TECHNISCHE UNIVERSITÄT MÜNCHEN

Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt

Lehrstuhl für Allgemeine Lebensmitteltechnologie

Impact of Cross-Breeding on the Phytic Acid Contents and the Metabolite Profiles of *Low Phytic Acid* Rice (*Oryza sativa* L.) Mutants

Chenguang Zhou

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften
(Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Rudi F. Vogel
Prüfer der Dissertation: 1. Prof. Dr. Karl-Heinz Engel
2. apl. Prof. Dr. Philippe Schmitt-Kopplin

Die Dissertation wurde am 15.11.2018 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 21.02.2019 angenommen.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

I

ABBREVIATIONS

CE	Capillary electrophoresis
CV	Coefficient of variance
DIMS	Direct injection mass spectrometry
DNA	Deoxyribonucleic acid
DW	Dry weight
EI	Electron impact ionization
ESI	Electrospray ionization
eV	Electron Volt
FAME	Fatty acid methyl ester
FAO	Food and Agriculture Organization of the United Nations
FDR	False discovery rate
FFA	Free fatty acid
FT-ICR	Fourier transform ion cyclotron resonance
GABA	γ-Aminobutyric acid
GC	Gas chromatography
GM	Genetically modified
HCA	Hierarchical clustering analysis
HILIC	Hydrophilic interaction liquid chromatography
HPIC	High pressure ion chromatography
IAEA	International Atomic Energy Agency
ID	Internal diameter
InsP	Inositol phosphate
IS	Internal standard
LC	Liquid chromatography
LOD	Limit of detection
LOQ	Limit of quantitation
Lpa	*Low phytic acid*

MRP	Multi-drug resistance-associated protein
MS	Mass spectrometry
MSI	Metabolomics Standards Initiative
MSTFA	*N*-Methyl-*N*-(trimethylsilyl) trifluoroacetamide
MTBE	Methyl *tert*-butyl ether
NMR	Nuclear magnetic resonance
OPLS-DA	Orthogonal partial least squares-discriminate analysis
PA	Phytic acid
PC	Principal component
PCA	Principal component analysis
PGK	Phosphoglycerate kinase
PLS-DA	Partial least squares-discriminate analysis
RSD	Relative standard deviation
RT	Retention time
SPE	Solid phase extraction
TCA	Tricarboxylic acid cycle
TMSIM	*N*-Trimethylsilylimidazole

SUMMARY

Phytic acid (PA) is the major storage form of phosphorus in cereal grains. However, it is considered as antinutrient in food and feed. Accordingly, several *low phytic acid* (*lpa*) rice (*Oryza sativa* L.) mutants have been generated through chemical or physical mutagenesis. *Lpa* mutant crops often exhibit inferior agronomic performance, e.g., low germination rate, decreased field emergence and reduced grain weight, and cross-breeding of *lpa* mutants with commercial cultivars is being applied to improve the agronomic performance of *lpa* mutants. However, information regarding the impact of such crossing steps on the PA contents and the metabolite profiles of the resulting progeny rice seeds was missing.

Therefore, in the first part of the study, three previously developed *lpa* rice mutants (*Os-lpa*-XQZ-1, *Os-lpa*-XS110-2 and *Os-lpa*-MH86-1) differing in mutation type were crossed with three commercial rice cultivars, and the PA contents in the resulting progenies were determined via targeted analysis for various generations (F_4 to F_8) grown at different locations. The results showed that the PA contents of the *lpa* progenies were differently expressed for the investigated mutation-types and were dependent on environment and/or PA contents of the crossing parents. Nevertheless, for all three mutants, the homozygous *lpa* mutant progenies always displayed significantly lower PA contents than the original wild-types subjected to the mutation. This demonstrated that the *lpa* trait, i.e. the significantly reduced content of phytic acid, remained nearly unaffected by the crossing step and was consistently expressed in *lpa* progenies over generations, independent from the environmental influence. This result fulfilled a first major prerequisite for the implementation of cross and selection breeding with commercial cultivars as part of the procedure to generate *lpa* rice seeds.

The *lpa* rice mutant *Os-lpa*-MH86-1 was used as example to follow metabolic changes in *lpa* mutants induced by disruption of the putative sulfate transporter *OsSULTR3;3* gene. A non-targeted metabolite profiling approach enabled the analyses of a broad spectrum of low molecular weight rice constituents ranging from lipophilic (e.g. fatty acid methyl esters, free fatty acids, fatty alcohols and phytosterols) to polar (e.g. sugars, sugar alcohols, acids, amino acids and amines) metabolites. Multivariate and univariate analyses revealed that the mutation of *OsSULTR3;3* resulted not only in a

pronounced reduction of the PA content, but also in increased contents of constituents from various classes, such as sugars, sugar alcohols, amino acids, phytosterols and biogenic amines, among them some nutritionally relevant compounds. In addition, statistical assessments of the metabolite profiling data demonstrated that this mutation-induced metabolite signature in *lpa* mutants was nearly unaffected by cross-breeding of *Os-lpa*-MH86-1 with the commercial cultivar JH99, and was consistently expressed over several generations at the same field trial. This indicates that even for complex metabolic changes resulting from a mutation, cross-breeding can be employed as a tool to generate *lpa* progeny rice seeds stably exhibiting the mutation-induced traits.

To extend the scope of the applicability of cross-breeding as a tool to generate *lpa* progeny rice seeds exhibiting the mutation-specific metabolite signature, the *Os-lpa*-MH86-1 mutant was crossed with another commercial rice cultivar JH218, showing distinct metabolic differences compared to JH99. By following the metabolite profiles of the F_8 progenies of the two crosses *Os-lpa*-MH86-1 × JH99 and *Os-lpa*-MH86-1 × JH218 over three independent field trials, the *OsSULTR3;3* mutation-induced metabolite signature was shown to be consistently expressed in homozygous *lpa* mutant progenies, independent from the crossing parent cultivars. The superimposition of the metabolite profiles of the homozygous *lpa* mutant progenies by the lipid profiles of the crossing parents did not hamper the mutation-induced metabolite signature determined by the polar constituents. The metabolite profiles of the homozygous *lpa* progenies were also shown to be strongly impacted by the environment. Tropical and subtropical growing locations resulted in clearly different metabolite patterns. However, for each individual field trial, the mutation-specific metabolite signature was maintained upon cross-breeding.

The elaborated data are valuable and promising from a breeder's point of view. They demonstrate that cross-breeding of *lpa* rice mutants with commercial cultivars does not compromise the intended phytic acid reduction as well as the mutation-specific metabolite signature of homozygous *lpa* mutant progenies. Therefore, cross-breeding is a useful strategy to generate *lpa* progeny rice seeds stably exhibiting the mutation-induced metabolic traits.

ZUSAMMENFASSUNG

Phytinsäure (PA) ist die Hauptspeicherform von Phosphor in Getreide. Sie gilt jedoch als antinutritive Substanz in Lebens- und Futtermitteln. Folglich wurden bereits zahlreiche Reismutanten (*Oryza sativa* L.) mit erniedrigten Phytinsäuregehalten (*low phytic acid, lpa*) mit Hilfe chemischer oder physikalischer Mutagenese erzeugt. *Lpa* Mutanten weisen häufig verminderte agronomische Leistungen auf, z. B. geringe Keimrate, verminderte Emergenz im Feldversuch und reduziertes Getreidegewicht. *Lpa* Mutanten werden daher mit kommerziellen Sorten gekreuzt, um ihre agronomischen Eigenschaften zu verbessern. Die Auswirkungen solcher Kreuzungsschritte auf die PA Gehalte und die Metabolitenprofile der resultierenden Nachkommen sind bisher allerdings nicht untersucht worden.

Daher wurden im ersten Teil der Arbeit drei zuvor entwickelte *lpa* Reismutanten (*Os-lpa*-XQZ-1, *Os-lpa*-XS110-2 und *Os-lpa*-MH86-1), denen unterschiedliche Mutationstypen zugrunde lagen, mit drei kommerziellen Reissorten gekreuzt und die PA Gehalte der resultierenden Nachkommen mittels zielgerichteter Analyse über mehrere Generationen (F_4 bis F_8) und an verschiedenen Anbauorten untersucht. Die Ergebnisse zeigten, dass sich die PA Gehalte in den *lpa* Nachkommen der untersuchten Mutationstypen unterschieden und von der Umwelt und/oder den PA Gehalten der Kreuzungspartner beeinflusst wurden. Trotzdem zeigten die homozygoten *lpa* Mutanten aller drei Mutationstypen stets signifikant erniedrigte PA Gehalte im Vergleich zu den ursprünglichen Wildtypen, aus denen die Mutanten generiert worden waren. Dies zeigte, dass der *lpa* Phänotyp, d.h. der signifikant erniedrigte PA Gehalt, kaum vom Kreuzungsschritt beeinflusst und konsistent sowie unabhängig von Umwelteinflüssen in den *lpa* Nachkommen über Generationen ausgebildet wurde. Somit ist eine wesentliche Voraussetzung für den Einsatz der Kreuzungs- und Selektionszüchtung mit kommerziellen Sorten zur Entwicklung von *lpa* Reis erfüllt.

Die *lpa* Reismutante *Os-lpa*-MH86-1 wurde exemplarisch verwendet, um die metabolischen Veränderungen in *lpa* Mutanten mit Störungen im Sulfattransporter *OsSULTR3;3* Gen zu verfolgen. Ein nicht zielgerichteter Metabolite Profiling Ansatz ermöglichte die Analyse eines breiten Spektrums an niedermolekularen Verbindungen

in Reis, welches von lipophilen (z. B. Fettsäuremethylester, freien Fettsäuren, Fettalkohole und Phytosterole) bis hin zu polaren Verbindungen (z. B. Zucker, Zuckeralkohole, Säuren, Aminosäuren und Amine) reichte. Multi- und univariate Analysen zeigten, dass die Mutation von *OsSULTR3;3* nicht nur zu einer deutlichen Reduktion des PA Gehalts führte, sondern auch zu erhöhen Gehalten an teilweise ernährungsphysiologisch relevanten Verbindungen aus verschiedenen Klassen, wie beispielsweise Zuckern, Zuckeralkoholen, Aminosäuren, Phytosterolen und biogenen Aminen. Darüber hinaus ergaben statistische Auswertungen der Daten, dass dieses mutationsbedingte Metabolitenprofil in *Ipa* Mutanten durch die Kreuzung von *Os-Ipa*-MH86-1 mit der kommerziellen Sorte JH99 im selben Feldversuch über mehrere Generationen nahezu unverändert blieb. Dies verdeutlicht, dass Kreuzungszüchtung sogar im Fall komplexer metabolischer Veränderungen eingesetzt werden kann, um *Ipa* Reis mit stabilen mutationsbedingten Eigenschaften zu erzeugen.

Um den Anwendungsbereich von Kreuzungszüchtungen zur Erzeugung von Nachkommen aus *Ipa* Reismutanten mit mutationsspezifischem Metabolitenprofil zu erweitern, wurde die Mutante *Os-Ipa*-MH86-1 mit einer anderen kommerziellen Reissorte (JH218) gekreuzt, die im Vergleich zu JH99 deutliche metabolische Unterschiede aufwies. Das mutationsbedingte Metabolitenprofil von *OsSULTR3;3* war in allen homozygoten *Ipa* F_8 Nachkommen der beiden Kreuzungen *Os-Ipa*-MH86 1 × JH99 und *Os-Ipa*-MH86 1 × JH218 in allen drei Feldversuchen unabhängig vom Kreuzungspartner nachweisbar. Die Überlagerung der Metabolitenprofile der homozygoten *Ipa* Mutanten durch die Lipidprofile der Kreuzungspartner führte nicht zu einer Beeinträchtigung der mutationsspezifischen, metabolischen Signatur, die von polaren Inhaltsstoffen bestimmt wurde. Die Metabolitenprofile der homozygoten *Ipa* Nachkommen wurden außerdem stark durch Umweltfaktoren beeinflusst. Tropische und subtropische Anbaugebiete führten zu deutlich unterschiedlichen Metabolitenprofilen, dennoch blieb für jeden individuellen Feldversuch die mutationsspezifische Signatur nach der Kreuzung erhalten.

Die erarbeiteten Daten sind aus der Sicht von Züchtern wertvoll und vielversprechend. Sie zeigen, dass die Kreuzung von *Ipa* Reismutanten mit kommerziellen Sorten die beabsichtigte Reduzierung der Phytinsäure sowie die mutationsspezifischen

Metabolitenprofile in homozygoten *lpa* Nachkommen nicht beeinträchtigen. Daher stellt Kreuzungs- und Selektionszüchtung eine nützliche Strategie dar, um *lpa* Reis mit konsistenten, mutationsbedingten metabolischen Merkmalen zu erzeugen.

1 INTRODUCTION AND OBJECTIVES

Phosphorus (P) is an important element for plant growth and development. In cereal grains and legume seeds, up to 85% of P is stored in the form of phytic acid (PA, *myo*-inositol-1,2,3,4,5,6-hexa*kis*phosphate or InsP$_6$) (Raboy, 2003). However, phytic acid is considered as an antinutrient in food and feed, since it may form indigestible salts of cations such as Zn^{2+} and Fe^{2+}, and thus lead to a reduced bioavailability of these minerals for humans and monogastric animals (Kumar *et al.*, 2010; Humer *et al.*, 2015). In addition, undigested phytate from the manure of monogastric animals contributes to water eutrophication in the environment (Abdel-Megeed *et al.*, 2015). In order to minimize these disadvantageous effects, during the past years various *low phytic acid* (*lpa*) crops comprising maize, barley, rice, wheat and soybean have been obtained through genetic engineering and chemical/physical mutagenesis (Raboy, 2007).

Rice (*Oryza sativa* L.) is one of the major cereals serving as staple food for a large part of the world's population (FAO, 2017). Accordingly, there have been various attempts to produce *lpa* rice mutants. For example, five non-lethal *lpa* rice mutant lines have been developed with PA contents 35 to 65% lower than those of the corresponding wild-types (Liu *et al.*, 2007). *Lpa* mutant crops often exhibit inferior agronomic traits, e.g., low germination rate, decreased field emergence and reduced grain weight, and cross and selection breeding of *lpa* mutants with commercial cultivars has been applied to improve the agronomic performance of *lpa* mutants (Zhao *et al.*, 2008a). However, the impact of such crossing steps on the PA contents in the resulting progenies has not been investigated. A limited amount of information regarding *lpa* crossbred lines is only available for *lpa* soybean mutants (Spear *et al.*, 2007; Averitt *et al.*, 2017; Wiggins *et al.*, 2018).

Therefore, three previously developed *lpa* rice mutants differing in the type of mutation were crossed with three commercial rice cultivars, and the PA contents in the resulting progenies were determined for various generations grown at different locations. The first objectives of this study were: (i) to investigate the impact of cross-breeding of *lpa* rice mutants with commercial cultivars on the PA contents of the resulting progenies, and (ii) to assess the stability of the *lpa* trait in crossbred progenies over several generations with the consideration of environmental influence.

Induced mutations in *lpa* mutants have been shown to result not only in decreased contents of phytic acid, but also in changed levels of other metabolites. For example, a GC/MS-based metabolite profiling study revealed consistent changes of the contents of *myo*-inositol, raffinose, galactose and galactinol in *lpa* rice mutants compared to the wild-type cultivars (Frank *et al.*, 2007). Therefore, the *lpa* rice mutant *Os-lpa*-MH86-1 was employed as an example to investigate the metabolic changes induced by the mutation effect. In addition, the crossbred progeny rice seeds resulting from the cross of *Os-lpa*-MH86-1 with a commercial cultivar were subjected to a GC/MS-based metabolite profiling approach. The objectives of the second part of the present study were: (i) to observe the consistent metabolic changes in the *Os-lpa*-MH86-1 *lpa* mutants compared to the respective wild-types by following their compositions over several independent field trials; (ii) to investigate the impact of cross-breeding of *Os-lpa*-MH86-1 with a commercial cultivar on the metabolite profiles of the resulting *lpa* mutant progenies, and (iii) to assess the stability of the mutation-specific metabolite signature in the *lpa* progenies from several generations harvested at the same field trial.

It has been demonstrated that metabolite profiles of cereal grains, e.g. maize and barley, were influenced by both genetics and growing environment (Frank *et al.*, 2012b; Lanzinger *et al.*, 2015; Wenzel *et al.*, 2015). Therefore, another commercial wild-type cultivar expected to exhibit a distinct metabolite profile was selected as additional crossing parent. In addition, progenies resulting from the two crosses were grown at three independent field trials, two of them in the subtropical and one in the tropical region. Based on the application of a GC/MS-metabolite profiling approach followed by multivariate and univariate statistical analyses, the objectives of this third part of the study were: (i) to investigate the impact of the crossing parent on the metabolite profiles of the resulting progenies, and (ii) to assess the stability of the metabolite signature of the homozygous *lpa* progenies in the light of environmental influence.

This thesis is structured as follows:
- Following these introductory remarks outlining the objectives of the study (Chapter 1), the background of the thesis is presented in Chapter 2.
- Materials and methods, results of the studies and their specific discussions are presented in the three publications included in Chapter 3.

− An overall discussion highlighting key overarching aspects and the significance of the results in the light of existing literature knowledge is presented in Chapter 4.

2 BACKGROUND

2.1 *Low phytic acid* crops

2.1.1 Natural occurrence of phytic acid

Phytic acid (*myo*-inositol-1,2,3,4,5,6-hexa*kis*phosphate, $InsP_6$, PA) is the major storage form of phosphorus in seeds of mature plants. Approximately, between 65 and 85 % of total phosphorus in mature seeds is found in this compound (Raboy, 2003). The chemical structure of phytic acid is shown in Figure 1.
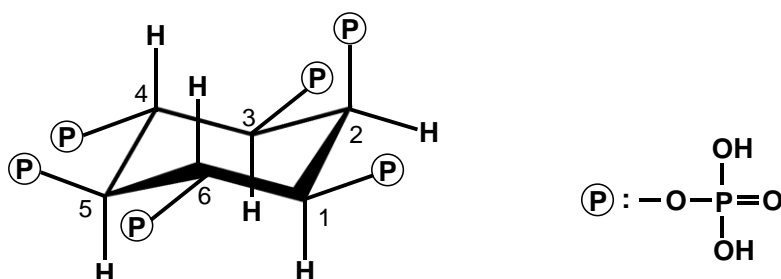
Figure 1. Chemical structure of phytic acid. Numbers 1 to 6 refer to the carbon atoms in the *myo*-inositol backbone.

Phytic acid typically occurs as mixed salts of the mono- and divalent cations K^+, Mg^{2+}, Ca^{2+}, Fe^{2+} and Zn^{2+}, which is referred to as phytate. In mature seeds, phytate is often deposited as globular inclusions (globoids) and is stored in protein storage vacuoles (Lott *et al.*, 2000; Raboy, 2003). Generally, phytic acid contents increase significantly during seed development and reach the highest concentration in mature seeds (Reddy, 2002). The distributions of phytic acid in crop seeds can be very different depending on species. For rice and wheat grains, more than 80 % of phytic acid was found in the aleurone layer, and only a small amount (approximately 10%) is stored in the embryo. On the contrary, almost 90% of the phytic acid is accumulated in the germ of maize grains (O'Dell *et al.*, 1972). For legume seeds, such as common beans, most of the phytic acid (94 – 98 %) is located in the cotyledons (Ariza-Nieto *et al.*, 2007). During seed germination, phytic acid is degraded by phytases into phosphate and *myo*-inositol (Ins) which are necessary for seedling growth (Raboy, 2003).

2.1.2 Biosynthesis of phytic acid

The biosynthetic pathway leading to phytic acid is shown in Figure 2. The biosynthesis of phytic acid consists of three major steps: (i) Synthesis of *myo*-inositol (Ins) and *myo*-inositol 3-phosphate, i.e. Ins(3)P_1; (ii) conversion of Ins and/or Ins(3)P_1 to InsP_6 (phytic acid) and (iii) transportation and storage of phytic acid (Raboy, 2009; Sparvoli *et al.*, 2015).

Synthesis of Ins(3)P_1 and myo-inositol

In the first step, glucose-6-phosphate is converted to Ins(3)P_1 by *myo*-inositol 3-phosphate synthase (*MIPS*). Then the conversion of Ins(3)P_1 to free *myo*-inositol is catalyzed by *myo*-inositol-phosphate monophosphatase (*IMP*). This reaction can be reversed by catalysis of the *myo*-inositol kinase (*MIK*) (Raboy, 2009; Sparvoli *et al.*, 2015).

Conversion of Ins and/or Ins(3)P_1 to InsP_6

Ins acts as the initial substrate in the biosynthesis of phytic acid via two different routes: the lipid-dependent and the lipid-independent pathway (Figure 2). In the lipid-dependent pathway, Ins is first converted to the lipid phosphatidylinositol (PtdIns) with Ins as the head-group. PtdIns is further phosphorylated via phosphatidylinositol kinases (*PtdInsK*) to generate PtdIns(4,5)P_2 (Raboy, 2009). The hydrolysis of PtdIns(4,5)P_2 is catalyzed by the PtdInsP-specific phospholipase C and results in Ins(1,4,5)P_3 and diacylglycerol (Odom *et al.*, 2000). Ins(1,4,5)P_3 is sequentially phosphorylated to generate phytic acid, and these reactions involve three types of Ins polyphosphate kinases: inositol 1,4,5-trisphosphate kinase (*IPK2*), inositol 1,3,4-trisphosphate 5/6-kinase (*ITPK*) and inositol 1,3,4,5,6 penta*kis*phosphate 2-kinase (*IPK1*) (Sparvoli *et al.*, 2015).

The lipid-independent pathway differs from the lipid-dependent pathway mainly in the step of generating InsP_3 (Sparvoli *et al.*, 2015). The lipid-independent route begins with the conversion of Ins to Ins(3)P_1 catalyzed by *myo*-inositol kinase (*MIK*). Ins(3)P_1 is expected to be converted to Ins(3,4)P_2 and Ins(3,4,6)P_3 by Ins phosphate kinases (Raboy, 2009). The subsequent steps of phosphorylation from Ins(3,4,6)P_3 to phytic acid are similar to those in the lipid-dependent route.
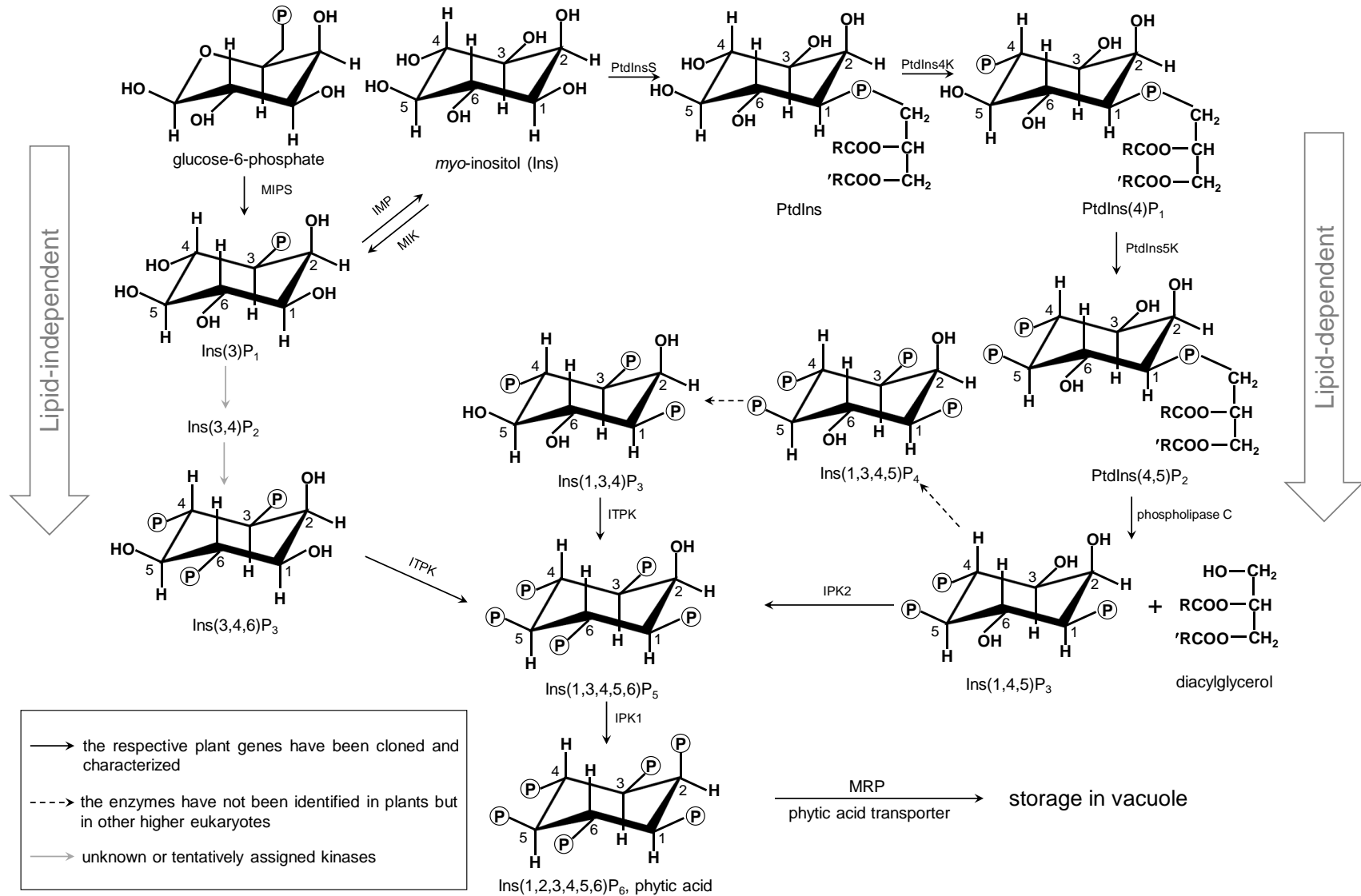
Figure 2. (Caption on the following page)

Figure 2. Biosynthetic pathway of phytic acid in plant metabolism. *MIPS*, *myo*-inositol 3-phosphate synthase; *IMP*, *myo*-inositol-phosphate monophosphatase; *MIK*, *myo*-inositol kinase; *PtdInsS*, phosphatidylinositol synthase; *PtdIns*, phosphatidylinositol; *PtdIns4K*, phosphatidylinositol 4-kinase; *PtdIns5K*, 1-phosphatidylinositol-4-phosphate 5-kinase; *PtdInsP*, phosphatidylinositol phosphate(s); *ITPK*, inositol 1,3,4-trisphosphate 5/6-kinase; *IPK2*, inositol 1,4,5-trisphosphate kinase; *IPK1*, inositol 1,3,4,5,6 penta*kis*phosphate 2-kinase; *MRP*, multidrug-resistance-associated protein ATP-binding cassette.

Transportation and storage of phytic acid

After synthesis steps, phytic acid is transported by the multi-drug resistance associated protein (*MRP*) from the cytosol into vacuoles and stored as globoids (Sparvoli *et al.*, 2014).

2.1.3 Generation of *low phytic acid* crops

Phytic acid is considered as an antinutrient in food and feed, since it can form indigestible salts of cations such as Ca^{2+}, Zn^{2+} and Fe^{2+}, and thus lead to reduced bioavailability of these minerals for humans and monogastric animals (Kumar *et al.*, 2010; Humer *et al.*, 2015). In addition, undigested phytic acid from the manure of monogastric animals contributes to water eutrophication in the environment (Abdel-Megeed *et al.*, 2015). Preliminary studies indicated that consumption of *lpa* grains and beans improves the utilization of mineral nutrients for humans (Mendoza *et al.*, 1998; Petry *et al.*, 2013) and livestock (Veum *et al.*, 2002; Veum *et al.*, 2016). Therefore, it is of importance to develop *low phytic acid* (*lpa*) crops with reduced phytic acid levels in seeds.

In the past years, genetic engineering techniques have been applied to produce various *lpa* crops including rice (Feng *et al.*, 2004; Kuwano *et al.*, 2006; Ali *et al.*, 2013), soybean (Nunes *et al.*, 2006; Punjabi *et al.*, 2018), maize (Shi *et al.*, 2007) and wheat (Bhati *et al.*, 2016). However, the development of genetically modified (GM) crops has raised concerns regarding food safety, environmental effects and socio-economic issues (Singh *et al.*, 2006). Alternatively, mutation breeding, which is considered as a conventional breeding technique, has been applied to develop *lpa* mutant crops. During the mutation breeding process, seeds are subjected to physical mutagens (e.g., X-rays, γ-rays and ion beam) and/or chemical mutagens (e.g. sodium azide and

hydroxylamine), followed by the steps of mutation screening and mutant confirmation (Oladosu *et al.*, 2016).

The first documented mutant was barley (Stadler, 1928) and up to now, more than 3200 crop mutants have been officially released worldwide (FAO/IAEA, 2018). γ-Irradiation and chemical mutagenesis have been employed to generate *lpa* cereal grains and legume seeds, such as barley (Rasmussen *et al.*, 1998; Dorsch *et al.*, 2003), maize (Raboy *et al.*, 2000; Shi *et al.*, 2003; Cerino Badone *et al.*, 2012), rice (Larson *et al.*, 2000; Liu *et al.*, 2007; Kim *et al.*, 2008), soybean (Wilcox *et al.*, 2000; Hitz *et al.*, 2002; Yuan *et al.*, 2007), wheat (Guttieri *et al.*, 2004) and common bean (Campion *et al.*, 2009). These *lpa* mutant seeds exhibited significantly decreased levels of phytic acid compared to the corresponding wild-types, ranging from -30% to -90%. Correspondingly, remarkably increased contents of inorganic phosphorus were observed in *lpa* mutants (Sparvoli *et al.*, 2015).

Although a few *lpa* barley mutants showed comparable agronomic performance compared to the corresponding wild-types (Raboy *et al.*, 2015), most *lpa* mutant crops exhibited inferior agronomic traits, e.g. low seed germination rate (Pilu *et al.*, 2005), decreased field emergence (Meis *et al.*, 2003) and reduced grain weight (Guttieri *et al.*, 2004; Zhao *et al.*, 2008a), which limited the practical use of *lpa* cereal grains and legume seeds to breeders. To date, the challenge of breeding high-yielding *lpa* mutant crops still remains.

2.1.4 *Lpa* mutation types and target genes

Based on the affected steps of the biosynthetic pathway leading to phytic acid (Figure 2), *lpa* mutations are divided into three types as shown in Table 1 (Raboy, 2009; Ye *et al.*, 2013; Sparvoli *et al.*, 2015). For the Type 1 *lpa* mutants, mutations occur in the genes that are involved in the initial stage of the biosynthetic pathway, i.e., steps from glucose-6-phosphate to *myo*-inositol (Ins) and $Ins(3)P_1$. This type of mutation was reported for the soybean mutants LR33 and *Gm-lpa*-TW-1 (Hitz *et al.*, 2002; Yuan *et al.*, 2007), as well as for the rice mutant *Os-lpa*-XS110-1 (Zhao *et al.*, 2013), which exhibited decreased contents of phytic acid and a molar equivalent increase in the levels of inorganic phosphorus.

8

The mutations of the target genes in Type 2 *lpa* mutants can result in the perturbation of biosynthesis steps from $Ins(3)P_1$ to phytic acid (Figure 2), and the decreases of phytic acid content are often accompanied by the accumulations of lower inositol phosphates ($InsP_3$, $InsP_4$ and $InsP_5$). For example, the *lpa* soybean mutant *Gm-lpa-ZC-2* with the mutation of the *IPK1* gene exhibited significantly increased levels of $InsP_3$, $InsP_4$ and $InsP_5$, compared to the corresponding wild-type Zhechun no.3 (Yuan *et al.*, 2007; Frank *et al.*, 2009b). Similar results regarding the accumulation of lower inositol phosphates were observed in the *ITPK* mutant maize seeds (Shi *et al.*, 2003) and the *ITPK* mutant rice seeds (Kim *et al.*, 2014).

The third type of *lpa* mutants is classified as the mutations affecting the transportation and storage of phytic acid to vacuoles. This type of *lpa* mutants, including the previously reported rice mutant *Os-lpa*-XS110-2 (Frank *et al.*, 2007; Liu *et al.*, 2007), the soybean mutant *CX1834* (Gao *et al.*, 2008) and the common bean mutant *lpa1* (Panzeri *et al.*, 2011), are characterized by a pronounced reduction of the phytic acid content, as well as increased levels of *myo*-inositol and inorganic phosphorus. In addition, no accumulation of lower inositol phosphates was detected for this type of *lpa* mutants.

Table 1. Classification of *lpa* mutations, the affected steps in the biosynthetic pathway to phytic acid, and the target genes in *lpa* mutants.

Mutation Type	Perturbation in the pathway	Target gene
Type 1	Synthesis of $Ins(3)P_1$ and Ins	*MIPS, IMP, MIK*
Type 2	Conversion of Ins and $Ins(3)P_1$ to $InsP_6$	*2PGK, IPK2, ITPK, IPK1*
Type 3	Transportation and storage of $InsP_6$	*MRP*

2.2 Metabolomics

2.2.1 Definition

The term 'metabolome' was introduced to the scientific community in 1998. It represents the set of all low-molecular-weight (typically < 1000 Dalton) metabolites in a biological system or biofluid (Oliver *et al.*, 1998). The resulting discipline, which was referred to as 'metabonomics' (Nicholson *et al.*, 1999), is therefore aiming to profile all small molecules in biological samples. Later, the term 'metabolomics' was introduced (Fiehn, 2002) and is most commonly used at present.

Metabolite fingerprinting and metabolite profiling are two major approaches to metabolomics research (Lindon *et al.*, 2007). Metabolite fingerprinting is often conducted via nuclear magnetic resonance (NMR) spectroscopy to obtain the unassigned NMR spectra of the investigated samples and to compare these 'fingerprints' between different groups. This approach ignores the identification of individual metabolites, and it allows high throughput measurements for rapid screening and classification of the samples (Nicholson *et al.*, 1999; Krishnan *et al.*, 2005). In contrast, metabolite profiling focuses on identification and quantitation of a broad spectrum of individual metabolites in biological samples. Although this approach is time-consuming and more challenging, the obtained profiles are expected to provide more detailed information on the biological system or biofluid (Fiehn, 2002).

2.2.2 Analytical techniques for plant metabolomics

Plant metabolomes include a broad spectrum of constituents exhibiting diverse physiochemical properties. It is estimated that there might be up to 200,000 metabolites occurring in the 'plant kingdom' (Fiehn, 2002; Goodacre *et al.*, 2004) and 5000 – 10,000 in specific species (Moritz *et al.*, 2008). In addition, the concentrations of these constituents can vary over a wide range of magnitude (Wishart, 2011). Owing to this complexity, no single analytical technique is capable of covering the overall metabolome, and therefore several analytical technologies are used to complement each other. The two most commonly used technologies are nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) coupled to separation techniques, such as gas chromatography (GC), liquid chromatography (LC) and

capillary electrophoresis (CE) (Zhang *et al.*, 2012; Larive *et al.*, 2015; Jorge *et al.*, 2016).

2.2.2.1 MS-based plant metabolomics

GC/MS. One of the first reports on plant metabolomics focused on potato tubers (Roessner *et al.*, 2000). Another pioneering study is the metabolite profiling of *Arabidopsis thaliana* leaves (Fiehn *et al.*, 2000). Both studies were conducted using gas chromatography coupled to mass spectrometry (GC/MS). To date, GC/MS is still one of the most widely employed analytical technique in plant metabolomics studies (Jorge *et al.*, 2015). For chromatographic separation, capillary columns with stationary phases of 100 % dimethylpolysiloxane or (5%-Phenyl)-methylpolysiloxane are often employed. As the most common hard ionization method, electron impact (EI) ionization is widely used in GC/MS to ionize the molecules before mass-spectrometric detection and analysis (Kopka *et al.*, 2004; Lisec *et al.*, 2006).

After separation on the GC column, the molecules flow into a mass spectrometer via a transfer line and are bombarded by a fixed high-energy electron beam set to 70 eV (electron Volts) under a high vacuum environment. EI ionization generates highly reproducible fragmentation patterns and has minimal variability between different instruments, regardless of the manufacturer. This allows the construction of commercial GC/MS libraries containing tens of thousands of mass-spectral data, which enable the identifications of plant metabolites and the integration of metabolomics data worldwide. Single Quadruple (Q), Triple Quadruple (QQQ), Time-of-flight (TOF) and Ion Trap are the common mass analyzers in GC/MS. Generally, TOF mass analyzers exhibit higher mass accuracy and faster acquisition rate than Q, QQQ and Ion Trap (Kopka *et al.*, 2004; Dunn *et al.*, 2005; Lisec *et al.*, 2006; Glauser *et al.*, 2013; Jorge *et al.*, 2015). An overview on the common mass analyzers is shown in Table 2.

Despite the good reproducibility of GC/MS, a major limitation is the prerequisite of volatility and thermal stability of the analytes. Therefore, derivatization steps are very essential during the process of sample preparation (Kopka, 2006). Derivatization types typically include silylation, alkylation and acylation, among which the silylation method is predominantly employed for metabolite profiling of plant extracts. *N*-Methyl-*N*-(trimethylsilyl) trifluoroacetamide (MSTFA) and methoxyamine hydrochloride are the most commonly used derivatization reagents (Lisec *et al.*, 2006).

Table 2. Overview on the common mass analyzers for GC/MS.

Mass analyzer	Mass accuracy (ppm) [a]	Mass resolution ($\times 10^3$) [b]	Scan rate (Hz)	Reference
Quadruple	Low	3-5	2-10	(Douglas, 2009)
Ion Trap	50-500	4-20	2-10	(March, 2009)
Triple Quadruple	5-50	> 7.5	10-20	(Douglas, 2009)
Time-of-flight	1-5	10-60	10-50	(Guilhaus *et al.*, 2000)
Quadruple-time-of-flight	1-5	> 60	>100	(Douglas *et al.*, 2005)

[a] Mass accuracy = (mass error) / (true mass) $\times 10^6$; mass error = (true mass) − (measured mass). [b] Mass resolution is defined as the ability to distinguish peaks with similar m/z ratios in a mass spectrum. Mass resolution = $m/\Delta m_{50}$, in which m refers to mass and Δm_{50} refers to the peak width at half of the peak height (Balogh, 2004).

LC/MS. Lipid chromatography coupled to mass spectrometry (LC/MS) is an important complementary technique to GC/MS. It does not necessitate the derivatization step during sample preparation and is capable of analyzing primary and secondary plant metabolites of interest which are high-molecular-weight and thermolabile (Allwood *et al.*, 2010). In most cases, reversed-phase (RP) octyl (C8) or octadecyl (C18) columns are used in LC/MS analysis for plant metabolomics studies. However, highly polar metabolites in plant extracts are not well retained on classical RP columns and elute along with the void volume (Dettmer *et al.*, 2007; Allwood *et al.*, 2010). As an alternative, hydrophilic interaction liquid chromatography (HILIC) is employed (Tang *et al.*, 2016). Different from hard ionization (i.e. EI) in GC/MS, a typical soft ionization method used in LC/MS is electrospray ionization (ESI) (Fenn *et al.*, 1989). This type of ionization can occur via protonation (ESI+) or deprotonation (ESI-) and leads to single and/or multiple ions (Kopka *et al.*, 2004).

Although the application of LC/MS technique in plant metabolomics exhibits several advantages, a lower reproducibility is often observed compared to GC/MS (t'Kindt *et al.*, 2009). Another disadvantage is that the application of the ESI method in LC/MS could cause ion suppression effects, i.e. mutual interference between compounds, leading to reduced detection capability and repeatability, as well as to the non-

detection of analytes or the under/over-estimation of their real concentrations (Antignac *et al.*, 2005; Furey *et al.*, 2013).

CE/MS. For capillary electrophoresis coupled to mass spectrometry (CE/MS) analysis, metabolites are separated based on their charge-to-size ratios and electrophoretic mobility, followed by the detection and the quantification in the mass spectrometer according to the specific mass-to-charge (*m/z*) ratio (Ramautar *et al.*, 2009; Jorge *et al.*, 2015). CE/MS shows great advantages in analyzing polar and/or charged metabolites including amino acids, organic acids, sugar phosphates and nucleotides (Ramautar *et al.*, 2017). Despite much less reports of CE/MS-based metabolomics studies compared to GC/MS and LC/MS, the advantages of this technique, such as simplicity of sample pre-treatment and low-consumption of organic solvents, will facilitate its increasing role in plant metabolomics research (Ramautar *et al.*, 2018).

FT-ICR/MS. Another MS-based technique is Fourier transform ion cyclotron resonance mass spectroscopy (FT-ICR/MS). The increasing popularity of this technique in metabolomics studies is attributed mainly to its ultra-high resolution (> 1,000,000), and the currently highest available mass accuracy (< 0.2 ppm) (Ghaste *et al.*, 2016). These outstanding advantages also allow the reliable assignment of molecular formulas to the detected ions for the characterization of unknown metabolites (Brown *et al.*, 2005). However, as a direct injection (DI) MS technique, the approach lacks prior chromatographic separation of sample constituents, making it difficult to distinguish isomeric molecules. The principles of FT-ICR/MS and its potential in metabolomics have been thoroughly summarized in a review (Moritz *et al.*, 2013). Unfortunately, the high costs of FT-ICR/MS instruments so far limit a wider application of this promising technique in metabolomics research (Dettmer *et al.*, 2007).

2.2.2.2 NMR-based plant metabolomics

Nuclear magnetic resonance (NMR) spectroscopy has been well established in plant metabolomics studies over the past years. Despite the lower sensitivity compared to MS, ^{1}H-NMR spectroscopy enables the identification and quantification of metabolites in samples via a non-destructive way (Rolin *et al.*, 2013). This technique requires few sample preparation steps (none at all in certain cases), and a high throughput analysis is achievable owing to the short acquisition time per sample (Kim *et al.*, 2011). The

highly reproducible results obtained from NMR also facilitate the collaborative metabolomics data collection across different laboratories/research groups (Ward *et al.*, 2010). In addition, NMR is also a powerful tool for structural elucidation of unknown metabolites (Boiteau *et al.*, 2018). Hyphenated NMR techniques such as liquid chromatography LC-NMR (Spraul *et al.*, 2015) as well as the so-called hybrid MS/NMR (Bingol, 2018) are expected to exhibit their usefulness for research in plant metabolomics in the near future.

2.2.3 Workflow of metabolite profiling approach

The methodological pipeline of metabolite profiling typically consists of sample collection, data acquisition via the analytical platform, data processing and metabolite identification, statistical analysis and biological interpretation of the results. This general workflow is illustrated in Figure 3. More detailed descriptions of each section in this workflow are given in the following chapters.
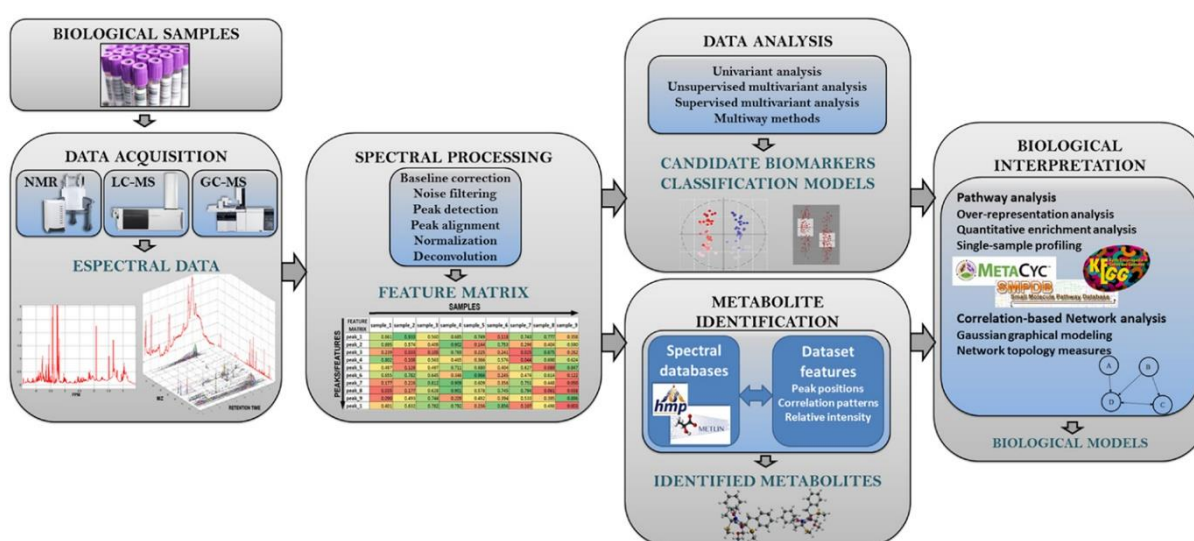
Figure 3. Exemplary workflow of the metabolite profiling approach; adapted with permission from (Alonso *et al.*, 2015)

2.2.4. Feature detection and metabolite identification

For each analyzed sample, the raw data acquired by MS-based metabolomics platforms/instruments are initially pre-processed to generate a data matrix of mass-to-

charge (*m/z*) ratio versus retention time (RT) versus ion intensity (Dunn *et al.*, 2011). This step is usually performed via vendor software coupled to the analytical instruments. In addition, some open-source tools are frequently used, e.g. XCMS (Smith *et al.*, 2006) and MZmine (Katajamaa *et al.*, 2006). Parameter setting of these feature-detecting tools showed a significant impact on the quality and the number of features that can be acquired (Yu *et al.*, 2014). Therefore, the parameters of the feature-detecting tools are suggested to be case-specific.

The metabolite identification is performed by comparing the retention times (RT) / retention indexes (RI) and spectral data of the detected features to those of authentic standards analyzed under the identical experimental conditions (Lisec *et al.*, 2006). However, authentic standards are not always commercially available or too expensive. In this case, a variety of public available metabolomics databases, e.g. Golm Metabolome Database (Kopka *et al.*, 2005), METLIN Metabolite Database (Smith *et al.*, 2005), BioMagResBank (Ulrich *et al.*, 2008), FiehnLib Library (Kind *et al.*, 2009), MassBank (Horai *et al.*, 2010) and MetaboLights Database (Haug *et al.*, 2013), would be alternatives to assist in putative annotation of metabolites.

The Metabolomics Standards Initiative (MSI) defined a standardized reporting of metabolomics-based metabolite identifications, in which the identification results were classified into four levels (Sumner *et al.*, 2007). A detailed description of the classification is shown in Table 3.

2.2.5 Data processing techniques for metabolomics

2.2.5.1 Peak alignment

Despite the good reproducibility of GC/MS analysis under optimized conditions, there are always some shifts in retention times for the same peaks over different sample batches, which can be due to the changes in the stationary phase of the capillary GC column, small variations in solvent compositions, gas flow and oven temperatures (Jonsson *et al.*, 2004). Therefore, peak alignment should be performed to eliminate this variability. A number of alignment software packages which are free, such as Chrom*pare* (Frenzel *et al.*, 2003), MZmine (Katajamaa *et al.*, 2006) and XCMS (Smith *et al.*, 2006), as well as commercially available tools including MarkerLynx (Waters Corporation), MassHunter (Agilent Technologies) and Sieve (Thermo Fisher Scientific) have been summarized (Katajamaa *et al.*, 2007).

Table 3. Classification of metabolomics-based metabolite identifications according to Sumner *et al.* (2007).

Level	Definition	Description
1	Identified compounds	Two or more independent and orthogonal data (e.g., retention time and mass spectrum, retention time and NMR spectrum, accurate mass and isotope pattern) relative to the authentic compound analyzed under identical experimental conditions
2	Putatively annotated compounds	Without authentic compounds, only based on commercial and/or public spectral libraries
3	Putatively characterized compound classes	Based on physicochemical properties of a chemical class of compounds, or by spectral similarity to known compounds from a chemical class
4	Unknown compounds	Unidentified or unclassified, but can still be differentiated and quantified based on spectral data

2.2.5.2 Data integrity check and missing value imputation

Sometimes missing values (peak intensities) in MS-based metabolomics datasets can be observed, which are mainly due to the following reasons: (i) the metabolite is detected in one sample but is absent in another sample; (ii) the metabolite is detected in one sample but its concentration in another sample is lower than the limit of detection (LOD); (iii) the metabolite is detected but the data processing software coupled to the instrument does not report the detection. It is very essential to replace these missing values by substitutes to maintain the integrity of the datasets in the data processing pipeline prior to univariate and multivariate data analysis (Hrydziuszko *et al.*, 2012; Armitage *et al.*, 2015; Di Guida *et al.*, 2016).

To date, a number of methods for missing value imputation are available. Zero, Mean and Median Replacement correspond to replace the missing values with zero, average values and median values of the corresponding metabolites/features, respectively (Steuer *et al.*, 2007). The method K-Nearest Neighbors (KNN) was initially developed

for the data matrix of gene expression (Troyanskaya *et al.*, 2001). In metabolomics datasets, KNN was applied to find k nearest samples via Euclidean metric, and then the missing value is replaced by the average values of those 'nearest neighbors'. The Random Forest (RF) imputation method is based on the machine-learning algorithm random forest (Stekhoven *et al.*, 2012). This method builds a mathematical model using selected metabolites with no missing values, and then those missing values of the target metabolites are predicted by the established model (Wei *et al.*, 2018). The computations of these methods were usually conducted by *R* package. Recently, a user-friendly and easily-accessible web tool MetImp (https://metabolomics.cc.hawaii.-edu/software/MetImp/) was developed, including most of the commonly used methods of missing value imputation (Wei *et al.*, 2018). The imputation performance of the different methods was evaluated using a GC/MS metabolomics dataset (Gromski *et al.*, 2014). The results suggested that the Random Forest (RF) method was the most robust approach, while the Mean Replacement showed the poorest performance.

2.2.5.3 Data filtering

Data filtering is employed to remove the noise signals or low-quality variables in a metabolomics dataset. By comparing the mean or median values, those variables showing very low or close-to-zero intensities are regarded as baseline noise and could be discarded. Low-quality variables often exhibit very close values throughout the whole dataset. These variables could be selected and filtered based on their low standard deviations (SD). This data-processing step is strongly recommended for metabolomics datasets containing a large number of variables (typically > 1,000), which could significantly improve the statistical power in the downstream multivariate and univariate statistical analyses (Gentleman *et al.*, 2004; Hackstadt *et al.*, 2009; Xia *et al.*, 2016).

2.2.5.4 Data standardization and normalization

The standardization, also referring to as Internal Standard (IS) Correction, is the step to reduce systematic bias between different measurements. Usually, the peak intensities of each metabolite in the dataset are standardized by correcting to the response of the single IS compound, or to the responses of multiple IS compounds corresponding to different chemical classes (Bijlsma *et al.*, 2006).

In biological samples, the concentrations of metabolites can vary over a wide range of magnitude. From a biological point of view, metabolites exhibiting high contents are not necessarily more important than those showing low levels (Wishart, 2011). In order to improve the consistency of the overall dataset and, at the same time, to retain the meaningful biological variations, normalization of the dataset is always performed. Normalization methods are divided into two strategies, i.e., data transformation and data scaling. Log and cube root transformations are the two most frequently used transformation methods. For data scaling, a number of different methods including auto scaling, pareto scaling, range scaling and vast scaling are commonly utilized (Craig *et al.*, 2006; Sysi-Aho *et al.*, 2007). A typical example of the data normalization effects on the original data obtained from GC/MS metabolite profiling is shown in Figure 4 (Berg *et al.*, 2006).

It is noteworthy that there is no consensus which transformation and/or scaling methods are most suitable for metabolomics datasets. It is suggested to perform different methods separately or to combine the transformation and scaling methods to assess the "bell shape" of the Gaussian distribution for the normalized data (Xia *et al.*, 2016).

2.2.6 Statistical analysis of metabolomics data

2.2.6.1 Multivariate statistical analysis

Mass spectrometry-based metabolomics studies often generate huge and complex data matrixes, which would be very difficult to be summarized and interpreted without appropriate statistical tools. Multivariate statistical methods are the most frequently utilized chemometric tools to summarize and visualize high-dimensional metabolomics data matrixes (Eriksson *et al.*, 2004). This kind of pattern recognition method could be subdivided into two groups: unsupervised and supervised statistical methods.
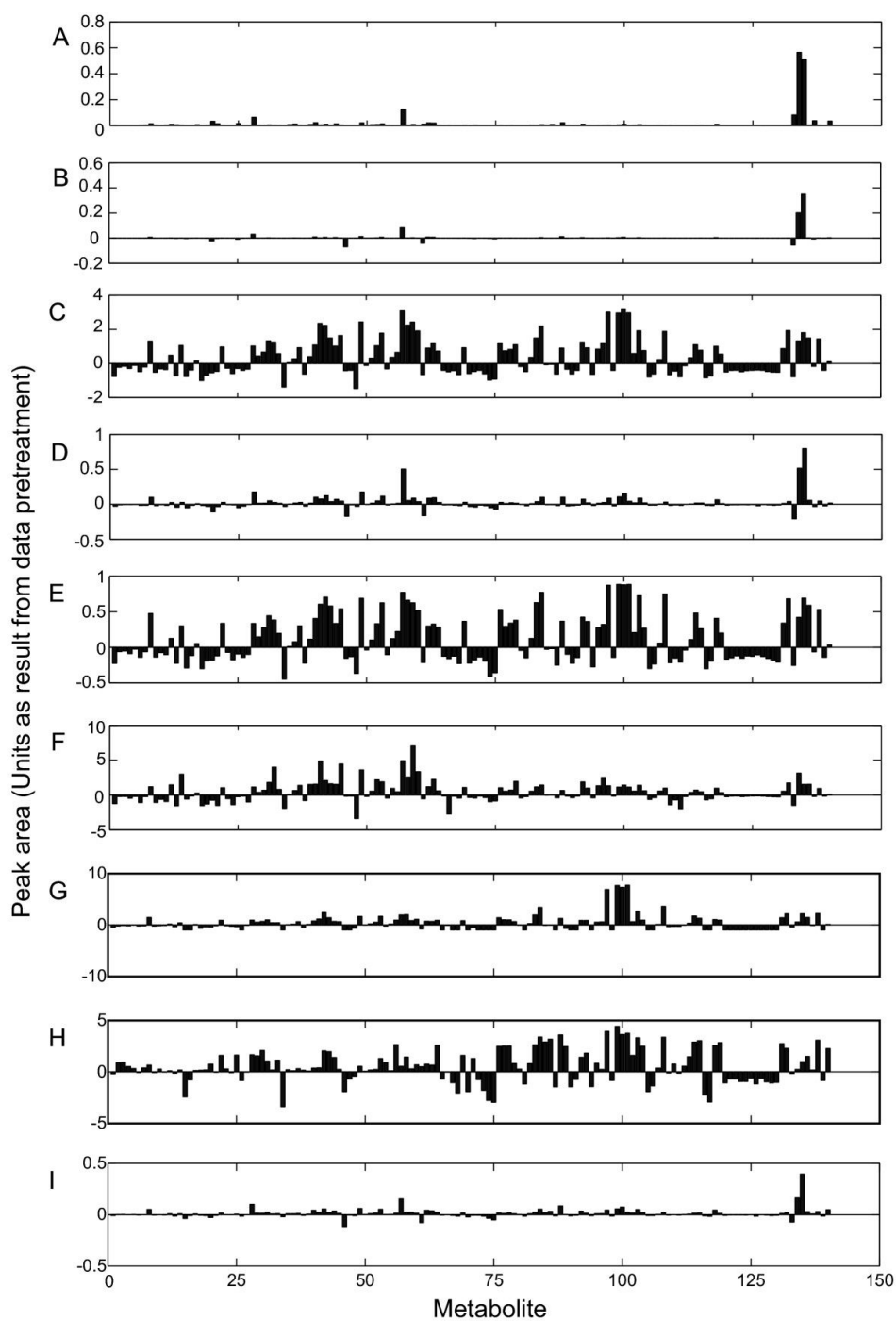
Figure 4. Effect of data normalization on the original data. Original data of experiment (A), and data after centering (B), autoscaling (C), pareto scaling (D), range scaling (E), vast scaling (F), level scaling (G), log transformation (H), and power transformation (I); adapted with permission from Berg *et al.* (2006).

Unsupervised statistical analysis, such as principal component analysis (PCA) and hierarchical clustering analysis (HCA), is often employed to generate an unbiased overview of the overall dataset with reduced dimensions, as well as to preliminarily explore the trends, clustering and outliers among the samples (Trygg *et al.*, 2007). In addition, PCA is also regarded as a primary assessment of the necessity of conducting supervised statistical analysis (Worley *et al.*, 2016).

Partial least squares-discriminate analysis (PLS-DA) is one of the most extensively performed supervised statistical methods. Samples are classified into different groups prior to being subjected to dimensionality reduction algorithms. Therefore, compared to PCA, PLS-DA is usually more favorable to reduce less important sources of data variance and to exhibit better differentiation between classes (Barker *et al.*, 2003). However, data variations that are uncorrelated to pre-defined group labels are also present in PLS-DA scores, which would lead to difficulties in identifying biological markers if the within-group variation is relatively high (Worley *et al.*, 2013; Mastrangelo *et al.*, 2015). In this regard, orthogonal partial least squares-discriminate analysis (OPLS-DA), another supervised statistical method, is more powerful in separating different sample groups by incorporating the Orthogonal Signal Correction (OSC) filter (Bylesjo *et al.*, 2006; Wiklund *et al.*, 2008). An illustration of the difference between PLS-DA and OPLS-DA models is shown in Figure 5.
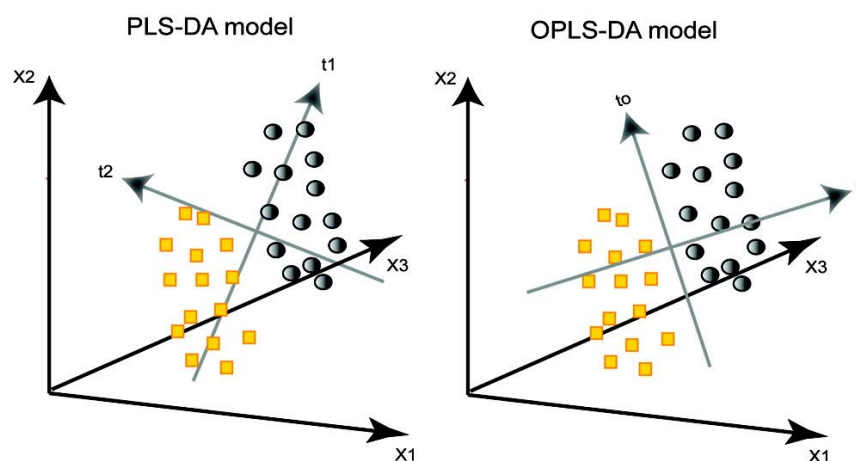
Figure 5. Geometrical illustration of the difference between PLS-DA and OPLS-DA models. The OPLS-DA model is rotated compared with the PLS-DA model. The class variation is found in the predictive component t_p, and the variation within each class is observed in the first y-orthogonal components t_o; adapted with permission from Wiklund *et al.* (2008).

The disadvantage of the supervised statistical method is the high probability of over-fitting, which aggressively forces separations between biological samples. As a result, the good classification between groups is obtained sometimes by chance (Westerhuis *et al.*, 2008). Therefore, rigorous validation procedures, such as cross-validation and permutation test, are very essential to evaluate the reliability and robustness of the established (O)PLS-DA models (Triba *et al.*, 2015).

2.2.6.2 Univariate statistical analysis

Univariate statistical analysis is a classic statistical approach to compare the individual metabolites from different groups. Prior to the statistical analysis, the statistical properties of the metabolites need to be assessed. Based on results of normality test (e.g. Shapiro-Wilk and Jarque-Bera tests) and homogeneity test of variances (e.g. Levene's and Bartlett's tests) for the metabolomics derived-data, parametric tests such as Student's *t*-test and ANOVA with post hoc test are performed for those metabolites exhibiting a normal distribution and homogeneity of variance. Otherwise, the non-parametric Mann-Whitney test (for two-group comparison) and the Kruskal-Wallis test (for multiple-group comparison) are conducted (Korman *et al.*, 2012; Vinaixa *et al.*, 2012).

The parallel application of univariate statistical analysis to single metabolites has the advantage of being easy to use and to interpret. However, this method can lead to considerable false positive results if multiple tests are performed in parallel for hundreds of metabolites. For example, the *p*-value threshold of 0.05 implies that there is 5% possibility to obtain a false positive result of the statistically significant difference between compared samples, which is acceptable for a single metabolite test. However, the number of false positive result will drastically increase to 100, when 2,000 variables from a metabolomics dataset are simultaneously subjected to univariate statistical analysis (Vinaixa *et al.*, 2012). Therefore, multiple testing correction methods need to be considered as a fundamental principle of the univariate statistical analysis of the metabolomics derived-data (Storey *et al.*, 2003; Broadhurst *et al.*, 2006).

One of the solutions of this multiple-testing problem is the False discovery rate (FDR) correction, which computes the *p*-corrected value (also referred to as *q*-value) for each tested metabolite (Pike, 2011). An FDR-adjusted, *p*-corrected value of 0.05 means that only 5% of test results with a statistically significant difference will be false positive, which considerably reduces the possibilities of false positive cases. Bonferroni

correction is another approach to handle the multiple-testing problem, by strictly controlling the false positive rate at the expense of false negative rate (Broadhurst *et al.*, 2006; Vinaixa *et al.*, 2012).

2.2.7 Application of plant metabolomics

In recent years, metabolomics studies in various fields of plant science research have been extensively reported, such as the investigation of plant responses to environmental impact, the exploration of phytochemical diversities of plant species, the evaluation of postharvest quality of horticultural plants, the substantial equivalence assessment of genetically modified (GM) crops and the assistance in plant breeding research (Shulaev *et al.*, 2008; Beckles *et al.*, 2012; Kusano *et al.*, 2012; Herrmann *et al.*, 2013). In the following, some examples of metabolomics studies in plant research are presented.

2.2.7.1 Metabolic changes in response to stress

The GC/MS-based metabolite profiling approach has been applied to investigate the drought stress-induced metabolic changes in barley grains obtained from different genotypes, including both drought resistant and susceptible cultivars (Lanzinger *et al.*, 2015; Wenzel *et al.*, 2015). Different metabolic responses were observed in the seedling roots of wild, semi-wild and cultivated soybean lines under salt stress (Li *et al.*, 2017). Metabolite profiling via FT-ICR/MS revealed comprehensive rearrangement of the leaf metabolome of poplars grown under high UV-B radiation. Several hundred metabolites were up- and down-regulated over various pathways, including increased levels of flavonoids, anthocyanins and polyphenols, and decreased levels of steroids (Kaling *et al.*, 2015). In addition to these abiotic stresses, metabolite profiling was also employed to study the metabolic changes of plants induced by biotic factors. For example, by comparing the metabolite profiles of citrus fruits from both healthy and symptomatic trees suffering from greening disease, citric acid, nobiletin, malic acid and phenylalanine were identified as the metabolic biomarkers for the citrus trees infected with the specific bacterium *Candidatus* Liberibacter (Hung *et al.*, 2018).

2.2.7.2 Exploration of phytochemical diversities

Several metabolomics-based studies have revealed the diversities of phytochemical compositions of plants. For red and blue potato tubers, significant differences in the profiles of anthocyanins and other polyphenols were observed via LC/MS metabolite profiling (Oertel *et al.*, 2017). Similarly, different metabolite profiles were reported for red, black and colorless rice grains (Frank *et al.*, 2012a; Kim *et al.*, 2015) as well as for several onion cultivars with different colors (Bottcher *et al.*, 2018). Metabolic variation between the two major rice subspecies, *japonica* and *indica*, was investigated based on the complementary LC/MS and GC/MS platforms, and correlations between metabolic phenotype and geographic origin of the rice seeds were revealed (Hu *et al.*, 2014). The obtained information on phytochemical characteristics in plants resulting from biodiversity may provide useful information to breeders for generating novel plant cultivars.

2.2.7.3 Evaluation of postharvest quality

The postharvest qualities of fruits and vegetables, including visual, textural, flavor and compositional qualities, are crucial determinants of profitability for producers and of acceptability by consumers (Brasil *et al.*, 2018). A number of metabolomics-based studies were conducted to understand the impact of postharvest treatment on the qualities of fruits and vegetables. The metabolite profiles of apple fruits from two varieties were observed to respond differently to low oxygen storage condition, which were reflected by differently changed levels in both the volatile compounds (e.g. ethyl acetate, ethyl propanoate and ethyl tiglate) and non-volatile compounds (e.g. proline, glycine, serine and palmitic acid) (Brizzolara *et al.*, 2017). The metabolic changes of banana during postharvest senescence at five stages were elucidated by 1D and 2D NMR-based metabolite profiling approaches (Yuan *et al.*, 2017). Similar metabolomics-based studies on postharvest quality evaluation were also carried out for fruits such as grapes (Toffali *et al.*, 2011), peaches (Lauxmann *et al.*, 2014) and kiwi (Lim *et al.*, 2017) as well as for vegetables such as tomatoes (Luengwilai *et al.*, 2012), radish roots (Jahangir *et al.*, 2014) and lettuce (Garcia *et al.*, 2016).

2.2.7.4 Safety and authenticity assessment of plant-derived food

Metabolite profiling techniques are also useful for safety and authenticity assessments of plant-derived foods. The kernels of transgenic and non-GM maize were subjected

to a GC/MS-based metabolite profiling approach, and the results indicated that the environmental factors (growing location and season) contributed to much higher numbers of significant differences (up to 42%) in metabolite levels than the genetic modification (up to 4%) (Frank *et al.*, 2012b). To reveal the unintended metabolic changes induced by the mutation, metabolite profiling was employed to compare the metabolic differences between *low phytic acid* (*lpa*) mutant seeds and the corresponding wild-types both for rice (Frank *et al.*, 2007) and soybean (Frank *et al.*, 2009b). For the authenticity assessment, a metabolic approach based on large-scale sample size was applied to study the differences of metabolite compositions of carrots produced via conventional or organic agricultural systems. The OPLS-DA model established on the basis of the metabolite profiling data was shown to successfully distinguish the origin of the agricultural systems of the harvested carrots (Cubero-Leon *et al.*, 2018). An FT-ICR/MS-based metabolomics study on 18 Chardonnay white wines revealed that the wines could be discriminated according to the concentrations of SO_2 added to the must during the winemaking process. In addition, there was a correlation between the wine metabolomics picture and the type of stopper used (Roullier-Gall *et al.*, 2017).

3. RESULTS

3.1 Publication I

Zhou, C. G.; Tan, Y. Y.; Goßner, S.; Li, Y. F.; Shu, Q. Y.; Engel, K. H.
Impact of cross-breeding of *low phytic acid* Rice (*Oryza sativa* L.) mutants with commercial cultivars on the phytic acid contents.
Eur. Food Res. Technol. 2019, 245, 707-716
Reprinted with permission from Springer Nature (Copyright 2018).

Phytic acid (PA) is the major storage form of phosphorus in rice (*Oryza sativa* L.) seeds. Owing to its antinutritional properties, a number of *low phytic acid* (*lpa*) rice mutant lines have been obtained via mutation breeding. To improve the agronomic performance (e.g. grain yield and seed viability) of the *lpa* rice mutants, cross and selection breeding with commercial cultivars is applied. To investigate the impact of such cross-breeding steps on the PA contents of the resulting progenies, three previously developed *lpa* rice mutants, i.e. *Os-lpa*-XQZ-1, *Os-lpa*-XS110-2 and *Os-lpa*-MH86-1, carrying different mutation targets were crossed with the three commercial rice cultivars JX081, JH218 and JH99. The crossbred progenies of generations F_4 to F_8 grown at different field trials were subjected to a high-pressure ion chromatography (HPIC)-based approach for determination of the PA contents. The results revealed the impact of mutation types, PA contents of the crossing parents and the environmental conditions on the PA contents in homozygous *lpa* mutant progenies. However, despite the variability of the PA contents, the *lpa* trait was consistently expressed in homozygous *lpa* progenies over several generations harvested from different growing locations. From an agronomic point of view, the study demonstrated that a major prerequisite for the commercial production of elite *lpa* cultivars via cross and selection breeding is being fulfilled.

Candidate's contribution: Implementation of a High Pressure Ion Chromatography (HPIC) method for the analysis of phytic acid in rice; quantitation of phytic acid in all investigated rice samples; statistical assessment of the elaborated analytical data and of the agronomic performance data provided by the cooperation partner, writing and revision of the complete manuscript and of the Supplementary Material.

ORIGINAL PAPER

CrossMark

Impact of cross-breeding of *low phytic acid* rice (*Oryza sativa* L.) mutants with commercial cultivars on the phytic acid contents

Chenguang Zhou[1] · Yuanyuan Tan[2] · Sophia Goßner[1] · Youfa Li[3] · Qingyao Shu[2] · Karl-Heinz Engel[1]

Abstract

Phytic acid (PA) is considered as antinutrient in food and feed. Several *low phytic acid* (*lpa*) rice (*Oryza sativa* L.) mutants have been generated through induced mutation. *Lpa* mutant lines often exhibit inferior agronomic performance, and cross-breeding is applied to minimize these effects. The impact of such crossing steps on the PA contents in the resulting progenies is unknown. Therefore, three *lpa* rice mutants differing in mutation type were crossed with commercial rice cultivars, and PA contents in the progenies were determined for various generations grown at different locations. The PA contents of the *lpa* progenies were differently expressed for the investigated mutation-types and were dependent on environment and/or PA contents of the crossing parents. Nevertheless, for all three mutants, the homozygous *lpa* progenies always displayed significantly lower PA contents than the original wild-types subjected to the mutation. This demonstrated that cross-breeding of *lpa* rice mutants with commercial cultivars does not compromise the intended PA reduction and is a useful tool to obtain mutants stably expressing the intended *lpa* trait.

Keywords Rice (*Oryza sativa* L.) · Induced mutation · *Low phytic acid* (*lpa*) rice · Cross-breeding

Introduction

Phytic acid (*myo*-inositol-1,2,3,4,5,6-hexa*kis*phosphate, PA) is the major storage form of phosphorus in cereal grains and legume seeds [1]. However, it is considered as an antinutrient in food and feed, since it may form indigestible salts of cations such as Zn^{2+} and Fe^{2+}, and thus lead to a reduced bioavailability of these minerals for humans and monogastric animals [2, 3]. During the past years, various *low phytic* acid (*lpa*) crops comprising maize, barley, rice, wheat and soybean have been obtained through chemical/physical mutagenesis or genetic engineering [4]. Nutritional studies have revealed that consumption of *lpa* grains and beans improves the utilization of mineral nutrients for humans [5, 6] and livestock [7, 8].

Rice (*Oryza sativa* L.) is one of the major cereal grains serving as staple food for a large part of the world's population [9]. Accordingly, there have been various attempts to produce *lpa* rice crops. For example, five non-lethal *lpa* rice mutant lines have been developed with PA contents 35–65% lower than those of the corresponding wild-types [10]. *Lpa* mutant lines often exhibit lower grain yield and seed viability compared with the wild-type parents, and there have been attempts to minimize these negative effects by cross and selection breeding of *lpa* rice mutants [11]. However, the impact of these crossing steps on the PA contents in the resulting progenies has not been investigated. A limited amount of information regarding *lpa* crossbred lines is only available for *lpa* soybean mutants [12–14].

Therefore, the objective of this study was to investigate in more detail the impact of cross-breeding of *lpa* rice mutants on the PA contents of the resulting progenies. Three *lpa* rice mutants differing in the type of mutation were crossed with

✉ Karl-Heinz Engel
k.h.engel@wzw.tum.de

[1] Chair of General Food Technology, Technical University of Munich, Maximus-von-Imhof-Forum 2, 85354 Freising-Weihenstephan, Germany

[2] State Key Laboratory of Rice Biology and Zhejiang Provincial Key Laboratory of Plant Germplasm, Institute of Crop Sciences, Zhejiang University, Hangzhou 310058, China

[3] Jiaxing Academy of Agricultural Sciences, Jiaxing 314016, China

🖄 Springer

three commercial rice cultivars, and the PA contents in the resulting progenies were determined for various generations grown at different locations.

Materials and methods

Chemicals

Phytic acid dodecasodium salt and iron(III) nitrate nonahydrate were purchased from Sigma-Aldrich (Steinheim, Germany), hydrochloric acid (0.5 M, analytic grade) from Merck (Darmstadt, Germany), perchloric acid (0.33 M, HPLC grade) from Bernd Kraft (Duisburg, Germany), and deionized water (LC–MS grade) from VWR International (Darmstadt, Germany).

Sample materials

The *lpa* rice mutants *Os-lpa*-XQZ-1, *Os-lpa*-XS110-2 and *Os-lpa*-MH86-1 have previously been generated from the corresponding wild-types XQZ-B, XS110 and MH86, respectively, via physical and chemical mutagenesis [10]. In *Os-lpa*-XQZ-1, γ-irradiation resulted in a mutated homolog of the 2-phosphoglycerate kinase (*2-PGK*) gene [15]. The mutation of *Os-lpa*-XS110-2 was attributed to a single base pair change (C/G–T/A transition) of the multi-drug resistance-associated protein ABC transporter gene 5 (*OsMRP5*) [16]. In *Os-lpa*-MH86-1, a 1-bp deletion in a putative sulfate transporter gene (*OsSULTR3;3*) was shown to be responsible for the *lpa* phenotype [17]. The employed wild-type crossing parents were the two commercial rice cultivars Jiaxian 081 (JX081) and Jiahe 218 (JH218), and the restorer line (JH99) of the hybrid cultivar Jiayou 99, showing yields of 8–9 t/ha and meeting the Chinese national quality standards. In total, progenies were obtained from six crosses, i.e. *Os-lpa*-XQZ-1×JX081, *Os-lpa*-XQZ-1×JH218; *Os-lpa*-XS110-2×JX081, *Os-lpa*-XS110-2×JH218; *Os-lpa*-MH86-1×JH99, *Os-lpa*-MH86-1×JH218. F_2 seeds were bulk-harvested and grown into plants. For each cross, 200 F_2 plants were genotyped for the *lpa* mutations and classified into three types, i.e. homozygous wild-type, homozygous *lpa* type, and heterozygous progenies (Figure S1). Plants were also segregating for other agronomic traits, e.g. flowering time, plant height and fertility; therefore, five F_2 plants each of the homozygous and ten each of the heterozygous genotype were chosen based on similarity of flowering time, and their F_3 seeds were harvested. While seeds of homozygous F_2 plants were stored for growth until 2013, those from heterozygous F_2 plants were grown into $F_{2:3}$ plots and their F_3 plants were genotyped for the *lpa* mutation and classified into three types as in F_2. Similar work was performed in F_4–F_6, and seeds harvested from homozygous F_2–F_5 plants

were grown in a common paddy field in 2013 in Jiaxing for the production of sample materials (F_4–F_7). The homozygous F_8 seeds were harvested at three field trials in Jiaxing 2014, Wuxi 2014 and Hainan 2015. For all field trials, in addition to the progenies, the respective wild-types, the progenitor *lpa* mutants and the crossing parents were harvested. At each field trial, the same standard agronomic practice (fertilization and water management) was applied.

Owing to the limited amount of the sample materials, for each cross and each generation rice seeds from 3 to 4 plants of each phenotype were pooled. After dehulling and grinding with a cyclone mill equipped with a 500-μm sieve (Foss, Rellingen, Germany), the flour was freeze-dried for 48 h and stored at − 20 °C until analysis.

Genotyping of progenies

For the genotyping, genomic DNA was extracted from leaf tissues following a previously described method [18]. For *Os-lpa*-XS110-2 and *Os-lpa*-MH86-1, high-resolution melting curves were used to distinguish between homozygous *lpa* mutant, homozygous wild-type and heterozygous progenies. The genotyping of progenies from *Os-lpa*-XQZ-1 crosses was performed using gel-based molecular markers [18].

Phytic acid analysis

Extraction and analysis of PA were performed according to a procedure described for dried distillers grains [19]. For each sample, three aliquots (100 mg) of freeze-dried rice flour were weighed into 50 mL polystyrene centrifuge tubes. Twenty milliliters of 0.5 M HCl was added to each aliquot. The extractions of PA were performed under sonication for 20 min after the mixture had been thoroughly vortexed. After centrifugation at 2000g for 20 min, the supernatants were filtered through a 0.22 μm polyether sulfone filter disk, and 100 μL was subjected to ion chromatography.

PA analysis was performed on a Thermo Scientific Dionex ICS-5000 high-pressure ion chromatography (HPIC) system equipped with a CarboPac PA100 guard column (4×50 mm) and a CarboPac PA100 analytical column (4×250 mm) constantly thermostatted to 30 °C. A gradient elution was performed by two mobile phases, i.e. deionized water and 0.5 M HCl, at a flow rate of 1 mL/min. The initial proportion of gradient HCl of 5% was increased linearly to 100% within 4 min, and then held for 10 min. Post-column derivatization was based on reaction with ferric nitrate (solution of 0.1% Fe(NO$_3$)$_3$ in 0.33 M perchloric acid) at a flow rate of 0.4 mL/min; peak detection was performed at 290 nm. Quantification of PA was based on an external standard calibration curve (R^2 = 0.9997) with a recovery rate of 98.8%. The limit of detection (LOD) and the limit of

27

quantification (LOQ) were 1.0 mg/L and 2.9 mg/L, respectively [20].

Statistical analysis

Values were expressed as means $\pm$ standard deviations (SD) resulting from the analysis of three aliquots of freeze-dried flour. Student's t test and ANOVA analysis were performed by XLSTAT (version 19.5, France).

Results

Progenies resulting from crosses of the *lpa* mutants *Os-lpa*-XQZ-1, *Os-lpa*-XS110-2 and *Os-lpa*-MH86-1 with commercial rice cultivars were genotyped using previously developed high-resolution melting curve analyses and gel-based molecular markers, respectively [18]. An overview of the investigated samples is given in Fig. 1.

The PA contents and their percentage reductions in the *lpa* mutants determined in the field trials performed in this study (Table 1) were in the same order of magnitude as

Fig. 1 Overview on investigated rice samples

previously reported data [10, 17, 21, 22]. The PA levels of the three commercial rice cultivars used as crossing parents varied depending on growing location and year. JX081 consistently exhibited the highest PA contents, followed by JH218 and JH99 (Table S1).

Phytic acid contents in progenies of crosses of *Os-lpa*-XQZ-1

In the homozygous *lpa* mutant of generation F_4 of cross *Os-lpa*-XQZ-1$\times$JX081, the PA content as well as its percentage decrease (Table 2) was in good agreement with the results determined for the progenitor *Os-lpa*-XQZ-1 mutant (Table 1). The homozygous *lpa* mutants from the following generations F_5 and F_7 exhibited significantly lower PA levels than the F_4 *lpa* mutant, resulting in more pronounced percentage reductions of PA. All decreases of PA contents observed for the homozygous *lpa* mutant progenies are in agreement with data formerly reported for the *lpa* mutants of this mutation type [10, 21]. Similar results regarding the PA content and its percentage reduction were observed for the homozygous *lpa* mutants of cross *Os-lpa*-XQZ-1$\times$JH218.

Phytic acid contents in progenies of crosses of *Os-lpa*-XS110-2

For the cross *Os-lpa*-XS1102$\times$JX081, the PA content in the homozygous *lpa* mutant from generation F_4 (Table 2) was similar to the average content determined for the progenitor *Os-lpa*-XS110-2 mutants in the field trials performed in the present study (Table 1) as well as to the PA contents reported for this type of mutant in the literature [10, 21, 22]. The percentage reduction of PA was also within the previously reported range (-11.7 to -33.9%) [10, 21, 22]. PA contents in the homozygous *lpa* progenies of the following generations F_5–F_7 were significantly lower than in the homozygous

Table 1 Comparison of the PA contents (mg/g dry matter) in progenitor *lpa* mutants (*Os-lpa*-XQZ-1, *Os-lpa*-XS110-2 and *Os-lpa*-MH86-1) and the corresponding wild-types (XQZ-B, XS110 and MH86)

Field trial	XQZ-B	*Os-lpa*-XQZ-1	Decrease (% change)[a]	XS110	*Os-lpa*-XS110-2	Decrease (% change)[a]	MH86	*Os-lpa*-MH86-1	Decrease (% change)[a]
Jiaxing 2013	6.70 ± 0.05	5.30 ± 0.05	$-20.8*$	n.a.[b]	5.38 ± 0.10	n.a.	9.69 ± 0.20	6.27 ± 0.05	$-35.5*$
Jiaxing 2014	8.45 ± 0.09	6.19 ± 0.04	$-26.8*$	7.77 ± 0.03	6.33 ± 0.08	$-18.5*$	9.15 ± 0.10	5.27 ± 0.10	$-42.4*$
Wuxi 2014	7.56 ± 0.07	6.35 ± 0.10	$-16.0*$	n.a.	3.66 ± 0.55	n.a.	9.98 ± 0.10	6.24 ± 0.03	$-37.5*$
Hainan 2015	n.a.	5.72 ± 0.27	n.a.	n.a.	n.a.	n.a.	n.a.	6.60 ± 0.10	n.a.
Mean	7.57 ± 0.76	5.89 ± 0.44	$-21.2*$	7.77 ± 0.03	5.12 ± 1.18	$-18.5*$	9.60 ± 0.39	6.06 ± 0.51	$-37.7*$

Values represent means $\pm$ standard deviations resulting from the analysis of three aliquots of freeze-dried flour

[a]Asterisks indicate statistically significant differences (Student's t-test, $p < 0.05$) between the *lpa* mutant and the corresponding wild-type

[b]Samples not available

Table 2 PA contents (mg/g dry matter) in homozygous wild-type and *lpa* mutant progenies (generations F_4–F_7) of crosses of *lpa* mutants (*Os-lpa*-XQZ-1, *Os-lpa*-XS110-2 and *Os-lpa*-MH86-1) with commercial rice cultivars (JX081, JH218 and JH99) grown in Jiaxing in 2013

Generation	*Os-lpa*-XQZ-1×JX081		% Change[b]	*Os-lpa*-XQZ-1×JH218		% Change[b]
	Wild-type	*Lpa* mutant[a]		Wild-type	*Lpa* mutant[a]	
F_4	7.95 ± 0.18	6.23 ± 0.06 a	− 21.6	7.5 ± 0.14	5.95 ± 0.06 a	− 20.7
F_5	6.59 ± 0.07	4.56 ± 0.07 b	− 30.8	6.7 ± 0.03	5.27 ± 0.10 b	− 21.4
F_6	n.a.[c]	n.a.	n.a.	n.a.	n.a.	n.a.
F_7	7.88 ± 0.12	4.61 ± 0.14 b	− 41.5	7.02 ± 0.08	4.61 ± 0.14 c	− 34.4
Mean	7.47 ± 0.67	5.13 ± 0.83	− 31.3	7.07 ± 0.36	5.28 ± 0.59	− 25.3
Generation	*Os-lpa*-XS110-2×JX081		% Change[b]	*Os-lpa*-XS110-2×JH218		% Change[b]
	Wild-type	*Lpa* mutant[a]		Wild-type	*Lpa* mutant[a]	
F_4	7.94 ± 0.09	5.53 ± 0.04 a	− 30.4	5.24 ± 0.08	4.16 ± 0.12 c	− 20.6
F_5	6.91 ± 0.09	4.93 ± 0.10 c	− 28.6	4.77 ± 0.02	3.97 ± 0.03 c	− 16.7
F_6	7.52 ± 0.06	5.34 ± 0.01 b	− 29.1	5.33 ± 0.05	4.43 ± 0.09 b	− 17.0
F_7	7.18 ± 0.03	5.31 ± 0.07 b	− 26.1	6.42 ± 0.15	4.98 ± 0.08 a	− 22.5
Mean	7.39 ± 0.41	5.28 ± 0.23	− 28.6	5.44 ± 0.64	4.38 ± 0.40	− 19.5
Generation	*Os-lpa*-MH86-1×JH99		% Change[b]			
	Wild-type	*Lpa* mutant[a]				
F_4	7.13 ± 0.03	4.54 ± 0.05 c	− 36.3			
F_5	9.79 ± 0.03	6.12 ± 0.01 a	− 37.5			
F_6	7.70 ± 0.04	5.38 ± 0.11 b	− 30.2			
F_7	9.27 ± 0.07	6.08 ± 0.06 b	− 34.4			
Mean	8.47 ± 1.14	5.53 ± 0.68	− 34.7			

Values represent means ± standard deviations resulting from the analysis of three aliquots of freeze-dried flour

[a]Different letters behind the values indicate statistically significant differences (ANOVA with Tukey's test, $p < 0.05$) among homozygous *lpa* mutants of different generations of the same cross

[b]For each generation, the changes between homozygous *lpa* mutant and homozygous wild-type progenies are statistically significant (Student's t-test, $p < 0.05$)

[c]Samples not available

F_4 *lpa* progeny; however, there was no consistent trend with progressing generations.

The PA content of the homozygous F_4 *lpa* progeny resulting from the cross *Os-lpa*-XS110-2×JH218 was significantly lower than that of the homozygous F_4 *lpa* progeny from the cross *Os-lpa*-XS110-2×JX081. However, the percentage reduction of PA in the homozygous F_4 *lpa* progeny remained in the range previously reported for the progenitor mutant *Os-lpa*-XS1102 [10, 21, 22], since the PA content in the corresponding homozygous wild-type progeny was also significantly lower than that in the homozygous F_4 wild-type progeny of the cross with JX081. In addition, no consistent changes in PA contents and their percentage reductions were observed with progressing generations in *lpa* progenies of the cross *Os-lpa*-XS110-2×JH218.

Phytic acid contents in progenies of crosses of *Os-lpa*-MH86-1

As shown in Table 2, in the homozygous F_4 *lpa* progeny of the cross *Os-lpa*-MH86-1×JH99, the PA content was significantly lower than that in the progenitor *Os-lpa*-MH86-1 mutant grown at the same field trial in Jiaxing 2013 (Table 1). However, the reduction degree remained comparable to that for the progenitor mutant *Os-lpa*-MH86-1, owing to the lower PA content in the homozygous F_4 wild-type progeny. For the homozygous *lpa* progenies of the following generations from F_5 to F_7, the PA contents and their percentage reductions were all consistently in accordance with data previously provided for the progenitor *Os-lpa*-MH86-1 mutant [10, 17].

Environmental impact on the phytic acid contents in homozygous *lpa* mutant progenies

To investigate the environmental impact on the PA contents, for each cross F_8 progenies from several field trials were assessed (Table 3). The homozygous F_8 *lpa* mutants of cross *Os-lpa*-XQZ-1×JX081 exhibited remarkable percentage reductions of PA in the three field trials, ranging from −41.5 to −50.6% (Table 3). There were no statistically significant differences in the PA contents of the *lpa* mutant progenies depending on the field trial. This is in agreement with the fact that the PA contents determined in the progenitor *Os-lpa*-XQZ-1 mutants at the field trials Jiaxing 2014 and Wuxi 2014 (Table 1) were very similar. Pronounced reductions of PA were also found for the homozygous F_8 *lpa* progenies of cross *Os-lpa*-XQZ-1×JH218. There were statistically significant differences of the PA contents depending on the field trial. However, the patterns of both, the PA contents and the percentage decreases, were quite comparable between the two crosses. Also remarkable were the nearly identical PA contents determined in the homozygous wild-type progenies for both crosses. Overall, the data suggest that for this type

of mutant resulting from a mutation of the *2-PGK* gene [15], the PA contents are rather unsusceptible to environmental impact and that this is not altered by cross-breeding with commercial cultivars.

For the cross *Os-lpa*-XS110-2×JX081, the homozygous F_8 *lpa* progenies grown in Jiaxing 2014 showed a PA content comparable to that in the progenitor *Os-lpa*-XS110-2 mutant grown at the same field trial (Table 1), and the PA reduction of −30.3% was also in good agreement with the data of the crossbred *lpa* mutants from generations F_4–F_7 (Table 2). The homozygous F_8 *lpa* progenies from another field trial Wuxi 2014 displayed a significantly lower PA content than that of the F_8 *lpa* progenies from Jiaxing 2014. Similar differences depending on the field trials were observed for the homozygous F_8 *lpa* progenies of cross *Os-lpa*-XS110-2×JH218. The progenitor *lpa* mutant *Os-lpa*-XS110-2 from the field trials Jiaxing 2014 and Wuxi 2014 also exhibited the same pattern regarding the PA contents (Table 1). The results indicate that the observed environment-related effect on the PA contents in the *lpa* progenies is specific for this type of mutant resulting from a mutation of the *OsMRP5* gene [16], and not a result of the employed cross-breeding.

Table 3 PA contents (mg/g dry matter) in homozygous wild-type and *lpa* mutant progenies (generation F_8) of crosses of *lpa* mutants (*Os-lpa*-XS110-2, *Os-lpa*-XQZ-1 and *Os-lpa*-MH86-1) with commercial rice cultivars (JX081, JH218 and JH99) grown in different field trials

Generation	*Os-lpa*-XQZ-1×JX081		% Change[b]	*Os-lpa*-XQZ-1×JH218		% Change[b]
	Wild-type	*Lpa* mutant[a]		Wild-type	*Lpa* mutant[a]	
Jiaxing 2014	9.67±0.07	4.98±0.11 a	−48.5	9.54±0.11	5.12±0.06 b	−46.3
Wuxi 2014	8.59±0.09	5.03±0.03 a	−41.5	8.64±0.10	5.67±0.04 a	−34.4
Hainan 2015	9.41±0.16	4.65±0.24 a	−50.6	9.42±0.06	4.67±0.08 c	−50.4
Mean	9.22±0.50	4.89±0.22	−47.0	9.20±0.43	5.15±0.44	−44.0
Generation	*Os-lpa*-XS110-2×JX081		% Change[b]	*Os-lpa*-XS110-2×JH218		% Change[b]
	Wild-type	*Lpa* mutant[a]		Wild-type	*Lpa* mutant[a]	
Jiaxing 2014	9.43±0.07	6.57±0.06 a	−30.3	8.17±0.06	6.46±0.11 a	−20.9
Wuxi 2014	8.29±0.08	5.49±0.09 b	−33.7	6.53±0.07	4.74±0.08 b	−27.5
Hainan 2015	n.a.[c]	n.a.	n.a.	n.a	n.a.	n.a.
Mean	8.86±0.63	6.03±0.60	−31.9	7.35±0.90	5.60±0.95	−23.8
Generation	*Os-lpa*-MH86-1×JH99		% Change[b]	*Os-lpa*-MH86-1×JH218		% Change[b]
	Wild-type	*Lpa* mutant[a]		Wild-type	*Lpa* mutant[a]	
Jiaxing 2014	6.98±0.14	4.31±0.04 c	−38.3	7.91±0.11	6.18±0.06 c	−22.0
Wuxi 2014	7.56±0.13	4.93±0.07 b	−34.8	8.87±0.03	7.02±0.07 b	−20.9
Hainan 2015	10.07±0.15	5.99±0.19 a	−40.5	9.09±0.16	7.34±0.07 a	−19.2
Mean	8.20±1.43	5.08±0.75	−38.0	8.62±0.55	6.84±0.53	−20.6

Values represent means ± standard deviations resulting from the analysis of three aliquots of freeze-dried flour

[a]Different letters behind the values indicate statistically significant differences (ANOVA with Tukey's test, $p < 0.05$) in phytic acid levels among homozygous *lpa* mutants of different generations of the same cross

[b]For each field trial, the changes between homozygous *lpa* mutant and homozygous wild-type progenies are statistically significant (Student's t-test, $p < 0.05$)

[c]Samples not available

🙋 Springer

For the cross *Os-lpa*-MH86-1×JH99, the highest PA level in homozygous F$_8$ *lpa* mutants was found for the field trial Hainan 2015, followed by Wuxi 2014 and Jiaxing 2014 (Table 3). The PA contents in the corresponding wild-type progenies exhibited a similar trend among the three field trials, thus leading to comparable reductions of PA contents in the *lpa* mutants, ranging from −34.8 to −40.5%, which were in accordance with those observed in generations F$_4$–F$_7$ (Table 2) and data previously provided for the progenitor *Os-lpa*-MH86-1 mutant [10, 17]. The PA contents in the homozygous F$_8$ *lpa* mutants of cross *Os-lpa*-MH86-1×JH218 were also found to be the highest in the field trial Hainan 2015, followed by those in Wuxi 2014 and Jiaxing 2014. The percentage reductions in these *lpa* mutants were consistent (−19.2 to −22.0%) but less pronounced than that for the cross *Os-lpa*-MH86-1×JH99. A similar pattern of environmental impact has been determined for the progenitor *Os-lpa*-MH86-1 mutant, with PA contents ranging from the highest value for samples from Hainan 2015 to the lowest for samples from Jiaxing 2014 (Table 1). This indicates that also in this case the mutant rather than the employed crossing step is the driver for the environment-related differences observed in the PA contents of the homozygous *lpa* progenies.

Impact of the phytic acid content of the crossing parent

The data presented in Fig. 2a demonstrate that for the mutant *Os-lpa*-XQZ-1 there was no consistent correlation between the PA contents of the crossing parents (JX081 and JH218) and those of the corresponding homozygous F$_8$ *lpa* progenies. Despite the higher PA content of JX081, the PA contents of the resulting homozygous *lpa* progenies were either the same (Jiaxing 2014, Hainan 2015) or they showed the opposite ratio (Wuxi 2014) compared to the crossing parents.

For *Os-lpa*-XS110-2 the differences in PA contents of the two crossing parents were reflected in corresponding statistically significant differences in the homozygous F$_8$ *lpa* mutants at one of the investigated field trials (Fig. 2b).

The most obvious correlation between the PA contents of the crossing parents and those of the homozygous *lpa* progenies was observed for the crosses of *Os-lpa*-MH86-1. As shown in Fig. 2c, the PA content in the crossing parent JH218 was significantly higher than that in JH99. This difference was clearly reflected in the resulting homozygous F$_8$ *lpa* progenies. For each field trial, the PA content in the homozygous *lpa* progeny resulting from the cross of *Os-lpa*-MH86-1 with JH218 was significantly higher than in the *lpa* progeny resulting from the cross with JH99. In consequence, the PA contents in the *lpa* progenies of the cross with JH218 were even consistently higher than those in the

progenitor *lpa* mutant grown under the respective field trial conditions. Nevertheless, for the field trials Jiaxing 2014 and Wuxi 2014, it could be demonstrated that also for this mutant resulting from a disruption of the putative sulfate transporter gene *OsSULTR3;3* [17], the PA contents of the homozygous F$_8$ *lpa* progenies remained significantly lower than those in the original wild-type, independent from the employed crossing parents.

Discussion

For all three mutation-types, the progenies resulting from cross-breeding of the *lpa* mutants with commercial cultivars exhibited the *lpa* trait, i.e. there were significant reductions of the PA contents in the homozygous *lpa* progenies compared to the homozygous wild-type progenies. The reductions of PA contents in crossbred *lpa* progenies remained stable over several generations or even increased with progressing generations (Table 2). A similar study with soybeans demonstrated that after crossing a soybean *lpa* mutant with a normal-phytate (NP) line, the PA contents in the *lpa* progenies remained comparable to the progenitor *lpa* mutant and were significantly lower than those in the NP-line at five locations [12].

However, the impact of cross-breeding of the *lpa* rice mutants with commercial cultivars on the PA contents of the resulting homozygous *lpa* progenies differed depending on the type of mutation. To assess the importance of different sources to the variability of PA contents observed in the homozygous *lpa* mutants, a two-way ANOVA was performed considering the PA content of the commercial crossing parent, the environment, i.e. locations and years of the field trials, and their interaction as potential sources of variation (Table 4). The coefficients of variance demonstrated that the PA contents of the *lpa* progenies of *Os-lpa*-MH86-1 showed the highest variability, followed by the progenies of *Os-lpa*-XS110-2 and *Os-lpa*-XQZ-1. This is in agreement with the previously reported higher variability of PA contents in the progenitor *lpa* mutant *Os-lpa*-XS110-2 compared with *Os-lpa*-XQZ-1 at nine independent field trials [21].

The types of *lpa* mutants investigated in this study were selected to reflect different principles of PA reduction [23]. The observed differences in the variability of the PA contents of *lpa* progenies in response to the factors environment and phytic acid content of the crossing parent correlate with differences in complexity of the underlying mechanisms. *Os-lpa*-XQZ-1 represents a type of mutation where a step directly involved in the biosynthesis of PA in the pathway downstream of *myo*-inositol-3-phosphate is affected. The mutation of a homolog of the *2-PGK* gene influences the rate limiting step from InsP$_1$ to InsP$_2$ [15, 23]. Taking this

Fig. 2 PA contents of the original wild-types, the progenitor *lpa* mutants, the employed commercial crossing parents and the F_8 *lpa* progenies.
a *Os-lpa*-XQZ-1; **b** *Os-lpa*-XS110-2; **c** *Os-lpa*-MH86-1. The data were assessed by one-way ANOVA; different letters indicate significant differences (Tukey's test at $p < 0.05$) in the phytic acid levels of the rice seeds from the same field trial
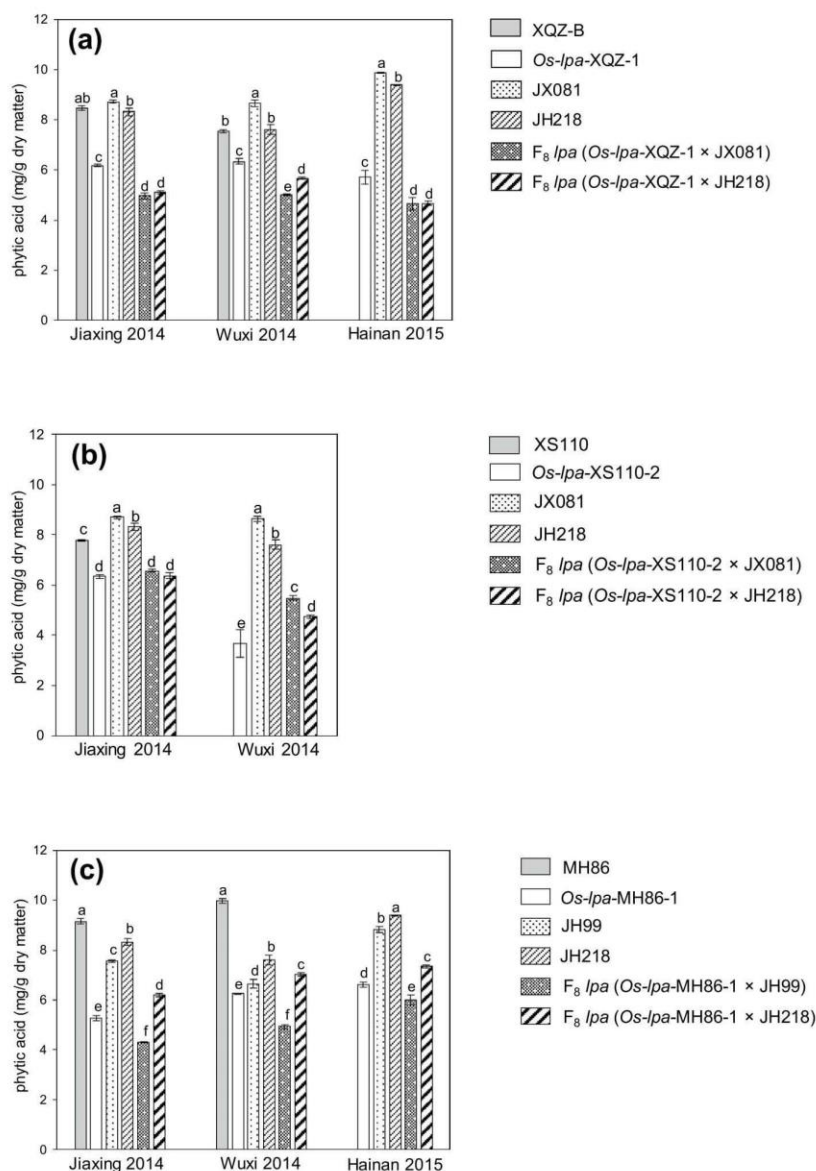
nature of the mutation into account, it is comprehensible that the impact of factors such as environment and crossing parent on the PA content in progenies of *Os-lpa*-XQZ-1 is rather low. In contrast, the reduced PA content in *Os-lpa*-XS110-2 is due to the mutation of the *OsMRP5* gene which may affect tissue compartmentation of PA and/or its transport and storage to the vacuole [16, 24]. This complex mechanism going beyond a sole direct disruption of the biosynthetic flow to PA is expected to be more susceptible to the investigated impact factors. This complexity is even further increased in *Os-lpa*-MH86-1. The disruption of the putative sulfate transporter gene *OsSULTR3;3* is thought to

directly or indirectly play a role in the cross-talk between sulfate and phosphate homeostasis and/or signaling [17]. Considering this scenario, the influence of factors such as environment and crossing parent on the PA contents of *lpa* progenies is highly probable. The differences in complexity of the mechanisms underlying the PA reduction are also reflected by the significant metabolic changes of the concentrations of a broad spectrum of low-molecular-weight rice constituents and up- and down-regulations of gene expressions involved in PA metabolism, as well as phosphate and sulfate homeostasis reported in *Os-lpa*-MH86-1 [17]. On the other hand, in *Os-lpa*-XS110-2 the reduction of the PA

content was accompanied by only a few metabolic changes compared with the wild-type XS110 [22].

For *Os-lpa*-XS110-2 and *Os-lpa*-MH86-1 both the PA content of the crossing parent and the environmental conditions are significant contributors (Table 4). The comparison of the PA contents of the homozygous *lpa* progenies (Table 3) and of the progenitor *lpa* mutants (Table 1) from the same field trials revealed that for *Os-lpa*-XS110-2 and *Os-lpa*-MH86-1 the contribution of the environment as a source of variation is already imprinted in the type of mutant; the subsequent cross-breeding with a commercial cultivar only results in slight modulations of this effect. The observed influence of the type of mutation on the PA contents is in agreement with the reported impact on the lysophospholipid-phosphorus (LPL-P) contents of *lpa* rice grain [25]. Defects in sulfate transporter (*OsSULTR3;3*) and *myo*-inositol kinase (*OsMIK*) genes led to pronounced decreases of LPL-P contents in rice grain, whereas defective *OsMRP5* and *Os2-PGK* genes resulted in significant increases of individual LPL components. The effect of the *Os2-PGK* gene on the LPL accumulation was also demonstrated for breeding lines derived from crosses of the KBNT-*lpa* mutant with JH218 [25].

From a breeder's point of view, it is essential that, independent from individual fluctuations depending on the underlying mutation-type, cross-breeding of *lpa* mutants with commercial cultivars does not compromise the intended *lpa* trait, i.e. the reduction of the PA content compared to the original wild-type. The box plots illustrated in Fig. 3 demonstrate that this important prerequisite is fulfilled. For all three mutation-types, the investigated homozygous *lpa* mutant progenies displayed significantly lower PA contents than the original wild-types subjected to the mutation, independently from cross-breeding steps, crossing parents or environmental conditions. This is a key result from an agronomic point of view demonstrating that a major requirement

for the commercial production of elite *lpa* cultivars via cross and selection breeding is being met.

Results reported for soybeans indicated that backcrossing is successful for developing *lpa* lines with normal field emergence [12]. The low number of plants resulting from the field trials performed in this study did not allow an appropriate statistical assessment of the agronomic performance of the *lpa* mutant progenies. The limited data available indicated that the cross-breeding of the *lpa* rice mutants had no impact on thousand grain weights and seed setting ratios (Tables S2 and S3). These data should, however, only be considered as preliminary. Larger field trials yielding more plant materials would be required to elaborate robust data and to draw valid conclusions.

Notwithstanding this pending final evaluation of the agronomic performance of the generated *lpa* progenies, the elaborated data are valuable and promising from a breeder's point of view. They demonstrate that crossing of *lpa* rice mutants with commercial rice cultivars may influence the PA contents in the resulting *lpa* progenies. The extent of the changes depends on the mutant-type. Of course, the mutants investigated in this study do not cover the complete spectrum of presently known approaches applied to achieve the *lpa* trait, and the results obtained for the three exemplary mutant-types should not be unconditionally extrapolated to other applications. Nevertheless, the data revealed that different scenarios may be expected: On the one hand, there are *lpa* mutants such as *Os-lpa*-XQZ-1, for which a mutation involved in the direct biosynthetic pathway leading to PA results in a stable PA reduction that is rather unsusceptible to changes upon cross-breeding as well as to environmental impact. On the other hand, *Os-lpa*-MH86-1 represents an example of a mutant for which the metabolic complexity underlying the *lpa* trait results in pronounced variability of the PA contents of *lpa* progenies depending on the crossing parents and the environment.

Table 4 Results of two-way ANOVA and coefficients of variance (CV) obtained for the PA contents of homozygous F_8 *lpa* mutant progenies

Source of variation	CV (%)	Sum of squares	Mean squares	F	p
Os-lpa-XQZ-1	7.1				
PA content of crossing parent		1.90	1.90	111	<0.0001
Environment		0.06	0.03	1.6	0.23
Interaction		0.90	0.04	2.6	0.12
Os-lpa-XS110-2	13.0				
PA content of crossing parent		3.48	3.48	328	<0.0001
Environment		3.25	3.25	306	<0.0001
Interaction		0.04	0.04	3.4	0.10
Os-lpa-MH86-1	18.1				
PA content of crossing parent		8.56	8.56	146	<0.0001
Environment		11.36	5.68	97	<0.0001
Interaction		0.08	0.04	0.6	0.55
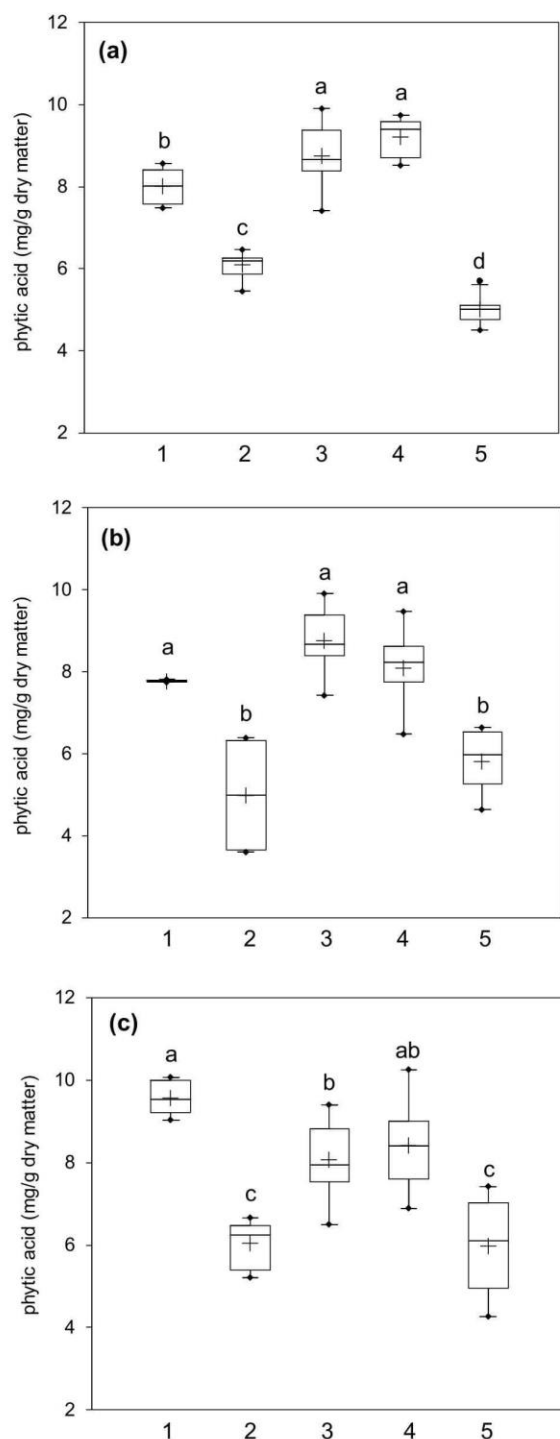
◀ **Fig. 3** Box-and-whisker plots of the PA contents for all rice samples from the field trials Jiaxing 2014, Wuxi 2014 and Hainan 2015. Plus symbol: mean value; line in the box: median value; bottom and top of the box: first quartile (Q1) and third quartile (Q3); upper whisker: top of box + 1.5 × interquartile range (IQR = Q3 − Q1) (maximum value); lower whisker: bottom of box − 1.5 × IQR (minimum value). **a** *Os-lpa*-XQZ-1; **b** *Os-lpa*-XS110-2; **c** *Os-lpa*-MH86-1. Lane 1: original wild-type, 2: progenitor *lpa* mutant, 3: crossing parents, 4: homozygous wild-type progenies, 5: homozygous *lpa* mutant progenies. The data were assessed by one-way ANOVA; different letters indicate significant differences (Tukey's test at $p < 0.05$) in the PA levels of the rice seeds from the same mutation type

in the *atipk-1-1*, 35% in the *atipk2β-1*, and > 95% in the *atipk1-1 atipk2β-1* double mutant [26]. *Lpa* soybean double mutants carrying two *IPK1* mutation targets on chromosome 6 and 14 also showed drastic PA reductions up to − 88% [27]. Therefore, the generation of double mutants seems a useful strategy that might also be applicable to obtain *lpa* rice seeds with trace PA contents.

It is promising that despite the variability of the PA contents of *lpa* progenies observed for the three types of mutants investigated in this study, the *lpa* progenies always exhibited lower PA contents than the original wild-types used for the mutation. This demonstrates that cross-breeding with commercial cultivars can be used as tool to obtain mutants stably expressing the intended *lpa* trait.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

Compliance with ethics requirements This article does not contain any studies with human or animal subjects.

References

1. Raboy V (2003) Myo-inositol-1,2,3,4,5,6-hexa*kis*phosphate. Phytochemistry 64:1033–1043
2. Kumar V, Sinha AK, Makkar HPS, Becker K (2010) Dietary roles of phytate and phytase in human nutrition: a review. Food Chem 120:945–959
3. Humer E, Schwarz C, Schedle K (2015) Phytate in pig and poultry nutrition. J Anim Physiol Anim Nutr 99:605–625
4. Raboy V (2007) Seed phosphorus and the development of low-phytate crops. In: Turner B, RichardsonE LAE, Mullaney J (eds) Inositol phosphates: linking agriculture and the environment, 1st ed. CAB International, Wallingford, pp 111–132
5. Mendoza C, Viteri FE, Lonnerdal B, Young KA, Raboy V, Brown KH (1998) Effect of genetically modified, low-phytic acid maize on absorption of iron from tortillas. Am J Clin Nutr 68:1123–1127
6. Petry N, Egli I, Campion B, Nielsen E, Hurrell R (2013) Genetic reduction of phytate in common bean (*Phaseolus vulgaris* L.) seeds increases iron absorption in young women. J Nutr 143:1219–1224

In the course of the generation of phytate-free *Arabidopsis* seeds through disruption of inositol polyphosphate kinases (*IPK*), seed phytate levels were reduced by 83%

7. Veum TL, Ledoux DR, Bollinger DW, Raboy V, Cook A (2002) Low-phytic acid barley improves calcium and phosphorus utilization and growth performance in growing pigs. J Anim Sci 80:2663–2670

8. Veum TL, Raboy V (2016) Hulled and hull-less barley grains with the genetic trait for low-phytic acid increased the apparent total-tract digestibility of phosphorus and calcium in diets for young swine. J Anim Sci 94:1000–1011

9. FAO (Food and Agriculture Organization of the United Nations) (2017) http://www.fao.org/faostat/en/#data/QC/visualize. Accessed 16 Dec 2017

10. Liu QL, Xu XH, Ren XL, Fu HW, Wu DX, Shu QY (2007) Generation and characterization of low phytic acid germplasm in rice (*Oryza sativa* L.). Theor Appl Genet 114:803–814

11. Zhao HJ, Liu QL, Fu HW, Xu XH, Wu DX, Shu QY (2008) Effect of non-lethal low phytic acid mutations on grain yield and seed viability in rice. Field Crops Res 108:206–211

12. Spear JD, Fehr WR (2007) Genetic improvement of seedling emergence of soybean lines with low phytate. Crop Sci 47:1354–1360

13. Averitt B, Shang C, Rosso L, Qin J, Zhang MC, Rainy KM, Zhang B (2017) Impact of *mips1, lpa1*, and *lpa2* Alleles for low phytic acid content on agronomic, seed quality, and seed composition traits of soybean. Crop Sci 57:2490–2499

14. Wiggins SJ, Smallwood CJ, West DR, Kopsell DA, Sams CE, Pantalone VR (2018) Agronomic performance and seed inorganic phosphorus stability of low-phytate soybean line TN09-239. J Am Oil Chem Soc 95:787–796

15. Zhao HJ, Liu QL, Ren XL, Wu DX, Shu QY (2008) Gene identification and allele-specific marker development for two allelic low phytic acid mutations in rice (*Oryza sativa* L.). Mol Breed 22:603–612

16. Xu XH, Zhao HJ, Liu QL, Frank T, Engel KH, An GH, Shu QY (2009) Mutations of the multi-drug resistance-associated protein ABC transporter gene 5 result in reduction of phytic acid in rice seeds. Theor Appl Genet 119:75–83

17. Zhao HJ, Frank T, Tan YY, Zhou CG, Jabnoune M, Arpat AB, Cui HR, Huang JZ, He ZH, Poirier Y, Engel KH, Shu QY (2016)

18. Disruption of *OsSULTR3;3* reduces phytate and phosphorus concentrations and alters the metabolite profile in rice grains. New Phytol 211:926–939

18. Tan YY, Fu HW, Zhao HJ, Lu S, Fu JJ, Li YF, Cui HR, Shu QY (2013) Functional molecular markers and high-resolution melting curve analysis of low phytic acid mutations for marker-assisted selection in rice. Mol Breed 31:517–528

19. Oates K, De Borba B, Rohrer J (2014) Determination of inositol phosphates in dried distillers grains with solubles. Thermo Fisher Scientific, Sunnyvale, pp 1–10

20. Guideline IHT (2005) Validation of analytical procedures: text and methodology Q2 (R1). In: International conference on harmonization, Geneva, Switzerland, pp 11–12

21. Frank T, Habernegg R, Yuan FJ, Shu QY, Engel KH (2009) Assessment of the contents of phytic acid and divalent cations in *low phytic acid (lpa)* mutants of rice and soybean. J Food Compost Anal 22:278–284

22. Frank T, Meuleye BS, Miller A, Shu QY, Engel KH (2007) Metabolite profiling of two *low phytic acid (lpa)* rice mutants. J Agric Food Chem 55:11011–11019

23. Sparvoli F, Cominelli E (2015) Seed biofortification and phytic acid reduction: a conflict of interest for the plant? Plants 4:728–755

24. Sparvoli F, Cominelli E (2014) Phytate transport by MRPs. In: Plant Abc transporters, pp 19–38

25. Tong C, Chen YL, Tan YY, Liu L, Waters DLE, Rose TJ, Shu QY, Bao JS (2017) Analysis of lysophospholipid content in low phytate rice mutants. J Agric Food Chem 65:5435–5441

26. Stevenson-Paulik J, Bastidas RJ, Chiou ST, Frye RA, York JD (2005) Generation of phytate-free seeds in Arabidopsis through disruption of inositol polyphosphate kinases. Proc Natl Acad Sci USA 102:12612–12617

27. Vincent JA, Stacey M, Stacey G, Bilyeu KD (2015) Phytic acid and inorganic phosphate composition in soybean lines with independent IPK1 mutations. Plant Genome 8:1–10

3.2 Publication II

Zhou, C. G.; Tan, Y. Y.; Goßner, S.; Li, Y. F.; Shu, Q. Y.; Engel, K. H.

Stability of the Metabolite signature resulting from the *OsSULTR3;3* mutation in *low phytic acid* rice (*Oryza sativa* L.) seeds upon cross-breeding.

J. Agric. Food Chem. 2018, 66, 9366-9376.

Reprinted with permission from American Chemical Society (Copyright 2018).

To investigate the potential impact of the cross-breeding of *lpa* mutants with commercial cultivars on the metabolite profiles of the resulting progenies, the previously generated *lpa* rice mutant *Os-lpa*-MH86-1, resulting from the disruption of a putative sulfate transporter gene *OsSULTR3;3*, was crossed with the commercial rice cultivar JH99. A non-targeted metabolite profiling approach enabling the analyses of a wide spectrum of low molecular weight rice constituents was employed to investigate the impact of the cross-breeding step on the metabolite profiles of the progenies, and to assess their stability over several generations. The results of the statistical assessments revealed that the mutation-induced metabolic changes in the *lpa* mutant was not hampered by cross-breeding with a commercial cultivar, and that the mutation-specific metabolite signature was consistently observed in homozygous *lpa* mutant progenies of generations F_4 to F_7. The results underline that cross-breeding of *lpa* mutants with commercial cultivars could be used as a suitable tool to generate *lpa* progenies maintaining metabolic traits induced by mutation.

Candidate's contribution: Performance of the complete sample work-up sequence required for the analysis of the rice samples; independent performance of the GC/MS-based metabolite profiling for all investigated rice samples; implementation of quality control analyses for the metabolite profiling; multivariate and univariate statistical assessments of the elaborated metabolite profiling data; interpretation of the results via pathway analyses; statistical assessment of the agronomic performance data provided by the cooperation partner; writing and revision of the complete manuscript and of the Supporting Information.

Stability of the Metabolite Signature Resulting from the *OsSULTR3;3* Mutation in *Low Phytic Acid* Rice (*Oryza sativa* L.) Seeds upon Cross-breeding

Chenguang Zhou,[†] Yuanyuan Tan,[‡] Sophia Goßner,[†] Youfa Li,[§] Qingyao Shu,[‡] and Karl-Heinz Engel*,[†]

[†]Chair of General Food Technology, Technical University of Munich, Maximus-von-Imhof-Forum 2, D-85354 Freising-Weihenstephan, Germany

[‡]State Key Laboratory of Rice Biology and Zhejiang Provincial Key Laboratory of Plant Germplasm, Institute of Crop Sciences, Zhejiang University, Hangzhou 310058, China

[§]Jiaxing Academy of Agricultural Sciences, Jiaxing 314016, China

Ⓢ *Supporting Information*

ABSTRACT: The *low phytic acid* (*lpa*) rice (*Oryza sativa* L.) mutant *Os-lpa*-MH86-1, resulting from the mutation of the putative sulfate transporter gene *OsSULTR3;3*, was crossed with a commercial rice cultivar. The obtained progenies of generations F_4 to F_7 were subjected to a nontargeted metabolite profiling approach allowing the analyses of a broad spectrum of lipophilic and hydrophilic low-molecular-weight constituents. The metabolite profiles of the homozygous *lpa* progenies were characterized not only by a decreased concentration of phytic acid but also by increased contents of constituents from various classes, such as sugars, sugar alcohols, amino acids, phytosterols, and biogenic amines. Statistical assessments of the data via multivariate and univariate approaches demonstrated that this mutation-induced metabolite signature was nearly unaffected by the cross-breeding step and consistently expressed over several generations. The data demonstrate that even for complex metabolic changes resulting from a mutation, cross-breeding can be employed as a tool to generate progeny rice seeds stably exhibiting the mutation induced traits.

KEYWORDS: *low phytic acid mutant, metabolite profiling, rice (Oryza sativa L.), cross-breeding, OsSULTR3;3*

■ INTRODUCTION

Phytic acid (*myo*-inositol-1,2,3,4,5,6-hexa*kis*phosphate) constitutes the major storage form of phosphorus (P) in cereal grains.[1] Owing to the chelation of divalent cations, it may reduce the bioavailability of nutritionally important minerals such as Zn^{2+} and Fe^{2+}.[2] Additionally, the excretion of undigested phytate from the manure of monogastric animals contributes to P pollution of the environment.[3] Therefore, during the past years several *low phytic acid* (*lpa*) crops have been developed using genetic engineering as well as mutation breeding via chemical mutagenesis and γ-irradiation.[4]

Induced mutations in *lpa* crops have been shown to result not only in the intended decreased levels of phytic acid but also in other metabolic changes. In *lpa* soybean seeds, for example, decreased levels of *myo*-inositol, stachyose, raffinose, and galactosyl cyclitols and increased levels of sucrose compared to the wild-type cultivar have been detected via targeted analysis[5] as well as via metabolite profiling.[6] Application of metabolite profiling to *lpa* rice mutants also revealed consistent changes of the contents of *myo*-inositol, raffinose, galactose, and galactinol compared with the wild-type cultivars.[7]

Recently, it has been demonstrated that the disruption of *OsSULTR3;3*, an ortholog of the sulfate transporter family group 3 gene, resulted not only in a pronounced reduction of the phytic acid content in the *lpa* rice mutant *Os-lpa*-MH86-1 but also in significant changes of the metabolite profile of the *lpa* rice grain

compared with the corresponding wild-type MH86.[8] The mutation-induced metabolite signature comprised altered levels of a broad spectrum of constituents, for example, reduced content of cysteine, increased concentrations of various amino acids (e.g., serine, threonine and isoleucine), organic acids (e.g., citric acid) and other nutritionally relevant compounds, such as γ-aminobutyric acid (GABA).

Lpa mutant crops often show inferior agronomic performance (e.g., lower grain yield and seed viability compared with the wild-types), and therefore, approaches such as cross and selection breeding are applied to minimize these negative effects.[9] Studies with soybean (*Glycine max* L. Merr.) lines showed that the field emergence of *lpa* progenies was significantly improved by crossing parental *lpa* lines with elite cultivars.[10,11]

However, information on the potential impact of the cross-breeding of *lpa* mutants with commercial cultivars on the metabolite profiles of the resulting *lpa* mutant progenies is lacking. Therefore, based on the application of a nontargeted metabolite profiling approach, the objective of this study was to use the *lpa* rice mutant *Os-lpa*-MH86-1 as an example (i) to investigate the impact of cross-breeding with a commercial
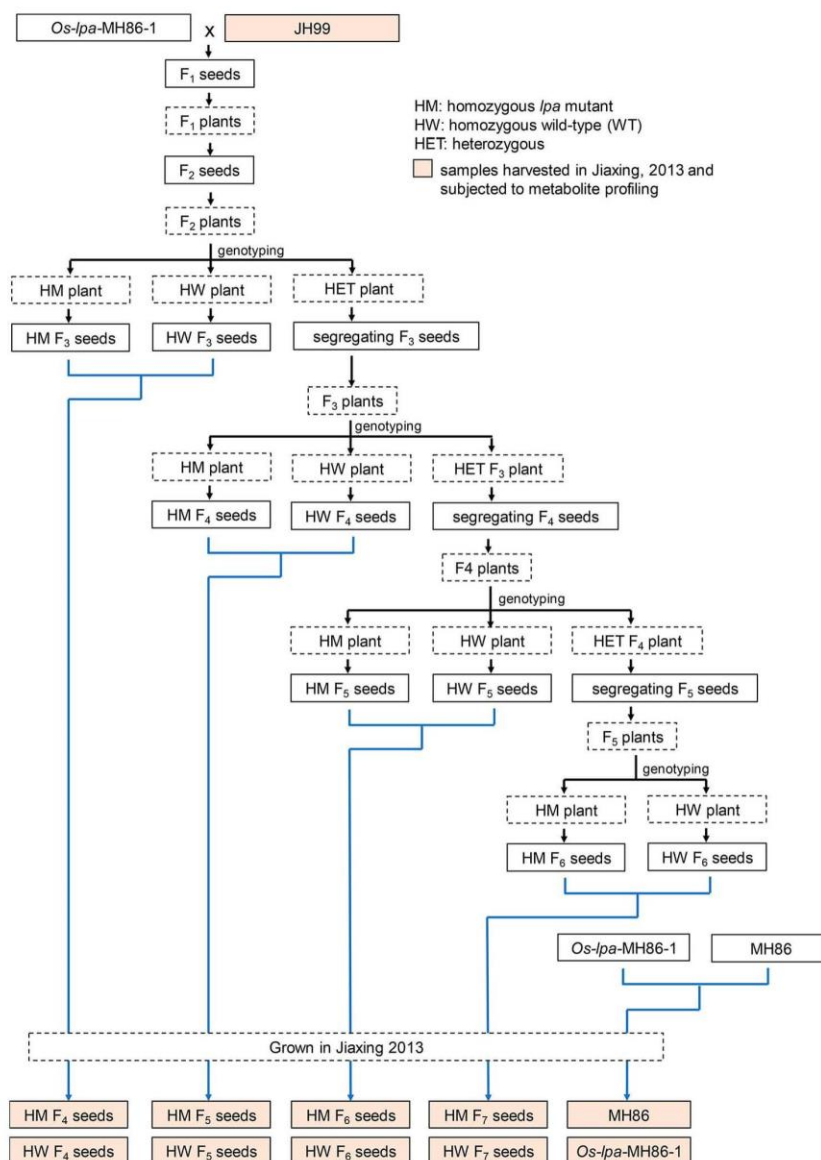
Figure 1. Flowchart of the cross-breeding steps applied to produce the rice materials.

cultivar on the metabolic phenotype of the homozygous *lpa* mutant and (ii) to assess the stability of the mutation-specific metabolite signature in the *lpa* progenies over several generations.

■ MATERIALS AND METHODS

Chemicals. Internal standards tetracosane, 5α-cholestan-3β-ol, phenyl-β-D-glucopyranoside and 4-chloro-L-phenylalanine were purchased from Fluka (Buchs, Switzerland). Reference compounds were supplied by VWR International (Darmstadt, Germany), Fluka (Buchs, Switzerland), Sigma-Aldrich (Steinheim, Germany), and Roth (Karlsruhe, Germany). All other reagents and solvents were obtained from VWR International (Darmstadt, Germany), Sigma-Aldrich (Steinheim, Germany), and Merck (Darmstadt, Germany) as HPLC grade.

Development of Sample Materials. The *lpa* rice mutant *Os-lpa*-MH86-1 has been previously generated from the corresponding wild-type MH86 via γ-irradiation.[12] The mutation of *Os-lpa*-MH86-1 was attributed to a 1-bp deletion of the putative sulfate transporter gene (*OsSULTR3;3*).[8] In the present study, the wild-type line Jiahui 99 (JH99), the restorer line of the hybrid cultivar Jiayou 99, was used to cross with *Os-lpa*-MH86-1 and to generate wild-type and *lpa* progenies, which were used for evaluation of the effect of genetic background on the metabolite signature of the *lpa* mutants. F_2 seeds were bulk-harvested and grown into plants. Two hundred F_2 plants were genotyped for the *lpa* mutation; consequently, F_2 plants were classified into three types: homozygous wild-type, homozygous *lpa* type, and heterozygous progenies (Figure 1). Plants were also segregated for other agronomic traits: flowering time, plant height, and fertility; therefore, 5 F_2 plants each of the homozygous and 10 each of the heterozygous genotype were chosen on the basis of similarity of flowering time, and their F_3 seeds were harvested. While seeds of

38

homozygous F_2 plants were stored for growth until 2013, those from heterozygous F_2 plants were grown into $F_{2:3}$ plots and their F_3 plants were genotyped for the *lpa* mutation and classified into three types as in F_2 (Figure 1). Similar work was performed in F_4–F_6, and seeds harvested from homozygous F_2 to F_5 plants were grown in a common paddy field in 2013 in Jiaxing for the production of sample materials (F_4 to F_7, as well as the two parental lines) for metabolite profiling (Figure 1).

For the genotyping, genomic DNA was extracted from leaf tissues following a modified CTAB method, and high-resolution melting curve analysis was utilized to distinguish among homozygous wild-types, homozygous *lpa* mutants and heterozygous progenies, according to a previously described procedure.[13]

The homozygous F_4 to F_7 seeds, as well as the original wild-type MH86, the progenitor *lpa* mutant *Os-lpa*-MH86-1 and the crossing parent JH99 were harvested at a field trial in Jiaxing, 2013. For each generation of the progenies, seeds from 3 to 4 plants were available for analysis. Owing to the limited amount of the sample materials, for each generation rice seeds from 3 to 4 plants of each phenotype were pooled. After dehulling and grinding with a cyclone mill equipped with a 500 μm sieve (Cyclotec, Fodd, Germany), the flour was immediately freeze-dried for 48 h and then stored at −20 °C until analysis.

Phytic Acid Content Analysis. Extraction and analysis of phytic acid were performed according to a procedure described for dried distillers grains.[14] For each sample, three aliquots (100 mg) of freeze-dried rice flour were weighed into 50 mL polystyrene centrifuge tubes. Twenty milliliters of 0.5 M HCl were added to each aliquot. The extraction of phytic acid was performed under sonication for 20 min after the mixture had been vortexed thoroughly. After centrifugation at 2000*g* for 20 min, the supernatants were filtered through a 0.22 μm polyether sulfone filter disk, and 100 μL were subjected to ion chromatography.

The analysis of phytic acid was performed on a Thermo Scientific Dionex (Dreieich, Germany) ICS-5000 HPIC (high pressure ion chromatography) system equipped with a CarboPac PA100 guard column (4 × 50 mm) and a CarboPac PA100 analytical column (4 × 250 mm) constantly thermostated to 30 °C. The flow rate of the column was 1 mL/min, and a gradient elution was performed by two mobile phases (i.e., deionized water and 0.5 M HCl). The initial proportion of gradient HCl of 5% was increased linearly to 100% within 4 min and then held for 10 min. The flow rate of postcolumn derivatization with ferric nitrate (solution of 0.1% $Fe(NO_3)_3$ in 0.33 M perchloric acid) was 0.4 mL/min, and peak detection was performed at 290 nm. Quantification of phytic acid was based on an external standard calibration curve ($R^2 = 0.9997$); the recovery rate was 98.8%, and the limits of detection and quantification were 0.7 mg/L and 2.2 mg/L, respectively.[15]

Metabolite Profiling. Extraction and fractionation of the rice flour were performed as previously described.[7] Briefly, for each sample, three aliquots (600 mg) of freeze-dried rice flour were weighted into empty cartridges (3 mL volume; Supelco, Munich, Germany). The lipophilic constituents of the flour were eluted with 4 mL of dichloromethane. The polar constituents were eluted with 10 mL of a mixture of methanol and deionized water (80 + 20, v/v). After adding 50 and 100 μL of internal standard solution I (1.5 mg/mL of tetracosane in hexane) and II (0.6 mg/mL of 5α-cholestan-3β-ol in dichloromethane) to the lipophilic extract, respectively, the lipids were transesterified with sodium methylate and separated by solid-phase extraction into fraction I containing fatty acid methyl esters (FAME) and hydrocarbons, and fraction II containing minor lipids (free fatty acids, fatty alcohols, and sterols). Prior to fractionation of the polar constituents, 50 and 100 μL of internal standard solution III (1.6 mg/mL of phenyl-β-D-glucopyranoside in deionized water) and IV (0.8 mg/mL of 4-chloro-L-phenylalanine in deionized water) were added to the polar extract, respectively. Fraction II and IV were silylated with 50 μL of N-methyl-N-(trimethylsilyl)-trifluoroacetamide by heating at 70 °C for 15 min, fraction III with 100 μL trimethylsilylimidazole by heating at 70 °C for 20 min. Selective hydrolysis of silylated derivatives was applied to separate the polar extracts into fraction III (sugars and sugar alcohols) and fraction IV (acids, amino acids and amines). The obtained four

fractions were analyzed by capillary gas chromatography (GC) coupled with a flame ionization detector and a mass spectrometer (MS), respectively, under the previously described conditions.[7]

The identification of metabolites was based on the comparison of retention times and mass spectra with those of reference compounds, with data from the NIST08 mass spectra library and the literature.[16,17] The amounts of identified metabolites from fractions I−IV were expressed as relative peak intensities (i.e., metabolite peak intensity/internal standard peak intensity × 100) based on the internal standards.

Quality Control. Commercially available rice seeds were analyzed regularly together with actual samples as reference material. The results of the quality control analysis were evaluated on the basis of the reproducibility of selected representative compounds from each fraction: C16:0 FAME, C18:0 FAME, and squalene for fraction I; C16:0 FFA, C18:0 FFA, campesterol, and β-sitosterol for fraction II; sucrose and raffinose for fraction III; and serine, aspartic acid, glutamic acid, and citric acid for fraction IV. In accordance with previous studies,[18,19] the data were considered acceptable when the relative standard deviation (RSD) of these selected compounds did not exceed 25%.

Data Processing and Statistical Analysis. Chromatographic data were acquired and integrated by Chrom-Card 2.3 (Thermo Electron). Peaks below noise level were discarded on the basis of a threshold of 1% peak height relative to the internal standards. Standardization of peak heights and retention time alignment were performed by Chrom*pare* 1.1 (http://www.chrompare.com).[20] For data pretreatment, log transformation and Pareto scaling were carried out.[21] Principal component analysis (PCA) and heat map-analysis were conducted using XLSTAT (version 19.5, France). Orthogonal Partial Least Squares−Discriminant Analysis (OPLS−DA) was performed using the web-based tool MetaboAnalyst (version 4.0; http://www.metaboanalyst.ca/faces/home.xhtml). The OPLS−DA model was validated by a 10-fold cross validation and a permutation test (1000 times).[22] In accordance with a guideline on the univariate statistical analysis of metabolomics-derived data,[23] Student's *t*-test ($p < 0.05$) and ANOVA with Tukey's post hoc test ($p < 0.05$) were performed for those metabolites exhibiting a normal distribution and homogeneity of variance. Only for phosphoric acid, which did not fulfill these requirements, the nonparametric Mann−Whitney test ($p < 0.05$) and the Kruskal−Wallis test with Dunn's post hoc test ($p < 0.05$) were performed by XLSTAT.

■ RESULTS

Multivariate Analysis of Rice Seed Metabolite Profiles. Homozygous *lpa* mutant, homozygous wild-type, and heterozygous progenies of F_2−F_6 plants resulting from the cross of the *lpa* mutant *Os-lpa*-MH86-1 with the commercial rice cultivar JH99 were genotyped by high-resolution melting curve analysis.[13] Seeds of homozygous wild-type and homozygous *lpa* mutant progenies of generations F_4−F_7 were obtained from the same field trial (Figure 1) and subjected to metabolite profiling. The previously established GC/MS-based approach[7,24] covered a broad spectrum of lipophilic and hydrophilic low molecular weight rice constituents and resulted in four fractions containing fatty acid methyl esters (FAMEs) and hydrocarbons (fraction I); free fatty acids (FFAs), fatty alcohols, and sterols (fraction II); sugars and sugar alcohols (fraction III); and amino acids, acids, and amines (fraction IV). A total of 279 peaks were detected; 118 metabolites were identified by comparison of retention times and mass spectral data to those of reference compounds and/or data from mass spectra libraries (Tables S1 and S2, Figure S2).

A PCA score plot based on the combined fractions I−IV revealed a clear separation not only of the *lpa* progenitor *Os-lpa*-MH86-1 and the corresponding wild-type MH86, but also a comparable and consistent separation of the homozygous *lpa* mutant and the homozygous wild-type progenies from
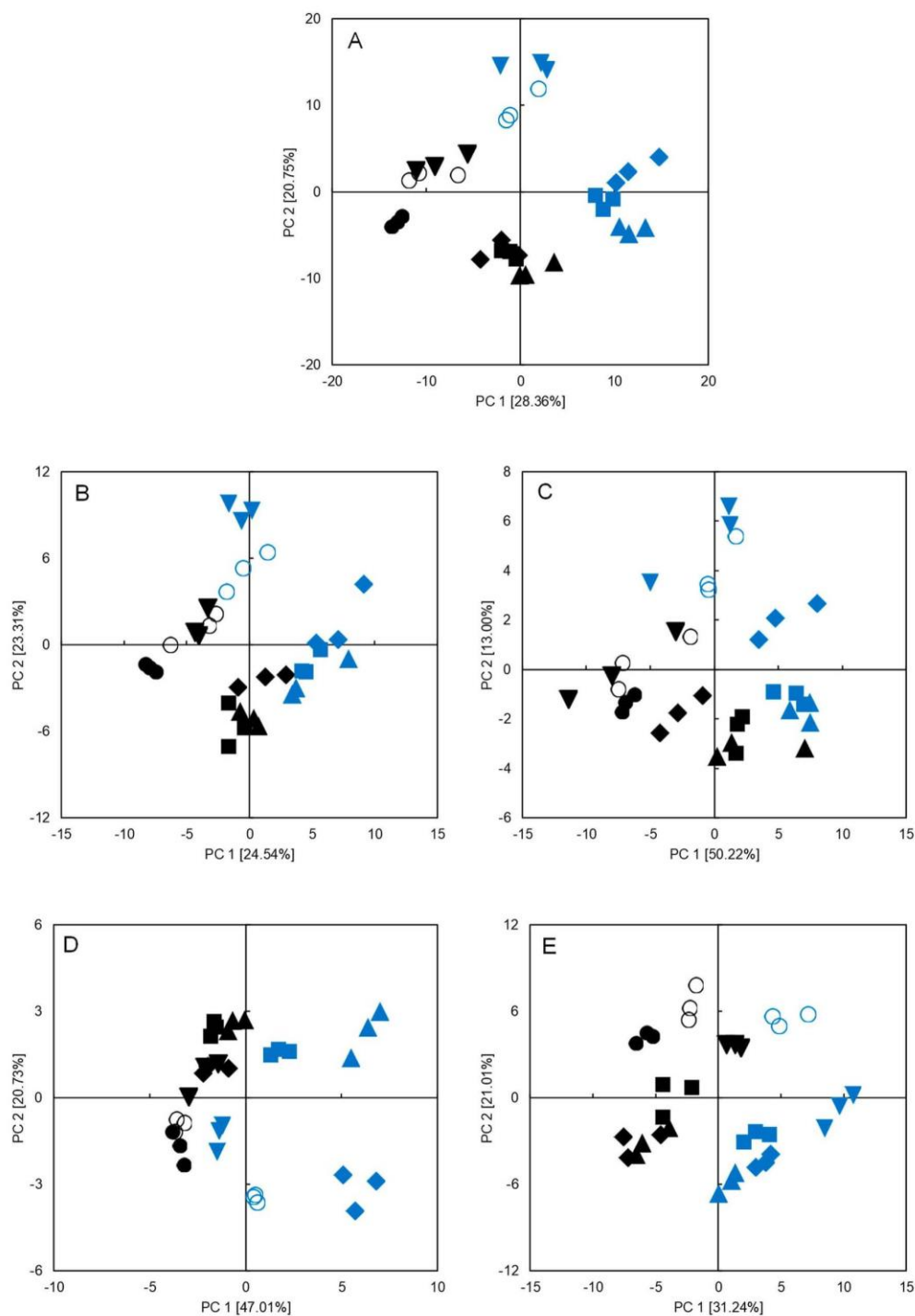
Figure 2. PCA score plots of metabolite profiling data of combined fractions I–IV (A) and single fractions I (B), II (C), III (D), and IV (E) from wild-type (black) and *lpa* mutant (blue) rice seeds: original wild-type MH86 (open circles) and *Os-lpa*-MH86-1 mutant (open circles); crossing parent JH99 (black solid circles); homozygous progenies of generations F_4 (inverted triangles), F_5 (diamonds), F_6 (squares), and F_7 (triangles).

generations F_4 to F_7 (Figure 2A). The homozygous *lpa* mutant progenies of generation F_4 clustered closely to the progenitor *lpa* mutant and the wild-type progenies closely to the original wild-type. In addition, the crossing parent JH99 clustered together with the wild-type MH86 and the F_4 wild-type progenies, and exhibited a clear separation from *Os-lpa*-MH86-1 and F_4 *lpa*

40

Article
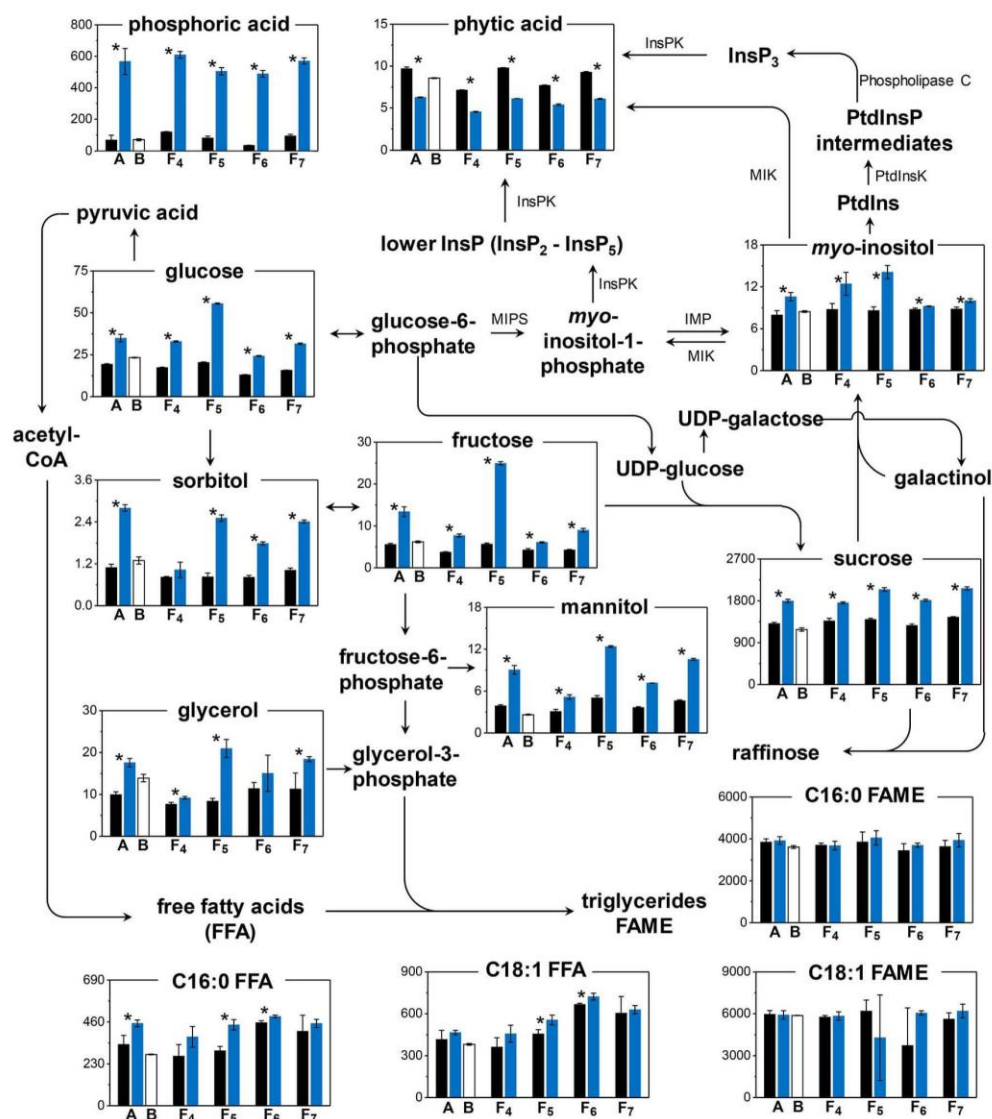
Figure 3. Simplified biosynthetic pathways of selected rice seed metabolites involved in phytic acid, sugar, and lipid metabolism. The bars are displayed in the following order: (A) original wild-type MH86 (black) and progenitor *Os-lpa*-MH86-1 mutant (blue); (B) crossing parent JH99 (white); homozygous wild-type progenies (black) and homozygous *lpa* mutant progenies (blue) of generations F_4, F_5, F_6 and F_7. The phytic acid contents are expressed in mg/g of dry matter. All other metabolites are expressed as relative peak intensities (i.e., metabolite peak intensity/internal standard peak intensity × 100.) Values represent means ± standard deviations resulting from the analysis of three aliquots of freeze-dried flour. Asterisks represent statistically significant differences (Student's *t*-test or Mann–Whitney test, $p < 0.05$) between the wild-type and the corresponding *lpa* mutant. *MIPS*, 1D-*myo*-inositol 3-phosphate synthase; *IMP*, *myo*-inositol monophosphatase; *MIK*, *myo*-inositol kinase; *PtdIns*, phosphatidyl inositol; *PtdInsK*, phosphatidyl inositol phosphate kinase; *PtdInsP*, phosphatidyl inositol phosphate; *InsPK*, inositol phosphate kinase; *UDP*, uridine diphosphate.

progenies. These results indicated that the mutation-specific metabolite signature in *lpa* mutants was nearly unaffected by the cross-breeding step.

For the progenies of generations F_5 to F_7, there were shifts along PC1 and PC2 compared with generation F_4; however, the separation between homozygous *lpa* mutant and homozygous wild-type progenies remained, which demonstrated the stability of the mutation-induced metabolite signature in *lpa* progenies over generations.

PCA score plots of the single fractions are shown in Figure 2B−E. The score plots of fractions III (Figure 2D) and IV (Figure 2E) also showed clear separations of the homozygous wild-type and the homozygous *lpa* progenies. This demonstrated that the constituents contained in these polar fractions are the main contributors to the differentiation of these two phenotypes seen in the score plot of the combined fractions (Figure 2A). In contrast, the lipophilic fractions I (Figure 2B) and II (Figure 2C) showed separations between progenies from generation F_4 and those from generations $F_5−F_7$. This indicated
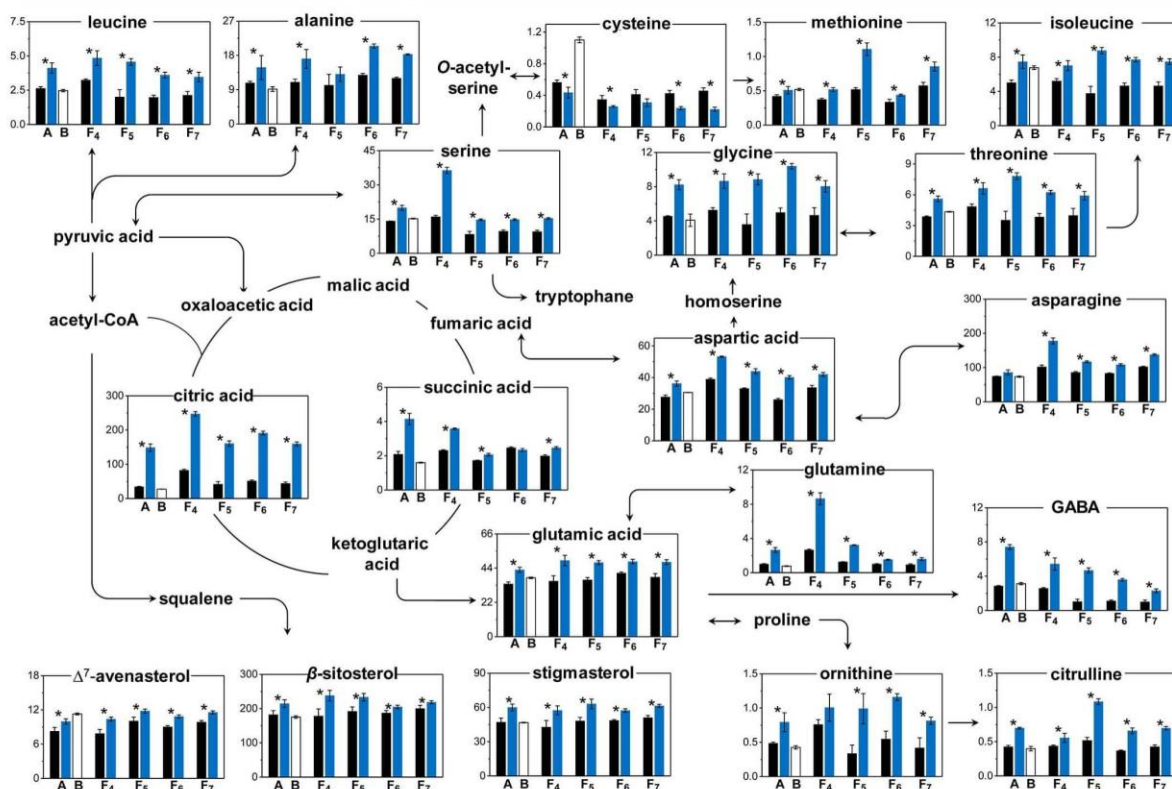
41

Figure 4. Simplified biosynthetic pathways of selected rice seed metabolites involved in amino acid metabolism, phytosterol metabolism, and tricarboxylic acid cycle. The bars are displayed in the following order: (A) original wild-type MH86 (black) and progenitor *Os-lpa*-MH86-1 mutant (blue); (B) crossing parent JH99 (white); homozygous wild-type progenies (black) and homozygous *lpa* mutant progenies (blue) of generations F_4, F_5, F_6, and F_7. The metabolites are expressed as relative peak intensities (i.e., metabolite peak intensity/internal standard peak intensity × 100). Values represent means ± standard deviations resulting from the analysis of three aliquots of freeze-dried flour. Asterisks represent statistically significant differences (Student's *t*-test, $p < 0.05$) between the wild-type and the corresponding *lpa* mutant.

that these fractions are the main contributors to the shifts between generations observed in the score plot of the combined fractions (Figure 2A) for both wild-type and *lpa* progenies.

Metabolic Differences between Homozygous Wild-Type and Homozygous *lpa* Mutant Progenies. In order to determine individual metabolites responsible for the observed clustering (Figure 2A), the respective PCA loading plot (Figure S3) was used. The metabolites were quantitated (relative peak intensities based on the internal standards in each fraction), and mapped in simplified biosynthetic pathways adapted from the KEGG pathway database.[25] Figures 3 and 4 show comparisons of the contents of metabolites in (i) the original wild-type MH86 and the *lpa* mutant *Os-lpa*-MH86-1, (ii) the crossing parent JH99, and (iii) the homozygous wild-type and homozygous *lpa* mutant progenies of generations F_4 to F_7, resulting from the cross-breeding of *Os-lpa*-MH86-1 with JH99.

As shown in Figure 3, the phytic acid content in the homozygous F_4 *lpa* mutant (4.54 ± 0.05 mg/g), determined separately via targeted analysis using HPIC, was significantly lower than that in the corresponding wild-type progeny (7.13 ± 0.03 mg/g). For the following generations F_5 to F_7, despite the variations of phytic acid contents, the homozygous *lpa* mutant progenies from each generation always exhibited significantly lower levels of phytic acid than the corresponding wild-type progenies. In addition, the phytic acid contents of the

homozygous *lpa* mutant progenies of generations F_4 to F_7 were consistently lower than that of the original wild-type MH86.

The metabolite signatures of the homozygous *lpa* mutant progenies resulting from the cross-breeding were characterized not only by this intended reduction of the phytic acid content but also by changes in the levels of a broad spectrum of metabolites from different classes. Statistically significant differences in the levels of metabolites previously reported to discriminate between MH86 and its *lpa* mutant *Os-lpa*-MH86-1[8] were similarly observed for the homozygous wild-type and the homozygous *lpa* mutant progenies of generation F_4 resulting from the crossing of *Os-lpa*-MH86-1 with JH99. These included sugars related to the biosynthesis of phytic acid (*myo*-inositol, glucose, fructose, and sucrose), a broad spectrum of amino acids, organic acids (citric acid, succinic acid), sugar alcohols (sorbitol, mannitol), and biogenic amines (GABA) (Figures 3 and 4).

In addition, the mutation-specific signature was consistently expressed in *lpa* progenies from the generations F_4 to F_7. Although the absolute concentrations of metabolites varied depending on the generations, in nearly all cases, the differences between the homozygous wild-type and the homozygous *lpa* mutant progenies were statistically significant for each generation.

Table 1. Contents of Phytic Acid and Relative Peak Intensities of Selected Metabolites in the Original Wild-Type MH86 and the *lpa* Progenies of Generations F_4, F_5, F_6, and F_7 Obtained from Field Trial Jiaxing, 2013.[a,b,c]

	original wild-type		homozygous *lpa* mutant progenies							
metabolite	MH86		F_4		F_5		F_6		F_7	
phytic acid	9.69 ± 0.2	a	4.54 ± 0.05	d	6.12 ± 0.01	b	5.38 ± 0.11	c	6.08 ± 0.06	b
phosphoric acid	67.6 ± 30.0	b	609.3 ± 21.1	a	504.4 ± 23.3	a	488.2 ± 21.1	a	569.8 ± 19.5	a
glucose	19.3 ± 0.3	e	32.9 ± 0.4	b	55.5 ± 0.4	a	24.2 ± 0.3	d	31.5 ± 0.4	c
sucrose	1319 ± 25.9	c	1782 ± 15.9	b	2158 ± 37.2	a	1859 ± 30.3	b	2219 ± 47.5	a
mannitol	3.9 ± 0.1	e	5.2 ± 0.3	d	12.4 ± 0.1	a	7.2 ± 0.02	c	10.5 ± 0.1	b
glycine	4.5 ± 0.1	c	8.6 ± 0.9	b	8.8 ± 0.6	ab	10.4 ± 0.3	a	8.0 ± 0.7	b
leucine	2.6 ± 0.1	c	4.8 ± 0.5	a	4.6 ± 0.2	a	3.6 ± 0.2	b	3.5 ± 0.4	b
isoleucine	5.0 ± 0.3	c	7.0 ± 0.6	b	8.8 ± 0.4	a	7.7 ± 0.3	ab	7.5 ± 0.3	b
threonine	3.9 ± 0.1	c	6.6 ± 0.6	b	7.8 ± 0.3	a	6.2 ± 0.2	b	5.9 ± 0.4	b
cysteine	0.56 ± 0.03	a	0.26 ± 0.01	bc	0.31 ± 0.05	b	0.24 ± 0.02	bc	0.22 ± 0.03	c
aspartic acid	27.7 ± 1.2	d	53.2 ± 0.4	a	43.9 ± 1.6	b	40.1 ± 1.1	c	42.0 ± 1.2	bc
glutamic acid	33.8 ± 1.3	b	48.9 ± 3.3	a	47.6 ± 1.3	a	48.2 ± 1.5	a	47.9 ± 1.5	a
asparagine	74.7 ± 0.4	d	177.8 ± 8.8	a	116.8 ± 2.8	c	108.2 ± 3.2	c	137.6 ± 2.4	b
ornithine	0.49 ± 0.02	b	1.01 ± 0.20	a	0.99 ± 0.22	a	1.16 ± 0.05	a	0.81 ± 0.05	a
citrulline	0.43 ± 0.02	d	0.56 ± 0.07	c	1.09 ± 0.04	a	0.66 ± 0.04	bc	0.70 ± 0.02	b
GABA	2.9 ± 0.1	c	5.4 ± 0.7	a	4.7 ± 0.3	a	3.6 ± 0.1	b	2.3 ± 0.2	d
citric acid	34.2 ± 1.4	d	247.0 ± 6.6	a	160.6 ± 7.8	c	191.2 ± 5.5	b	159.5 ± 5.8	c
stigmasterol	47.0 ± 3.7	b	57.4 ± 3.9	a	63.1 ± 4.2	a	57.3 ± 1.5	a	61.4 ± 1.3	a

[a]The phytic acid contents are expressed in mg/g of dry matter. All other metabolites are expressed as relative peak intensities, i.e. metabolite peak intensity/internal standard peak intensity ×100. [b]Values represent means ± standard deviations resulting from the analysis of three aliquots of freeze-dried flour. [c]For each metabolite, different letters indicate significant differences (ANOVA with Tukey's post hoc test or Kruskal−Wallis test with Dunn's post hoc test, Benjamini-Hochberg corrected $p < 0.05$) among the original wild-type MH86 and the homozygous *lpa* mutant progenies of generations F_4, F_5, F_6, and F_7.
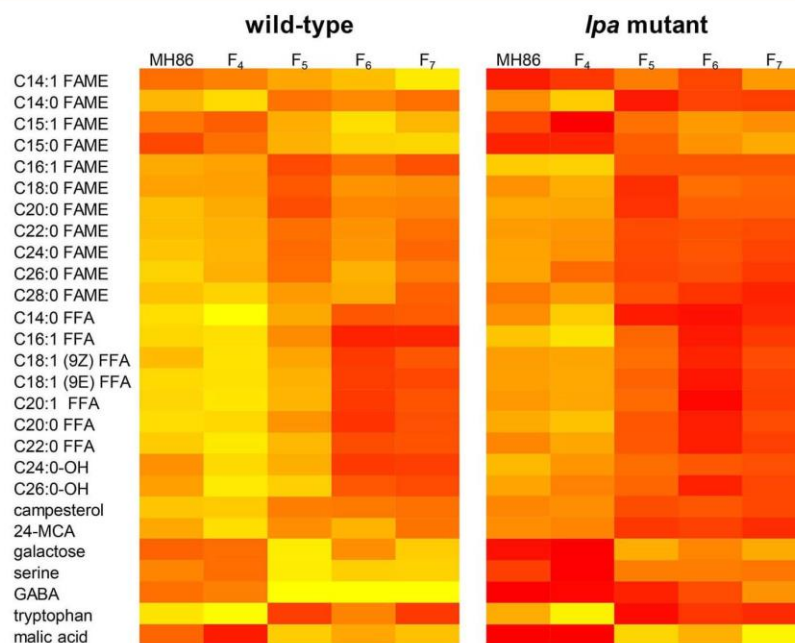
Figure 5. Heat map of metabolites contributing to the metabolic shift among the wild-type and *lpa* mutant rice seeds of different generations.

Univariate comparisons of the metabolite levels were also carried out between the original wild-type MH86 and the homozygous *lpa* mutant progenies of generations F_4–F_7. There were a few sporadic examples where increases observed for metabolites of the *lpa* progenies compared with the original wild-type were not statistically significant for one of the

43

Article

generations. However, in most cases the differences between MH86 and the homozygous *lpa* mutant progenies were consistently expressed over all investigated generations (Table 1).

Metabolic Differences between Progenies Depending on the Generation. Apart from the clear differentiation of the metabolite profiles of the homozygous *lpa* and the homozygous wild-type progenies, there was also a metabolic shift among crossbred progenies of generations from F_4 to F_7 for both wild-types and *lpa* mutants on the PCA score plot along PC 1 and PC 2, respectively (Figure 2A). The extent of the shift was most pronounced from F_4 to F_5 and then decreased with ascending generations from F_5 to F_7. The dynamics of the changes of these metabolites depending on the generations are illustrated by the heat map shown in Figure 5. The metabolic changes over the generations were mainly attributable to lipophilic compounds. Apart from a few exceptions, most of the FAMEs, FFAs, and fatty alcohols showed increasing levels with ascending generations both for homozygous wild-type and *lpa* mutant progenies. This is in agreement with the PCA results which showed similar metabolic shifts along ascending generations of progenies both in fractions I and II (Figure 2B,C). On the contrary, polar constituents contributed to a lesser extent to the differentiation according to the generation; in fractions III and IV galactose, malic acid, GABA, and serine exhibited reduced levels, and only the tryptophan levels were increased from the early to the late generations.

Separation of Homozygous Wild-Type and *lpa* Mutant Progenies via Supervised Statistical Assessment. The results of the multivariate analysis via PCA gave a first indication that it is possible to consistently differentiate between wild-type MH86 and homozygous wild-type progenies resulting from the cross-breeding of the *Os-lpa*-MH86-1 mutant with the commercial cultivar JH99, on the one hand, and *Os-lpa*-MH86-1 mutants and the homozygous *lpa* mutant progenies resulting from the cross-breeding step, on the other hand. To confirm this observation, a supervised OPLS-DA was performed, taking into account all presently available metabolite profiling data on MH86 and *Os-lpa*-MH86-1 mutants.

Figure 6A shows the OPLS-DA score plot of metabolite profiling data of MH86 and the *Os-lpa*-MH86-1 mutant generated in this study of MH86 and *lpa* mutants from three other field trials previously performed at other locations and in different years, applying the same standard agronomic practice (fertilization and water management).[8] There were similar separation patterns within the wild-type and the *lpa* mutant groups predefined in the supervised model on the orthogonal component. Despite this variability, a pronounced separation between wild-types and *lpa* mutants could be observed along the predictive component. This separation remained after incorporation of the metabolite profiling data from homozygous wild-type and homozygous *lpa* mutant progenies from different generations obtained after cross-breeding of the *Os-lpa*-MH86-1 mutant with JH99 performed in this study (Figure 6B). For both plots, the high values of R^2 and Q^2 demonstrated the good fitness and the predictability of the employed OPLS-DA models.[26] In addition, the results of permutation tests ($n = 1000$) demonstrated the models to be robust without overfitting (Figure S5).

The OPLS-DA loadings S-plot[22] was employed to reveal the metabolites contributing to the differentiation between wild-type and *lpa* mutants. For the main metabolites responsible for the differentiation seen in Figure 6B, boxplots depicting the
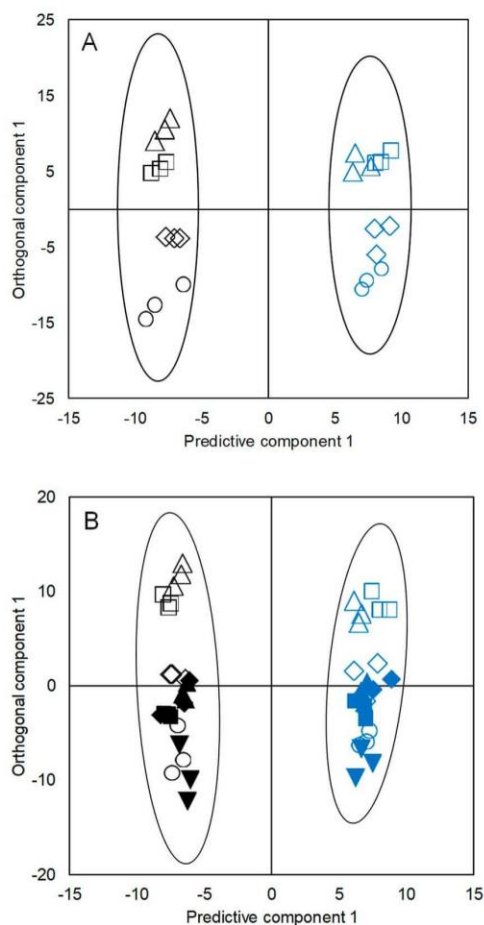
Figure 6. OPLS-DA score plots of metabolite profiling data. (A) Original wild-type MH86 (black) and *Os-lpa*-MH86-1 mutants (blue) from the four field trials Hangzhou, 2011 (open diamonds); Hainan, 2012 (open triangles); Hainan, 2013 (open squares); and Jiaxing, 2013 (open circles). (B) Same sample set as in (A) plus homozygous wild-type progenies (solid black) and homozygous *lpa* mutant progenies (solid blue) of generations F_4 (inverted triangles), F_5 (diamonds), F_6 (squares), and F_7 (triangles). The boundaries of the clusters of wild-type and *lpa* mutant seeds correspond to the 95% Hotelling's T2 ellipses.

statistically significant differences in the mean concentrations (expressed as relative peak intensities) of samples in the wild-type and in the *lpa* mutant OPLS-DA groups are shown in Figure 7.

■ DISCUSSION

The phytic acid contents of homozygous *lpa* mutant progenies resulting from the cross-breeding of the *Os-lpa*-MH86-1 mutant with JH99 were lower than those of the corresponding homozygous wild-type progenies and of the original wild-type MH86 (Figure 3, Table 1). This demonstrated that the *lpa* trait (i.e., the significantly reduced content of phytic acid) remained nearly unaffected by the crossing step and was consistently expressed in *lpa* progenies over generations (F_4 to F_7). This result fulfilled a first major prerequisite for the implementation

44
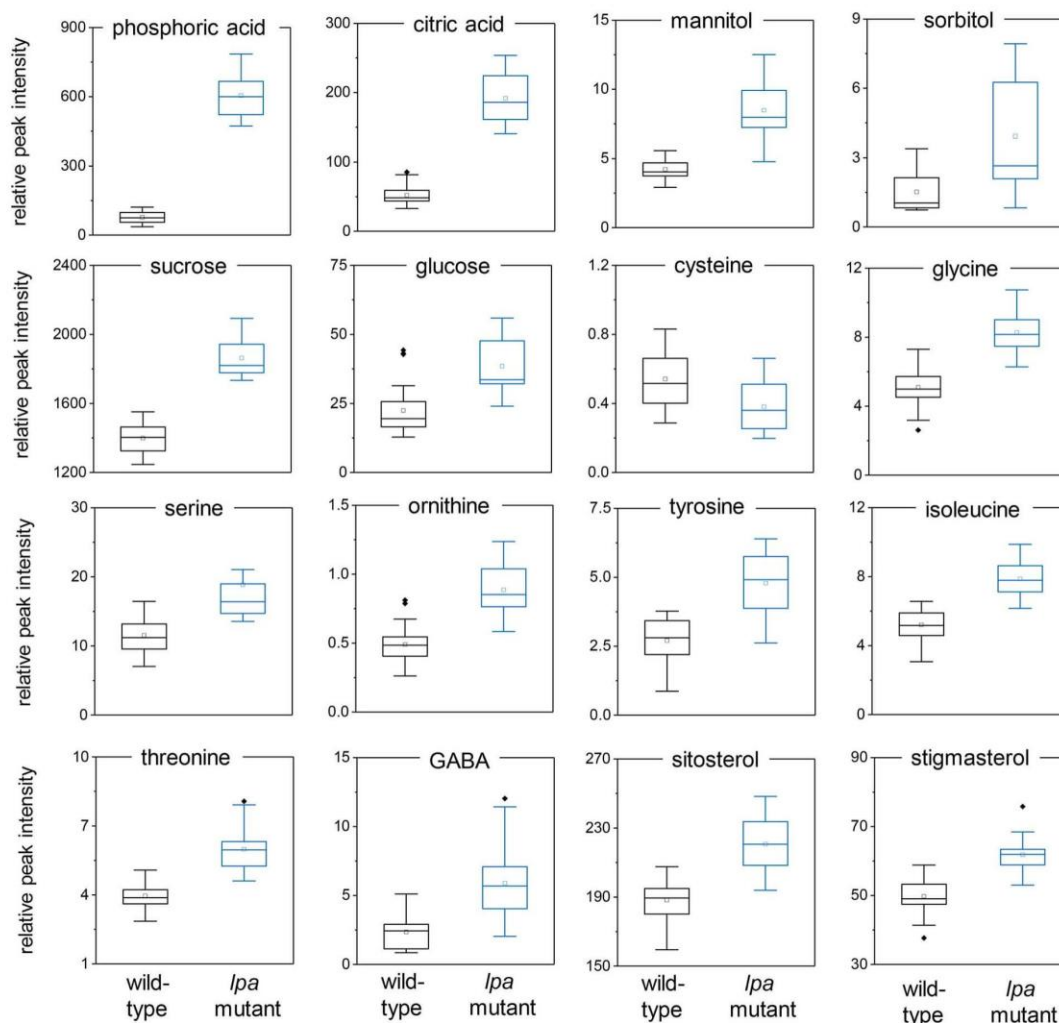
Figure 7. Metabolites contributing to the OPLS-DA separation of wild-type and *lpa* mutant rice seeds. All metabolites exhibited significantly different levels between wild-type and *lpa* mutant rice seeds (Benjamini−Hochberg adjusted-*p* < 0.05). For each metabolite, the relative peak intensity of wild-type was the mean value of MH86 from four field trials (three aliquots per field trial) and homozygous wild-type progenies of generations F_4 to F_7 (three aliquots per generation); the relative peak intensity of *lpa* mutants was the mean value of *Os-lpa*-MH86-1 from four field trials (three aliquots per field trial) and homozygous *lpa* progenies of generations F_4 to F_7 (three aliquots per generation).

of cross and selection breeding with commercial cultivars as part of the procedure to generate *lpa* rice seeds.

The metabolic differences between the *lpa* mutant *Os-lpa*-MH86-1 and the respective wild-type MH86 have been shown to result from the mutation of the putative sulfate transporter gene *OsSULTR3;3*.[8] The mutation affects the expression of a number of genes involved in grain phosphorus and sulfur homeostasis and metabolism. The reduced concentration of cysteine and the increased concentration of its precursor serine suggest a deficiency in sulfide supply. Overall, the *lpa* mutant exhibited a metabolite profile similar to that reported in plants grown under phosphorus deficiency and/or sulfate starvation (e.g., increased concentrations of sugars, sugar alcohols, free amino acids, and GABA).[27,28] The complex metabolite signature has been discussed to be the result of a cross-talk between phosphate and sulfate homeostasis and signaling.[8]

This study demonstrated that these mutation-induced metabolic differences could also be observed between homozygous wild-type and homozygous *lpa* progenies resulting from the cross-breeding of the *Os-lpa*-MH86-1 mutant with JH99. The consistent metabolic differences observed for individual metabolites linked via the biosynthetic pathways shown in Figures 3 and 4 form the molecular basis for the separations seen in the PCA plots (Figure 2). The metabolite signature was shown to be superimposed by metabolic changes depending on the generation. The shift in metabolite profiles between F_4 seeds, on the one hand, and F_5−F_7 seeds, on the other hand, observed in the PCA score plot (Figure 2A) can be explained by the differences in genetic diversity to be expected from the employed breeding scheme (Figure 1). All investigated rice seeds originated from F_2 plants. Homozygous wild-type and homozygous *lpa* mutant F_4 seeds were directly obtained by growing the F_3 seeds from the homozygous wild-type and the

homozygous *lpa* mutant F$_2$ plants. In contrast, F$_5$ to F$_7$ seeds were obtained by growing the segregating F$_3$ seeds of the heterozygous F$_2$ plant. The subsequent process then included two further stages of growing the segregating seeds of F$_4$ and F$_5$ plants, respectively. The resulting metabolic differences between progenies of generations F$_4$ and F$_5$–F$_7$, respectively, were mainly attributable to changes in the concentrations of lipophilic constituents. They occurred in both the homozygous wild-type and the homozygous *lpa* progenies; they were therefore reflected in a parallel translation of the respective clusters in the PCA score plot (Figure 2A) and thus did not hamper their discrimination. The mechanism underlying this effect on the concentrations of lipophilic metabolites still needs further investigation.

The application of a supervised statistical assessment approach demonstrated that it was possible to assign the progenies resulting from the employed cross-breeding step to the respective wild-type and *lpa* mutant groups. The OPLS-DA score plots reflected variability within these groups regarding environmental impact. For example, the metabolite profiles of the rice samples obtained from the field trial performed in this study (Jiaxing, 2013) clustered together with those from a field trial performed at another location in the temperate zone (Hangzhou, 2011) but were separated from those of two field trials in the tropical zone (Hainan, 2012 and 2013). The OPLS-DA score plots also showed differences between the homozygous wild-type and homozygous *lpa* mutant progenies depending on the generations. However, none of these factors finally interfered with the clear separation of the wild-type and *lpa* mutant groups in the supervised statistical assessment.

The specific mutation-induced metabolite signature expressed in the *lpa* mutant *Os-lpa*-MH86-1 as well as in the homozygous *lpa* progenies encompasses the discriminating metabolites shown in Figure 7. Except for cysteine, all other metabolites exhibited significantly increased levels in *lpa* mutant rice seeds compared with wild-type seeds. It is noteworthy that these metabolites also include some nutritionally relevant compounds (e.g., essential amino acids, phytosterols, and the biogenic amine GABA, a well-known neurotransmitter).

In conclusion, the results demonstrate that both the *lpa* trait and the concomitant metabolic changes resulting from the *OsSULTR3;3* mutation were nearly unaffected by cross-breeding of the *lpa* mutants with a commercial cultivar. Despite variations of the metabolite levels, the mutation-induced metabolite signature in homozygous *lpa* mutants was consistently expressed over several generations.

For the *lpa* mutant *Os-lpa*-MH86-1, the actual mechanism underlying the reduction of the concentration of phytic acid as a result of the disruption of the sulfate transporter gene has not been fully elucidated.[8] Considering the broad spectrum of metabolic alterations observed in the *lpa* mutant compared to the wild-type MH86, it is noteworthy that the resulting complex mutation-induced metabolite signature remained stable upon the employed cross-breeding step. Of course, the results elaborated in this study for the metabolite profiles of progenies obtained upon cross-breeding of the *lpa* mutant *Os-lpa*-MH86-1 with the commercial cultivar JH99 should not be unconditionally extrapolated to any other *lpa* mutant. Depending on the induced genetic disruptions, for each type of mutation specific features of changes in the metabolic network and their susceptibility to a cross-breeding step will have to be taken into account. However, the stability of a rather complex metabolite signature shown in this study is encouraging and

indicates that cross-breeding might be a suitable tool to obtain *lpa* crops retaining the desired metabolic phenotype induced by the mutation.

As shown in Figures 3 and 4, the metabolite profile of the crossing parent JH99 employed in this study was similar to that of the wild-type MH86. Therefore, it might also be of importance to follow the impact of cross-breeding on the mutation-induced metabolite signature with a commercial cultivar showing more pronounced own metabolic characteristics compared with the MH86.

Owing to the limited amount of plant materials available from the field trials, only preliminary assessments of the agronomic parameters could be performed (Table S3). They indicated improved field emergence of the *lpa* progenies; however, larger field trials yielding more plant materials would be required to elaborate robust data and to draw valid conclusions regarding the agronomic performance.

Nevertheless, from a breeder's point of view, the achieved metabolite profiling data are promising. The *lpa* mutant *Os-lpa*-MH86-1 exhibited its favorable potential as useful germplasm resource to be exploited for generating valuable *lpa* rice cultivars. The data indicate that even for complex metabolic changes resulting from a mutation, cross-breeding can be employed to obtain progeny rice seeds stably exhibiting the mutation-induced traits. Further studies with different crossing parents should be performed to substantiate the usefulness of this approach.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jafc.8b03921.

> Compounds identified in fractions I–IV of rice seeds; agronomic parameters of selected rice materials; genotyping results via high-resolution melting (HRM) analysis; GC chromatograms of fractions I–IV of MH86; PCA score and loading plots of metabolite profiling data; and permutation tests of OPLS-DA models (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Tel.: +49 8161 714250; Fax +49 8161 714259; E-mail: k.h. engel@wzw.tum.de.

ORCID ⦿

Karl-Heinz Engel: 0000-0001-8932-5175

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Raboy, V. *myo*-Inositol-1,2,3,4,5,6-hexa*kis*phosphate. *Phytochemistry* **2003**, *64*, 1033–1043.

(2) Kumar, V.; Sinha, A. K.; Makkar, H. P. S.; Becker, K. Dietary roles of phytate and phytase in human nutrition: A review. *Food Chem.* **2010**, *120*, 945–959.

(3) Abdel-Megeed, A.; Tahir, A. Reduction of Phosphorus Pollution from Broilers Waste through Supplementation of Wheat Based Broilers Feed with Phytase. *J. Chem.* **2015**, *2015*, 867014.

(4) Raboy, V. Seed phosphorus and the development of low-phytate crops. In *Inositol phosphates: Linking agriculture and the environment*, 1st ed.; Turner, B. L., Richardson, A. E., Mullaney, E. J., Eds.; CAB International: Wallingford, U.K., 2007; pp 111−132.

(5) Hitz, W. D.; Carlson, T. J.; Kerr, P. S.; Sebastian, S. A. Biochemical and molecular characterization of a mutation that confers a decreased raffinosaccharide and phytic acid phenotype on soybean seeds. *Plant Physiol.* **2002**, *128*, 650−660.

(6) Frank, T.; Nörenberg, S.; Engel, K. H. Metabolite Profiling of Two Novel Low Phytic Acid (*lpa*) Soybean Mutants. *J. Agric. Food Chem.* **2009**, *57*, 6408−6416.

(7) Frank, T.; Meuleye, B. S.; Miller, A.; Shu, Q. Y.; Engel, K. H. Metabolite profiling of two low phytic acid (*lpa*) rice mutants. *J. Agric. Food Chem.* **2007**, *55*, 11011−11019.

(8) Zhao, H.; Frank, T.; Tan, Y.; Zhou, C.; Jabnoune, M.; Arpat, A. B.; Cui, H.; Huang, J.; He, Z.; Poirier, Y.; Engel, K.-H.; Shu, Q. Disruption of *OsSULTR3;3* reduces phytate and phosphorus concentrations and alters the metabolite profile in rice grains. *New Phytol.* **2016**, *211*, 926−939.

(9) Zhao, H. J.; Liu, Q. L.; Fu, H. W.; Xu, X. H.; Wu, D. X.; Shu, Q. Y. Effect of non-lethal low phytic acid mutations on grain yield and seed viability in rice. *Field Crops Res.* **2008**, *108*, 206−211.

(10) Boehm, J. D.; Walker, F. R.; Bhandari, H. S.; Kopsell, D.; Pantalone, V. R. Seed Inorganic Phosphorus Stability and Agronomic Performance of Two Low-Phytate Soybean Lines Evaluated across Six Southeastern US Environments. *Crop Sci.* **2017**, *57*, 2555−2563.

(11) Spear, J. D.; Fehr, W. R. Genetic improvement of seedling emergence of soybean lines with low phytate. *Crop Sci.* **2007**, *47*, 1354−1360.

(12) Liu, Q. L.; Xu, X. H.; Ren, X. L.; Fu, H. W.; Wu, D. X.; Shu, Q. Y. Generation and characterization of low phytic acid germplasm in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* **2007**, *114*, 803−814.

(13) Tan, Y. Y.; Fu, H. W.; Zhao, H. J.; Lu, S.; Fu, J. J.; Li, Y. F.; Cui, H. R.; Shu, Q. Y. Functional molecular markers and high-resolution melting curve analysis of low phytic acid mutations for marker-assisted selection in rice. *Mol. Breed.* **2013**, *31*, 517−528.

(14) Oates, K.; De Borba, B.; Rohrer, J. *Determination of Inositol Phosphates in Dried Distillers Grains with Solubles*; Thermo Fisher Scientific: Sunnyvale, CA, 2014; pp 1−10.

(15) Guideline, I. H. T. *Validation of analytical procedures: text and methodology Q2 (R1)*; International Conference on Harmonization: Geneva, Switzerland, 2005; pp 11−12.

(16) Kamal-Eldin, A.; Appelqvist, L. Å.; Yousif, G.; Iskander, G. M. Seed lipids of *Sesamum indicum* and related wild species in Sudan. The sterols. *J. Sci. Food Agric.* **1992**, *59*, 327−334.

(17) Xu, Z. M.; Godber, J. S. Purification and identification of components of gamma-Oryzanol in rice bran oil. *J. Agric. Food Chem.* **1999**, *47*, 2724−2728.

(18) Wenzel, A.; Frank, T.; Reichenberger, G.; Herz, M.; Engel, K.-H. Impact of induced drought stress on the metabolite profiles of barley grain. *Metabolomics* **2015**, *11*, 454−467.

(19) Mastrangelo, A.; Ferrarini, A.; Rey-Stolle, F.; Garcia, A.; Barbas, C. From sample treatment to biomarker discovery: A tutorial for untargeted metabolomics based on GC-(EI)-Q-MS. *Anal. Chim. Acta* **2015**, *900*, 21−35.

(20) Frenzel, T.; Miller, A.; Engel, K. H. A methodology for automated comparative analysis of metabolite profiling data. *Eur. Food Res. Technol.* **2003**, *216*, 335−342.

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

(22) Chong, J.; Soufan, O.; Li, C.; Caraus, I.; Li, S.; Bourque, G.; Wishart, D. S.; Xia, J. MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* **2018**, *46*, W486−W494.

(23) Vinaixa, M.; Samino, S.; Saez, I.; Duran, J.; Guinovart, J. J.; Yanes, O. A Guideline to Univariate Statistical Analysis for LC/MS-Based Untargeted Metabolomics-Derived Data. *Metabolites* **2012**, *2*, 775.

(24) Frenzel, T.; Miller, A.; Engel, K. H. Metabolite profiling - A Fractionation method for analysis of major and minor compounds in rice grains. *Cereal Chem.* **2002**, *79*, 215−221.

(25) Kanehisa, M.; Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **2000**, *28*, 27.

(26) Worley, B.; Powers, R. Multivariate Analysis in Metabolomics. *Curr. Metabolomics* **2013**, *1*, 92−107.

(27) Hammond, J. P.; White, P. J. Sucrose transport in the phloem: integrating root responses to phosphorus starvation. *J. Exp. Bot.* **2008**, *59*, 93−109.

(28) Nikiforova, J. V.; Bielecka, M.; Gakière, B.; Krueger, S.; Rinder, J.; Kempa, S.; Morcuende, R.; Scheible, W.-R.; Hesse, H.; Hoefgen, R. Effect of sulfur availability on the integrity of amino acid biosynthesis in plants. *Amino Acids* **2006**, *30*, 173−183.

3.3 Publication III

Zhou, C. G.; Tan, Y. Y.; Goßner, S.; Li, Y. F.; Shu, Q. Y.; Engel, K. H.

Impact of crossing parent and environment on the metabolite profiles of progenies generated from a *low phytic acid* rice (*Oryza sativa* L.) Mutant.

J. Agric. Food Chem. 2019, 67, 2396-2407

Reprinted with permission from American Chemical Society (Copyright 2019).

It is well known that metabolite profiles of cereal grains can be influenced by both genetics and environment. To investigate these phenomenon, the *lpa* rice mutant *Os-lpa*-MH86-1, which has been reported to exhibit complex metabolic alterations induced by the *OsSULTR3;3* mutation, was crossed with two commercial rice cultivars JH99 (*indica*) and JH218 (*japonica*) with different genetic backgrounds. In addition, the field trials for the resulting progenies of generation F_8 from the two crosses were performed at three growing locations, two of them in the subtropical and one in the tropical region. The crossbred progenies as well as the progenitor *lpa* mutant, the two crossing parents and the original wild-type were all subjected to a GC/MS metabolite profiling, followed by multivariate and univariate statistical analyses. The results revealed a strong impact of environment on the metabolite profiles of rice seeds. However, for each individual field trial, despite an influence of the crossing parents on the lipid profiles of the generated *lpa* progenies, the mutation-induced metabolite signature was consistently expressed in the homozygous *lpa* mutant progenies for both crosses. The elaborated molecular data are promising from a breeder's point of view. They demonstrate that independent from the subspecies and/or cultivar of the crossing parents and independent from the environmental influence, cross and selection breeding could be a useful strategy to produce *lpa* rice seeds stably exhibiting the mutation-induced metabolic traits.

Candidate's contribution: Performance of the complete sample work-up sequence required for the analysis of the rice samples; independent performance of the GC/MS-based metabolite profiling for all investigated rice samples; implementation of quality control analyses for the metabolite profiling; multivariate and univariate statistical assessments of the elaborated metabolite profiling data; interpretation of the results via pathway analyses; writing and revision of the complete manuscript and of the Supporting Information

Impact of Crossing Parent and Environment on the Metabolite Profiles of Progenies Generated from a *Low Phytic Acid* Rice (*Oryza sativa* L.) Mutant

Chenguang Zhou,[†] Yuanyuan Tan,[‡] Sophia Goßner,[†] Youfa Li,[§] Qingyao Shu,[‡] and Karl-Heinz Engel*,[†]

[†]Chair of General Food Technology, Technical University of Munich, Maximus-von-Imhof-Forum 2, D-85354 Freising-Weihenstephan, Germany

[‡]State Key Laboratory of Rice Biology and Zhejiang Provincial Key Laboratory of Plant Germplasm, Institute of Crop Sciences, Zhejiang University, Hangzhou 310058, P. R. China

[§]Jiaxing Academy of Agricultural Sciences, Jiaxing 314016, P. R. China

Ⓢ *Supporting Information*

ABSTRACT: The *low phytic acid* (*lpa*) rice mutant *Os-lpa*-MH86-1, exhibiting a mutation-induced metabolite signature (i.e., increased levels of sugars, sugar alcohols, amino acids, phytosterols, and biogenic amines), was crossed with two commercial wild-type cultivars. The resulting progenies of generation F_8 harvested at three independent field trials were subjected to a GC/MS-based metabolite profiling approach. Statistical assessments via multivariate and univariate analyses demonstrated that the environment had a strong impact on the metabolite profiles of the resulting progenies. In addition, the metabolites of homozygous *lpa* progenies were significantly influenced by the lipid profiles of the wild-type cultivars employed as the crossing parents. However, for each individual field trial, both the *lpa* trait and the mutation-specific metabolite signature were consistently expressed in the homozygous *lpa* mutant progenies of the two crosses. The data underline that cross-breeding can be employed as a tool to generate *lpa* progeny rice seeds stably exhibiting the mutation-induced metabolic traits.

KEYWORDS: *low phytic acid rice (Oryza sativa L.) mutant, metabolite profiling, cross-breeding, crossing parent, OsSULTR3;3*

■ INTRODUCTION

Phosphorus (P) is an important element for plant growth and development. In cereal grains and legume seeds, up to 85% of P is stored in the form of phytic acid (PA, *myo*-inositol-1,2,3,4,5,6-hexa*kis*phosphate or InsP₆).[1] PA can form complexes with mineral cations such as Zn^{2+}, Fe^{2+}, and Ca^{2+}. Thus, it may adversely affect the absorption and bioavailability of these minerals for human nutrition.[2] In addition, undigested phytate from the manure of monogastric animals contributes to water eutrophication in the environment.[3] In order to minimize these disadvantageous effects, genetic engineering and mutation breeding have been applied to develop *low phytic acid* (*lpa*) crops comprising maize, barley, rice, wheat, and soybean.[4] However, *lpa* mutant crops often exhibit inferior agronomic traits, e.g., low germination rate, decreased field emergence, and reduced grain weight. Therefore, cross and selection breeding of *lpa* mutants with commercial cultivars has been applied to improve the agronomic performance of *lpa* mutants.[5−7]

Recently, it has been demonstrated that after cross and selection breeding of the *lpa* rice mutant *Os-lpa*-MH86-1 with the commercial cultivar JH99, the *lpa* trait as well as the *OsSULTR3;3* mutation-specific metabolite signature, i.e., increased levels of a broad spectrum of low-molecular-weight metabolites (sugars, sugar alcohols, amino acids, organic acids, biogenic amines, and phytosterols), were nearly unaffected and consistently expressed in homozygous *lpa* progenies over

several generations (F_4 to F_7).[8] However, the crossing had been performed with only one commercial cultivar, and the progenies had been grown at a single field trial.

It is well-known that metabolite profiles of cereal grains can be influenced by both genetics and environment, as shown for rice,[9] maize,[10] and barley.[11,12] The previous study has demonstrated that the metabolite profiles of MH86 (the original wild-type of the *lpa* mutant *Os-lpa*-MH86-1) and the crossing parent JH99 were similar.[8] Therefore, another commercial wild-type cultivar (JH218) expected to exhibit a distinct metabolite profile compared to MH86 and JH99 was selected as additional crossing parent. In addition, the field trials for the resulting progenies were extended to three growing locations, two of them in the subtropical and one in the tropical region.

Rice seeds of the crossbred progenies as well as the original wild-type MH86, the progenitor *lpa* mutant *Os-lpa*-MH86-1, and the crossing parents JH99 and JH218 harvested at the three field trials were subjected to a GC/MS-based metabolite profiling approach. On the basis of multivariate and univariate statistical analyses of the resulting data, the objectives of the study were: (i) to investigate the impact of the crossing parents
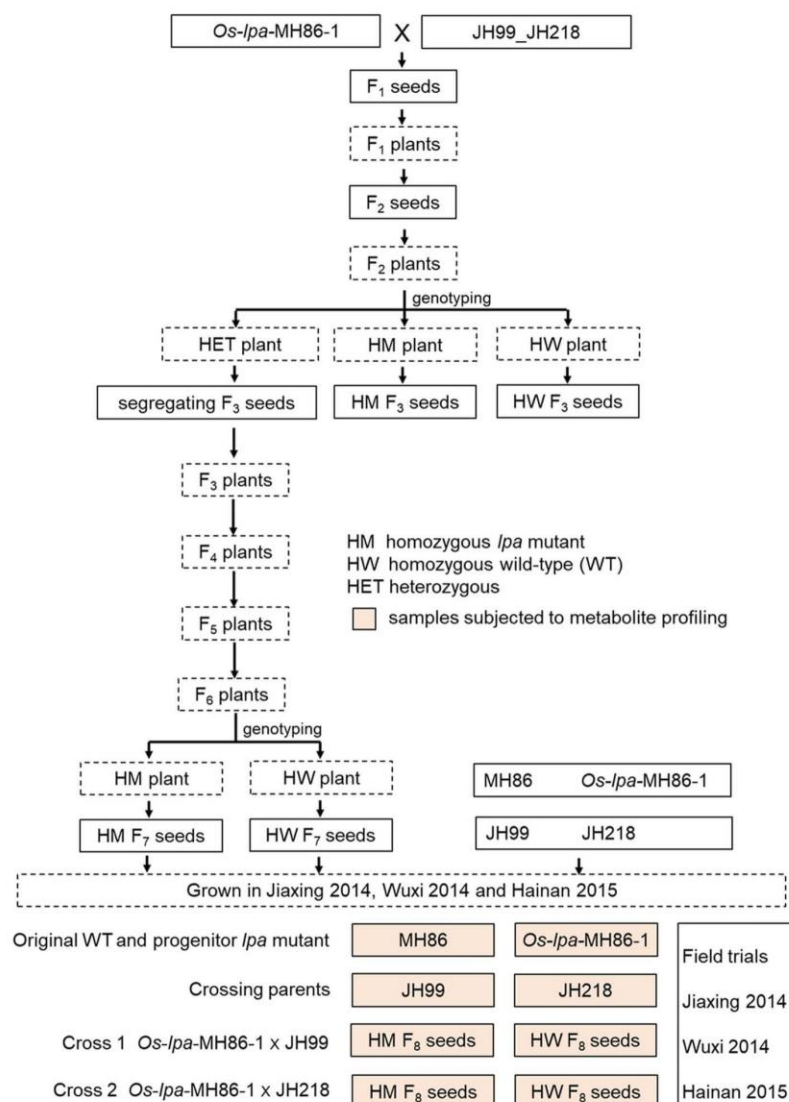
Figure 1. Overview on the cross-breeding steps and the investigated rice samples.

on the metabolite profiles of the generated progenies, and (ii) to assess the stability of the *OsSULTR3;3* mutation-induced metabolite signature of the homozygous *lpa* progenies in the light of environmental influence.

■ **MATERIALS AND METHODS**

Chemicals. Reference compounds were purchased from Sigma-Aldrich (Steinheim, Germany), Roth (Karlsruhe, Germany), VWR International (Darmstadt, Germany), and Fluka (Buchs, Switzerland). Internal standards tetracosane, 5α-cholestane-3β-ol, phenyl-β-D-glucopyranoside and 4-chloro-L-phenylalanine were obtained from Fluka (Buchs, Switzerland). All other solvents and reagents were supplied as HPLC grade by Sigma-Aldrich (Steinheim, Germany), VWR International (Darmstadt, Germany), and Merck (Darmstadt, Germany).

Development of Sample Materials. The wild-type MH86 had been subjected to γ-irradiation to obtain the progenitor *lpa* mutant *Os-lpa*-MH86-1 as previously described.[13] The commercial rice

cultivars JH99 (*indica*) and JH218 (*japonica*), showing yields of 8−9 tons/ha and meeting the Chinese national quality standards, were employed as crossing parents.

The cross and selection breeding steps leading to the crosses *Os-lpa*-MH86-1 × JH99 and *Os-lpa*-MH86-1 × JH218 were performed as previously described.[8] Briefly, F_2 seeds were bulk-harvested and grown into F_2 plants. High-resolution melting curve (HRM) analysis was utilized to genotype the F_2 plants into homozygous wild-type, homozygous *lpa* mutant and heterozygous progenies.[14] Plants were also segregated for other agronomic traits: flowering time, plant height, and fertility; to this end, five homozygous and ten heterozygous F_2 plants, all with reasonable fertility (>70% seed-set), were selected on the basis of similarity of flowering time, and their F_3 seeds were harvested. The same was done for F_3 to F_7 plants. The homozygous wild-type and the homozygous *lpa* mutant progeny rice seeds of generation $F_7$ of the two crosses, as well as the original wild-type MH86, the progenitor *lpa* mutant *Os-lpa*-MH86-1 and the crossing parents JH99 and JH218 were grown at three independent field trials: Jiaxing 2014, Wuxi 2014, and Hainan 2015 (Figure 1).

Due to an experimental error, no material of MH86 from Hainan 2015 was available. For each field trial, homozygous wild-type and homozygous *lpa* progeny rice seeds from 3 to 4 plants were obtained. Owing to this limited amount of sample materials, for each cross rice seeds of the same phenotype were pooled. Rough rice grains were hulled and ground into flour by a cyclone mill with a 500 μm sieve (Foss, Rellingen, Germany). The milled rice flour was freeze-dried for 48 h and stored at −20 °C until analysis.

Phytic Acid Analysis. After acidic extraction, the analysis of phytic acid was performed via high-pressure ion chromatography using an HCl gradient and post column derivatization with ferric nitrite and UV detection, as previously described.[8]

Metabolite Profiling. The metabolites in freeze-dried rice flour were extracted and fractionated in accordance with previously described procedures.[15] Briefly, six aliquots (600 mg) of rice flour were weighted into empty cartridges with 3 mL volume (Supelco, Munich, Germany) which were connected to a vacuum manifold. The lipophilic components of the flour were eluted with 4 mL of dichloromethane within 15 min, the polar metabolites with 10 mL of a mixture of methanol and deionized water (80 + 20, v/v) within 30 min. Fifty microliters of internal standard solution I (1.5 mg/mL of tetracosane in hexane) and 100 μL of internal standard solution II (0.6 mg/mL of 5α-cholestane-3β-ol in dichloromethane) were added to 4 mL of the lipophilic extract. After transesterification with sodium methylate (30% solution in methanol), the lipids were separated by solid-phase extraction (SPE) cartridges into fraction I (fatty acid methyl esters and hydrocarbons) and fraction II (free fatty acids, fatty alcohols, and sterols). For the polar extract, 50 μL of internal standard solutions III (1.6 mg/mL of phenyl-β-D-glucopyranoside in deionized water) and IV (0.8 mg/mL of 4-chloro-L-phenylalanine in deionized water) were added. One microliter of polar extract was silylated with 100 μL of trimethylsilylimidazole by heating at 70 °C for 20 min. Selective hydrolysis of silylated derivatives was applied to obtain fraction III containing sugars and sugar alcohols. After silylation with 50 μL of N-methyl-N-(trimethylsilyl)-trifluoroacetamide and selective hydrolysis of another 4 mL of the polar extract, fraction IV containing acids, amino acids and amines was obtained.

Capillary gas chromatography (GC) was performed using a TraceGC Ultra (Thermo Fisher Scientific, Austin, U.S.A.) equipped with a flame ionization detector (FID) and an ISQ single quadrupole mass spectrometer (MS), under the previously described conditions.[15] The metabolites were identified by comparing the retention times and mass spectra with those of authentic compounds, the data from the NIST08 mass spectra library and the literature.[16,17] The concentrations of identified metabolites from fractions I−IV were expressed as relative peak intensities, i.e. (metabolite peak intensity/internal standard peak intensity) × 100.

Quality Control (QC). QC samples were prepared by pooling equal amounts (2 g) of all rice seed samples investigated in the study. Their analysis was regularly conducted in triplicate together with batches of six actual samples.

Data Processing and Statistical Analysis. Acquisition and integration of chromatographic data were performed by Xcalibur 2.1 (Thermo Fisher Scientific). For each fraction, the detected peaks with peak heights less than 1% relative to the respective internal standards were discarded as noise signals. Alignment of retention time and standardization of peak intensities were performed using the previously developed tool Chrom*pare* 1.1 (http://www.chrompare.com).[18]

For data pretreatment, log transformation and Pareto scaling were applied. Principal component analysis (PCA) and heat-map analysis were performed via software XLSTAT (version 19.5, France). Orthogonal Partial Least Squares−Discriminant Analysis (OPLS−DA) was performed using the web-based tool MetaboAnalyst (version 4.0; https://www.metaboanalyst.ca/faces/home.xhtml).[19] The Variable Importance in Projection values for the Predictive Component (VIP_ pred) of the OPLS-DA model were calculated using the web-based tool Workflow4Metabolomics (version 3.2; https://galaxy.workflow4metabolomics.org).[20] The number of cross-validation segments was set to 10, and the number of permutation testing was

set to 1000 for the OPLS-DA model validation. According to a guideline regarding the univariate statistical analysis of metabolomics-derived data,[21] the identified metabolites showing a normal distribution and homoscedasticity were subjected to Student's *t*-test and ANOVA with Tukey's post hoc test; otherwise the nonparametric Mann−Whitney test and the Kruskal−Wallis test with Dunn's post hoc test were performed.
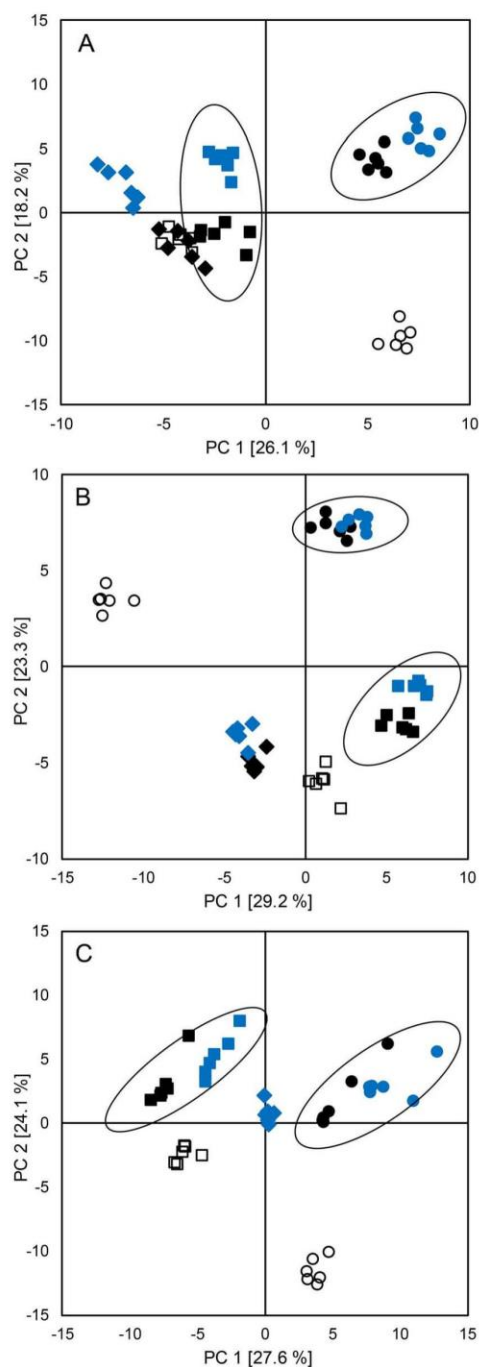
■ RESULTS

Homozygous wild-type and homozygous *lpa* mutant progeny rice seeds of generation F_8 of the crosses *Os-lpa*-MH86-1 × JH99 and *Os-lpa*-MH86-1 × JH218, as well as the original wild-type MH86, the progenitor *lpa* mutant *Os-lpa*-MH86-1 and the two crossing parents JH99 and JH218 were harvested at three independent field trials (Figure 1), and were subjected to a GC/MS-based metabolite profiling approach.[15,22] For each sample, four fractions were obtained: fraction I containing fatty acid methyl esters (FAMEs) and hydrocarbons; fraction II containing free fatty acids (FFAs), fatty alcohols and phytosterols; fraction III containing sugars and sugar alcohols, and fraction IV containing amino acids, acids and amines. The untargeted metabolite profiling resulted in a total of 262 peaks; 118 metabolites were identified (Tables S1 and S2 and Figure S3).

Metabolite Profiles of the Lipophilic Fractions. At first, the metabolites of the lipophilic fractions I−II were assessed via PCA. As shown in Figure 2A for the rice seeds harvested at the field trial Jiaxing 2014, none of the two crosses showed a clear separation of homozygous wild-type and homozygous *lpa* progenies. This is in line with the close clustering of the wild-type MH86 and the corresponding *lpa* mutant *Os-lpa*-MH86-1. Similar results were observed at the field trials Wuxi 2014 (Figure 2B) and Hainan 2015 (Figure 2C).

There were remarkable separations of JH99 (open squares) and JH218 (open circles) at all field trials, indicating distinct lipid profiles of the two crossing parents (Figure 2). This differentiation is imprinted on the corresponding progenies, i.e., there were comparable separations of the progenies of the crosses *Os-lpa*-MH86-1 × JH99 (solid squares) and *Os-lpa*-MH86-1 × JH218 (solid circles). In addition, the progenies of the cross *Os-lpa*-MH86-1 × JH99 clustered closely to the corresponding crossing parent JH99, whereas a pronounced segregation was observed for the progenies of cross *Os-lpa*-MH86-1 × JH218 and the corresponding crossing parent JH218.

To identify the metabolites responsible for the consistently observed discrimination of JH99 and JH218, as well as of the respective progenies, loading plots (Figure S4) of the corresponding PCAs and peak-by-peak comparisons by Chrom*pare*[18] were employed. The discriminating metabolites were quantitated based on relative peak intensities; they are shown in Table 1 for the field trial Jiaxing 2014 as an example (Tables S3 and S4 for the other two field trials). Univariate analysis of each metabolite was performed for the crossing parents JH99 and JH218, and for the homozygous *lpa* progenies of the two crosses.

The crossing parent JH99 exhibited higher levels of several lipophilic metabolites than JH218, and these differences were also reflected in the corresponding *lpa* progenies. For example, as one of the major free fatty acids (FFA) in rice seeds, oleic acid (C18:1 FFA) exhibited statistically significantly higher levels in JH99 and the resulting homozygous *lpa* progenies than those in JH218 and the corresponding *lpa* progenies.

Figure 2. PCA score plots of metabolite profiling data of the lipophilic fractions I−II of wild-type (black) and *lpa* mutant (blue) rice seeds harvested at field trials (A) Jiaxing 2014, (B) Wuxi 2014, and (C) Hainan 2015: original wild-type MH86 (diamonds) and *Os-lpa*-MH86-1 mutant (diamonds); crossing parents JH99 (open squares) and JH218 (open circles); homozygous wild-type and homozygous *lpa* mutant progenies of the crosses *Os-lpa*-MH86-1 × JH99 (solid squares) and *Os-lpa*-MH86-1 × JH218 (solid circles). The encircled clusters (boundaries correspond to the 95% Hotelling's T2 ellipses) include the progenies generated from the respective crossing parent.

Similar results were also observed for a number of other metabolites including FAMEs (e.g., C16:0 and C18:1), FFAs (e.g., C16:0 and C18:1), fatty alcohols (e.g., C24:0-OH and C26:0-OH), and phytosterols (e.g., sitosterol). In contrast, for only a few metabolites, e.g., squalene and phytol, higher concentrations were observed in JH218, and the resulting *lpa* progenies than in JH99 and the corresponding *lpa* progenies.

Metabolite Profiles of the Polar Fractions. For the polar fractions III−IV of the rice seeds harvested at Jiaxing 2014, a clear separation of MH86 and the *Os-lpa*-MH86-1 mutant was observed (Figure 3A). This was in agreement with the previous results for MH86 and the *Os-lpa*-MH86-1 mutant harvested at the field trial Jiaxing 2013.[8] A similar separation was also observed for the homozygous wild-type and the homozygous *lpa* mutant progenies of generation F_8 (Figure 3A). Despite significantly different positions of the crossing parents JH218 and JH99 in the PCA score plot, the homozygous wild-type and the homozygous *lpa* progenies, respectively, were within the same clusters. Similar patterns were also observed for the field trials Wuxi 2014 (Figure 3B) and Hainan 2015 (Figure 3C).

In order to elucidate the metabolites consistently contributing to the discriminations of MH86 and the *Os-lpa*-MH86-1 mutant, as well as of homozygous wild-type and homozygous *lpa* mutant progenies of generation F_8 at each field trial (Figure 3), the loading plots (Figure S5) of the corresponding PCAs and peak-by-peak comparisons via Chrom*pare*[18] were used. The metabolites were quantitated based on relative peak intensities, and simplified biosynthetic pathways were constructed based on the KEGG pathway database.[23] Figures 4 and 5 show comparisons of the levels of metabolites in (i) the original wild-type MH86 and the progenitor *lpa* mutant *Os-lpa*-MH86-1, (ii) the homozygous wild-type and the homozygous *lpa* mutant progenies of generation F_8 of the cross *Os-lpa*-MH86-1 × JH99, and (iii) the homozygous wild-type and the homozygous *lpa* mutant progenies of generation F_8 of the cross *Os-lpa*-MH86-1 × JH218.

In agreement with the previous results,[24] the *Os-lpa*-MH86-1 mutant consistently exhibited statistically significant differences in the levels of a broad spectrum of metabolites compared to the wild-type MH86 at each field trial. The metabolic changes in the *Os-lpa*-MH86-1 mutant included increased contents of sugars and sugar alcohols (e.g., fructose, glucose, sucrose, and *myo*-inositol), decreased contents of cysteine, increased levels of amino acids (e.g., alanine, serine and glycine), organic acids (e.g., citric acid, succinic acid), phytosterols (e.g., β-sitosterol) and the biogenic amine γ-aminobutyric acid (GABA). After crossing the *Os-lpa*-MH86-1 mutant with JH99, these metabolic changes were shown to be maintained in the homozygous *lpa* progenies of generation F_8 at three independent field trials (Figures 4 and 5). The homozygous *lpa* progenies of generation F_8 from the cross *Os-lpa*-MH86-1 × JH218 also exhibited this mutation-specific metabolite signature at each field trial.

Environmental Influence on the Rice Seed Metabolite Profiles. Multivariate analyses of the metabolite profiling data of the combined fractions I−IV were performed for all progeny rice seeds harvested at the three field trials. A PCA score plot of the combined fractions I−IV of the homozygous wild-type and the homozygous *lpa* progenies of the two crosses is shown in Figure 6A. The rice seeds harvested at growing locations in the subtropical zone (Jiaxing and Wuxi) clustered together; they were clearly separated from the rice seeds

Article

Table 1. Metabolites Contributing to the PCA Separation of the Lipophilic Fractions of the Crossing Parents JH99 and JH218 as well as of the Resulting Homozygous *lpa* Progenies of the Two Crosses with *Os-lpa*-MH86-1 at the Field Trial Jiaxing 2014[a]

| | crossing parents | | | | homozygous *lpa* mutant progenies | | | |
metabolites	JH99		JH218		cross 1		cross 2	
FAMEs and hydrocarbons								
C11:0	2.0 ± 0.1	b	1.2 ± 0.2	d	2.2 ± 0.1	a	1.6 ± 0.1	c
C14:0	130 ± 9.4	a	69 ± 1.8	c	140 ± 10.0	a	109 ± 6.4	b
C15:1	7.5 ± 0.4	a	4.5 ± 0.1	c	7.7 ± 0.6	a	6.2 ± 0.4	b
C15:0	8.7 ± 0.3	c	11.8 ± 0.4	b	7.5 ± 0.5	d	12.9 ± 0.3	a
C16:1	57.4 ± 5.2	a	41.2 ± 1.3	c	58.2 ± 3.5	a	47.0 ± 1.7	b
C16:0	4838 ± 176	a	3931 ± 110	b	4650 ± 203	a	3492 ± 224	c
C18:1	8064 ± 372	a	6742 ± 179	c	7524 ± 262	b	5952 ± 348	d
C20:0	194 ± 6.8	a	157 ± 6.8	b	190 ± 10	a	154 ± 8.4	b
C22:1	9.5 ± 0.3	a	7.4 ± 0.3	b	9.8 ± 0.7	a	7.3 ± 0.3	b
C22:0	84.3 ± 1.6	b	64.6 ± 2.0	d	90.1 ± 4.2	a	78.0 ± 3.0	c
C24:1	5.2 ± 0.2	a	4.4 ± 0.1	b	5.6 ± 0.5	a	4.1 ± 0.2	b
C24:0	143 ± 4.8	b	107 ± 2.9	d	160 ± 12	a	130 ± 5.0	c
squalene	28.8 ± 0.7	c	58.8 ± 3.7	a	28.6 ± 1.0	c	46.5 ± 1.3	b
FFAs								
C14:0	22.7 ± 1.5	a	12.5 ± 0.5	c	22.3 ± 1.4	a	18.4 ± 1.3	b
C16:0	664 ± 42	a	453 ± 18	c	664 ± 25	a	535 ± 29	b
C18:1	969 ± 29	b	717 ± 23	c	1255 ± 138	a	694 ± 64	c
C18:0	81.6 ± 2.1	a	54.7 ± 2.9	c	77.3 ± 4.7	a	60.3 ± 3.4	b
fatty alcohols and sterols								
C22:0	0.8 ± 0.05	a	0.5 ± 0.05	c	0.9 ± 0.1	a	0.6 ± 0.04	b
C24:0	7.8 ± 0.3	b	2.2 ± 0.2	d	9.5 ± 0.3	a	4.9 ± 0.2	c
C26:0	14.3 ± 1.3	a	2.9 ± 0.1	c	14.9 ± 0.5	a	10.1 ± 0.4	b
phytol	0.5 ± 0.05	d	2.6 ± 0.1	a	1.1 ± 0.1	c	1.7 ± 0.1	b
sitosterol	289 ± 12.1	a	195 ± 5.3	b	316 ± 23	a	223 ± 23	b
campestanol	1.5 ± 0.07	c	3.1 ± 0.1	b	1.5 ± 0.1	c	4.2 ± 0.3	a

[a]Relative peak intensities (i.e., metabolite peak intensity/internal standard peak intensity ×100) were used to express the levels of the metabolites. Values represent means ± standard deviations resulting from the analysis of six aliquots of freeze-dried flour. For each metabolite, different letters indicate statistically significant differences (Benjamini-Hochberg corrected $p < 0.05$) among the crossing parents JH99, JH218 and homozygous *lpa* mutant progenies of the two crosses. Cross 1: *Os-lpa*-MH86-1 × JH99; Cross 2: *Os-lpa*-MH86-1 × JH218.

harvested at the field trial in the tropical area (Hainan). Within each climate-zone, differentiations of the progenies depending on the phenotypes (wild-type and *lpa* mutant) and the crosses were observed. Similarly, the environment was still the dominant factor driving the overall metabolite profiles of progeny rice seeds when only the field trials Jiaxing and Wuxi located in the same climate zone were considered (Figure 6B).

The metabolites responsible for the climate zone-dependent clustering were selected based on the variable importance in projection (VIP) values (>1.0) via a supervised OPLS-DA model (Figure S6). As shown in the heat-map depicted in Figure 7, a broad spectrum of rice constituents exhibited changed levels in response to the different growing environments. For example, except for linoleic acid (C18:2 FFA), a number of FAMEs and FFAs showed increased levels in the rice seeds harvested at Hainan, a growing location with a higher temperature. This is in accordance with previously reported positive correlations between the temperature during the rice plant ripening stage and the concentrations of myristic (C14:0 FFA), palmitoleic (C16:1 FFA), stearic (C18:0 FFA), and oleic (C18:1) acids in rice seeds; negative correlations have been reported for linoleic (C18:2 FFA) and linolenic (C18:3 FFA) acids.[25] The levels of several sugars, sugar alcohols, and amino acids were also increased in the rice seeds from Hainan. On the contrary, only a few metabolites exhibited lower levels in Hainan than in Jiaxing and Wuxi,

e.g., trehalose, glutamic acid, asparagine, and threonic acid (Figure 7).

Separation of Wild-Types and *lpa* Mutants via Supervised Statistical Assessment. The separate investigations of the lipophilic and the polar fractions indicated that the metabolite profiles of the crossbred progenies were determined by superimposition effects, i.e., an impact of the crossing parent on the lipophilic compositions and an impact of the mutation on the polar constituents (Figure S7). A supervised multivariate analysis via OPLS-DA was performed to check the possibility to distinguish wild-type and *lpa* mutant rice seeds based on the metabolite profiling data of combined fractions I−IV. Taking the observed predominant impact of environment on the metabolite profiles into account (Figure 6), the investigations were conducted for each field trial separately.

Figure 8A shows the OPLS-DA score plot of metabolite profiling data of homozygous wild-type and homozygous *lpa* progenies of the two crosses, as well as of the original wild-type MH86 and the progenitor *lpa* mutant *Os-lpa*-MH86-1 harvested at the field trial Jiaxing 2014. A pronounced separation between wild-types and *lpa* mutants was observed along the predictive component. In addition, within the two clusterings, the homozygous wild-type and *lpa* progenies of cross *Os-lpa*-MH86-1 × JH99 were grouped more closely to MH86 and *Os-lpa*-MH86-1 along the orthogonal component, respectively, than those of the cross *Os-lpa*-MH86-1 × JH218

53
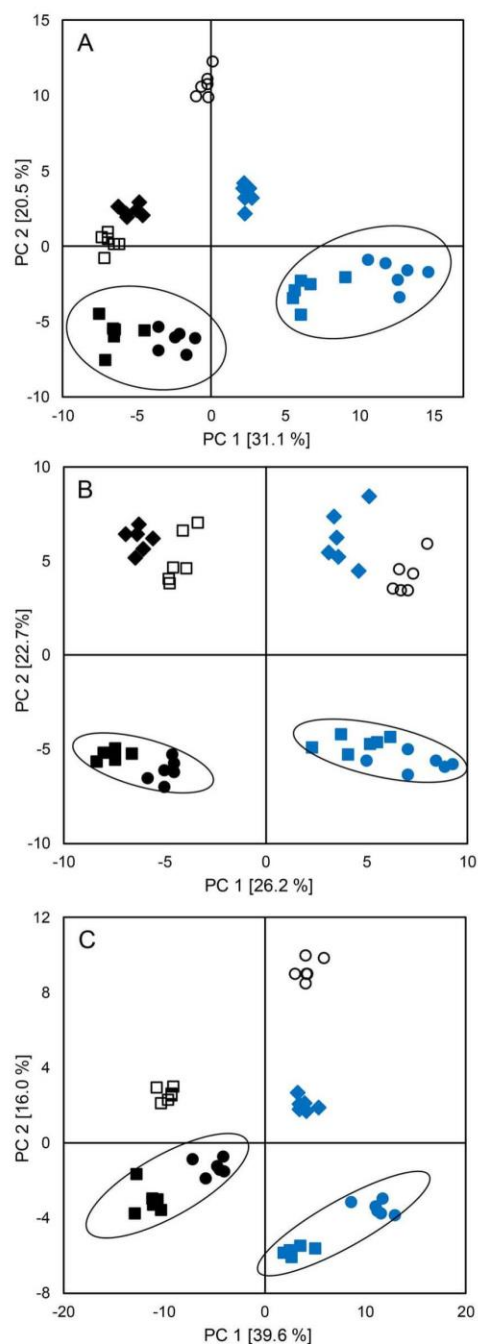
Figure 3. PCA score plots of metabolite profiling data of polar fractions III−IV of wild-type (black) and *lpa* mutant (blue) rice seeds harvested at field trials (A) Jiaxing 2014, (B) Wuxi 2014, and (C) Hainan 2015: original wild-type MH86 (diamonds) and *Os-lpa*-MH86-1 mutant (diamonds); crossing parents JH99 (open squares) and JH218 (open circles); homozygous wild-type and homozygous *lpa* mutant progenies of the crosses *Os-lpa*-MH86-1 × JH99 (solid squares) and *Os-lpa*-MH86-1 × JH218 (solid circles). The encircled clusters (boundaries correspond to the 95% Hotelling's T2 ellipses) include wild-type and *lpa* mutant progenies, respectively, from both crosses.

(Figure 8A). The OPLS-DA score plots at the field trials Wuxi (Figure 8B) and Hainan (Figure 8C) exhibited the same clusterings of wild-types and *lpa* mutants along the predictive components, as well as similar distributions along the orthogonal components within each cluster. The high values of R^2 and Q^2 and the results of permutation tests (Figure S8) demonstrated that these OPLS-DA models were robust and not overfitted.

The metabolites consistently exhibiting VIP values higher than 1.0 in three OPLS-DA models were regarded as the discriminating metabolites for the separation of the phenotypes wild-type and *lpa* mutant (Table 2). They included organic acids (e.g., citric acid), sugars (e.g., glucose and sucrose), sugar alcohols (e.g., mannitol and *myo*-inositol), amino acids (e.g., glycine, leucine, and proline), and the biogenic amine GABA. For each field trial, the mean concentrations (expressed as relative peak intensities) of these metabolites always exhibited statistically significantly increased levels in the *lpa* mutant rice seeds compared to those in the wild-type rice seeds.

■ DISCUSSION

A 1-bp deletion in the putative sulfate transporter gene *OsSULTR3;3* is responsible for the *lpa* phenotype of the *Os-lpa*-MH86-1 mutant.[24] This gene was shown to directly or indirectly play a role in the cross-talk between sulfate and phosphate homeostasis and/or signaling, since the expression of several genes involved in grain phosphorus and sulfur homeostasis and metabolism were up- and/or down-regulated upon the disruption of *OsSULTR3;3*.[24] As a result, the *Os-lpa*-MH86-1 mutant exhibited a mutation-specific metabolite signature which was similar to that reported in plants suffering from sulfate deficiency, e.g., reduction of cysteine content and increased level of its precursor serine, increased concentrations of a wide range of amino acids, sugars, sugar alcohols, and GABA.[26,27] Potential mechanisms underlying this type of *lpa* mutation, metabolic pathways depicting the complex metabolic changes, and possible nutritional implications have previously been discussed.[24]

In addition, it has been demonstrated that this metabolite signature was nearly unaffected by crossing the *Os-lpa*-MH86-1 mutant with the commercial wild-type JH99 and was consistently expressed in homozygous *lpa* progenies over generations from F_4 to F_7 at the same field trial (Jiaxing 2013).[8] The present study demonstrated that the differentiation of homozygous wild-type and homozygous *lpa* mutant progenies resulting from this cross was also expressed in seeds grown under other environmental conditions. The previous study had shown metabolic shifts from the crossing parents and the crossbred progenies of generation F_4 to those of generations F_5−F_7, for both wild-types and *lpa* mutants.[8] In accordance with these results, also in the PCA score plots obtained in this study the progenitor *lpa* mutant *Os-lpa*-MH86-1 was separated from the homozygous *lpa* mutant progenies of generation F_8, and the wild-types MH86 and JH99 were separated from the homozygous wild-type progenies of generation F_8 (Figure 3). The uniformities observed within the two clusters of homozygous *lpa* mutant and homozygous wild-type progenies may be explained by the fact that, owing to the limited amount of sample materials, for each cross rice seeds of the same phenotype were pooled for the metabolite profiling analysis.

Both the crossing parent JH99 and the original wild-type MH86 belong to the *indica* rice subspecies and exhibited
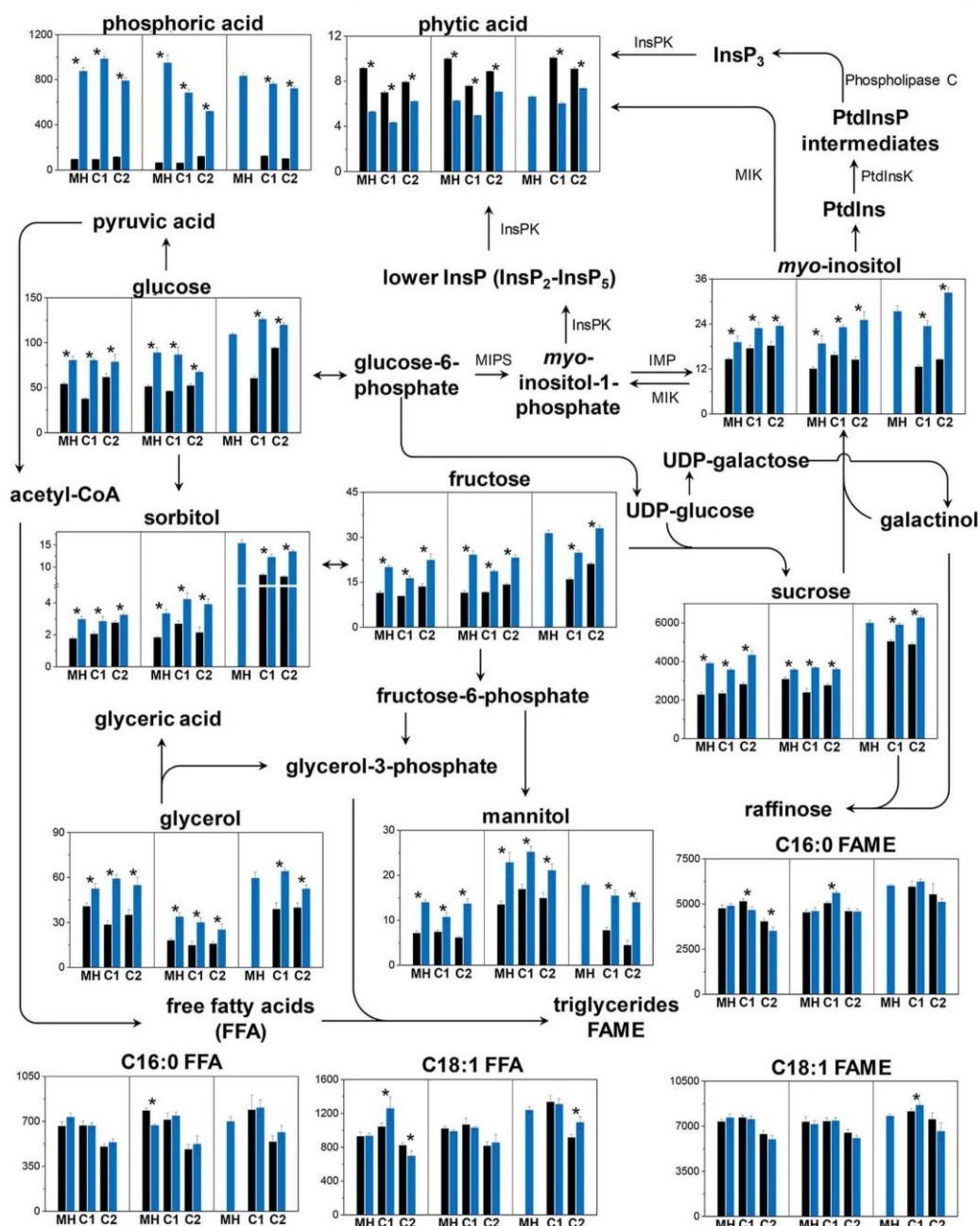
Figure 4. Simplified biosynthetic pathways of selected metabolites involved in phytic acid, sugar, and lipid metabolism. The wild-type rice seeds are shown in black and the *lpa* mutants in blue. The results for the field trials Jiaxing 2014, Wuxi 2014, and Hainan 2015 are successively shown in the three cells. For each cell, the bars are displayed in the following order: (MH) original wild-type MH86 and progenitor *Os-lpa*-MH86-1 mutant; (C1) homozygous wild-type and *lpa* progenies of cross 1 *Os-lpa*-MH86-1 × JH99; (C2) homozygous wild-type and *lpa* progenies of cross 2 *Os-lpa*-MH86-1 × JH218. The phytic acid contents are expressed in mg/g dry matter. All other metabolites are expressed as relative peak intensities. Values represent means ± standard deviations resulting from the analysis of six aliquots of freeze-dried flour. Asterisks represent statistically significant differences (Benjamini−Hochberg adjusted-$p < 0.05$) between the wild-type and the corresponding *lpa* mutant. *MIPS*, 1D-*myo*-inositol 3-phosphate synthase; *IMP*, *myo*-inositol monophosphatase; *MIK*, *myo*-inositol kinase; *PtdIns*, phosphatidyl inositol; *PtdInsK*, phosphatidyl inositol phosphate kinase; *PtdInsP*, phosphatidyl inositol phosphate; *InsPK*, inositol phosphate kinase; and *UDP*, uridine diphosphate.

similar metabolite profiles.[8] Therefore, it seemed reasonable to use another wild-type cultivar expected to show distinct

metabolic differences compared to JH99 and MH86 as additional crossing parent; accordingly, a commercial cultivar
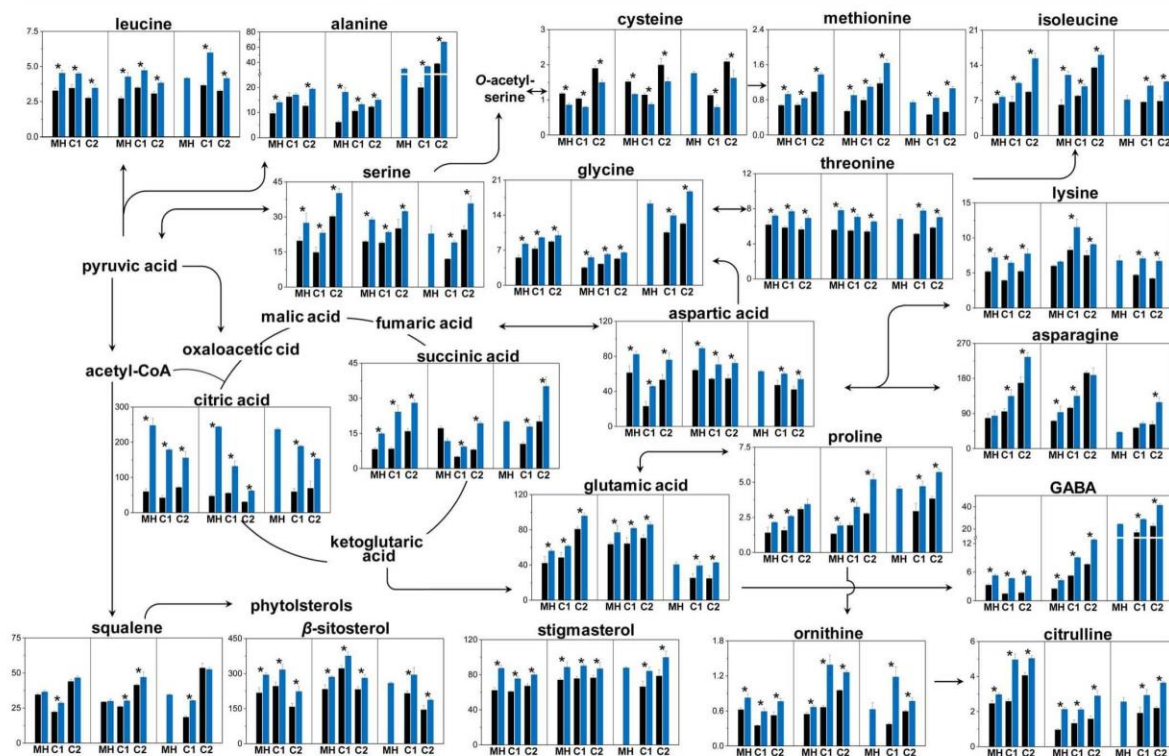
Figure 5. Simplified biosynthetic pathways of selected metabolites involved in amino acid metabolism, phytosterol metabolism, and tricarboxylic acid cycle. The wild-type rice seeds are shown in black and *lpa* mutants in blue. The results for the field trials Jiaxing 2014, Wuxi 2014, and Hainan 2015 are successively shown in the three cells. For each cell, the bars are displayed in the following order: (MH) original wild-type MH86 and progenitor *Os-lpa*-MH86-1 mutant; (C1) homozygous wild-type and homozygous *lpa* mutant progenies of cross 1 *Os-lpa*-MH86-1 × JH99; (C2) homozygous wild-type and homozygous *lpa* mutant progenies of cross 2 *Os-lpa*-MH86-1 × JH218. Metabolites are expressed as relative peak intensities, i.e., metabolite peak intensity/internal standard peak intensity ×100. Values represent means ± standard deviations resulting from the analysis of six aliquots of freeze-dried flour. Asterisks represent statistically significant differences (Benjamini−Hochberg adjusted-$p < 0.05$) between the wild-type and the corresponding *lpa* mutant.

(JH218) belonging to the *japonica* rice subspecies was selected. The crossing parents JH99 and JH218 indeed exhibited substantial metabolic differences regarding both lipophilic and polar constituents (Figures 2 and 3; Tables 1, S3, and S4). There is some information on metabolic variation between *japonica* and *indica* rice subspecies.[28] However, it is not possible to conclude whether the differences observed between JH99 and JH218 in this study result from the inherent variability of the two subspecies *japonica* and *indica*, or whether they just reflect individual properties specific for these two cultivars.

Notwithstanding these distinct differences in the metabolite profiles of the two crossing parents, the mutation-induced metabolic differences were consistently observed in the course of pairwise comparisons of homozygous wild-type and homozygous *lpa* progenies of generation F$_8$ for each cross (Figures 4 and 5).

The geographic locations of the three field trials investigated in this study and their climate conditions were very different. Hainan (latitude 18°30′ N) is located in the tropical zone while Jiaxing (30°45′ N) and Wuxi (31°29′ N) are in the subtropical zone. A remarkably higher annual average temperature and longer sunshine duration were recorded in Hainan compared to Jiaxing and Wuxi during the growing periods.[29,30]

Despite these differences of the environmental conditions, the homozygous *lpa* progenies of each cross expressed the mutation-specific metabolite signature compared to the corresponding homozygous wild-type progenies at each field trial (Figures 4 and 5). Only in very few cases, e.g., for asparagine and proline (Figure 5), the differences were not consistently statistically significant. This is a crucial result since it demonstrates that *OsSULTR3;3* mutation-induced metabolic changes are relatively stable and are not compromised in *lpa* mutant progenies by different growing conditions.

However, it is noteworthy that when comparing the metabolite profiles of all homozygous wild-type and *lpa* progeny rice seeds harvested from different field trials, the factor environment had a far more significant impact on the metabolite profiles of rice grains than the mutation effect (Figure 6). This means that on the one hand, the concentrations of lipophilic metabolites, such as C16:1 FAME, C18:0 FAME, and C18:2 FFA, were not affected by the mutation, but were strongly dependent on the growing environment (Figure 7). However, as shown, for example, for sorbitol (Figure 4), there were mutation-induced differences in concentrations between the homozygous wild-type and the homozygous *lpa* progenies at each single field trial. However, this feature of the signature was "masked" by the environ-

56
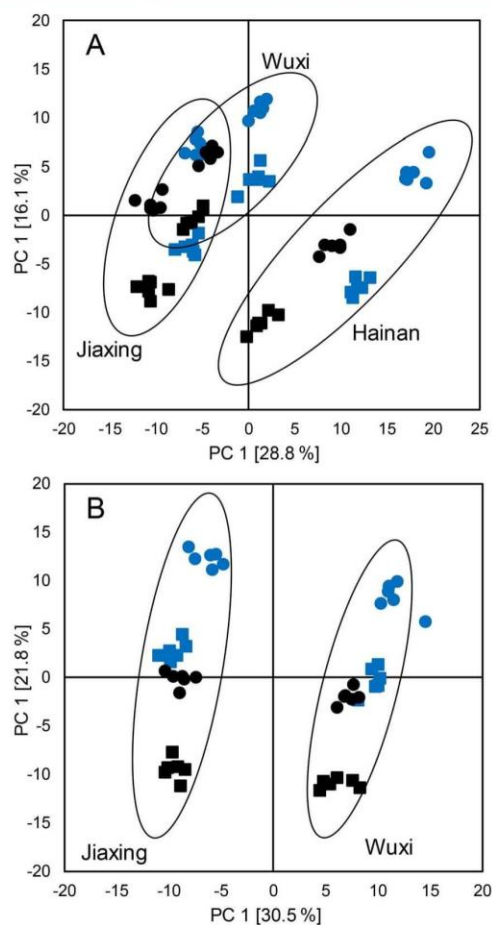
Figure 6. PCA score plot of metabolite profiling data of combined fractions I–IV of homozygous wild-type (black) and homozygous *lpa* mutant (blue) progenies of the crosses *Os-lpa*-MH86-1 × JH99 (solid squares) and *Os-lpa*-MH86-1 × JH218 (solid circles). Data are displayed for (A) field trials Jiaxing 2014, Wuxi 2014 and Hainan 2015; (B) field trials Jiaxing 2014 and Wuxi 2014. The encircled clusters (boundaries correspond to the 95% Hotelling's T2 ellipses) include all progenies harvested at the respective field trial.

Figure 7. Heat map of the metabolites responsible for the climate zone-dependent clustering of the homozygous wild-type and the homozygous *lpa* mutant progenies of the crosses *Os-lpa*-MH86-1 × JH99 and *Os-lpa*-MH86-1 × JH218. Lanes 1 and 2: homozygous wild-type and homozygous *lpa* progenies of cross *Os-lpa*-MH86-1 × JH99, respectively; and lanes 3 and 4: homozygous wild-type and homozygous *lpa* progenies of cross *Os-lpa*-MH86-1 × JH218, respectively.

mental impact, since the absolute levels of sorbitol in rice seeds from Hainan, regardless of wild-type or *lpa* phenotype, were all significantly higher than those from Jiaxing and Wuxi (Figure 4). Similar results could also be observed for a number of other metabolites, e.g., glucose, sucrose, alanine, glycine, asparagine, proline, glutamic acid, and GABA (Figures 4 and 5). Only for phosphoric acid, phytic acid, and citric acid, the absolute levels in the *lpa* rice seeds exhibited consistent differences compared to those in the wild-type rice seeds, independent from the environmental impact. Previous studies on rice,[9,31] maize,[10,32] and barley[12] also confirmed that growing environment was an important parameter driving the variability of the grain metabolite profiles.

In accordance with the previous results,[8,24] the *OsSULTR3;3* mutation was shown to have little impact on the lipophilic constituents in the progenitor *lpa* mutant *Os-lpa*-MH86-1 as well as in the homozygous *lpa* progenies of the two crosses, and no clear differentiation of lipophilic metabolite profiles was

observed between wild-types and *lpa* mutants for each cross (Figure 2). For the cross *Os-lpa*-MH86-1 × JH99, the previous study had shown metabolic shifts, in particular for the lipophilic fractions, from the original wild-type, the two crossing parents and the progenies of generation F_4 to the progenies of generations F_5–F_7. This could be explained by the genetic diversity to be expected from the employed breeding scheme.[8] Homozygous wild-type and homozygous *lpa* mutant F_4 seeds had been directly generated by growing the homozygous F_3 seeds. Alternatively, F_5 to F_7 seeds had been obtained from the segregating F_3 seeds,[8] a route also employed to produce the F_8 seeds investigated in the present study. The resulting genetic diversity was reflected in the obtained score plots (Figure 2), in which the wild-type MH86 as well as the two crossing parents *Os-lpa*-MH86-1 and JH99 were separated from the progenies of generation F_8 obtained for the cross *Os-lpa*-MH86-1 × JH99. Again, the uniformities observed within the two progeny clusters may be explained by the pooling of rice seeds of the same phenotype prior to metabolite profiling analysis.

The lipid profiles of the *lpa* progenies were significantly impacted by the corresponding crossing parents (Figure 2). In most cases, JH99 exhibited higher levels of individual lipophilic metabolites than JH218, and the homozygous *lpa* progenies resulting from JH99 also displayed lipid-enriched compositions
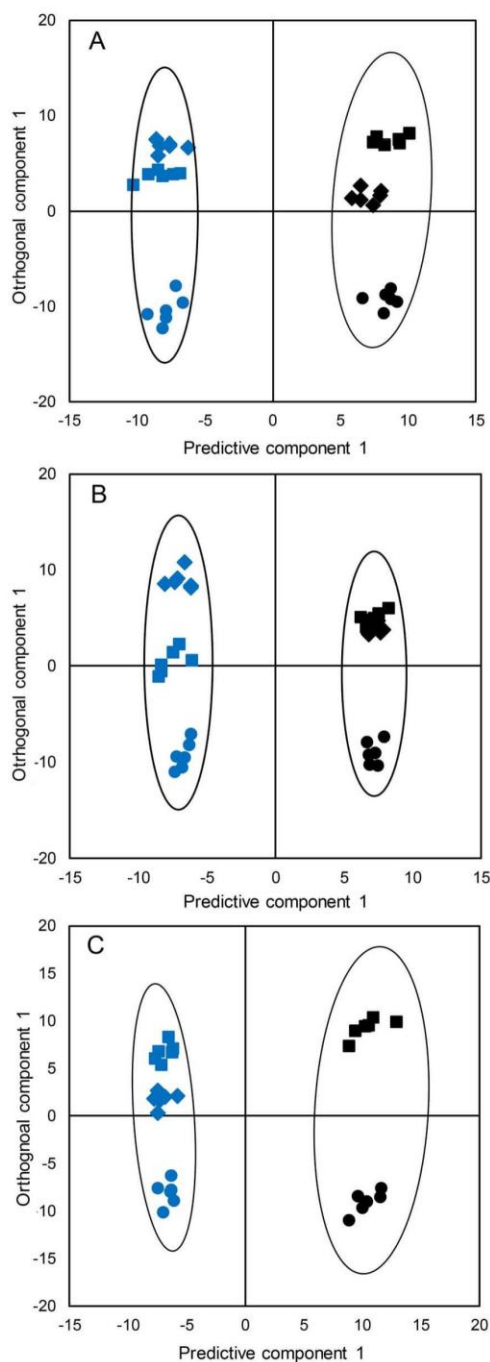
Figure 8. OPLS-DA score plots of metabolite profiling data of combined fractions I−IV of wild-type (black) and *lpa* mutant (blue) rice seeds harvested at field trials (A) Jiaxing 2014, (B) Wuxi 2014, and (C) Hainan 2015: original wild-type MH86 (diamonds) and *Os-lpa*-MH86-1 mutant (diamonds); homozygous wild-type and homozygous *lpa* mutant progenies of crosses *Os-lpa*-MH86-1 × JH99 (solid squares) and *Os-lpa*-MH86-1 × JH218 (solid circles). The boundaries of the clusters correspond to the 95% Hotelling's T2 ellipses.

compared to the homozygous *lpa* progenies resulting from JH218 (Tables 1, S3, and S4). This indicates that after crossing the *Os-lpa*-MH86-1 mutant with different commercial cultivars, the lipid compositions of the resulting *lpa* progenies would be strongly dependent on the lipid profiles of the crossing parent.

As reflected in the PCA score plots of the combined fractions I−IV for each field trial (Figure S7), there was a superimposition effect on the overall metabolite profiles of the resulting progenies after crossing *Os-lpa*-MH86-1 mutant with a commercial cultivar, i.e., the impact of the crossing parents on lipophilic constituents (Figure 2, Tables 1, S3, and S4) and of the mutation effect on polar constituents (Figures 3-5). The supervised multivariate statistical assessment OPLS-DA was shown to be able to classify the crossbred progenies into the respective wild-type and *lpa* mutant phenotypes based on their overall metabolite profiles (Figure 8). This further confirmed a stably expressed metabolite signature in *lpa* mutant rice seeds after cross-breeding.

The discriminating metabolites obtained from the OPLS-DA models are summarized in Table 2. All these metabolites exhibited statistically increased levels in the *lpa* mutants compared with the wild-type rice seeds. These metabolic changes in *lpa* mutants were stable and independent from the impact of crossing parents as well as from the genetic background variations in *lpa* progenies, and were consistently observed in different growing environments. It is noteworthy that these metabolites also include nutritionally relevant compounds, such as essential amino acids (e.g., leucine and methionine) and the biogenic amine GABA, a well-known neurotransmitter.[33] The spectrum of metabolites covered in this study was constricted by the employed GC/MS-based approach. The application of additional platforms might enable to investigate whether the *lpa* mutation also affects other nutrients and secondary metabolites or whether quality parameters such as flavor are influenced.

In conclusion, the results demonstrated that both the *lpa* trait and the metabolite signature of the *Os-lpa*-MH86-1 mutant were not hampered by cross-breeding with different commercial wild-type cultivars. The profiles of lipophilic metabolites of the homozygous *lpa* progenies were impacted by the lipid profiles of the crossing parents. However, this superimposition did not compromise the mutation-induced metabolite signature in homozygous *lpa* progenies. In addition, the metabolite profiles of the crossbred progenies were also shown to be strongly impacted by the environment. Tropical and subtropical growing locations resulted in clearly different metabolite patterns. However, for each individual field trial, the mutation-specific metabolite signature was maintained upon cross-breeding.

The elaborated data are valuable and promising from a breeder's point of view. They show that cross-breeding of *lpa* rice mutants with commercial cultivars does not compromise the intended phytic acid reduction and does not alter the mutation-specific metabolite signature of homozygous *lpa* mutant progenies. The results further strengthen the applicability of cross-breeding as a practical approach to produce *lpa* rice seeds. They demonstrate that independent from the subspecies and/or cultivar of the crossing parents and independent from the environmental impact, cross and selection breeding can be employed as a tool to generate *lpa* rice seeds stably exhibiting the mutation-induced metabolic traits.

Table 2. Metabolites Consistently Exhibiting VIP Values >1 for the OPLS-DA Separations of Wild-Type and *lpa* Mutant Rice Seeds at Each Field Trial[a]

metabolite	Jiaxing 2014 wild-type	Jiaxing 2014 *lpa* mutant	Wuxi 2014 wild-type	Wuxi 2014 *lpa* mutant	Hainan 2015 wild-type	Hainan 2015 *lpa* mutant
phosphoric acid	103 ± 11	881 ± 86	84 ± 28	714 ± 187	114 ± 13	770 ± 52
citric acid	57.7 ± 14	194 ± 43	44.1 ± 11	146 ± 77	64.3 ± 16	192 ± 36
glucose	51.1 ± 11	79.7 ± 5.5	49.8 ± 3.4	80.8 ± 12	77.3 ± 18	118 ± 7.4
fructose	11.7 ± 1.5	19.5 ± 2.9	12.4 ± 1.4	21.9 ± 2.6	18.5 ± 2.8	29.6 ± 3.8
sucrose	2473 ± 270	3925 ± 354	2743 ± 325	3606 ± 86	4957 ± 120	6047 ± 192
sorbitol	2.2 ± 0.4	3.0 ± 0.3	2.2 ± 0.4	3.8 ± 0.5	8.0 ± 0.3	13.7 ± 1.5
mannitol	6.9 ± 0.7	12.7 ± 1.8	15.1 ± 1.8	23.0 ± 2.4	6.1 ± 1.9	15.7 ± 1.9
myo-inositol	16.7 ± 1.8	21.8 ± 2.4	14.1 ± 1.7	22.2 ± 3.2	13.5 ± 1.1	27.6 ± 4.0
glycerol	34.7 ± 5.9	55.4 ± 4.8	16.1 ± 2.3	29.6 ± 4.9	39.3 ± 3.8	58.6 ± 5.7
glycine	7.2 ± 1.4	9.2 ± 0.9	4.3 ± 0.8	6.1 ± 0.5	11.4 ± 0.9	16.3 ± 2.1
leucine	3.2 ± 0.4	4.2 ± 0.5	3.1 ± 0.3	4.3 ± 0.4	3.5 ± 0.2	4.8 ± 0.9
threonine	5.9 ± 0.4	7.3 ± 0.4	5.5 ± 0.2	7.1 ± 0.6	5.5 ± 0.4	7.2 ± 0.6
aspartic acid	45.8 ± 18	67.8 ± 17	57.6 ± 5.5	77.1 ± 10	44.5 ± 5.2	58.8 ± 4.5
ornithine	0.5 ± 0.1	0.7 ± 0.1	0.7 ± 0.2	1.1 ± 0.3	0.5 ± 0.1	0.9 ± 0.3
citrulline	3.0 ± 0.8	4.3 ± 1.0	1.3 ± 0.3	2.4 ± 0.4	2.1 ± 0.3	3.0 ± 0.5
proline	2.0 ± 0.8	2.7 ± 0.6	2.0 ± 0.6	3.4 ± 1.5	3.4 ± 0.6	5.0 ± 0.6
methionine	0.8 ± 0.1	1.0 ± 0.2	0.8 ± 0.3	1.2 ± 0.3	0.5 ± 0.0	0.9 ± 0.1
GABA	2.1 ± 0.9	5.0 ± 0.3	5.1 ± 2.2	8.7 ± 3.6	19.9 ± 3.5	31.4 ± 7.5

[a]For each metabolite, the relative peak intensity (i.e., metabolite peak intensity/internal standard peak intensity ×100) of wild-type was the mean value of original MH86 and homozygous wild-type progenies of two crosses; the relative peak intensity of *lpa* mutant was the mean value of *Os-lpa*-MH86-1 and homozygous *lpa* mutant progenies of two crosses. All metabolites exhibited statistically significantly different (Benjamini−Hochberg adjusted-$p < 0.05$) levels between wild-type and *lpa* mutant rice seeds at each field trial.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jafc.8b06696.

Compounds identified in fractions I−IV of MH86 rice seeds; levels of metabolites in progenies depending on the crossing parents; genotyping results via high-resolution melting (HRM) analysis; GC chromatograms of fractions I−IV of MH86; loading plots of metabolite profiling data for lipophilic and polar fractions; OPLS-DA score plot and permutation tests of resulting progenies from three field trials; PCA score plots of combined fractions I−IV at each field trial; and permutation tests of OPLS-DA models (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +49 8161 714250. Fax: +49 8161 714259. E-mail: k.h. engel@wzw.tum.de.

ORCID ⓞ

Karl-Heinz Engel: 0000-0001-8932-5175

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Raboy, V. myo-Inositol-1,2,3,4,5,6-hexakisphosphate. Phytochemistry 2003, 64, 1033−1043.

(2) Kumar, V.; Sinha, A. K.; Makkar, H. P. S.; Becker, K. Dietary roles of phytate and phytase in human nutrition: A review. Food Chem. 2010, 120, 945−959.

(3) Abdel-Megeed, A.; Tahir, A. Reduction of Phosphorus Pollution from Broilers Waste through Supplementation of Wheat Based Broilers Feed with Phytase. J. Chem. 2015, 3, 1−3.

(4) Raboy, V., Seed phosphorus and the development of low-phytate crops. In Inositol Phosphates: Linking Agriculture and the Environment, 1st ed.; Turner, B. L., Richardson, A. E., Mullaney, E. J., Eds.; CAB International: Wallingford, UK, 2007; pp 111−132.

(5) Spear, J. D.; Fehr, W. R. Genetic improvement of seedling emergence of soybean lines with low phytate. Crop Sci. 2007, 47, 1354−1360.

(6) Averitt, B.; Shang, C.; Rosso, L.; Qin, J.; Zhang, M. C.; Rainy, K. M.; Zhang, B. Impact of mips1, lpa1, and lpa2 Alleles for Low Phytic Acid Content on Agronomic, Seed Quality, and Seed Composition Traits of Soybean. Crop Sci. 2017, 57, 2490−2499.

(7) Wiggins, S. J.; Smallwood, C. J.; West, D. R.; Kopsell, D. A.; Sams, C. E.; Pantalone, V. R. Agronomic Performance and Seed Inorganic Phosphorus Stability of Low-Phytate Soybean Line TN09−239. J. Am. Oil Chem. Soc. 2018, 95, 787−796.

(8) Zhou, C. G.; Tan, Y. Y.; Gossner, S.; Li, Y. F.; Shu, Q. Y.; Engel, K. H. Stability of the metabolite signature resulting from the OsSULTR3;3 mutation in low phytic acid rice (Oryza sativa L.) seeds upon cross-breeding. J. Agric. Food Chem. 2018, 66, 9366−9376.

(9) Zhang, J. J.; Zhao, C. X.; Zeng, Z. D.; Luo, P.; Zhao, Y. N.; Zhao, J. Y.; Li, L. L.; Lu, X.; Xu, G. W. Sample-directed pseudotargeted method for the metabolic profiling analysis of rice seeds based on liquid chromatography with mass spectrometry. J. Sep. Sci. 2016, 39, 247−255.

(10) Frank, T.; Röhlig, R. M.; Davies, H. V.; Barros, E.; Engel, K.-H. Metabolite profiling of maize kernels - genetic modification versus environmental influence. J. Agric. Food Chem. 2012, 60, 3005−3012.

(11) Lanzinger, A.; Frank, T.; Reichenberger, G.; Herz, M.; Engel, K.-H. Metabolite profiling of barley grain subjected to induced drought stress: responses of free amino acids in differently adapted cultivars. J. Agric. Food Chem. 2015, 63, 4252−4261.

(12) Wenzel, A.; Frank, T.; Reichenberger, G.; Herz, M.; Engel, K. H. Impact of induced drought stress on the metabolite profiles of barley grain. Metabolomics 2015, 11, 454−467.

(13) Liu, Q. L.; Xu, X. H.; Ren, X. L.; Fu, H. W.; Wu, D. X.; Shu, Q. Y. Generation and characterization of low phytic acid germplasm in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* **2007**, *114*, 803−814.

(14) Tan, Y. Y.; Yu, X. M.; Shu, Q. Y.; Zhang, H. L.; Wang, S. g.; Yuan, F. J.; Shi, C. H. Development of an HRM-based, safe and high-throughput genotyping system for two low phytic acid mutations in soybean. *Mol. Breed.* **2016**, *36*, 101.

(15) Frank, T.; Meuleye, B. S.; Miller, A.; Shu, Q. Y.; Engel, K. H. Metabolite profiling of two *low phytic acid* (*lpa*) rice mutants. *J. Agric. Food Chem.* **2007**, *55*, 11011−11019.

(16) Kamaleldin, A.; Appelqvist, L. A.; Yousif, G.; Iskander, G. M. Seed lipids of *Sesamum indicum* and related wild species in Sudan. The Sterols. *J. Sci. Food Agric.* **1992**, *59*, 327−334.

(17) Xu, Z. M.; Godber, J. S. Purification and identification of components of gamma-oryzanol in rice bran oil. *J. Agric. Food Chem.* **1999**, *47*, 2724−2728.

(18) Frenzel, T.; Miller, A.; Engel, K. H. A methodology for automated comparative analysis of metabolite profiling data. *Eur. Food Res. Technol.* **2003**, *216*, 335−342.

(19) Chong, J.; Soufan, O.; Li, C.; Caraus, I.; Li, S. Z.; Bourque, G.; Wishart, D. S.; Xia, J. G. MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* **2018**, *46*, W486−W494.

(20) Giacomoni, F.; Le Corguille, G.; Monsoor, M.; Landi, M.; Pericard, P.; Petera, M.; Duperier, C.; Tremblay-Franco, M.; Martin, J. F.; Jacob, D.; Goulitquer, S.; Thevenot, E. A.; Caron, C. Workflow4Metabolomics: a collaborative research infrastructure for computational metabolomics. *Bioinformatics* **2015**, *31*, 1493−1495.

(21) Vinaixa, M.; Samino, S.; Saez, I.; Duran, J.; Guinovart, J. J.; Yanes, O. A Guideline to Univariate Statistical Analysis for LC/MS-Based Untargeted Metabolomics-Derived Data. *Metabolites* **2012**, *2*, 775.

(22) Frenzel, T.; Miller, A.; Engel, K. H. Metabolite profiling - A fractionation method for analysis of major and minor compounds in rice grains. *Cereal Chem.* **2002**, *79*, 215−221.

(23) Kanehisa, M.; Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **2000**, *28*, 2730.

(24) Zhao, H.; Frank, T.; Tan, Y.; Zhou, C.; Jabnoune, M.; Arpat, A. B.; Cui, H.; Huang, J.; He, Z.; Poirier, Y.; Engel, K.-H.; Shu, Q. Disruption of *OsSULTR3;3* reduces phytate and phosphorus concentrations and alters the metabolite profile in rice grains. *New Phytol.* **2016**, *211*, 926−939.

(25) Kitta, K.; Ebihara, M.; Iizuka, T.; Yoshikawa, R.; Isshiki, K.; Kawamoto, S. Variations in lipid content and fatty acid composition of major non-glutinous rice cultivars in Japan. *J. Food Compos. Anal.* **2005**, *18*, 269−278.

(26) Hammond, J. P.; White, P. J. Sucrose transport in the phloem: integrating root responses to phosphorus starvation. *J. Exp. Bot.* **2007**, *59*, 93−109.

(27) Nikiforova, J. V.; Bielecka, M.; Gakière, B.; Krueger, S.; Rinder, J.; Kempa, S.; Morcuende, R.; Scheible, W.-R.; Hesse, H.; Hoefgen, R. Effect of sulfur availability on the integrity of amino acid biosynthesis in plants. *Amino Acids* **2006**, *30*, 173−183.

(28) Hu, C.; Shi, J.; Quan, S.; Cui, B.; Kleessen, S.; Nikoloski, Z.; Tohge, T.; Alexander, D.; Guo, L.; Lin, H.; Wang, J.; Cui, X.; Rao, J.; Luo, Q.; Zhao, X.; Fernie, A. R.; Zhang, D. Metabolic variation between *japonica* and *indica* rice cultivars as revealed by non-targeted metabolomics. *Sci. Rep.* **2015**, *4*, 5067.

(29) World Weather Online Database; https://www.worldweatheronline.com (June 15, 2018 updated).

(30) Zeng, X.; Zhao, N.; Ma, Y. In *Distributed Modeling of Surface Solar Radiation Based on Aerosol Optical Depth and Sunshine Duration in China*; IOP Conference Series: Earth and Environmental Science, IOP Publishing: 2018.

(31) Song, E. H.; Jeong, J.; Park, C. Y.; Kim, H. Y.; Kim, E. H.; Bang, E.; Hong, Y. S. Metabotyping of rice (*Oryza sativa* L.) for understanding its intrinsic physiology and potential eating quality. *Food Res. Int.* **2018**, *111*, 20−30.

(32) Röhlig, R. M.; Eder, J.; Engel, K.-H. Metabolite profiling of maize grain: differentiation due to genetics and environment. *Metabolomics* **2009**, *5*, 459.

(33) Moore, R. Y.; Speh, J. C. GABA is the Principal Neuro-transmitter of the Circadian System. *Neurosci. Lett.* **1993**, *150*, 112−116.

4. DISCUSSION

The objectives of the studies underlying this dissertation were to investigate the stability of both the reductions of phytic acid (PA) contents and the mutation-induced metabolite signature in homozygous *low phytic acid* (*lpa*) mutant progenies obtained by crossing *lpa* rice mutants with commercial wild-type rice cultivars. Overarching key aspects that were crucial for the respective study design and that were taken into account in the interpretation of the elaborated data will be discussed in the following sections.

4.1 Methodology

Phytic acid analysis

Early methods employed for the analysis of phytic acid were based on the formation of Fe^{3+}–phytate precipitates and colorimetry-based approaches (Heubner *et al.*, 1914; McCance *et al.*, 1935). These were later replaced by ion-exchange methods (Harland *et al.*, 1977) and high performance liquid chromatography (HPLC) ion-pair methods (Lehrfeld, 1994). Nowadays high-pressure ion chromatography (HPIC) is one of the most frequently employed approaches for the simultaneous determination of PA and of lower inositol phosphates ($InsP_2$ – $InsP_5$). After simple sample preparations (extraction, centrifugation and filtration), the extract is directly subjected to ion chromatography. The separation of PA and $InsP_2$ – $InsP_5$ is performed by a linear gradient elution, and the detection is achieved by an ultraviolet absorbance detector after post-column derivatization. The HPIC method exhibits high sensitivity and good capability in separating and analyzing PA as well as stereoisomers (excluding enantiomers) of $InsP_2$ – $InsP_5$ (Chen *et al.*, 2003; Chen, 2004).

In the present study, an HPIC-based approach (Oates *et al.*, 2014) was adapted for the determination of PA contents in the rice seeds. This method would have also allowed to obtain quantitative information on the contents of $InsP_2$ – $InsP_5$, as demonstrated in a parallel study on the contents of inositol phosphates in progenies of soybean *lpa* mutants (Goßner *et al.*, 2019). The potential accumulation of lower inositol phosphates is of interest, since they have been shown to possess important physiological functions, e.g., $Ins(1,3,4,5)P_4$ regulates the immune cell development

(Sauer *et al.*, 2010), as well as nutritional properties, e.g., no inhibitory effect on the absorption of minerals was observed for $InsP_4$ and $InsP_5$ compared to phytic acid (Lonnerdal *et al.*, 1989; Sandberg *et al.*, 1999). However, in the rice *lpa* mutant types investigated in this thesis such an accumulation of lower inositol phosphates was not observed, which is in line with previously reported results (Frank *et al.*, 2007; Zhao *et al.*, 2016).

Metabolite profiling

The GC-metabolite profiling approach applied in this study is based on the principles of a methodology originally developed using rice as model crop (Frenzel *et al.*, 2002). It comprises the preparation of extracts containing lipophilic and polar rice constituents and their subsequent fractionation, employing transesterification and solid phase extraction for the lipophilic metabolites and selective hydrolysis of silylated derivatives for the polar constituents. The resulting four fractions are subsequently subjected to GC/MS, a platform which offers high chromatographic resolution (separation of peaks), sufficient reproducibility of peak retention times, and the possibility to use commercially available mass spectral data bases. This technique allows the simultaneous detection, identification and relative quantification of a broad spectrum of low molecular weight (< 1000 Da) rice constituents, ranging from lipophilic to hydrophilic metabolites for both major (e.g. fatty acid methyl esters and sugars) and minor (e.g. free fatty acids and free amino acids) compound classes (Frenzel *et al.*, 2002).

This untargeted GC-based metabolite profiling approach has also been shown to be suitable for the analysis of other crops, e.g. maize, soybean, mung bean and barley. Investigations employing this methodology have not only been used to investigate *lpa* crops (Frank *et al.*, 2007; Frank *et al.*, 2009b), but have also covered a wide range of topics, e.g., influence of farming practices (conventional vs. organic farming) on metabolite profiles of maize (Röhlig *et al.*, 2010), metabolic changes during the sprouting of mung beans (Na Jom *et al.*, 2011) and the malting of barley (Frank *et al.*, 2011), metabolic differences between colored rice (Frank *et al.*, 2012a), and metabolic responses of barley to drought stress (Lanzinger *et al.*, 2015; Wenzel *et al.*, 2015).

Despite the broad applicability and the good reproducibility of the employed GC/MS-based metabolite profiling, some limitations of this method should be noted. The applied fractionation steps, on the one hand, deliver very useful information regarding the contributions of specific classes of metabolites. For example, in the studies

performed for this thesis, the fractionation greatly assisted in assigning the roles of metabolite classes regarding the impact of crossing partner and environment and pinpointed those metabolite classes contributing to the differentiation of wild-types and *lpa* rice mutants. On the other hand, the procedure is time-consuming, and this hampers its application as high-throughput screening of metabolic phenotypes.

In addition, the scope of metabolites covered by GC is inherently limited to those that are volatile or that can be easily transformed into volatile derivatives (Kopka *et al.*, 2004). Important secondary metabolites (e.g. flavonoids and alkaloids) are not accessible to analysis via this technique. Liquid chromatography coupled to mass spectrometry (LC/MS) would be an important complementary technique covering this gap and enabling the analysis of high-molecular weight and/or thermolabile metabolites (Allwood *et al.*, 2010). Therefore, the application of alternative analytical platforms might extend the scope of metabolites analyzed in this study and thus allow to make correlations to other quality and nutrition-related parameters of the *lpa* rice mutants that might be affected by the mutation and/or the cross-breeding step.

Statistical assessments

The utilization of appropriate statistical tools is a crucial step in interpreting the huge and complex data matrices resulting from the metabolite profiling. In the present study, the unsupervised multivariate method of principal component analysis (PCA) was employed as initial step to generate an unbiased overview of the metabolomics-data matrices with reduced dimensions. In order to identify the individual metabolites contributing to the differentiations of metabolite phenotypes between the wild-types and the *lpa* mutants, the supervised multivariate statistical assessment of partial least squares-discriminate analysis (PLS-DA) was attempted. However, due to the significant variations in metabolite profiles of rice seeds resulting from generation effects and from the environmental impact, the models of the generated PLS-DAs were always shown to be over-fitted. Therefore, orthogonal partial least squares-discriminate analysis (OPLS-DA) was used as another supervised statistical tool. As a result, the metabolite profiling data with pronounced variations that are uncorrelated to pre-defined group labels (e.g. wild-type and *lpa* mutant) were filtered by the Orthogonal Signal Correction (OSC) and were reflected in the OPLS-DA score plot along the Orthogonal Component axis (y-axis) (Bylesjo *et al.*, 2006). The established models proved to be robust and not over-fitted by permutation tests and cross-validations.

Finally, the metabolites exhibiting high Variable Importance in Projection values for the Predictive Component (VIP_ pred) of the OPLS-DA model were selected, followed by univariate analyses. These metabolites were further mapped in simplified biosynthetic pathways to illustrate their contributions to the mutation-induced metabolite signature of the *lpa* mutants and their progenies.

4.2 *Lpa* mutation-types

During the past years, several *lpa* rice mutant lines have been developed via mutation breeding (Larson *et al.*, 2000; Rutger *et al.*, 2004; Liu *et al.*, 2007; Kim *et al.*, 2008; Kim *et al.*, 2014). They differ in the achieved percentage reductions of phytic acid contents, ranging from −11.8% to −68.8%, as well as in the mutation targets. Reductions in the concentrations of phytic acid in induced *lpa* mutants have been reported to occur via three routes: (i) mutations in genes directly involved in the biosynthesis of phytic acid (Zhao *et al.*, 2008b); (ii) mutations perturbing transportation and storage of phytic acid in plant tissues (Xu *et al.*, 2009); and (iii) mutations in genes involved in a cross-talk between sulfate and phosphate homeostasis (Zhao *et al.*, 2016). In order to cover these different types of mutation, the three previously developed *lpa* rice mutants *Os-lpa*-XQZ-1, *Os-lpa*-XS110-2 and *Os-lpa*-MH86-1 carrying the respective mutation targets were employed.

In *Os-lpa*-XQZ-1, the mutation of a homolog of the 2-phosphoglycerate kinase (*2-PGK*) gene, which is directly involved in the biosynthetic step from $InsP_1$ to $InsP_2$, was shown to be responsible for the *lpa* phenotype (Zhao *et al.*, 2008b; Sparvoli *et al.*, 2015). The PA contents in this progenitor *lpa* mutants ranged from 4.32 mg/g to 7.5 mg/g, and they showed percentage reductions from −11.8% to −46.6% at different field trials (Liu *et al.*, 2007; Frank *et al.*, 2009a).

In *Os-lpa*-XS110-2, the mutation was attributed to a single base pair change (C/G−T/A transition) of the multi-drug resistance-associated protein ABC transporter gene 5 (*OsMRP5*). The *OsMRP5* mutation in *Os-lpa*-XS110-2 induces the PA reductions by triggering a disturbance in tissue compartmentation of PA and/or its transport and storage to the vacuole (Xu *et al.*, 2009; Sparvoli *et al.*, 2014). The previously reported ranges for the PA contents and the percentage reductions observed in *Os-lpa*-XS110-2

lpa mutants were 4.0 mg/g to 6.6 mg/g and –33.8% to –63.5%, respectively (Frank *et al.*, 2007; Liu *et al.*, 2007; Frank *et al.*, 2009b).

In *Os-lpa*-MH86-1, γ-irradiation resulted in a one base pair deletion in a putative sulfate transporter gene (*OsSULTR3;3*). This gene was shown to play a role in the cross-talk between sulfate and phosphate homeostasis and/or signaling (Zhao *et al.*, 2016). The *Os-lpa*-MH86-1 *lpa* mutants exhibited significantly lower levels of PA (5.1 mg/g to 6.7 mg/g) as well as pronounced percentage reductions (–35.1% to –43.9%) compared to the corresponding wild-types MH86 at different field trials (Liu *et al.*, 2007; Zhao *et al.*, 2016).

The employed three *lpa* mutants represent different principles of PA reductions. The complexities of the underlying mechanisms gradually increase in the order *Os-lpa*-XQZ-1, *Os-lpa*-XS110-2 and *Os-lpa*-MH86-1. This was reflected in the different dependencies of the phytic acid contents of the generated homozygous *lpa* mutant progenies on parameters, such as crossing partner and environmental conditions.

Previously reported metabolic changes in *lpa* rice mutants had been limited to changes in the levels of a few metabolites related to the biosynthesis of phytic acid, e.g. *myo*-inositol, galactose and raffinose (Frank *et al.*, 2007; Frank *et al.*, 2009b). The *lpa* mutant *Os-lpa*-MH86-1 was the first example of an *lpa* rice mutant, in which the mutation of a putative sulfate transporter gene *OsSULTR3;3* resulted in a broad spectrum of metabolic changes compared to the corresponding wild-type (Zhao *et al.*, 2016). These changes comprised not only reduced contents of the sulfur-containing amino acid cysteine, but also increased contents of phytosterols, sugars, sugar alcohols, acids and amino acids. As examples, the metabolites that showed significantly different levels in the *lpa* mutant *Os-lpa*-MH86-1 and the corresponding wild-type MH86 rice seeds harvested at the field trial Hangzhou in 2011 are listed in Table 4.

The complex changes in the metabolite profile of the *lpa* mutant *Os-lpa*-MH86-1 are similar to those reported in plants grown under phosphorus deficiency and/or sulfate starvation, e.g. decreased level of cysteine, increased levels of glucose, fructose, sucrose, serine, glycine and GABA (Table 4) (Nikiforova *et al.*, 2006; Hammond *et al.*, 2008).

Table 4. Contents of phytic acid (mg/g of dry matter) and relative peak intensities of selected metabolites in wild-type MH86 and *lpa* mutant *Os-lpa*-MH86-1 harvested in Hangzhou 2011. [a,b,c]

metabolite	MH86	Os-lpa-MH86-1		metabolite	MH86	Os-lpa-MH86-1	
phytic acid	11.4±0.1	6.6±0.1	*	glycine	6.1±0.4	7.2±0.6	
phosphoric acid	56.2±1.6	616±7.7	*	alanine	19.8±0.5	27.1±0.8	*
citric acid	46.7±1.4	168±19.3	*	serine	12.2±0.2	18.5±0.7	*
				valine	5.5±0.2	6.8±0.4	*
glucose	19.5±0.3	33.5±0.2	*	leucine	2.6±0.2	3.6±0.4	*
fructose	8.4±0.1	16.3±0.2	*	isoleucine	6.0±0.3	9.6±0.6	*
sucrose	1538±38.3	1798±23.2	*	threonine	3.6±0.2	4.9±0.3	*
trehalose	4.5±0.2	9.3±0.4	*	cysteine	0.6±0.1	0.4±0.1	
myo-inositol	10.3±0.3	11.0±0.9		aspartic acid	31.5±1.4	39.9±0.8	*
sorbitol	2.4±0.1	7.9±0.1	*	asparagine	72.8±1.3	76.8±2.9	
mannitol	4.0±0.0	8.0±0.1	*	lysine	3.1±0.7	4.8±1.1	
				ornithine	0.5±0.1	0.8±0.3	*
ß-sitosterol	188±3.2	226±16.1	*	citrulline	0.4±0.1	0.5±0.1	
stigmasterol	48.1±0.9	63.2±4.9	*	tryptophan	5.1±0.1	7.7±0.4	*
campesterol	61.6±1.4	72.1±4.4	*	GABA [d]	3.1±0.3	11.3±1.0	*

[a] The Table is based on the metabolite profiling data elaborated by Zhao *et al.* 2016 (permission from John Wiley & Sons, Copyright 2016). [b] The phytic acid contents are expressed in mg/g dry matter. All other metabolites are expressed as relative peak intensities, i.e. metabolite peak intensity / internal standard peak intensity $\times$ 100. [c] Values represent means ± standard deviations resulting from the analysis of three aliquots of freeze-dried flour. Asterisks represent statistically significant differences (Benjamini-Hochberg adjusted-$p < 0.05$) between the wild-type and the corresponding *lpa* mutant. [d] *γ*-aminobutyric acid.

It is noteworthy that the mutation-induced metabolite signature observed in the *Os-lpa*-MH86-1 *lpa* mutant from the field trial Hangzhou 2011 was also consistently expressed at other field trials (Hainan 2012, Hainan 2013 and Jiaxing 2013). Accordingly, a clear separation in the PCA score plot of metabolite profiling data between the wild-type MH86 and the *Os-lpa*-MH86-1 *lpa* mutant from different field trials was achieved (Zhao *et al.*, 2016; Figure 6). These results suggested that *Os-lpa*-MH86-1 would be an ideal *lpa* mutant candidate to investigate the impact of cross-breeding with commercial rice cultivars on the stability of a mutation-induced metabolite signature in homozygous *lpa* mutant progenies.
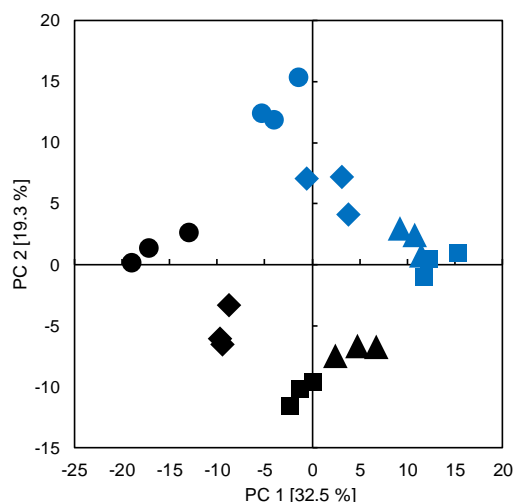
Figure 6. PCA score plot of metabolite profiling data of combined fractions I-IV of the wild-type MH86 (black) and the *Os-lpa*-MH86-1 *lpa* mutant (blue) harvested at the field trials Hangzhou 2011 (diamonds), Hainan 2012 (triangles), Hainan 2013 (squares) and Jiaxing 2013 (circles). The Figure is based on the metabolite profiling data elaborated by Zhao *et al.* 2016 (permission from John Wiley & Sons, Copyright 2016).

4.3 Crossing parents

In the present study, the two commercial rice cultivars Jiaxian 081 (JX081) and Jiahe 218 (JH218) were used as crossing parents. They are known to exhibit good agronomic performance with yields of 8-9 tons/ha and meet the Chinese national quality standards. In addition, another wild-type rice cultivar JH99, the restorer line of the hybrid cultivar Jiayou 99, was employed.

Results of preliminary studies showed that these commercial rice cultivars differed in their phytic acid contents. JX081 consistently exhibited the highest PA contents at several field trials, followed by JH218 and JH99. Therefore, they were considered suitable to investigate the impact of the phytic acid content of the crossing partner on the phytic acid contents and the percentage phytic acid reductions, respectively, in the homozygous *lpa* progenies obtained after cross-breeding with *lpa* mutants. For the *Os-lpa*-XQZ-1 *lpa* mutant there was no consistent correlation between the PA contents of the crossing parents (JX081 and JH218) and those of the corresponding homozygous *lpa* mutant progenies. On the contrary, for *Os-lpa*-XS110-2 the same order of differences in PA contents between the two crossing parents JX081 and JH218 were also observed between the corresponding *lpa* progenies. This phenomenon was even more pronounced for the *lpa* progenies resulting from the *Os-lpa*-MH86-1 *lpa* mutant.

Considering the simple and direct disruption of the biosynthetic flow to phytic acid in the *Os-lpa*-XQZ-1 mutant, it is not surprising that the PA contents in the homozygous *lpa* progenies were rather stable upon cross-breeding. On the contrary, the complex mechanism underlying the PA reduction in *Os-lpa*-MH86-1 is reflected in the increased susceptiblity of the PA content to that of the crossing parent, compared with *Os-lpa*-XQZ-1 and *Os-lpa*-XS110-2. This is also in line with the reported significantly changed expression levels of genes involved in phosphate and sulfate homeostasis (Zhao *et al.*, 2016).

Rice cultivars are classified into the two subspecies *indica* and *japonica* according to their geographical distribution, plant architecture and physiological features (Ikehashi, 2009). The rice seeds from these two subspecies are varying largely in genetic backgrounds (Garris *et al.*, 2005; Zhang *et al.*, 2011) as well as in metabolic compositions (Lou *et al.*, 2011; Hu *et al.*, 2014). Therefore, both an *indica* (JH99) and a *japonica* (JH218) rice type were selected as crossing parents.

The results revealed that the lipid profiles of the crossing parents had a significant impact on the lipophilic constituents of the resulting *lpa* mutant progenies. A number of individual lipophilic metabolites were shown to exhibit higher levels in JH99 compared to JH218. Accordingly, the homozygous *lpa* mutant progenies resulting from JH99 also exhibited higher levels of these lipophilic metabolites than the *lpa* progenies from JH218. This demonstrated that for the *Os-lpa*-MH86-1 *lpa* mutant the lipophilic constituents of the crossbred *lpa* progenies are dependent on the lipid profiles of the crossing parent.

However, despite the fact that substantial metabolic differences were observed between the crossing parents JH99 and JH218, the data do not allow to conclude whether the distinctions of metabolite profiles in the *lpa* progenies result from the inherent differences between the rice subspecies *indica* and *japonica*, or whether they just reflect individual differences between the two rice cultivars selected for this study.

4.4 Environmental variability

The PA contents in rice seeds are significantly influenced by the growing environment (Perera *et al.*, 2018). For example, increases in the levels of PA from 4 % to 31 %,

were observed for rice seeds grown in controlled environment facilities at higher temperature (daily mean temperature 33 °C vs. 22 °C) (Su *et al.*, 2014). In addition, there was also a pronounced impact of growing environment on the metabolite profiles of rice seeds (Huo *et al.*, 2017). Therefore, in order to investigate the environmental influence on the PA contents as well as on the metabolite profiles of the homozygous *lpa* mutant progenies, the rice plants were grown at three field trials (Jiaxing 2014, Wuxi 2014 and Hainan 2015). The climate conditions of these three locations were distinct due to the significant geographic differences. Jiaxing (latitude 30°45′ N) and Wuxi (31°29′ N) are located in the subtropical zone and Hainan in the tropical zone (latitude 18°30′ N). Therefore, the rice plants grown in Hainan were exposed to higher temperature and longer sunshine duration in the course of cultivation.

Again, the PA contents in the homozygous *lpa* mutant progenies were expressed differently depending on mutation types in response to the impact factor growing environment. For *Os-lpa*-XQZ-1, the variations of PA contents in homozygous *lpa* mutant progenies were rather low in light of the environmental influence. This variability was increased for the *Os-lpa*-XS110-2 progenies, and was most significantly expressed for the *lpa* progenies resulting from the *Os-lpa*-MH86-1 *lpa* mutant. The different variabilities of PA contents for three mutation-types in response to environments were similar to those in response to the crossing parents, which further confirmed that the susceptibilities of PA contents in crossbred *lpa* mutant progenies are correlated with the complexities of the mechanisms underlying the *lpa* mutations.

Growing environment has been shown to be the most predominant factor driving the variations of the metabolite profiles of cereal grains such as rice (Zhang *et al.*, 2016), maize (Frank *et al.*, 2012b) and barley (Wenzel *et al.*, 2015). Similarly, in the present study the environment had a far more pronounced influence on the metabolite profiles of homozygous *lpa* mutant progenies than the crossing parent and the mutation-type. The most remarkable differentiations of metabolite profiling data in the PCA score plot were observed between the rice seeds harvested from the tropical zone (Hainan) and those from the subtropical zone (Jiaxing and Wuxi), irrespective of the mutation-type or the genetic background inherited from the different crossing parents. In addition, the separation of metabolite profiles between rice seeds still remained significant in the light of environmental impact, even when only the field trials Jiaxing and Wuxi from the same climate zone were considered. Univariate analysis demonstrated that different growing environments induced changed levels of a broad spectrum of metabolites in

crossbred progenies both for wild-type and *lpa* mutants, ranging from lipophilic metabolites, e.g. fatty acid methyl esters and free fatty acids, to polar metabolites such as sugars, sugar alcohols, organic acids and amino acids.

4.5 Agricultural significance of the elaborated data

Phytic acid contents

From a breeder's point of view, it would be an essential requirement that despite the observed fluctuations of the phytic acid contents in the *lpa* mutant progenies the intended *lpa* trait, i.e. the significant reduction of the content of phytic acid compared with the original wild-type, remained unaffected. Figure 7, showing the distributions of the phytic acid contents in all rice seeds samples investigated in this study, demonstrates that this crucial prerequisite is fulfilled.
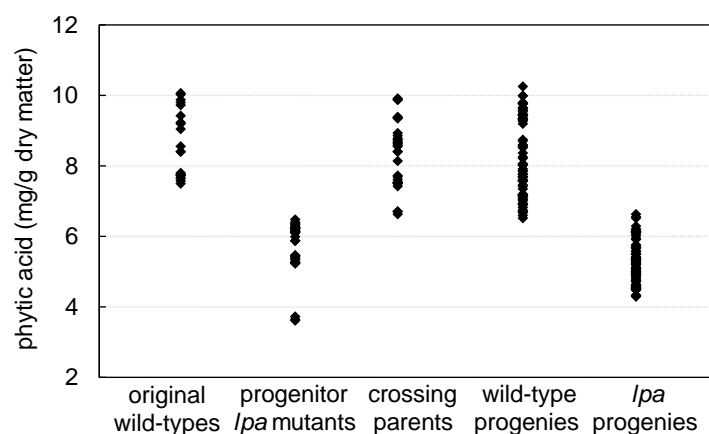
Figure 7. Distributions of the phytic acid contents in all rice samples investigated in this study

The data encompass the phytic acid contents in the original wild-types, the commercial wild-type cultivars employed as crossing parents, the progenitor *lpa* mutants, the homozygous wild-type and the homozygous *lpa* mutant progenies of generations F_4 to F_8, harvested at four different field trials (Jiaxing 2013, Jiaxing 2014, Wuxi 2014 and Hainan 2015). The data demonstrate that the obtained homozygous *lpa* mutant progenies always exhibited significantly lower levels of phytic acid than the original wild-types, independent from mutation types, cross-breeding steps, generations, crossing parents or environmental conditions. The phytic acid contents in the *lpa* progenies were in a similar order of magnitude as in the progenitor *lpa* mutants.

Similarly, a parallel study on soybeans also showed that after crossing the *lpa* soybean mutants TW-1-M-*lpa* (mutation of the *D-myo*-inositol 3-phosphate synthase gene 1, *MIPS1*) and ZC-*lpa* (mutation of the inositol 1,3,4,5,6-penta*kis*phosphate 2-kinase, *IPK1*) with commercial wild-type cultivars, the *lpa* trait was still consistently expressed in the resulting homozygous *lpa* mutant progenies over several generations at different field trials (Goßner *et al.*, 2019).

For future developments of *lpa* rice mutants, another aspect might be of relevance. In this study, the percentage reductions of PA contents in *lpa* rice mutants with single mutation targets ranged from -19.5% to -46.3%. For *lpa* mutants carrying double mutation targets, much more pronounced PA reductions have been reported. For example, double *lpa* soybean mutants with two *IPK1* mutation targets on chromosomes 6 and 14 showed drastic PA reductions between 79% and 88% (Vincent *et al.*, 2015). Another double *lpa* mutant generated by crossing of two single *lpa* soybean mutants ZC-*lpa* (*IPK1* mutation) and TW-1-M-*lpa* (*MIPS1* mutation) exhibited up to 87% of PA reductions (Goßner *et al.*, 2019).

The development and the investigation of double *lpa* rice mutants have never been reported. It is foreseeable that the generation of double *lpa* mutants of rice by crossing of two single *lpa* rice mutants could be an efficient way to generate rice seeds exhibiting PA contents at only very low levels.

Metabolite profiles

The *OsSULTR3;3* mutation-induced metabolite signature (i.e., increased levels of sugars, sugar alcohols, organic acids, amino acids, and biogenic amines) was nearly unaffected by the crossing-breeding of *Os-lpa*-MH86-1 *lpa* mutant with the commercial rice cultivar JH99. Despite the variations in the levels of lipophilic constituents depending on generations, the mutation-specific metabolite signature was consistently observed in homozygous *lpa* mutant progenies of generations F_4 to F_7 at the same field trial Jiaxing 2013.

The results revealed that the growing environment played a more important role in driving the variability of the metabolite profiles of crossbred progenies than the mutation-type and the crossing parents. As a result, the absolute levels of a broad spectrum of metabolites, ranging from lipophilic to polar constituents (e.g. hydrocarbons, fatty acid methyl esters, free fatty acids, sugars, sugar alcohols, organic

acids and amino acids), were increased or decreased in crossbred progenies depending on the different field trials (i.e. Jiaxing 2014, Wuxi 2014 and Hainan 2015), irrespective of the mutation effect and the genetic backgrounds of the crossing parents. Notwithstanding the significant metabolic changes in crossbred progenies resulting from environmental impact, for each field trial the *OsSULTR3;3* mutation-induced metabolic differences were consistently observed between the homozygous wild-type and the corresponding homozygous *lpa* mutant progenies of generation F_8 for both the crosses *Os-lpa*-MH86-1 × JH99 and *Os-lpa*-MH86-1 × JH218. In addition to the mutation-induced metabolic changes mainly in polar metabolites, the lipophilic compositions in *lpa* progenies were shown to be influenced by the lipid profiles of the crossing parents. The *lpa* progenies resulting from the crossing parent with lipid-enriched profiles also exhibited higher levels of individual lipophilic constituents. However, this did not compromise the mutation-specific metabolic differentiations between the wild-type and the *lpa* mutant progenies for each cross. As a result, a superimposition effect was observed for the overall metabolite profiles of the crossbred *lpa* mutant progenies, i.e. the impact of mutation effect on polar metabolites and of the crossing parents on lipophilic metabolites.

The stability of the mutation-induced metabolite signature in *lpa* mutants is depicted in Figure 8. For each metabolite investigated at individual field trials, the boxplots for the wild-types represent the data resulting from both the original wild-type MH86 as well as the homozygous wild-type progenies of the two crosses; those for the *lpa* mutants represent the data from both the progenitor *lpa* mutant *Os-lpa*-MH86-1 and the homozygous *lpa* progenies of the two crosses. All listed metabolites exhibited statistically significantly different levels between wild-type and *lpa* mutant rice seeds at each field trial, and these differences are in good agreement with the results previously reported in the progenitor *lpa* mutant *Os-lpa*-MH86-1 (Zhao *et al.*, 2016). This key result demonstrated that the *OsSULTR3;3* mutation-specific metabolite signature, including increased levels of nutritionally relevant compounds (e.g. *γ*-aminobutyric acid), was not affected by the crossing-breeding of *Os-lpa*-MH86-1 mutant with different commercial rice cultivars, and was consistently expressed over several generations as well as under different environmental conditions.
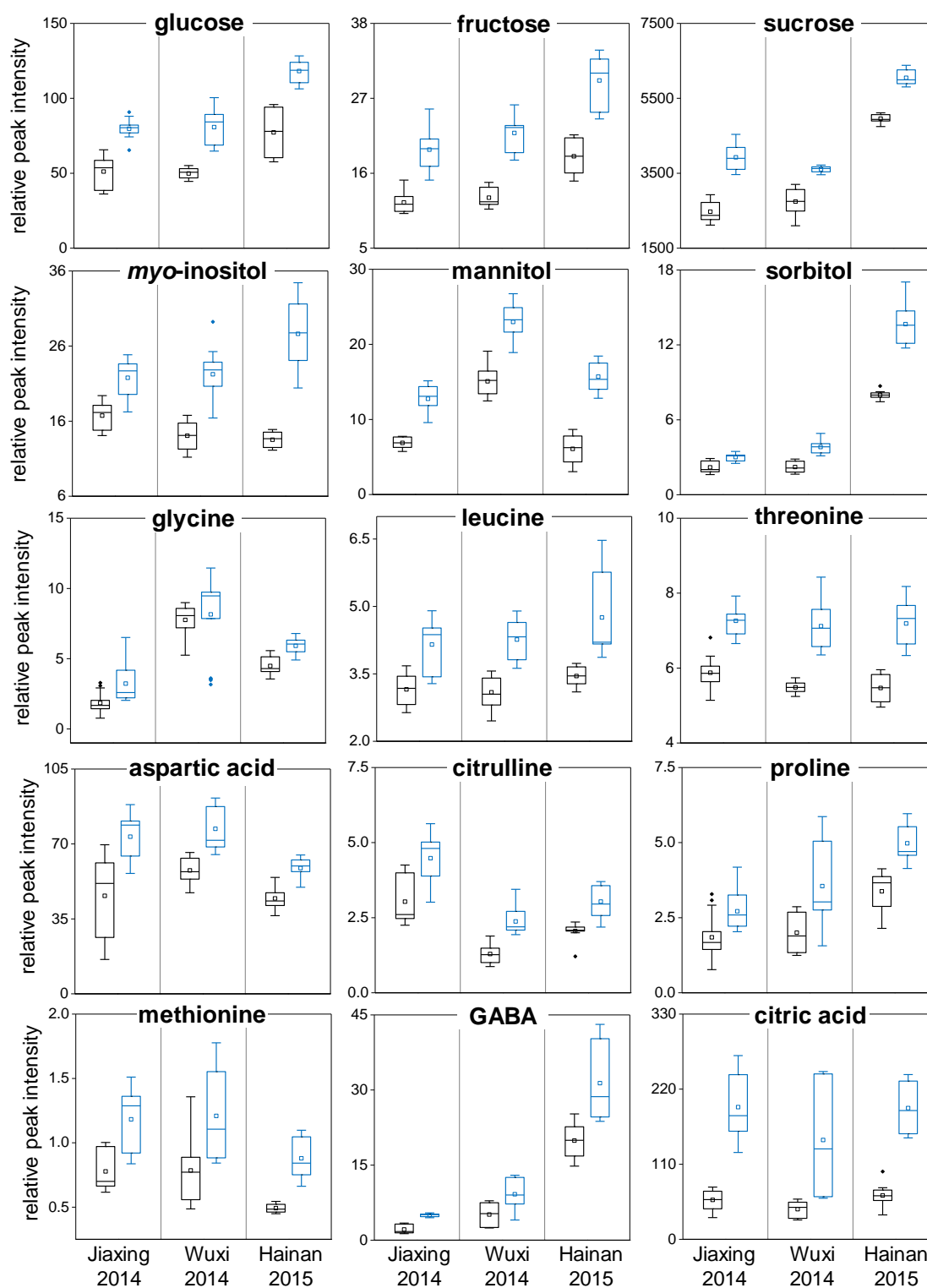
Figure 8. Boxplots of the metabolites contributing to the mutation-specific metabolite signature of *lpa* mutants. For each metabolite, the relative peak intensity (i.e., metabolite peak intensity/internal standard peak intensity ×100) of wild-type (black) represents the mean value of original wild-type MH86 and homozygous wild-type progenies of two crosses; the relative peak intensity of *lpa* mutant (blue) represents the mean value of *Os-lpa*-MH86-1 and homozygous *lpa* mutant progenies of two crosses. All metabolites exhibited statistically significantly different (Benjamini–Hochberg adjusted-$p < 0.05$) levels between wild-type and *lpa* mutant rice seeds at each field trial.

Of course, it is noteworthy that the results obtained in the present study should not be unconditionally extrapolated to any other *lpa* mutants. Other mutation-types might result in different effects upon cross-breeding and might be differently impacted by the crossing partner or by the environmental conditions. Nevertheless, the elaborated metabolite profiling data are encouraging for breeders. The mutants investigated in this study cover the main types of mutation presently known for *lpa* rice mutants. Even for such complex metabolic changes as induced in *Os-lpa*-MH86-1 by the *lpa* mutation, cross-breeding could be employed as a tool to generate progeny rice seeds stably exhibiting the mutation-specific metabolite signature.

4.6 Agronomic performance

One of the reasons for crossing induced *lpa* mutants with commercial cultivars is to overcome potential agronomic deficiencies (Zhao *et al.*, 2008a). Therefore, it would have been valuable to have comprehensive data regarding parameters such as germination rate, field emergence and yield for the generated homozygous *lpa* progenies. Unfortunately, for each cross and each generation only a low number of rice plants were obtained which did not allow an appropriate evaluation of the agronomic performance of the *lpa* mutant progenies. The limited data available indicated that after cross-breeding with the commercial rice cultivars, no significant improvements in thousand grain weights and seed setting ratios were observed for the *lpa* rice mutants. However, these results should only be considered as preliminary. Agronomic data obtained from large-scale field trials would be necessary to elaborate robust data allowing to draw valid conclusions.

Such agronomic data on the *lpa* progenies are definitely required in follow-up studies. In addition to the results elaborated in this thesis, i.e. that the *lpa* trait is stably inherited and the mutation-specific metabolite signature is consistently expressed in crossbred *lpa* mutant progenies, the demonstration of good agronomic performance would be another important cornerstone to demonstrate the usefulness of cross-breeding of progenitor *lpa* mutants with commercial cultivars to develop promising elite *lpa* rice cultivars.

5 REFERENCES

Abdel-Megeed, A.; Tahir, A., Reduction of phosphorus pollution from broilers waste through supplementation of wheat based broilers feed with phytase. *J Chem* **2015,** *3,* 1-3.

Ali, N.; Paul, S.; Gayen, D.; Sarkar, S. N.; Datta, K.; Datta, S. K., Development of low phytate rice by RNAi mediated seed-specific silencing of Inositol 1,3,4,5,6-Penta*kis*phosphate 2-kinase gene (*IPK1*). *PLoS One* **2013,** *8.*

Allwood, J. W.; Goodacre, R., An introduction to liquid chromatography-mass spectrometry instrumentation applied in plant metabolomic analyses. *Phytochem. Anal.* **2010,** *21*, 33-47.

Alonso, A.; Marsal, S.; Julià, A., Analytical methods in untargeted metabolomics: state of the art in 2015. *Frontiers in bioengineering and biotechnology* **2015,** *3*, 23.

Antignac, J. P.; de Wasch, K.; Monteau, F.; De Brabander, H.; Andre, F.; Le Bizec, B., The ion suppression phenomenon in liquid chromatography-mass spectrometry and its consequences in the field of residue. *Anal. Chim. Acta* **2005,** *529*, 129-136.

Ariza-Nieto, M.; Blair, M. W.; Welch, R. M.; Glahn, R. P., Screening of iron bioavailability patterns in eight bean (*Phaseolus vulgaris* L.) genotypes using the Caco-2 cell in vitro model. *J. Agric. Food Chem.* **2007,** *55*, 7950-7956.

Armitage, E. G.; Godzien, J.; Alonso-Herranz, V.; Lopez-Gonzalvez, A.; Barbas, C., Missing value imputation strategies for metabolomics data. *Electrophoresis* **2015,** *36*, 3050-3060.

Averitt, B.; Shang, C.; Rosso, L.; Qin, J.; Zhang, M. C.; Rainy, K. M.; Zhang, B., Impact of *mips1*, *lpa1*, and *lpa2* alleles for low phytic acid content on agronomic, seed quality, and seed composition traits of soybean. *Crop Sci.* **2017,** *57*, 2490-2499.

Balogh, M. P., Debating resolution land mass accuracy in mass spectrometry. *Spectroscopy* **2004,** *19*, 34-40.

Barker, M.; Rayens, W., Partial least squares for discrimination. *J. Chemometrics* **2003,** *17*, 166-173.

Beckles, D. M.; Roessner, U., Plant metabolomics: Applications and opportunities for agricultural biotechnology. In *Plant biotechnology and agriculture*, Elsevier: 2012; pp 67-81.

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

Bhati, K. K.; Alok, A.; Kumar, A.; Kaur, J.; Tiwari, S.; Pandey, A. K., Silencing of *ABCC13* transporter in wheat reveals its involvement in grain development, phytic acid accumulation and lateral root formation. *JExB* **2016,** *67*, 4379-4389.

Bijlsma, S.; Bobeldijk, L.; Verheij, E. R.; Ramaker, R.; Kochhar, S.; Macdonald, I. A.; van Ommen, B.; Smilde, A. K., Large-scale human metabolomics studies: A strategy for data (pre-) processing and validation. *AnaCh* **2006,** *78*, 567-574.

Bingol, K., Recent Advances in Targeted and Untargeted Metabolomics by NMR and MS/NMR Methods. *High-Throughput* **2018,** *7*, 9.

Boiteau, R. M.; Hoyt, D. W.; Nicora, C. D.; Kinmonth-Schultz, H. A.; Ward, J. K.; Bingol, K., Structure Elucidation of Unknown Metabolites in Metabolomics by Combined NMR and MS/MS Prediction. *Metabolites* **2018,** *8*.

Bottcher, C.; Krahmer, A.; Sturtz, M.; Widder, S.; Schulz, H., Effect of Cultivar and Cultivation Year on the Metabolite Profile of Onion Bulbs (Allium cepa L.). *J. Agric. Food Chem.* **2018,** *66*, 3229-3238.

Brasil, I. M.; Siddiqui, M. W., Postharvest Quality of Fruits and Vegetables: An Overview. In *Preharvest Modulation of Postharvest Fruit and Vegetable Quality*, Elsevier: 2018; pp 1-40.

Brizzolara, S.; Santucci, C.; Tenori, L.; Hertog, M.; Nicolai, B.; Sturz, S.; Zanella, A.; Tonutti, P., A metabolomics approach to elucidate apple fruit responses to static and dynamic controlled atmosphere storage. *Postharvest Biol. Technol.* **2017,** *127*, 76-87.

Broadhurst, D. I.; Kell, D. B., Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics* **2006,** *2*, 171-196.

Brown, S. C.; Kruppa, G.; Dasseux, J. L., Metabolomics applications of FT-ICR mass spectrometry. *Mass Spectrom. Rev.* **2005,** *24*, 223-231.

Bylesjo, M.; Rantalainen, M.; Cloarec, O.; Nicholson, J. K.; Holmes, E.; Trygg, J., OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *J. Chemometrics* **2006,** *20*, 341-351.

Campion, B.; Sparvoli, F.; Doria, E.; Tagliabue, G.; Galasso, I.; Fileppi, M.; Bollini, R.; Nielsen, E., Isolation and characterisation of an *lpa* (*low phytic acid*) mutant in common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* **2009,** *118*, 1211-1221.

Cerino Badone, F.; Amelotti, M.; Cassani, E.; Pilu, R., Study of *low phytic acid1-7* (*lpa1-7*), a new *ZmMRP4* mutation in maize. *J. Hered.* **2012,** *103*, 598-605.

Chen, Q. C., Determination of phytic acid and inositol pentakisphosphates in foods by high-performance ion chromatography. *J. Agric. Food Chem.* **2004,** *52*, 4604-4613.

Chen, Q. C.; Li, B. W., Separation of phytic acid and other related inositol phosphates by high-performance ion chromatography and its applications. *J. Chromatogr.* **2003,** *1018*, 41-52.

Craig, A.; Cloareo, O.; Holmes, E.; Nicholson, J. K.; Lindon, J. C., Scaling and normalization effects in NMR spectroscopic metabonomic data sets. *AnaCh* **2006,** *78*, 2262-2267.

Cubero-Leon, E.; De Rudder, O.; Maquet, A., Metabolomics for organic food authentication: Results from a long-term field study in carrots. *Food Chem.* **2018,** *239*, 760-770.

Dettmer, K.; Aronov, P. A.; Hammock, B. D., Mass spectrometry-based metabolomics. *Mass Spectrom. Rev.* **2007,** *26*, 51-78.

Di Guida, R.; Engel, J.; Allwood, J. W.; Weber, R. J. M.; Jones, M. R.; Sommer, U.; Viant, M. R.; Dunn, W. B., Non-targeted UHPLC-MS metabolomic data processing methods: a comparative investigation of normalisation, missing value imputation, transformation and scaling. *Metabolomics* **2016,** *12*.

Dorsch, J. A.; Cook, A.; Young, K. A.; Anderson, J. M.; Bauman, A. T.; Volkmann, C. J.; Murthy, P. P. N.; Raboy, V., Seed phosphorus and inositol phosphate phenotype of barley *low phytic acid* genotypes. *Phytochemistry* **2003,** *62*, 691-706.

Douglas, D. J., Linear Quadrupoles in Mass Spectrometry. *Mass Spectrom. Rev.* **2009,** *28*, 937-960.

Douglas, D. J.; Frank, A. J.; Mao, D. M., Linear ion traps in mass spectrometry. *Mass Spectrom. Rev.* **2005,** *24*, 1-29.

Dunn, W. B.; Bailey, N. J. C.; Johnson, H. E., Measuring the metabolome: current analytical technologies. *Analyst* **2005,** *130*, 606-625.

Dunn, W. B.; Broadhurst, D.; Begley, P.; Zelena, E.; Francis-McIntyre, S.; Anderson, N.; Brown, M.; Knowles, J. D.; Halsall, A.; Haselden, J. N.; Nicholls, A. W.; Wilson, I. D.; Kell, D. B.; Goodacre, R.; C, H. S. M. H., Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* **2011,** *6*, 1060-1083.

Eriksson, L.; Antti, H.; Gottfries, J.; Holmes, E.; Johansson, E.; Lindgren, F.; Long, I.; Lundstedt, T.; Trygg, J.; Wold, S., Using chemometrics for navigating in the large data sets of genomics, proteomics, and metabonomics (gpm). *Anal. Bioanal. Chem.* **2004,** *380*, 419-429.

FAO (Food and Agriculture Organization of the United Nations). http://www.fao.org/faostat/en/#data/QC/visualize (accessed 16 December 2017).

FAO/IAEA (Food and Agriculture Organization of the United Nations/International Atomic Energy Agency) Mutant Variety Database. https://mvd.iaea.org/ (accessed 10 October, 2018).

Feng, X.; T. Yoshida, K., Molecular approaches for producing *low-phytic-acid* grains in rice. *Plant Biotechnol.* **2004,** *21*, 183-189.

Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M., Electrospray Ionization for Mass-Spectrometry of Large Biomolecules. *Sci* **1989,** *246*, 64-71.

Fiehn, O., Metabolomics - the link between genotypes and phenotypes. *Plant Mol. Biol.* **2002,** *48*, 155-171.

Fiehn, O.; Kopka, J.; Dormann, P.; Altmann, T.; Trethewey, R. N.; Willmitzer, L., Metabolite profiling for plant functional genomics. *Nat. Biotechnol.* **2000,** *18*, 1157-1161.

Frank, T.; Habernegg, R.; Yuan, F. J.; Shu, Q. Y.; Engel, K. H., Assessment of the contents of phytic acid and divalent cations in *low phytic acid* (*lpa*) mutants of rice and soybean. *J. Food Compost. Anal.* **2009a,** *22*, 278-284.

Frank, T.; Meuleye, B. S.; Miller, A.; Shu, Q. Y.; Engel, K. H., Metabolite profiling of two *low phytic acid* (*lpa*) rice mutants. *J. Agric. Food Chem.* **2007,** *55*, 11011-11019.

Frank, T.; Nörenberg, S.; Engel, K. H., Metabolite profiling of two novel *low phytic acid* (*lpa*) soybean mutants. *J. Agric. Food Chem.* **2009b,** *57*, 6408-6416.

Frank, T.; Reichardt, B.; Shu, Q. Y.; Engel, K. H., Metabolite profiling of colored rice (*Oryza sativa* L.) grains. *J Cereal Sci* **2012a,** *55*, 112-119.

Frank, T.; Rohlig, R. M.; Davies, H. V.; Barros, E.; Engel, K. H., Metabolite Profiling of Maize Kernels-Genetic Modification versus Environmental Influence. *J. Agric. Food Chem.* **2012b,** *60*, 3005-3012.

Frank, T.; Scholz, B.; Peter, S.; Engel, K. H., Metabolite profiling of barley: Influence of the malting process. *Food Chem.* **2011,** *124*, 948-957.

Frenzel, T.; Miller, A.; Engel, K. H., Metabolite profiling - A Fractionation method for analysis of major and minor compounds in rice grains. *Cereal Chem.* **2002,** *79*, 215-221.

Frenzel, T.; Miller, A.; Engel, K. H., A methodology for automated comparative analysis of metabolite profiling data. *Eur. Food Res. Technol.* **2003,** *216*, 335-342.

Furey, A.; Moriarty, M.; Bane, V.; Kinsella, B.; Lehane, M., Ion suppression; A critical review on causes, evaluation, prevention and applications. *Talanta* **2013,** *115*, 104-122.

Gao, Y.; Biyashev, R. M.; Maroof, M. A. S.; Glover, N. M.; Tucker, D. M.; Buss, G. R., Validation of low-phytate QTLs and evaluation of seedling emergence of low-phytate soybeans. *Crop Sci.* **2008,** *48*, 1355-1364.

Garcia, C. J.; Garcia-Villalba, R.; Garrido, Y.; Gil, M. I.; Tomas-Barberan, F. A., Untargeted metabolomics approach using UPLC-ESI-QTOF-MS to explore the metabolome of fresh-cut iceberg lettuce. *Metabolomics* **2016,** *12*.

Garris, A. J.; Tai, T. H.; Coburn, J.; Kresovich, S.; McCouch, S., Genetic structure and diversity in *Oryza sativa* L. *Genetics* **2005,** *169*, 1631-1638.

Gentleman, R. C.; Carey, V. J.; Bates, D. M.; Bolstad, B.; Dettling, M.; Dudoit, S.; Ellis, B.; Gautier, L.; Ge, Y. C.; Gentry, J.; Hornik, K.; Hothorn, T.; Huber, W.; Iacus, S.; Irizarry, R.; Leisch, F.; Li, C.; Maechler, M.; Rossini, A. J.; Sawitzki, G.; Smith, C.; Smyth, G.; Tierney, L.; Yang, J. Y. H.; Zhang, J. H., Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **2004,** *5*.

Ghaste, M.; Mistrik, R.; Shulaev, V., Applications of Fourier Transform Ion Cyclotron Resonance (FT-ICR) and Orbitrap Based High Resolution Mass Spectrometry in Metabolomics and Lipidomics. *Int. J. Mol. Sci.* **2016,** *17*.

Glauser, G.; Boccard, J.; Wolfender, J. L.; Rudaz, S., Metabolomics: application in plant sciences. *Metabolomics in Practice: Successful Strategies to Generate and Analyze Metabolic Data* **2013**, 313-343.

Goodacre, R.; Vaidyanathan, S.; Dunn, W. B.; Harrigan, G. G.; Kell, D. B., Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends Biotechnol.* **2004,** *22*, 245-252.

Goßner, S.; Yuan, F. J.; Zhou, C. G.; Tan, Y. Y.; Shu, Q. Y.; Engel, K. H., Impact of cross-breeding of *low phytic acid MIPS1* and *IPK1* soybean (*Glycine max* L. Merr.) mutants on their contents of inositol phosphate isomers. *J. Agric. Food Chem.* **2019,** *67*, 247-257.

Gromski, P.; Xu, Y.; Kotze, H.; Correa, E.; Ellis, D.; Armitage, E.; Turner, M.; Goodacre, R., Influence of Missing Values Substitutes on Multivariate Analysis of Metabolomics Data. *Metabolites* **2014,** *4*, 433.

Guilhaus, M.; Selby, D.; Mlynski, V., Orthogonal acceleration time-of-flight mass spectrometry. *Mass Spectrom. Rev.* **2000,** *19*, 65-107.

Guttieri, M.; Bowen, D.; Dorsch, J. A.; Raboy, V.; Souza, E., Identification and characterization of a *low phytic acid* wheat. *Crop Sci.* **2004,** *44*, 418-424.

Hackstadt, A. J.; Hess, A. M., Filtering for increased power for microarray data analysis. *BMC Bioinformatics* **2009,** *10*.

Hammond, J. P.; White, P. J., Sucrose transport in the phloem: integrating root responses to phosphorus starvation. *JExB* **2008,** *59*, 93-109.

Harland, B. F.; Oberleas, D., A modified method for phytate analysis using an ion-exchange procedure: Application to textured vegetable proteins. *Cereal Chem.* **1977,** *54*, 827-832.

Haug, K.; Salek, R. M.; Conesa, P.; Hastings, J.; de Matos, P.; Rijnbeek, M.; Mahendraker, T.; Williams, M.; Neumann, S.; Rocca-Serra, P.; Maguire, E.; Gonzalez-Beltran, A.; Sansone, S. A.; Griffin, J. L.; Steinbeck, C., MetaboLights-an

open-access general-purpose repository for metabolomics studies and associated meta-data. *NAR* **2013,** *41*, D781-D786.

Herrmann, A.; Schauer, N., Metabolomics-Assisted Plant Breeding. *The handbook of plant metabolomics* **2013**, 245-254.

Heubner, W.; Stadler, H., Über eine Titrationsmethode zur Bestimmung des Phytins. *Biochem. Z.* **1914,** *64*, 422–437.

Hitz, W. D.; Carlson, T. J.; Kerr, P. S.; Sebastian, S. A., Biochemical and molecular characterization of a mutation that confers a decreased raffinosaccharide and phytic acid phenotype on soybean seeds. *Plant Physiol.* **2002,** *128*, 650-660.

Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K.; Oda, Y.; Kakazu, Y.; Kusano, M.; Tohge, T.; Matsuda, F.; Sawada, Y.; Hirai, M. Y.; Nakanishi, H.; Ikeda, K.; Akimoto, N.; Maoka, T.; Takahashi, H.; Ara, T.; Sakurai, N.; Suzuki, H.; Shibata, D.; Neumann, S.; Iida, T.; Tanaka, K.; Funatsu, K.; Matsuura, F.; Soga, T.; Taguchi, R.; Saito, K.; Nishioka, T., MassBank: a public repository for sharing mass spectral data for life sciences. *J. Mass Spectrom.* **2010,** *45*, 703-714.

Hrydziuszko, O.; Viant, M. R., Missing values in mass spectrometry based metabolomics: an undervalued step in the data processing pipeline. *Metabolomics* **2012,** *8*, S161-S174.

Hu, C.; Shi, J.; Quan, S.; Cui, B.; Kleessen, S.; Nikoloski, Z.; Tohge, T.; Alexander, D.; Guo, L.; Lin, H.; Wang, J.; Cui, X.; Rao, J.; Luo, Q.; Zhao, X.; Fernie, A. R.; Zhang, D., Metabolic variation between *japonica* and *indica* rice cultivars as revealed by non-targeted metabolomics. *Sci. Rep.* **2014,** *4*, 5067.

Humer, E.; Schwarz, C.; Schedle, K., Phytate in pig and poultry nutrition. *J. Anim. Physiol. Anim. Nutr.* **2015,** *99*, 605-625.

Hung, W. L.; Wang, Y., Metabolite Profiling of Candidatus Liberibacter Infection in Hamlin Sweet Oranges. *J. Agric. Food Chem.* **2018,** *66*, 3983-3991.

Huo, Y. Q.; Kamal, G. M.; Wang, J.; Liu, H. L.; Zhang, G. N.; Hu, Z. Y.; Anwar, F.; Du, H. Y., NMR-based metabolomics for discrimination of rice from different geographical origins of China. *J Cereal Sci* **2017,** *76*, 243-252.

Ikehashi, H., Why are there indica type and japonica type in rice?—History of the studies and a view for origin of two types. *Rice Science* **2009,** *16*, 1-13.

Jahangir, M.; Abdel-Farid, I. B.; Mahmood, Z.; Jamil, M.; Rey, P. J.; Choi, Y. H.; Verpoorte, R., Metabolomic Approach: Postharvest Storage Stability of Red Radish (Raphanus sativus L.). *J. Chem. Soc. Pak.* **2014,** *36*, 931-938.

Jonsson, P.; Gullberg, J.; Nordstrom, A.; Kusano, M.; Kowalczyk, M.; Sjostrom, M.; Moritz, T., A strategy for identifying differences in large series of metabolomic samples analyzed by GC/MS. *AnaCh* **2004,** *76*, 1738-1745.

Jorge, T. F.; Mata, A. T.; Antonio, C., Mass spectrometry as a quantitative tool in plant metabolomics. *Philos T R Soc A* **2016,** *374*.

Jorge, T. F.; Rodrigues, J. A.; Caldana, C.; Schmidt, R.; van Dongen, J. T.; Thomas-Oates, J.; Antonio, C., Mass spectrometry-based plant metabolomics: Metabolite responses to abiotic stress. *Mass Spectrom. Rev.* **2015,** *35*, 620-649.

Kaling, M.; Kanawati, B.; Ghirardo, A.; Albert, A.; Winkler, J. B.; Heller, W.; Barta, C.; Loreto, F.; Schmitt-Kopplin, P.; Schnitzler, J. P., UV-B mediated metabolic rearrangements in poplar revealed by non-targeted metabolomics. *Plant Cell and Environment* **2015,** *38*, 892-904.

Katajamaa, M.; Miettinen, J.; Oresic, M., MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data. *Bioinformatics* **2006,** *22*, 634-636.

Katajamaa, M.; Oresic, M., Data processing for mass spectrometry-based metabolomics. *J. Chromatogr.* **2007,** *1158*, 318-328.

Kim, G.; Jung, E.; Lee, S.; Lim, S.-H.; Ha, S.-H.; Lee, C., Combined Mass Spectrometry-Based Metabolite Profiling of Different Pigmented Rice (Oryza sativa L.) Seeds and Correlation with Antioxidant Activities. *Molecules* **2015,** *19*, 15673.

Kim, H. K.; Choi, Y. H.; Verpoorte, R., NMR-based plant metabolomics: where do we stand, where do we go? *Trends Biotechnol.* **2011,** *29*, 267-275.

Kim, S. I.; Andaya, C. B.; Newman, J. W.; Goyal, S. S.; Tai, T. H., Isolation and characterization of a *low phytic acid* rice mutant reveals a mutation in the rice orthologue of maize *MIK. Theor. Appl. Genet.* **2008,** *117*, 1291-1301.

Kim, S. I.; Tai, T. H., Identification of novel rice *low phytic acid* mutations via TILLING by sequencing. *Mol. Breed.* **2014,** *34*, 1717-1729.

Kind, T.; Wohlgemuth, G.; Lee, D. Y.; Lu, Y.; Palazoglu, M.; Shahbaz, S.; Fiehn, O., FiehnLib: Mass Spectral and Retention Index Libraries for Metabolomics Based on Quadrupole and Time-of-Flight Gas Chromatography/Mass Spectrometry. *AnaCh* **2009,** *81*, 10038-10048.

Kopka, J., Current challenges and developments in GC-MS based metabolite profiling technology. *J. Biotechnol.* **2006,** *124*, 312-322.

Kopka, J.; Fernie, A.; Weckwerth, W.; Gibon, Y.; Stitt, M., Metabolite profiling in plant biology: platforms and destinations. *Genome Biol.* **2004,** *5*.

Kopka, J.; Schauer, N.; Krueger, S.; Birkemeyer, C.; Usadel, B.; Bergmuller, E.; Dormann, P.; Weckwerth, W.; Gibon, Y.; Stitt, M.; Willmitzer, L.; Fernie, A. R.; Steinhauser, D., The Golm Metabolome Database. *Bioinformatics* **2005,** *21*, 1635-1638.

Korman, A.; Oh, A.; Raskind, A.; Banks, D., Statistical Methods in Metabolomics. In *Evolutionary Genomics: Statistical and Computational Methods, Volume 2*, Anisimova, M., Ed. Humana Press: Totowa, NJ, 2012; pp 381-413.

Krishnan, P.; Kruger, N. J.; Ratcliffe, R. G., Metabolite fingerprinting and profiling in plants using NMR. *JExB* **2005,** *56*, 255-265.

Kumar, V.; Sinha, A. K.; Makkar, H. P. S.; Becker, K., Dietary roles of phytate and phytase in human nutrition: A review. *Food Chem.* **2010,** *120*, 945-959.

Kusano, M.; Saito, K., Role of metabolomics in crop improvement. *J. Plant Biochem. Biotechnol.* **2012,** *21*, 24-31.

Kuwano, M.; Ohyama, A.; Tanaka, Y.; Mimura, T.; Takaiwa, F.; Yoshida, K. T., Molecular breeding for transgenic rice with *low-phytic-acid* phenotype through manipulating *myo*-inositol 3-phosphate synthase gene. *Mol. Breed.* **2006,** *18*, 263-272.

Lanzinger, A.; Frank, T.; Reichenberger, G.; Herz, M.; Engel, K. H., Metabolite profiling of barley grain subjected to induced drought stress: responses of free amino acids in differently adapted cultivars. *J. Agric. Food Chem.* **2015,** *63*, 4252-4261.

Larive, C. K.; Barding, G. A.; Dinges, M. M., NMR Spectroscopy for Metabolomics and Metabolic Profiling. *AnaCh* **2015,** *87*, 133-146.

Larson, S. R.; Rutger, J. N.; Young, K. A.; Raboy, V., Isolation and genetic mapping of a non-lethal rice (*Oryza sativa* L.) *low phytic acid* 1 mutation. *Crop Sci.* **2000,** *40*, 1397-1405.

Lauxmann, M. A.; Borsani, J.; Osorio, S.; Lombardo, V. A.; Budde, C. O.; Bustamante, C. A.; Monti, L. L.; Andreo, C. S.; Fernie, A. R.; Drincovich, M. F.; Lara, M. V., Deciphering the metabolic pathways influencing heat and cold responses during post-harvest physiology of peach fruit. *Plant Cell and Environment* **2014,** *37*, 601-616.

Lehrfeld, J., Hplc separation and quantitation of phytic acid and some inositol phosphates in foods - problems and solutions. *J. Agric. Food Chem.* **1994,** *42*, 2726-2731.

Li, M. X.; Guo, R.; Jiao, Y.; Jin, X. F.; Zhang, H. Y.; Shi, L. X., Comparison of Salt Tolerance in Soja Based on Metabolomics of Seedling Roots. *Frontiers in Plant Science* **2017,** *8*.

Lim, S.; Lee, J. G.; Lee, E. J., Comparison of fruit quality and GC-MS-based metabolite profiling of kiwifruit 'Jecy green': Natural and exogenous ethylene-induced ripening. *Food Chem.* **2017,** *234*, 81-92.

Lindon, J. C.; Nicholson, J. K.; Holmes, E., *The handbook of metabonomics and metabolomics.* Elsevier: 2007.

Lisec, J.; Schauer, N.; Kopka, J.; Willmitzer, L.; Fernie, A. R., Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nat. Protoc.* **2006,** *1*, 387-396.

Liu, Q. L.; Xu, X. H.; Ren, X. L.; Fu, H. W.; Wu, D. X.; Shu, Q. Y., Generation and characterization of *low phytic acid* germplasm in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* **2007,** *114*, 803-814.

Lonnerdal, B.; Sandberg, A. S.; Sandstrom, B.; Kunz, C., Inhibitory effects of phytic acid and other inositol phosphates on zinc and calcium-absorption in suckling rats. *J. Nutr.* **1989,** *119*, 211-214.

Lott, J. N. A.; Ockenden, I.; Raboy, V.; Batten, G. D., Phytic acid and phosphorus in crop seeds and fruits: a global estimate. *Seed Sci Res* **2000,** *10*, 11-33.

Lou, Q.; Ma, C.; Wen, W.; Zhou, J.; Chen, L.; Feng, F.; Xu, X.; Lu, X.; Luo, L.; Mei, H.; Xu, G., Profiling and Association Mapping of Grain Metabolites in a Subset of the Core Collection of Chinese Rice Germplasm (Oryza sativa L.). *J. Agric. Food Chem.* **2011,** *59*, 9257-9264.

Luengwilai, K.; Saltveit, M.; Beckles, D. M., Metabolite content of harvested Micro-Tom tomato (Solanum lycopersicum L.) fruit is altered by chilling and protective heat-shock treatments as shown by GC-MS metabolic profiling. *Postharvest Biol. Technol.* **2012,** *63*, 116-122.

March, R. E., Quadrupole Ion Traps. *Mass Spectrom. Rev.* **2009,** *28*, 961-989.

Mastrangelo, A.; Ferrarini, A.; Rey-Stolle, F.; Garcia, A.; Barbas, C., From sample treatment to biomarker discovery: A tutorial for untargeted metabolomics based on GC-(EI)-Q-MS. *Anal. Chim. Acta* **2015,** *900*, 21-35.

McCance, R. A.; Widdowson, E. M., Phytin in human nutrition. *Biochem. J.* **1935,** *29*, 2694.

Meis, S. J.; Fehr, W. R.; Schnebly, S. R., Seed source effect on field emergence of soybean lines with reduced phytate and raffinose saccharides. *Crop Sci.* **2003,** *43*, 1336-1339.

Mendoza, C.; Viteri, F. E.; Lonnerdal, B.; Young, K. A.; Raboy, V.; Brown, K. H., Effect of genetically modified, *low phytic acid* maize on absorption of iron from tortillas. *Am. J. Clin. Nutr.* **1998,** *68*, 1123-1127.

Moritz, F.; Forcisi, S.; Harir, M.; Kanawati, B.; Lucio, M.; Tziotis, D.; Schmitt-Kopplin, P., The Potential of Ultrahigh Resolution MS (FTICR-MS) in Metabolomics. *Metabolomics in Practice: Successful Strategies to Generate and Analyze Metabolic Data* **2013**, 117-136.

Moritz, T.; Johansson, A. I., Plant Metabolomics. In *Metabolomics, Metabonomics and Metabolite Profiling*, The Royal Society of Chemistry: 2008; pp 254-272.

Na Jom, K.; Frank, T.; Engel, K. H., A metabolite profiling approach to follow the sprouting process of mung beans (*Vigna radiata*). *Metabolomics* **2011,** *7*, 102-117.

Nicholson, J. K.; Lindon, J. C.; Holmes, E., 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* **1999,** *29*, 1181-1189.

Nikiforova, J. V.; Bielecka, M.; Gakière, B.; Krueger, S.; Rinder, J.; Kempa, S.; Morcuende, R.; Scheible, W.-R.; Hesse, H.; Hoefgen, R., Effect of sulfur availability on the integrity of amino acid biosynthesis in plants. *Amino Acids* **2006,** *30*, 173-183.

Nunes, A. C. S.; Vianna, G. R.; Cuneo, F.; Amaya-Farfan, J.; de Capdeville, G.; Rech, E. L.; Aragao, F. J. L., RNAi-mediated silencing of the *myo*-inositol-1-phosphate synthase gene (*GmMIPS1*) in transgenic soybean inhibited seed development and reduced phytate content. *Planta* **2006,** *224*, 125-132.

O'Dell, B. L.; De Boland, A. R.; Koirtyohann, S. R., Distribution of phytate and nutritionally important elements among the morphological components of cereal grains. *J. Agric. Food Chem.* **1972,** *20*, 718-723.

Oates, K.; De Borba, B.; Rohrer, J., Determination of inositol phosphates in dried distillers grains with solubles. Thermo Fisher Scientific, Sunnyvale, CA, USA. **2014**, 1-10.

Odom, A. R.; Stahlberg, A.; Wente, S. R.; York, J. D., A role for nuclear inositol 1,4,5-trisphosphate kinase in transcriptional control. *Sci* **2000,** *287*, 2026-2029.

Oertel, A.; Matros, A.; Hartmann, A.; Arapitsas, P.; Dehmer, K. J.; Martens, S.; Mock, H. P., Metabolite profiling of red and blue potatoes revealed cultivar and tissue specific patterns for anthocyanins and other polyphenols. *Planta* **2017,** *246*, 281-297.

Oladosu, Y.; Rafii, M. Y.; Abdullah, N.; Hussin, G.; Ramli, A.; Rahim, H. A.; Miah, G.; Usman, M., Principle and application of plant mutagenesis in crop improvement: a review. *Biotechnol Biotec Eq* **2016,** *30*, 1-16.

Oliver, S. G.; Winson, M. K.; Kell, D. B.; Baganz, F., Systematic functional analysis of the yeast genome. *Trends Biotechnol.* **1998,** *16*, 373-378.

Panzeri, D.; Cassani, E.; Doria, E.; Tagliabue, G.; Forti, L.; Campion, B.; Bollini, R.; Brearley, C. A.; Pilu, R.; Nielsen, E.; Sparvoli, F., A defective ABC transporter of the *MRP* family, responsible for the bean *lpa1* mutation, affects the regulation of the phytic acid pathway, reduces seed *myo*-inositol and alters ABA sensitivity. *New Phytol.* **2011,** *191*, 70-83.

Perera, I.; Seneweera, S.; Hirotsu, N., Manipulating the phytic acid content of rice grain toward improving micronutrient bioavailability. *Rice* **2018,** *11*.

Petry, N.; Egli, I.; Campion, B.; Nielsen, E.; Hurrell, R., Genetic reduction of phytate in common bean (*Phaseolus vulgaris* L.) seeds increases iron absorption in young women. *J. Nutr.* **2013,** *143*, 1219-1224.

Pike, N., Using false discovery rates for multiple comparisons in ecology and evolution. *Methods Ecol. Evol.* **2011,** *2*, 278-282.

Pilu, R.; Landoni, M.; Cassani, E.; Doria, E.; Nielsen, E., The maize *lpa*241 mutation causes a remarkable variability of expression and some pleiotropic effects. *Crop Sci.* **2005,** *45*, 2096-2105.

Punjabi, M.; Bharadvaja, N.; Jolly, M.; Dahuja, A.; Sachdev, A., Development and evaluation of *low phytic acid* soybean by siRNA triggered seed specific silencing of inositol polyphosphate 6-/3-/5-kinase gene. *Frontiers in Plant Science* **2018,** *9*.

Raboy, V., *myo*-Inositol-1,2,3,4,5,6-hexa*kis*phosphate. *Phytochemistry* **2003,** *64*, 1033-1043.

Raboy, V., Seed phosphorus and the development of low-phytate crops. In *Inositol phosphates: Linking agriculture and the environment*, 1st ed.; Turner, B. L.; Richardson, A. E.; Mullaney, E. J., Eds. CAB international: Wallingford, UK, 2007; pp 111-132.

Raboy, V., Approaches and challenges to engineering seed phytate and total phosphorus. *Plant Sci.* **2009,** *177*, 281-296.

Raboy, V.; Gerbasi, P. F.; Young, K. A.; Stoneberg, S. D.; Pickett, S. G.; Bauman, A. T.; Murthy, P. P. N.; Sheridan, W. F.; Ertl, D. S., Origin and seed phenotype of maize *low phytic acid 1-1* and *low phytic acid 2-1*. *Plant Physiol.* **2000,** *124*, 355-368.

Raboy, V.; Peterson, K.; Jackson, C.; Marshall, J.; Hu, G.; Saneoka, H.; Bregitzer, P., A substantial fraction of barley (*Hordeum vulgare* L.) *low phytic acid* mutations have little or no effect on yield across diverse production environments. *Plants* **2015,** *4*, 225.

Ramautar, R.; Somsen, G. W.; de Jong, G. J., CE-MS in metabolomics. *Electrophoresis* **2009,** *30*, 276-291.

Ramautar, R.; Somsen, G. W.; de Jong, G. J., CE-MS for metabolomics: Developments and applications in the period 2014-2016. *Electrophoresis* **2017,** *38*, 190-202.

Ramautar, R.; Somsen, G. W.; Jong, G. J., CE-MS for metabolomics: Developments and applications in the period 2016–2018. *Electrophoresis* **2018**.

Rasmussen, S. K.; Hatzack, F., Identification of two low-phytate barley (*Hordeum vulgare* L.) grain mutants by TLC and genetic analysis. *Hereditas* **1998,** *129*, 107-112.

Reddy, N. R., Occurrence, distribution, content, and dietary intake of phytate. *Food phytates* **2002**, 25-51.

Roessner, U.; Wagner, C.; Kopka, J.; Trethewey, R. N.; Willmitzer, L., Simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry. *Plant J.* **2000,** *23*, 131-142.

Röhlig, R. M.; Engel, K. H., Influence of the input system (conventional versus organic farming) on metabolite profiles of maize (*Zea mays*) kernels. *J. Agric. Food Chem.* **2010,** *58*, 3022-3030.

Rolin, D.; Deborde, C.; Maucourt, M.; Cabasson, C.; Fauvelle, F.; Jacob, D.; Canlet, C.; Moing, A., Chapter One - High-Resolution 1H-NMR Spectroscopy and Beyond

to Explore Plant Metabolome. In *Adv. Bot. Res.*, Rolin, D., Ed. Academic Press: 2013; Vol. 67, pp 1-66.

Roullier-Gall, C.; Hemmler, D.; Gonsior, M.; Li, Y.; Nikolantonaki, M.; Aron, A.; Coelho, C.; Gougeon, R. D.; Schmitt-Kopplin, P., Sulfites and the wine metabolome. *Food Chem.* **2017,** *237*, 106-113.

Rutger, J.; Raboy, V.; Moldenhauer, K.; Bryant, R.; Lee, F.; Gibbons, J., Registration of KBNT *lpa*1-1 *low phytic acid* germplasm of rice. *Crop Sci.* **2004,** *44*, 363-364.

Sandberg, A. S.; Brune, M.; Carlsson, N. G.; Hallberg, L.; Skoglund, E.; Rossander-Hulthen, L., Inositol phosphates with different numbers of phosphate groups influence iron absorption in humans. *Am. J. Clin. Nutr.* **1999,** *70*, 240-246.

Sauer, K.; Cooke, M. P., Regulation of immune cell development through soluble inositol-1,3,4,5-tetra*kis*phosphate. *Nature Reviews Immunology* **2010,** *10*, 257-271.

Shi, J. R.; Wang, H. Y.; Schellin, K.; Li, B. L.; Faller, M.; Stoop, J. M.; Meeley, R. B.; Ertl, D. S.; Ranch, J. P.; Glassman, K., Embryo-specific silencing of a transporter reduces phytic acid content of maize and soybean seeds. *Nat. Biotechnol.* **2007,** *25*, 930-937.

Shi, J. R.; Wang, H. Y.; Wu, Y. S.; Hazebroek, J.; Meeley, R. B.; Ertl, D. S., The maize *low-phytic acid* mutant *lpa2* is caused by mutation in an inositol phosphate kinase gene. *Plant Physiol.* **2003,** *131*, 507-515.

Shulaev, V.; Cortes, D.; Miller, G.; Mittler, R., Metabolomics for plant stress response. *Physiol. Plant.* **2008,** *132*, 199-208.

Singh, O. V.; Ghai, S.; Paul, D.; Jain, R. K., Genetically modified crops: success, safety assessment, and public concern. *Appl. Microbiol. Biotechnol.* **2006,** *71*, 598-607.

Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G., METLIN - A metabolite mass spectral database. *Ther. Drug Monit.* **2005,** *27*, 747-751.

Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G., XCMS: Processing mass spectrometry data for metabolite profiling using Nonlinear peak alignment, matching, and identification. *AnaCh* **2006,** *78*, 779-787.

Sparvoli, F.; Cominelli, E., Phytate Transport by MRPs. *Signal Commun Plants* **2014**, 19-38.

Sparvoli, F.; Cominelli, E., Seed biofortification and phytic acid reduction: a conflict of interest for the plant? *Plants* **2015,** *4*, 728-755.

Spear, J. D.; Fehr, W. R., Genetic improvement of seedling emergence of soybean lines with low phytate. *Crop Sci.* **2007,** *47*, 1354-1360.

Spraul, M.; Braumann, U.; Godejohann, M.; Daolio, C.; Tseng, L.-H., Application of LC-NMR to the Study of Natural Products. In *Modern NMR Approaches to the Structure Elucidation of Natural Products*, 2015; pp 71-92.

Stadler, L. J., Mutations in barley induced by X-rays and radium. *Sci* **1928,** *68*, 186-187.

Stekhoven, D. J.; Buhlmann, P., MissForest-non-parametric missing value imputation for mixed-type data. *Bioinformatics* **2012,** *28*, 112-118.

Steuer, R.; Morgenthal, K.; Weckwerth, W.; Selbig, J., A Gentle Guide to the Analysis of Metabolomic Data. In *Metabolomics: Methods and Protocols*, Weckwerth, W., Ed. Humana Press: Totowa, NJ, 2007; pp 105-126.

Storey, J. D.; Tibshirani, R., Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U. S. A.* **2003,** *100*, 9440-9445.

Su, D.; Lei, B. T.; Li, Z. W.; Cao, Z. Z.; Huang, F. D.; Pan, G.; Ding, Y.; Cheng, F. M., Influence of high temperature during filling period on grain phytic acid and its relation to spikelet sterility and grain weight in non-lethal low phytic acid mutations in rice. *J Cereal Sci* **2014,** *60*, 331-338.

Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W. M.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reily, M. D.; Thaden, J. J.; Viant, M. R., Proposed minimum reporting standards for chemical analysis. *Metabolomics* **2007,** *3*, 211-221.

Sysi-Aho, M.; Katajamaa, M.; Yetukuri, L.; Oresic, M., Normalization method for metabolomics data using optimal selection of multiple internal standards. *BMC Bioinformatics* **2007,** *8*.

t'Kindt, R.; Morreel, K.; Deforce, D.; Boerjan, W.; Van Bocxlaer, J., Joint GC-MS and LC-MS platforms for comprehensive plant metabolomics: Repeatability and sample pre-treatment. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2009,** *877*, 3572-3580.

Tang, D. Q.; Zou, L.; Yin, X. X.; Ong, C. N., HILIC-MS for metabolomics: An attractive and complementary approach to RPLC-MS. *Mass Spectrom. Rev.* **2016,** *35*, 574-600.

Toffali, K.; Zamboni, A.; Anesi, A.; Stocchero, M.; Pezzotti, M.; Levi, M.; Guzzo, F., Novel aspects of grape berry ripening and post-harvest withering revealed by untargeted LC-ESI-MS metabolomics analysis. *Metabolomics* **2011,** *7*, 424-436.

Triba, M. N.; Le Moyec, L.; Amathieu, R.; Goossens, C.; Bouchemal, N.; Nahon, P.; Rutledge, D. N.; Savarin, P., PLS/OPLS models in metabolomics: the impact of permutation of dataset rows on the K-fold cross-validation quality parameters. *Mol. Biosyst.* **2015,** *11*, 13-19.

Troyanskaya, O.; Cantor, M.; Sherlock, G.; Brown, P.; Hastie, T.; Tibshirani, R.; Botstein, D.; Altman, R. B., Missing value estimation methods for DNA microarrays. *Bioinformatics* **2001,** *17*, 520-525.

Trygg, J.; Holmes, E.; Lundstedt, T., Chemometrics in metabonomics. *J. Proteome Res.* **2007,** *6*, 469-479.

Ulrich, E. L.; Akutsu, H.; Doreleijers, J. F.; Harano, Y.; Ioannidis, Y. E.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z.; Nakatani, E.; Schulte, C. F.; Tolmie, D. E.; Wenger, R. K.; Yao, H. Y.; Markley, J. L., BioMagResBank. *NAR* **2008,** *36*, 402-408.

Veum, T. L.; Ledoux, D. R.; Bollinger, D. W.; Raboy, V.; Cook, A., *Low phytic acid* barley improves calcium and phosphorus utilization and growth performance in growing pigs. *J. Anim. Sci.* **2002,** *80*, 2663-2670.

Veum, T. L.; Raboy, V., Hulled and hull-less barley grains with the genetic trait for *low phytic acid* increased the apparent total-tract digestibility of phosphorus and calcium in diets for young swine. *J. Anim. Sci.* **2016,** *94*, 1000-1011.

Vinaixa, M.; Samino, S.; Saez, I.; Duran, J.; Guinovart, J. J.; Yanes, O., A Guideline to Univariate Statistical Analysis for LC/MS-Based Untargeted Metabolomics-Derived Data. *Metabolites* **2012,** *2*, 775-795.

Vincent, J. A.; Stacey, M.; Stacey, G.; Bilyeu, K. D., Phytic acid and inorganic phosphate composition in soybean lines with independent *IPK1* mutations. *Plant Genome-Us* **2015,** *8*.

Ward, J. L.; Baker, J. M.; Miller, S. J.; Deborde, C.; Maucourt, M.; Biais, B.; Rolin, D.; Moing, A.; Moco, S.; Vervoort, J.; Lommen, A.; Schafer, H.; Humpfer, E.; Beale, M. H., An inter-laboratory comparison demonstrates that [H-1]-NMR metabolite fingerprinting is a robust technique for collaborative plant metabolomic data collection. *Metabolomics* **2010,** *6*, 263-273.

Wei, R. M.; Wang, J. Y.; Su, M. M.; Jia, E.; Chen, S. Q.; Chen, T. L.; Ni, Y., Missing Value Imputation Approach for Mass Spectrometry-based Metabolomics Data. *Sci. Rep.* **2018,** *(8): 663*.

Wenzel, A.; Frank, T.; Reichenberger, G.; Herz, M.; Engel, K. H., Impact of induced drought stress on the metabolite profiles of barley grain. *Metabolomics* **2015,** *11*, 454-467.

Westerhuis, J. A.; Hoefsloot, H. C. J.; Smit, S.; Vis, D. J.; Smilde, A. K.; van Velzen, E. J. J.; van Duijnhoven, J. P. M.; van Dorsten, F. A., Assessment of PLSDA cross validation. *Metabolomics* **2008,** *4*, 81-89.

Wiggins, S. J.; Smallwood, C. J.; West, D. R.; Kopsell, D. A.; Sams, C. E.; Pantalone, V. R., Agronomic performance and seed inorganic phosphorus stability of low-phytate soybean line TN09-239. *J. Am. Oil Chem. Soc.* **2018,** *95*, 787-796.

Wiklund, S.; Johansson, E.; Sjostrom, L.; Mellerowicz, E. J.; Edlund, U.; Shockcor, J. P.; Gottfries, J.; Moritz, T.; Trygg, J., Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. *AnaCh* **2008,** *80*, 115-122.

Wilcox, J. R.; Premachandra, G. S.; Young, K. A.; Raboy, V., Isolation of high seed inorganic P, low-phytate soybean mutants. *Crop Sci.* **2000,** *40*, 1601-1605.

Wishart, D. S., Advances in metabolite identification. *Bioanalysis* **2011,** *3*, 1769-1782.

Worley, B.; Powers, R., Multivariate Analysis in Metabolomics. *Curr Metabol* **2013,** *1*, 92-107.

Worley, B.; Powers, R., PCA as a Practical Indicator of OPLS-DA Model Reliability. *Curr Metabol* **2016,** *4*, 97-103.

Xia, J.; Wishart, D. S., Using MetaboAnalyst 3.0 for Comprehensive Metabolomics Data Analysis. *Current Protocols in Bioinformatics* **2016,** *55*, 14.10.1-14.10.91.

Xu, X. H.; Zhao, H. J.; Liu, Q. L.; Frank, T.; Engel, K. H.; An, G. H.; Shu, Q. Y., Mutations of the multi-drug resistance-associated protein ABC transporter gene 5 result in reduction of phytic acid in rice seeds. *Theor. Appl. Genet.* **2009,** *119*, 75-83.

Ye, H.; Li, C.; Bellgard, M.; Lance, R.; Wu, D., Genes controlling *low phytic acid* in plants: Identifying targets for barley breeding. In *Advance in Barley Sciences*, Springer: 2013; pp 185-198.

Yu, T. W.; Jones, D. P., Improving peak detection in high-resolution LC/MS metabolomics data using preexisting knowledge and machine learning approach. *Bioinformatics* **2014,** *30*, 2941-2948.

Yuan, F. J.; Zhao, H. J.; Ren, X. L.; Zhu, S. L.; Fu, X. J.; Shu, Q. Y., Generation and characterization of two novel low phytate mutations in soybean (*Glycine max* L. Merr.). *Theor. Appl. Genet.* **2007,** *115*, 945-957.

Yuan, Y. F.; Zhao, Y. P.; Yang, J. L.; Jiang, Y. M.; Lu, F.; Jia, Y. X.; Yang, B., Metabolomic analyses of banana during postharvest senescence by H-1-high resolution-NMR. *Food Chem.* **2017,** *218*, 406-412.

Zhang, A. H.; Sun, H.; Wang, P.; Han, Y.; Wang, X. J., Modern analytical techniques in metabolomics analysis. *Analyst* **2012,** *137*, 293-300.

Zhang, J. J.; Zhao, C. X.; Zeng, Z. D.; Luo, P.; Zhao, Y. N.; Zhao, J. Y.; Li, L. L.; Lu, X.; Xu, G. W., Sample-directed pseudotargeted method for the metabolic profiling analysis of rice seeds based on liquid chromatography with mass spectrometry. *J. Sep. Sci.* **2016,** *39*, 247-255.

Zhang, P.; Li, J. Q.; Li, X. L.; Liu, X. D.; Zhao, X. J.; Lu, Y. G., Population structure and genetic diversity in a rice core collection (*Oryza sativa* L.) investigated with SSR Markers. *PLoS One* **2011,** *6*.

Zhao, H. J.; Cui, H. R.; Xu, X. H.; Tan, Y. Y.; Fu, J. J.; Liu, G. Z.; Poirier, Y.; Shu, Q. Y., Characterization of *OsMIK* in a rice mutant with reduced phytate content reveals an insertion of a rearranged retrotransposon. *Theor. Appl. Genet.* **2013,** *126*, 3009-3020.

Zhao, H. J.; Frank, T.; Tan, Y. Y.; Zhou, C. G.; Jabnoune, M.; Arpat, A. B.; Cui, H. R.; Huang, J. Z.; He, Z. H.; Poirier, Y.; Engel, K. H.; Shu, Q. Y., Disruption of *OsSULTR3;3* reduces phytate and phosphorus concentrations and alters the metabolite profile in rice grains. *New Phytol.* **2016,** *211*, 926-939.

Zhao, H. J.; Liu, Q. L.; Fu, H. W.; Xu, X. H.; Wu, D. X.; Shu, Q. Y., Effect of non-lethal *low phytic acid* mutations on grain yield and seed viability in rice. *Field Crops Res.* **2008a,** *108*, 206-211.

Zhao, H. J.; Liu, Q. L.; Ren, X. L.; Wu, D. X.; Shu, Q. Y., Gene identification and allele-specific marker development for two allelic *low phytic acid* mutations in rice (*Oryza sativa* L.). *Mol. Breed.* **2008b,** *22*, 603-612.

PUBLICATIONS AND PRESENTATIONS

PUBLICATIONS (peer-reviewed)

Zhou, C. G.; Tan, Y. Y.; Goßner, S.; Li, Y. F.; Shu, Q. Y.; Engel, K. H., Impact of crossing parent and environment on the metabolite profiles of progenies generated from a *low phytic acid* rice (*Oryza sativa* L.) Mutant. *J. Agric. Food Chem.* 2019, 67, 2396-2407.

Zhou, C. G.; Tan, Y. Y.; Goßner, S.; Li, Y. F.; Shu, Q. Y.; Engel, K. H., Impact of Cross-Breeding of *Low Phytic Acid* Rice (*Oryza sativa* L.) Mutants with Commercial Cultivars on the Phytic Acid Contents. *Eur. Food Res. Technol.* 2019, 245, 707-716.

Zhou, C. G.; Tan, Y. Y.; Goßner, S.; Li, Y. F.; Shu, Q. Y.; Engel, K. H., Stability of the Metabolite signature resulting from the *OsSULTR3;3* mutation in low phytic acid rice (*Oryza sativa* L.) seeds upon cross-breeding. *J. Agric. Food Chem.* 2018, 66, 9366-9376.

Goßner, S.; Yuan, F.J; **Zhou, C. G.**; Tan, Y. Y.; Shu, Q. Y.; Engel, K. H., Impact of cross-breeding of *low phytic acid MIPS1* and *IPK1* soybean (*Glycine max* L. Merr.) mutants on their contents of inositol phosphate isomers. *J. Agric. Food Chem.* 2019, 67, 247-257.

Zhao, H. J.; Frank, T.; Tan, Y. Y.; **Zhou, C. G.**; Jabnoune, M.; Arpat, A. B.; Cui, H. R.; Huang, J. Z.; He, Z. H.; Poirier, Y.; Engel, K. H.; Shu, Q. Y., Disruption of *OsSULTR3;3* reduces phytate and phosphorus concentrations and alters the metabolite profile in rice grains. *New Phytol.* 2016, 211, 926-939.

PRESENTATIONS

Zhou, C. G.; Tan, Y. Y.; Goßner, S.; Shu, Q. Y.; Engel, K. H., Impact of cross-breeding on the metabolite signature resulting from the *OsSULTR3;3* mutation in *low phytic acid* rice seeds (Oral). FAO/IAEA International Symposium on Plant Mutation Breeding and Biotechnology, August 27 – 31, 2018, Vienna, Austria.

Zhou, C. G.; Frank, T.; Tan, Y. Y.; Zhao, H. J.; Shu, Q. Y.; Engel, K. H., Changes in the metabolite profile of rice upon disruption of the sulfate transporter gene *OsSULTR3;3* (Poster). Max Rubner Conference 2016 – Food Metabolomics, October 10 – 12, 2016, Karlsruhe, Germany.

CURRICULUM VITAE

PERSONAL INFORMATION

Name:	Chenguang Zhou
Date of birth:	20.08.1990
Nationality:	Chinese

SCIENTIFIC ACTIVITY

10/2014 - present:	PhD candidate/Research assistant Chair of General Food Technology (Univ.-Prof. Dr. Karl-Heinz Engel) Technical University of Munich Freising, Germany

ACADEMIC STUDIES AND EDUCATION

09/2012 - 07/2014:	Master in Food Biotechnology Department of Food Science and Nutritional Engineering China Agricultural University
09/2008 - 07/2012:	Bachelor in Food Quality and Safety Department of Wine and Food Engineering Guizhou University