# Mapping genetic diversity of *Phyllanthus emblica*: Forest gene banks as a new approach for *in situ* conservation of genetic resources

**R. Uma Shaanker and K. N. Ganeshaiah***

Departments of Crop Physiology and *Genetics and Plant Breeding, University of Agricultural Sciences, Bangalore 560 065, India

**We have initiated efforts in identifying the 'hot-spots' of genetic variation of a few important medicinal plant species in south India. In this paper, we present our work on *Phyllanthus emblica*, one of the most important medicinal plant species in the subcontinent. Genetic diversity of four populations of *P. emblica* from south India were studied using isozyme analysis. Based on our results, we propose strategies for the long-term *in situ* conservation of the genetic resources of *P. emblica*. We introduce the concept of 'Forest gene banks' as a viable alternative to the existing methods of *in situ* conservation of genetic resources. In these gene banks, genetic diversity from different sources can be maintained *in situ* and allowed to 'evolve' with the exchange of gene pools within and among populations in the natural habitats.**

THE Indian subcontinent constitutes a rich repository of medicinal plants that are used by various indigenous health care systems. Over 7000 species of plants are estimated to be used for medicinal purposes[1–3]. They form an important source of income to many forest dwelling communities[4–6]. However, because of an indiscriminate use of these resources over time and fragmentation of habitats, many of these species are increasingly threatened and face the risk of becoming genetically impoverished. In south India alone, it is estimated that about 70 to 80 out of 300 species are either severely endangered or threatened (pers. commun., Foundation for the Revitalization of Local Health Traditions, Bangalore). It is imperative that viable strategies to conserve the surviving populations and their genetic resources of at least the critically important species are formulated to arrest further loss. However, for an effective *in situ* conservation, it is important to identify the 'hot-spots' of genetic diversity; this can be done in two steps. First, an extensive geographic distribution map of the species needs to be developed to identify sites with viable population sizes. Second, among these sites, populations that are genetically rich need to be identified. Such sites can then be considered for *in situ* conservation of genetic resources. While there have been attempts to map the geographic distribution of the medicinal plants, that of identifying the 'hot-spots' of genetic variation has been singularly lacking.

Recently we have begun to map the distribution of genetic diversity of a set of medicinal plant species, namely, *Phyllanthus emblica*, *Terminalia bellerica*, *T. chebula*, *Sida rhombifolia* and *Embelia ribes* over south India. In this paper, we focus on *Phyllanthus emblica*, one of the most important medicinal plant species. Based on our results, we propose a viable long-term strategy for the *in situ* conservation of genetic resources of *Phyllanthus emblica* in particular and of other medicinal plants in general.

## Materials and methods

### *System*

*Phyllanthus emblica* L. (syn. *Emblica officinalis* Gaertn., Euphorbiaceae), a small to moderately sized tree, is one of the most important medicinal plant species[1–3]. It is indigenous to tropical south east Asia and occurs mainly in the dry or moist deciduous forests of central and southern India[7–10]. Its fruits along with those of *Terminalia bellerica* and *T. chebula* constitute the well-known ayurvedic drug 'Triphala', used as a purgative. The trees flower during January–March and fruit in November–December. The flowers are predominantly pollinated by wind[11,12]. The fruit is a berry with six nutlets.

### *Sampling*

Thirty Medicinal Plant Conservation Areas (MPCAs) spread over the three southern states, Karnataka, Tamilnadu and Kerala have been established by the Foundation for the Revitalization of Local Health Traditions (FRLHT), Bangalore, for *in situ* conservation of medicinal plant species. Of these, 13 are located in the dry or moist deciduous forest, the likely habitat for *Phyllanthus*. We selected seven of these sites such that they are

geographically isolated and represent diverse bio-geographic strata; of these, viable populations were found only in four sites (Figure 1).

At each of the four MPCA sites, 10 to 15 fruits were collected from at least 25–50 randomly selected trees. The fruits were sun dried and the seeds from each fruit processed separately.

### *Isozyme analysis*

At least ten randomly selected seeds from each tree of every site were sown in petri dishes containing moistened filter paper. Five or six days after germination the seedlings were extracted in Chase extraction buffer[13]. The extracts were absorbed on wicks (Whatman no. 3) and stored in eppendorf tubes at 35°C.

Several enzyme systems (ACO, ADH, ALD, AAT, DIA, GlDH, GuDH, GDH, GLY-3-P, HK, IDH, MDH, ME, MR, NADH DH, NADPH DH, 6-PGDH, PGM, PGI, SKDH, SuDH, TPI) were assayed. Of these, six enzyme systems (PGI, 6-PGDH, IDH, MDH, MR and TPI; for abbreviations see Table 1) that exhibited consistently good staining and polymorphism were selected for the analysis of the populations of *Phyllanthus* over the four MPCAs.

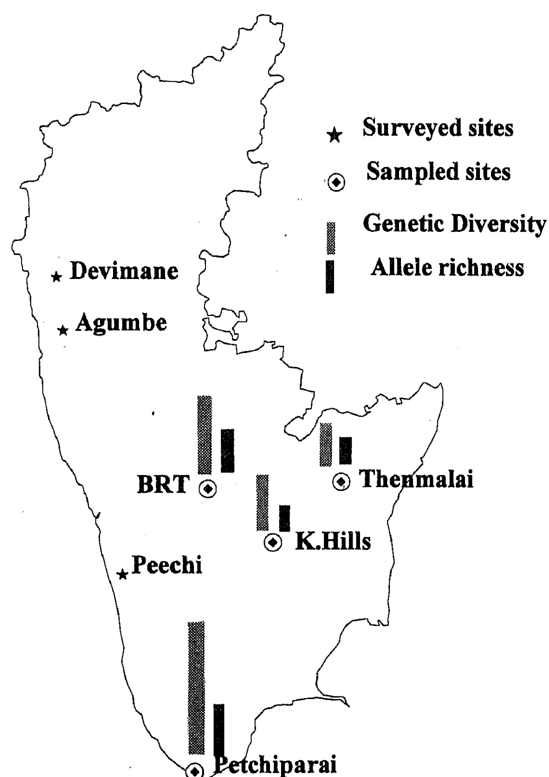


**Figure 1.** Medicinal Plant Conservation Areas (MPCAs) in south India showing the sites surveyed and from where samples were drawn for estimating the genetic diversity of *Phyllanthus emblica*. The relative magnitudes of genetic diversity and allele richness are depicted as histograms.

**Table 1.** Frequency of alleles of the various isozymes for populations of *Phyllanthus emblica*

| | | Allele frequency | | | | |
|---|---|---|---|---|---|---|
| Alleles | | BRT | Then | Kol | Pet | Mean |
| 1 | pgi1 | 0.61 | 0.44 | 0.97 | 0.89 | 0.69 |
| 2 | pgi2 | 0.88 | 0.86 | 0.72 | 0.73 | 0.81 |
| 3 | pgi3 | 0.34 | 0.00 | 0.00 | 0.21 | 0.13 |
| 4 | pgi4 | 0.00 | 0.00 | 0.00 | 0.05 | 0.006 |
| 5 | pgi5 | 0.77 | 0.65 | 0.86 | 0.68 | 0.74 |
| 6 | pgi6 | 0.97 | 0.83 | 0.97 | 0.94 | 0.93 |
| 7 | pgd1 | 0.27 | 0.41 | 0.64 | 0.73 | 0.47 |
| 8 | pgd2 | 0.93 | 0.72 | 0.91 | 0.89 | 0.86 |
| 9 | pgd3 | 0.72 | 0.76 | 0.97 | 0.89 | 0.82 |
| 10 | pgd4 | 0.04 | 0.02 | 0.00 | 0.00 | 0.02 |
| 11 | idh1 | 0.38 | 0.25 | 0.08 | 0.05 | 0.22 |
| 12 | idh2 | 0.45 | 0.58 | 0.64 | 0.52 | 0.55 |
| 13 | idh3 | 0.29 | 0.23 | 0.27 | 0.42 | 0.28 |
| 14 | mdh1 | 0.90 | 0.04 | 0.00 | 0.78 | 0.39 |
| 15 | mdh2 | 0.02 | 0.90 | 0.97 | 0.63 | 0.61 |
| 16 | mdh3 | 0.93 | 1.00 | 1.00 | 0.94 | 0.97 |
| 17 | mdh4 | 0.72 | 0.60 | 0.75 | 0.84 | 0.71 |
| 18 | mdh5 | 0.06 | 0.00 | 0.02 | 0.05 | 0.03 |
| 19 | mr1 | 0.27 | 0.13 | 0.81 | 0.39 | 0.38 |
| 20 | mr2 | 0.04 | 0.18 | 0.16 | 0.23 | 0.14 |
| 21 | mr3 | 0.38 | 0.39 | 0.37 | 0.39 | 0.38 |
| 22 | mr4 | 0.97 | 0.95 | 0.97 | 0.92 | 0.96 |
| 23 | mr5 | 0.43 | 0.20 | 0.59 | 0.65 | 0.43 |
| 24 | mr6 | 0.09 | 0.00 | 0.08 | 0.07 | 0.05 |
| 25 | tpi1 | 0.52 | 0.47 | 0.62 | 0.63 | 0.54 |
| 26 | tpi2 | 1.00 | 0.93 | 0.97 | 0.94 | 0.96 |
| 27 | tpi3 | 1.00 | 0.81 | 0.94 | 0.94 | 0.92 |
| 28 | tpi4 | 0.84 | 1.00 | 0.94 | 0.94 | 0.93 |
| 29 | tpi5 | 0.56 | 0.74 | 0.72 | 0.73 | 0.68 |
| 30 | tpi6 | 0.97 | 0.88 | 0.94 | 0.94 | 0.93 |
| 31 | tpi2a | 0.02 | 0.00 | 0.00 | 0.36 | 0.05 |
| 32 | tpi1a | 0.00 | 0.00 | 0.00 | 0.39 | 0.05 |
| 33 | pgi1a | 0.00 | 0.00 | 0.00 | 0.05 | 0.006 |
| Diversity index | | 1.373 | 1.340 | 1.353 | 1.426 | |
| Richness | | 30 | 26 | 26 | 32 | |
| Evenness | | 0.929 | 0.947 | 0.956 | 0.948 | |

pgi, phosphoglucose isomerase; pgd, 6, phosphogluconic acid; idh, isocitrate dehydrogenase; mdh, malate dehydrogenase; mr, menadione reductase; tpi, triose phosphate isomerase. Suffixes to the enzyme indicate the bands (alleles). The numbers are from slowest moving (cathodal) to fastest moving (anodal) bands. BRT, BR Hills; Then, Thenmalai; Kol, Kolihills; Pet, Petchiparai.

Samples were electrophoresed at 4°C using 9 per cent starch gel (Sigma, USA) at 150–200 volts and 50 mA for about six hours (gels were allowed to over run to facilitate a better separation of the isozymes). The gel buffer was histidine-HCl (pH 7.0) while the tray buffer was tris-citrate (pH 7.0) (ref. 14). After electrophoresis, the gels were horizontally sliced and stained for the enzymes using the agar-overlay method[15]. The gels were fixed in acetic acid : alcohol and scored for the isozyme bands.

### *Data analyses*

Analysis of the zymograms was done using electromorphs (alleles) as the isozyme descriptors. The presence of a band was scored as 1 and its absence as 0. In a few cases, where the bands could not be scored unambiguously, a value of 0.5 was assigned. From the six polymorphic enzyme systems, 33 different allelomorphs were detected. TPI yielded the highest number of allelomorphs (8), followed by PGI (7), MR (6), MDH (5), 6-PGDH (4) and IDH (3). In the Petchiparai population, three rare alleles, one of PGI and the other two of TPI, were recovered (Table 1).

### *Diversity analysis*

The frequency of different electromorphs (alleles) for the six enzyme systems (totally 33 bands or alleles) was computed for each population. Diversity, evenness and richness indices were computed following Ludwig and Reynolds[16] by using allele frequency for that of species ($p_i$) in the respective formulae.

Squared Euclidean distance (SED) was computed between all combinations of individuals among the four populations following Ludwig and Reynolds[16] as:

$$\mathrm{SED}_{jk} = \mathrm{SQR}\left(\sum_{i=1}^{1}(X_{ij} - X_{ik})\hat{}2\right),$$

where, $X_{i,j}$ and $X_{i,k}$ are the abundance of the $i$th allele in population $j$ and $k$ respectively and 1 is the total number of alleles. From the matrix of these distances, a dendrogram of the four populations was constructed following a minimum variance technique.

The allelomorph data of the populations was subjected to principal component analysis to examine the dispersion of individuals belonging to different sites on a few principal component axes.

### *Relative abundance of alleles*

For each site, the relative abundance of an allele was computed as a ratio of its frequency in the site to that in all the sites. A ratio of 1 suggests that the site is similar to the whole population; if it exceeds 1, then such populations have a relatively higher frequency of the alleles. The coefficient of variation (CV) for the occurrence of an allele within a site was also computed. The relative abundance of an allele was then plotted against their respective CVs. From the point of conservation, populations with high value of relative abundance and high CV are most desired.

### *Mean abundance of all alleles in a site*

We also computed the mean relative abundance of all alleles together in a site and the mean CV for their frequency occurrence. Sites with high mean relative abundance of all alleles but with low CV would merit a greater conservation value than others as they are relatively rich in genetic diversity. Such sites are likely to have a higher representation of most of the alleles in relatively invariant proportions.

## Results and discussion

### *Allele diversity*

The population at Petchiparai (PET) had the highest allelic diversity followed by that from BRT Hills (BRT), Kolihills (KOL) and Thenmalai (THEN; Figure 1, Table 1). Allele richness was also highest for PET (32), followed by BRT Hills (30), KOL (26) and THEN (26). The greater number of alleles in PET seems to be contributed by the occurrence of three additional alleles in this population, one of PGI and two of TPI (Table 1). Thus based on both diversity and richness index, the population at Petchiparai appears potentially promising for any conservation efforts. The evenness index was highest for population from KOL followed by that in PET, THEN and BRT Hills (Table 1).

The first four component axes explained 42.4% of the total variance in the data set. There was a clear segregation of the individuals belonging to BRT Hills from those belonging to THEN and KOL. The population from Petchiparai was distributed across the whole space (Figure 2). The population at BRT and THEN appeared to complement each other in the dispersion. The dispersion pattern was essentially similar when plotted on the other principal component axes. The squared Euclidean distances among the four populations suggest that the populations at THEN and KOL are more closely related than any other pairwise combination of populations. Furthermore, the population at PET pervaded the entire spectrum of characteristics of the populations at BRT Hills, THEN and KOL. Cluster analysis indicated two distinct clusters or groups (Figure 3). The first cluster comprised the populations at THEN and KOL while the second cluster comprised populations at BRT Hills and PET.

### *Relationship between CV for occurrence of an allele and mean relative abundance of an allele*

Relative abundance of most of the alleles was greater at Petchiparai than in other populations (Figure 4 *a–d*). For four alleles (nos 4, 31, 32, 33), the mean ratio of allele

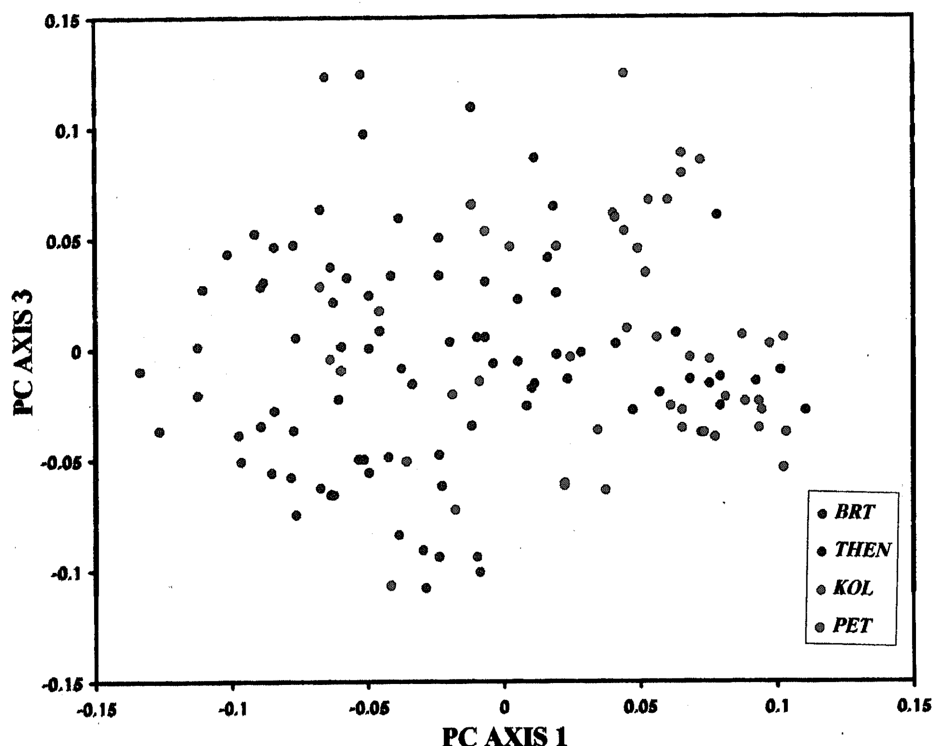


**Figure 2.** Principal component analysis of the populations of *Phyllanthus emblica*. Note that the population from Petchiparai is dispersed all over the principal component space.

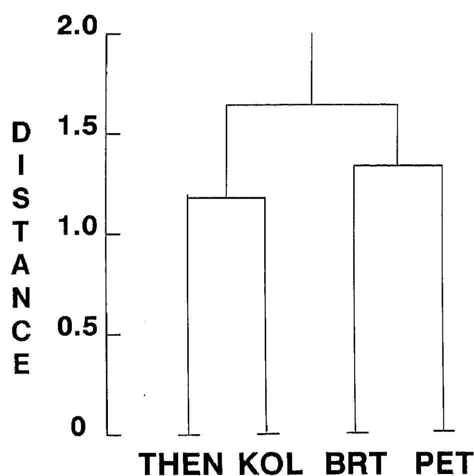


**Figure 3.** Dendrogram of the clustering of various populations of *Phyllanthus emblica*.

frequency was highest compared to other populations; in fact the alleles 31, 32, 33 are present only in the Petchiparai population. Further, a number of alleles were found to exhibit a substantial CV for their frequency. Eleven of the 32 alleles from the PET population had a CV > 100% and a mean ratio of relative abundance > 1. Thus it appears that the population at Petchiparai is genetically very rich yet diverse with respect to the occurrence of the various alleles. This is also apparent from the relation between the mean relative abundance of all the alleles and their CV (Figure 5). The population from PET had the highest mean frequency occurrence of the alleles with a CV corresponding nearly to that of the entire population. Thus PET population appears to be the most important for *in situ* conservation.

### *Implications for in situ conservation of genetic diversity*

Our study shows that on several criteria, the population at Petchiparai is most diverse genetically and represents the entire spectrum of variability existing in other populations. Thus, the Petchiparai population can be regarded as a potential 'hot-spot' of genetic variation of *Phyllanthus* and can be considered for *in situ* conservation. Our studies are concentrated mostly on the MPCAs of FRLHT. We are not sure if there could be other sites more potential than PET. Nevertheless conservation plans for the MPCAs could be based on these results. Accordingly we propose the following alternative strategies for *in situ* conservation of genetic diversity of *Phyllanthus emblica*.

*Strategy I.* As a priority, the population at PET may be targeted for conservation. This would ensure the conser-

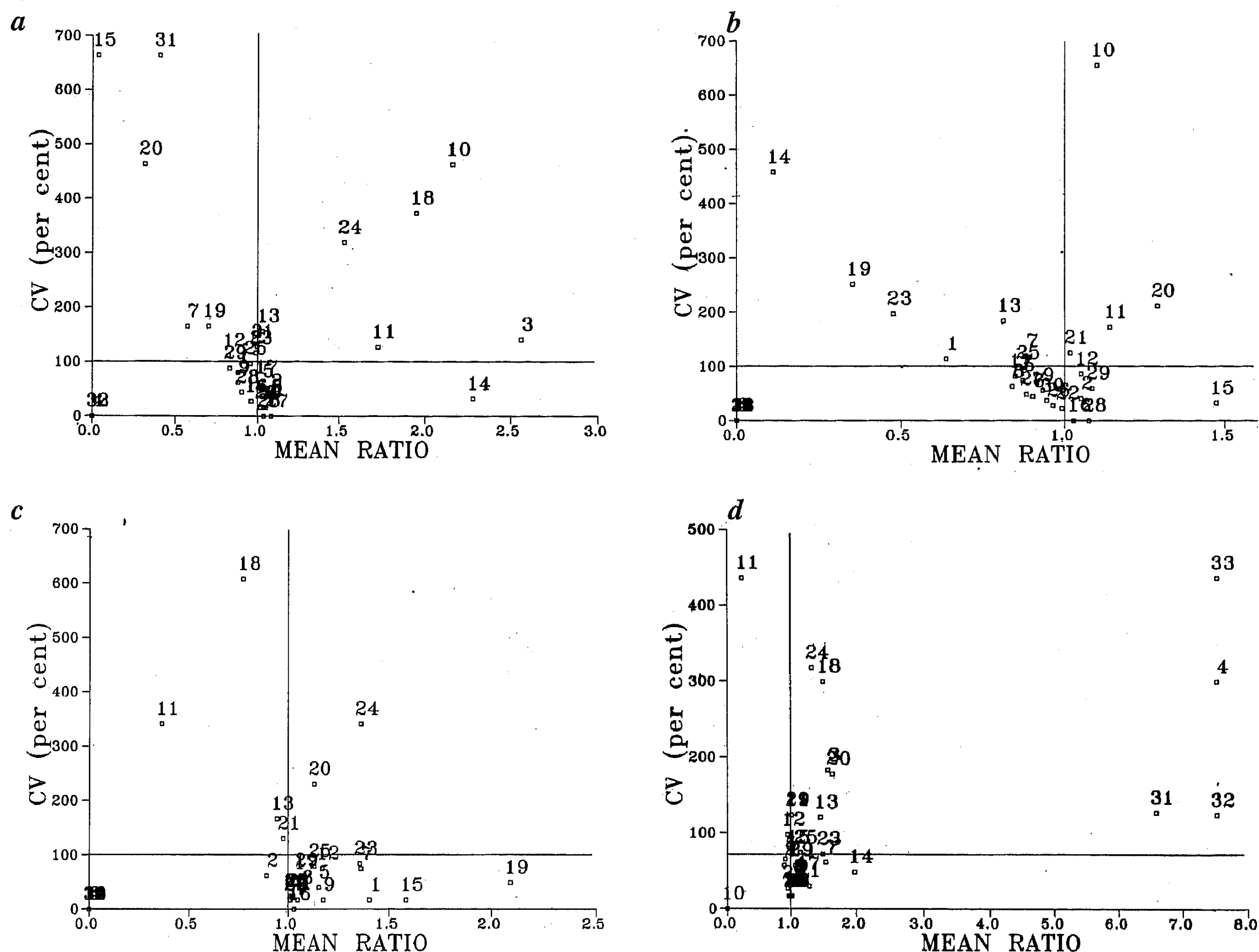


**Figure 4.** Relation between the mean ratio of relative abundance of an allele with the coefficient of variation for the frequency occurrence of the allele, (*a*) BRT Hills, (*b*) Thenmalai, (*c*) Kolihills, and (*d*) Petchiparai. The figures show an arbitary cut-off of coefficient of variation (CV) > 100 % and mean ratio > 1. The alleles in the top right quadrant are most variable with respect to their occurrence within a population but with a higher mean allele abundance relative to the entire population. The numbers refer to alleles (for details see Table 1).

vation of a relatively rich proportion of genetic diversity representative of that existing in other populations. Besides, because of the presence of a few alleles exclusively in this population, conservation at this site also ensures preservation of rare alleles.

*Strategy II.* In the event that the population at PET cannot be conserved, the population at BRT Hills and THEN may be conserved together to represent the entire array of variability.

## Forest gene banks

The above strategies, however, may not still ensure conservation of the complete gene pool of the species. We hence propose establishment of 'Forest gene banks' as a viable alternative to the conventional *in situ* conservation sites, *ex-situ* gardens and field gene banks. We envisage these forest gene banks as *in situ* sinks into which gene pools from various source sites are introduced and maintained so as to serve as a repository of the gene pool of the species. In these banks, apart from maintaining the 'global' allelic set of the species, there would be a continuos turnover of the genetic material within and among populations. In other words, such banks combine the virtues of both *in situ* conservation sites as well as those of the *ex situ* gardens and field gene banks. In these banks, genetic diversity would be allowed to 'evolve' as it would in any other natural habitat.

*Ex-situ* gardens and field gene banks are particularly of limited use in maintaining the gene pool of species that are long-lived and cross-pollinated. Because of the relative insulation of the populations in such conservation sites, mating would be constrained, leading to decline in genetic diversity over time. Further, successive generations at these sites may also express high degree of inbreeding depression. Forest gene banks attempt to overcome these limitations.

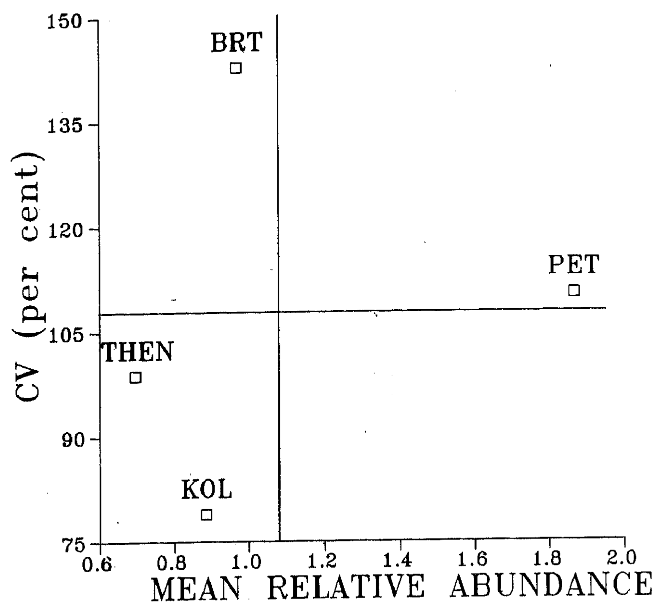


**Figure 5.** Relation between mean abundance of all alleles at a site with the coefficient of variation for the frequency occurrence of the alleles (see text for details).

## *Establishment of forest gene banks*

The establishment of forest gene bank for a species will involve three steps. First, it requires mapping the geographic distribution and genetic diversity of the species. Second, from among these sites, it is necessary to identify the sites which can serve as donors (from where gene pool could be imported) and recipient (a genetic hot-spot and a sink to receive the gene pool from outside). The choice of the donors and recipients can be guided among several criteria, by the extent of genetic variation in the population, the presence of rare alleles, population size, local heterogeneity, etc. Ideally, recipients could be those which are allele rich and have a broad genetic base while donors could be those that contain rare alleles. The actual process of gene exchange can occur either through transfer of seeds and/or pollen depending upon the species. The interval at which such gene exchange can occur may also depend upon the longevity of the species and its breeding system. A number of medicinal plant species face the prospect of being endangered and threatened. In these species, because native populations are highly fragmented and rapidly shrinking, conventional methods of conservation of genetic resources is besieged with problems. In these situations, forest gene banks can serve as a practical and profitable strategy of conserving and importantly enriching the extant genetic resources.

## *Forest gene banks for Phyllanthus emblica*

There are virtually no systematic efforts for the conservation of genetic resources of *Phyllanthus* in the country. For example, as a part of a programme on conservation of genetic resources of arid fruit crops, 16 germplasm accessions of *Phyllanthus emblica* are being maintained in 'field gene banks' in parts of central India[17]. Such limited number of accessions in the gene banks, apart from not being representative of the population genetic variability, shall also lead to severe inbreeding depression over time in cross pollinated species such as *Phyllanthus emblica*. Based on the results of our study, we propose that forest gene bank be established at BRT Hills or Thenmalai. The gene pool from other sources can be incorporated into such forest gene banks. This can be achieved by introducing either pollen or seed material into these sites from Petchiparai. Because of its cross pollinated nature, dusting of pollen grains may be worthwhile though techniques for pollen storage may have to be standardized. The genetic diversity could be monitored with frequent enrichment from other sources.

1. Nadkarni, K. M., *Indian Materia Medica*, Popular Prakashan, Bombay, 1976, vols 1 and 2.
2. Dastur, J. F., *Medicinal Plants of India and Pakistan*, D. B. Taraporevala and Sons, Bombay, 1952.
3. Chopra, R. N., Nayar, S. L. and Chopra, I. C., *Glossary of Indian Medicinal Plants*, Council for Scientific and Industrial Research, New Delhi, 1956.
4. Murali, K. S., Uma Shankar, Uma Shaanker, R., Ganeshaiah, K. N. and Bawa, K. S., *Econ. Bot.*, 1996, **50**, 252–269.
5. Uma Shankar, Murali, K. S., Uma Shaanker, R., Ganeshaiah, K. N. and Bawa, K. S., *Econ. Bot.*, 1996, **50**, 270–279.
6. Sekar, C., Rai, R. S. V. and Ramaswamy, C. J., *Trop. Forest Sci.*, 1996, **8**, 280–288.
7. Firminger, T. A., *Firminger Manual of Gardening for India*. Thacker Soink and Co., Calcutta, 1947.
8. Morton, J. E., *Econ. Bot.*, 1960, **14**, 119–127.
9. Pathak, R. K. and Pathak, R. A., in *Advances in Horticulture* (eds Chadha, K. L. and Pareek, O. P.), Malhotra Publishing House, New Delhi, 1993, vol. 1, pp. 407–421.
10. Arora, R. K. and Nayar, E. R., *Wild Relatives of Crop Plants in India*, NBPGR Sci. Monogr. No. 7, 1984, p. 90.
11. Bajpai, P. N., *Indian J. Hort.*,1968, **7**, 38–67.
12. Reddi, C. S. and Janaki Bai, *New Bot.*, 1979, **4**, 155–160.
13. Murawski, D. A. and Bawa, K. S., *Am. J. Bot.*, 1994, **81**, 155–160.
14. Cheliak, W. M. and Pitel, J. A., *Techniques for Starch Gel Electrophoresis of Enzymes from Forest Tree Species*, Petawa National Forestry Institute, Canadian Forestry Service, 1984.
15. Wickneswari, R. and Norwati, M., in *Breeding Technologies for Acacias* (ed. Aker, C.), ACIAR Proceedings no. 37, 1992.
16. Ludwig, J. A. and Reynolds, J. F., *Statistical Ecology*, John Wiley and Sons, New York, 1988.
17. Pareek, O. P. and Sharma, S., in *Advances in Horticulture* (eds Chadha, K. L. and Pareek, O. P.), Malhotra Publishing House, New Delhi, 1993, vol. 1, pp. 189–225.

ACKNOWLEDGEMENTS. This work was supported by a research grant from the Foundation for Revitalization of Local Health Traditions, Bangalore. Partial support from CIFOR, IPGRI and Department of Science and Technology, Govt. of India is also acknowledged.