ORIGINAL ARTICLE

Ethnomedicine Claim Directed in *Silico* Prediction of Anticancer Activity

Biniyam Girma[1]*, Eshetu Mulisa[1], ShibruTessema[1], Wote Amelo[1]

Affiliation and Correspondence:
[1]School of Pharmacy, Faculty of Health Sciences, Institute of Health, Jimma University, Jimma, Ethiopia
*Email: Semayat.girma@gmail.com

ABSTRACT

BACKGROUND: The merits of ethnomedicine-led approach to identify and prioritize anticancer medicinal plants have been challenged as cancer is more likely to be poorly understood in traditional medicine practices. Nonetheless, it is also believed that useful data can be generated by combining ethnobotanical findings with available scientific studies. Thus, this study combined an ethnobtanical study with ligand based in silico screening to identify relevant medical plants and predict their anticancer potential based on their phytoconstiutents reported in scientific literatures.

METHODS: First, relevant medicinal plants were identified through an ethnobotanical survey. A list of phytochemicals was prepared based on literature review of articles which reported on the natural products of identified medicinal plants. Then, their phytochemicals were subjected to in silico evaluation, which included a hybrid score similarity measure, rule of five, Ghose–Viswanadhan–Wendoloski (GVW)-indices and structural features criteria, to predict their anticancer activity and drugability.

RESULTS: A total of 18 medicinal plants and 265 phytoconstituents were identified. The natural product pool constituted 109(41.13%) terpenoids, 67(25.28%) phenolics, 29(10.94%) simple and functionalized hydrocarbons, 26(9.81%) alkaloids, 25(9.43%) glycosides and 9(3.40%) compounds belonging to different phytochemical classes. The similarity measure using CDRUG identified 34(12.73%) phytochemicals with high (p-Value < 0.05) and 35(13.21%) with moderate possibility (p-Value < 0.1) of anticancer activity. In fact, three of the predicted compounds had the same structure with known anticancer compounds (HSCORE=1). The 80% GVW-indices based antineoplastic drugabilityranges were all mate by 25 of the predicted compounds. Predicted compounds were also shown to have ring structures and functional groups deemed important for anticancer activity.

CONCLUSIONS: Given the findings, there is a promising anticancer activity by the traditionally used medicinal plants and a potential for the predicted phytochemicals to be pursued as possible hits or me-too drugs.
KEYWORDS: Ethnobotany, phytochemicals, in silico, anticancer, medicinal plants

INTRODUCTION

Cancer is a growing public health concern in many developing countries. For example, in Africa only, there were about 715,000 new cancer cases and 542,000 deaths in 2008. These figures are also expected to double in the next twenty years (1).

Currently, the lack of access to modern cancer care coupled with other factors makes many cancer patients of the developing world dependent on traditional medicines (2). Given the challenges of finding novel anticancer medications which are effective, safe, affordable and accessible, ethnobotanical and ethnopharmacological studies play important roles in identifying relevant medicinal plants that can be developed further.

Medicinal plants, which formed the basis of traditional medicine practices, have long been used by mankind for preventive and curative purposes (3,4,5). Historically, ethnobotanical studies have resulted in drug discovery for numerous ailments including cancer (6,7). However, the extrapolation into western medicine has no always been direct like in the case of vincristine where therapeutic efficacy was remotely correlated with the indigenous use (8). Apart from the general success of the ethnobotanical approach, its values and performance vary among different diseases. Its scientific importance has also been challenged in disease like cancer which are believed to be poorly characterized in traditional medicine practices (9).

In spite of ethnobotanical study based claims of anticancer activity made by many, the National Cancer Institute (NCI) argues that it is unlikely for cancer to be established as a single disease entity in traditional medicine, particularly those internal organ neoplasms such as lung, colon, ovarian and prostate cancers (10).

Such a challenge to the ethnomedicine based approach has been further complicated by the lack of theory based practice in many of the traditional medicine practices which makes the information on diseases and plants used less reliable. Hence, the ethnobotanincal data gathered will in general be inadequate to identify relevant medicinal plants for further expensive biological investigations (11).

Still, ethnobotanical reports can provide reliable information in external tumors like skin cancer, external growths and swellings, tumors of the oral cavity, genital tumors despite the general issue of establishing cancer as a specific diseases in ethnomedicine (10). Besides, their relevance in selecting plants for further investigation in the area of cancer can be enhanced by corresponding the findings with existing scientific literature (11). Virtual screening methods, which are cost effective, fast and reliable, could be utilized to correlate findings from existing scientific studies and ethnobotanical surveys (12). In retro, such methodological combination could also help understand the indigenously defined malady in terms of its links to cancer.

To this end, our study combined an ethnomedicne study with a ligand based *in silico* method. In doing so, the study aimed at providing a scientific basis for the existing traditional medicine practice in a complex disease like cancer and exploiting its potential for drug discovery and development. Accordingly, we studied the potential anticancer activity and drugability of phytochemicals found in plants identified based on an ethnobotanical survey.

MATERIALS AND METHODS

Ethno-botanical part

Study area: The survey was conducted in Harari Region. Harari Region is located in eastern part of Ethiopia about 500km from Addis Ababa. The regions estimated area is about 311.25 square kilometres. It is located at $09^018'33''$ N latitude and $42^007'32''$ E longitude, and situated at an elevation of about 1861m above sea level. The mean annual rainfall is about 858mm. The

region is divided into 19 urban and 17 rural kebeles. Harari region is one of the most diversified regions of Ethiopia with different ethnic and religious groups inhabiting it.

Data collection and analysis: We have collected ethnomedicine data on a malady locally called *'Gofla'*. It is a term used by the local people to express a variety of body part swellings, and also claimed to include neoplastic tumors. The survey included 23 respondents among which 10 were traditional healers, 6 were incense and herbal medicine sellers and 7 were elderly persons.Verbal consent was sought from interviewees before data collection. Finally, taxonomic identification of collected specimens was done at Haramaya University herbarium. The data were analysed and presented in a table containing description of the medicinal plants and their use.

In Silico Part

Identification phytochemicals in the medicinal plants: After taxonomic identification, a review was conducted to determine the different phytoconstituents reported in each herb. The search for relevant articles was made on Google Scholar (http://scholar.google.com/) and PubMed(https://www.ncbi.nlm.nih.gov/pubmed). Once the search was finalized, reported phytochemicals were extracted and compiled in to a list.

Prediction of anticancer activity: The anticancer activity was studied *in silico*using a web server named Cancer Drug (CDRUG; http://bsb.kiz.ac.cn/CDRUG/).CDRUG uses relative frequency-weighted fingerprint and a hybrid score*(HSCORE)* to determine the similarity between compound of interest and those found within its dataset. It computes*p-values* for the similarity measures based on which anticancer activity is predicted as highly possible, possible and less possible. It has the ability to hit 65% positive results at a rate 0.05 false-positive (13). Since it requires the Simplified Molecular Input Line Entry String (SMILES) of compounds to be investigated, PubChem(https://pubchem.ncbi.nlm.nih.gov/) and FoodDB(http://foodb.ca/) databases were used as sources of required entries. For a small

number of compounds, line strings were also determined on Molinspiration Cheminformatics (www.molinspiration.com/). Ultimately, the SMILES and a single word name of the query compounds were fed in to CDRUG as inputs for the prediction analysis.

Evaluation of general and anticancer drug-likeness: The analysis to discriminate between drugable and non-drugable compounds was done only for those compounds predicted to have anticancer activity. Initially, they were evaluated against Lipinski's rule (14). Next, Ghose–Viswanadhan–Wendoloski (GVW) drug like indices, which are based on the distribution of physicochemical properties and chemical components in Comprehensive Medicinal Chemistry (CMC) database, were used for the evaluation. Among the indices determined for general drug likeness and seven disease based drug classes, we used the physicochemical property based ranges (Preferred and Qualifying range) particularly set for antineoplastic drugability. We also determined the presence of relevant functional groups, as determined by *Ghose et al* (15). All required physicochemical properties were determined online by Molinspiration (www.molinspiration.com/) and E-Dragon(http://www.vcclab.org) using SMILES and SDF files as required (16). Finally, additional structure features of drug-likeness like aromatic rings and proportion of heavy atoms (R-value) were assessed(17).

RESULTS

The initial ethnobotanical survey gathered information from twenty three participants, and we were able to identify eighteen different medicinal plants and five additives used in the preparation of medications for swellings in different parts of the body, locally called *'Gofla'*. The plants mainly belonged to *Asteraceae* and *Lamiaceae* families**.** The most commonly reported plant was *Senecionevadensis* (9) followed by *Hydnorajohannis* and 'Shefewein' each being reported eight times by respondents. *Viscumschimperi*, *Aloe vera* and *Brassica carinata* were the third most frequently reported medicinal plants.These plants were used to treat

Table 1: List of medicinal plants used for the treatment of swellings to different body parts*('Gofla' locally)*in Harari region

Voucher number	Scientific name	Family	Local name	Use description	Swelled body part
2765	*Aloysia triphylla*	Verbenaeae	Kute Nana	-Leaves and stems are crushed, boiled with water, tea or milk and taken orally	Neck, breast, abdomen
662	*Hydnora johannis*	Hydnoraceae	DechMerech	-Root of *Hydnorajohannis* are crushed, boiled with water containing exocarp/peel of coffee beans and taken orally	Breast, abdomen
**			Shefewein	-Shefewein are crushed, boiled with water containing exocarp/peel of coffee beans and taken orally	
1146	*Senecio nevadensis*	Asteraceae	Balkutel	-Leaves of *Senecioabyssinicus*are boiled with water containing exocarp/peel of coffee beans and taken orally	
2970	*Echinops kebericho*	Asteraceae	Kebericho	- All three plants mentioned above and roots of*Echinopskebericho are* also combined, boiled together with exocarp/peel of coffee beans and taken orally.	
7491	*Punica granatum*	Punicaceae	Roman	-Rind of the fruit is boiled with water and taken orally	Abdomen
9836	*Boswellia papyrifera*	Burseraceae	Habeket	-Gum of the plant is macerated in cold water which forms a thick gel taken orally	Abdomen
2053	*Osyridocarpus schimperanus*	Santacaceae	Muka- Gofela	-Stems are crushed, rolled in a small paper like a cigar and smoked. The smoke is held around the neck area and exhaled	Neck
2075	*Aloe harlana*	Aloeaceae	Rate	- Leaves are expressed and filtrate is taken orally	Abdomen
2775	*Brassica carinata*	Cruciferae	Goman-zer	-Seeds are crushed, mixed with honey and taken orally. Also applied on breast externally	Breast
3227	*Zaleya pentandra*	Aizoaceae	Wachara-haree	-Leaves are boiled with water and taken orally	Liver
5960	*Sphaeranthus suaveolens*	Asteraceae	Hadeshshadye	-Leaves and stems *Sphaeranthussuaveolens*are crushed, the juice squeezed and mixed with the concoction of *Senecionendensis* leaves	Neck, breast, abdomen
2857	*Senecio abyssinicus*	Asteraceae	Generas		
2225	*Lasiocorys abyssinica*	Lamiaceae	Mukawakwak	-leaves and stems of the plant is crushed, boiled with water and taken orally	Breast
				-leaves and stems are powdered, mixed in cold water and applied externally on the breast	
3147	*Fagera usambarensis*	Rutaceae	Geda	-Leaves and stems are crushed and boiled with water containing exocarp/peel of coffee beans and taken orally	Neck, breast, abdomen
697	*Amaranthus spp*	Amaranthaceae	AnenaKuti	-Leafs and stems are boiled with tea or milk and then taken orally	Neck, breast, abdomen
7643	*Eulophia petersii*	Orchidaceae	Shenkurtagara	-Bulbs are cooked and eaten	Abdomen
7850	*Viscum schimperi*	Loranthac	Digelo	-Leaves and stems are crushed, boiled with water and taken orally	Neck, breast, abdomen

swellings in parts like the neck, the breast, the abdomen and the liver. Leaves and stems, which collectively accounted for over 60%, were commonly used parts. The use of roots was also relatively common, 3(15.8%). The four additives were exocarp/peel of coffee beans, water, tea, milk and honey. The medications were almost always administered orally that may indicate understanding of the need for systemic administration in treating the tumours (Table 1). Even though we initially determined 18 medicinal plants, only 14 were taxonomically identified. The following review managed to retrieve relevant literatures for 13 of the identified species and a total of 265 phytochemicals. This initial pool included 109(41.13%) terpenes/terpenoids, 67(25.28%) phenolic compounds, 29(10.94%) saturated and unsaturated hydrocarbons, 26(9.81%) alkaloids, 25(9.43%) glycosides, and 9(3.40%) other phytochemicals. Their analysis on CDRUG predicted 34 (12.73%) of them as highly likely to possess anticancer activity *(P-value < 0.05)*. Another 35(13.21%) of the phytoconstituents were also predicted to have possible anticancer activity *(P-value < 0.1)*. The mean logGI50 (the 50% growth inhibition concentration) values of the natural products with high possibility of antineoplastic activity ranged from -4.681 to -5.981, and the logGI50 of those constituents with possible bioactivity ranged from -5.000 to -7.914. Among the 34 compounds predicted to have high possibility of anticancer activity, 17(14 triterpenoids, 2 sesquterpenoid & 1 diterpenoid) were terpenes/terpenoids 9 (1 phenolic acid, 6 flavanoids, 1 phenylpropanoid &1 anthraquinone) were phenolic compounds, 5 (4

isoquinolones & 1 indole) were alkaloids, 2 were steroidal glycosides, and 1 was unclassified. Similarly, the possible group contained 17(4 monoterpenoids, 11 sesquterpenoids & 2 diterpenoids) terpenes/terpenoids, 14(5 phenolic acids, 4 flavanoids, 2 stilbenoids & 1 phenylpropanoid) phenolic compound, 3(2 isoquinolone &1 indole) alkaloids, and 1 phenolic glycoside.

The drugability of the compounds was assessed by combing methods that are based on physicochemical properties and structural features. First, the general drug-likeness of predicted phytochemicals was evaluated using the Lipinski's rule of five. The rule states that most of orally absorbed drugs possess hydrogen bond donors (HBDs) ≤ 5, hydrogen bond acceptors (HBAs) ≤ 10, a molecular weight (MW) ≤ 500 Daltons and octanol/water partition coefficient (LogP) value of ≤ 5 (18). The analysis of these molecular properties for the highly possible compounds have shown that 13(38.2%) and 18(52.9%) of the predicted active compounds have no or one violation, respectively. Among the possible compounds, 16(45.7%) have no violation and 16(45.7%) have only one violation. Thirty-one of the highly possible compounds and thirty of the possible compounds have molecular weight less than 500 Daltons. The LogP distribution showed that 16(47.1%) of the highly possible compounds and 24 (68.6%) of the possible compounds were under the limit set by the Lipinski rule. Over seventy-five percent of predicted compounds on both groups possessedfive HBDs or fewer. Over ninety percent of the predicted active compounds were with ten HBAs or fewer (Figure 1).
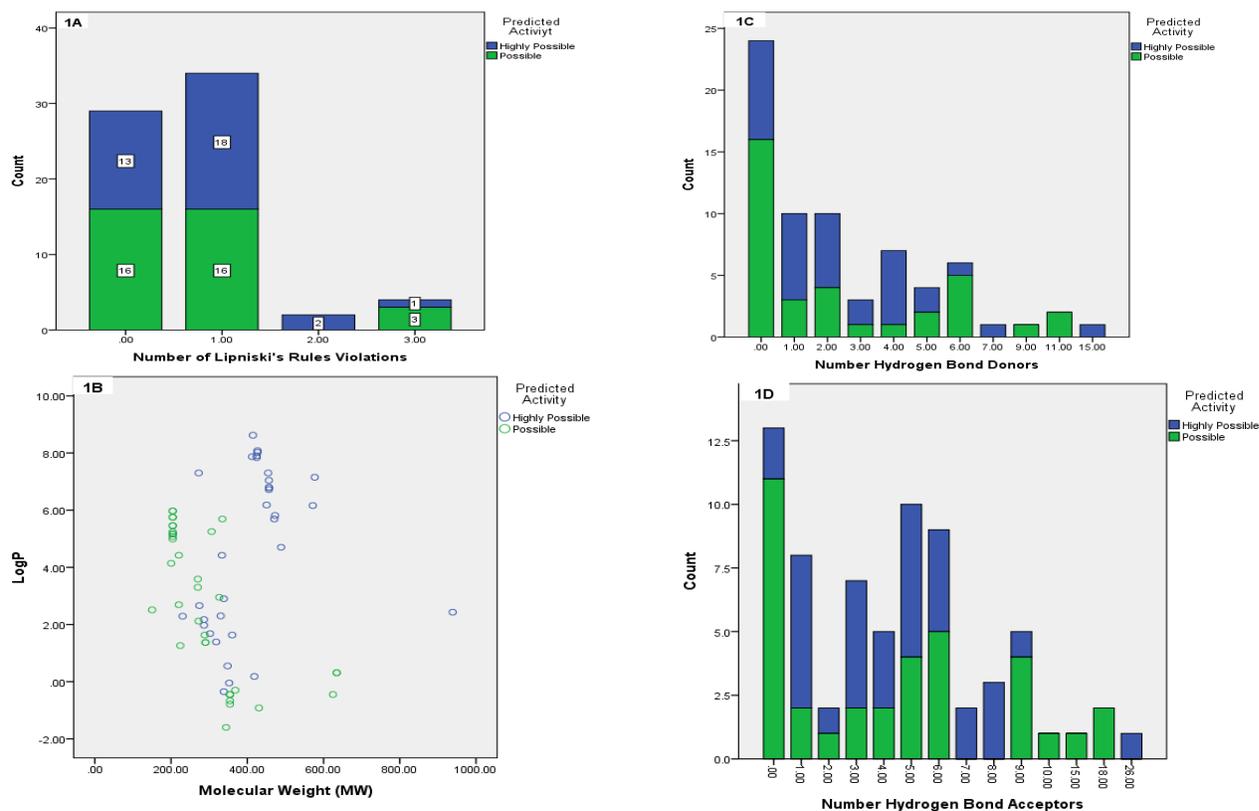
Figure 1: *Drug-likeness of predicted phytochemicals based on Lipinski's rule of five (1A): Distribution of predicted phytochemicals based on the number of violations of the rule of five. (1B,1C,1D): Distribution of molecular weight, LogP, number of hydrogen bond donors and number of hydrogen bond acceptors of predicted phytochemicals.*

According to *Ghose et al.,*Ghose–CrippenLogP (ALogP), Ghose–Crippen molar refractivity (AMR), molecular weight (MW) and number of atoms (nA) can be used collectively to determine drug-likeness (19). Their analysis showed that none of the predicted active compounds fulfilled all the criteria of the preferred range. The individual violations of the preferred range constituted 40(58.0%), 39(56.5%), 22(31.9), 47(68.1%) for MW, ALogP, AMR and nA, respectively. Unlike the preferred range, a total of 25 phytochemicals both from the highly possible [12(35.3%)] and the possible [13(37.1%)] lists fall within the limits set for the qualifying range (Figure 2).

Analysis of rings and functional groups showed that 18(52.9%) of the highly possible natural products were with two or more of the structural features believed to be important for drug-likeness. A similar feature was also observed in 19(54.3%) of the possible phytochemicals. Heterocyclic rings (mainly aliphatic), alcohol and benzene were the most common groups in both sets of phytoconstituents. Functional groups like carboxyester, keto, aliphatic amine and hetero aromatic rings were also found in the predicted compounds. Nonetheless, no carboxamide group was found in any of the compounds. Besides, over seventy percent of the predicted phytochemicals were found to possess proportion of heavy atoms, i.e. a ratio of non-hydrocarbon atoms to the number of all non-hydrogen atoms *(R-value)*, between 0.05 and 0.50 (Figure 2).
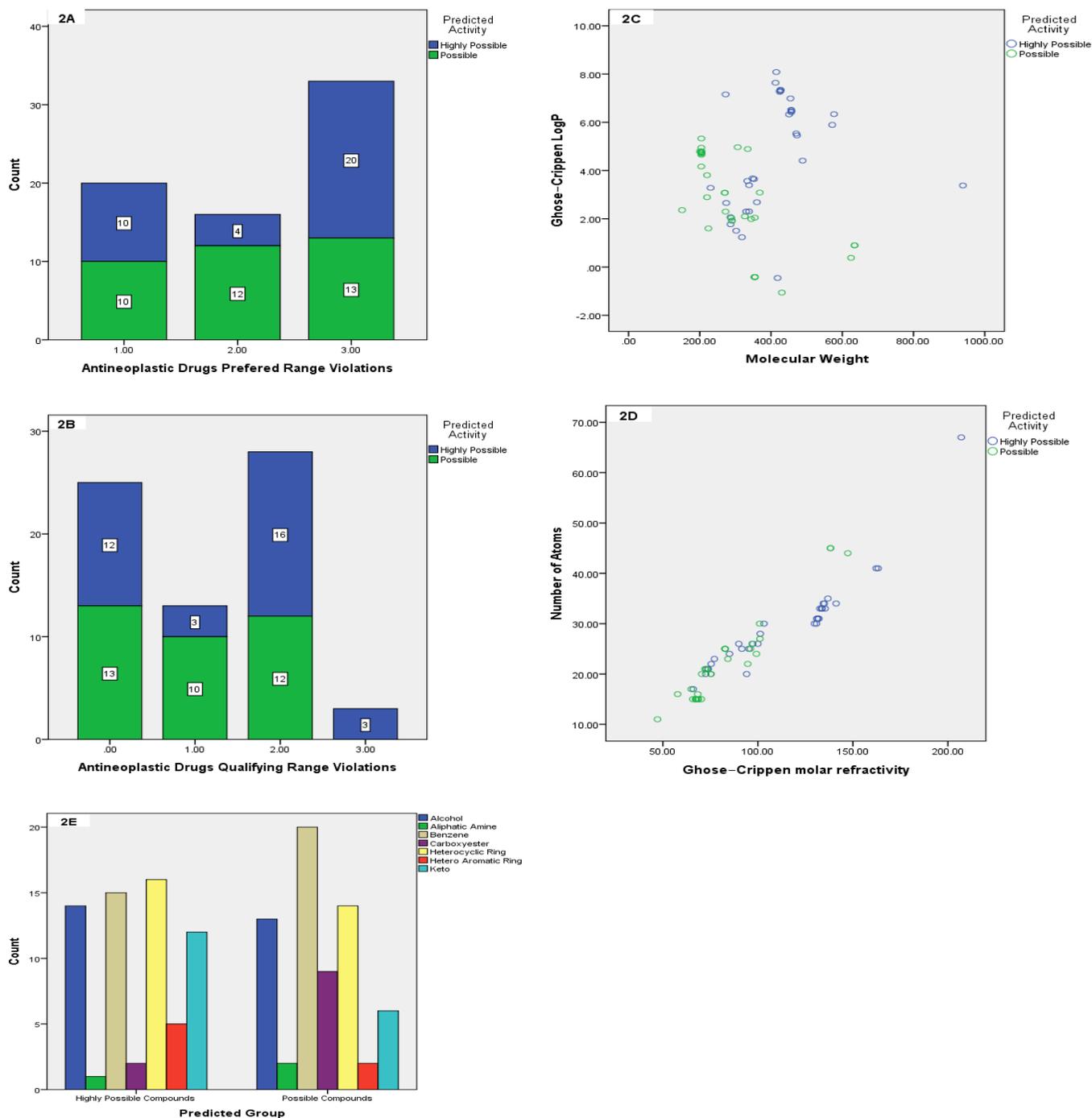
Figure 2: Antineoplastic drug-likeness of predicted phytochemicals according to GVW-indices based ranges. (2A, 2B): Bar graphs of number of violations to the preferred and qualifying antineoplastic drug ranges by predicted phytochemicals. (2C, 2D): Distribution of molecular weight, Ghose–CrippenLogP, Ghose–Crippen molar refractivity and number of atoms (nA) of predicted phytochemicals. (2E): Histogram of relevant structural features found in the predicted phytochemicals

DISCUSSION

Traditional medicines, which are largely based on the use of plants as drugs, are estimated to be used approximately by 80% of the population in developing countries as primary healthcare sources (20). This estimation may be higher for diseases like cancer where treatment options are usually inaccessible and/or unaffordable. Given this magnitude of use, it is in the public's health interest to identify these medicinal plants and generate scientific evidence on their efficacy. Hence, we conducted an ethnobotanical survey followed by a ligand based computational study to identify and evaluate their therapeutic value claimed in the traditional practice.

Medicinal plants have played a huge role in anticancer drug discovery. They contributed to the discovery of around 48% of all anticancer drugs between 1940 and 2002 (21). Many of these plants were used in different traditional medicines for ages. Although the importance ethnobotanical directed drug discovery efforts have been challenged due to their limited impact in the last two decades and unsatisfactory performance when compared to random screening approach, such approaches still play a useful role in drug discovery programs (22).We do believe that an *in silico* method imparts cost and time effectiveness given the limitations of the approach while harnessing its potential. Accordingly, a ligand-based *in silico* approach was used to identify phytochemicals as possible hits or me-too drugs in addition to scientifically evaluating the herbal medicines traditional use.

According to a study on traditional medicine database, CDRUG was able to identify 5278 compound out of which 75% were similar to approved anticancer drugs (Tanimoto coefficient $\geq$ 0.70, MACCS fingerprint), and the top 346 compounds it identified were identical to compounds with proven anticancer activity on 60 cell lines (23). Given such performance of CDRUG, our finding can be taken as a preliminary evidence of anticancer activity by many of the medicinal plants used for treating *'Gofla'*. Among the 13 species whose phytochemicals were analysed, 11(84.6%) contained one or more compounds predicted to

possess high possibility of anticancer activity. One or more phytoconstituents predicted to have a moderate possibility for bioactivity were also found in 12 (92.3%) of medicinal plants. The use of these plants in combinations in the traditional medicine practice may also enhance their activities by increasing both the amount and type of the predicted anticancer compounds in the preparations. Species such as *Punica granatum*, *Boswellia papyrifera*, *Viscum schimperi* and *Fagara usambarensis* contained higher number of compounds predicted to have anticancer activity in both categories. Nonetheless, these species should not be considered as having superior anticancer activity over the rest since the number of phytochemicals analysed for each of species varies.

Drug discovery is a challenging, time consuming and costly process. In small molecule drug development, the initial emphasis is largely on efficacy which makes addressing the issue of drugability secondary. This practice often leads to high failure rates and development costs (17). Hence, we also assessed the general and antineoplastic drugability of relevant phytochemicals in addition to a similarity based prediction of anticancer activity. The initial evaluation using Lipinski's rule of five determined that over ninety percent of the predicted compounds to be drug-like as they contained no or only oneviolation (24).

Since the Lipinski rule serves little to discriminate between drugs and non-drugs other than predicting oral bioavailability, we used additional rules that not only test the general drug-likness but also the antineoplastic drugability (24). The physicochemical properties used by *Ghoseet al*, are classified into two sets of ranges. The first one is the preferred range (50%) defined as the range within the qualifying range occupied by approximately 50% of the antineoplastic drugs in CMC database. Compounds with ALogP 0.0-3.7, AMR 60-107, MW 258-388 andnA 30-55 belong to this range. The fact that we found no compound of interest, which meet all the 50% criteria, may indicate the lack of novel chemical entities in our finding as this range is considered an efficient way

of searching for new drugs. The second one is the qualifying range which covers 80% of the drugs and provides the advantage of reducing the chance of missing anticancer drug-like compounds. Compounds with ALogP -1.5-4.7, AMR 43-128, MW 180-475 and nA 21- 63 are included to this range (25). Contrary to the preferred range, we found 25 chemicals meeting all the criteria of this range. This augments the result from CDRUG in proving the anticancer activity of the predicted phytochemicals.

Structural features are important criteria in the assessment of drug-likeness. The combined presence of certain structures like benzene ring, heterocyclic rings (both aliphatic and aromatic), keto, alcohol, aliphatic amine, carboxamide and carboxyester groups, may signal antineoplastic and other drug likenesses (15). Many of the bioactive natural products determined in this study contained all these features, which may be considered as an indication of good drugability in addition to efficacy. Aromatic rings either alone or combined with non-aromatic rings are relevant for antineoplastic activity, structural features possessed by over half of the predicted compounds. In addition, over two-third of identified phytochemicals possessed proportion of heavy atoms (*R- value)* between 0.05–0.50, which is thought to indicate good antineoplastic drug-likeness. However, the fact that only 3(4.3%) of the predicted compounds contained an amine (-NH_2) group, a functional group thought to be beneficial for anticancer activity, may undermine the prediction of being both efficacious and drugable(17). Generally, the evaluation based on important structural features found at least half of the predicted compounds to have good anticancer drugability**.**

Our investigation of both anticancer efficacy and drugability, which combined a hybrid score based similarity measure, rule of five, GVW-drug-likeness indices and structural features criteria, have determined a number of possible antineoplastic phytochemicals. These compounds have the potential for further study as possible hits or me-too drug sources of cancer treatment. Given the distribution of these phytochemicals, there is also a potential for many of the medicinal plants to

possess anticancer activity. As far as the traditional medicine practice is concerned, there is an *in silico* evidence to suggest anticancer efficacy of the traditional treatments. The efficacy of these treatments may be even higher due to the practice of combining the medicinal plants, which can increase both the amount and type of predicted anticancer compounds within preparations.

ACKNOWLEDGEMENTS

REFERENCES

1. Jemal A, Bray F, Forman D, et al. Cancer burden in Africa and opportunities for prevention. Cancer. 2012; 118(18):4372-4387.
2. Woldeamanuel YW, Girma B, Teklu AM. Cancer in Ethiopia. *Lancet Oncol* 2013; 14(4):289-290.
3. Rajandeep K, Suman K, Anil SK. Medicinal plants for the treatment of sexual transmitted diseases. *International Journal of Pharmaceutical Innovation* 2012; 2(3).
4. Gurib-Fakim A. Medicinal plants: Traditions of yesterday and drugs of tomorrow. *Mol Aspects Med* 2006; 27:1–93.
5. Woldegerima B. Review on the importance of documenting ethnopharmacological information on medicinal plants. *Afr J Pharm Pharmacol* 2009; 3(9):200-403.
6. Itharat A, Ooraikul B. Research on Thai medicinal plants for cancer treatment. In:Acharya SN, Thomas JE, editors. Advances in Medicinal Plant Research. Trivandrum: PUB CAB: 2007; 287-371.
7. Cragg GM, Newman DJ. Plants as a source of anticancer agents. *J Ethnopharmacol* 2005; 100:72-79.
8. Cox PA. The ethnobotanical approach to drug discovery: strengths and limitations. 1994 Ethnobotany and the search for new drugs. In: 185 CFS, editor. Ethnobotany and search for new drugs. Chichester: Wiley: 1994; 25-41.
9. Sawadogo WR, Schumacher M, Teiten M-H, Dicato M, Diederich M. Traditional West African pharmacopeia, plants and derived

compounds for cancer therapy. *Biochem Pharmacol* 2012; 84:1225–1240.

10. Cragg, GM, Boyd MR, Cardellina JH, Newrnan DJ, Snader KM, McCloud TG. Ethno botany and drug discovery: the experience of the US National Cancer Institute. In: 185 CFS, editor. Ethnobotany and search for new drugs. Chichester: Wiley: 1994; 178-176.

11. Farnsworth NR. Ethnopharmacology and drug development. . In: 185 CFS, editor. Ethnobotany and search for new drugs. Chichester: Wiley: 1994; 42-59.

12. Prada-Gracia D, Huerta-Yépez S, Moreno-Vargas LM. Application of computational methods for anticancer drug discovery, design, and optimization. *Bol Med Hosp Infant Mex* 2016; 73(6):411- 423.

13. Li G-H, Huang J-F. CDRUG: a web server for predicting anticancer activity of chemical compounds. *Bioinformatics* 2012; 28(24) 3334–3335.
doi:10.1093/bioinformatics/bts625

14. Lipinski CA. Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov Today Technol* 2004; 1(4):337-341. doi: 10.1016/j.ddtec.2004.11.007

15. Ghose AK, Viswanadhan VN, Wendoloski JJ.A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *J Comb Chem* 1999; 1:55-68.

16. Tetko IV, Gasteiger J, Todeschini R, et al. Virtual computational chemistry laboratory - design and description. *J Comput Aid Mol Des* 2005; 19:453-463.

17. Mao F, Ni W, Xu X, et al.Chemical structure-related drug-like criteria of global approved drugs. *Molecules* 2016; 21(75). doi:10.3390/molecules21010075

18. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev*, 2001; 46:3–26.

19. Walters WP, Murcko MA. Prediction of 'drug-likeness'. *Adv Drug Deliv Rev* 2002; 54:255–271.

20. Karou D, Nadembega WMC, Ouattara L, et al. African ethnopharmacology and new drug discovery. *Med Aromat Plant SciBiotechnol*2007; 1(1).

21. Balunasa MJ, Kinghorn AD. Drug discovery from medicinal plants. *Life Sci* 2005; 78:431 – 441.

22. Albuquerquea UP, Medeiros PM, Ramos MA, et al. Are ethnopharmacological surveys useful for the discovery and development of drugs from medicinal plants? *Rev Bras Farmacogn* 2014; 24:110-115. http://dx.doi.org/10.1016/j.bjp.2014.04.003

23. Dai S-X, Li W-X, Han F-F, et al. In silico identification of anti-cancer compounds and plants from traditional Chinese medicine database. *Sci Rep* 2016; 6:25462. doi: 10.1038/srep25462

24. Tian S, Wang J, Li Y, Li D, Xu L, Hou T.The application of in silico drug-likeness predictions in pharmaceutical research. *Adv Drug Deliv Rev* 2015. doi: 10.1016/j.addr.2015.01.009

25. Todeschini R, Consonniet V, eds. Molecular descriptors for chemoinformatics. 2nd ed. Weinheim: WILEY-VCH Verlag GmbH & Co. KGaA; 2009.