

Ethnobotanical approach versus random approach in the search for new bioactive compounds: support of a hypothesis

C. Gyllenhaal¹, M. R. Kadushin^{1,2}, B. Southavong³, K. Sydara³, S. Bouamanivong⁴, M. Xaiveu³, L. T. Xuan⁵, N. T. Hiep⁵, N. V. Hung⁵, P. K. Loc,⁵ L. X. Dac,⁵ T. Q. Bich⁶, N. M. Cuong⁶, H. M. Ly, H. J. Zhang¹, S. G. Franzblau¹, H. Xie,⁷ M. C. Riley¹, B. G. Elkington¹, H. T. Nguyen¹, D. P. Waller,¹ C. Y. Ma,⁸ P. Tamez,⁹ G. T. Tan,¹⁰ J. M. Pezzuto,¹⁰ D. D. Soejarto^{1,2,*}.

¹Department of Medicinal Chemistry and Pharmacognosy, College of Pharmacy, University of Illinois at Chicago, 833 S. Wood St., Chicago, IL 60612, ²Botany Department, Field Museum, 1400 S. Lake Shore Dr., Chicago, IL 60605, USA; ³Institute of Traditional Medicine, Ministry of Health, Vientiane, Lao People's Democratic Republic, ⁴National Herbarium of Laos, National Science Council, Vientiane, Lao People's Democratic Republic, ⁵Vietnam Academy of Science and Technology, Hanoi, Vietnam, ⁶Cuc Phuong National Park, Ninh Binh Province, Vietnam, ⁷Center for Clinical and Translational Sciences, University of Illinois at Chicago, Chicago IL 60612, ⁸U.S. Pharmacopoeia, 12601 Twinbrook Parkway Rockville, Maryland 20852, ⁹University of Notre Dame, Notre Dame IN 46556, and ¹⁰College of Pharmacy, University of Hawaii, Hilo, Hilo HI 96720.

*Corresponding author Djaja Doel Soejarto, dds@uic.edu.

¹Department of Medicinal Chemistry and Pharmacognosy
College of Pharmacy, University of Illinois at Chicago,
833 S. Wood Street, Room 539, mc 781
Chicago, IL 60612
Telephone 312-996-8889; Fax 312-996-7107.

Abstract

Context: Whether natural product drug discovery programs should rely on wild plants collected “randomly” from the natural environment, or whether they should also include plants collected on the basis of use in traditional medicine remains an open question.

Objective: This study analyzes whether plants with ethnomedical uses from Vietnam and Laos have a higher hit rate in bioassay testing than plants collected from a national park in Vietnam with the goal of maximizing taxonomic diversity (“random” collection).

Materials and Methods: All plants were extracted and subjected to bioassay in the same laboratories. Results of assays of *plant collections* and *plant parts* (samples) were scored as active or inactive based on whether any extracts had a positive result in a bioassay. Contingency tables were analyzed using chi-square statistics.

Results: Random collections had a higher hit rate than ethnomedical collections, but for samples, ethnomedical plants were more likely to be active. Ethnomedical collections and samples had higher hit rates for tuberculosis, while samples, but not collections, had a higher hit rate for malaria. Little evidence was found to support an advantage for ethnomedical plants in HIV, chemoprevention and cancer bioassays. Plants whose ethnomedical uses directly correlated to a bioassay did not have a significantly higher hit rate than random plants.

Discussion: Plants with ethnomedical uses generally had a higher rate of activity in some drug discovery bioassays, but the assays did not directly confirm specific uses.

Conclusion: Ethnomedical uses may contribute to a higher rate of activity in drug discovery screening.

Keywords: Traditional medicine, Drug discovery, Screening, Vietnam, Laos, Tuberculosis, Malaria, HIV, Cancer, Chemoprevention.

Introduction

One of the persistent questions in natural product drug discovery has been whether testing plants on the basis of their use for specific diseases in traditional medicine would lead to a higher rate of discovering active lead compounds than simply testing plants “randomly” from the natural environment, ie., opportunistically collecting whatever plants are available without regard to traditional use (Farnsworth, 1994). For example, Rios and Recio (2005) suggest that random selection of plants be avoided in the search for new antimicrobials, and that plants collected on the basis of ethnopharmacological information be emphasized. Accordingly, Svetaz et al. (2010) assayed plant species with a history of medicinal uses from Latin America for bioactivity against fungal pathogens, and discovered that plants with ethnomedical (e.g. traditional medicine) uses were significantly more likely to inhibit growth of pathogenic fungi than randomly collected plants (40% vs 21%). Coelho de Souza et al. (2004), however, found little correlation of traditional use and activity of plants in antibacterial assays. Cragg et al. (2009) list some anticancer drugs and lead compounds discovered based upon traditional medicine use. Spjut (2005) found that plants used as anthelmintics and arrow

poisons had a higher percentage activity than plants assessed randomly in the National Cancer Institute records of plants screened against cancer (cancer cell cytotoxicity is thought to be one of the areas poorly targeted by ethnomedical use because of the poor definition of cancer signs, symptoms and disease concepts in traditional medicine). If ethnomedical information is significantly correlated with activity against such drug development targets, it is hypothesized that the number of plants that would need to be collected and evaluated in order to discover a successful natural product drug could be significantly reduced. This would lead to financial savings and efficiencies in research. Further data on this question are thus needed.

The Vietnam-Laos International Cooperative Biodiversity Group (ICBG) project “Studies of Biodiversity of Vietnam and Laos” has operated since 1998 with funding from the Fogarty International Center of the U.S. National Institutes of Health. It has been described in detail elsewhere (Soejarto et al. 1999, 2002, 2004a, 2006, 2007, 2009). A specific aim of the project is the development of new natural product drugs from the biodiversity of Vietnam and Laos. Major collaborators include the Vietnam Academy of Science and Technology, Cuc Phuong National Park in Vietnam, and the Institute of Traditional Medicine in Laos. In the project, plant materials have been collected from locations in Vietnam and Laos, and subjected to bioassay-guided fractionation to isolate active lead compounds. The project included both plants collected based on ethnomedical information in Vietnam and Laos and plants collected on a “random” or biodiversity-maximized basis from the tropical forest environment of Cuc Phuong National Park in Ninh Binh Province of Vietnam. The availability of data on both ethnomedical and

random plant collections made by a single project and analyzed using the same bioassay techniques gives us the opportunity to test the hypothesis that plants collected on the basis of ethnomedical use will result in a higher hit rate, or percentage of plants active in bioassays, than plants collected randomly. It also allows us to assess the corollary proposition that drug development assays will confirm the supposed clinical activity of the traditional medicine plants.

Methods

Ethnobotany-guided plant collections

To collect plants with ethnomedical information, interviews with traditional healers in Laos and Vietnam were conducted, and the plants they listed as being used for different diseases were collected from nearby field sites. Interviews were conducted chiefly by project staff at the Vietnam and Lao collaborating institutions, who received training based on questionnaires and interview methods developed by UIC project coordination staff.

The plants in Laos were collected in collaboration with the Institute of Traditional Medicine (ITM, formerly the Traditional Medicine Research Center) in Vientiane. The Laos collection process is described in detail elsewhere in this issue (Soejarto et al. 20XX). The ITM maintains a network of Traditional Medicine Stations throughout Laos (Soejarto et al., 1999, 2009; Riley 2001). The Stations typically work with noted local

healers and maintain active communication between the healers and the Institute. Project interviews were conducted with these healers, other nearby healers and local residents, with the coordination of Traditional Medicine Station staff. The project collected plants throughout Laos, in a variety of different environments and landscapes. Medicinal plants in Laos are obtained from forested areas near villages, as well as from gardens and cultivated plants.

The ethnomedical collections in Vietnam were based on interviews with traditional healers and other community members in Cuc Phuong Commune, a village located near Cuc Phuong National Park, inhabited by members of the Muong minority population who had previously lived in the area of the Park. Medicinal plants used by commune members are primarily collected in the forested buffer zone surrounding Cuc Phuong National Park, with some also being harvested from gardens and fields. The interviews were conducted by staff of the Vietnamese Academy of Science and Technology (VAST).

Interviews were conducted under UIC Institutional Review Board protocols # 1997-1056 and # 2003-0636, #2003-0217, 2004-0197 and # 2007-0369. Prior informed consent was sought at ministerial, provincial and local levels for the project, and individual informed consent was obtained before commencing interviews. The project also is governed by a comprehensive benefit-sharing agreement in order to comply with the Convention on Biodiversity. For most plants, multiple uses were recorded. In addition to recording traditional medicine uses of each plant species, project staff

collected both herbarium specimens used to identify the plant species and serve as a permanent record of the plants collected, together with bulk samples of plant material (e.g. 0.5 kg dry weight) to be used for extraction and bioassay testing. In most cases, only the plant parts mentioned as being useful by the healers were collected.

Biodiversity-based [“random”] collection

Plants for the “random” or biodiversity-maximized collection were obtained in Cuc Phuong National Park, in collaboration with Park scientific staff. Cuc Phuong National Park was inaugurated in 1962. It is located in a limestone karst area of the Annamite Mountains, and has a nearly 600 m elevational range. More than 2000 vascular plant species are found in the park (Soejarto et al. 2004b), reflecting the large number of available environments. Several areas in the park were inhabited by Muong minority communities until the late 1980s, but following government-sponsored relocation program in the 1990’s, all Muong households are now settled in areas surrounding the Park, and the areas impacted by them is minor. The goal of the ICBG collection was to obtain as high a biodiversity of plant species as possible. Thus, in the field collection process, plant species that had already been collected or for which we had already collected multiple species in the same genus were avoided. This collection is referred to in the current paper as a random collection since it was gathered without regard to traditional medicine use; it does not constitute a true random sample of Park vegetation, but is nonetheless consistent with typical collection practices in the drug discovery field.

In the field collection process, both herbarium specimens (and bulk samples of plant parts were collected from a single individual plant. Typically multiple plant parts

were collected as bulk samples for each species. For both the ethnomedical and the random collections, herbarium specimens were deposited at the herbaria of Cuc Phuong National Park (CPNP) and the Institute of Ecology and Biological Resources of the Vietnam Academy of Science and Technology in Hanoi (HN) in Vietnam, and at the herbarium of the Institute of Traditional Medicine in Vientiane in Laos, as well as at the Field Museum (F) in Chicago. Taxonomic identification was performed at these institutions, as well as with the assistance of taxonomic specialists in some cases. For the present analysis, only plant collections that were identified to species level, or at least to genus, were included in order to avoid undetected duplication of species collected due to unidentified specimens.

Post-collection processing and data management

Following the field interviews and collections, all bulk plant samples were dried and extracted according to ICBG project's extraction protocols. They were then subjected to a variety of bioassays at the UIC and VAST laboratories of the ICBG program, as well as at the Pasteur Institute in Hanoi (antimycobacterial assay), Purdue University (cancer chemoprevention assay). Each sample was tested in multiple bioassays for an average of approximately 7 bioassays per plant sample. Based on these bioassays, bioactive compounds have been isolated and their chemical structures and modes of action have been published (Soejarto et al., 20xx, this issue) or are currently being elucidated. Assays included 6 cancer chemoprevention assays, 8 cancer treatment assays using well-known cancer cell lines, 2 antiplasmodial (malaria) assays, an HIV

assay and an antimycobacterial (tuberculosis) assay. Data from these assays, along with the botanical and ethnomedical data from fieldwork, were entered into the ICBG program's data management system, NAPIS® Laboratory (Natural Products Information System; <http://www.wps2.com>).

The final type of accession was bioassay test results. In addition to quantitative results, tests results were assigned qualitative descriptors of active (A), questionable (Q) or inactive (I) based on standard cutoff points established for each bioassay. Qualitative and quantitative data were entered into NAPIS. The data entered into NAPIS thus included plant collection information, plant sample tracking number, plant identification, ethnomedical use information, bioassay results and compound isolation results. Data handling for this analysis was conducted using queries from NAPIS Laboratory and processing the downloaded datasets in Microsoft Access and Excel. Table 1 shows the project accessions based on data in NAPIS.

Plant collections and plant samples

Accessions were analyzed at the level of collections (a single plant species collected in one locality at one time) and samples (a single plant part or a set of pooled parts from a collection; collections could have more than one plant part sampled, e.g. fruits, seeds, leaves plus stems, or other parts). Figure 1 shows graphically the difference between plant collections and plant samples, as well as the eventual destination of the voucher specimens that represent each collection, and the bulk samples for laboratory analysis.

We distinguish collections versus samples to help clarify the different aspects of effort in obtaining and processing plant materials for drug discovery screening projects. In the process of *collection*, whether ethnomedical or random, two types materials were gathered, and prepared from one single individual plant: a) **voucher herbarium specimens** (a series of 3-7 specimens) intended for taxonomic identification and other botanical studies; b) one or more **bulk samples** each consisting of one plant part or a combination of parts, f.e., leaf plus twigs, stem bark, or root, in the amount of approximately 1000 g fresh weight (\approx 400 g dry), intended for bioassay and chemical analysis. The two plant materials are linked by a common identifier, normally the name and number of the collector. The destination of voucher herbarium specimens is a Herbarium institution, where taxonomic identification is performed, whereas the destination of the bulk samples is the laboratory, where the samples are extracted, then the extract submitted to biological assay, which may be followed by bioassay-guided fractionation, in the event of active extracts. ~~involved the location of an appropriate individual of a plant species, obtaining ethnomedical information if necessary, collecting voucher specimens and procuring botanical identification. On the other hand, each sample represented a separate effort in the bulk collection, extraction, bioassay test(s) and, where appropriate, fractionation. Thus, the collections were considered to relate more to the effort of field acquisition whereas the samples relate more to the effort of laboratory analyses.~~

[Figure 1 about here]

Accessions were further analyzed by plant uses. The basic data point for a plant use stored in NAPIS is an excerpt from the field interview, usually a description of a

disease or symptom for which the plant is used, such as cough, malaria, stomach-ache, diabetes or other condition. Because of the diversity of such descriptions, for this analysis we assigned a further data element, ethnomedical use category. These categories were assigned based on groups of use descriptions that were potentially relevant to the bioassay targets of the overall project. Twelve use categories were developed: bacterial, cancer, cardiovascular, central nervous system (CNS), infection, inflammation, malaria, miscellaneous, parasite, tuberculosis (TB) and viral. Uses describing a symptom or disease that characterizes one of these use categories were assigned to that category. Thus, malaria, fever and chronic fever are assigned to the malaria ethnomedical use category. Tuberculosis, cough, fatigue and lung disease are assigned to the tuberculosis ethnomedical use category. Hepatitis, dengue fever, measles and other viral diseases are assigned to the viral category. Diabetes, stomach-aches, folk illnesses and a number of other common complaints that did not appear to fit well into mechanisms relevant to the project's bioassay targets were assigned to the miscellaneous category.

The ethnomedical use categories were related to the bioassay targets of the study on a mechanistic basis. The malaria use category could be related to antiplasmodial assays. The bacterial, parasitic and TB categories could be related to the antimycobacterial assays. The viral category, in contrast to the TB and malaria categories, included multiple common viral diseases. This was thought to be related to HIV assays to the extent that plants used for viral diseases might show some general antiviral activity. There were very few attributions of use for HIV specifically by our informants. Cancer prevention could be related to the inflammation and cancer categories; no concept of cancer prevention was encountered among our informants, but

inflammation is thought to be strongly related to carcinogenesis (Tan et al. 2011). The relationship of plants in the cancer category is also questionable since cancer is not a common disease in rural populations of Laos and Vietnam.

Statistical analysis

Data were analyzed using contingency tables and percentage calculations of hit rate or positive activity based on qualitative bioassay results of “A”. Statistical significance was tested for the contingency tables, using chi square analysis, or, for tables with small numbers of samples, Fisher’s exact test (the latter does not yield a chi-square statistic as will be seen in the tables below). We used an alpha of $p=0.05$. However, because we performed multiple statistical tests, we used the Bonferroni correction for multiple comparisons, resulting in a significant p-value for this study of .00135.

For the initial analyses we grouped bioassay results at the levels of *collections* (a species collected on a certain date from a particular locality) and *samples* (plant parts) to determine if ethnomedical collections or samples had a higher hit rate (percentage “A” qualitative result in bioassay) than random samples or collections. As a followup assessment, we tested whether the plants with an ethnomedical use category that is directly related to the assay target had a higher hit rate than plants in other ethnomedical use categories or than random accessions. In investigating the results from this assay, we also plotted proportional frequency graphs for different groups of ethnomedical plant

accessions (Laos vs Vietnam) and different bioassays, to see if the distribution of use categories differed.

Bioassay results were grouped as follows: for collections level analyses, a single qualitative was assigned to a collection based on the best qualitative result for any sample or test arising from that collection, with active ranked as the best, questionable in the middle, and inactive as the worst result (A>Q>I). Thus, in an assessment of any single bioassay or group of bioassays, if any sample tested for one collection has a qualitative result of A, that collection is assigned the value of A for that bioassay. If none was active but one or more were questionable, the collection is assigned a value of Q. If no samples were active or questionable, the collection is assigned a value of I. A similar procedure is used to assign a qualitative result to a sample when sample groupings are tested.

For tuberculosis and HIV we used only a single bioassay each, so qualitative results refer only to these test results. For the target diseases in which we used multiple bioassays (antiplasmodial, cancer chemoprevention and cancer treatment), we assessed activity based on both individual bioassays and the entire group of bioassays for the target conditions.

[Table 2 about here]

A sample complete contingency table is shown in Table 2. This table shows the analysis by collections for tuberculosis, for which we used the bioassay MABA-H37RA.

In this table, A and I are the qualitative results (no “questionable” cutoff was established for this bioassay). The collection accession and rationale (ethnomedical or random) are shown in the first column. This table analyzes data at the level of collections. The number of collections with A qualitative results and I qualitative results are shown in the main body of the table, and in the totals for the rows and columns of the table. The hit rates or percent activity are calculated as the number of active plants for a given collection accession divided by the row total for that collection accession. The chi-square, degrees of freedom and p-value for the contingency table were calculated using an interactive online calculator (Preacher 2001). The p-value for this table is significant (.00134). Note that the hit rates or percent activities are shown not for calculation of statistical significance but for purposes of easy heuristic assessment of the study hypothesis. For the tables in the results section, we show only the hit rates, total number of samples or collections in the contingency table, and the chi-square or Fisher’s exact test for the contingency table, rather than exhibiting all contingency tables. Degrees of freedom for the contingency tables vary because some assays lacked questionable results, and because in some instances two accession types (i.e. all ethnomedical versus all random) and in some three accession types (i.e. Vietnam samples versus Laos samples versus random samples) were tested.

Results

The results of the overall assessment of whether all ethnomedical collections and samples had a higher hit rate than all random collections and samples are summarized in Table 3.

[Table 3 about here]

The results of the contingency table analysis and the hit rates indicate a contrasting situation between the collections analysis and the samples analysis. For both collections and samples the results are highly significant. However, in the case of collections, the random collections have a somewhat higher hit rate (31% vs 28.5%) whereas for the samples, the ethnomedical samples have a higher hit rate (28.1% vs 22.2%).

[Table 4 about here]

Table 4 shows the results for all five bioassay targets, grouped according to collections for Laos ethnomedical, Vietnam ethnomedical and random accessions. Note in this table that the only ethnomedical collections tested in the HIV assay that have been included in the NAPIS database are those from Laos, while the only collections tested in the chemoprevention assay are from Vietnam. The chi-square values are significant for all targets except cancer chemoprevention. However, of the hit rates associated with significant chi-square values, only tuberculosis shows a higher hit rate for collections with ethnomedical rationale than random collections. For the malaria, HIV and cancer treatment, the random collections had a higher hit rate than the ethnomedical collections.

The hit rate for the Vietnam ethnomedical collections appears to be lower than that for the Laos collections in tuberculosis and cancer treatment bioassays.

[Table 5 about here]

Table 5 shows the results for all five bioassay targets, grouped according to samples for Laos ethnomedical, Vietnam ethnomedical and random accessions. The targets with significant chi-square tests are antimycobacterial, HIV and cancer treatment. Antiplasmodial activity and cancer chemoprevention are non-significant. In the case of antiplasmodial assays, the hit rates of the samples collected with different rationales are quite similar, though this is not the case with cancer chemoprevention. In the case of the samples, the bioassay results indicate that for antimycobacterial and cancer the hit rates favor ethnomedical samples, specifically the samples from Laos. For HIV the hit rates favor random samples, though the difference is quite small.

[Table 6 about here]

Table 6 shows the results for all five bioassay targets, grouped according to samples for the Laos ethnomedical, Vietnam ethnomedical and random accessions. In this table the results for different tests are presented separately. For antimycobacterial and HIV assays, because there was only one bioassay for each of these targets, the results are the same as in Table 5. For antiplasmodial activity, the results of the two assays, D6 and W2, are notably different. For D6, the chi-square is non-significant, and the hit rates are

similar for all accession types. For W2, however, the chi-square is significant, and the hit rate is much higher than for the D6 assay. D6 was used as the primary screen for these samples, while W2 was used primarily but not exclusively as a confirmatory screen, so that samples that had been active in the D6 assay were submitted to the W2 assay as a secondary screen. Additionally, W2 has been observed to be more sensitive than D6 (Yuan et al. 2009). Another feature of the W2 assay is that the hit rate of the ethnomedical samples (75-80%) is much higher than the hit rate of the random samples (45%). In the case of cancer chemoprevention, none of the contingency tables was significant, while in the case of cancer treatment, only HL-60 and LNCaP were significant, but the percentage differences in hit rates were quite small. In tuberculosis and cancer assays, we again see the plants of Vietnam with a lower hit rate than those of Laos or the random plants.

[Table 7 about here]

Table 7 shows the results of assessing plant collections with usage categories most related to the different bioassay targets against plants with other ethnomedical uses and randomly collected plants, to see if the ethnomedical uses accurately predict bioassay results. For all the disease targets with more than one bioassay, i.e. antiplasmodial, cancer chemoprevention and cancer treatment, all assays are included, similar to Tables 3-5. The ethnomedical uses were assigned as follows: antimycobacterial bioassay + tuberculosis plus infection use categories; antiplasmodial bioassay + malaria and infection use categories; HIV bioassay + viral use category; cancer chemoprevention and

cancer treatment + cancer and inflammation use categories. In none of these cases do the plants with ethnomedical use categories have a higher hit rate than random plants. In three cases the contingency table is significant, but in favor of the random accessions.

As presented above, these results raise some questions about the nature of the relationship between ethnomedical use and bioassay results. An initial question is why the Vietnam samples have a lower hit rate in the antimycobacterial and cancer bioassays. One possible reason for the difference in the antimycobacterial assays is that they were performed in two different laboratories, one at UIC and one in Hanoi. While there are other possible explanations that are somewhat speculative, an explanation that could be addressed using our dataset is whether the types of ethnomedical uses for plants collected in Vietnam and Laos were similar or different.

[Figure 2 about here]

Fig. 2 shows the frequencies of ethnomedical use categories for the plants collected from Laos, those from Vietnam and those in the total collection (both countries). It is clear from this graph that the uses of plants collected in Vietnam and Laos were different. In Vietnam, larger frequencies of plants in the categories of inflammation and miscellaneous were collected. In Laos, larger frequencies of plants in the categories of bacterial, cancer, central nervous system, malaria, and tuberculosis were collected. These differences in plant uses might underlie the lower hit rates for tuberculosis and cancer in the above analyses, especially if the plants collected

differentially in Vietnam versus Laos have some relation to the ethnomedical use categories that were most active in our bioassays.

[Figure 3 about here]

To explore the question of the *most active use categories*, we created similar frequency graphs, shown in Fig. 3, for each of the project's disease targets. These graphs show the frequencies of ethnomedical use categories for the collections that were found to be active in each of the assay systems, charted alongside the frequencies of ethnomedical use categories in the total collection (the same data as for the total collection as shown in Fig. 2). These charts may suggest which ethnomedical use categories were actually the most likely to be active in the five bioassay targets. Inspecting Fig. 3A, it can be observed that while the frequency of active plants in the use category of tuberculosis (n=33) was higher than the frequency of plants used for tuberculosis in the entire collection, the use category that was by far the most likely to be active was central nervous system plants, although this sample size is rather small. The active central nervous system plants had ethnomedical uses of nervine, tonic and pain. Among the plants active in the malaria bioassays (Fig. 3B, n=112), most use categories, including malaria, occurred at about the same frequency as the total collection, with the exception of plants used for viral disease and to a smaller extent, central nervous system uses, which were more frequent than would be expected in the total collection. In the plants active in the HIV bioassay (Fig. 3C, n=19), plants used for malaria and cancer occurred more frequently than in the total collection, although this sample size was also

rather small. Among plants active in the chemoprevention bioassays (Fig. 3D, n=246), plants in the cardiovascular, miscellaneous, parasite and viral groups occurred more frequently than in the total collection. Among plants active in the cancer treatment bioassays (Fig. 3E, n=284), plants in the cancer, malaria and miscellaneous groups occurred more frequently than in the total collection.

Discussion

These results cast some light on the questions of whether testing of traditional medicine plants may result in a higher hit rate in drug discovery programs than testing plants collected without regard to traditional medicine use, and whether drug discovery bioassays may support the supposed clinical uses of traditional medicine plants. Table 3 showed contradictory results. When analyzed according to collections, random plants were about 3 percentage points more likely to result in at least one active bioassay among all the test results. When analyzed according to samples (plant parts), the ethnomedical plants were about 6 percentage points more likely to result in at least one active bioassay. This could simply reflect the situation that random plants had more plant parts collected than ethnomedical samples. With more samples, and thus more test results, contributing to the possibility of a single plant collection receiving an “active” designation, the random collections likely had a greater probability of being active than the ethnomedical collections, which usually had only one plant part tested. Thus, it is likely that the ethnomedical collections actually had a higher hit rate on a per sample basis than the randomly collected species. But how did the ethnomedical uses actually correlate with

the bioassay results? To address this question each of the main disease targets is discussed separately.

Tuberculosis

In the case of tuberculosis, both the analysis by collections and the analysis by samples had higher hit rates than the random accessions. This indicates that ethnomedical sampling did accurately predict antimycobacterial activity. However, Table 7 and Fig. 3A suggest that the ethnomedical use of plants for tuberculosis did not predict positive bioassay results that well. Among the plants that were in the ethnomedical use category of “tuberculosis,” only five were actually active in the MABA-H37RA bioassay; the specific uses attributed to them were cough in 4 cases and fatigue in one case. The ethnomedical category that was clearly the most active in this bioassay was “central nervous system,” composed of uses such as nervine, tonic and pain. For the most part, these central nervous system species did not have other uses attributed to them, i.e., they were not additionally used for tuberculosis, infections or other conditions. The elevated rate of activity among these central nervous system plants could simply reflect a random variability in this rather small sample, and be unrelated to the use of the plants. One could also speculate that such plants were being used as tonics to address the fatigue common among tuberculosis patients, but actually or additionally were acting against the tuberculosis bacillus.

Malaria

For the antiplasmodial assays, the random plants clearly had a higher hit rate when assessed at the level of collections (species) but not when addressed at the level of samples. In the assessment of the W6 confirmatory bioassay, the ethnomedical plants had a significantly and obviously higher hit rate than the random plants. This suggests that the antiplasmodial activity was more robust in the ethnomedical plants than in the random plants, and that they might thus be better sources of active compounds. This may suggest confirmation of the ethnomedical application of these plants. Plants classified in the ethnomedical use category of malaria (used for malaria, fever, etc), however, were not over-represented among active plants in the antiplasmodial assays (Fig. 3B) – plants used for viruses and central nervous system uses were more likely to be active.

HIV

In the case of HIV bioassays, random accessions were more likely than ethnomedical accessions to be active both at the collections level and at the samples level. The plants that were most likely to be active in this assay were used for cancer, malaria and parasites (Fig. 3C); however, the number of ethnomedical samples that were active against HIV is quite small (n=19) and this may represent a random fluctuation. A complication in relating plants in the viral ethnomedical use category to anti-HIV bioassays is the fact that HIV is a retrovirus, unlike the causative agents of the common viral diseases that were most often encountered in our ethnomedical data.

Cancer chemoprevention

In the case of the cancer chemoprevention bioassays, the contingency tables were not significant for either collections or samples. The plants that were most likely to be active in this bioassay were those used for cardiovascular, miscellaneous, parasite and viral use categories. This was a large sample, but the differences between the uses of the active accessions and the total collection were not that marked in any of these instances, making discussion of possible mechanisms rather speculative.

Cancer treatment

In the case of cancer treatment bioassays (cancer cell culture lines), the contingency tables were significant for both the collections and samples. In both cases, the hit rates favored the random plants over the Laos ethnomedical plants, but part of the significance of the tables may have been due to the low hit rate of the Vietnam ethnomedical plants. The categories most likely to have been active in this assay were cancer, malaria and miscellaneous plants. Thus, there may be some rationale for collecting plants traditionally used for cancer to test against cancer cell lines, but, like Spjut (2005), our data suggest that other uses may be more likely to give positive results.

Laos versus Vietnam results

The question of why the ethnomedical plants of Laos had a higher recorded hit rate than the plants of Vietnam in the tuberculosis and cancer treatment bioassays does

appear to be most likely attributable to the different plant uses predominant in the two collections. The plants that were active in this project's bioassays were most likely to be in the ethnomedical use categories of cancer, central nervous system, malaria, miscellaneous, tuberculosis and viral. The Laos ethnomedical collection had a higher frequency of accessions in each of these use categories than the Vietnam collection, with the exception of the miscellaneous and viral groups. The Laos collectors emphasized the main target diseases of the project to a somewhat greater extent than the Vietnam group in their interviews (central nervous system assays were also part of our bioassay targets for a short time). They may also have been more successful in obtaining targeted collections because they interviewed more prominent healers over a wider geographic area, with the collaboration of the Traditional Medicine Stations. For the diversity of ethnomedical uses of plants collected in Laos, see Soejarto et al. (20xx), this issue. The Vietnam samples were collected from residents of one small area, in which there may have been less access to specialized healers, resulting in a more generalized collection.

Support for the ethnomedical hypothesis

As mentioned at the start of this discussion, the analysis of all ethnomedical samples versus all random samples in all bioassays ultimately suggests that ethnomedical samples are indeed more likely to have positive results in drug discovery bioassays. The higher hit rate among all random collections may be attributable at least in part to the larger number of samples analyzed for each random collection than for the ethnomedical collections. This level of support does not appear to extend, however, to the corollary

hypothesis that drug discovery bioassays will tend to affirm the mechanisms of specific ethnomedical uses of plants. A higher hit rate in the W6 confirmatory bioassay for malaria does suggest that the ethnomedical plants do have a more antimalarial activity than the randomly collected plants. In only two instances, however, tuberculosis and cancer treatment, were the plants in a specific ethnomedical use category more likely to be among the active species in the bioassays with which they were most obviously correlated. Other ethnomedical use categories that were over-represented among active plants were central nervous system, malaria, viral and miscellaneous. These, however, were not correlated mechanistically with the assays that corresponded to them. The higher hit rate of the Laos collection versus the Vietnam collection in antimycobacterial and cancer bioassays, and the higher frequencies of cancer, central nervous system, malaria and tuberculosis in the Laos collection, also suggest that general ethnomedical collections may result in an elevated percentage of plants that are biologically active in drug discovery bioassays, though not necessarily directly correlated with the ethnomedical uses of the plants. This pattern was also noted by Spjut (2005) in the National Cancer Institute collection.

Another means of evaluating whether plants collected for ethnomedical purposes may contribute to natural product drug discovery programs is to determine whether, in fact, any ethnomedical plants proceeded further along the path towards clinical testing than simply giving positive results in bioassays. Our group has published articles on active compounds from 13 species to date (He et al. 2005, 2006; Jutiviboonsuk et al. 2005; Ma et al. 2005, 2006, 2008; Nguyen et al. 2004; Truong et al. 2011 ; Zhang et al.

2002, 2003a, 2003b, 2004, 2006). Of these, 5 are from plants with ethnomedical uses. The correlation of active compound activity with ethnomedical use in these papers was not straightforward?. For instance, the species *Rourea minor* (Gaertn.) Alston (Connaraceae) is used for dengue fever in Laos; the publication by He et al. (2006) presents antiplasmodial compounds isolated from this species. An interesting correlation to be noted is the observation in this paper that antiviral plants were over-represented among the plants active against malaria. *Asparagus cochinchinensis* (Lour.) Merr. (Asparagaceae) was recorded to be used in Laos for chronic fever, although uses for fever, cough, kidney disease and benign breast tumors have been recorded from China. Chronic fever is a symptom that could relate to multiple diseases such as infections (including HIV), cancers, autoimmune diseases, inflammation and other conditions. The cytotoxic compounds isolated from *A. cochinchinensis* proved to have poor antiviral activity against HIV but moderate activity against cancer cells. Other instances from Laos are recounted in this issue by Soejarto et al. (20XX). These are thus cases of plants with ethnomedical uses that contain active compounds isolated on the basis of bioassay results that have questionable or no relevance to their ethnomedical uses. Gertsch (2009) notes the lack of successful ethnobotanically based drug discovery in recent years. Our data do not contradict his observations, and in fact suggest that the ethnomedical hit rate for actual drug discovery lead compounds is less impressive than would be desired.

The choice and decision to pursue further testing of active compounds in drug discovery programs are complicated, and depend on a number of factors including the novelty of the compound and the record of other compounds in the same class in more

advanced clinical testing, among several others. Traditional medicine plants may not fit well in this framework in many instances. Compounds isolated from traditional medicine plants are often not novel. The problem of “frequent hitters” (Gertsch 2009), phytochemicals such as quercetin, tannins and curcumin which have multiple bioactivities, may complicate interpretation of bioassay results or possibly even impede the detection of drug leads. The activity of a medicinal plant may arise through synergistic mechanisms of several compounds, none of which is sufficiently active to merit attention as a lead compound for drug development. Healers may also use medicinal plants for mechanisms not contemplated in specific industrially-oriented drug discovery programs, such as immune support and symptom relief. That a drug discovery program does not lend much support to the ability of such plants to affect diseases does not mean that they are not relevant in their actual treatment context. The clinical relevance of ethnomedical plants would ,however, need to be determined by specifically designed preclinical and clinical studies.

Study limitations

A major limitation of this study is that the quantitative assessment of the contribution of ethnomedical samples is limited to their activity in initial bioassay screening for a drug discovery program. An alternative assessment on the basis of chemical compounds isolated, which we have presented in brief immediately above, could prove useful in investigating whether traditional medicine plants are more likely to yield active compounds in drug discovery programs. The “random” collections in this

analysis, while typical of drug discovery collections based on acquisition of wild plants, are not a truly random sample of local vegetation, and there may be an inconsistency in comparing plants of Laos with wild vegetation of Vietnam. We are still completing bioassay and phytochemistry work on some of the samples for the ICBG program, and these in-progress data are not included in this analysis. In some cases the sample sizes of specific analyses are rather small, and we did perform a large number of analyses on this dataset. However, the dataset itself is quite large (over 24 000 bioassay results form the basis of the calculations), and we have addressed the specific statistical concerns for multiple comparisons and contingency tables with small expected cell sizes. Nonetheless, some of the results based on small sample sizes should be taken as tentative. The assignment of symptoms and diseases reported in interviews to ethnomedical use categories may have been subject to inaccuracy, lack of clarity or misunderstanding of culture-bound concepts of disease, leading to uncertainty in the circumscriptions of these categories.

Further, this study cannot address the question of whether ethnomedical plants are actually selected by healers based on experience with their bioactivity, or whether other factors influence their selection. Moerman et al. (1999) note that ethnic groups in the Northern Hemisphere that are widely separated by a history of migrations, but are nevertheless culturally and genetically related, have strikingly similar medicinal floras. This suggests that cultural influences may affect the selection of herbs, as well as direct experience with their bioactivity. Leonti (2011) further notes that written texts that are based on a variety of cultural meanings unrelated to biochemistry (e.g. humors, doctrine of signatures, doshas) may influence selection of medicinal plants. Certainly medicinal

plant texts that date back a minimum of a few hundred years are known from South Asia, such as the palm-leaf manuscripts that were produced in Laotian monasteries (Elkington et al. 2009). Cultural processes may thus result in the medicinal use of plants for reasons other than their pharmacological properties.

Conclusions

This study is based on analysis of plants collected on an ethnomedical basis from Vietnam and Laos in comparison with plants collected on a “random” or biodiversity-maximized basis from Cuc Phong National Park in Vietnam, and subjected to a set of bioassays carried out by the same laboratory team as part of an International Cooperative Biodiversity Group collaboration. The overall analysis suggests that plants collected based on ethnomedical use may in fact have a somewhat higher rate of positive bioassays on a per-species or per-sample basis, although a portion of these assay results may be due to ubiquitous bioactive compounds. Ethnomedical collections in general may nevertheless play a useful role in drug discovery programs due to this elevated rate of bioactivity.

We noted that the published compounds from ethnomedical plants arising from this project did not all have activities that coincided with the uses of the plants. The correlation of specific ethnomedical use categories with positive activity in bioassays that should in theory be related to those uses was far from perfect. In general, plants used to treat cancer, central nervous system conditions, malaria, parasites, tuberculosis, viral disease and miscellaneous conditions were among those likely to be active against the

project's main disease targets. The lack of confirmation of ethnomedical uses by bioassay results suggests that traditional medicine plants may have an elevated rate of bioactivity, but that, if they are clinically effective in their original context, their effects may be due to mechanisms not included among our target bioassays.

Declaration of Interest

The authors have no declarations of interest.

Acknowledgements

We wish to express our deepest thanks to all healers and community members in Vietnam and Laos who have graciously collaborated with us in the present ICBG-sponsored research for the period of 1999 through 2008. We also thank the governments of Vietnam and Laos for their support and auspices in the planning and implementation of this research, and for the issuance of permits for plants collections. Thanks are expressed to all members of the Vietnam-Laos ICBG consortium, 1998-2011, for their generous professional collaboration in making this ICBG a successful endeavor. The Vietnam-Laos ICBG is supported by grants administered by the Fogarty International Center, U.S. National Institutes of Health (NIH), with funding supplied by NIH, National Science Foundation and the United States Department of Agriculture, as follows: 1 UO1 TW001015-01; 2 UO1 TW001015-06; 3 UO1 TW 001015-10S1; 3 UO1 TW001015-10S2; 2 UO1 TW001015-11A1.

References

Cragg GM, Grothaus PG, Newman DJ. (2009). Impact of natural products on developing new anticancer agents. *Chem Rev*, 109, 3012-3043.

[Elkington BG](#), [Southavong B](#), [Sydara K](#), [Souliya O](#), [Vanthanouvong M](#), [Nettavong K](#), [Thammachack B](#), [Pak DH](#), [Riley MC](#), [Franzblau SG](#), [Soejarto DD](#). (2009). Biological evaluation of plants of Laos used in the treatment of tuberculosis in Lao traditional medicine. *Pharm Biol*, 47, 26-33.

Farnsworth NR. (1994). Ethnopharmacology and drug development. *Ciba Found Symp*. 1994, 185, 42-51.

Gertsch J. (2009) How scientific is the science in ethnopharmacology? Historical perspectives and epistemological problems. *J Ethnopharmacol*, 122, 177-183.

He ZD, Ma CY, Zhang HJ, Tan GT, Tamez P, Sydara K, Bouamanivong S, Southavong B, Soejarto DD, Pezzuto JM, Fong HH. (2005). Antimalarial constituents from *Nauclea orientalis* (L.) L. *Chem Biodivers*, 2, 1378-1386.

He ZD, Ma CY, Tan GT, Sydara K, Tamez P, Southavong B, Bouamanivong S, Soejarto DD, Pezzuto JM, Fong HH, Zhang HJ. (2006). Rourinoside and rouremin, antimalarial constituents from *Rourea minor*. *Phytochemistry*, 67, 1378-1384.

Jutiviboonsuk A, Zhang H, Tan GT, Ma C, Van Hung N, Manh Cuong N, Bunyapraphatsara N, Soejarto DD, Fong HH. (2005). Bioactive constituents from roots of *Bursera tonkinensis*. *Phytochemistry*, 66, 2745-2751.

Leonti M. (2011). The future is written: Impact of scripts on the cognition, selection, knowledge and transmission of medicinal plant use and its implications for ethnobotany and ethnopharmacology. *J Ethnopharmacol*, 134, 542-55.

Ma C, Case RJ, Wang Y, Zhang HJ, Tan GT, Van Hung N, Cuong NM, Franzblau SG, Soejarto DD, Fong HH, Pauli GF. (2005). Anti-tuberculosis constituents from the stem bark of *Micromelum hirsutum*. *Planta Med*, 71, 261-267.

Ma C, Zhang HJ, Tan GT, Hung NV, Cuong NM, Soejarto DD, Fong HH. (2006). Antimalarial compounds from *Grewia bilamellata*. *J Nat Prod*, 69, 346-350.

Ma CY, Musoke SF, Tan GT, Sydara K, Bouamanivong S, Southavong B, Soejarto DD, Fong HH, Zhang HJ. (2008). Study of antimalarial activity of chemical constituents from *Diospyros quaesita*. *Che.. Biodivers*, 5, 2442-2448.

Moerman DE, Pemberton RW, Kiefer D, Berlin B. (1999). A comparative analysis of five medicinal floras. *J Ethnobiol*, 19, 49-67.

Nguyen QC, Nguyen VH, Santarsiero BD, Mesecar AD, Nguyen MC, Soejarto DD, Pezzuto JM, Fong HH, Tan GT. (2004). New 3-O-acyl betulinic acids from *Strychnos vanprukii* Craib. *J Nat Prod*, 67, 994-998.

Preacher, KJ. (2001). Calculation for the chi-square test: An interactive calculation tool for chi-square tests of goodness of fit and independence [Computer software]. Available at <http://quantpsy.org> and <http://quantpsy.org/chisq/chisq.htm>. Accessed on 9/22/2011.

Riley MC. (2001). Traditional Medicine Research Center (TMRC): A potential tool for protecting Traditional and Tribal Medicinal Knowledge in Laos. *Cultural Survival Quart*, 24, 21-24.

Soejarto DD, Gyllenhaal C, Regalado JC, Pezzuto JM, Fong HHS, Tan GT, Hiep NT, Xuan LT, Binh DQ, Hung NV, Bich TQ, Thin NN, Loc PK, Vu BM., Southavong BH, Sydara K, Bouamanivong S, O'Neill MJ, Lewis J, Xie X, Dietzman G. (1999). Studies on biodiversity of Vietnam and Laos: A UIC-based ICBG Program. *Pharm. Biol* 37 (Supplement), 1-14. Available online at <http://www.uic.edu/pharmacy/depts/ICBG/UIC-ICBG-Pbiol-1999-.pdf>. Accessed on 9/22/2011.

Soejarto DD, Gyllenhaal C, Regalado JC, Pezzuto JM, Fong HHS, Tan GT, Hiep NT, Xuan,LT, Hung NV, Bich TQ, Loc PK, Vu BM, Southavong BH, Sydara K, Bouamanivong S, O'Neill MJ, Dietzman G. (2002). An international collaborative

program to discover new drugs from tropical biodiversity of Vietnam and Laos. *Nat Prod Sci*, 8,1-15.

Soejarto DD, Hiep NT, Loc PK, Cuong NM, Bien LK, Dai TD, Kadushin, MR. (2004a) Seed Plants of Cuc Phuong National Park, Vietnam. A Documented Checklist. Published by Cuc Phuong National Park, Ninh Binh, Vietnam. Agricultural Publishing House, Hanoi, Vietnam,. i-xxxiv, 1-760, plates I-XCIX.

Soejarto DD, Gyllenhaal C, Fong HHS, Xuan LT, Hiep NT, Hung NV, Bich TQ, Southavong B, Sydara K, Pezzuto JM. (2004b). The UIC ICBG (University of Illinois at Chicago International Cooperative Biodiversity Group) Memorandum of Agreement: A model of benefit-sharing arrangement in natural products drug discovery and development. *J Nat Prod*, 67, 294-299.

Soejarto DD, Fong HHS, Tan GT, Zhang HJ, Ma CY, Franzblau SG, Gyllenhaal C, Riley MC, Kadushin MR, Pezzuto JM, Xuan LT, Hiep NT, Hung NV, Vu BM, Loc PK, Dac LX, Binh LT, Chien NQ, Hai NV, Bich TQ, Cuong NM, Southavong B, Sydara K, Bouamanivong S, Ly HM, Thuy TV, Rose WC, Dietzman GR. (2006). Studies on Biodiversity of Vietnam and Laos 1998-2005: Examining the impact. *J Nat Prod*, 69 (Special Issue), 473-481. Available online at <http://pubget.com/paper/16562860>. Accessed on 9/22/2011.

Soejarto DD, Gyllenhaal C, Tarzian Sorensen JA, Fong HHS, Xuan LT, Binh LT, Hiep NT, Hung NV, Bich TQ, Vu BM, Southavong BV, Sydara K, Pezzuto JM, Riley MC. (2007). Bioprospecting Arrangements: Cooperation Between the North and the South. In: *Intellectual Property Management in Health and Agricultural Innovation: A Handbook of Best Practices*, Krattiger A, Mahoney RT, Nelsen L, et al., Eds. MIHR: Oxford, U.K., and PIPRA: Davis, U.S.A.. Available online at www.ipHandbook.org. Accessed on 9/22/2011.

Soejarto DD, Southavong B, Sydara K, Bouamanivong S, Riley MC, Libman A, Kadushin MR, Gyllenhaal C. (2009). Studies on medicinal plants of Laos: a collaborative program between the University of Illinois at Chicago (UIC) and the Traditional Medicine Research Center (TMRC), Lao PDR: Accomplishments. In: *Contemporary Lao Studies: Research on Development, Culture, Language, and Traditional Medicine*, pp. 307-323. Center for Laos Studies, San Francisco, and Center for Southeast Asian Studies, Northern Illinois University, De Kalb, Illinois. [Selected Papers from the First International Conference on Lao Studies, De Kalb, Illinois, May 20-22, 2005.]

Soejarto DD, Gyllenhaal C, Kadushin MR, Southavong B, Sydara K, Bouamanivong S, Xaiveu M, Zhang HJ, Rong L, Franzblau SG, Fong HHS, Riley MC, Elkington BG, Waller DP. (20XX). An ethnobotanical survey of medicinal plants of Laos toward the discovery of bioactive compounds as potential candidates for pharmaceutical development. *Pharmaceutical Biology* [present issue].

Spjut RW. (2005). Relationships between plant folklore and antitumor activity: an historical review. *Sida*, 2 (4), 2205-2241.

Svetaz L, Zuljan F, Derita M, Petenatti E, Tamayo G, Cáceres A, Cechinel Filho V, Giménez A, Pinzón R, Zacchino SA, Gupta M. (2010). Value of the ethnomedical information for the discovery of plants with antifungal properties. A survey among seven Latin American countries. *J Ethnopharmacol*, 127(1), 137-1358.

[Tan AC](#), [Konczak I](#), [Sze DM](#), [Ramzan I](#). (2011). Molecular pathways for cancer chemoprevention by dietary phytochemicals. *Nutr Cancer*, 63, 495-505.

Truong NB, Pham CV, Doan HT, Nguyen HV, Nguyen CM, Nguyen HT, Zhang HJ, Fong HH, Franzblau SG, Soejarto DD, Chau VM. (2011). Antituberculosis cycloartane triterpenoids from *Radermachera boniana*. *J Nat Prod*, 74, 1318-1322

Yuan J, Johnson RL, Huang R, Wichterman J, Jiang H, Hayton K, Fidock DA, Wellems TE, Inglese J, Austin CP, Xu X-Z. (2009). Genetic mapping of targets mediating differential chemical phenotypes in *Plasmodium falciparum*. *Nature Chem Biol*, 5, 765-771.

Zhang H-J, Tamez PA, Aydogmus Z, Tan GT, Saikawa Y, Hashimoto K, Nakata M, Hung NV, Xuan LT, Cuong NM, Soejarto DD, Pezzuto JM, Fong HHS. (2002) Antimalarial agents from plants. III. Trichothecenes from *Ficus fistulosa* and *Rhaphidophora decursiva*. *Planta Med*, 68, 1088-1091

Zhang H-J, Tan GT, Hoang VD, Hung NV, Cuong NM, Soejarto DD, Pezzuto JM, Fong

HHS. (2003a). Natural anti-HIV agents. Part III. Litseaverticillols A-H, novel sesquiterpenes from *Litsea verticillata*. *Tetrahedron* 59, 141-148.

Zhang H-J, Tan GT, Hoang VD, Hung NV, Cuong NM, Soejarto DD, Pezzuto JM, Fong HHS. (2003b). Natural anti-HIV agents. Part IV. Anti-HIV constituents from *Vatica cinerea*. *J Nat Prod*, 66, 263-8.

Zhang HJ, Sydara K, Tan GT, Ma C, Southavong B, Soejarto DD, Pezzuto JM, Fong HH. (2004). Bioactive constituents from *Asparagus cochinchinensis*. *J Nat Prod*, 67, 194-200.

Zhang HJ, Ma C, Nguyen VH, Nguyen MC, Tan GT, Santarsiero BD, Mesecar AD, Soejarto DD, Pezzuto JM, Fong HH. (2006). Miliusanes, a class of cytotoxic agents from *Milium sinensis*. *J Med Chem*, 49, 693-708.

Table 1. Accessions in the Vietnam-Laos ICBG project. Includes collections (unit of 1 species collected at one locality on one date), samples (plant parts, often more than one per collection), medicinal uses (unit of 1 use recorded for 1 species in 1 interview) and bioassay results (unit of 1 bioassay test on one sample).

	Laos Ethnomedical	Vietnam Ethnomedical	Random (Vietnam)	Total
Collections	374	318	1124	1816
Samples	376	326	2547	3249
Uses	661	457	--	1118
Bioassay results	2876	2526	19450	24,852

Table 2. Contingency table, hit rate and statistical analysis for plant **collections** analyzed for antimycobacterial assay (MABA-H37RA). Statistical analysis refers to contingency table.

Accession type and rationale	A	I	Total	% active	Contingency table chi-square, df, p-value
Laos ethnomedical	19	183	202	9.4	13.229, df=2, -p=.00134
Vietnam ethnomedical	4	224	228	1.8	
Random (Vietnam)	36	704	740	4.9	
Total	59	1111	1170		

Table 3. Hit rates for **collections** and for **samples**, all random versus all ethnomedical. Active designation based on highest level of qualitative result (A>Q>I) for any test result pertaining to accession.

Accession type	Rationale (n)	% active	Chi-square	df	p
Collections	Ethnomedical (692)	28.5%			
	Random (1124)	31.0%	26.083	2	.00000217
Samples	Ethnomedical (703)	28.1%			
	Random (2551)	22.2%	19.347	2	.000063

Table 4. Hit rate for **collections**, sorted by rationale and country, for all 5 bioassay systems. Active designation based on the best qualitative results (A>Q>I) for any test result pertaining to collection.

Assay type (n)	% active, Laos ethnomedical	% active, Vietnam ethnomedical	% active, Random (Vietnam)	Contingency table chi-square	df	P, chi-square or Fisher's
Antimycobacterial (1168)	9.4	1.8	4.9	13.229	2	.000134
Antiplasmodial (1154)	7.0	6.5	13.1	32.332	4	.00000172
HIV (835)	5.2	--	8.4	36.442	2	.000000001
Chemoprevention (400)	--	42.1	54.1	5.95	1	.018
Cancer (1421)	12.5	4.1	18.2	na, Fisher's test	4	.000000001 47

Table 5. Hit rate for **samples**, sorted by rationale and country, for all 5 bioassay systems. Active designation based on the best qualitative results (A>Q>I) for any test result pertaining to sample.

Assay type (n)	% active, Laos ethnomedical	% active, Vietnam ethnomedical	% active, Random (Vietnam)	Contingency table chi-square	df	P, chi-square or Fisher's
Antimycobacterial (2342)	9.3	1.8	2.5	29.85	2	.000000033
Antiplasmodial (2098)	7.0	6.2	7.0	10.469	4	.033
HIV (2682)	5.2	--	5.9	13.583	2	.00112
Chemoprevention (682)	--	41.2	54.1	3.646	1	.0562
Cancer (2633)	14.2	3.9	11.8	na, Fisher's test	4	.0000025

Table 6. Hit rate for **samples**, sorted by rationale and country, for individual bioassay systems. Active designation based on the best qualitative results (A>Q>I) for any test result pertaining to sample.

Bioassay (n)	% active, Laos ethnomedical	% active, Vietnam ethnomedical	% active, Random (Vietnam)	Contingency table chi-square	df	p, chi-square or Fisher's
Antimycobacterial						
MABA-H37RA (2342)	9.3	1.7	2.5	29.85	2	.000000033
Antimycobacterial						
D6 (2088)	7.6	6.6	7.3	10.424	4	.033
W2 (359)	80.0	75.0	45.2	na, Fisher's test	4	.00000004
HIV						
HOG.R5 (2682)	5.2	--	5.9	13.583	2	.00112
Chemoprevention						
ARE (642)	--	10.6	9.6	0.162	1	.68
ARO (641)	--	9.3	4.6	5.34	1	.0208
COX-1 (637)	--	0.8	1.0	0.03	1	.8624
COX-2 (640)	--	14.0	1.0	3.27	1	.0703
NFκB (641)	--	2.5	3.7	0.635	1	.4255
QR (637)	--	2.2	13.9	7.29	1	.00694
Cancer						
COL-2 (1712)	5.3	--	7.0	1.038	1	.308
HL-60 (1645)	1.9	--	1.2	22.518	1	.000013
HTERT (1774)	6.1	--	9.4	3.641	1	.0564
HUVEC (1763)	6.6		8.6	1.316	1	.2513
KB (1985)	5.7	--	7.0	0.647	1	.421
LNCaP (2401)	7.3	0.4	8.0	16.386	2	.00028
LU-1 (2401)	8.6	3.0	7.4	7.404	2	.0247
MCF-7 (631)	--	1.3	4.7	4.878	1	.027

Abbreviations of chemoprevention bioassays: ARE = antioxidant response element; ARO = aromatase inhibition;

COX-1 = cyclooxygenase-1; COX-2 = cyclooxygenase-2; NFκB = nuclear factor kappa B; QR = quinone reductase.

Cancer bioassays: COL-2: human colon cancer; HL-60:human promyelocytic leukemia; HTERT:human telomerase reverse transcriptase; HUVEC: human umbilical vein endothelial cells; KB: oral carcinoma; LNCaP:human prostate carcinoma; LU-1:lung cancer; MCF-7:human breast cancer.

Table 7. Hit rate for **collections**, sorted by specific versus other ethnomedical use, for all 5 bioassay systems (all bioassays included for malaria, chemoprevention and cancer treatment). Active designation based on the best qualitative results (A>Q>I) for any test result pertaining to collection.

Assay type (n)	% active, specific ethnomedical use category	% active, other ethnomedical use categories	% active, Random (Vietnam)	Contingency table chi-square	df	P, chi-square or Fisher's
Tuberculosis (n)	3.4	6.3	4.9	1.308	2	0.52
Malaria (n)	7.1	6.7	13.1	22.408	4	.00016
HIV (n)	6.5	5.0	8.4	33.949	4	000000077
Chemoprevention (n)	10.1	8.3	18.2	na, Fisher's test	4	2.2×10^{-37}
Cancer (n)	36.1	44.4	54.1	6.026	3	.049

Figure 1. Graphic representation of the difference between collections and samples. A collection is a single individual plant collected on a certain date in one locality. For every collection, two plant materials are gathered: (a) voucher herbarium specimens; (b) one or more bulk samples each composed of selected plant parts. The two types of plant materials are linked by a common tracking number.

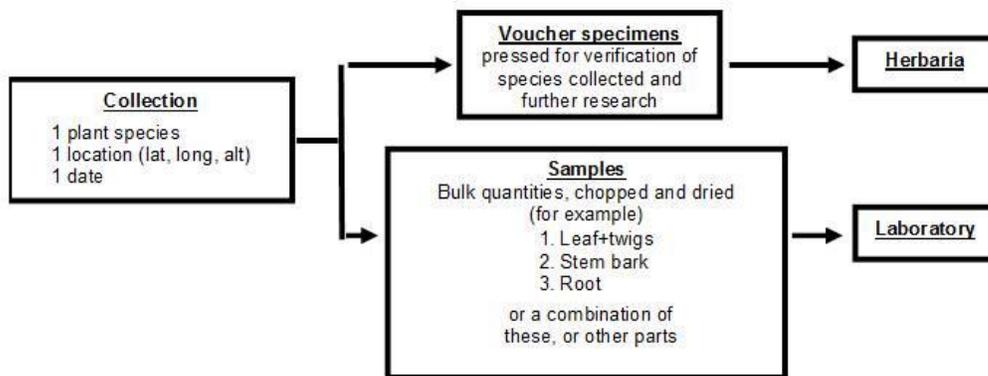


Figure 2. Distribution of ethnomedical use categories for total collection (Laos + Vietnam), Laos collections and Vietnam collections.

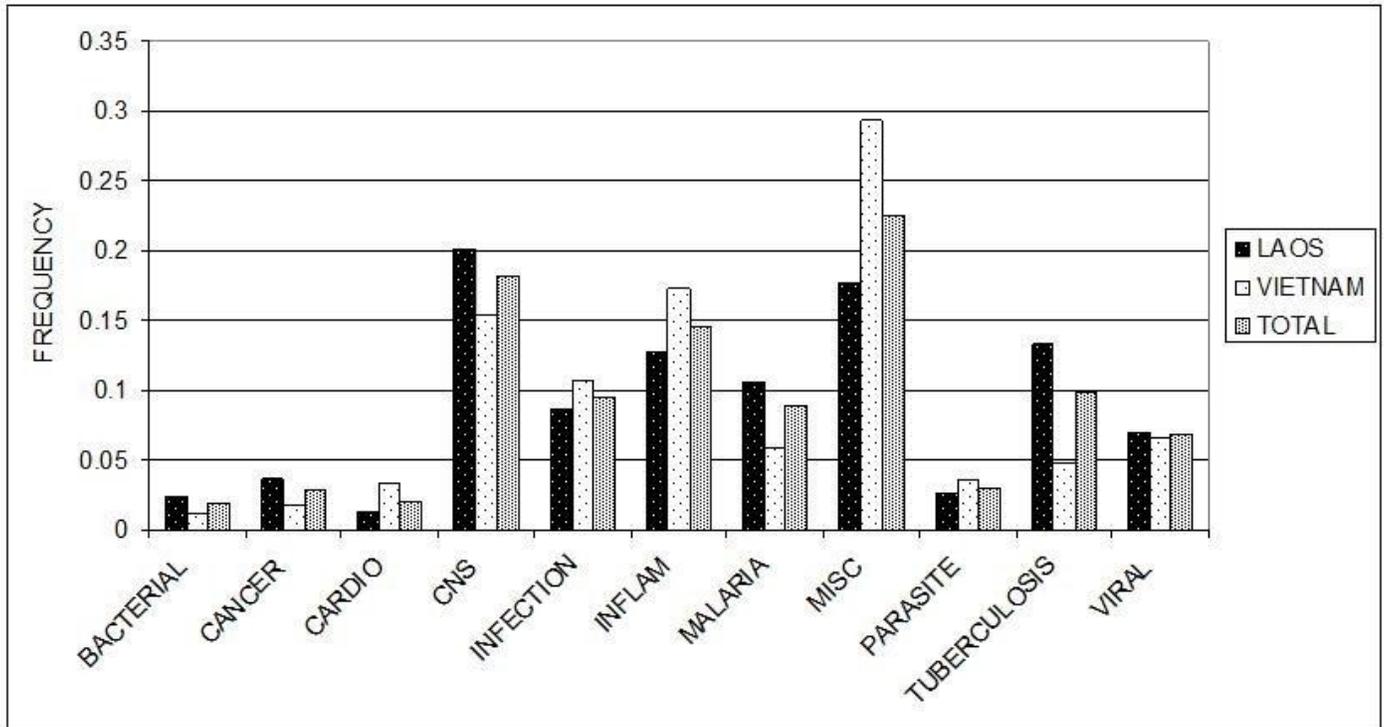


Figure 3. Frequencies of ethnomedical use categories in the active collections and in the total collection. A. Antimycobacterial actives (n=33). B. Antiplasmodal actives (n=112). C. HIV actives (n=19). D. Cancer chemoprevention actives (n=246). E. Cancer treatment actives (n=284).

