Empowering breeding programs with new approaches to overcome constraints for selecting superior quality traits of rice

Mariafe N. Calingacion

Empowering breeding programs with new approaches to overcome constraints for selecting superior quality traits of rice

Mariafe N. Calingacion

Empowering breeding programs with new approaches to overcome constraints for selecting superior quality traits of rice

Mariafe N. Calingacion

Thesis committee

Promoters

Prof. Dr. R. D. Hall
Professor of Plant Metabolomics
Wageningen University

Prof. Dr. M. A. Fitzgerald
Professor of Agriculture and Food Sciences
University of Queensland

Co-promoter

Dr. R. Mumm
Researcher and Project Leader, Plant Research International
Wageningen University and Research Centre

Other members

Prof. P. Struik, Wageningen University
Prof. A. Dobermann, Rothamsted Research, UK
Prof E. Woltering, Wageningen University
Dr Y. Tikunov, Plant Research International, Wageningen University and Research Centre

This research was conducted under the auspices of the Graduate School of Experimental Plant Sciences.

Empowering breeding programs with new approaches to overcome constraints for selecting superior quality traits of rice

Mariafe N. Calingacion

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 9th January 2015
at 4 p.m. in the Aula.

Mariafe N. Calingacion
Empowering breeding programs with new approaches to overcome constraints for selecting superior quality traits of rice.
198 pages.

PhD thesis, Wageningen University, Wageningen, NL (2015)
With references and propositions, with summaries in Dutch and English

ISBN 978-94-6257-218-8

The research for this PhD dissertation has been funded by Monsanto Beachell-Borlaug International Scholarship Program, International Rice Research Institute and University of Queensland.

Financial support of Wageningen University for printing this thesis is gratefully acknowledged.

Table of Contents Page

Chapter 1

General Introduction

Rice is the staple food for nearly half of the world's population. It belongs to genus *Oryza* comprising of 24 species distributed across the world but only two species are cultivated namely *Oryza sativa* L. and *Oryza glaberrima* S. (Vaughan, Morishima et al. 2003). *O. sativa* was domesticated in Asia, but is now cultivated around the world, whereas *O. glaberrima* has been cultivated exclusively in Africa (Rosell and Marco 2008). Because of its long history of cultivation and selection under diverse environments, remarkable diversity exists within the rice crop (Childs 2004). The largest collection of rice germplasm is found at the International Rice Research Institute, where more than 100,000 rice accessions are held in the International Rice Genebank (McNally, Bruskiewich et al. 2006).

Rice accounts for 35–60% of the calories consumed by 3 billion people in Asia (Brar, Khush et al. 2013). Rice is also a cheap source of protein, and in Asia, rice supplies 20% of the dietary protein intake (Rosell and Marco 2008). The majority of rice consumed globally is white rice which is categorised as being completely milled rice and/or parboiled (placed under steam pressure before milling). Brown rice i.e. unmilled grains, previously accounted for a much smaller share in the market. However, brown rice is now gaining popularity because of its nutritional benefits (Childs 2004). Rice is also consumed in the form of noodles, puffed rice, fermented sweet rice, breakfast cereals, candy, and crackers (Wilkinson and Champagne 2004). It is also used in making beer, rice wine, sake and vinegar (Yoshizawa and Ogawa 2004).

Rice is grown on 159 million ha across the world, with annual production reaching 690 million tonnes (Brar, Khush et al. 2013). Although rice consumption per capita in many Asian countries is decreasing due to changing dietary habits, mainly as a result of western influence, involving for example, increased consumption of dairy, meat and fast foods, as well as general economic development (Pingali 2007), rice consumption still exceeds 100 kg per capita annually (Seck, Diagne et al. 2012). Western countries have also started eating rice more regularly and the continued growth of Asian communities within western countries including America and Europe has significantly increased the overall market size for rice as well as creating growing awareness and appreciation of rice-based meals (Suwannaporn and Linnemann 2008). Furthermore, the demand for rice is rapidly growing in Sub-Saharan Africa due to population growth, and a shift in consumer preference for rice, especially in urban areas (Balasubramanian, Sie et al. 2007). However, despite efforts to develop higher yielding varieties by rice improvement programs, current projections of population growth and proportionally more consumers eating rice every day in the world

indicate that an additional 8 million tonnes of rice will need to be produced each year in Asia (Mohanty 2013).

Yield

During the 'Green Revolution', the adoption of semi-dwarf, high-yielding varieties coupled with efficient production technologies was able to maintain rice production ahead of population growth (Zeigler and Barclay 2008). Rice production has increased by 130% from 257 million tons in 1966 to 600 million tons in 2000 (Khush 2005). However, the total population of rice-consuming countries continues to grow. But the challenges is even greater because this increased demand will also have to be met (i) by using fewer hectares - as more agricultural land is being lost to urbanisation, (ii) by using less labour due to significant migration of workers from farms as the younger generation moves to more rural areas, and (iii) following cultivation under sub-optimal conditions mainly due to the impact of climate change in the most important rice-producing areas. To meet this challenge we require more / novel rice varieties with higher yield potential and greater yield stability. Currently, various strategies for increasing rice yield potential are being developed and employed in research and breeding programs including: (1) conventional hybridization and selection procedures (Khush 2005), (2) new ideotype breeding (Peng, Khush et al. 2008), (3) hybrid breeding (Xie and Hardy 2009), (4) diversifying selection strategies (McCouch 2004) and (5) re-engineering rice photosynthesis (Sheehy and Mitchell 2013).

To date, progress in yield improvement of crops has been accomplished through conventional hybridization and selection procedures (Khush 2005). For rice, since the release of the first high-yielding modern rice variety in 1966, an annual 1% increase in rice yield has been obtained, i.e. 75-81 kg ha^{-1} (Peng, Laza et al. 2000). This has mainly been due to a combination of improved harvest index and increased total biomass.

The new ideotype approach has been used in breeding programs at the International Rice Research Institute (IRRI) and in China to improve rice yield potential (Khush 2005). Crop ideotype is an idealized plant type with model characteristics favourable for photosynthesis, growth, and grain production (Donald 1968). The new plant type (NPT) for rice was proposed to have a growth duration of 100-130 days, with few unproductive tillers, 200-250 grains per panicle, 90 –100 cm plant height, sturdy and thick stems, thick, dark green, and erect leaves, a vigorous root system, and an increased harvest index (Peng, Khush et al. 2008). Although the first generation NPT lines did not reach the expected yield due to low

biomass production and poor grain-filling, yield of second generation NPT lines improved significantly due to increased panicle number and improved grain filling (Laza, Peng et al. 2003). China has also used an ideotype breeding approach and launched the 'super rice mega project' which focused on developing super hybrid rice (Cheng, Cao et al. 2007). Several NPT lines have also been released in Indonesia, and other National Agricultural Research Systems are still evaluating and further improving their NPT lines (Khush 2005). However, up until now, NPT varieties have not produced a significant yield benefit.

China has made significant progress in developing hybrid rice technology based on heterosis (Wang, Xue et al. 2005). Heterosis, or hybrid vigor, refers to the greater performance of a hybrid in relation to its parents (Zhou, Chen et al. 2012). In 2006, China released a total of 34 super hybrids that were planted in a total area of 13.5 million ha and produced 6.7 billion kg more rice annually since 1998 (Cheng, Cao et al. 2007). However, most super hybrid rices require very high input of both chemical fertilizers and pesticides. As well as undesirably higher production costs, this can result in serious environmental pollution and hazards to farmer and consumer health (Zhang 2007). Consequently, the 'green super rice program' has since been conceived within which new rice varieties are currently being developed that not only give good yield and quality, but also are more resistant to drought and multiple insect pests and microbial pathogens. These new varieties should also have high nutrient efficiency, require less pesticides, chemical fertilizers, and water, for delivery to resource-poor farmers particularly in sub-Saharan Africa and Asia (http://thegsr.org). However, some challenges still remain as there is still a lot of work to be done to address for example, poor milling, eating, and cooking qualities of hybrid rice (Tan, Xing et al. 2000).

To widen gene pools in breeding programs, crosses between existing crop cultivars with wild species have frequently been exploited for improving many agronomic traits including yield potential (Khush 2005). In such crosses, alleles that were not selected / retained during the domestication process may be restored into the cultivated gene pool (McCouch 2004). Backcross derivatives, for example, from a cross between an *Oryza rufipogon* accession and cultivated rice delivered lines giving 18% more yield than the recurrent parent (Xiao, Grandillo et al. 1996). Improved yield and yield components such as increased panicle length, number of panicles per plant, percentage seed set, grains per plant, and grain weight, were also observed in a backcross population developed from a cross between *Oryza sativa* (cv IR64) and *O. rufipogon* (IRGC 105491) (Septiningsih, Prasetiyono

et al. 2003). Transgressive segregation for traits related to yield from crosses of cultivated and wild species, indeed suggests that wild rice contains genes that can improve quantitative traits including yield (McCouch 2004).

In addition to the more conventional breeding strategies as outlined above, scientists are also developing alternative approaches. For example, it is aimed to achieve a 30-50% increase in yield potential in rice through harnessing the C4 photosynthetic pathway, as found in crops such as maize, sorghum and sugar cane (Zeigler and Barclay 2008). Therefore, the central goal of the International C4 Rice Consortium is to identify and engineer the genes that are important for incorporating C4 photosynthesis into rice (www.c4.irri.org). Modelling and empirical analysis have shown that maximum crop yield is influenced by the level and proportion of light which the crop captures during its growth, its radiation use efficiency, and the harvest index (Sheehy and Mitchell 2013). Currently, a subset of genes needed for metabolite transporters that are involved in the C4 pathway have been engineered and coupled to promoters to give cell-specific expression (Kajala, Covshoff et al. 2011). However, while promising, it is estimated that another 15 years of research will still be needed to optimise the phenotype and conduct field testing of C4 rice before it becomes ready for release (von Caemmerer, Quick et al. 2012).

It is true that the intrinsic (genetic) characteristics of the rice plant principally affect yield. However, certain crop management practices and technologies have proven highly effective in improving rice productivity. Alternate wetting and drying (AWD) irrigation strategies has been shown to increase yield by 4–6% compared to plants grown in continuous submergence (Belder, Spiertz et al. 2005). Moreover, AWD, when combined with site-specific nitrogen management (SSNM) could also increase grain yield by 14% compared to the yield of those plants that were grown under SSNM or AWD alone (Liu, Chen et al. 2013). However, to narrow the yield gap between research and farmers' field, we still need a better understanding of the cultural, social, and economic factors that drive adoption and technological adaptation (Zeigler and Barclay 2008).

As indicated above, several approaches in breeding programs have already proven efficient in improving rice yield potential and stability. However, yield is just one part of the equation for successful varietal adoption by both farmers and consumers alike. Without consumer acceptance, new varieties have little chance to succeed. Such preference is largely driven by desires for premium grain quality and an acceptable price / quality parameter (Fitzgerald,

McCouch et al. 2009). Attention to quality aspects has however often been overseen or indeed, ignored.

Grain Quality

Many countries still prefer to grow rice varieties that were released several decades ago. Examples are varieties such as Khao Dawk Mali 105, Swarna and IR64 which were selected in 1958, 1979 and 1985, respectively (Fitzgerald, McCouch et al. 2009). One reason why farmers have not widely replaced these traditional varieties with more recently bred varieties is that consumers have often not accepted the quality of the grains of these (higher or reliable yielding) varieties (Boualaphanh, Calingacion et al. 2011). Consumer acceptance of a variety of rice is driven initially by the physical appearance of the grains (Fitzgerald, McCouch et al. 2009) and is then further influenced by cooking and sensory quality attributes of the rice (Cramer, Wailes et al. 1993). More recently, nutritional value has also been influential. Physical properties include yield of edible and marketable polished grain, kernel length, uniform shape as well as translucence and a lack of chalkiness (Fitzgerald, McCouch et al. 2009). These traits are visible and immediately obvious to consumers and consequently are major factors defining market value. Sensory qualities typically include the aroma, flavour, taste and the texture of the rice in the mouth, often referred to as 'mouth-feel' (Champagne, Bett et al. 1999). Cooking qualities, on the other hand, are largely influenced by the starch content which makes up to 90% weight of milled rice (Graham 2002). Starch properties are measured in terms of gelatinisation temperature, amylose content and gel consistency. Another important trait related to cooking quality is the aroma emitted during rice cooking and which fills the kitchen.

The main physical quality traits of the rice grain are immediately recognisable even to an untrained eye – these are grain length, width and shape (ratio length / width), translucence and chalkiness (Graham 2002). These are crucial aspects of grain quality not only for consumers but also for millers, wholesalers and retailers (Fitzgerald, McCouch et al. 2009). Grain dimensions and appearance therefore also play a pivotal role in selecting varieties that are to be developed for commercial production (Juliano 2003). The classes of grain length of milled rice are defined as short (<5.5 mm), medium (5.51–6.6 mm), long (6.61–7.5 mm) and extra-long (>7.51 mm). Grain shape, the ratio of length and width, is classified as bold (<2), medium (2.1 – 3), and slender (>3). Grain dimensions are usually measured using a caliper, calibrated scanner or a kernel digital analysis equipment.

Chalk, an opaque region in the grain, affects the visual appearance of white rice (Lisle, Martin et al. 2000). Chalk may also influence the tendency of a grain to break during polishing (Swamy and Bhattacharya 1982), thus decreasing the final yield of edible rice. Thus, chalkiness lowers the value of rice in the market and indeed, most markets will not even accept rice that contains >2% chalky grains (Lisle, Martin et al. 2000). Chalkiness in rice is caused by several factors and occurs when rice is harvested at very high moisture levels (Koutroubas, Mazzini et al. 2004). A closer look at the chalky areas of the rice grain reveals these comprise loosely-packed starch granules, whereas more tightly packed granules are to be found in the translucent areas (Chun, Song et al. 2009). It has been proposed that the processes involved in the packing of the starch granules are susceptible to stress in varieties which are susceptible to developing chalky grains (Lisle, Martin et al. 2000). Therefore, the identification of pathways which lead to chalk formation in rice is a major focus of rice quality research as this influences not only eating quality but also overall, final yield.

There are a number of measurable traits that indicate cooking quality of rice. Gelatinisation temperature (GT) is the cooking temperature at which water is absorbed by the rice grain and coincides with the crystalline structures of the starch beginning to melt (Graham 2002). GT is used in rice breeding programs as an indicator of the cooking time of rice samples (Cuevas, Daygon et al. 2010). The GT of rice varieties may be classified as low (55 to 69 °C), intermediate (70 to 74 °C), and high (>74°C) (Juliano 2003). GT is commonly determined by the alkali spreading value (ASV) (Little, Hilder et al. 1958), differential scanning calorimetry (DSC) (Cuevas, Daygon et al. 2010), deducting 3°C from the pasting temperature derived from the rapid visco analyser (RVA) (Juliano 2003) or by nuclear magnetic resonance (NMR) analysis method (Ritota, Gianferri et al. 2008). ASV measures the degree of disintegration of milled rice in 1.7% potassium hydroxide solution after 23 hr and uses a scale of 1-7 proportionate to the degree of spreading (Little, Hilder et al. 1958). DSC, on the other hand, measures the temperature at which the crystalline regions of the starch granules irreversibly melt i.e. the temperature from the peak of the endotherm (Cuevas, Daygon et al. 2010). RVA measures the viscosity of rice starch during a temperature cycle and the temperature at which the starch granules begin to absorb water and swell is denoted as the pasting temperature and is directly associated with GT (Cuevas and Fitzgerald 2012). In NMR, the proton transverse magnetisation decay curve of a partly heated and cooled rice flour/water mixture is measured and the relaxation values denote GT (Gomi, Fukuoka et al. 1998). GT also gives an insight into the carbon footprint of rice since

lowering the GT of the rice grain could decrease average cooking times by up to 4 min. Such a reduction has been calculated to save >10,000 years of cooking time each year globally and hence, would represent a major saving in often scarce fuel resources (Fitzgerald, McCouch et al. 2009).

All the textural traits of freshly cooked rice and the potential of cooked rice to retrograde after cooking are largely influenced by the amylose content (AC) (Cuevas and Fitzgerald 2012). Thus, AC is used as a selection tool for eating quality in the early stages of rice breeding programs (Fitzgerald, Bergman et al. 2009). Amylose are mainly linear polymers of glucose units linked primarily by α-1,4 linkages and exist in starch granules along with amylopectin as highly branched high molecular weight molecules (Fitzgerald 2004). Based on amylose content, rice varieties are classified as being high (>25%), intermediate (20 – 25%), low (10 – 19%), very low (3 – 9%), or waxy (0 – 2%). AC is determined by a standard iodine colourimetric method (Fitzgerald, Bergman et al. 2009), size exclusion chromatography (SEC) (Ward, Gao et al. 2006) and by the genotyping of the Waxy (*Wx*) gene (Virk, Ford-Lloyd et al. 1995). Amylose, when mixed with iodine, forms an intense blue colour that can be quantitatively determined using a spectrophotometer (Cuevas and Fitzgerald 2012). In SEC, debranched amylose and amylopectin molecules separate because of the differences in structure of the chains that constitute them (Ward, Gao et al. 2006). Thus, in a debranched chain-length distribution curve, the proportion of amylose and amylopectin present in rice can be determined by measuring the peak areas of the polymers from about 30 to 3,000 degrees of polymerisation (DP) and those that are greater than 3,000 DP, respectively (Fitzgerald 2004). The amylose classes are associated with polymorphisms in the *Wx* gene that encodes the granule-bound starch synthase (GBSSI) enzyme responsible for amylose synthesis (Sano 1984). Two functional alleles Wx^a and Wx^b are associated with a G-T polymorphism at the 5' splice site at intron 1 resulting in differences between low amylose rice and high and intermediate amylose rice (Ayres, McClung et al. 1997). Wx^{in}, on the other hand, is an A-C polymorphism in exon 6 discriminating intermediate from high amylose classes (Chen, Bergman et al. 2008). Moreover, a null mutation with a 23-bp duplication in exon 2 of waxy rice resulted in a non-functional GBSSI protein (Wanchana, Toojinda et al. 2003) and no amylose (Fitzgerald, Bergman et al. 2009). Rice grains with low amylose are soft when cooked, sticky, and glossy, while higher amylose varieties tend to have grains that are dry, fluffy and separate (Juliano 2003). AC is therefore an important quality attribute strongly linked to regional consumer preferences for specific types of rice.

Gel consistency is used to measure the firmness of cooked rice and was developed primarily to differentiate texture of varieties with amylose contents above 25% (Juliano 2003). This is a standard assay that measures the distance travelled by a gel after the rice flour is cooked in 0.2M KOH, and is used to classify rice into hard, medium and soft classes (Cagampang, Perez et al. 1973). Recently, a single nucleotide polymorphism in exon 10 has been associated with hard and soft gel consistency of high AC rice (Tran, Daygon et al. 2011). Cooked rices with a hard gel consistency harden faster (Tang, Khush et al. 1991) but it was later found that this firmness does not change over 24 h, whereas rices with a soft gel consistency phenotype retrograde significantly over 24 h (Tran, Daygon et al. 2011).

The nutritional quality of rice is mainly determined by the protein content of the grain (Koutroubas, Mazzini et al. 2004). However, significant effort has also been made to elevate other nutritional components such as micronutrients – both minerals and vitamins – in rice grains in the past decade. For example, by enhancing the capacity of grains to accumulate iron through expression of the ferritin gene from soybean, a two- fold increase in the levels of iron and zinc were obtained in the endosperm of transgenic rice grains (Vasconcelos, Datta et al. 2003). Overexpression of nicotianamine synthase genes resulted in a 2-4-fold increase in the levels of iron and zinc in grains (Johnson, Kyriacou et al. 2011). Overexpression of a functional iron [Fe (II)]- and manganese [Mn(II)]-nicotianamine complex transporter not only increased the level of iron but also of manganese in polished rice (Ishimaru, Masuda et al. 2010). To help combat vitamin A deficiency, field trials are now being carried out with 'Golden rice' (www.goldenrice.org), a transgenic rice containing elevated levels of beta carotene (pro-vitamin A) obtained by introducing the phytoene synthase gene from maize (Zea mays) in combination with the *Erwinia uredovora* carotene desaturase (Paine, Shipton et al. 2005). Furthermore, rice with a low glycaemic index (GI) has also been considered a potential source of carbohydrates to assist in the management of type II diabetes (Fitzgerald, Rahman et al. 2011).

The aroma of fragrant rices is an issue of particular importance as it is not only a factor determining market price but also is a trait with clear local, cultural and national identity (Fitzgerald, Sackville Hamilton et al. 2008). Jasmine and basmati rices are highly preferred by consumers in countries of SE Asia South and Central Asia, respectively (Suwannaporn and Linnemann 2008). There are more than one hundred volatile metabolites present in rice but only a handful of these are considered to influence significantly and define rice aroma (Buttery, Turnbaugh et al. 1988, Jezussek, Juliano et al. 2002, Laguerre, Mestres et al.

2007, Yang, Lee et al. 2010, Bryant and McClung 2011, Mathure, Jawali et al. 2014). Of particular importance is 2-acetyl-1-pyrroline (2AP), which is long considered to be the major fragrance compound in rice and is associated with a popcorn-like aroma. 2AP has a very low odour threshold and hence can still influence the aroma phenotype and be detected by consumers at very low concentrations (Buttery, Turnbaugh et al. 1988). 2AP is currently being used in rice breeding programs to select for fragrant progeny (Fitzgerald, Sackville Hamilton et al. 2008). 2AP is normally determined either subjectively by sniffing a gram of rice grains that have been soaked for an hour in potassium hydroxide (Sood and Sidiq 1978), or quantitatively, using gas chromatography (Bergman, Delgado et al. 2000). Alternatively, as the genetic basis of 2AP synthesis is now known, the ability of a genotype to accumulate 2AP can be predicted by detecting the mutation in Betaine aldehyde dehydrogenase2 (Badh2) gene which has been directly linked to the presence of 2AP (Bradbury, Fitzgerald et al. 2005) and not to the quantity that accumulates.

The sensory properties of rice as experienced by consumers determine whether these consumers continue to purchase particular varieties of rice. Hence, a detailed understanding of sensory quality in rice is of paramount importance in assessing and designing new strategies for generating rice varieties with improved quality attributes. The flavour and aroma of rice can be described and measured analytically by panellists trained in descriptive sensory analysis (Meilgaard, Civille et al. 2007). In a descriptive sensory analysis, rice flavour and aroma are typically characterised by trained panellists using a lexicon with 10–12 descriptors (Champagne 2008). The panellists also measure the intensities of the flavour and aroma attributes using a universal scale for all foods (Meilgaard, Civille et al. 2007). Champagne et al (2005) employing an intensity scale of 0-15 with flavour components of US named brand products, found that the rating of aroma and flavour for rice varieties is generally between 1-3 with a maximum rating of 5. This poses challenges to sensory panellists where differences in aroma and flavour of rices can be difficult to detect because of low intensity.

Current phenotyping tools used in most quality evaluation programs find it difficult to discriminate between rice varieties that are within a particular quality class (Fitzgerald, McCouch et al. 2009). Nevertheless, a trained sensory panel recently employed a descriptive sensory analysis and successfully characterised and differentiated the flavour and aroma of premium and second best varieties from 9 rice-consuming countries (Champagne, Bett-Garber et al. 2010). Since flavour and aroma are defined in terms of mixtures of more than one hundred volatile compounds, it would be valuable to be able to

correlate descriptive sensory scores with volatile compounds detected in rice grains and eventually to determine which compounds can serve as markers for the most important flavour attributes (Chambers and Koppel 2013). However, 2AP to date, is the only volatile compound for which the relationship between its concentration in the rice grain and its sensory intensity has been fully established (Champagne 2008). More work is therefore needed in order to be able to generate a more detailed picture of the relationship between biochemical profiles on the one hand and sensory profile / phenotype on the other.

Climate change affects yield and quality

More than 90% of all rice is grown in Asia – but under a considerable range of agroclimatic conditions (Brar, Khush et al. 2013). About 50% of rice is grown under intensive irrigation systems i.e. farming that heavily relies on and ample water supply. Modern, high-yielding varieties have been designed for and can be grown very effectively under these conditions (Zeigler and Barclay 2008). The rest of the rice production area is rainfed which means it depends exclusively on natural rainfall and hence is strongly influenced by seasonal perturbations. Rainfed rice can grow on steeply sloping lands such as in the mountainous areas of Southeast Asia as well as on the flat lowlands of the delta and coastal areas of South and Southeast Asia (Mackill, Ismail et al. 2012).

Unfortunately, crop yield in many Asian countries has declined significantly due to more extreme weather patterns and increasing temperatures (Cruz, Harasawa et al. 2007). It is increasingly likely that more periods of drought, erratic rainfall distribution, flooding and a rise in sea level will be experienced due to climate change (Tester and Langridge 2010). The occurrence of climate-related diseases and heat stress in Central, East, South and South-East Asia has also increased due to variability in rainfall and a rise in temperature (Cruz, Harasawa et al. 2007). Currently, the most serious stresses affecting rice production in Asia are drought, submergence (flooding), and salt stress. Annually, these abiotic stresses affect about 23, 20, and 15 million ha, respectively (Brar, Khush et al. 2013) and have significant influence on global crop yield.

Salinity is a common problem in coastal areas due to the invasion of seawater onto agricultural land areas (Gregorio, Senadhira et al. 1997). Furthermore, overuse of underground water and inappropriate management of irrigation can also cause salt accumulation in inland rice production areas (Leung 2008). Although rice is considered moderately sensitive to salinity, salt stress affects rice particularly during the early seedling

and reproductive stages. This can result in decreased spikelet and tiller number and consequently, lower grain yield (Zeng, Lesch et al. 2003). Plant shoot height, root length, fresh weight, dry weight and leaf area are also inhibited in salt stressed rice plants (Cha-um, Ashraf et al. 2010). Starch synthase activity (α1-4-glucan glucosyle transferases) in developing rice grains was also inhibited significantly under sodium chloride salinity (Abdullah, Khan et al. 2001). The ability to produce mature seeds was also significantly inhibited in fragrant rice compared to non-fragrant varieties when exposed to a high salt treatment (Fitzgerald, Waters et al. 2010).

Heat stress can also cause serious yield and quality loss in rice production (Cao, Duan et al. 2008). Global average land and ocean surface temperature has increased by about 0.85ºC during the period 1880-2012 and will continue to do so (Hewitson, Janetos et al. 2014). It has been estimated that each 1°C increase minimum temperature would result in a 10% decline in grain yield during the dry cropping season (Peng, Huang et al. 2004). While an average daily temperature above 35°C lasting for a number of days can result in a significant yield loss, especially when this elevation occurs during the reproductive stage of growth (Cao, Duan et al. 2008). This is mainly due to induced spikelet sterility, a reduced number of spikelets per panicle, a lower percentage of filled spikelets and a reduced 1000-grain weight (Krishnan, Ramakrishnan et al. 2011). Exposure to high temperature also reduced amylose content and the gel consistency of rice grains (Lin, Li et al. 2010). Chalk is also induced by high temperature through decreasing the time needed to supply the substrate in the rice panicle during grain-filling (Fitzgerald and Resurreccion 2009). The initiation and packing of amyloplasts present in chalky grains has been shown to be susceptible to high temperatures (Lisle, Martin et al. 2000). In addition, for every 1% increase in chalkiness results in a decrease of 1% in head rice yield (Zhao and Fitzgerald 2013).

Annually, low temperatures can also cause crop yield losses of about 10% (Jiang, Jin et al. 2011). In rice, cold injury can be caused by exposure to periods of low temperature or the incorrect use of cold irrigation water during the different developmental stages of the plant (Cruz, Sperotto et al. 2013). This resulted in delays in germination and vegetative stages, lower rate of germination, lower stature and a decreased number of tillers (Fujino, Sekiguchi et al. 2004). During the grain filling stage, chilling temperature may result in delayed heading, an extended flowering period, and incomplete panicle exertion and spikelet sterility (Ye, Fukai et al. 2009). Rice plants grown under a cold treatment also show significant

decreases in grain width, chalkiness, 1000- grain weight and viscosity, whereas head rice yield, gel consistency and protein content increased (Zhao, Liu et al. 2009).

Most of the lowland and coastal rice production areas in the tropics and subtropics especially in South, East, and Southeast Asia, are considered to be flood-prone areas and are becoming increasingly vulnerable to a rise in sea level as well as an increased frequency and intensity of storms (Mackill, Ismail et al. 2012). Around 22 million hectares per year have been reportedly affected by flooding at various stages of the rice growth cycle (Jantaboon, Siangliw et al. 2011). It has been estimated that farmers can suffer up to 77% production losses when rice is submerged for two weeks in stagnant water especially when the water depth is more than 100 cm (Manzanilla, Paris et al. 2011). Diffusion of oxygen, carbon dioxide, and ethylene are often limited in plants under flooded conditions (Mackill, Ismail et al. 2012). During submergence, a lower intensity of light also impairs photosynthesis resulting in reduced biomass and increased plant mortality (Ram, Singh et al. 2002). Increased plant mortality was also observed as the percentage of silt increased in turbid floodwater, severely inhibiting the transmission of light reaching the submerged plants (Das, Panda et al. 2009). Submergence also resulted in delayed flowering and maturity, as well as a reduction in biomass, lower numbers of grains per panicle, reduced grain weight, and a lower harvest index (Singh, Mackill et al. 2009).

Cultivation of rice requires at least double the amount of water needed to grow other crops such as maize and wheat. On average, around 1,400L of water is needed to produce 1kg of rice (Bouman 2013). Thus, drought is considered to be a major cause of yield loss on an average of at least 23 million hectares of land (Pandey, Bhandari et al. 2007). Drought affects the growth of rice by reducing the levels of water and nutrients in the soil that are available to the plants (Fukai, Pantuwan et al. 1999). Reduced plant height and spikelet fertility were observed when rice is grown under drought conditions (Venuprasad, Dalid et al. 2009). Importantly, rice is once again most sensitive to drought stress during the period of reproductive growth, when it can delay flowering and hinder grain filling. This can result in an overall reduction in grain yield of up to 80% (Pantuwan, Fukai et al. 2002, Kumar, Bernier et al. 2008). Drought stress has also been shown to decrease plant biomass, number of panicles, leaf area and harvest index (Bernier, Kumar et al. 2007). Furthermore, the incidence of leaf blast infection was also observed to increase resulting in less leaf area for grain filling (Bocco, Lorieux et al. 2012).

In order to cope with these increasingly harsh, sub-optimal conditions brought about by climate change, breeding programs have been designed to develop more robust and climate-ready rice varieties which are less sensitive to environmental perturbation. Several salt and submergence tolerant varieties have been released in Lao PDR, Bangladesh, India, Vietnam and the Philippines (Mackill, Ismail et al. 2012, Gregorio, Islam et al. 2013). However, adoption of these new rice varieties by farmers even in stress prone areas has been remarkably low (Mackill, Ismail et al. 2012). One reason for this is that these tolerant varieties do not possess the same desirable traits as the existing "mega" varieties, i.e. varieties that are currently widely grown in major rice-growing areas in Asia and which are popular because of their high yield and superior grain quality (Fitzgerald, McCouch et al. 2009). Recently, submergence-tolerant versions of IR64, Sambha mashuri, Swarna, BR11 and TDK were released in Southeast Asia, and acceptability among farmers is favourable (Manzanilla, Paris et al. 2011). Attempts to develop other stress-tolerant improved varieties or varieties that possess multiple abiotic traits (salinity- submergence-drought) are now underway (Gregorio, Islam et al. 2013).

New opportunities through -Omics technologies

The rapid development in 'omics' technologies provides new opportunities to understand the complexity of biochemical networks in crop species (Langridge and Fleury 2011). In recent decades, whole genome sequencing and whole genome expression studies in model plants have led to the discovery of multiple phenotyype-relevant genes and their function (Mochida and Shinozaki 2011). New tools for functional genomics have also emerged in recent years, including high throughput methods for transcriptomic, proteomic, metabolomic and ionomic analyses. These appraoches have once again, significantly improved our ability to analyse physiological traits (Sanchez, Siahpoosh et al. 2008). With the aid of dedicated bioinformatic tools, complex fingerprinting and profiling approaches have been used to generate predictive models for gene expression studies and biochemical phenotype analysis (Shulaev, Cortes et al. 2008).

Genomics

Several important agronomic and quality-related traits in rice are considered as quantitative traits (Collard, Jahufer et al. 2005). The genes associated with a particular quantitative trait are located in regions within the genome known as quantitative trait loci (QTLs). The majority of these genes are still unknown, thus, QTLs are commonly used as molecular tools for marker-assisted selection (MAS) wherein the presence/ absence of QTLs/ markers

are used to substitute/ assist in phenotypic selection in a more efficient, effective, reliable and cost-effective way than conventional breeding (Collard, Jahufer et al. 2005). This can be done without the full knowledge of the actual genes responsible. Since the completion of rice genome sequence (IRGSP 2005), high density molecular maps have been developed that have led to major breakthroughs in mapping genes/ QTLs in rice (Brar, Khush et al. 2013). In addition, a variety of molecular marker assay platforms with different throughputs have also been developed to further assist the rice breeder in the more rapid design of new, improved varieties (Chen, He et al. 2013).

Restriction fragment length polymorphism (RFLP) and randomly amplified polymorphic DNA (RAPD) comprised the first generation markers. RFLPs have been the workhorse of genetic mapping studies since the approach was introduced several decades ago (Rieseberg 1998). RFLP is robust, reliable, and transferrable across populations (Collard, Jahufer et al. 2005). However, application requires large amounts of DNA and is time-consuming, laborious and expensive. Using a doubled haploid population derived from an indica/ japonica hybrid, several QTLs were identified for AC and gel consistency. Furthermore, a QTL associated with the major alkali degeneration gene, alk, on chromosome 6 for GT has also been identified (Gao, Zeng et al. 2011). In addition, 13 QTLs linked to seedling vigor traits such as length of shoots, roots, coleoptile and mesocotyl, have been identified and were estimated to account for phenotypic variance of up to 38% (Redoña and Mackill 1996). The RAPD method, on the other hand, is simple, inexpensive, and a large number of loci can be screened quickly (Rieseberg 1998). However, the majority of RAPD loci are dominant and this results in low linkage sensitivity and accuracy in F2 mapping populations. There are also reports of spurious bands and a lack of reproducibility for certain RAPD results (Rieseberg 1998). However, RAPD markers have been useful in identifying a major QTL, Sub1 on chromosome 9 that is involved in the regulation of the submergence response of rice (Xu and Mackill 1996).

Simple sequence repeats (SSRs) and amplified fragment length polymorphism (AFLP) comprise the second generation marker approaches. SSR markers, commonly referred to as microsatellites, are still frequently used by plant breeders to screen genotypes for specific traits (Chen, He et al. 2013). More than 500 microsatellite markers have been developed and are well distributed throughout the rice genome (McCouch, Temnykh et al. 2001). SSRs are highly reproducible, co-dominant in inheritance, simple and inexpensive (Collard and Mackill 2008). However, the development of SSR markers is labour-intensive and expensive (Rieseberg 1998). The use of SSR markers also requires polyacrylamide gel

electrophoresis and shows limited information for a single locus per assay (Collard and Mackill 2008). Using SSR markers, 4 QTLs which have a large effect on grain yield under severe drought stress were identified: qDTY2.1 (Venuprasad, Dalid et al. 2009), qDTY12.1 (Bernier, Kumar et al. 2007), qDTY2.2, and qDTY9.1 (Dixit, Swamy et al. 2012). Two of the 11 QTLs identified on chromosome 12 conferring cold tolerance at the booting stage of rice, qCTB2a and qCTB3, affect cold-induced wilting and necrosis (Andaya and Mackill 2003). The major QTL for grain length and weight, GS3, was also identified using SSR markers from a progeny produced by a cross between Minghui 63 and Chuan 7 (Fan, Yu et al. 2009). Polymorphic SSRs were also used to genotype F2 individuals from a mapping population of Kyeema and Gulfmont that were closely linked to a recessive gene *fgr* on chromosome 8 (Bradbury, Fitzgerald et al. 2005). AFLP markers, on the other hand, use restriction fragments with selective amplification (Zhu, Gale et al. 1998), and multiple AFLP markers can be assayed on a single gel making it an attractive choice for fine-scale mapping (Mackill, Zhang et al. 1996). However, large amounts of DNA are required (Collard, Jahufer et al. 2005). Several AFLP markers linked to thermosensitive male sterility trait (Dong, Subudhi et al. 2000) and the erect panicle trait (Kong, Wang et al. 2007) were identified.

The most recent developments in high throughput genotyping opportunities are DNA array platforms and next generation sequencing (NGS) technologies. Several array-based genotyping technologies have been developed such as Restriction Site-Associated DNA (RAD), Single Feature Polymorphism (SFP) and Single Nucleotide Polymorphism (SNP) (Chen, Xie et al. 2013). Due to the overwhelming number of SNPs in rice, these have been heavily used in linkage mapping and MAS (Collard, Jahufer et al. 2005, Thomson, Zhao et al. 2012). SNPs are usually bi-allelic which results from substitution, point mutation, insertion, deletion of nucleotides and can be detected when similar genomic regions from different genotypes are aligned (Chen, Xie et al. 2013). SNP genotyping is highly efficient and fast and is also cheaper compared to the previously-mentioned marker platforms (Thomson, Zhao et al. 2012). The first genome-wide SNP discovery pool was constructed from japonica (Nipponbare) then indica (93-11) varieties (Shen, Jiang et al. 2004). This was followed by the discovery of 160,000 high-quality SNPs using 20 rice varieties from the OryzaSNP project (McNally, Childs et al. 2009) and an even larger number of SNPs using 125 diverse rice varieties through the Rice SNP Consortium (www.ricesnp.org). Using the subsequently-developed SNP genotyping chips, SNP markers that were selected from the SNP discovery pools now provide valuable SNP frequency data within and between

germplasm groups (Thomson, Zhao et al. 2012). Using 1,536 SNPs in an Illumina GoldenGate SNP chip, 395 diverse rice varieties were categorised consistently into 5 major germplasm groups indica, aus, tropical japonica, temperate japonica and Group V (Zhao, Wright et al. 2010). In addition, 44,100 SNPs were used in a genome-wide association study of 413 diverse rice varieties from 82 countries that were phenotyped for 34 traits (Zhao, Tung et al. 2011). Recently, multiple, low-cost SNP sets of 384 SNPs from 1,536 and 44,100 SNPs were designed and optimised for application with different germplasm groups i.e. indica/ indica, indica/ japonica, japonica/ japonica, indica/ O. rufipogon, and japonica/ O. rufipogon. Using a 384 indica/ indica SNP chip, 2 putative QTLs, qHTSF1.1 and qHTSF4.1, associated with spikelet fertility under high temperature were identified (Ye, Argayoso et al. 2012). Among the NGS technologies, genotyping by sequencing (GBS) has become the marker genotyping platform of choice because of its robust, cost-effective, highly multiplexed sequencing approach (Beissinger, Hirsch et al. 2013, Mir, Hiremath et al. 2013). GBS involves the use of sites from restriction enzyme digestion, for barcode attachment and primer annealing (Spindel, Wright et al. 2013). Several QTLs for leaf width and aluminium tolerance were identified using 384-plex GBS with 30,984 markers in an indica × japonica mapping population (Spindel, Wright et al. 2013). In addition, genome-wide association mapping of 4 Multi-parent Advanced Generation Inter-Cross (MAGIC) populations using GBS, identified several known major genes and QTLs associated with blast and bacterial blight resistance, salinity and submergence tolerance, and grain quality (Bandillo, Raghavan et al. 2013). Clearly, the development of these genotyping tools are already having a major impact on rice breeding and it is anticipated that each new development shall further empower the rice breeder to select more rapidly for new varieties with the desired combinations of multigenic traits.

Metabolomics

It has been estimated that there are between 90,000–200,000 different metabolites in the plant kingdom (Fiehn 2002). Metabolites are the final products of cellular processes and characterise the response of biological systems to genetic or environmental changes (Shu, Frank et al. 2008). Metabolomics is the measurement of metabolites in biological systems under a given set of conditions, utilising high resolution analytical platforms such as mass spectrometry (MS) coupled with capillary electrophoresis (CE), gas chromatography (GC) or liquid chromatography (LC), or alternatively, detection using nuclear magnetic resonance (NMR). Dedicated chemometric statistical tools such as principal component analysis (PCA), partial least squares (PLS) and other non-canonical multivariate analyses

are then used to mine the data obtained to identify discriminatory metabolites and to help correlate metabolite profile and (biochemical) plant phenotype. Metabolomics studies are often used to define a set of biomarkers that reliably captures the metabolic changes of a plant species occuring in relation to development, environment or genotype (Tarpley, Duran et al. 2005). 'Targeted' metabolomics approaches aim to identify and quantify specific biochemically-relevant metabolites while 'non-targeted' metabolite profiling focuses on providing comprehensive metabolic information on a specific set of samples (Zhao, Wang et al. 2013).

Metabolomic studies have been used to reveal dynamic changes involved in different developmental stages of rice. Shu et al (2008) employed a time-dependent metabolite profiling approach to determine dynamic changes in the course of rice germination. They identified 174 metabolites including sugars, organic acids, amino acids, amines, fatty alcohols, acid methyl esters, hydrocarbons and sterols and found that more pronounced changes were particularly observed among the polar metabolites during the germination stage (Shu, Frank et al. 2008). In addition, 21 biomarker metabolites comprising primarily organic acids, sugars and amino acids, were identified that reliably captured 83% of the metabolite variation during the tillering or branching stage of rice growth (Tarpley, Duran et al. 2005). Recently, the parallel use of CE-MS and CE-diode array detection successfully measured dynamic changes at hourly intervals over a 24-hr period in the levels of 56 key metabolites involved in glycolysis, the tricarboxylic acid cycle (TCA), the pentose phosphate pathway, photorespiration and amino acid biosynthesis in rice leaves (Sato, Arita et al. 2008). Using Kohonen's self-organizing maps, metabolites involved in biochemical pathways influenced by the light and dark cycle were identified as were potential bottleneck enzymes controlling development-related metabolic networks (Sato, Arita et al. 2008).

Metabolite profiling studies have also led to new knowledge on how rice responds to abiotic stresses. Flowering of inferior spikelets was significantly reduced under severe drought conditions and this appeared to be correlated to decreased levels of spermidine and spermine while ethylene, 1-aminocylopropane-1-carboxylic acid and hydrogen peroxide increased (Chen, Xu et al. 2013). This led to the proposal that there is a potential interaction between ethylene and polyamines under dry soil conditions and this mediates the grain-filling of inferior spikelets in rice (Chen, Xu et al. 2013). Grain filling is also impaired by high temperature which appears linked to an inhibition of the accumulation of sugar phosphates and organic acids involved in glycolysis and the TCA cycle and increase in the levels of

sucrose and pyruvate/ oxaloacetate-derived amino acids (Yamakawa and Hakata 2010). Using H^1-high-resolution magic angle spinning and liquid-state NMR experiments, significant accumulation of amino acids and sugars were detected in shoots and roots of rice grown under drought and salt stress (Fumagalli, Baldoni et al. 2009), whereas pyruvic, citric, aconitic, malic, 2-oxoglutaric, shikimic, quinic acids and TCA cycle intermediates in roots were observed to be depleted under salt stress (Zuther, Koehl et al. 2007). Sucrose and amino acids that are synthesised from glycolysis intermediates and pyruvate were also observed to accumulate when rice is submerged for several days (Barding, Fukao et al. 2012). In subsequent work, combined H^1 NMR and GC- time of flight (TOF)-MS provided broader coverage of metabolites including e.g. S-methyl methionine and alanylglycine, and quantitation of sugars, organic acids and amino acids, that are essential in understanding the complex biochemical and molecular responses of rice following environmental perturbation (Barding, Béni et al. 2013). In a targeted profiling of soluble carbohydrates combined with measurements of oxidative products and antioxidative enzymes, in chilling-tolerant and sensitive rice varieties grown under chilling stress conditions, tolerance mechanisms have been predicted to be associated with the accumulation of osmoprotectants such as glucose, trehalose and mannitol and a more effective reactive oxygen species (ROS) scavenging system (Morsy, Jouve et al. 2007). Metabolomic analysis of rice roots treated with chromium showed increased levels of proline, ornithine, lactate, fructose, uracil and alanine suggesting that the modulation of the sucrose degradation pathway acts as a rescue mechanism when respiration is inhibited (Dubey, Misra et al. 2010). Air pollutants such as ozone (O_3) which is known to reduce photosynthesis, growth and yield and cause leaf injury and senescence, has been shown to enhance the accumulation of gamma-aminobutyric acid, some amino acids and glutathione in exposed leaves (Cho, Shibato et al. 2008).

Several metabolomic studies have revealed certain nutritional and functional properties of rice. Kusano et al (2007) developed a comprehensive method combining one-dimensional (1D) and two-dimensional (GC×GC) gas chromatography-time-of-flight (TOF)–mass spectrometry in phenotyping metabolic variants in 70 unpolished rices from the world rice core collection. Alpha-tocopherol, gamma-aminobutyric acid, glycerol-3-phosphate, myristate, fructose, indole-1-acetic acid, inositol-1-phosphate, trehalose, cholesterol, and raffinose were detected as discriminatory compounds contributing to significant differences between japonica and indica varieties (Kusano, Fukushima et al. 2007). Ten contrasting varieties of cooked rice also clustered according to their subspecies

i.e. indica, japonica and Aus based on data on phenolics, vitamin E, phytosterols, and linolenic acid from a total of 3,097 metabolites measured (Heuberger, Lewis et al. 2010). Alpha- tocopherol and gamma-aminobutyric acid were also found to be present in high levels in red and black colored rices (Frank, Reichardt et al. 2012). Interestingly, the black rice also contained higher levels of fatty acid methy esters, free fatty acids, organic acids and amino acids when compared to white and red rice (Frank, Reichardt et al. 2012). Ferulic and p-coumaric acid, total flavonoid content and antioxidant activity were also found to be higher in red and black Thai rice when compared to the white rices analysed (Vichapong, Sookserm et al. 2010). In a health-related study, remarkably, the bioactive properties of rice bran in preventing chronic disease, appeared linked to phytochemicals detected from rice bran from 3 rice varieties that had been fermented with *Saccharomyces cerevisiae* var. *boulardii* which reduced the growth of human B lymphomas compared to the nonfermented control indicating that fermentation differentially changed the profile of bioactive compounds (Ryan, Heuberger et al. 2011).

Metabolite profiling studies are useful for identifying metabolites associated with quality traits in rice. Shen et al (2009) showed that the flavonoid content was positively correlated with grain length and shape, and negatively with the 100-grain weight among 481 rice accessions. In addition, 100-grain weight had negative correlations with phenolic content and antioxidant capacity (Shen, Jin et al. 2009). Similarly, in 67 rice varieties from a rice diversity research set, grain length and shape were positively correlated with succinate, glucose-6-phosphate, and glycine, while putrescine was strongly correlated with a number of traits i.e. ear emergence day and amylose content, even though these traits are not correlated with each other (Redestig, Kusano et al. 2011). Moreover, amylose content showed negative correlations with glycerol, linoleic acid, palmitic acid, and phosphate, although the biology describing such correlation is still unclear (Kusano, Fukushima et al. 2012).

Metabolomic studies have also started to help us in understanding the complexity of fragrance and taste in rice. Significant differences were found between 3 japonica cultivars in the volatile profiles that were quantitatively and qualitatively affected by the degree of milling, with most of the aldehydes located in the endosperm of the grain (Yang, Lee et al. 2008). Dimethyl sulphide, phenolic compounds and most lipid-oxidation products increased significantly when brown rice was allowed to germinate which is a new cereal product gaining popularity in China and Japan because of its nutritional value and taste (Wu, Yang et al. 2011). Because rice is generally consumed as cooked rice, profiling volatile

compounds liberated during cooking has also been investigated. Zeng et al. (2009) monitored the compounds emitted in the four stages of cooking and found nonanal and hexanal were the major components detected in the first stage of cooking i.e. from the time of the first heating to the time steam was emitted from the rice cooker. Moreover, hexadecanoic acid was the predominant volatile compound detected at the second stage of cooking, from the start to the end of steam leaving the rice cooker (Zeng, Zhang et al. 2009). Key odorant compounds such as (E)-2-nonenal, (E,E)-2,4-decadienal, 2-methoxy-4-vinylphenol, indole, and vanillin were detected in the last two stages of cooking, from the end of steam coming out of the rice cooker to the stage when rice is kept warm for a couple of minutes (Zeng, Zhang et al. 2009). Studies on volatile components of different types of rice have also been conducted. Guaiacol and 2-AP were the major components giving black rice its unique character (Yang, Lee et al. 2007). Non-fragrant rices contained more n-hexanal, (E)-2-heptenal, 1-octen-3-ol, n-nonanal, (E)-2-octenal, (E)-2,(E)-4-decadienal, 2-pentylfuran, 4-vinylguaiacol and 4-vinylphenol, than fragrant rices (Widjaja, Craske et al. 1996, Bryant and McClung 2011), whereas fragrant rice had higher amounts of 2-AP, guaiacol, indole, hexanal, nonanal, decanal, benzyl alcohol, vanillin, 2-phenylethanol, oct-1-en-3-ol, octan-1-ol, hexan- 1-ol, butanoic acid, and hexanoic acid (Widjaja, Craske et al. 1996, Maraval, Mestres et al. 2008, Bryant and McClung 2011, Mathure, Wakte et al. 2011).

Recently, metabolite profiling has played a significant role in identifying metabolic markers that are now being employed in association with mapping of specific genes related to compositional quality of one or more compounds in crops (Fernie and Schauer 2008). Metabolic profiles and association mapping of 48 rice varieties from a Chinese core collection led to the identification of 20 marker loci that are associated with the levels of amino acids and sugars (Lou, Ma et al. 2011). SNPs associated with total phenolics and tocopherol, have also been identified after profiling metabolites in cooked brown rice (Heuberger, Lewis et al. 2010). In addition, more than a hundred metabolome quantitative trait loci (mQTL) were identified in Sasanishiki x Habataki back- crossed inbred lines using CE-TOF-MS, GC-TOF-MS, LC-Q-TOF-MS and LC-IT-TOFMS, used to analyse an extensive range of polar, primary, and secondary metabolites, and lipids (Matsuda, Okazaki et al. 2012). An mQTL hotspot on chromosome 3 was identified that was associated with coordinated control of amino acid and triacylglycerol levels (Matsuda, Okazaki et al. 2012), perhaps suggesting a regulatory process. Currently, little is actually known about the genetic basis of aroma in rice aside from 2AP (Bradbury, Fitzgerald et al. 2005, Chen,

Yang et al. 2008, Kovach, Calingacion et al. 2009). Starting with a well-chosen mapping population, metabolomic profiling of flavour compounds is a potentially valuable tool to help further characterise rice quality in relation to aroma and flavour. This is true also for varieties that are so-called 'non-fragrant', but which, like all rice, do have a characteristic aroma on cooking, when the lid of the cooking pot is lifted, and when the rice approaches the human sensory organs on a fork, or between chopsticks.

Post-genomic breeding

The outlook is hugely promising. The completion of the resequencing of a core collection of 3,000 rice accessions from 89 countries (The3000ricegenomesproject 2014), will offer untold opportunities for the large-scale discovery of novel alleles for important rice phenotypic traits. With the dynamic incorporation of other 'omics technologies in a coordinated complementary manner in rice genomic studies, coupled with powerful dedicated bioinformatics tools, obstacles on the way to 'breeding by design' shall increasingly be overcome one by one (Chen, He et al. 2013). The breeder will become increasingly empowered to encorporate such approaches as part of future crop improvement strategies. For example, gene and QTL pyramiding is just one promising approach where combining several QTLs, each associated with a trait controlled by different mechanism, can enhance its trait effect, and can help develop optimal combinations of QTLs for several traits (Thomson 2009). But there is also huge potential for linking genetics and genomics methodologies with other omics approaches. Opportunities for using, for example, proteomics and metabolomics to help first define phenotype in chemical terms and then, identify the genetics behind crop-relevant phenotypic (chemical) differences are of particular promise. Developments in these omics approaches somewhat lag behind what is now possible in genomics, certainly in terms of speed and scale. More attention is needed to maximise the potential of these technologies and demonstrate their potential and feasibility. That was one of the central goals of the work reported herein, and then specifically, for the field of metabolomics.

Thesis outline

The overarching goal of this PhD research project was to develop and utilise a multi-disciplinary approach to understand quality traits in rice. This work aims to enhance current phenotyping tools in rice quality evaluation programs to help breeders select for new varieties that capture what consumers highly prefer in terms of taste, aroma etc. We set out to: determine if metabolomics can become a valid phenotyping tool to associate with

new genotyping tools; to determine whether metabolomics can be really used as a tool to assign quality; to identify taste and flavour compounds that correlate with traits which are either desired by consumers or alternatively are not desired by consumers because of their 'off flavour' attributes. The materials used in the research were selected as they are related to both high / low quality aspects as well as other environmentally agronomic traits related to yield under suboptimal conditions. The work presented in this thesis uses multi-disciplinary approach to understand rice grain quality traits but such an approach could also prove to have broad applicability to many other crop species.

To reach a deeper understanding of how consumer preferences drive the market and differ across rice consuming regions, Chapter 2 describes the results of an extensive analysis of regional preferences for the main features of rice quality that are routinely measured in quality evaluation programs – namely, size and shape of the grain, amylose content, gelatinisation temperature, gel consistency and fragrance.

In Chapter 3, a novel multiplatform metabolomic and mineral approach with genome-wide genotyping has been used to investigate the extent of biochemical and genetic diversity of three premium waxy rice cultivars of great importance in Lao PDR and how this may or may not be related to the use of artificial fertilizer during cultivation and to determine whether there was any basis for associating metabolomic data with genomic information.

Chapters 4 and 5 are focused on our aim to gain a profound understanding of the influence of water availability to the metabolome of drought tolerant rice. For this we have focused on 2 contrasting rice varieties and their progeny population in terms of flavour and aroma and drought tolerance. In Chapter 4 the varieties Apo and IR64 have been evaluated. Metabolite profiling, sensory analysis and genome-wide genotyping approaches have been used on grains of Apo and IR64 obtained from plants grown in stressed and non-stressed environments. Multivariate statistical analysis was performed on the data obtained to identify traits that were affected by water availability. A progeny population obtained from Apo and IR64 was used in Chapter 5 to identify metabolites that contribute to flavour and aroma of rice. Sensory and genetic analyses were also conducted to identify flavour and aroma descriptors and QTLs associated with aroma.

In Chapter 6 a General Discussion has been compiled to draw attention to the overall main conclusions of this experimental work and to relate these both to previous knowledge and

the state of the art as well as the future potential (and limitations) for the use of metabolomics approaches in the context of rice crop quality parameters.

References

Abdullah, Z., M. Khan and T. Flowers (2001). Causes of sterility in seed set of rice under salinity stress. Journal of Agronomy and Crop Science **187**(1): 25-32.

Andaya, V. and D. Mackill (2003). QTLs conferring cold tolerance at the booting stage of rice using recombinant inbred lines from a japonica × indica cross. Theoretical and Applied Genetics **106**(6): 1084-1090.

Ayres, N., A. McClung, P. Larkin, H. Bligh, C. Jones and W. Park (1997). Microsatellites and a single-nucleotide polymorphism differentiate apparent amylose classes in an extended pedigree of US rice germplasm. Theoretical and Applied Genetics **94**(6-7): 773-781.

Balasubramanian, V., M. Sie, R. Hijmans and K. Otsuka (2007). Increasing rice production in Sub-Saharan Africa: Challenges and opportunities. Advances in Agronomy. L. Donald, Academic Press. **94:** 55-133.

Bandillo, N., C. Raghavan, P. Muyco, M. Sevilla, I. Lobina, C. Dilla-Ermita, C. Tung, S. McCouch, M. Thomson, R. Mauleon, R. Singh, G. Gregorio, E. Redona and H. Leung (2013). Multi-parent advanced generation inter-cross (MAGIC) populations in rice: Progress and potential for genetics research and breeding. Rice. **6:** 11-26.

Barding, G., S. Béni, T. Fukao, J. Bailey-Serres and C. Larive (2013). Comparison of GC-MS and NMR for metabolite profiling of rice subjected to submergence stress. Journal of Proteome Research **12**(2): 898-909.

Barding, G., T. Fukao, S. Béni, J. Bailey-Serres and C. Larive (2012). Differential metabolic regulation governed by the rice SUB1A gene during submergence stress and identification of alanylglycine by ^{1}H NMR spectroscopy. Journal of Proteome Research **11**(1): 320-330.

Beissinger, T., C. Hirsch, R. Sekhon, J. Foerster, J. Johnson, G. Muttoni, B. Vaillancourt, C. Buell, S. Kaeppler and N. de Leon (2013). Marker density and read depth for genotyping populations using genotyping-by-sequencing. Genetics **193**(4): 1073-1081.

Belder, P., J. Spiertz, B. Bouman, G. Lu and T. Tuong (2005). Nitrogen economy and water productivity of lowland rice under water-saving irrigation. Field Crops Research **93**(2–3): 169-185.

Bergman, C., J. Delgado, R. Bryant, C. Grimm, K. Cadwallader and B. Webb (2000). Rapid gas chromatographic technique for quantifying 2-acetyl-1-pyrroline and hexanal in rice (*Oryza sativa*, L.). Cereal Chemistry **77**(4): 454-458.

Bernier, J., A. Kumar, R. Venuprasad, D. Spaner and G. Atlin (2007). A large-effect QTL for grain yield under reproductive-stage drought stress in upland rice. Crop Science **47**(2): 507-516.

Bocco, R., M. Lorieux, P. Seck, K. Futakuchi, B. Manneh, H. Baimey and M. Ndjiondjop (2012). Agro-morphological characterization of a population of introgression lines derived from crosses between IR 64 (*Oryza sativa* indica) and TOG 5681 (*Oryza glaberrima*) for drought tolerance. Plant Science **183**: 65-76.

Boualaphanh, C., M. Calingacion, R. Cuevas, D. Jothityangkoon, J. Sanitchon and M. Fitzgerald (2011). Yield and quality of traditional and improved Lao varieties of rice. Scienceasia **37**(2): 89-97.

Bouman, B. (2013) Does rice really use too much water? IRRI.

Bradbury, L., T. Fitzgerald, R. Henry, Q. Jin and D. Waters (2005). The gene for fragrance in rice. Plant Biotechnology Journal **3**(3): 363-370.

Brar, D., G. Khush, S. Gurbir and K. Navreet (2013). Chapter 8 - Biotechnological approaches for increasing productivity and sustainability of rice production. Agricultural Sustainability. San Diego, Academic Press**:** 151-175.

Bryant, R. and A. McClung (2011). Volatile profiles of aromatic and non-aromatic rice cultivars using SPME/GC–MS. Food Chemistry **124**(2): 501-513.

Buttery, R., J. Turnbaugh and L. Ling (1988). Contribution of volatiles to rice aroma. Journal of Agricultural and Food Chemistry **36**(5): 1006-1009.

Cagampang, G., C. Perez and B. Juliano (1973). A gel consistency test for eating quality of rice. Journal of Science, Food and Agriculture **24**: 1589-1594.

Cao, Y., H. Duan, L. Yang, Z. Wang, S. Zhou and J. Yang (2008). Effect of heat stress during meiosis on grain yield of rice cultivars differing in heat tolerance and its physiological mechanism. Acta Agronomica Sinica **34**(12): 2134-2142.

Cha-um, S., M. Ashraf and C. Kirdmanee (2010). Screening upland rice (*Oryza sativa* L. ssp. indica) genotypes for salt-tolerance using multivariate cluster analysis. African Journal of Biotechnology **9**(30): 4731-4740.

Chambers, E. and K. Koppel (2013). Associations of volatile compounds with sensory aroma and flavor: The complex nature of flavor. Molecules **18**(5): 4887-4905.

Champagne, E. (2008). Rice aroma and flavor: A literature review. Cereal Chemistry **85**(4): 445-454.

Champagne, E., K. Bett-Garber, M. Fitzgerald, C. Grimm, J. Lea, K. Ohtsubo, S. Jongdee, L. Xie, P. Bassinello, A. Resurreccion, R. Ahmad, F. Habibi and R. Reinke (2010). Important sensory properties differentiating premium rice varieties. Rice **3**: 270–281.

Champagne, E., K. Bett, B. Vinyard, A. McClung, F. Barton, K. Moldenhauer, S. Linscombe and K. McKenzie (1999). Correlation between cooked rice texture and rapid visco analyser measurements. Cereal Chemistry **76**(5): 764-771.

Chen, H., H. He, F. Zhou, H. Yu and X. Deng (2013). Development of genomics-based genotyping platforms and their applications in rice breeding. Current Opinion in Plant Biology **16**(2): 247-254.

Chen, H., W. Xie, H. He, H. Yu, W. Chen, J. Li, R. Yu, Y. Yao, W. Zhang, Y. He, X. Tang, F. Zhou, X. Deng and Q. Zhang (2013). A high-density SNP genotyping array for rice biology and molecular breeding. Molecular Plant **7**(3): 541-553.

Chen, M., C. Bergman, S. Pinson and R. Fjellstrom (2008). Waxy gene haplotypes: Associations with apparent amylose content and the effect by the environment in an international rice germplasm collection. Journal of Cereal Science **47**(3): 536-545.

Chen, S., Y. Yang, W. Shi, Q. Ji, F. He, Z. Zhang, Z. Cheng, X. Liu and M. Xu (2008). Badh2, encoding betaine aldehyde dehydrogenase, inhibits the biosynthesis of 2-acetyl-1-pyrroline, a major component in rice fragrance. The Plant Cell **20**(7): 1850-1861.

Chen, T., Y. Xu, J. Wang, Z. Wang, J. Yang and J. Zhang (2013). Polyamines and ethylene interact in rice grains in response to soil drying during grain filling. Journal of Experimental Botany **64**(8): 2523-2538.

Cheng, S., L. Cao, J. Zhuang, S. Chen, X. Zhan, Y. Fan, D. Zhu and S. Min (2007). Super hybrid rice breeding in China: Achievements and prospects. Journal of Integrative Plant Biology **49**(6): 805-810.

Childs, N. (2004). Production and utilization of rice. Rice: Chemistry and Technology. E. Champagne. St. Paul, Minnesota, USA, American Association of Cereal Chemists, Inc.**:** 1-24.

Cho, K., J. Shibato, G. Agrawal, Y. Jung, A. Kubo, N. Jwa, S. Tamogami, K. Satoh, S. Kikuchi, T. Higashi, S. Kimura, H. Saji, Y. Tanaka, H. Iwahashi, Y. Masuo and R. Rakwal (2008). Integrated transcriptomics, proteomics, and metabolomics analyses to survey ozone responses in the leaves of rice seedling. Journal of Proteome Research **7**(7): 2980-2998.

Chun, A., J. Song, K. Kim and H. Lee (2009). Quality of head and chalky rice and deterioration of eating quality by chalky rice. Journal of Crop Science and Biotechnology **12**(4): 239-244.

Collard, B., M. Jahufer, J. Brouwer and E. Pang (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica **142**(1-2): 169-196.

Collard, B. and D. Mackill (2008). Marker-assisted selection: An approach for precision plant breeding in the twenty-first century. Philosophical Transactions of the Royal Society B: Biological Sciences **363**(1491): 557-572.

Cramer, G., E. Wailes and S. Shui (1993). Impacts of liberalizing trade in the world rice market. American Journal of Agricultural Economics **75**(1): 219-226.

Cruz, R., H. Harasawa, M. Lal, S. Wu, Y. Anokhin, B. Punsalmaa, Y. Honda, M. Jafari, C. Li and N. Huu Ninh (2007). Asia. Climate Change 2007: Impacts, Adaptation and Vulnerability. Fourth Assessment Report of the Intergovernmental Panel on Climate Change. M. Parry, O. Canziani, J. Palutikof, P. van der Linden and C. Hanson. Cambridge, UK, Cambridge University Press**:** 469-506.

Cruz, R., R. Sperotto, D. Cargnelutti, J. Adamski, T. de FreitasTerra and J. Fett (2013). Avoiding damage and achieving cold tolerance in rice plants. Food and Energy Security **2**(2): 96-119.

Cuevas, R., V. Daygon, H. Corpuz, L. Nora, R. Reinke, D. Waters and M. Fitzgerald (2010). Melting the secrets of gelatinisation temperature in rice. Functional Plant Biology **37**(5): 439-447.

Cuevas, R. and M. Fitzgerald (2012). Genetics of rice quality. Genetic Diversity. S. Junjicic. Rijeka, Croatia, In-Tech. **3:** 285-310.

Das, K., D. Panda, R. Sarkar, J. Reddy and A. Ismail (2009). Submergence tolerance in relation to variable floodwater conditions in rice. Environmental and Experimental Botany **66**(3): 425-434.

Dixit, S., B. Swamy, P. Vikram, H. Ahmed, M. Sta Cruz, M. Amante, D. Atri, H. Leung and A. Kumar (2012). Fine mapping of QTLs for rice grain yield under drought reveals sub-QTLs conferring a response to variable drought severities. Theoretical and Applied Genetics **125**(1): 155-169.

Donald, C. (1968). The breeding of crop ideotypes. Euphytica **17**(3): 385-403.

Dong, N., P. Subudhi, P. Luong, V. Quang, T. Quy, H. Zheng, B. Wang and H. Nguyen (2000). Molecular mapping of a rice gene conditioning thermosensitive genic male sterility using AFLP, RFLP and SSR techniques. Theoretical and Applied Genetics **100**(5): 727-734.

Dubey, S., P. Misra, S. Dwivedi, S. Chatterjee, S. Bag, S. Mantri, M. Asif, A. Rai, S. Kumar, M. Shri, P. Tripathi, R. Tripathi, P. Trivedi, D. Chakrabarty and R. Tuli (2010). Transcriptomic and metabolomic shifts in rice roots in response to Cr (VI) stress. BMC Genomics **11**(1): 648-667.

Fan, C., S. Yu, C. Wang and Y. Xing (2009). A causal C–A mutation in the second exon of GS3 highly associated with rice grain length and validated as a functional marker. Theoretical and Applied Genetics **118**(3): 465-472.

Fernie, A. and N. Schauer (2008). Metabolomics-assisted breeding: a viable option for crop improvement? Trends in Genetics **25**(1): 39-48.

Fiehn, O. (2002). Metabolomics – the link between genotypes and phenotypes. Plant Molecular Biology **48**: 155–171.

Fitzgerald, M. (2004). Starch. Rice: Chemistry and Technology. E. Champagne. St. Paul, Minnesota, USA, American Association of Cereal Chemists, Inc.: 109-142.

Fitzgerald, M., C. Bergman, A. Resurreccion, J. Möller, R. Jimenez, R. Reinke, M. Martin, P. Blanco, F. Molina, M. Chen, V. Kuri, M. Romero, F. Habibi, T. Umemoto, S. Jongdee, E. Graterol, K. Reddy, P. Bassinello, R. Sivakami, N. Rani, S. Das, Y. Wang, S. Indrasari, A. Ramli, R. Ahmad, S. Dipti, L. Xie, N. Lang, P. Singh, D. Toro, F. Tavasoli and C. Mestres (2009). Addressing the dilemmas of measuring amylose in rice. Cereal Chemistry **86**(5): 492-498.

Fitzgerald, M., S. McCouch and R. Hall (2009). Not just a grain of rice: the quest for quality. Trends in Plant Science **14**(3): 133-139.

Fitzgerald, M., S. Rahman, A. Resurreccion, J. Concepcion, V. Daygon, S. Dipti, K. Kabir, B. Klingner, M. Morell and A. Bird (2011). Identification of a major genetic determinant of glycaemic index in rice. Rice **4**(2): 66-74.

Fitzgerald, M. and A. Resurreccion (2009). Maintaining the yield of edible rice in a warming world. Functional Plant Biology **36**(12): 1037-1045.

Fitzgerald, M., N. Sackville Hamilton, M. Calingacion, H. Verhoeven and V. Butardo (2008). Is there a second fragrance gene in rice? Plant Biotechnology Journal **6**(4): 416-423.

Fitzgerald, T., D. Waters, L. Brooks and R. Henry (2010). Fragrance in rice (*Oryza sativa* L.) is associated with reduced yield under salt treatment. Environmental and Experimental Botany **68**(3): 292-300.

Frank, T., B. Reichardt, Q. Shu and K. Engel (2012). Metabolite profiling of colored rice (*Oryza sativa* L.) grains. Journal of Cereal Science **55**(2): 112-119.

Fujino, K., H. Sekiguchi, T. Sato, H. Kiuchi, Y. Nonoue, Y. Takeuchi, T. Ando, S. Lin and M. Yano (2004). Mapping of quantitative trait loci controlling low-temperature germinability in rice (*Oryza sativa* L.). Theoretical and Applied Genetics **108**(5): 794-799.

Fukai, S., G. Pantuwan, B. Jongdee and M. Cooper (1999). Screening for drought resistance in rainfed lowland rice. Field Crops Research **64**(1–2): 61-74.

Fumagalli, E., E. Baldoni, P. Abbruscato, P. Piffanelli, A. Genga, R. Lamanna and R. Consonni (2009). NMR techniques coupled with multivariate statistical analysis: Tools to analyse *Oryza sativa* metabolic content under stress conditions. Journal of Agronomy and Crop Science **195**(2): 77-88.

Gao, Z., D. Zeng, F. Cheng, Z. Tian, L. Guo, Y. Su, M. Yan, H. Jiang, G. Dong, Y. Huang, B. Han, J. Li and Q. Qian (2011). ALK, the key gene for gelatinization temperature, is a modifier gene for gel consistency in rice. Journal of Integrative Plant Biology **53**(9): 756-765.

Gomi, Y., M. Fukuoka, T. Mihori and H. Watanabe (1998). The rate of starch gelatinization as observed by PFG-NMR measurement of water diffusivity in rice starch/water mixtures. Journal of Food Engineering **36**(4): 359-369.

Graham, R. (2002). A proposal for IRRI to establish a grain quality and nutrition research center. IRRI Discussion Paper Series No. 44. IRRI, Los Baños (Philippines)**:** 1-18.

Gregorio, G., M. Islam, G. Vergara and S. Thirumeni (2013). Recent advances in rice science to design salinity and other abiotic stress tolerant rice varieties. Sabrao Journal of Breeding and Genetics **45**(1): 31-41.

Gregorio, G., D. Senadhira and R. Mendoza (1997). Screening rice for salinity tolerance. IRRI Discussion Paper Series No. 22. Los Banos, Philippines, International Rice Research Institute**:** 1-31.

Heuberger, A., M. Lewis, M. Chen, M. Brick, J. Leach and E. Ryan (2010). Metabolomic and functional genomic analyses reveal varietal differences in bioactive compounds of cooked rice. PLoS One **5**(9): e12915.

Hewitson, B., A. Janetos, T. Carter, F. Giorgi, R. Jones, W. Kwon, L. Mearns, E. Schipper and M. van Aalst (2014). Regional Context In Climate Change 2014: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the IPCC Fifth Assessment Report - Changes to the Underlying Scientific/Technical Assessment. T. Downing and P. Duffy. Cambridge, United Kingdom and New York, NY, USA, Cambridge University Press. **2:** 1-97.

Ishimaru, Y., H. Masuda, K. Bashir, H. Inoue, T. Tsukamoto, M. Takahashi, H. Nakanishi, N. Aoki, T. Hirose, R. Ohsugi and N. Nishizawa (2010). Rice metal-nicotianamine transporter, OsYSL2, is required for the long-distance transport of iron and manganese. Plant Journal **62**(3): 379-390.

Jantaboon, J., M. Siangliw, S. Im-mark, W. Jamboonsri, A. Vanavichit and T. Toojinda (2011). Ideotype breeding for submergence tolerance and cooking quality by marker-assisted selection in rice. Field Crops Research **123**(3): 206-213.

Jezussek, M., B. Juliano and P. Schieberle (2002). Comparison of key aroma compounds in cooked brown rice varieties based on aroma extract dilution analyses. Journal of Agricultural and Food Chemistry **50**(5): 1101-1105.

Jiang, W., Y. Jin, J. Lee, K. Lee, R. Piao, L. Han, J. Shin, R. Jin, T. Cao, H. Pan, X. Du and H. Koh (2011). Quantitative trait loci for cold tolerance of rice recombinant inbred lines in low temperature environments. Molecules and Cells **32**(6): 579-587.

Johnson, A., B. Kyriacou, D. Callahan, L. Carruthers, J. Stangoulis, E. Lombi and M. Tester (2011). Constitutive overexpression of the OsNAS gene family reveals single-gene strategies for effective iron- and zinc-biofortification of rice endosperm. PLoS One **6**(9): e24476.

Juliano, B. (2003). Rice Chemistry and Quality. Munoz, Nueva Ecija, Philippine Rice Research Institute.

Kajala, K., S. Covshoff, S. Karki, H. Woodfield, B. Tolley, M. Dionora, R. Mogul, A. Mabilangan, F. Danila, J. Hibberd and W. Quick (2011). Strategies for engineering a two-celled C4 photosynthetic pathway into rice. Journal of Experimental Botany **62**(9): 3001-3010.

Khush, G. (2005). What it will take to Feed 5.0 Billion Rice consumers in 2030. Plant Molecular Biology **59**(1): 1-6.

Kong, F., J. Wang, J. Zou, L. Shi, M. De Jin, Z. Xu and B. Wang (2007). Molecular tagging and mapping of the erect panicle gene in rice. Molecular Breeding **19**(4): 297-304.

Koutroubas, S., F. Mazzini, B. Pons and D. Ntanos (2004). Grain quality variation and relationships with morpho-physiological traits in rice (*Oryza sativa* L.) genetic resources in Europe. Field Crops Research **86**(2–3): 115-130.

Kovach, M., M. Calingacion, M. Fitzgerald and S. McCouch (2009). The origin and evolution of fragrance in rice (*Oryza sativa* L.). Proceedings of the National Academy of Sciences **106**(34): 14444-14449.

Krishnan, P., B. Ramakrishnan, K. Reddy, V. Reddy and L. Donald (2011). Chapter 3 - High-temperature effects on rice growth, yield, and grain quality. Advances in Agronomy, Academic Press. **111:** 87-206.

Kumar, A., J. Bernier, S. Verulkar, H. Lafitte and G. Atlin (2008). Breeding for drought tolerance: Direct selection for yield, response to selection and use of drought-tolerant donors in upland and lowland-adapted populations. Field Crops Research **107**(3): 221-231.

Kusano, M., A. Fukushima, N. Fujita, Y. Okazaki, M. Kobayashi, N. Oitome, K. Ebana and K. Saito (2012). Deciphering starch quality of rice kernels using metabolite profiling and pedigree network analysis. Molecular Plant **5**(2): 442-451.

Kusano, M., A. Fukushima, M. Kobayashi, N. Hayashi, P. Jonsson, T. Moritz, K. Ebana and K. Saito (2007). Application of a metabolomic method combining one-dimensional and two-dimensional gas chromatography-time-of-flight/mass spectrometry to metabolic phenotyping of natural variants in rice. Journal of Chromatography B **855**(1): 71-79.

Laguerre, M., C. Mestres, F. Davrieux, J. Ringuet and R. Boulanger (2007). Rapid discrimination of scented rice by solid-phase microextraction, mass spectrometry, and multivariate analysis used as a mass sensor. Journal of Agricultural and Food Chemistry **55**(4): 1077-1083.

Langridge, P. and D. Fleury (2011). Making the most of 'omics' for crop breeding. Trends in Biotechnology **29**(1): 33-40.

Laza, M., S. Peng, S. Akita and H. Saka (2003). Contribution of biomass partitioning and translocation to grain yield under sub-optimum growing conditions in irrigated rice. Plant Production Science **6**(1): 28-35.

Leung, H. (2008). Stressed genomics—bringing relief to rice fields. Current Opinion in Plant Biology **11**(2): 201-208.

Lin, C., C. Li, S. Lin, F. Yang, J. Huang, Y. Liu and H. Lur (2010). Influence of high temperature during grain filling on the accumulation of storage proteins and grain quality in rice (*Oryza sativa* L.). Journal of Agricultural and Food Chemistry **58**(19): 10545-10552.

Lisle, A., M. Martin and M. Fitzgerald (2000). Chalky and translucent rice grains differ in starch composition and structure and cooking properties. Cereal Chemistry **77**(5): 627-632.

Little, R., G. Hilder and E. Dawson (1958). Differential effect of dilute alkali on 25 varieties of milled white rice. Cereal Chemistry **35**(111-126).

Liu, L., T. Chen, Z. Wang, H. Zhang, J. Yang and J. Zhang (2013). Combination of site-specific nitrogen management and alternate wetting and drying irrigation increases grain yield and nitrogen and water use efficiency in super rice. Field Crops Research **154**: 226-235.

Lou, Q., C. Ma, W. Wen, J. Zhou, L. Chen, F. Feng, X. Xu, X. Lu, L. Luo, H. Mei and G. Xu (2011). Profiling and association mapping of grain metabolites in a subset of the core collection of chinese rice germplasm (*Oryza sativa* L.). Journal of Agricultural and Food Chemistry **59**(17): 9257-9264.

Mackill, D., A. Ismail, U. Singh, R. Labios, T. Paris and L. Donald (2012). Chapter 6 - Development and rapid adoption of submergence-tolerant (Sub1) rice varieties. Advances in Agronomy, Academic Press. **115:** 299-352.

Mackill, D., Z. Zhang, E. Redoña and P. Colowit (1996). Level of polymorphism and genetic mapping of AFLP markers in rice. Genome **39**(5): 969-977.

Manzanilla, D., T. Paris, G. Vergara, A. Ismail, S. Pandey, R. Labios, G. Tatlonghari, R. Acda, T. Chi, K. Duoangsila, I. Siliphouthone, M. Manikmas and D. Mackill (2011)."Submergence risks and farmers' preferences: Implications for breeding Sub1 rice in Southeast Asia. Agricultural Systems **104**(4): 335-347.

Maraval, I., C. Mestres, K. Pernin, F. Ribeyre, R. Boulanger, E. Guichard and Z. Gunata (2008). Odor-active compounds in cooked rice cultivars from Camargue (France) analyzed by GC−O and GC−MS. Journal of Agricultural and Food Chemistry **56**(13): 5291-5298.

Mathure, S., N. Jawali, R. Thengane and A. Nadaf (2014). Comparative quantitative analysis of headspace volatiles and their association with BADH2 marker in non-basmati scented, basmati and non-scented rice (*Oryza sativa* L.) cultivars of India. Food Chemistry **142**: 383-391.

Mathure, S., K. Wakte, N. Jawali and A. Nadaf (2011). Quantification of 2-acetyl-1-pyrroline and other rice aroma volatiles among Indian scented rice cultivars by HS-SPME/GC-FID. Food Analytical Methods **4**(3): 326-333.

Matsuda, F., Y. Okazaki, A. Oikawa, M. Kusano, R. Nakabayashi, J. Kikuchi, J. Yonemaru, K. Ebana, M. Yano and K. Saito (2012). Dissection of genotype-phenotype associations in rice grains using metabolome quantitative trait loci analysis. Plant Journal **70**(4): 624-636.

McCouch, S. (2004). Diversifying selection in plant breeding. PLoS Biology **2**(10): e347.

McCouch, S., S. Temnykh, A. Lukashova, J. Coburn, G. DeClerck, S. Cartinhour, S. Harrington, M. Thomson, E. Septiningsih, M. Semon, P. Moncada and J. Li (2001). Microsatellite markers in rice: Abundance, diversity, and applications. Rice genetics IV. G. Khush, D. Brar and B. Hardy. International Rice Research Institute, Los Baños (Philippines), Science Publishers, Inc.**:** 117-136.

McNally, K., R. Bruskiewich, D. Mackill, C. Buell, J. Leach and H. Leung (2006). Sequencing multiple and diverse rice varieties. Connecting whole-genome variation with phenotypes. Plant Physiology **141**(1): 26-31.

McNally, K., K. Childs, R. Bohnert, R. Davidson, K. Zhao, V. Ulat, G. Zeller, R. Clark, D. Hoen, T. Bureau, R. Stokowski, D. Ballinger, K. Frazer, D. R. Cox, B. Padhukasahasram, C. Bustamante, D. Weigel, D. Mackill, R. Bruskiewich, G. Rätsch, C. Buell, H. Leung and J. Leach (2009). Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. Proceedings of the National Academy of Sciences **106**(30): 12273-12278.

Meilgaard, M., G. Civille and B. Carr (2007). Sensory evaluation techniques. Boca Raton, FL, USA, CRC Press.

Mir, R., P. Hiremath, O. Riera-Lizarazu and R. Varshney (2013). Evolving molecular marker technologies in plants: From RFLPs to GBS. Diagnostics in Plant Breeding. T. Lübberstedt and R. Varshney, Springer Netherlands: 229-247.

Mochida, K. and K. Shinozaki (2011). Advances in omics and bioinformatics tools for systems analyses of plant functions. Plant and Cell Physiology **52**(12): 2017-2038.

Mohanty, S. (2013). Trends in global rice production. Rice Today **13**: 44-45.

Morsy, M., L. Jouve, J. Hausman, L. Hoffmann and J. Stewart (2007). Alteration of oxidative and carbohydrate metabolism under abiotic stress in two rice (*Oryza sativa* L.) genotypes contrasting in chilling tolerance. Journal of Plant Physiology **164**(2): 157-167.

Paine, J., C. Shipton, S. Chaggar, R. Howells, M. Kennedy, G. Vernon, S. Wright, E. Hinchliffe, J. Adams, A. Silverstone and R. Drake (2005). Improving the nutritional value of Golden Rice through increased pro-vitamin A content. Nature Biotechnology 23(4): 482-487.

Pandey, S., H. Bhandari and B. Hardy (2007). Economic Costs of Drought and Rice Farmers' Coping Mechanisms: A Cross-country Comparative Analysis. Los Banos, Philippines, International Rice Research Institute.

Pantuwan, G., S. Fukai, M. Cooper, S. Rajatasereekul and J. O'Toole (2002). Yield response of rice (*Oryza sativa* L.) genotypes to drought under rainfed lowlands: 2. Selection of drought resistant genotypes. Field Crops Research **73**(2–3): 169-180.

Peng, S., J. Huang, J. Sheehy, R. Laza, R. Visperas, X. Zhong, G. Centeno, G. Khush and K. Cassman (2004). Rice yields decline with higher night temperature from global warming. Proceedings of the National Academy of Sciences **101**(27): 9971-9975.

Peng, S., G. Khush, P. Virk, Q. Tang and Y. Zou (2008). Progress in ideotype breeding to increase rice yield potential. Field Crops Research **108**(1): 32-38.

Peng, S., R. Laza, R. Visperas, A. Sanico, K. Cassman and G. Khush (2000). Grain yield of rice cultivars and lines developed in the Philippines since 1966. Crop Science **40**(2): 307–314.

Pingali, P. (2007). Westernization of Asian diets and the transformation of food systems: Implications for research and policy. Food Policy **32**(3): 281-298.

Ram, P., B. Singh, A. Singh, P. Ram, P. Singh, H. Singh, I. Boamfa, F. Harren, E. Santosa, M. Jackson, T. Setter, J. Reuss, L. Wade, V. Pal Singh and R. Singh (2002). Submergence tolerance in rainfed lowland rice: physiological basis and prospects for cultivar improvement through marker-aided breeding. Field Crops Research **76**(2–3): 131-152.

Redestig, H., M. Kusano, K. Ebana, M. Kobayashi, A. Oikawa, Y. Okazaki, F. Matsuda, M. Arita, N. Fujita and K. Saito (2011). Exploring molecular backgrounds of quality traits in rice

by predictive models based on high-coverage metabolomics. BMC Systems Biology **5**(176):1-11.

Redoña, E. and D. Mackill (1996). Mapping quantitative trait loci for seedling vigor in rice using RFLPs. Theoretical and Applied Genetics **92**(3-4): 395-402.

Rieseberg, L. (1998). Genetic mapping as a tool for studying speciation. Molecular systematics of plants II. D. Soltis, P. Soltis and J. Doyle, Springer US**:** 459-487.

Ritota, M., R. Gianferri, R. Bucci and E. Brosio (2008). Proton NMR relaxation study of swelling and gelatinisation process in rice starch–water samples. Food Chemistry **110**(1): 14-22.

Rosell, C. and C. Marco (2008). Rice. Gluten-free cereal products and beverages. E. Arendt and F. Bello. San Diego, Academic Press**:** 81.

Ryan, E., A. Heuberger, T. Weir, B. Barnett, C. Broeckling and J. Prenni (2011). Rice bran fermented with *Saccharomyces boulardii* generates novel metabolite profiles with bioactivity. Journal of Agricultural and Food Chemistry **59**(5): 1862-1870.

Sanchez, D., M. Siahpoosh, U. Roessner, M. Udvardi and J. Kopka (2008). Plant metabolomics reveals conserved and divergent metabolic responses to salinity. Physiologia Plantarum **132**(2): 209-219.

Sano, Y. (1984). Differential regulation of waxy gene expression in rice endosperm. Theoretical and Applied Genetics **68**(5): 467-473.

Sato, S., M. Arita, T. Soga, T. Nishioka and M. Tomita (2008). Time-resolved metabolomics reveals metabolic modulation in rice foliage. BMC Systems Biology **2**(51):1-13.

Seck, P., A. Diagne, S. Mohanty and M. Wopereis (2012). Crops that feed the world: Rice. Food Security **4**(1): 7-24.

Septiningsih, E., J. Prasetiyono, E. Lubis, T. Tai, T. Tjubaryat, S. Moeljopawiro and S. McCouch (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**(8): 1419-1432.

Sheehy, J. and P. Mitchell (2013). Designing rice for the 21st century. Los Baños (Philippines), International Rice Research Institute.

Shen, Y., H. Jiang, J. Jin, Z. Zhang, B. Xi, Y. He, G. Wang, C. Wang, L. Qian, X. Li, Q. Yu, H. Liu, D. Chen, J. Gao, H. Huang, T. Shi and Z. Yang (2004). Development of genome-wide DNA polymorphism database for map-based cloning of rice genes. Plant Physiology **135**(3): 1198-1205.

Shen, Y., L. Jin, P. Xiao, Y. Lu and J. Bao (2009). Total phenolics, flavonoids, antioxidant capacity in rice grain and their relations to grain color, size and weight. Journal of Cereal Science **49**(1): 106-111.

Shu, X., T. Frank, Q. Shu and K. Engel (2008). Metabolite profiling of germinating rice seeds. Journal of Agricultural and Food Chemistry **56**: 11612–11620.

Shulaev, V., D. Cortes, G. Miller and R. Mittler (2008). Metabolomics for plant stress response. Physiologia Plantarum **132**(2): 199-208.

Singh, S., D. Mackill and A. Ismail (2009). Responses of SUB1 rice introgression lines to submergence in the field: Yield and grain quality. Field Crops Research **113**(1): 12-23.

Sood, B. and E. Sidiq (1978). A rapid technique for scent determination in rice. Indian Journal of Genetics and Plant Breeding **38**: 268–271.

Spindel, J., M. Wright, C. Chen, J. Cobb, J. Gage, S. Harrington, M. Lorieux, N. Ahmadi and S. McCouch (2013). Bridging the genotyping gap: using genotyping by sequencing (GBS) to add high-density SNP markers and new value to traditional bi-parental mapping and breeding populations. Theoretical and Applied Genetics **126**(11): 2699-2716.

Suwannaporn, P. and A. Linnemann (2008). Rice-eating quality among consumers in different rice grain preference countries. Journal of Sensory Studies **23**(1): 1-13.

Swamy, Y. and K. Bhattacharya (1982). Breakage of rice during milling. 1. Types of cracked and immature grains. Journal of Food Science and Technology **19**: 106–111.

Tan, Y., Y. Xing, J. Li, S. Yu, C. Xu and Q. Zhang (2000). Genetic bases of appearance quality of rice grains in Shanyou 63, an elite rice hybrid. Theoretical and Applied Genetics **101**(5-6): 823-829.

Tang, S., G. Khush and B. Juliano (1991). Genetics of gel consistency in rice (*Oryza sativa* L.). Journal of Genetics **70**(2): 69-78.

Tarpley, L., A. Duran, T. Kebrom and L. Sumner (2005). Biomarker metabolites capturing the metabolite variance present in a rice plant developmental period. BMC Plant Biology **5**(1): 1-12.

Tester, M. and P. Langridge (2010). Breeding technologies to increase crop production in a changing world. Science **327**(5967): 818-822.

The3000ricegenomesproject (2014). The 3,000 rice genomes project. GigaScience **3**(1): 1-6.

Thomson, M. (2009). Marker-assisted breeding for Abiotic Stress Tolerance in Rice: Progress and Future Perspectives. International Symposium on Rice Research in the Era of Global Warming**:** 43-55.

Thomson, M., K. Zhao, M. Wright, K. McNally, J. Rey, C. Tung, A. Reynolds, B. Scheffler, G. Eizenga, A. McClung, H. Kim, A. Ismail, M. de Ocampo, C. Mojica, M. Reveche, C. Dilla-Ermita, R. Mauleon, H. Leung, C. Bustamante and S. McCouch (2012). High-throughput single nucleotide polymorphism genotyping for breeding applications in rice using the BeadXpress platform. Molecular Breeding **29**(4): 875-886.

Tran, N., V. Daygon, A. Resurreccion, R. Cuevas, H. Corpuz and M. Fitzgerald (2011). A single nucleotide polymorphism in the Waxy gene explains a significant component of gel consistency. Theoretical and Applied Genetics **123**(4): 519-525.

Vasconcelos, M., K. Datta, N. Oliva, M. Khalekuzzaman, L. Torrizo, S. Krishnan, M. Oliveira, F. Goto and S. Datta (2003). Enhanced iron and zinc accumulation in transgenic rice with the ferritin gene. Plant Science **164**(3): 371-378.

Vaughan, D., H. Morishima and K. Kadowaki (2003). Diversity in the Oryza genus. Current Opinion in Plant Biology **6**(2): 139-146.

Venuprasad, R., C. Dalid, M. Del Valle, D. Zhao, M. Espiritu, M. Sta Cruz, M. Amante, A. Kumar and G. Atlin (2009). Identification and characterization of large-effect quantitative trait loci for grain yield under lowland drought stress in rice using bulk-segregant analysis. Theoretical and Applied Genetics **120**(1): 177-190.

Vichapong, J., M. Sookserm, V. Srijesdaruk, P. Swatsitang and S. Srijaranai (2010). High performance liquid chromatographic analysis of phenolic compounds and their antioxidant activities in rice varieties. Lwt- Food Science and Technology **43**(9): 1325-1330.

Virk, P., B. Ford-Lloyd, M. Jackson and H. Newbury (1995). Use of RAPD for the study of diversity within plant germplasm collections. Heredity **74**: 170–179.

von Caemmerer, S., P. Quick and R. Furbank (2012). The development of C4 rice: Current progress and future challenges. Science **336**(6089): 1671-1672.

Wanchana, S., T. Toojinda, S. Tragoonrung and A. Vanavichit (2003). Duplicated coding sequence in the waxy allele of tropical glutinous rice (*Oryza sativa* L.). Plant Science **165**(6): 1193-1199.

Wang, Y., Y. Xue and J. Li (2005). Towards molecular breeding and improvement of rice in China. Trends in Plant Science **10**(12): 610-614.

Ward, R., Q. Gao, H. de Bruyn, R. Gilbert and M. Fitzgerald (2006). Improved methods for the structural analysis of the amylose-rich fraction from rice flour. Biomacromolecules **7**(3): 866-876.

Widjaja, R., J. Craske and M. Wootton (1996). Comparative studies on volatile components of non-fragrant and fragrant rices. Journal of the Science of Food and Agriculture **70**(2): 151-161.

Wilkinson, H. and E. Champagne (2004). Value-added rice products. Rice: Chemistry and Technology. E. Champagne. St. Paul, Minnesota, USA, American Association of Cereal Chemists, Inc.: 473-494.

Wu, F., N. Yang, H. Chen, Z. Jin and X. Xu (2011). Effect of germination on flavor volatiles of cooked brown rice. Cereal Chemistry **88**(5): 497-503.

Xiao, J., S. Grandillo, S. Ahn, S. McCouch and S. Tanksley (1996). Genes from wild rice improve yield. Nature **384**: 223-224.

Xie, F. and B. Hardy (2009). Accelerating hybrid rice development. Los Baños (Philippines), International Rice Research Institute.

Xu, K. and D. Mackill (1996). A major locus for submergence tolerance mapped on rice chromosome 9. Molecular Breeding **2**(3): 219-224.

Yamakawa, H. and M. Hakata (2010). Atlas of rice grain filling-related metabolism under high temperature: Joint analysis of metabolome and transcriptome demonstrated inhibition of starch accumulation and induction of amino acid accumulation. Plant and Cell Physiology **51**(5): 795-809.

Yang, D., K. Lee, O. Jeong, K. Kim and S. Kays (2007). Characterization of volatile aroma compounds in cooked black rice. Journal of Agricultural and Food Chemistry **56**(1): 235-240.

Yang, D., K. Lee and S. Kays (2010). Characterization and discrimination of premium-quality, waxy, and black-pigmented rice based on odor-active compounds. Journal of the Science of Food and Agriculture **90**(15): 2595-2601.

Yang, D., K. Lee, K. Kim and S. Kays (2008). Site of origin of volatile compounds in cooked rice. Cereal Chemistry **85**(5): 591-598.

Ye, C., M. Argayoso, E. Redoña, S. Sierra, M. Laza, C. Dilla, Y. Mo, M. Thomson, J. Chin, C. Delaviña, G. Diaz and J. Hernandez (2012). Mapping QTL for heat tolerance at flowering stage in rice using SNP markers. Plant Breeding **131**(1): 33-41.

Ye, C., S. Fukai, I. Godwin, R. Reinke, P. Snell, J. Schiller and J. Basnayake (2009). Cold tolerance in rice varieties at different growth stages. Crop and Pasture Science **60**(4): 328-338.

Yoshizawa, K. and Y. Ogawa (2004). Rice in brewing. Rice: Chemistry and Technology. E. Champagne. St. Paul, Minnesota, USA, American Association of Cereal Chemists, Inc.: 541-568.

Zeigler, R. and A. Barclay (2008). The relevance of rice. Rice **1**(1): 3-10.

Zeng, L., S. Lesch and C. Grieve (2003). Rice growth and yield respond to changes in water depth and salinity stress. Agricultural Water Management **59**(1): 67-75.

Zeng, Z., H. Zhang, T. Zhang, S. Tamogami and J. Chen (2009). Analysis of flavor volatiles of glutinous rice during cooking by combined gas chromatography–mass spectrometry with modified headspace solid-phase microextraction method. Journal of Food Composition and Analysis **22**(4): 347-353.

Zhang, Q. (2007). Strategies for developing Green Super Rice. Proceedings of the National Academy of Sciences **104**(42): 16402-16409.

Zhao, G., J. Liu, S. Yang, J. Yea, X. Liao, Z. Su, R. Shi, C. Jiang and L. Dai (2009). Effect of cold-water irrigation on grain quality traits in japonica rice varieties from Yunnan Province, China. Rice Science **16**(3): 201-209.

Zhao, K., C. Tung, G. Eizenga, M. Wright, M. Ali, A. Price, G. Norton, M. Islam, A. Reynolds, J. Mezey, A. McClung, C. Bustamante and S. McCouch (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. Nature Communications **2**: 467-477.

Zhao, K., M. Wright, J. Kimball, G. Eizenga, A. McClung, M. Kovach, W. Tyagi, M. Ali, C. Tung, A. Reynolds, C. Bustamante and S. McCouch (2010). Genomic diversity and introgression in *O. sativa* reveal the impact of domestication and breeding on the rice genome. PLoS One **5**(5): e10780.

Zhao, N., G. Wang, A. Norris, X. Chen and F. Chen (2013). Studying plant secondary metabolism in the age of genomics. Critical Reviews in Plant Sciences **32**(6): 369-382.

Zhao, X. and M. Fitzgerald (2013). Climate change: Implications for the yield of edible rice. PLoS One **8**(6): e66218.

Zhou, G., Y. Chen, W. Yao, C. Zhang, W. Xie, J. Hua, Y. Xing, J. Xiao and Q. Zhang (2012). Genetic composition of yield heterosis in an elite rice hybrid. Proceedings of the National Academy of Sciences **109**(39): 15847-15852

Zhu, J., M. Gale, S. Quarrie, M. Jackson and G. Bryan (1998). AFLP markers for the study of rice biodiversity. Theoretical and Applied Genetics **96**(5): 602-611.

Zuther, E., K. Koehl and J. Kopka (2007). Comparative metabolome analysis of the salt response in breeding cultivars of rice. Advances in Molecular Breeding Toward Drought and Salt Tolerant Crop. M. Jenks, P. Hasegawa and S. Jain. Berlin, Springer-Verlag**:** 285–315.

Chapter 2

Diversity of global rice markets and the science required for consumer-targeted rice breeding

M. Calingacion[1,2,3], A. Laborte[1], A. Nelson[1], A. Resurreccion[1,3], J. Concepcion[1,3,a], V. Daygon[1,3,a], R. Mumm[4,5], R. Reinke[3,6,b], S. Dipti[3,7], P. Bassinello[3,8], J. Manful[3,9], S. Sophany[3,10], K. Lara[3,11], J. Bao[3,12], L. Xie[3,13], K. Loaiza[3,14], A. El-hissewy[3,15], J. Gayin[3,16], N. Sharma[3,17], S. Rajeswari[3,18], S. Manonmani[3,18], N. Rani[3,19], S. Kota[3,19], S. Indrasari[3,20], F. Habibi[3,21], M. Hosseini[3,21], F. Tavasoli[3,21], K. Suzuki[3,22], T. Umemoto[3,23], C. Boualaphanh[3,24], H. Lee[3,25], Y. Hung[25], A. Ramli[3,26], P. Aung[3,27], R. Ahmad[3,28], J. Wattoo[3,29], E. Bandonill[3,30], M. Romero[3,30], C. Brites[3,31], R. Hafeel[3,32], H. Lur[3,33], K. Cheaupun[3,34], S. Jongdee[3,35], P. Blanco[3,36], R. Bryant[3,37], N. Lang[3,38], R. D. Hall[2,3,4,5], and M. Fitzgerald[1,3,a]

Published, PlosOne (2014) 9 (2): e85106; doi: 10.1371/journal.pone.0085106.

[1]International Rice Research Institute, Los Banos, Laguna, Philippines; [2]Laboratory of Plant Physiology, Wageningen University, Wageningen, The Netherlands; [3]International Network for Quality Rice; [4]Plant Research International, Wageningen, The Netherlands, [5]Centre for BioSystems Genomics, Wageningen, The Netherlands; [6]Yanco Agricultural Institute, NSW Department of Industry and Investment, PMB, Yanco, New South Wales, Australia; [7]Grain Quality and Nutrition Division, Bangladesh Rice Research Institute (BRRI), Gazipur, Bangladesh; [8]EMBRAPA Rice and Beans, Santo Antonio de Goia´s, GO, Brazil; [9]Africa Rice Center (AfricaRice), Cotonou, Republic of Benin, [10]Cambodian Agricultural Research and Development Institute, Phnom Penh, Cambodia; [11]Mejoramiento Genetico de Arroz INIA CRI Quilamapu, Vicente Mendez, Chile; [12]Institute of Nuclear Agricultural Sciences, Zhejiang University, Hua Jiachi Campus, Hangzhou, China; [13]China National Rice Research Institute, Hangzhou, China; [14]Laboratorio de Calidad FLAR-CIAT, CIAT, Cali-Palmira, Colombia; [15]Rice Research Section, Field Crops Research Institute, A.R.C., Rice Research and Training Center, Sakha, Kafr El-Shiekh, Egypt; [16]CSIR-Food Research Institute, Accra, Ghana; [17]Rice Section, Department of Plant Breeding and Genetics, Punjab Agricultural University Ludhiana, Ludhiana, India; [18]Department of Rice, Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University Coimbatore, Tamil Nadu, India; [19]Crop Improvement Section, Directorate of Rice Research, Rajendra Nagar, Hyderabad, AP, India; [20]Indonesian Center for Rice Research (ICRR) BB Padi, Sukamandi, Subang Jawa

Barat, Indonesia; [21]Rice Research Institute of Iran (RRII), Rasht, I.R. Iran; [22]NARO Institute of Crop Science, 2-1-18 Kannondai, Tsukuba, Ibaraki, Japan, [23]NARO Hokkaido Agricultural Research Centre, Toyohira, Sapporo, Hokkaido, Japan; [24]Rice and Cash Crop Research Institute, NAFRI, Vientiane, Lao PDR; [25]Faculty of Agriculture and Food Science, Universiti Putra Malaysia, Bintulu Sarawak, Malaysia; [26]Pusat Penyelidikan Padi dan Tanaman Industri, MARDI Seberang Perai Beg Berkunci, Seberang Perai Pulau Penang, Malaysia; [27]Plant Biotechnology Center, Myanmar Agriculture Service, Ministry of Agriculture and Irrigation, Yangon, Myanmar; [28]Rice Programme, National Agricultural Research Centre, Islamabad, Pakistan, [29]National Institute for Biotechnology and Genetic Engineering, Faislabad, Pakistan, [30]Rice Chemistry and Food Science Division, Philippine Rice Research Institute, Maligaya, Science City of Munoz, Nueva Ecija, Philippines; [31]Instituto Nacional de Investigacao Agraria e Veterinaria, Quinta do Marques, Oeiras, Portugal; [32]Rice Research Station, Department of Agriculture, Ambalantota, Sri Lanka, [33]Department of Agronomy, National Taiwan University, Taiwan, [34]Pathumthani Rice Research Centre, Bureau of Rice Research and Development, Thailand; [35]Khon Kaen Rice Research Center, Khon Kaen, Thailand; [36]Rice Research Program, National Agricultural Research Institute, INIA Treinta y Tres, Treinta y Tres, CP, Uruguay; [37]USDA-ARS, Dale Bumpers National Rice Research Center, Stuttgart, Arkansas, United States of America; [38]Genetic & Plant Breeding Division, Cuulong Delta Rice Research Inst., Can Tho, Viet Nam; [a](current address) School of Agriculture and Food Science, University of Queensland, St Lucia, Queensland, Australia; [b](current address) International Rice Research Institute, National Institute of Crop Science, Gwonsun-gu, Suwon, Republic of Korea.

Abstract

With the ever-increasing global demand for high quality rice in both local production regions and with Western consumers, we have a strong desire to understand better the importance of the different traits that make up the quality of the rice grain and obtain a full picture of rice quality demographics. Rice is by no means a 'one size fits all' crop. Regional preferences are not only striking, they drive the market and hence are of major economic importance in any rice breeding / improvement strategy. In this analysis we have engaged local experts across the world to perform a full assessment of all the major rice quality trait characteristics and importantly, to determine how these are combined in the most preferred varieties for each of their regions. Physical as well as biochemical characteristics have been monitored and this has resulted in the identification of no less than 18 quality trait combinations. This complexity immediately reveals the extent of the specificity of consumer preference. Nevertheless, further assessment of these combinations at the variety level reveals that several groups still comprise varieties which consumers can readily identify as being different. This emphasises the shortcomings in the current tools we have available to assess rice quality and raises the issue of how we might correct for this in the future. Only with additional tools and research will we be able to define directed strategies for rice breeding which are able to combine important agronomic features with the demands of local consumers for specific quality attributes and hence, design new, improved crop varieties which will be awarded success in the global market.

Introduction

Economic growth in Asia over the past thirty years has increased incomes and lifted millions out of poverty (Lee and Hong 2012). Globally confidence is growing that the first Millennium Development Goal (to cut extreme hunger and poverty by half) will be met, and that economic growth will continue, and lift millions more from poverty (FAO, WFP et al. 2012). It is projected that by 2050 Asia will account for half of the world's economic output (Government 2012), and be home to the largest proportion of middle classes, who will increasingly demand higher quality food, including higher quality rice (ADB 2012, ANZ 2012, Foundation 2012, Government 2012). However, coupled with economic growth in Asia, population pressure is increasing, agricultural land is being lost to urbanisation, and there is significant loss of workers from farms as the younger generation moves to cities (FAO, WFP et al. 2012). Furthermore, climate change is predicted to have a severe impact on agricultural production in Asian countries (Stern 2006).

Rice consumption per capita in many Asian countries is decreasing steadily due to changing dietary habits as a result of western influence, such as increased intake of dairy, meat and fast foods, and economic development (Pingali 2007). Despite this decline, current projections of population growth indicate that an additional 8 million tonnes of rice must be produced each year in Asia (Mohanty 2013). Furthermore, western countries have started eating rice regularly (Supakornchuwong and Suwannaporn 2012), and continued growth of Asian communities within western countries has increased the market size for rice as well as creating awareness and appreciation of rice-based meals (Suwannaporn and Linnemann 2008). Taken together, for a developing Asia these factors all indicate that resource use efficiency must increase significantly in future rice production, and must meet the dual requirements of increasing production with fewer resources, while meeting the market demands of increasingly discerning consumers.

Many countries still grow traditional varieties of rice such as Khao Dawk Mali 105, selected in 1958 in Thailand (Fitzgerald, McCouch et al. 2009), or popular, long-standing, improved varieties such as IR64 released in 1985 in the Philippines (Fitzgerald, McCouch et al. 2009) and Swarna released in 1979 in India. This is despite significant national investment in rice improvement programs and continual release of new, high-yielding varieties. Adoption of these improved varieties by rice farmers is conditional upon consumer acceptance of the sensory and cooking properties of the grain (Boualaphanh, Daygon et al. 2011). People prefer a specific type of rice for a number of reasons. A survey conducted in Cambodia revealed that different quality traits were prioritised by different actors in the rice value chain. These traits included grain shape and appearance, aroma, texture and lack of chalk (ACIAR CSE-2009-005 Rice Market Survey). Consumers are readily able to determine if the sensory properties of a new variety are acceptable. However when a new variety is not acceptable, it is difficult for consumers to describe why, and this makes it difficult for rice research programs to conduct the type of research that will deliver relevant selection tools to breeding programs.

Most of the studies conducted on rice consumer preferences have focused on the requirements of a particular country (Tomlins, Manful et al. 2005, Rachmat, Thahir et al. 2006, Azabagaoglu and Gaytancioglu 2009, Diako, Sakyi-Dawson et al. 2010, Anang, Adjetey et al. 2011, Musa, Othman et al. 2011, Abazari, Mojaverian et al. 2012, Minten, Murshid et al. 2013), but the rice market is now global, and there is a need for a more holistic perspective on consumer preferences relating to rice grain quality and its geographic

variability. This will allow for a more targeted approach towards developing and disseminating new rice varieties which have an increased probability of adoption and acceptance.

The establishment of the International Network for Quality Rice (INQR), made it possible to conduct a survey of consumer preferences for rice quality in different rice-consuming regions. The INQR serves as the platform for all experts working on rice grain quality to: exchange information; establish new and standardised protocols for measuring different quality parameters; and identify new traits of physical, sensory and nutritional quality (https://inqr.irri.org). Through the INQR, experts in 23 countries have participated in the survey reported here.

In the present paper we describe an extensive analysis of regional preferences for the main features of rice quality that are routinely measured in quality evaluation programs – length and shape of the grain, amylose content, gelatinisation temperature, gel consistency and aroma. Our objectives are to understand the specific market requirements for rice in each country in Asia as close to the provincial level as possible, and in some of Asia's export markets, and to identify gaps that investment and research must fill to increase the ability of rice breeders to develop high-yielding, climate-ready varieties with the quality traits required by the consumers in their local markets.

Methods

Data collection and analysis

A survey was conducted among the members of the INQR. They were asked to identify the three most popular varieties of rice from their respective countries, and where possible disaggregated to the provincial, regional or state level (Table S1). Information was supplied for the commonly measured quality traits of grain length and shape, amylose content, gel consistency, gelatinisation temperature and aroma. For countries unable to measure all traits, samples of the chosen varieties were sent to the International Rice Research Institute (IRRI) for phenotyping of those traits. The consumer preferences for grain quality traits presented here refer to the characteristics of the three most popular varieties identified per region (country, state or province). Although variations in preferences exist within a region, only the dominant preferences are captured here.

Data for each quality trait was mapped in units based on administrative level (e.g., whole country, state, province) using boundaries denoted in the Global Administrative Areas (GADM) database (www.gadm.org). There are a total of 92 spatial units for which we have data on preference for at least two grain quality traits (Figure S1).

Data from the survey were classified as previously defined (Khush, Paule et al. 1979, Juliano 1985). The classes of grain length of milled rice in Asian rice improvement programs are defined as short (<5.5 mm), medium (5.51-6.6 mm), long (6.61-7.5 mm) and extra-long (>7.51 mm) (Khush, Paule et al. 1979). These definitions differ from those described in the CODEX standard (198-1995) and by the European Commission (Commission 2003), both of which define long grains as being 6 mm or more. For the purposes of this paper, which focuses on Asian rice, we have retained the classes for grain dimensions defined by Khush et al. (1979), and used the classifications for all other traits that were described by Juliano et al. (1985). Grain shape (length/width) was classified as bold (<2), medium (2 – 3) or slender (>3). Amylose content was classified as waxy (~ 0%), low (2 – 19%), intermediate (20-25%) or high (>25%). Categories for gel consistency were soft (>60 mm), intermediate (40-60 mm) or hard (<40 mm). Gelatinisation temperature was classified as low (<70 °C), intermediate (70-74 °C) or high (>74 °C).

The potential market share of each country was estimated based on FAO data on rice consumption per capita over a period of 20 years (Table S2) from 1990-2009 (the most recent available year: (FAOSTAT 2013)). The countries in the INQR dataset account for over 90% of the global rice consumption, indicating that the data cover a substantial proportion of the global rice market.

Length and shape of the grain

At IRRI, the length and width of grains were measured using a Cervitec Grain Inspector 1625 (FOSS Analytical, Hoganas, Sweden). For grain lengths reported from elsewhere, length and width were measured using calipers or a calibrated scanner.

Amylose content

A. Colourimetric method

Amylose content was measured using the standard iodine colourimetric method ISO 6647-2-2011 (Standardization 2011). Briefly, ethanol (1mL, 95%) and 1M sodium hydroxide (9mL) were added to rice flour (100mg), and this was heated in a boiling water bath until

gelatinisation of the starch occurred. After cooling, 1M acetic acid (1mL) and iodine solution (2mL) were added and the volume was made up to 100 mL with Millipore water. The iodine solution was prepared by dissolving 0.2g iodine and 2.0g potassium iodide in 100mL Millipore water. Absorbance of the solution was measured using an Auto Analyser 3 (Bran+Luebbe, Norderstedt, Germany) at 600 nm. Amylose content was quantified from a standard curve generated from absorbance values of 4 well-known standard rice varieties (IR65, IR24, IR64 and IR8).

B. Size exclusion chromatography

Rice samples with low, intermediate and high amylose were selected and analysed using SEC exactly as previously described (Ward, Gao et al. 2006). Rice flour (50mg) was gelatinised, then debranched with isoamylase (Pseudomonas, Megazyme, Wicklow, Ireland) at 50 °C for 2h, with regular agitation. An aliquot of each debranched solution (40µL) was analysed using SEC (Waters, Alliance 2695, Waters, Milford, USA) equipped with an Ultrahydrogel 250 column (Waters).

C. Genotyping the Waxy gene

Genomic DNA was extracted from rice leaves of the low amylose varieties from Cambodia, Thailand, Australia, and Japan using the method described previously (Virk, Ford-Lloyd et al. 1995). Leaf samples were frozen in liquid nitrogen and ground into fine powder prior to addition of extraction buffer. PCR amplification was performed using a G-Storm Thermal Cycler (model GS1, Gene Technologies Ltd, Essex, UK) in a 20-µL reaction volume containing forward and reverse primers (Table 1), designed with Primer 3 software (Rozen and Skaletsky 1999), and components from KAPA HiFi HotStart PCR Reagent Kit (KAPA Biosystems, Boston, Massachusetts, USA). Twenty microliters of the PCR product were electrophoresed through a 1.2% agarose gel, stained with SybrSafe nucleic acid stain (Invitrogen, Carlsbad, CA, USA), and visualized using a non-ultraviolet transilluminator (Dark Reader DR195M, Clare Chemicals, Dolores, CO, USA). PCR fragments from agarose gels were purified using the QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany) according to the manufacturer's instruction. Sequencing of the PCR products was done by Macrogen Inc., Seoul, South Korea. Nucleotide sequences were retrieved and aligned using BIOEDIT software (Hall 1999) with the Nipponbare rice *Wx* gene sequence from GRAMENE (Ware, Jaiswal et al. 2002) as the reference.

Gelatinisation temperature

Gelatinisation temperature was measured using a DSC Q100 instrument (TA Instrument, New Castle, DE, USA) exactly as previously described (Cuevas, Daygon et al. 2010). Rice flour (4mg) was mixed with Millipore water (8mg) in hermetic aluminium pans which were then sealed. The pans were heated under pressure from 25 to 120°C at 10 °C min^{-1}. Some countries reported gelatinisation temperature as alkali spreading values, which were determined as previously described (Little, Hilder et al. 1958), and values ascribed to high, intermediate or low by the correlations reported previously (Cuevas, Daygon et al. 2010).

Table 1. Primers used to sequence Wx gene of low AAC varieties.

Primer Name	Sequence (5'-3')	Size (bp)
Wx_ex1F	CAATGCAACGTACGCCAAG	19
Wx_ex1R	CCTGGGTGTGTTTCTCTAGACTC	23
Wx_ex2aF	GTGGGCTAGCTGACCTAGATTTG	23
Wx_ex2aR	TGTTTAAGGTTTGGTGAGCCTA	22
Wx_ex2bF	CCAAGAAACTGCTCCTTAAGTCC	23
Wx_ex2bR	GTACCTGTCTGCAACCTTGATCT	23
Wx_ex3-6F	GCATTGGATGGATGTGTAATGT	22
Wx_ex3-6R	GGCTGGTAGTTGTTCTTCAGGT	22
Wx_ex6-9F	GGAAGCATCACGAGTTTACCAT	22
Wx_ex6-9R	CTTGCCTTGTCAGAATCAATCA	22
Wx_ex10-11F	CAACCACGGTAAGAACGAATG	21
Wx_ex10-11R	AGGGCTGGAGAAATCAACAAG	21
Wx_ex12-13F	CTGCAGGGGATGAGATACG	19
Wx_ex12-13R	TGTAGATCTCAGGCTCTTCAAGG	23
Wx_ex14F	TGTTTGCAACATGGATTTCAGGG	23
Wx_ex14R	TCCTGAGTCAAACTACTGCTCCT	23

Gel consistency

Gel consistency was determined exactly as previously described (Cagampang, Perez et al. 1973). Rice flour (100 mg) was mixed with ethyl alcohol (0.2mL) containing 0.025% thymol blue and 0.2M potassium hydroxide (2mL) and heated in a boiling water bath for 8 min. After heating, the sample tubes were allowed to cool in an ice-water bath and immediately laid horizontally on the table. Gel consistency was measured by the length of the cold rice paste in the culture tube held horizontally for one hour. Hard, medium and soft gel standards, IR48, PSBRC9 and IR42, respectively, were included in every set.

Aroma

The current definition of aromatic rice is the presence of the volatile compound 2-acetyl-1-pyrroline (2AP). This was quantified at IRRI using gas chromatograph (Agilent 6890N, Santa Clara, CA, USA) equipped with a flame ionisation detector (Fitzgerald, Sackville Hamilton et al. 2008). For those rice samples not measured at IRRI, aroma was determined by smelling and tasting cooked grains.

Volatile analysis of aromatic rice by gas chromatography- mass spectrometry (GCMS)

Volatile compounds in the aromatic rice from Iran, Pakistan, India and the Greater Mekong Sub-region (GMS) were analysed. Headspace volatile compounds of selected aromatic rice were collected by solid phase microextraction using a 65-mm polydimethylsiloxane-divinylbenzene fibre (Supelco, Bellefonte, USA) and analysed using GC-MS (GC 8000, Fisons Instruments, Cheshire, UK) (Calingacion, Boualaphanh et al. 2012). GCMS raw data were processed using MetAlign (Lommen and Kools 2012) to extract and align the mass signals, and MSClust (Tikunov, Laptenok et al. 2012) to remove signal redundancy per metabolite and reconstruct mass spectra. The PCA plot was constructed using SIMCA-P 12.0 (Umetrics AB, Umea°, Sweden).

Results and Discussion

Without an understanding of consumer preference for rice grain quality, wide adoption of any newly developed rice variety is not guaranteed. Hence, identifying the grain traits that govern acceptance is important to guide a successful breeding program. Quality attributes of the most popular rice varieties consumed in the countries and provinces of Asia, as well as for some of the rice- growing countries in other continents have been collected.

Currently, rice grain quality is classified in terms of the physical, cooking and sensory characteristics. The physical appearance of the grain defines its price in the market, whereas the cooking and sensory properties determine the reputation of the variety (Fitzgerald 2010). Grain appearance, the first thing a consumer sees, is defined by length, width and shape (ratio of length and width); and chalk, which consumers dislike (Zhao and Fitzgerald 2013) unless the variety is targeted towards use in paella or risotto. The physical traits of the grain are immediately recognisable even to an untrained eye. Cooking quality is a measure of the time and fuel required to cook the rice, and this is indirectly measured by gelatinisation temperature (Cuevas, Daygon et al. 2010). Sensory properties are influenced by amylose content (Fitzgerald, Bergman et al. 2009), gel consistency (Cagampang, Perez et al. 1973, Tran, Daygon et al. 2011) and also gelatinisation temperature (Umemoto, Horibata et al. 2008).

Length and shape of the grain

The defined classes of grain length differ between Asian and European standards (Khush, Paule et al. 1979, Commission 2003), but for the purposes of this work, the Asian classification of length, as defined by Khush, et al. (1979) was used. Rice consumers in parts of South East Asia - Thailand, Lao PDR, Cambodia, Malaysia, and Philippines prefer long and slender grains (Figure 1). Consumers in Indonesia and Bangladesh prefer grains that are medium in length and slender. In North Asia, the Japanese, Taiwanese and South Koreans eat short and bold rice grains, whereas in several states of India, and Sri Lanka both short, medium grains are popular. In most of Pakistan and the Indian states of Punjab and Haryana, the extra long grains characteristic of basmati rice are popular, and long grains are also popular in most of Iran and parts of Pakistan (Figure 1).

Rice consumers in the large countries of China, and India have regional differences in preferences for grain length and shape. Consumers in China prefer rice grains that are either short or bold, or are long in size but with the grain shape ranging from medium to slender depending on province. Indian rice preferences, on the other hand, are composed of an even more varied set of grain types- from short to medium, long and extra-long in length with shapes ranging from bold to slender (Figure 1).

Understanding the genetic basis of length should assist in harmonising the definitions of grain length between the European and Asian standards. Seven quantitative trait loci (QTLs) for grain length have been reported and three for grain width (Huang, Jiang et al.

2012). Biallelic variation at the *GS3* locus appears to exert the most control over grain length (Fan, Xing et al. 2006). All the short and medium grains genotyped carry the C-allele and all the long and extra-long grains carry the A allele (Fan, Yu et al. 2009, Takano-Kai, Jiang et al. 2009). However the association studies done for the gene *GS3* were all on paddy rice, and not on polished rice (Fan, Xing et al. 2006, Fan, Yu et al. 2009, Takano-Kai, Jiang et al. 2009), so the range of white grain length of those carrying the C and A allele is not yet known, and this information would enable the differences between the standards to be resolved.

Figure 2a shows four clear groups for grain length. Koshihikari from Japan and Pandan Wangi from Indonesia, both of which are likely to be in the grain length range for allele C of *GS3*, differ in grain length and do so reliably. Furthermore, the basmati rices around 8.0 mm and KDML105 at 6.8 mm both carry the A-allele of *GS3* (Fan, Yu et al. 2009). Another gene associated with grain length, *GL3*, has recently been cloned (Zhang, Wang et al. 2012), but its presence has not been associated widely with grain length in diverse germplasm. We therefore have four reliable phenotypes of grain length (Figure 2a), and two genotypes at the *GS3* loci, but as other QTLs and genes such as *GL3* are further understood, the complete genetic regulation of grain length should become known.

Grain shape is often used to describe the physical dimensions of grains. This is the ratio of length and width, and three classes are defined: bold (<2), medium (2.1 – 3), and slender (>3). Confusingly, these terms could also describe width. Figure 2b shows the percentage of each grain shape category for each length class for all the varieties in the present paper. Interestingly, as grain length increases, the proportion of slender grains increases and the proportion of bold grains decreases. However, Fitzgerald et al. (2009) showed a grain with length and width 11 and 3.8 mm respectively, which places an extra-long grain into the medium shape classification. This suggests that selection has driven the trend seen in Figure 2b, but the diversity of rice could encompass all grain shapes for all length classes. If the rice community intends to continue using grain shape to describe the dimensions of the grain, it would be useful to also always include the length in order to avoid confusion.

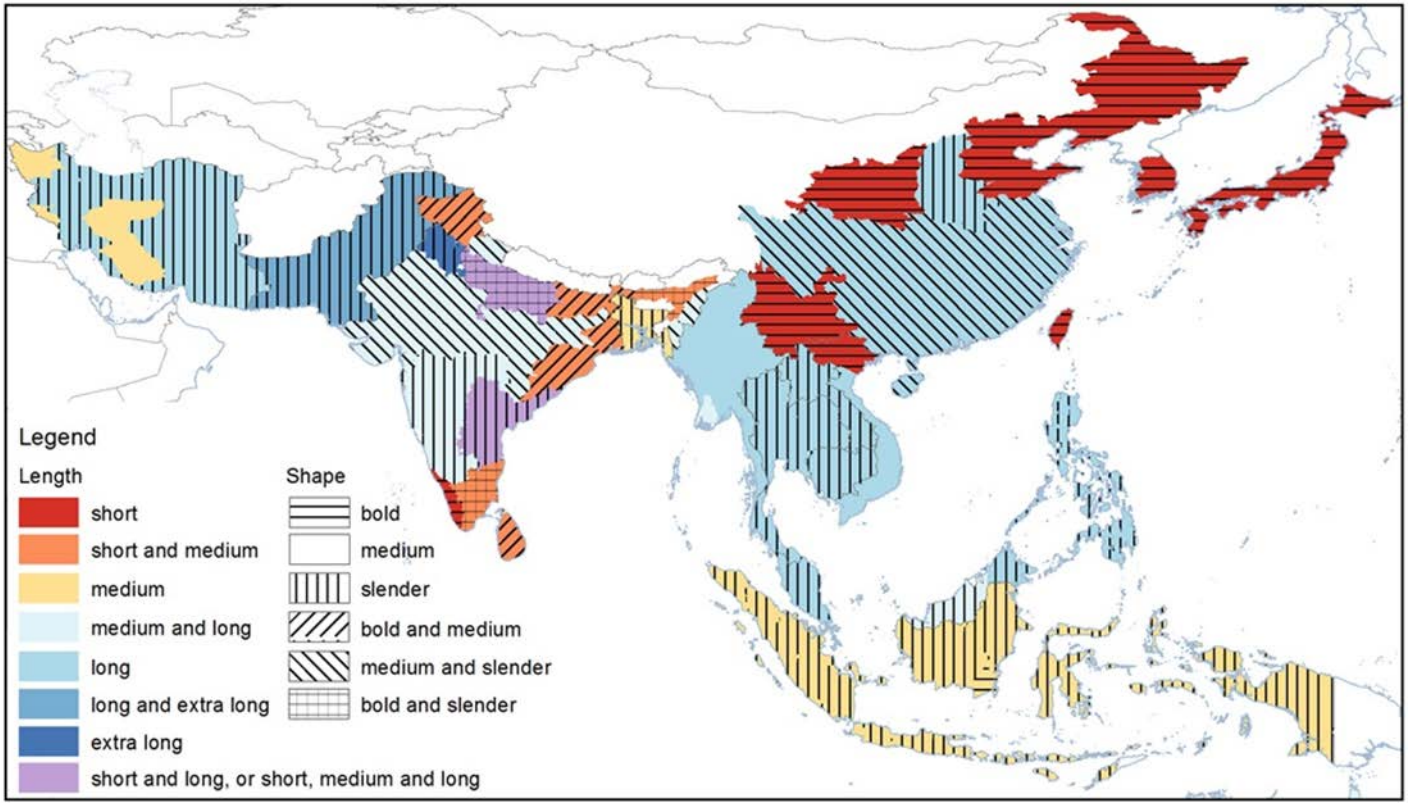


Figure 1. Regional variation in rice length and shape (length/width) of the three most popular varieties in the countries, states, and provinces of Asia. In some regions, more than one type of grain lengths and shapes are preferred. Colours represent length, and lines represent the shape. Additional information for other regions can be found in Table 2. Data were obtained from INQR representatives from each region.

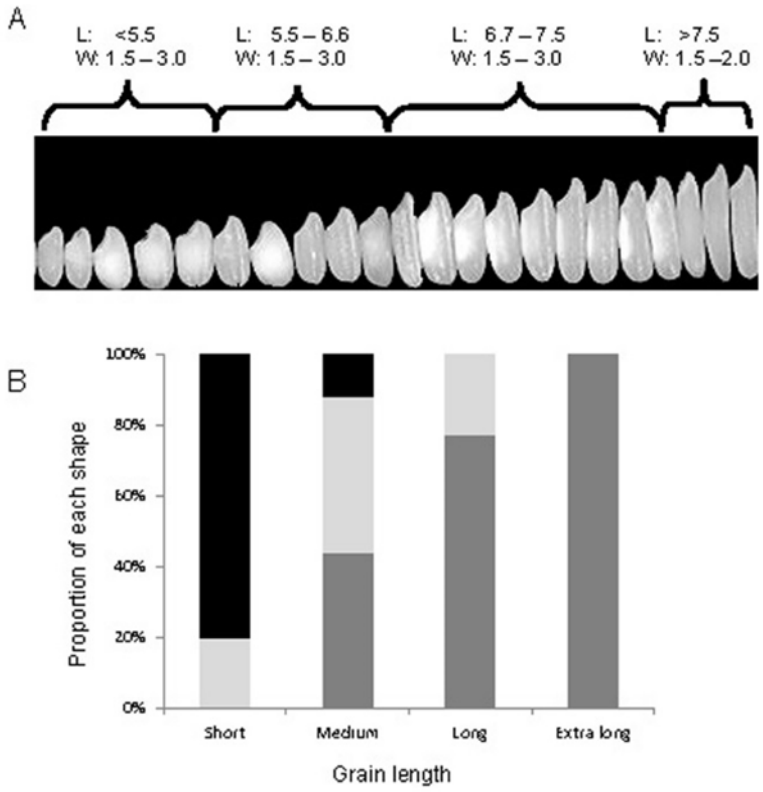


Figure 2: A: Grains ordered from shortest to longest, showing different widths (mm) in each class except the extra-long class. B: Histogram showing the proportion of bold (black), medium (light grey) and slender (dark grey) grain shapes within each length class for all the popular rices discussed in the present paper.

Amylose content

Amylose, a linear polymer of glucose units linked primarily by α-1,4 linkages, influences texture and the potential of cooked grains to retrograde after cooking. This is also one of the major traits used in the selection process for eating quality among rice breeding programs (Fitzgerald, McCouch et al. 2009). Hence, most of the INQR members were familiar with the amylose classes popular in their region.

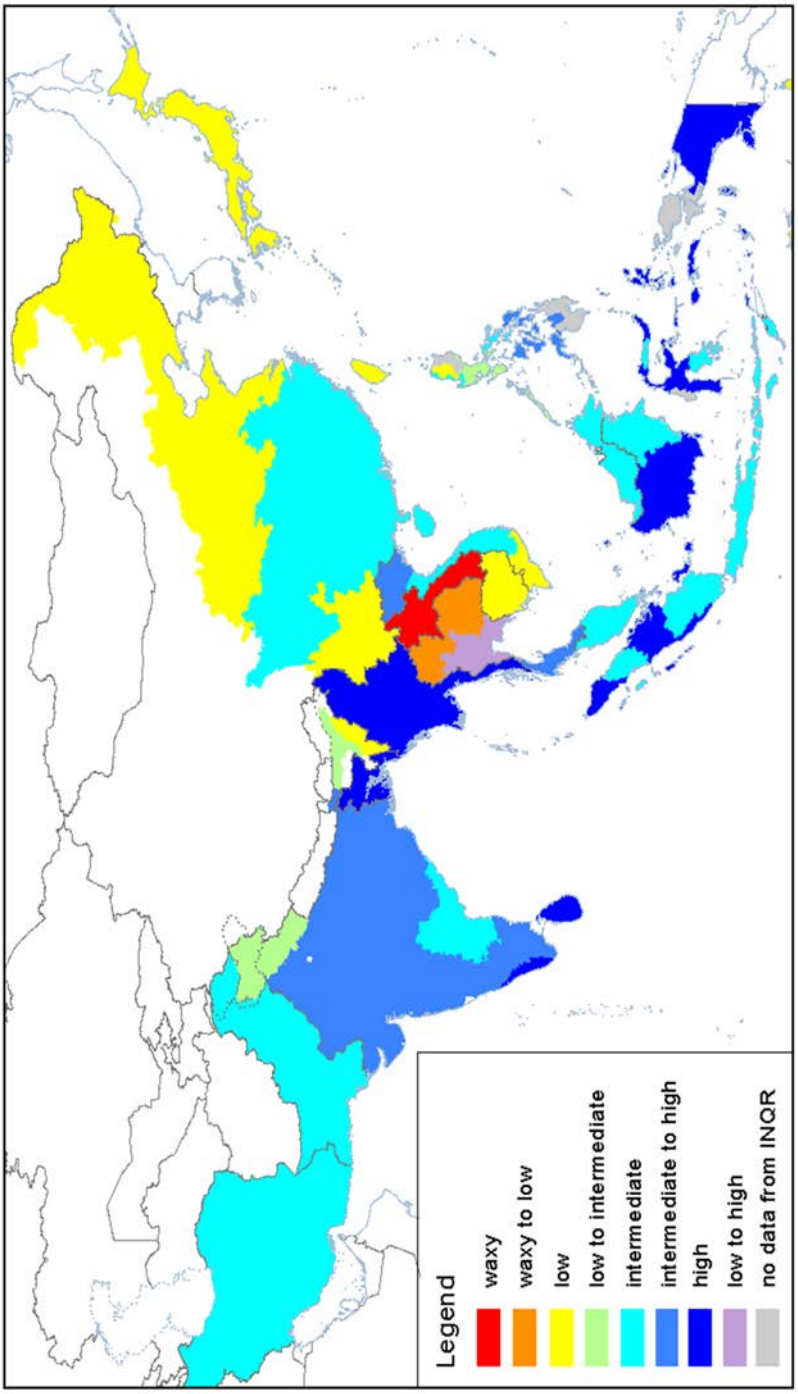


Figure 3. Regional variation in amylose content of the three most popular varieties in the countries, states, and provinces of Asia. In some regions, two types of amylose class are preferred. Additional information for other regions can be found in Table 2. Data were obtained from INQR representatives from each region

Consumers in Lao PDR and the Isan region of Thailand prefer waxy or sticky rice (Figure 3). In Japan, Taiwan, Cambodia, Thailand, parts of Lao PDR, Egypt and Australia, consumers prefer low amylose rice, as do consumers in the northern and south-western provinces of China and southern Vietnam (Figure 3, Table 2). Rice with intermediate amylose content is preferred in Iran, Pakistan, Malaysia, Philippines, many states in India and some provinces of China, Vietnam, Indonesia and Uruguay. High amylose varieties are popular in Myanmar, Sri Lanka, provinces of Indonesia, and many states of India (Figure 3). High AC rice is also preferred in Ghana, Senegal, Suriname, Colombia and parts of Uruguay (Table 2).

The amylose classes are associated with polymorphisms in the Waxy gene (Wx) (Sano 1984) which encodes for the granule-bound starch synthase (GBSSI) enzyme that is responsible for amylose synthesis (Smith, Denyer et al. 1997). A G-T polymorphism at the 5' splice site at intron 1 results in two functional alleles Wxa and Wxb, which differentiate low amylose from high and intermediate classes (Ayres, McClung et al. 1997). An A-C polymorphism in exon 6, Wxin discriminates intermediate from high amylose (Chen, Bergman et al. 2008, Mikami, Uwatoko et al. 2008). The waxy phenotype is a null mutation with a 23-bp duplication in exon 2 which causes a frame-shift resulting in non-functional GBSSI protein (Wanchana, Toojinda et al. 2003) and no amylose (Fitzgerald, Bergman et al. 2009, Cuevas and Fitzgerald 2012).

Each class of amylose is a range of 4 – 12 percentage points. Figure 3 has been plotted on the basis of the traditionally-defined classes. However, data from different countries indicated that popular rice varieties span each range. For example, the low amylose varieties from Thailand and Cambodia are 12-15% amylose, whereas the low amylose varieties from Japan, China, Korea and Australia are about 18-19% amylose. These varieties are all the same Waxy haplotype, with the single-nucleotide polymorphism (SNP) at intron 1. The same variation within class was found for the intermediate and high amylose classes.

Table 2: Grain quality traits in non-Asian rice-growing countries. Some countries do not measure all traits.

Country	Length	Shape	Amylose	Aroma	Gel consistency	Gel temperature
Australia	Medium and long	Medium	Low	1 of 3		Low and intermediate
Egypt	Long	Slender and medium	Low			
Ghana	Medium and long	Slender	High	1 of 3	Soft and intermediate	Low and intermediate
Uganda	Medium and long	Slender and medium	Intermediate and high	1 of 3	Soft and intermediate	Intermediate and high
Senegal	Medium and long	Slender and medium	High		Soft	Intermediate and high
Portugal	Medium and long	Slender and medium	Intermediate			Intermediate and high
Suriname	Extra long	Slender	High			low
Chile	Short and long	Bold	Intermediate			low
Colombia	Long	Slender	High			Low and intermediate
Brazil	Long		Intermediate and high		Intermediate	Intermediate and high
Uruguay	Long		High			Low and intermediate
USA	Medium and long	Slender and medium	Low and intermediate			Low and intermediate

In order to visualise the amylose content more clearly, varieties with amylose content across the range of each class were analysed using size exclusion chromatography (SEC). The SEC traces clearly demonstrate that within a typical amylose class, different levels of amylose and amylopectin were observed (Figure 4). For the low amylose class there was a clear group of lower amylose varieties, which were the Thai and Cambodian indica varieties, and a clear group with higher amylose which are the Australian and Japanese temperate japonica varieties. However, the range in the other two classes is due to a difference in the amount of amylopectin chains (Figure 4). Those at the upper end of the intermediate and high amylose class had a significantly lower percentage of amylopectin than those at the lower end of the amylose class (Figure 4). Since only the low amylose haplotype showed differences in the amylose chains, the coding region of the *Wx* gene was sequenced for each, to determine the presence of another allele. However, no sequence differences were found, suggesting either that the difference in amylose is due to loci other than *Wx*, interactions relating to amylose that are specific to either the indica or tropical japonica germplasm class, or that the amylose content is lower when this haplotype is grown under tropical conditions. This allele of the *Wx* gene leading to low amylose is sensitive to high temperature (Larkin and Park 2003). However, it is possible that main climate difference between the countries could be humidity, although no relationship between amylose content and humidity has yet been demonstrated. For intermediate and high amylose, the data indicate that amylopectin synthesis is also involved in amylose content, and presents an opportunity to discover the biochemical and genetic bases of varying amylopectin content.

The range in amylose content of varieties within each class suggests that an additional sub-categorisation within each amylose class would be valuable to define clearly the differences within rice that belong to the same amylose haplotype. Low amylose rice could be further classified into subclass A and B with amylose content 11-15% and 16-19%, respectively. Similarly, the intermediate amylose group may be divided into subclass A with amylose contents between 20-22%, and subclass B with 23-24% amylose. High amylose can also be categorised into subclass A and B with amylose content 25-27% and >27%, respectively. Although the underlying genetics for each sub-class has not been elucidated, it would be valuable to adopt these descriptors to facilitate understanding in the global rice literature.

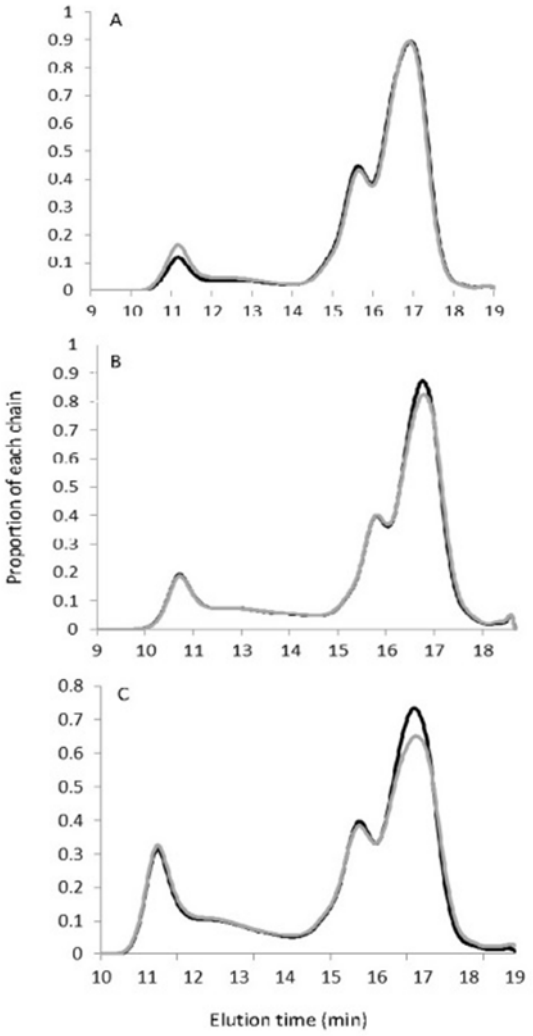


Figure 4. Size exclusion chromatograms of low (A), intermediate (B) and high (C) amylose varieties. Each curve is the average of ten different rice varieties from each end of the range in each class. In each class, those with high amylose are shown by the grey curves, and those with lower amylose are shown by the black curve. Chains of amylose elute before 14 min and amylopectin chains after 14 min.

Gel consistency

Gel consistency is a test that was developed to differentiate between the texture of cooked rice from the high and intermediate amylose classes (Cagampang, Perez et al. 1973). It measures the distance travelled by a gel after cooking and was developed for varieties of 24-30% amylose (Cagampang, Perez et al. 1973). Gel consistency is not measured in a number of countries because high amylose rices are not part of the rice improvement program.

It was expected that only the rice improvement programs developing high and intermediate amylose would be using a gel consistency test, such as Iran, Pakistan, Thailand, Philippines, Tamil Nadu in India, Sarawak in Malaysia and Kalimantan in Indonesia and some others (Figure 3). However other countries also reported values for gel consistency (Figure 5, Table 1), even those where low amylose varieties are developed. Most countries that reported gel consistency values prefer soft or intermediate gel consistency, which indicates soft-texture in varieties with high amylose. The genetic basis of gel consistency is a SNP in exon 10 of the Waxy gene (Tran, Daygon et al. 2011). This SNP has only been associated with gel consistency for varieties with a functional intron 1 (Tran, Daygon et al. 2011), which excludes the lower amylose varieties.

Gelatinisation temperature

Gelatinisation temperature describes the temperature that starch granules begin to melt, and the grains begin to cook (Slade and Levine 1989). It ranges from 55 – 85 °C in domesticated rice, and correlates with the cooking time of rice and the final cooked texture (Cuevas, Daygon et al. 2010). In the countries of South and Central Asia, gelatinisation temperature is intermediate to high, which means the rice takes about 4 min longer to cook than rice with low gelatinisation temperature, such as is found in some Southeast Asian countries (Figure 6).

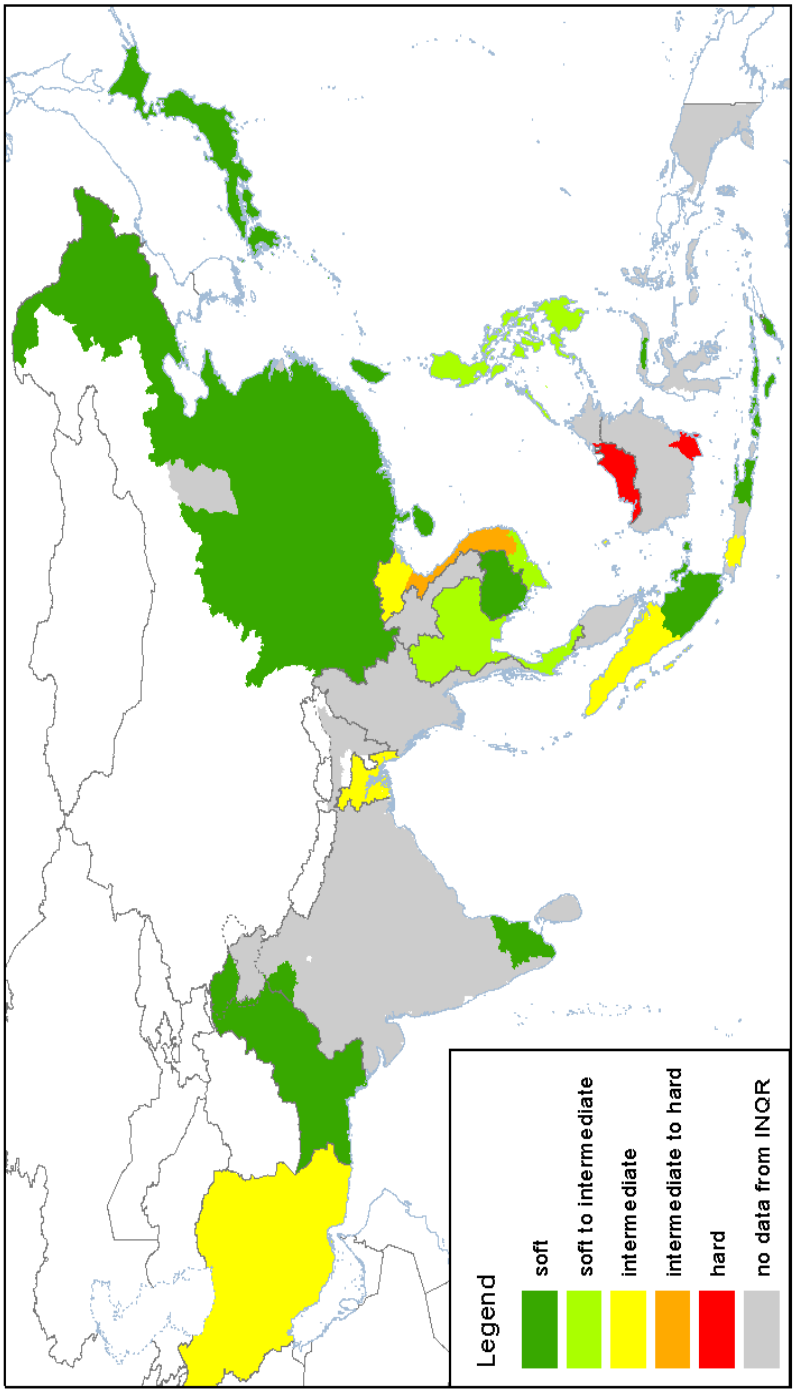


Figure 5. Consumer preferences for texture based on gel consistency values. In many countries and regions, gel consistency is not measured (grey). Additional information for the other regions can be found in Table 2. Data obtained from INQR representatives from each region.

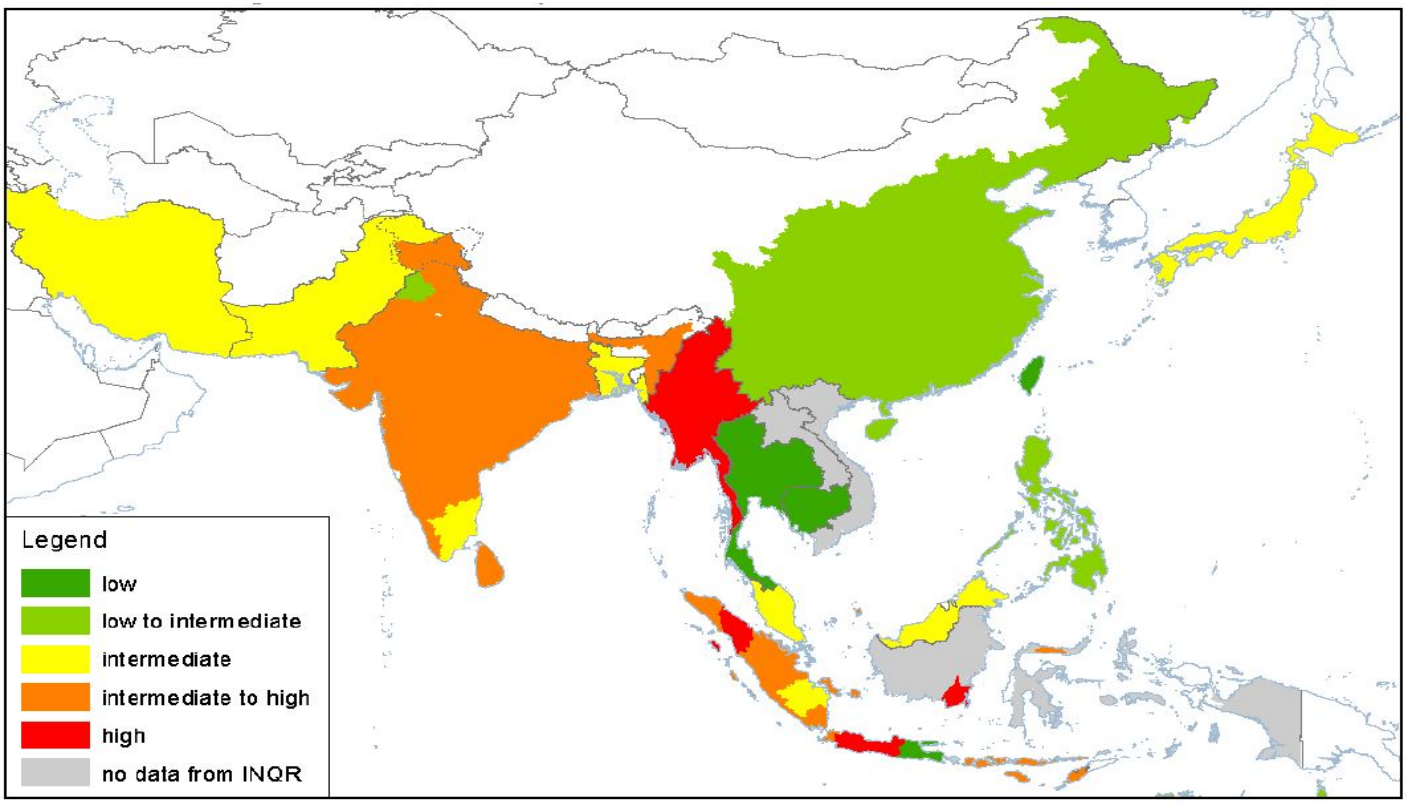


Figure 6. Regional variation in gelatinisation temperature of the three most popular varieties in the countreis, states, and provinces of Asia. In some regions, two classes are preferred. Additional information for the other regions can be found in Table 2. Data obtained from INQR representatives from each region.

Gelatinisation temperature is divided into two groups by allelic variation in SSIIa (Umemoto, Yano et al. 2002, Bao, Corke et al. 2006, Waters, Henry et al. 2006). The SNPs in SSIIa define four haplotypes (Umemoto and Aoki 2005, Waters, Henry et al. 2006) and two haplotypes associate with high and two with low gelatinisation temperature (Cuevas, Daygon et al. 2010). Intermediate gelatinisation temperature varieties are found in all haplotype groups (Cuevas, Daygon et al. 2010), suggesting that another locus interacts with SSIIa to produce the intermediate phenotype. Many of the popular varieties have been genotyped for the SSIIa locus (Cuevas, Daygon et al. 2010), and the genotype associates with the reported phenotypes (Figure 6) in all cases.

The additional cooking time required for rice of high gelatinisation temperature leads to an enormous amount of extra energy used when expressed on a population basis. It has been estimated that each minute less cooking time globally represents 2500 years cooking time saved per day (Fitzgerald, McCouch et al. 2009). Thus if all countries could lower the gelatinisation time, and thereby cooking time of their popular varieties, this would lead to significant savings of fuel, and a measureable reduction in the carbon footprint of rice.

Aroma

Aromatic rice varieties are of great interest in the market because they command a higher price than non- aromatic rice. The two types of aromatic rice are basmati and jasmine. Aromatic rice is characterised by the presence of a popcorn / baked bread-like flavour compound called 2-acetyl-1-pyrroline (2AP). Although generally only present in very low amounts, its low odour threshold means that 2AP is easily detected by consumers (Buttery, Ling et al. 1983).

In both jasmine and basmati rice breeding programs, the presence of 2AP is used to select fragrant progeny. 2AP is determined subjectively either by sniffing a gram of rice grains that have been soaked for an hour in potassium hydroxide, quantitatively by gas chromatography, or by detecting the mutation in the gene that associates with aroma (Fitzgerald, Sackville Hamilton et al. 2008).

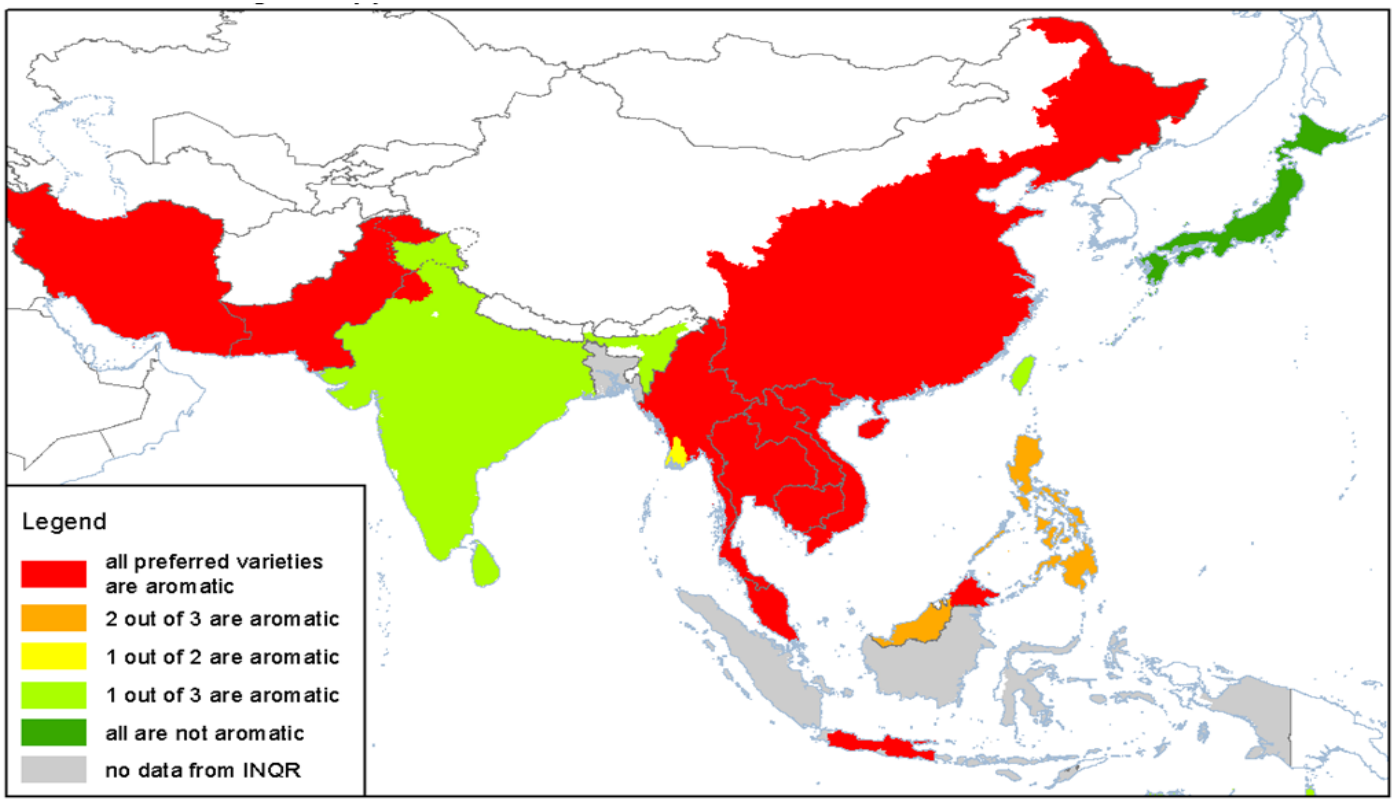


Figure 7. Consumer preferences for aromatic rice for the countries, states, and provinces of Asia. In some regions, all popular rices are aromatic, in others, one or two is aromatic, and in some countries, aromatic rice is not popular. Additional information for other regions can be found in Table 2. Data were obtained from INQR representatives from each region.

Figure 7 shows that aromatic rice is very important in the Greater Mekong Subregion (GMS), Malaysia, Iran, Pakistan, the states of Punjab and Haryana of India and some provinces of Indonesia. In Iran, the aromatic varieties are sadri types, which group together genetically with the basmati rices of Pakistan and India (Courtois, Frouin et al. 2012). Most of those varieties are traditional, and the improved varieties grown in Pakistan and India are derived from traditional basmati parents. The fragrant rices in the GMS are indica jasmine types, and these are mostly traditional varieties. The taste and aroma of South and Central Asian and the GMS jasmine rices are distinctly different, although all have 2AP and all have been shown to carry the common mutation in the fragrance gene (Fitzgerald, Sackville Hamilton et al. 2008).

The fragrance of rice is a function of the volatile compounds emitted from the grains. Deep analysis of a selection of the sadri rices from Iran, the basmati rices from Pakistan, Punjab and the jasmine rices from the GMS shows that there are distinct differences in the volatile profile between the three germplasm classes (Figure 8). Moreover, in a sensory analysis previously carried out using the same popular varieties discussed here, the sadri rices were characterised as sweet dairy, the Thai indica rices as popcorn and the basmati as haylike (Champagne, Bett-Garber et al. 2010), consistent with the different volatile signatures observed in Figure 8.

Rice fragrance is almost always defined based on the presence of 2AP. However, Figure 8, showing the metabolomic signatures of the varieties in this paper, and other studies working with jasmine and basmati types (Limpawattana and Shewfelt 2010, Bryant and McClung 2011) had shown that other compounds and fragrance descriptors also contribute to the aroma of jasmine and basmati rices. Given the importance of these three classes of rice to Iran, Pakistan and to some states of India, and the countries in the GMS, it seems that improving yields in those regions requires a deeper understanding of the important compounds of fragrance for each type, and research to deliver robust phenotyping tools that are able to distinguish between the three types of aroma.

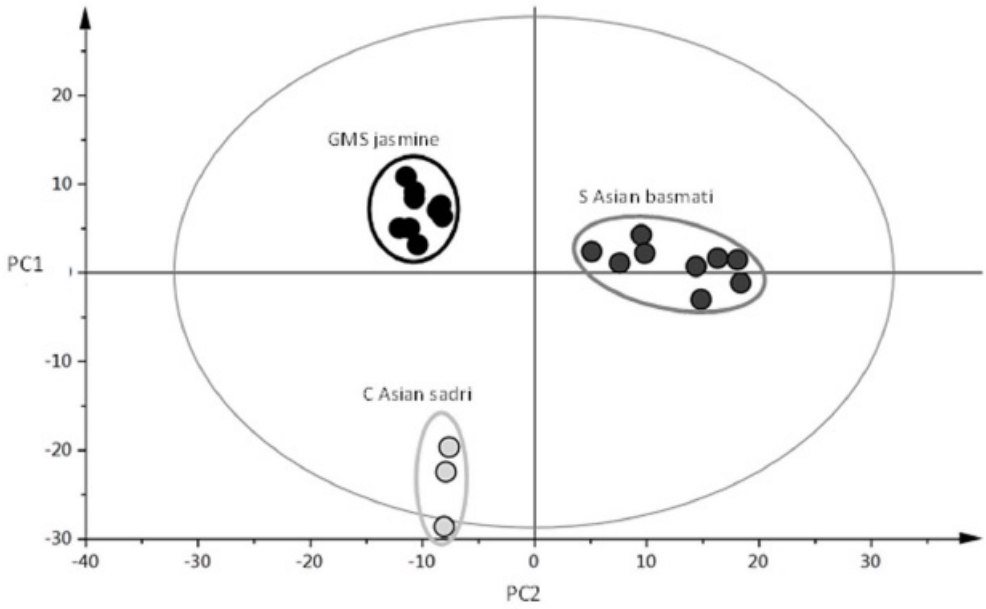


Figure 8: Principal Components Analysis of the volatile metabolomic signature of the traditional indica varieties from the Greater Mekong Subregion (GMS), basmati varieties from South Asia, and the sadri varieties from Iran. PC1 explains 27% and PC2 explains 21% of variation.

Grain quality combinations

Using the information for the quality traits we currently measure, the number of different combinations of quality were determined (Figure 9). Immediately, the complexity of local preference becomes evident. Based on the combinations, there were 18 grain types found, spanning Asia, Australia, and the main rice producing countries of Europe, Africa and the Americas. For short grain, there are just two types of quality; for medium length grain, there are seven types, with one type being aromatic, represented by Pandan Wangi from Indonesia and Paw San Hmwe from Myanmar. For the extra long grains, there is just one type, and these are the basmati rices from Pakistan and the northwestern states of India. The highest number of combinations was found in the long grain class (Figure 9), and five of the eight combinations are aromatic.

Figure 9 shows that consumer preferences from several countries are described by exactly the same combinations of grain quality. Some of these make sense, such as the long, slender, waxy, aromatic varieties from Isan in Thailand and Lao PDR since these two regions spent many years together in the ancient Kingdom of Siam. However, other common combinations cannot be explained by common heritage. For example, BRS Primavera from Brazil has exactly the same combinations of quality as IR64 from the Philippines (Figure 9). However, a previous study undertaken by the INQR included these two varieties and found that their taste, flavour and texture differ significantly (Champagne, Bett-Garber et al. 2010). There are many such examples in Figure 9 that can be compared in the same way. This indicates that our suite of phenotyping tools for grain quality is insufficient to distinguish fully the properties that consumers use to make decisions.

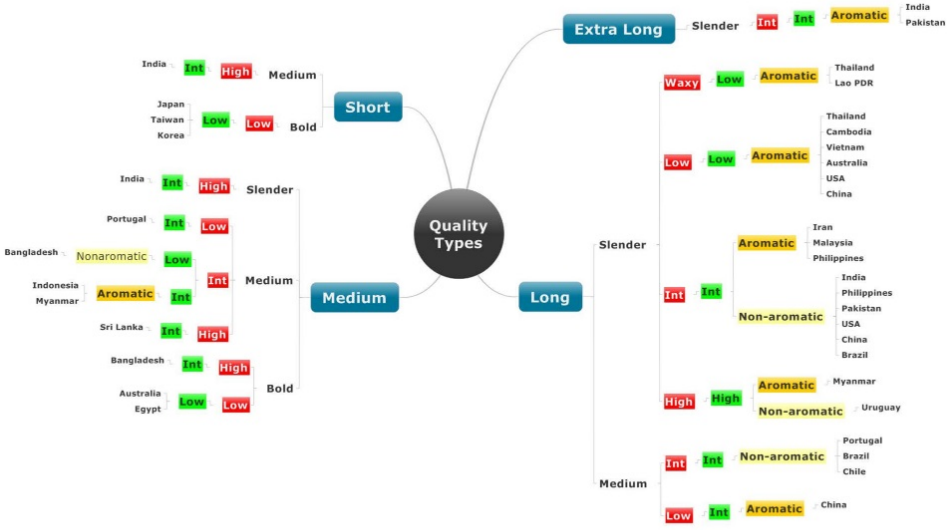


Figure 9: Number of different quality types in each grain length class based on combinations of the current tools for measuring quality: grain length (blue), shape, amylose content (red), gelatinisation temperature (green), and the presence of aroma.

Conclusions

By mapping the different quality traits for each country, the INQR has shown that there are at least 18 different quality types of rice that are favoured around the world. The two most popular combinations, that are found all over the world, are both long and slender, while one type has low amylose, low gelatinisation temperature and is aromatic, and the other has intermediate amylose and gelatinisation temperature and is not aromatic. Our previous study showed that the same combinations of measured quality traits are still different when eaten by consumers (Champagne, Bett-Garber et al. 2010). This strongly indicates that quality evaluation teams are not able to provide sufficient information to enable breeders to select for market quality, and suggests that new science needs to be brought to rice quality to develop new trait combinations of aroma, flavour and texture that, match market demands. This paper presents a framework for classifying the global variation in rice grain quality, which can be further enumerated as required, and reinforces the need for improved objective measurement within each quality category to capitalise on the considerable investment in variety development and to meet better the needs of a growing population of discerning consumers demanding high quality rices.

Acknowledgements

The authors thank the International Rice Research Institute (IRRI) and the International Network for Quality Rice (INQR) for funding the survey of consumer preference. MC also thanks the Monsanto Beachell-Borlaug International Scholarship Program. RDH acknowledges additional funding from the Netherlands Genomics Initiative. AL and AN thank the Japan Rice Breeding Program and the CGIAR Global Rice Science Partnership (GRiSP).

Supporting information

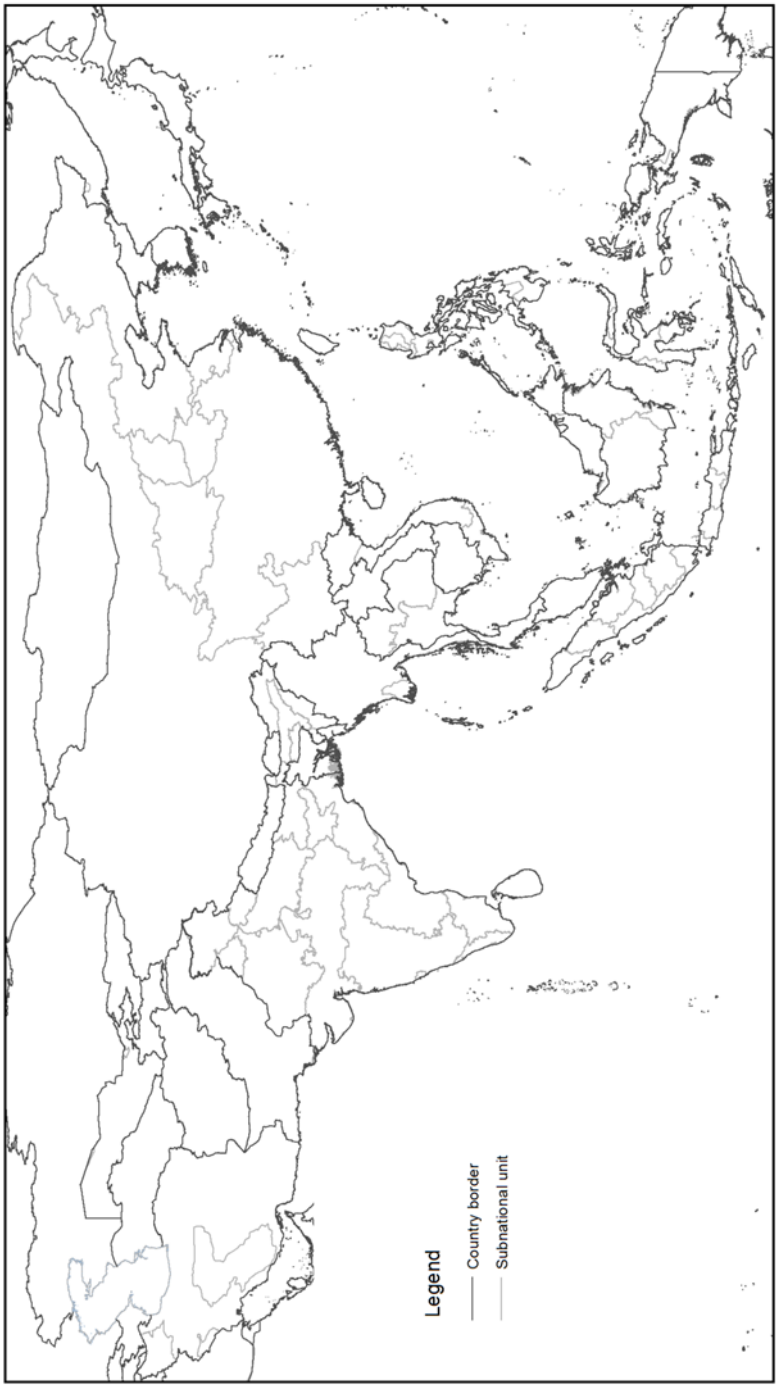


Figure S1. Spatial units depicting level of detail of data on preferred rice traits.

Table S1 Varieties nominated by each country as most popular.

Country	Nominated varieties	Country	Nominated varieties
Punjab (North India)	Pusa basmati 1121	Vietnam	Jasmine 85
	Basmati 386		OM 4900
	Punjab basmati 2		OM 7347
Central (India)	Taroari basmati	Iran	Hashemi
	Pusa 1121		Shirodi
	BPT 5204		Tarom
Tamil Nadu (India)	ADT 43		Fajv
	ADT 45	Cambodia	Phka malis
	Improved white ponni		Phka romduol
Andra Pradesh (India)	Sambha masuri		Phka kngnei
	Swarna	Egypt	Sakha 105
Pakistan	Super basmati		Sakha 106
	Basmati 515		Sakha 107
	Basmati 2000	Ghana	Wakatsuki
China	Zhongzheyou1		Sikamo
	Zhongjian 2		Nerica 1
	Xiangwanxian 13	Uganda	Jinja
	Ningjing 43		Supa
Bangladesh	BR11		Pearl Kenya pishori
	BRRI Dhan 28		Kaisho
	BRRI Dhan 29		Nerica
Indonesia	Ciherang	Senegal East	Sahel 177
	Pandan wangi		Sahel 328
	Rojolele		Nerica S19
Japan	Koshihikari		Nerica S21
	Hitomebore		ITA 150
	Hinohikari		WAB 56-50
Lao PDR	KDML 105	Australia	Amaroo
	Homesaven		Langi
Malaysia	MRQ 74		Reiziq
	MR 219	Portugal	Ariete
	MR 220		Albatross

Table S1 continued

Country	Nominated varieties	Country	Nominated varieties
Malaysia - Sarawak	Bario	Suriname	ADRON 125
	Bajong		ADRON 128
	Biris		ADRON 130
Myanmar	Paw San Hmwe	Uruguay	INIA Tacuarir
	Manawthukha		INIA Olimar
	Aye Yarr Min		El Paso 144
Philippines	Sinandomeng	Chile	Diamante INIA
	Dinorado		Zafiro INIA
	IR64		Oro
Sri Lanka	Bg 300	Colombia	Oryzica 1
	Bg 352		Fedearroz 50
	Bg 358		Fedearroz 60
Taiwan	Koshihikari	Brazil	BRS Primavera
	Taikeng 9		Epagri 108
	Tainung 71		IRGA 417
Thailand	KDML 105	USA	Wells
	PTT 1		CL 151
	RD 6		Jupiter

Table S2: Rice consumption per capita per country from 2004 – 2009 (FAOStat 2013).

Geographic entity			*kg per capita per year (5 year averages)*				millions of tonnes of rice per year (5 year averages)				As share of global total			
Country	Region	Continent	1990-1994	1995-2000	2000-2004	2004-2009	1990-1994	1995-2000	2000-2004	2004-2009	1990-1994	1995-2000	2000-2004	2004-2009
China	Eastern Asia	Asia	76.7	78.5	77.4	76.2	92.16	99.34	101.79	102.96	32.5%	32.1%	31.1%	29.7%
India	Southern Asia	Asia	75.6	73.6	69.2	70.0	68.78	73.62	75.34	82.18	24.2%	23.8%	23.0%	23.7%
Indonesia	South-Eastern Asia	Asia	127.1	130.7	125.1	125.7	24.20	26.80	27.40	29.22	8.5%	8.7%	8.4%	8.4%
Bangladesh	Southern Asia	Asia	158.7	157.3	173.3	171.1	17.47	19.26	23.26	24.63	6.2%	6.2%	7.1%	7.1%
Viet Nam	South-Eastern Asia	Asia	133.5	141.8	144.9	143.2	9.34	10.79	11.67	12.17	3.3%	3.5%	3.6%	3.5%
Philippines	South-Eastern Asia	Asia	89.3	96.8	108.4	125.1	5.77	7.01	8.75	11.09	2.0%	2.3%	2.7%	3.2%
Thailand	South-Eastern Asia	Asia	99.9	104.7	111.7	123.2	5.81	6.39	7.22	8.35	2.0%	2.1%	2.2%	2.4%
Japan	Eastern Asia	Asia	64.2	61.7	58.3	55.7	7.90	7.71	7.35	7.05	2.8%	2.5%	2.2%	2.0%
Myanmar	South-Eastern Asia	Asia	160.1	165.0	157.6	143.3	6.48	7.15	7.18	6.72	2.3%	2.3%	2.2%	1.9%
Brazil	South America	Americas	39.9	37.3	36.5	35.0	6.17	6.23	6.54	6.64	2.2%	2.0%	2.0%	1.9%
Sth Korea	Eastern Asia	Asia	98.1	94.5	81.3	78.3	4.29	4.27	3.77	3.72	1.5%	1.4%	1.2%	1.1%
Egypt	Northern Africa	Africa	34.3	39.9	40.8	38.3	2.02	2.56	2.87	2.95	0.7%	0.8%	0.9%	0.9%
USA	Northern America	Americas	7.2	8.5	8.5	8.4	1.88	2.31	2.46	2.54	0.7%	0.7%	0.8%	0.7%
Pakistan	Southern Asia	Asia	14.6	13.7	14.8	15.4	1.73	1.84	2.22	2.54	0.6%	0.6%	0.7%	0.7%
Cambodia	South-Eastern Asia	Asia	165.4	158.3	159.0	159.1	1.69	1.85	2.04	2.17	0.6%	0.6%	0.6%	0.6%
Malaysia	South-Eastern Asia	Asia	84.8	87.1	76.9	74.5	1.63	1.90	1.88	2.02	0.6%	0.6%	0.6%	0.6%
Sri Lanka	Southern Asia	Asia	90.4	91.3	92.6	99.2	1.60	1.68	1.77	2.01	0.6%	0.5%	0.5%	0.6%
Iran	Southern Asia	Asia	34.5	34.9	31.6	28.1	1.97	2.16	2.13	2.00	0.7%	0.7%	0.7%	0.6%
Colombia	South America	Americas	30.6	29.9	34.9	32.5	1.05	1.13	1.43	1.44	0.4%	0.4%	0.4%	0.4%
Lao PDR	South-Eastern Asia	Asia	164.7	163.6	161.8	163.5	0.73	0.82	0.89	0.97	0.3%	0.3%	0.3%	0.3%
Senegal	Western Africa	Africa	62.0	66.5	73.0	73.1	0.48	0.59	0.73	0.84	0.2%	0.2%	0.2%	0.2%
Ghana	Western Africa	Africa	16.2	10.7	21.5	25.1	0.25	0.19	0.43	0.57	0.1%	0.1%	0.1%	0.2%
Australia	Australia and New Zealand	Oceania	6.4	7.7	9.6	10.6	0.11	0.14	0.19	0.22	0.0%	0.0%	0.1%	0.1%
Portugal	Southern Europe	Europe	15.0	15.5	17.0	15.9	0.15	0.16	0.18	0.17	0.1%	0.1%	0.1%	0.0%
Chile	South America	Americas	7.4	6.8	7.3	9.0	0.10	0.10	0.12	0.15	0.0%	0.0%	0.0%	0.0%
Uganda	Eastern Africa	Africa	2.2	3.5	4.4	4.6	0.04	0.08	0.11	0.14	0.0%	0.0%	0.0%	0.0%
Uruguay	South America	Americas	10.2	11.0	15.9	18.5	0.03	0.04	0.05	0.06	0.0%	0.0%	0.0%	0.0%
Suriname	South America	Americas	83.2	79.8	71.2	67.1	0.03	0.04	0.03	0.03	0.0%	0.0%	0.0%	0.0%
Summary of countries in the INQR dataset							***263.88***	***286.16***	***299.81***	***315.58***	***93.0%***	***92.5%***	***91.7%***	***90.9%***

References

Abazari, A., S. Mojaverian and S. Hosseini-Yekani (2012). Effective factors on rice consumers' preferences: A case study of urban consumers in Mazandaran province of Iran. American Journal of Scientific Research (53): 70-79.

ADB (2012). Food security and poverty in Asia Pacific. Key Challenges and Policy Issues. Philippines, Asian Development Bank.

Anang, B., S. Adjetey and S. Abiriwe (2011). Consumer preferences for rice quality characteristics and the effects on price in the Tamale metropolis, northern region, Ghana. International Journal of AgriScience **1**(2): 67-74.

ANZ (2012). Greener pastures: The global soft commodity opportunity for Australia and New Zealand. ANZ Insights. **3:** 1-80.

Azabagaoglu, M. and O. Gaytancioglu (2009). Analyzing consumer preferences to different rice varieties in Turkey. Agricultura Tropica Et Subtropica **42**(3): 118-125.

Boualaphanh, C., V. Daygon, M. Calingacion, J. Sanitchon, D. Jothityangkoon, R. Mumm, R. Hall and M. Fitzgerald (2011). Use of new generation single nucleotide polymorphism genotyping for rapid development of near-isogenic lines in rice. Crop Science **51**(5): 2067-2073.

Cagampang, G., C. Perez and B. Juliano (1973). A gel consistency test for eating quality in rice. Journal of the Science of Food and Agriculture **24**: 1589-1594.

Calingacion, M., C. Boualaphanh, V. Daygon, R. Anacleto, R. Sackville Hamilton, B. Biais, C. Deborde, M. Maucourt, A. Moing, R. Mumm, R. de Vos, A. Erban, J. Kopka, T. Hansen, K. Laursen, J. Schjoerring, R. Hall and M. Fitzgerald (2012). A genomics and multi-platform metabolomics approach to identify new traits of rice quality in traditional and improved varieties. Metabolomics **8**(5): 771-783.

Champagne, E., K. Bett-Garber, M. Fitzgerald, C. Grimm, J. Lea, K. Ohtsubo, S. Jongdee, L. Xie, P. Bassinello, A. Resurreccion, R. Ahmad, F. Habibi and R. Reinke (2010). Important sensory properties differentiating premium rice varieties. Rice **3**: 270–281.

Commission, E. (2003). Council Regulation (EC) No 1785/2003 on the common organisation of the market in rice. Official Journal of the European Union: 1-18.

Cuevas, R., V. Daygon, H. Corpuz, L. Nora, R. Reinke, D. Waters and M. Fitzgerald (2010). Melting the secrets of gelatinisation temperature in rice. Functional Plant Biology **37**(5): 439-447.

Diako, C., E. Sakyi-Dawson, B. Bediako-Amoa, F. Saalia and J. Manful (2010). Consumer perceptions, knowledge and preferences for aromatic rice types in Ghana.

Fan, C., Y. Xing, H. Mao, T. Lu, B. Han, C. Xu, X. Li and Q. Zhang (2006). GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice encodes a putative transmembrane protein. Theoretical and Applied Genetics **112**(6): 1164-1171.

Fan, C., S. Yu, C. Wang and Y. Xing (2009). A causal C–A mutation in the second exon of GS3 highly associated with rice grain length and validated as a functional marker. Theoretical and Applied Genetics **118**(3): 465-472.

FAO, WFP and IFAD (2012). The State of Food Insecurity in the World 2012. Economic growth is necessary but not sufficient to accelerate reduction of hunger and malnutrition. Rome FAO.

FAOSTAT (2013).

Fitzgerald, M. (2010). Cereal grains Rice: Characteristics and quality requirements. Woodhead Publishing Series in Food Science, Technology and Nutrition**:** 212-236.

Fitzgerald, M., C. Bergman, A. Resurreccion, J. Möller, R. Jimenez, R. Reinke, M. Martin, P. Blanco, F. Molina, M. Chen, V. Kuri, M. Romero, F. Habibi, T. Umemoto, S. Jongdee, E. Graterol, K. Reddy, P. Bassinello, R. Sivakami, N. Rani, S. Das, Y. Wang, S. Indrasari, A. Ramli, R. Ahmad, S. Dipti, L. Xie, N. Lang, P. Singh, D. Toro, F. Tavasoli and C. Mestres (2009). Addressing the dilemmas of measuring amylose in rice. Cereal Chemistry **86**(5): 492-498.

Fitzgerald, M., S. McCouch and R. Hall (2009). Not just a grain of rice: the quest for quality. Trends in Plant Science **14**(3): 133-139.

Fitzgerald, M., N. Sackville Hamilton, M. Calingacion, H. Verhoeven and V. Butardo (2008). "Is there a second fragrance gene in rice?" Plant Biotechnology Journal **6**(4): 416-423.

Foundation, G. (2012). Feeding Asia and the World Australia has the potential to be a global food superpower. from http://www.globalfoundation.org.au/en/program-activities.html#20121018feeding.

Government, A. (2012). Australia in the Asian Century.

Hall, T. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acid Symposium Series no. 41, Oxford University Press.

Huang, R., L. Jiang, J. Zheng, T. Wang, H. Wang, Y. Huang and Z. Hong (2012). Genetic bases of rice grain shape: so many genes, so little known. Trends in Plant Science: 1-9.

Juliano, B. (1985). Criteria and Tests for Rice Grain Qualities. Rice: Chemistry and Technology. St. Paul, MN, American Association of Cereal Chemists, Inc,**:** 43-524.

Khush, G., C. Paule and N. de la Cruz (1979). Rice grain quality evaluation and improvement at IRRI. Workshop on Chemical Aspects of Rice Grain Quality, Los Banos, International Rice Research Institute.

Larkin, P. and W. Park (2003). Association of *waxy* gene single nucleotide polymorphisms with starch characteristics in rice (*Oryza sativa* L.). Molecular Breeding **12**(4): 335–339.

Lee, J. and K. Hong (2012). Economic growth in Asia: Determinants and prospects. Japan and the World Economy **24**(2): 101-113.

Little, R., G. Hilder and E. Dawson (1958). Differential effect of dilute alkali on 25 varieties of milled white rice. Cereal Chemistry **35**(111-126).

Lommen, A. and H. Kools (2012). MetAlign 3.0: performance enhancement by efficient use of advances in computer hardware. Metabolomics **8**(4): 719-726.

Minten, B., K. Murshid and T. Reardon (2013). Food quality changes and implications: Evidence from the rice value chain of Bangladesh. World Development **42**(1): 100-113.

Mohanty, S. (2013). Trends in global rice production. Rice Today **13**: 44-45.

Musa, M., N. Othman and F. Fatah (2011). Determinants of consumers purchasing behavior for rice in Malaysia. American International Journal of Contemporary Research **1** (3): 159-167.

Pingali, P. (2007). Westernization of Asian diets and the transformation of food systems: Implications for research and policy. Food Policy **32**(3): 281-298.

Rachmat, R., R. Thahir and M. Gummert (2006). The empirical relationship between price and quality of rice at market level in West Java. Indonesian Journal of Agricultural Science **7**(1): 27-33.

Rozen, S. and H. Skaletsky (1999). Primer3 on the WWW for general users and for biologist programmers. Bioinformatics methods and protocols. S. Misener and S. Krawetz, Humana Press**:** 365-386.

Standardization, I. O. f. (2011). ISO/DIS 6647-2: 2011 Rice - Determination of amylose content Part 2: Routine methods. 10.

Stern, N. (2006). The Economics of Climate Change, The Stern Review. B. G. Office of Climate Change, Cambridge**:** 1-700.

Supakornchuwong, C. and P. Suwannaporn (2012). Attitudes toward rice compared with potatoes and pasta among british, french, dutch and belgian consumers. Journal of Sensory Studies **27**(2): 71-77.

Suwannaporn, P. and A. Linnemann (2008). Rice-eating quality among consumers in different rice grain preference countries. Journal of Sensory Studies **23**(1): 1-13.

Takano-Kai, N., H. Jiang, T. Kubo, M. Sweeney, T. Matsumoto, H. Kanamori, B. Padhukasahasram, C. Bustamante, A. Yoshimura, K. Doi and S. McCouch (2009). Evolutionary history of GS3, a gene conferring grain length in rice. Genetics **182**(4): 1323-1334.

Tikunov, Y., S. Laptenok, R. Hall, A. Bovy and R. de Vos (2012). MSClust: a tool for unsupervised mass spectra extraction of chromatography-mass spectrometry ion-wise aligned data. Metabolomics **8**(4): 714-718.

Tomlins, K., J. Manful, P. Larwer and L. Hammond (2005). Urban consumer preferences and sensory evaluation of locally produced and imported rice in West Africa. Food Quality and Preference **16**(1): 79-89.

Tran, N., V. Daygon, A. Resurreccion, R. Cuevas, H. Corpuz and M. Fitzgerald (2011). A single nucleotide polymorphism in the Waxy gene explains a significant component of gel consistency. Theoretical and Applied Genetics **123**(4): 519-525.

Umemoto, T., T. Horibata, N. Aoki, M. Hiratsuka, M. Yano and N. Inouchi (2008). Effects of variations in starch synthase on starch properties and eating quality of rice. Plant Production Science **11**(4): 472–480.

Virk, P., B. Ford-Lloyd, M. Jackson and H. Newbury (1995). Use of RAPD for the study of diversity within plant germplasm collections. Heredity **74**: 170–179.

Ward, R., Q. Gao, H. de Bruyn, R. Gilbert and M. Fitzgerald (2006). Improved methods for the structural analysis of the amylose-rich fraction from rice flour. Biomacromolecules **7**(3): 866-876.

Ware, D., P. Jaiswal, J. Ni, I. Yap, X. Pan, K. Clark, L. Teytelman, S. Schmidt, W. Zhao, K. Chang, S. Cartinhour, L. Stein and S. McCouch (2002). Gramene, a tool for grass genomics. Plant Physiology **130**(4): 1606-1613.

Zhang, X., J. Wang, J. Huang, H. Lan, C. Wang, C. Yin, Y. Wu, H. Tang, Q. Qian, J. Li and H. Zhang (2012). Rare allele of OsPPKL1 associated with grain length causes extra-large grain and a significant yield increase in rice. Proceedings of the National Academy of Sciences **109**(52): 21534-21539.

Zhao, X. and M. Fitzgerald (2013). Climate change: Implications for the yield of edible rice. PLoS One **8**(6): e66218.

Chapter 3

A genomics and multi-platform metabolomics approach to identify new traits of rice quality in traditional and improved varieties

M. Calingacion[1,2], C. Boualaphanh[3], V. Daygon[1], R. Anacleto[1], R. Sackville Hamilton[1], B. Biais[4], C. Deborde[4,5], M. Maucourt[5,6], A. Moing[4,5], R. Mumm[7,8], R. de Vos[7,8], A. Erban[9], J. Kopka[9], T. Hansen[10], K. Laursen[10], J. Schjoerring[10], R. D. Hall[2,7,8], M. Fitzgerald[1]

Published, Metabolomics (2012) 8:771–783; doi: 10.1007/s11306-011-0374-4

[1]International Rice Research Institute, Los Banos, Laguna, Philippines; [2]Laboratory of Plant Physiology, Wageningen University, Wageningen, The Netherlands; [3]Rice and Cash Crop Research Institute, Vientiane, Lao PDR; [4]INRA—UMR1332 Fruit Biology and Pathology, Centre INRA de Bordeaux, 33140 Villenave d'Ornon, France; [5]Metabolome-Fluxome Facility of Bordeaux Functional Genomics Center, IBVM, Centre INRA de Bordeaux, 33140 Villenave d'Ornon, France; [6]Universite´ de Bordeaux—UMR1332 Fruit Biology and Pathology, 33140 Villenave d'Ornon, France; [7]Plant Research International, Wageningen, The Netherlands, [8]Centre for BioSystems Genomics, Wageningen, The Netherlands; [9]Max-Planck-Institute of Molecular Plant Physiology (MPIMP), Am Mu¨hlenberg 1, 14476 Potsdam-Golm, Germany; [10]Plant and Soil Science Section, Department of Agriculture and Ecology, Faculty of Life Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg C Copenhagen, Denmark.

Abstract

Using a novel approach combining four complementary metabolomic and mineral platforms with genome-wide genotyping at 1536 single nucleotide polymorphism (SNP) loci, we have investigated the extent of biochemical and genetic diversity in three commercially relevant waxy rice cultivars important to food production in the Lao People's Democratic Republic (PDR). Following cultivation with different nitrogen fertiliser regimes, multiple metabolomic data sets, including minerals, were produced and analysed using multivariate statistical methods to reveal the degree of similarity between the genotypes and to identify discriminatory compounds supported by multiple technology platforms. Results revealed little effect of nitrogen supply on metabolites related to quality, despite known yield differences. All platforms revealed unique metabolic signatures for each variety and many discriminatory compounds could be identified as being relevant to consumers in terms of nutritional value and taste or flavour. For each platform, metabolomic diversity was highly associated with genetic distance between the varieties. This study demonstrates that multiple metabolomic platforms have potential as phenotyping tools to assist breeders in their quest to combine key yield and quality characteristics. This better enables rice improvement programs to meet different consumer and farmer needs, and to address food security in rice-consuming countries.

Introduction

Since the turn of the century, many rice-producing countries have gained access to export markets, bringing significant income to their countries (Dodsworth 1997). This is due to both economic development and yield benefits delivered by the Green Revolution. The rice of the Green Revolution was of the non-waxy type, meaning that it contains amylose. In the Lao People's Democratic Republic (PDR), waxy rices, which do not contain amylose, are the staple food (Bounphanousay, Appa Rao et al. 2000). As a result, the Lao PDR gained no benefit from the Green Revolution, and continuing to grow low-yielding traditional varieties significantly compromised food security until the late 1990s. At this time, breeding programs tailored to waxy rices led to the release of improved varieties such as Thadokkham 1 (TDK1) and Thasano 1 (TSN1) (Inthapanya, Boualaphan et al. 2006). Despite higher yield and acceptable quality, these cultivars do not have the same traits of quality, or command the same price premium, as the two most popular traditional varieties, Kai Noy Leuang (KNL) and Hom Nang Nouane (HNN). As the Lao PDR negotiates its entry into the World Trade Organisation, export opportunities are emerging for its prized traditional waxy varieties of rice. If the quality of the two traditional varieties can be combined with the

agronomic adaptation of high-yielding improved varieties, such as TDK1 or TSN1, this could have considerable positive economic and social impact on poor Lao rice farmers.

In order to combine quality with yield, it is necessary to understand important quality traits, be able to select for them in a breeding program, and to find a suitable agronomically adapted background. A previous study indicates that TSN1, which is responsive to fertilisers, yields twice as much as the traditional varieties, and has better milling quality than TDK1 (Boualaphanh, Calingacion et al. 2011), is such a suitable background. However, identifying quality traits is complex. Current tools of quality evaluation have evolved around the more commonly consumed non-waxy types of rice, and particularly centre around the effects of amylose (Fitzgerald, McCouch et al. 2009). These tools, therefore, are potentially inappropriate for waxy types.

Over the past few years, genotyping technology has advanced rapidly, to the point where identification of important loci has become significantly easier (Newton-Cheh and Hirschhorn 2005, Shirasawa, Shiokai et al. 2006, Raghavan, Naredo et al. 2007, Hyten, Smith et al. 2009, Lee, Koh et al. 2009, Wright, Tung et al. 2010) when an appropriate mapping population and accurate phenotyping tools are available (Tarpley and Roessner 2007) . Aside from sensory panels, there is no phenotyping tool that adequately describes the taste of rice (Champagne, Bett-Garber et al. 2010). In the last decade, technologies for the detection and identification of small molecules in plant tissue has undergone rapid development (Hall 2006) to the point where these ‘metabolomic’ tools could be utilised to reveal biochemically important metabolites that affect the taste and nutritional value of the foods we eat (Hall, Brouwer et al. 2008).

The objectives of the present study are to develop a unique multi-platform metabolomics and ionomics approach to (i) identify differences in primary metabolites, volatile compounds and mineral elements between the grains of TSN1, KNL and HNN grown at four different N fertiliser regimes; (ii) determine the effect of N on the metabolic signature of each variety; and (iii) determine whether differences in grain metabolites correlate with genetic differences between the varieties and with known differences in taste and flavour that Lao rice-consumers prize. The over-arching objective is to explore the possibility of developing a new generation of selection tools for rice improvement programs, which could also have broader application for all grain crop improvement programs.

Materials and methods

Plant material

Three varieties of rice, *Oryza sativa* L., were used in this study. TSN1 is in the *indica* germplasm class and is an improved variety with valuable traits of agronomic adaptation, HNN is a traditional landrace and is also *indica*, and KNL is a traditional landrace and in the *tropical japonica* germplasm class. Both traditional varieties have valuable traits of quality. Twenty five seedlings of TSN1, HNN and KNL were planted as subplots in a split plot design within four main plots at the Agriculture Research Centre, Vientiane, Lao PDR. N was applied at either 0, 30, 60 or 90 kg ha^{-1} to each main plot, and 30 kg ha^{-1} each of P and K were applied to the four main plots (Boualaphanh, Calingacion et al. 2011). Each variety was replicated three times in each main plot. Grain from each subplot was harvested at maturity and sun-dried. Paddy from each sample was dehulled (Satake Rice Machine, Tokyo, Japan), brown grain was polished (Grainman 60-230-60-2AT, Grain Machinery Mfg. Corp., Miami, FL, USA), and the milled rice from each subplot (three varieties and three replicates) was then dispatched to collaborators for profiling of metabolites and mineral elements.

SNP genotyping

DNA was extracted from leaves of the three varieties exactly as previously described (Cuevas, Daygon et al. 2010). The DNA extracts were quantified to 50ng μl^{-1} prior to genotyping with Nanodrop 1000 (Thermo Scientific, Wilmington, DE, USA). Single nucleotide polymorphism (SNP) genotyping was carried out at 1536 loci (Zhao et al. 2010) using a BeadXpress (Illumina, San Diego, CA, USA). SNP calls were analysed using Alchemy software (Wright, Tung et al. 2010). SNP maps were generated and genetic distances were calculated using the software GGT 2.0: Graphical Genotyping (van Berloo 2008). Euclidean distance was used as a similarity coefficient to determine the genetic distance between each variety.

Metabolite and elemental profiling

Milled rice of each variety, replicate and N treatment was ground in liquid nitrogen and stored at -80^0C. Primary metabolites in polar extracts were profiled by proton-Nuclear Magnetic Resonance (^{1}H-NMR) and gas chromatography-electron impact ionization/time of flight-mass spectrometry (GC-EI/TOF-MS). Volatile compounds were measured by gas chromatography-mass spectrometry (GC-MS) of headspace sampled by solid phase microextraction (SPME) and mineral elemental analysis was done by inductively coupled

plasma-mass spectrometry (ICP-MS). The mineral data were autoscaled with full cross validation, and the three metabolite datasets were mean-centred and Pareto scaled using either Amix software v. 3.5 (*Bruker BioSpin* GmbH, Rheinstetten, Germany) or SIMCA-P 11.0 (Umetrics AB, Umea°, Sweden) (Eriksson, Johansson et al. 2006) for subsequent analysis by PCA. The tendencies highlighted with PCA for the discriminant compounds were verified using univariate analyses (Analysis of Variance, ANOVA).

^{1}H-NMR profiling of polar compounds

^{1}H-NMR profiling was performed on ethanolic extracts after precipitation of residual starch. Two replicates for each ground rice flour sample were extracted as previously described for fruit and Arabidopsis (Moing, Maucourt et al. 2004), with modifications. Fifty mg of lyophilised powder were extracted successively with 2 ml of ethanol/water mixtures: 80/20, 50/50 (v/v) and pure water (4 ml) for 15 min at 80^{o}C. After centrifugation, the supernatants were combined, dried under vacuum and lyophilised. The lyophilised extracts were mixed with 500 µl of 100 mM potassium phosphate buffer pH 6.0, 1 mM ethylene diamine tetraacetic acid disodium salt (EDTA), in D_2O, titrated with KOD solution to pH 6 when necessary. In order to precipitate residual starch, nine volumes of absolute ethanol (4.5 ml) were added to rice extracts (500 µl), which were then mixed vigorously for 30 s. Samples were stored at 4^{0}C for 24 h and then centrifuged at 30,000 x *g* for 20 min. The supernatant was then collected and dried under vacuum. Dried extracts were solubilised with 500 µl of D2O and 5 µl of solution of sodium salt of deuterated trimethylsilylpropionic acid (TSP-*d4*, 0.01% final concentration) were added. The mixture was centrifuged at 10,000 x *g* for 5 min at room temperature. The supernatant was then transferred into a 5 mm NMR tube for acquisition.

^{1}H-NMR spectra were recorded on a Bruker ™ Avance Spectrometer (Bruker BioSpin, Wissembourg, France) at 500.162 MHz and 300 K using a 5 mm Broad Band Inverse probe (BBI). Spectra were acquired using a classical monodimensional liquid pulse sequence set with a 90^{o} pulse, 32 K data points, a spectral width of 6000 Hz, 20 s relaxation delay and 64 scans. The acquired spectra were line broadened (0.3 Hz), and manually phased using TOPSPIN v1.3 software (Bruker BioSpin, Wissembourg, France) and manually baseline corrected. They were then aligned with the TSP-d4 signal at d 0.00 ppm. Before statistical analysis, data reduction of the entire spectra (10–0.5 ppm with exclusion of 4.9–4.5 ppm leading to 456 integrated buckets of 0.02 ppm width) was done, followed by normalisation against total intensity. Attribution of each significant bucket to one compound, or a mixture

of compounds, was performed using comparison to a library of authentic standards and standard spiking. The ^{1}H-NMR spectra were converted into JCAMP-DX (the Joint Committee on Atomic and Molecular Physical data - Data Exchange format) standard exchange format and have been deposited, with associated metadata and compound list, into the Metabolomics Repository of Bordeaux MeRy-B (http://www.cbib.u-bordeaux2.fr/MERYB/projects/query_exp.php?project_id=37).

Derivatised GC-MS analysis of polar compounds

Metabolite profiling of polar compounds was performed as detailed previously (Allwood, Erban et al. 2009) using GC-EI/TOF-MS. An Agilent 6890N24 gas chromatograph (Agilent Technologies, Boblingen, Germany) was equipped with a FactorFour VF-5 ms capillary column, 30-m length, 0.25-mm inner diameter, 0.25-µm film thickness (Varian-Agilent Technologies, Boblingen, Germany), which was connected to a Pegasus III time-of-flight mass spectrometer (LECO, St. Joseph, MI, USA). Fifty mg ground rice flour samples were standardised by 30µl of 0.02 mg ml^{-1} $^{13}C_6$-sorbitol in water and subjected to a two-step methanol (15 min at 70^{0}C) and chloroform (5 min at 37^{0}C) extraction. The polar metabolite fraction was obtained by liquid partitioning into water/methanol using a final methanol/water/chloroform solvent phase system (330:400:230, v/v/v). 80 µl aliquots of the polar metabolite fraction were dried by vacuum concentration and chemically derivatised by sequential methoxyamination and trimethylsilylation prior to GC-EI/TOF-MS analysis (Allwood et al. 2009). Retention indices (RIs) were calibrated by addition of a C_{10}, C_{12}, C_{15}, C_{18}, C_{19}, C_{22}, C_{28}, C_{32}, and C_{36} n-alkane mixture to each sample immediately prior to splitless GC injection.

GC-EI/TOF-MS chromatograms were acquired, visually controlled, baseline corrected and exported in NetCDF file format using ChromaTOF software (Version 4.22; LECO, St. Joseph, USA). GC-MS chromatography data were converted into a RI-aligned standardized numerical data matrix using the TagFinder software (Luedemann, Strassburg et al. 2008, Allwood, Erban et al. 2009). Compounds were identified within the TagFinder software by mass spectral and RI matching to the reference collection of the Golm Metabolome Database (GMD, http://gmd.mpimpgolm.mpg.de/; (Hummel, Strehmel et al. 2010). Guidelines for this manually supervised metabolite identification process were the presence of at least three specific mass fragments per compound and a retention index deviation <1.0% (Strehmel, Hummel et al. 2008). All mass features of an experiment were normalised to sample weight and to internal standard prior to statistical analysis.

Headspace GC-MS analysis of volatile compounds

Headspace volatiles were collected by SPME using a 65-mmpolydimethylsiloxane-divinylbenzene fibre (Supelco, Bellefonte, USA) as described in detail (Verhoeven, Jonker et al. 2011). The volatile compounds were thermally desorbed at 250^0C by inserting the fiber for 1 min into the GC injection port (GC 8000, Fisons Instruments, Cheshire, UK). The released compounds were transferred onto the analytical column (HP-5, 30 m 9 0.25 mm ID, 1.05 µm – film thickness) in splitless mode. The temperature program started at 45^0C (2-min hold) and rose 5^0C min^{-1} to 250^0C (5-min hold). The column effluent was ionised by electron impact (EI) ionisation at 70 eV (MD800 electron impact MS, Fisons Instruments, Cheshire, UK). Mass scanning was done from 35 to 400 m/z with a scan time of 2.8 scans s^{-1}. GC-MS raw data were processed by using MetAlign software (Lommen 2009) to extract and align the mass signals (s/n ≥ 3). Mass signals that were below s/n of 3 were randomized between 2.4 and 3 times the calculated noise value. Mass signals that were present in ≤ 6 samples were discarded. Signal redundancy per metabolite was removed by means of clustering and mass spectra were reconstructed (Tikunov, Lommen et al. 2005). Metabolites were identified by matching the mass spectra of obtained metabolites to authentic reference standards and the NIST08, Wiley, and Wageningen Natural compounds spectral library and by comparison with RIs in the literature (Strehmel, Hummel et al. 2008).

ICP-MS analysis of mineral elements

The ground rice samples were digested in 100 ml closed vessels in a microwave oven (Multiwave 3000, Anton Paar, Graz, Austria) for 50 min at 210^0C with a maximum pressure of 40 bar. The digestion medium consisted of 250 mg dry sample, 5 ml 65% ultrapure HNO_3 (J.T. Baker Instra-Analysed Reagent), and 5 ml 15% H_2O_2 (30% Extra-Pure, Riedel de Ha¨en, Selze, Germany). After digestion, the samples were diluted to 3.5% v/v HNO_3 with ultrapure water (Milli-Q Element, Millipore, Massachusetts, USA) (Hansen, Laursen et al. 2009). Multi-elemental analysis was performed using ICP-MS (Agilent 7500ce, Agilent Technologies, Manchester, UK) tuned in standard mode. The plasma power was operated at 1500 W and the argon carrier and make-up gases were set at 0.82 and 0.17 l min^{-1}, respectively. Sample uptake was maintained at approximately 0.6 ml min^{-1} by a perfluoroalkoxy micro-flow nebulizer. Elimination of spectral interferences was obtained by the use of an octopole ion guide with the cell gasses helium or hydrogen as described previously (Laursen, Hesselhøj Hansen et al. 2009). Seven replicates of certified reference material NIST 8436 (durum wheat, particle size <200 µm; National Institute of Standards and Technology, Gaithersburgh, MD, USA) were included to validate digestion efficiency

and analytical accuracy. Only element concentrations deviating less than ±10% from the certified mean reference values were accepted. Limit of detection (LOD) was determined as three times the standard deviation of minimum 7 blanks and only data above LOD was included for chemometrics.

Association analysis

In order to determine the extent that phenotypic divergence reflects genetic divergence, the genetic distance between each variety was compared to the distance between each genotype cluster from each phenotyping platform. For each set of metabolite or mineral element profiling data, mean scores were calculated for each of the three varieties on the first two principal components (PC). Genetic distances between varieties were determined as the Euclidean distance calculated from 1536 SNPs assayed on one plant per variety using GGT 2.0 software (van Berloo 2008), and converted to principal coordinates. Procrustes Analysis was used to rotate and scale the variety mean scores on the metabolite principal components to give the best fit to the genetic principal coordinates. The principal coordinates were then used to generate triangles whereby each vertex of the triangle is a variety, and the length of each side is equal to the scaled Euclidean distance.

Results

Genetic differences between varieties

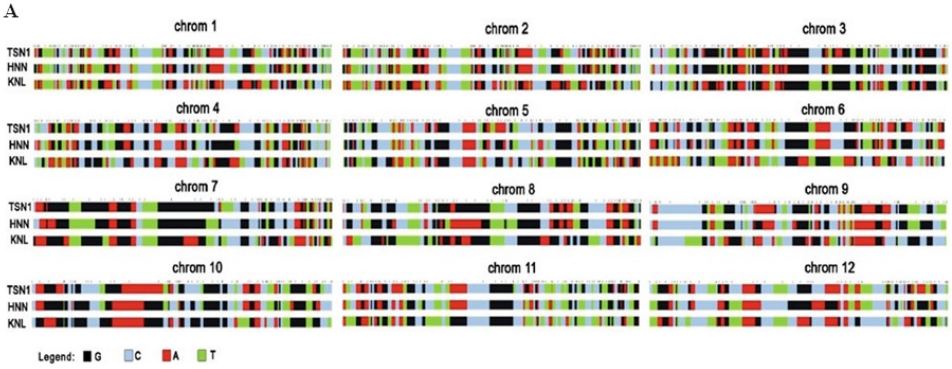
The three varieties studied include an improved indica variety, TSN1, a traditional indica variety, HNN, and KNL, a traditional variety from the tropical japonica class. Genome-wide genotyping of SNPs at 1536 loci (Fig. 1) shows that the genetic similarity between the traditional and improved indica is about 80%, whereas KNL differs from both the indica varieties by about 50% at these SNP loci. However, it can be seen in Fig. 1b that the differences between KNL and HNN are not at the same loci as the differences between KNL and TSN1. The Euclidean distance between each genotype reflects the genetic similarity and is shown in Table 3.

Discriminating between the quality of each variety

A set of standard quality evaluation data, usually obtained to indicate cooking and sensory quality, was previously determined for each of the three varieties grown in each N treatment (Boualaphanh, Calingacion et al. 2011). The data include tests of gelatinisation temperature; data derived from viscosity curves, in this case breakdown, setback and retrogradation (Fitzgerald, Martin et al. 2003); and hardness and stickiness of

the cooked grains. PCA of this dataset of six quality evaluation variables shows that the first two principal components together explain 98% of the variability. Examination of the score plot (Fig. 2) shows two distinct clusters, but clustering is not on the basis of either genotype or N treatment.

A



B

Location and number of unique polymorphic loci

	chrom 1	chrom 2	chrom 3
TSN1	15	15	11
HNN	10	7	8
KNL	56	26	31

	chrom 4	chrom 5	chrom 6
TSN1	8	9	3
HNN	12	7	7
KNL	27	37	41

TSN1	chrom 7	chrom 8	chrom 9
HNN	6	6	6
KNL	5	8	4
	22	25	32

	chrom 10	chrom 11	chrom12	TOTAL
TSN1	11	9	8	99
HNN	7	3	12	98
KNL	14	30	21	362

Figure 1. Genetic differences between the rice varieties. (a) Genetic variability between TSN1, KNL and HNN rice varieties at 1536 SNP loci across all 12 chromosomes. Different shades of gray indicate differences in nucleic acid. At these loci, the genetic similarity between TSN1 and KNL is 53.2%, and between TSN1 and HNN is 80.7% and between KNL and HNN is 53.6%. (b) The number of SNP loci on each chromosome (chrom) that are uniquely polymorphic to one variety, when that same locus is monomorphic for the other two varieties.

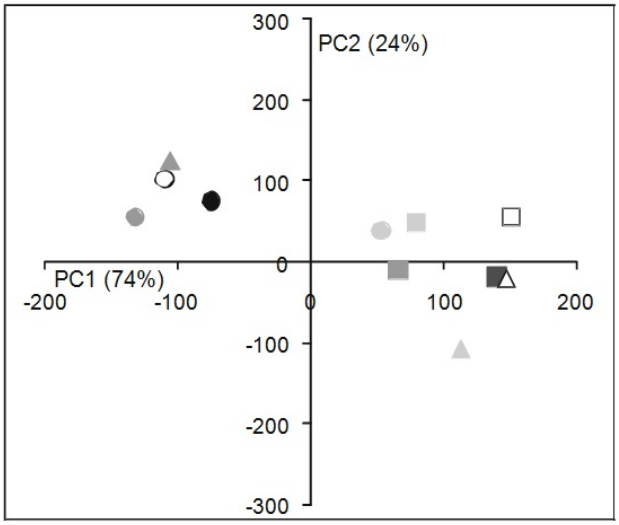


Figure 2. PCA scores plots of data from six routine quality evaluation of grain traits of the three rice varieties cultivated under four different N regimes showing two clusters unrelated to N or genotype. Genotypes: HNN, squares; KNL, circles; TSN1, triangles. N regime: 0 kg N ha^{-1}, unfilled; 30 kg N ha^{-1}, light grey; 60 kg N ha^{-1}, dark grey; 90 kg N ha^{-1}, black.

Discriminating between the grain metabolome of each variety

Four metabolite or elemental profiling technologies were applied to the analysis of polished grains from the three varieties grown under different nitrogen fertiliser regimes in order to try and separate the varieties, and find compounds and minerals that could explain the known flavour differences.

First, 14 mineral elements detected by Inductively Coupled Plasma Mass Spectroscopy (ICP-MS) were quantified in the polished grains of each variety from each N treatment. Observation of the PCA scores plot shows that all samples of KNL formed one distinct cluster, though significant overlap was seen between HNN and TSN1 (Fig. 3a). PC1 explains 47% of the variation and PC2 explains 21%. No clustering or sub-clustering was seen on the basis of N treatment. Several macro- and micronutrient minerals differed significantly between genotypes, with most significant differences between KNL and the other two varieties (Table S1).

PCA analysis of proton Nuclear Magnetic Resonance (^{1}H-NMR) signatures of polar extracts showed that the samples clustered on the basis of genotype (Fig. 3b). The major

metabolites contributing to separation of TSN1 along PC1 were sucrose and raffinose on the negative side and mainly lipids, betaine and choline on the positive side (Table 1). PC2, however, explaining 18% of the total variability, separated KNL and HNN from TSN1 (Fig. 3b). The major metabolites contributing to this separation along PC2 were choline on the negative side and lipids on the positive side (Tables 1, S2). A clear separation between the lowest and highest N levels in the PC1 x PC2 plane could be observed within the TSN1 cluster, but the effect of N was not clearly seen for the other two genotype clusters (Fig. 3b). The relative intensities of selected spectra regions corresponding to the compounds discriminating the three varieties in Table 1 and their ANOVA data are listed in Table S2.

Since detection of primary compounds is more sensitive by gas chromatography-electro impact-time of flight mass spectroscopy (GC-EI-TOF-MS) than ^{1}H-NMR, the data in Fig. 3b were complemented with GC-EI-TOF-MS of polar extracts. PCA was performed using 43 variables, and samples of each variety clustered together (Fig. 3c). PC1 accounted for 50% of the variation, and along that axis, samples of KNL were separated from those of TSN1 and HNN. The PC2 axis, accounting for 25% of the variation, separated the TSN1 samples from the other two varieties (Fig. 3c). ANOVA showed that the main compounds driving the separation along PC1 were cysteine, 5-oxoproline, ribonic acid, glycerol, threonine, and putrescine, tyrosine and trehalose (Table 1, SIII). The major metabolites that separated TSN1 along the PC2 axis were threitol, arabinonic acid, proline, azelaic acid, glycerol, fumarate, and allantoin (Table 1, S3). Many of the other compounds leading to the separation along the PC axes were identified as sugar alcohols and hexose sugars (Table 1). A selection of the most discriminating compounds and the ANOVA data are shown in Table S3.

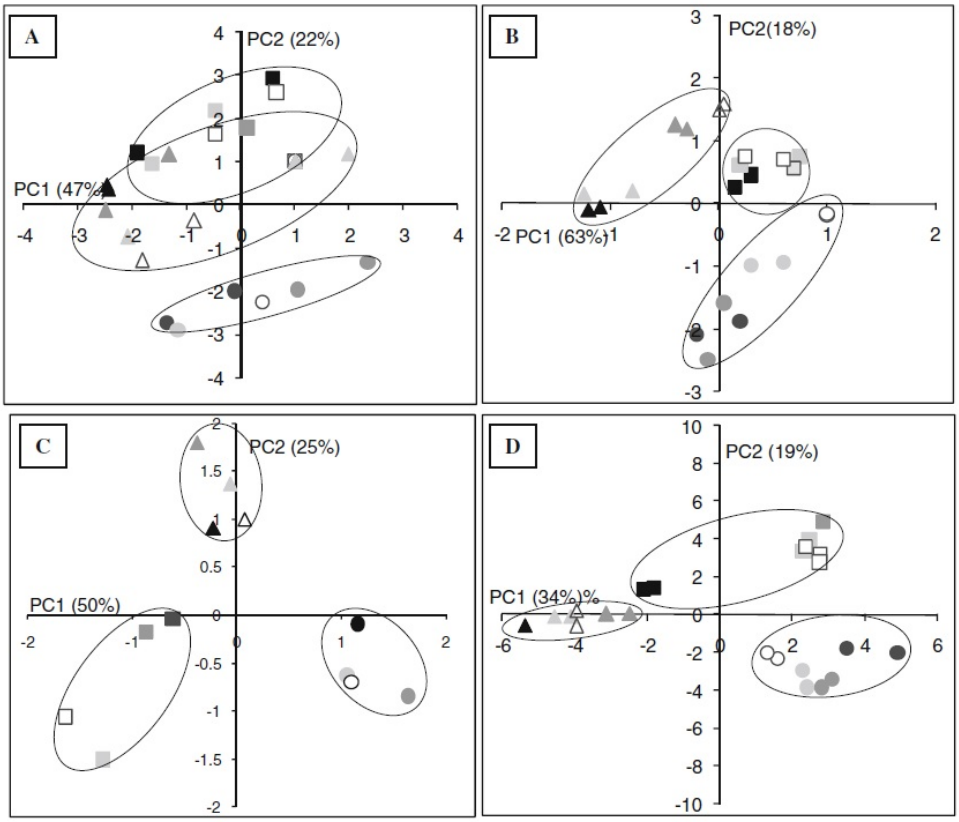


Figure 3. Phenotypic variability between rice varieties.(a) Mineral nutrients determined by ICP-MS, (b) polar metabolite profiling with 1H-NMR fingerprints, (c) polar metabolites determined by GC-TOF-MS, (d) volatiles determined by SPME GC-MS. Genotypes: HNN, squares; KNL, circles; TSN1, triangles. N regime: 0 kg N ha-1, unfilled; 30 kg N ha-1, light grey; 60 kg N ha-1, dark grey; 90 kg N ha-1, black.

Table 1 Polar compounds discriminating HNN, KNL or TSN1 detected by [1]H-NMR[a] or derivatised GC-EI-TOF-MS[b] of extracts of polished grains

Compounds	HNN	KNL	TSN1
Lipids			
lipids-1[a]	++	+	+
lipids-2[a]	+++	+	++
linoleic acid[b]	++	+	+
Sugars			
glucose[a,b]	+	+	++
sucrose[a]	+	++	+
sucrose and raffinose[a]	++	+++	+
raffinose[b]	++	++	+
galactose[b]	++	+	+
maltose[b]	+	++	+
psicose[b]	+	+	++
trehalose[b]	+	++	+
Sugar alcohols			
glycerol[a,b]	+	+	++
tagatose[b]	+++	+	+++
erythritol[b]	+	+	++
mannitol[b]	+	+	++
threitol[b]	+	++	+++
galactitol[b]	++	+	+++
inositol[b]	++	++	+
xylitol[b]	+	+++	+
ribitol[b]	+	++	+
Organic acids			
malate[a]	++	+	+++
fumarate[a,b]	++	+	++
succinate and unknown[a]	++	+	+++
succinate	++	++	++
azelaic acid	++	+	+
lyxonic acid	+	++	+
benzoate [b]	+	++	+

+/++ Indicate the relative peak intensities between varieties

Table 1 continued

Compounds	HNN	KNL	TSN1
gluconate[b]	+	+	++
ribonate[b]	+	++	+
arabinonate[b]	++	+	+
Amino compounds			
alanine[b]	+	++	+
alanine + lipids[a]	++	+	++
asparagine[b]	+++	+	++
aspartate[b]	++	+	++
betaine + unknown[a]	++	+	+++
choline[a]	++	+	+++
choline-o-sulfate[a]	++	+	+++
proline[b]	++	+	+
GABA + valine + unknown[a]	++	+	++
putrescine[b]	++	+++	+
glutamate[b]	+	++	+
valine + lipids[a]	++	+	++
cysteine[b]	+	++	+
isoleucine[b]	+	+	+
lysine[b]	+	+	+++
adenine[b]	+	++	+
5-oxoproline[b]	+	++	+
serine[b]	+	++	+
mannosamine[b]	++	+	++
phenylalanine[b]	++	+	+
tyrosine[b]	+	++	+
indole[b]	++	+	++
allantoin[b]	+++	++	+
Inorganic acid			
phosphoric acid[b]	++	++	+

+/++ Indicate the relative peak intensities between varieties.

Volatile compounds are also major determinants of taste and flavour and these were determined for each variety using GC-MS. The PCA scores plot shows that volatile compounds determined by solid phase micro extraction (SPME) GC-MS clearly separated the three varieties (Fig. 3d). PC1 explained 34% of total variability, and TSN1 was separated along the PC1 axis from HNN and KNL. PC2 explained 19% of total variability and clearly separated the samples of HNN from the other two varieties (Fig. 3d). Four significant principal components were extracted, cumulatively explaining 77% of variation. The three varieties were separated by 54 compounds (Table 2), though not all could be identified by the spectral libraries. The major discriminating compound was 2-acetyl-1-pyrroline (2AP). TSN1 had undetectable levels of 2AP while the other two varieties both contained 2AP. The elevated concentration of 2AP in KNL contributed to the separation of those samples from HNN. HNN grains are characterised by ketones, KNL grains by 2AP and several alcohols, and TSN1 grains, by only three volatiles compounds that could be identified, such as butane 2,3 diol (Table 2) and twelve that could not. Interestingly, several of the discriminating volatile compounds unique to each variety have quite low odour thresholds and unique flavour notes (Table 2), and the relative importance of each is shown in Table S4.

Relating the biochemical phenotype of the grain with the genome

For each metabolite and mineral element profiling platform, Euclidean distances between varieties were calculated from Fig. 3, using the first two principal components (Table 3). Based on 1536 SNP loci of the three varieties, the genetic distance between HNN and TSN1 was smallest, and was similar between KNL and HNN and between KNL and TSN1 (Table 3). The residual sum of squares from the Procrustes rotation of the association between Euclidean distances based on the genome and metabolome or ionome shows that each biochemical profile associates with genotype (Table 3). Comparing the triangles generated from each set of normalised coordinates shows that mineral elements, polar metabolites and volatile compounds all associate very well with the genetic distance between each (Fig. 4). All platforms showed a smaller distance between HNN and TSN1, and a larger distance between KNL and the other two varieties.

Table 2. Volatile compounds, determined by SPME GC-MS, that discriminated each variety, their odour thresholds and flavour.

Compounds	Odour	Threshold	HNN	KNL	TSN1
Ketones					
2-acetyl-1-pyrroline	popcorn	0.1	+	++	
2-heptanone	fruity	140	++		
2-octanone	herbal	50	++		
2-hexanone			++		
2,3 heptadione	cheese*		++		
3-octen-2-one	berry*		++		
3,5-octadiene-2-one	herby		++		
Aldehydes					
hexanal	grassy	5	++		
pentanal	floral	12	++		
2-heptenal	fatty	13	++		
2-octenal	herby	3	++		
2-hexenal	fruity	17	++		
2-butyl,2-octenal	green tea	2	++		
benzaldehyde	almond	350	++		
Alcohols					
1-octen-3-ol	mushrooom	1	++		
1-pentanol	plastic	4000	++		
1-heptanol	citrus*			++	
1-hexanol	grassy*	2500		++	
ethanol				++	
1-octanol	citrus	110		++	
2,3-butanediol	creamy*				++
2-ethyl1-hexanol	rose				++

The odour/flavour of each compound and its threshold (ppb) in water is also shown.

*Indicate compounds that are used commercially at 5ppm or less to add flavour or fragrance to manufacture foods or perfumes.

+/++ Indicate the tendencies between varieties based

Table 2 continued

Compounds	Odour	Threshold	HNN	KNL	TSN1
Hydrocarbons					
pentadecane	waxy		++		
undecane	herbal			++	
longicyclene	floral				++
dl-limonene	citrus	10		++	
1-ethyl-3-methyl benzene				++	
Furans					
2-propylfuran	fruity	6000	++		
2-butylfuran	wine	10000	++		
2-pentylfuran	beany	2000	++		
Carboxylic acids					
Nonanoic acid	rancid	3000		++	
Esters					
ethyl benzoate	cherry	60		++	

The odour/flavour of each compound and its threshold (ppb) in water is also shown.

+/++ Indicate the tendencies between varieties based.

Table 3 Procrustes rotation of correspondence between Euclidean distances between varieties from each analytical platform, based on principal components 1 and 2, and Euclidean genetic distance between each variety

Pairs	Primary metabolites		Volatiles	Minerals	Genes
	GC-EI-TOF-MS	^{1}H-NMR	SPME GC-MS	ICP-MS	SNPs
TSN1-HNN	2.67	1.29	8.06	4.13	12.1
HNN-KNL	2.87	2.56	7.35	5.96	18.1
TSN1-KNL	2.82	2.94	9.14	4.21	18.4
	Procrustes rotation				
	0.04	0.02	0.06	0.06	

Procrustes rotation is calculated as the residual sum of squares scaled so that the total sum of squares is 1.

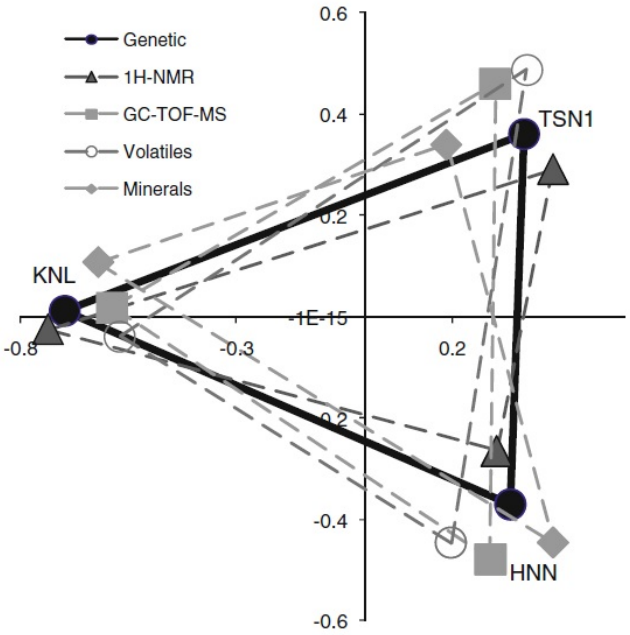


Figure 4. Association between genetic and phenotypic differences between rice varieties. Comparison of standardised coordinates of the Procrustes rotation (Table 3) of Euclidean distances separating genotypes in Fig. 1a and the metabolomic and elemental data from Fig. 3, plotted as triangles showing excellent agreement between genotype and all metabolic profiles of each variety.

Discussion

In many rice-consuming countries, flavour of the rice is considered just as important as texture (Schiller et al. 2006). For example, the rices used in this study differ in taste (Bounphanousay 2007), and the two traditional varieties (HNN and KNL) command a 20% price premium in Lao markets over TSN1. Standard tools used to measure grain quality could not differentiate the three varieties (Fig. 2), indicating a need to bring new science to rice quality. In the present study, a unique combination of analytical metabolomics and elemental profiling platforms has been used to determine volatile and polar primary metabolites and minerals in polished grains. The data obtained were analysed to determine if (i) unique profiles existed for each variety that might explain the

differences in flavour, (ii) N fertiliser affected these profiles, and (iii) associations could be found between the metabolome of each variety and genotype.

The three varieties were well separated by all metabolite and elemental platforms (Fig. 3a–d), suggesting that each variety has a unique metabolite signature. Most of the detected primary polar metabolites were present in all three varieties (Table 1), but at differing levels, so discrimination between the varieties from these platforms is mainly quantitative, which is possibly expected with primary metabolism. By contrast, Table 2 shows that particular volatile compounds were often unique to each variety, suggesting that these contribute more than the polar primary metabolites to the uniqueness of each metabolic signature.

No consistent effect of N regime was seen for any of the datasets (Figs. 2, 3a–d), and in all cases, clustering was based on genotype. Some association with N was seen within the TSN1 genotype cluster for volatile and primary compounds, but not for the other two varieties. Previous data on the same samples (Boualaphanh, Calingacion et al. 2011) are consistent with these findings. The yield components mostly affected by N were increased tiller number and grain number per panicle, and no change in grain weight or protein content (Boualaphanh, Calingacion et al. 2011). Traditional varieties are not usually responsive to N, so it is therefore unsurprising that N did not induce consistent variation in metabolites.

The three varieties show distinct genetic differences (Fig. 1a). The two indica varieties share the most genetic similarity, but both indica varieties share only about 50% homology with the tropical japonica variety, though the differences are not at the same loci (Fig. 1b). These genetic differences are consistent with known groupings within domesticated rice (Garris, Tai et al. 2005). Table 3 and Fig. 4 show that metabolomes and the ionome associated excellently with genetic distances, indicating that (i) discriminating compounds follow genetic differences between germplasm classes, (ii) sufficient polymorphic loci existed to associate with metabolites, and (iii) metabolite data from each platform can be associated with mapping populations, using discriminating SNPs to find candidate loci. In the future, this will enable discovery of the genes responsible for the synthesis of important taste, flavour and nutritional compounds, such as those detected in the present paper.

Nutritionally valuable compounds

Many of the metabolites that were detected in the milled grains have attributed nutritional benefits (Table 1). Examples are GABA, found in all varieties, which is involved in a host of cortical functions in mammals (Sanacora, Mason et al. 1999), and is involved in signalling and osmotic regulation in plants (Bouche´ and Fromm 2004). Choline was detected in all three varieties, and is a methyl donor involved in brain health and liver function (Zeisel, Da Costa et al. 1991). Glucosamine was found in HNN and is known to contribute to the formation of cartilage and is used to alleviate and prevent pain due to arthritis and bone stress (Hughes and Carr 2002). Azelaic acid, found in HNN, is an organic compound that contributes to the health of hair and skin (Nguyen and Bui 1995). A significant amount of lysine was detected in extracts of TSN1, and none in the other two (Table 1, S3). Lysine, an essential amino acid, has multiple physiological functions (Baker 2005). Cereal proteins are low in lysine, and efforts to increase lysine centre around altering the balance of storage proteins that accumulate (Shih 2004) but perhaps non-protein lysine, such as that found in TSN1, reveals an additional mechanism for increasing lysine in cereals.

Improving the amount of Zn in polished rice is the subject of a large international effort under Harvestplus (www.harvestplus.com) to address hidden hunger. In the present study, the higher concentration of Zn in KNL discriminated it from the other two varieties (Table SI). Consistent with this, KNL was found to have high Zn content in a study using 56 Lao varieties (Bounphanousay 2007). The concentration of Zn in KNL approaches the ambitious targets set by Harvestplus. Consequently, KNL would be a useful starting point for further varietal improvement where sensory properties are to be combined with enhanced nutritional properties.

All three varieties contain hexose sugars such as tagatose and psicose, and sugar alcohols such as xylitol and erythritol (Table 1). These sweeteners do not induce a blood sugar response, which could explain some of the variability found in the glycaemic index of different rices (Wheeler and Pi-Sunyer 2008) and be useful in the management of diseases like Type 2 diabetes (Association 2004).

It has generally been assumed that polished rice provides mostly calories, and that the bran layer contains all the compounds with potential nutritional benefit (Yokoyama 2004, Sharif and Butt 2006, Sakamoto, Hayashi et al. 2007, Yu, Nehus et al. 2007, Heinemann,

Xu et al. 2008, Shen, Jin et al. 2009, Butsat and Siriamornpun 2010). This is the first demonstration of the extensive biochemical diversity that occurs in polished rice grains (Tables 1, 2), many of which have a role in human physiology and health.

Compounds of taste and flavour

Consumers have difficulty describing the taste of rice, so testing for taste is generally not part of rice quality programs (Fitzgerald, McCouch et al. 2009, Champagne, Bett-Garber et al. 2010). However, a recent sensory study, using popular varieties grown throughout South and Southeast Asia, revealed a characteristic flavour profile for those varieties commonly grown and consumed in Southeast Asia consisting of sweet, floral, grassy and dairy notes (Champagne, Bett-Garber et al. 2010). KNL and HNN were found to contain polar and volatile com- pounds that give sweet, floral, fruity and grassy flavours (Tables 1, 2). HNN contained a number of ketones, many of which are used commercially in low concentrations to impart fruity, nutty, floral and butter/dairy aromas and flavours in foods. KNL contained compounds that are commercially used in low concentrations to impart grassy fresh aroma notes. Several of the volatile compounds detected have an odour threshold detectable by humans (Table 2), based on the detection limit of the GC-MS. The strongest of these is 2AP, but some of the other compounds also have low odour thresholds (Table 2). KNL and HNN can be differentiated by consumers in taste trials (Bounphanousay 2007), and the data presented here suggest that each has different compounds that could contribute to flavours that humans can detect. They both also contain unpleasant compounds (off-flavours) such as putrescine, which was first found in putrefying flesh (Olle 1986), but which is considered to be a precursor in the pathway of aroma in rice (Bradbury, Gillies et al. 2008). Putrescine was not found in TSN1 (Table 1), consistent with a possible role in 2AP synthesis in the two fragrant rices (Bradbury, Gillies et al. 2008). However TSN1 contained the highest levels of succinate, fumarate and malate (Table 1), which could indicate oxidation of lipids, leading to aromas of rancidity.

Further analyses are required to correlate the current findings with the quality of cooked rice. The presence of amino acids and reducing sugars among the primary metabolites in the raw grains of each variety leads to the possibility that Maillard reaction products may also be produced during cooking (De Kimpe and Keppens 1996) leading to additional contributions to flavour. Cysteine was found in KNL (Table 1), and this amino acid is an efficient precursor of a-acetyl N heterocycles, which have a very low odour

threshold and contribute to the roasted aroma of cooked rice (Kerler, vanderVen et al. 1997).

Three volatile compounds found in TSN1 could be identified (Table 2), and another 12 could not, and this was also found for the other two varieties. Therefore more research is needed in order to build a more complete picture of flavour and fragrance of rice. Nevertheless, this study is the first to show that metabolomic profiling can be a valuable new tool which leads to unique signatures. The generic nature of the approaches also means that such knowledge is not just restricted to the varieties used here. Instead our study shows feasibility and benefit of metabolomic analyses from which all rice and crop improvement programs should profit.

The identification of important compounds in rice should be the subject of further work tuned to find genes that lead to their synthesis and accumulation. An expanded palette of both DNA and chemical markers for such compounds would greatly enhance the capacity of rice improvement programs to select actively for delicious and nutritious rice.

Acknowledgments

All the authors acknowledge funding from the EU for the project META-PHOR (FOOD-CT-2006-036220) which was part of FP VI. RDH, RCHV and RM acknowledge additional support from the Centre for BioSystems Genomics, an initiative under the auspices of the Netherlands Genomics Initiative. MAF acknowledges additional support from the International Rice Research Institute (IRRI). CB acknowledges support from IRRI, KKU and NAFRI. University of Copenhagen acknowledges additional support from HarvestPlus to the Danish Plant Research Network entitled 'Supporting the Foundation for Iron and Zinc Biofortification in Beans, Rice and Wheat'. Bordeaux Metabolome Facility acknowledge additional support from IBiSA and INRA CNOC, and Daniel Jacob for developing and maintaining the MeRy-B database.

Supporting Information

Table S1. Observed differences in mineral concentrations between the three varieties.

Minerals	Mineral concentration in µg g^{-1} dry matter[a]			Significance (p values) of mineral changes obtained from ANOVA[b]		
	KNL	HNN	TSN1	HNN vs KNL	HNN vs TSN1	KNL vs TSN1
B	0.13±0.06	0.16±0.02	0.22±0.12	>0.05	>0.05	>0.05
Mg	249±64	226±28	205±27	>0.05	>0.05	>0.05
P	716±167	646±67	569±58	>0.05	>0.05	**0.035**
S	879±120	763±117	862±56	>0.05	>0.05	>0.05
K	925±151	892±63	773±54	>0.05	>0.05	**0.016**
Ca	28±5.39	19±3.26	18±7.70	**0.018**	>0.05	**0.008**
Mn	7.16±1.52	7.52±1.09	5.47±0.30	>0.05	**0.003**	**0.014**
Fe	2.99±0.86	3.06±0.40	2.52±0.43	>0.05	>0.05	>0.05
Cu	0.75±0.10	1.30±0.17	1.01±0.11	**<0.001**	**<0.001**	**0.002**
Zn	18±2.65	15±1.94	15.42±0.85	**0.015**	>0.05	>0.05
Se	0.04±0.01	0.05±0.01	0.06±0.01	>0.05	>0.05	**0.008**
Mo	0.69±0.18	0.63±0.07	0.68±0.17	>0.05	>0.05	>0.05
Cd	0.00±0.00	0.07±0.01	0.04±0.02	**<0.001**	>0.05	**<0.001**
Ba	0.07±0.02	0.02±0.00	0.02±0.00	**<0.001**	>0.05	**<0.001**

[a] Values represent mean (n=8) S.D.

[b] Values in bold represent significant differences.

Table S2. Observed relative intensities of selected spectra regions corresponding to a known metabolite or metabolite mixture identified using 1H-NMR of polar extracts after elimination of starch residue.

	Relative intensity of characteristic metabolite regions from NMR dataset (arbitrary units)			Significance (P values) of changes obtained from ANOVA*		
	KNL	HNN	TSN1	HNN vs KNL	HNN vs TSN1	KNL vs TSN1
Alanine+ unknown+lipids (1.49 ppm)	48.74 ± 2.85	57.15 ± 1.33	57.15 ± 3.47	< 1 x 10^{-4}	.96 x 10^{-1}	3.92 x 10^{-4}
Asparagine (2.93 ppm)	6.13 ± 0.74	14.04 ± 0.52	11.62 ± 0.92	< 1 x 10^{-4}	< 1 x 10^{-4}	< 1 x 10^{-4}
Aspartate (2.83 ppm)	7.55 ± 0.55	13.05 ± 0.64	12.51 ± 0.76	< 1 x 10^{-4}	.72 x 10^{-1}	< 1 x 10^{-4}
Betaine+unknown (3.27 ppm)	23.40 ± 1.45	27.08 ± 1.17	33.11 ± 2.12	2.19 x 10^{-4}	< 1 x 10^{-4}	< 1 x 10^{-4}
Choline (3.19 ppm)	43.99 ± 3.27	49.22 ± 0.96	56.16 ± 3.91	1.45 x 10^{-3}	.41 x 10^{-4}	< 1 x 10^{-4}
Choline-o-Sulfate (3.25 ppm)	30.66 ± 1.23	36.20 ± 1.60	44.42 ± 3.37	< 1 x 10^{-4}	< 1 x 10^{-4}	< 1 x 10^{-4}
Fumarate (6.51 ppm)	0.15 ± 0.74	1.96 ± 0.43	2.07 ± 0.46	1.13 x 10^{-4}	.34 x 10^{-1}	< 1 x 10^{-4}
Glucose (5.23 ppm)	8.56 ± 1.05	8.38 ± 0.56	11.81 ± 1.83	6.87 x 10^{-1}	.19 x 10^{-4}	2.07 x 10^{-3}
Glutamate+Malate (2.37 ppm)	17.38 ± 0.73	19.99 ± 0.44	21.29 ± 1.32	< 1 x 10^{-4}	.67 x 10^{-2}	< 1 x 10^{-4}
Glycerol (3.63 ppm)	78.12 ± 5.78	70.78 ± 4.62	86.78 ± 6.81	0.25 x 10^{-1}	.48 x 10^{-4}	2.90 x 10^{-2}
Lipids (0.87 ppm)	87.83 ± 6.28	108.07 ± 5.08	95.01 ± 9.95	< 1 x 10^{-4}	.94 x 10^{-3}	1.05 x 10^{-1}
Lipids (1.55 ppm)	53.82 ± 2.94	71.44 ± 1.63	64.61 ± 5.59	< 1 x 10^{-4}	.78 x 10^{-3}	9.21 x 10^{-4}

Values represent mean (n = 8 for HNN and TSN1, n=7 for KNL) ± standard deviation. The values in parentheses correspond to the centre of each 0.02 ppm-width spectra region. ANOVA calculations indicate significant differences in spectra region intensity between varieties.

* Values in bold represent significant changes.

Table S2 continued.

	Relative intensity of characteristic metabolite regions from NMR dataset (arbitrary units)			Significance (P values) of changes obtained from ANOVA*		
	KNL	HNN	TSN1	HNN vs KNL	HNN vs TSN1	KNL vs TSN1
Succinate+unknown (2.41 ppm)	17.99 ± 0.76	21.28 ± 0.61	25.18 ± 1.55	< 1 x 10^{-4}	< 1 x 10^{-4}	< 1 x 10^{-4}
Sucrose (3.47 ppm)	117.01 ± 12.15	86.75 ± 3.66	83.77 ± 9.67	< 1 x 10^{-4}	.58 x 10^{-1}	1.05 x 10^{-4}
Sucrose+Raffinose (5.43 ppm)	112.20 ± 10.44	83.98 ± 3.98	71.25 ± 9.48	< 1 x 10^{-4}	.52 x 10^{-3}	< 1 x 10^{-4}
Valine+Lipids (1.03 ppm)	19.05 ± 1.23	26.28 ± 0.73	26.28 ± 2.05	< 1 x 10^{-4}	.95 x 10^{-1}	< 1 x 10^{-4}
Valine+GABA+unknown (2.27 ppm)	22.93 ± 1.50	26.98 ± 1.62	28.73 ± 3.49	4.55 x 10^{-4}	.50 x 10^{-1}	2.23 x 10^{-3}

Values represent mean (n = 8 for HNN and TSN1, n=7 for KNL) ± standard deviation. The values in parentheses correspond to the centre of each 0.02 ppm-width spectra region. ANOVA calculations indicate significant differences in spectra region intensity between varieties.

* Values in bold represent significant changes.

Table S3. Observed relative intensities of significantly different metabolites detected with GC-EI-TOF-MS, extending the list from those identified using 1H-NMR.

	Concentration of metabolites in each (ppm)			Significance (P values) of changes obtained from ANOVA*		
	KNL	HNN	TSN1	HNN vs KNL	HNN vs TSN1	KNL vs TSN1
Adenine	84.24 ±12.85	55.9 ±13.37	59.78 ±8.34	**8.57E-03**	7.70E-01	**2.61E-03**
Allantoin	55.58 ±28.01	74.92 ±19.67	39.19 ±5.15	1.82E-01	**2.35E-03**	1.49E-01
Arabinonic acid	45.82 ±8.77	74.64 ±21.45	48.63 ±10.02	**3.06E-02**	**3.62E-02**	9.67E-01
Azelaic acid	69.52 ±13.63	87.36 ±12.48	58.34 ±7.59	**3.80E-02**	**3.10E-03**	**4.22E-02**
Benzoate	90.64 ±10	69.5 ±10.65	78.95 ±11.19	**9.45E-03**	4.14E-01	5.20E-02
Butyrate	82.68 ±19	13.41 ±3.55	18.81 ±6.13	**3.09E-05**	2.15E-01	**3.07E-04**
Cysteine	83.17 ±11.86	57.69 ±9.61	43.13 ±8.28	**2.32E-03**	**3.19E-02**	**3.46E-04**
Erythritol	60.38 ±6.05	53.78 ±11.83	85.55 ±12.96	1.89E-01	**1.37E-02**	**2.26E-02**
Galactitol	44.17 ±6.25	65.65 ±19.92	89.79 ±9.35	**4.67E-02**	8.74E-02	**3.36E-05**
Gluconate	43.15 ±7.24	50.49 ±10.6	87.84 ±11.15	3.47E-01	**4.15E-03**	**8.46E-04**
Glycerol	76.33 ±18.56	58.73 ±9.43	45.75 ±6.94	5.02E-02	**3.92E-02**	**2.42E-03**
Indole	46.56 ±9.68	77.79 ±22.8	83.11 ±7.42	**2.64E-02**	7.29E-01	**7.01E-04**
Linoleic acid	52.51 ±12.36	74.66 ±26.85	35.31 ±3.53	2.05E-01	**6.35E-03**	**1.71E-02**
Xylitol	85.31 ±10.87	29.14 ±7.19	41.24 ±7.83	**5.67E-05**	7.43E-02	2.45E-04
Trehalose	78.2 ±29.03	47.63 ±4.19	30.03 ±10.21	**1.02E-01**	3.14E-02	**1.43E-02**

Values represent mean (n = 8 ± standard deviation). ANOVA calculations indicate significant differences in spectra region intensity between varieties.

* Values in bold represent significant changes

Table S3 continued

	Concentration of metabolites in each (ppm)			Significance (P values) of changes obtained from ANOVA*		
	KNL	HNN	TSN1	HNN vs KNL	HNN vs TSN1	KNL vs TSN1
Inositol	75.77 ±16.35	74.17 ±17.95	47.14 ±5.98	7.02E-01	**7.71E-03**	**9.92E-04**
Lysine	21.42 ±7.83	15.39 ±1.87	68.42 ±22.73	8.45E-02	**1.98E-04**	**2.04E-03**
Lyxonic acid	88.49 ±13.69	57.34 ±13.24	60.39 ±11.46	**7.65E-03**	9.27E-01	**1.05E-02**
Maltose	72.49 ±20.84	29.43 ±9.23	34.69 ±10.45	**5.23E-03**	6.06E-01	**1.07E-02**
Mannitol	15.36 ±3.99	14.44 ±0.76	48.84 ±34.82	**4.51E-01**	**1.97E-02**	**2.76E-02**
Mannosamine	42.67 ±7.09	81.28 ±20.65	83.91 ±9.7	**3.76E-03**	9.36E-01	**9.96E-05**
Phenylalanine	75.07 ±12.28	86.92 ±9.95	61.75 ±8.97	1.45E-01	**4.79E-03**	**4.24E-02**
Proline	35.98 ±5.59	70.57 ±31.97	24.18 ±4.02	7.54E-02	**7.01E-03**	**1.53E-02**
Psicose	55.72 ±7.08	59.35 ±13.56	90.33 ±7.84	7.89E-01	**1.02E-02**	**1.21E-04**
Putrescine	79.3 ±18.35	59.43 ±7.34	17.83 ±6.54	**2.89E-02**	**6.86E-04**	**3.31E-04**
Ribitol	71.03 ±20.54	46.16 ±9.92	51.3 ±7.12	**1.58E-02**	5.54E-01	**2.34E-02**
Ribonic Acid	85.47 ±13.43	58.33 ±8.88	50.64 ±12.28	**3.25E-03**	2.45E-01	**8.17E-03**
Serine	77.3 ±17.3	26.95 ±6.13	36.77 ±5.42	**3.70E-04**	8.64E-02	**7.04E-04**
Tagatose	17.82 ±5.1	67.35 ±21.54	85.05 ±14.54	**8.07E-04**	2.57E-01	**9.99E-05**
Threitol	61.56 ±6.29	36.19 ±8.59	96.06 ±3.35	2.08E-03	**1.20E-04**	**3.38E-05**

Values represent mean (n = 8 ± standard deviation). ANOVA calculations indicate significant differences in spectra region intensity between varieties.

* Values in bold represent significant changes

Table S4. Observed relative intensities of significantly different metabolites detected with GC--MS. Values represent mean (n = 8 ± standard deviation). ANOVA calculations indicate significant differences in spectra region intensity between varieties.

	Relative metabolite intensities of selected volatile compounds from GC-MS (arbitrary units)			Significance (P values) of changes obtained from ANOVA*		
	KNL	HNN	TSN1	HNN vs KNL	HNN vs TSN1	KNL vs TSN1
Pentanal	42966 ± 3463	70468 ± 7238	46111 ± 4978	**1.37E-07**	**1.73E-06**	1.65E-01
1-Pentanol	106729 ± 8382	141537 ± 8369	75983 ± 7422	**8.75E-07**	**1.35E-10**	**1.93E-06**
Hexanal	1220518 ± 81256	2016335 ± 168036	1349003 ± 111910	**8.78E-09**	**2.14E-07**	**1.99E-02**
unknown	113572 ± 23754	136695 ± 45867	174208 ± 33561	2.26E-01	8.30E-02	**9.42E-04**
1-Hexanol	242162 ± 50667	143260 ± 15867	61540 ± 13058	**1.19E-04**	**2.14E-08**	**1.26E-07**
2-Heptanone	201426 ± 26321	335546 ± 35373	196436 ± 13340	**5.82E-07**	**5.68E-08**	6.40E-01
2-Acetyl-1-pyrroline	27714 ± 6181	11718 ± 1222	350 ± 122	**4.70E-06**	**2.72E-13**	**5.42E-09**
2-Heptenal	44587 ± 6944	68904 ± 7627	47276 ± 6991	**1.06E-05**	**3.79E-05**	4.53E-01
unknown	38457 ± 2244	66320 ± 3483	59814 ± 4520	**2.13E-11**	**6.12E-03**	**9.68E-09**
1-Octen-3-ol	139574 ± 13033	275013 ± 36043	160853 ± 8909	**9.40E-08**	**5.11E-07**	**1.90E-03**
2-Pentylfuran	191132 ± 52759	323182 ± 63352	196398 ± 35557	**4.71E-04**	**2.19E-04**	8.18E-01
unknown	353290 ± 165079	183465 ± 82957	47274 ± 9629	**2.10E-02**	**4.03E-04**	**1.26E-04**
Octanal	91789 ± 37981	110137 ± 28221	71532 ± 21632	2.91E-01	**8.30E-03**	2.11E-01

Values represent mean (n = 8 ± standard deviation). ANOVA calculations indicate significant differences in spectra region intensity between varieties.

* Values in bold represent significant changes

Table S4 continued.

	Relative metabolite intensities of selected volatile compounds from GC-MS (arbitrary units)			Significance (P values) of changes obtained from ANOVA*		
	KNL	HNN	TSN1	HNN vs KNL	HNN vs TSN1	KNL vs TSN1
unknown	43708 ± 19457	27345 ± 9970	9115 ± 1970	5.27E-02	**1.70E-04**	**1.93E-04**
3-Octen-2-one	49604 ± 9226	107657 ± 30438	55437 ± 21882	**1.44E-04**	**1.48E-03**	4.99E-01
unknown	29370 ± 12386	19312 ± 8069	5743 ± 2130	7.49E-02	**4.14E-04**	**1.09E-04**
unknown	28008 ± 11605	19754 ± 8227	6320 ± 2417	1.23E-01	**5.70E-04**	**1.41E-04**
unknown	26484 ± 3391	46251 ± 10008	26172 ± 3704	**1.14E-04**	**1.08E-04**	8.63E-01
unknown	47374 ± 7527	43028 ± 9889	25932 ± 3080	3.39E-01	**3.62E-04**	**3.08E-06**
nonanal	207408 ± 13898	216655 ± 28944	187579 ± 10070	4.29E-01	**1.78E-02**	**5.61E-03**
dodecane	285481 ± 67368	253423 ± 81271	158473 ± 40622	4.05E-01	**1.04E-02**	**4.40E-04**
tridecane	121247 ± 12919	103836 ± 10868	89409 ± 10659	**1.13E-02**	**1.79E-02**	**9.77E-05**
2-butyl-2-octenal	6961 ± 4585	65234 ± 10445	16429 ± 1982	**8.34E-10**	**3.38E-09**	**1.00E-04**

Values represent mean (n = 8 ± standard deviation). ANOVA calculations indicate significant differences in spectra region intensity between varieties.

* Values in bold represent significant changes

References

Allwood, J., A. Erban, S. Koning, W. Dunn, A. Luedemann, A. Lommen, L. Kay, R. Löscher, J. Kopka and R. Goodacre (2009). Inter-laboratory reproducibility of fast gas chromatography–electron impact–time of flight mass spectrometry (GC–EI–TOF/MS) based plant metabolomics. Metabolomics **5**(4): 479-496.

Association, A. (2004). Nutrition principles and recommendations in diabetes. Diabetes Care **27**: s36.

Baker, D. (2005). Comparative nutrition and metabolism: Explication of open questions with emphasis on protein and amino acids. Proceedings of the National Academy of Sciences **102**(50): 17897-17902.

Boualaphanh, C., M. Calingacion, R. Cuevas, D. Jothityangkoon, J. Sanitchon and M. Fitzgerald (2011). Yield and quality of traditional and improved Lao varieties of rice. Scienceasia **37**(2): 89-97.

Bouche´, N. and H. Fromm (2004). GABA in plants: Just a metabolite? Trends in Plant Science **9**(3): 110-115.

Bounphanousay, C. (2007). Use of phenotypic characters and DNA profiling for classification of the genetic diversity in black glutinous rice of the Lao PDR Khon Kaen University.

Bounphanousay, C., S. Appa Rao, K. Kanyavong, B. Sengthong, J. Schiller and M. Jackson (2000). Conservation of Lao rice germplasm at the International Rice Genebank. Part 5. Los Banos, Philippines: IRRI.

Bradbury, L., S. Gillies, D. Brushett, D. Waters and R. Henry (2008). Inactivation of an aminoaldehyde dehydrogenase is responsible for fragrance in rice. Plant Molecular Biology **68**(4-5): 439-449.

Butsat, S. and S. Siriamornpun (2010). Antioxidant capacities and phenolic compounds of the husk, bran and endosperm of Thai rice. Food Chemistry **119**(2): 606-613.

Champagne, E., K. Bett-Garber, M. Fitzgerald, C. Grimm, J. Lea, K. Ohtsubo, S. Jongdee, L. Xie, P. Bassinello, A. Resurreccion, R. Ahmad, F. Habibi and R. Reinke (2010). Important sensory properties differentiating premium rice varieties. Rice **3**: 270–281.

Cuevas, R., V. Daygon, H. Corpuz, L. Nora, R. Reinke, D. Waters and M. Fitzgerald (2010). Melting the secrets of gelatinisation temperature in rice. Functional Plant Biology **37**(5): 439-447.

De Kimpe, N. and M. Keppens (1996). Novel syntheses of the major flavor components of bread and cooked rice. Journal of Agricultural and Food Chemistry **44**(6): 1515-1519.

Dodsworth, J. (1997). How Indochina's economies took off. Finance and Development **34**(1): 20–23.

Eriksson, L., E. Johansson, N. Kettaneh-Wold, C. Wikstrom, J. Trygg and S. Wold (2006). Multi- and megavariate data analysis. Basic principles and applications. Umea, Sweden, Umetrics Academy.

Fitzgerald, M., M. Martin, R. Ward, W. Park and H. Shead (2003). Viscosity of rice flour: A rheological and biological study. Journal of Agricultural and Food Chemistry: 2295-2299.

Fitzgerald, M., S. McCouch and R. Hall (2009). Not just a grain of rice: the quest for quality. Trends in Plant Science **14**(3): 133-139.

Garris, A., T. Tai, J. Coburn, S. Kresovich and S. McCouch (2005). Genetic structure and diversity in Oryza sativa L. Genetics **169**: 1631–1638.

Hall, R. (2006). Plant metabolomics: from holistic hope, to hype, to hot topic. New Phytologist **169**(3): 453-468.

Hall, R., I. Brouwer and M. Fitzgerald (2008). Plant metabolomics and its potential application for human nutrition. Physiologia Plantarum **132**(2): 162-175.

Hansen, T., K. Laursen, D. Persson, P. Pedas, S. Husted and J. Schjoerring (2009). Micro-scaled high-throughput digestion of plant tissue samples for multi-elemental analysis. Plant Methods **5**: 12.

Heinemann, R., Z. Xu, J. Godber and U. Lanfer-Marquez (2008). Tocopherols, tocotrienols, and c-oryzanol contents in japonica and indica subspecies of rice (Oryza sativa L.) cultivated in Brazil. Cereal Chemistry **85**(2): 243-247.

Hughes, R. and A. Carr (2002). A randomized, double-blind, placebocontrolled trial of glucosamine sulphate as an analgesic in osteoarthritis of the knee. Rheumatology **41**(3): 279-284.

Hummel, J., N. Strehmel, J. Selbig, D. Walther and J. Kopka (2010). Decision tree supported substructure prediction of metabolites from GC-MS profiles. Metabolomics **6**(322-333).

Hyten, D., J. Smith, R. Frederick, M. Tucker, Q. Song and P. Cregan (2009). Bulked segregant analysis using the GoldenGate Assay to locate the Rpp3 locus that confers resistance to soybean rust in soybean. Crop Science **49**(1): 265-271.

Inthapanya, P., C. Boualaphan, Hatsadong and J. Schiller (2006). The history of lowland rice variety improvements in Laos. Rice in Laos. M. Schiller, M. Chanphengxay, B. Linquist and S. Appa Rao. International Rice Research Institute:Philippines**:** 325-358.

Kerler, J., J. G. M. vanderVen and H. Weenen (1997). alpha-acetyl-N-heterocycles in the Maillard reaction. Food Reviews International **13**(4): 553-575.

Laursen, K., T. Hesselhøj Hansen, D. Persson, J. Schjoerring and S. Husted (2009). Multi-elemental fingerprinting of plant tissue by semi-quantitative ICP-MS and chemometrics. Journal of Analytical Atomic Spectroscopy **24**: 1198–1207.

Lee, G., H. Koh, H. Chung, A. Dixit, J. Chung and K. Ma (2009). Development of SNP-based CAPS and dCAPS markers in eight different genes involved in starch biosynthesis in rice. Molecular Breeding **24**(1): 93-101.

Lommen, A. (2009). MetAlign: Interface-driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing. Analytical Chemistry **81**(8): 3079-3086.

Luedemann, A., K. Strassburg, A. Erban and J. Kopka (2008). TagFinder for the quantitative analysis of gas chromatography-mass spectrometry (GC-MS) based metabolite profiling experiments. Bioinformatics **24**(5): 732–737.

Moing, A., M. Maucourt, C. Renaud, M. Gaudille`re, R. Brouquisse and B. Lebouteiller (2004). Quantitative metabolomic profiling by 1-dimensional 1H-NMR analyses: Application to plant genetics and functional genomics. Functional Plant Biology **31**: 889-902.

Newton-Cheh, C. and J. Hirschhorn (2005). Genetic association studies of complex traits: Design and analysis issues. Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis **573**(1-2): 54-69.

Nguyen, Q. and T. Bui (1995). Azelaic acid - pharmacokinetic and pharmacodynamic properties and its therapeutic role in hyperpigmentary disorders and acne. International Journal of Dermatology **34**(2): 75-84.

Olle, H. (1986). Putrescine, spermidine, and spermine. Physiology **1**(1): 12-15.

Raghavan, C., M. Naredo, H. Wang, G. Atienza, B. Liu and F. Qiu (2007). Rapid method for detectin SNPs on agarose gels and its application in candidate gene mapping. Molecular Breeding **19**(2): 87-101.

Sakamoto, S., T. Hayashi, K. Hayashi, F. Murai, M. Hori and K. Kimoto (2007). Pre-germinated brown rice could enhance maternal mental health and immunity during lactation. European Journal of Nutrition **46**(7): 391-396.

Sanacora, G., G. Mason, D. Rothman, K. Behar, F. Hyder and O. Petroff (1999). Reduced cortical {gamma}-aminobutyric acid levels in depressed patients determined by proton magnetic resonance spectroscopy. Archives of General Psychiatry: 1043-1047.

Sharif, K. and M. Butt (2006). Preparation of fiber and mineral enriched pan bread by using defatted rice bran. International Journal of Food Properties **9**(4): 623-636.

Shen, Y., L. Jin, P. Xiao, Y. Lu and J. Bao (2009). Total phenolics, flavonoids, antioxidant capacity in rice grain and their relations to grain color, size and weight. Journal of Cereal Science **49**(1): 106-111.

Shih, F. (2004). Rice Proteins. Rice chemistry and technology. E. Champagne. St Paul, Minnesota, American Association of Cereal Chemists Inc**:** 143-157.

Shirasawa, K., S. Shiokai, M. Yamaguchi, S. Kishitani and T. Nishio (2006). Dot-blot-SNP analysis for practical plant breeding and cultivar identification in rice. Theoretical and Applied Genetics **113**(1): 147–155.

Strehmel, N., J. Hummel, A. Erban, K. Strassburg and J. Kopka (2008). Retention index thresholds for compound matching in GC–MS metabolite profiling. Journal of Chromatography B **871**(2): 182-190.

Tarpley, L. and U. Roessner (2007). Metabolomics: Enabling systems level phenotyping in rice functional genomics. Rice functional genomics: Challenges, progress and prospects. N. Upadhyaya. New York, Springer**:** 91-107.

Tikunov, Y., A. Lommen, C. de Vos, H. Verhoeven, R. Bino and R. Hall (2005). A novel approach for nontargeted data analysis for metabolomics. Large-scale profiling of tomato fruit volatiles. Plant Physiology **139**(3): 1125–1137.

van Berloo, R. (2008). GGT 2.0: Versatile software for visualization and analysis of genetic data. Journal of Heredity **99**(2): 232-236.

Verhoeven, H., H. Jonker, R. De Vos and R. Hall (2011). Solid phase micro-extraction GC–MS analysis of natural volatile components in melon and rice. Plant Metabolomics. N. Hardy and R. Hall, Humana Press. **860:** 85-99.

Wheeler, M. and F. Pi-Sunyer (2008). Carbohydrate issues: Type and amount. Journal of the American Dietetic Association **108**(4): S34–S39.

Wright, M., C. Tung, K. Zhao, A. Reynolds, S. McCouch and C. Bustamante (2010). ALCHEMY: A reliable method for automated SNP genotype calling for small batch sizes and highly homozygous populations. Bioinformatics **26**(23): 2952–2960.

Yokoyama, W. (2004). Nutritional properties of rice and rice bran. Rice chemistry and technology. E. Champagne. Minneapolis, American Association of Cereal Chemists, Inc.**:** 595-606.

Yu, S., Z. Nehus, T. Badger and N. Fang (2007). Quantification of vitamin E and y-oryzanol components in rice germ and bran. Journal of Agricultural and Food Chemistry **55**: 7308-7313.

Zeisel, S., K. Da Costa, P. Franklin, E. Alexander, J. Lamont and N. Sheard (1991). Choline, an essential nutrient for humans. The FASEB Journal **5**(7): 2093-2098.

Chapter 4

Delving deeper into technological innovations to understand differences in rice quality.

M. Calingacion[1,2,3], L. Fang[3], L. Quiatchon-Baeza[1,4], R. Mumm[5], A. Riedel[3], R. D. Hall[2,5], and M. Fitzgerald[3]

Rice, in press.

[1]International Rice Research Institute, Los Banos, Laguna, Philippines; [2]Laboratory of Plant Physiology, Wageningen University, Wageningen, The Netherlands; [3]School of Agriculture and Food Science, University of Queensland, St Lucia 4072 Qld, Australia; [4]Department of Crop Sciences, College of ACES, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA; [5]BU Bioscience, Plant Research International, Wageningen, The Netherlands.

Abstract

Increasing demand for better quality rice varieties, which are also more suited to growth under sub-optimal cultivation conditions, is driving innovation in rice research. Here we have used a multi-disciplinary approach, involving SNP-based genotyping together with phenotyping based on yield analysis, metabolomic analysis of grain volatiles, and sensory panel analysis to determine differences between two contrasting rice varieties, APO and IR64. Plants were grown under standard and drought-induced conditions. Results revealed important differences between the volatile profiles of the two rice varieties and we relate these differences to those perceived by the sensory panel. Apo, which is the more drought tolerant variety, was less affected by the drought condition concerning both sensory profile and yield; IR64, which has higher quality but is drought sensitive, showed greater differences in these characteristics in response to the two growth conditions. Metabolomics analyses using GCxGC-MS, followed by multivariate statistical analyses of the data, revealed a number of discriminatory compounds between the varieties, but also effects of the difference in cultivation conditions. Results indicate the complexity of the rice volatile profile, and how metabolomics can be used to help link changes in aroma profile with the sensory phenotype. Our outcomes also suggest valuable multi-disciplinary approaches which can be used to help define the aroma profile in rice, and its underlying genetic background, in order to support breeders in the generation of improved rice varieties combining high yield with high quality, and tolerance of both these traits to climate change.

Introduction

Asia currently faces two major challenges that are predicted to have significant impacts on food security in the region: rapid and significant population growth (Nations 2014), and climate change (IPCC 2014). The Asian population is growing at a much faster rate than those in other regions, and is predicted to peak by 2050. Rice is the staple food of Asia, and to protect Asia against food shortages in the future, it is imperative that rice improvement programs develop new varieties with much higher yield. However, finding an acceptable solution is more complex than increasing yield potential alone.

Socioeconomic patterns within the Asian population are changing rapidly. Research is helping to meet the United Nations Millennium Development goals, lifting people out of poverty (Fanzo and Pronyk 2011), Asia is currently undergoing economic transformation (Kharas and Gertz 2010), and is predicted to become home to the biggest middle class by 2030. Following from this economic growth, it is expected that the majority of Asians will

have more discretionary income and greater financial capacity to make food choices based on assessment of quality and safety (Goodman and Robison 2013). Therefore, the impact of population growth in Asia, on both rice production and rice improvement programs, is that new high-yielding rice varieties must contain the quality traits that increasingly-discriminatory consumers in different Asian markets require (Calingacion, Laborte et al. 2014). Combining yield and quality in rice is not trivial, because the tools for measuring quality are unable to discriminate in a manner that explains the sensory experiences of aroma, taste, flavour and texture (Fitzgerald, McCouch et al. 2009, Daygon and Fitzgerald 2013).

Further challenging the need for more and better rice, the effects of climate change are predicted to impact most strongly in the latitudes and longitudes relevant to Asia (IPCC 2014). Climate change is predicted to lead to more drought events, higher temperatures (Peng, Huang et al. 2004) and more unpredictable weather patterns (Knutson and Tuleya 2004, Trenberth 2005). In anticipation of this dramatically changing environment, rice improvement programs now investigate varietal selection for climate-ready varieties of rice (Zeigler and Barclay 2008). The breeding and domestication of rice has created variation among genotypes that enables cultivation in a number of less favourable environments, including salinity-affected coastal areas, water-scarce upland areas, and lowland monsoonal regions that are prone to inundation and flood. These adaptations demonstrate that the species as a whole houses genes that enable a high degree of plasticity, and thus permit productive growth in different environments (Wassmann, Jagadish et al. 2009). Identifying the genetic basis of different stress tolerances, in order to mobilise causal genes into production of elite varieties is a research priority. Recent examples of gains in managing the need for stress-tolerant rices are identifying genes for submergence tolerance for rice grown in monsoonal regions (Xu, Xu et al. 2006), and genes for drought tolerance for varieties grown in conditions of water scarcity or drought (Zeng, Zhong et al. 2006, Kumar, Dixit et al. 2014).

New rice varieties with tolerance to abiotic stresses offer a clear yield advantage to farmers. However, this does not always translate into profit because next to agronomic traits a major driver of widespread uptake of a new variety is not farmer adoption, but consumer acceptance (Fitzgerald, McCouch et al. 2009, Champagne, Bett-Garber et al. 2010, Boualaphanh, Calingacion et al. 2011, Daygon and Fitzgerald 2013, Calingacion, Laborte et al. 2014). A clear example of this exists in the case of the two varieties Apo (also known as

PSBRc 9) and IR64. Apo repeatedly gives good yield under drought conditions (Venuprasad, Lafitte et al. 2007); it has been released in several drought-prone areas of Asia, but has not been widely adopted, because consumers consider it to have poor eating quality (Fitzgerald, McCouch et al. 2009). IR64, on the other hand, is susceptible to many abiotic stresses, including drought (Venuprasad, Lafitte et al. 2007), but regardless, due to its exceptional quality, has been grown annually on over a million hectares in major rice-producing countries since its release in 1985 (Fitzgerald, McCouch et al. 2009).

Apo and IR64 have been assessed in a paired comparison by a sensory evaluation panel, comprising another 11 similarly matched pairs of rice varieties, where both varieties in each pair had similar physico-chemical traits of quality, but differed in their consumer acceptance/popularity (Champagne, Bett-Garber et al. 2010). The flavour of Apo rice was found to be dominated by sewer/animal, astringent, and water-like metallic notes, whereas IR64 was predominantly characterised by sweet and corn notes (Champagne, Bett-Garber et al. 2010). The compounds that individually, or in combination, are responsible for such flavours and aromas have not yet been identified. In order to assist rice improvement programs to develop varieties that meet consumer taste and flavour requirements as well as meeting the need for increased yield, phenotyping techniques such as metabolomics can be used to identify these compounds of taste and flavour in foods (Hall, Brouwer et al. 2008, Calingacion, Boualaphanh et al. 2012).

Our objectives here were to (i) confirm that flavour notes previously reported by Champagne et al. (2010) for Apo and IR64 grains are reproducible, and are therefore robust trait characteristics of the two varieties; (ii) determine how water scarcity affects these flavour notes, because the cultivation of Apo germplasm was originally developed for water-scarce environments; and (iii) identify a panel of target compounds with low odour thresholds that differ between Apo and IR64, and investigate their association with the phenotypic aromas previously reported. To build up the most detailed picture of these two important rice varieties, and to see how key traits are influenced by water availability, we have used a highly multi-disciplinary approach involving genotyping (Single Nucleotide Polymorphism (SNP)-based) and phenotyping (metabolomics, taste panel sensory analysis, yield measurements) supported by advanced multivariate statistical analyses.

Materials and Methods

Plant material

Apo and IR64 were planted at the Experimental Station of the International Rice Research Institute, Philippines in the dry season of 2012. For each variety, 100 plants, at 1 plant per hill, were planted at 15cm within and between rows in 3 blocks under irrigated and 3 blocks under drought conditions, with 4 replicated plots in each block, in an alpha lattice design. Inorganic fertiliser nitrogen: phosphorus: potassium (NPK) was applied to the field before transplanting at a ratio of 40:40:40 kg ha^{-1} (NPK), and the plants were top-dressed with urea 30 and 55 days after transplanting at a level of 30:0:0 kg ha^{-1}. Drought stress was artificially imposed by draining the field when plants had been assessed to have achieved maximal tillering, so that water stress coincided with the reproductive stage of the plant. Irrigated plots were maintained at a water level of approximately 5 cm until harvest, at which time they were drained. Mature grains from all 100 plants were harvested to determine total yield.The grains were dried in an oven until a moisture content of 12-14% was reached and stored at 21^{o}C until analysis. Grains were dehulled (Otake FCY2 Dehusker, Oharu, Japan), polished (test tube mill) and cryo-ground (2010 SPEX SamplePrep Geno/Grinder) for 3 min at 1750 rpm. We compared yields of the four treatment classes (2 varieties x 2 treatments) in a two factor linear mixed effects model that assessed the presence of variety x treatment interactions.

Sensory evaluation of flour

Eight trained panellists participated in this study. Rice samples were prepared by placing rice flour (1 g) in screw capped vials and heating in a water bath at 80°C for 10 min. Rice was presented in this way to replicate the way the sample and its volatile compounds were later presented for headspace analysis by gas chromatography. Samples were presented to panellists unidentified and in randomised order. Panellists opened the lid of the vial carefully and smelt the aroma of the heated rices. Panellists then scored the intensity or absence of eight flavour notes on a scale of 1-2 indicated by the number of stars. The flavour notes considered were: sweet taste, corn, sweet aromatic, astringent, water-like metallic, sewer/animal, sour/silage and hay-like/musty (Champagne et al 2010). Panellists were free to record additional aromas they detected.

Headspace analysis using two-dimensional gas chromatography time of flight-mass spectrometry (GCxGC TOF-MS)

A. Headspace extraction

Rice flour (1 g) was placed in a sealed vial and allowed to equilibrate overnight at room temperature. The sample was then heated at 80°C for 10 min while being agitated at 500 rpm in an agitator of GCMS. Volatile compounds in the headspace (1 mL) were collected in a syringe and 1.5 mL was injected in a GCxGC TOF-MS system consisting of an Agilent 7890 gas chromatograph (Agilent Technologies, Palo Alto, CA) and Pegasus IV TOF-MS mass spectrometer (LECO, St Joseph, MI). The primary column used was RXI-5sil (29 m x 0.25 mm id x 0.25 µm film thickness) while the secondary column was RxI-17 (1.7 m x 0.1 mm id x 0.18 µm film thickness). The secondary column was placed inside the secondary oven after the thermal modulator. The flow rate of the helium carrier gas was set to a constant flow of 1 mL min^{-1}. The primary column was set at an initial temperature of 45°C for 1 min, then ramped at 10°C min^{-1} to 250°C. The secondary column was programmed with an initial temperature of 60°C for 1 min which was then then increased to 250°C at a rate of 10°C min^{-1}. The thermal modulator was set at 70°C then ramped to 260°C at a rate of 10°C min^{-1} with a 5 s modulation period. The MS mass range was 35-400 m/z with an acquisition rate of 200 spectra s^{-1}. The ion source temperature was set at 250°C and the detector voltage was 1.5kV with electron energy of -70eV and the temperature of the transfer line was 250°C. Headspace of a standard mixture of known compounds (benzene, toluene, ethylbenzene, p-xylene) was also injected as a quality control sample to check if peak identification are accurate and reproducible.

B. Data processing

The raw GCxGC data were processed using ChromaTOF software (version 4.5, LECO, St Joseph, MI). We used the automated peak finding and spectral deconvolution algorithm with baseline offset of 0.5 and signal to noise ratio of 6. Automated retention time alignment was also performed using the statistical comparison function in ChromaTOF. All samples were divided among the two water treatment classes, then data were aligned within and between the classes. Only the compounds that were present in more than 50% of the samples in a class were considered. Mass spectra were compared and matched against the NIST 2011 mass spectral library (www.NIST.gov).

C. Statistical Analysis

All statistical analyses were done in R (RCoreTeam 2014), using a number of packages. First, we assessed treatment effects and interactions on yield, using a simple univariate model. For our multivariate data, to remedy the skew of several variables, they were first log transformed. To moderate the influence of large-scale variation among some compounds, and thereby allow smaller-scale variation in other metabolites to play a greater role in the overall characterisation of rice aromatic qualities, we Pareto scaled all data (van den Berg, Hoefsloot et al. 2006): for each variable, data was centred on a mean of zero and the variance scaled by dividing by the square root of the standard deviation for each variable (R package: 'MetabolAnalyze', (Gift, Gormley et al. 2010)). We used Principal Components Analysis (PCA) (R package: 'FactoMineR', (Husson, Josse et al. 2013)) to extract the first two dimensions in our data that explained the greatest amount of variation. To support our description of the treatment classes by PCA, we extracted the PC scores for the 48 individual samples on these first two dimensions, and treated these condensed descriptions of the principal variation in the data as response variables in a linear model. Because our data included experimental blocks nested within growing condition (irrigated vs. drought), in all analyses we used linear mixed effects modelling to partition out the random effects from our analysis of treatment effects on the metabolic profile encapsulated in the first two principal components (R package: 'lme4', (Bates, Maechler et al. 2014)); as recommended for mixed models, we used Kenward-Rogers estimation of the denominator degrees of freedom (Saxton 2004).We established the significance of effects in our models in the usual manner (Quinn and Keough 2002): log-likelihood ratio tests compared a full model of treatment factors, and the random experimental effects, with a nested model omitting successive effects.

In the second part of our analysis, we looked more closely at a shortlist of compounds of known low odour threshold (below 0.1), as these have greatest influence on rice quality and consumer preference. We created a data set comprising 65 compounds of the original 187, and again explored the treatment effects on chemical composition using PCA of those compounds that are known to influence aroma.

High dimensional data, such as is generated by GCxGC MS, is typically explored using PCA; PC loadings are one method to extract which compounds show the greatest variation, and therefore may be having greatest influence in distinguishing groups. However, high dimensional data presents a challenge in discerning which variables are more important,

and the criteria for acceptance or rejection of putatively important contributions can be arbitrary and less reliable when based on the, typically small, sample sizes of these methods. Here, using a method that gives greater statistical underpinning to the choice of important variables, we used sPLS-DA (R package 'mixOmics'; (Dejean, Gonzalez et al. 2013)), to identify those compounds likely to be most influential in discriminating between treatment effects on the two rice varieties in k-1 dimensions, i.e., three in this instance. We constrained the model to identify the 20 compounds, each with a low odour threshold, that were most influential in discriminating between the treatment groups, for each dimension in turn, extending to three dimensions, which were then extracted from the sPLS-DA model.

Genotyping using 384 SNP chip

In order to gain insight into genetic similarity between Apo and IR64, DNA was extracted from leaf tissue of both using the modified CTAB DNA extraction method (Murray and Thompson 1980). DNA in the extracts was quantified and diluted to 50 ng uL-1 using a Thermo Scientific Nanodrop 1000. Genotyping was performed at 384 SNP loci using an Illumina BeadXpress GoldenGate Genotyping Assay (Illumina, San Diego, CA). Single nucleotide polymorphism calls were analysed using Alchemy software. Genetic distance was estimated using Jaccard's similarity coefficient option in GGT software (van Berloo 2008).

Results

Response to drought treatment

Drought significantly lowered yield for Apo and IR64 by 27% and 38%, respectively (Effect of removing treatment from the model: $X_2 1 = 20.855$, $p = <0.0001$) (Figure 1a). There was a significant difference in overall yield between the two varieties (Effect of removing variety from the model: $X_2 1 = 6.5484$, $p = <0.0105$). The mean yield of IR64 was lower than that of Apo, but showed no statistically significant interaction between genotype and treatment, meaning that neither variety was significantly more affected by treatment than the other. Over the duration of the experiment, there was 3.5 mm average rainfall per day, with some rain falling during grain-filling (Figure 1b), but the artificial drought imposed on the plots by draining at panicle initiation and just before flowering, was enough to significantly affected yield (Figure 1a).

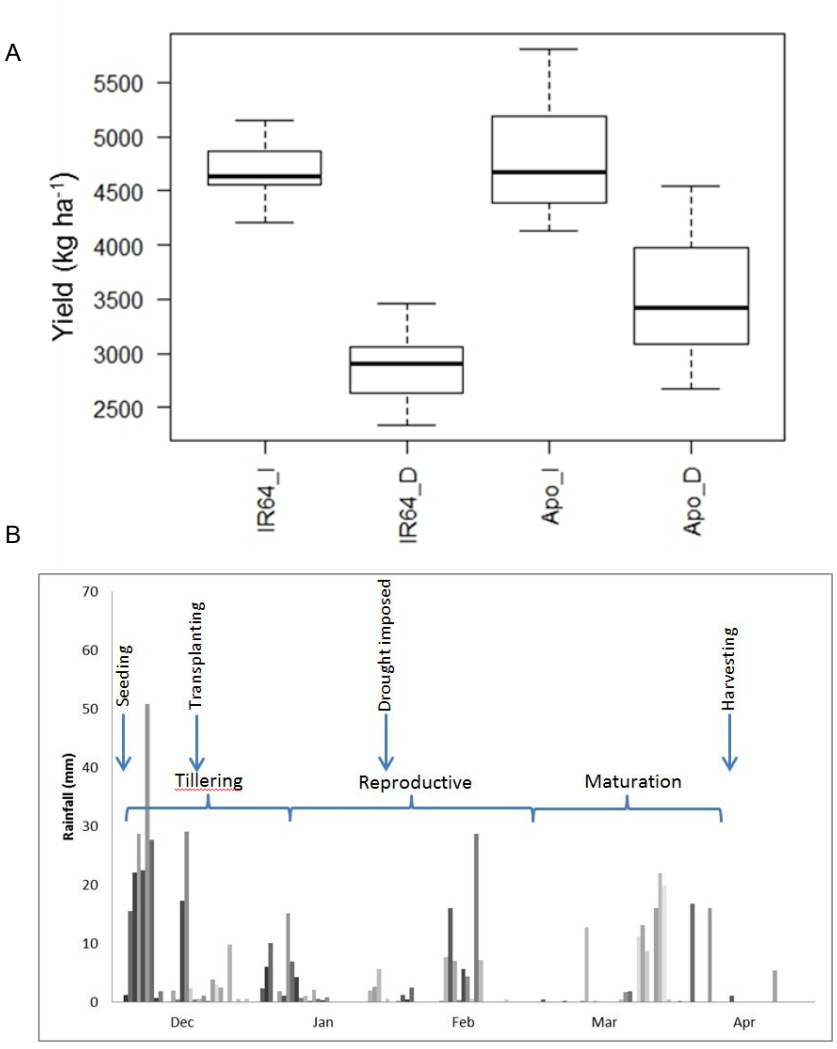


Figure 1. (A) Yields (kg ha^{-1}) of Apo and IR64 rices grown in irrigated and drought treatments (Dry Season 2012). Boxplots represent the median and interquartile range (25-75%), and the whiskers the 5th and 95th percentiles. n=48. (B) Rainfall (mm) at the Experimental Station of International Rice Research Institute, Philippines during the growth of the rice samples (Dry Season 2012). Bars indicate stages of rice development and cultivation. Weather data were supplied by the Climate Unit, Crop and Environment Sciences Division, IRRI.

Sensory evaluation

The sensory panel reproducibly detected aromas of sweet taste, corn and sweet aromatic flavours in the grains of IR64 from irrigated plots, and water-like metallic, astringent, sour silage, sewer animal and hay-like/musty aromas in Apo grains from irrigated plots (Table 1). When comparing the effects of drought within varieties, no difference in aroma was detected for grains of Apo (Table 1). For the two IR64 samples, the levels of corn and sweet aromatic aromas were similar in both treatments. However, higher levels of sweet taste and hay-like/musty were perceived in IR64 from the irrigated than the drought treatment. In addition, water-like metallic, astringent and sour/silage aromas were detected in IR64 grains from the drought treatment.

Table 1. Comparison of flavour attributes in Apo and IR64 from irrigated and drought treatments with sensory qualities/attributes as previously reported by Champagne et al (2010). Eight panellists indicated the strength of the perceived aromatic characteristic of the sample by the number of stars, 1-2.

Flavour	Apo[a]	IR64[a]	Apo		IR64	
			Irrigated	Drought	Irrigated	Drought
Sweet taste	+	++	+	+	++	+
Corn		+			+	+
Sweet aromatic		+			+	+
Astringent	++	+	+	+		+
Water like metallic	++	+	++	++		+
Sewer/animal	++		++	++		
Sour/silage	++		+	+		+
Hay-like/musty			++	++	++	+

[a]Adapted from Champagne et al (2010).

Analysis of headspace samples by GCxGC-TOFMS

We detected 187 volatile compounds in Apo and IR64 (Figure 2a). Principal components analysis (PCA) extracted the main axes of variation in the relative abundance of volatile compounds detected in Apo and IR64 from irrigated and drought treatments. PC1 explained 11% of the variation in metabolite concentrations, with IR64 forming two distinct clusters corresponding to water availability during growth (Figure 2a). Apo from both irrigation treatments clustered closely and showed considerably less variation in PC1 than IR64. Our linear mixed model analysis found no significant interactions between treatment and block

effects (Effect of removing block x variety interaction from the model: $X_2 3 = 5.1799$, $p = 0.1591$) in PC1, so this term was removed from the model before assessing the fixed treatment effects, but we retained the random block component to partition out variances derived from the experimental design. Critically, despite discerning no differing effects on yields of the two varieties, we found a substantive part of variation in metabolites in PC1 was associated with variation in the differing treatment responses of the two rice varieties (Effect of removing two-way interaction, variety x treatment, from the model: $X_2 1 = 22.271$, $p = <0.0001$).

PC2, encapsulating 9% of the total variation modelled, further distinguished samples according to their variety: all Apo samples loaded negatively on PC2, all IR64 samples loaded positively. Our linear mixed model found no variation between varieties among blocks (Effect of removing block x variety interaction from the model: $X_2 3 = 0$, $p = 1$). Variation in scores for PC2 was found to be associated solely with differences between the two rice varieties (Effect of removing variety from the model: $X_2 1 = 90.93$, $p = < 2.2e\text{-}16$).

The second PCA, of selected compounds with aromatic descriptors and low odour threshold, showed some characteristics in common with our first PCA (Figure 2b). PC1, explaining 17% of the total variation in low odour threshold metabolites, described broad separation of IR64 grown in irrigated conditions, with negative scores in this dimension for all samples. The same variety grown under drought conditions showed both positive and negative scores on PC1, where both Apo treatments were closely clustered and had predominantly positive scores. PC2 (12.5%) allocated positive scores to the IR64 drought-treated samples, negative scores to both Apo sample groups, and widely scattered scores for IR64 grown in irrigated conditions whose mean score values were nominally negative on PC2.

Sparse Partial Least Squares-Discriminant Analysis (sPLS-DA) identified 20 compounds for each direction through the data that discriminated between treatment groups (Table 2). The first axis discriminated samples from the IR64 irrigated treatment, the second distinguishing IR64 drought samples from the two Apo treatment classes, while the third described those compounds that differed in extent between the two treatments applied to Apo.

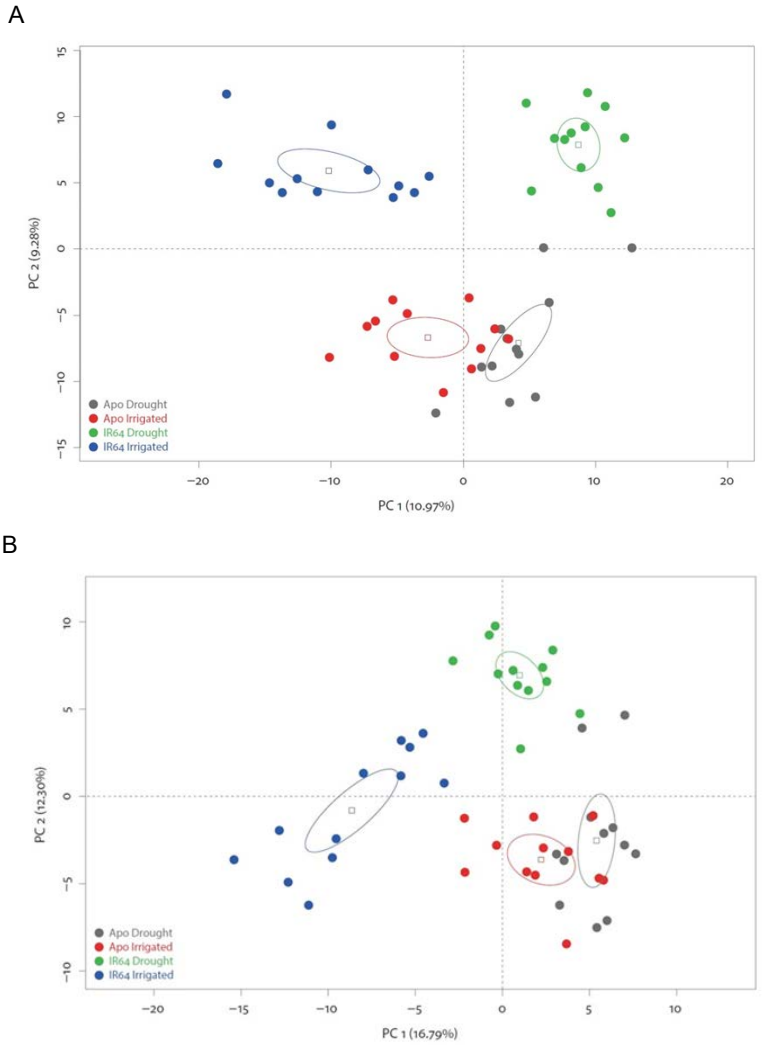


Figure 2. (A) PCA of metabolites detected in samples of Apo and IR64 from irrigated and drought treatments. The square in each ellipse represents the mean value on PC1 and PC2 for each group; the edges of the ellipses represent confidence limits based on the sample variances for each group in each dimension. n=48. (B) PCA of metabolites with known flavour and low odour threshold detected in Apo and IR64 grown in irrigated and drought treatments. The square in each ellipse represents the mean value on PC1 and PC2 for each group; the edges of the ellipses represent confidence limits based on the sample variances for each group in each dimension. n=48.

Table 2. Discriminating compounds, of known flavour and low odour threshold, in Apo and IR64 grown in irrigated and drought treatments. These compounds have been extracted from each of the three axes of a sparse Partial Least Squares-Discriminant Analysis.

First axis		Second axis		Third axis	
Compound	Loading	Compound	Loading	Compound	Loading
Limonene	-0.391	Methyl benzoate	-0.406	1,2,3,4-Tetramethylbenzene	0.434
n-Propylbenzene	-0.372	Camphene	-0.401	Tridecane	0.425
1-Ethyl-3-methylbenzene	-0.358	Benzonitrile	-0.379	Phenethyl acetate	0.366
Indole	0.338	Tetradecane	0.368	Benzophenone	-0.361
2-Heptanone	-0.297	Phenol	-0.334	3-Methylbenzaldehyde	0.262
2-Heptenal	-0.249	Isopropyl dodecanoate	-0.227	Pentadecanoic acid	0.243
Hexanal	-0.233	Decanal	-0.200	Isomethyl ionone	0.225
1-Hepten-3-ol	-0.231	Benzaldehyde	-0.199	Caryophyllene	-0.200
2-Pentylfuran	-0.216	Octanal	-0.194	Naphthalene	0.199
Methylstyrene	-0.201	Heptanal	-0.193	o-Cymene	0.194
Ethyl octanoate	-0.175	1-Propene-1-thiol	0.167	Benzaldehyde	-0.161
2-Propylfuran	-0.173	Acetophenone	-0.102	1-Hexanol	0.139
α-Phellandrene	-0.168	4-Hydroxy-4-methyl,2-pentanone	-0.096	α-Farnesene	0.122
o-Cymene	-0.118	2-Pentylfuran	0.093	2-Methylheptane	0.073
B-Ocimene	-0.107	2-Methylheptane	0.091	2-ethylhexanal	-0.045
2-Octenal	-0.071	(-)-Carvone	-0.087	p-Xylene	-0.041
2-Butylfuran	-0.027	Hexanal	0.065	4-Hydroxy-4-methyl,2-pentanone	0.036
1,3-Dimethylbenzene	-0.014	2,4-Hexadien-1-ol	-0.042	Undecane	0.028
Undecane	-0.009	Pentadecanoic acid	-0.009	Phenol	0.019
α-Pinene	-0.001	Indole	0.007	α-Phellandrene	0.006

Genetic differences between Apo and IR64

The coefficient of parentage of Apo and IR64 is 0.13 indicating that these two varieties are genetically different. Furthermore, analysis of the pedigree tree of each shows no common parents back at least three generations (Figure S1; www.irri.org/iris). Genome-wide

genotyping of SNPs at 384 loci showed that the genetic difference between Apo and IR64 is about 32% (Figure 3). Major differences were found in chromosomes 1, 8, and 11 while chromosomes 6 and 7 of Apo and IR64 were genetically similar.

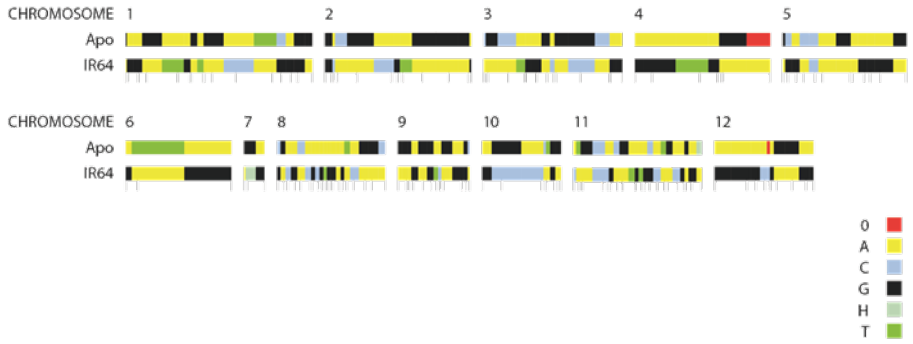


Figure 3. SNP map across the chromosomes of Apo and IR64. Colours represent different nucleotide bases.

Discussion

As climate change exerts its influence on rice-producing countries, it is essential that science delivers new phenotyping tools to support the development and release of new varieties of rice for new and challenging environments. These new varieties must combine stress tolerance for both yield and grain quality in order to ensure consumer acceptability (Calingacion, Laborte et al. 2014). The difficulty in combining both traits and then adding stress tolerance is demonstrated by the variety Apo, which has been released for drought tolerance (Espino and Cinense 2010), despite unacceptable cooking and eating quality. Breeding programs have focused on using genetic pyramiding techniques to create stress-tolerant versions of already popular varieties of rice that can be grown in more challenging environments (Sundaram, Vishnupriya et al. 2008, Septiningsih, Pamplona et al. 2009, Win, Korinsak et al. 2012). This results in a longer breeding cycle by inserting rate-limiting selection steps. However there is still a need to understand and control cooking and eating quality to ensure these varieties are accepted by consumers, and to be able to respond if pleiotropic or epistatic interactions from the introgression have an effect on the phenotypes of quality (Li, Song et al. 2014). As scientific technologies develop, genetic selection tools become more accessible, and the need to push through current yield barriers becomes more pressing. Bringing new science to rice quality to identify traits that consumer values is

a more purposeful pathway forward. In this study, two rice varieties, the unpopular, drought tolerant Apo and the highly popular, drought susceptible IR64, were evaluated for yield and drought tolerance, sensory flavour and volatile metabolites in the grain, and finally genetic differences. Data obtained were used to (i) identify a panel of compounds that may be linked to the cooking and eating quality of the rices, (ii) determine the effect of drought stress on the compounds; and (iii) evaluate the potential of using a mapping population developed from these two varieties to enable future identification of genes that underlie the aromas emitted during cooking and consumption of the rices. The ultimate goal envisaged is that this work will lead to the identification of QTLs for key quality-related traits which will then facilitate the breeder in the quest for new, improved varieties with enhanced sensory characteristics.

Response to drought treatment

Drought during the reproductive stage can reduce spikelet fertility and panicle exsertion which both decrease yield (Wassmann, Jagadish et al. 2009). Yield can therefore be used as a proxy to measure the degree of drought tolerance (Kumar, Bernier et al. 2008). Several studies have demonstrated the drought tolerance of Apo (Venuprasad, Lafitte et al. 2007, Venuprasad, Dalid et al. 2009, Venuprasad, Bool et al. 2012) and susceptibility of IR64 (Venuprasad, Lafitte et al. 2007, Kumar, Bernier et al. 2008, Guan, Serraj et al. 2010, Vikram, Swamy et al. 2011, Ghimire, Quiatchon et al. 2012, Palanog, Swamy et al. 2014). However, in our experiment, the yield of both varieties under drought was reduced, but surprisingly there was no statistical difference found between the two varieties (Figure 1a, Table S1). Furthermore, drought significantly affected the aromas and flavours of the grain from IR64, but it did not have such an effect on the samples of grain from the Apo plots (Figures 2a and b).

Sensory analysis of the grains

Consumer acceptance of a variety of rice is driven initially by the physical appearance of the grains (Fitzgerald, McCouch et al. 2009), and further cemented by, first, the cooking and second, the eating quality of rice (Cramer, Wailes et al. 1993). Characteristics of cooking rice that are important to consumers are (i) the aromas generated by the processes of cooking rice, for example the sweet and floral aromas emitted when jasmine rice is cooking; (ii) cooking time; and (iii) the appearance of the cooked rice when ready for serving and eating (Bett-Garber, Champagne et al. 2012). Eating quality describes the aroma, flavour, taste and texture of the rice in the mouth (Fitzgerald, McCouch et al. 2009). Differences in

the aroma of Apo and IR64 during and after cooking are likely to explain much of the differences in consumer acceptance of the two varieties. The aroma of IR64 grains has previously been evaluated as being high in sweet taste, and low in astringent and water-like metallic attributes, whereas Apo was described as high in sewer/animal, water-like metallic, astringent and sour/silage and low in sweet taste (Champagne, Bett-Garber et al. 2010). In that study, the rice had been grown in irrigated conditions in the same country as the rice evaluated in the present research, but 4 years earlier than this study, and in a different region of the island of Luzon. The results presented here for the samples from irrigated treatments accord well with those of the previous study. In addition, we found high levels of hay-like/musty flavours in Apo that were not previously reported by Champagne et al. (2010). Slight differences between our results with those of previous reports may derive from the harvest and storage conditions of the samples tested, but in general, the aromatic qualities of each variety are reproducible.

Both the sensory analysis and aromatic profile of IR64 were affected by the drought treatment. Biochemical changes in the grain may explain much of this, either due to changes in gene expression in the grain, or changes to the regulation of source-sink transport rom the plant to the grain. In IR64 grains from the drought treatment, sweetness was decreased and a water-like metallic flavour was detected which was absent in grains from the irrigated treatment. Though we cannot demonstrate the role of the underlying metabolic processes in this study, we have linked the stress caused by drought to changes in the metabolite pattern, and therefore the flavour profile of IR64. Samples of Apo from both irrigated and drought fields were indistinguishable in their characterisation by water-like metallic, sewer/animal, and hay-like/musty aromas at high levels. One previous study has investigated the effect of drainage on the sensory aroma of rice and the variety used in that study showed no difference in flavour/aroma compounds (Champagne, Bett-Garber et al. 2005). Our data suggest that a response to moisture stress in the metabolomic profile of the grain is variety dependent, as is often observed in crops e.g., fruit response to water stress that is cultivar dependent (Giné Bordonaba and Terry 2010). Understanding different varietal responses and determining the basis of stress response is important in developing climate-ready varieties.

Sensory analysis correlates with volatile metabolite profiles

Our analysis has highlighted a variety x treatment interaction, with IR64 grains showing greater environmental sensitivity than Apo grains. The differences in IR64 according to

differences in water availability, highlight how water stress not only affects grain yield, but also the metabolites present in the grains, and we found this to be reflected in the responses from sensory profiling. Water stress changed a number of the metabolites detected in IR64; by contrast, volatile metabolites present in Apo showed less variation between the water treatments, an outcome also reflected in the uniform characterisation of Apo rice by the sensory panel. The clear differences between Apo and IR64 in their volatile profiles, which are supported by the findings of the sensory panel, reinforces that metabolite profiling is indeed capable of distinguishing between varieties (Calingacion, Boualaphanh et al. 2012), providing a tool to identify and potentially quantify important compounds of quality.

Aromatic volatile compounds can only be considered important in rice quality if the odour threshold, which defines the lowest concentration of a given compound that humans can perceive, is low (Leonardos, Kendall et al. 1969). Of the 187 compounds detected, 65 have been putatively annotated as aromatic with a low odour threshold. The PCA of this shortlist of 65 closely reflected the PCA of 187 compounds, suggesting that much of the variation between varieties is captured with less, but more relevant data. Statistical analysis was further able to identify those flavour compounds that are most influential in discriminating between Apo and IR64. Samples of IR64 from the irrigated treatment formed a distinct cluster in the first axis, with indole as the compound most strongly correlated with that dimension (Table 2). Indole has a sweet flavour and is used at low concentrations to form synthetic sweet floral aromas (Jezussek, Juliano et al. 2002), so it is a promising candidate for the sweeter taste perceived in the grains of IR64 from the irrigated treatment, and in agreement its previous detection at concentrations above its odour threshold in two aromatic and one non-aromatic rice cultivar grown in France (Maraval, Mestres et al. 2008).

Samples of IR64 from the drought treatment were discriminated by strong loadings of octanal and decanal, which are both described as having a low odour threshold of fruity or waxy odours (Jezussek, Juliano et al. 2002, Yang, Shewfelt et al. 2008). Hexanal, heptanal, octanal and decanal were also present, as were their primary alcohols and several alkanes, possibly representing oxidation ladders (Monsoor, Proctor et al. 2004), with beneficial outcomes for the aromatic profile of the rice.

Sensory differences between irrigated and drought IR64 identified an astringent aroma (Table 1). Phenol is present in IR64 drought and Apo grains (Table 2), and this is a compound commonly associated with astringency (Bahar and Altug 2009). Undesirable

aromas such as sewer animal and sour/silage were found in Apo alone, often/typically associated with the presence of propene thiol and ocimene (Loper, Flath et al. 1971, O'Neill and Phillips 1992). Hay-like/musty off-flavour aromas were most strongly detected by the sensory evaluation in drought-affected IR64 and both samples of Apo. Hexanal and heptanal could explain those aromas (Lam and Proctor 2003). The strong similarity between the sensory data of the samples and the pattern of loadings in Figure 2b indicates that delving deeper into metabolomics to explain aroma could lead to the development of a new suite of tools to select for aromatic profiles.

Genetic difference of Apo and IR64 valuable to mapping population

Our study again shows that over a hundred volatile metabolites are present in rice and, while only a handful may be key to characterising and defining rice aroma, their relative importance needs to be better understood to enable breeders to select for this trait (Buttery, Turnbaugh et al. 1988, Jezussek, Juliano et al. 2002, Laguerre, Mestres et al. 2007, Yang, Lee et al. 2010, Bryant and McClung 2011, Calingacion, Boualaphanh et al. 2012, Mathure, Jawali et al. 2014). However, little is known about the genetic basis of aroma in rice, or the inheritance patterns of these key flavour compounds, aside from 2-acetyl-1-pyrroline, which is associated with the popcorn-like aroma of fragrant rices (Bradbury, Fitzgerald et al. 2005, Chen, Yang et al. 2008, Kovach, Calingacion et al. 2009). Here we have taken substantial steps in identifying additional specific compounds associated with favourable and unfavourable consumer response to rice aromatic quality. With a better understanding of the flavour traits that may influence rice acceptance, and tying these to the specific compounds that drive this acceptance, we gain the potential to take advantage of the genetic difference between Apo and IR64 (Figure 3) for discovery of causative loci. SNP genotyping showed 32% polymorphism between Apo and IR64. These allelic differences, mainly on chromosomes 1, 8 and 11, could present a testable explanation of their divergent response to water stress and sensory attributes such as aroma. Chromosome 1 has been previously reported to have qDTY1.1, a large-effect QTL for grain yield under drought (Ghimire, Quiatchon et al. 2012), and chromosome 3 has been shown to house another large effect QTL for drought qDTY3.1 and this has previously been identified in Apo (Kumar, Dixit et al. 2014). Development of a mapping population between Apo and IR64 could facilitate the identification of quantitative trait loci (QTLs) associated with specific components of rice aroma. Furthermore this work makes it possible to develop varieties that take the favourable aromatic/flavour characteristics of one variety, and at the same time select for the drought

tolerant traits of another if the quality of IR64 can be captured into the agronomically favourable background of Apo.

Conclusions

A panel of compounds identified here that are strongly associated with the pleasant cooking and eating experience of IR64 are indole, octanal, decanal, and phenol, which give sweet, fruity aromas, while Apo is identified with off-flavour compounds such as propene thiol, ocimene, heptanal and hexanal which give sewer/ animal, sour/silage and hay-like/musty aromas. These compounds all have low odour thresholds and describe the aromas detected by sensory analysis in this and previous studies. To identify the genes underlying these flavour notes, requires identifying QTLs associated with the specific compounds that appear to drive consumer responses through the use of a mapping population.

Acknowledgments

The authors thank F Salisi, R Ilagan, L Samadio and A Madrid for their assistance in conducting the field experiment. MC also thanks the Monsanto Beachell-Borlaug International Scholarship Program and University of Queensland for additional funding.

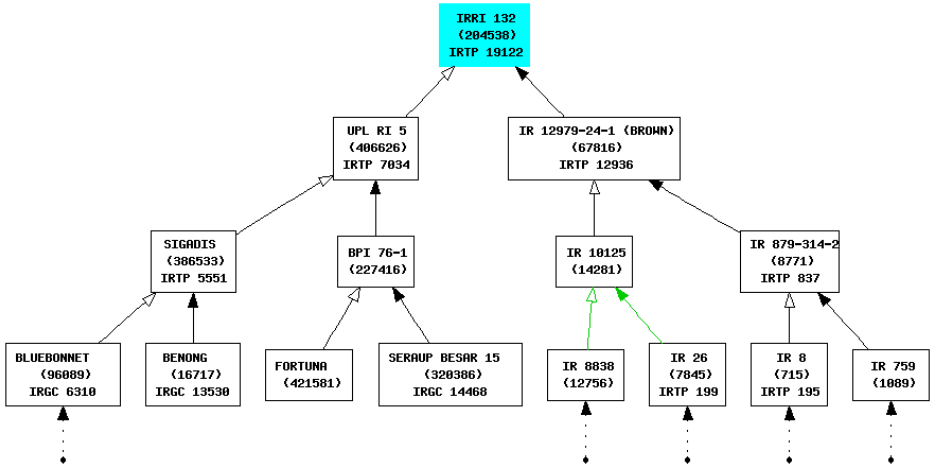
Supporting Information

Table S1. Log-likelihood ratio tests: effect of removing fixed effect terms for linear mixed models of yield (A), PC1 (B) and PC2 (C).

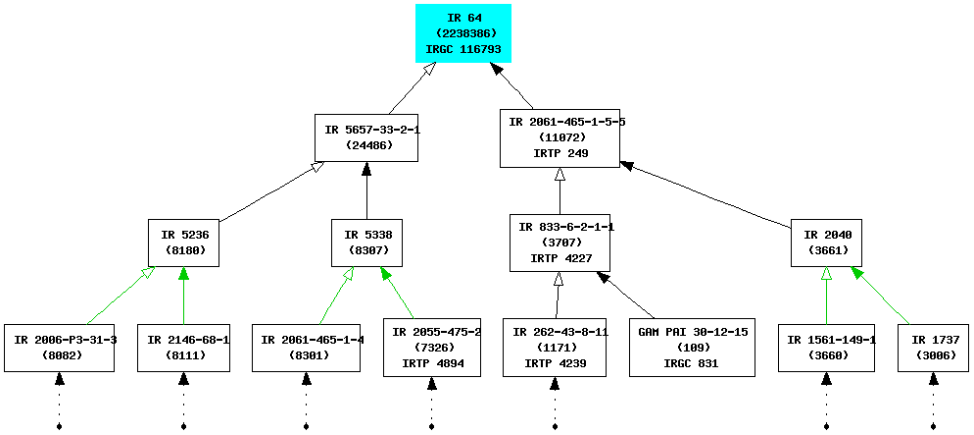
A				
Full model	Nested model	X^2	Df	p
Variety x Environment	Variety + Environment			
-332.54	-333.96	2.8468	1	0.09156
Variety + Environment	Environment			
-333.96	-337.24	6.5484	1	0.0105
Variety + Environment	Variety			
-333.96	-344.39	20.855	1	4.953e-06
B				
Full model	Nested model	X^2	Df	p
Variety x Environment	Variety + Environment			
-133.59	-144.72	22.271	1	2.368e-06
C				
Full model	Nested model	X^2	Df	p
Variety x Environment	Variety + Environment			
-118.17	-119.21	2.0699	1	0.1502
Variety + Environment	Environment			
-119.21	-164.67	90.93	1	< 2.2e-16
Variety + Environment	Variety			
-119.21	-119.65	0.8744	1	0.3497

Figure S1. Pedigree tree of (A) Apo (IRRI 132) and (B) IR64 showing no common parents back at least three generations (www.irri.org/iris).

A



B



References

Bahar, B. and T. Altug (2009). Flavour characterization of sumach (*Rhus coriaria* L.) by means of GC/MS and sensory flavour profile analysis techniques. International Journal of Food Properties **12**(2): 379-387.

Bates, D., M. Maechler, B. Bolker and S. Walker (2014). lme4: Linear mixed-effects models using Eigen and S4. R package.

Bett-Garber, K., E. Champagne, J. Thomson and J. Lea (2012). Relating raw rice colour and composition to cooked rice colour. Journal of the Science of Food and Agriculture **92**(2): 283-291.

Boualaphanh, C., M. Calingacion, R. Cuevas, D. Jothityangkoon, J. Sanitchon and M. Fitzgerald (2011). Yield and quality of traditional and improved Lao varieties of rice. Scienceasia **37**(2): 89-97.

Bradbury, L., T. Fitzgerald, R. Henry, Q. Jin and D. Waters (2005). The gene for fragrance in rice. Plant Biotechnology Journal **3**(3): 363-370.

Bryant, R. and A. McClung (2011). Volatile profiles of aromatic and non-aromatic rice cultivars using SPME/GC–MS. Food Chemistry **124**(2): 501-513.

Buttery, R., J. Turnbaugh and L. Ling (1988). Contribution of volatiles to rice aroma. Journal of Agricultural and Food Chemistry **36**(5): 1006-1009.

Calingacion, M., C. Boualaphanh, V. Daygon, R. Anacleto, R. Sackville Hamilton, B. Biais, C. Deborde, M. Maucourt, A. Moing, R. Mumm, R. de Vos, A. Erban, J. Kopka, T. Hansen, K. Laursen, J. Schjoerring, R. Hall and M. Fitzgerald (2012). A genomics and multi-platform metabolomics approach to identify new traits of rice quality in traditional and improved varieties. Metabolomics **8**(5): 771-783.

Calingacion, M., A. Laborte, A. Nelson, A. Resurreccion, J. Concepcion, V. Daygon, R. Mumm, R. Reinke, S. Dipti, P. Bassinello, J. Manful, S. Sophany, K. Lara, J. Bao, L. Xie, K. Loaiza, A. El-hissewy, J. Gayin, N. Sharma, S. Rajeswari, S. Manonmani, N. Rani, S. Kota, S. Indrasari, F. Habibi, M. Hosseini, F. Tavasoli, K. Suzuki, T. Umemoto, C. Boualaphanh, H. Lee, Y. Hung, A. Ramli, P. Aung, R. Ahmad, J. Wattoo, E. Bandonill, M. Romero, C. Brites, R. Hafeel, H. Lur, K. Cheaupun, S. Jongdee, P. Blanco, R. Bryant, N. Thi Lang, R. D. Hall and M. Fitzgerald (2014). Diversity of global rice markets and the science required for consumer-targeted rice breeding. PLoS One **9**(1): e85106.

Champagne, E., K. Bett-Garber, M. Fitzgerald, C. Grimm, J. Lea, K. Ohtsubo, S. Jongdee, L. Xie, P. Bassinello, A. Resurreccion, R. Ahmad, F. Habibi and R. Reinke (2010). Important sensory properties differentiating premium rice varieties. Rice **3**: 270–281.

Champagne, E., K. Bett-Garber, J. Thompson, R. Mutters, C. Grimm and A. McClung (2005). Effects of Drain and Harvest Dates on Rice Sensory and Physicochemical Properties. Cereal Chem. **82**(4): 369–374.

Chen, S., Y. Yang, W. Shi, Q. Ji, F. He, Z. Zhang, Z. Cheng, X. Liu and M. Xu (2008). Badh2, encoding betaine aldehyde dehydrogenase, inhibits the biosynthesis of 2-acetyl-1-pyrroline, a major component in rice fragrance. The Plant Cell **20**(7): 1850-1861.

Cramer, G., E. Wailes and S. Shui (1993). Impacts of liberalizing trade in the world rice market. American Journal of Agricultural Economics **75**(1): 219-226.

Daygon, V. and M. Fitzgerald (2013). Application of metabolomics for providing a new generation of selection tools for crop improvement. Hot Topics in Metabolomics: Food and Nutrition, Future Science Ltd**:** 6-16.

Dejean, S., I. Gonzalez, K. Le Cao, P. Monget, J. Coquery, F. Yao, B. Liquet and F. Rohart (2013). mixOmics: Omics data integration project. R package.

Fanzo, J. and P. Pronyk (2011). A review of global progress toward the Millennium Development Goal 1 Hunger Target. Food & Nutrition Bulletin **32**(2): 144-158.

Fitzgerald, M., S. McCouch and R. Hall (2009). Not just a grain of rice: the quest for quality. Trends in Plant Science **14**(3): 133-139.

Ghimire, K., L. Quiatchon, P. Vikram, B. Swamy, S. Dixit, H. Ahmed, J. Hernandez, T. Borromeo and A. Kumar (2012). Identification and mapping of a QTL (qDTY1.1) with a consistent effect on grain yield under drought. Field Crops Research **131**: 88-96.

Gift, N., I. Gormley, L. Brennan and t. R. C. team (2010). MetabolAnalyze: probabilistic principal components analysis for metabolomic data. R package.

Giné Bordonaba, J. and L. Terry (2010). Manipulating the taste-related composition of strawberry fruits (Fragaria × ananassa) from different cultivars using deficit irrigation. Food Chemistry **122**(4): 1020-1026.

Goodman, D. and R. Robison (2013). Economic development, social status and political consciousness. New York, Routledge.

Guan, Y., R. Serraj, S. Liu, J. Xu, J. Ali, W. Wang, E. Venus, L. Zhu and Z. Li (2010). Simultaneously improving yield under drought stress and non-stress conditions: A case study of rice (*Oryza sativa* L.). Journal of Experimental Botany **61**(15): 4145-4156.

Hall, R., I. Brouwer and M. Fitzgerald (2008). Plant metabolomics and its potential application for human nutrition. Physiologia Plantarum **132**(2): 162-175.

Husson, F., J. Josse, S. Le and J. Mazet (2013). FactoMineR: Multivariate exploratory data analysis and data mining with R. R package.

IPCC (2014). IPCC Fifth Assessment Report: Working Group II Report 'Impacts, adaptation and vulnerability'.

Jezussek, M., B. Juliano and P. Schieberle (2002). Comparison of key aroma compounds in cooked brown rice varieties based on aroma extract dilution analyses. Journal of Agricultural and Food Chemistry **50**(5): 1101-1105.

Kharas, H. and G. Gertz, Eds. (2010). The new global middle class: A crossover from West to East. China's emerging middle class: beyond economic transformation, Brookings Institution Press.

Knutson, T. and R. Tuleya (2004). Impact of CO2-induced warming on simulated hurricane intensity and precipitation: Sensitivity to the choice of climate model and convective parameterization. Journal of Climate **17** (18): 3477-3495.

Kovach, M., M. Calingacion, M. Fitzgerald and S. McCouch (2009). The origin and evolution of fragrance in rice (*Oryza sativa* L.). Proceedings of the National Academy of Sciences **106**(34): 14444-14449.

Kumar, A., J. Bernier, S. Verulkar, H. Lafitte and G. Atlin (2008). Breeding for drought tolerance: Direct selection for yield, response to selection and use of drought-tolerant donors in upland and lowland-adapted populations. Field Crops Research **107**(3): 221-231.

Kumar, A., S. Dixit, T. Ram, R. Yadaw, K. Mishra and N. Mandal (2014). Breeding high-yielding drought-tolerant rice: genetic variations and conventional and molecular approaches. Journal of Experimental Botany: 1-14.

Laguerre, M., C. Mestres, F. Davrieux, J. Ringuet and R. Boulanger (2007). Rapid discrimination of scented rice by solid-phase microextraction, mass spectrometry, and multivariate analysis used as a mass sensor. Journal of Agricultural and Food Chemistry **55**(4): 1077-1083.

Lam, H. and A. Proctor (2003). Milled rice oxidation volatiles and odor development. Journal of Food Science **68**(9): 2676-2681.

Leonardos, G., D. Kendall and N. Barnard (1969). Odor threshold determinations of 53 odorant chemicals. Journal of the Air Pollution Control Association **19**(2): 91-95.

Loper, G., R. Flath and J. Webster (1971). Identification of Ocimene in Alfalfa Flower Aroma by Combined GC-Mass Spectrometry1. Crop Sci. **11**(1): 61-63.

Maraval, I., C. Mestres, K. Pernin, F. Ribeyre, R. Boulanger, E. Guichard and Z. Gunata (2008). Odor-active compounds in cooked rice cultivars from Camargue (France) analyzed by GC−O and GC−MS. Journal of Agricultural and Food Chemistry **56**(13): 5291-5298.

Mathure, S., N. Jawali, R. Thengane and A. Nadaf (2014). Comparative quantitative analysis of headspace volatiles and their association with BADH2 marker in non-basmati scented, basmati and non-scented rice (*Oryza sativa* L.) cultivars of India. Food Chemistry **142**: 383-391.

Monsoor, M., A. Proctor and T. Siebenmorgen (2004). Surface lipid and free fatty acids (FFA) content of head and broken rice produced by milling after different drying treatments. Cereal Chemistry **81**(6): 705–709.

Murray, M. and W. Thompson (1980). Rapid isolation of high molecular weight plant DNA. Nucleic Acids Research **8**(19): 4321-4326.

Nations, U. (2014). World Population Prospects: The 2012 Revision P. D. Department of Economic and Social Affairs. New York, United Nations. Working Paper No. ESA/P/WP.235.

O'Neill, D. and V. Phillips (1992). A review of the control of odour nuisance from livestock buildings: Part 3. Properties of the odorous substances which have been identified in livestock wastes or in the air around them. Journal of Agricultural Engineering Research **53**: 23-50.

Palanog, A., B. Swamy, N. Shamsudin, S. Dixit, J. Hernandez, T. Boromeo, P. Cruz and A. Kumar (2014). Grain yield QTLs with consistent-effect under reproductive-stage drought stress in rice. Field Crops Research **161**: 46-54.

Peng, S., J. Huang, J. Sheehy, R. Laza, R. Visperas, X. Zhong, G. Centeno, G. Khush and K. Cassman (2004). Rice yields decline with higher night temperature from global warming. Proceedings of the National Academy of Sciences **101**(27): 9971-9975.

Quinn, G. and J. Keough (2002). Experimental design and data analysis for biologists, Cambridge Univesity Press.

RCoreTeam (2014). R package.

Saxton, A. (2004). Genetic analysis of complex traits using SAS. SAS. SAS Institute, Inc., Cary, NC.

Trenberth, K. (2005). Uncertainty in hurricanes and global warming. Science **308**(5729): 1753-1754.

van Berloo, R. (2008). GGT 2.0: Versatile software for visualization and analysis of genetic data. Journal of Heredity **99**(2): 232-236.

van den Berg, R., H. Hoefsloot, J. Westerhuis, A. Smilde and M. van der Werf (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data." BMC Genomics **7**(142).

Venuprasad, R., M. Bool, L. Quiatchon and G. Atlin (2012). A QTL for rice grain yield in aerobic environments with large effects in three genetic backgrounds. Theoretical and Applied Genetics **124**(2): 323-332.

Venuprasad, R., C. Dalid, M. Del Valle, D. Zhao, M. Espiritu, M. Sta Cruz, M. Amante, A. Kumar and G. Atlin (2009). Identification and characterization of large-effect quantitative trait loci for grain yield under lowland drought stress in rice using bulk-segregant analysis. Theoretical and Applied Genetics **120**(1): 177-190.

Venuprasad, R., H. Lafitte and G. Atlin (2007). Response to direct selection for grain yield under drought stress in rice. Crop Science **47**(1): 285-293.

Vikram, P., B. Swamy, S. Dixit, H. Ahmed, M. Teresa Sta Cruz, A. Singh and A. Kumar (2011). QDTY1.1, a major QTL for rice grain yield under reproductive-stage drought stress with a consistent effect in multiple elite genetic backgrounds. BMC Genetics **12**: 89-104.

Wassmann, R., S. Jagadish, S. Heuer, A. Ismail, E. Redona, R. Serraj, R. Singh, G. Howell, H. Pathak and K. Sumfleth (2009). Climate change affecting rice production: The physiological and agronomic basis for possible adaptation strategies. Advances in Agronomy. D. Sparks, Academic Press. **101:** 59-122.

Xu, K., X. Xu, T. Fukao, P. Canlas, R. Maghirang-Rodriguez, S. Heuer, A. Ismail, J. Bailey-Serres, P. Ronald and D. Mackill (2006). Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. Nature **442**(7103): 705-708.

Yang, D., K. Lee and S. Kays (2010). Characterization and discrimination of premium-quality, waxy, and black-pigmented rice based on odor-active compounds. Journal of the Science of Food and Agriculture **90**(15): 2595-2601.

Yang, D., R. Shewfelt, K. Lee and S. Kays (2008). Comparison of odor-active compounds from six distinctly different rice flavor types. Journal of Agricultural and Food Chemistry **56**(8): 2780-2787.

Zeigler, R. and A. Barclay (2008). The relevance of rice. Rice **1**(1): 3-10.

Zeng, H., Y. Zhong and L. Luo (2006). Drought tolerance genes in rice. Functional & Integrative Genomics **6**(4): 338-341.

Chapter 5

A multidisciplinary phenotyping and genotyping analysis of a mapping population from an Apo x IR64 cross enables quality to be combined with yield in rice

M. Calingacion[1,2,3], L. Quiatchon-Baeza[1,4], K. Tan[3], S. Prakash[3], R. Mumm[5], M. Fitzgerald[3], and R. D. Hall[2,5]

Submitted to Rice

[1]International Rice Research Institute, Los Banos, Laguna, Philippines; [2]Laboratory of Plant Physiology, Wageningen University, Wageningen, The Netherlands; [3]School of Agriculture and Food Science, University of Queensland, St Lucia 4072 Qld, Australia; [4]Department of Crop Sciences, College of ACES, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA; [5]BU Bioscience, Plant Research International, Wageningen, The Netherlands.

Abstract

In this study an F_8 mapping population of a cross between the commercial rice varieties Apo and IR64 has been used to both genotype and phenotype approximately 200 individual progeny. A genotype-by-sequencing approach was first used to identify 2681 polymorphic SNP markers which gave dense coverage of the genome with a good distribution across all 12 chromosomes with few small gaps. The coefficient of parentage was also very low, at 0.13, confirming that the parents were genetically distant from each other. All these lines, together with both parents, were grown under irrigated and drought-induced conditions in a random block design and all grain was harvested to determine variation in yield across the population. The grains were then polished following standard procedures prior to performing the phenotyping analyses. A Gas Chromatography - Mass Spectrometry approach was used to determine the volatile biochemical profiles of each line and after data curation and processing, discriminatory metabolites were putatively identified based on in house and commercial spectral libraries. These data were used to predict their potential role in determining aroma differences between genotypes. A number of QTLs for yield and individual metabolites could be identified.The same grain materials were also used in a sensory analysis performed by a trained sensory panel. Clear biochemical and sensory differences were observed between the parents with IR64 generally being described in terms of favourable aroma attributes such as corn and sweet aromatic attributes while Apo was found to have musty and sour silage notes. Following these analyses it proved possible to identify a number of lines which appeared to combine the favourable aroma attributes of IR64 with the favourable (higher) yield potential of Apo. As such, these lines are excellent candidates to assess further as potential genotypes to work up into a new variety of rice which has both good yield and good quality potential, thus meeting the needs of both farmer and consumer alike.

Introduction

Flavour, the blend of aroma and taste, is one of the most important factors influencing the quality of rice as perceived by the consumer, and consequently, flavour dictates to a great extent, consumer preference (Del Mundo and Juliano 1981, Fitzgerald, McCouch et al. 2009). Aroma is generally the major contributor to the overall flavour in rice (Cho, Nuijten et al. 2014). As a consequence, the identification of compounds contributing to rice aroma, the factors affecting the aroma profile as well as the genetic basis of its main components have been major research concerns. Such goals are not only limited to rice breeding programs but also are of general relevance to research programs on many food crops where

advanced instrumentation for the detection of volatile aromatic compounds is becoming more widely used (Hall, Brouwer et al. 2008).

Over 100 volatile compounds have been reported to be present in rice, while only few may be key to characterizing and defining rice aroma and in particular, odour threshold (human limit of detection), which plays a major part in whether individual compounds play a role in the aroma phenotype (Buttery, Turnbaugh et al. 1988, Jezussek, Juliano et al. 2002, Yang, Lee et al. 2010, Bryant and McClung 2011, Calingacion, Boualaphanh et al. 2012, Mathure, Jawali et al. 2014). However, little is known about the genetic basis of aroma in rice, or the inheritance patterns of any of these key flavour compounds. The only exception to this is 2-acetyl-1-pyrroline (2AP), which is associated with the popcorn-like / floral aroma of fragrant rices and for which the genetic background has, at least in part, been elucidated and established (Bradbury, Fitzgerald et al. 2005, Chen, Yang et al. 2008, Kovach, Calingacion et al. 2009).

Clearly the aroma of rice is not as simple as the presence or absence of 2AP and thus is not the consequence of just a single determinant. There are many other naturally-occurring compounds that are volatile, or which can arise from lipid oxidation products or which occur as a result of Maillard reactions, thus contributing to complex aroma profiles comprising multiple aromatic compounds, including alcohols, alkanes, alkenes, aldehydes and other reductones. While these compounds may only be present at low concentrations, nevertheless, many have also a low odour threshold and hence can play an influential role in detectable aroma (Demyttenaere, Macura et al. 2003). Furthermore, it is well known that sensory panels are readily able to distinguish between the aromas of different varieties of cooked rice (Champagne, Bett-Garber et al. 2010) using descriptors of both pleasant and unpleasant notes. For example, IR64, a popular variety of rice has been described as having a pleasant aroma, whereas Apo, a less popular variety has been described as having a number of less pleasant aroma attributes (Champagne, Bett-Garber et al. 2010).

IR64 is variety of rice that is popular with consumers in many Asian countries (Mackill, Ismail et al. 2012). Ever since its release in 1985 it has been grown annually on more than one million hectares of land (Fitzgerald, McCouch et al. 2009). IR64 has been adopted by farmers and accepted by consumers, mainly due to its excellent eating quality (Champagne, Bett-Garber et al. 2010). However, IR64 is susceptible to a number of abiotic and biotic stresses which can significantly limit yield potential and entail seasonal risks to production

(Venuprasad, Lafitte et al. 2007). In contrast, Apo is tolerant to many stresses, especially drought, and is reported to give high yield under the drought conditions characteristic of upland areas as well as in lowland, well-watered areas. This aspect of 'reliable' yield is of growing importance in rice production. If we could capture the favoured agronomic and quality traits of both varieties in a single genotype, this could lead to the release of a new variety that would be valuable to farmers and consumers alike, in many areas of Asia. We have previously shown that the volatile metabolomes of Apo and IR64 grains contain many metabolites that associate with the sensory descriptors of each, in both positive and negative terms, and we have developed a panel of compounds that discriminate between the aromatic quality of IR64 and Apo (Calingacion, Fang et al. 2014).

Crop improvement programs are moving with increasing rapidity towards using customized genotyping techniques for progeny selection (Collard and Mackill 2008). Identification of genetic markers has become significantly less complex with the rapid evolution of genotyping technologies, such as the one million SNP chip, and genotyping-by-sequencing. Methods to phenotype specifically for aroma have also been developed in recent years, with metabolomic profiling now reaching a stage where it can be reliably used as an advanced phenotyping tool for objectives such as understanding plant aroma and food flavour (Dunemann, Ulrich et al. 2009, Mathieu, Cin et al. 2009, Inui, Tsuchiya et al. 2013).

For this study we have used a mapping population derived from IR64 and Apo, and have conducted a genotyping-by-sequencing approach to characterize the progeny. We have used this population to address the following objectives: (i) to screen aroma of the progeny by a combination of sensory profiling (sensory panel analysis) and metabolomic profiling using gas chromatography with mass spectrometry (GCMS); (ii) to identify QTLs that associate with the yield of Apo; (iii) to identify QTLs that associate with the major discriminatory metabolites of low odour threshold and that may help define the flavour characteristics of each variety; and (iv) to identify lines displaying the yield potential of Apo combined with the metabolomic profile and grain quality of IR64.

Materials and Methods

Plant material

Apo, IR64 and 213 recombinant-inbred lines (F_5) derived from a cross between Apo and IR64 were planted at the Experimental Station of the International Rice Research Institute (IRRI), Philippines in the dry season of 2011. The plants were carefully monitored, and a

single panicle (F_6) was harvested from each. The selected samples were then planted at IRRI in the next season for seed increase.

During the dry season of 2012, 150 seeds of each of the 213 lines (F_7) from seed increase of previous season were sown in seedling trays. Before transplanting, inorganic fertilizer, nitrogen: phosphorus: potassium (NPK) was applied to the field at a ratio of 40:40:40 kg ha^{-1}. Seedlings were transplanted in 6 blocks of 100 plants each in a random block design, with 3 blocks under irrigated and 3 blocks under drought conditions. For each rice sample, one plant per hill was planted at a spacing of 15cm between and within rows. After transplanting, the plants were top-dressed with urea after 30 and 55 days at a level of 30:0:0 kg ha^{-1}. A small piece of leaf material was taken from one plant in the centre of each block grown under irrigation for extraction of DNA. For blocks under drought conditions, stress was artificially imposed by draining the field when the plants were at the maximum tillering stage so that drought stress overlapped with the reproductive stage of the plant. Irrigated blocks, on the other hand, were maintained at a water level of approximately 5 cm until harvest, at which time they were drained. Mature grains from plants in all blocks were harvested (F_8), yield was measured, and the grains were dried in an oven until a moisture content of 12-14% was reached for milling. Grains were dehulled (Otake FCY2 Dehusker, Oharu, Japan), polished in a paint shaker with aluminum oxide and cryo-ground (IKA A11b basic analytical mill) with liquid nitrogen. Samples were stored at -80^{0}C until further experimentation.

Genotyping

A. Coefficient of parentage

In order to determine the diversity of the rice varieties that were used in developing the mapping population, the coefficient of parentage (COP) was calculated between Apo and IR64 using the COP function in the International Rice Information System (IRIS) database (http://irri.org/tools-and-databases/international-rice-information-system).

B. DNA preparation and SNP scans

DNA was extracted from leaf tissue of Apo, IR64 and all 213 lines of the population using the modified CTAB DNA extraction method (Murray and Thompson 1980). DNA in the extracts was quantified and diluted to 50 ng uL^{-1} using a Thermo Scientific Nanodrop 1000. Genotyping by sequencing was conducted in sets of 96 samples per lane using an Illumina HiSeq at Cornell University (http://www.igd.cornell.edu/index.cfm/page/projects /GBS.htm). SNP calls were made using Nipponbare as reference. As the confidence level of calling the

heterozygote state was low, all were considered as missing data. Only 0.65% of the data points were heterozygotes. The sites were filtered at a maximum count of 170 of 213 which accounts for sites where 80% of the lines have a call and a minimum frequency of 0.25 for the minor allele. The above criteria resulted in 2,681 filtered SNPs which were used for QTL mapping. A circular archaeopteryx tree showing all lines, Apo and IR64 was generated using the cladogram function in Trait Analysis by the Association Evolution and Linkage (TASSEL) program (Bradbury, Zhang et al. 2007).

Metabolomic profiling of volatile compounds

A. Headspace extraction

Rice flour (1g) of each of the samples was placed in a 10 mL glass vial and capped. Volatile compounds in the headspace were collected by solid phase microextraction (SPME) using a 65-mm polydimethylsiloxane-divinylbenzene fiber (Supelco, Bellefonte, PA, USA), as previously described (Verhoeven, Jonker et al. 2011, Calingacion, Boualaphanh et al. 2012). The volatile compounds were thermally desorbed at 250°C by inserting the SPME fiber for 1 min into the GC injection port of a GC8000 instrument (Fisons Instruments, Cheshire, UK) with an HP-5 column (30 m x 0.25 mm id x 1.05 µm film thickness) in split mode. The temperature program started at 45°C and remained at this temperature for 2 min, was then increased by 5°C min^{-1} to 250°C, which was then maintained for 5 minutes. Mass spectra were acquired over the range 35 to 400 m/z (mass-to-charge ratio) at 2.8 scans sec^{-1}, with electron impact ionization at 70 eV (MD800 electron impact MS, Fisons Instruments, Cheshire, UK).

B. Data processing

Raw data from the GC-MS analyses were processed using MetAlign software (Lommen 2009) to extract and align mass signals with a signal-to-noise ratio of $\geq$ 3. Only mass signals that were present in at least ten samples were retained for analysis; all others were discarded. Signal redundancy per metabolite was removed by means of clustering and mass spectra were reconstructed as previously described (Tikunov, Laptenok et al. 2012). Metabolites were putatively identified by matching the mass spectra of obtained metabolites against in house as well as NIST08 (www.NIST.gov) and Wiley spectral libraries, and by comparison with retention indexes of reference standards published in the literature (Strehmel, Hummel et al. 2008). Data were mean-centered, log_2 transformed, and pareto scaled. The processed, relative quantities of volatile metabolites were subjected to Principle Components Analysis (PCA) using SIMCA-P 13.0 (Umetrics AB, Umea, Sweden). The

number of significant PCs was determined by cross-validation (Eriksson, Johansson et al. 2006).

Sensory evaluation of rice flour

For sensory evaluation using quantitative descriptive analysis (QDA), a subset of 26 samples was randomly selected in order to identify compounds that characterised the flavour of the parents. Six trained panelists participated in this study. Rice samples were prepared by placing rice flour (1 g) in 20 ml screw-capped vials. Samples were heated in a water bath at 80°C for 10 min and presented immediately to the panelists for aroma analysis in randomised order. Panelists opened the lid of the vial carefully and evaluated the presence of 10 aroma notes using the training reference standards based on the work of Champagne et al. (Champagne, Bett-Garber et al. 2010) (Table S1). Panelists quantitatively scored the presence and intensity of each aroma note based on a modified universal scale of flavour intensity (Meilgaard, Civille et al. 2007) (Table S2). Six samples were evaluated by all panelists at each session. A total of ten sessions were held, with a standard rice sample (commercially available, long grain, non-aromatic) given to each panelist in every session, as a blind sample, to measure consistency of the panelists across the sessions. Reference standards for each attribute were also available at all sessions (Table S1).

Marker-trait association and QTL mapping

After processing of the genotyping data and validating against known genes for amylose content and gelatinisation temperature (data not shown), QTL mapping could be carried out using a subset of 184 progeny and the parents Apo and IR64. The generated report and map file were used for QTL analysis by using composite interval mapping (CIM) using the QGene software V4.3.8 (Joehanes and Nelson 2008). The genetic distance between SNP markers was estimated from the physical map based on the genomic sequence available at GRAMENE (www.gramene.org), with genetic distance (cM) = Physical distance (kb)/250. CIM was performed using the standard model with a walk speed of 2 cM. Cofactor selection was set to auto. Permutation tests were performed for each trait with composite interval mapping and 1000 permutations (Churchill and Doerge 1994). Marker-trait association was conducted by using TASSEL program (Bradbury, Zhang et al. 2007). The filtered sites which were polymorphic among the parents were then used for association analysis using a general linear model (GLM). In this study, only QTLs with a significance threshold of $p < 0.0001$ ($-\log_{10}$p-value=3.0) identified for yield under irrigation and drought, and for discriminating metabolites were used. Genotypic and phenotypic data were used for QTL

mapping using Qgene software as described above. The identified QTLs were then named using CGSNL nomenclature (McCouch 2008).

Results

Genotyping

The data obtained from genotyping by sequencing (GBS) was assembled, annotated and filtered, and resulted

in 2,681 polymorphic SNPs. These give dense coverage of the genome, with very few gaps seen in any of the chromosomes (Figure 1). The calculated coefficient of parentage between Apo and IR64 is 0.13. Using all the genotype data, a circular archaeopteryx tree was constructed (Figure 2). This tree shows two main branches, with Apo in B and IR64 in A. There are similar numbers of progeny in each of these main branches A and B (116 and 97). After the first cluster break, giving the branches A and B, there are several sub-clusters. IR64 was in sub-cluster C along with another eight lines, while 10 lines were in sub-cluster O with Apo (Figure 2).

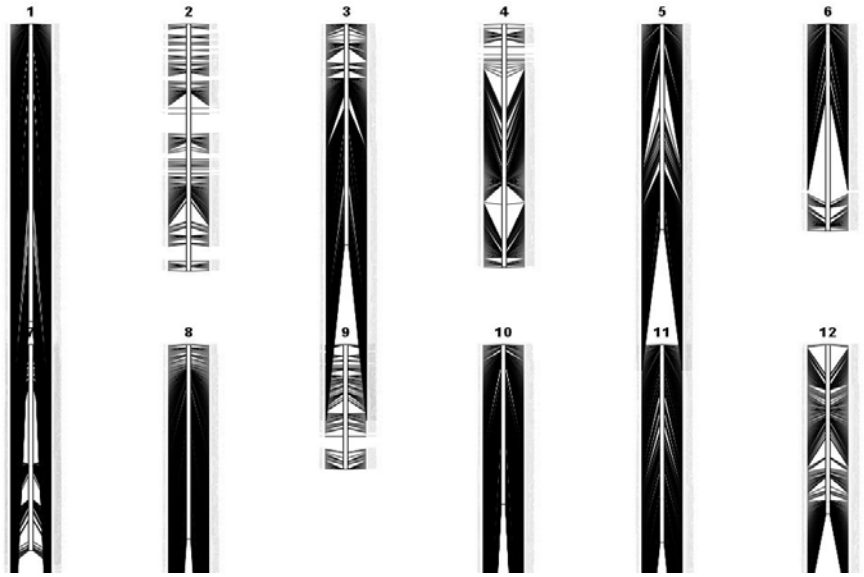


Figure 1. Genetic linkage map showing the distribution of 2681 polymorphic markers across the 12 chromosomes in the F8 recombinant-inbred lines derived from Apo and IR64, generated by QGene version 4.3.8 (Joehanes and Nelson 2008).

Yield under irrigated and drought conditions

The yield of Apo under irrigated and drought conditions was higher than the yield obtained from IR64 grown under the same conditions (Figure 3). Yield of more than half of the progeny was higher than yield of either Apo or IR64 by an average of 17% under both irrigation and drought (Figure 3). Line 83 (arrowed), which is in the same cluster as IR64 in Figure 2, has the second highest yield under irrigation and also gives a high yield under drought. The lines most genetically similar to either IR64 or Apo all show higher yield under drought than the parents, with Line 83 being the only one with a significantly higher yield under irrigation (Figure 3).

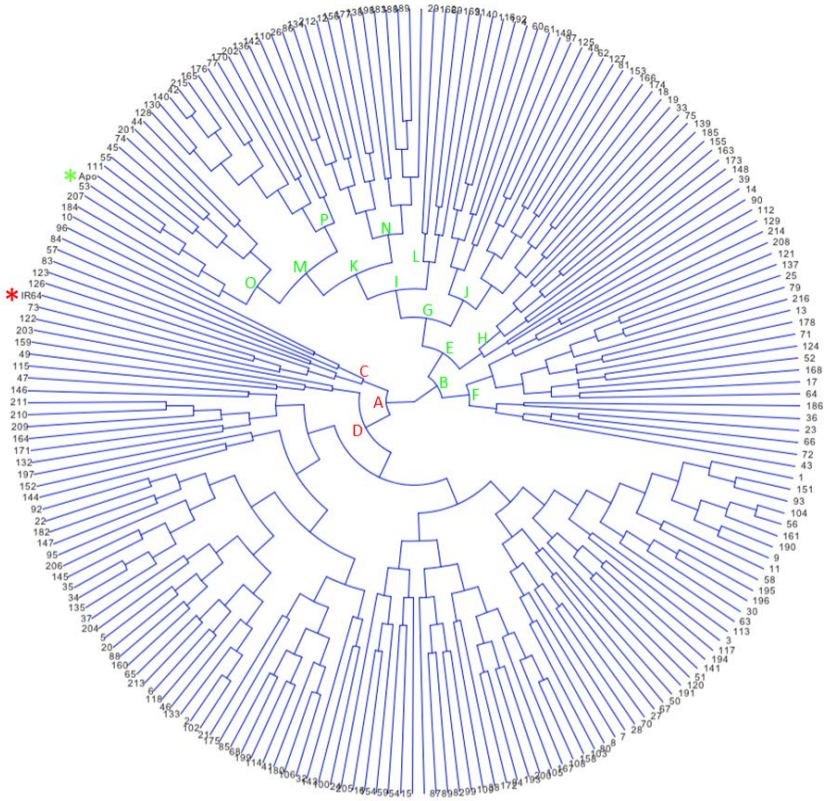


Figure 2. Circular archaeopteryx tree showing the 213 F8 recombinant-inbred lines of the cross between Apo and IR64, generated using Trait Analysis by the Association Evolution and Linkage (TASSEL) program (Bradbury, Zhang et al. 2007) showing two main cluster breaks associated with each parent and followed by lower level dissociation into several subclusters. Clusters are indicated in letters. Stars show Apo (green) and IR64 (red).

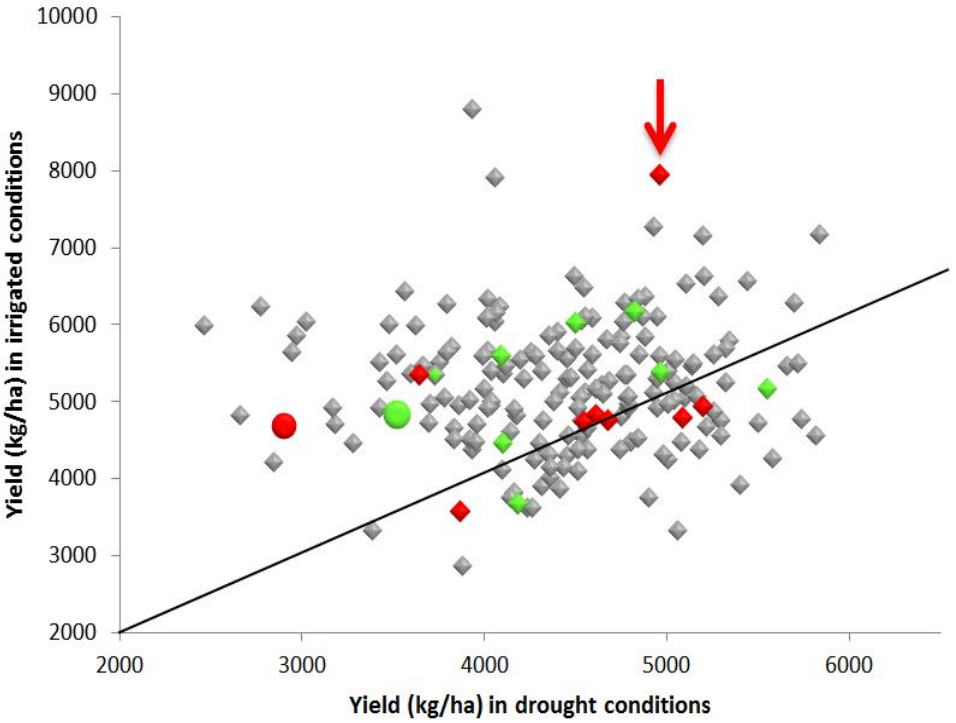


Figure 3. Yield (kg ha^{-1}) of 213 F8 inbred lines of the population grown under irrigated and drought conditions (Dry season 2012) at International Rice Research Institute. Samples of the parents, Apo (green) and IR64 (red) are shown in circles. Lines that fell in the same sub-cluster in Figure 2 as Apo are given in green and those falling in the same subcluster (Fig. 2) as IR64 are in red. Line 83 (arrowed red) has the second highest yield under irrigation and gives high yield under drought, and lies in the same cluster as IR64 in the archaeopteryx tree. A bar indicating the 1:1 ratio of yield under irrigated and drought is also shown.

One significant QTL, on the short arm of chromosome 3, was found for yield under drought in this population (Figure 4a). This QTL spanned the interval 1.3-13.3 cM of chromosome 3, flanked by SNP markers S3_346683 and S3_3337815. The QTL peak, at SNP marker S3_1849851, was found at 7.3 cM with a LOD score of 6.814 and F statistic score of 33.8. All progeny carrying this QTL had an average yield under drought of 4570 kg ha^{-1} which was significantly higher than the average yield of 4224 kg ha^{-1} for progeny without this QTL (X_2 = 10.919, df = 1, p = 0.0009).

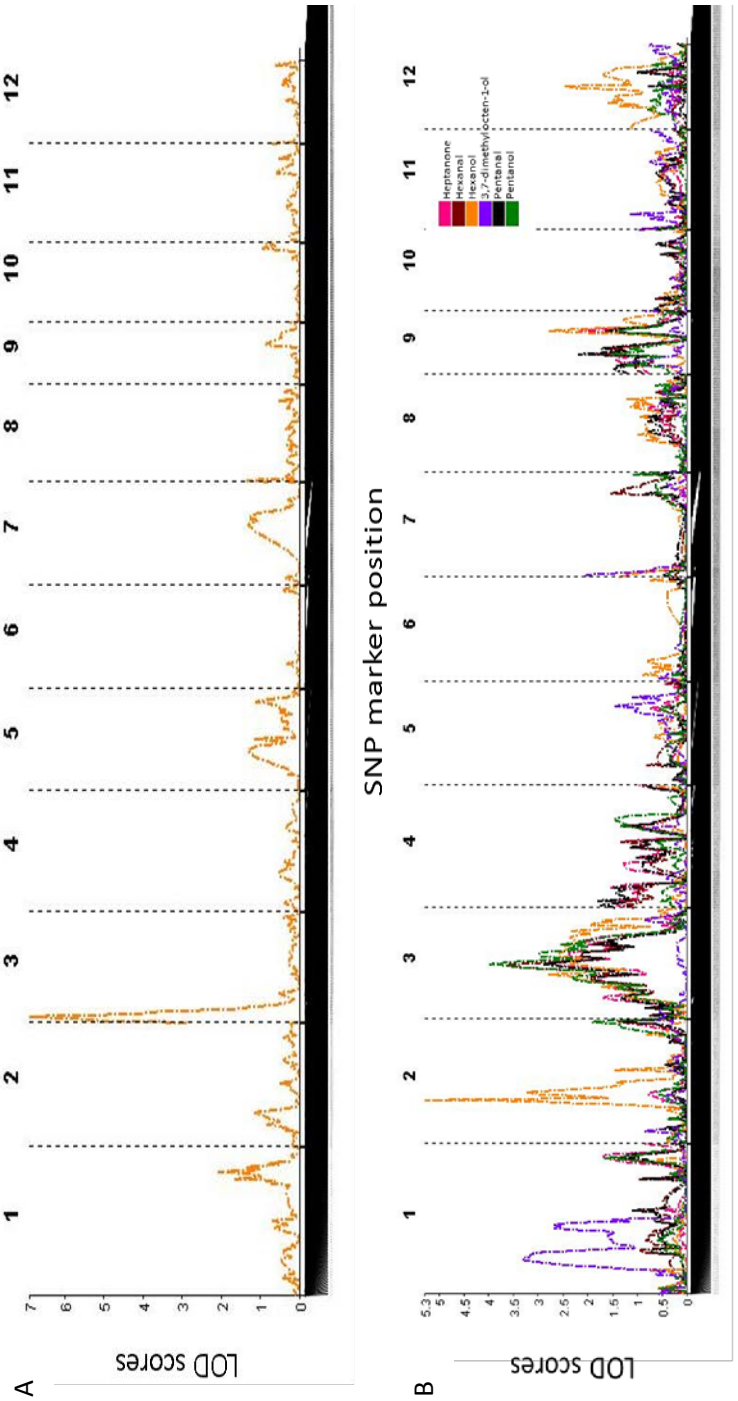


Figure 4. (A) LOD score curve denoting a strong QTL on chromosome 3 for yield under drought condition. (B) LOD score curve indicating metabolite QTLs on chromosome 1 for 3,7-dimethyl,7-octen-1-ol, chromosomes 2 and 3 for hexanol, and chromosomes 3 for heptanone, hexanal, pentanal and pentanol.

Grain Quality – Aroma

Using headspace sampling and GCMS, 105 compounds were detected in Apo, IR64 and subset of 184 lines of the population (Figure 5, Table 1). PC1 and PC2 explained 55.6% of the variation in the metabolite profiles with many of the lines clustering in between the Apo and IR64 parent values. The ten lines of the population that grouped with Apo in sub-cluster O and the eight that clustered with IR64 in sub-cluster C (Figure 2) did not cluster in the same way based on the metabolomic profile of the grains (Figure 5A). Interestingly, most of the lines in subcluster O based on genotype data (Figure 2), showed a metabolomic profile closer to that of IR64 than Apo. Most of the lines in subcluster C with IR64 (Figure 2) showed a metabolomic profile in between that of both the parents (Figure 5A).

Most of the compounds that were high in Apo and lines of the population that were clustered with Apo were alcohols, aldehydes, and ketones (Table 1). On the other hand, the compounds putatively identified as DL-2,3- butanediol and butanol were the compounds detected at high levels in both IR64 and lines of the population that were clustered with IR64 in Figure 5a (Table 1).

Metabolite QTLs (mQTLs) that are related to rice aroma were detected on chromosome 1, 2 and 3 (Figure 4B). The compound which could be annotated as 3,7-dimethyl-7-octen-1-ol was found to associate with the regions spanning the SNP markers S1_5944962 (23.7 cM) and S 1_14444337 (57.7 cM) positions. The QTL peak at chromosome 1 was mapped at 41.7 cM with an LOD score of 3.298 and F statistic score of 16.121 (Figure 4b, Table 1). The annotated compound as hexanol was found to be linked to regions spanning SNP markers S2_27187174 (28.7 cM) and S2_23698655 (94.7 cM). The peak QTL for the annotated compound hexanol was found at chromosome 2 at the position 50.7 cM with LOD score of 5.283 and F statistic score of 25.705. The annotated compound as hexanal associated with the regions spanning the SNP markers S3_12847023 (51.3 cM) and S3_16837834 (67.3 cM) positions. A QTL peak at chromosome 3 was mapped at 61.3 cM with an LOD score of 3.631 and F statistic score of 17.304 ($p<0.01$) (Figure 4b, Table 1). A QTL for the compound annotated as heptanone was also found in the regions encompassing the markers S3_12847023 (51.3 cM) and S3_16837834 (67.3 cM) positions, with the QTL peak located at 63.3 cM, providing a LOD score of 2.834 and F statistic score of 13.374 (Figure 4b, Table 1). The annotated compound pentanol was also mapped to the regions encompassing the markers S3_12847023 (51.3 cM) and S3_16837834 (67.3 cM) positions, with QTL peak located at 63.3 cM, providing a LOD score of 3.985 and F statistic

score of 19.078 (Figure 4b, Table 1). QTLs could not be found for either butanol or DL-2,3 butanediol, the two main compounds that clustered with IR64.

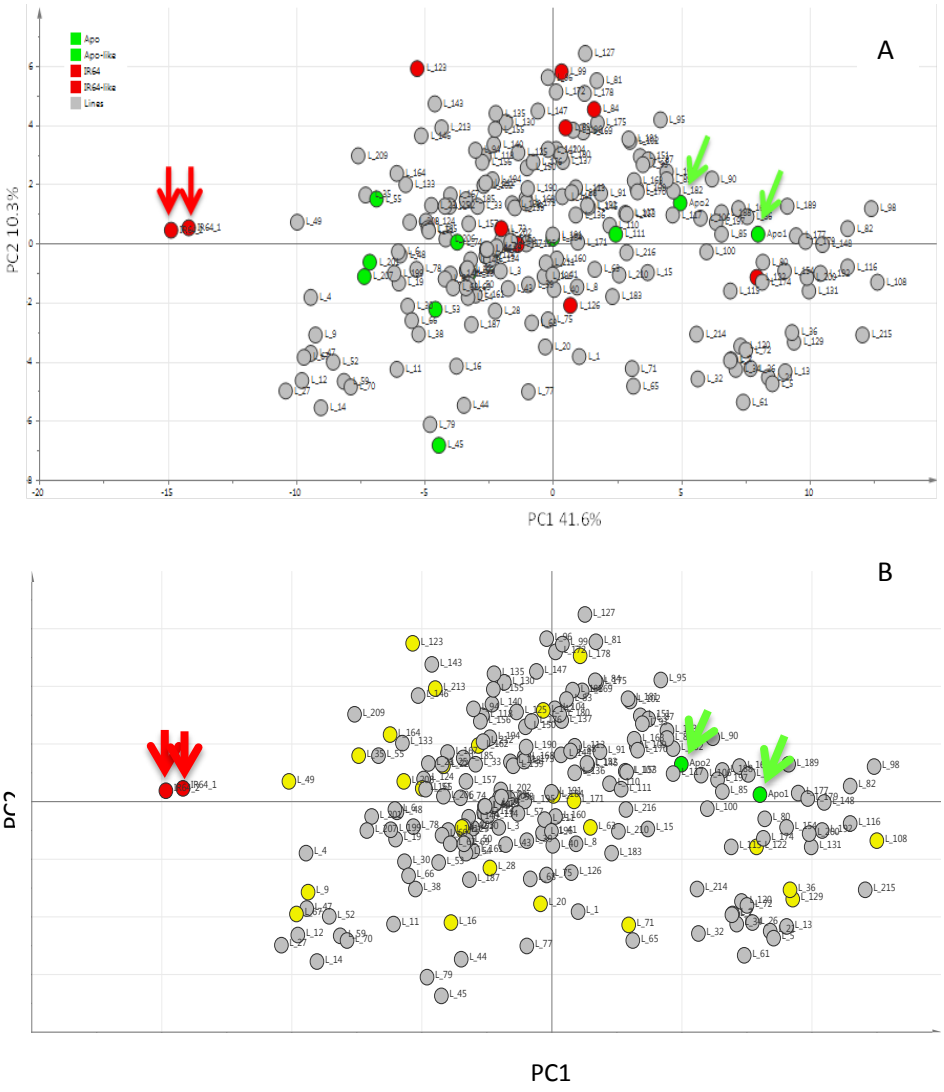


Figure 5. (A) Principal components analysis of metabolites detected in the headspace of Apo, IR64 and 184 inbred lines derived from Apo and IR64 that were grown under irrigated conditions. Arrows indicate the Apo replicates (green) and IR64 (red). Those lines in the same subcluster in the archaeopteryx tree with either the Apo or the IR64 are coloured green and red respectively (Fig.2); (B) The lines highlighted in yellow were selected for sensory analysis.

Table 1. List of putatively-identified metabolites detected in the headspace of Apo, IR64 and subset of 184 lines derived from a cross between Apo and IR64. Where known flavour attributes and QTLs we detected associated with the metabolites are also indicated. The metabolites have been grouped according to their chemical classes.

Compound	Aroma description	QTL
Alcohols		
1-Butanol	medicinal	
1-Pentanol	sweet	Chr 3 (63.3cM)
1-Hexanol	grassy	Chr 2 (50.7 cM)
		Chr 3 (51.3 cM)
1-Heptanol	citrus	
1-Octanol	citrus	
1-Octen-3-ol	mushroom	
2-Butoxyethanol		
6-Methyl-1-heptanol		
2-Ethyl-1-hexanol		
5-Methyl-2-(1-methylethyl)-1-hexanol		
1-Hepten-4-ol		
2-Butyl-1-octanol		
2-Methylpenten-3-ol		
3,7-Dimethyl-7-octen-1-ol		Chr1 (41.7cM)
Dimethylbenzenemethanol		
DL-2,3-butanediol	buttery	
Aldehydes		
Pentanal	floral	
Hexanal	grassy	Chr3 (61.3cM)
Heptanal	fruity	
Octanal	fatty	
Nonanal	floral	
Decanal	soapy	
Benzaldehdye	nutty	
(Z)-3-Hexenal	leaf-like	
(Z)-2-Heptenal	fatty, grassy	
(E)-2-Octenal	green, herby	
(E)-2-Nonenal	fatty	
(E)-2-Decenal	green	
3-Methylbutanal		

Table 1 continued.

Compound	Aroma description	QTL
Ketones		
2-Heptanone	fruity	Chr3 (63.3cM)
Cyclopentanone		
3-Penten-2-one	Fruity to fishy	
1-Cyclopropyl-1-propanone		
1-Decalone		
Acetophenone		
Aromatics		
2-Pentylfuran	nutty, beany	
p-xylene		
Indene		
Mesitylene		
Styrene		
Limonene	citrus	
Phenylethyne		
Naphthalene	mothball	
α-Phellandrene	spicy	
Hydrocarbons		
Decane		
Undecane		
Dodecane		
Tridecane		
Tetradecane		
Pentadecane		
Hexadecane		
Heptadecane		
Nonadecane		
Undecene		
1-Dodecene		
3-Dodecene		
3-Methylpentane		
3-Methylheptane		
2-Methylnonane		
5-Methylundecane		
2,4-Dimethylheptane		
2,5-Dimethlheptane		

Table 1 continued.

Compound	Aroma description	QTL
3,5-Dimethyloctane		
2,3-Dimethyloctane		
3,3-Dimethyloctane		
3,5-Dimethylundecane		
2,8-Dimethylundecane		
2,2,3-Trimethylhexane		
2,6,10-Trimethyldodecane		
2,2,3,4-Tetramethylpentane		
2,2,4,4,6,8,8-Heptamethylnonane		
3-Ethyl-2-methylheptane		
2,2'-diethyl-1,1'-Biphenyl		
1-Methylpropylcyclohexane		

Sensory evaluation

Twenty six of the 213 lines of the population were randomly selected for sensory evaluation (Figure 5B). Sensory profiling was carried out on these using 10 aroma attributes (Table S1). PCA of the 26 selected samples was performed based on the 10 sensory attributes. PC1 and PC2 explained 46.5% of the variation in the sensory profiles identified by the panelists. Notes of hay-like/musty and sour silage clustered together in the PCA biplot jointly showing the correlation structure of the samples and sensory attributes (Figure 6). Grassy, sewer animal, and grainy/starchy aroma as well as sweet aromatic, dairy and corn, formed two separate groups. Floral aroma was located opposite (negatively correlated) to hay-like/musty and sour silage in the biplot (Figure 6).

Among the aroma descriptors evaluated, IR64 was observed by the panelists to have more of a dairy, sweet aromatic and corn aroma than the other descriptors evaluated (Figure 6). Apo, on the other hand, was described as having more of a hay-like/musty and sour silage aroma with no floral scent. Of the progeny, three lines were observed by the panelists to have aroma similar to that of IR64 – Lines 164, 20 and 165, while lines 171, 22, 16 and 9 were observed to have similar aroma to that perceived in Apo. Lines 28, 63 and 123 also have similar aroma to that of IR64 and also have floral aroma. One sample was observed to have high levels of floral aroma, while Lines 122, 184 and 29 were observed to have more aroma of grassy, sewer animal and grainy/starchy notes than the other aroma descriptors evaluated.

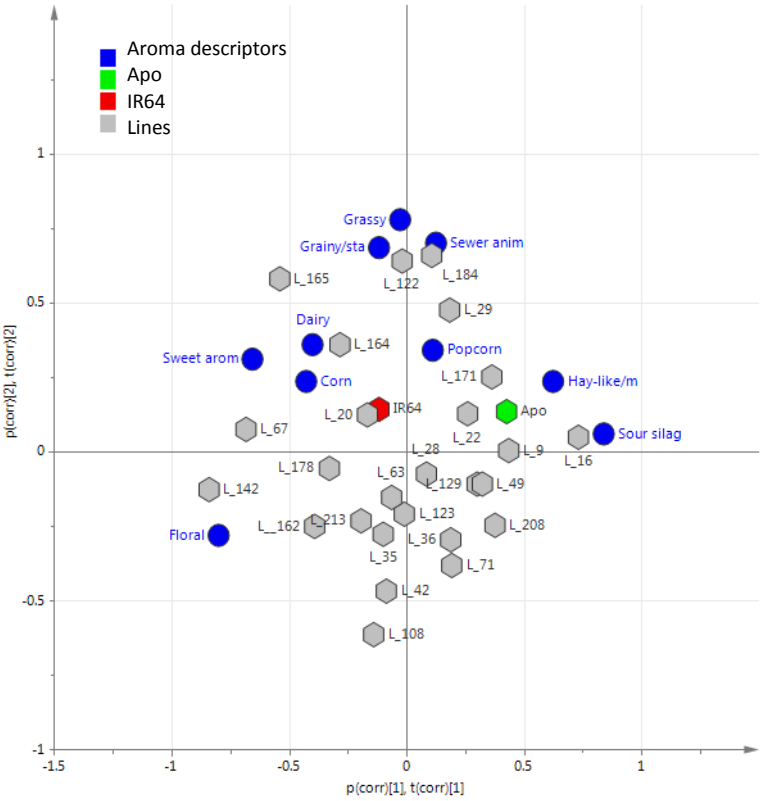


Figure 6. Biplot showing sensory attributes perceived in flour of Apo, IR64 and 26 selected lines, based on metabolomic similarity.

Discussion

The population derived from IR64 and Apo underwent genotyping by sequencing and final processing of the data revealed 2,681 polymorphic SNPs which were distributed across the 12 chromosomes (Figure 1). Chromosomes 8 and 10 were the most densely covered, along with the long arm of chromosome 1 and the short arm of chromosome 6 and 11 (Figure 1). Figure 2 shows that based on the genotyping data, the progeny separate into two main clusters, with IR64 in cluster A and Apo in cluster B. The low coefficient of parentage indicates that the parents are genetically quite distant and therefore represent a significant opportunity for recombination in a mapping population (Wang and Lu 2006). This conclusion is also supported by the results presented in Figure 3.

Agronomy: Yield under Drought

The parents of this Apo x IR64 mapping population differ in terms of grain quality and yield under drought. Drought has been shown not to significantly decrease yield in Apo (Venuprasad, Bool et al. 2012) and this was also observed in this investigation (Figure 3). In contrast, IR64 is susceptible to drought, and in Figure 3 we can observe that the yield of IR64 halved under drought conditions. In Figure 3 we also observe significant transgressive segregation, whereby many of the progeny had higher yield in drought than Apo, including the ten lines clustering with Apo in sub-cluster O and the eight with IR64 in sub-cluster C (Figure 2). Line 83, which is in the same sub-cluster C as IR64 (Figure 2), is positioned in the group of highest yield under drought, and gives the second highest yield under irrigation. Line 28, which had the highest yield under drought (Figure 3) is in the IR64 main cluster of the archaeopteryx tree but is positioned in sub-cluster D (Figure 2).

The large amount of transgressive segregation for yield (Figure 3) suggests that recombination has occurred at many loci that govern yield and stress tolerance. Recently, several varieties of rice have been resequenced, including IR64 and over 1000 genes for drought stress have been identified (Jain, Moharana et al. 2014), indicating the highly multi-genic nature of drought stress. Therefore, in our population there are many loci where recombination could have occurred and this could explain the high degree of transgressive segregation observed.

QTL analysis of drought stress identified one major QTL on the long arm of chromosome 3 (Figure 4A). QTLs for drought stress in rice have previously been identified on chromosome 12 (Bernier, Kumar et al. 2007), chromosomes 3 and 6 (Dixit, Singh et al. 2014) and chromosomes 1 and 2 (Sandhu, Singh et al. 2014). Thirteen QTLs have been identified on these chromosomes, and many of them have been detected in a number of different mapping populations (Kumar, Dixit et al. 2014). Two QTLs have been mapped onto chromosome 3 for yield under drought (Sandhu, Singh et al. 2014). The first, $qDTY_{3.1}$, was detected in a population derived from Apo and Swarna. $qDTY_{3.1}$ is located in the interval between 9.1–11.0 cM. The QTL peak was located at 10.0 cM and flanked by microsatellite markers RM520 (9.1 cM) and RM416 (10.0 cM). In the present study, the QTL is mapped in the interval of 1.3-13.3 cM of chromosome 3, flanked by SNP markers S3_346683 and S3_3337815, and the QTL peak is found at 7.3 cM, with a LOD score of 6.814 and F statistic score of 33.8. Even though the QTL peak found in this study does not fall within the interval reported by Venuprasad et al. (2012), it is likely that the dense marker coverage

from GBS enabled us to locate the QTL more accurately, and that it is in fact the same as $qDTY_{3.1}$. All the lines with $qDTY_{3.1}$ showed significantly higher yield under drought, indicating the importance of this QTL, at least in this population.

Quality: Aroma and Flavour

The aroma of rice is usually only discussed in the context of 2AP (Buttery, Ling et al. 1983). However, there are many other volatile compounds in rice, and many of these have flavour descriptors and are known to have low odour thresholds, meaning that the human nose can detect them at relatively low concentrations. Such compounds are therefore phenotypically relevant. Examples of these include alcohols, alkanes, alkenes, substituted alkanes and alkenes, saturated and unsaturated aldehydes (Buttery, Turnbaugh et al. 1988, Jezussek, Juliano et al. 2002, Laguerre, Mestres et al. 2007, Yang, Lee et al. 2010, Bryant and McClung 2011, Calingacion, Boualaphanh et al. 2012, Mathure, Jawali et al. 2014). Aroma is therefore a highly complex trait unlikely to be described by one or a small number of compounds. For example, the aroma of 2AP has been described as the roasted cracker smell of baking bread (Deblander, Van Aeken et al. 2014), but rice containing 2AP is usually described as having a 'floral' aroma (Buttery and Nam 1999, Champagne 2008, Mathure, Wakte et al. 2011). This indicates that aroma can be determined by a suite of compounds that may combine additively or synergistically and that individual compounds may contribute differently in different matrices or within different biochemical profiles.

It is well known that different varieties of rice have distinctly different aromas, and furthermore, that the environment can have an impact on the aroma of the polished grains obtained following different (regional) cultivation conditions and also as a result of seasonal fluctuations (Yoshihashi, Kabaki et al. 2002, Champagne, Bett-Garber et al. 2005). However, in a rice improvement program, breeders have only ever been able to select qualitatively and quantitatively for 2AP, because currently, tools were not available to facilitate selection for any other aromatic compound or aroma profile.

The current complex demographics of global population growth, the expansion of the Asian middle class, urbanisation and climate change (ADB 2012, ANZ 2012, Mohanty 2013), make it imperative that rice improvement programs release new varieties of rice which have both high yield potential in situations of drought, flood and climate volatility, as well containing the genes for the quality traits desired by consumers in each particular region (Calingacion, Laborte et al. 2014). This quest for so-called 'reliable yield' needs to be

closely linked to specific quality traits identified as being important for rice if we are to design sustainable varieties which will be accepted and which will continue to meet the needs of increasingly demanding consumers as well as help solve the anticipated problems relating to the predicted expansion of the rapidly-growing, rice-eating population (Calingacion, Laborte et al. 2014). These quality traits include aroma of the raw rice and of the rice when it is cooked. The variety Apo, in spite of its high yield in both favourable and unfavourable environments, has not been accepted by consumers in any country where it has been released (www.irri.org), and the reasons for this are likely to lie in both its aroma during cooking and its texture (Champagne, Bett-Garber et al. 2010).

Rice breeding programs are moving increasingly towards using tools of genetic selection centred on many different, complementary platforms (Chen, He et al. 2013, Li, Jiang et al. 2013). For example, breeding programs regularly combine QTLs or SNPs that associate with desired traits on one chip to assist in progeny selection (Dilla, Reveche et al. 2011, Hoffmann, Kvale et al. 2011, Fadista and Bendixen 2012, Johnston, Lindqvist et al. 2013, Li, Jiang et al. 2013, Mullen, McClure et al. 2013). Using particular platforms and those QTLs and SNPs which have been identified as being particularly important, whole breeding populations can be relatively rapidly screened for a set of commercially valuable genetic polymorphisms. Currently, a number of these customized selection platforms are being developed to assist in breeding programs for new varieties of rice (McCouch, Zhao et al. 2010, Tung, Zhao et al. 2010, Dilla, Reveche et al. 2011, Thomson, Zhao et al. 2012). In order to make use of customized selection chips which also include markers for rice quality, the most important traits defining this quality must first be properly described, after which a robust and relevant phenotyping tool must become available to measure the variability in these traits.

The aromatic quality of rice can be measured using new metabolomic profiling techniques such as GCMS that are e.g. able to detect volatile compounds of importance to aroma (Hall 2006), and this data is able to be used in QTL mapping and genetic associations to identify QTLs. In the current study, metabolomic profiling of the parents and progeny of the population derived from IR64 and Apo has shown that the parents are separated by a significant distance along PC1 based on the 105 volatile compounds detected (Figure 5). The progeny data generally distribute the lines between the parents. However, the ten lines in the same sub-cluster with Apo and the eight line close to IR64 were found not to cluster with the relevant parent in terms of the metabolomic profile of the grains (Figure 5). This

might suggest that some of the metabolites are not direct genetic products, but for example, may be present due to oxidative chemistry occurring post-harvest during grain processing. Indeed, many of the compounds detected are alcohols, alkanes and aldehydes, which are known products of fatty acid oxidation (Lam and Proctor 2003). Furthermore, many of these compounds have aroma descriptors (Table 1) and low odour thresholds suggesting that they will likely play a true role in the aroma phenotype.

The compounds that were found to discriminate IR64 and the genetically similar lines have been putatively identified on the basis of fragmentation data and retention index to be DL-2,3-butanediol and butanol (Table 1). DL-2,3-butanediol is an alcohol with a pleasant buttery and creamy aroma (Buttery, Orts et al. 1999). It has also been detected in black rice (Ajarayasiri and Chaiseri 2008). IR64 was also perceived by the panelists to have high levels of dairy, corn and sweet aromatic aromas. Butanol, on the other hand, is described as having a malty aroma and has also been found in other rice varieties (Buttery, Turnbaugh et al. 1988). Unfortunately, in this study we were unable to identify strong QTL for DL-2,3-butanediol and butanol. Moreover, because of high odour threshold of butanol (Czerny, Christlbauer et al. 2008) and DL-2,3-butanediol (Buttery, Orts et al. 1999), these compounds are unlikely to contribute to the aroma of IR64 as detected by humans.

Several compounds were detected that discriminated Apo from both IR64 and the lines that were clustering with it. These compounds were also associated with aroma descriptors and have low odour threshold. Pentanol and heptanone have been linked to a sweet and fruity aroma. We have identified a mQTL for heptanone on chromosome 3. Moreover, hexanol and hex-3-enal are associated with a grassy and leafy-like smell, which may explain the panelists perceiving high levels of hay-like and sour silage aromas.

Interestingly, there were lines with similar metabolite profiles as IR64, which were also perceived to have aroma descriptors similar to those perceived in IR64 by the panelists. Lines 164 and 165 were located in the PCA near IR64 (Figure 5B) and were also described by the panelists to have corn, sweet aromatic and dairy aroma similar to that of IR64. On the other hand, Line 20 that was located in the middle of the PCA was also perceived to have similar aroma descriptors as those that were perceived in IR64 by the panelists. Similarly, Lines 171, 22 and 16 that were perceived by the panelists to have similar aroma descriptors to Apo i.e. high in hay-like/musty and sour silage, were also located in the middle of the PCA (Figure 5B). It should be noted that the PCA was based on all the volatile

compounds detected in the headspace of the rice sample that may or may not contribute directly to the aroma being perceived by the panelists in the sensory evaluation. The aroma that was perceived by the panelists is determined only by those volatile compounds with low odour threshold.

QTLs were found for few metabolites which may suggest that many of the metabolites of rice aroma detected are products arising after harvest. Many of the compounds observed were alcohols, alkanes and aldehydes, suggesting that they may have resulted from oxidation of fatty acids. The unsaturated aldehydes generally have pleasant aroma descriptors, whereas the saturated aldehydes are generally described by less pleasant odours. Most of the unsaturated aldehydes were more associated with IR64 whereas the saturated aldehydes were associated more with Apo. Considering their odour thresholds, some of these compounds are likely to explain the phenotypic differences between the two varieties.

It is of equal importance in breeding programs to achieve increased levels of yield potential and premium grain quality even, or especially, under a stress environment. In this study we identified lines that we have in the population of which have similar metabolomic and sensory properties to IR64 and also had fairly similar yield values as Apo. The yield of Lines 20, 164 and 28 under both irrigated and drought conditions were similar to that of the Apo under the same conditions. Interestingly, Line 28, which had the highest yield under drought and has the QTL associated with yield under drought on chromosome 3, is also located in the main cluster A of the archaeopteryx tree where IR64 is also located.

Conclusions

This study offers valuable information for developing new varieties with specific aroma traits as desired by consumers, through marker-assisted breeding approaches and consumer-validated phenotyping. Using a population derived from Apo and IR64, we were able to identify lines that had similar metabolomic and sensory properties to IR64 and had comparable yield values to Apo. Lines 20, 164 and 28 were perceived by the panelists to have high levels of corn, dairy and sweet aromatic features, and their yields under both irrigated and drought conditions were similar to that of the Apo parent under the same conditions. These lines were also located in the main cluster A of the archaeopteryx tree where IR64 is also located. Interestingly, Line 28 had the highest yield under drought and

carries the QTL associated with yield under drought on chromosome 3. These three lines warrant further testing in multi-location trials for potential variety release.

Six novel mQTLs for volatile compounds in rice were identified. Using a highly dense genetic map, 4 major QTLs for the metabolites which were annotated as pentanol, hexanol, hexanal, and heptanone were mapped to the same region in chromosome 3. Moreover, 1 QTL was detected in chromosome 1 for 3,7-dimethyl-octen-1-ol and 1 QTL for hexanol in chromosome 2. The importance of these QTLs in influencing metabolite variation can be validated in the future using other rice varieties and populations.

Acknowledgements

The authors thank F Salisi, R Ilagan, L Samadio and A Madrid for their assistance in conducting the field experiment, C Raghavan for her guidance in GBS data filtering, J Concepcion for her assistance in QTL mapping, W Eusebio for his assistance in IRIS database, A Riedel for his assistance in the statistical analysis and V Daygon for her assistance in the sensory studies and H Li, E Ang, T Stanley, F Widjaja, M Vegafria, and A Kanon as panelists in the sensory studies. MC also thanks the Monsanto Beachell-Borlaug International Scholarship Program, International Rice Research Institute and University of Queensland for additional funding.

Supporting Information

Table S1. Flavour descriptions and reference used in the sensory evaluation of Apo, IR64 and 26 selected lines.

Flavour	Standard used	Description
Sewer/animal	Hard-boiled egg	An immediate and distinct pungent aromatic in the flavour characterized as sulfur-like and generic animal. Animal aromatic in the flavour can sometimes be identified as “piggy”.
Grain/starchy	Flour mixture	A general term used to describe the aromatics in the flavour associated with grains such as corn, oats and wheat. It is an overall grainy impression characterized as sweet, brown, sometimes dusty, and sometimes generic nutty or starchy.
Floral	Potpourri	Aromatics associated with dried flowers, such as lilac or lavender. This aromatic is characterized as spicy floral as in an “old fashioned sachet”.
Hay-like/ musty	Hay	A dry, dusty, slightly brown aroma/flavour with a possible trace of musty.
Corn	Canned creamed corn	The sweet aromatics of the combination of corn kernels, corn milk, and corn germ.
Grassy	Green beans	A dried, green, slightly earthy, slightly sweet aroma/flavour including grassy and fresh green bean aroma/flavour.
Sour/silage	Alfalfa	A sour fermented vegetation aroma/flavour, not decaying vegetation.
Sweet aromatic	Fairy floss	A sweet impression such as cotton candy, caramel, or sweet fruity that may appear in the aroma and or aromatics.
Dairy	Milk	A general term associated with aromatics of pasteurized cow’s milk. Most apparent just before swallowing.
Popcorn	Popcorn	A dry, dusty, slightly toasted and slightly sweet aromatic in the flavour that can be specifically identified as popcorn.

Table S2. Intensity scales used in the sensory evaluation of rice.

Descriptor	Reference	Intensity
Sweet	1% sucrose solution	1
Oil	Thin's Potato Chips	2
Orange complex	Just Juice's Orange Drink	3
Sweet	Nabisco's Ritz Cracker	4
Sugar	Arnott's Scotch Finger	5
		6
Grape	Berri's Grape Juice	7
		8
		9
Apple	Three 333 Three's Apple Sauce	10

References

ADB (2012). Food security and poverty in Asia Pacific. Key Challenges and Policy Issues. Philippines, Asian Development Bank.

Ajarayasiri, J. and S. Chaiseri (2008). Comparative study on aroma-active compounds in Thai, black and white glutinous rice varieties. Kasetsart Journal: Natural Sciences **42**(4): 715-722.

ANZ (2012). Greener pastures: The global soft commodity opportunity for Australia and New Zealand. ANZ Insights. **3:** 1-80.

Bernier, J., A. Kumar, R. Venuprasad, D. Spaner and G. Atlin (2007). A large-effect QTL for grain yield under reproductive-stage drought stress in upland rice. Crop Science **47**(2): 507-516.

Bradbury, L., T. Fitzgerald, R. Henry, Q. Jin and D. Waters (2005). The gene for fragrance in rice. Plant Biotechnology Journal **3**(3): 363-370.

Bradbury, P., Z. Zhang, D. Kroon, T. Casstevens, Y. Ramdoss and E. Buckler (2007). TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics **23**(19): 2633-2635.

Bryant, R. and A. McClung (2011). Volatile profiles of aromatic and non-aromatic rice cultivars using SPME/GC–MS. Food Chemistry **124**(2): 501-513.

Buttery, R., L. Ling, B. Juliano and J. Turnbaugh (1983). Cooked rice aroma and 2-acetyl-1-pyrroline. Journal of Agricultural and Food Chemistry **31**(4): 823-826.

Buttery, R. and Y. Nam (1999). Flavor volatiles of rice and rice products: Some recent studies. Abstracts of Papers of the American Chemical Society **218**: U30-U30.

Buttery, R., J. Turnbaugh and L. Ling (1988). Contribution of volatiles to rice aroma. Journal of Agricultural and Food Chemistry **36**(5): 1006-1009.

Calingacion, M., C. Boualaphanh, V. Daygon, R. Anacleto, R. Sackville Hamilton, B. Biais, C. Deborde, M. Maucourt, A. Moing, R. Mumm, R. de Vos, A. Erban, J. Kopka, T. Hansen, K. Laursen, J. Schjoerring, R. Hall and M. Fitzgerald (2012). A genomics and multi-platform metabolomics approach to identify new traits of rice quality in traditional and improved varieties. Metabolomics **8**(5): 771-783.

Calingacion, M., L. Fang, L. Quiatchon-Baeza, R. Mumm, A. Riedel, R. Hall and M. Fitzgerald (2014). Delving deeper into technological innovations to understand differences in rice quality. Rice: In press.

Calingacion, M., A. Laborte, A. Nelson, A. Resurreccion, J. Concepcion, V. Daygon, R. Mumm, R. Reinke, S. Dipti, P. Bassinello, J. Manful, S. Sophany, K. Lara, J. Bao, L. Xie, K. Loaiza, A. El-hissewy, J. Gayin, N. Sharma, S. Rajeswari, S. Manonmani, N. Rani, S. Kota, S. Indrasari, F. Habibi, M. Hosseini, F. Tavasoli, K. Suzuki, T. Umemoto, C. Boualaphanh, H. Lee, Y. Hung, A. Ramli, P. Aung, R. Ahmad, J. Wattoo, E. Bandonill, M. Romero, C. Brites, R. Hafeel, H. Lur, K. Cheaupun, S. Jongdee, P. Blanco, R. Bryant, N. Thi Lang, R. D. Hall and M. Fitzgerald (2014). Diversity of global rice markets and the science required for consumer-targeted rice breeding. PLoS One **9**(1): e85106.

Champagne, E. (2008). Rice aroma and flavor: A literature review. Cereal Chemistry **85**(4): 445-454.

Champagne, E., K. Bett-Garber, M. Fitzgerald, C. Grimm, J. Lea, K. Ohtsubo, S. Jongdee, L. Xie, P. Bassinello, A. Resurreccion, R. Ahmad, F. Habibi and R. Reinke (2010). Important sensory properties differentiating premium rice varieties. Rice **3**: 270–281.

Champagne, E., K. Bett-Garber, J. Thompson, R. Mutters, C. Grimm and A. McClung (2005). Effects of drain and harvest dates on rice sensory and physicochemical properties. Cereal Chemistry **82**(4): 369–374.

Chen, H., H. He, F. Zhou, H. Yu and X. Deng (2013). Development of genomics-based genotyping platforms and their applications in rice breeding. Current Opinion in Plant Biology **16**(2): 247-254.

Chen, S., Y. Yang, W. Shi, Q. Ji, F. He, Z. Zhang, Z. Cheng, X. Liu and M. Xu (2008). Badh2, encoding betaine aldehyde dehydrogenase, inhibits the biosynthesis of 2-acetyl-1-pyrroline, a major component in rice fragrance. The Plant Cell **20**(7): 1850-1861.

Cho, S., E. Nuijten, R. Shewfelt and S. Kays (2014). Aroma chemistry of African *Oryza glaberrima* and *Oryza sativa* rice and their interspecific hybrids. Journal of the Science of Food and Agriculture **94**(4): 727-735.

Churchill, G. and R. Doerge (1994). Empirical threshold values for quantitative trait mapping. Genetics **138**(3): 963-971.

Collard, B. and D. Mackill (2008). Marker-assisted selection: An approach for precision plant breeding in the twenty-first century. Philosophical Transactions of the Royal Society B: Biological Sciences **363**(1491): 557-572.

Czerny, M., M. Christlbauer, M. Christlbauer, A. Fischer, M. Granvogl, M. Hammer, C. Hartl, N. Hernandez and P. Schieberle (2008). Re-investigation on odour thresholds of key food aroma compounds and development of an aroma language based on odour qualities of defined aqueous odorant solutions. European Food Research and Technology **228**(2): 265-273.

Deblander, J., S. Van Aeken, A. Adams, N. De Kimpe and K. Abbaspour Tehrani (2014). New short and general synthesis of three key Maillard flavour compounds: 2-acetyl-1-pyrroline, 6-acetyl-1,2,3,4-tetrahydropyridine and 5-acetyl-2,3-dihydro-4H-1,4-thiazine. Food Chemistry **168**: 327-331.

Del Mundo, A. and B. Juliano (1981). Consumer preference and properties of raw and cooked milled rice. Journal of Texture Studies **1**: 133-140.

Demyttenaere, J., S. Macura, N. De Kimpe and R. Verhe (2003). Production of pyrazines and 2-acetyl-l-pyrroline by *Bacillus cereus* strains. Flavour Research at the Dawn of the Twenty-First Century: 344-349.

Dilla, C., M. Reveche, S. Carandang, J. Rey, C. Mojica, M. De Ocampo, D. Lozada, P. Capistrano, E. Tandayu and R. Mauleon (2011). SNP marker technology in the molecular breeding pipeline. Philippine Journal of Crop Science

Dixit, S., A. Singh, M. Sta Cruz, P. Maturan, M. Amante and A. Kumar (2014). Multiple major QTL lead to stable yield performance of rice cultivars across varying drought intensities. BMC Genetics **15**(1): 1-13.

Dunemann, F., D. Ulrich, A. Boudichevskaia, C. Grafe and W. Weber (2009). QTL mapping of aroma compounds analysed by headspace solid-phase microextraction gas chromatography in the apple progeny 'Discovery' × 'Prima'. Molecular Breeding **23**(3): 501-521.

Eriksson, L., E. Johansson, N. Kettaneh-Wold, C. Wikstrom, J. Trygg and S. Wold (2006). Multi- and megavariate data analysis. Basic principles and applications. Umea, Sweden, Umetrics Academy.

Fadista, J. and C. Bendixen (2012). Genomic position mapping discrepancies of commercial SNP chips. PloS One **7**(2): e31025.

Fitzgerald, M., S. McCouch and R. Hall (2009). Not just a grain of rice: the quest for quality. Trends in Plant Science **14**(3): 133-139.

Hall, R. (2006). Plant metabolomics: from holistic hope, to hype, to hot topic. New Phytologist **169**(3): 453-468.

Hall, R., I. Brouwer and M. Fitzgerald (2008). Plant metabolomics and its potential application for human nutrition. Physiologia Plantarum **132**(2): 162-175.

Hoffmann, T., M. Kvale, S. Hesselson, Y. Zhan, C. Aquino, Y. Cao, S. Cawley, E. Chung, S. Connell and J. Eshragh (2011). Next generation genome-wide association tool: Design and coverage of a high-throughput European-optimized SNP array. Genomics **98**(2): 79-89.

Inui, T., F. Tsuchiya, M. Ishimaru, K. Oka and H. Komura (2013). Different beers with different hops. Relevant compounds for their aroma characteristics. Journal of Agricultural and Food Chemistry **61**(20): 4758-4764.

Jain, M., K. Moharana, R. Shankar, R. Kumari and R. Garg (2014). Genomewide discovery of DNA polymorphisms in rice cultivars with contrasting drought and salinity stress response and their functional relevance. Plant Biotechnology Journal **12**(2): 253-264.

Jezussek, M., B. Juliano and P. Schieberle (2002). Comparison of key aroma compounds in cooked brown rice varieties based on aroma extract dilution analyses. Journal of Agricultural and Food Chemistry **50**(5): 1101-1105.

Joehanes, R. and J. Nelson (2008). QGene 4.0, an extensible Java QTL-analysis platform. Bioinformatics **24**(23): 2788-2789.

Johnston, S., M. Lindqvist, E. Niemelä, P. Orell, J. Erkinaro, M. Kent, S. Lien, J. Vähä, A. Vasemägi and C. Primmer (2013). Fish scales and SNP chips: SNP genotyping and allele frequency estimation in individual and pooled DNA from historical samples of Atlantic salmon (*Salmo salar*). BMC Genomics **14**(1): 439.

Kovach, M., M. Calingacion, M. Fitzgerald and S. McCouch (2009). The origin and evolution of fragrance in rice (*Oryza sativa* L.). Proceedings of the National Academy of Sciences **106**(34): 14444-14449.

Kumar, A., S. Dixit, T. Ram, R. Yadaw, K. Mishra and N. Mandal (2014). Breeding high-yielding drought-tolerant rice: genetic variations and conventional and molecular approaches. Journal of Experimental Botany: 1-14.

Laguerre, M., C. Mestres, F. Davrieux, J. Ringuet and R. Boulanger (2007). Rapid discrimination of scented rice by solid-phase microextraction, mass spectrometry, and

multivariate analysis used as a mass sensor. Journal of Agricultural and Food Chemistry **55**(4): 1077-1083.

Lam, H. and A. Proctor (2003). Milled rice oxidation volatiles and odor development. Journal of Food Science **68**(9): 2676-2681.

Li, H., L. Jiang, J. Youn, W. Sun, Z. Cheng, T. Jin, X. Ma, X. Guo, J. Wang, X. Zhang, F. Wu, C. Wu, S. Kim and J. Wan (2013). A comprehensive genetic study reveals a crucial role of CYP90D2/D2 in regulating plant architecture in rice (*Oryza sativa*). New Phytologist.

Lommen, A. (2009). MetAlign: Interface-driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing. Analytical Chemistry **81**(8): 3079-3086.

Mackill, D., A. Ismail, U. Singh, R. Labios, T. Paris and L. Donald (2012). Chapter 6 - Development and rapid adoption of submergence-tolerant (Sub1) rice varieties. Advances in Agronomy, Academic Press. **115:** 299-352.

Mathieu, S., V. Cin, Z. Fei, H. Li, P. Bliss, M. Taylor, H. Klee and D. Tieman (2009). Flavour compounds in tomato fruits: identification of loci and potential pathways affecting volatile composition. Journal of Experimental Botany **60**(1): 325-337.

Mathure, S., N. Jawali, R. Thengane and A. Nadaf (2014). Comparative quantitative analysis of headspace volatiles and their association with BADH2 marker in non-basmati scented, basmati and non-scented rice (*Oryza sativa* L.) cultivars of India. Food Chemistry **142**: 383-391.

Mathure, S., K. Wakte, N. Jawali and A. Nadaf (2011). Quantification of 2-acetyl-1-pyrroline and other rice aroma volatiles among Indian scented rice cultivars by HS-SPME/GC-FID. Food Analytical Methods **4**(3): 326-333.

McCouch, S. (2008). Gene nomenclature system for rice. Rice **1**(1): 72-84.

McCouch, S., K. Zhao, M. Wright, C. Tung, K. Ebana, M. Thomson, A. Reynolds, D. Wang, G. DeClerck, M. Ali, A. McClung, G. Eizenga and C. Bustamante (2010). Development of genome-wide SNP assays for rice. Breeding Science **60**(5): 524-535.

Meilgaard, M., G. Civille and B. Carr (2007). Sensory evaluation techniques. Boca Raton, FL, USA, CRC Press.

Mohanty, S. (2013). Trends in global rice production. Rice Today **13**: 44-45.

Mullen, M., M. McClure, J. Kearney, S. Waters, R. Weld, P. Flynn, C. Creevey, A. Cromie and D. Berry (2013). Development of a custom SNP chip for dairy and beef cattle breeding, parentage and research. Interbull Bulletin(47):1-8.

Murray, M. and W. Thompson (1980). Rapid isolation of high molecular weight plant DNA. Nucleic Acids Research **8**(19): 4321-4326.

Sandhu, N., A. Singh, S. Dixit, M. Cruz, P. Maturan, R. Jain and A. Kumar (2014). Identification and mapping of stable QTL with main and epistasis effect on rice grain yield under upland drought stress. BMC Genetics **15**(1): 63.

Strehmel, N., J. Hummel, A. Erban, K. Strassburg and J. Kopka (2008). Retention index thresholds for compound matching in GC-MS metabolite profiling. Journal of Chromatography B **871**(2): 182-190.

Thomson, M., K. Zhao, M. Wright, K. McNally, J. Rey, C. Tung, A. Reynolds, B. Scheffler, G. Eizenga, A. McClung, H. Kim, A. Ismail, M. de Ocampo, C. Mojica, M. Reveche, C. Dilla-Ermita, R. Mauleon, H. Leung, C. Bustamante and S. McCouch (2012). High-throughput single nucleotide polymorphism genotyping for breeding applications in rice using the BeadXpress platform. Molecular Breeding **29**(4): 875-886.

Tikunov, Y., S. Laptenok, R. Hall, A. Bovy and R. de Vos (2012). MSClust: a tool for unsupervised mass spectra extraction of chromatography-mass spectrometry ion-wise aligned data. Metabolomics **8**(4): 714-718.

Tung, C., K. Zhao, M. Wright, M. Ali, J. Jung, J. Kimball, W. Tyagi, M. Thomson, K. McNally, H. Leung, H. Kim, S. Ahn, A. Reynolds, B. Scheffler, G. Eizenga, A. McClung, C. Bustamante and S. McCouch (2010). Development of a research platform for dissecting phenotype–genotype associations in rice (*Oryza spp.*). Rice **3**(4): 205-217.

Venuprasad, R., M. Bool, L. Quiatchon and G. Atlin (2012). A QTL for rice grain yield in aerobic environments with large effects in three genetic backgrounds. Theoretical and Applied Genetics **124**(2): 323-332.

Venuprasad, R., H. Lafitte and G. Atlin (2007). Response to direct selection for grain yield under drought stress in rice. Crop Science **47**(1): 285-293.

Verhoeven, H., H. Jonker, R. De Vos and R. Hall (2011). Solid phase micro-extraction GC–MS analysis of natural volatile components in melon and rice. Plant Metabolomics. N. Hardy and R. Hall, Humana Press. **860:** 85-99.

Wang, S. and Z. Lu (2006). Genetic diversity among parental lines of indica hybrid rice (*Oryza sativa* L.) in China based on coefficient of parentage. Plant Breeding **125**(6): 606-612.

Yang, D., K. Lee and S. Kays (2010). Characterization and discrimination of premium-quality, waxy, and black-pigmented rice based on odor-active compounds. Journal of the Science of Food and Agriculture **90**(15): 2595-2601.

Yoshihashi, T., N. Kabaki, T. Nguyen and H. Inatomi (2002). Formation of flavor compound in aromatic rice and its fluctuations with drought stress. Research Highlights - JIRCAS: 32-33.

Chapter 6

General discussion

Global food security is currently one of the most critical issues demanding attention in relation to projected population growth in the next decades (Fitzgerald, Waters et al. 2010, ADB 2012, ANZ 2012). It is predicted that the global population will reach 9 billion before the year 2050 (FAOSTAT 2013) entailing ca 2 billion extra mouths will need to be fed. Most of these people will have rice as their staple food. Constraints to food production are climate change, urbanisation, and shrinkage of agricultural land; these compounding issues will ensure that global food security will remain on political and strategic agendas for the foreseeable future (Zeigler and Mohanty 2010).

The key issues of increasing overall yield potential as well as securing 'reliable yield' (i.e. ensuring reasonable yields even during sub-optimal seasons) are gaining top priority and especially so for the staple food crops such as wheat, corn and rice. Areas under the most extreme threat of climate change include large parts of South and Central Asia; pockets of Southeast Asia; East, West and Sub-Saharan Africa; and Latin America (IPCC 2014). In much of these regions, the staple food is rice. It is therefore essential that rice improvement programs in the regions at the highest level of risk include targets of developing climate-ready varieties that can produce sufficient and sustainable yield in increasingly less favourable conditions. In addition, new climate-tolerant varieties must meet the needs of both farmers and consumers alike - in terms of agronomic traits as well as traits meeting consumer requirements relating for example, to nutritional value and other quality attributes.

Agronomy: Yield under Drought

The domestication of rice has produced variation among genotypes that has enabled cultivation in a wide range of environments, including salinity-affected coastal areas, water-scarce upland areas, and lowland monsoonal regions that are prone to inundation and flood (Wassmann, Jagadish et al. 2009). These adaptations demonstrate that the species as a whole houses genes that enable a high degree of crop plasticity, and which therefore permit productive growth under a range of environmental conditions (Wassmann, Jagadish et al. 2009). Identifying the genetic basis of tolerance to the different abiotic stresses, in order to introgress causal genes into elite varieties, has been and remains a central research priority. Recent examples of enhancing our capacity to fulfil the need for stress-tolerant rices include the identification of genes for submergence tolerance for rice grown in monsoonal regions (Xu, Xu et al. 2006), genes for grain yield under phosphorus-deficient soil (Gamuyao, Chin et al. 2012).

Efforts to develop high-yielding drought-tolerant rice varieties has revealed fourteen large-effect quantitative trait loci (QTL) for grain yield under drought conditions in chromosomes 1 and 2 (Sandhu, Singh et al. 2014), and 3, 6 and 12 (Bernier, Kumar et al. 2007). Six of these: qDTY1.1, qDTY2.2, qDTY3.1, qDTY3.2, qDTY6.1, and qDTY12.1, are associated with higher yields under drought in both lowland and upland rice areas (Kumar, Dixit et al. 2014), but the different combinations of QTLs in different varieties suggests that QTLs for drought tolerance work differently in different genetic backgrounds. In the study reported in this thesis, a population derived from two well-known varieties, IR64 and Apo, important respectively for quality and yield under drought, underwent genotyping by sequencing (GBS). A densely covered single nucleotide polymorphism (SNP) map was generated and used to identify QTLs associated with yield under irrigation and drought. Drought did not significantly decrease yield in Apo while yield of IR64 was reduced by half (Chapter 5). A large amount of transgressive segregation for yield was also observed indicating that recombination occurred at many loci that govern yield and stress tolerance. Consistent with this, a recent study has identified that over 1000 genes are up or down-regulated in rice during drought (Huang, Zhao et al. 2012) suggesting that drought tolerance is a feature of several metabolic pathways. Interestingly, the major QTL for drought tolerance identified in the present study, on chromosome 3, is likely to be qDTY3.1 as it was previously detected in a population derived from Apo as the drought tolerant parent (Dixit, Singh et al. 2014). All the lines carrying this QTL showed significantly higher yield under drought than those without it, indicating the potential importance of this QTL (Chapter 5) in drought tolerance. However, there is a large amount of transgressive segregation seen for yield under drought, from 4 – 6 tonnes/ha, indicating that there are many other genes perhaps that are not individually strong QTLs that also contribute to yield under drought.

New rice varieties with tolerance to abiotic stresses offer a clear yield advantage to farmers, particularly in those regions of climate instability. However, the release of such varieties does not automatically lead to adoption. For example, 17 drought tolerant varieties have been released by the International Rice Research Institute (Kumar, Dixit et al. 2014), but adoption of these by farmers, even in stress prone areas, has been notably low (Mackill, Ismail et al. 2012). The difficulty rice improvement programs face in releasing varieties with stress adaptation is that stresses, such as submergence and drought, occur widely over Asia, in regions with a wide range in consumer preferences (Chapter 2). Interestingly, many of the most popular widely-grown and widely-accepted, “mega” varieties, are not tolerant to abiotic stresses. As a way around selecting for the multiple traits of megavarieties, the

submergence tolerance gene has been introgressed into many of the mega varieties of Asia, using new generation genotyping for rapid background and foreground selection. This ensures that each mega variety carries the submergence tolerance gene, in its own genetic background. Recently, submergence-tolerant versions of IR64, Sambha mashuri, Swarna, BR11 and TDK were released in Southeast Asia, and acceptability among farmers is favourable (Manzanilla, Paris et al. 2011). However by focussing on megavarieties, the gene for submergence tolerance does not reach farmers in areas where mega varieties are not grown. These areas tend to be characterised by poverty and subsistence farming of traditional varieties. Previously, Apo was released in the upland areas of the Lao PDR, where drought is common, and where a range of low-yielding, traditional varieties are grown. Despite its high yield, Apo was rejected. While this illustrates a difficult problem of mobilising genes of stress tolerance to the poorest farmers, it reveals that understanding the factors governing consumer acceptance, and enabling rapid selection of those traits, would assist enormously in the mobilisation of stress tolerance genes to areas that do not grow megavarieties. The ultimate driver of widespread and long term uptake and acceptance of a new variety is not only farmer adoption but also consumer acceptance (Fitzgerald, McCouch et al. 2009, Champagne, Bett-Garber et al. 2010, Boualaphanh, Calingacion et al. 2011). If consumers do not like a new variety, for whatever reason, and are in a position to choose, then that variety will have limited market potential and the financial feedback to farmers will mean that they quickly return to growing those varieties which consumers provide a market for.

Rice quality: Aroma and Flavour

The widespread adoption of “upgraded” mega varieties, suggests that it is currently easier to fine-tune existing varieties than to create new ones. Part of the reason for this likely lies in our current inability to identify those (quality) traits that characterise consumer preference. Without this knowledge, target breeding strategies to develop and release stress tolerant varieties are likely to fail. In order to understand the diversity of consumer preferences more fully, a survey was conducted among members of the International Network for Quality Rice (INQR), who are local experts involved in grain quality evaluation programs from 25 countries. These experts are best positioned to provide consumer and market information on local and regional preferences for particular varieties of rice as well as information on their quality attributes. The main grain quality traits of the major varieties preferred by consumers in these 25 countries, including for the larger countries, were identified (Chapter 2). Interestingly, an analysis of the data showed that across this region, there are 18

combinations of the quality traits that are currently screened for, such as grain length and shape, amylose content, gelatinisation temperature, gel consistency, and aroma which could be identified. This complexity immediately revealed the extent of the specificity of consumer preference and also that these quality traits were not sufficiently discriminating between varieties. The two most popular combinations of traits are long and slender grains either with 1) low amylose, low gelatinisation temperature and which are aromatic or alternatively 2) intermediate amylose and gelatinisation temperature and which are not aromatic. However, here it has also been demonstrated that this is an oversimplification because further assessment of these two types reveals that those varieties included within each category can be readily identified by consumers as being different. For example, BRS Primavera and IR64 are popular varieties in Brazil and the Philippines, respectively. These varieties have the same combination of quality traits: long grain, intermediate amylose, medium gel consistency and intermediate gelatinisation temperature. Nevertheless, sensory panellists still readily found their taste, flavour and texture to be significantly different (Champagne, Bett-Garber et al. 2010). This reinforces the complexity of quality attributes as well as our need for improved screening techniques to assess rice quality in depth.

Consumer preference is initially driven by the physical appearance of the dry grains as experienced at the moment of purchase. Physical properties, including grain length and shape, grain translucence and the lack of chalk, are quality components immediately visible to consumers and so consequently, these are also major factors determining consumer choice and hence, help define market value (Fitzgerald, McCouch et al. 2009). Consumer preference is then further cemented by aspects related to the cooking and eating qualities. Sensory and mouth-feel qualities typically include the aroma, flavour, taste and the texture of rice grains in the mouth. Cooking qualities, on the other hand, are largely influenced by the properties of the specific starch composition. The importance of such properties has been recognised in the past and analytical methods have been developed and are commonly used to measure these properties in terms of gelatinisation temperature, amylose content and gel consistency. In many countries rice is cooked two or three times a day and the aroma emitted when rice is being cooked is also a quality attribute of great importance and is again, strongly linked to local cultural preferences.

The sensory characteristics of rice are commonly measured by descriptive sensory analyses. This is often a subjective approach used to characterise and analytically measure traits of aroma, flavour and texture of foods by a panel specifically trained for this purpose

(Meilgaard, Civille et al. 2007). For example, Champagne et al. (2008) developed an approach based upon a lexicon with 10-12 descriptors to define differences between a series of diverse rice varieties. The panellists are able to measure and quantify the intensities of the flavour and aroma attributes using a universal scale for all foods (Meilgaard, Civille et al. 2007). This approach has been used extensively to determine the effect of different growing, processing and storage conditions on the sensory properties of rice grains (Meullenet, Marks et al. 2000, Moi, Srzednicki et al. 2003, Champagne, Bett-Garber et al. 2009, Shobana, Malleshi et al. 2011). However, earlier, Champagne et al, (2005) employing an intensity scale of 0-15 with flavour components of US name brand products, found that the rating of aroma and flavour for rice using this approach was generally only 1-3 with a maximum rating of 5. This poses a significant challenge to sensory panellists when working with rice, where differences in aroma and flavour are relatively small and difficult to define. Furthermore, sensory analysis using trained experts is challenging, time-consuming and expensive to conduct. The training and retraining of the panellists take an appreciable amount of time and differences between panellists can be problematic (Yang, Shewfelt et al. 2008). For these reasons, it is always more attractive to identify an objective, instrument – based approach which gives a more robust analysis. The difficulty with this though, is ensuring that the data produced, and the interpretation, are relevant to the sensory properties as experienced by consumers.

Several instrumental methods have been developed to evaluate individual compounds that affect the aroma of food. These methods are usually based on separation, identification, and quantification of metabolites found to be present in the headspace around a food product or within the actual food matrix (Chambers and Koppel 2013). However, the sensory characteristics of rice and aroma perception, like for other foods, are not simple stimulus-response processes but rather, are much more complex such that each aroma is usually composed of a few, to many, key volatile compounds (Buettner and Schieberle 2000). Recent advances in the detection and identification of volatile compounds through the use of non-targeted metabolomics approaches (Hall, Brouwer et al. 2008, Fitzgerald, McCouch et al. 2009) have led to the suggestion that metabolomic profiling could enable us to define better the compounds that are present in different varieties of rice, and to help us identify those that make an important contribution to aroma and flavour. Adopting this approach as reported here, four complementary metabolomic platforms were used to study a set of commercially-relevant rice materials. Unique metabolomic signatures were identified for each of three waxy rice (so – called ‘sticky rice’) varieties from the Lao

People's Democratic Republic that had been grown in the Lao PDR under different nitrogen (N) fertiliser treatments (Chapter 3). Many of the compounds identified, which discriminated the three varieties, were found to be associated with aroma and nutritional value. Indeed, each metabolomics platform tested, which each targeted a different chemistry of metabolites, was individually capable of showing different varietal profiles. This is in contrast to the more routine quality evaluation tools used for traits such as gelatinisation temperature, viscosity measured in terms of breakdown, setback and retrogradation, hardness and stickiness of the cooked grains, which were unable to distinguish the three Lao varieties (Chapter 3). Interestingly, although changes in individual metabolites were observed in the metabolic profiles of rice grains obtained from the different nitrogen regimes, the overall profiles were relatively similar for each variety (Figure 3, Chapter 3). This suggests that aroma and flavour of each genotype may not be greatly affected by the growing conditions.

As indicated above, the aroma of rice is an important feature of rice which determines to some extent its quality and it can also play a role in consumer preference and be related to contrasting consumer preferences (Chapter 2). Much previous research conducted towards understanding rice aroma was focused solely on 2-acetyl-1-pyrroline (2AP) (Champagne, Bett-Garber et al. 2009, Kovach, Calingacion et al. 2009, Maraval, Sen et al. 2010, Vanavichit and Yoshihashi 2010). 2AP is the characteristic aroma compound present in popular fragrant varieties including both the jasmine and basmati types (Fitzgerald, Sackville Hamilton et al. 2008). However there are many other volatile compounds in rice that have known aroma descriptors and which can be predicted to contribute also to rice fragrance. It is therefore likely to be these compound mixtures which determine the true albeit subtle and detectable differences between rice varieties. Among these groups of compounds are alcohols, alkanes, alkenes, substituted alkanes and alkenes as well as saturated and unsaturated aldehydes (Buttery, Turnbaugh et al. 1988, Jezussek, Juliano et al. 2002, Laguerre, Mestres et al. 2007, Yang, Lee et al. 2010, Bryant and McClung 2011, Calingacion, Boualaphanh et al. 2012, Mathure, Jawali et al. 2014). A suite of these compounds are known to be perceived either by orthonasal or retronasal olfaction (Tieman, Bliss et al. 2012). In orthonasal olfaction, the volatiles are perceived through the nostrils, whereas in retronasal olfaction, the volatiles are directed from the mouth behind the palate into the nasal cavity while the food containing the volatiles is being chewed and swallowed (Tieman, Bliss et al. 2012). Each volatile compound must be at a certain minimum level in order to be perceived. This level can differ by many orders of magnitude for different

compounds and each minimum level is termed as the odour threshold for that particular compound (Leonardos, Kendall et al. 1969). Therefore, volatiles that are present in concentrations above their odour threshold can contribute to the overall aroma of a product (Buttery, Turnbaugh et al. 1988). Previous research was centred on estimating the significance of each volatile compound by calculating the aroma value (AV) or the ratio of its concentration and odour threshold (Lam and Proctor 2003). In other studies aimed at describing the aroma of individual compounds, a volatile compound was able to be correlated to a particular aroma by smelling the volatile as it exits the sniffing port of a gas chromatograph (GC) (Czerny, Christlbauer et al. 2008).

In the study described here (Chapter 4 and 5), two contrasting and commercially-important rice varieties, Apo and IR64, have been used to develop and analyse a population segregating for good yield under water-deficient environments as well as good grain quality. Apo (also known as PSBRc 9) repeatedly gives good yield under drought conditions (Dixit, Singh et al. 2014) and it has been released in several drought-prone areas of Asia. However, this variety has not been widely adopted, because consumers consider it to have poor sensory quality. IR64, on the other hand, is susceptible to a number of biotic and abiotic stresses, including drought (Venuprasad, Lafitte et al. 2007), but due to its premium grain quality traits, it has been grown annually on over a million hectares in major rice-producing countries since its release in 1985 (Fitzgerald, McCouch et al. 2009). The coefficient of parentage between Apo and IR64 has been determined to be 0.13 indicating that these two varieties are genetically quite different (Wang and Lu 2006), which suggests that extensive recombination could occur in a mapping population. Indeed, the genotyping data separated the mapping population into 2 main clusters and several sub-clusters. Furthermore, analysis of the pedigree tree of each shows there are no common parents back at least three generations (Chapter 4, Figure S1).

The differences in aroma between Apo and IR64 have been described previously (Champagne, Bett-Garber et al. 2010). In that study, INQR members submitted eleven matched pairs of rice varieties (where both varieties in each pair had similar physico-chemical traits of quality, but differed in their consumer acceptance/popularity). We contributed Apo and IR64 based on conversations with consumers from regions where Apo had been released and rejected. Consumers particularly described Apo as smelling rancid. The sensory analysis conducted by Champagne et al. (2010) confirmed the consumers' description of Apo, and identified a number of pleasant flavour notes in IR64. Following on

from this, in a sensory analysis conducted as part of this investigation (Chapter 4), trained panellists were able to reproducibly detect aromas of 'corn' and 'sweet aromatic flavours' in IR64 and 'water-like metallic', 'astringent', 'sour silage', 'sewer animal' and 'hay-like/musty' aromas in samples of Apo, in agreement with the previous INQR study (Champagne, Bett-Garber et al. 2010). Similarly, these findings were further supported (Chapter 5) when 26 selected lines from a population derived from Apo and IR64 were evaluated by a sensory panel using the same aroma attributes (Chapter 5).

Interestingly, three lines from the mapping population observed by the panellists to have similar aroma as IR64 while four lines were observed to have similar aroma as that perceived in Apo. Only the Lines 20 and 164 have similar aroma to IR64 and also lie in the same cluster in the archaeopteryx tree containing IR64. Line 165, on the other hand, has similar aroma to IR64 but lies in the other cluster where Apo is also located. Similarly, Lines 9, 16, 22 and 171 that were observed by the panellists to have similar aroma as that perceived in Apo lie in the cluster where IR64 is and not in the cluster of Apo. This reveals that overall genetic similarity of the lines to the parents does not directly translate to aroma similarities.

Given the reproducibility of the sensory panels, and the distinct descriptions of aroma, it is likely that consumers detect these aromas either nasally, retronasally, consciously or subconsciously. This means that there would be considerable value to rice improvement programs to be able to select for these aromatic descriptors in early generations of the breeding cycle. The first step towards selection tools is to detect, and optimally identify, the compounds that define these aromas. Tools to separate and detect volatile compounds have increased in complexity, robustness and accuracy enormously over the past decade. Separation of gases can now be done in two dimensions, with resolution maximised by both separation on columns and computing power. As separation science progresses, detection science does too. The better the separation between compounds, the easier it is for a mass spectrometer to identify daughter fragments from individual compounds, rather than trying to separate fragments informatically from co-eluted compounds. The ability to detect compounds reproducibly that associate with, for example IR64, could be exploited as a breeding tool. However, in order to understand the role of the compound in the flavour and aroma of IR64 we require gains in understanding of chemistry and mass spectra in order to identify the compound, its descriptor and its odour threshold.

In this study two metabolomic platforms were used to identify the volatile compounds present in Apo, IR64 and progeny of the mapping population derived from them (Chapter 5). Using headspace sampling and gas chromatograph-mass spectrometry, 105 compounds were detected in Apo, IR64 and the population. The first two principal components explained more than 50% of the variation in the metabolite profiles with many of the progeny lines clustering between Apo and IR64. Metabolomics analysis revealed that most of the discriminatory compounds, putatively identified on the basis of retention time indexes and by comparison of mass spectra with those of spectral reference libraries, that were associated with Apo and those lines of the population that were clustered close to Apo were alcohols, aldehydes, and ketones. DL-2,3-butanediol and butanol, on the other hand, were more prevalent in IR64 and also in those lines of the population that were clustered closely with IR64.

Previous research from the perfume industry and from food flavour research has provided us with a corpus of literature of aroma descriptors and odour thresholds for a wide range of natural, plant-based volatile compounds. This information as well as published odour threshold values, were used to identify compounds from the total of 105 that might be associated with rice aroma as described by the sensory panel. Several saturated and unsaturated aldehydes, alcohols, ketones, and hydrocarbons were detected in the headspace of the grains which have low odour threshold and which were found to correlate with sensory description. However, assessing the strength of correlations between individual compounds and the responses of the sensory panel provides only a portion of the real picture. Firstly, several compounds can have a different aroma at different concentrations and our understanding of the relationship between concentration and perceived aroma is still low. Secondly, the aroma of a less intense compound can be completely suppressed in the presence of a compound that has stronger aroma intensity (Buettner and Schieberle 2000). Thirdly, when compounds are mixed together the resulting perceived aroma could be something altogether different from the individual descriptors (Masanetz, Guth et al. 1998).

Because Apo is drought-tolerant while IR64 is susceptible, the volatiles were analysed which were present in both their grains that were grown in irrigated and drought conditions using a comprehensive GCxGC-time of flight-mass spectrometer (TOF-MS) and conducted a sensory evaluation of both samples (Chapter 4). The sensory analysis and aromatic profile of IR64 were both affected by water condition i.e. principal component analysis

(PCA) of the metabolites showed that IR64 formed two clusters based on water condition and panellists were able to perceive aroma differences in IR64 that were grown in irrigated and drought conditions. On the other hand, Apo from irrigated and drought plots both have high levels of 'water-like metallic', 'sewer/animal', and 'hay-like/musty' aromas and formed only one cluster in the PCA. This suggests that response to water stress in the metabolomic profile of the grain is variety dependent.

In the case of rice aroma, with the exception of 2AP, little is yet known about the genes responsible for the synthesis of other aromatic compounds or about the inheritance of these genes. Our sensory analysis of IR64 and Apo and the progeny indicates that the aromatic description is reproducible and is heritable. As a first step to finding genes for the compounds identified that correlate with the sensory descriptions a dense linkage map based on GBS was used. The work reported here has led to the identification of six novel metabolite QTLs for volatile compounds in rice (Chapter 5). One QTL was detected on chromosome 1 for 3,7-dimethyl-7-octen-1-ol and 1 QTL for hexanol on chromosome 2. Interestingly, four major QTLs associated with pentanol, hexanol, hexanal, and heptanone were mapped to almost the same region in chromosome 3. It is possible that these compounds are from the product of oxidation of fatty acids, catalysed by lipoxygenase (LOX) and alcohol dehydrogenase (ADH) (Dudareva, Negre et al. 2006) rather than direct genetic products. Several QTLs for lipoxenase genes have been mapped on several chromosomes in rice, including four putative LOX genes on chromosome 3 (Umate 2011). The location of our QTL does not fall within the previously mapped QTLs. This could suggest that aldehydes are more likely to be products of oxidation rather than being direct genetic products. Many of the aldehydes have low odour thresholds and pleasant aroma, and should be considered key candidates for selection of aroma profiles. In order to determine whether the alcohols, alkanes, alkenes and aldehydes are the product of individual genes or auto-oxidation, further studies need to be conducted on tracing the volatile metabolome from the point of harvest, through postharvest processing and by using different storage conditions. By understanding this, rice programs can be supported with additional genetic selection tools and postharvest management techniques.

Combining Yield and Quality

As emphasised earlier, rice breeding programmes need to be equipped with the tools needed to facilitate the creation of varieties which combine both yield potential (in sub-optimal environmental conditions) and quality traits relevant to consumers. To be able to

capture the premium grain quality of IR64 and agronomic adaptation of Apo, the two contrasting genotypes chosen for this study, a population derived from their mapping population was grown under irrigated and drought conditions. As previously described in Chapter 4, yield of Apo under drought conditions was not reduced significantly, nor did the aromatic quality appear to change. In contrast, yield of IR64 halved under drought, and the aromatic quality changed substantially. In addition, the yield of Apo was higher than the yield obtained from IR64 grown under irrigated conditions. Moreover, the yield of >50% of the progeny was higher than yield of either Apo or IR64 under both irrigation and drought, indicating significant transgressive segregation (Chapter 5), which suggests that recombination has occurred at many of the loci that govern yield and drought tolerance.

In terms of volatile compounds, several lines in the population were identified which had similar metabolomic and sensory properties to the high quality parent, IR64. Line 20 and 164 were both perceived by the panellists to have high levels of favourable features (corn, dairy and sweet aromatic) which are similar to IR64. Importantly, the yield of these lines under both irrigated and drought conditions was similar to that of the higher yielding parent, Apo, grown under the same conditions. Moreover, Line 28, which had the highest yield under drought and which has an important QTL associated with yield under drought on chromosome 3, was also perceived by the panellists to have similar aroma to that of IR64. The excellent yield potential of these three lines, together with the aromatic quality that they possess warrants introduction into the replicated phase of breeding cycles for large, multi-environment testing.

Conclusions

It is already evident that the direction rice breeding programs are taking, is strongly influenced by the growing volume of accumulated information arising from rice genomics research (Li, Zhang et al. 2011). As is true also of a growing number of other crops, marker-assisted selection through the use of selected DNA markers linked to the phenotypic traits of interest has enormously improved the efficiency and precision of breeding programs and has decreased the time to market of new varieties (Collard and Mackill 2008). Furthermore, this is an important and exciting time in rice breeding with the growing capacity to (re)sequence rice genomes and to generate population-size datasets of rice genome data. Programmes such as the 3000 Rice genomes project are anticipated to revolutionise how we proceed in the future with rice breeding. The more tools we have available to generate complementary data to those coming from the genomics approaches, the stronger position

we will gain and the greater discriminative power we will have to define better breeding targets and translate knowledge into selection tools.

In this study, the huge potential of a multi-disciplinary approach involving genotyping, metabolomics and other phenotyping tools such as sensory analysis, and supported by advanced multivariate statistical analyses has been demonstrated to generate a better understanding of the aroma of rice. With the completion of the resequencing of 3,000 rice accessions (The3000ricegenomesproject 2014), QTLs and probably genes associated with compounds that are responsible for distinct aromas of rice will be helpful in advancing our understanding of the complexity of undertaking a molecular-assisted aroma breeding programme.

Three lines have been identified in this study which carry the agronomic adaptation and stress tolerance of Apo, and the grain quality of IR64. So, in just three years, with the use of next generation genotyping and advanced phenotyping, these three lines have been identified and can now be introduced into the late stages of a breeding program for replicated multi-location testing. By proceeding with a population derived from genetically distinct parents, targeted research aims, advanced field scale and laboratory scale phenotyping, it should be possible to shed years from the conventional breeding cycle.

References

ADB (2012). Food security and poverty in Asia Pacific. Key Challenges and Policy Issues. Philippines, Asian Development Bank.

ANZ (2012). Greener pastures: The global soft commodity opportunity for Australia and New Zealand. ANZ Insights 3: 1-80.

Bernier, J., A. Kumar, R. Venuprasad, D. Spaner and G. Atlin (2007). A large-effect QTL for grain yield under reproductive-stage drought stress in upland rice. Crop Science **47**(2): 507-516.

Boualaphanh, C., M. Calingacion, R. Cuevas, D. Jothityangkoon, J. Sanitchon and M. Fitzgerald (2011). Yield and quality of traditional and improved Lao varieties of rice. Scienceasia **37**(2): 89-97.

Bryant, R. and A. McClung (2011). Volatile profiles of aromatic and non-aromatic rice cultivars using SPME/GC–MS. Food Chemistry **124**(2): 501-513.

Buettner, A. and P. Schieberle (2000). Influence of mastication on the concentrations of aroma volatiles - some aspects of flavour release and flavour perception. Food Chemistry **71**(3): 347-354.

Buttery, R., J. Turnbaugh and L. Ling (1988). Contribution of volatiles to rice aroma. Journal of Agricultural and Food Chemistry **36**(5): 1006-1009.

Calingacion, M., C. Boualaphanh, V. Daygon, R. Anacleto, R. Sackville Hamilton, B. Biais, C. Deborde, M. Maucourt, A. Moing, R. Mumm, R. de Vos, A. Erban, J. Kopka, T. Hansen, K. Laursen, J. Schjoerring, R. Hall and M. Fitzgerald (2012). A genomics and multi-platform metabolomics approach to identify new traits of rice quality in traditional and improved varieties. Metabolomics **8**(5): 771-783.

Chambers, E. and K. Koppel (2013). Associations of volatile compounds with sensory aroma and flavor: The complex nature of flavor. Molecules **18**(5): 4887-4905.

Champagne, E., K. Bett-Garber, M. Fitzgerald, C. Grimm, J. Lea, K. Ohtsubo, S. Jongdee, L. Xie, P. Bassinello, A. Resurreccion, R. Ahmad, F. Habibi and R. Reinke (2010). Important sensory properties differentiating premium rice varieties. Rice **3**: 270–281.

Champagne, E., K. Bett-Garber, J. Thomson and M. Fitzgerald (2009). Unraveling the impact of nitrogen nutrition on cooked rice flavor and texture. Cereal Chemistry **86**(3): 274-280.

Collard, B. and D. Mackill (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Philosophical Transactions of the Royal Society B: Biological Sciences **363**(1491): 557-572.

Czerny, M., M. Christlbauer, M. Christlbauer, A. Fischer, M. Granvogl, M. Hammer, C. Hartl, N. Hernandez and P. Schieberle (2008). Re-investigation on odour thresholds of key food aroma compounds and development of an aroma language based on odour qualities of defined aqueous odorant solutions. European Food Research and Technology **228**(2): 265-273.

Dixit, S., A. Singh, M. Sta Cruz, P. Maturan, M. Amante and A. Kumar (2014). Multiple major QTL lead to stable yield performance of rice cultivars across varying drought intensities. BMC Genetics **15**(1): 1-13.

Dudareva, N., F. Negre, D. Nagegowda and I. Orlova (2006). Plant volatiles: Recent advances and future perspectives. Critical Reviews in Plant Sciences. **25**: 417-440.

FAOSTAT (2013).

Fitzgerald, M., S. McCouch and R. Hall (2009). Not just a grain of rice: the quest for quality. Trends in Plant Science **14**(3): 133-139.

Fitzgerald, M., N. Sackville Hamilton, M. Calingacion, H. Verhoeven and V. Butardo (2008). Is there a second fragrance gene in rice? Plant Biotechnology Journal **6**(4): 416-423.

Fitzgerald, T., D. Waters, L. Brooks and R. Henry (2010). Fragrance in rice (Oryza sativa) is associated with reduced yield under salt treatment. Environmental and Experimental Botany **68**(3): 292-300.

Gamuyao, R., J. Chin, J. Pariasca-Tanaka, P. Pesaresi, S. Catausan, C. Dalid, I. Slamet-Loedin, E. Tecson-Mendoza, M. Wissuwa and S. Heuer (2012). The protein kinase Pstol1 from traditional rice confers tolerance of phosphorus deficiency. Nature **488**(7412): 535-539.

Hall, R., I. Brouwer and M. Fitzgerald (2008). Plant metabolomics and its potential application for human nutrition. Physiologia Plantarum **132**(2): 162-175.

Huang, X., Y. Zhao, X. Wei, C. Li, A. Wang, Q. Zhao, W. Li, Y. Guo, L. Deng, C. Zhu, D. Fan, Y. Lu, Q. Weng, K. Liu, T. Zhou, Y. Jing, L. Si, G. Dong, T. Huang, T. Lu, Q. Feng, Q. Qian, J. Li and B. Han (2012). Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. Nature Genetics **44**(1): 32-39.

IPCC (2014). IPCC Fifth Assessment Report: Working Group II Report Impacts, adaptation and vulnerability.

Jezussek, M., B. Juliano and P. Schieberle (2002). Comparison of key aroma compounds in cooked brown rice varieties based on aroma extract dilution analyses. Journal of Agricultural and Food Chemistry **50**(5): 1101-1105.

Kovach, M., M. Calingacion, M. Fitzgerald and S. McCouch (2009). The origin and evolution of fragrance in rice (Oryza sativa L.). Proceedings of the National Academy of Sciences **106**(34): 14444-14449.

Kumar, A., S. Dixit, T. Ram, R. Yadaw, K. Mishra and N. Mandal (2014). Breeding high-yielding drought-tolerant rice: genetic variations and conventional and molecular approaches. Journal of Experimental Botany 1-14.

Laguerre, M., C. Mestres, F. Davrieux, J. Ringuet and R. Boulanger (2007). Rapid discrimination of scented rice by solid-phase microextraction, mass spectrometry, and multivariate analysis used as a mass sensor. Journal of Agricultural and Food Chemistry **55**(4): 1077-1083.

Lam, H. and A. Proctor (2003). Milled rice oxidation volatiles and odor development. Journal of Food Science **68**(9): 2676-2681.

Leonardos, G., D. Kendall and N. Barnard (1969). Odor threshold determinations of 53 odorant chemicals. Journal of the Air Pollution Control Association **19**(2): 91-95.

Li, J., H. Zhang, D. Wang, B. Tang, C. Chen, D. Zhang, M. Zhang, J. Duan, H. Xiong and Z. Li (2011). Rice omics and biotechnology in China. Plant Omics **4**(6): 302-317.

Mackill, D., A. Ismail, U. Singh, R. Labios, T. Paris and L. Donald (2012). Chapter Six - Development and rapid adoption of submergence-tolerant (Sub1) rice varieties. Advances in Agronomy, Academic Press **115**: 299-352.

Manzanilla, D., T. Paris, G. Vergara, A. Ismail, S. Pandey, R. Labios, G. Tatlonghari, R. Acda, T. Chi, K. Duoangsila, I. Siliphouthone, M. Manikmas and D. Mackill (2011). Submergence risks and farmers' preferences: Implications for breeding Sub1 rice in Southeast Asia. Agricultural Systems **104**(4): 335-347.

Maraval, I., K. Sen, A. Agrebi, C. Menut, A. Morere, R. Boulanger, F. Gay, C. Mestres and Z. Gunata (2010). Quantification of 2-acetyl-1-pyrroline in rice by stable isotope dilution assay through headspace solid-phase microextraction coupled to gas chromatography–tandem mass spectrometry. Analytica Chimica Acta **675**(2): 148-155.

Masanetz, C., H. Guth and W. Grosch (1998). Fishy and hay-like off-flavours of dry spinach. Zeitschrift für Lebensmitteluntersuchung und -Forschung A **206**(2): 108-113.

Mathure, S., N. Jawali, R. Thengane and A. Nadaf (2014). Comparative quantitative analysis of headspace volatiles and their association with BADH2 marker in non-basmati scented, basmati and non-scented rice (Oryza sativa L.) cultivars of India. Food Chemistry **142**: 383-391.

Meilgaard, M., G. Civille and B. Carr (2007). Sensory evaluation techniques. Boca Raton, FL, USA, CRC Press.

Meullenet, J., B. Marks, J. Hankins, V. Griffin and M. Daniels (2000). Sensory quality of cooked long-grain rice as affected by rough rice moisture content, storage temperature, and storage duration. Cereal Chemistry **77**(2): 259-263.

Moi, L., G. Srzednicki and J. Craske (2003). The effects of high-temperature drying on fragrant rice. Advances in Stored Product Protection 439-447.

Sandhu, N., A. Singh, S. Dixit, M. Cruz, P. Maturan, R. Jain and A. Kumar (2014). Identification and mapping of stable QTL with main and epistasis effect on rice grain yield under upland drought stress. BMC Genetics **15**(1): 63.

Shobana, S., N. Malleshi, V. Sudha, D. Spiegelman, B. Hong, F. Hu, W. Willett, K. Krishnaswamy and V. Mohan (2011). Nutritional and sensory profile of two Indian rice varieties with different degrees of polishing. International Journal of Food Sciences and Nutrition **62**(8): 800-810.

The3000ricegenomesproject (2014). The 3,000 rice genomes project. GigaScience **3**(1): 1-6.

Tieman, D., P. Bliss, L. McIntyre, A. Blandon-Ubeda, D. Bies, A. Odabasi, G. Rodríguez, E. van der Knaap, M. Taylor, C. Goulet, M. Mageroy, D. Snyder, T. Colquhoun, H. Moskowitz,

D. Clark, C. Sims, L. Bartoshuk and H. Klee (2012). The chemical interactions underlying tomato flavor preferences. Current Biology **22**(11): 1035-1039.

Umate, P. (2011). Genome-wide analysis of lipoxygenase gene family in Arabidopsis and rice. Plant Signaling and Behavior **6**(3): 335-338.

Vanavichit, A. and T. Yoshihashi (2010). Molecular aspects of fragrance and aroma in rice. Advances in Botanical Research **56**: 50-73.

Venuprasad, R., H. Lafitte and G. Atlin (2007). Response to direct selection for grain yield under drought stress in rice. Crop Science **47**(1): 285-293.

Wang, S. and Z. Lu (2006). Genetic diversity among parental lines of Indica hybrid rice (Oryza sativa L.) in China based on coefficient of parentage. Plant Breeding **125**(6): 606-612.

Wassmann, R., S. Jagadish, S. Heuer, A. Ismail, E. Redona, R. Serraj, R. Singh, G. Howell, H. Pathak and K. Sumfleth (2009). Climate change affecting rice production: The physiological and agronomic basis for possible adaptation strategies. Advances in Agronomy. D. Sparks, Academic Press. **101**: 59-122.

Xu, K., X. Xu, T. Fukao, P. Canlas, R. Maghirang-Rodriguez, S. Heuer, A. Ismail, J. Bailey-Serres, P. Ronald and D. Mackill (2006). Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. Nature **442**(7103): 705-708.

Yang, D., K. Lee and S. Kays (2010). Characterization and discrimination of premium-quality, waxy, and black-pigmented rice based on odor-active compounds. Journal of the Science of Food and Agriculture **90**(15): 2595-2601.

Yang, D., R. Shewfelt, K. Lee and S. Kays (2008). Comparison of odor-active compounds from six distinctly different rice flavor types. Journal of Agricultural and Food Chemistry **56**(8): 2780-2787.

Zeigler, R. S. and S. Mohanty (2010). Support for international agricultural research: Current status and future challenges. New Biotechnology **27**(5): 565-572.

Summary

Most rice breeding programs have focused on improving agronomic traits such as yield, while enhancing grain quality traits such as flavour and aroma, especially of non-fragrant rices, has not been given high priority. In this study, we utilised a multi-disciplinary approach to understand better quality traits of aroma and flavour in rice grains, and to determine whether good flavour in the grain could be combined with stress tolerant genotypes.

To understand what factors drive rice preferences, an extensive survey among members of the International Network for Quality Rice who are local experts in grain quality evaluation programs in 25 countries was conducted (Chapter 2). The objective was to identify the grain quality characteristics of the popular rice varieties in each region. Eighteen combinations of size and shape of the grain, amylose content (AC), gelatinisation temperature (GT) and fragrance were identified. These trait combinations reveal the complexity of consumer preferences. The two most popular combinations both have long and slender grains, while one has low amylose, low GT and is aromatic, and the other has intermediate AC and intermediate GT and is non-aromatic. Further evaluation of varieties having the same combination of grain quality traits showed that consumers readily identify differences between these varieties. For example, BRS Primavera and IR64 that are popular in Brazil and in the Philippines, respectively, have the same combination of all 18 traits, however, panellists of sensory evaluation can easily perceive differences in aroma and flavour of BRS Primavera and IR64. This emphasises that the current tools we have available to assess rice quality are unable to capture all the quality traits consumers are looking for in rice.

In Chapter 3, a novel multiplatform metabolomic and ionomic approach with genome-wide genotyping was utilised to investigate the effect of different nitrogen fertiliser regimes on the biochemical profile of three premium waxy rice varieties, Hom Nang Nouane (HNN), Kai Noi Leuanag (KNL) and Tha Sa No (TSN) from Lao PDR. The current tools used to phenotype grain quality such as GT, values from viscosity curves, and hardness and stickiness, were unable to differentiate between HNN, KNL and TSN either on the basis of nitrogen treatment nor genotype. However, metabolite profiling of metabolites and minerals followed by multivariate statistical methods readily separated the genotypes on each platform, and discriminatory compounds that were identified were relevant to consumers in terms of flavour, taste and nutrition. However, despite yield differences, nitrogen treatment did not significantly affect the overall metabolite and mineral profiles of the samples. Using 1536

single nucleotide polymorphism (SNP) loci, the Euclidean distance between each variety was calculated and compared to the distance between each variety for each metabolomic platform. Procrustes analysis was used to rotate and scale the variety mean scores on the metabolite principal components to give the best fit to the genetic principal coordinates. Comparing the triangles whereby each vertex of the triangle is a variety and the length of each side is equal to the scaled Euclidean distance, mineral elements, polar metabolites and volatile compounds all associate very well with the genetic distance between each variety. This study highlights that multiple metabolomic platforms are potential phenotyping tools to characterise rice quality in a comprehensive and efficient way, and in a way that provides data that is relevant to consumers.

To gain insights on the influence of water availability to the metabolomic profile of drought tolerant rice, two contrasting varieties, Apo and IR64, and a mapping population derived from them were extensively characterised in Chapters 4 and 5. Apo is drought tolerant but has unacceptable grain quality while IR64 is drought susceptible with premium grain quality. Apo and IR64 were grown under irrigated and drought conditions. Yield of Apo from both water conditions was higher than yield of IR64 under the same conditions. Moreover, metabolite profiling and sensory analysis showed that grains of Apo were not affected by drought conditions i.e. panellists perceived no difference in the aroma of Apo from both conditions and Principal Components Analysis (PCA) of the volatiles showed one cluster of Apo from both conditions. However, grains of IR64 formed two clusters based on water condition in the PCA and panellists were able to perceive ‘water-like metallic’ aroma in IR64 that was grown under drought conditions but this was not detected in grains from the irrigated treatment. This suggests that response to water stress in the metabolomic profile of the grain is variety dependent.

In Chapter 5, a mapping population derived from Apo and IR64 was grown, with the parents, under irrigated and drought conditions. The yield of more than half of the population was higher than the yield of Apo and IR64 under both irrigated and drought conditions; this indicates significant transgressive segregation. Using a dense linkage map based on genotyping by sequencing data, quantitative trait loci (QTL) analysis of drought stress identified one major QTL on chromosome 3 that is likely to be $qDTY_{3.1}$ which was previously detected in a population derived from Apo as the drought tolerant parent. All the lines of the population carrying this QTL showed significantly higher yield under drought than those without it, indicating the potential importance of this QTL in drought tolerance.

Metabolite profiling and sensory analysis were also conducted in the grains of the population. More than a hundred volatiles were detected in the headspace of rice samples and PC1 and PC2 explained 55.6% of the variation in the metabolite profiles with many of the lines clustering in between the Apo and IR64 parent values. Six novel metabolite QTLs for volatile compounds were identified - 1 QTL was detected in chromosome 1 for 3,7-dimethyl-octen-1-ol, 1 QTL for hexanol in chromosome 2, and 4 QTLs for pentanol, hexanol, hexanal, and heptanone in chromosome 3. Interestingly, three lines were observed by the panellists to have similar aroma as IR64 while four lines were observed to have similar aroma as that perceived in Apo. Lines 20, 164 and 28 were perceived by the panellists to have high levels of corn, dairy and sweet aromatic features. Moreover, the yield of these 3 lines under both irrigated and drought conditions was similar to that of the Apo parent under the same conditions with Line 28 yielding the highest under drought and has the QTL associated with yield under drought on chromosome 3.

Finally, the potential of metabolomics as a phenotyping tool in characterising grain quality is further highlighted in Chapter 6. Combining metabolomics with high throughput genotyping and sensory analysis offers new breadth of approach in understanding grain quality of rice. Three lines identified that carry IR64 quality along with high yield in both irrigation and drought, are recommended to enter a rice breeding program at the stage of advanced replicated and multi-location testing. By using advanced tools of phenotyping and genotyping, with validation by sensory panels, these three advanced lines have been selected in just three years.

Samenvatting

Rijstveredelingsprogramma's zijn tot nu toe vooral gericht geweest op verbetering van agronomisch belangrijke kenmerken als opbrengst en korrelvorm, en nauwelijks op verbetering van kwaliteitseigenschappen van de korrel zélf, zoals smaak en aroma. In dit onderzoek is gebruik gemaakt van een multi-disciplinaire aanpak om smaak en aroma van rijst beter te karakteriseren en te bepalen of de goede aroma eigenschappen aanwezig in sommige genotypen kunnen samengaan met droogteresistentie van niet-geurende genotypen.

Allereerst is er een omvangrijke enquête uitgevoerd onder vertegenwoordigers van het Internationaal Netwerk voor Rijstkwaliteit, de lokale experts op het gebied van korrelkwaliteit, om te bestuderen welke eigenschappen een rol spelen bij de consumentenvoorkeur voor een bepaald type rijst in 25 landen (Hoofdstuk 2). Het doel was om per regio de belangrijkste kwaliteitseigenschappen van de meest populaire rijstsoorten vast te stellen. Er is gekeken naar verschillende combinaties van 18 verschillende uiterlijke kenmerken (fenotype), waaronder grootte en vorm van de korrel, het amylosegehalte (AC), gelering temperatuur (GT) en geur, om de variatie in consumentenvoorkeuren te kunnen beschrijven. De twee meest populaire combinaties van eigenschappen, beide aanwezig in lange en dunne korrels, waren I) aromatisch samen met lage AC en GT, en II) niet-aromatisch met middelmatige GT. Consumenten konden echter duidelijk verschil proeven tussen verschillende variëteiten met dezelfde combinatie van eigenschappen. Bijvoorbeeld de variëteiten BRS Primavera en IR64, populair in respectievelijk Brazilië en de Filippijnen, hebben ongeveer dezelfde combinatie van kwaliteitseigenschappen, maar de panelleden waren desondanks goed in staat deze variëteiten te onderscheiden op basis van smaak. Deze studie geeft aan dat de algemene kenmerken die momenteel gebruikt worden bij de beoordeling van de kwaliteit van rijst onvoldoende zijn om alle consumentenvoorkeuren goed te ondervangen.

Hoofdstuk 3 beschrijft het gebruik van uitgebreide metabolietprofielering technieken ("metabolomics"), uitgevoerd met verschillende analytische instrumenten om zoveel mogelijk metabolieten te kunnen detecteren, en mineraalanalyse om gedetailleerd de effecten van stikstofbemesting op de biochemische kwaliteit van rijstkorrels vast te stellen. Hierbij werden 3 verschillende variëteiten gebruikt die alle een hoge kwaliteit rijstkorrel leveren: Hom Nang Nouane (HNN), Kai Noi Leuanag (KNL) en Tha Sa No (TSN), alle

afkomstig uit Laos. Terwijl de huidige kwalificeringsmethode op basis van algemen kenmerken onvoldoende bleek om onderscheid te maken tussen de 3 rijstvariëteiten, gaven de profielen van zowel metabolieten, onafhankelijk van gebruikte analyseplatform, als mineralen duidelijke verschillen aan. De bekende verschillen in voedingskwaliteit van deze rijstvariëteiten konden worden gerelateerd aan de meest differentiële metabolieten, waaronder geur-, smaak- en voedingsstoffen. Stikstofbemesting bleek geen significant effect te hebben op de samenstelling van metabolieten en mineralen in de korrels. Op basis van 1536 genetische merkers ("single nucleotide polymorphism", SNP) kon de genetische afstand tussen de 3 variëteiten worden berekend en vergeleken met de verschillen in de biochemische profielen. De genetische afstand correleerde goed met de verschillen in zowel mineralen als polaire metabolieten en vluchtige stoffen. Deze studie toont aan dat deze uitgebreide metabolietanalyses goede mogelijkheden bieden om rijstvariëteiten te onderscheiden en te fenotyperen, bijvoorbeeld ten behoeve van het vaststellen van hun consumentkwaliteit.

Om meer inzicht te krijgen in de rol van waterbeschikbaarheid op de metabolietsamenstelling van rijst, werden 2 contrasterende variëteiten, Apo en IR64, geteeld met en zonder irrigatie en uitgebreid gekarakteriseerd (Hoofdstukken 4 en 5). Apo is droogte-tolerant maar de korrels hebben matige korrelkwaliteit, terwijl IR64 droogte-gevoelig is met hoge, aromatische korrelkwaliteit. Onder beide waterregimes was de opbrengst van Apo hoger dan die van IR64. Analyse van metabolieten en smaak gaven aan dat droogte geen aantoonbaar effect had op de biochemische samenstelling van de korrels van droogte-resistente variëtiet Apo, in tegenstelling tot die van droogte-gevoelige IR64: de panelleden konden een duidelijke "waterige metaalsmaak" waarnemen in IR64-rijst geteeld onder droge condities. Dit geeft aan dat het effect van droogtestress op het biochemische profiel van de rijskorrel afhankelijk is van genotype.

Hoofdstuk 6 beschrijft het effect van waterbeschikbaarheid op de rijstkwaliteit in een kruisingspopulatie van deze contrasterende ouderlijnen Apo en IR64. Bij meer dan de helft van de nakomelingen was de opbrengst hoger dan die van beide ouderlijnen, onafhankelijk van waterregime. Dit resultaat toont een significante transgressief effect aan. Door de gevonden fysiologische verschillen onder invloed van waterbeschikbaarheid te koppelen aan genetische variatie, verkregen door sequentieanalyses, konden er locaties op het genoom ("quantitative trait loci", of kortweg QTL) worden aangewezen die mogelijk een rol spelen in droogteresistentie van rijst. Een belangrijke QTL lag op chromosoom 3, zeer

waarschijnlijk overeenkomstig met qDTY3,1, een QTL dat onlangs is ontdekt in een andere Apo-populatie. Alle lijnen van onze Apo-IR64 populatie die dit QTL bevatten gaven significant meer opbrengst onder droogtestress dan lijnen zonder dit QTL. Dit resultaat geeft aan dat dit QTL een belangrijke rol kan spelen bij droogteresistentie in rijst.

Dezelfde kruisingspopulatie werd verder uitgebreid biochemisch gekarakteriseerd met behulp van metabolomics en smaakproeven (Hoodstuk 7). Meer dan honderd vluchtige stoffen afkomstig van de rijstkorrels konden worden gedetecteerd, en de variatie in de samenstelling van deze vluchtige stoffen in de populatie lag grotendeels tussen de twee extreme ouderlijnen. Met behulp van QTL analyse op al deze vluchtige stoffen konden 6 nieuwe genoomlocaties worden aangewezen die mogelijk een rol spelen bij de ophoping van deze stoffen in de rijstekorrel: 1 QTL op chromosoom 1 voor 3,7-dimethyl-octen-1-ol, 1 QTL op chromosoom 2 voor hexanol, en 4 QTLs op chromosome 3 voor respectievelijk pentanol, hexanol, hexanal en heptanone. Van 3 lijnen werd het aroma door de panelleden evengoed gewaardeerd als de sterk-aromatische ouderlijn IR64, terwijl 4 lijnen hetzelfde aroma vertoonden als de niet-geurende ouderllijn Apo. De lijnnummers 20, 164 en 28 hadden een zeer specifiek aroma dat door de panelleden werd beoordeeld met hoge scores voor popcorn, zuivel en aromatisch. De opbrengst van deze 3 lijnen was even goed als die van de Apo ouderllijn bij zowel droge als geirrigeerde groeicondities, terwijl lijn 28 zelfs beter produceerde onder droge dan onder geirrigeerde condities. Deze lijn bezit ook de bovengenoemde QTL op chromosoom 3 voor droogteresisitentie.

Tenslotte is dieper ingegaan op de potentie van metabolomics als fenotyperingsmiddel voor het karakteriseren van de kwaliteit van rijstkorrels. De combinatie van metabolomics met ‘high-throughput’ genotypering en sensorische analyse biedt een geheel nieuwe benadering om rijstkwaliteit te beschrijven en te begrijpen. De 3 lijnen die het IR64 genetische eigenschappen voor hoge rijstkwaliteit bevatten én een hoge opbrengst gaven bij zowel irrigatie als droogte, zijn goede kandidaten om te gebruiken en uitgebreid te testen in een rijstveredelingsprogramma. Door gebruik te maken van de modernste technieken voor fenotypering en genotypering, en validatie met smaakpanels, konden 3 veelbelovende nieuwe lijnen binnen een tijdsbestek van 3 jaar worden geselecteerd.

Acknowledgment

This research would not have been possible without the help of so many people to whom I am truly grateful.

I sincerely express my gratitude to Monsanto Beachell-Borlaug International Scholarship Program, International Rice Research Institute and University of Queensland who generously provided funds for this research.

I am honoured and truly grateful for the guidance and supervision of Profs Robert Hall and Mellisa Fitzgerald. This research would not have been what it is right now without your valuable insights. Thank you for encouraging me to focus on the goals and patiently pushing me to be on track. Thank you for being my mentor professionally and personally for which I will forever be in your debt. I also thank Dr Roland Mumm for all the intellectual discussions and hands on supervision. I would also like to thank Harro Bouwmeester for accommodating me in Plant Physiology group and assisting me in all the formalities of WUR as my prinicipal promoter for the first part of my PhD.

I would also like to thank Chanthakhone Boualaphanh and our METAPHOR collaborators for helping out in analysing Lao rice varieties using different metabolomics and ionomics platforms. I greatly appreciate all the members of International Network for Quality Rice for actively participating in the consumer preference survey. Thank you to Alice Laborte and Andrew Nelson for patiently constructing the consumer preference maps. I am also grateful to Lu Fang for analysing the 'parents samples' in GCxGC TOF MS.

I would also like to express my warm thanks to everyone who helped me in the lab and field. I thank Harry Jonker, Jacques Davies, Bert Schipper, Jeroen van Arkel and Francel Verstappen for rendering their assistance as I conducted experiments in Wageningen. I also thank Pieter Lukasse for always accommodating my inquiries in using Galaxy. I would not have survived 2 sun-scorching and storm- wrecking seasons growing rice in IRRI fields without the help of Ferdie Salisi, Roldan Ilagan, Lucy Cuevas, Jun Madrid, Cindy Llorente, Rochie Cuevas and Lenie Baeza – thank you for all those extra hardworking hands. I also thank Dory Resurreccion, Cindy Llorente and Roslen Anacleto for facilitating all the paperwork needed so my rice samples would arrive at their destination promptly and safely. Thank you to Chitra Raghavan for patiently teaching me how to analyse my GBS data. I

also thank Lenie Baeza and Crystal Concepcion for helping me look for the elusive metabolite QTLs. I thank Arthur Riedel for sharing his statistical and artistic skills. I greatly appreciate Sangeeta Prakash, Kevin Tan and Dara Daygon for helping out in conducting the sensory analysis under a very tight schedule. I also thank Hongyan Li, Tamara Stanley, Ervinna Ang, Fariz Widjaja, Maria Cristina Vegafria and Alexander Kanon for taking the challenge of sniffing rice flour for our sensory evaluation studies.

I am also grateful to our Dutch friend Ric de Vos for translating the summary of this work to his native tongue, Nghia Tran (Khang) for capturing the beauty of a bowl of rice that I used as the cover photo of this book, and my husband Riz for doing the lay out of this cover in a simple yet artistic way. I would also like to thank Connie Remoroza and Yuanyuan Zhang for accepting the role of being my paranymphs and helping out in all the preparations before the big day.

I would like to thank Ana Genil, Anilyn Maningas and Froilan Fule of IRRI, Rina Anthonijsz, Douwe Zuidema, Marion Rodenburg and Rudie Jullens of WUR, and Lyndall Dooley of UQ for coordinating all documents and fund transfer needed as I moved from one lab to another. I also convey my thanks to the MMM group especially Ric de Vos, Roland Mumm, Bert Schipper, Desalegn Etalo, Yury Tikunov, Natalia Carreno-Quintero, Carlos Carollo, Jeroen van Arkel, and Anna Undas for the valuable discussions. I've learned a lot from your research projects and appreciated metabolomics research even more.

And finally, thank you to all the friends I met in so many places and helped keep me sane throughout my PhD. Thank you to GQNC family for cheering for me everytime I felt homesick, "Pere and friends" for sharing the good times in Netherlands, and FilOz and ITTC groups for welcoming me in Australia. Thank you to Nova, Elenette, May, Archie, Babes, Tetet, Bec, Jhoana, Riza and Mercy for pulling me up when I felt really low. Words cannot express how thankful I am to Mommy, Tatay, Sobun, Lola Sofia, Lola Lita and Tita Matet for all the encouragements in pursuing my dreams and care you have showered to Hannah while I'm away. Most importantly, I wholeheartedly thank Riz and Hannah for all the love, understanding and support. It has been quite a journey and you are my constant source of inspiration.

Biography

Mariafe Navarro Calingacion was born on September 19th, 1976 in Quezon City, Manila, Philippines. She attended elementary and high school at the Post Elementary School and Colegio de San Juan de Letran Calamba, respectively, and finished with honours.

She successfully pursued a Bachelor of Science in Chemistry at University of the Philippines Los Banos (UPLB) and studied coconut coir dust as potential cation exchanger for heavy metal remediation as her undergraduate thesis. She was awarded a BSc in Chemistry in 1998. Her passion for research led her to take on a research assistant position at Chemical Control, Pesticide Toxicology and Chemistry Laboratory of National Crop Protection Center (NCPC) in UPLB where she conducted pesticide residue analysis and published a number of papers on this work. While working in NCPC, she took Master of Science in Agricultural Chemistry also at UPLB and continued working on bioremediation studies of pesticide residues for her graduate Masters thesis.

She moved to Grain Quality and Nutrition Centre of International Rice Research Institute (IRRI) as a researcher in 2005. She was then heavily involved in fragrance and micronutrient studies. Later she was granted a Professional Growth Committee award that allowed her to expand her knowledge and undergo a three months training period in metabolomics at Plant Research International, The Netherlands in collaboration with the Metabolomics for Plant, Health and Outreach program (META-PHOR) funded by the European Union. Immediately after this training, she was awarded a Monsanto Beachell-Borlaug International Scholars Fellowship which funded her research to pursue a PhD degree in Wageningen University (WUR). She gave up her permanent position to be able to carry out this PhD project under the WUR PhD Sandwich program and became an IRRI Scholar for the period of her PhD work. She conducted her field and grain quality studies in IRRI, carried out metabolite profiling using state-of-the-art platforms in WUR and performed sensory analysis with a trained panel in University of Queensland (UQ).

Currently, she is based in Brisbane with her husband, Riz and daughter, Hannah, and is enjoying her stint in UQ as a postdoc research fellow. Her current project focused on understanding the biochemical pathway of fragrance in rice.

List of publications

Calingacion, M, Fang, L, Quiatchon-Baeza, L, Mumm, R, Riedel, A, Hall, R and Fitzgerald, M. 2014. Delving deeper into technological innovations to understand differences in rice quality. Rice, in press. [This thesis]

Calingacion, M, Laborte, A, Nelson, A, Resurreccion, A, Concepcion, C, Daygon, V, Mumm, R, Russell, R, Dipti, S, Bassinello, P, Manful, J, Sophany, S, Lara, K, Bao, J, Xie, L, Loaiza, K, El-hissewy, A, Gayin, J, Sharma, N, Rajeswari, S, Manonmani, S, Shobha Rani, N, Kota, S, Indrasari, S, Habibi, F, Hosseini, M, Tavasoli, F, Suzuki, K, Umemoto, T, Boualaphanh, C, Lee, H, Hung, Y, Ramli, A, Aung, P, Ahmad, R, Wattoo, J, Bandonill, E, Romero, M, Brites, C, Hafeel, R, Lur, H, Cheaupun, K, Jongdee, S, Blanco, P, Bryant, R, Thi Lang, N, Hall, R and Fitzgerald, M. 2014. Diversity of global rice markets and the science required for consumer-targeted rice breeding. PlosOne. doi: 10.1371/journal.pone.0085106. [This thesis]

Calingacion, M, Boualaphanh, C, Daygon, V, Anacleto, R, Sackville Hamilton, R, Biais, B, Deborde, C, Maucourt, M, Moing, A, Mumm, R, de Vos, R, Erban, A, Kopka, J, Hansen, T, Laursen, K, Schjoerring, J, Hall, R and Fitzgerald, M. 2012. A genomics and multi-platform metabolomics approach to identify new traits of rice quality in traditional and improved varieties. Metabolomics 8:771-783. [This thesis]

Bajet, C, Kumar, A, **Calingacion, M**, and Narvacan, T. 2012. Toxicological assessment of pesticides used in the Pagsanjan-Lumban catchment to selected non target aquatic organisms in Laguna Lake, Philippines. Agricultural Water Management, 106: 42-49.

Boualaphanh, C, Daygon, V, **Calingacion, M**, Sanitchon, J, Jothiyangkoon, D, Mumm, R, Hall, R, and Fitzgerald, M. 2011. Use of new generation single nucleotide polymorphism genotyping for rapid development of near-isogenic lines in rice. Crop Science 51:2067-2073.

Boualaphanh, C, **Calingacion, M**, Cuevas, R, Jothiyangkoon, D, Sanitchon, J, and Fitzgerald, M. 2011. Yield and quality of traditional and improved Lao varieties of rice. Science Asia 37(2):89-97.

Carvalho, F, Villeneuve, J, Cattini, C, Tolosa, I, Bajet, C and **Calingacion, M**. 2010. PCBs in sediments and oysters of Manila Bay, Philippines. International Journal of Environmental Health Research 20 (4):259-269.

Kovach, M, **Calingacion, M**, Fitzgerald, M and McCouch, S. 2009. Where's the aroma coming from? The origin and evolution of fragrance in rice (Oryza sativa L). Proceedings of the National Academy of Sciences 106(34):14444-14449.

Carvalho, F, Villeneuve, J, Cattini, C, Bajet, C and **Calingacion, M**. 2009. Chlorinated hydrocarbons in sediments from Manila Bay, the Philippines. International Journal of Environmental Studies 67 (4):493-504.

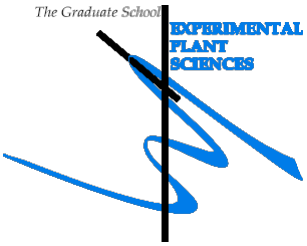
Carvalho, F, Villeneuve, J, Cattini, C, Tolosa, I, Bajet, C and **Calingacion, M**. 2009. Organic contaminants in the marine environment of Manila Bay, Philippines. Archives of Environmental Contamination and Toxicology 57:348-358.

Fitzgerald, M, Hamilton, R**, Calingacion, M** and Butardo, V. 2008. Is there a second fragrance gene in rice? Plant Biotechnology Journal 6:416-423.

Bajet, C and **Navarro, M**. 2002. Degradation, release and bioavailability of ^{14}C DDT and ^{14}C DDE sediment residues to oysters and mussels. Environmental Technology 23:1293-1302.

Education Statement of the Graduate School

Experimental Plant Sciences



Issued to: **Mariafe Calingacion**
Date: **9 January 2015**
Group: **Laboratory of Plant Physiology**
University: **Wageningen University & Research Centre**

1) Start-up phase	*date*
► **First presentation of your project**	
A taste of rice metabolomics	Jun 06, 2012
► **Writing or rewriting a project proposal**	
Empowering rice breeders to overcome constraints by identifying ways to select for the quality trait of taste	2012
► **Writing a review or book chapter**	
► **MSc courses**	
► **Laboratory use of isotopes**	
Subtotal Start-up Phase	*7.5 credits**

2) Scientific Exposure	*date*
► **EPS PhD student days**	
EPS PhD student day, University of Amsterdam	Nov 30, 2012
► **EPS theme symposia**	
EPS theme 3 symposium 'Metabolism and Adaptation', University of Amsterdam	Mar 22, 2013
EPS theme 1 symposium 'Developmental Biology of Plants', Leiden University	Jan 08, 2015
► **NWO Lunteren days and other National Platforms**	
► **Seminars (series), workshops and symposia**	
Managing your research data	Jun 12, 2012
Parenting in plants:maternal control of seed dormancy	Jun 12, 2012
Genomic-enabled prediction in plant breeding when modeling GxE interaction	Jun 14, 2012
Use of resurrection plants as models to understand how plants tolerate extreme water loss	Jun 26, 2012
Genetic modification for iron biofortification and drought tolerance in rice	Jun 29, 2012
Recognition of social identity in ants: pheromones and signature mixtures	Sept 20, 2012
ABA signalling networks in Arabidopsis	Nov 14, 2012
► **Seminar plus**	
► **International symposia and congresses**	
Norman E. Borlaug International Symposium, USA	Oct 13-15, 2010
Assoc of American Cereal Chemists Annual Symposium, USA	Oct 24-27, 2010
3rd International Rice Congress, Vietnam	Nov 08-12, 2010
International Network for Quality Rice Annual Symposium, Thailand	Aug 17-19, 2011
Norman E. Borlaug International Symposium, USA	Oct 12-14, 2011
Assoc of American Cereal Chemists Annual Symposium, USA	Oct 16-19, 2011
International Rice Genetics Symposium, Philippines	Nov 05-08, 2013
Metabolomics 2014, Japan	Jun 23-26, 2014
Australian Cereal Chemistry Conference, Australia	Aug 25-27, 2014
► **Presentations**	
Oral presentation at Assoc of American Cereal Chemists Annual Symposium, USA	Oct 27, 2010
Oral presentation at 26th Philippine Chemistry Congress, Philippines	April 14, 2011
Oral presentation at International Network for Quality Rice Symposium, Thailand	Aug 17, 2011
Poster presentation at 3rd International Rice Congress, Vietnam	Nov 12, 2011
Oral presentation at International Rice Genetics Symposium, Philippines	Nov 08, 2013
Poster presentation at Metabolomics 2014, Japan	Jun 25, 2014
Poster presentation at Australian Grain Science Symposium 2014, Australia	Aug 26, 2014
► **IAB interview**	
Meeting with a member of the International Advisory Board of EPS	Nov 14, 2012
► **Excursions**	
Subtotal Scientific Exposure	*19.2 credits**

3) In-Depth Studies	*date*
► **EPS courses or other PhD courses**	
Introduction to R: Data Manipulation and Statistical Analysis, IRRI	Jul 26-30,2010
Intermediate R course, IRRI	Sep 13-16, 2010
Systems biology: statistical analysis of ómics data	Dec 10-14, 2012
► **Journal club**	
Participation in a literature discussion group	2012-2013
► **Individual research training**	
► **EPS courses or other PhD courses**	
Subtotal In-Depth Studies	*5.0 credits**

4) Personal development	*date*
► **Skill training courses**	
Writing a research article for international publication, IRRI	Nov 28-Dec 02, 2011
Leadership Development, USA	Aug 20-24, 2012
EPS creative thinking and seminar workshop	Feb 02, 2013
► **Organisation of PhD students day, course or conference**	
► **Membership of Board, Committee or PhD council**	
Subtotal Personal Development	*3.3 credits**

TOTAL NUMBER OF CREDIT POINTS*	**35.0**
Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS which comprises of a minimum total of 30 ECTS	

** A credit represents a normative study load of 28 hours of study.*

Cover photo taken by Nghia Tran
Cover lay out by Lorizty Calingacion
Cover design by the author