Genetic Resources and Crop Evolution **46:** 587–598, 1999.

The genetic structure and conservation of aus, aman and boro rices from Bangladesh

B.J. Parsons[1], H.J. Newbury[1,*], M.T. Jackson[2] & B.V. Ford-Lloyd[1]
[1]*School of Biological Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, U.K.;* [2]*International Rice Research Institute, P.O. Box 3127, Makati Central Post Office, 1271 Makati City, Philippines (***Author for correspondence. Fax: 21 4145925; E-mail: h.j.newbury@bham.ac.uk)*

Received 3 November 1998; accepted in revised form 5 March 1999

Key words: diversity, genetic resources, GIS, *Oryza sativa*, RAPD, rice

Abstract

A diverse set of 115 rice varieties from Bangladesh was surveyed using 35 polymorphic RAPD (randomly amplified polymorphic DNA) markers and the genetic structure of this germplasm, encompassing the principal rice ecotypes of Bangladesh (aus, aman and boro), was determined using multivariate analysis. The level of genetic diversity was evaluated and compared with the levels of diversity found within other rice growing areas of the world. Geographical information systems analysis using Atlas-GIS was employed to analyse and present the geographic distribution of genetic diversity across Bangladesh, and cluster analysis was used to test the efficiency of selection of material for a core collection.

Introduction

Among the major cereals, rice germplasm is probably the most intensively evaluated and the best conserved (Chang, 1989; SGRP, 1996, 1997; Jackson, 1997). Over 102,700 rice accessions are maintained in the International Rice Genebank (IRG) at the International Rice Research Institute (IRRI) in the Philippines. However, the sheer number of accessions maintained increases the complexity of management, characterization and evaluation of the collection (Jackson, 1997).

Morphological and physiological characters, sexual affinity and serodiagnostic techniques have all been used to assess variation in *O. sativa* L. and to subsequently classify rice genetic resources (Kato, 1930; Oka, 1953; Morinaga, 1954). However, morphological and physiological traits are under complex genetic control (Oka, 1953) and subject to environmental effects and hence such markers may not completely represent underlying genetic diversity.

Isozymes have traditionally been used to evaluate genetic variation in *O. sativa*. Glaszmann (1986) presented a detailed evaluation of the genetic structure of *O. sativa* using biochemical variation at 21 isozyme loci. His classification scheme based on multilocus associations among isozyme genes permits a rapid assessment of Asian cultivated rice varieties into six varietal groups (I–VI) with varying environmental and macrogeographic distributions (Glaszmann, 1987, 1988). In our study, RAPD markers (Williams et al., 1990) have been used to assess genetic diversity within a set of rice varieties from Bangladesh. Bangladeshi rice germplasm is particularly worthy of investigation as it provides a rich source of genetic diversity for resistance to a range of biotic and abiotic factors. As part of an FAO land resources appraisal (FAO, 1988), Bangladesh has been divided into a number of growing period zones for both kharif (wet period from approx. March to October) and rabi (dry period from approx. November to February) seasons. A total of 12 kharif growing periods were delineated varying from 180–200 days in the extreme west, where the mean annual rainfall is 1250 mm, to 280–290 days in the extreme east, where mean annual rainfall is 6000 mm. We have focused on kharif zones as most of the varieties considered in this study are grown during the kharif season.

This kharif-rabi seasonal moisture regime has played an important role in shaping the diversity of rice varieties which are differentiated into three prin-

cipal ecotypes in Bangladesh: aus, aman and boro. During the kharif season, the photoperiod-insensitive aus crop is grown. These varieties are usually sown in the pre-monsoon season (March/April) and harvested between July and August, a growing period of 80–120 days. Aus rices are dwarf in stature, thermosensitive, and perform best under summer conditions (Alim, 1974). They are grown under rainfed conditions and are prone to both drought and flooding. Aus crops have low yield and poor quality.

Traditionally, aman varieties have been the main rice crop in Bangladesh. They are sown (broadcast) in March or transplanted following the aus harvest and mature throughout the kharif season. They are photoperiod-sensitive and flower in October/November. Since they have a longer growth duration (120–160 days), aman varieties are more productive than the aus rices and produce high quality, fine white grains.

Boro rices are photoperiod insensitive and are adapted to mild winter conditions (Zaman, 1980) showing some degree of cold tolerance. They are similar to transplanted aman both in their method of cultivation and crop habit. The boro crop is sown in October/November, transplanted around December/January and harvested in the spring. Traditionally, they have only been grown on land which retains sufficient water throughout the rabi season to support crop growth. However, with improved irrigation, these high yielding varieties are increasingly being adopted by Bangladeshi farmers.

Based on crossability and morphological studies, aman varieties are considered as *indica* rices (isozyme group I), whereas the aus varieties are intermediate between *indica* and *japonica* (isozyme group VI) (Morinaga & Kuriyama, 1955, 1958; Takahashi & Hamza, 1983; Ueno et al., 1990). In Bangladesh, all six isozyme groups are encountered and correspondence with local ecotypes has been demonstrated (Glaszmann, 1987) although cultivation of *japonica* rices in Bangladesh is limited. Group II rices are exclusive to South and West Asia and correspond to the aus ecotype and some boro rices. Varieties in group II, together with groups III, IV (minor groups composed of Bangladeshi deepwater varieties) and V (high quality rices), are thought to constitute an alternative intermediate gene pool that is distinct from the rices differentiated into *indica* and *japonica*.

Differentiation of *O. sativa* into aus and aman ecotypes in Bangladesh is thought to have resulted from a mutation for photoperiod reaction, acting as a seasonal isolation mechanism (Morinaga & Kuriyama, 1958) and it has been further speculated that *japonica* rices subsequently evolved from the aus varieties of Bangladesh. However, whether the intermediate types classified by Glaszmann (1986, 1987) and including aus rices constitute primitive *O. sativa* forms which have not yet differentiated into *indica* or *japonica*, or whether they represent the products of *indica-japonica* hybridization is still unclear (Chang et al., 1991). What is evident, however, is that the composition of rice diversity in Bangladesh is complex and hence a reliable means of assessing diversity levels to facilitate germplasm management is needed.

Further to investigating genetic diversity within Bangladeshi rice varieties, we have explored the potential for using GIS (Geographical Information Systems) analysis to investigate the geographical distribution of this diversity across Bangladesh. While the geographical distribution of plant and animal species is increasingly being studied with the aid of GIS, to the best of our knowledge, this paper is the first report where attribute data generated using DNA-based markers have been incorporated into a GIS for spatial analysis.

Finally, core collections are regarded as being a limited set of germplasm accessions from a larger germplasm collection**,** selected on the basis that they are representative of the diversity within the whole collection, and aim at improving the efficiency of management and use of large germplasm collections (Brown, 1995). There are a number of ways that a core set of germplasm can be selected, and in our work we have tested, on a small scale, the efficiency of one possible method.

The aims of this research were therefore: (i) to assess the genetic structure within a diverse set of rice germplasm encompassing the ecotypes of Bangladesh using RAPD markers; (ii) to evaluate the level of genetic diversity existing within this material; (iii) to compare this estimate with the levels of diversity found within other rice growing areas of the world; (iv) to investigate the relationship between the observed genetic variation within this Bangladeshi material and its geographic distribution, and (v) to undertake a preliminary investigation of how a core collection for rice might be developed.

Materials and methods

Plant material and DNA isolation

An ecogeographically diverse set of 115 *O. sativa* varieties originating from Bangladesh was obtained from the IRG (Table 1). The geographic distribution of the diverse rice material across Bangladesh is illustrated in Figure 1. Each variety was selected on the basis of the availability of (i) accurate locational data (latitude-longitude co-ordinates), and (ii) passport, evaluation and characterization data from the International Rice Genebank Collection Information System (IRGCIS). Plants were grown in a glasshouse in Birmingham maintained at 27±2 °C, with a 16 h daylength and 80% relative humidity for approximately 2 weeks. Two mg of leaf material were sampled from each of 10 seedlings per accession, yielding 20 mg of material for DNA isolation. The DNA extraction method adopted was the 4% CTAB procedure described by Virk et al. (1995). DNA was diluted to a final concentration of 2 ng μl^{-1}in SDW and stored at −20 °C.

RAPD protocol

A set of 8 decamer primers (Operon primers K01, K02, K04, K08, K10, K12, K17 and C14) of arbitrary sequence were used to initiate amplifications of PCR products. Primers were selected from a total of 19 decanucleotides pre-screened across a set of 19 rice accessions (data not shown). Selection criteria included: number of RAPD bands amplified and the proportion of which were polymorphic; clarity of the amplification profile; ability to generate markers diagnostic of isozyme crossability groups; and whether bands could be mapped in a doubled haploid population (IR64 × Azucena; Parsons et al., 1997). The RAPD reaction mix, with a total volume of 25 μl, contained ca. 5 ng DNA, with 0.4 μM of each decanucleotide (Operon Technologies, Inc.), 200 μM of each dNTP (Pharmacia Biotech) and 1U *Taq* polymerase (Bioline), maintained in a 1× incubation buffer (Bioline), containing a total of 2.5 mM $MgCl_2$ (Bioline). Samples were covered with 35 μl of mineral oil (Sigma) and processed, using a Hybaid Omnigene thermocycler, through 47 simulated tube-controlled temperature cycles as described in Parsons et al. (1997). All reactions were performed at least twice to check PCR consistency. Amplification products were subjected to 1.4% agarose gel electrophoresis in 0.5 × TBE buffer and were detected by ethidium bromide staining, viewed by fluorescence under UV light. Gels were photographed using a digital imaging system (Flowgen) or Polaroid film.

Data analysis

Specific PCR products that were reproducible in successive amplifications were selected and marker bands were defined by their molecular weights, estimated using 1 kb DNA ladder (Gibco-BRL) size standards. Amplification products were scored as present (1), absent (0) and in a very few cases, ambiguous (9), across all 115 rice accessions. Estimates of diversity were obtained using Nei's gene diversity index (H) (Nei, 1973) and Shannon's Information index (H_s) (Bowman et al., 1971). Average per locus 'gene diversity' indices were calculated by dividing over k polymorphic loci. Indices of similarity between accessions were calculated using Jaccard's coefficient, and the UPGMA clustering algorithm computed using NTSYS-pc (Rohlf, 1992), was used to group accessions. Relationships between accessions were portrayed graphically in the form of a dendrogram.

GIS analysis

To obtain a digitised map outline of Bangladesh, the map sheet (scale 1:750,000) of the 1988 FAO Climatic Resources Inventory (Report 4, Volume I) was digitised using the Arc Digitising System of pc-ARC/INFO (Version 3.4.2; Environmental Systems Research Institute (ESRI) Inc., 1993). Kharif growing period zones were also digitised from this base map. Map coverages were imported into Atlas-GIS (Strategic Mapping Inc., 1994) for analysis of the distribution of genetic diversity of rice accessions across Bangladesh. ARC/INFO and Atlas-GIS user commands are detailed in their respective reference manuals (ESRI Inc., 1993; Strategic Mapping Inc., 1994).

Distribution of diversity of aus rices across Bangladesh

As the majority (70%) of the Bangladeshi rices considered were of the aus ecotype, the pattern of diversity of aus rices only was investigated across Bangladesh. Indices of genetic diversity (H and H_s) were calculated for accessions collected from the different kharif growing period zones.

Table 1. Diverse set of 115 *O. sativa* accessions from Bangladesh

Code	IRGC[a] No	Variety name	Code	IRGC No.	Variety name	Code	IRGC No.	Variety name
S1	25835	Bailam	S42	64768	Bhuita gara	S83	66801	Kaberi
S2	25838	Bathuri	S43	64769	Binna sopa	S84	66802	Kaisha manja
S3	25845	Chandarhat	S44	64771	Chikon shoni	S85	66803	Kalimbom
S4	25850	Da29	S45	64773	Dharia	S86	66804	Kalo buri
S5	25851	Dhalashaita	S46	64774	Dharia boalia	S87	66805	Kalo chengri
S6	25852	Dumai	S47	64775	Dheki shaita	S88	66806	Kalo kuchi
S7	25854	Garia	S48	64778	Holoi bash (soloi bash)	S89	66808	Kharai murali
S8	25865	Jabarshail	S49	64780	Kal shoni	S90	66809	Khoia boro
S9	25868	Jhum fulbadam	S50	64781	Kat gimi	S91	66811	Koi murali
S10	25874	Jhum sonalichikon	S51	64785	Lal agalia oynna	S92	66813	Lakkhi kajal
S11	25877	Kalabail	S52	64786	Lalchi aus	S93	66814	Leja gomvir
S12	25885	Lakhsnikajal	S53	64787	Mi natik	S94	66815	Lenja murali
S13	25892	Mikhudeb	S54	64788	Mi ronguchan	S95	66816	Moisdol
S14	25893	Mikotchu	S55	64789	Moshur	S96	66817	Moshia bhadoi
S15	25897	Mimidam	S56	64790	Moysha-irri	S97	66818	Moyna moti
S16	25898	Mimidim	S57	64791	Mulla dari	S98	66819	Munshi murali
S17	25906	Molladigha	S58	64792	Narikel jhupi	S99	66821	Pathor bhati
S18	25911	Pankhiraj	S59	64793	Rakhoil	S100	66822	Rangpuri aus
S19	25920	Sampatti	S60	64795	Shada aus	S101	66823	Shada kuchi
S20	25924	Sultanjata	S61	64796	Shada shaita	S102	66824	Shandha moni
S21	25925	Tepakain	S62	64799	Sreerampur shaita	S103	66825	Shol buna
S22	26298	Bashiraj	S63	64800	Uttari aus	S104	66826	Siruti
S23	26772	Pura binni	S64	66763	Agir jail biroi	S105	66827	Taraf mon
S24	27513	Dholi boro	S65	66764	Akhni sail	S106	66828	Tepi borua
S25	27515	Ghuni boro	S66	66768	Aus nagra	S107	66829	Vasha murali
S26	27522	Ashmber	S67	66771	Bir mazla	S108	66830	Bhoria aus
S27	27526	Bazail	S68	66774	Chengri murali	S109	66831	Jabor sail
S28	27539	Buri katari	S69	66776	Dhal kachai	S110	66810	Khudey borra
S29	27540	Chakila	S70	66781	Dumai kalo	S111	27510	Bhaturi
S30	27564	Laksmilota	S71	66782	Dumai lal	S112	34711	Lakhi jhota
S31	27581	Shitki saita	S72	66786	Gobir sail	S113	37362	Badali
S32	31611	Noroi	S73	66787	Gochi boro	S114	66794	Hizuli
S33	31754	Chandra har	S74	66788	Gomvir	S115	64767	Bholanath
S34	34682	Boteshawar	S75	66789	Gorba 1			
S35	34712	Lakhi puri	S76	66790	Gorbai			
S36	34722	Mery	S77	66791	Gorbai			
S37	34730	Rangoon	S78	66792	Gul murali			
S38	34732	Serety	S79	66793	Hizli aus			
S39	34737	Bawoi	S80	66795	Honuman jata			
S40	37005	Battiboro	S81	66796	Inda			
S41	64766	Baulan	S82	66800	Juma			

[a] International Rice Genebank Collection.

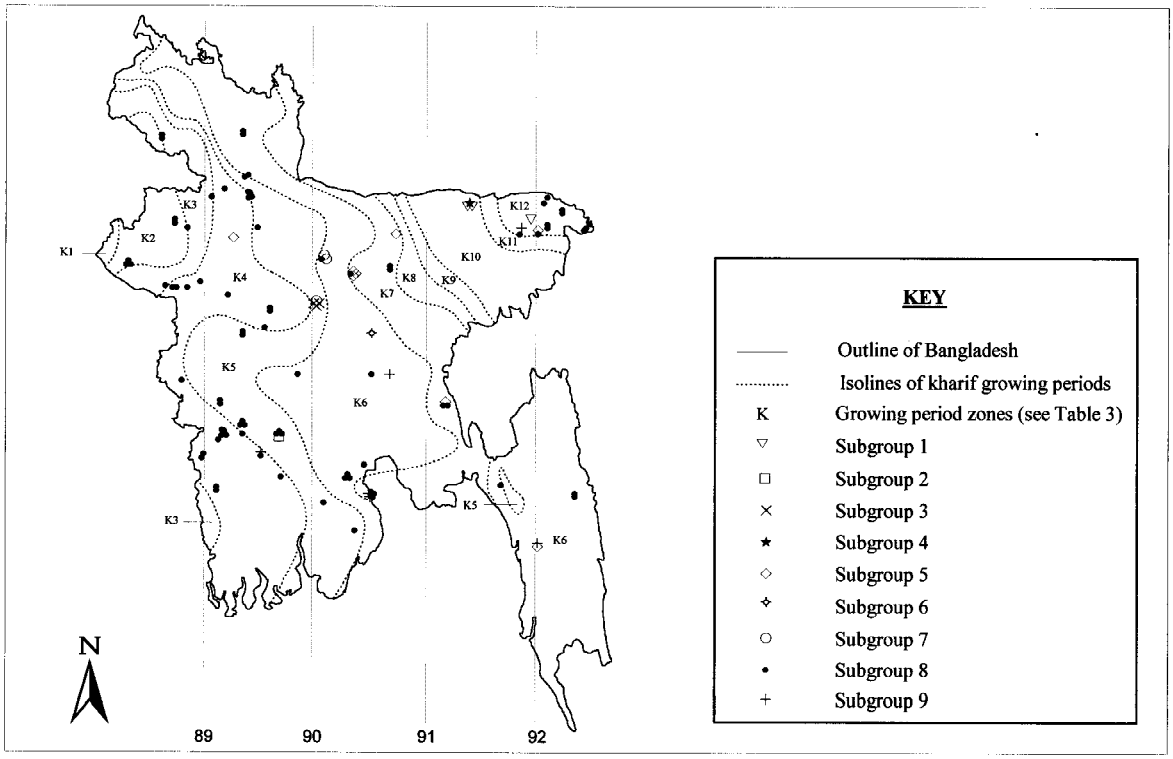

Figure 1. Map of Bangladesh displaying the collection sites of 115 rice varieties in their nine subgroups (see Figure 3) and the mean kharif growing period zones (K1 –K12).

Aus ecotype core collection

A core collection of accessions comprising 30% of the total sample of aus rices referred to above was selected using a hierarchical sampling strategy. By truncating the hierarchy at a similarity level of 0.83, 24 end groups resulted. A single accession from each end group was randomly sampled. Nei's average 'gene diversity' index (H) and Shannon's Information index (H_s) were used to quantify the extent of diversity within the core collection. By way of comparison, the levels of diversity within three randomly chosen sets of 24 accessions were evaluated and the average level of diversity for a randomly-selected core was calculated.

Results

Quantification of genetic diversity within Bangladesh

Sixty-nine RAPD markers (ranging in size between 0.3–2.75 kb) were generated, of which 35 (50.7%) were polymorphic. Figure 2 illustrates the amplification profiles generated across 12 duplicate pairs of accessions using primer OPK-10. The 35 polymorphic bands were scored across the 115 rice accessions and used to estimate the level of genetic diversity. Average genetic diversity within the Bangladeshi rice material was quantified as 0.200 or 0.332 according to Nei's or Shannon's indices, respectively.

Cluster analysis

RAPD data, comprising 35 polymorphic markers scored across the diverse Bangladeshi rice material were subjected to multivariate analysis. Information recorded in the IRGCIS database revealed that 7 accessions (S56, S58, S60, S63, S83, S97 and S100) were 'exotic' (i.e. not originating from the location at which they were collected), so these accessions were subsequently excluded from multivariate analysis. Levels of similarity between the remaining 108 accessions were calculated using Jaccard's coefficient and accessions were grouped according to the UPGMA clustering algorithm. Accessions ranged in their

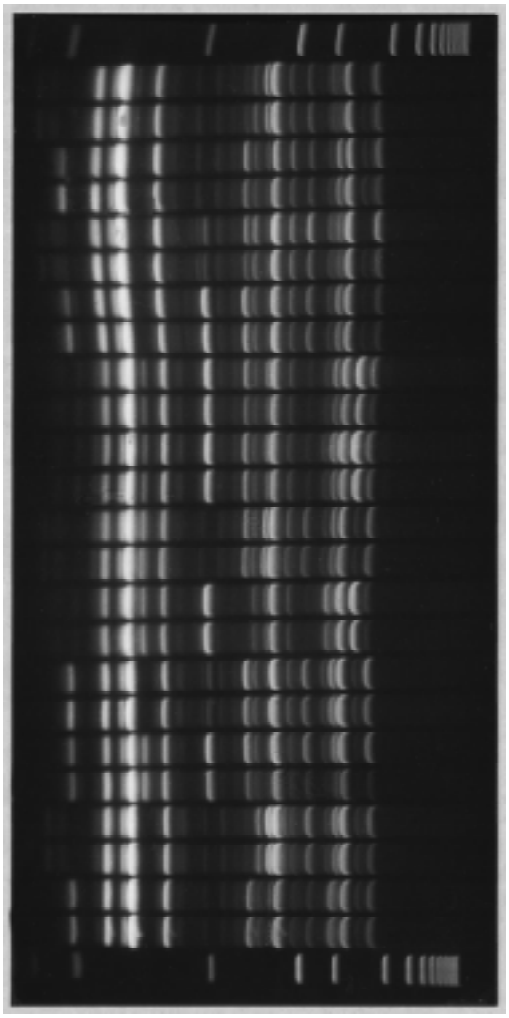

Figure 2. Agarose gel electrophoresis of RAPD products amplified using primer OPK-10 demonstrating the reproducibility of the RAPD reaction and the level of genetic variation within the diverse test material. Molecular 1 kb size standards are shown in lanes 1 and 26. Lanes 2–25 contain duplicate amplifications of 12 accessions (S49–S60; see Table 1).

similarities to each other from 27% similar to being identical on the basis of 35 molecular markers. The predominant split in the hierarchical classification occurred at a similarity level of 0.27 (see Figure 3). On the left side of the dendrogram, seven accessions (S13, S15, S16, S33, S72, S73 and S106) formed a group, clearly distinct from the bulk of the germplasm. Accessions S13, S15 and S16 were characterized as *javanica* (ja) varieties on the basis of morphology (IRRI, 1980); *javanica* rices are tropical forms of *japonica* rice (isozyme group VI). Accessions S72, S73 and S106 are cultivated under irrigated conditions; their close association with the *javanica* rices suggests that they may be of hybrid origin with an *indica* × *japonica* genetic base. No useful data were available for the classification of S33.

Rice ecotypes

By accessing data regarding accession ecotype, sowing and harvest date and varietal maturity from the IRGCIS, it was possible to characterize most *indica* accessions into their respective cultural types, namely aus (a), aman (m) or boro (b). Accessions S9, S10, S14, S17, S19, S21 and S28 could only be tentatively characterized as aus rices. Aus and aman rices grouped separately (groups A and B; see Figure 3) at a similarity level of 0.45; boro rices could not, however, be easily differentiated from aus and aman cultural types on the basis of 35 polymorphic RAPD markers. Excluding S1, S14, S28, S39, S53 and S54, aus rices show a high degree of genetic relatedness, with a similarity of approx. 68%. Accessions S53 and S54 are garo tribal rices which may account for their distinctness from the bulk of the aus germplasm.

Correspondence of aus and aman ecotypes with isozyme groups I and II, respectively, was previously demonstrated by Glaszmann (1987). Although the vast majority of accessions considered here have not been classified into specific isozyme groups, several indicator accessions of known isozyme group were included in the analysis; it was thus possible to predict the likely isozyme group status of the remaining Bangladeshi material. As accessions S13, S15 and S16 are *javanicas*, they were classified as isozyme group VI. Accession S115 in subgroup A was classified as isozyme group I (aman) whilst accessions S99, S111, S113 and S114 in subgroup B were classified as isozyme group II (aus). Therefore, at a similarity level of 0.27, rice varieties encompassing three isozyme groups were differentiated from each other.

The level of genetic diversity for each rice ecotype was quantified using Nei's average 'gene diversity' index (H) and Shannon's information index (H_s), averaged over the 35 loci, the results of which are presented in Table 2.

Distribution of diversity across Bangladesh

At an arbitrarily defined similarity value of 0.61 (selected to achieve a reasonable number of groups), the rice material can be separated into nine subgroups, designated 1 to 9 (see Figure 3). By assigning accessions within each subgroup a symbol code, and plot-

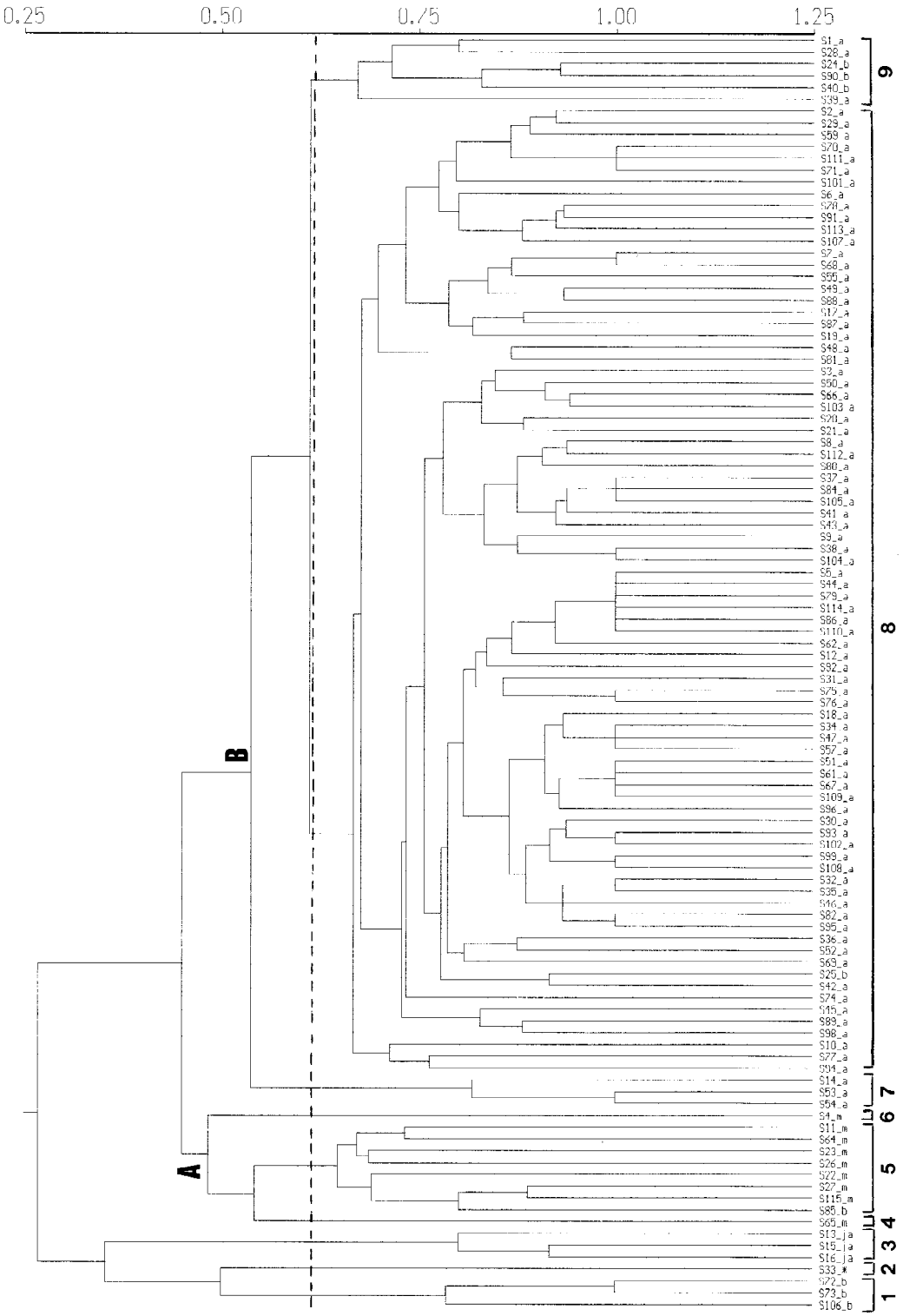

Figure 3. Dendrogram generated by UPGMA cluster analysis (Rohlf, 1992) of 1-F values based on Jaccard's coefficient determined using 35 polymorphic RAPD markers showing clustering of 108 Bangladeshi rice accessions. Accessions are listed according to their reference numbers in Table 1. The nine subgroups (1–9), arbitrarily defined at a similarity level of 0.61, correspond to the coded subgroups illustrated in Figure 1.

Table 2. Summary of genetic diversity estimates in Asian cultivated rice

Origin of accessions	No. of accessions	No. of loci (polymorphic)	H	H_s	Reference
Bangladesh	115	35 RAPD	0.200	0.332	This study
Asia	114	35 RAPD	0.344	0.511	Virk, pers. comm.
Asia	1688	15 isozymes	0.360	–	Glaszmann, 1988
North East India	289	14 isozymes	0.341	–	Glaszmann et al., 1989
Yunnan Province of China	252	6 isozymes	0.270	–	Nagamine et al., 1992
Chinese Provinces	19	6 isozymes	0.120	–	Nagamine et al., 1992
Japan	21	6 isozymes	0.050	–	Nagamine et al., 1992
Asia	42	6 isozymes	0.320	–	Nagamine et al., 1992
Asia	101	6 isozymes	0.493	–	Sano and Morishima, 1992
Himalayan hilly areas	151	6 isozymes	0.471	–	Sano and Morishima, 1992
Worldwide	468	25 isozymes	0.230	–	Second, 1982
Asia	140	10 microsatellites	0.656	–	Yang et al., 1994
Asia	26	43 RFLP probes	–	0.438	Zhang et al., 1992
Aus (isozyme II)	80[a]	21 RAPD	0.126	0.206	This study
Aman (isozyme I)	9[a]	17 RAPD	0.154	0.237	This study
Boro	8[a]	20 RAPD	0.221	0.325	This study
Aus core (hierarchical)	24[b]	20 RAPD	0.149	0.237	This study
Aus core (random)[c]	24[b]	16.3 (mean) RAPD	0.125	0.197	This study

[a] Only accessions which could be unambiguously classified into their respective ecotypes were included in the diversity analysis.
[b] Represents 30% of the total sample of aus rices.
[c] Diversity indices are average values from 3 samples of 24 randomly selected accessions.

ting accessions onto the outline map of Bangladesh according to their latitude-longitude co-ordinates, it was possible to visualize the spatial distribution of genetic diversity. Figure 1 is an Atlas-GIS plot illustrating the geographic distribution of these nine subgroups across Bangladesh; dotted vertical lines represent positions of longitude. Varieties sampled from the West (up to 90° E) are closely related with 95% of the accessions clustering within subgroup 8 (indicated by black dots). Only three accessions (S33, S39 and S85) are classified outside of this subgroup. Diversity appears much richer in the Central and Eastern part of the country (90–92° E), where all of the nine subgroups defined in Figure 3 are present, with subgroup 8 representing 57% of the total. These observations are in agreement with numerical diversity analyses. Average genetic diversity within the Bangladeshi material located in the West (up to 90° E; $n = 52$ accessions) was calculated as 0.12 or 0.20 according to Nei's (H) or Shannon's (H_s) indices respectively. For accessions ($n = 53$) located in central and eastern parts (90–92.5° E), H and H_s were calculated as 0.24 and 0.39 respectively. Furthermore, when the indices of genetic diversity (H and H_s) were calculated only for the aus ecotype accessions collected from the different kharif growing period zones (see Figure 1), the results also supported the hypothesis that Central Bangladesh shows the highest levels of diversity. These results are summarised in Table 3.

Aus ecotype core collection

Cluster analysis was carried out with the 80 aus rices and resulted in the dendrogram presented as Figure 4. By truncating the hierarchy at a similarity level of 0.83, 24 end groups resulted. A single accession from each end group was randomly sampled and these 24 accessions were pooled to produce a core collection. Selected accessions are indicated in Figure 4 with an asterisk. 20 of the 35 RAPD markers (57.1%)

Figure 4. Dendrogram generated by UPGMA cluster analysis of 1-F values based on Jaccard's coefficient determined using 23 polymorphic RAPD markers, showing clustering of 80 accessions within the aus ecotype. Hierarchically sampled core accessions are indicated with an asterisk.

that were polymorphic in the set of 115 accessions used earlier were polymorphic in this aus core. Nei's average 'gene diversity' index (H) and Shannon's Information index (H_s) were used to quantify the extent of diversity within the core collection, and these results are recorded in Table 2. By way of comparison,

Table 3. Levels of genetic diversity according to kharif length of growing period

Kharif length of growing period (K100) zone[a]	No. of accessions	H	H_s
K2 & K3 (180–200 days)[b]	10	0.098	0.148
K4 (200–210 days)	22	0.086	0.138
K5 (210–220 days)	16	0.108	0.166
K6 (220–230 days)	14	0.134	0.211
K7 (230–240 days)	7	0.089	0.135
K12 (280–290 days)	11	0.103	0.158

the level of diversity within three randomly chosen sets of 24 accessions was evaluated and the average level of diversity for a randomly selected core was calculated. For the three randomly selected aus cores, the average number of polymorphic bands was 16.3. The amount of diversity sampled using the hierarchical approach was found to be greater (H = 0.149; H_s = 0.237) than that achieved by taking the average of three random selections (H = 0.125; H_s = 0.197).

Discussion

The International Rice Genebank conserves over 5000 rice accessions originating from Bangladesh. This germplasm has long been recognised for its useful variation, including submergence tolerance (deepwater and aus varieties) and disease and pest resistances to rice blast, bacterial blight, rice tungro virus and stem borers (Jackson & Huggan, 1993; Glaszmann et al., 1996). In this study, a subset of this germplasm was evaluated using 35 polymorphic RAPD loci. The average genetic diversity was quantified at 0.200 and 0.332 using Nei's 'gene diversity' index and Shannon's information index, respectively. Table 2 summarises the work of eight independent studies where *O. sativa* germplasm has been evaluated for its genetic diversity. As diversity estimates were calculated from different data sets using contrasting marker types, it is not possible to evaluate precisely how levels of rice genetic diversity in Bangladesh compare with diversity levels in other rice growing areas. Our results do, however, substantiate the observation that Bangladeshi germplasm is relatively rich in genetic variation and represents an important genetic resource for rice breeders.

Cluster analysis of the Bangladeshi material revealed its genetic structure, consisting of *javanica* and *indica* rices, which could be further differentiated into aus, aman and to some extent, boro ecotypes. In agreement with Glaszmann's observations, aman rices were more diverse (H = 0.154) than varieties of the aus ecotype (H = 0.126); boro rices exhibited the greatest level of 'gene diversity' (H = 0.221). Levels of diversity quantified within aman and boro ecotypes are however tentative, as sample sizes were small.

Hierarchical cluster analysis is considered to be a useful approach for establishing core collections, by reducing the *ad hoc* nature of selecting accessions (Brown, 1989). Van Hintum (1995) used molecular characterization data to construct a diversity tree of accessions, by clustering them into related subgroups; diversity within the core was maximized by selecting a set number of genetically diverse accessions within each subgroup. This approach was used to sample a core collection comprising 30% of the aus varieties. The extent of genetic diversity sampled using the hierarchical approach was greater than could be achieved by randomly sampling within the aus ecotype accessions, which supports the assertion of van Hintum (1995).

Hierarchical cluster analysis, used in combination with GIS was used to visualize the spatial distribution of genetic diversity across Bangladesh. Rice germplasm was separated into a number of groups by truncating the hierarchy at an arbitrarily defined similarity level. Accessions within each group were subsequently assigned a group specific symbol and were plotted, within a GIS, onto an outline map of Bangladesh according to the co-ordinates of their collection sites. Diversity richness patterns were visualized as a function of the number of different symbol types residing within a given region. Diversity was greatest in the East and Central regions of Bangladesh, and much more limited in the West of the country (see Figure 1). The results demonstrate that the novel approach of combining hierarchical cluster analysis with GIS is a useful tool for accurately visualizing the spatial distribution of existing genetic diversity within a sample of germplasm in a given geographic area. This technique shows particular promise for studying diversity richness patterns in germplasm which is less well characterized and could potentially be employed for identifying 'hotspots' of diversity and targeting regions where further sampling of germplasm is required for species conservation.

To eliminate the complication of the non-random sampling of ecotypes across Bangladesh, the geographic pattern of diversity within a single rice ecotype

was investigated. Diversity indices were re-calculated exclusively for the sample of aus rices. Results indicate that aus rices originating from Central Bangladesh (between 90 and 90.5° E) show the highest levels of diversity (H = 0.136); less was detected in varieties from the West of the country (88.5–89° E; H = 0.089).

A considerable gradient in annual rainfall exists across Bangladesh, from 1250 mm in the extreme West to 6000 mm in the Northeast. To test the hypothesis of Nevo et al. (1979) of a curvilinear association between the level of genetic diversity and rainfall, the distribution of genetic diversity across Bangladesh according to the length of kharif growing period was performed. Kharif growing period mirrors the gradient in rainfall across the country, from a minimum of 180 days in the Northwest (Zone K2, see Figure 1), where annual rainfall averages approximately 1250 mm, to a maximum of 290 days in the Northeast (Zone K12), where annual rainfall averages approximately 6000 mm. In Central Bangladesh (zone K6) where rainfall is intermediate, the length of growing period is approximately 220 days. Our results demonstrate that, in general agreement with Nevo et al. (1979), genetic diversity is least at the geographic extremes and is at a maximum (H = 0.134) where rainfall is intermediate (K6; Central Bangladesh).

In conclusion, we have been able to show that it is possible not only to identify crossability groups within *O. sativa* but also the aus, aman and boro ecotypes. Our approach provides preliminary indications that it is possible to identify associations between levels of genetic diversity and ecogeographic variables, and the efficiency of GIS for visualizing the spatial pattern of diversity is demonstrated. We have shown that user-friendly GIS-software, which is widely available at relatively low cost, provides an ideal platform for integrating molecular and agro-ecological data, facilitating increasingly sensitive and sophisticated studies of species diversity patterns, which would be lengthy to perform manually. Finally, the hierarchical approach of cluster analysis not only allows the effective identification of ecogeographical patterns of diversity, but can also be used as a basis for developing diversity-rich core collections.

Acknowledgements

We are grateful for the technical assistance of Fay Hughes during this work.

References

Alim, A., 1974. An introduction to Bangladesh agriculture. (Dacca: Ministry of Agriculture).

Bowman, K.O., K. Hutcheson, E.P. Odum & L.R. Shenton, 1971. Comments on the distribution of indices of diversity. Stat. Ecol. 3: 315–366.

Brown, A.H.D., 1989. Core collections: a practical approach to genetic resources management. Genome 31: 818–824.

Brown, A.H.D., 1995. The core collection at the cross roads. In: Hodgkin, T., A.H.D. Brown, Th.J.L. van Hintum & E.A.V. Morales (Eds), Core Collections of Plant Genetic Resources. John Wiley and Sons, New York, NY.

Chang, T.T., 1989. The management of rice genetic resources. Genome 31: 825–831.

Chang, T.T., Y. Pan, Q. Chu, R. Peiris & G.C. Loresto, 1991. Cytogenetic, electrophoretic, and root studies of *javanica* rices. In: Rice Genetics II, IRRI, ed. (Manila: IRRI), pp. 21–31.

ESRI Inc., 1993. Understanding GIS: The ARC/INFO Method. Longman Scientific & Technical, Harlow.

FAO, 1988. Land resources appraisal of Bangladesh for agricultural development. Report 4 Hydro-climatic resources. (Rome: FAO).

Glaszmann, J.C., 1986. A varietal classification of Asian cultivated rice (*Oryza sativa* L.) based on isozyme polymorphism. In: Rice Genetics, IRRI, ed. (Manila: IRRI), pp. 119–130.

Glaszmann, J.C., 1987. Isozymes and classification of Asian rice varieties. Theor. Appl. Genet. 74: 21–30.

Glaszmann, J.C., 1988. Geographic pattern of variation among Asian native rice cultivars (*Oryza sativa* L.) based on fifteen isozyme loci. Genome 30: 782–792.

Glaszmann, J.C., P. Benyayer & M. Arnaud, 1989. Genetic divergence among rices from Northeast India. Rice Genet. Newslett. 6: 65.

Glaszmann, J.C., T. Mew, H. Hibino, C.K. Kim, T.I. Vergel de Dios-Mew, C.H. Vera Cruz, J.L. Notteghem & J.M. Bonman, 1996. Molecular variation as a diverse source of disease resistance in cultivated rice. In: Rice Genetics III, G.S. Khush ed. (Manila: IRRI), pp. 460–465.

International Rice research Institute, 1980. Descriptors for rice *Oryza sativa* L. IBPGR-IRRI Advisory Committee. P.O. Box 933, Manila, Philippines.

Jackson, M. & R. Huggan, 1993. Sharing the diversity of rice to feed the world. Diversity 9: 22–25.

Jackson, M., 1997. Conservation of rice genetic resources – the role of the International Rice Genebank at IRRI. Plant Mol. Biol. 35: 61–67.

Kato, S., 1930. On the affinity of cultivated varieties of rice plants *Oryza sativa* L. Journal of the Department of Agriculture, Kyushu Imperial University 2: 241–276.

Morinaga, T., 1954. Classification of rice varieties on the basis of affinity. Jpn. J. Breeding 4: 1–19.

Morinaga, T. & H. Kuriyama, 1955. *Japonica* type rice in the subcontinent of India and Java. Jpn. J. Breeding 5: 148–153.

Morinaga, T. & H. Kuriyama, 1958. Intermediate type of rice in the subcontinent of India and Java. Jpn. J. Breeding 7: 253–258.

Nagamine, T., J.H. Xiong & Q. Xiao, 1992. Genetic variation in several isozymes of indigenous rice varieties in the Yunnan province of China. Jpn. J. Breeding 42: 507–513.

Nei, M., 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA 70: 3321–3323.

Nevo, E., D. Zohary, A.H.D. Brown & M. Haber, 1979. Genetic diversity and environmental associations of wild barley, *Hordeum spontaneum*, in Israel. Evolution 33: 815–833.

Oka, H.I., 1953. Phylogenetic differentiation of cultivated rice, 1. Variations in various characters and character combinations among rice varieties. Jpn. J. Breeding 3: 33–34.

Parsons, B.J., H.J. Newbury, M.T. Jackson & B.V. Ford-Lloyd, 1997. Contrasting genetic diversity relationships are revealed in rice (*Oryza sativa* L.) using different marker types. Mol. Breeding 3: 115–125.

Rohlf, F.J., 1992. NTSYS-pc Numerical Taxonomy and Multivariate Analysis System Version 1.70. State University of New York, Stony Brook, NY.

Sano, R. & H. Morishima, 1992. *Indica-japonica* differentiation of rice cultivars viewed from variations in key characters and isozymes, with special reference to landraces from the Himalayan hilly areas. Theor. Appl. Genet. 84: 266–274.

Second, G., 1982. Origin of the genic diversity of cultivated rice (*Oryza* spp.): study of the polymorphism scored at 40 isozyme loci. Jpn. J. Genetics 57: 25–57.

SGRP, 1996. Report of the internally commissioned external review of the cgIAR genebank operations. International Plant Genetic Resources Institute, Rome, Italy.

SGRP, 1997. Report of the internally commissioned external review of the cgIAR genebank operations. Annex: Centres' responses to the external review. International Plant Genetic Resources Institute, Rome, Italy.

Strategic Mapping Inc., 1994. Atlas GIS reference manual. Strategic Mapping Inc., Santa Clara.

Takahashi, N. & A.H.A. Hamza, 1983. Differentiation of ecotypes in *Oryza sativa* L. I. Re-examination for colour reaction with phenol. Jpn. J. Breeding 33: 243–250.

Ueno, K., T. Sato & N. Takahashi, 1990. The *indica-japonica* classification of Asian rice ecotypes and Japanese lowland and upland rice (*Oryza sativa* L.). Euphytica 46: 161–164.

van Hintum, Th.J.L., 1995. Hierarchical approaches to the analysis of genetic diversity in crop plants. In: Hodgkin, T., A.H.D. Brown, Th.J.L. van Hintum & E.A.V. Morales (Eds), Core Collections of Plant Genetic Resources, IPGRI, Rome, pp. 23–34.

Virk, P.S., B.V. Ford-Lloyd, M.T. Jackson & H.J. Newbury, 1995. Use of RAPD for the study of diversity within plant germplasm collections. Heredity 74: 170–179.

Williams, J.G.K., A.R. Kubelik, K.J. Livak, J.A. Rafalski & S.V. Tingey, 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18: 6531–6535.

Yang, G.P., M.A. Saghai Maroof, C.G. Xu, Q. Zhang & R.M. Biyashev, 1994. Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice. Molecular and General Genetics 245: 187–194.

Zaman, S.M.H., 1980. Ecology of transplant aman rice. In: Proceedings of the International seminar on photoperiod-sensitive transplant rice 1977. (Dacca, Bangladesh Rice Research Institute), pp. 1–7.

Zhang, Q., M.A. Saghai Maroof, T.Y. Lu & B.Z. Shen, 1992. Genetic diversity and differentiation of *indica* and *japonica* rice detected by RFLP analysis. Theor. Appl. Genet. 83: 495–499.