Biodiversity and Chemodiversity: Future Perspectives in Bioprospecting

Ramesha B.T.[1,2], Jürg Gertsch[3], Ravikanth G.[2,5], Priti V.[2], Ganeshaiah K.N.[2, 4, 5] and Uma Shaanker R.[*,1,2,5]

[1]*Department of Crop Physiology, University of Agricultural Sciences, GKVK, Bangalore 560 065, India*

[2]*School of Ecology and Conservation, University of Agricultural Sciences, GKVK, Bangalore 560 065, India*

[3]*Institute of Biochemistry and Molecular Medicine, University of Bern, CH-3012 Bern, Switzerland*

[4]*Department of Forestry and Environmental Sciences, University of Agricultural Sciences, GKVK, Bangalore 560 065, India*

[5]*Ashoka Trust for Research in Ecology and the Environment, Royal Enclave, Sriramapura, Jakkur, Bangalore 560 064, India*

Abstract: Biological diversity and its constituent chemical diversity have served as one of the richest sources of bioprospecting leading to the discovery of some of the most important bioactive molecules for mankind. Despite this excellent record, in the recent past, however, bioprospecting of biological resources has met with little success; there has been a perceptible decline in the discovery of novel bioactive compounds. Several arguments have been proposed to explain the current poor success in bioprospecting. Among them, it has been argued that to bioprospect more biodiversity may not necessarily be productive, considering that chemical and functional diversity might not scale with biological diversity.

In this paper, we offer a critique on the current perception of biodiversity and chemodiversity and ask to what extent it is relevant in the context of bioprospecting. First, using simple models, we analyze the relation among biodiversity, chemodiversity and functional redundancies in chemical plans of plants and argue that the biological space for exploration might still be wide open. Second, in the context of future bioprospecting, we argue that brute-force high throughput screening approaches alone are insufficient and cost ineffective in realizing bioprospecting success. Therefore, intelligent or non-random approaches to bioprospecting need to be adopted. We review here few examples of such approaches and show how these could be further developed and used in the future to accelerate the pace of discovery.

Keywords: Biodiversity, bioprospecting, chemodiversity, data mining, evolutionary rationale, phylogenetic rationale, search strategies, traditional knowledge.

INTRODUCTION

Biological diversity and its constituent chemical diversity have served as one of the richest sources of bioprospecting thus leading to the discoveries of some of the most important bioactive molecules that mankind have ever known [1-5]. From the antibiotic penicillin (from *Penicillium* sp.), the aspirin precursor scaffold salicylic acid (from *Salix* sp.), or the anticancer agent taxol (from *Taxus brevifolia*) to the antimalarial artemisinin (from *Artemisia annua*), the chemical diversity contained in plants, animals, insects and microorganisms has directly or indirectly contributed to the development of approximately 75 per cent of all known pharmaceutical compounds [6, 7]. These discoveries have not only profoundly contributed to human health but also in large part have shaped the political geography of the world [8, 9]. Many of these spectacular discoveries made during the last 150 years owe to a) the naturally available and biologically active compounds found in nature, b) the fortitude of indigenous communities, who discovered the cures residing in plants and animals and shared their knowledge among communities and across generations and c) to the synthetic chemists who used their skills to elucidate the chemical structures to reconstruct and develop the most interesting bioactive natural products [5, 10].

Motivated by these rich findings, and with the hope of emulating the early successes, interest in chemically exploring biological resources began to be taken up earnestly by natural product chemists for most part of the last century [11, 12]. Scores of organisms including plants, microorganisms, insects and animals like never before, were chemically profiled; numerous natural products were discovered and described. Unfortunately, not many of these went on to make any significant impact let alone contribute to making blockbuster drugs [10, 13].

The lack of success and general despair (and perhaps even a lack of deep conceptualization in the area of natural product chemistry) gave way to the science of synthetic and combinatorial chemistry [13, 14]. Deriving momentum from the new found tools, hundreds of thousands of new compounds were synthesized in the hope that these would unleash a flood of potentially useful molecules. Barely two or three decades into this new found activity and hard labour, it was soon clear that by just creating compounds by their

*Address correspondence to this author at the Department of Crop Physiology, University of Agricultural Sciences, GKVK campus, Bangalore 560 065, India; Tel: 91-80-23636350; Fax: (+91)-80-23530070; E-mail: umashaanker@gmail.com

1389-4501/11 $58.00+.00

hundred thousands did not necessarily lead to success [15-17]. Hard work or brute force alone is insufficient to cause success in drug discovery [17, 18]. Further, as someone put it rather uncharitably, synthetic and combinatorial chemistry only helped increase the size of the hay stack and made our search for the proverbial needle that much more difficult [19]. However, recently the development of natural-product-like libraries has shifted the focus from quantity to quality [20].

With the general disappointment of early combinatorial chemistry, interest has once again shifted back to bioprospecting, i.e. the search for economically interesting constituents from nature [21-23]. Moreover, there is an interest to build combinatorial libraries based on promising natural product scaffolds. The implicit belief is that if one is going to find biologically active and relevant compounds, the best bet would be to look at the chest of compounds that organisms have adapted to produce during the course of their evolutionary biochemical interactions. But as we are on the threshold of what appears to be another long stint of exploration of biological resources and their natural products, there are some concerns that need to be addressed.

In this paper, we reflect upon the current perception of biodiversity and chemodiversity in drug discovery and ask to what extent it is relevant in the context of the renewed interest in bioprospecting. First, we critically review the relationship between biodiversity and chemodiversity and show that if anything, the biological space for exploration is still wide open. Second, considering the fact that discovery rates from biological resources, e.g. expressed as new chemical entities, have dampened (or at least generally perceived to be so) over the last years, we suggest that newer and smarter strategies should be developed that can not only make bioprospecting more focused, but also more rapid. Third and finally, we discuss potential algorithms that could make the search for biologically useful molecules more concerted and rapid. A few examples demonstrating these algorithms are presented. We conclude that bioprospecting is still one of the most productive of approaches in the identification of lead structures for drug discovery. The current loss of biodiversity through human destruction and climate change will therefore directly affect bioprospecting because it is correlated with natural product diversity as discussed by Pietra, 2002 [24].

BIODIVERSITY AND CHEMODIVERSITY

A central assumption in bioprospecting the plant kingdom or any other group of organisms is that they are storehouses of secondary metabolites, evolved over millions of years within a biochemical environment. Thus, these compounds are predestined to have biological or pharmacological relevance [25]. An obvious corollary that follows is that bioprospecting greater parts of biodiversity is likely to yield newer and potentially novel compounds that may have novel utilities. *Is it really so?* In the last decade, there have been some mixed opinions, some endorsing this view [1, 11, 12, 26] and yet others stridently opposing it [27], each of course based on their own points of view. For example, the former view emerges from the rich success that bioprospecting has met since the last couple of hundreds of years. In fact, it is almost a *cliche* to mention that biodiversity and its constituent chemical diversity have contributed to drug discovery second to none. The opposing view, comes from the observation that initial successes aside, bioprospecting of plants in the last few decades has actually dampened off, with fewer and fewer discoveries. In this section, we briefly address these apparent contradictions by analyzing how biodiversity, as we understand it, is related to the underlying chemical and functional diversity.

What is the Relation between Biological Diversity and Chemical Diversity?

We use the terms, biological and chemical space to reflect respectively, the relative amplitude of biological and chemical diversity in a given group of organisms or taxa and ask how they might be related. We propose two simple models to explain the possible relationship [Fig. (**1**)]. In the first model, the chemical space scales linearly with the biological space; that is with increase in biological diversity of the taxa in question, the chemical diversity contained in the taxa also increases. This model assumes that with increasing biological space, more novel chemical scaffolds are made, probably to address newer challenges and adaptations of species [28-30]. For example, it is known that total number of alkaloids scales directly with the number of species and genera examined [31, Fig. (**2A, B**)]. Evidence exists to suggest that the structural diversity of compounds has also increased with speciation [28, 32]. Under this model, it would be appropriate to bioprospect greater biodiversity to harness a greater pool of chemical diversity.

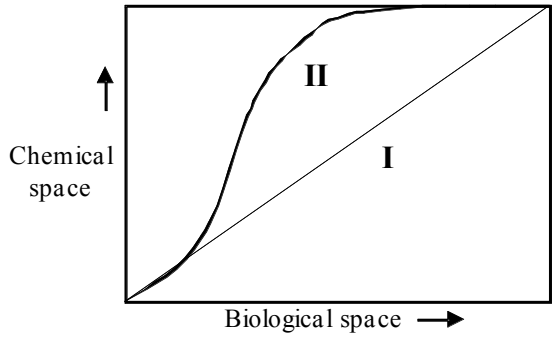

Fig. (1). Relationship between biological space (biological diversity) and chemical space (chemical diversity). Please see text for explanation.

In the second model, the chemical space rises rapidly with increase in biological space and then dampens off [Fig. (**1**)]. In this model we assume that, chemical plans, akin to, let us say, body plans in animals, are finite, and therefore, soon after most of the chemical space or the number of potential chemical scaffolds have been recovered, there is a dampening of the chemical space with further increase in biological space. A specific example to illustrate this could be sought from the diversity of alkaloids that are reported from plants. Over all the major plant groups, monocots and dicots (in angiosperms) and gymnosperms, only about 22 different alkaloid scaffolds have been recovered. For example, the family Asteraceae with the highest species diversity (21000 species) is known to produce only 14 of the 22 known structural classes of alkaloids and Fabaceae with 16400 species is known to produce as many as 17 of the 22

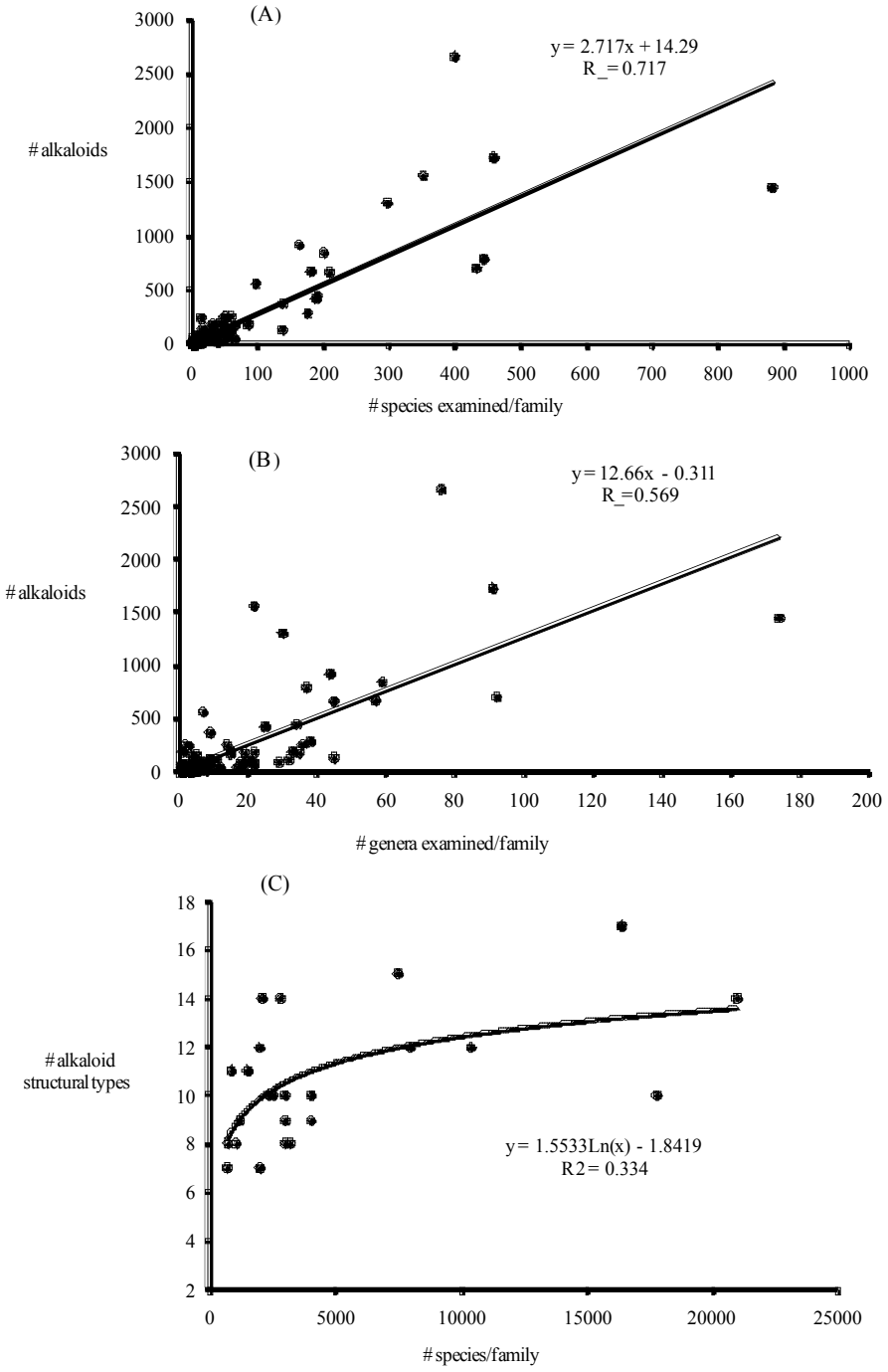

Fig. (2). Relationship between number of alkaloids discovered and (**A**) per family and number of genera per family (**B**) and number of alkaloid structural types and number of species per family (**C**). Graphs redrawn from Cordell *et al.* (2001) [31].

classes of alkaloids. Overall, there seems to be dampening off of the number of structural classes of alkaloids with number of species per family [Fig. (**2C**), 31].

Thus, it is unlikely that fine-combing of the global plant resources (or for the matter any biological resources), which share similar biological niches, would increase this number substantially. This model also emerges from a phylogenetic-evolutionary consideration; phylogenetic inertia for chemical compounds may far outweigh the rate at which newer compounds or structural plans evolve. In an interesting paper, Becerra *et al.* (2009) [28] studied the macroevolutionary patterns of species and chemical diversity in the genus *Bursera*, which is rich in terpenoids. As predicted by co-evolutionary theory, in the genus *Bursera,* descendant taxa (new species diverged over time) had significantly higher number of natural product classes than its ancestral taxa, clearly suggesting that, as new species diverged over time, they tended to be equipped with more compounds per species, presumably to serve newer adaptations [28]. However, the rates at which chemical diversity increased was much slower than the rate at which species diversity increased [28]. A similar picture is obtained on analyzing the microbial flora. Even though species diversity is extremely

high, it does not appear to correlate linearly with chemical scaffold diversity [24].

An obvious implication of this model is that the marginal pay-offs of bioprospecting additional biodiversity is bound to decline. Furthermore, the model suggests that rather than just increase the number of species bioprospected, it would be better if bioprospecting is done across clades or phylogenetic groups, and species rich families that would maximize the chemical space.

What is the Relation between Chemical Space and Functional Space?

From the point of view of bioprospecting, how much of the chemical space (diversity) is actually biologically useful? In other words, what is the relation between chemical space and the functional space? We discuss two possible scenarios based on purely theoretical considerations. In the first scenario, the functional space scales linearly with the chemical space [Fig. (**3**), R = 0]. This scenario assumes no redundancy among the secondary metabolites biosynthesized by plants. Every compound produced (within the chemical space) is unique and therefore contributes to a unique functionality. Thus, from a bioprospecting perspective, it pays to search more and more of the chemical space. However, clearly this is not necessarily true as there are numerous cases known where very diverse natural products appear to have the same function or no apparent function (*see below*).

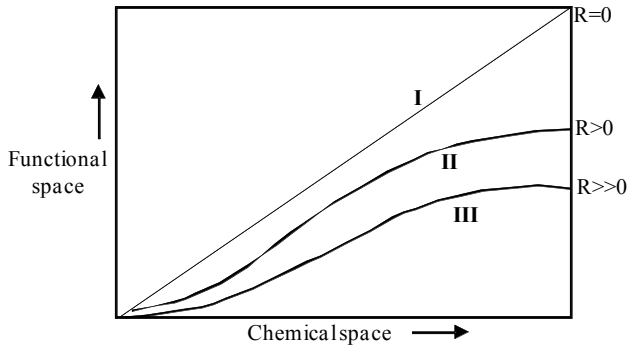

Fig. (3). Relationship between chemical space (chemical diversity) and functional space (functional diversity). Please see text for explanation. R refers to the extent of redundancy.

In the second scenario, initially the functional space increases with chemical space but soon attains an asymptote [Fig. (**3**), R>0]. Here, we assume a functional redundancy for the chemical compounds made by organisms. Thus, while the chemical structures might vary, they are assumed to still share common receptors. Several empirical examples lend support to this assumption. For example, as of now, besides taxol, there are at least eight different compounds belonging to different classes of natural products that have been shown to induce 'microtubule stabilization' [33]. However, microtubule-stabilizing natural products known to date share the feature of a flexible macrocyclic structure with the exception of the lactone discodermolide [34]. Accordingly, the relationship between the functional and chemical space would depend upon the extent of redundancy; a greater redundancy would weaken the relationship between the functional and chemical space [Fig. (**3**), R>>0]. The attendant implication of such a relationship on bioprospecting is quite obvious. However, before this model is accepted, it has to be borne in mind that the domains of functional space are a function of our tool-kits. With the ongoing increase of our knowledge of the receptor sites, of structural features of proteins and their folding, of enzymes and their expressions, of genes and their regulation, etc., one can expect the functional space to expand and therefore relate differently to the chemical space. For example, in certain situations, multiple functionalities of the same natural product could be assumed. Gossypol, a sesquiterpene dimer found in cotton is available as a mixture of (+)- and (-)-enantiomers. Only (-)-gossypol is known to induce apoptosis in non-ruminant animals such as rodents, rabbits and humans by inhibiting mitochondrial membrane bound antiapototic protein Bcl-X_L. However, the (+)- enantiomer completely lacks this activity and is least toxic. On the contrary, both (+)- and (-)-enantiomers are shown to be equally effective on insects and fungi, clearly suggesting multiple functionalities of the natural product gossypol [35]. The same may be true for mammalian secondary metabolites, which are probably best studied. While peptides may be rather selective towards their respective receptors, lipids and other secondary metabolites like steroids often target different receptors to various degrees (e.g. seen with endocannabinoids, neurosteroids, etc.).

Summing-up: Is it Promising to Mine Biological Diversity?

In recent years, there has been some rather acrimonious debate on the utility of biodiversity in bioprospecting. Most of these debates have been prompted by the new found tool of synthetic and combinatorial chemistry as well as the rather poor track record of returning discoveries from natural products research. For instance, Tulp and Bohlin (2002) [36], while acknowledging that natural sources can contribute to identification of leads for novel targets, argue that is not necessary that one has to literally scan the entire of biodiversity, for the compounds are likely to be found in more than one species as important molecular mechanisms are likely to be ubiquitous [37, 38]. Further, since peptide receptor ligands tend to be highly conserved in nature, evolutionary pressures on plants and animals could have shaped the evolution of "similar" chemicals to comply with these "conserved" receptor ligands. For instance, recently, besides vertebrates, cannabinoid receptors (CB_1 and CB_2) have been shown to be present in several other species; interestingly, the ligands are also similar across species [39, 40]. On the other hand, plants which do not express cannabinoid receptors are able to synthesize active cannabinoid receptor ligands [41]. Thus searching endlessly among the bioresources is unlikely to take us further than we are. In this context, Tulp and Bohlin (2002) [27] argue that there might be no obvious advantage in emphasizing the need for bioprospecting in mega diverse countries more than in any other landscape for the matter. As they mention, "in all its complexity, life is probably simpler than we think" [27].

However, and in contrast to such disapproval, it has to be borne in mind that natural product libraries (as currently available for biological testing) are far from investigated [35,

42]. Too often, researchers keep their small amounts of isolated compounds and they are not accessible for broader screening. Thus, unique bioactive natural products may be published but never been screened for a meaningful biological activity. Moreover, only a fraction of all natural products is currently known. Less than 20 per cent of all higher plants have been systematically investigated [31] and it remains to be determined whether this is enough to draw statistical conclusion. In an excellent review, Cordell *et al.* (2001) [31] highlighted the diversity of alkaloids in plants and argued how even for this well known natural product class, there still remains a large degree of ignorance. As a group of compounds, alkaloids are distributed in most plants. Till date, alkaloids have been recovered from 7231 species of higher plants from 1730 genera within 186 families [31]. However, there remain about 153 families with about 674 genera and 5835 species which have yet to be explored for the presence of alkaloids. Altogether the chemical plans for this array of alkaloids fit into 22 structural classes. Overall, the richness of alkaloids seems to be restricted to about 20 families and is not uniformly distributed. Of about 21,120 alkaloids that are described from higher plants, a mere 2291 have been evaluated using a single bioassay, 1995 have been evaluated in less than 10 bioassays [31]. Only about 167 have been used in 20 or more assays. In other words, over one-third (35.9%) of the alkaloids that have been examined biologically in 20 or more assays are pharmaceutically significant. Therefore, a significant 76.4 per cent of the alkaloids have not been evaluated even in a single assay. As argued by Cordell *et al.* (2001) [31], this vast chemical space awaits discovery of newer and novel pharmacophoric properties. Thus, it seems that a major problem is that natural products are not generally tested enough. Besides plants, less than one per cent of all microorganisms are perhaps studied. Only recently have other groups such as cyanobacteria, insects, snails and other marine organisms, amphibians etc. even started to be investigated [22]. Thus, there not only exists a rich theoretical basis to bioprospect the remaining of our bioresources but also a substantial fraction (about 85 per cent of all plants) that has barely been investigated.

In summary, careful analysis of the above arguments and propositions suggests two possible relationships between the biological and functional space. First, by increasing our knowledge of biological space, we are likely to discover newer functionalities, i.e. the functional space would be linearly correlated with the biological space [Fig. (**4**)]. Second, because of phyletic/chemical inertia and functional redundancies, functional space would tend to increase rapidly with the biological space and then dampen off [Fig. (**4**)]. But, in either case, it appears that there is a promise for bioprospecting a greater biological space, even if it means going to the ends of the earth or depths of the oceans to search new biological functionalities [12]. In a rather candid admission, Caporale (1995) [25] of Merck Research Laboratories mentioned, "All of the drugs discovered at the Merck Research Laboratories that became available to patients in the last decade emerged from programs that benefited from knowledge of biological diversity". Studies such as these, reiterate that hope lost on bioprospecting biodiversity is not well placed, rather there appears to be hope again to bioprospect the richness of plants, animals, insects and microorganisms. However to a large extent, the promise of bioprospecting notwithstanding, the success of bioprospecting will depend upon how we strategize our search for these functionalities. In the next section we briefly review the various search strategies that are in place for bioprospecting and discuss the hope they hold in enhancing the probability of success.

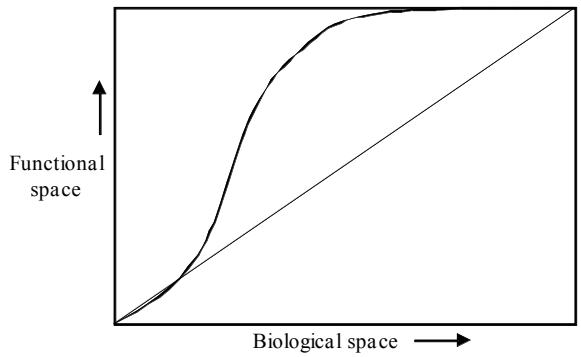

Fig. (4). Relationship between biological space (biological diversity) and functional space (chemical diversity). Please see text for explanation.

BIOPROSPECTING BIODIVERSITY: STRATEGIZING THE SEARCH

"If you don't know where you are going, any road will get you there."- Lewis Carroll

We are not sure if Lewis Carroll got his quote right! One of the major drawbacks in bioprospecting is the extremely low percentage of successful "hits" [42-46]. An often cited example of this is the now classical search effort by the NCI, USA [46]. Of the 35,000 plants species screened by the NCI, only about 4.3 per cent had any anti-cancer activity and only 0.07 per cent of the plants yielded compounds that had high anti-cancer activity [45]. One of these hits led to the discovery of taxol from *Taxus brevifolia* [46, 47]. The Central Drug Research Institute (CDRI), India screened over 2000 plant species for a number of biological activities but failed to get any significant leads [46]. From about 10 million microbes that have been screened, only about 2000 antibiotics (success rate of about 0.002 %) have been identified [17]. Of the 7000 isolates of *Bacillus thuringiensis* that were screened against grass grub, only 0.07 per cent showed substantially high activity [43]. From a data mining analysis of the activity of plant crude extracts for acetycholine esterase inhibition (against diseases such as Alzheimer's), we found that only 0.05 per cent of the plants had an IC_{50} of less than 50 μg/ml [*unpublished*]. A number of other studies have also reported similar low percentages of hits [43, 46, 48]. As a consequence, given that natural product research is time and cost intensive, many pharmaceutical companies have abandoned their natural product programs and closed their natural product units in the last two decades.

Why is the probability of success very low? A key rate limiting step in bioprospecting natural resources is the choice of the starting material. Poor and an inappropriate choice of the starting material can derail bioprospecting success. Most

of the more recent bioprospecting programs were conducted on a random collection of plants and other organisms and hence were essentially shots in the dark [12, 46]. Most of these screening programmes suffered from a mis-match between the choice of biological material and the end-screens. In other words, for most part, the screening was like trying to push a cylindrical peg through a square hole. Quite expectedly, such programmes did woefully little in advancing our search for a match between the biological material and the biological activity or functional activity. Because of the fact that there was no *a priori* rationale, the probability of finding successful "hits" were extremely low, often one in 10,000 [32]. Translation of these screening successes was limited both in developing relevant leads as well as in their applications to drug discovery research. Jones and Firn (1991) [43] argue that the generally low percent success of bioprospecting efforts could be explained by plants possessing a large number of compounds arising by chance, but having an inherently low probability of possessing any biological activity.

The low probability of success, however, is not typical only of conventional bioprospecting. Large scale combinatorial chemistry efforts that generate compounds randomly are also known to have an abysmally low percentage of successful hits [14]. In a recent analysis of search for polyketides, Li and Vederas (2009) [13] showed that compared to a mere 0.001% successful hits from synthetic compound libraries, two orders higher magnitude of discovery (0.3%) were obtained from natural products. In other words, it appears that despite the low percentage of success, bioprospecting efforts even with a random collection of plants or species seem to fare as well or even better than that scored by combinatorial chemistry. This has also been recognized by the pharmaceutical industry and few big pharma companies like e.g. Novartis still hold on to their natural product programs. Nevertheless, it is increasingly realized that efforts should be made to enhance the success rates, if bioprospecting has to become a profitable enterprise. A pointer in this direction was interestingly offered by the US Congressional committee that evaluated the discovery of taxol. The committee observed that it would have been better if the NCI had not restricted its bioprospecting efforts to only species selected at random; rather NCI could have also used other criteria including indigenous traditional knowledge of people to advice on the choice of species for screening [46, 47, 49, 50]; such an approach could have substantially hastened the discovery process. However, the unsuccessful example of Shaman Pharmaceuticals, a company which has exclusively focused on plants with ethnopharmacological background, clearly shows that this is not always an easy solution. Currently, there are several programmes for drug discovery from traditional Chinese medicine (TCM) and Ayurveda herbal medicines [51, 52].

In the recent past, several efforts have been made to consciously strategize the selection of species with the aim of maximizing the probability of successful "hits" and thus making bioprospecting a cost effective exercise. This is much the same as efforts that have moved from screening massive combinatorial libraries to smaller and focused libraries of compound. In the following section we review some of these approaches and highlight the recent developments in this area.

SEARCH ALGORITHMS

"My take on natural products based drugs is similar to that on extra-terrestrial intelligence. There's probably a lot of them out there, but how do you find them" Derek L. [53]

As opposed to a random-walk approach to bioprospecting, which is both tedious and cost ineffective, efforts have been made from time to time to employ what may be regarded as targeted or non-random strategies to obtaining leads with a greater degree of certainty and faster. Broadly these strategies can be classified in to four classes [Fig. (**5**)]

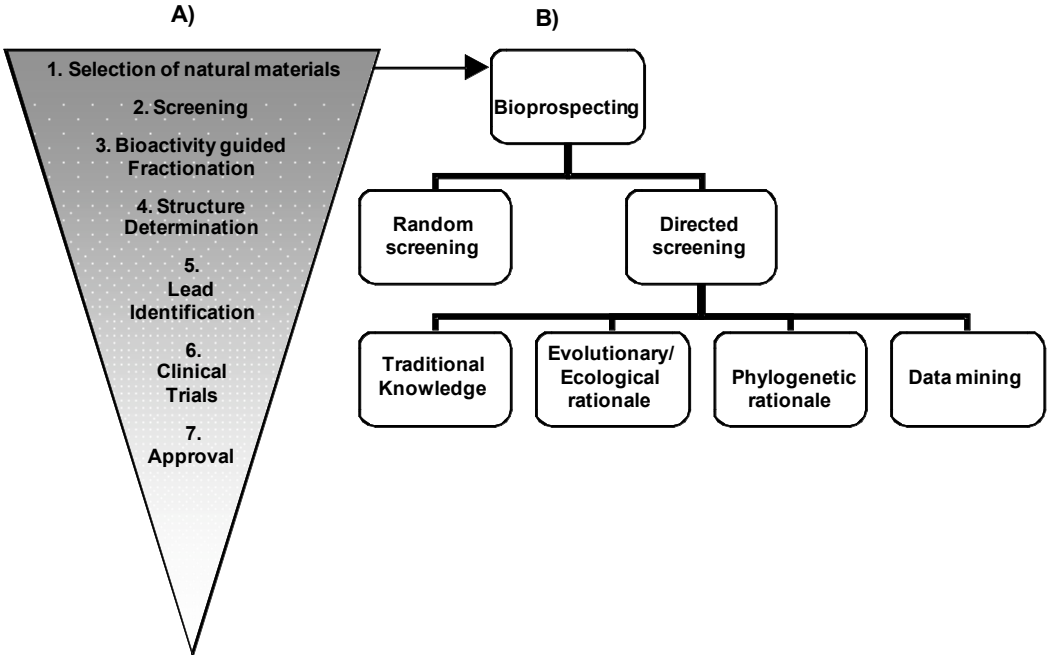

Fig. (5). Schematic diagram of (**A**) drug discovery pipeline and (**B**) different non-random approaches to bioprospecting.

and include those based on indigenous traditional knowledge (ITK) or ethnobotanical information, evolutionary rationale and ecological associations, phylogenetic rationale and data mining. In this section we review each of these search strategies and highlight a few salient applications.

Ethnomedical knowledge

Bioprospecting based on ethnomedical knowledge has been one of the earliest non-random approaches to exploring plants for bioactive properties and compounds [46, 51, 52, 54, 55]. Pioneered predominantly by European chemists, it essentially followed a reductionist approach from folklore and traditional knowledge [46, 54, 55]. One of the earliest examples of such an approach led to the isolation of opium alkaloids including morphine [56]. Later a number of high profile compounds such as quinine (from *Cinchona*) [57], colchicine [58] (from *Colchicum autumnale*), salicin (from the willow tree) [59] etc. were isolated by putting chemistry to work on ethnobotanical and ethnomedical information held by local communities. The plant *Strophantus hispidus* (Apocynaceae) used by several African ethnic groups to poison their arrow heads, was found to contain toxic alkaloids and cardiac glycosides g-strophanthin (syn. ouabain), k-strophanthin and e-strophanthin [60]. Strophanthin's effect on the heart and blood circulation was similar to that of digitalis glycosides derived from common foxglove in Europe [60].

How reliable and successful has been the search based on ethnomedical information? It is often claimed that from an ethnomedical prespective, 65 to 90 per cent of the realized assay results correlate with that projected by ethnomedical information [46, 61]. That is, ethnomedical information shall be a reliable index of the putative properties of the plants. In the recent past a few workers have actually been able to demonstrate that indigenous and traditional knowledge (ITK) based selection of plants resulted in a higher probability of successful bioprospecting "hits" than from those chosen randomly. In a screening program for vasoactivity, Slish *et al.* (1999) [62] showed that plants selected based on an earlier ITK had a greater probability of returning successful hits (10 to 12 %) compared to no success at all from plants selected randomly (with no prior ITK). Similar differences were reported for plants for their anti-malarial effects; plants traditionally used by the Aguaruna ethnic group for malarial fever were more effective than those chosen randomly, in killing the plasmodium [63]. Crude extracts of plants used by healers in Belize (Brazil) produced four times as many positive results in lab tests for anti-HIV activity than did specimens collected randomly [64]. Spjut and Perdue (1976) [65] were able to show greater frequencies (19.9%) of primary activity in anti-neoplastic screens using plants from four families known to treat cancer in traditional medicine (Fabaceae, Liliaceae, Rubiaceae, and Rutaceae) versus a background average of 10.4% in a NCI survey that analyzed 20,525 species during 1960-1980. In another study evaluating plants for their anti-cancer activity, Pandey, (1998) [49] showed that compared to about 10.4 % successful hits obtained from plants selected randomly, for those based on traditional medicine it ranged from 19.9 to 52.5 %, even though there was no direct link to its traditional use. There are several examples which show traditional phytotherapy links with modern pharmacology. The use of *Desmodium adscendens* in Ghana for treatment of asthma provided a long sought agonist of the maxi-K (potassium) channel [66]. Thus, the pharmacological activity of a secondary metabolite of *D. adscendens* is consistent with the use of the herb associated with smooth muscle contraction.

The above studies in many ways confirm the potential of ethnomedicinal information as a bioprospecting strategy [67, 68]. In an interesting paper, Douwes *et al.* (2008) [32] examined the often neglected null hypothesis that plants used by traditional healers are no different from a random collection of plants. Using a regression analysis of ethnomedicinal plants of southern Africa against the total number of taxa in an order, they showed respectively that certain plant orders are either over represented ("hot") or underrepresented ("cold") in traditional medicine than would be expected by chance. In other words, certain taxa or orders seem to have actually been "selected" by native healers compared to other taxa. Regardless of how (by trial or error, or by heuristics) this discrimination came to being, does this mean that bioprospecting the "hot" taxa as opposed to a random collection of taxa would result in a higher probability of bioprospecting hits?

In summary, ethnomedical information based approach has served as a useful search engine; culturally significant plant species may be more efficacious than species collected randomly [3, 54, 55, 64, 68-70] for driving bioprospect-ing especially in biologically and culturally rich regions of the world. Though sometimes looked upon as archaic and unscientific [71], more drugs have been discovered today from ITK than by any other approach [46, 61]. Even as recent as 2001-2005, five drugs obtained from ITK leads were approved by the FDA [72]. However, it has to be admitted that these findings are no match for the heyday that ethnopharmacology witnessed in the last 150 years. Have most of the low hanging fruits been plucked? In other words have we dried up most of the important leads that ethnopharmacology had to offer? In a recent commentary, Gertsch [10] laments that all is not too well with ethnopharmacology as it is practiced today; in fact he wonders if there is enough science in the field of ethnopharamcology. Hand waving instead of rigorous evaluations, sloppy questions, lack of hypotheses or ill-conceived hypotheses and an ad-hoc approach to bioprospecting are some of the characteristics of modern day ethnopharmacology which may inhibit continous drug discovery. Consequently, despite a flurry of papers, there is little incremental translation of these findings in to what may be regarded as leads [10]. Thus, the challenge in using ethnomedical/ethnobotanical information in guiding bioprospecting lies in incorporating good science, developing predictive theories and besides of course in designing appropriate assays/evaluation that make a subjective match with the traditional knowledge. Finally it is important to bear in mind that ethnomedical information based search is merely an approach or a toolkit and by itself is not laden with any intrinsic intelligence or scientific rationale. To this extent, ethnomedical information based bioprospecting will always fall short of discovering and delivering newer and novel processes that might otherwise have profound implications in drug discovery processes. Clearly the solution to the problem (of bioprospecting) lies in developing algorithms that are intelligent and that are based on a deep seated

scientific rationale. One such example would be to strategically study all plants toxic to mammals to explore their molecular mechanisms of action in the context of drug discovery. Given that toxic plants are bioactive per se, the underlying biochemical mechanisms may help to elucidate novel targets for therapies.

Evolutionary Rationale – from HTS to Primary Observations

"Nothing in biology makes sense except in the light of evolution"- Th. Dobzhansky

In a recent visit to a forest site in the Western Ghats, southern India, one of the authors, accidentally brushed his hand against leaves of the Devil's Nettle (*Laportea crenulata*, Urticaceae). In less than 10 minutes, he experienced excruciating pain spreading below and above the site where the leaves had touched his hand. In less than 30 minutes the pain spread to the whole hand and lead to the swelling of the lymph nodes under the arm-pit. He suffered intense pain for the next 12 hours and had to be administered pain killers. What are the underlying proximate biochemical and molecular mechanisms responsible in inducing pain? Literature suggests that the leaves of Devil's Nettle have very fine microscopic hypodermic needle-like hairs that sting in serotonin and acetylcholine [73, 74]. Why do plants have neurotransmitters, when they do not have a neural system? What is the evolutionary or adaptive significance of this behavior? Answers to these and other similar questions are clearly in the realm of the evolutionary history of the plant and its evolved relationships with other organisms. Over the last hundred years, scores of plant products have been extensively used as curatives for a wide range of human ailments by both traditional and allopathic health care systems [46, 55, 61, 75]. Yet the functional significance of most of these products to plants themselves remains enigmatic [32, 76]. Understanding the evolutionary significance of these products to plants can facilitate bioprospecting in a manner that is intelligent and hence more directed. In recent years there have been some attempts to use evolutionary logic in making predictions about processes that might contribute to bioprospecting leads. We briefly review a few of them here.

Uma Shaanker *et al.* (1997) [76] addressed the evolutionary significance of laxatives in plants. Several compounds including certain mucopolysaccharides and anthraquinones are known to be present in plants and responsible for their laxative property [76, 77]. Why do plants contain laxatives, when they have no bowels to move? It was hypothesized [76] that plants might be selected to possess laxative causing compounds in their seeds and fruits, perhaps to optimize the gut passage time of seeds taken in by dispersal agents, such that the life-time fitness of the plant is maximized. Too long a stay in the gut would over-scarify seeds and too less would be ineffective in removing seed coat constraints to germination [Fig (**6**)]. Thus, plants might be selected to pack laxatives in their seeds to ensure an optimal gut passage time for seeds. Accordingly, Uma Shaanker *et al.* (1997) [76] predicted that laxatives should be present in animal dispersed species more frequently than expected by chance compared to species dispersed by non-animal means (wind, water and other passive dispersal modes). They evaluated their prediction using a data set of 114 species reported for their laxative property and found that an overwhelming 60 per cent of the species were indeed animal dispersed. On the contrary there were fewer wind and passively dispersed species with laxative properties than expected. This analysis demonstrated a far reaching implication for targeted bioprospecting: that is, if one were bioprospecting for laxative yielding compounds, then searching animal dispersed species alone would not only fetch a higher probability of successful hits but also would be faster. In the Western Ghats forest, a mega diversity hotspot in south India there are an estimated 6000 plant species, of which roughly 30 per cent are animal dispersed [78]. Applying the evolutionary logic mentioned above, it would be sufficient to bioprospect only 1800 of the 6000 species (a 70% reduction in tedium) for laxative yielding compounds.

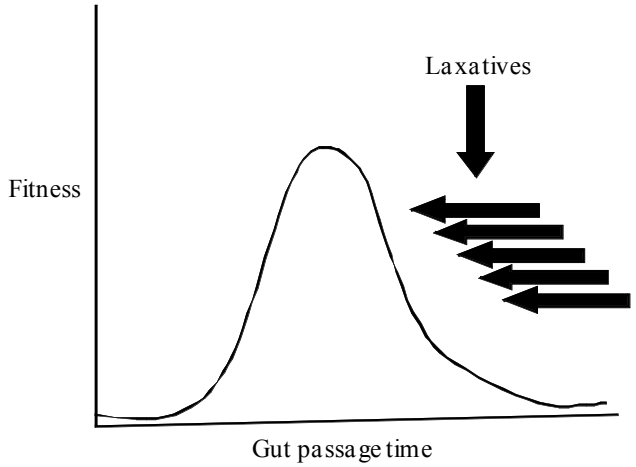

Fig. (6). Relationship between the gut passage time of seeds and fitness. Too slow or too fast a gut passage time will be selected against. Plants would be selected to provision laxatives in their seeds that help achieve an optimum gut passage time.

Several other studies from the same laboratory have used similar evolutionary rationale in predicting and bioprospecting plants for proteinase inhibitors [79] and for anti-helminitic activity [*unpublished*]. In a related context, Lokesha *et al.* (1992) [80] predicted that in prospecting for seed oil content a higher "hit" would be obtained by mining wind dispersed species. For a given constant calorific requirement of seeds, wind dispersed species will be selected to preferentially pack its energy more in the form of fat than in the form of starch (to save space and hence to attain a higher degree of buoyancy for dispersal) than a species whose seeds are passively dispersed (the dispersal in such species is not dependent on seed mass). Analyzing the seed oil content of a large database of plants, Lokesha *et al.* (1992) [80] were indeed able to confirm their predictions: wind dispersed species tend to have a significantly high seed oil content compared to seeds of species dispersed by other means. Though this study was done much before the recent interest on biofuels in plants, it is nevertheless clear that bioprospecting for biofuels following the evolutionary rationale suggested by Lokesha *et al.* (1992) [80] would have tremendous implications in the search for biofuel rich plants. These studies are perhaps some of the first to demonstrate comprehensively the effectiveness of bioprospecting using an evolutionary rationale.

Besides the examples described above, it appears that several studies in literature might have actually used an evolutionary rationale without in fact explicitly realizing it. We discuss here few examples and argue how application of an evolutionary/ecological logic might have driven these discoveries.

The first example is drawn from the now historic discovery of HMG-CoA reductase inhibitors [81, 82]. Soon after the biosynthetic pathway of cholesterol was elucidated by Fisher [83, 84], it was evident that one of the ways by which the total cholesterol content could be lowered was by inhibiting the enzyme, HMG co-reductase, which catalyses the conversion of HMG-CoA to mevalonate. Akira Endo looked up to fungi as possible sources of HMG-CoA reductase inhibitors [81, 82]. The immediate motivation seems to have been derived from the knowledge that a) one of the components of the fungal cell wall is ergosterol, a chemical relative of cholesterol, and b) that the pathway of synthesis of ergosterol could be similar to that of cholesterol. Accordingly, he hypothesized that fungi, such as the common rot fungi, would be selected to secrete compounds that inhibit the synthesis of ergosterol in competing co-occurring fungi. True to his expectation, it was found that *Penicillium citrinum*, indeed secreted compounds that inhibited cholesterol biosynthesis. In retrospect, it appears that this was a brilliant deduction of an evolutionary strategy at work – research in the last couple of decades on such a premise, opened up what may be regarded as one of the biggest discoveries that lead to the development of the multi-billion dollar statins in the world market [85].

Another example pertains to vampire bats and the discovery of the potent plasminogen activators (PA) from their salivary secretions [86]. Vampire bats depend on a diet of fresh blood which they obtain by inflicting an apparently painless wound on their victims. These wounds continue to bleed for several hours. Why do the wounds not heal at a time comparable with other bleeding injuries? From an evolutionary perspective, it is relatively straightforward to propose that animals such as vampire bats would be selected to inject certain anti-blood clotting factors when they bleed their prey – in order that they have an uninterrupted supply of their forage, the blood. Indeed it is now known that vampire bats contain highly potent plasminogen activators (PA), which are specialized, in rapid hydrolysis of fresh blood clots [86]. Several studies have indeed suggested that these activators are superior over human counterparts in their pharmacological and toxicological properties [87]. It follows from the above, that if one is interested in bioprospecting for compounds that interfere with hemostasis, the probable candidates to look for would be organisms such as the vampire bats and other blood sucking creatures; in fact the discovery of hirudin from the medicinal leech is a case in point in this direction [25].

In a number of instances, cues about the underlying processes could be obtained by mere simple natural history observations. For example, the bird Corsican blue tits (*Parus caeruleus*) bring each evening bits and pieces of leaves of several aromatic plants to their nests. Could this intriguing behavior have a larger ecological drive that might help foster bioprospecting efforts? Lafuma *et al.* (2001) [88] examined the nature of the leaves and found that these effectively repel blood sucking flying insects, including mosquitoes – a strategy clearly aimed at protecting the vulnerable nestlings. Clearly, the birds seem to have adopted this behavior in order to maximize their brood fitness. From the point of bioprospecting however, it is easy now to relate how a perceived evolutionary/ecological feature could be incorporated into the process of bioprospecting.

The association between ants (Attini: Formicidae) and their fungal gardens has been known for years. The ants bring in green leaves into their fungal gardens and cultivate the fungi for food. Most often the fungal gardens are maintained free of any contamination, though at times they are infected with a specialized and virulent parasitic micro fungus in the genus, *Escovopsis*. Under some conditions, *Escovopsis* can completely overrun the fungal gardens. How do ants protect their fungal gardens from the marauding *Escovopsis* fungi? Answer to this question came in the form of a discovery of one more mutualistic partner, an actinobacteria that produces dentigerumycin, a selective and new antibiotic which defends the fungal gardens from *Escovopsis* [89]. Some what similar to the ant-fungal association is the association between beewolf and its symbiotic bacteria. Beewolf hunt bees to feed their larvae which live in warm and humid environment in the soil tunnel made by the adult females. The larval habitat makes them vulnerable to attack by pathogenic bacteria and fungi. Recently, Kroiss *et al.* (2010) [90] demonstrated that beewolf wasps cultivate specific symbiotic bacteria (*Streptomyces spp*.) along with their larvae. The symbiotic bacterium produces a mixture of 9 different antibiotics that help ward off any infection from bacteria or fungi. In both these examples, local exigencies during the evolutionary history of the respective organisms have favoured mechanisms that help enhance the lifetime fitnesses of the respective organisms. Closer examination of such case studies can lead to generic extrapolations that might have implications for bioprospecting.

Off late there has been some interest in bioprospecting insects for their anti-microbial activity. A number of chemical molecules, peptide and non-peptides alike, have been found in insects which have both, antibacterial and antiviral properties. A few cytotoxic chemicals have been also reported which have anti-tumour activity. The most notable among these are Alloferon-2, an antiviral and anti-tumour compound isolated from the blowfly, *Calliphora vicina*; pyrrhocoricin, an antibacterial peptide from *Pyrrhocoris apteris*, a bug; ETD 151, an antifungal compound extracted from a South American butterfly larva and Cecropin, the very first antibacterial compound from Cecropia moth. The search for antimicrobials from insects has been so successful that a candidate compound with activity against the superbug, *Staphylococcus aureus* has been reported from Australia [91]. From a bioprospecting perspective, how does one go about prioritizing insects in search of anti-microbial activity? Chandrashekara *et al.* (2009) [92] suggested several criteria that were ecologically rooted. Among others he proposed that insects living under "challenging" conditions such as microbe rich detritus habitats and in crowded conditions such as seen in social insects could be expected to be selected to possess anti-microbial activity compared to insects that do not experience the above challenges.

Our final example relates to the recent discovery of glucagon-like-peptide (GLP) in the Gila monsters [93]. Gila monsters spend most of their time resting in the nest and feed much less frequently compared to other reptiles. Given these features of their basal metabolism and feeding habits, is it possible that Gila monsters have evolved a strategy to regulate their sugar metabolism? A few years ago, it was discovered that Gila monsters posses GLP-4 (a glucagon-like-peptide), in addition to normal GLP-1. Normally GLP-1 increases glucose dependant secretion of insulin by pancreatic β-cells, slows down gastric emptying time and reduce food intake. Unlike GLP-1, GLP-4 of Gila monsters is expressed only in the saliva and was shown to increase rapidly after feeding, peaking within 2 h before decreasing to pre-feeding levels within 24·h [93]. Although, the precise role of GLP-4 in Gila monsters is not yet known, it is speculated that GLP-4 may be involved in modulating the gut morphology to accommodate large volumes of food, some times almost equal to its body weight and at lean times. GLP-4 has a longer half-life *in vivo*. Thus from a purely evolutionary responses, it could have been predicted that there could be something queer in the sugar metabolism in Gila monster that can potentially have some bioprospecting value. Indeed, GLP-4 with the trade name, exenatide-4 has now been released to treat type II diabetes and obesity [94].

In summary, though one of the least explored, evolutionary/ecological rationale based approach offers a rich science-based method to prospect biological diversity. Several earlier workers have indeed recognized this feature, though not as explicitly as stated here [4, 95, 96]; for example, both Wynne-Edwards (2000) [96] and Coley *et al.* (2003) [4] argued that the evolutionary biology of plant defenses against herbivory could hold great promise in directing drug discovery and bioprospecting efforts. An obvious advantage of the evolutionary rationale is the high degree of selectivity in assay results. Because evolutionary inter-relationships are likely to have shaped the evolution of a suite of chemicals/chemistry, bioprospecting based on these leads would elicit a larger proportion of receptor matches than those arising from random screening or for the matter from combinatorial chemistry [97]. In fact, Muller *et al.* (2004) [97] attribute the high selectivity due to "evochemistry" or evolutionary chemistry. Further, the evolutionary adaptations of organisms and their relationships in the ecological web of life may have also set up numerous ecological correlates, an investigation of which can hold great promise in bioprospecting. But it remains to be reiterated that this approach is not always obvious for use; as stated by Caporale (1995) [25] it is very important to ask better questions and further on understand the evolutionary history of the process in question; only then can exciting evolutionary interactions between and among organisms be explored for obtaining interesting bioprospecting leads.

Phylogenetic Rationale

Phylogenetic rationale is based on the fact that phylogeny (relation among species through genetic descent) establishes an evolutionary link between the different taxa, and thereby provides the grounds for the discovery of common metabolic pathways and empirical biomedical prospecting for natural compounds. In fact, in the absence of any other information, the phylogenetic search algorithm (based on the taxonomic neighbourhood) can be regarded as the first route to discovery of alternative sources of existing secondary metabolites, novel biological molecules and activities [98, 99]. The success of this approach rests on the assumption that there is a phylogenetic inertia for the compounds being searched. Several earlier workers have successfully used the phylogenetic route to discover alternative sources for existing high value metabolites [100].

A commonly cited example of the predictive value of the phylogenetic approach concerns the cancer treatment drug, taxol [100]. Taxol was originally extracted from the bark of the Pacific yew, *Taxus brevifolia* Nutt. High demand for taxol required discovery of another source of the compound. An alternative source of taxol (the precursor baccatin III) was found quickly and efficiently by searching for the compound in *T. baccata*, a close relative of *T. brevifolia* [100]. In this example, a naive, random screening for taxol that might have taken years was streamlined into a simple task by incorporating the principle of genetic descent with modification.

In the spiny rayed fishes, previously unreported occurrences of venom have been discovered using a phylogenetic approach [101]. Using explicit phylogeny of spiny rayed fishes and their prior knowledge on distribution of venomous species, an attempt was made to predict the occurrence of venoms in previously unreported species. Using the phylogenetic strategy, in addition to already known 200 venomous species, >1200 species from 11 different clades were predicted to be venomous. To test the effectiveness of the phylogenetic prediction, museum specimens of spiny rayed fishes were examined for the presence or absence of venom delivery structure and a conspicuous venom gland. Interestingly clades that were phylogenetically predicted to be venomous indeed had conspicuous venom glands and delivery structures suggesting all species from these clades produce venoms [101]. In this case, the phylogenetic approach not only helped in discovery of additional venomous species but was accomplished in an efficient manner.

In a similar context an attempt was made to determine the pattern of distribution of several well known plant secondary metabolites by mapping their occurrences on to the angiosperm phylogenetic trees and examine if these were monophyletic or polyphyletic in origin [99]. Occurrence of quinolizidine alkaloids, for example, was found to be restricted to Papilionoidae of Fabaceae; Steroidal alkaloids, on the other hand, are predominant in the genera *Solanum* and *Lycopersicon*, iridoids in Lamioidae, volatile monoterpenes and sesquiterpenes in Nepetoideae; in other words most plant secondary metabolites were phyogenetically conserved in their distribution [25, 98, 99, 102]. From the point of bioprospecting, it is easy to visualize how further search efforts can be effectively conducted by just targeting the additional species or populations within those monophyletic groups for discovery of newer compounds or its derivatives.

The phylogenetic rationale, in principle, is analogous to a special case of statistical sampling referred to as "cluster or adaptive" sampling [103]. Simply stated, if one has struck a rare event/sample, say an individual of rare tree species, the

probability of encountering the next individual of the same species would be highest closest to where the initial sample was found. Analogously, if a novel molecule or bioactivity has been obtained in a certain organism or taxa, it pays to intensify further bioprospecting search in the taxonomic neighborhood or clade of the organisms than elsewhere. In a way it reminds us of James Black's prescription for drug discovery, "the most fruitful basis of discovery of a new drug is to start with an old drug".

In summary, the phylogenetic rationale could be put to use to bioprospect for a range of products including searching for known chemical moieties in unknown floristic systems to unraveling newer sources of a certain bioactivity. In recent years this has been greatly facilitated by the development of molecular phylogenetic tools and the subsequent rapid build up of angiosperm phylogenetic trees (APG III) [104]. An unparalled access to information on the phylogenetic relatedness of plants coupled with the fact that the basic structural and controlling elements (metabolites and their converting enzymes; transcription factors) and systems (allosteric regulation) are conserved [105] provides an interesting arena for directed bioprospecting.

Data-Mining

"I would rather be vaguely right than be precisely wrong"- J. Milton Keynes

Data mining or knowledge discovery is a process of analyzing data with the aim of realization of some useful information or pattern. While certainly not an exact substitute for primary data, it nevertheless helps set broad patterns that could be useful. In the recent past, especially with the dawn of chemo-informatics, data mining has been increasingly used in the area of drug discovery and drug design [106]. For example, attempts have been made to identify the pharmacophore properties of compounds by data mining chemical libraries based on ligand pharmacophores and natural product diversity [107, 108]. However its use in the process of bioprospecting itself is not yet very widespread. In the recent past several opportunities have opened up for bioprospecting using data mining approaches thanks to the steady development of a large number of data-base libraries due to an increasingly accessible chemical space [109, 110]. Among a few major global data sources are the NAPRALERT, NCI discovery resources, Plants for Future, Plant derived drugs database, Rainforest plant database etc. The NAPRALERT data base for instance is one of the largest available relational databases of plants pro-viding information of the ethnomedical, chemical, pharma-cological features of about 45,000 plant species. Fabricant and Farnsworth (2001) [46] used the database to prospect for plant species containing sweet tasting principles. Cordell *et al.* (2001) [31] data mined information on alkaloids in plants and reported some very interesting patterns that have the potential to accelerate the pace of discovery. In a more recent study, Srirama *et al.* (2008) [111] examined the relation between crown gall infection and its association with anti-cancer activity using a data-mining approach as a potential tool to short list plants for their anti-cancer activity. They showed that plants with anti-cancer activity have a higher proportion of species resistant to crown gall than randomly selected species. Thus, it appears that plants intrinsically resistant to crown gall infection could, in principle at least, also be associated with anti-cancer activity. These results have immense exploratory potential in the search for newer plants as sources of anti-cancer activity.

An interesting recent application of data mining exercise revealed that people with Down's syndrome rarely get tumours [112]. Based on this output, investigation showed that an extra copy of the gene DSCR1 on chromosome 21 (that is duplicated in the Down's syndrome people) inhibits the spread of mouse and human tumours. The gene suppresses the growth of new blood vessels by blocking the activity of the protein calcineurin. In summary the simple data mining exercise has now opened up the possibility of devising a new target for future cancer drugs [112]. In yet another study, Epstein (2009) [113] showed how relatively simple data mining can help formulate drug development hypotheses and thus reduce empirical reliance on expensive pre-clinical and early-phase clinical trials.

An important pre-requisite to data-mining of course is the need for a hypothesis that would drive the data to yield useful information. Thus, it is desirable to arm ourselves with good questions which can then be turned over to data mining for answers. In the recent past there have been many developments in the area of data mining that could be potentially used in the area of bioprospecting. For example, besides, manual data mining techniques, one can resort to techniques such as statistical, machine learning or even neural network models. Each of them would of course depend on the need and the magnitude of number crunching that would be required. Finally it needs to be borne in mind that data mining tools can only be suggestive and hence would require further studies to own up the relation brought forth by data mining exercise.

CONCLUSIONS

"The fox knows many things, but the hedgehog knows one big thing" Archilochus

Sir Alexander Fleming [114] in his Nobel lecture said quite candidly, "In my first publication I might have claimed that I had come to the conclusion, as a result of serious study of the literature and deep thought, that valuable antibacterial substances were made by moulds and that I set out to investigate the problem. That would have been untrue and I preferred to tell the truth that penicillin started as a chance observation. My only merit is that I did not neglect the observation and that I pursued the subject as a bacteriologist. My publication in 1929 was the starting-point of the work of others who developed penicillin especially in the chemical field".

We have come a long way since Alexander Fleming made that famous remark in his Nobel lecture. *Or have we really?* While preparing this manuscript, the authors had an opportunity to critically review the state of bioprospecting research in the recent past and analyze the underlying process through which these were actually conducted. A significant proportion of the papers had no *a priori* hypothesis; in other words these papers appeared to merely hoping that they would land some "hits" by sheer chance. In case some of these papers indeed had an underlying

rationale, this was not explicitly spelt out. It is strange but true that strategizing search on well accepted principles of scientific methodology – either inductive or deductive reasoning has not been the hallmark of bioprospectors. Moreover, original observations, such as the one made by Fleming, are widely lacking these days, probably due to a lack of biochemical insight and valuable associations. While we will refrain here from discussing the cause of this shortcoming, it needs to be mentioned that this lack of approach has clearly costed bioprospecting its scientific

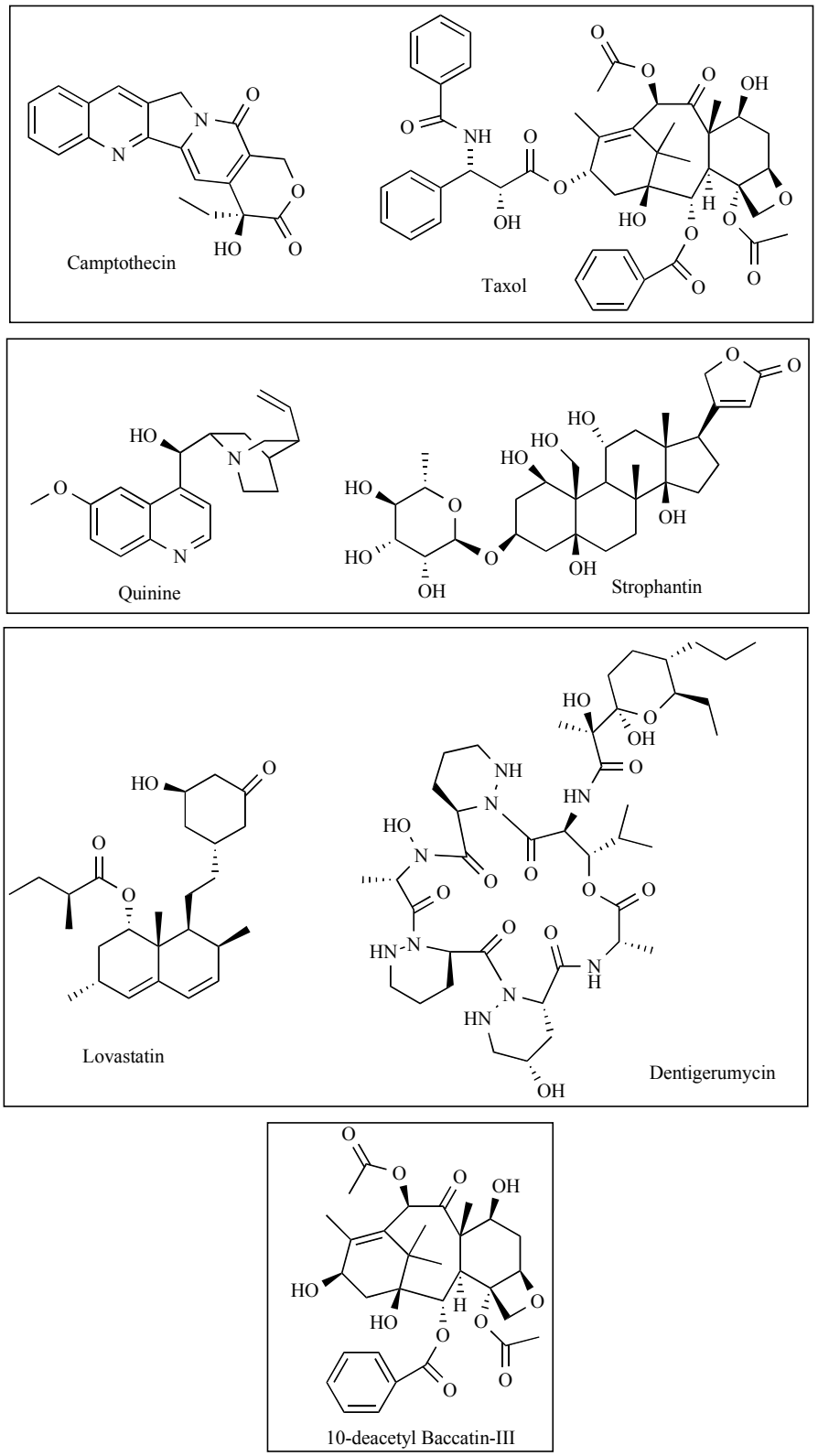

Fig. (7). Representative compounds whose discovery is attributed to different search strategies (random, traditional knowledge, evolutionary rationale, phylogenetic rationale). See text for explanation.

element. This situation was only aggravated with the arrival of brute force, high throughput screening programs, a domain catalyzed by synthetic and combinatorial chemistry. The latter further drained whatever little science that was associated with bioprospecting. It appeared that for bioprospecting, the end mattered more than the means. Thus the random or directionless search methods were often clueless with respect to the search for novel bioactivities or compounds. With the low hanging fruits already plucked, serendipitous and occasional burst of enlightenment are not among the best of strategies to hope for in pursuit of the high hanging fruits.

It is against this general backdrop, that we have made a case for formally strategizing search algorithms for bioprospecting. Given biological resources with a hypothetical chemical domain, we assume that it should be possible to arrive at a plausible hypothesis and thereafter to verify the hypothesis towards exploiting such information for bioprospecting. Among the various strategies that have been discussed, we would like to in particular emphasize that, based on the evolutionary rationale or ecological associations [Fig. (**7**)]. This rationale is firmly entrenched in the assumption that all chemistry is shaped by an organism's evolutionary history, interactions and adaptation [96-99]. Thus, by understanding the evolutionary significance of the compound, in theory, it should be possible to make sound predictions of its modes of action and extrapolate it to the bioprospecting platform. At least to the extent current science allows, one can extrapolate scenarios for bioprospecting. For example, could one have predicted the existence of omega conotoxins from cone snails that block calcium channels – that offer succor to neuropathic pain in AIDS and cancer patients? [115]. In recent years, efforts are being made to move away from a largely random or chance process to more targeted approach. For example, a genome mining approach is being applied to decode chemical structures from genome and hence to bioprospect for organisms that may have these chemicals [116].

Not all search strategies need to necessarily be straightjacketed to conform to one of the four categories that have been discussed. Following Gould who said, "Nature's oddities are my bread and butter" – virtually anything that makes an impression and appears interesting is worthy of pursuit, though one might not always be sure if it would lead to a breakthrough. Take the case of the short-tail shrew that with its bite is able to paralyze (not kill) its prey which can be stored for a few weeks. How does the shrew manage this? Stewart and his co-workers found that this is due to venom that the shrew injects into its victim along with its saliva. The venom was later found to contain a "paralytic peptide", which is now finding application in the control of neuropathic pain [117]. It is now known that certain digger wasps provision their nests before laying eggs with caterpillars that have been paralyzed (not dead). Are the processes similar and could this also have a bioprospecting application? We strongly believe that chemical ecology should be at the forefront of drug discovery and there are many important observations to be made.

The smell of cooked Basmati rice (grown in parts of India) is uncannily similar to that emanating from tiger's urine and has been attributed to 2-acetyl pyrroline [118]. While a stretch of imagination suggests that the tiger is likely to be using this to mark his territory, it is not clear why basmati rice should have this chemical. Investigation into the process and seeking the evolutionary significance of the compound, in the light of positive and negative selection pressures, can inform possible bioprospecting scenarios.

So "Are miracle cures out there in rainforest, coral reefs or in the deep oceans?" We believe that the answers to such questions reside in our ability to ask questions that provoke the surfacing of the evolutionary significance of the systems and their chemistries. We agree with Triggle (2009) [119] who laments that the fault in the discovery process "lies not in our molecules, but in our way of approaching them". In summary, just as mammoth random combinatorial libraries are slowly giving way to more smaller and focused libraries, there is a need to be more precise in handling biodiversity for bioprospecting. A random selection of plants or other organisms is likely to lead to more noise than signal and would be highly cost ineffective. On the other hand, as Rausser and Small (2000) [120] argue, the probabilities of successful "hits" are likely to increase if better and informed science directs the screening programme.

ACKNOWLEDGEMENTS

This article has originated based on a discussion between US and JG at the ETH Zurich, Switzerland. Many of the ideas expressed in the article are due to interactions and discussions with a number of research students and lab colleagues at the School of Ecology and Conservation, University of Agricultural Sciences, Bangalore. In particular, the authors wish to acknowledge the inputs from a brainstorming session on bioprospecting (Nov 13-14, 2009) held at Bangalore. Work of US reported here, have been supported by the Department of Biotechnology, Government of India.

REFERENCES

[1] Balick MJ, Elisabetsky E, Laird SA. Medicinal resources of the tropical forest: biodiversity and its importance to human health. New York: Columbia University Press 1996.

[2] Craig GM, Newman DJ. Discovery and development of antineoplastic agents from natural sources. Cancer Invest 1999; 17: 153-63.

[3] Cox PA. Ensuring equitable benefits: the Falealupo covenant and the isolation of anti-viral drug prostratin from a Samoan medicinal plant. Pharm Biol 2001; 39: 33-40.

[4] Coley PD, Heller MV, Aizprua R, *et al.* Using ecological criteria to design plant collection strategies for drug discovery. Front Ecol Environ 2003; 1(8): 421-8.

[5] Nicolaou KC, Chen JS, Dalby SM. From nature to the laboratory and into the clinic. Bioorg Med Chem 2009; 17(6): 2290-303.

[6] Fransworth NR, Bingel AS. In: Wagner H, Wolff P, Eds. Problems and prospects of discovering new drugs from higher plants by pharmacological screening. Berlin, Springer. 1977; 1-22.

[7] Stevenson WG, Friedman PL. In: Hennekens CH, Ed. Clinical trials in cardiovascular disease. Philadelphia, WB Saunders Co. 1999; 217-30.

[8] Macilwain C. When rhetoric hits reality in debate on bioprospecting. Nature 1998; 392: 535-40.

[9] Soejarto DD. In: Kinghorn AD, Balandrin MF Eds. Logistics and politics in plant drug discovery. Washington, American Chemical Society 1993; 96-111.

[10] Gertsch J. How scientific is the science in ethnopharmacology? Historical perspectives and epistemological problems. J Ethnopharmacol 2009; 122(2): 177-83.

[11] Newman DJ, Cragg GM, Snader KM. Natural products as sources of new drugs over the period 1981-2002. J Nat Prod 2003; 66(7): 1022-37.

[12] Newman DJ, Cragg GM. Natural products as sources of new drugs over the last 25 years. J Nat Prod 2007; 70(3): 461-77.

[13] Li JW, Vederas JC. Drug discovery and natural products: end of an era or an endless frontier? Science 2009; 325(5937): 161-5.

[14] Service RF. Surviving the blockbuster syndrome. Science 2004; 303(5665): 1796-9.

[15] Garcia P. Discovery by serendipity: a new context for an old riddle. Found Chem 2009; 11: 33-42.

[16] Verheij H, Robeson BL. The role of chemical diversity in drug discovery, Genomic/Proteomic Technology-2 September/October 2002; pp 32, 34-35.

[17] Baltz RH. Marcel Faber Roundtable: Is our antibiotic pipeline unproductive because of starvation, constipation or lack of inspiration? J Ind Microbiol Biotechnol 2006; 33: 507-13.

[18] Taleb NN. Fooled by randomness: the hidden role of chance in life and in the markets. Random house publishing group, 2008.

[19] Floyd CD, Leblanc C, Whittaker M. Combinatorial chemistry as a tool for drug discovery. Prog Med Chem 1999; 36: 91-168.

[20] Kaiser M, Wetzel S, Kumar K, Waldmann H. Biology-inspired synthesis of compound libraries. Cell Mol Life Sci 2008; 65(7-8): 1186-201.

[21] Tan G, Gyllenhaal C, Soejarto DD. Biodiversity as a source of anticancer drugs. Curr Drug Targets 2006; 7(3): 265-77.

[22] Liu X, Ashforth E, Ren B, *et al.* Bioprospecting microbial natural product libraries from the marine environment for drug discovery. J Antibiot (Tokyo) 2010; 63: 415-22.

[23] Gani OA, Engh RA. Protein kinase inhibition of clinically important staurosporine analogues. Nat Prod Rep 2010; 27(4): 489-98.

[24] Pietra F. Biodiversiy and Natural Product Diversity, 2002, Pergamon, Tetrahedron Organic Chemistry Series, pp 366.

[25] Caporale LH. Chemical ecology: A view from the pharmaceutical industry. Proc Natl Acad Sci U S A 1995; 92: 75-82.

[26] Harvey AL, Waterman PG. The continuing contribution of biodiversity to drug discovery. Curr Opin Drug Discovery Dev 1998; 1: 71-6.

[27] Tulp M, Bohlin L. Functional *versus* chemical diversity: is biodiversity important for drug discovery? Trends Pharmacol Sci 2002; 23(5): 225-31.

[28] Becerra JX, Noge K, Venable DL. Macroevolutionary chemical escalation in an ancient plant-herbivore arms race. Proc Natl Acad Sci U S A 2009; 106(43): 18062-6.

[29] Becerra JX. Synchronous coadaptation in an ancient case of herbivory. Proc Natl Acad Sci U S A 2003; 100(22): 12804-7.

[30] Becerra JX. Insects on plants: macroevolutionary chemical trends in host use. Science 1997; 276: 253-6.

[31] Cordell GA. Quinn-Beattie ML, Farnsworth NR. The potential of alkaloids in drug discovery. Phytother Res 2001; 15(3): 183-205.

[32] Douwes E, Crouch NR, Edwards TJ, Mulholland DA. Regression analyses of southern African ethnomedicinal plants: informing the targeted selection of bioprospecting and pharmacological screening subjects. J Ethnopharmacol 2008; 119(3): 356-64.

[33] He L, Orr GA, Horwitz SB. Novel molecules that interact with microtubules and have functional activity similar to Taxol. Drug Discovery Today 2001; 6: 1153-64.

[34] Altmann KH, Gertsch J. Anticancer drugs from nature-natural products as a unique source of new microtubule-stabilizing agents. Nat Prod Rep 2007; 24: 327-57.

[35] Gershenzon J, Dudareva N. The function of terpene natural products in the natural world. Nat Chem Biol 2007; 3: 408-14.

[36] Tulp M, Bohlin L. Rediscovery of known natural compounds: Nuisance or goldmine? Bioorg Med Chem 2005; 13: 5274-82.

[37] Forbey JS, Harvey AL, Huffman MA, Provenza FD, Sullivan R, Tasdemir D. Exploitation of secondary metabolites by animals: A response to homeostatic challenges. Integr Comp Biol 2009; 49(3): 314-28.

[38] McGrath P, Li CQ. Zebrafish: a predictive model for assessing drug-induced toxicity. Drug Discovery Today 2008; 13: 394-401.

[39] Lam CS, Rastegar S, Strahle U. Distribution of cannabinoid receptor-1 in the CNS of Zebrafish. Neuroscience 2006; 138: 83-95.

[40] Chianese R, Cobellis G, Pierantoni R, Fasano S, Meccariello R. Non-mamalian vertebrate models and the endocannabinoid system: relationships with gonadotropin-releasing hormone. Mol Cell Endocrinol 2008; 286: 846-51.

[41] Gertsch J, Pertwee RG, Di Marzo V. Phytocannabinoids beyond the Cannabis plant - do they exist? Br J Pharmacol 2010; 160(3): 523-9.

[42] Verpoorte R. Exploration of nature's chemodiversity: the role of secondary metabolites as leads in drug development. Drug Discovery Today 1998; 3: 232-8.

[43] Jones CG, Firn RD, Malcolm SB. On the evolution of secondary chemical diversity. Philos Trans R Soc London Ser B 1991; 333: 273-80.

[44] Bhakuni DS, Dhar ML, Dhar MM, Dhavan BN, Gupta B, Srimal RC. Screening of Indian plants for biological activity-III. Indian J Exp Biol 1971; 9: 91-102.

[45] Suffness M, Douros J. Current status of the NCI plant and animal product program. J Nat Prod 1982; 45: 1-14.

[46] Fabricant DS, Farnsworth NR. The value of plants used in traditional medicine for drug discovery. Environ Health Perspec 2001; 109: 69-75.

[47] Wall ME, Wani MC. Camptothecin and taxol: from discovery to clinic. J Ethnopharmacol 1996; 51: 239-54.

[48] Morgan RE, Westwood NJ. Screening and synthesis: high throughput technologies applied to parasitology. Parasitology 2004; 128: 871-9.

[49] Pandey RC. Prospecting for potentially new pharmaceuticals from natural sources. Med Res Rev 1998; 18(5): 333-46.

[50] Spjut RW. Relationships between plant folklore and antitumor activity: an historical review. Sida 2005; 21(4): 2205-41.

[51] Patwardhan B, Vaidya ADB, Chorghade M. Ayurveda and natural products drug discovery. Curr Sci 2004; 86(6): 789-99.

[52] Patwardhan B, Warude D, Pushpangadan P, Bhatt N. Ayurveda and traditional Chinese medicine: A comparative overview. eCAM 2005; 2(4): 465-73.

[53] Derek L. Merck bails on natural products. Available at: http://pipeline.corante.com/archives/2008/05/08/merck_bails_on_natural_products.php [accessed May 8, 2008].

[54] Farnsworth NR. The role of ethnopharmacology in drug development. Ciba Found Symp 1990; 154: 2-11.

[55] Farnsworth NR. Ethnopharmacology and drug development. Ciba Found Symp 1994; 185: 42-51.

[56] Hamilton GR, Baskett TF. In the arms of Morpheus: the development of morphine for postoperative pain relief. Can J Anaesth 2000; 47: 367-74.

[57] Sneader W. Drug prototypes and their exploitation. UK; Wiley. 1996.

[58] Pelletier PS, Caventon J. [Title unknown]. Ann. Chim. Phys 1820; 14: 69.

[59] Gerhardt C. Untersuchungen über die wasserfreien organischen Säuren. Annalen der Chemie und Pharmacie 1853; 87: 149-79.

[60] McKenzie AG. The rise and fall of Strophanthin. International Congress Series 2002; 1242: 95-100.

[61] Farnsworth NR, Akerele O, Bingel AS, Soejarto DD, Guo Z. Medicinal plants in therapy. Bull World Health Organ 1985; 63(6): 965-81.

[62] Slish DF, Ueda H, Arvigo R, Balick MJ. Ethnobotany in the search for vasoactive herbal medicines. J Ethnopharmacol 1999; 66 (2): 159-165.

[63] Carvalho LH, Kettli AU. Antimalarial chemotherapy with natural products and chemically defined molecules. Mem Inst Oswaldo Cruz 1991; 86(2): 181-4.

[64] Cox PA, Balick MJ. The ethnobotanical approach to drug discovery. Sci Am 1994; 270(6): 82-7.

[65] Spjut RW, Perdue RE Jr. Plant folklore: a tool for predicting sources of antitumor activity? Cancer Treat Rep 1976; 60(8): 979-85.

[66] McManus OB, Harris GH, Glangiacorno KM, *et al.* An activator of calcium-dependent potassium channels isolated from a medicinal herb. Biochemistry 1993; 32: 6128-33.

[67] Hamburger M, Hostettmann K. Bioactivity in plants: the link between phytochemistry and medicine. Phytochemistry 1991; 30: 3864-74.

[68] Cox PA. In: Chadwick DJ, Marsh J Eds, The ethnobotanical approach to drug discovery: strengths and limitations. Ciba Found Symp. Chichester, United Kingdom. John Wiley & Sons. 1994; 25-36, discussion 36-41.

[69] Khafagi IK, Dewedar A. The efficiency of random versus ethno-directed research in the evaluation of Sinai medicinal plants for bioactive compounds. J Ethnopharmacol 2000; 71(3): 365-76.

[70] Cordell GA, Colvard MD. Some thoughts on the future of ethnopharmacology. J Ethnopharmacol 2005; 100: 5-14.

[71] Cordell GA. Changing strategies in natural product chemistry. Phytochemistry 1995; 40(6): 1585-1612.

[72] Chin YW, Balunas MJ, Chai HB, Kinghorn AD. Drug discovery from natural sources. AAPS J 2006; 8(2): E239-53.

[73] Thurston EL, Lersten NR. The morphology and toxicology of plant stinging hairs. Bot Rev 2008; 35(4): 393-412.

[74] Fu HY, Chen SJ, Chen RF, Ding WH, Kuo-Huang LL, Huang RN. Identification of oxalic acid and tartaric acid as major persistent pain-inducing toxins in the stinging hairs of the nettle, *Urtica thunbergiana*. Ann Bot 2006; 98: 57-65.

[75] Alves RRN, Rosa IML. Biodiversity, traditional medicine and public health: where do they meet? J Ethnobiol Ethnomed 2007; 3: 14.

[76] Uma Shaanker R, Ravishankar KV, Ganeshaiah KN. Why do plants posses laxatives? Curr Sci. 1997; 73(8): 646-7.

[77] Ebadi MS. Pharmacodynamic basis of herbal medicine. CRC Press, 2007.

[78] Gunawardene NR. *et al.* A brief overview of the Western Ghats-Sri Lanka biodiversity hotspot. Curr Sci 2007; 93(11): 1567-72.

[79] Hebbar R, Sashidhar VR, Uma Shaanker R, Udaya Kumar M, Sudarshana L. Dispersal mode of species influences the trypsin inhibitor levels in fruits. Naturwissenschaften 1993; 80: 519-21.

[80] Lokesha R, Hegde SG, Uma Shaanker R, Ganeshaiah KN. Dispersal mode as a selective force in shaping chemical composition of seeds. Am Nat 1992; 140(3): 520-5.

[81] Endo A, Kuroda M, Tsujita Y. ML-236A, ML-236B, and ML-236C, new inhibitors of cholesterogenesis produced by *Penicillium citrinum*. Jpn J Antibiot 1976; 29: 1346-8.

[82] Stossel TP. The discovery of statins. Cell 2008; 134(6): 903-5.

[83] Bloch K, Rittenberg D. On the utilization of acetic acid for cholesterol formation. J Biol Chem 1942; 145: 625-36.

[84] Bloch K. The biological conversion of cholesterol to pregnanediol. J Biol Chem 1945; 157: 661-6.

[85] Tobert JA. Lovastatin and beyond: the history of the HMG-CoA reductase inhibitors. Nat Rev Drug Discovery 2003; 2: 517-26.

[86] Hawkey C. Plasminogen activator in the saliva of the vampire bat *Desmodus rotundus*. Nature 1966; 211: 434-5.

[87] Schleuninh WD. Vampire bat plasminogen activator DSPA-Alpha-1(Desmoteplase): A thrombolytic drug optimized by natural selection. Haemostasis 2001; 31: 118-22.

[88] Lafuma L, Lambrechts MM, Raymond M. Aromatic plants in bird nests as a protection against blood sucking flying insects. Behav Processes 2001; 56: 113-20.

[89] Oh DC, Poulsen M, Currie CR, Clardy J. Dentigerumycin: a bacterial mediator of an ant-fungus symbiosis. Nat Chem Biol 2009; 5: 391-3.

[90] Kroiss J, Kaltenpoth M, Schneider B, *et al.* Symbiotic Streptomycetes provide antibiotic combination prophylaxis for wasp offspring. Nat Chem Biol 2010; 6: 261-3.

[91] Chernysh S, Kim SI, Bekker G, *et al.* Antiviral and antitumor peptides from insects. Proc Natl Acad Sci U S A 2002; 99(20): 12628-32.

[92] Chandrashekara K, Jayappa J, Kumar ARV. Insects as source of novel antimicrobial peptides: towards an ecological and evolutionary rationale for bioprospecting insects. International Symposium on Innnovative Agricultural Biotechnology, Oct 5-9, 2009, National Pingtung University, Pingtung, Taiwan.

[93] Christel CM, DeNardo DF, Secor SM. Metabolic and digestive response to food ingestion in a binge-feeding lizard, the Gila monster (*Heloderma suspectum*). J Exp Biol 2007; 210(19): 3430-9.

[94] Mason EE. Gila monster's guide to surgery for Obesity and Diabetes. J Am Coll Surg 2008; 206(2): 357-60.

[95] Beattie AJ, Ehrlich PR. Wild solutions: how biodiversity is money in the bank. New Haven, CT: Yale University Press 2001.

[96] Wynne-Edwards EK. Evolutionary biology of plant defences against herbivory and their predictive implications for endocrine disruptor susceptibility in vertebrates. Environ Health Perspect 2001; 109(5): 443-8.

[97] Müller WE, Schröder HC, Wiens M, Perović-Ottstadt S, Batel R, Müller IM. Traditional and modern biomedical prospecting: Part II-the benefits: approaches for a sustainable exploitation of biodiversity (Secondary Metabolites and Biomaterials from Sponges). Evid Based Complement Alternat Med 2004; 1(2): 133-44.

[98] Dahlgren RMT, Rosendal-Jensen S, Nielsen BJ. In: Young DA, Seigler DS, Eds. A revised classification of the angiosperms with comments on correlation between chemical and other characters. NY, Praeger, 1981; 149-204.

[99] Wink M. Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. Phytochemistry 2003; 64: 3-19.

[100] Kingston DGI. Recent advances in the chemistry of Taxol. J Nat Prod 2000; 63: 726-34.

[101] Smith WL, Wheeler WC. Venom evolution widespread in Fishes: a phylogenetic road map for the bioprospecting of *Piscine* venoms. J Hered 2006; 97(3): 206-17.

[102] Cronquist A. The evolution and classification of flowering plants. New York Botanical Garden, NY, USA 1988.

[103] Thompson SK, Seber GAF. Adaptive Sampling. John Wiley and Sons, Inc., New York, 1996.

[104] THE ANGIOSPERM PHYLOGENY GROUP. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Bot J Linn Soc 2009: 161: 105-21.

[105] Muller WEG, Batel R, Schröder HC, Muller IM. Traditional and modern biomedical prospecting: Part I-the history and sustainable exploitation of biodiversity (Sponges and Invertebrates) in the Adriatic Sea in Rovinj (Croatia). Evid Based Complement Alternat Med 2004; 1(1): 71-82.

[106] Warr WA, Some Trends in Chem(o)informatics. Methods Mol Biol 2011; 672: 1-37.

[107] Markt P, Feldmann C, Rollinger JM, *et al.* Discovery of novel CB2 receptor ligands by a pharmacophore-based virtual screening workflow. J Med Chem 2009; 52(2): 369-78.

[108] Rollinger JM, Schuster D, Danzl B, *et al.* In silico target fishing for rationalized ligand discovery exemplified on constituents of *Ruta graveolens*. Planta Med 2009; 75(3): 195-204. Erratum in: Planta Med 2009; 75(3): 293.

[109] Nguyen KT, Syed S, Urwyler S, Bertrand S, Bertrand D, Reymond JL. Discovery of NMDA glycine site inhibitors from the chemical universe database GDB. ChemMedChem 2008; 3(10): 1520-4.

[110] Wagner BK, Clemons PA. Connecting synthetic chemistry decisions to cell and genome biology using small-molecule phenotypic profiling. Curr Opin Chem Biol 2009; 13(5-6): 539-48.

[111] Srirama R, Ramesha BT, Ravikanth G, Uma Shaanker R, Ganeshaiah KN. Are plants with anti-cancer activity resistant to crown gall? : A test of hypothesis. Curr Sci 2008; 95: 1407-8.

[112] Baek KH, Zaslavsky A, Lynch RC, *et al.* Down's syndrome suppression of tumor growth and the role of the calcineurin inhibitor DSCR1. Nature. 2009; 459(7250): 1126-30.

[113] Epstein RJ. Unblocking blockbusters: Using Boolean text-mining to optimize clinical trial design and timeline for novel anticancer drugs. Cancer Informatics 2009; 7: 231-8.

[114] Fleming A. Penicillin, Nobel Lecture, 11 Dec 1945. In Nobel Lectures: Physiology or Medicine 1942-1962 (1964), 83.

[115] Nelson L. Venomous snails: One slip and you're dead. Nature 2004; 429: 798-9.

[116] Bachmann BO. Decoding chemical structures from genomes. Nat Chem Biol 2005; 1: 244-5.

[117] Kita M, Nakamura Y, Okumura Y, *et al.* Blarina toxin, a mammalian lethal venom from the short-tailed shrew *Blarina brevicauda*: Isolation and characterization. Proc Natl Acad Sci USA 2007; 101: 7542-7.

[118] Brahmachary RL, Sarkar MP, Dutta J. The aroma of rice and tiger. Nature 1990; 344: 26.

[119] Triggle DJ. The chemist as astronaut: Searching for biologically useful space in the chemical universe. Biochem Pharmacol 2009; 78(3): 217-23.

[120] Rausser G, Small A. Valuing research leads: bioprospecting and the conservation of genetic resources. J Polit Econ 2000; 108(1): 173-206.

Received: June 29, 2010 Revised: January 5, 2011 Accepted: February 02, 2011

PMID: 21561422