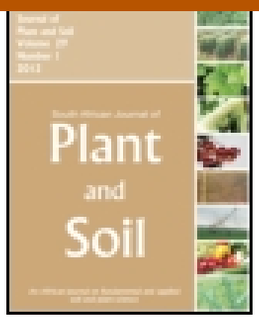

South African Journal of Plant and Soil

ISSN: 0257-1862 (Print) 2167-034X (Online) Journal homepage: http://www.tandfonline.com/loi/tjps20

Assessment of genetic diversity of rice based on SNP markers for selection of parents for sheath rot (Sarocladium oryzae) resistance breeding

Simon M Mvuyekure, Julia Sibiya, John Derera, John Nzungize & Germain Nkima

To cite this article: Simon M Mvuyekure, Julia Sibiya, John Derera, John Nzungize & Germain Nkima (2017): Assessment of genetic diversity of rice based on SNP markers for selection of parents for sheath rot (Sarocladium oryzae) resistance breeding, South African Journal of Plant and Soil, DOI: 10.1080/02571862.2017.1333636

To link to this article: http://dx.doi.org/10.1080/02571862.2017.1333636

South African Journal of Plant and Soil 2017: 1–9
Printed in South Africa — All rights reserved

This is the final version of the article that is published ahead of the print and online issue

SOUTH AFRICAN JOURNAL OF PLANT AND SOIL
ISSN 0257-1862 EISSN 2167-034X
http://dx.doi.org/10.1080/02571862.2017.1333636

Assessment of genetic diversity of rice based on SNP markers for selection of parents for sheath rot (*Sarocladium oryzae*) resistance breeding

Simon M Mvuyekure[1,2]*, Julia Sibiya[1], John Derera[1], John Nzungize[2,3] and Germain Nkima[2]

[1] *African Centre for Crop Improvement, School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, Pietermaritzburg, South Africa*
[2] *Rwanda Agriculture Board, Kigali, Rwanda*
[3] *International Crops Research Institute for the Semi-Arid Tropics, Bamako, Mali*
* *Corresponding author, email: msmartin202@gmail.com*

Sheath rot of rice, caused by *Sarocladium oryzae*, is an important emerging rice disease not only in Rwanda, but also in other rice-growing countries. Given that cultivar resistance is a sustainable management strategy for small-scale farmers, the aim of this study was to identify genetically distant parental materials for sheath rot resistance breeding. Ten resistant and fifteen susceptible accessions were analysed using 94 single nucleotide polymorphism (SNP) markers. The number of alleles amplified per locus ranged from 1 to 4 with a mean of 2.01 and a total of 189 alleles detected from the 25 genotypes. The number of observations per marker locus ranged from 11 to 25 with an average of 23. The mean major allele frequency was 76.2%, whereas the mean polymorphic information content was 0.263, and gene diversity was estimated at 0.325. Consequently, the markers were highly informative and revealed good estimates of genetic diversity among the studied accessions. Genetic distances ranged from 0 to 0.63 and a UPGMA dendrogram distinguished resistant and susceptible genotypes. This study revealed the possibility of improving resistance to sheath rot with minimum risk of genetic depression or reduced variability among progenies through hybridisation of locally adapted germplasm.

Keywords: genetic diversity, germplasm, resistance, sheath rot of rice, SNPs

Introduction

Sheath rot of rice (ShR), caused by the seed-borne fungal pathogen *Sarocladium oryzae* (Sawada) W.Gams & D.Hawksw., is one of the most important emerging and devastating diseases in rice-growing regions (Lanoiselet 2012; Hittalmani et al. 2016). Losses range between 26% and 50%, but higher yield losses of up to 85% have been recorded (Sakthivel 2001). The disease is currently among many diseases that were formerly considered as minor, but have recently acquired the status of major diseases (Ngala and Adeniji 1986). This is probably due to changes in cultivation practices resulting from the Green Revolution, on one hand, and from the apparent climate change effects, on the other hand (Madhav et al. 2013). In addition, crop intensification practices, such as increased plant density, high application rates of nitrogen fertilisers, and the use of semi-dwarf and photoperiod-insensitive cultivars, favour the susceptibility of rice to some diseases including ShR (Bigirimana et al. 2015).

Management of ShR relies on the integration of chemicals with cultural practices. However, according to Ayyadurai et al. (2005), fungicide treatments have not been effective under some farming conditions or are very expensive as well as harmful to the environment. In the same context, biological control has been of limited effectiveness due to variability of antagonists under field conditions (Gnanamanickam 2009). Therefore, the most sustainable solution is the development and deployment of resistant cultivars, as this is easy for farmers to adopt at no additional cost and is environmentally friendly.

Although a number of resistant cultivars have been developed in different countries (Lakshmanan and Velusamy 1991; Pearce et al. 2001), most of them have failed to adapt to harsh environmental conditions of ecosystems into which they are introduced (Linares 2002). There is, therefore, a need to develop resistant genotypes using locally adapted parents. In this regard, breeding for sheath rot resistance requires the identification of sufficiently genetically distant parental materials for hybridisation. This approach aims at avoiding both genetic depression and a reduction of genetic variability in subsequent progenies.

Consequently, based on this observation, an assessment of genetic diversity, relationships and structure within a given set of germplasm is useful in plant breeding. This would provide information to assist in (1) selection of parental combinations for development of progenies with maximum genetic variability for genetic mapping or further selection, (2) determination of the level of genetic variability when defining core subsets selected for specific traits, and (3) estimation of possible loss of genetic diversity during conservation or selection programs (Reif et al. 2005).

Morphological characterisation of rice germplasm has been regarded as a central component of plant breeding

programs that facilitates selections, study of traits genetics, association of markers with traits and understanding traits diversity (Nascimento et al. 2011). Despite their usefulness, morphometric markers lead to more reliable indications when coupled with molecular markers (Kilian and Graner 2012). Molecular markers are particularly useful for the evaluation of genetic diversity in various crop species with a narrow genetic base (Soleimani et al. 2002). More recently, single nucleotide polymorphism (SNP) markers have acquired significant consideration in genetic diversity studies. This is because they are bi-allelic in nature and occur at a much higher frequency in the genome than any other markers (Ren et al. 2013).

This study was, therefore, undertaken to (1) provide substantial information to maintain and use locally available rice genetic resources in breeding, and (2) identify genetically distant parental materials to be utilised in various cultivar improvement programs with regard to resistance to ShR using SNP markers.

Materials and methods

Plant materials, DNA extraction and SNP genotyping

Plant materials used in this study comprised 25 rice accessions that were selected in cultivar improvement programs directed towards resistance to ShR. The selection of the genotypes was based on agromorphological attributes, reaction to ShR of rice as well as farmer and consumer preferences. The key agronomical characteristics of the assessed accessions, which were derived from a separate study on germplasm screening for resistance to ShR by Mvuyekure (2016), are given in Table 1. The latter study was carried out at the Rwanda Agriculture Board's rice research site located in Rurambi (02°02′23.53″ S, 30°10′58.92″ E; 1 340 m above sea level).

Leaf samples for DNA extraction were collected from 30-day-old seedlings using the LGC genomics plant sample collection kit (http://www.lgcgroup.com/plant-kit/#.Vsb0KFR97IU) and shipped to LGC Genomics, Hoddesdon, UK. DNA extraction and all SNPs genotyping processes were performed by LGC genomics, according to their validated protocol and working conditions. Genetic diversity among 10 ShR resistant and 15 susceptible cultivars was assessed using 94 SNPs that were obtained from the Integrated Plant Breeding Platform (https://www.integratedbreeding.net/544/communities/genomics-crop-info/crop-information/gcp-kaspar-snp-markers/crop-snp-markers/rice?map=1).

Selection of the SNPs was guided by factors that included an even distribution along all 12 linkage groups corresponding to the 12 rice chromosomes. While each linkage group contains between 100 and 120 markers, 7–9 markers were randomly and evenly chosen from each linkage group for this study.

A list of the selected SNPs is given in Table 2.

Data analysis

Genotyping data were analysed using Power Marker 3.25 for estimation of SNPs summary statistics, including allele number, major allele frequency, heterozygosity, number of observed genotypes per marker locus, gene diversity, polymorphic information content (PIC) and Nei frequency-based distance (Nei et al. 1983), as described by Liu and Muse (2005). Based on this distance, an unweighted pair group method with arithmetic mean (UPGMA) dendrogram was constructed using Mega 5.2 software for cluster analysis (Hall 2013) to evaluate relationship groups among the accessions.

Results

Polymorphism of SNP markers

The summary SNP statistics are presented in Table 2. Results indicated that 90.4% of marker alleles (85 markers out of 94) were polymorphic, whereas nine markers out of the 94 SNPs were monomorphic. The number of observations for a marker locus or the number of non-missing genotypes observed in the sample ranged from 11 to 25 with an average of 23. A genotype is generally regarded as missing if one of its two alleles is missing. The number of alleles amplified per locus varied between one and four. A total of 189 alleles were amplified with an average of 2.01 alleles per locus in the 25 accessions.

The major allele frequency ranged from 50% to 100% with an average of 76.2%. More than 60% of the polymorphic loci showed a major allele frequency higher than 70% and 11 loci showed more than 90%. The mean PIC value for markers was 0.263 with a range between 0 (monomorphism) and 0.555. On the basis of the marker informative levels established by Botstein et al. (1980), 48 markers (51.06%) were highly informative, 21 (22.3%) reasonably informative and 12 (15.8%) were not informative.

Heterozygosity, which is a measure of allelic diversity at a locus, ranged from 0.018 to 0.160, and its expected estimations or gene diversity ranged from 0 to 61.2% with a mean gene diversity of 32.5%. Therefore, the high allelic richness coupled with estimates of gene diversity indicated a high level of genetic diversity among the studied genotypes.

Genetic distance among accessions

The average genetic distances between and within indica/japonica rice groups are presented in Table 3. The least distance (0) was recorded between Ndamirabana (G18) and Gakire (G27); both cultivars belong to the indica group. This was followed by the distance between Rumbuka (G53) and Yunkeng (G40), which belong to the indica and japonica groups, respectively. The greatest distance (0.63) was recorded between Yunyine (G24) and Tetep (G60), from the japonica and indica groups, respectively. Yunyine recorded a high genetic distance (>0.5) with most of the genotypes (19 genotypes) except Zongeng (G7), Yunertian (G6), Yunkeng (G40), Fac 56 (G12), Rumbuka (G53) and Moroberekan (G59). In general, genetic distance clearly distinguished japonica and indica accessions at a broad scale, but also revealed the least distances within subspecies and greatest distances between subspecies.

Cluster analysis and relationship groups

The Nei frequency-based distance was used to assess similarities between the accessions and construct a UPGMA dendrogram for evaluation of relationship groups.

Table 1: Key agronomic features of sheath rot resistant and susceptible accessions used in the study

Code	Genotype	Subspecies	Cultivar reaction to ShR	Panicle exsertion	Plant stature	Tillering ability	Weight of 1 000 grains	Grain length
2	Intsindagirabigega	Indica	Moderately susceptible	Just exserted	Short	Intermediate	24.98	Long
4	Imbarurabukungu	Indica	Moderately susceptible	Partly exserted	Short	Intermediate	22.81	Long
6	Yunertian	Japonica	Resistant	Well exserted	Intermediate to long	Low	23.98	Medium
7	Zongeng	Japonica	Resistant	Moderately well exserted	Intermediate to long	Intermediate	22.53	Short
12	Fac 56	Indica	Susceptible	Partly exserted	Short	Intermediate	24.76	Long
15	Jyambere	Indica	Susceptible	Enclosed	Short	Intermediate	21.93	Long
16	Posiyani	Indica	Moderately susceptible	Moderately well exserted	Intermediate to long	Low	24.88	Short
18	Ndamirabana	Indica	Highly susceptible	Partly exserted	Short	Intermediate	22.6	Long
19	Fashingabo	Indica	Susceptible	Partly exserted	Short	Low	22.07	Long
24	Yunyine	Japonica	Resistant	Well exserted	Intermediate to long	Intermediate	23.96	Short
25	Nerica 1	*O. glaberrima* × *O. sativa*	Resistant	Moderately well exserted	Short to intermediate	Low	24.08	Medium
27	Gakire	Indica	Susceptible	Enclosed	Short	Low	24.74	Long
31	Ndengera	Indica	Susceptible	Partly exserted	Short	Intermediate	24.51	Medium
32	Buryohe	Indica	Highly susceptible	Partly exserted	Short	Intermediate	25.06	Long
33	Intsinzi	Indica	Susceptible	Enclosed	Short to intermediate	Low	23.61	Medium
34	Kimaranzara	Indica	Moderately susceptible	Just exserted	Short to intermediate	Low	25.69	Medium
40	Yunkeng	Japonica	Resistant	Well exserted	Intermediate to long	Low	25.66	Short
43	Nyiragikara	Unknown	Highly resistant	Well exserted	Intermediate to long	Intermediate	26.16	Short
44	Nerica 10	*O. glaberrima* × *O. sativa*	Moderately Resistant	Moderately well exserted	Short to intermediate	Low	22.2	Medium
48	Cyicaro	Indica	Resistant	Moderately well exserted	Short to intermediate	Intermediate	24.13	Short
50	Ndamirabahinzi	Indica	Highly susceptible	Partly exserted	Intermediate to long	Low	24.1	Medium
51	Mpembuke	Indica	Moderately susceptible	Partly exserted	Intermediate	Low	25.79	Long
53	Rumbuka	Indica	Highly susceptible	Just exserted	Intermediate to long	Low	20.34	Long
59	Moroberekan	*O. glaberrima*	Resistant	Well exserted	Intermediate to long	Intermediate	20.99	Short
60	Tetep	Unknown	Resistant	Just exserted	Short to intermediate	Intermediate	24.1	Medium

Table 2: Summary statistics for the SNPs used in the study. PIC = polymorphism information content

Marker	Major allele frequency	Genotype number	Sample size	Number of observations	Allele number	Gene diversity	Heterozygosity	PIC
K_id1002308	0.600	2	25	25	2	0.480	0	0.365
K_id1006954	0.667	2	25	24	2	0.444	0	0.346
K_id1008787	0.848	3	25	23	2	0.258	0.043	0.225
K_id1011568	0.560	3	25	25	2	0.493	0.080	0.371
K_id1014143	0.545	4	25	11	4	0.612	0	0.555
K_id1024233	0.545	3	25	22	2	0.496	0.091	0.373
K_id1025888	0.820	4	25	25	3	0.306	0.040	0.278
K_id1026656	0.960	2	25	25	2	0.077	0	0.074
K_id2000096	0.571	3	25	21	3	0.526	0	0.429
K_id2001992	0.841	3	25	22	2	0.268	0.045	0.232
K_id2004058	0.820	3	25	25	2	0.295	0.040	0.252
K_id2006621	0.720	2	25	25	2	0.403	0	0.322
K_id2007797	0.818	3	25	11	3	0.314	0	0.292
K_id2008480	0.880	2	25	25	2	0.211	0	0.189
K_id2010564	0.600	3	25	25	2	0.480	0.080	0.365
K_id2010969	0.696	3	25	23	3	0.446	0	0.378
K_id2013007	0.833	3	25	24	2	0.278	0.083	0.239
K_id3000111	0.700	3	25	25	2	0.420	0.040	0.332
K_id3002805	0.522	2	25	23	2	0.499	0	0.375
K_id3006808	0.708	2	25	24	2	0.413	0	0.328
K_id3007703	0.740	4	25	25	3	0.402	0.040	0.347
K_id3008390	0.714	3	25	21	3	0.431	0	0.370
K_id3010318	0.826	2	25	23	2	0.287	0	0.246
K_id3010628	0.913	2	25	23	2	0.159	0	0.146
K_id3013806	0.563	3	25	24	2	0.492	0.042	0.371
K_id3017084	0.905	3	25	21	3	0.177	0	0.169
K_id4001365	0.600	2	25	15	2	0.480	0	0.365
K_id4002780	0.543	3	25	23	2	0.496	0.043	0.373
K_id4004294	0.636	2	25	22	2	0.463	0	0.356
K_id4005120	0.636	2	25	22	2	0.463	0	0.356
K_id4005867	0.700	2	25	20	2	0.420	0	0.332
K_id4007444	0.587	3	25	23	2	0.485	0.043	0.367
K_id4010621	0.500	3	25	24	3	0.538	0	0.432
K_id4012434	0.714	2	25	21	2	0.408	0	0.325
K_id5000128	0.913	2	25	23	2	0.159	0	0.146
K_id5001534	0.920	2	25	25	2	0.147	0	0.136
K_id5003785	0.696	2	25	23	2	0.423	0	0.334
K_id5006332	0.565	2	25	23	2	0.491	0	0.371
K_id5007714	0.587	3	25	23	2	0.485	0.043	0.367
K_id5008723	1	1	25	18	1	0	0	0
K_id5011704	0.708	2	25	24	2	0.413	0	0.328
K_id5013100	0.696	2	25	23	2	0.423	0	0.334
K_id6000134	0.880	2	25	25	2	0.211	0	0.189
K_id6004862	1	1	25	22	1	0	0	0
K_id6007386	1	1	25	23	1	0	0	0
K_id6010534	0.909	2	25	22	2	0.165	0	0.152
K_id6012080	0.740	3	25	25	2	0.385	0.040	0.311
K_id6012658	0.860	3	25	25	2	0.241	0.040	0.212
K_id6016125	0.708	2	25	24	2	0.413	0	0.328
K_id6002535	0.880	2	25	25	2	0.211	0	0.189
K_id7000063	0.750	2	25	24	2	0.375	0	0.305
K_id7001596	0.913	2	25	23	2	0.159	0	0.146
K_id7002534	0.913	2	25	23	2	0.159	0	0.146
K_id7003748	0.600	2	25	25	2	0.480	0	0.365
K_id7004442	0.960	2	25	25	2	0.077	0	0.074
K_id7005111	0.842	2	25	19	2	0.266	0	0.231
K_id7005689	0.860	3	25	25	2	0.241	0.040	0.212
K_id8000131	0.625	2	25	24	2	0.469	0	0.359
K_id8001667	0.760	2	25	25	2	0.365	0	0.298
K_id8003220	0.761	3	25	23	2	0.364	0.043	0.298
K_id8004986	0.520	3	25	25	2	0.499	0.160	0.375
K_id8006032	0.750	2	25	24	2	0.375	0	0.305

Table 2 (cont.)

Marker	Major allele frequency	Genotype number	Sample size	Number of observations	Allele number	Gene diversity	Heterozygosity	PIC
K_id8006950	0.826	2	25	23	2	0.287	0	0.246
K_id8007951	0.810	2	25	21	2	0.308	0	0.261
K_id9000045	1	1	25	19	1	0	0	0
K_id9001558	0.739	2	25	23	2	0.386	0	0.311
K_id9002532	0.604	3	25	24	2	0.478	0.042	0.364
K_id9003471	1	1	25	25	1	0	0	0
K_id9004347	0.739	2	25	23	2	0.386	0	0.311
K_id9005089	0.870	2	25	23	2	0.227	0	0.201
K_id9006757	1	1	25	19	1	0	0	0
K_id9007001	0.640	2	25	25	2	0.461	0	0.355
K_id9007259	0.700	3	25	25	2	0.420	0.040	0.332
K_id10000028	0.565	2	25	23	2	0.491	0	0.371
K_id10001624	0.580	3	25	25	2	0.487	0.040	0.369
K_id10002912	0.660	3	25	25	2	0.449	0.040	0.348
K_id10004275	1	1	25	25	1	0	0	0
K_id11000399	0.804	3	25	23	2	0.315	0.043	0.265
K_id11001993	1	1	25	25	1	0	0	0
K_id11003845	0.935	2	25	23	2	0.122	0.130	0.114
K_id11005657	0.813	2	25	16	2	0.305	0	0.258
K_id11006897	0.960	2	25	25	2	0.077	0	0.074
K_id11007625	0.848	3	25	23	2	0.258	0.130	0.225
K_id11008403	1	1	25	25	1	0	0	0
K_id11008862	0.640	2	25	25	2	0.461	0	0.355
K_id11010309	0.771	3	25	24	2	0.353	0.125	0.291
K_id12000266	0.750	2	25	24	2	0.375	0	0.305
K_id12001996	0.773	2	25	22	2	0.351	0	0.290
K_id12004271	0.583	2	25	24	2	0.486	0	0.368
K_id12005822	0.917	2	25	24	2	0.153	0	0.141
K_id12006560	0.708	2	25	24	2	0.413	0	0.328
K_id12008285	0.650	2	25	20	2	0.455	0	0.351
K_id12008894	0.875	2	25	24	2	0.219	0	0.195
K_id12006515	0.739	2	25	23	2	0.386	0	0.311
Mean	0.762	2.2872	25	23	2.0106	0.325	0.018	0.263

Table 3: Nei frequency-based distances among the studied germplasm

	G12	G15	G16	G18	G19	G2	G24	G25	G27	G31	G32	G33	G34	G4	G40	G43	G44	G48	G50	G51	G53	G59	G6	G60	G7
G12																									
G15	0.28																								
G16	0.25	0.04																							
G18	0.34	0.22	0.25																						
G19	0.25	0.20	0.21	0.32																					
G2	0.23	0.17	0.23	0.26	0.24																				
G24	0.26	0.61	0.61	0.57	0.60	0.54																			
G25	0.32	0.28	0.29	0.28	0.20	0.24	0.61																		
G27	0.29	0.20	0.24	0.00	0.29	0.24	0.56	0.29																	
G31	0.31	0.27	0.28	0.25	0.25	0.19	0.60	0.07	0.26																
G32	0.26	0.27	0.26	0.29	0.25	0.28	0.57	0.29	0.27	0.28															
G33	0.35	0.19	0.21	0.28	0.19	0.19	0.58	0.30	0.26	0.31	0.24														
G34	0.36	0.28	0.33	0.33	0.28	0.31	0.59	0.30	0.32	0.29	0.31	0.33													
G4	0.26	0.14	0.18	0.17	0.22	0.16	0.61	0.25	0.17	0.22	0.26	0.21	0.25												
G40	0.27	0.46	0.45	0.48	0.41	0.45	0.36	0.50	0.47	0.46	0.39	0.44	0.45	0.44											
G43	0.12	0.23	0.24	0.29	0.19	0.20	0.55	0.23	0.26	0.22	0.22	0.25	0.31	0.26	0.45										
G44	0.23	0.19	0.22	0.23	0.17	0.12	0.56	0.21	0.22	0.23	0.19	0.22	0.32	0.23	0.44	0.15									
G48	0.24	0.16	0.21	0.20	0.23	0.20	0.52	0.30	0.19	0.24	0.28	0.26	0.27	0.13	0.41	0.24	0.22								
G50	0.20	0.11	0.13	0.18	0.15	0.06	0.58	0.16	0.17	0.16	0.17	0.18	0.23	0.08	0.40	0.15	0.04	0.13							
G51	0.25	0.10	0.15	0.21	0.22	0.15	0.60	0.27	0.19	0.22	0.26	0.21	0.26	0.03	0.42	0.26	0.21	0.10	0.05						
G53	0.34	0.49	0.47	0.49	0.42	0.45	0.42	0.47	0.47	0.42	0.41	0.40	0.53	0.46	0.18	0.43	0.43	0.45	0.39	0.43					
G59	0.32	0.48	0.48	0.51	0.44	0.47	0.39	0.52	0.50	0.48	0.44	0.47	0.50	0.46	0.05	0.50	0.47	0.43	0.43	0.45	0.20				
G6	0.15	0.60	0.57	0.57	0.53	0.52	0.13	0.56	0.56	0.55	0.51	0.59	0.59	0.56	0.30	0.51	0.51	0.54	0.51	0.55	0.38	0.31			
G60	0.25	0.18	0.22	0.26	0.22	0.12	0.63	0.22	0.22	0.21	0.22	0.24	0.31	0.24	0.48	0.16	0.05	0.20	0.05	0.20	0.43	0.51	0.57		
G7	0.23	0.58	0.59	0.56	0.58	0.51	0.19	0.57	0.55	0.56	0.53	0.62	0.61	0.56	0.37	0.54	0.50	0.57	0.50	0.56	0.39	0.39	0.08	0.55	

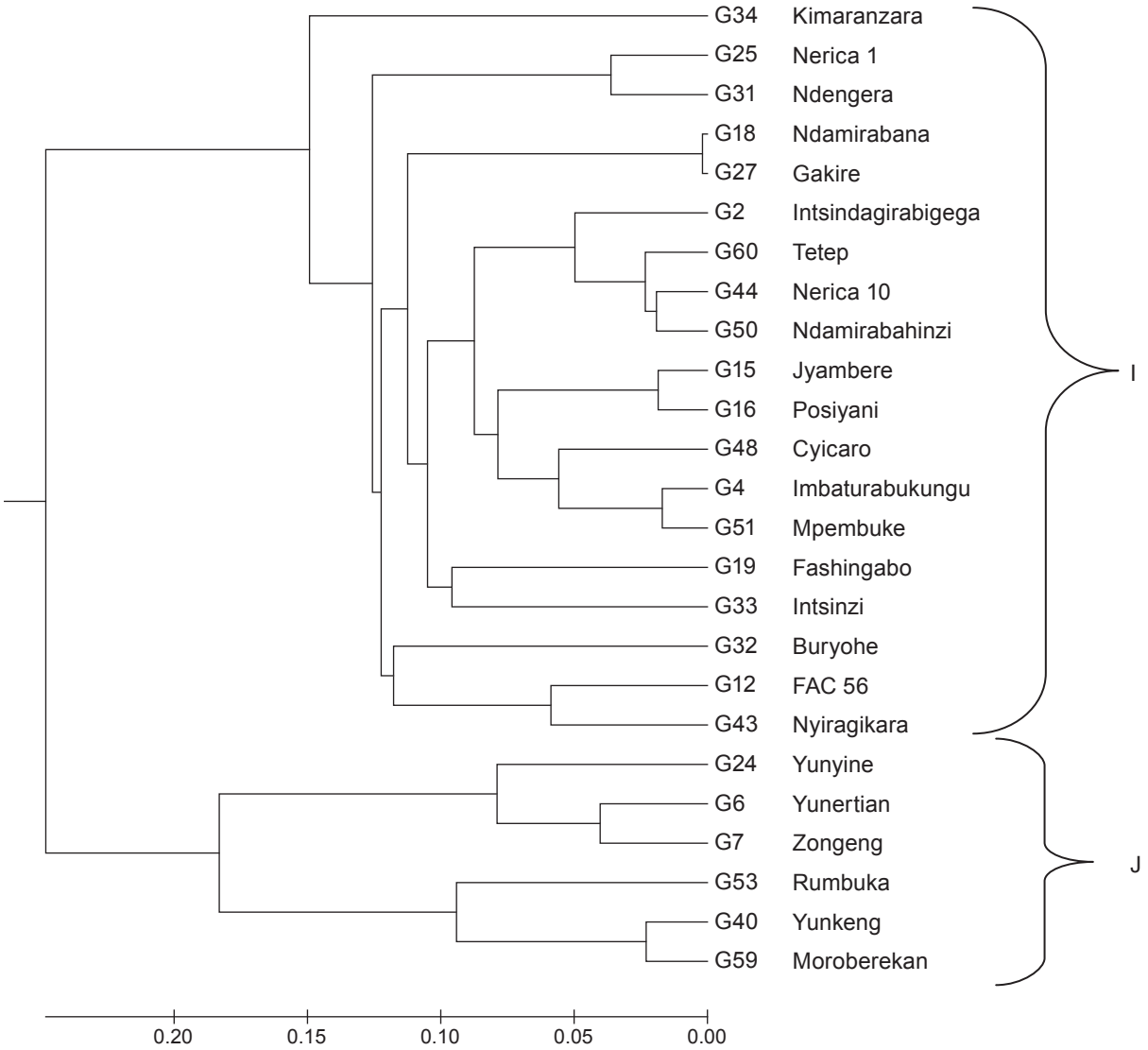

Figure 1: UPGMA dendrogram of sheath rot resistant and susceptible genotypes of rice based on Nei genetic distance

The dendrogram in Figure 1 summarises evolutionary relationships among the genotypes and categorises them into distinct genetic groups. In the dendrogram, nodes represent different genotypes, whereas branches are graphical estimates of the genetic distance between the genotypes, thus indicating genetic relationships between the genotypes. In the dendrogram, accessions were separated into two major groups. Accessions placed in cluster I were morphologically similar to indica rice, whereas accessions placed in cluster J were morphologically close to japonica rice.

However, a number of exceptions to these groupings were observed. For instance, accessions such as Nerica 1 and Nerica 10 placed in cluster I are not indica rice cultivars but hybrids between *Oryza glaberrima* and *Oryza sativa.* These accessions are reported to be resistant to sheath rot of rice (Mvuyekure 2016), in contrast to the remainder of the group. Similarly, Moroberekan was genetically similar to japonica-type accessions, although it is an accession of *Oryza glaberrima.* Japonica genotypes are easily distinguished by their tall plant stature, well-exserted panicles and short grains, and more particularly by their resistance to ShR compared with indica-type cultivars. Susceptible genotypes of the indica group in the same cluster were more genetically similar, whereas those in different clusters were genetically dissimilar. This was the case, for instance, for Rumbuka and Kimaranzara, which are both indica types but showed a high degree of dissimilarity.

Based on morphological characteristics of the studied genotypes and the Nei similarity matrix (Table 3), the dendrogram in Figure 1 revealed two accessions, Ndamirabana and Gakire, that were previously believed to be different cultivars and showed a high degree of similarity with a very small genetic distance between them. Although Rumbuka and Moroberekan were clustered with the japonica rice types, they actually belong to the indica type and *Oryza glaberrima,* respectively. As far as resistance to ShR is concerned, the most susceptible cultivars were placed in

cluster I, except for Nerica 1 and 10 and Nyiragikara, which are resistant to ShR. Most of the resistant genotypes were placed in cluster J except for Rumbuka.

Discussion

Accurate identification of genetic relationship and divergence of genetic resources is most useful for efficient choice of parental materials in breeding and genetic conservation strategies (Guimarães 2009). This would assist in minimising the use of closely related parents in breeding programs, which might lead to a high risk of genetic depression and reduced genetic variation (Weddell 2002). The present investigation was, therefore, conducted to establish the genetic variability and relationships among 25 selected rice accessions or which are 10 resistant and 15 susceptible to ShR. This was an attempt to identify potential parental materials suitable for various hybridisation processes, with particular regard to resistance to ShR. To this end, SNPs were used because of their low cost per data point, high genomic abundance, locus-specificity, co-dominance, high-throughput analysis and lower genotyping error rates (Rafalski 2002).

Singh et al. (2013) indicated that polymorphism frequencies are an important criterion that can be used to assess the value of markers for germplasm characterisation. In this regard, 85 out of the 94 (90.4%) SNP markers used provided adequate informative polymorphism to evaluate genetic diversity of the studied accessions. Moreover, high values of heterozygosity and PIC statistics are a sign of marker informativeness, which is a desirable property in linkage association tests (Boopathi 2013). As heterozygosity is a measure of genetic variation within a population, its high average at a locus could be expected to correlate with high levels of genetic variation at loci with critical importance for adaptive response to environmental changes (Ojango et al. 2011). The results of this study are in close agreement with findings by Chen et al. (2011), who performed SNP genotyping on more than 300 rice inbred lines and obtained an average mean PIC value of 0.277 and 0.35 gene diversity compared with 0.25 and 0.32 for PIC and gene diversity, respectively, in the present study. However, mean allele number and PIC values were relatively low compared with other genetic diversity studies where SNPs were used. One of the reasons for low allelic variation may be due to the use of already released and locally adapted cultivars that have been exposed to high selection pressure, instead of landraces, wild relatives or segregating populations (Ram et al. 2007; Thomson et al. 2007).

In their study comparing the effectiveness of SSR and SNP markers in estimation of genetic diversity in rice, Singh et al. (2013) obtained PIC values ranging from 0.03 to 0.37 with an average PIC of 0.23 using SNP markers. These values were slightly below the values obtained in the present study. However, Singh et al. (2013) indicated that because of the bi-allelic nature of SNPs, PIC values can range from 0 to 0.5, compared with SSR markers which are multi-allelic and can have PIC values above 0.5 and up to 1.0. Consequently, results from the present study demonstrated that the set of SNPs used were sufficiently informative and can thus be used as a tool for large-scale genotyping in rice molecular breeding research involving japonica × japonica, indica × japonica and indica × indica crosses.

Based on the cluster analysis, SNP markers in the present study were useful in revealing two distinct major genetic groups. This may enable breeders to design targeted crosses for development of ShR-resistant genotypes, while conserving genetic diversity. The polymorphism observed in this study can also be attributed to the fact that the germplasm consisted of morphologically diverse subspecies of *Oryza sativa*, that is, subspp. *japonica* and *indica*. Obviously, indica and japonica have evolved from two partially isolated gene pools (Vaughan et al. 2008). Being adapted to different environments, indica and japonica cultivars have diverse morphological, agronomical, physiological and molecular characteristics (Oka and Morishima 1982; Lin et al. 2012) that provide valuable genetic resources for hybridisation in various breeding programmes.

Clear SNPs-based distinction between the *indica* and *japonica* subspecies of *Oryza sativa* has been previously described by Feltus et al. (2004) and Chen et al. (2011). In these studies, SNP markers revealed some common and contrasting patterns of haplotype diversity along different rice chromosomes in the indica and japonica accessions, which suggest different evolutionary forces possibly acting in specific regions of the rice genome during domestication and evolution of rice. In a different study, subgroups within the indica group based on SSR and SNP markers were reported by Singh et al. (2013).

Within the indica and japonica types, variability is probably a result of pedigrees that evolved in different gene pools in the same subspecies, as suggested by Lu et al. (2009). The indica cultivars from the International Rice Research Institute, The Philippines, were closely related because of selection under similar environments for specific breeding aims (Lin et al. 2012). Moreover, in the present study Rumbuka and Moroberekan were clustered with japonica accessions, but are accessions of indica rice and *Oryza glaberrima*, respectively'. The Nerica cultivars were indicated to have a close relationship with indica accessions but actually are crosses between *Oryza glaberrima* and *O. sativa*. This finding indicates that the Nerica cultivars may possess an indica genome in their parental genomic make up. In the same context, Rumbuka is indicated to possess, in its pedigree, a parental genome very close to subsp. *japonica* and, therefore, is doubtfully classified as an indica rice type, as previously reported by Ndikumana and Gasore (2010). The close relationship between Moroberakan and japonica cultivars has also been reported by McNally et al. (2009) and Arai-Kichise et al. (2014). This is an indication that morphological classification must be coupled with molecular characterisation to avoid biased assumptions. This corroborates suggestions by Lu et al. (2009), according to which molecular markers are not developed specifically based on differences between the two subspecies of rice and, therefore, diverse results are often obtained when molecular markers are used for identification of indica and japonica rice cultivars. Despite this observation, molecular markers are still useful in clustering the two subspecies into two separate groups.

Regarding ShR disease, given that indica types are generally susceptible and japonica cultivars show different levels of resistance, the genetic diversity among both groups as confirmed by SNP markers is a new development towards cultivar improvement in rice for resistance to the disease. Cultivars at both ends of the dendrogram can be considered as potential parental materials for this purpose in any breeding strategy used. For instance, hybridisation programs targeting improvement of indica cultivars used in this study for ShR, except Rumbuka, should involve cultivars such as Yunyine, Yunertian and Moroberekan due to their genetic distances. This could lead to reduced risks of genetic depression and increased diversity in the resulting progenies.

Conclusion

This study concluded that 85 out of the 94 SNPs markers used were highly informative and sufficiently polymorphic to distinguish relationship groups from 25 rice cultivars that are being considered as potential parental lines in breeding for sheath rot resistance programs. The studied accessions revealed the existence of high genetic variability that can be exploited for crop improvement with minimised risks of genetic depression and reduced diversity among progenies. The information generated will contribute significantly to further breeding studies mainly in determination of gene action and nature of inheritance governing resistance to sheath rot. It will also be helpful to design an adequate breeding strategy to introgress sheath rot resistance genes into popular cultivars. As a result, three cultivars, namely Yunyine, Yunertian and Moroberekan, are recommended as good candidate sources of resistance genes for improvement of most of the indica cultivars, except Rumbuka. The improvement of Rumbuka, as a cultivar of high yield potential, should lead to better results when hybridised with Nyiragikara. In conclusion, these are valuable findings that will give a head start to the rice breeding program in Rwanda towards breeding for resistance to sheath rot disease and the sources of resistance can also be shared with other breeding programs where the disease is becoming a problem.

Acknowledgements — The Alliance for a Green Revolution in Africa (AGRA), African Centre for Crop Improvement (ACCI) and University of KwaZulu-Natal are gratefully acknowledged for funding and supervising this study. Special thanks also go to the Rwanda Agriculture Board (RAB) for facilitating this study.

References

Arai-Kichise Y, Shiwa Y, Ebana K, Shibata-Hatta M, Yoshikawa H, Yano M, and Wakasa K. 2014. Genome-wide DNA polymorphisms in seven rice cultivars of temperate and tropical japonica groups. *PLoS ONE* 9: e86312.

Ayyadurai N, Kirubakaran SI, Srisha S, Sakthivel N. 2005. Biological and molecular variability of *Sarocladium oryzae*, the sheath rot pathogen of rice (*Oryza sativa* L.). *Current Microbiology* 50: 319–323.

Bigirimana VdP, Hua KHG, Nyamangyoku OI, Höfte M. 2015. Rice sheath rot: an emerging ubiquitous destructive disease complex. *Frontiers in Plant Science* 6: 1066–1082.

Boopathi NM. 2013. G*enetic mapping and marker assisted selection: basics, practice and benefits*. New Delhi: Springer India.

Botstein D, White RL, Skolnick M, Davis RW. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32: 314–331.

Chen H, He H, Zou Y, Chen W, Yu R, Liu X, Yang Y, Gao YM, Xu JL, Fan LM, Li Y, Li ZK, Deng XW. 2011. Development and application of a set of breeder-friendly SNP markers for genetic analyses and molecular breeding of rice (*Oryza sativa* L.). *Theoretical and Applied Genetics* 123: 869–879.

Feltus FA, Wan J, Schulze SR, Estill JC, Jiang N, Paterson AH. 2004. A SNP resource for rice genetics and breeding based on subspecies indica and japonica genome alignments. *Genome Research* 14: 1812–1819.

Gnanamanickam SS. 2009. *Biological control of rice diseases*. Dordrecht: Springer.

Guimarães PE. 2009. Rice breeding. In: Carena JM (ed.), *Cereals*. New York: Springer. pp 99–126.

Hall BG. 2013. Building phylogenetic trees from molecular data with MEGA. *Molecular Biology and Evolution* 30: 1229–1235.

Hittalmani S, Mahesh HB, Mahadevaiah C, Prasannakumar MK. 2016. *De novo* genome assembly and annotation of rice sheath rot fungus *Sarocladium oryzae* reveals genes involved in helvolic acid and cerulenin biosynthesis pathways. *BMC Genomics* 17: 271–284.

Kilian B, Graner A. 2012. NGS technologies for analyzing germplasm diversity in genebanks. *Briefings in Functional Genomics* 11: 38–50.

Lakshmanan P, Velusamy R. 1991. Resistance to sheath rot (ShR) of breeding lines derived from *Oryza officinalis*. *International Rice Research Newsletter* 16: 8–9.

Lanoiselet V, You MP, Li YP, Wang CP, Shivas RG, Barbetti MJ. 2012. First report of *Sarocladium oryzae* causing sheath rot on rice (*Oryza sativa*) in Western Australia. *Plant Disease* 96: 1382.

Lin H-Y, Wu Y-P, Hour A-L, Ho S-W, Wei F-J, Hsing C, Lin Y-R, 2012. Genetic diversity of rice germplasm used in Taiwan breeding programs. *Botanical Studies* 53: 363–376.

Linares OF. 2002. African rice (*Oryza glaberrima*): history and future potential. *Proceedings of the National Academy of Sciences of the USA* 99: 16360–16365.

Liu K, Muse SV. 2005. PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21: 2128–2129.

Lu BR, Cai X, Xin J. 2009. Efficient Indica and Japonica rice identification based on the InDel molecular method: its implication in rice breeding and evolutionary research. *Progress in Natural Science* 19: 1241–1252.

Madhav MS, Laha GS, Padmakumari AP, Somasekhar N, Mangrauthia SK, Viraktamath BC. 2013. Phenotyping rice for molecular plant breeding. In: Panguluri SK, Kumar AA (eds), *Phenotyping for plant breeding*. New York: Springer. pp 1–40.

McNally KL, Childs KL, Bohnert R, Davidson RM, Zhao K, Ulat VJ, Zeller G, Clark RM, Hoen DR, Bureau TE, Stokowski R, Ballinger DG, Frazer KA, Cox DR, Padhukasahasram B, Bustamante CD, Weigel D, Mackill DJ, Bruskiewich RM, Rätsch G, Buell CR, Leung H, Leach JE. 2009. Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. *Proceedings of the National Academy of Sciences of the USA* 106: 12273–12278.

Mvuyekure SM. 2016. Increasing the resilience of elite rice cultivars to sheath rot (*Sarocladium oryzae* [(Sawada) W. Gams & D. Hawksw]) in Rwanda through breeding for resistance. PhD thesis, University of KwaZulu Natal, Pietermaritzburg, South Africa.

Nascimento WFD, Silva EFd, Veasey EA. 2011. Agro-morphological characterization of upland rice accessions. *Scientia*

Agricola 68: 652–660.
Ndikumana I, Gasore ER. 2010. Description of new rice varieties introduced in Rwanda. Kigali: Institut des Sciences Agronomiques du Rwanda.
Nei M, Tajima F, Tateno Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* 19: 153–170.
Ngala GN, Adeniji MO. 1986. Sheath rot disease in tropical Africa. In: Juo ASR, Lowe JA (eds), *The wetlands and rice in subsaharan Africa*. IITA: Ibadan. pp 161–167.
Ojango J, Mpofu N, Marshall K, Andersson-Eklund L. 2011. Quantitative methods to improve the understanding and utilisation of animal genetic resources In: Ojango JM, Malmfors B, Okeyo AM (eds), *Animal genetic training resource*, version 3. International Livestock Research Institute, Nairobi. pp 1–38.
Oka HI, Morishima H. 1982. Phylogenetic differentiation of cultivated rice: potentiality of wild progenitors to evolve the indica and japonica types of rice cultivars. *Euphytica* 31: 41–50.
Pearce DA, Bridge PD, Hawksworth DL. 2001. Species concept in *Sarocladium*, the causal agent of sheath rot in rice and bamboo blight. In: Sreenivasaprasad S, Johnson R (eds), *Major fungal diseases of rice: recent advances*. Dordrecht: Kluwer Academic Publishers. pp 285–292.
Rafalski A. 2002. Applications of single nucleotide polymorphisms in crop genetics. *Current Opinion in Plant Biology* 5: 94–100.
Ram SG, Thiruvengadam V, Vinod KK. 2007. Genetic diversity among cultivars, landraces and wild relatives of rice as revealed by microsatellite markers. *Journal of Applied Genetics* 48: 337–345.
Reif JC, Zhang P, Dreisigacker S, Warburton ML, van Ginkel M, Hoisington D, Bohn M, Melchinger AE. 2005. Wheat genetic diversity trends during domestication and breeding. *Theoretical and Applied Genetics* 110: 859–864.
Ren J, Sun D, Chen L, You FM, Wang J, Peng Y, Nevo E, Sun D, Luo MC, Peng J. 2013. Genetic diversity revealed by single nucleotide polymorphism markers in a worldwide germplasm collection of durum wheat. *International Journal of Molecular Sciences* 14: 7061–7088.
Sakthivel N. 2001. Sheath rot disease of rice: current status and control strategies. In: Sreenivasaprasad S, Johnson R (eds), *Major fungal diseases of rice: recent advances*. Dordrecht: Kluwer Academic Publishers. pp 271–283.
Singh N, Choudhury DR, Singh AK, Kumar S, Srinivasan K, Tyagi RK, Singh NK, Singh R. 2013. Comparison of SSR and SNP markers in estimation of genetic giversity and population structure of indian rice varieties. *PLoS ONE* 8: e84136.
Soleimani DV, Baum RB, Johnson AD. 2002. AFLP and pedigree-based genetic diversity estimates in modern cultivars of durum wheat [*Triticum turgidum* L. subsp. *durum* (Desf.) Husn.]. *Theoretical and Applied Genetics* 104: 350–357.
Thomson M, Septiningsih E, Suwardjo F, Santoso T, Silitonga T, McCouch S. 2007. Genetic diversity analysis of traditional and improved Indonesian rice (*Oryza sativa* L.) germplasm using microsatellite markers. *Theoretical and Applied Genetics* 114, 559–568.
Vaughan D, Lu BR, Tomooka N. 2008. Was Asian rice (*Oryza sativa*) domesticated more than once? *Rice* 1: 16–24.
Weddell BJ. 2002. *Conserving living natural resources: in the context of a changing world*. Cambridge: Cambridge University Press.

Received 4 March 2016, revised 9 May 2017, accepted 13 May 2017
Associate Editor: Toi J Tsilo