Structural and functional characterization of natural alleles of potato (*Solanum tuberosum* L.) starch phosphorylases associated with tuber quality traits

Inaugural Dissertation

zur

Erlangung des Doktorgrades

Der Mathematisch-Naturwissenchaftlichen Fakultät

Der Universität zu Köln

Vorgelegt von

Anna Camila Nader Nieto

aus Bogotá, Kolumbien

Köln, 28.06.2011

Die vorliegende Arbeit wurde am Max-Planck-Institut für Züchtungsforschung in der Abteilung für Pflanzenzüchtung und Genetik (Direktor Prof. Dr. Maarten Koornneef) angefertigt.

Berichterstatter: PD. Dr. Christiane Gebhardt

Prof. Dr. Martin Hülskamp

Tag der mündlichen Prüfung: 30.06.2011

To my family Demetrio, María Luisa and María Fernanda whom I owe the happiest moments of my life, I dedicate this thesis.

-Nana-

TABLE OF CONTENTS

ABBREVIATIONS.....IV

FIGURES.....V

TABLES.....VI

INTRODUCTION

1.1 POTATO GENETICS.....1

1.2 STARCH METABOLISM IN TUBERS.....3

1.3 GENETIC BASIS OF COLD SWEETENING.....6

1.4 ALLELIC VARIATION ASSOCIATED WITH TUBER QUALITY TRAITS.....7

1.5 STARCH PHOSPHORYLASES.....8

1.5.1 STARCH PHOSPHORYLASE PLASTIDIAL ISOFORMS *PHO1A* AND *PHO1B*.....9

1.6 OBJECTIVES.....11

MATERIALS AND METHODS

2.1 MATERIALS.....12

2.1.1 EQUIPMENT AND CHEMICALS.....12

2.1.2 ENZYMES.....12

2.1.3 ANTIBIOTICS.....13

2.1.4 VECTORS.....13

2.1.5 OLIGONUCLEOTIDES.....13

2.1.6 BACTERIA.....16

2.1.7 PLANT MATERIAL.....16

2.1.8 BUFFERS AND CULTURE MEDIA.....17

2.1.9 SOFTWARE AND DATABASES.....17

2.2 METHODS.....18

2.2.1 GROWTH CONDITIONS.....18

2.2.2 NUCLEIC ACIDS ISOLATION.....18

2.2.3 POLYMERASE CHAIN REACTION (PCR).....19

2.2.4 BAC-Library screening 19
2.2.5 cDNA sSynthesis 20
2.2.6 RT-qPCR 20
2.2.7 Pyrosequencing 20
2.2.8 Cloning of full length *Pho1* cDNA for allele mining 21
2.2.9 Preparation of competent bacteria cells and transformation 21
2.2.10 Sequencing 22
2.2.11 Allele mining 22
2.2.12 *In silico* protein modeling and simulation of point mutations 23
2.2.13 Native protein extraction 23
2.2.14 Native PAGE 23
2.2.15 Phosphorylase activity test 23
2.2.16 2-D SDS electrophoresis and mass spectrometry 23

RESULTS

3.1 THE STARCH PHOSPHORYLASE LOCUS ON CHROMOSOME III 25
3.1.1 Allele mining on the *Pho1a* locus uncovered the allele accountable for the *Stp28-8b* SSCP polymorphism associated with better tuber quality traits 25
3.1.2 *Pho1a-H_A is a*ssociated with quality traits 32
3.1.2.1 Pho1a-H_A is associated with higher tuber starch content and tuber bruising 32
3.1.2.2 Pho1a-H_A is associated with lower reducing sugar contents upon cold storage 34
3.1.3 *Pho1a* promoter analysis 36
3.1.4 Expression analysis of *Pho1a* and *Pho1a-H_A* specific expression 40
3.1.4.1 Pho1a expression profile upon cold storage 40
3.1.4.2 Expression analysis of the Pho1a-H_A allele upon cold storage 42
3.1.5 Functional characterization of *Pho1a* alleles 44
3.1.5.1 The PHO1a-H_A protein isoform is present in the tuber proteome. 49
3.1.6 Starch phosphorylase activity upon H_A allele dosage 52

3.2 THE STARCH PHOSPHORYLASE *PHO1B* 54
3.2.1 *Pho1b* allele mining revealed high diversity in this locus. 54

DISCUSSION

4.1 LIMITS OF ASSOCIATION MAPPING FOR THE *PHO1A* AND *PHO1B* LOCI AND THE NEED FOR VALIDATION OF CANDIDATE GENES 59

4.2 IDENTIFICATION OF NATURAL ALLELES OF *PHO1A* ASSOCIATED WITH TUBER QUALITY TRAITS 60

4.2.1 THE *PHO1A-H_A* ALLELE IS ASSOCIATED WITH HIGHER STARCH CONTENT AND BETTER CHIP QUALITY AFTER COLD STORAGE 62

4.3 *PHO1A-H_A* INFLUENCES DIFFERENT TUBER QUALITY TRAITS 62

4.4 GENOTYPE-DEPENDENT *PHO1A* EXPRESSION AND REGULATION UPON COLD STORAGE 64

4.5 ALLELE SPECIFIC TRANSCRIPTIONAL ANALYSIS SHOWS THAT A *PHO1A-H^A* IS NOT DIFFERENTIALLY REGULATED 65

4.6 THE *PHO1A-H_A* ALLELE AFFECTS PHO1A ENZYME ACTIVITY 65

4.7 AMINO ACID SUBSTITUTIONS IN THE *PHO1A-H_A* ALLELE MAY AFFECT DIMERIZATION 66

4.8 A PUTATIVE ROLE FOR PHO1A IN COLD SWEETENING 68

4.9 DIVERSITY OF THE *PHO1B* LOCUS IN CHROMOSOME V 69

CONTRIBTIONS OF CO-WORKERS TO THIS PROJECT 71

LITERATURE 72

ABSTRACT 82

ZUSAMENFASSUNG 84

APPENDIXES 86

ACKNOWLEDGEMENTS 96

ERKLÄRUNG 98

RESUME 99

Abbreviations

A	Adenine
AFLP	Amplified Fragment Length Polymorphism acknowledgements
AtPhs1	*Arabidopsis thaliana* starch phosphorylase gene
BAC	Bacterial artificial chromosome
Bio	biotin
bp	Base pair
C	Cytosine
cDNA	Complementary DNA
cM	centimorgans
cv	Cultivar
E. coli	*Escherichia coli*
EST	expressed sequence tag
Fru	Fructose
G	Guanine
GF	Glucose and fructose
Glc1P / G1P	Glucose-1-phosphate
Glu	Glucose
GP	Glycogen phosphorylase
Indel	Insertion / deletion
Kb	Kilo base pair
LD	Linkage disequilibrium
MALP	Maltodextrin phosphorylase
ORF	Open reading frame
PHO	Starch phosphorylase
Pho1a	Potato starch phosphorylase gene on chromosome III
Pho1b	Potato starch phosphorylase gene on chromosome V
qRT-PCR	quantitative reverse transcription polymerase chain reaction
QTL	Quantitative trait locus
RFLP	restriction fragment length polymorphism
rRNA	Ribosomal RNA
RS	Reducing sugars
S. tuberosum	*Solanum tuberosum*
SNP	Single-nucleotide polymorphism
SSCP	Single strand conformational polymorphism
SSR	Single sequence repeat
Sta4	*Chlamydomonas reinhardtii* starch phosphorylase gene
Stp23	Potato starch phosphorylase on chromosome III
StpL/Stp1	Potato starch phosphorylase on chromosome V
T	Tymidine
YGP	Yeast glycogen phosphorylase
WCS	Weeks of cold storage

Figures

Figure 1. Simplified scheme of sucrose to starch conversion in storage cells. 4
Figure 2. Starch degradation in storage cells. 5
Figure 3. Phenotypes of starch phosphorylase mutants in different species. 11
Figure 4. BLAST results of *Pho1a* cDNA against the most recent draft of the potato genome (v3) 28
Figure 5. Glucose and fructose (GF) content of potato 34
Figure 6. Cold induced accumulation of GF in potato tubers after cold storage. 35
Figure 7. Sugar content means in 40 potato cultivars according to the presence of the H_A haplotype 36
Figure 8. Distribution of the DPO1 PCR fragment in the *Pho1a* promoter. 37
Figure 9. Distribution of the *Pho1a* promoter polymorphisms *DP01* and *DP02B* in the BIOSOL population. 39
Figure 10. Distribution of the *Pho1a* promoter polymorphism DP02B in the BIOSOL population 39
Figure 11.Dosage scoring for the $A_{824}G$ SNP (green arrow) in cv 26 (nuliplex for H_A) and 17 (duplex).. 41
Figure 12. expression of *Pho1a* in five low-sugars (dark blue) and five high-sugars potato cultivars, respectively 41
Figure 13. Two SNPs of the *Pho1a*-H_A allele used to determine allele dosage in the individuals chosen for pyrosequencing assays. 42
Figure 14. Differential expression of the *Pho1A*-H_A allele during cold storage 43
Figure 15. MUSCLE alignment of plant PHO1a protein sequences. 45
Figure 16. Comparison between protein sequences of glucan phosphorylases from approx. 200 different species 46
Figure 17. *In silico* mutation of the Ala56 residue in rabbit muscle glycogen phosphorylase 48
Figure 18. Lys-275-Arg amino acid substitution in PHO1a 48
Figure 19. Amino acid substitutions introduced by the *Pho1a*-H_A allele 49
Figure 20. 2D-SDS Electrophoresis of total protein extract from the potato cultivar 17 of the BIOSOL population.50
Figure 21. Peptide Mass Fingerprinting (PMS) of PHO1a peptides.. 51
Figure 22. PHO1a activity in four genotypes bearing the *Pho1a*-H_A allele in different dosages. 53
Figure 23. Starch phosphorylase activity in 40 genotypes (BIOSOL population) 53
Figure 24. *Pho1b* expression is only detected in tuber but not inleaf tissue in cDNA from four potato genotypes 55
Figure 25. Overview of the location of SNPs along the *Pho1b* sequence. 58
Figure 26. Structure of the bacterial maltodextrin phosphorylase (MALP) dimer complex 67

Tables

TABLE 1. MAJOR FEATURES OF POTATO STARCH PHOSPHORYLASES 9
TABLE 2. ANTIBIOTICS 13
TABLE 3. LIST OF OLIGONUCLEOTIDES 14
TABLE 4. BACTERIAL STRAINS 16
TABLE 5. *PHO1A* SSCP MARKERS ASSOCIATED WITH CHIP QUALITY TRAITS (LI ET AL., 2008) 25
TABLE 6. NUMBER OF FULL LENGTH CLONES OBTAINED FROM FOUR TETRAPLOID POTATO CULTIVARS 26
TABLE 7. SUMMARY OF *PHO1A* CDNA ALLELES CLONED FROM FOUR TETRAPLOID POTATO CULTIVARS. 27
TABLE 8. COSEGREGATION OF *PHO1A* CDNA SNPS AND SSCP *STP* MARKERS FROM LI ET AL. (2008) 30
TABLE 9. CDNA SNPS OF *PHO1A* IN COMMON WITH THE SOLCAP DATABASE 31
TABLE 10. ASSOCIATION OF THE *PHO1A-C* SSCP MARKER (CORRESPONDING TO THE *PHO1A-H_A* HAPLOTYPE) WITH TUBER QUALITY TRAITS IN THE TASK POPULATION (URBANY ET AL., 2011) 33
TABLE 11. SNPS IN THE *PHO1A* GENE UNDERLYING BOTH THE *PHO1A-C* (URBANY ET AL., 2011) AND *STP23-8B* (LI ET AL., 2008) SSCP POLYMORPHISMS (FOR COMPLETE SEQUENCE SEE APPENDIX 5). 33
TABLE 12. DNA POLYMORPHISMS IN THE *PHO1A* PROMOTER. 38
TABLE 13. PEPTIDES OF DIFFERENT MASSES BETWEEN PHO1A-H_R AND PHO1A-H_A ISOFORMS AS OBTAINED BY *IN SILICO* TRYPSIN DIGESTION 50
TABLE 14. STRONGEST PEPTIDE PEAKS OF THE PHO1A MASS SPECTRA 52
TABLE 15.*PHO1B* SSCP MARKERS ASSOCIATED WITH CHIP QUALITY TRAITS (LI ET AL., 2008). 54
TABLE 16. DISTRIBUTION OF THE *PHO1B* SSCP MARKERS IN FOUR TETRAPLOID POTATO GENOTYPES 55
TABLE 17. NON-SYNONYMOUS SNPS OF ALLELES OF THE *PHO1B* GENE PRESENT IN 4 TETRAPLOID POTATO GENOTYPES 57

1 Introduction

The origin of cultivated potatoes (Solanaceae sect. *Petota)* can be traced to a single domestication event originating from *Solanum bukasovii* (Spooner et al., 2005) near Lake Titicaca in the southern Peruvian highlands. The world largest collection of potato germplasm held at the International Potato Center (CIP) in Peru includes about 3,800 samples of Andean cultivated potatoes and 1,500 wild relatives. About 4,000 edible varieties have been historically cultivated by farmers, being spread from southern USA to Argentina (Hijmans and Spooner, 2001). *Solanum tuberosum*, the major cultivated potato, is divided into two subspecies: *andigena*, adapted to tropical and temperate regions and grown mainly in the Andes, and *tuberosum*, a subsequent adaptation to long day conditions. The latter is today the main cultivated potato worldwide. Potatoes were introduced into Europe in the 16th century and rapidly became a staple food for a fast growing population. Nowadays, potatoes are grown globally being the number three food crop worldwide after wheat, and rice (FAO Crops statistics database).

1.1 Potato genetics

Cultivated potatoes as well as wild relatives range in ploidy levels from diploid to pentaploid (Hijmans and Spooner, 2001). *Solanum tuberosum* is an autotetraploid with a basic number of 12 chromosomes present in four copies (2n=4x=48). It displays a complex inheritance pattern derived from the multivalent pairing of homologous chromosomes during meiosis (Bradshaw, 1994; Gavrilenko, 2007), where up to four different alleles can freely recombine and segregate. The inheritance scenario includes the possibility of two sister chromatids ending up in the same gamete as a result of double-reduction (Bradshaw, 1994). Therefore, potatoes are highly heterozygous individuals. Inbred lines are usually not available due to pronounced inbreeding depression which prevents the generation of homozygous inbred lines. Lines with genetic potential for breeding purposes can be easily maintained via clonal propagation.

Classic investigation of gene function is often based on the usage of recessive mutants that are analyzed in homozygous state. In potato, the ploidy, heterozygocity, and inbreeding depression make it difficult to recover recessive genotypes at the tetraploid level and therefore to infer the genotype from the phenotype. Lack of dominant morphological markers

and the fact that multiple homologous chromosomes mask the recessive haplotypes, impeded the construction of potato linkage maps based on morphological characters (Gebhardt, 2007).

Genetic linkage maps were constructed taking advantage of the possibility to generate diploid lines either by interspecific crosses that induced the parthenogenetic development of female gametes or by anther culture (Dunwell and Sunderlan, 1973; Powell and Uhrig, 1987). Due to gametophytic incompatibility, diploid parents are highly self-incompatible, but crosses between two diploid lines generate F1 segregating offspring. Several reference molecular maps were constructed for all potato chromosomes using isozymes, RFLPs , SSRs and AFLPs markers among others (Bonierbale et al., 1988; Gebhardt et al., 1989; Gebhardt et al., 1991; Bonierbale et al., 1992; Tanksley et al., 1992; Jacobs et al., 1995; van Eck et al., 1995; Leister et al., 1996; Milbourne et al., 1998; Gebhardt and Valkonen, 2001; Gebhardt et al., 2003). Moreover, an ultra-dense genetic map based on 10,000 AFLP markers was constructed (van Os et al., 2006). The genetic molecular map of potato is one of the most elaborate maps in crop species.

At the tetraploid level, however, genetic linkage studies are less frequently reported. Molecular markers that are fully informative in diploid species fail to be so in polyploids, as one allele can be present in more than one copy. A number of statistical models for the analysis of genetic linkages of molecular markers were developed for polyploid species. Such models, however, failed to integrate the critical notions of tetrasomic inheritance, considering only bivalent pairing of homologous chromosomes (Cao et al., 2005; Cervantes-Flores et al., 2008). Recently, new models were developed for two-locus (Luo et al., 2006) and multi-locus (Leach et al., 2010) linkage analyses in autotetraploids that take into consideration both, allele segregation and recombination, under tetrasomic inheritance.

Further construction of molecular-function maps allowed the identification of functional candidate genes co-localizing with quantitative trait loci (QTLs) for key traits such as pathogen resistance (Gebhardt and Valkonen, 2001) and carbohydrate metabolism and transport (Schäfer-Pregl et al., 1998; Chen et al., 2001). A candidate gene-based QTL map for cold-induced sugar accumulation identified a total of 24 QTLs distributed along the 12 potato chromosomes contributing to the phenotypic variance, confirming that cold sweetening is a truly quantitative trait (Menendez, 2002).

The outcrossing mating habits and genetic heterogeneity observed in potato resemble the genetic inheritance patterns in humans. Consequently, approaches that are used in human

genetics can be used to uncover the genetic bases of complex traits in potato, with the advantage that each cross produces a high number of siblings that can be used to evaluate linkages. One of those approaches is association mapping, which takes advantage of historical meiotic recombinations of individuals related by descent to detect the extent to which two loci are linked and how certain loci relate to phenotypic traits (Balding, 2006; Morton, 2005).

Recently, genome wide association mapping for the detection of QTLs was achieved at the tetraploid level. The authors observed genome-wide linkage disequilibrium (LD) up to approximately 5 cM using r^2 higher than 0.1 as a criterion for significant LD. Nevertheless, within-group LD, increased on average to 10 cM when compared to overall LD (D'Hoop et al., 2010).

Amplicon sequencing allows an approximate estimation of allelic dosages in a given genotype by comparing the height ratio of the base calls in a polymorphic position (4:0, 3:1, 2:2, 1:3, 0:4). From the dosages haplotype blocks can be inferred by computational approaches (Neigenfind et al., 2008), however, this method cannot be used in high throuput approaches yet, as manual curation of dosage scoring is still required. There is a high degree of natural variation at the DNA level in tetraploid potatoes (Gebhardt et al., 1989) and SNPs can occur in very high frequencies (Rickert et al., 2003; Draffehn et al., 2010). This makes it difficult to accurately phase SNPs into large haplotype blocks. Similar to human genetics, haplotyping could potentiate the power of genome-wide association studies to find candidate genes underlying complex traits (Simko, 2004; Fan et al., 2011; Kitzman et al., 2011).

1.2 Starch metabolism in tubers

Starch metabolism is a complex process consisting mainly of a mosaic of nuclear encoded genes of eukaryotic and cyanobacterial origin (Deschamps et al., 2008a). The different components of starch metabolism have been extensively studied not only for their evolutionary implications but also because of its economic importance. For instance in potato, starch and sugar metabolism are major determinants of quality traits.

Carbon fixed during the day time trough the Calvin cycle is transported from green tissues to specialized heterotrophic tissues (such as the tubers) via the phloem in the form of sucrose. In the cytosol sucrose is cleaved to form hexose-phosphates which are imported into the plastids and further processed until their final conversion into starch, which accumulates

during tuber development (reviewed in Neuhaus and Emes, 2000 and Lytovchenko et al., 2007). The synthesis of starch in heterotrophic cells is illustrated in Figure 1.

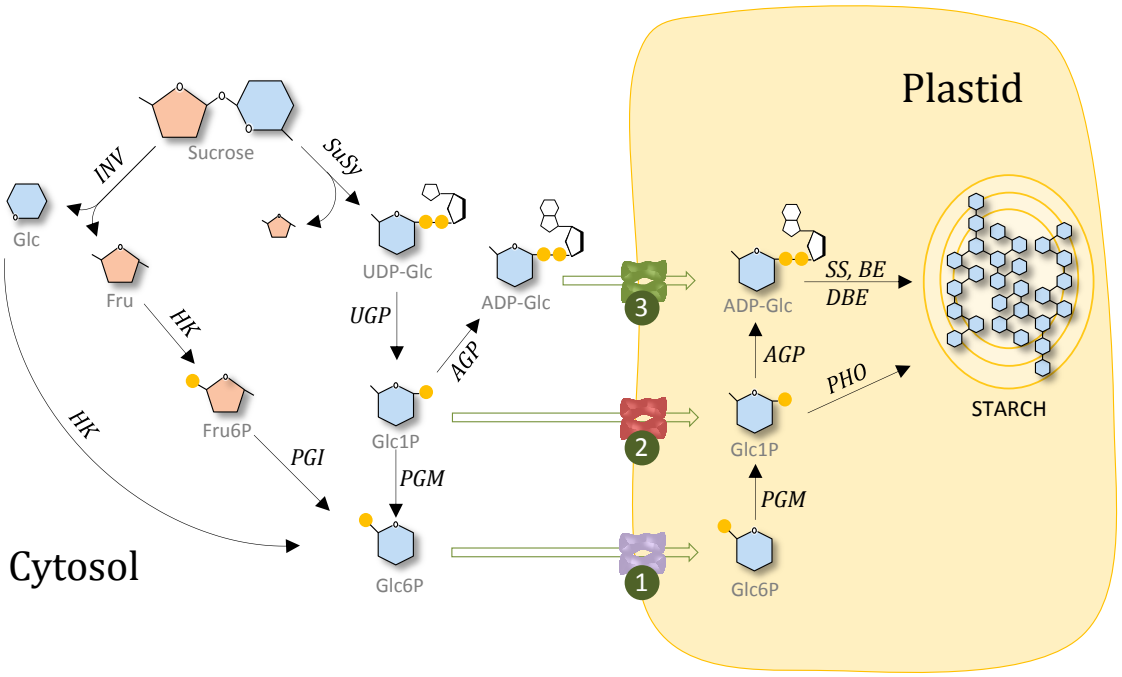


Figure 1. Simplified scheme of sucrose to starch conversion in storage cells. Sucrose molecules that are imported into storage cells are cleaved by sucrose synthase (**SuSy**) to produce nucleotide-sugars (UDP-Glc and ADP-Glc), which are further converted into sugar phosphates (Glc1P and Glc6P) by the action of UDP-glucose pyrophosphorylase (**UGP**), cytosolic phosphoglucomutase (**PGM**) and ADP-glucose pyrophosphorylase (**AGP**). Alternatively, invertases (**INV**) can hydrolyze sucrose to produce glucose and fructose molecules that are converted into sugar phosphates by hexokinases (**HK**) and phosphogluco-isomerase (**PGI**). Sugar-phosphates and nucleotide sugars are imported into the plastids through membrane transporters: **(1)** Glc6P-transporter (GPT); **(2)** Glc1P transporter; **(3)** ADP-Glc transporter. In the plastid, starch synthesis takes place. Hexose molecules are integrated into the starch granule by the action of starch synthase (**SS**), branching enzyme (**BE**), debranching enzyme (pullulanase and isoamylase, **DBE**), and alternatively by starch phosphorylase (**PHO**). This scheme has been adapted from Lytovchenko et al. (2007)

When tubers reach maturity, the function of carbohydrate metabolism is to remobilize starch to provide metabolites that support maintenance respiration. The process of the interconversion of starch into sugars starts by the relaxation of the tightly packed insoluble starch granules by the action of water-glucan dikinase (GWD) and phospho-glucan water dikinase (PWD), exposing the glucan chains. Debranching enzymes linearize the glucan chains which are further hydrolyzed by β-amylases and exported to the cytoplasm trough maltose transporters (MEX1), or phosphorolytically cleaved by starch phosphorylases (PHO)

to release glucose-1-phosphate (Glc1P) residues. Phosphoglucomutase (PGM) converts Glc1P into glucose-6-phosphate (Glc6P), which is then exported to the cytoplasm through Glc6P transporters (Kammerer et al., 1998). Once in the cytoplasm, hexose-phosphates undergo further processing to the final production of sucrose, and the reducing sugars glucose and fructose (Zeeman et al., 2004a; Zeeman et al., 2007; Deschamps et al., 2008b; Jeon et al., 2010). The starch degradation pathway is illustrated in Figure 2.

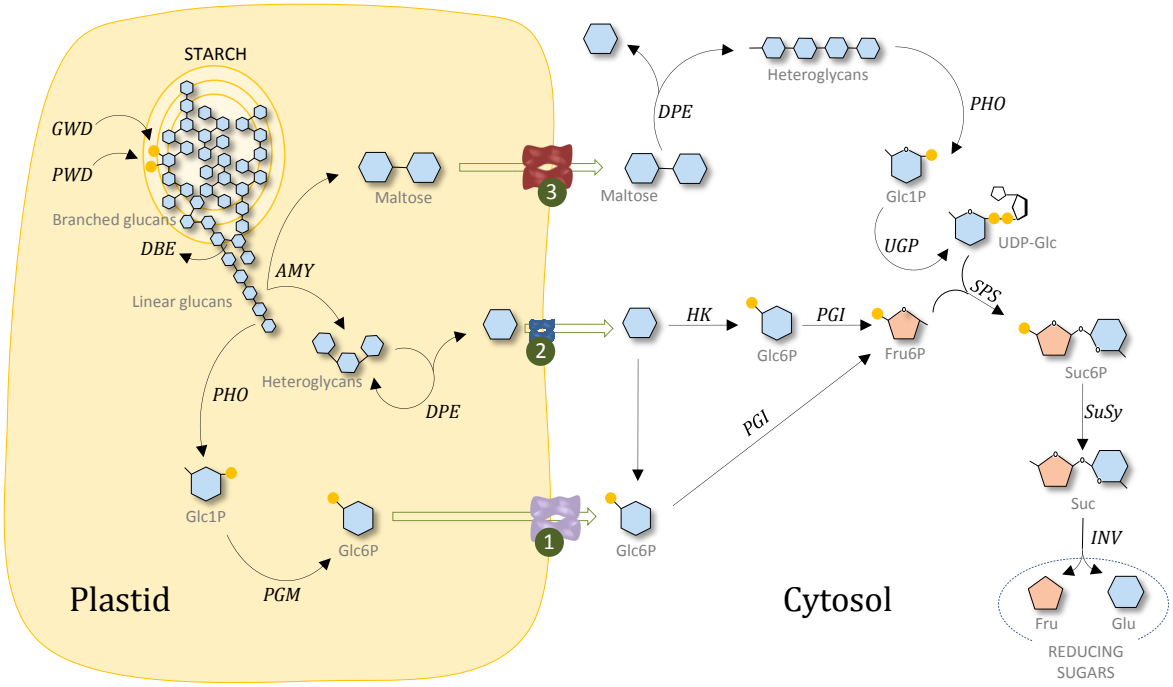


Figure 2. Starch degradation in storage cells. **GWD**, Glucan-water dikinase; **PWD**, phosphoglucan-water dikinase; **DBE**, debranching enzyme; **PHO**, starch phosphorylase; **PGM**, phosphoglucomutase; **βAMY**, β-amylase; **DPE**, disproportionating enzyme; **HK**, hexokinase; **PGI**, phosphoglucose isomerase; **SPS**, sucrose phosphate synthase; **SuSy**, sucrose synthase; **UGP**, UDP-glucose pyrophosphorylase. Membrane transporters: **(1)** ADP-Glc transporter; **(2)** Glucose transporter; **(3)** Maltose transporter (MEX1). Scheme adapted from Lu and Sharkey (2006) and Lytovchenko et al. (2007). Details described in text.

The final products of the starch degradation pathway are reducing sugars that feed different metabolic pathways. The accumulation of reducing sugars in tubers has been extensively studied given their direct impact in quality traits, which is discussed in the subsequent section.

1.3 Genetic basis of cold sweetening

Many plants respond to exposure to low temperatures by accumulating reducing sugars (RS) at the expense of storage starch (Müller-Thurgau, 1882). This physiological process known as cold sweetening (Burton, 1969) has a negative impact on potato processing quality and is one of the most challenging problems in potato breeding. Soluble sugars accumulated during cold storage can originate from respiration (Gounaris, 2001; Sowokinos, 2001), impaired starch synthesis, starch breakdown, glycolysis, or hexogenesis (Sowokinos, 2001).

Frequently, after harvesting potatoes, tubers are stored at temperatures below 10°C to prevent sprouting and the proliferation of diseases. Consequently, enzymatic metabolism of starch is altered leading to the accumulation of reducing sugars (RS, mainly glucose and fructose) in the tuber, that increase proportionally to the storage temperature (Malone et al., 2006). The amount of reducing sugars in potato tubers is directly correlated with after-cooking darkening (Shallenberger et al., 1959). In the presence of heat, RS react in a non-enzymatic Maillard reaction with free amino acids in the cytosol. This is of particular relevance for potato chips production, where, as a result of the Maillard reaction, unacceptably dark-colored and bitter chips are produced. Moreover, the amount of RS is correlated with the concentration of acrylamide in heat-processed tubers, constituting a potential risk for human health (Mottram et al., 2002; Stadler et al., 2002).

Tuber quality indicators (starch content, chip color, sugar content, bruising index), just as most traits relevant for potato crop production, are complex traits controlled by natural variation at several loci and environmental factors. Studies on the heritability of chip color after cold storage have revealed that this is a highly heritable trait, controlled by both additive and non-additive effects (Cunningham and Stevenson, 1963; Loiselle et al., 1990; Pereira et al., 1993; Jakuczun and Zimnoch-Guzowska, 2004). Wild germplasm has been surveyed for resistance to cold sweetening and resistance from wild species has been introgressed into tetraploid potato (Hamernik et al., 2009; Luthra et al., 2009; McCann et al., 2010; Bhardwaj et al., 2011). Several attempts have been made to alter the expression of single genes involved in the carbohydrate metabolism, thought to play a crucial role in cold induced sugar accumulation in potato. A decrease in the amount of reducing sugars after cold storage has been achieved by transcriptional inhibition of several genes of the carbohydrate metabolism including invertases (Zrenner et al., 1996; Greiner et al., 1999; Bhaskar et al., 2010; Ye et al., 2010) and invertase inhibitors (Greiner et al., 1999), sucrose phosphatase

(Chen et al., 2008), water-glucan dikinase (Lorberth et al., 1998), sucrose-phosphate synthase (Krause et al., 1998), plastidial starch phosphorylase (Yan et al., 2006), and multigenic silencing of polyphenol oxidase, water-glucan dikinase, and starch phosphorylase (Rommens et al., 2006). In contrast, inhibition of disproportionating (debranching) enzyme (Lloyd et al., 2004) did not show any effect on cold sweetening. Antisense inhibition of the cytosolic starch phosphorylase affected sprouting but had no significant impact on carbohydrate metabolism (Duwenig et al., 1997). Inhibition of UDP-glucose pyrophosphorylase resulted in an increase of sucrose that was not translated into the accumulation of hexoses (Spychalla et al., 1994; Borovkov et al., 1996). Transgenic lines expressing the 6-phosphofructokinase (PFK) gene from *Lactobacillus bulgaricus* under the control of the patatin promoter produced 60% less sugars than non-transgenic plants after cold storage, although the amount of sugars before storage was higher than in the controls in most cases (Navratil et al., 2007). A potential regulation of invertases by invertase inhibitors has also been proposed (Liu et al., 2010). Futhermore, Allelic variants of UGPase showed differences in kinetic properties (Gupta et al., 2008), being one of the few allelic functional studies that discuss genetic natural variation in the context of cold sweetening (Draffehn, unpublished).

A heterologous microarray experiment provided a comprehensive analysis of early events associated with cold sweetening giving evidence to the participation of redox and hormone-signalling networks in addition to sugars metabolism (Bagnaresi et al., 2008).

Taken together, potato carbohydrate metabolism and in particular cold sweetening is a field that attracts extensive scientific attention, due to the paramount importance of these mechanisms for potato yield and quality.

1.4 Allelic variation associated with tuber quality traits

Tuber quality, just as most characteristics relevant to the potato crop, is controlled by natural variation at several loci and environmental factors. Recently, association mapping of candidate genes involved in carbohydrate metabolism provided valuable information for the understanding of complex traits such as starch content and cold induced sweetening (Li et al., 2005; Li et al., 2008; Urbany et al., 2011) Two loci encoding starch phosphorylase genes in potato (*Pho1a* and *Pho1b*) co-localized with QTLs for starch content and chip quality (Schäfer-Pregl et al., 1998; Chen et al., 2001). Meeting both, the functional and positional criteria, these loci were selected and allelic variants were tested for association with the

traits of interest in a population of 240 tetraploid breeding clones (Li et al., 2008). Allelic polymorphisms at the two loci were found to be associated with chip quality after cold storage.

1.5 Starch phosphorylases

Starch (1-4-alpha glucan) phosphorylases (PHOs, EC 2.4.1.1) are highly conserved proteins among living organisms. They catalyze the reversible removal of a terminal glucose moiety at the non-reducing end of a glucan molecule, thereby playing an important role in the metabolism of reserve polysaccharides. The discovery of glycogen phosphorylase dates back to 1936 (Cori and Cori, 1936). It was, 20 years later, the first enzyme shown to undergo covalent (phosphorylation) and allosteric regulation (Krebs and Fischer, 1956; Helmreich and Cori, 1964). Specific PHOs are named after the natural substrate they act on. Accordingly, glycogen-, maltodextrin- and starch- phosphorylases are the respective isozymes present in animal/yeast, bacterial and plant tissues.

In contrast to animal and yeast glycogen phosphorylases, plant starch phosphorylases are always present in their active form, devoid of either allosterical or covalent regulation. They are not affected by AMP nor modified by phosphorylase kinase from rabbit muscle (Fukui et al., 1982). The same applies for the bacterial isozyme, although transcriptional regulation of *malP*, the operon encoding maltodextrin phosphorylase in *Escherichia coli,* has been described (Debarbouille et al., 1982; Decker et al., 1998).

Two types of starch phosphorylases have been identified in plants. They differ in sub-cellular compartmentalization, molecular size and substrate affinity. A cytosolic isoform (PHO2) resembles animal PHOs in size and in its high affinity for branched glucans (Gerbrandy, 1974). This isoform is also known as the **H**-type (after its high affinity to glycogen). The second isoform (PHO1) has a lower affinity to highly branched glucans and therefore is also referred to as type-**L** starch phosphorylase. Its sequence includes an N-terminal transit-peptide that targets the enzyme to the plastids, as well as a sequence of mostly hydrophilic amino acids inserted in the center of the mature protein. This central peptide insertion is not found in *E. coli* or animal glycogen phosphorylase and accounts for the larger molecular size and the different substrate affinity (Nakano and Fukui, 1986; Nakano et al., 1989). PHO1 shows a high degree of conservation for most of its amino acid sequence across plants, with

the exception of the regions corresponding to the N-terminal transit peptide and the central insertion characteristic of this isoform type (Nakano et al., 1989).

1.5.1 Starch phosphorylase plastidial isoforms *Pho1a* and *Pho1b*

In potato, plastidial starch phosphorylases are encoded at two loci: *Pho1a* on chromosome III and *Pho1b* on chromosome V (Stp-23 and Stp-1/Stp-L in Chen et al., 2001); Li et al., 2008). The major characteristics of these two loci, as well as of *Pho2* (StpH in Chen et al., 2001; Li et al., 2008) are summarized in Table 1.

Table 1. Major features of potato starch phosphorylases

	PHO-H	PHO1a	PHO1b
Sub cellular localization	cytoplasm	plastid	Plastid
Chromosome	IX	III	V
ORF (bp)	2514	2898	2922
Primary structure (a.a.)	838	966	974
N-terminal transit peptide	no	yes	Yes

The *Pho1a* gene extends over ~16,4 kb from which only 17% are coding sequences (Nakano et al., 1989). The 966 amino acids protein is encoded in 15 exons. It has been suggested that this gene is present in a single copy per haploid genome (Camirand et al., 1990). The molecular evolution of plant phosphorylases has been discussed based on the presence of the central peptide insertion in comparison to other glucan phosphorylases. It has been proposed that the insertion is the result of a change in the splicing site of an intron (Nakano and Fukui, 1986; Camirand et al., 1990). This hypothesis, however, has been argued based on the lack of consistency in the position of the introns for instance in human brain PHOs as well as in *Vicia faba* (Buchner et al., 1996).

The second isoform, *Pho1b,* was discovered more than ten years later in potato leaves (Sonnewald et al., 1995). It was found to be 80-84% similar to *Pho1a* with the exception of the N-terminal target peptide and the central insertion which were only 22% and 29% similar, respectively.

Regions containing positive and negative regulators have been described in the *Pho1a* promoter (St. Pierre and Brisson, 1995). In potato, homodimers $(Pho1a)_2$ and heterodimers (*Pho1a*-*Pho1b*) of the two isoforms occur (Albrecht et al., 1998). In tubers, levels of *Pho1b*

transcript are extremely low and the gene product cannot be detected in crude extracts so far. However, Albrecht et al., (2001) described tissue-printing experiments revealing *Pho1b* expression restricted to a small number of cells in the close vicinity of the vascular tissue, while *Pho1a* is constitutively expressed (Albrecht et al., 2001).

The essential role of *Pho1* in starch synthesis was demonstrated in rice and *Chlamydomonas reinhardtii* starch phosphorylase mutants (Reinhold et al., 2007; Satoh et al., 2008). Seeds of *Pho1* mutant rice lines generated by N-methyl-N-nitrosourea (MNU) treatment, varied between shrunken to pseudonormal (Figure 3A). The cells had modified amylopectin structure and smaller starch grains accumulated in the endosperm (Figure 3C). The mutant phenotypes were temperature dependent, being more striking at 20°C than at 30°C (Satoh et al., 2008). Similarly, in *C. reinhardtii STA4* mutants Dauvillée *et al*. (2006) observed modified amylopectin structure and high amylase content in addition to abnormal starch granules (Figure 3D).

Zeeman et al. (2004b), suggested that starch phosphorylases were not required for starch degradation in leaves as indicated by analysis of *Arabidopsis* mutants. Nevertheless, upon abiotic stress, starch phosphorylase mutants formed lesions in the edges of the leaves and showed excessive wilting (Figure 3B). The authors suggested a role for starch phosphorylase in tolerance to abiotic stress. It has been hypothesized that starch phosphorylase may have an important role for the supply of ATP in cases of limited availability in the plastids (Gounaris, 2001; Reinhold et al., 2007).

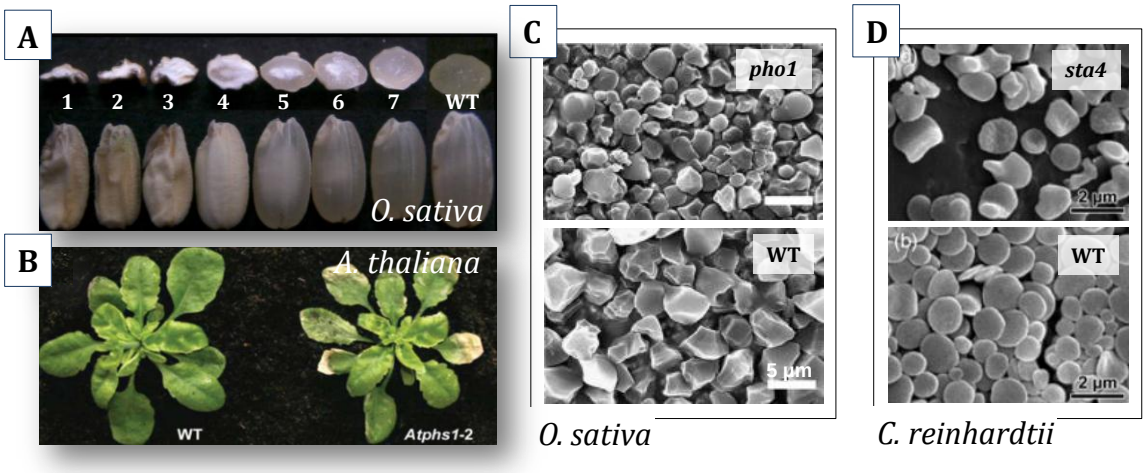


Figure 3. Phenotypes of starch phosphorylase mutants in different species. **A**. Rice seeds from *Pho1* mutants vary from shrunken (1) to low weight seeds (7) in comparison to the wild-type (Satoh et al., 2008). **B**. *Arabidopsis thaliana Atphs-1* mutants showed higher susceptibility to abiotic stress (Zeeman et al., 2004). **C**. Smaller starch granules with modified amylopectin structure of rice *pho1* mutants (Satoh et al., 2008). **D**. Abnormal starch granules of *Chlamydomonas reinhardtii sta4* mutants (Dauvillé et al., 2006).

1.6 Objectives

A candidate gene-based association study identified SSCP polymorphisms at two starch phosphorylase loci (*Pho1a* and *Pho1b*) associated with potato tuber quality traits (Li et al., 2008). The aims of this work were to identify and characterize the genetic variation within these loci and to determine whether the allelic diversity has direct impact on the enzyme function, which contributes to the phenotypic variation observed in tuber quality traits. The strategy followed to accomplish the objectives included (i) a survey of starch phosphorylase natural alleles in four tetraploid genotypes, (ii) identification of alleles underlying SSCP polymorphisms associated with tuber quality traits, (iii) verification of the association of the alleles with quality traits in different populations, (iv) expression analyses of *Pho1a* alleles, and (v) characterization of enzyme functions in 40 potato genotypes. This work provides strong evidence for the direct effect of natural *Pho1a* alleles in tuber quality traits, as opposed to indirect associations resulting from large blocks of linkage disequilibrium. Additionally, it contributes to the understanding of the molecular basis underlying cold sweetening in potato and empowers plant breeding by providing new insights in these complex traits.

2 Materials and Methods

2.1 Materials

2.1.1 Equipment and chemicals

Most chemicals were obtained from:

Applied Biosystems / Ambion (Austin, USA)
Carl Roth GmbH (Karlsruhe, Germany)
GE Healthcare (Munich, Germany)
Invitrogen GmbH (Carlsbad, USA)
MBI Fermentas GmbH (St. Leon-Rot, Germany)
Merck (Darmstadt, Germany)
Millipore™ (Billerica, MA, USA)
New England BioLabs, Inc. (Ipswich, USA)
PEQLAB Biotechnologie GmbH (Erlangen, Germany)
Promega GmbH (Mannheim, Germany)
QIAGEN GmbH (Hilden, Germany)
Roche Applied Science (Mannheim, Germany)
Sigma-Aldrich Chemie GmbH (Munich, Germany)

2.1.2 Enzymes

The main sources of enzymes were:

Ampliqon (Skovlunde, Denmark)
Applied Biosystems / Ambion (Austin, USA)
Invitrogen GmbH (Carlsbad, USA)
MBI Fermentas GmbH (St. Leon-Rot, Germany)
New England BioLabs, Inc. (Ipswich, USA)
PEQLAB Biotechnologie GmbH (Erlangen, Germany)
Promega GmbH (Mannheim, Germany)
Roche Applied Science (Mannheim, Germany)
USB (Cleveland, USA)

2.1.3 Antibiotics

Antibiotic stock solutions were dissolved in ddH_2O and sterilized by filtration through a 0.22 µm Millipore™ filter. The antibiotics and concentrations used are listed in Table 2.

Table 2. Antibiotics

Antibiotic	Stock solution	Final concentration
Ampicillin (Amp)	100 mg/mL	100 µg/mL
Gentamycin (Gent)	50 mg/mL	50 µg/mL
Kanamycin (Kan)	50 mg/mL	50 µg/mL

2.1.4 Vectors

The following vectors were used:

Cloning:

pGEM-T Easy, Promega GmbH (Mannheim, Germany)
Zero Blunt® TOPO®, Invitrogen, (Carlsbad, USA)
TOPO® XL, Invitrogen, (Carlsbad, USA)

Gateway® Donor and Expression vectors (Invitrogen, Carlsbad, USA)

pDONR™221
pBAD-DEST49

2.1.5 Oligonucleotides

Primers used in this project were custom synthesized by commercial sources (Sigma, Invitrogen, and Metabion) and are listed in Table 3.

Table 3. List of oligonucleotides

Primer purpose	Forward primer	Primer sequence 5' - 3'	Reverse primer	Primer sequence 5' - 3'	Annealing temperature
SCCP primers					
	Stp23-8_F[1]	gcaacagctcaaagtgttcg	Stp23-8_R[1]	cacctcctcctgaccatctt	57
	StpL-3F[1]	gttcagagacatcatggcaac	StpL-3R[1]	agtccagaaagcaagaagca	58
Full length *Pho1a* cDNA synthesis					
	D02_F	atcactctcattcgaaaagctagat	D02_R	tgcctttgttatttttcattcacttc	65
	D03_F	gctagatttgcatagagagcaca	D03_R	gtggtaataacatcatcctcttacac	66
Pho1a Sanger sequencing primers					
	D02_F	atcactctcattcgaaaagctagat	D02_R	tgcctttgttatttttcattcacttc	-
	Di01_F	aaagatggtcaggaggaggtgg	Di02_R	tggcacaaagactaccttcaacag	-
Pho1a SSCP-SNP specific primers (Li, L. unpublished)					
	Stp23-8b_F	cgcatcagaaaaaacctcgg	Stp23-8b_R	acctcctcctgaccatcttt	58 (TD)[2]
Primers for *Pho1a* amplicon sequencing					
	DsnpC22T_F	gaggaccgaacacaacacactt	Stp23-8_R[1]	cacctcctcctgaccatctt	59
	Dpyr1(A108T)_F	ccatcccattactgaacaagg	Dpyr1(A108T)_R	aggagcgaatcacgaacact	59
	DgSNPs01_F	gcgacctgagttcttttgct	DgSNPs01_R	catgctttggtgtgctctcc	59
	DgSNPs02_F	cctcctcttagtgccatcata	DgSNPs02_R	ccctccactcattttgaagat	59
	DgSNPs03_F	gctcatgaaattgcagggcttag	DgSNPs03_R	gcttcaatgttccaaatgtctt	59
	Dpyr3(D926N)_F	gccgtgctgactatttccttt	Dpyr3(D926N)_R	tccacagcttcaatgttcca	59
Pho1a alleles SNP specific primers					
	Dc208_F	agttcttttgctcctgatgcc	Dsnp824G_R	gaacatcatacgcaactgtcc	54 (TD)
	Dc2478_F	tgaaataagggaagaggttgga	Dsnp2776A_R	aacaaactggacttactttctgatt	54 (TD)
	Dsnp2578A_F	cttagaaaagaaagagctggca	Di14-50_R	caaagtcaatagcatacacaaaacat	54 (TD)

Table 3 (continued)

Primer purpose	Forward primer	Primer sequence 5' - 3'	Reverse primer	Primer sequence 5' - 3'	Annealing temperature
Pho1a promoter					
	DP01_F	agttgatttaagagtcattgaacattacat	DP01_R	aagagagtgaggattgtgaattattttatt	62
	DP02b_F	ccaatatccaacatgccaaca	DP02_R	tgcggttctttcgtttttga	62 (TD)
Pho1a RT-qPCR					
	Dpyr3(D926N)_F	gccgtgctgactatttcctt	Dpyr3(D926N)_R	(bio)tccacagcttcaatgttcca	60
	EF1a_F	attggaaacggatatgctcca	EF1a_R	tccttacctgaacgcctgtca	60
Pho1a pyrosequencing					
	Dpyr2(K275R)_F	gaggaggtggctgaagattg	Dpyr2(K275R)_R	(bio)aaggctgattgtggttctgg	60
	DpSeq2(K275R)S	ttggattggtggagaggata			-
	Dpyr3(D926N)_F	(bio)gccgtgctgactatttcctt	Dpyr3(D926N)_R	(bio)tccacagcttcaatgttcca	60
	DpSeq3(D926N)S	ttgtccacctttctggt			-
Full lenght *Pho1b* cDNA synthesis					
	StpL-3F[1]	gttcagagacatcatggcaac	X01_R	tggctttcatcctttaactgg	56
Pho1b SSCP-SNP specific primers					
	StpL-3b_F	gaagaagtttctctgttagccac	StpL-3b_R	cgagtgacgtctgtagttatactagt	58 (TD)
	StpL-3e_L	ggacccttgtattttcaggat	StpL-3e_R	aaagctcttccctgaaagaac	

[1] Primer from Li et al., (2008)
[2] TD. Touchdown-PCR: Initial annealing temperature is 5°C higher than noted, and decreases 0.5°C per cycle during the following 10 cycles. Last 25 cycles run at the noted annealing temperature

2.1.6 Bacteria

Bacterial strains of *Escherichia coli* are listed in Table 4. Strains were obtained from the *E. coli* Genetic Stock Center at Yale (CGSC; http://cgsc.biology.yale.edu) or purchased from a commercial source (Invitrogen).

Table 4. Bacterial strains

Strain	Genotype	Source	Reference
DH5α™	F- ϕ80*dlac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(*rk*-, *mk*+) *pho*A *sup*E44 λ- *thi*-1 *gyr*A96 *rel*A1	Invitrogen	(Hanahan, 1983)
One Shot® ccdB Survival™	F- *mcr*A Δ(*mrr*-*hsd*RMS-*mcr*BC) ϕ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*Δ139 Δ(*ara*-*leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G *fhu*A::*IS*2	Invitrogen	(Bernard and Couturier, 1992; Bernard et al., 1993)
HfrG6[a]	Hfr λ- *his*A323	T.S. Matney	(Matney et al., 1964)
HfrG6MD2	Hfr λ- *his*A323 Δ(*bio*H-*asd*)29	M. Schwartz	(Schwartz, 1966)
BW25113[b]	F- Δ(*ara*D-*ara*B)567 Δ*lac*Z4787(::*rrn*B-3) λ- *rph*-1 Δ(*rha*D-*rha*B)568 *hsd*R514	B.L Wanner	(Datsenko and Wanner, 2000; Baba et al., 2006)
JW3381-3	F- Δ(*ara*D-*ara*B)567 Δ*lac*Z4787(::*rrn*B-3) λ- Δ*mal*T752::*kan rph*-1 Δ(*rha*D-*rha*B)568, *hsd*R514	Keio collection	(Baba et al., 2006)
JW5689-1	F- Δ(*ara*D-*ara*B)567 Δ*lac*Z4787(::*rrn*B-3) λ- Δ*mal*P751::*kan rph*-1 Δ(*rha*D-*rha*B)568 *hsd*R514	Keio collection	(Baba et al., 2006)

[a] HfrG6 is the parent strain of HfrG6MD2 lacking the Δ(*bio*H-*asd*)29 mutation
[b] BW25113 is the parent strain for the Keio collection single gene knockouts JW3381-3 and JW5689-1

2.1.7 Plant Material

Potato (*Solanum tuberosum* L.) tetraploid genotypes used in this study were the populations GABI-CHIPS (Li et al., 2008), TASK (Urbany et al., 2011), and BIOSOL (Fischer, M. and Schreiber, L., unpublished). The latter comprises 40 commercial cultivars selected for good and bad chipping quality after cold storage (20 each). Tetraploid genotypes used for allele mining experiments were cultivars Diana (A), Theresa (E), Satina (I), and Saturna (U). Additionally, the diploid lines H80.577/1 (P_3), H80.576/16 (P_{38}), H80.696/4 (P_{40}), and H81.839/1 (P_{54}) (Gebhardt et al., 1989; Menendez et al., 2002) were included in this study. Tubers were provided by the breeding companies BIOPLANT (Ebstorf), SaKa Pflanzenzucht GbR (Windeby), and NORIKA GmbH (Groß Lüsewitz). Sterile cultures for propagation were obtained from the *Solanum tuberosum* collection of the Max Planck Institute for Plant Breeding Research (Köln).

2.1.8 Buffers and culture media

Buffers and culture media were prepared as described by Sambroock and Russell (2001) and/or obtained from commercial sources (Carl Roth, Merck, Sigma). All solutions were made using purified Milli-Q-water (Millipore™). Growth media included Lysogeny Broth (LB-Luria), Super Optimal Broth (SOC, Invitrogen), and M9 Minimal Media for bacteria (Sambrook and Russell, 2001) unless otherwise specified. Agar was added to a final concentration of 1% for solid media. When needed, solutions were autoclaved for 10 or 20 minutes at 121°C or sterile filtered using Millipore™ filters.

2.1.9 Software and databases

The following software tools were employed:

Statistical analysis software:

SPSS Statistics (IBM Corporation, Somers NY, USA)

Primer design:

NetPrimer (PREMIER Biosoft International, Palo Alto, CA, USA)

Primer3 (Rozen and Skaletsky, 2000)

Nucleic acids and protein sequence analysis tools:

Bioedit (Hall, 1999)

CLC Genomics Workbench (CLC Bio, Aarhus, Denmark)

ClustalW2 and ClustalX2 (Larkin et al., 2007)

DNASTAR Lasergene® 8 (DNASTAR, Inc., Madison WI, USA)

MEGA Version 4 (Tamura et al., 2007)

NovoSNP (Weckx et al., 2005)

Protein structure software:

AUTO-MUTE (Masso and Vaisman, 2010; http://proteins.gmu.edu/automute/)

The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC

Toolbox compilation resources:

EMBL-EBI Toolbox (McWilliam et al., 2009); http://www.ebi.ac.uk/tools/)

MPI Bioinformatics Toolkit (Biegert et al., 2006); http://toolkit.tuebingen.mpg.de/)

Statistics software:

SPSS Statistics (IBM Corporation, Somers NY, USA)

Databases:

EcoGene (http://www.ecogene.org/)

E. coli Genetic Stock Center (http://cgsc.biology.yale.edu/)

Potato Genome Sequencing Consortium (http://www.potatogenome.net/)

TAIR - The Arabidopsis Information Resource (Swarbreck et al., 2008)
http://www.arabidopsis.org/)

The NCBI BioSystems database(Geer et al., 2010); http://www.ncbi.nlm.nih.gov/)

The SOL Genomics Network (Mueller et al., 2005); http://solgenomics.net/)

2.2 Methods

2.2.1 Growth conditions

Unless otherwise indicated, *E. coli* cultures were incubated in LB-Luria or M9 supplemented with 2% carbon source (glucose, maltose, or maltodextrin) at 37°C over night. Liquid cultures were placed on a shaker at 220 RPM.

Plants were regenerated from sterile shoot culture and subsequently grown in plastic pots (25 cm diameter) under green house conditions or in a Saran-house under natural climatic conditions from May to September 2008 and 2009. Leaf and mature tuber tissues were harvested and kept at -80°C until processed for molecular analysis.

2.2.2 Nucleic acids isolation

2.2.2.1 DNA

E. coli plasmid DNA was isolated using the Plasmid Mini Kit (Qiagen). Plant genomic DNA was isolated from leaf or tuber tissue either by the CTAB method (Doyle and Doyle, 1987) or using the DNeasy® Plant Kit (Qiagen) according to the manufacturer's instructions. Purification of gel-extracted DNA was done using the QIAquick Gel Extraction Kit (Qiagen). PCR products were cleaned-up with ExoSAP-it® (USB) or QIAquick PCR Purification Kit (Qiagen) prior to sequencing. DNA was quantified with a NanoDrop UV-Vis spectrophotometer (Thermo Scientific).

2.2.2.2 RNA

RNA from leaves tissue was extracted using the ToTALLY RNA™ Total RNA Isolation Kit (Ambion).

Tuber tissue RNA was isolated either with the PureLink™ Plant RNA Reagent (Invitrogen) or using the protocol described by Kumar et al. (2007) using acid-phenol:chloroform (Ambion) instead of phenol-chloroform. RNA was eluted in DEPC-treated water. RNA was quantified with a Qubit® fluorometer (Invitrogen) or a NanoDrop UV-Vis spectrophotometer (Thermo Scientific).

2.2.3 Polymerase Chain Reaction (PCR)

Standard PCR reactions were setup in a 25 µL total volume containing 1x PCR buffer (10 mM Tris-HCL pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.1% Triton X-100); 0.2 mM each dNTP; 1 µM each forward and reverse primers, 1.5 U Taq-Polymerase and 40 ng of template DNA (Sambrook and Russell, 2001). For standard PCR, either Taq-Polymerase prepared as described by Pluthero (1993) or Ampliqon III Taq DNA Polymerase (Ampliqon) was used. Reactions were generally performed in a Labcycler (SensoQuest GmbH, Göttingen) under the following cycling conditions: Initial denaturation (2 min at 94°C), 20 to 35 cycles of denaturation (30 sec at 94°C), annealing (30 sec at 53 to 65°C, see Table 3), extension (1 min/Kb at 72°C), final extension (5 min at 72°C).

When high accuracy was required (cloning, allele mining), the following enzymes were used, adapting the protocols of the manufacturer: KAPA2G™ Fast PCR Kit (PEQLAB Biotechnologie), FastStart High Fidelity PCR System (Roche), AccuPrime™ Pfx DNA Polymerase (Invitrogen).

2.2.4 BAC-Library screening

The *Pho1b* 1,8 Kb radioactive labeled probe for BAC library screening was generated with the primers StpL3_F / StpL3_R (Li et al., 2008 Table 3). The DNA template used was from the diploid genotype P_{40}, parent of the genotype P_6/210 used to construct the "BC" library (Ballvora et al., 2007).

The PCR product was treated with EXOSAP-IT® and diluted to a final concentration of 200 ng. Nick translation was performed in a 25 µL reaction volume with buffer A1 (0.2 mM of each dTTP, dCTP, and dGTP; 1.5 U DNase I plus 1 U DNA-polymerase; 3 µL ^{32}P dATP) by incubating for one hour at 16°C. Purification of the probe was done in a Sephadey-G50 column (Amersham Biosciences) following the recommendations of the manufacturer.

Finally, the probe was heated at 95°C for 5 minutes. Filter hybridization and autoradiography was performed as described in Gebhardt et al. (1989)

2.2.5 cDNA synthesis

Contaminating DNA from the RNA samples was removed with DNaseI using the DNA-free™ Kit (Ambion). First strand cDNA was synthesized from 1 µg of total RNA using the Transcriptor First Strand cDNA Synthesis Kit (Roche) following the manufacturer's recommendations. Sequence-specific or anchored-oligo $(dT)_{18}$ primers were used to obtain first strand cDNA. Further amplification of the cDNA was performed by PCR as described above.

2.2.6 RT-qPCR

First strand cDNA was diluted 1:100, and 5 µL of the dilution were used as template for quantitative amplification using the Power SYBR® Green PCR Master Mix (Applied Biosytems) system. Amplification of *Pho1a* transcripts was done using the Dpyr3(D926N)_F/Dpyr3(D926N)_R primer pair. The gene used for standardization was *elongation factor 1*-α (*efl-α*), the stability of which upon cold storage has been previously described (Nicot et al., 2005). Primers used for *efl-α* amplification were EfIa_F/EfIa_R. Reactions were performed in a Mastercycler® ep Realplex (Eppendorf). Cycling conditions were as follows: (1) 95°C for 10 min; (2) 40 cycles of [95°C for 15 sec, 60°C for 30 sec, and 60°C for 45 sec]; (3) 95°C/15 sec; (4) 60°C/15 sec. Finally, δ20 minutes for melting curve analysis finished by 95°C/15 seconds. Relative expression of the *Pho1a* transcripts was calculated by the standard curve method (Applied Biosystems, User Bulletin #2, 2001), generally with three biological and three or four technical replicates.

2.2.7 Pyrosequencing

Two sequencing templates were generated by amplification of *Pho1a* cDNA using the primer pairs Dpyr3(D926N)_F/Dpyr3(D926N)_R(bio) and Dpyr2(K275R)_F(bio)/ Dpyr2(K275R)_R (Table 3). The cycling conditions for the primary PCR were (1) 95°C for 2 min; (2) 40 cycles of [95°C for 15 sec, 60°C for 20 sec, and 60°C for 30 sec]; 60 for 50min.

Solid phase preparation of single strand cDNA template was prepared using the Pyrosequencing Vacuum Prep Tool according to the manufacturer's instructions (Sample

Preparation Guidelines for the PSQ 96MA System, Biotage). The reaction set up was as follows: 20 µL PCR product, 15 µL H20, 10 µL binding buffer (Biotage), and 5 µL streptavidin sepharose (GE Healthcare).

The pyroseqeuencing reactions were carried out in the PSQ™96 MA System (Biotage), using Mark™ Gold Q96 Reagents Kit (Biotage) following the manufacturer's instructions. Pyrosequencing primers were DpSeq3(D926N)_S and DpSeq2(K275R)_S (Table 3).

2.2.8 Cloning of full length *Pho1* cDNA for allele mining

Starch phosphorylase alleles were cloned from cultivars Diana (A), Teresa (E), Satina (I) and Saturna (U). Full length *Pho1a* cDNA was independently obtained from leaf and tuber tissue after selective amplification with primer pairs D02_F/D02_R or D03_F/D03_R (Table 3). Full length *Pho1b* cDNA was cloned from leaf tissue after selective amplification with primer pair StpL_F/X01_R (Table 3). The sequences were subsequently cloned into *E. coli* plasmid vectors pGEM-T Easy (Promega) in a 1:1 (insert:vector) molar ratio using T4 DNA ligase (Promega) or TOPO® XL (Invitrogen), the latter being more suitable for large size products. Following the manufacturer's instructions, the ~3 Kb PCR fragments were previously gel-purified using crystal violet (Rand, 1996) to avoid DNA damage under UV light. Fragments generated with proof-reading DNA polymerase were cloned into Zero Blunt® TOPO® (Invitrogen).

2.2.9 Preparation of competent bacteria cells and transformation

E. coli mutants and parental lines were prepared for calcium chloride chemical transformation following the protocol described by Seidman et al. (2001) with a few modifications: One colony from an LB-plate of the desired genotype was inoculated into 2 mL LB-medium and incubated at 37°C overnight with shaking at 220 RPM. 1 mL of the overnight cell culture was inoculated into 100 mL of LB medium in a 500 ml flask and incubated under conditions described before for ~2 hours to an OD_{600} of 0.25-0.3. Cultures were chilled on ice for 15 minutes and subsequently centrifuged for 10 minutes at 4000 RPM. Pellets were re-suspended in 40 mL ice-cold 0.1 M $CaCl_2$ and incubated on ice for 30 minutes. Finally cells were centrifuged, carefully re-suspended in 6 mL 0.1 M CaCl2 + 15% glycerol, and 50 µL aliquots were immediately used for transformation or stored at -80°C until used.

For transformation 50 to 100 ng of plasmid DNA and 50 µL of competent cells were incubated on ice for 15 minutes. Cells were heat-shocked in a heat block at 42°C for 5 minutes and

placed on ice immediately after. 1 mL SOC was added to the cells. Expression of the antibiotic resistance genes, if any, was allowed incubating the transformed cells for 1 hour at 37°C on a roller drum at 40 RPM. Aliquots of 20 µL to 100 µL were plated onto LB or LB-selective media.

Strains DH5α™ and One Shot® ccdB Survival™ (Invitrogen) are commercially available and were transformed according to the manufacturer's instructions.

Toothpick samples from transformed cells were diluted in 20 µL sterile water. 1 µL of the dilution was used for colony PCR in order to confirm the presence of the insert of interest (reaction set up as described in 2.2.3, using the bacteria dilution as a source of DNA).

2.2.10 Sequencing

Sequencing of PCR products or plasmid DNA was carried out by the MPIPZ DNA core facility on Abi Prism 377, 3100 and 3730 sequencers (Applied Biosystems) using BigDye-terminator v3.1 chemistry.

2.2.11 Allele mining

Clones obtained from the different cultivars and tissues, were sequenced as described above. *Pho1a* and *Pho1b* alleles were sequenced with four primers each, to cover the entire ~3 kb fragment. Sequences were entered into an MS-Excel sheet, one nucleotide per column. A consensus for each clone was created and inconsistencies, when they occurred, were verified manually using the trace files. Each consensus sequence was treated as a potential haplotype.

The reference sequences used for SNP detection were available from GenBank (D00520 for *Pho1a* and X52385 for *Pho1b*). For SNP discovery, each sequence was compared to the reference. Non-polymorphic positions were discarded, as well as those where the variants were not consistent in at least two independent clones. Commands and functions were automated with macros to avoid errors. Remaining SNPs were classified according to their frequency and verified using the software SeqMan (DNASTAR) to visualize trace files.

Clones were then sorted by common SNPs at different positions and putative allelic sequences were determined.

2.2.12 *In silico* protein modeling and simulation of point mutations

The structure of PHO1a was modeled with SWISS-MODEL using the crystal structure of the yeast (*Saccharomyces cerevisiae*) glycogen phosphorylase (YGP) as a template (PDB: 1YGP). Graphic outputs and *in silico* mutations were obtained with PyMOL.

The AUTO-MUTE algorithm was used to evaluate the effect of point mutations in the structure of the rabbit muscle glycogen phosphorylase (PDB: 6GPB).

2.2.13 Native protein extraction

Deep frozen tuber tissue (~1 g) was homogenized in 500 µL grinding buffer (100 mM Hepes-NaOH, pH 7.5, 1 mM EDTA, 5 mM DTT, 10% (v/v) glycerol). Homogenates were centrifuged at maximum speed (12000 rpm) for 15 minutes at 4°C and the supernatants containing the protein were quantified with a Qubit® Fluorometer (Invitrogen) and stored at -20°C for native PAGE.

2.2.14 Native PAGE

40 ug or 50 ug of total native protein were loaded into NativePAGE™ Novex® 3-12% Bis-Tris Gels and electrophoresed at 200 V and 4°C in an XCell *SureLock*® Mini-Cell (Invitrogen), following the manufacturer's instructions.

2.2.15 Phosphorylase activity test

Phosphorylase activity was tested as described by Manchenko (2003). Electrophoresed gels were placed for 15 minutes in 100 mM citrate-NaOH (pH 6.5) at room temperature, and then incubated for at least 4 hours at 37°C in a 100 mM citrate, 20 mM G1P (pH 6.5) and 0.2% (w/v) soluble starch or glycogen. Glucans formed in the separation gel by phosphorylase activity were stained with iodine (Lugol solution).

2.2.16 2-D SDS electrophoresis and mass spectrometry

Mass spectrometry was performed in colaboration with the Masss Spectrometry Group of the Max Planck Institute for Plant Breeding Research.

Standard 2-D SDS electrophoresis was carried out as described by Laemmli (1970). 2-D Gels were stained using Coomassie PageBlue (Fermentas). Spots were selected visually and punched robotically using the Proteineer spII system (Bruker Daltonics, Bremen). The excised spots were digested with trypsin and spotted on AnchorChip™ targets by a Proteineer dp robot (Bruker Daltonics, Bremen) using a standard HCCA thin layer preparation. Peptide mass fingerprint (PMF) data and subsequent peptide fragmentation fingerprint (PFF) data were collected on an UltraflexIII MALDI ToF/ToF mass spectrometer (Bruker Daltonics, Bremen) in a two-stage process as outlined by Suckau et al. (2003). Eight peptide precursors with sufficient signal intensity were selected for fragmentation.

3 Results

3.1 The starch phosphorylase *Pho1a*

3.1.1 Allele mining on the *Pho1a* locus uncovered the allele accountable for the *Stp28-8b* SSCP polymorphism associated with better tuber quality traits.

According to Li et al. (2008), two SSCPs (single-strand conformation polymorphism) in the starch phosphorylase locus *Pho1a* on chromosome III were associated with potato chip quality traits (Table 5).

Table 5. *Pho1a* SSCP markers associated with chip quality traits (Li et al., 2008)

SSCP marker	Marker frequency	Chip quality at harvest	Chip quality after cold storage	Tuber starch content	Tuber starch yield
Stp23-8b	0.27	↑ (3.9)	↑ (9.0)	↑ (11.4)	↑ (6.7)
Stp23-8a	0.53	–	↑ (5.2)	↑ (9.6)	↑ (5.9)

Markers *Stp23-8b* and *Stp23-8a* are two different SSCP fragments obtained from one PCR amplicon[1]. The frequency indicates the proportion of the population that contains the allele. The direction of the arrow depicts a positive or negative effect of the marker on a given trait. The number in parenthesis indicates the amount of variance (in %) explained by the marker.

As a first step in the identification of the *Pho1a* alleles accountable for the associated SSCP fragments, four tetraploid potato genotypes were selected from the GABI-CHIPS population used in the aforementioned study. The genotypes were chosen based on the presence or absence of the associated SSCP markers *Stp23-8b* and *Stp23-8a* (Table 5). The cultivars Diana (A), Theresa (E) and Saturna (U) were positive for the presence of both SSCP markers, while in cultivar Satina (I) the markers were absent. The latter was therefore chosen as a control genotype. Additionally, four diploid parental lines used as parents in a QTL mapping population for tuber starch and sugar content (P_3, P_{38}, P_{40} and P_{54}; Schäfer-Pregl et al., 1998) were included in the analysis.

[1] Primer pair Stp23-8_F / Stp23-8_R (Table 3, materials and methods)

To obtain the *Pho1a* allele sequences of each genotype, total RNA was extracted from tuber and leaf tissues. Reverse transcription followed by gene specific PCR amplification[2] produced full length *Pho1a* cDNAs that were further cloned and sequenced for haplotype determination. A total of 161 full length (2901 bp) cDNA clones were obtained from the four tetraploid cultivars out of 19 independent amplifications (Table 6).

Table 6. Number of full length clones obtained from four tetraploid potato cultivars

Cultivar	Diana		Theresa		Satina		Saturna		Total
Tissue	Leaf	Tuber	Leaf	Tuber	Leaf	Tuber	Leaf	Tuber	
PCR replications	2	3	3	3	2	2	2	2	19
Clones [a]	26	20	22	37	16	13	14	13	
Total clones per cultivar	46		59		29		27		161

[a] Each PCR product was ligated into cloning vectors and transformed into *E. coli* (Section 2.2.8). The number of clones represents the amount of colonies selected whose plasmids were sequenced per transformation.

After comparison of the sequences, a SNP was considered factual if found in at least three different clones from two independent amplification reactions. PCR replications were carried out to detect and discard false SNPs originated from amplification of errors occurring at early stages during the PCR (Meyerhans et al., 1990; Pienaar et al., 2006). Singletons (or polymorphic positions that appeared only in one clone) were discarded from the analysis. The polymorphic positions found in cDNA clones were confirmed in genomic DNA clones and also by direct PCR sequencing. Consistent polymorphisms among the clones led to nine different sequences (Table 7). Out of 15 SNPs, seven introduced an amino acid change, whereas the remaining eight SNPs were silent mutations.

At least one copy of 'Sequence 1' was present in each of the four tested genotypes. This sequence is identical to those for *Pho1a* in the GenBank database (D00520 and X52385). Hence, 'Sequence 1' was referred to as the Reference Haplotype "H_R". The very few SNPs found between Satina clones were not reproducible. Therefore, the *Pho1a* locus in this genotype was assumed to be H_R-homozygous.

[2] D02_F / D02_R or D03_F / D03_R (Table 3)

Table 7. Summary of *Pho1a* cDNA alleles cloned from four tetraploid potato cultivars.

← Clone sequences	SNP	$C_{22}T$	$G_{322}A$	$C_{504}T$	$A_{534}G$	$A_{720}C$	$A_{824}G$	$G_{1923}T$	$G_{1939}T$	$G_{2578}A$	$G_{2616}A$	$G_{2754}A$	$A_{2765}T$	$G_{2776}A$	$C_{2795}A$	$G_{2823}C$
	A.A.	His-$_{8}$-Tyr	Ala -$_{08}$-Thr	-	-	-	Lys-$_{275}$-Arg	-	Ala-$_{647}$-Ser	Gly-$_{860}$-Arg	-	-	Glu-$_{922}$-Val	Asp-$_{926}$-Asn	Thr-$_{932}$-Lys	-
1 (H_R)		C	G	C	A	A	A	G	G	G	G	G	A	G	C	G
2		T	A	C	G	C	A	G	G	G	G	G	A	G	A	C
3		T	A	C	G	C	A	G	G	G	G	G	A	A	A	C
4		T	A	C	G	C	G	G	G	G	G	G	A	G	C	G
5 (H_A)		T	A	C	G	C	G	T	T	G	G	G	A	A	A	C
6		T	A	C	G	C	G	T	T	G	G	G	A	G	A	C
7		C	G	T	G	A	A	G	G	A	G	A	A	G	A	C
8		C	G	T	G	A	A	G	G	A	G	A	A	G	A	G
9		C	G	T	G	A	A	G	G	A	A	G	T	G	A	G

SNPs in *Pho1a* sequences at the given position with respect to the AUG codon in the mRNA sequence. A.A. – Amino acid exchange introduced by the nucleotide substitution (the number indicates the position from the start methionine). The numbers of the first column name the different sequences found in cDNA clones. Sequences were clustered into three haplotype groups (red, green, and blue), according to the first four SNPs. H_R and H_A stand for Reference and Associated Haplotype, respectively. For complete alignment see Appendix 1

Since the potato genotypes used for allele mining were tetraploid, a maximum of four different alleles was expected to be found for a single copy gene, as it was proposed for *Pho1a* (Camirand et al., 1990). Nevertheless, annealing and comparison of the clones suggested the presence of at least eight different sequences in addition to H_R (Appendix 1). Most of them were consistently obtained in cDNA clones from Diana (sequences 1, 2, 4, 5, 6, 9), Theresa (sequences 1, 2, 3, 5, 6, 9), and Saturna (sequences 1, 2, 4, 5, 7, 8, 9) in independent experimental replications. The presence of more than four alleles per genotype was an indication of either additional copies of this gene or technical artifacts.

To discard the presence of a second copy of *Pho1a* or a pseudogene in the potato genome, the exons and partial intron sequences obtained from genomic DNA sequencing were aligned using the BLAST-algorithm against the recently released drafts of the potato genome sequence that became publicly available during the development of this work (Potato Genome Sequencing Consortium). The best homology to the *Pho1a* exome was found in four different mini-scaffolds (Figure 4) (Potato Genome Consortium; Version 3). The second and third best hits for the *Pho1a* cDNA sequence corresponded to the closest homologues *Pho1b* and *Pho2,* respectively. This information strongly supported that *Pho1a* is a single gene and

no similar sequences were likely to be amplified by the primers chosen to produce full length *Pho1a* cDNA.

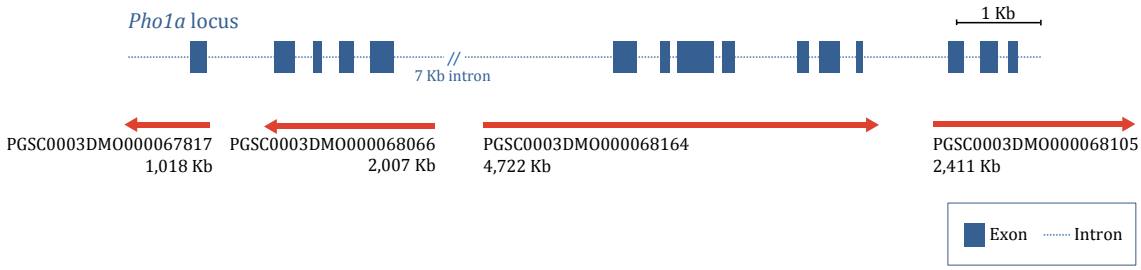


Figure 4. BLAST results of *Pho1a* cDNA against the most recent draft of the potato genome (v3). The red arrows represent the positions of the aligned four mini-scaffolds matching *Pho1a*. The arrow indicates the direction of the scaffold.

The increased amount of haplotypes could be explained by the shuffling of SNPs in the 2 Kb region between positions 824 and 2823, suggesting the possibility of chimeric sequences derived from different alleles in the PCR amplification or incomplete cDNA synthesis.

To approach the problem of the high number of putative alleles, cDNA sequences were clustered into three haplotype groups according to the first four SNPs that consistently appeared together in the cDNA clones: Red or H_R [C_{22},G_{322},C_{504},A_{534}], green [T_{22},A_{322},C_{504},G_{534}], and blue [C_{22},G_{322},T_{504},G_{534}] (Table 7). A subset of 34 potato genotypes from the GABI-CHIPS population (referred to as Standards) was then used to determine which of the conflicting SNPs were in linkage disequilibrium and constituted real alleles. The Standards subpopulation included genotypes Diana, Theresa, Saturna and Satina.

Genomic DNA amplicon sequencing of the individuals from the Standards subpopulation, allowed at first the identification of 12 out of 34 Standard genotypes carrying the sequence from the green group, as manifested by the presence of the SNPs T_{22}[3] and A_{322}[4]. From the 12 individuals, 11 carried the allele in a single dosage (simplex). In the remaining genotype (Saturna) the allele is present in a duplex dosage (Table 8).

[3] DsnpC22T_F / Stp23-8_R (Table 3)
[4] Dpyr1(A108T)_F / Dpyr1(A108T)_R and DgSNPs01_F / DgSNPs01_R (Table 3)

Furthermore, the prior amplicon sequences led to the identification of an additional SNP, $T_{i3\text{-}1}$ (SNP #1 -out of 3- of intron 3, located 41 bases after the 3' boundary of exon 2), present in linkage disequilibrium with A_{322}. The former intron-SNP was previously used to design specific primers representative of the SSCP marker *Stp23-8b*[5] (Li Li, Personal communication). The LD between this two SNPs, which was further confirmed by sequencing genomic DNA clones from Saturna, suggested that the allele accountable for the *Stp23-8b* SSCP polymorphism was contained within the green group.

Consequently, this group was further scrutinized. SNP specific primers for G_{824}[6] and A_{2776}[7] (Table 7) were designed under the rationale that the resolution of these two positions would help to distinguish the real haplotypes from possible artifacts in this group. It was observed that in the Standards subpopulation, the presence of G_{824} occurred always associated with A_{2776} (Table 8). The result that these two SNPs co-segregated in this subpopulation (and in additional 40 genotypes described in Section 3.1.2.2), demonstrated their physical linkage. This information allowed discarding the spurious sequences 3, 4 and 6 (Table 7). In this way, chimeric sequences of the green group could be discarded and 'Sequence 5' (Table 7) was identified as a real haplotype for this group.

Comparing the distribution of the presence of the SNPs of 'Sequence 5' among the individuals of the Standards subpopulation with that of the *Stp23-8b* SSCP marker, confirmed that the *Pho1a* 'Sequence 5' is in agreement, with one exception (St18), with the *Stp23-8b* SSCP fragment (Table 8). Therefore, this allele is referred to as the Associated Haplotype "H_A".

Individuals containing alleles from the blue group were identified with SNP specific primers for A_{2578}[8] (Table 8). The distribution of A_{2578} in the Standards subpopulation was different from that of the associated SSCP fragments (*Stp23-8a* or *Stp23-8b*). In contrast, with one exception, its distribution was in agreement with the distribution of an additional SSCP polymorphism, *Stp23-7c* (Table 8, not published), which was not associated with quality traits in the association study (Li et al., 2008). Consequently, the blue group was not further analyzed. A cDNA haplotype for the *Stp23-8a* marker was not determined.

[5] Stp23-8b_F / Stp23-8b_R (Table 3)
[6] c208_F / Dsnp824G_R (Table 3)
[7] Dc2478_F / Dsnp2776A_R (Table 3)
[8] Dsnp2578A_F / Di14-50_R (Table 3)

Table 8. Cosegregation of *Pho1a* cDNA SNPs and SSCP *STP* markers from Li et al. (2008)

Genotype	Stp23-8a (SSCP)[a]	Stp23-8b (SCCP)[a]	$G_{i3\text{-}1}T$[b]	$C_{22}T$[b]	$G_{322}A$[b]	G_{824}[c]	A_{2776}[c]	Genotype	Stp23-7c (SSCP)[a]	A_{2578}[c]
St01	0	0	GGGG	CCCC	GGGG	0	0	St01	0	0
St02	1	0	GGGG	CCCC	GGGG	0	0	St02	1	1
St03	0	0	GGGG	CCCC	GGGG	0	0	St03	0	0
St04	0	0	GGGG	CCCC	GGGG	0	0	St04	0	0
St05	0	0	GGGG	CCCC	GGGG	0	0	St05	0	0
St06	1	1	TGGG	TCCC	AGGG	1	1	St06	0	0
St07	1	1	TGGG	TCCC	AGGG	1	1	St07	1	1
St08	1	1	TGGG	TCCC	AGGG	1	1	St08	0	1
St09	1	1	TGGG	TCCC	AGGG	1	1	St09	0	0
Theresa	1	1	TGGG	TCCC	AGGG	1	1	Theresa	1	1
St11	1	1	TGGG	TCCC	AGGG	1	1	St11	0	0
Saturna	1	1	TTGG	TTCC	AAGG	1	1	Saturna	1	1
St14	0	0	GGGG	CCCC	GGGG	0	0	St14	1	1
St15	0	0	GGGG	CCCC	GGGG	0	0	St15	1	1
St16	1	0	GGGG	CCCC	GGGG	0	0	St16	0	0
St18	0	0	TGGG	TCCC	AGGG	1	1	St18	0	0
St19	1	0	GGGG	CCCC	GGGG	0	0	St19	1	1
St20	1	1	TGGG	TCCC	AGGG	1	1	St20	0	0
St21	0	0	GGGG	CCCC	GGGG	0	0	St21	1	1
St22	1	0	GGGG	CCCC	GGGG	0	0	St22	1	1
St23	1	0	GGGG	CCCC	GGGG	0	0	St23	1	1
St24	1	0	GGGG	CCCC	GGGG	0	0	St24	1	1
St25	1	0	GGGG	CCCC	GGGG	0	0	St25	1	1
St26	0	0	GGGG	CCCC	GGGG	0	0	St26	0	0
St27	0	0	GGGG	CCCC	GGGG	0	0	St27	0	0
Diana	1	1	TGGG	TCCC	AGGG	1	1	Diana	1	1
St29	0	0	GGGG	CCCC	GGGG	0	0	St29	0	0
St30	0	0	GGGG	CCCC	GGGG	0	0	St30	0	0
St31	0	0	GGGG	CCCC	GGGG	0	0	St31	0	0
St32	1	1	TGGG	TCCC	AGGG	1	1	St32	1	1
Satina	0	0	GGGG	CCCC	GGGG	0	0	Satina	0	0
St34	0	0	GGGG	CCCC	GGGG	0	0	St34	0	0
St35	1	1	TGGG	TCCC	AGGG	1	1	St35	0	0
St36	0	0	GGGG	CCCC	GGGG	0	0	St36	0	0

Distribution of the SSCP polymorphisms *Stp23-8b*, *-8a* and *-7c* (Li et al., 2008) and the *Pho1a* cDNA SNPs in a subset of 34 genotypes from the GABI-CHIPS population (Standards). The color codes indicate SNPs of the green or blue group. No SNPs are cosegregating with *Stp23-8a* (grey).

[a] SSCP polymorphism scored as present (1) or absent (0)

[b] SNPs as scored by amplicon sequencing.

[c] SNP scored by SNP-specific primers as present (1) or absent (0)

Given the importance of having accurate haplotype sequences for this study, the cDNA-based SNPs found in the different clones were further compared with those annotated by the Solanaceae Genomics Resources Database from Michigan State University (http://solanaceae.plantbiology.msu.edu/analyses_snp.php) (Table 8). This database uses EST sequences from the PlantGDB-assembled unique transcripts (PUTs) database

(http://www.plantgdb.org/prj/ESTCluster/index.php) for SNP discovery. *Pho1a* corresponds to the assembly PUT-157a-Solanum_tuberosum-72573151. Thus, the presence of the T_{1923} and T_{1939} nucleotides in the same DNA molecule (*Pho1a-H$_A$*), as well as A_{2776} and C_{2823} was confirmed.

Table 9. cDNA SNPs of *Pho1a* in common with the SolCAP database

cDNA SNP Position[a]		$G_{1923}T$	$G_{1939}T$	$G_{2578}A$	$G_{2616}A$	$A_{2765}T$	$G_{2776}A$	$C_{2795}A$	$G_{2823}C$	EST GenBank gi	EST library	Cultivar
SGR DATABASE	SGR SNP Loc.	3825	3841	4480	4518	4667	4678	4697	4725			
	Ref.	G (68)	G (65)	G (51)	G (48)	A (50)	G (49)	C (39)	G (45)			
	SNP	T (6)	T (6)	A (2)	A (3)	T (2)	A (3)	A (10)	C (5)			
										11035743	Mature tuber lambda ZAP	FG Kuras
										78747651	Cold Sweetening C	Shepody
										52826129	After-Cooking Darkening	Shepody
		T	T							53696606	Common Scab-Challenged Tubers	Shepody
		T	T							53695095	Common Scab-Challenged Tubers	Shepody
		T	T							46294374	Suspension culture	Shepody
		T	T							52825579	After-Cooking Darkening	Shepody
		T	T							52824723	After-Cooking Darkening	Shepody
		T	T							52824525	After-Cooking Darkening	Shepody
							A	A	C	78748664	Cold Sweetening C	Shepody
							A	A	C	53696091	Common Scab-Challenged Tubers	Shepody
							A	A	C	62909246	Common Scab-Challenged Tubers	Shepody
								A	C	53702668	Developing Tubers	Shepody
								A		46830105	Tuber Skin	Shepody
								A		15185629	Mature tuber lambda ZAP	FG Kuras
								A		15185610	Mature tuber lambda ZAP	FG Kuras
								A		13179036	Mature tuber lambda ZAP	FG Kuras
				A	A	A		A		78747040	Cold Sweetening C	Shepody
				A	A	A		A		31324489	C33B1bc LIPTC	11379-03
					A					13614315	Sprouting eyes from tubers	Kennebec

SNPs found in the PUT-157a-Solanum_tuberosum-72573151 Multiple Sequence Alignment of the SGR Database.

[a] SNP position from the ATG (Table 7)

[b] Data from the SGR database. The numbers in parenthesis indicate the number times in which the particular base was found in the ESTs. SNPs marked in green and blue show SNP positions characteristic of the groups of the same color described in Table 7

After the correct allele sequences were determined, cDNA clones obtained with different polymerases were compared in order to judge the performance of the enzymes, and to examine which one is less likely to produce the errors previously discussed. The sequences of 12 clones, which were obtained from one genotype (Saturna) using a fast synthesis polymerase (KAPA fast ,1 Kb per second), were not chimerical. Hence, the use of a fast

synthesis polymerase is likely to result in less spurious combinations and is a good choice for cDNA synthesis of large transcripts in studies of this type.

3.1.2 *Pho1a-H_A is a*ssociated with quality traits

The validation of the *Pho1a-H_A* allele as the one underlining the *Stp23-8b* SSCP polymorphism, implied its association with traits including chip quality at harvest and after cold storage, tuber starch content, and tuber starch yield in the GABI-CHIPs population (Table 5). To test whether this correlation was maintained across different populations, further association analyses for the H_A allele were conducted in a Tuber Bruising (TASK) population (Urbany et al., 2011) and a Reducing Sugars Content (BIOSOL) population (Matthias Fischer and Lena Schreiber, personal communication).

3.1.2.1 *Pho1a-H_A* is associated with higher tuber starch content and tuber bruising

The distribution of the *Pho1a-H_A* allele was determined in the 205 tetraploid potato genotypes of the TASK population, previously used to determine marker associations to unravel the genetic basis underlying tuber bruising susceptibility. In this population, three SSCP markers of the *Pho1a* gene (*Pho1A-a, b,* and *c*) were found associated with bruising index, specific gravity (starch content), starch corrected bruising, and tuber shape (Table 10, Urbany et al., 2011).

Specific primers for the SNP A_{2776}[9] (*Pho1a-H_A*) were tested on the 205 individuals of this TASK population. In 97.7% of the cases, the distribution of the *Pho1a-H_A* allele was in agreement with the distribution of the marker allele *Pho1A-c*, with a frequency of appearance in the population of 0.21. The effect of the marker was an increase in starch content and in tuber bruising susceptibility (Table 10, Urbany et al., 2011).

[9] Dsnp2776A_R / Dc2478_F (Table 3)

Table 10. Association of the *Pho1A-c* SSCP marker (corresponding to the *Pho1a-H_A* haplotype) with tuber quality traits in the TASK population (Urbany et al., 2011)

Trait	Allele Effect	p Value
Bruising index	↑ 13.30 ± 3.30	$6.3.10^{-5}$
Specific gravity	↑ 1.21 ± 0.38	0.00203
Starch corrected bruising	↑ 17.60 ± 2.40	0.00172
Tuber shape	↓ 10.45 ± 0.14	0.00225

The direction of the arrow indicates an increasing or decreasing effect of the marker allele on the trait value.

At the genomic level, the SNPs accountable for the *Stp23-8b* (Table 5) and *Pho1A-c* markers were determined by sequencing of genomic DNA amplicons[10] as well as genomic DNA clones. With the exception of SNP $G_{322}A$, which is included only in the amplicon used to generate the Stp23-8 SSCPs, both markers *Stp23-8b* and *Pho1A-c* were based on the same haplotype (Table 11) and therefore represent the same allele (*Pho1a-H_A*).

Table 11. SNPs in the *Pho1a* gene underlying both the *Pho1A-c* (Urbany et al., 2011) and *Stp23-8b* (Li et al., 2008) SSCP polymorphisms (for complete sequence see Appendix 5).

Location of the SNP	SNP denotation	Surrounding sequences[a]	Base call in *Pho1a-H_A*
Exon 2[b]	$G_{322}A$		A
Intron 2	$T_{2\text{-}1}G$	TTAATTTTATG	G
Exon 3	$C_{504}T$		C
Intron 3	$T_{3\text{-}1}G$	CCTAGTTTTCC	T
	$G_{3\text{-}2}T$	TCTGTGTTTTA	G
	$ATTC_{3\text{-}3}GAAT$	TGTGTATTCATTTG	ATTC
	$T_{3\text{-}4}C$	TATAATATTCA	C
Exon 4	$A_{534}G$		G

[a] Surrounding sequences are provided for the intron-SNPs. Positions of the exon-SNPs are indicated in the SNP denotation.
[b] Only in *Pho1A-c*

[10] DgSNPs01_F / DgSNPs01_R (Table 3)

3.1.2.2 *Pho1a-H_A* is associated with lower reducing sugar contents upon cold storage

The distribution of the *Pho1a-H_A* allele was tested with SNP specific primers (A_{2776}[11] polymorphism) on a population of 40 commercial cultivars pre-selected for their superior (cases) and inferior (controls) chipping quality characteristics (BIOSOL population). Total content of reducing sugars was measured for each of the 40 genotypes at different time points of cold storage (Matthias Fischer, personal communication). The H_A allele was mostly present in the genotypes with lower sugar contents at harvest time (7 out of 10, Figure 7).

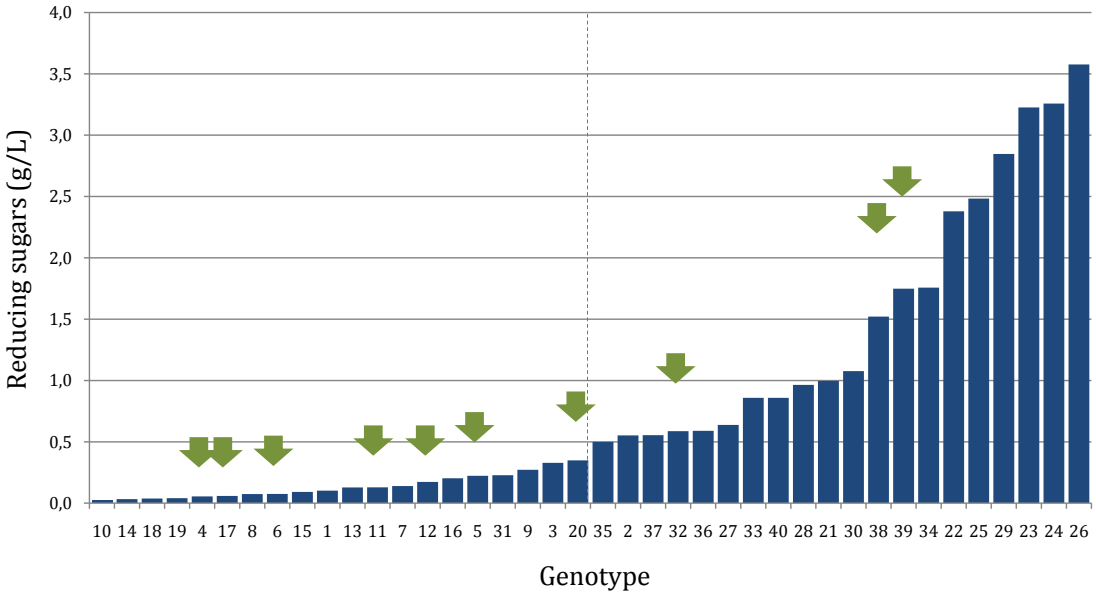


Figure 5. Glucose and fructose (GF) content of potato tubers of different genotypes measured at harvest time in a population of 20 good (numbers 1 to 20) and 20 bad (numbers 21 to 40) chipping quality cultivars. Genotypes bearing the *Pho1a-H_A* allele are indicated with a green arrow. Genotypes are sorted by increasing total content of reducing sugars (glucose and fructose), as measured by enzymatic hydrolysis of sucrose from freeze dried tuber tissue (Matthias Fischer, personal communication).

The accumulation of glucose and fructose as a result of enzymatic degradation during cold storage was investigated by assessing the amount of reducing sugars present after two and four weeks of cold storage (WCS, Figure 6). Generally, as expected, reducing sugars levels were found to increase during cold storage. Most of the genotypes bearing the associated

[11] Dsnp2776A_R / Dc2478_F

allele clustered to the good cold storage genotypes (low sugar increase, Figure 6), suggesting a positive association of the marker with this trait.

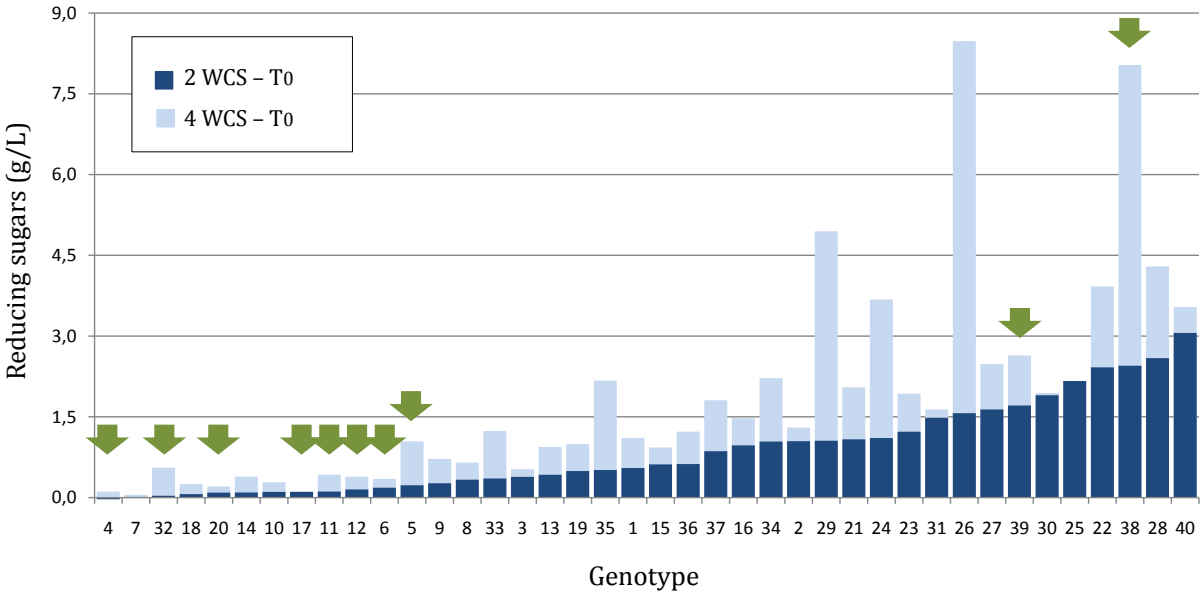


Figure 6. Cold induced accumulation of GF in potato tubers after cold storage. Each value is calculated as the GF amount at 2 and 4 weeks after colds storage (WCS) minus the GF content at harvest time. Genotypes are sorted by sugar accumulation at 2 WCS. Distribution of the H_A allele is represented with a green arrow.

Cultivars bearing the associated allele H_A contain on average significantly lower amounts of sugars as confirmed by means comparison analysis (Figure 7B). Mean comparisons were performed on the entire population, as well as for the subgroup of the 20 good chipping cultivars. Since the bad chipping cultivars group only contained two individuals with the marker, mean comparisons were not done for this subgroup. The analysis confirmed that among genotypes which accumulated low sugars; the average of those containing the associated allele was still significantly lower. Analysis of variance suggested that the H_A allele explained between approx. 14% and 22% of the variance in sugar contents at 2 and 4 weeks after cold storage in low sugars genotypes.

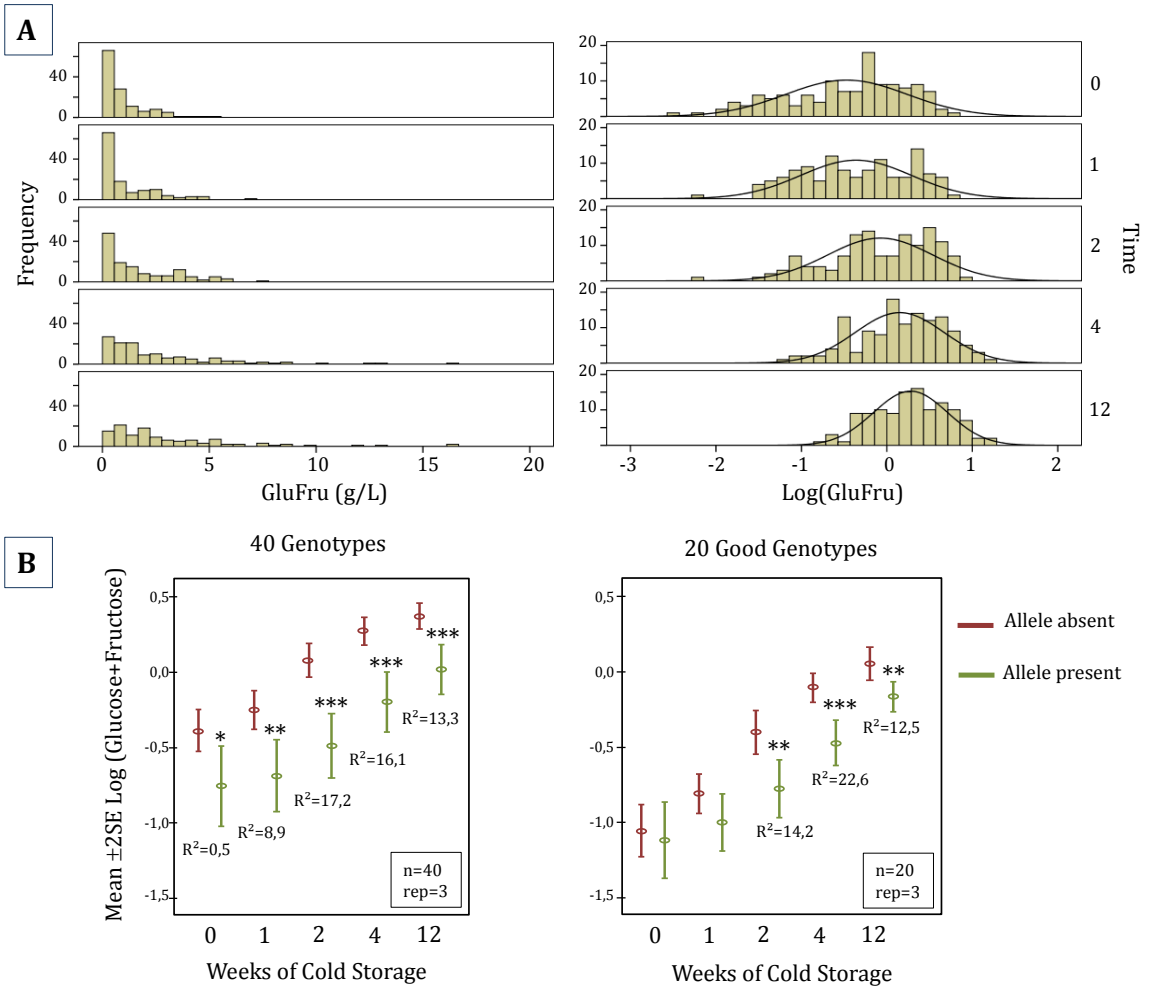


Figure 7. **A**. Distribution of the original sugar measurements (left panel) and the transformed data (right panel) at the different time points after cold storage (0 to 12 weeks). The fitted normal curves are displayed. **B**. Sugar content means in 40 potato cultivars according to the presence of the H_A haplotype (left panel), and in the 20 good chipping cultivars (right panel). Significance levels are displayed by stars and the R^2 values are given in percentage.

3.1.3 *Pho1a* promoter analysis

In order to find sequence polymorphisms in the promoter region linked to the H_A haplotype, PCR primers were designed to amplify a ~2 Kb fragment[12] upstream of the *Pho1a* coding region. For this purpose, the sequence information available in GeneBank (AF143202.2) was used. Promoter amplicons were obtained from cultivars Diana, Theresa, Saturna and Satina as a control (negative for *Pho1a*-H_A).

[12] DP01_F / DP01_R

In addition to the expected 2 Kb PCR fragment; a 1 Kb polymorphic fragment (DP01) was amplified in cultivars Diana, Theresa and Saturna. This polymorphic band was not associated with tuber quality traits in the GABI-CHIPS population. The distribution of the marker in the population corresponded to that of the *Stp23-7c* SSCP marker represented by the group of alleles classified in the blue group (Figure 8, Table 7).

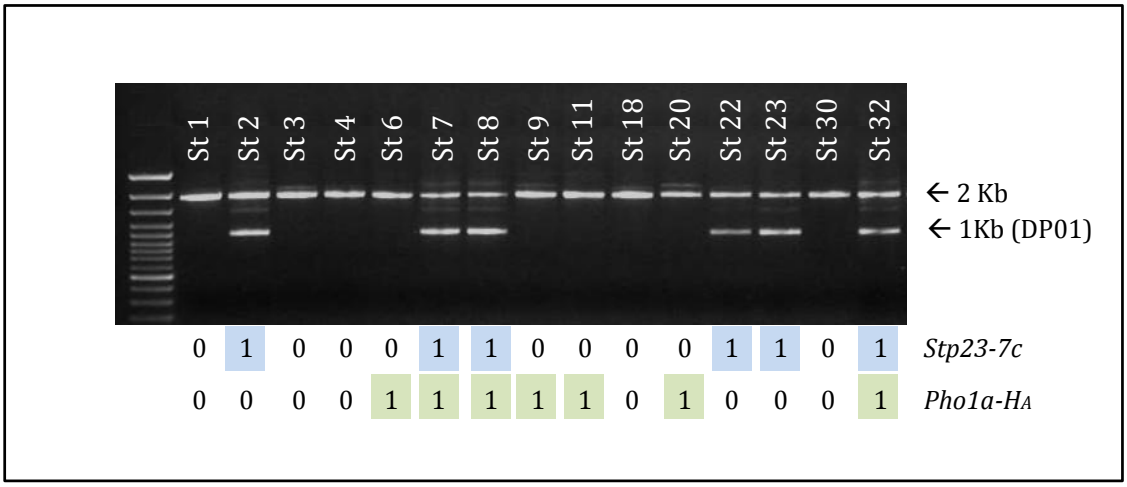


Figure 8. Distribution of the DPO1 PCR fragment in the *Pho1a* promoter. The 15 individuals tested belong to the Standards subpopulation. The distribution of this marker is in agreement with that of the SSCP marker *Stp23-7c*, but differs from that of the *Pho1a-H*$_A$ allele (Table 8).

The promoter PCR products from cultivars Diana, Theresa, Saturna, and Satina were sequenced. Aligning and comparison of the sequences revealed four Indels (one insertion [8 bp] and three deletions [1 Kb, 7 bp and 5 bp]), in addition to 10 nucleotide substitutions (Table 12). Cultivar Satina was homozygous (Appendix 4).

The distributions of the promoter polymorphisms DP01[13] and DP02b[14] (Table 12) were tested in the BIOSOL population for their associations with sugars content (Figure 9). No obvious correlation was found between the DP01 deletion and the reducing sugars content at any time point of cold storage. The DP01b polymorphism was more frequent (6 out of 9) in genotypes with higher sugar content at harvest time. However, no relation to sugar

[13] DP01_F / DP01_R
[14] DP02b_F / DP02_R

accumulation during cold storage could be found (Figure 10). Thus, the polymorphism was not further analyzed.

Table 12. DNA polymorphisms in the *Pho1a* promoter.

Location[a]	Mutation[b]	Type
-1888	A/G	Substitution
-1845	CCAACATG	Insertion (DP02b)
-1828	G/A	Substitution
-1730	G/A	Substitution
-1633	G/A	Substitution
-1533	C	Conflict [c]
-1520	G	Conflict
-1464	C/T	Substitution
-1464 to -590	1Kb	Deletion (DP01)
-1437	C	Conflict
-562 to -567	T(6)/A(6)	Substitution
-557 to -560	T(3)-A(3)	Substitution
-556	T/A	Substitution
-452	G/A	Substitution
-412	T/C	Substitution
-335	T/C	Substitution
-304	G/A	Substitution
-299	T	Deletion
-262 to -268	AGAGTGT	Deletion
-225	A/C	Substitution
-180	C/A	Substitution
-143 to -147	AATTT	Deletion

[a] Position upstream of the ATG codon
[b] Sequence polymorphisms between alleles from Diana, Theresa and Saturna in comparison with that from Satina
[c] Conflict are bases in the sequences that differ from the GenBank sequence (AF143202.2). For complete alignment see Appendix 4.

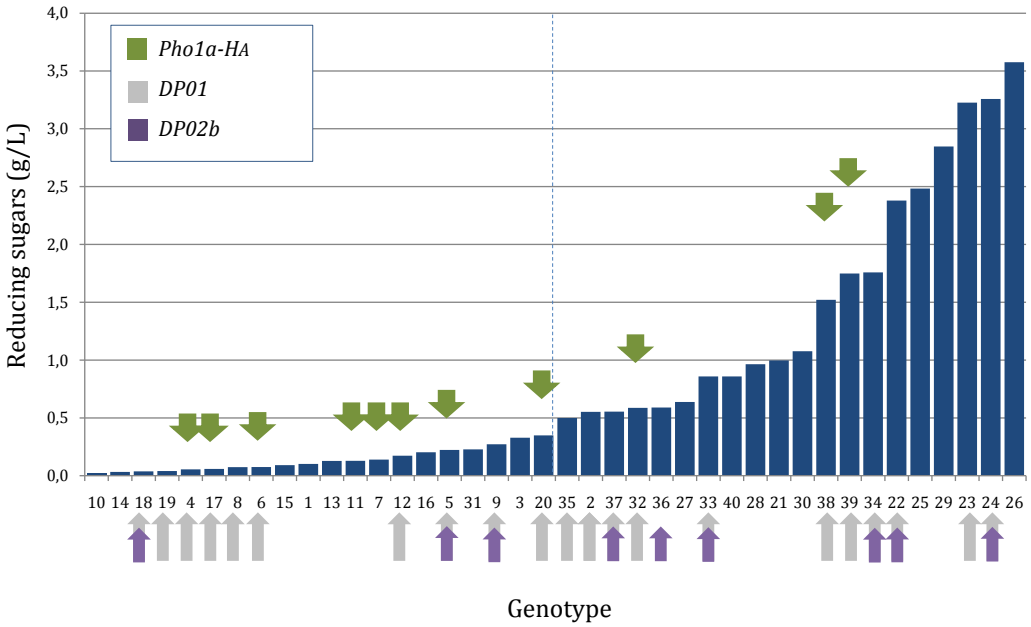


Figure 9. Distribution of the *Pho1a* promoter polymorphisms *DP01* and *DP02b* in the BIOSOL population. Individuals were sorted according to their reducing sugar content at harvest time.

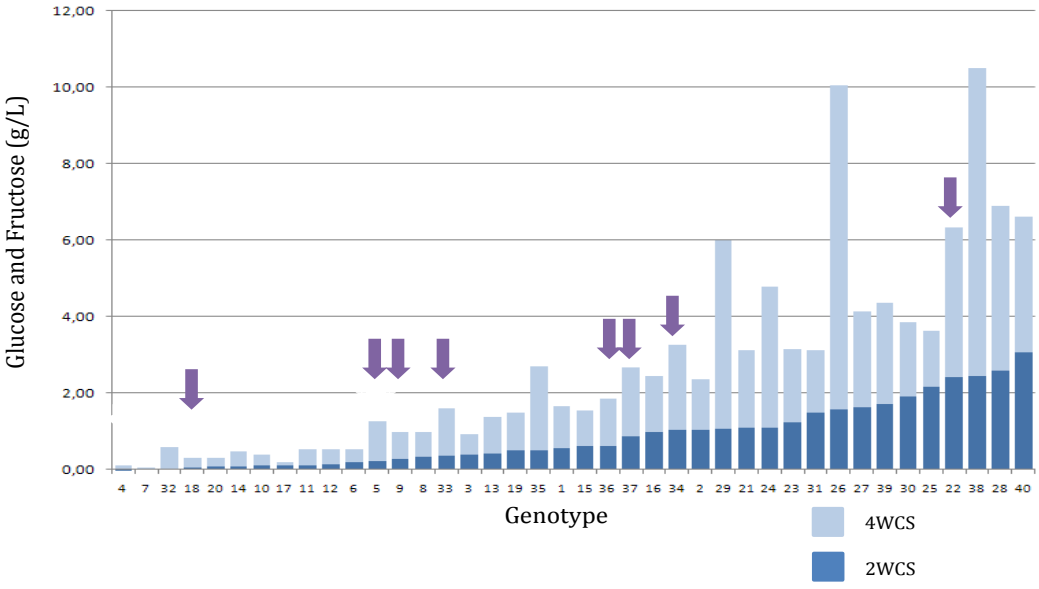


Figure 10. Distribution of the *Pho1a* promoter polymorphism DP02b in the BIOSOL population. Individuals were sorted according to the reducing sugar content accumulated after 2 weeks of cold storage (WCS).

3.1.4 Expression analysis of *Pho1a* and *Pho1a-H_A* specific expression

3.1.4.1 *Pho1a* expression profile upon cold storage

To determine the overall expression levels of the starch phosphorylase on chromosome III, the transcription level of *Pho1a*[15] was analyzed by quantitative reverse transcription PCR (qRT-PCR). For this purpose, expression levels were compared on two low- and high-sugar genotypes (cultivars 17 and 26, respectively; Figure 5) from the BIOSOL population, at 0, 1, 2, and 4 weeks after cold storage. The low-sugar genotype bears the H_A allele in duplex dose as indicated by scoring of genomic DNA sequences[16] (Figure 11B)

Higher *Pho1a* expression levels were observed at time point 0 (before cold storage) compared to later time points in both genotypes, more pronounced in Cv17 than in Cv26. At the initial time point also comparison between the two genotypes showed significant differences in *Pho1a* expression levels ($P<0.05$), with the high-sugar genotype having approximately a twofold higher relative expression than the low-sugar genotype. The down-regulation of *Pho1*a transcripts after cold storage is therefore stronger in the high-sugar genotype, but no relevant differences were significative among the genotypes after cold storage.

The difference observed before cold storage and similarities in subsequent time points are consistent with results of the sugars content measurements (Figure 5 and Figure 6), which showed that the overall reducing sugar content before cold storage is a determinant of the further accumulation of sugars after cold storage. Therefore, the expression profile of *Pho1a* at this time was further analyzed in eight additional genotypes, in a total of five low-sugar (4, 7, 10, 17 and 18) and five high-sugar cultivars (22, 24, 26, 29, 38) (Figure 12).

[15] Dpyr3(D926N)_F/ Dpyr3(D926N)_R
[16] Pyr1_F/Pyr1_R, gSNPs3_F/gSNPs3_R, and Pyr3_F/Pyr3_R in genomic DNA

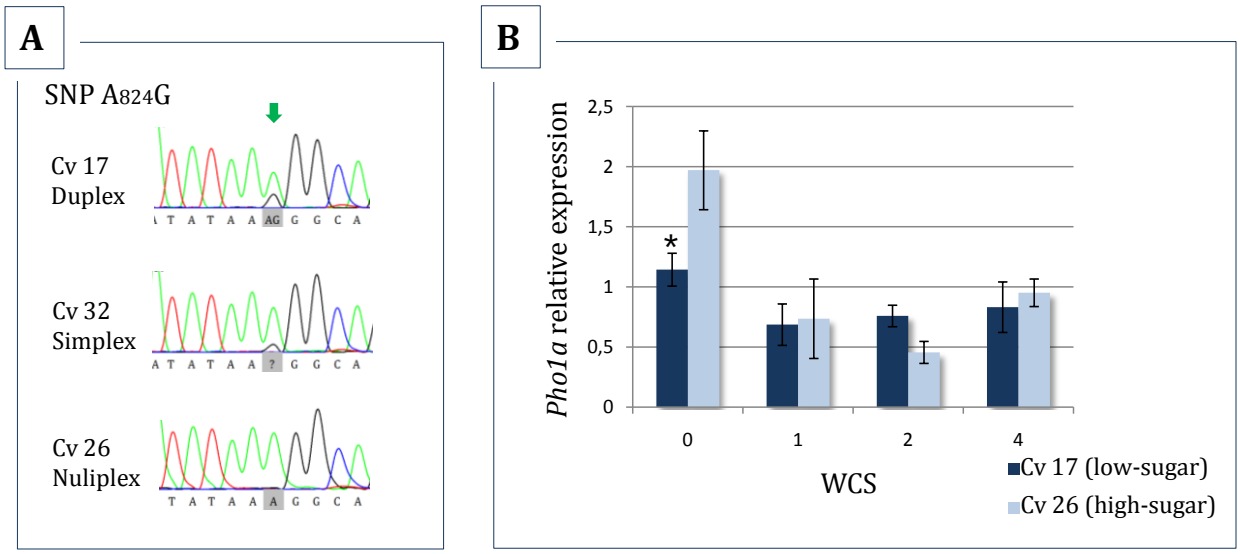


Figure 11. **A**. Dosage scoring for the $A_{824}G$ SNP (green arrow) in cv 26 (nuliplex for H_A) and 17 (duplex). Cultivar 32 was included for the visualization of a simplex SNP. **B**. Expression of *Pho1a* in a low (cv 17) and a high sugar (cv 26) cultivar during cold storage. Each bar represents the mean of three biological and three technical replications, normalized with *EF1-a* expression. Standard error is displayed for each sample.

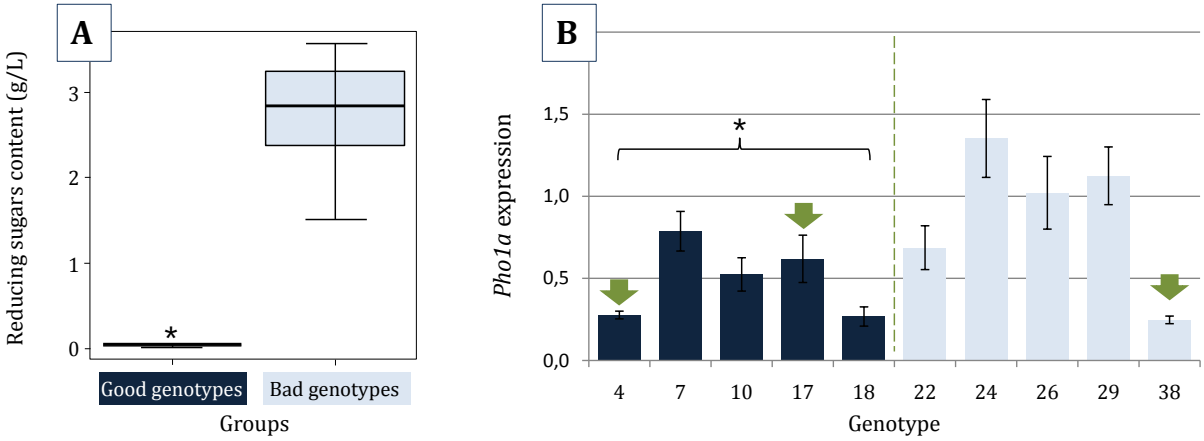


Figure 12. Relation of sugar content and *Pho1a* expression in five good quality (low sugars) and five bad quality (high sugars) potato genotypes before cold storage. **A**. Mean differences in the sugar content between groups (n is the mean of five genotypes and three (or four) technical replicates. **B**. Normalized expression of *Pho1a* in five low-sugars (dark blue) and five high-sugars potato cultivars, respectively. The green arrows denote the presence of the *Pho1a-H_A* allele in the genotype.

T-test for equality of means showed a significant difference among the mean relative expressions of the genotypes bearing the *Pho1a-H_A* allele and the ones that do not contain it ($P < 0.05$). However, the fourty-fold difference observed between sugar content means of good and bad quality genotypes (Figure 12A) was not proportional to the overall expression differences of starch phosphorylase (Figure 12B).

3.1.4.2 Expression analysis of the *Pho1a-H_A* allele upon cold storage

To address if the H_A allele is differentially regulated upon cold storage, the differential expression between the *Pho1a-H_A* and other *Pho1a* alleles was determined by pyrosequencing. Two cultivars bearing the H_A allele in a simplex (cv 4) and duplex (cv 17,) dosage (Figure 13) were used for comparison of transcript levels in cold stored tubers. Cultivars 10 and 7, which do not contain the H_A allele, were included as a control.

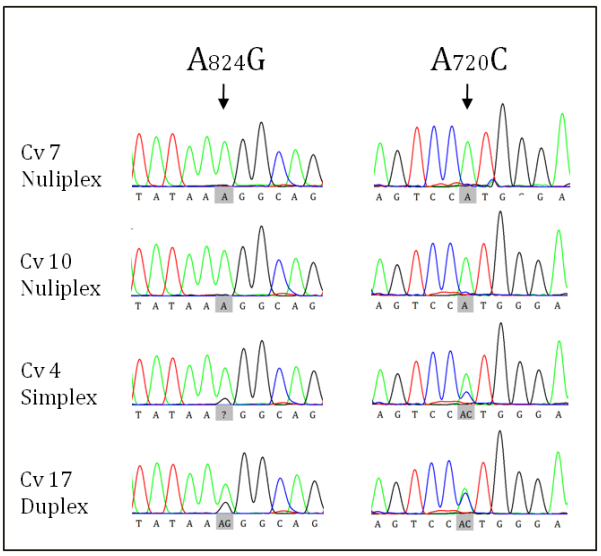


Figure 13. Two SNPs of the *Pho1a-H_A* allele used to determine allele dosage in the individuals chosen for pyrosequencing assays.

Pyrosquencing was performed for two SNPs of the H_A allele, A_{2776}[17] and G_{824}[18] (Figure 14). In a tetraploid genotype, a frequency of 25% is expected for a simplex allele that is not undergoing regulation at the transcription level. Likewise, a duplex allele is expected to appear in approx. 50% of the total transcripts. However, the expression levels for the A_{2776}

[17] Amplicon primers Dpyr3(D926N)_F / Dpyr3(D926N)_R(biotin). Sequencing primer DpSeq3(D926N)_S (Table 3)
[18] Amplicon primers Dpyr2_F(K275R)(biotin) / Dpyr2_R(K275R). Sequencing primer DpSeq2(K275R)_S (Table 3)

SNP reached only 12% and 30% in the simplex and duplex genotypes, respectively (Figure 14). The A_{2776} sequencing primer included part of intron 14 which impeded its evaluation in the genomic DNA. Pyrosequencing of the SNP $G_{824}A$ allowed the comparison of expression levels and genomic DNA dosage (Figure 14B). Allele ratios in cDNA did not vary in comparison with the rations in genomic DNA. Because the expression of the H_A-haplotype did not change upon cold storage, it was deduced that the *Pho1a*-H_A allele is not regulated at the transcriptional level.

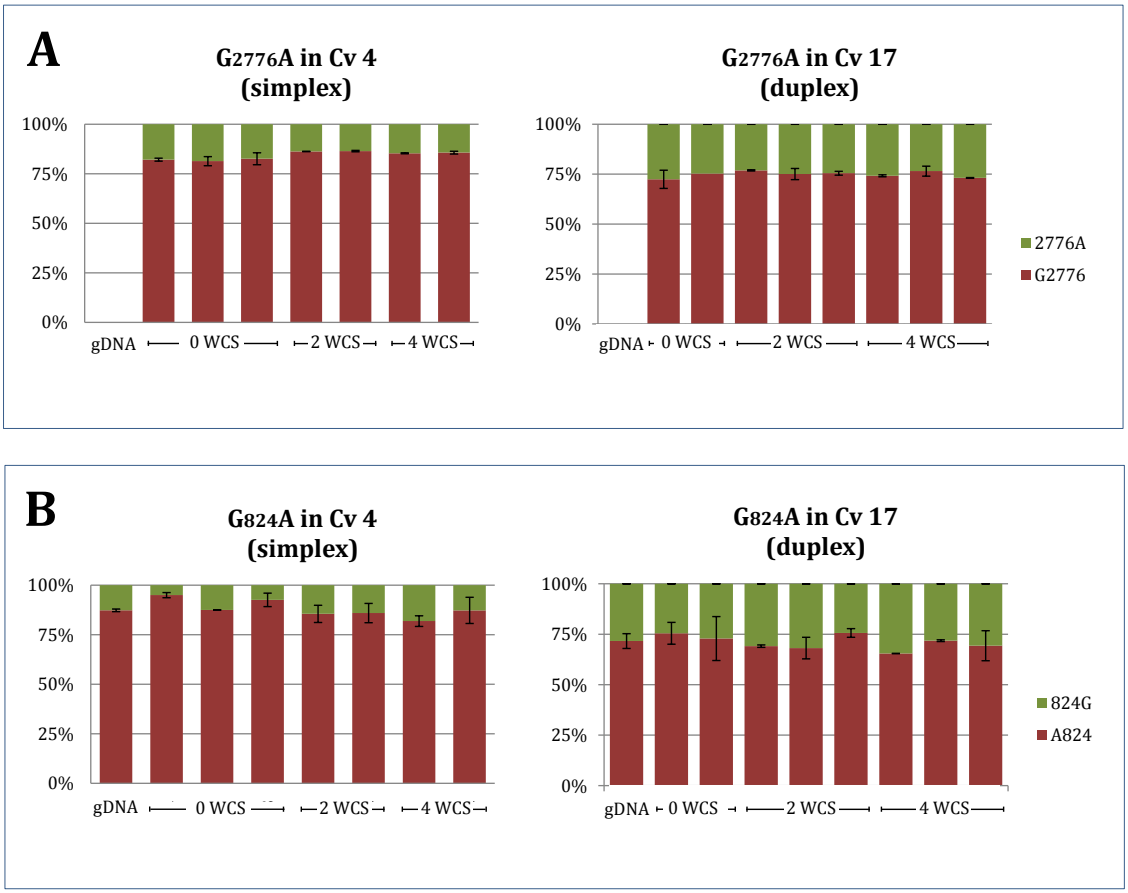
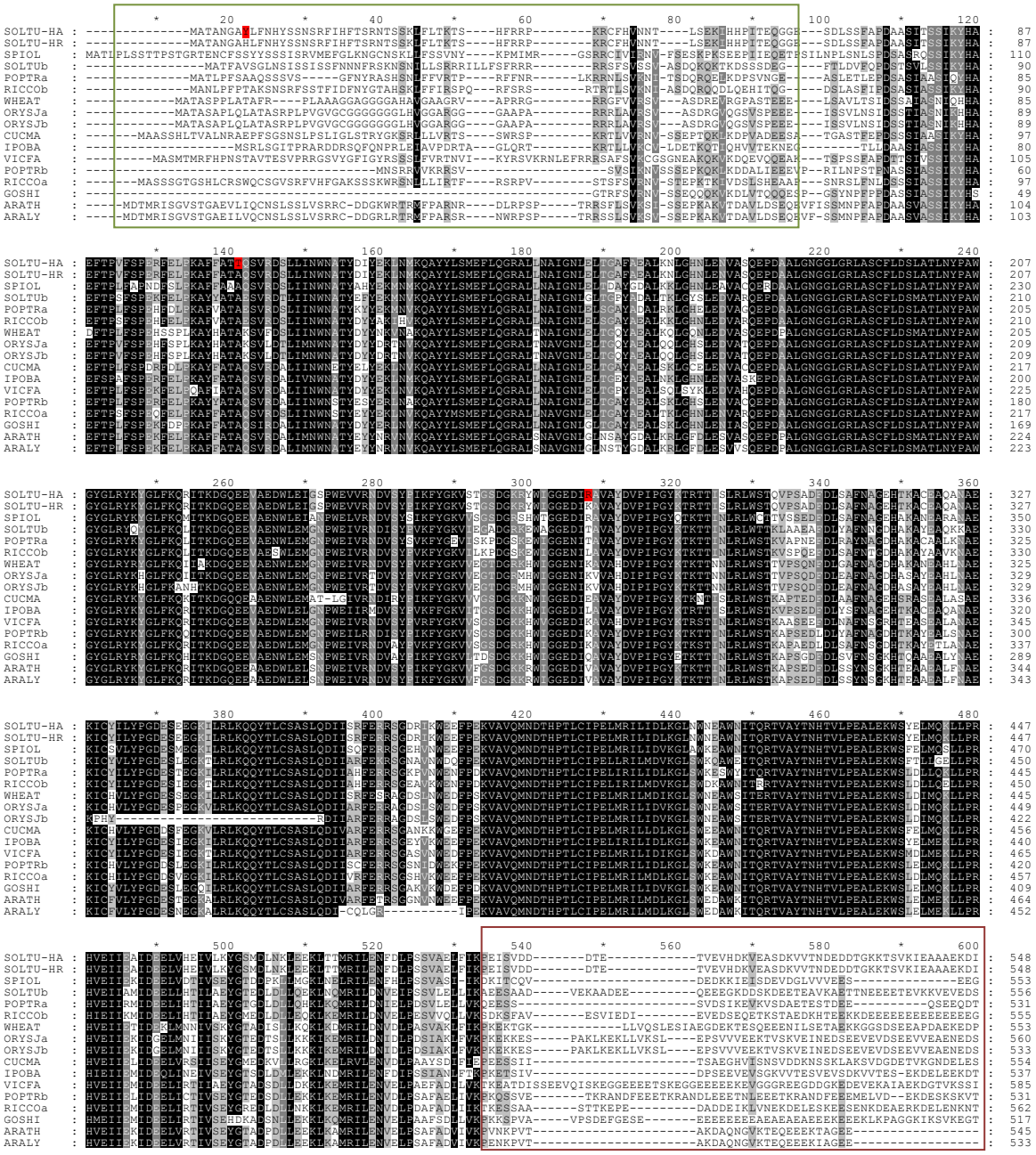


Figure 14. Differential expression of the *Pho1A*-H_A allele during cold storage (WCS; biological replicates marked for each case) in one H_A-simplex (cv 4) and one H_A-duplex (cv 17) genotype, as resolved by pyrosequencing. Green bars (average of three technical replicates; error bars displayed) represent the proportion of the H_A-SNP within the total *Pho1a* transcript. In **A**) and **B**), SNPs $G_{2776}A$ and $G_{824}A$ of the H_A allele are presented, Genomic DNA is denoted as gDNA.

3.1.5 Functional characterization of *Pho1a* alleles

To investigate whether the amino acid changes introduced by the nucleotide polymorphisms were located in areas of high or low conservation in the PHO1a protein, all available plant plastidial phosphorylase protein sequences were retrieved from Uniprot (The UniProt Consortium). A MUSCLE sequence alignment of these sequences (Edgar, 2004) showed a high degree of conservation of this protein across species, with the exception of the regions corresponding to the N-terminal signal peptide and the area that has been described as the 'central insertion' (Nakano and Fukui, 1986) (Figure 15).



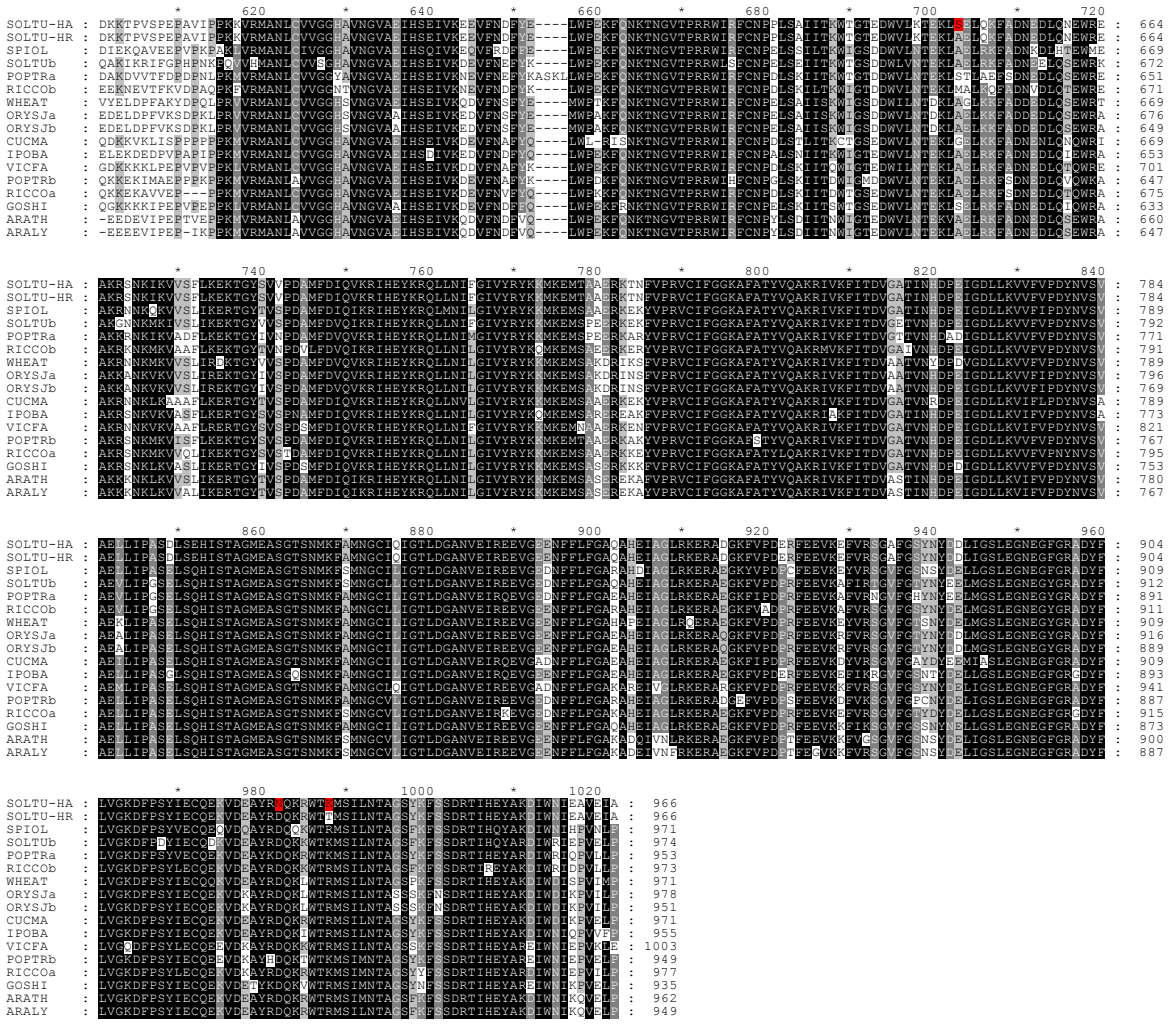


Figure 15. MUSCLE alignment of plant PHO1a protein sequences. SOLTU, *Solanum tuberosum* (P04045); SPIOL, *Spinacea oleracea* (O24363); POPTR, *Populus tricocarpa* (a, B9HXL0; b, B9H0D3); RICCO, *Ricinus communis* (a, B9SJB6; b, B9RCW0); WHEAT, *Triticum aestivum* (B2LXU4); ORYSJ, *Oryza sativa* subsp. Japonica (a, B3IYE3; b, Q9AUX8); CUCMA, *Cucurbita maxima* (B2DG13); IPOBA, *Ipomea batatas* (P27598); VICFA, *Vicia faba* (P53536); GOSHI, *Gossypium hirsutum* (D2D337); ARATH, *Arabidopsis thaliana* (Q9LIB2); ARALY, *Arabidopsis lyrata* (D7LNX4). The sequence of the protein encoded by *Pho1a-*H_A is placed in the first row and the amino acid changes with respect to *Pho1a-*H_R are indicated in red. Green and blue boxes indicate the transit peptide and the central insertion, respectively.

The first amino acid substitution from the *Pho1a-*H_A allele occurs in the in the N-terminal transit peptide (Figure 15, Nakano et al., 1989). The $C_{22}T$ nucleotide change results in the substitution of a histidine residue by tyrosine in the eighth position (His-8-Tyr) from the starting methionine.

The second amino acid change is an Ala-108-Thr substitution at position 108. Alanine is a 100% conserved residue not only in plant starch phosphorylases (Figure 16) but in animal alpha-glucan phosphorylases described so far. Sequence alignment of more than 200 amino acid sequences from plants, animals, fungi, and bacteria (retrieved from NCBI by the BLASTn algorithm), showed that this substitution has no precedents in alpha-glucan phosphorylases (Figure 16). In this position, a serine residue was occasionally found in proteins from fungal and bacterial origin. The strong conservation observed at this position makes the Ala-108-Thr substitution a likely candidate for an altered enzyme activity.

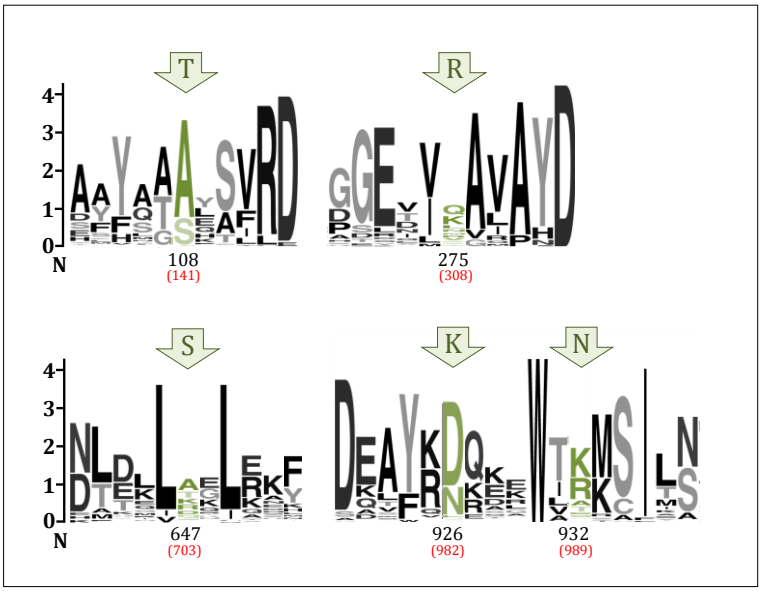


Figure 16. Comparison between protein sequences of glucan phosphorylases from approx. 200 different species (plants, animals, bacteria, fungi). Each position showing SNPs between H_A and H_R sequences is indicated by an arrow, showing the substitution introduced by the *Pho1a*-H_A allele (the black number below points out its position in *Pho1a*; the red number its position in Figure 15. The substitutions are surrounded by 5 amino acids on each side. The relative frequency with which an amino acid appears at a specific position, is given by the height of its one-letter code in the logo (Crooks et al., 2004). The total height at a given position is proportional to the level of sequence conservation. The mutation (Ala-$_{108}$-Thr) has no precedents in alpha-glucan phosphorylases. Sequence alignments were manually edited using Mega 3.0 (Kumar et al., 2004)

Potential effects of this point mutation were analyzed *in silico*. The crystal structure of plant starch phosphorylase has not been determined; however, a variety of structures for starch phosphorylase homologues, such as glycogen- and maltodextrin-phosphorylases from animals and bacteria, respectively, are available in the protein data base. Predictions of the

functional consequences of the alanine to threonine substitution at the 56th residue of the rabbit glycogen phosphorylase B (PDB: G6PB), which corresponds to the Ala-108 residue in PHO1a, were evaluated by the AUTO-MUTE algorithm (Masso and Vaisman, 2010). Results suggested with 84% confidence that the mutation decreases the stability of the protein.

The Ala-56-Thr mutation in rabbit muscle glycogen phosphorylase (GP) and Ala-108-Thr mutation in PHO1a (modeled from yeast glycogen phosphorylase (YGP), Section 2.2.12; sequence alignment in Appendix 3) were simulated with PyMOL to visualize the possible cause of the deleterious effect of the mutation (Figure 17). The residue is located on the second alpha-helix. Protein dimers are formed by the interaction of the subunits at two different interfaces: (1) the cap region of the first subunit interacts with the α2-helix and the β7-strand of the second subunit (denoted as the cap'-α2-β7 interface) and (2) the tower'-tower (α7'-α7) helices interface (Watson et al., 1997). It is possible that the introduction of a hydrophilic amino acid in a predominantly hydrophobic environment may alter the position or conformation of the alpha-helix by establishing novel polar interactions.

The Lys-275-Arg substitution occurred in a non-conserved position in the protein (Figure 17). Polar interactions remained unchanged when the residue was mutated, as lysine and arginine are both positively charged amino acids (Figure 18). The residue is located at the interface of the tower region. Therefore it may not affect protein stability but still interfere with dimerization.

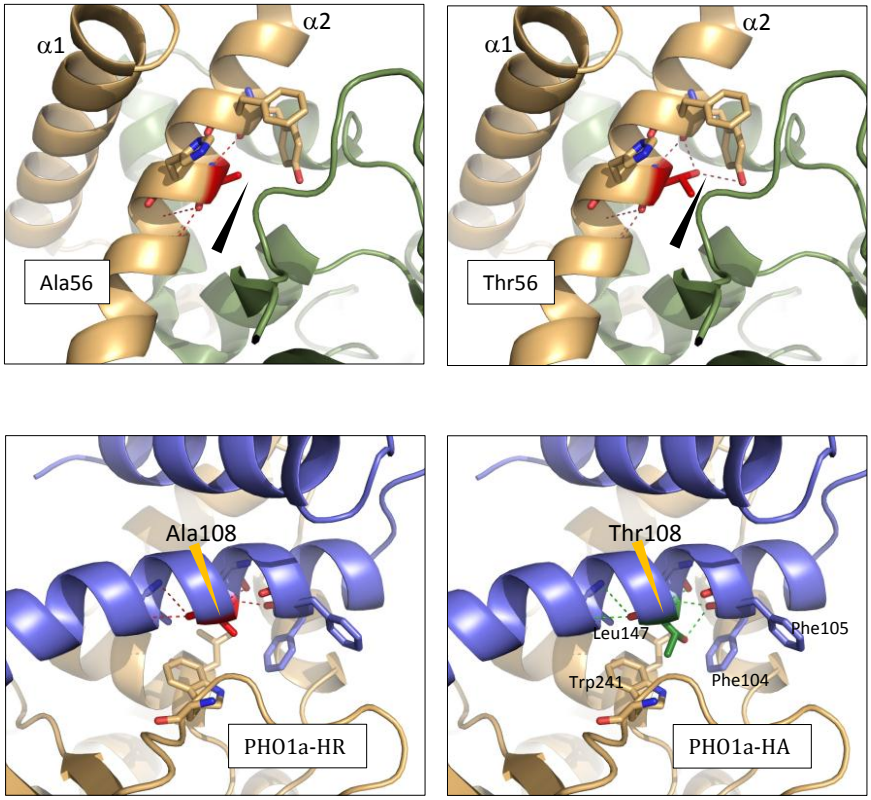


Figure 17. *In silico* mutation of the Ala56 residue in rabbit muscle glycogen phosphorylase (GP, upper panels) and the Ala108 residue in potato starch phosphorylase (PHO1a, lower panels). The black arrows show novel polar interactions introduced by the mutation in GP. Yellow arrows point at the amino acids at the 108 position in PHO1a

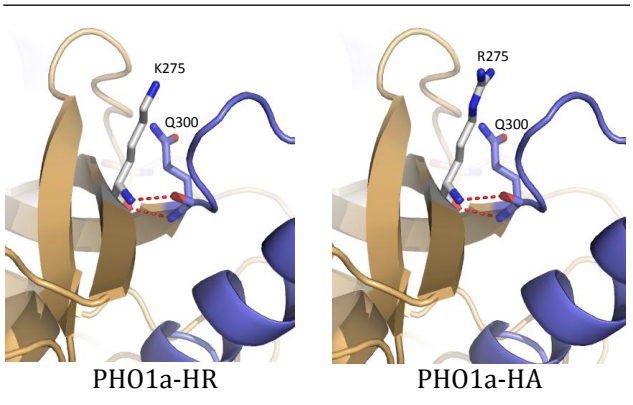


Figure 18. Lys-275-Arg amino acid substitution in PHO1a

The last three mutations on the *Pho1a-H$_A$* allele are Ala-647-Ser, Asp-926-Asn, and Thr-932-Lys. Despite the 285 residues that separate Ala-647 and Thr-932, they are in close proximity in the protein. These residues are located on the protein surface. The positions where the mutations are located are well conserved across plant phosphorylases, but not so in alpha glucan phosphorylases in general (Figure 15 and Figure 16). Protein modeling shows that polar interactions are modified which may loosen the contact between the helices (Figure 19).

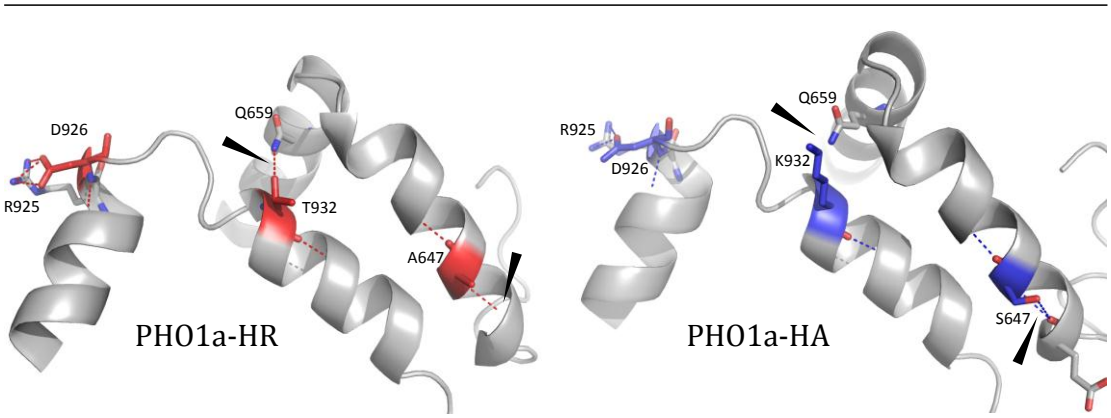


Figure 19. Amino acid substitutions introduced by the *Pho1a-H$_A$* allele. Ala-647-Ser, Asp-926-Asn, and Thr-932-Lys. The black arrows indicate polar interactions that are likely to be modified in the two isoforms

3.1.5.1 The PHO1a-H$_A$ protein isoform is present in the tuber proteome.

To verify whether the protein encoded by the *Pho1a-H$_A$* allele is translated and present in the tuber proteome, total tuber protein extract from the genotype cv 17 (BIOSOL population), carrying the H$_A$ allele in duplex dosage, was separated by two-dimensional (2D)-SDS PAGE. The protein spots corresponding to the plastidial starch phosphorylase (Figure 20) was identified by mass spectrometry (MS) with the collaboration of the Mass Spectrometry facility at the Max Planck Institute for Plant Breeding Research.

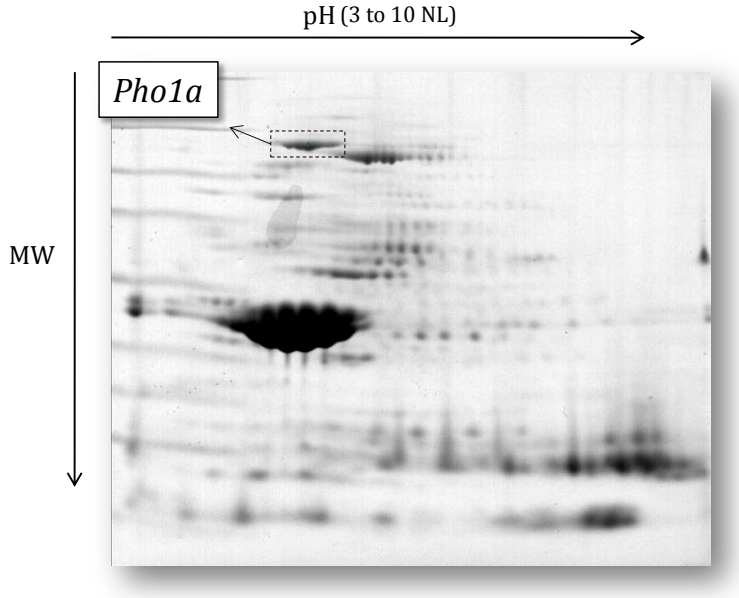


Figure 20. 2D-SDS Electrophoresis of total protein extract from the potato cultivar 17 of the BIOSOL population (duplex for the *Pho1a-H$_A$* allele). Proteins were separated in the first dimension by IEF and in the second dimension by molecular weight. The spots for the PHO1a protein are marked in a dashed box.

In silico trypsin digestion of isoforms PHO1a-H$_R$ and -H$_A$ resulted in 74 peptides, five of which allow the discrimination between the isoforms due to the changes in masses or in the restriction patterns introduced by the amino acid changes (Table 13).

Table 13. Peptides of different masses between PHO1a-H$_R$ and PHO1a-H$_A$ isoforms as obtained by *in silico* trypsin digestion

PHO1a-H$_R$			PHO1a-H$_A$		
Mass	Position	Peptide sequence	Mass	Position	Peptide sequence
1977.9	1-18	MATANGAHLFNHYSSNSR	2003.9	1-18	MATANGAYLFNHYSSNSR
1097.6	103-112	AFFATAQSVR	1127.6	103-112	AFFATTQSVR
1080.5	267-275	YWIGGEDIK	1108.5	267-275	YWIGGEDIR
701.4	646-651	LAELQK	717.4	646-651	LSELQK
1572.8	930-943	WTTMSILNTAGSYK	1184.6	933-943	MSILNTAGSYK

When digested with trypsin, the gel-extracted spot yielded 52 monoisotopic peptides that matched the accessions gi|130173 and gi|21579 from GeneBank, with a protein score of 322.

Results from the MS/MS experiment show that all main peaks in the spectra could be traced back to peptides of the PHO1a protein (Figure 21, Table 14). The masses of the two peaks that were not assigned to peptides in the reference database (peaks of mass 1108 and 1127) (Table 14) corresponded to those of the peptides YWIGGEDIR and AFFATTQSVR from the PHO1a-H_A isoform. Thus, the data confirmed the presence of the PHO1a-H_A protein isoform in the tuber proteome.

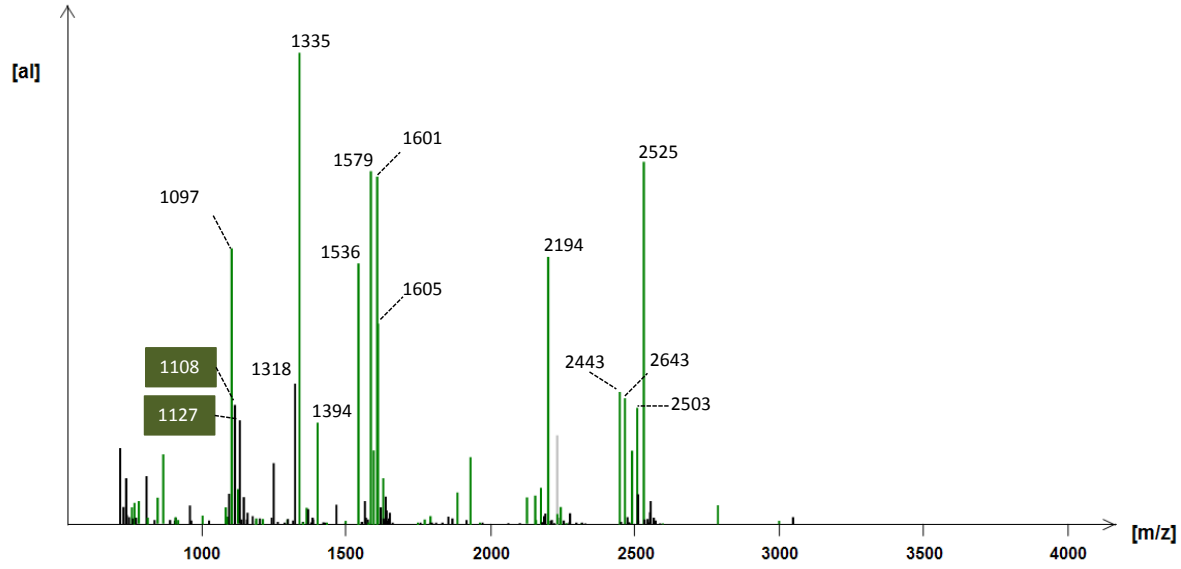


Figure 21. Peptide Mass Fingerprinting (PMS) of PHO1a peptides. All main peaks were traced back to peptides of the PHO1a protein (Table 14). The two peaks marked in a green box fit peptides from the *Pho1a-H_A* allele. The numbers in the mass spectrum show the precise m/z (mass-to-charge ratio) values of the detected peptides.

Table 14. Strongest peptide peaks of the Pho1a mass spectra

m/z acq[a]	m/z theor[b]	Error [ppm]	Height	Protein isoform	Sequence
1335.7	1335.8	-26.2	48107	H_R/H_A	QLLNIFGIVYR
2525.1	2525.1	-20.1	36976	H_R/H_A	SGAFGSYNYDDLIGSLEGNEGFGR
1579.7	1579.8	-22.2	36009	H_R/H_A	YHAEFTPVFSPER
1601.8	1601.8	-19.1	35485	H_R/H_A	GLNWNEAWNITQR
1097.5	1097.6	-28.9	28150	H_R	AFFATAQSVR
2194.0	2194.1	-24.1	27310	H_R/H_A	YHAEFTPVFSPERFELPK
1536.6	1536.7	-21.4	26657	H_R/H_A	FADNEDLQNEWR
1605.7	1605.8	-23.7	20453	H_R/H_A	QAYYLSMEFLQGR
2443.1	2443.1	-26.3	13562	H_R/H_A	DGQEEVAEDWLEIGSPWEVVR
2463.1	2463.2	-28.6	12877	H_R/H_A	EEVGEENFFLFGAQAHEIAGLR
1108.5[c]			12154	H_A	YWIGGEDIR
2503.2	2503.2	-27.1	11861	H_R/H_A	NLGHNLENVASQEPDAALGNGGLGR
1127.6[c]			10599	H_A	AFFATTQSVR
1394.7	1394.7	-21.4	10440	H_R/H_A	FVPDERFEEVK
2225.0	2225.1	-37.9	9131	Trypsin	LGEHNIDVLEGNEQFINAAK + Methyl

Matched peptides ranked by intensity.
[a, b] mass-to-charge ratio acquired, calibrated, and theoretical, respectively
[c] Masses of peptides of the *Pho1a-H_A* allele explain the two strongest unexplained peaks.

3.1.6 Starch phosphorylase activity upon H_A allele dosage

Once the presence of the protein encoded by the *Pho1a-H_A* allele was confirmed, experiments were performed to determine whether potato genotypes bearing this allele showed differential PHO1a activity in comparison to genotypes where the allele is not present. For this purpose, total native protein was extracted (before cold storage) from tubers of cultivars Diana, Theresa and Saturna which contain the *Pho1a-H_A* allele and from cultivar Satina as a control. Native-PAGE separation followed by a phosphorylase activity assays (Section 2.2.15) measured the ability of the PHO1a proteins contained in the tuber extracts to synthesize starch using glucose-1-phosphate. Results showed that the duplex genotype (Saturna) had barely detectable levels of activity when compared to the control (Satina) (Figure 22).

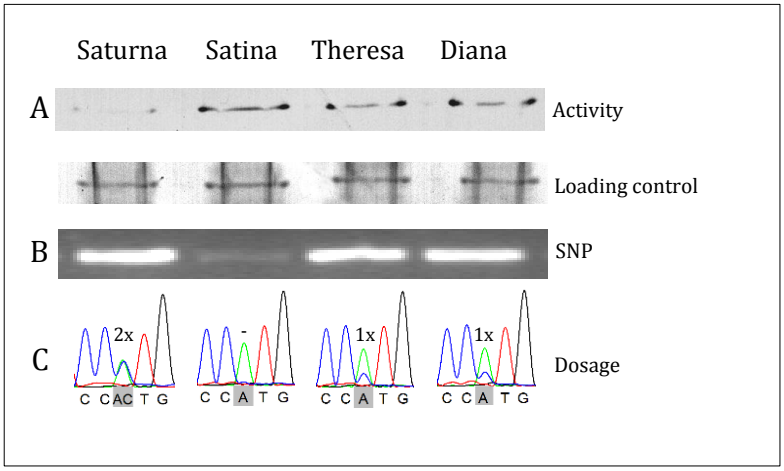


Figure 22. PHO1a activity in four genotypes bearing the *Pho1a-H*$_A$ allele in different dosages. **A**) Activity as measured by the amount of starch (band) synthesized from G1P. **B**) Presence or absence of the *Pho1a-H*$_A$ allele as scored by SNP specific primers. **C**) Estimation of the dosage of the A$_{720}$C SNP

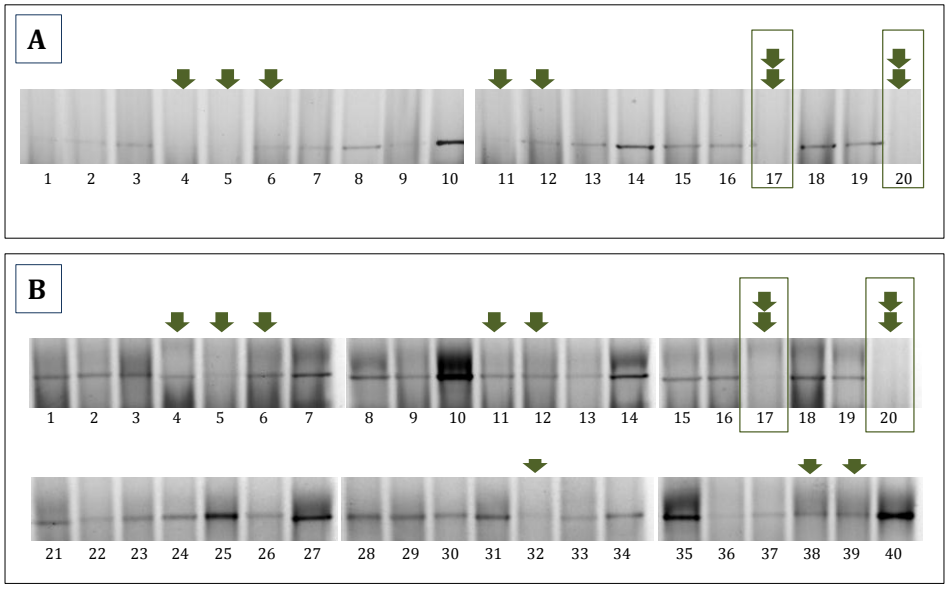


Figure 23. Starch phosphorylase activity in 40 genotypes (BIOSOL population). Total protein was separated by Native-PAGE. Gels were incubated with G1P and starch to prime the reaction. Starch synthesized by starch phosphorylase was stained with iodine and visualized as a band. Results were consistent in two replications. Each green arrow indicates one copy of the *Pho1a-H*$_A$ allele in the respective genotype. **A**) 20 good chipping quality genotypes before cold storage (40 ug of total protein per lane). **B**. 20 good (1-20) and 20 bad (21-40) chipping quality genotypes two weeks after cold storage (WCS) (50 ug of total protein per lane). Genotypes 17 and 20 contain two copies of this allele.

Given the activity differences observed among the four genotypes, the assay was extended to the 40 genotypes from the BIOSOL population at two weeks after cold storage. Genotypes 17 and 20 from this population are duplex for the *Pho1a-H_A* allele. These two duplex genotypes confirmed the results obtained with cultivar Saturna, demonstrating that in the presence of two copies of the associated allele leads to a dramatic loss of PHO1a activity.

3.2 The starch phosphorylase *Pho1b*

The *Pho1b* locus corresponds to the second isoform of plastidial starch phosphorylase. Li et al. (2008) found three SSCP polymorphisms in this locus associated with tuber quality traits (Table 15).

Table 15.*Pho1b* SSCP markers associated with chip quality traits (Li et al., 2008).

SSCP marker	Marker frequency	Chip quality at harvest	Chip quality after cold storage	Tuber starch content	Tuber starch yield
StpL-3b	0.51	↓ (5.3)	↓ (7.3)	↓ (7.0)	-
StpL-3e	0.54	↑ (6.8)	↑ (12.6)	↑ (9.3)	↓ (7.7)
StpL-3c	0.31	–	↓ (2.5)	↓ (6.4)	

Three SSCP *Pho1b* (StpL) fragments were associated with tuber quality traits. The frequency indicates the proportion of the population that contains the allele. The direction of the arrow depicts a positive or negative effect of the marker in a given trait. The number in parenthesis indicates the amount of variance (in %) explained by the marker.

3.2.1 *Pho1b* allele mining revealed high diversity in this locus.

Allele mining of the *Pho1b* locus on chromosome V was performed following the same procedures described for the *Pho1a* locus. The presence of each marker allele in the tetraploid cultivars Diana, Theresa, Satina and Saturna was confirmed in SSCP gels (Table 16). The diploid parent line P_{38} was positive for the *StpL-3b* allele.

Table 16. Distribution of the *Pho1b* SSCP markers in four tetraploid potato genotypes

Genotype	StpL-3b ↓	StpL-3e ↑	StpL-3c ↓
Theresa	0	1	0
Saturna	0	1	1
Diana	1	1	0
Satina	1	0	0

Numbers 1 or 0 stand for the presence and absence of the SSCP allele, respectively.

Full length *Pho1b* cDNA was synthesized[19] from leaf mRNA. After several attempts, *Pho1b* cDNA could not be obtained from tubers, indicating that this gene is not expressed in tuber tissue (Figure 24).

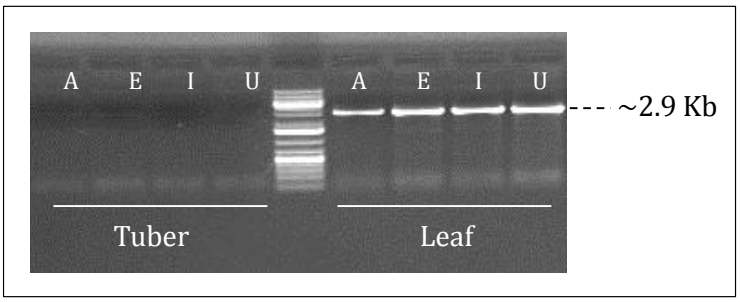


Figure 24. *Pho1b* expression is only detected in tuber but not inleaf tissue in cDNA from four potato genotypes, applying 35 PCR cycles. A, Diana; E, Theresa; I, Satina; U, Saturna. Ladder, Fermentas 1Kb Plus.

Pho1b cDNA from two independent PCR amplifications per genotype was cloned. A total of 55 Pho1b clones (17 from Diana, 14 from Theresa, 10 from Satina and 14 from Saturna) were sequenced for haplotype discovery.

The copy number of the *Pho1b* gene was determined by screening the 'BC' BAC library (Ballvora et al., 2007) with a 1800 bp PCR probe[20] generated from the diploid genotype P_{40}, parent of the F1 hybrid $P_{6/210}$ used to construct the library. BAC Clones BC70m7, BC104k19 and BC112i23 hybridized with the probe and the presence of the *Pho1b* insert was verified by PCR using the primer pair used to generate the probe. An additional PCR amplicon with

[19] StpL-3_F / X01_R
[20] StpL3_F / StpL3_R

primers designed at the 3' end of the gene confirmed full length inserts in BC70m7 and BC104k19 but no fragment was obtained for BC112i23. Since the 'BC' library is expected to have an 8-fold coverage of the potato genome, the presence of two to three positive BACs indicated that *Pho1b* was present as a single copy gene in the potato genome. This information was afterwards confirmed when the potato genome sequence was released (Assembly version 3. Scaffold PGSC003DMS000000944; Potato Genome Sequencing Consortium).

The variation found within the coding sequence of the *Pho1b* locus was very high with a total of 37 synonymous and 13 non-synonymous SNPs. The high amount of SNPs did not allow the accurate assembly of cDNA haplotypes. Because of this limitation, the haplotype calling was done only at the protein level, with those SNPs that introduced amino acid changes in the protein sequence (Table 17 and Figure 25).

Two protein haplotypes were assigned to two of the associated alleles, *StpL-3b* and *StpL-3e* (Li et al., 2008), based on the presence of the SNPs A_{155}-C_{156}, and T_{417}, respectively. These two SNPs were used to design PCR based SNP-specific primers representative of the *StpL-3b* and *StpL-3e* SSCP polymorphisms (Li, unpublished). T_{417} however, did not allow the discrimination of the *StpL-3e* allele, since two different sequences shared this particular SNP (Table 17).

Table 17. Non-synonymous SNPs of alleles of the *Pho1b* gene present in 4 tetraploid potato genotypes.

POSITION / AMINO ACID SUBSTITUTION	125 Arg/Lys	138 -(Arg Arg)	151 Ser/Cys	155 Ser/Asn	156 Ser/Asn	379 Glu/Lys	417 ∅	651 Gln/His	814 Ala/Thr	901 Thr/Ser	914 Ala/Gly	934 Tyr/Hys	1625 Glu/Ala	2084 Pro/Arg	2303 Glu/Ala	2925 3' URT
PHo1b	G	AAG	A	G	T	G	C	A	G	A	C	T	A	C	A	A
Pho1b-1	G	0	T	A	T	A	C	A	G	T	C	C	G	G	C	A
Pho1b-2	A	AAG	A	G	T	G	C	A	G	A	C	T	A	C	A	A
StpL-3b	G	AAG	A	A	C	A	C	C	G	T	C	C	A	C	C	T
StpL-3e2	A	AAG	A	G	T	G	T	A	G	A	C	C		C	C	A
StpL-3e1	A	AAG	A	G	T	G	T	A	A	T	G	C		C	A	A

Non-synonimous SNPs of the *Pho1b* cDNA sequences from four potato cultivars.The upper row shows the amino acid substitutions induced by the SNP. The number in the reference is the position of the SNP from the ATG codon in the RNA sequence. Pho1b is the reference sequence. StpL-3b, StpL-3e2, and *Stpl-3e1* are alleles which represent SSCP polymorphysms from Li et al. (2008). Two different sequences group to the StpL-3e SSCP marker. The SNPs enclosed in a box, were used to assign each allele to a SSCP polymorphism.

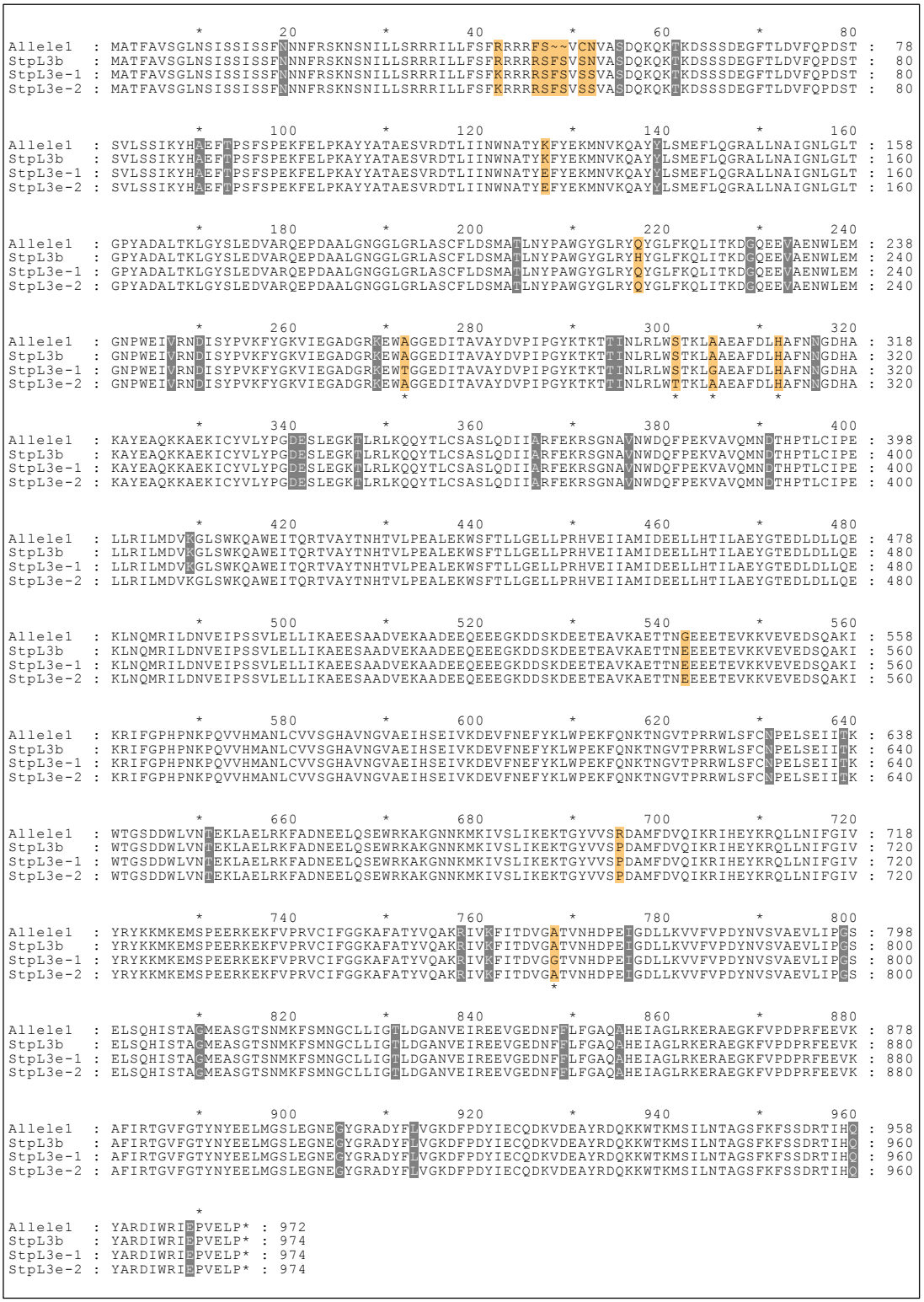


Figure 25. Overview of the location of SNPs along the *Pho1b* sequence. In grey are marked the positions where synonymous SNPs are located. Non-synonimous SNPs are shown in orange.

4 Discussion

4.1 Limits of association mapping for the *Pho1a* and *Pho1b* loci and the need for validation of candidate genes

A candidate gene-based association mapping study led to the identification of marker-alleles at two starch phosphorylase loci (*Pho1a* on chromosome III and *Pho1b* on chromosome V) associated with relevant tuber quality traits such as chip color after cold storage and starch content (Li et al., 2008). In general, association mapping based on candidate genes detects alleles in linkage disequilibrium with the trait of interest (Balding, 2006; Morton, 2005) and has thereby a great potential to identify the actual genes responsible for the investigated phenotypes. Still, a major limitation is that the locus itself may be causal for the QTL effect, or may as well just be proximal to a locus responsible for the variation, resulting in an indirect association. The likelihood of recombination between the marker-allele and the gene causal of the phenotypic effect will depend on the extent of linkage disequilibrium.

D`Hoop et al. (2010) reported that LD across population groups to decay on average at about 5 cM throughout the potato genome, which based on the estimated size of the potato genome (~840 Mbp / 800 cM (van Os et al., 2006; Bennett and Leitch, 2010), roughly corresponds to a physical distance of ~5.3 Mbp (D'Hoop et al., 2008). Large regions of LD facilitate the development of markers for precision-breeding given that the linkage between the marker and the trait is more likely to remain after meiotic recombinations occur. Yet, large LD blocks complicate the identification of the genes responsible for the phenotypic variation, limiting the understanding of molecular basis underlying a complex trait.

On chromosome III *Pho1a* alleles were found in LD with alleles from the *Pain 1* locus (vacuolar acid invertase), located 6 cM apart in the proximal direction, constituting a large haplotype block in this particular region of the potato genome (Li et al., 2008). Even though a draft of the potato genome sequence has been generated (Potato Genome Sequencing Consortium) the arrangement of contigs on the upper arm of chromosome III did not include the *Pho1a* locus. Hence, based on the potato sequence, it is still not possible to identify which loci involved in starch metabolism may be within LD distance.

The analyses during this study were required to identify the natural allelic variants of the two *Pho1* genes and to determine, whether they have a direct effect on the variation in tuber

quality traits or if the association is the result of the genetic linkage to a neighboring functional allele within a large LD block. Validation of the candidate genes allows the generation of molecular markers in the causal gene and also provides a valuable contribution to the understanding of the nature of these complex traits.

4.2 Identification of natural alleles of *Pho1a* associated with tuber quality traits

In the previous association study (Li et al., 2008), three different SSCP (single strand conformational polymorphisms) allelic fragments were identified for the *Pho1a* gene (*Stp23* in Li et al., 2008). The SSCP fragment *Stp23-8b*, showed a positive association with quality traits explaining ~10% of the phenotypic variance in chip quality after cold storage and tuber starch content (Table 5). The second SSCP fragment, *Stp23-8a*, also showed significant positive associations although with a smaller effect (~5%). The remaining allele (*Stp23-7c*) was not found in association with the evaluated traits.

The first step towards the identification of the *Pho1a* haplotypes underlying the associated SSCP polymorphisms was the characterization of full length allele sequences from genotypes bearing the associated SSCP markers. Full length cDNAs from four different cultivars were cloned and sequenced for this purpose. The nucleotide sequences of 161 clones were compared to a reference *Pho1a* mRNA sequence (GenBank: D00520.1). Of the 700 SNPs scored in the 161 sequences, singletons (SNPs that appeared only in one clone) and SNPs occurring only twice (which could not be confirmed in different amplifications), made up for 38% and 8% of the total amount of variants scored, respectively. The described method resulted in a high rate of errors, as only 2.1% of the total SNPs scored were reproducible in clones from independent experimental replications. Comparable error rates were also obtained for cDNA clones from potato invertase alleles (Draffehn, 2010). Nucleotide miss-calls in the sequencing were likely to be responsible for the high rate of singletons. To compensate for miss-calls, sequencing of both ends of a PCR product is a common procedure. However, full coverage of the *Pho1a* clones required four different sequencing reactions and double sequencing of the clones was considered too laborious. To a lesser extent, errors in the cDNA synthesis or those introduced by the PCR at late cycles of amplification (Sun, 1995) could have appeared in clones at very low frequencies. Therefore, a base call was considered factual only when found in at least three different clones from two independent amplification reactions.

Sorting of remaining revised sequences by genotype revealed that more than four different sequences had been retrieved from each of the cultivars Dina, Theresa, and Saturna (Table 7). The output was certainly unexpected for what was proposed to be a single copy gene (Camirand et al., 1990), and it was an indication of either additional copies of this gene or technical artifacts. More than one region of high similarity to the *Pho1a* sequence was not found in the most recent release of the potato genome (V.3., Potato Genome Sequencing Consortium) supporting that *Pho1a* is a single copy gene and that no other elements in the genome were likely to be co-amplified by the primers chosen to produce full length *Pho1a* cDNA.

The possibility remained that the high amount of allele sequences was a product of spurious combinations of non-linked SNPs during PCR. Because the success of the structural and functional characterization of the *Pho1a* alleles depends on the accuracy of their sequences, phasing of the SNPs (i.e. determining which SNPs are inherited together constituting a haplotype) was carried out.

In the cDNA clones, SNPs seemed to be correctly phased in small regions (e.g. the first 700 bp), since within these regions SNPs consistently appeared together in the same clones. Beyond that range, allele sequence combinations increased the number of putative alleles. This suggested PCR-introduced chimerical sequences derived from the combination of partial PCR or cDNA synthesis products priming the PCR reaction. Artificial chimeras are of common occurrence between similar templates in the PCR mixture and increase with the similarity between the sequences (Meyerhans et al., 1990). Up to 30% of chimeric sequences were identified after 30 cycles of amplification in rRNA genes (Wang and Wang, 1996). Considering the presence of maximum four different copies per genotype and given that *Pho1a* alleles are 95.5% identical, the DNA template is highly likely to result in artificial chimeras.

As SNPs in LD will be certain to have the same distribution in the population, scoring a subpopulation of 34 individuals was a good approach to resolve the problem of artificial allele sequences and also to correlate SSCP markers with their respective cDNA alleles (Table 8). Comparison of the distribution of SNPs diagnostic for each haplotype group (green and blue, Table 7), showed that (1) the distribution of SNPs from the green group and the SNP A_{2578} from the blue group was different, confirming that they belong to different haplotypes and (2) that SNPs G_{824} and A_{2776} are in complete LD (green group, Table 8).

Two SCCP markers (*Stp23-7c* and *Stp23-8b*) could be assigned to cDNA allele sequences based on their distribution in the population, but an allele for the SSCP marker *Stp23-8a* was not identified. It is possible that *Stp23-8a* is encoded in one of the groups for which the SNP combinations were not resolved. Scoring of the A_{2754} and A_{2616} (or A_{2765}) SNP distribution may help to clarify this issue (Table 7).

From the problems encountered along the process of finding phased SNPs in long cDNA sequences, it can be concluded that alone, the cloning of PCR products from heterozygous templates is not sufficient (in spite of the use of high fidelity reverse transcriptase and DNA polymerases) to accurately phase SNPs, as it would be assumed if each clone contained a single "clean" DNA molecule. This method needs to be complemented with the use of additional techniques to produce accurate results.

4.3 The *Pho1a-H$_A$* allele is associated with higher starch content and better chip quality after cold storage.

Here we describe the *Pho1a* allele corresponding to the SSCP polymorphism *Stp23-8b*, which had a positive effect on tuber quality traits explaining 9%, 11.4% and 6.7% of the phenotypic variance in chip color after cold storage, tuber starch content, and tuber starch yield, respectively, in the GABI-CHIPs population (Li et al., 2008). The associated haplotype was designated as "*Pho1a-H$_A$*" (Appendix 2). The SNP specific primer pairs used to score this allele (8Dc208_F / Dsnp824G_R and Dc2478_F / Di14-50_R, Table 3) are PCR based and proved highly reliable under a touch-down PCR protocol. They can be used for marker assisted selection in breeding programs.

4.4 *Pho1a-H$_A$* influences additional tuber quality traits

The association of the *Pho1a-H$_A$* allele with quality traits was also observed in two additional populations evaluating different traits correlated with starch metabolism: The TASK population, in which 205 tetraploid genotypes were used to study the genetic basis underlying tuber bruising susceptibility (Urbany et al., 2011), and the BIOSOL population, consisting of 40 individuals, used to determine the genetic basis underlying the accumulation of sugars upon cold storage (unpublished). In these populations, *Pho1a-H$_A$* is positively associated with specific gravity (starch content) and increases the susceptibility to tuber

bruising. Moreover, potato genotypes bearing *Pho1a-H_A* accumulate less reducing sugars after cold storage (Sections 3.1.2.1).

Tuber bruising is the enzymatic discoloration on localized areas of the potato tuber caused by mechanical damage during harvesting. Bruised tubers represent considerable economic losses as they are rejected in the market for their appearance, flavor and texture. Tuber bruising occurs probably as a consequence of the disruption of the amyloplastic membranes caused by mechanical damage. As a result, polyphenol oxidases are released and react with cytoplasmic and vacuolar phenolic substrates causing a dark coloration (Urbany et al., 2011). Given the strong positive correlation between tuber starch content and bruising susceptibility, it has been proposed that amyloplast with high loads of starch are more likely to suffer disrupted membranes when subjected to mechanical impact (Urbany et al., 2011). Our data show that the *Pho1a-H_A* allele is the DNA sequence causal of the SSCP marker *PHO1A-c,* reported in the above mentioned study. The effect of this allele in the TASK population was an increase in tuber starch content but also an increase in bruising susceptibility (Table 10). A general decrease of the *Pho1a* activity may explain the higher starch content in tubers bearing the *Pho1a-H_A* allele, either resulting from a reduction of transcription or enzymatic activity, thereby leading to increased tuber bruising susceptibility.

Accumulation of sugars upon cold storage is directly linked to starch degradation (Isherwood, 1973). The association of the *Pho1a-H_A* marker with sugar content before cold storage in the BIOSOL population was positive (Section 3.1.2.2), as most of the genotypes containing the *Pho1a-H_A* marker contained on average less reducing sugars than genotypes where the allele was not present (Figure 6). The presence of the H_A haplotype in two bad cold storage cultivars (high sugar accumulation, Figure 6), may be explained by unfavorable alleles at different loci affecting the same trait. Cold sweetening is a complex trait regulated by multiple genes, and quality depends upon a combination of favorable alleles.

The association of *Pho1a-H_A* with reduced cold sweetening appeared more pronounced as the time of cold storage increased (Figure 7B). This suggests an important role of this allele and that genotypes bearing the H_A allele might be less efficient in breaking down starch.

In summary these findings demonstrate that besides traits such as chip color, tuber starch content, and tuber starch yield, variation within the *Pho1* locus contributes to tuber bruising susceptibility and accumulation of reducing sugars at cold temperatures. All these traits are correlated to tuber starch content. The association with *Pho1a-H_A* remained strong in every

population. This adds yet more evidence to the role of *Pho1a* in starch degradation in different contexts. It has been argued that *Pho1* may have important roles in starch synthesis but not in degradation in leaves (Zeeman et al., 2004). The consistent correlation of this locus with traits that depend on a lower rate of starch degradation is a good indication that this gene is in fact an important factor in starch catabolism and that the *Pho1a-*H_A allele represents a reduced PHO1a activity.

4.5 Genotype-dependent *Pho1a* expression and regulation upon cold storage

To test whether the decreased degradation capability of PHO1a was a consequence of a lower amount of *Pho1a* transcripts, we analyzed the structure of its promoter and its expression levels.

The analysis of the promoter sequences of four genotypes revealed a large 1 Kb deletion (DP01) and two 8 bp Indels. The DP01 deletion encompasses a region previously reported to contain two negative and one positive regulators of gene expression as determined by 5' deletion construct analysis (St. Pierre and Brisson, 1995). This opened the possibility that some of the alleles were differentially regulated because of the characteristics of their promoter regions, causing an effect on the phenotypic variation. However, experiments revealed that the DP01 polymorphism in the promoter had the same distribution as the SSCP *Stp23-7c* polymorphism, which was not associated with quality traits. Further analysis of these variants could provide useful information in the regulation and organ specific localization of the *Pho1a* gene.

qRT-PCR experiments were performed to determine how the overall expression of *Pho1a* may be related to sugar content at different time points. Down-regulation of *Pho1a* expression upon cold storage (Figure 11) supports the findings of different studies that report a rapid drop off of *Pho1* transcripts levels when placed in cold temperatures that slowly start rising as the time of cold storage increases. (Bagnaresi et al., 2008). Therefore, a possible physiological role for *Pho1a* at later stages after cold storage has been suggested to involve the degradation of starch to provide a source of sugars readily available for sprouting. The results from this experiment support that hypothesis as the transcript levels were increasing at four weeks after cold storage.

Expression levels before cold storage are significantly different between low and high sugar genotypes, where the bad genotypes have higher *Pho1a* expression (Figure 12B). The differences in expression are much lower than the four-fold difference in sugar means between good and bad genotypes before cold storage (Figure 12A), indicating that the expression differences only have a minor influence on sugar content . It must be considered, however, that tuber sugar content is a complex trait and many different genes contribute to the phenotypic variation.

During cold storage expression levels were similar in good and bad genotypes, indicating that sugar accumulation during cold storage is not influenced by the overall *Pho1a* expression levels.

4.6 Allele specific transcriptional analysis shows that a *Pho1a-H^A* is not differentially regulated

The analysis of *Pho1* transcripts did not discriminate between alleles. Provided that significant expression differences were found between genotypes containing the *Pho1a-H_A* allele (Section 3.1.4.1), it was necessary to investigate whether this allele was being differentially regulated in comparison to the total *Pho1a* expression. The polymorphisms found in the promoter regions supported the possibility of allelic differential transcriptional regulation. Pyrosequencing experiments carried out to study relative regulation of the *Pho1a-H_A* allele upon cold storage showed that the proportions of this allele were maintained at different time points and were as well proportional to the allele dosages in genomic DNA (Figure 14). From these expression experiments, it was concluded that the effect of the *Pho1a-H_A* allele does not rely on transcriptional regulation.

4.7 The *Pho1a-H_A* allele affects PHO1a enzyme activity

In-gel activity assays with 40 genotypes were carried out to determine, whether the *Pho1a* allelic variants had any effect on the overall activity of the enzyme complexes on starch degradation and if there were differences in the activity before and after cold storage.

Tuber starch phosphorylases were tested for their ability to synthesize starch as an indicator of the activity of the enzyme complexes. In general, low PHO1a activity was observed in genotypes that contained the *Pho1a-H_A* allele. Moreover, a striking result was to observe that the three H_A-duplex genotypes evaluated (Saturna, Figure 22; and cv 17 and cv 20, Figure 23) displayed no detectable activity. This suggested that *Pho1a-H_A* is a dominant allele, which affects the overall activity of the enzyme complex.

Glucan phosphorylases are highly conserved proteins among living organisms (Appendix 3, Gerbrandy and Doorgeest, 1972; Nakano and Fukui, 1986; Johnson, 1989; Watson et al., 1997) The areas involved in the stabilization of dimers are among the most conserved along with the areas involved in the catalytic site around the PLP domain (Tanizawa et al., 1994; Palm et al., 1985; Fletterick and Madsen, 1980; Krebs and Fischer, 1956; Johnson, 1989). It is well documented that the activity of starch phosphorylases depends on the formation of dimers (Nakano and Fukui, 1986; Johnson, 1989; Watson et al., 1997). In this context, it seems reasonable that dimerization of the PHO1a-H_A isoform affected, forming non-functional homo-dimers and hetero-dimers with different isoforms and that this detrimental effect increases upon the allele dosage.

4.8 Amino acid substitutions in the *Pho1a-H_A* allele may affect dimerization

At the sequence level, various differences were identified between *Pho1a* alleles. As no changes were found on the transcriptional level, sequence changes within the coding region were investigated in detail. The coding region of the *Pho1a-H_A* allele contains four conservative and two non-conservative amino acid substitutions with respect to the reference allele *Pho1a-H_R* (Table 7).

The Ala-108-Thr substitution has no precedents in α-glucan phosphorylases as revealed by alignment and comparison of 200 different PHO sequences (Figure 16). This alanine residue is 100% conserved among animal and plant PHOs, suggesting that it has an important function in the PHO enzyme. Moreover, an *in silico* mutation of this residue was predicted (with 84% confidence) to affect the enzyme stability in the rabbit muscle glycogen phosphorylase (Figure 17).

α-glucan phosphorylases depend on the formation of dimers for activity. Subunit to subunit interactions are mostly occurring at the first 280 residues in the N-terminal domain (Nakano and Fukui, 1986; Johnson, 1992). The motifs contributing to the dimerization interface are the "cap", α2-helix, and β7-strand (Johnson, 1989). These regions are highly conserved blocks in α-glucan phosphorylases. *Pho1a-H_A* amino acid exchange Ala-108-Thr is localized in the α2-helix and Lys-275-Arg is in the interface of the β7 strand; hence, it is possible that the combination of these two amino acids in one allele affects dimerization and compromise the enzymatic activity. Like plant starch phosphorylases, bacterial maltodextrin phosphorylases (MALP) lack allosteric regulation and the sites of contact between subunits are very similar. The location of the Ser56 and Thr222 residues in MALP, correspondent to the Ala108 and Lys275 residues in PHO1a, show the critical position of these amino acid for the dimer complex (Figure 26).

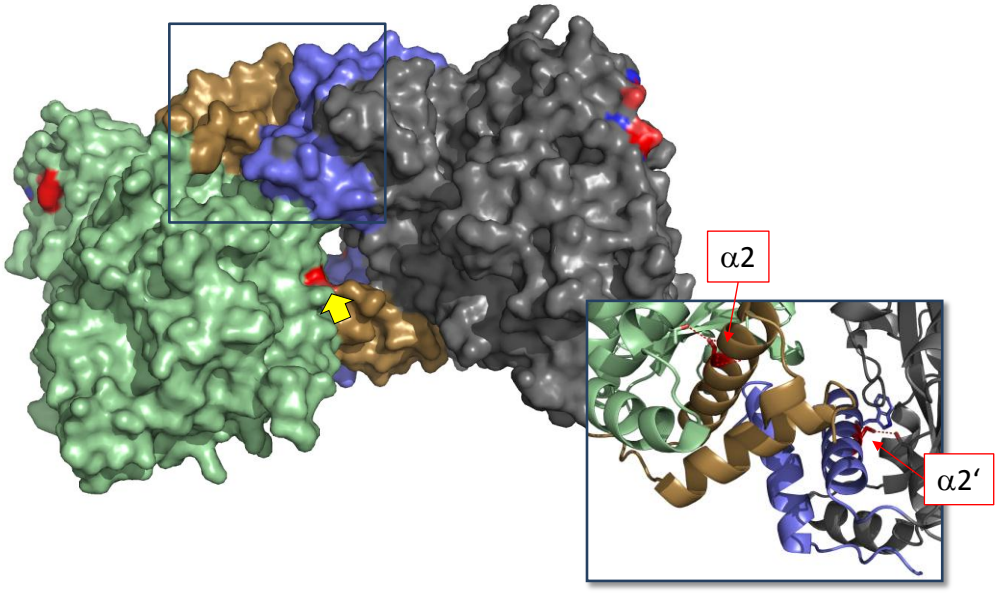


Figure 26. Structure of the bacterial maltodextrin phosphorylase (MALP) dimer complex (PDB: 2ECP, Watson et al., 1997). Colors brown and blue represent the regions for dimer contact (for the green and gray subunits, respectively). A close-up of the helices forming the Cap-α2-β7 interface is shown. The position in the α2 helix of the Ser56 residue in MALP (Ala108 in PHO1a) is indicated with red arrows. The yellow arrow points out the location of the Thr222 residue (Lys275 in PHO1a) in the proximity of the tower-tower' interface. The remaing red residues correspond to the position of the mutations described in section 3.1.5

The effect of the mutations introduced by the *Pho1a-H_A* allele appears to be strong enough to dramatically reduce the overall activity in the genotypes where the allele is found in duplex dosage. This indicates that the PHO1a-H_A isoform not only is inactive, but it seems reasonable to think that it also renders the H_A/H_R dimers dysfunctional, which would explain the reduced activity observed and supports the hypothesis of a dominant allele with deleterious effect. Production of recombinant proteins of the *Pho1a-H_R* and *Pho1a-H_A* alleles is in progress to determine the biochemical properties of the isoforms and the *Pho1a* allelic heterodimers.

4.9 A putative role for PHO1a in cold sweetening

Pho1 has been proposed as a candidate to be involved in tuber quality traits, particularly sugar accumulation upon cold storage. Structural and functional analysis added evidence for a direct role of the Pho1a locus on chromosome II.

The importance of starch phosphorylase in starch synthesis has been shown in rice endosperm (Satoh et al., 2008) as well as in *Chlamidomonas reinhardtii* (Dauvillé et al., 2006). In potato, the plastidic starch phosphorylase is one of the predominant proteins present in the amyloplasts of mini-tubers (Stensballe et al., 2008), which supports its role in starch synthesis in potato. On the other hand, lack of starch phosphorylase did not compromise the normal starch degradation rates in *A. thaliana AtPHS1* (*Pho1a* homologue) mutants (Zeeman et al., 2004). Nevertheless, the mutants proved to be more susceptible to abiotic stress. Dauvillé et al. (2006) pointed out that as long as the supply of ATP is unlimited in the leaf, the phosphorolysis is not favored over hydrolysis for starch degradation. Despite the fact that starch phosphorylase has been long proposed to have an active role in cold sweetening (Sowokinos, 2001), the expression profile of *Pho1a* was not in line with the pattern of sugar accumulation upon cold storage suggesting instead a possible role at later stages (Bagnaresi et al., 2008). The results of this work, however, added evidence for a role of starch phosphorylase in cold sweetening at the protein level by showing that genotypes that exhibit lower activity of starch phosphorylases accumulate less reducing sugars upon cold storage.

Gounaris (2001) proposed a possible scenario, in which the accumulation of sugars in cold stored tissues occurs as a response to a decline of plastid ATP levels caused by the inhibition of plastid and mitochondrial ATPases at low tempetaures. In this context, a shift of starch phosphorylase activity from starch synthesis to degradation would be stimulated by the higher availability of Pi to avoid ATP depletion upon long periods of cold storage. In fact,

Geigenberger et al. (2001) showed that plastids defective of plastid ATP membrane transporters in potato tubers showed lower levels of starch and higher levels of free sugars and hexose phosphates.

4.10 Diversity of the *Pho1b* locus in chromosome V

Screening for haplotypes of the *Pho1b* locus was carried out following the same approach as used for *Pho1a* alleles. Albrecht et al. (2001) reported extremely low expression levels of *Pho1b* in potato tubers that were detectable only after 40 cycles of PCR amplification and that the protein remained undetectable also in crude tuber extracts. At the cellular level, however, they were able to detect the *Pho1b* protein in few cells in the vicinity of the vascular tissue. In our experiments we could not obtain *Pho1b* transcripts from tuber tissue after 35 cycles of amplification. Moreover, no *Pho1b* sequences were identified in Roche 454-sequencing data of four different sets of tuber transcripts (unpublished data), whereas transcripts matching *Pho1a* were abundant. Our data suggest that *Pho1b* is not expressed in tubers. *Pho1b* transcripts for allele discovery were therefore obtained from leaf tissue.

The variation within the *Pho1b* locus (in average, 1 SNP every 60 bases of the coding sequence) was 3.2 times higher than that observed at the *Pho1a* locus (1 SNP every 193 bases). An accurate phasing of the 50 SNPs to predict haplotypes was therefore not possible. *Pho1b* clone sequences were divided into six groups of similarity that represent the allelic diversity of this locus. The cDNA diversity met the expectations derived from the association study that provided the basis of this work (Li et al., 2008), predicting five different allelic variants based on the number of SSCP fragments scored that had segregated independently in the population analyzed.

The SSCP polymorphisms *StpL-3b* and *StpL-3e* (having a negative and positive association, respectively) were identified within the groups, and four isoforms remained unassigned. The *StpL-3e* SSCP fragment, however, is represented by two different alleles (Table 17). This occurred because the amplicon sequence from which the SSCP markers were obtained is the same for both alleles and they diverge in a different region of the gene. This finding may change the associations of this marker as two haplotypes should be considered instead of one. Testing the original population with SNP specific primers for the A_{814}, T_{901}, or G_{914} SNPs will clarify this issue and predict accurately the association of the alleles. This clarification is important for further applicability of the *StpL-3e* marker for breeding purposes, since the

linkage of the marker with the phenotype may be lost in populations where the two alleles segregate separately.

Not all the SNPs from the cDNA sequences could be phased, but haplotyping of those SNPs that introduced amino acid exchanges allowed the assembly of protein variants with decent accuracy. This analysis provides information on SNPs, from which PCR-base markers can be designed for breeding purposes.

Urbany et al. (2011) also found associations of *Pho1b* alleles with tuber bruising susceptibility and specific gravity (starch content). The question arises how the *Pho1b* gene influences tuber quality traits if it is not expressed in the tubers. The accumulation of starch in tubers depends on a successful flux of sugars from source to sink tissues. If the starch degradation in the leaf is either increased or impaired as the result of the activity of an aberrant protein, a different amount of sugars may be mobilized to the tubers. This could result in a different amount of substrate (sugars) from which starch is synthesized, thereby explaining the effect of *Pho1b* on starch quality traits in potato tubers.

Contribution of co-workers to this project

Matthias Fischer and Lena Schreiber provided the plant material of the BIOSOL population for the functional analysis of starch phosphorylase, central for the development of this work.

Sugars measurements were done by Markus Kuckenberg.

Mass spectrometry was done with the support of Tom Colby from the Mass Spectrometry group of the Max Planck Institute for Plant Breeding Research.

Li Li and Claude Urbany provided the DNA from the GABI-CHIPS and the TASK population.

Renier van der Hoorn for his help with the protein modeling

5 Literature

Albrecht, T., Greve, B., Pusch, K., Kossmann, J., Buchner, P., Wobus, U., and Steup, M. (1998). Homodimers and heterodimers of Pho1-type phosphorylase isoforms in S*olanum tuberosum* L. as revealed by sequence-specific antibodies. Eur. J. Biochem. **251**, 343-352.

Albrecht, T., Koch, A., Lode, A., Greve, B., Schneider-Mergener, J., and Steup, M. (2001). Plastidial (Pho1-type) phosphorylase isoforms in potato (*Solanum taberosum* L.) plants: expression analysis and immunochemical characterization. Planta **213**, 602-613.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. **2**, 2006.0008 16738554.

Bagnaresi, P., Moschella, A., Beretta, O., Vitulli, F., Ranalli, P., and Perata, P. (2008). Heterologous microarray experiments allow the identification of the early events associated with potato tuber cold sweetening. Bmc Genomics **9**.

Balding, D.J. (2006). A tutorial on statistical methods for population association studies. Nature Genetics. **7**, 781-791

Ballvora, A., Jocker, A., Viehover, P., Ishihara, H., Paal, J., Meksem, K., Bruggmann, R., Schoof, H., Weisshaar, B., and Gebhardt, C. (2007). Comparative sequence analysis of Solanum and Arabidopsis in a hot spot for pathogen resistance on potato chromosome V reveals a patchwork of conserved and rapidly evolving genome segments. BMC Genomics **8,** 112.

Bennett, M.D., and Leitch, I.J. (2010). Plant DNA C-values Database (Release 5.0, Dec. 2010). http://www.kew.org/cvalues/.

Bernard, P., and Couturier, M. (1992). Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. J. Mol. Biol. **226**, 735-745.

Bernard, P., Kézdy, K.E., Van Melderen, L., Steyaert, J., Wyns, L., Pato, M.L., Higgins, P.N., and Couturier, M. (1993). The F Plasmid CcdB Protein Induces Efficient ATP-dependent DNA Cleavage by Gyrase. J. Mol. Biol. **234**, 534-541.

Bhardwaj, V., Manivel, P., and Gopal, J. (2011). Screening potato (*Solanum* spp) for cold-induced sweetening. Indian J. Agr. Sci. **81**, 20-24.

Bhaskar, P.B., Wu, L., Busse, J.S., Whitty, B.R., Hamernik, A.J., Jansky, S.H., Buell, C.R., Bethke, P.C., and Jiang, J.M. (2010). Suppression of the Vacuolar Invertase Gene Prevents Cold-Induced Sweetening in Potato. Plant Physiol. **154**, 939-948.

Biegert, A., Mayer, C., Remmert, M., SÃ¶ding, J., and Lupas, A.N. (2006). The MPI Bioinformatics Toolkit for protein sequence analysis. Nucleic Acids Res. **34**, W335-W339.

Bonierbale, M., Plaisted, R., and Tanksley, S. (1992). Genetic-Mapping and Utilization of Quantitative Trichome Mediated Insect Resistance in Potato. Neth. J. Plant Pathol. **98**, 211-214.

Bonierbale, M.W., Plaisted, R.L., and Tanksley, S.D. (1988). RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics **120**, 1095-1103.

Borovkov, A.Y., McClean, P.E., Sowokinos, J.R., Ruud, S.H., and Secor, G.A. (1996). Effect of expression of UDP-glucose pyrophosphorylase ribozyme and antisense RNAs on the enzyme activity and carbohydrate composition of field-grown transgenic potato plants. J.Plant Physiol. **147**, 644-652.

Bradshaw, J.E. (1994). Quantitative genetics theory for tetrasomic inheritance. Potato genetics**,** 71-99.

Buchner, P., Borisjuk, L., and Wobus, U. (1996). Glucan phosphorylases in *Vicia faba* L. cloning, structural analysis and expression patterns of cytosolic and plastidial forms in relation to starch. Planta **199**, 64-73.

Burton, W.G. (1969). The Sugar Balance in Some British Potato-D Varieties During Storage II the Effects Tuber Age Previous Storage Temperature and Intermittent Refrigeration Upon Low Temperature Sweetening. European Potato Journal **12**, 81-95.

Camirand, A., St. Pierre, B., Marineau, C., and Brisson, N. (1990). Occurrence of a Copia-Like Transposable Element in One of the Introns of the Potato Starch Phosphorylase Gene. Mol. Gen. Genet. **224**, 33-39.

Cao, D., Craig, B.A., and Doerge, R.W. (2005). A Model Selection-Based Interval-Mapping Method for Autopolyploids. Genetics **169**, 2371-2382.

Cervantes-Flores, J., Yencho, G., Kriegner, A., Pecota, K., Faulk, M., Mwanga, R., and Sosinski, B. (2008). Development of a genetic linkage map and identification of homologous linkage groups in sweetpotato using multiple-dose AFLP markers. Mol. Breeding **21**, 511-532.

Chen, S., Hajirezaei, M.R., Zanor, M.I., Hornyik, C., Debast, S., Lacomme, C., Fernie, A.R., Sonnewald, U., and Bornke, F. (2008). RNA interference-mediated repression of sucrose-phosphatase in transgenic potato tubers (*Solanum tuberosum*) strongly affects the hexose-to-sucrose ratio upon cold storage with only minor effects on total soluble carbohydrate accumulation. Plant Cell Environ. **31**, 165-176.

Chen, X., Salamini, F., and Gebhardt, C. (2001). A potato molecular-function map for carbohydrate metabolism and transport. Theor. Appl. Genet. **102**, 284-295.

Cori, C.F., and Cori, G.T. (1936). Mechanism of formation of hexosemonophosphate in muscle and isolation of a new phosphate ester. Proc. Soc. Exp. Biol. Med. **34**, 702-705.

Crooks, G.G., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: A Sequence Logo Generator. Genome Res. **14**, 1188-1190.

Cunningham, C., and Stevenson, F. (1963). Inheritance of factors affecting potato chip color and their association with specific gravity. Am. J. Potato Res. **40**, 253-265.

D'hoop, B. João Paulo, M. Kowitwanich, K., Sengers, M., Visser, R. van Eck, H.J. and van Eeuwijk, F.A. (2010) Population structure and linkage disequilibrium unravelled in tetraploid potato. Theor Appl Genet. **121**(6): 1151–1170.

Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci USA **97**, 6640-6645.

Dauvillée, D., Chochois, V., Steup, M., Haebel, S., Eckermann, N., Ritte, G., Ral, J.P., Colleoni, C., Hicks, G., Wattebled, F., Deschamps, P., d'Hulst, C., Lienard, L., Cournac, L., Putaux, J.L., Dupeyre, D., and Ball, S.G. (2006). Plastidial phosphorylase is required for normal starch synthesis in *Chlamydomonas reinhardtii*. Plant J. **48**, 274-285.

Debarbouille, M., Cossart, P., and Raibaud, O. (1982). A DNA-Sequence Containing the Control Sites for Gene Malt and for the Malpq Operon. Mol. Gen. Genet. **185**, 88-92.

Decker, K., Plumbridge, J., and Boos, W. (1998). Negative transcriptional regulation of a positive regulator: the expression of malT, encoding the transcriptional activator of the maltose regulon of *Escherichia coli*, is negatively controlled by Mlc. Mol. Microbiol. **27**, 381-390.

Deschamps, P., Haferkamp, I., d'Hulst, C., Neuhaus, H.E., and Ball, S.G. (2008a). The relocation of starch metabolism to chloroplasts: when, why and how. Trends Plant Sci. **13**, 574-582.

Deschamps, P., Moreau, H., Worden, A.Z., Dauvillée, D., and Ball, S.G. (2008b). Early gene duplication within chloroplastida and its correspondence with relocation of starch metabolism to chloroplasts. Genetics **178**, 2373-2387.

D'Hoop, B.B., Paulo, M.J., Mank, R.A., van Eck, H.J., and van Eeuwijk, F.A. (2008). Association mapping of quality traits in potato (*Solanum tuberosum* L.). Euphytica **161**, 47 - 60.

Doyle, J.J., and Doyle, J.L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin **19**, 11-15.

Draffehn, A.M. (2010) Structural and functional characterization of natural alleles of potato (*Solanum tuberosum* L.) invertases associated with tuber quality traits. Inaugural – Dissertation zur Erlangung des Doktorgrades der Mathematisch-Naturwissenschaftlichen Fakultät der Universität zu Köln.

Draffehn, A.M., Meller, S., Li, L., and Gebhardt, C. (2010). Natural diversity of potato (*Solanum tuberosum*) invertases. Bmc Plant Biology **10**.

Dunwell, J.M. and Sunderlan, N. (1973). Anther culture of *Solanum tuberosum* L. Euphytica **22**: 317-323.

Duwenig, E., Steup, M., Willmitzer, L., and Kossmann, J. (1997). Antisense inhibition of cytosolic phosphorylase in potato plants (*Solanum tuberosum* L.) affects tuber sprouting and flower formation with only little impact on carbohydrate metabolism. Plant J. **12**, 323-333.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**, 1792-1797.

Fan, H.C., Wang, J.B., Potanina, A., and Quake, S.R. (2011). Whole-genome molecular haplotyping of single cells. Nat. Biotechnol. **29**, 51-57.

FAO Crops statistics database, http://faostat.fao.org/

Fletterick, R.J., and Madsen, N.B. (1980). The Structures and Related Functions of Phosphorylase-A. Annual Review of Biochemistry **49,** 31-61.

Fukui, T., Shimomura, S., and Nakano, K. (1982). Potato and Rabbit Muscle Phosphorylases - Comparative Studies on the Structure, Function and Regulation of Regulatory and Non-Regulatory Enzymes. Mol. Cell. Biochem. **42**, 129-144.

Gavrilenko, T. (2007). Potato cytogenetics. In Potato biology and biotechnology. Advances and perspectives., D. Vreugdenhil, J.E. Bradshaw, C. Gebhardt, F. Govers, D.K.L. MacKerron, M.A. Taylor, and H. Ross, eds (Oxford, Amsterdam: Elsevier), pp. 203-216.

Gebhardt, C. (2007). Molecular markers, maps and population genetics. In Potato biology and biotechnology. Advances and perspectives., D. Vreugdenhil, J.E. Bradshaw, C. Gebhardt, F.

Govers, D.K.L. MacKerron, M.A. Taylor, and H. Ross, eds (Oxford, Amsterdam: Elsevier), pp. 77-89.

Gebhardt, C., and Valkonen, J.P.T. (2001). Organization of genes controlling disease resistance in the potato genome. Annu. Rev. Phytopathol. **39**, 79 - 102.

Gebhardt, C., Ballvora, A., Walkemeier, B., Oberhagemann, P., and Schuler, K. (2004). Assessing genetic potential in germplasm collections of crop plants by marker-trait association: a case study for potatoes with quantitative variation of resistance to late blight and maturity type. Mol. Breed. **13**, 93 - 102.

Gebhardt, C., Ritter, E., Barone, A., Debener, T., Walkemeier, B., Schachtschabel, U., Kaufmann, H., Thompson, R.D., Bonierbale, M.W., Ganal, M.W., Tanksley, S.D., and Salamini, F. (1991). RFLP Maps of Potato and Their Alignment with the Homoeologous Tomato Genome. Theor. Appl. Genet. **83**, 49-57.

Gebhardt, C., Ritter, E., Debener, T., Schachtschabel, U., Walkemeier, B., Uhrig, U., and Salamini, F. (1989). RFLP analysis and linkage mapping in *Solanum tuberosum*. Theor. Appl. Genet. **78**, 65 - 75.

Gebhardt, C., Walkemeier, B., Henselewski, H., Barakat, A., Delseny, M., and Stuber, K. (2003). Comparative mapping between potato (*Solanum tuberosum*) and *A. thaliana* reveals structurally conserved domains and ancient duplications in the potato genome. Plant J. **34**, 529 - 541.

Geer, L.Y., Marchler-Bauer, A., Geer, R.C., Han, L., He, J., He, S., Liu, C., Shi, W., and Bryant, S.H. (2010). The NCBI BioSystems database. Nucleic Acids Res. **38,** D492-D496.

Geigenberger, P., Stamme, C., Tjaden, J., Schulz, A., Quick, P.W., Betsche, T., Kersting, H.J., and Neuhaus, H.E. (2001). Tuber physiology and properties of starch from tubers of transgenic potato plants with altered plastidic adenylate transporter activity. Plant Physiology **125,** 1667-1678.

Gerbrandy, S.J. (1974). Glycogen-Phosphorylase of Potatoes - Purification and Thermodynamic Properties of Adsorption on Glycogen. Biochim. Biophys. Acta **370**, 410-418.

Gounaris, Y. (2001). A qualitative model for the mechanism of sugar accumulation in cold-stressed plant tissues. Theor. Biosci. **120**, 149-165.

Greiner, S., Rausch, T., Sonnewald, U., and Herbers, K. (1999). Ectopic expression of a tobacco invertase inhibitor homolog prevents cold-induced sweetening of potato tubers. Nat. Biotechnol. **17**, 708-711.

Gupta, S.K., Sowokinos, J.R., and Hahn, I.S. (2008). Regulation of UDP-glucose pyrophosphorylase isozyme UGP5 associated with cold-sweetening resistance in potatoes. J. Plant Physiol. **165,** 679-690.

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. **41**, 95-98.

Hamernik, A.J., Hanneman, R.E., and Jansky, S.H. (2009). Introgression of Wild Species Germplasm with Extreme Resistance to Cold Sweetening into the Cultivated Potato. Crop Sci. **49**, 529-542.

Hanahan, D. (1983). Studies on Transformation of *Escherichia coli* with Plasmids. J. Mol. Biol. **166**.

Helmreich, E. and Cori, C.F. (1964). Role of Adenylic Acid in Activation Phosphorylase. Proc. Natl. Acad. Sci. USA. **51**, 131-138.

Hijmans, R.J., and Spooner, D.M. (2001). Geographic distribution of wild potato species. Am. J. Bot. **88**, 2101-2112.

Isherwood, F.A. (1973). Starch-sugar interconversion in *Solanum tuberosum* L. Phytochemistry **12**: 2579

Jacobs, J.M.E., van Eck, H.J., Arens, P., Verkerkbakker, B., Hekkert, B.T.L., Bastiaanssen, H.J.M., Elkharbotly, A., Pereira, A., Jacobsen, E., and Stiekema, W.J. (1995). A Genetic-Map of Potato (*Solanum tuberosum*) Integrating Molecular Markers, Including Transposons, and Classical Markers. Theor. Appl. Genet. **91**, 289-300.

Jakuczun, H. and Zimnoch-Guzowska, E. (2004). Inheritance of glucose content in tubers of diploid potato families. Am. J. Potato Res. **81,** 359-370.

Jeon, J.S., Ryoo, N., Hahn, T.R., Walia, H., and Nakamura, Y. (2010). Starch biosynthesis in cereal endosperm. Plant Physiol. Bioch. **48**, 383-392.

Johnson, L. (1989). Glycogen phosphorylase: A multifaceted enzyme. Carlsberg Res. Commun. **54**, 203-229.

Johnson, L.N. (1992). Glycogen-Phosphorylase - Control by Phosphorylation and Allosteric Effectors. Faseb J. **6**, 2274-2282.

Kammerer, B., Fischer, K., Hilpert, B., Schubert, S., Gutensohn, M., Weber, A., and Flügge, U.-I. (1998). Molecular Characterization of a Carbon Transporter in Plastids from Heterotrophic Tissues: The Glucose 6-Phosphate/Phosphate Antiporter. Plant Cell **10**, 105-118.

Kitzman, J.O., MacKenzie, A.P., Adey, A., Hiatt, J.B., Patwardhan, R.P., Sudmant, P.H., Ng, S.B., Alkan, C., Qiu, R.L., Eichler, E.E., and Shendure, J. (2011). Haplotype-resolved genome sequencing of a Gujarati Indian individual. Nat. Biotechnol. **29**, 59-63.

Krause, K.P., Hill, L., Reimholz, R., Nielsen, T.H., Sonnewald, U., and Stitt, M. (1998). Sucrose metabolism in cold-stored potato tubers with decreased expression of sucrose phosphate synthase. Plant Cell Environ. **21**, 285-299.

Krebs, E.G., and Fischer, E.H. (1956). The Phosphorylase-B to Phosphorylase-a Converting Enzyme of Rabbit Skeletal Muscle. Biochim. Biophys. Acta **20**, 150-157.

Krebs, E.G., and Fischer, E.H. (1956). The Phosphorylase-B to Phosphorylase-a Converting Enzyme of Rabbit Skeletal Muscle. Biochimica Et Biophysica Acta **20,** 150-157.

Kumar, G.N.M., Lyer, S., and Knowles, N.R. (2007). Extraction of RNA from fresh, frozen, and lyophilized tuber and root tissues. Journal of Agricultural and Food Chemistry **55,** 1674-1678.

Kumar, S., Tamura, K., and Nei, M. (2004). MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings in Bioinformatics **5**, 150-163.

Laemmli, U.K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Nature **227**, 680-685.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., and Higgins, D.G. (2007). Clustal W and Clustal X version 2.0. Bioinformatics **23**, 2947-2948.

Leach, L.J., Wang, L., Kearsey, M.J., and Luo, Z. (2010). Multilocus tetrasomic linkage analysis using hidden Markov chain model. Proc. Natl. Acad. Sci. **107**, 4270-4274.

Leister, D., Ballvora, A., Salamini, F., and Gebhardt, C. (1996). A PCR-based approach for isolating pathogen resistance genes from potato with potential for wide application in plants. Nat. Genet. **14**, 421-429.

Li, L., Paulo, M.J., Strahwald, J., Lubeck, J., Hofferbert, H.R., Tacke, E., Junghans, H., Wunder, J., Draffehn, A., van Eeuwijk, F., and Gebhardt, C. (2008). Natural DNA variation at candidate loci is associated with potato chip color, tuber starch content, yield and starch yield. Theor. Appl. Genet. **116**, 1167-1181.

Li, L., Strahwald, J., Hofferbert, H.R., Lubeck, J., Tacke, E., Junghans, H., Wunder, J., and Gebhardt, C. (2005). DNA variation at the invertase locus invGE/GF is associated with tuber quality traits in populations of potato breeding clones. Genetics **170**, 813 - 821.

Liu, X., Song, B., Zhang, H., Li, X.-Q., Xie, C., and Liu, J. (2010). Cloning and molecular characterization of putative invertase inhibitor genes and their possible contributions to cold-induced sweetening of potato tubers. MGG Molecular Genetics and Genomics **284**.

Lloyd, J.R., Blennow, A., Burhenne, K., and Kossmann, J. (2004). Repression of a novel isoform of disproportionating enzyme (stDPE2) in potato leads to inhibition of starch degradation in leaves but not tubers stored at low temperature. Plant Physiol. **134**, 1347-1354.

Loiselle, F., Tai, G., and Christie, B. (1990). Genetic components of chip color evaluated after harvest, cold storage and reconditioning. Am. J. Potato Res. **67**, 633-646.

Lorberth, R., Ritte, G., Willmitzer, L., and Kossmann, J. (1998). Inhibition of a starch-granule-bound protein leads to modified starch and repression of cold sweetening. Nat. Biotechnol. **16**, 473-477.

Lu, Y., and Sharkey, T.D. (2006). The importance of maltose in transitory starch breakdown. Plant Cell Environ. **29,** 353-366.

Luo, Z.W., Zhang, Z., Leach, L., Zhang, R.M., Bradshaw, J.E., and Kearsey, M.J. (2006). Constructing genetic linkage maps under a tetrasomic model. Genetics **172**, 2635-2645.

Luthra, S.K., Gopal, J., Kumar, D., Singh, B.P., and Pandey, S.K. (2009). Solanum Wild and Cultivated Species as Source of Resistance to Cold Induced Sweetening. Potato Journal (Shimla) **36**, 115-120.

Lytovchenko, A., Sonnewald, U., and Fernie, A.R. (2007). The complex network of non-cellulosic carbohydrate metabolism. Curr. Opin. Plant Biol. **10**, 227-235.

Malone, J.G., Mittova, V., Ratcliffe, R.G., and Kruger, N.J. (2006). The response of carbohydrate metabolism in potato tubers to low temperature. Plant Cell Physiol. **47**, 1309-1322.

Manchenko, G.P. (2003). Handbook of detection of enzymes on electrophoretic gels. (Boca Ratón: CRC Press).

Masso, M., and Vaisman, I.I. (2010). AUTO-MUTE: web-based tools for predicting stability changes in proteins due to single amino acid replacements. Protein Eng. Des. Sel. **23**, 683-687.

Matney, T.S., E.P., G., Erwin, N.S., and Scroggs, R.A. (1964). A preliminary map of genomic sites for F-attachment in *Escherichia coli* K12. Biochem. Biophys. Res. Commun. **17**, 278-281.

McCann, L.C., Bethke, P.C., and Simon, P.W. (2010). Extensive Variation in Fried Chip Color and Tuber Composition in Cold-Stored Tubers of Wild Potato (*Solanum*) Germplasm. J. Agr. Food Chem. **58**, 2368-2376.

McWilliam, H., Valentin, F., Goujon, M., Li, W., Narayanasamy, M., Martin, J., Miyar, T., and Lopez, R. (2009). Web services at the European Bioinformatics Institute-2009. Nucleic Acids Res. **37**, W6-W10.

Menendez, C.M., Ritter, E., Schafer-Pregl, R., Walkemeier, B., Kalde, A., Salamini, F., and Gebhardt, C. (2002). Cold sweetening in diploid potato: Mapping quantitative trait loci and candidate genes. Genetics **162**, 1423-1434.

Meyerhans, A., Vartanian, J.-P., and Wain-Hobson, S. (1990). DNA recombination during PCR. Nucleic Acids Res. **18**, 1687-1691.

Milbourne, D., Meyer, R.C., Collins, A.J., Ramsay, L.D., Gebhardt, C., and Waugh, R. (1998). Isolation, characterization and mapping of simple sequence repeat loci in potato. Mol. Gen. Genet. **259**, 233 - 245.

Morton, N. (2005). Linkage disequilibrium maps and association mapping. The Journal of Clinical Investigations **6**, 1425-1430.

Mottram, D.S., Wedzicha, B.L., and Dodson, A.T. (2002). Acrylamide is formed in the Maillard reaction. Nature **419**, 448-449.

Mueller, L.A., Solow, T.H., Taylor, N., Skwarecki, B., Buels, R., Binns, J., Lin, C., Wright, M.H., Ahrens, R., Wang, Y., Herbst, E.V., Keyder, E.R., Menda, N., Zamir, D., and Tanksley, S.D. (2005). The SOL Genomics Network. A Comparative Resource for Solanaceae Biology and Beyond. Plant Physiol. **138**, 1310-1317.

Müller-Thurgau, H. (1882). Über Zuckerhäufung in Pflanzenteilen in Folge niedriger Temperatur. Landwirtschaftliche Jahrbücher **11,** 751-828.

Nakano, K., and Fukui, T. (1986). The Complete Amino-Acid-Sequence of Potato Alpha-Glucan Phosphorylase. J. Biol. Chem. **261**, 8230-8236.

Nakano, K., Mori, H., and Fukui, T. (1989). Molecular-Cloning of cDNA-Encoding Potato Amyloplast Alpha-Glucan Phosphorylase and the Structure of Its Transit Peptide. Journal of Biochemistry **106**, 691-695.

Navratil, O., Fischer, L., Cmejlova, J., Linhart, M., and Vacek, J. (2007). Decreased amount of reducing sugars in transgenic potato tubers and its influence on yield characteristics. Biol. Plant. **51,** 56-60.

Neigenfind, J., Gyetvai, G., Basekow, R., Diehl, S., Achenbach, U., Gebhardt, C., Selbig, J., and Kersten, B. (2008). Haplotype inference from unphased SNP data in heterozygous polyploids based on SAT. BMC Genomics **9**, 356.

Neuhaus, H.E., and Emes, M.J. (2000). Nonphotosynthetic metabolism in plastids. Annu. Rev. Plant Phys. **51**, 111-140.

Nicot, N., Hausman, J.F., Hoffmann, L., and Evers, D. (2005). Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. J. Exp. Bot. **56**, 2907-2914.

Palm, D., Goerl, R., and Burger, K.J. (1985). Evolution of Catalytic and Regulatory Sites in Phosphorylases. Nature **313,** 500-502.

Pereira, A., Coffin, R., Yada, R., and Machado, V. (1993). Inheritance patterns of reducing sugars in potato tubers after storage at 12 C and 4 C followed by reconditioning. Am. J. Potato Res. **70**, 71-76.

Pienaar, E., Theron, M., Nelson, M., and Viljoen, H.J. (2006). A Quantitative Model Of Error Accumulation During PCR Amplification. Comput. Biol. Chem. **30**, 102-111.

Pluthero, F.G. (1993). Rapid Purification of High-Activity Taq DNA-Polymerase. Nucleic Acids Res. **21**, 4850-4851.

Powell, W. and Uhrig, H. (1987). Anther culture of *Solanum* genotypes. *Plant Cell Tissue Organ Cult.* **11**, (1987), pp. 13–24.

Rand, K.N. (1996). Crystal violet can be used to visualize DNA bands during gel electrophoresis and to improve cloning efficiency. Elsevier Trends Journals Technical Tips Online **T40022**.

Reinhold, T., Ali, A., Grimm, B., Beran, K.C., Jahns, P., Conrath, U., Bauer, J., Reiser, J., Melzer, M., Jeblick, W., and Neuhaus, H.E. (2007). Limitation of nocturnal import of ATP into Arabidopsis chloroplasts leads to photo oxidative damage. Plant J. 5**0**, 293-304.

Rickert, A.M., Kim, J.H., Meyer, S., Nagel, A., Ballvora, A., Oefner, P.J., and Gebhardt, C. (2003). First-generation SNP/InDel markers tagging loci for pathogen resistance in the potato genome. Plant Biotechnology Journal **1,** 399-410.

Rommens, C.M., Ye, J.S., Richael, C., and Swords, K. (2006). Improving potato storage and processing characteristics through all-native DNA transformation. J. Agr. Food Chem. **54**, 9882-9887.

Rozen, A., and Skaletsky, H.J. (2000). Primer3 on the WWW for general users and for biologist programmers. In Bioinformatics Methods and Protocols: Methods in Molecular Biology, S. Krawetz and S. Misener, eds (Totowa: Humana Press), pp. 365-386.

Sambroock, J., and Russell, D. (2001). Molecular cloning: A laboratory manual. (New York: Cold Spring Harbor Laboratory Press).

Satoh, H., Shibahara, K., Tokunaga, T., Nishi, A., Tasaki, M., Hwang, S.-K., Okita, T.W., Kaneko, N., Fujita, N., Yoshida, M., Hosaka, Y., Sato, A., Utsumi, Y., Ohdan, T., and Nakamura, Y. (2008). Mutation of the Plastidial Alpha-Glucan Phosphorylase Gene in Rice Affects the Synthesis and Structure of Starch in the Endosperm. Plant Cell **20**, 1833-1849.

Schäfer-Pregl, R., Ritter, E., Concilio, L., Hesselbach, J., Lovatti, L., Walkemeier, B., Thelen, H., Salamini, F., and Gebhardt, C. (1998). Analysis of quantitative trait loci (QTLs) and quantitative trait alleles (QTAs) for potato tuber yield and starch content. Theor. Appl. Genet. **97**, 834-846.

Schwartz, M. (1966). Location of Maltose A and B Loci on Genetic Map of *Escherichia Coli*. Journal of Bacteriol. **92**, 1083-&.

Seidman, C.E., Struhl, K., Sheen, J., and Jessen, T. (2001). Introduction of Plasmid DNA into Cells. (John Wiley & Sons, Inc.).

Shallenberger, R.S., Smith, O., and Treadway, R.H. (1959). Food Color Changes, Role of the Sugars in the Browning Reaction in Potato Chips. J. Agr. Food Chem. **7**, 274-277.

Simko, I. (2004). One potato, two potato: haplotype association mapping in autotetraploids. Trends Plant Sci. **9**, 441-448.

Sonnewald, U., Basner, A., Greve, B., and Steup, M. (1995). A 2nd L-Type Isozyme of Potato Glucan Phosphorylase - Cloning, Antisense Inhibition and Expression Analysis. Plant Mol. Biol. **27**, 567-576.

Sowokinos, J.R. (2001). Biochemical and molecular control of cold-induced sweetening in potatoes. Am. J. Potato Res. **78**, 221-236.

Spooner, D.M., McLean, K., Ramsay, G., Waugh, R., and Bryan, G.J. (2005). A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. Proc. Natl. Acad. Sci. USA **102**, 14694-14699.

Spychalla, J.P., Scheffler, B.E., Sowokinos, J.R., and Bevan, M.W. (1994). Cloning, Antisense RNA Inhibition, and the Coordinated Expression of Udp-Glucose Pyrophosphorylase with Starch Biosynthetic Genes in Potato-Tubers. Plant Physiol. **144**, 444-453.

St. Pierre, B., and Brisson, N. (1995). 5'-Deletion Analysis of the Potato Starch Phosphorylase Gene - an Upstream Sequence Defines Distal Regulatory Elements and a Proximal Organ-Dependent Promoter. Plant Sci. **110**, 193-203.

Stadler, R.H., Blank, I., Varga, N., Robert, F., Hau, J., Guy, P.A., Robert, M.C., and Riediker, S. (2002). Acrylamide from Maillard reaction products. Nature **419**, 449-450.

Stensballe, A., Hald, S., Bauw, G., Blennow, A., and Welinder, K.G. (2008). The amyloplast proteome of potato tuber. FEBS Journal **275,** 1723-1741.

Suckau, D., Resemann, A., Schuerenberg, M., Hufnagel, P., Franzen, J., and Holle, A. (2003). A novel MALDI LIFT-TOF/TOF mass spectrometer for proteomics. Analytical and Bioanalytical Chemistry **376**, 952-965.

Sun, F. (1995). The polymerase chain reaction and branching processes. J Comput. Biol. **2**, 63-86.

Swarbreck, D., Wilks, C., Lamesch, P., Berardini, T.Z., Garcia-Hernandez, M., Foerster, H., Li, D., Meyer, T., Muller, R., Ploetz, L., Radenbaugh, A., Singh, S., Swing, V., Tissier, C., Zhang, P., and Huala, E. (2008). The Arabidopsis Information Resource (TAIR): gene structure and function annotation. Nucleic Acids Res. **36**, D1009-D1014.

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.

Tanizawa, K., Mori, H., Tagaya, M., and Fukui, T. (1994). Higher plant phosphorylases. Molecular aspects of enzyme catalysis**,** 107-126.

Tanksley, S.D., Ganal, M.W., Prince, J.P., Devicente, M.C., Bonierbale, M.W., Broun, P., Fulton, T.M., Giovannoni, J.J., Grandillo, S., Martin, G.B., Messeguer, R., Miller, J.C., Miller, L., Paterson, A.H., Pineda, O., Roder, M.S., Wing, R.A., Wu, W., and Young, N.D. (1992). High-Density Molecular Linkage Maps of the Tomato and Potato Genomes. Genetics **132**, 1141-1160.

The UniProt Consortium. Ongoing and future developments at the Universal Protein Resource. Nucleic Acids Res. **39**, D214-D219.

Urbany, C., Stich, B., Schmidt, L., Simon, L., Berding, H., Junghans, H., Niehoff, K.H., Braun, A., Tacke, E., Hofferbert, H.R., Lubeck, J., Strahwald, J., and Gebhardt, C. (2011). Association

genetics in *Solanum tuberosum* provides new insights into potato tuber bruising and enzymatic tissue discoloration. Bmc Genomics **12**.

van Eck, H.J., Vandervoort, J.R., Draaistra, J., Vanzandvoort, P., Vanenckevort, E., Segers, B., Peleman, J., Jacobsen, E., Helder, J., and Bakker, J. (1995). The Inheritance and Chromosomal Localization of AFLP markers in a Noninbred Potato Offspring. Mol. Breeding **1**, 397-410.

van Os, H., Andrzejewski, S., Bakker, E., Barrena, I., Bryan, G.J., Caromel, B., Ghareeb, B., Isidore, E., de Jong, W., van Koert, P., Lefebvre, V., Milbourne, D., Ritter, E., van der Voort, J., Rousselle-Bourgeois, F., van Vliet, J., Waugh, R., Visser, R.G.F., Bakker, J., and van Eck, H.J. (2006). Construction of a 10,000-marker ultradense genetic recombination map of potato: Providing a framework for accelerated gene isolation and a genomewide physical map. Genetics **173**, 1075-1087.

Vidal, M. (1997). The Reverse Two-Hybrid System. In The Two-Hybrid System, P. Bartel and S. Fields, eds (New York: Oxford University Press), pp. 109.

Wang, G.C.Y., and Wang, Y. (1996). The frequency of chimeric molecules as a consequence of PCR co-amplification of 16S rRNA genes from different bacterial species. Microbiology **142**, 1107-1114.

Watson, K.A., Schinzel, R., Palm, D., and Johnson, L.N. (1997). The crystal structure of Escherichia coli maltodextrin phosphorylase provides an explanation for the activity without control in this basic archetype of a phosphorylase. Embo Journal **16,** 1-14.

Weckx, S., Del-Favero, J., Rademakers, R., Claes, L., Cruts, M., De Jonghe, P., Van Broeckhoven, C., and De Rijk, P. (2005). novoSNP, a novel computational tool for sequence variation discovery. Genome Res. **15**, 436-442.

Yan, H., Chretien, R., Ye, J., and Rommens, C.M. (2006). New Construct Approaches for Efficient Gene Silencing in Plants. Plant Physiol. 1**41**, 1508-1518.

Ye, J., Shakya, R., Shrestha, P., and Rommens, C.M. (2010). Tuber-Specific Silencing of the Acid Invertase Gene Substantially Lowers the Acrylamide-Forming Potential of Potato. J. Agr. Food Chem. **58**, 12162-12167.

Zeeman, S.C., Delatte, T., Messerli, G., Umhang, M., Stettler, M., Mettler, T., Streb, S., Reinhold, H., and Kotting, O. (2007). Starch breakdown: recent discoveries suggest distinct pathways and novel mechanisms. Functional Plant Biology **34**, 465-473.

Zeeman, S.C., Smith, S.M., and Smith, A.M. (2004a). The breakdown of starch in leaves. New Phytol. **163**, 247-261.

Zeeman, S.C., Thorneycroft, D., Schupp, N., Chapple, A., Weck, M., Dunstan, H., Haldimann, P., Bechtold, N., Smith, A.M., and Smith, S.M. (2004b). Plastidial alpha-glucan phosphorylase is not required for starch degradation in arabidopsis leaves but has a role in the tolerance of abiotic stress. Plant Physiol. **135**, 849-858.

Zrenner, R., Schuler, K., and Sonnewald, U. (1996). Soluble acid invertase determines the hexose-to-sucrose ratio in cold-stored potato tubers. Planta **198**, 246-252.

Abstract

The enzymatic degradation of tuber starch and the subsequent accumulation of reducing sugars upon cold storage, also known as cold sweetening, are complex traits that negatively affect potato processing quality, constituting a challenge for breeding programs. In an effort to decipher the molecular basis of cold sweetening, a previous association mapping study identified DNA polymorphisms in functional genes of the carbohydrate metabolism, which were associated with tuber starch content and chip quality before and after cold storage. Two of the associated alleles encoded the plastidial starch phosphorylases *Pho1a* and *Pho1b* located on chromosomes III and V, respectively. The main objective of this research was to functionally characterize allelic variants of starch phosphorylase that possibly contribute to the variation of quality traits.

Full length *Pho1a* cDNA's were cloned from four tetraploid potato cultivars. The *Pho1a-H_A* allele was identified, which represented the molecular markers associated with tuber quality traits. The association of *Pho1a-H_A* was further confirmed in two additional populations, in which genotypes bearing the allele accumulated fewer reducing sugars during cold storage and had higher contents of starch. The *Pho1a-H_A* allele was negatively associated with resistance to tuber bruising. Despite the presence of polymorphisms in areas of positive and negative regulation of the *Pho1a* promoter, the expression in tubers of the *Pho1a-H_A* allele was constant during cold storage and corresponded to the allele dosage, showing that the allele is not subjected to transcriptional regulation. The presence of the PHO1a-H_A isoform in protein extracts was confirmed by 2D-gel electrophoresis and mass spectrometry. This isoform contains six amino acid substitutions from which one has no precedents in all known glucan phosphorylases. Activity assays demonstrated that genotypes bearing the PHO1a-H_A isoform in a duplex allele dosage had no detectable starch phosphorylase activity, suggesting that *Pho1a-H_A* is an allele which compromises functional starch phosphorylase enzyme complexes. The negative effect of the novel substitutions in key regions for protein dimerization, shown by *in silico* mutagenesis in the crystal structure of the rabbit muscle glycogen phosphorylase, supports this hypothesis. The association of *Pho1a-H_A* with quality traits in different populations and the detrimental effect of the allele on enzyme activity support the hypothesis that natural variants of the *Pho1a* gene contribute to the phenotypic variation in tuber quality traits and provides further insights into the understanding of sugar accumulation at low temperatures.

Additionally, natural alleles of the *Pho1b* gene were identified. The large amount of SNPs limited the phasing and haplotype assembly. Non-synonymous SNPs however, allowed the clustering of the alleles which were assigned to the different SSCP polymorphisms associated with quality traits.

The broad genetic diversity of potato cultivars and wild relatives offers a wealth of natural resources, which genetic potential can be exploited to improve agronomic performance. The PCR-based molecular markers developed in this work can be integrated by breeders in their selection programs.

Zusammenfassung

Der enzymatische Abbau von Stärke in Kartoffelknollen während der Lagerung bei niedrigen Temperaturen und die dadurch auftretende Zuckeranreicherung wird als „Cold Sweetening“ bezeichnet. Diese Eigenschaft, die von verschiedenen Genen beeinflusst wird, hat einen negativen Einfluss auf die industrielle Weiterverarbeitung und wird daher in Züchtungsprogrammen berücksichtigt. In einer zuvor durchgeführten Assoziationsstudie, um die molekularen Vorgänge während des „cold Sweetening“-Prozesses zu verstehen, wurden DNA Polymorphismen in Genen des Kohlenhydratmetabolismus identifiziert, und Assoziationen mit dem Stärkegehalt von Kartoffelknollen und der Verarbeitbarkeitsqualität vor und nach der Kühlhauslagerung gefunden. Zwei der assoziierten Allele kodieren für die Stärke-Phosphorylasen *Pho1a* und *Pho1b,* die auf den Chromosomen III und V lokalisiert sind. Das Ziel dieser Studie war die funktionelle Charakterisierung der Allelvarianten der Stärke-Phosphorylase und ihr Einfluss auf Qualitätsmerkmale.

Eine Analyse von *Pho1a* cDNA-Sequenzen, die aus tetraploiden Kartoffelsorten isoliert wurden, führte zur Identifikation des *Pho1a-H$_A$*-Allels, welches mit Qualitätsmerkmalen assoziiert ist. Diese Assoziation konnte in weiteren Populationen bestätigt werden, wobei Pflanzen die das *Pho1a-H$_A$*-Allel trugen während der Kühlhauslagerung weniger Zuckeranreicherung zeigten und höhere Stärkekonzentrationen aufwiesen. Des Weiteren konnte ein negativer Zusammenhang mit der Resistenz gegenüber Schwarzfleckingkeit/Schlagschäden belegt werden. Obwohl Polymorphismen in *Pho1a*-Promotorbereichen gefunden wurden, zeigten Expressionsanalysen, dass das *Pho1a-H$_A$* auch während der Kühllagerung im Vergleich zu anderen Allelen nicht transkriptionell reguliert wird. Die Anwesenheit der PHO1a-H$_A$-Isoform in Proteinextrakten konnte durch 2D-Gelelektrophoreseanalysen und massensprektrometrisch bestätigt werden. Diese Isoform enthält sechs Amionsäuresubstitutionen, von denen eine in keiner bekannten Glucan-Phosphorylase auftritt. Untersuchungen der Phosphorylaseaktivität von verschiedenen Sorten zeigte, dass in Pflanzen mit zwei *Pho1a-H$_A$* Allelen keine messbare Aktivität mehr festgestellt werden konnte, was darauf hindeutet, dass das *Pho1a-H$_A$*-Allel einen negativen Einfluss auf die Funktionalität von Stärke-Phosphorylasekomplexen hat. Diese Hypothese wird durch *in silico*-Mutageneseanalysen an der Kristallstruktur der Kaninchenmuskel-Glycogen-Phosphorylase unterstützt, die zeigen, dass die Substitutionen in *Pho1a-H$_A$* in Schlüsselregionen der Protein-Dimerisierung liegen. Die Assoziation von *Pho1a-H$_A$* mit

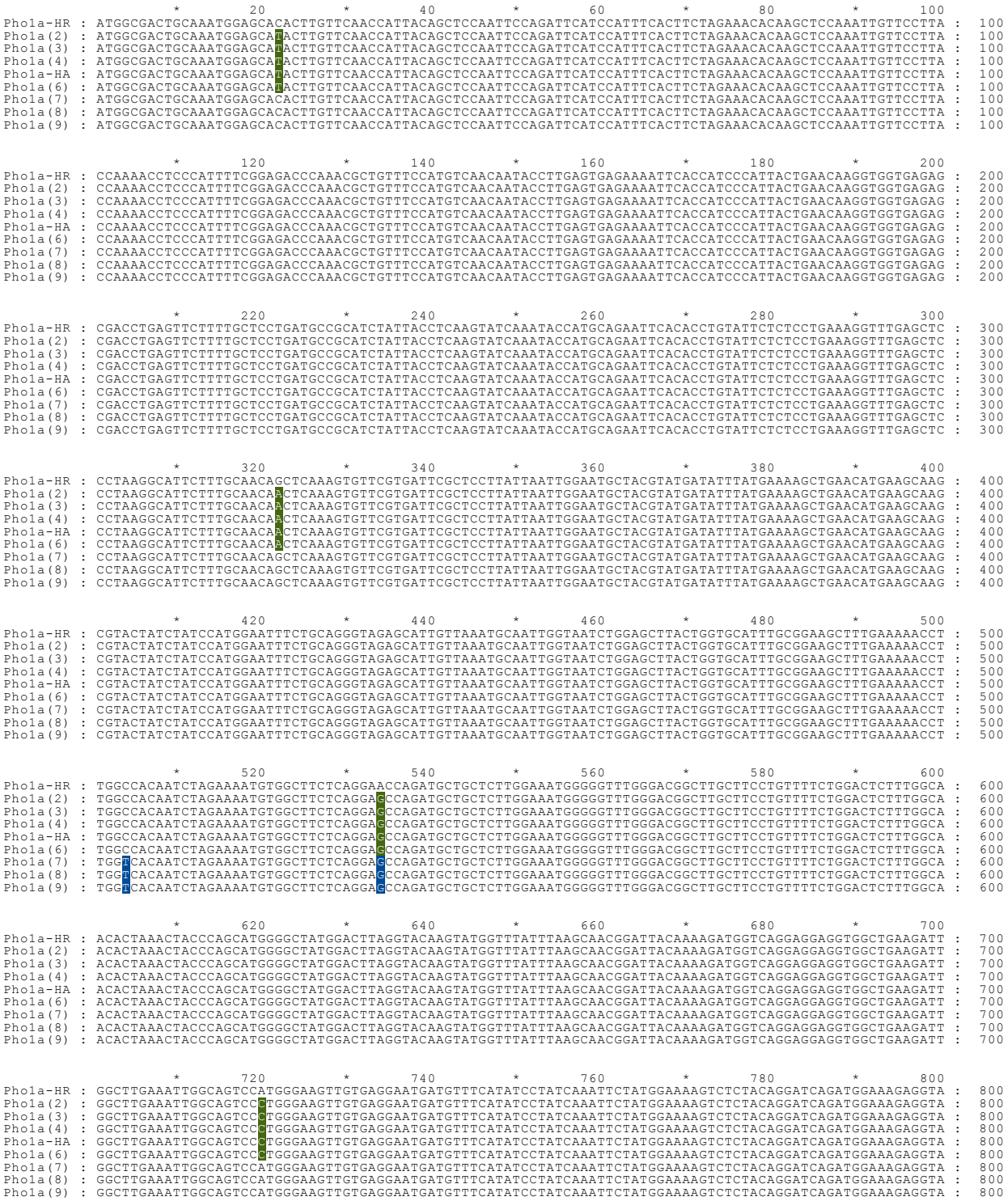
Qualitätsmerkmalen und der negative Effekt auf die Phosphorylaseaktivität befürworten die Hypothese, dass natürliche Varianten des *Pho1a*-Gens zur phänotypischen Vielfalt beitragen, und erweitern das Verständnis des Prozesses der Zuckeranreicherung bei niedrigen Temperaturen.

Des Weiteren wurden natürliche Allele des *Pho1b*-Gens identifiziert. Die große Anzahl von SNPs behinderte hier die Identifizierung von Haplotypen. Allerdings konnten durch die Analyse nicht-synonymer SNPs Proteinsequenzen mehrerer Allele ermittelt werden, welche dann verschiedenen SSCP-Polymorphismen, die mit Qualitätsmerkmalen assoziiert sind, zugeordnet werden konnten.

Die ausgeprägte genetische Vielfalt von Kartoffelsorten und Wildformen bietet eine Fülle von Ressourcen, deren genetisches Potential zur Verbesserung der landwirtschaftlichen Leistungsfähigkeit genutzt werden kann. Die PCR-basierten molekularen Marker, die im Rahmen dieser Arbeit entwickelt wurden, können in Zuchtprogrammen eingesetzt werden.

APPENDICES

Appendix 1. Sequence alignment of *Pho1a* cDNA putative alleles



```
                  *       820         *       840         *       860         *       880         *       900
Phola-HR : TTGGATTGGTGGAGAGGATATAAAGGCAGTTGCGTATGATGTTCCCATACCAGGGTATAAGACCAGAACCACAATCAGCCTTCGACTGTGGTCTACACAG :  900
Phola(2) : TTGGATTGGTGGAGAGGATATAAAGGCAGTTGCGTATGATGTTCCCATACCAGGGTATAAGACCAGAACCACAATCAGCCTTCGACTGTGGTCTACACAG :  900
Phola(3) : TTGGATTGGTGGAGAGGATATAAAGGCAGTTGCGTATGATGTTCCCATACCAGGGTATAAGACCAGAACCACAATCAGCCTTCGACTGTGGTCTACACAG :  900
Phola(4) : TTGGATTGGTGGAGAGGATATAAGGGCAGTTGCGTATGATGTTCCCATACCAGGGTATAAGACCAGAACCACAATCAGCCTTCGACTGTGGTCTACACAG :  900
Phola-HA : TTGGATTGGTGGAGAGGATATAAGGGCAGTTGCGTATGATGTTCCCATACCAGGGTATAAGACCAGAACCACAATCAGCCTTCGACTGTGGTCTACACAG :  900
Phola(6) : TTGGATTGGTGGAGAGGATATAAGGGCAGTTGCGTATGATGTTCCCATACCAGGGTATAAGACCAGAACCACAATCAGCCTTCGACTGTGGTCTACACAG :  900
Phola(7) : TTGGATTGGTGGAGAGGATATAAAGGCAGTTGCGTATGATGTTCCCATACCAGGGTATAAGACCAGAACCACAATCAGCCTTCGACTGTGGTCTACACAG :  900
Phola(8) : TTGGATTGGTGGAGAGGATATAAAGGCAGTTGCGTATGATGTTCCCATACCAGGGTATAAGACCAGAACCACAATCAGCCTTCGACTGTGGTCTACACAG :  900
Phola(9) : TTGGATTGGTGGAGAGGATATAAAGGCAGTTGCGTATGATGTTCCCATACCAGGGTATAAGACCAGAACCACAATCAGCCTTCGACTGTGGTCTACACAG :  900

                  *       920         *       940         *       960         *       980         *      1000
Phola-HR : GTTCCATCAGCGGATTTTGATTTATCTGCTTTCAATGCTGGAGAGCACACCAAAGCATGTGAAGCCCAAGCAAACGCTGAGAAGATATGTTACATACTCT : 1000
Phola(2) : GTTCCATCAGCGGATTTTGATTTATCTGCTTTCAATGCTGGAGAGCACACCAAAGCATGTGAAGCCCAAGCAAACGCTGAGAAGATATGTTACATACTCT : 1000
Phola(3) : GTTCCATCAGCGGATTTTGATTTATCTGCTTTCAATGCTGGAGAGCACACCAAAGCATGTGAAGCCCAAGCAAACGCTGAGAAGATATGTTACATACTCT : 1000
Phola(4) : GTTCCATCAGCGGATTTTGATTTATCTGCTTTCAATGCTGGAGAGCACACCAAAGCATGTGAAGCCCAAGCAAACGCTGAGAAGATATGTTACATACTCT : 1000
Phola-HA : GTTCCATCAGCGGATTTTGATTTATCTGCTTTCAATGCTGGAGAGCACACCAAAGCATGTGAAGCCCAAGCAAACGCTGAGAAGATATGTTACATACTCT : 1000
Phola(6) : GTTCCATCAGCGGATTTTGATTTATCTGCTTTCAATGCTGGAGAGCACACCAAAGCATGTGAAGCCCAAGCAAACGCTGAGAAGATATGTTACATACTCT : 1000
Phola(7) : GTTCCATCAGCGGATTTTGATTTATCTGCTTTCAATGCTGGAGAGCACACCAAAGCATGTGAAGCCCAAGCAAACGCTGAGAAGATATGTTACATACTCT : 1000
Phola(8) : GTTCCATCAGCGGATTTTGATTTATCTGCTTTCAATGCTGGAGAGCACACCAAAGCATGTGAAGCCCAAGCAAACGCTGAGAAGATATGTTACATACTCT : 1000
Phola(9) : GTTCCATCAGCGGATTTTGATTTATCTGCTTTCAATGCTGGAGAGCACACCAAAGCATGTGAAGCCCAAGCAAACGCTGAGAAGATATGTTACATACTCT : 1000

                  *      1020         *      1040         *      1060         *      1080         *      1100
Phola-HR : ACCCTGGGGATGAATCAGAGGAGGGAAAGATCCTTCGGTTGAAGCAACAATATACCTTATGCTCGGCTTCTCTCCAAGATATTATTTCTCGATTTGAGAG : 1100
Phola(2) : ACCCTGGGGATGAATCAGAGGAGGGAAAGATCCTTCGGTTGAAGCAACAATATACCTTATGCTCGGCTTCTCTCCAAGATATTATTTCTCGATTTGAGAG : 1100
Phola(3) : ACCCTGGGGATGAATCAGAGGAGGGAAAGATCCTTCGGTTGAAGCAACAATATACCTTATGCTCGGCTTCTCTCCAAGATATTATTTCTCGATTTGAGAG : 1100
Phola(4) : ACCCTGGGGATGAATCAGAGGAGGGAAAGATCCTTCGGTTGAAGCAACAATATACCTTATGCTCGGCTTCTCTCCAAGATATTATTTCTCGATTTGAGAG : 1100
Phola-HA : ACCCTGGGGATGAATCAGAGGAGGGAAAGATCCTTCGGTTGAAGCAACAATATACCTTATGCTCGGCTTCTCTCCAAGATATTATTTCTCGATTTGAGAG : 1100
Phola(6) : ACCCTGGGGATGAATCAGAGGAGGGAAAGATCCTTCGGTTGAAGCAACAATATACCTTATGCTCGGCTTCTCTCCAAGATATTATTTCTCGATTTGAGAG : 1100
Phola(7) : ACCCTGGGGATGAATCAGAGGAGGGAAAGATCCTTCGGTTGAAGCAACAATATACCTTATGCTCGGCTTCTCTCCAAGATATTATTTCTCGATTTGAGAG : 1100
Phola(8) : ACCCTGGGGATGAATCAGAGGAGGGAAAGATCCTTCGGTTGAAGCAACAATATACCTTATGCTCGGCTTCTCTCCAAGATATTATTTCTCGATTTGAGAG : 1100
Phola(9) : ACCCTGGGGATGAATCAGAGGAGGGAAAGATCCTTCGGTTGAAGCAACAATATACCTTATGCTCGGCTTCTCTCCAAGATATTATTTCTCGATTTGAGAG : 1100

                  *      1120         *      1140         *      1160         *      1180         *      1200
Phola-HR : GAGATCAGGTGATCGTATTAAGTGGGAAGAGTTTCCTGAAAAAGTTGCTGTGCAGATGAATGACACTCACCCTACACTTTGTATCCCTGAGCTGATGAGA : 1200
Phola(2) : GAGATCAGGTGATCGTATTAAGTGGGAAGAGTTTCCTGAAAAAGTTGCTGTGCAGATGAATGACACTCACCCTACACTTTGTATCCCTGAGCTGATGAGA : 1200
Phola(3) : GAGATCAGGTGATCGTATTAAGTGGGAAGAGTTTCCTGAAAAAGTTGCTGTGCAGATGAATGACACTCACCCTACACTTTGTATCCCTGAGCTGATGAGA : 1200
Phola(4) : GAGATCAGGTGATCGTATTAAGTGGGAAGAGTTTCCTGAAAAAGTTGCTGTGCAGATGAATGACACTCACCCTACACTTTGTATCCCTGAGCTGATGAGA : 1200
Phola-HA : GAGATCAGGTGATCGTATTAAGTGGGAAGAGTTTCCTGAAAAAGTTGCTGTGCAGATGAATGACACTCACCCTACACTTTGTATCCCTGAGCTGATGAGA : 1200
Phola(6) : GAGATCAGGTGATCGTATTAAGTGGGAAGAGTTTCCTGAAAAAGTTGCTGTGCAGATGAATGACACTCACCCTACACTTTGTATCCCTGAGCTGATGAGA : 1200
Phola(7) : GAGATCAGGTGATCGTATTAAGTGGGAAGAGTTTCCTGAAAAAGTTGCTGTGCAGATGAATGACACTCACCCTACACTTTGTATCCCTGAGCTGATGAGA : 1200
Phola(8) : GAGATCAGGTGATCGTATTAAGTGGGAAGAGTTTCCTGAAAAAGTTGCTGTGCAGATGAATGACACTCACCCTACACTTTGTATCCCTGAGCTGATGAGA : 1200
Phola(9) : GAGATCAGGTGATCGTATTAAGTGGGAAGAGTTTCCTGAAAAAGTTGCTGTGCAGATGAATGACACTCACCCTACACTTTGTATCCCTGAGCTGATGAGA : 1200

                  *      1220         *      1240         *      1260         *      1280         *      1300
Phola-HR : ATATTGATAGATCTGAAGGGCTTGAATTGGAATGAAGCTTGGAATATTACTCAAAGAACTGTGGCCTACACAAACCATACTGTTTTGCCTGAGGCACTGG : 1300
Phola(2) : ATATTGATAGATCTGAAGGGCTTGAATTGGAATGAAGCTTGGAATATTACTCAAAGAACTGTGGCCTACACAAACCATACTGTTTTGCCTGAGGCACTGG : 1300
Phola(3) : ATATTGATAGATCTGAAGGGCTTGAATTGGAATGAAGCTTGGAATATTACTCAAAGAACTGTGGCCTACACAAACCATACTGTTTTGCCTGAGGCACTGG : 1300
Phola(4) : ATATTGATAGATCTGAAGGGCTTGAATTGGAATGAAGCTTGGAATATTACTCAAAGAACTGTGGCCTACACAAACCATACTGTTTTGCCTGAGGCACTGG : 1300
Phola-HA : ATATTGATAGATCTGAAGGGCTTGAATTGGAATGAAGCTTGGAATATTACTCAAAGAACTGTGGCCTACACAAACCATACTGTTTTGCCTGAGGCACTGG : 1300
Phola(6) : ATATTGATAGATCTGAAGGGCTTGAATTGGAATGAAGCTTGGAATATTACTCAAAGAACTGTGGCCTACACAAACCATACTGTTTTGCCTGAGGCACTGG : 1300
Phola(7) : ATATTGATAGATCTGAAGGGCTTGAATTGGAATGAAGCTTGGAATATTACTCAAAGAACTGTGGCCTACACAAACCATACTGTTTTGCCTGAGGCACTGG : 1300
Phola(8) : ATATTGATAGATCTGAAGGGCTTGAATTGGAATGAAGCTTGGAATATTACTCAAAGAACTGTGGCCTACACAAACCATACTGTTTTGCCTGAGGCACTGG : 1300
Phola(9) : ATATTGATAGATCTGAAGGGCTTGAATTGGAATGAAGCTTGGAATATTACTCAAAGAACTGTGGCCTACACAAACCATACTGTTTTGCCTGAGGCACTGG : 1300

                  *      1320         *      1340         *      1360         *      1380         *      1400
Phola-HR : AGAAATGGAGTTATGAATTGATGCAGAAACTCCTTCCCAGACATGTCGAAATCATTGAGGCGATTGACGAGGAGCTGGTACATGAAATTGTATTAAAATA : 1400
Phola(2) : AGAAATGGAGTTATGAATTGATGCAGAAACTCCTTCCCAGACATGTCGAAATCATTGAGGCGATTGACGAGGAGCTGGTACATGAAATTGTATTAAAATA : 1400
Phola(3) : AGAAATGGAGTTATGAATTGATGCAGAAACTCCTTCCCAGACATGTCGAAATCATTGAGGCGATTGACGAGGAGCTGGTACATGAAATTGTATTAAAATA : 1400
Phola(4) : AGAAATGGAGTTATGAATTGATGCAGAAACTCCTTCCCAGACATGTCGAAATCATTGAGGCGATTGACGAGGAGCTGGTACATGAAATTGTATTAAAATA : 1400
Phola-HA : AGAAATGGAGTTATGAATTGATGCAGAAACTCCTTCCCAGACATGTCGAAATCATTGAGGCGATTGACGAGGAGCTGGTACATGAAATTGTATTAAAATA : 1400
Phola(6) : AGAAATGGAGTTATGAATTGATGCAGAAACTCCTTCCCAGACATGTCGAAATCATTGAGGCGATTGACGAGGAGCTGGTACATGAAATTGTATTAAAATA : 1400
Phola(7) : AGAAATGGAGTTATGAATTGATGCAGAAACTCCTTCCCAGACATGTCGAAATCATTGAGGCGATTGACGAGGAGCTGGTACATGAAATTGTATTAAAATA : 1400
Phola(8) : AGAAATGGAGTTATGAATTGATGCAGAAACTCCTTCCCAGACATGTCGAAATCATTGAGGCGATTGACGAGGAGCTGGTACATGAAATTGTATTAAAATA : 1400
Phola(9) : AGAAATGGAGTTATGAATTGATGCAGAAACTCCTTCCCAGACATGTCGAAATCATTGAGGCGATTGACGAGGAGCTGGTACATGAAATTGTATTAAAATA : 1400

                  *      1420         *      1440         *      1460         *      1480         *      1500
Phola-HR : TGGTTCAATGGATCTGAACAAATTGGAGGAAAAGTTGACTACAATGAGAATCTTAGAAAATTTTGATCTTCCCAGTTCTGTTGCTGAATTATTTATTAAG : 1500
Phola(2) : TGGTTCAATGGATCTGAACAAATTGGAGGAAAAGTTGACTACAATGAGAATCTTAGAAAATTTTGATCTTCCCAGTTCTGTTGCTGAATTATTTATTAAG : 1500
Phola(3) : TGGTTCAATGGATCTGAACAAATTGGAGGAAAAGTTGACTACAATGAGAATCTTAGAAAATTTTGATCTTCCCAGTTCTGTTGCTGAATTATTTATTAAG : 1500
Phola(4) : TGGTTCAATGGATCTGAACAAATTGGAGGAAAAGTTGACTACAATGAGAATCTTAGAAAATTTTGATCTTCCCAGTTCTGTTGCTGAATTATTTATTAAG : 1500
Phola-HA : TGGTTCAATGGATCTGAACAAATTGGAGGAAAAGTTGACTACAATGAGAATCTTAGAAAATTTTGATCTTCCCAGTTCTGTTGCTGAATTATTTATTAAG : 1500
Phola(6) : TGGTTCAATGGATCTGAACAAATTGGAGGAAAAGTTGACTACAATGAGAATCTTAGAAAATTTTGATCTTCCCAGTTCTGTTGCTGAATTATTTATTAAG : 1500
Phola(7) : TGGTTCAATGGATCTGAACAAATTGGAGGAAAAGTTGACTACAATGAGAATCTTAGAAAATTTTGATCTTCCCAGTTCTGTTGCTGAATTATTTATTAAG : 1500
Phola(8) : TGGTTCAATGGATCTGAACAAATTGGAGGAAAAGTTGACTACAATGAGAATCTTAGAAAATTTTGATCTTCCCAGTTCTGTTGCTGAATTATTTATTAAG : 1500
Phola(9) : TGGTTCAATGGATCTGAACAAATTGGAGGAAAAGTTGACTACAATGAGAATCTTAGAAAATTTTGATCTTCCCAGTTCTGTTGCTGAATTATTTATTAAG : 1500

                  *      1520         *      1540         *      1560         *      1580         *      1600
Phola-HR : CCTGAAATCTCAGTTGATGATGATACTGAAACAGTAGAAGTCCATGACAAAGTTGAAGCTTCCGATAAAGTTGTGACTAATGATGAAGATGACACTGGTA : 1600
Phola(2) : CCTGAAATCTCAGTTGATGATGATACTGAAACAGTAGAAGTCCATGACAAAGTTGAAGCTTCCGATAAAGTTGTGACTAATGATGAAGATGACACTGGTA : 1600
Phola(3) : CCTGAAATCTCAGTTGATGATGATACTGAAACAGTAGAAGTCCATGACAAAGTTGAAGCTTCCGATAAAGTTGTGACTAATGATGAAGATGACACTGGTA : 1600
Phola(4) : CCTGAAATCTCAGTTGATGATGATACTGAAACAGTAGAAGTCCATGACAAAGTTGAAGCTTCCGATAAAGTTGTGACTAATGATGAAGATGACACTGGTA : 1600
Phola-HA : CCTGAAATCTCAGTTGATGATGATACTGAAACAGTAGAAGTCCATGACAAAGTTGAAGCTTCCGATAAAGTTGTGACTAATGATGAAGATGACACTGGTA : 1600
Phola(6) : CCTGAAATCTCAGTTGATGATGATACTGAAACAGTAGAAGTCCATGACAAAGTTGAAGCTTCCGATAAAGTTGTGACTAATGATGAAGATGACACTGGTA : 1600
Phola(7) : CCTGAAATCTCAGTTGATGATGATACTGAAACAGTAGAAGTCCATGACAAAGTTGAAGCTTCCGATAAAGTTGTGACTAATGATGAAGATGACACTGGTA : 1600
Phola(8) : CCTGAAATCTCAGTTGATGATGATACTGAAACAGTAGAAGTCCATGACAAAGTTGAAGCTTCCGATAAAGTTGTGACTAATGATGAAGATGACACTGGTA : 1600
Phola(9) : CCTGAAATCTCAGTTGATGATGATACTGAAACAGTAGAAGTCCATGACAAAGTTGAAGCTTCCGATAAAGTTGTGACTAATGATGAAGATGACACTGGTA : 1600

                  *      1620         *      1640         *      1660         *      1680         *      1700
Phola-HR : AGAAAACTAGTGTGAAGATAGAAGCAGCTGCAGAAAAAGACATTGACAAGAAAACTCCCGTGAGTCCGGAACCAGCTGTTATACCACCTAAGAAGGTACG : 1700
Phola(2) : AGAAAACTAGTGTGAAGATAGAAGCAGCTGCAGAAAAAGACATTGACAAGAAAACTCCCGTGAGTCCGGAACCAGCTGTTATACCACCTAAGAAGGTACG : 1700
Phola(3) : AGAAAACTAGTGTGAAGATAGAAGCAGCTGCAGAAAAAGACATTGACAAGAAAACTCCCGTGAGTCCGGAACCAGCTGTTATACCACCTAAGAAGGTACG : 1700
Phola(4) : AGAAAACTAGTGTGAAGATAGAAGCAGCTGCAGAAAAAGACATTGACAAGAAAACTCCCGTGAGTCCGGAACCAGCTGTTATACCACCTAAGAAGGTACG : 1700
Phola-HA : AGAAAACTAGTGTGAAGATAGAAGCAGCTGCAGAAAAAGACATTGACAAGAAAACTCCCGTGAGTCCGGAACCAGCTGTTATACCACCTAAGAAGGTACG : 1700
Phola(6) : AGAAAACTAGTGTGAAGATAGAAGCAGCTGCAGAAAAAGACATTGACAAGAAAACTCCCGTGAGTCCGGAACCAGCTGTTATACCACCTAAGAAGGTACG : 1700
Phola(7) : AGAAAACTAGTGTGAAGATAGAAGCAGCTGCAGAAAAAGACATTGACAAGAAAACTCCCGTGAGTCCGGAACCAGCTGTTATACCACCTAAGAAGGTACG : 1700
Phola(8) : AGAAAACTAGTGTGAAGATAGAAGCAGCTGCAGAAAAAGACATTGACAAGAAAACTCCCGTGAGTCCGGAACCAGCTGTTATACCACCTAAGAAGGTACG : 1700
Phola(9) : AGAAAACTAGTGTGAAGATAGAAGCAGCTGCAGAAAAAGACATTGACAAGAAAACTCCCGTGAGTCCGGAACCAGCTGTTATACCACCTAAGAAGGTACG : 1700
```

```
                  *      1720          *      1740          *      1760          *      1780          *      1800
Phola-HR : CATGGCCAACTTGTGTGTTGTGGGCGGCCATGCTGTTAATGGAGTTGCTGAGATCCATAGTGAAATTGTGAAGGAGGAGGTTTTCAATGACTTCTATGAG : 1800
Phola(2) : CATGGCCAACTTGTGTGTTGTGGGCGGCCATGCTGTTAATGGAGTTGCTGAGATCCATAGTGAAATTGTGAAGGAGGAGGTTTTCAATGACTTCTATGAG : 1800
Phola(3) : CATGGCCAACTTGTGTGTTGTGGGCGGCCATGCTGTTAATGGAGTTGCTGAGATCCATAGTGAAATTGTGAAGGAGGAGGTTTTCAATGACTTCTATGAG : 1800
Phola(4) : CATGGCCAACTTGTGTGTTGTGGGCGGCCATGCTGTTAATGGAGTTGCTGAGATCCATAGTGAAATTGTGAAGGAGGAGGTTTTCAATGACTTCTATGAG : 1800
Phola-HA : CATGGCCAACTTGTGTGTTGTGGGCGGCCATGCTGTTAATGGAGTTGCTGAGATCCATAGTGAAATTGTGAAGGAGGAGGTTTTCAATGACTTCTATGAG : 1800
Phola(6) : CATGGCCAACTTGTGTGTTGTGGGCGGCCATGCTGTTAATGGAGTTGCTGAGATCCATAGTGAAATTGTGAAGGAGGAGGTTTTCAATGACTTCTATGAG : 1800
Phola(7) : CATGGCCAACTTGTGTGTTGTGGGCGGCCATGCTGTTAATGGAGTTGCTGAGATCCATAGTGAAATTGTGAAGGAGGAGGTTTTCAATGACTTCTATGAG : 1800
Phola(8) : CATGGCCAACTTGTGTGTTGTGGGCGGCCATGCTGTTAATGGAGTTGCTGAGATCCATAGTGAAATTGTGAAGGAGGAGGTTTTCAATGACTTCTATGAG : 1800
Phola(9) : CATGGCCAACTTGTGTGTTGTGGGCGGCCATGCTGTTAATGGAGTTGCTGAGATCCATAGTGAAATTGTGAAGGAGGAGGTTTTCAATGACTTCTATGAG : 1800

                  *      1820          *      1840          *      1860          *      1880          *      1900
Phola-HR : CTCTGGCCGGAAAAGTTCCAAAACAAAACAAATGGAGTGACTCCAAGAAGATGGATTCGTTTCTGCAATCCTCCTCTTAGTGCCATCATAACTAAGTGGA : 1900
Phola(2) : CTCTGGCCGGAAAAGTTCCAAAACAAAACAAATGGAGTGACTCCAAGAAGATGGATTCGTTTCTGCAATCCTCCTCTTAGTGCCATCATAACTAAGTGGA : 1900
Phola(3) : CTCTGGCCGGAAAAGTTCCAAAACAAAACAAATGGAGTGACTCCAAGAAGATGGATTCGTTTCTGCAATCCTCCTCTTAGTGCCATCATAACTAAGTGGA : 1900
Phola(4) : CTCTGGCCGGAAAAGTTCCAAAACAAAACAAATGGAGTGACTCCAAGAAGATGGATTCGTTTCTGCAATCCTCCTCTTAGTGCCATCATAACTAAGTGGA : 1900
Phola-HA : CTCTGGCCGGAAAAGTTCCAAAACAAAACAAATGGAGTGACTCCAAGAAGATGGATTCGTTTCTGCAATCCTCCTCTTAGTGCCATCATAACTAAGTGGA : 1900
Phola(6) : CTCTGGCCGGAAAAGTTCCAAAACAAAACAAATGGAGTGACTCCAAGAAGATGGATTCGTTTCTGCAATCCTCCTCTTAGTGCCATCATAACTAAGTGGA : 1900
Phola(7) : CTCTGGCCGGAAAAGTTCCAAAACAAAACAAATGGAGTGACTCCAAGAAGATGGATTCGTTTCTGCAATCCTCCTCTTAGTGCCATCATAACTAAGTGGA : 1900
Phola(8) : CTCTGGCCGGAAAAGTTCCAAAACAAAACAAATGGAGTGACTCCAAGAAGATGGATTCGTTTCTGCAATCCTCCTCTTAGTGCCATCATAACTAAGTGGA : 1900
Phola(9) : CTCTGGCCGGAAAAGTTCCAAAACAAAACAAATGGAGTGACTCCAAGAAGATGGATTCGTTTCTGCAATCCTCCTCTTAGTGCCATCATAACTAAGTGGA : 1900

                  *      1920          *      1940          *      1960          *      1980          *      2000
Phola-HR : CTGGTACAGAGGATTGGGTCCTGAAAACTGAAAAGTTGGCAGAATTGCAGAAGTTTGCTGATAATGAAGATCTTCAAAATGAGTGGAGGGAAGCAAAAAG : 2000
Phola(2) : CTGGTACAGAGGATTGGGTCCTGAAAACTGAAAAGTTGGCAGAATTGCAGAAGTTTGCTGATAATGAAGATCTTCAAAATGAGTGGAGGGAAGCAAAAAG : 2000
Phola(3) : CTGGTACAGAGGATTGGGTCCTGAAAACTGAAAAGTTGGCAGAATTGCAGAAGTTTGCTGATAATGAAGATCTTCAAAATGAGTGGAGGGAAGCAAAAAG : 2000
Phola(4) : CTGGTACAGAGGATTGGGTCCTGAAAACTGAAAAGTTGGCAGAATTGCAGAAGTTTGCTGATAATGAAGATCTTCAAAATGAGTGGAGGGAAGCAAAAAG : 2000
Phola-HA : CTGGTACAGAGGATTGGGTCCTTAAAACTGAAAAGTTGTCAGAATTGCAGAAGTTTGCTGATAATGAAGATCTTCAAAATGAGTGGAGGGAAGCAAAAAG : 2000
Phola(6) : CTGGTACAGAGGATTGGGTCCTTAAAACTGAAAAGTTGTCAGAATTGCAGAAGTTTGCTGATAATGAAGATCTTCAAAATGAGTGGAGGGAAGCAAAAAG : 2000
Phola(7) : CTGGTACAGAGGATTGGGTCCTGAAAACTGAAAAGTTGGCAGAATTGCAGAAGTTTGCTGATAATGAAGATCTTCAAAATGAGTGGAGGGAAGCAAAAAG : 2000
Phola(8) : CTGGTACAGAGGATTGGGTCCTGAAAACTGAAAAGTTGGCAGAATTGCAGAAGTTTGCTGATAATGAAGATCTTCAAAATGAGTGGAGGGAAGCAAAAAG : 2000
Phola(9) : CTGGTACAGAGGATTGGGTCCTGAAAACTGAAAAGTTGGCAGAATTGCAGAAGTTTGCTGATAATGAAGATCTTCAAAATGAGTGGAGGGAAGCAAAAAG : 2000

                  *      2020          *      2040          *      2060          *      2080          *      2100
Phola-HR : GAGCAACAAGATTAAAGTTGTCTCCTTTCTCAAAGAAAAGACAGGGTATTCTGTTGTCCCAGATGCAATGTTTGATATTCAGGTAAAACGCATTCATGAG : 2100
Phola(2) : GAGCAACAAGATTAAAGTTGTCTCCTTTCTCAAAGAAAAGACAGGGTATTCTGTTGTCCCAGATGCAATGTTTGATATTCAGGTAAAACGCATTCATGAG : 2100
Phola(3) : GAGCAACAAGATTAAAGTTGTCTCCTTTCTCAAAGAAAAGACAGGGTATTCTGTTGTCCCAGATGCAATGTTTGATATTCAGGTAAAACGCATTCATGAG : 2100
Phola(4) : GAGCAACAAGATTAAAGTTGTCTCCTTTCTCAAAGAAAAGACAGGGTATTCTGTTGTCCCAGATGCAATGTTTGATATTCAGGTAAAACGCATTCATGAG : 2100
Phola-HA : GAGCAACAAGATTAAAGTTGTCTCCTTTCTCAAAGAAAAGACAGGGTATTCTGTTGTCCCAGATGCAATGTTTGATATTCAGGTAAAACGCATTCATGAG : 2100
Phola(6) : GAGCAACAAGATTAAAGTTGTCTCCTTTCTCAAAGAAAAGACAGGGTATTCTGTTGTCCCAGATGCAATGTTTGATATTCAGGTAAAACGCATTCATGAG : 2100
Phola(7) : GAGCAACAAGATTAAAGTTGTCTCCTTTCTCAAAGAAAAGACAGGGTATTCTGTTGTCCCAGATGCAATGTTTGATATTCAGGTAAAACGCATTCATGAG : 2100
Phola(8) : GAGCAACAAGATTAAAGTTGTCTCCTTTCTCAAAGAAAAGACAGGGTATTCTGTTGTCCCAGATGCAATGTTTGATATTCAGGTAAAACGCATTCATGAG : 2100
Phola(9) : GAGCAACAAGATTAAAGTTGTCTCCTTTCTCAAAGAAAAGACAGGGTATTCTGTTGTCCCAGATGCAATGTTTGATATTCAGGTAAAACGCATTCATGAG : 2100

                  *      2120          *      2140          *      2160          *      2180          *      2200
Phola-HR : TACAAGCGACAACTGTTAAATATCTTCGGCATCGTTTATCGGTATAAGAAGATGAAAGAAATGACAGCTGCAGAAAGAAAGACTAACTTCGTTCCTCGAG : 2200
Phola(2) : TACAAGCGACAACTGTTAAATATCTTCGGTATAAGAAGATGAAAGAAATGACAGCTGCAGAAAGAAAGACTAACTTCGTTCCTCGAG : 2200
Phola(3) : TACAAGCGACAACTGTTAAATATCTTCGGCATCGTTTATCGGTATAAGAAGATGAAAGAAATGACAGCTGCAGAAAGAAAGACTAACTTCGTTCCTCGAG : 2200
Phola(4) : TACAAGCGACAACTGTTAAATATCTTCGGCATCGTTTATCGGTATAAGAAGATGAAAGAAATGACAGCTGCAGAAAGAAAGACTAACTTCGTTCCTCGAG : 2200
Phola-HA : TACAAGCGACAACTGTTAAATATCTTCGGCATCGTTTATCGGTATAAGAAGATGAAAGAAATGACAGCTGCAGAAAGAAAGACTAACTTCGTTCCTCGAG : 2200
Phola(6) : TACAAGCGACAACTGTTAAATATCTTCGGCATCGTTTATCGGTATAAGAAGATGAAAGAAATGACAGCTGCAGAAAGAAAGACTAACTTCGTTCCTCGAG : 2200
Phola(7) : TACAAGCGACAACTGTTAAATATCTTCGGCATCGTTTATCGGTATAAGAAGATGAAAGAAATGACAGCTGCAGAAAGAAAGACTAACTTCGTTCCTCGAG : 2200
Phola(8) : TACAAGCGACAACTGTTAAATATCTTCGGCATCGTTTATCGGTATAAGAAGATGAAAGAAATGACAGCTGCAGAAAGAAAGACTAACTTCGTTCCTCGAG : 2200
Phola(9) : TACAAGCGACAACTGTTAAATATCTTCGGCATCGTTTATCGGTATAAGAAGATGAAAGAAATGACAGCTGCAGAAAGAAAGACTAACTTCGTTCCTCGAG : 2200

                  *      2220          *      2240          *      2260          *      2280          *      2300
Phola-HR : TATGCATATTTGGGGGAAAAGCTTTTGCCACATATGTGCAAGCCAAGAGGATTGTAAAATTTATCACAGATGTTGGTGCTACTATAAATCATGATCCAGA : 2300
Phola(2) : TATGCATATTTGGGGGAAAAGCTTTTGCCACATATGTGCAAGCCAAGAGGATTGTAAAATTTATCACAGATGTTGGTGCTACTATAAATCATGATCCAGA : 2300
Phola(3) : TATGCATATTTGGGGGAAAAGCTTTTGCCACATATGTGCAAGCCAAGAGGATTGTAAAATTTATCACAGATGTTGGTGCTACTATAAATCATGATCCAGA : 2300
Phola(4) : TATGCATATTTGGGGGAAAAGCTTTTGCCACATATGTGCAAGCCAAGAGGATTGTAAAATTTATCACAGATGTTGGTGCTACTATAAATCATGATCCAGA : 2300
Phola-HA : TATGCATATTTGGGGGAAAAGCTTTTGCCACATATGTGCAAGCCAAGAGGATTGTAAAATTTATCACAGATGTTGGTGCTACTATAAATCATGATCCAGA : 2300
Phola(6) : TATGCATATTTGGGGGAAAAGCTTTTGCCACATATGTGCAAGCCAAGAGGATTGTAAAATTTATCACAGATGTTGGTGCTACTATAAATCATGATCCAGA : 2300
Phola(7) : TATGCATATTTGGGGGAAAAGCTTTTGCCACATATGTGCAAGCCAAGAGGATTGTAAAATTTATCACAGATGTTGGTGCTACTATAAATCATGATCCAGA : 2300
Phola(8) : TATGCATATTTGGGGGAAAAGCTTTTGCCACATATGTGCAAGCCAAGAGGATTGTAAAATTTATCACAGATGTTGGTGCTACTATAAATCATGATCCAGA : 2300
Phola(9) : TATGCATATTTGGGGGAAAAGCTTTTGCCACATATGTGCAAGCCAAGAGGATTGTAAAATTTATCACAGATGTTGGTGCTACTATAAATCATGATCCAGA : 2300

                  *      2320          *      2340          *      2360          *      2380          *      2400
Phola-HR : AATCGGTGATCTGTTGAAGGTAGTCTTTGTGCCAGATTACAATGTCAGTGTTGCTGAATTGCTAATTCCTGCTAGCGATCTATCAGAACATATCAGTACG : 2400
Phola(2) : AATCGGTGATCTGTTGAAGGTAGTCTTTGTGCCAGATTACAATGTCAGTGTTGCTGAATTGCTAATTCCTGCTAGCGATCTATCAGAACATATCAGTACG : 2400
Phola(3) : AATCGGTGATCTGTTGAAGGTAGTCTTTGTGCCAGATTACAATGTCAGTGTTGCTGAATTGCTAATTCCTGCTAGCGATCTATCAGAACATATCAGTACG : 2400
Phola(4) : AATCGGTGATCTGTTGAAGGTAGTCTTTGTGCCAGATTACAATGTCAGTGTTGCTGAATTGCTAATTCCTGCTAGCGATCTATCAGAACATATCAGTACG : 2400
Phola-HA : AATCGGTGATCTGTTGAAGGTAGTCTTTGTGCCAGATTACAATGTCAGTGTTGCTGAATTGCTAATTCCTGCTAGCGATCTATCAGAACATATCAGTACG : 2400
Phola(6) : AATCGGTGATCTGTTGAAGGTAGTCTTTGTGCCAGATTACAATGTCAGTGTTGCTGAATTGCTAATTCCTGCTAGCGATCTATCAGAACATATCAGTACG : 2400
Phola(7) : AATCGGTGATCTGTTGAAGGTAGTCTTTGTGCCAGATTACAATGTCAGTGTTGCTGAATTGCTAATTCCTGCTAGCGATCTATCAGAACATATCAGTACG : 2400
Phola(8) : AATCGGTGATCTGTTGAAGGTAGTCTTTGTGCCAGATTACAATGTCAGTGTTGCTGAATTGCTAATTCCTGCTAGCGATCTATCAGAACATATCAGTACG : 2400
Phola(9) : AATCGGTGATCTGTTGAAGGTAGTCTTTGTGCCAGATTACAATGTCAGTGTTGCTGAATTGCTAATTCCTGCTAGCGATCTATCAGAACATATCAGTACG : 2400

                  *      2420          *      2440          *      2460          *      2480          *      2500
Phola-HR : GCTGGAATGGAGGCCAGTGGAACCAGTAATATGAAGTTTGCAATGAATGGTTGTATCCAAATTGGTACATTGGATGGCGCTAATGTTGAAATAAGGGAAG : 2500
Phola(2) : GCTGGAATGGAGGCCAGTGGAACCAGTAATATGAAGTTTGCAATGAATGGTTGTATCCAAATTGGTACATTGGATGGCGCTAATGTTGAAATAAGGGAAG : 2500
Phola(3) : GCTGGAATGGAGGCCAGTGGAACCAGTAATATGAAGTTTGCAATGAATGGTTGTATCCAAATTGGTACATTGGATGGCGCTAATGTTGAAATAAGGGAAG : 2500
Phola(4) : GCTGGAATGGAGGCCAGTGGAACCAGTAATATGAAGTTTGCAATGAATGGTTGTATCCAAATTGGTACATTGGATGGCGCTAATGTTGAAATAAGGGAAG : 2500
Phola-HA : GCTGGAATGGAGGCCAGTGGAACCAGTAATATGAAGTTTGCAATGAATGGTTGTATCCAAATTGGTACATTGGATGGCGCTAATGTTGAAATAAGGGAAG : 2500
Phola(6) : GCTGGAATGGAGGCCAGTGGAACCAGTAATATGAAGTTTGCAATGAATGGTTGTATCCAAATTGGTACATTGGATGGCGCTAATGTTGAAATAAGGGAAG : 2500
Phola(7) : GCTGGAATGGAGGCCAGTGGAACCAGTAATATGAAGTTTGCAATGAATGGTTGTATCCAAATTGGTACATTGGATGGCGCTAATGTTGAAATAAGGGAAG : 2500
Phola(8) : GCTGGAATGGAGGCCAGTGGAACCAGTAATATGAAGTTTGCAATGAATGGTTGTATCCAAATTGGTACATTGGATGGCGCTAATGTTGAAATAAGGGAAG : 2500
Phola(9) : GCTGGAATGGAGGCCAGTGGAACCAGTAATATGAAGTTTGCAATGAATGGTTGTATCCAAATTGGTACATTGGATGGCGCTAATGTTGAAATAAGGGAAG : 2500

                  *      2520          *      2540          *      2560          *      2580          *      2600
Phola-HR : AGGTTGGAGAAGAAAACTTCTTTCTCTTTGGTGCTCAAGCTCATGAAATTGCAGGGCTTAGAAAAGAAAGAGCTGACGGAAAGTTTGTACCTGATGAACG : 2600
Phola(2) : AGGTTGGAGAAGAAAACTTCTTTCTCTTTGGTGCTCAAGCTCATGAAATTGCAGGGCTTAGAAAAGAAAGAGCTGACGGAAAGTTTGTACCTGATGAACG : 2600
Phola(3) : AGGTTGGAGAAGAAAACTTCTTTCTCTTTGGTGCTCAAGCTCATGAAATTGCAGGGCTTAGAAAAGAAAGAGCTGACGGAAAGTTTGTACCTGATGAACG : 2600
Phola(4) : AGGTTGGAGAAGAAAACTTCTTTCTCTTTGGTGCTCAAGCTCATGAAATTGCAGGGCTTAGAAAAGAAAGAGCTGACGGAAAGTTTGTACCTGATGAACG : 2600
Phola-HA : AGGTTGGAGAAGAAAACTTCTTTCTCTTTGGTGCTCAAGCTCATGAAATTGCAGGGCTTAGAAAAGAAAGAGCTGACGGAAAGTTTGTACCTGATGAACG : 2600
Phola(6) : AGGTTGGAGAAGAAAACTTCTTTCTCTTTGGTGCTCAAGCTCATGAAATTGCAGGGCTTAGAAAAGAAAGAGCTGACGGAAAGTTTGTACCTGATGAACG : 2600
Phola(7) : AGGTTGGAGAAGAAAACTTCTTTCTCTTTGGTGCTCAAGCTCATGAAATTGCAGGGCTTAGAAAAGAAAGAGCTGACAGAAAGTTTGTACCTGATGAACG : 2600
Phola(8) : AGGTTGGAGAAGAAAACTTCTTTCTCTTTGGTGCTCAAGCTCATGAAATTGCAGGGCTTAGAAAAGAAAGAGCTGACAGAAAGTTTGTACCTGATGAACG : 2600
Phola(9) : AGGTTGGAGAAGAAAACTTCTTTCTCTTTGGTGCTCAAGCTCATGAAATTGCAGGGCTTAGAAAAGAAAGAGCTGACAGAAAGTTTGTACCTGATGAACG : 2600
```

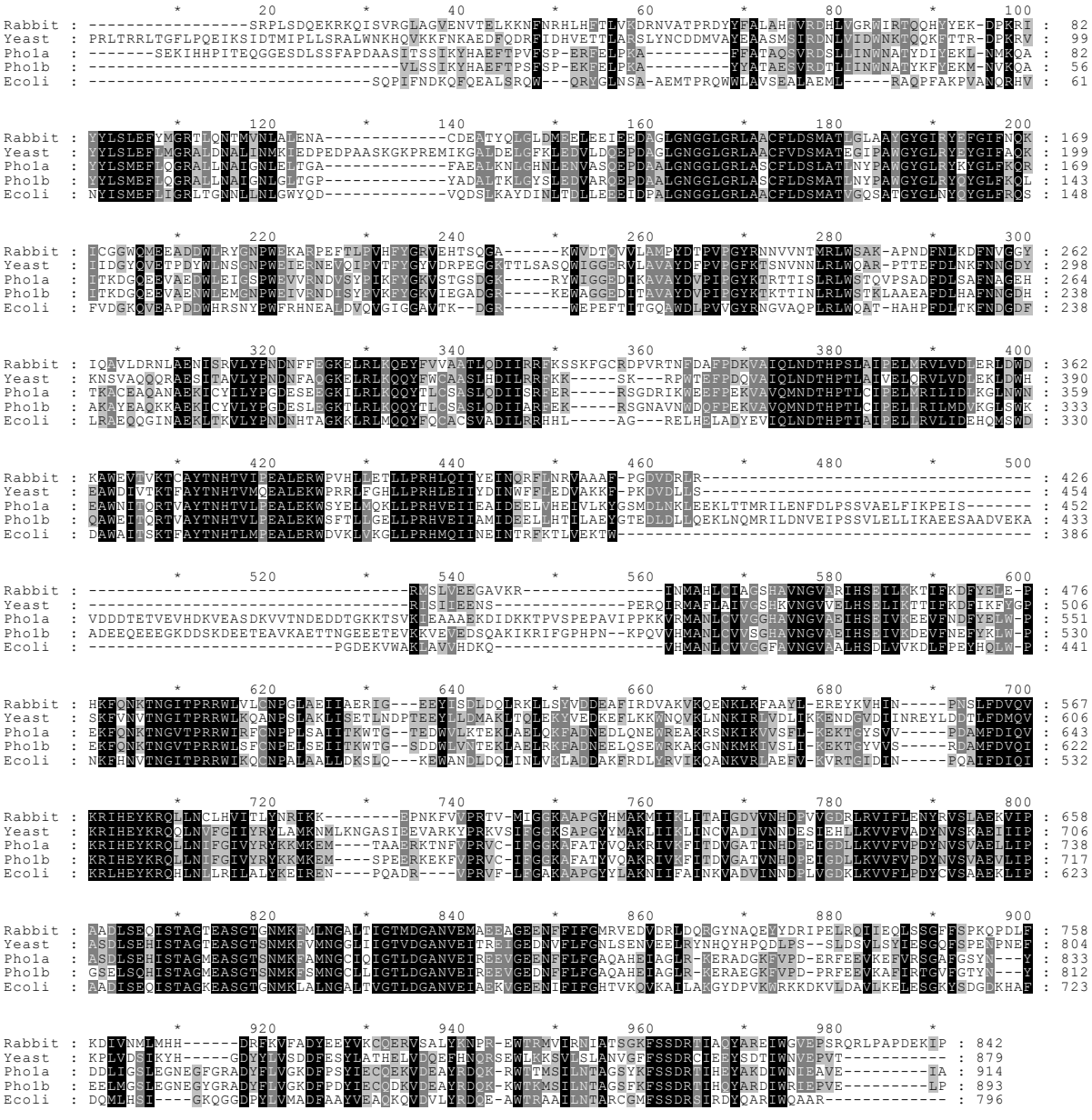
```
                  *      2620         *      2640         *      2660         *      2680         *      2700
Phola-HR : TTTTGAAGAGGTGAAGGAATTTGTTAGAAGCGGTGCTTTTGGCTCTTATAACTATGATGACCTAATTGGATCGTTGGAAGGAAATGAAGGTTTTGGCCGT : 2700
Phola(2) : TTTTGAAGAGGTGAAGGAATTTGTTAGAAGCGGTGCTTTTGGCTCTTATAACTATGATGACCTAATTGGATCGTTGGAAGGAAATGAAGGTTTTGGCCGT : 2700
Phola(3) : TTTTGAAGAGGTGAAGGAATTTGTTAGAAGCGGTGCTTTTGGCTCTTATAACTATGATGACCTAATTGGATCGTTGGAAGGAAATGAAGGTTTTGGCCGT : 2700
Phola(4) : TTTTGAAGAGGTGAAGGAATTTGTTAGAAGCGGTGCTTTTGGCTCTTATAACTATGATGACCTAATTGGATCGTTGGAAGGAAATGAAGGTTTTGGCCGT : 2700
Phola-HA : TTTTGAAGAGGTGAAGGAATTTGTTAGAAGCGGTGCTTTTGGCTCTTATAACTATGATGACCTAATTGGATCGTTGGAAGGAAATGAAGGTTTTGGCCGT : 2700
Phola(6) : TTTTGAAGAGGTGAAGGAATTTGTTAGAAGCGGTGCTTTTGGCTCTTATAACTATGATGACCTAATTGGATCGTTGGAAGGAAATGAAGGTTTTGGCCGT : 2700
Phola(7) : TTTTGAAGAGGTGAAGGAATTTGTTAGAAGCGGTGCTTTTGGCTCTTATAACTATGATGACCTAATTGGATCGTTGGAAGGAAATGAAGGTTTTGGCCGT : 2700
Phola(8) : TTTTGAAGAGGTGAAGGAATTTGTTAGAAGCGGTGCTTTTGGCTCTTATAACTATGATGACCTAATTGGATCGTTGGAAGGAAATGAAGGTTTTGGCCGT : 2700
Phola(9) : TTTTGAAGAGGTGAAAGAATTTGTTAGAAGCGGTGCTTTTGGCTCTTATAACTATGATGACCTAATTGGATCGTTGGAAGGAAATGAAGGTTTTGGCCGT : 2700

                  *      2720         *      2740         *      2760         *      2780         *      2800
Phola-HR : GCTGACTATTTCCTTGTGGGCAAGGACTTCCCCAGTTACATAGAATGCCAAGAGAAAGTTGATGAGGCATATCGCGACCAGAAAAGGTGGACAACGATGT : 2800
Phola(2) : GCTGACTATTTCCTTGTGGGCAAGGACTTCCCCAGTTACATAGAATGCCAAGAGAAAGTTGATGAGGCATATCGCGACCAGAAAAGGTGGACAAAGATGT : 2800
Phola(3) : GCTGACTATTTCCTTGTGGGCAAGGACTTCCCCAGTTACATAGAATGCCAAGAGAAAGTTGATGAGGCATATCGCAACCAGAAAAGGTGGACAAAGATGT : 2800
Phola(4) : GCTGACTATTTCCTTGTGGGCAAGGACTTCCCCAGTTACATAGAATGCCAAGAGAAAGTTGATGAGGCATATCGCGACCAGAAAAGGTGGACAACGATGT : 2800
Phola-HA : GCTGACTATTTCCTTGTGGGCAAGGACTTCCCCAGTTACATAGAATGCCAAGAGAAAGTTGATGAGGCATATCGCAACCAGAAAAGGTGGACAAAGATGT : 2800
Phola(6) : GCTGACTATTTCCTTGTGGGCAAGGACTTCCCCAGTTACATAGAATGCCAAGAGAAAGTTGATGAGGCATATCGCGACCAGAAAAGGTGGACAAAGATGT : 2800
Phola(7) : GCTGACTATTTCCTTGTGGGCAAGGACTTCCCCAGTTACATAGAATGCCAAGAAAAAGTTGATGAGGCATATCGCGACCAGAAAAGGTGGACAAAGATGT : 2800
Phola(8) : GCTGACTATTTCCTTGTGGGCAAGGACTTCCCCAGTTACATAGAATGCCAAGAAAAAGTTGATGAGGCATATCGCGACCAGAAAAGGTGGACAAAGATGT : 2800
Phola(9) : GCTGACTATTTCCTTGTGGGCAAGGACTTCCCCAGTTACATAGAATGCCAAGAGAAAGTTGATGTGGCATATCGCGACCAGAAAAGGTGGACAAAGATGT : 2800

                  *      2820         *      2840         *      2860         *      2880         *      2900
Phola-HR : CAATCTTGAATACAGCGGGATCGTACAAGTTCAGCAGTGACAGAACAATCCATGAATATGCCAAAGACATTTGGAACATTGAAGCTGTGGAAATAGCATA : 2900
Phola(2) : CAATCTTGAATACAGCGGGATCCTACAAGTTCAGCAGTGACAGAACAATCCATGAATATGCCAAAGACATTTGGAACATTGAAGCTGTGGAAATAGCATA : 2900
Phola(3) : CAATCTTGAATACAGCGGGATCCTACAAGTTCAGCAGTGACAGAACAATCCATGAATATGCCAAAGACATTTGGAACATTGAAGCTGTGGAAATAGCATA : 2900
Phola(4) : CAATCTTGAATACAGCGGGATCGTACAAGTTCAGCAGTGACAGAACAATCCATGAATATGCCAAAGACATTTGGAACATTGAAGCTGTGGAAATAGCATA : 2900
Phola-HA : CAATCTTGAATACAGCGGGATCCTACAAGTTCAGCAGTGACAGAACAATCCATGAATATGCCAAAGACATTTGGAACATTGAAGCTGTGGAAATAGCATA : 2900
Phola(6) : CAATCTTGAATACAGCGGGATCCTACAAGTTCAGCAGTGACAGAACAATCCATGAATATGCCAAAGACATTTGGAACATTGAAGCTGTGGAAATAGCATA : 2900
Phola(7) : CAATCTTGAATACAGCGGGATCCTACAAGTTCAGCAGTGACAGAACAATCCATGAATATGCCAAAGACATTTGGAACATTGAAGCTGTGGAAATAGCATA : 2900
Phola(8) : CAATCTTGAATACAGCGGGATCGTACAAGTTCAGCAGTGACAGAACAATCCATGAATATGCCAAAGACATTTGGAACATTGAAGCTGTGGAAATAGCATA : 2900
Phola(9) : CAATCTTGAATACAGCGGGATCGTACAAGTTCAGCAGTGACAGAACAATCCATGAATATGCCAAAGACATTTGGAACATTGAAGCTGTGGAAATAGCATA : 2900


Phola-HR : A : 2901
Phola(2) : A : 2901
Phola(3) : A : 2901
Phola(4) : A : 2901
Phola-HA : A : 2901
Phola(6) : A : 2901
Phola(7) : A : 2901
Phola(8) : A : 2901
Phola(9) : A : 2901
```

Appendix 2. Nucleotide sequence of the *Pho1a-H_A* allele

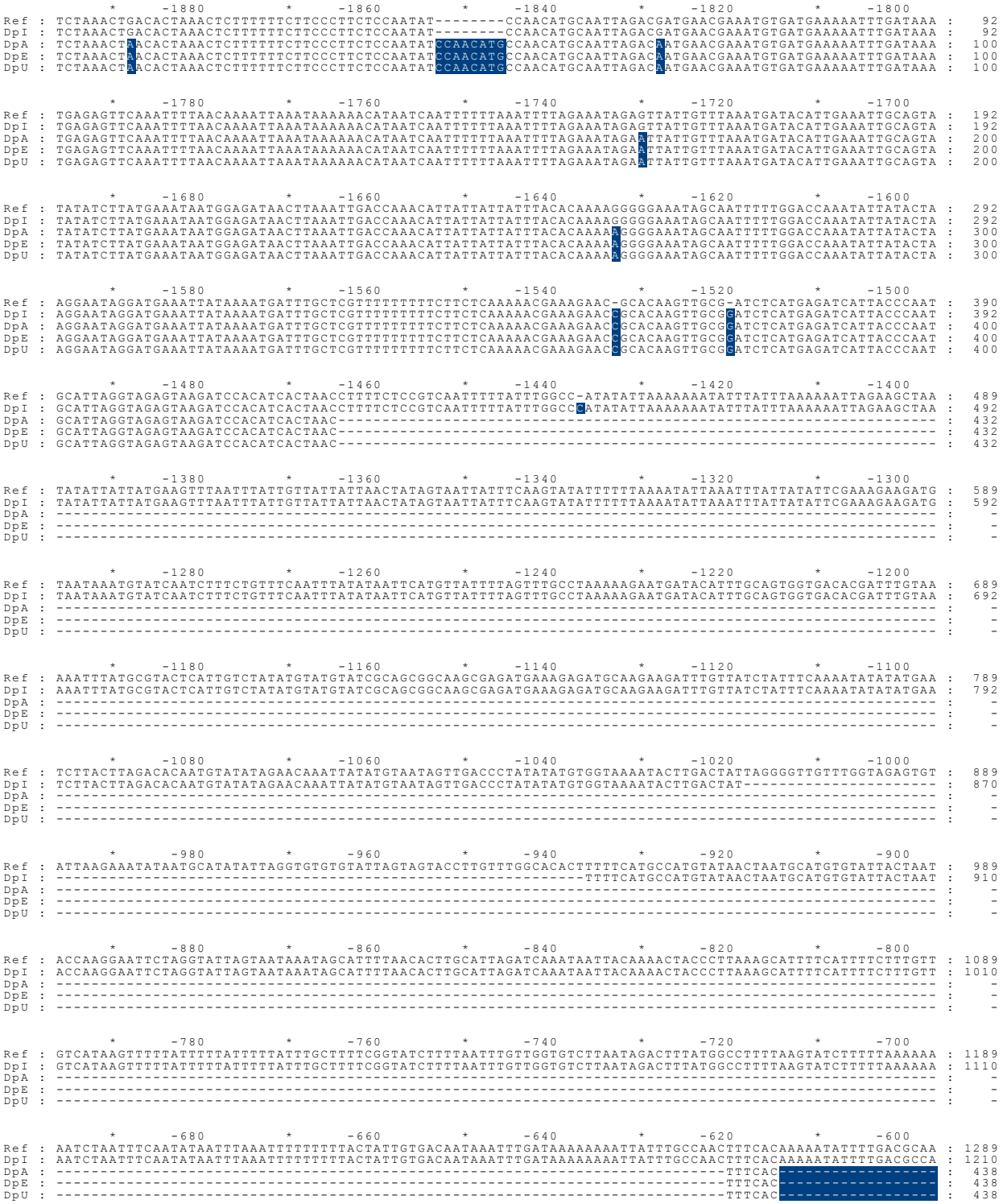
> *Pho1a-H_A*

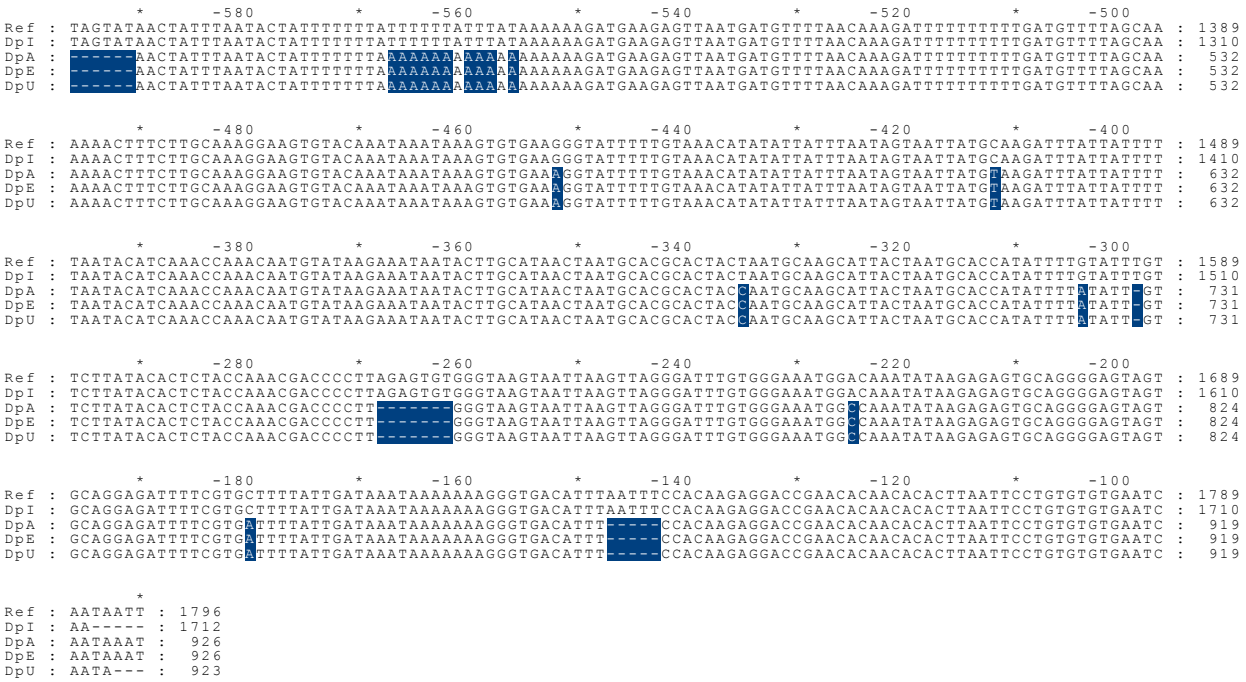
ATGGCGACTGCAAATGGAGCAtACTTGTTCAACCATTACAGCTCCAATTCCAGATTCATCCATTTCACTTCTA
GAAACACAAGCTCCAAATTGTTCCTTACCAAAACCTCCCATTTTCGGAGACCCAAACGCTGTTTCCATGTCAA
CAATACCTTGAGTGAGAAAATTCACCATCCCATTACTGAACAAGGTGGTGAGAGCGACCTGAGTTCTTTTGCT
CCTGATGCCGCATCTATTACCTCAAGTATCAAATACCATGCAGAATTCACACCTGTATTCTCTCCTGAAAGGT
TTGAGCTCCCTAAGGCATTCTTTGCAACAaCTCAAAGTGTTCGTGATTCGCTCCTTATTAATTGGAATGCTAC
GTATGATATTTATGAAAAGCTGAACATGAAGCAAGCGTACTATCTATCCATGGAATTTCTGCAGGGTAGAGCA
TTGTTAAATGCAATTGGTAATCTGGAGCTTACTGGTGCATTTGCGGAAGCTTTGAAAAACCTTGGCCACAATC
TAGAAAATGTGGCTTCTCAGGAgCCAGATGCTGCTCTTGGAAATGGGGGTTTGGGACGGCTTGCTTCCTGTTT
TCTGGACTCTTTGGCAACACTAAACTACCCAGCATGGGGCTATGGACTTAGGTACAAGTATGGTTTATTTAAG
CAACGGATTACAAAAGATGGTCAGGAGGAGGTGGCTGAAGATTGGCTTGAAATTGGCAGTCCcTGGGAAGTTG
TGAGGAATGATGTTTCATATCCTATCAAATTCTATGGAAAAGTCTCTACAGGATCAGATGGAAAGAGGTATTG
GATTGGTGGAGAGGATATAAgGGCAGTTGCGTATGATGTTCCCATACCAGGGTATAAGACCAGAACCACAATC
AGCCTTCGACTGTGGTCTACACAGGTTCCATCAGCGGATTTTGATTTATCTGCTTTCAATGCTGGAGAGCACA
CCAAAGCATGTGAAGCCCAAGCAAACGCTGAGAAGATATGTTACATACTCTACCCTGGGGATGAATCAGAGGA
GGGAAAGATCCTTCGGTTGAAGCAACAATATACCTTATGCTCGGCTTCTCTCCAAGATATTATTTCTCGATTT
GAGAGGAGATCAGGTGATCGTATTAAGTGGGAAGAGTTTCCTGAAAAAGTTGCTGTGCAGATGAATGACACTC
ACCCTACACTTTGTATCCCTGAGCTGATGAGAATATTGATAGATCTGAAGGGCTTGAATTGGAATGAAGCTTG
GAATATTACTCAAAGAACTGTGGCCTACACAAACCATACTGTTTTGCCTGAGGCACTGGAGAAATGGAGTTAT
GAATTGATGCAGAAACTCCTTCCCAGACATGTCGAAATCATTGAGGCGATTGACGAGGAGCTGGTACATGAAA
TTGTATTAAAATATGGTTCAATGGATCTGAACAAATTGGAGGAAAAGTTGACTACAATGAGAATCTTAGAAAA
TTTTGATCTTCCCAGTTCTGTTGCTGAATTATTTATTAAGCCTGAAATCTCAGTTGATGATGATACTGAAACA
GTAGAAGTCCATGACAAAGTTGAAGCTTCCGATAAAGTTGTGACTAATGATGAAGATGACACTGGTAAGAAAA
CTAGTGTGAAGATAGAAGCAGCTGCAGAAAAAGACATTGACAAGAAAACTCCCGTGAGTCCGGAACCAGCTGT
TATACCACCTAAGAAGGTACGCATGGCCAACTTGTGTGTTGTGGGCGGCCATGCTGTTAATGGAGTTGCTGAG
ATCCATAGTGAAATTGTGAAGGAGGAGGTTTTCAATGACTTCTATGAGCTCTGGCCGGAAAAGTTCCAAAACA
AAACAAATGGAGTGACTCCAAGAAGATGGATTCGTTTCTGCAATCCTCCTCTTAGTGCCATCATAACTAAGTG
GACTGGTACAGAGGATTGGGTCCTtAAAACTGAAAAGTTGtCAGAATTGCAGAAGTTTGCTGATAATGAAGAT
CTTCAAAATGAGTGGAGGGAAGCAAAAAGGAGCAACAAGATTAAAGTTGTCTCCTTTCTCAAAGAAAAGACAG
GGTATTCTGTTGTCCCAGATGCAATGTTTGATATTCAGGTAAAACGCATTCATGAGTACAAGCGACAACTGTT
AAATATCTTCGGCATCGTTTATCGGTATAAGAAGATGAAAGAAATGACAGCTGCAGAAAGAAAGACTAACTTC
GTTCCTCGAGTATGCATATTTGGGGGAAAAGCTTTTGCCACATATGTGCAAGCCAAGAGGATTGTAAAATTTA
TCACAGATGTTGGTGCTACTATAAATCATGATCCAGAAATCGGTGATCTGTTGAAGGTAGTCTTTGTGCCAGA
TTACAATGTCAGTGTTGCTGAATTGCTAATTCCTGCTAGCGATCTATCAGAACATATCAGTACGGCTGGAATG
GAGGCCAGTGGAACCAGTAATATGAAGTTTGCAATGAATGGTTGTATCCAAATTGGTACATTGGATGGCGCTA
ATGTTGAAATAAGGGAAGAGGTTGGAGAAGAAAACTTCTTTCTCTTTGGTGCTCAAGCTCATGAAATTGCAGG
GCTTAGAAAAGAAAGAGCTGACGGAAAGTTTGTACCTGATGAACGTTTTGAAGAGGTGAAGGAATTTGTTAGA
AGCGGTGCTTTTGGCTCTTATAACTATGATGACCTAATTGGATCGTTGGAAGGAAATGAAGGTTTTGGCCGTG
CTGACTATTTCCTTGTGGGCAAGGACTTCCCCAGTTACATAGAATGCCAAGAGAAAGTTGATGAGGCATATCG
CaACCAGAAAAGGTGGACAAaGATGTCAATCTTGAATACAGCGGGATCcTACAAGTTCAGCAGTGACAGAACA
ATCCATGAATATGCCAAAGACATTTGGAACATTGAAGCTGTGGAAATAGCATAA

Appendix 3. Alignment of PHOs

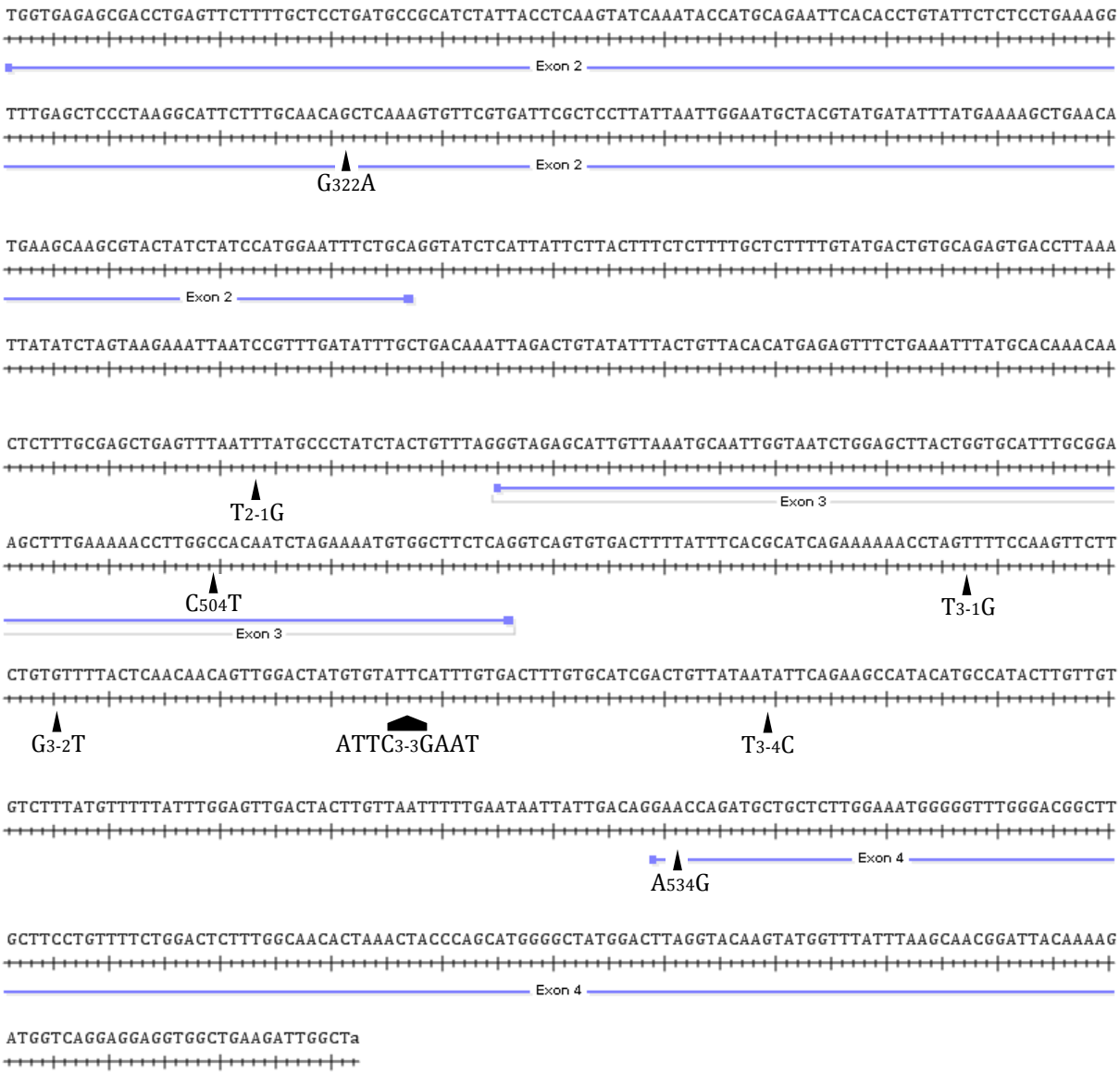


Appendix 4. Alignment of *Pho1a* promoter sequences. 1Kb fragment from cultivar Satina (DpI) and ~1Kb fragment from Diana (DpA), Theresa (DpE), and Saturna (DpU).




* -580 * -560 * -540 * -520 * -500
Ref : TAGTATAACTATTTAATACTATTTTTTTATTTTTTATTTATAAAAAAGATGAAGAGTTAATGATGTTTTAACAAAGATTTTTTTTTTGATGTTTTAGCAA : 1389
DpI : TAGTATAACTATTTAATACTATTTTTTTATTTTTTATTTATAAAAAAGATGAAGAGTTAATGATGTTTTAACAAAGATTTTTTTTTTGATGTTTTAGCAA : 1310
DpA : -------AACTATTTAATACTATTTTTTTTAAAAAAAAAAAAAAAAAAAGATGAAGAGTTAATGATGTTTTAACAAAGATTTTTTTTTTGATGTTTTAGCAA : 532
DpE : -------AACTATTTAATACTATTTTTTTTAAAAAAAAAAAAAAAAAAAGATGAAGAGTTAATGATGTTTTAACAAAGATTTTTTTTTTGATGTTTTAGCAA : 532
DpU : -------AACTATTTAATACTATTTTTTTTAAAAAAAAAAAAAAAAAAAGATGAAGAGTTAATGATGTTTTAACAAAGATTTTTTTTTTGATGTTTTAGCAA : 532
* -480 * -460 * -440 * -420 * -400
Ref : AAAACTTTCTTGCAAAGGAAGTGTACAAATAAATAAAGTGTGAAGGGTATTTTTGTAAACATATATTATTTAATAGTAATTATGCAAGATTTATTATTTTT : 1489
DpI : AAAACTTTCTTGCAAAGGAAGTGTACAAATAAATAAAGTGTGAAGGGTATTTTTGTAAACATATATTATTTAATAGTAATTATGCAAGATTTATTATTTTT : 1410
DpA : AAAACTTTCTTGCAAAGGAAGTGTACAAATAAATAAAGTGTGAAAGGTATTTTTGTAAACATATATTATTTAATAGTAATTATGTAAGATTTATTATTTTT : 632
DpE : AAAACTTTCTTGCAAAGGAAGTGTACAAATAAATAAAGTGTGAAAGGTATTTTTGTAAACATATATTATTTAATAGTAATTATGTAAGATTTATTATTTTT : 632
DpU : AAAACTTTCTTGCAAAGGAAGTGTACAAATAAATAAAGTGTGAAAGGTATTTTTGTAAACATATATTATTTAATAGTAATTATGTAAGATTTATTATTTTT : 632
* -380 * -360 * -340 * -320 * -300
Ref : TAATACATCAAACCAAACAATGTATAAGAAATAATACTTGCATAACTAATGCACGCACTACTAATGCAAGCATTACTAATGCACCATATTTTGTATTTGT : 1589
DpI : TAATACATCAAACCAAACAATGTATAAGAAATAATACTTGCATAACTAATGCACGCACTACTAATGCAAGCATTACTAATGCACCATATTTTGTATTTGT : 1510
DpA : TAATACATCAAACCAAACAATGTATAAGAAATAATACTTGCATAACTAATGCACGCACTACCAATGCAAGCATTACTAATGCACCATATTTTATATT-GT : 731
DpE : TAATACATCAAACCAAACAATGTATAAGAAATAATACTTGCATAACTAATGCACGCACTACCAATGCAAGCATTACTAATGCACCATATTTTATATT-GT : 731
DpU : TAATACATCAAACCAAACAATGTATAAGAAATAATACTTGCATAACTAATGCACGCACTACCAATGCAAGCATTACTAATGCACCATATTTTATATT-GT : 731
* -280 * -260 * -240 * -220 * -200
Ref : TCTTATACACTCTACCAAACGACCCCTTAGAGTGTGGGTAAGTAATTAAGTTAGGGATTTGTGGGAAATGGACAAATATAAGAGAGTGCAGGGGAGTAGT : 1689
DpI : TCTTATACACTCTACCAAACGACCCCTTAGAGTGTGGGTAAGTAATTAAGTTAGGGATTTGTGGGAAATGGACAAATATAAGAGAGTGCAGGGGAGTAGT : 1610
DpA : TCTTATACACTCTACCAAACGACCCCTT-------GGGTAAGTAATTAAGTTAGGGATTTGTGGGAAATGGCCAAATATAAGAGAGTGCAGGGGAGTAGT : 824
DpE : TCTTATACACTCTACCAAACGACCCCTT-------GGGTAAGTAATTAAGTTAGGGATTTGTGGGAAATGGCCAAATATAAGAGAGTGCAGGGGAGTAGT : 824
DpU : TCTTATACACTCTACCAAACGACCCCTT-------GGGTAAGTAATTAAGTTAGGGATTTGTGGGAAATGGCCAAATATAAGAGAGTGCAGGGGAGTAGT : 824
* -180 * -160 * -140 * -120 * -100
Ref : GCAGGAGATTTTCGTGCTTTTATTGATAAATAAAAAAAGGGTGACATTTAATTTCCACAAGAGGACCGAACACAACACACTTAATTCCTGTGTGTGAATC : 1789
DpI : GCAGGAGATTTTCGTGCTTTTATTGATAAATAAAAAAAGGGTGACATTTAATTTCCACAAGAGGACCGAACACAACACACTTAATTCCTGTGTGTGAATC : 1710
DpA : GCAGGAGATTTTCGTGATTTTATTGATAAATAAAAAAAGGGTGACATTT-----CCACAAGAGGACCGAACACAACACACTTAATTCCTGTGTGTGAATC : 919
DpE : GCAGGAGATTTTCGTGATTTTATTGATAAATAAAAAAAGGGTGACATTT-----CCACAAGAGGACCGAACACAACACACTTAATTCCTGTGTGTGAATC : 919
DpU : GCAGGAGATTTTCGTGATTTTATTGATAAATAAAAAAAGGGTGACATTT-----CCACAAGAGGACCGAACACAACACACTTAATTCCTGTGTGTGAATC : 919
*
Ref : AATAATT : 1796
DpI : AA----- : 1712
DpA : AATAAAT : 926
DpE : AATAAAT : 926
DpU : AATA--- : 923

Appendix 5. SNPs underlying the PHO1Ac and Stp23-8 SSCP polymorphisms (Table 11).



Acknowledgements

As this project is close to come to a happy end, I would like to express my sincere gratitude to the people who have helped me grow professionally and all those who made my life in this country an unforgettable experience.

My sincere gratitude to Christiane Gebhardt for opening the doors of her group and giving me the opportunity to work in the genetics of this challenging crop. The experiences I have obtained during my work in your group will help me support my future career. Muchas gracias to you.

To Maarten Koornneef reviewing the progress of my project, the discussions and for the funding during the last period of this project.

Claude Urbany for all his help during these 3+ years. Thanks a lot for taking the time to discuss my project, for answering my never ending questions, and for your help in the correction of this manuscript. I'm not your lawyer, but you would do a terrific group leader.

Matthias Fischer for introducing me to the mysterious (until then) world of proteomics. For your valuable corrections on this manuscript and constructive comments on how to improve scientific presentations, many thanks. Your positive attitude was always much appreciated.

To the always welcoming and helpful Birgit Walkemeier for all her support and for keeping up the spirit of her co-workers.

Li Li for providing the basis for this project and for all her collaboration.

Special thanks to Lena Schreiber who was always willing to help and with whom I shared successes and failures during this Ph.D. Thanks a lot for taking care of the survival of this Colombian in Germany.

Astrid Draffehn for sharing with me her experiences during the development of her (twin) project and for her discussions on how to deal with the "endless" potato allelic variation.

Sandra Maurer for instructing me on SSCP gels on the beginning of my PhD and Markus Kuckenberg for all his help with the protein work.

To Elske Schönhals for your cheerful attitude and predisposition to help. I truly believe you are a good addition to this group.

To the former scientific members of this group, Dr. Damaris Odeny, Dr. Agim Ballvora and Dr. Gabor Gyetvai to whom I owe all my gratitude not only because of their contributions to my project, but for all their personal support ever since my arrival to this country.

Gabriela Ronquillo was of great help in correcting my manuscript. Also, for the always interesting conversations on maize genetics, plant breeding and QTL mapping. Muchas gracias 'mijita'.

Special thanks to my dear *profe* Teresa Mosquera. This experience exceeded all my expectations. Thanks a lot for all your trust and your support always.

Last but not least, to Bodo Raatz, whom I can't thank enough for all his help during the writing of this manuscript, his corrections, his unlimited support, and of course for never losing his *cool* in the most stressful moments. Ahí estás tú, mirando siempre el lado brillante de la vida.

5.1 Erklärung

Die vorliegende Arbeit wurde am Max-Planck-Institut für Züchtungsforschung in Köln-Vogelsang durchgeführt.

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von PD Dr. Christiane Gebhardt betreut worden.