly involve the use of multicomponent prescriptions. Ethiopian traditional medical practice recognizes the synergistic interactions and potentiating effects among the various medicinal materials used in a prescription from multiple sources [13]. While there are several modern literatures on the therapeutic use of individual medicinal materials, information on the therapeutic use of multicomponent prescriptions remains limited. The variations in the use of constituent medicinal materials among different traditional medicine practitioners, diverse cultural practices, and geographical locations contribute to this limitation. Additionally, the availability of specific medicinal materials for prescription is influenced by the geo-climatic conditions of a given area. These factors collectively determine whether particular medicinal materials are included in a multicomponent prescription or not.

For this analysis, prescription data was manually collected from various literature sources, including research articles, books, and theses. A significant portion of the data was sourced from the book Medicinal Plants and Enigmatic Health Practices of Northern Ethiopia [14], where information on northern Ethiopian traditional medicine prescriptions was compiled from diverse sources, including ancient manuscripts and microfilms from the British Library. Prescriptions containing single medicinal materials, veterinary medicines, and non-illness related enigmatic practices, such as remedies for evil eye, attack deterrent, thief deterrent, warding off sorcery stealing, anti-poltergism, intelligence booster, love potions, courage, hair dye, phobia, misfortune, etc., were excluded from the study.

In total, 505 prescriptions from some combinations of 567 medicinal materials for treating 106 human diseases were used in this study. To identify associations between medicinal materials and diseases from the prescription data, a binary matrix was constructed, where columns represented human diseases and medicinal materials, rows represented prescriptions, and each cell indicated the presence (1) or absence (0) of an item. Figure 1 depicts an overview of the method used to identify associations between medicinal materials and diseases from the prescription data.

2.2. Association rule mining

Association rule mining is a fundamental data mining concept extensively investigated to identify associations between itemsets in vast databases. Its objective is to identify sets of items that frequently occur together in transactional databases. The association rule mining algorithm was first introduced by Agrawal et al. [15], primarily applied to analyze basket data models, where each basket represents a transaction, and the items purchased within the basket are termed as itemsets. The association rule mining algorithm employed in this study is adopted from the algorithm introduced in Agrawal et al. [15]. In essence, this algorithm generates association rules satisfying user-defined minimum support and minimum confidence thresholds.

For an association rule X (antecedent) ⇒ Y (consequent), the support of the rule is calculated as the proportion of transactions in in the data that contain both antecedent and consequent, representing the frequency of occurrence of the rule.

$$\mathrm{Support}\ (\mathrm{X} \Rightarrow \mathrm{Y}) = \frac{\mathrm{Numberoftransactionscontainingboth} \mathrm{X} \wedge \mathrm{Y}}{\mathrm{Totalnumberoftransactions}}$$

Confidence is a measure of how often a rule is true in the data. It is the ratio of the number of transactions that contain both the antecedent and the consequent of the rule, to the number of transactions that contain only the antecedent [16]. A high confidence means that the rule is reliable and consistent. A confidence value of 0 indicates that none of the transactions with the antecedent itemset have the consequent itemset, while a confidence value of 1 signifies that all transactions with the antecedent itemset include the consequent itemset.

$$\mathrm{Confidence}\left(\mathrm{X} \Rightarrow \mathrm{Y}\right) = \frac{\mathrm{Numberoftransactionscontainingboth X} \wedge \mathrm{Y}}{\mathrm{Numberoftransactionscontaining X}}$$

Lift is a measure of how much more likely the consequent of a rule is to occur when the antecedent is present, compared to when it is absent. It is the ratio of the confidence of the rule, to the frequency of the consequent in the whole dataset [17]. For example, for the association rule X (antecedent) ⇒ Y (consequent), the lift is the confidence of the rule, divided by the proportion of transactions that have Y. A high lift value indicates a significant and interesting rule, highlighting a strong association between the antecedent and the consequent. The lift value of 1 means that antecedent and consequent itemsets are completely independent, which can be achieved by the pure random rules, and the lift value of 2 means that antecedent and consequent itemsets are two times more dependent than the pure random rules [17].

$$\mathrm{lift}\left(\mathrm{X} \Rightarrow \mathrm{Y}\right) = \frac{\left(\mathrm{Numberoftransactionscontainingboth X} \wedge \mathrm{Y}\right) / \left(\mathrm{Numberoftransactionscontaining Y}\right)}{\mathrm{Fractionoftransactionscontaining Y}}$$

In this study, association rule mining was employed to identify the associations between diseases and therapeutic medicinal materials using the arules package with the apriori algorithm in R [18]. The association rules took the form of 'antecedent ⇒ consequent', where the antecedent represented the disease and the consequent represented the medicinal material. Statistical parameters, including support, confidence, and lift, were used to assess the significance of the associations. In this analysis, we set minimum support and confidence thresholds of 0.3% and 10%, respectively. Any generated rule that met or exceeded these thresholds for both support and confidence was deemed statistically significant.

When configuring the association rule parameters for this specific analysis, the support of an item (either disease or medicinal material) is determined by the fraction of transactions (in this context, prescriptions) in which the item is present. The support value of an association rule is calculated as the ratio of the number of prescriptions containing both the disease and medicinal materials to the total number of prescriptions. The confidence value of an association rule is calculated as the ratio of prescriptions that include both the disease and the medicinal material to the prescriptions that include the disease alone. This value provides an indication of how often the rule has been observed to be true. The lift is a measure of how much more likely the medicinal material occurs when the related disease is present, compared to when the disease is absent [19]. It indicates the dependence between the disease and medicinal material. In this analysis, a high lift value and lower confidence suggest that the medicinal material is specifically used for treating that particular disease. On the other hand, a high confidence value and relatively low lift value indicate that the medicinal material is not specific to the disease, and it is utilized to treat various diseases in addition to the related disease.

2.3. Network analysis

A network, also known as a graph in mathematics, consists of nodes or vertices connected by edges or lines [20]. Networks find applications in various fields such as social networks, transportation systems, communication

networks, and bioinformatics. Several network measures like degree, centrality, and modularity are used to characterize and analyze networks [21]. The degree of a node is the number of connections that the node has. It is the most fundamental network measure and most other measures are ultimately linked to node degree [22]. In our analysis, we represented disease and medicinal materials as nodes in a bipartite network, where edges indicate that the medicinal material is part of the prescription used to treat the disease. By employing the Python NetworkX library version 2.5, we constructed, analyzed, and visualized the networks, allowing for a better understanding and visualization of the associations between the interacting items.

2.4 Investigating the pharmacological activity of R. abyssinicus Jacq phytochemicals through target identification, Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, and network analysis

After identifying the association between medicinal materials and human disease, we explored molecular level association between the medicinal materials and associated diseases. Specifically, we focused on R. abyssinicus Jacq, a medicinal material strongly associated with wound healing.

The process flow diagram, as shown in Fig. 2, illustrates the following steps: First, we compiled the phytochemical constituents of Rumex abyssinicus Jacq from various sources, including the ETM-DB [6] and other literature. Then, we identified functionally related therapeutic targets from the Compound Combination-Oriented Natural Product Database with Unified Terminology (COCONUT) database [5]. Next, to identify diseases associated with these targets, we utilized the Context-Oriented Directed Associations (CODA) network [23]. By constructing a heterogeneous network encompassing all types of biological interactions within CODA, we traced the converging information flow into disease nodes, starting from the targets, using the Random-Walk with Restart (RWR) algorithm. As a result, we ranked the disease nodes in the network based on their RWR scores.

To gain further insights into the therapeutic effects of R. abyssinicus Jacq phytochemicals, we analyzed the KEGG[24] pathways and GO terms[25] associated with the identified targets. For a comprehensive understanding of the complex interactions among compounds, targets, pathways, and biological processes, we conducted an enrichment test using the enrichr function in the gseapy Python library, utilizing the internal gene set named 'KEGG_2021_Human'. Furthermore, we conducted molecular docking and molecular dynamics simulation to gain insights into molecular interactions and the stability of the phytochemicals and their associated targets. We also conducted a comprehensive literature review to provide additional evidence for the inferred therapeutic effects of R. abyssinicus Jacq phytochemicals.

2.5. Protein–protein interaction (PPI) network, molecular docking and molecular dynamics (MD) simulations

We constructed a PPI network from CODA, comprising a total of 17,358 protein nodes. To focus on the therapeutic targets related to R. abyssinicus Jacq phytochemicals, we selected proteins with relationships involving gene products. The nodes in the network were converted into gene symbols for better understanding. From this PPI network, we extracted a hub network containing the therapeutic targets associated with R. abyssinicus Jacq phytochemicals. To identify the hub targets, we used two network measures, namely degree centrality and eigenvector centrality, and defined hub targets as nodes within the top 1% of the PPI network in terms of both degree centrality and eigenvector centrality.

We performed structure-based molecular docking to assess the binding interactions between the phytochemicals and their associated targets. For this purpose, we retrieved the PDB files of five selected hub proteins from the CODA PPI network. Additionally, the ligands' SDF files were obtained from the PubChem database [26]. The protein-

ligand blind molecular docking was conducted using the CB-Dock web server (http://clab.labshare.cn/cb-dock/php/index.php). CB-Dock incorporates cavity detection, docking, and homologous template matching. To perform the structure-based molecular docking at the detected candidate pockets, CB-Dock utilizes AutoDock Vina [27]. Finally, we visualized the 3D and 2D poses of the protein-ligand complex using Discovery Studio Visualizer.

In this research, we also conducted molecular dynamics (MD) simulation to evaluate the stable binding affinity between phytochemicals and their target protein. Emodin, which displayed a relatively low binding energy score of -8.9 kcal/mol with the target protein ESR1 in docking studies, was chosen for the simulation. Hence, we performed a 100 nanosecond (ns) MD simulation using the GROningen MAchine for Chemical Simulations (GROMACS) 2021.4 software package [28]. For the preparation of protein topology, the official SwissParam (https://www.swissparam.ch/)[29] was used, and for the preparation of ligand topology, the CHARMM27 all-atom forcefield [30]. A triclinic box of the default Simple Point Charge (SPC) water model was utilized as the solvation method, and explicit solvent periodic boundary conditions were used. For the solvated complexes, charge neutralization was done using sodium and chloride ions. By using the steepest descent approach over 5000 steps, the systems were subjected to energy minimization in order to remove any steric clashes and bad contacts. A series of equilibration was performed to relax the system and bring it to an appropriate starting state. This involved utilizing a Berendsen thermostat to equilibrate the system under constant number of particles, volume, and temperature (NVT) for 100 picoseconds (ps) [31]. Additionally, re-equilibration was carried out using the Parrinello-Rahman barostat with a time step of 2 femtoseconds (fs) for each equilibration round for a further 100 ps at constant number of particles, pressure and temperature (NPT) [32]. The MD production phase was then performed for 100 ns with a time step of 2 fs, a constant temperature of 300 K, and a constant pressure of 1 atm. Using GROMACS' "trjconv" function, the complex was re-centered and re-wrapped into the unit cells for MD trajectories investigation. The RMSD of the protein, ligand, and protein-ligand complex in relation to its initial position and the radius of gyration (Rg) was then used to analyze the trajectories. Data charting was done using the GRaphing, Advanced Computation and Exploration of data (GRACE) tool (https://plasma-gate.weizmann.ac.il/Grace/).

3. Results and discussion

3.1. Associations between disease and medicinal materials

In this study, we employed the apriori algorithm of the association rule mining to identify significant associations between diseases and medicinal materials in Ethiopian traditional medicine prescriptions. Using statistical metrics such as support, confidence, and lift, we assessed the strength of these associations. Table 1 displays the support, confidence, and lift values of the association rules between common diseases (antecedent) and their associated medicinal materials (consequent) with a confidence value of at least 10%.

The strength of association rules can be assessed using lift and confidence values. Strong associations exist between the diseases and medicinal materials with relatively high confidence and lift values. For example, the rule wound ⇒ Rumex abyssinicus Jacq has high confidence (17.95) and lift (5.33), showing that Rumex abyssinicus Jacq is the most frequently used herbal material for treating wounds. In a study conducted by [33], they evaluated the wound healing and anti-inflammatory properties of Rumex abyssinicus Jacq by utilizing its hydroalcoholic extract on mouse wound models. The results showed significant improvements in wound contraction, epithelization time, tissue breaking strength, and hydroxyproline content, supporting the traditional claims associated with this medicinal material. Hence, association rules can aid in pinpointing essential medicinal materials, while additional experimental studies can corroborate their traditional applications.

The rules Mental illness ⇒ Clerodendrum myricoides (Hochst.) R. Br. ex Vatke and Mental illness ⇒ Ruta chalepensis L have higher confidence values while their lift value is relatively low. This shows that Clerodendrum myricoides (Hochst.) R. Br. ex Vatke and Ruta chalepensis L are frequently used to treat mental illness and other illnesses as well. But, the rule Mental illness ⇒ Trigonella foenum-graecum L have relatively high lift value (7.01) and low confidence value (11.11) showing that Trigonella foenum-graecum L is specifically used to treat mental illness.

The association rules Sexual stimulation (aphrodisiacs) ⇒ Tragia pungens (Forssk.) Muell. Arg, sexual stimulation (aphrodisiacs) ⇒ Asparagus Africanus Lam, and sexual stimulation (aphrodisiacs) ⇒ Ferula communis L have high confidence, implying that these medicinal materials are commonly used as aphrodisiacs. Interestingly, the rules, sexual stimulation (aphrodisiacs) ⇒ Tragia pungens (Forssk.) Muell. Arg, and, sexual stimulation (aphrodisiacs) ⇒ Asparagus Africanus Lam have higher confidence and lift values indicating that Tragia pungens (Forssk.) Muell. Arg. and Asparagus Africanus Lam are specific herbal materials used as aphrodisiacs. It is also interesting to note that the rule sexual stimulation (aphrodisiacs) ⇒ Habenaria sp. have higher lift value (28.06) and lower confidence value (11.11) implying that Habenaria sp. is a medicinal material specifically used as aphrodisiacs.

For rabies infection, the rule Rabies ⇒ Cucumis ficifolius A. Rich have high confidence (43.75) and relatively low lift values (4.80). This implies that Cucumis ficifolius A. Rich is frequently used to treat rabies infection as well as other diseases such as frequent miscarriage, sexual dysfunction, eczema, and syphilis as shown in Table 1. Hence, Cucumis ficifolius A. Rich can be considered as a generally used medicinal material. Conversely, the rule Rabies ⇒ Sida cuneifolia Roxb. or S. ovata Forssk, despite its relatively low confidence (12.50), has high lift value (10.52) indicating that this herb is specifically used for treating rabies infection.

The lift value (10.52) for the rule Frequent miscarriage ⇒ Periploca linearifolia A. Rich. & Quart.-Dill is relatively higher indicating its specific traditional therapeutic use to prevent frequent miscarriage. Similarly, the rule Menorrhagia ⇒ Protea gaguedi Gmel has exceptionally higher lift value (36.07) indicating its very specific therapeutic use for treating menorrhagia. It may be worthwhile to further investigate the effect of treating menorrhagia with Protea gaguedi Gmel.

Inspection of the rules for leprosy also reveals some interesting patterns. For example, the rules Leprosy ⇒ Plumbago zeylanica L, Leprosy ⇒ Withania somenifera (L.) Dunal, Leprosy ⇒ Euclea schimperi (DC.) Dandy, and Leprosy ⇒ Capparis tomenitosa Lam have high confidence and low lift values indicating that these medicinal materials are commonly used to treat not only leprosy, but also other diseases. On the other hand, the rules Leprosy ⇒ Sylvicapra grimmia, Leprosy ⇒ Maesa lanceolata Forssk, and Leprosy ⇒ Ranunculus multifidus Forssk have higher lift and relatively lower confidence values indicating that they are mostly used to treat leprosy. In the same way, important relationships can be inferred for other diseases in Table 1. The results were visualized by the radar charts shown in Fig. 3. The individual charts for confidence (Fig. 3(a)) and lift (Fig. 3(b)) shows which of the medicinal materials have high values of confidence and lift for leprosy. Fig. 3(c) shows the values of confidence and lift simultaneously by using their standardized values. It can be seen that Sylvicapra grimmia, Maesa lanceolate Forssk, and Ranunculus multifidus Forssk have larger lift values compared to their confidence values; the opposite is true for Plumbago zeylanica L, Withania somenifera (L.), Dunal Euclea schimperi (DC.) Dandy, and Capparis tomenitosa Lam. To confirm the associations between diseases and medicinal materials inferred in this study, in vitro and in vivo investigations of the effects of the medicinal materials and/or their extracts on the corresponding pathogens and diseases should be carried out.

Table 1 Support, confidence, and lift values between diseases and associated medicinal materials.

Disease	Medicinal material	Support (%)	Confidence (%)	Lift
Wound	Rumex abyssinicus Jacq.	1.39	17.95	5.33
	Conyza steudelii Sch.-Bip. ex A. Rich.	1.19	15.38	5.55
	Plumbago zeylanica L.	0.99	12.82	2.02
	Calpurnia aurea (Ait.) Benth.	0.79	10.26	2.35
Mental illness	Clerodendrum myricoides (Hochst.) R. Br. ex Vatke	1.58	29.63	3.48
	Ruta chalepensis L.	1.19	22.22	3.87
	Clausena anisata (Willd.) Benth.	0.79	14.81	4.68
	Trigonella foenum-graecum L.	0.59	11.11	7.01
	Cannabis sativa L.	0.59	11.11	5.10
	Withania somenifera (L.) Dunal	0.59	11.11	2.81
	Gladiolus psittacinus Hook.	0.59	11.11	3.74
	Verbena officinalis L.	0.59	11.11	2.95
	Lepidium sativum L.	0.59	11.11	2.95
	Asparagus africanus Lam.	0.59	11.11	2.44
	Securidaca longepedunculata Fresen.	0.59	11.11	2.44
	Capparis tomenitosa Lam.	0.59	11.11	2.67
	Adhatoda schimperiana (Hochst) Nees.	0.59	11.11	2.34
	Myrrh (Commiphora cf. crenulata or other C. spp.)	0.59	11.11	2.08
	Cucumis ficifolius A. Rich.	0.59	11.11	1.22
Sexual stimulation (aphrodisiacs)	Tragia pungens (Forssk.) Muell. Arg.	1.19	33.33	28.06
	Asparagus Africanus Lam.	0.99	27.78	23.38
	Ferula communis L.	0.99	27.78	9.35
	Thalictrum rhynchocarpum Dill. & A. Rich.	0.79	22.22	4.32
	Phoenix reclinata Jacq.	0.59	16.67	7.65
	Sida cuneifolia Roxb.	0.59	16.67	7.01
	Olea europaea L. subsp. africana (Mill.) P.S. Green	0.59	16.67	6.47
	Gladiolus psittacinus Hook.	0.59	16.67	5.61
	Cucumis ficifolius A. Rich.	0.59	16.67	1.83
	Habenaria sp.	0.40	11.11	28.06

	Sida cuneifolia Roxb. or S. ovata Forssk.	0.40	11.11	9.35
	Tragia pungens (Forssk) Muell. Arg.	0.40	11.11	8.02
	Catha edulis (Vahl) Forssk. ex Endl.	0.40	11.11	9.35
	Stylochiton kerensis N.E. Brown	0.40	11.11	6.23
	Cannabis sativa L.	0.40	11.11	5.10
	Zehneria scabra (L.f.) Sonder	0.40	11.11	2.67
	Clerodendrum myricoides (Hochst.) R. Br. ex Vatke	0.40	11.11	1.30
Eye disease	Hagenia abyssinica (Bruce) J.F. Gmel.	0.59	18.75	7.28
	Premna schimperi Engl.	0.40	12.50	9.02
	Artemisia afra Jacq. ex Willd.	0.40	12.50	7.89
	Vitis vinifera L.	0.40	12.50	5.74
	Solanum marginatum L.f.	0.40	12.50	5.26
Rabies	Cucumis ficifolius A. Rich.	1.39	43.75	4.80
	Phytolacca dodecandra L'Herit.	0.99	31.25	10.52
	Cyphostemma junceum (Webb) Descoings ex Wild & Drummond	0.79	25.00	10.52
	Solanum cf. adoens Hochst.	0.79	25.00	4.86
	Adhatoda schimperiana (Hochst) Nees.	0.59	18.75	3.95
	Sida cuneifolia Roxb. or S. ovata Forssk.	0.40	12.50	10.52
	Senecio myriocephalus Sch. Bip. ex A.Rich.	0.40	12.50	7.01
	Vitis vinifera L.	0.40	12.50	5.74
	Kalanchoe petitiana A. Rich.	0.40	12.50	5.74
	Mandragora officinarum/Luffa aegyptica	0.40	12.50	4.86
Frequent miscarriage	Osyris quadripartita Decn.	0.59	18.75	5.57
	Clutia abyssinica Jaub. & Spach.	0.59	18.75	6.31
	Solanum cf. adoens Hochst.	0.59	18.75	3.64
	Cucumis ficifolius A. Rich.	0.59	18.75	2.06
	Periploca linearifolia A. Rich. & Quart.-Dill.	0.40	12.50	10.52
	Aloe sp.	0.40	12.50	6.31
	Rumex nervosus Vahl	0.40	12.50	4.86
Leprosy	Plumbago zeylanica L.	0.99	33.33	5.26
	Withania somenifera (L.) Dunal	0.59	20.00	5.05

	Euclea schimperi (DC.) Dandy	0.59	20.00	5.61
	Capparis tomenitosa Lam.	0.59	20.00	4.81
	Sylvicapra grimmia	0.40	13.33	16.83
	Maesa lanceolata Forssk.	0.40	13.33	22.44
	Ranunculus multifidus Forssk.	0.40	13.33	22.44
	Olea europaea L. subsp. africana (Mill.) P.S. Green	0.40	13.33	5.18
	Osyris quadripartita Decn.	0.40	13.33	3.96
	Clematis simensis Fresen.	0.40	13.33	5.18
	Brucea antidysenterica J.F. Mill.	0.40	13.33	3.96
	Rumex abyssinicus Jacq	0.40	13.33	3.54
	Jasminum floribundum R. Br. ex Fresen.	0.40	13.33	3.06
	Acokanthera schimperi (DC.) Oliv.	0.40	13.33	3.21
	Cucumis ficifolius A. Rich.	0.40	13.33	1.46
Eczema	Hagenia abyssinica (Bruce) J.F. Gmel.	0.59	21.43	8.32
	Cucumis ficifolius A. Rich.	0.59	21.43	2.35
	Lichen	0.40	14.29	14.43
	Gossypium barbadens L.	0.40	14.29	12.02
	Mandragora officinarum/Luffa aegyptica	0.40	14.29	5.55
	Zehneria scabra (L.f.) Sonder	0.40	14.29	3.44
Menorrhagia	Brassica nigra L.	0.59	21.43	12.02
	Protea gaguedi Gmel.	0.40	14.29	36.07
	Cordia africana lam.	0.40	14.29	10.31
	Nigella sativa L.	0.40	14.29	6.56
	Achyranthes aspera L.	0.40	14.29	3.61
	Asparagus africanus Lam.	0.40	14.29	3.14
Malaria	Dodonaea viscosa (L.) Jack.	0.59	25.00	15.78
	Ekebergia capensis Sparm.	0.40	16.67	12.02
	Kanahia laniflora (Forssk.) R. Br.	0.40	16.67	7.65
	Jasminum floribundum R. Br. ex Fresen.	0.40	16.67	3.83
	Solanum cf. adoens Hochst.	0.40	16.67	3.24
	Clerodendrum myricoides (Hochst.) R. Br. ex Vatke	0.40	16.67	1.96

Epidemic	Ekebergia capensis Sparm.	0.59	25.00	18.04
	Coccinia abyssinica / Cucurbita pepo	0.40	16.67	9.35
	Stephania abyssinica (Dill. & A. Rich.) Walp.	0.40	16.67	7.65
	Brucea antidysenterica J.F. Mill.	0.40	16.67	4.95
	Clausena anisata (Willd.) Benth.	0.40	16.67	5.26
	Thalictrum rhynchocarpum Dill. & A. Rich.	0.40	16.67	3.24
	Croton macrostachyus Del.	0.40	16.67	2.48
Rectal prolapse	Brucea antidysenterica J.F. Mill.	0.59	27.27	8.10
	Croton macrostachyus Del.	0.59	27.27	4.05
	Af goma	0.40	18.18	18.36
	Euclea schimperi (DC.) Dandy	0.40	18.18	5.10
	Myrrh (Commiphora cf. crenulata or other C. spp.)	0.40	18.18	3.40
	Clerodendrum myricoides (Hochst.) R. Br. ex Vatke	0.40	18.18	2.14
Syphilis	Conyza steudelii Sch.-Bip. ex A. Rich.	0.79	36.36	12.24
	Cucumis ficifolius A. Rich.	0.59	27.27	2.99
	Cussonia ostinii Chiov.	0.40	18.18	22.95
	Vitis vinifera L.	0.40	18.18	8.35
	Phytolacca dodecandra L'Herit.	0.40	18.18	6.12
Rheumatic pain	Clerodendrum myricoides (Hochst.) R. Br. ex Vatke	0.59	27.27	3.20
	Tragia pungens (Forssk) Muell. Arg.	0.40	18.18	13.12
	Otostegia integrifolia Benth.	0.40	18.18	8.35
	Osyris quadripartita Decn.	0.40	18.18	5.40
	Achyranthes aspera L.	0.40	18.18	4.59
	Jasminum floribundum R. Br. ex Fresen.	0.40	18.18	4.17
Sorcery poisoning (stomach infection)	Verbena officinalis L.	0.79	36.36	9.67
	Achyranthes aspera L.	0.59	27.27	6.89
	Sida cuneifolia Roxb.	0.40	18.18	7.65
	Jasminum floribundum R. Br. ex Fresen.	0.40	18.18	4.17
	Asparagus africanus Lam.	0.40	18.18	3.99
	Thalictrum rhynchocarpum Dill. & A. Rich.	0.40	18.18	3.53

Migraine	Salvia nilotica Juss. ex Jacq.	0.79	44.44	22.44
	Echinops sp.	0.40	22.22	10.20
	Clerodendrum myricoides (Hochst.) R. Br. ex Vatke	0.40	22.22	2.61
Epilepsy	Myrrh (Commiphora cf. crenulata or other C. spp.)	0.59	37.50	7.01
	Brassica oleracea L. cf. var. capitata	0.40	25.00	63.13
	Adhatoda schimperiana (Hochst) Nees.	0.40	25.00	5.26
Hemorrhoid	Clematis simensis Fresen.	0.79	50.00	19.42
	Euphorbia abyssinica Gmel.	0.40	25.00	10.52
	Rumex steudelii Hochst. ex A. Br.	0.40	25.00	9.71
Coughs	Rubia cordifolia L.	0.40	25.00	10.52
	Calpurnia aurea (Ait.) Benth.	0.40	25.00	5.74
Ascaris	Rumex nervosus Vahl	0.40	33.33	12.95
	Rumex abyssinicus Jacq	0.40	33.33	8.86
	Croton macrostachyus Del.	0.40	33.33	4.95
Deafness	Millettia ferruginea (Hochst.) Bak.	0.006	0.59	0.50
	Myrica salicifolia Hochst. ex A. Rich.	0.004	0.40	0.33
Hematuria	Artemisia afra Jacq. ex Willd.	0.004	0.40	1.00
	Anethum graveolens L.	0.004	0.40	1.00
Hemorrhoid	Clematis simensis Fresen.	0.008	0.79	0.50
	Euphorbia abyssinica Gmel.	0.004	0.40	0.25
	Rumex steudelii Hochst. ex A. Br.	0.004	0.40	0.25
Headache	Gladiolus psittacinus Hook.	0.004	0.40	0.40
	Myrrh (Commiphora cf. crenulata or other C. spp.)	0.004	0.40	0.40
Infertility	Securidaca longepedunculata Fresen.	0.004	0.40	0.29
	Verbascum sinaiticum Benth.	0.004	0.40	0.29
Measles	Clausena anisata (Willd.) Benth.	0.004	0.40	0.50
	Clerodendrum myricoides (Hochst.) R. Br. ex Vatke	0.004	0.40	0.50
Oxytocic	Rumex steudelii Hochst. ex A. Br.	0.004	0.40	0.25
	Achyranthes aspera L.	0.004	0.40	0.25
Relapsing fever	Dovyalis abyssinica (A. Rich.) Warburg	0.004	0.40	0.40

	Urtica simensis Hochst. ex Steud.	0.004	0.40	0.40
	Coccinia abyssinica / Cucurbita pepo	0.004	0.40	0.40
	Clausena anisata (Willd.) Benth.	0.004	0.40	0.40
	Zehneria scabra (L.f.) Sonder	0.004	0.40	0.40
Scabies	Brassica nigra L. sensu lato	0.004	0.40	0.50
	Euclea schimperi (DC.) Dandy	0.004	0.40	0.50
	Securidaca longepedunculata Fresen.	0.004	0.40	0.50
	Myrrh (Commiphora cf. crenulata or other C. spp.)	0.004	0.40	0.50
	Plumbago zeylanica L.	0.004	0.40	0.50
Snake bite	Carissa edulis (Forssk.) Vahl	0.006	0.59	0.60
	Plumbago zeylanica L.	0.006	0.59	0.60
	Cardiospermum halicacabum L.	0.004	0.40	0.40
	Calpurnia aurea (Ait.) Benth.	0.004	0.40	0.40
Vitiligo	Osyris quadripartita Decn.	0.006	0.59	0.50
	Hagenia abyssinica (Bruce) J.F. Gmel.	0.004	0.40	0.33

3.2. Network analysis of Ethiopian traditional medicine prescription

Fig. 4 illustrates a network diagram displaying the associations between 20 sets of diseases and medicinal materials. For instance, Cucumis ficifolius A. Rich is represented by a large green circle, indicating its extensive usage in treating various diseases. Similarly, the broad light blue square designating aphrodisiacs suggests the utilization of multiple therapeutic materials for this purpose. Notably, the network diagram also shows that leprosy and mental illness are linked to a significant number of medicinal materials, indicating their diverse treatment approaches

The network graph illustrates the associations between a set of 20 diseases and the medicinal materials. In the graph, diseases are represented by light blue squares, while medicinal materials are denoted by green circles. The size of the nodes indicates their relative degree, and the edges between nodes signify that the prescriptions for treating specific conditions include the respective therapeutic materials. The width of the edges reflects the strength of the association, measured by the lift, between the medicinal materials and the diseases.

The network graph illustrates the associations between a set of 20 diseases and the medicinal materials. In the graph, diseases are represented by light blue squares, while medicinal materials are denoted by green circles. The size of the nodes indicates their relative degree, and the edges between nodes signify that the prescriptions for

treating specific conditions include the respective therapeutic materials. The width of the edges reflects the strength of the association, measured by the lift, between the medicinal materials and the diseases.

3.3. Investigating the pharmacological activity of R. abyssinicus Jacq phytochemicals through target identification, GO, KEGG pathways, and network analysis

In this study, we successfully mapped 10 phytochemicals from R. abyssinicus Jacq to the Compound Combination-Oriented Natural Product Database with Unified Terminology (COCONUT) database. These phytochemicals included emodin, chrysophanol, oleanolic acid, physcion, helminthosporin, lupeol, citreorosein, chrysophanein, stigmastane-3,6-dione, and stigmastane-3,6-dione [5]. The COCONUT database is a comprehensive resource containing information on herbs, compounds, targets, and phenotypes from traditional herbal medicines, functional foods, and conventional drugs [34]. Through the COCONUT database, we identified 756 human targets related to these phytochemicals.

The study revealed that the phytochemicals from R. abyssinicus Jacq targeted 756 human target proteins which are significantly enriched in 196 KEGG pathways (adjusted P-value ≤ 0.05) (Supplementary Table S1 online). Among these pathways, Fig. 5 (right) highlights the top 20 enriched KEGG pathways, encompassing 208 target proteins. To better understand the relationships, a network of target genes and pathways associated with the phytochemicals was constructed using the Python NetworkX package version 2.5, as depicted in Fig. 5 (left). Several KEGG pathways, including lipid and atherosclerosis, chemical carcinogenesis, pathways in cancer, kaposi sarcoma-associated herpesvirus infection, AGE-RAGE signaling pathway in diabetic complications, MAPK signaling pathway, human cytomegalovirus infection, C-type lectin receptor signaling pathway, and others, were related to most of the R. abyssinicus Jacq phytochemicals. This KEGG enrichment analysis provided insights into the targets and pathways through which these phytochemicals could potentially exert their therapeutic benefits [35]. However, further research is required to discover the specific mode of action of R. abyssinicus Jacq phytochemicals on the underlying molecular targets.

Gene Ontology (GO) analyses were performed to identify the biological processes associated with the target genes [36]. The 756 human targets of the phytochemicals from R. abyssinicus Jacq were significantly enriched in 918 biological processes (P-values ≤ 0.05) (Supplementary File S2 online). The functional enrichment analysis of the top 20 biological processes revealed, as shown in Fig. 6 (right), that the majority of the related targets were linked to proteolysis, cellular protein modification processes, dephosphorylation, estrogen metabolic processes, cortical cytoskeleton organization, ras protein signal transduction, and other processes. Furthermore, a phytochemical-target-biological process network was constructed using the NetworkX package in Python version 2.5, showcasing that 361 genes predominantly enriched the top 20 biological processes (Fig. 6 (left)).

In addition, a network analysis using CODA network was conducted to identify diseases related to the targets. Through this analysis, phytochemicals from R. abyssinicus Jacq have been linked to various diseases, including a number of cancer types, metabolic and endocrine disorders (like diabetes mellitus and hypertension), inflammation, immune system diseases, congenital malformations, respiratory diseases, and nervous system diseases. Most of these pharmacological activities align with the traditionally reported therapeutic uses of R. abyssinicus Jacq. For instance, the anti-Alzheimer's effect of R. abyssinicus Jacq is attributed to its isolated secondary metabolite, helmintosporin, which acts as a dual inhibitor of the enzymes AChE and BChE [37]. Additionally, crude extracts of R. abyssinicus possess antibacterial [38, 39], anticancer [39], antiviral [38], anti-inflammatory [38, 33], antioxidant [40], wound

healing [33], antimalarial [41], diuretic, and analgesic [42] activities. The diseases associated with the phytochemicals identified through in silico and network analysis can be found in Supplementary Table S2 online.

In order to support the association rule mining results and the reported traditional uses of medicinal materials, we thoroughly investigated an in silico analysis to identify any associations between the medicinal materials and the associated diseases revealed in this analysis. Notably, a strong association was found between wound healing and R. abyssinicus Jacq. The outcomes of the in silico analysis indicate that most of the phytochemicals found in R. abyssinicus Jacq are linked to various disease categories, including inflammatory diseases, immune system diseases, and infectious diseases, all of which are directly or indirectly associated with wound healing (Fig. 7). Inflammation, marked by characteristic signs like heat, redness, pain, swelling, increased body temperature, and fever, constitutes the initial stage of the wound healing process [43]. Infection is also a significant factor that can impact wound healing negatively [44]. Hence, this finding provides further support for the traditional use of R. abyssinicus Jacq in treating wounds, which aligns with the results obtained from the association rule mining.

3.4. PPI network, molecular docking and molecular dynamics simulations

From a Context-Oriented Directed Associations (CODA) protein-protein interaction (PPI) network (Fig. 8a) [23], we extracted hub target genes associated with phytochemicals in R. abyssinicus Jacq. These hub genes' distribution of degree and eigenvector centralities was then examined within the CODA PPI network, which comprises a total of 17,358 protein nodes. The degree centrality values for the top 1%, 5%, and 10% of the nodes were equal to or greater than 243, 101, and 67, respectively. Similarly, the eigenvector centrality values for the top 1%, 5%, and 10% of nodes were equal to or greater than 0.0313, 0.0138, and 0.008, respectively. Out of the R. abyssinicus Jacq phytochemical-associated targets, a total of 22 targets were discovered to be present within the top 1% for both degree and eigenvector centralities (Fig. 9). These targets include APP, EGFR, CUL3, MCM2, TP53, COPS5, FN1, ESR1, CDK2, HDAC1, BRCA1, EED, CSNK2A1, AKT1, MAPK1, SRC, YWHAE, RELA, HSPB1, CTNNB1, MAPK14, and PCNA (Fig. 8b).

To further substantiate the association rule mining results and the reported traditional uses of medicinal materials, we explored the relationship between the phytochemical associated hub genes (identified via PPI network analysis), and the associated disease (wound, as identified by association rule mining). Interestingly, several of the identified hub genes were found to be linked to the wound healing process. Studies have reported the expression of beta-amyloid precursor protein (APP) and its homologue, amyloid precursor-like protein 2 (APLP2), during skin wound repair using a mice full-thickness skin excision wound healing model [45]. The EGFR signaling pathway also plays a crucial role in maintaining normal skin integrity and promoting wound healing [46]. Moreover, Estrogen Receptor 1 (ESR1) is enriched in patients with diabetic wounds (DWs) and has been shown to regulate human skin fibroblasts (HSFs) [46].

In this study, structure-based molecular docking was used to investigate the interaction between R. abyssinicus Jacq phytochemicals and the identified targets. The results showed successful binding with low binding energies, confirming the interactions between the phytochemicals and the drug targets. Fig. 10 presents 3D and 2D views of the interactions of selected phytochemicals with selected hub genes/proteins in the PPI network, while Table 2 provides details on binding affinity energies, interacting amino acid residues, docking center, and docking size for the selected phytochemicals and receptors.

Table 2. Binding affinity energies, interacting amino acid residues, docking center and docking size between selected R. abyssinicus Jacq phytochemicals and receptors.

Phytochemical	Target (PDB ID)	Docking score (Kcal/mol)	Interacting amino acid residues	Docking center (x, y, z)	Docking size (x, y, z)
Emodin	ESR1 (3OS8)	-8.9	Chain D: ET343, LEU346, THR347, LEU349, ALA350, GLU353, LEU384, LEU387, MET388, GLY390, LEU391, ARG394, PHE404, LEU428, GLY521, MET522, HIS524, LEU525, VAL533	16, 17, -61	35, 32, 31
Physcione	ESR1 (3OS8)	-8.2	Chain D: MET343, LEU346, THR347, LEU349, ALA350, GLU353, TRP383, LEU384, LEU387, MET388, LEU391, ARG394, PHE404, MET421, ILE424, LEU428, GLY521, MET522, HIS524, LEU525, VAL533	27, 6, -62	28, 20, 20
Emodin	EGFR (3G5Z)	-7.4	Chain A: PRO120, SER121, SER122, GLU123, LEU125, Chain B: TYR125, RO126, LEU127, ALA128, LYS212, LEU213, GLU214, PRO215, ARG216, LA217	11, 39, 50	19, 25, 19
Oleanolic Acid	CUL3 (6I2M)	-7.2	Chain B: LYS274, VAL277, GLU278, MET279, GLU280, ASN281, SER282, GLY283, VAL285, HIS286, LYS289, GLU315, CYS316, SER319	-73, 7, -5	23, 23, 23
Oleanolic Acid	APP (5Z6E)	-6.2	Chain A: LYS51, ILE53, TRP55, GLU60, TYR64, GLU66, ARG84, VAL86, ASP87, ASP89, THR90, ARG91	6, 35, 2	23, 23, 23

In this study, molecular dynamics (MD) simulation was performed to assess the stable binding affinity of phytochemicals with their target protein. Emodin, a ligand with a reasonably low binding energy score (-8.9 kcal/mol) with the target protein ESR1, was selected for the simulation based on docking studies. The root-mean-square deviation (RMSD) parameter was utilized to examine the stability of the protein-ligand complex over 100 ns. The RMSD of the atoms in the ligand, ligand fit to protein backbone, and protein backbone exhibited similar patterns of fluctuation during most of the simulation period. The RMSD of ligand fit to protein backbone and protein backbone displayed initial increase followed by stability throughout the simulation. Notably, the RMSD of the ligand (Fig. 11, black) was slightly lower than that of the ligand fit to protein backbone (Fig. 11, green), indicating a conformational change of the target protein in the protein-ligand complex upon binding. The protein-ligand complex (Fig. 11, green) remained relatively stable throughout the simulation, indicating tight binding between the ligand and the receptor.

Additionally, the radius of gyration (Rg) analysis confirmed that the protein structures were largely unaffected by the binding region over the simulation period (Fig. 12). These findings suggest the stable interaction between the phytochemical and the target protein, supporting the potential therapeutic significance of the compound.

4. Conclusions

In this study, our primary objective was to uncover significant associations between human diseases and medicinal materials utilized in Ethiopian traditional medicine prescriptions, ultimately aiming to identify reliable therapeutic

medicines for common human diseases. Through association rule mining, we successfully identified notable associations, including a significant correlation between R. abyssinicus Jacq and wound healing.

Moreover, employing in-depth in silico analysis, we explored into molecular-level associations between phytochemicals present in R. abyssinicus Jacq and wound healing. This comprehensive analysis unveiled promising therapeutic potentials, with 756 associated targets enriched across various KEGG pathways and GO terms, indicating therapeutic potential against a broad spectrum of diseases including cancer, metabolic diseases, endocrine diseases, inflammation, immune system diseases, congenital malformations, respiratory diseases, and neurological diseases.

Additionally, network analysis utilizing CODA network pinpointed hub target genes closely linked with the phytochemicals, particularly accentuating their role in the wound healing process. Molecular docking and molecular dynamics simulations further reinforced the stable binding interactions between the phytochemicals and hub target genes.

In summary, our in silico and network analyses provide robust validation for the findings from association rule mining and the traditional therapeutic applications of R. abyssinicus Jacq in wound treatment. Nonetheless, to substantiate these outcomes, controlled experimental investigations on the pharmacological effects of R. abyssinicus Jacq phytochemicals are imperative.

Abbreviations

TCAM: Traditional, Complementary and Alternative Medicine

TM: Traditional Medicine

COCONUT: Compound Combination-Oriented Natural Product Database with Unified Terminology

ETM-DB: Integrated Ethiopian Traditional Herbal Medicine and Phytochemicals Database

GO: Gene Ontology

CODA: Context-Oriented Directed Associations

KEGG: Kyoto Encyclopedia of Genes and Genomes

PPI: Protein-Protein Interaction

MD: Molecular Dynamics

Declarations

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation of Korea (NRF) funded by the Korean government, Ministry of Science & ICT (MSIT) (2022M3A9B6017511). This work was also partly supported by the National Research Foundation of Korea (NRF) funded by the Korean government, MSIT (2020R1A6A1A03047729). The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

The authors gratefully acknowledge the helpful feedback from the staffs of the School of Chemical and Bio Engineering at Addis Ababa Institute of Technology (AAiT). They also thank Prof. Tsige Gebre-Mariam from the School of Pharmacy at Addis Ababa University (AAU) for his valuable insights that enhanced the paper. They express their special appreciation to the team at the Traditional and Modern Medicine Research Directorate (TMMRD) of the Ethiopian Public Health Institute (EPHI), especially Mr. Ashenif Tadele, for providing essential resources that supported this work.

Data Availability Statement

All relevant data are within the paper and Supplementary Files.

References

1. Ekor M. The growing use of herbal medicines: Issues relating to adverse reactions and challenges in monitoring safety. *Front Neurol*. 2014;4 JAN(January):1-10. doi:10.3389/fphar.2013.00177
2. Che CT, George V, Ijinu TP, Pushpangadan P, Andrae-Marobela K. Traditional Medicine. In: *Pharmacognosy: Fundamentals, Applications and Strategy*. Elsevier Inc.; 2017:15-30. doi:10.1016/B978-0-12-802104-0.00002-0
3. Kayne SB. Introduction to traditional medicine. *Traditional Medicine: A Global Perspective*. Published online 2010:25-39.
4. Bird, Matilda van den Bosch W. *Oxford Textbook of Nature and Public Health: The Role of Nature in Improving the Health of a Population*. First Edit. Oxford University Press; 2018. doi:10.1093/med/9780198725916.001.0001
5. Yoo S, Ha S, Shin M, Noh K, Nam H, Lee D. A data-driven approach for identifying medicinal combinations of natural products. *IEEE Access*. 2018;6(58106):58106-58118. doi:10.1109/ACCESS.2018.2874089
6. Bultum LE, Woyessa AM, Lee D. ETM-DB: Integrated Ethiopian traditional herbal medicine and phytochemicals database. *BMC Complement Altern Med*. 2019;19(1):1-11. doi:10.1186/s12906-019-2634-1
7. Kassaye K, Amberbir A, Getachew B, Mussema Y. A historical overview of traditional medicine practices and policy in Ethiopia. *Ethiopian Journal of Health Development*. 2007;20(2). doi:10.4314/ejhd.v20i2.10023
8. WHO. *WHO Global Report on Traditional and Complementary Medicine*.; 2019.
9. Hailu F, Cherie A, Gebreyohannis T, Hailu R. Determinants of traditional medicine utilization for children: a parental level study in Tole District, Oromia, Ethiopia. *BMC Complement Med Ther*. 2020;20(1):125. doi:10.1186/s12906-020-02928-1
10. Negasu G, Banchiamlak N, Mihirat M. Composition, distribution and economic importance of insect pests of prioritized aromatic plants in some growing of Ethiopia. *International Journal of Advanced Biological and Biomedical Research*. 2016;4(1):1-9. doi:10.18869/IJABBR.2016.1

11. Bekele E. Study on Actual Situation of Medicinal Plants in Ethiopia. *Japan Association for International Collaboration of Agriculture and Forestry*. Published online 2007:1-5.
12. Fullas F. *Ethiopian Traditional Medicine: Common Medicinal Plants in Perspective*. F. Fullas; 2001.
13. Suleman S, Alemu T. A survey on utilization of ethnomedicinal plants in Nekemte town, East Wellega (Oromia), Ethiopia. *J Herbs Spices Med Plants*. 2012;18(1):34-57. doi:10.1080/10496475.2011.645188
14. Abebe D, Ayehu A. Medicinal plants and enigmatic health practices of Northern Ethiopia. Published online 1993.
15. Agrawal R, Imieliński T, Swami A. Mining association rules between sets of items in large databases. *ACM SIGMOD Record*. 1993;22(2):207-216. doi:10.1145/170036.170072
16. Larose DT, Larose CD. *Discovering Knowledge in Data: An Introduction to Data Mining*. John Wiley & Sons, Inc.; 2014.
17. Garg A. Complete guide to Association Rules. Published 2018. Anisha Garg
18. Hornik K, Grün B, Hahsler M. arules-A computational environment for mining association rules and frequent item sets. *J Stat Softw*. 2005;14(15):1-25.
19. Srikant R, Agrawal R. Fast Algorithms for Mining Association Rules. *Proceedings of the 20th VLDB Conference Santiago, Chile, 1994*. Published online 1994:487-499.
20. Newman MEJ. *Networks: An Introduction*. Oxford University Press, Oxford; 2010. doi:http://dx.doi.org/10.1093/acprof:oso/9780199206650.001.0001
21. Memon N, Xu JJ, Hicks DL, Chen H. *Social Network Data Mining: Research Questions, Techniques, and Applications.*; 2010. doi:10.1007/978-1-4419-6287-4_1
22. Bullmore E, Sporns O. Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*. 2009;10(3):186-198. doi:10.1038/nrn2575
23. Yu H, Jung J, Yoon S, et al. CODA: Integrating multi-level context-oriented directed associations for analysis of drug effects. *Sci Rep*. 2017;7(1):7519. doi:10.1038/s41598-017-07448-6
24. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*. 1999;27(1):29-34. doi:10.1093/nar/27.1.29
25. Harris MA, Clark J, Ireland A, et al. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res*. 2004;32(Database issue):D258-61. doi:10.1093/nar/gkh036
26. Bolton EE, Wang Y, Thiessen PA, Bryant SH. PubChem Substance and PubChem Compound. *Annu Rep Comput Chem*. 2008;4(D1):217-241. doi:10.1016/S1574-1400(08)00012-1
27. Liu Y, Grimm M, Dai W tao, Hou M chun, Xiao ZX, Cao Y. CB-Dock: a web server for cavity detection-guided protein–ligand blind docking. *Acta Pharmacol Sin*. 2020;41(1):138-144. doi:10.1038/s41401-019-0228-6
28. Abraham MJ, Murtola T, Schulz R, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 2015;1-2:19-25. doi:https://doi.org/10.1016/j.softx.2015.06.001
29. Zoete V, Cuendet MA, Grosdidier A, Michielin O. SwissParam: a fast force field generation tool for small organic molecules. *J Comput Chem*. 2011;32(11):2359-2368. doi:10.1002/jcc.21816
30. Bjelkmar P, Larsson P, Cuendet MA, Hess B, Lindahl E. Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models. *J Chem Theory Comput*. 2010;6(2):459-466. doi:10.1021/ct900549r
31. Golo VL, Shaĭtan K V. [Dynamic attractor for the Berendsen thermostat an the slow dynamics of biomacromolecules]. *Biofizika*. 2002;47(4):611-617.

32. Tuble SC, Anwar J, Gale JD. An Approach to Developing a Force Field for Molecular Simulation of Martensitic Phase Transitions between Phases with Subtle Differences in Energy and Structure. *J Am Chem Soc*. 2004;126(1):396-405. doi:10.1021/ja0356131
33. Mulisa E, Asres K, Engidawork E. Evaluation of wound healing and anti-inflammatory activity of the rhizomes of Rumex abyssinicus J. (Polygonaceae) in mice. *BMC Complement Altern Med*. 2015;15(1):1-10. doi:10.1186/s12906-015-0878-y
34. Lee D. CONET: a virtual human system-centered platform for drug discovery. *Front Comput Sci*. 2018;12(1):1-3. doi:10.1007/s11704-017-7902-y
35. Shahid M, Azfaralariff A, Law D, et al. Comprehensive computational target fishing approach to identify Xanthorrhizol putative targets. *Sci Rep*. 2021;11(1):1594. doi:10.1038/s41598-021-81026-9
36. Consortium TGO. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res*. 2019;47(D1):D330-D338. doi:10.1093/nar/gky1055
37. Augustin N, Nuthakki VK, Abdullaha Mohd, Hassan QP, Gandhi SG, Bharate SB. Discovery of Helminthosporin, an Anthraquinone Isolated from Rumex abyssinicus Jacq as a Dual Cholinesterase Inhibitor. *ACS Omega*. 2020;5(3):1616-1624. doi:10.1021/acsomega.9b03693
38. Getie M, Gebre-Mariam T, Rietz R, et al. Evaluation of the anti-microbial and anti-inflammatory activities of the medicinal plants Dodonaea viscosa, Rumex nervosus and Rumex abyssinicus. *Fitoterapia*. 2003;74(1):139-143. doi:https://doi.org/10.1016/S0367-326X(02)00315-5
39. Tamokou J de D, Chouna JR, Fischer-Fodor E, et al. Anticancer and Antimicrobial Activities of Some Antioxidant-Rich Cameroonian Medicinal Plants. *PLoS One*. 2013;8(2):e55880.
40. Mohammed SA, Panda RC, Madhan B, Demessie BA. Extraction of bio-active compounds from Ethiopian plant material Rumex abyssinicus (mekmeko) root—A study on kinetics, optimization, antioxidant and antibacterial activity. *J Taiwan Inst Chem Eng*. 2017;75:228-239. doi:https://doi.org/10.1016/j.jtice.2017.03.004
41. Muganga R, Angenot L, Tits M, Frédérich M. Antiplasmodial and cytotoxic activities of Rwandan medicinal plants used in the treatment of malaria. *J Ethnopharmacol*. 2010;128(1):52-57. doi:https://doi.org/10.1016/j.jep.2009.12.023
42. Mekonnen T, Urga K, Engidawork E. Evaluation of the diuretic and analgesic activities of the rhizomes of Rumex abyssinicus Jacq in mice. *J Ethnopharmacol*. 2010;127(2):433-439. doi:https://doi.org/10.1016/j.jep.2009.10.020
43. Collier M. Understanding wound inflammation. *Nurs Times*. 2003;99(25):63-64.
44. Guo S, Dipietro LA. Factors affecting wound healing. *J Dent Res*. 2010;89(3):219-229. doi:10.1177/0022034509359125
45. Kummer C, Wehner S, Quast T, Werner S, Herzog V. Expression and potential function of beta-amyloid precursor proteins during cutaneous wound repair. *Exp Cell Res*. 2002;280(2):222-232. doi:10.1006/excr.2002.5631
46. Bodnar RJ. Epidermal Growth Factor and Epidermal Growth Factor Receptor: The Yin and Yang in the Treatment of Cutaneous Wounds and Cancer. *Adv Wound Care (New Rochelle)*. 2013;2(1):24-29. doi:10.1089/wound.2011.0326

Figures

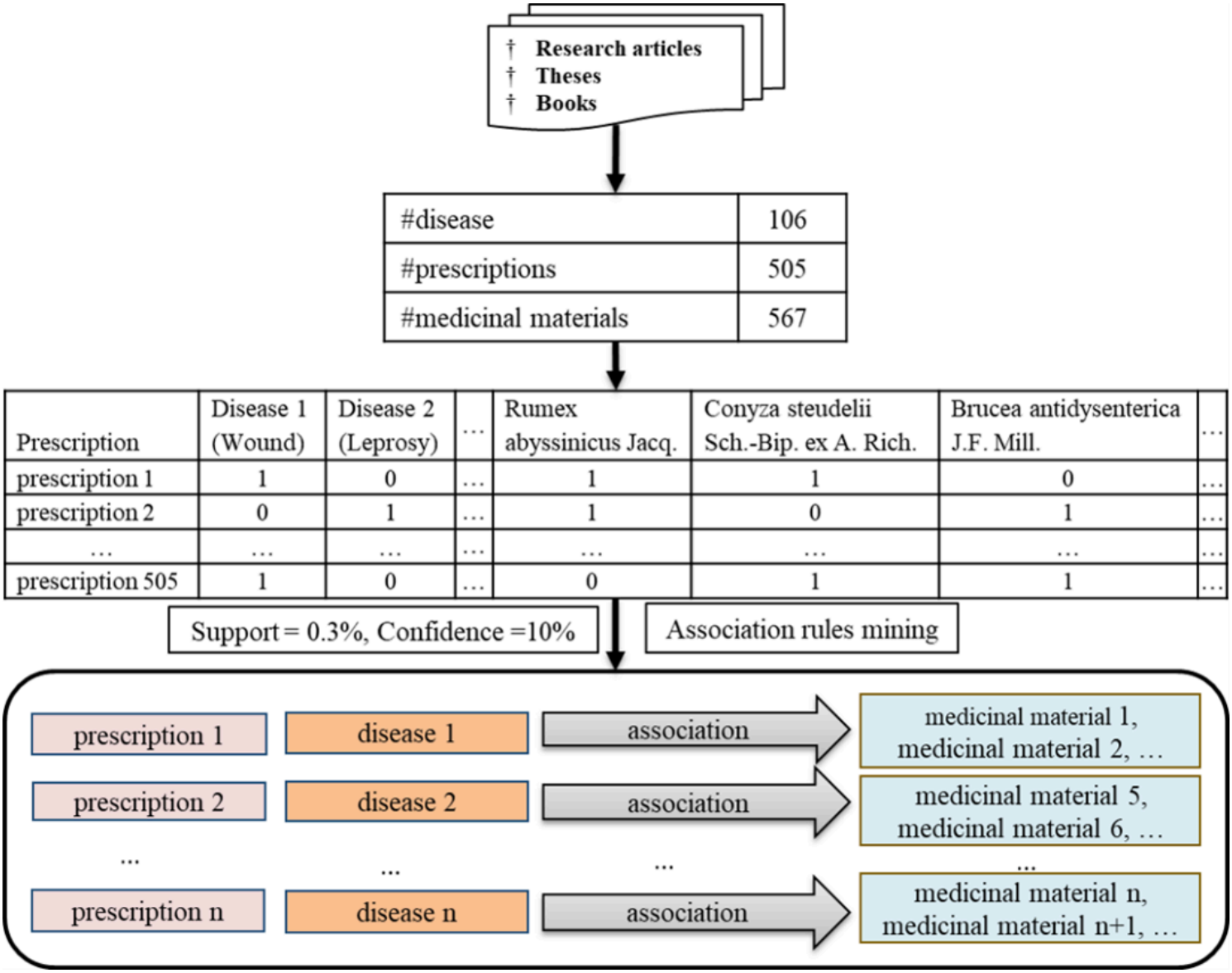

Figure 1

Overview of the process to identify associations between medicinal materials and human diseases from prescription data. Multicomponent prescriptions, constituent medicinal materials, and related diseases were organized in a transaction database. Association rule mining was then employed to extract significant associations between medicinal materials and diseases from the multicomponent prescriptions.

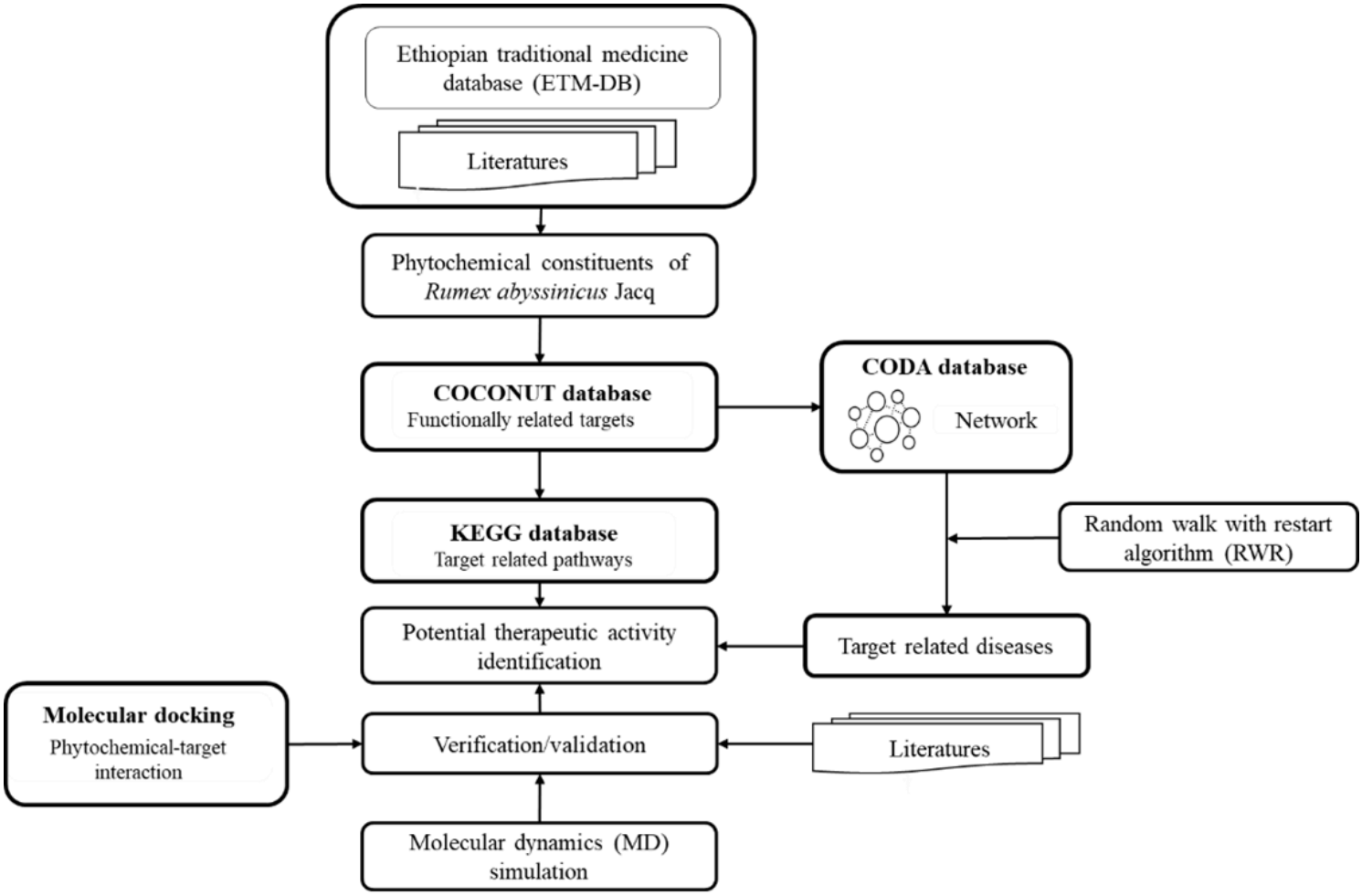

Figure 2

Process flow diagram to identify potential therapeutic activity of R. abyssinicus Jacq phytochemicals.

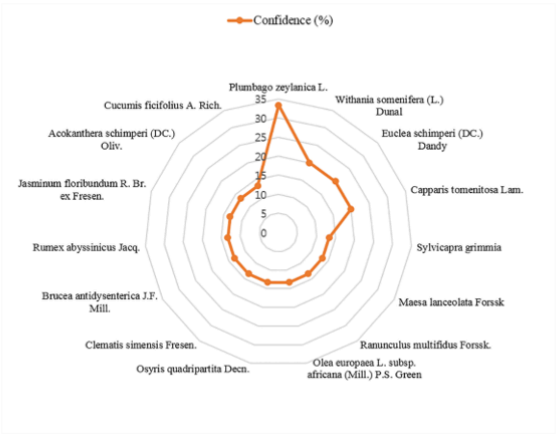

Fig. 3(a)

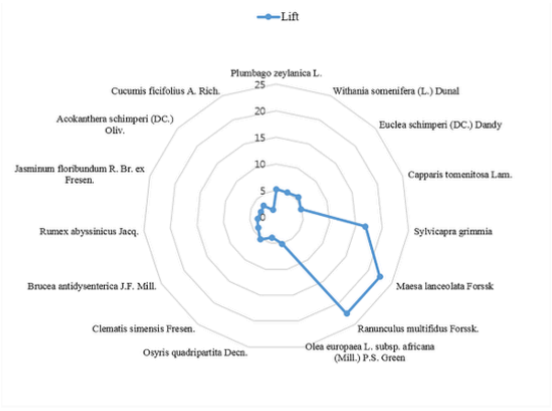

Fig. 3(b)

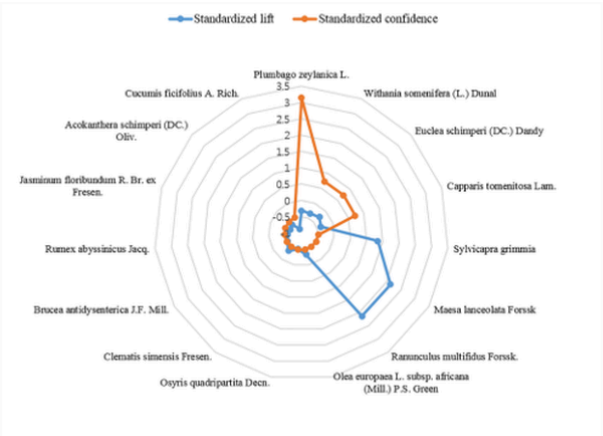

Fig. 3(c)

Figure 3

(a) Radar charts of confidence values for association rules between leprosy and its associated medicinal materials.

(b) Radar charts of lift values for association rules between leprosy and its associated medicinal materials.

(c) Radar charts of standardized confidence and lift values for association rules between leprosy and its associated medicinal materials.

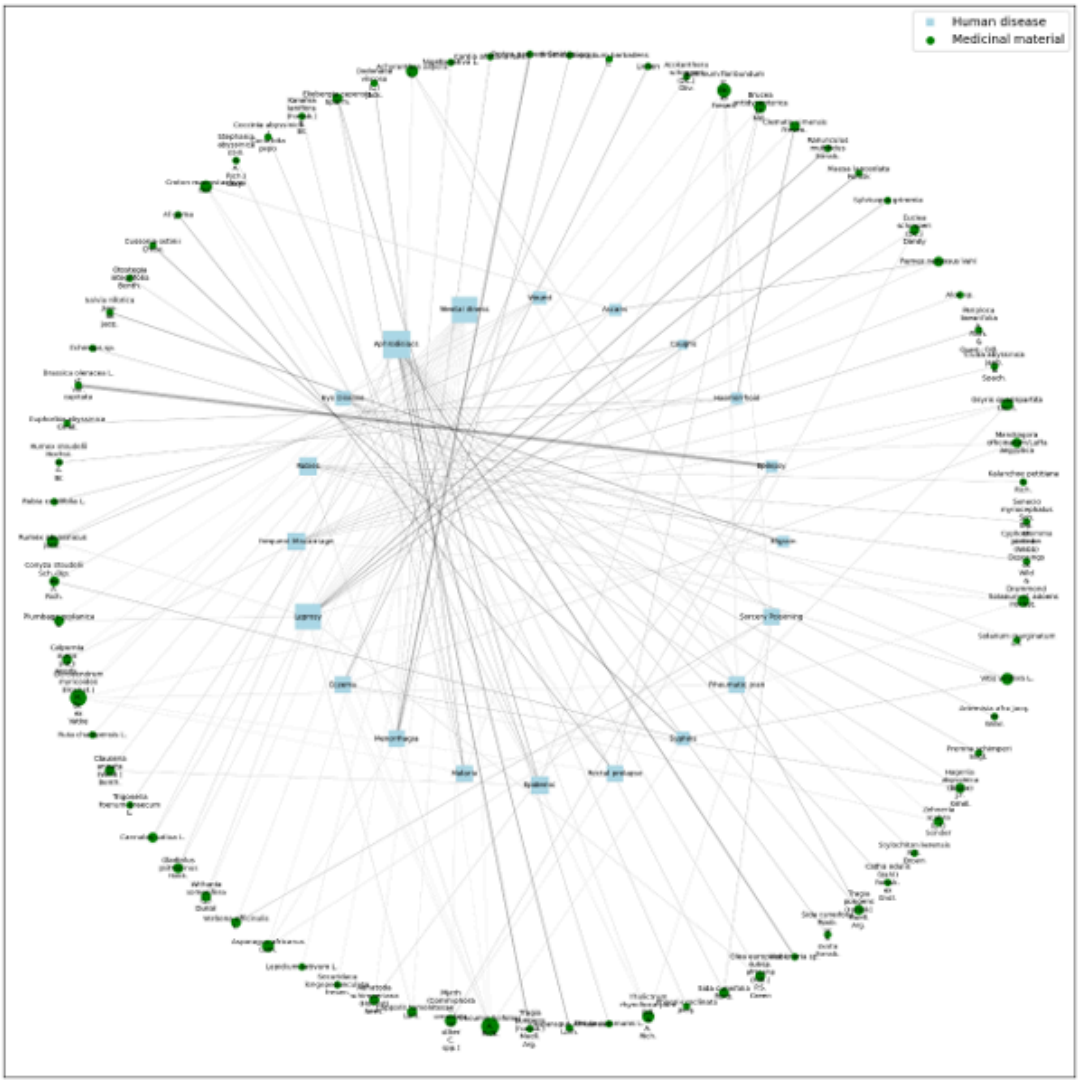

Figure 4

The network graph illustrates the associations between a set of 20 diseases and the medicinal materials. In the graph, diseases are represented by light blue squares, while medicinal materials are denoted by green circles. The size of the nodes indicates their relative degree, and the edges between nodes signify that the prescriptions for treating specific conditions include the respective therapeutic materials. The width of the edges reflects the strength of the association, measured by the lift, between the medicinal materials and the diseases.

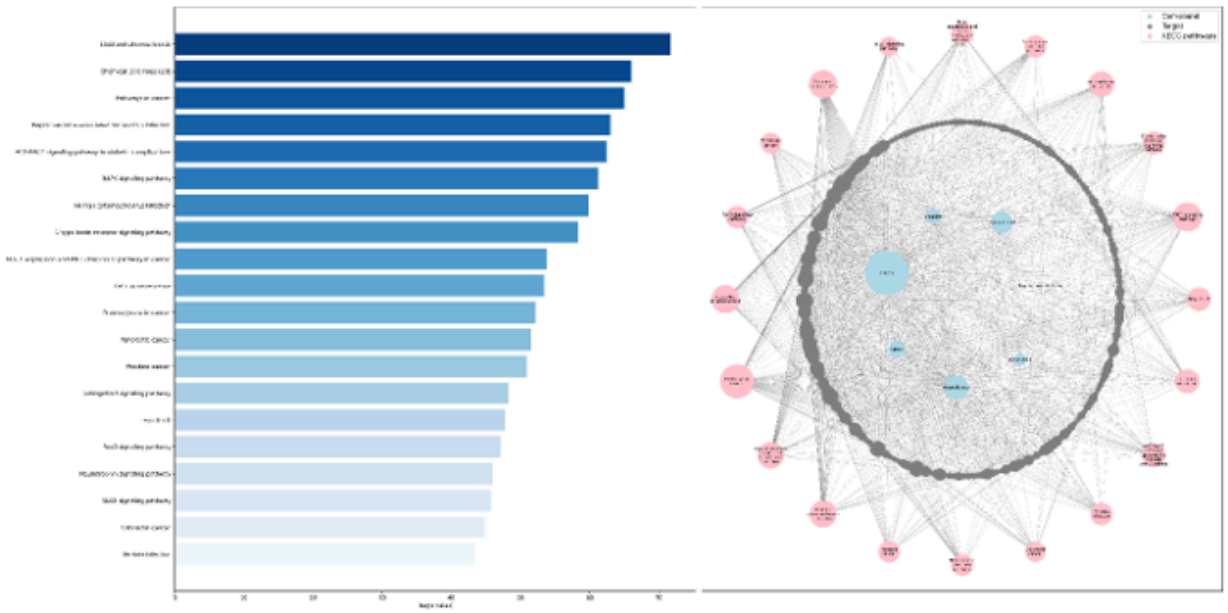

Figure 5

A bar graph showing the top 20 enriched pathways from KEGG analysis (right). Phytochemical-target-pathway network constructed by python NetworkX library illustrated pathways and their associated genes (Adjusted P-value ≤ 0.05) (left).

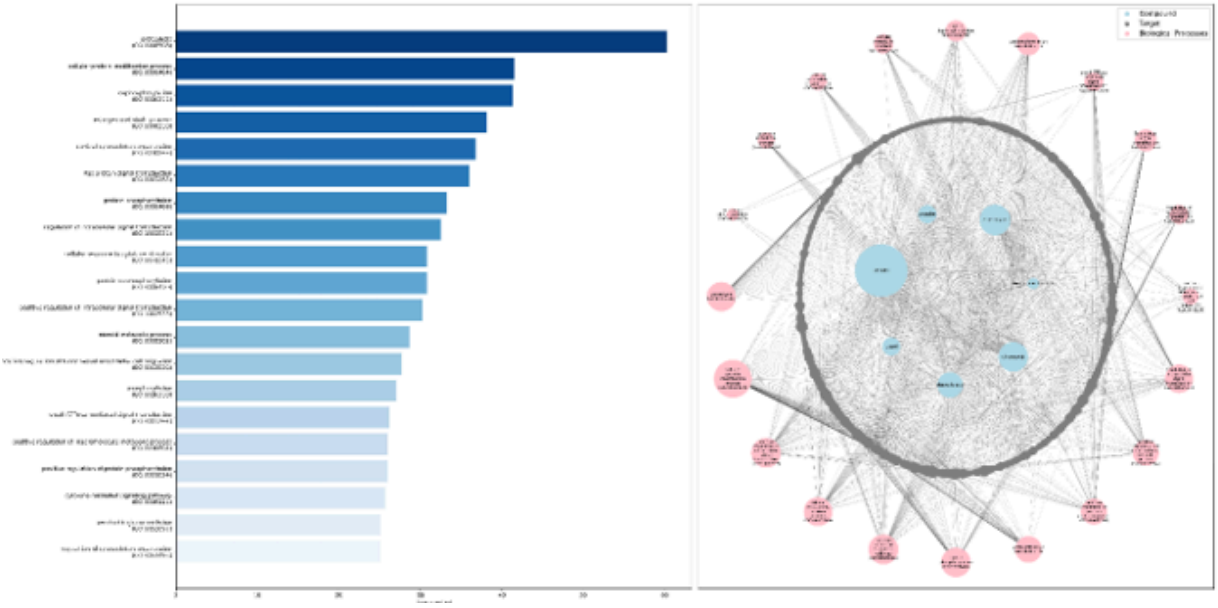

Figure 6

A bar graph showing the top 20 enriched biological processes from the GO analysis (right). Phytochemical-target-biological process network created using the Python NetworkX package shows pathways and the genes related to the phytochemicals (adjusted P-value ≤ 0.05) (left).

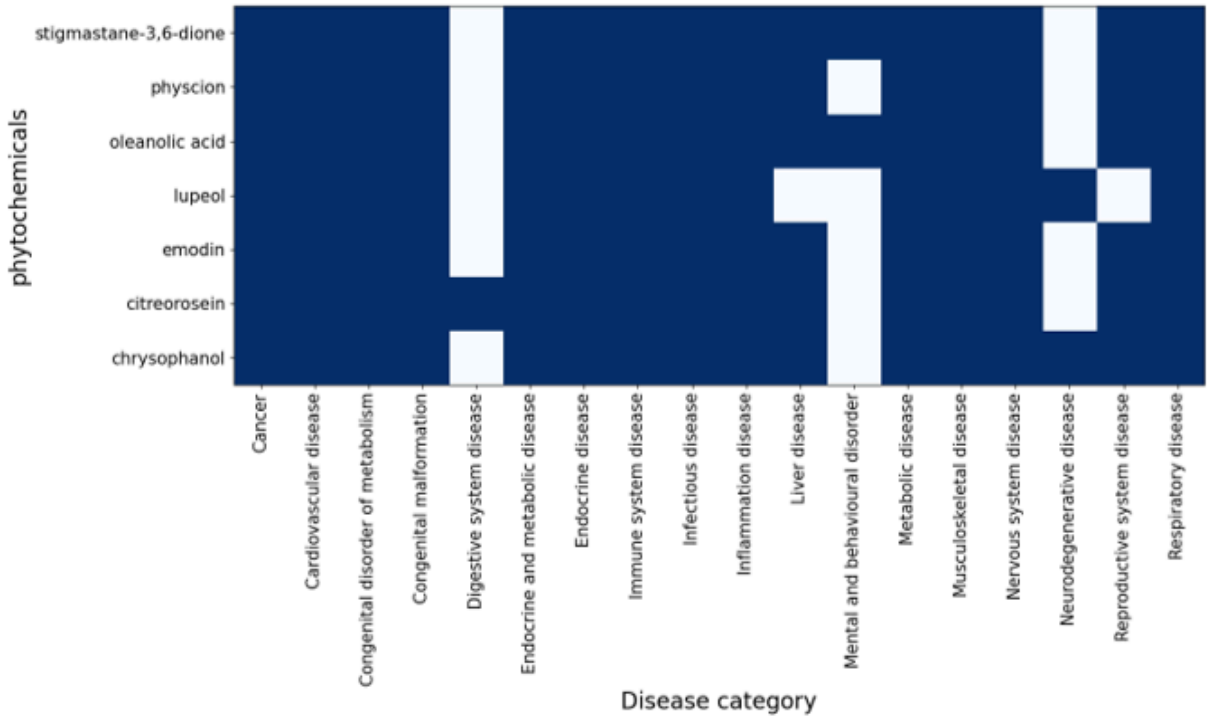

Figure 7

Image map of phytochemicals and related diseases.

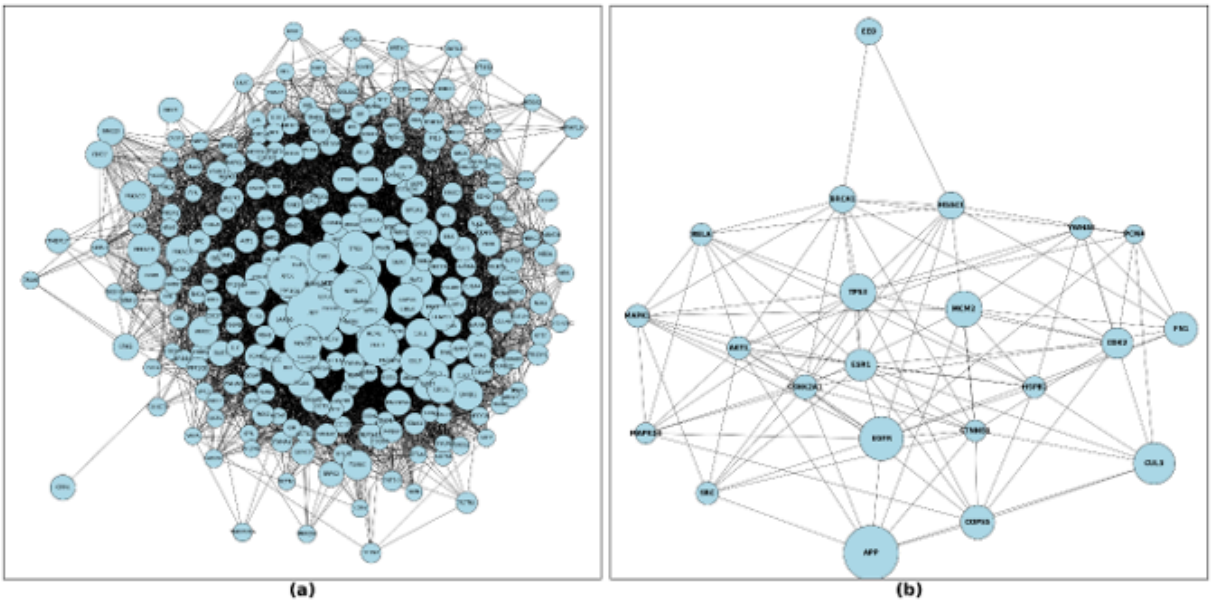

Figure 8

(a) PPI network constructed with CODA. (b) Network of R. abyssinicus Jacq phytochemicalsrelated hub genes.

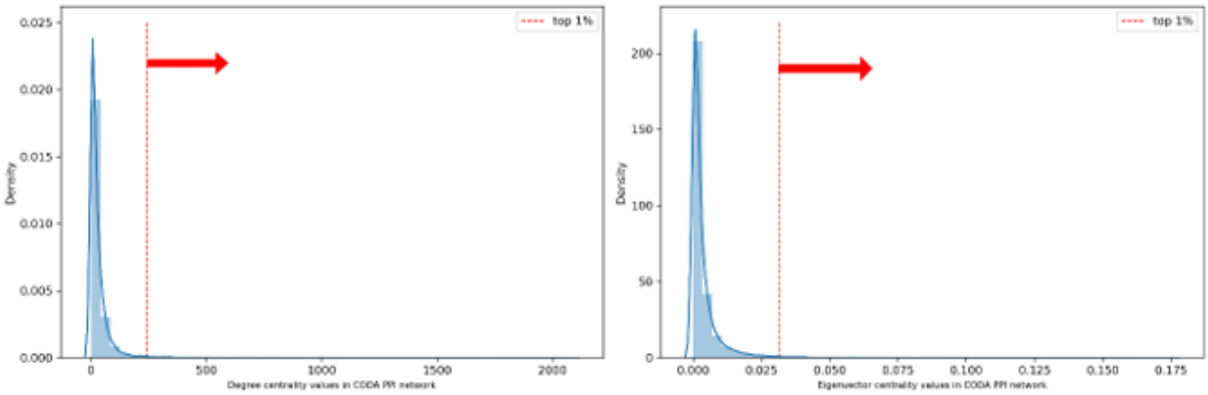

Figure 9

Distribution of R. abyssinicus jacq phytochemicals-related hub genes centralities in a CODA PPI network.

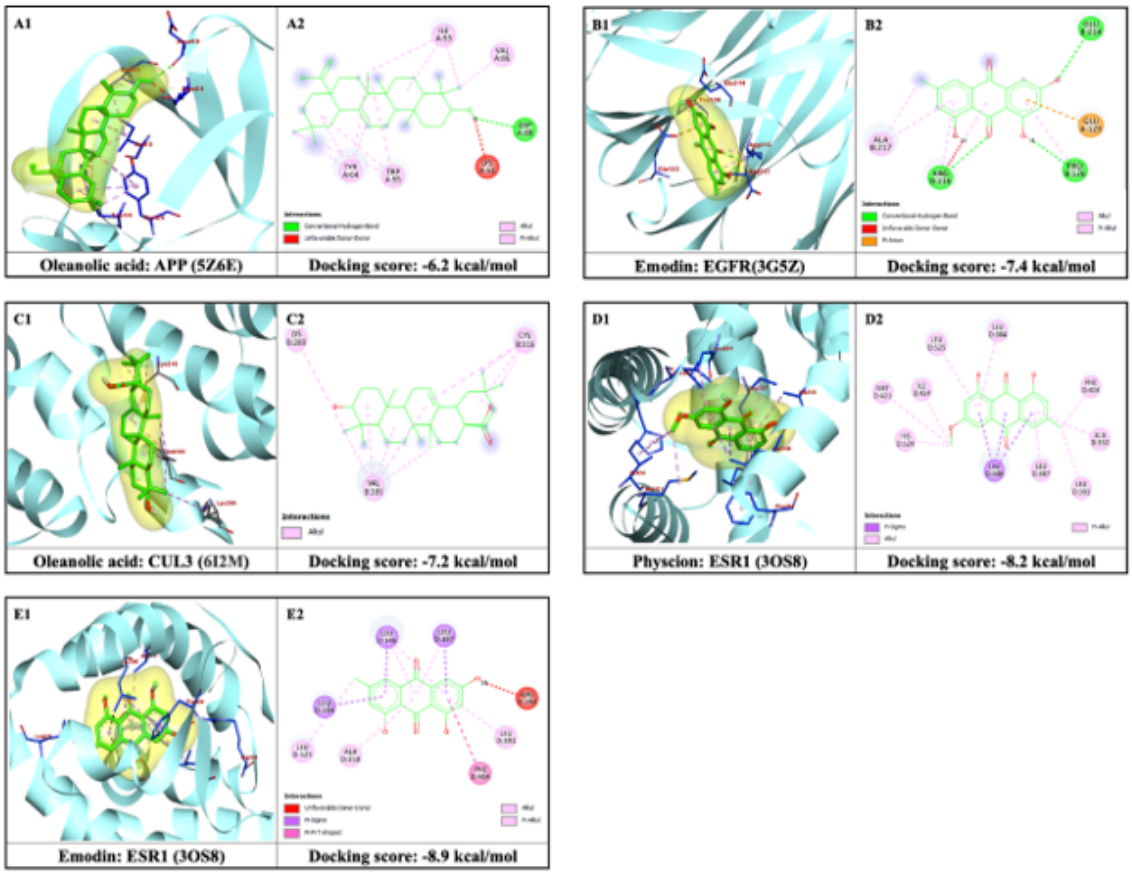

Figure 10

Molecular docking analysis of select phytochemicals from R. abyssinicus Jacq with their corresponding targets. (A1-E1) 3D visualization of phytochemical and target interactions. (A2-E2) 2D pose views depicting phytochemicals and their interactions with targets.

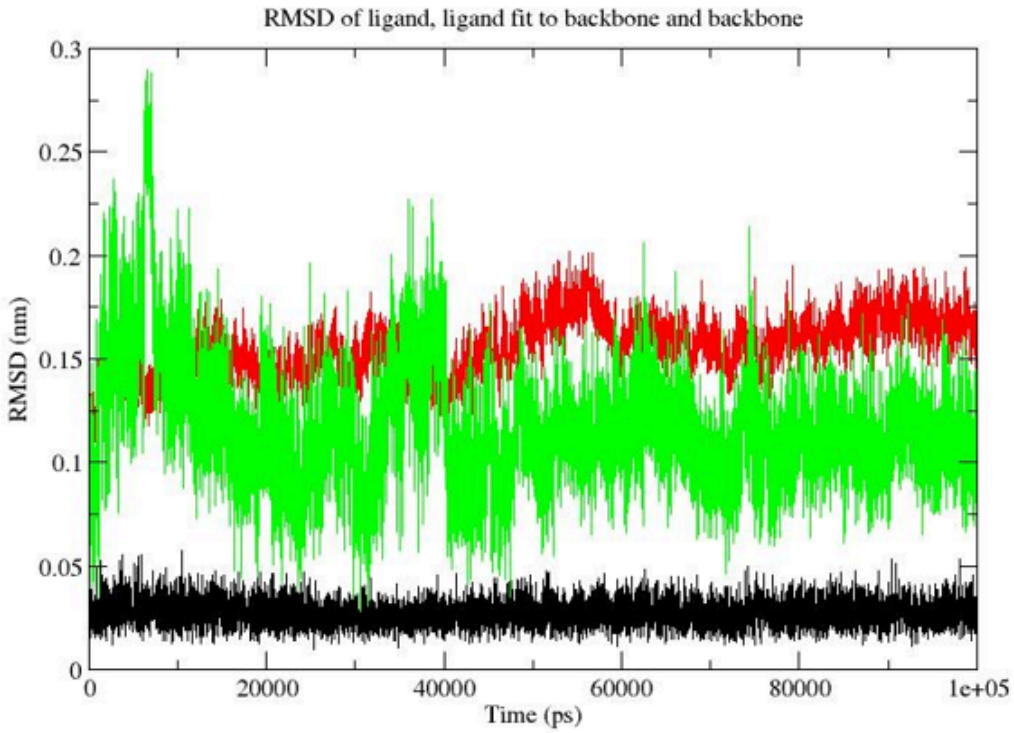

Figure 11

RMSD plots of the ligand (emodin) (black), target protein backbone (ESR1) (red), and ligand fit to backbone (green) over 100 ns MD simulation.

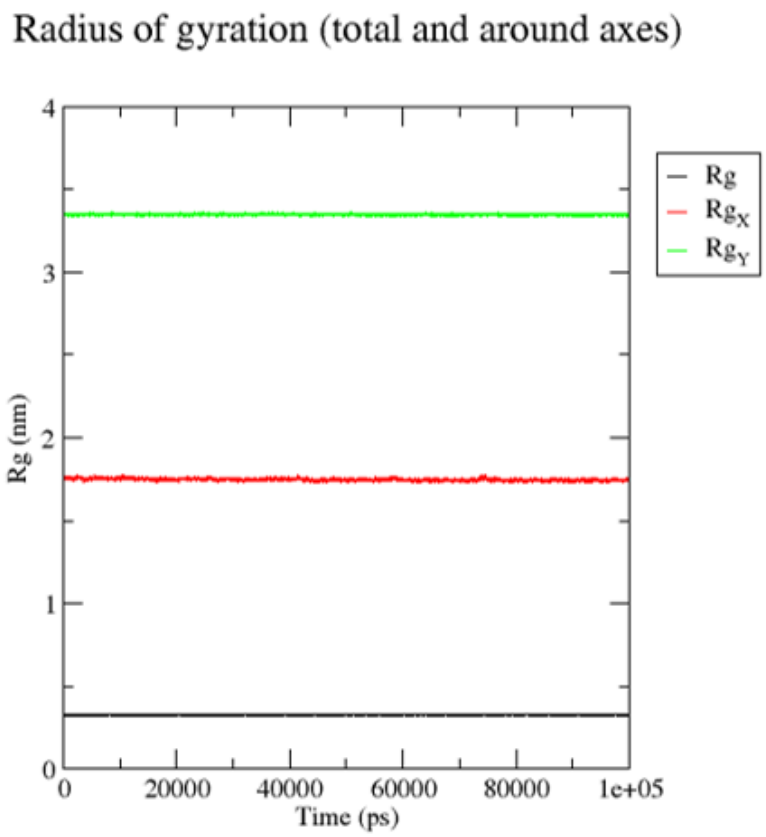

Figure 12

Radius of gyration, Rg (nm) plots of the ligand (emodin) (Rg), target protein backbobne (ESR1) (RgX) and the system (RgY) over 100 ns simulation.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1.xlsx
- TableS2.xlsx