A GENETIC STUDY FOR RESISTANCE TO AFRICAN RICE GALL MIDGE IN WEST AFRICAN RICE CULTIVARS

By

Nasser Kouadio Yao

M.Sc. Genetics and Plant Breeding (Genetics and Plant Breeding)

University of Abidjan-Cocody, Côte d'Ivoire

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD) in Plant Breeding

School of Agricultural, Earth and Environmental Sciences (SAEES)

College of Agriculture, Engineering and Science

University of KwaZulu-Natal

Republic of South Africa

April 2012

Thesis Abstract

The African Rice Gall Midge (AfRGM), *Orseolia oryzivora* Harris and Gagné (Diptera: Cecidomyiidae), is an endemic rice pest found throughout Africa. The failure of most other control methods imposes the need to use crop resistance. This study was initiated: (1) to develop an accurate method for assessing damage caused by AfRGM; (2) to determine AfRGM resistance genes' modes of action, the heritability estimates of their resistance to AfRGM and the behavioural pattern of progenies with resistance to AfRGM attack; (3) to reveal convergent evolution of same or similar resistance gene(s) in geographically distinct landraces, or divergent evolution of genotypes carrying the same gene, by analysing the genetic diversity among five AfRGM parental lines; (4) to build a core sample of progenies to be used as a reduced mapping population, largely reflecting the entire genome of the whole population, after an estimate of the heritability of 15 agro-morphological descriptors and; (5) determine Simple Sequence Repeat (SSR) markers flanking genes or quantitative trait loci (QTLs) linked to resistance to AfRGM.

A method of accurately assessing damage caused by AfRGM was determined by comparing four methods of assessment including the International Rice Research Institute's (IRRI) Standard Evaluation System (SES) for rice and three methods based on resistance index (RI) assessments differing in the computing of the percentage of tillers with galls on a resistant check variety. The RI-based assessment (RI-BA) methods consistently provided a better evaluation of AfRGM damage than the SES, regardless of the trial size. Within RI-BA methods, RI-BA2 was always more accurate than RI-BA1 and RI-BA3 when the plot was large. RI-BA2 and RI-BA3 were equally accurate when the plot size was small, and they provided better estimates than RI-BA1. When the plot was of medium size, RI-BA2 was more accurate than RI-BA3; RI-BA3 also surpassed RI-BA1. Overall, the best method of assessing AfRGM damage was RI-BA2, regardless of the plot size.

Five rice populations including F_1, F_2 and F_3 generations involving ITA306, a susceptible variety of *Oryza sativa* subsp. *indica*, and four varieties having different reactions against AfRGM were used to determine the genetic basis of resistance and estimate the heritability of resistance to AfRGM. All the F_1s were susceptible, suggesting recessive gene inheritance. The F_2 generations' segregation pattern of 1R:15S in both ITA306-TOS14519 and

ITA306-TOG7106 crosses as well as the segregation of 1R:8Seg:7S in ITA306-TOS7106 F_3 families indicated that the AfRGM resistance expression being studied is governed by two genes. The deviation of the segregation patterns of crosses involving ITA306 and the tolerant parental lines from Mendelian segregation ratios suggests that the tolerance to AfRGM shown by BW348-1 and Cisadane is under complex mechanisms of control rather than under simple genetic control. The narrow-sense heritability estimates of resistance to AfRGM were low in populations involving tolerant varieties and were high in populations involving resistant varieties. They ranged from 0.086 in the ITA306-Cisadane population, to 0.4 in the ITA306-TOG7106 population. Conversely, the broad-sense heritability estimates ranged from 0.23 (ITA306-Cisadane) to 0.63 (ITA306-TOS14519).

The behavioural patterns of progenies against AfRGM attack were evaluated for 532, 413 and 479 F_2 progenies from ITA306-BW348-1, ITA306-Cisadane and ITA306-TOS14519 crosses, respectively, in addition to 90 BC_1F_2 progenies from the ITA306 and TOG7106 cross. One F_3 generation of 649 families from a cross between ITA306 and TOS14519 was also tested. Four types of behavioural pattern categories were observed: (1) progenies were more resistant than the resistant check entry at 45 DAT and 70 DAT; (2) progenies were more resistant at 45 DAT and became susceptible at 70 DAT; (3) progenies were susceptible at both 45 DAT and 70 DAT; (4) progenies were susceptible at 45 DAT but reverted to resistant at 70 DAT. The first three categories were the most frequently observed and occurred in all cross combinations. The last category was observed only for a few progenies from the ITA306-TOS14519 F_2 and F_3 generations and, surprisingly, many from the ITA306 and BW348-1 cross.

Heritability estimates were calculated for 15 major traits in an F_3 population in order to predict the genetic gain associated with each trait, together with the resistance to AfRGM and to estimate the influence of the environment on phenotypic values. Broad-sense heritability (H^2) estimates were high for the penultimate leaf length (PLL) - 0.99, penultimate leaf width (PLW) – 1.0, flag leaf length (FLL) - 0.99, flag leaf width (FLW) – 1.0, ligule length (LigL) - 0.99, tillering ability (Til) - 0.99, number of days to booting (DB) - 0.95, number of days to first heading (DFH) - 0.96, number of days to heading (DH) - 0.89, number of days to maturity (DM) - 0.98, culm length (CL) - 0.99, plant height (PH) - 0.99, panicle length (PanL) - 0.95, secondary branching (SB) - 0.95 and the thousand grains weight (TGW) - 0.71. Conversely,

narrow-sense heritability estimates were very low (nearly 0) in PLL, FLL, Lig, DB, DFH, DM and SB or low (at most 0.267) in PLW, FLW, DH and PH, with a high value of 0.727 for TGW. Inheritance of the traits studied was therefore under non-additive gene effects rather than additive genetic effects and can therefore be improved using pedigree breeding schemes along with breeding for AfRGM resistance.

Fine genetic evaluation of five AfRGM parental lines was studied in terms of polymorphisms using 303 SSR primers covering the rice genome. Of the 178 polymorphic primers identified, 60 were highly polymorphic and informative. The number of alleles amplified by these primers ranged from one to five for a total of 1,041 alleles. The polymorphism rate was globally high, ranging from 45.2% to 66.8%. The mean of the polymorphism information content (PIC) was 0.553. Factorial analysis, based on the allelic diversity, demarcated the parental lines into *Oryza glaberrima* Steud, *Oryza sativa* subsp. *japonica* and *O. sativa* subsp. *indica* groups, while a cluster analysis distinguished them into four groups: AfRGM resistant, susceptible, moderately resistant and tolerant. BW348-1 and Cisadane showed the least diversity, despite their distant geographical origins. TOS14519 and TOG7106 showed more divergence to ITA306 despite their common West African origin. This variability amongst the genotypes tested is the result of farmer-based selection for AfRGM resistance rather than direct breeding efforts through breeder intervention.

A method of selecting individuals for a mapping population, based on a core sample, was developed in order to speed up the mapping procedure. A diversity study amongst F_2 and F_3 generations involving 15 quantitative and 26 qualitative agro-morphological characters was carried out and led to the dropping of seven non-discriminant descriptors. The diversity index (H) was calculated for each remaining character and the discriminant descriptors were selected based on a diversity index threshold value above 0.4. Four descriptors of H values less than 0.35 were therefore dropped. The sizing of the core collection of 64 individuals and the selection of these individuals were done using MSTRAT version 4.1 package in redundancy mode, a construction run of 100 times with an iteration number of 500. The core sample was similar to the whole population for clustering pattern, minimum and maximum quantitative values and diversity index, while mean values and coefficient of variation distinguished them. The core sample, which represents 10% of the whole population, also revealed the same

phenotypic variation and the same genotypic segregation according to two SSR markers. It can therefore efficiently reflect the whole population as a mapping population.

Finally, a study was undertaken to identify flanking markers to the gene/QTL involved in the resistance against AfRGM using bulked segregant analysis (BSA). A polymorphism study between ITA306 and TOS14519 displayed 145 polymorphic SSR markers, which were used to screen the bulks that originated from the two tails, and depicted only two SSRs as candidate markers linked to gall midge resistance. These markers included RM317 and RM17303 which displayed strong significance after an analysis of variance using an F test, meaning that they were segregating with the resistant alleles.

Declaration

I, **KOUADIO NASSER YAO** declare that:

i. The research reported in this thesis, except where otherwise indicated, is my original work.

ii. This thesis has not been submitted for any degree or examination at any other university.

iii. This thesis does not contain any other persons' data, pictures, graphs or information unless specifically acknowledged as being sourced from other persons.

iv. This thesis does not contain any other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

a) their words have been rewritten but the general information attributed to them has been referenced;

b) where the exact words have been used, their writing has been placed inside quotation marks, and referenced.

v. Where I have reproduced a publication of which I am an author, co-author or editor, I have indicated in detail which part of the publication was actually written by myself alone and have fully referenced such publications.

vi. This thesis does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged and the sources detailed in the thesis and references section.

…… Date……

As the candidate's supervisors, we agree to the submission of this thesis.

…… Date……

Professor Mark Laing (Supervisor)

…… Date……

Dr Marie-Noëlle Ndjiondjop (Co-supervisor)

…… Date……

Dr Francis Nwilene (Co-supervisor)

Acknowledgements

This study was made possible by a grant from United States Aid (USAID) through AfricaRice funding.

I gratefully appreciate the supervisory team of Dr Marie-Noelle Ndjiondjop, Dr Francis Nwilene and Professor Mark D. Laing who found the time to support this study regardless of their tight time tables.

I hereby acknowledge the support of Mr Kehinde Adebayo from AfricaRice-Ibadan and Mrs Euloge Ahanhanzo, Appolinaire Yavoga and Roland Bocco from AfricaRice-Cotonou for their invaluable assistance in data collection and field maintenance. The lab staff of AfricaRice, in particular, Ms Marie Adjoua Gouete, Cissoko Mamadou, Olufisayo Kolade, Blandine Fatondji, Mrs Loris Onésime Mahuna Gbedjissi and Emmanuel Wilfrid Mehigbeto are acknowledged for their support during this study at the AfricaRice Main Biotechnology Laboratory, Cotonou, Benin. I also appreciate the effort and assistance of AfricaRice for providing the facilities, the genetic materials used in this study and the technical assistance. I also express my heartfelt gratitude to "Mon Ami" Dr Eklou A. Somado for his support and prayers.

I would also like to thank all my PhD colleagues and office mates including Bella Teeken, David Koffi Montcho, Esperance Zossou, El-Hassimi Sow Mounirou, Florent Okry, Gustave Djedatin, Honoré Kam, Issaka Souley and "my dearest and lovely spouse" Yonnelle Dea Moukoumbi for the conviviality of their interaction with me and their invaluable input on scientifically related subjects.

Above all, I thank the Almighty God for the gift of "iron health", my dearest wife Sylvie Koffi for always being a patient wife; and pay tribute to my son, Ryan Nathan Moyet YAO ASSAI, for having suffered from my long and repetitive absence throughout the duration of the PhD. I am also grateful to my parents: the KOUADIO YAO family in Bouaké, the NANGLE family in Toumodi and Daloa and the KOUAME Alfred family in Abidjan for their financial and moral support without which I would not have been able to get educated. Last but not the least, my gratitude also goes to the SEHI's and KOFFI's families. "Mon Pied" Koffi Mathurin, this thesis is yours. You "initiated" it and you have never failed to support me.

For others whose names could not be mentioned here, I say thank you.

Dedication

This thesis is dedicated to my late father,

who would have loved to see the final version of this PhD.

Table of Contents

Thesis Introduction

0.1 Importance, production and consumption of rice worldwide

Rice (*Oryza sativa* L.) is cultivated on all continents except Antarctica. Together, rice and wheat represent the major cereal crop production in the world (Mendez, 2010) and are the basic food for 75% of the world's population (Ju-Kon and Krishnan, 2002; Correa *et al*., 2007). It provides food for almost half of the world's population (Schalbroeck, 2001; Jena and Mackill, 2008). In developing countries, rice is the major source of calories for about 820 million households (FAO, 2006). Rice production worldwide is of about 710 million tons of paddy rice (FAO, 2010) over an area of more than 155 million ha (IRRI 2009).

In Asia, around 90% of world rice is consumed and produced (FAOSTAT, 2009), with China and India growing more than half of the total production.

In Africa, rice accounts for 24% of food intake in households (FAO, 2004) but represents just 6% (9.4 million ha) of the world rice-growing area (156.7 million ha) (FAOSTAT, 2008), and only 3.2% in terms of world production (22.7 of 710 million tons).

Rosegrant *et al*. (1995) estimated the actual sub-Saharan Africa (SSA) rice production (17.2 million tons annually) to be about 22.23 million tons by the year 2025 while the estimates of future increase in world demand for the same period are 75% (Khush, 2005). This suggests that SSA need to increase rice production from the current 17.2 to almost 38.2 million in fifteen years (IRRI, 2009).

In West Africa, rice consumption is increasing at an annual rate of 4.5% (AfricaRice, 2010). More than seventeen (17.2) million tons of milled rice is consumed annually, while the production for the same period is 13.2 million tons (AfricaRice, 2008). Therefore, West African countries have to resort to imports to meet 40% of their rice demand and have spent at least US$ 2 billion per year on rice purchase over the last five years (AfricaRice, 2010), with a record of US$ 3.6 billion in 2008 (Diagne *et al.*, 2010). In 2006, all Africa accounted for about 30% of world rice imports (FAOSTAT, 2009) for only 14% of the world's population (PopulationData, 2010).

0.2 Constraints associated to rice production in West Africa

Rice farming in West Africa is subjected to many abiotic and biotic constraints. Abiotic constraints include (1) global warming which effect increase drought, temperature and flooding, (2) water scarcity and shrank area of arable rice lands due to urbanisation and industrialisation. Biotic constraints comprise weeds, occurrence of new biotypes of diseases and insect pests, including the African rice gall midge (AfRGM), *Orseolia oryzivora* Harris and Gagné (Diptera: Cecidomyiidae).

AfRGM is a destructive insect pest of lowland rice, and causes yield losses between 45-80% in farmers' fields in West Africa (Nwilene *et al*., 2006). AfRGM was first reported in Sudan (Joyce, 1954), before being described and illustrated morphologically 28 years later in 1982 by Harris and Gagné as a distinct species from the Asian species, *Orseolia oryzae*. According to Williams *et al.* (2002), 20 African countries including twelve West African, two Central African and six East and Southern African countries are affected by AfRGM, which has not yet been recorded outside of the continent. In West Africa, this pest is widespread in the forest zone, the Guinea savannah, and the Sudan savannah, where its infestation level is low compared to the two first zones (Harris, 1960, 1980; Alam and Alluri., 1985; Ukwungwu *et al*., 1989; Williams *et al.*, 2002).

Damage caused by AfRGM was initially restricted to lowland rice (Peter *et al.*, 1991) with little damage to upland rice (Bonzi, 1980). However, the AfRGM has spread to the upland cropping system (F.E. Nwilene personal communication), increasing its impact. In the most affected countries (Burkina Faso, Cameroon, Mali, Nigeria and Sierra Leone) field infestations have been evaluated at above 40% incidence (Ukwungwu and Joshi, 1992; Dakouo *et al.*, 1988). Depending on weather conditions, a 1% incidence of infested tillers equal a 2% yield loss (Nacro *et al.* 1996) and a further increase of 1% in infestation results in a 2.9% yield loss (Williams *et al.*, 1999). In certain regions, severe attacks led to a total loss of the harvest (Ethel *et al.*, 1993). In the face of a rapidly growing population and the need to reduce poverty among the smallholder farmers, pest-induced losses due to AfRGM must be tackled.

0.3 New breeding directions to sustain rice production in West Africa

Despite, its production of 13.2 million tons, constraints on rice sector in West Africa have been tackled to substantially reduce gap between production and consumption. To make rice available to the millions of people that rely on it for their daily intake, efforts using

molecular-based technologies (Nguyen, 2005) and low cost methodologies for varietal release and adoption in Africa must be undertaken.

In West Africa, genetic improvement of rice, especially the development of the New Rice for Africa named NERICA, and improved *O. sativa* varieties, have been tested using a participatory varietal selection (PVS) approach (AfricaRice, 2010). These developments have accounted for the growth of rice production to 9.9 million tons in 2008 (FAOSTAT, 2009). NERICA alone contributed up to a 50% yield increase for two million hectares of African rice production (Sarla and Mallikarjuna Swamy, 2005). The increased production, however, has not been enough to meet the demand over the last decade (FAOSTAT, 2009; AfricaRice, 2010). Therefore, the constraints, associated with molecular-based methodologies for varietal release in West Africa, must be addressed using genomics, transformation and molecular breeding tools. The recent progresses of new breeding approaches and the availability of new tools plead in favour of such approach. These progresses and tools include:

1) Cultivars development through tissue culture, transgenesis and genetic transformation of *O. sativa* subsp. *indica* and *japonica* cultivars (Hiei *et al*., 1999; Waterhouse and Upadhyaya, 1999; Klöti and Potrykus, 1999);

2) Construction of a comprehensive genetic and physical map, a high density molecular map for gene mapping and gene cloning;

3) Development of bacterial artificial chromosome (BAC) and yeast artificial chromosome (YAC) libraries (Kurata *et al.*, 2002; Jena and Mackill, 2008);

4) Investigation of synteny or co-linearity of rice genes along the chromosomes with other cereals such as wheat (*Triticum aestivum* L.), pearl millet (*Pennisetum glaucum* L.), maize (*Zea mays* L.), sorghum (*Sorghum bicolor* L.) and barley (*Hordeum vulgare* L.) (Moore *et al*., 1995; Gale and Devos, 1998a; 1998b);

5) Development of the *Oryza* Map Alignment followed by the complete sequencing of the rice genome in *O. sativa* subsp. *indica* and *japonica* cultivars, and the annotation of gene sequences (International Rice Genome Sequencing Project, 2005).

0.4 Justification for the study

Several control measures have been advocated for the management of AfRGM. Of these, varietal resistance/tolerance is seen as the most farmer-friendly option. Highly resistant varieties, all from the African species *O. glaberrima*, have been found but are not popular with

farmers because of their susceptibility to lodging and their low yield due to spontaneous shattering. Other screenings have identified AfRGM resistance in some *O. glaberrima* varieties and have found moderately resistant varieties of *O. sativa* (Maji *et al*., 1998; Ukwungwu *et al.*, 1998; Nwilene *et al*., 2002). Multi-location studies conducted by Williams *et al.* in 1999 and later by Nwilene *et al*. (2002) revealed resistant varieties, including the traditional variety TOS14519 from the Gambia. With regard to mean susceptibility in farmers' fields, Cisadane from Indonesia is a tolerant variety released in Nigeria as FARO 51 (Omoloye *et al.*, 2002). This variety of *O. sativa* subsp. *indica* produces an acceptable yield for farmers (Williams *et al.*, 1999). However, Cisadane is very susceptible to iron toxicity, which is one of the abiotic constraints prevalent in the lowland cropping system of West Africa. BW348-1 from Sri Lanka is both tolerant to AfRGM and iron toxicity (Singh, 1998), but is less productive. Transferring AfRGM resistance genes from *O. glaberrima* or resistant *O. sativa* lines into high potential yielding *O. sativa* varieties could prevent damage by this pest and could complement other control methods such as cultural practices and biological control. However, to date, no *O. sativa* nor *O. glaberrima* variety has been found to be highly resistant or highly susceptible to AfRGM, respectively, limiting progress on marker assisted breeding (MAB) for resistance to the pest. A study on the genetic/molecular basis of resistance to AfRGM would help in choosing the appropriate techniques for transferring the genes. Mapping and tagging the resistance genes followed by their transfer into an *O. sativa* background using MAB could provide an approach to AfRGM management. In the present study, molecular mapping approaches used in breeding for pest resistance were used as instruments to help in developing new AfRGM resistant varieties suitable for rice farmers in West Africa.

0.5 Goal and objectives of the study

The overall goal of the study was to identify, map and transfer the AfRGM resistance gene into African elite varieties using both marker assisted selection (MAS) and conventional breeding approaches, to minimise the impact of AfRGM in rice-based systems for the benefit of smallholder farmers.

The specific objectives of the study .include:

1. Developing a new AfRGM assessment methods to find the best method to accurately estimate the degree of pest infestation during phenotyping;

2. To use this method of screening to determine the genetic basis and the heritability estimate of known resistance to AfRGM and to estimate the behavioural pattern of parents and progenies from early filial generation;
3. Estimating the genetic diversity among rice genotypes with a range of resistance expression to AfRGM;
4. To investigate the heritability of common agro-morphological traits, along with the resistance to AfRGM;
5. To develop a novel mapping approach using a core collection that represents a reduced mapping population;
6. Apply the new mapping approach to identify flanking SSR markers associated to the resistance to AfRGM.

0.6 Structure of the thesis

The thesis introduction outlines the economic importance of rice and points out its major constraints including AfRGM. It then establishes the primary goals of the thesis. Chapter 1 provides a review of literature and reveals that relatively little research has been done and information is lacking, most particularly on the genetic base of resistance to AfRGM. Alternative means for assessing AfRGM damage accurately were analysed and compared in Chapter 2, as a basis for accurate phenotyping. Based on these assessment methods, the nature and the inheritance of resistance to AfRGM were assessed through an appropriate population (Chapter 3). In parallel to the genetic study in Chapter 3, heritability estimates of agro-morphological traits were investigated to see if these traits could be transmitted along with AfRGM resistance (Chapter 4). The response to AfRGM attack for both parental lines and progenies were investigated in Chapter 5, while a diversity study among the parental lines is covered in Chapter 6. The diversity study aimed to reveal convergent evolution of identical or similar resistance gene(s) in geographically distinct landraces, or conversely the divergent evolution of genotypes carrying the same AfRGM resistance gene. Chapter 8 deals with the identification of molecular markers flanking quantitative trail loci (QTL) linked to AfRGM resistance using a mapping population previously derived from a core collection based on the MSTRAT approach (Chapter 7). The Thesis Overview is presented as the summary of research findings with reference to the initial goal, the implication of the findings on AfRGM resistance rice breeding in Africa, and future recommendations.

References

AfricaRice (WARDA). (2008). *Africa Rice Trends 2007*. Africa Rice Center (WARDA), Cotonou, Benin. 84 p.

AfricaRice (2010). *New Breeding Directions at AfricaRice: Beyond NERICA*. Africa Rice Center, Cotonou, Benin: 24 p.

Alam, M.S. and Alluri, K. (1985). Gall midge *Orseolia oryzivora* Harris and Gagné in Zambia. *International Rice Research Newsletter* 10:15-16.

Bonzi, S.M. (1980). Wild host plants of the rice gall midge *Orseolia oryzae* W.M. (Diptera, Cecidomyiidae) in Upper Volta. *WARDA Technical Newsletter 2*:5-6.

Correa, P.C., Schwanz da Silva, F., Jaren, C., Afonso Junior, P.C. and Arana, I. (2007). Physical and mechanical properties in rice processing. *Journal of Food Engineering* 79:137-142.

Dakouo, D. Nacro, S. and Sié, M. (1988). Evolution saisonnière des infestations de la cécidomyie du riz, *Orseolia oryzivora* Harris and Gagné (Diptera, Cecidomyiidae) dans le Sud-ouest du Burkina Faso. *Insect Science Application* 9:467-472.

Diagne, A., Midingoyi, S.-K.G., Wopereis, M. and Akintayo, I. (2010). *The NERICA Success Story: Development, Achievements and Lessons Learned.* Africa Rice Center (AfricaRice), Cotonou, Benin. 29 p

Ethel, D., Umeh, N. and Joshi, R.C. (1993). Aspect of biology and natural biological control of the African rice gall midge, *Orseolia oryzivora* Harris and Gagné (Dipteria Cecidomyiidae) in south east Nigeria. *Applied Entomology* 116:391-398

FAO (2004) Vingt-troisième conférence régionale pour l'Afrique. FAO, Johannesburg 2004. http://www.fao.org/rice *2004/en/world.htm/* accessed August 2010.

FAO (2006). The state of food insecurity in the World 2006. Eradicating world hunger: taking stock ten years after the World Food Summit. 40p http://www.fao.org/icatalog/inter-e.htm/ last accessed June 2010.

FAO (2010). Preliminary 2009 data now available for selected countries and products, FAO Statistics Division, Rome.

FAOSTAT (2008). Database, FAO, Rome. 22 Sep 2008/ last accessed June 2010.

FAOSTAT (2009). (www.faostat.fao.org)/Agriculture and Food Trade/ last accessed June 2010.

Gale, M.D. and Devos, K.M. (1998a). Plant comparative genetics after 10 years. *Science* 282:656-659.

Gale, M.D. and Devos K.M. (1998b). Comparative genetics in the grasses. *Proceeding of the National Academy of Science, USA* 95:1971-1974.

Harris, K.M. (1960). Notes on gall midge Cecidomyidae on Nigeria crops with a description of a new species of *Thomasiella rubsaamen. Bulletin of Entomological Research 50*:661-666.

Harris, K.M. (1980). Family Cecidomyiidae in Crossekey. *Catalogue of the Diptera of the Afrotropical Region.* British Museum (Natural History), London, United Kingdom 18:238-251.

Hiei, Y., Komari T., and Kubo T. (1999). *Agrobacterium*-mediated transformation. In: *Molecular Biology of Rice*, Shimamoto, K. (Ed). Springer-Verlag,Tokyo, Japan. p. 235-256.

International Rice Genome Sequencing Project (2005). The map-based sequence of the rice genome. *Nature* 436:793-800.

IRRI World Rice Statistics, (2009). http://www.irristat.org/ last accessed August 2010.

Jena, K.K. and Mackill D.J. (2008). Molecular marker and their use in marker-assisted selection in rice. *Crop Science* 48:1266-1276.

Joyce, R.J.V. (1954.) The rice gall fly, a new record. *FAO Plant Protection Bulletin* 2:122.

Ju-Kon, K. and Krishnan, H.B. (2002). Making rice a perfect food: turning dreams into reality. *Journal of Crop Production* 5:93-130.

Khush, G.S. (2005). What it will take to feed five billions rice consumers by 2030. *Plant Molecular Biology* 59:1-6.

Klöti, A. and Potrykus, I. (1999). Rice improvement by genetic transformation. In: *Molecular Biology of Rice*. Shimamoto, K. (Ed). Springer-Verlag, Tokyo, Japan. p. 283-301.

Kurata, N., Nonomura, K. and Harushima, Y. (2002). Rice genome organization: the centromere and genome interactions. *Annals of Botany* 90:427-435.

Maji, A.T., AKen'ova, M.E. and Singh, B.N. (1998). Variability studies in African rice *Oryza glaberima* Steud. Using principal component analysis. *Nigerian Journal of Genetic* 111: 34-42.

Mendez, D.V.P. (2010). Les marchés mondiaux. La renaissance du Palais d'Eté. In: *Cyclope 2010*. Economica, Paris, France. p. 235-240.

Moore, G., Devos, K.M. Wang, Z. and Gale, M.D. (1995). Cereal genome evolution, grasses, line up and form a circle. *Current Biology* 5:737-739.Nacro, S., Dakouo, D. and Heinrichs, E.A. (1995). Population dynamics, host plant damage and parasitism associated with the African rice gall midge in southern Burkina Faso. *Insect Science Application* 16:251-257.

Nacro, S., Heinrichs, E.A. and Dakouo, D. (1996). Estimation of rice yield losses due to the African rice gall midge, *Orseolia oryzivora* Harris and Gagné. *International Journal of Pest Management* 42:331-334.

Nguyen, N. (2005) http://www.fao.org/newsroom/en/news/2006/1000379/index.html/ accessed November 2010.

Nwilene, F.E., Williams, C.T., Ukwungwu, M.N., Dakouo, D., Nacro, S., Hamadoun, A., Kamara, S.I., Okhidievbie, O, Abamu, F.J. and Adam, A. (2002). Reactions of differential rice genotypes to African rice gall midge in West Africa. *International Journal of Pest Management* 48:195-201.

Nwilene, F.E., Harris, K.M., Okhidievbie, O., Onasanya, A., Sere, Y. and Ingelbrecht. I. (2006). Morphological diversity and genomic DNA fingerprinting of the African rice gall midge *Orseolia oryzivora* (Diptera: Cecidomyiidae) and of two other species of African *Orseolia*. *International Journal of Tropical Insect Science* 26:256-265.

Omoloye, A.A., Odebiyi, J.A., Williams C.T., and Singh, B.N. (2002). Tolerance indicators and responses of rice cultivars to infestation by the African rice gall midge, *Orseolia oryzivora. Journal of Agricultural Science* 139:335-340.

Peter, I., Mark, N., Ukwungwu, M.N. and Joshi, C. (1991). Criblage au champ pour la resistance à la cecidomyie africaine du riz. *Guide recherche de l'IITA,* No. 34.

PopulationData, (2010). http//www.populationdata.net/last accessed November 2010.

Rosegrant, M.W., Agacoili-Sombilla M. and Perez. N. (1995). Global food projections to 2020: Implications for Investments. IFPRI, Washington, D.C., USA.

Sarla, N. and Mallikarjuna Swamy, B.P. (2005). *O. glaberrima*: a source for the improvement of *O. sativa*. *Current Science* 89:955-963.

Schalbroek, J.J. (2001). Rice *Oryza sativa* L. In: *Crop Production in Tropical Africa.* Romain H. Raemaekers (Ed). DGIC, Ministry of Foreign Affairs, External trade and International Cooperation, Brussels, Belgium, p 59-78.

Singh, B.N. (1998). Breeding for rice gall midge (*Orseolia oryzivora* Harris and Gagné) resistance at WARDA. *Paper presented during the WARDA/NARS group discussion and planning meeting on Bioecology and management of African rice gall midge, 12-16 October at IITA, Ibadan, Nigeria*.

Ukwungwu, M.N., Winslow, M.D. and John, V.T. (1989). Severe outbreak of rice gall midge in savana zone, Nigeria. *International Rice Research Newsletter* 14:36-37.

Ukwungwu, M.N. and Joshi, R.C. (1992). Evaluation of Asian gall midge (AsGM)-resistant varieties for the African rice gall midge (ARGM). *International Rice Research Newsletter* 17:12-14.

Ukwungwu, M.N., Williams, C.T. and Okhidievbie, O. (1998). Screening of African rice, *Oryza glaberrima* Steud, for resistance to the African rice gall midge *Orseolia oryzivora* Harris and Gagné. *Insect Science Application* 18:167-170.

Waterhouse, P.M. and Upadhyaya, N.M. (1999). Genetic engineering of virus resistance. In: *Molecular Biology of Rice.* Shimamoto, K. (Ed.). Springer-Verlag, Tokyo, Japan. p. 257-281.

Williams, C.T., Okhidievbie, O., Harris, K.M. and Ukwungwu, M.N. (1999). The host plant range, annual cycle and parasitoids of the African rice gall midge *Orseolia oryzivora* (Diptera: cecidomyiidae) in central and southeast Nigeria. *Bulletin of Entomological Research* 89:589-597.

Williams, C.T., Harris, K.M., Ukwungwu, M.N., Nacro, S., Dakouo, D., Nwilene, F.E., Singh, B.N. and Okhidievbie, O. (2002). *African rice gall midge research guide*. Bouaké, Côte d'Ivoire: West Africa Rice Development Association. CAB International, Wallingford, United Kingdom 28 p.

Chapter 1

Review of the literature

1.1 Origins, genetics, and domestication of rice

Consumption of Asian rice (*Oryza sativa* L.) is believed to have originated in Asia about 10,000 years ago (Wang *et al*., 1999). Its cultivation begins as early as 4000 B.C. (Hiroshi, 1993) and was introduced to the Greeks around 326 B.C. From there, it spread across Europe over the following 2000 years (Vaughan, 1994). These rice crops swept south and westward, before reaching the west coast (Chang, 1976). It has become an important crop in several American states, with the development of irrigation methods (Porteres, 1956).

Rice belongs to the genus, *Oryza,* one of the 12 genera within the tribe Oryzeae (Vaughan, 1994) of the *Poaceae* family. The number of species is not clear and it varies from author to author. The genus *Oryza* has about 20-24 wild species (Chang, 1976; Vaughan, 1994) and two cultivated species (Vaughan, 1994; Khush, 1997). Species are divided into two big groups (Tateoka, 1962; IRRI, 1964): (1) the *Sativa* group that includes species of AA diploid (2n = 24) genome type; (2) the *Latifolia* group that include species having B, C and D genomes in the diploid and tetraploid state (2n = 24; 2n = 48). Cultivated rice belongs to two species *Oryza sativa* L. and *Oryza glaberrima* Steud of the S*ativa* complex. *O. sativa,* of genotype AA, originally from Asia, was introduced to Africa by the Portuguese in 1500 B.C. (Porteres, 1956). Its ancestor is *Oryza rufipogon* Griff. (Morishima, 1969; Nayar, 1973; Second, 1982) and comprises (a) the *Oryza sativa* subsp. *indica* group that originated from tropical Asia, characterised by high tillering ability, long and thin grains; and b) the *Oryza sativa* subsp. *japonica* group that originated from the temperate and sub-tropical zone of Asia. Its tillering ability is medium and it has short and round grains. Japonica types are well adapted to the rainfed, upland conditions in West Africa.

O. glaberrima, with the genotype AgAg, has a centre of origin in West Africa (van Andel, 2010). Domesticated in the central delta of the Niger River in Mali (Bonjean and Picard, 1990), its cultivation has been spread from Cap Vert to Chad (Porteres, 1956), the Gambia, Casamance and the Sokoto Basin (Carpenter, 1978). Cultivated from at least 3500 years ago, *O. glaberrima* landraces are gradually being displaced by *O. sativa* (Sié, 1991). However, it is still

widely cultivated in West Africa because of its resistance to drought and excess water, ability to weather acidic soil, iron toxicity and heat. It can also resist blast, *Helminthosporium* leaf spot, rice yellow mottle virus (RMYV), insects and nematodes. With its high level of plant vigour (Pham, 1992, Adeyemi and Vodouhe, 1996, Jones *et al*., 1997), it has the ability to contain and tolerate weeds. Its ancestor is *Oryza barthii* A. Chev. (previously *Oryza breviligulata* Chev.).

Apart from their different geographic origins, *O. glaberrima* and *O. sativa* are morphologically different. The ligule is short, roundish and tough in *O. glaberrima* but is pointed, bifid and long in *O. sativa*. At maturity, *O. glaberrima* shows an open panicle, whereas it is closed in *O. sativa*. In general, a higher number of spikelets are observed on the panicle of the Asian species (Porteres, 1956) (Table 1-1).

Table 1-1: Comparison among the Asian species *O. sativa* and the African species *O. glaberrima*

Traits	*Oryza sativa*	*Oryza glaberrima*
Distribution	Cosmopolitan	Endemic to West Africa
Cycle	Annual	Annual
Varietal difference	High	Low
Lodging	Low	High
Drought tolerance	Weak	Strong
Ligule type	Pointed, bifid and long	Short, round and tough
Secondary branching	High	Low
Number of grains	High (250/panicle)	Low (75-100/panicle)
Threshing	Easy	Difficult
Seed dormancy	Low	High
Shattering	Low	High
Protein content	Low	High

Source (Sarla and Mallikarjuna Swamy, 2005)

Their resemblances concern mainly their parallel biology. According to McDonald (1979), the two species are typical cereal grasses, either rainfed rice, when water is only provided by rain, or irrigated when water is supplied artificially. In the first case, these cultivars are mostly cultivated in upland fields and in the latter case, in the irrigated lowlands and uplands.

O. glaberrima and *O. sativa* have the same vegetative system composed of a fibrous root, erect culms and long flat leaves. According to Angladette (1966), roots that have already differentiated in the embryo stage are called primary roots, have limited growth and live up to

one month. They are progressively replaced by secondary roots 15 days after germination. Mishra and Rathore (1997) reported that around 46% of the roots are found in the first 150 mm of the ground with some of them reaching more than one metre in length. The development of the root system is related to genotype and farming conditions. There are about 2000 roots/hill. Each root has 10 to 30 sub-roots, each having many root hairs (Mishra and Ratore, 1997). The growth of the roots is rapid during the vegetative phase. They become inactive when grain maturity begins, even if the soil remains moist (Angladette, 1966).

The two cultivated species have also tillers which consist of a culm, leaves and panicles. The culm is constituted of nodes and hollow internodes which length-diameter ratio increase up to the upper top of the culm. The length of the stem is critical in lodging and depends both on the length and the number of internodes of the variety, or the type of rice cultivation: 1-3 metres for the *indica*, 0.5 metre for the dwarf *japonica* in a rainfed system, while the *indica* plants could easily measure 5-6 metres in the floating cultivating system.

After primary tillers emerge from nodes near the base of the main culm, secondary and tertiary tillers are produced subsequently. In rainfed rice varieties, these tillers are produced close to the principal stem during growth. The vigorous ones produce inflorescences and flowers at the same time as the principal stem. The tillering ability is, however, controlled by many factors including the sowing depth, the distance between hills, soil quality and the climate.

Leaves are developed alternately on the culm and consist of a sheath and a flat leaf blade joined by a collar, inside which a ligule and two auricles are also developed. Leaves length, width, colour and pubescence distinguish cultivars. Generally pubescent in *O. sativa* and glabrous in *O. glaberrima*, the limb is much longer in *indica* than in *japonica* and shows a length/width ratio ranging between 15 and 20.

The reproductive system includes only the panicle that emerges from the uppermost node of a culm, within a flag-leaf sheath and contains the flowers. It consists of a central rachis with up to four primary branches at each node. Primary and secondary panicle branches bear the spikelets. Each spikelet has a single floret and two glumes and is enclosed by a lemma, which can be awned or awnless, and a smaller palea. The stigma surrounds the ovary which, when fertilised, produce a caryopsis, a single-seeded dry fruit, called grain which consists of an embryo, endosperm, pericarp and testa surrounded by the husk or hull (lemma and palea). Grain length varies within cultivars between 5 and 7 mm, and grains can be round, bold or slender.

1.2 Growth and rice cycle

N'cho (1992) divided the growth of rice into three phases: (1) the vegetative phase beginning from germination to panicle initiation; (2) the reproductive phase from panicle initiation to flowering; and (3) the maturity phase lasting from flowering to the ripening. Each phase is divided into several steps. The cycle can last from three to six months, regardless of the varieties. Rice has an intermediate cycle because it is not strictly annual nor strictly perennial. Some varieties, like *Oryza barthii* A. Chev and *Oryza meridionalis* Ng. are annual (Cheng *et al*., 2002); others such as *Oryza longistaminata* A. Chev. and Röhr. and *Oryza nivara* Sharma and Shastry are perennial (Chang, 1985; Khush, 1997; Vaughan and Morishima, 2003), whereas *Oryza rufipogon* Griff. and *O. glumaepatula* Steud vary in perenniality (Cheng *et al*., 2002). The variable life cycle is determined by the vegetative phase while the reproductive and the maturity phases are quite constant.

1.3 Rice taxonomy

The history of the classification of the genus *Oryza* until 1963 was summarised by Angladette (1966). Although the nomination of the species *O. sativa* was made in the early eighteenth century, the first classification was only published at the end of the nineteenth century by Baillon. Roschevicz (1931) distinguished 19 species based on the form of the spikelet, the form and the nature of the glumella, the form and the importance of the glumes. However, Chevalier (1932) and Chatterjee (1948) modified this classification, distinguishing 22 and 23 species, respectively. Based on characteristics studied on herbarium, Tateoka (1962) recognised the existence of five complexes and species. He separated the *Sativa* and *Glaberrima* complexes, each comprising one of the two cultivated species based on biological data, while Sampath (1966), grouped the cultivated rice and their related close parents in the *Sativa* group. Moreover, unlike Roschevicz who divided the different species of the genus *Oryza* into four sections (*Sativae*, *Granulatae*, *Coarctatae* and *Rhynchoryza*); Sampath grouped them in four big series (*Sativae*, *Latifoliae*, *Angustifoliae* and *Granulatae*).

The most widely used classification of the genus *Oryza* was proposed by Second in 1985, taking in account the classifications of Chevalier (1932), Chatterjee (1948), Tateoka (1962) and Sampath (1966). It included the geographical origin of the species, the number of the chromosomes and the genomic group, which was determined by observing the chromosomes during the meiosis phase.

Table 1-1: The genus *Oryza* species*:* geographic origin, biological type (BT), reproduction system (RS), number of chromosomes and genomic groups (G)

	Geographic origin	BT	RS	2n	G
SATIVA GROUP					
Cultivated species					
O. sativa (two sub-species *indica* and *japonica*)	Asia	I	U (sometimes I)	24	AA
O. glaberrima	Africa	A	U	24	AA
Wild species					
O. rufipogon (complex species)	Asia, Australia, America	A-I-P	U-I-L+V	24	AA
O. longistaminata	Africa	P	L+V	24	AA
O. breviligulata	Africa	A	U	24	AA
LATIFOLIA GROUP					
O. officinalis (complex species)	South and South-east Asia, Southern China, New Guinea Complex species	P	U+V	24	CC+DD
O. latifolia *O. alata* *O. grandiglumis*	Central and South America Africa (+ Sri Lanka)	P	U+V	48	CCDD
O. eichingeri (complex species)	Africa	P	U	24	CC
O. punctata diploid	Africa	A	U	24	BB
O. punctata tetraploid	South-east Asia	P	U	48	BBCC
O. minuta	North Australia	P	U	48	BBCC
O. australiensis		P	U+V	24	EE
RIDLEYI COMPLEX					
O. ridleyi	South-east Asia	P	U ou I	48	--
O. longiglumis	New Guinea	P	U ou I	48	--
MEYERIANA COMPLEX					
O. meyeriana (complex species)	South-east Asia, South China	P	U	24	--
ISOLATED SPECIES					
O. brachyantha	Africa	A	U	24	FF
O. schelchteri	New Guinea	--	--	--	--
ISOLATED GENERA					
Rynchoryza subulata	Temperate South America	P	U or I	24	--
Porteresia coaretata	Indian under continent	P	U or I	48	--

Biological type : A = Annual P = Perennial I = Intermediate

Reproduction system : U = autogamous largely predominant

L = allogamous largely predominant

I = Intermediate

V = Vegetative

2n : Diploid

G : Genome symbols

Source: Second (1985)

He included the closed species *Rhynchoryza subulata* Bail (*Oryza coarctata* Nees) and *Porteresia coarctata* Tateoka (*Oryza coarctata* Roxb.) (Table 1-1).

1.4 Other uses of rice

Rice is produced as a major food staple for the world but has also been utilised in several other ways. The kernels are dissolved and used in the manufacture of glue. Different types of paper are made from rice plants. Rice starch is often incorporated into beers. Paints are made of powdered rice. The dried stems are used in feeding animals and making straw hats and sandals. The Mende people of Sierra Leone use palm oil soaked rice for ritual sacrifices to their ancestors (Richards, 1996). Felupos people of Casamance in the west corner of Senegal, use traditional African rice for ritual link between crops and ancestors (Linares, 2002).

Rice is also believed to have medicinal properties and has been used in many countries for medicinal purposes. In the Philippines, polished rice residue is used as a source of Vitamin B to prevent and cure beri-beri. Malayans used greened boiled rice to treat eye and acute inflammation of the inner body tissues while a mixture of dried, powdered rice is applied to treat certain skin ailments (Muzaffar *et al.*, 2005). In Cambodia, the husks of mature rice plants are used to treat dysentery, while the hulls of a three-month old rice plant are thought to have diuretic properties.

The Chinese believe that rice strengthens the spleen, "weak stomach", increases appetite and cures indigestion. Dried, sprouted rice grains were once used as an external medicine to aid in digestion, to give tone to muscles and to reject gas from the stomach and intestines. Rice water is prescribed by the Pharmacopoeia of India as an ointment to treat inflamed skin surfaces (Medicinal uses of rice: htpp://www.rice-trade.com/medicinal-uses-of-rice.html).

1.5 GMO issues in rice improvement

According to Nguyen (2005), rice has benefited greatly from the molecular biotechnologies. Rice was the first plant crop to be sequenced and to be fully mapped (IRGSP, 2005). The International Rice Genome Sequencing Project (IRGSP), a consortium of 10 developed countries funded research institution, identified 37,544 genes, many of which are duplicated. Of these genes, 34,685 are found in *Arabidopsis thaliana* L., a genetic model plant which was completely sequenced in 2000. A total of 2,859 rice genes are different from genes

found in *A. thaliana*, and may discriminate monocots and dicots (International Rice Genome Sequencing Project, 2005).

Haploid production and embryo isolation and culture have been researched and exploited for rice production.

Despite the strong opposition of many NGOs, genetically modified organisms (GMO) technology has been applied to rice as well as many other crops including brassicas, sunflowers, potatoes and soybeans (James, 2007). Genetic engineering has produced a genetically modified line of rice that contains beta-carotene, a component that is transformed into Vitamin A in the human body (Al-Babili and Peter, 2005). Scientists have also created transgenic rice with a high iron content, aiming at improving the health of billions of poor people in the Asia-Pacific region by the combination of the proVitamin A and iron in rice (ETH Zurich, 2009). Rice lines from genetic engineering, resistant to tungro virus and bacterial blight, have been developed at the International Rice Research Institute (IRRI). However, the adoption of these GMOs by farmers failed due to the burden of biosafety regulations and the high cost in developing and transitional countries (James, 2007). In Europe, most of the governments of the European Union have put a blanket ban on GMOs, allegedly for safety concerns. Only two GMO crops have been approved to be grown by EU farmers, (James, 2009). Despites this strong opposition, 25 countries are currently growing GMO crops. These countries are, in decrease order of area (million ha): USA (62.5), Argentina (21.0), Brazil (15.8), India (7.6), Canada (7.6), China (3.8), Paraguay (2.7), South Africa (1.8), Uruguay, Bolivia, Philippines, Australia, Mexico, Spain, Chile, Colombia, Honduras, Burkina Faso, Czech Republic, Romania, Portugal, Germany, Poland, Slovakia and Egypt (James 2008). According to James (2010), the non-adoption of GMOs is mostly due to incorrect understanding that people have of GMOs.

Another challenge is to create C_4 rice. Scientists expect to increase rice photosynthesis capacity by transforming normal C_3 rice into C_4 rice, which may lead to yield increases of 15-20% (Hiroshi *et al*., 2001). Marker Assisted Selection using genetic fingerprint techniques to identify genetic variations of pest and disease resistance, and protoplast transformation to produce transgenic organisms for the introgression of new genes have also been investigated (Akio *et al*., 1990; Sharma, 2003; Pan *et al.*, 2009).

Development of GMO constructs involves several steps including: (1) genetic construction that consists of identifying, isolating and integrating the genes of interest in a

genetic construct; (2) development of marker genes and cloning that consist of multiplying a plasmid (circular DNA molecules of unique chromosomes present in bacteria such as *Escherichia coli*) or genetic construction in a host bacterium (usually *Agrobacterium tumefaciens).* The soil bacterium *A. tumefaciens* causes a tumour called neck gall in the infected plant, which is followed by the natural or biologic transfer of its plasmid T-DNA (for transferred DNA) part into the plant chromosomes. The usual practice is to modify its Ti plasmid to avoid gall formation, but to keep the transfer, integration and virulence functions of the desired gene in the plant genome by constructing disarmed Ti plasmids (delete the genes situated on the T-DNA, responsible for the pathogenicity of the bacterium).

1.6 Rice production constraints in Africa

Despite rice becoming an important cereal and staple food crop in Madagascar, West, East, Central and Southern Africa (Balasubramanian *et al*., 2007), rice production in Africa is still the lowest in the world and cannot meet the increasing demand for rice in many African countries (Hossain, 2006). Unless this situation is reversed, food dependency will increase in much of the continent. This situation is due to both abiotic and biotic constraints on rice production, and many technical, management, socioeconomic, health, and policy constraints that restrict rice production (Balasubramanian *et al*., 2007).

1.6.1 Biotic and abiotic constraints in Africa

In Africa, rice is affected by a wide range of abiotic and biotic factors that harm productivity. The abiotic constraints include: drought, flooding and variable rainfall, extreme temperatures, salinity, acidity/alkalinity and poor soils, soil erosion and high phosphorus fixation. The biotic constraints include weeds, blast leaf spots, at least 15 viruses (Ou, 1985; Hibino, 1989, Ndjiondjop, 1999), and African rice gall midge (AfRGM). According to Balasubramanian (2007), among biotic factors, weeds are the most serious, followed by blast and brown spot diseases. Estimated yield losses due to weeds range from 30 to 100% (Akintayo *et al*., 2008). Weed infestation and loss of nitrogen reduce yields by 25% on intensive dryland rice farms of West Africa (Becker and Johnson, 2001). Nematodes are a serious problem in continuously monocropped dryland rice fields (Plowright and Hunt, 1994; Coyne *et al*., 2004) and can reduce yields by up to 30%. Termites are critical in some areas. Rodents and birds

damage rice crops in all ecosystems. Stem borers, leaf hoppers and rice midges are the major insect pests.

Another rice production constraint is global warming. Methane (CH_4) emission is becoming a major concern with flooded rice fields because rice plants act as gas vents, releasing large amounts of CH_4 (about 115 million tons) annually into the atmosphere, increasing CH_4 concentrations each year by 1% (Greenhouse Gas Online, 2008). It is projected that by the end of the twenty first century, a 2-8^oC rise in global temperatures will be caused by greenhouse gases (Crutzen *et al.*, 2008). This greenhouse gas, along with CO_2, contributes to the trapping of more heat in the atmosphere. Studies by Lin (2005) revealed that CH_4 traps 20 times more energy than CO_2. To reduce CH_4 emission, water management actions such as moving crops from flooded conditions to upland sites, and soil aeration, are being researched.

1.6.2 African rice gall midge (AfRGM)

AfRGM was first reported in Sudan and is now spreading throughout Africa. Souleymane (1998) reported that the AfRGM had been found in 12 West African, two Central African, and five East and Southern African countries. In West Africa, this pest is widely spread in the Guinea Savannah, the Wet Forest Savannah, and the Sudan Savannah (Harris, 1960, 1980; Alan *et al.*, 1985; Ukwungwu *et al*., 1989). Unlike the Asian rice gall midge (AsRGM), for which more than ten biotypes have been identified so far (Behura *et al.*, 1999; Katiyar and Bennett, 2001), the number of biotypes of AfRGM is unknown. However, the variability of AfRGM reaction to the host's resistance genes and the peculiarity of varietal resistance with respect to the site, show that there is more than one biotype of the AfRGM. Fernando (1972) reported that the existence of different gall midge biotypes is shown by the differential reactions of a given set of different crop varieties submitted to pest populations in many areas at a narrow or high scale. Based on this differential assay, Nwilene *et al*. (2006a) identified *Orseolia nwanzei* Harris and Nwilene, another species distinct from *Orseolia oryzivora* Harris and Gagné. Moreover, the susceptibility of *O. glaberrima* to AsRGM and high resistance to AfRGM, and vice versa, shows that the strains of AsRGM could be different from the strains of AfRGM. In addition, a 69 base pairs transposon was identified in the AfRGM, which is not detected in the AsRGM, signifying a possible evolution of the latter from the former.

1.6.2.1 Economic importance of AfRGM

The AfRGM, one of the major dipteran pests affecting rice (Lai *et al*. 1984; Ukwungwu *et al*. 1989; Ukwungwu and Joshi 1992; Dale 1994; Taylor *et al*. 1995), causes damage to irrigated and rainfed lowland rice crops (Peter *et al.*, 1991). It rarely affects rainfed rice crops (Bonzi, 1980). A recent communication suggests that AfRGM also occurs in upland agrosystem (F.E. Nwilene, personal communication). The most affected countries are Burkina Faso, Mali and Nigeria, with farmers' field infestation being above 40% (Ukwungwu and Joshi, 1992; Dakouo *et al.*, 1988). A total crop loss has been observed at severe outbreaks in Nigeria (Ukwungwu *et al*., 1989; Ogah *et al*., 2009). Insectary cage studies suggest that even moderate infestation can cause high yield reductions (Nacro *et al*., 1996). There is very little information on the relationship between the AfRGM infestation level and crop yield under field conditions. However, studies by Nacro *et al.* (1996) show that 1% of infested tillers caused a 2% yield loss. Similar studies in Nigeria (Williams *et al.*, 1999a; 2002) show that an increase of 1% in infestation resulted in a 2.9% yield loss. In certain regions, severe attacks lead to total loss of the harvest (Ethel *et al.*, 1993).

1.6.2.2 Biology of the AfRGM

AfRGM, like AsRGM, is a species of the genus *Orseolia,* and is a small mosquito-like insect (Peter *et al.*, 1991). The African *Orseolia* includes three species: *O. oryzivora* Harris and Gagné, *O. bonzii* Harris and *O. nwanzei* Harris and Nwilene. There are a few morphological differences between the adult and immature stages, and differences are especially evident in pupal characteristics. *O. oryzivora* and *O. nwanzei* are directly harmful to rice, while *O. bonzii* is the Paspalum gall midge (PGM) and is an alternate host for *Platygaster diplosisae* (Hymenoptera: Platygastridae) and *Aprostocetus procerae* (Hymenoptera: Eulophidae), the two main parasitoids of AfRGM. However, genomic DNA fingerprinting of these three insect species, using random amplified polymorphic DNA (RAPD) and sequence-characterised amplified region (SCAR) methods have revealed that *O. oryzivora* and *O. bonzii* are closely related and *O. nwanzei* is genetically more distant (Nwilene *et al*., 2006a).

1.6.2.3 AfRGM cycle of reproduction

The mode of reproduction is similar to that of its counterpart, the Asian *Orseolia*. The male mates with 2-4 females which mate only once (Bennett *et al.*, 2004). The female adult,

with a robust, reddish-brown abdomen, lays eggs (100 to 400/day) on rice leaf blades and ligules. After hatching, the larvae move to the base of the stem through the sheath. The larvae then bore into the stem by feeding on it, leaving a cavity that turns the leaf blade into swollen onion like leaf galls. The rice plant is therefore stunted and the infested tillers do not produce panicles. At the end of its development, the larvae turn into pupae from which emerge adult midges, with a usual sex ratio of one male to three females (Umeh *et al.*, 1992), which can become 1:6 or 1:0.5 in cold area (Bennett *et al*., 2004). Adult gall midges are mostly nocturnal and are attracted by artificial light. The life cycle of the insect from egg to adulthood varies from three to five weeks (Bouchard *et al.*, 1992, Nwilene *et al*., 2006b). The male has a slender, brown abdomen and longer antennae than the female (Figure 1-1). The life span of the adult male is about three days.

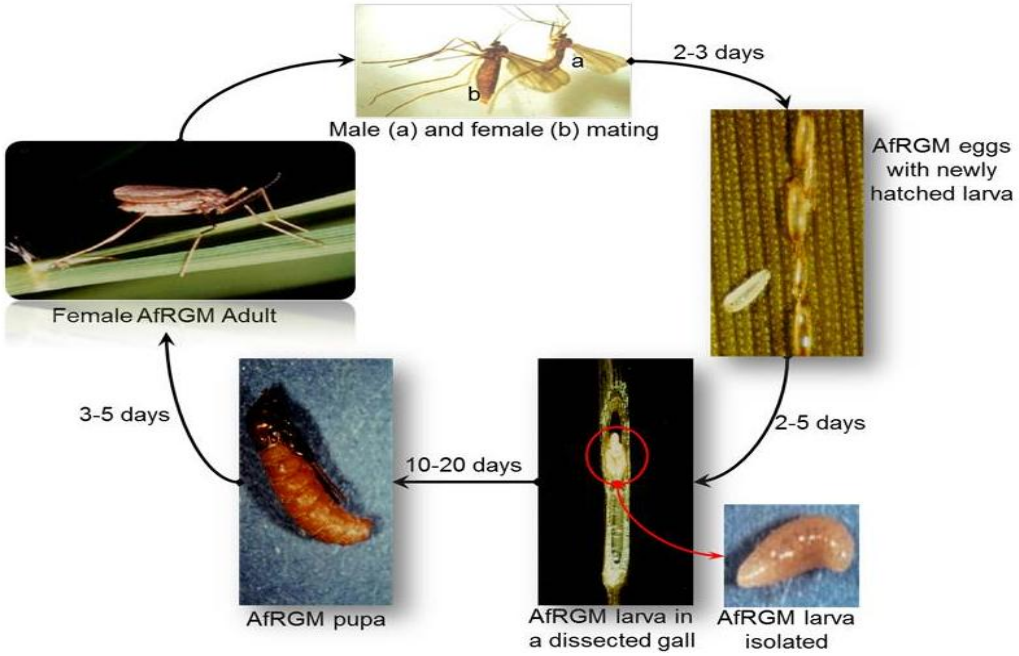


Figure 1-1: African rice gall midge cycle of reproduction (adapted from Nwilene *et al*., 2006b)

1.6.2.4 AfRGM symptoms

Crop damage is caused by the larvae of AfRGM (Ba *et al*., 2004), which infests rice tillers at the vegetative growth stage, inducing the transformation of leaf sheaths into a tubular "silver shoot" galls (Figure 1: 2-a, 1-b, 1-c and 1-d). A galled tiller is unable to initiate growth of more leaves or a panicle. Early gall infestation (Figure 1-2-a) results in a stunted rice plant with bushy appearance having up to 50 small tillers. These extra tillers, produced in response to

attack by AfRGM, often become infested as well. Gall formation is a hypersensitive response of defense for the rice plant (Omoloye, 2010) to confine the insect in space. Galls cannot be pulled out of the rice tillers, unlike dead hearts caused by stem borers. Later infestation, however, leads to oval shapes with ragged edge holes from which adults emerge, leaving behind empty, transparent pupal skins protruding from the exit hole (Nwilene *et al.*, 2006b).

Gall formation is due to secretions from the insect's salivary gland called cecidogen (Chen *et al.*, 2008), which is a combination of many chemicals including plant regulators (IAA and tryptophan) that act as a gall inducer, polyphenol oxidase (PPO) and phenolic compounds (Ananthakrishnan, 1998). Small galls from floral structures called panicle galls are produced when plants are attacked in the reproductive phase (Rajamani *et al.*, 1979; Kittur and Agrawal, 1983). However, the vegetative phase is extended when the rice plant is repeatedly infested. The gall tissue acts as a physiological sink that feed the growing insect.

Figure 1-2-a: Galls symptom in a nursery bed

Figure 1-2-b: Galls symptom 14 days after transplanting

Figure 1-2-c: Gall symptom 45 days after transplanting

Figure 1-2-d: Galls symptom 70 days after transplanting

Figure 1-2: Gall midge symptoms at different rice growth stages

1.6.2.5 AfRGM epidemiology and conditions favouring transmission

Since Harris and Gagné (1982) described and illustrated AfRGM as a species different from the Asian type, it has been the subject of many studies, particularly in West Africa, the insect's zone of activity (Omoloye *et al*., 2002; Ba *et al*., 2004; Nwilene *et al*., 2006a,b; Nacro and Nénon, 2009; Ogah *et al*., 2009; Maji *et al*., 2010; Omoloye, 2010). The insect causes damage during the rainy season. Infestations are low at the beginning of the rice cycle, then reach a peak at panicle initiation or maximum tillering and decrease afterwards (Heinrichs *et al.*, 1995; Dakouo *et al.*, 1988). The gall midge needs a certain level of humidity to develop. So, during the dry season, the insect remains dormant in the pupal stage, causing low infestations. It becomes active again when the buds start growing after the rains. In the absence of paddy fields of *O. sativa* or *O. glaberrima*, wild rice plants and new growth serve as relay hosts. Among these wild plants are *O. longistaminata* and *Paspalum scrobiculatum* L. (Bonzi, 1980), *O. barthii, Oryza punctata* Kotschy ex Steud. and *Oryza stapfii* Roschev. (Nwilene *et al.*, 2006b). A wide range of factors contribute to increasing the AfRGM population density and transmission. These include cloudy or rainy weather, cultivation of high-tillering varieties, intensive management practices and low parasitism of the midges (Bennett *et al*., 2004).

Nwilene *et al*. (2006b) reported that conditions favouring AfRGM transmission are due both to abiotic constraints and cultivation habits. Abiotic constraints include: (1) Changes in weather with high rainfall, excessive cloud cover and humidity favouring outbreaks, and (2) Higher risk of occurrence in rainfed lowland and hydromorphic agrosystems than upland and mangrove agrosystem. Certain cultural practices increase the risk of AfRGM. Some of them are: (1) Use of a wide range of planting dates leading to higher infestation of late-planted fields; (2) Planting of new high-yielding AfRGM-susceptible varieties, generally more susceptible than the traditional landraces they are replacing; (3) Increased use of fertiliser, which increase AfRGM infestation.

Site and field characteristics are also associated with higher risks of AfRGM infestation and concern planting methods (direct-seeded rice less infected than transplanted rice), wet-season weather patterns (cloudy, humid weather with frequent rain or mist favours AfRGM) and presence of alternate hosts (*O. longistaminata* and volunteer rice) that build-up midge population prior rice planting.

1.6.2.6 Strategies to control AfRGM

1.6.2.6.1 Biological control

Natural pests linked with AfRGM have been identified. The major ones are parasitoids such as *Platygaster diplosisae* Risbec (Platygastridae) and *Aprostocetus procerae* Risbec (Eulophidae) (Williams *et al.*, 1999a; Nacro and Nénon, 2009), predators like *Cyrtorhinus viridis* Linnavuori (Miridae), *Conocephalus longipennis* de Haan (Tettigoniidae), *Anaxipha longipennis* Serville (Gryllidae) and ladybird beetles (Coccinellidae) (Nwilene *et al.*, 2006b). Biological control studies show that these pests can develop a maximum parasitism of 72% in farmers' fields (Umeh and Joshi., 1993). Unfortunately, this parasitism occurs on the larval stage in the stem and cannot therefore prevent yield losses. Moreover, the parasitoid populations build up too late to prevent heavy AfRGM infestation (Bennett *et al*., 2004). Natural biological control has also been attempted and consisted of: (1) Manipulating habitat to increase the carry-over of the two main parasitoids, from the *Paspalum* gall midge (PGM), *O. bonzii*, a related gall midge to AfRGM; (2) Cultivating in the dry-season for *P. scrobiculatum* abundance early in the wet season, to encourage the prevalence of the predators of midges (Ba *et al*., 2004).

1.6.2.6.2 Chemical control

The use of insecticides to control AfRGM is not efficient because of cost, risk to human health and the destruction of natural AfRGM enemies (in the case of foliar sprays) (Nacro and Nénon, 2009). Foliar sprays have not been efficient because the larvae are not accessible because they live within the plant (Nwilene *et al.*, 2006b). However, granular systemic insecticides, such as the carbofuran (Furadan), are used at outbreak levels with: (1) Good water control that prevents water from running off the field, carrying away some of the insecticide; (2) An infestation level of at least 0.3 galls per hill, 5 to 10% of hills with galls or 1 gall per square metre, for the application of the insecticide to be economically viable. The insecticide can also be applied to seedlings before transplanting using granules, or spraying on the nursery bed, or soaking the roots in the container of insecticide solution for several hours prior to transplanting (Bennett *et al*., 2004). With this approach, advantages such as the need for less insecticide to protect a given area of crop, early protection of plants and avoidance of killing

natural enemies to AfRGM, are a trade-off with a risk of heavy AfRGM damage occurring late in the crop cycle. However, this risk is not predictable (Williams *et al*., 2002).

1.6.2.6.3 Cultural practices

Cultural practices that involve changing the way the crop is grown in order to reduce pest damage (Bennett *et al.*, 2004) have been sought for AfRGM control and include: (1) Favouring early and synchronised planting which suffer less damage than those planted late; (2) Destruction of alternative hosts (rice ratoons, volunteer rice and *Oryza longistaminata*; (3) Use of moderate levels of NPK fertiliser (e.g. 60 kg ha^{-1} for field derived from savannah) instead of high doses of nitrogen fertiliser that increase AfRGM infestation; (4) Discouraging movement of seedlings (transplanting) to prevent seedling infestation by AfRGM in the nursery and later infestation spread; (5) Avoidance of close plant spacing to avoid a micro-environment favouring AfRGM (Nwilene *et al.*, 2006b)

1.6.2.6.4 Varietal resistance

A range of studies on the control of AfRGM have used biological, cultural (Ukwungwu, 1987; Nacro *et al.*, 1995; Nacro et Nénon, 2009) and chemical (Ukwungwu, 1990; Lafleur, 1994**)** control strategies. These studies show that host plant resistance is the single most effective means of controlling this insect (Heinrichs, 1994; Kaityar *et al*., 1995, Nwilene *et al*., 2006) because chemical control of AfRGM has not been successful and biological control is still in its infancy (Singh *et al*., 1997), while cultural practices remain difficult to implement.

1.6.2.7 Sources of resistance to AfRGM

Many of rice varieties currently available to farmers are highly susceptible to AfRGM (Maji *et al.*, 2010). Improving varietal resistance appears to be one of the most promising options for managing the pest especially since, in Asia, resistant varieties have been used with considerable success against the closely related AsRGM *Orseolia oryzae* Wood-Mason (Heinrichs, 1994). Therefore, since the early-1980s, rice lines have been screened for resistance to AfRGM in Nigeria by the National Cereals Research Institute (NCRI), in collaboration with the Africa Rice Centre (AfricaRice formerly known as WARDA), IRRI and the International Institute of Tropical Agriculture (IITA). Despite intensive screening, no *O. sativa* lines have been found with very strong resistance under high AfRGM pressure (Ogah *et al.*, 2009).

However, a number of lines with moderate AfRGM resistance have been identified with superior resistance to currently released varieties (Williams *et al*., 1998, Nwilene *et al*., 2006b; Ogah *et al*., 2009). Most of these lines are traditional varieties which are unsuitable for direct release to farmers because of their low yields.

Cisadane, an improved *O. sativa* subsp. *indica* variety from Indonesia, first identified as moderately resistant to AfRGM (Ukwungwu and Alam, 1991), was the highest yielding test entry in yield trials of the more promising lines carried out in Nigeria, in 1991 (Singh, 1992). Consequently it was considered as potentially useful for rice farmers in the country's AfRGM prone areas without further varietal improvement. Cisadane is, however, very susceptible to iron toxicity. The varieties that is both resistant to AfRGM and tolerant to iron toxicity are BW348-1 (Singh, 1998) and TOS14519 (Nwilene *et al.*, 2002). However, these varieties are less productive.

1.6.2.8 Mode of inheritance

Very little is known about genetic control of resistance to AfRGM. The genetic basis of resistance to AfRGM has not been established. Pedigree selection has, however, been used to develop resistant varieties to AsRGM (Maji *et al*, 1998). Resistant genes to AsRGM have been identified in some varieties. Resistance to AsRGM is under the control of at least ten resistance genes, eight of which have been tagged and mapped (Kumar *et al.*, 2005; Himabindu *et al*., 2007). Flanking markers have been used to identify the resistance genes *Gm1* and *Gm2* in various rice cultivars (Himabindu *et al*., 2007). Some of these genes were found on Chromosome 4 (Mohan *et al.*, 1994) and Chromosome 9 (Biradar *et al*., 2004). Some studies report resistance genes to AsRGM being related to the resistance genes against the brown leafhopper and the brown stripes of paddy rice (Prasad *et al.*, 1975).

1.6.2.9 Mechanisms of resistance

The mechanism of resistance against AfRGM has not been specifically investigated yet, unlike the resistance to its related AsRGM. Breeding for resistance has been limited. One recent project is to introgress the resistance from TOS14519 into promising agronomic parent varieties (AfricaRice Center (WARDA), 2005). *O. glaberrima* varieties are included in this project. These efforts to breed for resistance to AfRGM could enable the AfRGM mechanisms of resistance to be studied. However, close similarity with its related Asian species and the

possible evolution of the AsRGM from the AfRGM suggest that the mechanism of resistance against both midges could be similar. Three mechanisms of resistance have been proposed in the *O. sativa*/AsRGM system.

1.6.2.9.1 Antixenotic resistance

Antixenosis is the resistance mechanism used by the plants to deter or reduce colonisation of insects which orient themselves towards the plant for food, mating, ovipositioning and shelter (Korgan and Ortman, 1978). Antixenotic characteristics are either biophysical, biochemical or a combination of the two, and reduce the initial number of insects in subsequent generations. It utilises morphological, physical and structural qualities that include pubescence, tissue hardness that limits insect mobility (Alvarez *et al.*, 2006), plant stimuli (Liu *et al*., 2007;), repellents and anti-feedents (Wu *et al*., 2008). Initiation and completion of key activities of the insects is believed to be influenced by inappropriate stimuli (Subramanyam *et al.*, 2008) such as production of the ovipositioning inhibitor, pentadecanal, in the rice cultivar TKM6 against *Chilo suppressalis* Walker (Pyralidae) (Saxena, 1986) and biophysical factors such as plant pubescences, as observed in the comparison of *Capsicum annuum* L. (Solanaceae) and *C. pubescens* for antixenosis as a means of aphid resistance (Bosland and Ellington, 1996). Frego bracts in cotton also helped in the reduction of the number of eggs laid and subsequent damage by boll weevils, *Anthonomus grandis* Boheman (Curculionidae) (Jenkins and Parrot, 1971). There are also some visual factors such as colour (Harris and Miller, 1982).

1.6.2.9.2 Antibiotic resistance or antibiosis

The primary component of varietal resistance against AsRGM is antibiosis (Kalode, 1980). Antibiosis is the mechanism brought into operation after the insect has colonised the plant (Rawat *et al*., 2010). Insect growth, development, reproduction and survival become affected as a result of increased nymphal mortality, reduced nymphal growth extension, and reduced adult longevity and fecundity (Liu *et al.*, 2010). These conditions also reduce resistance against its insect enemies. Factors that provide an antibiotic defense include toxins which act as insecticides (Zarember *et al*. 2009). Other factors include reduction in sugar content needed by the feeding insect at critical growth stages (Giovanini *et al.*, 2007). An association has been identified between phenolics and the resistance (Dubui and Haass, 2007; Staudt *et al*., 2010). Phenols play a significant role in the expression of the resistance conferred

by the *Gm2* gene (Amudhan *et al*. 1999). Furthermore, a wide range of allelochemical compounds, while playing an important defensive role against insects, have not been associated with any resistance to pest infestation (Vidyachandra *et al*. 1981, Rajamani, 1982, Joshi and Venugopal, 1984). The same conclusion was made in a study on total polyphenol, total nitrogen, soluble nitrogen, and soluble sugar content of 29 AsRGM resistant and susceptible rice varieties (Sain and Kalode, 1994).

1.6.2.9.3 Tolerance

Tolerance is the genetic trait of a plant that protects it against a parasite population which would have damaged a susceptible variety; to preserve an acceptable yield or the quality of the plant marketable product (Suwa and Maherali, 2008). Tolerance cannot be accurately evaluated at the vegetative or flowering stage, unlike antibiosis and antibiotic resistance (Bennett *et al*., 2004) Some traits, including the number of productive tillers (tillers having panicles with grain), extra tillers-short plants and few panicles, panicle m^{-2} and uninfested grain yield and number of uninfested tillers at 63 days after transplanting (DAT) and at harvest can be used to predict tolerance (Omoloye, 2002). Tolerance mechanisms are different from antixenosis and antibiosis in that tolerance does not lower the insect population or exerts evolutionary pressure on insect pests to develop new strains, but it could raise the threshold levels (Soroush *et al*., 2011).

Tolerance manifests itself as: (1) A compensatory reaction (e.g. rice cultivar Cisadane, against AfRGM, as reported by Omoloye and Odebiyi (2006)); (2) In a sink-source relationship (reduction of leaf area due to leaf mining by *Liriomyza sativae* Blanchard (Agromyzidae) without affecting the tomato yield (Johnson *et al*., 1983)); (3) By photosynthesis enhancement (partial defoliation of tobacco when attacked by the tobacco budworm (Kolodny-Hirsch *et al*., 1986)); (4) In resource allocation (early attack by the sorghum shoot-fly, *Atherigona varia scotia* Rondani (Muscidae), on the main shoots of sorghum induces the production of synchronous tillers growing rapidly producing harvestable heads (Rana *et al*. 1985)).

1.7 Use of molecular markers in breeding for AfRGM resistance

Various molecular techniques have been applied in breeding for gall midge resistance. Only the most commonly used techniques are reviewed.

1.7.1 Restriction fragment length polymorphism (RFLP)

RFLP is the most widely used hybridisation-based molecular marker because it provides less complex profiles, it facilitates genetic fingerprinting and genetic mapping, it is reliable and the results obtained are reproducible. However, the technique is difficult to perform and the transfer and hybridisation steps block automation of the technique (Kumar *et al*., 2009). RFLP markers were first used in 1975 to identify DNA sequence polymorphisms in adenovirus serotypes (Grodzicker *et al*., 1975), then used for human genome mapping (Botstein *et al*., 1980), and adopted a decade later for plant genomes (Helentjaris *et al*. 1986; Weber and Helentjaris, 1989).

RFLP analysis has been used to map *Gm2*, a dominant gene conferring resistance to Biotype 1 of the AsRGM (Nair *et al.*, 1995). The gene is found on Chromosome 4 at 1.3 cM from Marker RG329 and 3.4 cM from Marker RG476. It has been used for DNA fingerprinting of individual AsRGM insects, using a particular repetitive DNA sequence as a probe (Ehtesham *et al*. 1995). This is despite the method being slow, costly, and generally limited to a single restriction enzyme digest and a single DNA blot for each individual.

RFLP uses a gel fingerprint of DNA, treated with restriction enzymes to reveal a pattern that differentiates between DNA fragment sizes in individual organisms. Each DNA sequence modification (mutation, addition, deletion, translocation, inversion or duplication) leads to a re-organisation of the restriction sites. Hence, digestion of DNA with restriction enzymes results in the modification of the number or the size of fragments and leads to a polymorphism. In maize, 95% of the probes used are polymorphic, while in wheat, an autogamous plant, only 5 to 10% of the probes are polymorphic. Perez de la Vega (1993), Terachi (1993) and Landry (1994) described the procedure of the RFLP technique is as follows: (a) Extraction of the DNA; (b) Digestion of the DNA with one or more restriction enzyme(s); (c) Separation of the restriction fragments in an agarose gel; (d) Transfer of separated fragments from the agarose gel to a filter by Southern blotting; (e) Detection of individual fragments by nucleic acid hybridisation with a labelled probe(s); and (f) Visualisation by autoradiography.

1.7.2 Random amplified polymorphic DNA (RAPD)

RAPD was, with AP-PCR (arbitrarily primed PCR) and DAF (DNA amplification fingerprinting), collectively termed multiple arbitrary amplicon profiling (MAAP) (Caetano-Annolles, 1994) and also the first techniques used to amplify DNA fragments from any species

without prior sequence information (Caetano-Anolles *et al*., 1992; Semagn *et al*., 2006). RAPD analysis has been performed on gall midge for DNA fingerprinting (Behura *et al*., 1999) even though the small amount of total genomic DNA, extracted from a single insect, and the small number of informative bands, obtained with each amplification, pose severe limitations. However, because it is simple, fast and provides more polymorphic loci than RFLP, this technique has been widely used for genomic analysis (Jena *et al*., 1998: Zheng *et al*., 2008; Lin *et al*., 2009), gene mapping (Nair *et al*., 1996; Milton *et al*., 2007), and biotype characterisations (Behura *et al*., 1999; Bennett *et al*., 2004) of the rice gall midge.

The RAPD technique consists of performing a PCR using a short primer of about ten nucleotides and of arbitrary sequence that match randomly in the genome. If two sites of matching are close together and are situated on two DNA strands, there is amplification. Conversely, if these sites are too far apart, there is no amplification. This type of polymorphism is called site primers matching polymorphism. The primers then constitute the markers. For the entire genome, an average of ten fragments is amplified. RAPD does not need digestion by a restriction enzyme, or a transfer on membrane, or even a radioactive probe preparation Arif *et al*., 2010). Therefore, it is fast and needs little technical expertise, compared to RFLP. However, it is not reproducible because the amplification is very sensitive to the polymerase chain reaction (PCR) process conditions (quality and quantity of DNA template, PCR buffer, concentration of magnesium chloride, primer to template ratio, annealing temperature, Taq DNA polymerase brand or source, and thermal cycler brand (Wolff *et al*., 1993)), which differ from one lab to another. It is usually used for fast preliminary analyses, and may reveal markers near a gene of interest.

1.7.3 Amplification fragment length polymorphism (AFLP)

In contrast to RFLP and RAPD, amplified fragment length polymorphism (AFLP) analysis (Vos *et al.*, 1995) generates unlimited numbers of DNA fragments from less than 1µg of genomic DNA and is the friendliest technique dedicated to DNA fingerprinting. AFLP analysis has been used to assess the biodiversity of the Asian and the African rice midges by Kathiar *et al*. (1999) and Nwilene *et al*. (2006a), respectively. It was also used to identify a DNA marker linked to the Hessian fly resistance gene (Bennett *et al*., 2004). AR257, an AFLP marker, was also associated with *Gm8*, a gene conferring resistance to AsRGM (Jain *et al*., 2004). In addition, it has been used in several genome studies, including bacteria (Janssen *et*

al., 1996; Keim *et al*., 1997), yeasts (Majer *et al*., 1996), plants (Thomas *et al*., 1995; Gerats *et al*., 1995; Cho *et al*., 1996), and animals (Folkertsma *et al*., 1996; Joerg *et al*., 1996).

The AFLP technique combines the power of RFLP with the flexibility of PCR-based technology by ligating primer recognition sequences (adaptors) to the restricted DNA (Lynch and Walsh, 1998) and is based on each restriction site polymorphism with a matching polymorphism of an arbitrary sequence primer. Romero *et al*. (2009) described the AFLP technique as follow. DNA is digested by restriction enzymes and produces varying fragment lengths, depending on the enzymes used. Then, specific nucleotide adaptors (of known sequences) of the restriction enzymes are added at the borders of the restriction fragments. The fragments are then amplified by PCR, using as primer, an oligonucleotide that is complementary to the adaptor sequence, prolonged by few arbitrary nucleotides (1 to 3) called border bases. Only these fragments, with complementary bases of the arbitrary bases, are amplified. The primers are then selective primers that reduce the number of amplified fragments to hundreds, eliminating thousands of fragments that do not have these border sequences.

The limitations of AFLP include: (a) An increased number of steps to produce the result; (b) The DNA template must be free of inhibitor compounds that interfere with the restriction enzyme; (c) The use of polyacrylamide gel in combination with $AgNO_3$ staining, radioactivity, or fluorescent methods of detection, which is more expensive and laborious than agarose gels; (d) Additional cost to purchase both restriction and ligation enzymes as well as adapters; (e) As RAPD, most AFLP loci are dominant, which does not differentiate dominant homozygotes from heterozygotes. This reduces the accuracy of AFLP markers in population genetic analysis, genetic mapping, and MAS (Semagn *et al*., 2006).

1.7.4 Microsatellite markers: simple sequence repeat (SSR)

Microsatellites (Litt and Luty, 1989), are known as simple sequence repeats (SSRs) (Tautz *et al*., 1986), short tandem repeats (STRs) or simple sequence length polymorphisms (SSLPs) (McDonald and Potts, 1997). They are the smallest class of simple repetitive DNA sequences and are defined as 2-8 bp repeats (Armour *et al*., 1999), 1-6 bp repeats (Goldstein and Pollock, 1997, Arif *et al*., 2010) or even 1-5 bp repeats (Schlotterer, 1998). Microsatellites are common sequences of 1 to 4 nucleotides repeated units. The most common are (A)n, (TC)n, (TAT)n and (GATA)n. N values vary from few units to several tens (Arif *et al*., 2010). The

nucleotides are repeated in tandem or SSRs. SSR markers have been intensively used during the last decade in different ways in breeding for gall midge resistance.

The SSR markers RM316, RM444 and RM219, located on Chromosome 9, were linked to *Gml* at genetic distances of 8.0, 4.9 and 5.9 cM, respectively, and flanked the gene locus (Biradar *et al*., 2004) while the markers RM28574 and RM28706, on Chromosome 12, tag the new rice gall midge resistance gene *Gm11t* within 4.4 and 3.8 cM, respectively (Himabindu *et al*., 2010). Himabindu *et al*. (2007) also reported a flanking SSR marker for an allelism test for the AsRGM resistance genes. The recessive AsRGM gene *gm3* was also fine mapped onto Chromosome 4L and flanked by two SSR markers, RM17473 and RM17480, at a genetic distance of 3.2 cM and 2.0 cM, respectively (Sama *et al*., 2010). Genetic diversity amongst rice genotypes resistant and susceptible to the AsRGM SSR markers was reported by Himabindu *et al*. (2009).

The value of microsatellites lies in their polymorphism due to the variation in the number of repeated sequences constituting the microsatellite (Noda *et al*., 2009). Microsatellite polymorphism is revealed by PCR. One pair of specific primers of the left and right borders of one microsatellite amplifies the same microsatellite in different individuals. In fact, each microsatellite is bordered by unique sequences that are specific to it. Amplification fragments are revealed by electrophoresis. One individual, having more repeated sequences than another, shows amplification products that migrate more slowly. The SSR technique needs a heavy pre-preparation to know, synthesise, and test the primers bordering the microsatellite. In contrast, it is easy to perform because it uses only the PCR technique. It helps to develop many markers, particularly for maize, cabbage and rice. However, this technique is not applicable to all species; tomato, for instance, does not show high levels of polymorphisms with microsatellites (Bredemeijer *et al*., 1998).

1.7.5 Single nucleotide polymorphism (SNP)

Recent development in DNA sequence analysis and the development of high speed methodologies have made possible the identification and the analysis of nucleotidic variations on a large scale (Agarwal *et al*., 2008; Sarwat *et al*., 2008; Heubl, 2010). Single nucleotide polymorphism (SNP) is defined as any polymorphism between two genomes that is based on a single nucleotide change (Arif *et al*., 2010). The use of SNP (or Snip) has never been reported

for any gall midge study. However, it can provide fine results to differentiate alleles between individuals.

In plant breeding, it has been used for line selection and for the saturation of a given region near the gene, for its cloning (for positioning in this region a high number of markers) (Foster *et al*., 2010). SNPs are excellent markers for association mapping of genes controlling complex traits, and provide the highest level of map resolution (Brookes, 1999; Bhattramakki *et al*., 2002; Botstein and Risch, 2003). SNPs are a robust technique and polymorphisms are identifiable and detectable with several methods. SNPs are the most frequent type of variation found in DNA (Brookes, 1999; Cho *et al*., 1999); and their discovery together with insertions/deletions, is the basis of most differences between alleles (Gomez and Maloo, 2009).

The SNP technique involves the annealing onto the target DNA of a complementary probe carrying a fluorescent molecule (fluorophore). Each probe is specific to a given DNA sequence. The oligonucleotides are thereafter added at the extremity of the primers by the Taq polymerase and then the fluorophores, fixed on the probes, are released. The nucleotidic polymorphism results in different levels of fluorescence which is observed by visualisation, stimulation and quantification of the fluorophore, at its wave length (Arif *et al*., 2010).

SNP technology avoids electrophoresis and therefore presents a technology adaptable to automation. It is more powerful than RFLP, RAPD, AFLP and SSR technologies. It can then be performed at high speeds, on a large number of samples (Arif *et al*., 2010). However, SNP is costly, needs very good DNA quality, the parental chips should also be available and has a bi-allelic nature (Engle *et al*., 2006).

1.7.6 Properties of good molecular markers

A good marker must be: (1) Neutral: its different alleles should not have any effect on the phenotype of the individual; (2) Polymorphic: have many alleles that allow the technique to characterise between individuals; (3) Co-dominant: the heterozygous individual can be distinguished because it shows simultaneously the trait of its two homozygous parents; (4) Not sensitive to the environment; (5) Non epistatic; and (6) Multiallelic (Semagn *et al*., 2006; Kumar *et al*., 2009).

1.7.7 Genetic characteristics revealed by molecular markers

RFLP markers and microsatellites are locus specific: one RFLP probe reveals one locus (polymorph or not), and microsatellite primers amplify only one microsatellite (Varshney *et al*., 2007). Conversely, in the RAPD or AFLP techniques, there is random amplification of one sequence, and they then reveal simultaneously many loci (Semagn *et al*. 2006).

RFLP markers and microsatellites are co-dominant, i.e., it is possible to differentiate heterozygous individuals from the two homozygous parental lines (Hokanson *et al*., 1998). Indeed, the heterozygous individuals are visualised by two bands. In contrast, the RAPD or AFLP techniques, reveal one polymorphism, showing the presence or absence of hybridisation sites but not a fragment length polymorphism (Semagn *et al*., 2006). In this case, there is or there is no amplification of the fragment.

1.7.8 Optimal utilisation of different molecular markers

RFLP and microsatellite techniques, from their genetic characteristics, are used for mapping and in quantitative genetics for QTL detection (Hokanson *et al*., 1998; Noda *et al*., 2009). RAPD and AFLP techniques are used to saturate a particular region of the genome (Lowe *et al*., 1996). The objective is to saturate the region near a gene of interest, and to use the information for position cloning or backcrossing.

1.8 Mapping techniques and methods

1.8.1 Mapping techniques

1.8.1.1 Association mapping or linkage disequilibrium mapping

Association mapping (AM), also known as linkage disequilibrium mapping, is a method of mapping QTLs that uses old linkage disequilibria to link phenotypes to genotypes. AM is based on the idea that traits that have been introduced recently in a population will be found within a given haplotype rather than outside of it (Black *et al*., 2008). Therefore, AM seeks to find a particular and more common genetic marker (most often SNP) in a particular phenotype more often than expected. AM is based on scanning the entire genome for significant associations between a set of SNPs and a particular phenotype. These associations are independently checked to verify that they either: (1) Contribute to the trait of interest directly;

(2) Are linked to/in linkage disequilibrium with a QTL that contributes to the trait of interest (Yu *et al.*, 2008).

AM has been most widely applied to studies of human disease, specifically in the form of a genome-wide association study (GWAS). GWAS is performed by scanning an entire genome for SNPs associated with a particular human disease of interest (Gibson and Muse, 2009; Nussbaum *et al*., 2007). To date, thousands of GWAS studies have been performed on the human genome to identify SNPs associated with cancer, Alzheimer's disease and obesity (Edward *et al*., 2009).The advantage of AM over linkage analysis is that quantitative traits are mapped accurately with high resolution. AM, however, requires prior knowledge of SNPs within the genome of the organism of interest and cannot be performed on non-well-studied species or non-well-annotated genomes (Yu *et al*., 2008).

1.8.1.2 Nested association mapping (NAM)

NAM is a technique developed by the Buckler Laboratory at Cornell University to identify and dissect the genetic structure of complex traits in maize (*Zea mays* L.). NAM (unlike AM) cannot be performed outside of a specifically designed population in crops such as maize (Yu *et al*., 2008) and rice (Ndjiondjop, personal communication). The general goal of NAM is to associate a phenotype of interest to specific genotypes. QTL are identified using joint step-wise regression and joint inclusive composite interval mapping of the combined NAM families using the singular structure of the NAM population (Edward *et al*, 2009). The only study using NAM to identify QTLs was reported by the Buckler Laboratory on the genetic architecture of maize during flowering time (Edward *et al*., 2009). In this study, traits for nearly one million plants were scored, followed by single and joint step-wise regression and inclusive composite interval mapping (ICIM). Thirty nine (39) QTLs explaining 89% of the variance in days to silking and days to anthesis, and 29 QTLs explaining 64% of the variance in the silking-anthesis interval were identified (Edward *et al*., 2009).

A NAM population is created by crossing a reference line with other parental lines (for example, L lines), chosen in order to encompass the diversity of the species and to preserve historic linkage disequilibria. The resulting F_1 plants are self-fertilised for several generations (at least 6 to 8) in order to create Y homozygous recombinant inbred lines (RILs) per family and a total of L x Y RILs within the NAM population. Each RIL is genotyped with selected

common molecular markers for which the reference line has a rare allele, in order to identify recombination blocks. Each of the parental lines is either sequenced or high-density genotyped, and show the overlaying of the sequencing/genotyping and the recombination blocks identified for each RIL (Yu *et al*., 2008; Michael *et al*., 2009).

NAM is a tremendous useful tool to investigate agronomic traits. NAM is also statistically very powerful to identify QTLs for agriculturally relevant traits and to relate those QTLs to homolog and candidate genes in species (Edward *et al*., 2009), despite some disadvantages including several crosses, the lack of NAM populations coupled with a high cost. NAM combines the advantages and eliminates the disadvantages of linkage analysis and AM. NAM takes advantage of both old and recent recombination events, and requires low marker density, provides high allele richness and high mapping resolution, and has high statistical power, with none of the weaknesses of linkage analysis and AM (Yu *et al*., 2008; Michael *et al*., 2009).

1.8.1.3 Bulked segregant analysis (BSA)

BSA or selective genotyping is based on the comparison between two DNA bulks from a segregating population. Each of these bulks comprises DNA from individuals showing the extreme phenotypes (i.e. high or low) of the trait of interest (Becker *et al*., 2011). BSA has been used effectively, in conjunction with RFLP, RAPD and SSR markers, for mapping and tagging rice gall midge resistance genes (Bennett *et al*., 2004; Milton *et al*., 2007).

A major limitation of BSA is that if the experiment is aimed at analysing a number of traits, then by selecting the extremes of each trait one would select most of the population and thus no reduction in genotyping can be obtained. BSA is thus most appropriate for the cases where only one trait is being analysed. This assertion is valid when BSA is applied to QTL detection (Darvasi, 1997).

1.8.2 Mapping methods

1.8.2.1 Analysis of variance (ANOVA)

ANOVA is the simplest method among the QTL mapping methods. It is sometimes called “marker regression” at the marker loci (Soller *et al.*, 1976) and requires a backcross to compare the averages of the two marker genotype groups using T-statistic (Bailey, 1995). For

other types of crosses (such as the intercross), leading to more than two possible genotypes, a general form of ANOVA, which provides an F-statistic, is used (Broman, 2001). However, the ANOVA approach for QTL mapping has three important disadvantages: (1) QTL location and QTL effect are not separately estimated: QTL location is indicated only by markers that give the greatest differences between genotype group averages, and the apparent QTL effect at a marker is smaller than the true QTL effect because of recombination between the marker and the QTL; (2) The analysis does not accept individuals whose genotypes are missing at the marker and (3) When the markers are widely spaced, the QTL may be quite far from all other markers, therefore, the power for QTL detection decreases (Zeng, 1993; Broman, 2001).

1.8.2.2 From interval mapping to multiple interval mapping via composite interval mapping

The concept of using genetic markers to map QTLs has been reported since 1923 (Sax, 1923). With the advent of fine molecular techniques applicable to various organisms, the systematic mapping and analysis of individual QTL has been greatly facilitated.

Interval mapping (IM) has been first developed by Lander and Botstein (1989) to overcome the three disadvantages of ANOVA at marker loci. It is currently the most used-approach for QTL mapping and consists of using one marker interval at a time to construct a putative QTL by performing a likelihood ratio test at every position in the interval (Broman, 2001). This approach of IM considers one QTL at a time so that QTL identification and estimation are biased when multiple QTLs are located in the same linkage group (Haley and Knott, 1992; Zeng, 1994). To deal with this, IM was combined with multiple regression analysis (Jansen, 1993; Zeng, 1994) and named as composite interval mapping (CIM).

CIM consists of using a sub-set of markers loci as covariates to control other QTLs and to reduce the residual variance such that the test can be improved when testing for the putative QTL in an interval (Jansen and Stam, 1994). The key problem with CIM concerns the choice of suitable marker loci to serve as co-variates. Therefore, to have a more powerful and precise QTL detection technique, CIM was extended to a multiple QTL model for mapping multiple QTLs in a way that QTL is directly controlled to further improve QTL mapping (Broman, 2001). This approach was termed as multiple interval mapping (MIM) by Kao *et al*. (1999). MIM uses multiple marker intervals simultaneously to fit multiple putative QTLs directly in the model for mapping QTL**.**

1.9 Heritability and genetic variability

Heritability is the main measure of genetic variation in quantitative traits (Sabu *et al.*, 2009). The total variation of a trait in a population depend whether on genetic variation or environmental variation (Ahmad *et al*., 2008). Heritability is the proportion that is genetic, not environmental, out of that total. Heritability measures the genetic differences amongst individuals or the proportion of phenotypic variance due to genetic factors (Falconer and Mackay, 1996). It is not the fact of a trait to be inherited because if all individuals in the population inherit the same trait there is no genetic variation and therefore no heritability (Wray and Visscher, 2008). Values of heritability vary from 0 (none of the phenotypic variation in the population is genetic) to 1 (all of the phenotypic variation in the population is genetic).

Heritability has never been determined for resistance to AfRGM. However, heritability has been estimated for many agromophological traits in rice. These traits include tillering ability, number of days to first heading, thousand grains weight, plant height and panicle length, (Panwar *et al.*, 1997, Sarawgi *et al.*, 2000; Gannamani, 2001; Sao, 2002 and Rita *et al*., 2009). Moreover, narrow sense heritability has been estimated for penultimate leaf length, penultimate leaf width, flag leaf length, flag leaf width, ligule length, number of days to booting, number of days to maturity by Hoshikawa (1989), Moldenhauer *et al*. (1994), Gravois and Helms (1996), Wayne and Dilday (2003), Sabu *et al*. (2009) and Fahliani *et al*. (2010),

1.9.1 Different types of heritability

1.9.1.1 Broad-sense heritability

Broad-sense heritability is referred to as the proportion of a trait's variation that is due to genetics (Sleper and Poehlman, 2006). It is symbolised as H^2 (to indicate that the units of variance are squared) and is translated into mathematics as: $H = V_G / V_T$

Because of wide variations in its values for the same traits, many plant breeders have found that it does not accurately reflect the results of selection experiments (Falconer, 1989). One of the major constraints of broad-sense heritability comes from grouping all genetic phenomena into a single V_G factor which is paradoxically not directly or entirely inherited by an offspring from the parents. Genetic variance is split into additive genetic variance (V_A) and dominance genetic variance (V_D). $V_G = V_A + V_D$ (Sabu *et al*., 2009). Additive variance is the variance in a trait

that is due to the effects of each individual allele being added together, without any interactions with other alleles or genes (Falconer, 1989).

Dominance variance is the variance that is due to interactions between alleles: synergy effects due to two alleles interacting to make the trait greater (or lesser) than the sum of the two alleles acting alone. Dominance variance also includes both interactions between alleles of the same gene and interactions between different genes, which is referred to as epistasis variance (Rojas and Sprague, 1952). Dominance variance is not directly inherited by offspring from parent. It is due to the interaction of genes from both parents within the individual, and only one allele being passed from each parent to the offspring (Falconer, 1989).

1.9.1.2 Narrow-sense heritability

Given that dominance variance is not predictable and does not affect the mean or variance of the offspring of a selection cross in a systematic fashion, additive genetic variance is the useful measurable variance left (Falconer, 1989). Breeders and other scientists therefore focus on narrow-sense heritability, h, as a measure of heritability: $h = V_A / V_T$, V_T being the total variance.

Narrow-sense heritability can also be calculated directly from breeding experiments (Ulrich, 2007). For this reason, it is also called "realised heritability" and its value as measured by selection experiments, is remarkably constant over many generations. Selection can be continued for many generations without apparently running out of genetic effect (Visser *et al*., 2008).

1.9.2 Measure of heritability

A standard method for measuring narrow-sense heritability is a mid-parent offspring regression that consists of measuring the slope of the regression plot between the average trait values of the offspring of each family versus the average trait values of the two parents of each family (called the mid-parent value) (DeFries and Fulker, 1985). The principle that supports the test is that if the differences between the parents (first generation) are genetic, then they will be passed on to their offspring, which will tend to resemble their parents (Moose and Mumm, 2008).

1.9.3 Advantages and limitations of the method

The mid-parent offspring regression method is an easy way to estimate heritability but it is not perfect. Indeed, the trait under study should be normally distributed to fit regression analysis (DeFries and Fulker, 1985). Moreover, if parents and their offspring are in a similar environment that is different from other families, then parents and offspring may resemble each other for only environmental reasons. This is the case of higher heritability of the size of individuals, genetically the same, in the areas with good nutrients or poor nutrient (Hill *et al*., 2008). The mid-parent offspring regression would overestimate heritability because of environmental similarities between parents and offspring (Falconer and Mackay, 1996).

This problem could be overcome by raising the offspring of many families in constant environmental conditions (greenhouse) which is difficult to achieve.

References

Adeyemi, P. and Vodouhe, S.R. (1996). Amélioration de la productivité des variétés locales de *Oryza glaberrima* Steud. par des croisements intra et interspécifiques avec *Oryza sativa* L. *Hybridations interspécifiques au Benin, Rapport*. WARDA, Cotonou, Benin. p. 159-175

Africa Rice Center (WARDA). 2005. Africa Rice Center (WARDA) Annual Report 2003–2004: Towards new horizons. Cotonou, Benin, 56 pp.

Agarwal, M., Shrivastava, N. and Padh, H. (2008). Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Reports* 27:617-31.

Ahmad, W., Khan, N.U., Khalil, M.R., Parveen, A., Aimen, U., Saeed, M., Samiullah, A. and Shah, S.A. (2008). Genetic variability and correlation analysis in upland cotton. *Sarhad Journal of Agriculture* 24:573-580

Akintayo I., Cissé B. and Zadji L.D. (2008). *Guide pratique de la culture des NERICA de plateau*. AfricaRice. Cotonou, Benin. 28p.

Akio, H., Zhijian, L. and Norimoto, M. (1990). A polyethylene glycol-mediated protoplast transformation system for production of fertile transgenic rice plants. *Plant Physiology* 93:857-863.

Alvarez, A.E., Tjallingii, W.F., Garzo, E., Vleeshouwers, V., Dicke, M. and Vosman, B. (2006). Location of resistance factors in the leaves of potato and wild tuber-bearing *Solanum* species to the aphid *Myzus persicae*. *Entomologia Experimentalis et Applicata* 121:145-157.

Alan, M.S., Zan, K. and Alluri, K. (1985). Gall midge *Orseolia oryzivora* in Zambia. *International Rice Research Newsletter* 10:15-16.

Al-Babili, S. and Peter B. (2005). Golden rice: Five years on the road, five years to go? *Trends in Plant Science* 10:565-573.

Amudhan, S., Prasad Rao, U. and Bentur, J.S. (1999). Total phenol profile in some rice varieties in relation to infestation by Asian rice gall midge *Orseolia oryzae* (Wood-Mason). *Current Science* 76:1577-1580.

Ananthakrishnan, T.N. (1998). Insect gall system: patterns, process and adaptive diversity. *Current Science* 75:672-676.

Angladette, A. (1966). *Le Riz.* Maisonneuve et Larose, Paris, France.930 p.

Armour, J.A.L., Alegre, S.A., Miles, S., Williams, L.J. and Badge, R.M. (1999). Minisatellites and mutation processes in tandemly repetitive DNA. In: *Microsatellites: evolution and applications*. Goldstein, D.B. and Schlotterer, C. (Ed). Oxford University Press, Oxford, United Kingdom. p 24-33.

Ba, N.M., Dakouo, D., Nacro, S. et Ouédraogo, P.A. (2004). Variation saisonnière des populations pré-imaginales de la cécidomyie africaine du riz, *Orseolia oryzivora* et de ses dégâts en fonction des plantes hôtes dans le sud-ouest du Burkina Faso. *International Journal of Tropical Insect Science* 24:177-183.

Balasubramanian, V., Sie, M., Hijmans, R.J. and Otsuka, K. (2007). Increasing rice production in Sub-Saharan Africa: challenges and opportunities. *Advances in Agronomy* 94:55-103.

Becker, M. and Johnson, D.E. (2001). Cropping intensity effects on dryland rice yield and sustainability in West Africa. *Nutrient Cycling in Agroecosystems* 59:107-117.

Becker, A., Chao, D-Y., Zhang, X., Salt, D.E. and Baxter, I. (2011). Bulk segregant analysis using single nucleotide polymorphism microarrays. *PLoS ONE* 6: e15993.

Behura, S.K., Sahu, S.C., Rajamani, S., Devi, A., Mago, R., Nair, S. and Mohan, M. (1999). Differentiation of Asian rice gall midge, *Orseolia oryzae* (Wood-Mason), biotypes by sequence characterized amplified regions (SCARs). *Insect Molecular Biology* 8:391-397.

Bennett, J., Bentur, J.S., Pasalu, I.C. and Krishnaiah, K. (2004). *New approaches to gall midge resistance in rice.* Proceedings of the International Workshop, 22-24 Nov. 1998, Hyderabad, India. International Rice Research Institute and Indian Council of Agricultural Research, Los Baños, Philippines. 195 p.

Bhattramakki, D., Dolan, M., Hanafey, M., Wineland, R., Vaske, D., Register, J.C., Tingey, S.V. and Rafalski, A. (2002). Insertion-deletion polymorphisms in 3' regions of maize genes occur frequently and can be used as highly informative genetic markers. *Plant Molecular Biology* 48:539-547.

Biradar, S.K., Sundaram, R.M., Thirumurugan, T., Bentur, J.S., Amudhan, S., Shenoy, V.V., Mishra, B., Bennet, J. and Sarma, N.P. (2004). Identification of flanking SSR markers for a major rice gall midge resistance gene Gm1 and their validation. *Theoretical and Applied Genetics* 10:1468-1472.

Black, W.Ct., Gorrochetegui-Escalante, N., Randle, N.P. and Donnelly, M.J. (2008). The Yin and Yang of linkage disequilibrium: mapping of genes and nucleotides conferring insecticide resistance in insect disease vectors. *Advances in Experimental Medicine and Biology* 627:71-83.

Bonjean, A. et Picard, E. (1990). *Les céréales à paille: origine, histoire, économie, sélection*. Softword-Groupe ITM, Paris, France. 208 p.

Bonzi, S.M. (1980). Wild host plants of the rice gall midge *Orseolia oryzae* W.M. (Diptera, Cecidomyiidae) in Upper Volta. *West Africa Rice Development Association Technical Newsletter* 2:5-6.

Botstein, D. and Risch, N. (2003). Discovering genotypes underlying human phenotypes: past successes for Mendelian disease, future approaches for complex disease. *Nature Genetics* 33:228-237.

Botstein, D., White, R.L., Skolnick, M. and Davis, R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32:314 331.

Bouchard, D., Ouedraogo, A., Boivin, G. and Kinda, A. (1992). Mass rearing and life cycle of the African rice gall midge, *Orseolia oryzae* Harris and Gagné, in Burkina Faso. *Tropical Pest Management* 38:450-452.

Bredemeijer, G.M.M., Arens, P., Wouters, D. and Visser, D. (1998). The use of semi-automated fluorescent microsatellite analysis for tomato cultivar identification. *Theoretical and Applied Genetics* 97:584-590.

Brettel, J. H. (1980). Prospects for arthropods pest control utilizing cotton plant resistance. In: *Proceedings of cotton pest control workshop*. Union Carbide Agricultural Chemicals Newsprint, Johannesburg, South Africa. p 19-28.

Broman, K.W. (2001). Review of statistical methods for QTL mapping in experimental crosses. *Lab Animal* 30:44-52.

Brookes, A.J. (1999). The essence of SNPs. *Gene* 234:177-186.

Caetano-Anolles, G. (1994). MAAP- A versatile and universal tool for genome analysis. *Plant Molecular Biology* 25:1011-1026.

Caetano-Anolles, G., Bassam, B.J. and Gresshoff, P.M. (1992). Primer-template interactions during DNA amplification fingerprinting with single arbitrary oligonucleotides. *Molecular General Genetics* 235:157-165.

Carpenter, A.J. (1978). Rice history. In: *Rice in Africa.* Buddenhagen, I.W. and Persley, G.J. (Ed.). Academic Press, London, United Kingdom. p 3-10.

Chang, T. T. (1976). The origin, evolution, cultivation, dissemination and diversification of Asian and African rice. *Euphytica* 25:425-441.

Chang, T.T. (1985). Crop history and genetic conservation in rice: a case study. *Iowa State Journal of Research* 59: 445-455.

Chatterjee, M.F. (1948). Botany of the wild and cultivated rice. *Nature* 160:234-237.

Chen, M.S., Zhao, H.X., Zhu, Y.C., Scheffler, B., Liu, X.M., Liu, X., Hulbert, S. and Stuart, J.J. (2008). Analysis of transcripts and proteins expressed in the salivary glands of Hessian fly (*Mayetiola destructor*) larvae. *Journal of Insect Physiology* 54: 1-16.

Cheng, C., Tsuchimoto, Ohtsubo, S.H. and Ohtsubo, E. (2002). Evolutionary relationships among rice species with AA genome based on SINE insertion analysis. *Genes & Genetic Systems* 77:323-334.

Chevalier, A. (1932). Nouvelle contribution à l'étude systématique des *Oryza*. *Review of Botany and Application* 136:1014-1032.

Cho, R.J., Mindrinos, M., Richards, D.R., Sapolsky, R.J., Anderson, M., Drenkard, E., Dewdney, J., Reuber, T.L., Stammers, M., Federspiel, N., Theologis, A., Yang, W.H., Hubel, E., Au, M., Chung, E.Y., Lashkari, D., Lemieux, B., Dean, C., Lipshutz, R.J., Ausubel, F.M., Davis, R.W. and Oefner, P.J. (1999). Genome-wide mapping with biallelic markers in *Arabidopsis thaliana*. *Nature Genetics* 23:203-207.

Cho, Y. G., Blair, M.W., Panaud, O., McCouch, S.R. and Cho, Y.G. (1996). Cloning and mapping of variety-specific rice genomic DNA sequences: Amplified fragment length polymorphisms (AFLP) from silver-stained polyacrylamide gels. *Genome* 39:373-378.

Coyne, D.L., Sahrawat, K.L. and Plowright, R.A. (2004). The influence of mineral fertilizer application and plant nutrition on plant-parasitic nematodes in dryland and wetland rice in Côte d'Ivoire and its implications in long-term agricultural research trials. *Experimental Agriculture* 40:245-256.

Crutzen, P.J., Mosier, A.R., Smith, K.A. and Winiwater, W. (2008). N_2O release from agro-biofuel production negates global warming reduction by replacing fossil fuels. *Atmospheric Chemistry and Physics* 8:389-395.

Dubuis, C. and Haas, D. (2007). Cross-species GacA-controlled induction of antibiosis in pseudomonads. *Applied and Environmental Microbiology* 73:650-654.

Dakouo, D., Nacro, S. and Sié, M. (1988). Evolution saisonnière des infestations de la cécidomyie du riz, *Orseolia oryzivora* Harris and Gagné (Diptera, Cecidomyiidae) dans le Sud-Ouest du Burkina Faso. *Insect Science Application* 9: 467-472.

Dale, D. (1994). Insect pests of rice plant: Their biology and ecology. In: *Biology and Management of Rice Insects.* Heinrichs, E.A. (Ed.). Wiley Eastern, New York, USA. p 363- 485.

Darvasi, A. (1997). The effect of selective genotyping on QTL mapping accuracy. *Mammalian Genotype* 8:67-68.

DeFries, J.C and Fulker, D.W. (1985). Multiple regression analysis of twin data. *Behavior Genetics* 15: 467-473.

Edward, S.B., James, B.H., Peter J.B., Charlotte, B.A., Patrick, J.B., Chris, B., Elhan, E., Sherry, F.G., Arturo, G., Jeffrey, C.G., Major, M.G., Carlos, H., Kate, G., Dallas, E. K., Sara, L., Nicholas, K., Lepak, H.L., Sharon, E.M., Gael, P., Jason, A.P., Marco, O.R., Torbert, R.R., Cinta, R.M., Susan, R., Stella, S., Hector, S.V., Sofia da Silva, H., Qi, S., Feng, T., Narasimham, U., Doreen, W., Heather, Y., Jianming, Y., Zhiwu, Z., Stephen, K. and Michael, D.Mc. (2009). The genetic architecture of maize flowering time. *Science* 325:714-718.

Ehtesham, N.Z., Bentur, J.S. and Bennett, J. (1995). Highly repetitive DNA sequence elements from *Orseolia oryzae* (Wood-Mason) discriminate between the Indian isolates of the Asian rice gall midge and the paspalum midge. *Electrophoresis* 16:1762-1765.

Engle, L.J., Simpson, C.L. and Landers, J.E. (2006). Using high-throughput SNP technologies to study cancer. *Oncogene* 25:1594-1601.

Ethel, D., Umeh, N. and Joshi, R.C. (1993). Aspect of biology and natural biological control of the African rice gall midge, *Orseolia oryzivora* Harris and Gagné (Diptera Cecidomyiidae) in south east Nigeria. *Applied Entomology* 116:391-398.

ETH Zurich, (2009). Combating iron deficiency: Rice with six times more iron than polished rice kernels developed. *ScienceDaily*. Retrieved March 13, 2012, from http://www.sciencedaily.com /releases/2009/07/090721090129.htm.

Fahliani, R.A., Khodambashi, M. Houshmand S. and Arzani. A. (2010). Estimation of the heritability of agro-morphological traits in rice (*Oryza sativa* L.) using F2:3 families. *African Journal of Agricultural Research* 5:1297-1303.

Falconer, D.S., (1989). *Introduction to Quantitative Genetics*. Third edition. Longmans, London, United Kingdom. 438p.

Falconer, D.S. and Mackay, T. (1996). *Introduction to Quantitative Genetics*. Fourth edition. Longmans, Essex, England, United Kingdom.

FAOSTAT (2006). The state of food insecurity in the World. Eradicating world hunger: taking stock ten years after the World Food Summit. 40p http://www.fao.org/icatalog/intere.htm/ last accessed June 2010.

FAOSTAT (2009). (www.faostat.fao.org)/Agriculture and Food Trade/ last accessed June 2010.

Fernando, H.E. (1972). Biology and laboratory culture of the rice gall midge and studies on varietal resistance. In: *Rice Breeding*. International Rice Research Institute, Manila, Philippines. p. 343-351.

Folkertsma, R.T., Voort, J.N.A.M.R., Groot, K.E., Zandvoort, P.M., Schots, A., Gommers, F.J., Helder, J. and Bakker, J. (1996). Gene pool similarities of potato cyst nematode populations assessed by AFLP analysis. *Molecular Plant-Microbe Interaction* 9:47-54.

Foster, J.T., Allan, G.J., Chan, A.P., Rabinowicz, P.D., Ravel, J., Jackson, P.J. and Keim, P. (2010). Single nucleotide polymorphisms for assessing genetic diversity in castor bean (*Ricinus communis*). *BMC Plant Biology* 10:13-23.

Gannamani, N. (2001). *Study of Heterosis and Combining Ability by Utilizing Cytoplasmic Genetic Male Sterility and Fertility Restoration System in Rice (Oryza sativa L.),* M.Sc. (Ag.) Thesis, GAU, Raipur, India.

Gerats, T., de Keukeleire, P., Deblaere, R., van Montagu, M. and Zethof, J. (1995). Amplified fragment length polymorphism (AFLP) mapping in Petunia: A fast and reliable method for obtaining a genetic map. *Acta Horticulturae* 420:58-61.

Gibson, G. and Muse, S.V. (2009). *A Primer of Genome Science*, Third Edition. MA: Sinauer Associates, Sunderland, USA. 344 p.

Giovanini, M.P., Saltzmann, K.D., Puthoff, D.P., Gonzalo, M., Ohm, H.W. and Williams, C.E. (2007). A novel wheat gene encoding a putative chitin binding lectin is associated with resistance against Hessian fly. *Molecular Plant Pathology* 8: 69–82

Goldstein, D.B. and Pollock D.D. (1997). Launching microsatellites: A review of mutation processes and methods of phylogenetic inference. *Journal of Heredity*88:335-342.

Gómez, J.M.J., Maloo, J.N. (2009). Sequence diversity in three tomato species: SNPs, markers and molecular evolution. *BMC Plant Biology* 9:85-94.

Gravois, K. and Helms, R.S. (1996). Seeding rate effects on rough rice yield, head rice, and total milled rice. *Agronomy Journal* 88:82-84.

Greenhouse gas online (2008). Methane sources - rice paddies. http://www.ghgonline.org/methanerice.htm. Retrieved 6 March 2012.

Grodzicker, T., Williams, J., Sharp P. and Sambrook, J. (1975). Physical mapping of temperature sensitive mutants of adenovirus. *Quantitative Biology* 39:439-446.

Haley, C.S., and Knott, S.A. (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69**:** 315-324.

Harris, K.M. (1960). Notes on gall midge cecidomyidae on Nigeria crops with a description of a new species of *Thomasiella rubsaamen*. *Bulletin of Entomological Research* 50:661-666.

Harris, K.M. (1980). *Family Cecidomyiidae. Catalogue of the Diptera of the Afrotropical region.* Crossekey, R. W. (Ed.). British Museum (Natural History), London, United Kingdom. p 238-251.

Harris, K.M. and Gagné, R.J. (1982). Description of African rice gall midge, *Orseolia oryzivora* sp. with comparative notes on Asian rice gall midge, *O. oryzae* (Wood-Mason) (Diptera: Cecidomyiidae). *Bulletin of Entomological Research* 72: 467-472.

Harris, M.O. and Miller J.R. (1982). Synergism of visual and chemical stimuli in the oviposition behavior of Delta antique (Meigen) (Diptera: Anthromyhidae) in *Proceedings of the 5th International symposium on insect plant relationships*. Visser, S.H. and Minks, A.K. (Ed). PUDOC, Wageningen, Netherland. p 117-121.

Heinrichs, E.A. (1994). Host plant resistance. In: *Biology and management of rice insects.* Heinrichs, E.A. (Ed). Wiley Eastern Limited, New Delhi, India. p 517-547.

Heinrichs, E.A and Hamadoum, A. (1995). Rice arthropods of sahelian irrigated rice. In: *Irrigated rice in the Sahel: prospects for sustainable development*. Miezan, K.M., Wopereis, M.C.S., Dingkuhn, M., Deckens, J. and Randolph, T.F (Ed.). West Africa Rice Development Association, Bouaké, Côte d'Ivoire. p 289-300.

Helentjaris, T., Slocum, M., Wright, S., Schaefer, A. and Nienhuis, J. (1986). Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. *Theoretical and Applied Genetics* 61:650-658.

Heubl, G. (2010). New aspects of DNA-based authentication of Chinese medicinal plants by molecular biological techniques. *Planta Medica* 76:1963-1974.

Hibino, H. (1989). Insect-borne viruses of rice. In: *Advances in Disease Vector Research*. Harris, K.F. (Ed). Springer-Verlag, New York, USA.

Hill, W.G., Goddard, M.E. and Visscher, P.M. (2008). Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genetics* 4:2-12.

Himabindu, K., Sundaram, R.M, Neeraja, C.N., Mishra, B. and Bentur, J.S. (2007). Flanking SSR markers for allelism test for the Asian rice gall midge (*Orseolia oryzae*) resistance genes. *Euphytica* 157:267-279.

Himabindu, K., Sama, V.S.A.K, Naik, B.S., Pranitha J. Sinha, P., Sadasiva, N. and Bentur J.S. (2009). Genetic diversity in rice (*Oryza sativa* L.) genotypes resistant and susceptible to the Asian rice gall midge *Orseolia oryzae* (Wood-Mason). *Indian Journal of Plant Genetic Resources* 22:108-112.

Himabindu, K., Suneetha, K., Sama, V. S. A. K. and Bentur, J. S. (2010). A new rice gall midge resistance gene in the breeding line CR57-MR1523, mapping with flanking markers and development of NILs. *Euphytica* 174:179-187.

Hiroshi, F. (1993). Research into the history of rice cultivation using plant opal analysis. Pearsall, Deborah M. and Piperno, Dolores R. *MASCA: Current Research in Phytolith Analysis: Applications in Archaeology and Paleoecology* 10:147-158.

Hiroshi, F., Hiroko, T., Sakae, A., Mika, N., Haruko, O., Kazuko, O., Byung-Hyun, L., Sakiko, H., Seiichi, T., Maurice, S.B., Amane, M., Makoto, M. and Mitsue, M. (2001). Significant accumulation of C_4-specific pyruvate, orthophosphate dikinase in a C_3 plant, rice. *Plant Physiology* 127:1136-1146.

Hokanson, S.C., Szewc-McFadden, A.K., Lamboy, W.F. and McFerson, J.R. (1998). Microsatellite (SSR) markers reveal genetic identities, genetic diversity and relationships in a *Malus* x *domestica* Borkh. core subset collection. *Theoretical and Applied Genetics* 97:67-683.

Hoshikawa, K. (1989). *The Growing Rice Plant: An Anatomical Monograph*. Nobunkyo, Tokyo, Japan.

Hossain, M. (2006). Rice in Africa: Can rice help reduce hunger and poverty in Sub-Saharan Africa? *Rice Today* 5: 41.

International Rice Genome Sequencing Project (2005). The map-based sequence of the rice genome. *Nature* 436:793-800.

IRRI. (1964). *Rice genetics and cytogenetics*. Proceeding of the symposium on Rice genetic and cytogenetics. Amsterdam, Elsevier Pub. Co. International Rice Research Institute, Los Baños, Philippines.

Jain, A., Ariyadasa, R., Kumar, A., Srivastava, M.N., Mohan, M. and Nair, S. (2004). Tagging and mapping of a rice gall midge resistance gene, Gm8, and development of SCARs for use in marker-aided selection and gene pyramiding. *Theoretical and Applied Genetics* 109:1377-1384.

James, C. (2007). Global status of commercialized Biotech/GM crops in 2007. In: *ISAAA Brief 37*. ISAAA Ithaca, New York, USA.

James, C. (2008). Global status of commercialized Biotech/GM crops in 2008. In: *ISAAA Brief 39*. ISAAA Ithaca, New York, USA.

James, C. (2009). Global status of commercialized Biotech/GM crops in 2009. The first fourteen years, 1996 to 2009. In: *ISAAA Brief 41*. ISAAA Ithaca, New York, USA.

James, C. (2010). Global status of commercialized Biotech/GM crops in 2010. In: *ISAAA Brief 42*. ISAAA Ithaca, New York, USA.

Jansen, R.C. (1993). Interval mapping of multiple quantitative trait loci. *Genetics* 135**:**205-211.

Jansen, R.C. and Stam, P. (1994). High resolution of quantitative traits into multiple loci via interval mapping. *Genetics* 136:1447–55.

Janssen, P., Coopman, R., Huys, G., Swings, J., Bleeker, M., Vos, P., Zabeau, M. and Kersters, P. (1996). Evaluation of the DNA fingerprinting method AFLP as a new tool in bacteria taxonomy. *Microbiology* 142:1881-1893.

Jena, K.K., Pasalu, I.C. Krishniah, K. and Khush, G.S. (1998). A RAPD marker for the gene conferring resistance to Indian biotype of BPH. *Rice Genetic Newsletter* 15:133-134.

Jenkins, J.N. and Parrot, W.L. (1971). Effectiveness of frego bract as a boll weevil resistance character in cotton. *Crop Science* 11:739-743.

Joerg, H., Fries, H.R., Meijerink, E. and Stranzinger, G.F. (1996). Red coat color in Holstein cattle is associated with a deletion in the MSHR gene. *Mammalian Genome* 7:317-318.

Johnson, M.W., Welter, S.C., Toscano, N.S., Ting, I.P. and Trumble, J.T. (1983). Reduction of tomato leaflet photosynthesis rates by mining activity of *Liriomyza sativae* (Diptera Agromyzidae). *Journal of Economic Entomology* 76:1061-1063.

Jones, M.P., Audebert, A., Mande, S. and Aluko, K. (1997). Characterization and utilization of *Oryza glaberrima* Steud. in upland rice breeding. In: *Interspecific hybridization*: *Progress and Prospects,* Edited by Jones, M.P. Dingkuhn, M. Johnson, D.E. and Fagade, S.O. Proceedings of the workshop: Africa/Asia Joint Research on Interspecific Hybridization between the Africa and Asian Rice Species *(O. glaberrima and O. Sativa)*. WARDA, Mbé, Bouake, Côte d'Ivoire. p 43-59.

Joshi, R.C. and Venugopal, M.S. (1984). Mechanism of resistance in rice to gall midge *Orseolia oryzae* (Wood-Mason): Antibiosis. *Indian Journal of Entomology* 46:479-481.

Ju-Kon Kim and Hari B. Krishnan. (2002). Making Rice a perfect food: turning dreams into reality. *Journal of Crop Production* 5:93-130.

Kalode, M.B. (1980). The rice gall midge: varietal resistance and chemical control. In: *Rice Improvement in China and Other Asian Countries.* International Rice Research Institute, Manila, Philippines. p 173-193.

Kao, C-H., Zeng, Z-B. and Teasdale, R. D. (1999). Multiple interval mapping for quantitative trait loci. *Genetics* 152: 1203-1216.

Katiyar, S.K., Chandel, G., Tan, Y., Zhang, Y., Huang, B., Nugaliyadde, L., Fernando, K., Bentur, J.S., Inthavong, S., Constantino, S. and Bennett, J. (1995). Biodiversity of Asian rice gall midge (*Orseolia oryzae* Wood Mason) from five countries examined by AFLP analysis. *Genome* 43:322-332.

Katiyar, S.K., Verullcar, S.B., Chandel, G., Shrivastava, M.N., Kamar A. and Bennett J. (1999). Identification and molecular tagging of two new gall midge resistance genes (gm-3 and Gm-5) in rice. *Poster presented in the meeting of International Program on Rice Biotechnology*. October 20-24 1999, Phukhet, Thailand.

Katiyar, S.K. and Bennett, J. (2001). Biotechnology for gall midge resistance, molecular gene tagging and gene pyramiding. In*: International Rice Research Conference*. International Rice Research Institute (IRRI), Los Baños, Laguna, Philippines.

Keim, P.J., Schupp, J., Hill, K., Travis, S.E., Rickmond, K., Adair, D.M., Hughjones, M., Kuske, C.R. and Jackson, P. (1997). Molecular evolution and diversity in *Bacillus anthracis* as detected by amplified fragment length polymorphism markers. *Journal of Bacteriology* 179:818-824.

Khush, G.S. (1997). Origin, dispersal, cultivation and variation of rice. *Plant Molecular Biology* 35:25-34.

Kittur, S.U. and Agarwal, R.K. (1983). Gall midge damage at rice panicle stage. *International Rice Research Newsletter* 8:19.

Korgan, M. and Ortman ,E.E. (1978). Antixenosis: a new term proposed to replace Painter's "non-preference" modality of resistance. *Bulletin of Entomological Society of America* 24:175-176.

Kolodny-Hirsch, D.M., Saunder, J.A. and Harrison, F.P. (1986). Effects of stimulated tobacco hornworm (Lepidoptera: Sphingidae) defoliation on growth dynamics and physiology of tobacco as evidence of plant tolerance to leaf consumption. *Environmental Entomology* 15:1137-1144.

Kumar, A., Jain, A., Sahu, R.K., Shrivastava, M.N., Nair, S. and Mohan, M. (2005). Genetic analysis of resistance genes for the rice gall midge in two rice genotypes. *Crop Science* 45:1631-1635.

Kumar, P., Gupta, V.K., Misra, A.K., Modi, D.R. and Pandey, B.K. (2009). Potential of molecular markers in plant biotechnology. *Plant Omics Journal* 2:141-162.

Lafleur, G. (1994). Effet du carbofuran, de l'isofenphos et du neem sur les principaux ravageurs du riz et parasites de la cecidomyie du riz, au Burkina Faso. *Sahel PV info* 68:9-14.

Lai, K., Tan, Y. and Pan, Y. (1984). Rice gall midge biotypes in Guangdong Province. *International Rice Research Notes* 9:17-18.

Lander, E.S. and Botstein, D. (1989). Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185-199.

Landry, B.S. (1994). DNA mapping in plants. In: *Methods in Plant Molecular Biology and Biotechnology*. Glick, B.R., Thompson, J.E. (Ed). CRC Press, Boca Raton-Ann Arbor-London-Tokyo. p 269-285.

Lin, C., Ho, T.C., Chu, H., Yang, H., Martha, J.M., Krishnarajanagar, N., Chiou, P. and Hopper, J.R. (2005). A comparative study of US EPA 1996 and 1999 emission inventories in the west Gulf of Mexico coast region, USA. *Journal of Environmental Management* 75:303-313.

Lin, K.H., Lai, Y.C., Li, H.C., Lo, S.F., Chen, L.F.O. and Lo, H.F. (2009). Genetic variation and its relationship to root weight in the sweet potato as revealed by RAPD analysis. *Scientia Horticulturae* 120:2-7.

Linares, O.F. (2002). African rice (*Oryza glaberrima*): History and future potential. *Proceeding of the National Academy of Science, USA* 99:16360-16365.

Litt, M. and Luty, J.M. (1989). A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetic* 44:397-401.

Liu, X.L., Bai, J., Huang, L., Zhu, L., Liu, X., Weng, N., Reese, J.C., Harris, M., Stuart, J.J., Chen, M.S. (2007). Gene expression of different wheat genotypes during attack by virulent and avirulent Hessian fly (*Mayetiola destructor*) larvae. *Journal of Chemical Ecology* 33:2171-2194.

Liu, X., Williams, C.E., Nemacheck, J.A., Wang, H., Subramanyam, S., Zheng, C. and M.-S. Chen. (2010). Reactive oxygen species are involved in plant defense against a gall midge. *Plant Physiology*_152:985-999.

Lowe, A.J., Hanotte, O. and Guarino, L. (1996). Standardization of molecular genetic techniques for the characterization of germplasm collections: The case of random amplified polymorphic DNA (RAPD). *Plant Genetic Resources Newsletter* 107: 50-54.

Lynch, M., and Walsh, B. (1998). *Genetics and analysis of quantitative traits*. Sinauer Associates, Sunderland, United Kingdom.

Majer, D., Mithen, R., Lewis B.G., Vos, P., and Oliver, R.P. (1996). The use of AFLP fingerprinting for the detection of genetic variation in fungi. *Mycology Research* 100:1107-1111.

Maji A.T. (1998). Genetics of resistance to rice gall midge (*Orseolia* spp.) Paper presented during the WARDA/NARS group discussion and planning meeting on Bioecology and management of African rice gall midge, 12-16 October at IITA, Ibadan, Nigeria.

Maji A.T., Gana A.S. and Ukwungwu M.N. (2010). Responses of *Oryza glaberrima* accessions to rice stresses and their morphological characteristics. *African Journal of Agricultural Research* 6:732-737.

McDonald, D.J. (1979). Rice. Chapter 3. In: *Australian field crop Vol. 2: Tropical cereals, oilseeds, grains legumes and other crops.* Angus and Robertson, London, United Kingdom. P 70-94.

McDonald, D.B. and Potts, W.K. (1997). DNA microsatellites as genetic markers for several scales. In: *Avian Molecular Evolution and Systematics*. Mindell, D.P. (Ed). Academic Press, San Diego, USA. p. 29-49.

Medicinal uses of rice: htpp://www.rice-trade.com/medicinal-uses-of-rice.html. Retrieved 19 March, 2010.

Michael, D. Mc., Stephen, K., Hector, S.V., Peter, B., Huihui, L., Qi, S., Sherry, F.G., Jeffry, T., Charlotte, A., Christopher, B., Patrick, B., Chris, B., Magen, E., Kate, G., Carlos, H., Dallas, K., Nick, L., Sharon, E.M., Brooke, P., Gael, P., Susan, R., Marco, O.R., Stella, S., Heather, Y., Mark H., Elizabeth, J., Stephen, S., Jeffrey, C.G., Major, G., Doreen, W., James, B.H. and Edward, S.B. (2009). Genetic properties of the maize nested association mapping population. *Science* 325:737-740.

Milton, L.J., Abhishek Dass, S.C., Sahu, L.B. and Devendra K.C. (2007). A RAPD marker identified a susceptible specific locus for gall midge resistance gene in rice cultivar ARC5984. *Crop Protection* 26:1431-1435.

Mishra, H.S. and Rathore, T.R. (1997). Root growth, water potential, and yield of irrigated rice. *Irrigation Science* 17: 69-75.

Mohan, M., Nair, S., Bentur, U., Rao, P. and Bennett, J. (1994). RFLP and RAPD mapping of the rice Gm2 gene that confer resistance to biotype 1 of gall midge *Orseolia oryzae*. *Theoretical and Applied Genetics* 87:782-788.

Moldenhauer, K.A.K., Wells, B. and Helms, R. (1994). Rice growth stages. In: *Rice Production Handbook*: Helms, R.S. (Ed). Cooperative extension service print shop, Little Rock, Arkansas, USA. p 5-12.

Moose, S.P. and Mumm R.H. (2008). Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiology* 147: 969-977.

Morishima, H. (1969). Phenetic and phylogenetic relationship among strains of *Oryza perennis*, estimated by methods of numerical taxonomy. *Evolution* 23:249-443

Muzaffar, D.T.A., Khoo, K.K. and Selvamany, G. (2005). *The History of Medicine in Malaysia, the Foundation Years*. Academy of medicine, Kuala Lumpur, Malaysia. 120 p.

Nacro, S., Dakouo, D. and Heinrichs, E.A. (1995). Population dynamics, host plant damage and parasitism associated with the African rice gall midge in southern Burkina Faso. *Insect Science and its Application* 16:251-257.

Nacro, S., Heinrichs, E.A. and Dakouo, D. (1996). Estimation of rice yield losses due to the African rice gall midge, *Orseolia oryzivora* Harris and Gagné. *International Journal of Pest Management* 42:331-334.

Nacro, S. and Nénon, J.-P. (2009). Comparative study of the morphology of the ovipositor of *Platygaster diplosisae* (Hymenoptera: Platygasteridae) and *Aprostocetus procerae* (Hymenoptera: Eulophidae) two parasitoids associated with the African rice gall midge, *Orseolia oryzivora* (Diptera: Cecidomyiidae). *Psyche* 2009:1-7.

Nair, S., Bentur, J.S., Prasada, R. U. and Mohan, M. (1995). DNA markers tightly linked to a gall midge resistance gene (*Gm2*) are potentially useful for marker-aided selection in rice breeding. *Theoretical and Applied Genetics* 91:68-73.

Nair, S., Kumar, A., Srivastava, M.N. and Mohan, M. (1996). PCR-based DNA markers linked to a gall midge resistance gene, *Gm4t*, has potential for marker-aided selection in rice. *Theoretical and Applied Genetics* 92:660-665.

Nayar, N.M. (1973). Origin and cytogenetics of rice. *Advances in Genetics* 17:132-292

N'cho, A.L. (1992). Les stades de croissance de la plante de riz. *Rapport de stage sur la production du riz pluvial*. ADRAO, Bouaké, Côte d'Ivoire. 14 p.

Ndjiondjop, N.M-N. (1999*). La résistance du riz au virus de la panachure jaune: Hérédité, mécanismes et cartographie génétique*. Thèse de doctorat. Université de Montpellier, France. 137 p.

Nguyen, N. (2005). http://www.fao.org/newsroom/en/news/2006/1000379/index.html/ accessed November 2010.

Norris, D.M. (1986). Anti-feeding compounds. In: *Chemistry of Plant Protection.* Hang, G. and Hoffman (Ed). Springer-Verlag, Berlin, Germany. p 97-146.

Noda, A., Nomura, N., Mitsui, Y. and Setoguchi, H. (2009). Isolation and characterisation of microsatellite loci in *Calystegia soldanella* (Convolvulaceae), an endangered coastal plant isolated in Lake Biwa, Japan. *Conservation Genetics* 10:1077–1079.

Nussbaum, R.L., McInnes, R.R. and Willard, H.F. (2007). *Genetics in Medicine*. Saunders Elsevier, Philadelphia, USA.

Nwilene, F.E., Williams, C.T., Ukwungwu, M.N., Dakouo, D., Nacro, S., Hamadoun, A., Kamara, S.I., Okhidievbie, O., Abamu, F.J. and Adam, A. (2002). Reactions of differential rice genotypes to African rice gall midge in West Africa. *International Journal of Pest Management* 48:195-201.

Nwilene, F.E., Harris, K.M., Okhidievbie, O., Onasanya, A., Sere, Y. and Ingelbrecht, I. (2006a). Morphological diversity and genomic DNA fingerprinting of the African rice gall midge *Orseolia oryzivora* (Diptera: Cecidomyiidae) and of two other species of African *Orseolia*. *International Journal of Tropical Insect Science* 26:256-265.

Nwilene, F.E., Nwanze, K.F. and Okhidievbie, O. (2006b). *African Rice Gall Midge: Biology, Ecology and Control. Field Guide and Technical Manual*. Africa Rice Center (WARDA), Cotonou, Benin. 24 p.

Ogah, E.O., Nwilene, F.E., Ukwungwu, M.N., Omoloye, A.A. and Agunbiade, T.A. (2009). Population dynamics of the African rice gall midge *Orseolia oryzivora* (Diptera: Cecidomyiidae) and its parasitoids in the forest and southern Guinea savanna zones of Nigeria. *International Journal of Tropical Insect Science* 29:86-92.

Omoloye, A.A. (2010). Effects of genotype variation in the cultivated rice species on the biology and galling characteristics of the African rice gall midge, *Orseolia oryzivora* (Diptera: Cecidomyiidae). *Entomological Science* 13:320-327.

Ou, S.H. (1985). *Rice Diseases.* 2nd Ed. Commonwealth Mycology Institute, Surrey, United Kingdom. 360 p.

Pan, X-B., Chen, Z-X., Zuo, S-M., Zhang, Y-F., Wu, X-J., Ma, N., Jiang, Q-X., Que, J-H., and Zhou, C-H. (2009). New rice cultivar Wulingjing 1 resistant to *Rice stripe virus* developed by Marker Assisted Selection. *Acta Agronomica Sinica* 35:1851–1857.

Panwar, R.P., Ashvani, S.R., Sharma, K., Arya, K.P.S and Panwar, A. (1997). Genetic variability and inter-relationship in rice (*Oryza sativa* L.). *Advances in Plant Science* 10:29-32.

Perez de la Vega, M.P. (1993). Biochemical characterization of populations. In: *Plant Breeding: Principles and prospects*. Hayward, M.D., Bosemark, N.O. and Romagosa, I. (Ed). Chapman and Hall, London, United Kingdom. p 182-201.

Peter, I., Ukwungwu, M.N. and Joshi, R.C. (1991). Criblage au champ pour la resistance a la cecidomyie Africaine du riz. *Guide recherche de l'IITA No. 34*. Ibadan. Nigeria. 16 p.

Pham, J-L. (1992). *Evaluation des ressources génétiques des riz cultivés en Afrique par hybridation intra et interspécifique*. Thèse de doctorat. Université de Paris XI, Orsay, France. 233 p.

Plowright, R.A. and Hunt, D.J. (1994). Plant parasitic nematodes of dryland, hydromorphic and inland swamp rice ecosystems in Cote d'Ivoire. *Afro-Asian Journal of Nematology* 4:61-67.

Portères, R. (1956). Taxonomie agro-botanique des riz cultivés *Oryza sativa* L. et *O. glaberrima* Steud. *Journal d'Agriculture Tropicale et de Botanique Appliquée* 3:341-384.

Prasad, K., Chatterjee, S.M. and Misra, B.C. (1975). Inheritance of gall midge (*Pachydiplosis oryzae* Wood-Mason) resistance to rice with particular reference to cytoplasmic influence on its expression. *Current Science* 44:63-66.

Rajamani, S., Pasalu, I.C. and Mathur, K.C. (1979). Effect of gall midge attack in paddy at flowering stage. *Current Science* 48:832-841.

Rajamani, S. (1982). *Factors Affecting Gall Midge Resistance in Paddy*. Ph.D. thesis. Orissa University of Agriculture and Technology, Bhubaneswar, India. p

Rana, B.S., Singh, B.U. and Rao, N.G.P. (1985). Breeding for shootfly and stemborer resistance in sorghum. In: *Proceeding of International Sorghum Entomology Workshop*. Texas A.M. University College Station, TX/ICRISAT, Patoncheru. India. P. 347-360.

Rawat, N., Sinha, D.K., Rajendrakumar, P., Shrivastava, P., Neeraja, C.N., Sundaram, R.M., Nair, S. and Bentur, J.S. (2010). Role of pathogenesis-related genes in rice gall midge interactions. *Current Science* 99:1361–1368.

Richards, P. (1996). Ecology, Culture and Domestication. In: *Redefining Nature*, Ellen, R and Fukui, K. (Ed). Berg, Oxford, United Kingdom. 297 p.

Rita, B., Sarawgi, A.K. and Verulkar, S.B. (2009). Study of heritability, genetic advance and variability for yield contributing characters in rice. *Bangladesh Journal of Agricultural Research* 34:175-179.

Rojas, B.A. and Sprague, G.F. (1952). A comparison of variance components in corn yield trials: III. General and specific combining ability and their interaction with locations and years. *Agronomy Journal* 44:462-467.

Romero, G., Adeva, C. and Battad, Z. (2009). Genetic fingerprinting: Advancing the frontiers of crop biology research. *Philippines Science Letter* 2:8-13.

Roschevicz, R.J. (1931). A contribution to the knowledge of rice. *Bulletin of Applied Botanic and Genetic Plant Breeding* 27:1-133.

Sabu, K.K., Abdullah, M.Z., Lim, L.S. and Wickneswari, R. (2009). Analysis of heritability and genetic variability of agronomically important traits in *Oryza sativa* × *O. rufipogon* cross. *Agronomy Research* 7:97-102.

Sain, M. and Kalode, M.B. (1994). Greenhouse evaluation of rice cultivars for resistance to gall midge, *Orseolia oryzae* (Wood-Mason) and studies on mechanism of resistance. *Insect Science Application* 15:67-74.

Sama, V.S.A.K., Himabindu, K., Sundaram, R.M., Iraktamathb, C.V. and Bentur, J.S. (2010). Mapping of a rice gall midge resistance gene, *gm3*, in RP 2068-18-3-5 and *in silico* identification of candidate gene(s). *Rice Genetics Newsletter* 25:76-77.

Sampath, S. (1966). The genus *Oryza*: an evolutionary perspective. *Oryza* 3:30-34.

Sao, A. (2002). *Studies on Combining Ability and Heterosis in F1 Rice Hybrids Using Cytoplasmic Male Sterile Lines*. M.Sc. (Ag.) Thesis, IGAU, Raipur, India.

Sarawgi, A.K., Rastogi, N.K. and Soni, D.K. (2000). Studies on some quality parameters of indigenous rice in Madhya Pradesh. *Annual Agricultural Research* 21:258-261.

Sarla, N. and Mallikarjuna Swamy, B.P. (2005). *Oryza glaberrima*: a source for the improvement of *Oryza sativa*. *Current Science* 6:955-963.

Sarwat, M., Das, S. and Srivastava, P.S. (2008). Analysis of genetic diversity through AFLP, SAMPL, ISSR and RAPD markers in *Tribulus terrestris*, a medicinal herb. *Plant Cell Reports* 27:519-528.

Sax, K. (1923). The association of size difference with seed-coat pattern and pigmentation in *Phaseolus vulgaris. Genetics* 8**:**552-560.

Saxena, R.C. (1986). Biochemical bases of insect resistance in rice varieties. In: *Natural Resistance of Plants to Pests: Roles of Allelochemicals, ACS symposium s*eries. 296: 142-159. Green, B. and Hedin, P.A. (Ed). American Chemical Society, Washington, USA.

Schlotterer, C. (1998). Microsatellites. In: *Molecular Genetic Analysis of Populations: A Practical Approach*. Hoelzel, A.R. (Ed). IRL Press, Oxford, United Kingdom. p. 237-261.

Shin-Foon, C. (1980). The integrated control of rice gall midge. IRRI Thursday Seminar. International Rice Research Institute, Los Baños, the Philippines. 7 p.

Second, G. (1982). Origin of genetic diversity of cultivated rice (*Oryza* spp). Study of the polymorphism scored at 40 isozyme loci. *Japan Journal of Genetics* 57:25-57.

Second, G. (1985). *Relation Evolutives Chez le Genre Oryza et Processus de Domestication des Riz.* Edition de l'ORSTOM. Coll. Etudes et Thèses, Paris, France. 189 p.

Semagn, K., Bjørnstad, Å. and Ndjiondjop, M.N. (2006). An overview of molecular marker methods for plants. *African Journal of Biotechnology* 5:2540-2568.

Sharma, T.R. (2003). Molecular diagnosis and application of DNA markers in the management of fungal and bacterial plant disease. *Indian Journal of Biotechnology* 2:99-109.

Sié, M. (1991). *Prospection et Evaluation Génétique des Variétés Traditionnelles de Riz (Oryza* sp.) *du Burkina-Faso*. Thèse de doctorat. Université National de Côte d'Ivoire, Abidjan, Côte d'Ivoire. 118 p.

Singh, B.N. (1992). Breeding lowland rice varieties for insect pest resistance. West African Rice Development Association. *WARDA Annual Report* 1992:34-35.

Singh, B.N., Fagade, S., Ukwungwu, M.N., Williams, C.T., Jagtap, S.S., Oladimeji, O., Effisue, A. and Okhidievbie, O. (1997). Rice growing environments and biophysical constraints in different Agro-ecological zones of Nigeria. *Meteorology Journal* 2:35-44.

Singh, B.N. (1998). Breeding for rice gall midge (*Orseolia oryzivora* Harris and Gagné) resistance at WARDA. Paper presented during the WARDA/NARS Group Discussion and Planning Meeting on Bioecology and Management of African Rice Gall Midge, 12-16 October at IITA, Ibadan, Nigeria.

Sleper, D.A. and Poehlman, J.M. (2006). *Breeding Field Crops*. 5th Ed., Wiley-Blackwell Publication, Iowa, USA. 424 p.

Soller, M., Brody, T. and Genizi, A. (1976). On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. *Theoretical and Applied Genetics* 47:35–39.

Soroush, P., Sotelo, G. and Cardona, C. (2011). Characterizing herbivore resistance mechanisms: Spittlebugs on *Brachiaria* spp. as an example. *Journal of Visualized Experiments* 52:3047-3050.

Souleymane, N. (1998). Réunion de discussion et planning sur la bio-écologie et la gestion de la Cécidomyie africaine du riz, *Orseolia oryzivora*. IITA, Ibadan, Nigeria. 6 p.

Staudt, M., Jackson, B., El-Aouni, H., Buatois, B., Lacroze, J-P., Poëssel, J-L. and Sauge, M-H. (2010). Volatile organic compound emissions induced by the aphid *Myzus persicae* differ among resistant and susceptible peach cultivars and a wild relative. *Tree Physiology* 30:1320-1334.

Suwa, T. and Maherali, H. (2008). Influence of nutrient availability on the mechanisms of tolerance to herbivory in an annual grass, *Avena barbata* (Poaceae). *American Journal of Botany* 95:434-440.

Tateoka, T. (1962). Taxonomic studies of genus *Oryza* II. Several species complex. *Botany Magazine* 75: 455-461.

Tautz, D., Trick, M. and Dover, G.A. (1986). Cryptic simplicity in DNA is a major source of genetic variation. *Nature* 322:652-656.

Taylor, D.R., Fomba, S.N., Fannah, S.J. and Bernard, N.M. (1995). African midge pest in Sierra Leone. *International Rice Research Newsletter* 20: 27.

Terachi, T. (1993). The progress of DNA analyzing techniques and its impact on plant molecular systematics. *Journal of Plant Research* 106:75-79.

The Rice Genome Sequencing Project (2005). The map-based sequence of the rice genome. *Nature* 436:793-800.

Thomas, C.M., Vos, P., Zabeau, M., Jones, D.A., Norcott, K.A., Chadwick B.P. and Jones, D.G. (1995). Identification of amplified restriction fragment polymorphism (AFLP) markers tightly linked to tomato *Cf-9* gene for resistance to *Cladosporium fulvum*. *Plant Journal* 8:785-794.

Ukwungwu, M.N. (1987). Effect of spacing and nitrogen on infestation by the rice gall midge *Orseolia oryzivora* Harris and Gagné (Diptera: Cecidomyidae) in Nigeria. *Tropical Pest Management* 33:164-165.

Ukwungwu, M.N., Winslow, N.D. and John, V.T. (1989). Severe outbreak of rice gall midge (GM) in the Savannah zone of Nigeria. *International Rice Research Newsletter* 14:36-37.

Ukwungwu, M.N. (1990). Efficacy of granular isazofos against rice gall midge in Nigeria. *Crop Protection* 9:252-254.

Ukwungwu, M.N. and Alam M.S. (1991). Field resistance of rice cultivars to the gall midge *Orseolia oryzivora* in Nigeria. *International Rice Commission Newsletter* 40:42-43.

Ukwungwu, M.N. and Joshi R.C. (1992). Distribution of the African rice gall midge, *Orseolia oryzivora* Harris and Gagné and its parasitoids in Nigeria. *Tropical Pest Management* 38:241-244.

Ukwungwu, M.N., Williams, C.T. and Okhidievbie, O. (1998) Screening of African rice, *Oryza glaberrima* Steud, for resistance to the African rice gall midge *Orseolia oryzivora* Harris and Gagné. *Insect Science Application* 18:167-170.

Ullrich, S.E. (2007). Breeding field crops. *Crops Science* 47:900-904.

Umeh, E.D.N., Joshi, R.C. and Ukwungwu, M.N. (1992). Biology, status and management of rice insect pests in Nigeria. *Crop Protection* 11: 408-413.

Umeh, E.D.N. and Joshi, R.C. (1993). Aspect of biology, ecology and natural biological control of the African rice gall midge, *Orseolia oryzivora* Harris and Gagné (Diptera: Cecidomyiidae) in south east Nigeria. *Journal of Applied Entomology* 116:391-398.

van Andel, T. (2010). African rice (*Oryza glaberrima* Steud.): Lost crop of the enslaved Africans discovered in Suriname. *Economic Botany* 64:1-10.

Varshney, R.K., Chabane, K., Hendre, P.S., Aggarwal, R.K. and Graner, A. (2007). Comparative assessment of EST-SSR, EST-SNP and AFLP markers for evaluation of genetic diversity and conservation of genetic resources using wild, cultivated and elite barleys. *Plant Society* 173:638-649.

Vaughan, D.A. (1994). *The Wild Relative of Rice: A Genetic Resource Handbook*. International Rice Research Institute, Manila, the Philippines.

Vaughan, D.A. and Morishima, H. (2003). Biosystematics of the genus *Oryza*. In: *Rice: origin, history, technology, and production.* Smith W. (Ed.). John Wiley and Sons, New York, USA. p 27-65.

Vidyachandra, B., Roy, J.K. and Das, B. (1981). Chemical differences in rice varieties susceptible or resistant to gall midge and stem borer. *International Rice Research Newsletter* 6:7-8.

Visscher, P.M., Hill, W.G. and Wray, N.R. (2008). Heritability in the genomics era-concepts and misconceptions. *Nature Reviews Genetics* 9: 255–266.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee. T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., and Zabeau, M. (1995). AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research* 23:4407–4414.

Wang, X., Sun, C., Cai, H. and Zhang, J. (1999). Origin of the Chinese cultivated rice. *Chinese Science Bulletin* 44:295-304.

Watabe, T. (1985). Origin and dispersal of rice in Asia. *East Asian Culture Study* 24: 33-39.

Wayne, S.C. and Dilday, R.H. (2003). *Rice: Origin, History, Technology and Production*. Wiley Series in Crop Science, John Wiley and Sons, Inc. Hoboken, New Jersey, USA.

Weber, D. and Helentjaris, T. (1989). Mapping RFLP loci in maize using B-A translocations. *Genetics* 121:583-590.

Williams, C.T., Ukwungwu, M.N., Singh, B.N., Okhidievbie, O. and Nnabo, J. (1998). Farmer-managed trials in Southeast Nigeria to evaluate the rice variety Cisadane and estimate yield losses caused by the African rice gall midge, *Orseolia oryzivora* Harris and Gagné. *International Journal of Pest Management*. 45:117-124.

Williams, C.T., Okhidievbie, O., Unkwungwu, M.N., Dakouo, D., Nacro, S., Hamadou, A. and Kamara, S.I. (1999a). Multilocational screening of *Oryza sativa* and *Oryza glaberrima* for resistance to African Rice gall midge *Orseolia oryzivora* in West Africa. *International Rice Research Notes* 24:26-27.

Williams, C.T. and Ukwungwu, M.N. (1999b). Farmer-managed trials in south-east Niger to evaluate the rice variety Cisadane and estimate yield losses caused by the African rice gall midge, *Orseolia oryzivora* Harris and Gagné. *International Journal of Pest Management* 45:177-124.

Williams, C.T., Harris, K.M., Ukwungwu, M.N., Nacro, S., Dakouo, D., Nwilene, F.E., Singh, B.N. and Okhidievbie, O. (2002). African rice gall midge research guide. Bouaké, Côte d'Ivoire: West Africa Rice Development Association. CAB International, Wallingford, United Kingdom. 28 p.

Wray, N. and Visscher, P. (2008). Estimating trait heritability. *Nature Education* 1:1-1..

Wolff, K., Schoen, E.D. and Peters-Van Rijn, J. (1993). Optimizing the generation of random amplified polymorphic DNA in chrysanthemum. *Theoretical and Applied Genetics* 86:1033-1037.

Wu, J.X., Liu, X.M., Zhang, S.Z., Zhu, Y.C., Whitworth, R.J. and Chen, M.S. (2008). Differential responses of wheat inhibitor-like genes to hessian fly, *Mayetiola destructor*, attacks during compatible and incompatible interactions. *Journal of Chemical Ecology* 34: 1005-1012.

Yu, J., Holland, J.B., McMullen, M.D. and Buckler, E.S. (2008). Genetic design and statistical power of nested association mapping in maize. *Genetics* 178:539-551.

Zarember, K.A., Cruz, A.R., Huang, C-Y. and Gallin, J.I. (2009). Antifungal activities of natural and synthetic iron chelators alone and in combination with azole and polyene antibiotics against *Aspergillus fumigatus*. *Antimicrobial Agents and Chemotherapy* 53: 2654-2656.

Zeng, Z.-B. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. (1993). *Proceeding of the National Academy of Science, USA.* 90:10972–10977.

Zeng, Z-B. (1994). Precision mapping of quantitative trait loci. *Genetics* 136**:**1457-1468.

Zheng, W., Wang, L., Meng, L. and Liu, J. (2008). Genetic variation in the endangered *Anisodus tanguticus* (Solanaceae), an alpine perennial endemic to the Qinghai-Tibetan Plateau. *Genetica* 132:123-129.

Chapter 2

Novel comparative methods for assessment of damage caused by African rice gall midge (AfRGM) under controlled conditions.

Abstract

Gene mapping relies on good phenotyping and subsequently on an efficient assessment of the damage, mainly when dealing with genes conferring resistance to insect pests. To develop an accurate method of assessing AfRGM damage during phenotyping, studies were initiated during the dry 2009 and wet 2010 seasons at the research station of the International institute of tropical agriculture (IITA)-Ibadan, to compare four methods of assessment of damage caused by AfRGM in three different plot sizes. The four methods included the International rice research institute (IRRI)'s Standard Evaluation System (SES) for rice and three methods based on resistance index (RI) assessment. Each plot contained 50, 100 and 150 test entries for the Small (24.4 m^2), the Medium (48.8m^2) and the Large (73.2 m^2) plot sizes, respectively. The RI-based assessments (RI-BA) were named RI-BA1, RI-BA2 and RI-BA3 and differed in the computing of the percentage of tillers with galls on the resistant check variety. The RI-BA methods provided better evaluations of AfRGM damage than the SES method, regardless of the trial size. Within RI-BA methods, comparison based on the known reaction of ITA306 and TOS14519 (considered as control entries) against AfRGM showed that RI-BA2 was more accurate than RI-BA1 and RI-BA3 when the plot is large because it is the only assessment method which provides the reactions of the control entries. No significant difference was observed between the two novel methods, RI-BA2 and RI-BA3, when the plot size was small. However, RI-BA2 and RI-BA3 provided better resistance estimates than RI-BA1.When the plot was of medium size, RI-BA2 was significantly better than RI-BA3, and RI-BA3 was significantly better than RI-BA1. The best method of assessing AfRGM damage was found to be RI-BA2, regardless of the plot size.

Introduction

With the development of molecular technologies, gene/Quantitative Trait Loci (QTL) mapping and tagging have become routine in many plant breeding laboratories. However, gene/QTL

mapping has some requirements that must be met for a good mapping resolution. Genetic mapping requires that the researcher:

(1) Develops appropriate mapping population (Burr *et al*., 1988; He *et al*., 2001; Doerge, 2002);

(2) Decides on the sample size (a minimum of 200 individuals are required but the higher the number of individuals, the more precise is the mapping) (Ferreira *et al*., 2006) and the type of molecular marker(s) for genotyping the mapping population (Semagn *et al.*, 2006a);

(3) Genotype the mapping population (parents and progenies) after screening parents for marker polymorphism and then;

(4) Perform linkage analyses (calculate pair-wise recombination frequencies between markers, establish linkage groups, estimate map distances, and determine map order) using statistical programmes (Semagn *et al.*, 2006b).

Mapping resolution also depends on the number of recombinants present in the mapping population (Jeuken *et al*., 2001; Sharma *et al*., 2002; Crane and Crane, 2005) and on the effect size of the QTL (Tanksley *et al.*, 1995). The latter is frequently small, accounting for less than 10% of the phenotypic variance for the majority of behavioural traits, so that high resolution mapping of such QTL requires many thousands of recombinants, meaning a larger mapping population (Tanksley *et al.*, 1995, Valdar *et al*., 2003). Dealing with a large population also implies using appropriate experimental designs, such as an augmented design with randomised incomplete block to contain the whole population and deal with spatial variability (Stroup, 2002). Augmented randomised incomplete block design is attractive for large population size management as it is space efficient due to non-replication of test entries. However, one of its major limitations is that each entry is evaluated on a single plot basis and is therefore prone to experimental error. Moreover, resistance to insect pests is usually quantitative, and, differential reactions are observed between resistant and moderately resistant varieties. In such a case, evaluating the test entries according to a reference entry (resistant or susceptible check) reaction is the best option of assessing the level of resistance or susceptibility. In assessment of damage caused by AfRGM, the Resistance Index (RI) has been proven to be more accurate than the Standard Evaluation System (SES) for rice as used by IRRI since 1986. In particular, RI takes environmental impact into account and mainly states the tested entries' reaction relative to the check entries (Nwilene *et al*., 2002). However, calculation of the RI depends on the percentage of tiller infestation on a resistant check entry, which itself is subject to variation

across the trial and is therefore a source of result bias. Two approaches can be adopted to minimise this bias. The first consists of a controlled environment, implying the same availability of nutrient amount for each test entry. This can be achieved under greenhouse conditions. It also requires an equal population of midge throughout the trial. Apart from individual placement of larvae onto plants, these experimental conditions are impossible to realise when dealing with hundreds of test entries. The second solution is to assess AfRGM damage on the test entries with reference to a check entry grown in their vicinity, or by considering two checks that delimit an area in which the test entries are grown, in order to calculate the percentage tiller infestation of the check entry. This aspect of calculating the RI is the object of the present study. Four methods of AfRGM damage assessment, including the IRRI's SES for rice scores, and three methods based on RI assessment, were compared to record the accuracy of each method according to plot size.

2-1 Materials and method

2-1-1 Plant material

The plant material consisted of 649 F_3 families derived from a cross between ITA306 (a popular cultivar released by the IITA) and TOS14519 (an *Oryza sativa* subsp. *japonica* landrace from the Gambia). ITA306 and TOS14519 are susceptible and resistant to the AfRGM, respectively.

2-1-2 Development of the screening population

To efficiently assess damage caused by AfRGM, data records were made on 20 hills of each of the F_3 families. Validated F_1 plants were advanced by self-pollination to the F_2 in greenhouses. In order to break dormancy, harvested seeds from individual F_1 plants were first kept in the incubator for 14 days at 40^{o}C, then some seeds were dehulled manually before seeding. Naked F_1 seeds were first pre-germinated in petri dishes or *in vitro*. Three-week old seedlings were transplanted into 12-litre plastic buckets 3/4-filled with paddy soil, at the rate of one seedling per bucket. Daily watering was done until plant maturity.

Six hundred and forty nine F_2 seeds from one single F_2 plant were advanced by self-pollination to generate 649 F_3 families. One gram, each of N, P_2O_5 and K_2O was added to each pot at transplanting, using NPK 15-15-15 as source. The pots were hand-weeded whenever

necessary. At maturity, when each panicle ripened, F_2 individual plants were harvested separately and numbered for evaluation as F_3 families.

2-1-3 Phenotyping

2-1-3-1 Insect rearing

AfRGM insect culture was maintained on plants of the susceptible variety, ITA306. Planting was done in seed boxes and was timed to coincide the plant age for infestation with emergence of adult midges from the culture plants. Culture plants of five weeks bearing three-week old adult midges were transplanted to the screenhouse just prior to the transplanting of test entry.

2-1-3-2 Paddy greenhouse screening

The experiment was conducted between January and August 2010 at AfricaRice-Ibadan, hosted by the IITA-Ibadan station. Seeds for the entries were sown on a nursery bed. The F_3 families were transplanted, 14 days after seeding (DAS) in a paddy screenhouse, in a 2-m row with a space of 0.20 m within and between rows, with the two parents involved in the cross as a check after every 10 entries. The experimental plot was flanked by an infestation band comprising five rows of ITA306, also used as susceptible check. The equivalent of 40 kg each of N, P_2O_5 and K_2O per ha was applied at transplanting using NPK 15-15-15. Another 40 kg N/ha was applied at 21 days after transplanting, using urea. The plots were hand-weeded at 21 and 40 days after transplanting. The experimental design was an augmented design with 2 checks, replicated every ten lines.

2-1-4 Performing the phenotyping evaluation

Considering the large number of individuals in the F_3 families (649 excluding the checks), the uncertainty of guaranteeing an even attack of the midge on each individual plant and the difficulty of producing enough larvae to initiate larval infestation, resistant individuals were involved in a second screening with favourable pest-occurrence conditions to confirm the initial phenotyping and the level of resistance. This experiment was slightly modified from the first. The modifications included: (1) the infestation band was densely transplanted, (2) the experiment was conducted from April to July 2010 which is the rainy and hot period, and (3) the reduction of the distance between rows from 0.20 m to 0.18 m to increase the density of the planting.

2-1-5 Assessment of damage caused by AfRGM

Damage caused by AfRGM was assessed according to four different methods including the SES for rice and three Resistance Index-Based Assessments (RI-BA) namely RI-BA1, RI-BA2 and RI-BA3. Damage assessment on both F_2 populations and F_3 families was made for only the RI-BA methods. Because of AfRGM assessment on 20 hills by SES, only the F_3 generation which has many seeds per family was involved.

2-1-5-1 Damage assessment according to IRRI's Standard Evaluation System (SES)

Tiller damage levels were usually expressed as scores between the values of 0 and 9, according to the SES for rice (IRRI, 1986) as shown below (Table 2-1). Galls were counted on all the 20 hills in each row, 45 and 70 days after transplanting. Percentage tiller infestation was computed using the following formula:

$$\text{Infested tiller percentage} = \frac{\text{Number of infested tillers}}{\text{Total number of tillers}} \times 100$$

Table 2-1: Assessment of damage by AfRGM according to IRRI's SES (IRRI, 1986)

Score	Percentage tiller damage	Rating (reaction)
0	No damage	Highly resistant or immune
1	Less than 1%	Resistant
3	1-5%	Moderately resistant
5	6-10%	Moderately susceptible
7	11-25%	Susceptible
9	Above 25%	Highly susceptible

2-1-5-2 Damage assessment according to RI-BA1

The SES does not take into account the impact of the environment, and it does not adjust the test entries' reaction relative to the check. Nwilene *et al*. (2002) proposed, therefore, to assess damage using an RI that provides a more sensitive measure of resistance because, unlike the SES score, it takes into account variations in AfRGM population pressures in different trials (Williams, 1998). The RI formula is as follows:

$$\text{RI-BA1} = \frac{\text{\% of tillers with galls on test entry}}{\text{\% of tillers with galls on resistant check variety}} \times 100$$

Where % of tillers with gall on resistant check variety is the mean of tillers with gall on resistant check across the whole trial as a percentage.

2-1-5-3 Damage assessment according to RI-BA2 and RI-BA3

The RI-BA1 method of AfRGM damage assessment described above did not take into account plot scale variations in AfRGM population pressures within the overall trial. To cater for this, the RI was computed using an interval mean value or zone value of the resistant check variety, to compute the percentage of tillers with galls on resistant check variety in each block or zone instead of its mean value, across the whole trial. Interval mean value (IMV) is the mean value of two lines of the resistant check varieties separated by 10 rows of test entry lines. Zone values (ZV) are the percentage tiller infestations of the resistant check replicated after every ten lines of test entries. The Mean value (MV) of the resistant check variety is its percentage mean infestation across the experimental trial (Figure 2-1).

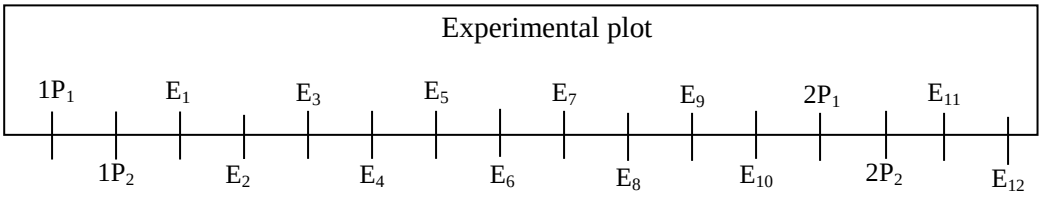


Figure 2-1: Front view of the schematic diagram of the experimental plot used for phenotyping

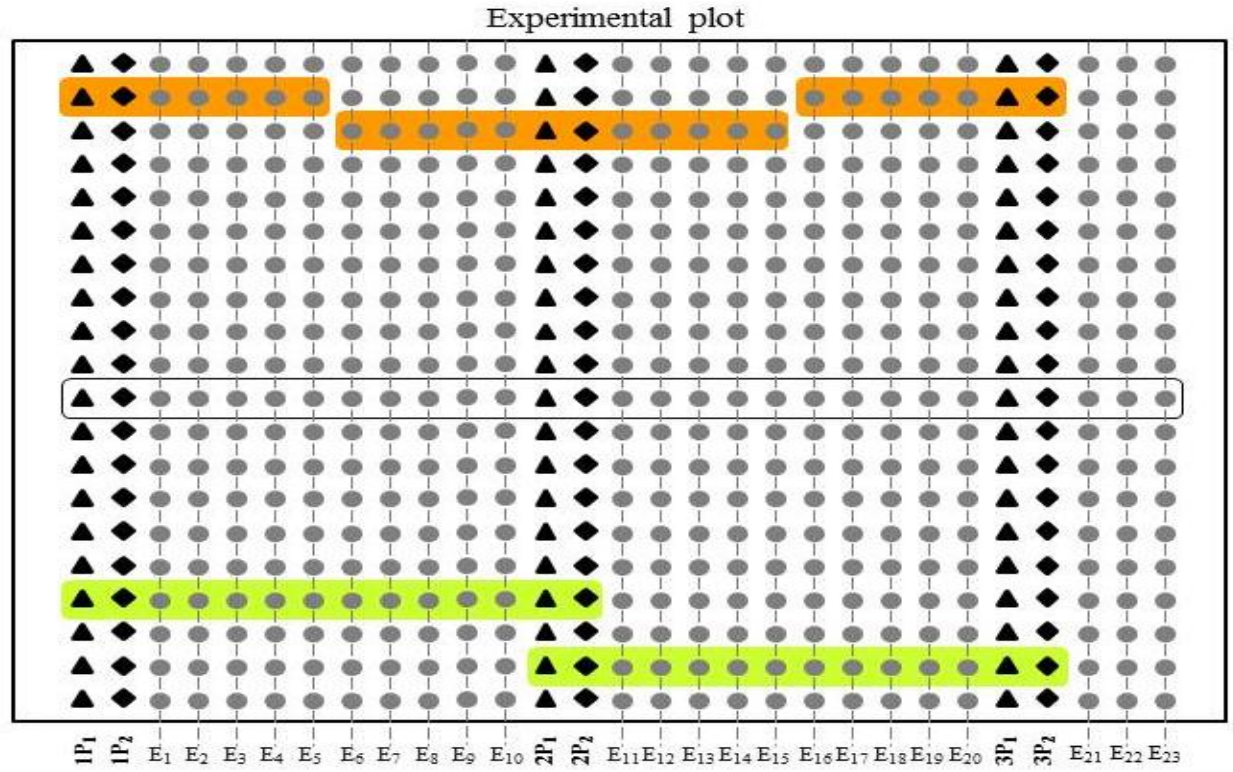


Figure 2-2: Top view of the schematic diagram of the experimental plot used for phenotyping

RI-BA1 is calculated for E_1-E_{23} using the mean value of $1P_1$, $2P_1$ and $3P_1$. RI-BA2 is calculated for E_1-E_5, E_6-E_{15} and E_{16}-E_{20} using the zone value of $1P_1$, $2P_1$ and $3P_1$ respectively. RI-BA3 is calculated for E_1-E_{10} and E_{11}-E_{20} using the interval mean value of $1P_1$-$2P_1$ and $2P_1$-$3P_1$ respectively.

Zone value = Individual PI of $1P_1$, $2P_1$, ……. and nP1

$$\text{Interval mean value} = \frac{\text{PI of } 1P_1 + \text{PI of } 2P_1}{2}$$

$$\text{Mean value} = \frac{\text{PI of } 1P_1 + \text{PI of } 2P_1 + \ldots\ldots + \text{PI of } nP_1}{n}$$

Where n = number of check entry, PI=Percent infestation, E=entry, P=parent

For each of the RI-BA, the index is 0 if there are no galls on the test entry, less than 1 if the test entry is more resistant than the resistant check, and more than 1 if the test entry is less resistant than the check (Williams, 1998).

2-1-6 Validation of the most accurate method

The principle of the validation process relies on the comparison of one variety (A) to itself. The null hypothesis Ho to test is then A=A. Conclusions such as A is equal to A or A is different from A will accept or reject Ho, respectively. The experiment was conducted based only on the assessment of varieties with known reactions against gall midge (resistance or susceptible). TOS14519 and ITA306 were used as tested entries and also as resistant and susceptible checks, respectively. The methods were compared based on whether they confirmed the assumption of the identical resistance or susceptibility of the entries. Three plot sizes were tested based on the space available in the tunnel paddy screenhouse. Each plot contained 50, 100 and 150 test entries (checks not included), respectively, for the Small (12.2m x 2m), Medium (24.4m x 2m) and Large (36.6m x 2m) plot sizes.

2-1-7 Statistical analyses

Data on percentage tiller and hill infestations were transformed prior to statistical analyses, to convert the data to a normal distribution:

(1) Data with a range of 30% to 70% were not transformed; (2) When all data were in the range between 0% and 50% or between 50% and 100%, the square root transformation was used; (3) For data that did not belong to any of the aforementioned ranges, the Arcsin transformation was used. RI and disease incidence were calculated using MS Excel 2009 and their values were adjusted to two decimal places (Gomez and Gomez (1984).

2-2 Results

2-2-1 Damage assessment according to SES method

For consistency with the three Resistance Index-Based assessment methods, SES ratings were converted either into resistant or susceptible varieties. The phenotypes highly resistant (HR), resistant (R) and moderately resistant (MR) were considered as resistant with reference to the resistant check TOS14519. The remaining phenotypes including moderately susceptible (MS), susceptible (S) and highly susceptible (HS) were designated as susceptible. Of the 649 entries tested, 72 were considered resistant by all the four methods considered altogether (Table 2-2). For SES, 62 (1HR, 6 R and 55 MR) entries were rated resistant with an infestation rate ranging

from 0% (Line 49) to 5.4% (Lines 604, 55, 35 and 570) with a mean percentage value of 3%. The disease incidence was, however, high with Lines 79, 582 and 604 with rate above 50%. The mean value for SES was 25.1% with extreme values of 0 and 63 % for Lines 49 and 604, respectively (Table 2-2 last column first part). Amongst the 62 classified as resistant by SES, 29, 27 and 27 lines were not resistant with the RI-BA1, RI-BA3 and RI-BA2 methods, respectively.

When considering the RI methods, 24 lines classified as resistant with the SES method were not resistant with any of the three RI methods, while 27 appeared resistant with all three RI methods. Twelve entries rated resistant either with one (RI-BA1 or RI-BA2 or RI-BA3) or two (RI-BA1 and RI-BA2 or RI-BA1 and RI-BA3 or RI-BA2 and RI-BA3) of the three RI methods (Table 2-3).

Apart from Line 49 which ranked first in all methods (Table 2-2, first line), the other entries considered resistant did not show the same ranking throughout the methods. Lines 484 and 546, that could be seen as two of the rare constant-ranking entries across methods, ranked 3rd, 4th, 5th and 6th respectively with SES, RI-BA2, RI-BA3 and RI-BA1 for the first (Table 2-3 line 4) and 7th, 5th, 3rd and 2nd respectively with SES, RI-BA2, RI-BA3 and RI-BA1 for the latter (Table 2-2 line 5).

Table 2-2: Distribution and ranking of 72 lines according to SES, RI-BA1, RI-BA2 and RI-BA3 methods of AfRGM damage assessment

Lines	TZV	TIMV	TMV	RI-BA1		RI-BA2		RI-BA3		SES		DI
				RI	Rk	RI	Rk	RI	Rk	PI	Rk	
49	**9.46**	**8.95**	**6.32**	**0.00**	**1**	**0.00**	**1**	**0.00**	**1**	**0.00**	**1**	**0.0**
234	10.06	9.77	6.32	0.21	3	0.13	2	0.13	2	1.31	12	15.0
565	2.74	1.67	1.79	0.40	10	0.26	3	0.42	13	0.71	5	5.0
484	**1.99**	**2.37**	**1.79**	**0.33**	**6**	**0.30**	**4**	**0.25**	**5**	**0.59**	**3**	**5.0**
546	**2.72**	**4.62**	**6.32**	**0.13**	**2**	**0.31**	**5**	**0.18**	**3**	**0.85**	**7**	**10.0**
181	3.57	2.85	1.79	0.62	19	0.31	6	0.39	10	1.12	10	10.0
204	2.13	2.85	1.79	0.39	9	0.33	7	0.24	4	0.69	4	5.6
331	13.41	12.84	6.32	0.69	22	0.33	8	0.34	9	4.39	48	25.0
464	9.48	10.87	6.32	0.50	14	0.33	9	0.29	6	3.15	32	15.0
87	9.46	7.25	6.32	0.57	17	0.38	10	0.49	18	3.57	40	30.0
443	13.41	10.19	6.32	0.82	29	0.39	11	0.51	19	5.19	57	30.0
412	6.52	4.62	6.32	0.40	11	0.39	12	0.55	21	2.55	22	25.0
613	7.43	6.23	6.32	0.52	15	0.44	13	0.53	20	3.29	36	25.0
606	**2.74**	**1.67**	**1.79**	**0.70**	**24**	**0.46**	**14**	**0.75**	**27**	**1.25**	**11**	**5.0**
381	1.68	1.87	1.61	0.48	12	0.46	15	0.42	11	0.78	6	10.0
34	6.97	7.70	6.32	0.55	16	0.50	16	0.45	14	3.45	39	25.0
587	2.72	4.62	6.32	0.24	4	0.56	17	0.33	7	1.51	14	15.0
527	1.99	2.37	1.79	0.62	20	0.56	18	0.47	16	1.12	9	5.0
192	5.04	7.25	6.32	0.48	13	0.61	19	0.42	12	3.06	28	25.0
557	2.74	2.37	1.79	1.01	38	0.66	20	0.76	29	1.81	16	15.8
120	3.57	2.1	1.79	1.42	-	0.71	21	1.21	-	2.55	23	25.0
83	6.97	7.70	6.32	0.80	27	0.73	22	0.66	24	5.05	54	25.0
150	10.06	9.77	6.32	1.16	-	0.73	23	0.75	28	7.32	-	35.0
585	2.74	1.67	1.79	1.13	-	0.74	24	1.21	-	2.03	18	15.8
107	3.57	2.1	1.79	1.47	-	0.74	25	1.26	-	2.64	24	20.0
298	5.56	7.00	6.32	0.65	21	0.74	26	0.59	22	4.14	44	25.0
55	6.97	7.70	6.32	0.85	30	0.77	27	0.70	25	5.38	60	42.1
567	2.72	4.62	6.32	0.34	7	0.78	28	0.46	15	2.12	19	30.0
18	6.97	7.70	6.32	0.90	33	0.81	29	0.74	26	5.66	-	40.0
578	2.72	4.62	6.32	0.36	8	0.83	30	0.49	17	2.27	20	25.0
366	6.52	4.62	6.32	0.98	35	0.95	31	1.35	-	6.23	-	50.0
647	0.59	1.67	1.79	0.32	5	0.96	32	0.34	8	0.57	2	5.0
35	5.56	5.68	6.32	0.85	31	0.97	33	0.95	33	5.38	61	45.0
601	2.74	1.67	1.79	1.57	-	1.03	34	1.68	-	2.81	26	15.0
605	7.43	6.23	6.32	1.22	-	1.03	35	1.23	-	7.69	-	40.0
286	2.13	2.85	1.79	1.33	-	1.12		0.84	30	2.39	21	33.3

Lines	TZV	TIMV	TMV	RI-BA1		RI-BA2		RI-BA3		SES		DI
				RI	Rk	RI	Rk	RI	Rk	PI	Rk	
593	2.74	1.67	1.79	1.74	-	1.14	-	1.87	-	3.13	31	15.0
183	3.57	2.85	1.79	2.38	-	1.20	-	1.50	-	4.28	46	31.6
125	3.57	2.1	1.79	2.38	-	1.20	-	2.04	-	4.28	47	45.0
328	2.13	2.06	1.79	1.53	-	1.29	-	1.33	-	2.75	25	20.0
305	4.84	4.45	6.32	1.00	37	1.31	-	1.42	-	6.33	-	26.3
191	5.04	7.25	6.32	1.11	-	1.40	-	0.97	35	7.03	-	40.0
535	2.74	2.37	1.79	2.13	-	1.40	-	1.62	-	3.83	42	20.0
442	1.99	2.06	1.79	1.70	-	1.54	-	1.48	-	3.06	29	25.0
138	2.94	1.69	1.61	2.84	-	1.55	-	2.70	-	4.57	51	46.2
543	2.72	4.62	6.32	0.70	23	1.62	-	0.95	34	4.40	49	40.0
407	1.99	2.06	1.79	1.81	-	1.64	-	1.58	-	3.25	34	25.0
447	1.99	2.06	1.79	1.83	-	1.65	-	1.59	-	3.28	35	45.0
582	2.72	4.62	6.32	0.75	25	1.75	-	1.03	36	4.75	52	55.0
299	2.13	2.06	1.79	2.13	-	1.80	-	1.86	-	3.83	43	35.0
632	0.59	1.67	1.79	0.61	18	1.85	-	0.65	23	1.09	8	10.0
599	2.72	4.62	6.32	0.80	28	1.86	-	1.10	-	5.07	55	50.0
411	1.68	0.84	1.61	2.01	-	1.93	-	3.86	-	3.24	33	45.0
570	2.74	2.37	1.79	3.00	-	1.97	-	2.28	-	5.39	62	30.0
418	1.68	0.84	1.61	2.10	-	2.02	-	4.04	-	3.38	38	35.0
598	2.72	4.62	6.32	0.89	32	2.06	-	1.21	-	5.60	-	45.0
503	1.99	2.37	1.79	2.48	-	2.24	-	1.88	-	4.46	50	15.0
517	2.84	2.84	6.32	1.03	40	2.29	-	2.29	-	6.49	-	10.0
228	2.63	2.63	6.32	0.96	34	2.30	-	2.30	-	6.06	-	20.0
637	0.59	1.67	1.79	0.80	26	2.41	-	0.86	32	1.43	13	15.0
604	2.14	1.07	1.61	3.34	-	2.51	-	5.03	-	5.37	59	63.6
372	1.99	2.06	1.79	2.80	-	2.53	-	2.44	-	5.03	53	20.0
408	2.44	4.06	6.32	1.03	39	2.66	-	1.60	-	6.48	-	35.0
82	0.63	2.1	1.79	1.00	36	2.87	-	0.85	31	1.79	15	20.0
648	0.59	1.67	1.79	1.06	-	3.23	-	1.14	-	1.91	17	15.0
28	0.63	2.1	1.79	1.73	-	4.98	-	1.48	-	3.11	30	26.3
631	0.59	1.67	1.79	1.86	-	5.63	-	2.00	-	3.33	37	25.0
109	0.63	2.1	1.79	2.12	-	6.10	-	1.81	-	3.81	41	33.3
74	0.45	1.69	1.61	1.81	-	6.49	-	1.73	-	2.92	27	30.0
643	0.59	1.67	1.79	2.38	-	7.21	-	2.56	-	4.27	45	33.3
634	0.59	1.67	1.79	2.91	-	8.84	-	3.13	-	5.23	58	25.0
79	0.45	1.69	1.61	3.20	-	11.45	-	3.05	-	5.16	56	55.0

TZV=TOS zone value, TIMV=TOS interval mean value, TMV=TOS mean value, RI=Resistance index, Rk=Rank, PI=Percent infestation, DI=Disease incidence, - means the entry is susceptible

Table 2-3: SES method resistant lines distribution according to various combinations of the other methods of damage assessment

	Non-resistant with						**Resistant with**										**Resistant with**						**Non-resistant with**	
	RI-BA1		**RI-BA2**		**RI-BA3**		**SES**				**RI-BA1**		**RI-BA2**		**RI-BA3**		**RI-BA13**		**RI-BA23**		**RI-BA123**		RI-BA123	
Resistant with SES and …	648	418	125	28	648	109	28	192	464	601	34	484	34	443	34	464	34	443	34	443	34	443	74	593
	585	109	138	74	585	535	34	204	484	604	35	527	35	464	35	484	35	464	35	464	35	464	138	503
	286	535	183	79	120	299	35	234	503	606	49	543	49	484	49	527	49	484	49	484	49	484	79	631
	120	299	286	82	107	643	49	286	527	613	55	546	55	527	55	543	55	527	55	527	55	527	411	125
	107	643	299	109	328	183	55	298	535	631	82	557	83	546	82	546	83	546	83	546	83	546	418	570
	328	183	328	372	601	125	74	299	543	632	83	565	87	557	83	557	87	557	87	557	87	557	604	372
	601	125	407	418	74	503	79	328	546	634	87	567	107	565	87	565	181	565	181	565	181	565	648	643
	74	503	411	503	442	138	82	331	557	637	181	578	120	567	181	567	192	567	192	567	192	567	328	634
	442	138	442	570	28	372	83	372	565	643	192	582	181	578	192	578	204	578	204	578	204	578	28	
	28	372	447	604	593	599	87	381	567	647	204	587	192	585	204	582	234	587	234	587	234	587	442	
	593	79	535	631	411	79	107	407	570	648	234	599	204	587	234	587	298	606	298	606	298	606	183	
	411	634	543	634	407	634	109	411	578		298	606	234	601	286	606	331	613	331	613	331	613	407	
	407	604	582	637	447	604	120	412	582		331	613	298	606	298	613	381	647	381	647	381	647	447	
	447	570	593	643	631	570	125	418	585		381	632	331	613	331	632	412	543	412		412		535	
	631		599	648	418		138	442	587		412	637	381	647	381	637							109	
			632				181	443	593		443	647	412		412	647							299	
							183	447	599		464				443									
Total	**29**		**31**		**29**		**62**				**33**		**31**		**33**		**28**		**27**		**27**		**24**	

2-2-2 Damage assessment according to RI-BA1

Table 2-4 depicted 40 lines, accounting for 6.16% of the total of 649 test entries, to be Resistant with the RI-BA1 assessment. Among them eight, seven and eleven entries were not rated Resistant by the SES, RI-BA3 and RI-BA2 methods, respectively. In addition, six (Lines 39, 228, 305, 408, 517 and 598) of them were rated Resistant only with the RI-BA1 method while 27 were constantly rated Resistant by the three other methods considered together. Lines 599 and 366, rated by RI-BA1 as Resistant, were not considered to be Resistant by both RI-BA2 and RI-BA3 and by both SES and RI-BA3.

Table 2-4: RI-BA1 method: distribution of resistant and susceptible lines relative to the other methods of evaluation.

	Susceptible lines with			**Resistant lines with**											
Resistant with RI-BA1 and …	**SES**	**RI-BA2**	**RI-BA3**	**RI-BA1**				**RI-BA2**				**RI-BA3**			
	18			49	412	298	35	234	464	587	567	546	565	181	83
	191	82	228												
	228	228	305	546	381	331	598	565	87	527	18	234	412	527	443
	305	305	366	234	192	543	18	484	443	192	578	587	381	298	55
	366	408	517	587	464	606	228	546	412	557	366	647	192	331	35
	408	598	598	647	613	582	366	181	613	83	647	484	464	543	18
	598	517	599	484	34	637	82	204	606	298	35	567	613	606	82
	517	543	408	567	87	83	305	331	381	55		578	34	582	557
		582		578	632	599	557					204	87	637	
		599		204	181	443	408								
		632		565	527	55	517								
		637													
Total	**8**	**11**	**7**	**40**				**27**				**33**			

2-2-3 Damage assessment according to RI-BA2

Thirty five entries, accounting for 5.39 % of the test entries, rated Resistant according to RI-BA2. Six (17.14%), 6 (17.14%) and 4 (11.43%) of them did not rate Resistant with RI-BA1, RI-BA2 and SES respectively, while 27 were also depicted as Resistant by all the other methods. Each of the resistant lines was either rated as Resistant by one or two, if not by all the other methods of assessment of AfRGM damage (Table 2-5).

Table 2-5: RI-BA2 method: distribution of resistant and susceptible lines relative to the other methods of evaluation.

	Susceptible with			Resistant with									
	SES	RI-BA1	RI-BA3	RI-BA2						RI-BA3			
Resistant with RI-BA2 and …..	18	150	585	49	331	606	120	55	647	234	331	578	55
	366	585	107	234	464	381	83	567	35	546	181	87	18
	605	107	366	565	87	34	150	18	601	204	381	443	606
	150	601	601	484	443	587	585	578	605	484	192	613	150
		605	605	546	412	527	107	366		464	565	412	557
		120	120	181	613	192	298			587	34	34	35
				204		557				647	567	567	
Total	**4**	**6**	**6**	**35**						**27**			

2-2-4 Damage assessment according to RI-BA3

Table 2-6 shows that 36 lines, representing 5.55% of the total test entries, rated Resistant according to RI-BA3. However, 8.33% (3 lines), 19.44% (7 lines) and 8.33% (3 lines) of them did not rate Resistant with RI-BA1, RI-BA2 and SES, respectively, while 27 were consistently assessed as Resistant by all the three other methods. Like RI-BA2, each of the resistant lines was either rated as resistant by one or two, if not by all the other methods of damage assessment.

Table 2-6: RI-BA3 method resistant lines distribution of resistant and susceptible lines relative to the other methods of evaluation.

	Susceptible with			Resistant with					
	SES	RI-BA1	RI-BA2	RI-BA3					
Resistant with RI-BA3 and…	18	150	286	49	587	34	613	606	35
	150	191	82	234	647	567	412	150	543
	191	286	637	546	331	527	298	557	191
			543	204	181	578	632	286	582
			191	484	381	87	83	82	565
			582 632	464	192	443	55	637	18
Total	**3**	**3**	**7**	**36**					

2-2-5 Comparison of AfRGM damage assessment methods according to plot size

2-2-5-1 Comparison of RI-BA1, RI-BA2 and RI-BA3 methods using small plot size

For small plot size, RI-BA2 and RI-BA3 showed no difference for resistance checking. Of the six tested entries, only one rated differently from TOS14519. However, both mentioned methods were different from RI-BA1, which rated two entries either identical to TOS14519, less resistant, or more resistant than TOS14519 (Table 2-7).

Table 2-7: Comparison of the RI-BA methods of resistance evaluation, when using a small plot size

Entry N^{o}	TOS14519				RI-BA2		RI-BA3		RI-BA1	
	PI	ZV	IMV	MV	RI	Rating	RI	Rating	RI	Rating
10.1	4.84	4.84	4.4	4.97	1	IT	1	IT	1	IT
20.1	4.05	4.05	2.0	4.97	1	IT	2	LRT	1	IT
30.1	8.43	8.43	7.1	4.97	1	IT	1	IT	2	LRT
40.1	1.41	1.41	1.5	4.97	1	IT	1	IT	0	MRT
50.1	1.64	1.64	1.4	4.97	1	IT	1	IT	0	MRT
60.1	9.46	9.46	7.2	4.97	1	IT	1	IT	2	LRT

ZV = Zone value, IMV = Interval mean value, MV=Mean value, RI = Resistance index, PI = Percent infestation, IT = Identical to TOS14519, LRT = Less resistant than TOS14519, MRT = More resistant than TOS14519

Table 2-8 shows the same trend with ITA306 for susceptibility evaluation, with the exception that RI-BA1 rated only one entry equivalent to ITA306. Among the remaining five tested entries, four were rated more susceptible and one, more resistant than ITA306. Apart from one entry, which was rated more resistant than ITA306 by RI-BA3, the others were rated identical to ITA306 by RI-BA3 and RI-BA2.

Table 2-8: Comparison of the RI-BA methods of susceptibility evaluation when using a small plot size

Entry No	ITA306				RI-BA3		RI-BA1		RI-BA2	
	PI	ZV	IMV	MV	SI	Rating	SI	Rating	SI	Rating
10.1	43.9	43.9	47.5	54.3	0.8	MRI	0.9	MRI	1	II
20.1	57.3	57.3	59.8	54.5	1.0	II	1.0	MSI	1	II
30.1	62.2	62.2	59.6	54.5	1.0	II	1.1	MSI	1	II
40.1	57.0	57.0	58.0	54.5	1.0	II	1.0	II	1	II
50.1	59.1	59.1	59.6	54.5	1.0	II	1.1	MSI	1	II
60.1	60.1	60.1	59.6	54.5	1.0	II	1.1	MSI	1	II

PI = Percent infestation, SI = Susceptibility index, II = Identical to ITA306, MSI = More susceptible than ITA306, MRI = More resistant than ITA306

2-2-5-2 Comparison of RI-BA1, RI-BA2 and RI-BA3 methods using a medium plot size

For resistance evaluation, when entries were laid in a medium sized plot, a clear difference was noticed between the three RI-BA methods. Only RI-BA2 rated the tested entries consistently identical to TOS14519. Both RI-BA1 and RI-BA3 only rated two entries as similar to TOS14519. However, although six and five entries were identical to TOS14519 respectively with RI-BA3 and RI-BA1, the entries were not rated the same, and therefore distinguished the methods. For RI-BA3, three and two entries were rated less resistant and more resistant than TOS14519, respectively, while RI-BA1 rated two and four entries less resistant and more resistant than TOS14519 (Table 2-9).

Table 2-9: Comparison of the RI-BA methods of resistance evaluation, when using a medium plot size

Entry N°	TOS14519				RI-BA2		RI-BA3		RI-BA1	
	PI	ZV	IMV	MV	RI	Rating	RI	Rating	RI	Rating
10.1	4.84	4.84	4.4	3.9	1.0	IT	1	IT	1	IT
20.1	4.05	4.05	2.0	3.9	1.0	IT	2	LRT	1	IT
30.1	8.43	8.43	7.1	3.9	1.0	IT	1	IT	2	LRT
40.1	1.41	1.41	1.5	3.9	1.0	IT	1	IT	0	MRT
50.1	1.64	1.64	1.4	3.9	1.0	IT	1	IT	0	MRT
60.1	9.46	9.46	7.2	3.9	1.0	IT	1	IT	2	LRT
70.1	0.01	0.01	2.8	3.9	1.0	IT	0	MRT	0	MRT
80.1	5.68	5.68	4.1	3.9	1.0	IT	1	IT	1	IT
90.1	2.44	2.44	1.2	3.9	1.0	IT	2	LRT	1	IT
100.1	0.01	0.01	2.6	3.9	1.0	IT	0	MRT	0	MRT
110.1	5.26	5.26	2.6	3.9	1.0	IT	2	LRT	1	IT

ZV = Zone value, IMV = Interval mean value, MV=Mean value, RI = Resistance index, PI = Percent infestation, IT = Identical to TOS14519, LRT = Less resistant than TOS14519, MRT = More resistant than TOS14519

In Table 2-10, susceptibility evaluation by the RI-BA2 method rated all the eleven tested entries as identical to ITA306 and conversely, RI-BA1 and RI-BA3 rated only four and seven entries to be the same as ITA306, respectively, confirming the superior accuracy of the RI-BA2 method over RI-BA3 and RI-BA3 over RI-BA1. With RI-BA1, three and four entries were rated as more resistant and more susceptible than ITA306, respectively, while RI-BA3 evaluated three entries as more resistant and one entry as more susceptible. Three entries were rated by both RI-BA1 and RI-BA3 as identical to ITA306, while three others were rated as more resistant than ITA306, by both methods.

Table 2-10: Comparison of the RI-BA methods of susceptibility evaluation when using a medium plot size

Entry N°	ITA306				RI-BA1		RI-BA2		RI-BA3	
	PI	ZV	IMV	MV	SI	Rating	SI	Rating	SI	Rating
10.1	43.94	43.9	47.5	54.5	0.9	MRI	1.0	II	0.8	MRI
20.1	57.26	57.3	59.8	54.5	1.0	MSI	1.0	II	1.0	II
30.1	62.25	62.2	59.6	54.5	1.1	MSI	1.0	II	1.0	II
40.1	56.96	57.0	58.0	54.5	1.0	II	1.0	II	1.0	II
50.1	59.05	59.1	59.6	54.5	1.1	MSI	1.0	II	1.0	II
60.1	60.06	60.1	59.6	54.5	1.1	MSI	1.0	II	1.0	II
70.1	43.94	43.9	47.5	54.5	0.8	MRI	1.0	II	0.9	MRI
80.1	51.06	51.1	44.7	54.5	0.9	II	1.0	II	1.1	MSI
90.1	38.26	38.3	46.7	54.5	0.7	MRI	1.0	II	0.8	MRI
100.1	55.24	55.2	54.9	54.5	1.0	II	1.0	II	1.0	II
110.1	54.64	54.6	55.7	54.5	1.0	II	1.0	II	1.0	II

PI = Percent infestation, SI = Susceptibility index, II = Identical to ITA306, MSI = More susceptible than ITA306, MRI = More resistant than ITA306

2-2-5-3 Comparison of RI-BA1, RI-BA2 and RI-BA3 methods when using a large plot size.

In Table 2-11 similar outcomes appear, as those observed for the small and the medium plot results. RI-BA2 again ranked the test entries consistently as similar to TOS14519 or ITA306, respectively, for resistance and susceptibility. However, with resistance evaluation, the differences between methods were more pronounced. Of the 34 entries that rated identical to TOS14519, 12 and 18 were rated identical to TOS14519 by RI-BA1 and RI-BA3, respectively. However, when they were considered together, RI-BA1 and RI-BA3 methods rated only 6 entries as identical to TOS14519. With RI-BA1, 13 entries were designated as more resistant than TOS14519 while 9 others as less resistant than TOS14519. With the RI-BA3 method, 8 entries were designated less resistant and 7 were designated more resistant than TOS14519.

Table 2-11: Comparison of the RI-BA methods of resistance evaluation, when using a large plot size

Entry N°	TOS14519				RI-BA1		RI-BA2		RI-BA3	
	PI	ZV	IMV	MV	RI	Rating	RI	Rating	RI	Rating
10.1	4.84	4.84	4.4	4.5	1.1	IT	1.0	IT	1	IT
20.1	4.05	4.05	2.0	4.5	0.9	IT	1.0	IT	2	LRT
30.1	8.43	8.43	7.1	4.5	1.9	LRT	1.0	IT	1	IT
40.1	1.41	1.41	1.5	4.5	0.3	MRT	1.0	IT	1	IT
50.1	1.64	1.64	1.4	4.5	0.4	MRT	1.0	IT	1	IT
60.1	9.46	9.46	7.2	4.5	2.1	LRT	1.0	IT	1	IT
70.1	0.01	0.01	2.8	4.5	0.0	MRT	1.0	IT	0	MRT
80.1	5.68	5.68	4.1	4.5	1.3	IT	1.0	IT	1	IT
90.1	2.44	2.44	1.2	4.5	0.5	IT	1.0	IT	2	LRT
100.1	0.01	0.01	2.6	4.5	0.0	MRT	1.0	IT	0	MRT
110.1	5.26	5.26	2.6	4.5	1.2	IT	1.0	IT	2	LRT
120.1	0.01	0.01	2.1	4.5	0.0	MRT	1.0	IT	0	MRT
130.1	4.11	4.11	2.1	4.5	0.9	IT	1.0	IT	2	LRT
140.1	0.01	0.01	5.0	4.5	0.0	MRT	1.0	IT	0	MRT
150.1	10.06	10.06	9.8	4.5	2.2	LRT	1.0	IT	1	IT
160.1	9.48	9.48	10.9	4.5	2.1	LRT	1.0	IT	1	IT
170.1	12.26	12.26	12.8	4.5	2.7	LRT	1.0	IT	1	IT
180.1	13.41	13.41	10.2	4.5	3.0	LRT	1.0	IT	1	IT
190.1	6.97	6.97	7.7	4.5	1.6	LRT	1.0	IT	1	IT
200.1	0.01	0.01	0.7	4.5	0.0	MRT	1.0	IT	0	MRT
210.1	5.81	5.81	5.7	4.5	1.3	IT	1.0	IT	1	IT
220.1	5.56	5.56	7.0	4.5	1.2	IT	1.0	IT	1	IT
230.1	8.44	8.44	9.0	4.5	1.9	LRT	1.0	IT	1	IT
240.1	1.10	1.10	0.6	4.5	0.2	MRT	1.0	IT	2	LRT
250.1	5.04	5.04	6.2	4.5	1.1	IT	1.0	IT	1	IT
260.1	7.43	7.43	7.0	4.5	1.7	LRT	1.0	IT	1	IT
270.1	6.52	6.52	4.6	4.5	1.5	IT	1.0	IT	1	LRT
280.1	2.72	2.72	1.6	4.5	0.6	IT	1.0	IT	2	LRT
290.1	0.45	0.45	1.7	4.5	0.1	MRT	1.0	IT	0	MRT
300.1	2.94	2.94	2.5	4.5	0.7	IT	1.0	IT	1	IT
310.1	2.07	2.07	1.9	4.5	0.5	MRT	1.0	IT	1	IT
320.1	1.68	1.68	1.1	4.5	0.4	MRT	1.0	IT	2	LRT
330.1	0.54	0.54	1.3	4.5	0.1	MRT	1.0	IT	0	MRT
340.1	2.14	2.14	1.3	4.5	0.5	MRT	1.0	IT	2	LRT

ZV = Zone value, IMV = Interval mean value, MV=Mean value, RI = Resistance index, PI = Percent infestation, IT = Identical to TOS14519, LRT = Less resistant than TOS14519, MRT = More resistant than TOS14519

With susceptibility evaluation, among the 34 tested entries listed in Table 2-12 and ranked identical to ITA306 by RI-BA2, 9 and 15 entries were not rated identical to ITA306 by RI-BA1 and RI-BA3, respectively. Surprisingly, 9 entries located at the end of the trial (bold entries in Table 2-13) rated identical to ITA306 for all three methods. However, 15 of the 25 entries

remaining, were rated either more resistant (7 entries) or more susceptible (8 entries) than ITA306 by the three methods together. Of these 15 progenies, 5 and 1 entries were rated more resistant and more susceptible than ITA306 by RI-BA3 and RI-BA1, respectively. When all the three methods were considered together, 17 entries were rated as similar to ITA306.

2-2-6 Validation of the most accurate method among the RI-BA methods

For the validation, TOS14519 was used as a test entry and as the resistant check for resistance evaluation and ITA306 was used as a test entry and as the susceptible check for susceptibility evaluation. Both TOS14519 and ITA306 were ranked equally resistant and susceptible to the test entry only by RI-BA2, consistently across the plot sizes. The RI value consistently equalled 1, meaning that the test entry and the resistant check had exactly the same reaction, and therefore confirmed the Ho hypothesis, A=A. Only one type of reaction was observed with RI-BA2, whereas RI-BA1 and RI-BA3 displayed three types. RI-BA1 and RI-BA3 ranked TOS14519 and ITA306 either equally resistant/susceptible (RI=1), or less resistant/susceptible (RI>1) or even more resistant/susceptible (RI<1) than themselves, which could not have been possible.

Table 2-12: Comparison of the RI-BA methods by susceptibility evaluation, when using a large plot size

Entry N°	ITA306				RI-BA1		RI-BA2		RI-BA3	
	PI	ZV	IMV	MV	SI	Rating	SI	Rating	SI	Rating
10	43.94	43.9	47.5	54.3	0.8	MRI	1.0	II	0.9	MRI
20	57.26	57.3	59.8	54.5	1.0	II	1.0	II	1.0	II
30	62.25	62.2	59.6	54.5	1.0	II	1.0	II	1.1	LRI
40	56.96	57.0	58.0	54.5	1.0	II	1.0	II	1.0	II
50	59.05	59.1	59.6	54.5	1.0	II	1.0	II	1.1	LRI
60	60.06	60.1	59.6	54.5	1.0	II	1.0	II	1.1	LRI
70	43.94	43.9	47.5	54.5	0.9	MRI	1.0	II	0.8	MRI
80	51.06	51.1	44.7	54.5	1.1	LRI	1.0	II	0.9	MRI
90	38.26	38.3	46.7	54.5	0.8	MRI	1.0	II	0.7	MRI
100	55.24	55.2	54.9	54.5	1.0	II	1.0	II	1.0	II
110	54.64	54.6	55.7	54.5	1.0	II	1.0	II	1.0	II
120	56.73	56.7	54.4	54.5	1.0	II	1.0	II	1.0	II
130	52.07	52.1	50.6	54.5	1.0	II	1.0	II	1.0	II
140	49.17	49.2	54.8	54.5	0.9	MRI	1.0	II	0.9	MRI
150	60.48	60.5	59.4	54.5	1.0	II	1.0	II	1.1	LRI
160	58.27	58.3	58.0	54.5	1.0	II	1.0	II	1.1	LRI
170	57.78	57.8	55.6	54.5	1.0	II	1.0	II	1.1	LRI
180	53.47	53.5	53.6	54.5	1.0	II	1.0	II	1.0	II
190	53.68	53.7	53.8	54.5	1.0	II	1.0	II	1.0	II
200	53.91	53.9	58.8	54.5	0.9	MRI	1.0	II	1.0	II
210	63.64	63.6	60.1	54.5	1.1	LRI	1.0	II	1.2	LRI
220	56.48	56.5	54.5	54.5	1.0	II	1.0	II	1.0	LRI
230	52.59	52.6	53.3	54.5	**1.0**	**II**	**1.0**	**II**	**1.0**	**II**
240	54.04	54.0	53.5	54.5	**1.0**	**II**	**1.0**	**II**	**1.0**	**II**
250	52.92	52.9	53.4	54.5	**1.0**	**II**	**1.0**	**II**	**1.0**	**II**
260	53.85	53.8	53.7	54.5	**1.0**	**II**	**1.0**	**II**	**1.0**	**II**
270	53.48	53.5	54.2	54.5	**1.0**	**II**	**1.0**	**II**	**1.0**	**II**
280	54.94	54.9	54.9	54.5	**1.0**	**II**	**1.0**	**II**	**1.0**	**II**
290	54.77	54.8	54.5	54.5	**1.0**	**II**	**1.0**	**II**	**1.0**	**II**
300	54.22	54.2	52.2	54.5	**1.0**	**II**	**1.0**	**II**	**1.0**	**II**
310	50.21	50.2	52.2	54.5	1.0	II	1.0	II	0.9	MRI
320	54.21	54.2	55.3	54.5	**1.0**	**II**	**1.0**	**II**	**1.0**	**II**
330	56.31	56.3	49.4	54.5	1.1	LRI	1.0	II	1.0	II
340	42.47	42.5	54.1	54.5	0.8	MRI	1.0	II	0.8	MRI

PI = Percent infestation, SI = Susceptibility index, II = Identical to ITA306, LRI = Less resistant than ITA306, MRI= More resistant than ITA306, ZV=Zone value, IMV=Interval mean value, MV=Mean value

2-3 Discussion

Using four different resistant evaluation methods: SES, RI-BA1, RI-BA2 and RI-BA2, resulted in four discrete sets of results, which indicate that these methods were different from each other. When a test entry's AfRGM infestation is nil, all four methods ranked the test entry as resistant regardless of the plot size. However, the methods provide different results when dealing with lines that carry partial resistance. The SES method selected nearly twice the number of resistant lines compared to each of the RI-BA methods. However, 24 lines of these lines were not identified as resistant by any of the three other evaluation methods. Nwilene *et al*. (2002) have demonstrated the accuracy of the RI-BA method relative to the SES. The large number of resistant lines found using SES, when compared to the other methods, was due to false resistant lines as reported earlier by Nwilene (2002), because the damage assessment is made with reference to the susceptible check. In addition, the SES assessment does not take into account the impact of the environment and does not adjust for variable pest distribution. The result is that SES frequently misclassifies susceptible lines as resistant. The resistance evaluation of Lines 647 and 484 illustrates this explanation. Line 647, with an infestation of 0.57%, was ranked second by the SES method, but was ranked 32nd out of 35 using RI-BA2. Line 484 with the same level of infestation (0.59%) ranked nearly the same using both the SES (third) and RI-BA2 (fourth) methods. The relative infestation of the resistant check entry (TOS14519) was 3.5 fold higher for Line 484 than for Line 647 (1.99% and 0.59% respectively) with RI-BA. Lines 647 and 484 were therefore ranked differently by the RI-BA method because of different values of infestation on the resistant check entry. With the SES assessment, the infestation was nearly the same for the two lines and infestation on the susceptible check entry (ITA306) for Lines 484 and 647 was 54.61% and 60.28%, respectively. Likewise, these two lines were closely ranked in the SES method because of the similarity in percentage infestation on the susceptible check entry. This difference in percentage infestation on check entries could be due to difference in insect pressure or due to difference in varietal response to pest attack. ITA306 is highly susceptible, so its level of infestation converges toward a maximum value, which is likely to remain the same throughout the trial. Therefore, assessing damage relative to this infestation is equivalent as referring only to direct percentage infestation on test entry. Thus, the infestation level of ITA306 at 45 DAT is highly variable (CV=0.19), ranging from 17.14% to 48.18% with a mean of 32.89% but escalated from 58% to

65% at 70 DAT with a twofold lower coefficient of variation (0.095), confirming the least variation of infestation at 70 DAT.

The RI-BA methods rely on the accurate assessment of the level of infestation of the resistant check entry. Therefore, conditions that impact upon its estimation influence the accuracy of the methods. The most accurate RI-BA method was confirmed using each of the two checks of known reaction against the AfRGM. Each check was compared to itself. In the case of the RI-BA2 method, the consistent evaluation of RI=1 means that the test entry and the check were of same value. This is the proof of a good evaluation since the two lines involved were the same. A consistent RI value was not obtained for the two other RI methods. Intermediate values between 0 and 1 were observed, reflecting a lack of accuracy for these two methods. RI-BA3 was more accurate than RI-BA1. In RI-BA2, the test entry is assessed relative to a nearby resistant check. This proximity minimises environmental impact, and also take into consideration the insect pressure that affects the test entry within this area.

Our results also showed that plot size is important in choosing the method of estimation. For small plot size, no significant difference was observed between RI-BA2 and RI-BA3, while RI-BA1 differed from the two aforementioned methods. This is because in small plot, variation within the trial is low, so test entries and checks are subjected to the same influences and show little difference of infestation level regardless of the method used. The environment being non-uniform may provide an explanation for the few lines that have scored differently by RI-BA1 and RI-BA2.

With medium plot sizes, a difference between methods was observed. RI-BA2 gave a better estimation of AfRGM damage than RI-BA1 and RI-BA3. With a large plot size, variation within plots tends to increase proportionally. RI-BA3, using the interval mean value calculated between two checks separated only by ten lines, reduced the variability more within the trial than RI-BA1 did and therefore showed more accuracy than RI-BA1. However, RI-BA3 showed less accuracy than RI-BA2 because, as the plot size becomes larger, checks were more repeated and more subjected to environmental variability. As the number of experimental plots to estimate RI-BA3 increased, the probability of environmental variability increased and therefore introduced differences.

Large plot size showed differences among methods of estimation following the same pattern as described above. The more a test entry is evaluated relative to a further check; the higher the variation between environments, and the lesser the accuracy of the method.

For the susceptibility index or RI calculated based on the susceptible parent, the difference between methods was not as highly pronounced as the one observed in resistance checking. In some parts of the trial, the three methods were surprisingly equivalent. This could be due to the fact that as the degree of infestation of ITA306 converged toward a maximum, the zone value, interval mean value and mean value also converged toward the same value and therefore lead to the same value of index. In this study, the mean infestation value for this part was 53.85%, with a low coefficient of variation (cv) of 0.014, confirming the least variation of infestation among entries.

To implement the RI-BA methods, the experimental design had to be slightly modified by replicating each check every ten lines instead of every twenty lines. This modification led to an increase in the total number of entries involved in the trial. This could be a limitation for screening large numbers of test entries especially if one wants to increase the accuracy by replicating each check every 5 lines, for example. For single plant assessment, as in F_2's populations, gaps within check rows should be avoided because each test entry is evaluated by referring to the check of the same column to which it belongs. However, these limitations are largely manageable. Moreover, the RI-BA methods allowed us to assess AfRGM damage in the F_2 generation on a single plant assessment basis, which is not possible with the SES method. The gain in saved time is substantial.

Overall, the RI-BA methods provided better evaluations of AfRGM damage than the SES method, regardless of the trial size or the type of population being investigated. Within RI-BA methods, RI-BA2 was more accurate than RI-BA1 and RI-BA3 when the plot was large because it consistently measures the reaction of the Control entries. No significant difference was observed between the two novel methods RI-BA2 and RI-BA3 when the plot size was small. However, RI-BA2 and RI-BA3 provided better resistance estimates than RI-BA1.When the plot was of medium size, RI-BA2 was significantly better than RI-BA3, and RI-BA3 was itself significantly better than RI-BA1. The best method of assessing AfRGM damage was found to be RI-BA2, regardless of the plot size.

References

Burr, B., Burr, F.A., Thompson, K.H., Albertson, M.C. and Stuber, C.W. (1988). Gene mapping with recombinant inbreds in maize. *Genetics* 118:519-526.

Crane, C.F. and Crane, Y.M. (2005). A nearest-neighboring-end algorithm for genetic mapping. *Bioinformatics* 21:1579-1591.

Doerge, R. (2002). Mapping and analysis of quantitative trait loci in experimental populations. *Nature Review* 3:43-52.

Ferreira, A, da Silva, M.F., da Costa e Silva, L. and Cruz, C.D. (2006). Estimating the effects of population size and type on the accuracy of genetic maps. *Genetics and Molecular Biology* 29:182-192.

Gomez, K.A. and Gomez, A.A. (1984). *Statistical Procedures for Agricultural Research*. 2nd ed. International Rice Research Institute, Los Baños, Philippines, 680 p.

He, P., Li, J.Z., Zheng, X.W., Shen, L.S., Lu, C.F., Chen, Y. and Zhu, L.H. (2001). Comparison of molecular linkage maps and agronomic trait loci between DH and RIL populations derived from the same rice cross. *Crop Science* 41:1240-1246.

IRRI, (1986). *Standard Evaluation System for Rice*. International Rice Research Institute, Los Baños, Philippines.

Jeuken, M., van Wijk, R., Peleman, J. and Lindhout, P. (2001). An integrated interspecific AFLP map of lettuce (*Lactuca*) based on two *L. sativa*×*L. saligna* F_2 populations. *Theoretical and Applied Genetics* 103:638-647.

Nwilene, F.E., Williams, C.T., Ukwungwu, M.N., Dakouo, D., Nacro, S., Hamadoun, A., Kamara, S.I., Okhidievbie, O., Abamu, F.J. and Adam A. (2002). Reactions of differential rice genotypes to African rice gall midge in West Africa. *International Journal of Pest Management* 48:195-201.

Semagn, K., Bjørnstad, Å. and Ndjiondjop, M.N. (2006a). An overview of molecular marker methods for plant. *African Journal of Biotechnology* 5:2540-2568.

Semagn, K., Bjørnstad, Å. and Ndjiondjop, M.N. (2006b). Principles, requirements and prospects of genetic mapping in plants. *African Journal of Biotechnology* 5:2569-2587.

Sharma, R., Aggarwal, R.A.K., Kumar, R., Mohapatra, T. and Sharma, R.P. (2002). Construction of RAPD linkage map and localization of QTLs for oleic acid level using recombinant inbreds in mustard (*Brassica juncea*). *Genome* 45:467-472.

Stroup, W. (2002). Power analysis based on spatial effects mixed models: A tool for comparing design and analysis strategies in the presence of spatial variability. *Journal of Agricultural Biology and Environmental Statistics* **7**:491-511.

Souleymane, N. and Nénon, J.-P. (2009). Comparative study of the morphology of the ovipositor of *Platygaster diplosisae* (Hymenoptera: Platygasteridae) and *Aprostocetus procerae* (Hymenoptera: Eulophidae) two parasitoids associated with the African rice gall midge, *Orseolia oryzivora* (Diptera: Cecidomyiidae). *Psyche* 2009:1-7.

Tanksley, S.D., Ganal, M.W. and Martin, G.B. (1995). Chromosome landing: a paradigm for map-based gene cloning in plants with large genomes. *Trends Genetics* 11:63-68.

Valdar, W.S., Flint, J. and Mott, R. (2003). QTL fine-mapping with recombinant-inbred heterogeneous stocks and in vitro heterogeneous stocks. *Mammalian Genome* 14:830-838.

Williams, C.T. (1998). Crop loss assessment methods for African rice gall midge, *Orseolia oryzivora* Harris and Gagne. Paper presented at WARDA/NARS Group Discussion and Planning Meeting Bio-Ecology and Management of the African Rice Gall Midge, October 12-16, 1998, IITA, Ibadan, Nigeria.

Chapter 3

Gene mode action and heritability estimates in early filial generations segregating for AfRGM resistance in view of predicting genetic gain from selection

Abstract

Plant breeding depends upon the understanding of the genetic basis of the inheritance of the trait of interest and the variability of this trait in the population. The wider the variability, the bigger is the genetic progress that can be achieved. The part of this variability due to genetics is known as heritability. The genetic basis of the resistance to African rice gall midge (AfRGM), *Orseolia oryzivora* Harris and Gagné (Diptera: Cecidomyiidae) and the heritability estimate were therefore calculated in this present study in order to design an optimal mating design to develop AfRGM resistance. Five populations were produced and evaluated for genetic and heritability. These included the F_1, F_2 and F_3 generations. The parents were ITA306, a susceptible *Oryza sativa* subsp. *indica*, and four varieties with a range of reactions to AfRGM attack. All the F_1s from ITA306xTOS14519, ITA306xTOG7106, ITA306xBW348-1 and ITA306xCisadane crosses were susceptible, suggesting that the mechanism of resistance involved a recessive gene. The segregation pattern in the F_2 generations of 1 resistant for 15 susceptible in both the ITA306xTOS14519 and ITA306xTOG7106 crosses, and the segregation of 1 resistant for 8 segregant and 7 susceptible in the ITA306xTOS7106 F_3 families, showed that AfRGM resistance is governed by two genes. However, the differences between the genes involved in the ITA306xTOS14519 and ITA306xTOG7106 combinations were not clear. Moreover, the segregation patterns observed in the ITA306xBW348-1 and ITA306xCisadane crosses were very dissimilar to those found in Mendelian segregations, suggesting that tolerance to AfRGM shown by BW348-1 and Cisadane were under complex genetic control rather than under simple genetic control. The narrow-sense heritability estimates were low in populations involving tolerant varieties, but were high in populations involving varieties resistant to AfRGM. It ranged from 0.086 in the ITA306xCisadane population to 0.4 in the ITA306xTOG7106 population. Intermediate values of 0.14 and 0.27 were observed for the ITA306xBW348-1 and ITA306xTOS14519 populations, respectively. Conversely, the broad-sense heritability estimates were as high as 0.63, with the lowest value of 0.23. These were

observed in the ITA306xTOS14519 and ITA306xCisadane populations, respectively. Intermediate values of 0.32 and 0.53 were observed in the ITA306xBW348-1 and ITA306xTOG7106 populations. The high values of the broad-sense heritability compared to the narrow-sense heritability suggested that the variability observed amongst the populations was mostly due to non-additive gene effects rather than additive gene effects.

Introduction

In modern plant breeding, there is an increasing trend to use a range of molecular breeding tools (Cooper *et al*., 2004; Nelson *et al*., 2004; Lorz and Wenzel, 2005; Varshney *et al*., 2006; Eathington *et al*., 2007; Mumm, 2007, Moose and Mumm, 2008). Breeding is no longer a speculative process where selection is made based only on the expressed phenotypes. It is also less subject to the impact of the environment or to sterility barrier constraints (Jones *et al*., 1997). To deal with the latter problem, transgenic crops were developed to allow for the exchange of genes between different genera or species (James, 2007). Selection has become more precise and more oriented as DNA, which conditions genetic differences amongst individuals, is investigated directly (Moose and Mumm, 2008). However, the costs related to this technology, its specificity of being a challenging technique to implement and the need for qualified technicians, causes this technology to be inacessible to the majority of scientists of the National Agricultural Research System (NARS) of Africa. One of the most effective ways to approach this situation is to develop low cost, accessible technologies with acceptable efficiency which can be used to produce high-yielding hybrid varieties. Marker-assisted selection (MAS) or marker-assisted breeding (MAB) is one of the most dependable options that has emerged over these two last decades (Gepts and Hancock, 2006; Bliss, 2007). However, the success of MAS depends upon several factors, including genetic features and facility constraints (availability of technical facilities and number of individual samples that can be analysed) (Semagn *et al*., 2006). The genetic features include the distance between the closest markers and the target gene, the number of target genes to be transferred, the type of molecular marker(s) used, the genetic base of the trait, and the genetic background into which the target gene has to be transferred (Weeden *et al*., 1992; Francia *et al*., 2005). Apart from the last characteristic mentioned, all the other characteristics involving genetic features are either directly or indirectly linked to, or related to, the genetic bases. Understanding the genetic bases

of the inheritance of the trait of interest is, therefore, the main prerequisite for production of improved varieties using MAS. The genetic basis of inheritance is sought through the first filial generations of inbreeding to the F_3 generation. It implies stating whether the trait is for resistance or susceptibility, the number of genes involved in the phenotypic expression of the trait and how the trait is inherited. The last implication was inappropriately termed heritability. Heritability is not a measure of whether a trait is inherited, but rather a measure of the degree to which the variance in distribution of a phenotype is due to genetic causes (Sleper and Poehlman, 2006; Ullrich, 2007), and gives an estimate of the contribution of each parent involved in a cross. This contribution is designated as their breeding value, and is described as either general combining ability (GCA) or specific combining ability (SCA). GCA and SCA values reflect the relative importance of additive genetic effects and non-additive effects, respectively (Rojas and Sprague, 1952, Moose and Mumm, 2008). The GCA of one parent provides an assessment of its breeding value, as judged by the mean performance of its progeny from crosses with other crossing partners. The SCA is the ability of a line or population to exhibit superiority or inferiority when combined with specific lines or populations (Saad *et al*., 2004).

One problem affecting the breeding for resistance to AfRGM is that heritability estimates for resistance in African rice population have not been calculated, limiting progress in breeding for resistance to AfRGM. Estimating the heritability in its narrow and broad-sense, in view of predicting genetic gain, is an important step towards breeding for resistance to AfRGM. Broad-sense heritability estimates include all genetic effects (additive, dominance and epistatic), while narrow-sense heritability estimates include only additive variance (Sprague and Tatum, 1942). Breeding for resistance to quantitative traits involves shifting the population mean toward resistance (Vacaro *et al*., 2002). In order to do this, breeding programmes rely heavily on additive genetic variation for successful population improvement toward more resistant phenotypes (Falconer and Mackay, 1996; Lynch and Walsh, 1998). However, since the inheritance of resistance to AfRGM in African rice is unknown, narrow-sense and broad-sense heritability would both be more useful to plant breeders. Heritability estimates for AfRGM have not yet been reported for any rice species. Therefore, this study was conducted to determine the narrow-sense and the broad-sense heritability of resistance to AfRGM in rice, after having determined its genetic basis. The objectives of this study were threefold: (1) to

determine the number of genes involved in the resistance to AfRGM and identify it as either dominant or recessive, in order to choose an appropriate method for mapping and tagging the genes, (2) to determine narrow-sense heritability estimates in view of predicting gain from selection for resistance to AfRGM in rice; and (3) to evaluate broad-sense heritability to determine the primary genetic effect (additive, dominant or epistatic) in the inheritance of resistance to AfRGM in early filial generations.

3-1 Materials and methods

3-1-1 Plant materials

The plant material was constituted of five populations comprising 479 F_2 plants and 649 F_3 families derived from a cross between ITA306 and TOS14519, 90 BC_1F_2 plants from an ITA306 and TOG7106 cross, 532 F_2 from an ITA306 and BW348-1 cross and 413 F_2 plants from an ITA306 and Cisadane cross. TOS14519 and TOG7106 are considered to be moderately resistant, and resistant, respectively, to AfRGM, according to IRRI's evaluations. Cisadane and BW348-1 are tolerant, while ITA306, a popular variety released in Nigeria as FARO57, is susceptible. The crosses involved one susceptible variety, used as female, and four males with different levels of resistance to AfRGM. Previously, field screenings were conducted by Nwilene *et al*. in 2000 during the July to September rainy season to identify these parents used in our study. The screening material includes *Oryza sativa* lines, *Oryza glaberrima* accessions and their interspecific crosses. These were screened for resistance to AfRGM in a paddy screenhouse, and field conditions across four West African countries (Burkina-Faso, Mali, Nigeria and Sierra Leone). The screening material was provided by the Genetic Resources Centre (GRU) of the International Institute for Tropical Agriculture (IITA) and the Africa Rice Centre (AfricaRice), Benin.

3-1-2 Genetic basis of the resistance to AfRGM

The level of damage caused by AfRGM was recorded on 20 hills representing the same number of F_1 individuals from each cross, to determine whether the gene for resistance to AfRGM is dominant or recessive. Segregation pattern was determined in the F_2 generation, to estimate the number of genes involved in the resistance. For the ITA306 and TOS14519 cross only,

confirmation of this segregation pattern was done in the F_3. F_1 and F_2 generations were evaluated on a single plant basis. Each plant was declared resistant when no galls were developed or susceptible when the tillers developed galls. For the F_3 generation, evaluation for each F_3 family was made on 20 hills. Resistance index (RI) was calculated based on zone value estimates of the level of tillers' infestation of resistant or susceptible checks. The resistance was assessed referring to percent of the level of infestation of the resistant check entry while the susceptibility was referenced to the susceptible check.

3-1-3 Development of F_1 generation and validation

Sowing: Based on their growth duration (Table 3-1), seeds of parents were sown in a scheduled to ensure that anthesis of the male parents coincided with the heading of the female parent. The female parent was sown into 5 pots of 12 liters filled 3/4 part with paddy soil, on weekly basis for a month prior to pollination. Male parents were sown either before (TOS14519), at the same time (Cisadane and BW348-1), or after (TOG7106) the female parent, depending on the duration of their crop cycle.

Crossing: Emasculation of each flower was done by excising spikelets with a pair of scissors at the top 1/3 and removing the anthers with small pliers, then covering the flowers with a seed envelope, working early in the morning between 08h30 and 10h00. Anthesis took place at around 12h00, at which time a panicle shedding pollen was gently dusted on the emasculated spikelet of the female parent to effect pollination. It was then covered with 80 x 130 mm waterproof transparent crossing envelopes and appropriately labelled.

Validation of F_1 plants: Successful crosses that created F_1s from ITA306xTOS14519 and ITA306xTOG14519 were identified through plant vigor assessment or/and through basal sheath anthocyanin colouration. The principle of this validation relies on the fact that F_1 plants are more vigorous than their parents, and that the F_1 plants' basal sheath colouration could only come from TOS14519 and TOG7106, used as male parents, since the female parent ITA306, is not coloured. For Cisadane and BW348-1, which are not coloured, the F_1 plants were identified with 3 highly polymorphic SSR markers for validation. Vigour assessment was also used to validate the F_1 plants.

Table 3-1: Pedigree, cycle and country of origin of the varieties involved

Variety	Pedigree	Cycle (days)	Country of origin
ITA306	TOX494-3-6-9-6/TOX711//BG6812	125-135	IITA
TOG7106	Landrace	90	Senegal
TOS14519	Landrace	160	The Gambia
BW348-1	Unknown	125	Sri Lanka
Cisadane	Pelita I-1/B2388	125	Nigeria

3-1-4 Development of filial generations

Ten F_1 seeds for each cross were sown, and validated F_1 plants were advanced to F_2 in the greenhouse. Dormancy was broken by keeping F_1 seeds in an incubator for 14 days at 40^{o}C and by dehulling them manually before sowing. F_1 seeds were pre-germinated in petri dishes under optimum germination conditions or *in vitro*. Three-week old seedlings were transplanted into 12-litre plastic buckets, which were 3/4-filled with paddy soil, at the rate of one seedling per bucket. Daily watering was done until plant maturity. Prior to transplanting, fresh leaves were harvested and DNA was extracted for the molecular identification of true F_1 plants using to the cetyl trimethyl ammonium bromide (CTAB) protocol, as described by Temnykh *et al*. (2000).

For the ITA306xTOS14519 cross, among the validated F_1 plants, two produced 503 and 689 F_2 seeds and were chosen to produce 479 and 649 F_2 plants, respectively. Each single F_2 plant of the latter was selfed to generate 649 F_3 families. Prior to sowing, 1g of compound fertiliser (15:15:15 NPK) at the rate of 100 kg.ha^{-1} was incorporated into each pot. The pots were hand-weeded whenever necessary. At maturity, F_2 individual plants were harvested separately and numbered for evaluation as F_3 families. The same principle was adopted to produce 532 F_2 plants derived from the ITA306xBW348-1 cross, and 413 F_2 plants were derived from the ITA306xCisadane cross. From the ITA306xTOG7106 cross, 10 validated F_1s were backcrossed to ITA306. One of them produced 154 grains BC_1F_1 which were individually seeded in plastic pots to produce BC_1F_2 plants. One of the BC_1F_1 plants was selfed and yielded a total of 94 grains BC_1F_2. From this total, 90 plants BC_1F_2 were grown.

3-1-5 Phenotyping

3-1-5-1 Paddy greenhouse screening

The experiment was conducted between January and August 2010 at AfricaRice-Ibadan, hosted by IITA, Ibadan Station, Nigeria. Seeds of the entries were sown in a nursery bed. The F_2 plants and the F_3 family were transplanted 14 days after sowing (DAS), in a paddy screenhouse, in a 2-m long row with a space of 0.20 m within and between rows, with the two parents of the cross planted after every 10 entries. The F_1 plants were transplanted 28 DAS to allow for their recovery from the leaf harvest which was done at 14 DAS. The experimental plot was flanked by five rows of ITA306 as an infestation band where the AfRGM population was maintained. The equivalent of 40 kg each of N, P_2O_5 and K_2O per ha was applied at transplanting, using NPK 15-15-15. Another 40 kg N/ha was applied at 21 days after transplanting (DAT), using urea. The plots were hand-weeded at 21 and 40 DAT. The experimental design was an augmented design with two checks replicated every ten lines. Two weeks after transplanting, daily water spraying was done every two hours from 8 am to 4 pm to facilitate larva movement.

3-1-5-2 Assessment of damage by AfRGM

Damage by AfRGM was assessed according to the resistance index-based assessment (RI-BA) as described in Chapter 2. The RI method of assessment followed was to use resistant check zone values to compute the percent tiller infestation relative to the resistant check, instead of each entry's mean value. The formula is as follows:

$$\text{Resistance Index (RI)} = \frac{\text{\% of tillers with galls on test entry}}{\text{\% of tillers with galls on resistant check variety}} \times 100$$

The zone value is the value measured on the nearest resistant check within an interval of ten lines. The resistance index is either 0 if there are no galls on the test entry, less than 1 if the test entry is more resistant than the resistant check, or more than 1 when the test entry is less resistant than the check (Williams, 1998).

3-1-6 Heritability estimates

Narrow-sense and broad-sense heritability (respectively h^2 and H^2) were estimated for each F_2 population and the F_3 family. Narrow-sense heritability (h^2) was determined from mid-parent versus progeny regression analysis, where the slope of the regression line is equal to h^2 (Fernandez and Miller, 1985). Mid-parent and progeny means utilised in the analysis are presented in Table 3-2. The pool of ten entry lines bordered by each of the two parents was used as family for the mean parent-offspring regression. Broad-sense heritability was estimated using the variance components method developed by Simmonds (1981). Broad-sense heritability (H^2) for AfRGM resistance was estimated as follows:

$$H^2 = \frac{VF_2 - \frac{(VP_1 + VP_2 + VF_1)}{3}}{VF_2} = \frac{VA + VD}{VA + VD + VE}$$

Where V = variance of the trait under consideration and H^2 is the broad-sense heritability. Variances were estimated using percent tiller infestation.

Table 3-2: Mid-parent and progeny mean used for narrow-sense heritability estimates according to population

F_2 ITAxTOS		F_2 ITAxCisadane		F_2 ITAxBW		F_2 ITAxTOG		F_3 ITA306xTOS	
PMI	OMI	PMI	OMI	PMI	OMI	PMI	OMI	PMI	OMI
31.59	38.95	60.40	54.01	61.70	50.89	27.01	54.90	33.96	39.52
30.99	37.86	60.30	54.63	54.90	52.67	31.70	62.01	32.70	38.92
23.02	36.49	53.84	45.91	55.87	50.76	27.33	52.42	22.83	36.49
28.85	41.00	59.86	42.25	53.90	47.31	35.24	55.46	30.09	41.43
29.99	39.14	64.69	53.95	54.00	52.88	33.95	64.34	32.80	41.21
30.54	38.96	59.38	52.30	63.00	55.54	25.57	53.41	33.94	39.56
31.62	42.14	56.30	58.14	57.25	51.35	30.03	60.48	35.62	42.43
30.87	41.03	58.31	56.66	59.06	52.09	28.37	55.84	34.27	40.97
27.00	38.16	57.69	50.35	58.97	48.05	34.52	61.18	37.70	42.65
30.54	36.99	62.74	55.48			30.48	59.40	35.65	38.11

PMI=Parent mean infestation, OMI=Offspring mean infestation

3-1-7 Statistical analyses

Data on the levels of tillers and hill infestations were transformed using the Arcsin transformation prior to the analysis of variance, as described by Gomez and Gomez (1984), as follows: (1) percentage data from the range 30% to 70% was not transformed; (2) when all data in the range was between 0% and 50% or between 50% and 100% the square root transformation was needed; (3) for data that did not belong to any of the aforementioned ranges, the Arcsin transformation was used. XLSTAT package version 2010 was utilised for analysis of variance of both the raw and transformed data.

3-2 Results

3-2-1 Genetic basis of the resistance to AfRGM

3-2-1-1 Screening of F_1 generation

The F_1s were all susceptible, irrespective of crosses. The maximum infestation level always corresponded to the maximum RI. Conversely, minimum infestation level did not correspond to the minimum RI except for ITA306xBW348-1. Overall, the lowest level of tiller infestation was observed in ITA306xTOS14519 (34.8%), followed by ITA306xBW348-1(59.7%), then by ITA306xTOG7106 (64.9), and finally by ITA306xCisadane (67.1%) for the highest infestation. However, the lowest index value was observed in the ITA306xBW348-1 combination (1.01), followed by ITA306xCisadane (1.45), then by ITA306xTOG7106 (16035.34), and finally by ITA306xTOS14519 (17981.17). Table 3-3 also shows that despite ITA306xTOS14519 progenies being the least infested, their corresponding index was the highest. The tail values of the infestation were not correlated to those of the RI, confirming that RI depends on the level of tiller infestation of the resistant check entry.

Table 3-3: Percentage tiller infestation and RI among F_1 progenies involving ITA306 and four parental lines of various degrees of resistance to AfRGM

F_1 generations	Percent infestation (%)			Resistance index (RI)		
	Min	Mean	Max	Min	Mean	Max
ITA306xTOS14519	15.85 (*3)	34.76	57.14 (*11)	1.28 (*4)	17981.17	57142.86 (*11)
ITA306xCisadane	50 (*3)	67.13	87.5 (*6)	0.96 (*1)	1.45	2.63 (*6)
ITA306xBW348-1	33.33 (*4)	59.7	91.67 (*12)	0.43 (*4)	1.01	1.92 (*12)
ITA306xTOG7106	44.44 (*6)	64.90	87.5 (*14)	7.35 (*2)	16035.34	31500 (*14)

(*) corresponds to the progeny on which the data was recorded

Concerning ITA306xTOS14519 F_1 progenies, the level of tiller infestation ranged from 15.85 (Line 3) to 57.14% (Line 11) with an infestation mean of 34.8%. The RIs associated with this cross were 1.28 (Line 4), 1798.17 and 57142.86 (Line 11) as a minimum, mean and maximum, respectively. Of the 20 hills involved in the data record, 13 individuals died due to high levels of infestation. At 45 DAT each hill was highly infested.

For the ITA306xBW348-1 cross, minimum and maximum level of infestation was associated with the corresponding RI. Percent tiller infestation ranged from 33.3% (Line 4) to 91.7% (Line 12), with an infestation mean of 59.7%. These corresponded to resistance indices of 0.43 (Line 4), 1.01 and 1.92 (Line 12) as a minimum, mean and maximum respectively. Of the 20 F_1 progenies, 12 were rated more resistant (RI<1) than BW348-1, one (Line 1) was rated as being of the same resistance (RI=1) and 7 (Lines 10, 12, 14, 15, 16, 18 and 20) were designated less resistant (RI>1) than BW348-1 (Table 3-4).

Table 3-4: RI calculated on twenty ITA306xBW348-1 F_1 progenies recorded at 70 DAT

Progenies	1	2	3	4	5	6	7	8	9	10
Resistance Index	0.73	0.74	0.92	0.43	0.73	0.64	0.83	0.86	0.86	1.43
Progenies	11	12	13	14	15	16	17	18	19	20
Resistance Index	0.80	1.92	0.93	1.53	1.33	1.19	0.74	1.11	1.00	1.49

In the progenies of ITA306xCisadane, tiller infestation ranged from 50% (Line 3) to 87.5% (Line 6) with a mean infestation of 67.1%. The RIs associated with this cross were 0.96 (Line 1), 1.45 and 2.63 (Line 6) as minimum, mean and maximum, respectively.

Concerning the ITA306xTOG7106 F_1 generation, the maximum tiller infestation of 87.5% and the maximum index of 31500, were recorded on Line 14. The minimum infestation was 44.4% (Line 6) with a mean infestation of 64.9%. The least and mean indices associated with this cross were 7.35 (Line 2) and 16035.34 respectively.

3-2-1-2 Screening of F_2 generation

Apart from ITA306xTOG7106 cross, in which maximum infestation was rated 90.9%, in the other cross combinations, minimum and maximum tiller infestation rated 0 and 100% respectively. For the mean infestation, ITA306xTOS14519 was the least infested (40.1%) while ITA306xCisadane was the most infested (67.1%). ITA306xBW348-1 and ITA306xTOG7106 had values of 51.4% and 58.6%, respectively, ranking them second and third. Table 3-3 also points out that minimum index was 0% in all crosses. In ITA306xBW348-1 and ITA306xTOG7106, this index was observed for many progenies while only one progeny had a 0 index in ITA306xTOS14519 (Line 4) and ITA306xCisadane (Line 142). Very high index values were also recorded for the mean and maximum indices for ITA306xTOS14519 (60360.18 and 200000) and ITA306xTOG7106 (14322.9 and 31500) while ITA306xCisadane and ITA306xBW348-1 rated 1.35 and 0.93, respectively. Overall, progenies from the ITA306xBW348-1 cross were more resistant (RI<1) than their parents whereas those of the other crosses appeared less resistant (RI>1) compared to the male parent. Apart from the ITA306xCisadane combination, where each tail value of infestation or index was consistently recorded on a single individual, in the other combinations, tails values were recorded on both single and multiple progenies (Table 3-5).

Table 3-5: Percentage tiller infestation and RI among F_2 progenies involving ITA306 and 4 parental lines of various degrees of resistance to AfRGM

F_2 generations	Percent infestation (%)			Resistance index (RI)		
	Min	Mean	Max	Min	Mean	Max
ITA306xTOS14519	0 (***)	40.1	100 (***)	0 (*4)	60360.18	200000 (***)
ITA306xCisadane	0 (*142)	67.1	93.8 (*135)	0 (*142)	1.35	6.70 (*135)
ITA306xBW348-1	0 (***)	51.4	100 (***)	0 (***)	0.93	2.15 (*500)
ITA306xTOG7106	0 (***)	58.6	90.9 (*39)	0 (***)	14322.9	31500 (*14)

(*) corresponds to the progeny on which the effect or the index was observed
(***) means that infestation effect or index was observed on multiple progenies

3-2-1-3 Screening of ITA306xTOS14519 F_3 generation

Table 3-6 shows that some lines were not affected among the ITA306xTOS14519 F_3 families, giving a corresponding RI of 0. Maximum tiller infestation was as high as 68.1% and was observed only on Line 3. Overall, around one-third of the tillers of each family was infested (28.5%). The mean index of 9.64 is distant from the maximum index of 123.94 observed on Progeny 81.

Table 3-6: Percentage tiller infestation and RI among F_3 families involving ITA306 and TOS14519

F_3 generation	Percent infestation			Resistance index (RI)		
	Min	Mean	Max	Min	Mean	Max
ITA306xTOS14519	0 (***)	28.5	68.1 (*3)	0 (***)	9.64	123.94 (*81)

(*) corresponds to the progeny on which the effect or the index was observed
(***) means that infestation effect or index was observed on multiple progenies

3-2-2: Segregation pattern of progenies in filial generations

All the F_1 progenies were susceptible to AfRGM, irrespective of the cross. Of the 20 hills evaluated, none recorded a nil value for disease incidence. In the F_2 generations, the

segregation patterns varied from a clear segregation of 1 resistant for 15 susceptible in ITA306-TOS14519 and ITA306xTOG7106 crosses, to an ambiguous one in ITA306xCisadane and ITA306xBW348-1 crosses. Indeed, a segregation of 1 resistant for 412 susceptible and 6 resistant for 526 susceptible was observed. These ratios do not correspond to any ratio of the known F_2 segregation pattern. However, these segregations (1 for 412 and 6 for 526) are statistically similar to the 0:1 ratio. An F_3 segregation pattern of 1 resistant for 8 segregant and 7 susceptible (1-8-7) was observed for the cross ITA306xTOS14519 (Table 3-7).

Table 3-7: F_2 and F_3 segregation patterns among ITA306 crosses with 4 parental lines of various degrees of resistance to AfRGM

Cross	F_1		F_2			F_3		
	R:S	Ratio	R:S	Ratio	Khi^2	R:Seg:S	Ratio	Khi^2
ITA306xTOS14519	0:20	0:1	34:445	1:15	0.66*	38:358:253	1:8:7	6.89*
ITA306xCisadane	0:20	0:1	1:412	0:1	0.002*	-	-	-
ITA306xBW 348-1	0:20	0:1	6:526	0:1	0.068*	-	-	-
ITA306xTOG7106	0:20	0:1	4:86	1:15	0.48*	-	-	-

- means the data was not recorded, U=Unknown, * = Not significant deviation from expected ratio at p= 0.001

3-2-3 Genetic variability among early generation using heritability estimates

3-2-3-1 Narrow-sense heritability estimates with mean parent offspring regression in the F_2 and F_3 generations

The results for the narrow-sense heritability estimates for resistance to AfRGM within early filial generations are presented in Figure 3-1. Narrow-sense heritability estimates ranged from 0.06 to 0.44 in the ITA306xBW348-1 and ITA306xTOG7106 crosses, respectively. Globally, it was low in populations involving ITA306 and the tolerant varieties, while it was high in populations where ITA306 was crossed with resistant varieties. Narrow-sense heritability estimates were as low as 0.06 and 0.13 in the ITA306xCisadane and ITA306xBW348-1 crosses, respectively but were as high as 0.264 and 0.44 in the ITA306xTOS14519 and ITA306xTOG7106 crosses.

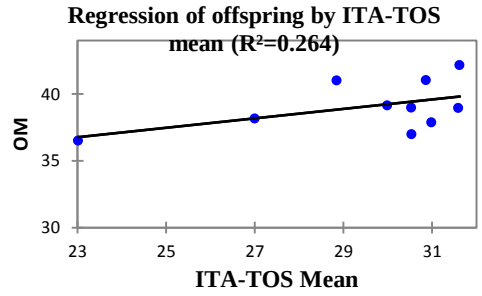


Figure 3-1a: h^2 estimate in ITA306xTOS14519 F_2 progenies

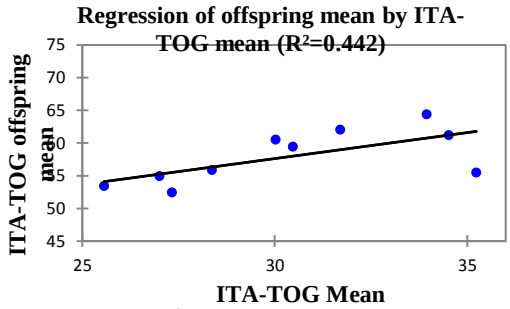


Figure 3-1b: h^2 estimate in ITA306xTOG7106 F_2 progenies

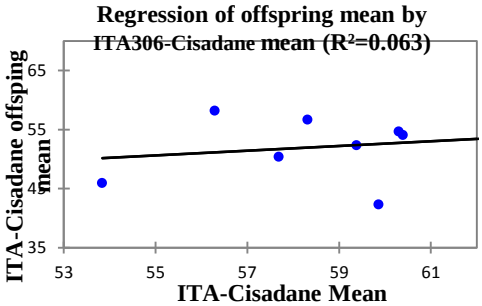


Figure 3-1c: h^2 estimate in ITA306xCisadane F_2 progenies

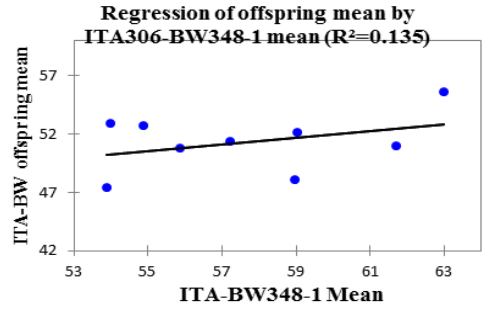


Figure 3-1d: h^2 estimate in ITA306xBW348-1 F_2 progenies

Figure 3-1 a-d: Narrow-sense heritability (h^2) estimates by mid-parent offspring regression within F_2 populations, involving ITA306 and 4 lines of various levels of resistance to AfRGM

In Figure 3-2, the narrow-sense heritability estimate at the F_3 level showed a higher value (0.4) than at the corresponding F_2 (0.26). This value remained low, amounting to less than 0.5.

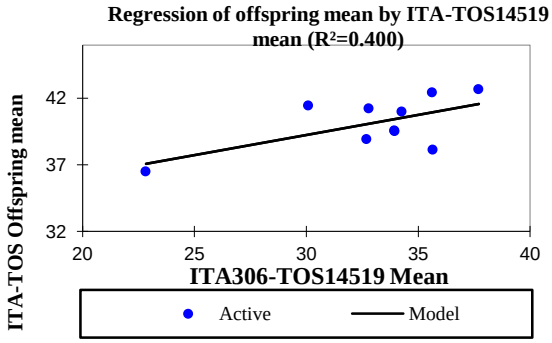


Figure 3-2: Narrow-sense heritability estimate within the ITA306xTOS14519 F_3 progenies

3-2-3-2 Broad-sense heritability estimates with variance components method in F_2 and F_3 generations

In the F_2 generations, broad-sense heritability estimates ranged from 0.23 within the ITA306xCisadane progenies to 0.63 amongst the ITA306xTOS14519 progenies. Intermediate

values of 0.32 and 0.53 were observed within the ITA306xBW348-1 and ITA306xTOG7106 progenies, respectively. Broad-sense heritability estimates were higher than those of narrow-sense heritability. The highest value of estimate was observed with ITA306xTOS14519 for H^2, whereas it was recorded in the ITA306xTOG7106 progenies for h^2. The lowest values were consistently recorded on the ITA306xCisadane progenies in both cases, but the broad-sense estimate was higher. Like narrow-sense heritability, H^2 estimates were higher in populations involving resistant parental lines than in tolerant parental lines. Variances also varied within populations and across generations. Within generations, F_2 had the highest variance values, regardless of the cross. Crosses involving resistant lines were more variable than crosses involving tolerant lines. The least variability was, however, observed with the male parents (P_2) of the ITA306xTOG7106 (5.87) and ITA306xTOS14519 crosses (72.88).

In the ITA306xTOS14519 F_3 generation, H^2 (0.34) was low and was in the same range as H^2 in the F_2 cross of ITA306xBW348-1 (0.32). It represented nearly half (0.34) of its corresponding F_2 (0.63) (Table 3.8 second part).

Table 3-8: Broad-sense heritability estimates in F_2 and F_3 generations involving ITA306 and different parental lines of various degrees of resistance to AfRGM

F_2 Cross	VF_2	VF_1	VP_1	VP_2	Broad-sense heritability (H^2)
ITA306xTOG7106	351.81	215.77	273.52	5.87	0.53
ITA306xTOS14519	502.74	256.24	224.14	72.88	0.63
ITA306xBW348-1	280.07	151.45	190.40	226.06	0.32
ITA306xCisadane	242.19	177.25	189.61	194.82	0.23
F_3 Cross	VF_3	VF_2	VP_1	VP_2	Broad-sense heritability (H^2)
ITA306xTOS14519	273.29	502.74	28.034	12.99	0.34

VF_3, VF_2, VF_1, $VP_1$and VP_2 are variances for F_3, F_2 , F_1, P_1 and P_2

3-3 Discussion

All the F_1 progenies exhibited susceptibility, irrespective of the crosses, proving that the gene involved in the resistance or tolerance to AfRGM studied here is recessive. In the F_2 generations, the segregation patterns varied. Crosses between ITA306 and rice varieties with tolerance to gall midge were susceptible but did not show any distinct segregation pattern,

suggesting a higher order of interaction. Himabindu *et al*. (2008) reported similar results for crosses between MR1523 (a resistant cultivar to Asian gall midge Biotypes 1, 3 and 4) and rice varieties with known and unknown gall midge resistance. However, in ITA306xTOS14519 and ITA306xTOG7106 crosses, a segregation of 1 resistant for 15 susceptible was noted. Inheritance of resistance to AfRGM is therefore controlled by two genes. This assertion was confirmed by the F_3 segregation pattern of 1 resistant for 8 segregants and 7 susceptibles (1-8-7) for the cross ITA306xTOS14519. Meanwhile, it is not clear whether the genes involved in the resistance in the ITA306xTOS14519 and ITA306xTOG7106 crosses are the same. An allelic test, involving the two parental lines, could help in deciding on the identity of these genes. This result that two genes governed a resistance also contradicts the paper by Maji (1998) who found that resistance in TOG7106 is controlled by a single recessive gene when he considered F_2 generation pooled data over locations. The segregation patterns, he documented, referring to F_1 and F_3 generations, did not conform to the 1R:3S and 1R:8S segregation models expected in the case of a single recessive gene inheritance. However, an allelic test involving two resistant lines (TOG7106 and TOG7442) each bearing, one recessive resistance gene, not only confirmed the existence of the same gene in the two lines, but also confirms that resistance in TOG7106 was under the control of a single recessive gene. The result reported here may differ from that of Maji because: (1) TOG7106 could have inherited the second gene from the traditional tropical *japonica* landrace TOS14519, which had been growing in close vicinity to TOG7106 during crossing for mapping population development, or during field screening, as reported by Asante *et al.* (2009) on the genetic origin of fragrance in Nerica1. Indeed, due to space limitation and experimental constraints, all the parental lines involved in the present study were grown side by side; (2) following the same explanation, the BC_1F_2 ITA306xTOG7106 population could have also exchanged genes with the varieties in its vicinity. Indeed, the population was surrounded by many other landraces, cultivars or varieties, including a collection of *glaberrima* landraces collected from Niger and Burkina-Faso; (3) The low number of 90 BC_1F_2 ITA306xTOG7106 progenies (due to interspecific incompatibility between ITA306 and TOG7106), could explain the difference because limited population sizes introduce bias into estimations. Maji (1998) reported similar constraint where the single location ratio did not fit the expected ratio but when data was pooled (meaning large population size) the segregating ratio conformed to the expected ratio. Therefore, apart from the allelic test

with TOS14519, the mapping of the gene and the use of a large population size could all be investigated to clarify the number of genes involved in TOG7106 resistance to AfRGM. However, the high value of the narrow-sense heritability recorded for the ITA306xTOG7106 population provides a strong indication of more than one resistance gene being present. Our results also point out significant deviations from the expected 1S:0R ratio, which was recorded for the tolerant parents and their F_1 progenies, indicating a possible role of modifiers or some degree of heterozygosity in the tolerant parents (Maji, 1998).

Heritability refers to the proportion of variation between individuals in a population that is influenced by genetic factors. Heritability describes the population, not individuals within that population. The determination of the extent of heritability is risky because the parameters defining heritability as low, medium or high have not clearly been stated. However, Singh (2005) suggested that the magnitude of heritability estimates for a highly heritable trait is more than 0.5. For a moderately heritable trait, heritability values fall between 0.2 and 0.5. In reference to the nature of the inheritance of the trait in this study, heritability estimates were globally considered as low to medium, ranking at most at 0.63. The overall low levels of heritability especially in the F_2 may be explained by the single plant data and the distribution of AfRGM (Edelson *et al*., 1989; Fournier *et al*., 1995). AfRGM are not randomly distributed throughout the tunnel paddy screenhouse. Therefore, any particular plant may have very few AfRGM infestations due to true genetic resistance or simply by chance. The heritability estimate changes according to the genetic and environmental variability present in the population. In this view, genetically inbred populations have much higher heritability estimates than outbred (genetically variable) populations under very homogeneous environments. Therefore, the lowest heritability estimates observed in populations involving varieties tolerant to AfRGM, compared to populations with resistant lines, may be explained by the similarity or the close relatedness of the tolerant lines to the susceptible lines ITA306, in terms of reaction against AfRGM. Indeed, the tolerant lines (BW348-1 and Cisadane) had the same range of infestation (59.1 and 52.5%) as ITA306 the susceptible parent (62.4%) (data not shown). The difference between them is due to the acceptable yield produced by the tolerant lines relative to the susceptible line, where almost all the plants died before reaching maturity. In such cases, crosses involving the tolerant lines (BW348-1 and Cisadane) and the susceptible line (ITA306)

as parental lines, led subsequently to a more homogeneous population and are therefore prone to exhibit low heritability. Indeed, several authors (Hill *et al*., 2006, Ahmad *et al*., 2008) reported the variability of the trait within the population as one of the requirements to have high heritability estimates irrespective of the material involved in the studies. Likewise, crosses involving distant parental lines according to their resistance to AfRGM produce a heterogeneous population in which heritability is consequently high.

Broad-sense heritability reflects all possible genetic contributions to a population's phenotypic variance. Included are effects due to allelic variation (additive variance), dominance variation, epistatic (multi-genic) interactions, and maternal and paternal effects, where individuals are directly affected by their parents' phenotype. The higher values of the broad-sense heritability relative to the narrow-sense values, suggested that the variability observed amongst the populations was mostly due to dominant and epistatic effects rather than additive effects. Conversely, narrow-sense heritability quantifies only the portion of the phenotypic variation that is additive (allelic). Variability was observed in the ITA306xTOS14519 F_3 population where narrow-sense heritability was preponderant over broad-sense heritability, and was mostly due to the additive effect.

Heritability estimates were also calculated in its narrow and broad senses, in the F_2 and F_3 generations, as per Painter (1951), who recommended evaluating F_3 families rather than F_2 plants for insect resistance research. Our result identified two trends. The first trend was in concordance with Fehr's (1991) assertion, which postulates that estimating heritability on a family basis increases the levels of heritability estimated. This trend was observed in the narrow-sense heritability estimates, where the estimate was higher in the F_3 than in the F_2 generation. Conversely, the second trend depicted a higher heritability estimate in the F_2 generation compared to the F_3 generation, in the broad-sense heritability estimates. This variation has been explained by many authors (Jensen, 1973a, b; Heimstein and Murray, 1994 and William and James, 2001), who claimed that heritability estimates change according to the genetic and environmental variability present in the population, and therefore they are not an absolute measurement of the contribution of genetic and environmental factors to a phenotype.

Overall, the low level of heritability for resistance to AfRGM in African rice cultivars indicates that selection on the basis of half sib or S_1 families would offer greater genetic gain

for this trait than single plant selections. Moreover, it is difficult to screen for resistance against AfRGM because of the low heritability (less than 0.5), due to the masking effect of the environment on genotypic effects, as reported by Singh (1993), and Singh *et al*. (1994), on a similar study conducted on sugarcane. However, the values of 26.4 and 44.2 suggest that resistance to AfRGM is a moderately heritable trait that can respond to selection. In addition, the study indicated that the resistance to AfRGM is governed by two recessive genes in both the landraces TOS14519 and TOG7106. However, their nature was not determined, and an allelic test, involving the two parental lines, would help in deciding on the identity of these genes.

References

Ahmad, W., Khan, N.U., Khalil, M.R., Parveen, A., Aimen, U., Saeed, M., Samiullah, A. and Shah, S.A. (2008). Genetic variability and correlation analysis in upland cotton. *Sarhad Journal of Agriculture* 24:573-580.

Asante, M., Kovach, M., Huang, L., Harrington, S., Dartey, P., Akromah, R., Semon, M. and McCouch, S. (2009). The genetic origin of fragrance in Nerica1. *Molecular Breeding* 26: 419-424.

Bliss, F.A. (2007). Education and preparation of plant breeders for careers in global crop improvement. *Crop Science* 47: 250-261.

Cooper, M., Smith, O.S., Graham, G., Arthur, L., Feng, L. and Podlich, D.W. (2004). Genomics, genetics, and plant breeding: a private sector perspective. *Crop Science* 44:1907-1913.

Eathington, S.R., Crosbie, T.M., Edwards, M.D., Reiter, R.D. and Bull, J.K. (2007). Molecular markers in a commercial breeding program. *Crop Science* 47:154-163.

Edelson, J.V., Cartwright, B. and Royer, T.A. (1989). Economics of controlling onion thrips (*Thysanoptera: Thripsidae*) on onions with insecticides in south Texas. *Journal of Economic Entomology* 82:561-564.

Falconer, D.S. and Mackay, T. (1996). *Introduction to Quantitative Genetics*. Longman, Essex, England, United Kingdom.

Fehr, W.R., (1991). *Principles of cultivar development. Vol. 1, Theory and technique*. MacMillan, New York, USA.

Fernandez, G. and Miller, J. (1985). Estimation of heritability by parent-offspring regression. *Theoretical and Applied Genetics* 70:650-654.

Fournier, F., Boivinand, G. and Stewart, R.K. (1995). Sequential sampling for *Thrips tabaci* on onions. In: *Thrips biology and management*. Parker, B.L., Skinner, M. and Lewis, T. (Eds), pp. 557-562. Plenum Press, New York, USA.

Francia, E., Tacconi, G., Crosatti, C., Barabaschi, D., Bulgarelli, D., Dall'Aglio, E. and Vale, G. (2005). Marker assisted selection in crop plants. *Plant Cell Tissue Organ Culture* 82: 317-342.

Gepts, P. and Hancock, J. (2006). The future of plant breeding. *Crop Science* 46:1630-1634.

.

Gomez, K.A. and Gomez, A.A. (1984). *Statistical procedures for agricultural research*. 2nd ed. International Rice Research Institute, Los Baños, Philippines. 680 p.

Heimstein, R.J., and Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. Simon and Schuster, New York, USA.

Hill, W.G., Barton N.H. and Turelli, M. (2006). Prediction of effect of genetic drift on variance component under a general model of epistasis. *Theoretical Population Biology* 70:56-62.

Himabindu, K., Suneetha, K., Sama, V.S.A.K., Cheralu, C., Rao, P.R.M. and Bentur, J.S. (2008). A new gene for gall midge resistance in rice variety MR1523. *Rice Genetics Newsletter* 25:73-75.

James, C. (2007). Global Status of Commercialized Biotech/GM Crops in 2007. In: *ISAAA Brief 37*. ISAAA, Ithaca, New York, USA.

Jensen, A. R. (1973a). *Educability and group differences.* Harper and Row, New York, USA.

Jensen, A. R. (1973b). *Educational differences.* Methuen, London, United Kingdom.

Jones, M.P., Audebert, A., Mande, S. and Aluko, K. (1997). Characterization and utilization of *Oryza glaberrima* Steud. In upland rice breeding. In: *Interspecific hybridization: Progress and Prospects*, Jones, M.P., Dingkuhn, M., Johnson D.E and Fagade, S.O. (Ed) Proceedings of the workshop: Africa/Asia Joint Research on Interspecific Hybridization between the Africa and Asian Rice Species *(O. glaberrima and O. Sativa)*. WARDA, Mbé, Bouaké, Côte d'Ivoire. p 43-59.

Lorz, H. and Wenzel, G. (2005). *Molecular Marker Systems in Plant Breeding and Crop Improvement*. Springer, New York, USA.

Lynch, M. and Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sinauer, Sunderland, United Kingdom.

Maji, A.J. (1998). Genetics of resistance to rice gall midge (*Orseolia* spp.) Paper presented during the WARDA/NARS group discussion and planning meeting on Bioecology and management of African rice gall midge, 12-16 October at IITA, Ibadan, Nigeria.

Mumm, R.H. (2007). Backcross versus forward breeding in the development of transgenic maize hybrids: theory and practice. *Crop Science* 47:164-171.

Moose, S.P. and Mumm R.H. (2008). Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiology* 147: 969-977.

Nelson, R.J., Naylor R.L. and Jahn M.M. (2004). The role of genomics research in improvement of ''orphan'' crops. *Crop Science* 44:1901-1904.

Painter, R.H., (1951). *Insect Resistance in Crop Plants*. The University Press of Kansas, Lawrence, USA.

Rojas, B.A. and Sprague, G.F. (1952). A comparison of variance components in corn yield trials: III. General and specific combining ability and their interaction with locations and years. *Agronomy Journal* 44:462-467.

Saad, I.M., Haq, N.M., Nasir, M.M. and Muhammad, M. (2004). General and specific combining ability studies in maize diallel crosses. *International Journal of Agriculture and Biology* 5:856-859.

Semagn, K., Bjørnstad, Å. and Ndjiondjop, M.N. (2006). Progress and prospects of marker assisted backcrossing as a tool in crop breeding programs. *African Journal of Biotechnology* 5:2588-2603.

Simmonds, N.W. (1981). *Principles of Crop Improvement.* Longman, London, United Kingdom. 408p.

Singh, B.D. (1993). *Plant Breeding*. 5th ed. Kalyani Publishers, Rajender Nagar, Ludhiana, India, p.102-104.

Singh, R.K., Singh, D.N., Singh, S.K. and Singh, H.N. (1994). Genetic variability and correlation studies in foreign commercial hybrids of sugarcane. *Agricultural Science Digest* 14:103-107.

Singh, B.D. (2005). *Plant Breeding*. Statistica. Stat Soft Inc., Kalyani Publishers, New Delhi, India. p 7.

Sleper, D.A. and Poehlman, J.M. (2006). *Breeding Field Crops*. 5th Ed., Wiley-Blackwell Publication, Iowa, USA. 424 p.

Sprague, G.F. and Tatum, L.A. (1942). General vs. specific combining ability in single crosses of corn. *Journal of American Society of Agronomy* 34: 923-932.

Temnykh, S.W., Park, D., Ayres, N., Cartingour, S., Hanck, N., Lipovich, L., Cho, Y.S. and Tshil, T. (2000). Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics* 100: 697-712.

Ullrich, S.E. (2007). Breeding field crops. *Crops Science* 47:900-904.

Vacaro, E., Fernandes, J., Neto, B., Pegoraro, D.G., Nuss, C.N. and Conceicao, L.H. (2002). Combining ability of twelve maize populations. *Pesquisa Agropecuara Brasileira* 37:67-72.

Varshney, R.V., Hoisington, D.A. and Tyagi, A.K. (2006). Advances in cereal genomics and applications in crop breeding. *Trends in Biotechnology* 24:490-499.

Weeden, N.R., Muehlbauer, F.J. and Ladizinsky, G. (1992). Extensive conservation of linkage relationships between pea and lentil genetic maps. *Journal of Heredity* 83:123-129.

Williams, C.T., (1998). Crop loss assessment methods for African rice gall midge, *Orseolia oryzivora* Harris and Gagné. Paper presented at WARDA/NARS Group Discussion and Planning Meeting `Bio-ecology and Management of the African Rice Gall Midge', October 12-16, 1998, IITA, Ibadan, Nigeria.

William, T.D. and James, R.F. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review* 108:346-369.

Chapter 4

Performance of progenies of F_2 and F_3 populations and parental lines of various degrees of resistance to Africa rice gall midge (AfRGM) exposed to AfRGM attack at different growth stages

Abstract

Plant responses to insect attack are influenced by environmental factors and the growth stage or the genotype of the plant. With African rice gall midge (AfRGM), this response is particularly pronounced because the pest is of seasonal and vegetative types. A study was conducted in two different filial populations at two growth stages. For each population, resistance index (RI) and disease incidence (DI) were calculated at 45 and 70 DAT, which represent the vegetative and reproductive stages, respectively. Four F_2 generations, involving ITA306 and BW348-1, Cisadane, TOS14519 and TOG7106 each with different levels of resistance to AfRGM, were evaluated. The plant material included 532, 413 and 479 F_2 progenies from ITA306xBW348-1, ITA306xCisadane and ITA306xTOS14519 crosses, respectively, in addition to 90 BC_1F_2 progenies from ITA306xTOG7106 cross. One F_3 generation of ITA306xTOS14519 cross, consisting of 649 F_3 families, was also tested. Progenies were grouped into four resistance categories: progenies that were more resistant than the check entry at 45 days after transplanting (DAT) and were still more resistant at 70 DAT; progenies that were more resistant at 45 DAT but became susceptible at 70 DAT; progenies that were susceptible at both 45 and 70 DAT; and finally, progenies that were susceptible at 45 DAT, but became resistant by 70 DAT. The first three categories were the most frequently observed and occurred in all combinations of crosses, while the last category was observed only for few progenies in the ITA306xTOS14519 F_2 and F_3 generations and for many progenies of the ITA306xBW348-1 cross. The last resistance category suggests that complex mechanisms are responsible for resistance to AfRGM. This may be a tolerance mechanism leading to the production of extra tillers after AfRGM attack, to compensate for the initial reduction in the number of tillers. With the number of non-infested tillers increasing relative to the number of infested tillers, the

percentage of infestation decreased afterwards. This may be combined with a physiological response by the plant to reduce levels of further AfRGM attack.

Introduction

The AfRGM, *Orseolia oryzivora* Harris and Gagné (Diptera: Cecidomyiidae) is a destructive insect pest of lowland rice in Africa (Umeh *et al*., 1992; Omoloye, 2000), causing yield losses between 45-80% in farmers' fields (Nwilene *et al*., 2006). Generally, the severity of the AfRGM attack is influenced by three major factors acting either independently or in combination. The first factor affecting AfRGM severity is related to the virulence of a given insect species or biotype, enabling it to overcome genes of resistance. Gunadasa *et al*. (2000) and Biradar *et al.* (2004) reported failure of resistance to the Asian rice gall midge (AsRGM) of all released rice lines in Sri Lanka and India, respectively. This was caused by the emergence of a new virulent AsRGM biotype, leading to the total and partial loss of harvest in farmers' fields, respectively.

AfRGM severity is also affected by environmental conditions, as reported by Williams *et al*. (2001), Omoloye (2002) and Nwilene *et al*. (2006). AfRGM severity is also affected by abiotic constraints and cultivation habits. Abiotic constraints include: (1) cloudy, humid weather with frequent rain or mist favours AfRGM build-up more than heavier, less frequent rainfall; (2) rainfed lowland and hydromorphic cropping systems favour AfRGM more than upland and mangrove cropping systems (Williams *et al*., 2002). Cultural practices include: (1) Wide range of planting dates, increasing the risk of AfRGM outbreaks. Late-planted fields are usually at higher risk; (2) Planting of new high-yielding varieties which are generally more susceptible to AfRGM than traditional landraces (Ukwungwu and Joshi, 1992); (3) Increase use of fertiliser, which has contributed to increased AfRGM infestation (Williams *et al*., 2002).

Site and field characteristics are also associated with higher risks of AfRGM. These concern planting method, wet-season weather pattern and presence of alternative hosts. Transplanted rice is at higher risk than direct-seeded rice. Alternate hosts such as *Oryza longistaminata* A. Chev. and Roehr. and volunteer rice, increases the risk, because they contribute to an early pest build-up in the wet season before rice crops have been planted.

Finally, the third factor affecting AfRGM severity is varietal response to AfRGM attack. Plant behaviour has been defined as a response to an event or environmental change during the

course of the lifetime of an individual (Silvertown and Gordon, 1989; Silvertown, 1998) and has been intensively studied and widely reviewed (Novoplansky, 2002; de Kroon and Mommer, 2005; de Kroon *et al*., 2005; Trewavas, 2005; Karban, 2008). These reviews have highlighted the ways in which plants collect resources, respond to physical stimuli, pathogens, pests and herbivory attacks. For example, plants develop leaves and roots non-randomly within their heterogeneous environments to allow them to actively modify their acquisition of essential nutrients, water and light. The growth or abscission of organs such as leaves or roots aim at recovering some of its investment to redirect resources to where the need is most critical (Pearcy and Sims, 1994; Ballare, 1999; Schmitt *et al*., 1999; Smith, 2000). Thus, plants respond to the attacks of pests and herbivores by changing many phenotypic traits (Karban and Baldwin, 1997; Agrawal *et al*., 1999), and initiating chemical, physiological, and morphological changes. It has also been reported that plants appear to anticipate, remember and communicate (Karban, 2008). In AfRGM, previous studies reported the differential reactions of lines against AfRGM attack (Ukwungwu and Alam, 1991; Ukwungwu and Joshi, 1992; Williams *et al*., 1998; Nwilene *et al*., 2002). This reaction is believed to be controlled by genetic factors involving antibiotic resistance, antixenotic resistance and tolerance. Understanding these factors may help breeders in their selection protocols when breeding for resistance to AfRGM. Moreover, no study has been reported previously on the resistance profiles of a segregating population derived from crosses involving two parental lines of different reactions to AfRGM. Furthermore, the response pattern between two growth stages of rice is also lacking. These aspects, in addition to evaluating the severity of AfRGM attack, and progenies' response patterns are investigated in the present study. The specific objectives were to: (1) determine the resistance reactions of progenies in two segregating generations against AfRGM attack at 45 and 70 DAT; (2) determine how much of the total infestation was achieved at 45 DAT for the susceptible and the resistant parents; and (3) discuss the different mechanisms of resistance observed in the resistance reactions. This study reports for the first time, the resistance reactions of a segregating population against AfRGM, and the extent of infestation at the end of the rice vegetative growth stage.

4-1 Materials and methods

4-1-1 Plant materials

The parent plants were constituted of ITA306, TOS14519, TOG7106 and BW348-1 which are susceptible, moderately resistant, resistant and tolerant to the AfRGM, respectively. The F_2 and F_3 progenies from the ITA306xTOS14519 crosses, comprising 479 and 649 individuals, respectively, were evaluated. The material also included 90 BC_1F_2 interspecific and 532 F_2 intraspecific progenies derived from the ITA306xTOG7106 and ITA306xBW348-1 crosses, respectively.

4-1-2 Methods

4-1-2-1 Screening of the F_2 and F_3 progenies

The progenies were screened in a tunnel paddy screenhouse at the International Institute for Tropical Agriculture (IITA) station at Ibadan from November 2009 to March 2010 for the F_3 generation, and from May 2010 to August 2010, for the F_2 generation. Seeds of the entries were sown in a nursery bed. The F_2 progenies and the F_3 families were transplanted 14 days after sowing (DAS) in a paddy screenhouse, in a 2-m row at a spacing of 200 mm within rows and between rows, with the two parents involved in the cross as a check after every 10 entries. The F_2 were transplanted on a single plant basis, while the F_3 were transplanted on family basis: each row of F_2 progenies consisted of twenty different individuals, while the row of the F_3 family consisted of twenty individuals of the same family. The experimental plot was flanked by an infestation band constituted of five rows of ITA306, also used as a susceptible check. The equivalent of 40 kg each of N, P_2O_5 and K_2O per ha was applied at transplanting, using NPK 15-15-15. Another 40 kg N/ha was applied 21 DAT, using urea. The plots were hand-weeded at 21 and 40 DAT. The experimental design was an augmented design with two checks, replicated every ten lines.

4-1-2-2 African rice gall midge damage assessment

AfRGM damage was assessed using previous established rating schemes, such as Resistance Index (RI) and Disease Incidence. RI-BA was calculated, as described in Chapter 2. The RI was calculated based on the resistant and the tolerant check's zone values to compute the

tillers' infestation relative to the resistant and tolerant checks. The formula used is from Williams (1998), with minor modifications, as follows:

$$\text{Resistance Index (RI)} = \frac{\text{\% of tillers with galls on test entry}}{\text{\% of tillers with galls on resistant/susceptible check variety}} \times 100$$

Where the % of tillers with galls on resistant/susceptible check variety is computed using zone values (ZV) which is the level of tillers' infestation of each resistant check, replicated after every 10 lines of test entry.

Disease incidence (DI) or percent hill infestation was calculated as follows:

$$\text{Disease incidence (DI)} = \frac{\text{Number of infested hills}}{\text{Total number of hills}} \times 100.$$

4-1-2-3 Determination of the resistance reaction of F_2 and F_3 generations of intraspecific and interspecific progenies to AfRGM

ITA306, TOS14519, TOG7106 and BW348-1 and the F_2 or F_3 progenies involving ITA306 and the three other parents were evaluated for their reaction to AfRGM at 45 and 70 DAT. Progenies' reactions were classified either as more resistant than the parent under test, identical to the parent, or less resistant than the parent. DI was calculated on a single plant basis for the F_2 generations whereas it was evaluated on family basis with the F_3 families.

4-1-2-4 Comparison of progenies and parents levels of infestation at vegetative and reproductive stages

The parents involved in the different crosses and their progenies were evaluated to state their level of infestation by AfRGM during the full growth cycle of rice. Percent infestation was calculated at 45 and 70 DAT, representing the vegetative and the reproductive stages, respectively. Comparison was made to state how much of the total infestation was achieved at the vegetative stage.

4-1-2-5 Statistical analysis

Data on the level of tiller and hill infestations were transformed prior to the analysis of variance, as described in Gomez and Gomez (1984), as follows: (1) percentage data in the range 30% to 70% were not transformed; (2) when all data in the range was between 0% and 50% or between 50% and 100%, the square root transformation was used; (3) for data that did not belong to any of the aforementioned ranges, the Arcsin transformation was used. XLSTAT package version 2009 was used for analysis of variance of both raw and the transformed data.

4-2 Results

The resistance, susceptibility and combined reactions in respect to the two evaluation dates reflect the behavioural patterns of the rice progenies. The identical reaction of progenies to each check entry was reported but was not discussed as it was considered as simple resistance.

4-2-1 Parental lines reaction against gall midge at 45 and 70 DAT

The reaction of parental lines was assessed using the standard evaluation system (SES) score. It shows two types of reaction patterns. Parental lines were either susceptible at both 45 and 70 DAT or resistant at both 45 and 70 DAT. The resistant (TOG7106) and the moderately resistant (TOS14519) parents were resistant at both growth stages while the susceptible (ITA306) and the two tolerant (BW348-1 and Cisadane) parents were susceptible at both 45 and 70 DAT. Apart from TOG7106 where the infestation at 70 DAT was lower (1.4%) than the level at 45 DAT (2.4), the percent infestation was always higher at 70 DAT than at 45 DAT among parental lines (Table 4-1).

Table 4-1: Percent mean infestation (PMI) of parental lines involved in cross combinations screened for African against rice gall midge resistance at 45 and 70 days after transplanting

Parental lines	45 Days after transplanting		70 Days after transplanting	
	PMI	SES rating	PMI	SES rating
ITA306	49.8	HS	62.4	HS
TOS14519	2.9	MR	3.5	MR
TOG7106	2.4	MR	1.4	R
BW348-1	49.4	HS	59.1	HS

Cisadane	34.1	HS	52.5	HS

PMI= Percent mean infestation, SES=Standard evaluation system, HS=Highly susceptible, MR=Moderately resistant, R=Resistant

4-2-2 F_2 progenies' reactions against AfRGM at 45 and 70 DAT

4-2-2-1 Behaviour of *Oryza sativa-Oryza glaberrima* (ITA306-TOG7106) progenies

No progenies were similar to TOG7106 at 70 DAT and none of the progenies that were more resistant than TOG7106 at 45 DAT (48, 50, 69 and 88) were confirmed at 70 DAT. However, four (40, 41, 63 and 89) of the 11 progenies that were similar to TOG7106 at 45 DAT were rated more resistant than TOG7106 at 70 DAT. All the progenies that developed less resistance than TOG7106 at 45 DAT were confirmed to be less resistant at 70 DAT. Of the 86 progenies rated less resistant than TOG7106 at 70 DAT, eleven were similar to TOG7106 at 45 DAT. Of the four types of reactions that a progeny can have, three were observed in this cross combination. The progenies were either: resistant at both 45 and 70 DAT (40, 41, 63 and 89), or resistant at 45 DAT and susceptible at 70 DAT (11, 36, 45, 53, 74, 82 and 85), or susceptible at both 45 and 70 DAT (Table 4-2). No progeny declared susceptible at 45 DAT became resistant at 70 DAT. The disease incidences were high amounting to 85.55% and 95.55% at 45 (77 infested hills) and 70 (86 infested hills) DAT, respectively (Table 4-2).

Table 4-2: Behavioural pattern of interspecific BC_1F_2 progenies derived from a cross between ITA306 and TOG7106 screened for resistance to AfRGM

Generation BC1F2	MRTg		ITg		LRTg									
	45 DAT	70 DAT	45 DAT	70 DAT	45 DAT					70 DAT				
ITA306-TOG7106	48	40	11	-	1	17	32	54	71	1	16	31	48	64	79
	50	41	36		2	18	33	55	72	2	17	32	49	65	80
	69	63	40		3	19	34	56	73	3	18	33	50	66	81
	88	89	41		4	20	35	57	75	4	19	34	51	67	82
			45		5	21	37	58	76	5	20	35	52	68	83
			53		6	22	38	59	77	6	21	36	53	69	84
			63		7	23	39	60	78	7	22	37	54	70	85
			74		8	24	42	61	79	8	23	38	55	71	86
			82		9	25	43	62	80	9	24	39	56	72	87
			85		10	26	44	64	81	10	25	42	57	73	88
			89		12	27	46	65	83	11	26	43	58	74	90
					13	28	47	66	84	12	27	44	59	75	
					14	29	49	67	86	13	28	45	60	76	
					15	30	51	68	87	14	29	46	61	77	
					16	31	52	70	90	15	30	47	62	78	
TOTAL	4	4	11	0	75					86					

MRTg= More resistant than TOG7106, ITg=Identical to TOG7106, LRTg = Less resistant than TOG7106, DAT= Days after transplanting

4-2-2-2 Behaviour of Intraspecific *O. sativa* subspecies *indica-O.sativa* subspecies *japonica* (ITA306-TOS14519) progenies

Overall, most progenies can be grouped into one of the two reactions. Progenies were either more or less resistant than TOS14519. Thus, only few progenies were similar to TOS14519 at both 45 (Progenies 29 and 210) and 70 (Progenies 59 and 90) DAT. Of the total of 479 progenies tested, Progenies 105 and 373 were more and less resistant than TOS14519 at 45 DAT respectively, whereas Progenies 50 and 427 developed the same reaction pattern at 70 DAT. However, in terms of behavioural pattern, four types of reactions were observed. Of the 107 resistant progenies at 45 DAT, 45 were confirmed resistant at 70 DAT, whereas the remaining 62 showed susceptibility at 70 DAT. Three hundred and seventy three entries were rated susceptible at 45 DAT, among which 366 were confirmed susceptible at 70 DAT, while 7 recovered to being resistant at 70 DAT. This ability to recover concerned Progenies 9, 39, 69, 99, 110, 201 and 456. The disease incidence was, however, 13% higher at 70 DAT than at 45 DAT. The numbers of non-infested hills were 97 and 34 at 45 and 70 DAT, respectively, corresponding to a disease incidence of 79.54% and 92.9%.

Table 4-3: Behavioural pattern of intraspecific F_2 *O. sativa* subsp. *indica-O.sativa* subsp. *japonica* progenies derived from a cross between ITA306 and TOS14519 screened for resistance to AfRGM

Generation	MRTs											ITs		LRTs	
F2	45 DAT							70 DAT				45 DAT	70 DAT	45 DAT	70 DAT
ITA306-TOS14519	**1**	61	133	180	264	351	414	1	***110**	260	403	29	59	TMTBL	TMTBL
	10	63	135	193	281	354	422	***9**	115	264	438	210	89		
	11	67	140	194	290	359	423	10	141	291	448				
	12	68	141	196	291	360	427	11	145	300	***456**				
	14	70	143	203	300	369	429	14	147	301	483				
	16	74	145	223	301	382	438	21	151	316					
	21	85	147	225	306	383	448	25	160	317					
	22	89	151	229	316	389	449	***39**	165	332					
	23	90	152	231	317	394	451	40	166	336					
	25	98	160	232	322	398	473	48	177	346					
	38	100	165	244	329	403	475	***69**	196	349					
	40	102	166	247	332	406	476	74	***201**	354					
	43	106	168	253	336	407	481	***99**	223	359					
	48	109	171	256	346	409	483	100	244	369					
	49	115	177	260	349	411	487	106	253	398					
TOTAL	105							50				2	2	373	427

MRTs= More resistant than TOS14519, ITs=Identical to TOS14519, LRTs = Less resistant than TOS14519, DAT= Days after transplanting, TMTBL= Too many to be listed. * represent lines that were not resistant at 45 DAT but became resistant at 70 DAT.

4-2-2-3 Behaviour of Intraspecific *O. sativa* subsp. *indica-O.sativa* subsp. *indica* (ITA306-BW348-1) progenies

Conversely to the other intraspecific and interspecific progenies, more intraspecific *O. sativa* subsp. *indica-O. sativa* subsp. *indica* progenies expressed similar reactions to the check (BW348-1). Twenty seven progenies reacted similarly to BW348-1, compared to 11, 2 and 8 progenies for ITA306xTOG7106, ITA306xTOS14519 and the F_3 families' progenies, respectively, at 45 DAT. Likewise, 59 progenies, similar to BW348-1, were matched by the resistance of 0, 2 and 5 progenies belonging to the same cross combinations at 70 DAT. Surprisingly, only one progeny (149) rated the same as BW348-1 at 45 DAT, maintained the same rating at 70 DAT. In addition, the number of progenies that were more resistant than BW348-1 (286 progenies) was much higher compared to the one of progenies of less resistance

than BW348-1 (222 progenies) at 45 DAT. Likewise, 293 progenies were more resistant compared to 181 progenies that were less resistant at 70 DAT. Among the four expected types of reactions, a high number of progenies that were susceptible at 45 DAT reverted to being resistant at 70 DAT. Sixty three (68.47%) of the 92 progenies concerned were less infested at 70 DAT than they were earlier at 45 DAT (Table 4-5 bold type), 6 (6.52%) had the same level of infestation (Table 4-5 type with asterisk), while 23 (25%) were more infested at 70 DAT than at 45 DAT (Table 4-5). DI was also very high, especially at 70 DAT, where only 6 hills out of the 532 plant stands were not infested, corresponding to a DI of 98.87%. At 45 DAT, where 85 hills were not infested, the DI was 80.01%.

Table 4-4: Behavioural pattern of intraspecific F_2 *O. sativa* subsp. *indica-O. sativa* subsp. *indica* progenies derived from a cross between ITA306 and BW348-1 screened for resistance to AfRGM.

Generation	MRB		IB										LRB	
F2	45 DAT	70 DAT	45 DAT			70 DAT							45 DAT	70 DAT
ITA306-BW348-1	TMTBL	TMTBL	25	***149**	403	2	79	158	277	400	513		TMTBL	TMTBL
			33	162	409	12	81	166	291	421	517			
			41	184	449	17	83	198	303	457	518			
			62	195	460	34	97	215	312	459	524			
			82	239	489	43	123	223	320	461	533			
			85	259	529	44	135	249	335	471	536			
			130	306	540	56	137	254	374	475	544			
			134	375		60	145	255	377	476	573			
			139	385		64	***149**	260	379	480	576			
			140	390		72	154	274	393	484				
TOTAL	286	293	27			59							222	181

MRB= More resistant than BW348-1, IB=Identical to BW348-1, LRB = Less resistant than BW348-1, DAT= Days after transplanting, TMTBL= Too many to be listed

Table 4-5: Phenotypic data on 92 susceptible progenies at 45 DAT and resistant at 70 DAT involving ITA306 and BW348-1 as AfRGM parental lines

Geno	IT 70	TT 70	PI 70	BWZV	RI70	IT 45	TT 45	PI 45	BWZV	RI45	Geno	IT 70	TT 70	PI 70	BWZV	RI70	IT 45	TT 45	PI 45	BWZV	RI45
11	18	37	49	86	0.6	4	6	67	40	1.7	**425**	11	21	52	58	0.9	4	6	67	33	2.0
14	5	11	45	83	0.5	1	2	50	29	1.8	**432**	9	18	50	78	0.6	4	5	80	33	2.4
22	6	13	46	59	0.8	2	3	67	50	1.3	**434**	9	17	53	60	0.9	5	5	100	38	2.7
28	6	19	32	76	0.4	3	5	60	20	3.0	**442**	5	15	33	43	0.8	3	6	50	21	2.3
39	11	24	46	85	0.5	6	8	75	70	1.1	**457**	3	9	33	38	0.9	1	2	50	29	1.8
44	5	16	31	56	0.6	2	3	67	60	1.1	**465**	7	15	47	58	0.8	3	4	75	33	2.3
45	3	8	38	65	0.6	3	4	75	50	1.5	**474**	5	9	56	60	0.9	2	3	67	38	1.8
46	8	14	57	62	0.9	4	5	80	56	1.4	**477**	4	14	29	38	0.7	2	4	50	29	1.8
66	7	13	54	62	0.9	3	4	75	56	1.3	**478**	4	10	40	57	0.7	2	3	67	0	VH
74	11	17	65	83	0.8	2	3	67	29	2.3	**484**	2	3	67	80	0.8	1	1	100	22	4.5
79	8	13	62	85	0.7	5	6	83	70	1.2	**485**	1	5	20	58	0.3	4	5	80	57	1.4
94	4	11	36	83	0.4	1	2	50	29	1.8	**504**	7	19	37	77	0.5	2	4	50	33	1.5
103	7	18	39	47	0.8	3	5	60	30	2.0	**515**	4	10	40	50	0.8	3	4	75	56	1.4
120	12	27	44	52	0.8	3	6	50	43	1.2	**528**	5	11	45	67	0.7	4	5	80	50	1.6
127	11	25	44	54	0.8	4	5	80	67	1.2	**538**	11	21	52	60	0.9	3	3	100	63	1.6
165	11	21	52	58	0.9	5	7	71	60	1.2	***47**	3	9	33	38	0.9	1	3	33	0	VH
179	4	11	36	60	0.6	2	3	67	50	1.3	***67**	2	10	20	38	0.5	1	5	20	0	VH
199	8	22	36	60	0.6	5	7	71	50	1.4	***352**	13	18	72	78	0.9	5	7	71	33	2.1
209	11	21	52	67	0.8	5	7	71	64	1.1	***413**	6	12	50	64	0.8	2	4	50	40	1.3
211	2	6	33	57	0.6	1	1	100	89	1.1	***463**	5	10	50	77	0.7	1	2	50	33	1.5
257	5	13	38	43	0.9	2	4	50	25	2.0	***525**	11	22	50	73	0.7	2	4	50	33	1.5
265	4	15	27	58	0.5	4	6	67	60	1.1	37	8	15	53	59	0.9	1	3	33	14	2.3
270	3	6	50	64	0.8	2	3	67	60	1.1	48	11	22	50	76	0.7	2	8	25	20	1.3
283	3	10	30	47	0.6	1	3	33	30	1.1	57	2	8	25	59	0.4	1	3	33	14	2.3
299	5	11	45	60	0.8	3	3	100	50	2.0	60	11	18	61	65	0.9	2	5	40	38	1.1
303	5	13	38	77	0.5	4	5	80	33	2.4	65	7	21	33	65	0.5	5	9	56	50	1.1
304	2	5	40	80	0.5	1	2	50	22	2.3	68	5	7	71	76	0.9	1	4	25	20	1.3
312	8	18	44	78	0.6	4	6	67	33	2.0	88	15	23	65	76	0.9	3	8	38	20	1.9
324	7	15	47	80	0.6	4	7	57	22	2.6	97	4	8	50	59	0.8	1	3	33	14	2.3
330	18	26	69	78	0.9	**6**	**7**	86	75	1.1	323	21	32	66	77	0.9	4	9	44	33	1.3
344	10	14	71	80	0.9	4	5	80	22	3.6	332	12	20	60	78	0.8	4	8	50	33	1.5
345	6	14	43	58	0.7	2	3	67	33	2.0	343	2	3	67	77	0.9	1	2	50	33	1.5
348	6	14	43	73	0.6	3	3	100	75	1.3	372	8	14	57	78	0.7	2	4	50	33	1.5
353	7	17	41	64	0.6	4	4	100	40	2.5	373	4	8	50	64	0.8	2	3	67	40	1.7
364	8	12	67	80	0.8	4	5	80	22	3.6	384	19	40	48	80	0.6	5	20	25	22	1.1
365	4	12	33	58	0.6	1	2	50	33	1.5	404	7	12	58	80	0.7	2	4	50	22	2.3
374	8	19	42	60	0.7	4	5	80	38	2.1	443	5	9	56	77	0.7	1	2	50	33	1.5
379	1	6	17	71	0.2	2	2	100	64	1.6	444	6	16	38	80	0.5	1	4	25	22	1.1
382	6	19	32	43	0.7	2	4	50	21	2.3	452	7	15	47	78	0.6	2	5	40	33	1.2
383	4	20	20	77	0.3	2	5	40	33	1.2	464	5	14	36	80	0.4	1	4	25	22	1.1
386	4	20	20	67	0.3	2	3	67	57	1.2	472	6	9	67	78	0.9	2	5	40	33	1.2
393	6	12	50	64	0.8	3	5	60	40	1.5	473	9	16	56	64	0.9	2	4	50	40	1.3
394	8	14	57	60	1.0	2	3	67	38	1.8	498	6	11	55	57	1.0	2	4	50	0	VH
405	2	6	33	58	0.6	1	2	50	33	1.5	533	4	7	57	62	0.9	2	4	50	25	2.0
412	8	16	50	78	0.6	3	3	100	33	3.0	539	10	20	50	67	0.8	2	6	33	17	2.0
414	4	10	40	60	0.7	2	4	50	38	1.3	542	7	13	54	92	0.6	2	4	50	40	1.3

IT=Infested tiller, TT=Total tiller, Pi=Percent infestation, BWZV=BW zone value, RI=Resistance index

4-2-3 Reaction of F_3 families from the cross of ITA306xTOS14519 against AfRGM at 45 and 70 DAT

Forty eight of the progenies rated more resistant than TOS14519 at 45 DAT. Of these, 13 (bold type in Column 2 and 3 Table 4-6) were confirmed at 70 DAT. Three (*366, *605 and *647) (Column 2 Table 4-6) were similar to TOS14519 at 70 DAT. None of the progenies similar to TOS14519 at 45 DAT were similar to TOS14519 at 70 DAT. However, 32 of the total of 48 progenies, rated more resistant than TOS14519 at 45 DAT, lost their resistance at 70 DAT. In addition, 580 out of the total of 649 progenies that ranked less resistant than TOS14519 at 45 DAT, were confirmed to be susceptible at 70 DAT. Of the 31 progenies more resistant than TOS14519 at 70 DAT, 13 were rated less resistant than TOS14519 at 45 DAT. Pest incidence was high at 70 DAT, ranging from 0 to 100 %, with a mean of 72.8 %. Its mean value was 58% at 45 DAT, also with a minimum and maximum of 0 and 100% respectively.

Table 4-6: Behavioural pattern of intraspecific F_3 *O. sativa* subsp. *indica-O. sativa* subsp. *japonica* progenies derived from a cross between ITA306 and TOS14519 screened for resistance to AfRGM

Generation F3	MRTs								ITs		LRTs	
	45 DAT					70 DAT			45 DAT	70 DAT	45 DAT	70 DAT
ITA306-TOS14519	12	163	**331**	**527**	**587**	18	204	**546**	28	35	TMTBL	TMTBL
	23	181	357	**542**	589	**34**	**234**	557	83	366		
	34	189	*366	543	593	49	298	565	120	601		
	70	196	390	**546**	598	55	**331**	**567**	133	605		
	82	210	**412**	554	*605	83	381	578	185	647		
	106	**234**	419	**567**	**613**	87	**412**	585	333			
	114	255	**464**	570	624	107	443	**587**	407			
	132	267	**484**	573	*647	120	**464**	606	447			
	150	268	499	578		**150**	**484**	**613**				
	153	289	526	582		181	**527**					
						192	**542**					
TOTAL	48					31			8	5	594	614

MRTs= More resistant than TOS14519, ITs=Identical to TOS14519, LRTs = Less resistant than TOS14519, DAT= Days after transplanting, TMTBL= Too many to be listed

4-2-4 Comparison of progenies and parental lines level of infestation at 45 and 70 DAT

4-2-4-1 Comparison among parental lines

Table 4-7 shows two types of reactions amongst the parental lines. Each landrace did not show significant difference (0.6 for TOS14519 and -1.0 for TOG7106) in the percent mean

infestation (PMI) at 45 and 70 DAT. However, for each cultivar, a clear difference was noticed. This difference ranged from 9.7 to 18.4, respectively, with BW348-1 and Cisadane. Within parental lines, the PMI at 45 DAT varied from 2.4 to 49.8, respectively, for TOG7106 and ITA306. No significant difference of infestation was observed between the two landraces TOS14519 and TOG7106 (2.9 vs. 2.4). Cultivars ITA306 (49.8) and BW348-1 (49.4) were equally infested at 45 DAT but were more infested than Cisadane (34.1). At 70 DAT, conversely to 45 DAT, a difference in PMI was noticed between landraces and cultivars. The PMI at 70 DAT for ITA306, TOS14519, TOG7106, BW348-1 and Cisadane were 62.4, 3.5, 1.4, 59.1 and 52.5, respectively. Overall, nearly 80% of the final infestation was reached by 45 DAT, the remaining 20% occurring between 45 and 70 DAT. The proportions of PMI reached at 45 DAT were 79.7, 83.5, 170, 83.6 and 65 for ITA306, TOS14519, TOG7106, BW348-1 and Cisadane, respectively. Surprisingly, the PMI at 70 DAT for the landrace TOG7106 was lower than its PMI at 45 DAT. The corresponding difference in percent infestation (DPI) at 45-70 DAT and the proportion of PMI reached at 45 DAT were 1% lower and 170%, respectively.

Table 4-7: Comparison of parental lines involved in cross combinations tested for AfRGM resistance at 45 and 70 DAT

Parental lines	Pedigree	PMI at 45 DAT	PMI at 70 DAT	DPI 45-70 DAT	Proportion of PMI reached at 45 DAT (%)
ITA306	Cultivar	49.8	62.4	12.6	79.7
TOS14519	Landrace	2.9	3.5	0.6	83.5
TOG7106	Landrace	2.4	1.4	- 1	170.7
BW348-1	Cultivar	49.4	59.1	9.7	83.6
Cisadane	Cultivar	34.1	52.5	18.4	65

PMI=Percent mean infestation, DPI= Difference percent infestation

4-2-4-2 Comparison among cross combinations

Each F_2 generation combination displayed differences in infestation at 45 and 70 DAT. The smallest DPI was 2.3% and was observed with the ITA306xCisadane cross while the highest value of 22.52% was noted for ITA306xTOG7106. Intermediate values of 3.5% and 7.5% were observed, respectively, for ITA306xTOS14519 and ITA306xBW348-1 crosses. Within cross combinations, no significant difference in PMI was observed between ITA306xTOS14519

(36.6) and ITA306xTOG7106 (36) at 45 DAT. PMI at 45 DAT, however, were different for ITA306xBW348-1 and ITA306xCisadane crosses, amounting to 43.9 and 54%, respectively. At 70 DAT, PMI was different among cross combinations and ranged from 40.1 to 58.6% with ITA306xTOS14519 and ITA306xTOG7106 crosses, respectively. Intermediate values of 51.4% and 56.3% were recorded for ITA306xBW348-1 and ITA306xCisadane crosses, respectively. In terms of level of infestation at 45 DAT, ITA306-Cisadane showed the highest infestation, followed by ITA306xTOS14519, then by ITA306xTOS14519 and finally by ITA306xTOG7106 (Table 4-8). The corresponding PMI values associated with these cross combinations were 95.9, 91.3, 85.5 and 61.5%, respectively.

Table 4-8: Comparison of F_2 generations involving parental lines of various degrees of resistance to AfRGM at 45 and 70 DAT

Cross	Pedigree	PMI at 45 DAT	PMI at 70 DAT	DPI 45-70 DAT	Proportion of PMI reached at 45 DAT (%)
ITA306-TOS14519	F_2	36.6	40.1	3.5	91.3
ITA306-TOG7106	BC_1F_2	36	58.6	22.5	61.5
ITA306-BW348-1	F_2	43.9	51.4	7.5	85.5
ITA306-Cisadane	F_2	54	56.3	2.3	95.9

PMI=Percent mean infestation, DPI 45-70 DAT= Difference percent infestation from 45 to 70 DAT

Table 4-9 shows a clear difference for the ITA306xTOS14519 F_3 generation in PMI at 45 and 70 DAT. It also shows a difference in infestation compared to the F_2 generations at 45 and 70 DAT. In each case, the PMI values recorded were lower than the ones observed in the F_2 generations. However, for a DPI 45-70 DAT of 12.9%, which is higher than the DPI values of three of the F_2 generations (ITA306-TOS14519, ITA306-BW348-1 and ITA306-Cisadane), the proportion of PMI reached at 45 DAT was the lowest compared to the corresponding proportion of PMI reached at 45 DAT of these three F_2 generations. The values associated with PMI at 45 and 70 DAT, DPI 45-70 DAT, and proportion of PMI reached at 45 DAT were 17.7, 30.6, 12.9 and 57.9%, respectively. Likewise, values associated with PMI at 45 DAT, PMI at 70 DAT, DPI 45-70 DAT and proportion of PMI reached at 45 DAT for TOS14519 in the F_3 screening conditions were 3.1, 5.1, 2 and 61.1%.

Table 4-9: Percent mean infestation (PMI) at 45 and 70 DAT, difference in percent infestation (DPI) 45-70 DAT and proportion of PMI reached at 45 DAT of the F_3 generation involving ITA306 and TOS14519

	Pedigree	PMI at 45 DAT	PMI at 70 DAT	DPI 45-70 DAT	Proportion of PMI reached at 45 DAT (%)
TOS14519	Landrace	3.1	5.1	2	61.1
ITA306-TOS14519	F_3	17.7	30.6	12.9	57.9

4-3 Discussion

In assessing resistance to AfRGM, more weight was given to the second scoring at 70 DAT because the infestation levels were higher, and the likelihood of an entry escaping infestation by chance was lower (Williams *et al*., 2001). The periods of 45 and 70 DAT, retained as a timing range to record damage caused by gall midge, corresponded to a two-week damage score. Indeed, the first generation of galls on the test entries started to appear at about 32 DAT and the second at about 55 DAT (Williams *et al.*, 2001). These two dates also correspond to the vegetative and the reproductive growth stage for a medium maturing rice variety. For the very early or very late maturing varieties, this period coincided either with their reproductive stages or their very early vegetative stages. In such cases, lines were not infested because the AfRGM is a vegetative insect pest. As such, the extra tillers produced in reaction to the initial AfRGM attack and late tillers produced by the very late maturing varieties, escaped from infestation.

Overall, four types of reactions were observed in the progenies. However, the parental lines exhibited only two types of behaviour**.** They were either resistant or susceptible at both growth stages studied. These types of reactions observed in parental lines were consistent with the fact that they are landraces or cultivars and are therefore fixed lines with constant and known reactions. Indeed, they were already designated as resistant, susceptible or tolerant to the AfRGM in previous studies conducted by Nwilene *et al*. (2002), Omoloye, (2002) and Williams *et al*. (2002). Our results therefore agree with these authors. The differential reaction of the parental lines could also be explained by three facts: (1) TOG7106 and TOS14519, being an African indigenous *O. glaberrima* and a traditional African *O. sativa* from the Gambia, respectively, may have co-evolved with the pest over years and have therefore accumulated horizontal resistance; (2) the *O. sativa* subsp. *indica* ITA306 is a newly released breeding line which has not co-evolved enough with the pest and is obviously poorly enriched in AfRGM

resistance genes, even though a previous study conducted by Omoloye *et al*. (2002) reported a certain amount of tolerance rather than susceptibility; and (3) the tolerant cultivars, Cisadane and BW348-1 from Indonesia and Sri Lanka, respectively, are popular cultivars widely cultivated in Nigeria, one of the area most affected by AfRGM. The tolerance mechanism could be assimilated to a horizontal resistance due to progressive accumulation of additive genes during the repetitive cropping of the cultivar, as documented by Robinson (1973).

On the other hand, progenies were either susceptible at both 45 and 70 DAT, or resistant at 45 DAT and became susceptible at 70 DAT, or were resistant at both 45 and 70 DAT, or were susceptible at 45 DAT but reverting to resistance at 70 DAT. The most frequent reactions were the first three, and were observed in all cross combinations, while the last one with susceptibility at 45 DAT and resistance at 70 DAT was rare. Progenies that were susceptible at both 45 and 70 DAT can be classified as purely susceptible, exhibiting no mechanism of resistance. Conversely, progenies that were resistant at 45 DAT but became susceptible at 70 DAT could be considered as partially resistant lines, exhibiting an early resistance or tolerance mechanism which was overcome with the increase of the gall midge pressure when progenies were exposed longer to the pest. Progenies that exhibited resistance at both growth stages could be considered as truly resistant lines, using a mixture of antixenotic and antibiotic resistance mechanisms, as reported by Vijaykumar *et al*. (2009). This persistent resistance throughout the growth stage may be due to repellants or morphological characteristics of the rice plant that prevented the adult insect from laying eggs or preventing the larvae from moving in the stem.

In the case of susceptible progenies that reverted to resistant progenies, three discrete explanations are the following: firstly, the progenies could have recovered from an early AfRGM damage if the attack was not sustained, as reported by Williams (1998). In this case, the non-sustainability of the attack could be due to low gall midge pressure due to a decrease in the number of insects because of a sex-ratio bias in the F_1 adult progeny, as reported by Omoloye (2006). The bias could have produced only males, leading in such a case, to few or no ovipositions; or could have favoured female prevalence, leading therefore to few matings or low fecundity. Secondly, the progenies could have produced more tillers than the newly infested ones so that the ratio between them was largely in favour of the new extra tillers produced (Omoloye *et al*., 2002). Thirdly, the damage was assessed relative to the infestation of the check entry. This check entry could have expressed a partial resistance mechanism which was overcome with longer exposure to AfRGM during the second growth stage. Thus, with the

level of tiller infestation of the check entry increasing, the RI proportionally decreased and the resistance of the entry appeared to increase. This phenomenon could be seen as a hypersensitive reaction aiming to suppress or reduce the AfRGM symptom on hills.

Our results also showed variation among cross combinations in the number of progenies susceptible at 45 DAT and resistant at 70 DAT. For the *O. sativa* subsp. *indica*-*O. sativa* subsp. *japonica* cross involving ITA306 and TOS14519, 7 progenies were concerned. This was due to an increase in the percent tiller infestation of the check entry while the level of tiller infestation at 45 and 70 DAT of test entries remained similar. Indeed, almost all (91.33 %) the PMI was already reached at 45 DAT, while a slight increase (0.58%) in TOS14519 PMI was noted at 70 DAT. Conversely, 14 susceptible progenies reverted to resistant within the F_3 families of the cross ITA306-TOS14519 with 3/5 of PMI reached at 45 DAT. The explanation could be that the 2/5 of the PMI remaining after 45 DAT either did not affect, or slightly affected, the 14 reverted progenies. The PMI remaining constant while the check entry recorded a slight increase in infestation, led to an apparent increase in resistance in these progenies. In addition, the F_3 being a more advanced generation, progenies were more fixed and therefore less prone to variation. The number of progenies reverting to resistance was twice as high in the F_3 combination (14) compared to the F_2 combination (7), probably because of the higher number of F_3 families (649) compared to the number of the F_2 progenies (479).

No progeny reverted to resistance for the *O. sativa* subsp. *indica*-*O. glaberrima* cross combination because with the check entry becoming more resistant at 70 DAT than at 45 DAT, the RI increased and therefore the progenies rated less resistant than the check entry. The type of resistance mechanism developed here by the resistant check entry TOG7106 could be a mixture of antixenosis and antibiosis, as described by Vijaykumar *et al*. (2009). For example, when TOG7106 was first attacked it reacted physiologically and stopped hence further AfRGM infestation. Subsequently, the tillering ability progressing with low or no infestation, hence the PMI decreased.

An unusually high number of susceptible progenies at 45 DAT that reverted to resistance at 70 DAT for the cross combination ITA306-BW348-1 was also observed. This situation can be explained by the following scenarios: (1) progenies were less infested at 70 DAT than they were at 45 DAT. Moreover, the PMI of the check entry was the highest at 70 DAT, apart from ITA306, the susceptible parental line. This led to a reduction of the RI and therefore implied that most progenies had more resistance than the check entry; (2) at equal level of PMI, the

high number of reverted progenies was due to the tolerance of the check entry. Indeed; because of its tolerance to AfRGM, BW348-1 can be highly infested without dying. At 70 DAT, the degree of infestation was higher than at 45 DAT due to longer exposure to the pest. The PMI of the check entry increased and therefore contributed to reduce the RI and subsequently increased the apparent level of resistance of the progenies. The same explanation is also valid when progenies susceptible at 45 DAT despite their higher infestation at 70 DAT than at 45 DAT, reverted to resistant at 70 DAT.

Despite the high infestation levels recorded on the tolerant cultivars used in our study, adequate numbers of seeds were harvested from F_2 plants from the ITA306xCisadane and ITA306xBW348-1 crosses. However, the F_2 plants from ITA306xTOG7106 and ITA306xTOS14519 crosses died before reaching maturity. Along with many authors (Louda, 1984; Parker, 1985; Young, 1985; Karban and Strauss, 1993; Omoloye *et al*., 1998*a*,*b*; Williams *et al*., 1998), we found Cisadane and BW348-1 to be tolerant of damaging levels of AfRGM, which led to harvesting seeds from their crosses, conversely to ITA306. Results also agree with Omoloye *et al*. (2002) on the tolerance of Cisadane, but contrary to his results, a higher level of tolerance was found for BW348-1 relative to ITA306.

The progenies can be grouped into four classes of behavioural patterns against African rice gall midge. The first class comprised progenies behaving similarly to ITA306, their susceptible parent, by being susceptible at 45 and 70 DAT. The second class comprised progenies behaving similarly to their resistant parent, TOS14519, by being resistant at 45 and 70 DAT. The third class bore progenies that were resistant at 45 DAT but became susceptible at 70 DAT. These progenies expressed partial resistance and could be improved using recurrent selection. Finally, the fourth class comprised of progenies that were susceptible at 45 DAT but which reverted to resistance at 70 DAT. Hence, they represent a good material suitable for initiating recurrent selection.

References

Agrawal, A.A., Tuzin, S. and Bent, E. (1999). *Induced Plant Defenses Against Pathogens and Herbivores*. APS Press, St Paul, USA.

Ballare, C.L. (1999). Keeping up with the neighbors: phytochrome sensing and other signalling mechanisms. *Trends in Plant Science* 4:97-102.

Biradar, S.K., Sundaram, R.M., Thirumurugan, T., Bentur, J.S., Amudhan, S., Shenoy, V.V., Mishra, B., Bennett, J. and Sarma, N.P. (2004). Identification of flanking SSR markers for a major rice gall midge resistance gene and their validation. *Theoretical and Applied Genetics* 109:1468-1473.

Gomez, K.A. and Gomez, A.A. (1984). *Statistical Procedures for Agricultural Research*. 2nd ed. International Rice Research Institute, Los Baños, Philippines, 680 p.

Gunadasa, A.H., Nugaliyade, L. and Edirisinghe, J.P. (2000). Development of gall midge resistant rice varieties. *Annals of the Sri-Lanka Department of Agriculture* 2:101-111.

Karban, R. and Strauss, S.Y. (1993). Effects of herbivore on growth and reproduction of their perennial host. *Erigeron glaucus Ecology* 74:39-46.

Karban, R. and Baldwin, I.T. (1997). *Induced Responses to Herbivory*. University of Chicago Press, Chicago, USA.

Karban, R. (2008). Plant behaviour and communication. *Ecology Letter* 11:727-739.

de Kroon, H. and Mommer, L. (2005). Root foraging theory put to the test. *Trends in Ecology and Evolution* 21:113-116.

de Kroon, H., Huber, H., Stuefer, J.F. and van Groenendael, J.M. (2005). A modular concept of phenotypic plasticity in plants. *New Phytologist* 166:73-82.

Louda, S.M. (1984). Herbivore effect on stature, fruiting and leaf dynamics of a native crucifer. *Ecology* 75:1379-1386.

Novoplansky, A. (2002). Developmental plasticity in plants: implications of non-cognitive behavior. *Evolution Ecology* 16:177-188.

Nwilene, F.E., Williams, C.T., Ukwungwu, M.N., Dakouo, D., Nacro, S., Hamadoun, A., Kamara, S.I., Okhidievbie, O., Abamu, F.J. and Adam A. (2002). Reactions of differential rice genotypes to African rice gall midge in West Africa. *International Journal of Pest Management* 48:195-201.

Nwilene, F.E., Harris, K.M., Okhidievbie, O., Onasanya, A., Sere, Y. and Ingelbrecht, I. (2006). Morphological diversity and genomic DNA fingerprinting of the African rice gall midge *Orseolia oryzivora* (Diptera: Cecidomyiidae) and of two other species of African Orseolia. *International Journal of Tropical Insect Science* 26:256-265.

Omoloye, A.A., Odebiyi, J.A., Singh, B.N. and Williams, C.T. (1998*a*). Trichome density and non-preference of rice cultivars for oviposition by the African rice gall midge, *Orseolia oryzivora* Harris and Gagné. *African Journal of Plant Protection* 8:29-36.

Omoloye, A.A., Odebiyi, J.A., Sing, B.N. and Williams, C.T. (1998*b*). Effects of antibiosis in resistant rice cultivars on the biology of African rice gall midge, *Orseolia oryzivora* Harris and Gagné. *African Journal of Plant Protection* 8:37-48.

Omoloye, A.A. (2000). Observation of a single sex F_1 progeny in the African rice gall midge, *Orseolia oryzivora* Harris and Gagné. *Annals of Agricultural Science* 1:6-13.

Omoloye, A.A., Odebiyi, J.A., Williams C.T., and Singh, B.N. (2002). Tolerance indicators and responses of rice cultivars to infestation by the African rice gall midge, *Orseolia oryzivora. Journal of Agricultural Science* 139:335-340.

Omoloye, A.A. (2006). Sex ratio bias in the F_1 adult progeny of African rice gall midge, *Orseolia oryzivora* Harris and Gagné. (Dipt., Cecidomyiidae). *Journal of Applied Entomology* 130:349-355.

Parker, M.A. (1985). Size-dependent herbivore attack and the demography of an arid grassland shrub. *Ecology* 66:850-860.

Pearcy, R.W. and Sims, D.A. (1994). Photosynthetic acclimation to changing light environments: scaling from the leaf to the whole plant. In: *Exploitation of Environmental Heterogeneity by Plants*. Caldwell, M.M. and Pearcy, R.W. (Eds.) Academic Press, San Diego, USA. p 145-174.

Robinson, R.A. (1973). Horizontal resistance. *Review of Applied Mycology* 52: 483-501.

Schmitt, J., Dudley, S.A. and Pigliucci, M. (1999). Manipulative approaches to testing adaptive plasticity: phytochrome-mediated shade-avoidance responses in plants. *The American Naturalist* 154:43-54.

Silvertown, J. (1998). Plant phenotypic plasticity and non-cognitive behaviour. *Trends in Ecology Evolution* 13:255-256.

Silvertown, J. and Gordon, D.M. (1989). A framework for plant behavior. *Annual Review of Ecology System* 20:349-366.

Smith, H. (2000). Phytochromes and light signal perception by plants: An emerging synthesis. *Nature* 407:585-591.

Trewavas, A. (2005). Plant intelligence. *Naturwissenschaften* 92:401-413.

Ukwungwu, M.N. and Alam M.S. (1991). Field resistance of rice cultivars to the gall midge *Orseolia oryzivora* in Nigeria. *International Rice Commission Newsletter* 40:42-43.

Ukwungwu, M.N. and Joshi, R.C. (1992). Distribution of the African rice gall midge, *Orseolia oryzivora* Harris and Gagné and its parasitoids in Nigeria. *Tropical Pest Management* 38:241-244.

Umeh, E.D.N., Joshi, R.C. and Ukwungwu, M.N. (1992). Biology, status and management of rice insect pests in Nigeria. *Crop Protection* 11:408-413.

Vijaykumar, L., Chakravarthy, A.K., Patil, S.U. and Rajanna, D. (2009). Resistance mechanism in rice to the midge *Orseolia oryzae* (Diptera: Cecidomyiidae). *Journal of Economic Entomology* 102:1628-1639.

Williams, C.T. (1998). Crop loss assessment methods for African rice gall midge, *Orseolia oryzivora* Harris and Gagné. Paper presented at WARDA/NARS Group Discussion and Planning Meeting: Bio-ecology and Management of the African Rice Gall Midge. IITA, Ibadan, Nigeria.

Williams, C.T., Ukwungwu, M.N., Singh, B.N., Okhidievbe, O. and Nnabo, J. (1998). Farmer managed trials in southeast Nigeria to evaluate the rice variety Cisadane and estimate yield losses caused by the African rice gall midge, *Orseolia oryzivora* Harris and Gagné. *International Journal of Pest Management* 45:117-124.

Williams, C.T., Okhidievbie, O., Harris, K.M. and Ukwungwu, M.N. (1999). The host plant range, annual cycle and parasitoids of the African rice gall midge *Orseolia oryzivora* (Diptera: cecidomyiidae) in central and southeast Nigeria. *Bulletin of Entomological Research* 89:589-597.

Williams, C.T., Ukwungwu, M.N., Singh, B.N. and Okhidevbie, O. (2001). Assessment of host plant resistance in *Oryza sativa* to the African rice gall midge, *Orseolia oryzivora* Harris and Gagné, with a description of a new method for screening under artificial infestation. *Journal of Applied Entomology* 125:341-345.

Williams, C.T., Harris K.M., Ukwungwu, M.N., Nacro, S., Dakouo, D., Nwilene, F.E., Singh, B.N. and Okhidievbie, O. (2002). *African Rice Gall Midge Research Guide*. WARDA and CAB International, Wallingford, United Kingdom. 28 p.

Young, T.P. (1985). *Lobelia telekii* herbivory, mortality and size at reproduction: variation with growth rate. *Ecology* 66:1879-1883.

Chapter 5

Diversity of agro-morphological traits in an F_3 population segregating for resistance to African Rice Gall Midge (AfRGM), determined through heritability estimates and correlation among traits.

Abstract

The extent to which genetic gain is possible is measured with the degree of variability that is controlled either by genetic or environmental factors. The success of most crop improvement programmes largely depends upon the heritability of desirable traits. Heritability estimates were calculated for 15 major traits in a segregating F_3 rice population, in order to predict the genetic gain associated with each trait, together with the resistance to AfRGM. Correlation of Pearson and Shannon-Weaver diversity index for these traits were also calculated. The high level of genetic variability was confirmed by the high diversity index (from 0.76 to 0.88) recorded for 8 of the 15 traits studied, and the intermediate values of 0.56, 0.57, 0.69, 0.64 and 0.49 recorded on tillering ability (Til), number of days to booting (DB), number of days to first heading (DFH), culm length (CL) and panicle length (PanL) respectively. Only ligule length (LigL) (0.17) and thousand grains weight (TGW) (0.22) were less diverse. Correlation among traits was only high (0.71 to 0.99) in traits associated either with plant maturity (DB, DFH, DH, and DM), plant stature (PH, CL), or plant dry weight (PLW, FLW and FLL). However, almost all the traits under study were highly heritable. Broad-sense heritability (H^2) estimates were high for the penultimate leaf length (0.99), the penultimate leaf width (1), the flag leaf length (0.99), the flag leaf width (1), the ligule length (0.99), the tillering ability (0.99), the number of days to booting (0.95), the number of days to first heading (0.96), the number of days to heading (0.89), the number of days to maturity (0.98), the culm length (0.99), the plant height (0.99), the panicle length (0.95), the secondary branching (0.95) and thousand grains weight (0.71). Conversely, narrow-sense heritability estimates were very low (nearly 0) in PLL, FLL, Lig, DB, DFH, DM and SB; low (at most 0.267) in PLW, FLW, DH and PH with a medium value of 0.727 for TGW. Inheritance of the traits investigated in the present study was therefore under dominance and epistatic gene effects rather than additive genetic effects and can therefore be improved using pedigree breeding schemes, along with breeding for AfRGM.

Introduction

Rice is the world's most important crop and a major source of nutrition for approximately 66% of mankind (Schalbroeck, 2001; Jena and Mackill, 2008). In West Africa, where 7.2 million tons of milled rice is consumed on average annually, the production for the same period is 13.2 million tons of paddy rice (AfricaRice, 2008) due to many constraints including African rice gall midge (AfRGM). AfRGM is one of the major pests causing extensive damage to both lowland and upland rice in Africa. The most effective way to control the pest is genetic resistance. This has led to breeding strategies for varietal resistance. Varietal resistance involves either looking for varieties already resistant to the pest or transferring resistance from an existing source of resistance into varieties without resistance. The first option was investigated and high level of AfRGM resistance was found only amongst *Oryza glaberrima* accessions (Nwilene *et al*, 2006, Maji *et al*., 1998) which were not very productive, or prone to shattering (Jones, 1997). Moreover, *O. glaberrima* was not susceptible enough, and *Oryza sativa* was not resistant enough, to use the new breeding strategy, tagged Marker-Assisted Breeding (MAB). The most susceptible *O. glaberrima* documented was susceptible at 60% (Nwilene *et al*, 2006, Maji *et al*., 1998). A few *O. sativa* lines, including the traditional *sativa* landrace TOS14519 from the Gambia, have moderate resistance to AfRGM (Nwilene *et al.*, 2002, Ukwungwu *et al*., 1998, Williams *et al.*, 1999), thus opening the opportunity of gene transfer to explore the second option of managing the pest. Crosses, involving ITA306 and TOS14519, were therefore made. In the early filial generations, variability among progenies was so pronounced that understanding how much of this variability was due to genetic factors and how much the traits responsible for the variability are correlated, is required for an understanding of whether or not important traits can be transmitted along with AfRGM resistance. This approach aims to prevent progenies from performing less well than their best parent. The proportion of the variability due to genetic factors is termed as heritability. Additive components on one hand, and dominance and epistatic components, on the other hand, are referred to as narrow-sense heritability and broad-sense heritability, respectively. According to the inheritance or the gene action of the trait of interest, one type of heritability is dominant over the other. However, in a case of multiple traits of unknown genetic pattern, both narrow-sense and broad-sense heritability are important. Two options were investigated: the first option involved estimating heritability of the segregating traits, to determine which of the

environmental or genetic factors is responsible for the variability. In the latter case, it was to determine which of the components of the genetic factors were preponderant over the others. Narrow-sense and broad-sense heritability were estimated for this issue. Narrow-sense heritability referred to additive genetic variance, while broad-sense heritability dealt with dominance and epistatic variances. The second approach was to determine the principal components of the traits responsible for the variability using correlation and cluster analyses as means.

The objectives of this study were to: (1) Determine the origin of the variability observed in the segregating population; (2) Determine the two types of heritability estimates and state which of the narrow-sense or broad-sense heritability is preponderant over the other; (3) Estimate correlation between traits involved in the study, in order to design appropriate breeding and selection strategies, that will best exploit any variability found, in order to create farmer-preferred varieties.

5-1 Material and method

5-1-1 Plant material

The plant material was constituted by F_2 and F_3 generations, each consisting of 652 individuals derived from a cross between ITA306, a high-yielding *Oryza sativa* subsp. *indica* cultivar, that is susceptible to AfRGM, which was used as the female parent, and TOS14519, an *Oryza sativa* subsp. *japonica* cultivar used as the male parent resistant to AfRGM. F_1 plants were validated prior to advancing them to F_2 and F_3. Routine care included watering and weeding until plant maturity and data collecting.

5-1-2 Field site description

The experiment was conducted from October 2009 to July 2010, in between the small dry season and the main rainy season, at the Africa Rice Center (AfricaRice) research station at Togoudo, Benin. Soil composition was of a fine-loamy type. Rainfall, evaporation, wind speed, solar radiation, temperature and relative humidity during the period of the experiment are summarised in Table 5-1.

Table 5-1: Meteorology data recorded on AfricaRice Togoudo Station from October 2009 to July 2010

Para-meters.	Units	Oct. 09	Nov. 09	Dec. 09	Jan. 10	Feb. 10	Mar. 10	Apr. 10	Ma. 10	Jun. 10	Jul. 10	Total	Mean
TRF	Mm	117.2	8.3	0.0	13.5	45.0	177.7	121.6	273.5	286.1	230.1	1273.0	127.3
TEv	Mm	110.1	124.1	517.4	112.9	141.7	126.7	128.1	91.0	75.7	98.4	1526.0	152.6
ARF	mm/day	3.8	0.3	0.0	0.4	1.6	5.7	4.1	8.8	9.5	7.4	41.7	4.2
Aev	mm/day	3.7	4.1	16.7	3.6	5.1	4.4	4.6	3.5	2.9	3.6	52.2	5.2
AWS	Km/h.	5.7	11.4	5.1	4.9	7.4	5.9	5.6	4.4	4.3	6.9	61.6	6.2
ASR	*	16.5	17.6	16.0	13.7	17.9	16.9	17.8	16.0	12.6	14.8	159.7	16.0
AmT	°C	22.8	22.8	23.3	23.6	25.1	25.3	24.8	24.0	23.5	22.9	238.1	23.8
AMT	°C	30.4	32.7	33.8	33.4	32.1	33.7	33.5	32.4	30.8	30.3	323.1	32.3
Am RH	%	70.6	52.8	48.2	53.1	49.5	54.4	55.9	63.7	73.6	72.4	594.1	59.4
AM RH	%	99.1	95.6	95.2	95.1	94.7	94.6	95.3	95.3	99.2	96.7	960.8	96.1
MAT	°C	28.3	27.7	28.6	28.7	30.3	29.4	29.2	28.2	27.1	26.6	284.1	28.4
MRH	%	84.9	74.2	71.7	74.5	72.2	74.5	75.6	79.5	86.4	84.5	778.0	77.8

RF: Total rainfall, TEv: Total evaporation, ARF: Average rainfall, AEv: Average evaporation, AWS: Average wind speed, ASR: Average solar radiation, AmT: Average minimum temperature, AMT: Average maximum temperature, MAT: Mean average temperature, MRH: Mean relative humidity, *:Gm-cal.cm^{-2}.day^{-1}

5-1-3 Methods

5-1-3-1 Production of the segregating population of the study

652 F_2 plants were grown in 12-litre plastic buckets at the AfricaRice Research Station at Cotonou, Benin. Basal fertilisation of 1g of each of N, P_2O_5 and K_2O was added in each pot at transplanting, using NPK 15-15-15 as source. One gram of urea was added at the late vegetative stage. The pots were hand-weeded whenever necessary. Upon maturity, F_2 plants were harvested and labelled for the F_3 study.

5-1-3-2 F_3 characterisation

The experimental design was laid in an augmented design (Federer, 1956) with two checks repeated in each block. ITA306 and TOS14519, which were the respective susceptible and the resistant parents to the population, were used as checks. Plot size for each accession was 600 mm x 3300 mm with 3 rows, with a spacing of 300 mm between and within rows and a distance of 400 mm maintained between plots. Direct seeding at the rate of two plants per hill was done. Plot number per block was 52 (50 progenies + 2 checks). A total number of 676 plots [650 + (2x13)] were shared between 13 blocks with a distance of 2 m between blocks. Fertiliser

used was NPK (10-18-18) as basal application, at the rate of 200 kg/ha during land preparation. Urea was also applied at the rate of 65 kg.ha^{-1} as top-dressing, first at tillering and a second time at booting. Observation was made on the ten hills of the central row.

5-1-3-3 Field investigated characters

A total of 15 quantitative characters were measured at various growth stages using the descriptors of wild and cultivated rice (*Oryza spp*.) (Bioversity International *et al.,* 2007). The quantitative characters were scaled randomly and included the penultimate leaf length (PLL), the penultimate leaf width (PLW), the flag leaf length (FLL), the flag leaf width (FLW), the ligule length (LigL), the tillering ability (Til), the number of days to booting (DB), the number of days to first heading (DFH), the number of days to heading (DH), the number of days to maturity (DM), the culm length (CL), the plant height (PH), the panicle length (PanL), the secondary branching (SB) and thousand grains weight (TGW).

5-1-3-4 Statistical analysis

The data was analysed using the XLSTAT 2010 package. Analysis of variance (ANOVA) and descriptive statistics were performed on characters studied to determine the broad-sense heritability (H^2). Mid-parent offspring regression mean was used to determine the narrow-sense heritability (h^2), where the mid-parent value of each block was plotted against the offspring mean of the same block. Broad-sense heritability was estimated using the block as the repetition and as environments. H^2 was calculated according to the variance component method suggested by Becker (1984) as follows:

$$H^2 = \frac{Vg}{Vp}$$

$$Vg = \frac{(rVi + Ve) - Ve}{r} = Vi$$

$$Vp = Vg + Ve$$

Where:

H^2: Broad-sense heritability, Vg: Genotypic variance, Ve: Environmental variance, Vi: Variance for progenies, r: number of repetition

Simple correlations among traits under study were also analysed using the formula of Gomez and Gomez (1984) as follows:

$$rxy = \frac{Xi - X \quad Yj - Y}{\sum Xi - X\ ^2 \sum (Yj - Y)^2}$$

Where:

r_{xy} : Correlation between the traits X and Y

X_i : X value

Y_j: Y value

X : Mean value of character X

Y : Mean value of character Y

The Shannon-Weaver diversity index was also computed using MS Excel according to the formula proposed by Jain *et al*. (1975) as follows:

$$H' = -\sum_{i=1}^{k} Pi \log 2 Pi$$

Where k is the number of phenotypic classes for a character and Pi is the proportion of the total number of entries (N) in the ith class. Each value of H'was standardised by dividing it by its maximum value ($\log_2 k$), in order to keep the values in the range of 0-1 (Abdi *et al*., 2002). For the calculation of H' the data was transformed into phenotypic classes. The scales for recording characters to phenotype frequencies were as indicated by the descriptors of wild and cultivated rice (Oryza sp.) (Bioversity International *et al.*, 2007) and Kayode *et al*., 2007, and are summarised in Table 5-2.

Table 5-2: List of characters and their respective classes used for the evaluation of diversity in an F_3 population segregating against AfRGM

Characters	Abbreviation	Code	Description	Characters state
Penultimate leaf length (mm)	PLL	1	< 400	Short
		5	410-600	Medium
		9	> 600	Long
Penultimate leaf width (mm)	PLW	1	< 12	Narrow
		5	12-16	Medium
		9	> 16	Broad
Flag leaf length (mm)	FLL	1	< 400	Short
		5	410-600	Medium
		9	> 600	Long
Flag leaf width (mm)	FLW	1	< 12	Narrow
		5	12-16	Medium
		9	> 16	Broad
Ligule length (mm)	LigL	1	< 5	Short
		3	5-15	Medium
		5	> 15	Long
Tillering ability	Til	1	> 25	Very high
		3	20-25	Good
		5	10-19	Medium
		7	5-9	Low
		9	< 5	Very low
Days to booting (days)	DB	1	< 45	Very early
		3	46-60	Early
		5	61-75	Medium
		7	76-90	Late
		9	> 90	Very late
Days to first heading (days)	DFH	1	< 60	Very early
		3	61-75	Early
		5	76-90	Medium
		7	91-105	Late
		9	> 105	Very late
Days to heading (days)	DH	1	< 75	Very early
		3	76-90	Early
		5	91-105	Medium
		7	106-120	Late
		9	> 120	Very late
Days to maturity (days)	DM	1	< 100	Very early
		3	101-115	Early
		5	116-130	Medium
		7	131-145	Late
		9	> 145	Very late

Table 5-2 continued

Characters	Abbreviation	Code	Description	Characters state
Culm length (mm)	CL	1	< 500	Very short
		2	510-700	Very short to short
		3	710-900	Short
		4	910-1050	Short to intermediate
		5	1060-1200	Intermediate
		6	1210-1400	Intermediate to long
		7	1410-1550	Long
		8	1560-1800	Long to very long
		9	> 1800	Very long
Plant height (mm)	PH	1	< 800	Very short
		2	810-1000	Very short to short
		3	1010-1200	Short
		4	1210-1350	Short to intermediate
		5	1360-1500	Intermediate
		6	1510-1700	Intermediate to tall
		7	1710-1850	Tall
		8	1860-2100	Tall to very tall
		9	> 2100	Very tall
Panicle length (mm)	PanL	1	< 110	Very short
		3	110-200	Short
		5	210-300	Medium
		7	310-400	Long
		9	> 400	Very long
Secondary branching	SB	1	< 45	Low
		3	46-60	Medium
		5	61-75	Dense
		7	76-90	Very dense
		9	> 90	Clustered
Thousand grains weight (g)	TGW	1	< 20	Light
		5	20-40	Intermediate
		9	> 40	Heavy

Source: Adapted from Bioversity, IRRI and WARDA (2007) and Kayode *et al*. (2007)

5-2 Results

5-2-1 Variances analysis

Report on variances analysis refers to environmental variance arising from differences between replications of each non-segregating genotype (parents and checks) and concerned only inter-block variances because checks were not replicated within blocks to allow such analysis. Inter-

block variances were generally low for all traits studied and are displayed in Table 5-3. The maximum and minimum values of 57.39 and 0.003 were recorded for number of DH and FLW. The corresponding standard deviation and coefficient of variation (CV) were 7.58% and 0.8%, respectively, for the maximum while they were the same value of 0.05, for the minimum. Apart from DB, DFH and DH, the standard errors recorded were less than 1. The CVs were also very low, at most 11%, and were observed in both SB and TGW. The lowest CV was 0.03% and was observed in PLL, DM, CL and PH variance estimates (Table 5-3).

Table 5-3: Descriptive statistics and interblock variances for all the traits under study

Statistic	PLL (mm)	PLW (mm)	FLL (mm)	FLW (mm)	LigL (mm)	Til (No.)	DB (days)	DFH (days)
Min	323.5	07.5	211.5	08.5	15.50	10.00	61.00	68.50
Max	362.5	09.5	254.5	10.0	19.00	14.00	77.00	83.50
Mean	344.5	08.3	230.1	09.2	17.00	12.23	65.42	73.31
Var	1.41	0.004	1.55	0.003	0.96	1.19	16.33	15.02
Std	1.19	0.06	1.25	0.05	0.98	1.09	4.04	3.88
CV	0.03	0.07	0.05	0.05	0.06	0.09	0.06	0.05
SE	0.33	0.02	0.35	0.01	0.27	0.30	1.12	1.07

Table 5-3: continued

Statistic	DH (days)	DM (days)	CL (mm)	PH (mm)	PanL (mm)	SB (No.)	TGW (g)
Min	79.50	104.00	830.0	1065.0	230.0	16.00	21.30
Max	103.50	114.00	930.0	1175.0	260.0	22.50	31.85
Mean	89.35	108.85	878.5	1116.9	238.5	18.50	25.51
Var	57.39	12.10	7.85	8.94	0.77	4.25	8.22
Std	7.58	3.48	2.80	2.99	0.88	2.06	2.87
CV	0.08	0.03	0.03	0.03	0.04	0.11	0.11
SE	2.10	0.96	0.78	0.83	0.24	0.57	0.80

Number of observation for each trait: 13, PLL=Penultimate leaf length, PLW= Penultimate leaf width, FLL=Flag leaf length, FLW=Flag leaf width, LigL=Ligule length, Til=Tillering, DB=Days to booting, DFH=Days to first heading, DH=Days to heading, DM=Days to maturity, CL=Culm length, PH=Plant height, PanL=Panicle length, SB=Secondary branching, TGW=1000 grains weight.

5-2-2 Genetic variation analysis

The diversity index was generally high, ranging up to 0.88 with a mean value of 0.66 (Table 5-4). However, a couple of particularly low values of 0.17 and 0.22 were noticed. The variables can be grouped into three diversity classes: two extreme classes, one with a high index value of

at least 0.8, and another with a low index value of at most 0.22. A medium class also existed bearing index values of 0.5 to 0.8. PLL, FLL, FLW, DH, DM, PH and SB recorded high index values amounting to 0.88, 0.88, 0.8, 0.79, 0.87, 0.81 and 0.83, respectively. Medium index values were recorded for PLW, Til, DB, DFH, CL and PanL. The corresponding indices were 0.76, 0.56, 0.57, 0.69, 0.64 and 0.49 respectively. Conversely, the low values of 0.17 and 0.22 were noticed for LigL and TGW, respectively. The lowest index values recorded for the high and medium classes were 0.79 and 0.49, and were recorded for DH and PanL, respectively.

Table 5-4: The Shannon-Weaver diversity (H') index according to 15 agro-morphological traits recorded on an F_3 population segregating against AfRGM

Variables	H'	Class	Variables	H'	Class
Penultimate leaf length	0.88	High	No of days to heading	0.79	Medium
Penultimate leaf width	0.76	Medium	No of days to maturity	0.87	High
Flag leaf length	0.88	High	Culm length	0.64	Medium
Flag leaf width	0.80	Medium	Plant height	0.81	High
Ligule length	0.17	Low	Panicle length	0.49	Medium
Tillering ability	0.56	Medium	Secondary branching	0.83	High
No of days to booting	0.57	Medium	Thousand grains weight	0.22	Low
No of days to first heading	0.69	Medium			
H' mean			0.66		

H'=Shannon index, Low ≤ 0.5; Medium > 0.5 to ≤ 0.8; High > 0.8

5-2-3 Simple correlations between traits

Results of simple correlations among traits are presented in Table 5-5. The p-values associated with these correlation coefficients were highly significant, ranging from 0 ($p<0.0001$) to 0.009 for an alpha level of 5%. High correlation was observed amongst traits, following a grouping pattern with respect to a particular metabolism mechanism of the rice plant. Thus, traits involved in dry weight mass (PLW, FLW and FLL) were highly correlated, as well as traits involved in maturity (DB, DFH, DH, and DM), and plant stature (PH, CL). In this regard, DM, DB, DFH and DH were positively and highly correlated with each other. Thus, DM was positively correlated to DH, DFH and DB with a correlation coefficient of at least 0.93. Likewise, DH was strongly correlated with DB and DFH while DFH and DB were correlated at 0.96. PH was also strongly correlated to CL at a correlation rate of 0.99. Slightly lower correlation coefficients of 0.81, 0.72 and 0.61 were also observed for PLW and FLW, FLL and PLL, and between LigL and FLL, respectively (Table 5-5).

Intermediate correlation coefficients ranging from 0.3 to 0.6 were observed and mainly concerned: (1) PLL and the other traits, excluding FLL, PanL and SB; and (2) FLL and the other traits, excluding LigL, Til, CL and TGW (Table 5-5). In this view, SB was correlated in the same range to CL, PH and PanL (0.34, 0.38 and 0.38, respectively), while PanL showed the same level of correlation to CL, PH, PLL and FLL (0.33, 0.46, 0.3 and 0.33 respectively) (Table 5-5 line 13).

Low correlation coefficients were also found and were evenly distributed among traits. These values were as low as 0 (FLW and TGW) and as high as 0.28 (FLW and CL) for the positive correlation. Poor and negative correlations were depicted among traits and involved DB, DFH, DH and DM, which were poorly correlated to PanL, SB and TGW. The correlation amongst some traits cannot be clearly stated as the p-value was not significant. Examples were FLW with DB, DFH, DH and DM. The same trend was observed for PH and CL with DB, DFH and DH, whose p-values were not significant (for an alpha of 0.05).

Table 5-5: Simple phenotypic correlation among traits recorded on an F_3 *O. sativa* population segregating for AfRGM resistance

Variables	PLL	PLW	FLL	FLW	LigL	Til	DB	DFH	DH	DM	CL	PH	PanL	SB
PLW	**0.46**													
FLL	**0.72**	**0.27**												
FLW	**0.45**	**0.81**	**0.40**											
LigL	**0.58**	**0.14**	**0.61**	**0.15**										
Til	**0.35**	**0.57**	**0.20**	**0.52**	**0.17**									
DB	**0.47**	**0.18**	**0.39**	-0.03	**0.32**	**0.18**								
DFH	**0.43**	**0.16**	**0.35**	-0.06	**0.29**	**0.18**	**0.96**							
DH	**0.44**	**0.15**	**0.39**	-0.04	**0.30**	**0.20**	**0.95**	**0.93**						
DM	**0.48**	**0.19**	**0.41**	0.00	**0.31**	**0.22**	**0.95**	**0.94**	**0.93**					
CL	**0.53**	**0.34**	**0.27**	**0.28**	**0.35**	**0.34**	0.08	0.07	0.05	**0.09**				
PH	**0.54**	**0.33**	**0.30**	**0.30**	**0.37**	**0.34**	0.05	0.04	0.03	0.06	**0.99**			
PanL	**0.30**	**0.12**	**0.33**	**0.23**	**0.23**	**0.15**	**-0.13**	**-0.14**	**-0.11**	**-0.12**	**0.33**	**0.46**		
SB	**0.28**	**0.29**	**0.29**	**0.40**	**0.12**	**0.14**	**-0.19**	**-0.20**	**-0.20**	**-0.18**	**0.34**	**0.38**	**0.38**	
TGW	-0.07	-0.04	**-0.09**	0.00	-0.01	0.06	**-0.10**	**-0.10**	**-0.08**	**-0.12**	0.04	0.05	0.07	-0.02

Bold values are significant and the corresponding p-values were significant at an alpha level of 0.05 ranging from 0 to 0.009. Non-bold=non-significant, PLL=Penultimate leaf length, PLW=Penultimate leaf width, FLL=Flag leaf length, FLW=Flag leaf width, LigL=Ligule length, Til=Tillering, DB=Number of days to booting, DFH=Number of days to first heading, DH=Number of days to heading, DM=Number of days to maturity, CL=Culm length, PH=Plant height, PanL=Panicle length, SB=Secondary branching, TGW=1000 grains weight

5-2-4 Cluster analysis

A principal component analysis (PCA) conducted on the 15 quantitative characters did not show any separation pattern among the progenies. The maximum variability explained by any two factor combinations was 58.24%, and was observed between the two first factors (Figure 5-1). The contributions of each trait to these factors were as follows: traits contribution ranged from 0.15% (TGW) to 13.66% (PLL) for Factor 1, while it ranged from 0.35% (TGW) to 11.25% (DFH) for Factor 2. PLL (62.95%), PLW (36.2%), FLL (24.42%), FLW (45.68%), LigL (31.9%), Til (38.31%), DB (70.66%), DFH (42.09%), DH (54.45%), DM (73.58%), CL (46.38%), PH (52.43%), PanL (44.85%), SB (28.6%) and TGW (68.89%) contributed the most to Factors 10, 11, 9, 11, 6, 8, 14, 12, 12, 13, 15, 15, 7, 6 and 5, respectively. An agglomerative hierarchical clustering using the principal component scores (PC score) of the first five factors explained up to 82% of the total variability. The automatic truncate proposed by the XLSTAT version 2010 depicted three groups of low intra-group variance (Figure 5-2). The three classes were statistically similar and contained 235, 211 and 206 individuals, respectively, for Classes 1, 2 and 3 (Table 6). The variance component estimates were low, amounting to 9.763, 6.797 and 6.144 for Classes 1, 2 and 3, respectively (Table 5-6).

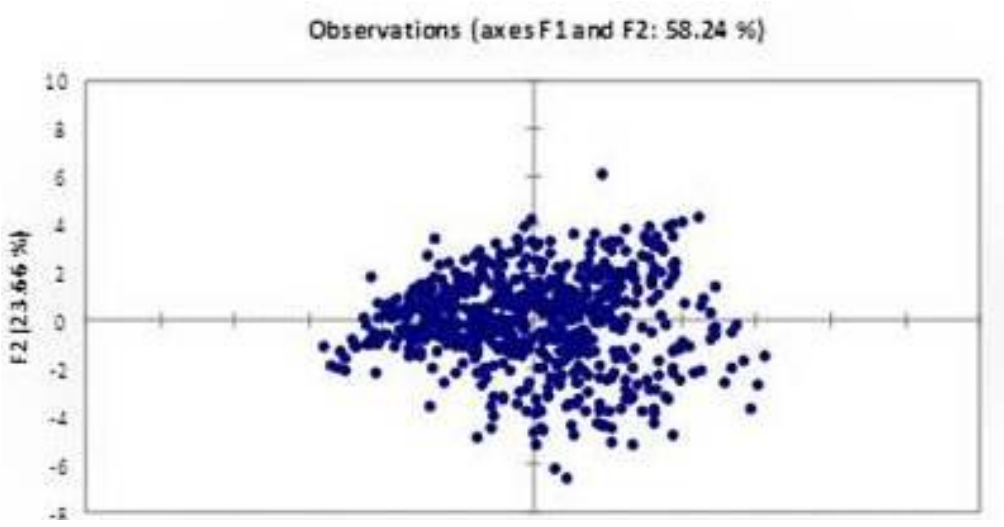


Figure 5-1: Principal component analysis of 15 agro-morphological traits segregating in a F_3 population based on Factor 1 and Factor 2

Table 5-6: Progenies distribution in cluster and within-class variance estimates obtained by agglomerative hierarchical clustering

Class	1	2	3
Number of individuals	235	211	206
Within-class variance	9.763	6.797	6.144
Minimum distance to centroid	0.956	0.788	0.554
Average distance to centroid	2.960	2.461	2.330
Maximum distance to centroid	5.886	6.451	4.997

5-2-5 Broad-sense and narrow-sense heritability estimates

Globally, heritability estimates were of two types: they were either very high or very low. They are shown in Table 5-7 and Figure 5-3. Broad-sense heritability (H^2) estimates were generally high, conversely to the narrow-sense heritability (h^2) estimates, which depicted low values. However, a surprising high narrow-sense heritability estimate of 0.72 was recorded for TGW (Table 5-7). Apart from this value, the highest h^2 was 0.267, followed by 0.258, 0.225 and 0.159, for DH, PH, PLW and CL, respectively (Figure 5-3). The other characters showed h^2 estimate values of at most 0.10. Among them, PLL, FLL, LigL, DFH and PanL showed very low h^2 estimates amounting to nearly 0%. Narrow-sense heritability estimates associated with these traits were 0.001, 0.003, 0.000, 0.000 and 0.001 respectively (Table 5-7 last column).

Conversely to narrow-sense heritability, broad-sense heritability (H^2) estimates were surprisingly high for all traits (Table 5-7). The lowest value was 0.71 and was recorded for TGW. Two traits, PLW and FLW, recorded the maximum value of 1.0 for H^2. Five other traits, PLL, FLL, Til, CL and PH, also recorded nearly the maximum (0.99). Amongst the eight traits remaining, two traits, DH and TGW, had H^2 estimates below 0.9 (0.89 and 0.71, respectively). Three characters (DB, PanL and SB) had an H^2 value of 0.95. Two traits (LigL and DFH) had H^2 values of 0.96. DM had an H^2 value of 0.98 (Table 5-7).

Table 5-7: Broad-sense and narrow-sense heritability estimates

Character	V_{prog}	V_e	V_p	H^2	h^2
Penultimate leaf length (PLL)	147.23	1.41	148.64	0.99	0.001
Penultimate leaf width (PLW)	0.06	0.004	0.06	1.00	0.225
Flag leaf length (FLL)	167.53	1.55	169.08	0.99	0.003
Flag leaf width (FLW)	0.06	0.003	0.06	1.00	0.105
Ligule length (LigL)	25.78	0.96	26.74	0.96	0.000
Tillering ability (Til)	173.76	1.19	174.95	0.99	0.101
No of days to booting (DB)	352.10	16.63	368.73	0.95	0.064
No of days to first heading (DFH)	389.04	15.02	404.06	0.96	0.000
No of days to heading (DH)	441.82	57.39	499.21	0.89	0.267
No of days to maturity (DM)	494.59	12.10	506.69	0.98	0.031
Culm length (CL)	541.50	7.85	549.35	0.99	0.159
Plant height (PH)	612.87	8.94	621.81	0.99	0.258
Panicle length (PanL)	14.35	0.77	15.12	0.95	0.001
Secondary branching (SB)	84.98	4.25	89.23	0.95	0.073
Thousand grain weight (TGW)	19.86	8.22	28.08	0.71	0.727

Vprog=Variance for progenies, Ve=Environmental variance, Vp=Phenotypic variance, H^2=Broad sense heritability estimate, h^2=Narrow sense heritability estimate

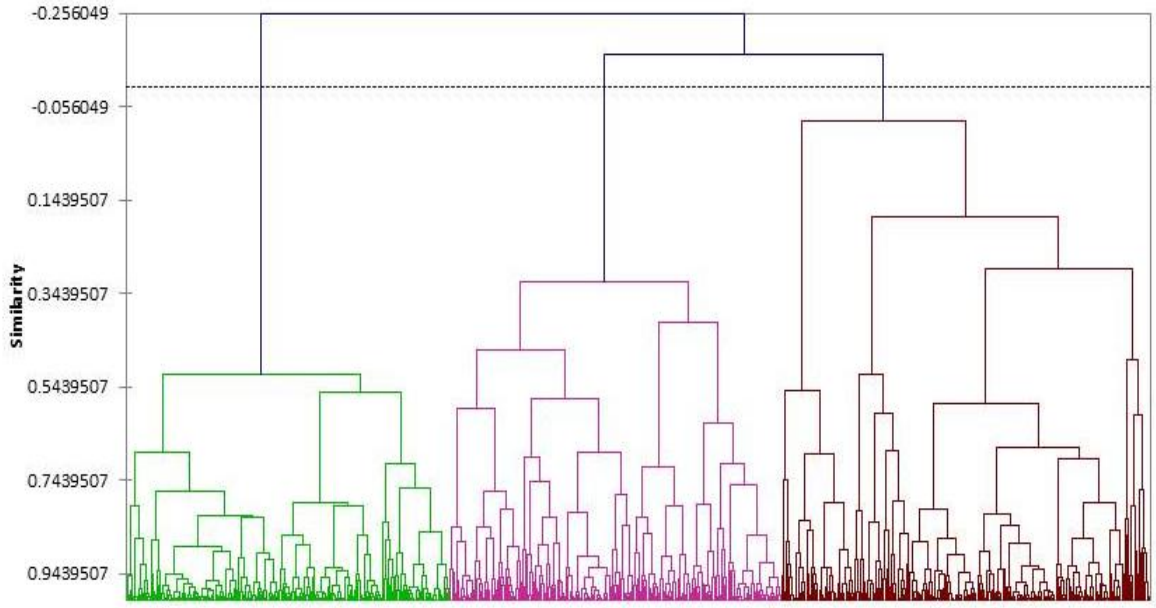
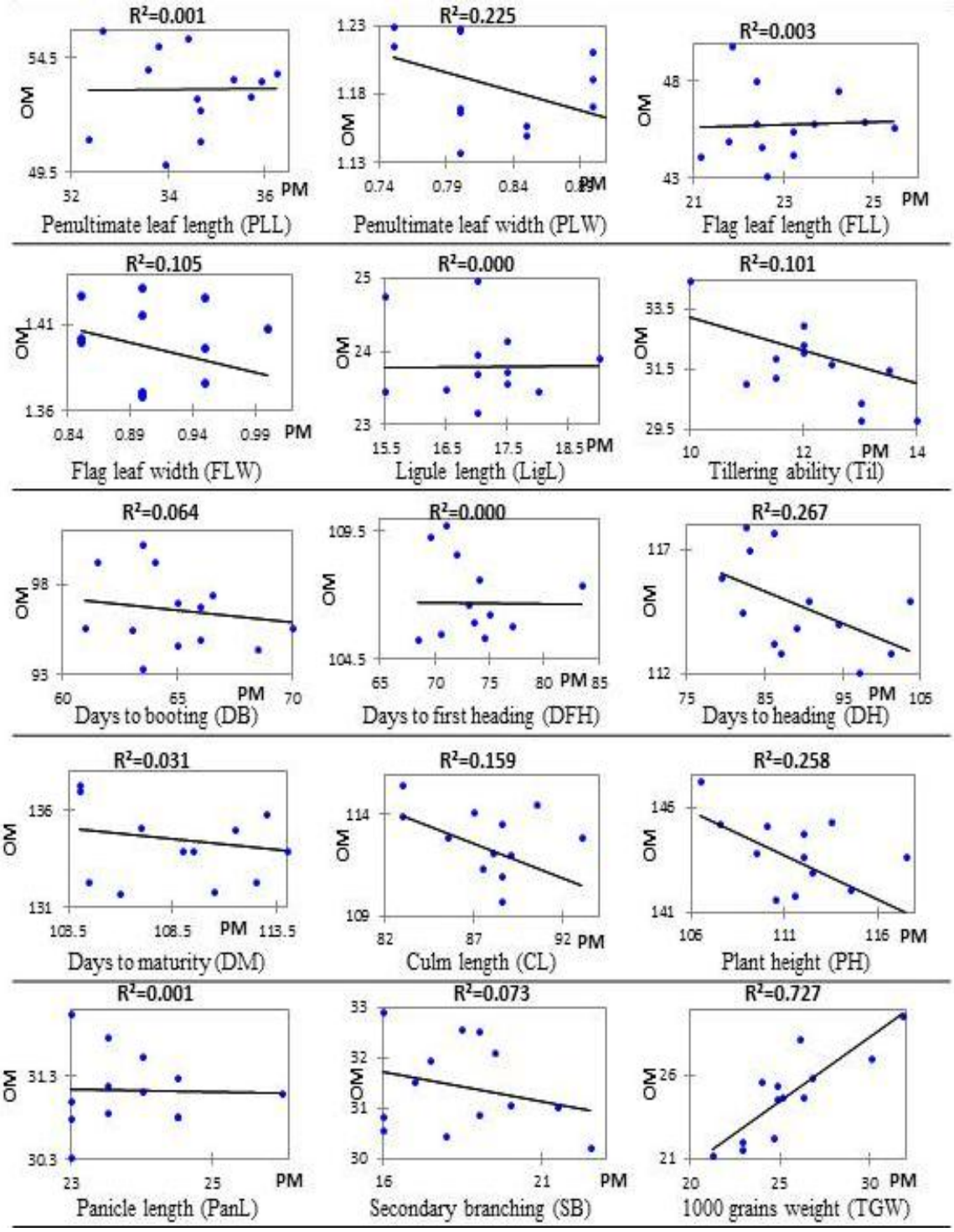


Figure 5-2: Dendrogram of 652 progenies obtained by agglomerative hierarchical clustering using the similarity coefficient of Pearson based on UPGM method



Legend: ● = Active, ——— = model, OM=Offspring mean, PM=Parent mean

Figure 5-3: Narrow-sense heritability (h^2) estimates for 15 agro-morphological traits measured in a F_3 population segregating against AfRGM resistance

5-3 Discussion

The genetic variability among progenies was checked by simple statistics (means, ranges, standard deviation and coefficient of variation), and was confirmed by the high Shannon index recorded for most of the traits under investigation apart from TGW and LigL. The fact that the inter-block variance was low, supported the assertion that the variability observed in the population was not due to environmental causes but was likely due to genetic causes (Villanueva and Moro, 2001). Our results also depicted variation in heritability estimates, irrespective of the trait measured, or the type of heritability computed. This agrees with Falconer (1989), who claimed that heritability estimates for agro-morphological traits in rice are sparse and contradictory, depending upon the genetic materials and methods used. Moreover, estimation of heritability are controversial, the magnitude of heritability estimates for a highly heritable trait is considered more than 0.5, and for a medium heritable trait it is between 0.2 and 0.5 (Singh, 2005). Indeed, in this present study, the broad-sense heritability was high in contrast to the narrow-sense heritability. This trend was observed in earlier studies conducted by Kato (1990), Rabiei *et al*. (2004), Vanaja and Babu (2006) and Fahliani *et al*. (2010). Broad-sense heritability was estimated by the variance component method developed by Becker (1984). Therefore, as reported above, the environmental variance being low, the phenotypic variance was proportionally low and its ratio under the genotypic variance was high. The other obvious explanation of this phenomenon of high broad-sense heritability (H^2) and low narrow-sense heritability (h^2) is that the increase of broad-sense heritability estimate was because of higher genetic variance rather than low environmental variance, as confirmed by the high Shannon index measured for the population for several of the traits investigated. The last explanation for this phenomenon is that, it could have been due to the control of the inheritance of the different traits by genetic dominance and/or epistatic effects, rather than by additive gene effects. This implies that the traits are under the control of either few genes or major genes. Indeed, the lowest broad-sense heritability was recorded for TGW. This character is believed to be a yield contributing character, which is also known to be a complex mechanism involving multiple genes (Abdus *et al*., 2009). High heritability estimates in their broad-sense were observed for DFH, PH, PanL, Til and TGW, indicating limited potential for genetic improvement of these characters through selection. Similar results have been reported

by Panwar *et al.* (1997), Sarawgi *et al.* (2000), Gannamani (2001), Sao (2002) and Rita *et al*. (2009).

All the explanations provided above also applied to narrow-sense heritability: low additive variance, high genetic variance and multiple genes interaction, acting independently or in combination resulting in low narrow-sense heritability values. However, one more reason for the low narrow-sense heritability estimates observed was due to the method of calculation. Narrow-sense heritability was estimated using the mid-parent offspring mean regression, which is sensitive to the normal distribution or the skewness of the data. Indeed, the preliminary normality tests prior to the regression showed that, apart from PLL and FLL, the other traits were not normally distributed, according to the normality tests of Shapiro-Wilk, Anderson-Darling, Lilliefors and Jarque-Bera from XLStat 2010 (data not shown).

Our results showed that leaf characters, LigL, maturity-related traits (excluding DH) and PanL had very low narrow-sense heritability, while plant stature-related traits (CL and PH) had low heritability. Therefore, breeding for these traits is difficult because low heritability suggests that the phenotype is not highly correlated with the genotype (Fahliani *et al*., 2010). Conversely, high heritability estimates imply a fairly easy selection of character due to a close correspondence between the genotype and phenotype, due to a relatively smaller contribution of the environment to the phenotype (Umadevi *et al*., 2009). According to Singh (2005), the contribution of environmental conditions for low heritable traits is relatively high. This assertion is in contradiction with our result, which showed low environmental variance despite the low narrow-sense heritability recorded for most of the studied traits. This could have been due to the preponderance of non-additive gene effects over additive gene effects, rather than the impact of environmental conditions, a possibility that was reinforced by the high broad-sense heritability recorded for these traits. The low broad-sense heritability estimate of PanL is, however, in agreement with previous results reported by Toriyama and Futsuhara (1958), Wells and Faw (1978), Jennings *et al*. (1979), Wayne and Dilday (2003), Sabu *et al*. (2009) and Fahliani *et al*. (2010). Likewise, low narrow-sense heritability of Til observed in our study agrees with previous reports documented by Hoshikawa (1989), Moldenhauer *et al*. (1994), Gravois and Helms (1996) and Fahliani *et al*. (2010).

TGW was found to have both high narrow (0.727) and broad (0.71) sense heritability estimates as similarly reported by Umadevi *et al*. (2009) with a broad-sense heritability estimate value of

0.77. Hence, TGW will respond to selection because a high level of narrow sense heritability suggests large additive variance. Moreover, the PLL (0.99) and the PLW (1) showed high broad-sense heritability estimates as confirmed by a similar study (0.88 vs. 0.79) reported by Mohapatra and Mohanty (1985).

Overall, correlations reported in the present study were of three types. Traits were either positively correlated or negatively correlated or non-correlated. For the first two types, the extent of the correlation was either high or low. High positive correlations were observed within traits involved in rice maturity, plant stature and weight dry mass. Rice cycle and maturity depend upon the vegetative phase because the reproductive and maturity phases remained quite constant (N'cho, 1992). Therefore, the cycle duration is influenced by the early steps of the reproductive stage including the booting stage. The sooner the booting occurs, the sooner plants mature. Likewise, the sooner the booting occurs, the shorter the Days to first heading and the Days to heading and consequently, the earlier the plant matures. However, some few cases were observed where, despite the precocity of the booting, the subsequent steps were delayed. This fact can be explained by environmental constraints. Indeed, according to Shah *et al*. (1999), the opening of the spikelets depends primarily on the prevailing atmospheric temperature, the light intensity and other climatic conditions. Plant height was highly correlated to Culm length, while it was moderately correlated to Panicle length. Plant height in rice is a complex character and is the end product of several genetically controlled factors called internodes (Cheema *et al*., 1987); their elongation is directly responsible for the Plant height. The moderate correlation could be explained by the fact that the plant, mobilising large stem reserve to support panicle growth, consequently auto-regulates the size of the panicle (Ali *et al*., 2000). Our results also surprisingly showed that Thousand grains weight was not correlated to PanL as reported by Zafar *et al*. (2004). Abbasi *et al.* (1995) claimed that although it contributes positively, PanL is not the only factor responsible for higher grain yield because a plant can have a maximum PanL, but due to lower grain fertility, can exhibit lower grain yield. Therefore, Panicle length alone does not determine the high grain yield because traits such as grain size, grain shape, higher numbers of tillers/plant, longer panicles and greater number of grains/panicle ultimately contribute to higher grain yield (Akram *et al*., 1994). TGW was not correlated with Panicle length, probably because, to reduce substantial yield losses, plants reduce their height to improve their resistance to lodging and therefore influence Panicle length

(Abbasi *et al*., 1995). Overall, TGW was either very poorly and negatively correlated or very poorly and positively correlated or not significantly correlated, irrespective of the traits. This observation was different from the findings documented earlier by Khedikar *et al*. (2004), Borbora *et al*. (2005), and Panwar and Mashiat Ali (2007). These authors observed that TGW was highly and positively correlated to the number of panicles per plant and secondary branches per panicle. In the same trend, highly significant negative correlation with PH was also observed by Rita *et al.* (2006), Panwar and Mashiat, (2007) and Umadevi *et al*. (2009).

The high and positive correlation between the penultimate leaf and the flag leaf width and length (traits involved in dry weight mass) is probably due to the fact that, being the top leaves, their greater length and width will result in better photosynthesis, providing more resource to meet the demands of the plant. A study conducted by Jain *et al*. (2001) on the flag leaf and penultimate leaf revealed that the amount and activity of RuPBCO, a chlorophyll precursor, in flag leaf is two-fold higher as compared to the penultimate leaf. They also mentioned that a reduction in leaf area led to less RuPBCO, inducing low photosynthesis activity for a low weight dry mass of the whole plant.

Apart from TGW, the PLL was consistently well correlated to the other traits studied. This trend was observed for PLW and FLL with an exception that the correlations were globally less than PLL. This fact is probably attributable to the involvement of these dry weight mass components in the metabolism of the traits under study. Indeed, as aforementioned in the Jain *et al.* (2001) study, these components are very active in photosynthesis, which conditions plant stature, maturity and dry weight mass.

This study reveals that the phenotypic traits investigated here had either low levels of narrow sense heritability, or high levels of broad sense heritability and cannot be improved through selection. Only Thousand grains weight (TGW) can respond to selection. However, this trait is either poorly negatively correlated, or poorly positively correlated or not correlated to the other traits investigated. Hence, selection for yield should be oriented toward yield component related traits such as number of tiller per plant, number of primary and secondary branching and panicle length.

References

Abbasi, F.M., M.A. Sagar, M. Akram and M. Ashraf, (1995). Agronomic and quality traits of some elite rice genotypes. *Pakistan Journal of Science and Industry Research* 38:348-350.

Abdus, S.K., Imran, M. and Ashfaq, M. (2009). Estimation of genetic variability and correlation for grain yield component in rice (*Oryza sativa* L.). *American-Eurasian Journal of Agriculture and Environmental Science* 6:585-590.

Abdi, A., Bekele, E., Asfaw, Z. and Teshome, A. (2002). Patterns of morphological variation of sorghum (Sorghum bicolor (L.) Moench) landraces in qualitative characters in North Shewa and South Welo, Ethiopia. *Hereditas* 137:161-172.

Africa Rice Center (WARDA). (2008). *Africa Rice Trends 2007*. Africa Rice Center, Cotonou, Benin. 84 p.

Akram, M., Abbasi, F.M., Sagar, M.A., and Ashraf, M. (1994). Increasing rice productivity through better utilization of germplasm. In: *Proceeding of a National Seminar on Genetic Resources of Cereals and Their Utilization.* Islamabad, Pakistan. p 107-14.

Ali, S.S., Jafri, S.J.H., Khan, T.Z., Mahmood, A. and Butt, M.A. (2000). Heritability of yield and yield components of rice. *Pakistan Journal of Agricultural Research* 16: 89-91.

Becker, W.A. (1984). *Manual of Quantitative Genetics*. 2nd Edn, Longman, London, United Kingdom.

Bioversity International, IRRI and WARDA. (2007). *Descriptors for Wild and Cultivated Rice (Oryza spp.)*. Bioversity International, Rome, Italy; International Rice Research Institute, Los Baños, Philippines; WARDA, Africa Rice Center, Cotonou, Benin.

Borbora, T.K., Hazarika, G.N. and Medhi, A.K. (2005). Correlation and path analysis for plant and panicle characters in rice. *Crop Research* 30:215-220.

Cheema, A.A., Awan, M.A. and. Iqbal, J. (1987). Improvement of plant height architecture in Basmati Rice. *Pakistan Journal of Agricultural Research* 8:371-374.

Fahliani, R.A., Khodambashi, M. Houshmand S. and Arzani. A. (2010). Estimation of the heritability of agro-morphological traits in rice (*Oryza sativa* L.) using F2:3 families. *African Journal of Agricultural Research* 5:1297-1303.

Falconer, D.S. (1989). *Introduction to Quantitative Genetics.* Longman, London, United Kingdom.

Federer, W.T. (1956). Augmented designs. *Hawaiian Planter's Record* 55:191-208.

Gannamani, N. (2001). *Study of Heterosis and Combining Ability by Utilizing Cytoplasmic Genetic Male Sterility and Fertility Restoration System in Rice (Oryza sativa L.),* M.Sc. (Ag.) Thesis, GAU, Raipur, India.

Gravois, K. and Helms, R.S. (1996). Seeding rate effects on rough rice yield, head rice, and total milled rice. *Agronomy Journal* 88:82-84.

Hoshikawa, K. (1989). *The Growing Rice Plant: An Anatomical Monograph*. Nobunkyo, Tokyo, Japan.

Jain, S.K., Qualset, C.O., Bhatt, G.M. and Wu, K.K. (1975). Geographical patterns of phenotypic diversity in a world collection of durum wheats. *Crop Science* 15:700-704.

Jain, V., Guru, S.K. and Abrol Y.P. (2001). Responses of penultimate and flag leaves of Wheat to different nitrogen supply. *Biologia plantarum* 44:373-378.

Jena K.K. and Mackill D.J. (2008). Molecular marker and their use in marker-assisted selection in rice. *Crop Science* 48:1266-1276.

Jennings, P.R., Coffman, W.R, Kauffman, H.E. (1979). *Rice Improvement*. International Rice Research Institute, Manila, the Philippines.

Jones, M.P., Audebert, A., Mande, S. and Aluko, K. (1997). Characterization and utilization of *Oryza glaberrima* Steud. in upland rice breeding. In: *Interspecific Hybridization: Progress and Prospects*, Jones, M.P., Dingkuhn, M., Johnson D.E and Fagade, S.O. (Ed) Proceedings of the workshop: Africa / Asia Joint Research on Interspecific Hybridization between the Africa and Asian Rice Species *(O. glaberrima and O. Sativa)*. WARDA, Mbé, Bouaké, Côte d'Ivoire. p 43-59.

Kato, T. (1990). Heritability for grain size of rice (*Oryza sativa* L.) from parent-offspring correlation and selection response. *Japanese Journal of Breeding* 40:313-320.

Kayode, A.S., Fawole, I., Guei, R.G., Ojo, D.K., Somado, E.A., Tia, D.D., Ogunbayo, S.A. and Sanchez. I. (2007). Geographical patterns of phenotypic diversity in *Oryza sativa* landraces of Côte d'Ivoire. *Euphytica* 160:389-400.

Khedikar, V.P., Bharose, A.A., Sharma, D., Khedikar, Y.P. and Khillare, A.S. (2004). Path coefficient analysis of yield components of scented rice. *Journal of Soils and Crops* 14:198-210.

Maji, A.J. (1998). Genetics of resistance to rice gall midge (*Orseolia* spp.). Paper presented during the WARDA/NARS group discussion and planning meeting on: Bioecology and Management of African Rice Gall Midge, 12-16 October at IITA, Ibadan, Nigeria.

Mohapatra, K.C. and Mohanty, H.K. (1985). Inheritance of some quantitative characters including heterosis in rice by combining ability analysis. In: *Rice Genetics: Proceeding of the International Rice Genetics Symposium.* IRRI (Ed), Manila, Philippines. p. 579-591.

Moldenhauer, K.A.K., Wells, B. and Helms, R. (1994). Rice growth stages. In: *Rice Production Handbook*: Helms, R.S. (Ed). Cooperative extension service print shop, Little Rock, Arkansas, USA. p 5-12.

Nwilene, F.E., Williams, C.T., Ukwungwu, M.N., Dakouo, D., Nacro, S., Hamadoun, A., Kamara, S.I., Okhidievbie, O., Abamu, F.J. and Adam, A. (2002). Reactions of differential rice genotypes to African rice gall midge in West Africa. *International Journal of Pest Management* 48:195-201.

Nwilene F.E., Harris, K.M., Okhidievbie, O. Onasanya, A. Sere Y. and Ingelbrecht I. (2006). Morphological diversity and genomic DNA fingerprinting of the African rice gall midge *Orseolia oryzivora* (Diptera: Cecidomyiidae) and of two other species of African *Orseolia*. *International Journal of Tropical Insect Science* 26**:**256-265.

N'cho, A.L. (1992). Les stades de croissance de la plante de riz. *Rapport de stage sur la production du riz pluvial.* ADRAO, Bouaké, Côte d'Ivoire. 14 p.

Panwar, R.P., Ashvani, S.R., Sharma, K., Arya, K.P.S and Panwar, A. (1997). Genetic variability and inter-relationship in rice (*Oryza sativa* L.). *Advances in Plant Science* 10:29-32.

Panwar, L.L. and Mashiat, A. (2007). Correlation and path analysis of yield and yield components in transplanted rice. *Oryza* 44:115-120.

Rabiei, B., Valizadeh, M., Ghareyazie, B. and Moghadam, M. (2004). Evaluation of selection indices for improving rice grain shape. *Field Crop Research* 89:359-367.

Rita, B., Motiramani, N.K. and Sarawgi, A.K. (2006). Association analysis and variability analysis in rice. *Mysore Journal of Agricultural Science* 40:375-380.

Rita, B., Sarawgi, A.K. and Verulkar, S.B. (2009). Study of heritability, genetic advance and variability for yield contributing characters in rice. *Bangladesh Journal of Agricultural Research* 34:175-179.

Sabu, K.K., Abdullah, M.Z., Lim, L.S. and Wickneswari, R. (2009). Analysis of heritability and genetic variability of agronomically important traits in *Oryza sativa* × *O. rufipogon* cross. *Agronomy Research* 7:97-102.

Sao, A. (2002). *Studies on Combining Ability and Heterosis in F1 Rice Hybrids Using Cytoplasmic Male Sterile Lines*. M.Sc. (Ag.) Thesis, IGAU, Raipur, India.

Schalbroek, J.J. (2001). Rice *Oryza sativa* L. In: *Crop Production in Tropical Africa*. RH. Raemaekers (Ed) DGIC, Ministry of Foreign Affairs, External trade and International Cooperation, Brussels, Belgium. p 59-78.

Shah, R., Sulemani, M.Z., Baloch, M.S. and Hassan, G. (1999). Performance of coarse rice genotypes in the plains of D.I. Khan, Pakistan. *Pakistan Journal of Biological Science* 2:507-515.

Singh, B.D. (2005). *Plant Breeding*. Kalyani Publishers Statistica *StatSoft Inc.*, New Delhi, India. 7 p.

Sarawgi, A.K., Rastogi, N.K. and Soni, D.K. (2000). Studies on some quality parameters of indigenous rice in Madhya Pradesh. *Annual Agricultural Research* 21:258-261.

Totiyama, K. and Futsuhara, Y. (1958). Estimates of heritabilities on individuals and lines in rice. *Japan Journal of Breeding* 7:208-210.

Umadevi, M., Veerabadhiran, P. and Manonmani, S. (2009). Genetic variability, heritability, genetic advance and correlation for morphological traits in rice genotypes. *Madras Agricultural Journal* 96:316-318.

Ukwungwu, M.N., Williams, C.T. and Okhidievbie, O. (1998). Screening of African rice, *Oryza glaberrima* Steud, for resistance to the African rice gall midge *Orseolia oryzivora* Harris and Gagné. *Insect Science and Its Application* 18:167-170.

Vanaja, T. and Babu, L.C. (2006). Variability in grain quality attributes of high yielding rice varieties (*Oryza sativa* L.) of diverse origin. *Journal of Tropical Agriculture* 44:61-63.

Villanueva, B. and Moro, J. (2001). Variance and efficiency of the combined estimator in incomplete block designs of use in forest genetics: a numerical study. *Canadian Journal of Forest Research* 31:71-77.

Wayne, S.C. and Dilday, R.H. (2003). *Rice: Origin, History, Technology and Production*. Wiley Series in Crop Science, John Wiley and Sons, Inc. Hoboken, New Jersey, USA.

Williams, C.T., Okhidievbie, O., Harris, K.M. and Ukwungwu, M.N. (1999). The host plant range, annual cycle and parasitoids of the African rice gall midge *Orseolia oryzivora* (Diptera: Cecidomyiidae) in central and southeast Nigeria. *Bulletin of Entomological Research* 89:589-597.

Wells, B.R. and Faw, W.F. (1978). Short-statured rice response to seeding and N rates. *Agronomy Journal* 70: 477-480.

Zafar, N., Aziz, S. and Masood, S. (2004). Phenotypic divergence for agro-morphological traits among landrace genotypes of rice (*Oryza sativa* L.) from Pakistan. *International Journal of Agricultural Biology* 6:335-339.

Chapter 6

Genetic variability in rice (*Oryza sp*.) genotypes with a range of resistance to the African Rice Gall Midge (AfRGM), *Orseolia oryzivora* Harris and Gagné

Abstract

Genetic variability studies on rice landraces conserve various objectives. Historically, its main use was to enable the creation of a core collection of rice for future breeding. Recently, with the use of new breeding approaches, genetic diversity has been studied for many purposes including biotype characterisation, phylogeny studies, population structures and genome-wide characterisation. Genetic variability among five gall midge *Orseolia oryzivora* Harris and Gagné resistant, tolerant and susceptible rice (*Oryza sativa* L.) genotypes, comprising of cultivars and landraces, was studied in terms of polymorphism, using 303 SSR primers covering the whole genome. Of the 178 polymorphic primers between cultivars and landraces, 60 were found to be highly polymorphic and informative. The number of alleles amplified by these primers ranged from one to five for a total of 1,041 alleles. The polymorphism rate was globally high ranging from 45.2 to 66.8%. The highest polymorphism rate was noted with primers from Chromosome 7 while the lowest was observed on Chromosome 9. The average of polymorphism information content (PIC) was 0.553. Factorial analysis based on the allelic diversity clearly demarcated the cultivars and landraces into *O. glaberrima*, *Oryza sativa* subsp. *japonica* and *Oryza sativa* subsp. *indica* groups. However, the cluster analysis distinguished the genotypes into gall midge resistant, susceptible, moderately resistant and tolerant groups. Two of the cultivars, BW348-1 and Cisadane, showed the least diversity, despite their diverse geographical origin. Likewise, two landraces (TOS14519 and TOG7106) and one cultivar (ITA306) showed more divergence, despite their common West African origin. This diversity among the genotypes tested is more evident with respect to natural evolution of genetic background or African rice gall midge (AfRGM) resistance rather than direct breeding efforts through breeder intervention.

Introduction

The AfRGM, *O. oryzivora* Harris and Gagné , is one of the more important insect pests of rice in the African lowland cropping system. Unlike its Asian counterpart where more than fifteen biotypes have been reported (Kumar *et al*., 1998, Sardesai *et al*., 2001, Bentur *et al*., 2003, Jain *et al*., 2004) with seven of them characterised (Vijaya *et al*., 2006), the number of biotypes of the African *Orseolia* is unknown to this day. However, the difference in reaction to multiple rice genotypes suggested the existence of more than one biotype. Furthermore, three insect species have been reported by Nwilene *et al*. (2006), who distinguished, based on genomic DNA fingerprinting using random amplified polymorphic DNA (RAPD) and sequence-characterized amplified region (SCAR), *Orseolia. nwanzei* Harris and Nwilene from the most closely related *O. oryzivora* Harris and Gagné and *Orseolia bonzii* Harris. Breeding and cultivation of resistant rice varieties are an important strategy to manage the pest. Field and paddy screenhouse screenings have led to the identification of resistant lines (Ukwungwu and Alam, 1991; Ukwungwu and Joshi, 1992; Maji, 1998; Williams *et al*., 1999; Nwilene *et al*., 2002). However, the resistant varieties found in these studies were mostly *Oryza glaberrima* landraces. Studies conducted by Williams *et al*. (1998), and more recently by Nwilene *et al.* (2002), identified TOS14519, a *O. sativa* subsp. *japonica* cultivar from the Gambia and the *O. glaberima* landrace TOG7106, as resistant lines to AfRGM. Like all the *O. glaberima* lines, they are less productive and could not be released to farmers. Varieties that have an acceptable yield because of tolerance to AfRGM are Cisadane from Indonesia, released in Nigeria as FARO 51, and BW348-1 from Sri Lanka. The latter is, however, susceptible to iron toxicity, one of the major constraints in African lowland cropping system (Singh, 1998). ITA306 is a breeding line released by the International Institute of Tropical Agriculture (IITA) but it has not been cultivated by farmers because it is highly susceptible to AfRGM (Williams *et al.*, 2002). Although their difference in reaction against AfRGM is known, the genome-wide diversity of these genotypes carrying either the same or different resistance genes has not been studied. Moreover, the genetic control involved in the tolerance of Cisadane and BW348-1 (Omoloye, 1998) to AfRGM is not known. The study conducted by Omoloye *et al.* (2002) only reported tolerance to infestations on the basis of a higher level of seedling survival, production of a significantly more fertile tillers, productive panicles and significantly higher grain yield per infested plant that ensured the acceptable yield noticed for the variety, despite the heavy

AfRGM infestation. Studying the variability of these lines reacting differently to AfRGM may reveal convergent evolution of same or similar resistance gene(s) in geographically distinct landraces or divergent evolution of genotypes carrying the same gene. Variability among populations has been historically addressed using agro-morphological traits. But with the major breakthrough in molecular breeding tools, molecular markers are being extensively used in the study of genetic diversity of crop plants to address questions of phylogeny, evolution and adaptation through domestication (Wang *et al.*, 1992; Joshi *et al*., 2000; Bautista *et al.,* 2001; Ren *et al.,* 2003). Simple Sequence Repeat (SSR) markers are the most commonly used in rice diversity because of their availability (over 25,000 SSR markers are available) (McCouch *et al.,* 2002) and the significant level of allelic diversity that they reveal (Ishii *et al.,* 2000). They are also particularly suitable for evaluating genetic diversity in different rice cultivars (Ren *et al.,* 2003) and different crops (Ellwood *et al.,* 2006). Germplasm analysis (Yong *et al.,* 2002), genetic relationships (Zhang *et al.,* 2003) and phylogenetic studies (Provan *et al.,* 2001) were also investigated by microsatellite SSR markers. The present study reports on a genetic diversity analysis based on genome wide polymorphism using SSR markers, with respect to five AfRGM resistant, tolerant and susceptible genotypes representing cultivars and landraces. Our results revealed the genomic diversity between cultivars and landraces and also distinguished them according to their reaction to AfRGM.

6-1 Material and method

6-1-1 Plant material

Five *O. sativa* and *O. glaberima* gall midge rice genotypes, including two landraces (TOS14519 and TOG7106), and three cultivars (BW348-1, Cisadane and ITA306), were used in this study. TOS14519 is an *O. sativa* subsp. *japonica* type from Gambia while TOG7106 is an *O. glaberima* accession from Casamance. BW348-1 is an *O. sativa* subsp. *indica* from Sri Lanka. Cisadane, from Indonesia, was released in Nigeria as FARO51. ITA306 is a high-yielding cultivar bred by the IITA. Details of these genotypes, along with their pedigree, their genotype type, their country of origin, their species type and their reaction against AfRGM attack are given in Table 6-1.

Table 6-1. Rice genotypes selected for AfRGM parental lines polymorphism study

Genotypes	Pedigree	Genotype type	Country of origin	Species type	Reaction* to AfRGM
BW348-1	Unknown	Cultivar	Sri Lanka	*indica*	Tolerant
Cisadane	Pelita I-1 / B2388	Cultivar	Indonesia	*indica*	Tolerant
ITA306	TOX494-3-6-9-6/TOX711//BG6812	Cultivar	Nigeria	*indica*	Susceptible
TOG7106	Landrace	Landrace	Casamance	*glaberrima*	Resistant
TOS14519	Landrace	Landrace	Gambia	*japonica*	MR

* MR= Moderately resistant

6-2 Methods

6-2-1 Choice of the SSR markers

A set of 303 SSR markers (Orjuela *et al*., 2009), covering the entire rice genome, was chosen based on the following criteria: (1) they express polymorphism according to two referenced populations including Azucena-IR64 and Azucena-TOG; (2) they derive from the Universal Core Genetic Map (UCGM) (Orjuela *et al*, 2009); (3) they belong to the 147 anchors evenly distributed along the rice genome; and (4) if one selected marker, belonging to an anchor and polymorphic for one population, was not polymorphic for the other population; another additional marker, polymorph for this second population, was chosen, making two markers equivalent for this anchor.

6-2-2 DNA extraction

Total genomic DNA was extracted using the cetyl trimethyl ammonium bromide (CTAB) method as described by Doyle and Dickson (1987), and later by Rusterucci *et al.* (2000), with minor modifications. Fresh leaves were collected 15 days after sowing from each individual. Three hundred mg of each sample was ground in liquid nitrogen, transferred into 25 ml tubes after being added to 7 ml of CTAB buffer, capped and incubated at 65^{o}C for one hour in a water bath and shaken intermittently during incubation. They were thereafter removed and cooled to room temperature. Seven ml of 24:1 chloroform: isoamyl alcohol was added to each sample and mixed by gently turning the tube upside down for 5 min and centrifuged for 15 min at 4000 revolutions per minute (rpm). The light green supernatant phase was carefully pipetted

into a pre-labelled 25 ml tube without disturbing the chloroform phase. Seven ml (i.e. about the volume of the light green phase) of 2-propanol (isopropanol) was added to the supernatant, mixed gently and incubated overnight. The samples were then spun for 10 min at 4000 rpm. The isopropanol was carefully decanted. The whitish pellet suspended on the surface of the tube was washed in 5 ml of 70% cold ethanol, transferred into a 1.5 ml Eppendorf tube and dried for 12 hours. The DNA pellet was then dissolved in Tris-ethanol (TE) for 24 hours.

6-2-3 DNA quality and quantity determination

The quality (level of non-degradation) of the extracted DNA was determined by spectrophotometer based on the ratio D260 to D280 and confirmed using an agarose gel electrophoresis method, using a 1% gel. The quantity of DNA was determined by spectrophotometer for reliability. A DNA ratio (OD260/OD280) of 1.8 to 2 is considered to be of good quality (Themo Scientific, 2010). DNA was diluted to 1/250 (2µl of DNA dissolved in 480 µl of water) and its concentration was assessed using its optical density, read at a wave length of 280 nm. Then 3 µl of diluted DNA mixed with 5µl of 10X loading dye (40% bromophenol blue plus 40% xylene cyanol FF and 50 % glycerol) was loaded into each well of the gel**.** The size of the amplified fragments was estimated using AlphaImager software with a 100-bp ladder (MBI Fermentas, Lithuania) as the size reference standard. DNA quality was verified by visualisation of the PCR product on an Alpha Imager gel documentation machine (Alpha Innotech, USA). A DNA sample with little or no smearing is considered as good quality. The values of the band size were used to perform a factorial analysis and to construct a dendrogram using DARwin Software version 5.0.155 (Perrier and Jacquemoud-Collet, 2006).

6-2-4 Genotyping

The susceptible female parent ITA306 and the four males of various levels of resistance were screened with the 303 SSR markers to reveal the polymorphism between them. A marker was considered polymorphic according to its polymorphism information content (PIC). A PIC above 0.7 was chosen as the polymorphism threshold. Polymorphism chain reaction (PCR) was performed as described by Temnykh *et al*. (2000) with slight modifications. Each reaction mixture consisted of a final volume of 15 µl containing 15 ng of template DNA, 1.3 X buffer (Tris-HCL 100 mM, 500 mM of KCl, Triton X-100 1%) already mixed with $MgCl_2$, 200µM d-

NTP, 0.1 µM of each primer, and 0.0250 unit of Taq polymerase. The mixture was placed in plate tubes covered with a thermal PCR seal to avoid evaporation. DNA amplification was performed on G-Storm thermocycler (Applied Bioscience, USA), programmed for an initial 3 min at 94°C, followed by 30 cycles of 30 seconds at 94°C, 45 seconds from 50 to 61°C according to the optimal annealing temperature for each SSR primer, and 1 min at 72°C, with a final extension of 5 min at 72°C to fill out any incompletely extended strands of DNA. The amplified products were separated by electrophoresis using a 4% non-denaturing polyacrylamide gel (Sambrook and Rusell 2001). When no or low amplification was observed, the PCR conditions were optimised by varying the annealing temperature of the primers (from 51 to 57°C), and the magnesium concentration (from 1X to 1.5X), in the PCR reaction.

6-2-5 Data analysis

The genetic distance between accessions was determined with DARwin software 5.0.155 (Perrier and Jacquemoud-Collet, 2006). The dissimilarity between individuals was calculated using the simple matching distance method based on the Sokal and Michener (1958) index as follows:

$$d_{ij} = 1 - \frac{1}{L}\sum_{i=1}^{L}\frac{m_l}{\pi}$$

With *dij*: dissimilarity between units *i* and *j*, *L*: number of loci, π: ploidy and m_l : number of matching alleles for locus *l*

The distances were used for factorial and cluster analyses. A dendrogram was constructed based on the dissimilarity allelic value according to the Unweighted Pair Group Method using the Arithmetic Mean (UPGMA) hierarchical clustering method. The PowerMarker 3.25 software package (Liu and Muse, 2005) was used to calculate the allele number per locus, major allele frequency, gene diversity and PIC. The PIC was calculated as $1-\sum p_i^2-\sum_i\sum_{j>i}2p_i^2p_j^2$; where p_i and p_j are the frequencies of the i and j alleles, respectively (Botstein *et al.*, 1980). The difference between the five AfRGM resistant, tolerant and susceptible lines was tested locus by locus, and overall significance was stated.

6-3 Results

6-3-1 Major allele frequency

Major allele frequency (MAF) ranged from 0.2 to 1 with a mean of 0.48. Five allelic classes were observed. Allelic Class 1 had a MAF of 0.2 and bore 5 alleles revealed by 19% (60 markers) of the total number of markers used. Allelic Class 2 had twice (0.4) the MAF of allelic Class 1 shared by four alleles revealed by 131 of the total markers. Class 3 showed a MAF of 0.5 - 0.6 and contained 3 alleles representing 20.46% (62 markers) of the total markers used to reveal these 3 alleles. Allelic Classes 4 and 5 displayed a MAF of 0.8 and 1, revealing only 2 and 1 alleles, respectively. The numbers of markers which revealed these alleles associated to these classes were both 25. The MAF was low in Class 1, medium in Classes 2 and 3 and high in Classes 4 and 5. Along with chromosomes, MAF ranged from 0.37 on Chromosome 7 to 0.59 on Chromosome 10.

6-3-2 Allelic diversity

A total of 1,047 alleles were amplified with all the SSR markers. Number of alleles per marker ranged from 1 to 5, with an average of 3.46. Within these two extremes, markers also revealed 4, 3 or 2 alleles. Across chromosomes, alleles detected ranged from 54 on Chromosome 9 to 144 on Chromosome 1, with a mean number of alleles per marker of 2.85 on chromosome 9 and 4.05 on Chromosome 7, respectively, as the minimum and maximum values.

Table 6-2: MAF, number of alleles per marker, and number of revealing marker distributions according to five allelic classes

	Allelic Classes				
	1	2	3	4	5
Major allele frequency	0.2	0.4	0.5-0.6	0.8	1
Number of alleles/marker	5	4	3	2	1
Number of revealing marker	60 (*19.8)	131 (*43.23)	62 (*20.46)	25 (*8.25)	25 (*8.25)

number in brackets is the corresponding percent of the number of revealing marker

6-3-3 Gene diversity

Gene diversity was as high as 0.8 and as low as zero, with a mean value of 0.6. Of the 303 primers used, 278 revealed diversity among the tested genotypes; the 25 remaining primers produced nil values for gene diversity. Between these extremes, intermediate gene diversity values were also observed, ranging from 0.32 to 0.72. In between this interval, other values of 0.48, 0.56, 0.63 and 0.64 were noted. Gene diversity distribution across chromosomes ranged from 0.5 on Chromosome 10 to 0.67 on Chromosome 7.

6-3-4 Polymorphism information content

Similarly to the gene diversity, 278 of the 303 SSR markers tested showed polymorphism among the test genotypes. Of these, 60 were the most informative in being highly polymorphic in reference to a threshold value of PIC higher than 0.7. The remaining 134 and 84 markers had moderate and low level of polymorphism with respect to threshold values of 0.5-0.7, and 0.3-0.5, respectively. Among the polymorphic markers, polymorphism varied from 26.9 to 76.8%. The mean of PIC was 0.553. The most polymorphic chromosome was Chromosome 7 (66.8%), while the least polymorphic was Chromosome 9 (45.2%).

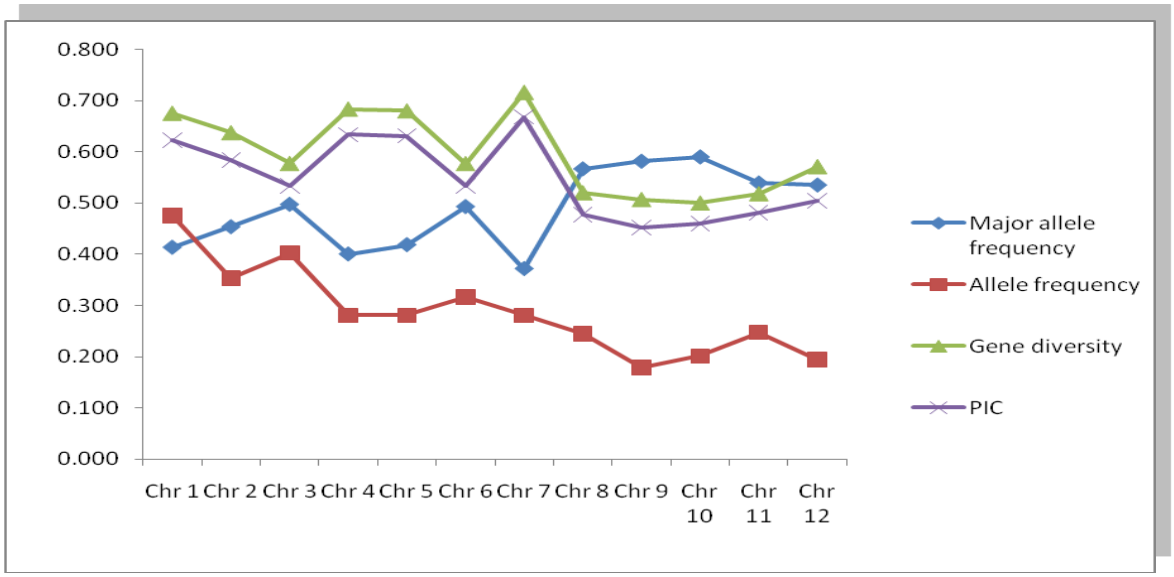


Figure 6-1: MAF, allele frequency, gene diversity distribution and PIC along the twelve chromosomes of rice

6-3-5 Correlation between genetic parameters

The summary table (Table 6-3) of the coefficient of determination (R^2) shows that the genetic parameters were highly correlated. The extent of correlation ranged from 0.833 between gene diversity and allele number to 0.99 between PIC and gene diversity.

Table 6-3: Coefficient of determination (R^2) between genetic parameters

Parameters	MAF	Allele No	Gene Diversity	PIC
MAF	1	0.920	0.931	0.945
Allele No	0.920	1	0.883	0.935
Gene Diversity	0.931	0.883	1	0.990
PIC	0.945	0.935	0.990	1

MAF= Major allele frequency, PIC = Polymorphism information content

Table 6-4 confirms not only the high correlation between genetic parameters but also the nature of the correlation. The parameters were correlated at 94% to 99.5%. Indeed, MAF was strongly, negatively correlated to allele number, gene diversity and PIC, whereas strong positive correlation was observed between allele number, gene diversity and PIC. The least correlation was observed between allele number and gene diversity (0.94) while the highest correlation was observed between PIC and gene diversity. A positive correlation of 0.967 between PIC and allele number was also measured.

Table 6-4: Correlation between genetic parameters according to the correlation matrix of Pearson

Parameters	MAF	Allele No	Gene Diversity	PIC
MAF	1	-0.959	-0.965	-0.972
Allele No	-0.959	1	0.940	0.967
Gene Diversity	-0.965	0.940	1	0.995
PIC	-0.972	0.967	0.995	1

MAF= Major allele frequency, PIC = Polymorphism information content, p-value <0.0001

6-3-6 Genetic distance between individuals

Genetic variability among the tested genotypes was analysed using the polymorphism data from 60 highly polymorphic and informative SSR markers. Factorial and cluster analyses

showed differential trend in genotype grouping. The factorial analysis performed on Axes 3 and 1 which explained 55.88% of the total variability, demarcated the test lines into two distinct groups, representing *O. sativa* and *O. glaberrima* accessions. The *O. glaberrima* group was represented only by TOG7106 while the *O. sativa* group included BW348-1, Cisadane, ITA306 and TOS14519. Among the *O. sativa* group, a clear demarcation was made between *O. sativa* subsp. *indica* types and the *O. sativa* subsp. *japonica* type (Figure 6-2). The *O. sativa* subsp. *japonica* TOS14519 was demarcated far away while the *O. sativa* subsp. *indica* varieties ITA306, BW348-1 and Cisadane, were closely associated. Axis 3 separated *O. glaberrima* from *O. sativa* while Axis 1 demarcated the *O. sativa* subsp. *japonica* TOS14519 from the *O. sativa* subsp. *indica* BW348-1, Cisadane and ITA306.

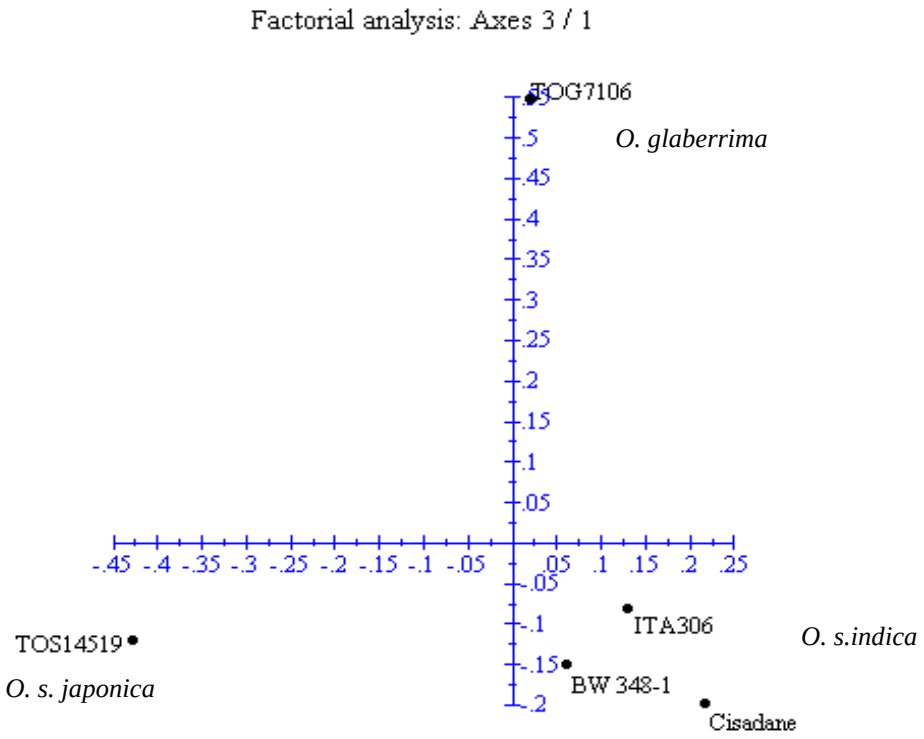


Figure 6-2: Factorial analysis on five rice genotypes resistant, moderately resistant, susceptible and tolerant to AfRGM, based on 303 SSR data projected on Axes 3/1

The dendrogram constructed based on the dissimilarity allelic value among the genotypes according to the UPGMA method at the threshold equality of 1% (Figure 6-3) grouped the tested genotypes into four distinct groups. The first group had only one member, the landrace

O. glaberrima TOG7106, which is resistant to AfRGM. The second group included only *the O. sativa* subsp. *indica* cultivar ITA306, susceptible to AfRGM. The third group comprised the *japonica* landrace TOS14519, which is moderately resistant to AfRGM, while the last group comprised two *O. sativa* subsp. *indica* cultivars (BW348-1 and Cisadane), which are tolerant to AfRGM. The furthest dissimilarity was observed between the resistant landrace TOG7106 and the tolerant cultivars BW348-1 and Cisadane, followed by the one between TOG7106 and the moderately resistant landrace TOS14519. ITA306, the susceptible cultivar, was the less genetically dissimilar to TOG7106. ITA306 showed less dissimilarity to TOS14519 than TOS14519 showed to BW348-1 and Cisadane. BW348-1 and Cisadane showed the highest similarity (Table 6.5).

Figure 6-3: Dendrogram resulting from UPGMA cluster analysis for five rice genotypes resistant, moderately resistant, susceptible and tolerant to AfRGM based on polymorphism data derived from 60 highly polymorphic SSR markers

T=Tolerant, MR=Moderately resistant, S= Susceptible, R= Resistant

Table 6.5: Dissimilarity index calculated for five rice genotype resistant, moderately resistant, susceptible and tolerant to AfRGM using Darwin software 5.0.155 based on simple matching

Parental lines	ITA306	TOG7106	TOS14519	BW348-1
TOG7106	0.801325			
TOS14519	0.677741	0.817881		
BW348-1	0.731788	0.818482	0.718543	
Cisadane	0.744186	0.82392	0.72	0.663333

6-4 Discussion

The MAF was inversely proportional to the number of alleles revealed by markers and the PIC. The higher the number of alleles, the lesser their individual frequency and the polymorphism were revealed. Likewise, the lesser the number of alleles, the bigger are their individual frequency and PIC value. Our results suggested that alleles revealed by each primer were equally represented in our population. This was confirmed by the mean MAF value of nearly fifty percent (0.48) and the nil value of heterozygosity (data not shown). This phenomenon could have been due to a consideration of a single allele during the polymorphism study rather than an equal frequency distribution of alleles. Thus, only alleles present in each of the genotypes which clearly demarcated and showed a visible band were considered during the genotyping.

Direct comparison of the extent of genetic variability in our study with those observed in previous studies is risky because it may be biased by the number and the type of accessions used to evaluate the genetic diversity. Only five accessions, including three cultivars and two landraces, were used. However, the following can be claimed:

The maximum number of alleles per primer (5) was lower than those reported by Barry *et al.* (2007), Zhang *et al.* (2007) and Semon *et al.* (2004) who found 15, 22 and 27 alleles, respectively, among the *O. sativa* and *O. glaberrima* landraces in the diversity study conducted on a country, province and regional scale, respectively. In another study on common wild rice of six Chinese provinces, more alleles were detected, on an average of 21.58 alleles per locus, with a maximum of 53 alleles, compared with ours (3.24 alleles) and previous reports by Gao *et al*. (2002) (3.1 alleles), Zhou *et al*. (2003) (10.6 alleles), and Song *et al*. (2003) (5 alleles). Our result is also slightly different from the one of Himabindu *et al.* (2009) who found a maximum

number of seven alleles for a similar study conducted on lines resistant and susceptible to the Asian gall midge. However, the mean values of alleles per marker were similar in the two studies (3.24 vs. 3.46 found in the present study). We also found 1,041 as the overall number of alleles, which is surprisingly, eight times more than the number of 129 found by Barry *et al.* (2007), but less than the 1,161 alleles documented by Himabindu *et al.* (2009). The differences with the Himabindu study can be explained by the limited number and type of accessions (mostly cultivars) used in our study. This also is probably owing to more representative materials being used in their study. Indeed, only five accessions were investigated here, in comparison to twelve used in the study conducted by Himabindu. The low level of 129 alleles observed was reported by Barry *et al.* as being due to the number and the choice of the SSR loci.

The PIC which refers to the value of a marker for detecting polymorphism within a population, depending on the number of detectable alleles; and the distribution of their frequency, was higher (0.553) than the PIC found by Semon *et al.* (0.34). However, it was lower than those reported by Barry *et al.* (0.8) and Himabindu *et al* (0.589).

Our gene diversity value (0.6) was also almost twice the value (0.27) reported by Semon *et al*. This is due to the fact that, accessions used in our study were mostly cultivars, representing the final product of plant breeders who have therefore broadened the cultivars' genetic background. Moreover, the composition of our accession set comprised of 1 *O. glaberima,* 1 *O. sativa* subsp. *japonica* and 3 *O. sativa* subsp. *indica* was more genetically diverse than the *glaberima* landraces used in Semon's study; landraces being least intervened. As reported by Glaszmann (1987), the *O. sativa* subsp. *indica* sub-species have very broad genetic diversity.

The lowest values of allelic diversity, gene diversity and PIC were consistently observed on Chromosome 9, which was reported by Biradar *et al*. (2004) as bearing the Asian gall midge resistance gene Gml, flanked by the SSR markers RM316, RM444 and RM219 at the genetic distances of 8.0, 4.9 and 5.9 cM, respectively**.**

Positive and negative strong correlations were measured between genetic parameters. A negative correlation was detected by the method of Pearson between allele number and the MAF. However, a consideration of this correlation on the chromosome scale, showed that the correlation was variable and of two types. The parameters were either effectively negatively

correlated on Chromosomes 2, 8, 9 and 11, or they were conversely positively correlated on the other chromosomes. In this latter case, the hypothesis of the existence of only one type of allele, considered as a major allele, can be postulated; and consequently, genes bearing these alleles can be considered to be major genes. Likewise the decrease in the number of alleles associated with the increase of gene diversity and PIC can be explained by genes bearing few alleles.

The factorial analysis distinguished three clusters of genetic variability among cultivars and landraces into *O. glaberima*, *O. sativa* subsp. *japonica* and *O. sativa* subsp. *indica* types, meaning that the markers used were effective and informative. Although, the *O. sativa indica* sub-species group was constituted only of cultivars, which are the final product of plant breeders and have gone through intensive selection pressure, their *O. sativa* subsp. *indica* background was not altered. Indeed, ITA306 is of BC_2F_6 generation whereas Cisadane is of BC_3 generation. The fact that selection narrows the genetic background could partly explain the least range of diversity among them.

However, the *O. sativa indica* sub-species were clearly separated in the cluster analysis, suggesting a divergence in their reaction to AfRGM. BW348-1 and Cisadane, which are both tolerant to AfRGM, showed no divergence, whereas a clear divergence was made between them and ITA306. This grouping trend, respective to AfRGM reaction was also confirmed by the nearest convergence between TOS14519, a moderately AfRGM resistant landrace and the two AfRGM tolerant cultivars. The cluster analysis separated TOG7106, the resistant landrace, from the others, which showed a differential reaction to AfRGM. The second clustering distinguished ITA306, the susceptible cultivar, from those with a low to moderate level of resistance, and the third clustering demarcated TOS14519, the most resistant landrace from BW348-1 and Cisadane, the tolerant cultivars. Surprisingly, our result depicted that resistance and susceptibility, although associated with extreme behavioural pattern against AfRGM, were the least divergent in terms of genetic diversity. However, resistance, moderate resistance and tolerance which are closer in terms of reactions against the AfRGM showed the furthest divergence in gene diversity. These suggested that resistance to AfRGM is rather of qualitative control than of quantitative. Breeding for resistance to AfRGM using introgression into promising lines should therefore be oriented at targeting a specific region rather than wide

genome screening. Marker assisted selection (MAS) could be a focused means to achieve this goal.

The geographical origin of the genotypes was not critical in the groupings, contrary to Zhou *et al*. (2003), who found that the genetic distances among twelve natural populations were related to their spatial distances; and Cai *et al*. (2004) who also reported the correlation between genetic distance and spatial isolation. Indeed, BW348-1 and Cisadane, respectively from Indonesia and Sri Lanka, were grouped together despite their wide geographical origins, while TOG7106, ITA306 and TOS14519, respectively, from Senegal, Nigeria and the Gambia from West Africa, were clustered. Our results are in line with the conclusion of Wang *et al.* (2008) on geographical genetic diversity and divergence of common wild rice (*Oryza rufipogon* Griff.) in China who claimed that, although isolation is a necessary factor for divergence in accessions such as cultivar and landraces, natural selection or other factors that substantially cause the loss and frequent change of alleles, would be the direct dynamics of divergence.

Our study using SSR markers showed four levels of clustering (resistant, susceptible, moderately resistant and tolerant) between cultivars and landraces, respective to their reaction against AfRGM and three distinct groups, respective to species and sub-species. Conversely, the grouping showed no pattern with their geographical origins. This finding is different from the one of Himabindu *et al.* (2009) who reported grouping only according to cultivars, with a low diversity noticed among them.

References

Barry, M., Pham, J., Noyer, J.L., Billot, C., Courtois, B. and Ahmadi, N. (2007). Genetic diversity of the two cultivated rice species (*O. sativa* and *O. glaberrima*) in Maritime Guinea. Evidence for interspecific recombination. *Euphytica* 154:127-137.

Bautista, N.S., Solis, R., Kamijima, O and Ishii, T. (2001). RAPD, RFLP and SSLP analysis of phylogenetic relationships between cultivated and wild species of rice. *Genes and Genetics System* 76:71-79.

Bentur, J.S., Pasalu, I.C., Sarma, N.P., Prasada Rao, U., Mishra, B. (2003). *Gall Midge Resistance in Rice.* DRR Research Paper series 01/2003, Directorate of Rice Research, Hyderabad, India. 20 p.

Biradar, S.K., Sundaram, R.M., Thirumurugan, T., Bentur, J.S., Amudhan, S., Shenoy, V.V., Mishra, B., Bennet, J., Sarma, N.P. (2004). Identification of flanking SSR markers for a major rice gall midge resistance gene Gm1 and their validation. *Theoretical and Applied Genetics* 10:1468-1472.

Botstein, D., White, R.L., Skolnick, M. and Davis. R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetic* 32:314-331.

Cai, H.W, Wang X.K. and Morishima, H. (2004). Comparison of population genetic structures of common wild rice (*Oryza rufipogon* Griff.), as revealed by analyses of quantitative traits, allozymes and RFLPs. *Heredity* 92: 409-417.

Doyle, J.J. and Dickson. E.E. (1987). DNA extraction using CTBA method. *Taxon* 36:715-722.

Ellwood, S.R., D'Souza, N.K., Kampheris, L.G., Burgess, T.I., Nair, R.M. and Oliver, R.P. (2006). SSR analysis of the *Medicago truncatula* SARDI core collection reveals substantial diversity and unusual genotype dispersal throughout the Mediterranean basin. *Theoretical and Applied Genetics* 112: 977-983.

Gao, L.Z., Schaal, B.A. and Zhang, C.H. (2002). Assessment of population genetic structure in common wild rice (*Oryza rufipogon* Griff.) using microsatellite and allozyme markers. *Theoretical and Applied Genetics* 106: 173-180.

Glaszmann, J. (1987). Isozymes and classification of Asian rice varieties. *Theoretical and Applied Genetics* 74:21-30.

Himabindu, K., Sama, V.S.A.K, Naik , B.S., Pranitha J. Sinha, P., Sadasiva, N. and Bentur J.S. (2009). Genetic Diversity in Rice (*Oryza sativa* L.) Genotypes resistant and susceptible to the Asian rice gall midge *Orseolia oryzae* (Wood-Mason). *Indian Journal of Plant Genetic Resources* 22:108-112.

Ishii, T. and McCouch, S.R. (2000). Microsatellites and microsynteny in the chloroplast genomes of *Oryza* and eight other Gramineae species. *Theoretical and Applied Genetics* 100: 1257-1266.

Jain, A., Ariyadasa, R., Kumar, A., Srivastava, M.N., Mohan, M. and Nair, S. (2004). Tagging and mapping of a rice gall midge resistance gene, Gm8 and development of SCARs for use in marker-aided selection and gene pyramiding. *Theoretical and Applied Genetics* 109: 1377-1384.

Joshi, S.P., Gupta, V.S. Aggarwal, R.K. Ranjekar, P.K. and Brar, D.S. (2000). Genetic diversity and phylogenetic relationship as revealed by inter simple sequence repeat (ISSR) polymorphism in the genus *Oryza*. *Theoretical and Applied Genetics* 100: 1311-1320.

Kumar, A., Srivastava, M.N., Shukla, B.C. (1998). Inheritance and allelic relationship of gall midge biotype-1 resistant gene(s) in some new donors. *Oryza* 35:70-73.

Liu, K., and Muse. S.V. (2005). PowerMarker: Integrated analysis environment for genetic marker data. *Bioinformatics* 21:2128-2129.

Maji, A. J. (1998). Genetics of resistance to rice gall midge (*Orseolia* spp.). Paper presented during the WARDA/NARS group discussion and planning meeting on Bioecology and management of African rice gall midge, 12-16 October at IITA, Ibadan, Nigeria.

McCouch, S.R., Teytelman, L., Xu, Y., Lobos, K.B., Clare, K., Walton, M., Fu, B., Maghirang, R., Li, Z., Xing, Y., Zhang, Q., Kano, I., Yano, M., Fjellstrom, R., De Clerk, G., Schneider, D., Cartinhour, S., Ware, D. and Stein, L. (2002). Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). *DNA Research* 9:199-207.

Nwilene, F.E., Williams, C. T., Ukwungwu, M. N., Dakouo, D., Nacro, S., Hamadoun, A., Kamara, S.I., Okhidievbie, O., Abamu, F.J. and Adam, A. (2002). Reactions of

differential rice genotypes to African rice gall midge in West Africa. *International Journal of Pest Management* 48:195-201.

Nwilene, F.E., Harris, K.M., Okhidievbie, O., Onasanya, A., Sere, Y. and Ingelbrecht, I. (2006). Morphological diversity and genomic DNA fingerprinting of the African rice gall midge *Orseolia oryzivora* (Diptera: Cecidomyiidae) and of two other species of African Orseolia. *International Journal of Tropical Insect Science* 26:256-265.

Omoloye, A.A. (1998). *Mechanisms of Resistance in Rice Cultivars to the African Rice Gall Midge, Orseolia oryzivora Harris and Gagné in Ibadan, Nigeria*. Ph.D. thesis, University of Ibadan, Nigeria.

Omoloye, A.A., Odebiyi, J.A., Williams, C.T. and Singh, B.N. (2002). Tolerance indicators and responses of rice cultivars to infestation by the African rice gall midge, *Orseolia oryzivora. Journal of Agricultural Science* 139:335-340.

Orjuela, J., Garavito, A., Bouniol, M., Arbelaez, J.D., Moreno, L., Kimball, J., Wilson, G., Rami, J.F., Tohme, J., McCouch, S.R. and Lorieux, M. (2009). A universal core genetic map for rice. *Theoretical and Applied Genetics* 120: 563-572.

Perrier, X., and Jacquemoud-Collet, J.P. (2006). DARwin software [Online]. Available by http://darwin.cirad.fr/darwin.

Provan, J, Powell, W. and Hollingsworth, P.M. (2001). Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trend in Ecology Evolution* 16:142-147.

Ren, F., Lu, B.R. Li, S. Huang, J. and Zhu, Y. (2003). A comparative study of genetic relationships among the AA genome *Oryza* species using RAPD and SSR markers. *Theoretical and Applied Genetics* 108:113-120.

Risterucci, A.M., Grivet, L. N'Goran, J.A.K., Pieretti, I., Flament, M.H and Lanaud. C. (2000). A high-density linkage map of *Theobroma cacao* L. *Theoretical and Applied Genetics* 101:948-955.

Sambrook, J. and Rusell, D.W. (2001). *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, New York, USA.

Sardesai, N., Kumar, A., Rajyashri, K.R., Nair, S. and Mohan, M. (2002). Identification and mapping of an AFLP marker linked to Gm7, a gall midge resistance gene and its conversion to a SCAR marker for its utility in marker-aided selection in rice. *Theoretical and Applied Genetics* 105:691-698.

Semon, M., Nielsen, R., Jones, P. and McCouch, S. (2004). The population structure of cultivated *Oryza glaberrima* (Steud): Evidence for elevated levels of LD caused by admixture with *O. sativa* and ecological adaptation. *Genetics* 169:639-1647.

Singh, B.N. (1998) Breeding for rice gall midge (*Orseolia oryzivora* Harris and Gagné) resistance at WARDA *Paper presented during the WARDA/NARS Group Discussion and Planning Meeting on 'Bioecology and Management of African Rice Gall Midge', 12-16 October at IITA, Ibadan, Nigeria*.

Sokal, R.R., and Michener, C.D. (1958). A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin* 38:1409-1438.

Song, Z.P., Xu, X., Wang, B., Chen, J.K. and Lu, B.-R. (2003). Genetic diversity in the northernmost *Oryza rufipogon* populations estimated by SSR markers. *Theoretical and Applied Genetics* 107: 1492-1499

Temnykh, S.W., Park, D., Ayres, N., Cartingour, S., Hanck, N., Lipovich, L., Cho, Y. S. and Tshil, T. (2000). Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics* 100: 697-712.

Thermo Scientific (2010). *T042-Technical Bulletin NanoDrop Spectrophotometers, 260/280 and 260/230 Ratios*. Available at: http://www.nanodrop.com/Library/T042-NanoDrop-Spectrophotometers-Nucleic-Acid-Purity-Ratios.pdf Retrieved 19 March 2012).

Ukwungwu, M.N. and Joshi R.C. (1992). Distribution of the African rice gall midge, *Orseolia oryzivora* Harris and Gagné and its parasitoids in Nigeria. *Tropical Pest Management* 38: 241-244.

Ukwungwu, M.N. and Alam M.S. (1991). Field resistance of rice cultivars to the gall midge *Orseolia oryzivora* in Nigeria. *International Rice Commission Newsletter* 40:42-43.

Vijaya, L.P., Amudhan, S. Himabindu, K. Cheralu, C. and Bentur, J.S (2006). A new biotype of the Asian rice gall midge *Orseolia oryzae* (Diptera: Cecidomyiidae) characterized from the Warangal population in Andhra Pradesh, India. *International Journal of Insect Science* 26:207-211.

Wang, Z., Second, G. and Tanksly, S.D. (1992) Polymorphism and phylogenetic relationships among species in the genus *Oryza* as determined by analysis of nuclear RFLPs. *Theoretical and Applied Genetics* 83:565-581.

Wang, M., Zhang, H., Zhang, D., Pan, D., Yuan Li, D., Fan, Z., Qi, Y., Sun, J., Yang, Q. and Chen, Li. (2008). Geographical genetic diversity and divergence of common wild rice (*O. rufipogon* Griff.) in China. *Chinese Science Bulletin* 53:3559-3566.

Williams, C.T., Okhidievbie O., Unkwungwu, M.N., Dakouo, D., Nacro, S., Hamadou, A., Kamara, S.I. (1999). Multilocational screening of *Oryza sativa* and *Oryza glaberrima* for resistance to African Rice gall midge *Orseolia oryzivora* in West Africa. *International. Rice Research Notes* 24:26-27.

Williams, C.T., Harris, K.M., Ukwungwu, M.N., Nacro, S., Dakouo, D., Nwilene, F.E., Singh, B.N. and Okhidievbie, O. (2002). *African Rice Gall Midge Research Guide*. West Africa Rice Development Association and CAB International, Wallingford, United Kingdom. 28pp.

Yong, Z., Yang, Kai Y., Cheema, A.A. and Jin, W.Y. (2002). Evaluation of germplasm using SSR markers of functional genes in rice. *Agricultural Science in China* 1: 477-480.

Zhang, D., Zhang, H., Wei, X., Qi, Y., Wang, M., Sun, J., Ding, L., Tang, S., Qiu, Z. and Cao, Y. (2007). Genetic structure and diversity of *Oryza sativa* L. in Guizhou, China. *Chinese Science Bulletin* 52:343-351.

Zhang, X.Y., Li, C.W., Wang, L.F., Wang, H.M., You, G.X. and Dong, Y.S. (2003). An estimation of the minimum number of SSR alleles needed to reveal genetic relationships in wheat varieties. I. Information from large-scale planted varieties and cornerstone breeding parents in Chinese wheat improvement and production. *Theoretical and Applied Genetics* 106:112-117.

Zhou, H.F., Xie, Z.W. and Ge, S. (2003) Microsatellite analysis of genetic diversity and population genetic structure of a wild rice (*Oryza rufipogon* Griff.) in China. *Theoretical and Applied Genetics* 107:332-339.

Chapter 7

Constitution of a core collection for a reduced mapping population size, bearing the total variability of the whole population, using agro-morphological traits in view of genome wide mapping

Abstract

Bulked segregant analysis and selective genotyping efficiently help in dealing with large population sizes. However, these mapping methods only permit the mapping of the major genes controlling the trait of interest, obliging breeders to use the tedious and time consuming procedure of genome wide mapping to monitor the effects of the minor quantitative trait loci (QTLs). A variant of these methods was studied, aiming to select individuals for a reduced mapping population based on a core sample that represents as much as possible the entire variability of the population. A diversity study amongst F_2 and F_3 generations involving 15 quantitative and 26 qualitative agro-morphological characters was carried out to discriminate between the individual previously used. First, seven morphological descriptors, including the leaf blade distribution of anthocyanin colouration, the leaf blade pubescence on blade surface, the leaf margin pubescence, the auricle colour, the collar colour, the ligule shape and the degree of panicle shattering, were discarded because they were not discriminative. The diversity index (H) was calculated for each remaining morphological character. The discriminant descriptors were selected based on a diversity index threshold value above 0.4. Therefore, ligule length, ligule colour and panicle main axis attitude of H values of 0.17, 0.34 and 0.34 respectively were also discarded. The variability was confirmed using a neighbour joining tree cluster analysis which stratified the 649 individuals into four distinct clusters that did not distinguish the descriptors. The size of the core collection was determined using MSTRAT version 4.1 package in redundancy mode. The 64 top ranking individuals, based on their frequency of appearance after a construction run of 100 times (number of replicates), and inferred with an iteration number of 500, were selected for the core collection. The comparison of the core sample to the entire population indicated a similarity between them for clustering pattern, minimum and maximum quantitative values, phenotypic variation, genotypic segregation and diversity index while mean values, and coefficient of variation distinguished them. The core

sample, which represents a 10% sample fraction of the whole population, can therefore efficiently represent it as a mapping population.

Introduction

As gene mapping becomes routine in many laboratories, requirements for its efficient operation constitute some of the major limitations and constraints to be overcome (Semagn *et al.,* 2006a, Semagn *et al*., 2006b and Semagn *et al*., 2006c). Historically, linkage mapping and gene mapping populations have consisted of large, randomly selected samples of progeny from a given pedigree, which is tedious and time consuming. Several alternative approaches were identified. For linkage mapping, the alternative approach is to use "selective mapping" as reported by Vision *et al*. (2000). They demonstrated that, to construct a map with high genome-wide marker density, it is neither necessary nor desirable to genotype all markers in every individual of a large mapping population. Instead they proposed the use of a reduced sample of individuals bearing complementary recombinational or radiation-induced breakpoints that can be selected for genotyping subsequent markers from a large, but sparsely genotyped, mapping population.

For gene mapping, the strategy relies on using: (1) Bulk segregant analysis (BSA) which consists of genotyping individuals selected from the low and high tails of phenotypic distribution of the trait under study (Lebowitz *et al*., 1987; Lander and Botstein, 1989); (2) Selective genotyping (SG), which involves individuals of contrasting phenotypes randomly chosen from the corresponding contrasting phenotype population (Darvasi and Soller 1994, Sun *et al.*, 2010); and (3) The technique of sequential bulk typing developed by Pérez-Enciso (1998), which is based on the pooling of DNA samples from consecutive pairs of individuals ranked phenotypically, i.e., pools are formed with individuals ranked (1^{st}, 2^{nd}), (3^{rd}, 4^{th}), (N^{-1st}, N^{th}). For gene mapping, apart from the sequential bulk typing developed by Pérez-Enciso, BSA and SG only allow for mapping of major QTLs, therefore minimising the minor QTLs.

Tagging minor QTLs involves implementing genome-wide mapping, for which the accuracy is commonly constrained to a large population size or a high density marker survey. The sequential bulk analysis is the only approach to date for the mapping of minor QTLs, without dealing with the entire population, by pooling consecutive individuals into a pool of at least two individuals. However, pooling individuals also has a major constraint in that DNA

should be accurately quantified with respect of the ratio of equimolarity of the pooled individuals. The higher the number of pooled individuals, the less even is the pooling process. Moreover, depending on the availability, the concentration and the quality of the DNA, pooling difficulties vary greatly. Pooling concentrated with dilute DNA, or high quality with low quality DNA, leads to unbalanced amplification of each individual of the pool. Therefore, one of the most efficient options to avoid pooling constraints is to individually genotype some representatives of a large mapping population. The chosen individuals need to represent the variability of the whole population from which they were selected, as far as possible. The accurate selection of the individual is therefore the key component of the success of the mapping.

Diversity studies have been dedicated to study variability among individuals. A diversity index is a statistic which intends to measure the diversity of a set consisting of various types of objects. Diversity indexes are used in many fields of study, to assess the diversity of any population in which each member belongs to a group, type or species. While categorising individuals, the diversity index quantifies the level of variability within these groups. A core collection is defined as a representative sample of the entire collection, with minimum repetitiveness and maximum genetic diversity of a crop species and its relatives (Zhang *et al*., 2010). Although core collections have been widely used in association mapping during the last five years, they have not been used as a mapping population, even though the theoretical feasibility of this approach has been documented by others (Richards *et al*., 2009). This approach was investigated in this study as a novel approach to genetic mapping of a large population size. It is quicker and more consistent because it uses a much-reduced number of samples. However, the diversity of the core sample approach that of the larger populations from which it was derived. This should allow for high-quality, high-density, genome-wide mapping. The objectives of this present study was to: (1) Create an accurate mapping population based on the construction of a core collection, representing the variability of the entire population; (2) Compare the genetic variability and validate the core collection, as a representation of the global mapping population.

7-1 Material and method

7-1-1 Development of F_2 and F_3 population after validation of F_1 generation

A plant population was constituted from F_2 plants and F_3 families derived from a cross between ITA306, a high-yielding *Oryza sativa* subsp. *indica* cultivar that is susceptible to African Rice Gall Midge (AfRGM), used as the female parent; and TOS14519, an *O. sativa* subsp. *indica* cultivar resistant to AfRGM used as the male parent. The F_1s obtained were confirmed by vigour assessment and basal sheath anthocyanin colouration, based on the vigour of true F_1 plants over their parents, and the presence of a discriminant character of the male parent in the progenies. Validated F_1 plants were advanced to F_2 in the greenhouse. Three-week old seedlings were transplanted into 12-litre plastic buckets, which were 3/4-filled with paddy soil, at the rate of one seedling per bucket. Daily watering was done until heading. 649 F_2 plants were advanced to F_3 generation. At maturity, F_2 individual plants were harvested separately and numbered for evaluation as F_3 families.

7-1-2 Agro-morphological characterisation

7-1-2-1 F_2 characterisation

649 F_2 plants were grown individually in 12-litre plastic buckets at the AfricaRice temporary headquarters in Benin, Cotonou. Fertilizer at the level of 1 g each of N, P_2O_5 and K_2O were added to each pot at transplanting, using NPK 15-15-15 as the source. The pots were hand-weeded whenever necessary. The experimental design was an augmented design (Federer 1956) with 6 checks. Each pot constituted the experimental plot. The rice accessions used as checks were: Moroberekan (drought resistant), Nerica4 (early maturing), RAM63 (late maturing) and B6144 (well adapted to the local site). ITA306 and TOS14519 were the susceptible and the resistant parents for resistance to AfRGM, respectively.

A total of thirteen quantitative and twenty six qualitative characters were measured at various growth stages using published descriptors of wild and cultivated rice (*Oryza* spp.) (Bioversity International *et al.*, 2007).

7-1-2-2 F_3 characterisation

The design used was the same as described above, but with minor modifications, including the checks, the basal fertiliser application, the plot and the block size and the number of characters measured. The number of checks was 8 instead of 6 because of two other progenies (500 and 467) that were added to balance the number of lines per block. Plot size for each accession was 0.6 m x 3.3 m occupied by three rows, with a spacing of 300 mm between and within rows and a distance of 400 mm between plots. Direct seeding at the rate of two plants per hill was done. Plot number per block was 58 (50 progenies + 8 checks) with a total number of 728 [650 + (6x13)] plots shared in 13 blocks. The fertiliser applied was NPK (10-18-18) at a rate of 200kg.ha^{-1}. Urea was also applied at the rate of 30 kg.ha^{-1} N two weeks after thinning and at panicle initiation. A diversity index was calculated for each character of the F_2 progenies. Only those found to discriminate with a high index diversity were used in the F_3 agro-morphological characterisation. Observations were made on the ten hills of the central row.

7-1-2-3 Crop characters investigated in field trials

A total of 15 quantitative characters and 26 qualitative characters were measured at various growth stages using the descriptors of wild and cultivated rice (*Oryza* spp.) (Bioversity International *et al.*, 2007). The quantitative characters included the penultimate leaf length (PLL), the penultimate leaf width (PLW), the flag leaf length (FLL), the flag leaf width (FLW), the ligule length (LigL), the tillering ability (Til), the number of days to booting (DB), the number of days to first heading (DFH), the number of days to heading (DH), the number of days to maturity (DM), the culm length (CL), the plant height (PH), the panicle length (PanL), the secondary branching (SB) and thousand grains weight (TGW). The qualitative characters included the basal leaf sheath colour, the leaf blade presence/absence of anthocyanin colouration, the leaf blade distribution of anthocyanin colouration, the leaf blade pubescence, the leaf blade pubescence on blade surface, the leaf margin pubescence, the auricle colour, the collar colour, the ligule shape, the ligule colour, the flag leaf attitude, the culm habit, the culm lodging resistance, the stigma colour, the lemma and palea color, the lemma colour of the apiculus, awn presence, awn distribution, awn colour, the awn length, the panicle arrangement of primary branches, the panicle attitude of main axis, the panicle attitude of branches, the panicle secondary branching, the panicle exertion and the panicle shattering.

7-1-3 Data analysis

The diversity index of both quantitative and qualitative character was calculated using the Shannon-Weaver diversity index (H'). The quantitative values were transformed into phenotypic classes, according to Kayode *et al*. (2007) and the descriptors of wild and cultivated rice (*Oryza spp*.) (Bioversity International *et al.,* 2007), prior to analysis, as described in Chapter 5. Jain *et al.* (1975) gave H' as:

$$H' = -\sum_{i=1}^{k} P_i \log_2 P_i$$

Where k is the number of phenotypic classes for a character and P_i is the proportion of the total number of entries in the i^{th} class.

7-1-4 Method for constitution of the mapping population

Progenies of the mapping population were chosen according to the following: The traits that showed no variability among progenies, because they exhibited one type of phenotype, were dropped. The diversity index was calculated for each remaining character to select the discriminant characters based on a diversity index threshold value above 0.4. The selected traits were thereafter involved in a neighbour joining tree cluster analysis to confirm the level of the variability revealed by the Shannon-Weaver index within the population. Once the variability was quantified, according to the retained traits, the entire population was analysed with MSTRAT (Gouesnard, 2001) to determine first, the number of the core collection samples, using its redundancy mode. The size of the core collection was run in a construction 100 times (number of replicates) and inferred with an iteration number of 500. The individuals of the core collections were, thereafter, chosen based on the 64 most frequent sampled progenies, in the 100 replications.

7-1-5 Validation of the core collection

The core collection was validated by comparing it to the entire population using descriptive statistics, the genotypic allelic segregation according to two SSR markers, RM317 and RM17303, and the Shannon-Weaver diversity index. These statistics included the minimum, the mean, the maximum and the coefficient of variation of each quantitative trait involved in the construction of the core collection. In addition, a clustering pattern of the core sample was

investigated, as well as broad-sense heritability estimates, in order to check the impact of the environment on the agro-morphological descriptors.

7-2 Results

7-2-1 Preliminary classification analysis of F_2 progenies according to quantitative and qualitative characters

The 15 quantitative characters were all capable of discriminating progenies. Seven of the 26 qualitative characters were dropped because they showed no variation amongst progenies. These characters included the leaf blade distribution of anthocyanin colouration where anthocyanin was distributed on the margin only, the leaf blade pubescence on blade surface where all progenies have hairs on their leaf upper surfaces, the leaf margin pubescence where all progenies have their leaf margin ciliated, the auricle colour where all auricles were of yellowish-green colour, the collar colour where all collars were green, the ligule shape which was of 2-cleft type and panicle shattering, which showed low shattering ability for all progenies.

7-2-2 Variability of traits within F_3 progenies

7-2-2-1 Genetic diversity

Genetic diversity was estimated according to the Shannon-Weaver diversity index. The index ranged from 0.99 for anthocyanin colouration to 0.17 for ligule length, with a mean value of 0.63. The variables were grouped into three categorical classes: one class of a high index value above 0.69, a median class with index values between 0.4 and 0.7, and a class of low index values, lower than 0.4. Traits of high index values included PLL, PLW, FLL, FLW, DH, DM, PH, number of secondary (NSB) branches, anthocyanin coloration (AnthCol), flag leaf attitude (FLfAt) and lodging (Lodg). The class of median index values included Til, DB, DFH, CL, PanL, TGW, anthocyanin on leave (AnthOL), leaf glabness (LeafGlab), stem attitude (StemAt), stigma colour (StigC), lemma palea colour (LemPalC), apiculus colour (ApiC), awn length (AwL), awn distribution (AwDist), panicle branches attitude (PanBrAt), secondary branches attitude (SBrAt) and panicle exertion (PanEx). Traits with low index values included LigL, ligule colour and panicle main axis attitude (PanMAxAt). These were dropped as they were

considered to be of low discriminant ability. Therefore, the minimum index value retained as diverse enough for the study was 0.41, and was recorded for lemma and palea colour (Table 7-1).

Table 7-1: Shannon-Weaver index calculated for 15 quantitative and 18 qualitative characters in a segregating F_3 population against AfRGM

Variables	H'	Variables	H'	Variables	H'
PLL	0.88	PH	0.81	StigC	0.44
PLW	0.76	PanL	0.49	LemPalC	0.41
FLL	0.88	NSB	0.83	ApiC	0.50
FLW	0.80	TGW	0.22	AwL	0.63
LigL	0.17	AnthCol	0.99	AwDist	0.50
Til	0.56	AntoL	0.69	Awn Color	0.50
DB	0.57	Leaf Glab	0.54	Awn Length	0.50
DFH	0.69	LiguleCol	0.34	PanMAxAt	0.34
DH	0.79	FLfAt	0.86	PanBrAt	0.49
DM	0.87	StemAt	0.53	SBrAt	0.50
CL	0.64	Lodg	0.71	PanEx	0.52

PLL: Penultimate leaf length, PLW: Penultimate leaf width, FLL: Flag leaf length, FLW: Flag leaf width, Til: Tillering ability, DB: Days to booting, DFH: Days to first heading, DH: Days to heading, DM: Days to maturity, CL: Culm length, PH: Plant height, PanL: Panicle length, NSB: Number of secondary branches, TGW: Thousand grain weight, AnthCol: Anthocyanin coloration, AnthCoL: Anthocyanin on leave, LeafGlab: Leaf glabness, FLfAt: Flag leaf attitude, StemAt: Stem attitude, Lodg: Lodging, StigC: Stigma colour, LemPalC: Lemma palea colour, ApiC: Apiculus colour, AwL: Awn length, AwDist: Awn distribution, PanMAxAt: Panicle main axis attitude, PanBrAt: Panicle branches attitude, SBrAt: Secondary branches attitude, PanEx: Panicle exertion.

7-2-2-2 Clustering pattern of the entire segregating population according to the UPGMA of the most diverse traits, according to DARwin software

Both quantitative and qualitative variables were transformed into classes prior to cluster analysis. Four distinct groups were observed, based on the neighbouring method of the unweighted pair group mean analysis. All of them showed secondary sub-divisions. A total number of 14 sub-groups were observed and included 5 sub-groups from Cluster A (SgA-1 to SgA-5), 4 sub-groups from Cluster B (SgB-1 to SgB-4), 2 sub-groups from Cluster C (SgC-1 and SgC-2) and 3 sub-groups from Cluster D (SgD-1 to SgD-3) (Figure 7-1).

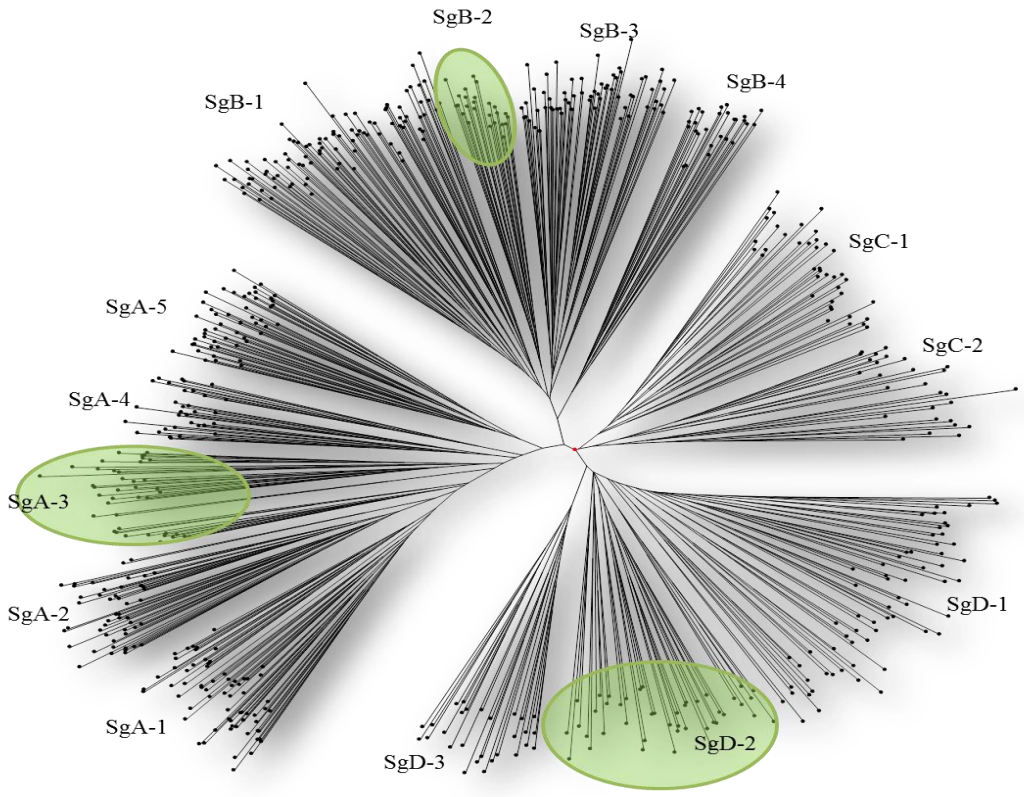


Figure 7-1: Diversity clustering pattern among 649 F_3 progenies of a ITA306-TOS14519 segregating population according to the UPGMA of the most diverse traits according to DARwin software

Globally, for each variable studied, none of the clusters depicted a clear grouping pattern. However, some exceptions were noticed, associating clusters or sub-clusters to a particular phenotype and can be summarised as follows:

Cluster A is the only cluster bearing individuals with short penultimate leaves and these individuals were spread within SgA-1 to SgA-3; with SgA-4 and SgA-5 including individuals with medium Penultimate leaf length. Clusters B, C and D included individuals of medium and long Penultimate leaf length. Only SgA-1, SgA-2, SgA-3 and SgC-2 included progenies with broad penultimate leaves, even though the frequency was low. Overall, Cluster A was characterised by short flag leaves with few individuals of medium flag leaf length in SgA-4 and SgA-5. The other clusters segregated for Flag leaf length. Progenies either segregated on short (very few), medium and broad flag leaf (Clusters B and D) or segregated in medium and broad

flag leaves (Cluster C). However, no clear clustering pattern was depicted for flag leaf width for Clusters A, B and D, which included progenies of narrow, medium and broad flag leaves. Cluster C, mostly in SgC-2, included progenies with broad flag leaves. Progenies with medium ligule length were only observed in Cluster A, and belonged to the three first sub-groups; the other clusters had progenies with long ligules. The Til was not distinguishable in Cluster A and B, ranging from low tillering to very good tillering ability. Conversely, Cluster C and D mostly included individuals with good or very high Til.

Apart from SgA-1, SgA-2 and SgA-4 in Cluster A, where some progenies were of medium booting ability, progenies booted either late or very late (Cluster C). Particularly in Cluster B and D, progenies booted very late, apart from SgB-4. The same clustering pattern was observed with days-to-first-heading and days-to-heading, with the exception that few progenies were early heading in Cluster A, medium heading in Cluster B and late heading in Cluster D. This was conversely to number of days-to-booting.

Apart from SgA-1, SgA-2 and SgA-4 in Cluster A, where progenies matured early, progenies were either medium or late maturing (Cluster C). Particularly in Clusters B and D, progenies were of late or very late maturing, apart from SgB-4.

Progenies belonging to Clusters A and C had Culm length between short to intermediate, while those from Clusters B and D had short culms and long culms respectively. In contrast, the clusters distinguished progenies for Plant height. Apart from SgD-3, which was characterised by tall and very tall progenies, the other sub-groups included progenies with variable height: very short and very tall progenies co-existed (SgB-3) and character states co-clustered (SgA-1, SgA-2, SgB-1 and SgD-1). The same trend was observed for panicle length, with the exception that the smallest panicle was of medium height. Likewise SgD-3 was characterised by long and very long panicles.

The level of secondary branching did not show any grouping pattern across clusters and sub-groups; panicle secondary branching ranged from low (<45) to highly branched (>90). Cluster D included a cluster bearing extreme panicle branching. SgD-3 was represented by a group of very dense branching (76-90) and clustered secondary branching. SgD-1 included a sub-group containing most of the progenies with little secondary branching.

In Cluster C, the progenies were of intermediate TGW, whereas the other clusters segregated either into light or intermediate TGW. However, SgA-1, SgA-2, SgB-1, SgB-4 and SgD-3 displayed the same pattern as Cluster C.

For the basal leaf sheath colouration, SgA-1 and SgA-2 represented sub-groups with green sheaths. In SgA-3, SgA-4, SgA-5 and the other clusters the sheath was either green with light purple, light purple and purple. However, SgD-3 included progenies with green sheaths with light purple colouration. In regard to leaf blade anthocyanin colouration, the SgD-1 sub-group represented this group. Overall, clusters did not segregate into either presence or absence of leaf blade anthocyanin colouration.

The leaf blade pubescence on blade leaf did not distinguish clusters. Progenies were either hairy on the upper surface or hairy on the lower surface across clusters and sub-groups.

For ligule colouration, progenies with white ligules were clustered in the sub-groups SgA-1 and SgA-2, while progenies with purple lines were grouped into the sub group SgA-5, the Clusters B, C and D except the sub-group SgD-3.

The flag leaf attitude, the culm habit and the culm lodging resistance also did not distinguish clusters.

For the flag leaf attitude, progenies had either erect, semi-erect, horizontal or descending flag leaves across clusters and sub-groups. However, SgD-2 displayed mostly semi-erect flag leaves. For the culm habit progenies had either erect ($<15^{o}$), semi-erect ($\sim20^{o}$), open ($\sim40^{o}$) or scarcely spreading ($>60\text{-}80^{o}$) culms across clusters and sub-groups. Concerning the culm lodging resistance, progenies had intermediate, strong or very strong culms across clusters and sub-groups. However, SgD-3 contained mostly progenies with weak culms.

Concerning the stigma color, apart from SgA-1, SgA-2, SgA-3 and SgA-4 in Cluster A, where progenies had white stigmas, progenies had either white or purple stigmas across clusters and sub-group.

Individuals from the SgA-1 and SgA-2 sub-groups mostly had yellowish-green palea and lemma, whereas the other clusters and sub-groups presented admixtures of two (SgD-2), three (Cluster C, SgA-5, SgD-2, SgD-3) or four (SgB-1) types of palea and lemma colouration. Progenies had brown furrows on green, blackish brown, yellowish green, and purple furrows on green palea and lemma.

Apiculus colour was white in SgA-1 and SgA-2, purple in SgB-1, SgB-2, SgB-3, Clusters C and D. It was an admixture of brown, purple, purple apex and occasionally black in the other remaining groupings.

Apart from progenies belonging to SgA-2 and Cluster B, which were partially awned, the other clusters and sub-groups were awnless. Only 4 and 2 progenies were partially awned in SgA-3 and Cluster C, respectively. The awn color and the awn length were distributed in clusters and sub-groups exactly the same as the awn presence is distributed. When awned, progenies had only awn on the tip of the panicle. However, the awn was white in SgA-2, purple for 4 progenies in SgA-3, red for 2 progenies in Cluster C and an admixture of white, purple and red awn in Cluster B and within Cluster B sub-groups. The awn length was evenly distributed, being very short (<5mm), short (~8mm), intermediate (~15mm), long (~30mm) or very long (>40mm) across clusters and sub-groups.

The panicle main axis attitude did not distinguish clusters. Progenies had mostly either slightly drooping or strongly drooping panicles within groupings, irrespective of the clusters and sub-groups. However, Clusters A and C could be considered as a strongly drooping panicle group.

Only Cluster C, which bore progenies with drooping panicle branching, distinguished progenies for panicle branching attitude. The other clusters presented a mixture of progenies of horizontal, drooping and, scarcely, spreading panicle branching.

For panicle secondary branching, SgA-1 and SgC-1 represented the dense and the clustered groups whereas the other sub-groups bore either dense or clustered panicle secondary branching.

For panicle exertion, most progenies had either “moderately well exerted” panicles or well exerted panicles with very few individuals having a just exerted panicle. The SgC-1 sub-group had exclusively individuals with well exerted panicles.

7-2-3 Constitution of the core collection

7-2-3-1 Determination of the size of the core collection

The redundancy curve (Figure 7-2) shows a convex pattern, confirming that progenies in the entire population were redundant in terms of allelic diversity richness. Moreover, it depicts an inflection point at a core size of 64 individuals. The maximum allelic richness of 391 was recorded at a core size of 85 individuals. In addition, the least variation of the allelic richness

was observed between the samples of 64 individuals and 85 individuals (380 and 391 respectively), suggesting that the optimum size of the core collection would be 64 individuals.

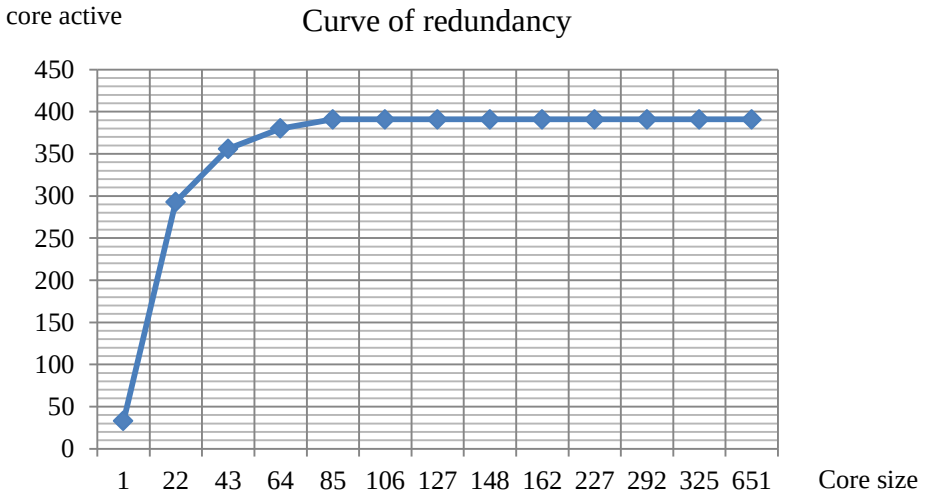


Figure 7-2: Curve of redundancy of 649 progenies showing the accumulation of the allelic richness according to the number of progenies representing the core size (100 iterations, 10 replicates)

7-2-3-2 Identification of the individuals of the core collection

The construction inferred at 500 iterations for a core size of 64 individuals and replicated 100 times, showed the frequency of sampling of the individuals throughout replicates. The first 64 top sampled individuals constituted the representative sample of the core collection. Their names and their frequency of sampling are provided in Table 7-2. The frequency of appearance ranged from 40 to 100 times, with a mean appearance of 81.77 times. Approximately 50% (30 progenies) of the total progenies (64 progenies) of the core collection were sampled in all the 100 replicates, while only 4 progenies appeared less than 50 times during replication. The least sampled progeny was Progeny 170 and it was sampled 40 times. Almost all the individuals (60 individuals representing 93.75% of the core sample) of the core collection, representing 10% of the entire population, were sampled at least one time out of two during the 100 replicates (Table 7-2).

Table 7-2: Frequency distribution of appearance of the genotypes after a construction run at 500 iterations and replicated 100 times according to MSTRAT 4.1 software

Geno	% sampling*	Geno	% sampling	Geno	% sampling	Geno	% sampling
37	100	352	100	488	91	545	63
73	100	353	100	581	88	81	60
89	100	411	100	116	85	2	58
111	100	420	100	164	85	502	57
133	100	452	100	43	83	53	56
142	100	494	100	586	80	551	55
148	100	501	100	28	79	178	54
182	100	507	100	23	78	159	53
204	100	509	100	65	71	225	53
221	100	525	100	566	71	654	53
230	100	541	100	74	69	49	52
232	100	587	100	6	66	340	50
278	100	602	100	375	66	70	47
318	100	604	100	418	66	463	45
347	100	466	96	518	66	51	42
349	100	487	91	500	64	170	40

Geno=Genotype, * represented the frequency of a given progenies to be sampled during the 100 replicates

7-2-3-3 Validation of the core collection

7-2-3-3-1 Comparison of the core collection with the entire population

Apart from PanL, SB and TGW, the diversity index in both the entire population and core collection were similar. This index ranged from 0.17 to 0.88 for the whole population and from 0.17 to 0.91 for the core collection. Their corresponding mean values were 0.66 and 0.67, respectively. Moreover, both samples were similar for minimum and maximum values of the traits under study. In contrast, mean values and coefficient of variation were different for the entire population and core collection, with values for the core collection being larger than those of the entire population. Variation within the core collection was above 20% for all traits measured, except PanL. Both populations were, however, similar for PLW (1.18 vs. 1.27), FLW (1.39 vs. 1.46), PanL (31.18 vs. 31.69) and TGW (24.84 vs. 24.09) (Table 7-3).

Table 7-3: Comparison of the minimum, maximum, mean, coefficient of variation and diversity index of the entire population and the core collection

Traits	Whole population					Core collection				
	Min	Max	Mean	CV(%)	H	Min	Max	Mean	CV(%)	H
PLL	21.5	91.5	53.23	22.72	0.88	24.2	91.5	57.73	25.69	0.91
PLW	0.7	2.0	1.18	19.92	0.76	0.7	2.0	1.27	20.69	0.77
FLL	12.8	83.0	45.90	28.13	0.88	12.8	82.5	49.72	31.75	0.97
FLW	0.9	2.2	1.39	17.93	0.80	0.9	2.2	1.46	20.41	0.89
Lig	11.0	45.0	23.85	21.05	0.17	11.0	45.0	26.28	23.82	0.17
Til	7.0	80.0	31.95	41.26	0.56	9.0	77.0	37.42	43.85	0.49
DB	64.0	162.0	96.76	19.28	0.57	66.0	162.0	104.00	23.75	0.49
DFH	42.0	173.0	106.99	18.59	0.69	42.0	173.0	112.94	24.21	0.67
DH	79.0	190.0	114.89	17.88	0.79	80.0	190.0	123.02	21.03	0.74
DM	88.0	217.0	134.87	16.40	0.87	88.0	217.0	142.17	20.42	0.88
CL	63.0	178.0	112.29	20.32	0.64	66.0	178.0	122.38	23.51	0.65
PH	90.0	214.0	143.46	16.94	0.81	92.0	214.0	153.94	20.53	0.81
PanL	15.0	54.0	31.18	12.12	0.49	15.0	53.0	31.69	19.85	0.66
SB	8.0	68.0	31.52	29.20	0.83	8.0	66.0	35.38	36.38	0.41
TGW	7.9	36.4	24.84	17.92	0.22	7.9	36.4	24.09	25.78	0.53

PLL: Penultimate leaf length, PLW: Penultimate leaf width, FLL: Flag leaf length, FLW: Flag leaf width, Til: Tillering ability, DB: Days to booting, DFH: Days to first heading, DH: Days to heading, DM: Days to maturity, CL: Culm length, PH: Plant height, PanL: Panicle length, NSB: Number of secondary branches, TGW: Thousand grains weight, CV=Coefficient of variation and H=Shannon-Weaver index.

7-2-3-3-2 Validation of the core collection, based on broad-sense heritability estimates

Broad-sense heritability (H^2) estimates were picked from Chapter 5. They were very high, with a mean value of 0.95. The minimum was 0.71 for TGW, while the maximum was 1, for PLW and FLW (Table 7-4).

Table 7-4: Broad-sense heritability estimates for 13 agro-morphological quantitative descriptors

Character	PLL	PLW	FLL	FLW	LigL	Til	DB
Broad-sense heritability	0.99	1	0.99	1	0.96	0.99	0.95

Table 7-4: continued

Character	DFH	DH	DM	CL	PH	PanL	SB	TGW
Broad-sense heritability	0.96	0.89	0.98	0.99	0.99	0.95	0.95	0.71

7-2-3-3-3 Clustering pattern of the core collection

The core collection was distributed in four clusters, similar to the clusters of the entire population. Only Cluster B clearly demarcated into two sub-clusters, allowing the core collection to be grouped into five distinct groups. As with the whole population, clustering in the core collection did not map onto phenotypic clustering. Many phenotypes co-existed between and within clusters, and even within sub-clusters. The following trends were noted:

Sixty one of the 64 individuals of the core sample exhibited long ligules. Only 3 had medium LigL. Clusters B, C and D bore progenies booting very late. Cluster A included progenies with late and medium booting, while only two (Clusters C and D) exhibited very late DFH, DH and DM. Cluster A mostly had progenies without anthocyanin. Only 11 progenies, spread across clusters, had anthocyanin on leaves. Progenies in Clusters B, C and D had purple apiculus, while Cluster A contained a mixture of brown, red apexes, purple, purple apexes, blackish and mostly white apiculi. Most progenies of the core sample were awnless, apart from those of Cluster C, which contained white, red and purple partially awned progenies. Most progenies were strongly drooping and only one-fourth were slightly drooping and spread in all clusters.

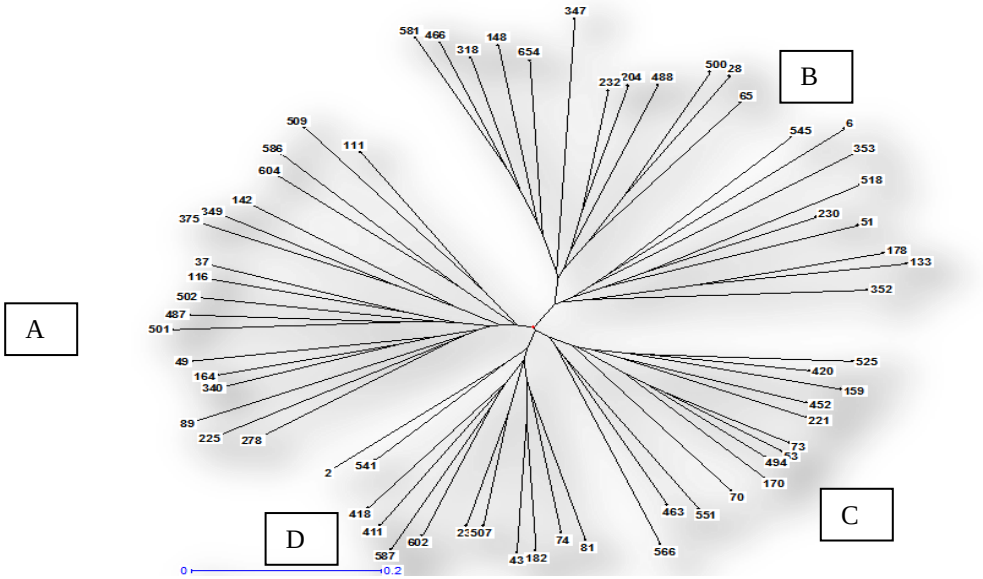


Figure 7-3: Unrooted neighbour joining tree of the 64 progenies of the core collection built based on the simple matching distances of the progenies according to agro-morphological traits

7-2-3-3-4 Validation of the core collection based on genotypic segregation

Both the core collection and the entire population showed a genotypic segregation of 1:2:1 for the two SSR markers. RM317 segregated for 156, 167 and 328 individuals homozygous for

Allele A, homozygous for Allele B and heterozygous, respectively, with a khi^2 value of 0.41 for the entire population. It segregated for 16, 13 and 35 individuals homozygous for Allele A, homozygous for allele B and heterozygous, respectively, with a khi^2 value of 0.84 for the core collection. Likewise, RM17303 segregated for 149, 148 and 345 individuals homozygous for Allele A, homozygous for Allele B and heterozygous, respectively, with a khi^2 value of 3.59 for the whole population. It also segregated for 12, 13 and 39 individuals homozygous for Allele A, homozygous for Allele B and heterozygous, respectively, with a khi^2 value of 3.09 for the core collection. For each SSR marker and for both populations, the khi^2 p-values were not significant, meaning that markers and alleles were not co-segregating (Table 7.5).

Table 7-5: Comparison of both entire population and core collection, based on genotypic segregation

Whole segregating population							
Markers	Homo A	Homo B	Hetero	N	khi^2	p-value	Segregation
RM 317	156	167	328	651	0.41	0.81459^{ns}	1:2:1
RM17303	149	148	345	642	3.59	0.16597^{ns}	1:2:1
Core collection							
RM 317	16	13	35	64	0.84	0.65582^{ns}	1:2:1
RM17303	12	13	39	64	3.09	0.21291^{ns}	1:2:1

Homo=Homozygous, Hetero=Heterozygous, N=Population size, ns=non significant

7-3 Discussion

A core collection is a sub-sample of a larger germplasm collection that contains, with a minimum of repetitiveness, the maximum possible genetic diversity of the species in question (Frankel, 1984). Using it as an alternative means to represent the random selection of the individuals of the mapping population, is a good method but its accuracy relies upon the method used to constitute this core. For instance, Brown (1989) proposed to hierarchically stratify the larger population into groups before creating a core collection. Hu *et al*. (2000) proposed to build the core collection based on genotypic values that retained larger genetic variability and had superior representatives than those based on phenotypic values. The first requirement was checked in our study by estimating the diversity index and by clustering

individuals, using the unweighted pair group method, using the arithmetic average (UPGMA method according to DARwin as reported by Sneath and Sokal (1973). Moreover, the MSTRAT procedure used in our study maximises the genetic diversity by maximising allelic or phenotypic richness between groups, using qualitative and quantitative variables (Gouesnard *et al*., 2001).

Phenotypic values used in research to construct core sub-sets has been reported earlier for various crop by many authors (Rodiño *et al*., 2003; Okpul *et al*., 2004; Volk *et al*., 2005; Yan *et al*., 2007). Hu *et al*. (2000) also reported the effects of interactions between genotype and the environment (GxE effects) to exist in phenotypic values. Therefore, a core sample based on phenotypic values as in this study may not accurately represent the genetic diversity of the initial germplasm group (Tanksley and McCouch, 1997). In this study, research reported in from Chapter 5 showed high broad-sense heritability estimates (ranging from 0.71 to 1.0) and postulated that the variability observed in the population was due to genetic causes rather than environmental causes. Moreover, the genetic cause was postulated as being a combination of dominance and epistatic gene effects, rather than additive genetic effects. The interaction between genotype and environment is therefore not strong enough to bias the composition of the core sample (McKhann *et al*., 2004).

The high value of the Shannon-Weaver index (Shannon and Weaver, 1949), as well as the clustering pattern displayed by the neighbour joining tree performed on the whole population, confirmed that the variability within the progenies was high. Moreover, the fact that generally, the clustering pattern did not distinguish between individuals of particular phenotypic traits, suggested that many traits were contributing concomitantly to the diversity, and subsequently, clusters could be assumed to have high allelic richness. Indeed, a principal component analysis study revealed that, for 60% of the variability explained by the first two axes, none of the individual contribution of the variables was above 14 and 12%, respectively, for Factor 1 and Factor 2 (See Section 5-2-4 in Chapter 5). Therefore, selecting the core sample based on these clusters also led to greater allelic diversity.

Previous findings by many authors (Kato, 1990; Rabiei *et al*., 2004; Vanaja and Babu, 2006; Fahliani *et al*., 2010) and Chapter 4, revealed that the quantitative descriptors used in the present study and the resistance to AfRGM were highly heritable. In addition, traits such as Ligule length, Culm length or Plant height, leaf length, leaf width and plant maturity, which are

positively associated with the AfRGM resistance (when they are of high value), were also correlated, (See Chapter 5, Table 5-5). Therefore, selecting for these traits also means selecting alleles involved in AfRGM resistance for the core collection. However, some representatives of the core collection did not bear alleles involved in AfRGM resistance because the core collection is also composed of contrasting individuals (Frankel and Brown, 1984) some of which, express susceptibility to AfRGM instead of resistance.

The comparison of the core collection to the entire population shows similarities for, the minimum and maximum values of the traits being study, genotypic allelic segregation with two SSR markers and the diversity index. This confirms that, not only the extreme values of the whole population were selected in the core sample, but also almost all the genetic diversity of the overall population had been represented in the core collection. Consequently, this also means that the individuals were accurately sampled. However, both populations were different for the mean phenotypic value and the coefficient of variation with larger values for the core sample values than the entire population values. This is probably due to the fact that, the core collection being of a reduced size (nearly 10% of the entire population) and bearing the extreme values, the variation within the core is more pronounced than the whole population where the greater number of individuals reduces contrast between progenies. The lower phenotypic mean value of the entire population may be due to increased redundancy of low phenotypic values which contribute to lower the overall mean of the population. This assertion is supported by the inbred nature of the population used here, which is subjected to less variation than natural germplasm. However, the fact of being an early generation population compensates the loss of variability by representing most of the segregating genes of their parents.

Comparing our core sample to others worldwide is risky because they have different biological characteristics, the original collection did not reach the same global diversity of the species, the analyses are seldom performed in the same way (Le Cunff *et al.*, 2008) and the nature of the population investigated is different. Moreover, with the size of the investigated population growing, the percentage of the core sample size decreases, as reported by Wang *et al*. (2006) in a study where the percentages of the core size were only 1% and 2% for a population size of 2710 accessions. Nevertheless, we can claim that our core sample size, representing 10% of the entire population, agrees to the proportion found by Brown (1989), which is considered to be a

good proportion to represent the whole population's variability (Gouesnard *et al*., 2001; Wang *et al*., 2008; Zhang *et al.*, 2010). Conversely, Diwan and McIntosh (1995) claimed that the best representation of an entire population by a core size was 17% in a similar study conducted on a collection of annual *Medicago* species.

To date, core collections of many plants have been constructed, including rice, wild bean, *Capsicum*, *Medicago*, peanut, barley, peach and strawberry (Diwan and McIntosh, 1995; Tohme and Gonzalez, 1996; Geibel *et al*., 2004; Zewide *et al*., 2004; Corley and Dung, 2005). In addition, their use for association mapping has become a routine during the last three years (Borba *et al*., 2010; Zhao *et al*., 2010). Conversely to association mapping, no study has been conducted concerning the use of a core collection as a basis for constituting a mapping population. However, the theoretical feasibility has been documented in few reports (Lou *et al*., 2007; Richards *et al*., 2009). Therefore, the present study represents the preliminary steps towards the use of a core collection for a gene/QTL mapping purpose in view of genome wide mapping.

Overall, genes can be mapped and tagged using a core collection as a reduce mapping population. However, to build a core collection using agro-morphological descriptors, one should: (1) Ascertain the variability within the population for further easy stratification; (2) Confirm that this variability is due to genetic causes rather than environmental causes by checking the variance components; (3) Retain as active variables only descriptors with high heritability, in order to avoid GxE interactions; (4) Use the MSTRAT 4.1 procedure (Gouesnard *et al.*, 2001) which is the most powerful software currently available for core collection building, using agro-morphological descriptors.

References

Bioversity International, IRRI and WARDA. (2007). *Descriptors for wild and cultivated rice (Oryza spp.).* Bioversity International, Rome, Italy; International Rice Research Institute, Los Baños, the Philippines; WARDA, Africa Rice Center, Cotonou, Benin.

Borba, T.C.d.O., Brondani, R.P.V. Breseghello, F., Coelho, A.S.G., Mendonça, J.A., Rangel, P.H.N. and Brondani, C. (2010). Association mapping for yield and grain quality traits in rice (*Oryza sativa* L.). *Genetics and Molecular Biology* 33:515-524.

Brown, A.H.D. (1989). A case for core collection. In: *The Use of Plant Genetic Resources* Brown, A.H.D., Frankel, O.H., Marshall, D.R. and Williams, J.T., (Eds). Cambridge University Press, Cambridge, United Kingdom. p 136-156.

Corley, H. and Dong, W. (2005). Development and evaluation of a mini core collection for the U.S. peanut germplasm collection. *Crop Science* 45:1540-1544.

Darvasi, A. and Soller, M. (1994). Selective DNA pooling for determination of linkage between a molecular marker and a quantitative trait. *Genetics* 138:1365-1373.

Diwan, N. and McIntosh B.G.R. (1995). Methods of developing a core collection of annual Medicago species. *Theoretical and Applied Genetics* 90: 755-761.

Fahliani, R.A., Khodambashi, M., Houshmand, S. and Arzani. A. (2010). Estimation of the heritability of agro-morphological traits in rice (*Oryza sativa* L.) using F2:3 families. *African Journal of Agricultural Research* 5:1297-1303.

Federer, W.T. (1956). Augmented designs. *Hawaiian Planter's Record* 55:191-208.

Frankel, O.H. (1984). Genetic perspectives of germplasm conservation. In: *Genetic Manipulation: Impact on Man and Society*. Arber, W., Llimensee, K., Peacock, W.J. and Starlinger, P. (Ed). Cambridge University Press, Cambridge, United Kingdom. p 161-170.

Frankel, O.H. and Brown, A.H.D. (1984). Plant genetic resources today: a critical appraisal. In: *Crop Genetic Resources: Conservation and Evaluation*. Holden, J.H.W. and Williams, J.T. (Ed). George Allen and Unwin, London, United Kingdom. p 249-257.

Geibel, M., Masny, A. and Coman, M. (2004). The European strawberry database and building up a European core collection. *Acta Horticulturae* 649:41-44.

Gouesnard, B., Bataillon, T.M., Decoux, G., Rozale C., Schoen, D.J. and David, J.L. (2001). MSTRAT: An algorithm for building germplasm core collections by maximizing allelic or phenotypic richness. *Journal of Heredity* 92:93-94.

Hu, J., Zhu, J. and Xu, H.M. (2000). Methods of constructing core collections by stepwise clustering with three sampling strategies based on the genotypic values of crops. *Theoretical and Applied Genetics* 101:264-268.

Jain, S.K., Qualset, C.O., Bhatt, G.M., Wu, K.K. (1975). Geographical patterns of phenotypic diversity in a world collection of durum wheats. *Crop Science* 15:700-704.

Kato, T. (1990). Heritability for grain size of rice (*Oryza sativa* L.) from parent-offspring correlation and selection response. *Japan Journal of Breeding* 40:313-320.

Kayode, A.S., Fawole, I., Guei, R.G., Ojo, D.K., Somado, E.A., Tia, D.D., Ogunbayo, S.A. and Sanchez, I. (2007). Geographical patterns of phenotypic diversity in *Oryza sativa* landraces of Cote d'Ivoire. *Euphytica* 160:389-400.

Lander, E.S. and Botstein, D. (1989). Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185-199.

Lebowitz, R.L., Soller M, Beckmann, J.S. (1987). Trait-based analysis for the detection of linkage between marker loci and quantitative trait loci in cross between inbred lines. *Theoretical and Applied Genetics* 73:556-562.

Le Cunff, L., Fournier-Level, A., Laucou, V., Vezzulli, S., Lacombe, T., Adam-Blondon, A-F., Boursiquot, J-M. and This, P. (2008). Construction of nested genetic core collections to optimize the exploitation of natural diversity in *Vitis vinifera* L. subsp. *sativa. BMC Plant Biology* 8:31-42.

Lou, P., Zhao, J., Song, X., Pino Del Carpio, D., Shen, S., Vreugdenhil, D., Wang, X., Koornneef, M., and Bonnema, A.B. (2007). Mapping quantitative trait loci for morphological traits in multiple populations of *Brassica rapa*. *Journal of Experimental Botany* 58:4005-4016.

McKhann, H.I., Camilleri, C. Bérard, A., Bataillon, T., David, J.L., Reboud, X., Le Corre, V., Caloustian, C., Gut, I.G. and Brunel, D. (2004). Nested core collections maximizing genetic diversity in Arabidopsis thaliana. *The Plant Journal* 38:193-202.

Okpul, T., Singh, D., Gunua, T., Wagih, M.E. (2004). Assessment of diversity using agro-morphological traits for selecting a core sample of Papua New Guinea taro (*Colocasia esculenta* (L.) Schott) collection. *Genetic Resources and Crop Evolution* 51:671-678.

Pérez-Enciso, M. (1998). Sequential bulked typing: a rapid approach for detecting QTLs. *Theoretical and Applied Genetics* 96:551-557.

Rabiei, B., Valizadeh, M., Ghareyazie, B. and Moghadam, M. (2004). Evaluation of selection indices for improving rice grain shape. *Field Crop Research* 89: 359-367.

Richards, CM., Volk, G.M., Reilley, A.A., Henk, A.D., Lockwood, D.R., Reeves, P.A. and Forsline, P.L. (2009). Genetic diversity and population structure in *Malus sieiersii,* a wild progenitor species of domesticated apple. *Tree Genetics and Genomes* 5:339-347.

Rodiño, A.P., Santalla, M., Ron, A.M.D. and Singh, S.P. (2003). A core collection of common bean from the Iberian peninsula. *Euphytica* 131:165-175.

Semagn, K., Bjørnstad, Å., Skinnes, H., Marøy, A.G., Tarkegne, Y. and William, M. (2006a). Distribution of DArT, AFLP and SSR markers in a genetic linkage map of a double haploid hexaploid wheat population. *Genome* 49:545-555.

Semagn, K., Bjørnstad, Å. and Ndjiondjop, M.N. (2006b). An overview of molecular marker methods for plants. *African Journal of Biotechnology* 5:2540-2568.

Semagn, K., Bjørnstad, Å. and Ndjiondjop, M.N. (2006c). Progress and prospects of marker assisted backcrossing as a tool in crop breeding programs. *African Journal of Biotechnology* 5:2588-2603.

Shannon, C.E. and Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, USA.

Sneath, P.H. and Sokal, R.M. (1973). *Numerical Taxonomy. The Principles and Practice of Numerical Classification.* Freeman W.H. and Co., San Francisco, USA.

Sun, Y., Wang, J., Crouch, J.H. and Xu, Y. (2010). Efficiency of selective genotyping for genetic analysis of complex traits and potential applications in crop improvement. *Molecular Breeding* 26:493-511.

Tanksley, S.D. and McCouch, S.R. (1997). Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 277:1063-1066.

Tohme, J. and Gonzalez, D. (1996). AFLP analysis of gene pools of a wild bean core collection. *Crop Science* 36: 1375-1384.

Vanaja, T. and Babu, L.C. (2006). Variability in grain quality attributes of high yielding rice varieties (*Oryza sativa* L.) of diverse origin. *Journal of Tropical Agriculture* 44:61- 63.

Vision, T.J., Brown, D.G., Shmoys, D.B., Durrett, R.T. and Tanksley, S.D. (2000). Selective mapping: A strategy for optimizing the construction of high-density linkage maps. *Genetics* 155: 407-420.

Volk, G.M., Richards, C.M., Reilley, A.A., Henk, A.D., Forsline, P.L. and Aldwinckle, H.S. (2005). Ex situ conservation of vegetatively propagated species: development of a seed-based core collection for *Malus sieversii*. *Journal of American Society of Horticultural Science* 130:203-210.

Wang, L., Guan, Y., Guan, R., Li, Y., Ma, Y., Dong, Z., Liu, X., Zhang, H., Zhang, Y., Liu, Z., Chang, R., Xu, H., Li, L., Lin, F., Luan, W., Yan, Z., Ning, X., Zhu, L., Cui, Y., Piao, R., Liu, Y., Chen, P., and Qiu, L. (2006). Establishment of Chinese soybean (Glycine max L.) core collections with agronomic traits and SSR markers. *Euphytica* 151:215-223.

Wang, J.C., Hu, J., Huang, X.X. and Xu, S.C. (2008). Assessment of different genetic distances in constructing cotton core subset by genotypic values. *Journal of Zhejiang University of Science* 9:356-362.

Yan, W.G., Ruter, J.N., Bryant, R.J., Bockelman, H.E., Fjellstrom, R.G., Chen, M.H., Tai, T.H. and McClung, A.M. (2007). Development and evaluation of a core subset of the USDA rice germplasm collection. *Crop Science* 47:869-876.

Zewdie, Y., Tong, N.K. and Bosland, P. (2004). Establishing a core collection of *Capsicum* using a cluster analysis with enlightened selection of accessions. *Genetic Resources and Crop Evolution* 51:147-151.

Zhang, J., Wang, Y., Zhang, X., Li, T., Wang, K., Xu, X. and Han, Z. (2010). Sampling strategy to develop a primary core collection of apple cultivars based on fruit traits. *African Journal of Biotechnology* 9:123-127.

Zhao, J., Artemyeva, A., Del Carpio, D.P., Basnet, R.K., Zhang, N., Gao, J., Li, F., Bucher, J., Wang, X., Visser, R.G. and Bonnema, G. (2010). Design of a *Brassica rapa* core collection for association mapping studies. *Genome* 53:884-898.

Chapter 8

Identification of flanking SSR markers associated with major QTL of resistance to African Rice Gall Midge (AfRGM) using bulked segregant analysis (BSA): progress and future prospects

Abstract

The ongoing impact of the damage caused by the African Rice Gall Midge (AfRGM), *Orseolia oryzivora* Harris and Gagné, to lowland rice and more recently to upland rice, drive the imperative to develop varietal resistance to manage the pest, have been proved. Therefore, mapping and tagging genes and QTLs for resistance to AfRGM could assist African rice breeders. A mapping population was developed using *indica* varieties, including ITA306 (susceptible) and TOS14519 (resistant). A study, looking for polymorphism involving 270 SSR markers, was performed amongst these parental lines, previously identified through a multi-location trial conducted in 4 sub-Saharan countries. Concurrently, paddy rice screenhouse screening of F_2 and F_3 populations was done at 70 days after transplanting (DAT) for phenotyping and genetic studies. The polymorphism study identified 145 polymorphic SSR markers, with an average of 10 markers per chromosome, with Chromosomes 10 and 11 showing the highest level of polymorphism, amounting to 66.67% and 72.22%, respectively. The least polymorphic chromosome was Chromosome 1, with a polymorphism level of 39.4%. The BSA identified RM317 and RM17303, two SSR markers as linked to AfRGM resistance. The BSA analysis performed on the 10 most susceptible and resistant lines confirmed that RM317 and RM17303 should be SSR markers flanking the QTLs for AfRGM resistance located on the long arm of Chromosome 4. The subsequent linkage analysis performed on the whole population revealed that these markers were linked to QTLs or genes involved in the resistance against AfRGM.

Introduction

The African rice gall midge (AfRGM), *Orseolia oryzivora* Harris and Gagné, is one of the most destructive insects of rice in the rice-growing areas in 20 west and central African countries (Ukwungwu *et al.*, 1989; Ukwungwu and Joshi, 1992; Dakouo *et al.*, 1988; Nwilene *et al.*,

2006). The economic importance of AfRGM has drawn the attention of agricultural scientists and the insect has been the subject of intensive investigation during the last decade. However, cultural practices, biological control and chemical control of AfRGM are often ineffective. Therefore, growing and breeding of resistant varieties is the most effective way to control the target pest (Heinrichs, 1994; Kaityar *et al*., 1995). Identification and utilisation of new genes conferring AsRGM resistance has been a major objective of rice breeding in Asia. So far, 11 gall midge resistance genes have been identified, tagged, mapped and named. These genes include 10 dominant genes (*Gm1, Gm2, Gm4, Gm5, Gm6, Gm7, Gm8, Gm9, Gm10* and *Gm11*), and one recessive gene (*gm3*). Nine of these genes have been tagged and mapped (Kumar *et al*., 2005; Himabindu *et al*., 2007 and Himabindu *et al*., 2010). Flanking markers have been used to identify the resistance genes *Gm1* and *Gm2* in various rice cultivars (Himabindu *et al*., 2007). However, no gene mapping or gene tagging for AfRGM has been reported to date. Genetic analyses reported by Maji *et al.* (1998) and our research finding from Chapter 3, suggested that a single dominant gene and two recessive genes/QTLs controlled the resistance to AfRGM in the rice landraces TOG7106 and TOS14519, respectively. TOG7106 and TOS14519 are from Casamance (Senegal) and the Gambia, respectively. Mapping and tagging these genes is an essential step for their efficient use in marker assisted breeding. The identification of pest resistance genes and their incorporation into adapted germplasm is a major challenge for plant breeders, because assays for pest resistance are often based upon qualitative responses and fail to identify specific resistance genes without the use of biotype testing. An easily scorable marker linked to the gene, conferring a particular resistance phenotype would therefore represent an important tool for plant breeders. Furthermore, information on the chromosomal and genetic map location of a gene provides an alternative route to gene isolation and cloning. Preliminary studies leading to gene mapping and tagging, such as the identification of flanking SSR markers associated with QTLs linked to AfRGM resistance, was conducted in this study. The objectives were to: (1) Identify polymorphic markers between the two parental lines; (2) Use and estimate the efficiency of a core collection, compared with a random choice of individuals, as an alternative to a reduced mapping population; and (3) Identify flanking SSR markers associated to genes/QTLs linked to AfRGM resistance.

8-1 Material and method

8-1-1 Screening experiments for identification of parents

Field screenings were conducted to identify lines to be used in our analysis. *Oryza sativa* lines, *Oryza glaberrima* accessions from the Genetic Resources Unit (GRU) of the International Institute for Tropical Agriculture (IITA) and Africa Rice Centre (AfricaRice), and their interspecific crosses were screened against the AfRGM in a paddy screenhouse and under field conditions across four West African countries in the year 2000, during the July to September rainy season (Nwilene *et al*., 2002). TOG7106, TOS14519, Cisadane and BW348-1 were found to be resistant, moderately resistant and tolerant to AfRGM for the last two, respectively. However, the last two were chosen for the study because they have been widely released and adopted in Nigeria. The choice of resistant varieties was directed toward TOS14519 and TOG7106 as no *O. sativa* or *O. glaberrima* accessions performed better than them.

8-1-2 Development and validation of F_1 generations

ITA306, a high-yielding *O. sativa* subsp. *indica* but highly susceptible to AfRGM, was used as the female parent and was crossed with TOS14519, an *O. sativa* subsp. *indica* used as the male parent resistant to AfRGM. To effect emasculation, mature spikelets from a selected panicle were incised with a pair of scissors at the top fourth of its length, to expose the anthers that were then removed with small pliers, and covered thereafter with a seed envelope. Emasculation was conducted early in the morning between 08h30 and 10h00. Anthesis took place between 10h30 and 12h00, during which time a panicle shedding pollen was gently dusted onto the emasculated spikelet of the female parent to effect pollination. It was then covered with 80 x 130 mm waterproof, transparent, crossing, envelope and appropriately labelled.

Later, the efficiency of the crosses of the F_1s generation was evaluated: (1) By morphological assessment by looking at the plant vigour or/and the basal sheath anthocyanin colouration; and (2) By molecular analysis by screening the progenies with three highly polymorphic SSR markers obtained from the polymorphism study among parental lines. The principle of the validation through vigour assessment and the basal sheath anthocyanin colouration rely on the fact that F_1 plants are more vigorous than their parents, and that the F_1 plants basal sheath

colouration could only come from TOS14519 (male parent) since ITA306, the female parent, does not have a coloured sheath.

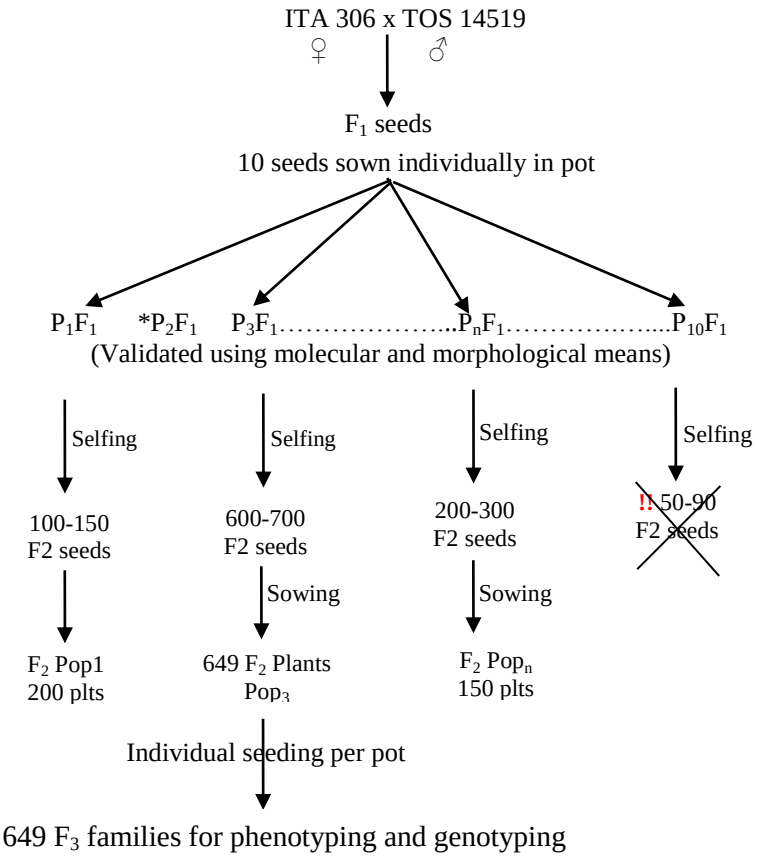
8-1-3 Development of the mapping populations and phenotyping

Validated F_1 plants were advanced to the F_2 in the greenhouse. Dormancy was broken by keeping the harvested seeds from individual F_1 plants in the incubator for 14 days at 45^{o}C. Then some seeds were dehulled manually, before germinating them in petri dishes or *in vitro*. Naked F_1 seeds were first pre-germinated in petri dishes for intraspecific crosses or *in vitro* for interspecific crosses. Three-week old seedlings were transplanted into 12-litre plastic buckets, which were 3/4-filled with paddy soil, at the rate of one seedling per bucket. Regular watering was done until plant maturity.

Six hundred and forty nine F_2 seeds from one single F_2 plant were advanced in the field to generate 649 F_3 families. Three-week old seedlings were transplanted in 12-litre plastic buckets. One gram of N, P_2O_5 and K_2O each were added in each pot at transplanting, using NPK 10-18-18 as source. The pots were hand-weeded whenever necessary. At maturity, individual F_2 plants were harvested separately and numbered for evaluation as F_3 families.

8-1-4 Choice of the mapping population

While developing the mapping populations, an analysis for polymorphisms was conducted amongst the parental lines. A total of 246 and 270 SSR markers from the core map, recognised as polymorphic for an Azucena-IR64 referenced population, were used and showed high levels of polymorphism for ITA306-TOG7106 and ITA306-TOS14519 crosses. Considering their high level of polymorphism and their resistance to AfRGM, ITA306-TOG7106 and ITA306-TOS14519 populations were advanced to backcross 2 (BC_2) and F_3 generation, respectively (Figure 8-1). Because of the low fertility of F_2 individuals due to the interspecific cross, the F_3 generation was chosen for the mapping and tagging of the AfRGM gene/QTLs for resistance, while the BC_2 was used to develop a BC_3 population for further investigations.



* means the plant was not validated as true F_1,

!! means the number of seeds was not enough for a mapping population as the minimum required is 150 seeds. At each step, seeds were harvested separately, per plant.

Figure 8-1: Scheme used to develop the mapping population ITA306 x TOS145198-1-5 Phenotyping

8-1-5 Phenotyping

8-1-5-1 Paddy greenhouse screening

The experiment was conducted between January and August 2010 at the AfricaRice-Ibadan facility, hosted by the IITA, Ibadan Station. Seeds for the entries were sown in a nursery bed. The F_3 families were transplanted 14 DAS in a paddy screenhouse, in a 2-m row with a space of 200 mm within and between rows, with the two parents involved in the cross as checks. The experimental plot was flanked by an infestation band consisting of five rows of ITA306 as the susceptible check, where the AfRGM population was maintained. The equivalent of 40 kg each of N, P_2O_5 and K_2O per ha was applied at transplanting using NPK 15-15-15 (% a.i.). Another 40

kg Nha^{-1} was applied 21 days after transplanting, using urea. The plots were hand-weeded twice at 21 and 40 days after transplanting. The experimental design was an augmented design with 2 checks, replicated every ten lines.

Galls, which are symptoms of insect damage or levels of infestation, were counted on all the 20 hills in each row 70 days after transplanting.

8-1-5-2 Phenotyping the F_2 population

Because of the large number of individuals (649) and because of difficulties associated with ensuring an even attack of the AfRGM on each individual plant, as well as the difficulty of individual placement of larvae on plant, the phenotyping and the levels of resistance were ascertained by exposing the resistant individuals to a second screening, in which conditions favoured the pest even more. The second screening was slightly modified from the first: (1) the spreader row infestation band was more densely transplanted; (2) the experiment was conducted from April to July 2010, which was a rainy and hot period; and 3) the reduction of the distance between rows to 180 mm to increase the density of the planting.

8-1-5-3 Assessment of AfRGM damage

Gall midge damage assessment was estimated according to the RI-BA system, as developed in Chapter 2. The Resistance index (RI) values were computed using the zone value of the resistant check variety to compute the percentage of tillers with galls on the resistant check variety, instead of using its mean value across the whole trial. Zone values are percentage tillers' infestation of each resistant check replicated after every 10 lines of test entries. The mean value of the resistant check variety is its mean infestation across the experimental trial.

$$\text{Resistance Index (RI)} = \frac{\text{tillers of test entry with galls (\%)}}{\text{tillers of resistant check variety with galls (\%)}} \times 100$$

The index is 0 if there are no galls on the test entry, less than 1 if the test entry is more resistant than the resistant check, and more than 1 if the test entry is less resistant than the check (Williams, 1998).

8-1-6 Genotyping

8-1-6-1 Analysis of parental lines for polymorphism study

8-1-6-1-1 Selection of the SSR markers

About 356 SSR markers were chosen based on the following criteria: (1) they express polymorphism according to two referenced populations including Azucena-IR64 and Azucena-TOG; (2) they derive from the Universal Core Genetic Map (UCGM) (Orjuela *et al*, 2009); (3) they belong to the 147 anchors evenly distributed along the rice genome; and (4) in the event that one selected marker from an anchor and polymorphic according to one population is not polymorphic for the other population; another additional marker, polymorphic for this second population, is chosen, giving therefore two markers equivalent candidates for this anchor.

8-1-6-1-2 DNA extraction

Total DNA was extracted using the cetyl trimethyl ammonium bromide (CTAB) method (Murray and Thompson, 1980; Doyle and Dickson, 1987) with minor modifications at various stages, as described below.

Fresh leaf tissue was collected 15 days after sowing from each individual and immediately put in pre-labelled 25 ml test tubes and placed in ice. Following this, 300mg of each sample was ground in liquid nitrogen using a pestle and mortar, and the fine powder was transferred to the same 25 ml tube after being added to 7 ml of CTAB buffer, capped and shaken vigorously before placing it back on ice. The mixture was then incubated at 65^{o}C for one hour in a water bath and shaken intermittently during incubation. The sample was thereafter removed and cooled at room temperature. Seven ml of 24:1 chloroform: isoamyl alcohol was added to each sample; the tube was capped and mixed gently by turning the tube upside down for 5 min, and then centrifuged for 15 min at 4000 rpm. The light green supernatant phase of the centrifuged liquid was carefully transferred into another pre-labelled 25ml tube by pipetting without disturbing the chloroform phase.

Seven ml (i.e., the equivalent of the aqueous phase) of 2-propanol (isopropanol) was added to the supernatant, mixed gently and incubated overnight. The samples were then spun for 10 min at 4000 rpm. The isopropanol was carefully decanted. The whitish pellet suspended on the surface of the tube was then washed in 5 ml of 70% cold ethanol, transferred into a 1.5 ml

Eppendorf tube and drained for 12 hours. The DNA pellet was dissolved in Tris-ethanol (TE) for 24 hours.

8-1-6-1-3 DNA quality and quantity determination

The quality (level of non-degradation) of the DNA was determined by a spectrophotometer (SmartSpec Plus, Bio-rad), based on the ratio of D260 by D280, and confirmed on agarose gel using a 1% gel. The quantity of DNA was only determined by spectrophotometer. A DNA sample with a ratio of OD260/OD280 of 1.8-2.0 is said to be of good quality (Thermo Scientific, 2010). DNA samples were diluted at 1/250 (2µl of DNA dissolved in 480 µl of water) and their concentrations were assessed from their optical density read at the wave-length of 280 nm. For the quality confirmation on agarose, DNA was run on 1% gel (1 g of agarose dissolved in 100ml of TBE 0.5X), stained with 3µl of ethidium bromide, for a total volume of 100 ml. Three µl of diluted DNA, mixed with 5µl of 10X loading dye (40% bromophenol blue plus 40% xylene cyanol FF and 50% glycerol), was loaded into each well of the gel**.** The DNA quality was checked by visualisation of the polymerase chain reaction (PCR) product of each sample on the Alpha Imager gel documentation machine (Alpha Innotech). A DNA sample with little or no smearing is of good quality.

8-1-6-1-4 Genotyping of the parental lines

The susceptible female parent ITA306 and the resistant male parent TOS14519 were screened with the same 356 SSR markers used to check for polymorphisms. Two individuals were rated polymorphic if they showed different levels of bands or a band size difference of 5 bp, from the Alpha Imager gel documentation estimate. PCR was performed, as described by Temnykh *et al*. (2000), with slight modifications. Each reaction mixture consisted of a final volume of 15 µl containing 15 ng of template DNA, 1.3 X buffer (Tris-HCL 100 mM, 500 mM of KCl, Triton X-100 1%) already mixed with $MgCl_2$, 200µM d-NTP, 0.1 µM of each primer, and 0.0250 units of Taq polymerase. The mixture was placed, in plate tubes, which were covered with a thermal PCR seal to avoid evaporation. DNA amplification was performed on G-Storm thermo cycler (Applied Bioscience, USA), programmed for an initial 3 min at 94°C, followed by 30 cycles of 30 seconds at 94°C, 45 seconds from 50 to 61°C according to the optimal annealing temperature for each SSR primer, and 1 min at 72°C, with a final extension of 5 min at 72°C to

fill out any incompletely extended strands of DNA. The amplified products were separated by electrophoresis using a 4% non-denaturing polyacrylamide gel (Sambrook and Rusell, 2001). In the case of no or low amplification, the PCR conditions were optimised by varying the annealing temperature of the primers (from 51 to 57°C) and the magnesium concentration (from 1X to 1.5X) in the PCR reaction. An additional step of 30 min at 72^{o}C was also added to discard the "plus A" effect (Breen *et al*., 1999).

8-1-6-2 Bulked segregant analysis (BSA)

Based on the phenotyping of the F_3 family population, ten individuals from each tail of the F_2 population were selected and their DNA was extracted and bulked to constitute two bulks samples differing in their level of AfRGM resistance. DNA sample was mixed so that each DNA concentration was equivalent to 5 ng.µl^{-1} and equimolar to the nine other samples of the bulk sample, to ensure that the final quantity of DNA in the PCR plate from a 3 µl of sample of the bulk sample would be 15 ng. The two extreme bulks samples and the two parents were genotyped with the polymorphic markers identified previously. Markers found polymorphic between the bulks were used to genotype individually each one of the 10 individuals of the bulk to confirm the resistance or the susceptibility of the selected individuals constituting the bulk.

8-1-6-3 Linkage analysis

The selected markers, identified from the BSA, were used to individually genotype each of the lines of the core collection, which represents the variability of the 649 progenies of the whole population. The susceptible parent band allele was scored as 1, the resistant parent allele band was scored as 2 and the heterozygous allele band was scored as 3. The linkage analysis was performed using MapDisto version 1.7 (Lorieux, 2007). The F-test was used to determine the type of linkage within markers and between markers and phenotype.

8-2 Results

8-2-1 Polymorphism study amongst parental lines

Table 8-1 shows that 145 of the 270 SSR markers used were found to be polymorphic for the varieties ITA306 and TOS14519, with an average of 10 markers per chromosome, with

Chromosomes 10 and 11 being the most polymorphic, amounting to 66.67% and 72.22 %, respectively. The least polymorphic chromosome was Chromosome 1, with a polymorphism level of 39.47%. The mean polymorphism level of the 12 chromosomes was 53.7%. The level of polymorphism ranged from 4.02 base pair difference (bpd) for RM463 located on Anchor 139 on chromosome 12 to 78.01 bpd for RM 2136 located on Anchor 131 on Chromosome 11 (Figure 8-2). The mean bpd was 12.38.

Table 8-1: Polymorphism rate distribution of SSR polymorphic markers between the parental lines, ITA306 and TOS14519, along the twelve rice chromosomes

Chromosomes	Polymorphism rate (%)
Chromosome 1	39.47
Chromosome 2	48.28
Chromosome 3	54.55
Chromosome 4	60
Chromosome 5	47.62
Chromosome 6	54.17
Chromosome 7	61.9
Chromosome 8	57.89
Chromosome 9	58.33
Chromosome 10	66.67
Chromosome 11	72.22
Chromosome 12	45

Type of marker	Number of markers
Monomorphic	125
Polymorphic	145
Total	270
Polymorphism rate (%)	53.7

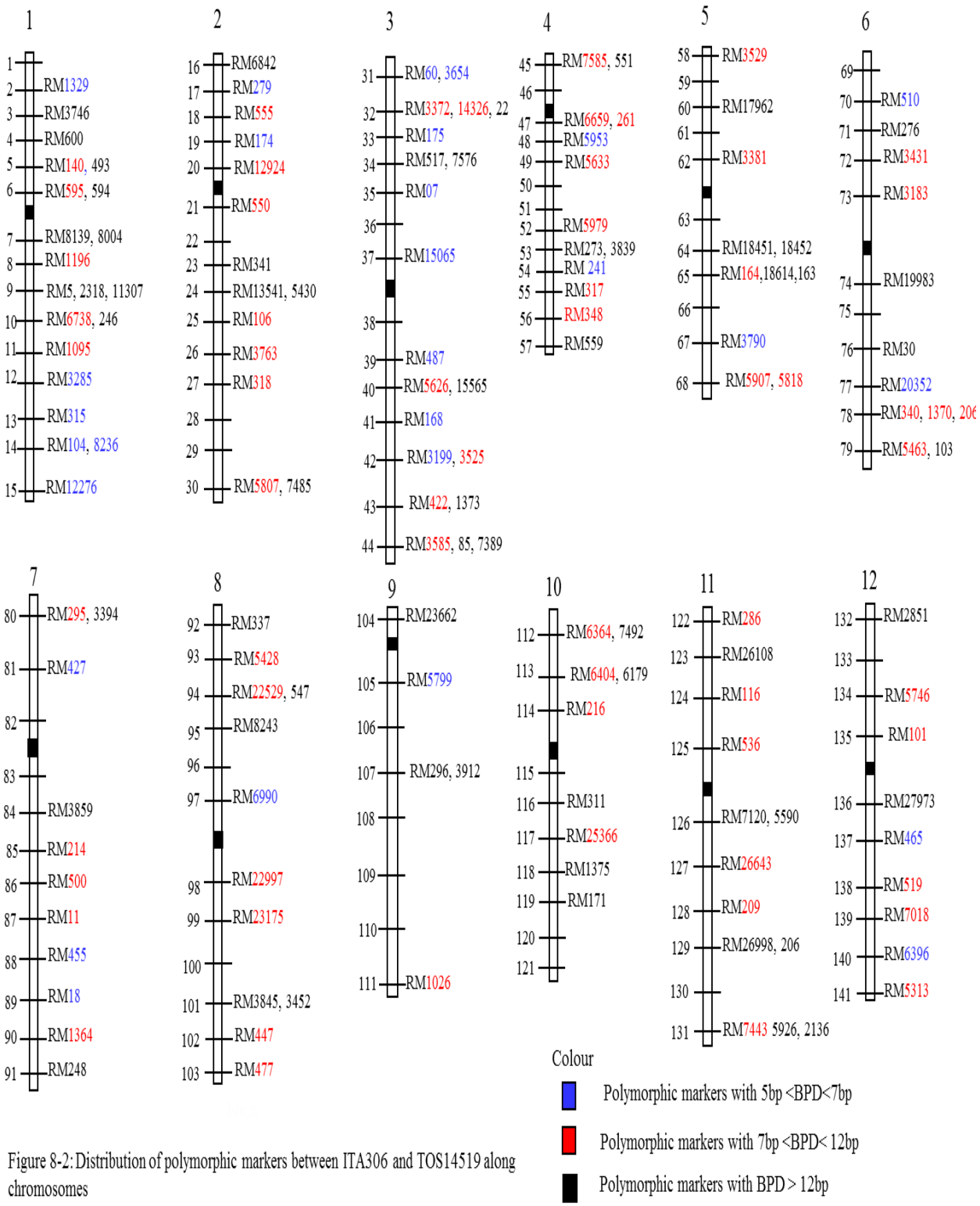


Figure 8-2: Distribution of polymorphic markers between ITA306 and TOS14519 along chromosomes

8-2-2 Phenotyping of the segregating population in a paddy screenhouse

Of the 649 F_3 progenies screened, 38 were resistant, 358 were segregating and 253 were susceptible (Table 8-2). Resistant lines are lines which have at most 5% of the hills infested; segregant lines have 5% to 90% of the hills infested while susceptible lines have above 90% of the hills infested. The 10 most resistant and 10 most susceptible lines were selected. These included 181, 204, 381, 484, 527, 557, 565, 585, 606 and 647 as resistant lines, and 3, 80, 176, 216, 290, 295, 369, 430, 490 and 548 as susceptible lines.

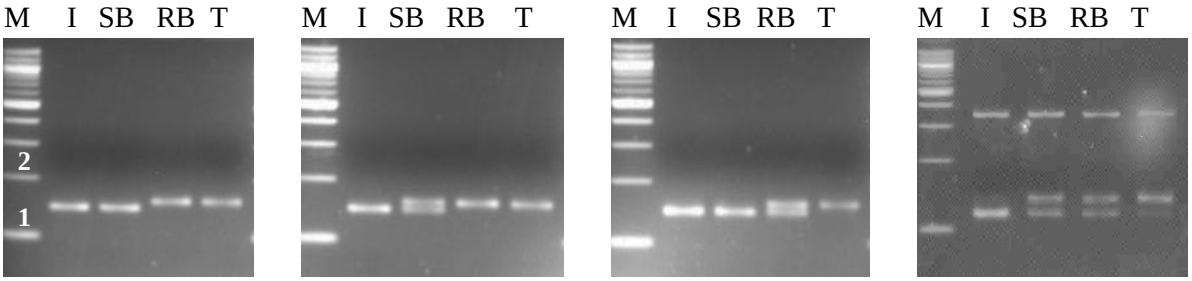
Table 8-2: Segregation pattern of F_2 and F_3 progenies derived from the cross between ITA306 and TOS 14519 in reaction of AfRGM attack

Population	Resistant	Segregant	Susceptible
F_2 ITA306-TOS14519	34	0	445
F_3 ITA306-TOS14519	38	358	253

8-2-3 Genotyping using BSA

8-2-3-1 Bulking segregation analysis

The genotyping of the two bulked samples along with the resistant and susceptible parents generated four main profile patterns: (1) Both resistant bulk and susceptible bulk samples were homozygous and had the band of the resistant and the susceptible parents, respectively (Figure 8-3-A); (2) The susceptible bulk sample was heterozygous, whereas resistant bulk sample was homozygous with the band of the resistant parent (Figure 8-3-B); (3) Both bulks were heterozygous (Figure 8-3-D) and (4) The susceptible bulk sample was homozygous while the resistant bulk sample was heterozygous (Figure 8-3-C). Type 1 and type 2 profiles were the ones where the markers were considered to be linked to the gene/QTL. Based on these profiles, two SSR markers, RM317 and RM17303, were selected as being linked to the resistance of AfRGM amongst the 145 polymorphic markers previously identified according to the polymorphism study amongst the parental lines.



A: RM317 B: RM17303 C: RM17377 D: RM1018

Figure 8-3: Profile patterns displayed by the polymorphic markers during the bulking analysis
M=Molecular weight, I=ITA306, SB=Susceptible bulk, RB=Resistant bulk, T=TOS14519, 1=100bp, 2=200bp

8-2-3-2 Debulking analysis

The use of the two SSR markers to genotype each individual of the bulks, confirmed the linkage of these markers to the resistance gene of AfRGM. Figure 8-4 shows that, apart from Individuals 204 (R_2) and 606 (R_9), all the remaining resistant individuals of the resistant bulk sample only showed the allele of the resistant parent, TOS14519. This trend was observed for both RM317 and RM17303. All the susceptible individuals of the susceptible bulk sample displayed only the allele of the susceptible parent, ITA306, for both markers.

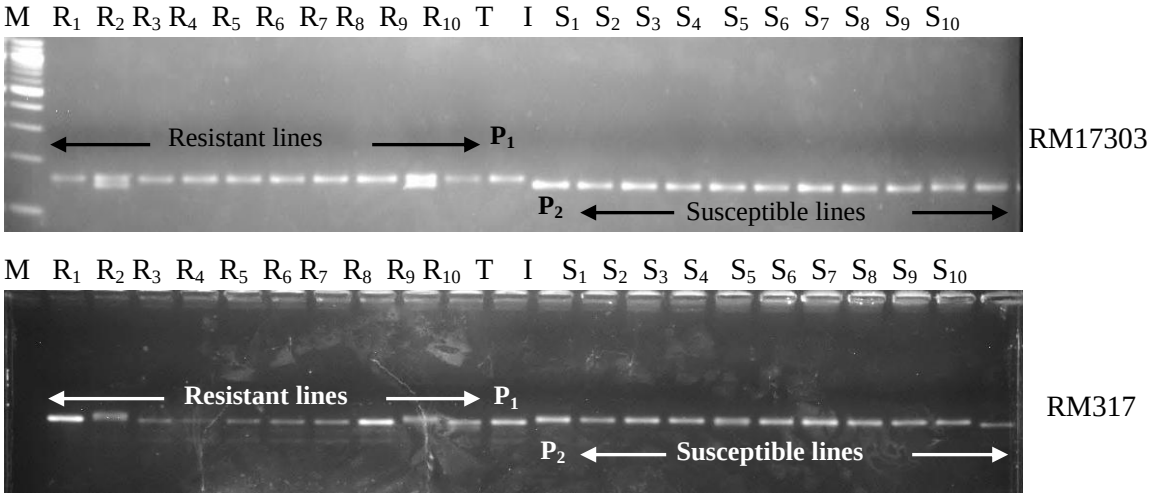


Figure 8-4 : Segregation of RM17303 and RM317 on the lines constituting the bulks
M=molecular weight, R_1=181, R_2=204, R_3=381, R_4=484, R_5=527, R_6=557, R_7=565, R_8=585, R_9=606, R_{10}=647, S_1=3, S_2=80, S_3=176, S_4=216, S_5=290, S_6=295, S_7=369, S_8=430, S_9=490, S_{10}=548

8-2-3-3 Linkage analysis of the whole population

The linkage analysis was performed on the whole population using a co-segregation analysis, based on an F-test. Table 8-3 shows that RM317 was located at 11.02 cM, which is far from RM17303. Both RM317 and RM17303 co-segregate with the phenotype at a very significant level. This means that these markers are tightly linked to the resistance to AfRGM. RM317 co-segregate with the resistance to AfRGM with an F value of 19.06 which is very significant. RM17303 also co-segregate with the phenotype with a value of 15.89 which is also very significant.

Figure 8-5 confirms the link between the two markers and the phenotype by associating a size-landmark to each marker relative to the degree of linkage of the marker to the phenotype. Therefore RM317 was more tightly link to the phenotype than RM17303.

Table 8-3: ANOVA table of all declared loci showing the markers and the phenotype segregation amongst the whole segregating population according to the SSR marker RM317 and RM17303

Loci number	Marker	Sequence	Location (cM)	F-test (Phenotype)	
1	RM317	1	0	19.06	********
2	RM17303	1	11.0241	15.89	******

Phenotype stand for resistance to African rice gall midge

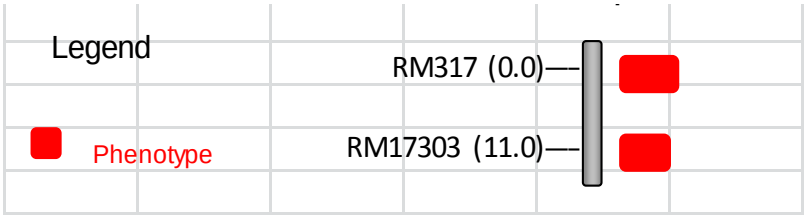


Figure 8-5: Genetic map displaying the link between markers and phenotype

Phenotype stand for resistance to African rice gall midge

8-3 Discussion

The markers selected for the polymorphism analysis in the parental lines were chosen from a reference population. However, the level of polymorphism was lower than that generated in the inter-subspecific cross, in which polymorphism ranged from 59.6% to 90%, as reported by previous studies (Moncada *et al*., 2001; Septiningsih *et al*., 2003; Thomson *et al*., 2003). Levels of less than 54% polymorphism were observed in this study. This could be attributed to the fact that ITA306 and TOS14519, being of *O. sativa* subsp. *indica* and *O. sativa* subsp. *japonica* types, like the reference population Azucena-IR64, are genetically different from the landrace Azucena and the breeding line IR64. ITA306 is a breeding line that has followed a mating scheme of improvement different from that of IR64. Moreover, Azucena and TOS14519 are of different origins, which have obviously contributed to their genetic distance. In addition, more than 80% of the primers designed for *O. sativa* successfully amplify in other AA-genome species (Lorieux *et al.,* 2000). In this present study, 46.3% of the markers selected were monomorphic. The use of a reference population as a prerequisite for choosing markers for a polymorphism study should be therefore oriented toward a specific genetic resemblance between individuals, rather than towards species or sub-species criterion. However, our study identified an average of 10 polymorphic markers per chromosome, which is considered by McCouch *et al*. (2002) and Orjuela *et al*. (2009) as a good genome covering, even though Chromosome 9 was poorly saturated.

BSA is an efficient approach for quick identification of molecular markers linked to target genes. BSA requires allelic difference at the target locus between two DNA pools (Michelmore *et al.*, 1991; Mohan *et al.*, 1994). According to this principle, Wu and Huang (2006) established a special BSA strategy for interactive genes where every pool pair consisted of a homozygous pool and a heterozygous pool. They also noted that BSA does not necessarily require two DNA pools, and that linked markers can also be identified by comparing the homozygous (AA or aa) pool with the parental DNAs. In this scenario, a co-dominant marker such as the SSR used in our study presents both parental bands in the homozygous pool, if it is not linked to the target gene, but presents only one parental band if it is linked. Therefore, constructing the homozygous pool is the key for BSA (Wu and Huang, 2006). Constructing the homozygous pools was difficult in this present study because this implies choosing individuals according to their immunity or their high susceptibility to AfRGM. Many progenies would show the high

susceptibility inherited from ITA306, whereas none would express immunity because TOS14519, from which they inherited their partial resistance, is not immune to AfRGM. Therefore, even progenies selected as resistant to gall midge, because they show the same or more resistance than TOS14519, still bear a degree of susceptibility. Thus, of the four types of profiles observed for the BSA, the marker which was heterozygous for the susceptible bulk sample was retained, because the resistance involved in the resistance of AfRGM was considered to be governed by two recessive genes (see Chapter 3). According to Wu and Huang (2006), susceptible individuals, in such case of interactive genes, present either one band or two bands for the homozygous susceptible individual or the heterozygous susceptible individual, respectively.

Our study also revealed that RM317 and RM17303 which were considered as linked to the resistance after the BSA and the debulking study were confirmed to co-segregate with the gene of interest after the linkage analysis.

In this study, we used the BSA procedure, which is based on pooling individuals of two contrasting tails. These pooling processes were done accurately, with the result that use of the 10 most susceptible and 10 most resistant lines selected for the debulking was enough to effectively reveal all the contrasting allele profiles, which were then used to reject or to select a given marker. This number of pooled individuals is in line with Zhang *et al*. (1994) and Govindaraj *et al.* (2005) who claimed that the use of few individuals (as low as 5-10) in the pool is beneficial in order to limit the appearance of heterozygotes, which hamper the detection of putative markers associated with the trait of interest. However, this low number contradicts the points of view of Michelmore *et al*. (1991), Wang and Patterson (1994); Labombarda *et al*. (2000) and Liu *et al*. (2010), who advocated the use of a higher number of individuals (20-100) in the pools to minimize false positive and false negative results.

The debulking study revealed that few individuals amongst the resistant lines constituting the resistant bulked sample showed recombination instead of homozygosity with the resistant parental line. This phenomenon was observed by Biradar *et al*. (2004) and could be explained, as reported by many authors (White and Fox, 1975; Stahl and Stahl, 1976; Zhao *et al*., 1995; McPeek, 1996, 1999), by a furthest distance separating the markers and the gene of interest, allowing therefore a recombination which is shown by the heterozygosity.

A low level of polymorphism was observed in this study, identifying only two SSR markers linked to the resistance out of the total of 145 previously identified. The explanation could be attributed to the use of BSA which, when applied to QTL identification, only detected the loci with large genetic effects (Grattapaglia, *et al*., 1996; Liu *et al*., 2008; Kanagaraj *et al*., 2010).

The study provided a genetic map of ITA306xTOS14519 cross, comprised of 145 polymorphic SSR markers, which can be used for further mapping studies of other traits. RM317 and RM17303 were found to be linked to the gene/QTL for resistance to AfRGM. This gene/QTL was mapped on the long arm of Chromosome 4 in the vicinity of Anchors 54 and 55 (old anchor) or Anchors 64 and 65 (new anchors) from the Universal Core Genetic Map for rice.

References

Bennett, J., Bentur, J.S., Pasalu, I.C. and Krishnaiah, K., editors. (2004). New approaches to gall midge resistance in rice. In: Proceedings of the International Workshop, 22-24 Nov. 1998, Hyderabad, India. International Rice Research Institute and Indian Council of Agricultural Research, Los Baños, Philippines. 195 p.

Biradar, S.K., Sundaram, R.M., Thirumurugan, T., Bentur, J.S., Amudhan, S., Shenoy, V.V., Mishra, B., Bennet, J. and Sarma, N.P. (2004). Identification of flanking SSR markers for a major rice gall midge resistance gene Gm1 and their validation. *Theoretical and Applied Genetics* 10:1468-1472.

Dakouo, D., Nacro, S. and Sié, M. (1988). Evolution saisonnière des infestations de la cécidomyie du riz, *Orseolia oryzivora* Harris and Gagné (Diptera, Cecidomyiidae) dans le sud-ouest du Burkina Faso. *Insect Science Application* 9:467-472.

Doyle, J.J. and Dickson, E.E. (1987). DNA extraction using CTBA method. *Taxon* 36: 715-722.

Grattapaglia, D., Bertolucci, F.L.G., Penchel, R. and Sederoff, R.R. (1996). Genetic mapping of quantitative trait loci controlling growth and wood quality traits in *Eucalyptus grandis* using maternal half-sib family and RAPD markers. *Genetics* 144:1205-1214.

Govindaraj, P., Arumugachamy, S. and Maheswaran, M. (2005). Bulked segregant analysis to detect main effect QTL associated with grain quality parameters in Basmati 370/ASD 16 cross in rice (*Oryza sativa* L) using SSR markers. *Euphytica* 144:61-68.

Heinrichs, E.A. (1994). Host plant resistance. In: *Biology and Management of Rice Insects.* Heinrichs, E.A. (Ed). Wiley Eastern Limited, New Delhi, India. p 517-547.

Himabindu, K., Sundaram, R.M., Neeraja, C.N., Mishra, B. and Bentur, J.S. (2007). Flanking SSR markers for allelism test for the Asian rice gall midge (*Orseolia oryzae*) resistance genes. *Euphytica* 157:267-279.

Himabindu, K., Suneetha, K., Sama, V.S.A.K. and Bentur, J.S. (2010). A new rice gall midge resistance gene in the breeding line CR57-MR1523, mapping with flanking markers and development of NILs. *Euphytica* 174:179-187.

Kanagaraj, P., Silvas Jebakumar, K., Annie Sheeba, J., Biji, K.R., Sheetal Babu, P., Senthil, A. and Chandra Babu, R. (2010). Microsatellite markers linked to drought resistance in rice (*Oryza sativa* L.). *Current Science* 98:836-839.

Katiyar, S.K., Chandel, G., Tan, Y., Zhang, Y., Huang, B., Nugaliyadde, L., Fernando, K., Bentur, J.S., Inthavong, S., Constantino, S. and Bennett, J. (1995). Biodiversity of Asian rice gall midge (*Orseolia oryzae* Wood Mason) from five countries examined by AFLP analysis. *Genome* 43:322-332.

Kumar, A., Jain, A., Sahu, R.K., Shrivastava, M.N., Nair, S. and Mohan. M. (2005). Genetic analysis of resistance genes for the rice gall midge in two rice genotypes. *Crop Science* 45:1631-1635.

Labombarda, P., Pupilli, F. and Arcioni, S. (2000). Optimal population size for RFLP-assisted cultivar identification in alfalfa (*Medicago sativa* L.). *Agronomie* 20:233-240.

Liu, G., Bernhardt, L., Jia, H., Wamishe, A. and Jia, Y. (2008). Molecular characterization of the recombinant inbred line population derived from a *japonica-indica* rice cross. *Euphytica* 159:73-82.

Liu, S., Chen, H.D., Makarevitch, I., Shirmer, R., Emrich, S.J., Dietrich, C.R., Barbazuk, W.B., Springer, N.M. and Schnable, P.S. (2010). High-throughput genetic mapping of mutants via quantitative single nucleotide polymorphism typing. *Genetics* 184:19-26.

Lorieux, M., Ndjiondjop, M-N. and Ghesquière, A. (2000). A first interspecific *Oryza sativa-Oryza glaberrima* microsatellite-based genetic linkage map. *Theoretical and Applied Genetics* 100:593-601.

Lorieux, M. (2007). MapDisto, a free user-friendly program for computing genetic maps. Computer demonstration given at the Plant and Animal Genome XV conference, Jan 13-17 2007, San Diego, CA. URL: http://mapdisto.free.fr/. Accessed December, 2010.

Maji, A.J. (1998). Genetics of resistance to rice gall midge (*Orseolia* spp.) Paper presented during the WARDA/NARS group discussion and planning meeting on bioecology and management of African rice gall midge, 12-16 October at IITA, Ibadan, Nigeria.

McPeek, M.S. (1996). An introduction to recombination and linkage analysis. In: *Genetic Mapping and DNA Sequencing.* Speed, T.P. and Waterman, M.S. (Ed). Springer-Verlag, New York, USA. p 1-14.

McPeek, M.S. (1999). Optimal allele-sharing statistics for genetic mapping using affected relatives. *Genetic Epidemiology* 16:225-249.

McCouch, S.R., Teytelman, L., Xu, Y., Lobos, K.B., Clare, K., Walton, M., Fu, B., Maghirang, R., Li, Z., Xing, Y., Zhang, Q., Kono, I., Yano M., Fjellstrom, R., DeClerck, G., Schneider, D., Cartinhour, S., Ware, D. and Stein, L. (2002) Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). *DNA Research* 9:257-279.

Michelmore, R.W., Paran, I. and Kesseli, R.V. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proceeding of the National Academy of Science, USA* 88:9828-9832.

Mohan, M., Nair, S., Bentur, J., Rao, U. and Bennett, J. (1994). RFLP and RAPD mapping of the rice *Gm2* gene that confers resistance to biotype 1 of Gall midge (*Orseolia Oryzae*). *Theoretical and Applied Genetics* 87:782-788.

Moncada, P., Martinez, C.P., Borrero, J., Chatel, M., Gauch, H. Jr., Guimaraes, E., Tohme, J. and McCouch, S.R. (2001). Quantitative trait loci for yield and yield components in an *Oryza sativa*×*Oryza rufipogon* BC_2F_2 population evaluated in an upland environment. *Theoretical and Applied Genetics* 102**:** 41-52

Murray, M.G. and Thompson, W.F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* 8:4321-4325.

Nwilene, F.E., Williams, C.T., Ukwungwu, M.N., Dakouo, D., Nacro, S., Hamadoun, A., Kamara, S.I., Okhidievbie, O., Abamu, F.J. and Adam, A. (2002). Reactions of

differential rice genotypes to African rice gall midge in West Africa. *International Journal of Pest Management* 48:195-201.

Nwilene, F.E., Harris, K.M., Okhidievbie, O., Onasanya, A., Sere, Y. and Ingelbrecht, I. (2006). Morphological diversity and genomic DNA fingerprinting of the African rice gall midge *Orseolia oryzivora* (Diptera: Cecidomyiidae) and of two other species of African *Orseolia*. *International Journal of Tropical Insect Science* 26:256-265.

Orjuela, J., Garavito, A., Bouniol, M., Arbelaez, J.D., Moreno, L., Kimball, J., Wilson, G., Rami, J.F., Tohme, J., McCouch, S.R. and Lorieux, M. (2009). A universal core genetic map for rice. *Theoretical and Applied Genetics* 120:563-572.

Sambrook, J., and Rusell, D.W. (2001). *Molecular Cloning a Laboratory Manual*. Cold Spring Harbor: NY, Cold Spring Harbor Laboratory Press 2001. New York, USA.

Septiningsih, E.M., Prasetiyono, J., Lubis, E., Tai, T.H., Tjubaryat, T., Moeljopawiro, S. and McCouch, S.R. (2003). Identification of quantitative trait loci for yield and yield components in an advanced backcross population derived from the *Oryza sativa* variety IR64 and the wild relative *O. rufipogon*. *Theoretical and Applied Genetics* 107:1419-1432.

Stahl, F.W and Stahl, M.M. (1976). On recombination between close and distant markers in phage lambda. *Genetics* 82:577-593.

Temnykh, S.W., Park, D., Ayres, N., Cartingour, S., Hanck, N., Lipovich, L., Cho, Y.S. and Tshil, T. (2000). Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics* 100: 697-712.

Thermo Scientific (2010). *T042-Technical Bulletin NanoDrop Spectrophotometers, 260/280 and 260/230 Ratios*. Available at: http://www.nanodrop.com/Library/T042-NanoDrop-Spectrophotometers-Nucleic-Acid-Purity-Ratios.pdf Retrieved 19 March 2012).

Thomson, M.J., Tai, T.H., McClung, A.M., Lai, X.H., Hinga, M.E., Lobo, K.B., Xu, Y., Martínez, R. and McCouch, S.R. (2003). Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between *Oryza rufipogon* and the *Oryza sativa* cultivar Jefferson. *Theoretical and Applied Genetics* 107:479-493.

Ukwungwu, M.N., Winslow, N.D. and John, V.T. (1989). Severe outbreak of rice gall midge (GM) in the Savannah zone of Nigeria. *International Rice Research Newsletter* 14:36-37.

Ukwungwu, M.N. and Joshi R.C. (1992). Distribution of the African rice gall midge, *Orseolia oryzivora* Harris and Gagné and its parasitoids in Nigeria. *Tropical Pest Management* 38:241-244.

Williams, C.T. (1998). Crop loss assessment methods for African rice gall midge, *Orseolia oryzivora* Harris and Gagné. Paper presented at WARDA/NARS Group discussion and planning meeting: Bio-ecology and management of the African rice gall midge, October 12-16, 1998, IITA, Ibadan, Nigeria.

Wang, G.L. and Paterson, A.H. (1994). Assessment of DNA pooling strategies for mapping of QTLs. *Theoretical and Applied Genetics* 88:355-361.

White, R.L. and Fox, M.S. (1975). Genetic heterozygosity in unreplicated bacteriophage alpha recombinants. *Genetics* 81:33-50.

Wu, W. and Huang, B. (2006). Strategy for the mapping of interactive genes using bulked segregant analysis method and Mapmaker/Exp software. *Chinese Science Bulletin* 51: 2619-2623.

Zhang, Q., Shen, B.Z., Dai, X.K., Mei, M.H., Saghai Maroof, M.A. and Li, Z.B. (1994). Using bulked extremes and recessive class to map genes for photoperiod-sensitive genetic male sterility in rice. *Proceeding of the National Academy of Science, USA* 91:8675-8679.

Zhao, H., Speed, T.P., and McPeek, M.S. (1995). A statistical analysis of crossing over interference using the Chi-square model. *Genetics* 139:1045-1056.

Thesis Overview

A Genetic study for resistance to African rice gall midge (AfRGM) in West African rice cultivars

Introduction

Developing new varieties to tackle hunger is a common approach adopted by breeders and the scientific community involved in poverty alleviation. Marker-assisted breeding (MAB) has become an established way of breeding despite its various requirements. This study was therefore initiated to map and tag the African rice gall midge (AfRGM) resistance gene of selected rice cultivars for its use in MAB. Gene mapping implies: (1) Having good parental lines contrasting for the trait of interest; (2) Developing a good mapping population by crossing accurately these parents; (3) Using a good phenotyping method to screen parents and progenies; (4) Having a good genotyping method by using an appropriate mapping system in relation with the type of population to genotype. Moreover, using a large population size is one of the additional limitations to gene mapping. Therefore, substituting this population with a narrow type population bearing the entire variability of the whole population would simplify the process and thus save substantially time as compared to time spent on a full Bulk Segregant Analysis (BSA). These requirements were considered and were investigated in this present study, with a focus on breeding resistance to AfRGM.

Findings from the study

Phenotyping relies either on the method of screening or less frequently on the assessment of the damage caused to plants. When the damage is due to an insect pest attack, as with AfRGM, this requirement is clear. The best screening method developed (RI-BA2) in this study more efficiently and accurately assesses AfRGM damage better than the International Rice Research Institute (IRRI)'s standard evaluation system (SES) method and two other resistance index-based assessment (RI-BA) methods. RI-BA2 takes into account the impact of the environment and measures the tested entries' reaction relative to the check. RI-BA2 provides a more sensitive measure of resistance which, unlike RI-BA1, RI-BA3 and the SES

score, takes into account variations in AfRGM infestation levels in the plants' vicinity. The study showed the similar accuracy of RI-BA2 to the other methods when the plot size was small or when the resistant check is an immune variety. This means that RI-BA2 is most useful when the plots are large and when the reaction of the resistant check entry to AfRGM is quantitative and varies across the trial. This method works well because AfRGM attack varies much across most trials, but by estimating AfRGM damage relative to the check in the immediate vicinity of the progenies, much more precise evaluation can be made. This method was developed using F_3 segregating families but could also be applied to individual plants in F_2 and F_1 generations. In these cases, AfRGM damage on a single progeny would be assessed relative to the resistant check grown to the same column of plants.

The RI-BA2 method was used to determine the mode of action of the AfRGM resistance gene, and the heritability estimates associated with AfRGM resistance. The study detected the existence of two recessive genes involved in the resistance to AfRGM. In addition, the heritability estimates showed that this resistance to AfRGM is highly heritable, and therefore resistance to AfRGM can be introgressed from TOS14519 into other promising lines. High heritability estimates confirmed that in addition to over-dominance, additive effects also played a role in influencing the resistance to AfRGM in rice genotypes. These could be exploited in the development of AfRGM resistant rice. The knowledge of the genetics of the resistance gene could help in choosing the appropriate mating design to breed for resistance to AfRGM. This finding is also important when attempting to map the relevant genes, orienting towards appropriate mapping methods. In the present case of two recessive genes, the mapping methodology would be directed toward mapping interactive genes rather than a single gene mapping. Therefore, the phenotype, the genotype and the profile observed in a BSA are expected to be as described by Wu and Huang (2006) in the case of two recessive genes segregating in an F_2 mapping population.

The behavioural pattern or resistance response of progenies derived from various crosses involving ITA306 and the four parental lines with varying levels of resistance to AfRGM, clustered the progenies' reaction into four behavioural classes. Three were commonly observed in plant response to insect attack. Notably plant in one class converted from being susceptible progenies at 45 days after transplanting (DAT) into resistant progenies at 70 DAT. Usually, susceptible progenies at 45 DAT are rejected because they are considered fully

susceptible. The finding from this study shows that this approach has to be reconsidered. The understanding of how these plants convert from susceptible to resistant may help to unravel AfRGM resistance expression. The unusually high number of progenies reverting from susceptibility to resistance for progenies of ITA306 x BW348-1, may be the result of the two following options: (1) less infestation of progenies at 70 DAT than they were at 45 DAT, associated with the highest infestation of check entry at 70 DAT. This would have led to a reduction of the resistance index (RI) and therefore implies more resistance than the check entry; (2) at a similar level of infestation of check entry at 45 and 70 DAT, the high number of reverted progenies was due to the tolerance aptitude of the check entry (BW348-1) which supports high infestation, because of its tolerance to gall midge. At 70 DAT, the level of infestation was higher than at 45 DAT due to longer exposure to the pest. Therefore, the percentage mean infestation (PMI) of the check entry (BW348-1) increased and contributed to reduce the RI value and subsequently the resistance appears to increase. The same explanation is consistent for susceptible progenies at 45 DAT, despite their higher infestation at 70 DAT than at 45 DAT, were scored as resistant at 70 DAT.

The high heritability estimates of the fifteen agro-morphological traits suggested that the variability observed within the segregating populations was to a large extent due to genetic effects rather than to environmental effects and suggested, that the genetic gain associated with each trait can be transmitted along with the resistance to AfRGM. Moreover, the high broad-sense heritability estimates and the contrasting low narrow-sense heritability estimates showed that inheritance of the traits studied was under dominance and epistatic gene effects more than additive genetic effects and can therefore be improved using pedigree breeding schemes.

Genetic diversity amongst the five AfRGM parental lines was studied in terms of polymorphism using 303 SSR primers covering the rice genome. Factorial analysis based on the allelic diversity clearly demarcated the parental lines into *Oryza glaberrima*, *Oryza sativa* subsp. *japonica* and *Oryza sativa* subsp. *indica* groups, while a cluster analysis distinguished them into AfRGM resistant, susceptible, moderately resistant and tolerant groups. BW348-1 and Cisadane showed the least allelic diversity despite the diverse geographical sources of their collection. TOS14519 and TOG7106 showed strong divergence from ITA306 despite their common West African origin. This allelic diversity among the genotypes tested must be due to

natural evolution, or farmer based mass selection for AfRGM resistance, rather than from direct breeding efforts by plant breeders.

A novel method of selecting individuals of a mapping population based on a core sample, which represents as much as possible the entire variability of the progenies, was developed. The core sample was similar to the whole population for allelic distribution, phenotypic variation, clustering pattern, minimum and maximum quantitative values and diversity index, while mean values and coefficient of variation distinguished them. The core sample, which represents a 10% sample of the whole population, can therefore efficiently represent it as a mapping population and save cost of maintenance.

Although the linkage analysis study did not confirm the co-segregation of any SSR marker with the gene of interest, 145 polymorphic markers between two contrasting parental lines, with strong and weak AfRGM resistance, were identified and can serve as a prerequisite for further mapping studies. The wide distribution of these markers along the twelve rice chromosomes within anchors indicates that the use of reference population as a prerequisite for choosing markers for a polymorphism study should be oriented toward a specific genetic resemblance between individuals rather than toward species or sub-species criteria. In addition, the assessment of the level of polymorphism by the allele band size provides useful information for multiplexing electrophoresis.

Conclusion and recommendations

The most accurate assessment method of AfRGM damage is RI-BA2. It is most useful when the resistant check is not immune, plot size are large and when AfRGM attacks are unevenly distributed. It can also be applied to a single plant stand with the respect to avoid gaps within the check lines. However, this method is space consuming if one wants to increase the accuracy by replicating each check very often, in order to reduce the distance between the check entry and the test entries. The segregation pattern of F_1, F_2 and F_3 generations revealed that two recessive genes are the basis for resistance to AfRGM in crosses derived from TOS14519. However, the difference or the similarity of these genes involved in the two crosses was not clear. Therefore, a further allelic test involving the two parental lines could help in deciding on the identity of these genes. Similarly, general combining ability (GCA) and specific combining ability (SCA) should be determined to clarify whether or not the transmission of AfRGM

resistance to progenies is favoured in specific genotype associations or if the transmission is not influenced by the parents of the cross.

To clearly elucidate the mechanism involved in the transformation of susceptible germplasm into resistant, additional studies on the epidemiology of infestation of progenies and a tolerant variety should be done under a strict experimental condition including an individual plant infestation by larvae instead of random infestation by adult flies. Moreover, the choice of the tested entries should be oriented toward fixed lines to avoid fluctuation of the behavioural pattern of the progenies due to the variability of the segregating population.

Our study revealed that the 15 quantitative agro-morphological traits can be transmitted along with AfRGM resistance. However, the correlation between these traits and AfRGM resistance was not determined. Therefore, an additional correlation study involving these traits and resistance to AfRGM should be addressed to see whether they are negatively correlated.

The five parental lines were clustered either according to their genetic origin or according to their reaction against AfRGM. The use of more *O. glaberrima, O. sativa* subsp. *indica* and *O. sativa* subsp. *japonica* lines on one hand and more resistant, susceptible and tolerant lines on the other hand, could help to confirm our findings from this study.

In a view of enhancing breeding for AfRGM using MAB, the future prospect should be oriented toward the fine mapping of the mapped genes. More sensitive markers in terms of polymorphism detection, such as SNP markers will be investigated. In addition, to accurately constitute the resistant homozygous pool without any bias in view of BSA, the choice of the resistant parental line to develop the mapping population should be oriented toward immune lines, or highly resistant lines.

References

Wu, W. and Huang, B. (2006). Strategy for the mapping of interactive genes using bulked segregant analysis method and Mapmaker/Exp software. *Chinese Science Bulletin* 51:2619-2623.