UNIVERSIDADE DE TRÁS-OS-MONTES E ALTO DOURO

TARGETING GENES FOR AL TOLERANCE IN PORTUGUESE BREAD WHEAT (*Triticum aestivum* L.)

TESE DE DOUTORAMENTO
EM GENÉTICA MOLECULAR COMPARATIVA E TECNOLÓGICA

ANA LUÍSA GARCIA OLIVEIRA

Orientadora: Professora Doutora Paula Martins Lopes
Co-orientador: Professor Doutor César Benito
Co-orientador: Professor Doutor Henrique Guedes Pinto

Vila Real, 2013

Universidade Trás-os-Montes e Alto Douro

Targeting genes for Al tolerance in Portuguese bread wheat (*Triticum aestivum* L.)

TESE APRESENTADA PARA OBTENÇÃO DO GRAU DE DOUTOR EM GENÉTICA MOLECULAR COMPARATIVA E TECNOLÓGICA

ORIENTADORA: PROFESSORA PAULA MARTINS LOPES

CO-ORIENTADORES: PROFESSOR CÉSAR BENITO, PROFESSOR HENRIQUE GUEDES PINTO

Ana Luísa Garcia Oliveira

Vila Real
2013

Ana Luísa Garcia Oliveira

Targeting genes for Al tolerance in Portuguese bread wheat (*Triticum aestivum* L.)

Thesis submitted to Universidade de Trás-os-Montes e Alto Douro in fulfillment of the requirements for the degree of Doctor of Philosophy, in the field of Technologic, Comparative Molecular Genetics, under the guidance of Professor(s) Paula Martins Lopes, César Benito and Henrique Guedes Pinto.

UNIVERSIDADE DE TRÁS-OS-MONTES E ALTO DOURO

Targeting genes for Al tolerance in Portuguese bread wheat (*Triticum aestivum* L.)

The present PhD thesis was carried out under the supervision of Professor Paula Martins Lopes (CGB/IBB-UTAD), Professor César Benito (UCM), and Professor Henrique Guedes Pinto (CGB/IBB-UTAD).

Funding was provided by FCT (Fundação para a Ciência e Tecnologia de Portugal) in a form of PHD grant (SFRH/BD/45556/2008).

Governo da República Portuguesa

UNIÃO EUROPEIA
Fundo Social Europeu

In loving memory of my late grandmother and late great grandmother …

Joaquina Garcia Pereira (9/07/1940 –17/04/2011)

Antonia Garcia Pereira (14/05/1921 – 24/06/2011)

...

começar com um abrir de olhos ...
e de cada vez que os cerramos
numa memória incessante,
vem a expectância, a euforia e também o desapontamento
como numa correria à procura do que não encontra,
uma breve resposta pronta para um sem fim de perguntas
e quem sabe um vislumbre momentaneo,
jogando com o certo e com o errado
numa dança cuja saia não é mais que um infinito inumerado de folhos
que revimos ... e lembramos com humildade.

E, se por vezes a ironia e a arrogância pedante
interrompe a história que lemos no apontamento
é para nos fazer lembrar que na linha de conta,
do nosso conhecimento, todas as pontes têm juntas
as tábuas que construímos no infinito espontâneo
e nos vários Outonos à descoberta do Inverno pasmado,
a caída das folhas a serem amontoadas em molhos
não são mais que lições apreendidas e a serem revividas com saudade.

...

Dedicated to All that I love, especially to my grandfather

Acknowledgements

'When the student is ready, the master appears …' (Buddhist quote)

In this special moment I couldn't forget to acknowledge all the persons that encourage me along with meaningful words and cherish gestures during my last four years.

I'm very thankful to Prof. Henrique Guedes Pinto, Dr. Paula Martins Lopes, Prof. César Benito, Dr. Pilar Prieto, Prof. Claudina Rodrigues-Pousada and Prof. Charlotte Poschrenreider to accept and receiving me in their labs to endorse and complete successfully my PhD degree. All useful advices, teachings and kindness showed by all are here sincerely acknowledged. I'm extremely thankful to Prof. César for allowing my staying in his lab for almost the four years of my PhD, and all the teachings. To Prof. Claudina, Dr Pilar and Prof. Charllote thank you so much for all the kind gestures, help and keen advices. I'm very thankful to my supervisor Dr. Paula for giving me the opportunity to explore, as possible; better ways to endorse the research work and being helpful all the time. I'm also very thankful to Prof. Henrique Guedes-Pinto for keen advices.

I express my appreciation to all colleagues and personnel from Genetics Department, and Genomic and Proteomic Unit, UCM (Madrid), IAS/CSIC (Córdoba), ITQB (Oeiras), FD/UAB (Barcelona), CGB/IBB (Vila Real) and FD/CITAB (Vila Real) for kindness and sympathy. Special thanks to Ms Mari Carmen, Dr. Raquel Chaves, Dr. Fernando Nunes, Dr. Roser Tolrá, Prof. Eunice Bacelar, Ms Helena Ferreira, Ms Donzilia Costa and Ms Ana Fraga. My sincere thanks to all.

I am glad to thank to those you provide me PhD Fellowship namely *Science and Technology Foundation*, Portugal. Without the financial support of FCT from Portugal this PhD, as it is, wouldn't exist.

I can't forget to extend my deep gratitude to colleagues that became truly friends. Thanks to all that allow me to share good moments together with respective beloved ones. Thanks to 'Those' for essential support in crucial moments.

All of this would not be easy to carry on without the support and understanding of my family and my friends. Thank you so much for being always with me, among distances! Each was always on the right time. Thank you, thank you, and thank you.

Thank you so much to all that knowingly or unknowingly gave me hope and string.

Ana

List of abbreviations

a.a.- Amino Acids

AACT1- Aluminium activated citrate transporter

Al - Aluminium

ALMT - Aluminium-Activated Malate Transporter

ALS - Al-sensitive

AltSB- Aluminium tolerance Sorghum bicolor

ART - Al resistance transcription factor

Ca^{2+}/Ca - calcium

cDNA - complementary DNA

Cys2His2 - Cysteine2/Hystidine2

DCF-DA - 2',7'-dichlorfluorescein-diacetate

DNA: deoxyribonucleic acid

dNTP – Deoxynucleotide triphosphates

EST – Expressed sequence tag

FISH - Fluorescent *in situ* hybridization

GAA- GAA microsatellite

GFP- Green Fluorescent Protein

h – Hours

Hv – *Hordeum vulgare*

Kb – Kilobase

kD - Dissociation constant

MATE - Multidrug and Toxic compound Exudation

Mg - Magnesium

mM - Milimolar

µmol - Micromolar

µM - Micrometer

Nrat1 - Nramp aluminium transporter 1

OA – Organic acid

ORF - Open reading frame

P - Probability

P - Phosphorus

pAs1 – Clone isolated from the D genome of *Ae. squarrosa*

PCR – Polymerase chain reaction

pH – Decimal logarithm of the reciprocal of the hydrogen ion activity

pI – Isoelectric point

QTL - Quantitative trait loci/locus

RACE - Rapid amplification of cDNA ends

RNA – Ribonucleic acid

RT-PCR – Reverse transcription-PCR

ROS - Reactive oxygen species

SD - Standard deviation

SNP - Single nucleotide polymorphism

STAR - Sensitive to Al rhizotoxicity

STOP - Sensitive to proton rhizotoxicity

UTR – Untranslated region

ZF - Zinc finger domains

GENERAL ABSTRACT

Aluminium (Al) toxicity is one of the most challenging features when dealing with crop production on acid soils. In Portugal, wheat holds among the cereals a special place in the sense that the majority of its production occurs on acid soils. The aim of the present research was to characterize the Portuguese bread wheat genotype Barbela 7/72/92 for Al tolerance, in low pH conditions, at the physiological and molecular levels. Through erichrome cynanin R, hematoxylin and morin assays it was possible to demonstrate that Barbela 7/72/92 is a highly Al resistant genotype and its resistance mechanism mainly relies on external Al exclusion. Schiff's reagent and Evans blue procedures also showed absence of Al induced lipid peroxidation and rare physical injuries in the Al treated roots of Barbela 7/72/92. Furthermore, better root hair growth in Barbela 7/72/92 under Al stress also indicated that this Portuguese bread wheat genotype may acquire several mechanisms of Al tolerance in order to adapt to acidic soil conditions. The histochemical assays herein presented clearly demonstrated that Al toxicity induces multiple stresses in the bread wheat's root system.

In the present investigation, three candidate genes, associated with Al tolerance in bread wheat were cloned, namely *TaMATE1* and *TaMATE2* (MATE family), and *TaSTOP1* (transcription factor family). PCR based mapping of these genes using Chinese Spring aneuploid stocks revealed that *TaMATE1*, *TaMATE2* and *TaSTOP1* are located on the long arms of homoeologous group 4, 1 and 3 chromosomes, respectively. Furthermore, the results of *TaSTOP1* mapping were also validated with *in situ* hybridization in Barbela 7/72/92.

Tissue specific expression analysis clearly revealed the higher transcript expression of *TaMATE1* and *TaALMT1* genes in the roots, whereas *TaMATE2* showed higher transcript levels in the shoot tissues. Contrarily, *TaSTOP1* seemed to be equally express in both wheat root and shoot tissues. Interestingly, homoeologue specific transcript expression clearly revealed the unequal expression of the *TaMATE1* and *TaSTOP1* homoeologues in bread wheat. Barbela 7/72/92 exhibited the preponderance of *TaMATE1-4B* transcripts followed by *TaMATE1-4D*, whereas homoeologue *TaMATE1-4A* seemed to be silenced. On the other hand, significantly higher transcript expression of *TaSTOP1-A* was noticed over *TaSTOP1-D* and *TaSTOP1-B* homoeologues. In addition our results also indicate that *TaSTOP1-A* trans-activation potential is constitutive and may not depend on the presence/absence of Al at least in yeast. The individual *TaMATE2* homoeologues transcript expression quantification was

not possible due to the high similarity among the nucleotide sequences present in the different homologues groups for this gene. However, the unequal transcript expression verified in *TaMATE1* and *TaSTOP1*, and their different response to Al toxicity indicate that homoeologue(s) of these genes, in bread wheat, may differentially contribute to Al tolerance under stress conditions.

The promoter analysis of *TaMATE1* homoeologues revealed that Al resistant genotype Barbela 7/72/92 has a Sukkula like transposon in 25 bp upstream of *TaMATE1-4B* homoeologue. Further, *TaMATE1-4B* promoter analysis in a collection of Portuguese bread wheat genotypes revealed a correlation between the presence of this transposon and Al tolerance. Furthermore, *TaALMT1* promoter sequencing also showed that Barbela 7/72/92 had a type VI promoter, whereas type I promoter was observed in Anahuac. Type VI promoter of *TaALMT1* was previously reported as being associated to high Al tolerance levels in bread wheat. These results were supported by the transcript expression of both genes, suggesting that Barbela 7/72/92 has novel alleles of *TaMATE1* and *TaALMT1* genes that may enable it to exclude the Al toxic forms in acid soils in a more efficient way. In addition, the co-localization of *TaMATE1*, *TaMATE2* and *TaSTOP1* with previously reported genomic regions implicated with Al tolerance in bread wheat on chromosomes 4 (B genome), 1 (A and B genomes) and 3 (B genome), respectively, suggest that these genes could be potential used to improve Al tolerance in bread wheat.

In conclusion, physiological and molecular studies performed in the present investigation clearly exhibited that the Portuguese bread wheat genotype Barbela 7/72/92 is highly resistant to Al stress, having developed several mechanisms to adapt to acid soils. Herein we also demonstrated that Barbela 7/72/92 possesses novel alleles, which could be used in breeding programmes to improve Al tolerance for the development of cultivars suitable for acid soils. On the other hand this study also allowed a better understanding of the molecular mechanisms involved in Al tolerance in bread wheat.

Key words: Al, aluminium, tolerance, bread wheat, *Triticum aestivum* L., candidate genes, *TaSTOP1*, *TaMATE1*, *TAMATE2*

Resumo geral

A tolerância à toxicidade ao aluminio (Al) é um dos maiores desafios para a produção agrícola em solos ácidos. Em Portugal, a cultura do trigo ocupa um lugar especial entre os cereais, ocorrendo a maioria do seu cultivo neste tipo de solos. O objectivo do presente estudo foi o de caracterizar, quer fisiológica quer molecularmente o genótipo Barbela 7/72/92 em relação à tolerância ao Al quando em presença de condições de pH baixo. As técnicas fenotípicas baseadas na ericromocianina R, na hematoxilina e na morina demonstraram que o Barbela 7/72/92 é um genótipo muito resistente ao Al e que o seu mecanismo de tolerância baseia-se principalmente na destoxificação externa deste elemento tóxico. Técnicas baseadas nos reagentes de Schiff e Evans blue também evidenciaram um baixo nível de peroxidação lipídica bem como a quase inexistência de danos físicos nas raízes de Barbela 7/72/92 tratadas com Al. De salientar, que o crescimento de pêlos radiculares presentes no Barbela 7/72/92 quando na presença de Al poderá indicar que este genótipo de trigo mole activa diversos mecanismos de resposta à toxicidade deste elemento. As técnicas histoquímicas utilizadas neste estudo demonstraram claramente que a toxicidade pelo Al induz múltiplos efeitos nas raízes de trigo mole.

Na presente investigação, três genes candidatos, para a tolerância ao Al em trigo hexaplóide, foram clonados nomeadamente o *TaMATE1* e o *TaMATE2* (membros da família MATE) e o *TaSTOP1* (membro da família de fatores de transcrição). O mapeamento através do recurso a linhas aneuploides de trigo da cultivar Chinese Spring, revelou que o *TaMATE1*, o *TaMATE2* e o *TaSTOP1* estão localizados no braço longo dos cromossomas 4, 1 e 3, respectivamente. No caso do gene *TaSTOP1* a localização no cromossoma 3 foi igualmente confirmada por hibridação *in situ* no genótipo Barbela 7/72/92.

A análise da expressão dos genes *TaMATE1* e *TaALMT1* revelou que ambos os transcriptos existem em maior abundância nas raízes comparativamente aos rebentos, contrariamente ao que foi verificado para o gene *TaMATE2*. Em relação ao gene *TaSTOP1*, o gene expressou-se equitativamente em ambos os tecidos (raízes e rebentos). A análise da expressão específica para os genes *TaMATE1* e *TaSTOP1* revelou claramente a transcrição não uniforme destes genes, em ambos os tecidos de trigo hexaplóide. A expressão do homólogo *TaMATE1-4B* foi preponderante à expressão do *TaMATE1-4D* em Barbela 7/72/92. Quanto ao homólogo *TaMATE1-4A* a sua expressão parece estar silenciada. Para o gene *TaSTOP1*, os níveis de expressão do

homólogo *TaSTOP1-A* encontraram-se significativamente elevados em relação aos encontrados para *TaSTOP1-B* e *TaSTOP1-D*, em ambos os tecidos (raízes e rebentos). Os nossos resultados indicaram igualmente que o gene *TaSTOP1-A* possui um potencial de trans-activação constitutivo e que pode não depender da ausência/ presença do stress por Al. Em relação ao *TaMATE2*, não foi possível a quantificação individual de cada um dos homólogos devido à elevada similaridade verificada entre as suas sequências nucleotídicas. A expressão desigual encontrada para os genes *TaMATE1* e *TaSTOP1*, bem como a diferente resposta à toxicidade por Al indica que os homólogos destes genes, em trigo mole, parecem de forma diferencial na presença de Al.

A análise dos promotores dos homólogos do gene *TaMATE1* revelou que no genótipo Barbela 7/72/92, a 25 bp a jusante, existe um transposão do tipo Sukkula no homólogo *TaMATE1-4B*. A subsequente análise desta região numa coleção de trigo mole Portugueses revelou uma correlação positiva entre a presença deste transposão e a tolerância ao Al. De igual modo, a sequenciação do promotor do gene *TaALMT1,* no Barbela 7/72/92 e no Anahuac, revelou a presença dos promotores do Tipo VI e I, respectivamente. O promotor *ALMT1* tipo VI foi previamente descrito como tendo uma correlação positiva com a tolerância ao Al em trigo. Estes resultados são suportados pelas análises de expressão de ambos genes, sugerindo que o Barbela 7/72/92 possui novos alelos para os genes referidos permitindo a este genótipo um melhor desempenho quando na presença do stress pelo Al através da sua exclusão. A co-localização do *TaMATE1*, *TaMATE2* e *TaSTOP1* com regiões genómicas implicadas na tolerância ao Al nos cromossomas 4 (genoma B), 1 (genomas A e B) e 3 (genoma B), respectivamente, sugerem que estes genes podem ser um potencial alvo para a tolerância ao Al neste cereal.

Em conclusão, os estudos fisiológicos e moleculares realizados neste estudo claramente demonstram que o genótipo de trigo mole Português Barbela 7/72/92 é altamente tolerante ao stress pelo Al, o qual desenvolveu diversos mecanismos de forma a adaptar-se aos solos ácidos onde é cultivado. Neste trabalho, demonstramos que o Barbela 7/72/92 possui novos alelos para múltiplos genes para a tolerância ao Al, os quais poderão vir a ser posteriormente utilizados em programas de melhoramento de trigo.

Palavras chave: Al, aluminio, tolerância, trigo mole, *Triticum aestivum* L., genes candidatos, *TaSTOP1*, *TaMATE1*, *TaMATE2*

Table of contents

INTRODUCTION

Adverse environmental conditions cause more than 50% crop loss worldwide (Bray *et al.*, 2000). Nevertheless the degree of losses due to abiotic stresses varies and depends upon the intensity and the stress duration. For instance, a large proportion of the earth's land is acidic, occurring mainly in the northern (cold and temperate) and in the southern (tropical) belts (Figure 1A), and only 4.5% of the area belonging to these acidic soils is used for arable crops (von Uexküll and Mutert, 1995). In Portugal, most of the regions have acidic soil with pH lower than 5.5 (Figure 1B). However, in Northern Portugal, particularly in the inner region of Trás-os-Montes and Alto Douro, the majority of soils are of leptosols and cambisols types and characterized as being very acidic with pH values as low as 4.5 (Coutinho, 1989).

In all countries where acidic soils are widespread, the sub-optimal plant performance can usually be attributed to a combination of abiotic stress factors, such as, aluminium (Al), manganese (Mn) and protons (H). In the earth crust, Al is the third most abundant element after oxygen and silicon. Thus, it is a major constituent of the soils and under low pH it dissolves in various ionic forms [Al^{+3}, $Al(OH)^{+2}$ and $Al(OH)_2^{+}$]. Among these ionic forms, Al^{3+} is the most toxic for rhizosphere of plants.

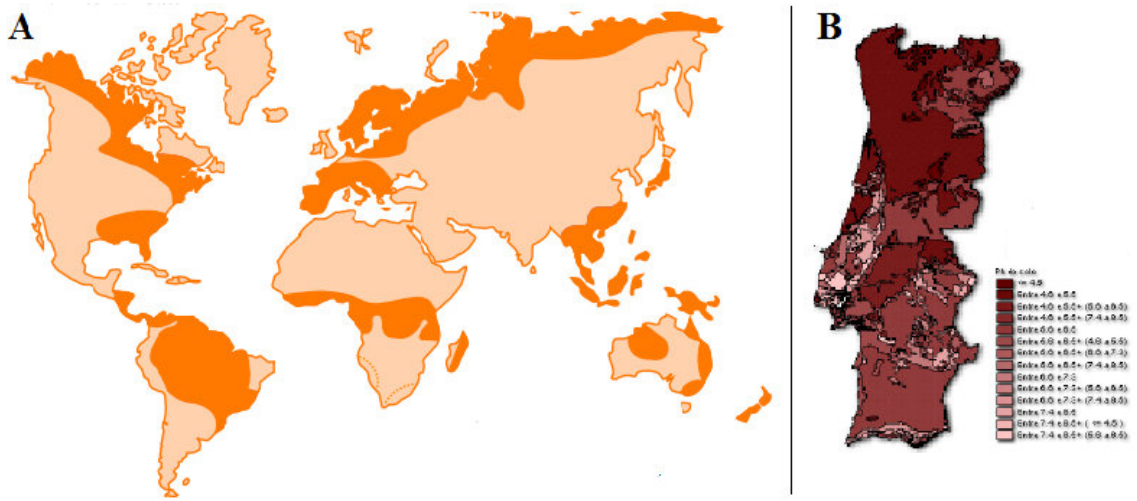


Figure 1. Distribution of acid soils. (A) Global areas of predominant acid soils are highlighted in colour. (B) Portugal soil pH gradient Note: light colour depicts higher soil pH whereas dark colour shows low pH. (Source: von Uexkúll and Mutert, 1995 and http://www.iambiente.pt/)

Agronomically, lime (calcium carbonate) application is used to change the soils' pH and it is considered to be the best way to correct soils acidity. However, this is not always economically or physically feasible because the amount of lime required greatly depends upon the soil's pH and texture. Thus, the development of Al-resistant cultivars is the most durable and eco-friendly approach in order to enhance food production in acid soils.

Al toxicity in wheat and its effect

During 1920s, scientists reported the first evidence of visual differences among wheat cultivars grown in acid soils in Brazil, although it had not yet been established that the acid soils themselves were causing these differences. Poor plant growth and a burning effect in wheat on acid soils were observed and it was termed as "crestamento" which in Portuguese literally means 'burning' or 'toasting'. In 1925, crestamento tolerant Brazilian cultivars, such as Fronteira, Surpresa, Minuano, Jesuita and Guarany were developed from crosses between the crestamento-tolerant Alfredo Chaves and Polyssu cultivars (Rajaram *et al.*, 1988). By 1942, "crestamento" was found to be caused by Al toxicity present in these acid soils.

The primary response to Al stress in roots, particularly in the cell division and elongation zone has been described in a wide variety of plant species (Bergmann, 1992). Al phytotoxicity on the dynamics of root growth inhibition is well known resulting in a reduced and damaged root system and subsequently limited water and mineral nutrient uptake in different taxa (Kochian *et al.*, 2004). In addition, Al also causes extensive damage in other cellular components of wheat such as inhibition of DNA synthesis (Wallace and Anderson, 1984), alteration of cell membrane potential (Kinraide, 1988), and reduction of root apex H^+ efflux (Ryan *et al.*, 1992).

In wheat, toxic levels of Al begin to inhibit root elongation within 1-3 hours from its exposure (Ownby and Popham, 1990; Ryan *et al.*, 1992) and within 6 hours, Al accumulates at higher concentration (about 7-8 fold) in the root's terminal region (2 mm) of sensitive cultivar in comparison to tolerant cultivars (Rincón and Gonzales, 1992). However, the remarkable differences in the Al uptake by the roots of Al-tolerant and

sensitive wheat genotypes could be observed within 4 hours of exposure to Al stress (Delhaize *et al.*, 1993a) correlating the degree of Al sensitivity with the differential behaviour in relation to Al accumulation in the growing root tissue in wheat (Samuels *et al.*, 1997). An Al uptake kinetics study revealed that the Al uptake by wheat roots is biphasic under Al stress, consisting of an initial rapid uptake phase (first half hour) followed by a linear uptake phase (Zhang and Taylor, 1989 and 1990). Further analysis of Al uptake during the linear phase also suggested two mechanisms: metabolism dependent Al binding in the apoplasm and permeation of the cell membrane (Zhang and Taylor, 1989 and 1990).

Phenotyping for Al toxicity tolerance

An accurate and precise phenotyping is not only the essential component in any crop improvement programme, but also has paramount importance to dissect the genetic architecture of quantitatively inherited complex traits. From the breeder's point of view, phenotyping technique should be flexible, simple and rapid which also favoured the costs within a reasonable limit. In the second half of the last century, the field of Al tolerance phenotyping in plants has evolved a number of techniques that vary from the field evaluation to the soil bioassays and nutrient solution culture. By far the most common phenotyping medium for Al tolerance is the nutrient solution culture, which provides an easy access to the root systems, tight control over nutrient availability and pH, and non-destructive measurements of tolerance (Carver and Ownby, 1995).

In the past a numerous of histochemical assays such as hematoxylin, morin, Schiff's reagent, eriochrome cyanine R and Evans blue staining have been reported to investigate the Al toxicity in different plants. Hematoxylin and morin assays are used to localize Al in the plant tissues. The hematoxylin assay is based on the formation of colored complex between hematoxylin and the root-bounded Al (Polle *et al.*, 1978); whereas morin is a fluorochrome which forms a fluorescent complex with Al (Eticha *et al.*, 2005). Eriochrome cyanine R staining has been extensively used for the measurement of root re-growth under Al stress and shows that if root apical meristem is irreversibly damaged, the root tips

remained intensively stained with purple colour, whereas the part of the root which grows after exposure to Al stress remains unstained (Aniol, 1995). In general, Schiff's reagent staining technique is widely used for the assessment of Al-induced lipid peroxidation in the plants' root system detecting aldehyde functions that are originated from the peroxidation of the membrane's lipids and are bound to the membrane's proteins. Schiff's reagent procedure was originally performed to investigate the lipid peroxidation in liver tissue of bromobenzene-intoxicated mice (Pompella *et al.*, 1987) and is based on the intensity of the pink colour development (Yamamoto *et al.*, 2002). In addition, the loss of plasma membrane integrity, induced by the toxic Al concentration, could be visualized in the roots by Evans blue staining technique (Yamamoto *et al.*, 2001). Among these histochemical assays, hematoxylin and eriochrome cyanine R staining are widely used to screen for Al tolerance in wheat because of their consistent reliability in the staining patterns (Delhaize *et al.*, 1993b; Luo and Dvořák, 1996; Raman *et al.*, 2005; Zhou *et al.*, 2007; Cai *et al.*, 2008; Martins-Lopes *et al.*, 2009; Navakode *et al.*, 2009; Raman *et al.*, 2010). However, the Al tolerance mechanisms identified in the histochemical assays could then be confirmed in the soil bioassay and such tandem phenotyping approach can be helpful for the better understanding of both seedling as well as adult plant tolerance, keeping in mind that the ultimate goal is to enhance the economic yield under acid soils.

Genetics of Al tolerance in wheat

On the basis of segregation pattern observed in F_2 generation derived from a cross between Al tolerant and sensitive genotypes, Beckman (1954) suggested the presence of a single dominant gene for Al tolerance in bread wheat. Similarly, Kerridge and Kronstad (1968) also reported that a moderately Al-tolerant cultivar, Druchamp differed from a sensitive cultivar, Brevor, by a single gene governing seedling root growth under Al stress. Subsequently, in the last quarter of the 20th century, the wider recognition of Al toxicity as the predominant growth-limiting factor in acid soils has swung the pendulum of attention to Al tolerance and its genetic control (Carver and Ownby, 1995). Contrarily to earlier speculations, wide genetic range of Al tolerance observed in the wheat germplasm

indicated that Al tolerance inheritance was much more complex (Lafever *et al.*, 1977). However, most of studies suggested that Al tolerance in bread wheat is mainly governed by one to two dominant genes (Choudhry, 1978; Camargo 1981 & 1984; Campbell and La fever, 1981).

Chinese Spring aneuploids stocks lacking a specific chromosome or chromosome arms helped the geneticist to locate the Al tolerance genes more precisely on individual chromosomes in bread wheat. Although Chinese Spring is considered as moderately Al tolerant, surprisingly several genes/loci were identified and located on chromosome arms 2DL, 3DL, 4BL, 4DL, 6AL, 7AS and chromosome 7D (Aniol and Gustafson, 1984). In addition, recent findings in DNA technology and in the development of robust statistical analysis methods facilitated the understanding of the gene action and their relative contributions in governing Al tolerance in different plant species. In the previous decade, several genomic regions located on different chromosomes and their relative contribution to Al tolerance in bread wheat have been demonstrated in numerous studies (Luo and Dvořák, 1996; Raman *et al.*, 2005; Zhou *et al.*, 2007; Cai *et al.*, 2008; Navakode *et al.*, 2009; Raman *et al.*, 2010).

Processes of Al detoxification in plants

The physiological basis of Al tolerance has been extensively investigated in several crop and model plant species and these mechanisms can be broadly divided into two categories: (A) external detoxification of Al or exclusion of Al or Al resistance and (B) internal detoxification of Al or inclusion of Al or Al tolerance mechanisms.

A. Mechanisms of external detoxification of Al

The external detoxification of Al includes all mechanisms that limit the Al uptake by roots and also reduce the harmful interactions of toxic forms of Al with sensitive sites in the apoplast, such as the plasma membrane and the cell wall. The evidence of these mechanisms implicates cell wall chemistry, release of mucilage, efflux of organic anions and secondary metabolites from root apices, and the active export of Al from cells

(Delhaize *et al.*, 2012). Among the Al exclusion strategies, efflux of organic anions from root apices is one of the most demonstrated by strong genetic and molecular studies in different plant species, including wheat. Malate and citrate has been shown to be constitutively released from the roots of different Al tolerant wheat genotypes, but the efflux of these anions is triggered by external Al (Delhaize *et al.*, 1993b; Ryan *et al.*, 2009). These anions form harmless complexes with Al in the apoplast that protect the sensitive root apex and reduce uptake into the roots.

Root mucilage is a gelatinous polysaccharide which is exuded from the outer layers of the root cap. Horst and collaborators (1982) observed a higher Al sensitivity of cowpea roots without mucilage. They correlated the level of Al contents in the root tip tissues with the differential amount of root mucilage indicating the protective function of the mucilage against Al injury. Similarly, Henderson and Ownby (1991) also found association between the amounts of mucilage produced by the wheat roots and the level of Al tolerance, and suggested that the mucilage acted as a diffusion barrier to Al or through the concentration of organic acids that chelated toxic Al ions. However, the mucilage mechanism of Al exclusion in wheat is not unequivocal, but it cannot be ruled out.

Besides these factors, there are also indications that phosphate (Pi) exudation could be an important Al exclusion mechanism through the formation of Al-phosphate complexes, and protonation that might increase pH in the acidic rhizosphere (Taylor, 1991). Pellet and co-workers (1997) also observed a constitutive phosphate exudation from the root apex of Al tolerant genotype Atlas seedlings and suggested that Al tolerance in wheat can be mediated by multiple exclusion mechanisms. So far no conclusive data associating Pi exudation with Al tolerance have been presented in different plant species. Similarly, recent studies also indicate that phosphate efflux does not play a significant role in Al tolerance in wheat (Tang *et al.*, 2002).

Another hypothesis for differential tolerance to Al in plants relates to the negative electrical charges present at the cell surface. In general, it is usually believed that binding of Al to charged sites on the cell surface is a prerequisite for its uptake and toxicity. Earlier

reports exhibited the negative correlation between root cation exchange capacity (CEC) and Al tolerance in genotypes of some plant species, including wheat cultivars (Foy *et al.*, 1967; Blamey *et al.*, 1990; Blamey *et al.*, 1993). Contrarily, Kinraide and collaborators (1992) suggested that such mechanisms do not play significant role in differential Al tolerance in wheat.

B. **Mechanisms of internal detoxification of Al**

On the other hand, internal detoxification mechanisms allow plants to cope with Al once it enters the cell and rely on either on the formation of harmless complexes with organic ligands, by sequestering them into organelles, or by a rapid repair of any lesions incurred, including those resulting from oxidative stress (Delhaize *et al.*, 2012). Some Al accumulator species such as tea (*Camelia sinensis* (L.) Kuntze), buckwheat (*Fagopyrum esculentum* Moench) and Hydrangea (*Hydrangea macrophylla* Thunb.) are good examples where internal detoxification mechanisms are acting, as they contain Al in their leaves without adverse symptoms and even their growth can be stimulated by its presence (Tsuji *et al.*, 1994; Ma *et al.*, 1997; Zheng *et al.*, 1998; Ma and Hiradate, 2000; Tolrà *et al.*, 2011).

Molecular aspects of Al tolerance in wheat

The advent of novel technologies in the field of molecular biology has not only greatly improved the ability to understand the Al tolerance mechanism in plants but also provided powerful means to supplement and complement the conventional methods of crop improvement. DNA based techniques allowed to dissect the complex traits, particularly both the locations and numbers of genomic regions, and their relative contribution to a particular phenotype (Vinh and Paterson, 2005). Recent genome wide association and QTL mapping studies not only enabled to confirm the involvement of previously reported loci (Aniol and Gustafson, 1984), but also identified the new genomic regions for Al tolerance in bread wheat which were previously not possible to detect with the traditional approaches (Luo and Dvořák, 1996; Raman *et al.*, 2005; Zhou *et al.*, 2007; Cai *et al.*, 2008; Navakode *et al.*, 2009; Raman *et al.*, 2010). However, most of the studies reported the use of random

DNA markers associated with Al tolerance in wheat, mainly due to the unavailability of *priori* candidate genes in bread wheat.

During the last century, collaborative research, involving the Brazilian wheat breeding programme and CIMMYT (International Maize and Wheat Improvement Center) has made considerable strides in developing Al tolerant lines/cultivars which were distributed worldwide, and helped in the development of segregating populations and near-isogenic lines for Al tolerance in wheat. Delhaize and collaborators (1993a) demonstrated that Al resistance in a wide range of bread wheat genotypes relied on the Al-activated efflux of malate from root apices and it was mainly governed by single locus. Almost one decade later, the same group cloned the first wheat transporter gene associated with Al resistance and it was named *ALMT1* (Aluminium Activated Malate Transporter) (Sasaki *et al.*, 2004) being the first Al resistance gene identified in any plant species (Ryan *et al.*, 2011). *ALMT1* gene has been extensively studied for Al tolerance in different plant species. Transgenic barley, wheat and Arabidopsis over-expressing wheat *ALMT1* enhanced malate efflux and exhibited greater relative root growth by 20-, 8-, and 4-fold, respectively than the control (Delhaize *et al.*, 2004; Pereira *et al.*, 2010; Ryan *et al.*, 2011). In addition, the knock-out mutant of *ALMT1* gene also clearly demonstrated that *ALMT1* plays an important role in Al detoxification by releasing the malate efflux in Arabidopsis roots (Hoekenga *et al.*, 2006).

In bread wheat *ALMT1* gene is located on the long arm of chromosome 4D and the absence or loss of this gene coincided with the loss of both Al tolerance and Al-activated malate efflux (Raman *et al.*, 2005). Promoter analysis showed the specific pattern of variation (types I–VI) in the upstream region of the *ALMT1* gene which was correlated with the relative level of *TaALMT1* expression in wheat (Sasaki *et al.*, 2006). Type I pattern revealed to have the simplest structure, while the other patterns presented blocks of sequence in duplicates and triplicates. These tandem repeats present in the *TaALMT1* promoter generated significantly greater expression of a GFP (encoding green fluorescent protein) reporter gene in callus and regenerated roots (Ryan *et al.*, 2010).

It is well known that in sorghum and barley, Al resistance mainly relies on the citrate efflux. In the previous decade, the fine mapping of the major loci implicated in citrate efflux in each of these species, allowed the identification of a second transporter gene (*SbMATE* in sorghum and *HvAACT1* in barley) for Al tolerance which belongs to the MATE (Multidrug and Toxic Compound Exudation) family (Furukawa *et al.*, 2007; Magalhães *et al.*, 2007). It is interesting to note that a tourist-like miniature inverted repeat transposable elements (MITEs) in the upstream of *SbMATE* and a 1023 bp insertion in the 5'UTR of *HvAACT1* were observed in the Al-tolerant genotypes of sorghum and barley, respectively (Magalhães *et al.*, 2007; Fujii *et al.*, 2012).

Although, Al resistance associated with citrate efflux has also been reported in bread wheat genotype Carazinho and the analysis of segregating populations for citrate efflux also suggested that it was governed by a single locus, located on long arm of chromosome 4B (Ryan *et al.*, 2009). To date, no other transporter except *TaALMT1* has been cloned.

Like other abiotic stresses, model plant species such as *Arabidopsis* and rice provided the evidence that regulatory genes also play an important role in Al tolerance in plants. Recently, the zinc finger transcription factors, *STOP1* (Sensitive To Proton Rhizotoxicity1) in Arabidopsis and *ART1* (Aluminium Resistance Transcription factor 1) in rice were identified by mutant analysis (Iuchi *et al.*, 2007; Yamaji *et al.*, 2009), sharing a significant sequence similarity. Transcriptome analysis under Al stress revealed that these transcription factors coordinate the expression of Al tolerance genes in the Arabidopsis Stop1 (Sawaki *et al.*, 2009) and rice ART1 mutants (Yamaji *et al.*, 2009). So far none of the regulatory gene has been reported in bread wheat.

Present Investigation: Global and Portugal Perspectives

Nowadays, directly or indirectly, wheat is central to the food supply of over half the world's population. Globally, it is grown on more land area than any other commercial crop due to its ability to grow in a wide range of climatic environments and geographic regions (from near arctic regions to equator and from sea level to plains of Tibet) (Dixon *et al.*,

2009). Wheat ranks second, after rice, as the main human food crop. Worldwide, the European Union (EU) holds the first position in total annual wheat production followed by China, India and USA. Wheat in EU occupied 11.84% of the global acreages accounting for 19.89% of the world's wheat production, according to 2011 FAO's data.

Among cereals wheat is one of the major crop in Portugal. The majority of the wheat is sown in October - November and harvested in the months of June and July of the following year. It is a typical dryland crop and survives in most parts of Portugal on rainfall which mainly occurs in the winter between November and March. However, wheat is also grown in some areas with assured irrigation. In 2011, Portugal harvested about 58.63 thousand tonnes of wheat from an estimated area of 42.89 thousand hectares (http://faostat.fao.org). Among cereals, wheat is the most preferred staple food in Portugal, being the annual consumption *per* capita of 105 kg. For the fulfilment of domestic wheat demand, Portugal relies on import from other countries. Thus, the low productivity of wheat in Portugal needs to be improved in order to be self-sufficient.

As mentioned in the beginning of this chapter, most of the Portuguese soils are acidic in nature; therefore, Al toxicity seems to be one of the major constraints for wheat productivity in Portugal. Barbela is a well known Al tolerant Portuguese bread wheat landrace which has been grown for more than one century (Lapa, 1865; Coutinho, 1884). Barbela may be a good source of Al tolerance genes, especially since it has not yet been well utilized in wheat improvement (Hede *et al.*, 2001). Compared with rice, the molecular basis of Al tolerance in bread wheat is still poorly understood, because of its complex genomic architecture. In model plant species such as Arabidopsis and rice, numerous candidate genes including regulatory and structural genes for Al tolerance have been cloned and functionally validated. However, the identification of such candidate genes in bread wheat except *TaALMT1* is far behind than other counterparts in cereal crops.

Therefore, keeping in view the importance of bread wheat in the Portuguese diet, the objectives of the present investigation were:

1. To characterize the bread wheat genotype Barbela 7/72/92 for Al tolerance, previously described to be derived from the Barbela landrace.
2. To clone novel candidate genes for Al tolerance in Barbela 7/72/92 and,
3. To characterize the cloned genes at the molecular level, in order to make possible their use for wheat improvement on acid soils in the future.

References:

Aniol, A.M. (1995). Physiological aspects of aluminium tolerance associated with the long arm of chromosome 2D of the wheat (*Triticum aestivum* L.) genome. *Theor Appl Genet*, 91(3): 510-516.

Aniol, A. and Gustafson, J.P. (1984). Chromosome location of genes controlling aluminum tolerance in wheat, rye, and triticale. *Can J Genet Cytol*, 26: 701-705.

Beckman, I. (1954). Sobre o cultivo e melhoramento do trigo (triticum vulgare, Vill) no sul do Brasil. *Agron SulRioGrandense*, 1:64-72.

Bergmann, W. (1992). Nutiritional disorders of plants: Development, visual and analytical diagnosis. Gustav Fischer Verlag, Jena, Germany.

Blamey, F.P.C.; Edmeades, D.C. and Wheeler, D.M. (1990). Role of root cation exchange capacity in different aluminum tolerance of Lotus species. *J. Plant Nutr*, 13:729-744.

Blamey, F.P.C.; Robinson, N.J. and Asher, C.J. (1993). Interspecific differences in aluminium tolerance in relation to root cation-exchange capacity. In Genetic Aspects of Plant Mineral Nutrition. *Develop Plant Soil Sci*, 50: 91-96.

Bray, E.A.; Bailey-Serres, J. and Weretilnyk, E. (2000). Responses to abiotic stresses. In: Gruissem W, Buchannan B, Jones R, eds. *Responses to abiotic stresses.* Rockville, MD: *American Society of Plant Physiologists*, 1158-1249.

Cai, S.; Bai, G.H. and Zhang, D. (2008). Quantitative trait loci for aluminum resistance in Chinese wheat landrace FSW. *Theor Appl Genet*, **117**(1): 49-56.

Camargo, C.E.O. (1981). Wheat improvement. I. The heritability of tolerance to aluminum toxicity. *Bragantia*, 40: 33-45. (In Portuguese).

Camargo, C.E.O. (1984). Wheat improvement. VI. Heritability studies on aluminum tolerance using three concentrations of aluminum in nutrient solutions. *Bragantia*, 43: 279-291. (In Portuguese).

Campbell, L.G. and Lafever, H.N. (1981). Heritability of aluminum tolerance in wheat. *Cereal Res. Comm*, 9: 281-287.

Carver, B.F. and Ownby, J.D. (1995) Acid soil tolerance in wheat. *Adv Agron*, **54**:117-173.

Choudhry, M.A. (1978). Genetic differences for aluminum tolerance in five wheat crosses. Agr. Abstr. Madison, *Am Soc Agr*, 151.

Coutinho, A.X.P. (1884). Os fenos espontâneos e as palhas de trigo, em Portugal. Estudos Agronómicos, Lisboa: Imprensa Nacional.

Coutinho, J.F. (1989). Acidez do solo e toxicidade do aluminio. Tese de Doutoramento, UTAD, Vila Real, pp. 709.

Delhaize, E.; Craig, S.; Beaton, C.D.; Bennet, R.J.; Jagadish, V.C. and Randall, P.J. (1993a). Aluminum tolerance in wheat (*Triticum aestivum* L.) (I. Uptake and distribution of aluminum in root apices). *Plant Physiol*, 103(3): 685-693

Delhaize, E.; Ryan, P.R. and Randall, P.J. (1993b). Aluminum tolerance in wheat (*Triticum aestivum* L.) (II. Aluminum-stimulated excretion of malic acid from root apices). *Plant Physiol*, 103(3): 695-702.

Delhaize, E.; Ryan, P.R.; Hebb, D.M.; Yamamoto, Y.; Sasaki, T. and Matsumoto, H. (2004). Engineering high-level aluminum tolerance in barley with the ALMT1 gene. *Pro Nat Acad Sci USA*, 101: 15249-15254.

Delhaize, E.; Ma, J.F. and Ryan, P.R. (2012). Transcriptional regulation of aluminium tolerance genes. *Trends Plant Sci*, 17(6): 341-348.

Dixon, J.; Braun, H.J. and Crouch J. (2009). Overview: Transitioning wheat research to serve the future needs of the developing world. In Wheat facts and futures 2009, eds. Dixon, J., Braun, H.J., Kosina, P., and Crouch J., pp 1-25. Mexico, D.F.: CIMMYT.

Eticha, D.; Stass, A. and Horst, W.J. (2005). Localization of aluminium in the maize root apex: can morin detect cell wall-bound aluminium? *J Exp Bot*, 56(415): 1351-1357.

Foy, C.D.; Armiger, W.H.; Fleming, A.L. and Lewis, C.F. (1967). Differential tolerance of cotton varieties to an acid soil high in exchangeable aluminum. *Agron J*, 29:415-418.

Fujii, M.; Yokosho, K.; Yamaji, N.; Saisho, D.; Yamane, M.; Takahashi, H.; Sato, K.; Nakazono, M. and Ma, J.F. (2012). Acquisition of aluminium tolerance by modification of a single gene in barley. *Nat Commun*, 3: 713.

Furukawa, J.; Yamaji, N.; Wang, H.; Mitani, N.; Murata, Y.; Sato, K.; Katsuhara, M.; Takeda, K. and Ma, J.F. (2007). An aluminum-activated citrate transporter in barley. *Plant Cell Physiol*, 48: 1081-1091.

Hede, A.R.; Skovmand, B. and López-Cesati, J. (2001). Acid soils and aluminum toxicity. In Application of Physiology in Wheat Breeding, ed. MP Reynolds, JI Ortiz-Monasterio, A McNab, pp. 172-82. Mexico:CIMMYT

Henderson, M. and Ownby, J.D. (1991). The role of root cap mucilage secretion in aluminum tolerance in wheat. *Biochem Curr Topics Plant Physiol*, 10: 134-141.

Hoekenga, O.A.; Maron, L.G.; Piñeros, M.A.; Cançado, G.M.A.; Shaff, J.; Kobayashi, Y.; Ryan, P.R.; Dong, B.; Delhaize, E.; Sasaki, T.; Matsumoto, H.; Yamamoto, Y.; Koyama, H. and Kochian, L.V. (2006). AtALMT1, which encodes a malate transporter, is identified as one of several genes critical for aluminum tolerance in Arabidopsis. *Proc Natl Acad Sci USA*, 103(25):9738-9743.

Horst, W.J.; Wagner, A. and Marschner, H. (1982). Mucilage protects root meristems from aluminium injury. *Z Pflanzenphysiol*, 105: 435-444.

Iuchi, S.; Koyama, H.; Iuchi, A.; Kobayashi, Y.; Kitabayashi, S.; Ikka, T.; Hirayama, T.; Shinozaki, K. and Kobayashi, M. (2007). Zinc finger protein STOP1 is critical for proton tolerance in Arabidopsis and coregulates a key gene in aluminum tolerance. *Proc Nat Acad Sci USA*, 104: 9900-9905.

Kerridge, P.C. and Kronstad, W. E. (1968). Evidence of genetic resistance to aluminium toxicity in wheat (*Triticum aestivum vill.*, Host). *Agron J*, 60: 710-711.

Kinraide, T.B. (1988). Proton extrusion by wheat roots exhibiting severe aluminium toxicity symptoms. *Plant Physiol*, 88: 418-423.

Kochian, L.V.; Hoekenga, O.A. and Piñeros, M.A. (2004). How do crop plants tolerate acid soils? Mechanisms of aluminum tolerance and phosphorous efficiency. *Annu Rev Plant Biol*, 55:459-93.

Lafever, H.N.; Campbell, L.G. and Foy, C.D. (1977). Differential response of wheat cultivars to Al. *Agron J*, 69: 563-568.

Lapa JIF. (1865). Memória sobre o estudo industrial e químico dos trigos Portugueses. Lisboa, Portugal: Tipografia Academia.

Luo, M.C. and Dvořák, J. (1996). Molecular mapping of an aluminum tolerance locus on chromosome 4D of Chinese Spring wheat. *Euphytica*, 91(1): 31-35

Ma, J.F.; Hiradate, S.; Nomoto, K.; Iwashita, T. and Matsumoto, H. (1997). Internal detoxification mechanism of al in Hydrangea (Identification of Al form in the leaves). *Plant Physiol*, 113(4): 1033-1039.

Ma, J.F. and Hiradate, S. (2000). Form of aluminium for uptake and translocation in buckwheat (*Fagopyrum esculentum* Moench). *Planta*, 211(3): 355-360.

Magalhães, J.V.; Liu, J.; Guimarães, C.T.; Lana, U.G.; Alves, V.M.; Wang, Y.H.; Schaffert, R.E.; Hoekenga, O.A.; Piñeros, M.A.; Shaff, J.E.; Klein, P.E.; Carneiro, N.P.; Coelho, C.M.; Trick, H.N. and Kochian, L.V. (2007). A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. *Nat Genet*, 39: 1156-1161.

Martins-Lopes P.; Maçãs B. and Guedes-Pinto H (2009). Portuguese bread wheat germplasm evaluation for aluminium tolerance. *Cereal Res Commun*, 37(2): 179-188.

Navakode, S.; Weidner, A.; Lohwasser, U.; Röder M.S. and Börner, A. (2009). Molecular mapping of quantitative trait loci (QTLs) controlling aluminium tolerance in bread wheat. *Euphytica*, **166**(2): 283-290.

Ownby, J.D. and Popham H.R. (1990). Citrate reverses the inhibition of wheat root growth caused by aluminum. *J Plant Physiol*, 135(5): 588–591.

Pellet, D.M.; Papernick, L.A.; Jones, D.L.; Darrah, P.R.; Grunes, D.L. and Kochian, L.V. (1997). Involvement of multiple aluminium exclusion mechanisms in aluminium tolerance in wheat. *Plant Soil*, 192: 63-68.

Pereira, J.F.; Zhou, G.; Delhaize, E.; Richardson, T. and Ryan, P.R. (2010). Engineering greater aluminium resistance in wheat by over-expressing TaALMT1. *Ann Bot*, 106; 205-214.

Polle, E.; Konzak, C.F. and Kittrick, J.A. (1978). Visual detection of aluminium tolerance levels in wheat by hematoxylin staining of seedling roots. *Crop Sci*, 18: 823-827.

Pompella, A.; Maellaro, E.; Casini, A.F. and Comporti, M. (1987). Histochemical detection of lipid peroxidation in the liver of bromobenzene–poisoned mice. *Am J Pathol*, 129: 295-301.

Rajaram, S.; Pfeiffer, W. and Singh, R. (1988). Developing bread wheat for acid soils through shuttle breeding. In Kohli and Rajaram eds. Wheat breeding in acid soils. Review of Brazilian/CIMMYT Collaboration, 1974-1986 pp: 51-58.

Raman, H.; Stodart, B.; Ryan, P.R.; Delhaize, E.; Emebiri, L.; Raman, R.; Coombes, N. and Milgate, A. (2010). Genome-wide association analyses of common wheat (*Triticum aestivum* L.) germplasm identifies multiple loci for aluminium resistance. *Genome*, **53**(11): 957-966.

Raman, H.; Zhang, K.R.; Cakir, M.; Appels, R.; Garvin, D.F.; Maron, L.G.; Kochian, L.V.; Moroni, J.S.; Raman, R.; Imtiaz, M.; Drake-Brockman, F.; Waters, I.; Martin, P.; Sasaki, T.; Yamamoto, Y.; Matsumoto, H.; Hebb, D.M.; Delhaize, E. and Ryan, P.R. (2005). Molecular characterization and mapping of ALMT1, the aluminium-tolerance gene of bread wheat (*Triticum aestivum* L.). *Genome*, **48**: 781-791.

Rincón, M. and Gonzales, R.A. (1992). Aluminum partitioning in intact roots of aluminum-tolerant and aluminum-sensitive wheat (*Triticum aestivum* L.) cultivars. *Plant Physiol*, 99(3): 1021-1028.

Ryan, P.R.; Raman, H.; Gupta, S.; Horst, W.J. and Delhaize, E. (2009). A second mechanism for aluminum resistance in wheat relies on the constitutive efflux of citrate from roots. *Plant Physiol*, 149: 340-351.

Ryan, P.R.; Tyerman, S.D.; Sasaki, T.; Furuichi, T.; Yamamoto, Y.; Zhang, W.H. and Delhaize, E. (2011). The identification of aluminium-resistance genes provides opportunities for enhancing crop production on acid soils. *J Exp Bot*, 62(1): 9-20.

Ryan, P.R.; Raman, H.; Gupta, S.; Sasaki, T.; Yamamoto, Y. and Delhaize, E. (2010). The multiple origins of aluminium resistance in hexaploid wheat include *Aegilops tauschii* and more recent *cis* mutations to TaALMT1. *Plant J*, 64: 446-455.

Ryan, P.R.; Shaff, J.E. and Kochian, L.V. (1992). Correlation among ionic currents, ion fluxes, and root elongation in aluminum-sensitive and aluminum-tolerant wheat cultivars. *Plant Physiol*, 99(3): 1193-1200.

Samuels, T.D.; Kucukakyuz, K. and Rincon-Zachary, M. (1997). Al partitioning patterns and root growth as related to al sensitivity and al tolerance in wheat. *Plant Physiol*, 113(2): 527-534.

Sasaki, T.; Yamamoto, Y.; Ezaki, B.; Katsuhara, M.; Ahn, S.J.; Ryan, P.R.; Delhaize, E. and Matsumoto, H. (2004). A wheat gene encoding an aluminum-activated malate transporter. *Plant Journal*, 37: 645-653.

Sasaki, T.; Ryan, P.R.; Delhaize, E.; Hebb, D.M.; Ogihara, Y.; Kawaura, K.; Noda, K.; Kojima, T.; Toyoda, A.; Matsumoto, H. and Yamamoto, Y. (2006). Sequence Upstream of the Wheat (*Triticum aestivum* L.) ALMT1 Gene and its Relationship to Aluminum Resistance. *Plant Cell Physiol*, 47(10): 1343-1354.

Sawaki, Y.; Iuchi, S.; Kobayashi, Y.; Ikka, T.; Sakurai, N.; Fujita, M.; Shinozaki, K.; Shibata, D.; Kobayashi, M. and Koyama, H. (2009). STOP1 regulates multiple genes

that protect Arabidopsis from proton and aluminum toxicities. *Plant Physiol*, 150: 281-294.

Tang, Y.; Garvin, D.F.; Kochian, L.V.; Sorrells, M.E. and Carver, B.F. (2002). Physiological genetics of aluminum tolerance in the wheat cultivar Atlas 66. *Crop Sci*, 42(5): 1541-1546.

Taylor, G.J. (1991). Current views of the aluminum stress response; the physiological basis of tolerance. *Curr Top Plant Biochem Physiol*, 10: 57-93.

Tolrà, R.; Vogel-Mikuš, K.; Hajiboland, R.; Kump, P.; Pongrac, P.; Kaulich, B.; Gianoncelli, A.; Babin, V.; Barceló, J.; Regvar, M.; Poschenrieder, M. (2011). Localization of aluminium in tea (*Camellia sinensis*) leaves using low energy X-ray fluorescence spectro-microscopy. *J Plant Res,* 124(1): 165-172.

Tsuji, M.; Kuboi, T. and Konishi, S. (1994). Stimulatory effects of aluminum on the growth of cultured roots of tea. *Soil Sci Plant Nutr*., 40:3, 471-476.

Vinh, N.T. and Paterson, A.H. (2005). Genome mapping and its implication for stress resistance in plants. In: Ashraf M, Harris PJC, eds. Abiotic stresses: plant resistance through breeding and molecular approaches.

von Uexkull, HR and Mutert, E. (1995). Global extent, development and economic impact of acid soils. In *Plant-Soil Interactions at Low pH:Principles and Management*, ed. RA Date, NJ Grundon, GE Raymet, ME Probert, pp. 5–19. Dordrecht, The Neth:Kluwer Academic.

Wallace, S.U. and Anderson, I.C. (1984). Aluminum toxicity and DNA synthesis in wheat roots. *Agron J*, 76: 5-8.

Yamaji, N.; Huang, C.F.; Nagao, S.; Yano, M.; Sato, Y.; Nagamura, Y. and Ma, J.F. (2009). A zinc finger transcription factor ART1 regulates multiple genes implicated in aluminum tolerance in rice. *Plant Cell*, 21: 3339-3349.

Yamamoto, Y.; Kobayashi, Y.; Devi, R.S., Rikiishi, S. and Matsumoto, H. (2002). Aluminum toxicity is associated with mitochondrial dysfunction and the production of reactive oxygen species in plant cells. *Plant Physiol*, 128: 63-72.

Zhang, G. and Taylor, G.J. (1989). Kinetics of aluminum uptake by excised roots of aluminum-tolerant and aluminum-sensitive cultivars of *Triticum aestivum* L. *Plant Physiol*, 91(3): 1094- 1099.

Zhang, G. and Taylor, G.J. (1990). Kinetics of aluminum uptake in *Triticum aestivum* L. Identity of the linear phase of aluminum uptake by excised roots of aluminum-tolerant and aluminum-sensitive cultivars. *Plant Physiol*, 94(2): 577- 584.

Zheng, S.J.; Ma, J.F. and Matsumoto, H. (1998). High aluminum resistance in buckwheat: I. Al-induced specific secretion of oxalic acid from root tips. *Plant Physiol*, 117:745-51.

Zhou, L.L.; Bai, G.H.; Ma, H.X. and Carver, B.F. (2007). Quantitative trait loci for aluminum resistance in wheat. *Mol Breed*, **19**(2): 153-161.

Morphological and physiological response to aluminium stress in bread wheat (*Triticum aestivum* L.) roots

ABSTRACT

Background

Al is toxic to many plant species, and globally, it is the major limiting factor for plant productivity in acid soils. Inhibition of root growth is one of the common symptoms of Al stress which is extensively studied in different plant species. In the past, various mechanisms of Al tolerance or resistance have been advocated which could simultaneously operate in plants to cope with Al toxicity. At the physiological level, toxic concentrations of Al seem to induce multiple stresses. Therefore, the aims of the present study were to investigate fundamental characteristics of Al toxicity in the roots of two diverse bread wheat genotypes using different histochemical assays, and also to explore the utility of these techniques in bread wheat screening for Al tolerance.

Results

Portuguese bread wheat genotype Barbela 7/72/92 clearly exhibited a higher root re-growth under Al stress compared to a Polish genotype Anahuac. Hematoxylin and morin assays illustrated that Barbela 7/72/92 is an Al resistant genotype whose Al resistance mechanism seems to rely on the exclusion of Al from its roots. Contrarily, a high concentration of Al was observed in the roots of Anahuac under Al stress.

In addition, Schiff's reagent and Evan blue procedures also indicated that a higher level of Al induced lipid peroxidation and physical injuries in the roots of Anahuac in comparison to Barbela 7/72/92, respectively. It is interesting to note that Al stress seems to a have positive effect on the rate of root hair development in genotype Barbela 7/72/92. These histochemical techniques revealed that genotype Barbela 7/72/92 switched on several mechanisms to cope with Al toxicity.

Conclusions

Various histochemical assays clearly exhibited that bread wheat genotype Barbela 7/72/92 is highly resistant to Al toxicity. The results obtained in the present

investigation also suggested the advantage of histological techniques over biochemical procedures as these assays detect the actual localization of Al in the root and the sites of Al induced stresses. These results clearly indicated that toxic concentration of Al induce multiple effects in wheat roots such as lipid peroxidation, loss of plasma membrane integrity and physical injuries in the form of cell ruptures and death. The callose formation and root hair development also indicated their important role to cope with Al toxicity in Barbela 7/72/92 genotype. Finally, the histochemical techniques results presented on the bread wheat root responses upon exposure to toxic concentration of Al provides an indication that Barbela 7/72/92 is not only a useful experimental material to further elucidate the complex mechanisms of Al toxicity in wheat, but also could be used in breeding programmes to develop improved cultivars for acidic soils.

Key words: Aluminium toxicity, bread wheat, physiological assays.

INTRODUCTION

Plants exposed to stress exhibit broad range of morphological, physiological and biochemical responses. Aluminium (Al) is phytotoxic affecting plant growth and productivity in regions where acid soils are more prevalent. Acid soils are widespread on the earth's surface and have been estimated to occur on about 40% of the world's arable land (Conner and Meredith, 1985). Al toxicity is one of the most important factors in these acid soil areas (Eswaran *et al.*, 1997). The primary site of Al toxicity is the roots where toxic Al concentration led to the disruption of cellular and molecular processes that result in complete or partial inhibition of the main root growth and also restriction of lateral root and poor root hair development (Delhaize and Ryan, 1995; Yamamoto et al., 2003). Subsequently, it affects water and nutrient uptake leading to a short and stunted root system (Barcelo and Poschenrieder, 2002).

In the past decades, extensive research efforts have paved the way for substantial advances in our understanding on several adverse effects of Al phytotoxicity in plants, such as, disruption of signal transduction pathways, inhibition of cell division and ion fluxes (i.e. cytoplasmic Ca^{2+} homeostasis, disruption of cytoskeletal dynamics, induced generation of reactive oxygen species (ROS), induction of lipid peroxidation, disturbance of plasma membrane stability and function, inhibition of DNA synthesis, and induction of callose accumulation in the plasmodesmata of root cell that could block the cell to cell trafficking) (Wallace and Anderson 1984; Jones and Kochian, 1995; Blancaflor *et al.*, 1998; Yamamoto *et al.*, 2001 and 2002; Kochian *et al.*, 2004; Ma, 2007).

Substantial inter and intra-specific variation for Al resistance has been reported in the members of Poaceae family (Magnavaca *et al.*, 1987; Jan and Petterson, 1989; Poschenrieder *et al.*, 2008; Martins-Lopes *et al.*, 2009). Among cereal crops, wheat is one of the major staple food crops and is considered to be sensitive to Al toxicity. However, in some parts of the world, particularly Brazil and Portugal, vast expanses of land have acid soils. In these regions, Al toxicity is one of the major abiotic stress which affects wheat productivity. Barbela is a collective name for a Portuguese bread wheat landrace that has been cultivated for more than one century and several lines exhibiting varying levels of tolerance to Al toxicity were developed from the original Barbela

collection using single seed descent method (Guedes-Pinto *et al.*, 1998; Martins-Lopes *et al.*, 2009).

Germplasm screening is a pre-requisite step in any plant improvement programme. In order to perform an efficient screening, the availability of an easy and reliable phenotyping technique for complex traits of interest, particularly Al tolerance is most critical from the breeder's point of view. Hematoxilin staining is one of the widely used techniques for the Al resistance assessment in plants due to its quick and non-destructive nature, but the reliability of this technique to classify plants according to their Al-sensitivity is also genotype and plant species dependent (Macêdo *et al.*, 2009). Previously, selected Portuguese bread wheat genotypes along with some exotic lines were screened under Al stress and exhibited varying degrees of tolerance to Al toxicity (Guedes-Pinto *et al.*, 1998; Martins-Lopes *et al.*, 2009). These genotypes could provide unique opportunities to better understand the Al tolerance/resistance mechanisms in wheat.

The aim of the present work was to visualize the effect of Al toxicity in two diverse bread wheat genotypes using different physiological markers, and also to explore the utility and assess the impacts of these techniques in the screening of bread wheat genotypes for Al tolerance/resistance which could be applied in the development of elite cultivars for acid-affected soils.

RESULTS

Root morphological response to Al exposure

Two wheat genotypes, Barbela 7/72/92 and Anahuac, were characterized for A1 tolerance using percentage of root re-growth as an index using eriochrome cyanine R staining method. Genotype Barbela 7/72/92 exhibited the highest percentage of root re-growth, 38.7 mm, under 74 µM of Al. Contrarily, no root re-growth was observed under Al stress (74 µM) in genotype Anahuac, after 24 h. In the roots of the Al sensitive genotype, Anahuac, the stain intensity was greater in the apical area (0-3 mm), whereas in Barbela 7/72/92 the stain showed root re-growth basal to this region (Figure 1).

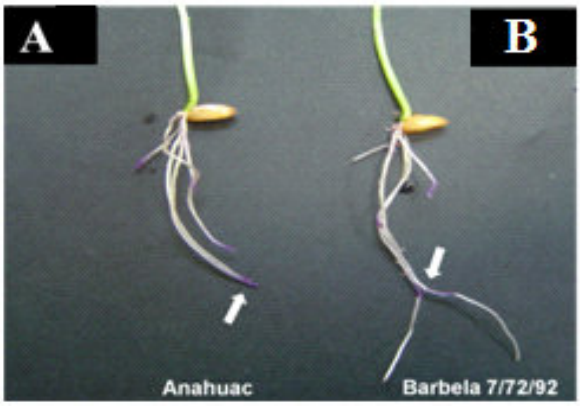


Figure 1: Physiological characterization of root re-growth of two bread wheat genotypes Anahuac (A) and Barbela 7/72/92 (B) using eriochrome cyanine R assay after exposure to Al (74 µM) stress.

Al localization

In the absence of Al, the roots of both genotypes could not be differentiated because untreated root did not stain. Al treated Barbela 7/72/92 roots exhibited a typical staining pattern by minimal staining at the root tip, whereas Anahuac roots could be easily differentiated by more intensely stained of a large root extension (Figure 2).

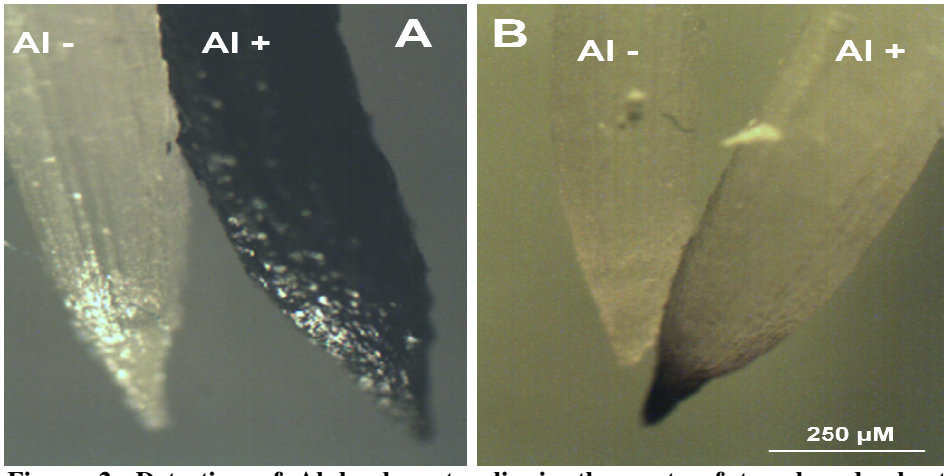


Figure 2: Detection of Al by hematoxylin in the roots of two bread wheat genotypes Anahuac and Barbela 7/72/92. In both pictures the dark colour in the roots represents the accumulation of Al [(A, Anahuac) and (B, Barbela 7/72/92)]. Note: Al with – and + sign indicate the samples grown under control (0 µM) and Al (74 µM) stress, respectively.

The roots of both genotypes (Barbela 7/72/92 and Anahuac) were compared for the strength of their morin signal under control and Al stress. Figure 3 illustrates that Al is distributed uniformly in all the root structure of sensitive genotype Anahuac, whereas Al tolerant genotype, Barbela 7/72/92, presented less Al accumulation and retained only the stain at the root tip. Contrarily, in the absence of Al, no fluorescent dye could be seen on the roots (Figure 3).

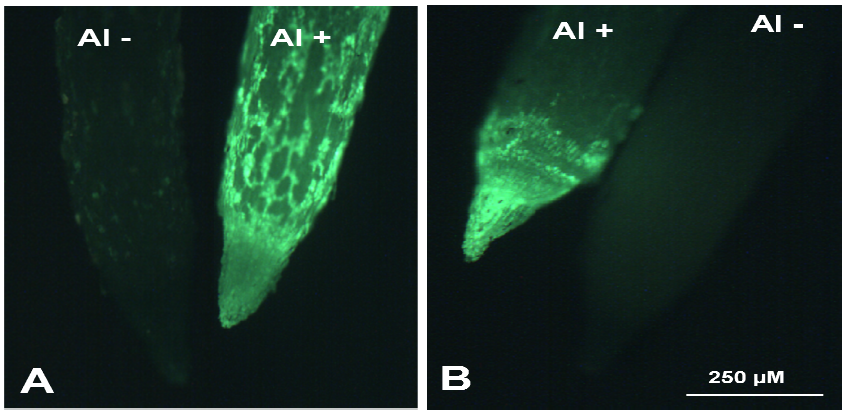


Figure 3: Al accumulation analyzed by fluorescent morin dye in the roots of two bread wheat genotypes Anahuac (A) and Barbela 7/72/92 (B). Control represents without Al (-) stress which do not present the fluorescent morin dye; whereas roots exposed to 74 µM of Al (+) stress present accumulation of the fluorescent dye.

Callose formation under Al accumulation

To understand the effect of Al accumulation on root morphology, the roots of both Al tolerant and sensitive genotypes were visualized with aniline blue staining. In comparison with the control, Al-treated Barbela 7/72/92 roots showed more callose accumulation in the root tip region, whereas, the Anahuac presented a similar fluorescence signal in the control and Al treated samples (Figure 4). Interestingly, Anahuac roots also exhibited Al induced injuries.

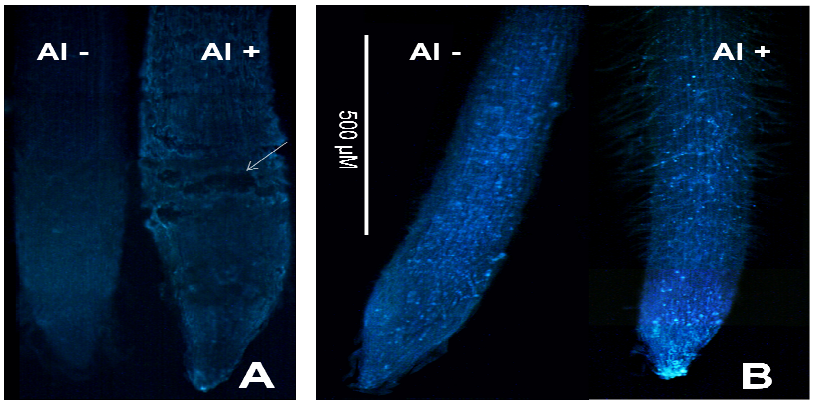


Figure 4: Detection of callose formation and root injury with aniline blue staining in the roots of bread wheat genotypes Anahuac (A) and Barbela 7/72/92 (B). Control represents without Al (-) stress which was compared with roots exposed to 74 µM of Al (+) stress. The arrow indicates injuries caused by Al stress.

Reactive Oxygen Species (ROS) activity

To study the effect of Al treatment on ROS accumulation in wheat root tissue, the untreated and Al treated roots of an Al sensitive genotype Anahuac (A) and Al tolerant genotype Barbela 7/72/92 were visualized by fluorescence microscopy. Figure

5 illustrated that an increase of ROS accumulation in the roots of both genotypes can be clearly visualized by DCF-DA.

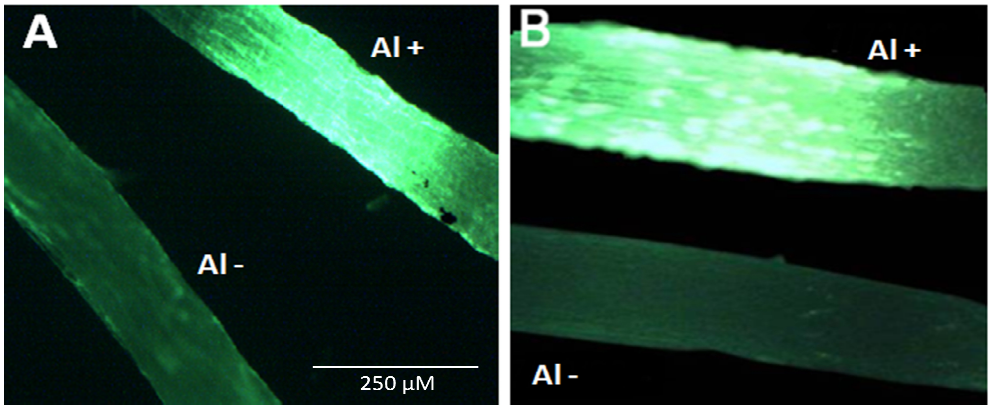


Figure 5: Imaging of peroxide accumulation using DCFDA fluorescence staining in the roots of bread wheat genotypes Anahuac (A) and Barbela 7/72/92 (B). Control root samples (without Al stress; represented by - sign) shows no ROS (DCFDA fluorescent dye) whereas roots exposed to 74 µM of Al stress (represented by + sign) illustrates ROS activity by accumulation fluorescent dye.

Lipid peroxidation

In plants, the Al-induced lipid peroxidation has been widely assessed by Schiff's reagent staining technique which is based on the intensity of the pink colour development (Yamamoto *et al.*, 2001). Under control conditions, no pink colour was developed in the roots of both genotypes. The level of lipidic peroxidation was remarkably higher in the roots of Al sensitive genotype Anahuac which can be clearly indicated due to the development of highly intense pink colour in the roots upon Al exposure. Interestingly, a faint pink ring was visible in the roots of Al tolerant genotype Barbela 7/72/92 which is indicating the lower level of lipidic peroxidation.

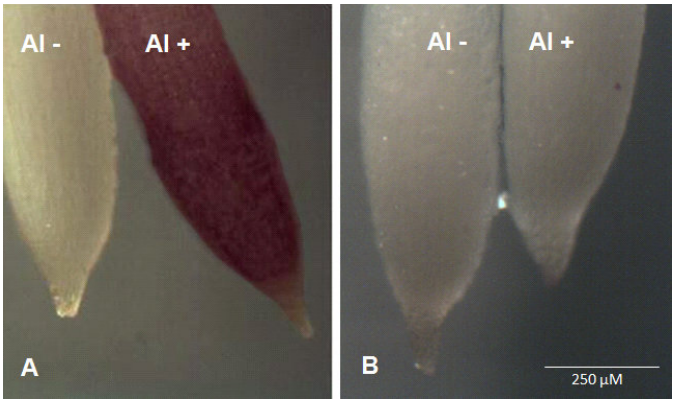


Figure 6: Lipidic peroxidation evaluation in roots of bread wheat genotypes Anahuac (A) and Barbela 7/72/92 (B). Controls without Al (-) shows no staining whereas roots exposed to 74 µM of Al stress (+) present pink coloration on root tips.

Plasma membrane integrity loss

The loss of plasma membrane integrity induced by toxic Al concentration was visualized in the roots of two diverse bread wheat genotypes using Evans blue staining technique (Yamamoto *et al.*, 2001). The control roots (in the absence of Al) exhibited almost no Evans blue staining and the roots periphery was smooth. After Al exposure, the Al sensitive genotype Anahuac stained with Evans blue indicated ruptures formation and cell death in both the cell differentiation and elongation zones of the roots. While the roots of tolerant genotype Barbela 7/72/92 maintained better plasma membrane integrity in comparison to the Anahuac roots (Figure 7).

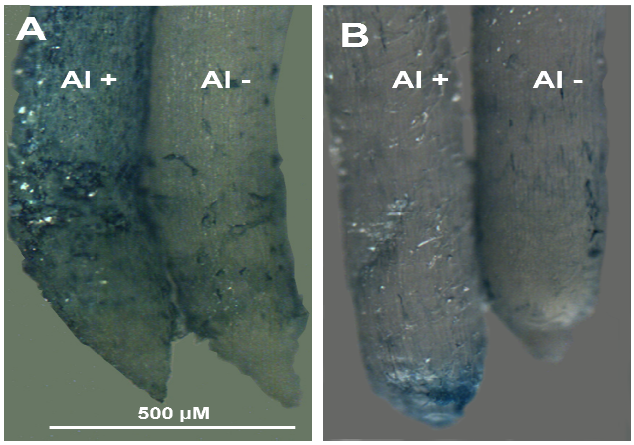


Figure 7: Cell death accessd by Evans blue in the roots of bread wheat genotypes Anahuac (A) and Barbela 7/72/92 (B). Darker regions indicate cell death. Control roots are indicated with Al- whereas Al treated roots (74 µM) are indicated by Al+.

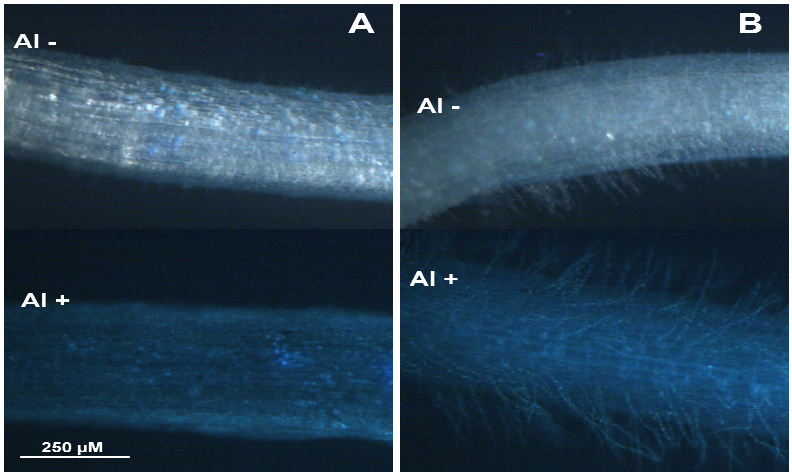


Figure 8: Roots stained with aniline blue dye and observed under fluorescent microscope without (-) and with (+) imposed Al stress (74 µM) on diverse hexaploid wheat cultivars: Anahuac (A) and Barbela 7/72/92 (B).

Root hairs under Al stress

Plant's root hairs play a major role in the nutrient and water uptake and aid to anchor the plant in the soil. Root hairs are the fastest elongating plant cells which are generally very sensitive to Al exposure. Keeping in view, the root hair development was analysed in two bread wheat genotypes under microscope. Remarkable differences were observed in root hair development between the Al sensitive and tolerant genotypes (Figure 8). In the absence of Al, Barbela 7/72/92 showed less root hairs development. However, it is interesting to verify that in the presence of toxic Al concentration, root hairs development and their elongation in tolerant genotype Barbela 7/72/92 were much higher. Contrarily, it seemed that sensitive genotype Anahuac also had minute root hairs under control conditions, but no hairs were observed on Al treated roots.

DISCUSSION

Considerable advances in the understanding of Al tolerance mechanism in different plants have been made in the last two decades. Although, numerous aspects of Al toxicity are still to be clarified, but undoubtedly, previous studies represent that root apex is the first target of Al toxicity. In the present work several parameters associated with Al tolerance mechanisms in bread wheat were assessed using different physiological assays. Eriochrome cyanine R staining was extensively used in cereal species for the evaluation of root re-growth after exposure to Al stress (Gallego and Benito, 1997; Ma *et al.*, 1997, Guedes-Pinto *et al.*, 1998). The assay clearly shows if the root apical meristem has been irreversibly damaged, through an intensively root tip purple color staining; whereas the root section that grows after Al stress remains unstained (white color) (Aniol, 1995). In the present investigation, the roots submitted to eriochrome cyanine R staining clearly demonstrated that Barbela 7/72/92 genotype is highly Al tolerant whereas Anahuac is sensitive to Al stress (Figure 1A). Subsequently, hematoxylin assay which is based on the formation of colored complex between hematoxylin and root-bound Al (Polle *et al.*, 1978) also clearly demonstrated that Barbela 7/72/92 is tolerant to high Al toxicity levels when compared to Anahuac (Figure 2). Hematoxylin assay provides a semi-quantitative estimation of resistance to Al toxicity which allows to classify the investigated genotypes as Al-sensitive, resistant/tolerant and intermediate based on the intensity of staining (stained, unstained and partially stained).

Morin is a fluorochrome which forms a fluorescent complex with Al and is thus used to localize Al in plant tissues (Eticha *et al.*, 2005). Morin staining of the Al-treated roots indicated that Barbela 7/72/92 accumulated less Al than Anahuac in their root tip region, particularly the apical tip which includes the cell division and cell elongation zones (Figure 3). The minor morin staining in the tolerant/resistant genotype may be that the Al uptake into the root tips may occur as the result of Al toxicity (Ezaki *et al.*, 2000). Thus, the higher resistance to Al toxicity observed in Barbela 7/72/92 genotype can be reflected in the reduced Al content present in the root tips. However, the results obtained with morin staining need to be interpreted with cautions as not all Al forms such as Al bound to the cell walls, are accessible to the dye (Eticha *et al.*, 2005).

Al-induced callose (1,3-β-D-glucan) accumulation has also been implied as an Al resistance mechanisms in the roots of several plant species and is considered as an indicator of Al sensibility (Tahara *et al.*, 2005). Similar to pea (Yamamoto *et al.*, 2001), it is also interesting to note that callose production are caused directly by the interaction of Al with the root surface in the Barbela 7/72/92 genotype suggesting that callose production depends on the tissues sensitivity and could be a mechanism to prevent Al from penetrating into the apoplast (Figure 4). Evidence of the protective role of root cap mucilage under Al stress has been frequently reported, this mucilage is exuded from the outer layers of the root cap and it thought to aid in forming a diffusion barrier to Al (Archambault *et al.*, 1996). Both genotypes, Barbela 7/72/92 and Anahuac, did not form mucilage. Although, a previously characterized Al resistant bread wheat genotype Viloso Mole (Martins-Lopes *et al.*, 2009) showed mucilage formation on the root cap under Al stress [Supplementary file 1]. Nevertheless, the mucilage formation was not present in all the roots of a plant. Thus, more studies are required to clarify whether mucilage was a reaction to Al exposure or natural phenomena present in the Viloso Mole.

Schiff's reagent procedure was originally performed to investigate the lipid peroxidation in liver tissue of bromobenzene-intoxicated mice (Pompella *et al.*, 1987) through the detection of aldehyde functions that are originated from the peroxidation of the membrane lipids and bound to the membranes' protein. In agreement with previous studies in maize (Vardar *et al.*, 2011), pea (Yamamoto *et al.*, 2001), soybeans (Cakmak and Horst, 1991), tobacco (Yin *et al.*, 2010) and wheat (Hossain *et al.*, 2005), this histochemical technique also seems to be reliable for the detection of lipid peroxidation

in plant roots, as positive staining with Schiff's reagent was detected only in Al-treated roots (Figure 6). Although, we did not compare the results with other biochemical procedures, the histological technique has an advantage over the biochemical procedure as it reveals the localization of the Al-enhanced peroxidation of lipids *in situ* on the root surface with higher sensitivity. Figure 6 clearly detected a higher level of Al-enhanced peroxidation of lipids in the Al sensitive genotype, Anahuac, when compared to Barbela 7/72/92. Similar to peas (Yamamoto *et al.*, 2001) and soybeans (Cakmak and Horst, 1991), our results also indicate that lipids could be one of the primary cellular targets of oxidative stress under Al toxicity in wheat. In addition, these results also indicate that ROS formations can be facilitated by the loss of membrane integrity in the sensitive wheat genotype when compared to the resistant wheat, as the peroxidase activity is considered as an evidence for enhanced production of free oxygen radicals (H_2O_2/O_2).

Generally, Al-induced lipid peroxidation causes membrane breakdown which may provoke plant cell death. Evan blue dye can be used as an indicator of the loss of plasma membrane's integrity as well as an indication of cell death, which is a non-permeating dye that leaks through ruptured plasma membranes and stains the dead cells' contents. In the present work, the loss of plasma membrane's integrity assessed by Evans blue indicated that membrane damage induced by Al could be due to the mechanical disruption of cells in Al sensitive genotype Anahuac (Figure 7). Whether this is due to necrosis or programmed cell death remains to be determined, but there remains a possibility that Al may cause a minor alteration of the plasma membrane permeability.

It has been speculated that in acidic soil, when the roots reach a region containing highly toxic levels of Al, a mechanism could be activated to destroy the root tip cells, to break the apical dominance and thus induce the formation of secondary roots to explore other soil portions with lower Al levels (Boscolo *et al.*, 2003). Root hairs are lateral extensions of a single cell which are found only in the maturation region of the root. Root hairs are among the fastest elongating plant cells and are considered very sensitive to Al toxicity (Care, 1995). The suppression in root hairs growth under Al toxic concentrations has been reported in *Arabidopsis thaliana*, soybean (*Glycine max*) and *Limnobium stoloniforum* (Brady *et al.*, 1993; Jones *et al.*, 1995; 1998). It is interesting to note that Al resistant genotype Barbela 7/72/92 showed lower number of hairy roots under control condition and that the Al treatment seemed to a have positive

effect on the rate of root hair development (Figure 8). The positive effect of Al on the root hair length can be explained by an alleviation of proton toxicity by Al as occurs for root growth (Kinraide, 1993) and the root hairs might play an important role for the efficient uptake of water and nutrients, particularly under Al stress in acid soils (Delhaize *et al.*, 2012).

In agreement with the previous studies on Al tolerance in different plant species, our results also advocate the usefulness of eriochrome cyanine R staining technique over other histochemical techniques in terms of its inexpensiveness, high accuracy (can assess the quantitative level of Al tolerance in wheat) and easy to be used in the screening of large sample number (germplasms accessions) in a limited time frame. Our results clearly show that exposure to Al toxic concentrations induces different response patterns and mechanisms of Al tolerance/resistance which affects several morphological and physiological traits in wheat, suggesting that this trait could be governed by multiple genes. In addition, Al toxicity seems to induce the correlated response among physiological parameters, such as, lipidic peroxidation and inhibition of root elongation (Camak and Horst, 1991).

Conclusions

Summarizing, histochemical assays clearly exhibited that Barbela 7/72/92 is a highly Al resistant genotype and its resistance mechanism mainly relies on Al exclusion. Furthermore, the results of different physiological parameters presented here on the responses of wheat roots to Al exposure provides a clue that Barbela 7/72/92 genotype is a useful experimental material to understand the mechanisms of Al toxicity in wheat and also contains elite alleles for the candidate genes contributing to Al tolerance/resistance in bread wheat which could be used to develop improved cultivar for acidic regions.

MATERIAL AND METHODS

Plant material and growth conditions

Seeds of bread wheat genotypes screened in relation to Al tolerance were obtained from the CGB/UTAD germplasm bank (Vila Real, Portugal). Seedlings of these genotypes were grown in hydroponic solution (0.4 mM $CaCl_2$, 0.65 mM KNO_3,

0.25 mM $MgCl_2.6H_2O$, 0.1 mM $(NH_4)_2SO_4$ and 0.04 mM NH_4NO_3) at pH 4.0 during four days in controlled growth chamber under 14 h/26ºC day and a 10 h/22ºC night regime, with a light intensity of 150 µmol photons $m^{-2}s^{-1}$ and a relative humidity of 65 *per cent*. For Al stress assay, 74 µM of Al in the form of $AlCl_3 \cdot 6H_2O$ was applied to the nutritive solution. In all the following assays at least six plants (4 to 6 individual roots) were tested.

Determination of Al accumulation

To investigate the Al uptake, hematoxylin staining assay was performed (Polle *et al.*, 1978). Control and Al treated roots from ten seedlings of each genotype were rinsed twice in fresh deionized water for 30 min. Roots were stained with hematoxylin solution [0.2% w/v hematoxylin (Sigma, Germany), 0.01% w/v KIO_3 (Sigma, Germany)] for 30 min. Subsequently, the roots were washed with deionized water for an additional 30 min and hematoxylin observations were made under a light microscope.

Similarly, for morin staining, control and Al treated roots were washed for 10 min in 5 mM solution of ammonium acetate (NH_4OAc, Sigma, Germany) buffer (pH 5.0), at room temperature and stained with 100 µM morin (Sigma, Germany) in the same buffer for 1 h (Tice *et al.*, 1992). Morin fluorescence was visualized on the roots surface of different plants from each genotype using an Olympus inverted fluorescent microscope (Olympus CX31, Japan) connected with camera (U-LH50HG, Japan).

Assessment of plasma membrane integrity loss

Cell viability loss or cell death was evaluated using an impermeant dye (Evans blue, Sigma, Germany) retained by the cells. Freshly harvested roots were stained with 0.25% (v/v) of Evans blue aqueous solution during 10 to 15 min, at room temperature (Yamamoto *et al.*, 2001). The roots were washed with $CaCl_2$ 100 µM (pH 5.6) 3 times and observed under a light microscope.

Determination of cell peroxidation

Structures rich in polysaccharides, mucopolysacharides or glycolipids are the principal tissues stained with Schiff's reagent after acid degradation. These tissues

include mucous glands, basement membranes, reticular tissue fibers, granules in some pituitary cells. In order to perform this assay, ¼ of the root tips were left 20 min in Schiff's reagent (Merck, Germany) at room temperature, after which it was discarded and a solution of 0.5% $K_2S_2O_5$ and 0.05 M HCl was added to clean the roots (Yin *et al.*, 2010).

Callose detection

To detect callose formation, approximately ten seedling roots of each genotype were excised and stained with few drops of 0.1% (w/v) aniline blue [Aniline Blue diammonium salt (Sigma-Aldrich, Germany)] diluted in 1 mL Gly/NaOH buffer (pH 9.5). The slides were observed directly on an inverted fluorescent microscope (Olympus CX31, Japan) with camera U-LH50HG.

Distribution and accumulation of H_2O_2 in the roots – DCF-DA

Distribution of hydrogen peroxide (H_2O_2) in the wheat roots was detected by DCF-DA (Wako Pure Chemical; Jones *et al.*, 2006). Root tips were excised and placed into a solution containing 200 µM $CaCl_2$ (pH 4.4) and 10 µM DCF-DA for 15 min (Yin *et al.*, 2010). The DCF-DA fluorescence was then detected under an inverted fluorescent microscope (Olympus CX31, Japan) with fluorescence camera U-LH50HG.

Image treatment

Images were treated using Adobe photoshop C4.

REFERENCES

Aniol, A.M. (1995). Physiological aspects of aluminium tolerance associated with the long arm of chromosome 2D of the wheat (*Triticum aestivum* L.) genome. *Theor Appl Genet*, 91(3): 510-516.

Archambault, D.J.; Zhang, G.C. and Taylor, G.J. (1996). Accumulation of Al in root mucilage of an Al-resistant and an Al-sensitive cultivar of wheat. *Plant Physiol*, 112: 1471-1478.

Barceló J. and Poschenrieder, C. (2002). Fast root growth responses, root exudates, and internal detoxification as clues to the mechanisms of aluminium toxicity and resistance: a review. *Environ Exp Bot*, 48(1): 75-92.

Blancaflor, E.B.; Jones, D.L. and Gilroy, S. (1998). Alterations in the cytoskeleton accompany aluminium-induced growth inhibition and morphological changes in primary roots of maize. *Plant Physiol*, 118: 159-172.

Boscolo, P.R.S.; Menossi, M. and Jorge, R.A. (2003). Aluminum-induced oxidative stress in maize. *Phytochem*, 62: 181-189.

Brady, D.J.; Edwards, D.G.; Asher, C.J. and Blamey, F.P.C. (1993). Calcium amelioration of aluminium toxicity effects on root hair development in soybean *[Glycine max* (L.) Merr.]. *New Phytol*, 123(3): 531-538.

Cakmak, I. and Horst, W.J. (1991). Effect of Aluminum on lipid peroxidation, superoxide dismutase, catalase, and peroxidise activities in root tips of soybean (*Glycine max*). *Physiol Plant*, 83: 463-468.

Care, D.A. (1995). The effect of aluminium concentration on root hairs in white clover (*Trifolium repens* L.). *Plant Soil*, 171: 159-162.

Conner, A.J. and Meredith, C.P. (1985). Simulating the mineral environment of aluminium toxic soils in plant cell culture. *J Exp Bot*, 36: 870-880.

Delhaize, E and Ryan, P.R. (1995). Aluminum toxicity and tolerance in plants. *Plant Physiol*, 107: 315-321.

Delhaize, E.; James, R.A. and Ryan, P.R. (2012). Aluminium tolerance of root hairs underlies genotypic differences in rhizosheath size of wheat (*Triticum aestivum*) grown on acid soil. *New Phytol*, 195(3): 609-619.

Eswaran, H.; Reich, P. and Beinroth, F. (1997). Global distribution of soils with acidity. In: *Plant-Soil Interactions at Low pH.* Moniz, A.C. et al. (eds.). Brazilian Soil Science Society, pp. 159-164.

Eticha, D.; Stass, A. and Horst, W.J. (2005). Localization of aluminium in the maize root apex: can morin detect cell wall-bound aluminium? *J Exp Bot*, 56(415): 1351-1357.

Ezaki, B.; Gardner, R.C.; Ezaki, Y. and Matsumoto, H. (2000). Expression of aluminum induced genes in transgenic Arabidopsis plants can ameliorate aluminum stress and/or oxidative stress. *Plant Physiol*, 122(3): 657-666.

Gallego, F.J. and Benito, C. (1997). Genetic control of aluminium tolerance in rye (Secale cereale L.). *Theor Appl Genet*, 95: 393-399.

Guedes-Pinto, H.; Ribeiro-Carvalho, C.; Pinto-Carnide, O.; Igrejas, G.; Nascimento, M.; Carnide, V.P.; Harrison, G.; Heslop-Harrison, J.S.; Gateau, I. and Branlard, G. (1998). Studies of Barbela wheat, an old Portuguese landrace with rye introgression. In: Lelley T, ed. *Current topics in plant cytogenetics related to plant improvement*. Vienna: Universitätsverlag, 175-181.

Hossain, M.A.; Hossain, A.K.M.Z.; Kihara, T.; Koyama, H. and Hara, T. (2005). Aluminum-induced lipid peroxidation and lignin deposition are associated with an increase in H_2O_2 generation in wheat seedlings. *Soil Sci Plant Nutr*, 51(2): 223-230.

Jan, F. and Pettersson, S. (1989). Varietal diversity of upland rice in sensitivity to aluminum. *J Plant Nutr*, 12: 973-993.

Jones, D.L. and Kochian, L.V. (1995) Aluminum inhibition of the inositol 1,4,5-triphosphate signal transduction pathway in wheat roots: a role in aluminum toxicity. *Plant Cell*, 7: 1913-1922.

Jones, D.L.; Blancaflor, E.B.; Kochian, L.V. and Gilroy, S. (2006). Spatial coordination of aluminum uptake, production of reactive oxygen spcies, callose production and wall rigidification in maize roots. *Plant Cell Environ,* 29: 1309-1318.

Jones, D.L.; Gilroy, S.; Larsen, P.B.; Howell, S.H. and Kochian, L.V. (1998). Effect of aluminum on cytoplasmic Ca^{2+} homeostasis in root hairs of *Arabidopsis thaliana* (L). *Planta*, 206: 378-387.

Jones, D.L.; Shaff, J.E. and Kochian, L.V. (1995). Role of calcium and other ions in directing root hair tip growth in *Limnobium stoloniferum*. I. Inhibition of tip growth by aluminium. *Planta*, 197(4): 672-680.

Kinraide, T.B. (1993). Aluminum enhancement of plant growth in acid rooting media. A case of reciprocal alleviation of toxicity by two toxic cations. *Physiol Plant*, 88: 619-625.

Kochian, L.V.; Hoekenga, O.A. and Piñeros, M.A. (2004). How do crops plants tolerate acid soils? Mechanisms of aluminum tolerance and phosphorous efficiency. *Ann Rev Plant Biol*, 55: 459-493.

Ma, J.F. (2007). Syndrome of aluminum toxicity and diversity of aluminum resistance in higher plants. *Int Rev Cytol*, 264: 225-252.

Ma, J.F.; Zheng, S.J.; Li, X.F.; Takeda, K. and Matsumoto, H. (1997). A rapid hydroponic screening for aluminium tolerance in barley. *Plant Soil*, 191: 133-137.

Macêdo, C.E.C.; Jan, V.V.S.; Kinet, J.M. and Lutts, S. (2009). Effects of aluminium on root growth and apical root cells in rice (Oryza sativa L.) cultivars. Reliability of screening tests to detect Al resistance at the seedling stage. *Acta Physiol Plant*, 31: 1255-1262.

Magnavaca, R.; Gardner, C.O. and Clark, R.B. (1987). Evaluation of inbred maize lines for aluminum tolerance in nutrient solution. In HW Gabelman, BC Loughman, eds, Genetic Aspects of Plant Mineral Nutrition. Martinus Nijhoff, Dordrecht, The Netherlands, pp 255-265.

Martins-Lopes, P.; Maçãs, B. and Guedes-Pinto, H. (2009). Portuguese bread wheat germplasm evaluation for aluminium tolerance. *Cereal Res Commun*, 37(2): 179-188.

Polle, E.; Konzak, C.F. and Kittrick, J.A. (1978). Visual detection of aluminium tolerance levels in wheat by hematoxylin staining of seedling roots. *Crop Sci*, 18: 823-827.

Pompella, A.; Maellaro, E.; Casini, A.F. and Comporti, M. (1987). Histochemical detection of lipid peroxidation in the liver of bromobenzene–poisoned mice. *Am J Pathol*, 129: 295-301.

Poschenrieder, C.; Gunsé, B.; Corrales, I. and Barceló, J. (2008). A glance into aluminum toxicity and resistance in plants. *Sci Total Environ*, 400(1-3): 356-368.

Tahara, K.; Norisada, M.; Hogetsu, T. and Kojima, K. (2005). Aluminum tolerance and aluminum-induced deposition of callose and lignin in the root tips of Melaleuca and Eucalyptus species. *J Forest Res*, 10(4): 325-333.

Tice, K.R.; Parker, D.R. and DeMason, D.A. (1992). Operationally defined apoplastic and symplastic aluminum fractions in root tips of aluminum-intoxicated wheat. *Plant Physiol*, 100: 309-318.

Vardar, F.; Ismailoğlu, I.; Inan, D. and Ünal, M. (2011). Determination of stress responses induced by aluminum in maize (Zea mays). *Acta Biol Hung*, 62(2): 156-170.

Wallace, S.U. and Anderson, I.C. (1984). Aluminium toxicity and DNA synthesis in wheat roots. *Agron J*, 76: 5-8.

Yamamoto, Y.; Devi, S.R.; Rikiishi, S. and Matsumoto, H. (2003). Oxidative stress triggered by aluminum in plant roots. *Plant Soil*, 255: 239-243.

Yamamoto, Y.; Kobayashi, Y. and Matsumoto, H. (2001). Lipid peroxidation is an early symptom triggered by aluminum, but not the primary cause of elongation inhibition in pea roots. *Plant Physiol*, 125: 199-208.

Yamamoto, Y.; Kobayashi, Y.; Devi, R.S.; Rikiishi, S. and Matsumoto, H. (2002). Aluminum toxicity is associated with mitochondrial dysfunction and the production of reactive oxygen species in plant cells. *Plant Physiol*, 128: 63-72.

Yin, L.; Mano, J.; Wang, S.; Tsuji, W. and Tanaka, K. (2010). The involvement of lipid peroxide-derived aldehydes in aluminum toxicity of tobacco roots. *Plant Physiol*, 152: 1406-1417.

Appendix-I

Mucilage production in Al exposed root tip of bread wheat

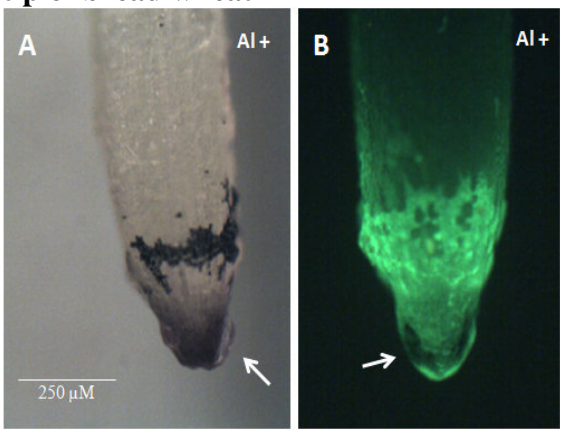


Figure 9: Evidence of mucilage production in Al exposed root tip of bread wheat genotype Viloso Mole accessed by hematoxylin (a) and morin (b). Arrows indicate the presence of mucilage.

Molecular characterisation of *TaMATE1* gene and the analysis of organic acids genes expression under Al stress in bread wheat (*Triticum aestivum* L.)

ABSTRACT

Background:

The efflux of organic anions from roots is considered as one of the major mechanisms for Al detoxification in different plants. The role of ALMT1 which is associated with Al-activated efflux of malate from roots has been extensively studied in bread wheat. In addition, the importance of citrate efflux for Al tolerance has also been reported in cereal species such as maize, barley, sorghum and including some Brazilian wheat cultivars. In the recent past, *HvAACT1* and *AltSB* have shown to be members of MATE family, conferring Al tolerance by Al-activated citrate efflux in barley and sorghum, respectively. In view of the importance of the citrate transporter genes in Al tolerance in different cereals, the objective of the present investigation was to clone and characterise at the molecular level a candidate gene in the MATE1 family for Al tolerance in bread wheat. Furthermore, the upstream variations in citrate and malate transporter genes were also investigated in attempt to gain understanding of Al tolerance mechanism in bread wheat.

Results:

In bread wheat, *TaMATE1* gene consists of 13 exons and 12 introns, located on the long arm of the homoeologous group 4 chromosomes [4AL (*TaMATE1-4A*), 4BL (*TaMATE1-4B*) and 4DL (*TaMATE1-4D*)]. *TaMATE1* homoeologues transcript expression study exhibited the preponderance of homoeologue *TaMATE1-4B* followed by *TaMATE1-4D,* whereas homoeologue *TaMATE1-4A* seemed to be silenced. *TaMATE1* homoeologues transcript levels were observed quite stable under Al stress. Noticeably, higher levels of *TaMATE1,* particularly homoeologue *TaMATE1-4B* and *TaALMT1* gene transcript were detected in the root apices in comparison to the shoots in bread wheat under control and Al stress conditions. Al responsive *TaALMT1* transcript in the root and shoot tissues seemed to depend on the wheat genotype. Barbela7/72/92 presented a very high basal level of both *TaALMT1* and *TaMATE1* transcripts, in both tissues (roots and shoots), contrarily to what was observed in

Anahuac Al sensitive genotype. A Sukkula like transposon was also observed in the promoter of *TaMATE1-4B* homoeologue in different Al tolerant Portuguese bread wheat genotypes including in the Barbela derived lines.

Conclusion:

TaMATE1 belongs to multidrug transporter protein family which is located on the long arm of the homoeologous group 4 chromosomes in bread wheat. A high basal level of *TaMATE1* transcripts in Al tolerant genotype Barbela 7/72/92 and its quite stable expression observed under Al stress condition in both genotypes (Barbela 7/72/92 and Anahuac) suggests that *TaMATE1* is associated with the constitutive efflux of citrate from roots in bread wheat. Remarkably, a higher basal transcript expression of both genes (*TaMATE1* and *TaALMT1*) observed in Al tolerant genotype (Barbela 7/72/92) in comparison with the sensitive genotype, indicating that alleles of these genes play a major role in the Al tolerance of Barbela 7/72/92. The presence of Sukkula like transposon in the upstream of *TaMATE1-4B* homoeologue and Type VI promoter in *TaALMT1* showed a strong positive correlation with Al tolerance in bread wheat genotypes which indicates that selection pressure of Al stress could give rise to tolerant genotypes. Finally, the results clearly show how different mechanisms for Al tolerance can operate simultaneously in a single Portuguese bread wheat genotype Barbela 7/72/92.

Key words: Bread wheat, *Triticum aestivum* L., aluminium tolerance, organic acids, MATE, ALMT, Sukkula like transposon

INTRODUCTION

Soil acidity severely constrains the production of non-adapted crop plants. Besides deficiency of essential nutrients, like P, Ca, and Mg; the high availability of metal ions particularly aluminium (Al) can cause toxicity (Kochian *et al.*, 2004; Poschenrieder *et al.*, 2008). Al is a ubiquitous element and its high levels in acidic soils are extremely toxic for the normal development of plants. Phytotoxic levels of Al primarily block root growth, resulting in small and brittle root systems (Barceló and Poschenrieder, 2002). At the cellular level, Al causes extensive damage not only to cellular processes like cell division and elongation, but also targets various cell components such as cell wall, plasma membrane surface, the cytoskeleton and the nucleus (Panda *et al.*, 2009). Moreover, Al seems to induce callose formation in the plasmodesmata of root cells in bread wheat (*Triticum aestivum* L.) (Schreiner *et al.*, 1994).

The selection pressure of Al toxicity on plants in acid soils is driving the development of new strategies to cope with high levels of Al. This results in considerable genetic variation at inter- and intra-specific levels (Famoso *et al.*, 2012). In plants, two different mechanisms are likely to confer Al tolerance either by external or internal detoxification (Delhaize *et al.*, 2012). So far, organic acid (OAs) exudation from root tips is the best known mechanism for Al exclusion. Moreover, inside cells, OAs regulate the cells' pH and chelate the Al, therefore, protect the roots from Al toxicity (Taylor, 1991). Recent physiological evidence advocates that wheat have either the Al-activated efflux of malate or the constitutive citrate efflux from the roots to cope with Al toxicity (Ryan *et al.*, 1995; Ryan *et al.*, 2009).

At the molecular level, the Al toxicity mechanism is extremely complex and is still not fully understood. In cereals, numerous genetic studies clearly revealed that Al tolerance is governed by the expression of multiple genes that have minor to moderate levels of phenotypic effects, making the identification of these genes difficult. A thorough investigation of both the genetics and physiology of Al resistance played a pivotal role in the identification of major loci associated with Al tolerance in cereals such as *ALMT1* (Al-activated malate transporter) in wheat (Sasaki *et al.*, 2004), and *MATE1* (multidrug and toxic compound exudation) in barley (*Hordeum vulgare* L.) (Furukawa *et al.*, 2007) and sorghum (*Sorghum bicolour* L.) (Magalhães *et al.*, 2007).

In addition, minor loci such as *OsSTAR1*/*OsSTAR2* (sensitive to Al rhizotoxicity) and *ART1* (Al resistance transcription factor1) (Huang *et al.*, 2009; Yamaji et al., 2009) involved with Al tolerance have also been cloned by exploiting the available genomic information coupled with mutational analyses in model plant rice (*Oryza sativa* L.). The high genomic similarity among cereals facilitated further identification of these genes based on their homology in other members of *Poaceae* family such as *ALMT1* in rye (*Secale cereal* L.) (Collins *et al.*, 2008), *MATE1* in maize (*Zea mays* L.) (Maron *et al.*, 2010) and *STOP1* (sensitive to proton rhizotoxicity) in wheat (Garcia-Oliveira *et al.*, 2013). Among these genes, *ALMT1* encodes for a membrane-localized protein and *MATE*/*AACT1* encode transport proteins which control the malate and citrate efflux from roots, respectively (Sasaki *et al.*, 2004; Furukawa *et al.*, 2007; Magalhães *et al.*, 2007).

Among cereal species, wheat is the staple food for 35% of the world's population (Paux *et al.*, 2008). Its high socio-economic relevance as a crop species and the long history of systematic breeding imposed to this crop has resulted in the availability of extensive information on the genetic control of a wide range of traits, including Al tolerance. Apparently, hexaploid wheat exhibits a moderate level of tolerance to Al toxicity among the *Poaceae* family members. Although, some bread wheat genotypes such as Brazilian cultivar BH 1146 and Portuguese landrace Barbela have high degree of tolerance to Al, nevertheless not as high as rye (Pinto-Carnide and Guedes-Pinto, 1999). Recent advances in genomics research have led the way for diploid plant species and the information has also been transferable to other species due to the high level of genome similarity, particularly between wheat, rye and barley. Still, the situation is not as clear in the case of wheat due to its allohexaploid nature as unusual patterns of gene expression can occur in polyploids (Langridge *et al.*, 2006).

In the present study, we report the cloning and molecular characterisation of candidate gene *TaMATE1* homoeologues for Al tolerance which is member of *MATE1* family in bread wheat. Furthermore, we investigated the temporal pattern of *TaMATE1* and *TaALMT1* genes transcripts response under Al stress in the root apices and shoots of two diverse bread wheat genotypes, Barbela 7/72/92 and Anahuac, which exhibit remarkable differences in relation to Al tolerance. Finally, the role of *TaMATE1* promoter was also assessed for Al tolerance in bread wheat. Organic acid exudation by

the roots was analyzed in order to connect the genetic background of both varieties to functional differences in Al tolerance.

RESULTS

TaMATE1 cloning and structure analysis

To obtain the cDNA sequence of *TaMATE1* in bread wheat, two wheat ESTs showing the highest similarities with *ScFRDL1*/*ScMATE1* from rye were aligned and the sequence of *TaMATE1* was obtained from bread wheat genotypes Barbela 7/72/92 and Anahuac using PCR. The full-length coding region was determined using a combination of RT-PCR and 5'- and 3'-RACE. Finally, the structure of *TaMATE1* gene was confirmed by several rounds of amplifications and cloning with specific pairs of primers in both genotypes [Appendix-I]. Multiple alignments of *TaMATE1* clones suggested the mixed amplification from distinct wheat genomes (A, B, and D). The comparison of the *TaMATE1* genomic with cDNA sequences revealed that *TaMATE1* genes consists of 13 exons and 12 introns, with highest intron size of about 1.0 kb (2nd intron) depending upon the wheat genotype and the respective homoeologue (A or B or D) (Figure 1 and Appendix-II). The deduced TaMATE1-4D peptide from Barbela 7/72/92 genotype is a 542 amino acids protein having 58.5 kD molecular weight and 7.90 isoelectric point (pI).

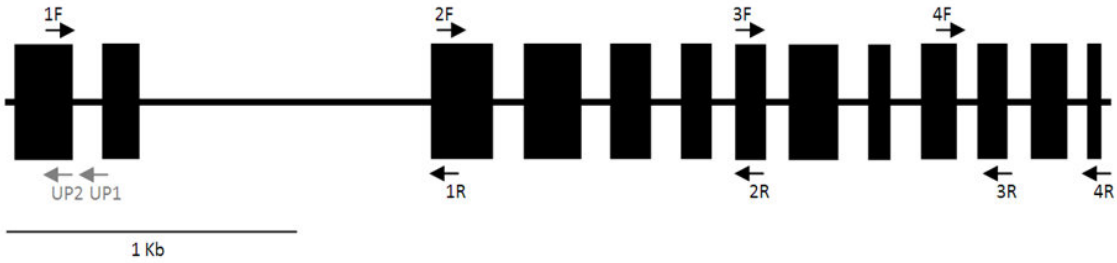


Figure 1. The structure of the genomic sequence of *TaMATE1* homoeologues in bread wheat. Exons shown in boxes are separated by lines representing the introns. Arrows shows the approximate position of the primers used in order to amplify the genomic sequences.

Phylogenetic relations among MATE1 proteins

The phylogenetic tree was constructed by comparison of full-length deduced amino acid sequences of TaMATE1 from bread wheat genotype Barbela 7/72/92 with those of protein sequences of other MATE1-type transporters from different plant

species, including MATE1 gene (ScAACT1/ScFRDL1) in rye (Yokosho *et al.*, 2010; Silva-Navas *et al.*, 2012). The phylogenetic analysis clustered the 29 MATE1 type proteins from 24 plant species into two distinct MATE groups which represent monocots and eudicots groups (Figure 2). The homoeologues TaMATE1 sequences identified in bread wheat form a sub-group together with other members of *Poaceae* family in monocots group. Among the *Poaceae* family, TaMATE1 from bread wheat shared the highest identity with MATE1 from rye followed by barley and *Brachypodium distachyum* with identities of 96, 94 and 89%, respectively, and seems to be more closely related than other members of this family.

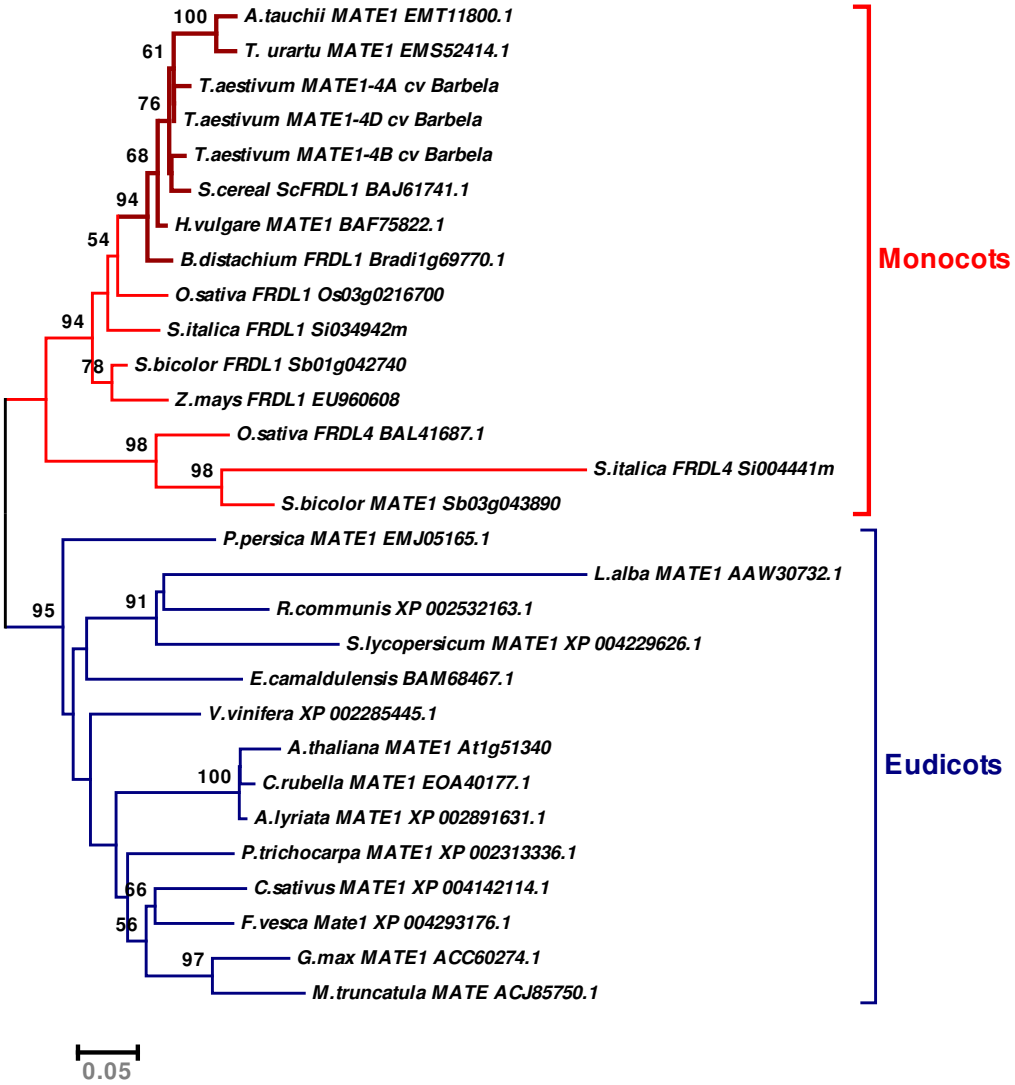


Figure 2. Phylogenetic tree constructed on the basis of amino acid sequences of known and putative MATE1 genes showing the relationship of TaMATE1 with MATE1 type proteins from various plant species. The deduced amino acid sequences were aligned with CLUSTALW program and tree was constructed by the software MEGA version 3.1 with neighbour joining algorithm.

Genetic localisation of *TaMATE1*

To localise the *TaMATE1* on homoeologous chromosomes in wheat, locus specific pair of primers were designed to differentiate *TaMATE1* homoeologues and were further amplified in a series of nullitetrasomic lines of Chinese Spring wheat, whereas Chinese Spring was used as a positive control. On the basis of PCR amplifications, *TaMATE1* was located on homoeologous chromosomes 4A, 4B and 4D in wheat, and named *TaMATE1-4A*, *TaMATE1-4B* and *TaMATE1-4D*, respectively (Figure 3). Furthermore, ditelosomic lines for homoeologous group 4 chromosomes were used to re-assign the *TaMATE1* homoeologues on specific chromosomal arms. PCR based amplification further confirmed that all the three loci are located on the long arms of homoeologous group 4 chromosomes (Figure 3D).

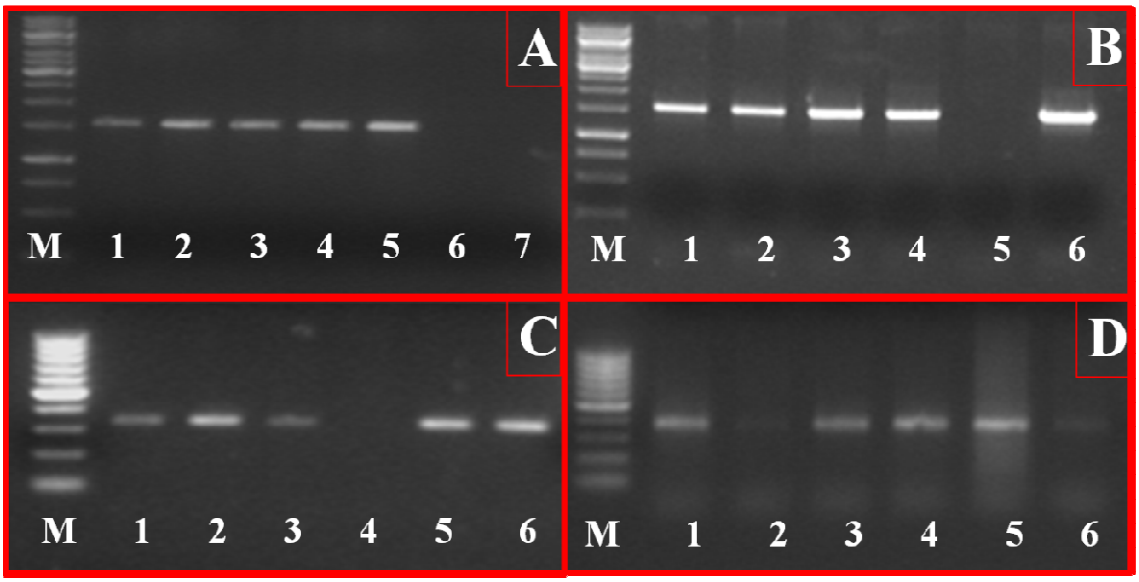


Figure 3. Mapping of *TaMATE1* on homoeologous group 4 chromosomes and their arms in bread wheat. *TaMATE1* mapping on wheat chromosome from genome D using nullitetrasomics lines [A]. Note: M (Marker); 1 ('*Chinese Spring*' control); 2 (N4AT4B); 3 (N4A4D); 4 (N4BT4A); 5 (N4BT4D); 6 (N4DT4A); 7 (N4DT4B). Arms localization of *TaMATE1-4D* [B], *TaMATE1-4B* [C] and *TaMATE1-4A* [D] using ditelosomic lines. NOTE: M (Marker); 1 ('*Chinese Spring*' control); 2 (Dt4AS); 3 (Dt4AL); 4 (Dt4BS); 5 (Dt4DS); 6 (Dt4DL).

TaMATE1 and *TaALMT1* transcript levels under Al stress

Considering the role of *MATE1* and *ALMT1* in Al tolerance in different plant species, relative transcript levels of both genes were evaluated under control and Al stress conditions in the root apices and shoots of two diverse bread wheat genotypes,

Barbela 7/72/92 and Anahuac, which exhibited contrasting response to Al stress [Appendix-III]. The transcript levels of all the three *TaMATE1* homoeologues using homoeologue-specific primers were analysed, whereas the transcript level of *TaALMT1* was quantified as a sum total of the entire set of *TaALMT1* homoeologues using the primers described by Sasaki *et al.* (2004). All the data were normalized with the lowest gene transcript level under control conditions using endogenous *18S RNA* expression as an internal control.

The transcript levels of these genes, associated with transporter of organic anions, were much lower in shoots than in root apices of both bread wheat genotypes (Figure 4). The highest transcript level was observed in *TaMATE1-4B* homoeologue followed by *TaMATE1-4D*, whereas homoeologue *TaMATE1-4A* seemed to be silenced, as we were not able to amplify and quantify the *TaMATE1-4A* homoeologue in these genotypes. The levels of both homoeologues *TaMATE1-4B* and *TaMATE1-4D* in Al tolerant genotype Barbela 7/72/92 was about six folds higher in the root apices than in the shoots, whereas relatively similar transcript levels of these homoeologues were noticed in both tissues of the Al sensitive genotype Anahuac, under both control and Al stress conditions (Figure 4A).

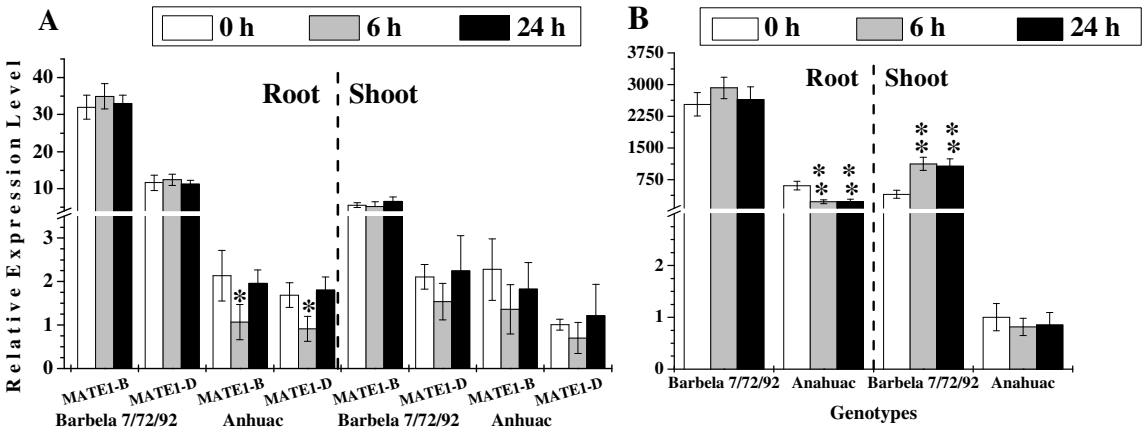


Figure 4. Relative transcript level of *TaMATE1* homoeologues (*TaMATE1-4B* and *TaMATE1-4D*) (A) and *TaALMT1* (B) in root and shoot tissues of two diverse bread wheat genotype Barbela 7/72/92 (Al tolerant) and Anahuac (Al sensitive) under Al stress. Data were normalized using the *18SRNA* gene transcript levels. The lower Δ Ct value obtained from *TaMATE1-4D* homoeologue and *TaALMT1* in shoot of Anahuac under control condition (0 h) was used as calibrator for *TaMATE1* and *TaALMT1* genes, respectively. Error bars represent SD of results from three replicates. Different numbers of asterisks (0 to 2) vertically placed above the columns in the graphs indicate the data sets that are statistically different (P value < 0.05) from 0 h time point (control).

Noticeably, the transcript levels of both *TaMATE1-4B* and *TaMATE1-4D* homoeologues were higher in the Barbela 7/72/92 root apices and shoots in comparison to Anahuac under both control and Al stress conditions (Figure 4A). In the root apices of Al tolerant genotype Barbela 7/72/92, the transcript levels of both *TaMATE1-4B* and *TaMATE1-4D* homoeologues were relatively stable under control and Al stress conditions. Contrarily, in Anahuac transcript levels of these homoeologues appeared to be slightly reduced within 6 h under Al stress, but within 24 h the expression reached their basal level in the root apices of Anahuac. However, no significant differences were noticed in the transcript expression of *TaMATE1* homoeologues (*TaMATE1-4B* and *TaMATE1-4D*) in the shoots of both wheat genotypes under control and Al stress conditions.

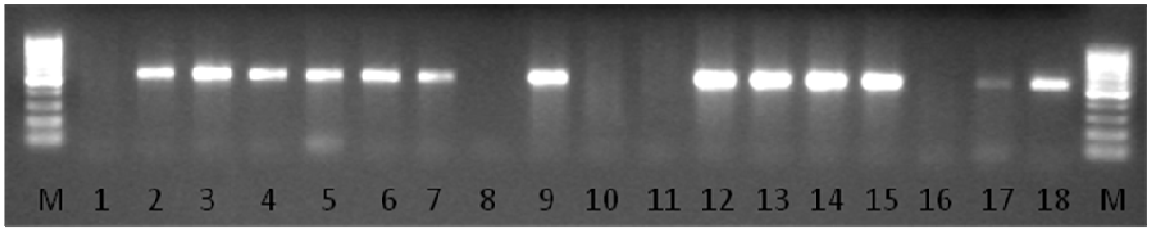
Similar to *TaMATE1* homoeologues, the transcript of *TaALMT1* gene was also highly expressed under control and Al stress conditions in both tissues of Al tolerant genotype Barbela 7/72/92 (Figure 4B). In response to Al stress, a quite stable transcript expression of *TaALMT1* was observed in the root apices of Al tolerant genotype (Barbela 7/72/92); whereas in Al sensitive genotype (Anahuac) the expression of *TaALMT1* in the root apices seemed to be repressed under Al stress at 6 hours, maintaining significantly lower expression compared to the control conditions. Contrarily, a significant induction was noticed in the transcript level of *TaALMT1* in the shoots of Barbela 7/72/92 under Al stress at 6 hours and remained at the same level for 24 hours time point (Figure 4B). Interestingly, a relatively higher *TaALMT1* transcript level was noticed in the shoots of Barbela 7/72/92 than in the root apices of Anahuac.

TaMATE1-4B promoter analysis in diverse bread wheat genotypes

The sequencing of 5'UTR of *TaMATE1* homoeologues in Barbela 7/72/92 and Anahuac revealed that one of the *TaMATE1* homoeologue has a completely different sequence in the upstream region of *TaMATE1* between these genotypes. In addition, nullitetrasomic and ditelosomic lines analysis confirmed the localization of the concerned *TaMATE1* homoeologue in the B genome. Subsequently, sequence analysis suggested that 5'UTR sequence of *TaMATE1-4B* in Barbela 7/72/92 differed from Anahuac by having a transposable element insert at the position of 25 bp upstream of the start codon. The analysis of 5'UTR sequence of *TaMATE1-4B* from Barbela 7/72/92 and its alignment with 5'UTR sequence from Carazinho, a Brazilian Al tolerant bread

wheat genotype, also further confirmed the presence of similar Sukkula like- transposon element in the Al tolerant genotype Barbela 7/72/92 [Appendix-IV] (Tovkach *et al.*, 2012).

In order to see whether this transposon is also present in other Al tolerant bread wheat genotypes from Portugal, additional diverse bread wheat genotypes, including some Barbela derived lines, characterised for Al tolerance were selected (Martins-Lopes et al., 2009). The presence or absence of the Sukkula like- transposon in upstream of *TaMATE1-4B* homoeologue was determined using specific pair of primers which harbors part of transposon and partially the second exon of *TaMATE1-4B* homoeologue (Figure 5). Interestingly, we observed the presence of this transposon in the promoter of *TaMATE1-4B* homoeologue from all the Barbela derived lines which exhibited high tolerance to Al toxicity. In addition, most of the Al tolerant genotypes from Portugal, excepting Sacho and Nabão also exhibited the presence of this transposon in their genome. Noticeably, none of the bread wheat genotypes from outside Portugal, including two well-known Al tolerant genotypes Atlas66 (USA) and BH1146 (Brazil), investigated in the present study, showed the transposon in their *TaMATE1-4B* homoeologue promoters [Appendix-IV].



Figures 5. Detection of Sukkula-like transposon in the promoter of TaMATE1-4B homoeologue from diverse Bread wheat genotypes. Note:- M (Marker: 100 bp); 1 (Anahuac); 2 (Barbela 7/72/92); 3 (Ruivo); 4 (Viloso Mole); 5 (Barbela 55/88/92); 6 (Barbela 6/93/92); 7 (Ardito); 8 (BH1146); 9 (Mocho Rapado); 10 (Chinese Spring); 11 (Fronteiriço); 12 (Barbela 61/94/92); 13 (Barbela 58/70/92); 14 (Barbela 16/95/92); 15 (Barbela 15/92/92); 16 (Sacho); 17 (Mocho de Espiga Branca); 18 (Barbela 35/94/92).

Root exudation of organic acids in diverse bread wheat genotypes

Different patterns of organic acid exudation were observed in the different wheat genotypes. In the Al sensitive genotype Anahuac organic acid exudation was too small to be detectable, excepting after 2 h of Al exposure where 0.6 nmol root^{-1} of citric acid was found in the reaction medium. Contrastingly, control root tips from the Al tolerant Barbela exhibited a constitutively high exudation rate for citrate between 1.8 to 3.8

nmol root^{-1} over the assay period (2-24 h). Detectable concentrations of citrate exudated from Al treated Barbela root tips were found after 6 h exposure; the citrate concentrations increased to 3.5 nmol root^{-1} after 24 h exposure [Appendix-V]. Moreover, small concentrations of oxalate and tartaric acid, but no malate, were found in the exudates. Aluminium-induced malate and oxalate exudation of around 1.8 and 0.7 nmol root^{-1} were observed in the genotypes Ruivo and Viloso Mole. Viloso Mole, but not Ruivo produced Al-induced citrate exudation in the root tips at a rate of 0.2 to 0.6 nmol root^{-1} over the experimental period (2-24 h). In Ruivo, succinic acid was the most abundant organic acid in root tip exudates of Al treated plants reaching close to 23 nmol root^{-1} after 24 h.

DISCUSSION

The important role of organic acid anions such as citrate, malate and oxalate has been demonstrated for the detoxification of Al in different plant species (Yang *et al.*, 2013). The knock-out mutants of *ALMT1* and *MATE1* genes clearly demonstrated that these genes play an important role in Al detoxification by releasing malate and citrate efflux in *Arabidopsis* roots, respectively (Hoekenga *et al.*, 2006; Fujii *et al.*, 2012). The type of organic acids secreted from root tips under Al stress differs among cereal species. It is well established that malate is secreted in Al tolerant bread wheat and rye genotypes, whereas citrate is released in sorghum and barley. The role of *ALMT1* in wheat and *MATE1* in sorghum and barley has been demonstrated for malate and citrate efflux, respectively (Ryan *et al.*, 2011). Recently, however, *MATE1* was reported to be involved in the Al tolerance mechanism in a Brazilian bread wheat genotype Carazinho (Tovkach *et al.*, 2013). Here we report the cloning and molecular characterisation of *MATE1* gene in bread wheat and transcript response of *MATE1* and *ALMT1* genes under Al stress in two contrasting bread wheat genotypes exhibiting differences for Al tolerance.

The MATE proteins comprise large family proteins which were first characterised for their capacity to confer drug resistance to microbes by an efflux mechanism (Hvorup *et al.*, 2003). Some members of this family have been shown to function as drug/cation antiporters that eliminate toxic compounds and secondary metabolites from the cytosol by exporting them out of the cell or sequestering them to the vacuole (Omote *et al.*, 2006). The characterisation of two *MATE* genes, *Frd3* and

AltSB, has established that citrate is also a substrate of some members of this family which is associated with Al tolerance in *Arabidopsis* and sorghum (Magalhães, 2010).

In bread wheat genotypes Barbela 7/72/72 and Anahuac, *TaMATE1* gene consists of 13 exons and 12 introns, which encode a protein size of 543 (TaMaTE1-4A) and 542 (TaMaTE1-4B or TaMaTE1-4D) amino acids. Multiple alignments of MATE1 like proteins identified weak conserved sequence among monocots and eudicots plant species; but members from *Poaceae* family clearly showed consensus regions. The predicted secondary structure suggested the presence of 10 to 12 transmembrane helices in TaMATE1 homoeologues protein in bread wheat [Appendix-VI]. Most of MATE family members' protein length in prokaryotes and eukaryotes ranges from about 400 to 700 amino acids with 12 transmembrane helices which share about 40% sequence similarity with poorly conserved region (Omote *et al.*, 2006).

Phylogenetic analysis suggested that TaMATE1 was a member of the large and complex family of transporters which can clearly differentiated MATE1 like proteins in monocots and eudicots (Figure 2). Recently, 58 and 40 putative MATE transporter proteins have been reported in the genomes of *Arabidopsis* and *Medicago truncatula*, respectively (Magalhães, 2010) which can also indicate the presence of more members of this family in bread wheat. Among the members of *Poaceae* family, TaMATE1 homoeologues from bread wheat appear to be closer to rye, barley and *Brachypodium* MATE1 transporters (Figure 2).

In the past, the ability of allopolyploids to tolerate structural and numerical changes of chromosomes facilitated the identification of chromosomes and their arms associated with Al tolerance, such as chromosome arms 2DL, 3DL, 4BL, 4DL, 6AL, 7AS and chromosome 7D in bread wheat (Aniol and Gustafson, 1984). In the present investigation, chromosome localisation by analyzing the amplification products of *TaMATE1* genes from genomic DNA of nullitetrasomic and ditelosomic lines of Chinese Spring unraveled that three copies of *TaMATE1* were located on the long arms of homoeologoues group 4 chromosomes. In the post genomic era, nullitetrasomic (missing a pair of chromosomes that is replaced by an extra pair of homoeologous chromosomes) and ditelosomics lines (one arm in a given chromosome is missing) have been used to assign individual homoeologue of a gene to specific chromosomes in hexaploid wheat (Garcia-Oliveira *et al.*, 2013).

Recent molecular findings suggest an asymmetric genomic expression pattern in bread wheat (Feldman *et al.*, 2012). However, the predominant transcript expression of one genome over the other genome(s) varies from gene to gene (Garcia-Oliveira *et al.*, 2013). Keeping this in mind, the relative contribution of the three homoeologues of *TaMATE1* gene at the transcript expression level was determined under Al stress in two bread wheat genotypes showing contrasting Al tolerance. Interestingly, *TaMATE1* homoeologues showed a biased transcript expression under control and Al stress with the highest transcript expression of homoeologue *TaMATE1-4B* followed by *TaMATE1-4D*, whereas homoeologue *TaMATE1-4A* seemed to be silenced (Figure 4A). Among the three homoeologues, silencing of at least one homoeologue has also been previously reported for approximately 20-29% of the genes in bread wheat (Mochida *et al.*, 2004; Bottley *et al.*, 2006); while expression of about 19% of the studied genes showed homoeologue-specific non-additive up- or down-regulation in synthetic allohexaploid wheat (Akhunova *et al.*, 2010).

Noticeably, *TaMATE1* homoeologues transcript levels appear to be quite stable under Al stress. In contrast to the Al sensitive genotype Anahuac, the Al tolerant genotype Barbela 7/72/92 had a very high basal level of *TaMATE1* homoeologues, *TaMATE1-4B* and *TaMATE1-4D*, transcript expression in the root tips and of *TaMATE1-4B* in the shoots. This suggests a role of *TaMATE1-4B* and *TaMATE1-4D* in Al tolerance in bread wheat genotype Barbela 7/72/92 (Figure 4A). Furthermore, these differences in expression levels between Barbela 7/72/92 and Anahuac are reflected in the clearly distinct patterns of root exudation of organic acids between both genotypes. Similarly, an aluminium-activated citrate transporter 1 (*HvAACT1*) belonging to the MATE family is also constitutively expressed in barley (Furukawa *et al.*, 2007). Contrarily, Al responsive expression of *MATE1* gene has been observed in maize (Maron *et al.*, 2010), rye (Yokosho *et al.*, 2010), Arabidopsis (Liu *et al.*, 2009), rice (Yokosho *et al.*, 2011), rice bean (*Vigna umbellate* L.) (Yang *et al.*, 2011) and common bean (*Phaseolus vulgaris* L.) (Eticha *et al.*, 2011). Furthermore, the significantly higher level of *TaMATE1* genes transcript in the root apices than in the shoots of both bread wheat genotypes (Figure 4A) is consistent with previous reports on *MATE1* in *Arabidopsis* (Liu *et al.*, 2009), barley (Furukawa *et al.*, 2007) and river red gum (*Eucalyptus camaldulensis*) (Sawaki *et al.*, 2013).

Similar to *TaMATE1*, the transcript level of *TaALMT1* gene was also higher in the root as well as shoots of Al tolerant bread wheat genotype Barbela 7/72/92, when compared with the Anahuac (Al sensitive) under control and Al stress conditions. Noticeably, both genotypes showed a higher level of *TaALMT1* transcript in the root than in the shoot tissues (Figure 4B). Surprisingly, Al seems to modulate the *TaALMT1* transcript expression in the root and in the shoot tissues of Al sensitive (Anahuac) and tolerant (Barbela 7/72/92) genotypes, respectively. Al responsive transcript expression of *AtALMT1* has also been reported in *Arabidopsis* (Kobayashi *et al.*, 2007) whereas a constitutive expression was described in wheat (Sasaki *et al.*, 2004). In the past, specific pattern of variation (Types I–VI) in the upstream region of the *ALMT1* gene was described and correlated with transcript expression of *TaALMT1* in wheat (Sasaki *et al.*, 2006). Type I pattern had shown the simplest structure, while the others had duplicated and triplicated blocks of sequence in different arrangements that generated significantly greater expression of a reporter gene in the callus and regenerated roots (Ryan *et al.*, 2010). The *TaALMT1* promoter analysis from both genotypes revealed the higher basal expression of *TaALMT1* in Al tolerant genotype Barbela 7/72/92 which contain a Type VI promoter having three tandem repeats of a 205 bp region, whereas Al sensitive genotype, Anahuac, contains only a single copy of these regions [Appendix-VII].

Numerous Barbela lines, including genotype Barbela 7/72/92, studied in the present investigation were derived from Barbela landrace through single seed descent method which had shown tolerance to Al stress (Martins-Lopes *et al.*, 2009). Barbela is a collective name for a Portuguese bread wheat landrace that has been cultivated for more than one century. Large germplasm collection of the 'Barbela' wheat landrace were derived on the bases of individual spikes collected from farmers' fields (Guedes-Pinto *et al.*, 1998). The presence of Sukkula like transposon in the promoter of *TaMATE1-4B* homoeologue of most of bread wheat genotypes from Portugal showed a strong positive correlation with tolerance. This suggests that the selection pressure of Al stress in soils in North Portugal with pH as low as 4.5, could give rise to tolerant genotypes (Coutinho, 1989). These results are in good agreement with what was also observed from Brazilian bread wheat genotype Carazinho (Ryan *et al.*, 2009). Interestingly, we did not detect the transposon in some Al tolerant genotypes from Portugal (Nabão and Sacho) which could be explained by the fact that *ALMT1* might be playing a major role in these Al tolerant Portuguese bread wheat genotypes (Riede and

Anderson, 1996; Zhou *et al.*, 2006) Noticeably, the higher level of Al tolerance in Barbela 7/72/92 seems to be the result of favourable alleles at both loci which is illustrated by remarkably high basal expression of both *TaMATE1* and *TaALMT1* genes.

Conclusions

TaMATE1, a malate transporter gene localised on the long arm of homoeologous group 4 chromosomes exhibited tissues and genotype specific transcript levels in bread wheat. In contrast to Al sensitive genotype Anahuac, Al tolerant Barbela 7/72/92 exhibited a very high basal transcript level of both citrate and malate transporter genes *TaMATE1* homoeologues *TaMATE1-4B, TaMATE1-4D,* and *TaALMT1,* respectively. Recently, the presence of Sukkula like-transposon in the *TaMATE1-B* promoter was somehow correlated with Al tolerance in Brazilian bread wheat genotype Carazinho (Tovkach *et al.*, 2013). The existence of an identical transposon element in Barbela 7/72/92, derived from a Portuguese landrace adapted to acid soils, supports the view that it might be responsible for the higher transcript expression of *TaMATE1-4B* in the root tips of wheat where it confers citrate efflux providing enhanced Al tolerance. Furthermore, the detection of this transposon in other Al tolerant bread wheat genotypes from Portugal also strongly supports this role.

Additionally, the higher basal expression of *TaALMT1* gene in Barbela 7/72/92 was also revealed by the presence of duplicated tandem repeats in its upstream region (Type VI promoter in *TaALMT1*) which was previously correlated with higher level of Al tolerance in bread wheat (Sasaki et al., 2006). Moreover, the higher transcript of these genes in shoots suggests that both genes could be involved not only in external but also in the internal detoxification of Al in Barbela 7/72/92. The transcript expression and sequence analysis of both *TaMATE1* and *TaALMT1* genes in Al tolerant and sensitive genotypes clearly illustrates that Barbela 7/72/92 possesses novel alleles for these transporters which could be used for developing elite cultivars for acid-affected soils through genomic assisted breeding. Finally, it would be very interesting in the future to assess the relative contribution of these novel alleles for the transporter genes in Al tolerant genotype Barbela 7/72/92 either through functional or segregation analysis.

MATERIAL AND METHODS

Plant material and growth conditions

In the present investigation, a set of 30 diverse bread wheat genotypes [(different provinces of Portugal (28), Brazil (1), Poland (1), China (1) and USA (1)] including some Barbela derived lines were selected from the previously characterised bread wheat genotypes for Al tolerance at the Centre of Genomics and Biotechnology (CGB), University of Trás-os-Montes and Alto Douro (UTAD), Vila Real, Portugal (Guedes-Pinto *et al.*, 1998, Martins-Lopes *et al.,* 2009) [Appendix-IV]. For gene cloning and expression studies, bread wheat seedlings were raised in hydroponic culture [0.4 mM $CaCl_2$, 0.65 mM KNO_3, 0.25 mM $MgCl_2.6H_2O$, 0.1 mM $(NH_4)_2SO_4$ and 0.04 mM NH_4NO_3] under controlled growth chamber conditions as previously described (Garcia-Oliveira *et al.*, 2013).

Al stress assay

Among the 30 diverse genotypes of bread wheat, two genotypes namely Barbela 7/72/92 (Al resistant) and Anahuac (Al sensitive) were selected for further physiological characterisation under Al stress [Appendix-III]. Approximately, 20 seedlings from each genotype were raised in a continuously aerated nutritive solution for four days. For Al accumulation assessment, the seedlings were transferred into a continuously aerated nutritive solution with (74 µM Al as $AlCl_3.6H_2O$) or without Al, at pH 4.0 for 24 h. The roots were carefully separated from shoots and washed with 5 mM of ammonium acetate buffer (pH 5.0) for 10 min, followed by staining with a 100 µM aqueous solution of morin ($C_{15}H_{10}O_7$) for 60 min as described by Tice *et al.* (1992). The dye solution was prepared just prior to the beginning of the assay. Root samples were transferred to a glass slide for observation using a fluorescence microscope (Olympus CX31, Japan) equipped with camera (U-LH50HG, Japan). Furthermore, hematoxylin (Polle *et al.*, 1978) and eriochrome assay for detection of Al accumulation and re-growth in roots of both genotypes under Al stress were performed [Appendix-III].

DNA and RNA extraction

Total genomic DNA was isolated from young seedlings using DNeasy Plant Mini Kit (Qiagen, Germany) and RNA was eliminated by the addition of ribonuclease

A during extraction. Total RNA was extracted using Trizol reagent (Invitrogen; http://www.invitrogen.com) followed by purification using PureLinkTM RNA Mini Kit (Ambion, Invitrogen, USA). The possible contamination of genomic DNA in each sample of RNA was removed by DNAseI digestion. The quality of DNA and RNA samples was assessed by electrophoresis on agarose gel and further quantification of each sample was computed by spectrophotometry. The first-strand cDNA was synthesized in a final volume of 20 µl reaction containing: 1 µg RNA, 2 µl 10× RT buffer, 0.8 µl of 25× dNTP mix (100 mM), 2 µl of 10× RT random primers and 1 µl of Multiscribe reverse transcriptase (Applied Biosystems, USA), accordingly manufacturer protocol.

Cloning of full-length *TaMATE1* cDNA

Previously published *ScFRDL1*/*ScMATE1* (Yokosho *et al.*, 2010) nucleotide sequence (NCBI: AB571881.1) from rye was selected as query sequence and searched with this in wheat expressed sequence tag (EST) database from NCBI (http://www.ncbi.nlm.nih.gov/Blastn). Two homologous wheat ESTs with a nucleotide sequence of 592 (NCBI: BE498331.1) and 416 (NCBI: BE605049.1) bases were retrieved. Specific primers for both wheat ESTs showing the highest similarities with *MATE1*, particularly *ScMATE1,* were designed to clone the *TaMATE1* cDNA from Al treated root tissues of two bread wheat genotypes Barbela 7/72/92 and Anahuac. Full-length *TaMATE1* cDNA including 5'UTR and 3'UTR were obtained using 5'RACE (Rapid Amplification of cDNA Ends) and 3'RACE, respectively (SMARTer™ RACE cDNA Amplification Kit, clontech, USA) [Appendix-I]. Further, full length genomic sequence of *TaMATE1* was acquired from genomic DNA of the both bread wheat genotypes [Appendix-II].

MATE1 sequence analysis, phylogenetic tree construction and secondary structure prediction

Sequences of nucleic acid and amino acids were analyzed using Chromas Lite 1.0 (Technelysium Pty Ltd). Twenty-nine MATE1 like protein sequences from 24 different plant species including homoeologues from bread wheat were assessed from the public database using the Basic Local Alignment Search Tool (BLAST) programme from NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) [Appendix-VIII]. Multiple

alignment among MATE1 like protein sequences from different plant species were performed using the CLUSTALW algorithm (http://www.ebi.ac.uk/Tools/clustalw2/index.html) and phylogenetic tree was constructed by MEGA4 software using neighbor-joining method (Tamura *et al.*, 2007). Bootstrap analysis was computed using 1000 replicates and excluding positions with gaps.

The predicted translated amino-acids sequences of TaMATE1 homoeologues from bread wheat genotypes Barbela 7/72/92 and Anahuac were used to identify the protein transmembrane helices using RHYTHM-web server programme (http://proteinformatics.charite.de/rhythm) (Rose *et al.*, 2009). Further, these data were used to create the two-dimensional visual representation of the transmembrane structure of protein by TMRPres2D tool (http://biophysics.biol.uoa.gr/TMRPres2D) (Spyropoulos *et al.*, 2004).

Chromosomal mapping of *TaMATE1* homoeologues in bread wheat

A nullitetrasomic and ditelosomic series of Chinese Spring was used to locate the *TaMATE1* locus on a particular chromosome and its arm in bread wheat, respectively. Each line of the nullitetrasomic series lacks a given pair of homoeologous A, B or D genome chromosomes (nullisomic condition) that are replaced by the corresponding homoeologous chromosome pair (tetrasomic condition) whereas a *ditelosomic* stock lacks both chromatids of one chromosome arm. For this purpose, different pairs of genome specific primers were designed and *TaMATE1* locus was localised in nullitetrasomic/ditelosomic line based on the presence/absence of the expected PCR amplification product.

Extraction and measurement of root organic acid exudation

Organic acid exudates from intact root tips were assayed in control (no Al) and 50 µM of Al stress solution according to Delhaize *et al.* (2004). Root tips were continuously and gently agitated in flasks during the experiment. Exudates were collected at 0, 6, 12, and 24 h time point and lyophilized prior to analysis. Organic acids were quantified by high performance liquid chromatography (Shimadzu Corporation, Kyoto, Japan) attached with a SPD-10AV vp UV-vis detector, auto sampler SIL-10AF, system control SCL-10A VP, and software LC solution. Organic acids were separated

on a YHC- Pack ODS-A column (Ultrapure silica C18, 250 x 4.6 mm D. S-5µm.12 nm) (Waters, USA) kept at 25-27 ºC temperature which was attached with a guard-column and detected at 210 nm wavelength. The mobile phase consisting of 0.12 % phosphoric acid and acetonitrile (H_3PO_4 + CH_3CN) at pH 2.0-2.5 with a flow rate of 0.8 ml/min was used. Identification and quantification of organic acids in the samples were achieved with comparison of retention times and by use of external standards using six-point standard curves, respectively. Values are means of 2 replicates per time sample and treatment.

TaMATE1 and *TaALMT1* transcript expression analysis

Seedlings of both the Al tolerant (Barbela 7/72/92) and sensitive (Anahuac) genotypes of bread wheat grown for four days after germination in nutrient solution were shifted to fresh solution with (74 µM Al in the form of $AlCl_3 \cdot 6H_2O$; stress treatment) or without Al (control treatment). For expression assay, both root apices and shoot tissues were harvested in liquid nitrogen separately after treatment at specific time points (0, 6 and 24 hours) from control and Al stress imposed seedlings. Relative transcript level of *TaMATE1* homoeologues and *TaALMT1* genes was analyzed using the SYBR Premix Ex Taq (Takara, Japan) and the ABI 7500 Real-Time FAST PCR System (Applied Biosystems, USA) from three distinct biological and two technical replicates of each sample collected under control and Al stress conditions in the root and shoot tissues of both bread wheat genotypes.

The primers for amplification of respective *TaMATE1* homoeologue were designed to harbor specific nucleotide polymorphisms at their 3'-ends and the PCR efficiencies of both genes were computed by performing a 10-fold serial dilution of positive control template. The housekeeping gene of wheat *18S RNA* was used as an internal control reference gene. Relative expression levels were calculated by the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). For each gene, expression values were normalized against the tissue sample of a cultivar with the lowest transcript level under control conditions, where in case of *TaMATE1* and *TaALMT1* was Anahuac shoot samples.

REFERENCES

Akhunova, A.R.; Matniyazov, R.T.; Liang, H. and Akhunov, E.D. (2010). Homoeolog-specific transcriptional bias in allopolyploid wheat. *BMC Genomics*, 11:505, doi:10.1186/1471-2164-11-505.

Aniol, A. and Gustafson, J.P. (1984). Chromosome location of genes controlling aluminum tolerance in wheat, rye, and triticale. *Can J Genet Cytol*, 26: 701-705.

Barceló, J. and Poschenrieder, C. (2002). Fast root growth responses, root exudates, and internal detoxification as clues to the mechanisms of aluminium toxicity and resistance: A review. *Environ Exp Bot*, 48: 75-92.

Bottley, A.; Xia, G.M. and Koebner, R.M.D. (2006). Homoeologous gene silencing in hexaploid wheat. *Plant J*, 47: 897-906.

Collins, N.C.; Shirley, N.J.; Saeed, M.; Pallotta, M. and Gustafson, J.P. (2008). An ALMT1 gene cluster controlling aluminum tolerance at the Alt4 locus of rye (*Secale cereale* L.). *Genetics*, 179: 669-692.

Coutinho, J.F. (1989). Acidez do solo e toxicidade do aluminio. Tese de Doutoramento, UTAD, Vila Real, pp. 709.

Delhaize, E.; Ma, J.F. and Ryan, P.R. (2012). Transcriptional regulation of aluminium tolerance genes. *Trends Plant Sci*, 17(6): 341-48.

Delhaize, E; Ryan, P.R.; Hebb, D.M.; Yamamoto,Y.; Sasaki, T. and Matsumoto, H. (2004). Engineering high-level aluminum tolerance in barley with the ALMT1 gene. *Proc Natl Acad Sci USA*, 101(42): 15249-15254.

Eticha, D.; Zahn, M.; Bremer, M.; Yang, Z.; Rangel, A.F.; Rao, I.M. and Horst, W.J. (2011). Transcriptomic analysis reveals differential gene expression in response to aluminium in common bean (*Phaseolus vulgaris*) genotypes. *Ann Bot*, 105: 1119-1128.

Famoso, A.N.; Zhao, K.; Clark, R.T.; Tung, C.W.; Wright, M.H.; Bustamante, C.; Kochian, L.V. and McCouch, S.R. (2011). Genetic architecture of aluminum tolerance in rice (*Oryza sativa* L.) determined through genome-wide association analysis and QTL mapping. *PLoS Genet*, 7(8): e1002221. doi:10.1371/journal.pgen.1002221.

Feldman, M.; Levy, A.A.; Fahima, T. and Korol, A. (2012). Genomic asymmetry in allopolyploid plants: wheat as a model. *J Exp Bot*, 63(14): 5045-5059.

Fujii, M.; Yokosho, K.; Yamaji, N.; Saisho, D.; Yamane, M.; Takahashi, H.; Sato, K.; Nakazono, M. and Ma, J.F. (2012). Acquisition of aluminium tolerance by modification of a single gene in barley. *Nat Commun*, 6(3): DOI: 10.1038/ncomms1726.

Furukawa, J.; Yamaji, N.; Wang, H.; Mitani, N.; Murata, Y.; Sato, K.; Katsuhara, M.; Takeda, K. and Ma, J.F. (2007). An aluminum-activated citrate transporter in barley. *Plant Cell Physiol*, 48: 1081-1091.

Garcia-Oliveira, A.L.; Benito, C.; Prieto, P.; Menezes R.A.; Rodrigues-Pousada, C.; Guedes-Pinto, H. and Martins-Lopes, P. (2013). Molecular characterization of *TaSTOP1* homoeologues and their response to aluminium and proton (H^+) toxicity in bread wheat (*Triticum aestivum* L.). *BMC Plant Biol*, 13: 134, DOI: 10.1186/10.1186/1471-2229-13-134 .

Guedes-Pinto, H.; Ribeiro-Carvalho, C.; Pinto-Carnide, O.; Igrejas, G.; Nascimento, M.; Carnide, V.P.; Harrison, G.; Heslop-Harrison, J.S.; Gateau, I. and Branlard, G. (1998). Studies of Barbela wheat an old Portuguese landrace with rye introgression. In: Lelley, T. (Ed.), *Current Topics in Plant Cytogenetics Related to Plant Improvement*. Universitätsverlag, Vienna, 175-181.

Hoekenga, O.A.; Maron, L.G.; Piñeros, M.A.; Cançado, G.M.A.; Shaff, J.; Kobayashi, Y.; Ryan, P.R.; Dong, B.; Delhaize, E.; Sasaki, T.; Matsumoto, H.; Yamamoto, Y.; Koyama, H. and Kochian, L.V. (2006). AtALMT1, which encodes a malate transporter, is identified as one of several genes critical for aluminum tolerance in Arabidopsis. *Proc Natl Acad Sci USA*, 103(25): 9738-9743.

Huang, C.F.; Yamaji, N.; Mitani, N.; Yano, M.; Nagamura, Y. and Ma J.F. (2009). A bacterial-type ABC transporter is involved in aluminum tolerance in rice. *Plant Cell*, 21: 655-667.

Hvorup, R.N.; Winnen, B.; Chang, A.B.; Jiang, Y.; Zhou, X.F. and Saier, M.H. (2003). The multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) exporter superfamily. *Eur J Biochem*, 270: 799-813.

Kobayashi, Y.; Hoekenga, O.A.; Itoh, H.; Nakashima, M.; Saito, S.; Shaff, J.E.; Maron, L.G.; Piñeros, M.A.; Kochian, L.V. and Koyama, H. (2007). Characterization of *AtALMT1* expression in aluminum-inducible malate release and its role for rhizotoxic stress tolerance in Arabidopsis. *Plant Physiol*, 145(3): 843-852.

Kochian, L.V.; Hoekenga, O.A. and Piñeros, M.A. (2004). How do crop plants tolerate acid soils? Mechanisms of aluminum tolerance and phosphorous efficiency. *Annu Rev Plant Biol*, 55: 459-493.

Langridge, P.; Paltridge, N. and Fincher, G. (2006). Functional genomics of abiotic stress tolerance in cereals. *Brief Funct Genom Proteom*, 4(4): 343-354.

Ligaba, A.; Katsuhara, M.; Ryan, P.R.; Shibasaka, M. and Matsumoto, H. (2006). The BnALMT1 and BnALMT2 genes from Brassica napus L. encode aluminum-activated malate transporters that enhance the aluminum resistance of plant cells. *Plant Physiol*, 142: 1294-1303.

Liu, J.P.; Magalhaes, J.V.; Shaff, J. and Kochian, L.V. (2009). Aluminum-activated citrate and malate transporters from the MATE and ALMT families function independently to confer Arabidopsis aluminum tolerance. *Plant J*, 57: 389-399.

Liu, M.Y.; Chen, W.W.; Xu, J.M.; Fan, W.; Yang, J.L. and Zheng, S.J. (2013). The role of VuMATE1 expression in aluminium-inducible citrate secretion in rice bean (*Vigna umbellata*) roots. *J Exp Bot*, doi:10.1093/jxb/ert039.

Livak, K.J. and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta}C_T$. *Method*, 25(4): 402-408.

Magalhães, J.V. (2010). How a microbial drug transporter became essential for crop cultivation on acid soils: aluminium tolerance conferred by the multidrug and toxic compound extrusion (MATE) family. *Ann Bot*, 106: 199-203.

Magalhães, J.V.; Liu, J.; Guimarães, C.T.; Lana, U.G.; Alves, V.M.; Wang, Y.H.; Schaffert, R.E.; Hoekenga, O.A.; Pineros, M.A.; Shaff, J.E.; Klein, P.E.; Carneiro, N.P.; Coelho, C.M.; Trick, H.N. and Kochian, L.V. (2007). A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. *Nat Genet*, 39: 1156-1161.

Maron, L.G.; Piñeros, M.A.; Guimarães, C.T.; Magalhães, J.V.; Pleiman, J.K.; Mao, C., Shaff, J.; Belicuas, S.N. and Kochian, L.V. (2010). Two functionally distinct

members of the MATE (multi-drug and toxic compound extrusion) family of transporters potentially underlie two major aluminum tolerance QTLs in maize. *Plant J*, 61: 728-740.

Martins-Lopes, P.; Maçãs, B. and Guedes-Pinto, H. (2009). Portuguese bread wheat germplasm evaluation for aluminium tolerance. *Cereal Res Commun*, 37: 179-188.

Mochida, K.; Yamazaki, Y. and Ogihara, Y. (2004). Discrimination of homoeologous gene expression in hexaploid wheat by SNP analysis of contigs grouped from a large number of expressed sequence tags. *Mol Genet Genomics*, 270: 371-377.

Omote, H.; Hiasa, M.; Matsumoto, T.; Otsuka, M. and Moriyama, Y. (2006). The MATE proteins as fundamental transporters of metabolic and xenobiotic organic cations. *Trends Pharmacol Sci*, 27(11): 587-593.

Panda, S.K.; Baluška, F. and Matsumoto, H. (2009). Aluminum stress signaling in plants. *Plant Signal Behav*, 4(7): 592-597.

Paux, E.; Sourdille, P.; Salse, J.; Saintenac, C.; Choulet, F.; Leroy, P.; Korol, A.; Michalak, M.; Kianian, S.; Spielmeyer, W.; Lagudah, E.; Somers, D.; Kilian, A.; Alaux, M.; Vautrin, S.; Bergès, H.; Eversole, K.; Appels, R.; Safar, J.; Simkova, H.; Dolezel, J.; Bernard, M.; Feuillet, C. (2008). A Physical Map of the 1-Gigabase Bread Wheat Chromosome 3B. *Science*, 322 (5898): 101-104.

Pinto-Carnide, O. and Guedes-Pinto, H. (1999). Aluminum tolerance variability in rye and wheat Portuguese germplasm. *Genet Resour Crop Evol*, 46: 81-85.

Polle, E.; Konzak, C.F. and Kittrick, J.A. (1978). Visual detection of aluminium tolerance levels in wheat by hematoxylin staining of seedling roots. *Crop Sci*, 18: 823-827.

Poschenrieder, C.; Gunsé, B.; Corrales, I. and Barceló; J. (2008). A glance into aluminum toxicity and resistance in plants. *Sci Total Environ*, 400(1-3): 356-368.

Riede, C.R. and Anderson, J.A. (1996). Linkage of RFLP markers to an aluminum tolerance gene in wheat. *Crop Sci*, 36: 905-909.

Rose, A.; Lorenzen, S.; Goede, A.; Gruening, B. and Hildebrand, P.W. (2009). RHYTHM -a server to predict the orientation of transmembrane helices in channels and membrane-coils. *Nucleic Acids Res*, 37: W575-80.

Ryan, P.R.; Delhaize, E. and Randall, P.J. (1995). Malate efflux from root apices and tolerance to aluminium are highly correlated in wheat. *Aust J Plant Physiol*, 22: 531-536.

Ryan, P.R. and Delhaize, E. (2010). The convergent evolution of aluminium resistance in plants exploits a convenient currency. *Funct Plant Biol*, 37: 275–284.

Ryan, P.R.; Raman, H.; Gupta, S.; Horst, W.J. and Delhaize E (2009). A second mechanism for aluminum resistance in wheat relies on the constitutive efflux of citrate from roots. *Plant Physiol*, 149: 340-351.

Ryan, P.R.; Tyerman, S.D.; Sasaki, T.; Furuichi, T.; Yamamoto, Y.; Zhang, W.H. and Delhaize, E. (2011). The identification of aluminium-resistance genes provides opportunities for enhancing crop production on acid soils. *J Exp Bot*, 62(1): 9-20.

Ryan, P.R.; Raman, H.; Gupta, S.; Sasaki, T.; Yamamoto, Y. and Delhaize, E. (2010). The multiple origins of aluminium resistance in hexaploid wheat include *Aegilops tauschii* and more recent cis mutations to TaALMT1. *Plant J*, 64: 446-455.

Sasaki, T.; Yamamoto, Y.; Ezaki, B.; Katsuhara, M.; Ahn, S.J.; Ryan, P.R., Delhaize, E. and Matsumoto, H. (2004). A wheat gene encoding an aluminum-activated malate transporter. *Plant J*, 37: 645-653.

Sasaki, T.; Ryan, P.R., Delhaize, E.; Hebb, D.M.; Ogihara, Y.; Kawaura, K.; Noda, K.; Kojima, T.; Toyoda, A.; Matsumoto, H. and Yamamoto, Y. (2006). Sequence Upstream of the Wheat (*Triticum aestivum* L.) ALMT1 Gene and its Relationship to Aluminum Resistance. *Plant Cell Physiol*, 47(10): 1343-1354.

Sawaki, Y.; Kihara-Doi, T.; Kobayashi, Y.; Nishikubo, N.; Kawazu, T.; Kobayashi, Y.; Koyama, H. and Sato, S. (2013). Characterization of Al-responsive citrate excretion and citrate-transporting MATEs in *Eucalyptus camaldulensis. Planta*, 237(4): 979-989.

Schreiner, K.A.; Hoddinott, J. and Taylor, G.J. (1994). Aluminum-induced deposition of (1,3)-β-glucans (callose) inTriticum aestivum L. *Plant Soil*, 162 (2): 273-280.

Silva-Navas, J.; Benito, C.; Téllez-Robledo, B.; Abd El-Moneim, D. and Gallego, F.J. (2012). The ScAACT1 gene at the Qalt5 locus as a candidate for increased aluminum tolerance in rye (*Secale cereale L.*). *Mol Breed*, 30(2): 845-856.

Spyropoulos, I.C.; Liakopoulos, T.D.; Bagos, P.G. and Hamodrakas, S.J. (2004). TMRPres2D: high quality visual representation of transmembrane protein models. *Bioinformatics*, 20(17): 3258-3260.

Tamura, K.; Dudley, J.; Nei, M. And Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol*, 24: 1596-1599.

Taylor, G.J. (1991). Current views of the aluminum stress response: the physiological basis of tolerance. *Curr Top Plant Biochem Physiol*, 10: 57-93.

Tice, K.R.; Parker, D.R. and DeMason, D.A. (1992). Operationally defined apoplastic and symplastic aluminum fractions in root tips of aluminum-intoxicated wheat. *Plant Physiol*, 100: 309-318.

Thompson, J.D.; Higgins, D.G. and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*, 22: 4673-4680.

Tovkach, A.; Ryan, P.R., Richardson, A.E.; Lewis, D.C.; Rathjen, T.M.; Ramesh, S.; Tyerman, S.D. and Delhaize, E. (2013). Transposon-mediated alteration of TaMATE1B expression in wheat confers constitutive citrate efflux from root apices. *Plant Physiol*, 161(2): 880-892.

Yamaji, N.; Huang, C.F.; Nagao, S.; Yano, M.; Sato, Y.; Nagamura, Y. and Ma, J.F. (2009). A zinc finger transcription factor ART1 regulates multiple genes implicated in aluminum tolerance in rice. *Plant Cell*, 21: 3339-3349.

Yang, L.T.; Qi, Y.P.; Jiang, H.X. and Chen, L.S. (2013). Roles of organic acid anion secretion in aluminium tolerance of higher plants. *Biomed Res Int*, 173682: doi: 10.1155/2013/173682.

Yang, X.Y.; Yang, J.L.; Zhou, Y.; Piñeros, M.A.; Kochian, L.V.; Li, G.X. and Zheng, S.J. (2011). A de novo synthesis citrate transporter, Vigna umbellata multidrug and toxic compound extrusion, implicates in Alactivated citrate efflux in rice bean (Vigna umbellata) root apex. *Plant Cell Environ*, 34: 2138-2148.

Yokosho, K.; Yamaji, N. and Ma, J.F. (2011). An Al-inducible MATE gene is involved in external detoxification of Al in rice. *Plant J*, 68: 1061-1069.

Yokosho, K.; Yamaji, N. and Ma, J.F. (2010). Isolation and characterisation of two MATE genes in rye. *Funct Plant Biol*, 37:296-303.

Zhou, L.; Bai, G.; Ma, H. and Brett, C.F. (2006). Quantitative trait loci for aluminum resistance in wheat. *Molecular Breed*, 19:153-161.

Appendix-I

Table: Detail of primers used in present investigation.

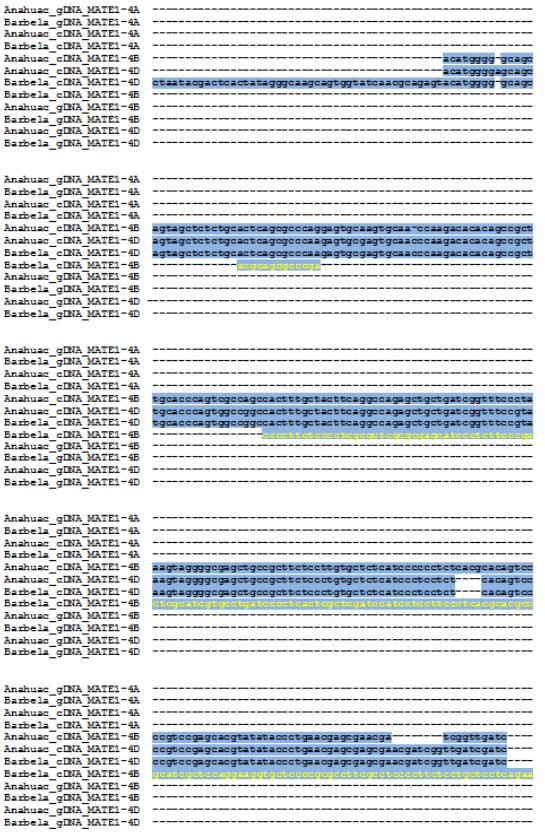
Assay	Primer name	Sequence (5'-3')
ORF amplification	TaMATE1_ORF_F	GTGGTGCTCGGTCTCTTCAT
	TaMATE1_ORF_R	CAAAGCAATAAAGCCCCATCTA
	TaALMT1_ORF_F	CATGAACATCAGGAGCGTCTT
	TaALMT1_ORF_R	TGGAGAGATCAGCACACTCTGT
5' UTR RACE	TaMATE1-5UTR	ACGAAGAGGTACATCCCTGTCGTCCTG
3'UTR RACE	TaMATE1-3UTR	GCATATGGGTCGCATTGACCATCTACA
5' and 3' UTR RACE	UNIVERSAL PRIMER*	CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT
Chromosomal mapping	TaMATE1-4A_F	CGCTCGTTAGCGTAACAACA
	TaMATE1-4A_R	CCGAGAAATGAGCCAACAAT
	TaMATE1-4B_F	GGCCGTCGACGTTCCTG
	TaMATE1-4B_R	CGGCGGCCAGAGCAA
	TaMATE1-4D_F	GTCGACGTTCCCACGGACGCTG
	TaMATE1-4D_R	GGGCACATTGCAGGCAT
Quantitative PCR	qTaMATE1_A_F	TTCCCGCAGACGCTGCCAC
	qTaMATE1_A_F	GAGCTCGTCGAGCTTGAAGACGAA
	qTaMATE1_B_F	CGTGTTCTGCAGGTGAGCATC
	qTaMATE1_B_R	GTTTATCGTCTGCGTGCCGGC
	qTaMATE1_D_F	GTTCTGGGGATGGGTCTCACA
	qTaMATE1_D_R	ATCGTCTGCGTGCCGGCGACA
	*qTa*ALMT1_*F***	ACACTTCTTGCGGACTTGGTTGATA
	*qTa*ALMT1_*R*	TTGTGTCCTCGGGGTTCTTA
	18SRNA_F	TCCACGAGGAATGCCTAGTAAGC
	18SRNA_R	ACAAAGGGCAGGACGTAGTC
cDNA 'Barbela 7/72/92' 4BL	TaMATE1_Barb_F	GCGCGAGCATCCCTCTTC
	TaMATE1_Barb_R	CCATCTCCGGCACAGAACAA
***TaALMT1* Promotor 'Anahuac'**	*TaALMT1-Prom_A_F***	CCTGGTTTTCTTGATGGGGGCACACAC
	TaALMT1-Prom_A_R	TGCTCACCATCTCGCCGTCGATCTCTCT
***TaALMT1* Promotor 'Barbela 7/72/92'**	*TaALMT1-Prom_B_F***	GGCCAAATGGAGACGCTC
	TaALMT1-Prom_B_R	CGCGCTCACATTCCTCCT

***From SMARTer™ RACE cDNA Amplification Kit (Clonetech, USA)**

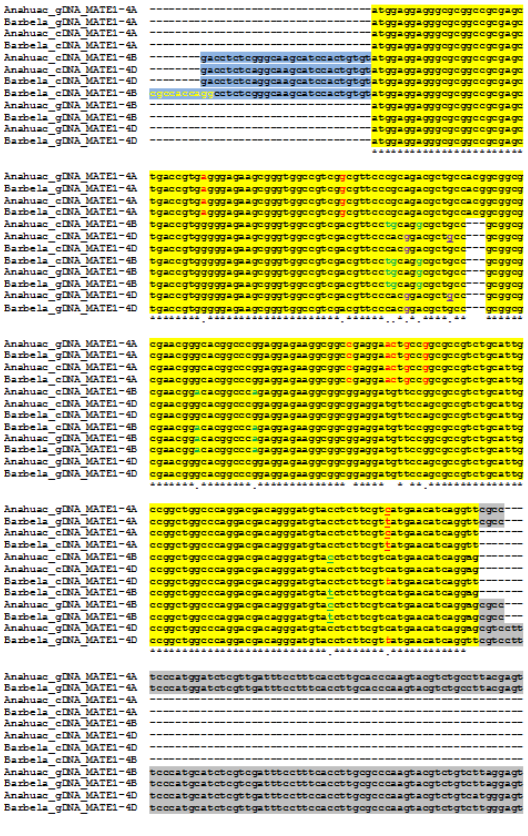
**** From Sasaki *et al.* (2004)**

Appendix-II

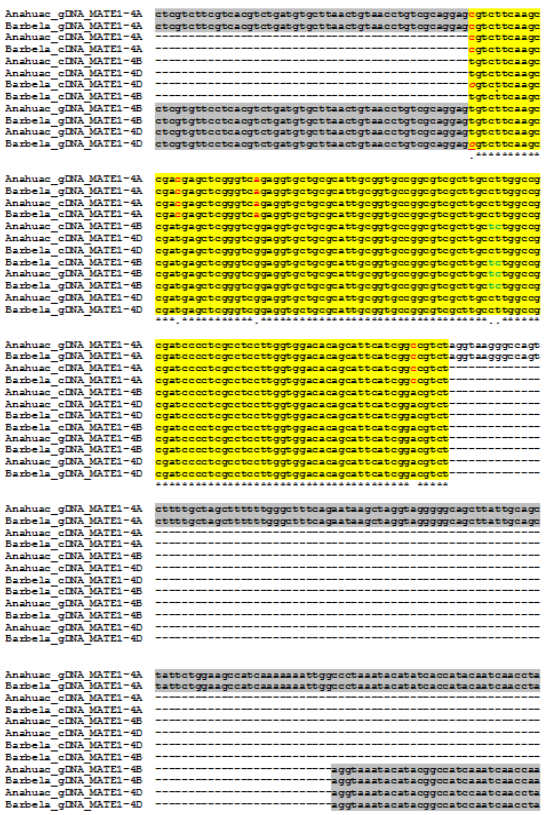
Multiple alignment of genomic and cDNA sequences of *TaMATE1* homoeologues in bread wheat genotype Anahuac and Barbela 7/72/92 (Highlighted in yellow, grey, blue and pink colour represents exon, introns, 5 UTR whereas coloured letters represent differences among the respective homoeologues)



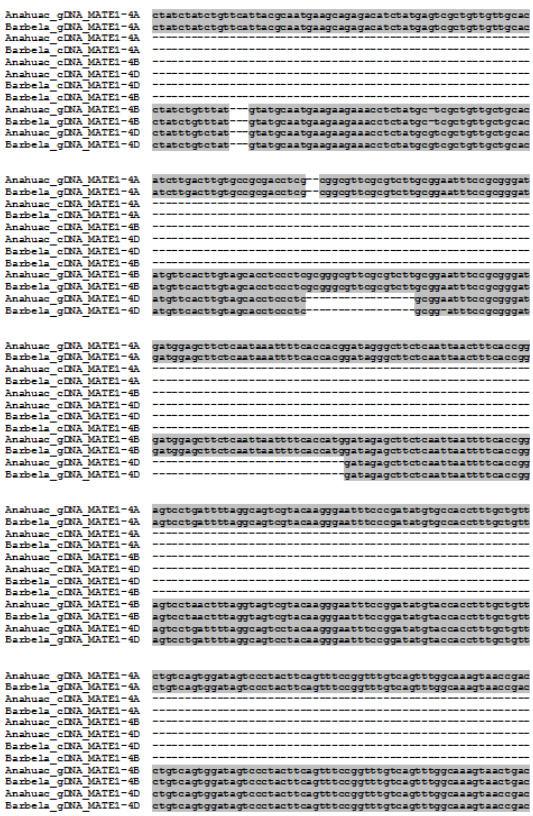
A



B

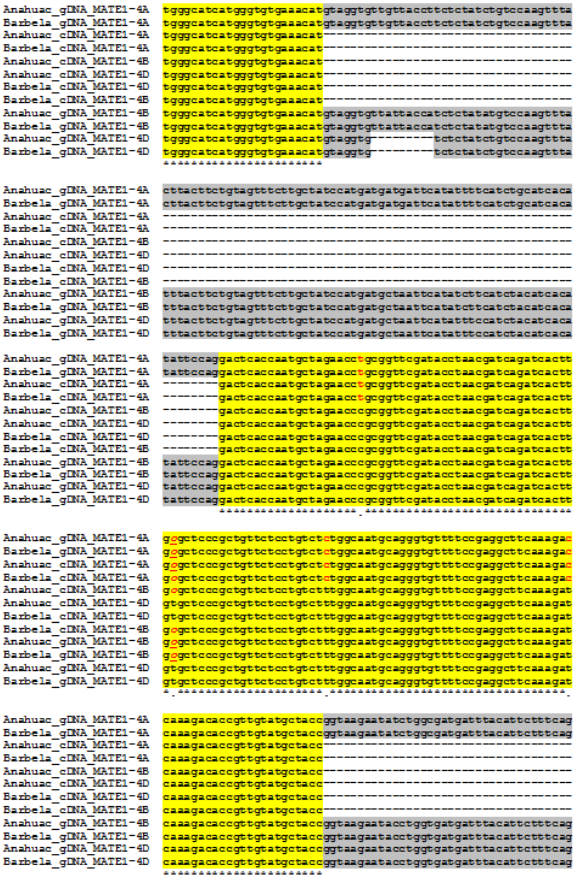


C

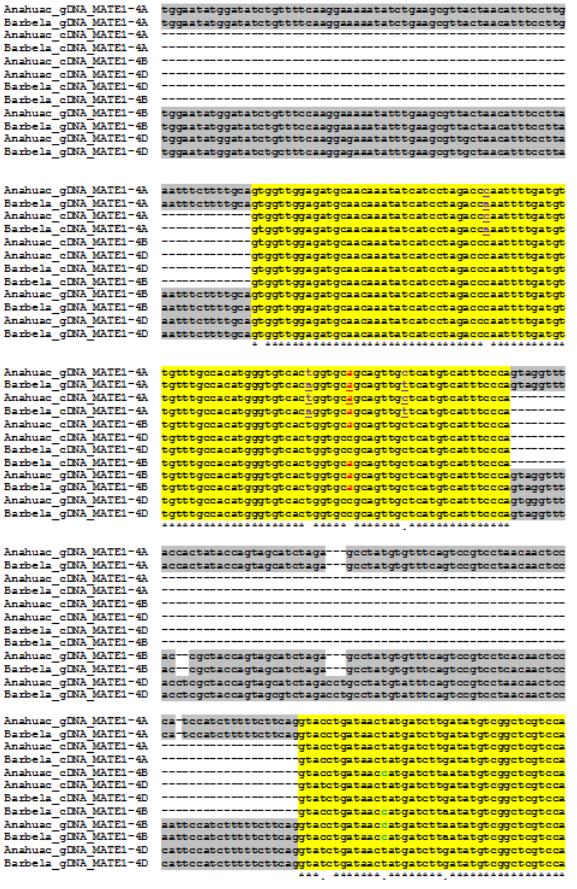


D

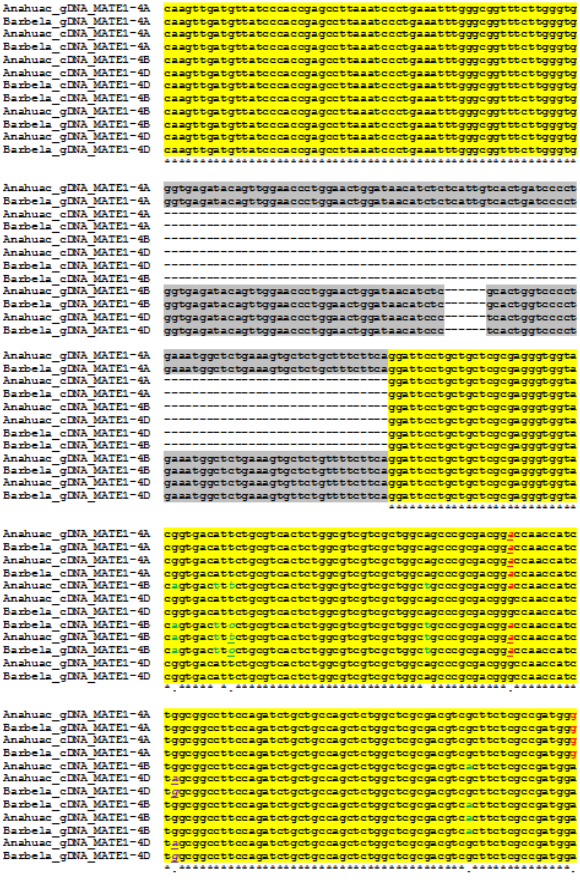

E
F
G
H


Anahuac_gDNA_MATE1-4A
Barbela_gDNA_MATE1-4A
Anahuac_cDNA_MATE1-4A
Barbela_cDNA_MATE1-4A
Anahuac_cDNA_MATE1-4B
Anahuac_cDNA_MATE1-4D
Barbela_cDNA_MATE1-4D
Barbela_cDNA_MATE1-4B
Anahuac_gDNA_MATE1-4B
Barbela_gDNA_MATE1-4B
Anahuac_gDNA_MATE1-4D
Barbela_gDNA_MATE1-4D

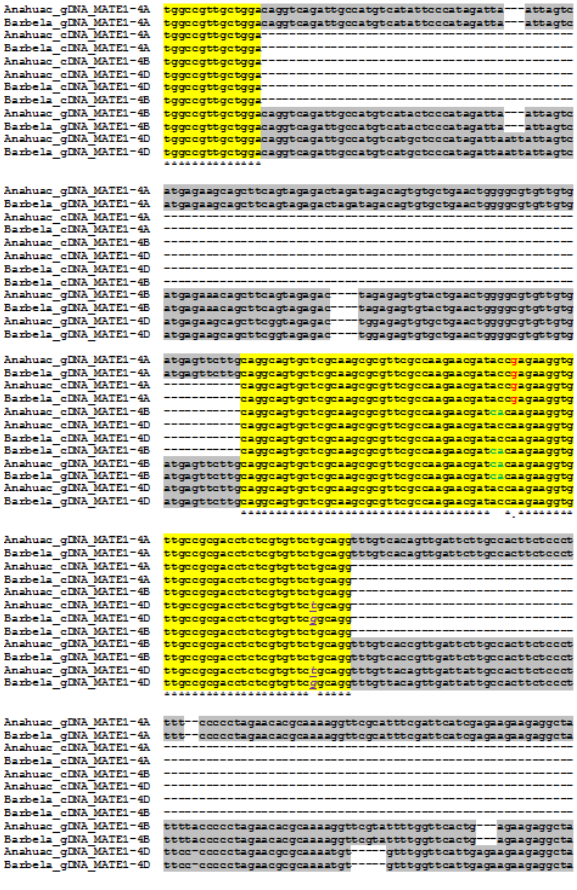
I


Anahuac_gDNA_MATE1-4A
Barbela_gDNA_MATE1-4A
Anahuac_cDNA_MATE1-4A
Barbela_cDNA_MATE1-4A
Anahuac_cDNA_MATE1-4B
Anahuac_cDNA_MATE1-4D
Barbela_cDNA_MATE1-4D
Barbela_cDNA_MATE1-4B
Anahuac_gDNA_MATE1-4B
Barbela_gDNA_MATE1-4B
Anahuac_gDNA_MATE1-4D
Barbela_gDNA_MATE1-4D

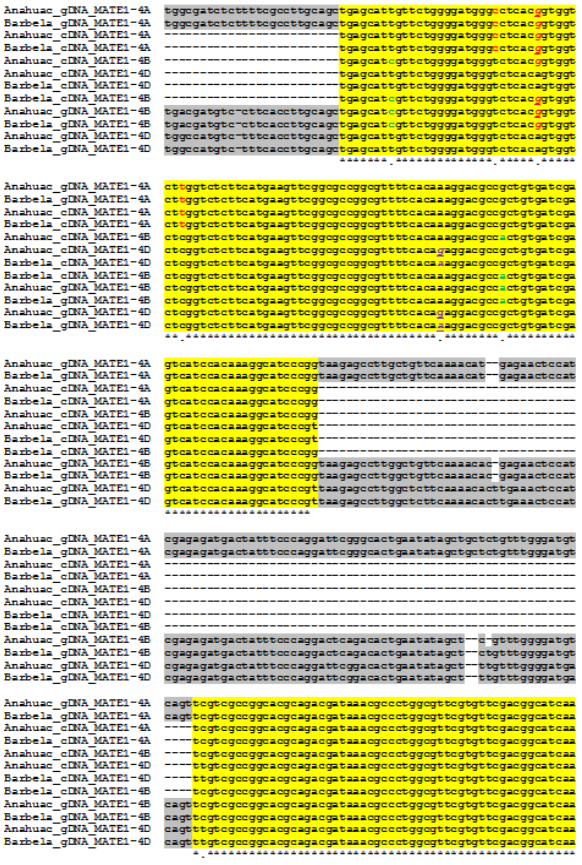
J


Anahuac_gDNA_MATE1-4A
Barbela_gDNA_MATE1-4A
Anahuac_cDNA_MATE1-4A
Barbela_cDNA_MATE1-4A
Anahuac_cDNA_MATE1-4B
Anahuac_cDNA_MATE1-4D
Barbela_cDNA_MATE1-4D
Barbela_cDNA_MATE1-4B
Anahuac_gDNA_MATE1-4B
Barbela_gDNA_MATE1-4B
Anahuac_gDNA_MATE1-4D
Barbela_gDNA_MATE1-4D

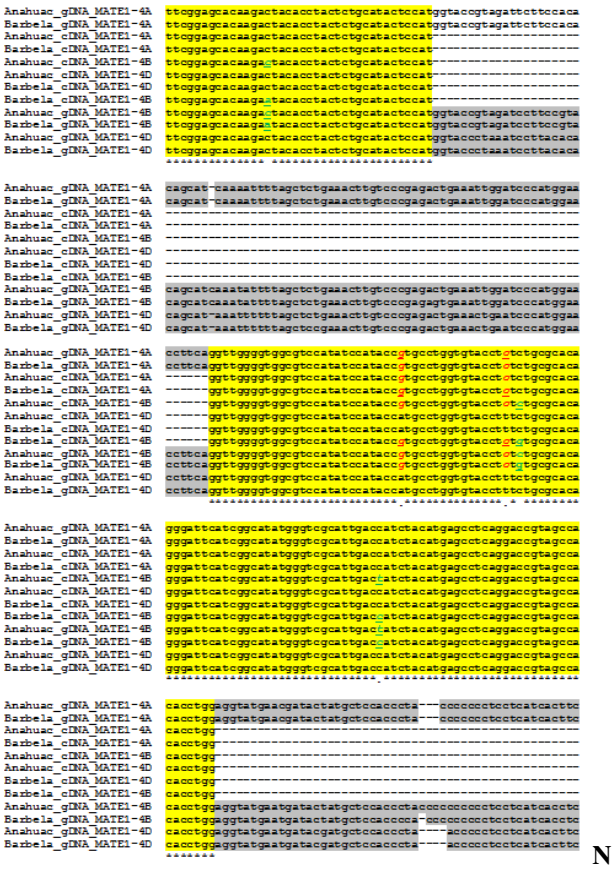
K


Anahuac_gDNA_MATE1-4A
Barbela_gDNA_MATE1-4A
Anahuac_cDNA_MATE1-4A
Barbela_cDNA_MATE1-4A
Anahuac_cDNA_MATE1-4B
Anahuac_cDNA_MATE1-4D
Barbela_cDNA_MATE1-4D
Barbela_cDNA_MATE1-4B
Anahuac_gDNA_MATE1-4B
Barbela_gDNA_MATE1-4B
Anahuac_gDNA_MATE1-4D
Barbela_gDNA_MATE1-4D

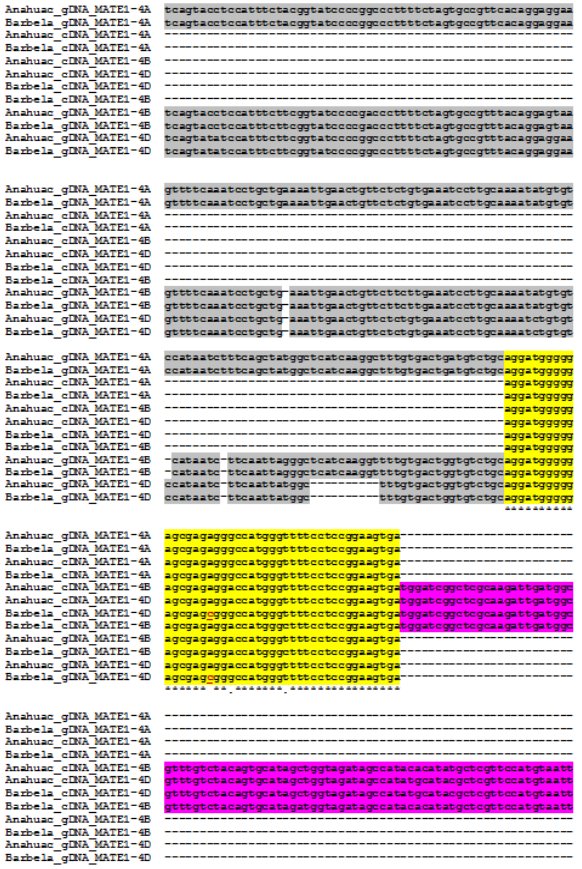
L



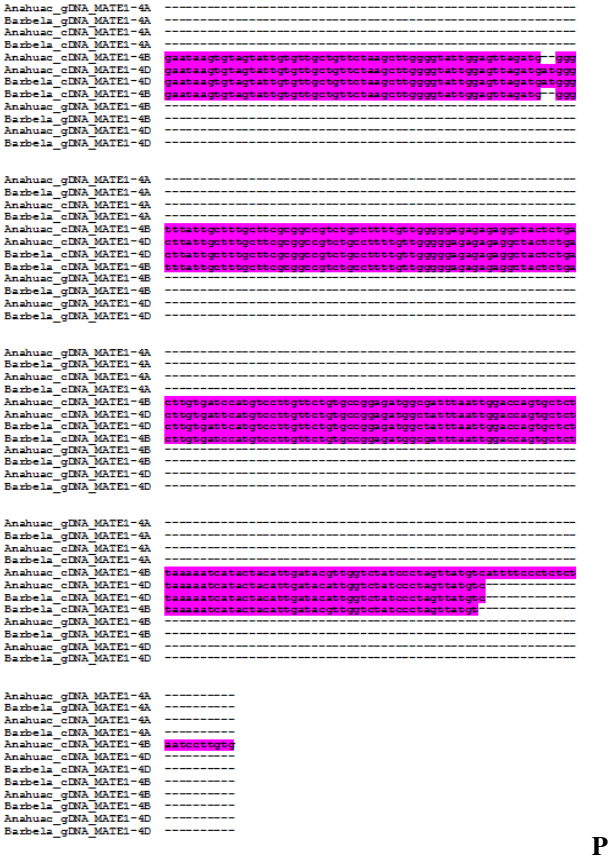
M



N



O



P

Appendix-III

Physiological characterisation of two bread wheat genotypes for Al stress

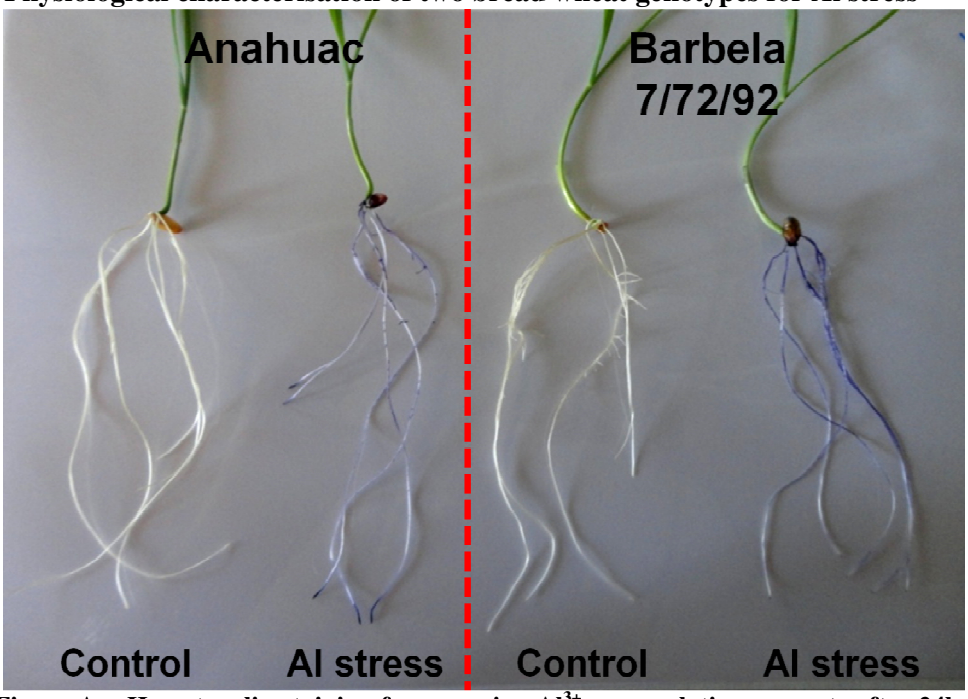


Figure A – Hematoxylin staining for accessing Al^{3+} accumulation on roots after 24h exposure in 74 uM of Al stress.

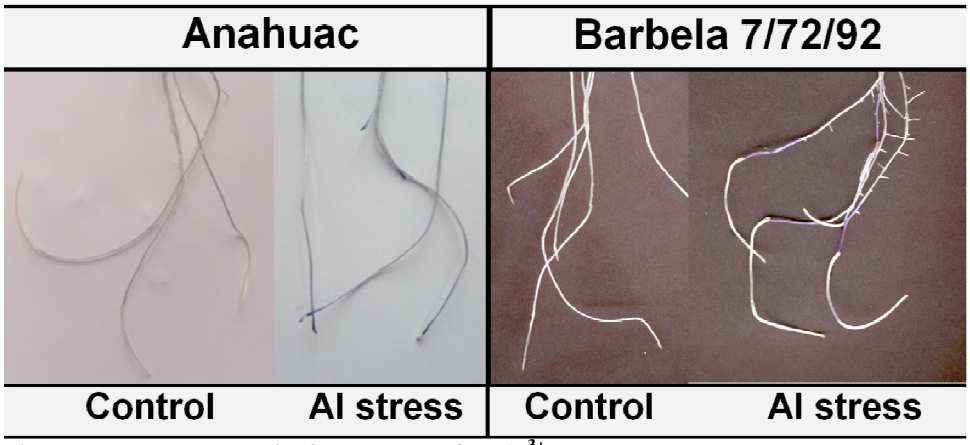


Figure B- Erychromocianin R assay for Al^{3+} tolerance assay based on root re-growth. Roots were exposed with (74 µM) or without (control) Al stress for 24h. After 24h plants were transferred to a nutritive solution without Al stress for 72h.

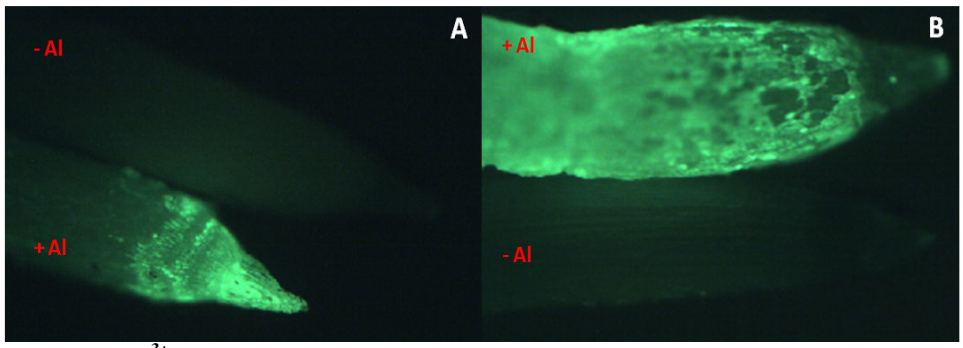


Figure C- Al^{3+} accumulation accessed by Morin dye in root tips with (74 µM) and without stress (0 µM). A) Resistant 'Barbela 7/72/92' roots without (- Al) and with (+ Al) stress and B) sensitive 'Anahuac' roots with (+ Al) and without (- Al) Al stress.

Appendix-IV

Table: Characterisation of selected thirty bread wheat genotypes for the presence of Sukkula-like transposon and their response under Al stress (% root re-growth under 2ppm and 5ppm Al).

Triticum	Genotypes	(% root re-growth)		Sukkula like-transposon	Origin
		2ppm	5ppm		
aestivum	Alva	1.9	1.1	Absent	Portugal
aestivum	Almansor	0.0	0.0	Absent	Portugal
aestivum	Anahuac	0.0	0.0	Absent	Poland
ferrugineum	Ardito	8.5	0.0	Absent	Portugal
aestivum	Barbela 7/72/92	31.4	40.4	**Present**	Portugal
aestivum	Barbela 55/88/92	45.9	41.4	**Present**	Portugal
aestivum	Barbela 61/94/92	41.2	38.3	**Present**	Portugal
aestivum	Barbela 58/70/92	22.0	41.2	**Present**	Portugal
aestivum	Barbela 6/93/92	40.9	31.0	**Present**	Portugal
aestivum	Barbela 35/94/92	44.8	35.1	**Present**	Portugal
aestivum	Barbela 16/95/92	38.0	17.9	**Present**	Portugal
aestivum	Barbela15/92/92	25.1	16.3	**Present**	Portugal
aestivum	BH1146	37.7	27.6	Absent	Brazil
aestivum	Egipcio	0.0	0.0	Absent	Portugal
aestivum	Eufrates	0.1	0.0	Absent	Portugal
aestivum	Fronteiriço	0.0	0.0	Absent	Portugal
aestivum	Jordão	3.4	0.0	Absent	Portugal
aestivum	Magueija	40.1	42.6	**Present**	Portugal
lutescens	Mocho de Espiga Branca	22.2	15.4	**Present**	Portugal
lutescens	Mocho Rapado	15.1	0.0	**Present**	Portugal
aestivum	Nabão	20.9	9.4	Absent	Portugal
aestivum	Roxo	0.9	0.0	Absent	Portugal
ferrugineum	Ruivo	41.3	45.0	**Present**	Portugal
aestivum	Sacho	37.3	2.7	Absent	Portugal
ferrugineum	Saloio	0.0	0.0	Absent	Portugal
aestivum	Sever	0.0	0.0	Absent	Portugal
hostianum	Viloso Mole	37.0	42.8	**Present**	Portugal
libycum	Mourisco preto	2.1	1.0	Absent	Portugal
aestivum	Atlas66*			Absent	USA
aestivum	Chinese Spring*			Absent	China

(Source: Pinto-Carnide and Guedes-Pinto, 1999; Martins-Lopes *et al.*, 2009)

*Atlas66 is well known Al tolerant genotype whereas Chinese Spring was considered as moderately tolerant. The DNA samples for transposon analysis of these two genotypes were received from Prof. César Benito laboratory, Departamento de Genética, Facultad de Biología, Universidad Complutense de Madrid (UCM), 28040-Madrid, Spain.

Note: For Sukkula-like transposon analysis, Anahuac and Barbela 7/72/92 were used as –ve (absent) and +ve (present) control, respectively.

Appendix-V

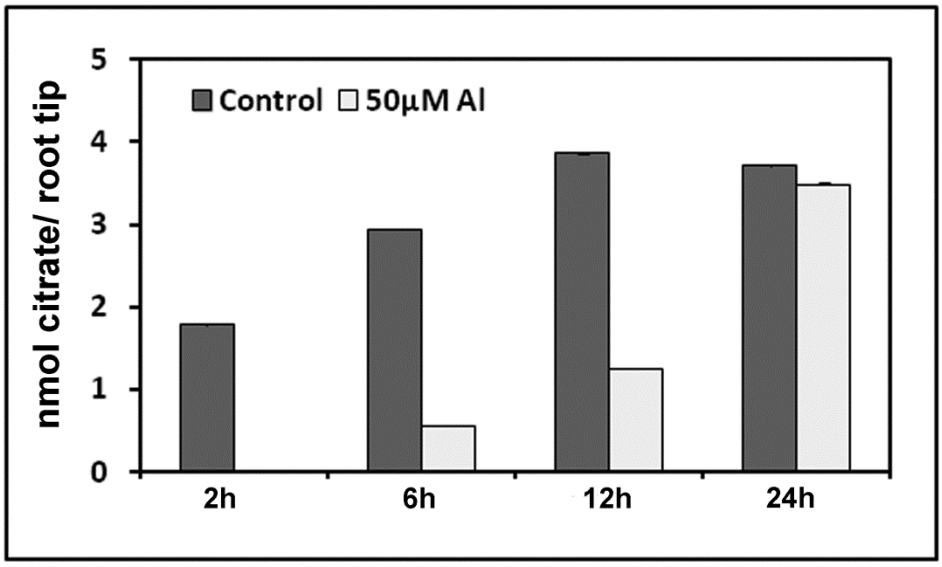


Figure. Citrate efflux from intact root of genotype Barbela 7/72/92 under control and Al stress

Appendix-VI

Secondary structure of TaMATE1 homoeologues in bread wheat

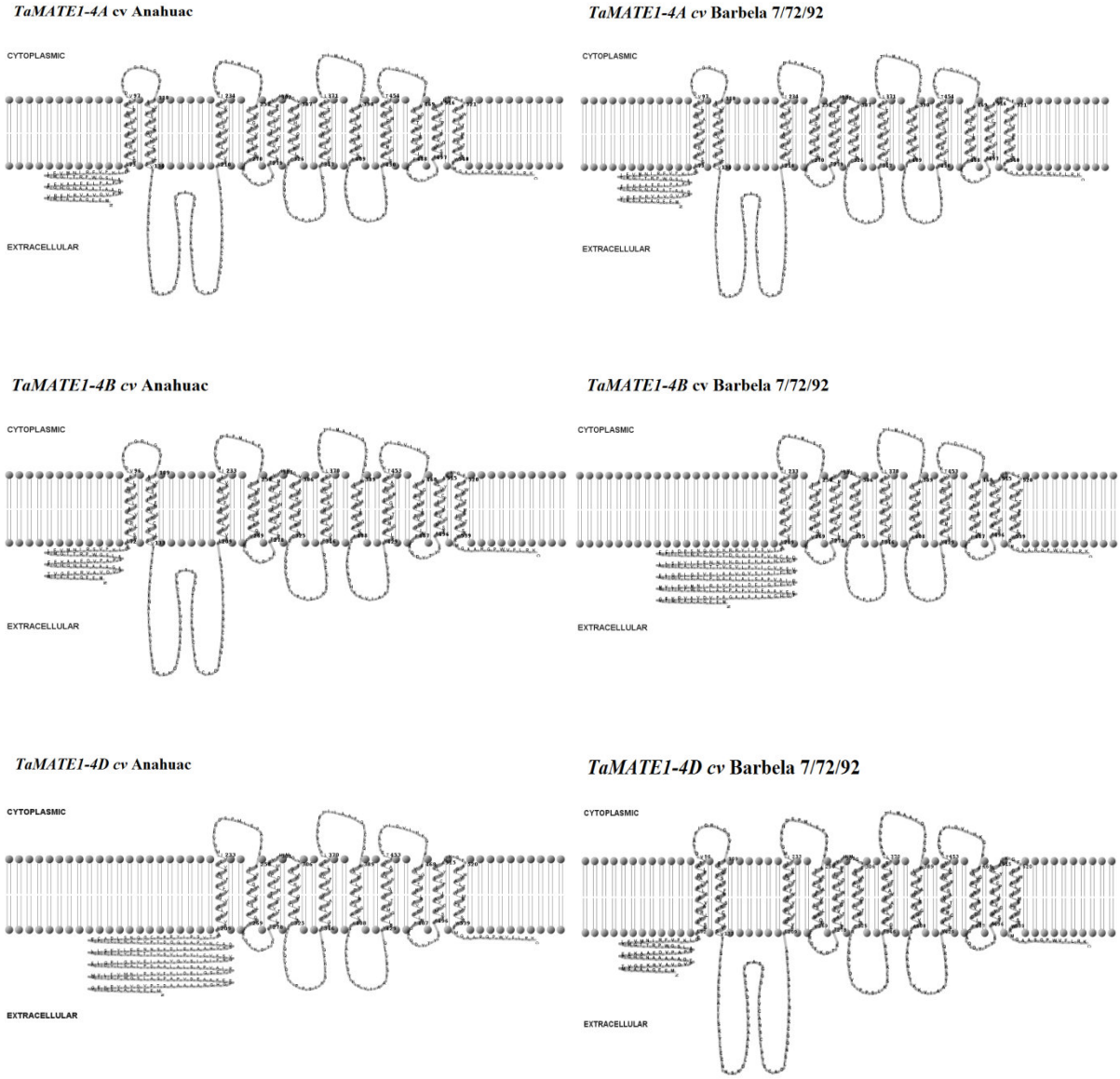


Figure: Secondary structure of TaMATE1 homoeologue from bread wheat genotypes Anahuac and Barbela 7/72/92. Hydropathy analysis predicts 10- 12 transmembrane domains in TaMATE1 homoeologues proteins.

Appendix-VII

The cDNA sequences of TaALMT1 from Al tolerant (Barbela 7/72/92) and sensitive (Anahuac) genotypes of bread wheat. The highlighted region with different colours represent the conserved sequence blocks in the upstream [(where light blue, red, green and dark blue colours represent block A, B, C and D, respectively (Sasaki et al., 2006)] of the *ALMT1* coding region (ORF is highlighted with yellow colour).

>cDNA_TaALMT1_var_ Barbela 7/72/92

GGCCAAATGGAGACGCTCCTACCACTATGGTTGCGCTCCACACGCACGAGCAGCAGGCAGTAGCACAGCGAACGCA
GATCGGGTCCATGATATTCATTCCCCAGGCCCCCAAATAAAACACGTACGGACGGTCGCAGGGGCGAGGTCGTATCT
GGCAGCGGCGTGCCCTGAGGAGGTCGGATCCGGCGGAGGCGCGCGCTCGGAGAGGCCGTATCCAGCGGAGGCGACC
GGCAGGGGGGGGGTCCACACGCACGAGCAGCAGGCAGTAGCACAGCGAACGCAGATCGGGTCCATGATATTCATTCC
CCAGGCCCCCAAATAAAACACGTACGGACGGTCGCAGGGGCGAGGTCGTATCTGGCAGCGGCGTGCCCTGAGGAGG
TCGGATCCGGCGGAGGCGCGCGCTCGGAGAGGCCGTATCCAGCGGAGGCGACCGGCAGGGGGGGGGTCCACACGCAC
GAGCAGCAGGCAGTAGCACAGCGAACGCAGATCGGGTCCATGATATTCATTCCCCAGGCCCCCAAATAAAACACGTA
CGGACGGTCGCAGGGGCGAGGTCGTATCTGGCAGCGGCGTGCCCTGAGGAGGTCGGATCCGGCGGAGGCGCGCGCT
CGGAGAGGCCGTATCCAGCGGAGGCGACCGGCAGGGGGGGGCACGCCGGCTCGTACGTAGCGCCGTCGTGGTGTCC
CCTGGCGACTGATTTGGGTAGCGCGGTGGATGGGTTAGGAGGAATGTGAGCGCGCCATGTGTTTGTCCGCCAGTGCC
TAACTGCCGCACTGCCTCAAAAGGCGCGTGCTAGGGTACTGTACTTAATTAGCAGCGCCGTACGCGCTTGAGCTTGCA
TTGCATCTGCCATGGATATTGATCACGGCAGAGAGATCGACGGCGAGATGGTGGGCACCATCGCCAGCTGCGGGCTG
CTGCTCCACTCGCTTCTCGCCGGGCTCGGGCGTCGCGCCGCCGGGTTCGCCCGGAAGGTGGGCGGCGCCGCGCGGGA
GGACCCGAGGCGGGTGGCGCACTCGCTCAAAGTCGGCCTGGCGCTCGCGCTGGTGTCCGTCGTCTACTTCGTCACGCC
GCTCTTCAACGGCCTCGGGGTGTCCGCGATATGGGCCGTGCTCACCGTCGTCGTCGTCATGGAGTACACCGTCGGTGC
CACGCTGAGTAAAGGCTTGAACAGAGCCTTGGCGACGTTGGTGGCTGGCTGCATCGCCGTCGGAGCTCATCAGTTAG
CTGAATTAGCTGAACGCTGTGGTGATCAGGGAGAGCCCATAGTGCTCACCGTGCTCGTCTTCTTCGTAGCGTCAGCGG
CGACGTTCTTGCGCTTCATCCCGGAGATCAAGGCCAAGTACGACTACGGCGTGACCATCTTCATACTGACCTTCGGTC
TGGTGGCCGTGTCGAGCTACAGAGTGGAGGAGCTCATCCAGCTCGCGCACCAGCGGTTCTACACCATAGCCGTCGGC
GTCTTCATCTGCATCTGCACCACCGTCTTCCTCTTCCCCGTCTGGGCCGGAGAGGACGTCCACAAGCTCGCCTCCGGC
AACCTCGACAAACTCGCTCAGTTCATTGAAGGAATGGAATTCAACTGCTTTGGCGAAAACAGTGTTGCAAATAATTTT
GGGGGAAAGGATTCCCCCCAAATGCACAAGAGCGTCCTTAATTCGAAGGCCACTGAGGACTCTTTGTGCACCTTTGC
CAAATGGGAGCCTCGTCATGGCCAGTTCAGATTTCGACACCCATGGAGCCAATACCAGAAGCTGGGAACTCTTTGTC
GCCAATGTGCGTCTTCTATGGAGGCTCTTGCTTCATATGTCATCACAACCTCAAAAACCCAGTGCCCTGCTGCAGCCA
ACCCTGAGCTATCATGTAAGGTTCGAAAAACATGTGGCGAAATGAGCTTGCATTCCTCCAAGGTGCTTAGGGATCTC
GCAATGGCAACTCGAACAATGACTGTGCCGTCTCCAGTGAATATCACCATGGCTACAGCCGTGAAAGCAGCGGAAAG
CCTCAGAAGCGAGCTTGCAGAGAACACGGCTCTGTTGCAAGTGATGCATGTGGCCGTCACCGCAACACTTCTTGCGG
ACTTGGTTGATAGGGTGAAGGAAATCGCGGAATGTGTTGATGTCCTAGCAAGACTGGCGCACTTTAAGAACCCCGAG
GACACAAAAAATGTCGTTGTCAGTACCGTGAGTCGAGGGATAGACGAACCTTTGCCTGACGTGGTTATTTTGTAA

>cDNA_TaALMT1_var_ Anahuac

CCTGGTTTTCTTGATGGGGGCACACACATCGACGATGGACGAGCAGGGGAGCGCGGCCTCATCTTGAACCAGATGTC
GGGACCGGCGACGACCACCGCCCCTACGCCCAAGCCGAGCCCGAGCACTGAAGCTATGCCCACGCCGATGCCGAGC
CCGAGCACTGAAGCTGCGCCCACGCCGACGCCGAGCCAGAGCACTGAAACTGCGCACACGCCGACCTCGGCCAGCCT
TGCGGTAGAGGAGCCTGCCGTTTGATGAGCTCCATGTCTATTGTTCCTACCCTTTTGATTGCCGGACTTTGGCTATGTT
CGCTGGCCGACCCTGTGGCATGGCGGTCGCCGAACTGTGTCATTTTTTGGCAGCGGCAAAGTGAAGTTTGAATTTTTT
ATGCGTTTGGGGGCCGAAACCTGGGGGCGCGGCTAGGAATTCGATCACCCCCAAATCGAGAAAACCGTCGGCGCAA
ACGCCAAAAACCCATCACCAGGTGCAAGGTTGTTAGAAGACTCAGAATCGTGATTCTGAATCTGATCGGAGCTCCGT
TGGTAAGATCGCAAATCGTAGAATCTTCACTATCAGAATCGTAAAAACGTTATTCTAGATACTAAAATCGTATAATCA
AAAAAGTTACTTTGGATTGTAAAGTCGTAGAATCGTAAAACAAAATCGCGATTCTAAGAATCATGGCTGGGTGCGCT
GCGGGCCAAATGGCTGGAGATGCTCCTACCACTATGGTTGCGCTCCACACGCACGAGCAGCAGGCAGTAGCACAGCG
AATGCAGATCGGGTCCATGATATTCATTCCCCAGGCCCCCAAATAAAACACGTACGGACGGTCGCAGGGGCGAGGTC
GTATCTGGCAGCGGCGTGCCCTGAGGAGGTCGGATCCGGCGGAGGCGCGCGCTCGGAGAGGCCGTATCCAGCGGAG
GCGACCGGCAGGGGGGGGCACGCCGGCTCGTACGTAGCGCCGTCGTGGTGTCCCCTGGCGACTGATTTGGGCAGCGCG
GTGGATGGGTTAGGAGGAATGTGAGCGCGCCATGTGTTTGTCCGCCAGTGCCTAACTGCCGCACTGCCTCAAAAGGC
GCGTGCTAGGGTACTGTACTTAATTAGCAGCGCCGTACGCGCTTGAGCTTGCATTGCATCTGCCATGGATATTGATCA
CGGCAGAGAGATCGACGGCGAGATGGTGGGCACCATCGCCAGCTGCGGGCTGCTGCTCCACTCGCTTCTCGCCGGGC
TCGGGCGTCGCGCCGCCGGGTTCGCCCGGAAGGTGGGCGGCGCCGCGCGGGAGGACCCGAGGCGGGTGGCGCACTC
GCTCAAAGTCGGCCTGGCGCTCGCGCTGGTGTCCGTCGTCTACTTCGTCACGCCGCTCTTCAACGGCCTCGGGGTGTC
CGCGATATGGGCCGTGCTCACCGTCGTCGTCGTCATGGAGTACACCGTCGGTGCCACGCTGAGTAAAGGCTTGAACA
GAGCCTTGGCGACGTTGGTGGCTGGCTGCATCGCCGTCGGAGCTCATCAGTTAGCTGAATTAGCTGAACGCTGTGGTG
ATCAGGGAGAGCCCATAATGCTTACCGTGCTCGTCTTCTTCGTAGCGTCAGCGGCGACGTTCTTGCGCTTCATCCCGG
AGATCAAGGCCAAGTACGACTACGGCGTGACCATCTTCATACTGACCTTCGGTCTGGTGGCCGTGTCGAGCTACAGA
GTGGAGGAGCTCATCCAGCTCGCGCACCAGCGGTTCTACACCATAGCCGTCGGCGTCTTCATCTGCCTCTGCACCACC
GTCTTCCTCTTCCCCGTCTGGGCCGGAGAGGACGTCCACAAGCTCGCCTCCGGCAACCTCGACAAACTCGCTCAGTTC
ATTGAAGGAATGGAATTCAACTGCTTTGGCGAAAACAGTGTTGCAAATAATTTTGGGGGAAAAGATTTCCCCCAAAT
GCACAAGAGCGTCCTTAATTCGAAGGCCACTGAGGACTCTTTGTGCACCTTTGCCAAATGGGAGCCTCGTCATGGCCA
GTTCAGATTTCGACACCCATGGAGTCAATACCAGAAGCTGGGAACTCTTTGTCGCCAATGTGCGTCTTCTATGGAGGC
TCTTGCTTCATATGTCATCACAACCTCAAAAACCCAGTGCCCTGCTGCAGCCAACCCTGAGCTATCATGTAAGGTTCG
AAAAACATGTGGCGAAATGAGCTTGCATTCCTCCAAGGTGCTTAGGGATCTCGCAATGGCAACTCGAACAATGACTG
TGCCGTCTCCAGTGAATATCACCATGGCTACAGCCGTGAAAGCAGCGGAAAGCCTCAGAAGCGAGCTTGCAGAGAAC
ACGGCTCTGTTGCAAGTGATGCATGTGGCCGTCACCGCAACACTTCTTGCGGACTTGGTTGATAGGGTGAAGGAAAT
CGCGGAATGTGTTGATGTCCTAGCAAGACTGGCGCACTTTAAGAACCCCGAGGACACAAAAAATGTCGTTGTTAGTA
CCGTGAGTCGAGGGATAGACGAACCTTTGCCTGACGTGGTTATTTTGTAA

Appendix-VIII

Detail of MATE1 like proteins in different species used for phylogenetic analysis.

S. No.	Scientific Name	Sequence name and Protein ID	Protein length
1	*Arabidopsis thaliana*	A.thaliana_MATE1_At1g51340	507
2	*Arabidopsis lyrata*	A.lyriata_MATE_XP_002891631.1	514
3	*Capsella rubella*	C.rubella_MATE1_EOA40177.1	509
4	*Brachypodium distachyon*	B.distachium_FRDL1_ Bradi1g69770.1	559
5	*Cucumis sativus*	C.sativus_MATE1_XP_004142114.1	521
6	*Eucalyptus camaldulensis*	E.camaldulensis_BAM68467.1	502
7	*Fragaria vesca subsp. Vesca*	F.vesca_Mate1_XP_004293176.1	492
8	*Glycine max*	G.max_MATE1_ACC60274.1	555
9	*Hordeum vulgare*	H.vulgare_AACT1-BAF75822.1	555
10	*Lupinus alba*	L.alba_MATE1_AAW30732.1	531
11	*Medicago truncatula*	M.truncatula_MATE_ACJ85750.1	507
12	*Oryza sativa*	O.sativa_FRDL1_Os03g0216700	571
		O.sativa_FRDL4_BAL41687.1	599
13	*Populus trichocarpa*	P.trichocarpa_MATE_XP_002313336.1	493
14	*Prunus persica*	P.persica_MATE1_EMJ05165.1	494
15	*Ricinus communis*	R.communis__XP_002532163.1	546
16	*Secale cereal*	S.cereal_ScFRDL1_BAJ61741.1	554
17	*Setaria italica*	S.italica_FRDL1_Si034942m	564
		S.italica_FRDL4_Si004441m	382
18	*Sorghum bicolor*	S.bicolor_FRDL1_Sb01g042740	565
		S.bicolor_MATE1_Sb03g043890	600
19	*Solanum lycopersicum*	S.lycopersicum_MATE1_XP_004229626.1	525
20	*Triticum aestivum*	T.aestivum_Barbela_genome_A	554
		T.aestivum_Barbela_genome_B	553
		T.aestivum_Barbela_genome_D	553
21	*Triticum urartu*	T. urartu_MATE1_EMS52414.1	524
22	*Vitis vinifera*	V.vinifera_XP_002285445.1	513
23	*Zea mays*	Z.mmays_FRDL1_EU960608	380
24	*Aegilops tauschii*	A.Tauchii_MATE1_ EMS52414.1	463

>A.thaliana_MATE1_At1g51340
MATTQIFQETLYTFSVISVLKFDELGLEIARIALPAALALTADPIASLVDTAFIGQIGPVELAAVGVSIALFNQVSRIAIFPLVSITTSFVAEE
DACSSQDTVRDHKECIEIGINNPTEETIELIPEKHKDSLADEFKTSSSIFSISKPPAKKRNIPSASSALIIGGVLGLFAVFLISAAKPLLSFMG
VKHDSPMMRPSQRYLSLRSLGAPAVLLSLAAQGVFRGFKDTTTPLFATVIGDVTNIIDLPIFIFVFRLGVTGAATAHVISQYLMCGILLWKLMG
QVDIFNMSTKHLQFCRFMKNGFLLLMRVIAVTFCVTLSASLAAREGSTSMAAFQVCCLQVWLATSLLADGYAVAGQAILASAFAKKDYKRAAAT
ASRVLQLGLVLGFVLAVILGAGLHFGARVFTKDDKVLHLISIGLPFVAGTQPINALAFVFDGVNFGASDFGYAAASLVMVAIVSILCLLFLSST
HGFIGLWFGLTIYMSLRAAVGFWRIGTGTGPWSFLRS

>A.lyriata_MATE1_XP_002891631.1
MSEDGYNTDFPRNPLKIFFSDFRSVFKLDELGLEIVRIALPAALALTADPIASLVDTAFIGQIGPVELAAVGVSIALFNQVSRIAIFPLVSITT
SFVAEEDACSSQQDTVQDHKECIEAGINNPTEETQELIPEKNKDSLSDEFKTGSSIFSISKPPAKKRNIPSASSALIIGGFLGLFQAVFLISAA
KPLLSFMGVKHDSPMLRPAQRYLSLRSLGAPAVLLSLAAQGVFRGFKDTTTPLFATVIGDVTNIILDPIFIFVFRLGVTGAATAHVISQYLMCG
ILLWKLMGQVDIFNMSTKHLQLCRFMKNGFLLLMRVIAVTFCVTLSASLAAREGSTSMAAFQVCLQVWLATSLLADGFAVAGQAILASAFAKKD
YKRAAATASRVLQLGLVLGFVLAVILGAGLHFGARVFTKDDKVLHLISIGLPFVAGTQPINALAFVFDGVNFGASDFGYAAASLVMVAIVSILC
LLFLSSTHGFIGLWFGLTIYMSLRAAVGFWRIGTGTGPWSFLRS

>C.rubella_MATE1_EOA40177.1
MATTKISQETLFTSSLVISVLKFDELGLEIARIALPAALALTADPIASLVDTAFIGQIGPVELAAVGVSIALFNQVSRIAIFPLVSITTSFVAE
EDACSSQQNTVQDHKESIETGTNNTMEETQELIPENSKDSSSDESKTISSIFSVSKPPAKKRNIPSASSALIIGAFLGLFQAVFLISAAKPLLS
FMGVKHDSPMMRPAQRYLSLRSLGAPAVLLSLAAQGVFRGFKDTTTPLFATVIGDVTNIILDPIFIFVFRLGVTGAATAHVISQYLMCGILLWK
LMGQVDVFNMSTKHLQLCRFMKNGFLLLMRVIAVTFCVTLSASLAAREGSTSMAAFQVCLQVWLATSLLADGFAVAGQAILASAFAKKDYKRAA
ATASRVLQLGLVLGFVLAVILGAGLHFGARVFTKDDKVLHLISIGLPFVAGTQPINALAFVFDGVNFGASDFGYAAASLVMVAIVSIMCLLFLS
STLGFIGLWFGLTIYMSLRAAVGFWRIGTGTGPWSFLRS

>B.distachium_FRDL1_Bradi1g69770.1
MEDGAPTTMTQEKRLAVVVVAPEDPAAAPVLLANGAAAEEHKAGEDPPAPPTALLSGWPRRTGLYLFVMNVRSVFKLDELGSEVLRIAVPASLA
LAADPLASLVDTAFIGRLGSVEIAAVGVSIAIFNQVSKVCIYPLVSVTTSFVAEEDAIISKYLEENSSQDLEKAPVDSEASNVPVSGAECVNSC
IPTECTDLSNQGCKRKYIPSVTSALIVGSFLGLVQAVFLIFSAKVVLGIMGVKHDSPMLEPAVRYLKIRSLGAPAVLLSLAMQGVFRGFKDTKT
PLYATVVGDAANIILDPILMFVCHMGVTGAAVAHVVSQYLITMILICRLVQQVDVIPPSLKSLKFGRFLGCGFLLLARVVAVTFCVTLASSLAA
RDGATIMAAFQICCQLWLATSLLADGLAVAGQAVLASAFAKNDNNKVVAATSRVLQLSIVLGMGLTVVLGLFMKFGAGVFTKDAAVIDVIHRGI
PFVAGTQTINALAFVFDGINFGASDYTYSAYSMVGVAAISIPCLVYLSAHNGFIGIWVALTIYMSLRTIASTWRMGAARGPWVFLRKGW

>C. sativus_MATE1_XP_004142114.1
MAFSIMSEEDDPYPSWDKTKTPIRIFFKNARHVFKLDELGREIAQIALPAALALAADPVASLVDTAFIGQIGSVELAAVGVAIALFNQVSRIAI
FPLVSVTTSFVAEEDTIGSVSIEAEDNNDMESGFFTNDEKSSMIPQNGKGEDAHHSRKPLEKKFENSKVENGRRYIPSASSALVIGGVLGLIQA
IFLISGARPLLNFMGVKSDSLMMTPAQQYLTLRSLGAPAVLLSLAIQGVFRGFKDTKTPLYATVAGDATNIILDPIFIFVFRLGVSGAAIAHVI
SQYLIALILFWRLMGQVDLLPPSIKHLQFSRFLKNGFLLLMRVIAVTFCVTLAASLSARQGSTSMAAFQVCLQVWLTTSLLADGLAVAGQAILA
TAFAQNDHDKATAAASRVLQLGLFLGLMLAVFLGVGMTFGARLFTSDVDVLRLIGIGIPFVAATQPINALAFVFDGINFGASDFAYSACSMVLV
AIISIFCLFILSSTQGFIGIWVALTIYMSLRTLAGFGRIGTGTGPWYFLRI

>E.camaldulensis_BAM68467.1
MPLSMFFKDARNVFKKDELGIEIAQIALPAALALAADPVASLIDTAFIGHIGPTELAAVGVAIAIFNQVSKVAIFPLVSITTSFVAEEDTKERL
HIEAQKDENGDKWFPVSKEKDVEMEELLPQSDSTSKSSFTDTSFGKMADLDNKRRYIPSASSALVIGSILGILQTLFLIFAAKPILNYMGVKSD
SPMLMPAQKYLTLRSLGAPAVLLSLAMQGIFRGFKDTKTPLYATVVGDASNIILDRLFIFDFRMGISGAAIAHVISQYLISLILLWRLMSQVDL
LPPSIKDLKFERFLKNGLLLLVRVIAVTFCVTLAASLAARHGATSMAAFQVCLQIWLATSLLADGLAVAGQAILASAFARADYEKAMSTASRVL
QLGLAMGLVLSVVLVTGLQFASRLFTEDAGVLHLISVGIPFVAVTQPINALAFVFDGINYGASDFAYSAYSMVLVAVVSIACLCALSSTYGFIG
IWIALTIYMTLRTFAGFLRIGAGMGPWGFLNN

>F.vesca_Mate1_XP_004293176.1
MGRYVFNLDELGLEIATIALPAALALTADPIASLIDTAFIGQLGSVELAAVGVSIALFNQVSRIAIFPLVSVTTSFVAEEDAIGRVRLESNENI
HLESDSPTNGETKQLLPETDTGKNAYNSKLVNVSFDTVTTGQKKRYIPSASSAMVIGGILGFIQAIFLISGAKPLLSFMGVGSDSPMLKPAQQY
LMLRSLGAPAVLLSLAMQGVFRGFKDTKTPLHATVAGDATNIILDPIFIFVFRFGVNGAAIAHVISQYLMCIILFLRLMAQVDLIPPSAKHLQF
GRFLKNGFLLLMRVIAVTFCVTLAASLSARQGPIAMAAFQVCLQVWLATSLLADGLAVAGQAILASAFAKNDHDKAAVTAARVLQLGLALGLIL
AFLLGVGLQYGARLFTKDVNVLHLISIGIPFVAATQPINALAFVFDGVNFGASDFAYAALSMVLVAIVSIIFLFILSYTSGFIGIWVALTIYMS
LRSFAGFWRIGTGTGPWQFLRT

>G.max_MATE1_ACC60274.1
MQDIGAPSVSLQLCITSLNKLPLFVWFLKQHNSEFGVSDMAEKESIHSFGDWRRIPICTFFQDARLVFKADSLGREILSIALPAAMALTADPIA
SLVDTAFIGQIGPVELAAVGVSIALFNQVSRIAIFPLVSVTTSFVAEEDTLSGANPQTEEGRCLEAGQPTDTETKELLPQKVTGGNVHNSDFVG
ESFNIAKEERKRRHIPSASSAIFIGGILGLIQAIFLISAAKPLLNFMGVTSDSPMLHPAKQYLKLRSLGAPAVLLSLAMQGVFRGFKDTKTPLY
ATVAGDVTNIALDPLFMFVFRLGVSGAAIAHVISQYLISVILLWRLLEQVDLIPPSINHLQLDRFLKNGFLLLMRVIAVTFCVTLAASLAARQG
PTSMAAFQVCLQVWLAVSLLADGLAVAGQAILAGAFANKDFNRATATASRVLQMGLVLGLALAFILGTGSHFGAKIFTQDANVLHLIQIGIPFI
AATQPLNSLAFVFDGVNFGASDFAYSAFSMVVVAILSIISLLILLSAGGFIGIWVALTIYMGLRAFAGFLRIGTGSGPWEFLRSS

>*H.vulgare_MATE1*_BAF75822.1
MEEGAAASMMTGDKKWVAVVDVPADADAATAANGHGPEEKAAEDLPAALSGCPRTTGLYLFVMNIRSVFKLDELGSEVLRIAVPASLALAADPL
ASLVDTAFIGRLGSVEIAAVGVSIAIFNQVSKVCIYPLVSVTTSFVAEEDAIISKYLEENSSQDLEKASHVHSDACNVPASGPDTPVCANSCIP
TECTDLSNQGCKKRYIPSVTSALIVGSFLGLVQAVFLIFSAKFVLGIMGVKHDSPMLEPAVRYLTIRSLGAPAVLLSLAMQGVFRGFKDTKTPL
YATVVGDATNIILDPILMFVCHMGVTGAAVAHVISQYLITMILICRLVQQVDVIPPSLKSLKFGRFLGCGFLLLARVVAVTFCVTLASSLAARD
GPTIMAAFQICCQLWLATSLLADGLAVAGQAVLASAFAKNDHKKVIAATSRVLQLSIVLGMGLTVVLGLFMKFGAGVFTRDADVINVIHKGIPF
VAGTQTINALAFVFDGINFGAQDYTYSAYSMVGVASISIPCLVYLSAHKGFIGIWVALTIYMSLRTVASTWRMGAARGPWVFLRK

>L.alba_MATE1_AAW30732.1
MAENGTVHAPIKSSWMRPLLIFFKDARLAFKIDSLTKEILVIGIPAAIAVAADPLASLIDTAFIGHLGPVQLAATGASIALFNQALRITVFPLV
SITTSFVAEEDTKEKINALAAEKKLAEIIKADELEKGVTKENNNETPKESLAVNGEIKVLVDGTSKNVTNKGNAGKKKRRLASASTALLFGLLL
GLFQAAILILLEKPLLYAMGLKHNSPMLVPAGKYLRLRALGSPAVLLSMVMQGIFRGFKDATTPLYVILSGYALNVALDPILIFYCKLGIEGAA
ISHVLAQYVMALALLLILMKKMVLLPPGLKDLQIFRFLKNGGLVLARVVAVTFCITLSASLASRLGPIKMAGFQVCLQVWLTSSLLADGLAVAV
QAILACSFTEKNNEKVAAAAARTLQLGFILGVILFIFVGAGLYFGAGMFSNSILVVQFIKIGMPYVAATQPINSFAFVFDGIYYGSSDFAYSAY
SLILASIVSIGSLFLLYESNGFVGIWIGLTINMSLRMLAGVWRMGTGTGPWRFLRSHPMSV

>M.truncatula_MATE_ACJ85750.1
MAEKESLFSIGDWMRIPICTFFKDARLVFKLDDLGREILSIALPAAMALTADPIASLVDTAFIGQLGPVELAAVGVSIALFNQASRIFIFPLVS
VTTSFVAEEDALSDASSQVEENGCLEAATPPDAETKEFLPQKNSVVESFNVVKDDQHKRRQIPSASSALYFGGVLGLVQATILISAAKPLLNFM
GVTSDSPMLHHAQQYLKLRSLGAPAVLLSLAMQGVFRGFKDTKTPLYATVAGDLTNIALDPLFIFVFRMGVNGAAIAHVISQYLLSAIHLWSLN
KQVDLIPPSIKHMQFDRFAKNGFLLFMRVIAVTFCVTLSASLAAHHGSTSMAAFQVCLQVWLAVSLLADGLAVAGQAILAGAFANKDYEKASTT
ATRVLQMGMVLGLALAFILGTGLHFGAKLFTKDIDVLHLIRVGVPFVALTQPLNCLAFVFDGVNFGASDFAYSAFSMVIVAIISIICLLILSSA
GGFIGIWVALTIYMSLRAFAGFLRIGTGSGPWEFLRS

>O.sativa_FRDL1_Os03g0216700
MAGLKKMEEVTAAAAAVAASSTAEKRAAAVVVPDAALTMNGAAGAEEKTAAAAAAPEDLPAPAALSGWPRRVGLYLFVMNIRSVFKLDELGSEV
LRIAVPASLALAADPLASLVDTAFIGRLGSVEIAAVGVSIAIFNQVSKVCIYPLVSVTTSFVAEEDAIISKCIEENSSQDLEKASPVDSETNNL
PVSGPDKVECVNSCIPTECTNPSDQGCKRKYIPSVTSAVIVGSFLGLLQAVFLVFSAKFVLNIMGVKNDSPMLRPAVRYLTIRSLGAPAVLLSL
AMQGVFRGFKDTKTPLYATVVGDAANIILDPILMFVCHMGVTGAAVAHVISQYLITMILLCRLIRQVDVIPPSLKSLKFGRFLGCGFLLLARVV
AVTFCVTLASSLAARHGPTIMAAFQICCQLWLATSLLADGLAVAGQAVLASAFAKNDKGKVVVATSRVLQLSIVLGMGLTVVLGVGMKFGAGIF
TKDIDVIDVIHKGIPFVAGTQTINSLAFVFDGINFGASDYTYSAYSMVGVAAISIPCLVYLSAHNGFIGIWIALTIYMSLRTIASTWRMGAARG
PWVFLRK

>O.sativa_FRDL4_BAL41687.1
MARSSSAPESMEADHQLTATVTAAASGDMPVTEQQQQKQCMVAVAAPPATENADAAGGGGGDNGDHLPPTATSLLGGARRTGLHLFVLNARSAL
RLDELGAEVLRIAVPASLALTADPLASLIDTAFIGRIGSVEIAAVGVAIAVFNQVMKVCIYPLVSVTTSFVAEEDAILSKGAAGADDDNDDGHD
AKGHGASAAAVADPEKQQVVGVDSAETNGAEVSTAAVRTTDDKKAAAAGVGVGKCRRRFVPSVTSALIVGAFLGLLQAVFLVAAGKPLLRIMGV
KPGSPMMIPALRYLVVRSLGAPAVLLSLAMQGVFRGFKDTKTPLYATVTGDLANIALDPILIFTCRFGVVGAAIAHVISQYLITLIMLCKLVRK
VDVIPSSLKSLKFRRFLGCGFLLLARVVAVTFCVTLAASLAARHGATAMAAFQICAQVWLASSLLADGLAVAGQALLASAFAKKDHYKVAVTTA
RVLQLAVVLGVGLTAFLAAGMWFGAGVFTSDAAVISTIHRGVPFVAGTQTINTLAFVFDGVNFGASDYAFAAYSMVGVAAVTIPCLVLLSSHGG
FVGIWIALAIYMSVRAFASTWRMGAARGPWKFLRK

>P.trichocarpa_MATE1_XP_002313336.1
MGRYVLKLDELGLEIAQIAFPAALAFTADPIASLVDTAFIGQIGPVELAAVGVSIALFNQVSRIAIFPLVSVTTSFVAEEDAIGRVSPEEEDSE
SLETGSTVNSENKELIPQNYSAEGPCKAKSPVSSFGIDKIENERRCIPSASSALVIGAILGLIQATFLISGAKPLLNFMGVGSDSPMLGPAQQY
LTLRSLGAPAVLLSLAMQGVFRGFKDTKTPLYATVAGDVTNIILDPIFMFVFGLGVRGAAIAHVLSQYLISVILLWRLMKQVDLLPPSIKHLRL
GQFLRNGLLLLMRVVAVTFCVTLSASLAARQGSTSMAAFQVCLQVWLATSLLADGLAVAGQAILASAFAKKDYEKATATATRVLQLGLLLGLML
AAVLGLGLRFGARLFTSDADVLHMISIGIPFVAGTQPINALAFVFDGVNFGASDFAYSAYSMESLCSLIIPCLFLLSSSHKFIGIWVALTIYMS
LRALAGFWRIGTGTGPWNFLRNL

>P.persica_MATE1_EMJ05165.1
MGNMPLFVFFKDIRNVFKADELGLEIVHIALPTLLALLADPVASLVDTAFVGHIGPVELAAVGISIAVFNQVSKIAIFPLVSITTSFVAEEDAT
EELRANEQVHENVENGTPISMEMEMEELVPLVGKSSSTEMVKLECVRRHIPSASSALVVGSILGFIQAVFLIFAANPVLNYMGVDSNSPMLKPA
RQYLTLRSLGAPAVLLSLAVQGVFRGFKDTKTPLYATIVGDVANIILDPILMFVFHMGVRGAAIAHVVSQYLISLILLWKLNKQVDLLTSGVKD
LRFGQFLKNGFLLLVRVIAATFCVTLAAALAARQGPTTMAAFQVCLQIWLAASLLADGLAVAGQAILASAFARKDHSKAVATASRVLQLALVLG
LMLSIILMVVLQFGSRIFTKDINVLQLISLGIPFVAVTQPINALAFVFDGVNYGASDFAYSAYSMVLVALVSILCLFLLSSHGFVGIWVALSVF
MSLRALVGFWRIGTGMGPWSFLRE

>R.communis_XP_002532163.1
MAENGALQLTERKWKIPLLVFFRDARLVFKTDELGLEILRIAVPAAMALAADPVASLIDTAFIGHLGPVEIAAVGVSIAIFNQASKVTIFPLVS
ITTSFVAEEDTFQRITNESQNGEGSEKDLPKTRDIKEVVPEDVMLENLEKGSATDSEKKDSIPGDANCKATTCKSPSFFEGKSIKDEQKNNKGR
RHIPSASTALIVGGILGLVQAIFLIFCAKPLLSIMGVKSGSPMLTPARKYLTLRALGSPAVLLSLAMQGVFRGFKDTKTPLYATVAGDLANIIL
DPIFIFTCRLGVSGAAIAHVLSQYLISLILLWRLMKNVDLLPPSPKDLQFGRFLKNGFLLLARVIAATICVTLAASRAARLGSTRMAAFQICLQ
VWLTSSLLADGLAVAGQAIIACAFAEKDYQKATTAATRVLQMSFVLGLGLAAVVGIGLHFGDGIFSKDPNVLDIISIGIPFVAATQPINSIAFV
FDGVNFGASDFAYSAYSMVLVAVASIAAIFVLSKTGGFVGIWIALTIFMGLRTFAGVWRMGTGTGPWNFLRGKLLP

>S.cereal_ScFRDL1_BAJ61741.1
MEEGAAASMTVREKRVAVGVPADAATAAANGHGPEEKAAEELPAPSALSGWPRTTGMYLFVMNIRSVFKLDELGSEVLRIAVPASLALAADPLA
SLVDTAFIGRLGSVEIAAVGVSIAIFNQVSKVCIYPLVSVTTSFVAEEDAIISKYLEENNSKDLEKAAHVHSDACNVPASGGDTPVCANSCIPT
ECADPSNQGCKRRYIPSVSSALIVGSFLGLVQAVFLIFSAKVVLGIMGVKRDSPMLEPAVRYLTIRSLGAPAVLLSLAMQGVFRGFKDTKTPLY
ATVVGDATNIILDPILMFVCHMGVTGAAVAHVISQYLITMILICRLVQQVDVIPPSLKSLKFGRFLGCGFLLLARVVAVTFCVTLASSLAARDG
PTIMAAFQICCQLWLATSLLADGLAVAGQAVLASAFAKNDTKKVIAATSRVLQLSIVLGMGLTVVLGLFMKFGAGVFTKDAAVIDVIHRGIPFV
AGTQTINALAFVFDGINFGAQDYTYSAYSMVGVASISIPCLVYLSAHKGFIGIWVALTIYMSLRTIASTWRMGAARGPWAFLRK

>S.italica_FRDL1_Si034942m
MEDNNSAAAAASVVTTSVPAEKHVAIPAAAVAAMTNGGGSAAEESKAEEDMLPPAPALPCGPRKTGLHLFIMNIRSVFKLDELGSEVLRIAVPA
SLALAADPLASLIDTAFIGRLGSVEIAAVGVSIAIFNQVSKVCIYPLVSVTTSFVAEEDAIISKAVEENSSHDLEKASHVDSETNNSPVSGPDI
AECVNTCIPTECTELSNHGSKKKYIPSVTSALIVGSILGLLQAIFLVFSAKFVLSIMGVKSGSPMQGPAVRYLTIRSLGAPAVLLSLAMQGVFR
GFKDTKTPLYATVVGDATNIILDPILMFVCHMGVTGAAVAHVASQYLITLILLCRLVQQVDVIPPSIKSLKFGRFLGCGFLLLARVVAVTFCVT
LAASLAARQGPTIMAGFQICCQLWLATSLLADGLAVAGQAVLASAFAKNDNKKVVAATSRVLQLSIVLGMCLTVVLGLAMKFGAGIFTKDLPVI
EVIHKGIPFVAGTQTINSLAFVFDGINFGASDYTYSAYSMVAVASVSIPCLVYLSAHNGFIGIWIALTIYMSLRTIASTWRMGAARGPWTFLRK

>S.italica_FRDL4_Si004441m
MEEHQSAAHATPEAEQQPQVELATAVAPAAEEDGDGPAEGGDGSPAAAETALLGRPRRTGLHLLVMNIRSVFKLDELGAEVLGIAVPASLALTA
DPLASLIDTAFIGRLGSVEIAAVGVAIAVFNQVMKVCTYPLVSVTTSFVAEQDAILSKGAAAKVVDGEEEGENPGQHAAVAATDPEKQQSPEEA
AKNGDSNAEPSEAPPAELAGAEECAPAVIGRKGCKNRKFVSSVTSALIVGAFLGLFQTVLVAAGKPPLRLIGVKPGSSMMIPALRYLTLRALGA
PAVLLSLAMQGVFRGFKDTKTPLYATVAGDLANIMLDPILIFGCRIMGVIGAAIAHVLRSSGGKYHRHLTGDRTRHCASLPCTDFPKRTRDGQQ
LCYFHP

>S.bicolor_FRDL1_Sb01g042740
MEDNGAAAANAAPATSMLPAAEKHVAIPDDAAAAAAATNATAGEENKVEEEDAPEPALLACGPRKTGLHLFVMNIRSVFKLDDLGSEVLRIAVP
ASLALAADPLASLVDTAFIGRLGSVEIAAVGVSIAIFNQVSKVCIYPLVSVTTSFVAEEDAIISKAIEEKSSQDLEKASHVDSETNNLPASGPD
LAECVNSCIPTECTDLPNQGCKKRYIPSVTSALIVGSILGLLQAVFLVFSAKFVLNIMGVKSGSPMQKPAVRYLTIRSLGAPAVLLSLAMQGVF
RGFKDTKTPLYATVVGDAANIILDPILMFVCHMGVTGAAIAHVVSQYLITLILLCRLVQQVDVIPPSIKSLKFGRFLGCGFLLLARVVAVTFCV
TLAASLAARHGPTIMAGFQICCQLWLATSLLADGLAVAGQAVLASAFAKNDKKKVVAATSRVLQLSIVLGMGLTVVLGLAMRFGAGIFTSDLPV
IEVIHKGIPFVAGTQTINSLAFVFDGINFGASDYTYSAYSMVAVASVSIPCLVYLSVHNGFIGIWIALTIYMSLRTIASTWRMGAARGPWKFLR
K

>S.bicolor_MATE1_Sb03g043890
MEEHRSPAHAKPEAEQPPQQQVPAAMAVAVAVDVAAPAALQNSTAAPAENGDVAAAGAAENGTAASAANGDGGGSELLGGPRWTGLHLFVMNIR
SVFKLDELGAEVLGIAVPASLALTADPLASLIDTAFIGRLGSVEIAAVGVAIAVFNQVMKVCIYPLVSVTTSFVAEEDAVLSKGGAKVIDNGEE
EEELEAGQVGPEKHTAAAGADPEKQQQPADEEAAKNGGEGCAPAVVAGRSSGKKSGNRRFVPSVTSALIVGALLGLFQTVFLVAAGKPLLRLMG
VKPGSPMVMPALRYLTLRALGAPAVLLSLAMQGVFRGFKDAKTPLYAIVAGDAANIVLDPILIFGCRLGVIGAAIAHVLSQYLITLIMLSKLVR
KVDVVPPSLKCLKFRRFLGCGFLLLARVVAVTFCVTLAASLAARHGPTAMAAFQICTQVWLATSLLADGLAVAGQAMIASAFAKEDRYKVAATA
ARVLQLGVVLGAALTALLGLGLQFGAGVFTSDAAVIKTIRKGVPFVAGTQTLNTLAFVFDGINFGASDYAFSAYSMIGVAAVSIPSLIFLSSHG
GFVGIWVALTIYMGVRALASTWRMAAAQGPWKFLRQ

>S.lycopersicum_MATE1_XP_004229626.1
MAEVGNVACTQRWKIPIFVFFQDIWNVFKFDNLGMEILRIAFPAALALAADPIASLIDTAFIGHLGAVEIAAVGVSIAIFNQASKVTIFPLVNI
TTSFVAEEDTVRRMNEKEGIADLEKNDSTKQLVLTDGNNNDKTETKETVLAQECKTTKCDSLELKRVKRHIPSASTAILMGCILGVLQTIFLIF
LAKPILSLMGVKSGSAMLSPAKKYLTLRAIGAPAVLLSLAMQGVFRGFKDTKTPLFATVVGDLTNIILDPIFIFVFHWGVSGAAIAHVLSQYLI
SIILLCKLMTEVHLLPPSTKDLQFSKFLKNGFWLLARVIAVTFCVTLSASLAARLGTTQMAAFQVCLQIWLTSSLLADGLAVAGQAILATSFAE
KDFQKAKAAGVRVLQMGLVLGFGLAAVVAIGLYFGSGVFSKDKNVIRLIIIAIPFVAGTQPINSLAFVLDGVNFGASDFAYSAYSMALVGALTI
TSEFVLSKTNGYIGIWIALTIFMVLRTFAGLWRMGTGTGPWRFLRIPVMSVEAKS

>T.aestivum_MATE1-4A_cv_Barbela
MEEGAAASMTVREKRVAVGVPADAATAAANGHGPEEKAAEELPAPSALSGWPRTTGMYLFVMNIRFVFKLDELGSEVLRIAVPASLALAADPLA
SLVDTAFIGRLGSVEIAAVGVSIAIFNQVSKVCIYPLVSVTTSFVAEEDAIISKYLEENNSKDLEKAAHVHSDACNVPASGGETPVCANSCIPT
ECADLSNQGCKRRYIPSVTSALIVGSFLGLVQAVFLIFSAKVVLGIMGVKHDSPMLEPAVRYLTIRSLGAPAVLLSLAMQGVFRGFKDTKTPLY
ATVVGDATNIILDQILMFVCHMGVTGAAVVHVISQYLITMILICRLVQQVDVIPPSLKSLKFGRFLGCGFLLLARVVAVTFCVTLASSLAARDG
PTIMAAFQICCQLWLATSLLADGLAVAGQAVLASAFAKNDTEKVIAATSRVLQVSIVLGMGLTVVLGLFMKFGAGVFTKDAAVIDVIHKGIPVV
AGTQTINALAFVFDGINFGAQDYTYSAYSMVGVASISIPCLVYLSAHKGFIGIWVALTIYMSLRTVASTWRMGAARGPWVFLRK

>T.aestivum_MATE1-4B_cv_Barbela
MEEGAAASMTVGEKRVAVDVPAGAAAAANGHGPEEKAAEDVPAPSALSGWPRTTGMYLFVMNIRSVFKLDELGSEVLRIAVPASLALAADPLAS
LVDTAFIGRLGSVEIAAVGVSIAIFNQVSKVCIYPLVSVTTSFVAEEDAIISKYIEENNSKDLEKAAHVHSDACNVPASGGDTPVCANSCIPTE
CADPSNQGCKRRYIPSVTSALIVGSFLGLVQAVFLIFSAKVVLGIMGVKHDSPMLEPAVRYLTIRSLGAPAVLLSLAMQGVFRGFKDTKTPLYA
TVVGDATNIILDPILMFVCHMGVTGAAVAHVISQYLITMILICRLVQQVDVIPPSLKSLKFGRFLGCGFLLLARVVAVTSCVTLASSLAARDGP
TIMAAFQICCQLWLATSLLADGLAVAGQAVLASAFAKNDHKKVIAATSRVLQVSIVLGMGLTVVLGLFMKFGAGVFTKDATVIDVIHKGIPVVA
GTQTINALAFVFDGINFGAQDYTYSAYSMVGVASISIPCLVYLSAHKGFIGIWVALTIYMSLRTVASTWRMGAARGPWVFLRK

>T.aestivum_MATE1-4D_cv_Barbela
MEEGAAASMTVGEKRVAVDVPTDAAAAANGHGPEEKAAEDVPAPSALSGWPRTTGMYLFVMNIRFVFKLDELGSEVLRIAVPASLALAADPLAS
LVDTAFIGRLGSVEIAAVGVSIAIFNQVSKVCIYPLVSVTTSFVAEEDAIISKYLEENNSKDLEKAAHVHSDACNVPASGGDTPVCANSCIPTE
CADLSNQGCKRRYIPSVTSALIVGSFLGLVQAVFLIFSAKVVLGIMGVKHDSPMLEPAVRYLTIRSLGAPAVLLSLAMQGVFRGFKDTKTPLYA
TVVGDATNIILDPILMFVCHMGVTGAAVAHVISQYLITMILICRLVQQVDVIPPSLKSLKFGRFLGCGFLLLARVVAVTFCVTLASSLAARDGP
TIMAAFQICCQLWLATSLLADGLAVAGQAVLASAFAKNDTKKVIAATSRVRQVSIVLGMGLTVVLGLFMKFGAGVFTKDAAVIDVIHKGIPFVA
GTQTINALAFVFDGINFGAQDYTYSAYSMVGVASISIPCLVYLSAHKGFIGIWVALTIYMSLRTVASTWRMGAASGPWVFLRK

>T. urartu_MATE1_EMS52414.1
MPPTPAGEAVPYVSPRPSQRWSVFKLDELGSEVLRIAVPASLALAADPLASLVDTAFIGRLGSVEIAAVGVSIAIFNQVSKVCIYPLVSVTTSF
VAEEDAIISKYLEENNSKDLEKAAHVHSDACNVPASGGETPVCANSCIPTECADLSNQGCKRRYIPSVTSALIVGSFLGLVQAVFLIFSAKVVL
GIMGVKHDSPMLEPAVRYLTIRSLGAPAVLLSLAMQGVFRGFKDTKTPLYATVVGDATNIILDPILMFVCHMGVTGAAVAHVISQYLITMILIC
RLVQQVDVIPPSLKSLKFGRFLGCGEIQLEPWNWITSLIVTDPLEMTLKVLCFLQASSLAARDGPTIMAAFQICCQLWLATSLLADGLAVAGQA
VLASAFAKNDTKKVIAATSRVLQLSIVLGMGLTVVLGLFMKFGAGVFTKDAAVIDVIHKGIPFVAGTQTINALAFVFDGINFGAQDYTYSAYSM
VGVASISIPCLVYLSAHKGFIGIWVALTIYMSLRTVASTWRMGAARGPWAFLRK

>V.vinifera__XP_002285445.1
MAEEDGLCQSRDGKGIPLWIFLKDARFAFKMDELGLEIAQIAFPAALALTADPIASLVDTAFIGQLGPVELAAVGVSIAVFNQVSRIAIFPLVS
VTTSFVAEEDTIGILDSEPEVSKSVEMGSAVNGETKKLIPKGSGERPYDLEMHGSGHDTPKFESKRHIPSASAALVVGGILGLIQAIFLISGAK
PILNFMGVHSDSPMLAPAQEYLTLRSLGAPAVLLSLAMQGVFRGFKDTKTPLYATVAGDVTNIILDPIFMFVFHMGVGGAAIAHVISQYIISVI
LFWKLMQQVELLPPSTKVLRFGRFLKNGLLLLMRVIAVTFCVTLAASLAARQGPTSMAAFQVCLQVWLATSLLADGLAVAGQAILASAFAKQDY
SKATAAASRVLQLGLVLGLVLSSILGTGMQSAAKLFTKDLSVLHLISIGIPFVAVTQPINSLAFVFDGVNFGASDFAYSAYSMVLVAIVSILCL
FILSSSFGFVGIWVALTIYVTLRAFAGFWRIGTGSGPWSFLRS

>Z.mays_FRDL1_EU960608
MAECVNSCIPTECAADPSGRQGRCEKRYVPSVTSALIVGSILGLLQAVFLVLSARFVLNIMGVKSGSPMQGPAVRYLTIRSLGAPAVLLSLAMQ
GVFRGFKDTKTPLYATVVGDAANIILDPILMFVCHMGVTGAAIAHVVSQYMITLILLCRLVQRVHVIPPSIKSLKFGRFLGCGFLLLARVVAVT
FCVTLAASLAARHGPTVMAGFQICCQLWLATSLLADGLAVAGQAVLASAFAKNDSKKVAAATSRVLQLSIVLGMGLTVVLGLAMRFGAGIFTSD
VPVIQVIHRGIPFVAGTQTINSLAFVFDGINFAASDYRYSAYSMVAVASVSIPCLLYLSAHNGFIGIWIALTIYMSLRTIASTWRMGAARGPWT
FLRN

>A.tauchii_MATE1_EMT11800.1
MYLFVMNIRSVFKLDELGSEVLRIAVPASLALAADPLASLVDTAFIGRLGSVEIAAVGVSIAIFNQVSKVCIYPLVSVTTSFVAEEDAIISKYL
EENNSKDLEKAAHVHSDACNVPASGGDTPVCANSCIPTECADLSNQGCKRRYIPSVTSALIVGSFLGLVQAVFLIFSAKVVLGIMGVKHDSPML
EPAVRYLTIRSLGAPAVLLSLAMQGVFRGFKDTKTPLYATVVGDATNIILDPILMFVCHMGVTGAAVAHVISQYLITMILICRLVQQVDVIPPS
LKSLKFGRFLGCASVETGECAELGRVVDEFLQAVLASAFAKNDTKKVIAATSRVLQLSIVLGMGLTVVLGLFMKFGAGVFTKDAAVIDVIHKGI
PFVAGTQTINALAFVFDGINFGAQDYTYSAYSMVGVASISIPCLVYLSAHKGFIGIWVALTIYMSLRTVASTWRMGAASGPWAFLRK

Molecular characterisation of *TaMATE2* homoeologues genes in bread wheat (*Triticum aestivum* L.)

ABSTRACT

Background:

At low pH, Al inhibits root growth and its toxic level may induce mineral nutrients deficiency in plants which restricts the crop productivity in acidic soils. Recently, members of MATE family have been implicated in Al tolerance by facilitating citrate efflux in plants. In addition, MATE genes may play specific physiological functions by transporting a narrow-range of endogenous metabolites in plants. The aim of the work was to clone and map the MATE2 gene in bread wheat and to further proceed to its molecular characterisation for Al tolerance in two diverse bread wheat genotypes, Barbela 7/72/92 and Anahuac, which exhibit contrasting response to Al toxicity.

Results:

In the present investigation, we cloned a member of the MATE gene family in bread wheat, named *TaMATE2*, which showed the highest homology with *ScFRDL2/ScMATE2* from rye. The amplification in Chinese Spring nullitetrasomic and ditelosomic lines revealed that *TaMATE2* is located on the long arms of homoeologous group 1 chromosomes. Multiple alignments of *TaMATE2* homoeologues cDNAs illustrated very high similarity among these genes. Thus, collective transcript expression of all the three *TaMATE2* homoeologues was quantified and a similar level of *TaMATE2* transcript was observed in both root apices and shoot tissues of Al resistant genotype Barbela 7/72/92. Contrarily, Al sensitive genotype Anahuac showed a higher expression level in the shoot tissues than in the root apices. Interestingly, the *TaMATE2* transcripts were gradually upregulated in response to Al in the Barbela 7/72/92 shoot tissues.

Conclusion:

TaMATE2 belongs to the MATE family and is located on the long arms of homoeologous group 1 chromosomes in bread wheat. *TaMATE2* genes have a high similarity with *ScFRDL2* from rye and also showed the typical secondary structure of MATE-type transporters maintaining all the 12 transmembrane domains. A time-course analysis of *TaMATE2* expression revealed its constitutive level in root apices of both bread wheat genotypes under Al stress. Al responsive transcript expression of *TaMATE2* in shoot tissues of the Al resistant genotype, Barbela 7/72/92, indicated that *TaMATE2* could play a major role in the shoot tissue in comparison to the roots of bread wheat. Considering the Al responsive *TaMATE2* expression in the present research and in previously reported loci for Al resistance on homoeologous chromosomes 1A and 1B, there is an indication that it plays a minor role in Al resistance in bread wheat.

Key words: *Triticum aestivum* L., Al, aluminium tolerance, *TaMATE2*, homoeologues, transcripts

INTRODUCTION

Aluminium (Al) is one of the most abundant metals in the earth crust and is considered as a rhizotoxic element under acidic conditions in plants. Thus, Al toxicity has been recognized as a major abiotic stress in low pH soils, which account for more than 40% of the earth's arable land and remains a serious obstacle for sustainable food production worldwide (von Uexkúll and Mutert, 1995). The most easily recognized symptom of Al toxicity is the inhibition of root growth particularly root apex (root cap, meristem, and elongation zone). Plant's roots under Al toxicity become impaired in water and mineral nutrient uptake, resulting in stunted and brittle root systems and poor root hair development (Barceló and Poschenrieder, 2002). The principal cause of Al toxicity has not yet been unambiguously identified, which is associated with alterations in a number of physiological processes and biochemical pathways in plants (Samac and Tesfaye, 2003).

Over the past decades, substantial advances in our understanding have been made on the physiological mechanism of Al tolerance in plants. Broadly, the strategies of plants to cope with Al toxicity have been categorized into external and internal Al detoxification (Taylor, 1991). Although, there has been considerable speculation that multiple Al tolerance mechanism are employed by different plant species to accomplish such strategies (Barceló and Poschenrieder, 2002). Among these mechanisms, organic acid anions (OA) of low molecular weight seem to play an important role in external and internal detoxification of Al in different plant species (Ryan et al., 2001; Kochian et al., 2004). These organic acids chelate the Al, and therefore either protect the roots (chelation in rhizhosphere) or cellular components (chelation in the cytosol) from the phytotoxic effects of Al (Kochian *et al.*, 2004).

Among the commonly secreted organic acids from roots of different plant species, malic and citric acids have been found in the bread wheat (Delhaize *et al.*, 1993; Ryan *et al.*, 2009). During the last two decades numerous studies revealed the role of two major gene families, namely ALMT (aluminium activated malate transporter) and MATE (multidrug and toxic compound extrusion) in Al tolerance in several plant species which encode membrane proteins and facilitate malate and citrate efflux, respectively (Ryan *et al.*, 2011).

Among multidrug efflux transporter families, members of the MATE family are extensively categorized and are widespread in both prokaryotic and eukaryotic organisms (Kuroda and Tsuchiya, 2009). Recent, genome wide analysis suggested the presence of at least 58 and 40 members of the MATE family in *Arabidopsis* and *Medicago truncatula*, respectively, most of which remain to be functionally characterised (Li *et al.*, 2002; Zhao and Dixon, 2009). So far, only few MATE homologues have been genetically identified in the *Poaceae* family members , such as: sorghum (*SbMATE1*; Magalhães *et al.*, 2007), barley (*HvAACT1*; Furukawa *et al.*, 2007), rice (*OsFRDL1*; Yokosho *et al.*, 2009; *OsFRDL4*; Yokosho *et al.*, 2011), rye (*ScFRDL1* and *ScFRDL2*; Yokosho *et al.*, 2010), maize (*ZmMATE1* and *ZmMATE2*; Maron *et al.*, 2010) and wheat (*TaMATE1*; Tovkach *et al.*, 2013; Garcia-Oliveira *et al.*, 2013) which are implicated in the constitutive or Al-activated citrate secretion from roots. These citrate transporters exhibit varying degrees of constitutive or element responsive (Al-activated/Fe-deficiency) expression and may play a role in the detoxification of Al in the rhizosphere or mineral nutrient acquisition and transport such as Fe and Zn from the roots to the shoots (Yang *et al.*, 2013).

Wheat is an essential component of the global food security which provides one fifth of the total calories of the world's population. So far the major genetic gain in performance of field crops is owed to conventional plant breeding (Reynolds *et al.*, 2011). In the future, consistent genetic gain in any crop relies on the accumulation of favourable alleles for highly complex traits which will depend on the knowledge of *priori* candidate genes. There has been an increasing focus on the identification of candidate genes associated with Al tolerance in wheat. Nevertheless, candidate gene discovery in bread wheat presents special challenges due to its huge polyploid genome size, abundance of repetitive sequences and very high level of sequence similarity among homoeologues of an individual gene (Lai *et al.*, 2012; Garcia-Oliveira *et al.*, 2013). We had previously report on the molecular characterisation of a candidate gene *TaMATE1* (Chapter 3) for Al tolerance in bread wheat. Here, we also report on the molecular characterisation of a novel candidate gene *TaMATE2* in bread wheat which might be responsible for Al tolerance in this crop.

RESULTS

Cloning of TaMATE2 and its secondary structure

```
            5′ AGCTCGGCTCTCCATCTCATGCGCGCGCGCTGCTCCCTGCCGCCTGCTTCCGCGAGGAGAGCCT
CCTCCTCCTCCACCAGCCCGCGGCGGGCGCCATGCACCTGCTCGGCGTCTTCTTCCATGGCGCGACGTGCGTAGGCCGG
AGGCAGCCCCGCTTATTACCCGATCCCTCCCGCCTCTGAGCTTTTCTTTTCCTCTAATTTGGTGTGTACGTTCTTCAGG
ATGCACCTGCTCGGC GTCTTCTTCCATGGC GCGACGCTCGCGTTC GAGCGGGACGACCTC GGGAAGGAGATCATG
 M  H  L  L  G   V  F  F  H  G   A  T  L  A  F   E  R  D  D  L   G  K  E  I  M
GGGATCGCGGTGCCG GGCGCGCTCGCGCTC ATGGCCGACCCGCTC GCGTCTCTCGTCGAC ACCGCCTTCATCGGC
 G  I  A  V  P   G  A  L  A  L   M  A  D  P  L   A  S  L  V  D   T  A  F  I  G
CGCATAGGTCCAGTA GAGATTGCAGCTGTA GGTGTATCTATTGTC GTGTTCAATCAAGTA ACAAGAATTGCAGTA
 R  I  G  P  V   E  I  A  A  V   G  V  S  I  V   V  F  N  Q  V   T  R  I  A  V
TTCCCCCTTGTGAGC GTCACAACATCATTT GTTGCAGAGGAAGAT GCTACTTCCAGTGAC AGAAACAAAGATGAA
 F  P  L  V  S   V  T  T  S  F   V  A  E  E  D   A  T  S  S  D   R  N  K  D  E
ATAAGTGGCGAGAAT GAACACAATGTTAGC GAAATGGATGAACTG ATTACTCATGAAGAG AATAATGCCACATCC
 I  S  G  E  N   E  H  N  V  S   E  M  D  E  L   I  T  H  E  E   N  N  A  T  S
GGCAAATCGTCTTTC GAAACTGACTCAAGT GAGATCAATACTGAG CATAAGCGGAAAAAA ATTCCATCAGTTTCT
 G  K  S  S  F   E  T  D  S  S   E  I  N  T  E   H  K  R  K  K   I  P  S  V  S
ACAGCATTACTACTT GGTGTGGTGCTTGGT CTTGTTGAAACTCTG CTGCTTGTTTCCTGT GCAAAGCCTATCTTA
 T  A  L  L  L   G  V  V  L  G   L  V  E  T  L   L  L  V  S  C   A  K  P  I  L
GACTTCATGGGTGTG AAAGCGGACACTGGA ATGTTGAATCCTGCA TTACAGTACTTAGTA CTCAGATCTTTAGGT
 D  F  M  G  V   K  A  D  T  G   M  L  N  P  A   L  Q  Y  L  V   L  R  S  L  G
GCTCCTGCTGTTCTT TTATCTCTGGCAATG CAAGGGGTATTTCGT GGACTTAAAGATACA AGGACGCCTCTATAT
 A  P  A  V  L   L  S  L  A  M   Q  G  V  F  R   G  L  K  D  T   R  T  P  L  Y
GCAACTGTGGCTGGA GATGCAATCAATATA GTTTTGGATCCAATA TTTATGTTTGTGTTT CAGTATGGTGTCAGT
 A  T  V  A  G   D  A  I  N  I   V  L  D  P  I   F  M  F  V  F   Q  Y  G  V  S
GGTGCAGCAGTTGCT CATGTTATATCACAA TATTTCATTGCTGCC ATACTCCTATGTAGA CTACGTCTGCAAGTT
 G  A  A  V  A   H  V  I  S  Q   Y  F  I  A  A   I  L  L  C  R   L  R  L  Q  V
GAACTACTGCCACCC AACTTGAAACATCTG CCGATTGGTCGGTTC CTTAAAAACGGTTCT CTGTTACTTGCCAGA
 E  L  L  P  P   N  L  K  H  L   P  I  G  R  F   L  K  N  G  S   L  L  L  A  R
GTTATTGCCGCAACA TGCTGTGTAACACTA TCTGCATCAATGGCT GCACGGCTAGGTTCA ACTCAAATGGCTGCA
 V  I  A  A  T   C  C  V  T  L   S  A  S  M  A   A  R  L  G  S   T  Q  M  A  A
TTCCAGATTTGCTTG CAGATCTGGTTGGCA TCTTCCCTTCTTGCT GATGGATTGGCTTTT GCTGGACAGGCTATA
 F  Q  I  C  L   Q  I  W  L  A   S  S  L  L  A   D  G  L  A  F   A  G  Q  A  I
CTTGCAAGTGCATTT GCTCGTAAGGACCAT TCGAAGGCCAAGGCC ACAGCTTCCCGCTTA CTGCAGCTTGGGTTG
 L  A  S  A  F   A  R  K  D  H   S  K  A  K  A   T  A  S  R  L   L  Q  L  G  L
ATCTTGGGGCTTCTC CTGGGCCTACTTCTA GGAGTTGGCCTCCAT ACAGGTTCGAGATTA TTTACAGAAGACCAG
 I  L  G  L  L   L  G  L  L  L   G  V  G  L  H   T  G  S  R  L   F  T  E  D  Q
GGTGTACTGCATCAT ATCTACGTGGCAATA CCGTTTGTCGCTCTC ACTCAACCGATCAAC GCTTTAGCTTTTGTT
 G  V  L  H  H   I  Y  V  A  I   P  F  V  A  L   T  Q  P  I  N   A  L  A  F  V
TTCGACGGTGTCAAT TATGGTGCATCTGAT TTTGCATATGCTGCT TATTCAATGATACTT GTGGCGATTGTCAGC
 F  D  G  V  N   Y  G  A  S  D   F  A  Y  A  A   Y  S  M  I  L   V  A  I  V  S
ATTGCTTTCATTGTC ACCCTCGCCAATTAC AGTGGTTTCATTGGA ATCTGGATAGCCTTG TCGATCTACATGTGC
 I  A  F  I  V   T  L  A  N  Y   S  G  F  I  G   I  W  I  A  L   S  I  Y  M  C
CTCCGCATGTTTGCG GGGTTATGGAGGATT GGGACTGCACGAGGG CCATGGGCTTTCCTT CGCAGCTGAGCAAGC
 L  R  M  F  A   G  L  W  R  I   G  T  A  R  G   P  W  A  F  L   R  S  *
ATGCATGCATGTATATTCTCGTAGATAGTCTGACTGTAGGTAGAATATTTACACTTGTAAACAACTTTGGAGTAACTGT
GTTACTTGTATGCCTCAAGCCTGCCACCATGATCACCAGAGACCAGACTGATATATCTTTAACAATGCAACCTAGCACA
AGAATGACATATTGTACAAACGAATAAACAAAAAGAGTTGCACTTTAACTTTCC 3′
```

Figure 1. Nucleotide sequence of *TaMATE2-1D* from Barbela 7/72/92 and its deduced protein (amino acid sequences). The nucleotide sequence was presented over the deduced amino acid sequence. Note: translation stop codon was noted by *.

Recently, two MATE transporters *ScFRDL1*/*ScMATE1* and *ScFRDL2*/*ScMATE2* have been cloned in rye (Yokosho *et al.*, 2010). *ScFRDL2*/*ScMATE2* was used as query sequence in NCBI database to identify wheat expressed sequenced tags (ESTs). For the cloning of MATE2 in bread wheat, primers from wheat EST showing highest similarity with *Secale cereale ScFRDL2*/*ScMATE2* were used to amplify the *TaMATE2* in bread wheat using cDNA template from Al exposed seedlings of Barbela 7/72/92. The multiple alignments of putative TaMATE2 fragments suggested the mixed amplification from distinct wheat genomes (A, B, and D), because the coding sequences of all three genome sequences (A, B and D) seemed to be extremely similar and difficult to distinguish. In order to clone the complete ORF of TaMATE2, 5' and 3' UTR ends

were amplified in genotype Barbela 7/72/92 using RACE as described in material and methods. The bread wheat *TaMATE2* (homoeologue from genome D) nucleotide sequence had an open reading frame of 1494 bp which encodes a putative polypeptide of 497 amino acids, with a predicted molecular mass of 53.098 kD and a theoretical pI of 6.42.

TaMATE2 protein analyses with the RHYTHM-web server programme predicted twelve transmembrane domains in all the three homoeologues corresponding to amino acid residues 24-43, 48-67, 152-171, 190-209, 224-245, 250-269, 296-315, 332-355, 370-389, 406-426, 439-458 and 463-482 (Figure 2).

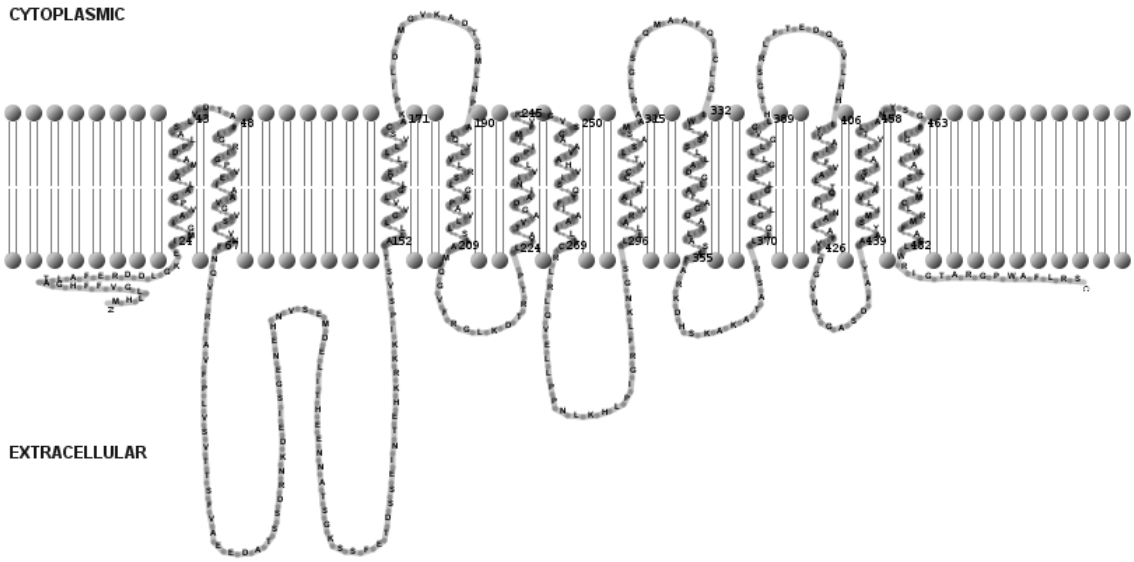


Figure 2: Secondary structure of TaMATE2-1D homoeologue from bread wheat genotype Barbela 7/72/92. Hydropathy analysis predicts 12 transmembrane domains.

Phylogenetic analysis of TaMATE2

The phylogenetic analysis of TaMATE2 homoeologues was performed by comparison of full length amino acids of TaMATE2 with those of MATE2-like transporters from other plant species. MATE transporters from higher plants can be clustered into two main groups: group I (MATE2 proteins from monocots) and group II (MATE2 proteins from eudicots) (Figure 3). Among members of Poaceae family, TaMATE2 (particularly from genome D) exhibited the highest homology to that of ScFRDL2/ScMATE2 from rye (97), followed by *Triticum urartu* (90%) and *Brachypodium distachyon* (82%). The MATE2 sequences from sorghum (*S. bicolor*), maize (*Z. mays*) and foxmillet (*S. italica*) formed a subgroup in the main group of

monocots. On the other hand, a second group consists of MATE2 like sequences from different eudicots which clearly subgrouped the members in different family.

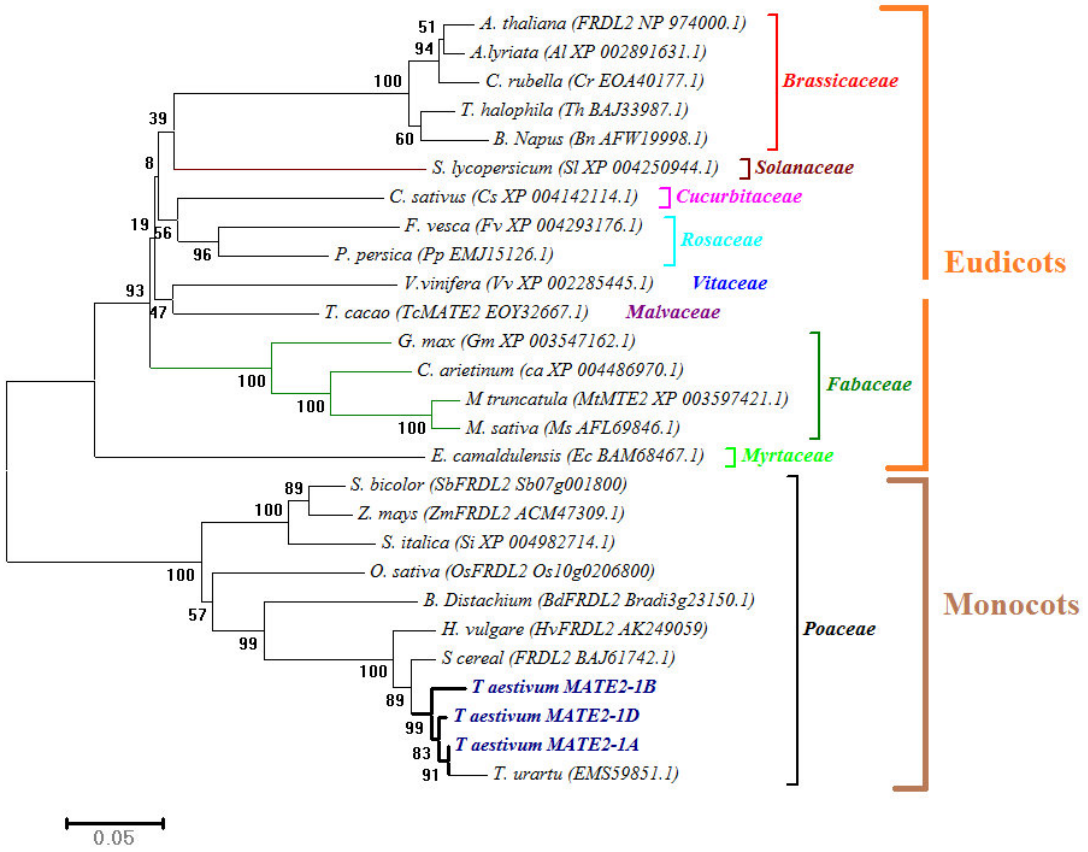


Figure 3: Phylogenetic tree based on the full amino acid sequences length showing the relationship of TaMATE2 with MATE2 like proteins from plant. The deduced amino acid sequences were aligned with MEGA 4.0. Evolutional distances were computed using the neighbor-joining method (Saitou and Nei, 1987). The scale bar indicates the distances as calculated from the multiple alignments.

Genetic localisation of TaMATE2

For localisation of *TaMATE2* on homoeologous chromosomes, locus specific primer pairs were designed [Appendix 1]. *TaMATE2* homoeologues were amplified from genomic DNA of a series of wheat nullitetrasomic lines with Chinese Spring as a positive control. On the basis of the presence or absence of PCR products visualized in agarose gels each *TaMATE2* homoeologue was assigned to the A, B and D genomes, respectively. Figure 4 illustrated that *TaMATE2* homoeologues were located at homoeologous chromosomes 1A, 1B and 1D and named *TaMATE2-1A*, *TaMATE2-1B* and *TaMATE2-1D*, respectively. Furthermore, ditelosomic lines for homoeologous group 1 chromosomes were used to assign *TaMATE2-1A*, *TaMATE2-1B* and *TaMATE1-D* on chromosomal arms. Homoeologue *TaMATE2-1A*, *TaMATE2-1B* and *TaMATE2-1D* specific primers did not amplify on ditelosomic lines Dt1AS, Dt1BS and Dt1DS,

respectively which confirmed that TaMATE2 homoeologues were located on the long arms of homoeologous group 1 chromosomes (Figure 4).

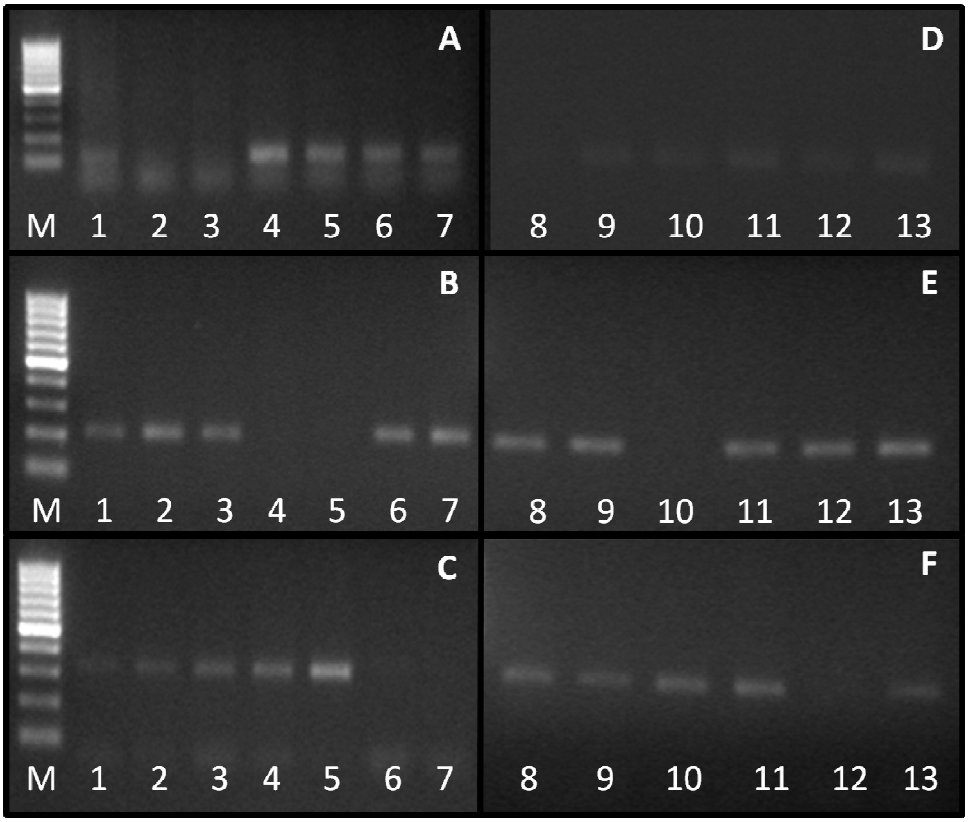


Figure 4. Mapping of *TaMATE2* on homoeologous group 1 chromosomes in hexaploid wheat. *TaMATE2* mapping on wheat chromosome from genomes A, B and D using nullitetrasomics lines [A, B and C, respectively]. Note: M (Marker); 1 (Chinese Spring control); 2 (N1AT1B); 3 (N1AT1D); 4 (N1BT1A); 5 (N1BT1D); 6 (N1DT1A); 7 (N1DT1B). Arms localization of *TaMATE2-1A* [D], *TaMATE2-1B* [E], *TaMATE2-1D* [F] using ditelosomic lines. NOTE: 8 (Dt1AS); 9 (Dt1AL); 10 (Dt1BS); 11 (Dt1BL); 12 (Dt1DS); 13 (Dt1DL).

Relative expression pattern of *TaMATE2* mRNA levels

In order to investigate the effect of Al stress on the transcription of *TaMATE2*, the expression of *TaMATE2* gene under different time-course of Al stress was quantified by real-time qPCR. A single primer pair for *TaMATE2* transcripts quantification was designed, once the *TaMATE2* homoeologues sequences are very similar. Thus, the quantified transcript results in the sum of all the three *TaMATE2* homoeologues transcription levels. Noticeably, the level of *TaMATE2* transcripts accumulated in the shoot tissues are higher when compared to the root apices in the Al sensitive genotype Anahuac whereas similar levels of transcripts were observed in both tissues of Al resistant genotype Barbela 7/72/92 (Figure 5). In addition, a significantly higher level of transcript ($P>0.01$) was observed in both root apices and shoot tissues of Barbela 7/72/92 compared with Anahuac. However, the relative transcript level

differences between these genotypes were much higher in the root apices tissue than in the shoot tissues. The gene expression time-course analysis revealed a constitutive transcript level of *TaMATE2* gene in the root apices of both studied genotypes under Al stress (Figure 5). Interestingly, the expression of *TaMATE2* increased gradually and reached its maximum at the 24 h time point in the shoot tissues of Al resistant genotype Barbela 7/72/92. Contrarily, a quite stable transcript level was observed in the shoot tissue of Al sensitive genotype Anahuac under Al stress during the first 12 h, followed by a down-regulation at the 24 h time-point.

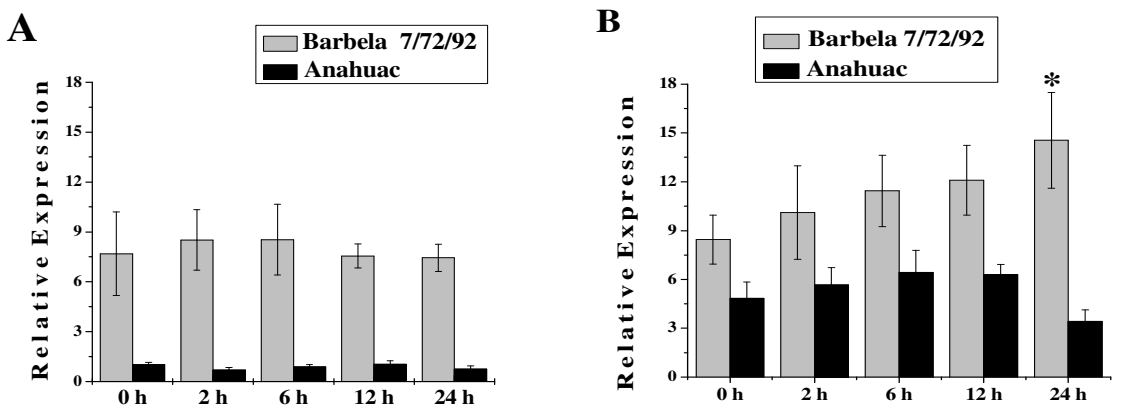


Figure 5: Relative transcript level of *TaMATE2* in root apices (A) and shoot tissues (B) of bread wheat genotypes Barbela 7/72/92 (Al resistant) and Anahuac (Al sensitive) with (74 µM) and without Al stress. Data were normalized using the *18SRNA* gene transcript levels. Values are means ± SD of three independent experiments. Asterisks indicate significant difference between the control and treatment sample in the respective tissue of a genotype (Student´s *t* test, * P < 0.05).

Discussion

During the last decades, major loci associated with genetic variability for Al tolerance have been identified in several crop species due to the rapid progress in the understanding of the physiological and molecular basis for Al resistance/tolerance mechanisms in plants (Ryan *et al.*, 2011). In the post genomic era, major emphasis is to explore the candidate genes which may be involved in Al tolerance mechanism and subsequently, their utilization for the development of Al-tolerant cultivars. At the physiological level, the role of organic anions efflux has seemed to be one of the Al tolerance mechanisms in different plants. Thus, the availability of genes encoding organic anions transport proteins could provide opportunities for further crop improvement considering yield incensement on acid soils. So far only ALMT and MATE families of membrane transporters have been implicated in the secretion of organic acid anions from plant roots in response to Al (Yang *et al.*, 2013).

In the present investigation, we cloned and mapped the *TaMATE2* homoeologues that are very similar to the Al-activated citrate transporter from rye (Yokosho *et al.*, 2010). Secondary structure analysis also indicated that wheat TaMATE2 maintained all the 12 transmembrane domains and share similar membrane topologies with the typical secondary structure of MATE-type transporters (Omote *et al.*, 2007). Furthermore, phylogenetic analysis of MATE2 like protein sequences from different plant species also indicated that TaMATE2 was a member of the large and complex family of transporters leading to the division of the MATE2 proteins into two groups (Figure 3). Group 1 comprises monocot MATE2 type transporters, representing the members of *Poaceae* family and, Group II consists of eudicots MATE2 like transporters.

With the advent of PCR technology, bread wheat aneuploids, such as nullitetrasomic and ditelosomic lines, have been extensively employed to assign genes and molecular markers to individual chromosomes and chromosome arms (Qi *et al.*, 2003). In this study, the analysis of nullitetrasomic Chinese Spring lines using PCR based amplification of *TaMATE2* genes from genomic DNA confirmed that *TaMATE2* genes are located on the homoeologous group 1 chromosomes (1A, 1B and 1D). Subsequently, the ditelosomic lines, which have 20 pairs of normal chromosomes and one represented by a pair of either short or long arms of a specific Chinese Spring chromosome revealed that *TaMATE2* genes were located on the long arm of homoeologous group 1 chromosomes (Figure 4).

However, in spite of the high similarity of wheat *TaMATE2* with the *ScFRDL2* (rye), the role of *TaMATE2* in Al tolerance can be different. The *ScFRDL2* was only expressed in the rye root tissues being significantly up-regulated by Al treatment, whereas in the shoot tissues no *ScFRDL2* expression was observed (Yokosho *et al.*, 2010). The individual *TaMATE2* homoeologues transcript was not investigated due to the high sequences similarity among these homoeologues from genome A, B and D. Although, a collective transcript expression of all the three *TaMATE2* homoeologues suggested that *TaMATE2* is expressed in both root apices and shoot tissues of bread wheat (Figure 5). Under control condition, a higher transcript level of *TaMATE2* was observed in the shoot tissues rather than in the root apices of Al sensitive genotype Anahuac; whereas approximately similar levels of transcripts were noticed in both tissues of Barbela 7/72/92, the Al resistant genotype. Interestingly, *TaMATE2*

expression seemed to be gradually up-regulated in the shoot tissues of Al resistant genotype Barbela 7/72/92 under Al stress. Al resistant genotype Barbela 7/72/92 has also showed the ability to accumulate higher amount of Al in its shoot tissues under Al stress (Silva *et al.*, 2010). Thus our results indicate that *TaMATE2* might be involved in internal Al detoxification mechanism through translocation of Al from the roots to the shoots of Barbela 7/72/92.

To date, no Al tolerance mechanisms other than malate and citrate efflux have been identified in wheat (Delhaize *et al.*, 1993; Ryan *et al.*, 2009). In the past, some physiological studies indicated that flavonoid-type phenolics could impart Al tolerance by dual mechanisms of Al chelation and anti-oxidant scavenging (Kidd *et al.*, 2001; Barcelo and Poschenrieder, 2002; Kochian *et al.*, 2004). In addition, Al induced high levels of flavonoid-type phenolic compounds in the shoot tissues of *Rumex acetosa* and in the roots of Al tolerant maize cultivar also suggested that these compounds may contribute to detoxification of the Al that have surpassed the exclusion barriers (Tolrá *et al.*, 2005; Tolrá *et al.*, 2009). Recent evidences are clearly indicating that the members of MATE family have been implicated in the export of a wide range of substrates in plant species and some MATE proteins may have specifically diverged to confer Al tolerance by a physiological mechanism based on phenolics (Yazaki *et al.*, 2008; Zhao and Dixon, 2009; Magalhães, 2010; Zhao *et al.*, 2011).

Conclusion

Recently, the genome-wide association mapping has detected loci associated with Al resistance on chromosomes 1A and 1B in bread wheat (Raman *et al.*, 2010). The candidate genes underlie these loci are still unknown. Therefore, it would be interesting to investigate whether there is a linkage between *TaMATE2* identified in the present investigation and Al resistance in bread wheat by further genotyping the previously studied bread wheat accessions using *TaMATE2* derived marker. In addition, the further research should be directed towards the identification of metabolite that *TaMATE2* transports and how it could be related to Al tolerance in wheat. Finally, the cloning of *TaMATE2* in the present research can provide a very useful reference to the understanding of wheat Al tolerance mechanisms, whose molecular bases are still poorly understood.

MATERIAL AND METHODS

Plant material

Seed of bread wheat (*Triticum aestivum*) genotypes Barbela 7/72/92, Anahuac, and Chinese Spring were collected from the Centro de Genómica e Biotecnologia, Instituto de Biotecnologia e Bioengenharia (CGB/IBB), Universidade de Trás-os-Montes e Alto Douro (UTAD), Vila Real, Portugal. Seeds of each genotype were surface sterilized and grown in hydroponic solution in growth chamber under controlled conditions (14 h/26ºC day and 10 h/22ºC night regime with photosynthetically active radiation (PAR) 150 µmol $m^{-2}s^{-1}$ and relative humidity of 65 %) (Garcia-Oliveira *et al.*, 2013).

Nucleic acid extraction and cDNA synthesis

Seedlings of each genotype were ground in liquid nitrogen and genomic DNA was extracted using DNeasy Plant Mini Kit, (Qiagen, Germany). Total RNA was isolated using the Trizol solution (Invitrogen, USA) followed by purification with PureLinkTM RNA Mini Kit (Ambion, Invitrogen, USA) as described in the manufacturer's instructions. DNA and RNA quality of each sample was visualized by electrophoresis in agarose gel. Prior to synthesize first-strand cDNA, concentration of each sample was quantified with a NanoDrop® ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, DE). The first-strand cDNA was synthesized in a final volume of 20 µl reaction containing: 1 µg RNA, 2 µl 10× RT buffer, 0.8 µl of dNTP mix (10 mM), 2 µl of 10× RT random primers and 1 µl of Multiscribe reverse transcriptase (Applied Biosystems, USA) following the manufacturer's protocols.

Gene Structure and Cloning Full-length cDNA

A 743 bp candidate expressed sequence tag of putative *TaMATE2* (NCBI: CJ862898.1) was retrieved from wheat EST database by using the *ScFRDL2/ScMATE2* (NCBI: BAJ61742.1) (Yokosho *et al.*, 2010) protein sequence from rye on the web page of NCBI (http://www.ncbi.nlm.nih.gov/tblastn). Full-length cDNA of MATE2 orthologue in wheat was amplified from the Al treated seedling cDNA template of bread wheat genotype Barbela 7/72/92 using 5' and 3'RACE (Rapid Amplification of cDNA Ends) (SMARTer™ RACE cDNA Amplification Kit, clontech, USA).

Subsequently, *TaMATE2* specific primers were designed to perform the sequential analysis in bread wheat genotypes Barbela 7/72/92, Anahuac and Chinese Spring and several positive clones were sequenced. The oligonucleotide primer sequences are shown in Appendix-I. Putative protein sequence was predicted in MEGA 4.0 whereas molecular weight and pI was calculated using bioinformatics resource portal ExPASy (http://web.expasy.org/compute_pi/).

Secondary structure prediction and construction of phylogenetic tree

The predicted translated amino-acids sequences of *TaMATE2* (homoeologue from genome D) gene from bread wheat genotypes Barbela 7/72/92 was used to identify the protein transmembrane helices using RHYTHM-web server programme (http://proteinformatics.charite.de/rhythm) (Rose *et al.*, 2009). Further, these data were used to create the two-dimensional visual representation of the transmembrane structure of protein by TMRPres2D tool (http://biophysics.biol.uoa.gr/TMRPres2D) (Spyropoulos *et al.*, 2004).

For phylogenetic relationship of TaMATE2 and other transporters, sixteen MATE2 like protein sequences were assembled from seven different plant species using the Basic Local Alignment Search Tool (BLAST) programme from NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) [Appendix-III]. Phylogenetic relationships among different TaMATE2 proteins were employed using MEGA 4.0 program (Tamura *et al.*, 2007) by Neighbor-Joining clustering method. Bootstraps with 1000 replicates were performed to infer the consensus tree. Branches corresponding to partitions reproduced in less than <50% bootstrap replicates were collapsed.

Chromosome arm assignment of TaMATE2 homoeologues

To obtain genomic sequence of *TaMATE2,* partial genomic fragment amplification strategy was approached as no amplification of the complete genomic fragment was obtained using the flanking primers. For *TaMATE2* localisation, common wheat variety Chinese Spring and aneuploid lines derived from this variety were used. On the basis of PCR amplification, *TaMATE2* homoeologues were assigned to specific chromosomes in Chinese Spring nullisomic-tetrasomic lines and to specific arms of chromosomes using Chinese Spring ditelocentric lines.

TaMATE2 transcript expression assay under Al stress

Four days old seedlings of two diverse bread wheat genotypes Barbela 7/72/92 (Al tolerant) and Anahuac (sensitive) grown in control nutrient solution (without Al) were shifted to fresh solution with Al (74 µM Al in the form of $AlCl_3 \cdot 6H_2O$; stress treatment) and without Al (control treatment). For transcript expression assay, root apices and shoot samples from approximately 20 seedlings in triplicate of each genotype were excised separately after treatment at specific time points (0, 2, 6, 12 and 24 hours) from control and Al stress exposed seedlings and immediately stored in liquid nitrogen.

The relative transcript level of *TaMATE2* genes in both tissues of each genotype (Al resistant and sensitive) was analyzed by Real Time qRT-PCR. PCR reactions were performed in 20 µl reactions containing 10 µl Fast SYBR® Green Master Mix (Applied Biosystems). For normalization of the data, endogenous *18 SRNA* gene was used as reference. Data analyses were performed according to relative standard curve method (Applied Biosystems). Readings were made in duplicate from three different biological repetitions for each time point.

REFERENCES

Barceló, J. and Poschenrieder, C. (2002). Fast root growth responses, root exudates, and internal detoxification as clues to the mechanisms of aluminium toxicity and resistance: A review. *Environ Exp Bot,* 48: 75-92.

Delhaize, E.; Ryan, P.R. and Randall, P.J. (1993) Aluminum tolerance in wheat *(Triticum aestivum* L.) (II. Aluminum-stimulated excretion of malic acid from root apices. *Plant Physiol*, 103: 695-702

Furukawa, J.; Yamaji, N.;Wang, H.; Mitani, N.; Murata, Y.; Sato, K.; Katsuhara, M.; Takeda, K. and Ma, J.F. (2007). An aluminum-activated citrate transporter in barley. *Plant Cell Physiol*, 48: 1081-1091.

Garcia-Oliveira, A.L.; Benito, C.; Prieto, P.; Menezes, R.A.; Rodrigues-Pousada, C.; Guedes-Pinto, H. and Martins-Lopes, P. (2013). Molecular characterization of *TaSTOP1* homoeologues and their response to aluminium and proton (H^+) toxicity in bread wheat (*Triticum aestivum* L.). *BMC Plant Biol*, 13: 134, DOI: 10.1186/10.1186/1471-2229-13-134.

Kidd, P.S.; Llugany, M.; Poschenrieder, C.; Gunsé, B. and Barceló, J. (2001). The role of root exudates in aluminium resistance and silicon-induced amelioration of aluminium toxicity in three varieties of maize (*Zea mays* L.). *J Exp Bot*, 52 (359): 1339-1352.

Kochian, L.V.; Hoekenga, O.A. and Piñeros, M.A. (2004). How do crop plants tolerate acid soils? Mechanisms of aluminum tolerance and phosphorous efficiency. *Annu Rev Plant Biol*, 55: 459-493.

Kuroda, T. and Tsuchiya, T. (2009). Multidrug efflux transporters in the MATE family. *Biochim Biophys Acta*, 1794(5):763-768.

Lai, K.; Duran, C.; Berkman, P.J.; Lorenc, M.T.; Stiller, J.; Manoli, S.; Hayden, M.J.; Forrest, K.L.; Fleury, D.; Baumann, U.; Zander, M.; Mason, A.S.; Batley, J. and Edwards, D. (2012). Single nucleotide polymorphism discovery from wheat next-generation sequence data. *Plant Biotechnol J*, 10:743-749.

Li, L.; He, Z.; Pandey, G.K.; Tsuchiya, T. and Luan, S. (2002). Functional cloning and characterization of a plant efflux carrier for multidrug and heavy metal detoxification. *J Biol Chem*, 277: 5360-5368.

Magalhães, J.V. (2010). How a microbial drug transporter became essential for crop cultivation on acid soils: aluminium tolerance conferred by the multidrug and toxic compound extrusion (MATE) family. *Ann Bot*, 106 (1): 199-203.

Magalhães, J.V.; Liu, J.; Guimarães, C.T.; Lana, U.G.P.; Alves, V.M.C.; Wang, Y.H.; Schaffert, R.E.; Hoekenga, O.A.; Piñeros, M.A. and Shaff, J.E.; Klein P.E.; Carneiro, N.P.; Coelho, C.M.; Trick, H.N. and Kochian, L.V. (2007). A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. *Nat Genet*, 39: 1156-1161.

Maron, L.G.; Piñeros, M.A.; Guimarães, C.T.; Magalhães, J.V.; Pleiman, J.K.; Mao, C.; Shaff, J.; Belicuas, S.N. and Kochian, L.V. (2010). Two functionally distinct members of the MATE (multi-drug and toxic compound extrusion) family of transporters potentially underlie two major aluminum tolerance QTLs in maize. *Plant J*, 61(5):728-40.

Omote, H.; Hiasa, M.; Matsumoto, T.; Otsuka, M. and Moriyama, Y. (2006). The MATE proteins as fundamental transporters of metabolic and xenobiotic organic cations. *Trends Pharmacol Sci*, 27(11): 587-593.

Qi, L.; Echalier, B.; Friebe, B. and Gill, B.S. (2003). Molecular characterization of a set of wheat deletion stocks for use in chromosome bin mapping of ESTs. *Funct Integr Genomics*, 3:39-55.

Raman, H.; Stodart, B.; Ryan, P.R.; Delhaize, E.; Emebiri, L.; Raman, R.; Coombes, N. and Milgate, A. (2010). Genome-wide association analyses of common wheat (*Triticum aestivum* L.) germplasm identifies multiple loci for aluminium resistance. *Genome,* 53(11): 957-966.

Reynolds, M.; Bonnet, D.; Chapman, S.C.; Furbank, R.T.; Manes, Y.; Mather, D.E. and Parry, M.A. (2011). Raising yield potential of wheat. I. Overview of a consortium approach and breeding strategies. *J Exp Bot*, 62 (2): 439-452.

Rose, A.; Lorenzen, S.; Goede, A.; Gruening, B. and Hildebrand, P.W. (2009). RHYTHM -a server to predict the orientation of transmembrane helices in channels and membrane-coils. *Nucleic Acids Res*, 37: W575-80.

Ryan, P.R.; Delhaize, E. and Jones, D.L. (2001). Function and mechanism of organic anion exudation from plant roots. *Annu Rev Plant Physiol Plant Mol Biol*, 52: 527-560.

Ryan, P.R.; Raman, H.; Gupta, S.; Horst, W.J. and Delhaize, E. (2009). A second mechanism for aluminum resistance in wheat relies on the constitutive efflux of citrate from roots. *Plant Physiol*, 149: 340-351.

Ryan, P.R.; Tyerman, S.D.; Sasaki, T.; Furuichi, T.; Yamamoto, Y.; Zhang, W.H. and Delhaize, E. (2011). The identification of aluminium-resistance genes provides opportunities for enhancing crop production on acid soils. *J Exp Bot*, 62(1): 9-20.

Samac, D.A. and Tesfaye, M. (2003). Plant improvement for tolerance to aluminum in acid soils - a review. *Plant Cell Tiss Org Cult*, 75:189-207.

Silva, S.; Pinto-Carnide, O.; Martins-Lopes, P.; Matos, M.; Guedes-Pinto, H. and Santos, C. (2010) Differential aluminium changes on nutrient accumulation and root differentiation in an Al sensitive vs. tolerant wheat. *Environ Exp Bot*, 68(1): 91-98.

Spyropoulos, I.C.; Liakopoulos, T.D.; Bagos, P.G. and Hamodrakas, S.J. (2004). TMRPres2D: high quality visual representation of transmembrane protein models. *Bioinformatics*, 20(17):3258-3260.

Tamura, K.; Dudley, J.; Nei, M. and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol*, 24:1596.

Taylor, G.J. (1991). Current views of the aluminum stress response: the physiological basis of tolerance. *Curr Top Plant Biochem Physiol*, 10: 57-93.

Tolrà, R.; Barceló, J. and Poschenrieder, C. (2009). Constitutive and aluminium-induced patterns of phenolic compounds in two maize varieties differing in aluminium tolerance. *J Inorganic Biochem*, 103(11): 1486-1490

Tolrà, P.R.; Poschenrieder, C.; Luppi, B. and Barcelo, J. (2005). Aluminium-induced changes in the profiles of both organic acids and phenolic substances underlie Al tolerance in *Rumex acetosa* L. *Environ Exp Bot*, 54 (3): 231-238.

Tovkach, A.; Ryan, P.R.; Richardson, A.E.; Lewis, D.C.; Rathjen, T.M.; Ramesh, S.; Tyerman, S.D. and Delhaize, E. (2013). Transposon-mediated alteration of TaMATE1B expression in wheat confers constitutive citrate efflux from root apices. *Plant Physiol*, 161(2):880-92.

von Uexkull, H.R. and Mutert, E. (1995). Global extent, development and economic impact of acid soils. *Plant Soil*, 171: 1-15.

Yang, L.T.; Qi, Y.P.; Jiang, H.X. and Chen, L.S. (2013). Roles of organic acid anion secretion in aluminium tolerance of higher plants. *Biomed. Res. Int.*,173682: doi: 10.1155/2013/173682.

Yazaki, K.; Sugiyama, A.; Morita, M. and Shitan, N. (2008). Secondary transport as an efficient membrane transport mechanism for plant secondary metabolites. *Phytochem Rev*, 7: 513-524.

Yokosho, K.; Yamaji, N.; Ueno, D.; Mitani, N. and Ma, J.F. (2009). OsFRDL1 is a citrate transporter required for efficient translocation of iron in rice. *Plant Cell*, 149(1): 297-305.

Yokosho, K.; Yamaji, N. and Ma, J.F. (2010). Isolation and characterisation of two MATE genes in rye. *Funct. Plant Biol*, 37:296-303.

Yokosho, K.; Yamaji, N. and Ma, J.F. (2011). An Al-inducible MATE gene is involved in external detoxification of Al in rice. *Plant J*, 68(6):1061-1069.

Zhao, J. and Dixon, R. (2009). MATE Transporters facilitate vacuolar uptake of epicatechin 3′-O-glucoside for proanthocyanidin biosynthesis in *Medicago truncatula* and *Arabidopsis*. *Plant Cell*, 21: 2323-2340.

Zhao, J.; Huhman, D.; Shadle, G.; He, X.Z.; Summer, L.W.; Tang, Y.H. and Dixon, R. (2011). MATE2 mediates vacuolar sequestration of flavonoid glycosides and glycoside malonates in *Medicago truncatula*. *Plant Cell*, 23: 1536-1555.

Appendix-I

Table: Detail of primers used in present investigation.

Assay	Primer name	Sequence (5'-3')
TaMATE2_ORF amplification	*TaMATE2_ORF*	ATGCACCTGCTCGGCGTCTTC
		TCAGCTGCGAAGGAAAGC
TaMATE2 chromosomal assignment	*TaMATE2-1A*	GCGTGATCTCTCCCGCCTCTCTA
		ATGATCTCCCTCCCGAGGTC
	TaMATE2-1B	TGGAAATGGTAAATACGTAC
		TGCTTCTTGCCACAT
	TaMATE2-1D	TGTTTATGCCCTTTATATACCT
		CACCTTCCGGCACAA
TaMATE2 expression studies	*18sRNA*	TCCACGAGGAATGCCTAGTAAGC
		ACAAAGGGCAGGACGTAGTC
	TaMATE2_qPCR	GCTGCATTCCAGATTTGCTTGC
		GAGAAGCCCCAAGATCAATCC

Appendix-II

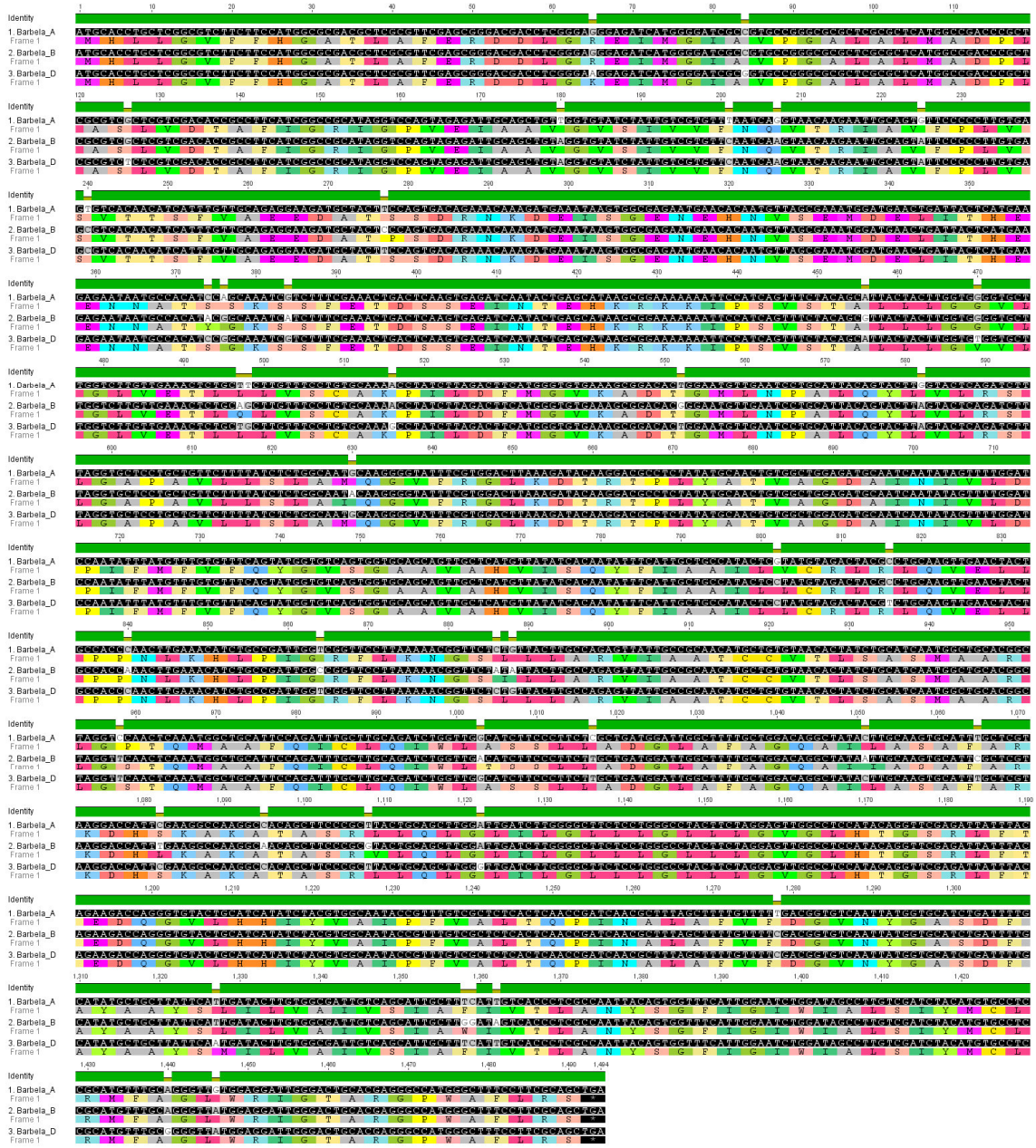


Figure: Multiple alignments of *TaMATE2* homoeologues nucleotide and their deduced amino acid sequences from bread wheat genotype Barbela 7/72/92.

Appendix-III

Detail of MATE2 like proteins in different species used for phylogenetic analysis.

S. No.	Scientific Name	Sequence name and Protein ID	a.a. length
1	*Arabidopsis thaliana*	*A.thaliana*_ NP_974000.1	515
2	*Arabidopsis lyriata*	*A.lyriata*_ XP_002891631.1	514
3	*Capsela rubella*	*C. rubella*_ EOA40177.1	509
4	*Thellungiella halophila*	*T. halophila_FRD3*_ BAJ33987.1	515
5	*Brassica napus*	*B. Napus*_ AFW19998.1	519
6	*Solanum lycorpersicum*	*S. lycopersicum*_ XP_004250944.1	507
7	*Cucumis sativus*	*C sativus*_ XP_004142114.1	521
8	*Fragaria vesca*	*F. vesca*_ XP_004293176.1	492
9	*Prunus persica*	*P. persica*_EMJ15126.1	515
10	*Vitis vinifera*	V.vinifera_ XP_002285445.1	513
11	*Theobroma cacao*	*T. cacao*_ EOY32667.1	516
12	*Glycine max*	*G. max*_ XP_003547162.1	552
13	*Cicer arietinum*	*C arietinum*_XP_004486970.1	506
14	*Medicago truncatula*	*M. truncatula*_ XP_003597421.1	507
15	*Medicago_sativa*	*M. sativa*_ AFL69846.1	507
16	*Eucalyptus camaldulensis*	*E. camaldulensis*_ BAM68467.1	502
17	*Sorghum bicolor*	*S. bicolor*_SbFRDL2_Sb07g001800	525
18	*Zea mays*	*Z. mays*_ ACM47309.1	563
19	*Setaria italica*	*S. italica*_ XP_004982714.1	532
20	*Oryza sativa*	*O. sativa* _Os10g0206800	537
21	*Brachypodium distachyon*	*B. distachyon*_ Bradi3g23150.1	502
22	*Hordeum vulgare*	*H. vulgare*_ AK249059	494
23	*Secale cereal*	*S. cereal*_FRDL2_BAJ61742.1	497
24	*Triticum urartu*	*T.urartu*_MATE_ EMS59851.1	559
25	*Triticum aestivum* *Cv Barbela 7/72/92*	*T. aestivum*_TaMATE2-1A	497
		*T. aestivum*_TaMATE2-1B	497
		*T. aestivum*_TaMATE2-1D	497

>*A. thaliana* _FRDL2_ NP_974000.1

MMSEDGYNTDFPRNPLYIFFSDFRSVLKFDELGLEIARIALPAALALTADPIASLVDTAFIGQIGPVELAAVGVSIALFNQVSRIAIFPLVSIT
TSFVAEEDACSSQQDTVRDHKECIEIGINNPTEETIELIPEKHKDSLSDEFKTSSSIFSISKPPAKKRNIPSASSALIIGGVLGLFQAVFLISA
AKPLLSFMGVKHDSPMMRPSQRYLSLRSLGAPAVLLSLAAQGVFRGFKDTTTPLFATVIGDVTNIILDPIFIFVFRLGVTGAATAHVISQYLMC
GILLWKLMGQVDIFNMSTKHLQFCRFMKNGFLLLMRVIAVTFCVTLSASLAAREGSTSMAAFQVCLQVWLATSLLADGYAVAGQAILASAFAKK
DYKRAAATASRVLQLGLVLGFVLAVILGAGLHFGARVFTKDDKVLHLISIGLPFVAGTQPINALAFVFDGVNFGASDFGYAAASLVMVAIVSIL
CLLFLSSTHGFIGLWFGLTIYMSLRAAVGFWRIGTGTGPWSFLRS

>*C. rubella*_ EOA40177.1

MATTKISQETLFTSSLVISVLKFDELGLEIARIALPAALALTADPIASLVDTAFIGQIGPVELAAVGVSIALFNQVSRIAIFPLVSITTSFVAE
EDACSSQQNTVQDHKESIETGTNNTMEETQELIPENSKDSSSDESKTISSIFSVSKPPAKKRNIPSASSALIIGAFLGLFQAVFLISAAKPLLS
FMGVKHDSPMMRPAQRYLSLRSLGAPAVLLSLAAQGVFRGFKDTTTPLFATVIGDVTNIILDPIFIFVFRLGVTGAATAHVISQYLMCGILLWK
LMGQVDVFNMSTKHLQLCRFMKNGFLLLMRVIAVTFCVTLSASLAAREGSTSMAAFQVCLQVWLATSLLADGFAVAGQAILASAFAKKDYKRAA
ATASRVLQLGLVLGFVLAVILGAGLHFGARVFTKDDKVLHLISIGLPFVAGTQPINALAFVFDGVNFGASDFGYAAASLVMVAIVSIMCLLFLS
STLGFIGLWFGLTIYMSLRAAVGFWRIGTGTGPWSFLRS

>*T. halophila_FRD3*_ BAJ33987.1

MMSEDGYSTDFPRNPLCIFFTDFRSVLKFDELGLEIARIALPAALALTADPIASLVDTAFIGQIGPVELAAVGVSIALFNQVSRIAIFPLVSIT
TSFVAEEDACSSQQNIVQDHKECIETGSNNTNEETQELIPENNKDSTSDESKTSSSIFSVSKTPAKKRNIPSASSALIIGGILGLLQAVLLISA
AKPLLSFMGVKHDSPMLRPAQRYLSLRSLGAPAVLLSLATQGVFRGFKDTTTPLYATVIGDATNIILDPIFIFVFRLGVTGAATAHVISQYLMC
GILLWKLMGQVDIFSLSTKHLQLCRFMKNGFLLLMRVIAVTFCVTLSASLAAREGSISMAAFQVCLQVWLATSLLADGFAVAGQAILASAFAKK
DYKRAAATASRVLQLGLVLGLLLAVILGAGLHFGARLFTKDDKVLHLISIGLPFVAGTQPINALAFVFDGVNFGASDFGYAAASLVMVAIVSIL
CLVFLSSTHGFIGLWFGLTIYMSLRAAVGFWRIGTGTGPWSFLRS

>*B. Napus*_ AFW19998.1

MSEDVGYNKETPCDFPRNPLCIFLSDFKSVLKFDELGLEITRIALPAALALTADPIASLVDTAFIGQIGPVELAAVGVSIALFNQVSRIAIFPL
VSITTSFVAEEDACSSQENTVQDHKECIETGINNTKEETQELIPEINKDSLPDESKISSSIFSVNKSSVKKRNIPSASSALIIGAILGLLQAAF
LISTARPLLSFMGVKHDSPMLGPAQRYLSLRSLGAPAVLLSLATQGVFRGFKDTTTPLYATVIGDATNIILDPIFIFVFRLGVTGAATAHVLSQ
YLMCGILLWKLMGQVDIFNLSTKHLQFSRFMKNGFLLLMRVIAVTFCVTLSASLAAREGSISMAAFQVCLQVWLATSLLADGFAVAGQAILASA
FAKKDYKRAAATASRVLQLGLVLGFLLAIILGAGLHFGARLFTKDDKVLHLISIGLPFVAGTQPINALAFVFDGVNFGASDFGYAAASLVMVAI
VSILCLVLLSSTHGFIGLWFGLTIYMSLRAAVGFWRIGTATGPWSFLRR

>*S. lycopersicum*_ XP_004250944.1

MAEKEVPFCIFFREARSVFKLDELGREIARIALPAALALTADPIASLVDTAFIGQIGPVELAAVGISIAVFNQASRIAIFPLVSVTTSFVAEED
TITKVSSMPRDNEIQDIQSQDVEGLDTESQSNSENKELIPQNRSVYKSETTATSFEVVKPKPEKRHIPSASSALIIGAILGFIQAAFLISGAKP
ILSFMGVKHGSPMLKPAQEYLKLRSLGAPAVLLSLAMQGVFRGFKDTKTPLFATVAGDLTNIILDPIFIFAFHMGVRGAAIAHVISQYLISAIL
FWRLMEKVDLLPPSLKYMQFSRFLTNGFLLLMRVIAVTFCVTLAASLAARLGPTEMAAFQVCLQVWLTVSLLADGLAVAGQAILASAFAQNDYN
RATATASRVLQLGLVLGLVLAVILGVGLHFGARLFTEDVNVIHLIGIGMPFVAATQPINALAFVFDGVNFGASDFAYSAYSMVTVALFSILFLF
ILSSSYKFVGIWVALTIYMSLRALAGFWRIGTRTGPWKFL

>*C. sativus*_ XP_004142114.1

MAFSIMSEEDDPYPSWDKTKTPIRIFFKNARHVFKLDELGREIAQIALPAALALAADPVASLVDTAFIGQIGSVELAAVGVAIALFNQVSRIAI
FPLVSVTTSFVAEEDTIGSVSIEAEDNNDMESGFFTNDEKSSMIPQNGKGEDAHHSRKPLEKKFENSKVENGRRYIPSASSALVIGGVLGLIQA
IFLISGARPLLNFMGVKSDSLMMTPAQQYLTLRSLGAPAVLLSLAIQGVFRGFKDTKTPLYATVAGDATNIILDPIFIFVFRLGVSGAAIAHVI
SQYLIALILFWRLMGQVDLLPPSIKHLQFSRFLKNGFLLLMRVIAVTFCVTLAASLSARQGSTSMAAFQVCLQVWLTTSLLADGLAVAGQAILA
TAFAQNDHDKATAAASRVLQLGLFLGLMLAVFLGVGMTFGARLFTSDVDVLRLIGIGIPFVAATQPINALAFVFDGINFGASDFAYSACSMVLV
AIISIFCLFILSSTQGFIGIWVALTIYMSLRTLAGFGRIGTGTGPWYFLRI

>*F. vesca*_ XP_004293176.1

MGRYVFNLDELGLEIATIALPAALALTADPIASLIDTAFIGQLGSVELAAVGVSIALFNQVSRIAIFPLVSVTTSFVAEEDAIGRVRLESNENI
HLESDSPTNGETKQLLPETDTGKNAYNSKLVNVSFDTVTTGQKKRYIPSASSAMVIGGILGFIQAIFLISGAKPLLSFMGVGSDSPMLKPAQQY
LMLRSLGAPAVLLSLAMQGVFRGFKDTKTPLHATVAGDATNIILDPIFIFVFRFGVNGAAIAHVISQYLMCIILFLRLMAQVDLIPPSAKHLQF
GRFLKNGFLLLMRVIAVTFCVTLAASLSARQGPIAMAAFQVCLQVWLATSLLADGLAVAGQAILASAFAKNDHDKAAVTAARVLQLGLALGLIL
AFLLGVGLQYGARLFTKDVNVLHLISIGIPFVAATQPINALAFVFDGVNFGASDFAYAALSMVLVAIVSIIFLFILSYTSGFIGIWVALTIYMS
LRSFAGFWRIGTGTGPWQFLRT

>*P. persica*_EMJ15126.1

MAEEDDSYSNGNKSKTPIYILFKDCRFVFNLDKLGLEIASIALPAALALTADPIASLVDTAFIGQIGPVELAAVGVSIALFNQASRIAIFPLVS
VTTSFVAEEDTIGTASPEENQNDYLETGSSINGETRQLIPERGTDQNAYNSKPVGASFEIVTTNHQKRYIPSASSAMVIGSILGLIQAIFLISA
AKPLLNFMGVSSDSPMLKPAQQYLILRSLGAPAVLLSLAMQGVFRGFKDTKTPLYATVAGDVTNIILDPIFMFVFRLGVNGAAISHVISQYLIC
VILLWRLMAQVDLLPPSIKHLQFGRFLKNGFLLLMRVIAVTFCVTLAASLAARQGPTPMAAFQVCLQVWLATSLLADGLAVAGQAILASAFAKK
DHDKATATASRVLQLGLVLGLMLAVILGVGLQYGARLFTKDVDVLHLISIGIPFVAATQPINALAFVFDGVNFGASDFAYSAFSMVMVAIVSIF
VLFILSSTNGFIGIWVALTIYMSLRAFAGFWRIGTGTGPWEFLRA

>*V.vinifera*_ XP_002285445.1

MAEEDGLCQSRDGKGIPLWIFLKDARFAFKMDELGLEIAQIAFPAALALTADPIASLVDTAFIGQLGPVELAAVGVSIAVFNQVSRIAIFPLVS
VTTSFVAEEDTIGILDSEPEVSKSVEMGSAVNGETKKLIPKGSGERPYDLEMHGSGHDTPKFESKRHIPSASAALVVGGILGLIQAIFLISGAK
PILNFMGVHSDSPMLAPAQEYLTLRSLGAPAVLLSLAMQGVFRGFKDTKTPLYATVAGDVTNIILDPIFMFVFHMGVGGAAIAHVISQYIISVI
LFWKLMQQVELLPPSTKVLRFGRFLKNGLLLLMRVIAVTFCVTLAASLAARQGPTSMAAFQVCLQVWLATSLLADGLAVAGQAILASAFAKQDY
SKATAAASRVLQLGLVLGLVLSSILGTGMQSAAKLFTKDLSVLHLISIGIPFVAVTQPINSLAFVFDGVNFGASDFAYSAYSMVLVAIVSILCL
FILSSSFGFVGIWVALTIYVTLRAFAGFWRIGTGSGPWSFLRS

>*T. cacao*_ EOY32667.1

MMAEEDDPYLSRVKMRLPIFIFFKDVRNVFKLDDLGSEIAQIALPAALALTADPIASLVDTAFIGQIGPVELAAVGVSIALFNQVSRIAIFPLV
SVTTSFVAEEDTIGRVSSEAQESECLETGSYVNNESKELIPQKESSEGAYQPKTLGGSFDIVKFEPERRHIPSASSALVIGGILGLLQAIFLIS
GAKPLLNFMGVSSDSPMLNPAQQYLTLRSLGAPAVLLSLAMQGVFRGFKDTKTPLYATVAGDVTNIILDPIFMFVFHLGVSGAAIAHVISQYLI

SVILLWKLMSQVDLLPPSLKHLQFSRFLKNGFLLLIRVMAVTFCITLSASMAARQGSTSMAAFQVCLQVWLATSLLADGLAVAGQAILASAFAK
GDHEKATATASRVLQLGLVLGLILAVVLGGGLSFGAKLFTKDVNVLHLIGTGIPFVAATQPINSLAFVFDGVNFGASDFAYSAFSLVLVAIVSI
ICLSILSSSRGFIGLWIALTIYMSLRAFAGFWRIGTGTGPWKFLRV

>*G. max*_ XP_003547162.1

MPNIGAPSVSLQLCITSLNKLPLLVWFLKQHSEFGFSNMAEKESMYSLGDWRRIPICTFFKDARLVFKADSLGREILSIALPAAMALTADPIAS
LVDTAFIGQIGPVELAAVGVSIALFNQVSRIAIFPLVSVTTSFVAEEDTLSGENPHIEEGRCLETGPPKDAETKELLPHKGGNNHNSDFVGECF
NIAKEEHKRRHIPSASSAIFIGGILGLIQAIFLISAAKPLLNFMGVTSDSPMLHPAKQYLKLRTLGAPAVLLSLAMQGVFRGFKDTKTPLYATV
AGDVTNIALDPLFMFVFRLGVSGAAIAHVISQYLISAILLWRLMEQVDLVPPSIKHLQLDRFLKNGFLLLMRVIAVTFCVTLAASLAARQGPTS
MAAFQVCLQVWLAVSLLADGLAVAGQAILAGAFANKDFDRATATASRVLQMGLVLGLALAFILGIGLHFGAKIFTQDANVLHLIQIGIPFVAVT
QPLNSIAFVFDGVNFGASDFAYSAFSMVVVAILSIICLLILSSAGGFIGIWVALTIYMGLRAFAGFLRIGTGSGPWEFLRSS

>*C. arietinum*_XP_004486970.1

MDDNGTSNNALKNKWMLPLSVFFKDARLVFKMDSLAKELLGIAFPSALAVAADPIASLIDTAFIGHLGPVELAAAGVSIALFNQASRITIFPLV
SITTSFVAEEDTINKINIRAAEKQFNDSIKTKSIEIMPGDHLLQDIEAGESKLNTEEKNGTLKNETKNGDDTNSNTSKSSIVTSNRNKSESKPI
RKKMHIASASTALLFGTFLGLIQTATLIFATKPLLGAMGLKYDSPMLIPAVKYLRLRAIGAPAVLLSLAMQGIFRGFKDTTTPLYVIVSGYALN
VAMDPLLIFYFKLGIKGAAISHVLSQYIMAFFLLLILMRKVDLLPPSMKDLQIFRFLKNGGLLLARVIAVTFCVTLSASLAARLGPIPMAAFQT
CLQVWMTSSLLADGLAVAIQAILACSFAEKDYNKVTAAATRVLQMSFVLGVGLSFVVGGGLFFGAGVFSKNVDVIHLIRLGIPFVAATQPINSL
AFVFDGVNYGASDFAYSAYSLVMVSLASVTSLFFLYKSKGFIGIWIALTIYMSLRMIAGIWRMGTGTGPWRFLRGHSMS

>*M. truncatula*_ XP_003597421.1

MAEKESLFSIGDWMRIPICTFFKDARLVFKLDDLGREILSIALPAAMALTADPIASLVDTAFIGQLGPVELAAVGVSIALFNQASRIFIFPLVS
VTTSFVAEEDALSDASSQVEENGCLEAATPPDAETKEFLPQKNSVVESFNVVKDDQHKRRQIPSASSALYFGGVLGLVQATILISAAKPLLNFM
GVTSDSPMLHHAQQYLKLRSLGAPAVLLSLAMQGVFRGFKDTKTPLYATVAGDLTNIALDPLFIFVFRMGVNGAAIAHVISQYLLSAILLWSLN
KQVDLIPPSIKHMQFDRFAKNGFLLFMRVIAVTFCVTLSASLAAHHGSTSMAAFQVCLQVWLAVSLLADGLAVAGQAILAGAFANKDYEKASTT
ATRVLQMGMVLGLALAFILGTGLHFGAKLFTKDIDVLHLIRVGVPFVALTQPLNCLAFVFDGVNFGASDFAYSAFSMVIVAIISIICLLILSSA
GGFIGIWVALTIYMSLRAFAGFLRIGTGSGPWEFLRS

>*M. sativa*_ AFL69846.1

MAEKESLFSIGDWMRIPICTFFEDARLVFKLDDLGREILSIALPAAMALTADPIASLVDTAFIGQLGPVELAAVGVSIALFNQASRIFIFPLVS
VTTSFVAEEDALSDASSQVEENGCLEAATPPDAETKEFLPQKNSVVESFNVVKVDGSKRRQIPSASSALYFGGILGLVQATLLISAAKPLLNFM
GVTSDSPMLHPAMQYLKLRSLGAPAVLLSLAMQGVFRGFKDTKTPLYATVAGDLTNIALDPFIFVFRMGVNGAAIAHVISQYLLSAILLWSLN
KQVDLIPPSIKHLQFDRFAKNGFLLFMRVIAVTFCVTLAASLAAHHGSTSMAAFQVCLQVWLAESLLADGLAVAGQAILARAFANKDYEKASTT
ATRALQMGMVLGLALAFILGTGLHFGAKLFTKDDDVLHLIRVGVPFVALTQPLNCLAFVFDGVNFGASDFAYSAFSMVIVAIISIICLLILSSA
GGFIGIWVALTIYMSLRAFAGFLRIGTGSGPWEFLRS

>*E. camaldulensis*_ BAM68467.1

MPLSMFFKDARNVFKKDELGIEIAQIALPAALALAADPVASLIDTAFIGHIGPTELAAVGVAIAIFNQVSKVAIFPLVSITTSFVAEEDTKERL
HIEAQKDENGDKWFPVSKEKDVEMEELLPQSDSTSKSSFTDTSFGKMADLDNKRRYIPSASSALVIGSILGILQTLFLIFAAKPILNYMGVKSD
SPMLMPAQKYLTLRSLGAPAVLLSLAMQGIFRGFKDTKTPLYATVVGDASNIILDRLFIFDFRMGISGAAIAHVISQYLISLILLWRLMSQVDL
LPPSIKDLKFERFLKNGLLLLVRVIAVTFCVTLAASLAARHGATSMAAFQVCLQIWLATSLLADGLAVAGQAILASAFARADYEKAMSTASRVL
QLGLAMGLVLSVVLVTGLQFASRLFTEDAGVLHLISVGIPFVAVTQPINALAFVFDGINYGASDFAYSAYSMVLVAVVSIACLCALSSTYGFIG
IWIALTIYMTLRTFAGFLRIGAGMGPWGFLNN

>*S. bicolor*_SbFRDL2_Sb07g001800

MMRGESESPLLLHRSADAMERGGEHHHHPLSVFLRDARLAFRWDELGQEIMKIAVPGALALMADPVASLVDTAFIGHIGPVELGAVGVSIAVFN
QVSRIAVFPLVSVTTSFVAEEDAMSNCRDNDKINQENECSVSVSEMEELISPEGASATTSISSFETDSCEVSVEQKRKNIPSVSTALLLGGVLG
LLETVLLVFSAKPILGYMGVTPDSAMMKPALQYLVLRSLGAPAVLLSLATQGVFRGFKDTKTPLYATVAGDAININVLDPIFIFVFQYGVSGAAI
AHVISQYFIASILLWRLRLHVDLLPPSFKHLQFGRFLKNGFLLLARVIAATFCVTLSASMAARQGSTPMAAFQICLQTWLACSLLADGLAFAGQ
AILASAFARKDYPKATATASRILQLALVLGLILSILLGIGLRIGSRLFTSDQGVLHHIYIGIPFVCLTQPINALAFVFDGINYGASDFGYAAYS
MVLVAIVSIICILTLESYSGFIGIWIALVIYMSLRMFAGFWRIGTAQGPWAYLRG

>*Z. mays*_ZmFRDL2_ACM47309.1

MHPPTHMRARAPIGNSGPDSLSYFIPLHPLLGVWTMHGGRERERESLLFLPDPSAGAMEGGGEHHHPLSVFLRDARLAFRWDELGQEIMRIAVP
GALALMADPVASLVDTAFIGHIGPVELGAVGVSIAVFNQVSRIAVFPLVSVTTSFVAEEDAMSNGRDNDKIHQQNERNVSVSEMDELIPPEGAS
ASTSISSFETDSCEVSVEQKRKNIPSVSTALLLGGVLGLLETLLLVLSAKPILGYMGVKPDSAMMKPALQYLVLRSLGAPAVLLSLAIQGVFRG
FKDTKTPLYATVAGDAININVLDPIFMFVFQYGVSGAAIAHVISQYFIASILLWRLRLHVDLLPPSFKHLQFGRFLKNGFLLLARVIAATCCVTL
SASMAARLGSTPMAAFQICLQTWLACSLLADGLAFAGQAILASAFARKDYPKATATASRILQLALVLGLLLSILLGVGLRIGSRLFTSDQGVLH
HIYIGIPFVCLTQPINALAFVFDGINYGASDFGYAAYSMVLVAVVSIICILTLESYGGFIGIWIALVIYMSLRMFAGFWRIGTAQGPWAYLRG

>*S. italica*_ XP_004982714.1

MHTGGACSREESLLDRSADDAAAMEGAGDYEHHHHPLRVFFRDARLAFSWDELGQEIMGIAVPGALALMADPVASLVDTAFIGHIGPVELGAVG
VAIAVFNQVSRIAIFPLVSVTTSFVAEEDAMSNGRDKDKINQENEHAVSNSEMEELISSEEASATTSKSSFETDSCEVNIEQKRKNIPSVSTAL
LLGGVLGLLETLLFVLSAKPILGYMGVTPDSAMMKPALQYLILRSLGAPAVLLSLAMQGVFRGLKDTKTPLYATVAGDAINIVLDPIFMFVFQY
GVRGAAIAHVVSQYFIASILLWRLRLHIDLLPPSLKHLQFGRFLKNGFLLLARVIAATCCVTLSASMAARLGSTPMAAFQICLQTWLACSLLAD
GLAFAGQAILASAFARKDYPKATATASRVLQLALVLGLLLSILLGIGLRIGSGLFTRDQGVLHHIYIGIPFVCLTQPINALAFVFDGINYGASD
FGYAAYSMVLVGIVSIICILTLASYSGFTGIWVALVIYMSLRMFAGFWRIGTAQGPWTFLRS

>*O. sativa*_OsFRDL2_Os10g0206800

MNGESLLDRSSSADAAMNDAVPGHRHHHPLSVFLRDARLAFRWDELGREIMGIAVPGALALMADPVASLVDTAFIGHIGPVELAAVGVSIAVFN
QVSRIAIFPLVSVTTSFVAEEDATSSDREKYEINGENEFNVSDSEMEELVSHEEASAAPSKSSFETDSSDVKIEHKRKNIPSVSTALLLGGVLG
LLQALLLVICAKPLLGYMGVKQGSAMLMPALKYLVVRSLGAPAVLLSLAMQGVFRGLKDTKTPLYATVAGDATNIVLDPIFMFVFQYGVSGAAI
AHVISQYFIASILLWRLRLHVDLLPPSFKHMQFSRFLKNGFLLLARVIAATCCVTLSASMAARLGSVPMAAFQICLQIWLASSLLADGLAFAGQ
AILASAFARQDHSKAAATASRILQLGLVLGLLLSIFLGIGLRLGSRLFTDDQDVLHHIYLGIPFVSLTQPINALAFVFDGINYGASDFGYAAYS
MILVAIVSIIFIVTLASYNGFVGIWIALTVYMSLRMLAGFLRIGTARGPWTFYAAQRMHSHEVVGLC

>*B. Distachium*_BdFRDL2_Bradi3g23150.1

MDGGRHPLSVLFQGASYGLAFKLDNLGKEIMGIAVPGALALMADPLASLVDTAFIGHIGPVELAAVGVSIAVFNQVSRIAIFPLVSVTTSFVAE
EDVTSSDRQKVETNKESEHNVSDSEMDELISSEDTSATSRKSSLSSLVNIEHNRKSIPSVSTSLLLAGVLGLLQTLLLVFYSKPILDFMGLKPD
SGMLNPALQYLVLRSLGAPATLLSLAMQGVFRGLKDTKTPLYATVAGDATNIVLDPIFMFVFKYGVSGAAIAHVISQYFIAAILLWRLRLHVDL
LQPSLKHLQIGRFLKNGFLLLARVIAATCCITLSASMAARLGSTPMAAFQICLQIWLASSLLSDGLAFAAQAILAGAFARKDYSKATVTASRVL
QLGFILGLLLSVLLGVGLRLGSRLFTEDKDVLHHIYVATPFVALTQPINALAFVSDGVSYGASDFAYAAYSMILVAIASVICIVTLTSYSGFLG
IWIALSIYMSLRMFAGFWRIGSARGPWAFLRS

>*H. vulgare*_HvFRDL2_AK249059

MDLPTVFFHGATLAFQRDDLGKEIMGIAVPGALALMADPLASLVDTAFIGHIGPVELAAVGVSIVVFNQVTRIAVFPLVSVTTSFVAEEDATSS
DRHKDEISGDNENNVSEMDELISHEETSGKSSFETGSSEINIEHTRKKIPSVSTALLLGGVLGLFETLLLVSCAKPILDFMGVKADTGMLKPAL
QYLVLRSLGAPAVLLSLAMQGVFRGLKDTRTPLYATVVGDAINIVLDPIFMFVFQYGVSGAAVAHVISQYFIAAILLCRLRLQVELLPPNLKHL
PIGRFLKNGSLLLARVIAATCCVTLSASMAARLGSTQMAAFQICLQIWLASSLLADGLAFAGQAILASAFARKDHSKAKATASRILQLGLILGL
LLSLLLGVGLHTGSRLFTEDKGVLHHIYVATPFVALTQPINALAFVFDGVNYGASDFAYAAYSLILVAIVSIACIVTLASYSGFVGIWIALSIY
MCLRMFAGLWRIGTARGPWVFLRS

>*S. cereal*_FRDL2 (BAJ61742.1)

MHLLGVFFHGATLTFERDDLGREIMGIAVPGALALMADPLASLVDTAFIGHIGPVEIAAVGVSIVVFNQVTRIAVFPLVSVTTSFVAEEDATSS
DRNKVEISGDNEHNVSEMDELITHEENNATSGKSSFETDSSEINTEHRRKKIPSVSTALLLGGVLGLVETLLLVFCAKPILDFMGVKADTGMLK
PALQYLVLRSLGAPAVLLSLAMQGVFRGLKDTRTPLYATVAGDAINIVLDPIFMFVFQYGVSGAAVAHVISQYFIAAILLCRLSLQVELLPPNL
KHLPIGRFLKNGSLLLARVIAATCCVTLSASMAARLGSTQMAAFQICLQIWLASSLLADGLAFAGQAILASAFARKDHSKAKATASRVLQLGLI
LGLLLGLLLGVGLHTGSRLFTEDQGVLHHIYVATPFVALTQPINALAFVFDGVNYGASDFAYAAYSLILVAIVSIACIVTLANYCGFIGIWIAL
SIYMSLRMFAGLWRIGTARGPWAFLRS

>*T. urartu*_MATE_ EMS59851.1

MDLLGVFFHGATLAFERDDLGREIMGIAVPGALALMADPLASLVDTAFIGRIGPVEIAAVGVSIVVFNQVTRIAVFPLVSVTTSFVAEEDATSS
DRNKDEISGENEHNVSEMDELITHEENNATSSKSSFETDSSEINTEHKRKKIPSVSTALLLGGVLGLVETLLLVSCAKPILDFMGVKADTGMLN
PALQYLVLRSLGAPAVLLSLAMQGVFRGLKDTRTPLYATVAGDAINIVLDPIFMFVFQYGVSGAAVAHVISQYFIAAILVCRLRLQVELLPPNL
KHLPIGRFLKNGSLLLARVIAATCCVTLSASMAARLGPTQMAAFQICLQIWLASSLLADGLAFAGQAILASAFARKDHSKAKATASRLLQLGLI
LGLLLGLLLGVGLHTGSRLFTEDQGVLHHIYVAIPILVAIVSIAFIVTLANYSGFIGIWIALSIYMCLRMFAGLWRLLVSVNVGSSKKPENSDG
QYPEAEHPATVVLEPHHDRAREEAAAIDGKVVPVVVELLILRYHRTGWLGMWEHKARCHPDPELREDGQGTLGRDCHRIRHFVQRKVSV

>*T. aestivum*_TaMATE2-1A_cv.Barbela

MHLLGVFFHGATLAFERDDLGREIMGIAVPGALALMADPLASLVDTAFIGRIGPVEIAAVGVSIVVFNQVTRIAVFPLVSVTTSFVAEEDATSS
DRNKDEISGENEHNVSEMDELITHEENNATSSKSSFETDSSEINTEHKRKKIPSVSTALLLGGVLGLVETLLLVSCAKPILDFMGVKADTGMLN
PALQYLVLRSLGAPAVLLSLAMQGVFRGLKDTRTPLYATVAGDAINIVLDPIFMFVFQYGVSGAAVAHVISQYFIAAILVCRLRLQVELLPPNL
KHLPIGRFLKNGSLLLARVIAATCCVTLSASMAARLGPTQMAAFQICLQIWLASSLLADGLAFAGQAILASAFARKDHSKAKATASRLLQLGLI
LGLLLGLLLGVGLHTGSRLFTEDQGVLHHIYVAIPFVALTQPINALAFVFDGVNYGASDFAYAAYSLILVAIVSIAFIVTLANYSGFIGIWIAL
SIYMCLRMFAGLWRIGTARGPWAFLRS

>*T. Aestivum*_TaMATE2-1B cv Barbela

MHLLGVFFHGATLAFERDDLGREIMGIAVPGALALMADPLASLVDTAFIGRIGPVEIAAVGVSIVVFNQVTRIAVFPLVSVTTSFVAEEDATPS
DRNKDEISGENEHNVSEMDELITHEENNATYGKSSFETDSSEINTEHKRKKIPSVSTALLLGGVLGLVETLQLVSCAKPILDFMGVKADTGMLN
PALQYLVLRSLGAPAVLLSLAIQGVFRGLKDTRTPLYATVAGDAINIVLDPIFMFVFQYGVSGAAVAHVISQYFIAAILLCRLRLQVELLPPNL
KHLPIGRFLKNGSILLARVIAATCCVTLSASMAARLGSTQMAAFQICLQIWLTSSLLADGLAFAGQAIIASAFARKDHLKAKATASRVLQLGLI
LGLLLGLLLGVGLHTGSRLFTEDQGVLHHIYVAIPFVALTQPINALAFVFDGVNYGASDFAYAAYSLILVAIVSIAWIVTLANYSGFIGIWIAL
SIYMCLRMFAGLWRIGTARGPWAFLRS

>*T. aestivum*_TaMATE2-1D_cv.Barbela

MHLLGVFFHGATLAFERDDLGKEIMGIAVPGALALMADPLASLVDTAFIGRIGPVEIAAVGVSIVVFNQVTRIAVFPLVSVTTSFVAEEDATSS
DRNKDEISGENEHNVSEMDELITHEENNATSGKSSFETDSSEINTEHKRKKIPSVSTALLLGVVLGLVETLLLVSCAKPILDFMGVKADTGMLN
PALQYLVLRSLGAPAVLLSLAMQGVFRGLKDTRTPLYATVAGDAINIVLDPIFMFVFQYGVSGAAVAHVISQYFIAAILLCRLRLQVELLPPNL
KHLPIGRFLKNGSLLLARVIAATCCVTLSASMAARLGSTQMAAFQICLQIWLASSLLADGLAFAGQAILASAFARKDHSKAKATASRLLQLGLI
LGLLLGLLLGVGLHTGSRLFTEDQGVLHHIYVAIPFVALTQPINALAFVFDGVNYGASDFAYAAYSMILVAIVSIAFIVTLANYSGFIGIWIAL
SIYMCLRMFAGLWRIGTARGPWAFLRS

Molecular characterization of *TaSTOP1* homoeologues and their response to aluminium and proton (H^+) toxicity in bread wheat (*Triticum aestivum* L.)

ABSTRACT

Background:

Aluminium (Al) toxicity is considered to be one of the major constraints affecting crop productivity on acid soils. Being a trait governed by multiple genes, the identification and characterization of novel transcription factors (TFs) regulating the expression of entire response networks is a very promising approach. Therefore, the aim of the present study was to clone, localize, and characterize the *TaSTOP1* gene, which belongs to the zinc finger family (Cys2His2 type) transcription factor, at molecular level in bread wheat.

Results:

TaSTOP1 loci were cloned and localized on the long arm of homoeologous group 3 chromosomes [3AL (*TaSTOP1-A*), 3BL (*TaSTOP1-B*) and 3DL (*TaSTOP1-D*)] in bread wheat. TaSTOP1 showed four potential zinc finger domains and the homoeologue TaSTOP1-A exhibited transactivation activity in yeast. Expression profiling of *TaSTOP1* transcripts identified the predominance of homoeologue *TaSTOP1-A* followed by *TaSTOP1-D* over *TaSTOP1-B* in root and only predominance of *TaSTOP1-A* in shoot tissues of two diverse bread wheat genotypes. Al and proton (H^+) stress appeared to slightly modulate the transcript of *TaSTOP1* homoeologues expression in both genotypes of bread wheat.

Conclusion:

Physical localization of *TaSTOP1* results indicated the presence of a single copy of *TaSTOP1* on homoeologous group 3 chromosomes in bread wheat. The three homoeologues of *TaSTOP1* have similar genomic structures, but showed biased transcript expression and different response to Al and proton (H^+) toxicity. These results indicate that *TaSTOP1* homoeologues may differentially contribute under Al or proton (H^+) toxicity in bread wheat. Moreover, it seems that *TaSTOP1-A* transactivation potential is constitutive and may not depend on the presence/absence of Al at least in yeast. Finally, the localization of *TaSTOP1* on long arm of homoeologous group 3 chromosomes and the previously reported major loci associated with Al resistance at chromosome 3BL, through QTL and genome wide association

mapping studies suggests that *TaSTOP1* could be a potential candidate gene for genomic assisted breeding for Al toxicity in bread wheat.

Key words: Aluminium, TaSTOP1, *Triticum aestivum L.*, *in situ* hybridization, transcription factor, transactivation, homoeologue, pH

INTRODUCTION

Aluminium (Al) toxicity is one of the major concerns for crop productivity in acidic soil, accounting for more than 50 % of the global arable land (von Uexküll, 1995). Al is one of the highly abundant elements in the earth crust and under low pH conditions (acidic soils), it is solubilised in the soil in a toxic ionic form that inhibits root elongation. The primary response of plants to Al toxicity is the rapid inhibition of root growth, particularly root apex by blocking the process of cell division along with cell elongation, and subsequently inefficient absorption of nutrients and water from soil, resulting in the reduction of plant growth and overall productivity (Kochian *et al.*, 2004).

Plant species differ in the level of Al resistance because they evolved different mechanisms to overcome the selective pressure of Al toxicity imposed under acid soils (Simões *et al.*, 2012). These mechanisms can be broadly divided into two categories: Al resistance and Al tolerance. The Al resistance mechanism is based on the Al exclusion from the root apex or the external detoxification of Al in the rhizosphere through the release of organic acid anions (oxalate, citrate and malate) from the roots limiting Al uptake (Ryan *et al.*, 2001). On the other hand, the tolerance mechanism involves the entrance of Al through the roots and its relocation (in the vacuole) or internal detoxification through Al chelating with organic acids (citrate and oxalate) (Ma, 2007). These mechanisms have been validated at the molecular level, particularly with the functional characterization of the major genes such as *ALMT* (Aluminium-Activated Malate Transporter) and *MATE* (Multidrug and Toxic compound Exudation) in bread wheat (*Triticum aestivum* L.) (Sasaki *et al.*, 2004) and sorghum (*Sorghum bicolor* L. Moench) (Magalhães *et al.*, 2004), respectively. In addition, genes encoding transporters have been identified through mutational analysis, especially ABC transporters type (*ALS1* and *ALS3*) in *Arabidopsis* (Larsen *et al.*, 2005, 2007) and bacterial-type ABC transporters (*STAR1* and *STAR2*) in rice (*Oryza sativa* L.) (Huang *et al.*, 2009). Recently, a plasma membrane localized transporter *Nrat1* (Nramp aluminium transporter 1) was also found to be associated with Al tolerance particularly to trivalent form of Al in rice (Xia *et al.*, 2010). These genes are essential for Al resistance, and seem to be a part of the pathway involved in the secondary level of protection through uptake and redistribution of Al to less sensitive tissues in plants, although their actual function in Al tolerance mechanism in plant is still unclear.

There is strong evidence supporting the important role of regulatory genes (transcriptional factors) in plant tolerance to abiotic stresses (Stockinger *et al.*, 1997; Liu *et al.*, 1998; Haake *et al.*, 2000; Yamaguchi-Shinozaki *et al.*, 2006). Recently, the zinc finger transcription factors, *STOP1* (sensitive to proton rhizotoxicity) and *ART1* (Al resistance transcription factor) have been identified in Al sensitive mutants of *Arabidopsis* and rice, respectively (Iuchi *et al.*, 2007; Yamaji *et al.*, 2009). Transcriptome analysis under Al stress revealed that 101 and 31 genes were down-regulated in the Arabidopsis *stop1* (Sawaki *et al.*, 2009) and rice *art1* mutants (Yamaji *et al.*, 2009) respectively. Interestingly, the major genes related to Al toxicity, particularly *ALMT1* and *MATE1* were regulated by these transcription factors (Sawaki *et al.*, 2009; Liu *et al*., 2009). Recently, the contribution of ART1 locus to the variation for Al tolerance in rice has also been identified in QTL analysis (Famoso *et al.*, 2011). The limited impact of single functional genes in plant stress tolerance has been associated with the polygenic nature of such traits. Thus, the identification and characterization of key regulatory genes that act as master regulators controlling entire response networks would be the most promising and sustainable approach to modify complex traits in plants as they coordinate the expression of many target genes (Century *et al.*, 2008).

Wheat is one of the most important natural allopolyploid species, as it is not only directly or indirectly contributing in the food supply for nearly half of the worlds' population but also can serve as a model plant for other economically important polyploid crop species. It is considered as one of the sensitive crop to Al stress among cereals. Bread wheat, with hexaploid nature comprised from three genomes (AABBDD) are organized in seven homoeologous groups, each homoeologous group has individual gene in triplicate form (from each of A, B and D genome). Consequently, it is of great interest to reveal how the expressions of homoeologues genes are regulated in hexaploid wheat because theoretically, all the three homoeologues of a gene are assumed to be uniformly expressed. In the previous decade, several studies have been conducted in order to identify the molecular markers (random DNA markers) linked to Al toxicity through QTL mapping and genome-wide association analyses (Raman *et al.*, 2005, 2010; Zhou *et al.*, 2007; Cai *et al.*, 2008; Navakode *et al.*, 2008) . So far only two candidate genes ALMT1 and MATE1 for Al tolerance in wheat have been identified and also mapped on chromosome 4DL using Chinese Spring deletion lines and 4BL through QTL mapping, respectively (Raman *et al.*, 2005; Ryan *et al.*, 2009; Delhaize *et al.*, 2012). Recently, the three homoeologues of TaMATE1 have been cloned (Tovkach *et al.*, 2013), although, no information is available on the expression of respective

homoeologues of these candidate genes for Al tolerance in hexaploid wheat. Therefore, in order to improve Al tolerance in bread wheat, identification of three homoeologues of the candidate gene is a promising strategy that could be utilized to develop functional markers for genomic assisted breeding programme in wheat. Herein, we report on the identification, physical localization, and molecular characterization of a novel transcription factor *TaSTOP1* homoeologues genes in bread wheat for Al and proton (H^+) toxicity.

RESULTS

Cloning and structure of TaSTOP1

In order to clone the *STOP1* in wheat, primers from wheat EST showing highest similarity with *Arabidopsis thaliana STOP1* were used to amplify the *TaSTOP1* in bread wheat genotype Barbela 7/72/92. Further, 5' and 3' UTR ends of *TaSTOP1* were amplified using RACE as described in material and methods. *TaSTOP1* was amplified in six different bread wheat genotypes and its multiple alignments suggested the mixed amplification from distinct wheat genomes (A, B, and D). Moreover, the comparison of the *TaSTOP1* cDNA sequence with genomic sequence revealed that *TaSTOP1* gene does not contain introns.

The coding region of *TaSTOP1* (*TaSTOP1-A*) from genome A has 1533 bp length and is differentiated from genomes B (*TaSTOP1-B*) and D (*TaSTOP1-D*) by having a 6-bp deletion (1276-1281 positions). *TaSTOP1-B* and *TaSTOP1*-D are differentiated by several SNPs in the respective open reading frame (ORF) [Appendix-IV]. *TaSTOP1-D* in bread wheat genotype Barbela 7/72/92 has a full-length cDNA of 1970 bp, containing a coding region of 1539 bp that encodes a polypeptide of 513 amino acids (Figure 1). *TaSTOP1-D* has a molecular weight of 55.8 kD, with a 5.64 isoelectric point (pI). InterProScan function domain analysis suggests that TaSTOP1 belongs to the Cys2His2 zinc finger family protein. Subcellular prediction analysis indicated that TaSTOP1 protein is localized in the nucleus.

Phylogenetic studies for STOP1 gene among plant species

Deduced amino acid sequence of TaSTOP1 consists of 511 (TaSTOP1-A) and 513 (TaSTOP1-B and TaSTOP1-D) amino acids in bread wheat and the phylogenetic relationship with other STOP-like proteins from 32 different plant species was studied based on their full length sequences. Phylogenetic analysis of STOP like proteins clearly formed two distinct clusters which implied that STOP2 (AT5G22890) was completely outgrouped from STOP1

(AT1G34370) in Arabidopsis. Furthermore, STOP1 and STOP2 protein from monocotyledons can be clearly distinguished from eudicots (Figure 2A). TaSTOP1 homoeologues from bread wheat showed high similarity with STOP1 from Barley (*Hordeum vulgare*) and *Brachypodium distachium* with identities of 92 and 87 %, respectively. Noticeably, the phylogenetic relationship of TaSTOP1 homoeologoues particularly TaSTOP1-A and TaSTOP1-D from bread wheat (AABBDD) displayed the maximum closeness with STOP1 like proteins from *Triticum urartu* (AA) and *Aegilops tauschii* (DD), respectively (Figure 2A).

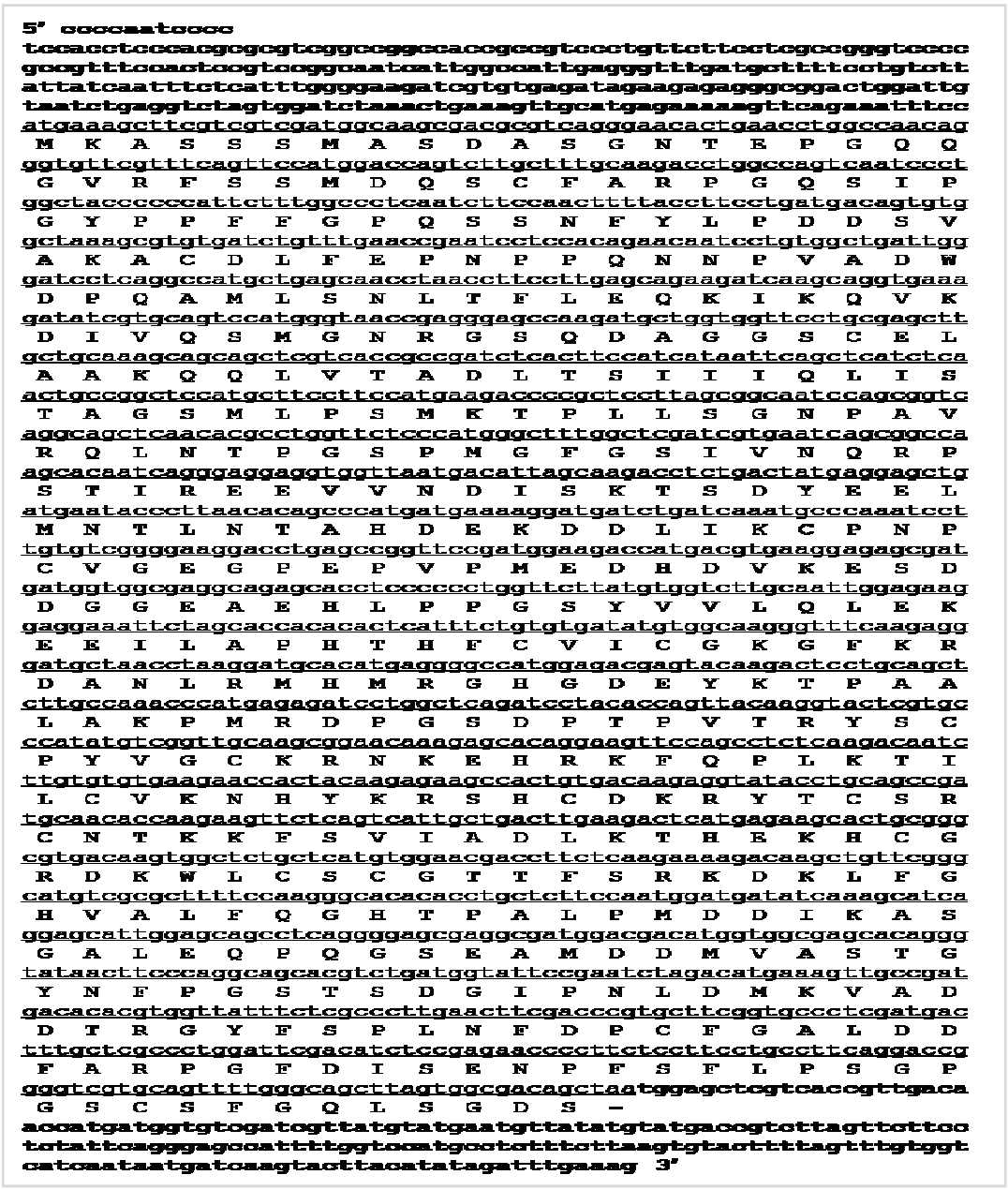


Figure 1. Nucleotide sequence of *TaSTOP1-D* from Barbela 7/7/92 and its deduced protein (amino acid sequences). The nucleotide sequence was presented over the deduced amino acid sequence. Note: translation stop codon was noted by -.

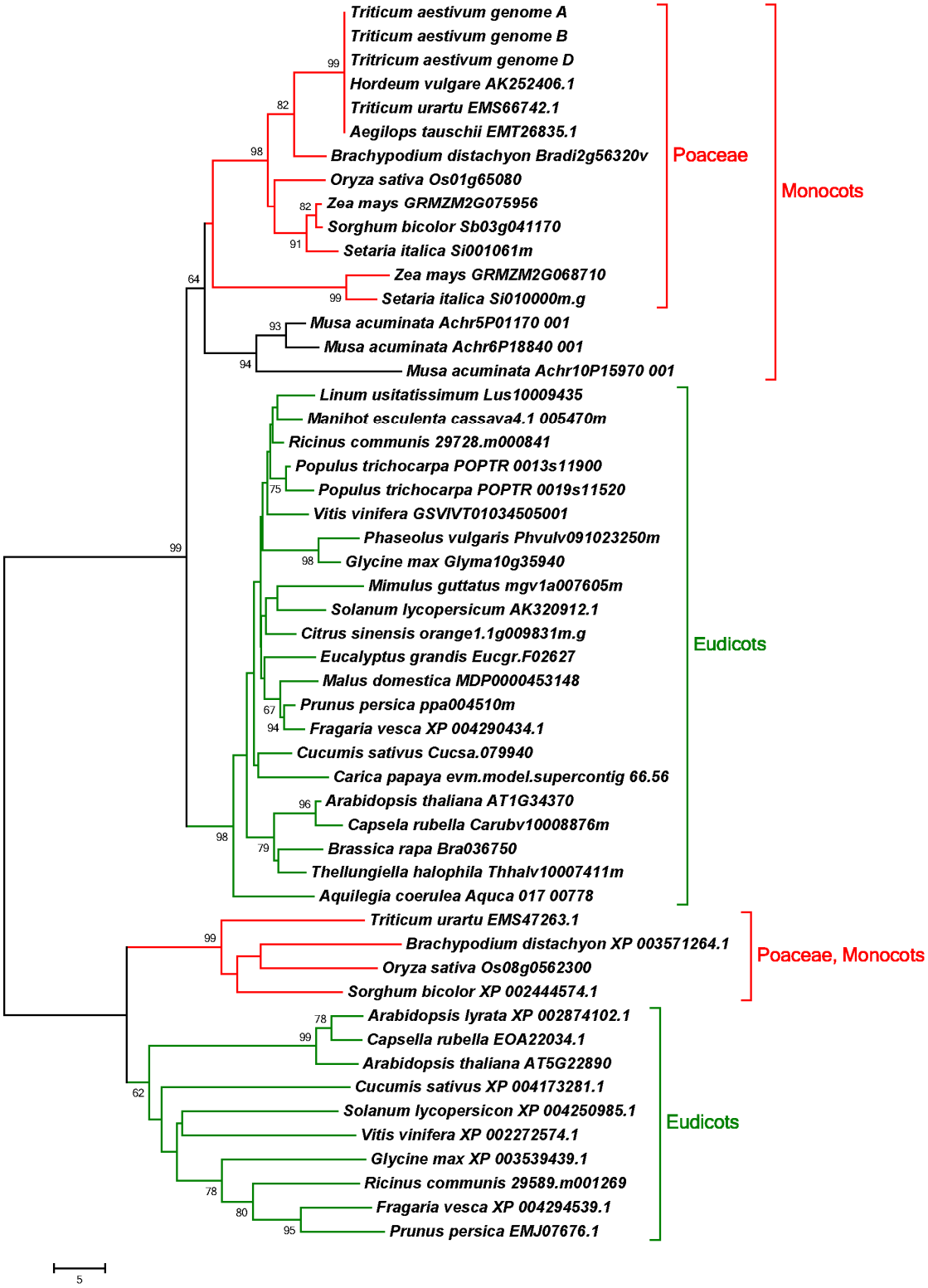


Figure 2A – Phylogenetic tree of STOP-like proteins. Phylogenetic tree based on amino acid sequences showing the relationship of TaSTOP1 with other plant STOP1 type proteins.

Sequence alignment of STOP1 protein from bread wheat with its homologous from cereals [*Brachipodium,* barley, rice, sorghum, maize (*Zea mays*) and foxmillet (*Setaria italic*)] and *Arabidopsis*, shows that *TaSTOP1* encodes a putative Cys2His2 zinc finger protein containing four potential zinc finger domains. Three zinc finger domains (ZF1, ZF2, and ZF4) are predicted as the Cys2His2 type, whereas ZF3 is predicted as the Cys2His-Cys or the Cys2His2 type (Figure 2B). Furthermore, sequence alignment revealed that STOP-like proteins share highly conserved regions in the ZF domains. In addition, the STOP1 protein also showed more conserved regions in C-terminus in comparison to N-terminus. The 6-bp deletion in *TaSTOP1-A* resulted in a loss of two amino acids [proline (P, Pro) and a glutamine (Q, Gln)] at position 426 that did not alter the protein frame. We also observed a SNP in the zinc finger domain of *TaSTOP1-A* and *TaSTOP1-D* that substituted the aspartic acid to aspargine in *TaSTOP1-B* in genotype Barbela 7/72/92 (Figure 2B).

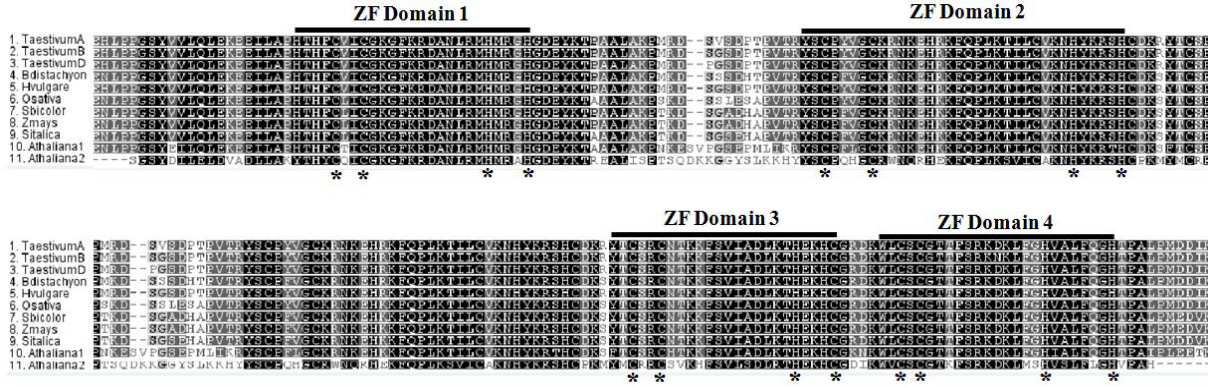


Figure 2B - Multiple alignments of potential Zinc Fingers domain of STOP-like proteins. The deduced amino acid sequences were aligned with MEGA 4.0. Comparison of ZFs domain of TaSTOP1 (TaSTOP1-A, TaSTOP1-B, TaSTOP1-D) with its homologous from Brachypodium (*B.distachyon*), Barley (*H.vulgare*), rice (*O.sativa*), sorghum (*S.bicolor*), maize (*Z.mays*) foxmillet (*S.italica*), and Arabidopsis (*A.thaliana*1 & *A.thaliana*2). Horizontal bars indicate ZFs domain and asterisks indicate conserved motif of Cys2His2 or Cys2His2-Cys.

TaSTOP1 localization on distinct wheat genomes

For localization of *TaSTOP1* on homoeologous chromosomes, locus specific primer pairs were designed [Appendix-I] and *TaSTOP1* homoeologues were amplified in a series of nullitetrasomic lines of Chinese Spring wheat along with Chinese Spring as a positive control. On the basis of the presence or absence of PCR products visualized in agarose gels, we observed that *TaSTOP1* genes are located on homoeologous chromosomes 3A, 3B and 3D, and named *TaSTOP1-A, TaSTOP1-B* and *TaSTOP1-D*, respectively (Figure 3A-C). Furthermore, ditelosomic lines of Chinese Spring for homoeologous group 3 chromosomes were also used to assign the *TaSTOP1* on chromosomal arms and confirmed that *TaSTOP1-A*,

TaSTOP1-B and *TaSTOP1-D* genes are located on the long arm of homoeologous chromosomes 3A, 3B and 3D, respectively (Figure 3D).

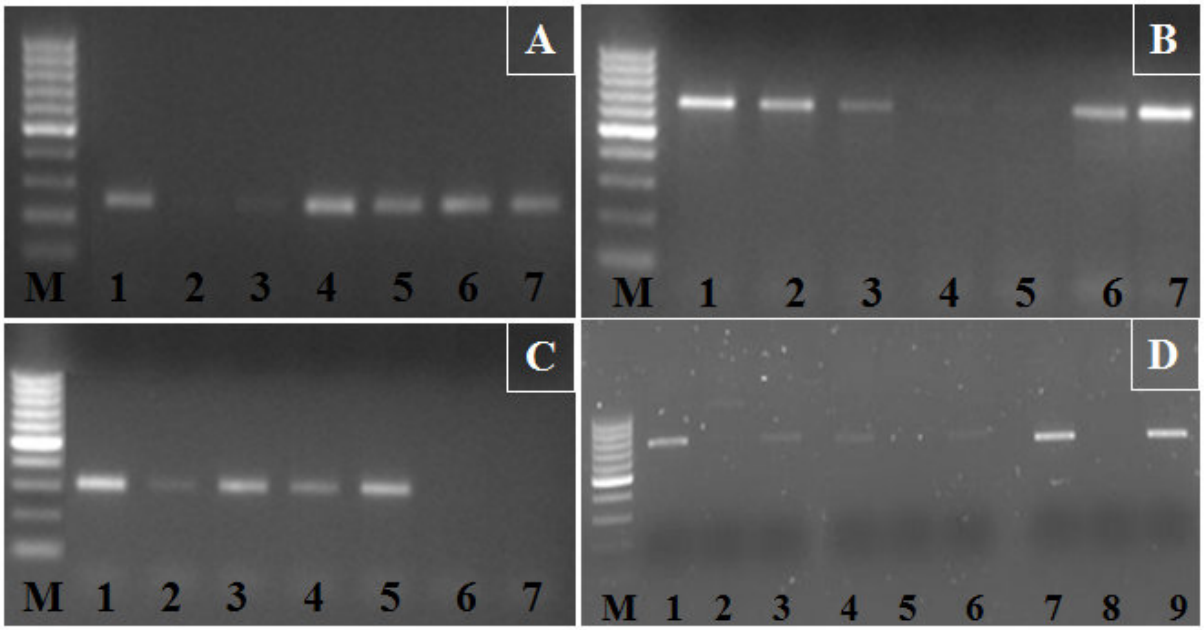


Figure 3. *TaSTOP1* mapping on homoeologous chromosomes and their arms in bread wheat. *TaSTOP1* mapping on wheat chromosomes from genome A (A), B (B) and D (C) using nullitetrasomics lines. M Molecular-weight marker (100bp ladder), 1 to 7: Chinese Spring as control, N3AT3B, N3AT3D, N3BT3A, N3BT3D, N3DT3A, N3DT3B. Arms mapping of *TaSTOP1-A*, *TaSTOP1-B* and *TaSTOP1-D* with ditelosomic lines of homoeologous group 3 (D). 1, 4 and 7: Chinese Spring as a control; 2: Dt3AS; 3: Dt3AL; 5: Dt3BS; 6: Dt3BL; 8: Dt3DS and 9: Dt3DL.

In order to validate these results, physical chromosomal localization of *TaSTOP1* was performed using Tyramide Signal Amplification FISH (Tyr-FISH) technique in the root-tip spreads of bread wheat genotype Barbela 7/72/92. For this purpose, a probe was developed by PCR amplification of *TaSTOP1* from genomic DNA of Barbela 7/72/92. FISH with the *TaSTOP1* specific probe on metaphase chromosomes depicted hybridization signals of *TaSTOP1* on both chromatids of the long arms of three chromosomes (Figure 4A) and simultaneous re-probing of chromosome preparations with GAA- (red) and pAs1 (green) identified homoeologous group 3 with positive signals on chromosomes 3AL, 3BL and 3DL (Figure 4B). Some background was also observed but it was distinguished from true hybridization signals as they usually were dots without any pattern and not in both chromatids of the same chromosome. Only those signals having the same size and appearance on the same position of both chromatids of one specific chromosome were analyzed as true-positive hybridization signals.

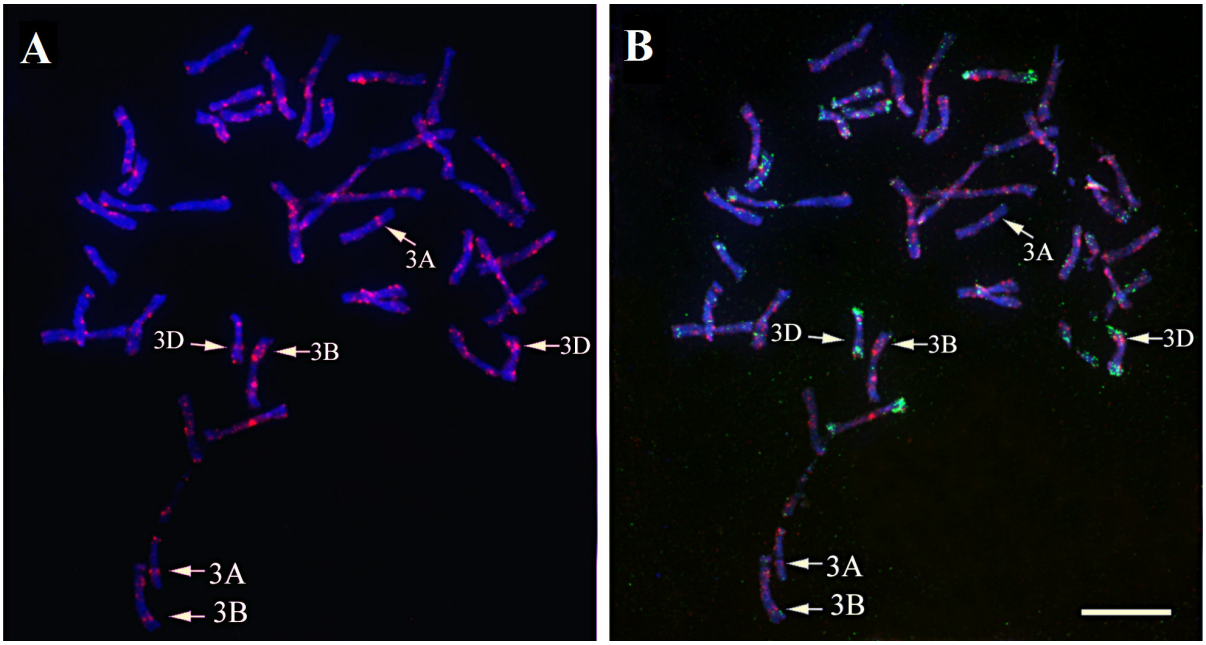


Figure 4. Physical localization of *TaSTOP1* on homoeologous chromosome in the root-tip spreads of bread wheat genotype Barbela 7/72/92 using TYR-FISH technique. (A) Detection of positive hybridization signals of *TaSTOP1* on both chromatids of the long arms of three chromosomes (B) and identification of chromosomes after re-probing with GAA- (red) and pAs1 (green). Note: Arrow indicates the position of *TaSTOP1* on respective chromosome (Scale bar = 10 microm).

TaSTOP1 Transactivation Activity

The nuclear localization of TaSTOP1 protein was predicted by WOLF-PSORT programme. *TaSTOP1* is a putative transcription factor of the Cys_2His_2-type Zinc fingers family. The three homoeologues of *TaSTOP1* showed similar genomic structure in bread wheat [Appendix-V]. Therefore, we only proceed to evaluate the activity of the transactivation potential of *TaSTOP1* located on genome A (*TaSTOP1-A*). To this aim, we performed a modified yeast one-hybrid assay (Menezes *et al.*, 2004) to evaluate the ability of the lexA-TaSTOP1-A fusion protein to activate the lexA-driven expression of the *lacZ* gene in a heterologous yeast system, the results of which are illustrated in Figure 5. It is shown that lexA-TaSTOP1-A has the potential to transactivate *lacZ* expression either in the presence or absence of Al, further corroborating its potential role as a transcription factor.

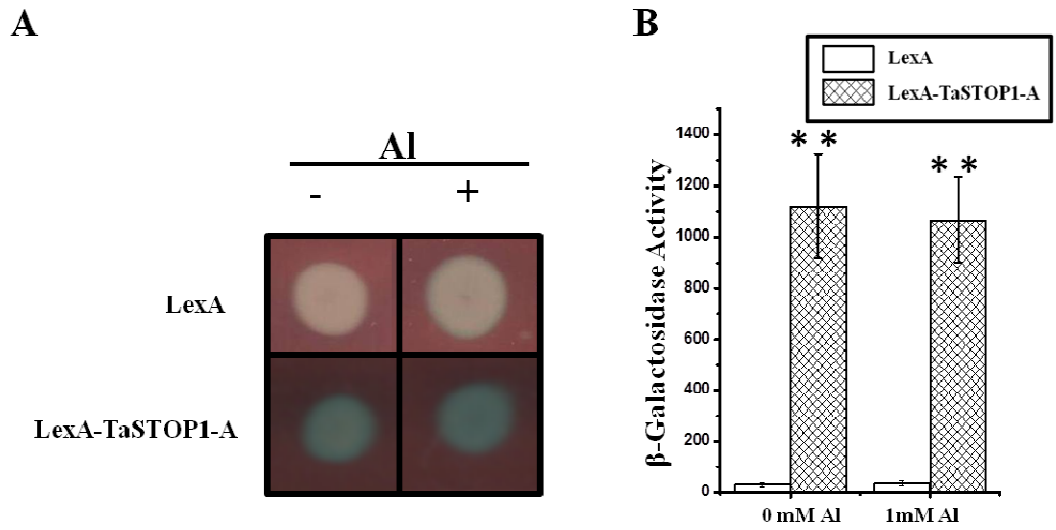


Figure 5. LexA-TaSTOP1-A exhibits transactivation potential in yeast. EGY48 cells co-expressing the lexA-TaSTOP1-A and a report cassette bearing the *lacZ* gene driven by a promoter containing lexA binding sites were grown in selective medium in the presence or absence of Al. Qualitative overlay-assay performed on solid medium (A) and quantitative measurements of β-galactosidase activity normalized to the A_{600} of cells (B). Values are means (± SD) of six independent transformants. Asterisks denote significant differences as identified by paired *t*-tests (: P < 0.001).**

Expression profile of three *TaSTOP1* homoeologues in bread wheat

To understand the homoeologue specific expression of *TaSTOP1* in bread wheat, we measured the temporal expression of *TaSTOP1-A, TaSTOP1-B* and *TaSTOP1-D* by Real-time qPCR in the root and shoot tissues of two bread wheat genotypes [Barbela 7/72/92 (Al resistant) and Anahuac (Al sensitive)] grown under Al stress (74 µM Al). The*18S RNA* expression was used as an internal control. We noticed a biased transcription of homoeologues of *TaSTOP1* gene in the root and shoot tissue of these diverse bread wheat genotypes (Figure 6). Under control as well as Al stress conditions, the transcript levels of homoeologue *TaSTOP1-A* were significantly higher than those of *TaSTOP1-D* and *TaSTOP1-B* in root and shoot tissues of both genotypes (depending upon genotypes, fold differences in transcript abundance were about 5-6 and 6-11 for homoeologue *TaSTOP1-A* versus *TaSTOP1-B* in root and shoot, respectively) whereas expression of homoeologue *TaSTOP1-D* was only two-fold higher than *TaSTOP1-B* in the root tissues of both genotypes (Figure 6A & B). In the root tissues, a slight but rapid induction (within the two hours to Al exposure) followed by return to basal level in the transcript expression of homoeologue *TaSTOP1-A* gene was observed in genotype Barbela 7/72/92 (Al resistant) whereas quite stable transcript expression was noticed in genotype Anahuac (Al sensitive) under Al stress (Figure 6A). The expression of *TaSTOP1-B* was also almost unaltered in the roots of both

genotypes under Al stress. Similarly, *TaSTOP1-D* was also constitutively expressed under Al stress in the roots of both genotypes except a rapid (2 h) and significant repression in the Al sensitive genotype Anahuac (Figure 6A).

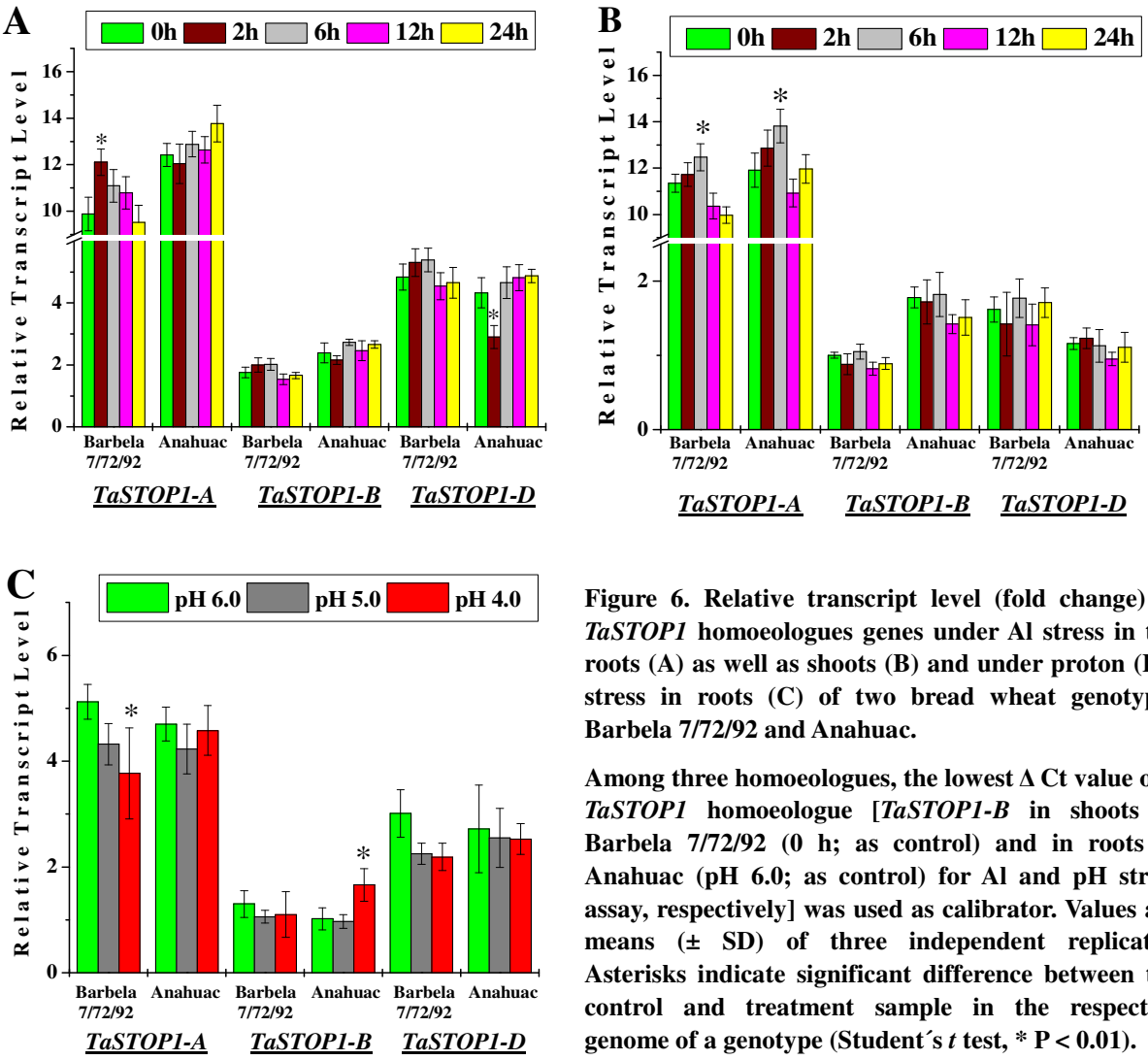


Figure 6. Relative transcript level (fold change) of *TaSTOP1* homoeologues genes under Al stress in the roots (A) as well as shoots (B) and under proton (H^+) stress in roots (C) of two bread wheat genotypes Barbela 7/72/92 and Anahuac.

Among three homoeologues, the lowest Δ Ct value of a *TaSTOP1* homoeologue [*TaSTOP1-B* in shoots of Barbela 7/72/92 (0 h; as control) and in roots of Anahuac (pH 6.0; as control) for Al and pH stress assay, respectively] was used as calibrator. Values are means (± SD) of three independent replicates. Asterisks indicate significant difference between the control and treatment sample in the respective genome of a genotype (Student´s *t* test, * P < 0.01).

Nevertheless, a quite stable transcript expression of *TaSTOP1* homoeologues was observed in the shoot of both genotypes except homoeologue *TaSTOP1-A* (Figure 6B). Interestingly, and similarly to roots, the expression of *TaSTOP1-A* was also slightly induced under Al stress in shoots of both genotypes, however, it's up-regulation was noticed only after 6 h and return to its basal level (Figure 6B). Noticeably, the transcript level of homoeologues *TaSTOP1-A* and *TaSTOP1-B* seems to be relatively higher in the root and shoot tissues of Al sensitive genotype Anahuac, respectively than the tolerant genotype Barbela 7/72/92 (Figure 6 A and B).

Moreover, to evaluate the transcript expression of *TaSTOP1* homoeologues under protons (H^+) toxicity, we also performed the relative quantification of homoeologue specific expression of *TaSTOP1* transcript in root tissues of both genotypes under different levels of pH (6.0, 5.0 and 4.0). A slight gradual repression in the transcript level of *TaSTOP1-A* homoeologue was observed in the roots of Barbela 7/72/92 under protons (H^+) toxicity, but a quite stable expression was observed in genotype Anahuac (Figure 6C). Furthermore, considerable up-regulation in homoeologue *TaSTOP1-B* and slight repression in homoeologue *TaSTOP1-D* transcript level was also observed under low pH (pH 4.0) in the roots of genotype Anahuac and Barbela 7/72/92, respectively (Figure 6C).

DISCUSSION

So far only *ALMT1* and *MATE* genes have been described as responsible for most of the genotypic variation for Al toxicity and are considered as major genes for Al resistance in bread wheat (Ryan et al., 2011). These genes have been identified in the model plant Arabidopsis (Liu *et al.*, 2009; Hoekenga *et al.*, 2006) and major cereals such as barley (Furukawa *et al.*, 2007), maize (Maron *et al.*, 2010), rye (Collins *et al.*, 2008; Yokosho *et al.*, 2010), sorghum (Magalhães *et al.*, 2007) and wheat (Sasaki *et al.*, 2004; Ryan *et al.*, 2009). The mechanisms underlying Al-tolerance in plant species have yet to be fully elucidated, as it seems that Al targets multiple cellular sites such as cell walls, plasma membranes, and cellular processes, like signal transduction pathways and homeostasis mechanisms (Kochian *et al.*, 2004; Ma, 2007; Barceló and Poschenrieder, 2002; Poschenrieder *et al.*, 2008). Recent evidence suggests that regulatory genes (transcription factors) play a key role in Al detoxification at different cellular levels (Iuchi *et al.*, 2007; Yamaji *et al.*, 2009).

In the present study, a novel transcription factor gene named *TaSTOP1* was cloned from bread wheat genotype Barbela 7/72/92 (Figure 1). TaSTOP1 belongs to a member of Cys2His2 zinc finger family protein that contains four potential zinc finger domains and has highly conserved regions in the zinc finger (ZF) domains (Figure 2B) like other C2H2-type zinc finger proteins contain more than one zinc finger motif with highly conserved amino acid sequence for DNA binding (Wolfe *et al.*, 2000; Iuchi *et al.*, 2007). Phylogenetic analysis clearly differentiated the STOP1 like proteins from monocots in a group (Figure 2A) and also suggested that genome A and B are more distant from genome D [Appendix-V], as genome D was incorporated recently in comparison to the A and B genomes in bread wheat (Huang *et al.*, 2002). In addition, our phylogenetic analysis results also support previous findings at

molecular level that the A and D genomes in bread wheat were derived from *T. urartu* and *A. tauschii*, respectively [Appendix-V] (Huang *et al.*, 2002).

In past, the nullitetrasomic and ditelosomic lines of Chinese Spring wheat have been extensively used in classical and molecular genetic studies for the identification of loci associated with numerous traits including Al tolerance (Aniol, 1984; Cai *et al.*, 2008)). Thus, gene mapping has practical implications in plant genetics and breeding, and the determination of its physical localization is even more precise as it enables us to confirm the exact position of a gene on a chromosome. PCR based mapping of *TaSTOP1* using nullitetrasomic and ditellosomic lines of Chinese Spring wheat revealed that *TaSTOP1* is located on the long arm of wheat homoeologous group 3 chromosomes (Figure 3).

In situ hybridization is not only the most direct method for physical localization of genes in chromosomes, but also it is a valuable tool for the identification of copy numbers of a gene in the species having a highly complex genome such as is the case of bread wheat. Due to its hexaploid nature, most genes can be found in triplicate with one copy on each genome (Clarke *et al.*, 2003; Danyluk *et al.*, 2003; Watanabe and Koval, 2003). Tyramide signal amplification FISH (Tyr-FISH) technique has been successfully used for the identification of low copy number DNA sequences in wheat (Perez *et al.*, 2009). However, for the successful localization of a gene using the FISH technique, the minimum length of the probe size seems to be crucial. Recently, *RD50* gene has been localized in bread wheat with a probe size of 2 kb (Perez *et al.*, 2009). Physical localization of *TaSTOP1* by Tyr-FISH technique not only confirmed the results obtained from PCR based mapping, but also indicated that *TaSTOP1* could be a single copy gene localized on the long arm of homoeologous group 3 chromosomes (Figure 4). In bread wheat, physical localization of single copy gene *Glu-1* and *RD50* with FISH technique has been demonstrated on long and short arms of the homoeologous group 1 chromosomes, respectively (Perez *et al.*, 2009; Cabrera *et al.*, 2002). In the present work, we were able for the first time to localize a gene with a probe size smaller than 2.0 kb demonstrating that the FISH technique can be used to simultaneously anchor homoeologous chromosomes with 1.5 kb probes even in bread wheat.

Subcellular prediction of TaSTOP1 protein in the nucleus is in agreement with what was previously reported (Sawaki *et al.*, 2009). Zinc finger motifs are thought to recognize and bind to target DNA sequence, but they are not required for transcriptional activity (Sakamoto *et al.*, 2004; Lin *et al.*, 2007). Our results clearly exhibited the transcriptional activation potential of *TaSTOP1-A* at least in yeast (Figure 5). In present investigation, we observed

highly similar genomic structure of *TaSTOP1* genes [Appendix- IV & V], therefore, it seems that *TaSTOP1* transactivation function is constitutive and may not depend on the presence/absence of Al. The SNPs observed among *TaSTOP1* homoeologues showed minor changes in respective amino acids of putative proteins which could alter the secondary structures by influencing the foldings of these proteins [Appendix-IV]. Therefore, the three homoeologues of the same locus in bread wheat share high sequence similarity could illustrates the flexibility of a polyploidy species in which due to the mutation in one homoeologue may be compensated for by the homoeologue (Tovkach *et al.*, 2013).

Allopolyploidy plays an important role in plant evolution that arises with the merging of two or more genomes into single nucleus which may contribute either equally or disproportionately. However, recent molecular findings confirmed the asymmetric genomic expression pattern in natural and synthetic allopolyploid plant species, but the predominant transcript expression of one genome over the other genome(s) vary from gene to gene (Mochida *et al.*, 2004). In hexaploid wheat, a uniform level of expression for all the three homoeologues has been reported for approximately 20% of the unigene loci (Mochida *et al.*, 2004); whereas 20-29% genes did not express at least one homoeoallele (Mochida *et al.*, 2004; Bottley *et al.*, 2006). Keeping in view, the relative contribution of the three homoeologues of *TaSTOP1* gene at transcript expression level was determined under Al and proton (H^+) stress in two bread wheat genotypes showing contrasting behaviour for Al toxicity. Transcript expression profiling of *TaSTOP1* homoeologues in root and shoot tissues identified the predominance of *TaSTOP1-A* homoeologue followed by *TaSTOP1-D* over *TaSTOP1-B* in root and only predominance of *TaSTOP1-A* in shoot tissues of both genotypes under control and stress (Al and pH) conditions (Figure 6). Similarly, among the three homoeologues, the higher expression of one homoeologue of transcription factor gene MAD box and Spa gene has been also observed in bread wheat (Shitsukawa *et al.*, 2007; Ravel *et al.*, 2009). Although, the presence of *cis* elements within the 5' untranslated region of a gene is unusual, it is not aberrant or abnormal (Daraselia *et al.*, 1996). Surprisingly, sequence analysis of 5' UTR of *TaSTOP1* genes differentiated the *TaSTOP1-A* homoeologue from *TaSTOP1-B* and *TaSTOP1-D* due to the presence of a pyrimidine-rich stretch and absence of light responsive element [Appendix-VI]. In tomato, the deletion of the 5' UTR containing pyrimidine-rich stretch from the HMG2 promoter reduced the level of HMG2 gene expression by a factor of 10 (Daraselia *et al.*, 1996). Therefore, among the three homogeologues in bread wheat, the up- or down-regulation of the expression of specific

homoeologue could be the result from either the dominancy of ancestral diploid donor parent or early polyploidization-cis-regulatory variation.

Interestingly, Al responsive transcript expression of *TaSTOP1* homoeologues observed in the root tissues of two contrasting bread wheat genotypes for Al toxicity at different time point in present investigation indicated the role of *TaSTOP1* in Al resistance in bread wheat (Figure 6A & B). Al responsive *STOP1* expression has also been reported in *Arabidopsis* (Iuchi *et al.*, 2007), alfalfa (Chen *et al.*, 2011) and common bean (Eticha *et al.*, 2010). Similar to Arabidopsis, the transcript expression of homoeologues of *TaSTOP1* in the roots of diverse bread wheat genotypes also slightly modulated under proton (H^+) toxicity (Iuchi *et al.*, 2007). It is noticeable that genotype Barbela 7/72/92 is highly resistant to Al toxicity compared with Anahuac, but in the absence of Al under low pH we did not observe significant differences between these genotypes for root growth [Appendix-VII]. Interestingly, the three homoeologues of *TaSTOP1* have similar genomic structures, but showed biased transcript expression and different response to Al and proton (H^+) toxicity. These results indicate that the homoeologues of TaSTOP1 may differentially contribute to Al or proton (H^+) tolerance in bread wheat.

Conclusions

Bread wheat has wide genotypic variation for Al resistance (Raman *et al.*, 2005; Martins-Lopes *et al.*, 2009) and the role of several chromosomes such as chromosome arms 2DL, 3DL, 4BL, 4DL, 6AL, 7AS and chromosome 7D, in Al tolerance has been revealed in classical genetic studies through chromosome manipulation in wheat (Anil, 1984). Furthermore, several QTL associated with Al toxicity in bread wheat have also been reported by many researchers (Raman *et al.*, 2005; Zhou *et al.*, 2007; Cai *et al.*, 2008; Navakode *et al.*, 2009). Contrarily to Arabidopsis and rice, due to the paucity of *a priori* candidate genes in wheat only two major QTL located on chromosome 4DL and 4BL have yet been elucidated at molecular level, showing the co-segregation with candidate genes *TaALMT1* homoeologue (4DL) and *TaMATE1* homoeologue (4BL), respectively (Raman *et al.*, 2005; Ryan *et al.*, 2009). The classical studies using chromosomal manipulation as well as recent QTL mapping and genome-wide association analysis have also detected loci associated with Al resistance on homoeologous group 3 chromosomes (3A, 3B and 3D) in bread wheat (Aniol, 1984; Raman *et al.*, 2010; Zhou *et al.*, 2007; Cai *et al.*, 2008). Recently, the role of a zinc finger transcription factor ART1 identified through mutational analysis in rice has also been shown

in natural variation of Al tolerance in rice, earlier which was suggested that it was not involved in Al tolerance (Famoso *et al.*, 2011; Ryan *et al.*, 2011). In the present investigation, we cloned and characterize the novel candidate genes *TaSTOP1* in bread wheat and the homoeologues of *TaSTOP1* have exhibited similar genomic structures, but showed biased transcript expression and different response to Al and proton (H^+) toxicity. Furthermore, *TaSTOP1-B* homoeologue from Al tolerant genotype Barbela 7/72/92 also showed highest similarity with respective homoeologue from an Al tolerant genotype Viloso Mole (Appendix-V) (Martins-Lopes *et al.*, 2009). Therefore, it would be very interesting either to functionally characterize or further verify the role of TaSTOP1 because gene(s) underlying the QTL on homoeologous group 3 chromosomes particularly 3BL has not been so far identified in this important cereal. Moreover, the transactivation potential of *TaSTOP1* can provide a very useful reference to understand comprehensively wheat Al resistance mechanisms, whose molecular bases are still poorly understood.

MATERIAL AND METHODS

Plant material and growth condition

Seedlings of bread wheat genotypes Barbela 7/72/92, Anahuac, Chinese Spring, Ruivo, Viloso Mole and Saloio were grown in hydroponic solution (0.4 mM $CaCl_2$, 0.65 mM KNO_3, 0.25 mM $MgCl_2.6H_2O$, 0.1 mM $(NH_4)_2SO_4$ and 0.04 mM NH_4NO_3) and kept in controlled growth chamber under 14 h/26ºC day and a 10 h/22ºC night regime, with a light intensity of 150 µmol photons $m^{-2}s^{-1}$ and a relative humidity of 65 %.

Al stress assay

Four days old seedlings of Al resistant bread wheat genotype Barbela 7/72/92 and Al susceptible genotype Anahuac growing in hydroponic solution with pH 4.0 were shifted in fresh nutritive solution having 74 µM Al whereas control seedling were raised in only fresh hydroponic solution. For root re-growth measurement, after 24 h, roots were immersed in 0.1 % eriochrome cyanine R (Sigma, Germany) dye solution for 10 minutes and prior to measurement; seedlings were allowed to grow in fresh nutritive solution for 48 h. For visual detection of Al accumulation in roots, hematoxylin assay was performed (Polle, 1978) and root samples were observed under microscope.

DNA and RNA extraction

Total genomic DNA was isolated from young leaf samples using the DNeasy Plant Mini Kit (Qiagen, Germany) and total RNA was extracted using Trizol method followed by purification using PureLink™ RNA Mini Kit (Ambion, Invitrogen, USA). The first-strand cDNA was synthesized in a final volume of 20 µl reaction containing: 1 µg RNA, 2 µl 10× RT buffer, 0.8 µl of 25× dNTP mix (100 mM), 2 µl of 10× RT random primers and 1 µl of Multiscribe reverse transcriptase (Applied Biosystems, USA).

Gene structure and cloning of full-length cDNA

The *Arabidopsis thaliana* AtSTOP1 (AT1G34370) protein sequence was retrieved from TAIR database (http://www.arabidopsis.org) and was used as the query sequence to search the GenBank wheat expressed sequence tag (EST) database. The homologous sequence of *STOP1* in wheat was obtained via tblastn in NCBI (http://www.ncbi.nlm.nih.gov). Homologous wheat EST clones showing highest similarity with *Arabidopsis STOP1* were retrieved. *TaSTOP1* (wheat EST) specific primers were designed to perform the sequential analysis in bread wheat genotype Barbela 7/72/92 [Appendix-I]. Total RNA from the seedlings of Barbela 7/72/92 and Anahuac was employed to obtain the full-length cDNA of *TaSTOP1* including 5' UTR and 3' UTR using Rapid Amplification of cDNA Ends (RACE) (SMARTer™ RACE cDNA Amplification Kit, clontech, USA). The genomic sequence of *TaSTOP1* was obtained in six different bread wheat genotypes, namely Barbela 7/72/92, Anahuac, Ruivo, Viloso Mole, Saloio and Chinese Spring [Appendix-II].

Molecular analysis and construction of phylogenetic tree

The Pfam (www.sanger.ac.uk/Software/Pfam/search,shtml) software was used to identify potential domains and WOLF-PSORT (http://wolfpsort.org) was used to predict the intracellular localization of the TaSTOP1 protein. Fifty two STOP1 like protein sequences were collected from 32 different plant species including TaSTOP1 homoeologues from bread wheat genotype Barbela 7/72/92 using the Basic Local Alignment Search Tool (BLAST) programme from NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) [Appendix-III]. Phylogenetic tree was constructed by MEGA4 software using neighbor-joining method (Tamura *et al.*, 2007).

Mapping of *TaSTOP1* gene in wheat genome

All the nullitetrasomic and ditelosomic lines of Chinese Spring wheat except for the homoeologous group 2, 4 and 6 chromosomes used in present investigation were available in our lab, whereas remaining stocks were kindly provide by Prof. Adam J. Lukaszewski at Department of Botany and Plant Sciences, University of California, USA. On the basis of *TaSTOP1* alignment, different pairs of specific primers for each *TaSTOP1* homoeologue were designed [Appendix-I]. *TaSTOP1* homoeologues (*TaSTOP1-A*, *TaSTOP1-B* and *TaSTOP1-D*) were amplified in the nullitetrasomic (N3AT3B, N3AT3D, N3BT3A, N3BT3D, N3DT3A and N3DT3B) and ditelosomic [Dt 3AL and 3AS (long and short arm), Dt 3BL and 3BS, Dt 3DL and 3DS] lines of Chinese Spring wheat. *TaSTOP1* homoeologues were assigned to the chromosomal arms based on the presence/absence of the PCR amplification products.

Physical localization of *TaSTOP1* was performed by *in situ* hybridization. Probe was prepared by PCR amplification of the 1.5 kb genomic region of *TaSTOP1* from Barbela 7/72/92. Chromosome spreads from the root tips of Barbela 7/72/92 seedlings at mitotic metaphase, probe labelling and hybridization mixture were carried as described by Prieto *et al.* [60]. Probe labelling was confirmed by dot-blot and detection of hybridization signals was carried out using the Tyramide Signal Amplification Kit (TSA^{TM}, PerkinElmer Life and Analytical Sciences, Inc., Waltham, MA, USA). To identify wheat chromosomes with positive signals, samples were re-hybridized using simultaneously the pAS1 repetitive sequence and GAA-satellite sequence as probes (Prieto *et al.*, 2001; Pedersen *et al.*, 1996). The GAA-satellite sequence identifies all the A and B wheat chromosomes (Pedersen et al., 1996) whereas the pAs1 identifies chromosomes from the D genome (Rayburn *et al.*, 1986). Individual slides were observed under a Nikon Eclipse 80i microscope (Nikon Instruments Europe BV, UK). Images were captured with a Nikon CCD camera using the appropriate Nikon 3.0 software and processed with Photoshop 4.0 software (Adobe Systems Inc., San Jose, California, USA).

Transactivation assay

TaSTOP1 was amplified by PCR from Barbela 7/72/92 using oligonucleotide primers containing restriction site for *Kpn* I [Appendix-I]. The *Kpn* I digested *TaSTOP1* fragment was cloned in frame with the *lexA* gene in the plasmid YCp91 (Menezes *et al.*, 2004). Transformation of the ligated products was performed in the *Escherichia coli* strain *XL1-Blue recA1 endA1 gyrA96 thi-hsdR17 supE44 relA1 lac* [*F'proAB lacI^{q}ZDM15 Tn10* (*Tet^{r}*)]

(Stratagene, USA) and the positive clones were sequenced. The recombinant plasmid YCp91-*TaSTOP1-A* was transformed in the yeast strain EGY48, which harbours the plasmid pSH18.84 encoding the *lacZ* gene under the control of a promoter containing LexA – responsive *cis* elements. Transformants were grown in selective media and β-galactosidase measurements were performed as previously described (Miller *et al.*, 1972; Azevedo *et al.*, 2007).

Analysis of *TaSTOP1* expression level

For Al assay, seedlings of two bread wheat genotypes Barbela 7/72/92 and Anahuac showing contrasting response to Al toxicity, were raised in hydroponic nutrient solution with pH 4.0 during four days and further, were shifted to fresh nutrient solution with (74 µM Al in the form of $AlCl_3 \cdot 6H_2O$; stress treatment) or without Al (control treatment). Both root and shoot tissues were collected separately after treatment at specific time points (0, 2, 6, 12 and 24 h) from control and Al stress imposed seedlings. For pH assay, four days old seedlings of both genotypes grown in hydroponic nutritive solution with pH 6.0 were transferred in fresh nutritive solutions having different levels of pH (pH 6.0, pH 5.0 and pH 4.0). Root samples of both genotypes from each treatment were collected after 0 h and 6 h exposure to different levels of pH.

Three biological replicates of each sample were prepared and duplicate quantitative assays were performed for each cDNA sample. *TaSTOP1* homoeologues gene expression pattern was determined using the SYBR Premix Ex Taq (Takara, Japan) and the ABI 7500 Real-Time FAST PCR System (Applied Biosystems, USA). The $2^{-\Delta\Delta C}{}_T$ method (Livak and Schmittgen, 2001), was used to quantify the relative expression levels of *TaSTOP1* homoeologues in comparison to *18SRNA* endogenous control.

Accession numbers

Sequence data from this article can be found in the GenBank database under the following accession numbers: Barbela 7/72/92 (TaSTOP1-A: GenBank number KF034793; TaSTOP1-B: GenBank number KF034794; TaSTOP1-D: GenBank number KF034795), Anahuac (TaSTOP1-A: GenBank number KF034796; TaSTOP1-B: GenBank number KF034797; TaSTOP1-D: GenBank number KF034798), Viloso Mole (TaSTOP1-A: GenBank number KF034801; TaSTOP1-B: GenBank number KF034802; TaSTOP1-D: GenBank number KF034803), Saloio (TaSTOP1-A: GenBank number KF034804; TaSTOP1-B:

GenBank number KF034805; TaSTOP1-D: GenBank number KF034806), Chinese Spring (TaSTOP1-A: GenBank number KF034799; TaSTOP1-B: GenBank number KF034800) and Ruivo (TaSTOP1-D: GenBank number KF034807).

REFERENCES

Aniol, A. and Gustafson, J.P. (1984). Chromosome location of genes controlling aluminum tolerance in wheat, rye, and triticale. *Can J Genet Cytol*, 26: 701-705.

Azevedo, D.; Nascimento, L.; Labarre, J.; Toledano, M.B. and Rodrigues-Pousada, C. (2007). The *S. cerevisiae* Yap1 and Yap2 transcription factors share a common cadmium-sensing domain. *FEBS Lett*, 581(2): 187-95.

Barceló, J. and Poschenrieder, C. (2002). Fast root growth responses, root exudates, and internal detoxification as clues to the mechanisms of aluminium toxicity and resistance: A review. *Environ Exp Bot*, 48: 75-92.

Bottley, A.; Xia, G.M. and Koebner, R.M.D. (2006). Homoeologous gene silencing in hexaploid wheat. *Plant J*, 47: 897-906.

Cabrera, A.; Martin, A. and Barro, F. (2002): *In-situ* comparative mapping (ISCM) of Glu-1 loci in *Triticum* and *Hordeum*. *Chromosome Res*, 10: 49-54.

Cai, S.; Bai, G.H. and Zhang, D. (2008). Quantitative trait loci for aluminum resistance in Chinese wheat landrace FSW. *Theor Appl Genet*, 117(1): 49-56.

Century, K.; Reuber, T.L. and Ratcliffe, O.J. (2008). Regulating the regulators: The future prospects for transcription-factor-based agricultural biotechnology products. *Plant Physiol*, 147: 20-29.

Chen, Q.; Zhang, X.D.; Wang, S.S.; Wang, Q.F.; Wang, G.Q.; Nian, H.J.; Li, K.Z.; Yu, Y.X. and Chen, L.M. (2011). Transcriptional and physiological changes of alfalfa in response to aluminium stress. *J Agric Sci*, 149: 737-751

Clarke, B.C.; Phongkham, T.; Gianibelli, M.C.; Beasley, H. and Bekes, F. (2003). The characterisation and mapping of a family of LMW-gliadin genes: Effects on dough properties and bread volume. *Theor Appl Genet,* 106: 629-635.

Collins, N.C.; Shirley, N.J.; Saeed, M.; Pallotta, M. and Gustafson J.P. (2008). An ALMT1 gene cluster controlling aluminum tolerance at the Alt4 locus of rye (*Secale cereale* L.). *Genetics*, 179: 669-692.

Danyluk, J.; Kane, N.A.; Breton, G.; Limin, A.E.; Fowler, D.B. and Sarhan, F. (2003): TaVRT-1, a putative transcription factor associated with vegetative to reproductive transition in cereals. *Plant Physiol,* 132: 1849-1860.

Daraselia, N.D.; Tarchevskaya, S. and Narita, J.O. (1996): The promoter for tomato 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase Gene 2 has unusual regulatory elements that direct high-leve1 expression. *Plant Physiol*, 112: 727-733

Delhaize, E.; Ma, J.F. and Ryan, P.R. (2012): Transcriptional regulation of aluminium tolerance genes. *Trends Plant Sci*, 17(6): 341-48.

Eticha, D.; Zahn, M.; Bremer, M.; Yang, Z.B.; Rangel, A.; Rao, I.M. and Horst, W. (2010): Transcriptomic analysis reveals differential gene expression in response to aluminium in common bean (*Phaseolus vulgaris*) genotypes. *Ann Bot*, 105: 1119-1128.

Famoso, A.N.; Zhao, K.; Clark, R.T.; Tung, C.W.; Wright, M.H.; Bustamante, C.; Kochian, L.V. and McCouch, S.R. (2011): Genetic Architecture of Aluminum Tolerance in Rice (Oryza sativa) Determined through Genome-Wide Association Analysis and QTL Mapping. *PLoS Genet*, 7(8): e1002221. doi:10.1371/journal.pgen.1002221

Feldman, M.; Levy, A.A.; Fahima, T. and Korol, A. (2012): Genomic asymmetry in allopolyploid plants: wheat as a model. *J Exp Bot*, 63(14): 5045-5059.

Furukawa, J.; Yamaji, N.; Wang, H.; Mitani, N.; Murata, Y.; Sato, K.; Katsuhara, M.; Takeda, K. and Ma, J.F. (2007): An aluminum-activated citrate transporter in barley. *Plant Cell Physiol*, 48: 1081-1091.

Haake, V.; Cook, D.; Riechmann, J.L.; Pineda, O.; Thomas, M.F.; and Zhang, J.Z. (2002). Transcription factor CBF4 is a regulator of drought adaptation in Arabidopsis. *Plant Physiol*, 130: 639-48.

Hoekenga, O.A.; Maron, L.G.; Piñeros, M.A.; Cançado, G.M.A., Shaff, J.; Kobayashi, Y., Ryan, P.R.; Dong, B.; Delhaize, E.; Sasaki, T.; Matsumoto, H.; Yamamoto, Y.; Koyama, H.; Kochian, L.V. (2006). AtALMT1, which encodes a malate transporter, is identified as one of several genes critical for aluminum tolerance in Arabidopsis. *Proc Natl Acad Sci USA*, 103: 9738-9743.

Huang, C.F.; Yamaji, N.; Mitani, N.; Yano, M.; Nagamura, Y. and Ma, J.F. (2009). A bacterial-type ABC transporter is involved in aluminum tolerance in rice. *Plant Cell*, 21: 655-667.

Huang, S.; Sirikhachornkit, A.; Su, X.J.; Faris, J.; Gill, B.; Haselkorn, R. and Gornicki, P. (2002): Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase

of the Triticum/Aegilops complex and the evolutionary history of polyploid wheat. *Proc Natl Acad Sci USA*, 99 (12): 8133-8138.

Iuchi, S.; Koyama, H.; Iuchi, A.; Kobayashi, A.; Kitabayashi, S.; Kobayashi, Y.; Ikka, T.; Hirayama, T.; Shinozaki, K. and Kobayashi, M. (2007): Zinc finger protein STOP1 is critical for proton tolerance in Arabidopsis and coregulates a key gene in aluminum tolerance. *Proc Natl Acad Sci USA*, 104: 9900-9905.

Kochian, L.V.; Hoekenga, O.A. and Piñeros, M.A. (2004). How do crop plants tolerate acid soils? Mechanisms of aluminum tolerance and phosphorous efficiency. *Annu Rev Plant Biol*, 55: 459-493.

Larsen, P.B.; Geisler, M.J.B.; Jones, C.A.; Williams, K.M. and Cancel, J.D. (2005). ALS3 encodes a phloem-localized ABC transporter-like protein that is required for aluminum tolerance in Arabidopsis. *Plant J*, 41: 353-363.

Larsen, P.B.; Cancel, J.; Rounds, M. and Ochoa, V. (2007). Arabidopsis ALS1 encodes a root tip and stele localized half type ABC transporter required for root growth in an aluminum toxic environment. *Planta*, 225: 1447-1458.

Lin, R.; Ding, L.; Casola, C.; Ripoll, D.R.; Feschotte, C. and Wang, H. (2007): Transposase-derived transcription factors regulate light signalling in Arabidopsis. *Science*, 318: 1302-1305.

Liu, J.P., Magalhães, J.V.; Shaff, J. and Kochian, L.V. (2009): Aluminum activated citrate and malate transporters from the MATE and ALMT families function independently to confer Arabidopsis aluminum tolerance. *Plant J*, 57: 389-399.

Liu, Q.; Kasuga, M.; Sakuma, Y.; Abe, H.; Miura, S.; Yamaguchi-Shinozaki, K. and Shinozaki, K. (1998). Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis. *Plant Cell*, 10: 1391-1406.

Livak, K.J. and Schmittgen, T.D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta}C_T$. *Method*, 25(4): 402-408.

Ma, J.F. (2007): Syndrome of aluminum toxicity and diversity of aluminum resistance in higher plants. Int Rev Cytol, 264: 225-252.

Magalhães, J.V.; Liu, J.; Guimarães, C.T.; Lana, U.G.; Alves, V.M.; Wang, Y.H.; Schaffert, R.E.; Hoekenga, O.A.; Piñeros M.A.; Shaff, J.E.; Klein, P.E.; Carneiro, N.P.; Coelho, C.M.; Trick, H.N. and Kochian L.V. (2007). A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. *Nat Genet*, 39: 1156-1161.

Maron, L.G.; Piñeros, M.A.; Guimarães, C.T.; Magalhães, J.V.; Pleiman, J.K.; Mao, C.; Shaff, J.; Belicuas, S.N. and Kochian, L.V. (2010). Two functionally distinct members of the MATE (multi-drug and toxic compound extrusion) family of transporters potentially underlie two major aluminum tolerance QTLs in maize. *Plant J*, 61: 728-740.

Martins-Lopes, P.; Maçãs, B. and Guedes-Pinto, H. (2009). Portuguese bread wheat germplasm evaluation for aluminium tolerance. *Cereal Res Commun*, 37(2): 179-188.

Menezes, R.A.; Amaral, C.; Delaunay, A.; Toledano, M. and Rodrigues-Pousada, C. (2004). Yap8p activation in *Saccharomyces cerevisiae* under arsenic conditions. *FEBS Lett,* 566(1): 141-146.

Miller, J.H. (1972). Experiments in Molecular Genetics: Assay of β-Galactosidase, p. 352-355. CSH Laboratory Press, Cold Spring Harbor, NY.

Mochida, K.; Yamazaki, Y. and Ogihara, Y. (2004). Discrimination of homoeologous gene expression in hexaploid wheat by SNP analysis of contigs grouped from a large number of expressed sequence tags. *Mol Genet Genomics*, 270: 371-377.

Navakode, S.; Weidner, A.; Lohwasser, U.; Röder, M.S. and Börner, A. (2009). Molecular mapping of quantitative trait loci (QTLs) controlling aluminium tolerance in bread wheat. *Euphytica*, 166(2): 283-290.

Pedersen, C.; Rasmussen, S.K. and Linde-Laursen, I. (1996). Genome and chromosome identification in cultivated barley and related species of the Triticeae (Poaceae) by in situ hybridization with the GAA-satellite sequence. *Genome*, 39: 93-104.

Pérez, R.; de Bustos, A.; Jouve, N. and Cuadrado, A. (2009). Localization of *Rad50,* a single-copy gene, on group 5 chromosomes of wheat, using a FISH protocol employing Tyramide for Signal Amplification (Tyr-FISH). *Cytogenet Genome Res*, 125: 321-328.

Polle, E.; Konzak, C.F. and Kittrick, J.A. (1978). Visual detection of aluminium tolerance levels in wheat by hematoxylin staining of seedlings roots. *Crop Sci*, 18: 823-827.

Prieto, P.; Ramirez, M.C.; Ballesteros, J. and Cabrera, A. (2001). Identification of intergenomic translocations involving wheat, Hordeum vulgare and Hordeum chilense chromosomes by FISH. *Hereditas,* 135: 171-174.

Poschenrieder, C.; Gunse, B.; Corrales, I. and Barceló, J. (2008). A glance into aluminum toxicity and resistance in plants. *Sci Total Environ*, 400: 356-368.

Raman, H.; Zhang, K.R.; Cakir, A.M.; Appels, R.; Garvin, D.F.; Maron, L.G.; Kochian, L.V.; Moroni, J.S.; Raman, R.; Imtiaz, M.; Drake-Brockman, F.; Waters, I.; Martin, P.; Sasaki, T.; Yamamoto, Y.; Matsumoto, H.; Hebb, D.M.; Delhaize, E. and Ryan, P.R. (2005). Molecular characterization and mapping of ALMT1, the aluminium-tolerance gene of bread wheat (*Triticum aestivum* L.). *Genome*, 48: 781-791.

Raman, H.; Stodart. B.; Ryan, P.R.; Delhaize, E.; Emebiri, L.; Raman, R.; Coombes, N. and Milgate, A. (2010). Genome-wide association analyses of common wheat (*Triticum aestivum* L.) germplasm identifies multiple loci for aluminium resistance. *Genome*, 53(11): 957-966.

Ravel, C.; Martre, P.; Romeuf, I.; Dardevet, M.; El-Malki, R.; Bordes J.; Duchateau, N.; Brunel, D.; Balfourier, F. and Charmet, G. (2009). Nucleotide polymorphism in the wheat transcriptional activator Spa influences its pattern of expression and has pleiotropic effects on grain protein composition, dough visco elasticity, and grain hardness. *Plant Physi*l, 151: 2133-2144.

Rayburn, A.L. and Gill, B.S. (1986). Isolation of a D-genome specific repeated DNA sequence from Aegilops squarrosa. *Plant Mol Biol Rep*, 4: 102-109.

Ryan, P.R.; Delhaize, E. and Jones, D.L. (2001). Function and mechanism of organic anion exudation from plant roots. *Annu Rev Plant Physiol Plant Mol Biol*, 52: 527-560.

Ryan, P.R.; Raman, H.; Gupta, S.; Horst, W.J. and Delhaize, E. (2009). A second mechanism for aluminum resistance in wheat relies on the constitutive efflux of citrate from roots. *Plant Physiol*, 149: 340-351.

Ryan, P.R.; Tyerman, S.D.; Sasaki, T.; Furuichi, T.; Yamamoto, Y.; Zhang, W.H. and Delhaize, E. (2011). The identification of aluminium-resistance genes provides opportunities for enhancing crop production on acid soils. *J Exp Bot*, 62(1): 9-20.

Sakamoto, H.; Maruyama, K.; Sakuma, Y.; Meshi, T.; Iwabuchi, M.; Shinozaki, K. and Yamaguchi-Shinozaki, K. (2004). Arabidopsis Cys2/His2-type zinc-finger proteins

function as transcription repressors under drought, cold, and high-salinity stress conditions. *Plant Physiol*, 136: 2734-2746.

Sasaki, T.; Yamamoto, Y.; Ezaki, B.; Katsuhara, M.; Ahn, S.J.; Ryan, P.R.; Delhaize, E. and Matsumoto, H. (2004). A wheat gene encoding an aluminum-activated malate transporter. *Plant J*, 37: 645-653.

Sawaki, Y.; Iuchi, S.; Kobayashi, Y.; Ikka, T.; Sakurai, N.; Fujita, M.; Shinozaki, K.; Shibata, D.; Kobayashi, M. and Koyama, H. (2009). STOP1 regulates multiple genes that protect Arabidopsis from proton and aluminum toxicities. *Plant Physiol*, 150: 281-294.

Shitsukawa, N.; Tahira, C.; Kassai, K.; Hirabayashi, C.; Shimizu, T.; Takumi, S.; Mochida, K.; Kawaura, K.; Ogihara, Y. and Murai, K. (2007). Genetic and epigenetic alteration among three homoeologous genes of a class E MADS box gene in hexaploid wheat. *Plant Cell*, 19: 1723-1737.

Simões, C.C.; Melo, J.O.; Magalhães, J.V. and Guimarães, C.T. (2012). Genetic and molecular mechanisms of aluminum tolerance in plants. *Genet Mol Res*, 11(3): 1949-1957.

Stockinger, E.J.; Gilmour, S.J. and Thomashow, M.F. (1997). *Arabidopsis thaliana CBF1* encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Proc Natl Acad Sci USA,* 94: 1035-1040.

Tamura, K.; Dudley, J.; Nei, M. and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol*, 24: 1596-1599.

Tovkach, A.; Ryan, P.R.; Richardson, A.E.; Lewis, D.C.; Rathjen, T.M.; Ramesh, S.; Tyerman, S.D. and Delhaize, E. (2013). Transposon-mediated alteration of *TaMATE1B* expression in wheat confers constitutive citrate efflux from root apices. *Plant Physiol*, 161(2): 880-892.

von Uexküll, H.R. and Mutert, E. (1995). Global extent, development and economic impact of acid soils. In Plant-Soil Interactions at Low pH: Principles and Management, ed. RA Date, NJ Grundon, GE Raymet, ME Probert, pp. 5-19. Dordrecht, The Neth: Kluwer Academic.

Watanabe, N. and Koval, S.F. (2003). Mapping of chlorina mutant genes on the long arm of homologous group 7 chromosomes in common wheat with partial deletion lines. *Euphytica*, 129: 259-265.

Wolfe, S.A.; Nekludova, L. and Pabo, C.O. (2000). DNA recognition by Cys(2)His(2) zinc finger proteins. *Annu Rev Biophys Biomol Struct*, 29: 183-212.

Xia, J.; Yamaji, N.; Kasai, T. and Ma, J.F. (2010). Plasma membrane-localized transporter for aluminum in rice. *Proc Natl Acad Sci USA*, 107: 18381-18385.

Yamaguchi-Shinozaki, K. and Shinozaki, K. (2006). Transcriptional Regulatory Networks in Cellular Responses and Tolerance to Dehydration and Cold Stresses. *Annu Rev Plant Biol*, 57: 781-803.

Yamaji, N.; Huang, C.F.; Nagao, S.; Yano, M.; Sato, Y.; Nagamura, Y. And Ma, J.F. (2009). A zinc finger transcription factor ART1 regulates multiple genes implicated in aluminum tolerance in rice. *Plant Cell*, 21: 3339-3349.

Yokosho, K.; Yamaji, N. and Ma, J.F. (2010). Isolation and characterization of two MATE genes in rye. *Func Plant Biol*, 37: 296-303.

Zhou, L.L.; Bai, G.H.; Ma, H.X. and Carver, B.F. (2007). Quantitative trait loci for aluminum resistance in wheat. *Mol Breed*, 19(2): 153-161.

Appendix-I

Table: Detail of primers used in present investigation.

Assay	Primer name	Sequence (5'-3')
EST amplification	TaSTOP_EST_F	TGAGGGTTTGATGCTTTTCC
	TaSTOP_EST_R	CAAGGAAGGTTAGGTTGCTCA
5'UTR RACE	TaSTOP1-5UTR	CAAGGAAGGTTAGGTTGCTCAGCATGG
3'UTR RACE	TaSTOP1-3UTR	GACATCTCCGAGAACCCCTTCTCCTTC
5' and 3' UTR RACE	UNIVERSAL PRIMER*	CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT
Chromosome Mapping	TaSTOP1_A_F	GCAGAGGAGCGAGGCGATGGACGAC
	TaSTOP1_A_R	GCTGCAAGAACCCGGTCC TGAAG
	TaSTOP1_B_F	CAAATGCCCAAATCCTTGTGTC
	TaSTOP1_B_R	CCCCTGAGGCTGCTCCG
	TaSTOP1_D_F	ATGATATCAAAGCATCAGGAGCATT
	TaSTOP1_D_R	TCAACGGTGACGAGCTCCATT
Chromosomal Arm mapping	TaSTOP1_ditA_F	GAATACCCTTAATCCAGCCATGAT
	TaSTOP1_ditA_R	CCATCGCCTCGCTCCTCTGC
	TaSTOP1_ditB_F	CAAATGCCCAAATCCTTGTGTC
	TaSTOP1_ditB_R	CCCCTGAGGCTGCTCCG
	TaSTOP1_ditD_F	CCTTAACACAGCCCATGATG
	TaSTOP1_ditD_R	CTGCTCCAATGCTCCTGATGCTTTG
Transactivation assay	TaSTOP1_TA_F	AAAGCTTCGTCGTCGATGG
	TaSTOP1_TA_R	ggtaccTCAGCTGTCTCCACTAAG
Expression studies (Realtime PCR)	qTaSTOP1_expA_F	GAAAGGACAAGCTGTTCGGC
	qTaSTOP1_expA_R	CATCGCCTCGCTCCTCTG
	qTaSTOP1_ expB_F	CTTCGGGACCGGGTTCC
	qTaSTOP1_ expB_R	CCCTGAATAGAGGAAGAACTGAGATGA
	qTaSTOP1_expD_F	TCAGGAGCATTGGAGCAGCCTC
	qTaSTOP1_expD_R	CCTGGGAAGTTATACCCTGTGCTCG
	18SRNA_F	TCCACGAGGAATGCCTAGTAAGC
	18SRNA_R	ACAAAGGGCAGGACGTAGTC

***From SMARTer™ RACE cDNA Amplification Kit (Clonetech, USA)**

Appendix-II

Nucleotide sequence *TaSTOP1* homoeologues gene in different bread wheat genotypes

>COMPLETE_SEQ_TaSTOP1_BARBELA 7/72/92 _Genome A

CCCCAATCCCCTCCACCACCCACGCGCGTGGGCCGGCCGGCCACCGCCGCCCCCTGCTCTTCCTCGCCGGGTCCCCGCCGTTTCCACTCCGTCCGGCAATCATTGGCCATTGA
GGGTTTGATGCTTTTCCTGTCCTATTATCAATTTCTCGTTTGGGGAAGATCGTGCGAGATAGAAGAAAGGGCGGACTGGATTGTAATCTGAGCTCTGGTGGATCTAGACTGGA
AGTTGCATGAGAAAAAGTTCAGAAATTTCCATGAAAGCTTCGTCGTCGATGGCAAGCGACGCTTCAGGGAACACTGACCCTGGCCAACAGGGTGCTCGTTTCAGTTCCATG
GACCAGTCTTGCTTTGCAAGACCTGGCCAGTCAATCCCTGGCTACCCCCCCATTCTTTGGCCCTCAGTCTTCCAACTTTACCTTCCTGATGACAGTGTGGCTAAAGCGTGTGA
TCCGTTCGAGCCGAATCCTCCTCAGAACAATCCTGTGGCAGATTGGGACCCTCAGGCCATGCTGAGCAACCTAACCTTCCTTGAGCAGAAGATCAAGCAGGTGAAAGATAT
CGTGCAGTCCATGGGTAACCGAGGGAGCCAAGATGTTGGTGGTTCCTGCGAGCTTGCCGCAAAGCAGCAGCTCGTCACCGCTGATCTCACTTCCATCATAATTCAGCTCATC
TCGACTGCCGGCTCCATGCTTCCTTCCATGAAGACCCCGCTGCTTAGCAGCAATCCAGCGGTCAGGCAGCTCAACACGCCTGGTTCTCCCATGGGCTTTGGCTCGATTGTGA
ATCAGCGGCCAAGCACAGTCAGGGAGGAGATGGTTCCTGACATTACCAAGACCCCTGACTATGAGGATCTGATGAATACCCTTAATCCAGCCCATGATGAAAAGGATGATCT
GATCAAATGCCCAAATCCTTGTGTTGGGGAAGGGCCTGAGCCGGTTCCGATGGAAGACCATGACGTGAAGGAGAGCGATGATGGTGGCGAGGCAGAGCATCTCCCCCCTG
GTTCTTATGTGGTCTTGCAATTGGAGAAGGAGGAGATTTTAGCACCACACACTCATTTCTGTGTGATATGTGGCAAGGGTTTCAAGAGGGATGCTAACCTAAGGATGCACAT
GAGGGGCCATGGAGACGAGTACAAAACTCCCGCAGCTCTTGCCAAACCCATGAGAGATTCTGTCTCAGATCCTACACCAGTTACAAGGTACTCGTGCCCATATGTCGGTTGC
AAGCGCAACAAAGAGCACAGGAAGTTCCAGCCCCTCAAGACAATCTTGTGTGTGAAGAACCACTACAAGAGAAGCCATTGCGACAAGAGGTATACCTGCAGCCGATGCA
ACACCAAGAAGTTCTCAGTCATTGCGGACTTGAAGACTCATGAGAAGCACTGTGGGCGTGACAAGTGGCTCTGCTCATGTGGAACAACTTTCTCAAGAAAGGACAAGCTG
TTCGGCCATGTCGCGCTTTTCCAAGGGCACACACCTGCTCTTCCAATGGATGATATTAAAGCAACAGGAGCATCGGAGCAGAGGAGCGAGGCGATGGACGACATGGTGGGG
AGCACAGGGTATAACTTCCCAGGCAGCACGTCTGATGGTATTCCGAATCTAGACATGAAAGTTGCCGATGACACACGTGGTTATTTCTCTCCCTTGAACTTCGACCCCTGCTT
CGGCGCCCTCGATGACTTCGCCCGACCTGGATTCGACATCTCCGAGAACCCCTTCTCCTTCCTGCCTTCAGGACCGGGTTCTTGCAGCTTTGGGCAGCTTAGTGGAGACAGC
TGATGGAGTTCGTCGCCGTAGACAACCATGATGGTGTCGATTGTTATGTATGAATGTTATATGTATGACCGTCTCAGTTCTTCCTCTATTCAGGGAGCCATTTTGGTCCATGCTT
CTTTCTTAAGTGTACTTTTAGTTTGTGGTCATCAATAATGATCTTACATATAGATTTGAAAG

>COMPLETE_SEQ_TaSTOP1_BARBELA 7/72/92 _GENOME B

CCCCAACCCCTCCACCTCCCACGCGCGTCGGCCGGCCGGCCACCGCCGTCCCTGTTCTTCCTCACCGGGTCCCCGCCGTTTCCACTCCGTCCGGCAATCATTGGCCATTGAG
GGTTTGATGCTATTCCTGTCTTATTGTCAATTTCTCATTTGGGGAGGATCGTGTGAGATAGAAGAGAGGGCGGACTGGATTGTAATCTGAGCTCTGGTGGATCTAGACTGAAA
GTTGCATGAGAAAAAGTTCAGAAATTTCCATGAAAGCTTCGTCGTCGATGGCAAGCGACGCGTCAGGGAACACTGAACCTGGCCAACAGGGCGTTCGTTTCAGTTCCATG
GACCAGTCTTGCTTTGCAAGACCTGGCCAGTCAATCCCTGGCTACCCCCCCATTCTTTGGCCCTCAATCTTCCAACTTTTACCTTCCTGATGACAGTGTGGCTAAAGCGTGTGATCT
TCCGTTTGAACCGAATCCTCCACAGAGCAATCCTGTGGCAGACTGGGATCCTCAGGCCATGCTGAGCAACCTAACCTTCCTTGAGCAGAAGATCAAGCAGGTGAAAGATAT
CGTGCAGTCCATGGGTAACCGAGGGAGCCAAGATGTTGGTGGTTCCTGCGAGCTTGCCGCAAAGCAGCAGCTCGTCACCGCTGATCTCACTTCCATCATAATTCAGCTCATC
TCGACTGCCGGCTCCATGCTTCCTTCCATGAAGACCCCGCTCCTTAGCGGCAATCCAGCGGTCAGGCAGCTCAACACGCCTGGTTCTCCCATGGGCTTTGGCTCGATTGTGA
ATCAGCGGCCAAGCACAGTCAGGGAGGAGATGGTTCCTGACATTAGCAAGACCTCTGATTATGAGGAGCTGATGAATACCCTTAACACAGCCGATGATGAAAAGGATGATC
TGATCAAATGCCCAAATCCTTGTGTCGGGGAAGGGCCTGAGCCGGTTCCGATGGAAGACCATGACGTGAAGGAGAGCGATGATGGTGGTGAGGCAGAGCACCTCCCCCCT
GGTTCTTATGTGGTCTTGCAGTTGGAGAAGGAGGAAATTTTAGCACCACATACTCATTTCTGTGTGATATGTGGGAAGGGTTTCAAGAGGGATGCTAACCTAAGGATGCACA
TGAGGGGCCATGGAGACGAGTACAAGACTCCCGCAGCTCTTGCCAAACCCATGAGAGATTCTGGCTCAGATCCTACACCAGTTACAAGGTACTCGTGCCCATATGTCGGTT
GCAAGCGGAACAAAGAGCACAGGAAGTTCCAGCCCCTCAAGACAATCTTGTGTGTGAAGAACCACTACAAGAGAAGCCACTGTGACAAGAGGTATACCTGCAGCCGATG
CAATACCAAGAAGTTCTCAGTCATTGCGGACTTGAAGACTCATGAGAAGCACTGCGGGCGTGACAAGTGGCTCTGCTCATGTGGAACAACTTTCTCAAGAAAGAACAAGC
TGTTCGGCCATGTCGCGCTTTTCCAAGGGCACACACCTGCTCTTCCAATGGATGATATTAAAGCAACAGGAGCATCGGAGCAGCCTCAGGGGAGCGAGGCGATGGACGACA
TGGTGGGGAGCACAGGGTATAACTTCCCAGGCAGCACGTCTGATGGTATTCCGAATCTAGACATGAAAGTTGCCGATGACACACGTGGTTATTTCTCGCCCTTGAACTTCGA
CCCGTGCTTCGGCGCCCTCGATGACTTCGCCCGCCCTGGATTCGACATCTCCGAGAACCCCTTCTCCTTCCTGCCTTCGGGACCGGGTTCCTGCAGCTTTGGGCAGCTTAGT
GGAGACAGCTGATGGAGATCGTCACCGTAGACAACCATGATGGTGTCGATCGTTATGTATGAATGTTATATGTATGATCATCTCAGTTCTTCCTCTATTCAGGGAGCCATTTTG
GTCCATGCCTCTTTCTTAAGTGTACTTTTAGTTTGTGGTCATCAATAATGATCTTACATATAGATTTGAAAG

>Complete_SEQ_ TaSTOP1_ BARBELA 7/72/92 _GENOME D

CCCCAATCCCCTCCACCTCCCACGCGCGTCGGCCGGCCACCGCCGTCCCTGTTCTTCCTCGCCGGGTCCCCGCCGTTTCCACTCCGTCCGGCAATCATTGGCCATTGAGGGT
TTGATGCTTTTCCTGTCTTATTATCAATTTCTCATTTGGGGAAGATCGTGTGAGATAGAAGAGAGGGCGGACTGGATTGTAATCTGAGGTCTAGTGGATCTAAACTGAAAGTT
GCATGAGAAAAAGTTCAGAAATTTCCATGAAAGCTTCGTCGTCGATGGCAAGCGACGCGTCAGGGAACACTGAACCTGGCCAACAGGGTGTTCGTTTCAGTTCCATGGAC
CAGTCTTGCTTTGCAAGACCTGGCCAGTCAATCCCTGGCTACCCCCCCATTCTTTGGCCCTCAATCTTCCAACTTTTACCTTCCTGATGACAGTGTGGCTAAAGCGTGTGATCT
GTTTGAACCGAATCCTCCACAGAACAATCCTGTGGCTGATTGGGATCCTCAGGCCATGCTGAGCAACCTAACCTTCCTTGAGCAGAAGATCAAGCAGGTGAAAGATATCGT
GCAGTCCATGGGTAACCGAGGGAGCCAAGATGCTGGTGGTTCCTGCGAGCTTGCTGCAAAGCAGCAGCTCGTCACCGCCGATCTCACTTCCATCATAATTCAGCTCATCTCA
ACTGCCGGCTCCATGCTTCCTTCCATGAAGACCCCGCTCCTTAGCGGCAATCCAGCGGTCAGGCAGCTCAACACGCCTGGTTCTCCCATGGGCTTTGGCTCGATCGTGAATC
AGCGGCCAAGCACAATCAGGGAGGAGGTGGTTAATGACATTAGCAAGACCTCTGACTATGAGGAGCTGATGAATACCCTTAACACAGCCCATGATGAAAAGGATGATCTGA
TCAAATGCCCAAATCCTTGTGTCGGGGAAGGACCTGAGCCGGTTCCGATGGAAGACCATGACGTGAAGGAGAGCGATGATGGTGGCGAGGCAGAGCACCTCCCCCCTGGT
TCTTATGTGGTCTTGCAATTGGAGAAGGAGGAAATTCTAGCACCACACACTCATTTCTGTGTGATATGTGGCAAGGGTTTCAAGAGGGATGCTAACCTAAGGATGCACATGA
GGGGCCATGGAGACGAGTACAAGACTCCTGCAGCTCTTGCCAAACCCATGAGAGATCCTGGCTCAGATCCTACACCAGTTACAAGGTACTCGTGCCCATATGTCGGTTGCA
AGCGGAACAAAGAGCACAGGAAGTTCCAGCCTCTCAAGACAATCTTGTGTGTGAAGAACCACTACAAGAGAAGCCACTGTGACAAGAGGTATACCTGCAGCCGATGCAA
CACCAAGAAGTTCTCAGTCATTGCTGACTTGAAGACTCATGAGAAGCACTGCGGGCGTGACAAGTGGCTCTGCTCATGTGGAACGACCTTCTCAAGAAAAGACAAGCTGT
TCGGGCATGTCGCGCTTTTCCAAGGGCACACACCTGCTCTTCCAATGGATGATATCAAAGCATCAGGAGCATTGGAGCAGCCTCAGGGGAGCGAGGCGATGGACGACATGG
TGGCGAGCACAGGGTATAACTTCCCAGGCAGCACGTCTGATGGTATTCCGAATCTAGACATGAAAGTTGCCGATGACACACGTGGTTATTTCTCGCCCTTGAACTTCGACCC
GTGCTTCGGTGCCCTCGATGACTTTGCTCGCCCTGGATTCGACATCTCTGAGAACCCCTTCTCCTTCCTGCCTTCAGGACCGGGGTCGTGCAGTTTTGGGCAGCTTAGTGGC
GACAGCTAATGGAGCTCGTCACCGTTGACAACCATGATGGTGTCGATCGTTATGTATGAATGTTATATGTATGACCGTCTTAGTTCTTCCTCTATTCAGGGAGCCATTTTGGTCC
ATGCCTCTTTCTTAAGTGTACTTTTAGTTTGTGGTCATCAATAATGATCTTACATATAGATTTGAAAG

>Complete_SEQ_Ta STOP1_ Anahuac _GENOME A

CCCCAATCCCCTCCACCACCCACGCGCGTGGGCCGGCCGGCCACCGCCGCCCCCTGCTCTTCCTCGCCGGGTCCCCGCCGTTTCCACTCCGTCCGGCAATCATTGGCCATTGA
GGGTTTGATGCTTTTCCTGTCCTATTATCAATTTCTCGTTTGGGGAAGATCGTGCGAGATAGAAGAAAGGGCGGACTGGATTGTAATCTGAGCTCTGGTGGATCTAGACTGGA
AGTTGCATGAGAAAAAGTTCAGAAATTTCCATGAAAGCTTCGTCGTCGATGGCAAGCGACGCTTCAGGGAACACTGACCCTGGCCAACAGGGTGCTCGTTTCAGTTCCATG
GACCAGTCTTGCTTTGCAAGACCTGGCCAGTCAATCCCTGGCTACCCCCCCATTCTTTGGCCCTCAGTCTTCCAACTTTACCTTCCTGATGACAGTGTGGCTAAAGCGTGTGA
TCCGTTCGAGCCGAATCCTCCTCAGAACAATCCTGTGGCAGATTGGGACCCTCAGGCCATGCTGAGCAACCTAACCTTCCTTGAGCAGAAGATCAAGCAGGTGAAAGATAT
CGTGCAGTCCATGGGTAACCGAGGGAGCCAAGATGTTGGTGGTTCCTGCGAGCTTGCCGCAAAGCAGCAGCTCGTCACCGCTGATCTCACTTCCATCATAATTCAGCTCATC
TCGACTGCCGGCTCCATGCTTCCTTCCATGAAGACCCCGCTGCTTAGCAGCAATCCAGCGGTCAGGCAGCTCAACACGCCTGGTTCTCCCATGGGCTTTGGCTCGATTGTGA
ATCAGCGGCCAAGCACAGTCAGGGAGGAGATGGTTCCTGACATTACCAAGACCCCTGACTATGAGGATCTGATGAATACCCTTAATCCAGCCCATGATGAAAAGGATGATCT
GATCAAATGCCCAAATCCTTGTGTTGGGGAAGGGCCTGAGCCGGTTCCGATGGAAGACCATGACGTGAAGGAGAGCGATGATGGTGGCGAGGCAGAGCATCTCCCCCCTG
GTTCTTATGTGGTCTTGCAATTGGAGAAGGAGGAGATTTTAGCACCACACACTCATTTCTGTGTGATATGTGGCAAGGGTTTCAAGAGGGATGCTAACCTAAGGATGCACAT
GAGGGGCCATGGAGACGAGTACAAAACTCCCGCAGCTCTTGCCAAACCCATGAGAGATTCTGTCTCAGATCCTACACCAGTTACAAGGTACTCGTGCCCATATGTCGGTTGC
AAGCGCAACAAAGAGCACAGGAAGTTCCAGCCCCTCAAGACAATCTTGTGTGTGAAGAACCACTACAAGAGAAGCCATTGCGACAAGAGGTATACCTGCAGCCGATGCA
ACACCAAGAAGTTCTCAGTCATTGCGGACTTGAAGACTCATGAGAAGCACTGTGGGCGTGACAAGTGGCTCTGCTCATGTGGAACAACTTTCTCAAGAAAGGACAAGCTG
TTCGGCCATGTCGCGCTTTTCCAAGGGCACACACCTGCTCTTCCAATGGATGATATTAAAGCAACAGGAGCATCGGAGCAGAGGAGCGAGGCGATGGACGACATGGTGGGG
AGCACAGGGTATAACTTCCCAGGCAGCACGTCTGATGGTATTCCGAATCTAGACATGAAAGTTGCCGATGACACACGTGGTTATTTCTCTCCCTTGAACTTCGACCCCTGCTT
CGGCGCCCTCGATGACTTCGCCCGACCTGGATTCGACATCTCCGAGAACCCCTTCTCCTTCCTGCCTTCAGGACCGGGTTCTTGCAGCTTTGGGCAGCTTAGTGGAGACAGC
TGATGGAGTTCGTCGCCGTAGACAACCATGATGGTGTCGATTGTTATGTATGAATGTTATATGTATGACCGTCTCAGTTCTTCCTCTATTCAGGGAGCCATTTTGGTCCATGCTT
CTTTCTTAAGTGTACTTTTAGTTTGTGGTCATCAATAATGATCTTACATATAGATTTGAAAG

>Complete_SEQ_ TaSTOP1_ Anahuac _GENOME B

CCCCAACCCCTCTACCTCCCACGCGCGTCGGCCGGCCGGCCACCGCCGTCCCTGTTCTTCCTCACCGGGTCCCCGCCGTTTCCACTCCGTCCGGCAATCATTGGCCATTGAG
GGTTTGATGCTATTCCTGTCTTATTGTCAATTTCTCATTTGGGGAGGATCGTGTGAGATAGAAGAGAGGGCGGACTGGATTGTAATCTGAGCTCTGGTGGATCTAGACTGAAA
GTTGCATGAGAAAAAGTTCAGAAATTTCCATGAAAGCTTCGTCGTCGATGGCAAGCGACGCGTCAGGGAACACTGAACCTGGCCAACAGGGCGTTCGTTTCAGTTCCATG
GACCAGTCTTGCTTCGCAAGACCTGGCCAGTCAATCCCTGGCTACCCCCCCATTCTTTGGCCCTCAATCTTCCAACTTTTACCTTCCTGATGACAGTGTGGCTAAAGCGTGTGA
TCCGTTTGAACCGAATCCTCCACAGAGCAATCCTGTGGCAGACTGGGATCCTCAGGCCATGCTGAGCAACCTAACCTTCCTTGAGCAGAAGATCAAGCAGGTGAAAGATAT

CGTGCAGTCCATGGGTAACCGAGGGAGCCAAGATGTTGGTGGTTCCTGCGAGCTTGCCGCAAAGCAGCAGCTCGTCACCGCTGATCTCACTTCCATCATAATTCAGCTCATC
TCGACTGCCGGCTCCATGCTTCCTTCCATGAAGACCCCGCTCCTTAGCAGCAATCCAGCGGTCAGGCAACTCAACACGCCTGGTTCTCCCATGGGCTTTGGCTCGATTGTGA
ATCAGCGGCCAAGCACAGTCAGGGAGGAGATGGTTCCTGACATTAGCAAGACCTCTGATTATGAGGAGCTGATGAATACCCTTAACACAGCGCATGATGAAAAGGATGATC
TGATCAAATGCCCAAATCCTTGTGTCGGGGAAGGGCCTGAGCCGGTTCCGATGGAAGACCATGACGTGAAGGAGAGCGATGATGGTGGTGAGGCAGAGCACCTCCCCCCT
GGTTCTTATGTGGTCTTGCAGTTGGAGAAGGAGGAAATTTTAGCACCACATACTCATTTCTGTGTGATATGTGGGAAGGGTTTCAAGAGGGATGCTAACCTAAGGATGCACA
TGAGGGGCCATGGAGACGAGTACAAGACTCCCGCAGCTCTTGCCAAACCCATGAGAGATTCTGGCTCAGATCCTACACCAGTTACAAGGTACTCGTGCCCATATGTCGGTT
GCAAGCGGAACAAAGAGCACAGGAAGTTCCAGCCCCTCAAGACAATCTTGTGTGTGAAGAACCACTACAAGAGAAGCCACTGTGACAAGAGGTATACCTGCAGCCGATG
CAATACCAAGAAGTTCTCAGTCATTGCGGACTTGAAGACTCATGAGAAGCACTGCGGGCGTGACAAGTGGCTCTGCTCATGTGGAACAACTTTCTCAAGAAAGGACAAGC
TGTTCGGCCATGTCGCGCTTTTCCAAGGGCACACACCTGCTCTTCCAATGGATGATATTAAAGCAACAGGAGCATCGGAGCAGCCTCAGGGGAGCGAGGCGATGGACGACA
TGGTGGGGAGCACAGGGTATAACTTCCCAGGCAGCACGTCTGATGGTATTCCGAATCTAGACATGAAAGTTGCCGATGACACACGTGGTTATTTCTCGCCCTTGAACTTCGA
CCCGTGCTTCGGCGCCCTCGATGACTTCGCCCGCCCTGGATTCGACATCTCCGAGAACCCCTTCTCCTTCCTGCCTTCGGGACCGGGTTCCTGCAGCTTTGGGCAGCTTAGT
GGAGACAGCTGATGGAGATCGTCACCGTAGACAACCATGATGGTGTCGATCGTTATGTATGATGAATGTTATATGTATGATCATCTCAGTTCTTCCTCTATTCAGGGAGCCATTTTG
GTCCATGCCTCTTTCTTAAGTGTACTTTTAGTTTGTGGTCATCAATAATGATCTTACATATAGATTTGAAAG

>Complete_SEQ_Ta STOP1_ Anahuac _GENOME D

CCCCAATCCCCTCCACCTCCCACGCGCGTCGGCCGGCCACCGCCGTCCCTGTTCTTCCTCGCCGGGTCCCCGCCGTTTCCACTCCGTCCGGCAATCATTGGCCATTGAGGGT
TTGATGCTTTTCCTGTCTTATTATCAATTTCTCATTTGGGGAAGATCGTGTGAGATAGAAGAGAGGGCGGACTGGATTGTAATCTGAGGTCTAGTGGATCTAAACTGAAAGTT
GCATGAGAAAAAGTTCAGAAATTTCCATGAAAGCTTCGTCGTCGATGGCAAGCGACGCGTCAGGGAACACTGAACCTGGCCAACAGGGTGTTCGTTTCAGTTCCATGGAC
CAGTCTTGCTTTGCAAGACCTGGCCAGTCAATCCCTGGCTACCCCCCATTCTTTGGCCCTCAATCTTCCAACTTTTACCTTCCTGATGACAGTGTGGCTAAAGCGTGTGATCT
GTTTGAACCGAATCCTCCACAGAACAATCCTGTGGCTGATTGGGATCCTCAGGCCATGCTGAGCAACCTAACCTTCCTTGAGCAGAAGATCAAGCAGGTGAAAGATATCGT
GCAGTCCATGGGTAACCGAGGGAGCCAAGATGCTGGTGGTTCCTGCGAGCTTGCTGCAAAGCAGCAGCTCGTCACCGCCGATCTCACTTCCATCATAATTCAGCTCATCTCA
ACTGCCGGCTCCATGCTTCCTTCCATGAAGACCCCGCTCCTTAGCAGCAATCCAGCGGTCAGGCAGCTCAACACGCCTGGTTCTCCCATGGGCTTTGGCTCGATCGTGAATC
AGCGGCCAAGCACAATCAGGGAGGAGGTGGTTAATGACATTAGCAAGACCTCTGACTATGAGGAGCTGATGAATACCCTTAACACAGCCCATGATGAAAAGGATGATCTGA
TCAAATGCCCAAATCCTTGTGTCGGGGAAGGACCTGAGCCGGTTCCGATGGAAGACCATGACGTGAAGGAGAGCGATGATGGTGGCGAGGCAGAGCACCTCCCCCCTGGT
TCTTATGTGGTCTTGCAATTGGAGAAGGAGGAAATTCTAGCACCACACACTCATTTCTGTGTGATATGTGGCAAGGGTTTCAAGAGGGATGCTAACCTAAGGATGCACATGA
GGGGCCATGGAGACGAGTACAAGACTCCCGCAGCTCTTGCCAAACCCATGAGAGATTCTGGCTCAGATCCTACACCAGTTACAAGGTACTCGTGCCCATATGTCGGTTGCA
AGCGGAACAAAGAGCACAGGAAGTTCCAGCCTCTCAAGACAATCTTGTGTGTGAAGAACCACTACAAGAGAAGCCACTGTGACAAGAGGTATACCTGCAGCCGATGCAA
CACCAAGAAGTTCTCAGTCATTGCTGACTTGAAGACTCATGAGAAGCACTGCGGGCGTGACAAGTGGCTCTGCTCATGTGGAACGACCTTCTCAAGAAAAGACAAGCTGT
TTGGGCATGTCGCGCTTTTCCAAGGGCACACACCTGCTCTTCCAATGGATGATATCAAAGCATCAGGAGCATTGGAGCAGCCTCAGGGGAGCGAGGCGATGGACGACATGG
TGGCGAGCACAGGGTATAACTTCCCAGGCAGCACGTCTGATGGTATTCCGAATCTAGACATGAAAGTTGCCGATGACACACGTGGTTATTTCTCGCCCTTGAACTTCGACCC
GTGCTTCGGTGCCCTCGATGACTTTGCTCGCCCTGGATTCGACATCTCCGAGAACCCCTTCTCCTTCCTGCCTTCAGGACCGGGGTCGTGCAGTTTTGGGCAGCTTAGTGGC
GACAGCTAATGGAGCTCGTCACCGTTGACAACCATGATGGTGTCGATCGTTATGTATGAATGTTATATGTATGACCGTCTTAGTTCTTCCTCTATTCAGGGAGCCATTTTGGTCC
ATGCCTCTTTCTTAAGTGTACTTTTAGTTTGTGGTCATCAATAATGATCAAGTACTTACATATAGATTTGAAAG

>SEQ_TaSTOP1_ ChineseSpring _Genome A

ATGAAAGCTTCGTCGTCGATGGCAAGCGACGCTTCAGGGAACACTGACCCTGGCCAACAGGGTGCTCGTTTCAGTTCCATGGACCAGTCTTGCTTTGCAAGACCTGGCCAG
TCAATCCCTGGCTACCCCCCATTCTTTGGTCCTCAGTCTTCCAACTTTTACCTTCCTGATGACAGTGTGGCTAAAGCGTGTGATCCGTTCGAGCCGAATCCTCCTCAGAACAA
TCCTGTGGCAGATTGGGACCCTCAGGCCATGCTGAGCAACCTAACCTTCCTTGAGCAGAAGATCAAGCAGGTGAAAGATATCGTGCAGTCCATGGGTAACCGAGGGAGCCA
AGATGTTGGTGGTTCCTGCGAGCTTGCCGCAAAGCAGCAGCTCGTCACCGCTGATCTCACTTCCATCATAATTCAGCTCATCTCGACTGCCGGCTCCATGCTTCCTTCCATGA
AGACCCCGCTCCTTAGCAGCAATCCAGCGGTCAGGCAACTCAACACGCCTGGTTCTCCCATGGGCTTTGGCTCGATTGTGAATCAGCGGCCAAGCACAGTCAGGGAGGAGA
GATGGTTCCTGACATTAGCAAGACCTCTGATTATGAGGAGCTGATGAATACCCTTAACACAGCGCATGATGAAAAGGATGATCTGATCAACTGCCCAAATCCTTGTGTCGGG
GAAGGGCCTGAGCCGGTTCCGATGGAAGACCATGACGTGAAGGAGAGCGATGATGGTGGTGAGGCAGAGCACCTCCCCCCTGGTTCTTATGTGGTCTTGCAGTTGGAGAA
GGAGGAAATTTTAGCACCACATACTCATTTCTGTGTGATATGTGGGAAGGGTTTCAAGAGGGATGCTAACCTAAGGATGCACATGAGGGGCCATGGAGACGAGTACAAGAC
TCCCGCAGCTCTTGCCAAACCCATGAGAGATTCTGGCTCAGATCCTACACCAGTTACAAGGTACTCGTGCCCATATGTCGGTTGCAAGCGGAACAAAGAGCACAGGAAGTT
CCAGCCCCTCAAGACAATCTTGTGTGTGAAGAACCACTACAAGAGAAGCCACTGTGACAAGAGGTATACCTGCAGCCGATGCAATACCAAGAAGTTCTCAGTCATTGCGGA
CTTGAAGACTCATGAGAAGCACTGCGGGCGTGACAAGTGGCTCTGCTCATGTGGAACAACTTTCTCAAGAAGGGACAAGCTGTTCGGCCATGTCGCGCTTTTCCAAGGGC
ACACACCTGCTCTAGCAATGGATGATATTAAAGCAACAGGAGCATCGGAGCAGCCTCAGGGGAGCGAGGCGATGGACGACATGGTGGGGAGCACAGGGTATAACTTCCCA
GGCAGCACGTCTGATGGTATTCCGAATCTAGACATGAAAGTTGCCGATGACACACGTGGTTATTTCTCGCCCTTGAACTTCGACCCGTGCTTCGGCGCCCTCGATGACTTCG
CCCCCCCTGGATTCGACATCTCCGAGAACCCCTTCTCCTTCCTGCCTTCGGGACCGGGTTCCTGCAGCTTTGGGCAGCTTAGTGGAGACAGCTGATGGAGATCGTCACCGTA
GACAACCATGATGGTGTCGATCGTTATGTATGAATGTTATATGTATGATCATCTCAGTTCTTCCTCTATTCAGGGAAATCACTAGTG

>SEQ_TaSTOP1_ Viloso Mole _Genome A

ATGAAAGCTTCGTCGTCGATGGCAAGCGACGCTTCAGGGAACACTGACCCTGGCCAACAGGGTGCTCGTTTCAGTTCCATGGACCAGTCTTGCTTTGCAAGACCTGGCCAG
TCAATCCCTGGCTACCCCCCATTCTTTGGCCCTCAGTCTTCCAACTTTTACCTTCCTGATGACAGTGTGGCTAAAGCGTGTGATCCGTTCGAGCCGAATCCTCCTCAGAACAA
TCCTGTGGCAGATTGGGACCCTCAGGCCATGCTGAGCAACCTAACCTTCCTTGAGCAGAAGATCAAGCAGGTGAAAGATATCGTGCAGTCCATGGGTAACCGAGGGAGCCA
AGATGTTGGTGGTTCCTGCGAGCTTGCCGCAAAGCAGCAGCTCGTCACCGCTGATCTCACTTCCATCATAATTCAGCTCATCTCGACTGCCGGCTCCATGCTTCCTTCCATGA
AGACCCCGCTGCTTAGCAGCAATCCAGCGGTCAGGCAGCTCAACACGCCTGGTTCTCCCATGGGCTTTGGCTCGATTGTGAATCAGCGGCCAAGCACAGTCAGGGAGGAG
ATGGTTCCTGACATTACCAAGACCCCTGACTATGAGGATCTGATGAATACCCTTAATCCAGCCCATGATGAAAAGGATGATCTGATCAAATGCCCAAATCCTTGTGTTGGGGA
AGGGCCTGAGCCGGTTCCGATGGAAGACCATGACGTGAAGGAGAGCGATGATGGTGGCGAGGCAGAGCATCTCCCCCCTGGTTCTTATGTGGTCTTGCAATTGGAGAAGG
AGGAGATTTTAGCACCACACACTCATTTCTGTGTGATATGTGGCAAGGGTTTCAAGAGGGATGCTAACCTAAGGATGCACATGAGGGGCCATGGAGACGAGTACAAAACTC
CCGCAGCTCTTGCCAAACCCATGAGAGATTCTGTCTCAGATCCTACACCAGTTACAAGGTACTCGTGCCCATATGTCGGTTGCAAGCGCAACAAAGAGCACAGGAAGTTCC
AGCCCCTCAAGACAATCTTGTGTGTGAAGAACCACTACAAGAGAAGCCATTGCGACAAGAGGTATACCTGCAGCCGATGCAACACCAAGAAGTTCTCAGTCATTGCGGAC
TTGAAGACTCATGAGAAGCACTGTGGGCGTGACAAGTGGCTCTGCTCATGTGGAACAACTTTCTCAAGAAAGGACAAGCTGTTCGGCCATGTCGCGCTTTTCCAAGGGCA
CACACCTGCTCTTCCAATGGATGATATTAAAGCAACAGGAGCATCGGAGCAGAGGAGCGAGGCGATGGACGACATGGTGGGGAGCACAGGGTATAACTTCCCAGGCAGCA
CGTCTGATGGTATTCCGAATCTAGACATGAAAGTTGCCGATGACACACGTGGTTATTTCTCTCCCTTGAACTTCGACCCCTGCTTCGGCGCCCTCGATGACTTCGCCCGACCT
GGATTCGACATCTCCGAGAAACCCCTTCTCCTTCCTGCCTTCAGGACCGGGTTCTTGGGCAGCTTAGTGGAGACAGCTGATGGAGTTCGTCGCCGTAGACAACCA
TGATGGTGTCGATTGTTATGTATGAATGTTATATGTATGACCGTCTCAGTTCTTCCTCTATTCAGGGAGCCATTTTGG

>SEQ_TaSTOP1_ Viloso Mole _Genome B

ATGAAAGCTTCGTCGTCGATGGCAAGCGACGCGTCAGGGAACACTGAACCTGGCCAACAGGGCGTTCGTTTCAGTTCCATGGACCAGTCTTGCTTTGCAAGACCTGGCCA
GTCAATCCCTGGCTACCCCCCATTCTTTGGCCCTCAATCTTCCAACTTTTACCTTCCTGATGACAGTGTGGCTAAAGCGTGTGATCCGTTTGAACCGAATCCTCCACAGAGCA
ATCCTGTGGCAGACTGGGATCCTCAGGCCATGCTGAGCAACCTAACCTTCCTTGAGCAGAAGATCAAGCAGGTGAAAGATATCGTGCAGTCCATGGGTAACCGAGGGAGCC
AAGATGTTGGTGGTTCCTGCGAGCTTGCCGCAAAGCAGCAGCTCGTTACCGCTGATCTCACTTCCATCATAATTCAGCTCATCTCGACTGCCGGCTCCATGCTTCCTTCCATG
AAGACCCCGCTCCTTAGCAGCAATCCAGCGGTCAGGCAACTCAACACGCCTGGTTCTCCCATGGGCTTTGGCTCGATTGTGAATCAGCGGCCAAGCACAGTCAGGGAGGA
GATGGTTCCTGACATTAGCAAGACCTCTGATTATGAGGAGCTGATGAATACCCTTAACACAGCGCATGATGAAAAGGATGATCTGATCAACTGCCCAAATCCTTGTGTCGGG
GAAGGGCCTGAGCCGGTTCCGATGGAAGACCATGATGTGAAGGAGAGCGATGATGGTGGTGAGGCAGAGCACCTCCCCCCTGGTTCTTATGTGGTCTTGCAGTTGGAGAA
GGAGGAAATTTTAGCACCACATACTCATTTCTGTGTGATATGTGGGAAGGGTTTCAAGAGGGATGCTAACCTAAGGATGCACATGAGGGGCCATGGAGACGAGTACAAGAC

TCCCGCAGCTCTTGCCAAACCCATGAGAGATTCTGGCTCAGATCCTACACCAGTTACAAGGTACTCGTGCCCATATGTCGGTTGCAAGCGGAACAAAGAGCACAGGAAGTT
CCAGCCCCTCAAGACAATCTTGTGTGTGAAGAACCACTACAAGAGAAGCCACTGTGACAAGAGGTATACCTGCAGCCGATGCAATACCAAGAAGTTCTCAGTCATTGCGGA
CTTGAAGACTCATGAGAAGCACTGCGGGCGTGACAAGTGGCTCTGCTCATGTGGAACAACTTTCTCAAGAAAGGACAAGCTGTTCGGCCATGTCGCGCTTTTCCAAGGGC
ACACACCTGCTCTTCCAATGGATGATATTAAAGCAACAGGAGCATCGGAGCAGCCTCAGGGGAGCGAGGCGATGGACGACATGGTGGGGAGCACAGGGTATAACTTCCCA
GGCAGCACGTCTGATGGTATTCCGAATCTAGACATGAAAGTTGCCGATGACACACGTGGTTATTTCTCGCCCTTGAACTTCGACCCGTGCTTCGGCGCCCTCGATGACTTCG
CCCGCCCTGGATTCGACATCTCCGAGAACCCCTTCTCCTTCCTGCCTTCGGGACCGGGTTCCTGCAGCTTTGGGCAGCTTAGTGGAGACAGCTGATGGAGATCGTCACCGTA
GACAACCATGATGGTGTCGATCGTTATGTATGAATGTTATATGTATGATCATCTCAGTTCTTCCTCTATTCAGGGAGCCATTTTGG

>SEQ_TaSTOP1_ Viloso Mole _Genome D

ATGAAAGCTTCGTCGTCGATGGCAAGCGACGCGTCAGGGAACACTGAACCTGGCCAACAGGGTGTTCGTTTCAGTTCCATGGACCAGTCTTGCTTTGCAAGACCTGGCCA
GTCAATCCCTGGCTACCCCCCATTCTTTGGCCCTCAATCTTCCAACTTTTACCTTCCTGATGACAGTGTGGCTAAAGCGTGTGATCTGTTTGAACCGAATCCTCCACAGAACA
ATCCTGTGGCTGATTGGGATCCTCAGGCCATGCTGAGCAACCTAACCTTCCTTGAGCAGAAGATCAAGCAGGTGAAAGATATCGTGCAGTCCATGGGTAACCGAGGGAGCC
AAGATGCTGGTGGTTCCTGCGAGCTTGCTGCAAAGCAGCAGCTCGTCACCGCCGATCTCACTTCCATCATAATTCAGCTCATCTCAACTGCCGGCTCCATGCTTCCTTCCATG
AAGACCCCGCTCCTTAGCAGCAATCCAGCGGTCAGGCAGCTCAACACGCCTGGTTCTCCCATGGGCTTTGGCTCGATCGTGAATCAGCGGCCAAGCACAATCAGGGAGGA
GGTGGTTAATGACATTAGCAAGACCTCTGACTATGAGGAGCTGATGAATACCCTTAACACAGCCCACGATGAAAAGGATGATCTGATCAAATGCCCAAATCCTTGTGTCGGG
GAAGGACCTGAGCCGGTTCCGATGGAAGACCATGACGTGAAGGAGAGCGATGATGGTGGCGAGGCAGAGCACCTCCCCCCTGGTTCTTATGTGGTCTTGCAATTGGAGAA
GGAGGAAATTCTAGCACCACACACTCATTTCTGTGTGATATGTGGCAAGGGTTTCAAGAGGGATGCTAACCTAAGGATGCACATGAGGGGCCATGGAGACGAGTACAAGAC
TCCCGCAGCTCTTGCCAAACCCATGAGAGATTCTGGCTCAGATCCTACACCAGTTACAAGGTACTCGTGCCCATATGTCGGTTGCAAGCGGAACAAAGAGCACAGGAAGTT
CCAGCCTCTCAAGACAATCTTGTGTGTGAAGAACCACTACAAGAGAAGCCACTGTGACAAGAGGTATACCTGCAGCCGATGCAACACCAAGAAGTTCTCAGTCATTGCTG
ACTTGAAGACTCATGAGAAGCACTGCGGGCGTGACAAGTGGCTCTGCTCATGTGGAACGACCTTCTCAAGAAAAGACAAGCTGTTCGGGCATGTCGCGCTTTTCCAAGGG
CACACACCTGCTCTTCCAATGGATGATATCAAAGCATCAGGAGCATTGGAGCAGCCTCAGGGGAGCGAGGCGATGGACGACATGGTGGCGAGCACAGGGTATAACTTCCCA
GGCAGCACGTCTGATGGTATTCCGAATCTAGACATGAAAGTTGCCGATGACACACGTGGTTATTTCTCGCCCTTGAACTTCGACCCGTGCTTCGGTGCCCTCGATGACTTTGC
TCGCCCTGGATTCGACATCTCTGAGAACCCCTTCTCCTTCCTGCCTTCGGGACCGGGGTCGTGCAGTTTTGGGCAGATTAGTGGCGACAGCTAATGGAGCTCGTCACCGTTG
ACAACCATGATGGTGTCGATCGTTATGTATGAATGTTATATGTATGACCGTCTTAGTTCTTCCTCTATTCAGGGAGCCATTTTGG

>SEQ_TaSTOP1_ Saloio _Genome A

ATGAAAGCTTCGTCGTCGATGGCAAGCGACGCTTCAGGGAACACTGACCCTGGCCAACAGGGTGCTCGTTTCAGTTCCATGGACCAGTCTTGCTTTGCAAGACCTGGCCAG
TCAATCCCTGGCTACCCCCCATTCTTTGGCCCTCAGTCTTCCAACTTTTACCTTCCTGATGACAGTGTGGCTAAAGCGTGTGATCCGTTCGAGCCGAATCCTCCTCAGAACAA
TCCTGTGGCAGATTGGGACCCTCAGGCCATGCTGAGCAACCTAACCTTCCTTGAGCAGAAGATCAAGCAGGTGAAAGATATCGTGCAGTCCATGGGTAACCGAGGGAGCCA
AGATGTTGGTGGTTCCTGCGAGCTTGCCGCAAAGCAGCAGCTCGTCACCGCTGATCTCACTTCCATCATAATTCAGCTCATCTCGACTGCCGGCTCCATGCTTCCTTCCATGG
AGACCCCGCTGCTTAGCAGCAATCCAGCGGTCAGGCAGCTCAACACGCCTGGTTCTCCCATGGGCTTTGGCTCGATTGTGAATCAGCGGCCAAGCACAGTCAGGGAGGAG
ATGGTTCCTGACATTACCAAGACCCCTGACTATGAGGATCTGATGAATACCCTTAATCCAGCCCATGATGAAAAGGATGATCTGATCAAATGCCCAAATCCTTGTGTTGGGGA
AGGGCCTGAGCCGGTTCCGATGGAAGACCATGACGTGAAGGAGAGCGATGATGGTGGCGAGGCAGAGCATCTCCCCCCTGGTTCTTATGTGGTCTTACAATTGGAGAAGGA
GGAGATTTTAGCACCACACACTCATTTCTGTGTGATATGTGGCAAGGGTTTCAAGAGGGATGCTAACCTAAGGATGCACATGAGGGGCCATGGAGACGAGTACAAAACTCC
CGCAGCTCTTGCCAAACCCATGAGAGATTCTGTCTCAGATCCTACACCAGTTACAAGGTACTCGTGCCCATATGTCGGTTGCAAGCGCAACAAAGAGCACAGGAAGTTCCA
GCCCCTCAAGACAATCTTGTGTGTGAAGAACCACTACAAGAGAAGCCATTGCGACAAGAGGTATACCTGCAGCCGATGCAACACCAAGAAGTTCTCAGTCATTGCGGACTT
GAAGACTCATGAGAAGCACTGTGGGCGTGACAAGTGGCTCTGCTCATGTGGAACAACTTTCTCAAGAAAGGACAAGCTGTTCGGCCATGTCGCGCTTTTCCAAGGGCACA
CACCTGCTCTTCCAATGGATGATATTAAAGCAACAGGAGCATCGGAGCAGAGGAGCGAGGCGATGGACGACATGGTGGGGAGCACAGGGTATAACTTCCCAGGCAGCACG
TCTGATGGTATTCCGAATCTAGACATGAAAGTTGCCGATGACACACGTGGTTATTTCTCTCCCTTGAACTTCGACCCCTGCTTCGGCGCCCTCGATGACTTCGCCCGACCTGG
ATTCGACATCTCCGAGAACCCCTTCTCCTTCCTGCCTTCAGGACCGGGTTCTTGCAGCTTTGGGCAGCTTAGTGGAGACAGCTGATGGAGTTCGTCGCCGTAGACAACCATG
ATGGTGTCGATTGTTATGTATGAATGTTATATGTATGACCGTCTCAGTTCTTCCTCTATTCAGGGAGCCATTTTGG

>SEQ_TaSTOP1_ Saloio _Genome B

ATGAAAGCTTCGTCGTCGATGGCAAGCGACGCGTCAGGGAACACTGAACCTGGCCAACAGGGCGTTCGTTTCAGTTCCATGGACCAGTCTTGCTTTGCAAGACCTGGCCA
GTCAATCCCTGGCTACCCCCCATTCTTTGGCCCTCAATCTTCCAACTTTTACCTTCCTGATGACAGTGTGGCTAAAGCGTGTGATCCGTTTGAACCGAATCCTCCACAGAGCA
ATCCTGTGGCAGACTGGGATCCTCAGGCCATGCTGAGCAACCTAACCTTCCTTGAGCAGAAGATCAAGCAGGTGAAAGATATCGTGCAGTCCATGGGTAACCGAGGGAGCC
AAGATGTTGGTGGTTCCTGCGAGCTTGCCGCAAAGCAGCAGCTCGTCACCGCTGATCTCACTTCCATCATAATTCAGCTCATCTCGACTGCCGGCTCCATGCTTCCTTCCATG
AAGACCCCGCTCCTTAGCAGCAATCCAGCGGTCAGGCAACTCAACACGCCTGGTTCTCCCATGGGCTTTGGCTCGATTGTGAATCAGCGGCCAAGCACAGTCAGGGAGGA
GATGGTTCCTGACATTAGCAAGACCTCTGATTATGAGGAGCTGATGAATACCCTTAACACAGCGCATGATGAAAAGGATGATCTGATCAAATGCCCAAATCCTTGTGTCGGG
GAAGGGCCTGAGCCGGTTCCGATGGAAGACCATGACGTGAAGGAGAGCGATGATGGTGGTGAGGCAGAGCACCTCCCCCCTGGTTCTTATGTGGTCTTGCAGTTGGAGAA
GGAGGAAATTTTAGCACCACATACTCATTTCTGTGTGATATGTGGGAAGGGTTTCAAGAGGGATGCTAACCTAAGGATGCACATGAGGGGCCATGGAGACGAGTACAAGAC
TCCCGCAGCTCTTGCCAAACCCATGAGAGATTCTGGCTCAGATCCTACACCAGTTACAAGGTACTCGTGCCCATATGTCGGTTGCAAGCGGAACAAAGAGCACAGGAAGTT
CCAGCCCCTCAAGACAATCTTGTGTGTGAAGAACCACTACAAGAGAAGCCACTGTGACAAGAGGTATACCTGCAGCCGATGCAATACCAAGAAGTTCTCAGTCATTGCGGA
CTTGAAGACTCATGAGAAGCACTGCGGGCGTGACAAGTGGCTCTGCTCATGTGGAACAACTTTCTCAAGAAAGGACAAGCTGTTCGGCCATGTCGCGCTTTTCCAAGGGC
ACACACCTGCTCTTCCAATGGATGATATTAAAGCAACAGGAGCATCGGAGCAGCCTCAGGGGAGCGAGGCGATGGACGACATGGTGGGGAGCACAGGGTATAACTTCCCA
GGCAGCACGTCTGATGGTATTCCGAATCTAGACATGAAAGTTGCCGATGACACACGTGGTTATTTCTCGCCCTTGAACTTCGACCCGTGCTTCGGCGCCCTTGATGACTTCGC
CCGCCCTGGATTCGACATCTCCGAGAACCCCTTCTCCTTCCTGCCTTCGGGACCGGGTTCCTGCAGCTTTGGGCAGCTTAGTGGAGACAGCTGATGGAGATCGTCACCGTAG
ACAACCATGATGGTGTCGATCGTTATGTATGAATGTTATATGTATGATCATCTCAGTTCTTCCTCTATTCAGGGAGCCATTTTGG

>SEQ_TaSTOP1_ Saloio _Genome D

ATGAAAGCTTCGTCGTCGATGGCAAGCGACGCGTCAGGGAACACTGAACCTGGCCAACAGGGTGTTCGTTTCAGTTCCATGGACCAGTCTTGCTTTGCAAGACCTGGCCA
GTCAATCCCTGGCTACCCCCCATTCTTTGGCCCTCAATCTTCCAACTTTTACCTTCCTGATGACAGTGTGGCTAAAGCGTGTGATCTGTTTGAACCGAATCCTCCACAGAACA
ATCCTGTGGCTGATTGGGATCCTCAGGCCATGCTGAGCAACCTAACCTTCCTTGAGCAGAAGATCAAGCAGGTGAAAGATATCGTGCAGTCCATGGGTAACCGAGGGAGCC
AAGATGCTGGTGGTTCCTGCGAGCTTGCTGCAAAGCAGCAGCTCGTCACCGCCGATCTCACTTCCATCATAATTCAGCTCATCTCAACTGCCGGCTCCATGCTTCCTTCCATG
AAGACCCCGCTCCTTAGCAGCAATCCAGCGGTCAGGCAGCTCAACACGCCTGGTTCTCCCATGGGCTTTGGCTCGATCGTGAATCAGCGGCCAAGCACAATCAGGGAGGA
GGTGGTTAATGACATTAGCAAGACCTCTGACTATGAGGAGCTGATGAATACCCTTAACACAGCCCACGATGAAAAGGATGATCTGATCAAATGCCCAAATCCTTGTGTCGGG
GAAGGACCTGAGCCGGTTCCGATGGAAGACCATGACGTGAAGGAGAGCGATGATGGTGGCGAGGCAGAGCACCTCCCCCCTGGTTCTTATGTGGTCTTGCAATTGGAGAA
GGAGGAAATTCTAGCACCACACACTCATTTCTGTGTGATATGTGGCAAGGGTTTCAAGAGGGATGCTAACCTAAGGATGCACATGAGGGGCCATGGAGACGAGTACAAGAC
TCCCGCAGCTCTTGCCAAACCCATGAGAGATTCTGGCTCAGATCCTACACCAGTTACAAGGTACTCGTGCCCATATGTCGGTTGCAAGCGGAACAAAGAGCACAGGAAGTT
CCAGCCTCTCAAGACAATCTTGTGTGTGAAGAACCACTACAAGAGAAGCCACTGTGACAAGAGGTATACCTGCAGCCGATGCAACACCAAGAAGTTCTCAGTCATTGCTG
ACTTGAAGACTCATGAGAAGCACTGCGGGCGTGACAAGTGGCTCTGCTCATGTGGAACGACCTTCTCAAGAAAAGACAAGCTGTTCGGGCATGTCGCGCTTTTCCAAGGG
CACACACCTGCTCTTCCAATGGATGATATCAAAGCATCAGGAGCATTGGAGCAGCCTCAGGGGAGCGAGGCGATGGACGACATGGTGGCGAGCACAGGGTATAACTTCCCA
GGCAGCACGTCTGATGGTATTCCGAATCTAGACATGAAAGTTGCCGATGACACACGTGGTTATTTCTCGCCCTTGAACTTCGACCCGTGCTTCGGTGCCCTCGATGACTTTGC
TCGCCCTGGATTCGACATCTCTGAGAACCCCTTCTCCTTCCTGCCTTCAGGACCGGGGTCGTGCAGTTTTGGGCAGCTTAGTGGCGACAGCTAATGGAGCTCGTCACCGTTG
ACAACCATGATGGTGTCGATCGTTATGTATGAATGTTATATGTATGACCGTCTTAGTTCTTCCTCTATTCAGGGAGCCATTTTGG

>SEQ_TaSTOP1_ Ruivo _Genome D

ATGAAAGCTTCGTCGTCGATGGCAAGCGACGCGTCAGGGAACACTGAACCTGGCCAACAGGGTGTTCGTTTCAGTTCCATGGACCAGTCTTGCTTTGCAAGACCTGGCCA
GTCAATCCCTGGCTACCCCCCATTCTTTGGCCCTCAATCTTCCAACTTTTACCTTCCTGATGACAGTGTGGCTAAAGCGTGTGATCTGTTTGAACCGAATCCTCCACAGAACA
ATCCTGTGGCTGATTGGGATCCTCAGGCCATGCTGAGCAACCTAACCTTCCTTGAGCAGAAGATCAAGCAGGTGAAAGATATCGTGCAGTCCATGGGTAACCGAGGGAGCC
AAGATGCTGGTGGTTCCTGCGAGCTTGCTGCAAAGCAGCAGCTCGTCACCGCCGATCTCACTTCCATCATAATTCAGCTCATCTCAACTGCCGGCTCCATGCTTCCTTCCATG
AAGACCCCGCTCCTTAGCAGCAATCCAGCGGTCAGGCAGCTCAACACGCCTGGTTCTCCCATGGGCTTTGGCTCGATCGTGAATCAGCGGCCAAGCACAATCAGGGAGGA
GGTGGTTAATGACATTAGCAAGACCTCTGACTATGAGGAGCTGATGAATACCCTTAACACAGCCCACGATGAAAAGGATGATCTGATCAAATGCCCAAATCCTTGTGTCGGG
GAAGGACCTGAGCCGGTTCCGATGGAAGACCATGACGTGAAGGAGAGCGATGATGGTGGCGAGGCAGAGCACCTCCCCCCTGGTTCTTATGTGGTCTTGCAATTGGAGAA
GGAGGAAATTCTAGCACCACACACTCATTTCTGTGTGATATGTGGCAAGGGTTTCAAGAGGGATGCTAACCTAAGGATGCACATGAGGGGCCATGGAGACGAGTACAAGAC
TCCCGCAGCTCTTGCCAAACCCATGAGAGATTCTGGCTCAGATCCTACACCAGTTACAAGGTACTCGTGCCCATATGTCGGTTGCAAGCGGAACAAAGAGCACAGGAAGTT
CCAGCCTCTCAAGACAATCTTGTGTGTGAAGAACCACTACAAGAGAAGCCACTGTGACAAGAGGTATACCTGCAGCCGATGCAACACCAAGAAGTTCTCAGTCATTGCTG
ACTTGAAGACTCATGAGAAGCACTGCGGGCGTGACAAGTGGCTCTGCTCATGTGGAACGACCTTCTCAAGAAAAGACAAGCTGTTCGGGCATGTCGCGCTTTTCCAAGGG
CACACACCTGCTCTTCCAATGGATGATATCAAAGCATCAGGAGCATTGGAGCAGCCTCAGGGGAGCGAGGCGATGGACGACATGGTGGCGAGCACAGGGTATAACTTCCCA
GGCAGCACGTCTGATGGTATTCCGAATCTAGACATGAAAGTTGCCGATGACACACGTGGTTATTTCTCGCCCTTGAACTTCGACCCGTGCTTCGGTGCCCTCGATGACTTTGC
TCGCCCTGGATTCGACATCTCTGAGAACCCCTTCTCCTTCCTGCCTTCAGGACCGGGGTCGTGCAGTTTTGGGCAGCTTAGTGGCGACAGCTAATGGAGCTCGTCACCGTTG
ACAACCATGATGGTGTCGATCGTTATGTATGAATGTTATATGTATGACCGTCTTAGTTCTTCCTCTATTCAGGGAGCCATTTTGG

Protein sequences of *TaSTOP1* homoeologues gene in different bread wheat genotypes

>TaSTOP1A_Barbela

MKASSSMASDASGNTDPGQQGARFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDPFEPNPPQNNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDV
GGSCELAAKQQLVTADLTSIIIQLISTAGSMLPSMKTPLLSSNPAVRQLNTPGSPMGFGSIVNQRPSTVREEMVPDITKTPDYEDLMNTLNPAHDEKDDLIKCPNPCVGEGPEPVPME
DHDVKESDDGGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDSVSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKN
HYKRSHCDKRYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKATGASEQRSEAMDDMVGSTGYNFPGSTSDGIPNLDMKVAD
DTRGYFSPLNFDPCFGALDDFARPGFDISENPFSFLPSGPGSCSFGQLSGDS

> TaSTOP1B_Barbela

MKASSSMASDASGNTEPGQQGVRFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDPFEPNPPQSNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDVG
GSCELAAKQQLVTADLTSIIIQLISTAGSMLPSMKTPLLSSNPAVRQLNTPGSPMGFGSIVNQRPSTVREEMVPDISKTSDYEELMNTLNTAHDEKDDLIKCPNPCVGEGPEPVPMED
HDVKESDDGGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDSGSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKNH
YKRSHCDKRYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKNKLFGHVALFQGHTPALPMDDIKATGASEQPQGSEAMDDMVGSTGYNFPGSTSDGIPNLDMKVA
DDTRGYFSPLNFDPCFGALDDFARPGFDISENPFSFLPSGPGSCSFGQLSGDS

> TaSTOP1D_Barbela

MKASSSMASDASGNTEPGQQGVRFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDLFEPNPPQNNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDA
GGSCELAAKQQLVTADLTSIIIQLISTAGSMLPSMKTPLLSGNPAVRQLNTPGSPMGFGSIVNQRPSTIREEVVNDISKTSDYEELMNTLNTAHDEKDDLIKCPNPCVGEGPEPVPME
DHDVKESDDGGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDPGSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKN
HYKRSHCDKRYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKASGALEQPQGSEAMDDMVASTGYNFPGSTSDGIPNLDMKV
ADDTRGYFSPLNFDPCFGALDDFARPGFDISENPFSFLPSGPGSCSFGQLSGDS

> TaSTOP1A_Anahuac

MKASSSMASDASGNTDPGQQGARFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDPFEPNPPQNNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDV
GGSCELAAKQQLVTADLTSIIIQLISTAGSMLPSMKTPLLSSNPAVRQLNTPGSPMGFGSIVNQRPSTVREEMVPDITKTPDYEDLMNTLNPAHDEKDDLIKCPNPCVGEGPEPVPME
DHDVKESDDGGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDSVSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKN
HYKRSHCDKRYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKATGASEQRSEAMDDMVGSTGYNFPGSTSDGIPNLDMKVAD
DTRGYFSPLNFDPCFGALDDFARPGFDISENPFSFLPSGPGSCSFGQLSGDS

> TaSTOP1B_Anahuac

MKASSSMASDASGNTEPGQQGVRFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDPFEPNPPQSNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDVG
GSCELAAKQQLVTADLTSIIIQLISTAGSMLPSMKTPLLSSNPAVRQLNTPGSPMGFGSIVNQRPSTVREEMVPDISKTSDYEELMNTLNTAHDEKDDLIKCPNPCVGEGPEPVPMED
HDVKESDDGGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDSGSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKNH
YKRSHCDKRYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKATGASEQPQGSEAMDDMVGSTGYNFPGSTSDGIPNLDMKVA
DDTRGYFSPLNFDPCFGALDDFARPGFDISENPFSFLPSGPGSCSFGQLSGDS

> TaSTOP1D_Anahuac

MKASSSMASDASGNTEPGRQGVRFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDLFEPNPPQNNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDA
GGSCELAAKQQLVTADLTSIIIQLISTAGSMLPSMKTPLLSSNPAVRQLNTPGSPMGFGSIVNQRPSTIREEVVNDISKTSDYEELMNTLNTAHDEKDDLIKCPNPCVGEGPEPVPME
DHDVKESDDGGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDSGSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKN
HYKRSHCDKRYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKASGALEQPQGSEAMDDMVASTGYNFPGSTSDGIPNLDMKV
ADDTRGYFSPLNFDPCFGALDDFARPGFDISENPFSFLPSGPGSCSFGQLSGDS

> TaSTOP1A_CS

MKASSSMASDASGNTDPGQQGARFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDPFEPNPPQNNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDV
GGSCELAAKQQLVTADLTSIIIQLISTAGSMLPSMKTPLLSSNPAVRQLNTPGSPMGFGSIVNQRPSTVREEMVPDITKTPDYEDLMNTLNPAHDEKDDLIKCPNPCVGEGPEPVPME
DHDVKESDDGGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMKGHGDEYKTPAALAKPMRDSVSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKN
HYKRSHCDKRYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKATGASEQRSEAMDDMVGSTGYNFPGSTSDGIPNLDMKVAD
DTRGYFSPLNFDPCFGALDDFARPGFDISENPFSFLPSGPGSCSFGQLSGDS

> TaSTOP1B_CS

MKASSSMASDASGNTEPGQQGVRFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDPFEPNPPQSNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDVG
GSCELAAKQQLVTADLTSIIIQLISTAGSMLPSMKTPLLSSNPAVRQLNTPGSPMGFGSIVNQRPSTVREEMVPDISKTSDYEELMNTLNTAHDEKDDLINCPNPCVGEGPEPVPMED
HDVKESDDGGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDSGSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKNH
YKRSHCDKRYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSREDKLFGHVALFQGHTPALAMDDIKATGASEQPQGSEAMDDMVGSTGYNFPGSTSDGIPNLDMKVA
DDTRGYFSPLNFDPCFGALDDFAPPGFDISENPFSFLPSGPGSCSFGQLSGDS

> TaSTOP1A_VilosoMole

MKASSSMASDASGNTDPGQQGARFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDPFEPNPPQNNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDV
GGSCELAAKQQLVTADLTSIIIQLISTAGSMLPSMKTPLLSSNPAVRQLNTPGSPMGFGSIVNQRPSTVREEMVPDITKTPDYEDLMNTLNPAHDEKDDLIKCPNPCVGEGPEPVPME
DHDVKESDDGGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDSVSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKN
HYKRSHCDKRYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKATGASEQRSEAMDDMVGSTGYNFPGSTSDGIPNLDMKVAD
DTRGYFSPLNFDPCFGALDDFARPGFDISENPFSFLPSGPGSCSFGQLSGDS

> TaSTOP1B_VilosoMole

MKASSSMASDASGNTEPGQQGVRFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDPFEPNPPQSNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDVG
GSCELAAKQQLVTADLTSIIIQLISTAGSMLPSMKTPLLSSNPAVRQLNTPGSPMGFGSIVNQRPSTVREEMVPDISKTSDYEELMNTLNTAHDEKDDLIKCPNPCVGEGPEPVPMED
HDVKESDDGGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDSGSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKNH
YKRSHCDKRYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKNKLFGHVALFQGHTPALPMDDIKATGASEQPQGSEAMDDMVGSTGYNFPGSTSDGIPNLDMKVA
DDTRGYFSPLNFDPCFGALDDFARPGFDISENPFSFLPSGPGSCSFGQLSGDS

> TaSTOP1D_VilosoMole

MKASSSMASDASGNTEPGQQGVRFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDLFEPNPPQNNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDA
GGSCELAAKQQLVTADLTSIIIQLISTAGSMLPSMKTPLLSSNPAVRQLNTPGSPMGFGSIVNQRPSTIREEVVNDISKTSDYEELMNTLNTAHDEKDDLIKCPNPCVGEGPEPVPME
DHDVKESDDGGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDSGSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKN
HYKRSHCDKRYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKASGALEQPQGSEAMDDMVASTGYNFPGSTSDGIPNLDMKV
ADDTRGYFSPLNFDPCFGALDDFARPGFDISENPFSFLPSGPGSCSFGQISGDS

> TaSTOP1A_Saloio

MKASSSMASDASGNTDPGQQGARFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDPFEPNPPQNNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDV
GGSCELAAKQQLVTADLTSIIIQLISTAGSMLPSMETPLLSSNPAVRQLNTPGSPMGFGSIVNQRPSTVREEMVPDITKTPDYEDLMNTLNPAHDEKDDLIKCPNPCVGEGPEPVPME
DHDVKESDDGGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDSVSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKN
HYKRSHCDKRYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKATGASEQRSEAMDDMVGSTGYNFPGSTSDGIPNLDMKVAD
DTRGYFSPLNFDPCFGALDDFARPGFDISENPFSFLPSGPGSCSFGQLSGDS

> TaSTOP1B_Saloio

MKASSSMASDASGNTEPGQQGVRFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDPFEPNPPQSNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDVG
GSCELAAKQQLVTADLTSIIIQLISTAGSMLPSMKTPLLSSNPAVRQLNTPGSPMGFGSIVNQRPSTVREEMVPDISKTSDYEELMNTLNTAHDEKDDLIKCPNPCVGEGPEPVPMED
HDVKESDDGGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDSGSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKNH
YKRSHCDKRYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKATGASEQPQGSEAMGDMVGSTGYNFPGSTSDGIPNLDMKVA
DDTRGYFSPLNFDPCFGALDDFARPGFDISENPFSFLPSGPGSCSFGQLSGDS

> TaSTOP1D_Saloio

MKASSSMASDASGNTEPGQQGVRFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDLFEPNPPQNNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDA
GGSCELAAKQQLVTADLTSIIIQLISTAGSMLPSMKTPLLSSNPAVRQLNTPGSPMGFGSIVNQRPSTIREEVVNDISKTSDYEELMNTLNTAHDEKDDLIKCPNPCVGEGPEPVPME
DHDVKESDDGGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDSGSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKN
HYKRSHCDKRYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKASGALEQPQGSEAMDDMVASTGYNFPGSTSDGIPNLDMKV
ADDTRGYFSPLNFDPCFGALDDFARPGFDISENPFSFLPSGPGSCSFGQLSGDS

>TaSTOP1D_Ruivo

MKASSSMASDASGNTEPGQQGVRFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDLFEPNPPQNNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDA
GGSCELAAKQQLVTADLTSIIIQLISTAGSMLPSMKTPLLSSNPAVRQLNTPGSPMGFGSIVNQRPSTIREEVVNDISKTSDYEELMNTLNTAHDEKDDLIKCPNPCVGEGPEPVPME
DHDVKESDDGGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDSGSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKN
HYKRSHCDKRYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKASGALEQPQGSEAMDDMVASTGYNFPGSTSDGIPNLDMKV
ADDTRGYFSPLNFDPCFGALDDFARPGFDISENPFSFLPSGPGSCSFGQLSGDS

Appendix- III

Table : Detail of STOP like proteins in different species used for phylogenetic analysis.

S. No.	Scientific Name	Sequence name and Protein ID	Amino acid length
1	*Manihot esculenta*	M.esculenta_cassava4.1_005470m	525
2	*Ricinus communis*	R.communis_29728.m000841	357
		R.communis_29589.m001269	416
3	*Linum usitatissimum*	L.usitatissimum_Lus10009435	469
4	*Populus trichocarpa*	P.trichocarpa_ POPTR_0013s11900	509
		P.trichocarpa_ POPTR_0019s11520	506
5	*Phaseolus vulgaris*	P.vulgaris_Phvulv091023250m	512
6	*Glycine max*	G.max_Glyma10g35940	507
		G. max_ XP_003539439.1	410
7	*Cucumis sativus*	C.sativus_Cucsa.079940	508
		C. sativus _XP_004173281.1	381
8	*Prunus persica*	P.persica_ppa004510m	505
		P. persica_ EMJ07676.1	388
9	*Malus domestica*	M.domestica_MDP0000453148	527
10	*Arabidopsis thaliana*	A.thaliana_AT1G34370	499
		A.thaliana_AT5G22890	373
11	*Capsella rubella*	C.rubella_Carubv10008876m	514
		C. rubella_ EOA22034.1	369
12	*Brassica rapa*	B.rapa_Bra036750	491
13	*Thellungiella halophila*	T.halophila_Thhalv10007411m	506
14	*Carica papaya*	C.papaya_evm.model.supercontig_66.56	466
15	*Citrus sinensis*	C.sinensis_orange1.1g009831m.g	524
16	*Eucalyptus grandis*	E.grandis_Eucgr.F02627	528
17	*Vitis vinifera*	V.vinifera_GSVIVT01034505001	527
		V. vinifera_ XP_002272574.1	423
18	*Mimulus guttatus*	M.guttatus_mgv1a007605m	402
19	*Aquilegia coerulea*	A.coerulea_Aquca_017_00778	535
20	*Sorghum bicolor*	S.bicolor_Sb03g041170	519
		S. bicolor_ XP_002444574.1	429
21	*Zea mays*	Z.mays_GRMZM2G068710	467
		Z.mays_GRMZM2G075956	519
22	*Setaria italica*	S.italica_Si001061m	517
		S.italic_Si010000m.g	477
23	*Oryza sativa*	O.sativa_Os01g65080	522
		O. sativa_Os08g0562300	385

S. No.	Scientific Name	Sequence name and Protein ID	Amino acid length
24	*Brachypodium distachyon*	B.distachyon_Bradi2g56320	525
		B.distachyon_ XP_003571264.1	387
25	*Triticum aestivum*	**>T.aestivum_genome_A**	510
	Cv Barbela 7/72/92	**>T.aestivum_genome_B**	512
		>T.aestivum_genome_D	512
26	*Solanum lycopersicum*	Solanum lycopersicum AK320912.1	513
		S. lycopersicum_ XP_004250985.1	384
27	*Hordeum vulgare*	Hordeum vulgare AK252406.1	512
28	*Musa acuminata*	M.acuminata_Achr5P01170_001	393
		M.acuminata_Achr10P15970_001	398
		M.acuminata_Achr6P18840_001	441
29	*Triticum urartu*	T. urartu_EMS66742.1	506
		T. urartu_EMS47263.1	330
30	*Aegilops tauschii*	Ae. tauschii_EMT26835.1	506
31	*Arabidopsis lyrata*	A.lyriata XP_002874102.1	368
32	*Fragaria vesca subsp. Vesca*	F. vesca_ XP_004294539.1	402
		F. vesca_ XP_004290434.1	402

>M.esculenta_cassava4.1_005470m

MDHKDRLCQETWANSSSSGNELPKNISTQKPSYANFNFNAQQHQQKWEDTSIFGVRTNSTFQEFNPQSETQSLLPSNNNDHIKIPDRESSRINETLQTNKVPDWDPKAMLNNLSFLEQKI
HQLQDLAHLIIGRRSQVLGRPDELLTQQQQLITADLTSIIVQLISTAGSLLPSVKHTLSVAMPPAGQLGGVLFPSGAGMNSSLQPQYSSASKVSDQSNQMDIAVNCGTEQNYSIEEHEMRD
EEDAEEGENLPPGSYEILQLEKEEILAPHTHFCTICGKGFKRDANLRMHMRGHGDEYKTPAALAKPNKEASSEPVLIKRYSCPFAGCKRNKDHKKFQPLKTILCVKNHYKRTHCDKSYT
CSRCNTKKFSVIADLKTHEKHCGKDKWLCSCGTTFSRKDKLFGHIALFQGHTPAIPPDETKATPGPSDKGDENEVANSVGSMNYSFGSGGANGSGAQDIMDVKGGVDDPASYFSPLNF
DTCNFGGFHEFPRPPFDDSESSFSFLISGLCNYSQKTGGESSSNNIH

>R.communis_29728.m000841

MDRKERLCTETWANSSSSGSGLPEHLSTEKPSFANFSSQQDQQKWEDTAIYGIRTDPPFQEFNLPSETQTLLPSNPDDQIKIMDEEDSHINEALQINKIPDWDPKAMLSNLSFLEQKIHQLQ
DLVHLIVGRRGQILERPDDLVAQQQQLITADLTSIIVQLISTAGSLLPSVKHTLFAATPGGQLGHLGEILFPSGMGMNSIVQPQHSSGNKLSHHSNQMDISGNCGTEQNYSVDEHELKDEED
VEEGENLPPGSYEILQLEKEEILAPHTHFCTICGKGFKRDANLRMHMRGHGDEYKTPAALAKPNKESSSEPVLKRYSCPFAGCKRNKDHKNYNLGIGYLSPPLNNLKINKMAI

>R.communis_29589.m001269

MIAGLSSRNFSNHHQQQHHHVLQVYPAASDDIVSTSLEASSSSEIGHANSLLYSLSVVKDKVNQVQSLVSILTSPDHHHQSYQQQNPEPNSIANLGTIIQEIIVTASSMMFTCQQMAIGSSN
SINNNASELHQPAPPQPKVAAPQENVLLPNSQDHQRGQSFFASENFDWYDNNNYSNNSCNGTENNQTASGMHNVISINTNKVERKELSENFSDVLQEGNNKSIKSYDIIELDASYLLAKY
THYCQVCGKGFKRDANLRMHMRAHGDEYKTSAALSNPMKNINTSAIGDSSEDSVMKLPRKYSCPQEGCRWNQKHAKFQPLKSMICVKNHYKRSHCPKMYVCKRCNRKQFSVLSDL
RTHEKHCGDLKWLCCCGTTFSRKDKLMGHVALFVGHTPAITSFNNPAKVESHQMQLDDR

>L.usitatissimum_Lus10009435

MDPKEVLCGEPWAKPSSSNDPDPKISLQNVPSLMDYGFRIDQPYQDRTENNNNKAPDWDPRAMLNNLSFLEQKIHQLQGVVHLIIGRKDQEVPPLGNRPDELVTQQQQQLVTADLTSII
VQLISTAGSLLPSVKHTLAATNPIGQFEQFGLIQPKPPTGTAPVPDQTDLMDLATNYSTEEHEMKDEEDTEDGENLPPGSYEILQLEKEEILAPHTHFCTICGKGFKRDANLRMHMRGHG
DEYKTPAALAKPPKEATSSEPTGAEPAMIKRYSCPFAGCKRNKDHKKFQPLKTILCVKNHYKRTHCDKSYICSRCHTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHIALF
QGHTPAIPPDETKALFEKGEDSNNIEAAAAANNKVGIGALDFNCLGSNSNIMDFKSVVAADDTAAATACFSPLSFDSFQEFNRQPFDDSDSSFSFLLSGSCNYSQKSCGDS

>P.trichocarpa_POPTR_0013s11900

MEPKDRPSLDAWANPSSSSNGLPRKMALENPSFTNFNLQQQEQKWDDTSILDYGIRIEPPFREFNQASESEYLLSSNCNNQPKVLDQEDGQNNEALQTSKLQDWDPRSMLNNLSFLEQK
IHHLQDLVHLIVGRKGQALGGQDQLVTQQQQLITADLTSIIVQLISTAGSLLPSVKHTFSTGTPNGQLGQLGGILFPPLAGMNCVPQPQHGSGSKVSDQCNQMDVTGNCGTEPNHSIEEHE
MKEEEDADEGENLPPGSYDILQLEKEEILAPHTHFCTICGKGFKRDANLRMHMRGHGDEYKTPAALAKPNKEPSSEPVIIKRYSCPFAGCKRNKDHKKFQPLKTILCVKNHYKRTHCDK
SYICSRCNTKKFSVMADLKTHEKHCGKDKWLCSCGTTFSRKDKLFGHIALFQGHTPAIPLEETKGPYVSCDKVDGNEASNKVGNTNFSLGSNAGSGSGAQNVIEVKEDADDPASYFSPL
SFDTCNFGGFHEFPRPPFDDSESSFSFLEAP

>P.trichocarpa_POPTR_0019s11520

MDLTDRLSVDVWADTSSSGNELPRRMSSENPSFTDFNLQQQQKQQQKWEDNIISPPFGGFNQASESGYLLLSNRNNQTKILDHQEDGKSIETLETNKVQDWDPRAMLSNLSFLEQKIHH
LQDLVHLFVGRKGQGLGGQDQLVTQQQQFLITADLTSIIVQLISTAGSLLPSVKHTLSTDTPNGHLGQLGGLLFPPVAGMNCSPKPQHGSGRRVSDQSDKMDGTGNCGTDQNHFIEEHET
KEEEDADEGDYLPPGSYDILQLEKEEILAPHTHFCTICGKGFKRDANLRMHMRGHGDEYKTPAALAKPNKESSSDPVVIKRYSCPFSGCKRNKDHKKFQPLKSILCVKNHYKRTHCDKS
YTCSRCNTKKFSVTADLKTHEKHCGKDRWLCSCGTTFSRKDKLFGHIALFQGHTPAIPLEETKGPAVSSDKVDGNEASNKVGNINFSFGSHAATGSGAKNVMEVNDDADDPSSCFSPLN
FDTCNFGGFHEFPRPPFDDSESSFSFLLS

>P.vulgaris_Phvulv091023250m

MDLEGGLCANNCTRSSSLTSPGNGLQTNLSSDPPSFYGLEIESPFNTFNPPSPPIQSVLPGQSNTDIEIPDQENCPLSDSSNTTKLQDWNPSAMLNNLCFLEEKIHQLRDLVHLIVNKKCQPF
GQSHELVTQEQQLVTADLTSIIVQLISTAGNLLPSVRHTLTNTNPLVGQMDQLHGINLPFGSEPSSGIRPQNSSGNKLFGKSTPNDRPNPCEMEQNYNMEEHEPKDEEDVDEGESLPPGSY
EILQLEKEEILAPHTHFCTICGKGFKRDANLRMHMRGHGDEYKTPAALAKPHKETGSEPKLIKRYSCPYAGCKRNKDHKKFLPLKTILCVKNHYKRTHCDKSYTCSRCNMKKFSVMAD
LKTHEKHCGKDKWLCSCGTTFSRKDKLFGHIALFQGHTPAIPLDDIKRVAEPPDPQNRECSNKVGSVNFCFGSKLSSENGVQNMMDMKGNIDDPMNYFSSLNFEGCNFGAFNEFPQTPF
EDSEGSFSFLMSGSFNYAPKFGGGESSTDNL

>G.max_Glyma10g35940

MDPKGSLCAKNCTRTSSLTSPGNGLQTNLSSDPPSFYGVRIESPFKEFIQPSPPTQSVLPGESNADIDIPDQENCPLSNSSHTSKLQDWNPSAMLNNLSFLEEKIHQLQDLVHVIVNKKCQP
FGQPHELVTQEQQLITADLTSIIVQLISTAGSLLPSVRHTLTNTNPLVGQLDQLHGINLPFGSEPSSGIRPQNNSGNKLFDQSTQNDLPNKLEMEQNYNMEEHEPKDEEDADEGENLPPGSY
EILQLEKEEILAPHTHFCTICGKGFKRDANLRMHMRGHGDKYKTPAALAKPHKETGSEPKLIKRYSCPYAGCKRNKDHKKFQPLKTILCVKNHYKRTHCDKSYTCSRCNTKKFSVMAD
LKTHEKHCGKDKWLCSCGTTFSRKDKLFGHIALFQGHTPAIPLDDTKGVAEPPDIQNRESNNKVESINFCFGSNPSTENVVQNIMDMKGNIDDPMNYFSSLNFEGCNFGAFNEFSQPPFE
DSEASFSFPMSGSYNYAPKFDSDNL

> G.max_ XP_003539439.1

MISRATSSFQSVHHQGLQMFPDAEDFVSPSVESGSGTSHSSSLLFNLSILKDKLSEVQTLVGVILSPNQSLPESTSMAISSMNSTIQEIVVTATSMMFTCQQMALTAPPGTNNNTNKLHHHQ
NKSLPHSSNINFATNGSCVLGNNIGTDNISDHHGLFSSTESETLDWFSESYNNNDDNSNSNIISKVAISENNIITRGGGGGVLSPRDEPNEEGLSPKMNSDDIIIELDAANLLAKYTHYCQV
CGKGFKRDANLRMHMRAHGDEYKTNAALSNPIKNQRDLECLMSVKPKRYSCPQEGCRWNQRHAKFQPLKSMICAKNHYKRSHCPKMYVCKRCNQKQFSVLSDLRTHEKHCGDLK
WLCSCGTSFSRKDKLMGHVALFVGHQPAINNNGLSYSGKLIRAAPHIANAD

>C.sativus_Cucsa.079940

MEPEDSLCSSEVWTKSSSFPNGNVQQPPQKWPDTSILDYAMDQPFQEFQEQTESKASLTCNPNQQIEILDKDSNQMNAALLASKIQDWDPRAMLNNLSFLEQKIHQLQELVHLIVGRRG
QVFGRPDELVVQQQQLITADLTSIIVQLISTAGSLLPSVKHNLSAAVPPVGQLEPFDKVIFASGPGTNGGVQSQHGDGTKLPELPTQVDGSSKCGKEQNMTVEEHESKDEEDADEHENLPP
GSYEILQLEKEEILAPHTHFCAICGKGFKRDANLRMHMRGHGDEYKTAAALAKPNKELGSETMLIKRYSCPFTGCKRNKDHKKFQPLKTILCVKNHYKRTHCDKSFTCSKCNSKKFSVI
ADLKTHEKHCGKDKWLCSCGTTFSRKDKLFGHIALFQGHTPAIPLDESKGATVSCDRGERYETTNKLGSINFSFGSIASGGSSVETIADAKGGVDDPLSYFSPLNFDSCNFGGFHEFPRPP
FENTDNAFSFLIPGSGNYTQKIWCRIKL

> C.sativus_ XP_004173281.1

MISNSQMLSLPSQHPLVSSSNDHHETLDNLLSSYSQSSNTLLYNLSVLKEKLHQIQSVVNIAVYNLQNSTESLPSAPPASSQSPVATAAVNSLIQELIMAASSMLFTCQQMDNLAVNASLHN
NNVGNNLQGQHHQHQQQGPVDDNNGRSNTLFNMANDHHHQRQDWYNTTTSNNYNKDNNNSRTIMTTKTTQDHRNPINIVELDASDLLAKYTHYCQICGKGFKRDANLRMHMR
AHGDEYKASGALSNPEKSHRKDLSNISKMGIKYSCPQEGCRWNQKHVKFQPLKSLICVKNHFKRTHCPKMYVCKLCSRKKFSVLSDLRTHEKHCGDVKWLCSCGTTFSRKDKLMGHV
ALFVGHTPAMGSSTKFLGKQEHEYVHSF

>P.persica_ppa004510m

MDHKEKLCADAWENPSSGNDLTGKISSDCPPFAIFKPQLHQQKWEDPSILDYGGIRIEPSFQEFNQPSETQTSLPCNSNNDTKMADREDGRMCETPQANKSQDWDPRTMLNNLSYLEQK
IHQLQDLVHLIVGRRGQVLGRPDELVAQQQQLITADLTSIIAQLISTAGSLLPSVKHTLSTTTPSIGQFGQLGGSFVPSGVGTDASVKLQNSNGSKLADQSNQTDLISNYGTEQNYSVEEHE
SKDEEDADEGENLPPGSYEILQLEKEEILAPHTHFCAICGKGFKRDANLRMHMRGHGDEYKTPAALAKPNKESSSEPTLIKRYSCPYAGCKRNKDHKKFQPLKTILCVKNHYKRTHCDK
SYTCSRCNTKKFSVIADLKTHEKHCGKDKWLCSCGTTFSRKDKLFGHIALFQGHTPAIPLDETKGTQGPADHGEGIQNLMDVKESINDPTSYFSPLNFETCNFDGFHEFPRPPFEDSESSFS
FLMPGSCNYTHKTGGEPNSNNLE

>P.persica_EMJ07676.1

MTEDPQVSSFSTSSGPENHSTSLLYSLSVLKEKVRQAKSLFSILISPDHHHHQSQPPESTSMAIASMNNVIQEIIVTASSMMFTCQQMALVSPPPPPPPPGNNNPTSTSNTIDHEEQLLHHQG
LNKPPPPRFSQPNFGSGLDNKIETPFYPTNDDQALDWFSDSYIDCTTRDTKTRHVQGKQQPISPKSFDIIELDAADLLAKYTHYCQVCGKGFKRDANLRMHMRAHGDEYKTTVALSNP
MKKNNSSTNNNNNHGAGGTDGNTDCSTNLPRKYSCPQVGCRWNQKHAKFQPLKSMICVKNHYKRSHCPKMYVCKRCNRKQFSVLSDLRTHEKHCGDLRWQCSCGTTFSRKDKLM
GHVALFTGHTPVVISSLASRLGKADQRSHGSQM

>M.domestica_MDP0000453148

MDPKERQWDTWENPSTGNDVTNTISSDHPSFTNFNSQQHQREWERPSVLDYEMRMEPSFLKFHQPSDSQMSYTCNSKNDTKIPDQEGGKMHEVQQPNKIQDWDARMTLNNLTFLEQ
KIHQLQDLVHVIVGRRGQVLGRPDELVAQQQQLITADLTSIIAQLISTAGSLLPSVKHTLSTTLPSTGQFGQLGGSFIPSAAGNDAGVKMQINSGSKLADQANQTDLISNYGTEHIEEHETK
DEEDADEGENLPPGSYEILQLEKEEILAPHTHFCAICGKGFKRDANLRMHMRGHGDEYKTAAALAKPNKESSSEPTLIKRYSCPYAGCKRNKDHKKFQPLKTILCVKNHYKRTHCDKSY
TCSRCNTKKFSVIADLKTHEKHCGIDKWLCSCGTTFSRKDKLFGHITLFQGHTPAIPLDETKGTLGPADHGEGSEASNRVGSISFSVSSTAPGGGGAAQSLMDVKESIDDPTSYFSPLNFET
CNFDGFQEFPRPPFEDTESSFSFLMPGSCNYTHKTGGGESNFNNLHRQ

>A.thaliana_AT1G34370

METEDDLCNTNWGSSSSKSREPGSSDCGNSTFAGFTSQQKWEDASILDYEMGVEPGLQESIQANVDFLQGVRAQAWDPRTMLSNLSFMEQKIHQLQDLVHLLVGRGGQLQGRQDELA
AQQQQLITTDLTSIIQLISTAGSLLPSVKHNMSTAPGPFTGQPGSAVFPYVREANNVASQSQNNNNCGAREFDLPKPVLVDEREGHVVEEHEMKDEDDVEEGENLPPGSYEILQLEKEEIL
APHTHFCTICGKGFKRDANLRMHMRGHGDEYKTAAALAKPNKESVPGSEPMLIKRYSCPFLGCKRNKEHKKFQPLKTILCVKNHYKRTHCDKSFTCSRCHTKKFSVIADLKTHEKHCG
KNKWLCSCGTTFSRKDKLFGHIALFQGHTPAIPLEETKPSASTSTQRGSSEGGNNNQGMVGFNLGSASNANQETTQPGMTDGRICFEESFSPMNFDTCNFGGFHEFPRLMFDDSESSFQ
MLIANACGFSPRNVGESVSDTSL

>A.thaliana_AT5G22890

MHIHMMNRDEHIAKKVEGSISSFSGETSTSSKQIYVNPVTTTGTKSMEDDDVSLSLLYNLSTLHEKVHQIQSLVSFYMVSTNNINQSSGSTSLAVANIGSLVQEIITAASSMLYTCQQLQIGS
NNNNNDIDNDQTVDAMVLEFSRQETDPGHDFVQESTNLFGVQERGQISFPDQNLDWYNTETINPKKDKHRSKPSSGSYDILELDVADLLAKYTHYCQICGKGFKRDANLRMHMRAHG
DEYKTREALISPTSQDKKGGYSLKKHYYSCPQHGCRWNQRHEKFQPLKSVICAKNHYKRSHCPKMYMCRRCSVKHFSVLSDLRTHEKHCGDIKWVCSCGTKFSRKDKLMSHVSLFLG
HVPAHGSSKPPTITLK

>C.rubella_Carubv10008876m

METEDDLCQQTNWGSSSSKLQEQGSSDGGGNQAFSGFTSEQKWEEASILDYEMDGVEPDLQVFDEMKASGEENVQANVDFLQGVRAQAWDPSTMLSNLSFMEQKIHELQELVHLLV
GRGEQVQGCQDELAAQQQQQLITTDLTSIIIQLISTAGSLLPSVKHNMSTAPGPYTGQPGSAMFPCAREANNVASQSQNNNNCGAQEFDLPKAVVVDERESHVVEEHEMKDEDDAEEGE
NLPPGSYEILQLEKEEILAPHTHFCTICGKGFKRDANLRMHMRGHGDEYKTAAALAKPNKESIPGSEPMLIKRYSCPFLGCKRNKEHRKFQPLKTILCVKNHYKRTHCDKSFTCSRCHTK
KFSVIADLKTHEKHCGKNKWLCSCGTTFSRKDKLFGHIALFQGHTPAIPIEETKPSASSSTQRENSEGGNNNQGMVGFNLGSPSNANQETAQFGMMDGNISFEDSFSPMNFETCSFGGFH
DFPRLMFDDSDSSFQTLIANTCGFSSPRNVGESVLDTSI

> C.rubella_ EOA22034.1

MMNQDDHLTSKGEGSMSSFSGTTATSSKQNYVNNTVTTAGVKSLEDDDASLSLLYNLSTLQDKVHQTQSLVSFYMISTNNINQSSGSTSLAVANIGSLVQEIITAASSMLFTCQQLNIGSN
NNNDNDQTADHAMVLEFSRQETEPGHNFVQESTNLFGVQERGQISFPDQNLEWFGTQTLNPKKEYHRSKQRSGNYEILELDVADLLAKYTHYCQICGKGFKRDANLRMHMRAHGDE
YKTREALISPISQDKKGDYSLKKHYYSCPQQGCRWNQRHEKFQPLKSVICAKNHYKRSHCPKMYMCRRCSVKHFSVLSDLRTHEKHCGDIKWVCSCGTRFSRKDKLMSHVSLFLGHV
PAHGSSAPTTTIKRS

>B.rapa_Bra036750

MEPEDDLCQNNWGGSSSSSSSKRREQVCFTSQHKWEDASILDYEMGMEEEPAFQENSNNNNGGQVNVDFLQGVRAQAWDPRTMLSNLSFMEEKIHELQDLVHLMVARNGQLQGRQE
QLVAQQQQLITTDLTSIIIQLISTAGSLLPSVKHHNMSTAPGPFTGSALFPYPREANNLASQTLNNNNNNTCEFDLPKPIVVEEEHEMKDEDDVEEGENLLPGSYEILQLEKEEILAPHTHFC
TICGKGFKRDANLRMHMRGHGDEYKTPAALAKPNKEAVPGSEPMLIKRYSCPFPGCKRNKDHKKFQPLKTILCVKNHYKRTHCDKSFTRSRCHTKKFSVIADLKTHEKHCGKNKWLC
SCGTTFSRKDKLFGHIALFQGHTPAIPLEETKPSAQKGSSACENSNNNNTGMVGFNLGSATNAIEEVAQPGFMDGKIRFEDSFSPLSFDTCNFGGFHEFPRPMFDDSESSFQMLISSACGFS
PRNGGESVSNTSL

>T.halophila_Thhalv10007411m

METEDDMCQNNWGSLSSSSKLEDGNQAFAGFTSQQKWEDAEILDYEMGIEKNLQAFDQMKPSSGQDNDNNNKNNNSTQANVDFLQGVKAQAWDPRTMLSNLSFMEQKIHELQDLV
HLIVGRNGLVAQQQQQLITTDLTSIIQLISTAGSLLPSVKHNMSSTAPGPYTGSAVFPCPREANNVASQSQNSNHQRGATQDFDLPKPIVVDERDSHVVEEHEMKDEDDAEEGVENLPPG
SYEILQLEKEEILAPHTHFCTICGKGFKRDANLRMHMRGHGDEYKTPAALAKPNKEAVPGTEPMLIKRYSCPFPGCKRNKDHKKFQALKTILCVKNHYKRTHCDKSFTCSRCHTKKFSV
IADLKTHEKHCGKNKWLCSCGTTFSRKDKLFGHIALFQGHTPAIPLEETKPSAQRGSSEGENNNTGMVGFNLGSATNADQETEQPGFMEEKISFEESFSPLSFDTCNFGGFHEFPRPMFD
DSESSFQMLISSACGFSPRNGGGGSVSNTIL

>C.papaya_evm.model.supercontig_66.56

MDTKERLCPDNWGNSSSSGNELYKKVSSDHRSFTGFGLQQNHQKWEDPSFLEYAMRIEPPFQEFNQPSQTQFSLCNQNDQMRILVQENGQENEMLESNKIQDWDPRATLNNLSFLEQK
IHQLQDLVHLIVGRRGQTLGHPDQLITQQQHLITADLTSIIVQLISTAGSLLPSARHSLSAANPSGAQIGQLGGVLFPHARGLHGEMQTQNNCGSKVSDSSTKIDPPTNYQVDENHIVEEHE
LKEEEDVDEGESLPPGSYEILQLEKDEILAPHTHFCTICGKGFKRDANLRMHMRGHGDEYKTAAALAKPVKESSSEPILIKRYSCPFPGCKRNKDHKKFQPLKTILCVKNHYKRTHCDKS
FLCSRCNTKKFSVIADLKTHEKHCGKDKWLCSCGTTFSRKDKLFGHIALFQGHTPAIPLDETKTSAGPSGRGENNEARYKVGSINFNFASNTPVLAVSTIMDVKGV

>C.sinensis_orange1.1g009831m.g

MDLKEGLRAETWAKSSLAENEMSKRISSNNPCFTDFNSQQYQQKWEDPSILDYDIRINPAFQEFNQPPQNQSSLPCDPTNQIKNPFQMNENLQTNKMQDWDPKAMLNNLSFLEQKIHEL
QDLVHLIVGRRSQILGRPDELVAQQQQLITADLTSIIVQLISTAGSLLPSMKHTLSAATPSMGQLGQFGGVPFPPGAGLSDSIQVQNSCAKKVSNQSNPIDLVGNSGTQQNHAVEEHDLKDE
DDADDGEQLPPGSYEILQLEKEEILAPHTHFCTICGKGFKRDANLRMHMRGHGDEYKTPAALAKPHKESSSEPMLIKRYSCPYAGCKRNKDHKKFQPLKTILCVKNHYKRTHCDKSYV
CSRCNTKKFSVIADLKTHEKHCGKDKWLCSCGTTFSRKDKLFGHIALFQGHTPAIPLDETIKGLAGPSDRREGNEATSKIGSTNFNFGSSVPNGTGVSDMDAKGNVDDPTNYFSPLNFDT
CNFDGFHEFPRPPFDDSENAFSFLIPGSCNYIQKTGGETSLNTLK

>E.grandis_Eucgr.F02627

MDLKMDLEERLRAETWGKPSSVNDLPQRVPPDRQTQFPNFASHKNLQKREDQDPSISNYGMRIEPSFSEFNRPSECQPPLPSNPISQDRGVQMNDILQLGKTQEWDPKAMLNNLSFLEQ
KIHQLQELVHAIVGRRGPVLGRPDELVAQQQQLITADLTSIIIQLISTAGSLLPSVKNSLSSASTPPIRQLGQLGGILNNSGSGIGLDSNLVLPSQGGSKVPDQSNQVDPMDQSAIDNLEDHES
KDDEDGDEGENLPPGSFEILQLEKEEILAPHTHFCTICGKGFKRDANLRMHMRGHGDEYKTPAALAKPHKEAGSEMMLIKRYSCPYAGCKRNKDHKKFQPLKTILCVKNHYKRTHCDK
SYTCSRCNTKKFSVIADLKTHEKHCGKDKWLCSCGTTFSRKDKLFGHITLFQGHTPAIPFDENKGGLSLQGEHNEDTNKVGNVSFSFGSSTPSSGGVQNIMEDVKGNVDDPSSFFSPLSF
EASNFGGFNEFTRSAFDDSEGAFSFLLQASCNYTQKNGGQSSSNNLE

>V.vinifera_GSVIVT01034505001

MDLKDRLSVDTWTQSSSSGNELPKDQQSFTNFSLQQQQHQHKWEDPSVLDYSVRIEPAFQEFNQTSATSSSLPCNPNSQIKNLNREDCQMNEMLEPNRIQDWDPRVMLSNLSFLEQKIH
QLQDLVHLIVGRRSQVLGRTNELVAQQQQLVTADLTSIIVQLITTAGTLLPSVKNTLSTASPSVGQLGQLGGVLFPSGTGMNGGGVAQSSGGSKVSDQSNQIDLTGACVIEQNNATEEHEL
KDEDDADEGENLPPGSYEILQLEKEEILAPHTHFCMICGKGFKRDANLRMHMRGHGDEYKTPAALAKPNKESSSEPVLIKRYSCPFAGCKRNKDHKKFQPLKTILCVKNHYKRTHCDK
SYTCSRCNTKKFSVIADLKTHEKHCGKDKWLCSCGTTFSRKDKLFGHIALFQGHTPAIPLDETKGSVGPSDRGEGNGAANKVGSVGFNFSSNASSGSGVQDMMMDVKRGADEPTGFF
SPLTFDPCSLVGFHEFPRPPFEDSESSFSFLVPGSCSYTRKTGGESSSNDLE

> V.vinifera_ XP_002272574.1

MIPVGTSNCFVNGSQGLQMYGMTAEGSVSSSLGGSSVLEAHSCSLVYSFSILKDKVHQVQSLISIFVPPNQGQPESMAMAVAGMGNLIQEVIAAASSMMYSCQQMGYGAAPGNSGTNH
GLPQQGVELSDGRVCGDTGVVQMGEERGQGFYSSDQSLDWYGDNHNNSNTNDHSRTIIVSNNDKVESRELPQGSTQMNEGLGGVLPKTFDIVELDAEDLLAKYTHYCQICGKGFKRD
ANLRMHMRAHGDEYKSNAALSNPTKNIGREMENKDDLIKLPRKYSCPQEGCRWNRKHAKFQPLKSMICVKNHYKRSHCPKMYICKRCNQKQFSVLSDLRTHEKHCGDLKWLCSCG
TTFSRKDKLMGHVALFVGHTPAINSMSKPPSVSLSKSWTDLLSPILAHELAQTWPMTGSKAARLELGQPGPFT

>M.guttatus_mgv1a007605m

MSDFQDEQIHGSHESNNMQDFDTKAMLNNLSFLEQKIHQLQEMVHLIVRQRGQADAILVQQQQLITADLTSVIVQLISTAGSLLPSLKTTLSSENPIGQFVANHHSAKVSDNNNNNNVVS
HKIEDHSDHEMMKSDEDADEGENLPPGSYEILQLEKEEILAPHTHFCTICGKGFKRDANLRMHMRGHGDEYKTPAALAKPNRESISDHPILVKRYSCPYIGCKRNKDHKKFQPLKTILCV
KNHYKRTHCDKNYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHIALFQGHTPAIPVEESKGFSGPPDRGQCSGVDQPEVNFKFDTPSTSACRNVIMDDDPTSYF
SPLDYEFPRSTFEDSESSFSFMLSESCDYLPKNGIYNGLNDME

>A.coerulea_Aquca_017_00778

MNPAVNLCSDTWAECSVRNKSSSEKFHSQQHHRKREDASSPDFTPPLSKFCKPSEPTSSHSGNLDNQRIIHDIVDGQAKDKSATAKTNEWNPNSILSNLSFVEQKIHQLQDLVHLMIGQD
KRIVNLSDELVAQQQQLVTADLTSIIVQLISTAGSLLPSVKSSHLVNPILGQLQSISGVGCATNLSVNNASGDDNNIKGETKLQCHLDQIEQVASWEREHKNAGKPIVSVSNVDEQIPKDEE
DAGDADNLPPGSYEVLQLEKEEILAPHTHFCLICGKGFKRDANLRMHMRGHGDEYKTPAALAKPNKEPGTEPMLAKRYSCPYTGCKRNKDHKKFQPLKTFLCVKNHYKRTHCDKSY
TCSRCNNKKFSVLADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPVEETKGACGTLNGEDSIEATNLENTGFNFSSGFSNESGVQNVIDMKGGIDDPAAFFSTLDFD
TCSFGGYEFPRPPFDISESSLSCLLSGSSDYVQNTGTESGSDAFRIEKWKESSLAEDF

>S.bicolor_Sb03g041170

MEGRMTSLEAAMKASSSMASSMARNADPDHQTLRSNSVEQFYFPRPGQSLPGIPPFFGPPSSSLYPPNDNEAKFGSQFESNPSQNTDWDPQAIVSNLTFLEQKIKQVKDIVQSMNNREN
QVAGGSSELAAKQQLITADLTSIIIQLISTAGSLLPSMKNPLLSSNPAVRQLGNTLGSPMGFGMNANQRPSVDSKTDIPDTGKTSDYDELMNSLNTAQDERDELIKCPNLCGGEGSEPTPME
DHDVKESDDGGEGENLPPGSYVVLQLEKEEILAPHTHFCLICGKGFKRDANLRMHMRGHGDEYKTPAALAKPTKDSGADHAPVTRYSCPFVGCKRNKEHKKFQPLKTILCVKNHYKR
SHCDKSYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMEDVKVSEASELPQDSEPTNEMARSMVYSFPSSSSDGISNLDMKMADDVRGYFS
PLNFDPCFGALDDFTRPGFDISENPFSFLPSGSCSYGQQNGDS

>S.bicolor_ XP_002444574.1

MIGRGGNPYYLQNQQQQLFHHGHALDATMDGGGFMAEPPTPASSGCSAADAQCHALLYYNLSVLRDKVQQLQPLVGLAVAHDGPGPVAAAPGAGAVIQEIIAAASSMMYAFQQLCG
HGGAVPASASATAAQAQQGGTSSGVVVADAAATCGAGDNHQHQQAAVIDHVMVMQQQWQQEHRYDGGYGGRIHHDSKTTTPVAAAAAMPSSSHPRPTTAAVMMAEEDEDVGVGV
AGGTIIELEATELLAKYTHYCQVCGKGFKRDANLRMHMRAHGDEYKSSAALANPAKAAAAAGGDAAAASTSSSRSLYSCPQEGCRWNRKHAKFQPLKSVICAKNHYKRSHCPKMYV
CNRCNRKHFSVLSDLRTHEKHCGDHRWLCSCGTSFSRKDKLVGHLALFTGHQPAVPLDRQANGGRRSSSSMSTSTQLDQTTNI

>Z.mays_GRMZM2G068710

MDNHIHSQPAANLYYQFGMGTEQQYPPPFASPPSTNIPHMDWNPATMLDNLTFIEEKIQQVKDVIRTMVDNAGQVQVQCHHQAGELLAQKQKQQVVNADLTCLIVQLISTAGSLLPSLK
NSSFLSHPHPAGHVDMANHVGPSSSLVPNAMTVSEDYEELFKGWTDGGIEVDDDGVLVEEQEITKDGPDSLSYELLQLEEDEILAPHTHFCTICGKGFKRDANLRMHMRGHGDEYKSPA
ALAKPPRDPGAEQEPAKRRYSCPFAGCKRNKLHKSFQPLKTILCVKNHYRRSHCEKSHTCGRCHAKKFSIVADLRTHEKHCGRRDRWVCSCGTSFSRKDKLFAHVALFQGHTPALSSPL
EEEPPKACSDQHQTGRAGSRREPAKVPAGSVGGGSGSSFVWGASSSGENSALDIKGFVDGCSGDFLSTANFGSFNLSFGPADGFTGEPSSGGSFSMLPPEHFQSAQKKGQN

>Z.mays_GRMZM2G075956

MEGRMTSLEATMKVSSSMASSISRNADPDQQTLRPNSVEQFYFPRPGQSLPGIPPFFGPPSSSLYLPNDNEAKFGNQFESNPSQNTDWDPQAIVSNLTFLEQKIKQVKDIVQSMSNRENQL
AGGSSELAAKQHLVTADLTSIIIQLISTAGSLLPSMKNPLLSSNPAVRQLGNTLGSPMGLGMNANQRPSVDSKTDIPDTGKTSDYDELMNSLNPAQDERDEMIKCPNPCGGEGSEPTPMED
HDVKESDDGGEGENLPPGSYVVLQLEKEEILAPHTHFCLICGKGFKRDANLRMHMRGHGDEYKTPAALAKPTKDSGADHAPVTRYSCPFVGCKRNKEHKKFQPLKTILCVKNHYKRS
HCDKSYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMEDVKVSEASEQPQDSEPMNEMARSNMYSFPCSSSDGISNLDMKMADDVRGYFS
PLNFDPCFGALDDFTRPGFDISENPFSFLPSGSCSYGQQNGDS

>S.italic_Si001061m

MEGGMSSLETAMKASSSIASGTARNAEPDQHPVCSNSLEQFYFPRPGQSLPGIPPFFGPPSSSLYLPNDNEAKVGNQFEPNPSQSTDWDPQAIVSNLTFLEQKIKQVKDVVQAMSNRENH
VAGGSCELAAKQQLITADLTSIIIQLITTAGSLLPSMKNPLSSNLAVRQLGNTLGSPMGFGMNTNQRPSVDSKTNIPDSGKVSDYEELINSLNTTQDERDEMIKCPNPCVGEGSEPTPMEDH
DVKESDDGEHEGENLPPGSYVVLQLEKEEILAPHTHFCLICGKGFKRDANLRMHMRGHGDEYKTAAALAKPTKDSGSEHAPVTRYSCPFVGCKRNKEHKKFQPLKTILCVKNHYKRS
HCDKSYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDVKISEASEQQQGSEPMNEISRSVGCFPCSSSDGISNLDMKMADDARGYYSPM
SFDPCFGTLDDFTRPGFDISEDPFSFLPSGCSYVQQNGDN

>S.italic_Si010000m.g

MDHSQNHIHGQSAANFCYQFGSDNPFLGMGVQQPFASFTSPFGATPSTNIPHMDWNPATMLDNLTFIEEKIRQVKDVIRTMVDDGGQRPGELAQQQQVVNADLTCLIVQLISTAGSLLPS
LKNSTFLSHPQAGHMGIVNHVGSSSSFVPNVTTISEENKEEMCGPEDYEELFKGFTDGALEGGIEIGNVLVEEQDAKDGDEGGDVGMDGESLPPGSYELLQLEKDEILAPHTHFCSICGK
GFKRDANLRMHMRGHGDEYKSPAALAKPPRDASAEHELVRRYSCPFVGCKRNKLHKNFQPLKTILCVKNHYKRSHCEKSYTCSRCHTKKFSVMADLKTHEKHCGRDKWLCSCGTSF
TRKDKLFAHVALFQGHTPALPTEEPKTSSDQISRVGSHQEPAKLPSSMGSSFVWGTSSGNENALDIKGVDGCSDDFLSTANFGSFNFSFGPADGFTGEPSGSSFSMLPSEHFQSAQKKGKN

>O.sativa_Os01g65080

MDSGLGRSSETSLKALPSMASNATRNTDPDQQGVRFSSMDQPPCFARPGQSFPAFPPLFGVQSSSLYLPDDIEAKIGNQFESNPSPNNPTMDWDPQAMLSNLSFLEQKIKQVKDIVQSMS
NRESQVAGGSSEAQAKQQLVTADLTCIIIQLISTAGSLLPSMKNPISSNPALRHLSNTLCAPMILGTNCNLRPSANDEATIPDISKTHDYEELMNSLNTTQAESDEMMNCQNPCGGEGSEPIP
MEDHDVKESDDGGERENLPPGSYVVLQLEKEEILAPHTHFCLICGKGFKRDANLRMHMRGHGDEYKTAAALAKPSKDSSLESAPVTRYSCPYVGCKRNKEHKKFQPLKTILCVKNHY
KRSHCDKSYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKVTGASEQPQGSEAMNTMVGSAGYNFPGSSSDDIPNLDMKMADDPRY
FSPLSFDPCFGGLDDFTRPGFDISENPFSFLPSGSCSFGQQNGDS

> O.sativa_Os08g0562300

MIPGGGGGGGGGGISPYLVQSQHGHGGGVDGMEMEEGGGFMGEQPQCHPLLYNLSVLKDRVQQLHPLVGLAVAHNAHAHGPLDVSAADAIIQEIVAAASSMMYAFQLLCDLGTAPTT
APSQETAAASAVVVKNNDHAADAGQMEDDHLMQQQWQQNGSRQHDYSSHAHAPPVFHSETAAPAGATSATDTIIELDAAELLAKYTHYCQVCGKGFKRDANLRMHMRAHGDEYK
SKAALSNPTKLLAKGGDETMAAAARKYSCPQEGCRWNRRHAKFQPLKSVICAKNHYKRSHCPKMYVCNRCGRKHFSVLSDLRTHEKHCGDHRWLCSCGTSFSRKDKLIGHVSLFAG
HQPVMPLDAPRAGKRQRSSSASVAGNIDDTTGIGMGAA

>B.distachyon_Bradi2g56320v

MESGMRRSSETSVKALSSMASEASRNTDPGQQGLRFNSTDQSYFARAGQSIPVFPPFFGPQSSSFCLPDGSGAKVGGQFEPNPSLNNPISDWDPQAMLSNLSFLEQKIKQVKDVVQSMS
NRGSQVVGGSGELAAKQQLVTADLTCIIIQLISTAGSMLPSMKSPLLSSNPAVRQLSNTLGSPMGFGSIANQRPSINKEQTIPDITKTSDYEELMNTINTTHDGKDDLIKCPNPCVGEGSDPI
PFEDHDVKESDDGGEGENLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMKDSSSDHTPVTRYSCPFVGCKRNKEHRKFQPLKTILCVKNH
YKRSHCDKSYTCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKGTCVSDQPEGSEVMDDMVGSTGYNFPGSASDGIPNLDMKVADDVR
GYFSPLSFDPCFGALDDFTRPGFDISENPFSFLPSGPGSCSFGQPSGDS

>B.distachyon_ XP_003571264.1

MFSAGDDSGASPYFQQLFDHGDLGGADGIGAFAGGGDAQCQALLYNLSVLREKVRLLHPLVGLAVHGRGGVAVAADAGAVVQEIVAAASSMMYAFQHLCAVSDAAMQAQDSVAAA
AGRAASNNAAGMAAAAAGCSDQQLQALEEDHEAATMQQWAHGGFYDDDGTSGSKPSAATAQQEAPAPAPGTKTRIIELDAAELLAKYTHYCKVCGKGFKRDANLRMHMRAHGDQ
YKSKAALSAVVSSSGASSSPAAMAASKYSCPQEGCRWNVRHARFTPLKSVICAKNHYRRSHCPKMYACSRCGRKQFSVLSDLRTHEKHCGDRRWLCSCGTTFSRKDKLAGHVSLFAG
HHPVVVGEGARQCKIDRSSLANSDQLLGNCTTSAGDGFSIT

>T.aestivum_genome_A

MKASSSMASDASGNTDPGQQGARFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDPFEPNPPQNNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDVGGSC
ELAAKQQLVTADLTSIIIQLISTAGSMLPSMKTPLLSSNPAVRQLNTPGSPMGFGSIVNQRPSTVREEMVPDITKTPDYEDLMNTLNPAHDEKDDLIKCPNPCVGEGPEPVPMEDHDVKES
DDGGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDVSVDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKNHYKRSHCDKRY
TCSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKATGASEQRSEAMDDMVGSTGYNFPGSTSDGIPNLDMKVADDTRGYFSPLNFDPCFG
ALDDFARPGFDISENPFSFLPSGPGSCSFGQLSGDS

>T.aestivum_genome_B

MKASSSMASDASGNTEPGQQGVRFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDPFEPNPPQSNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDVGGSCE
LAAKQQLVTADLTSIIIQLISTAGSMLPSMKTPLLSSNPAVRQLNTPGSPMGFGSIVNQRPSTVREEMVPDISKTSDYEELMNTLNTAHDEKDDLIKCPNPCVGEGPEPVPMEDHDVKESD
DGGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDSGSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKNHYKRSHCDKRYT
CSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKNKLFGHVALFQGHTPALPMDDIKATGASEQPQGSEAMDDMVGSTGYNFPGSTSDGIPNLDMKVADDTRGYFSPLNFDPCF
GALDDFARPGFDISENPFSFLPSGPGSCSFGQLSGDS

>T.aestivum_genome_D

MKASSSMASDASGNTEPGQQGVRFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDLFEPNPPQNNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDAGGSC
ELAAKQQLVTADLTSIIIQLISTAGSMLPSMKTPLLSGNPAVRQLNTPGSPMGFGSIVNQRPSTIREEVVNDISKTSDYEELMNTLNTAHDEKDDLIKCPNPCVGEGPEPVPMEDHDVKESD
DGGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDPGSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKNHYKRSHCDKRYT
CSRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKASGALEQPQGSEAMDDMVASTGYNFPGSTSDGIPNLDMKVADDTRGYFSPLNFDPCF
GALDDFARPGFDISENPFSFLPSGPGSCSFGQLSGDS

>S.lycopersicum_AK320912.1

MDPDDSLSEDPWIKSSSSGNELLKIMPSDNHSFTNLNLHAQKWEGSSYLDQQIRIEQQFSGFAKPKHTSEMDQQGNQRNENHDTTRIHDWDPRALLNNLSFLEQKIHQLQELVHLIVGR
RGQDEVQGNDLIVQQQQLITADLTSIIVQLISTAGSLLPTMKHTLSSAIPAASQLGQVGGVTVPSTAGTSAGGLTCNDGVAKLEDQSNHIDQLRDCGIEHNHAADEHELKDEDEAEEEENL
PPGSYEILQLEKEEILAPHTHFCTICGKGFKRDANLRMHMRGHGDEYKTPAALAKPHKEPSSEPTLIKRYSCPYVGCKRNKEHKKFQPLKTILCVKNHYKRTHCEKAYTCSRCNIKKFSV
IADLKTHEKHCGKDKWLCSCGTTFSRKDKLFGHLSLFQGHTPAVSPDETKGSAGTSDRGQTGEVTMKARQENYKVNASHGNEFQNPGVVKECPYNPSSYFSPLNFDTTNLNGFQEFPR
PPFDESDSSFSFLLSGSCEYPPHKAAKFMSSSEME

> S.lycopersicum_ XP_004250985.1

MIQGTSSSANNPQNIPLYSSNHHPNNNIQEDIFSSSSSSYQQNSFLFNLSLLKEKVHQVQSLATMFITPDNQTIIHPPPESISMIIANMGTLIQEIITTSSSLMFSCQKIVLDSTSLSQNSSRYREP
SQNDVGHGQGQGQVDHLLQDYDWYVDNYNSNCNTHEDNKNHVTSSSTIIASSTISHDNYYSKEFGKKELLLSTSKGKVVNEENNYDIVELDASDLLAKYTHYCQICGKGFKRDANLR
MHMRAHGDEYKSSAALSNPMKRINDLTSDGSLKSSSNTIKYSCPQEGCRWNKKHAKFQPLKSMVCVKNHYKRSHCPKMYVCKRCNKKSFSVLSDLRTHEKHCGDLKWQCSCGTTFS
RKDKLMGHVSLFVGHTPLIKQSAR

>H.vulgare_AK252406.1

MKASSSMAGDASGNTDPGQQGVRFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVGKACDPFEPNPPVNEPIADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQFGGASCE
LAAKQQLVTADLTSIIIQLISTAGSMLPSMKTPLLSSNPAVRQINTPGSTMGFGSTANQRPSATREEMIPDISKTSDYEELMNTLNTAHDEKDDQIKCPNPCGVEGSEPVPMEDHDVKESDD
GGEAEHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDSGSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKNHYKRSHCDKRYTC
SRCNTKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKATGASEQPQGSEAMDDMVGSTGYNFPGSTSDGVPNLDMKVADDTRGYFSPLNFDPCF
GALDDFARPGFDISENPFSFLPSGPGSCSFGQLSGDN

>M.acuminata_Achr5P01170_001

MLSNLSFLEQKIHQVQDIVRSIISQEGRANELAAQQQLVTADLTYIIIQLISTAGTLLPNIKNALLSATPPVAQTDHDIKDYDDGMDGENLPPGSYEVLQLEKEEILAPHTHFCLICGKGFKR
DANLRMHMRGHGDEYKTAAALAKPSKEASSEPAIIRRYSCPFVGCKRNKEHKKFQPLKTILCVKNHYKRSHCDKSYTCSRCKTKKFSVIADLKTHEKHCGRDKWICSCGTTFSRKDKL
FGHVALFQGHTPAIPMDDAKVSGMSDHVQVGETTNGMVANMGYNFSGNASDDAQSLDIKDVDDGQGYFSPMNFDACNFGGLDEFPRSAYDVSESPFSFLSTGTCTNIFLFACGGHYTS
FPKQKKKTIYDLLKGSSLFRSLPTLFIETCAFIN

>M.acuminata_Achr10P15970_001

MSSESRNLNQNQRATEFDARETEQSLPKFSSHLDNSFHQSNQAQRVAMNWDPRAVLNSLDILGQKIHQIQDIVRSTLSDERQLSIQPNEFAAQQQLINTDLTCTIIELISTAGTLLPLIKNAL
ASGINSSGQIDNDVKDNEDGVSGENLPLGSYEVLQLEKEEILAPHTHFCAICGKGFKRDANLRMHMRGHGDEYKSPSALAKPTKEGSSEPVPIKRYSCPFIGCKRNKEHKNFQPLKTILC
VKNHYKRSHCDKSYTCRRCNSKKFSVMADLKTHEKHCGSNKWLCSCGTTFSRKDKLFGHIALFQGHTPALSMDEVKSQGMSVQGQSDEVMAKEEDIDYLISGNVIEDTNFSGLNGTD
NDLGYFSSMNFDSFTFGGIDGLQQPSFDTSESLFSFHK

>M.acuminata_Achr6P18840_001

MLSNLSFLEQKIHQVQDIVRSIISHEDQFLDQANELAAQQQLVTADLTYIIIQLISTAGTLLPDIKNALLSTNPAVGRPGSNQKQHEILPVEVAKASQYDSLIKDLNSCGGKEDELIKCSNNS
GVEASEPIPIEDHDTKDNDDGMDGENLPPGSYEVLQLEEEEILAPHTHFCLICGKGFKRDANLRMHMRGHGDEYKSPAALAKPSKEASIEPQPLRRYSCPFVGCKRNKEHKKFQPLKTIL
CVKNHYKRSHCDKSYTCSRCRTKKFSIIADLKTHEKHCGRDKWICSCGTTFSRKDKLFGHVALFQGHAPAIPMDDAKVLVTSDKAQVSEASSGMPGNMGYNFSGNTSLDAQSPDIKGV
DDGHGCFSPMTLDACNIWGLAEFPHSAYEVSESTFSFLATGSCSSIQRNGENGFDISENPFSFLPSGPGSCSFGQLSGDN

>T.urartu_EMS66742.1

MASDASGNTDPGQQGARFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDPFEPNPPQNNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDVGGSCELAAKQ
QLVTADLTSIIIQLISTAGSMLPSMKTPLLSSNPAVRQLNTPGSPMGFGSIVNQRPSTVREEMVPDITKTPDYEDLMNTLNPAHDEKDDLIKCPNPCVGEGPEPVPMEDHDVKESDDGGEA
EHLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDSVSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKNHYKRSHCDKRYTCSRCN
TKKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKATGALEQPQGSEAMDDMVGSTGYNFPGSTSDGIPNLDMKVADDTRGYFSPLNFDPCFGALDD
FARPGFDISENPFSFLPSGPGSCSFGQLSGDS

>T.urartu_EMS47263.1

MFPDAGASPYFLQNQQQLFHGMNGIGVDAAFPGDGGGLAVPDSTHRSTLLYNLSVLKDKVQQLEPLVGHGHVVDPVVPGASAVVQEIIAAATSMLYALQHPYVLGALTTSGNHAATPE
GPTAVDTAIIELDAAELLAKYTHYCQVCGKGFKRDANLRMHMRAHGDEYKSKAALANPTTRLFATSGEDAAAGRPRNSKYSCPQDGCRWNRRHAKFQPLKSVICAKNHYKRSHCPK
MYVCNRCNRKHFSVLSDLRTHEKHCGDHRWICSCGTTFSRKDKLAGHVSLFAGHQPVAPLAPGSGRHGKRSSLLSPSSDDLAGNCTNTGFSITPTT

>Ae.Tauschii_EMT26835.1

MASDASGNTEPGQQGVRFSSMDQSCFARPGQSIPGYPPFFGPQSSNFYLPDDSVAKACDLFEPNPPQNNPVADWDPQAMLSNLTFLEQKIKQVKDIVQSMGNRGSQDAGGSCELAAKQ
QLVTADLTSIIIQLISTAGSMLPSMKTPLLSSNPAVRQLNTPGSPMGFGSIVNQRPSTIREEVVNDISKTSDYEELMNTLNTAHDEKDDLIKCPNPCVGEGPEPVPMEDHDVKESDDGGEAE
HLPPGSYVVLQLEKEEILAPHTHFCVICGKGFKRDANLRMHMRGHGDEYKTPAALAKPMRDSGSDPTPVTRYSCPYVGCKRNKEHRKFQPLKTILCVKNHYKRSHCDKRYTCSRCNT
KKFSVIADLKTHEKHCGRDKWLCSCGTTFSRKDKLFGHVALFQGHTPALPMDDIKASGALEQPQGSEAMDDMVASTGYNFPGSTSDGIPNLDMKVADDTRGYFSPLNFDPCFGALDDF
ARPGFDISENPFSFLPSGPGSCSFGQLSGDS

>A.lyriata_XP_002874102.1

MNRDEHITNKGEGSMSSFSGTTSTSSKQISLNPVTAAGAKSLEEDDVSLSLLYNLSTLQDKVHQIQSLVSFYMISSNNINQYSGSTSLAVANIGSLVQEIITAASSMLYTCQQLHIGSNNSND
IDNDHTVDAMVLEFSRQETDPGHDFVQESTNLFGVQERGHVSFPNQNHDWYDTETLNPKKDKHRSKPKPGNYEILELDVADLLAKYTHYCQICGKGFKRDANLRMHMRAHGDEYKT
REALISPTSQEKKGEYTLKKHYYSCPHQGCRWNQRHEKFQPLKSVICAKNHYKRSHCPKMYMCRRCSVKHFSVLSDLRTHEKHCGDIKWVCSCGTKFSRKDKLMSHVSLFLGHVPAH
GSSQPPTTITLK

>F.Vesca_XP_004294539.1

MNINQGSTSNCFESMYMTDRDAVSSFSLEPNSSSSGAVLSQDQNHSTSLLYSLSVLKEKVHQAKSLVGVLIAPADHHQSQQQQLPESTTMGAMASMSNVIQEIIVTASSMMFTCQQMSLV
APGNDTRSEPNFTISTAGRSNTRDAGFYPTHDQTLDWFSTESYNDCLPRDVISRSHRQVETKELPPRVQRNNALQVENPPPENSSYDIIELDAADLLAKYTHYCQVCGKGFKRDANLRM
HMRAHGDEYKTSIALSNPLKNNNMNSAMNVNGSSNAESSMKLPRKYSCPQEGCRWNQKHAKFQPLKSMICVKNHYKRSHCPKMYVCKRCNRKQFSVLSDLRTHEKHCGDLRWQCS
CGTTFSRKDKLMGHVALFVGHTPVIINSNFPIRHGRAESQAAAGGLMQ

>F.Vesca_XP_004290434.1

MDHKDKVCADTWAKDLRNKTCSDRPKFANFSAHQHQNKWEDPSILDYGIRIEPSFQKLSQPSEDQTSLPHNSNNEKTIADGEDVQMNEIFHASKIQDWDPSTMLTNLSFLEQKIHQLQD
LVHLIVGRRGQVLGRPDELVAQQQQLITADLTSIIAQLISTAGSLLPSVKHTLSSASASAVQFGQLGGSFVSSGAGTEASVKLQMNCGSKLPEQPNQTDPISNYGTELNYHIEEHESKDEED
ADEGENLPPGTYEILQLEKEEILAPHTHFCAICGKGFKRDANLRMHMRGHGDEYKTPAALAKPNKDSSSEPTLIKRYSCPYAGCKRNKDHKKFQPLKTILCVKNHYKRTHCDKSYTCSR
CNTKKFSVIADLKTHEKHCGKDKWLCSCGTTFSRKDKLFGHIALFQGHTPAIPLDETKTAVGASEHGEGSEAPNRVGSINFNFGSTLPSAGGLVQNIMDVKESVDDPTCYFSPLNFDTCN
FDGFHEFPRPPFEDSESSFSFLMPGSCNYTHKTGGESNSKQVE

Appendix- IV

Figure: Multiple alignments of the homoeologues of TaSTOP1 in bread wheat

Appendix-V

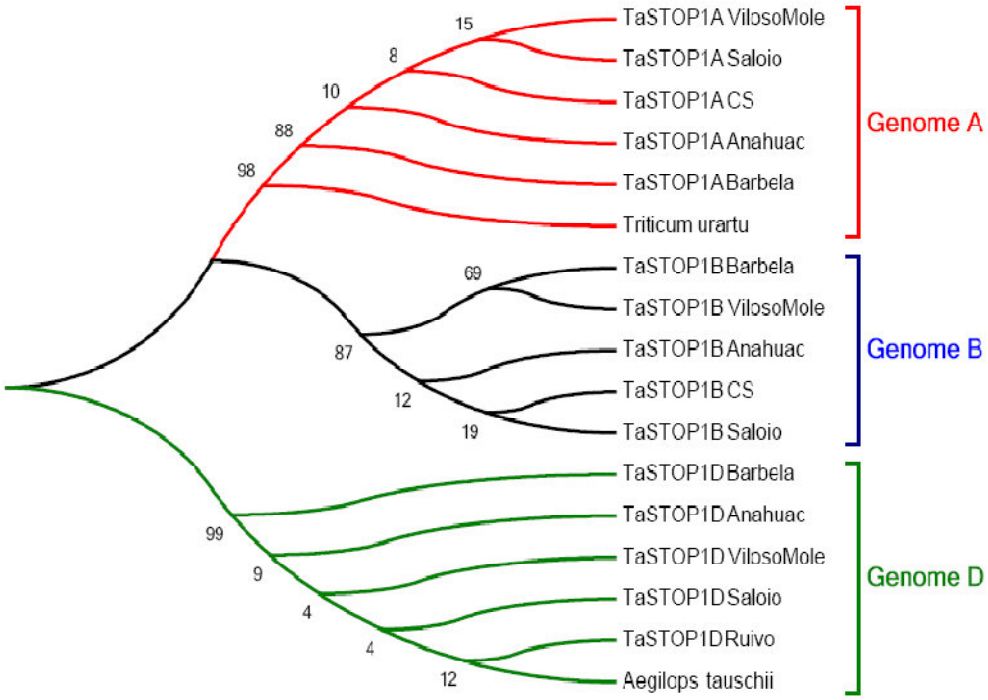


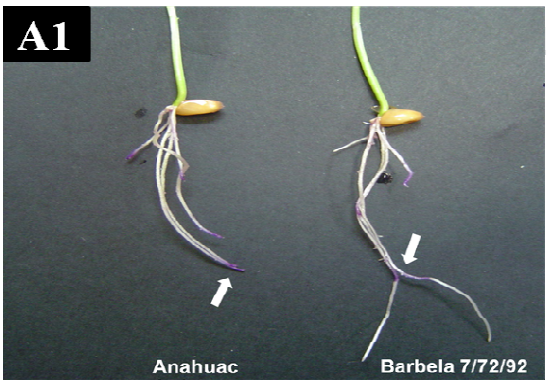
Figure: Phylogenetic analysis of *TaSTOP1* homoeologues genes in different species of wheat including some bread wheat genotypes.

Appendix-VI

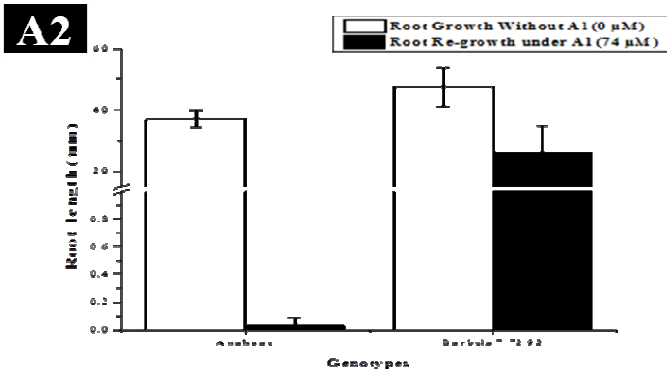
```
5UTR_TaSTOP1A_Barbela      CCCCAATCCCCTCCACCACCCACGCGCGTGGGCCGGCCGGCCACCGCCGC
5UTR_TaSTOP1A_Anahuac      CCCCAATCCCCTCCACCACCCACGCGCGTGGGCCGGCCGGCCACCGCCGC
5UTR_TaSTOP1D_Barbela      CCCCAATCCCCTCCACCTCCCACGCGCGT----CGGCCGGCCACCGCCGT
5UTR_TaSTOP1D_Anahuac      CCCCAATCCCCTCCACCTCCCACGCGCGT----CGGCCGGCCACCGCCGT
5UTR_TaSTOP1B_Barbela      -CCCCAACCCCTCCACCTCCCACGCGCGTCGGCCGGCCGGCCACCGCCGT
5UTR_TaSTOP1B_Anahuac      -CCCCAACCCCTCTACCTCCCACGCGCGTCGGCCGGCCGGCCACCGCCGT
                            ***.*:****** ***:***********    ****************

5UTR_TaSTOP1A_Barbela      CCCTGCTCTTCCTCGCCGGGTCCCCGCCGTTTCCACTCCGTCCGGCAATC
5UTR_TaSTOP1A_Anahuac      CCCTGCTCTTCCTCGCCGGGTCCCCGCCGTTTCCACTCCGTCCGGCAATC
5UTR_TaSTOP1D_Barbela      CCCTGTTCTTCCTCGCCGGGTCCCCGCCGTTTCCACTCCGTCCGGCAATC
5UTR_TaSTOP1D_Anahuac      CCCTGTTCTTCCTCGCCGGGTCCCCGCCGTTTCCACTCCGTCCGGCAATC
5UTR_TaSTOP1B_Barbela      CCCTGTTCTTCCTCACCGGGTCCCCGCCGTTTCCACTCCGTCCGGCAATC
5UTR_TaSTOP1B_Anahuac      CCCTGTTCTTCCTCACCGGGTCCCCGCCGTTTCCACTCCGTCCGGCAATC
                           ***** ********.**********************************

5UTR_TaSTOP1A_Barbela      ATTGGCCATTGAGGGTTTGATGCTTTTCCTGTCCTATTATCAATTTCTCG
5UTR_TaSTOP1A_Anahuac      ATTGGCCATTGAGGGTTTGATGCTTTTCCTGTCCTATTATCAATTTCTCG
5UTR_TaSTOP1D_Barbela      ATTGGCCATTGAGGGTTTGATGCTTTTCCTGTCTTATTATCAATTTCTCA
5UTR_TaSTOP1D_Anahuac      ATTGGCCATTGAGGGTTTGATGCTTTTCCTGTCTTATTATCAATTTCTCA
5UTR_TaSTOP1B_Barbela      ATTGGCCATTGAGGGTTTGATGCTATTCCTGTCTTATTGTCAATTTCTCA
5UTR_TaSTOP1B_Anahuac      ATTGGCCATTGAGGGTTTGATGCTATTCCTGTCTTATTGTCAATTTCTCA
                           ************************:******** ****.**********.

5UTR_TaSTOP1A_Barbela      TTTGGGGAAGATCGTGCGAGATAGAAGAAAGGGCGGACTGGATTGTAATC
5UTR_TaSTOP1A_Anahuac      TTTGGGGAAGATCGTGCGAGATAGAAGAAAGGGCGGACTGGATTGTAATC
5UTR_TaSTOP1D_Barbela      TTTGGGGAAGATCGTGTGAGATAGAAGAGAGGGCGGACTGGATTGTAATC
5UTR_TaSTOP1D_Anahuac      TTTGGGGAAGATCGTGTGAGATAGAAGAGAGGGCGGACTGGATTGTAATC
5UTR_TaSTOP1B_Barbela      TTTGGGGAGGATCGTGTGAGATAGAAGAGAGGGCGGACTGGATTGTAATC
5UTR_TaSTOP1B_Anahuac      TTTGGGGAGGATCGTGTGAGATAGAAGAGAGGGCGGACTGGATTGTAATC
                           ********.******* ***********.**********************

5UTR_TaSTOP1A_Barbela      TGAGCTCTGGTGGATCTAGACTGGAAGTTGCATGAGAAAAAGTTCAGAAA
5UTR_TaSTOP1A_Anahuac      TGAGCTCTGGTGGATCTAGACTGGAAGTTGCATGAGAAAAAGTTCAGAAA
5UTR_TaSTOP1D_Barbela      TGAGGTCTAGTGGATCTAAACTGAAAGTTGCATGAGAAAAAGTTCAGAAA
5UTR_TaSTOP1D_Anahuac      TGAGGTCTAGTGGATCTAAACTGAAAGTTGCATGAGAAAAAGTTCAGAAA
5UTR_TaSTOP1B_Barbela      TGAGCTCTGGTGGATCTAGACTGAAAGTTGCATGAGAAAAAGTTCAGAAA
5UTR_TaSTOP1B_Anahuac      TGAGCTCTGGTGGATCTAGACTGAAAGTTGCATGAGAAAAAGTTCAGAAA
                           **** ***.*********.****.**************************

5UTR_TaSTOP1A_Barbela      TTTCCATGAAAGCTTCGTCGTCG
5UTR_TaSTOP1A_Anahuac      TTTCCATGAAAGCTTCGTCGTCG
5UTR_TaSTOP1D_Barbela      TTTCCATGAAAGCTTCGTCGTCG
5UTR_TaSTOP1D_Anahuac      TTTCCATGAAAGCTTCGTCGTCG
5UTR_TaSTOP1B_Barbela      TTTCCATGAAAGCTTCGTCGTCG
5UTR_TaSTOP1B_Anahuac      TTTCCATGAAAGCTTCGTCGTCG
                           ***********************
```

Alignment of the *TaSTOP1* homoeologues genes 5' UTR regions in two bread wheat genotypes (Barbela 7/72/92 and Anahuac). The coding sequence of each gene is presented in red colour letters. The presence of 3-AF1 binding site (light responsive element) and pyrimidine-rich stretch within the 5' untranslated regions of *TaSTOP1* homoeologues genes is highlighted in yellow and blue colour, respectively. Positions that are conserved in *TaSTOP1* homoeologues genes of both genotypes are indicated by asterisks (*).

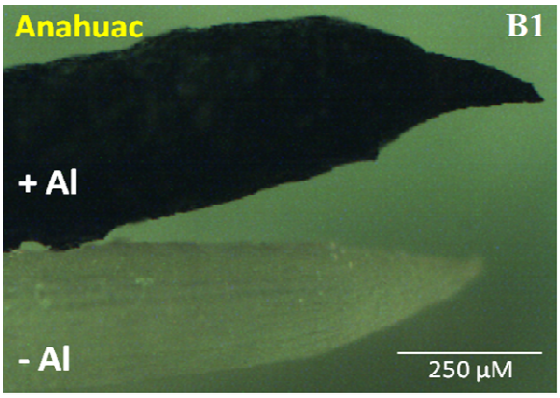
Appendix-VII



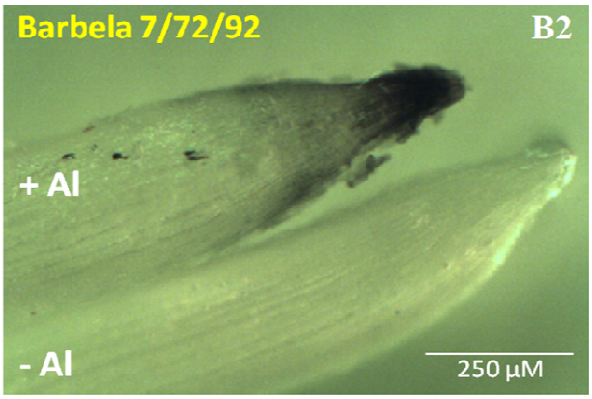
Eriochrome cyanine R assay



Root re-growth measurement



Hematoxylin assay



Visual detection of Al accumulation

Figure: Physiological characterization of roots of two bread wheat genotypes Anahuac and Barbela 7/72/92 for Al resistance using eriochrome cyanine R (A) and hematoxylin (B) assay. In picture A1, arrow indicates the eriochrome cyanine dye as a marker and the root after this arrow shows the re-growth whereas in figure A2 the root growth under low pH (pH 4.0) without Al (control) stress and root re-growth after exposure of Al (74 µM). In figure B, dark colour in roots represents the accumulation of Al [(B1, Anahuac) and (B2, Barbela 7/72/92). The upper and lower root samples in B1 and B2 were grown with and without (Control) Al.

Conclusions

Our understanding of Al toxicity mechanisms in plants has been progressing with the fast development of technology in plant science and intense research on Al tolerance. Although, Al toxicity mechanisms seem to be complex, as different plant species and/or genotypes within the same species are showing different Al tolerance mechanisms. At a broader level, these mechanisms in plants could be classified into external (exclusion of Al from root apex) and internal (tolerance of Al in the roots as well as shoots symplasm) detoxification of Al. Among all mechanisms implicated in Al toxicity, organic acids (OAs) exudates from roots is one of the best studied mechanisms for Al exclusion and it seems to play an important role to cope with Al toxicity under acidic conditions. In the past, several researchers reported that organic acids regulate the cell's pH and it is responsible for Al chelation, therefore, protecting the roots from Al toxicity. Among the organic acid anions, the role of citrate, malate and oxalate has been well documented in relation to the Al detoxification in different plant species. Generally, the type of organic acids secreted from the root tissues under Al stress differ among the plant species. In cereal crops, previous reports refer that malate is secreted in bread wheat whereas citrate is released in rye, sorghum and barley. Moreover, the level of a particular organic acid differs between the genotypes of a particular species, and also with the type of organic acid exudate, i.e., some Al tolerant bread wheat genotypes can exudate either malate or citrate.

At the molecular level, Al toxicity mechanism is extremely complex and still is not fully understood. In cereals, numerous genetic studies clearly revealed that Al tolerance is governed by multiple genes that cause minor to moderate level of phenotypic effects, making the identification of these genes difficult. A thorough understanding of both the genetics and physiology of Al tolerance played a pivotal role in the identification of the major loci such as *ALMT1* in wheat, and *MATE1* in barley and sorghum implicated in Al tolerance. Like other abiotic stresses, there are also strong evidences that regulatory genes (transcriptional factors) play important role in Al tolerance mechanism. Transcription factors are considered as master regulators in complex traits where they coordinate the expression of many target genes in entire response networks.

In order to better understand the Al tolerance mechanisms in bread wheat, we focused mainly on two diverse bread wheat genotypes, Barbela 7/72/92 and Anahuac, which were previously reported as tolerant and sensitive under Al stress, respectively. Barbela 7/72/92 was derived from an old Portuguese bread wheat landrace Barbela which has been cultivated by the farmers for more than one century due to its stable grain and straw yield under acidic soils. Besides Barbela 7/72/92 and Anahuac, some additional bread wheat genotypes were also used in parts of the work when requested.

In chapter 2, we described the effect of Al toxicity on bread wheat roots (in two diverse genotypes Barbela 7/72/92 and Anahuac) using various physiological parameters. From histochemical assays performed in the present research, it was clearly demonstrated that toxic concentration of Al induces multiple stresses in the roots of bread wheat, such as, lipid peroxidation, loss of plasma membrane integrity and physical injuries in the form of cell ruptures and death. The results obtained from hematoxylin and morin assays advocated that Barbela 7/72/92 is an Al resistant genotype and external detoxification of Al might be one of the major mechanisms of Al tolerance in this genotype. The results obtained from Schiff's reagent and Evan blue procedures were also in good agreement with hematoxylin and morin assays which detected a low level of Al induced lipid peroxidation and rare physical injuries in the Barbela 7/72/92 roots in comparison to the Anahuac roots when exposed to Al. Furthermore, better root hair growth in Barbela 7/72/92 under Al stress provided an additional proof that this genotype has developed several mechanisms of Al tolerance to adapt in acidic soils and probably operating in a combined way to cope with Al toxicity. Our histochemical assays clearly demonstrated that bread wheat genotype Barbela 7/72/92 is highly tolerant to Al toxicity and that it could be used to further elucidate the complex mechanisms of Al toxicity as well as to be included in plant breeding programmes aiming to develop improved wheat cultivars for acidic soils. In addition, our results indicated the advantage of histological techniques for Al-tolerance assessment as these assays detect the actual localization of Al in the root tissues and the sites of Al induced stresses, such as Al-enhanced *in situ* peroxidation of lipids on root surface with higher sensitivity than the biochemical procedures currently applied.

In chapters 3 and 4, we addressed the cloning and molecular characterization of two members of the MATE family for Al tolerance in bread wheat. Chapter 3 also describes the upstream variations in citrate and malate transporter genes for

comprehensive understanding of Al tolerance mechanism in bread wheat genotype Barbela 7/72/92. We illustrated that *TaMATE1* belongs to the multidrug transporter protein family. Chinese Spring nullitetrasomic and ditelosomic lines confirmed that *TaMATE1* genes are located on the long arms of the homoeologous group 4 chromosomes in bread wheat. A very high basal level of *TaMATE1* transcript expression was observed particularly on genome B (*TaMATE-4B*) in Al tolerant genotype Barbela 7/72/92 in comparison to the sensitive genotype Anahuac. Noticeably, *TaMATE1* presented a quite stable expression under Al stress condition in both genotypes suggesting that *TaMATE1* is implicated in the constitutive citrate efflux from roots in bread wheat. Besides *TaMATE1*, a higher basal transcript expression of *TaALMT1* gene in Barbela 7/72/92 indicated that it possesses novel alleles for both *TaMATE1* and *TaALMT1* genes. It is interesting to note that Al resistant genotype Barbela 7/72/92 has a Sukkula like transposon in the upstream of *TaMATE1-4B* homoeologue and also contained type VI promoter in the *TaALMT1* gene. Type VI promoter of *TaALMT1* had previously been reported strong positive correlation with Al tolerance in bread wheat. Similar to type VI promoter in *TaALMT1*, in the present work, a positive association of Sukkula like transposon in the upstream of *TaMATE1-4B* was associated to Al tolerance in bread wheat using a set of selected Al tolerant and sensitive genotypes of the Portuguese bread wheat collection. These results suggested that Barbela 7/72/92 has different Al tolerance mechanisms and both *TaMATE1* and *TaALMT1* are operating simultaneously to cope with Al toxicity in this particular genotype.

As described in chapter 4, we also cloned another member of the MATE family, named *TaMATE2* which is located on the long arms of homoeologous group 1 chromosomes in bread wheat. *TaMATE2* genes have a high similarity with *ScFRDL2* from rye and also showed typical secondary structure of MATE-type transporters, maintaining all the 12 transmembrane domains. Multiple nucleotide alignments of the *TaMATE2* homoeologues ORF (open reading frame) revealed very high similarity among these genes. Measurement of the collective transcript expression of *TaMATE2* homoeologues exhibited approximately similar level of *TaMATE2* transcript in both root apices and shoot tissues of Al resistant genotype Barbela 7/72/92; whereas Al sensitive genotype Anahuac showed a higher expression level in the shoot tissues in comparison to the root apices. Similar to *TaMATE1*, time-course analysis of *TaMATE2*

expression also revealed its constitutive level in the root apices of both genotypes of bread wheat under Al stress. Contrarily, Al responsive transcript expression of *TaMATE2* in the shoot tissues of Al resistant genotype Barbela 7/72/92 indicated that *TaMATE2* may play a major role in the shoot tissues. In addition, the higher transcript levels of *TaMATE2* in shoot tissues herein detected and its co-localisation with a previously reported minor QTL for Al tolerance located on homoeologous chromosomes 1A and 1B, suggests that *TaMATE2* is also a novel candidate gene for Al tolerance that may be implicated in internal detoxification of Al in bread wheat.

Chapter 5 reports on the cloning of *TaSTOP1* gene in bread wheat which contains four potential zinc finger domains and belongs to the zinc finger (Cys2His2 type) transcription factor family. PCR based mapping of *TaSTOP1* using Chinese Spring wheat aneuploid stocks revealed that *TaSTOP1* is located on the long arms of wheat homoeologous group 3 chromosomes. These results were also confirmed by *in situ* hybridization, which is the most direct method for physical localisation of genes in chromosomes, being used for the gene copy number identification in species with highly complex genomes such as bread wheat. Similar to the aneuploids results, *in situ* hybridisation also confirmed that *TaSTOP1* is localised on the long arms of homoeologous group 3 chromosomes and that it might be a single copy gene in bread wheat genotype Barbela 7/72/92. To the best of our knowledge, we were able for the first time to localise a gene with a probe size smaller than 2.0 kb demonstrating that the FISH technique can be used to simultaneously anchor homoeologous chromosomes with 1.5 kb probes even in bread wheat. We showed that *TaSTOP1* homoeologues have similar genomic structures. Expression profiling identified biased transcript expression of *TaSTOP1* homoeologues and seemed to differential respond to Al and proton (H^+) stress which indicated that *TaSTOP1* homoeologues may differentially contribute to Al or proton (H^+) stress in bread wheat. Furthermore, our results also indicated that *TaSTOP1-A* transactivation potential is constitutive and may not depend on the presence/absence of Al at least in yeast. Taken together, the localisation of *TaSTOP1* on long arm of homoeologous group 3 chromosomes in the present investigation and also previously reported major loci implicated with Al tolerance in bread wheat on the long arm of 3B chromosome, identified through QTL and genome wide association mapping studies, suggest that *TaSTOP1* could be a potential candidate gene for Al tolerance in bread wheat.

Follow the experiments elaborated in chapter 2 to 5; we conclude that Barbela 7/72/92 is a highly Al-resistant genotype of bread wheat which possesses a battery of novel genes that might operate together to cope with Al toxicity.

Future Perspectives:

Globally, wheat is one of the leading crops which is directly or indirectly contributing for the food supply of nearly half of the worlds' population. Additionally, the hexaploid nature of wheat makes it even more interesting for geneticist because it can serve as a relevant model for other polyploid species. Among cereals, wheat is considered as one of the sensitive crops to Al stress. However, conventional breeding for Al tolerance in wheat has been shown to be successful which led to expand its cultivation onto large areas of acid soils throughout the world, such as happened in Portugal and Brazil. To feed the burgeoning human population, a substantial rate of genetic gain and efficiency of conventional breeding systems need to be considerably improved, particularly to become more predictive. Molecular biologists are on the brink of identifying a number of genes that could play a key role in Al tolerance; identification of candidate genes will open up new avenues of examination into the molecular and genetic basis for Al tolerance, as well as provide new molecular resources for further improvements in crop Al tolerance through marker-assisted breeding. Keeping all these facts in view, following point should be considered:

1) Physiological study clearly demonstrated that Al induced lipid peroxidation and disruption of tissue integrity is extremely low in Al tolerant genotype Barbela 7/72/92. Thus, molecular mechanisms governing such traits should be addressed in future research.

2) Whether root hairs play any role in Al tolerance mechanism in bread wheat particularly, in genotype Barbela 7/72/92?

3) What is the actual function of *TaMATE2* in bread wheat? Is *TaMATE2* involved in internal detoxification of Al in bread wheat?

4) Does *TaSTOP1* interact with the malate and citrate transporter genes in bread wheat especially *TaALMT1* and *TaMATE1*?

5) The roles of H^+ and Mn toxicity as well as nutrient deficiencies associated with the Al toxicity in acid soil is well known. Thus, the possible function of these candidate genes in cation transportation needs to be addressed.

The extensive research during the last two or three decades has contributed enormously in our understanding of the Al toxicity mechanisms in plants. However we are just beginning to understand the molecular basis of Al tolerance mechanism in major food crops including bread wheat. Finally, the result obtained from the present investigation has raised as many questions for subsequent inquiry as it has answered.