# GENETIC STRUCTURE AND BOTTLENECK STUDIES AT THE MALATE DEHYDROGENASE AND PHOSPHOGLUCOISOMERASE (GLUCOSE-6-PHOSPHATE ISOMERASE) GENE LOCI OF NEEM (*AZADIRACHTA INDICA*)

I. K. ASANTE
*Department of Botany, P. O. Box LG55,*
*University of Ghana, Legon, Accra, Ghana*

## Abstract

Genetic structure of the neem at the malate dehydrogenase and phosphoglucoisomerase (glucose-6-phosphate isomerase) gene loci was studied in a total of 192 seedlings raised from seeds collected from eleven maternal plants. Both loci were polymorphic and the mean number of alleles per locus was $5.05 \pm 0.54$. Mean observed was $0.902 \pm 1.4$. Levene's and Nei's expected heterozygosities were $0.804 \pm 0.02$ and $0.801 \pm 0.02$, respectively. Wright's fixation index Fis estimates were negative for all the six alleles for the Mdh locus while for the Pgi locus the estimate was positive for two of the alleles. Analysis of molecular variance (AMOVA) indicated that 17.8 per cent of the total genetic diversity resided among families. The results showed that there was low level of population structure in the neem sample studied at the Mdh and Pgi gene loci and that the 11 families approximate a single panmictic unit. A normal 'L' shaped distribution of mode-shift test and three heterozygote excess tests suggested that there was no recent bottleneck in the population of neem at the two enzyme gene loci.

## Résumé

ASANTE, I. K. : *La structure génétique et l'étude de goulot d'étranglement à la malate dehydrogenase et phosphoglucoisomérase* [*glucose -6- phosphate isomérase*] *gène locus de margousier* (Azadirachta indica). La structure génétique du margousier à la malate dehydrogenase et phosphoglucoisomérase (glucose -6- phosphate isomérase) gène locus était étudiée au total de 192 jeune plantes élevées de graines prises de onze plantes maternelles. Les deux locus étaient polymorphes et le nombre moyen d' allèles par locus était $5.05 \pm 0.54$. Le moyen observé était $0.902 \pm 1.4$. Les hétérozygosites espérés de Levene et de Nei étaient respectivement $0.804 \pm 0.02$ et $0.801 \pm 0.02$. Les estimations Fis de l'index de fixation de Wright étaient négatives pour tous les six allèles pour le locus Mdh alors que pour le locus Pgi l'estimation était positive pour les deux d'allèles. Une analyse MOVAA indiquait que 17.8% de la diversité génétique totale residaient parmi les familles. Les résultats montraient qu'il y avait un niveau bas de la structure de population dans l'échantillon de margousier étudié au gène de locus à Mdh et à Pgi et que les familles onze font approximation d'une unité panmictique singulière. Une distribution en forme de 'L' normal de tests de changement de mode et trois tests de hétérozygote excédent suggéraient qu'il n'y avait pas d' étranglement récent dans la population de margousier aux deux locus de gène enzymatique.

## Introduction

Knowledge of the distribution of genetic variation within and among populations plays a vital role in conserving plant genetic resources (Hedrick, 1985; Brown *et al.,* 1990; Adams *et al.,* 1992).

Morphological and genetic characterization is the initial step towards formulating breeding policies and selecting the appropriate species for conservation in an effective and meaningful way (Gupt *et al.,* 2005). Isozyme technique has been extensively used in the study of many woody plants and crop plants (Hamrick & Godt, 1990; Hamrick *et al.*, 1992; Asante *et al*., 2003). Isozyme as genetic markers has played a major role in plant breeding and conservation of plant genetic resources. Along with DNA markers, isozyme markers have been found very effective in genetic studies. Isozyme technique is also inexpensive for the study of genetic variation in living organisms when compared to other biochemical or molecular techniques.

Neem is a member of the mahogany family, Meliaceae, and botanically known as *Azadirachta indica* A. Juss. Neem has high medicinal values, and insecticidal properties. Neem also produces industrial products, for instance, oil extracted from the seeds is used in the preparation of soaps, waxes and lubricants, as well as fuels for lighting and heating. The solid residue left after the oil is extracted from the kernels is used as a fertilizer and soil amendment. Neem wood is valued for construction, cabinetry and fuel. The bark is tapped for gum and extracted for tannins and dental-care products. The leaves are sometimes used for emergency livestock feed. The profuse flowers are a prized source of honey. There is existence of natural variability in the neem. For instance the level of azadirachtin extracted from the seed varies with source of collection (Morgan, 1981). Neem trees are widely distributed (Tewari, 1992) and this is also an indication of the presence of genetic variability.

Studies related to selection of superior trees, provenances and other genetic improvement in the neem is lacking (Tewari, 1992). There is, therefore, the need to carry out selection and genetic improvement programmes in the neem. However, knowledge of genetic structure and diversity is a prerequisite for selection and genetic improvement programmes. The study, therefore, aims at a preliminary investigation into the genetic structure in neem based on the Mdh and Pgi enzyme gene loci.

### Experimental

*Protein extraction*
A total of 192 seedlings were raised from seeds collected from eleven different maternal trees. Leaves were collected from seedlings at two-leaf stage and crushed in a small mortar with a pestle into a 'paste'. The crude squeezate in the 'paste' was absorbed into strips of Whatman No. 1 paper (about 8 mm × 4 mm), which were blotted carefully and inserted into a cut about 3.5 cm from the intended cathodal end of the starch gel.

*Preparation of starch gel*
Gel was prepared using 12.5 per cent mixture of hydrolyzed starch (from SIGMA) with Tris Continuous Citrate (CTC) buffer, *p*H 8.0, in a flask. The flask was whirled over a gas stove till the mixture was almost translucent gelly, usually at a point where the first large bubble formed from below. The cooked starch was poured quickly but carefully into a gel former. It was covered gently but quickly with a glass plate, and left to form. The gel was transferred into a refrigerator 2-3 h later.

*Running of gel*
The next day, the gel was loaded and placed on an electrophoretic bath assembly filled with CTC buffer. Gauze wicks were soaked in the buffer and connected to the gel at both ends to allow even flow of the current through the gel. The gel was covered with polyethylene sheet on which a glass plate was placed. Ice packs were placed on the glass plate. The covers and the ice pack were used to minimize heating of the gel during the run. The electrophoretic unit was placed in a refrigerator. The cathode electrode was connected

to the end where samples were loaded. Gels were subjected to between 148 and 171 V and a current of between 51 and 57 mA for 3 and 4 h.

*Slicing and staining of gel*

After running a gel, it was carefully trimmed at the sides with a scalpel and carefully transferred onto a glass plate for slicing. A gel was sliced horizontally into slices of 2-mm thick each and put into staining trays. The inner cut surface of gel slice was stained for specific enzyme activity with the following staining recipes: malate dehydrogenase (1.1.1.37): 150 mg L-malic acid, 10 mg NAD, 600 mg Tris, 30 ml distilled water, 6 m MTT, trace PMS; malic enzyme [$NADP^+$](1.1.1.40): 60 mg, DL-malic acid, 10 mg $MgCl_2$, 10 mg NADP, 30 ml Tris-HCl (0.2 *M*) *p*H 8.0, 7 mg MTT, trace PMS; phosphoglucoisomerase [glucose-6-phosphate isomerase](5.3.1.9): 20 mg fructose-6-phosphate, 10 mg $MgCl_2$, 10 mg NADP, 30 ml Tris-HCl (0.2*M*) *p*H 8.0, 50 µl G6PDH, 7 mg MTT, trace PMS.

*Computation and statistical analysis*

The computer program POPGENE (Yeh *et al.*, 1999) was used to estimate genotype frequencies, allele frequencies, heterozygosities (Nei, 1978), Wright's fixation index (Fis) and other genetic variables. The probability of random mating in the neem population was estimated by Chi-square ($\chi^2$) and the likelihood ratio ($G^2$) tests to examine Hardy-Weinberg equilibrium (HWE) at the two loci. Ewens-Watterson neutrality test was used to test the neutrality of the isozyme markers; statistics for the test were calculated by using the algorithm given by Manly (1985), using 1000 simulation samples.

Bottleneck events in the population were tested by two methods. The first method comprised three excess heterozygosity tests developed by Cornuet & Luikart (1996), namely (i) Sign test, (ii) Standardized differences test, and (iii) Wilcoxon sign-rank test. The probability distribution was established using 1000 simulations under three models, Infinite Allele Model (IAM), Stepwise Mutation Model (SMM) and Two-Phase Model of Mutation (TPM). The second method was graphical representation of the mode-shift indicator which was originally proposed by Luikart *et al.* (1998). Loss of rare alleles in bottlenecked populations is detected when one or more of the common allele classes have a higher number of alleles than the rare allele class (Luikart *et al.*, 1998). These two methods were conducted using Bottleneck v1.2.02 (Cornuet & Luikart, 1996). The computer program ALERQUIN (Stefan *et al.,* 2000) was used for analysis of molecular variance (AMOVA).

TABLE 1
*Overall genotype frequencies at the Mdh and Pgi gene loci*

| Genotype | Locus | |
|---|---|---|
| | Mdh | Pgi |
| A/A | 0.000 | 0.444 |
| B/A | 0.000 | 0.000 |
| B/B | 0.000 | 0.065 |
| C/A | 0.068 | 0.000 |
| C/B | 0.005 | 0.185 |
| C/C | 0.000 | 0.087 |
| D/A | 0.250 | 0.174 |
| D/B | 0.151 | 0.109 |
| D/C | 0.052 | 0.000 |
| D/D | 0.000 | 0.000 |
| E/A | 0.089 | 0.022 |
| E/B | 0.099 | 0.185 |
| E/C | 0.099 | 0.087 |
| E/D | 0.031 | 0.000 |
| E/E | 0.000 | 0.000 |
| F/A | 0.010 | 0.000 |
| F/B | 0.047 | 0.000 |
| F/C | 0.078 | 0.043 |
| F/D | 0.021 | 0.000 |
| F/E | 0.000 | 0.000 |
| F/F | 0.000 | 0.000 |

TABLE 2

*Overall allele frequency at the Mdh and Pgi loci*

| Allele\Locus | Mdh | Pgi |
|---|---|---|
| Allele A | 0.208 | 0.141 |
| Allele B | 0.151 | 0.304 |
| Allele C | 0.151 | 0.245 |
| Allele D | 0.253 | 0.141 |
| Allele E | 0.159 | 0.147 |
| Allele F | 0.078 | 0.022 |

### Results and discussion

Two soluble proteins involving two loci were assayed from the leaf extracts. Proteins were malate dehydrogenase (1.1.1.37) and phosphoglucoisomerase dehydrogenase. One polymorphic locus was scored in all 11 families

for each of the proteins.

A total of 13 out of 21 expected different genotype classes were observed at the Mdh locus, while 10 out of 21 expected different genotypic classes were observed at the Pgi locus (Table 1). The highest genotypic frequency for the Mdh locus was 0.25 (*Mdh-D/A*). At the Pgi, the highest genotype frequency was 0.444 (*Pgi-A/A*). Allele frequencies at the Mdh locus ranged between 0.078 (*Mdh-F*) and 0.253 (*Mdh-D*) (Table 2). At the Pgi locus, the range was 0.022 (*Pgi-F*) and 0.304 (*Pgi-B*) (Table 2).

Observed number of alleles at each locus was six, however, the effective number of allele was 5.43 and 4.66 for Mdh and Pgi, respectively, with a mean of $5.05 \pm 0.54$ (Table 3). Shannon's information indices for the Mdh and Pgi gene loci were 1.74 and 1.62, respectively, with a mean of $1.68 \pm 0.08$. Observed heterozygosity at the

TABLE 3

*Observed ($n_o$) and effective ($n_a$) number of alleles, observed heterozygosity ($H_o$), expected heterozygosity ($H_e$) and Shannon's information index (I)*

| Enzyme loci | No. of alleles | | | $H_e$ | | |
|---|---|---|---|---|---|---|
| | $n_o$ | $n_e$ | $H_o$ | Levene's | Nei's | I |
| Mdh | 6 | 5.4 | 1.000 | 0.818 | 0.816 | 1.734 |
| Pgi | 6 | 4.7 | 0.804 | 0.790 | 0.786 | 1.624 |
| Average | 6 | 5.0 | 0.902 | 0.804 | 0.801 | 1.681 |
| SD | 0 | 0.54 | 0.14 | 0.02 | 0.02 | 0.08 |

TABLE 4

*Chi-square and G-square probabilities (95% confidence level), and upper (U95) and lower (L95) 95% confidence limits of expected F values across the Mdh and Pgi gene loci in the neem*

| Enzyme locus | Hardy Weinberg Equilibrium test | | | Ewens-Watterson Neutrality test | | |
|---|---|---|---|---|---|---|
| | Degrees of freedom | Probability Chi-square | Probability G-square | Obs. F | L95 | U95 |
| Mdh | 50 | 0.000 | 0.000 | 0.184 | 0.257 | 0.895 |
| Pgi | 50 | 0.000 | 0.000 | 0.214 | 0.239 | 0.807 |

TABLE 5

*Wright's fixation index as a measure of heterozygote deficiency or excess*

| Allele\Locus | Mdh | Pgi |
|---|---|---|
| Allele A | -0.263 | 0.194 |
| Allele B | -0.178 | -0.130 |
| Allele C | -0.178 | 0.147 |
| Allele D | -0.338 | -0.165 |
| Allele E | -0.085 | -0.022 |
| Allele F | -0.226 | -0.024 |

Mdh and Pgi loci were 1.000 and 0.804, respectively, with a mean of $0.902 \pm 0.13$. Levene's and Nei's mean expected heterozygosities were $0.804 \pm 0.02$ and $0.801 \pm 0.02$, respectively. This information indicated a high genetic variability existing in the neem population at the Mdh and Pgi gene loci. Ewens-Watterson test for neutrality suggested that the two loci (Table 4, observed *F* values lie outside of the upper and lower limits of 95% confidence region of expected *F* values) deviated from neutrality and, probably, are linked to selected traits.

The Chi-square ($\chi^2$) and likelihood ratio ($G^2$) tests indicated that the two loci deviated from HWE (Table 4). Heterozygote excess analysis by

Wright's fixation index revealed deviation from HWE. The basis of this departure is, however, difficult to know. Wright's fixation index (Fis) estimates are presented in Table 5. At the Mdh locus the estimates were negative for all the six alleles, ranging between -0.338 (*Mdh-D*) and -0.085 (*Mdh-F*). At the Pgi locus the range was between -0.022 (*Pgi-F*) and 0.194 (*Pgi-1*). The estimate was, however, positive for alleles *Pgi-A* and *Pgi-C*. Generally, there is evidence of heterozygote excess at the two loci under study, indicating high levels of outbreeding at the two loci.

In the study no evidence for a bottleneck was detected with the two methods. Bottleneck characterization was based on three models of isozyme evolution. Using the Sign test, the expected number of loci with heterozygosity excess was 1.18 for the Infinite Allele Model (IAM) under null hypothesis (Table 6). The probability value in this case was 0.34807 and, therefore, accepts the null hypothesis of mutation-drift equilibrium under this model. However, the expected number of loci with heterozygosity excess was 1.19 and 1.16 under the Two-Phase Model of Mutation (TPM) and Stepwise Mutation Model (SMM) with probabilities 0.35462 and 0.33930, implying acceptance of the null hypothesis using the Sign test. These results

TABLE 6

*Test for null hypothesis under the isozyme evolution models for the Mdh and Pgi gene loci*

| Test/Model | I.A.M | | T.P.M | SMM |
|---|---|---|---|---|
| Sign test | Number of loci with heterozygosity excess (probability) | | | |
| | Expected = 1.18 (0.34807) | | 1.19 (0.35462) | 1.16 (0.33930) |
| | Observed = 2 | | 2 | 2 |
| Standard difference test | Ti (values probability) | | | |
| | 2.411 (0.00795) | | 2.210 (0.01355) | 1.198 (0.02624) |
| Wilcoxon rank test | (Probability of heterozygosity excess) | | | |
| | 0.12500 | | 0.12500 | 0.12500 |

indicate that the neem population studied has not undergone recent genetic bottleneck as a result of mutation-drift equilibrium at the Mdh and Pgi loci.

The standardized difference gave $T_2$ (probability) statistics which were equal to 2.411 (0.00795), 2.210 (0.01355) and 1.939 (0.02624) for the IAM, TPM and SMM models, respectively. The probability values were less than 0.05 in all cases, meaning that the null hypothesis was

Wilcoxon rank test. In view of the results from all the three tests together, it is quite clear that serious bottlenecks have not occurred in the neem tree at the Mdh and Pgi gene loci.

The Mode-shift indicator test was also used to detect potential bottleneck. This was based on the criterion that non-bottleneck populations that are near mutation-drift equilibrium are expected to have a large proportion of alleles with low frequency. This test discriminates many
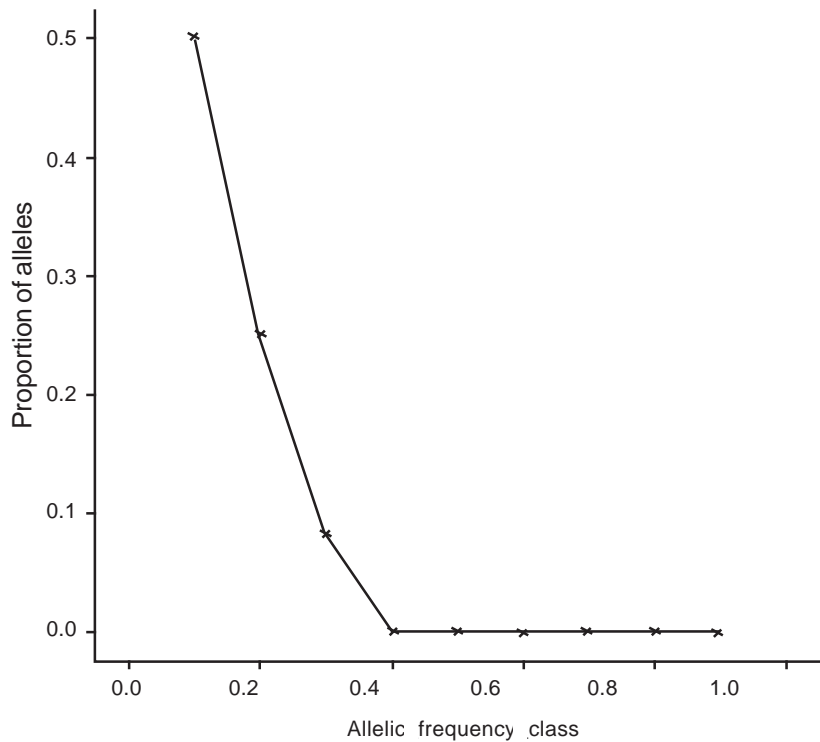


× × × × Proportion of alleles *v* Allelic frequency class

Fig. 1. Proportion of alleles and their distribution in neem

rejected indicating bottleneck under this model. Using the Wilcoxon rank test (a non-parametric test) the probability value was 0.12500 in all the three cases (IAM, TPM and SMM), indicating the acceptance of the null hypothesis under the

bottlenecked populations from stable populations (Luikart, 1997; Luikart & Cornuet, 1997). A graphical representation using allelic class and proportion of alleles showed a normal 'L' shaped distribution (Fig. 1). This distribution actually

TABLE 7

*Analysis of molecular variance for the 11 neem families at the Mdh and Pgi gene loci*

| Source of variation | df | SS | Variance component | % Variation |
|---|---|---|---|---|
| Among families | 10 | 6104.59 | 15.55 (Va) | 17.84 |
| Within family | 381 | 27283.66 | 71.61 (Vb) | 82.16 |
| Total | 391 | 3388.25 | 87.16 | |
| Fixation index | $F_{ST}$: | 0.1784 | | |

Significance tests (1023 permutations)

| Va ad $F_{ST}$ | P(random value > observed value) = 0.000 |
|---|---|
| | P(random value = observed value) = 0.000 |
| | P(random value ≥ observed value) = 0.000 |

shows that the neem population studied has not undergone any recent bottleneck.

The hierarchical AMOVA results indicated that 82.2 per cent of the variation was distributed within families, rather than among families (Table 7). The $F_{ST}$ value of 0.178 was significantly greater than zero; however, it indicates low level of population structure in the neem sample studied at the Mdh and Pgi loci. With reference to the Mdh and Pgi loci, the sharing of one gene pool among the studied population and a lack of barriers to gene flow in the *A. indica* has practical implications and must be considered. Also, the majority of the genetic variation within the studied populations and high genetic similarity among them must be taken into consideration when breeding programmes are established for *A. indica*.

## Acknowledgement

## References

ADAMS, W. T., STRAUSS, S. H., COPES, D. L. & GRIFFIN, A. R. (1992) *Population genetics of forest trees.* Kluwer Academic, Dordrecht and Boston.

ASANTE, I. K., LAING, E., DANKWA, E. Y. & OFFEI, S. K. (2003) Genetic structure at the isocitrate dehydrogenase and malate dehydrogenase enzyme gene loci in cowpea (*Vigna unguiculata* (L.) Walp ). *Ghana Jnl agric Sci.* **36**, 53-63.

BROWN, A. H. D., CLEGG, M. T., KAHLER, A. L. & WEIR, B. S. (1990) *Plant population genetics, breeding and genetic resources.* Sinauer Associates, Sunderland, Massachusett.

BUSH, R. M. & SMOUSE, P. E. (1992) Evidence for adaptive significance of allozymes in forest trees. *New For.* **6**, 179-196.

CORNUET, J. M. & LUIKART, G. (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**, 2001-2014.

Gupta, A. K., Chauhan, S.N., Tandon, S. N. & Sonia, S. N. (2005) Genetic diversity and bottleneck studies in the Marwari horse breed. *J. Gene.* **84** (3), 295-301.

Hamrick, J. K. & GoDt, M. J. W. (1990) Allozyme diversity in plant species. In *Plant population genetics, breeding and genetic resources.* (A. H. D. Brown, M. T. Clegg, A. L. Kahler and B. S. Weir, ed.) Sinauer Associates, Sunderland, Massachusett.

Hamrick, J. L., Godt, M. J. & Sherman-Broyles, S. L. (1992) Factors influencing levels of genetic diversity in woody plant species. *New For.* **6**, 95-124.

Hedrick, P. W. (1985) *Genetics of population.* Jones and Bartlett Publishers, Boston.

Knowles, P. & Grant, M. C. (1981) Genetic patterns associated with growth variability in ponderosa pine. *Am. J. Bot.* **68**, 942-946.

Knowles, P. & Mitton, J. B. (1980) Genetic heterozygosity and radial growth in *Pinus contorta. Silvae Genet.* **29**, 114-118.

Luikart, G. (1997) *Usefulness of molecular markers for detecting population bottlenecks and monitoring genetic change.* (PhD Thesis). University of Motana, Missoula, USA.

Luikart, G. & Cornet, J. M. (1997) Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conserv.*

*Biol.* **12**, 228-237.

Luikart, G. L., Allendorf, F. W., Cornuet, J. M. & Sherwin, W. B. (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *J. Heredity* **89**, 238-247.

Manly, B. F. J. (1985) The *statistics of natural selection.* Chapman and Hall, London.

Morgan, E. D. (1981) Strategy in the isolation of insect central substance from plants, *Proc. Ist. Int. Neem Conf.*, Rottach-Egern, 1980. pp. 43-52.

Nei, M. (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**, 583-590.

Tewari, D. N. (1992) *Monograph on Neem* (Azardirachta indica *A. Juss*). Hyerian Book. 287 pp.

Stefan, S., Roessli, D. & Excoffier, L. (2000) *Arlequin ver. 2000: A software for population genetics data analysis.* Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Weir, B. S. (1979) Inferences about linkage disequilibrium. *Biometrics* **25**, 235-254.

Yeh, F., Boyle, T., Ye, Z. & Xian, J. M. (1999) *POPGENE version 3.31, a Microsoft Windows based free ware for population genetic analysis.* University of Alberta, Edmonton.