UNIVERSITÀ DEGLI STUDI DI MILANO

---

PhD School in: Agriculture, Environment and Bioenergy-XXX CYCLE

# GENOME-WIDE ANALYSIS OF JAPONICA RICE PERFORMANCES UNDER ALTERNATE WETTING AND DRYING AND PERMANENT FLOODING CONDITIONS.

PhD program coordinator: Prof. Daniele Bassi

Supervisor: Prof. Gian Attilio Sacchi

Co-supervisor: Dr. Giampiero Valè

PhD candidate: Gabriele Orasen

Academic year 2016-2017

*Giampiero and Prof. Sacchi*

*Thank you very much for your supervision and support, for my professional training, for your guidance…..*

*…….and for believing in me.*

# Table of contents

## Abbreviation summary:

- As: Arsenic
- ABA: Abscissic Acid
- AWD: Alternate Wetting and Dryng
- Ca: Calcium
- Cd: Cadmium
- CHL: Chlorophyll content
- Cu: Copper
- DF: Days to Flowering
- DFM: Filling Time
- DM: Days to Maturing
- Fe: Iron
- FLA: Flavonoids Content
- FLL: Flag Leaf Length
- FLW: Flag Leaf Width
- GA: Gibberellin
- HGW: Grain Weight (one undred seeds)
- K: Potassium
- LA: Leaf Area
- Mg: Magnesium
- Mn: Manganese
- Na: Sodium
- NBI: Nitrogen Balance Index
- NHGW: Dehulled Grain Weight (one undred seeds)
- Ni: Nichel
- NSL: Dehulled Seed Length
- NSW: Dehulled Seed Width
- NSWLR: Dehulled Seed Width/Length Ratio
- P: Phosphorus
- PF: Permanent Flooding
- PH: Plant Total Height
- PL: Panicle Length
- PNH: Panicular Node Height
- PW: Yeld Index
- SL: Seed Length
- SW: Seed Width
- SWLR: Seed Width/Length Ratio
- TPM: Tillering
- Zn: Zinc

# ABSTRACT

A rice GWAS panel of 281 accessions of japonica rice was phenotypically characterized for traits related to phenology, plant and seed morphology, physiology, yield and grain ionome for two years in field conditions under permanent flooding (PF) or alternate wetting and drying(AWD). A genome-wide analysis approach uncovered a total of 360 significant marker-trait associations (MTAs), of which 105 were AWD-specific, 178 were PF-specific and 77 were in common between the two water management systems. AWD-specific associations were identified for several agronomic traits including days to maturation, days from flowering to maturation, leaf traits, plant height, panicle and seed traits, one hundred grain weight, yield, tillering, mineral nutrient and toxic trace elements level in grains. Significant MTAs were detected across all the 12 rice chromosomes. The analysis of genes annotated in the Nipponbare reference sequence and included in the regions associated to the analyzed traits allowed the identification of several loci known to affect the respective traits. The high number of MTAs identified open new perspectives for the development of functional genomic and breeding strategies.

# 1. INTRODUCTION

## 1.1.Rice, a staple crop

Rice is one of the most important staple crops for the world population, in fact it provides more than 40 % of the daily calories (Parengam et al., 2010; Zhang et al., 2014). Although rice is not rich of any particular mineral or vitamin, for those whose diet depends on it, rice can be an important source of these elements, as well as caloric energy (Zhang et al., 2014).

More than 90% of the world's rice is grown and consumed in Asia where 60% of the earth's people live. Rice accounts for 35 to 60% of the calories consumed by 3 billion Asians; it is sown on about 148 million hectares annually, i.e. on about 11% of the world's cultivated land. Wheat covers a slightly larger land area, but a considerable amount of this crop is used as animal feed. Rice is the only major cereal crop that is consumed almost exclusively by humans (Khush, 1997).

According to FAO forecast for the 2016/2017 period, rice will remain the third cereal in terms of production below wheat and coarse grains (corn, barley, sorghum, oats and rye) (FAO Trade and Market Division 2016). Nevertheless, rice stands out since for the 2016/2017 forecast period it is estimated that more than 80% of its production will be destined for direct human consumption, principally due to its high calorie content (FAO Rice Market Monitor 2016). In comparison, only 67 and 15% of the wheat and coarse grains production, respectively, will be destined for human consumption (FAO Trade and Market Division 2016). In addition, it has been proposed that rice will be one of the main calorie supplies in the forthcoming years (FAO Rice Market Monitor 2016).

Italy produces about 56% of Europe's rice(Conaf, 2015). In 2012-2014 years, world rice production was 4.1 million tons, 5.5 % of the world production (FAO Rice Market Monitor 2016).

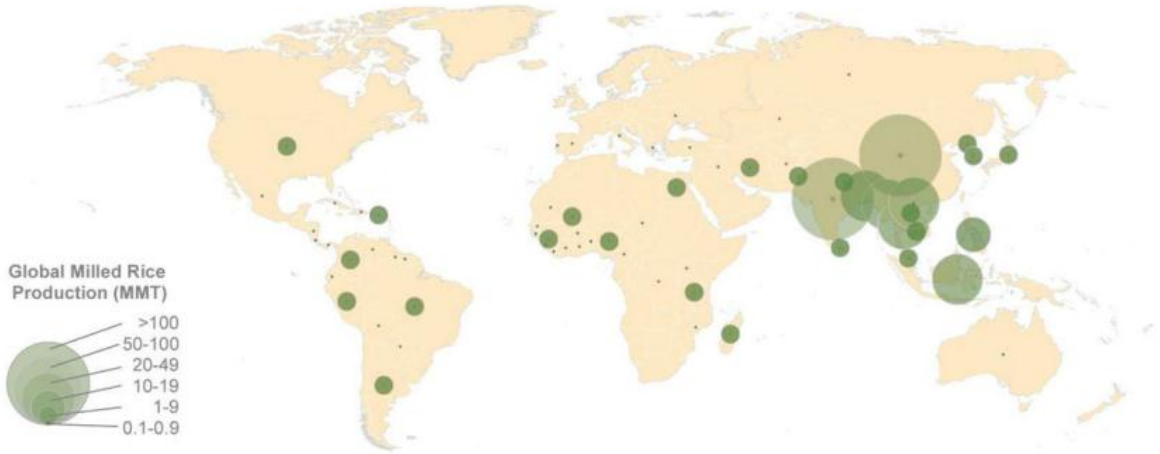



***Figure 1.1.*** *Rice milling 2011 (Muthayya, Sugimoto, Montgomery, & GF, 2014)*

## 1.2.Rice production ecosystems

Rice grow under different word's regions including temperate, sub-tropical and tropical climatic areas.

On the basis of soil-water condition different rice production ecosystems can be distinguished.

*Lowland rice*

In this ecosystem rice is grown in bunded fields (paddy fields) with assured water supply. Soil is submerged for part or all the growing season, with alternate wetting and drying (AWD), or permanent flooding (PF) irrigation system. Flooded conditions require large quantities of water, which are used not only by plant for growing, but also as management tool during cultivation (Brown et al., 1978; McCauley, 1990; Matsushima, 1976; Angelini et al., 2008). Although it is not a compulsory aquatic plant, rice efficiently takes root waterlogging by means of specific morphological and physiological adaptive processes (Colmer, 2003; Colmer et al. 2006; Kulichikhin et al, 2014; Shiono et al., 2014) making it a crop specie suitable for lowland ecosystem.

In the case of irrigated lowland rice, a complete water control is maintained along the whole crop growing season managing high volumes of water thus the soil is always under anaerobic condition, excluding the last 2-3 weeks when it is drained for harvesting. In the case of *rain fed lowland rice* the water is supplied only by frequent rainfalls, thus depending on weather soil could shift for periods from anaerobic to aerobic conditions.

Finally, in the case of *deepwater rice* plant grownin fields flooded to a sustained depth of at least 50 cm and for long period also large part of the shoots could be submerged.

About 90% of the world's annual rice production comes from lowland rice and in particular 75% by the irrigated ones (Tuong & Bouman, 2003; GRiSP, 2013).

*Upland rice*

In this ecosystem rice is grown in rainfed, naturally well drained soils without surface water accumulation, without phreatic water supply, and normally not bunded (IRRI, 1984). Upland rice is drought-prone, usually cultivated on sloping land with erosion problems, and the soils have both poor physical and chemical properties. Farmers in these environments are among the poorest and usually cannot afford to apply external inputs such as water and fertilizers. Upland rice varieties are drought tolerant, but have a low yield potential and tend to lodge under high levels of external inputs such as fertilizer and supplemental irrigation; they are considered as a low-yielding subsistence crop

Where farmers can afford to buy external inputs, have access to supplementary irrigation if rainfall was not sufficient, rice varieties combining drought-tolerant characteristics of upland varieties with the high-yielding characteristics of lowland varieties the so called *aerobic rice system* can be adopted. In essence, *aerobic rice* can be seen as "favorable" or "high yielding" upland rice. Achieving high yields under relatively favorable aerobic soil conditions requires new varieties. About 50% of

the world's rice land is irrigated (Fig. 1.2). Of the half that is not, rainfed lowland rice occupies more than all other cultural types combined.

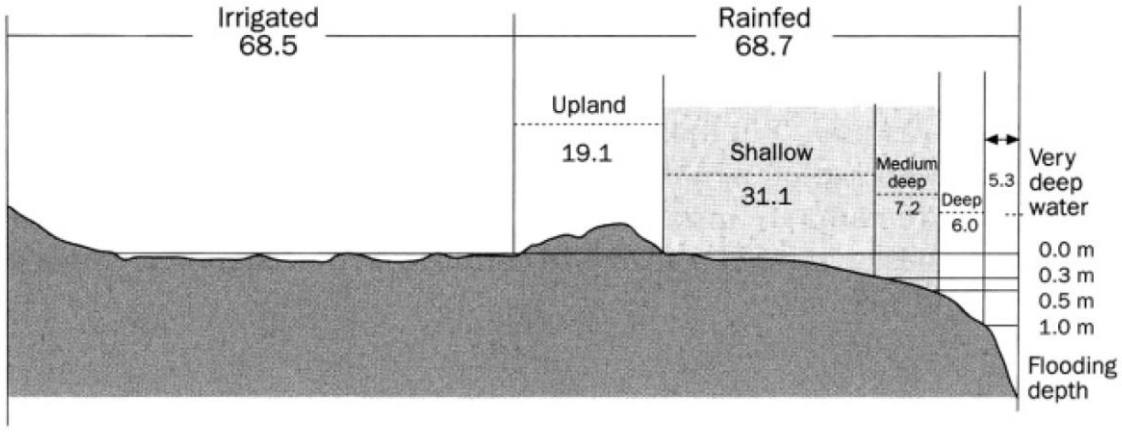



***Figure 1.2.*** *Different rice water management in the world. Width is proportional to the extent (million ha) of each area (Mackill et al., 1996).*

Irrigated lowland rice require more than 2,500 L for producing 1 Kg of rice grain and as a consequence about 40% of the global irrigation freshwater is employed in rice paddy fields (Maclean et al., 2002; Bouman et al., 2007; Bouman, 2009; Ringler and Zhu, 2015). In irrigated lowland rice the amount of water consumed through transpiration, evaporation and percolation phenomena is very high, and considering the amount of water required to maintain soil flooding under a regular hydric downflow, seasonal water volumes consumed for the cultivation of rice can vary between 17,000 $m^3$ $ha^{-1}$ in heavy soils to 42,000 $m^3$ $ha^{-1}$ in light soils (Angelini et al., 2008). These volumes of water are about 2-3 times higher than those required by other cereals (Zhi-Kang and JL, 2007), altough the average value of physiological water productivity (i. e. the ratio between yield and volume of water transpired by plants) of rice is comparable (about 1.1 g grain $kg^{-1}$ water) with the other major $C_3$ cereal crops (Tuong et *al*., 2005; Ringler and Zhu, 2015). Global warming challenge and, somewhere, competition between paddy rice and other seasonal crop, as well as industrial and/or civic requirements for water, might cause physic and/or economic water shortages for irrigated lowland rice systems (Tuong and Bouman, 2003; Wassmann et al., 2009). Thus it is expected that about 15-20 million hectares of irrigated rice will experience water scarcity within the 2050 (Bouman et *al*., 2007). This challenge concerns also Europe rice areas where, as in most of the temperate areas, the crop is grown almost exclusively under PF using irrigation waters from major rivers.

Lowland rice systems has a significant greenhouse footprint. Indeed, due to anaerobic fermentation of soil organic matter occurring during flooding periods paddies are important sources of atmospheric methane ($CH_4$), contributing approximately 15–20% of the global total anthropogenic emission of this gas (Sass and Fisher, 1997; Aulakh et al., 2001). Moreover, due to the combined effect of N fertilization and water paddies emit into atmosphere also substantial amounts of $N_2O$ (Nishimura et al., 2004; Yu et al., 2004), another greenhouse active gas (GHG). Overall flooded rice fields are a significant component of the total agricultural GHG emissions that contribute approximately 12% to total global anthropogenic GHG emissions.

It appears evident that for improving the sustainability of rice irrigated systems both the adoption of water-saving practices (Bouman et *al*., 2007; Farooq et *al*., 2009) and the constitution of new rice varieties more efficient in the use of water (Serraj et *al*., 2009) should be considered key strategies.

Among water-saving techniques, AWD is very valuable in reducing water input (about -35%) in comparison to PF (Price et al., 2013; Lampayan et al., 2015). In AWD water is applied up to flood the soil some days after ponded water is disappeared on soil surface (Siopongco et al., 2013). The water content of upper soil layers alternates thus from saturation to non-saturation, thus over and over shifting from anaerobic to aerobic conditions (Bouman and Tuong, 2001; Farooq et al., 2009).

As suggested by IRRI, the AWD practice commences at 1 to 2 week after transplanting, requires draining the field until the water level reaches about 15 cm below soil surface, after that the field is re-flooded to a ponded depth of around 10 cm before re-draining. In order to prevent yield penalty using lowland rice germplasm the particular irrigation scheme described, known as ***safe AWD*** (Lampayan et al 2009) should continue throughout the cropping season except from 1 week before and 1 week after flowering. Sometimes, after flowering the water management is return to PF. Since water could be not always available, as a consequence of environmental reasons or specific regional rules, the 15-cm subsurface water level threshold for re-flooding could be not obeyed and irrigation take place only when soil water potential decrease during the dry period below to very low values. In this case the term ***forced AWD***, or intermittent irrigation (IIR), is preferred.

Where AWD is practised as an alternative to PF the potential water saving has been evaluated to be in the range 25-50% (Lampayan et al., 2015). Nevertheless, not always this reduction in using water corresponds to an enhanced crop water use efficiency (the ratio between the yield and the total water needed including, other than seasonal transpiration, seasonal evaporation, percolation and runoff) since AWD may result in yield loss. Indeed, depending on the genotypes utilized and the number of the dry-wetting cycle along the season, the weather and the rice cultivar, AWD does not affect (Yao et al., 2012; Linquist et al., 2015), slightly lowers (Sudhir-Yadav et al., 2012; Shaibu et al., 2015) or even increases (Zhang et al. 2009) yield compared with PF. Thus, it is evident that the natural variation in adaptability to AWD existing in the rice germplasm, as well as its interaction with environmental variables, needs to be better explored and characterized.

During the aerobic periods of AWD plants risks to be exposed to mild drought stress (Bouman and Tuong, 2001) and this is much more likely for *forced* than *safe* AWD.

If rice plants experience water deficiency during their vegetative growth phase, lower sizes and reduction in the number of tillers as well as in the leaf surface occur, whereas if the stress take place during the reproductive phase (between panicle differentiation and flowering) both the number of spikelets as well as the fertility of flowers result negatively affected (Atlin et al., 2006; Cavigiolo et al., 2007; Venuprasad et al., 2007).

Since rather sensitive to drought high-yielding rice varieties bred for lowland cultivation systems are generally subjected to yield penalties in aerobic soils (Bouman and Tuong, 2001; Lafitte et al., 2002; Atlin et al., 2006; Matsuo et al., 2010). The identification among them of genotypes tolerant

recurring aerobic conditions of the soil could represent a good starting point for developing new rice genotypes to be cultivated under water saving condition.

Aside from water saving, another potential advantage provided by AWD is its contribution, or lack of, to greenhouse gases emission, especially methane (Wassmann et al., 2010; Li et al.,2006). In this regards it has been reported that multiple field aeration by AWD can potentially reduce $CH_4$ emissions by more than 50% and, more in general, the global warming potential (GWP of $CH_4$ and $N_2O$ emissions) by 45-90% compared to PF (Linquist et al., 2014).

Under PF the decomposition of soil organic matter (SOM) to $CH_4$ occurs in two phases: a) for the first week after flooding, as the soil redox potential drops and alternate electron acceptors [N-$NO_2$, $N_2O$, Mn (IV), $SO_4^{2-}$ and $Fe^{3+}$] are exhausted); b) during grain-filling (as fine roots and root exudates decompose). The drill-sowing, moving rice water management towards a more aerobic regime increases the soil redox potential (from about -300 mV to about +200 mV) ensuring that SOM will decompose to $CO_2$ rather than $CH_4$ and larger level in the soil of alternate electron acceptor make mitigation easier.

The GWP of nitrous oxides ($NO_x$) is more than 300 times that of $CO_2$. Consequently, small losses of nitrogen in the form of $NO_x$ have a significant greenhouse impact. $NO_x$ is commonly produced when soil transitions from aerobic to anaerobic states occur, as occurs in AWD. Under anaerobic condition nitrogen is mainly present in the reduced state $NH_4^+$. Progressive oxygenation of soil promotes the activity of nitrifier bacteria thus $NH_4^+$ is oxidized to $NO_3^-$. Subsequent soil re-flooding in turn triggers microbial denitrification reducing $NO_3^-$ in the volatile forms $N_2O$ and $NO_2$. Due to these transformations AWD systems contribute by $N_2O$ emission to GHG effect.

All the above considerations suggest that efficiently combining high-yield lowland genotypes with AWD could open perspective in improving the sustainability of rice cultivation. With this aim could be useful to identify within high-yield lowland germplasm collection genotypes well suited for AWD, as well as to identify the genetic traits sustaining their adaptability.

### 1.3.Water management and plant-soil relationships

Omission of oxygen from soil profiles causes physical, chemical and biological changes affecting availability of essential nutrients and consequently plant growth and yield. It is important to consider that flooding decreases water percolation rate. This take place because flooding promotes welling, disintegration and dispersion of soil aggregates, the reduction in size of soil pores where microbial activity and organic matter decomposition occur. All together these events limit the binding effect of aggregate and causes the tendency of the soil to seal off, a process known as pudding (Wickham and Singh, 1978; Soil Science Society of America, 2008). However, a positive consequence of decreased water percolation due to soil flooding is also a reduction of nutrient leaching toward the deep soil and groundwater.

As soon as soils are flooded the oxygen level begins to decline and thus aerobic microorganism become quiescent or die, while facultative and obligate anaerobic bacteria proliferate using gradually as electron acceptor for their respiration $NO_3^-$ ($E_h$ = +740 mV), $Mn^{4+}$ ($E_h$ =+400 mV), $Fe^{3+}$ ($E_h$ =- 190 mV), $SO_4^{2-}$ ($E_h$ =-210 mV), $CO_2$ ($E_h$ =-240 mV), $N_2$ ($E_h$ =-280mV) and $H^+$ ($E_h$ =+410 mVmV) producing $N_2$, $Mn^{2+}$, $Fe^{2+}$, $S^{2-}$, $CH_4$, $NH_3$ and $H_2$, respectively. Typically moving from upland to lowland condition the value of soil redox potential decrease from about +200 mV to lesser than -300 mV.

Other than causing changes in the availability of various nutrients, the processes of soil reduction generate a host of compounds known to be phytotoxic: the reduced forms of Fe and Mn, sulfide, ethanol, lactic acid, acetaldehyde and aliphatic acids such as formic, acetic, butyric acids, and cyanogenic compounds (Ponnamperuma, 1972).

Submergence induces alkalinisation or acidification of acidic or alkaline soils, respectively (Ponnamperuma, 1972). In other worlds flooding determines a convergence of soil pH to neutrality and thus benefits for rice plants through better availability of nutrient ($NH_4^+$, P, K) and exchangeable cations, which are mobilized in soil solution (Wade et al., 1998; Sahrawat, 2007). The presence of free water in paddy improves the availability and accessibility for the roots from the plant nutrients through mass flow and diffusion.

Under anaerobic condition nitrogen is mainly present in its reduced form $NH_4^+$ limiting the N losses by percolation and thus resulting in a higher (up to 80%) nitrogen use efficiency of paddy fields. Progressive oxygenation of soil promotes the activity of nitrifier bacteria and $NH_4^+$ is oxidized to $NO_3^-$; this accompanied by the physical changes induced on soil structure described above determines leaching of the nutrient. The subsequent soil re-flooding will provoke further N-losses since $NO_3^-$ is reduced to its volatile forms $NH_3$, $N_2O$ and $NO_2$. Due to its characteristic dry and wet in the AWD systems the nitrogen use efficiency fall also up to values lower than 50% (Tan et al., 2013). In order to overcome this negative effect a mixture of nitrate and ammonium nitrogen fertilization, as well as the adoption of slow release formulations of nitrogen are suggested for AWD.

Phosphorus under the aerobic phases of AWD could be not readily available since could be presents in the $Fe^{3+}$ less soluble forms (i.e. $FePO_4$). On the contrary Sulphur is more available in aerobic soil conditions since the root absorbs by root the oxidized $SO_4^{2-}$ form. Likewise, there should be no additional issues with $K^+$.

Soil submergence favors release of $Ca^{2+}$, $Mg^{2+}$ and $Na^+$ in soil solution, but since the amounts of these elements required by plants are relatively small, plants do not experience deficiencies also when their concentrations in soil solution decrease by soil drying.

Among micronutrients only zinc and iron availability is directly influenced by the redox condition of the soil. The shift from anaerobic to aerobic conditions of the soil can alter several soil chemical and physical properties that determine Zn availability, including pH, redox potential and organic matter content. In particular: a) $Zn^{2+}$ concentration in the soil solution is higher in aerobic conditions than in anaerobic ones since the cation is mainly bound with $SO_4^{2-}$ as soluble $ZnSO_4$ salt in the former,

whereas it is bond with $S^{2-}$ forming the insoluble ZnS salt in the latter; b) in aerobic soil the present $Fe^{3+}$ precipitates as $Fe(OH)_3$ onto which $Zn^{-}$ may be adsorbed ; c) aerobic condition often accelerates organic matter oxidation restricting $Zn^{2+}$ availability in soil solution; d) the reduction in soil water content as a result of the change in water management is likely to restrict transpiration and diffusion; as a consequence Zn transport towards roots and its movement within the plant can result restricted. In conclusion, the effect of water management on Zn availability can be quite complex.

Iron is the fourth most abundant element in the earth's crust, where ferric iron ($Fe^{3+}$) and ferrous iron ($Fe^{2+}$) are the most common forms. While $Fe^{3+}$ is insoluble and its uptake is difficult, $Fe^{2+}$ is soluble and readily available to plants (Guerinot and Yi, 1994). When the soil is aerated and its pH is in the alkaline range, Fe is oxidized as insoluble oxides, but in flooded soils the reduction of $Fe^{3+}$ to $Fe^{2+}$ occurs. This change is responsible for the lower availability of Fe in upland soils and for its high availability in flooded soils. Since $Fe^{2+}$ is highly reactive, the presence of high concentration of this ion (i.e. under anaerobic soil condition) could lead to a toxic accumulation of the metal in plant tissues (Winterbour, 1995). Formation of iron plaque (precipitates of iron oxides or hydroxides on the root surface) is generally considered as a rice adaptation to acclimate to anaerobic conditions, particularly when very high concentrations of $Fe^{2+}$, $Mn^{2+}$ and $S^{2-}$are present in the soil (Müller et al., 2017 and references therein). Iron plaque formation in rice root is related to Fe-excess and varies among cultivars (Pereira et al., 2014).

Plants have adopted two evolutionary strategies to acquire Fe from soil. Non-grass species activate a reduction-based Strategy I when starved for Fe, whereas the grass species activate a chelation-based strategy known as Strategy II (Marschner and Römheld, 1994). Strategy I plants extrude protons into the rhizosphere by a plasmalemma $H^{+}$-ATPase; this activity lowers the pH of rhizosphere and soil solution increasing the solubility of $Fe^{3+}$ (every one unit drop in pH $Fe^{3+}$ becomes a 1000-fold more soluble). $Fe^{3+}$ is then reduced to the more soluble and root absorbable $Fe^{2+}$ by a Ferric Reductase Oxidase (FRO) locate on root cell plasma-lemma (Connolly and Guerinot, 2002; Kobayashi and Nishizawa, 2012). Strategy II plants release a group of small molecular weight compounds known as phytosiderophores (PS) having high affinity for $Fe^{3+}$ and thus efficiently bind this ion in the rhizosphere. $Fe^{3+}$–PS complexes are then transported into the plant roots via a specific transport system (Kobayashi and Nishizawa, 2012). Strategy II plants can also take up $Fe^{2+}$ like Strategy I plants. Rice, for example, in addition to having the ability to transport Fe–PS complexes, is able to absorb $Fe^{2+}$ via a transporter known as *OsIRTs* (Ishimaru et al., 2006). Alternate shifts between anaerobic/aerobic soil condition induce changes in the $Fe^{3+}/Fe^{2+}$ratio availability in the rhizosphere thus involving differently the two Fe uptake strategies.

Mn deficiency is relatively rare especially in irrigated rice systems. It occurs frequently in upland rice, but is not common in rainfed or lowland systems rice because the solubility of Mn increases under submerged conditions when the element is mainly present as $Mn^{2+}$ that is taken up into the roots by a transporter (*Nramp5*) belonging to the *Nramp* family (Sasaki et al., 2012).

Contamination of agricultural soil with toxic metals and/or metalloids (As, Cd, Cr, Cu, Hg, Pb Sb, Co, and when present in excess, Zn and Ni) is one of the most pressing concerns in the debate

about food security and food safety in Europe (Tòt et al., 2016) and globally (Kong, 2014). It has recently emerged as a major health related issue almost across the globe. Concerning rice-growing areas Cd and As are the two-toxic element most worrying.

Cadmium is a toxic heavy metal naturally present in soils or anthropogenically released in natural and agricultural environments. Important sources of Cd soil contamination, other than short- or long-range atmospheric depositions, are both the use of phosphate fertilizers naturally containing traces of the metal and the agricultural use of manure, urban composts and industrial sludges (Alloway and Steinnes 1999; McLaughlin et al., 1999). Under anaerobic condition, the bioavailability of $Cd^{2+}$ gets reduced due to its transformation to insoluble CdS (Hu et al., 2013; Chaney, 2015). When soil become aerobic the oxidation of $S^{2-}$ to $SO_4^{2-}$ makes Cd, that in this condition is present as soluble $CdSO_4$, more avilable for root uptake.

The use of As contaminated groundwater for irrigation of crops has resulted in elevated concetrations of the element in agrigultural soils in Bangadlesh and West Bengala (India). Moreover,

The use of As contaminated groundwater for irrigation as in soil results from human activities including pesticide use. Arsenic is redox-sensitive (Roberts et al., 2010; Yamaguchi et al., 2011); in aerobic soil, As(V) is the dominant form, being strongly absorbed on soil (hydr)oxides and thus it is not much mobile (Takahashi et al., 2004). On the contrary in anaerobic soil As(III) is the dominant form and its sorption on soil minerals is very low in comparison to As(V). Moreover, considering that As(III) is more efficiently taken up by the rice roots (Zhao et al., 2010) it is not surprising that PF can promote accumulation of As in rice shoots and grain (Chen et al. 2005) in comparison with aerobic soil conditions. The effect of AWD on As accumulation in the shoot and then in the grain, might depend on the phenological stages during which the dried periods occur (Das et al., 2016).

Nickel (Ni) availability in soil varies as a function of pH. Plants require Ni in small quantities for normal development Knowledge on the redox geochemistry of Ni is behind in comparison to other metals. Usually mobilization of Ni increases at low soil $E_h$ in various soils, but is not a general rule as in some soils anoxic conditions lead to an increased mobilization of Ni. Those differences occur because the mobilization of Ni is often indirectly affected by $E_h$, e.g. through $E_h$-dependent pH changes, co-precipitation with Fe and Mn (hydr)oxides, complexation with soil organic carbon, similar position of Ni and Mg in the soil solid phase, and/or precipitation as sulphides (Rinklebe and Shaheen, 2017).

It is becoming more and more clear that the root architecture plays a central role in plant mineral nutrition. The architecture of rice roots is quite complex and determined by three different types of roots in term of functionality: 1) crown roots (CRs) which differentiate and emerge from the node of both stem and tillers; 2) large lateral roots (LLRs) which originate from CRs and characterized by indeterminate growth; 3) fine later roots (FLRs) originated from CRs and LLRS and characterized by determinate growth.

Rice root architecture and functionality is markedly influenced by environmental factors including water management (Kato and Okami, 2011; Vallino et al., 2014).Thus, it not possible to exclude that AWD might modify plant mineral nutrition also influencing root architecture differently than PF. In view of the expected future different environmental conditions, as well as in the case of AWD where a rotation of soil oxygen availability is imposed, a plasticity in root traits in terms of allocational, morphological, anatomical, or developmental plasticity could improve crop performance (Sandhu et al., 2016).

### 1.4. Rice grain ionome and human nutrition

Humans require more than 22 mineral elements, which can all be supplied by an appropriate diet. In many areas of the world, poor dietary quality and micronutrient deficiencies are more widespread problems than low energy intake. The diets of populations subsisting on rice often lack Fe, Zn, Ca, Mg, Cu, I or Se. This is because rice grains are not per se dense in essential micronutrients and, in addition, because they are usually consumed as white grain, i.e. after removing by milling the bran (the hard outer layers that consist of the combined aleurone and pericarp) and the embryo. The caryopses parts contain most of the grain inorganic microelement and vitamins present in the bran grain. Other than dietary supplementation and industrial rice fortification, the development of rice cultivars with increased concentration of nutrients in their edible portions has been proposed as a possible solution to alleviate global malnutrition (Graham et al., 1999, 2001; White and Broadley, 2005; Murgia et al.,2012).

The identification of the rice germplasms existing variability in the accumulation of nutritional microelements within the grain, together with the identification of quantitative trait loci (QTLs) and/or genetic elements controlling this trait, might be useful in order to plan rice genetic improvement focused on the increase of grain mineral microelement density.

Since, as above discussed, field water management affect the biogeochemical cycles of mineral nutrients in the soil and thus their availability for plant uptake, a deeper knowledge about the effects of AWD in comparison to PF on rice grain mineral nutrient concentration appears to be interesting.

Cadmium is one of the most toxic heavy metal elements, it is well known to have harmful effect on human and plant health. Rice accumulates in the grain more Cd than other cereals and thus it is the main source of dietary intake of this toxic metal in rice consuming populations (Grant et al., 2008; Arao et al., 2009; Meharg et al., 2013; Hu et al., , 2016). Dietary intake of cadmium through rice is very detrimental to the human health, as the patient suffers from pain, bone fractures, osteomalacia and kidney dysfunction (Yamagata and Shigematsu, 1970; Meharg et al., 2013). The Codex Alimentarius Commission of the Food and Agriculture Organization of the United Nations (FAO)/World Health Organization (WHO) determined the allowable levels of cadmium concentration as 0.4 mg $kg^{-1}$ in milled rice. The UE Commission fixed the limit at 0.2 mg $kg^{-1}$.

Although, the details of Cd uptake and allocation in rice are still unknown (Nocito et al., 2011; Fontanili et al., 2016) it is widely acknowledged that the mechanisms involved are associated with those of divalent metal cations including Fe and Zn. Consequently, adopting an enhanced divalent cations presence in rice kernel strategy should be carefully questioned in Cd contaminated soils.

Compared to other cereals, rice accumulates a larger amount of As because it is commonly grown under flooded conditions where As mobility is high (Takahashi et al., 2004; Xu et al., 2008) and also because rice is much more efficient assimilating arsenic into its grain than other staple cereal crops. Consequently, exposure to inorganic arsenic, a nonthreshold class 1 carcinogen, in populations not suffering from elevated arsenic in drinking water is dominated by the consumption of rice. UE Commission fixed the limit for As concentration in 0.2 mg $kg^{-1}$ and 0.1 mg $kg^{-1}$ for white rice and for rice for baby food, respectively.

Among rice germplasm a good variability in the ability to exclude As into the grain when grown on contaminated paddy fields exists (Meharg et al., 2009). Altough, not completely clarified the genetic and functional bases of As exclusion from rice grain are emerging (Song et al., 2014). More information about the variability exist in among European japonica rice germplasm should be necessary.

Recently, ionomic studies, consisting in the application of high-throughput elemental analysis technologies and their integration with both bioinformatic and genetic tools, have allowed the functional analysis of many genes and gene networks underlying mineral nutrient and trace element composition of living organisms (Huang & Salt, 2016). In particular, ionomic phenotyping have been exploited in genome wide association mapping for the exploration of loci controlling the essential minerals and potentially toxic elements in rice (Norton et al., 2010; Norton et al., 2014; Pinson et al., 2015). Not detailed ioniomic studies have been developed concerning European japonica rice accessions particularly when grown under AWD.

## 1.5.Rice as cereal model

*Rice, population structure and market classification*

It has been estimated that about 140,000 accessions of rice exist in the world today; more conservative estimates put the figure around 90,000. Included are wild species, landraces, and modern varieties. This estimation, mentioned by Evenson and Gollin (1997), it has certainly increased in the last twenty years.

Several Oryza species are wide spread in the world. The most important and largely cultivated are *Oryza sativa*, spread worldwide, and *Oryza glaberrima*, now confined almost exclusively to West

Africa. Many accession of African cultivated rice *Oryza glaberrima*, however have genes introgressed from *Oryza sativa*.

Asian cultivated rice (*O. sativa*) is thought to have been domesticated from divergent populations of Asian wild rice, *Oryza rufipogon* and *Oryza nivara*, >10,000 years ago. *Oryza rupfipogon* has many ecotypes and is widely distributed across Asia to Papua New Guinea and Australia, *O. nivara* is most common in the area of South and Southeast Asia with severe dry seasons. This is clearly seen in Sri Lanka where *Oryza nivara* is almost completely confined to the dry zone and *Oryza rufipogon* to the wet zone of the country (Vaughan et al.,2008).

According with Garris et al., (2004) and Vaughan et al., (2008), the *Oryza sativa* world population shall be classified in the following ecotypes:

- Aromatic
- Aswina
- Aus
- Indica
- Rayada
- Temperate japonica
- Tropical japonica

The Europan rice market classification, divide the *Oryza sativa*rice in the following subgroups (Regulation EC 3877/1987):

- Round grain: Length ≤5.2 mm; Length/Width Ratio <2mm
- Medium grain: Length >5.2 mm and <6 mm; Length/Width Ratio >3mm
- Long A grain: Length >6 mm; Length/Width Ratio >2 mm and <3mm
- Long B grain: Length >6mm; Length/Width Ratio ≥3 mm

*Rice Genome*

Rice genome is relatively small, in comparison with the other cereals, it accounts for about 390 Mb., (Chen et al., 2002; Devos, 2005; International Rice Genome Sequencing Project, 2005; Casella, 2011). MSU Rice Genome Annotation, Release 7 estimates about 66.000 gene models, included TE-related sequences (http://rice.plantbiology.msu.edu/analyses_facts.shtml), while rap-db database reports about 54200 genes (http://rapdb.dna.affrc.go.jp/rice_docs/docs_genes_statistics.html).

The Chromosomes size has been described by Chen et al., (2002) and reported in TabLE 1.1, with the number of loci for each chromosome (nuclear DNA only), taken out of rap-db database.

***Table 1.1**. Size of each chromosome, and number of Loci, by Rap-db database*

| Chr. | Chromosome Size (Chen et al., 2002) | Number of Loci (Rap-db) |
|---|---|---|
| 1 | 51.5 Mb | 4848 |
| 2 | 43.4 Mb | 3936 |
| 3 | 47.5 Mb | 4207 |
| 4 | 36.8 Mb | 3055 |
| 5 | 33.6 Mb | 2804 |
| 6 | 35.1 Mb | 2863 |
| 7 | 33.1 Mb | 2678 |
| 8 | 33.6 Mb | 2365 |
| 9 | 27 Mb | 1953 |
| 10 | 23.7 Mb | 1896 |
| 11 | 33.7 Mb | 2170 |
| 12 | 30.9 Mb | 2017 |
| **Total** | **430 Mb** | **34792** |

*Molecular diversity*

The completed rice genome sequence, available thanks to the continuous effort in updating sequences and annotation data coming from the two Rice Genome Sequencing projects (MSU and RAP-DB), lays the foundation for comparative genomics to the other grasses based on genome structure and individual gene (Devos, 2005; (International Rice Genome Sequencing Project, 2005; Casella, 2011).

Xu et al., (2012) resequenced the genomes of 40 cultivated accessions selected from the major groups of rice and 10 accessions of their wild progenitors *(Oryza rufipogon* and *Oryza nivara).* The work not only supports the hypothesis that japonica and indica were independently domesticated, but also further suggest *japonica* was domesticated from the Chinese strain of *Oryza rufipogon*. It appears therefore that *Oryza sativa japonica* rice was first domesticated from a specific population of *Oryza rufipogon* around the middle area of the Pearl River in southern China, and that *Oryza sativa indica*

rice was subsequently developed from crosses between japonica rice and a wild ancestor (probably *Oryza nivara),* (Huang et al., 2012).

The 3,000 rice genomes project (2014), re-sequenced 3,000 *Oryza sativa* accessions from 89 countries. The study confirm the clusterization in five groups:

- Indica
- Aus (aus/boro)
- Aromatic (basmati/sadri)
- Tropical japonica
- Temperate japonica

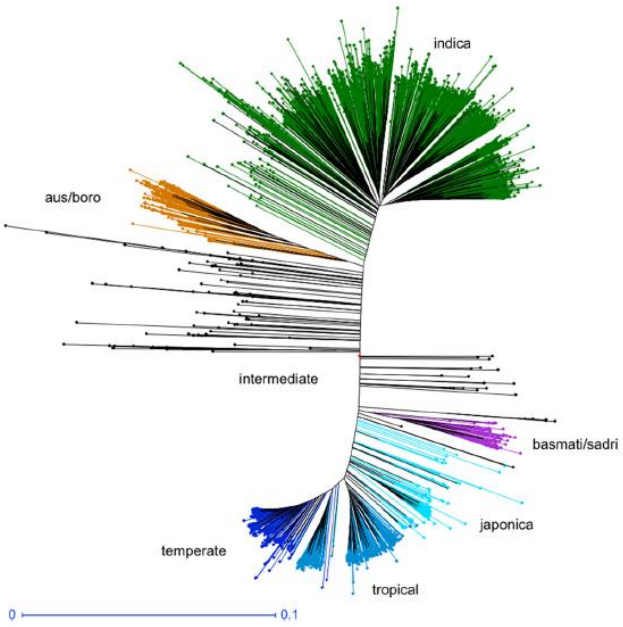



***Figure 1.3.*** *Classification of 3,000 rice accessions into five distinct varietal groups (3,000 rice genomes project, 2014).*

*Trait mapping*

During the last 10–20 years, the increasing availability of molecular markers has allowed researchers and breeders to track segments of the genome linked to specific phenotypes of interest in QTL-mapping and genome-wide association studies (Korte & Farlow, 2013). Recently developed genotyping techniques based on the reduction of the genome complexity, like Genotyping by Sequencing (GBS) (Elshire et al., 2011), are providing the marker density needed for genome-wide association study (GWAS), making the application of this procedure more feasible for different plant species (Biscarini et al., 2016).

QTL bi parental mapping: has proved, and remains, a powerful method to identify loci of the genome that cosegregate with a given trait either in F2-F3 populations or Recombinant Inbred Lines (RIL). Despite this success, QTL mapping suffers from two fundamental limitations; only allelic diversity that segregates between the two parents of the particular F2 cross or within the RIL population can be assayed and, second, the amount of recombination that occurs during the creation of the RIL population places a limit on the mapping resolution (Korte & Farlow, 2013).

GWAS Analysis and SNPs Markers: interesting markers, widely used in recent years are the Single Nucleotide Polymorphisms (SNPs), singular point mutations in the nucleotide sequence. These markers can be exploited for LD mapping, using genome-wide associations (GWAS) approaches. In this procedure, germplasm collections (where high numbers of historical recombination events are present) are used to identify statistically significant associations (LD) between marker polymorphisms (e.g. SNPs) and phenotypic variation of given traits across all chromosomes to map the genomic regions where the genes affecting the studied traits are localized (Hirschhorn & Daly, 2005; Korte & Farlow, 2013). This approach has become increasingly popular and powerful over the last few years thanks to the emergence of more cost-effective, high-throughput genotyping platforms and has become a widely adopted approach for QTL mapping in plants (Brachi et al., 2011).

*Impact of molecular genomics on plant breeding*

The above review emphasizes that despite recent advances and successful examples of molecular plant breeding, one of the current challenge in plant biology remains identifying those gene combinations that lead to significant crop improvement. This commentary closes by suggesting that the most effective approach to accelerate such efforts is to better integrate the different research disciplines and activities that form core components of molecular plant breeding (Moose & Mumm 2008).

SSRs or SNPs mapped to the genome are extensively used for Marker (or Gene) Assisted Selection (MAS, GAS) in plant breeding programs (Xu & Crouch, 2008). MAS and GAS have been successfully employed also in rice breeding (Kaur S, 2013), and molecular breeding applications bear the potential to help temperate rice contribute to the worldwide need for additional rice production in the next future (estimated 116 - 106 tons/ha by 2035); (Seck et al., 2012).

However, the success of new varieties is frequently not related to the presence of few loci with relevant phenotypic effect but also to the combination of several loci each having a small phenotypic effect. In these situations, MAs and GAS procedures can limit the breeding progress. Conversely, Genomic Selection (GS) allow the evaluation of the total effect of several loci on the phenotypic value during the selection process. GS calculate the genetic value of the individuals, which in turn, estimate the aptitude of the individuals in generating progenies with a superior phenotypic value. Molecular markers distributed across the whole genome are used to calculate genomic indexes

(Genomic Estimated Breeding Value, GEBV) which estimate the effect of the single markers (and of the different haplotypic combinations) on the traits of interest. The new high-throughput marker technologies and new statistical method, have recently and successfully introduced GS as a rice breeding instrument for improvement of complex traits (Spindel et al, 2015; Ben Hassen, 2017).

# 2. AIMS

The cultivation of high-yield lowland rice under AWD needs the knowledge of how this water management could influence quanti-qualitative traits of yield. Moreover, the knowledge of the genetic determinants of the potential adaptability of lowland rice accessions to AWD could be useful in planning the genetic improvement of rice toward water-safe field agro-technologies.

The objectives of the research were:

- to analyze the quanti-qualitative yield performances under AWD of a wide collection of japonica rice germplasm consisting in accessions cultivated in the past and/or currently in European rice areas;
- to identify QTLs and loci controlling the adaptability of japonica rice to AWD;
- to test how AWD influences rice grain ionome in comparison to PF;
- to identify QTLs and loci related to the accumulation and/or exclusion of mineral micronutrient and toxic trace elements (Cd and As) also in relation to the water management adopted.

# 3. MATERIALS AND METHODS

## 3.1. The accessions panel

The accession panel used in this study included originally 281 *Oryza sativa* varieties from the Rice Germplasm Collection maintained at the CREA-RIS (Research Centre for Cereal and Industrial Crops), (Vercelli, Italy). The panel was composed of 70 tropical japonica and 211 temperate japonica accessions (see below Most of these accessions (147) were developed in Italy, 32 from USA, 25 from Portugal, 19 from Spain, 10 from Bulgaria, 10 from Argentina, 6 from France and the remaining were developed elsewhere but considered well adapted to Italian agro-climatic conditions. Detailed information regarding the accessions are reported in the Annex 1.

## 3.2. Experimental design

Field trials were set up in the experimental fields of the CREA-Research Centre for Cereal and Industrial Crops in Vercelli (Piedmont, Italy), coordinates 45°19'204"N, 8°22'25,35"E (WGS84). According to the USDA classification the soil in the area utilized is a sub acid loam-type soil. In the Table 3.1a and 3.1b the chemical-physical characteristic of the soil the soil and its contents in mineral micronutrients and trace elements are reported, respectively. Sampling and the measure were carried out according to the document "*Metodi ufficiali di analisi chimica del suolo*" published in Gazzetta Ufficiale n. 248, 21.10.1999).

The trials, carried out during the 2012 and 2013 growing seasons, were conducted under two different soil water management: Alternate wetting and dry (AWD) and Permanent Flooding (PF). A completely randomized block design was adopted, with three replicates per watering condition. Each plot consisted of three 170 cm-long rows at a distance of 10 cm, each row contained about 60 plants (180 plants $plot^{-1}$). The following agronomic practices for rice growth were adopted: spring plow 25 cm; heavy tears; laser leveling; fertilization (23-0-10 NPK) 0.1 t $ha^{-1}$ at pre-sowing in pre-seed (no further nitrogen input was applied in the panicular differentiation phase); harrowing; sowing in lines; herbicide (Viper jigsaw according to the rate suggested by the producer); hand cleansing (for crown rice and any off-type in the sowing bed). The sowing was performed in dry conditions for both the water treatments. In the cases of PF, the field was flooded (10 cm water) when the majority of the cultivars reached the three-leaves stage (typically after 30 days) and it was kept in this condition throughout the whole growing season until 30 days before harvesting. In the case of AWD water was applied up to soil water potential ($\Psi_{H20}$) rise to zero only when $\Psi_{H20}$ values lower than -30 kPa at -20 cm depth. The values $\Psi_{H20}$ in the AWD were evaluated (two measures per week) by 6 tensiometers (60 cm 2710 ARL series) installed in the field.

***Table 3.1a**. Soil characteristics of the experimental fields. TKN: Total Kjeldhal Nitrogen; TOC:Total Organic Carbon; CEC: Cation Exchange Capability. Data are the mean of 6 determinations ± S.D.*

| | | |
|---|---|---|
| Sand | % | 47.80 ± 1.8 |
| Clay | % | 9.41 ± 0.7 |
| Loam | % | 42.80 ± 2.1 |
| pH $_{H2O}$ | | 6.36 ± 0.12 |
| TKN** | g $kg^{-1}$ | 0.79 ± 0.1 |
| TOC | g $kg^{-1}$ | 8.38 ± 0.4 |
| Organic matter | g $kg^{-1}$ | 14.45 ± 0.1 |
| P Olsen* | mg $kg^{-1}$ | 23.61 ± 0.6 |
| CEC | c $mol^{+}$ $kg^{-1}$ | 10.78 ± 0.5 |
| C/N ratio | | 18.85 ± 0.4 |
| Permanent drying point | $cm^3_{H2O}$ $cm^{-3}_{soil}$ | 0.09 ± 0.01 |
| Water field capability | $cm^3$ $_{H2O}$ $cm^{-3}_{soil}$ | 0.23 ± 0.02 |
| Apparent density | (g $cm^{-3}$) | 1.54 ± 0.07 |
| Saturation | ($cm^3$ $_{H2O}$ /$cm^{-3}_{soil}$) | 0.42 ± 0.01 |
| Hydraulic saturation conductivity | cm $h^{-1}$ | 2.99 ± 0.4 |
| Water available | mm $m^{-1}$ | 120.77 ± 11.3 |

***Table 3.1b**. Soil content in mineral nutrient and trace elements. For each element the total fraction has been evaluated. Data are the mean of 6 determinations ± SD.*

| | | |
|---|---|---|
| K | g $kg^{-1}$ | 2.34 ± 0.1 |
| Ca | g $kg^{-1}$ | 9.46 ± 0.4 |
| Mg | g $kg^{-1}$ | 9.78 ± 0.5 |
| Fe | g $kg^{-1}$ | 20.85 ± 1.2 |
| Zn | mg $kg^{-1}$ | 45.48 ± 1.5 |
| Cu | mg $kg^{-1}$ | 23.71± 1.7 |
| Mn | mg $kg^{-1}$ | 280.26 ± 11.1 |
| Ni | mg $kg^{-1}$ | 72.60 ± 4.6 |
| Na | mg $kg^{-1}$ | 824.25 ± 94.5 |
| Cd | mg $kg^{-1}$ | 0.14 ± 0.02 |
| As | mg $kg^{-1}$ | 2.57 ± 0.3 |

During the 2012 AWD soil dropped three times to a minimum value of about -40 kPa for not more than 48 h. On the contrary in 2013 no artificial irrigation was required for the AWD field.

The 2012 and 2013 field trials data yield were carried out and made available as raw data by the CREA-Ris research group. These data sets were elaborated at the beginning of the PhD program. Instead, all the post-harvest evaluations were performed in the framework of the PhD program.

## 3.3. Phenotypic evaluations

The 281 accessions were phenotyped in field for the different trait categories listed and described below.

### 3.3.1. Agronomic Traits

*Phenological traits (whole plot)*

- DAYS to FLOWERING (DF): the number of days between sowing and flowering day; the plot was considered flowered when at least 50% of the plants extruded about 1/3 of the panicles.

- DAYS to MATURITY (DM): the number of days between sowing and maturity; maturation date was assumed when at least 50% of panicles showed 2/3 of dried rachis.
- FILLING TIME (DFM): the number of days between DF and DM.

*Morphological traits (five plant per each plot)*

- FLAG LEAF LENGTH (FLL): measure of the flag leaf from ligule to leaf point.
- FLAG LEAF WIDTH (FLW): measure of the widest part of the flag leaf.
- LEAF AREA (LA): was calculated as FLL*FLW*0.75
- PLANT TOTAL HEIGHT (PH): measure of the height of the plant, from the ground to the apical seed, when the plant is fully grown.
- PANICULAR NODE HEIGHT (PNH): measure of the height of the plant, from the ground to the last developed internode, when the plant is fully grown.
- PANICLE LENGTH (PL): difference between PH and PNH.

*Post-harvest traits*

This group of measure was performed using 100 seeds randomly chosen for each plot. The scanned images (Fig. 3.2) of these seeds were analyzed by the software WinSeedlePro V.2011 (Regent Instruments).

The morphological descriptors were evaluated in both hulled and naked seeds. The traits evaluated were:

- SEED LENGTH (SL).
- SEED WIDTH (SW).
- SEED WIDTH/LENGTH RATIO (SWLR).
- DEHULLED SEEDS LENGTH (NSL).
- AVERAGE DEHULLED SEEDS WIDTH (NSW).
- AVERAGE DEHULLED SEEDS WIDTH/LENGTH RATIO (NSWLR).

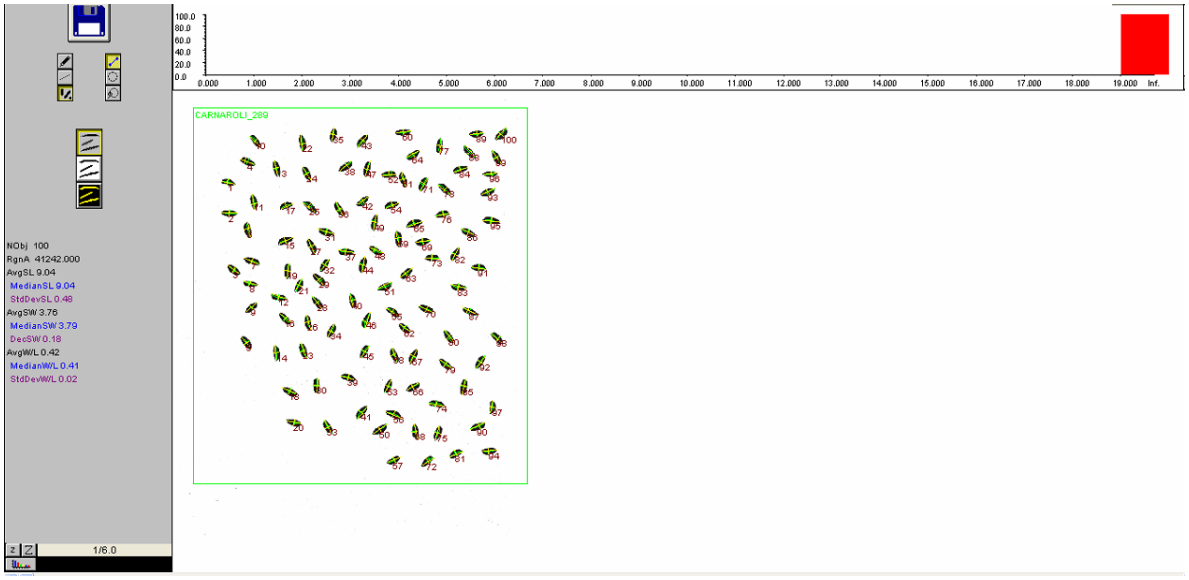



***Figure 3.1.*** *Scanned images of 100 seeds.*

*Yield traits (whole plot)*

- TILLERING (TPM): the number of tillers per linear meter.
- YIELD INDEX (PW): yield of 50 representative panicles.
- GRAIN WEIGHT (HGW and NHGW): the average weight of one hundred hulled (HGW) and de-hulled seeds.

*Physiological traits*

These traits were evaluated in the flag leaf 10 days after the flowering date, when it exerts maximum source activity on behalf of developing seeds and the nitrogen status of the plant can be considered stable, by using the Dualex® a sensor that measures flavonoids, anthocyanin and chlorophyll indices (see below). In each plot, three measures were performed on both the adaxial and the abaxial faces of the flag leaf in three plants representative of the plot; the 18 measurements thus obtained were finally averaged to obtain a plot level score. The physiology traits analyzed were:

- CHLOROPHYLL CONTENT (CHL).
- FLAVONOIDS CONTENT (FLA).
- NITROGEN BALANCE INDEX (NBI).

The Dualex® device allows instantaneous, non-destructive and simultaneous determination of the indexes CHL and FLA reflecting the chlorophyll and flavonoids leaf content, respectively. The instrument does not require both any preliminary calibration and sample preparation. The ratio between the two index (CHL/FLA) is automatically evaluated by the instrument and its value (NBI -

Nitrogen Balance Index) provides a reliable estimate of the nitrogen content of the plant (Goulas et al., 2004; Tremblay et al., 2012).

### 3.3.2 Ionomic traits

The ionome of the brown grains was defined measuring the concentrations of 12 elements: K, P, Ca, Mg, Cu, Fe, Mn, Na, Ni, Zn, Cd and As. To do this, samples of 300 mg (2 samples for each biological replicate) of flour obtained from completely milling brown grains were treated in Teflon tubes with 10 mL of 65% $HNO_3$ and then mineralized by a microwave digestion system (Anton Paar MULTIWAVE-ECO) applying a 10-min one-step temperature ramp (20-210 °C) and maintaining them at 210 °C for more than10 min. After 20 min cooling aliquots of the mineralized samples were diluted using MILLI-Q water (1: 40) in polypropylene test tubes. The concentrations of the 12 elements in the dilutes samples were determined by using an Inductively Coupled Plasma Mass Spectrometer (BRUKER Aurora-M90 ICP-MS). An aliquot of 2 $mgL^{-1}$ of an internal standard solution ($^{72}Ge$, $^{89}Y$, $^{159}Tb$) was added both to samples and calibration curve to give a final concentration of 20 g $L^{-1}$

Typical polyatomical analysis interferences were removed by using CRI (Collision-Reaction-Interface) with an $H_2$ flow of 70 mL $min^{-1}$ flown through skimmer cone.

## 3.4 Statistical and bio statistical analyses

The activities described in the following sections 3.4.2 – 3.4.5 were carried out in collaborating with the CREA-GPG AND CREA-RIS bio-informatic groups that the author of this thesis attend in order to acquire bioinformatics skills.

In detail the author participated in the activity of Filtering (3.4.2), PCoA Analysis (3.4.3), Structure Analysis (3.4.3), DAPC Analysis (3.4.3), LD decay (3.4.4), Haploview Analysis (3.4.5), GWA Analysis for Ionomic data (3.4.5), GWA Analysis for Ionomic data (3.4.5), silico Analysis for Ionomic data (3.4.5), and the script for automation of the process (3.4.5).

### 3.4.1. Phenotypic data

Boxplot of phenotypic data were tested in R environment. Analysis of variance (ANOVA) was performed using the "aov" function in R environment to assess significance of genotypes, year, Genotype x year interaction (GxE) and replicates within each environment (AWD and PF). Components of variance were estimated by fitting a mixed model by the Restricted Maximum Likelihood method, considering Genotype (G), year (E) and GxE as random factors. Broad sense heritability (H) was calculated according to (Nyquist, 1991):

$$H = \sigma^2 G / [\sigma^2 G + (\sigma^2 GE/E) + (\sigma^2 e/rE)]$$

where $\sigma^2 G$ is the genetic variance, $\sigma^2 GE$ is the genotype x environment interaction variance, $\sigma^2 e$ is the residual variance, $E$ is the number of environments, and $r$ the number of replicates.

### 3.4.2. Genotypic data and genetic diversity analysis

Most accessions (267) of the panel were genotyped-by-sequencing (GBS) following a pipeline described elsewhere by Biscarini, et al. (2016); thirty accessions were included using the same method of imputation (Annex 1), except for the number of tags required for the alignment on the Nipponbare sequence (1 instead of 5). The set of molecular markers provided by the GBS analysis at Cornel University included 246,084 SNPs, which were located on the Os-Nipponbare-Reference-IRGSP-1.0 pseudomolecule assembly (Kawahara et al., 2013) in order to define the percentage of markers located on gene regions (Volante et al., 2017).

*Filtering*

In order to perform a structure analysis, the original SNP dataset was filtered using the PLINK (http://pngu.mgh.harvard.edu/purcell/plink/, Purcell et al., 2007) software. Markers having call rate value lower than 95% and a minimum allele frequency (MAF) lower than 2.5% were discarded.

For GWA analysis a MAF lower than 5% were discarded to avoid the risk of biased detections due to rare alleles.

After filtering for call rate and MAF, a total of 31,421 SNPs emerged and were subsequently used for GWAS analyses and 37,383 SNPs for Structure analyses.

### 3.4.3. Analysis of population structure

Paleontological Statistics (PAST) ver.3 computer software was used for Principal Coordinates Analysis (PCoA), which is also known as metric multidimensional scaling (Fig. 3.1). This software uses a linear mapping of the distances between objects onto the ordination space and the algorithm attempts to explain most of the variance in the original data set (Gower, 1966).

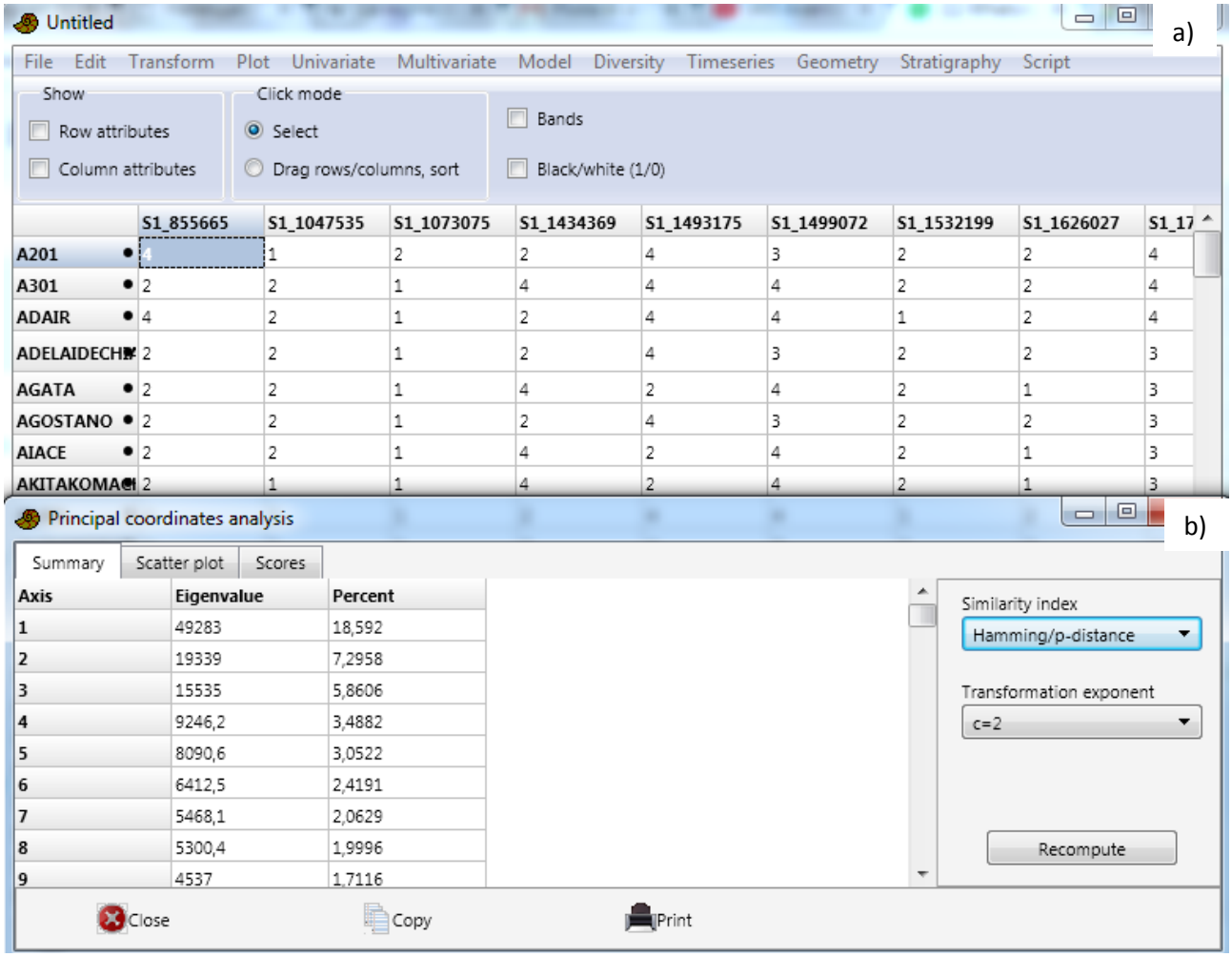



***Figure 3.5.*** *Screenshot of PAST application; (a) and (b) show the Principle Coordinate analysis.*

For the genetic diversity analyses, the number of polymorphic loci, the expected heterozygosity and the number of transitions and transvertions were computed using the Arlequin software, version3.5 (Excoffier & Lischer, 2010; Volante et al., 2017).

The whole sample and its partition into temperate or tropical *japonica* accessions were considered for the analyses, as well as the genetic groups highlighted by the Structure analysis. The divergence among the populations defined a priori according to the subspecies and among groups identified by Structure, was estimated as FST (Weir and Cockerham, 1984) with the Arlequin software, version3.5 (Excoffier and Lischer, 2010). The significance of the estimates was obtained through permutation tests, using 1000 permutations (Volante et al. 2017).

A phylogenetic tree or evolutionary tree is a branching diagram or "tree" showing the inferred evolutionary relationships among various biological species or other entities; their phylogeny is based upon similarities and differences in their physical or genetic characteristics. The taxa joined together in the tree are implied to have descended from a common ancestor. The unweighted Neighbour-joining tree was constructed in this study using a shared allele index based on a dissimilarity matrix estimated from the SNP dataset. Ape and Phyclust package of R software were used to construct the phylogenetic structure.

Discriminant Analysis of Principal Components (DAPC) using the Adegenet package 2.0.0 of R software (Jombart and Collins, 2015) was performed to identify and describe genetic clusters. Since DAPC analyzes the diversity between groups of individuals, genetic variability was decomposed using a standard multivariate ANOVA model as below:

$$\text{Total variance} = (\text{variance between groups}) + (\text{variance within groups})$$

One interesting feature of DAPC method is that it allows calculation of the contributions of alleles to the regions of the genome driving genetic divergence among groups (Jombart & Collins, 2015).

A further description of the population was obtained through neighbor-joining tree using MEGA7software (Kumar et al., 2015) to display the Jukes-Cantor algorithm that calculates the phylogenetic distance of the accessions. Finally, a model-based analysis was performed with the Structure software v2.3.4 (Pritchard et al., 2000). The parameters used in this analysis were:

- presence of admixture.
- allele frequencies correlated.
- burn-in period of 10,000 iterations.
- 20,000 Monte Carlo Markov Chain (MCMC) replications.
- K levels from 1 to 10, 5 runs per K value.

For the choice of the best number of clusters (K) the Evanno method of Δk was used, implemented in the free software Structure Harvester (Earl & von Holdt, 2012). Once defined the most probable K value, a final single run was performed using the same parameters listed above, except for:

- burn-in period of 100,000 iterations.
- 200,000 Monte Carlo Markov Chain (MCMC) replications.

Accessions with a minimum membership of 0.7 were assigned to a subpopulation, while the remaining were considered as admixed.

The phylogenetic tree, represented with iTOL (http://itol.embl.de/) was implemented with the results of the Structure analysis, together with the information relative to the varieties of the panel.

Following the structure analysis, 12 accessions were removed from the GWAS analysis, as a consequence all the filtering and structure analysis work was carried out again, with the new dataset.

### 3.4.4 Linkage disequilibrium analysis

The computation of pairwise Linkage Disequilibrium ($R^2$) among 5000 randomly selected markers was performed with the R package "LDcorSV v1.3.1" (Mangin et al., 2012), using the Structure membership matrix as a covariate.

The values were averaged in 10 kb windows as in (Biscarini et al., 2016). For each distance class, a mean value was obtained from the data of the 12 chromosomes; the resulting values were plotted against physical distance and fitted to a second-degree LOESS curve using an R script (Cleveland, 1979; Marroni et al., 2011).

A critical value of 0.2 was set as $R^2$ between unlinked loci. The physical distance corresponding to a LOESS curve value of 0.2 was assumed as LD decay in the rice panel (Volante et al., 2017).

### *3.4.5.* Association mapping

For the Genome-Wide Association analysis (GWAS), we calculated the least-square means by year for the as described by Spindel, et al. (2015).

A total number of 31,421 SNPs obtained through the filtering with plink, were used for the association mapping analysis.

Two separate GWAS analyses were performed for the two different growth conditions (PF and AWD) with the program Tassel, using a MLM that includes the kinship matrix (K) as random effect. The program was run with the following parameters: no compression, genetic and residual variance estimated for each marker (P3D OFF).

The MLM approach was fitted to the data:

$$Y_{ij} = \mu + g_i + y_j + gy_{ij} + e_{ij}$$

where $Y_{ij}$ is the phenotype, $\mu$, the overall mean, $g_i$, the genotype effect (fixed), $y_i$, the year effect (random), $gy_{ij}$, the effect of interaction genotype x year (random), and $e_{ij}$ the residual (random).

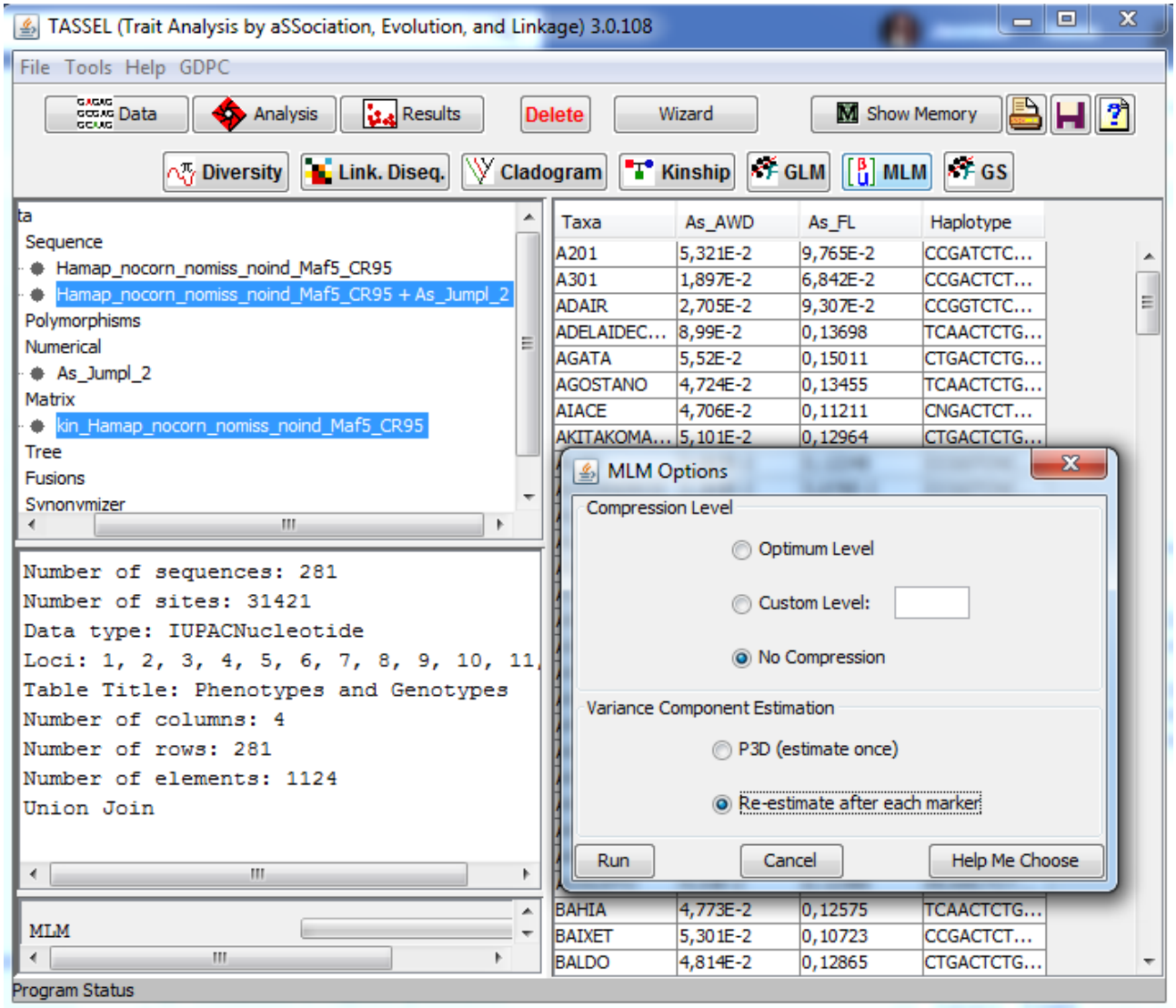



***Figure3.6.*** *Screenshot of Tassel window. An example of a run MLM configuration, with a Kinship matrix, like correction for the population structure, P3D off and no compression.*

For each marker a *p*-value of the association to the phenotypic traits was calculated; the significance threshold to declare a marker as associated was set to 0.05, after correction for multiple testing using the false discovery rate (FDR) method according to (Benjamini and Hochberg, 1995).

The False discovery rate (FDR) is one way of conceptualizing the rate of type I errors in null hypothesis testing when conducting multiple comparisons. FDR-controlling procedures are designed to control the expected proportion of rejected null hypotheses that were incorrect rejections (*false discoveries*; Benjamini and Hochberg, 1995). Thus, FDR-controlling procedures have greater power at the cost of increased rates of Type I errors. The FDR method first ranks all p-values from the smallest to the largest, and then adjusts each p-value accordingly.

$$\text{FDR corrected p} = \frac{\text{Number of associated marker with}\alpha}{\text{Total number or markers}} \text{x} \frac{\alpha(0.05)}{\sum \frac{1}{\text{Number of markers}}}$$

Manhattan plots and Q-Q plots of each trait were drawn using the R package "qqman" (Turner, 2014).

After identifying the associated markers with FDR approach, we calculated a research window with an LD blocks method.

The chromosome-wise local LD was calculated with the program Haploview v4.2 (Barrett et al., 2005). LD blocks were defined using the default settings, i.e. the method by (Gabriel et al., 2002), assuming 0.7 and 0.98 as D' lower and upper minima for strong LD, respectively. The regions associated to each trait were aligned with the results of the Haploview analysis, in order to detect adjacent associations possibly tagging a single LD block.

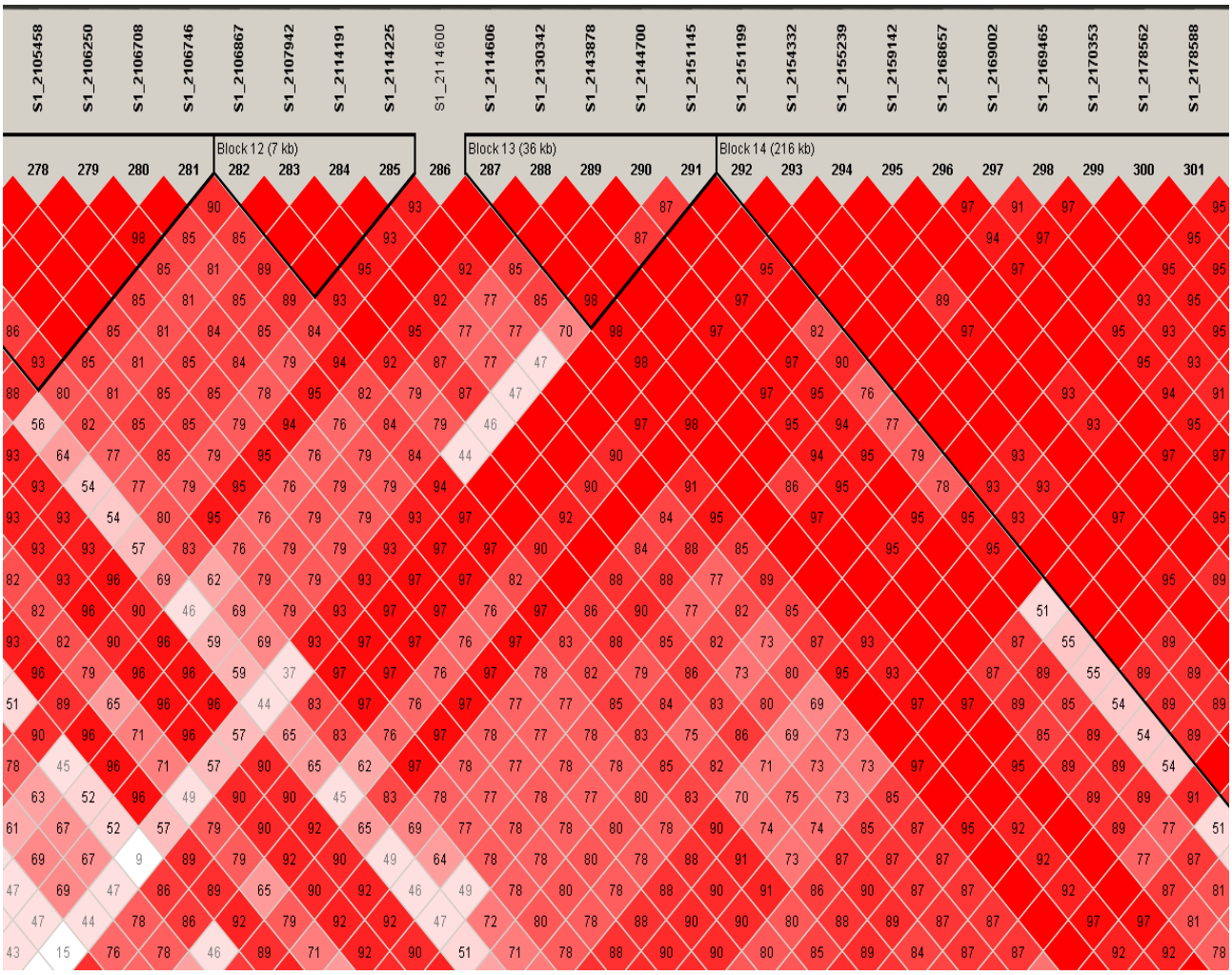



***Figure3.7.*** *Screenshot of Haploview window. The LD blocks were used to define the search windows around each associated marker.*

The regions defined by the peakmarker/region position ± 100 kbp (corresponding to an average LD decay of 0.5 as a trade-off between accuracy and power of the analysis) were screened to search for candidate genes underlying each trait.

These intervals were explored on the *Oryza sativa* reference genome (Os-Nipponbare-Reference-IRGSP-1.0, http://rapdb.dna.affrc.go.jp/download/irgsp1.html) and all the genes recorded.

Briefly, haploview groups all SNPs in several linkage blocks. When one or more SNPs considered as associated by FDR test, fell in a specific block, we considered all the block as an associated block, therefore, all the gene/locus known in literature (gene bank of Rice Annotation Project) within that block, were analyzed. For agronomic data, some very close blocks were grouped manually.

In order to validate the above results, all annotated genes included in the selected genomic windows were compared to the genes related to the phenotypic traits analyzed, available in the Oryzabase database (https://shigen.nig.ac.jp/rice/oryzabase/).

*Automation of the process*

Manually searching all the known genes in the RAP database contained in the associated blocks would have been a long and inaccurate work.

For this purpose, some scripts were created by the author thanks to the software R, which facilitated the process (see Annex 11a and 11b).

• RAP database packing scripts: This first script, using Haploview windows, divided all the genes / locus contained in the RAP database into the various linkage blocks by creating a file that we will call packaged RAP.

• Script for complete and automated analysis; this script performs the following steps

- it calculates the FDR of each single analysis by filtering those markers that are above it.
- it draws a Manhattan plot and a QQ plot with the "qqman" R package (Turner, 2014).
- it divides the associated markers into the various linkage blocks.
- it extracts the associated blocks from the packaged RAP file.
- it creates a list of candidate genes / locus.

# 4 RESULTS

Some of the results here reported have been recently published (Volante et al., 2017) with Gabrile Orasen as co-authors. Gabriele Orasen: a) partecipated in field and post-harvest evalutaion of phenotypic trait; b) permormed the ionomic analyses on the brown grains; c) attended to the GPG-CREA bioinformatic groups in order to acquire skills in bioinformatic analyses; in this framework collaborated to GWAS and gene discovery activities concerning agronomic traits; d) carried out GWAS and gene discovery activities related to ionomics.

## 4.1 Meteorological behavior along the two growing season

An Agroclimate and Information Systems unit is located 50 meters far from the fields. In Table 4.1. temperature (minimum, maximum, and average) and rainfall data for both the growing seasons are reported. The 2013 growing season was about two-fold more rainy and, on average, slightly colder than the 2012 ones.

***Table 4.1.*** *Temperatures and rainfall during the 2012 and 2013 growing season. Maximum and minimum temperatures are in bold.*

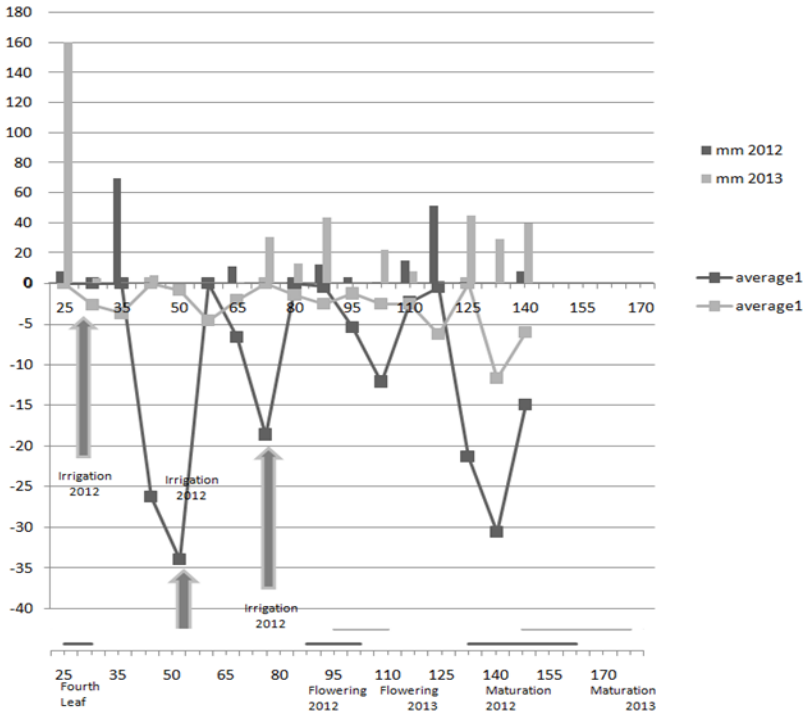
| | T-Max (°C) | | T-Min °C | | T-Average °C | | Rain fall (mm) | |
|---|---|---|---|---|---|---|---|---|
| | 2012 | 2013 | 2012 | 2013 | 2012 | 2013 | 2012 | 2013 |
| May | 31.00 | 25.50 | 7.20 | 6.50 | 18.72 | 16.27 | 12.6 | 191.2 |
| June | 32.90 | 34.30 | 13.40 | 12.00 | 23.09 | 22.11 | 71.0 | 9.0 |
| July | 35.80 | **34.80** | 5.60 | 10.00 | 24.73 | 25.35 | 12.6 | 44.8 |
| August | **38.40** | 34.60 | 5.50 | 12.90 | 25.23 | 23.25 | 39.4 | 74.4 |
| September | 33.00 | 30.80 | 5.30 | 7.50 | 19.43 | 18.94 | 68.2 | 85.6 |
| October | 27.40 | 23.10 | **0.30** | **3.80** | 14.78 | 13.99 | 60.8 | 86.6 |
| *Average* | - | - | - | - | 20.99 | 19.98 | - | - |
| *Total* | | | | | | | 264.6 | 491.6 |

In 2012 during the crop vegetative phase the soil $\Psi_{H2O}$ reached two times values lower than -30 kPa, and thus two irrigations were needed (Fig. 4.1). In the first year, 30 days after sowing the field was flooded in order to allow manual weeding). In 2013, due to highly rainfalls (Tab. 4.1), the value of soil $\Psi_{H2O}$ never reached values -15 kPa, and no irrigations were needed (Fig. 4.1).

## 4.2 Phenotyping: agronomic traits

Measures were carried out over the two-year for all the characters and the least-square means value by year was calculated as described by Spindel, et al. (2015). Results are summarized in Table 4.2. For each parameter its distribution within the accessions subjected to the two watering treatments is represented by a box-plot (Annex 2a-v), as well as its AWD/PF ratio by a histogram (Annex 4a-x).

***Figure 4.1.*** *Behavior of soil* $\Psi_{H2O}$ *in 2012 and 2013 growing seasons. Arrows indicate irrigations.*




.

From the comparison between AWD or PF results as AWD provokes, as tendency, a reduction of the mean values of almost all the parameters considered excluding DF, DM and FLA that instead increase and SWRL, NSWRL, FLW and DFM that do not result affected by the water-saving treatment (Tab. 4.2).

The analysis of variance within each treatment indicates that in the most cases the effect of the year was significant as well for all traits. The genotype effect was always highly significant for all traits in both watering techniques (see Supplementary Table S4 in Volante et al., 2017). The broad sense heritability (*H*) of traits was rather high suggesting that genetic factors significantly contribute to the variance of the measured traits.

Following the behaviour of each trait as a function of the two-field water management are described in detail.

*Physiological traits*

- CHL: CHLorophyll index.
- FLA: FLAvonoids index.
- NBI: Nitrogen Balance Index.

As is evident from the results reported in Table 4.2 and in Annex 2a and Annex 2c a tendency for lower flag leaf CHL and NBI indexes resulted in AWD with respect to PF; on the contrary FLA index resulted enhanced by AWD (Annex 2b).

However, the described behaviors are not absolute since, for all three parameters, several genotypes appeared to response differently: in AWD about 10% of the accessions showed both higher CHL and NBI than in PF and in about 31% of them a lower FLA index (Annex 2a and 2c).

*Yield-related traits*

- TPM: Total Panicles per linear Meter.
- HGW: Hundred Grain Weight.
- NHGW: Non-hulled Hundred Grain Weight.
- PW: Panicle (50 culms) Weight.

All the yield-related traits indicate that accession performances tend to be better in PF than in AWD (Annex 2d – 2g), but also in this case there are some exceptions: about 37.3% of the genotypes have a TPM higher in AWD, 2.1% of the genotypes showed HGW higher in AWD, while only 0.7 has a higher weight in AWD also for de-hulled seeds; finally, 2.8% of the genotypes have a weight of 50 panicles higher in AWD.

*Grain traits*

- SL: Seeds Length
- SW: Seeds Width
- SWLR: Seeds Width/Length Ratio
- NSL: Naked Seeds Length
- NSW: Naked Seeds Width
- NSWLR: Naked Seeds Width/Length Ratio

Even slightly AWD affected some the grain morphological traits. In detail the water-saving technique provoked tendency to a subtle but statistically significant reduction in length and width of both hulled and dehulled seeds (Annex 2h, 2i, 2k and 2l). However, overall, no important effect on both SWRL and NSWLR were observed (Annex 2j and 2m). Concerning the distribution of phenotypic value ratio for this group of traits (Annex 4h- 4m) it should be noted that: in AWD13% of the genotypes showed longer and wider hulled seeds than in PF; considering dehulled seeds this value falls to about 10%.

*Morphologic traits*

- FLL: Flag Leaf Length.
- FLW: Flag Leaf Width.
- LA: Leaf area.
- TH: Total Height.
- PNH: Panicular node height.
- PL: Panicle length.

For this data set (Annex 2n – Annex 2q), trend also pointed to decrease of the values of the traits in the case of AWD, excluding FLW that resulted unaffected with respect to PF.

Once again that described is not a constant rule. Indeed, 4 accessions showed higher FLL, 17 a larger LA and 7 a higher TH AWD. Finally, in AWD 4 accessions had increased values for both PNH and PL.

*Phenological traits*

- DF: Day to Flowering.
- DM: Day to Maturing.
- DFM: Days from Flowering to Maturity.

The results of the analysis of variance within each treatment (Tab. 4.2a for AWD and Tab. 4.2b for PF) indicates that, excluding NSL,PNH and FLL in AWD and SL in PF, the effect of the year was significant.

For all the accessions the flowering times were always longer in AWD than in PF (Annex 2t and 4v); the length of the delay was quite variable among the 281 genotypes. Also DM resulted generally affected by AWD (Annex 2u), but 31 genotypes showed a shorter DM periods in AWD than in PF (Annex 4z). Interesting, excluding 98 accessions the DFM period resulted higher in AWD than in PF (Annex 4x).

Pairise Pearson's coefficient of correlation among the traits indicates that (Fig.4.2 a and Fig.4.2b), as expected, several traits where higly correlated ($p < 0.01$).

***Table 4.2a.*** *Analysis of variance of agronomic traits in AWD (Volante et al., 2017).*

| **CHL_AV** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 22863 | 82 | 3.391 | <2e-16 | *** |
| Year | 1 | 7279 | 7279 | 302.24 | <2e-16 | *** |
| Variety:Year | 278 | 9863 | 35 | 1.473 | 1.13E-05 | *** |
| Residuals | 1071 | 25793 | 24 | | | |
| **NBI_AV** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
| Variety | 280 | 3206 | 11.5 | 3.78 | <2e-16 | *** |
| Year | 1 | 367 | 366.6 | 120.998 | <2e-16 | *** |
| Variety:Year | 278 | 1313 | 4.7 | 1.558 | 5.48E-07 | *** |
| Residuals | 1071 | 3245 | 3 | | | |
| **HGW** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
| Variety | 280 | 364 | 1.3 | 33.586 | <2e-16 | *** |
| Year | 1 | 33.2 | 33.15 | 856.573 | <2e-16 | *** |
| Variety:Year | 276 | 23 | 0.08 | 2.155 | <2e-16 | *** |
| Residuals | 1005 | 38.9 | 0.04 | | | |
| **TPM** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
| Variety | 280 | 316403 | 1130 | 4.343 | <2e-16 | *** |
| Year | 1 | 644049 | 644049 | 2475.509 | <2e-16 | *** |
| Variety:Year | 277 | 160511 | 579 | 2.227 | <2e-16 | *** |
| Residuals | 985 | 256266 | 260 | | | |
| **FLW** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
| Variety | 280 | 4311 | 15 | 13.029 | <2e-16 | *** |
| Year | 1 | 6237 | 6237 | 5278.126 | <2e-16 | *** |

| **FLA** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 36.53 | 0.13 | 4.117 | <2e-16 | *** |
| Year | 1 | 8.72 | 8.723 | 275.299 | <2e-16 | *** |
| Variety:Year | 278 | 12.85 | 0.046 | 1.459 | 1.81E-05 | *** |
| Residuals | 1072 | 33.97 | 0.032 | | | |
| **PW** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
| Variety | 280 | 1673405 | 5976 | 10.8 | <2e-16 | *** |
| Year | 1 | 248422 | 248422 | 448.94 | <2e-16 | *** |
| Variety:Year | 276 | 404559 | 1466 | 2.649 | <2e-16 | *** |
| Residuals | 939 | 519598 | 553 | | | |
| **NHGW** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
| Variety | 280 | 253.81 | 0.9065 | 48.922 | <2e-16 | *** |
| Year | 1 | 2.74 | 2.7381 | 147.778 | <2e-16 | *** |
| Variety:Year | 277 | 14.57 | 0.0526 | 2.839 | <2e-16 | *** |
| Residuals | 1016 | 18.83 | 0.0185 | | | |
| **FLL** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
| Variety | 280 | 2043430 | 7298 | 10.826 | <2e-16 | *** |
| Year | 1 | 1375 | 1375 | 2.04 | 0.153 | |
| Variety:Year | 280 | 320422 | 1144 | 1.698 | 2.13E-09 | *** |
| Residuals | 1078 | 726710 | 674 | | | |
| **LA** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
| Variety | 280 | 3.12E+08 | 1112961 | 9.059 | <2e-16 | *** |
| Year | 1 | 1.71E+08 | 1.71E+08 | 1391.393 | <2e-16 | *** |

| | | | | | | |
|---|---|---|---|---|---|---|
| Variety:Year | 280 | 535 | 2 | 1.616 | 5.69E-08 | *** |
| Residuals | 1078 | 1274 | 1 | | | |

| **PH** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 237357 | 847.7 | 21.403 | <2e-16 | *** |
| Year | 1 | 1190 | 1189.5 | 30.034 | 5.33E-08 | *** |
| Variety:Year | 278 | 15756 | 56.7 | 1.431 | 4.90E-05 | *** |
| Residuals | 1039 | 41151 | 39.6 | | | |

| **PL** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 10330 | 36.9 | 9.41 | <2e-16 | *** |
| Year | 1 | 919 | 919.3 | 234.46 | <2e-16 | *** |
| Variety:Year | 278 | 1777 | 6.4 | 1.63 | 4.02E-08 | *** |
| Residuals | 1039 | 4074 | 3.9 | | | |

| **SW** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 268.69 | 0.96 | 11.908 | <2e-16 | *** |
| Year | 1 | 5.91 | 5.909 | 73.328 | <2e-16 | *** |
| Variety:Year | 276 | 26.41 | 0.096 | 1.187 | 3.33E-02 | * |
| Residuals | 1004 | 80.91 | 0.081 | | | |

| **NSL** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 717.5 | 2.5626 | 14.259 | <2e-16 | *** |
| Year | 1 | 0 | 0 | 0 | 0.99381 | |
| Variety:Year | 277 | 65.8 | 0.2377 | 1.322 | 1.31E-03 | ** |
| Residuals | 1019 | 183.1 | 0.1797 | | | |

| **NSWLR** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 12.887 | 0.04602 | 29.618 | <2e-16 | *** |
| Year | 1 | 0.006 | 0.00622 | 4.006 | 0.0456 | * |
| Variety:Year | 277 | 0.774 | 0.00279 | 1.798 | 5.75E-11 | *** |
| Residuals | 1002 | 1.557 | 0.00155 | | | |

| **DM** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 113875 | 407 | 10.461 | <2e-16 | *** |
| Year | 1 | 89421 | 89421 | 2300.166 | <2e-16 | *** |
| Variety:Year | 276 | 42832 | 155 | 3.992 | <2e-16 | *** |
| Residuals | 1059 | 41169 | 39 | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Variety:Year | 280 | 59729540 | 213320 | 1.736 | 4.11E-10 | *** |
| Residuals | 1085 | 1.33E+08 | 122859 | | | |

| **PNH** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 186401 | 665.7 | 20.559 | <2e-16 | *** |
| Year | 1 | 17 | 17.4 | 0.537 | 0.464 | |
| Variety:Year | 278 | 13267 | 47.7 | 1.474 | 1.19E-05 | *** |
| Residuals | 1039 | 33643 | 32.4 | | | |

| **SL** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 1132.8 | 4.046 | 14.513 | <2e-16 | *** |
| Year | 1 | 23.8 | 23.848 | 85.547 | <2e-16 | *** |
| Variety:Year | 276 | 81.1 | 0.294 | 1.054 | 2.86E-01 | |
| Residuals | 1007 | 280.7 | 0.279 | | | |

| **SWLR** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 8.448 | 0.030171 | 21.094 | <2e-16 | *** |
| Year | 1 | 0.008 | 0.007994 | 5.589 | 0.0183 | * |
| Variety:Year | 276 | 0.61 | 0.00221 | 1.545 | 1.23E-06 | *** |
| Residuals | 987 | 1.412 | 0.00143 | | | |

| **NSW** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 180.17 | 0.6435 | 17.652 | <2e-16 | *** |
| Year | 1 | 1 | 1.0014 | 27.472 | 1.94E-07 | *** |
| Variety:Year | 277 | 11.19 | 0.0404 | 1.109 | 1.35E-01 | |
| Residuals | 1020 | 37.18 | 0.0365 | | | |

| **DF** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 112395 | 401 | 39.91 | <2e-16 | *** |
| Year | 1 | 74789 | 74789 | 7435.91 | <2e-16 | *** |
| Variety:Year | 280 | 14884 | 53 | 5.285 | <2e-16 | *** |
| Residuals | 1116 | 11224 | 10 | | | |

| **DFM** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 50644 | 180.9 | 5.353 | <2e-16 | *** |
| Year | 1 | 1735 | 1735.4 | 51.364 | 1.44E-12 | *** |
| Variety:Year | 276 | 29015 | 105.1 | 3.111 | <2e-16 | *** |
| Residuals | 1057 | 35713 | 33.8 | | | |

***Table 4.2b.*** *Analysis of variance of agronomic traits in PF (Volante et al., 2017).*

| **CHL_AV** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 27148 | 97 | 3.774 | <2e-16 | *** |
| Year | 1 | 2804 | 2803.8 | 109.125 | <2e-16 | *** |
| Variety:Year | 280 | 11117 | 39.7 | 1.545 | 7.57E-07 | *** |
| Residuals | 1104 | 28365 | 25.7 | | | |

| **NBI** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 4404 | 15.7 | 3.961 | <2e-16 | *** |
| Year | 1 | 599 | 598.9 | 150.828 | <2e-16 | *** |
| Variety:Year | 280 | 2010 | 7.2 | 1.808 | 1.69E-11 | *** |
| Residuals | 1104 | 4384 | 4 | | | |

| **NHGW** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 347 | 1.2393 | 172.26 | <2e-16 | *** |
| Year | 1 | 0.6 | 0.5522 | 76.75 | <2e-16 | *** |
| Variety:Year | 280 | 8.9 | 0.0318 | 4.42 | <2e-16 | *** |
| Residuals | 1079 | 7.8 | 0.0072 | | | |

| **TPM** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 443383 | 1584 | 5.412 | <2e-16 | *** |
| Year | 1 | 3354 | 3354 | 11.462 | 0.000739 | *** |
| Variety:Year | 278 | 172241 | 620 | 2.117 | <2e-16 | *** |
| Residuals | 968 | 283245 | 293 | | | |

| **FLW** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 6301 | 22.5 | 17.374 | <2e-16 | *** |
| Year | 1 | 2816 | 2815.9 | 2173.907 | <2e-16 | *** |
| Variety:Year | 280 | 783 | 2.8 | 2.159 | <2e-16 | *** |
| Residuals | 1101 | 1426 | 1.3 | | | |

| **PH** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 273430 | 977 | 47.589 | <2e-16 | *** |
| Year | 1 | 6631 | 6631 | 323.146 | <2e-16 | *** |
| Variety:Year | 280 | 7602 | 27 | 1.323 | 1.10E-03 | ** |
| Residuals | 1115 | 22880 | 21 | | | |

| **PL** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 12284 | 43.87 | 12.167 | <2e-16 | *** |
| Year | 1 | 199 | 198.88 | 55.158 | 2.21E-13 | *** |

| **FLA** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 41.54 | 0.15 | 5.882 | <2e-16 | *** |
| Year | 1 | 109.98 | 109.98 | 4360.556 | <2e-16 | *** |
| Variety:Year | 280 | 20.49 | 0.07 | 2.902 | <2e-16 | *** |
| Residuals | 1105 | 27.87 | 0.03 | | | |

| **PW** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 2299403 | 8212 | 16.076 | <2e-16 | *** |
| Year | 1 | 46050 | 46050 | 90.15 | <2e-16 | *** |
| Variety:Year | 279 | 421043 | 1509 | 2.954 | <2e-16 | *** |
| Residuals | 1044 | 533295 | 511 | | | |

| **HGW** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 476.9 | 1.703 | 79.38 | <2e-16 | *** |
| Year | 1 | 3.8 | 3.752 | 174.85 | <2e-16 | *** |
| Variety:Year | 280 | 16.5 | 0.059 | 2.74 | <2e-16 | *** |
| Residuals | 1085 | 23.3 | 0.021 | | | |

| **FLL** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 3913321 | 13976 | 17.438 | <2e-16 | *** |
| Year | 1 | 417822 | 417822 | 521.322 | <2e-16 | *** |
| Variety:Year | 280 | 540180 | 1929 | 2.407 | <2e-16 | *** |
| Residuals | 1102 | 883216 | 801 | | | |

| **LA** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 6.13E+08 | 2189211 | 13.549 | <2e-16 | *** |
| Year | 1 | 2.75E+08 | 2.75E+08 | 1700.676 | <2e-16 | *** |
| Variety:Year | 280 | 88604923 | 316446 | 1.958 | 1.93E-14 | *** |
| Residuals | 1099 | 1.78E+08 | 161582 | | | |

| **PNH** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 212677 | 760 | 42.033 | <2e-16 | *** |
| Year | 1 | 4479 | 4479 | 247.856 | <2e-16 | *** |
| Variety:Year | 280 | 5932 | 21 | 1.172 | 0.0422 | * |
| Residuals | 1114 | 20131 | 18 | | | |

| **SL** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 1629.3 | 5.819 | 178.925 | <2e-16 | *** |
| Year | 1 | 0 | 0.03 | 0.914 | 0.339 | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Variety:Year | 280 | 2097 | 7.49 | 2.077 | <2e-16 | *** |
| Residuals | 1114 | 4017 | 3.61 | | | |

| **SW** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 400.6 | 1.4307 | 197.378 | <2e-16 | *** |
| Year | 1 | 0.1 | 0.0896 | 12.363 | 0.000456 | *** |
| Variety:Year | 280 | 3.2 | 0.0113 | 1.553 | 6.16E-07 | *** |
| Residuals | 1071 | 7.8 | 0.0072 | | | |

| **NSL** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 1075.9 | 3.842 | 66.158 | <2e-16 | *** |
| Year | 1 | 1.7 | 1.749 | 30.114 | 5.08E-08 | *** |
| Variety:Year | 280 | 27.2 | 0.097 | 1.675 | 5.32E-09 | *** |
| Residuals | 1075 | 62.4 | 0.058 | | | |

| **NSWLR** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 16.362 | 0.05844 | 536.421 | <2e-16 | *** |
| Year | 1 | 0.013 | 0.01315 | 120.708 | <2e-16 | *** |
| Variety:Year | 280 | 0.101 | 0.00036 | 3.327 | <2e-16 | *** |
| Residuals | 1075 | 0.117 | 0.00011 | | | |

| **DM** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 103201 | 369 | 14.677 | <2e-16 | *** |
| Year | 1 | 39610 | 39610 | 1577.351 | <2e-16 | *** |
| Variety:Year | 280 | 23809 | 85 | 3.386 | <2e-16 | *** |
| Residuals | 1112 | 27924 | 25 | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Variety:Year | 280 | 27.1 | 0.097 | 2.976 | <2e-16 | *** |
| Residuals | 1071 | 34.8 | 0.033 | | | |

| **SWLR** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 11.068 | 0.03953 | 360.244 | <2e-16 | *** |
| Year | 1 | 0.001 | 0.00106 | 9.617 | 0.00198 | ** |
| Variety:Year | 280 | 0.073 | 0.00026 | 2.381 | <2e-16 | *** |
| Residuals | 1071 | 0.118 | 0.00011 | | | |

| **NSW** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 254.85 | 0.9102 | 309.507 | <2e-16 | *** |
| Year | 1 | 0.05 | 0.0507 | 17.244 | 3.55E-05 | *** |
| Variety:Year | 280 | 1.47 | 0.0052 | 1.781 | 6.37E-11 | *** |
| Residuals | 1074 | 3.16 | 0.0029 | | | |

| **DF** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 80854 | 289 | 47.834 | <2e-16 | *** |
| Year | 1 | 13942 | 13942 | 2309.438 | <2e-16 | *** |
| Variety:Year | 280 | 6086 | 22 | 3.601 | <2e-16 | *** |
| Residuals | 1121 | 6767 | 6 | | | |

| **DFM** | Df | Sum Sq | Mean Sq | F | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 280 | 55749 | 199 | 8.243 | <2e-16 | *** |
| Year | 1 | 6843 | 6843 | 283.275 | <2e-16 | *** |
| Variety:Year | 280 | 25788 | 92 | 3.813 | <2e-16 | *** |
| Residuals | 1110 | 26813 | 24 | | | |

## 4.3. Phenotyping: brown grain ionome

Evaluation of the concentrations of ions in brow grain of the 218 accessions grown under the two different watering techniques were carried out over the two-year. For each element the least-square means value by year was calculated as described by Spindel, et al. (2015). Results are summarized in Table 4.3. For each element the distribution of its concentration within the accessions subjected to the two watering treatments is represented by a box-plot (Annex 3a-l), as well as its AWD/PF ratio by a histogram (Annex 5a-n).

The results of the analysis of variance within each treatment (Tab. 4.3a for AWD and Tab. 4.3b for PF) indicates that excluding K, Mg and Cu, the effect of the year was significant. The genotype

effect was always highly significant for all the elements in both water management systems. In some cases, the broad sense heritability (H) of traits was very high suggesting that genetic factors significantly contribute to the variance of the measured element concentration in the brow grain.

Pairwise Pearson's coefficient correlation among the brown grain element concentrations were calculated for each water management system (Fig.4.2b and Fig. 4.2c). The analyses allowed to identify 44 strong correlations ($p<0,01$) in AWD and 47 in PF. In particular, in AWD positive strong correlations between Mg and P (r D 0.74), Mg and Fe (r D 0.61), K and P (r D 0.61) resulted.

In PF a high positive correlation between Mg and P (r D 0.78), Zn and Cu (r D 0.75), Zn and P (r D 0.7), K and P (r D 0.69), Mg and Zn (r D 0.68), Cu and P (r D 0.67), Cu and Mg (r D 0.66), Zn and Fe (r D 0.61), Mg and K (r D 0.61) were detected.

Following the behaviour of each element as a function of the two field water management is described in detail.

***Table 4.3a.*** *Analysis of variance of brown grain element concentrations in AWD.*

| **P** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 297 | 233393443 | 785837 | 14.2322 | <2.2e-16 | *** |
| Year | 1 | 1790061 | 1790061 | 32.4195 | 1.37E-08 | *** |
| Variety:Year | 286 | 88083032 | 307983 | 5.5778 | <2.2e-16 | *** |
| Residuals | 2755 | 152118745 | 55216 | | | |

| **Ca** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 297 | 711974 | 2397.2 | 25.595 | <2.2e-16 | *** |
| Year | 1 | 5431 | 5431.4 | 57.991 | 3.60E-14 | *** |
| Variety:Year | 286 | 277891 | 971.6 | 10.374 | <2.2e-16 | *** |
| Residuals | 2734 | 256064 | 93.7 | | | |

| **Fe** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 297 | 7872.7 | 26.5075 | 18.5927 | <2.2e-16 | *** |
| Year | 1 | 24.2 | 24.1916 | 16.9683 | 3.92E-05 | *** |
| Variety:Year | 286 | 2795.4 | 9.774 | 6.8556 | <2.2e-16 | *** |
| Residuals | 2619 | 3733.9 | 1.4257 | | | |

| **Mn** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 297 | 23438 | 79 | 10.185 | <2.2e-16 | *** |
| Year | 1 | 44847 | 44847 | 5788.218 | <2.2e-16 | *** |
| Variety:Year | 286 | 9041 | 32 | 4.08 | <2.2e-16 | *** |
| Residuals | 2626 | 20346 | 8 | | | |

| **Ni** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 297 | 432.43 | 1.456 | 8.6743 | <2.2e-16 | *** |
| Year | 1 | 52.51 | 52.515 | 312.8669 | <2.2e-16 | *** |
| Variety:Year | 286 | 215.19 | 0.752 | 4.4827 | <2.2e-16 | *** |
| Residuals | 2559 | 429.53 | 0.168 | | | |

| **Cd** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 297 | 6.9861 | 0.02352 | 26.64 | <2.2e-16 | *** |
| Year | 1 | 1.9738 | 1.97382 | 2235.44 | <2.2e-16 | *** |
| Variety:Year | 286 | 2.9597 | 0.01035 | 11.72 | <2.2e-16 | *** |
| Residuals | 4518 | 3.9892 | 0.00088 | | | |

| **K** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 297 | 128220959 | 431720 | 12.6965 | <2e-16 | *** |
| Year | 1 | 705 | 705 | 0.0207 | 0.8855 | |
| Variety:Year | 286 | 49915912 | 174531 | 5.1328 | <2e-16 | *** |
| Residuals | 2754 | 93644741 | 34003 | | | |

| **Mg** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 297 | 33814628 | 113854 | 16.459 | <2.2e-16 | *** |
| Year | 1 | 51908 | 51908 | 7.504 | 0.006196 | ** |
| Variety:Year | 286 | 9853190 | 34452 | 4.9804 | <2.2e-16 | *** |
| Residuals | 2749 | 19015962 | 6917 | | | |

| **Cu** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 297 | 1508.77 | 5.08 | 19.1327 | <2e-16 | *** |
| Year | 1 | 0.16 | 0.1558 | 0.5866 | 0.4438 | |
| Variety:Year | 286 | 394.9 | 1.3808 | 5.2004 | <2e-16 | *** |
| Residuals | 2697 | 716.09 | 0.2655 | | | |

| **Na** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 297 | 5834.2 | 19.64 | 14.2231 | <2.2e-16 | *** |
| Year | 1 | 613.8 | 613.78 | 444.4104 | <2.2e-16 | *** |
| Variety:Year | 286 | 2900.9 | 10.14 | 7.3441 | <2.2e-16 | *** |
| Residuals | 2497 | 3448.7 | 1.38 | | | |

| **Zn** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 297 | 33260 | 112 | 22.4148 | <2.2e-16 | *** |
| Year | 1 | 7316 | 7316 | 1464.323 | <2.2e-16 | *** |
| Variety:Year | 286 | 8254 | 28.9 | 5.7764 | <2.2e-16 | *** |
| Residuals | 2744 | 13709 | 5 | | | |

| **As** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 297 | 0.75574 | 0.00254 | 9.7749 | <2.2e-16 | *** |
| Year | 1 | 1.99339 | 1.99339 | 7657.458 | <2.2e-16 | *** |
| Variety:Year | 286 | 0.62382 | 0.00218 | 8.3789 | <2.2e-16 | *** |
| Residuals | 3975 | 1.03477 | 0.00026 | | | |

***Table 4.3b.*** *Analysis of variance of brown grain element concentrations in PF.*

| **P** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 298 | 401466435 | 1347203 | 32.2509 | <2.2e-16 | *** |
| Year | 1 | 5655225 | 5655225 | 135.3815 | <2.2e-16 | *** |
| Variety:Year | 294 | 87291164 | 296909 | 7.1078 | <2.2e-16 | *** |
| Residuals | 2844 | 118801005 | 41773 | | | |

| **Ca** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 298 | 558981 | 1875.8 | 34.2416 | <2.2e-16 | *** |
| Year | 1 | 4407 | 4407.1 | 80.4508 | <2.2e-16 | *** |
| Variety:Year | 294 | 93954 | 319.6 | 5.8336 | <2.2e-16 | *** |
| Residuals | 2832 | 155139 | 54.8 | | | |

| **Fe** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 298 | 9538.2 | 32.007 | 23.5008 | <2.2e-16 | *** |
| Year | 1 | 305.3 | 305.286 | 224.1506 | <2.2e-16 | *** |
| Variety:Year | 294 | 3062.1 | 10.415 | 7.6471 | <2.2e-16 | *** |
| Residuals | 2711 | 3692.3 | 1.362 | | | |

| **Mn** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 298 | 38280 | 128.457 | 38.1972 | <2.2e-16 | *** |
| Year | 1 | 303 | 303.37 | 90.2081 | <2.2e-16 | *** |
| Variety:Year | 294 | 5157 | 17.542 | 5.2161 | <2.2e-16 | *** |
| Residuals | 2821 | 9487 | 3.363 | | | |

| **Ni** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 298 | 43.617 | 0.146365 | 14.0602 | <2.2e-16 | *** |
| Year | 1 | 0.308 | 0.307983 | 29.5856 | 5.87E-08 | *** |
| Variety:Year | 294 | 12.723 | 0.043277 | 4.1573 | <2.2e-16 | *** |
| Residuals | 2507 | 26.098 | 0.01041 | | | |

| **Cd** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 298 | 4.5756 | 0.01535 | 89.996 | <2.2e-16 | *** |
| Year | 1 | 0.3872 | 0.3872 | 2269.449 | <2.2e-16 | *** |
| Variety:Year | 294 | 0.5437 | 0.00185 | 10.84 | <2.2e-16 | *** |
| Residuals | 5161 | 0.8805 | 0.00017 | | | |

| **K** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 298 | 174671260 | 586145 | 23.107 | <2.2e-16 | *** |
| Year | 1 | 295049 | 295049 | 11.631 | 0.000658 | *** |
| Variety:Year | 294 | 33202137 | 112932 | 4.452 | <2.2e-16 | *** |
| Residuals | 2843 | 72117433 | 25367 | | | |

| **Mg** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 298 | 40477080 | 135829 | 29.873 | <2.2e-16 | *** |
| Year | 1 | 79071 | 79071 | 17.39 | 3.14E-05 | *** |
| Variety:Year | 294 | 10061952 | 34224 | 7.527 | <2.2e-16 | *** |
| Residuals | 2836 | 12894864 | 4547 | | | |

| **Cu** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 298 | 701.25 | 2.35319 | 21.7569 | <2.2e-16 | *** |
| Year | 1 | 0.93 | 0.92875 | 8.5869 | 0.003413 | ** |
| Variety:Year | 294 | 132.42 | 0.45042 | 4.1644 | <2.2e-16 | *** |
| Residuals | 2776 | 300.25 | 0.10816 | | | |

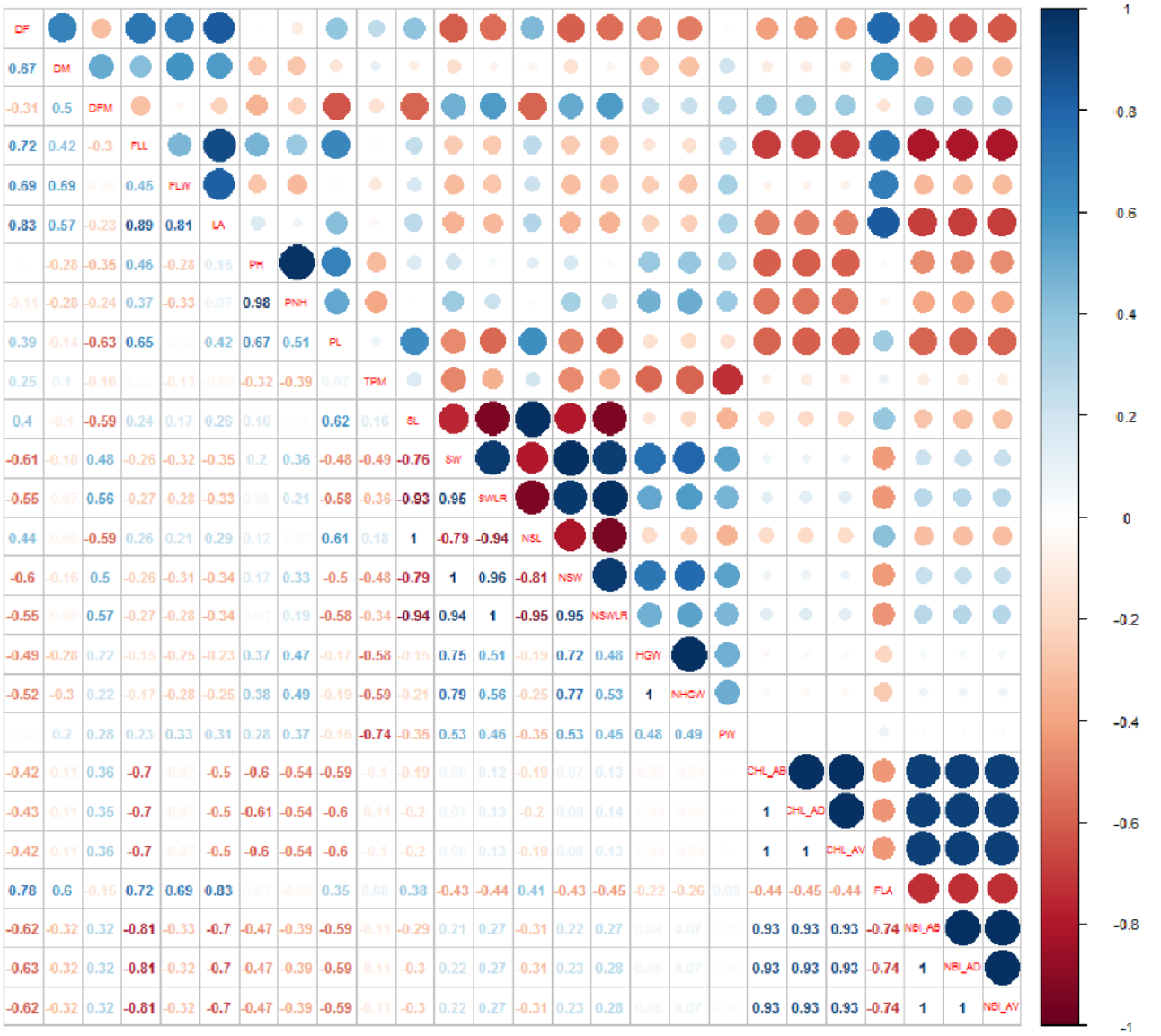
| **Na** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 298 | 8822.3 | 29.6 | 26.5043 | <2.2e-16 | *** |
| Year | 1 | 895.5 | 895.53 | 801.7401 | <2.2e-16 | *** |
| Variety:Year | 294 | 1828.1 | 6.22 | 5.5667 | <2.2e-16 | *** |
| Residuals | 2714 | 3031.5 | 1.12 | | | |

| **Zn** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 298 | 39890 | 133.9 | 49.1046 | <2.2e-16 | *** |
| Year | 1 | 4057 | 4056.7 | 1488.166 | <2.2e-16 | *** |
| Variety:Year | 294 | 4501 | 15.3 | 5.6165 | <2.2e-16 | *** |
| Residuals | 2842 | 7747 | 2.7 | | | |

| **As** | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Variety | 298 | 4.6284 | 0.01553 | 27.9886 | <2.2e-16 | *** |
| Year | 1 | 1.8611 | 1.86111 | 3353.777 | <2.2e-16 | *** |
| Variety:Year | 294 | 0.8593 | 0.00292 | 5.2669 | <2.2e-16 | *** |
| Residuals | 5508 | 3.0565 | 0.00055 | | | |

***Figure 4.2***. *Pearson correlations among agronomic traits used for GWAS recorded in the AWD (a) and PF (b) cultivation systems (Volante et al., 2017).*

**(a)**



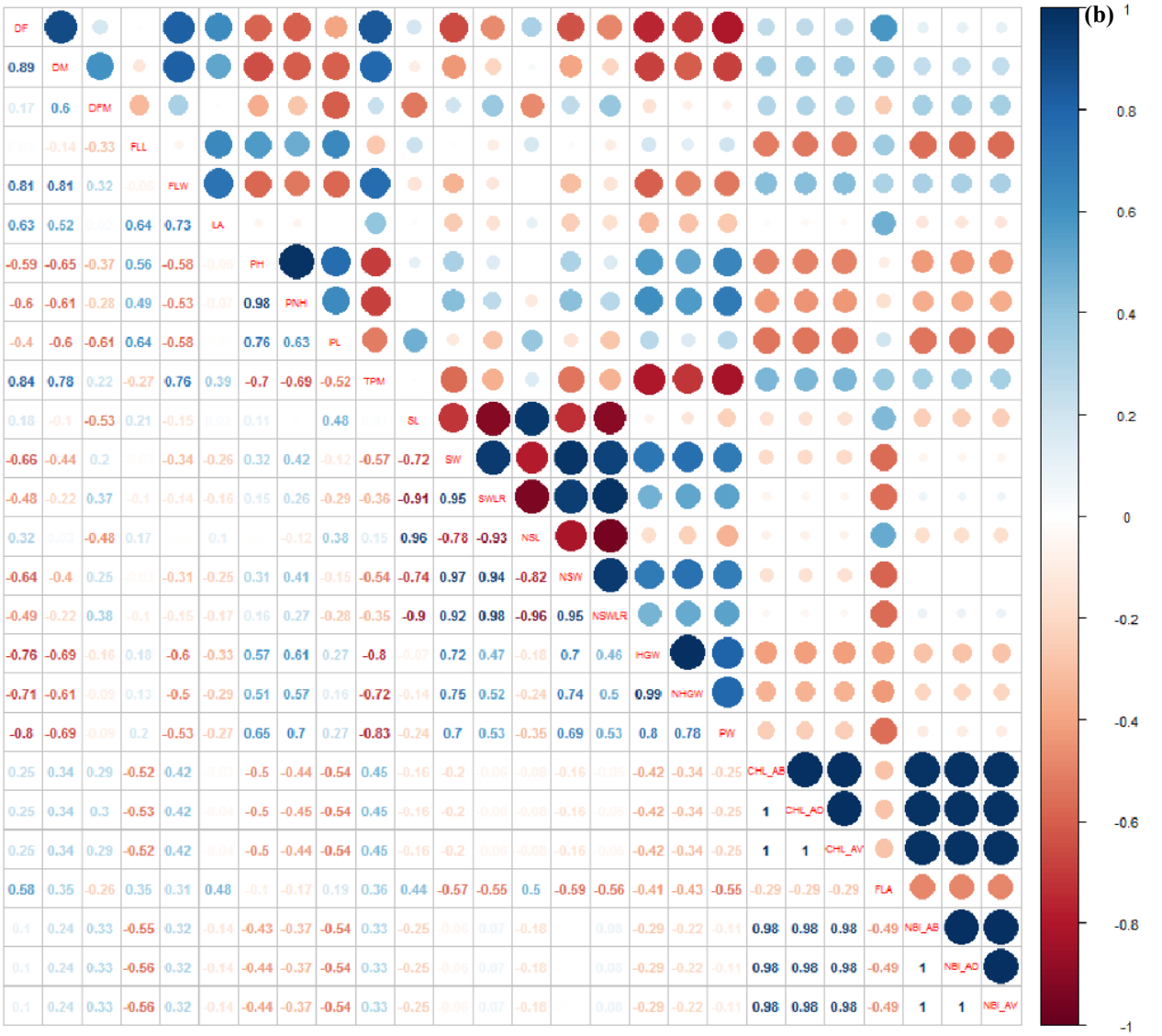
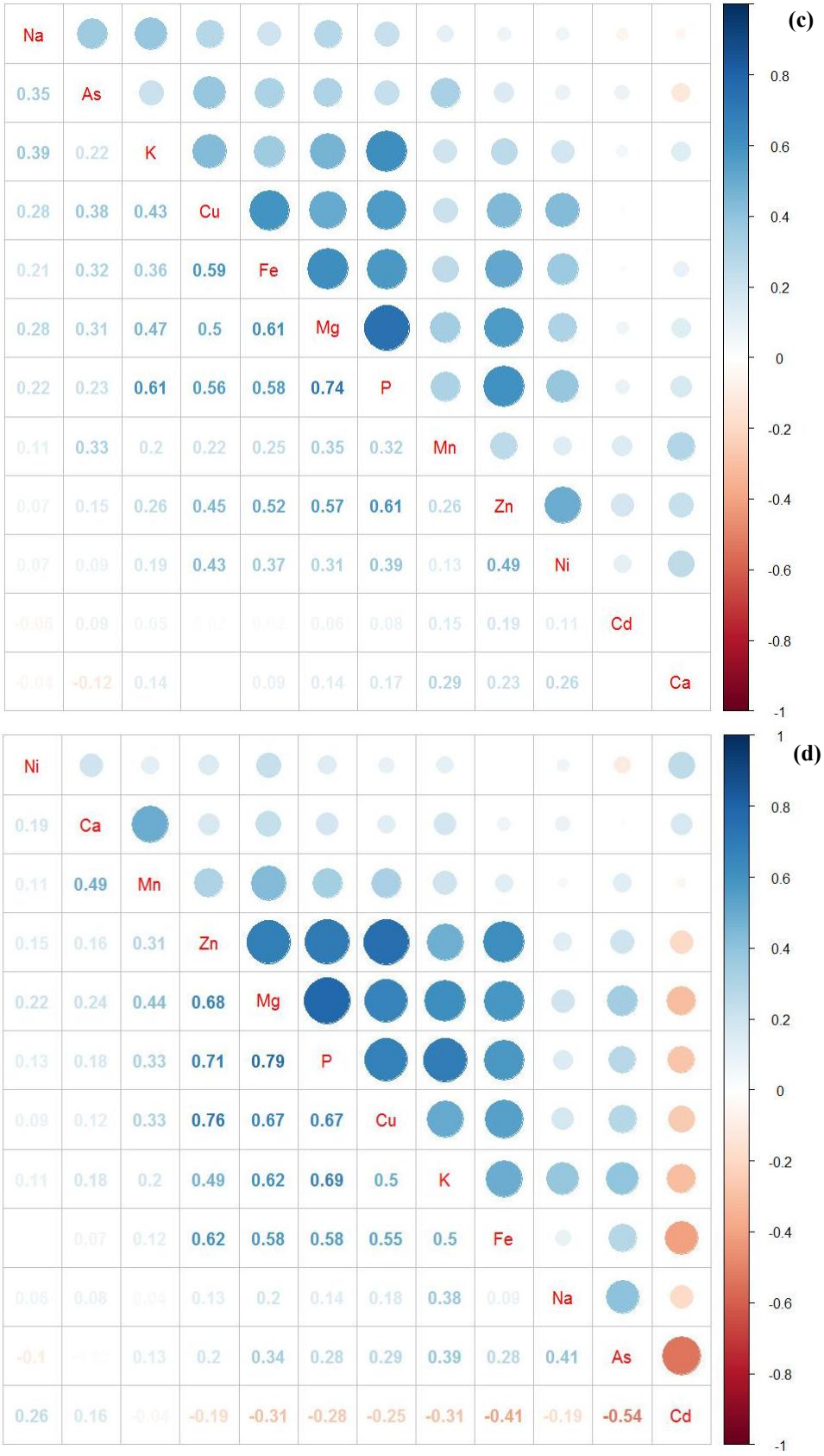

(b)
DF
DM
DFM
FLL
FLW
LA
PH
PNH
PL
TPM
SL
SW
SWLR
NSL
NSW
NSWLR
HGW
NHGW
PW
CHL_AB
CHL_AD
CHL_AV
FLA
NBI_AB
NBI_AD
NBI_AV
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1

***Figure 4.2***. *Pearson correlations among brown grain element concentrations used for GWAS recorded in the AWD (b) and PF (c) cultivation systems*



Following some details about the behaviour of each element in AWD or PF are reported

*Macronutrients (P, K, Ca, Mg)*

The average effect of AWD on the concentrations of both P and K in the brown grain was to induce a very slight increase compared to the values detected in PF (Annex 6b, Annex3a and 3b). However, in 61 and 100 accessions an opposite trend was observed (Annex 5a and 5b).

As quite similar behaviors were found concerning Ca and Mg (Annex 6b, Annex 3c and 3d). Ninety-seven and 75 accession showed an opposite behavior for Ca and Mg, respectively (Annex 5c and 5d)

*Micronutrients (Fe, Cu, Mn, Na, Zn, Ni)*

As a general behavior AWD strongly influenced the brown concentration of most of micronutrients (Tab. 4.3, Annex 3f-3j) In detail, in comparison to the value detected when plant grown in PF the effects of AWD were +11%, +59%, +6%, +473%, -25% concerning to Fe, Cu, Zn, Ni and Mn, respectively. In the case of Na AWD did not induced any change with respect to the average value observed in PF.

The AWD trends of increasing the brown grain concentration of Cu resulted for all the accession (Annex 5f), whereas in the case of Fe and Zn were not verified in the case of87 and 93 accessions, respectively. The negative effect of AWD on the accumulation of Mn in the brow grain appeared to be generalized since only in 4 accession it was not verified (Annex 5 g). Similarly, the positive effect of AWD on the Ni concentration in the brow grain was observed in each accession (Fig. 4.5l). For what concern Na in little less than 50% of the accessions its concentration resulted increased in AWD, whereas in the other about 50% resulted increased in comparison to the values measured in PF.

*Undesired trace elements (Cd, As)*

AWD acts in opposite way for what concern the accumulation of Cd and As in the brown grain (Tab. 3.3). For the former the water-saving technique induced a dramatic increase (+211%) in the concentration of the former toxic element with respect to the values detected for PF, whereas a consistent decrease in the concentration of the latter (-70%) was observed (Fig. 4.4m and 4.4n).

Although with variable intensity the effect of AWD on As was felt for each accession (Annex 5n), whereas in the case of Cd the observed increase did not occur in 26 accessions (Annex 5m).

## 4.4 SNPs panel and filtering

The initial panel of 246,084 SNPs, was filtered with Plink 2.0 provided the following panels:

A- A subset of 9,996 random SNP markers (833 markers/chromosome) was used to investigate the genetic stratification (structure analysis)
B- Tab second panel for LD and GWA analyses with call rate 95% and MAF 5%: 31,421 SNPs.

In Table 4.4, is reported the number of SNPs splitted in the chromosomes used for GWAS.

The set for GWAS analysis amounts to 31.421 SNPs, one SNPs for every 11.87 Kbp (ranging from 7.03 for chromosome 10 to 18.68 for chromosome 3).

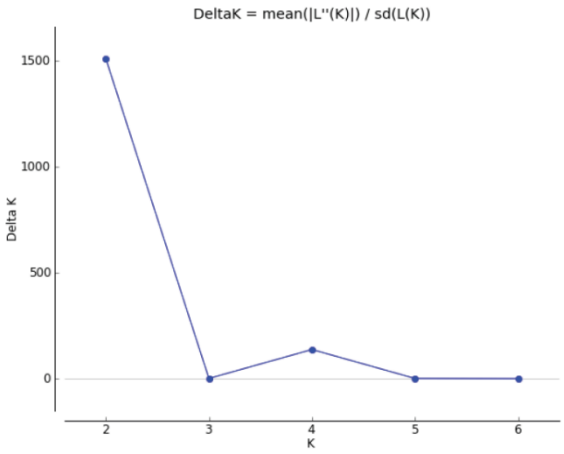
***Table 4.4.*** *Number of SNPs in the several chromosomes, for the filtering set B.*

| Chr | B |
|---|---|
| 1 | 3533 |
| 2 | 2303 |
| 3 | 1949 |
| 4 | 2814 |
| 5 | 2035 |
| 6 | 2300 |
| 7 | 2549 |
| 8 | 2727 |
| 9 | 1438 |
| 10 | 3301 |
| 11 | 3892 |
| 12 | 2580 |
| **Total** | **31421** |

## 4.5 Analysis of the population structure

*First analysis*

The first structure analysis showed that assuming $K = 4$ (the population is divided into four groups), some genotypes related to the subspecies indicated clustered separately from all the others. Even the neighbor joining tree (Fig.4.3b), grouped these genotypes into a separate branch.



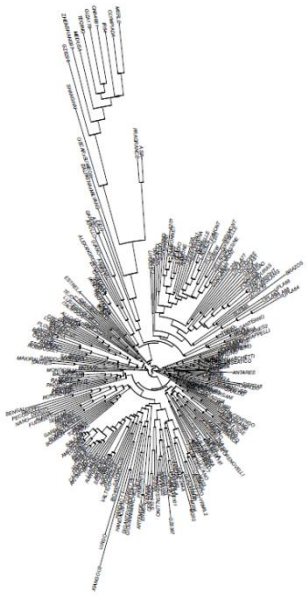



***Figure 4.3a.*** *Evanno curves.* ***Figure 4.3b.*** *Neighbor joining tree.*

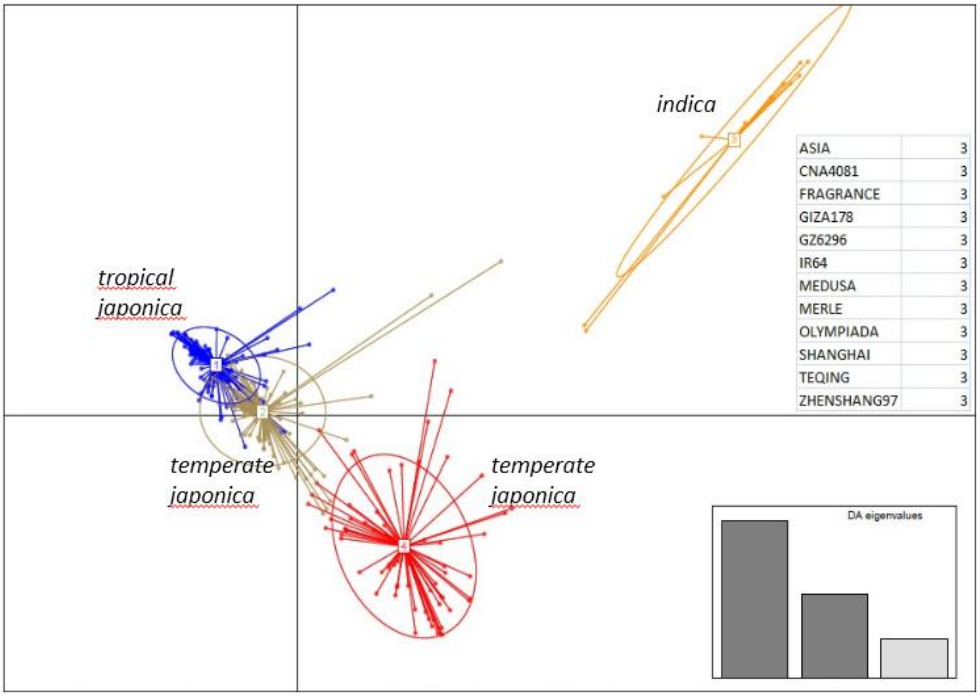



**Figure 4.4.** DAPC performed with Adegenet software in R environment The population is clustered in 4 subpopulations. The indica group in the table on the right.

A Discriminant Analysis of Principal Components (DAPC) performed with the *Adegenet* software in *R* environment pointed out (Fig. 4.4) that Asia, CNA4081, Fragrance, Giza178, Gz6296, IR64, Medusa, Merle, Olympiada, Shanghai, Tequing and Zhenshang 97 are part of this groupwhich is genuinely distant from the rest of the population.

*Second analysis*

Another parallel analysis was conducted removing the twelve genotypes listed above along with 4 genotypes with identified photoperiod problems (Ilang Ilang, Fortuna, Lady Wright and Zenith).

In this second analysis, the Evanno curve obtained from structure analysis suggested a $K = 2$ as the ideal solution. In this definitive analysis, neighbor joining tree and the structure *K* groups were graphically implemented with iTool in a single image (Fig.4.5).

The final analysis therefore suggested the following breakdown:

- K1- 49 genotypes predominantly belonging to the tropical Japonica group
- K2- 197 genotypes belonging to the tempered Japonica group
- Admixed genotypes belonging to both groups.

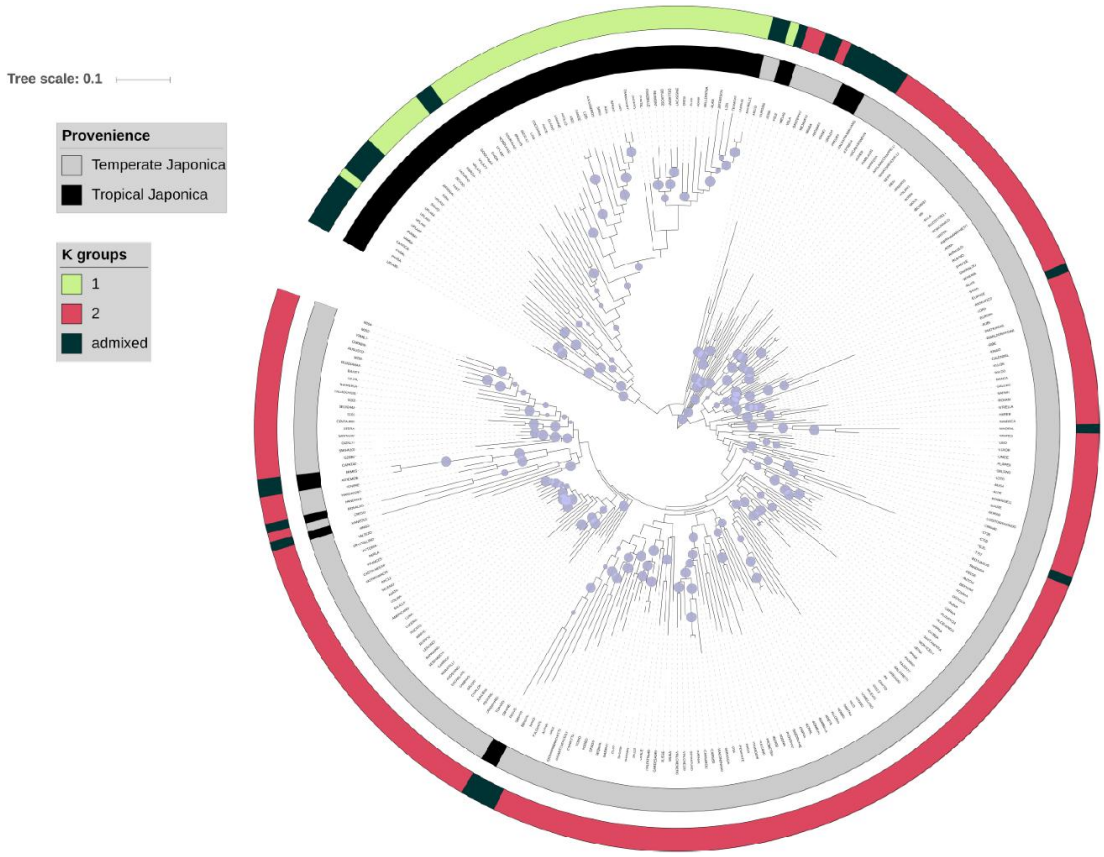



***Figure4.5.*** *Neighbor joining tree, overlapped with structure analysis The blue circles on each branch show the results of the bootstrap analysis, when higher than 0.7. The inner gray-and-black coded cycle represents the clustering of the different varieties of the panel according to O. sativa classification; the outer cycle (three-color scaled) reports the cluster organization resulting from the STRUCTURE analysis. The rice accessions showing LW/PF values for the PW, HGW, and both traits above the 95 percentile were, respectively, blue, red, and violet highlighted in the neighbor-joining tree. (from Volante et al., 2017).*

A second DAPC was carried out. DAPC is much less rigid and more elastic than the Bayesian analysis with Structure since not excluding the existence of admixed genotypes it forces each genotype into a group. The new DAPC (Fig. 4.6) highlighted the following distinctions:

- Cluster 1-99 genotypes predominantly belonging to Japan's temperate subspecies and to the Long A commodity class (mainly genotypes designated for domestic trade)
- Cluster 2-70 genotypes predominantly belonging to the tropical japonica subspecies, and to the Long B class
- Cluster 3-112 genotypes predominantly belonging to the temperate Japanese subspecies, and to Long A class (mainly parboiled genotypes), Medium and Round

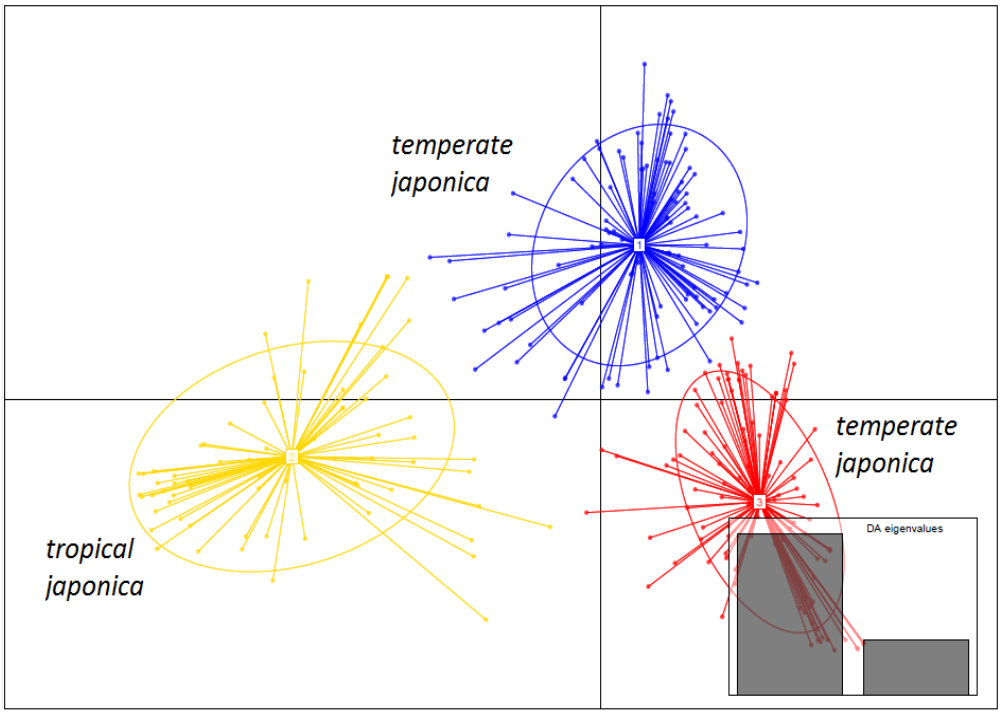



***Figure 4.6.****DAPC performed with Adegenet software in R environment The population is clustered in 3 subpopulations. The indica population has been removed.*

## 4.6. Linkage Disequilibrium and Haploview analyses

The pairwise SNP linkage disequilibrium (LD) among the 31,421 SNPs was estimated as the correlation existing between pairs of alleles across a pair of markers ($r^2$) and was calculated with R software (LDcoreSV package). The mean $r^2$ drops below 0.2 beyond approximately 900 kb (Tab. 4.4) inter-marker distance (Orasen et. al 2017).

An analysis with Haploview, performed on the panel without indica varieties (Tab. 4.5), divided the genome into 1476 linkage blocks, i.e. blocks in which the markers are in disequilibrium linkage with each other, with a D 'of at least 0.7, in accordance with the coefficients of Gabriel et. al (2002). In Table 4.5, we can see the number of blocks calculated by the program for each chromosome.

***Table 4.4.*** *LD decay in each chromosome for the whole population, for temperate japonica subgroup, and for tropical japonica subgroup*

| | | | |
|---|---|---|---|
| **Chr.** | **$R^2$ decay (Kbp)** | | |
| | **Total** | **Temperate Japonica** | **Tropical Japonica** |
| 1 | 645 | 975 | 975 |
| 2 | 715 | 955 | 855 |
| 3 | 905 | 1105 | 995 |
| 4 | 945 | 935 | 1515 |
| 5 | 1055 | 1245 | 1415 |
| 6 | 625 | 755 | 595 |
| 7 | 1215 | 1475 | 1605 |
| 8 | 1295 | 1695 | 1425 |
| 9 | 945 | 1055 | 1125 |
| 10 | 1055 | 1675 | 845 |
| 11 | 355 | 495 | 585 |
| 12 | 1185 | 1595 | 1865 |
| **Average** | 911.7 | 1163.3 | 1150 |

***Table 4.5.*** *LD blocks in each chromosomecalculated with Haploview.*

| **Chr** | **Blocks** |
|---|---|
| 1 | 157 |
| 2 | 146 |
| 3 | 81 |
| 4 | 146 |
| 5 | 57 |
| 6 | 108 |
| 7 | 103 |
| 8 | 138 |
| 9 | 63 |
| 10 | 120 |
| 11 | 271 |
| 12 | 86 |
| **tot** | **1476** |

## 4.7. GWAS: Agronomic traits

A total of 160 significant marker-trait associations (MTAs) were identified for the traits analyzed, obtained from the GWAS. It is important to remember that LSM was been used for average the data.

The $-\log(p_{value})$ for these associations ranged between 3.40 and 14.95 The lowest number of significant MTAs was detected for physiology-related traits while the plant morphology-related traits showed the highest one. Thirty-two MTAs were AWD-specific, 59 PF-specific and the remaining 69 were in common between the two watering treatments. Since 25 loci (single SNPs or chromosome regions) were associated with multiple (2 to 4) traits, the actual number of significant MTAs was 128.

Among them, 105 loci (66%) were defined by two or more markers (with an average size of the genomic regions of about 654 kbp) while the remaining associations were detected by single SNPs. In 23 MTAs no peak markers were defined due to the presence of a number of contiguous SNPs in full LD, thus showing exactly the same *q*-value in the association analysis. An analysis of genes underlying the genomic regions where significant MTAs were detected was carried out considering the genes annotated in the *Nipponbare* reference genome, in the Oryzabase and RAP data, as well as considering genes known as affecting the recorded traits in the literature. The analysis allowed the identification of positional relationships of several MTAs with genes known to affect the corresponding traits. The identified co-positional relationships are below indicated for each trait category.

For each phenotypical trait both Manhattan and Q-Q plot resulting from GWAS are reported, as well as for each trait MTAs and SNPs are summarized in a specific table.

Only the MTA-linked genes that literature somehow relates to the considered trait will be discussed later in the Discussion Section.

*Physiological traits (CHL, FLA, NBI)*

The analysis detected 82 SNPs and21 MTAs (3 AWD-specific, 13 PF-specific and 5 in common), with explained a variance ($R_2$) between 4.5 and 6.5% (Annex 7a). Eleven associations were revealed by two or more SNPs and the extent of the regions thus defined ranged between 5,607 bp and 537,739bp.

MTAs were found for CHL (3 MTAs in AWD and 5 in PF), FLA (2MTAs in AWD and 8 in PF) and NBI (3 MTAs in AWD and 5 in PF).

*Yield-related traits (TPM, HGW, NHGW, PW)*

The analysis detected 152 SNPs and 28 MTAs for yield-related traits (7 AWD-specific, 9 PF-specific and12 in common) with $R^2$ranging between 4.4 and 9.0%, the size of the regions varying between 8,562 bp and 1,588,290 bp (Annex 7b).

The MTAs detected for HGW and for the corresponding trait measured on de-hulled seeds (NHGW) was comparable (9 vs. 8).

Among the yield-related traits, we can see in Annex 7b, a congruent number of MTAs, for the weight of 50 panicles (5 AWD signals and 4 in PF), for the weight of 100 dehulled seeds (5 signals in AWD and 7 in PF), for the weight of 100 hulled seeds (6 signals in AWD and 7 in PF) and the number of culms per linear meter (3 signals in AWD and 3 in PF) were observed.

Among the MTAs, we can certainly highlight an MTA for PW on chromosome 4 and 8(AWD and PF), chromosome 4 for HGW and PW, and a common MTA in chromosome 5, for HGW and NHGW, finally on chromosome 1 for TPM (AWD and PF).

*Grain traits (SL, SW, SWLR, NSL, NSW, NSWLR)*

The analysis detected 1007 SNPs (including redundant ones), 43 MTAs (5 AWD-specific, and 15 PF-specific and the remaining shared by both conditions) with $R^2$ values ranging between 6.9 and 19.4%. Twenty-seven associations were detected by two or more markers with the highest number for NSWLR-5-1, SWLR-5-1, SW-5-1 see table Annex 7c (537, 604 and 619, respectively, including redundant) with region sizes ranging between 26,356 bp and 99,9998 bp.

In general, among the grain biometry traits we can see a high number of association signals for dressed biometrics, for SWLR (4 signals in AWD and 5 in PF), for SL (6 signals in AWD and 5 in PF), for SW (4 signals in AWD and 14 in PF) in Annex 7c.

Regarding the biometrics of the de-hulled seeds, we have different signals for NSWLR (3 signals in AWD and 3 in PF), for NSL (3 signals in AWD and 4 in PF) and NSW (7 signals in AWD and 6 in PF).

*Morphological traits (FLL, FLW, LA, TH, PNH, PL*)

A total of 526 SNPs and 50 MTAs were identified with R2 values ranging between 4.7 and 20.5% (Annex 7d). Among them, 35 were defined by two or more markers, with 5 associations (FLL-9-1, LA-4-1, FLW-4-1, PH-1-2 and PNH-1-3) showing a number of SNPs between 24 and 156. Moreover, the size of associated regions ranged from 3,587 bp to 3,504,074bp.

Among the morphological traits, a high number of association signals related to FLL (7 in AWD and 7 in PF), FLW (3 in AWD and 6 in PF), and for leaf area LA (4 in AWD and 7 in PF) were observed.

As for the grain traits, several signals for PNH (11 in AWD and 9 in PF) and total height (7 in AWD and 7 in PF) were observed.

Considering the redundant signals, four signals on chromosome 1, one redundant for FLL, PNH and TH (AWD and PF) exist. On chromosome 4 redundant signal for FLW and LA (AWD and PF) resulted.

*Phenological traits (DF, DM, DFM)*

The analysis highlighted 82 SNPs and 22 significant associations (5 AWD-specific, 11 PF specific and the remaining present in both conditions), with a percentage of the explained variance (R2) ranging between 4.9 and 11.4%. A total of 13 associations were detected through more than 1 SNP and 3 of these (DFM-11-1, DFM-3-1 and DFM-3-2) were identified by 11, 12 and 20 markers, respectively. Furthermore, the size of the associated regions ranged from 21,531 bp to 1,178,593 bp.

Among the phonological character traits, we can note in Annex 7e some association signals for DF (2 in AWD and 5 in PF), DM (5 in AWD and 7 in PF), and for DFM (4 in AWD and 5 in PF).

As for redundant signals, we can highlight only one on chromosome 10 for DF and DM (AWD and PF).

*Other results about agronomic traits*

Among the redundant signals found in several traits, further than those already mentioned, we have some in:

- TH, PNH, and FLA (chromosome 1).
- FLL, PH, PNH, FLA (chromosome 1).
- FLW, LA, NHGW, PW (chromosome 1).

## 4.7 GWAS: brown grain ionomic

We must start bysaying that all the MTAs were included in the results, because there are fewer studies concerning this topic in comparison to agronomic traits.

A total of 209 significant marker-trait associations (MTAs) were identified for the elements considered. The –log ($p_{value}$) for these associations ranged between 3.55 and 7.85. The lowest and highest number of significant MTAs was detected for Cd concentration. Seventy-one MTAs were AWD-specific, 129PF-specific and the remaining were in common between the two watering treatments. Since 28 loci (single SNPs or chromosome regions) were associated with multiple (2 or more) traits.

Among them, 168 loci (80.3%) were defined by two or more markers, while the remaining associations were detected by single SNPs. Also for Ionomic data, analysis of genes underlying the genomic regions where significant MTAs were detected was carried out considering the genes annotated in the *Nipponbare* reference genome, in the Oryzabase and RAP data, as well as considering genes known as affecting the recorded traits in the literature.

For each phenotypical trait both Manhattan and Q-Q plot resulting from GWAS are reported, as well as for each trait MTAs and SNPs are summarized in a specific table.

*Macro- and mesoelements (P, K, Ca, Mg)*

Concerning P and K the analysis detected 43 SNPs and 20 MTAs (9 AWD-specific, 9 PF-specific and 1 in common). Furthermore, the size of the SNPs interval ranged from 20 bp to 78,431 bp with $R^2$ values ranging between 4.9 and 8.8%. Two associations were detected by two or more markers with the highest number for P_04_1, P_04_2 Table 4.10

Among the MTAs some association signals related to P (6 in AWD and 7 in PF) and K (4 in AWD and 3 in PF) resulted (Annex 8a).As a redundant signal, we can highlight a MTA for P on chromosome 4 (AWD and PF).

Concerning Ca and Mg the analysis detected 78 SNPs and 28 MTAs (10 AWD-specific, 12 PF-specific and 3 in common), with $R^2$ values ranging between 4.49 and 8.34%. Furthermore, the size of the SNPs interval ranged from 5 bp to 521,996 bp.

Two associations were detected by two or more markers with the highest number for Ca_04_1, Mg_04_1 (Annex 8b).

Among the association signals related to the content of trace elements in the grain, as reported in Table 4.11, some association signals for Ca (9 in AWD and 8 in PF) and K (4 in AWD and 5 in PF) were observed. As for redundant signals, we can certainly highlight two signals for Ca on chromosome 4 (AWD and PF), and a signal for Mg on chromosome 4 (AWD and PF).

*Microelements (Fe, Cu, Mn, Na, Zn, Ni)*

The analysis detected 314 SNPs and 125 MTAs (33 AWD-specific, 86 PF-specific and 3 in common), with $R^2$ values ranging between 4.7 and 10.7%.

Three associations were detected by two or more markers with the highest number for Ni_01_1, Ni_01_2, Zn_01_1 (Annex 8c). Furthermore, the size of the SNPs interval ranged from 6 bp to 259,730 bp.

Among the MTAs related to microelement content in the grainsome Cu (4 in AWD and 10 in PF), Fe (7 in AWD and 2 in PF) and Mn (4 in AWD and 1 in PF) resulted (Tab. 4.12). Some MTAs for Na (7 in AWD and 5 in PF), Ni (9 in AWD and 67 in PF) and Zn (5 in AWD and 4 in PF) have also been found (Annex 8c).

Among the redundant MTAs, we can certainly highlight a MTA for Ni on chromosome 1 (AWD and PF), and a signal for chromosome 8 on Zn (AWD and PF).

*Undesired trace elements (As, Cd)*

The analysis detected 149 SNPs and 36MTAs (19 AWD-specific, 17 PF-specificand 3 in common). with $R^2$values ranging between 4.62 and10.85%.Furthermore, the size of the SNPs interval ranged from 13 bp to 469,527 bp.

Three associations were detected by one or more markers with the highest number for Ca_04_1, As_02_1, As_08_1, Cd_01_1 see Annex 8d

Among the association signals related to the content of toxic elements in the grain, as reported in Table 4.9, some association signals relating to As (4 in AWD and 2 in PF) and Cd (18 in AWD and

12 in PF) were found. Among the redundant signalst a signal for the CD on chromosome 2 (AWD and PF), and two signals for the CD on chromosome 8 (AWD and PF) resulted.

*Other results about brown grain ionomic*

Among the redundant signals found among the different elements, we have, in addition to those already mentioned:

• a signal for Cd and Zn on chromosome 11.

• a signal for CD and Mn on chromosome 11.

• a signal for Fe and P on chromosome 2.

• a signal for Fe and P on chromosome 7.

• a signal for Fe and P, chromosome 8.

• a signal for Ca and Ni on chromosome 3.

### 4.8. Gene identification

#### 4.8.1 Agronomics Traits

*Physiology parameters*

Several interesting candidate genes were found in the Nipponbare genomic regions where associations for physiology-related traits were detected. Following some details about these genes are reported.

- *Os03g0152400,* encodingthe enzyme 4-coumarate-CoA ligase-like (*OsAWDS1*) playing a key role in the polypropanoids biosynthesis (Falcone Ferreyra, 2012).This gene is included in the region associated with FLA on chromosome 3(FLA-3-1, PF-specific).
- *Os01g0102600,* encodingthe enzyme shikimate kinase 4 (*OsSK4*); it co-localizedn with the PF-specific locus FLA-1-1. Shikimate kinase is a key activity in shikimate pathway, the biosynthetic route to the biosynthesis of the aromatic amino acids phenylalanine, tyrosine, and tryptophan, necessary for the production of chorismate and thus in turn of flavonoids (Herrmann, 1995).

*Yield-related traits*

Some genes present in the Nipponbare reference sequence resulted in the significant associations for yield-related traits. The most interesting seem to be:

- *Os05g0187500*, *Os04g0663600* and *Os04g0615000* which are known to affect yield-related traits (trait ontology Oryzabase).
- *Os05g0187500* encoding for the *GW5* (*GRAIN WIDTH5*) protein that co-localizes with HGW-5-1 and NHGW-5-1, present in both AWD and PF conditions. This positional relationship confirms that GW5 represents a major QTL underlying rice width and weight. These gene appears to be involved in the regulation ofcell division during seed development, likely acting in the ubiquitin-proteasome pathway (Weng et al., 2008).
- *Os04g0663600* encoding the *WUSCHEL-LIKE HOMEOBOX 1* (*WOX1*); it co-localizes with the PF-specific TPM-4-1 on chromosome 4. Allelic variations in *WOX1* affected the expression of cytokinin regulators and the formation of axillary buds, which in turn deeply affect the tillering capacity (Lu et al., 2015).
- *Os04g0615000*or *Nal1* gene, described like associated to leaf-related traits (FLW and LA); it co-segregates also with grain weight (HGW-4-1, PF-specific) and panicle weight (PW-4-1, in both AWD and PF). Allelic variation in *Nal1* gene was demonstrated affecting yield acting on the number of spikelets (Fujita et al., 2013) and thus could therefore represent a candidate for the HGW-4-1 and PW-4-1 associations.

*Grain traits*

- *Os03g0407400*, *GS3,*involved in the definition of both the length and the weight of the seed(Fan et al., 2006). From the GWAS conducted in this work resulted that this gene is correlated with the traits SWLR (AWD and PF) and SL (AWD and PF).
- *Os05g0158500*, *GS5*, a gene involved in the grain width, filling and weight of the grain(Li et al., 2011); ccoherently with its putative function in the GWAS carried out the gene is associated with the traits SWLR (AWD and PF), SL (AWD and PF) and NSWLR (AWD and PF).
- Os05g0187500, *QSW5*, a gene controlling the width of the seed(Shomura et al., 2008);it was found associated with SWLR (AWD and PF), SL (AWD and PF), NSWLR (AWD and PF) andSW (PF).
- *Os05g0187500* encoding for the *GW5* protein; it has been found associated with SWLR (AWD and PF), SL (AWD and PF), NSWLR (AWD and PF) andSW (PF)

*Morphologic traits*

The analysis of the genes present in the Nipponbare reference genome and included in the regions associated to plant morphology traits, revealed some interesting co-localization relationships.

- *Os01g0118300*, identified in the PL-1-1 interval (present in both, AWD and PF) corresponds to the *SUI1 (SHORTENED UPPERMOST INTERNODE 1)* gene coding for a putative phosphatidyl serine synthase enzyme and representing a negative regulator of the uppermost internode elongation in rice (Zhu et al., 2011); since panicle traits are significantly correlated with internode elongation (Sunohara et al., 2003), the effect of this gene on internode elongation might affect PL.
- *Os08g0504700* (Oryzabase Gene Symbol Synonym *OsDOG*), identified in the PL_8_1 (PF specific) MTAs interval was identified. The gene encodes for an A20/AN1 zinc-finger protein which negatively regulates GA homeostasis and, thus, cell elongation in rice (Liu et al., 2011); it was demonstrated that rice transgenic lines with constitutive expression of *OsDOG* showed dwarf phenotypes, due to a deficiency in cell elongation.Moreover, these transgenic lines displayed reduction of Gibberellin (GA) due to a reduction of the expression of GA3ox2, which allows the production of biologically active GA, and enhanced expression of GA2ox1 and GA2ox3, which mediate conversion of active GA to inactive forms, leading to dwarfed plants and shorter panicles that did not completely emerge from the leaf sheath (Liu et al., 2011).
- *Os01g0625900*and *Os01g0626400* identified by PNH-1-2, AWD-specific, that were demonstrated to be involved in plant morphology; rice plants over-expressing *OsOFP2* (*Os01g0625900*) were reduced in height, as modulates NOX and BELL transcription factors which control vascular development, and showed a reduced expression of the GA biosynthetic enzyme GA 20-oxidase 7 reducing the GA level (Schmitz et al., 2015)
- For Os01g0626400, representing the *Dlf1* gene coding for the WRKY11 transcription factor, an association in the PNH-1-2 region was identified; mutations have pleiotropic effects on flowering time and plant height, thus the mutants exhibited semi-dwarf and late flowering phenotypes (Cai et al., 2014). In particular, *Dlf1* regulates plant height by altering cell size in internodes, supporting a putative implication of this gene in PNH determination.
- *Os01g0883800* encoding for a GA 20-oxidase and corresponding to the "green revolution gene" *Sd1*, whose mutations resulted in a reduced plant height in rice (Monna et al., 2012), was identified in the association detected for PNH-1-3, present in PF and AWD.
- TheOs04g0615000 (*Narrow leaf 1: Nal1*) locus is included in the region detected by FLW-4-1 and LA-4-1 (both present in AWD and PF); Rice *Nal1* mutants exhibited narrow leaves with a decreased number of longitudinal veins (Qi et al., 2008); altough the *Nal1* function is unknown, some evidences that mutations in this gene reduced polar auxin transport exist (Qi et al., 2008); moreover, it has been suggested that an *Nal1* allele (SPIKE – SPIKELET NUMBER) increased yield in modern *indica* cultivars since its presence resulted in a higher number of spikelets (Fujita et al., 2013).

*Phenological traits*

Four genes known to affect flowering time showed positional relationships with regions associated to phenology traits.

- *Os10g0463400* identified by two MTAs in DF-10-1 (PF) and DM-10-3 (AWD), representing the *Early heading date* (*Ehd1*) gene, which encodes a B-type response regulator that promotes flowering by activating the florigen gene *RICE FLOWERING LOCUS T 1 RFT1* (Sun et al., 2014).
- *Os02g0610500* corresponding to the *OsCOL4* gene that encodes for a protein activing *Ehd1* and belonging to the family CCT domain proteins; the gene is located closely to the DF MTAs (common to both, AWD and PF).
- *Os03g0309200* identified by a DFM MTAs (common to both, AWD and PF); the locus correspond to the gene *OsphyB* gene, encoding for a phytochrome and acting as a repressor of *OsCOL4* (Sun et al., 2014).
- Os10g0536100 in the region of DFM (AWD-specific); the locus correspond to a gene coding for the protein OsMADS56 acting as a repressor of *OsLFL1*, a putative B3 transcription factor whose over-expression decreases the expression *of Ehd1 resulting* in late flowering (Sun et al., 2014).

*Associations for multiple traits*

In several genomic regions, significant associations related to different traits were also identified.

As reported in Tab. 4.10 some locus, identified in several traits have been observed:

• FLL, TH, PNH and FLA on chromosome 4.

• HGW and HGW on chromosome 5.

• DF and DM on chromosome 10.

***Table 4.10.*** *Redundant associations for multiple agronomic traits*

| DF | DM | FLL | FLW | LA | PH | PNH | HGW | NHGW | PW | FLA | chr | ID |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | FLL_1_1 | | | PH_1_1 | PNH_1_1 | | | | FLA_1_1 | chr01 | Os01g0102600 |
| | | | FLW_4_1 | LA_4_1 | | | HGW_4_1 | | PW_4_1 | | chr04 | Os04g0615000 |
| | | | | | | | HGW_5_1 | NHGW_5_1 | | | chr05 | Os05g0187500 |
| DF_10_1 | DM_10_1 | | | | | | | | | | chr10 | Os10g0463400 |

### 4.8.2 Brown grain ionome

Among the several MTAs identified by GWAS in this thesis will be highlighted on discussed only resulting both in AWD and PF.

*Macro elements*

Two interesting genes are located in the genome windows linked to positive association concerning the brown grain P concentrations. In detail:

- Os04g0543600 identified both in AWD and PF, known as *OsCAT5* thatis indicated in GeOntology (GO) as coding for a cationic amino acid transporter;
- Os04g0543900 known as *OsGDH2* coding for a $NAD(P)^+$-dependent glutamate dehydrogenase activity involved in establish root architecture (Redillas et al., 2012); it well known as different root architecture can influence the efficency of P by the roots (White et al., 2013).
- *Os04g0605500* identified for brown grain Ca concentration, known with the synonyms of *osAWDA11, OsAWDA6, OsAWDA5 and OsPM4ATP8*, a gene encoding for the $Ca^{2+}$ P-type ATPase 11, a well-known $Ca^{2+}$ transporter (Bushart et al.,2013)
- *Os04g0543900* identified for brown grain Mg concentration known as *OsGDH2*, a glutamate dehydrogenase. This is a gene studied by Obara et al., (2011) for $NH_4^+$ uptake and nitrogen metabolism, in a root elongation study.

*Microelements*

Several candidate genes were found in the Nipponbare genomic regions in which associations for micro elements traits were detected. Below are listed only the most interesting identified loci resulting both in AWD and PF.

- Os01g0927000, identified for brown grain Ni concentration; this locus is occupied by a gene known as SDG714 involved in methyltransferase activity (Rao and Li, 2014)
- Os01g0926700, identified for brown grain Ni concentration known as OsIRX10, a glycosyl transferase gene family member (Zeng et al., 2010)
- Os08g0167000, identified for brown grain Zn concentration known as OsABCG18; this gene coding for a ABC transporter-like domain containing protein; this kind of protein are involved at root level in the uptake of forms of metals complexed by natural chelates produced and extruded into the soil by root (Matsuda et al., 2012).

*Undesired elements*

Several candidate genes were found in the Nipponbare genomic regions where associations for the toxic elements for As and Cd were detected. In the case of these two elements are listed genes linked to MTAs evidenced also in one of the two watering conditions.

- *Os07g0529000*, known as *OsEnS-107* has been identified for As concentration (AWD); this gene coding for an Isocitrate lyase activity; the same gene had been identified also in a study about Al tolerance in rice and resulted to be an orthologue of a maize isocitrate lyase involved in conferring Al tolerance in maize (Famoso et al., 2011). *OsEnS-17* has been identified after studies in biochemical and genome-wide modulation in the transcriptome of rice root exposed to cadmium (Cd), arsenate [As(V)], Pb) and chromium [Cr(VI)] in hydroponic condition (Chakrabarty et al., 2009)
- *Os02g0141000* identified in the case of brown grain As concentration (AWD) corresponding to the gene coding for a *AIG2-like* family domain containing protein; this gene has been identified by microarray analysis in rice among a group of defense and stress-responsive genes, transporters, heat-shock proteins, metallothioneins, sulfate-metabolizing proteins, and regulatory genes showing differential expression in seedlings exposed to arsenate (AsV) or arsenite (AsIII).
- *Os08g0327700.1*identified in the case of brown grain Cd concentration (AWD and PF) known as *OsCCC1* coding for a cation-chloride cotransporter expressed in the root and involved in root cell elongation and osmoregulation; a rice mutant for this gene showed dramatically increase the uptake of metals (Zn, Fe) into the roots (Chen at al., 2016), however at present no characterization of mutant for what concern its ability in Cd uptake exist.

*Associations for multiple traits*

The locus Os04g0543900, has been identified from the MTAs P and Mg.

# 5 DISCUSSION

## 5.1. Suitability of the phenotypical data

All results are based on two years of field trial only. This choice is due to the high costs of work needed to carry out this kind of experiment with hundreds of genotypes.

However, LMS approach were carried out for rectify this issue.

## 5.2 Suitability of the rice panel for GWAS analyses

A large set of SNP markers obtained by Genotyping by Sequencing was used for the genomic characterization of the rice panel. The size (about 37000 for structure analisys and 31000 for GWAS analysis), were similar to those of the study of Norton et al., (2014) with 36900 SNPs and Biscarini et al., (2016) with 57000 SNPs.

A relevant fraction of these markers (51.8%) was mapped in gene regions: this fraction was higher compared to other species where the same genotyping approach was applied (e.g. 39.5% in soybean and 20.5% in oat), due to short genome, in comparison with other cultivated species (Sonah et al., 2013; Huang et al., 2014).

The average LD decay detected in the rice panel is 912 kbp. Considering the temperate and tropical japonica ecotypes individually LD result to be 1,163 kbp and 1,150 kbp, respectively. These values are higher than those previously reported by Mather et al. (2007) of about 500 Kb and 150 Kb for temperate and tropical *japonica* rice, respectively. The discrepancies could be explained considering different factors such as SNP densities and/or kinship among accessions as previously discussed (Biscarini et al., 2016; Volante et al., 2017). However, higher LD values ranging from 600 kb up to 2 Mb were also observed in other researches for *japonica* and *indica* rice (Xu et al., 2011; Kumar et al., 2015). Moreover, for a germplasm collection of temperate *japonica* rice accessions related to the panel used in the present study values of LD decay of 1,250 kbp were reported (Biscarini et al., 2016).

The stratification analyses clearly evidenced a division of the population into two clusters, each corresponding to the two main japonica rice ecotypes. Such clustering is comparable to that observed by Biscarini et al. (2016) which used a panel containing a large proportion of accessions used in this work. Finally, Courtois et al. (2012) showed that the japonica subspecies clustered into tropical and temperate varieties also in a panel of 425 accessions belonging to four main rice varietal groups. The presence in the panel here studied of admixed varieties, defined as those that do not cluster in a defined group, suggests that these accessions were developed from interspecific breeding programs with significant gene exchange between temperate and tropical *japonica*.

A large phenotypic variation for all the traits investigated in the two field water management was observed in the panel. Most of the genotypes showed more favorable phenotypic values if grown under PF, as noted by Biscatini et al., (2016). Nevertheless, for all the traits analyzed the performances of some accessions in the AWD were not so much different to those obtained under PF. Only in the case of the brown grain concentrations of Cd, Ni and As, the effect of AWD was the same in all the accessions. Such differences between AWD and PF in ionomic data are comparable to that observed by Norton et al. (2017). This general behavior, together with the reasonably high broad sense heritability values calculated, suggested that the panel was adequate for investigating adaptability to AWD in temperate *japonica* rice.

In greater detail, considering PW as the most important yield-related trait, in some accessions the value of the ratio AWD/PF resulted higher than 1, suggesting that these accessions should be more adapted to AWD water management. Moreover, if the most promising accessions for the AWD/PF ratio in the PW trait are considered (e.g. the first 10), it is possible to highlight that these accessions also showed values of the ratios for other yield-, phenology- morphology-, and physiology-related traits that support their adaptation to AWD. For example:

Handao11 (AWD/PF ratio for PW 1.31) showed values slightly lower than 1 for NHGW and HGW (0.98 and 0.96, respectively) and values for DF, DM and DFM slightly higher than 1 (1.07, 1.04 and 1.06, respectively), indicating that flowering parameters in AWD are not so much affected;

- Cocodrie (AWD/PF ratio for PW 1.09) showed values of 1.17 and 0.95 for HGW and TPM, respectively;
- Campino, (AWD/PF ratio for PW 1.06), also had values of 1.01 for TPM;
- Sfera, (AWD/PF ratio for PW 1.03), showed values of 1 and 1.12 for HGW and TPM, respectively;
- Cigalon, (AWD/PF ratio for PW 0.99) recorded ratios of 0.91, 1.29 and 0.95 for HGW, TPM and PL, respectively;
- Escarlate (AWD/PF ratio for PW 0.99) showed values of 1.11 and 1.08 for TPM and PL, respectively. (Volante et al., 2017).

For ionomic data, a general trend for all elements in response to AWD don't really seem to exist; indeed, each element behaves differently, depending on genotype, but also on the effect of the water-save technology on its soil bioavailability. Generally speaking, for all the ions a good variability in their response to AWD was observed.

- Among the elements showing a decrease in their brown grain concentration in AWD with respect to PF are As, Mn;
- Na showed an AWD/PF ratio close to 1;
- Ca, Mg, P, and Zn ions have a slightly higher grain concentration in AWD;
- The grain concentration of Cd, Cu and Ni resulted to be strongly increased in AWD.

Arsenic is generally more concentrated in PF grains excluding Honduras (AWD/PF ratio 1.3), because if the soil is aerobic at grain filling, inorganic arsenic will be pre-dominantly present as arsenate, which has a reduced mobility and uptake by rice plants, while if the soil is flooded arsenite will be dominant, which is more mobile and rapidly accumulated by rice plants as noted by Norton et al., 2017. Manganese is generally more concentrated in PF grains excluding Allorio, Aiace, Alice, Savio (AWD/PF ratio 1, 1.02, 1.04, 1.15, respectively), according to Sasaki et al., (2012), that says the solubility of Mn increases under submerged conditions.

Norton et al., (2017), found a slightly higher content in AWD in an experiment in Bangladesh. This difference could be given by the different geographic area, different soils, and different genotypes used.

### 5.3. Validation of the system

Results recently published by McCouch et. al. (2016) sustain the validity of the panel's usefulness for GWA analysis. Analyzing grain rice biometric parameters, the authors evidenced the involvement of some gene emerged also in our work. In detail:

• GS3- involved in the definition of both length and weight of the seed (Fan et al., 2006) that in our work was found within MTAs for SWLR (AWD and PF), SL (AWD and PF);

• GS5- involved in establish grain width, fill and weight of grain characters (Li et al., 2011) that in our work it was found within MTAs for SWLR (AWD and PF), SL (AWD and PF), NSWLR (AWD and PF).

• GW5- involved in the width of the seed and very close to the qSW5 gene, implicated in the same parameters. In our research it was found within MTAs for SWLR (AWD and PF), SL (AWD and PF), NSWLR (AWD and PF), SW (PF).

### 5.4GWAS

#### 5.4.1 Agronomic traits

Physiological mechanisms controlling rice adaptation to soil aerobic conditions have been recently extensively reviewed (Price A. et al., 2013). They include the adaptability to different levels of oxygen availability, which decrease during flooding, leading to an increase of shoot ACC (1-aminocyclopropane-1-carboxylic acid, the ethylene precursor), and decrease of shoot ABA (abscisic acid) and cytokinin. Conversely, aerobic and drying increase shoot ABA (and possibly ACC) and decrease of shoot cytokinin. Limitation in water availability also affects the aerial part of the plant

since it is recognized that floral fertility in rice is extremely sensitive to water stress (Richards et al., 2010;Volante et al., 2017).

Considering the large number of parameters involved in adaptation to aerobic conditions, in this thesis a remarkable number of phenotypical traits of agronomical interest were monitored in 281 rice accessions grown in AWD or PF. Despite the high LD value resulting, that is considered a limit in the resolution of GWA studies, the pursued approach has allowed the identification of 32 AWD-specific MTAs and 69 MTAs in common between the two watering technologies. Among the 32 AWD-specific MTAs, 5 are related to phenological traits, 3 to physiological traits, 12 to plant morphology, 4 to seed morphology and 7 to yield-related traits.

AWD-specific MTAs are expected to derive from alleles present only in accessions that are more, or less adapted to this water condition (depending on the allelic contribution to the trait). Accumulation of these alleles conferring small fractions of improved phenotypic values for a given trait, as suggested by the estimated $R^2$ values, is expected to provide rice lines with improved adaptation to aerobic conditions. A similar output could be expected when loci showing associations detected in both the conditions (AWD and PF) are pyramided in an improved rice line. In this thesis, for yield-related traits, seven significant associations were identified in the AWD only and twelve significant associations were highlighted as significant under both water management systems.

Positional co-localization of the AWD-specific associations with previously identified QTLs for rice adaptation to reduced water availability can further support the role of the loci here identified in conferring adaptation to aerobic conditions.

Using as a selection criterion the grain yield under reproductive-stage drought, a number of large-effect QTLs for the trait in both AWD and PF have been identified and recently reported (Kumar et al., 2014; Volante et al., 2017).

To highlight positional relationships, the associations detected in this thesis concerning yield-related traits (TPM, HGW, NHGW and PW) were compared to QTLs described in Kumar et al. (2014). Among the associations specific for the AWD, PW-2-1 (from 24,941,038 bp to 24,962,679 bp), NHGW-10-1 (peak marker position at 4,699,332 bp) and HGW-12-1 (peak marker position at 16,580,963 bp; Tab. S2) showed co-positional relationships with the physical regions highlighted by:

- qDTY2.2 (2,020,512-25,865,568 bp).
- qDTY10.1 (5,352,766-18,655,769 bp).
- qDTY12.1 (14,106,460-18,155,593 bp).

In addition, several positional overlapping among MTAs identified in this thesis with position of drought-related QTLs identified through a meta-QTLs analysis in the Bala x Azucena rice mapping population exist (Khowaja et al., 2009).

The intervals defined by the AWD specific associations HGW-12-1 and HGW-12-2 (from 16,580,963 bp to 22,297,746 bp;) overlaps with a 20 Mb region defined by the markers RM247 and

RG543 (from 3,185,384bp to 23,775,332 bp) resulted related to drought avoidance by the meta-analysis.

Also in the case of PF specific associations detected for yield-related traits, relationships with previously mapped drought-related QTLs were identified. In detail, the following MTAs were related to grain yield:

- HGW-1-1 (peak marker at_23,753,836 bp).
- NHGW-1-1 (peak marker at 23,753,836 bp).
- NHGW-1-2 (peak marker at 33,101,881 bp).
- HGW-6-1 (peak marker at 18,125,239 bp).

In particular, the associations on chromosomes 1 are overlapping to qDTY1.2, qDTY1.3 3 (physical interval 7,445,627-36,734,272 bp and 24,807,508-36,734,272 bp), whereas HGW6.1 overlaps qDTY6.2 (7,592,066-19,514,172 bp).

Similar comparisons carried out for associations detected for other traits in the present thesis, highlighted additional positional overlapping. For traits related to phenology, it was observed that DFM-3-1 (PF-specific, with peak marker at 5,369,580 bp) and DFM-3-2 (present in AWD and PF, peak marker at 11,731,550 bp) overlapped with regions related to drought avoidance on chromosome 3 identified by the markers RZ474 and R1618 (from 25,128,239 bp to 30,720,415 bp). Similarly, DM-4-1 (PF-specific, peak marker at 24,756,463 bp) and DM-4-2 (AWD-specific, peak marker at 27,422,010 bp) overlapped with chromosome 4 regions involved in drought avoidance identified by the markers C513 and RM349 (from 22,349,484 bp to 32,499,619 bp).

Among the several associations detected for leaf morphology, FLW-4-1 (present in AWD and PF conditions, peak marker at 31,148,130 bp) overlapped with a 7 Kb region identified by Biscarini et al. (2016) with the marker S4_31080152, as associated to the same trait.

This locus co-localized with the Narrow leaf1 *(Nal1)* gene, which is involved in the regulation of cell division that affect leaf width and plant height by influencing auxin signaling (Bing et al., 2006; Jiang et al., 2015; Volante et al., 2017).

Furthermore, the association FLL-1-2 (AWD-specific, peak marker at 28,599,543 bp) was included in a LD block spanning the interval 28,186,751-29,184,329 bp, containing the S1_28597986 marker identified by Biscarini et al. (2016) as associated to flag leaf length.

Finally, the associations FLW-3-1 (AWD-specific, peak marker at 35,697,407 bp) and FLL-9-1 (present in AWD and PF conditions, peak marker at 14,598,793), were previously detected by Biscarini et al. (2016) as associated to the same trait.

The PH-related associations PH-6-1, PH-6-2, PNH-6-1 and PNH-6-2, all present in AWD and PF conditions, are included in a large LD block (22,367,525-22,678,707 bp) which contains the marker S6_22330734 associated to the same trait by Biscarini et al (2016). In this region the *D35* and

*HDA702* genes are localized. Allelic variation in the first one demonstrated to be implicated in the GA biosynthetic pathway, therefore influencing GA levels which in turn affect plant height (Itoh et al., 2004; Matusmoto et al.,2016).

*HDA702*, instead, encodes for a histone deacetylase which is involved in plant growth and architecture through an epigenetic repression of *OsNAWD6* (Chung et al., 2009; Volante et al., 2017).

The following associations overlapped respectively with three genomic regions on chromosome 1 (around the marker B1065E10: from 38,245,917 bp to 40,367,906 bp), chromosome 3 (flanked by markers RM3894 and R1618 from 1,116,926 bp to 30,720,415 bp) and chromosome 12 (flanked by markers RM247and RG543: from 3,185,384 bp to 23,775,332 bp) identified as related to drought avoidance through meta QTL analysis (Khowaja et al.,2009).

- FLW-1-2 (PF-specific, peak marker 41,251,872 bp).
- PNH-1-3 (present in AWD and PF, peak marker at 38,457,496 bp).
- PNH-3-1 (present in AWD and PF, peak marker at 1,248,941 bp).
- LA-3-1 (PF-specific, peak marker at 3,893,432 bp).
- FLW-3-1 (PF-specific, peak marker at 12,262,559 bp).
- FLW-3-2 (PF-specific, peak marker at 29,410,266 bp).
- PL-12-1 (AWD-specific, peak marker at 11,206,556 bp).

The interval 5,500,521-5,538,628 bp on chromosome 5 represents the peak of a large region associated to SW, SWLR and NSWLR traits (associations SW-5-1, SWLR-5-1 and NSWLR-5-1). It includes a 254-kb region previously identified by Biscarini et al. (2016) associated to the same traits (seed width and seed width/length ratio) with the S5_5401194 marker as a peak.

This region contains the major QTL that regulates cell proliferation during seed development, *GW5*, therefore negatively influencing grain width and weight (Weng et al., 2008), (Volante et al., 2017).

Furthermore, overlappings with drought avoidance-related regions described by Khowaja et al. (2009) were detected for the:

- AWD-specific association SL-2-1 (peak marker at 31,933,331 bp) with the chromosome 2 region defined by markers a18438 and C601 (from 22,596,168 bp to 30,270,847 bp);
- PF-specific SW-12-1 (peak marker at 4,190,069 bp) with the chromosome 12 region defined by markers RM247 and RG543 (from 3,185,384 bp to 23,775,332 bp);
- PF-specific NSL-12-1 and NSW-12-1 (peak markers at 17,866,204 bp and 17,780,115 bp, respectively) with the chromosome 12 region defined by markers RM247 and RG543;
- PF-specific SW-12-2 (peak marker at 21,652,306 bp) with the chromosome 12 region defined by markers RM247 and RG543.

In addition the association SW-1-3 (PF-specific, peak marker 42,726,259 bp) overlapped with a region defined by the marker B1065E10 (from 38,245,917 bp to 40,367,906 bp) detected as bearing drought avoidance-related effects by meta-QTLs analysis (Khowaja et al. 2009; Volante et al., 2017).

Also six associations detected in this work on chromosome 3, present in AWD and PF conditions and related to seed morphology (NSW-3-1, NSWLR3-1, SL-3-1, NSL-3-1 and SWLR-3-1, with peak markers ranging from 7,907,626 bp to 16,378,774 bp) overlapped with a 25Mb interval detected by metaQTL analysis and defined by markers RG191 and R1618 (from 5,729,669 bp to 30,720,415 bp) where regions involved in drought avoidance were identified (Khowaja et al., 2009).

Loci identified in this work as conferring advantage under AWD (namely AWD-specific and loci in common between the two conditions) could represent suitable target for Genomic Selection approaches addressed to improve yield under aerobic conditions. Similar approaches could be undertaken to improve the other phenotypic traits investigated in the present work.

The GWAS panel used in this study has recently been used as a reference population in a Genomic Selection approach to estimate the average genomic prediction accuracies within the reference population itself (cross validation) and genomic prediction of lines-progenies of bi-parental crosses involving accessions belonging to the reference population (across generations) for complex traits investigated also in this work, namely flowering date, nitrogen balance index and yield-related traits.

In this work it has been observed that using phenotypic and genotypic data from the reference population to train the prediction model allowed prediction of the performances, in both the approaches (cross validation and across generations) with accuracies superior to 0.5, even for complex traits such as grain yield, when the parameters that affect the accuracy are optimized (Ben Hassen et al., 2017). These preliminary results support that results obtained in the present study can be exploited to develop Genomic Selections models to assist breeding for rice adaptation to aerobic conditions.

### 5.4.2. Grain ionome

Irrigation systems of alternate aerobic and anaerobic conditions are also expected to alter the redox chemistry of soils and macro and micro nutrient availability and thus uptake by roots. Just as an example phosphorus is usually more available in anoxic than aerobic soils. Therefore, root-related processes are expected to play important roles in adaptation to changing levels of water availability. This because, depending on timing and intensity of water fluctuations soil redox potential, changes in the root activities can affect nutrient access and the signaling between root and shoot (Price et al., 2013).

In general, it was observed that rice grain had comparatively higher amounts of Cd and Zinc under AWD condition than PF condition (Oono et al., 2014). It appeared that the combination of cadmium with sulphur to form cadmium sulphide at low soil redox decreased the mobilization of cadmium in soil (Bingham, et a, 1976; Hu et al., 2013; Chaney, 2015), thereby decreasing the bioavailability of

cadmium for the plants. Because of the analogous resemblance of zinc with cadmium, perhaps zinc was also available in the form of zinc sulphide under PF condition (Yoneyama et al.,2015; Chaney, 2015; Slamet-Loedin et al.,2015), which ultimately resulted in decrement of zinc and cadmium content under PF condition (Arao et al., 2009). However, the results of the *anova* test carried out within the 281 accessions concerning the brown grain Cd concentration indicate the existence of significant genotypic effects.

It is well known that PF conditions increas the As mobilization in the soil, the absorption by the roots and then the accumulation in the grain. The higher mobilization of the element on soil under PF is due to the reduction of arsenate (AsV) to arsenite (III) that is released from the adsorption surfaces of iron oxyhydroxides into the solution phase (Hu et al., 2013 Sun et al., 2014). Arsenite is efficently absorbed by the root trough the LSi1 transporter, a protein belonging to the aquaporine family (Ma et al., 2006, 2008), whereas As(V) is absorbed by a $H^+$-phosphate cotransport mecahimis (Catarecha et al., 2007).

The opposite effect of AWD and PF on the accumulation of Cd and As into the brown grain limits the possibility to adopt one of the two watering technologies for producing safe grain when plant groan on a Cd/As contamined soil.

In this thesis a remarkable 12 inorganic elements were monitored in 281 rice accessions grown in AWD and PF. Despite the high extent of LD, which is known to decrease the resolution of GWA studies, the approach adopted has allowed the identification of 71 MTAs AWD-specific, 124 PF-specific and 10 associations which were in common between the two water systems.

Among the 71 AWD-specific associations, 43 were referred to micro elements, 9 to macroelements, and 19 for the undesired elements analyzed (Cd and As).

Among the 124 PF-specific associations, 86 were referred to microelements, 21 to macroelements and 17 to the toxic elements. Among the 10 non-specific associations, 3 were for microelements, 1 for macro elements, 3 for trace elements and 3 for toxic elements.

In particular, *SDG714* gene, identified in association with Ni (AWD and PF), was detected in chromosome 1 in a region ranging between 42,420,839 bp and 42,423,027 bp, identified by the SNP S1_42412011 (Garg et. al 2015). *OsIRX10* gene, is flanking to *SDG714* (Zeng et al., 2010).

The locus *Os02g0141000*, identified in association with arsenic in AWD, has been detected in chromosome 2, in a region ranging from 2,189,328 bp to 2,734,203 bp, identified by SNPs S2_2699366.

The *OsGDH2* and *OsCAT5* genes, detected in association with P (the first), and with P and Mg (the second), have in common SNPs S4_27790754 and S4_27790806, which are in a region of chromosome 4, ranging from 27,707,649 bp to 27,825,849 bp. (Obara et al. 2011)

Along with chromosome 4, about 3,000 kb, the content of Ca, the OsAWDA1 gene, identified by SNPs S4_31050939 and S4_31050953, was found in a region ranging from 31,050,899 bp to 31,225,422 bp (Redillas 2012).

On chromosome 7, the *OsEnS_107* gene, found associated with the content of Arsenic in AWD, was identified by SNP S7_21235989, contained in a chromosomal region ranging from 21,231,736 bp to 21,412,154 (Chakrabarty et al., 2009).

Finally, on chromosome 8, the *OsCCC1* genes (Matsuda et al. 2012; Chen at al., 2015), associated with Cd content, was found. Identified by SNPs S8_4096930 and S8_4099221, in a region ranging from 3,797,258 bp to 4,101,097 bp, while the latter was identified by SNPs S8_14616713, S8_14750851 and S8_14976342, in a region ranging from 14,051,981 bp to 14,997,761 bp.

# 6 CONCLUSIONS

The results obtained in the present work indicate that most of the 281 temperate and tropical rice accessions of our panel are penalized for several agronomically relevant traits when grown under AWD. However, performances of several accessions in the AWD were similar to those obtained under PF, suggesting that genetic variability for adaptation to aerobic conditions in our *japonica* rice panel exist.

Several traits provide a contribution to the yield values recorded for the accessions ranking at the highest position in the AWD system, suggesting that useful alleles for several traits conferring adaptation to the water-save technology as well as for improving the nutritional value and safety of the grains could be identified in this panel.

The GWAS analysis provided a large number of significant associations and among them 32 were AWD specific, whereas 69 were in common between AWD and PF. The robustness of several of these effects was assessed through the identification of genes whose allelic variation represents robust candidates for the phenotypic effect and through verification of positional relationships with QTLs previously identified as involved in rice adaptation to drought stress or reduced water availability.

There is no an unique trend for all elements in response to AWD. While for some elements (e.g. As, Ni and Cd) the effect of AWD in comparison to PF resulted to be very strong, for others (e.g. Na) the watering techniques do not differently influence their brown grain concentration. However, the genotype component of variability is marked. GWAS allowed the identification of 71 MTAs AWD-specific, 124 PF-specific and 10 associations which were in common between the two water systems-

Some interesting genes which allelic differences could explain the variability observed in accumulation of inorganic nutrients into the rice brown grain have been identified.

The GWAS panel used in this study has recently successfully been used in Genomic Selection study for AWD tolerance (Ben Hassen et al., 2017). These approaches involve cross validation and genomic prediction across generations; it is expected that accumulation of alleles here identified with better phenotypic values under AWD or both cumulated (AWD and PF) conditions, through Genomic Selection approaches, should allow selection of *japonica* rice lines with improved adaptation to aerobic conditions.

This work could be the starting point for Genomic Selection for ions, or for the development of molecular markers.

# REFERENCES


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

Alloway, BJ., Steinnes, E., (1999). Anthropogenic additions of cadmium to soils, in: McLaughlin, MJ., Singh , BR., (Eds.). *Cadmium in Soils and Plants, Kluwer Academic Publishers, Boston*, pp. 97–124.

Angelini, R., Ferrero, A., Tinarelli, A. (2008). Il riso. Coltura e Cultura. *Bayer CropScience*.

Arao, TKA. (2009). Effects of water management on cadmium and arsenic accumulation and dimethylarsinic acid concentrations in Japanese rice. *Environmental Science Technol*, 43:9361–9367.

Arao, T., Maejima, Y., Baba, K. (2009). Uptake of aromatic arsenicals from soil contaminated with diphenylarsinic acid by rice. *Environmental Science Technology*, 43(4): 1097-1101.

Atlin, G., Lafitte, H., Tao, D., Laza, M., Amante, M., Courtois, B. (2006). Developing rice cultivars for high-fertility upland systems in the Asian tropics. *Field* Crops Research, 97:43–52.

Aulakh, MS., Khera, TS., Doran, JW., and Bronson, KF. (2001). Denitrification, $N_2O$ and $CO_2$ fluxes in rice–wheat cropping system as affected by crop residues, fertilizer N, and legume green manure. *Biology and Fertility of Soils,* 34:375–389.

Barrett, J., Fry, B., Maller, J., Daly, M. (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, 21(2): 263-265.

Ben Hassen, M., Cao, T., Bartholomé, J., Orasen, G., Colombi, C., Rakotomalala, J., Ahmadi, N. (2017). Rice diversity panel provides accurate genomic predictions for complex traits in the progenies of biparental crosses involving members of the panel . *Theoretical and Applied Genetics*, 1-19.

Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. , Series B (Methodological),* 289-300.

Bing, Y., Wei, Y., Li-Jun, L., Yong-Zhong, X. (2006). QTL analysis for flag leaf characteristics and their relationships with yield and yield traits in rice. *Acta Genetica Sinica*, 33(9), 824–832.

Bingham, F., Page, A., MR., Ganje, T. (1976). Cadmium availability to rice in sludge-amended soil under "flood" and "nonflood" culture. *Soil Science Society of America Journal*, 40(5): 715-719.

Biscarini, F., Cozzi, P., Casella, L., Riccardi, P., Vattari, A., Orasen, G., Greco, R. (2016). Genome-Wide Association Study for traits related to plant and grain morphology, and root architecture in temperate rice accessions. *PLoS ONE* , 11(5): e0155425.

Bouman, BM., Tuong, TP. (2001). Field water management to save water and increase its productivity in irrigated rice. *Agricultural Water Management*, 49(1):11–30.

Bouman, BM., Lampayan, R. M., Tuong, T. P. (2007). Water Management in Irrigated Rice: Coping with Water Scarcity. ISBN 978-971-22-0219-3. *Los Baños, Philippines. Los Baños (Philippines): International Rice Research Institute.*

Bouman, BM. (2009). How much water does rice use. *Rice Today*, 8:28–29.

Brachi, B., Morris, G., Borevitz, J. (2011). Genome-wide association studies in plants: the missing heritability isin the field. *Genome biology*, 12(10):1–8.

Brown, KW., Turner, FT., Thomas, JC., Deuel, LE., Keener, ME. (1978). Water balance of flooded rice paddies. *Agricultural Water Management*, 1:277–291.

Bushart, T., Cannon, A., Haque, U. (2013). RNA-seq analysis identifies potential modulators of gravity response in spores of Ceratopteris (Parkeriaceae): evidence for modulation by calcium pumps and apyrase. *Journal of Plant Growth Regulation* , 21:137–145.

Cai, Y., Chen, X., Xie, K., Xing, Q., Wu, Y., Li, J., Guo, Z. (2014). *Dlf1, a WRKY* Transcription Factor, Is Involved in the Control of Flowering Time and Plant Height in Rice. *PLoS ONE*, 9(7): e102529.

Casella, L. (2011). SNP analysis for drought-related candidate genes in a germplasm collection and a tilling population of Italian rice. *Facoltà di Agraria.*

Cavigiolo, S., Greppi, D., Lupotto, E. (2007). Risparmio idrico in risaia con l'irrigazione turnata. CRA – Istituto sperimentale per la Cerealicoltura Sezione specializzata per la risicoltura di Vercelli. *L'informatore agrario*, n°21.

Catarecha, P., Segura, MD., Franco-Zorrilla, JM., García-Ponce, B., Lanza M., Solano, R., Paz-Ares J., Leyva, A. (2007). A mutant of the Arabidopsis phosphate transporter *PHT1*;1 displays enhanced arsenic accumulation*. Plant Cell 19:*1123–113310.1105.

Chakrabarty, D., Trivedi, P., Misra, P., Tiwari, M., Shri, M., Shukla, D., Tuli, R. (2009). Comparative transcriptome analysis of arsenate and arsenite stresses in rice seedlings, *Chemosphere*, 74:688–702.

Chaney, R. (2015). How does contamination of rice soils with Cd and Zn cause high incidence of human Cd disease in subsistence rice farmers. *Current Pollution Reports*, 1(1): 13-22.

Chen, M., Presting, G., et al., (2002). An integrated physical and genetic map of the rice genome. *The Plant Cell Online*, 14.3:537-545.

Chen, Z., Zhu,YG., Liu, WJ., Merhag, AA., (2005). Direct evidence showing the effect of root surface iron plaque on arsenite and arsenate uptake into rice (*Oryza sativa*) roots. *New Phytologist* Journal, 165 (1): 91–97.

Chen, G., Hu, Q., Luo, L., Yang, T., Zhang, S., et al. 2015. Rice potassium transporter *OsHAK1* is essential for maintaining potassium-mediated growth and functions in salt tolerance over low and high potassium concentration ranges. *Plant Cell Environ*, 38:2747–65

Chung, P., Kim, Y., Jeong, J., Park, S., Nahm, B., Kim, J. (2009). The histone deacetylase *OsHDAC1* epigenetically regulates the *OsNAC6* gene that controls seedling root growth in rice. Plant Journal 59(5): 764–776

Cleveland, W. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of American Statistical Association*, 74:829–836.

Colmer TD. (2003). Aerenchyma and an inducible barrier to radial oxygen loss facilitate root aeration in upland, paddy and deep-water rice (*Oryza sativa L.*). *Annals of Botany*, 91:301–309

Colmer TD, Cox MCH, Voesenek LACJ (2006). Root aeration in rice (Oryza sativa): evaluation of oxygen, carbon dioxide, and ethylene as possible regulators of root acclimatizations. *New Phytologist*, 170:767–778

Conaf. (2015). Riso, Italia prima in Europa per produzione. *Notiziario Conaf.*

Connolly, EL., Fett, JP., Guerinot, ML. (2002). Expression of *the IRT1* metal transporter is controlled by metals at the levels of transcript and protein accumulation. *Plant Cell* 14: 1347–1357

Courtois, B., Frouin, J., Greco, R., Bruschi, G., et al (2012). Genetic diversity and population structure in a European collection of rice. *Crop Science*, 52:1663–1675

Das, S., Chou, ML., Jean, JS., Liu, CC., Yang, HJ., (2016). Water management impacts on arsenic behavior and rhizosphere bacterial communities and activities in a rice agroecosystem. *Science of the Total Environmental*, 542 (A):642–652.

Devos, K. (2005). Updating the 'crop circle'. *Current opinion in Plant Biology*, 8: 155–162.

Duran, CN. (2008). AutoSNPdb: an annotated single nucleotide polymorphism database for crop plants. *Nucleic Acids Research*, 37: D951–953.

Earl, D., von Holdt, B. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetic Resources*, 4:359.

Elshire, R., Glaubitz, J., Sun, Q., Poland, J., Kawamoto, K., Buckler, E., et al. (2011). A robust, simple A robust, simple genotyping by-sequencing (GBS) approach for high diversity species. *PloS one*, 6(5):e19379.

Evenson, RE., Gollin, D., (1997). Genetic Resources, International Organizations, and Improvement in Rice Varieties. *Chicago Journals*, 45:471-500.

Excoffier, L., Lischer, H. (2010). Arlequin suite version 3.5: a new series of programs to perform population genetics analyses under linux and windows. *Molecular Ecology Resources* , 10:564–567.

Falcone Ferreyra, M. R. (2012). Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Frontiers in Plant Science*, 3:222.

Famoso, A., Zhao, K., Clark, R., Tung, C., Wright, M., Bustamante, C., McCouch, S. (2011). Genetic architecture of aluminum tolerance in rice (*Oryza sativa*) determined through genome-wide association analysis and QTL mapping. *PLoS Genetics*, 7(8): e1002221.

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

FAO Rice Market Monitor (2016) Rice Market Monitor. Trade and Markets Division. Food and Agriculture Organization of the United Nations.

FAO Trade and Market Division (2016) Food Outlook October 2016. *Food Outlook.*

Farooq, M., Wahid, A., Kobayashi, N., Fujita, D., Basra, SMA. (2009). Plant drought stress: effects, mechanisms and management. *Agronomy for Sustainable Development*, 29: 185-212.

Fontanili, L., Lancilli, C., Suzui, N., Dendena, B., Yin, Y., F. A., and Sacchi, G. (2016). Kinetic Analysis of Zinc/Cadmium Reciprocal Competitions Suggests a Possible Zn-Insensitive Pathway for Root-to-Shoot Cadmium Translocation in Rice. *Rice*, 9(1): 16.

Fujita, D., Trijatmiko, K., Tagle, A., Sapasap, M., Koide, Y., Sasaki, K., Kobayashia, N. (2013). *NAL1* allele from a rice landrace greatly increases yield in modern indica cultivars. . *Proceedings of the Natlional Academy of Sciences U.S.A.* , 110(51):20431–20436.

Gabriel, S., Schaffner, S., Nguyen, H., Moore, J., Roy, J., Blumenstiel, B., Liu-Cordero, S. (2002). The structure of haplotype blocks in the human genome. *Science*, 296(5576): 2225-2229.

Garg. R., Chevala, VVSN.., Shankar, R., Jain M. (2015). Divergent DNA methylation patterns associated with abiotic stress responses and regulation of gene expression in rice. *Scientific Reports*: 5, 14922

Garris, AJ., Tai, TH., Coburn, J., Kresovich, S., McCouch, S., (2004). Genetic Structure and Diversity *in Oryza sativa* L. *Genetics Society of America*, 1631-1638.

Goulas, Y., Cerovic, Z., Cartelat, A., Moya, I. (2004). Dualex: a new instrument for field measurements of epidermal ultraviolet absorbance by chlorophyll fluorescence. *OSA Publishing*, 448-4496.

Gower, J. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, 325-338.

Grant, C., Clarke, J., Duguid, S., Chaney, R. (2008). Selection and breeding of plant cultivars to minimize cadmium accumulation. *Science of the Total Environment*, 390(2): 301-310.

Global Rice Science Partnership (GRiSP). 2013. Rice almanac, 4th ed. International Rice Research Institute, Los Banos, *Philippines*, 283 p.

Graham, R. et al. (1999) Breeding for micronutrient density in edible portions of staple food crops: conventional approaches. *Field Crops Research,* 60:57–80.

Graham, R.D. et al. (2001). Addressing micronutrient malnutrition through enhancing the nutritional quality of staple foods: principles, perspectives and knowledge gaps. *Advances in Agronomy*, 70:77–142.

Guerinot, M.L., Yi, Y. (1994). Iron: Nutritious, Noxious, and Not Readily Available. *Plant Physiology*, 104:815-820.

Herrmann, K. (1995). The Shikimate Pathway: Early Steps in the Biosynthesis of Aromatic Compounds. *Plant Cell*, 7:907-919.

Hirschhorn, J., Daly, M. (2005). Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics*, 6(2):95–108.

Hu, P., Li, Z., Yuan, C., Ouyang, Y., Zhou, L., Huang, J., Wu, L. (2013). Effect of water management on cadmium and arsenic accumulation by rice (*Oryza sativa* L.) with different metal accumulation capacities. *Journal of Soils and Sediments*, 13(5): 916-924.

Hu, Y., Cheng, H., Tao, S. (2016). The challenges and solutions for cadmium-contaminated rice in china: A critical review. *Environment International*, 92:515-532.

Huang, X., Zhao, Y., Li, C., Wang, A., et al., (2012). Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nature Genetics*, 44:32-39.

Huang, X., Han, B. (2014). Natural variations and genome-wide association studies in crop plants. *Annual Review of Plant Biology*, 65: 531-551.

Huang, X., Salt, D. (2016). Plant ionomics: From elemental profiling to environmental adaptation. *Molecular Plant*, 9(6): 787-797.

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

IRRI (International Rice Research Institute). 1984. Upland rice in Asia. Pages 45-68 in An overview of upland rice research. Proceedings of the 1982 Bouake, Ivory Coast, upland rice workshop. International Rice Research Institute. *Los Baños, Philippine*

Ishimaru, Y., Takahashi, R., Bashir, K., Shimo, H., Senoura, T., Sugimoto, K., Nakanishi, H. (2012). Characterizing the role of rice NRAMP5 in manganese, iron and cadmium transport. . *Scientific Reports*, 2: 286.

Itoh, H., Tatsumi, T., Sakamoto, T., Otomo, K., Toyomasu, T., Kitano, H., Matsuoka, M. (2004). A rice semi-dwarf gene, *Tan-Ginbozu (D35)*, encodes the gibberellin biosynthesis enzyme, ent-kaurene oxidase. *Plant Molecular Biology*, 54(4): 533–547.

Jiang, D., Jingjing, F., Lamei, L., Jinfeng, Z., Shoujiang, Y., Liang, Y., Xueyong, L. (2015). Characterization of a null allelic mutant of the rice nal1 gene reveals its role inregulating cell division. *PLoS ONE* , 10(2):e0118169.

Jombart, T., Collins, C. (2015). A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0.0*. London: London: Imperial College London, MRC Centre for Outbreak Analysis and Modelling.*

Kato, Y., Okami, M. (2011). Root morphology, hydraulic conductivity and plant water relations of high-yielding rice grown under aerobic conditions. *Annals of Botany* 108, 575–583. doi:10.1093/ aob/mcr184

Kaur S, P. P. (2013). Simple Sequence Repeat Markers in Genetic Divergence and Marker Assisted Selection of Rice Cultivars. *Critical Reviews in Food Science and Nutrition.*

Kawahara, Y., de la Bastide, M., Hamilton, J., Kanamori, H., McCombie, W., Ouyang, S., al, e. (2013). Improvement of the *Oryza sativa Nipponbare* reference genome using next generation sequence and optical map data. *Rice*, 6:1-10.

Khowaja, F., Norton, G., Courtois, B., Price, A. (2009). Improved resolution in the position of drought-related QTLs in a single mapping population of rice by meta-analysis. . *BMC Genomics* , 10(1):2.

Khush, G. (1997). Origin, dispersal, cultivation and variation of rice. Plant Mol. Biol. , 35:25–34.

Kong, XB., (2014). China must protect high-quality arable land. *Nature* 506 (7).

Korte, A., Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review*. Plant methods*, 9(1):29.

Kulichikhin K., Yamauchi T.,Watanabe K. Nakazono M. (2014) Biochemical and molecular characterization of rice (*Oryza sativa L*.) roots forming a barrier to radial oxygen loss. *Plant, Cell Environm*ent 37:2406–2420.

Kumar, A., Dixit, S., Ram, T., Yadaw, R., Mishra, K., Mandal, N. (2014). Breeding high yielding drought-tolerant rice: genetic variations and conventional and molecular approaches. . *Journal of Experimental Botany*, 65(21):6265-6278.

Kumar, V., Singh, A., Mithra, S., Krishnamurthy, S., Parida, S., Jain, S., Mohapatra, T. (2015). Genome-wide association mapping of salinity tolerance in rice (*Oryza sativa*). *DNA Research*, 22(2):133–145 .

Lafitte, H. R., Courtois, B., Arraudeau, M. (2002). Genetic improvement of rice in aerobic systems: progress from yield to genes. *Field Crops Research*, 75:171–190.

Lampayan, RM., Palis, FG., Flor, RB., Bouman, BM., Quicho, ED., de Dios, JL., Espiritu, A., Sibayan, EB., Vicmudo, VR., Lactaoen, AT., Soriano, JB. (2009). Adoption and dissemination of “safe alternate

wetting and drying" in pump irrigated rice areas in the Philippines. In: Proceedings of the 60th International Executive Council Meeting and 5th Asian Regional Conference, 6-11 December 2009, New Delhi, India.

Lampayan , RM., Rejesus, RM., Singleton. GR., Bouman,BM. (2015) . Adoption and economics of alternate wetting and drying water management for irrigated lowland rice . *Field Crop Research*, 170 : 95 – 108 .

Li, C., Salas, W., DeAngelo, B., Rose, S., (2006). Assessing alternatives for mitigatingnet greenhouse gas emissions and increasing yields from rice production in China over the next twenty years. *Journal of Environmental Quality*, 35:1554–1565.

Li, Y., Fan, C., Xing, Y. J., Luo, L., Sun, L., Shao, D., Zhang, Q. (2011). Natural variation in *GS5* plays an important role in regulating grain size and yield in rice. *Nature genetics*, 43:1266–1269.

Linquist, B., Anders, M., Adviento-Borbe, M., Chaney, R., Nalley, L., daRosa, E., van Kessel, C. (2015). Reducing greenhouse gas emissions water use,and grain arsenic levels in rice systems. *Global Change Biology*, 21:407–417.

Linquist, B., Snyder, R., Anderson, F., 2015. Water balances and evapotranspirationin water- and dry-seeded rice systems. *Irrigation Science*, 33:375–385.

Liu, Y., Xu, Y., Xiao, J., Ma, Q., Li, D., Xue, Z., Chong, K. (2011). *OsDOG*, a gibberellin-induced A20/AN1 zinc-finger protein, negatively regulates gibberellin-mediated cell elongation in rice. *Journal of Plant Physiology*, 168(10):1098-1105.

Lu, Z., Shao, G., Xiong, J., Jiao, Y., Wang, J., Liu, G., Li, J. (2015). *MONOCULM 3*, an Ortholog of *WUSCHEL* in Rice, Is Required for Tiller Bud Formation. *Journal of Genetics and Genomics* , 42(2):71-78.

Ma, JF., Tamai, K., Yamaji, N., Mitani, N., Konishi, S., Katsuhara, M., Ishiguro, M., Murata, Y., Yano, M., (2006). A Si transporter in rice. *Nature* 440:688e691.

Ma, JF., et al., (2008). Transporters of arsenite in rice and their role in arsenic accumulation in rice grain, *Proceedings of the National Academy of Science U.S.A.*, 105:9931–9935.

Mackill, D., Coffman, W., Garrity, D. (1996). Rainfed Lowland Rice Improvement. *International Rice Research Institute, Los Banos, Philippines.*

Maclean, JL., Dawe, DC., Hardy, B., Hettel, GP., (2002). Rice Almanac: Source Book for the Most Important *Economic Activity on Earth, CABI Publishing*, 2002.

Mangin, B., Siberchicot, A., Nicolas, S., Doligez, A., This, P., Cierco-Ayrolles, C. (2012). Novel measures of linkage disequilibrium that correct the bias due to population structure and relatedness. *Heredity* , 108, 285–91.

Marschner, H., Römheld, V. (1994). Strategies of plants for acquisition of iron. *Plant Soil*, 165: 375–388.

Marroni, F., Pinosio, S., Zaina, G., Fogolari, F., Felice, N., Cattonaro, F., Morgante, M. (2011). Nucleotide diversity and linkage disequilibrium in Populus nigra cinnamyl alcohol dehydrogenase (*CAD4*) gene . *Tree Genetics & Genomes* , 7(5):1011–1023.

Matsuda, S., Funabiki, A., Furukawa, K., Komori, N., Koike, M., Tokuji, Y., Kato, K. (2012). Genome-wide analysis and expression profiling of half-size ABC protein subgroup G in rice in response to abiotic stress and phytohormone treatments. *Molecular Genetics and Genomics*, 10:819–835.

Matsushima, S. (1976). High-Yielding Rice Cultivation. *Japan Scientific Societies Press, Tokyo*, 367 p.

Mather, KA., Caicedo, AL., Polato, NR., Olsen, KM., McCouch, S., et al. (2007) The extent of linkage disequilibrium in rice (*Oryza sativa* L.). *Genetics*, 177: 2223–2232.

Matsuo, N., Ozawa, K., Mochizuki, T. (2010). Physiological and morphological traits related to water use by three rice (*Oryza sativa L*.) genotypes grown under aerobic rice systems. *Plant Soil*, 335:349–361.

Matusmoto, T., Yamada, K., Yoshizawa, Y., Oh, K. (2016). Comparison of Effect of Brassinosteroid and Gibberellin Biosynthesis Inhibitors on Growth of Rice Seedlings. *Rice Sci*ence, 23(1), 51–5.

Mackill DJ., Coffman WR., Garrity DP. (1996). Rainfed lowland rice improvement. International Rice Research *Institute, P.O. Box 933, Manila, Philippines*, 242 p.

McCauley, GN. (1990). Sprinkler vs. flood irrigation in traditional rice production regions of southeast Texas. *Agronomy Journal,* 82:677–683.

McCouch, S., Wright, M., Tung, C.-W., Maron, L., McNally, K., Fitzgerald, M., et al. (2016). Open access resources for genome-wide association mapping in rice. *Nature communications*.

Meharg, A., Williams, PN., Adomako, E., Lawgali, YY., Deacon, C., Villada, A., et al. (2009). Geographical variation in total and inorganic arsenic content of polished (white) rice. *Environmental Science Technology,* 43(5):1612–7.

Meharg, A., Norton, G., Deacon, C., Williams, P., Adomako, E., Price, A., Villada, A. (2013). Variation in rice cadmium related to human exposure. *Environmental Science Technology*, 47(11): 5613-5618.

Monna, L., Kitazawa, N., Yoshino, R., Suzuki, J., Masuda, H., Maehara, Y., Minobe, Y. (2012). Positional cloning of rice semidwarfing gene, *sd-1*: rice "green revolution gene" encodes a mutant enzyme involved in gibberellin synthesis. *DNA Research*, 9(1):11- 17.

Moos, SP., Mumm, RH., (2008). Molecular plant breeding as the foundation for 21st century crop improvement*. Plant physiolo*gy, 147.3:969-977.

Ministero per le Politiche Agricole - DM 1999.(1999). Metodi ufficiali di analisi chimica del suolo*. Gazzetta Ufficiale* n. 248, 21.10.1999.

Müller, C., Silveira, SF., Menezes Daloso D., Camargo Mendes, G., Merchant, A., Kuki, KN., Oliva, MA., Loureiro, ME., Almeida, AM. (2017). Ecophysiological responses to excess iron in lowland and upland rice cultivars. *Chemosphere,* 189:123-133.

Murgia, I., Arosio, P., Tarantino, D., Soave, C. (2012). Biofortification for combating 'hidden hunger' for iron. *Trends Plant Science*, 17:47–55

Muthayya, S., Sugimoto, J., Montgomery, S., GF, M. (2014). An overview of global rice production, supply, trade, and consumption. *Annals of the New York Academy of Sciences*, 1324:7e14.

Nishimura, AL. et al. (2004). A mutation in the vesicletrafficking protein VAPB causes late-onset spinal muscular atrophy and amyotrophic lateral sclerosis. *Archieve of "American Journal of Human Genetics"*, 75:822–831.

Nocito, F., Lancilli, C., Dendena, B., Lucchini, G., Sacchi, G. (2011). Cadmium retention in rice roots is influenced by cadmium availability, chelation and translocation. *Plant, Cell Environment*, 34(6): 994-1008.

Norton, G., Deacon, C., Xiong, L., Huang, S., Meharg, A., Price, A. (2010). Genetic mapping of the rice ionome in leaves and grain: identification of QTLs for 17 elements including arsenic, cadmium, iron and selenium. *Plant Soil*, 329:139–153.

Norton, G., Douglas, A., Lahner, B., Yakubova, E., Guerinot, M., Pinson, S., Islam, M. (2014). Genome wide association mapping of grain arsenic, copper, molybdenum and zinc in rice (Oryza sativa L.) grown at four international field sites. *PLoS One*, 9(2): p.e89685.

Norton, GJ., Shafaei, M.,Travis, AJ., Deacon, CM.,Danku, J.,Pond, D.,et al. 2017. Impact of alternate wetting and drying on rice physiology, grain production, and grain quality. *Field Crop Research*, 205:1–13.

Nyquist, H. (1991). Restricted Estimation of Generalized Linear Models. *Journal of the Royal Statistical Society,* 4(C):133-141.

Obara, M., Takeda, T., Hayakawa, T., Yamaya, T. (2011). Mapping quantitative trait loci controlling root length in rice seedlings grown with low or sufficient supply using backcross recombinant lines derived from a cross between Oryza sativa L. and Oryza glaberrima Steud. *Soil Science and Plant Nutrition* , 57:80–92.

Oono, Y., Yazawa, T., Kawahara, Y., Kanamori, H., Kobayashi, F., Sasaki, H., Itoh, T.(2014). Genome-wide transcriptome analysis reveals that cadmium stress signaling controls the expression of genes in drought stress signal pathways in rice. *PLoS One*, 9(5): p.e96946.

Orasen, G., Baldoni, E., Abruzzese, A., Pesenti, M., Maghrebi, M., Volante, A., Nocito, F.F., Dell'Orto, M., De Nisi, P., Valè, G., Sacchi, G.A. (2017). Genome-Wide Association Study for Mid-Salt Tolerance in Rice. *ITRC Annual Meeting-2017*

Parengam, M., Judprasong, K., Srianujata, S., Jittinandana, S., Laoharo¬janaphand, S., Busamongko, A. (2010). Study of nutrients and toxic minerals in rice and legumes by instrumental neutron activation analysis and graphite furnace atomic absorption spectropho¬tomerty. *Journal of Food Composition and Analysis* , 23:340–345.

Pereira, E.G., Oliva, M.A., Siqueira-Silva, A.I., Rosado-Souza, L., Pinheiro, D.T., Miyasaka, A., (2014). Tropical rice cultivars from lowland and upland cropping systems differ in iron plaque formation. *Journal of Plant Nutrition,* 37, 1373e1394.

Ponnamperuma, FN. (1972). The chemistry of submerged soil. *Advances in Agronomy*, 24:29-96.

Pinson, S., Tarpley, L., Yan, W., Yeater, K., Lahner, B., Yakubova, E., Salt, D. (2015). Worldwide genetic diversity for mineral element concentrations in rice grain. *Crop Science*, 55(1): 294-311.

Price, AH., Norton, GJ., Salt, DE., Ebenhoeh, O., Meharg, AA., Rafiqul Islam, M., Davies, WJ. (2013). Alternate wetting and drying irrigation for rice in Bangladesh: Is it sustainable and has plant breeding something to offer? *Food and Energy Security,* 2:120-129.

Price, A., Norton, G., Salt, D., Ebenhoeh, O., Meharg, A., Meharg, C., Davies, W. (2013). Alternate wetting and drying irrigation for rice in Bangladesh: Is it sustainable and has plant breeding something to offer? *Food and Energy Security*, 2:120-129.

Pritchard, J., Stephens, M., Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155:945–959.

Qi, J., Qian, Q., Bu, Q., Li, S., Chen, Q., Sun, J., Li, C. (2008). Mutation of the rice Narrow *leaf1* gene which encodes a novel protein, affects vein patterning and polar auxin transport. *Plant Physiology*, 147:1947-1959.

Rao, Y., Li, YQ. (2014). Recent progress on molecular breeding of rice in China. *Plant Cell Reports*, 33:551-564.

Redillas, M., Jeong, J., Kim, Y. J., Bang, S., Choi, Y., Ha, S., Kim, J. (2012). The overexpression of OsNAC9 alters the root architecture of rice plants enhancing drought resistance and grain yield under field conditions. *Plant Biotecnology Journal*.

Richards, R., Rebetzke, G., Watt, M., Condon, A., Spielmeye, W., Dolferus, R. (2010).Breeding for improved water productivity in temperate cereals: phenotyping, quantitative trait loci, markers and the selection environment. *Functional Plant Biology*, 37:85–97.

Ringler, C., Zhu, T. (2015). Water resources and food security. *Journal of Agronomy*, 106: 1–6.

Rinklebe, J.S., Shaheen (2017). Redox chemistry of nickel in soils and sediments: A review. *Chemosphere*, 179, 265-278.

Roberts, LC., Hug, SJ., Dittmar, J., Voegelin, A., Kretzschmar, R., Wehrli, B., Cirpka, OA., Saha, GC., Ali, MA., Badruzzaman, ABM. (2010). Arsenic release from paddy soils during monsoon flooding. *Nature Geoscience*, 3, 53–59.

Sahrawat, KL., Wani, S.P, Rego, TJ., Pardhasaradhi, G., Murthy, KVS. (2007). Widespread deficiencies of sulphur, boron and zinc in dryland soils of the Indian semiarid tropics. *Current Science* 93(10):1-6.

Sandhu, N., Raman, KA., Torres, R.O., Audebert, A., Dardou, A., Kumar, A. (2016). Rice Root Architectural Plasticity Traits and Genetic Regions for Adaptability to Variable Cultivation andStress Conditions. *American Society of Plant Biologists*,177: 2562-2576.

Sasaki, A., Yamaji, N., Yokosho, K., Ma, J. (2012). Nramp5 is a major transporter responsible for manganese and cadmium uptake in rice. . *The Plant Cell*, 24(5): 2155-2167.

Sass, R.L., Fisher, FM., (1997). Methane emissions from rice paddies: a process study summary. Nutrient *Cycling Agroecosystems*, 49:119–127.

Schmitz, A., Begcy, K., Sarath, G., Walia, H. (2015). Rice Ovate Family Protein 2 (*OFP2*) alters hormonal homeostasis and vasculature development. *Plant Science*, 241:177-188.

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

Serraj, R., Kumar, A., McNally, K.L., Slamet-Loedin, I., Bruskiewich, R., Mauleon, R., Cairns, J., Hijmans, RJ., (2009). Improvement of drought resistance in rice. *Advances Agronomy*, 103: 41–98.

Shaibu, AS., Adnan, AA., Rabiu, IU. (2015). Genetic correlation and contribution of some physiological traits to yield in some selected maize genotypes*. International Journal of Scientific & Technology Research*, 5(4):1–10.

Shomura, A., Izawa, T., Ebana, K., Ebitani, T., Kanegae, H., Konishi, S., Yano, M. (2008). Deletion in a gene associated with grain size increased yields during rice domestication. *Natural Genetics*, 40:1023–1028.

Slamet-Loedin, I., Johnson-Beebout, S., Impa, S., Tsakirpaloglou, N. (2015). Enriching rice with Zn and Fe while minimizing Cd risk. *Frontiers in Plant Science*, 6: 121.

Siopongco, JDLC, Wassmann, R, Sander, BO. (2013). Alternate wetting and drying in Philippine rice production: feasibility study for a clean development mechanism. *IRRI Technical Bulletin No. 17*. *Los Baños (Philippines): International Rice Research Institute*, 14 p.

Soil Science Society of America (2008). Soil Science Glossary Terms Committee. *Glossary of soil science terms*. USA: 1223.

Sonah, H., Bastien, M., Iquira, E., Tardivel, A., Légaré, G., et al., (2013). An improved genotyping by sequencing (GBS) approach offering increased versatility and efficiency of SNP discovery and genotyping. *PLoS ONE,* 8: e54603

Spindel, J., Begum, H., Akdemir, D., Virk, P., Collard, B., Redoña, E., McCouch, S. (2015). Genomic Selection and Association Mapping in rice (Oryza sativa): Effect of trait genetic architecture, training population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines. *Plos genetics*.

Shiono K., Yamauchi T., Yamazaki S., Mohanty B., Malik A.I., Nagamura Y.,. . . Nakazono M. (2014). Microarray analysis of laser-microdissected tissues indicates the biosynthesis of suberin in the outer part of roots

during formation of a barrier to radial oxygen loss in rice (Oryza sativa). *Journal of Experimental Botany*.

Song, WY., Yamaki, T., Yamaji, N., Ko, D., Jung, KH., et al. (2014). A rice ABC transporter, *OsABCC1*, reduces arsenic accumulation in the grain. PNAS 111:15699–704

Sudhir-Yadav, E., Humphreys, T. Li, G. Gill, SS. Kukal. (2012). Evaluation of tradeoffs in land and water productivity of dry seeded rice as affected by irrigation schedule. *Field Crops Research*, 128:180–190.

Sun, C., Chen, D., Fang, J., Wang, P., Deng, X., Chu, C. (2014). Understanding the genetic and epigeneticarchitecture in complex network of rice flowering pathways. *Protein Cell* , 5(12):869-898.

Sunohara, H., Satoh, H., Nagato, Y. (2003). Mutation in panicle development affect culm elongation in rice. *Breeding Science*, 59:109-117.

Takahashi, Y., Minamikawa, R., Hattori, K. H., Kurishima, K., Kihou, N., Yuita, K. (2004). Arsenic behaviour in paddy fields during the cycle of flooded and non-flooded periods. *Environmental Science and Technology*, 38:1038-1044.

Tan, X., Shao, D., Liu, H., Yang, F., Xiao, C., Yang, H., (2013). Effects of alternatewetting and drying irrigation on percolation and nitrogen leaching in paddyfields. *Paddy Water Environmental*, 11:381–395.

Kobayashi, T., Nishizawa, N.K. (2012). Iron Uptake, Translocation, and Regulation in Higher Plants. *Annual Review of Plant Biology*, 63:131-152

Tòt, G., Hermann, T., Da Silva, M.R., Montanarella, L. (2016). Heavy metals in agricultural soils of the European Union with implications for food safety. *Environmental International*, 88:299-309.

Tremblay, N., Wang, Z., Cerovic, Z.G. (2012). Sensing crop nitrogen status with fluorescence indicators. A review. *Agronomy for a Sustainable Development*, 32:451–464

Tuong, TP., Bouman, BA. (2003). Rice Production in Water-scarce environments. In Water Productivity in Agriculture: Limits and Opportunities for Improvement (Kijne JW, Barker R and D. Molden editors). *CAB International*, 53-67.

Tuong, TP., Bouman, B. A. M., and Mortimer, M. (2005). More rice, less water‐integrated approaches for increasing water productivity in irrigated rice‐based systems in Asia. *Plant Production Science*, 8: 231–241.

Turner, S. (2014). qqman: an R package for visualizing GWAS results using QQ and manhattan plots. *BioRxiv*, 005165.

Vallino, M., Fiorilli, V., Bonfante, P. (2014) Rice Flooding Negatively Impacts Root Branching and Arbuscular Mycorrhizal Colonization, but Not Fungal Viability. *Plant Cell and Environment,* 37, 557-572.

Vaughan, DA., Lu, BR., Tomooka, N., (2008). The evolving story of rice evolution. *Plant science*, 174(4):394-408.

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

Volante, A., Desiderio, F., Tondelli, A., Perrini, R., Orasen, G., Biselli, C., Valè, G., Riccardi, P., Vattari, A., Cavalluzzo, D., Urso, S., Ben HAssen, M., Fricano, A., Piffanell, P., Cozzi, P., Bisccarini, F., Sacchi, GA., Cattivelli, L., Valè, G. (2017). Genome-wide analysis of japonica rice performance under Limited Water and Permanent Flooding conditions. *Frontiers in Plant Science*, 8:1862.

Wade, L., George, T., Ladha, J., Singh, U., Bhuiyan, S., Pandy, S. (1998). Opportunities to manipulate nutrient-by-water interactions in rainfed lowland rice systems. *Field Crop Res*, 56:93–112.

Wassmann, R., Jagadish, SVK., Sumfleth, K., Pathak, H., Howell, G., Ismail, A., et al. (2009). Regional vulnerability of climate change impacts on Asian rice production and scope for adaptation. *Advances in Agronomy*. 102:91–133.

Wassmann, R., Nelson, G.C., Peng, S.B., Sumfleth, K., Jagadish, S.V.K., Hosen, Y.,Rosegrant, M.W., (2010). Rice and global climate change. In: Pandley, S., Byerlee,D., Dawe, D., Dobermann, A., Mohanty, S., Rozelle, S., Hardy, B. (Eds.), Rice inthe Global Economy: Strategic Research and Policy Issues for Food Security.*International Rice Research Institute Los Ba˜nos, Philippines*, pp. 411–432.

Weir, B., Cockerham, C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* , 38: 1358–1370.

Weng, J., Gu, S., Wan, X., Gao, H., Guo, T., Su, N., Wan, J. (2008). Isolation and initial characterization of *GW5*, a majorQTL associated with rice grain width and weight. *Cell Research*, 18(12):1199–1209.

White, PJ., Broadley, MR. (2005). Biofortifying crops with essential mineral elements. *Trends Plant Science*, 10:586–593.

White, PJ., George, TS., Gregory, PJ., Bengough, AG., Hallett, PD. et al. (2013) Matching roots to their environment. *Annals of Botany* 112, 207-222.

Wickham, TH. and Singh, VP. (1978) Water movement through wet soils. In: Soils and Rice. I*RRI, Los Baños, Philippines,* pp. 337–358.

Winterbour, CC. (1995) Toxicity of iron and hydrogen peroxide: the Fenton reaction. *Toxicology Letters,* 83:969–974

Xu, X., McGrath, S., Meharg, A., Zhao, F. (2008). Growing rice aerobicallymarkedly decreases arsenic accumulation. *Environmental Science Technology*, 42,5574–5579.

Xu, Y., Crouch, J. (2008). Marker-assisted selection in plant breeding: from publications to practice. *Crop Science*, 48(2):391–407.

Xu, X., Liu, X., Ge, S., Jensen, J., Hu, F., Li, X., Wang, W. (2011).Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nature Biotechnology*, 30: 105-111.

Yamagata, N., Shigematsu, I. (1970). Cadmium pollution in perspective. *The Bulletin of the Institute of Public Healt, Tokyo*, 19(1): 1-27.

Yao, FX., JL. Huang, KH., Cui, LX, Nie, J., Xiang, XJ, et al. 2012. Agronomic performance of high-yielding rice variety grown under alternate wetting and drying irrigation. *Field Crops Resea*rch, 126:16–22.

Yoneyama, T., Ishikawa, S., Fujimaki, S. (2015). Route and regulation of zinc, cadmium, and iron transport in rice plants (*Oryza sativa* L.) during vegetative growth and grain filling: metal transporters, metal speciation, grain Cd reduction and Zn and Fe biofortification. *Molecular Sciences*, 16(8):19111-19129.

Yu, K., Chen, G., Patrick, WH. Jr (2004) Redox window with minimum GWP contribution from rice soils. *Soil Science Society of America Journal*, 68:2086–2091.

Zeng, W., Jiang, N., Nadella, R., Killen, T., Nadella, V., Faik, A. (2010). A Glucurono(arabino)xylan Synthase Complex from Wheat Contains Members of the *GT43, GT47*, and *GT75* Families and Functions Cooperatively. *Plant Physiology*, 110.159749.

Zhang, H., Y. Xue, Z. Wang, J. Yang, Zhang, J. (2009). Alternate wetting and moderate soil drying improves root and shoot growth in rice. *Crop Science*, 49:2246–2260.

Zhang, M., Pinson, S., Tarpley, L., Huang, X., Lahner, B., Yakubova, E., Salt, D. (2014). Mapping and validation of quantitative trait loci associated with concentrations of 16 elements in unmilled rice grain. *Theoretical and Applied Genetics*, 127:137–165.

Zhang, Z., Ober, U., Erbe, M., Zhang, H., Gao, N., J, H., et al. (2014). Improving the accuracy of whole genome prediction for complex traits using the results of genome wide association studies. *PloS one*, 9(3): e93017.

Zhao, F., M. S., Meharg, A. (2010). Arsenic as a food chain contaminant:mechanisms of plant uptake and metabolism and mitigation strategies. *Annual Review of Plant Biology* , 61:535–559.

Zhi-Kang, L., JL, X. (2007). Breeding for drought and salt tolerant rice (oryza sativa l.). Progress and perspectives: advances in Molecular Breeding Toward Drought and Salt Tolerant Crops. *Springer Netherlands,* 21:531–564.

Zhu, L., Hu, J., Zhu, K., Fang, Y., Gao, Z., He, Y., Qian, Q. (2011). Identification and characterization of SHORTENED UPPERMOST INTERNODE 1, a gene negatively regulating uppermost internode elongation in rice. *Plant Molecular Biology*, 77, 475.

# PUBLICATIONS, EXTENDED ABSTRACTS AND PRESENTATIONS IN THE PhD COURSE

*Papers*

Biselli, C.,Bagnaresi, P.,Cavalluzzo, D.,Urso, S., Desiderio, F.,Orasen, G.,GianninettiA., Righettini, F.,Gennaro, M.,PerriniR.,, Ben Hassen, M.,Sacchi, GA.,Cattivelli L., Valè,G.(2015).Deep sequencing transcriptional fingerprinting of rice kernels for dissecting grain quality traits. *BMC genomics*, 16(1):1091.

Biscarini , F., Cozzi, P., Casella, L.,Riccardi, P., Vattari, A.,Orasen, G.,Perrini, R.,Tacconi, G., Tondelli, A., Biselli, C., Cattivelli, L., Spindel, J., McCouch, S., Abbruscato, P., Valé, G., Piffanelli, P., Greco, R.(2016).Genome-Wide Association Study for Traits Related to plant and Grain Morphology, and Root Architecture in Temperate Rice Accessions. *Plos One*, 11(5):30155425.

Volante, A., Desiderio, F., Tondelli, A., Perrini, R., Orasen, G., Biselli, C., Valè, G., Riccardi, P., Vattari, A., Cavalluzzo, D., Urso, S., Ben HAssen, M., Fricano, A., Piffanell, P., Cozzi, P., Bisccarini, F., Sacchi, GA., Cattivelli, L., Valè, G. (2017). Genome-wide analysis of japonica rice performance under Limited Water and Permanent Flooding conditions. *Frontiers in Plant Science*, 8:1862.

Ben Hassen, M., Cao, T., Bartholomé, J., Orasen, G., Colombi, C., Rakotomalala, J., Ahmadi, N. (2017). Rice diversity panel provides accurate genomic predictions for complex traits in the progenies of biparental crosses involving members of the panel . *Theoretical and Applied Genetics*, 1-19.

*Extended Abstract*

Oliver, V., Monaco, S., Volante, A., Cochrane, N., Gennaro, M., Orasen, G., Arn Teh, Y. (2016). Preliminary results on yield and CO2 fluxes when using alternate wetting and drying on different varieties of European rice. *Geophysical Research Abstracts*, (p. Vol. 18, EGU2016-9131-1).

Oliver, V., Monaco, S., Volante, A., Cochrane, N., Gennaro, M., Orasen, G., Arn Teh, Y. (2017). The effect of alternate wetting and drying on methane fluxes on different varieties of European rice. *Geophysical Research Abstracts*, (p. Vol. 19, EGU2017-19640).

Orasen, G., Baldoni, E., Abruzzese, A., Pesenti, M., Maghrebi, M., Volante, A., Nocito, F.F., Dell'Orto, M., De Nisi, P., Valè, G., Sacchi, G.A. (2017). Genome-Wide Association Study for Mid-Salt Tolerance in Rice. *ITRC Annual* Meeting-2017.

*Oral presentation*

Orasen, G., Volante, A., Tondelli, A., Sacchi, GA.(2017). GWAS: Genotyping By Sequency.*Annual meeting project NEURICE.St. Carles de la Rà*.

# ANNEX

*ANNEX 1: List of accessions used in the study with their geographical origin, taxonomical group, commercial class and group assigned by Structure analysis.*

| Accession | Origin | Group | Commercial Class | Structure Group |
|---|---|---|---|---|
| A201 | USA | tropical japonica | Long B | k2 |
| A301 | USA | tropical japonica | Long B | k2 |
| ADAIR | USA | tropical japonica | Long B | k2 |
| ADELAIDE CHIAPPELLI | ITALY | temperate japonica | Long A | k1 |
| AGATA | ITALY | temperate japonica | Long A | k1 |
| AGOSTANO | ITALY | temperate japonica | Long A | k1 |
| AIACE | ITALY | tropical japonica | Long A | k2 |
| AKITAKOMACHI | JAPAN | temperate japonica | Round | k1 |
| ALAN | USA | tropical japonica | Long B | k2 |
| ALEXANDROS | GREECE | tropical japonica | Long B | k2 |
| ALICE | ITALY | temperate japonica | Long A | k1 |
| ALLORIO | ITALY | temperate japonica | Long A | k1 |
| ALPE | ITALY | temperate japonica | Long A | k1 |
| ALPHA | ITALY | temperate japonica | Long A | k1 |
| AMERICANO | ITALY | temperate japonica | Round | k1 |
| ANSEATICO | ITALY | temperate japonica | Long A | admixed |
| ANTARES | ITALY | temperate japonica | Long A | admixed |
| ANTONI | BULGARIA | temperate japonica | Long A | k1 |
| APOLLO | ITALY | tropical japonica | Long B | k2 |
| ARBORIO | ITALY | temperate japonica | Long A | k1 |
| ARGO | ITALY | temperate japonica | Medium | k1 |
| ARIETE | ITALY | temperate japonica | Long A | k1 |
| ARSENAL | ITALY | tropical japonica | Long B | k2 |
| ARTEMIDE | ITALY | tropical japonica | Long B | admixed |
| AUGUSTO | ITALY | temperate japonica | Long A | k1 |

| | | | | |
|---|---|---|---|---|
| BAHIA | SPAIN | temperate japonica | Medium | k1 |
| BAIXET | SPAIN | temperate japonica | Long A | k1 |
| BALDO | ITALY | temperate japonica | Long A | k1 |
| BALILLA | ITALY | temperate japonica | Round | k1 |
| BALZARETTI | ITALY | temperate japonica | Medium | k1 |
| BARAGGIA | ITALY | temperate japonica | Round | k1 |
| BEIRAO | PORTUGAL | temperate japonica | Long A | k1 |
| BELLE PATNA | USA | tropical japonica | Long B | k2 |
| BENGAL | USA | temperate japonica | Long A | admixed |
| BERTONE | ITALY | temperate japonica | Long A | k1 |
| BIANCA | ITALY | temperate japonica | Long A | k1 |
| BOMBILLA | SPAIN | temperate japonica | Medium | k1 |
| BOMBON | SPAIN | temperate japonica | Medium | k1 |
| BONNI | ITALY | temperate japonica | Long A | k1 |
| BRAZOS | USA | tropical japonica | Long A | k2 |
| BURMA | ITALY | tropical japonica | Long A | admixed |
| CALENDAL | FRANCE | temperate japonica | Long A | k1 |
| CALMOCHI 101 | USA | temperate japonica | Medium | k1 |
| CAMPINO | PORTUGAL | temperate japonica | Medium | k1 |
| CAPATAZ | SPAIN | temperate japonica | Long A | k1 |
| CARINA | BULGARIA | temperate japonica | Round | k1 |
| CARIOCA | ITALY | tropical japonica | Long B | admixed |
| CARMEN | ITALY | temperate japonica | Long A | k1 |
| CARNAROLI | ITALY | temperate japonica | Long A | k1 |
| CARNISE | ITALY | temperate japonica | Long A | k1 |
| CARRICO | PORTUGAL | temperate japonica | Round | k1 |
| CASTELMOCHI | ITALY | temperate japonica | Round | k1 |
| CENTAURO | ITALY | temperate japonica | Round | k1 |
| CHIPKA | BULGARIA | temperate japonica | Round | k1 |
| CIGALON | FRANCE | temperate japonica | Medium | k1 |
| CINIA 40 | CILE | temperate japonica | Long A | k1 |

| | | | | |
|---|---|---|---|---|
| CLOT | SPAIN | temperate japonica | Medium | k1 |
| COCODRIE | USA | tropical japonica | Long B | k2 |
| COLINA | SPAIN | temperate japonica | Round | k1 |
| CORBETTA | ITALY | temperate japonica | Medium | k1 |
| CRESO | ITALY | temperate japonica | Long A | k1 |
| CRIPTO | ITALY | temperate japonica | Round | k1 |
| CT36 | COLOMBIA | temperate japonica | Long B | k1 |
| CT58 | COLOMBIA | temperate japonica | Long A | k1 |
| DELFINO | ITALY | temperate japonica | Long A | k1 |
| DELLMONT | USA | tropical japonica | Long B | k2 |
| DELLROSE | USA | tropical japonica | Long A | k2 |
| DIMITRA | GREECE | temperate japonica | Long A | k1 |
| DIXIEBELLE | USA | tropical japonica | Long A | k2 |
| DOURADAO | BRAZIL | tropical japonica | Long A | k2 |
| DRAGO | ITALY | temperate japonica | Long A | k1 |
| DREW | USA | tropical japonica | Long B | k2 |
| DUCATO | ITALY | temperate japonica | Round | k1 |
| ERCOLE | ITALY | temperate japonica | Long A | k1 |
| ERMES | ITALY | tropical japonica | Long B | admixed |
| ESCARLATE | PORTUGAL | temperate japonica | Round | k1 |
| ESTRELA | PORTUGAL | temperate japonica | Long A | admixed |
| EUROPA | ITALY | temperate japonica | Long A | k1 |
| EUROSE | ITALY | temperate japonica | Long A | k1 |
| EUROSIS | ITALY | temperate japonica | Long A | admixed |
| FAMILIA 181 | PORTUGAL | temperate japonica | Long A | k1 |
| FAST | | tropical japonica | Long B | k2 |
| FIDJI | PHILIPPINES | tropical japonica | Long B | admixed |
| FLIPPER | ITALY | temperate japonica | Long B | k1 |
| FRANCES | SPAIN | temperate japonica | Medium | k1 |
| FULGENTE | ITALY | temperate japonica | Medium | k1 |
| GALILEO | ITALY | temperate japonica | Long A | k1 |

| | | | | |
|---|---|---|---|---|
| GANGE | ITALY | tropical japonica | Long B | k2 |
| GARDE SADRI | TURKEY | temperate japonica | Long A | k1 |
| GIADA | ITALY | tropical japonica | Long B | k2 |
| GIANO | ITALY | tropical japonica | Long B | admixed |
| GIGANTE VERCELLI | ITALY | temperate japonica | Long A | k1 |
| GIOVANNI MARCHETTI | ITALY | temperate japonica | Medium | k1 |
| GITANO | ITALY | temperate japonica | Long A | k1 |
| GIZA 177 | EGYPT | temperate japonica | Medium | k1 |
| GLADIO | ITALY | tropical japonica | Long B | k2 |
| GLORIA | AUSTRALIA | temperate japonica | Long A | k1 |
| GOOLARAH | FRANCE | tropical japonica | Long B | k2 |
| GRAAL | FRANCE | tropical japonica | Long B | admixed |
| GRALDO | ITALY | tropical japonica | Long B | admixed |
| GREGGIO | ITALY | temperate japonica | Long A | k1 |
| GREPPI | ITALY | tropical japonica | Round | admixed |
| GRITNA | ITALY | temperate japonica | Long A | k1 |
| GUADIAMAR | SPAIN | temperate japonica | Medium | k1 |
| GZ8367 | EGYPT | temperate japonica | Round | k1 |
| HANDAO 11 | CHINA | temperate japonica | Round | k1 |
| HANDAO 297 | CHINA | temperate japonica | Round | k1 |
| HAREM | PORTUGAL | temperate japonica | Long A | k1 |
| HARRA | AUSTRALIA | temperate japonica | Round | k1 |
| HONDURAS | SPAIN | tropical japonica | Long A | k2 |
| IAC32 52 | BRAZIL | tropical japonica | Long B | k2 |
| IBO 380-33 | PORTUGAL | temperate japonica | Long A | k1 |
| IBO 400 | PORTUGAL | temperate japonica | Long A | k1 |
| ITALMOCHI | ITALY | temperate japonica | Medium | k1 |
| ITALPATNA 48 | ITALY | temperate japonica | Long A | k1 |
| ITALPATNAxMILYANG | PORTUGAL | temperate japonica | Long A | admixed |
| JACINTO | USA | tropical japonica | Long A | k2 |
| JEFFERSON | USA | tropical japonica | Long A | k2 |

| | | | | |
|---|---|---|---|---|
| JUBILIENI | BULGARIA | temperate japonica | Round | k1 |
| KARNAK | ITALY | temperate japonica | Long A | k1 |
| KING | ITALY | tropical japonica | Long B | k2 |
| KORAL | ITALY | temperate japonica | Long A | k1 |
| KRYSTALLINO | ITALY | temperate japonica | Round | k1 |
| KULON | RUSSIA | temperate japonica | Long A | k1 |
| L201 | USA | tropical japonica | Long B | k2 |
| L202 | USA | tropical japonica | Long B | k2 |
| L204 | USA | tropical japonica | Long B | k2 |
| L205 | USA | tropical japonica | Long B | k2 |
| LACASSINE | USA | tropical japonica | Long B | k2 |
| LAGRUE | USA | tropical japonica | Long A | k2 |
| LAMONE | ITALY | tropical japonica | Long B | k2 |
| LENCINO | ITALY | temperate japonica | Round | k1 |
| LIDO | ITALY | temperate japonica | Medium | k1 |
| LOMELLINO | ITALY | temperate japonica | Medium | k1 |
| LORD | ITALY | temperate japonica | Long A | k1 |
| LOTO | ITALY | temperate japonica | Long A | k1 |
| LUCERO | ITALY | temperate japonica | Round | k1 |
| LUNA | USA | temperate japonica | Medium | k1 |
| LUSITO IRRADIADO | PORTUGAL | temperate japonica | Long A | k1 |
| LUXOR | ITALY | temperate japonica | Long A | k1 |
| M202 | USA | temperate japonica | Medium | k1 |
| M203 | USA | temperate japonica | Long A | k1 |
| M204 | USA | temperate japonica | Long A | k1 |
| M6 | ITALY | temperate japonica | Long A | k1 |
| MAIORAL | PORTUGAL | temperate japonica | Long A | admixed |
| MANTOVA | ITALY | temperate japonica | Long A | k1 |
| MARATELLI | ITALY | temperate japonica | Medium | k1 |
| MARENY | SPAIN | temperate japonica | Long A | k1 |
| MARTE | ITALY | temperate japonica | Round | k1 |

| | | | | |
|---|---|---|---|---|
| MAYBELLE | USA | tropical japonica | Long B | k2 |
| MECO | FRANCE | temperate japonica | Long A | admixed |
| MEJANES | FRANCE | temperate japonica | Long B | admixed |
| MELAS | GREECE | temperate japonica | Long B | k1 |
| MIARA | ITALY | temperate japonica | Long B | k1 |
| MILEV 21 | BULGARIA | temperate japonica | Round | k1 |
| MOLO | ITALY | temperate japonica | Long A | k1 |
| MONTICELLI | ITALY | temperate japonica | Medium | k1 |
| MUGA | PORTUGAL | temperate japonica | Round | k1 |
| MUSA | ITALY | temperate japonica | Long A | k1 |
| NANO | ITALY | temperate japonica | Round | admixed |
| NEMBO | ITALY | temperate japonica | Long A | k1 |
| NILO | ITALY | temperate japonica | Long A | k1 |
| NOVARA | ITALY | temperate japonica | Medium | k1 |
| OLCENENGO | ITALY | temperate japonica | Long A | k1 |
| ONICE | ITALY | temperate japonica | Long A | k1 |
| OPALE | ITALY | temperate japonica | Long A | k1 |
| ORIGINARIO | ITALY | temperate japonica | Round | k1 |
| ORIONE | ITALY | temperate japonica | Long A | k1 |
| OSCARxSUWEON | PORTUGAL | temperate japonica | Long A | admixed |
| OSTIGLIA | ITALY | temperate japonica | Round | k1 |
| OTA | PORTUGAL | temperate japonica | Long A | k1 |
| P6 | ITALY | temperate japonica | Medium | k1 |
| PADANO | ITALY | temperate japonica | Long A | k1 |
| PANDA | ITALY | tropical japonica | Long B | admixed |
| PECOS | USA | tropical japonica | Medium | admixed |
| PEGONIL | SPAIN | temperate japonica | Medium | k1 |
| PELDE | AUSTRALIA | temperate japonica | | k1 |
| PERLA | ITALY | temperate japonica | Round | k1 |
| PIEMONTE | ITALY | temperate japonica | Long A | k1 |
| PIERINA MARCHETTI | ITALY | temperate japonica | Long A | k1 |

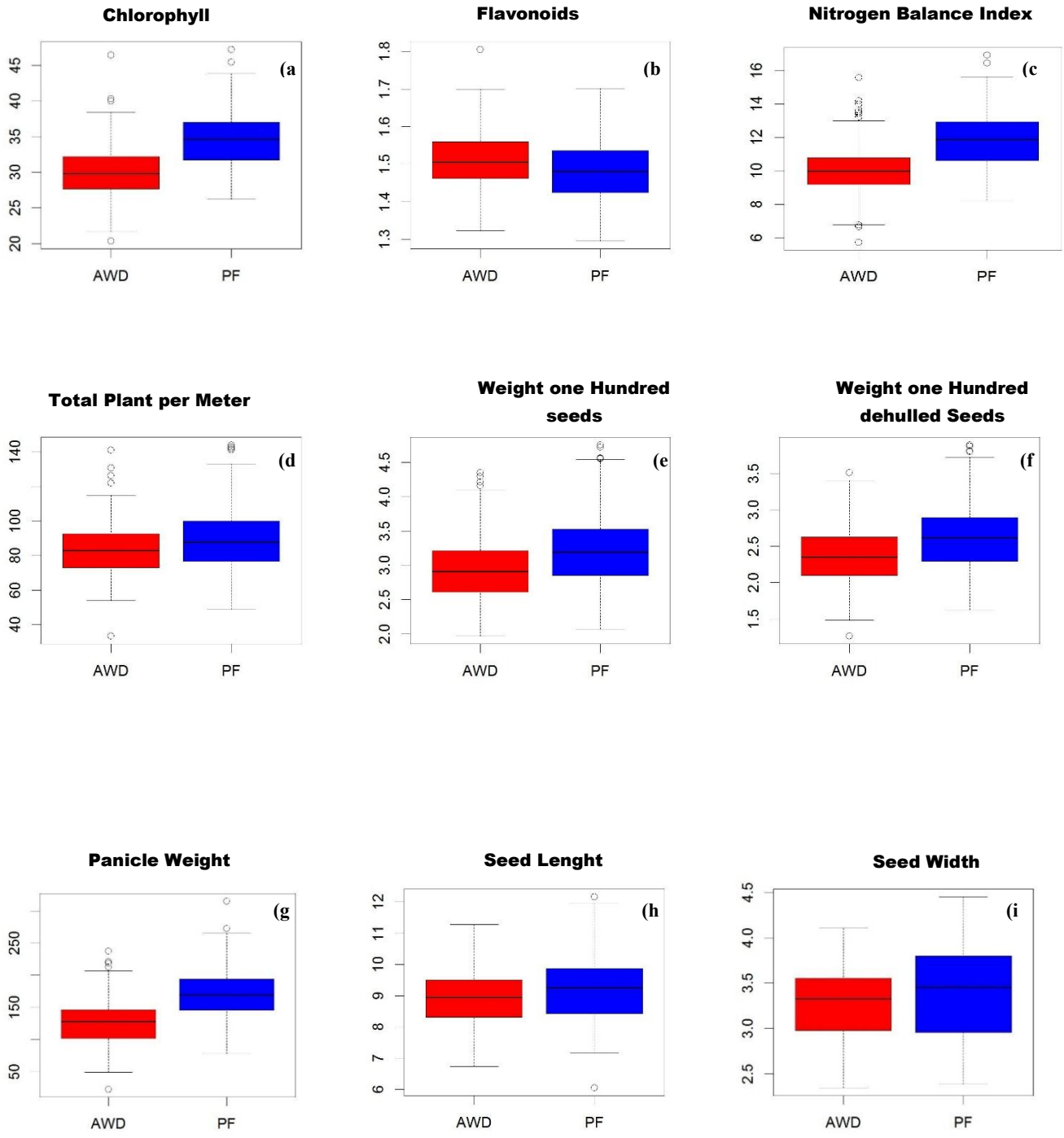
| | | | | |
|---|---|---|---|---|
| PLOVDIV 22 | BULGARIA | temperate japonica | Long A | k1 |
| PLOVDIV 24 | BULGARIA | temperate japonica | Round | k1 |
| PLUS | ITALY | tropical japonica | Long B | k2 |
| PREVER | ITALY | tropical japonica | Long B | admixed |
| PROMETEO | ITALY | temperate japonica | Medium | k1 |
| PUNTAL | SPAIN | tropical japonica | Long B | k2 |
| RANGHINO | ITALY | temperate japonica | Round | k1 |
| RAZZA 77 | ITALY | temperate japonica | Medium | k1 |
| REDI | ITALY | temperate japonica | Long A | k1 |
| REXMONT | USA | tropical japonica | Long B | k2 |
| RIBE | ITALY | temperate japonica | Long A | k1 |
| RINALDO BERSANI | ITALY | temperate japonica | Long A | k1 |
| RINGO | ITALY | temperate japonica | Long A | k1 |
| RIZZOTTO 51 1 | ITALY | temperate japonica | Long A | k1 |
| ROBBIO SEL1 | ITALY | temperate japonica | Long A | k1 |
| RODEO | ITALY | temperate japonica | Long A | k1 |
| RODINA | BULGARIA | temperate japonica | Round | k1 |
| ROMA | ITALY | temperate japonica | Long A | k1 |
| RONALDO | ITALY | tropical japonica | Long A | admixed |
| RONCAROLO | ITALY | temperate japonica | Medium | k1 |
| RONCOLO | ITALY | temperate japonica | Medium | k1 |
| ROTUNDUS | HUNGARY | temperate japonica | Long A | k1 |
| ROXANI | GREECE | temperate japonica | Long A | k1 |
| RPC 12 | CHINA | temperate japonica | Round | k1 |
| RUBI | PORTUGAL | temperate japonica | Long A | k1 |
| RUBINO | ITALY | temperate japonica | Round | k1 |
| RUSSO | ITALY | temperate japonica | | k1 |
| S101 | USA | temperate japonica | Medium | k1 |
| S102 | USA | temperate japonica | Medium | k1 |
| S102 2 | USA | temperate japonica | Medium | k1 |
| SAEDINENIE | BULGARIA | temperate japonica | Long A | k1 |

| | | | | |
|---|---|---|---|---|
| SAFARI | PORTUGAL | temperate japonica | Long A | k1 |
| SAGRES | PORTUGAL | temperate japonica | Long A | k1 |
| SAKHA 102 | EGYPT | temperate japonica | Medium | k1 |
| SAKHA 103 | EGYPT | temperate japonica | Round | k1 |
| SALOIO | PORTUGAL | temperate japonica | Long B | k1 |
| SALVO | ITALY | tropical japonica | Long B | admixed |
| SAMBA | ITALY | tropical japonica | Long A | admixed |
| SANDOCA | PORTUGAL | temperate japonica | Long B | k1 |
| SANDORA | HUNGARY | temperate japonica | Long A | k1 |
| SANT ANDREA | ITALY | temperate japonica | Long A | k1 |
| SANTERNO | ITALY | temperate japonica | Long B | admixed |
| SATURNO | ITALY | tropical japonica | Long B | k2 |
| SAVIO | ITALY | temperate japonica | Long B | k1 |
| SCUDO | ITALY | tropical japonica | Long B | k2 |
| SELENIO | ITALY | temperate japonica | Round | k1 |
| SELN 244A | AUSTRALIA | temperate japonica | Medium | k1 |
| SENATORE NOVELLI | ITALY | temperate japonica | Long A | k1 |
| SENIA | SPAIN | temperate japonica | Medium | k1 |
| SEQUIAL | SPAIN | temperate japonica | Medium | k1 |
| SESIA | ITALY | temperate japonica | Long A | k1 |
| SESIAMOCHI | ITALY | temperate japonica | Long A - Round | k1 |
| SETANTUNO | PORTUGAL | temperate japonica | Round | k1 |
| SFERA | ITALY | temperate japonica | Round | k1 |
| SHSS 381 | SPAIN | temperate japonica | Long A | k1 |
| SHSS 53 | SPAIN | temperate japonica | Long A | k1 |
| SILLA | ITALY | temperate japonica | Long A | k1 |
| SIRIO CL | ITALY | tropical japonica | Long A | k2 |
| SIS R215 | ITALY | tropical japonica | Long A | k2 |
| SLAVA | BULGARIA | temperate japonica | Medium | k1 |
| SMERALDO | ITALY | temperate japonica | Long A | k1 |
| SOURE | PORTUGAL | temperate japonica | Long A | k1 |

| | | | | |
|---|---|---|---|---|
| SPRINT | ITALY | tropical japonica | Long B | k2 |
| SR 113 | SPAIN | temperate japonica | Long A | k1 |
| STRELLA | ITALY | temperate japonica | Long A | k1 |
| SUPER | PORTUGAL | temperate japonica | Long A | k1 |
| T757 | INDIA | temperate japonica | Long A | k1 |
| TAICHUNG 65 | THAILAND | temperate japonica | Long A | k1 |
| TEJO | ITALY | temperate japonica | Long A | admixed |
| TEXMONT | USA | tropical japonica | Long A | k2 |
| THAIBONNET | ITALY | tropical japonica | Long B | k2 |
| THAIPERLA | | temperate japonica | Round | k1 |
| TITANIO | ITALY | temperate japonica | Long A | k2 |
| TOPAZIO | ITALY | temperate japonica | Medium | k1 |
| TORIO | PORTUGAL | temperate japonica | Long A | k1 |
| ULISSE | ITALY | temperate japonica | Long A | k1 |
| ULLAL | SPAIN | temperate japonica | Round | k1 |
| UPLA 32 | ARGENTINA | tropical japonica | Long B | k2 |
| UPLA 63 | ARGENTINA | tropical japonica | Long B | admixed |
| UPLA 64 | ARGENTINA | tropical japonica | Long B | admixed |
| UPLA 66 | ARGENTINA | tropical japonica | Long B | admixed |
| UPLA 68 | ARGENTINA | tropical japonica | Long B | k2 |
| UPLA 75 | ARGENTINA | tropical japonica | Long B | admixed |
| UPLA 77 | ARGENTINA | tropical japonica | Long B | admixed |
| UPLA 80 | ARGENTINA | tropical japonica | Long B | admixed |
| UPLA 91 | ARGENTINA | tropical japonica | Long B | k2 |
| VALTEJO | PORTUGAL | temperate japonica | Round | k1 |
| VELA | ITALY | temperate japonica | Long A | k1 |
| VENERE | ITALY | temperate japonica | Long B | k1 |
| VENERIA | ITALY | temperate japonica | Long A | k1 |
| VIALE | ITALY | temperate japonica | Long A | k1 |
| VIALONE 190 | ITALY | temperate japonica | Medium | k1 |
| VIALONE NANO | ITALY | temperate japonica | Medium | k1 |

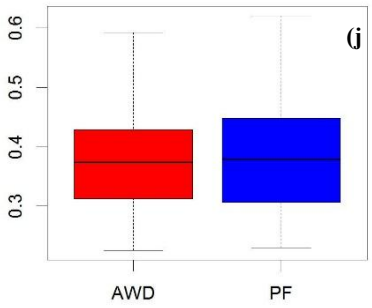
| | | | | |
|---|---|---|---|---|
| VIALONE NERO | ITALY | temperate japonica | Round | k1 |
| VICTORIA | ARGENTINA | temperate japonica | Round | k1 |
| VIRGO | ITALY | temperate japonica | Medium | k1 |
| VOLANO | ITALY | temperate japonica | Long A | k1 |
| VULCANO | ITALY | temperate japonica | Long A | k1 |
| XIANGOU2 | CHINA | tropical japonica | Medium | admixed |
| YRM 6 2 | AUSTRALIA | temperate japonica | Medium | k1 |
| ZENA | ITALY | tropical japonica | Long B | k2 |

*ANNEX 2: Box-plots of the phenotypic values of agronomi traits.In red the performances in AWD of 281 varieties are illustrated, in blue the performances in PF.*

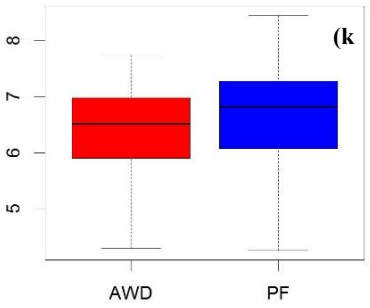


Seeed Width/length ratio

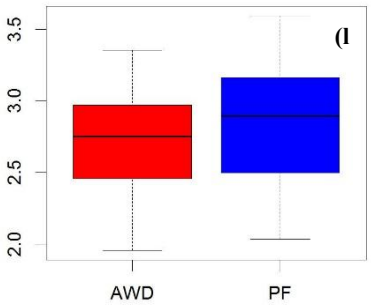
APRV



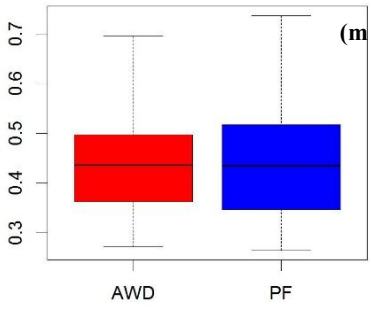

Dehulled Seed Length



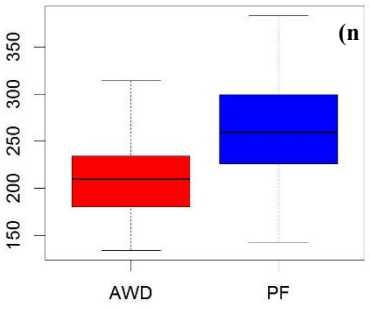

Dehulled Seed Width



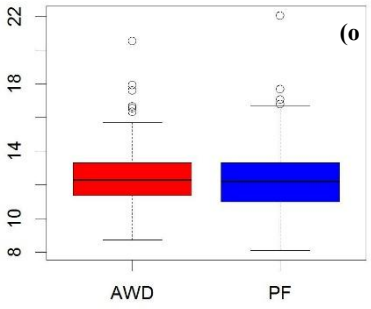

Dehulled SeedWidth/Length Ratio



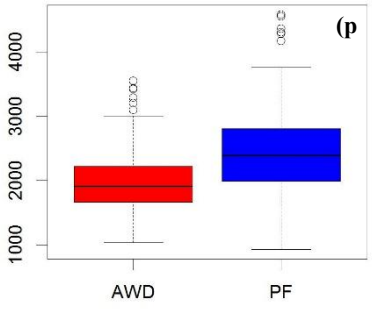

Flag Leaf Length



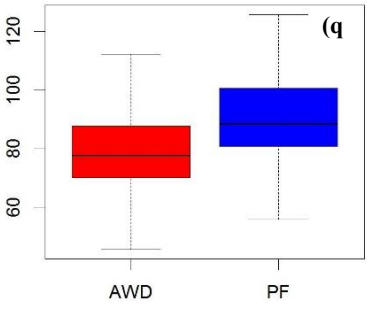

Flag Leaf Width



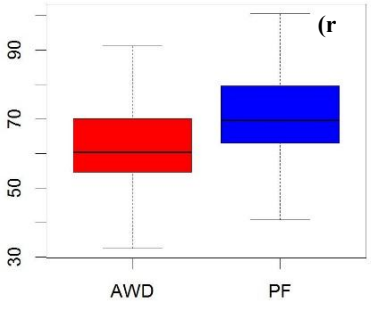

Leaf Area

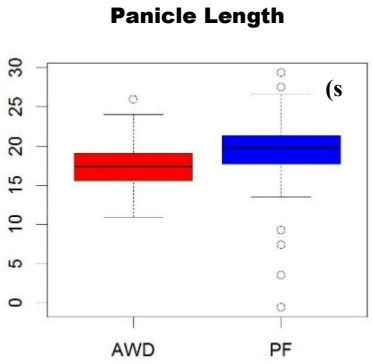


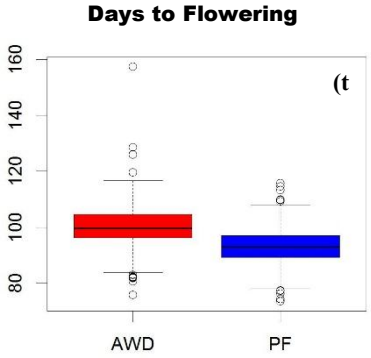

Panicle Height

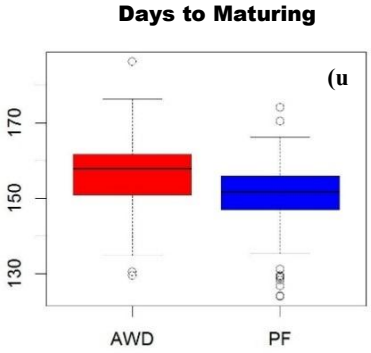


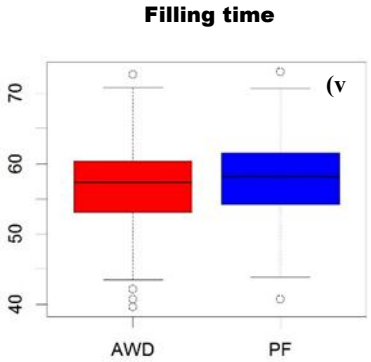

Panicle Neck Height



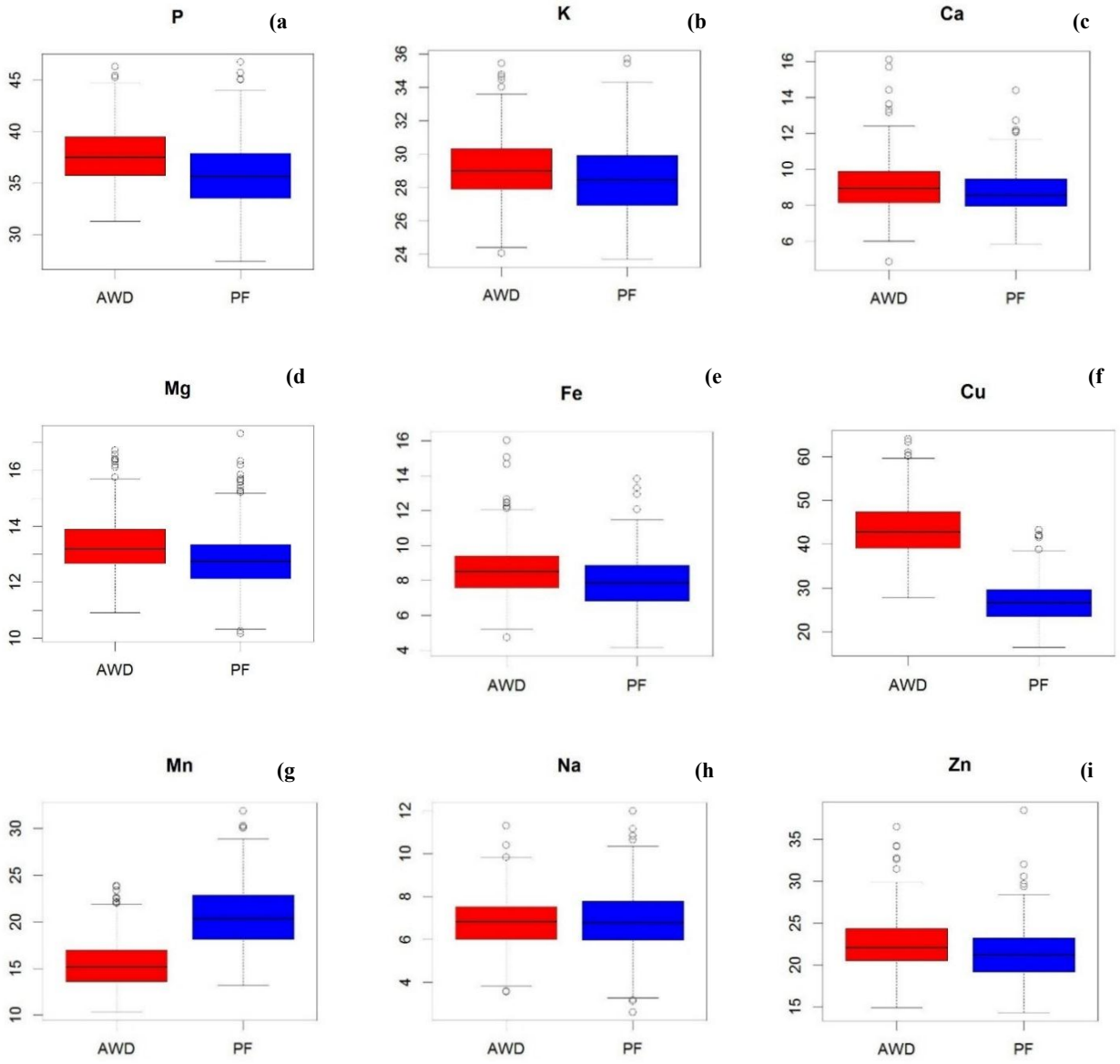

Panicle Length
(s
0
5
10
15
20
25
30
AWD
PF

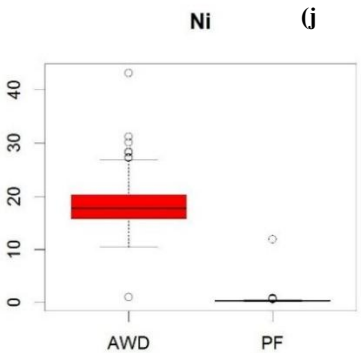

Days to Flowering
(t
80
100
120
140
160
AWD
PF


Days to Maturing
(u
130
150
170
AWD
PF

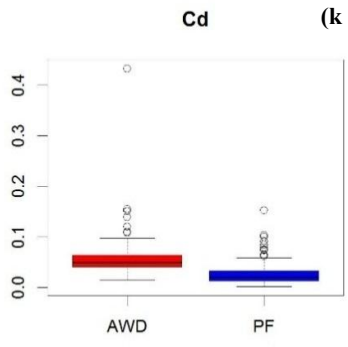

Filling time
(v
40
50
60
70
AWD
PF

*ANNEX 3:Box-plots of the phenotypic values of ionomics traits.In red the performances in AWD of 281 varieties are illustrated, in blue the performances in PF.*

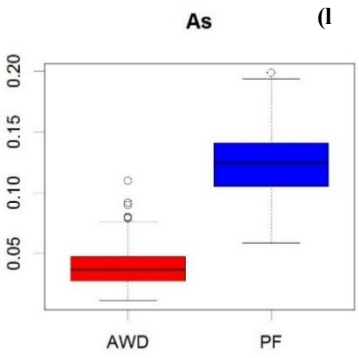



Ni
0
10
20
30
40
AWD
PF

(j

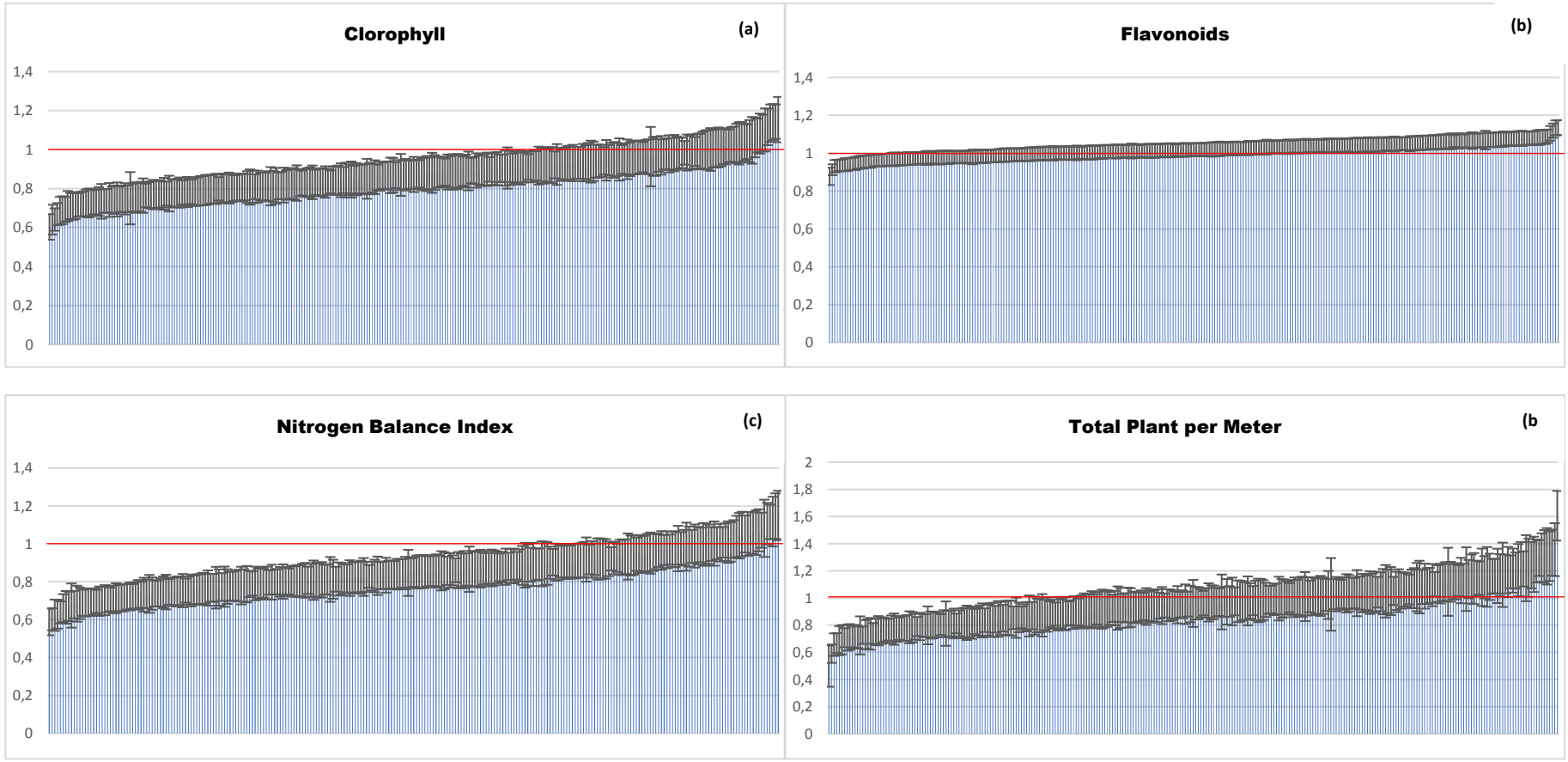

Cd
0.0
0.1
0.2
0.3
0.4
AWD
PF

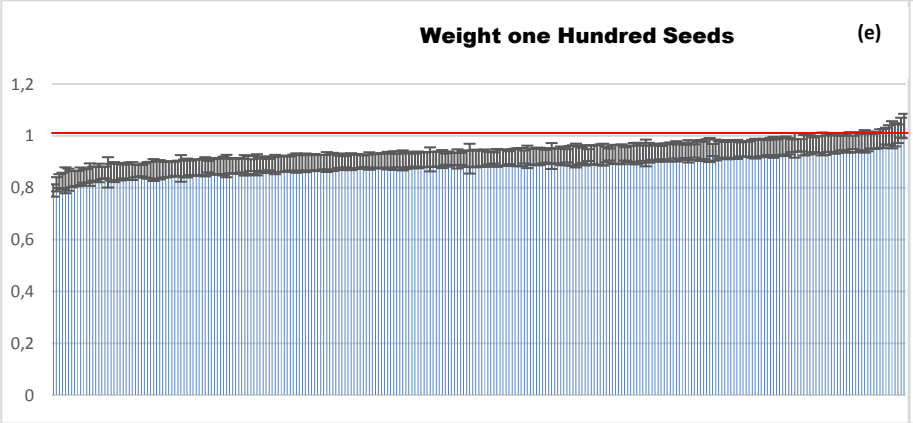
(k

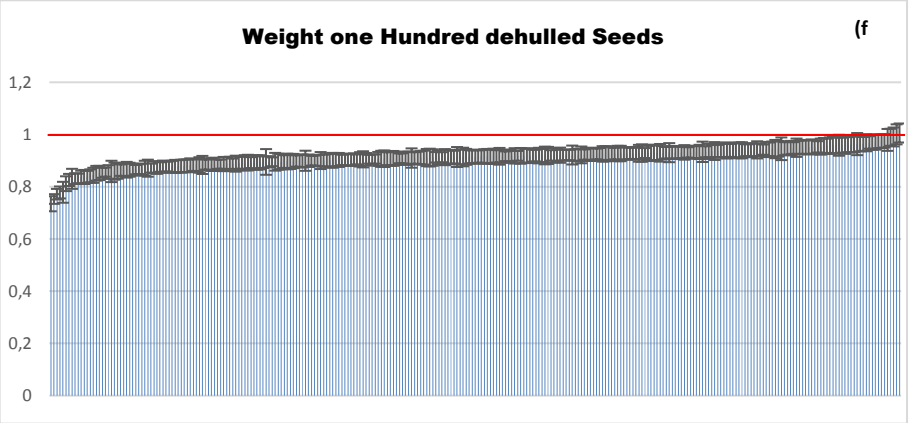

As
0.05
0.10
0.15
0.20
AWD
PF

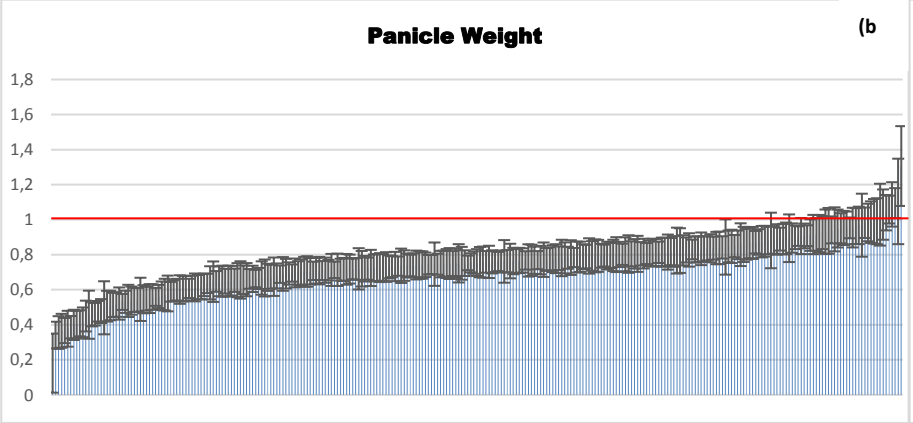
(l

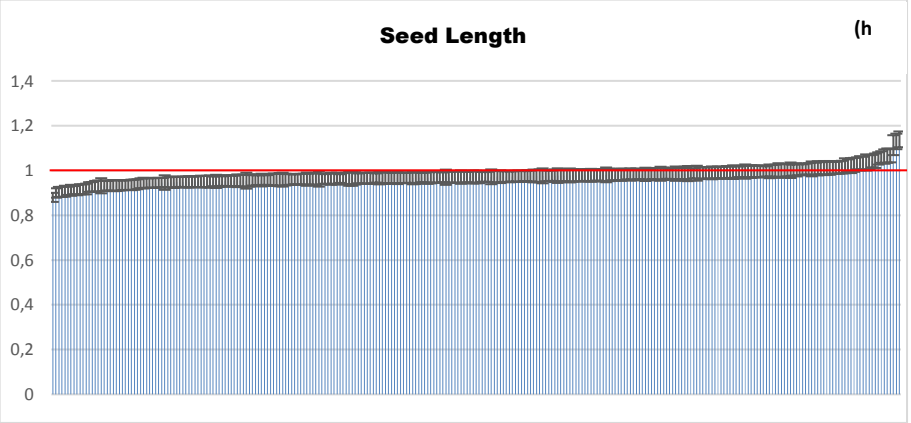
*ANNEX 4:Phenotypic value ratio (AWD/PF) for all the traits used for the GWAS. The red horizontal line represents the 1 value. Vertical bars represent the values of SD.*

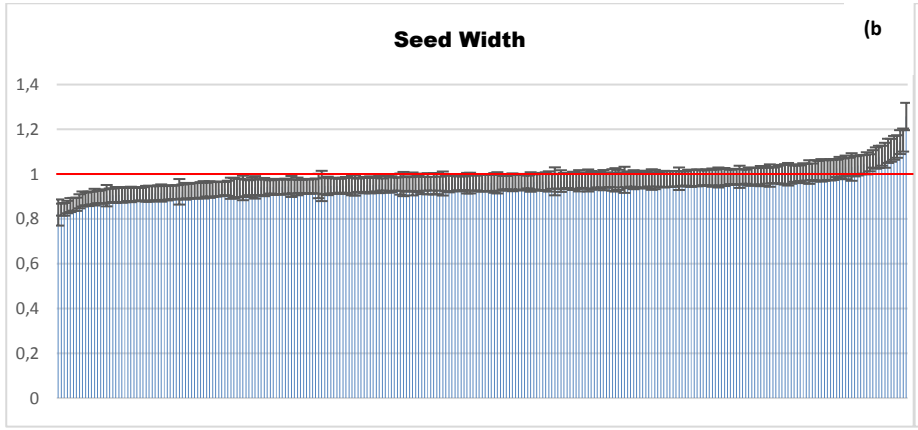


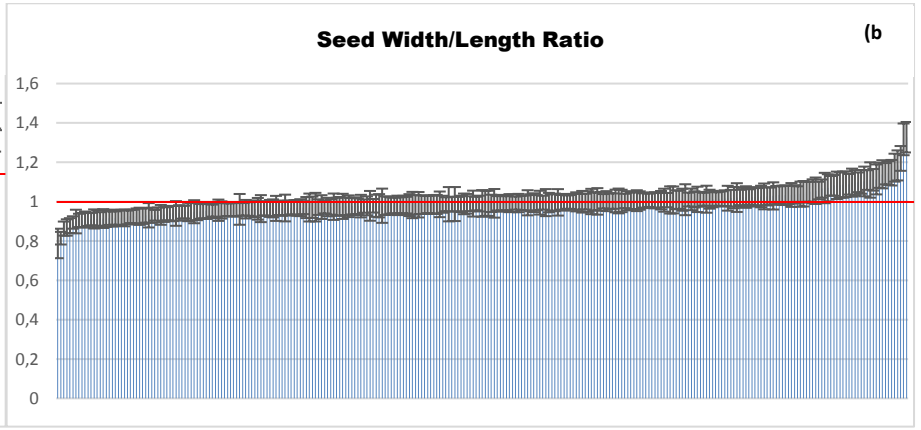

Weight one Hundred Seeds
(e)
1,2
1
0,8
0,6
0,4
0,2
0

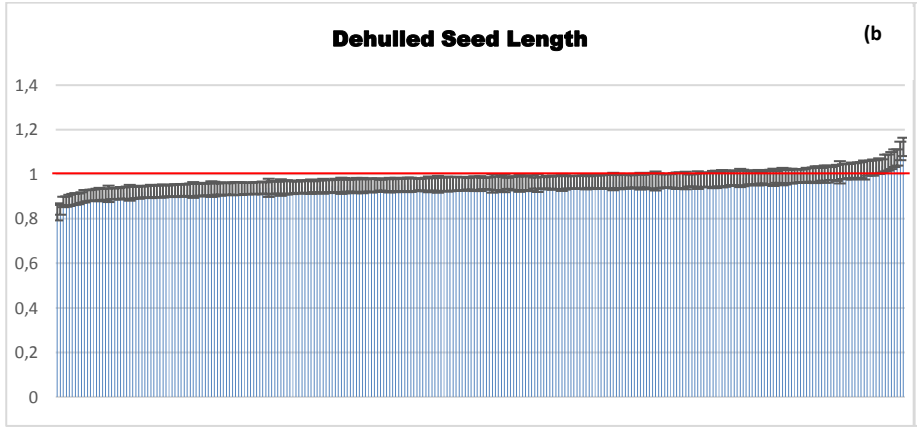

Weight one Hundred dehulled Seeds
(f
1,2
1
0,8
0,6
0,4
0,2
0

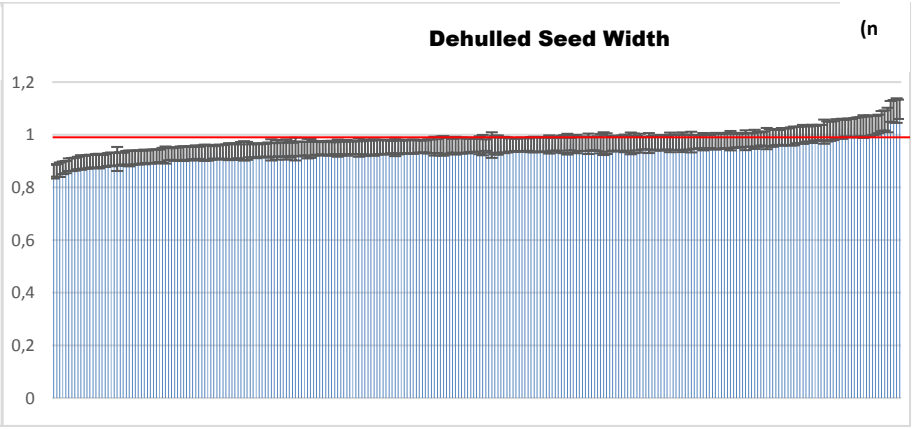

Panicle Weight
(b
1,8
1,6
1,4
1,2
1
0,8
0,6
0,4
0,2
0

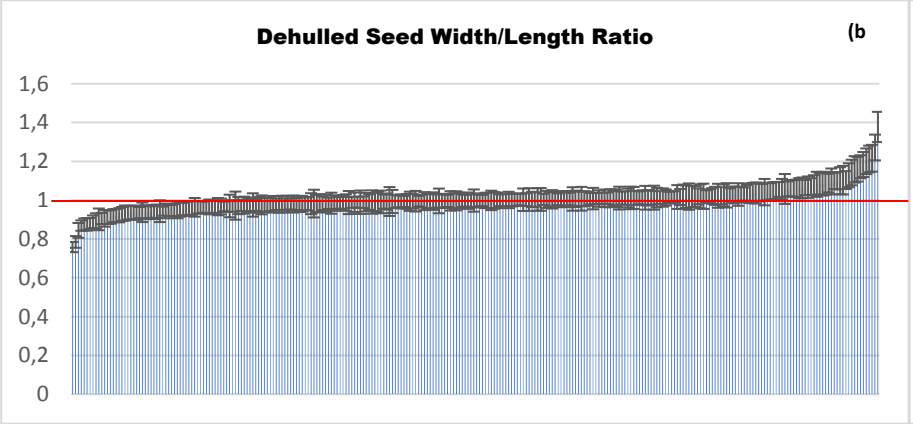

Seed Length
(h
1,4
1,2
1
0,8
0,6
0,4
0,2
0

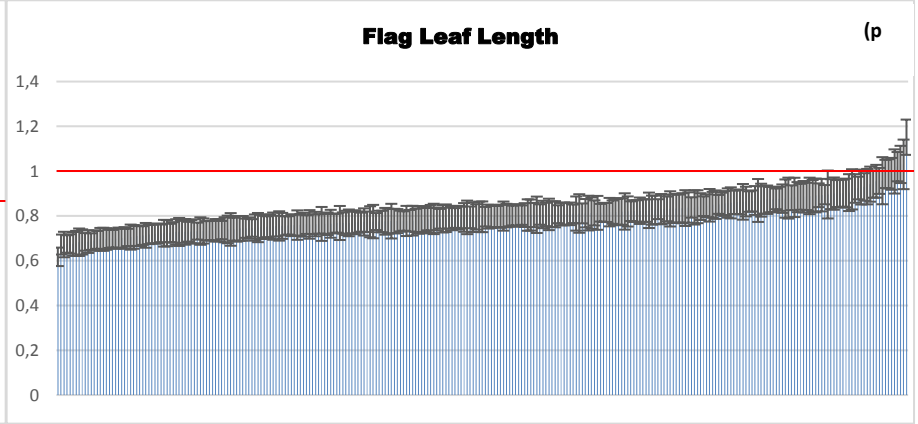

Seed Width
(b
1,4
1,2
1
0,8
0,6
0,4
0,2
0

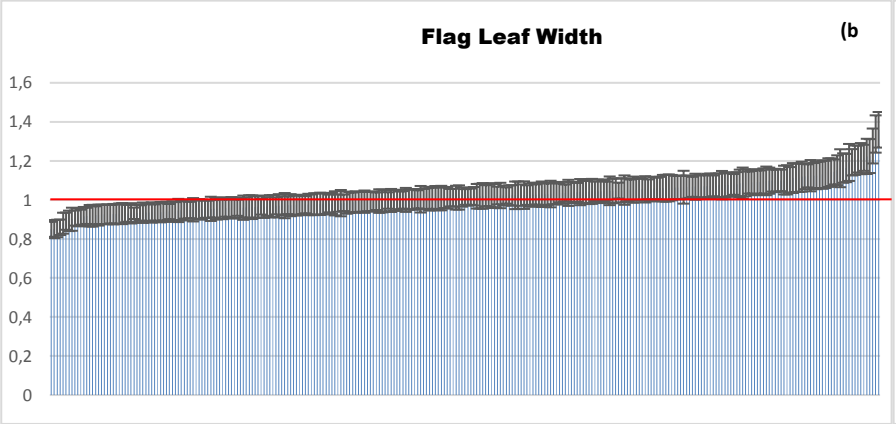

Seed Width/Length Ratio
(b
1,6
1,4
1,2
1
0,8
0,6
0,4
0,2
0

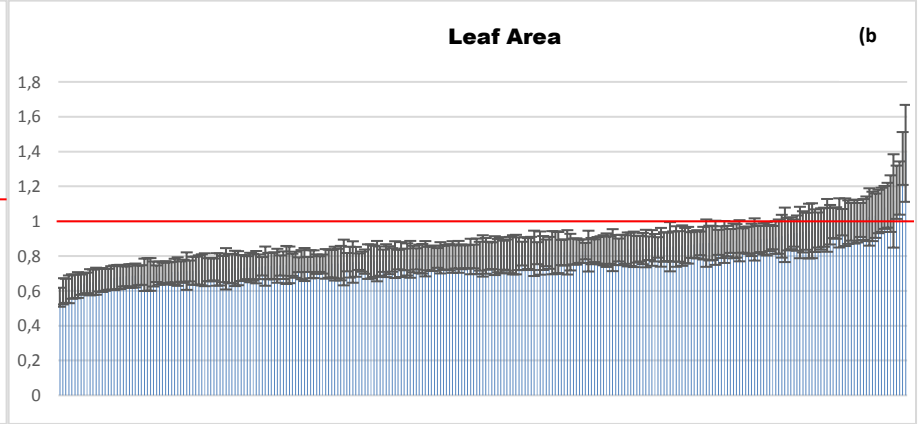

Dehulled Seed Length
(b
1,4
1,2
1
0,8
0,6
0,4
0,2
0

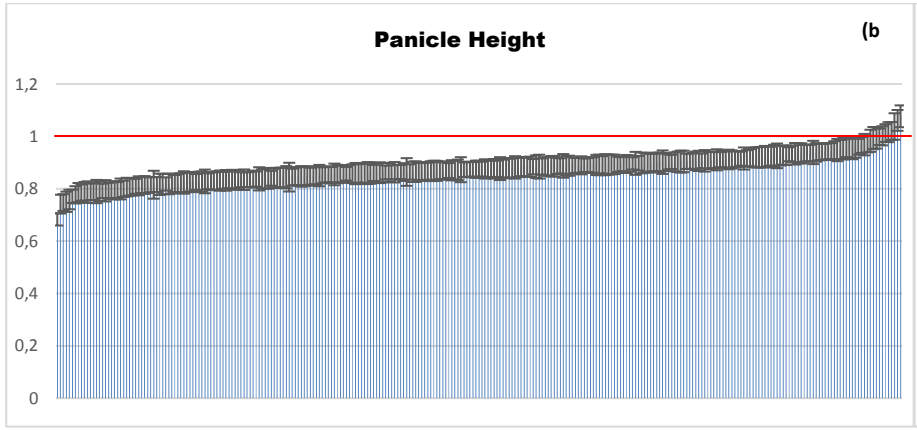

Dehulled Seed Width
(n
1,2
1
0,8
0,6
0,4
0,2
0

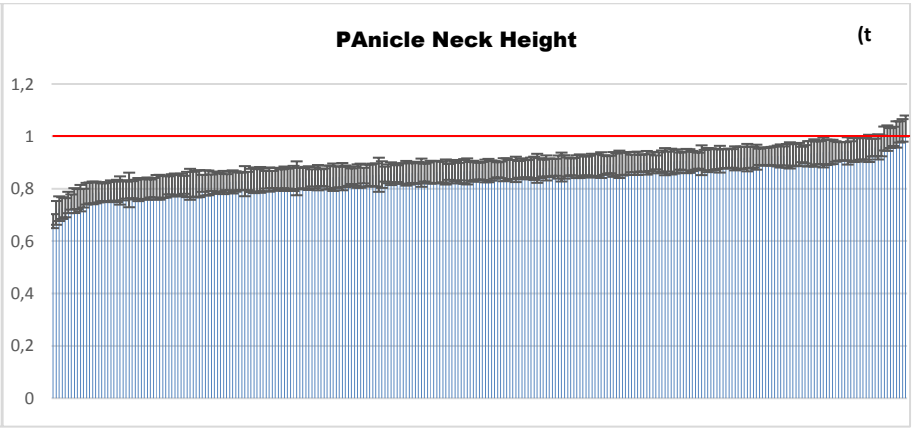

Dehulled Seed Width/Length Ratio
(b
1,6
1,4
1,2
1
0,8
0,6
0,4
0,2
0

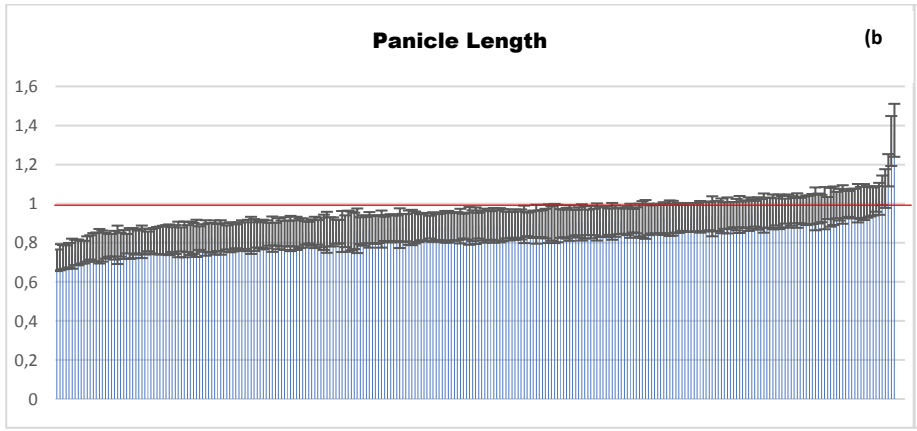

Flag Leaf Length
(p
1,4
1,2
1
0,8
0,6
0,4
0,2
0

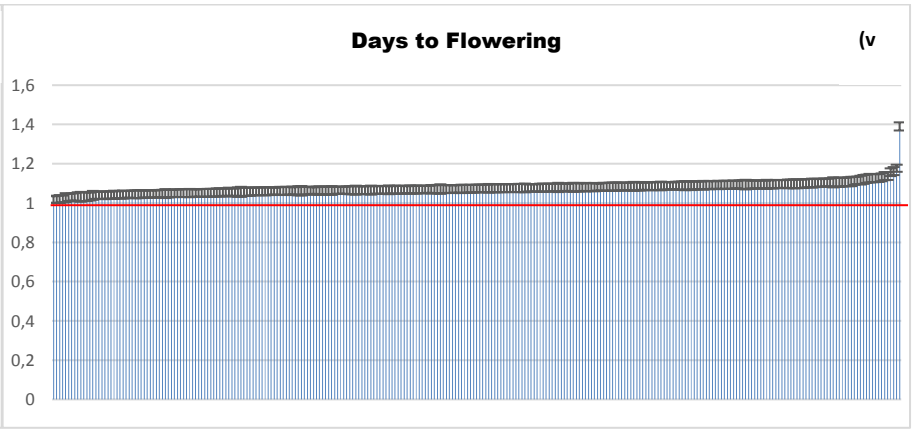

Flag Leaf Width
(b
1,6
1,4
1,2
1
0,8
0,6
0,4
0,2
0

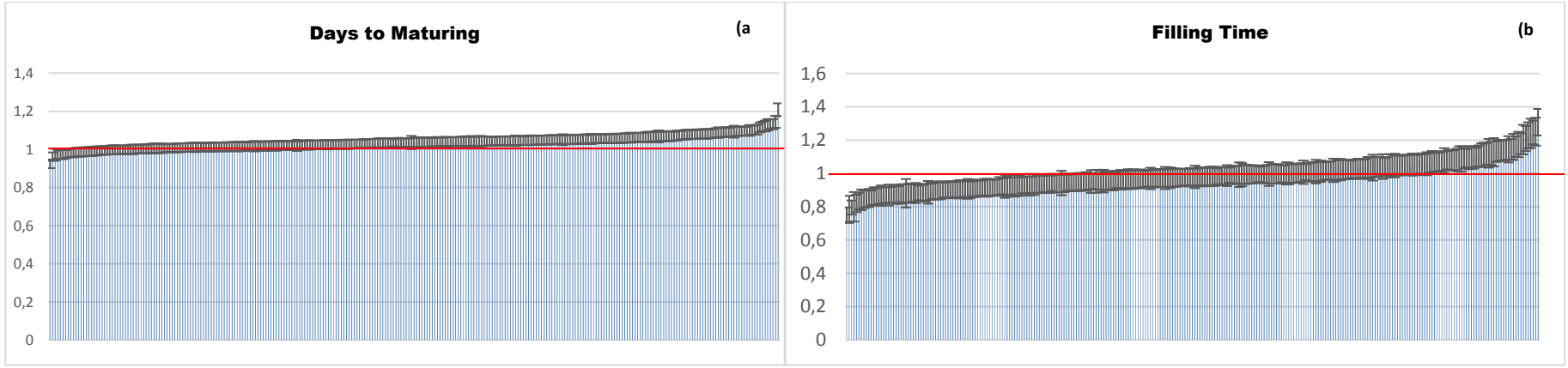

Leaf Area
(b
1,8
1,6
1,4
1,2
1
0,8
0,6
0,4
0,2
0


Panicle Height
(b
1,2
1
0,8
0,6
0,4
0,2
0

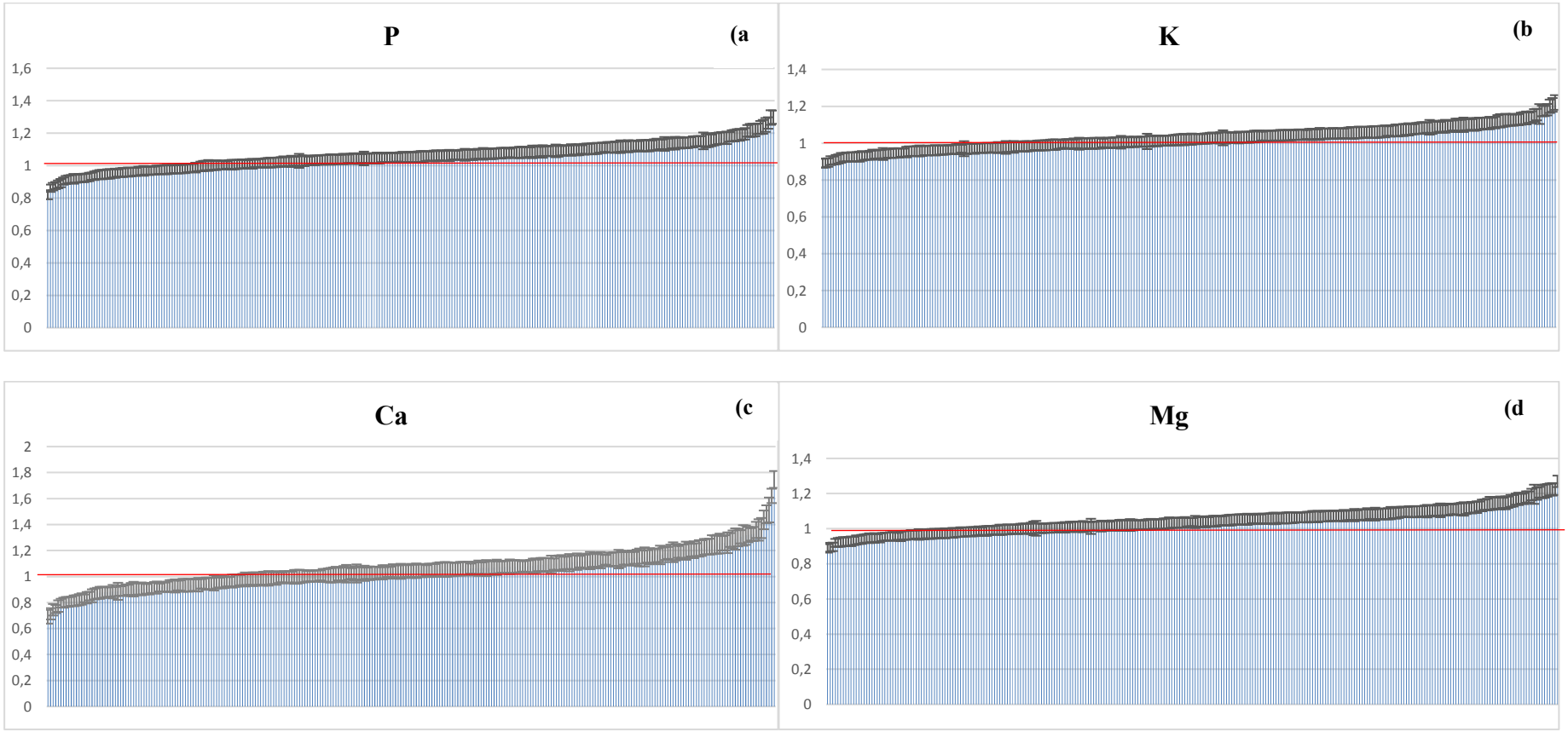

PAnicle Neck Height
(t
1,2
1
0,8
0,6
0,4
0,2
0

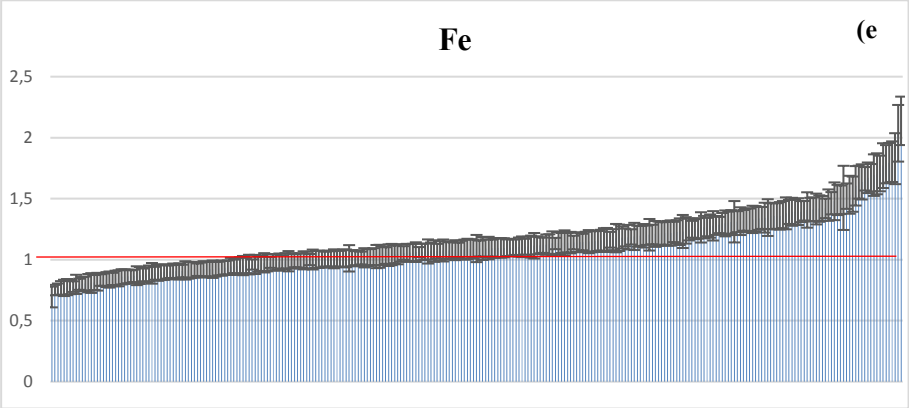

Panicle Length
(b
1,6
1,4
1,2
1
0,8
0,6
0,4
0,2
0

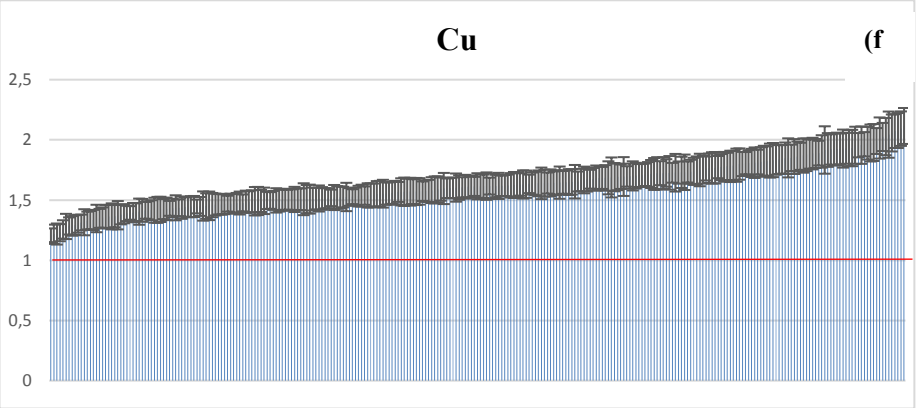

Days to Flowering
(v
1,6
1,4
1,2
1
0,8
0,6
0,4
0,2
0

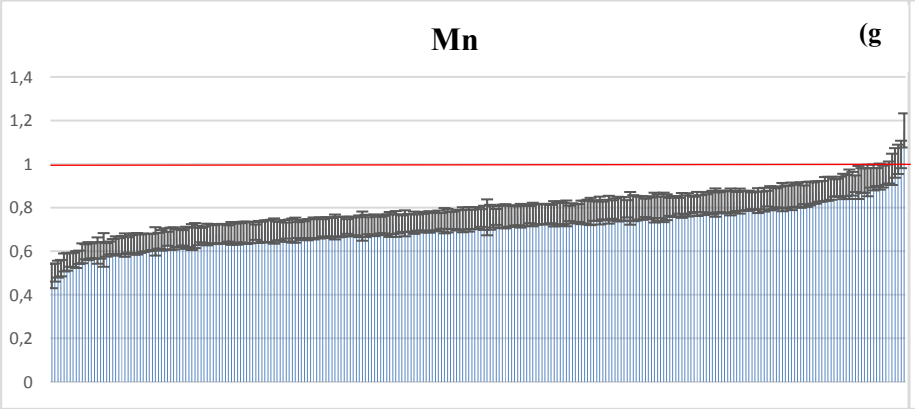

Days to Maturing
(a
1,4
1,2
1
0,8
0,6
0,4
0,2
0
Filling Time
(b
1,6
1,4
1,2
1
0,8
0,6
0,4
0,2
0

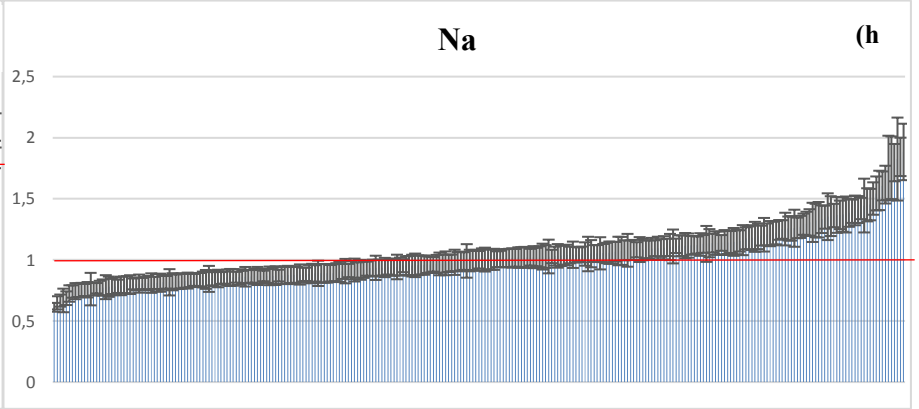
*ANNEX 5:Phenotypic value ratio (AWD/PF) for all the traits used for the GWAS. The red horizontal line represents the 1 value. Vertical bars represent the values of SD.*

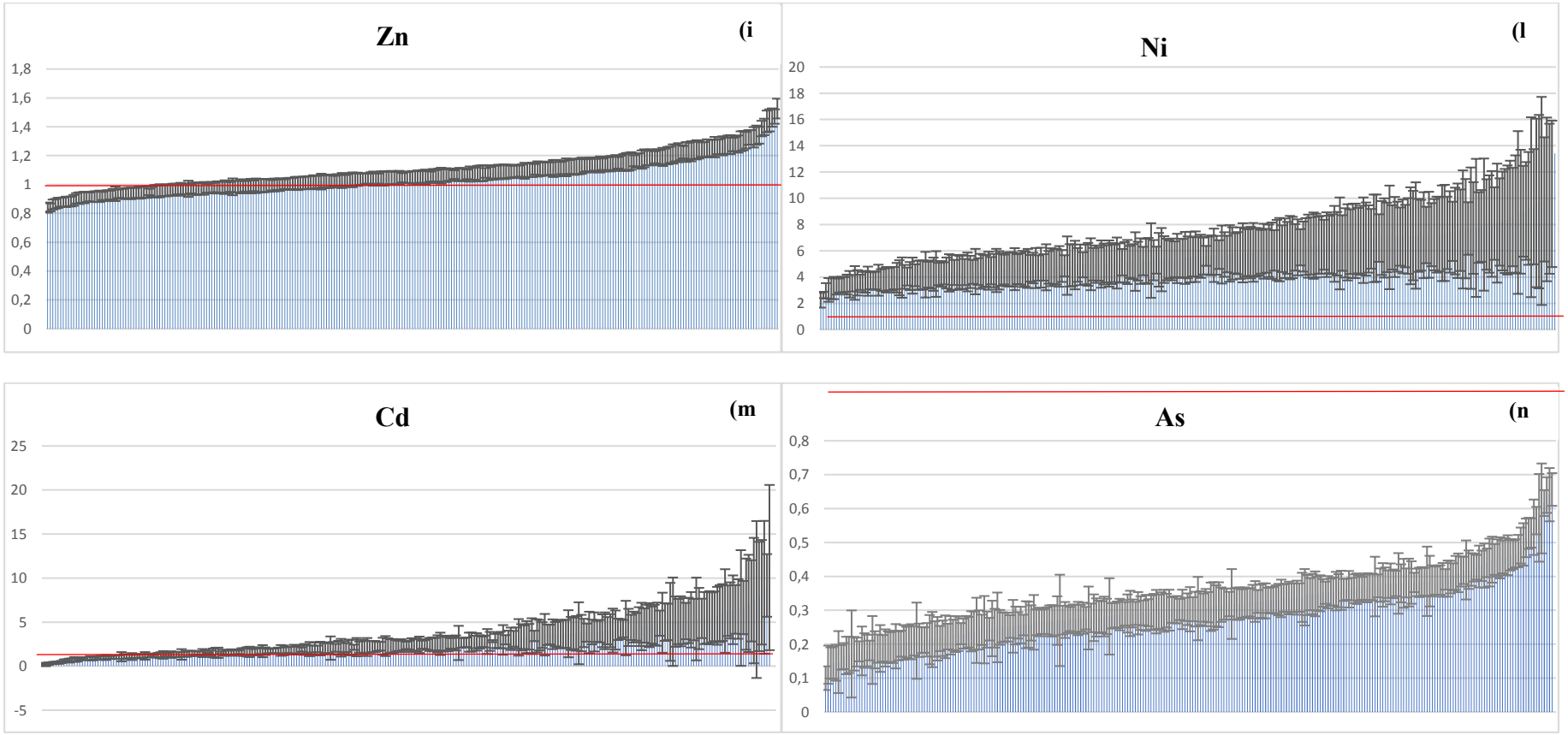



Fe
(e
2,5
2
1,5
1
0,5
0


Cu
(f
2,5
2
1,5
1
0,5
0


Mn
(g
1,4
1,2
1
0,8
0,6
0,4
0,2
0


Na
(h
2,5
2
1,5
1
0,5
0


Zn
(i
1,8
1,6
1,4
1,2
1
0,8
0,6
0,4
0,2
0
Ni
(l
20
18
16
14
12
10
8
6
4
2
0
Cd
(m
25
20
15
10
5
0
-5
As
(n
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0

.

*ANNEX 6a: Value of some agronomic traits in the rice accessions subjected to the two watering treatments. Data are reported as two-years least-square mean values.*

| | Mean value ± SD | | *p*_value | CV | | H | |
|---|---|---|---|---|---|---|---|
| Trait | AWD | PF | | AWD | PF | AWD | PF |
| CHL_AV | 30.1 ± 6 | 34.59 ± 6 | <.0001 | 0.21 | 0.19 | 0.57 | 0.59 |
| FLA | 3.0 ± 0.2 | 2.9 ± 0.3 | <.0001 | 0.08 | 0.12 | 0.65 | 0.49 |
| NBI_AV | 10.0 ± 2 | 11.8 ± 2 | <.0001 | 0.22 | 0.22 | 0.59 | 0.55 |
| TPM | 83.8 ± 29 | 88.7 ± 24 | <.0001 | 0.36 | 0.27 | 0.48 | 0.63 |
| HGW (*g*) | 2.9 ± 0.5 | 3.2 ± 0.6 | <.0001 | 0.18 | 0.17 | 0.94 | 0.97 |
| NHGW (*g*) | 2.4 ± 0.4 | 2.6 ± 0.5 | <.0001 | 0.18 | 0.18 | 0.94 | 0.97 |
| PW (c*m*) | 12.7 ± 4 | 17.2 ± 4 | <.0001 | 0.34 | 0.26 | 0.77 | 0.82 |
| SL (*mm*) | 8.9 ± 1 | 9.2 ± 1 | <.0001 | 0.11 | 0.11 | 0.93 | 0.98 |
| SW (*mm*) | 3.2 ± 0.4 | 3.4 ± 0.5 | <.0001 | 0.15 | 0.15 | 0.91 | 0.99 |
| SWLR | 0.37 ± 0.1 | 0.38 ± 0.1 | 0.1837 | 0.22 | 0.22 | 0.93 | 0.99 |
| NSL (*mm*) | 6.4 ± 0.7 | 6.7 ± 0.8 | <.0001 | 0.12 | 0.13 | 0.91 | 0.97 |
| NSW (*mm*) | 2.7 ± 0.3 | 2.8 ± 0.4 | <.0001 | 0.14 | 0.14 | 0.94 | 0.99 |
| NSWLR | 0.43 ± 0.1 | 0.43 ± 0.1 | 0.7776 | 0.23 | 0.23 | 0.94 | 0.99 |
| FLL (*mm*) | 208.9 ± 43 | 262.61 ± 58 | <.0001 | 0.21 | 0.22 | 0.85 | 0.86 |
| FLW (*mm*) | 12.5 ± 2 | 12.3 ± 2 | 0.0668 | 0.22 | 0.21 | 0.87 | 0.88 |
| LA ($cm^2$) | 19.6 ± 0.7 | 24.5 + 0.8 | <.0001 | 0.36 | 0.34 | 0.81 | 0.86 |
| PH (*cm*) | 79.3 ± 13 | 90.7 ± 13 | <.0001 | 0.17 | 0.15 | 0.93 | 0.97 |
| PNH (*cm*) | 61.9 ± 12 | 71.1 + 12 | <.0001 | 0.20 | 0.17 | 0.93 | 0.97 |
| PL (*cm*) | 17.4 ± 3 | 19.6 ± 3 | <.0001 | 0.19 | 0.17 | 0.83 | 0.83 |
| DF (*d*) | 100.2 ± 11 | 96.2 ± 8 | <.0001 | 0.11 | 0.08 | 0.87 | 0.83 |
| DM (*d*) | 156.5 ± 13 | 150.8 ± 10 | <.0001 | 0.09 | 0.07 | 0.64 | 0.93 |
| DFM (*d*) | 56.6 ± 8 | 57.6 ± 8 | 0.0015 | 0.15 | 0.14 | 0.42 | 0.54 |

*SD: Standard Deviation; CV: Coefficient of Variation of the panel; H: broad-sense heritability; P-value: significant difference among treatment (AWD and PF).*

*ANNEX 6b: Value of level of some inorganic elements in the brown grain of the rice accessions subjected to the two watering treatments. Data are reported as two-years least-square mean values.*

| Element | Mean value ± SD | | *p*_value | CV | | H | |
|---|---|---|---|---|---|---|---|
| | AWD | PF | | AWD | PF | AWD | PF |
| **P***(g kg$^{-1}$)* | 3.8 ± 0.3 | 3.6 ± 0.3 | <0.0001 | 0.07 | 0.10 | 0.98 | 0.99 |
| **K***(g kg$^{-1}$)* | 2.9 ± 0.2 | 2.8 ± 0.2 | <0.0001 | 0.07 | 0.08 | 0.99 | 0.99 |
| **Ca** *(mg kg$^{-1}$)* | 91.1 ± 14 | 87.4 ± 13 | <0.0001 | 0.16 | 0.15 | 0.99 | 0.99 |
| **Mg***(g kg$^{-1}$)* | 1.3 ± 0.1 | 1.2 ± 0.1 | <0.0001 | 0.07 | 0.08 | 0.99 | 0.99 |
| **Fe***(mg kg$^{-1}$)* | 8.7 ± 1 | 7.8 ± 1 | <0.0001 | 0.18 | 0.21 | 0.98 | 0.99 |
| **Cu***(mg kg$^{-1}$)* | 4.3 ± 0.7 | 2.7 ± 0.4 | <0.0001 | 0.16 | 0.17 | 0.37 | 0.82 |
| **Mn***(mg kg$^{-1}$)* | 15.5 ± 2 | 20.7 ± 3 | <0.0001 | 0.17 | 0.16 | 0.98 | 0.99 |
| **Na***(mg kg$^{-1}$)* | 6.9 ± 1 | 6.8 ± 1 | 0.617 | 0.19 | 0.22 | 0.99 | 0.99 |
| **Zn***(mg kg$^{-1}$)* | 22.6 ± 3 | 21.3 ± 3 | <0.0001 | 0.14 | 0.16 | 0.99 | 0.99 |
| **Ni***(mg kg$^{-1}$)* | 1.8 ± 0.4 | 0.38 ± 0.1 | <0.0001 | 0.21 | 0.34 | 0.97 | 0.51 |
| **Cd***(μg kg$^{-1}$)* | 54.5 ± 30 | 25.8 ± 1 | <0.0001 | 0.55 | 0.72 | 0.81 | 0.76 |
| **As***(μg kg$^{-1}$)* | 39 ± 10 | 123 ± 27 | <0.0001 | 0.39 | 0.22 | 0.99 | 0.98 |

*SD: Standard Deviation; CV: Coefficient of Variation of the panel; H: broad-sense heritability; P-value: significant difference among treatment (AWD and PF).*

*ANNEX 7a: MTAs for physiology parameters.*

| Trait | Condition | Association ID | Peak Marker/Region | Chr. | -log10(p) | SNPs | Associated region | | | Peak marker $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Start | End | Size | |
| CHL | AWD | CHL-5-1 | S5_21635103 | 5 | 3.684 | 1 | - | - | - | 4.90% |
| CHL | AWD | CHL-12-1 | S12_22218430 | 12 | 4.812 | 12 | 22,217,293 | 22,263,968 | 46,675 | 6.50% |
| CHL | BOTH | CHL-1-2 | S1_37826127 | 1 | 4.225 | 4 | 37,478,451 | 38,016,190 | 537,739 | 5.90% |
| CHL | PF | CHL-1-1 | S1_31694634 | 1 | 3.757 | 1 | - | - | - | 4.90% |
| CHL | PF | CHL-6-1 | S6_27160011 | 6 | 4.627 | 2 | 26,853,772 | 27,160,011 | 306,239 | 6.30% |
| CHL | PF | CHL-9-1 | S9_9433258 | 9 | 4.074 | 2 | 9,433,258 | 9,452,790 | 19,532 | 4.50% |
| CHL | PF | CHL-11-1 | S11_17562184 | 11 | 4.026 | 4 | 17,562,184 | 17,714,592 | 152,408 | 5.30% |
| FLA | BOTH | FLA-1-2 | S1_2073706 | 1 | 3.638 | 1 | - | - | - | 4.70% |
| FLA | BOTH | FLA-3-2 | S3_10946807 | 3 | 3.971 | 3 | 10,946,807 | 11,126,829 | 180,022 | 5.30% |
| FLA | PF | FLA-1-1 | S1_170244 | 1 | 3.914 | 1 | - | - | - | 5.20% |
| FLA | PF | FLA-2-1 | S2_8395817 | 2 | 3.752 | 1 | - | - | - | 4.90% |
| FLA | PF | FLA-3-1 | S3_2881274 | 3 | 3.78 | 1 | - | - | - | 5.10% |
| FLA | PF | FLA-4-1 | S4_33954928 | 4 | 4.311 | 1 | - | - | - | 5.80% |
| FLA | PF | FLA-6-1 | S6_10184497 | 6 | 3.67 | 1 | - | - | - | 5.00% |
| FLA | PF | FLA-11-1 | S11_24293732 | 11 | 3.807 | 1 | - | - | - | 5.00% |
| NBI | AWD | NBI-12-1 | S12_22218430 | 12 | 3.939 | 1 | - | - | - | 5.20% |
| NBI | BOTH | NBI-1-2 | S1_37826127 | 1 | 3.742 | 2 | 37,478,451 | 37,826,127 | 347,676 | 5.20% |
| NBI | BOTH | NBI-12-2 | S12_25921238 | 12 | 4.158 | 2 | 25,921,238 | 25,926,845 | 5,607 | 5.60% |
| NBI | PF | NBI-1-1 | S1_31694634 | 1 | 4.754 | 6 | 31,676,378 | 31,736,205 | 59,827 | 6.40% |
| NBI | PF | NBI-6-1 | S6_27160011 | 6 | 4.187 | 1 | - | - | - | 5.60% |
| NBI | PF | NBI-11-1 | S11_17714592 | 11 | 4.072 | 5 | 17,562,184 | 17,727,964 | 165,780 | 5.40% |

*ANNEX 7b: MTAs for yield relaited traits.*

| Trait | Condition | Association ID | Peak Marker/Region | Chr. | -log10(p) | SNPs | Associated region | | | Peak marker $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Start | End | Size | |
| HGW | AWD | HGW-4-2 | S4_31322465 | 4 | 4.194 | 6 | 31,178,449 | 31,322,465 | 144,016 | 5.60% |
| HGW | AWD | HGW-12-1 | S12_16580963 | 12 | 3.689 | 1 | - | - | - | 4.80% |
| HGW | AWD | HGW-12-2 | S12_22105068-S12_22297746 | 12 | 3.501 | 9 | 22,105,068 | 22,330,619 | 225,551 | 4.50% |
| HGW | BOTH | HGW-5-1 | S5_5285253-S5_5789766 | 5 | 4.963 | 20 | 4,926,453 | 6,191,511 | 1,265,058 | 6.60% |
| HGW | BOTH | HGW-8-1 | S8_26929484 | 8 | 4.188 | 20 | 26,856,646 | 26,944,705 | 88,059 | 5.50% |
| HGW | BOTH | HGW-9-1 | S9_21382154 | 9 | 6.513 | 6 | 21,225,371 | 22,196,951 | 971,580 | 9.00% |
| HGW | PF | HGW-1-1 | S1_23753836 | 1 | 3.628 | 1 | - | - | - | 4.70% |
| HGW | PF | HGW-4-1 | S4_31178449-S4_31243954 | 4 | 3.994 | 4 | 31,178,449 | 31,243,954 | 65,505 | 5.20% |
| HGW | PF | HGW-6-1 | S6_18125239 | 6 | 3.453 | 1 | - | - | - | 4.40% |
| HGW | PF | HGW-7-1 | S7_25416737 | 7 | 3.774 | 1 | - | - | - | 5.00% |
| NHGW | AWD | NHGW-10-1 | S10_4699332 | 10 | 3.513 | 1 | - | - | - | 4.50% |
| NHGW | BOTH | NHGW-4-1 | S4_31322465 | 4 | 4.848 | 9 | 31,178,449 | 31,342,633 | 164,184 | 6.50% |
| NHGW | BOTH | NHGW-5-1 | S5_5285253 | 5 | 5.185 | 21 | 4,926,453 | 5,789,766 | 863,313 | 7.00% |
| NHGW | BOTH | NHGW-8-1 | S8_26929484 | 8 | 3.771 | 11 | 26,856,646 | 26,929,672 | 73,026 | 4.90% |
| NHGW | BOTH | NHGW-9-1 | S9_21382154 | 9 | 5.87 | 4 | 21,225,371 | 22,196,951 | 971,580 | 8.10% |
| NHGW | PF | NHGW-1-1 | S1_23753836 | 1 | 3.741 | 3 | 23,662,182 | 23,753,836 | 91,654 | 4.90% |
| NHGW | PF | NHGW-1-2 | S1_33101881 | 1 | 3.467 | 1 | - | - | - | 4.50% |
| NHGW | PF | NHGW-7-1 | S7_23302632 | 7 | 3.91 | 2 | 22,983,134 | 23,302,632 | 319,498 | 5.00% |
| PW | AWD | PW-2-1 | S2_24962679 | 2 | 4.159 | 2 | 24,941,038 | 24,962,679 | 21,641 | 5.50% |
| PW | BOTH | PW-4-1 | S4_31257415 | 4 | 4.73 | 3 | 31,257,415 | 31,495,347 | 237,932 | 6.30% |
| PW | BOTH | PW-7-1 | S7_27661814 | 7 | 4.1 | 1 | - | - | - | 5.40% |
| PW | BOTH | PW-8-1 | S8_5279375 | 8 | 3.907 | 1 | - | - | - | 5.20% |
| PW | BOTH | PW-8-2 | S8_6313336 | 8 | 3.989 | 1 | - | - | - | 5.30% |
| TPM | AWD | TPM-4-2 | S4_33083637 | 4 | 5.921 | 11 | 31,495,347 | 33,083,637 | 1,588,290 | 8.10% |
| TPM | AWD | TPM-8-1 | S8_5272898 | 8 | 3.656 | 1 | - | - | - | 4.80% |
| TPM | BOTH | TPM-7-1 | S7_27788464 | 7 | 3.807 | 2 | 27,788,464 | 27,797,026 | 8,562 | 5.00% |
| TPM | PF | TPM-4-1 | S4_31406024 | 4 | 4.585 | 8 | 31,406,024 | 31,449,182 | 43,158 | 6.20% |
| TPM | PF | TPM-8-2 | S8_28049635 | 8 | 3.876 | 1 | - | - | - | 5.10% |

*ANNEX 7c: MTAs for grain traits.*

| Trait | Condition | Association ID | Peak Marker/Region | Chr. | -log10(p) | SNPs | Associated region | | | Peak marker $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Start | End | Size | |
| NSL | AWD | NSL-5-1 | S5_28291673 | 5 | 3.722 | 1 | - | - | - | 5.10% |
| NSL | BOTH | NSL-1-1 | S1_29909662 | 1 | 4.851 | 2 | 29,909,662 | 30,014,864 | 105,202 | 6.90% |
| NSL | BOTH | NSL-3-1 | S3_16378774 | 3 | 4.532 | 7 | 16,019,763 | 18,442,519 | 2,422,756 | 6.20% |
| NSL | PF | NSL-8-1 | S8_26913933 | 8 | 4.079 | 1 | - | - | - | 5.40% |
| NSL | PF | NSL-12-1 | S12_17866204 | 12 | 3.611 | 1 | - | - | - | 4.80% |
| NSW | AWD | NSW-8-1 | S8_5275795 | 8 | 3.797 | 1 | - | - | - | 5.00% |
| NSW | BOTH | NSW-3-1 | S3_7907626 | 3 | 6.807 | 28 | 7,770,213 | 10,142,408 | 2,372,195 | 9.50% |
| NSW | BOTH | NSW-7-1 | S7_25004532 | 7 | 6.442 | 2 | 25,004,532 | 25,211,630 | 207,098 | 9.00% |
| NSW | BOTH | NSW-12-1 | S12_17780115 | 12 | 3.841 | 3 | 17,063,073 | 17,780,115 | 717,042 | 5.10% |
| NSWLR | BOTH | NSWLR-3-1 | S3_15945566 | 3 | 3.949 | 5 | 15,945,452 | 15,964,532 | 19,080 | 5.20% |
| NSWLR | BOTH | NSWLR-5-1 | S5_5500521-S5_5538628 | 5 | 8.344 | 266 | 4,541,573 | 6,387,006 | 1,845,433 | 11.80% |
| NSWLR | BOTH | NSWLR-5-2 | S5_28104538 | 5 | 4.79 | 2 | 28,104,538 | 29,682,002 | 1,577,464 | 6.50% |
| SL | AWD | SL-2-1 | S2_31933331 | 2 | 4.092 | 2 | 31,933,331 | 32,044,550 | 111,219 | 5.40% |
| SL | AWD | SL-6-1 | S6_24522491 | 6 | 3.607 | 1 | - | - | - | 4.70% |
| SL | BOTH | SL-1-1 | S1_1504722 | 1 | 4.027 | 1 | - | - | - | 5.30% |
| SL | BOTH | SL-3-1 | S3_16378774 | 3 | 5.839 | 29 | 15,455,137 | 17,620,428 | 2,165,291 | 8.10% |
| SL | BOTH | SL-5-1 | S5_5500521-S5_5538628 | 5 | 5.842 | 11 | 5,474,539 | 5,540,758 | 66,219 | 7.40% |
| SL | BOTH | SL-8-1 | S8_26913933 | 8 | 6.213 | 38 | 26,767,608 | 28,380,474 | 1,612,866 | 8.30% |
| SL | PF | SL-3-2 | S3_26891194 | 3 | 3.882 | 1 | - | - | - | 5.20% |
| SW | BOTH | SW-4-1 | S4_31316330 | 4 | 3.858 | 5 | 31,316,330 | 31,342,633 | 26,303 | 5.00% |
| SW | BOTH | SW-5-1 | S5_5500521-S5_5538628 | 5 | 13.269 | 272 | 4,926,453 | 6,857,145 | 1,930,692 | 18.40% |
| SW | BOTH | SW-7-1 | S7_23302632 | 7 | 7.11 | 5 | 22,698,945 | 25,004,532 | 2,305,587 | 9.50% |
| SW | BOTH | SW-10-1 | S10_4699332 | 10 | 3.759 | 2 | 4,272,649 | 4,699,332 | 426,683 | 4.90% |
| SW | PF | SW-1-1 | S1_24768357 | 1 | 6.049 | 11 | 23,653,192 | 24,952,323 | 1,299,131 | 8.20% |
| SW | PF | SW-1-2 | S1_29770662 | 1 | 3.471 | 1 | - | - | - | 4.40% |
| SW | PF | SW-1-3 | S1_42726259 | 1 | 3.645 | 3 | 42,592,434 | 42,726,259 | 133,825 | 4.70% |
| SW | PF | SW-6-1 | S6_3991637 | 6 | 3.889 | 1 | - | - | - | 5.00% |
| SW | PF | SW-6-2 | S6_19272666 | 6 | 3.733 | 4 | 18,908,056 | 19,272,666 | 364,610 | 4.80% |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| SW | PF | SW-8-1 | S8_26865734-S8_26944705 | 8 | 4.157 | 8 | 26,865,734 | 26,944,853 | 79,119 | 5.40% |
| SW | PF | SW-9-1 | S9_21561117 | 9 | 4.542 | 5 | 19,584,784 | 21,892,523 | 2,307,739 | 6.00% |
| SW | PF | SW-10-2 | S10_10176295 | 10 | 3.401 | 1 | - | - | - | 4.40% |
| SW | PF | SW-12-1 | S12_4190069 | 12 | 3.417 | 1 | - | - | - | 4.40% |
| SW | PF | SW-12-2 | S12_21652306 | 12 | 4.031 | 1 | - | - | - | 5.30% |
| SWLR | AWD | SWLR-1-1 | S1_1504722 | 1 | 3.967 | 1 | - | - | - | 5.30% |
| SWLR | BOTH | SWLR-3-1 | S3_16378774 | 3 | 5.499 | 6 | 15,454,973 | 16,687,981 | 1,233,008 | 7.60% |
| SWLR | BOTH | SWLR-5-1 | S5_5500521-S5_5538628 | 5 | 14.036 | 269 | 5,181,849 | 7,022,613 | 1,840,764 | 19.40% |
| SWLR | BOTH | SWLR-8-1 | S8_26969268 | 8 | 4.27 | 3 | 26,913,933 | 27,195,766 | 281,833 | 5.60% |
| SWLR | PF | SWLR-1-2 | S1_3708821 | 1 | 5.695 | 2 | 3,624,642 | 3,708,821 | 84,179 | 7.80% |
| SWLR | PF | SWLR-7-1 | S7_22983134 | 7 | 6.121 | 4 | 19,935,696 | 23,302,632 | 3,366,936 | 7.80% |

*ANNEX 7d: MTAs for morphological traits.*

| Trait | Condition | Association ID | Peak Marker/Region | Chr. | -log10(p) | SNPs | Associated region | | | Peak marker $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Start | End | Size | |
| FLL | AWD | FLL-1-2 | S1_28599543-S1_29024851* | 1 | 3.762 | 1 | 28,186,751 | 29,184,329 | 997,578 | 5.00% |
| FLL | AWD | FLL-3-1 | S3_35697407 | 3 | 4.293 | 2 | 35,697,407 | 35,753,609 | 56,202 | 5.70% |
| FLL | BOTH | FLL-1-1 | S1_170244 | 1 | 3.767 | 1 | - | - | - | 5.00% |
| FLL | BOTH | FLL-1-3 | S1_31694634 | 1 | 3.623 | 1 | - | - | - | 4.70% |
| FLL | BOTH | FLL-2-2 | S2_15539466 | 2 | 4.607 | 4 | 14,499,656 | 16,059,865 | 1,560,209 | 6.20% |
| FLL | BOTH | FLL-8-1 | S8_15202256-S8_15606417* | 8 | 3.654 | 1 | 15,023,038 | 15,630,935 | 607,897 | 4.80% |
| FLL | BOTH | FLL-9-1 | S9_14598793 | 9 | 4.128 | 24 | 14,598,793 | 15,164,286 | 565,493 | 5.50% |
| FLL | PF | FLL-2-1 | S2_13011786-S2_13026551 | 2 | 4.377 | 2 | 13,011,786 | 13,026,551 | 14,765 | 5.80% |
| FLL | PF | FLL-2-3 | S2_25325801 | 2 | 4.082 | 1 | - | - | - | 5.40% |
| FLW | BOTH | FLW-1-1 | S1_34498117 | 1 | 6.159 | 13 | 34,498,117 | 35,187,397 | 689,280 | 8.60% |
| FLW | BOTH | FLW-4-1 | S4_31148130-S4_31153213 | 4 | 14.951 | 81 | 30,790,218 | 33,083,637 | 2,293,419 | 20.50% |
| FLW | BOTH | FLW-8-1 | S8_25121782 | 8 | 5.79 | 6 | 24,876,484 | 25,209,360 | 332,876 | 7.90% |
| FLW | PF | FLW-1-2 | S1_41251872 | 1 | 4.33 | 1 | - | - | - | 5.80% |
| FLW | PF | FLW-3-1 | S3_12262559-S3_29410266 | 3 | 3.745 | 2 | 12,262,559 | 12,266,766 | 4,207 | 4.90% |
| FLW | PF | FLW-3-2 | S3_29410266 | 3 | 5.577 | 1 | - | - | - | 7.60% |
| LA | AWD | LA-2-2 | S2_16059865 | 2 | 4.595 | 2 | 15,539,466 | 16,059,865 | 520,399 | 6.10% |
| LA | BOTH | LA-2-1 | S2_13011786-S2_13026551 | 2 | 5.252 | 4 | 13,011,786 | 14,613,002 | 1,601,216 | 7.10% |
| LA | BOTH | LA-4-1 | S4_31162467 | 4 | 7.119 | 26 | 30,858,396 | 33,083,637 | 2,225,241 | 9.80% |
| LA | BOTH | LA-7-1 | S7_27661814 | 7 | 4.717 | 2 | 27,577,919 | 27,661,814 | 83,895 | 6.30% |
| LA | PF | LA-1-1 | S1_2073706-S1_2073713 | 1 | 4.433 | 4 | 1,970,058 | 2,073,713 | 103,655 | 5.90% |
| LA | PF | LA-3-1 | S3_3893432-S3_3911954 | 3 | 4.117 | 11 | 3,893,432 | 4,022,997 | 129,565 | 5.50% |
| LA | PF | LA-8-1 | S8_3389674 | 8 | 4.73 | 4 | 3,389,674 | 3,460,692 | 71,018 | 6.40% |
| LA | PF | LA-8-2 | S8_25110888 | 8 | 4.917 | 3 | 25,110,888 | 25,209,360 | 98,472 | 6.90% |
| PH | AWD | PH-4-1 | S4_1967947 | 4 | 3.685 | 1 | - | - | - | 4.90% |
| PH | BOTH | PH-1-2 | S1_36281896-S1_36346718 | 1 | 5.159 | 111 | 36,103,426 | 38,570,119 | 2,466,693 | 7.00% |
| PH | BOTH | PH-2-1 | S2_24151602 | 2 | 4.216 | 2 | 24,151,602 | 27,429,693 | 3,278,091 | 5.60% |
| PH | BOTH | PH-4-2 | S4_30583938 | 4 | 5.27 | 13 | 30,483,967 | 30,613,803 | 129,836 | 7.20% |
| PH | BOTH | PH-5-1 | S5_17929906 | 5 | 3.919 | 1 | - | - | - | 5.20% |
| PH | BOTH | PH-6-1 | S6_22464644 | 6 | 4.919 | 2 | 22,464,644 | 22,489,397 | 24,753 | 6.60% |
| PH | BOTH | PH-6-2 | S6_24473578- | 6 | 3.743 | 1 | - | - | - | 4.90% |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| PH | PF | PH-1-1 | S1_170244 | 1 | 5.035 | 2 | 170,244 | 330,484 | 160,240 | 6.90% |
| PL | AWD | PL-2-1 | S2_8358936 | 2 | 4.858 | 2 | 8,358,936 | 8,591,671 | 232,735 | 6.60% |
| PL | AWD | PL-4-1 | S4_30582194 | 4 | 4.606 | 1 | - | - | - | 6.20% |
| PL | AWD | PL-6-1 | S6_13545500-S6_13588590 | 6 | 3.813 | 2 | 13,545,500 | 13,588,590 | 43,090 | 5.00% |
| PL | AWD | PL-9-1 | S9_15126957 | 9 | 4.038 | 2 | 15,011,225 | 15,126,957 | 115,732 | 5.40% |
| PL | AWD | PL-12-1 | S12_11206556 | 12 | 4.164 | 1 | - | - | - | 5.50% |
| PL | BOTH | PL-1-1 | S1_1059390 | 1 | 4.299 | 2 | 1,059,390 | 1,062,977 | 3,587 | 6.00% |
| PL | BOTH | PL-2-2 | S2_25831716 | 2 | 4.66 | 2 | 25,824,470 | 25,831,716 | 7,246 | 6.30% |
| PL | PF | PL-8-1 | S8_24876447 | 8 | 3.68 | 2 | 24,871,328 | 24,876,447 | 5,119 | 4.90% |
| PNH | AWD | PNH-1-2 | S1_24985034 | 1 | 3.71 | 1 | - | - | - | 4.90% |
| PNH | AWD | PNH-2-1 | S2_5785532 | 2 | 3.654 | 1 | - | - | - | 4.80% |
| PNH | AWD | PNH-4-1 | S4_30583938 | 4 | 4.898 | 13 | 30,483,967 | 30,613,803 | 129,836 | 6.60% |
| PNH | BOTH | PNH-1-1 | S1_170244 | 1 | 4.879 | 2 | 170,244 | 330,484 | 160,240 | 6.60% |
| PNH | BOTH | PNH-1-3 | S1_38457496 | 1 | 6.33 | 156 | 35,219,307 | 38,723,381 | 3,504,074 | 8.70% |
| PNH | BOTH | PNH-2-2 | S2_24151602 | 2 | 5.166 | 1 | - | - | - | 7.10% |
| PNH | BOTH | PNH-2-3 | S2_27429693 | 2 | 3.964 | 1 | - | - | - | 5.20% |
| PNH | BOTH | PNH-3-1 | S3_1248941 | 3 | 3.665 | 1 | - | - | - | 4.80% |
| PNH | BOTH | PNH-5-1 | S5_17929906 | 5 | 3.84 | 3 | 17,665,483 | 17,963,918 | 298,435 | 5.10% |
| PNH | BOTH | PNH-6-1 | S6_22464644 | 6 | 4.743 | 2 | 22,464,644 | 22,489,397 | 24,753 | 6.40% |
| PNH | BOTH | PNH-6-2 | S6_24473578 | 6 | 3.797 | 1 | - | - | - | 5.00% |

*ANNEX 7e: MTAs for phenological traits.*

| Trait | Condition | Association ID | Peak Marker/Region | Chr. | -log10(p) | SNPs | Associated region | | | Peak marker $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Start | End | Size | |
| DF | BOTH | DF-2-2 | S2_23232026 | 2 | 4.146 | 3 | 23,232,026 | 23,269,201 | 37,175 | 5.80% |
| DF | BOTH | DF-7-2 | S7_27784632 | 7 | 4.155 | 1 | - | - | - | 6.20% |
| DF | PF | DF-2-1 | S2_8591671 | 2 | 3.674 | 1 | - | - | - | 5.00% |
| DF | PF | DF-7-1 | S7_20912737 | 7 | 3.93 | 1 | - | - | - | 5.80% |
| DF | PF | DF-10-1 | S10_17323647 | 10 | 4.508 | 6 | 17,323,647 | 17,972,696 | 649,049 | 6.80% |
| DFM | AWD | DFM-10-3 | S10_21139023 | 10 | 3.926 | 1 | - | - | - | 5.30% |
| DFM | AWD | DFM-11-1 | S11_5403023 | 11 | 4.024 | 11 | 4,901,549 | 5,403,023 | 501,474 | 5.40% |
| DFM | AWD | DFM-11-2 | S11_6238153 | 11 | 3.943 | 1 | - | - | - | 5.20% |
| DFM | BOTH | DFM-3-2 | S3_11731550-S3_11731575 | 3 | 5.04 | 20 | 11,579,440 | 11,752,044 | 172,604 | 6.90% |
| DFM | PF | DFM-3-1 | S3_5369580 | 3 | 5.65 | 12 | 5,274,774 | 5,870,932 | 596,158 | 7.80% |
| DFM | PF | DFM-8-1 | S8_8220521 | 8 | 3.716 | 1 | - | - | - | 4.90% |
| DFM | PF | DFM-10-1 | S10_11841037 | 10 | 4.2 | 2 | 11,819,506 | 11,841,037 | 21,531 | 5.60% |
| DFM | PF | DFM-10-2 | S10_15368426 | 10 | 5.693 | 2 | 15,368,426 | 15,473,611 | 105,185 | 7.90% |
| DM | AWD | DM-4-2 | S4_27422010 | 4 | 3.844 | 1 | - | - | - | 5.10% |
| DM | AWD | DM-10-1 | S10_11152024 | 10 | 3.814 | 1 | - | - | - | 5.00% |
| DM | BOTH | DM-7-2 | S7_26728099 | 7 | 6.194 | 6 | 26,606,039 | 27,784,632 | 1,178,593 | 8.60% |
| DM | BOTH | DM-10-2 | S10_15368426 | 10 | 8.148 | 2 | 15,368,426 | 15,473,611 | 105,185 | 11.40% |
| DM | BOTH | DM-10-3 | S10_17323647 | 10 | 5.157 | 4 | 17,323,647 | 17,437,446 | 113,799 | 7.10% |
| DM | PF | DM-2-1 | S2_24315759 | 2 | 3.737 | 2 | 24,315,759 | 24,494,645 | 178,886 | 4.90% |
| DM | PF | DM-4-1 | S4_24756463 | 4 | 4.097 | 1 | - | - | - | 5.50% |
| DM | PF | DM-7-1 | S7_20912737 | 7 | 4.678 | 1 | - | - | - | 6.40% |
| DM | PF | DM-8-1 | S8_17809885 | 8 | 4.398 | 2 | 17,809,885 | 17,861,062 | 51,177 | 5.90% |

*ANNEX 8a: MTAs for macro elements (P and K).*

| Trait | Condition | Association ID | Peack Marker/Region | Chr | -log10(p) | SNPs | Associated region | | | markerR2 |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Start | End | Size | |
| K | AWD | K_8_1 | S8_9177046 | 8 | 4.07 | 3 | 9,177,046 | 9,194,662 | 17,616 | 5.42% |
| K | AWD | K_8_2 | S8_11025075 | 8 | 3.76 | 1 | . | . | . | 5.08% |
| K | AWD | K_9_1 | S9_2235559 | 9 | 5.00 | 1 | . | . | . | 6.82% |
| K | AWD | K_11_1 | S11_17295781 | 11 | 3.92 | 8 | 17,295,781 | 17,314,803 | 19,022 | 5.19% |
| K | PF | K_3_1 | S3_11168887 | 3 | 3.84 | 5 | 1,1168,887 | 11,184,504 | 15,617 | 5.08% |
| K | PF | K_10_1 | S10_12414990 | 10 | 3.90 | 1 | . | . | . | 5.16% |
| K | PF | K_12_1 | S12_24169856 | 12 | 3.73 | 1 | . | . | . | 4.92% |
| P | AWD | P_2_2 | S2_9777633 | 2 | 4.27 | 2 | 9,777,584 | 9,777,633 | 49 | 5.69% |
| P | AWD | P_2_1 | S2_9612122 | 2 | 3.85 | 1 | . | . | . | 5.06% |
| P | AWD | P_4_3 | S4_27790754 | 4 | 5.61 | 2 | 27,790,754 | 27790806 | 52 | 7.69% |
| P | AWD | P_8_1 | S8_8367471 | 8 | 4.43 | 2 | 8,367,471 | 8,422,304 | 54,833 | 5.94% |
| P | AWD | P_8_2 | S8_8563018 | 8 | 3.98 | 1 | . | . | . | 5.28% |
| P | AWD | P_11_1 | S11_24811815 | 11 | 3.81 | 1 | . | . | . | 5.01% |
| P | BOTH | P_4_1 and 2 | S4_27790754 | 4 | 5.61 | 2 | 27,790,754 | 27,790,806 | 52 | 7.69% |
| P | PF | P_3_1 | S3_9196982 | 3 | 6.38 | 4 | 9,195,427 | 9,196,982 | 1,555 | 8.79% |
| P | PF | P_6_1 | S6_25016747 | 6 | 4.29 | 2 | 25,016,747 | 25,095,178 | 78,431 | 5.75% |
| P | PF | P_7_1 | S7_20915767 | 7 | 4.94 | 3 | 20,915,752 | 20,915,772 | 20 | 6.71% |
| P | PF | P_7_2 | S7_22775107 | 7 | 4.12 | 1 | . | . | . | 5.49% |
| P | PF | P_10_1 | S10_12414990 | 10 | 5.87 | 1 | . | . | . | 8.08% |
| P | PF | P_10_2 | S10_17179400 | 10 | 4.27 | 1 | . | . | . | 5.68% |

*ANNEX 8b: MTAs for macro elements (Mg and Ca).*

| Trait | Condition | Association ID | Peack Marker/Region | Chr | -log10(p) | SNPs | Associated region | | | markerR2 |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Start | End | Size | |
| Ca | AWD | Ca_3_2 | S3_27768236 | 3 | 5.90 | 27 | 27,717,874 | 28,239,870 | 521,996 | 8.12% |
| Ca | AWD | Ca_3_1 | S3_16408420 | 3 | 3.58 | 1 | . | . | . | 4.69% |
| Ca | AWD | Ca_4_3 | S4_31497354 | 4 | 3.71 | 1 | . | . | . | 4.64% |
| Ca | AWD | Ca_4_2 | S4_33597687 | 4 | 3.63 | 1 | . | . | . | 4.75% |
| Ca | AWD | Ca_5_1 | S5_26991666 | 5 | 4.35 | 2 | 26,991,666 | 27,026,510 | 34,844 | 5.83% |
| Ca | AWD | Ca_6_1 | S6_1403579 | 6 | 3.60 | 1 | . | . | . | 4.49% |
| Ca | AWD | Ca_10_1 | S10_7579576 | 10 | 4.02 | 1 | . | . | . | 5.36% |
| Ca | BOTH | Ca_4_1 | S4_31068491 | 4 | 4.92 | 7 | 31,050,939 | 3,116,2467 | 111,528 | 6.63% |
| Ca | BOTH | Ca_4_4 | S4_31393661 | 4 | 5.52 | 2 | 31,390,637 | 31,393,661 | 3,024 | 7.54% |
| Ca | PF | Ca_4_6 | S4_32365853 | 4 | 3.65 | 2 | 32,260,467 | 32,365,853 | 105,386 | 4.77% |
| Ca | PF | Ca_4_7 | S4_33052629 | 4 | 3.78 | 5 | 33,052,629 | 33,114,563 | 6,1934 | 4.94% |
| Ca | PF | Ca_4_5 | S4_1216012 | 4 | 3.67 | 1 | . | . | . | 4.94% |
| Ca | PF | Ca_5_2 | S5_578074 | 5 | 3.69 | 1 | . | . | . | 4.86% |
| Ca | PF | Ca_6_2 | S6_21883579 | 6 | 4.10 | 1 | . | . | . | 5.43% |
| Ca | PF | Ca_10_2 | S10_2329639 | 10 | 3.61 | 3 | 2,329,639 | 2,329,658 | 19 | 4.74% |
| Mg | AWD | Mg_4_2 | S4_27790806 | 4 | 5.24 | 2 | 27,790,754 | 27,790,806 | 52 | 7.17% |
| Mg | AWD | Mg_4_1 | S4_27790806 | 4 | 5.24 | 2 | 27,790,754 | 27,790,806 | 52 | 7.17% |
| Mg | AWD | Mg_4_3 | S4_29940409 | 4 | 5.48 | 2 | 29,921,826 | 29,940,409 | 18583 | 7.51% |
| Mg | AWD | Mg_4_4 | S4_24155553 | 4 | 4.36 | 1 | . | . | . | 5.84% |
| Mg | PF | Mg_3_1 | S3_9196982 | 3 | 6.08 | 4 | 9,195,427 | 9,196,982 | 1555 | 8.34% |
| Mg | PF | Mg_3_2 | S3_16495531 | 3 | 4.30 | 2 | 16,495,531 | 16,495,789 | 258 | 5.73% |
| Mg | PF | Mg_4_5 | S4_27790806 | 4 | 4.12 | 2 | 27,790,754 | 27,790,806 | 52 | 5.50% |
| Mg | PF | Mg_9_1 | S9_17196410 | 9 | 3.73 | 2 | 17,196,410 | 17,196,415 | 5 | 4.89% |
| Mg | PF | Mg_10_1 | S10_2999963 | 10 | 3.69 | 1 | . | . | . | 5.08% |
| Mg | PF | Mg_10_2 | S10_12414990 | 10 | 5.01 | 1 | . | . | . | 6.81% |

*ANNEX 8c: MTAs for macro elements.*

| Trait | Condition | Association ID | Peack Marker/Region | Chr | -log10(p) | SNPs | Associated region | | | markerR2 |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Start | End | Size | |
| Cu | AWD | Cu_1_3 | S1_42414831 | 1 | 3.84 | 2 | 42,412,011 | 42,414,831 | 2,820 | 5.04% |
| Cu | AWD | Cu_1_1 | S1_23818658 | 1 | 3.66 | 1 | . | . | . | 4.86% |
| Cu | AWD | Cu_1_2 | S1_24808821 | 1 | 3.77 | 1 | . | . | . | 4.93% |
| Cu | AWD | Cu_12_1 | S12_24169856 | 12 | 3.77 | 1 | . | . | . | 4.97% |
| Cu | PF | Cu_3_1 | S3_9195427 | 3 | 4.13 | 2 | 9,195,427 | 9,196,982 | 1,555 | 5.50% |
| Cu | PF | Cu_7_3 | S7_20901517 | 7 | 4.17 | 2 | 20,901,229 | 20,901,517 | 288 | 5.57% |
| Cu | PF | Cu_7_4 | S7_20915767 | 7 | 4.16 | 5 | 2,0912,737 | 2,0915,822 | 3,085 | 5.55% |
| Cu | PF | Cu_7_5 | S7_20919987 | 7 | 3.98 | 6 | 20,919,987 | 2,1151,831 | 231,844 | 5.29% |
| Cu | PF | Cu_7_6 | S7_21235989 | 7 | 4.10 | 5 | 21,235,989 | 21,351,329 | 115,340 | 5.47% |
| Cu | PF | Cu_7_7 | S7_21440075 | 7 | 3.81 | 2 | 21,440,075 | 21,440,123 | 48 | 5.04% |
| Cu | PF | Cu_7_1 | S7_11035121 | 7 | 3.82 | 1 | . | . | . | 5.08% |
| Cu | PF | Cu_7_2 | S7_20693738 | 7 | 3.81 | 1 | . | . | . | 5.04% |
| Cu | PF | Cu_10_1 | S10_17323647 | 10 | 4.26 | 1 | . | . | . | 5.73% |
| Cu | PF | Cu_12_2 | S12_16452855 | 12 | 3.76 | 1 | . | . | . | 4.95% |
| Fe | AWD | Fe_1_1 | S1_30875567 | 1 | 4.02 | 1 | . | . | . | 5.32% |
| Fe | AWD | Fe_2_1 | S2_9777584 | 2 | 3.82 | 1 | . | . | . | 5.02% |
| Fe | AWD | Fe_8_1 | S8_8422304 | 8 | 4.54 | 2 | 8,367,471 | 8,422,304 | 54,833 | 6.11% |
| Fe | AWD | Fe_8_4 | S8_19591578 | 8 | 4.79 | 2 | 19,586,564 | 19,591,578 | 5,014 | 6.48% |
| Fe | AWD | Fe_8_2 | S8_8563018 | 8 | 4.99 | 1 | . | . | . | 6.78% |
| Fe | AWD | Fe_8_3 | S8_9447691 | 8 | 3.70 | 1 | . | . | . | 4.85% |
| Fe | AWD | Fe_11_1 | S11_23022078 | 11 | 4.887844 | 1 | . | . | . | 6.60% |
| Fe | PF | Fe_11_3 | S11_19770890 | 11 | 4.05 | 4 | 19,770,838 | 19,770,910 | 72 | 5.57% |
| Fe | PF | Fe_11_2 | S11_19748354 | 11 | 4.48 | 1 | . | . | . | 6.01% |
| Mn | AWD | Mn_11_4 | S11_24811815 | 11 | 4.27 | 1 | . | . | . | 5.69% |
| Mn | AWD | Mn_3_1 | S3_1434048 | 3 | 3.70 | 3 | 1,434,048 | 1,454,153 | 20,105 | 4.91% |
| Mn | AWD | Mn_3_2 | S3_8193361 | 3 | 3.75 | 3 | 8,193,361 | 8,198,376 | 5,015 | 4.95% |
| Mn | AWD | Mn_8_1 | S8_19594864 | 8 | 3.83 | 1 | . | . | . | 5.04% |
| Mn | AWD | Mn_11_1 | S11_21817552 | 11 | 4.88 | 1 | . | . | . | 6.60% |
| Mn | PF | Mn_4_1 | S4_27790754 | 4 | 4.20 | 1 | . | . | . | 5.60% |
| Na | AWD | Na_1_2 | S1_36963814 | 1 | 3.80 | 2 | 36,963,814 | 37,196,622 | 232,808 | 4.99% |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Na | AWD | Na_1_1 | S1_29879401 | 1 | 3.93 | 1 | . | . | . | 5.19% |
| Na | AWD | Na_1_3 | S1_40612679 | 1 | 4.14 | 1 | . | . | . | 5.56% |
| Na | AWD | Na_2_1 | S2_25007547 | 2 | 3.72 | 1 | . | . | . | 4.89% |
| Na | AWD | Na_4_1 | S4_18971731 | 4 | 3.59 | 1 | . | . | . | 4.69% |
| Na | AWD | Na_8_1 | S8_21801468 | 8 | 3.97 | 4 | 2,1801,468 | 21,801,588 | 120 | 5.24% |
| Na | AWD | Na_11_1 | S11_20264992 | 11 | 4.55 | 2 | 20,116,136 | 2,0264,992 | 148,856 | 6.11% |
| Na | PF | Na_2_2 | S2_11180804 | 2 | 3.76 | 1 | . | . | . | 4.93% |
| Na | PF | Na_3_1 | S3_12262559 | 3 | 3.65 | 2 | 12,262,559 | 12,266,766 | 4,207 | 4.78% |
| Na | PF | Na_7_1 | S7_28562982 | 7 | 3.68 | 1 | . | . | . | 4.83% |
| Na | PF | Na_11_2 | S11_19612704 | 11 | 4.46 | 14 | 19,577,792 | 19,625,046 | 47,254 | 5.96% |
| Na | PF | Na_11_3 | S11_19948652 | 11 | 3.70 | 1 | . | . | . | 4.84% |
| Ni | AWD | Ni_1_3 | S1_37839106 | 1 | 4.20 | 1 | . | . | . | 5.60% |
| Ni | AWD | Ni_1_4 | S1_42470312 | 1 | 4.83 | 1 | . | . | . | 6.41% |
| Ni | AWD | Ni_4_1 | S4_27781190 | 4 | 3.75 | 1 | . | . | . | 4.96% |
| Ni | AWD | Ni_8_1 | S8_116363 | 8 | 4.05 | 1 | . | . | . | 5.29% |
| Ni | AWD | Ni_9_1 | S9_16580832 | 9 | 3.63 | 1 | . | . | . | 4.79% |
| Ni | AWD | Ni_11_2 | S11_17388473 | 11 | 5.00 | 10 | 17,348,721 | 1,7389,068 | 40,347 | 6.80% |
| Ni | AWD | Ni_11_1 | S11_13868486 | 11 | 4.38 | 1 | . | . | . | 5.90% |
| Ni | BOTH | Ni_1_1 | S1_42414831 | 1 | 5.03 | 2 | 42,412,011 | 42,414,831 | 2,820 | 6.81% |
| Ni | BOTH | Ni_1_2 | S1_42420861 | 1 | 5.05 | 2 | 42,420,853 | 42,420,861 | 8 | 6.84% |
| Ni | PF | Ni_1_5 | S1_18637785 | 1 | 4.47 | 4 | 18,622,685 | 18,637,785 | 15,100 | 6.04% |
| Ni | PF | Ni_1_6 | S1_25004276 | 1 | 7.52 | 9 | 25,004,276 | 25,150,609 | 146,333 | 10.48% |
| Ni | PF | Ni_1_7 | S1_29076045 | 1 | 4.64 | 2 | 29,076,045 | 29,107,159 | 31,114 | 6.28% |
| Ni | PF | Ni_1_8 | S1_29909662 | 1 | 6.06 | 12 | 29,829,782 | 29,909,662 | 79,880 | 8.65% |
| Ni | PF | Ni_1_9 | S1_34512840 | 1 | 4.28 | 5 | 34,512,840 | 34,602,303 | 89,463 | 5.71% |
| Ni | PF | Ni_1_10 | S1_39308449 | 1 | 3.60 | 1 | . | . | . | 4.95% |
| Ni | PF | Ni_1_11 | S1_39582907 | 1 | 5.93 | 1 | . | . | . | 8.20% |
| Ni | PF | Ni_2_3 | S2_23103553 | 2 | 5.55 | 3 | 23,103,553 | 23,153,609 | 50,056 | 7.60% |
| Ni | PF | Ni_2_1 | S2_9921223 | 2 | 3.90 | 1 | . | . | . | 5.15% |
| Ni | PF | Ni_2_2 | S2_10216171 | 2 | 4.85 | 1 | . | . | . | 6.55% |
| Ni | PF | Ni_2_4 | S2_23559757 | 2 | 4.23 | 1 | . | . | . | 5.65% |
| Ni | PF | Ni_2_5 | S2_26126543 | 2 | 4.05 | 1 | . | . | . | 5.38% |
| Ni | PF | Ni_3_1 | S3_11540501 | 3 | 5.37 | 2 | 11,540,501 | 11,540,516 | 15 | 7.34% |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Ni | PF | Ni_3_4 | S3_27474350 | 3 | 4.21 | 3 | 27,474,350 | 27,509,648 | 35,298 | 5.89% |
| Ni | PF | Ni_3_2 | S3_12446399 | 3 | 3.59 | 1 | . | . | . | 4.69% |
| Ni | PF | Ni_3_3 | S3_26775536 | 3 | 3.83 | 1 | . | . | . | 5.05% |
| Ni | PF | Ni_3_5 | S3_28096677 | 3 | 4.35 | 1 | . | . | . | 5.89% |
| Ni | PF | Ni_4_2 | S4_1216012 | 4 | 4.84 | 2 | 1,216,012 | 1,221,362 | 5,350 | 6.73% |
| Ni | PF | Ni_4_3 | S4_1727665 | 4 | 4.93 | 2 | 1,727,665 | 1,811,257 | 83,592 | 6.66% |
| Ni | PF | Ni_4_5 | S4_5005781 | 4 | 5.80 | 3 | 5,005,639 | 500,5781 | 142 | 8.00% |
| Ni | PF | Ni_4_7 | S4_16642304 | 4 | 5.24 | 3 | 16,642,304 | 16,672,333 | 30,029 | 7.15% |
| Ni | PF | Ni_4_4 | S4_4924330 | 4 | 4.63 | 1 | . | . | . | 6.23% |
| Ni | PF | Ni_4_6 | S4_13928968 | 4 | 5.13 | 1 | . | . | . | 6.96% |
| Ni | PF | Ni_5_1 | S5_3623956 | 5 | 4.00 | 2 | 3,623,956 | 3,623,970 | 14 | 5.34% |
| Ni | PF | Ni_5_2 | S5_6132287 | 5 | 4.47 | 2 | 6,132,287 | 6,191,511 | 59,224 | 6.01% |
| Ni | PF | Ni_5_3 | S5_17237176 | 5 | 3.84 | 3 | 17,235,931 | 17,252,113 | 16,182 | 5.06% |
| Ni | PF | Ni_5_5 | S5_29485595 | 5 | 4.47 | 3 | 29,485,595 | 29,502,947 | 17,352 | 6.01% |
| Ni | PF | Ni_5_4 | S5_29205256 | 5 | 4.19 | 1 | . | . | . | 5.59% |
| Ni | PF | Ni_6_1 | S6_1349980 | 6 | 3.92 | 2 | 1,349,980 | 1,445,015 | 95,035 | 5.18% |
| Ni | PF | Ni_6_2 | S6_4926766 | 6 | 5.25 | 2 | 4,707,063 | 4,926,766 | 219,703 | 7.19% |
| Ni | PF | Ni_6_6 | S6_21703262 | 6 | 4.01 | 2 | 21,700,793 | 21,703,262 | 2,469 | 5.49% |
| Ni | PF | Ni_6_7 | S6_21750763 | 6 | 5.52 | 6 | 21,750,758 | 21,883,359 | 132,601 | 7.56% |
| Ni | PF | Ni_6_8 | S6_23410413 | 6 | 4.09 | 2 | 23,410,413 | 23,410,419 | 6 | 5.43% |
| Ni | PF | Ni_6_3 | S6_9667114 | 6 | 3.78 | 1 | . | . | . | 4.98% |
| Ni | PF | Ni_6_4 | S6_10279704 | 6 | 4.49 | 1 | . | . | . | 6.09% |
| Ni | PF | Ni_6_5 | S6_10791923 | 6 | 3.88 | 1 | . | . | . | 5.16% |
| Ni | PF | Ni_6_9 | S6_28285046 | 6 | 4.89 | 1 | . | . | . | 6.66% |
| Ni | PF | Ni_7_2 | S7_14937015 | 7 | 4.37 | 2 | 14,937,015 | 15,078,504 | 141,489 | 5.84% |
| Ni | PF | Ni_7_1 | S7_179912 | 7 | 4.19 | 1 | . | . | . | 5.58% |
| Ni | PF | Ni_7_3 | S7_15853917 | 7 | 3.89 | 1 | . | . | . | 5.13% |
| Ni | PF | Ni_7_4 | S7_21562734 | 7 | 4.56 | 1 | . | . | . | 6.12% |
| Ni | PF | Ni_7_5 | S7_24372816 | 7 | 4.02 | 1 | . | . | . | 5.32% |
| Ni | PF | Ni_8_3 | S8_5275795 | 8 | 4.24 | 2 | 5,275,364 | 5,275,795 | 431 | 5.67% |
| Ni | PF | Ni_8_2 | S8_2573992 | 8 | 4.47 | 1 | . | . | . | 5.99% |
| Ni | PF | Ni_8_4 | S8_26817327 | 8 | 3.78 | 1 | . | . | . | 5.00% |
| Ni | PF | Ni_9_2 | S9_416567 | 9 | 5.38 | 4 | 287,389 | 416,567 | 129,178 | 7.44% |

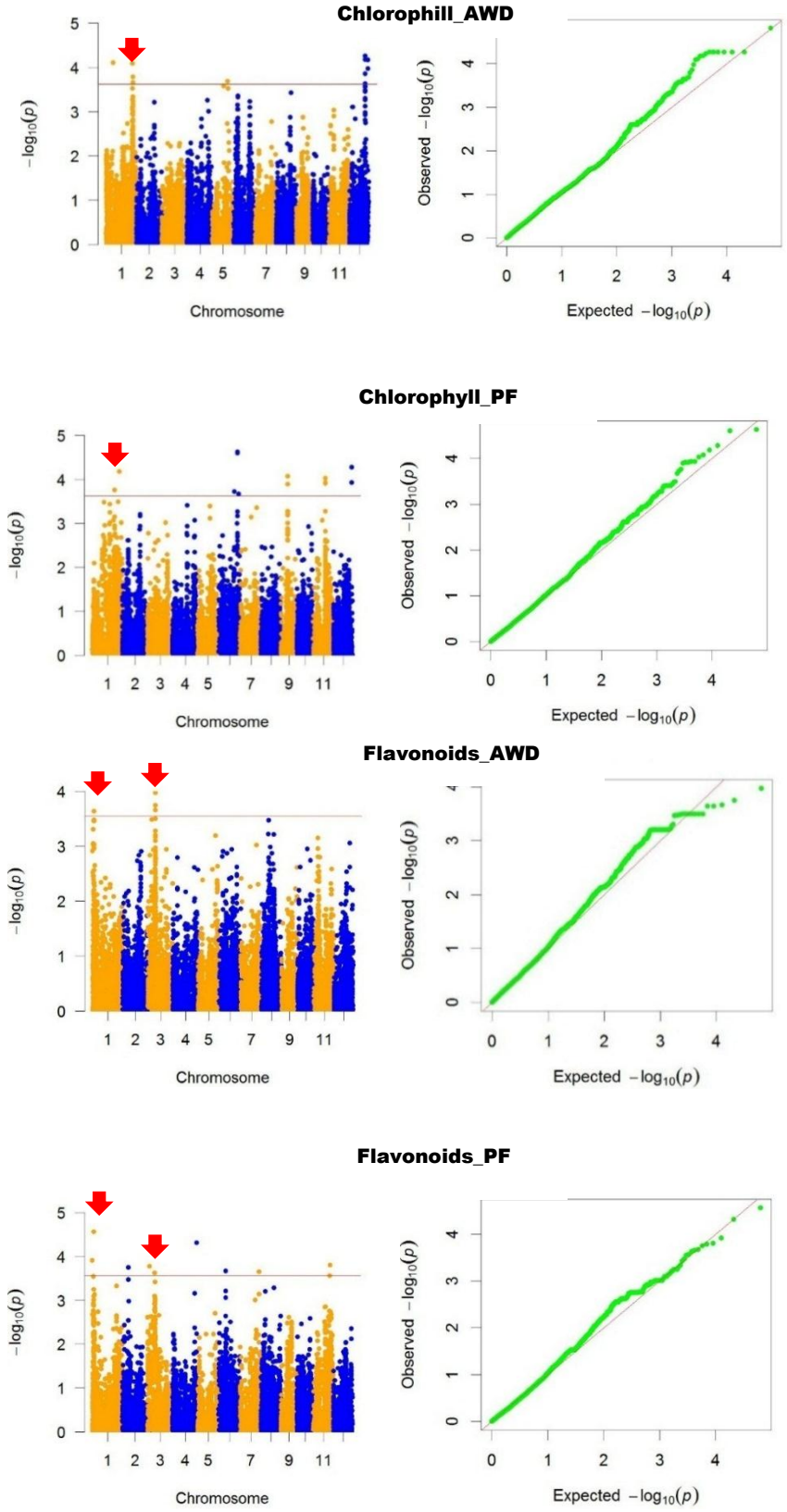
| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Ni | PF | Ni_9_3 | S9_9418065 | 9 | 3.72 | 1 | . | . | . | 4.88% |
| Ni | PF | Ni_9_4 | S9_9498065 | 9 | 5.37 | 1 | . | . | . | 7.31% |
| Ni | PF | Ni_9_5 | S9_10047397 | 9 | 3.78 | 1 | . | . | . | 5.01% |
| Ni | PF | Ni_10_3 | S10_3374071 | 10 | 6.27 | 8 | 3183428 | 3443158 | 259730 | 8.67% |
| Ni | PF | Ni_10_1 | S10_2999963 | 10 | 5.62 | 1 | . | . | . | 8.06% |
| Ni | PF | Ni_10_2 | S10_3024046 | 10 | 3.86 | 1 | . | . | . | 5.08% |
| Ni | PF | Ni_10_4 | S10_10161394 | 10 | 3.89 | 1 | . | . | . | 5.15% |
| Ni | PF | Ni_11_3 | S11_232528 | 11 | 3.90 | 2 | 232,528 | 250,151 | 17,623 | 5.16% |
| Ni | PF | Ni_11_4 | S11_3854031 | 11 | 3.94 | 21 | 3,689,937 | 3,8678,01 | 177,864 | 5.22% |
| Ni | PF | Ni_11_7 | S11_21855636 | 11 | 4.52 | 3 | 21,855,636 | 21,862,549 | 6,913 | 6.07% |
| Ni | PF | Ni_11_8 | S11_26088884 | 11 | 4.49 | 5 | 26,088,884 | 26,090,266 | 1382 | 6.02% |
| Ni | PF | Ni_11_5 | S11_17770399 | 11 | 4.15 | 1 | . | . | . | 5.52% |
| Ni | PF | Ni_11_6 | S11_20264992 | 11 | 3.66 | 1 | . | . | . | 4.81% |
| Ni | PF | Ni_12_1 | S12_6507018 | 12 | 3.58 | 4 | 6,507,018 | 650,7051 | 33 | 4.73% |
| Ni | PF | Ni_12_2 | S12_10636529 | 12 | 7.37 | 15 | 10,636,529 | 10,757,872 | 121,343 | 10.73% |
| Ni | PF | Ni_12_3 | S12_10761232 | 12 | 5.92 | 5 | 10,761,232 | 10,965,388 | 204,156 | 8.12% |
| Ni | PF | Ni_12_5 | S12_12806823 | 12 | 5.07 | 2 | 12,806,823 | 12,836,285 | 29,462 | 6.95% |
| Ni | PF | Ni_12_4 | S12_10996127 | 12 | 4.85 | 1 | . | . | . | 6.55% |
| Ni | PF | Ni_12_6 | S12_16331904 | 12 | 5.50 | 1 | . | . | . | 7.53% |
| Zn | AWD | Zn_2_1 | S2_24941038 | 2 | 5.32 | 2 | 24,941,038 | 24,962,679 | 21,641 | 7.24% |
| Zn | AWD | Zn_8_1 | S8_24914024 | 8 | 3.60 | 1 | . | . | . | 4.71% |
| Zn | AWD | Zn_10_1 | S10_4543916 | 10 | 3.60 | 2 | 4,543,916 | 4,543,968 | 52 | 4.68% |
| Zn | AWD | Zn_11_1 | S11_26877968 | 11 | 3.71 | 2 | 26,877,968 | 26,917,205 | 39,237 | 4.87% |
| Zn | BOTH | Zn_8_2 | S8_4099221 | 8 | 4.31 | 3 | 4,096,930 | 4,101,083 | 4,153 | 5.73% |
| Zn | PF | Zn_3_1 | S3_9196982 | 3 | 3.88 | 2 | 9,195,427 | 9,196,982 | 1,555 | 5.10% |
| Zn | PF | Zn_9_1 | S9_17196410 | 9 | 3.61 | 9 | 17,150,400 | 17,196,415 | 46,015 | 4.69% |
| Zn | PF | Zn_11_2 | S11_23942716 | 11 | 3.76 | 1 | . | . | . | 4.93% |

*ANNEX 8d: MTAs for macro undesired trace elements .*

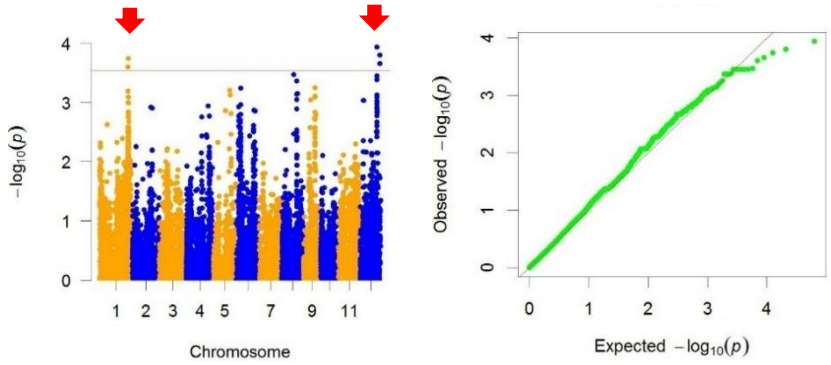
| Trait | Condition | Association ID | Peack Marker/Region | Chr | -log10(p) | SNPs | Associated region | | | markerR2 |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Start | End | Size | |
| As | AWD | As_2_1 | S2_2699366 | 2 | 3.82 | 1 | . | . | . | 5.02% |
| As | AWD | As_7_1 | S7_21235989 | 7 | 4.24 | 1 | . | . | . | 5.67% |
| As | AWD | As_10_1 | S10_17255433 | 10 | 5.86 | 4 | 17,255,433 | 17,378,597 | 123,164 | 8.09% |
| As | AWD | As_12_1 | S12_21791655 | 12 | 3.90 | 1 | . | . | . | 5.15% |
| As | PF | As_1_1+C3:C35 | S1_9603926 | 1 | 3.79 | 2 | 9,478,659 | 9,603,926 | 125,267 | 4.99% |
| As | PF | As_12_2 | S12_22194708 | 12 | 3.68 | 2 | 22,194,708 | 22,194,797 | 89 | 4.80% |
| Cd | AWD | Cd_1_1 | S1_2242781 | 1 | 4.15 | 8 | 2,242,781 | 2,263,709 | 20,928 | 5.54% |
| Cd | AWD | Cd_1_2 | S1_25290805 | 1 | 4.47 | 2 | 25,290,805 | 25,535,193 | 244,388 | 6.01% |
| Cd | AWD | Cd_2_1 | S2_27706195 | 2 | 4.12 | 2 | 27,706,195 | 27,739,329 | 33,134 | 5.48% |
| Cd | AWD | Cd_2_2 | S2_34741156 | 2 | 5.37 | 17 | 34,741,156 | 34,981,229 | 240,073 | 7.34% |
| Cd | AWD | Cd_3_2 | S3_16665467 | 3 | 6.42 | 3 | 16,659,994 | 1,666,5481 | 5,487 | 8.12% |
| Cd | AWD | Cd_3_1 | S3_6000810 | 3 | 4.72 | 1 | . | . | . | 6.41% |
| Cd | AWD | Cd_4_1 | S4_3238224 | 4 | 5.53 | 1 | . | . | . | 7.58% |
| Cd | AWD | Cd_7_1 | S7_57174 | 7 | 4.10 | 4 | 57,174 | 117,184 | 60,010 | 5.45% |
| Cd | AWD | Cd_7_2 | S7_18475741 | 7 | 4.07 | 2 | 18,475,741 | 18,475,875 | 134 | 5.43% |
| Cd | AWD | Cd_7_3 | S7_29552445 | 7 | 5.43 | 30 | 29,082,918 | 29,552,445 | 469,527 | 7.44% |
| Cd | AWD | Cd_7_4 | S7_29585802 | 7 | 5.43 | 12 | 29,580,230 | 29,658,494 | 78,264 | 7.44% |
| Cd | AWD | Cd_8_2 | S8_663578 | 8 | 3.54 | 1 | . | . | . | 4.62% |
| Cd | AWD | Cd_10_1 | S10_3553438 | 10 | 4.00 | 1 | . | . | . | 4.86% |
| Cd | AWD | Cd_11_1 | S11_24387741 | 11 | 4.13 | 1 | . | . | . | 5.48% |
| Cd | AWD | Cd_12_2 | S12_18353718 | 12 | 3.93 | 1 | . | . | . | 5.20% |
| Cd | BOTH | Cd_2_3 | S2_35099274 | 2 | 4.41 | 2 | 35,099,274 | 35,099,287 | 13 | 5.89% |
| Cd | BOTH | Cd_8_1 | S8_14976342 | 8 | 4.92 | 3 | 14,616,713 | 14,976,342 | 359,629 | 6.66% |
| Cd | BOTH | Cd_8_3 | S8_15319123 | 8 | 4.48 | 4 | 15,023,064 | 15,319,159 | 296,095 | 6.00% |
| Cd | PF | Cd_3_3 | S3_27021920 | 3 | 4.47 | 12 | 27,021,920 | 27,084,459 | 62,539 | 5.77% |
| Cd | PF | Cd_3_4 | S3_27161921 | 3 | 5.87 | 3 | 27,161,921 | 27,213,980 | 52,059 | 5.80% |
| Cd | PF | Cd_8_4 | S8_13653507 | 8 | 3.86 | 1 | . | . | . | 5.10% |
| Cd | PF | Cd_9_1 | S9_9777662 | 9 | 7.85 | 4 | 9,777,508 | 9,781,906 | 4,398 | 10.85% |
| Cd | PF | Cd_9_2 | S9_9867705 | 9 | 5.81 | 3 | 9,867,705 | 9,922,364 | 54,659 | 7.94% |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Cd | PF | Cd_9_3 | S9_10120723 | 9 | 4.60 | 3 | 9,979,504 | 10,120,723 | 141,219 | 6.16% |
| Cd | PF | Cd_9_4 | S9_19547273 | 9 | 5.04 | 4 | 19,419,424 | 19,564,391 | 144,967 | 7.14% |
| Cd | PF | Cd_11_2 | S11_21795392 | 11 | 3.71 | 2 | 21,795,392 | 21,800,488 | 5,096 | 4.90% |
| Cd | PF | Cd_11_3 | S11_24210972 | 11 | 3.86 | 1 | . | . | . | 5.10% |

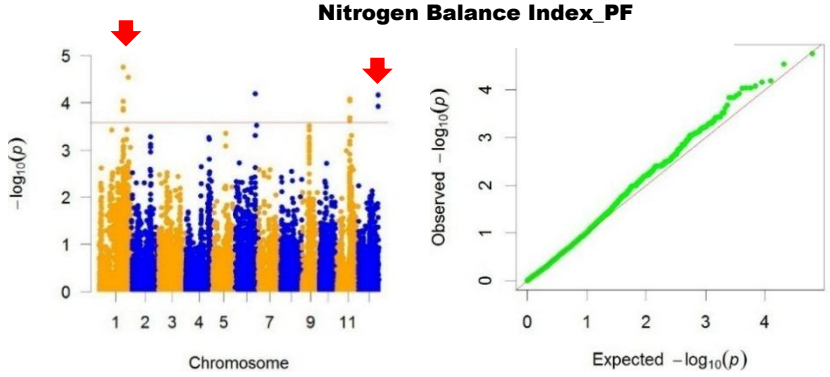
*ANNEX 9a: Manhattan (left) and QQ (right) plots for physiological traits. Red arrows indicate the most interesting association.*



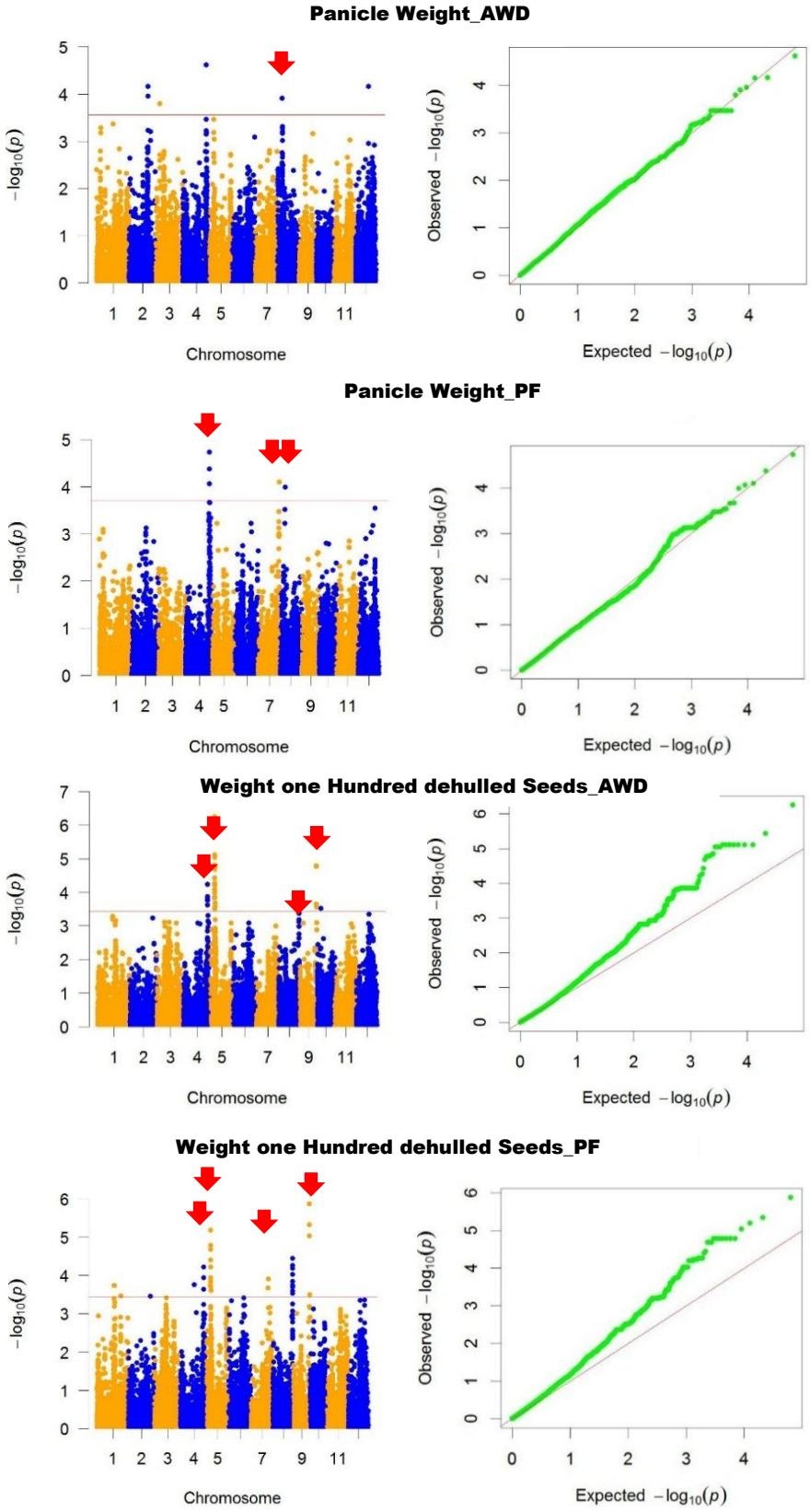
**Nitrogen Balance Index_AWD**



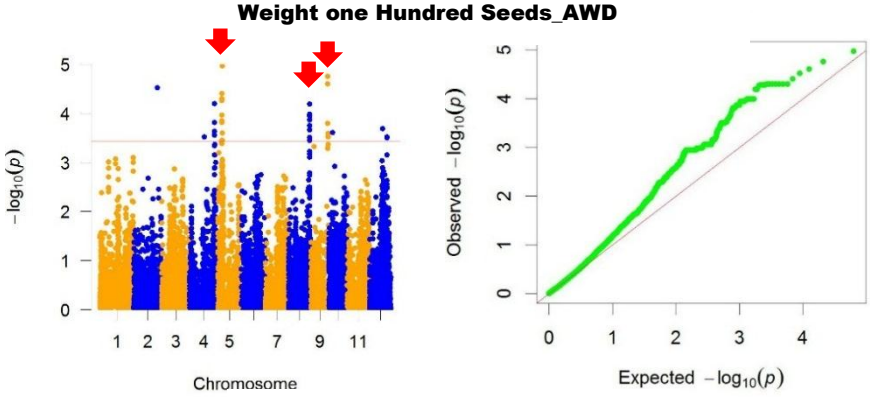

**Nitrogen Balance Index_PF**



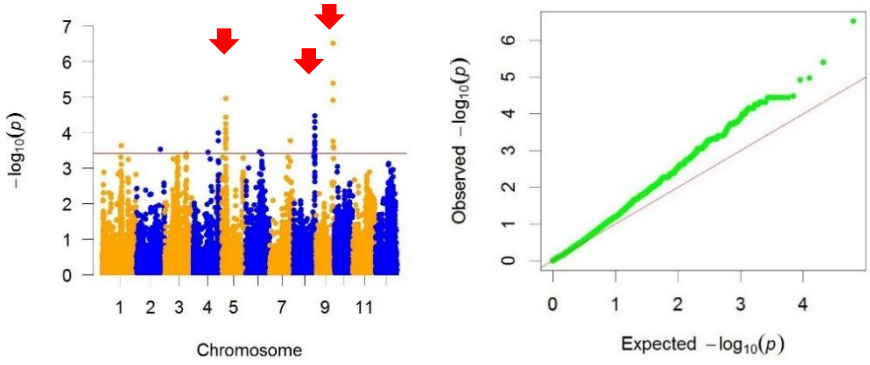
*ANNEX 9b: Manhattan (left) and QQ (right) plots for yield-related traits. Red arrows indicate the most interesting association.*



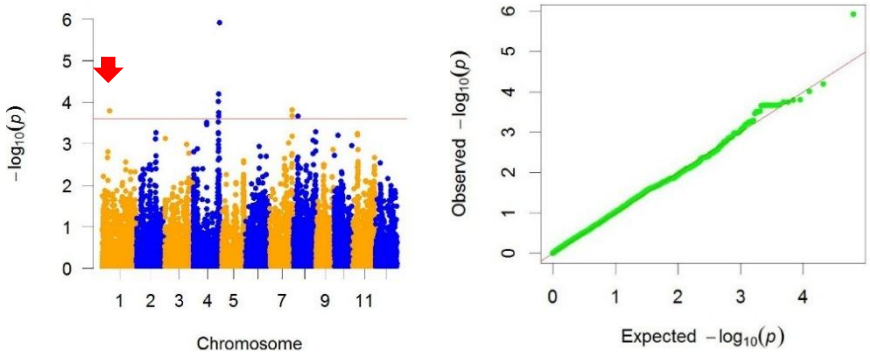
**Weight one Hundred Seeds_AWD**



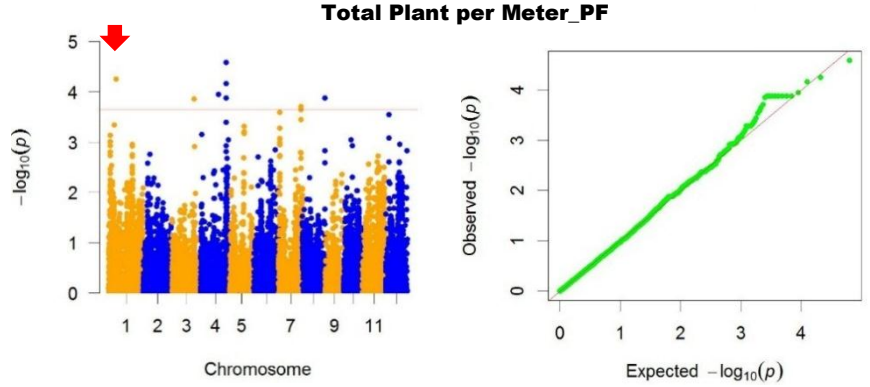

**Weight one Hundred Seeds_PF**




**Total Plant per Meter_AWD**

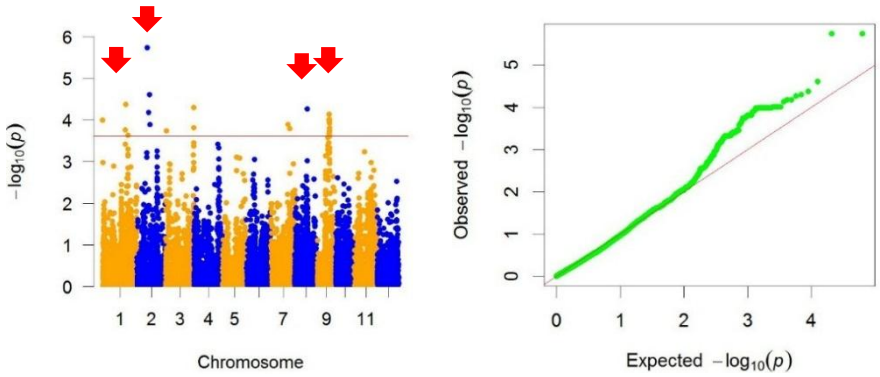



**Total Plant per Meter_PF**

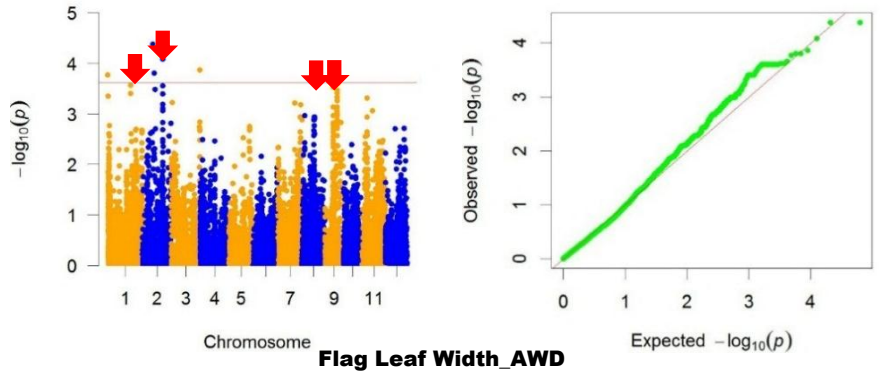


*ANNEX 9c: Manhattan (left) and QQ (right) plots for morphological traits. Red arrows indicate the most interesting association.*

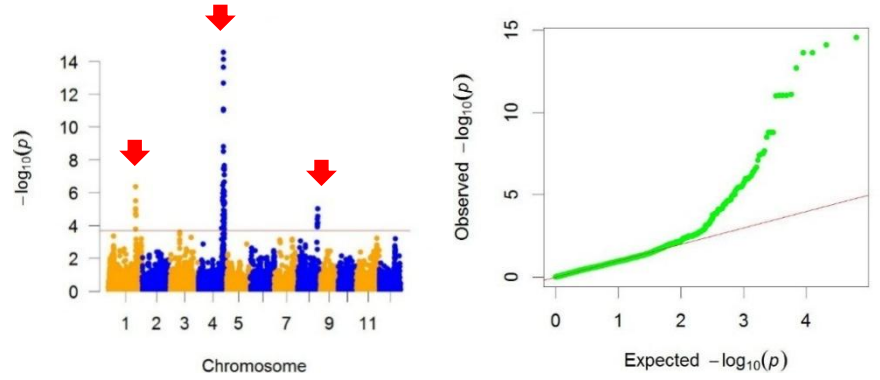
**Flag Leaf Length_AWD**



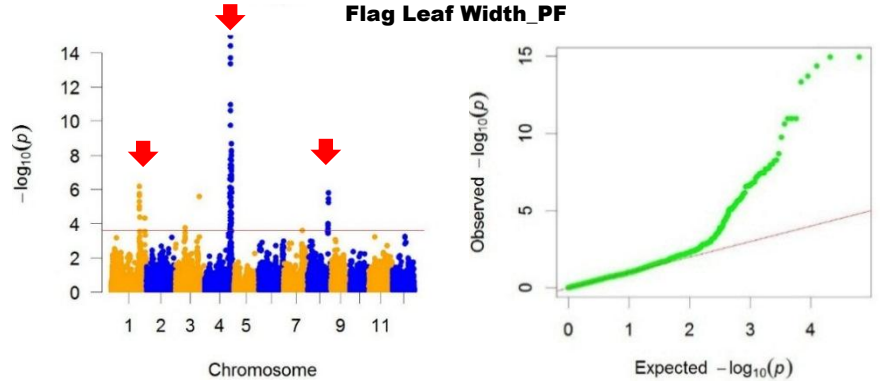

**Flag Leaf Length_PF**



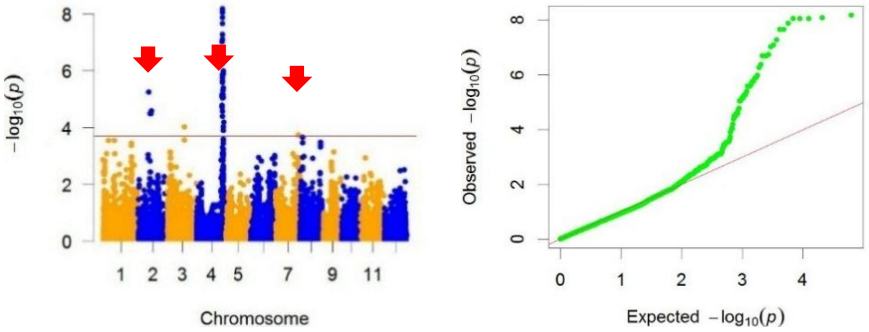

**Flag Leaf Width_AWD**



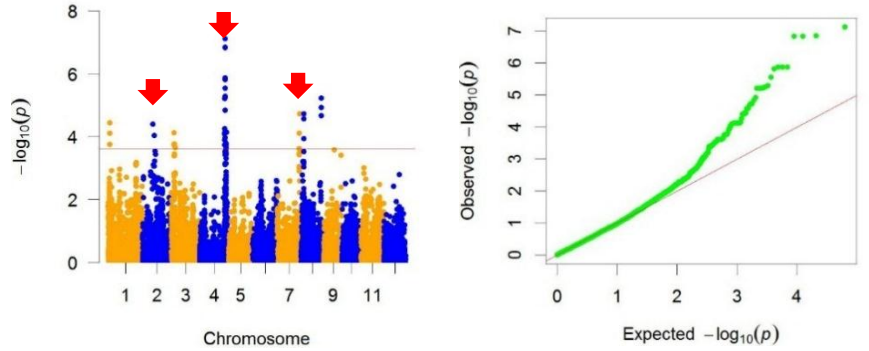

**Flag Leaf Width_PF**



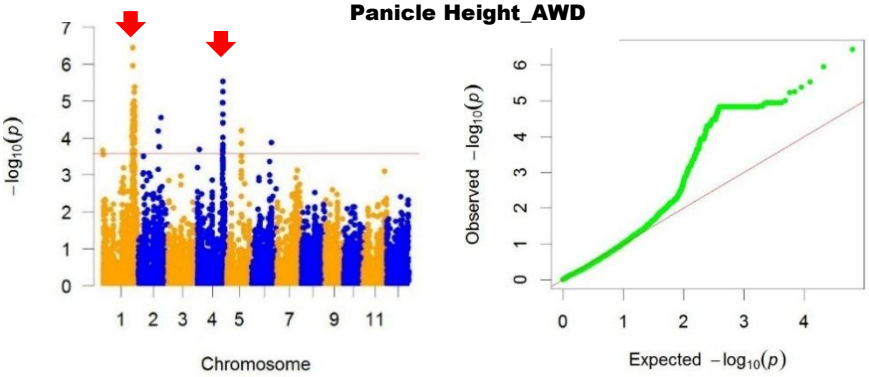
**Leaf Area_AWD**



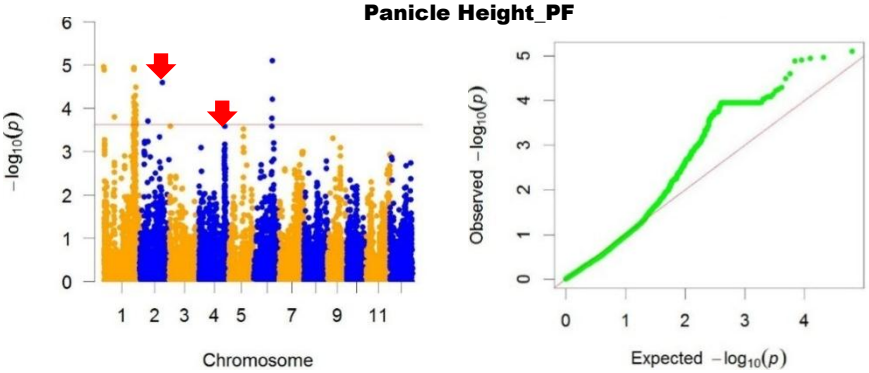

**Leaf Area_PF**

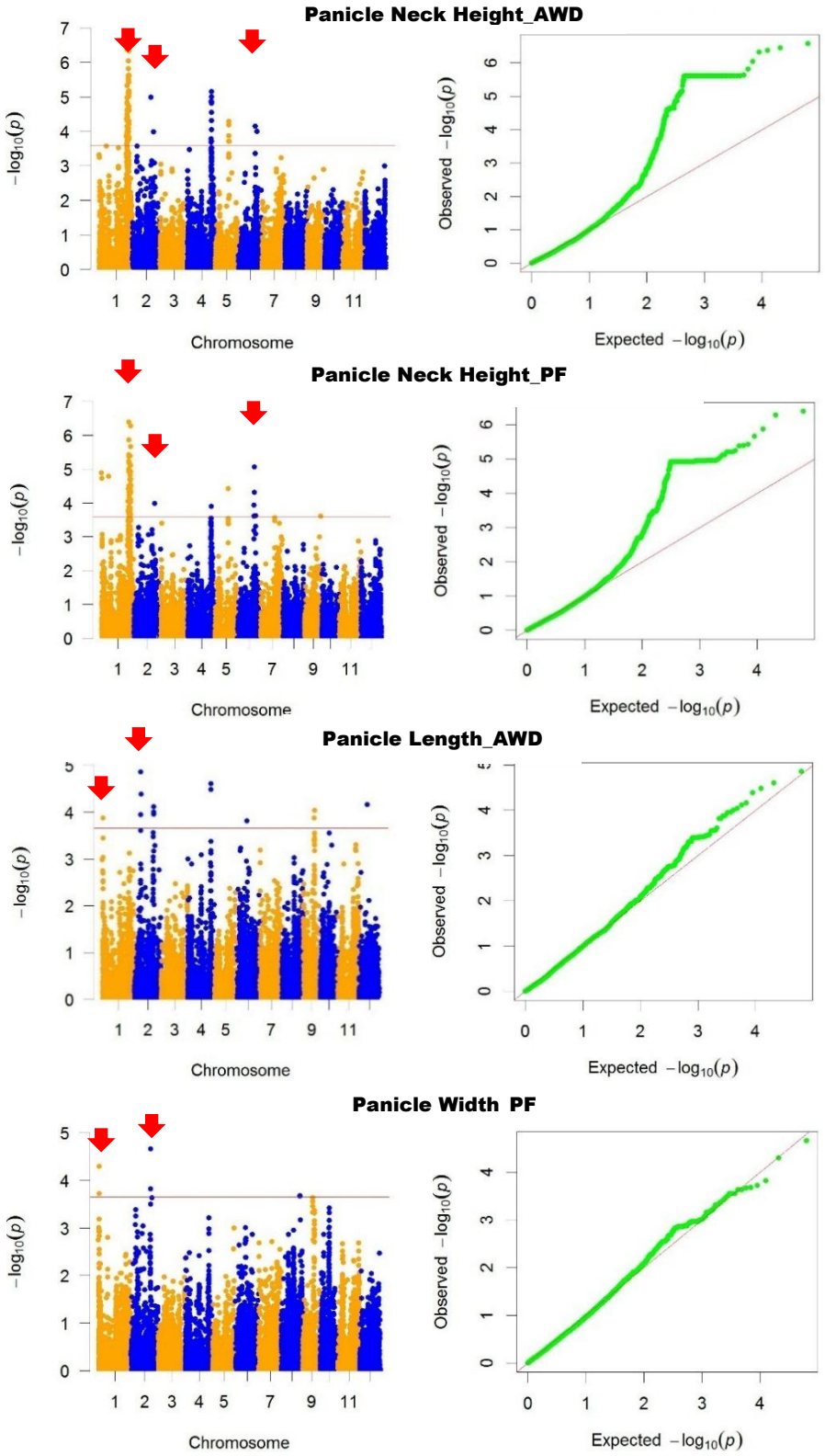



**Panicle Height_AWD**



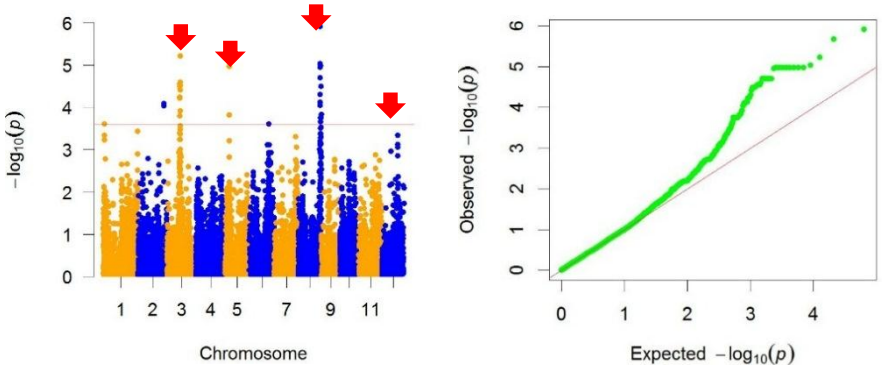

**Panicle Height_PF**



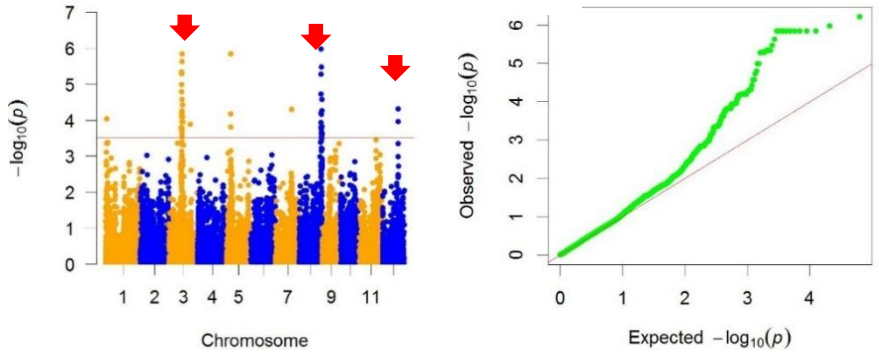

Panicle Neck Height_AWD
Panicle Neck Height_PF
Panicle Length_AWD
Panicle Width PF
−log10(p)
Chromosome
Observed −log10(p)
Expected −log10(p)

*ANNEX 9d: Manhattan (left) and QQ (right) plots for grain traits. Red arrows indicate the most interesting association.*

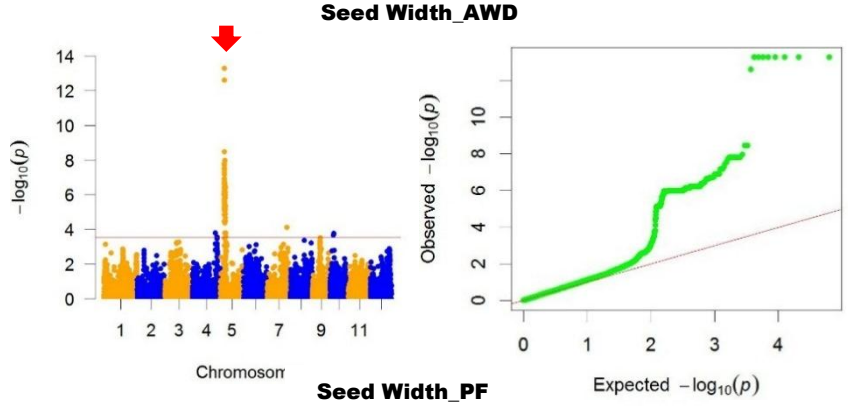
**Seed Length_AWD**



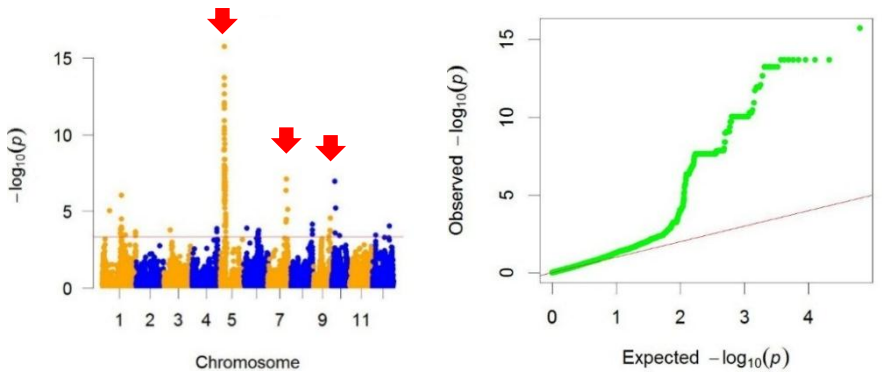

**Seed Length_PF**



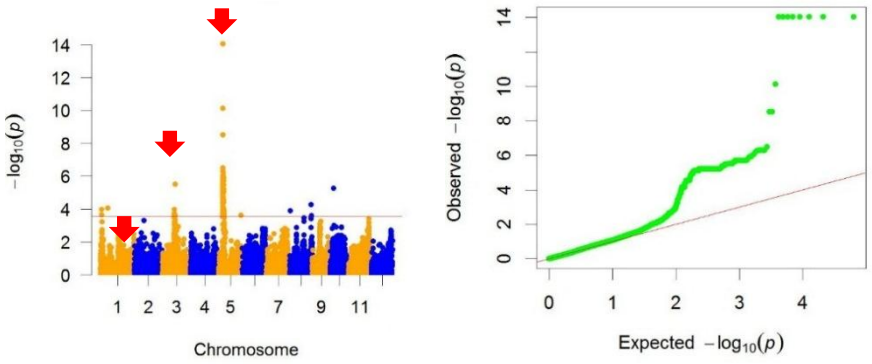

**Seed Width_AWD**



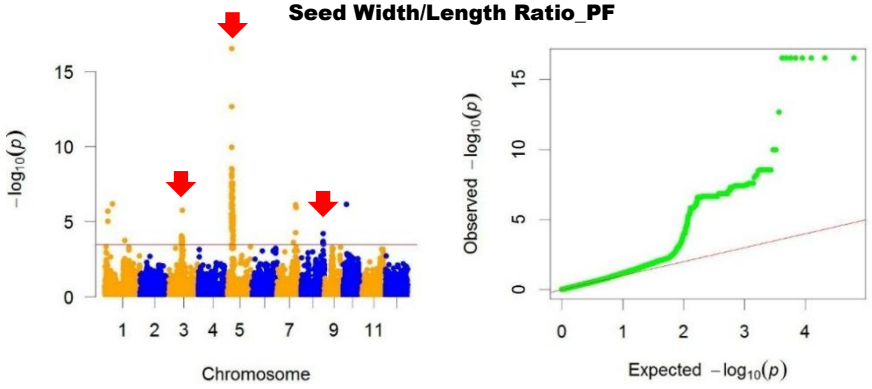

**Seed Width_PF**



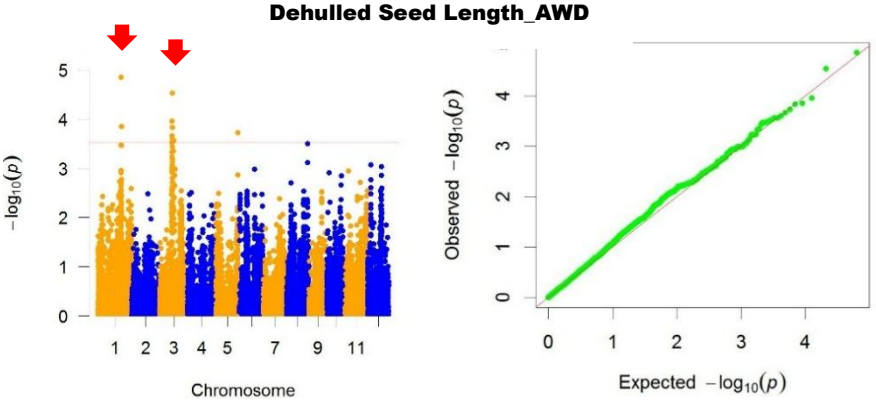
**Seed Width/Length Ratio_AWD**



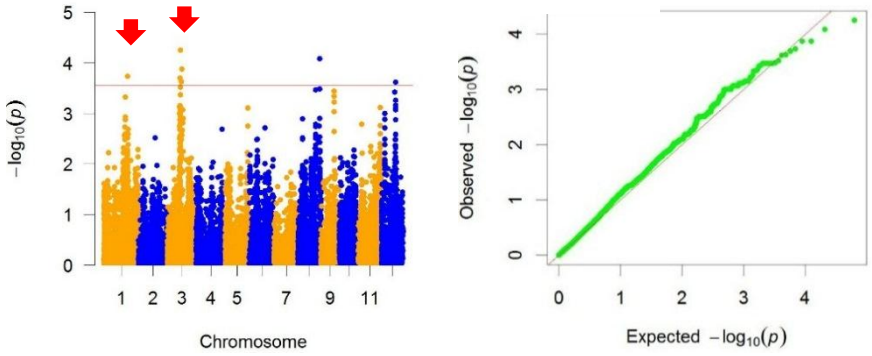

**Seed Width/Length Ratio_PF**



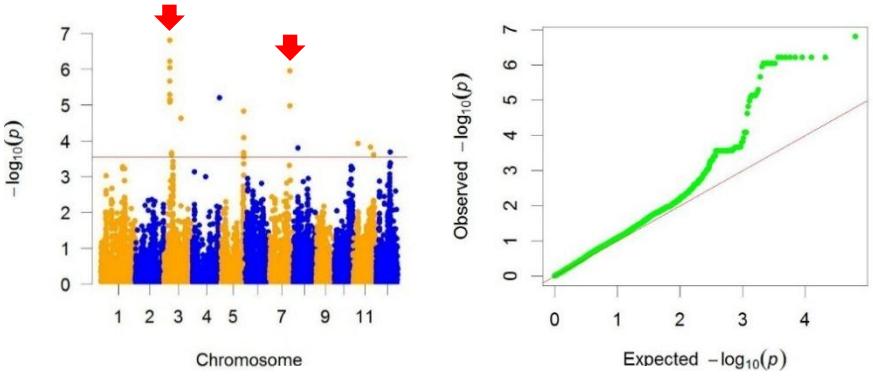

**Dehulled Seed Length_AWD**



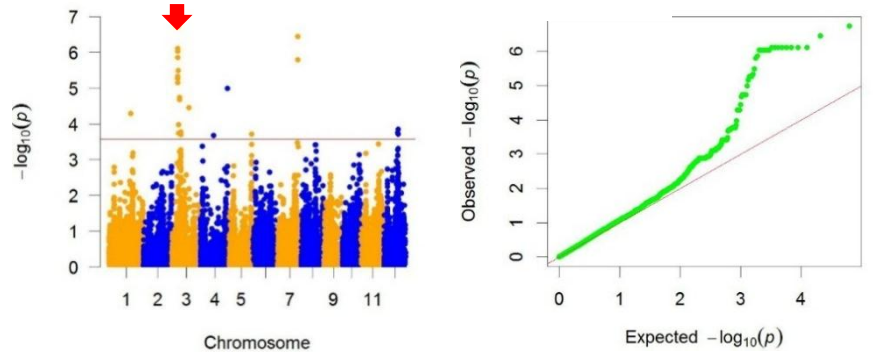

**Dehulled Seed Length_PF**



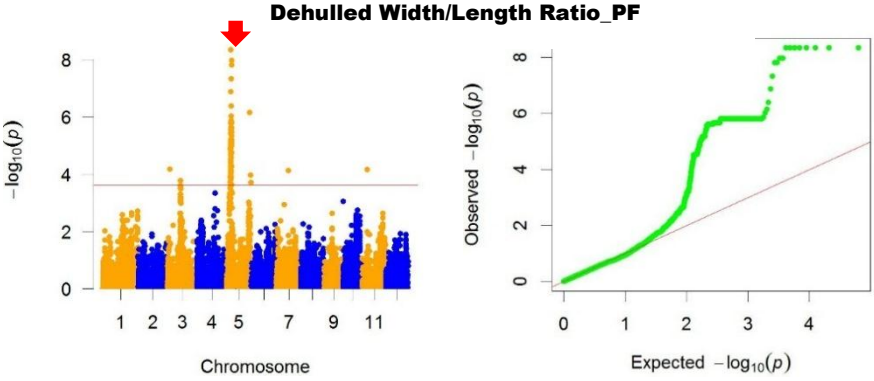
**Dehulled Seed Width_AWD**



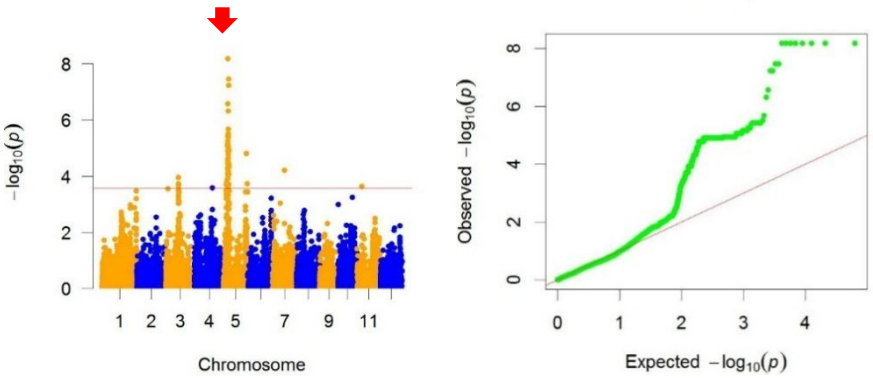

**Dehulled Seed Width_PF**



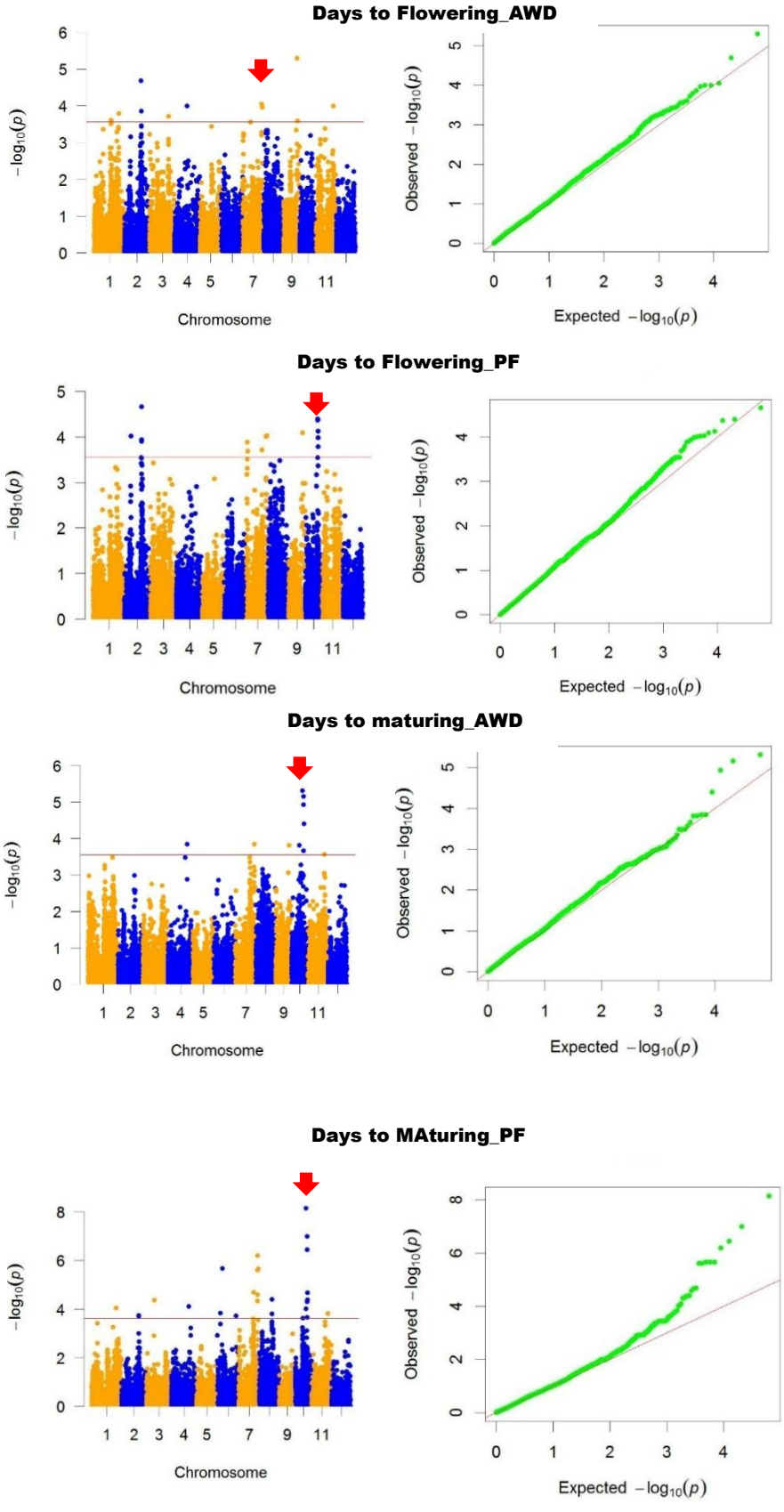

**Dehulled Width/Length Ratio_PF**



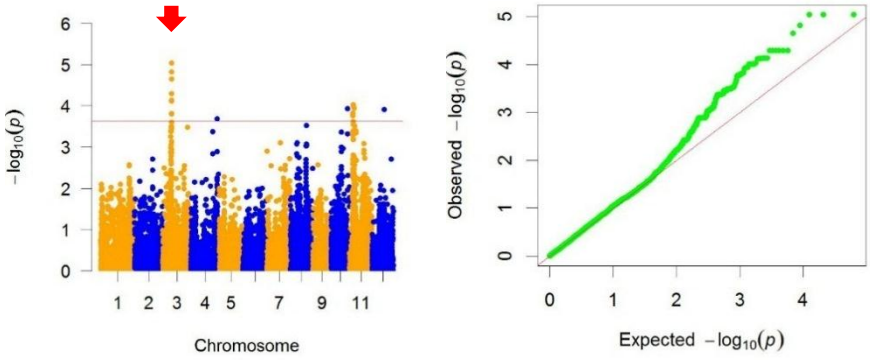

**Dehulled Seed Width/Length Ratio_PF**



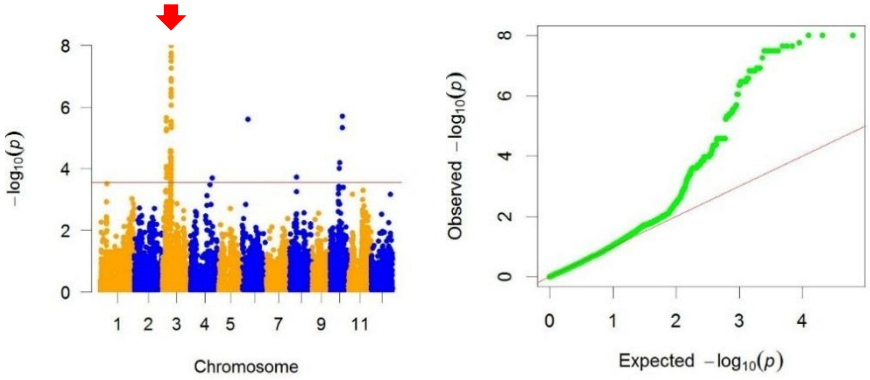
*ANNEX 9e: Manhattan (left) and QQ (right) plots for phenological traits. Red arrows indicate the most interesting association.*



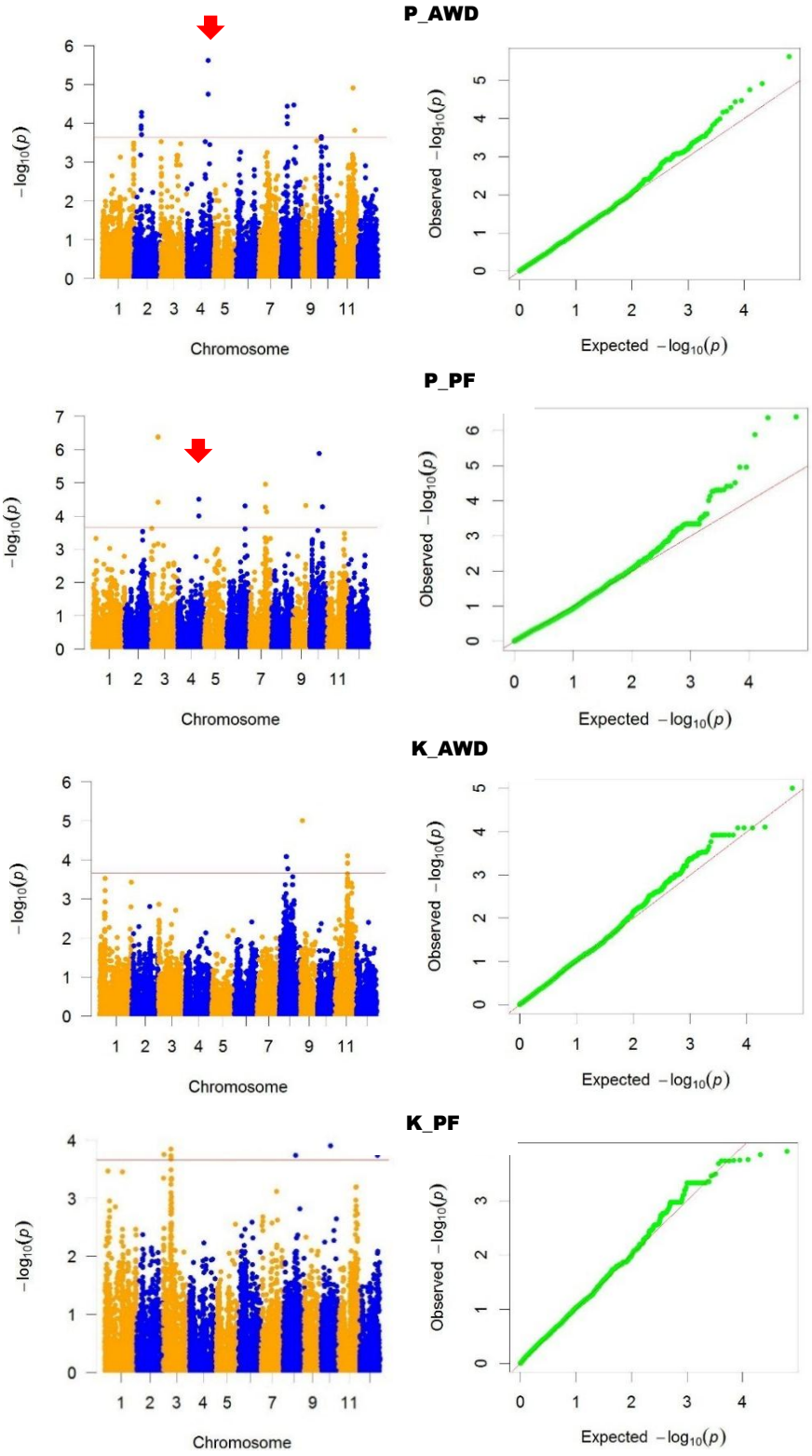
**Filling Time_AWD**



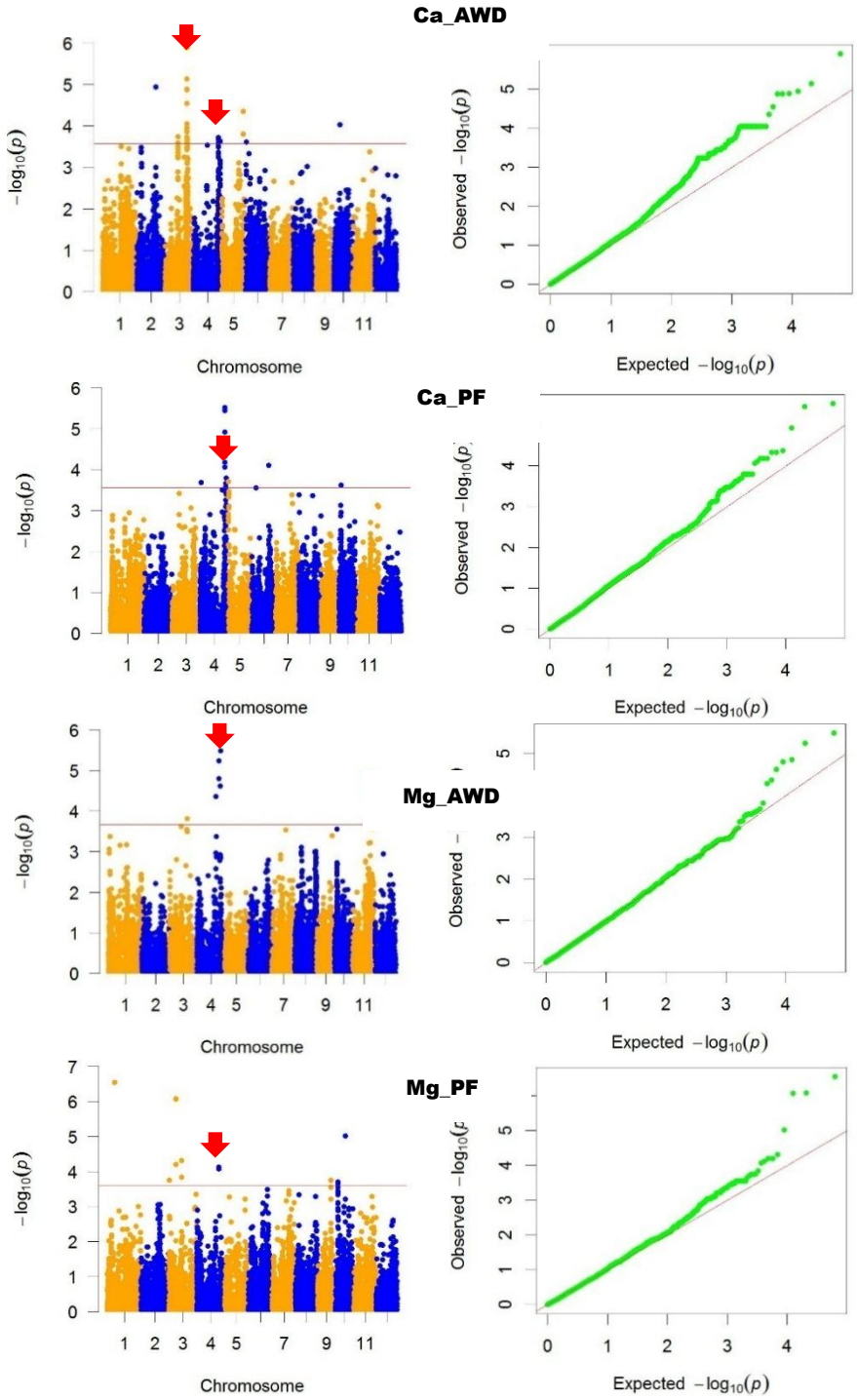

**Filling Time_PF**



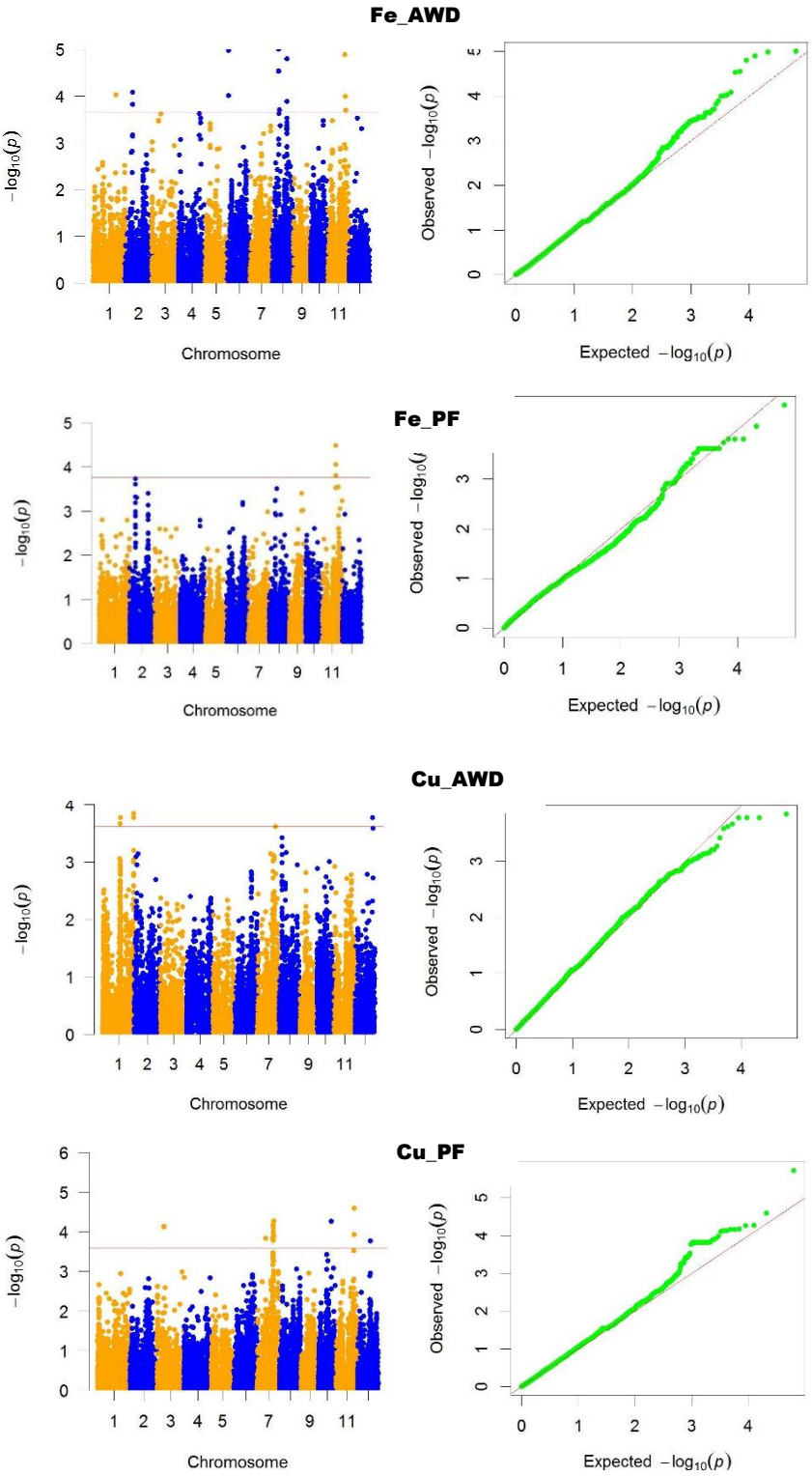
*ANNEX 10a: Manhattan (left) and QQ (right) plots for P and K. Red arrows indicate the most interesting association*

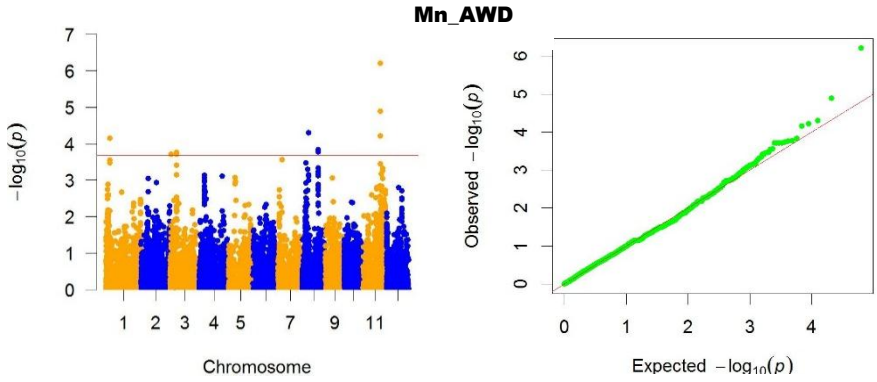


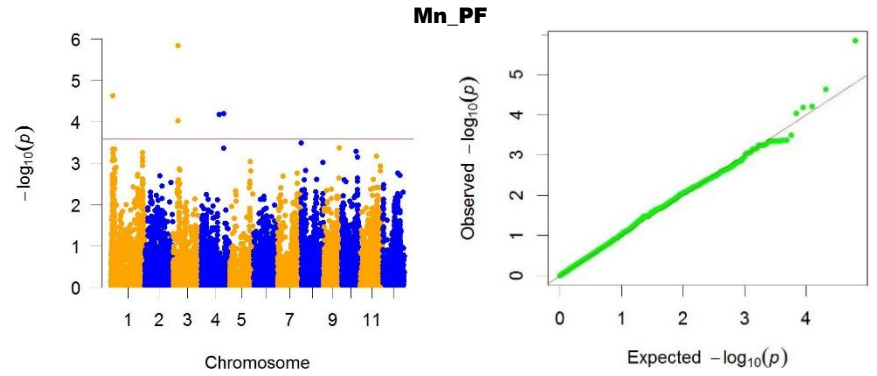
*ANNEX 10b: Manhattan (left) and QQ (right) plots for Ca and Mg. Red arrows indicate the most interesting association*

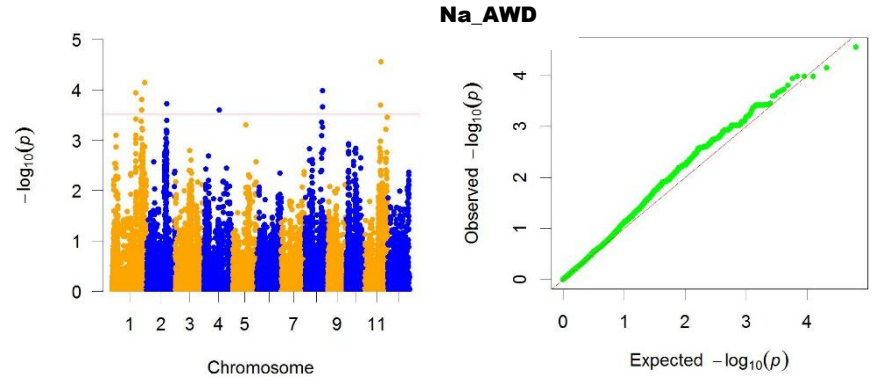


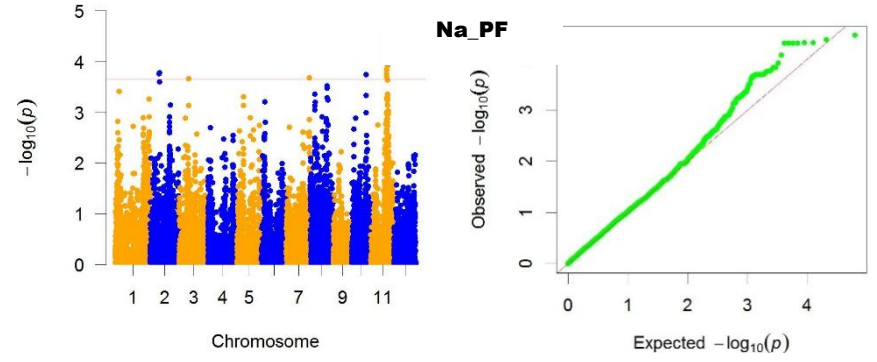
*ANNEX 10c: Manhattan (left) and QQ (right) plots for micro elements. Red arrows indicate the most interesting association*

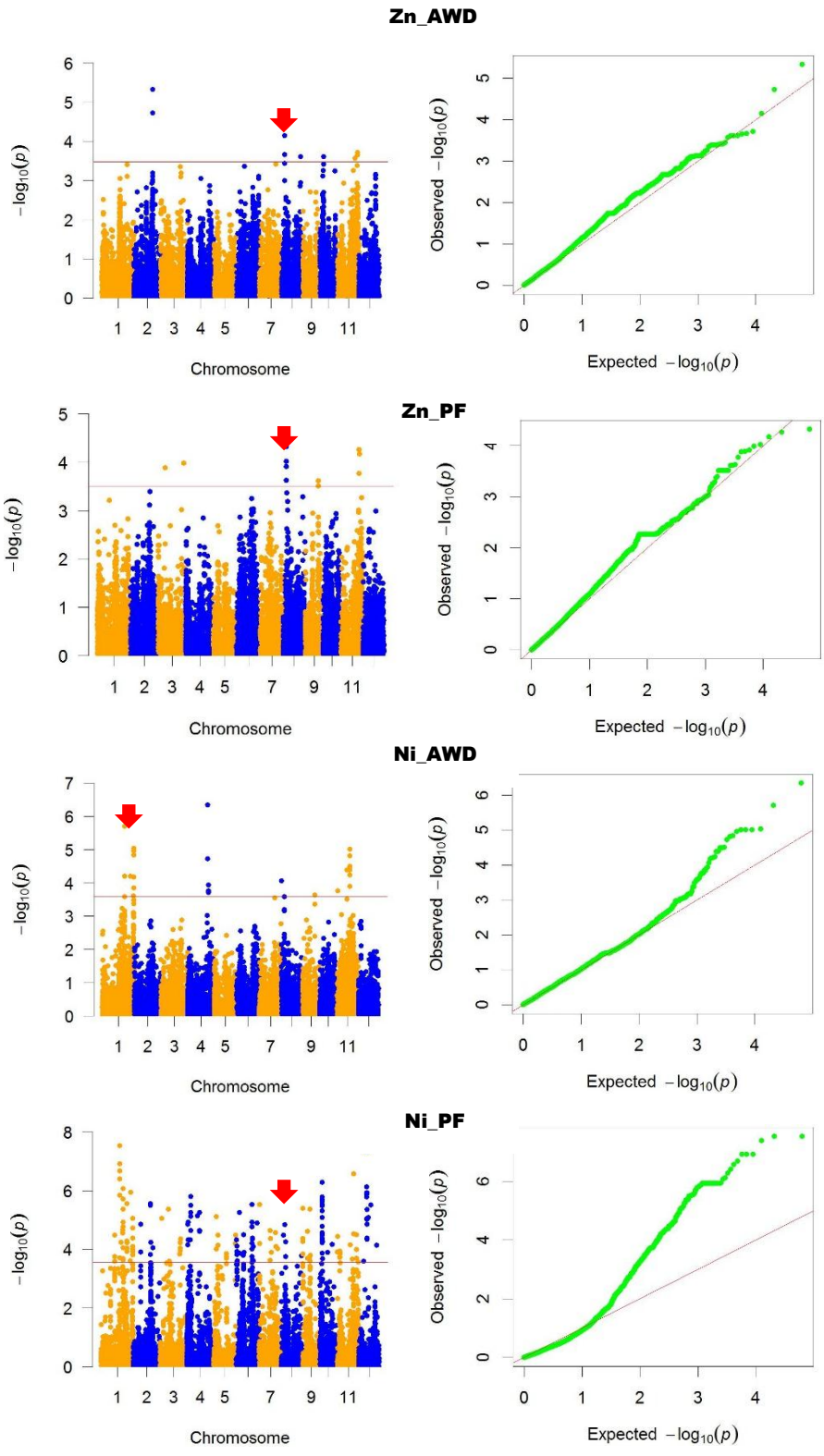



Mn_AWD
−log10(p)
Chromosome
Observed −log10(p)
Expected −log10(p)

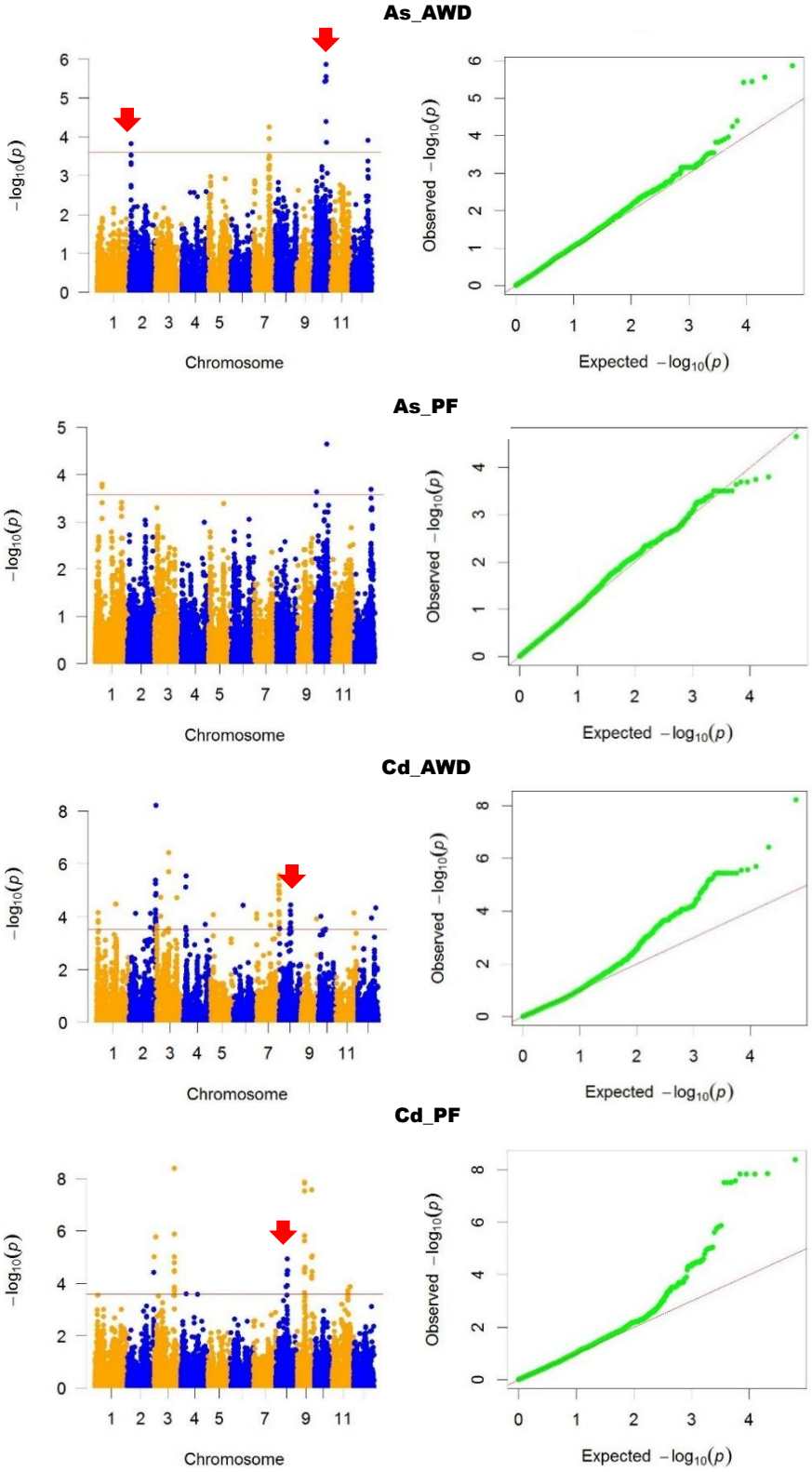

Mn_PF
−log10(p)
Chromosome
Observed −log10(p)
Expected −log10(p)


Na_AWD
−log10(p)
Chromosome
Observed −log10(p)
Expected −log10(p)


Na_PF
−log10(p)
Chromosome
Observed −log10(p)
Expected −log10(p)


Zn_AWD
Zn_PF
Ni_AWD
Ni_PF
−log10(p)
Chromosome
Observed −log10(p)
Expected −log10(p)

*ANNEX 10C: Manhattan (left) and QQ (right) plots for undesired trace elements. Red arrows indicate the most interesting association*



## *ANNEX 11.a: RAP database packing scripts:*

```
rm(list=ls())

setwd("C://Users//…//range")

range                                          <-
read.table("CH_tot_R.txt",sep="\t",header=TRUE
)

setwd("C://Users//…/GFF3_representative")



data<-
read.table("RAP_locus.txt",sep="\t",header=T)



prime<-data[ which(data$CH >13 ), ]

l<-length(range$z)

for (i in 1:l){

 z<-i

  min<-(range$From[z])

  max<-(range$To[z])

  chr<-range$CH[z]



  chrom<-data[ which(data$CH==chr ), ]

  newdata <- chrom[ which((chrom$BY > min &
chrom$BY< max )|(chrom$AT > min & chrom$AT< max
)), ]

  d<-length(newdata$BY)

  if(d>0 ){

    n<-cbind(newdata,z)

    prime <- rbind(prime, n)

  } else {




    newdatap<-matrix(1:9,1)

    newdatap<-newdatap*0



    dimnames(newdatap)[[2]]<-
c("CH","ORIGIN","TYPE","BY","AT","STRAND","ID"
,"NOTE","TRANSCRIPT_VARIANTS")

    n<-cbind(newdatap,z)

    prime<-rbind(prime,n)

  }



}

df<-prime

write.table(df,     "DATASET_LOCI_DIVIDED.txt",
quote = F, sep = "\t", row.names = F, col.names
= T)
```

## *ANNEX 11.b: Script for complete and automated analysis*

```
rm(list=ls())

setwd("C://Users//…//GFF3_representative")

data<-
read.table("DATASET_LOCI_DIVIDED_5_95_20170201
.txt",sep="\t",header=T

converter<-
read.table("RAP_MSU_converter.txt",sep="\t",he
ader=T)

setwd("C://Users//…//haplo//range")

range_table<-
read.table("CH_tot_R.txt",sep="\t",header=T

elements<-c("PL")

lel<-length(elements)

et<-1

for (et in 1:lel){

  el<-elements[et]

  setwd(file.path("C://Users//…//Preliminary
results//300_jumpl",el))

 #######          FDR       ##################

MLM                                            <-
read.table("MLM_S.txt",sep="\t",header=TRUE)

qbase<-as.matrix("q")

mainly<-unique(MLM$Trait)

mainly
```

```
MLMfil <- MLM[ which(MLM$p!=is.na(MLM$p)), ]


object<-as.matrix(unique(MLMfil$Trait))


#head(MLM)

#str(MLM)

lo<-length(object)

for (id in 1:lo){

  obj<-object[id]

  prime<-data[ which(data$CH>13), ]

  snps<-0

  snps<-as.matrix(snps)

  snps<-snps[ which(snps>13), ]

  single<-MLMfil[ which(MLMfil$Trait==obj), ]


  ###############FDR

  alpha <- 0.05

  m<-length(single$p)

  m

  alphasingle<-subset(single,subset=(p<alpha))

  malpha<-length(alphasingle$p)

  malpha

  lenghtlist<-c(1:m)

  unosun<-c(1/lenghtlist)

  summation<-sum(unosun)

  q<-(malpha/m)*(alpha/summation)

q

  singleFDRfiltered<-
subset(single,subset=(p<q))

  summary(singleFDRfiltered)

  write.table(singleFDRfiltered,
paste(obj,("_FDR.txt"),sep=""), quote = F, sep
= "\t", row.names = F, col.names = T)

q

  qbase<-rbind(qbase,q)#scrivo nel summary il
risultato di q

  dimnames(qbase)[[2]]<-c("q")

  ########################Grafici

  library(qqman)



cutoff<--log(q,10)

  evidence<-
as.vector(singleFDRfiltered$Marker)

  jpeg(paste(obj,("_mhan.jpg"),sep=""),
width=1000, height=1000, pointsize=40

  MANHATTAN<-manhattan(single,
chr="Locus",bp="Site", p="p",snp="Marker", col
= c("orange", "blue"),suggestiveline =F ,
genomewideline = cutoff, logp=T,main=obj)

    dev.off()


    jpeg(paste(obj,("_QQ.jpg"),sep=""),
width=1000, height=1000, pointsize=40


    QQPLOT1<-
qq(single$p,col=c("green"),main=obj)

      dev.off()


  pdf(paste(obj,("_Grafici1.pdf"),sep=""),
width=15, height=10)

MANHATTAN<-manhattan(single,
chr="Locus",bp="Site", p="p",snp="Marker", col
= c("orange", "blue"),suggestiveline =F ,
genomewideline = cutoff, logp=T,main=obj)


  MANHATTAN2<-manhattan(single,
chr="Locus",bp="Site", p="p",snp="Marker", col
= c("green", "blue"),suggestiveline =F ,
genomewideline = cutoff, logp=T,main=obj)

  QQPLOT1<-
qq(single$p,col=c("green"),main=obj)

  chFDR<-unique(singleFDRfiltered$Locus)
```

```
  for (ripe in 1:12){

    singleCH<-single[
which(single$Locus==ripe), ]

    MANHATTAN<-manhattan(singleCH,
chr="Locus",bp="Site", p="p",snp="Marker", col
= c( "blue"), suggestiveline =F ,genomewideline
=cutoff, logp=T,main=obj)

#ho dovuto eliminare gli snps evidence perchè
non sempre ce n'erano

}

  dev.off()



  ###############Silico Analysis

  lra<-length(singleFDRfiltered$Site)

  rname<-0

  for (t in 1:lra ){

    snp<-singleFDRfiltered$Site[t]

    ch<-singleFDRfiltered$Locus[t]

    locus_range<-                  range_table[
which((range_table$CH==ch                    &
range_table$From<snp & range_table$To>snp )), ]

  z<-locus_range$z

  rname<-rbind(rname,z)



    #parentesi fine ciclo for

  }

 rname<-rname[,1];rname<-
as.vector(rname);rname <- subset(rname, rname>0
)

  rname<-unique(rname)



for (u in rname){

  silico<- data[ which((data$z==u)), ]

  prime<-rbind(prime,silico)



}

  prime<-unique(prime[!duplicated(prime$ID),])

  if ((length(prime$ID))>0){



  unique_z<-unique(prime$z)

  unique_z



  z<-0

  nSNPs<-0

  lSNPs<-0

  second<-cbind(z,nSNPs,lSNPs)

  second<-second[which(second[1]>0),]

  l<-length(unique_z)

  for (t in 1:l){

    group<-unique_z[t]

    loc<-range_table[
which(range_table$z==group), ]

    min<-loc$From[1]

    max<-loc$To[1]

    Chrom<-loc$CH[1]

    snps<-singleFDRfiltered[
which(singleFDRfiltered$Locus==Chrom         &
singleFDRfiltered$Site>min                   &
singleFDRfiltered$Site<max), ]

    nSNPs<-length(snps$Site)

    lSNPs<-max-min

    row<-cbind(group,nSNPs,lSNPs)

    second<-rbind(second,row)



  }

  second<-        merge(prime,second,by.x="z",
by.y="z",all.x=T, all.y=F, incomparables=F)

  second<-    merge(second,converter,by.x="ID",
by.y="RAP",all.x=T, all.y=F, incomparables=F)



  write.table(second,
paste(obj,("_silico_loci_whole.xls"),sep=""),
```

```
quote = F, sep = "\t", row.names = F, col.names
= T)

  write.table(second,
paste(obj,("_silico_loci_whole.txt"),sep=""),
quote = F, sep = "\t", row.names = F, col.names
= T)

  }else{

    write.table(prime,
paste(obj,("_silico_loci_whole.xls"),sep=""),
quote = F, sep = "\t", row.names = F, col.names
= T)

    write.table(prime,
paste(obj,("_silico_loci_whole.txt"),sep=""),
quote = F, sep = "\t", row.names = F, col.names
= T)

  }

}

}
```

*ANNEX 12a: Agronomic synthetic phenotypic data.Mean value for AWD and Pf, calculated with LMS approach and p_value for the condition effect.*

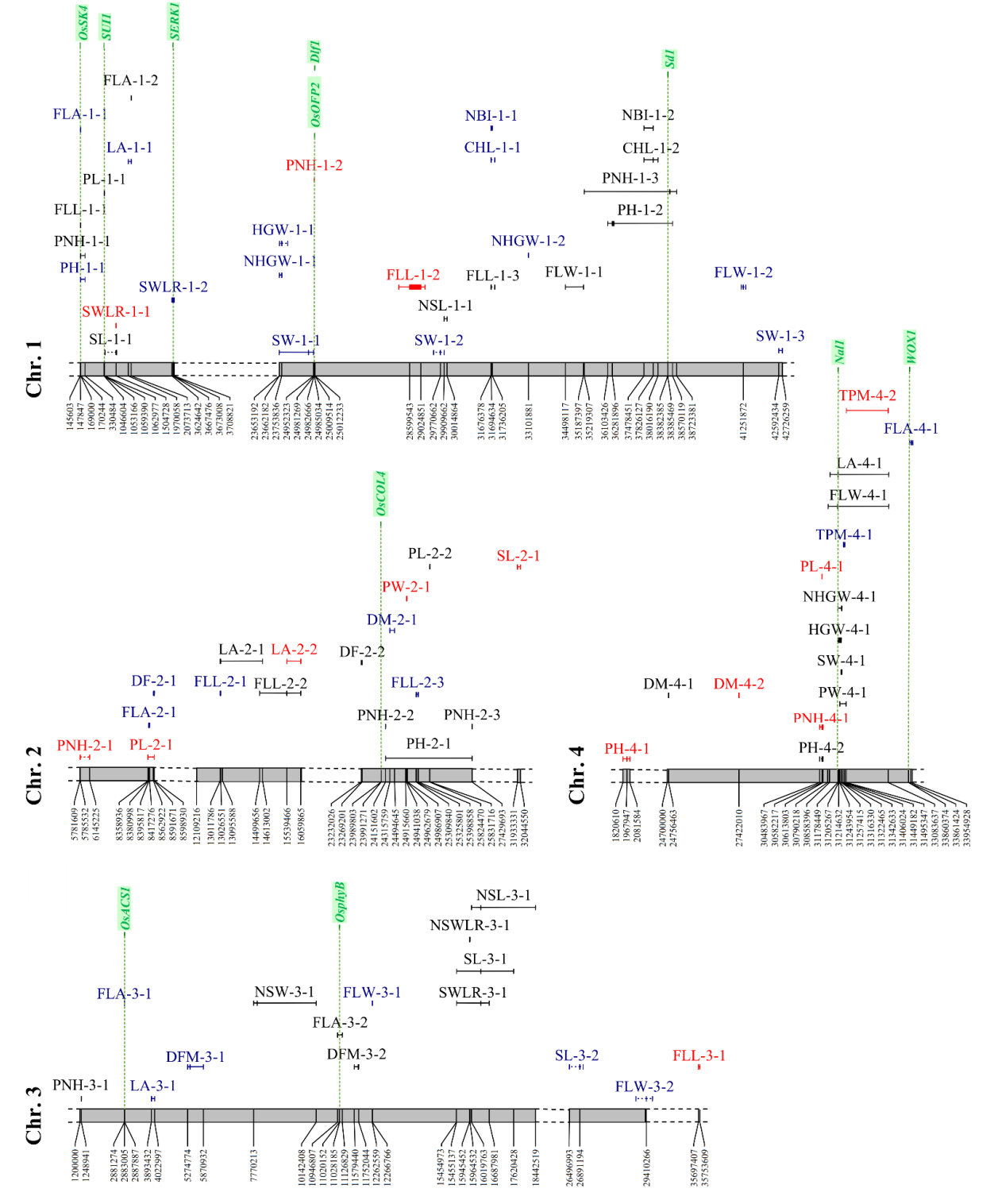
| | Mean value | | *p*_value |
|---|---|---|---|
| Trait | AWD | PF | |
| DF (*d*) | 100.2 ± 11 | 96.2 ± 8 | <.0001 |
| DM (*d*) | 156.5 ± 13 | 150.8 ± 10 | <.0001 |
| DFM (*d*) | 56.6 ± 8 | 57.6 ± 8 | 0.0015 |
| FLL (*mm*) | 208.9 ± 43 | 262.61 ± 58 | <.0001 |
| FLW (*mm*) | 12.5 ± 2 | 12.3 ± 2 | 0.0668 |
| LA ($mm^2$) | 1962.5 ± 702 | 2448.39 + 839 | <.0001 |
| PH (*cm*) | 79.3 ± 13 | 90.7 ± 13 | <.0001 |
| PNH (*cm*) | 61.9 ± 12 | 71.1 + 12 | <.0001 |
| PL (*cm*) | 17.4 ± 3 | 19.6 ± 3 | <.0001 |
| TPM | 83.8 ± 29 | 88.7 ± 24 | <.0001 |
| SL (*mm*) | 8.9 ± 1 | 9.16 ± 1 | <.0001 |
| SW (*mm*) | 3.2 ± 0.4 | 3.4 ± 0.5 | <.0001 |
| SWLR | 0.37 ± 0.08 | 0.38 ± 0.08 | 0.1837 |
| NSL (*mm*) | 6.4 ± 0.7 | 6.71 ± 0.8 | <.0001 |
| NSW (*mm*) | 2.7 ± 0.3 | 2.8 ± 0.4 | <.0001 |
| NSWLR | 0.43 ± 0.1 | 0.43 ± 0.1 | 0.7776 |
| HGW (*g*) | 2.9 ± 0.5 | 3.2 ± 0.6 | <.0001 |
| NHGW (*g*) | 2.4 ± 0.4 | 2.6 ± 0.5 | <.0001 |
| PW (*mm*) | 126.7 ± 43 | 171.6 ± 45 | <.0001 |
| CHL_AV | 30.1 ± 6 | 34.59 ± 6 | <.0001 |
| FLA | 3.0 ± 0.2 | 2.97 ± 0.3 | <.0001 |
| NBI_AV | 10.0 ± 2 | 11.8 ± 2 | <.0001 |

(Volante et al., 2017)

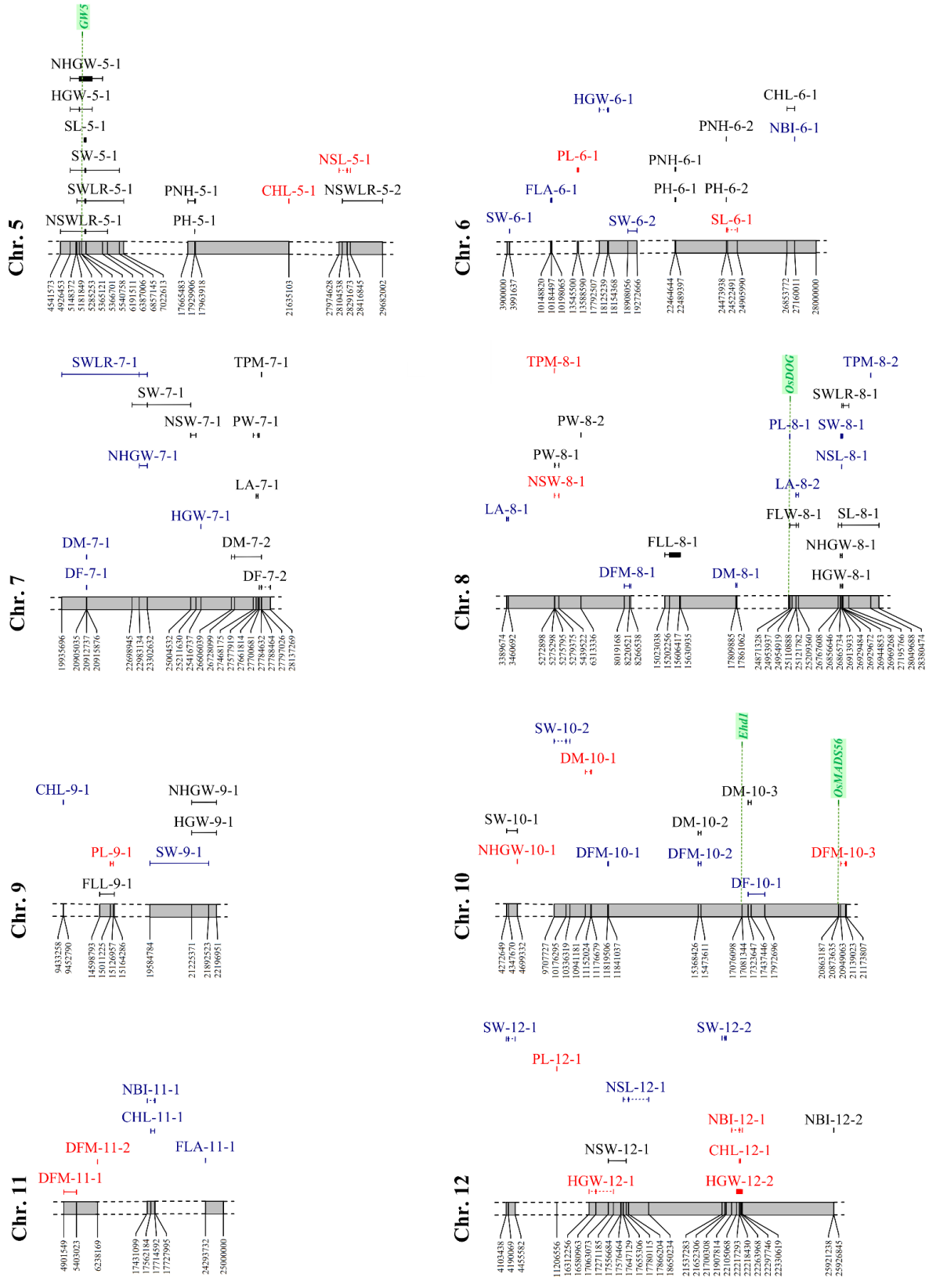
*ANNEX 12b : Ionomic synthetic phenotypic data. Mean value for AWD and Pf, calculated with LMS approach and p_value for the condition effect.*

| | Mean | | *P*_value |
|---|---|---|---|
| Element | AWD | PF | |
| As *(µg kg$^{-1}$)* | 39 ± 15 | 123 ± 27 | <0.0001 |
| Ca *(mg kg$^{-1}$)* | 91.1 ± 14 | 87.4 ± 12 | <0.0001 |
| Cd *(µg kg$^{-1}$)* | 54.5 ± 30 | 25.9 ± 18 | <0.0001 |
| Cu *(mg kg$^{-1}$)* | 4.3 ± 0.7 | 2.7 ± 0.5 | <0.0001 |
| Fe *(mg kg$^{-1}$)* | 8.7 ± 1 | 7.9 ± 1 | <0.0001 |
| K *(g kg$^{-1}$)* | 2.9 ± 0.2 | 2.8 ± 0.2 | <0.0001 |
| Mg *(g kg$^{-1}$)* | 1.3 ± 0.1 | 1.2 ± 0.1 | <0.0001 |
| Mn *(mg kg$^{-1}$)* | 15.5 ± 2 | 20.8 ± 3 | <0.0001 |
| Na *(mg kg$^{-1}$)* | 6.9 ± 1 | 6.8 ± 1 | 0.617 |
| Ni *(mg kg$^{-1}$)* | 1.8 ± 0.4 | 0.3 ± 0.1 | <0.0001 |
| P *(g kg$^{-1}$)* | 3.8 ± 0.2 | 3.6 ± 0.3 | <0.0001 |
| Zn *(mg kg$^{-1}$)* | 22.6 ± 3 | 21.3 ± 3 | <0.0001 |

*ANNEX 13 : Maps of agronomic data. Below the gray bar, we can see the associated snps, above the bar MTAs and the genes in green colour.*

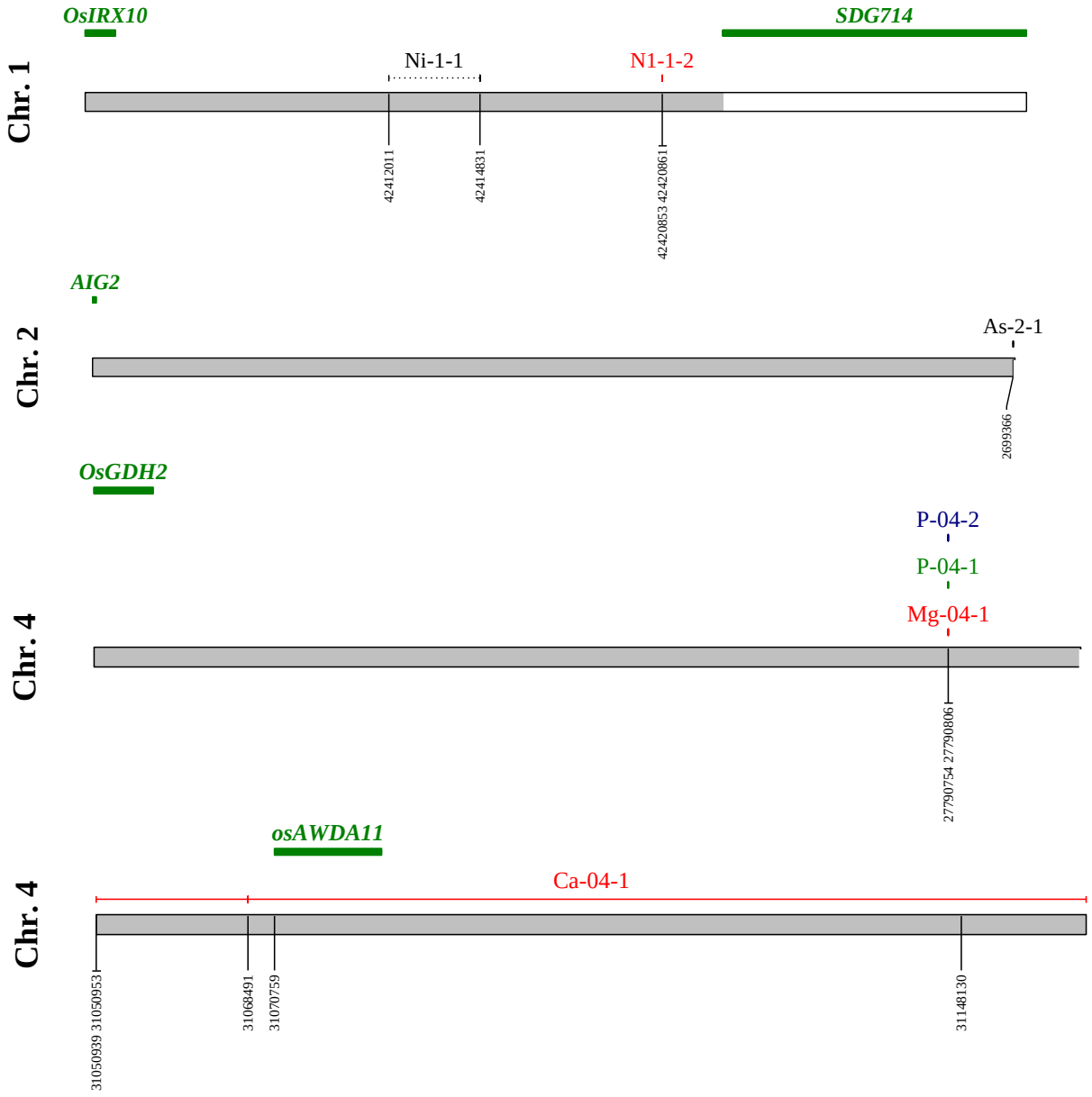


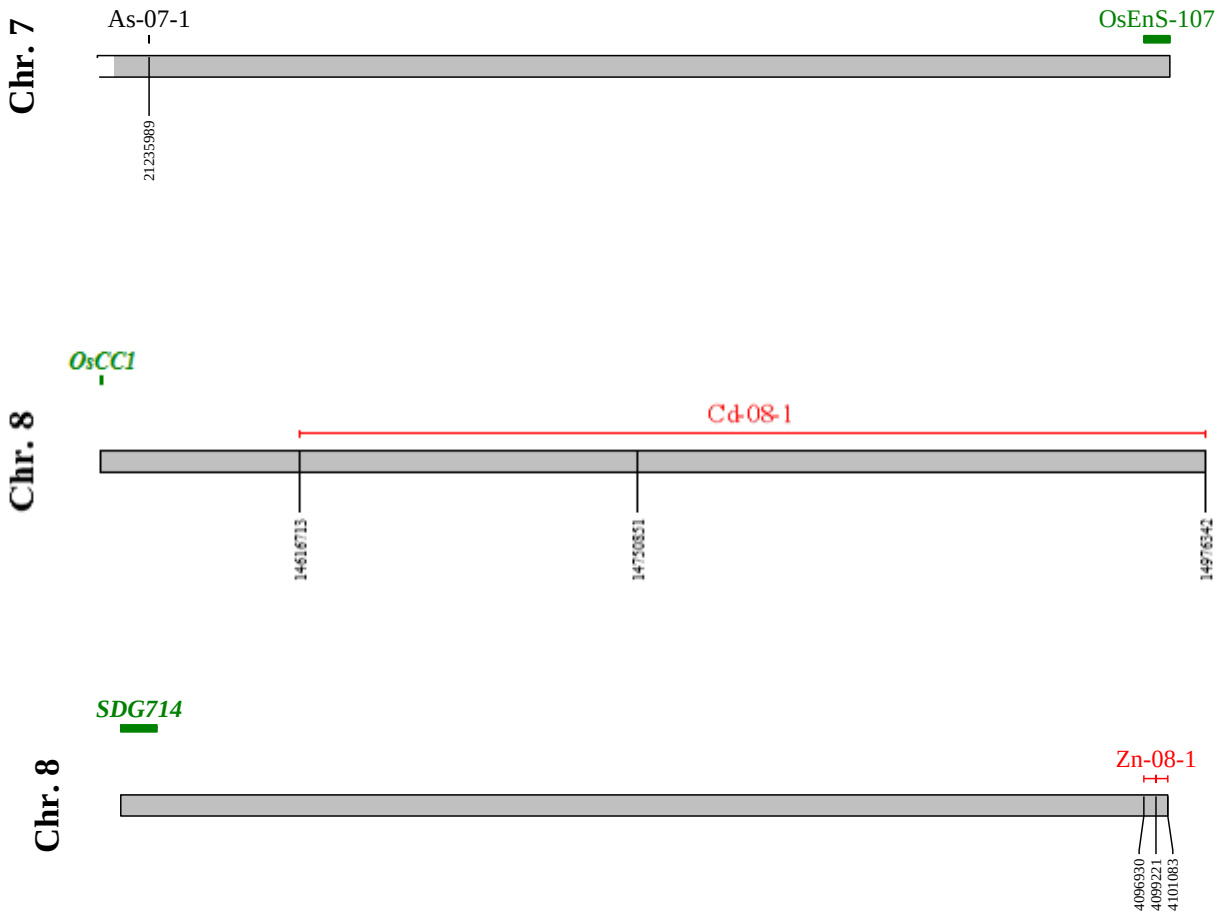

(Volante et al., 2017)




(Volante et al., 2017)

*ANNEX 14 : Maps of ionomic dataBelow the gray bar, we can see the associated snps, above the bar MTAs and the genes in green colour.*




Chr. 7
As-07-1
OsEnS-107
21235989
OsCC1
Cd-08-1
14616713
14750851
14976542
Chr. 8
SDG714
Zn-08-1
4096930
4099221
4101083
Chr. 8