Acta Universitatis Agriculturae Sueciae

Doctoral Thesis No. 2024:48
Faculty of Landscape Architecture, Horticulture and Crop Production Science

Gene editing for tuber protein utilization in potato (*Solanum tuberosum*)

Martin Friberg

Doctoral Thesis No. 2024:48 • Gene editing for tuber protein utilization in potato (*Solanum tuberosum*) • Martin Friberg

Gene editing for tuber protein utilization in potato (*Solanum tuberosum*)

Martin Friberg
Faculty of Landscape Architecture, Horticulture, and Crop Production Science
Department of Plant Breeding
Alnarp

SLU
SWEDISH UNIVERSITY OF AGRICULTURAL SCIENCES

DOCTORAL THESIS
Alnarp 2024

Acta Universitatis Agriculturae Sueciae
2024:48

Cover: Watercolour painting of a potato tuber unravelling into editable DNA.
Painted by Cecilia Ålander

ISSN 1652-6880
ISBN (print version) 978-91-8046-036-1
ISBN (electronic version) 978-91-8046-037-8
https://doi.org/10.54612/ a.3u7o51prv7

Swedish University of Agricultural Sciences, Department of Plant breeding, Alnarp, Sweden

Print: SLU Service/Repro, Alnarp 2024

Gene editing for improved tuber protein utilization in potato (*Solanum tuberosum*)

Abstract

As awareness of climate change and the need for a more plant based diet spreads, so has interest in novel protein sources. One crop that has been gaining interest in the protein field is potato. Especially potatoes grown for the starch industry have been discussed since tuber proteins are accumulated in one of the side streams of starch processing.

In this thesis, I discuss the potential of using precision editing technologies such as base editing and prime editing to modify potato proteins to make this under-utilized resource available. Starting with patatin, a protein that is already a by-product from starch industry. It was shown that several patatin genes can be targeted at once using CRISPR/Cas9, meaning the family is a viable target. Furthermore, expression analysis showed that a majority of transcripts were represented by only a few genes, suggesting that total gene coverage is unnecessary.

A second protein of interest, Pho1a, was investigated. While one of the most abundant proteins in potato tubers, its function remains the source of much debate. Mutant lines deficient in Pho1a activity were produced and analysed for phenotypic effects, showing low amylose content and small starch granules. Since these are detrimental traits in starch production, *Pho1a*, could be an interesting target for enhancing edits in the future.

Finally, a gene editing toolbox was developed, utilizing adapter proteins to allow modular assembly of fusion constructs for CRISPR/Cas editing. It was shown that addition of the adapter domains did not affect Cas9 efficiency. While initial experiments using prime editing were unsuccessful, the underlying principles remain robust.

Keywords: Gene editing, CRISPR/Cas, synthetic biology, protein engineering, protein stability, *Solanum tuberosum*, starch potato cultivars, tuber protein, patatin

Gen-editering för ökat utnyttjande av knölprotein från potatis (Solanum tuberosum)

Abstract

I takt med att intresset för klimatförändringar ökar, så väljer allt fler att skifta till en mer växtbaserad diet. Detta har lett till förnyat intresse i växtbaserade proteinkällor. En gröda som blivit uppmärksammad i detta sammanhang är potatis. I synnerhet stärkelsepotatisen har varit av intresse eftersom proteinerna anrikas i en av sidoströmmarna från stärkelseindustrin.

I denna avhandling diskuteras vilken potential precisions editerings metoder så som bas-editering och prime-editering har för att bättre utnyttja proteinresurserna som finns i potatisknölar. En första studie utfördes på patatin, ett protein som kan utvinnas som biprodukt från stärkelseindustrin. Lyckad editering av flera patatin gener med CRISPR/Cas9 verifierades *in planta*, vilket bekräftar att patatin familjen kan editeras. En analys av genuttryck visade även att variationen på RNA nivå domineras av några få gener. Vilket betyder att endast ett urval av gener behöver modifieras för att ge en effekt i en majoritet av protein poolen.

Ytterligare en lovande gen, *Pho1a*, undersöktes. Trots att Pho1a är ett av de vanligaste proteinerna i knölar, så är dess funktion omdebatterad. Gen-editerade linjer utan Pho1a aktivitet analyserades, och en låg nivå av amylos samt små stärkelsegranuler observerades. Eftersom dessa förändringar är olämpliga för industrin, skulle precisions editering för ökad effektivitet av *Pho1a* kunna va av intresse för framtida studier.

Slutligen utvecklades ett modulärt system baserat på adapter protein för montering av fusionsprotein för CRISPR/Cas-editering. En inledande analys av Cas9 med en adapter modul (Cas9:Doc1) visade god effektivitet, vilket är lovande för framtida utveckling av systemet.

Keywords: Gen editering, CRISPR/Cas, syntetisk biologi, protein design, protein stabilitet, *Solanum tuberosum*, stärkelsepotatissorter

“It's like the great stories, Mr. Frodo, the ones that really mattered. Full of darkness and danger they were, and sometimes you didn't want to know the end because how could the end be happy? How could the world go back to the way it was when so much bad has happened? But in the end, it's only a passing thing this shadow, even darkness must pass. A new day will come, and when the sun shines, it'll shine out the clearer. I know now folks in those stories had lots of chances of turning back, only they didn't. They kept going because they were holding on to something. That there's some good in this world, Mr. Frodo, and it's worth fighting for.”

— Samwise Gamgee, The Lord of the Rings, The Two Towers

Contents

List of publications 9

List of figures 11

Abbreviations 13

1. Introduction 15

2. Background 17
   - 2.1 Potato for starch industry 17
   - 2.2 Potato tuber proteins 18
     - 2.2.1 Patatin 18
     - 2.2.2 Protease inhibitors 20
     - 2.2.3 Plastidial starch phosphorylase 20
   - 2.3 Protein engineering 21
     - 2.3.1 Protein stability 21
     - 2.3.2 Adapter proteins 22
   - 2.4 Gene editing 22
     - A brief history of gene editing 22
     - 2.4.1 CRISPR/Cas 24
     - 2.4.2 Base editing 26
     - 2.4.3 Prime editing 26
     - 2.4.4 CRISPR/Cas application in plants 27

3. Aims and objectives 29

4. Research outcomes 31
   - 4.1 Patatin expression patterns in Kuras and Desirée 31
   - 4.2 Patatin mutation study 32
   - 4.3 Patatin engineering 34
   - 4.4 Systems for patatin expression and evaluation 37

4.5 Modular CRISPR/Cas system using adapter proteins 39
4.6 Alternative tuber protein targets 41

5. Conclusion 43

References 45

Popular science summary 59

Populärvetenskaplig sammanfattning 61

Acknowledgements 63

List of publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I. Modifying potato tuber storage protein Patatin for improved thermal stability and functionality, Martin Friberg, Shrikant Sharma, Folke Sitbon, Mariette Andersson and Per Hofvander, (Manuscript)

II. A modular genome editing system for increased CRISPR/Cas RNP functionality, Martin Friberg, Matías González, Paul Vogel, Niklas Olsson, Folke Sitbon, Mariette Andersson and Per Hofvander, (Manuscript)

III. *Pho1a* (plastid starch phosphorylase) is duplicated and essential for normal starch granule phenotype in tubers of *Solanum tuberosum* L., Shrikant Sharma, Martin Friberg, Paul Vogel, Helle Turesson, Niklas Olsson, Mariette Andersson and Per Hofvander, (Published, doi: 10.3389/fpls.2023.1220973)

The contribution of Martin Friberg to the papers included in this thesis was as follows:

I. Helped plan the study, analysed the protein structure, designed, and tested protein variants. Conducted protoplast isolation, transfection, and analysis of the patatin gene editing study.

II. Helped plan the study, lead the design of the constructs, took part in initial expression, purification, and interaction studies. Conducted protoplasts isolation, transfection, and analysis of these experiments

III. Performed Zymogram experiments and helped in the analysis of these experiments. Measured apparent Amylose content of purified starch.

List of figures

Figure 1: A schematic representation of industrial scale extraction of starch from potato. 17

Figure 2: Crystal structure of Pat17 from two angles, a) from the side, b) from beneath. The core of the protein is formed by six beta-strand that are surrounded by seven alpha helices. Beta sheets are coloured orange, alpha helices blue and unstructured regions are green. 19

Figure 3: A schematic representation of Cas9 (light blue) bound to DNA (grey) through hybridisation with the sgRNA (dark blue). The PAM site and the two nuclease domains (RuvC and HNH) are labelled. Adapted from J.J. Froehlich, original available from Wikimedia under a CC BY-SA 4.0 license. 25

Figure 4: Expression levels of patatin genes in two potato cultivars, Desirée and Kuras. 32

Figure 5: Characterisation of patatin lines with induced indels. a) High resolution fragment analysis of region flanking the target site shown in green, and size marker in orange. b) Distribution of patatin alleles in Desirée as characterised by amplicon sequencing. 33

Figure 6: Differences between pat17 from *S. cardiophyllum* and pat2-k2 from *S. tuberosum*. a) Sequence alignment with the differences highlighted. Red lines represent missmatches, white lines represent gaps. b) The predicted structure of pat2-k2 (light blue) superimposed on the crystal structure of pat17 (PDB: 1oxw, dark blue). The unfolding nuclei have been colored green

for pat2-k2 and orange for pat17, and the corresponding amino acid sidechains have been visualised as sticks. 35

Figure 7: Designed patatin variants, with substituted amino acids highlighted. 36

Figure 8: Calculated melting temperature of patatin variants as measured by DSF. Each bar represents the mean of four replicates and error bars are standard deviation. 37

Figure 9: Extraction tests for recombinant expression of patatin, extracts and fractions from IMAC purification. L is a protein size marker, Raw is clarified extract, FT is Flow through, W is Wash and E is Elute. Patatin is expected at 42kDa. a) Agrobacterium infiltration of *Nicotiana benthamiana*. b) Expression in *Saccharomyces cerviceae.* 38

Figure 10: Constructs assembled for the adapter protein toolbox. pTB1XX constructs are based on the pMJ806 backbone, while pTB2XX constructs are based on pET30-R5-GFP. The two fusion partners are connected to either the Cip7, or Doc1 domain, through a GGSG linker. All constructs carry a 6*His tag for purification, and are expressed from the T7 promoter, which is IPTG inducible. 39

Figure 11: Editing efficiencies as estimated by amplicon sequencing. a) Application of Cas9:Doc1 targeting *Pho1* as compared to a positive control (PC) consisting of commercially available Cas9. b) The efficiency of a toolbox prime editor based on Cip7:RTdH, compared to a similar system expressed transiently from a plasmid..... 40

Figure 12: Activity assay of starch phosphorylase for a selection of *Pho1a* knock-out lines (SPD), and wt (Desirée). ◁ marks Pho1a activity, ◀ marks Pho2 activity. a) On PAGE cast with 1.2% glycogen. b) On commercial 10% PAGE. 42

Abbreviations

Cas	CRISPR associated (protein)
CNA	Constraint network analysis
CRISPR	Clustered regularly interspaced palindromic repeats
DSB	Double stranded break
GFP	Green fluorescent protein
HDR	Homology directed repair
IMAC	Immobilised metal-ion affinity chromatography
LAH	Lipid acyl hydrolase
MW	Molecular weight
NHEJ	Non-homologous end joining
PAM	Protospacer adjacent motif
PEG	Poly ethylene glycol
PI	Protease inhibitor
PFJ	Potato fruit juice
RFU	Relative fluorescence units
RNP	Ribonucleoprotein
sgRNA	Single guide RNA

1. Introduction

During the last few years, we have seen an upswing in the general interest for plant-based proteins as more consumers are trying to decrease their environmental footprint by limiting meat consumption. This movement is sometimes called the protein shift, and it has led to the development of several new food products, such as textured pea-protein, and has increased the interest in protein crop species such as lupines (*Lupinus sp.*) (Lucas et al. 2015) and faba bean (*Vicia faba*) (Karkanis et al. 2018). These species, could act as a replacement for soybean in countries that currently relies heavily on soybean import.

With the acreage of cultivated arable land seemingly reaching its limits (Springmann et al. 2018), there has also been research directed at how to better utilize side-stream products, such as leafy green biomass (Møller et al. 2021), and the seedcake from oil crops (Singh et al. 2022). Something that is still in high demand, are plant products with good emulsifying properties. One group of proteins that have been suggested for this purpose are the patatins, a group of storage proteins found in potato (*Solanum tuberosum*) tubers (Ralet & Guéguen 2001). Potatoes are already being processed for starch extraction, during which patatin ends up in the wastewater side stream, making this an interesting by-product to isolate from the process stream. While stabilising mutations for patatin would be hard to breed for by traditional means, biotechnology methods are becoming available that allow for precise edits of genes.

With the rise of CRISPR/Cas technology, a slew of crops with novel traits have been developed. This includes traits such as resistance to disease (Wang et al. 2014; Tashkandi et al. 2018), or improved nutritional quality (Li et al. 2018; Do et al. 2019). The technology has even been applied to starch potatoes to generate high amylose lines, a trait that could be interesting for

the manufacture of resistant starch and bioplastics (Zhao et al. 2021b). While these studies are mostly done using gene-knockouts, the CRISPR/Cas-toolbox is ever evolving and methods for precision editing, such as prime editing and base editing, are already being applied in crops.

The ability to easily introduce targeted point mutations in plant genomes allow for a range of interesting developments where knock-outs will not suffice, and is a powerful new tool within the field of synthetic biology.

2. Background

2.1 Potato for starch industry

While a popular food crop for several hundred years now, potatoes are also grown for non-alimentary use in starch industry. The process of extracting starch is summarised in Fig. 1. First, the tubers are washed and grated into pulp. The pulp is centrifuged to remove the potato fruit juice (PFJ). The starch is then separated from the potato fibres, usually using a sieve, and washed. Finally, the starch is refined and dried before being packaged (Grommers & van der Krogt 2009).

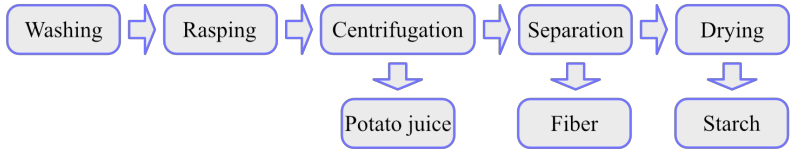

Figure 1: A schematic representation of industrial scale extraction of starch from potato.

Potato tubers contain only about 2% protein when looking at fresh weight, which is the most common way of consumption. However, the high yield of potato tubers means that the protein amount per hectare grown is considerable, around 400-600 kg per hectare[1]. This can be compared to the protein yield of soybean which has been estimated to be 700-1500 kg per hectare (Karges et al. 2022).

According to Starch Europe, an average of 8 million tonnes of potato tubers are processed by the starch industry in Europe each year. Using the

[1] Calculated using 2% protein by fresh weight and 20-30 tonnes yield per hectare as reported by FAOstat, 2022 and EUstat, 2022 respectively.

previously mentioned 2% of fresh weight protein, this corresponds to 160 000 tonnes of protein ending up in the side stream. Typically, the PFJ is treated with a combination of acid followed by heat, which coagulates the proteins. This allows for the rest of the liquid to be decanted. The protein mixture is then dried to a powder and sold primarily as animal feed since the high salt content and bitter taste makes it unsuitable for human consumption (Løkra & Strætkvern 2009).

The bitter taste is caused by glycoalkaloids, toxic compounds produced by many members of the nightshade family to defend against predators. In potato the two main compounds produced are α-solanine and α-chaconine, which are found primarily in leafs, and shoots, but can also accumulate in tubers as a result of different stresses such as light or wounding (Friedman et al. 1997). The fact that glycoalkaloids are also enriched in PFJ is currently the major hurdle to overcome for the production of food-grade potato protein.

2.2 Potato tuber proteins

Potato tuber proteins are usually divided into three major categories, protease inhibitors, patatin, and other proteins. The protease inhibitors represent about 30-50% of total protein, with a large diversity in both sizes and activity (Pouvreau et al. 2001). Patatin forms the second major group, making up around 25-45% of total protein (Racusen & Foote 1980; Barta et al. 2012a). The final 10-15% is made up of all other proteins, mainly kinases and enzymes involved in starch synthesis (Man et al. 1997; Pouvreau et al. 2001).

2.2.1 Patatin

The patatins are a group of glycoproteins found in several members of the *Solanaceae* family. In potato, the patatin genes are generally divided into two groups, denoted class-I, and class-II. The class-II patatins are distinguished by a 22nt up-stream region and are expressed at a low-level throughout the entire plant. By contrast, the class-I patatins lack the previously mentioned up-stream region, are expressed only in tubers, and at a much higher level than the class-II patatins (Pikaard et al. 1987; Prat et al. 1990).

The size of potato patatins vary between 40-42 kDa, and have a pI in the range of 4.8-5.2. The structure is a highly organised modified α/β hydrolase fold (Fig. 1), (Rydel et al. 2003; Wijeyesakere et al. 2014), with an even

charge distribution throughout the entire chain. In its native state, patatin forms dimers of around 80kDa through hydrophobic surface interactions (Alting et al. 2011).

The patatins are thought to function mainly as storage protein due to the high accumulation in tubers, and their localisation to the vacuole, but they also exhibit lipid acyl hydrolase (LAH) activity. This activity relies on a catalytic dyad comprised of a serine and an aspartic acid, to hydrolyse a variety of mono- and di-acylated glycerols as well as certain phospholipids (Galliard 1971; Hirschberg et al. 2001). This activity is thought to be involved in plant defence, either directly, as patatin has been shown to inhibit growth of certain larvae (Strickland et al. 1995), or through signalling triggered by patatin if it is released from the vacuole (Senda et al. 1996).

The patatins are of particular interest for the food industry since they display several properties that are rare among plant protein. An excellent amino acid composition together with high digestibility makes them interesting for nutritional purposes, while their foam-ability makes them a potential egg-substitute in plant-based food products (Ralet & Guéguen 2001; Løkra & Strætkvern 2009).

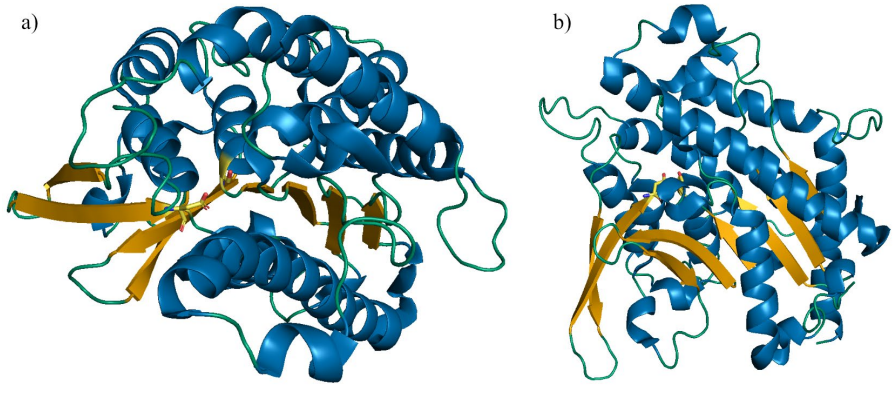

Figure 2: Crystal structure of Pat17 from two angles, a) from the side, b) from beneath. The core of the protein is formed by six beta-strand that are surrounded by seven alpha helices. Beta sheets are coloured orange, alpha helices blue and unstructured regions are green.

As discussed before, the tuber proteins in PFJ are usually treated with heat and acid and extracted as a coagulated mix. This treatment irreversibly dentures patatin, causing it to lose some of its desired properties (Pots et al. 1998; Ralet & Guéguen 2001). There are however, some methods developed for production of high purity patatin extracts. One that has been frequently

studied is ultrafiltration, although it has met some challenges. First of all is frequent filter fouling, leading to high maintenance costs (Eriksson & Sivik 1976), and second is the high glycoalkaloid content left in the final product (Wojnowska et al. 2006). The former has been solved by flocculation, using calcium, though the method has been patented and is thus not widely applicable on an industrial scale (Edens et al. 1999). The latter is usually solved by subjecting the final product to dialysis, which is a time-consuming process that further increases costs (Zwijnenberg et al. 2002).

There is also expanded bed chromatography, which is currently in use by Solanic, a subsidiary of the Dutch starch producer AVEBE (Strætkvern et al. 1998; Giuseppin et al. 2008). Expanded bed chromatography uses fluidised particles which allows for processing of crude extracts without previous treatment such as filtering or centrifugation which would normally be required for chromatographic methods such as size exclusion or ion exchange (Hjorth 1997; Vennapusa et al. 2008).

2.2.2 Protease inhibitors

Protease inhibitors are a large and diverse set of proteins with sizes range from 4 kDa, up to 22 kDa for some Kunitz type protease inhibitors (Bauw et al. 2006; Alting et al. 2011), and exhibiting a wide variety in both catalytic activity, and target specificity.

From a nutritional perspective they were initially discussed as anti-nutritional compounds, which is understandable since their natural function is as defence compounds (Ryan 1990). There have however also been some studies suggesting that protease inhibitors can possess anti-carcinogenic activity (Kennedy 1998), or that they could be used for dietary purposes since they limit energy uptake from protein-rich food (Hill et al. 1990).

2.2.3 Plastidial starch phosphorylase

After patatins and protease inhibitors, the remaining protein fraction consists mainly of enzymes involved in starch synthesis, one of the most highly expressed ones being plastidial starch phosphorylase (Pho1, or L type) (Sonnewald et al. 1995; Albrecht et al. 2001). Pho1 catalyses the reversible addition or removal of glucose-1-phosphate to an α-glucan chain. Despite the high expression, its physiological role for potato transitory and reserve starch remains a topic of debate. It has been suggested that it could be involved in transitory starch turnover, as has been shown in several other starch crops

(Schupp & Ziegler 2004; Hwang et al. 2016), or that it could be involved in starch granule initiation (Mérida & Fettke 2021).

Potato has been shown to carry two *Pho1* genes, *Pho1a* and *Pho1b*. *Pho1a* is the more ubiquitous one, with expression throughout the plant, and the highest protein level of the two in both tuber and leaf. *Pho1b* is expressed mainly in leaves, and this was the only tissue where it could be detected at protein level (Albrecht et al. 2001).

2.3 Protein engineering

2.3.1 Protein stability

Protein structure and its effect on stability has been an area of interest for many years. At its most simple, proteins are divided into two parts, the hydrophobic core, and a mostly hydrophilic surface. A tightly packed hydrophobic core without pockets or interference from polar residues will yield a more stable protein. Conversely, surface hydrophobic residues will destabilise structure, as will large unstructured regions (Anderson et al. 1993).

Protein secondary structure contribute to stability on two levels, local (through interactions with amino acids that are close in the protein sequence), or global (through interactions that are far apart in the protein sequence, or even on a separate peptide). On a local level, the most characteristic interaction is probably that of the alpha helix, where every turn is 2.7 amino acids long and kept together by H-bonds in the peptide backbone. Global interactions are more diverse, and can be both larger structures interacting, such as beta-strands forming sheets, or alpha helix bundles (Tompa et al. 2016). Or smaller, such as the di-sulphide bridges formed by cysteine. Global interactions can also be protein-protein interactions, or reversible binding of ligands such as structurally important metal ions, which can lead to stabilisation (Modarres et al. 2016). In addition to the protein sequence and structure, there are several post-translational modifications that can affect the stability of proteins. These include glycosylation, methylation and phosphorylation to name a few (Lee et al. 2023).

2.3.2 Adapter proteins

One of the many ways that nature has optimised chemistry within the cell is through proximity. This can be through localising enzymes to a specialised organelle, or thorough physical interaction. Adapter protein domains, such as SH2 and SH3, are often found as part of signal-transducing proteins where the modularity they offer leads to an additional layer of control (Horn & Sticht 2015).

Adapter protein domains have also been found in large enzyme complexes, where it is believed that they facilitate complex formation and thus rapid processing of intermediary products to speed up synthesis, a process known as substrate channelling (Conrado et al. 2008). Artificial substrate channelling has been successfully employed to facilitate the production of several compounds, such as glucaric acid (Moon et al. 2010) and resveratrol (Wang & Yu 2012).

One organism that has been extensively studied because of its efficient use of large protein complexes is *Clostridium thermocellum*, which uses a complex of enzymes to degrade cellulose. This cellulosome is built around a protein scaffold, CipA, a non-catalytic protein that is anchored to the cell surface. CipA contains a series of cohesion domains, to which the catalytic subunits can bind using a matching dockerin domain (Jeon et al. 2012; Blumer-Schuette et al. 2014). This dockerin-cohesin system has been suggested to be an interesting source for protein adaptors for use in biotechnology (Bayer et al. 1994), and has been successfully used for affinity chromatography (Craig et al. 2006)

2.4 Gene editing

A brief history of gene editing

The use of homing nucleases has been a staple of the gene editing field for many years. The method that paved the way for modern gene editing was the use of meganucleases. The meganucleases are self-splicing elements, found in all domains of life. They introduce double-stranded DNA breaks (DSBs), at recognition sequences of 14-40 bp (Belfort & Roberts 1997).

The generation of mutated lines then relies on errors made by the cell's endogenous DNA repair system. Most commonly, double-stranded breaks are repaired by non-homologous end joining (NHEJ), which does not utilize

a template for repair, but instead finds a way to put the two strands back together. This process is prone to errors, and short insertions and deletions (indels) are a common occurrence. When these indels are not triplets, it results in a frameshift, which leads to truncated or non-functional proteins being expressed (Rouet et al. 1994).

Another DNA-repair mechanism that is sometimes used by the cell is homology- directed repair (HDR). In this case, DNA is repaired using a homologous gene or sequence as template. By providing a DNA template flanked by sequences that are homologous to the target site, researchers can introduce any desired DNA sequence into the genome (Cohen-Tannoudji et al. 1998). Use of these repair pathways have been a staple of gene editing ever since.

While the long recognition sequences made the meganucleases very specific, the complex nature of the enzymes did not allow for targeting of tailored sequences. Instead, the recognition sequence of the meganuclease employed had to be introduced in the gene of interest if it was not already present (Stoddard 2014).

The first major breakthrough came with the discovery of zinc finger nucleases (ZFNs) (Kim et al. 1996). Zinc fingers are transcription factors that recognise 3-4 base sequences in the genome. By chaining several zinc fingers together, the recognition sequence could be tailored to your specific needs. To this recognition element was fused a nuclease (usually FokI), which would cut the DNA. By targeting both strands, DSBs could be easily obtained. There were however, major drawbacks with the method, not least of which was high toxicity due to sometimes rampant off target restriction (Carroll 2011).

The next evolution was transcription activator-like effector nucleases (TALENs). TALENS utilize 34 amino acid long conserved repeats derived from transcription activator-like effectors (TALEs) to guide their nucleases instead of zinc fingers (Christian et al. 2010; Mussolino et al. 2011). Each repeat contains two amino acids (position 12 and 13) that together recognises one of the four canonical DNA bases. By chaining these DNA-base recognising domains, any gene sequence can be targeted. TALENs gained popularity due to their higher specificity, and decreased off-target effect, but had some problems of their own related to their complex nature, which complicated construction, and their large size, which made delivery less efficient (Holkers et al. 2013).

During the last few years, applications of these technologies have largely been abandoned for the rising star of CRISPR/Cas9.

2.4.1 CRISPR/Cas

In 1987, while studying the sequence of the *iap*-gene in *Escherichia coli*, researchers at Osaka University in Japan reported a series of repeats in the regions flanking their gene of interest (Ishino et al. 1987). The same structure was later reported also in archaea (Mojica et al. 1993) suggesting that whatever these repeats were, they were important. The idea was finally posed that these were part of a microbial immune system (Bolotin et al. 2005; Mojica et al. 2005; Pourcel et al. 2005). Just before this, the structures had been given the name "clustered regularly interspaced palindromic repeats" (CRISPR), and at the same time a number of CRISPR associated (Cas) genes were discovered (Jansen et al. 2002).

The first studies of the mechanism showed how new spacers were added (Barrangou et al. 2007), and that immunity was achieved through targeting of viruses on the DNA level (Marraffini & Sontheimer 2008), and not through RNA interference as had been previously speculated. The final pieces of the puzzle came together when the two types of RNA required to make the protein functional, crRNA (Brouns et al. 2008) and tracrRNA (Deltcheva et al. 2011), were characterised. It would be the combination of these to RNAs into a single guide RNA (sgRNA) that unlocked the full potential of CRISPR/Cas for applications in biotechnology (Jinek et al. 2012). The system was then further optimised before being successfully applied in eukaryotic cells (Cong et al. 2013; Mali et al. 2013).

It is worth noting the Cas-proteins are divided into two classes, Class-I, which are multi-protein complexes, and Class-II, which consist of single, multi-domain proteins. For biotechnological applications, the focus has mainly been on the Class-II Cas-proteins (Makarova et al. 2011).

While many methods have been developed based on the CRISPR/Cas system, some of which are described below, the most common application remains the reverse genetics approach using *Streptococcus pyogenes* Cas9 (SpCas9). After introduction to the cell, the CRISPR/Cas9 complex binds the target DNA at the site recognised by the protospacer of the sgRNA (Fig. 3). The DNA is then cut upstream of a short recognition sequence, the "protospacer adjacent motif" (PAM). What PAM sequence is recognised depends on the Cas-protein used, but in the case of Cas9 it is NGG'. Cas9

houses two nuclease domains, RuvC which cuts the non-complementary strand, and HNH which digest the complementary strand, causing a double stranded break (Jinek et al. 2012). As mentioned before this will induce repair by either NHEJ or HDR, during which errors will result in edits. An additional feature of Cas9 is that edits will disrupt the recognition sequence; ensuring edited sequences are not continually targeted.

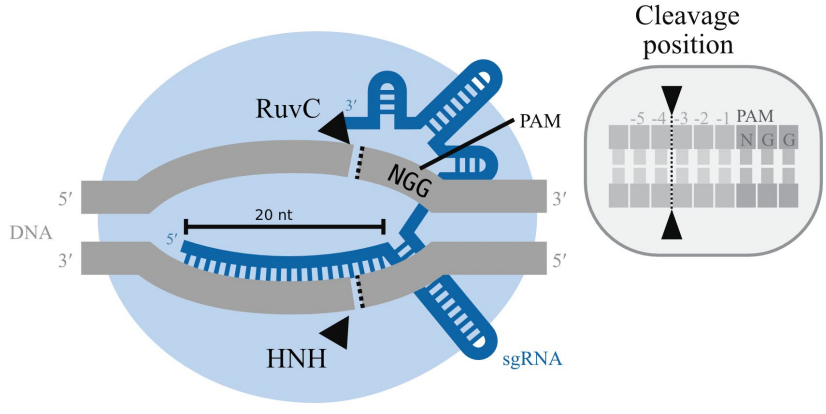

Figure 3: A schematic representation of Cas9 (light blue) bound to DNA (grey) through hybridisation with the sgRNA (dark blue). The PAM site and the two nuclease domains (RuvC and HNH) are labelled. Adapted from J.J. Froehlich, original available from Wikimedia under a CC BY-SA 4.0 license.

While SpCas9 is still the most common protein used, there are other options being investigated. For example, Cas12a, which produces a staggered cut, Cas12f, a smaller protein that simplified introduction to the cell, and Cas13, an RNA-recognising protein that has been used in diagnostics, perhaps most prominently for Covid-19 tests. There are also a slew of engineered variants, which aim to alter the PAMs recognised, or that eliminate the activity of either of both nuclease domains, creating nickases (nCas) or without nuclease activity (dCas). The dCas is particularly interesting since it can be used purely for DNA recognition. This property has been used for several purposes, such as chromosome tagging by GFP fusion (Chen et al. 2013), as well as gene regulation (Cheng et al. 2013; Gilbert et al. 2013; Qi et al. 2013).

2.4.2 Base editing

One of the first CRISPR/Cas-based precision editing technologies developed were base editors. Base editors utilize either a dead Cas-protein, or a nickase, linked to a DNA-modifying enzyme. The first base editors developed were cytosine and adenosine base editors, which utilizes deaminases to chemically alter the bases into uracil and inosine, respectively (Komor et al. 2016; Gaudelli et al. 2017). The induced mismatch will trigger DNA repair and when the edited strand is read, the uracil will read as thymine, and inosine will read as guanine. Causing the C→T and A→G exchanges, respectively. These have also been combined to generate dual base editors, which have been suggested as a tool to mimic error prone PCR for directed evolution (Grünewald et al. 2020; Li et al. 2020).

While not as prominent as the deaminase versions, constructs are being developed which allow for C→G and A →T transversions (Zhao et al. 2021a; Chen et al. 2023; Tong et al. 2023b). Together with the deaminases, these transversions allow for the changing of any base into any other base, as long as there are suitable PAMs on both strands.

While the availability of PAMs is of great importance for all CRISPR/Cas applications, they are particularly interesting in the use of base editors because of the editing window. Since the DNA-altering enzymes are not sequence-specific, they instead have an editing window, a range of amino acids that have a probability of being edited. While most optimisations aim to reduce the editing window, there are also applications where an extended window could be preferable, such as when applying saturation editing methods (Tong et al. 2023a).

The narrowing of the editing window also places a greater demand on the PAM, since it needs to be a set number of bases from the bases that are being targeted for edits. This problem is being tackled in two ways, the mutation of Cas9 to target alternative PAMs (Tan et al. 2019), and the development of base editors based on other Cas-proteins such as Cas12a (Wang et al. 2023)

2.4.3 Prime editing

One exciting development during the last few years has been the introduction of prime editors. These are nickases that cleave the non-target strand, linked to a reverse transcriptase (RT) (Anzalone et al. 2019). They also utilize an extended guide RNA, a so called prime editing guide RNA (pegRNA) that carry a primer-binding sequence as well as a modification sequence. The

primer-binding sequence binds the nicked DNA strand, which allows the reverse transcriptase to write the modification sequence in to the DNA. As with the previously described methods, the cell will recognise this anomaly and repair it. If the edited strand is retained and used as template, the edit is successfully maintained in the genome(Anzalone et al. 2019).

The first prime editors were developed using Cas9 and Moloney Murine Leukaemia Virus (MMLV)-RT, but the system has been reproduced using several other RTs including CamV-RT and human foamy virus-RT (HFV-RT) (Lin et al. 2020; Grünewald et al. 2023). Several mutated variants of MMLV have also been investigated, including truncated variants such as MMLV-RTΔH, where the C-terminal RNase H domain has been removed (Zheng et al. 2022).

While still primarily a fusion protein-based method, the size limitations of virus mediated transfection of human cells has led to the investigation of ways to divide the complex. One common way is using a split intein system (Gao et al. 2022; Zheng et al. 2022). Inteins are protein sequences that exist inside other proteins. They have the ability to join the two flanking protein regions, cutting itself out in the process. In the split intein system, the intein is divided between two separate proteins. When joined together the intein will become active and splice the two proteins together (Wang et al. 2022). There has also been an interesting study which show that prime editors remain efficient, even if entirely split up (Grünewald et al. 2023), suggesting that the pegRNA is enough to attract the two components .

2.4.4 CRISPR/Cas application in plants

Since the first applications in plants (Li et al. 2013; Nekrasov et al. 2013; Shan et al. 2013), several plant biology methods have been adapted for CRISPR/Cas delivery, using either stable transformation of a CRISPR/Cas-cassette, or through transient systems based on DNA, RNA or Ribonucleoprotein (RNP).

Stable transformation of plants was achieved in the early 1980's using *Agrobacterium tumefaciens* (Barton et al. 1983; Herrera-Estrella et al. 1983; Hoekema et al. 1983), and have been used in several studies for the introduction of a CRISPR/Cas-cassette expressing both Cas-protein and guide RNA. Stable transformation can also be achieved using biolistics, where microscopic particles, are coated in reagent and fired at plant tissue (Liang et al. 2018). Once introduced to the genome, the CRISPR/Cas-

cassette remains active until removed, something that leads to both high efficiency, and a major risk of off-target effects. Thus, it is common practise to segregate out the editing genes using crosses. Something that causes problems in clonally propagated crops as well as crops with long generation time (Chen et al. 2019).

An alternative to stable transformation is using transient systems. Both *Agrobacterium sp.* mediated delivery and biolistics can be used for delivery of transient systems. DNA based vectors in the case of *Agrobacterium* (Chen et al. 2018), while biolistics has also been used for delivery of RNA and RNP (Svitashev et al. 2016). A third option for transient expression is protoplast transfection. Protoplasts are plant cells that have been enzymatically treated to remove the cell wall. This leaves them susceptible to several different treatments that can be utilized to introduce editing vectors. Perhaps the most prevalent of which is PEG-mediated transfection (Woo et al. 2015).

When applying a transient system, it is important to consider what form to deliver the editing reagents in. DNA vectors are easy to prepare, but have been known to occasionally leave behind integrated fragments in the host genome. RNA based vectors circumvent this problem, but are often short-lived. The use of RNP requires the separate procurement of sgRNA and Cas-protein, but have been shown to be both highly efficient and does not risk unwanted integration of the vector (Kim et al. 2017; Chen et al. 2019).

3. Aims and objectives

The aim of this study was to identify mutant variants of patatin with improved thermostability, and to investigate available technologies with the long-term goal to introduce these mutations into patatin genes of commercially utilized starch potato varieties. This was divided into several objectives:

- To analyse the patatin structure and develop a series of mutants for in-depth analyses.
- To develop a simple expression and analysis pipeline for patatin mutant evaluation.
- To evaluate and improve available precision-editing methods in potato protoplasts.
- To express and purify the components required for precision editing for use as ribonucleoproteins in potato protoplasts.
- To identify future targets for modification, an investigation of plastidial starch phosphorylase (Pho1), another protein highly expressed in tubers, was also conducted.

Taken together these objectives strive to establish a pipeline for protein engineering in plants. Showing that plant proteins can be optimised using high throughput methods to identify beneficial mutations that can then be introduced into plants.

4. Research outcomes

4.1 Patatin expression patterns in Kuras and Desirée

In potato, many patatin genes are known to reside in a clustered locus. The genomic structure of this patatin region has proven hard to elucidate, with reports of the copy number ranging from 4 to 20 per genome. In 2023, a new genome assembly of Phureja DM was released based on long read sequencing that yielded an improved picture of chromosome 8, where a large number of patatin genes are clustered (Yang et al. 2023).

To investigate the expression patterns of patatin genes, we mapped short read RNA-seq data from two potato cultivars; Desirée and Kuras, onto this new DM8.1 assembly. As shown in Fig 4, the different patatin genes do not express at equal level, instead there seems to be some that form the major part of tuber expression. What specific genes are highly expressed seems to vary between the two cultivars, which is in line with what has been suggested in previous studies (Bauw et al. 2006; Barta et al. 2012b).

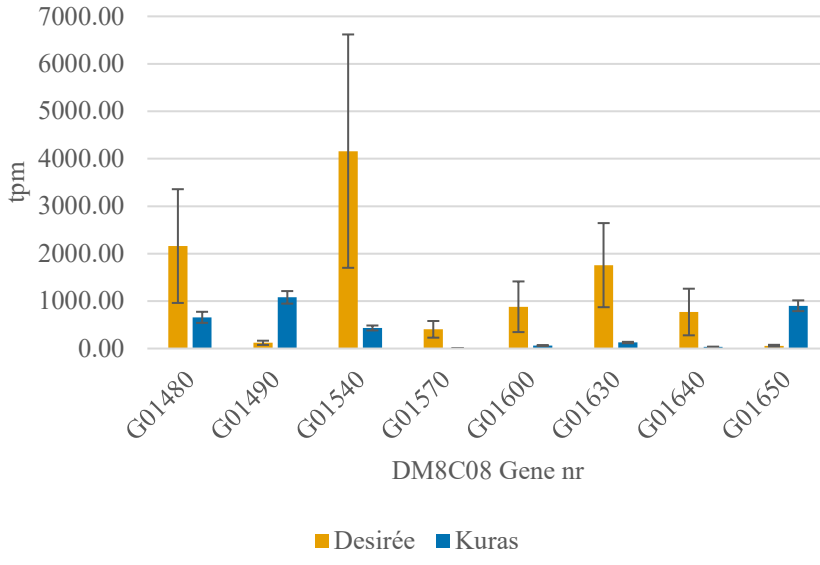

Figure 4: Expression levels of patatin genes in two potato cultivars, Desirée and Kuras.

Interestingly, only one of the expressed genes seems to be full length as well as have a preserved active site. While it is among the highly expressed genes in both investigated varieties, it is not the most abundant in either. This suggests that while some activity might be beneficial to the plant, i.e. through involvement in defence, it also seems like accumulation of catalytically active patatin for storage purposes may not be desired for the present function in the plant. Which in turn could lead to promoted expression of inactive variants. It was also discovered that one of the genes present in the chromosome 8 cluster was mainly expressed in Kuras leaves. This is interesting since the chromosome 8 cluster was previously believed to house only the tuber-localised class-I patatins.

4.2 Patatin mutation study

To investigate the potential of targeting patatin for gene editing, a mutation study of patatin in cultivar Desirée was conducted. The active site was chosen as target, and a sgRNA was designed. Protoplasts of Desirée were prepared and transfected with RNP complex using PEG treatment by a published method (Nicolia et al. 2021). Shoots were regenerated and

analysed using high-resolution fragment analysis to detect indel events (Fig. 5a).

Several events analysed showed indications of indels were induced, there was however, a large fraction of wild type (wt) alleles in all samples analysed. This was assumed to be due to the high number of allele variants, so a selection of events was sent for amplicon sequencing to get a better picture of how the edits were distributed. The sequencing data was processed using the CRISPResso2 web service and the frequency of edited alleles is summarised in Fig. 5b. Alleles were characterised by a sequence 4nt upstream of the Cas9 cut site, since the nucleotides closes to the cut site have been proven to be the most influential on editing efficiency (Jinek et al. 2012).

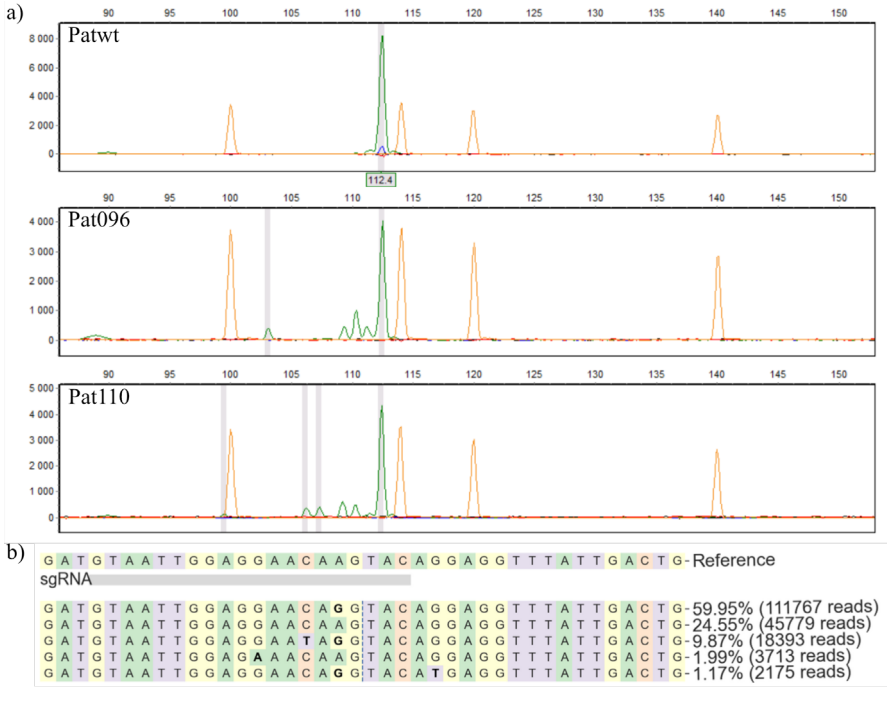

Figure 5: Characterisation of patatin lines with induced indels. a) High resolution fragment analysis of region flanking the target site shown in green, and size marker in orange. b) Distribution of patatin alleles in Desirée as characterised by amplicon sequencing.

Over all the results from the mutation study are promising. Using the HRFA data as basis, over 80% of regenerated events can be estimated to carry mutations at the desired site. Moreover, we can see that of an estimated 84 patatin copies in the tetraploid genome, over 10% are reliably edited (Paper I, Fig. 1c). We also see clear allele preference for editing, which is expected, but interestingly one of the alleles with a TAGG-sequence directly upstream the predicted cut site increased in frequency of reads compared to wt. This could be caused by PCR bias against some mutant alleles, which in turn would underestimate the editing efficiency (Paper I, Fig. 1d).

Comparing the deep sequencing data to our previous analysis of the DM8.1 genome assembly reveals some additional interesting features. Among the three sequence variants close to the PAM site defined, one (TAGG) is missing from DM8.1 further highlighting the patatin diversity among cultivars. Another layer is added to this when the RNA mapping data is added, showing that the different alleles are not evenly expressed between the two cultivars. This poses a question; will it be more efficient to multiplex for complete gene coverage, or to focus on targeting only those genes that are highly expressed in the studied cultivar?

4.3 Patatin engineering

Since there is no published structure available for potato patatin, the first step was to model the *S. tuberosum* patatin structure using the Phyre2 web service and the protein sequence for pat2-k2 from Kuras (Bauw et al. 2006; Kelley et al. 2015). The modelled structure was compared to that of the pat17 (1oxw) structure from *Solanum cardiophyllum*, a close relative which shares over 80% sequence identity on the protein level (Fig. 6). Both structures were fed into the CNAnalysis web service (Krüger et al. 2013), to identify unfolding nuclei (Radestock & Gohlke 2008). However, the modelled structure of potato patatin failed to produce reliable results, and thus the sites identified for pat17 were used. The unfolding nuclei identified were compared between the two structures using sequence alignment and were visualised on the superimposed structures (Fig. 6a and b respectively), so it was assumed that the CNA was valid also for potato patatins. This formed the basis for the protein design.

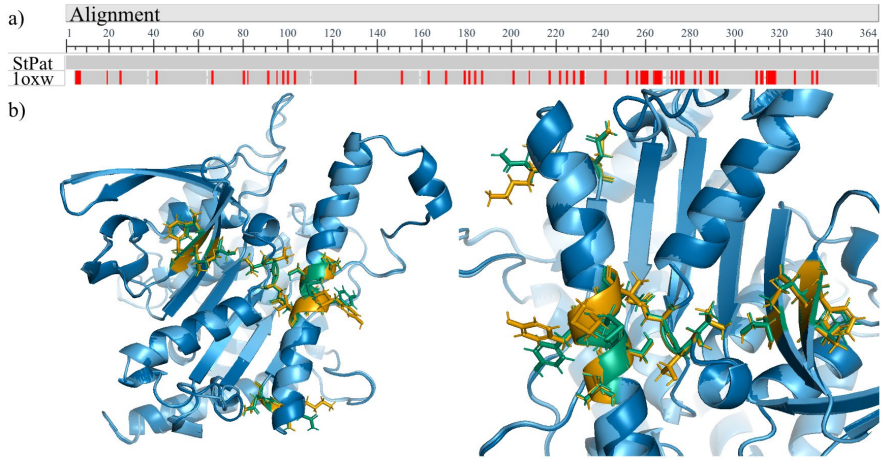

Figure 6: Differences between pat17 from *S. cardiophyllum* and pat2-k2 from *S. tuberosum*. a) Sequence alignment with the differences highlighted. Red lines represent missmatches, white lines represent gaps. b) The predicted structure of pat2-k2 (light blue) superimposed on the crystal structure of pat17 (PDB: 1oxw, dark blue). The unfolding nuclei have been colored green for pat2-k2 and orange for pat17, and the corresponding amino acid sidechains have been visualised as sticks.

A range of mutants was designed, testing a few of the common design principles in protein engineering. Two mutants introduced extra cysteine bonding pairs to try to stabilise the structure using di-sulphide bridges (M1 and M3 in Fig. 7), one mutant focused on reduced surface hydrophobicity in favour for polar residues with H-bonding partners within 2-3 Å (M2 in Fig. 7). The final mutant centred on a cluster of unstructured loops where three phenylalanines were introduced to allow for aromatic interactions such as pi-stacking (M4 in Fig. 7). A 6*His tag was added to all constructs to allow for IMAC purification.

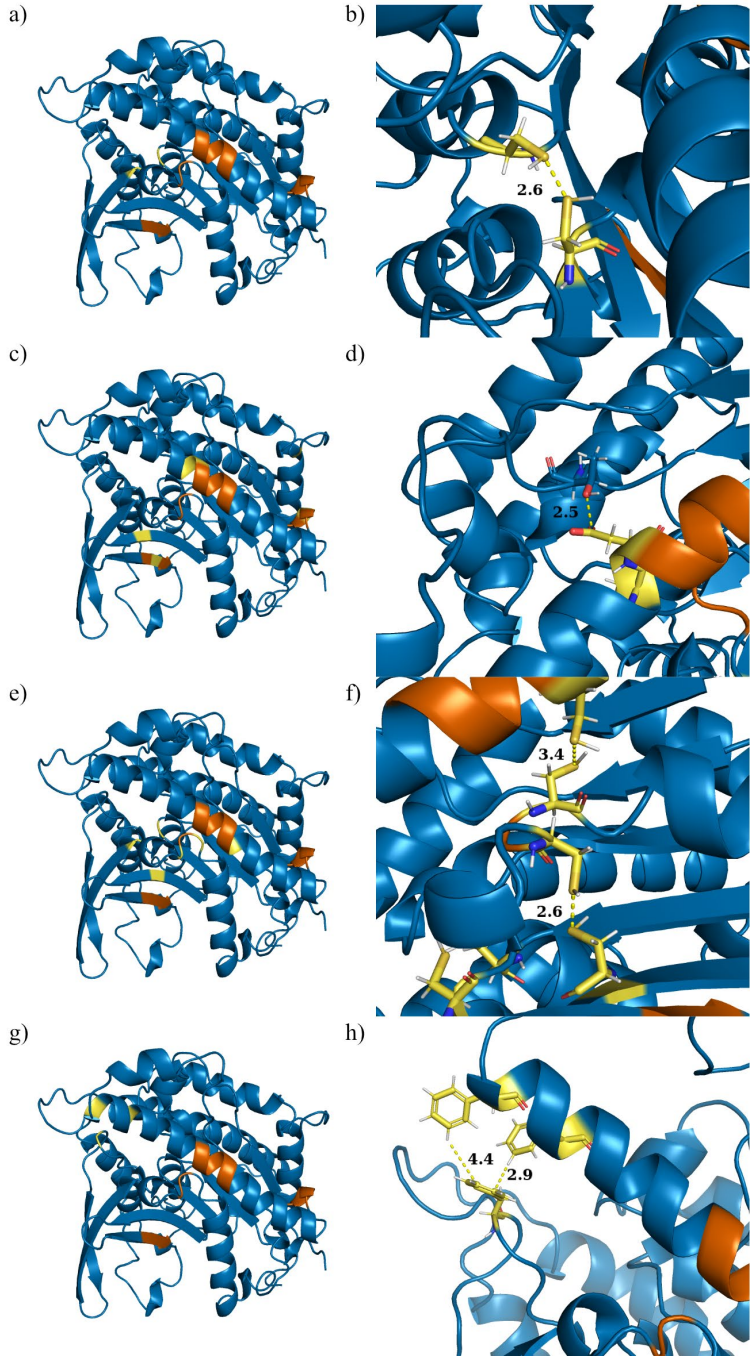

Figure 7: Designed patatin variants, with substituted amino acids highlighted.

The mutated patatin variants, as well as the wt protein, were expressed in *E. coli* where they were retrieved from inclusion bodies and refolded. They were then subjected to DSF to test the efficacy of the introduced modifications (Fig. 3 in Paper I). Initial inspection of the melt curves for M2 and M3 suggests that they are not properly folded. This was further supported when no melting point could be reliably identified using sigmoidal fitting. The remaining two samples, M1 and M4, have a melt curve more similar to that of wt and the melting temperature could be estimated by fitting of a sigmoidal curve. The average of four replicates are summarised in Fig. 8.

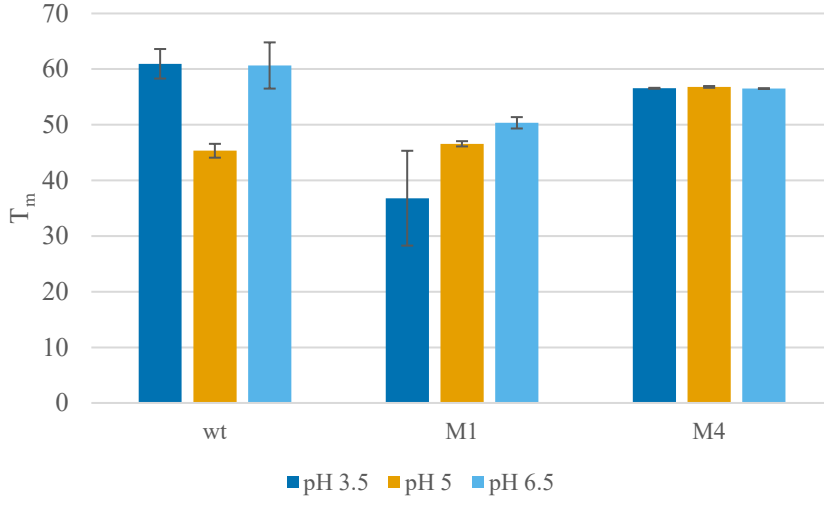

Figure 8: Calculated melting temperature of patatin variants as measured by DSF. Each bar represents the mean of four replicates and error bars are standard deviation.

4.4 Systems for patatin expression and evaluation

To be able to test the optimised patatin variants designed, there was a need for simple expression and evaluation pipeline. Since the effect of mutations on stability can be hard to predict, analysis of a larger set of mutants would increase the likelihood of finding promising candidates. Two methods were practically considered, *Agrobacterium tuberosum* infiltration of *Nicotiana benthamiana* and recombinant expression in *Saccharomyces cerevisiae* BY4742.

Initially, *N. benthamiana* was tested since this most closely mimics the natural expression conditions of patatin. Specifically the glycosylation was a

point of concern, since this is known to influence thermostability. Infiltration was done following (Snell et al. 2019), and protein extract was purified following (Carlsson et al. 2020). Example gel can be seen Fig. 9a. Infiltration effect was initially judged by co-transfection of GFP, which expressed successfully. However, no patatin could be detected in protein extracts. It is unclear whether this is due to poor expression or a high instability of the protein.

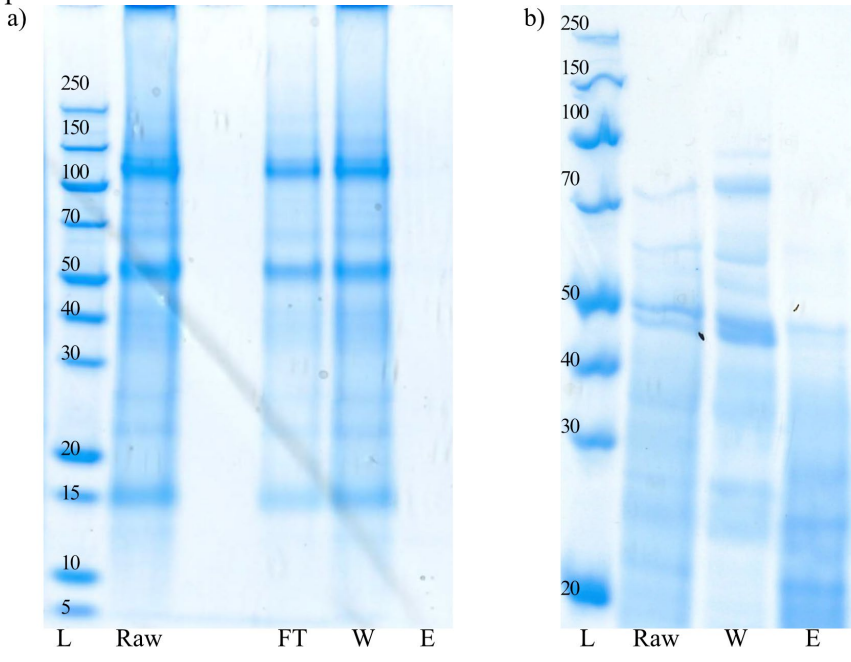

Figure 9: Extraction tests for recombinant expression of patatin, extracts and fractions from IMAC purification. L is a protein size marker, Raw is clarified extract, FT is Flow through, W is Wash and E is Elute. Patatin is expected at 42kDa. a) Agrobacterium infiltration of *Nicotiana benthamiana*. b) Expression in *Saccharomyces cerviceae.*

Yeast expression was tested as a second option, while not identical it does possess the ability to glycosylate proteins, and its microbial nature allow for easy handling of large mutant libraries for screening. Extraction was done using “Y-PER, Yeast protein extraction reagent” from Thermo Scientific, and the samples were analysed using SDS-PAGE (Fig. 9b). While initially promising bands were observed close to 50 kDa, these could not be purified using IMAC. Western blot of yeast extracts using anti-His-tag anti-bodies also failed to yield any signal, suggesting that the protein was not successfully expressed.

4.5 Modular CRISPR/Cas system using adapter proteins

As described earlier, there are fusion protein based technologies for precise editing using CRISPR/Cas. The most notable ones being base editors and prime editors. While proven to work in several species, neither technique has made it to market for use as RNP. To express and purify a construct of that size requires considerable time, which lead to the idea of a toolbox utilizing adapter proteins. Smaller proteins are usually easier to express and as more Cas-proteins are explored, they can be added to the toolbox and tested with several different fusion partners without having to optimise every different fusion construct.

The adapter proteins we chose to explore were two domains from the cohesin-dockerin system, from *Clostridium thermocellum*. The dockerin (Doc1) was taken from the *xynY* gene, and as partner was chosen the seventh domain of *CipA* (Cip7).

An initial toolbox was assembled containing Cas9, two nickases (nCas9) and one dead Cas9 (dCas9). The functional domains assembled were GFP, adenine deaminase (ABE), cysteine deaminase (CBE), and two modified reverse transcriptase (RT and RTΔH). These constructs are displayed in Fig. 10.

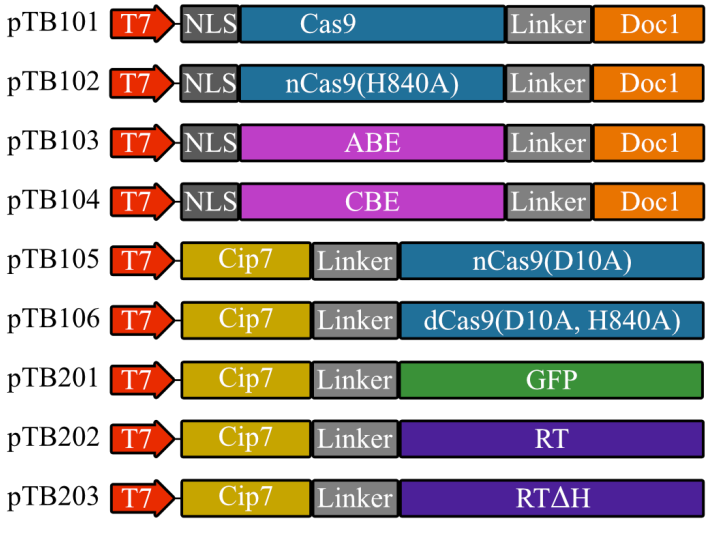

Figure 10: Constructs assembled for the adapter protein toolbox. pTB1XX constructs are based on the pMJ806 backbone, while pTB2XX constructs are based on pET30-R5-GFP. The two fusion partners are connected to either the Cip7, or Doc1 domain, through a GGSG linker. All constructs carry a 6*His tag for purification, and are expressed from the T7 promoter, which is IPTG inducible.

The pair of Cas9:Doc1 and Cip7:GFP were used to set up the initial expression and purification conditions. After being successfully purified, the activity of Cas9 was compared to that of commercially available Cas9 using *in vitro* digestion (Paper II, Fig. 2b), showing comparable activity between the two. This was followed by an initial interaction test using native PAGE, where we showed that there is a shift of the Cip7:GFP band to higher molecular weight upon addition of Cas9:Doc1. And that this shift is concentration dependent, with all GFP being shifted to the higher weight already at a 2:1 Cas9:GFP ratio. The same shift was not seen when commercial Cas9 was added to Cip7:GFP, showing us that the interaction observed is specific.

Finally, efficiency of Cas9:Doc1 as RNP was measured in potato protoplasts. Amplicon sequencing was done on protoplasts and efficiency was defined as "percent of reads carrying mutation. As seen in Fig. 11b, the efficiency was comparable to that of commercial Cas9 once more, and the efficiency of both Cas9 and Cas9:Doc1 remain unaffected by addition of Cip7:GFP. This shows great promise for the system, though co-localisation to the nucleus remains to be indisputably verified.

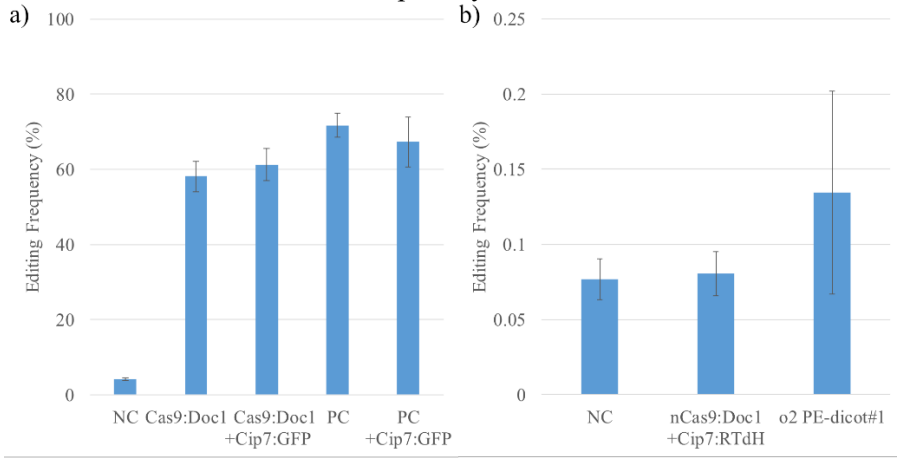

Figure 11: Editing efficiencies as estimated by amplicon sequencing. a) Application of Cas9:Doc1 targeting *Pho1* as compared to a positive control (PC) consisting of commercially available Cas9. b) The efficiency of a toolbox prime editor based on Cip7:RTdH, compared to a similar system expressed transiently from a plasmid.

Following up on the successful use of Cas9:Doc1 for RNP transfection, we decided to put the prime editing system to the test. Following the same experimental procedure as for the knock-out study, except the target was

changed from *Pho1a* to Acetohydroxyacid synthase (*AHAS*). A single nucleotide change in *AHAS* is known to result in herbicide resistance (Kolkman et al. 2004). A property that is useful to facilitate screening for desired mutations. As positive control was used a plasmid based system provided by INRAE "o2 PE-dicot#1". As shown in Fig. 11b, both prime editing systems perform in-line with previously published results in potato, with editing efficiency.

4.6 Alternative tuber protein targets

Apart from patatin, there are two major groups of proteins that could be of interest for precision editing in potato tubers. The first one would be the protease inhibitors (PTI); however, their large diversity means they could not be targeted as a whole. Meaning any edits would target only one type of PTI, limiting the impact the project could have.

The other group would be enzymes involved in starch synthesis. Since potatoes are mainly grown for their starch content, whether for food or industry, it is easy to see why these are such high values targets. Indeed, there have already been some gene editing studies targeting a few of the starch synthases and branching enzymes (Andersson et al. 2018; Zhao et al. 2021b). In contrast, the highly expressed plastidial starch phosphorylases (Pho1) are still poorly understood.

To further characterise the importance of Pho1 in tubers, the major variant, Pho1a was knocked out using CRISPR/Cas9 RNP in potato protoplasts. While analysing mutant events, it was discovered that there was a potential duplication of the *Pho1a* target. This was further established using sequencing revealing two tandem copies, *Pho1.1a* and *Pho1.2a*.

Nine mutant lines were analysed for phosphorylase activity using Zymogram (Paper III, Fig. 7). Showing that only one line retains detectable Pho1a activity, while all lines show the distinct presence of Pho2 as seen by the high MW band on the glycogen containing PAGE gel.

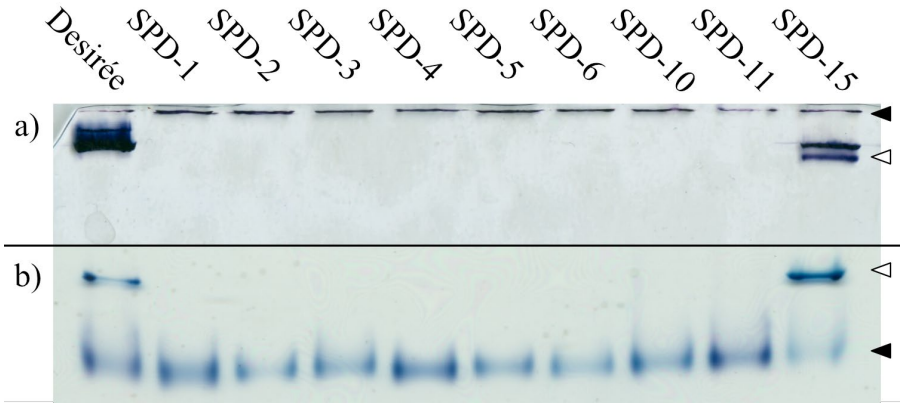

Figure 12: Activity assay of starch phosphorylase for a selection of *Pho1a* knock-out lines (SPD), and wt (Desirée). ◁ marks Pho1a activity, ◀ marks Pho2 activity. a) On PAGE cast with 1.2% glycogen. b) On commercial 10% PAGE.

When grown in greenhouse, the mutant lines produced more, but smaller tubers with a slight decrease in total starch content. When the starch granule phenotype was investigated, the mutant lines showed a high number of small and round granules, that when analysed for apparent amylose were shown to contain a significantly lower amount of amylose (14-16%, as compared to 22% for Desirée).

With such an impact on both tuber and starch yields, it is unlikely that knock-out lines of starch phosphorylase would be of commercial interest. This opens up the question of whether there are other modifications that could prove beneficial. Could *Pho1a* for example be up regulated using precision editing to increase the amylose content in tubers? If so, it could be a valuable addition to previously edited lines that are deficient in starch branching enzymes.

5. Conclusion

In this thesis, I have explored the suitability of potato tuber proteins as targets for precision gene editing and tested the concept of *in planta* protein engineering through CRISPR/Cas9.

We have shown that patatin can be successfully targeted, though the diverse genetic make-up makes it important to study the chosen cultivar beforehand. In Paper I we show that patatin expression in both analysed cultivars is dominated by a few genes, and that these genes are cultivar dependent. This in turn makes RNA sequencing a valuable tool to determine what gene targets are likely to give the largest effect on the protein level.

Though the suggested pipeline for heterologous patatin expression and evaluation could not be established, a small subset of patatin mutants were retrieved from *E. coli* inclusion bodies. Analysis of the melting temperature revealed that mutants carrying additional cysteines more readily form aggregates, which occludes the use of di-sulphide bridges to stabilise the structure. In contrast, the introduction of aromatic sidechains to unstructured loops seems to have made unfolding less pH dependent. A trait that would be interesting to study further.

In order to implement the desired type of edits, we developed an adapter protein based approach for expressing and purifying large fusion constructs. In Paper II, we showed that addition of an adapter protein domain does not alter the efficiency of Cas9 *in vitro* nor *in vivo*. We further showed that Cas9:Doc1 successfully could bind to two different Cip7 partners.

Finally, we have evaluated *Pho1a* as a future target for precision editing. The knock-out lines generated in Paper III show an aberrant phenotype, namely small and misshapen starch granules, and a decrease in apparent amylose. Suggesting that *Pho1a* could be an interesting target for up-

regulation, especially interesting would be the study of synergistic effects in amylose rich lines.

These studies together shows that the humble potato tuber still holds great potential and can be further enhanced for further use in sustainable food and industrial applications. They also show that there is potential of applying a synthetic biology approach to crop development. Although synthetic biology is still in its infancy, and specific precision engineering techniques of crops remain impractical at a larger scale, the rapid continued development of CRISPR/Cas methods will pave the way for a multitude of interesting applications in the near future.

References

Albrecht, T., Koch, A., Lode, A., Greve, B., Schneider-Mergener, J. & Steup, M. (2001). Plastidic (Pho1-type) phosphorylase isoforms in potato (Solanum tuberosum L.) plants: expression analysis and immunochemical characterization. *Planta*, 213 (4), 602–613. https://doi.org/10.1007/s004250100525

Alting, A.C., Pouvreau, L., Giuseppin, M.L.F. & van Nieuwenhuijzen, N.H. (2011). 12 - Potato proteins. In: Phillips, G.O. & Williams, P.A. (eds) *Handbook of Food Proteins*. Woodhead Publishing. 316–334. https://doi.org/10.1533/9780857093639.316

Anderson, D.E., Hurley, J.H., Nicholson, H., Baase, W.A. & Matthews, B.W. (1993). Hydrophobic core repacking and aromatic-aromatic interaction in the thermostable mutant of T4 lysozyme Ser 117-->Phe. *Protein Science : A Publication of the Protein Society*, 2 (8), 1285–1290

Andersson, M., Turesson, H., Olsson, N., Fält, A.-S., Ohlsson, P., Gonzalez, M.N., Samuelsson, M. & Hofvander, P. (2018). Genome editing in potato via CRISPR-Cas9 ribonucleoprotein delivery. *Physiologia Plantarum*, 164 (4), 378–384. https://doi.org/10.1111/ppl.12731

Anzalone, A.V., Randolph, P.B., Davis, J.R., Sousa, A.A., Koblan, L.W., Levy, J.M., Chen, P.J., Wilson, C., Newby, G.A., Raguram, A. & Liu, D.R. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*, 576 (7785), 149–157. https://doi.org/10.1038/s41586-019-1711-4

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A. & Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science (New York, N.Y.)*, 315 (5819), 1709–1712. https://doi.org/10.1126/science.1138140

Barta, J., Bartova, V., Zdrahal, Z. & Sedo, O. (2012a). Cultivar Variability of Patatin Biochemical Characteristics: Table versus Processing Potatoes (Solanum tuberosum L.). *Journal of Agricultural and Food Chemistry*, 60 (17), 4369–4378. https://doi.org/10.1021/jf3003448

Barta, J., Bartova, V., Zdrahal, Z. & Sedo, O. (2012b). Cultivar Variability of Patatin Biochemical Characteristics: Table versus Processing Potatoes (Solanum tuberosum L.). *Journal of Agricultural and Food Chemistry*, 60 (17), 4369–4378. https://doi.org/10.1021/jf3003448

Barton, K.A., Binns, A.N., Matzke, A.J.M. & Chilton, M.-D. (1983). Regeneration of intact tobacco plants containing full length copies of genetically engineered T-DNA, and transmission of T-DNA to R1 progeny. *Cell*, 32 (4), 1033–1043. https://doi.org/10.1016/0092-8674(83)90288-X

Bauw, G., Nielsen, H.V., Emmersen, J., Nielsen, K.L., Jorgensen, M. & Welinder, K.G. (2006). Patatins, Kunitz protease inhibitors and other major proteins in tuber of potato cv. Kuras. *Febs Journal*, 273 (15), 3569–3584. https://doi.org/10.1111/j.1742-4658.2006.05364.x

Bayer, E.A., Morag, E. & Lamed, R. (1994). The cellulosome — A treasure-trove for biotechnology. *Trends in Biotechnology*, 12 (9), 379–386. https://doi.org/10.1016/0167-7799(94)90039-6

Belfort, M. & Roberts, R.J. (1997). Homing endonucleases: keeping the house in order. *Nucleic Acids Research*, 25 (17), 3379–3388. https://doi.org/10.1093/nar/25.17.3379

Blumer-Schuette, S.E., Brown, S.D., Sander, K.B., Bayer, E.A., Kataeva, I., Zurawski, J.V., Conway, J.M., Adams, M.W.W. & Kelly, R.M. (2014). Thermophilic lignocellulose deconstruction. *FEMS Microbiology Reviews*, 38 (3), 393–448. https://doi.org/10.1111/1574-6976.12044

Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S.D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology (Reading, England)*, 151 (Pt 8), 2551–2561. https://doi.org/10.1099/mic.0.28048-0

Brouns, S.J.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J.H., Snijders, A.P.L., Dickman, M.J., Makarova, K.S., Koonin, E.V. & Oost, J. van der (2008). Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science (New York, N.Y.)*, 321 (5891), 960. https://doi.org/10.1126/science.1159689

Carlsson, M.L.R., Kanagarajan, S., Bülow, L. & Zhu, L.-H. (2020). Plant based production of myoglobin - a novel source of the muscle heme-protein. *Scientific Reports*, 10 (1), 920. https://doi.org/10.1038/s41598-020-57565-y

Carroll, D. (2011). Genome Engineering With Zinc-Finger Nucleases. *Genetics*, 188 (4), 773–782. https://doi.org/10.1534/genetics.111.131433

Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G.-W., Park, J., Blackburn, E.H., Weissman, J.S., Qi, L.S. & Huang, B. (2013). Dynamic Imaging of Genomic Loci in Living Human Cells

by an Optimized CRISPR/Cas System. *Cell*, 155 (7), 1479–1491. https://doi.org/10.1016/j.cell.2013.12.001

Chen, K., Wang, Y., Zhang, R., Zhang, H. & Gao, C. (2019). CRISPR/Cas Genome Editing and Precision Plant Breeding in Agriculture. *Annual Review of Plant Biology*, 70 (1), 667–697. https://doi.org/10.1146/annurev-arplant-050718-100049

Chen, L., Hong, M., Luan, C., Gao, H., Ru, G., Guo, X., Zhang, D., Zhang, S., Li, C., Wu, J., Randolph, P.B., Sousa, A.A., Qu, C., Zhu, Y., Guan, Y., Wang, L., Liu, M., Feng, B., Song, G., Liu, D.R. & Li, D. (2023). Adenine transversion editors enable precise, efficient A•T-to-C•G base editing in mammalian cells and embryos. *Nature Biotechnology*, 1–13. https://doi.org/10.1038/s41587-023-01821-9

Chen, L., Li, W., Katin-Grazzini, L., Ding, J., Gu, X., Li, Y., Gu, T., Wang, R., Lin, X., Deng, Z., McAvoy, R.J., Gmitter, F.G., Deng, Z., Zhao, Y. & Li, Y. (2018). A method for the production and expedient screening of CRISPR/Cas9-mediated non-transgenic mutant plants. *Horticulture Research*, 5 (1), 1–12. https://doi.org/10.1038/s41438-018-0023-4

Cheng, A.W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T.W., Rangarajan, S., Shivalila, C.S., Dadon, D.B. & Jaenisch, R. (2013). Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Research*, 23 (10), 1163–1171. https://doi.org/10.1038/cr.2013.122

Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J. & Voytas, D.F. (2010). Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*, 186 (2), 757–761. https://doi.org/10.1534/genetics.110.120717

Cohen-Tannoudji, M., Robine, S., Choulika, A., Pinto, D., El Marjou, F., Babinet, C., Louvard, D. & Jaisser, F. (1998). I-SceI-Induced Gene Replacement at a Natural Locus in Embryonic Stem Cells. *Molecular and Cellular Biology*, 18 (3), 1444–1448

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A. & Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science (New York, N.Y.)*, 339 (6121), 819–823. https://doi.org/10.1126/science.1231143

Conrado, R.J., Varner, J.D. & DeLisa, M.P. (2008). Engineering the spatial organization of metabolic enzymes: mimicking nature's synergy. *Current Opinion in Biotechnology*, 19 (5), 492–499. https://doi.org/10.1016/j.copbio.2008.07.006

Craig, S.J., Foong, F.C. & Nordon, R. (2006). Engineered proteins containing the cohesin and dockerin domains from Clostridium thermocellum provides a reversible, high affinity interaction for biotechnology applications. *Journal of Biotechnology*, 121 (2), 165–173. https://doi.org/10.1016/j.jbiotec.2005.07.005

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J. & Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, 471 (7340), 602–607. https://doi.org/10.1038/nature09886

Do, P.T., Nguyen, C.X., Bui, H.T., Tran, L.T.N., Stacey, G., Gillman, J.D., Zhang, Z.J. & Stacey, M.G. (2019). Demonstration of highly efficient dual gRNA CRISPR/Cas9 editing of the homeologous GmFAD2–1A and GmFAD2–1B genes to yield a high oleic, low linoleic and α-linolenic acid phenotype in soybean. *BMC Plant Biology*, 19 (1), 311. https://doi.org/10.1186/s12870-019-1906-8

Edens, L., Lee, J.A.B.V.D. & Plijter, J.J., *Novel food compositions*. EP0912105A1. (1999). https://patents.google.com/patent/EP0912105A1/en [2024-04-18]

Eriksson, G. & Sivik, B. (1976). Ultrafiltration of potato process water—influence of processing variables. *Potato Research*, 19 (3), 279–287. https://doi.org/10.1007/BF02365722

Friedman, M., McDonald, G.M. & Filadelfi-Keszi, P.M. (1997). Potato Glycoalkaloids: Chemistry, Analysis, Safety, and Plant Physiology. *Critical Reviews in Plant Sciences*,. https://doi.org/10.1080/07352689709701946

Galliard, T. (1971). The enzymic deacylation of phospholipids and galactolipids in plants. Purification and properties of a lipolytic acyl-hydrolase from potato tubers. *Biochemical Journal*, 121 (3), 379–390. https://doi.org/10.1042/bj1210379

Gao, Z., Ravendran, S., Mikkelsen, N.S., Haldrup, J., Cai, H., Ding, X., Paludan, S.R., Thomsen, M.K., Mikkelsen, J.G. & Bak, R.O. (2022). A truncated reverse transcriptase enhances prime editing by split AAV vectors. *Molecular Therapy: The Journal of the American Society of Gene Therapy*, 30 (9), 2942–2951. https://doi.org/10.1016/j.ymthe.2022.07.001

Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I. & Liu, D.R. (2017). Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*, 551 (7681), 464–471. https://doi.org/10.1038/nature24644

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., Lim, W.A., Weissman, J.S. & Qi, L.S. (2013). CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell*, 154 (2), 442–451. https://doi.org/10.1016/j.cell.2013.06.044

Giuseppin, M.L.F., Van Der Sluis, C. & Laus, M.C., *Native Potato Protein Isolates*. (2008). https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2008069650 [2024-02-27]

Grommers, H.E. & van der Krogt, D.A. (2009). Chapter 11 - Potato Starch: Production, Modifications and Uses. In: BeMiller, J. & Whistler, R. (eds) *Starch (Third Edition)*. Academic Press. 511–539. https://doi.org/10.1016/B978-0-12-746275-2.00011-2

Grünewald, J., Miller, B.R., Szalay, R.N., Cabeceiras, P.K., Woodilla, C.J., Holtz, E.J.B., Petri, K. & Joung, J.K. (2023). Engineered CRISPR prime editors with compact, untethered reverse transcriptases. *Nature biotechnology*, 41 (3), 337–343. https://doi.org/10.1038/s41587-022-01473-1

Grünewald, J., Zhou, R., Lareau, C.A., Garcia, S.P., Iyer, S., Miller, B.R., Langner, L.M., Hsu, J.Y., Aryee, M.J. & Joung, J.K. (2020). A dual-deaminase CRISPR base editor enables concurrent adenine and cytosine editing. *Nature Biotechnology*, 38 (7), 861–864. https://doi.org/10.1038/s41587-020-0535-y

Herrera-Estrella, L., Depicker, A., Van Montagu, M. & Schell, J. (1983). Expression of chimaeric genes transferred into plant cells using a Ti-plasmid-derived vector. *Nature*, 303 (5914), 209–213. https://doi.org/10.1038/303209a0

Hill, A.J., Peikin, S.R., Ryan, C.A. & Blundell, J.E. (1990). Oral administration of proteinase inhibitor II from potatoes reduces energy intake in man. *Physiology & Behavior*, 48 (2), 241–246. https://doi.org/10.1016/0031-9384(90)90307-P

Hirschberg, H.J.H.B., Simons, J.-Willem.F.A., Dekker, N. & Egmond, M.R. (2001). Cloning, expression, purification and characterization of patatin, a novel phospholipase A. *European Journal of Biochemistry*, 268 (19), 5037–5044. https://doi.org/10.1046/j.0014-2956.2001.02411.x

Hjorth, R. (1997). Expanded-bed adsorption in industrial bioprocessing: Recent developments. *Trends in Biotechnology*, 15 (6), 230–235. https://doi.org/10.1016/S0167-7799(97)01045-7

Hoekema, A., Hirsch, P.R., Hooykaas, P.J.J. & Schilperoort, R.A. (1983). A binary plant vector strategy based on separation of vir- and T-region

of the Agrobacterium tumefaciens Ti-plasmid. *Nature*, 303 (5913), 179–180. https://doi.org/10.1038/303179a0

Holkers, M., Maggio, I., Liu, J., Janssen, J.M., Miselli, F., Mussolino, C., Recchia, A., Cathomen, T. & Gonçalves, M.A.F.V. (2013). Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic Acids Research*, 41 (5), e63. https://doi.org/10.1093/nar/gks1446

Horn, A.H.C. & Sticht, H. (2015). Synthetic Protein Scaffolds Based on Peptide Motifs and Cognate Adaptor Domains for Improving Metabolic Productivity. *Frontiers in Bioengineering and Biotechnology*, 3. https://www.frontiersin.org/articles/10.3389/fbioe.2015.00191 [2024-02-16]

Hwang, S.-K., Singh, S., Cakir, B., Satoh, H. & Okita, T.W. (2016). The plastidial starch phosphorylase from rice endosperm: catalytic properties at low temperature. *Planta*, 243 (4), 999–1009. https://doi.org/10.1007/s00425-015-2461-7

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. & Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *Journal of Bacteriology*, 169 (12), 5429–5433

Jansen, R., Embden, J.D.A. van, Gaastra, W. & Schouls, L.M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*, 43 (6), 1565–1575. https://doi.org/10.1046/j.1365-2958.2002.02839.x

Jeon, S.D., Lee, J.E., Kim, S.J., Kim, S.W. & Han, S.O. (2012). Analysis of selective, high protein–protein binding interaction of cohesin–dockerin complex using biosensing methods. *Biosensors and Bioelectronics*, 35 (1), 382–389. https://doi.org/10.1016/j.bios.2012.03.023

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. & Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *SCIENCE*, 337 (6096), 816–821. https://doi.org/10.1126/science.1225829

Karges, K., Bellingrath-Kimura, S.D., Watson, C.A., Stoddard, F.L., Halwani, M. & Reckling, M. (2022). Agro-economic prospects for expanding soybean production beyond its current northerly limit in Europe. *European Journal of Agronomy*, 133, 126415. https://doi.org/10.1016/j.eja.2021.126415

Karkanis, A., Ntatsi, G., Lepse, L., Fernández, J.A., Vågen, I.M., Rewald, B., Alsiņa, I., Kronberga, A., Balliu, A., Olle, M., Bodner, G., Dubova, L., Rosa, E. & Savvas, D. (2018). Faba Bean Cultivation – Revealing Novel Managing Practices for More Sustainable and Competitive European Cropping Systems. *Frontiers in Plant Science*, 9, 375969. https://doi.org/10.3389/fpls.2018.01115

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. & Sternberg, M.J.E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, 10 (6), 845–858. https://doi.org/10.1038/nprot.2015.053

Kennedy, A.R. (1998). Chemopreventive Agents: Protease Inhibitors. *Pharmacology & Therapeutics*, 78 (3), 167–209. https://doi.org/10.1016/S0163-7258(98)00010-2

Kim, H., Kim, S.-T., Ryu, J., Kang, B.-C., Kim, J.-S. & Kim, S.-G. (2017). CRISPR/Cpf1-mediated DNA-free plant genome editing. *Nature Communications*, 8 (1), 14406. https://doi.org/10.1038/ncomms14406

Kim, Y.G., Cha, J. & Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences of the United States of America*, 93 (3), 1156–1160

Kolkman, J.M., Slabaugh, M.B., Bruniard, J.M., Berry, S., Bushman, B.S., Olungu, C., Maes, N., Abratti, G., Zambelli, A., Miller, J.F., Leon, A. & Knapp, S.J. (2004). Acetohydroxyacid synthase mutations conferring resistance to imidazolinone or sulfonylurea herbicides in sunflower. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 109 (6), 1147–1159. https://doi.org/10.1007/s00122-004-1716-7

Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A. & Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*, 533 (7603), 420–424. https://doi.org/10.1038/nature17946

Krüger, D.M., Rathi, P.C., Pfleger, C. & Gohlke, H. (2013). CNA web server: rigidity theory-based thermal unfolding simulations of proteins for linking structure, (thermo-)stability, and function. *Nucleic Acids Research*, 41 (W1), W340–W348. https://doi.org/10.1093/nar/gkt292

Lee, J.M., Hammarén, H.M., Savitski, M.M. & Baek, S.H. (2023). Control of protein stability by post-translational modifications. *Nature Communications*, 14 (1), 201. https://doi.org/10.1038/s41467-023-35795-8

Li, C., Zhang, R., Meng, X., Chen, S., Zong, Y., Lu, C., Qiu, J.-L., Chen, Y.-H., Li, J. & Gao, C. (2020). Targeted, random mutagenesis of plant genes with dual cytosine and adenine base editors. *Nature Biotechnology*, 38 (7), 875–882. https://doi.org/10.1038/s41587-019-0393-7

Li, J.-F., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M. & Sheen, J. (2013). Multiplex and homologous recombination–mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. *Nature Biotechnology*, 31 (8), 688–691. https://doi.org/10.1038/nbt.2654

Li, X., Wang, Y., Chen, S., Tian, H., Fu, D., Zhu, B., Luo, Y. & Zhu, H. (2018). Lycopene Is Enriched in Tomato Fruit by CRISPR/Cas9-Mediated Multiplex Genome Editing. *Frontiers in Plant Science*, 9, 332568. https://doi.org/10.3389/fpls.2018.00559

Liang, Z., Chen, K., Zhang, Y., Liu, J., Yin, K., Qiu, J.-L. & Gao, C. (2018). Genome editing of bread wheat using biolistic delivery of CRISPR/Cas9 in vitro transcripts or ribonucleoproteins. *Nature Protocols*, 13 (3), 413–430. https://doi.org/10.1038/nprot.2017.145

Lin, Q., Zong, Y., Xue, C., Wang, S., Jin, S., Zhu, Z., Wang, Y., Anzalone, A.V., Raguram, A., Doman, J.L., Liu, D.R. & Gao, C. (2020). Prime genome editing in rice and wheat. *Nature Biotechnology*, 38 (5), 582–585. https://doi.org/10.1038/s41587-020-0455-x

Løkra, S. & Strætkvern, K. (2009). Industrial Proteins from Potato Juice. A Review. *Food*, 3, 88–95

Lucas, M.M., Stoddard, F.L., Annicchiarico, P., Frias, J., Martinez-Villaluenga, C., Sussmann, D., Duranti, M., Seger, A., Zander, P.M. & Pueyo, J.J. (2015). The future of lupin as a protein crop in Europe. *Frontiers in Plant Science*, 6, 160197. https://doi.org/10.3389/fpls.2015.00705

Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J.M., Wolf, Y.I., Yakunin, A.F., van der Oost, J. & Koonin, E.V. (2011). Evolution and classification of the CRISPR–Cas systems. *Nature Reviews Microbiology*, 9 (6), 467–477. https://doi.org/10.1038/nrmicro2577

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. & Church, G.M. (2013). RNA-Guided Human Genome Engineering via Cas9. *Science (New York, N.Y.)*, 339 (6121), 823–826. https://doi.org/10.1126/science.1232033

Man, A.L., Purcell, P.C., Hannappel, U. & Halford, N.G. (1997). Potato SNF1-related protein kinase: molecular cloning, expression analysis

and peptide kinase activity measurements. *Plant Molecular Biology*, 34 (1), 31–43. https://doi.org/10.1023/A:1005765719873

Marraffini, L.A. & Sontheimer, E.J. (2008). CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA. *Science (New York, N.Y.)*, 322 (5909), 1843–1845. https://doi.org/10.1126/science.1165771

Mérida, A. & Fettke, J. (2021). Starch granule initiation in Arabidopsis thaliana chloroplasts. *The Plant Journal*, 107 (3), 688–697. https://doi.org/10.1111/tpj.15359

Modarres, H.P., Mofrad, M.R. & Sanati-Nezhad, A. (2016). Protein thermostability engineering. *RSC Advances*, 6 (116), 115252–115270. https://doi.org/10.1039/C6RA16992A

Mojica, F.J., Juez, G. & Rodríguez-Valera, F. (1993). Transcription at different salinities of Haloferax mediterranei sequences adjacent to partially modified PstI sites. *Molecular Microbiology*, 9 (3), 613–621. https://doi.org/10.1111/j.1365-2958.1993.tb01721.x

Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J. & Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution*, 60 (2), 174–182. https://doi.org/10.1007/s00239-004-0046-3

Møller, A.H., Hammershøj, M., Dos Passos, N.H.M., Tanambell, H., Stødkilde, L., Ambye-Jensen, M., Danielsen, M., Jensen, S.K. & Dalsgaard, T.K. (2021). Biorefinery of Green Biomass─How to Extract and Evaluate High Quality Leaf Protein for Food? *Journal of Agricultural and Food Chemistry*, 69 (48), 14341–14357. https://doi.org/10.1021/acs.jafc.1c04289

Moon, T.S., Dueber, J.E., Shiue, E. & Prather, K.L.J. (2010). Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered *E. coli*. *Metabolic Engineering*, 12 (3), 298–305. https://doi.org/10.1016/j.ymben.2010.01.003

Mussolino, C., Morbitzer, R., Lütge, F., Dannemann, N., Lahaye, T. & Cathomen, T. (2011). A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Research*, 39 (21), 9283–9293. https://doi.org/10.1093/nar/gkr597

Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J.D.G. & Kamoun, S. (2013). Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. *Nature Biotechnology*, 31 (8), 691–693. https://doi.org/10.1038/nbt.2655

Nicolia, A., Fält, A.-S., Hofvander, P. & Andersson, M. (2021). Protoplast-Based Method for Genome Editing in Tetraploid Potato. In: Tripodi,

P. (ed.) *Crop Breeding: Genetic Improvement Methods*. Springer US. 177–186. https://doi.org/10.1007/978-1-0716-1201-9_12

Pikaard, C.S., Brusca, J.S., Hannapel, D.J. & Park, W.D. (1987). The two classes of genes for the major potato tuber protein, patatin, are differentially expressed in tubers and roots. *Nucleic Acids Research*, 15 (5), 1979–1994. https://doi.org/10.1093/nar/15.5.1979

Pots, A.M., De Jongh, H.H.J., Gruppen, H., Hamer, R.J. & Voragen, A.G.J. (1998). Heat-induced conformational changes of patatin, the major potato tuber protein. *European Journal of Biochemistry*, 252 (1), 66–72. https://doi.org/10.1046/j.1432-1327.1998.2520066.x

Pourcel, C., Salvignol, G. & Vergnaud, G. (2005). CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology (Reading, England)*, 151 (Pt 3), 653–663. https://doi.org/10.1099/mic.0.27437-0

Pouvreau, L., Gruppen, H., Piersma, S.R., van den Broek, L.A.M., van Koningsveld, G.A. & Voragen, A.G.J. (2001). Relative Abundance and Inhibitory Distribution of Protease Inhibitors in Potato Juice from cv. Elkana. *Journal of Agricultural and Food Chemistry*, 49 (6), 2864–2874. https://doi.org/10.1021/jf010126v

Prat, S., Frommer, W.B., Höfgen, R., Keil, M., Kossmann, J., Köster-Töpfer, M., Liu, X.J., Müller, B., Peña-Cortés, H. & Rocha-Sosa, M. (1990). Gene expression during tuber development in potato plants. *FEBS letters*, 268 (2), 334–338. https://doi.org/10.1016/0014-5793(90)81281-r

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P. & Lim, W.A. (2013). Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell*, 152 (5), 1173–1183. https://doi.org/10.1016/j.cell.2013.02.022

Racusen, D. & Foote, M. (1980). A Major Soluble Glycoprotein of Potato Tubers1. *Journal of Food Biochemistry*, 4 (1), 43–52. https://doi.org/10.1111/j.1745-4514.1980.tb00876.x

Radestock, S. & Gohlke, H. (2008). Exploiting the Link between Protein Rigidity and Thermostability for Data-Driven Protein Engineering. *Engineering in Life Sciences*, 8 (5), 507–522. https://doi.org/10.1002/elsc.200800043

Ralet, M.-C. & Guéguen, J. (2001). Foaming Properties of Potato Raw Proteins and Isolated Fractions. *LWT - Food Science and Technology*, 34 (4), 266–269. https://doi.org/10.1006/fstl.2000.0684

Rouet, P., Smih, F. & Jasin, M. (1994). Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Molecular and Cellular Biology*, 14 (12), 8096–8106

Ryan, C.A. (1990). Protease Inhibitors in Plants: Genes for Improving Defenses Against Insects and Pathogens. *Annual Review of Phytopathology*, 28 (Volume 28, 1990), 425–449. https://doi.org/10.1146/annurev.py.28.090190.002233

Rydel, T.J., Williams, J.M., Krieger, E., Moshiri, F., Stallings, W.C., Brown, S.M., Pershing, J.C., Purcell, J.P. & Alibhai, M.F. (2003). The Crystal Structure, Mutagenesis, and Activity Studies Reveal that Patatin Is a Lipid Acyl Hydrolase with a Ser-Asp Catalytic Dyad. *Biochemistry*, 42, 6696–6708. https://doi.org/10.2210/pdb1oxw/pdb

Schupp, N. & Ziegler, P. (2004). The Relation of Starch Phosphorylases to Starch Metabolism in Wheat. *Plant and Cell Physiology*, 45 (10), 1471–1484. https://doi.org/10.1093/pcp/pch170

Senda, K., Yoshioka, H., Doke, N. & Kawakita, K. (1996). A Cytosolic Phospholipase A2 from Potato Tissues Appears to Be Patatin. *Plant and Cell Physiology*, 37 (3), 347–353. https://doi.org/10.1093/oxfordjournals.pcp.a028952

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., Zhang, K., Liu, J., Xi, J.J., Qiu, J.-L. & Gao, C. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. *Nature Biotechnology*, 31 (8), 686–688. https://doi.org/10.1038/nbt.2650

Singh, R., Langyan, S., Sangwan, S., Rohtagi, B., Khandelwal, A. & Shrivastava, M. (2022). Protein for Human Consumption From Oilseed Cakes: A Review. *Frontiers in Sustainable Food Systems*, 6, 856401. https://doi.org/10.3389/fsufs.2022.856401

Snell, P., Grimberg, Å., Carlsson, A.S. & Hofvander, P. (2019). WRINKLED1 Is Subject to Evolutionary Conserved Negative Autoregulation. *Frontiers in Plant Science*, 10, 387. https://doi.org/10.3389/fpls.2019.00387

Sonnewald, U., Basner, A., Greve, B. & Steup, M. (1995). A second L-type isozyme of potato glucan phosphorylase: cloning, antisense inhibition and expression analysis. *Plant Molecular Biology*, 27 (3), 567–576. https://doi.org/10.1007/BF00019322

Springmann, M., Clark, M., Mason-D'Croz, D., Wiebe, K., Bodirsky, B.L., Lassaletta, L., de Vries, W., Vermeulen, S.J., Herrero, M., Carlson, K.M., Jonell, M., Troell, M., DeClerck, F., Gordon, L.J., Zurayk, R., Scarborough, P., Rayner, M., Loken, B., Fanzo, J., Godfray, H.C.J., Tilman, D., Rockström, J. & Willett, W. (2018). Options for keeping

the food system within environmental limits. *Nature*, 562 (7728), 519–525. https://doi.org/10.1038/s41586-018-0594-0

Stoddard, B.L. (2014). Homing endonucleases from mobile group I introns: discovery to genome engineering. *Mobile DNA*, 5, 7. https://doi.org/10.1186/1759-8753-5-7

Strætkvern, K.O., Schwarz, J.G., Wiesenborn, D.P., Zafirakos, E. & Lihme, A. (1998). Expanded bed adsorption for recovery of patatin from crude potato juice. *Bioseparation*, 7 (6), 333–345. https://doi.org/10.1023/A:1008161618599

Strickland, J., Orr, G. & Walsh, T. (1995). Inhibition of Diabrotica Larval Growth by Patatin, the Lipid Acyl Hydrolase from Potato-Tubers. *Plant Physiology*, 109 (2), 667–674. https://doi.org/10.1104/pp.109.2.667

Svitashev, S., Schwartz, C., Lenderts, B., Young, J.K. & Mark Cigan, A. (2016). Genome editing in maize directed by CRISPR–Cas9 ribonucleoprotein complexes. *Nature Communications*, 7 (1), 13274. https://doi.org/10.1038/ncomms13274

Tan, J., Zhang, F., Karcher, D. & Bock, R. (2019). Engineering of high-precision base editors for site-specific single nucleotide replacement. *Nature Communications*, 10 (1), 439. https://doi.org/10.1038/s41467-018-08034-8

Tashkandi, M., Ali, Z., Aljedaani, F., Shami, A. & Mahfouz, M.M. (2018). Engineering resistance against Tomato yellow leaf curl virus via the CRISPR/Cas9 system in tomato. *Plant Signaling & Behavior*, 13 (10), e1525996. https://doi.org/10.1080/15592324.2018.1525996

Tompa, D.R., Gromiha, M.M. & Saraboji, K. (2016). Contribution of main chain and side chain atoms and their locations to the stability of thermophilic proteins. *Journal of Molecular Graphics & Modelling*, 64, 85–93. https://doi.org/10.1016/j.jmgm.2016.01.001

Tong, H., Liu, N., Wei, Y., Zhou, Y., Li, Y., Wu, D., Jin, M., Cui, S., Li, H., Li, G., Zhou, J., Yuan, Y., Zhang, H., Shi, L., Yao, X. & Yang, H. (2023a). Programmable deaminase-free base editors for G-to-Y conversion by engineered glycosylase. *National Science Review*, 10 (8), nwad143. https://doi.org/10.1093/nsr/nwad143

Tong, H., Wang, X., Liu, Y., Liu, N., Li, Y., Luo, J., Ma, Q., Wu, D., Li, J., Xu, C. & Yang, H. (2023b). Programmable A-to-Y base editing by fusing an adenine base editor with an N-methylpurine DNA glycosylase. *Nature Biotechnology*, 41 (8), 1080–1084. https://doi.org/10.1038/s41587-022-01595-6

Vennapusa, R., Hunegnaw, S.M., Cabrera, R.B. & Fernández-Lahore, M. (2008). Assessing adsorbent–biomass interactions during expanded

bed adsorption onto ion exchangers utilizing surface energetics. *Journal of Chromatography A*, 1181 (1), 9–20. https://doi.org/10.1016/j.chroma.2007.11.078

Wang, H., Liang, J., Chen, L., Deng, B., Gu, D., Liu, X., Jin, S., Xu, R., Qin, R., Zhu, Y., Zhao, L., Kou, D., Chen, Y., Jiang, Y., Li, J. & Wei, P. (2023). Development of plant cytosine base editors with the Cas12a system. *The Crop Journal*, 11 (5), 1451–1457. https://doi.org/10.1016/j.cj.2023.03.002

Wang, H., Wang, L., Zhong, B. & Dai, Z. (2022). Protein Splicing of Inteins: A Powerful Tool in Synthetic Biology. *Frontiers in Bioengineering and Biotechnology*, 10, 810180. https://doi.org/10.3389/fbioe.2022.810180

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C. & Qiu, J.-L. (2014). Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nature Biotechnology*, 32 (9), 947–951. https://doi.org/10.1038/nbt.2969

Wang, Y. & Yu, O. (2012). Synthetic scaffolds increased resveratrol biosynthesis in engineered yeast cells. *Journal of Biotechnology*, 157 (1), 258–260. https://doi.org/10.1016/j.jbiotec.2011.11.003

Wijeyesakere, S.J., Richardson, R.J. & Stuckey, J.A. (2014). Crystal Structure of Patatin-17 in Complex with Aged and Non-Aged Organophosphorus Compounds. *Plos One*, 9 (9), e108245. https://doi.org/10.1371/journal.pone.0108245

Wojnowska, I., Poznanski, S. & Bednarski, W. (2006). Processing of Potato Protein Concentrates and Their Properties. *Journal of Food Science*, 47, 167–172. https://doi.org/10.1111/j.1365-2621.1982.tb11051.x

Woo, J.W., Kim, J., Kwon, S.I., Corvalán, C., Cho, S.W., Kim, H., Kim, S.-G., Kim, S.-T., Choe, S. & Kim, J.-S. (2015). DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nature Biotechnology*, 33 (11), 1162–1164. https://doi.org/10.1038/nbt.3389

Yang, X., Zhang, L., Guo, X., Xu, J., Zhang, K., Yang, Y., Yang, Y., Jian, Y., Dong, D., Huang, S., Cheng, F. & Li, G. (2023). The gap-free potato genome assembly reveals large tandem gene clusters of agronomical importance in highly repeated genomic regions. *Molecular Plant*, 16 (2), 314–317. https://doi.org/10.1016/j.molp.2022.12.010

Zhao, D., Li, J., Li, S., Xin, X., Hu, M., Price, M.A., Rosser, S.J., Bi, C. & Zhang, X. (2021a). Glycosylase base editors enable C-to-A and C-to-G base changes. *Nature Biotechnology*, 39 (1), 35–40. https://doi.org/10.1038/s41587-020-0592-2

Zhao, X., Jayarathna, S., Turesson, H., Fält, A.-S., Nestor, G., González, M.N., Olsson, N., Beganovic, M., Hofvander, P., Andersson, R. & Andersson, M. (2021b). Amylose starch with no detectable branching developed through DNA-free CRISPR-Cas9 mediated mutagenesis of two starch branching enzymes in potato. *Scientific Reports*, 11 (1), 4311. https://doi.org/10.1038/s41598-021-83462-z

Zheng, C., Liang, S.-Q., Liu, B., Liu, P., Kwan, S.-Y., Wolfe, S.A. & Xue, W. (2022). A flexible split prime editor using truncated reverse transcriptase improves dual-AAV delivery in mouse liver. *Molecular Therapy*, 30 (3), 1343–1351. https://doi.org/10.1016/j.ymthe.2022.01.005

Zwijnenberg, H.J., Kemperman, A.J.B., Boerrigter, M.E., Lotz, M., Dijksterhuis, J.F., Poulsen, P.E. & Koops, G.-H. (2002). Native protein recovery from potato fruit juice by ultrafiltration. *Desalination*, 144 (1), 331–334. https://doi.org/10.1016/S0011-9164(02)00338-7

Popular science summary

To combat the effect of climate change, it is often said that it is necessary for humanity to replace part of their meat consumption with plant-based alternatives. This idea of a "protein shift" has led to several novel food products, such as textured pea-proteins. It has also increased interest in fringe protein crops, such as faba bean. Another interesting avenue is found where the field of plant based food overlaps with industrial waste products. While "Industrial waste" might not be the first thing we think of when talking about food. There are examples, such as sausage manufacturing, where side streams are already effectively utilized for food production.

In the case of plants, there has been research conducted on harvest waste, such as broccoli leafs and stem, and on side streams from industry, such as the seedcake left after rapeseed is pressed for oil. The potato starch industry has its own side stream of interest, the potato fruit juice. The potato fruit juice is comprised mainly of water, proteins and glycoalkaloids. Glycoalkaloids are toxic compounds, and thus the protein must be purified before it can be consumed by humans. One of the most abundant proteins in potato are the patatins. Being both highly nutritious, and a good foaming agent the patatins are often compared to eggs, a product that has been especially challenging to replace in plant based foods. Patatin is also largely flavourless, something that has been a challenge for products based on both beans and peas. Unfortunately, the patatins have proven prone to coagulation during processing, making them impossible to extract at high enough quality.

While breeding crops for processing traits is common practice in some crops, generation of new potato varieties through crosses is impractical. There are however good biotechnology applications available which can be used to achieve the desired traits. With the development of new CRISPR/Cas methods, researchers have gained the ability to make more precise edits than

ever before. This thesis explores the idea of using precision editing methods based on the CRISPR/Cas system to enhance protein properties in potato tubers. We studied two potential targets, the patatins that were mentioned above, and another protein found in potato tubers called Pho1a, which is thought to be involved in starch granule formation. The analysis of patatin showed that modification is possible, and several patatins can be modified at once. The patatins are a large group of proteins that all come from different genes. It was also discovered that these genes are not equally represented on the protein level, but rather a few genes were responsible for producing a majority of the protein. These high efficiency genes would be a great starting point for future modifications.

The second target investigated was Pho1a. When working with genes, it is often easier to remove something than to add it. This is why we started by producing Pho1a deficient potatoes. These produced small malformed starch granules, suggesting that elimination of the protein is not desirable, and that this could be a good target for the harder to achieve enhancing modifications.

Populärvetenskaplig sammanfattning

Ett byte till en mer växtbaserad diet är ett vanligt förekommande förslag på vad vi kan göra för att motverka klimatförändringar. Det här ”protein skiftet” har lett till ett uppsving i nya vegetariska produkter, såsom texturerat ärtprotein. Det har också ökat intresset för nya proteingrödor, t.ex. bondbönor.

Ytterligare ett fält som utforskats är hur vi bättre kan använda biprodukter från etablerade industrier. Inom växtbiologin gäller det både skördeavfall, så som blad och stam från broccoli som lämnas i fältet, men också rester från industrin som t.ex. frökakan som blir kvar efter att raps pressas för olja. Produktionen av potatisstärkelse har sin egen biprodukt som kan vara av intresse, potatissaft. Potatissaften består huvudsakligen av protein, vatten och glykoalkaloider. Eftersom glykoalkaloider är giftiga så måste proteinet renas ytterligare innan det kan konsumeras av människor.

Ett av de rikligaste proteinerna i potatis knölar är patatin. Patatin har ett mycket högt näringsvärde, och en god förmåga att stabilisera skum. Två egenskaper som gjort att det ofta jämförs med ägg, en produkt som varit särskilt knepig att ersätta i växtbaserad mat. Patatin saknar dessutom smak, till skillnad från liknande produkter baserade på ärtor och bönor. Tyvärr har framställning av patatin visat sig problematiskt eftersom patatinets låga stabilitet gör att det ofta koagulerar och då tappar många av sina goda egenskaper.

Att förädla växter för önskade egenskaper är vanligt förekommande, men korsningar av potatis är mycket ineffektiva. Det finns däremot mycket effektiva bioteknologiska metoder för att uppnå samma resultat. Bland annat har pågående utvecklingar av CRISPR/Cas-baserade metoder givit forskare mer kontroll över vilken typ av ändringar de kan göra i genomet. I den här avhandlingen utforskar jag möjligheten att använda CRISPR/Cas-baserade

metoder för att förbättra potatis proteiners egenskaper. Två kandidater undersöks, patatin vilka diskuterades tidigare, och Pho1a, ett protein som tros vara inblandat i tidig utveckling av stärkelsegranuler. Eftersom patatin är en stor grupp proteiner, gjordes en analys av hur många av dem som kan modifieras i ett svep. Vi såg att några få gener producerar en majoritet av proteinerna, vilket betyder att en stor effekt kan nås genom att fokusera på att modifiera dessa effektiva gener snarare än att försöka träffa allihopa.

Den andra kandidaten som analyserades var Pho1a. För att få en bättre insikt i vilken betydelse Pho1a har i potatis producerades plantor utan Pho1a. Vi såg att dessa potatisar producerade stärkelsegranuler som var små och hade en konstig form. Detta antyder att Pho1a är av stor betydelse för stärkelsegranulernas naturliga utveckling, och att ökad Pho1a aktivitet skulle kunna leda till en boost i deras tidiga utveckling. En egenskap som skulle vara intressant att undersöka vidare i framtiden.

Acknowledgements

There are so many people without whose help this thesis could never have come into being. My supervisors, Per Hofvander, Mariette Andersson and Folke Sitbon. Thank you for your patience with me and for your encouragement when things have not gone to plan. I sometimes wonder if you were to kind to me, and sometimes if I could have ever made it if you were not. Thank you for the guidance and for allowing me to take on this project with you!

To my family, my parents and my sister, for keeping in touch when I was far away and could not see them as often as I wished. For showing interest in my passion, without being overly nosey when I did not want to talk about it. Thank you, your encouragement has meant the world to me.

To Ying, for always offering help when you had some extra time. For your generous distribution of snacks. And for listening to me when I needed to complain.

To Hannah, Johanna and Oliver, the custodians of the coffee machine, thank you for the laughs, for enduring my sometimes unending ability to talk, and for starting a fresh pot when I was late to the office.

To Paul, the valiant master student. Thank you for taking on the adapter project with me, and for always being eager to learn and discover new techniques. You were a joy to work with, and I hope you learned as much during your time with us at SLU as I did.

To Shrikant, thank you for helping me through two of my first manuscripts. Thank you for dealing with the mess that is patatin genomics, and for confirming that it is indeed as complicated as I feared.

To Helle, Niklas and Matías, for generously sharing your knowledge of all things practical. For helping me with both design and execution of a wide variety of different experiments. And for always showing kindness.

To Ann-Sofie, for teaching me the art of protoplast isolation, and the craft that is tissue culture. Thank you for all the jokes and laughs, and for the occasional potato cake.

To Mirella, Marina, and Josefine. Thank you for keeping the sterile lab in order, and for lending a hand when I had too many things going on. I could not start to guess how many shoots you have saved for mould or starvation.

To the “Lemming herd” back in Uppsala. Thank you all, for staying in touch, for offering up your homes so that Cissi and me had somewhere to stay when we visited. Thank you for inviting us to all the things, even when you knew we could not attend, just to show that you cared. Thank you for all your encouragement, the support and the good-natured teasing. I wish everyone could have friends as good as you!

To the open source community, including but not limited to the developers of gnuplot, Notepad++. GIMP, PyMol, RosettaFold, Zotero, YASARA, TSA_CRAFT. You made my work easier and more enjoyable. I have cited you where I can, but that is not even half the recognition you deserve.

Finally, in the place of honour, Cecilia. Thank you for coming with me on this adventure, for encouraging me the whole way thorough. For taking care of me when I needed it. And for all the hundreds of tiny things you do, that makes my life a joy. I love you.

III

frontiers | Frontiers in Plant Science

TYPE Hypothesis and Theory
PUBLISHED 09 August 2023
DOI 10.3389/fpls.2023.1220973

OPEN ACCESS

EDITED BY
Angel Merida,
Spanish National Research Council (CSIC), Spain

REVIEWED BY
Joerg Fettke,
University of Potsdam, Germany
Ermelinda Botticella,
National Research Council (CNR), Italy

*CORRESPONDENCE
Shrikant Sharma
shrikant.sharma@slu.se
Per Hofvander
per.hofvander@slu.se

RECEIVED 11 May 2023
ACCEPTED 18 July 2023
PUBLISHED 09 August 2023

CITATION
Sharma S, Friberg M, Vogel P, Turesson H, Olsson N, Andersson M and Hofvander P (2023) *Pho1a* (plastid starch phosphorylase) is duplicated and essential for normal starch granule phenotype in tubers of *Solanum tuberosum* L.
Front. Plant Sci. 14:1220973.
doi: 10.3389/fpls.2023.1220973

Pho1a (plastid starch phosphorylase) is duplicated and essential for normal starch granule phenotype in tubers of *Solanum tuberosum* L.

Shrikant Sharma*, Martin Friberg, Paul Vogel, Helle Turesson, Niklas Olsson, Mariette Andersson and Per Hofvander*

Department of Plant Breeding, Swedish University of Agricultural Sciences, Alnarp, Sweden

Reserve starch from seeds and tubers is a crucial plant product for human survival. Much research has been devoted to quantitative and qualitative aspects of starch synthesis and its relation to abiotic factors of importance in agriculture. Certain aspects of genetic factors and enzymes influencing carbon assimilation into starch granules remain elusive after many decades of research. Starch phosphorylase (Pho) can operate, depending on metabolic conditions, in a synthetic and degradative pathway. The plastidial form of the enzyme is one of the most highly expressed genes in potato tubers, and the encoded product is imported into starch-synthesizing amyloplasts. We identified that the genomic locus of a Pho1a-type starch phosphorylase is duplicated in potato. Our study further shows that the enzyme is of importance for a normal starch granule phenotype in tubers. Null mutants created by genome editing display rounded starch granules in an increased number that contained a reduced ratio of apparent amylose in the starch.

KEYWORDS

starch phosphorylase, reserve starch, sink tissue, starch granule, tuber, potato, CRISPR, gene editing

1 Introduction

Plant α-glucan phosphorylase or starch phosphorylase is a somewhat enigmatic enzyme where both cytosolic and plastid forms can be found that are encoded by independent genes (Mori et al., 1991; Sonnewald et al., 1995; Albrecht et al., 2001). The role of the enzyme forms has been investigated in several plant species over a long time with different findings regarding a role in biosynthetic or degradative pathways and importance of spatio/temporal or abiotic conditions that have been comprehensively reviewed (Rathore

et al., 2009; Hwang et al., 2020; Li et al., 2021; Shoaib et al., 2021). The plastid directed *Pho1a* is one of the highest expressed genes in potato tuber (Sonnewald et al., 1995; Albrecht et al., 2001; Van Harsselaar et al., 2017), but its precise role regarding reserve starch accumulation, if any, remains elusive.

Starch phosphorylase (EC 2.4.1.1) is a member of the GT35-gycosyl transferase superfamily and carries out a transfer of glucosyl unit to or from an α-1,4-glucan chain in a reversible reaction (Hanes, 1940; Brisson et al., 1989; Rathore et al., 2009; Cuesta-Seijo et al., 2017). The release or incorporation of glucose-1-phosphate (G-1-P) is suggested to be dependent on multiple factors including the ratio of G-1-P to inorganic phosphate (Pi) (Preiss and Levi, 1980; Kruger and ap Rees, 1983; Schupp and Ziegler, 2004; Satoh et al., 2008; Rathore et al., 2009; Tiessen et al., 2012; Tetlow and Bertoft, 2020). Two forms of starch phosphorylase are reported in higher plants, localized to plastid (Pho1: L type) and cytosol, respectively (Pho2: H type) (Brisson et al., 1989; Sonnewald et al., 1995; Albrecht et al., 2001). The cellular compartmentalization of respective isozymes subjects them to different metabolic effectors, redox environments, and protein turnover factors (Albrecht et al., 1998; Albrecht et al., 2001; Hwang et al., 2020). In addition to a peptide for plastid localization, a major difference in Pho1 from Pho2 is an internal, approximately 78–82 amino acid domain (L80). This domain forms an extended auxiliary loop of unordered structure and is suggested to define substrate specificity based on stereological hindrance in binding to large polysaccharides (Nakano and Fukui, 1986; Albrecht et al., 1998; Chen et al., 2002; Young et al., 2006; Tickle et al., 2009; Hwang et al., 2016b; Cuesta-Seijo et al., 2017; Nakamura et al., 2017). The L80 domain was found not to be of any importance for catalytic functions of Pho1 in rice (Hwang et al., 2010). However, it contains a highly variable set of negatively charged amino acids, phosphorylation sites, and a PEST motif, which is reported to be a substrate of proteasomes to modulate the degradation of Pho1 in sweet potato (Chen et al., 2002; Young et al., 2006; Lin et al., 2012). The exact role of Pho isozymes in starch metabolism in higher plants has been debated over decades. Although it is generally accepted that the plastidial form of Pho1 is involved in maltooligosaccharides (MOs) metabolism, cytosolic Pho2 is generally involved in maltose metabolism resulting from starch degradation (Lu et al., 2006; Satoh et al., 2008; Hwang et al., 2010; Flores-Castellanos and Fettke, 2022).

Various studies have indicated that Pho1 is involved in transitory starch turnover in photosynthetic and reserve starch accumulation in sink organs of multiple species (cereal grains, roots, and tubers) by maintenance of plastidial maltodextrin pools (Chen et al., 2002; Schupp and Ziegler, 2004; Young et al., 2006; Subasinghe, 2013; Hwang et al., 2016b; Nakamura et al., 2017). In *Arabidopsis*, *PHS1*, a homolog of *Pho1* in *Arabidopsis*, is suggested to be part of the core set of evolutionary conserved genes, involved in starch granule initiation (Mérida and Fettke, 2021). Pho1 is also reported to form complexes with multiple starch synthases (SS) and starch-branching enzymes (SBEs) in several cereals including wheat, rice, maize, and barley, and has been shown to influence the starch synthesis (Tetlow et al., 2008; Liu et al., 2009; Nakamura et al., 2012; Ahmed et al., 2015; Crofts et al., 2017; Nakamura et al., 2017). Other complexes that have been reported are with disproportionating enzyme 1 (Dpe1, EC 2.4.1.25) in sweet potato and rice (Hwang et al., 2016a; Nakamura et al., 2017). The enzyme is suggested to serve in recycling of MOs released from the trimming of pre-amylopectin by debranching enzymes (DBEs) leading to the accumulation of G-1-P (Satoh et al., 2008; Hwang et al., 2010; Hwang et al., 2016a; Lin et al., 2017). The direct incorporation of G-1-P to the surface of native starch granules and soluble MOs in maltodextrin pool by Pho1 has been demonstrated in *in vitro* assays and with potato tuber disks (Fettke et al., 2010; Fettke et al., 2012; Flores-Castellanos and Fettke, 2022). Reserve starch accumulation in potato tubers has been suggested to follow two interacting pathways depending on the environmental conditions (Fettke et al., 2012). The high catalytic activity of Pho1 at lower temperatures as compared to ADP-glucose pyrophosphorylase (AGPase) has then been suggested to provide a complimentary pathway for starch biosynthesis at low temperatures in potato (Fettke et al., 2012; Slugina et al., 2020a; Mérida and Fettke, 2021).

Two isoforms of Pho1 are reported in potato, Pho1a and Pho1b (Brisson et al., 1989; Nakano et al., 1989; Sonnewald et al., 1995; Albrecht et al., 2001). The genes encoding the isoforms are suggested to be located on chromosome 3 and 5, respectively (Schreiber et al., 2014; Schönhals et al., 2016; Van Harsselaar et al., 2017; Slugina et al., 2020b) and localized on respective chromosomes in latest genome assemblies (Pham et al., 2020; Jayakody et al., 2023; Yang et al., 2023). The corresponding mRNAs are assembled from 15 and 14 exons, respectively, and an insertion of a 5,060-bp-long TE/Copia-like retrotransposon (*Tst1*) has been reported in the fifth intron of *Pho1a* (CamIrand et al., 1990). Both proteins are highly similar (81%–84%) in amino acid sequence identity, while the N-terminal transit peptide and L80 insertion domain are more diverse (Albrecht et al., 2001). Pho1a is reported to be ubiquitously present in both leaves and tubers, whereas the gene encoding Pho1b was mostly expressed in leaves and close vicinity to vascular tissue in tubers (Albrecht et al., 2001). Although expression of *Pho1b* is found to be higher in leaves and clearly detectable in tubers, at protein level, Pho1a is reported to be more abundant in leaves and tubers, whereas Pho1b was undetectable in tubers (Sonnewald et al., 1995; Albrecht et al., 1998; Albrecht et al., 2001). Pho1a has been detected as homo-$(Pho1a)_2$ and hetero-dimer (Pho1a–Pho1b) in leaves, but only as homodimer in tubers (Albrecht et al., 1998; Albrecht et al., 2001).

In this study, we show that the *Pho1a* gene is tandemly duplicated in the potato genome. We show that the transposon *Tst1* likely is not the driver of this duplication, as species in *Solanaceae* with duplication and no *Tst1* insertion exists. Furthermore, we show that knocking out the *Pho1a* genes, using genome editing, affects the starch granule phenotype in reserve starch of potato tubers. A decrease in measured amylose to amylopectin ratio was observed that could indicate a changed chain length distribution affecting the observed granule

phenotypical change. This shows that the major plastid starch phosphorylase activity of potato has a role in the organization of reserve starch structure but not for general starch accumulation capacity. No major differences in starch degradation into sugars could be found between null mutants and the parental variety upon cold storage of tubers.

2 Materials and methods

2.1 Plastid starch phosphorylase (*Pho1a*) mutagenesis

To determine a CRISPR target sequence, a partial genomic sequence of the plastidial starch phosphorylase was extracted from *Solanum tuberosum* cv. Desirée using primers listed in Table 1 and with a method previously described (Zhao et al., 2021). The target sequence to induce mutations in the *Pho1a* gene (Figure 1A) was selected using the CRISPR RGEN Tools (http://www.rgenome.net/cas-designer). The selected target, 5′-GCTGTTGCAAAGAATGCCTT-3′, was located on the negative strand on exon 2, adjacent the PAM site 5′-AGG-3′ (Figure 1B). Leaf tissue from *in vitro* grown plants of *S. tuberosum* L. cv Desirée was used for protoplast isolation, transfection, and shoot regeneration as described previously (Nicolia et al., 2021). Purified protoplasts were transfected with preassembled ribonucleoprotein complexes (RNPs hereafter) of 5 µg of Cas9 enzyme (Thermo Fisher Scientific, Waltham, MA, USA) and 0.1 nmol sgRNA, (Synthego, Redwood city, CA, USA), using 40% PEG 4000 (Sigma-Aldrich, Germany) with an incubation time of 30 min. Regenerated shoot isolation was limited to one from each callus for further analysis.

A primary analysis for screening of induced indels was made using high-resolution fragment analysis (HRFA hereafter) of PCR amplicons spanning the target site as described previously (Andersson et al., 2017) using primers listed in Table 1 and marked in Figure 1B. The mutations were confirmed by Sanger sequencing of PCR amplicons using primers listed in Table 1 and marked in Figure 1B, and indel distribution was analyzed using online ICE analysis (http://ice.synthego.com) (Figure 2). *In vitro* cuttings of selected events with indels (Table 2) and Desirée (as control) were planted in soil (Yrkesplantjord, SW Horto, Hammenhög, Sweden) as three to five biological replicates in 7.5-L pots. Plants were cultivated under controlled greenhouse conditions (16-h day length, 18/15°C day/night temperature, supplementary light intensity up to approximately 200 µmol s^{-1} m^{-2} photons, 50% relative humidity) for 5 months and were regularly fertilized with SW Bouyant RikaS 7-1-5 + mikro (SWHorto, Hammenhög, Sweden).

TABLE 1 List of primers used in this study for HFRA analysis, Sanger sequencing, and RT-qPCR based copy number estimation.

Application	Primer code	Amplicon	Primer sequence (5′–3′)	T_A (°C)	Amplicon Size (bp)
HFRA Analysis	HFRA_Fw	Pho1a_E2	AGAGCGACCTGAGTTCTTTT-FAM	62	209
	HFRA_Rw		GTACGCTTGCTTCATGTTCA		
Sanger Sequencing and ICE analysis	SangerSeq_Fw	Pho1a_E2-E3	CAGAAACTTGATGTATGGATCTTAGG	62	641
	SangerSeq_Rv		GCACCAGTAAGCTCCAGATT		
RT-qPCR based copy number estimation and *GBSS1* expression	qStgbssF	*StGBSS_qRT*	TTGCATAACTGGGATTGTGAATG	52	93
	qStgbssR		GACAGTGGTTATATCGTATTTGACATCTG		
	RT-qPCR_Fw	*StPho1a_qRT*	CCATGCAGAATTCACACCTG	52	95
	RT-qPCR_Rv		TAAGGAGCGAATCACGAACA		
	qSttubF	*StTUBB1_qRT*	GTTGGCAATTCAACCTCCAT	60	143
	qSttubR		ATGTTGCTCTCGGCTTCAGT		
Targetted Cloning and sequencing	PHO1.1a_F1	Amplicon 1	GTTTTAATTTGCGAGAGAGAGAGAG	62	601
	PHO1.1a_R1		TCTGTGAATGCCATGTCAGC		
	PHO1.1a_F2	Amplicon 2	TAGGGAGATGGTCACTGTTCCAG	62	904
	PHO1.1a_R2		TCACATCCCTTCACTTGTTCCTG		
	PHO1.1a_F2	Amplicon 3	TAGGGAGATGGTCACTGTTCCAG	62	786
	PHO1.2a_R2		GTGCTAAGACAAGAAGGAAGGTG		
	PHO1.2a_F1	Amplicon 4	ACCACATAATAAGAGATGAAGAGTCTC	62	895
	PHO1.2a_R1		TCAAATAGCCTCGCACTTACTC		
	PHO1.2a_F2	Amplicon 5	GAAGCTCATCCAAGATGACTATCTG	62	807
	PHO1.2a_R2		GTGCTAAGACAAGAAGGAAGGTG		

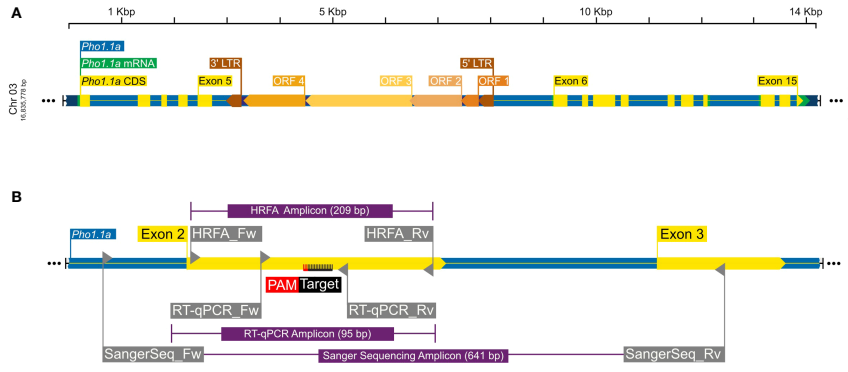

FIGURE 1
A graphical representation of *Pho1.1a* genomic loci. **(A)** Full-length *Pho1.1a* loci from *S. tuberosum* Group Phureja DM 1-3 516 R44 Genome Assembly (DM v6.1). **(B)** Enlarged view of *Pho1.1a* sequence representing location of Exon 2/3, RNP target adjacent to PAM sequence on negative strand, primers and respective amplicons for HFRA, Sanger sequencing, and RT-qPCR analysis. The sequences are color coded as follows: genomic as blue, mRNA as green, CDS/Exon as yellow, 5′ and 3′ LTRs as brown, ORFs of *Tst1* as shades of dark yellow, primers as gray, amplicons as purple, and RNP target/PAM site as black/red. The top scale represents length of *Pho1.1a* loci, and left/right numbers represent location of chromosome 3 of DM v6.1.

2.2 Characterization of *Pho1a* genomic loci

For the *in silico* characterization of the genomic *Pho1a* loci, a reference mRNA sequence of *Pho1a* (GenBank accession no. X52385.1) was blasted against high confidence gene models of long-read chromosome-scale genome assembly of doubled monoploid potato *S. tuberosum* Group Phureja DM 1-3 516 R44 Genome Assembly v6.1 (DM v6.1 hereafter) using blastn with

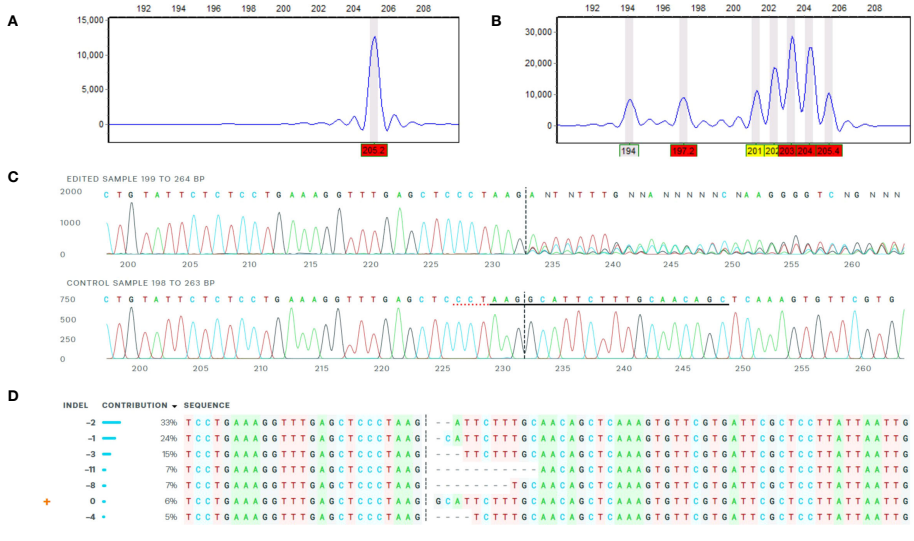

FIGURE 2
Example of confirmation of mutations by HFRA and ICE analysis. The figure represents HFRA and ICE analysis for confirmation of mutations in one representative mutational line, SPD-15 (1WT). **(A)** One peak detected in HFRA chromatogram representing the wild-type alleles in Desirée (control). **(B)** Seven peaks detected in HFRA chromatogram representing indels of seven different sizes in SPD-15. **(C)** Sanger sequencing trace data: top, SPD-15 and bottom, Desirée. The vertical dashed line represents Cas9 blunt cut, three nucleotides upstream from PAM site on negative strand. Red dashed line represents PAM site (AGG) on negative strand. Solid black line marks sgRNA guide (5′-GCTGTTGCAAAGAATGCCTT-3′) on negative strand. **(D)** Deduced sequence of six mutational and one wild-type alleles in SPD-15 and their relative abundance based on Sanger sequencing data. The calculated contribution indicate that two alleles have the same type of indel, either a −2 or a −1 deletion.

TABLE 2 Size of indels (bp) and mutational outcome in respective regenerated events.

Group	Line	Event ID	Size of indels (bp)	Mutational outcome
Group 1 (Full knockouts; FKO)	SPD-1	181057	-4;-2;-1	full knock, all alleles out of frame mutations
	SPD-2	181085	-5;-2;-1	
	SPD-3	181092	-2;-1	
	SPD-4	181113	-2;-1;1	
	SPD-5	181124	-10;-4;-2;-1	
Group 2 (In-Frame knockouts; IFM)	SPD-6	181004	-6;-5;-4;-2;-1	full knock, at least one allele in frame
	SPD-7	181011	-5;-3;-2;-1	
	SPD-8	181015	-6;-2;-1;	
	SPD-9	181027	-9;-5;-2;-1	
	SPD-10	181116	-8;-4;-3;-2;-1;1	
Group 3 (Partial knockouts; WTA)	SPD-11	181013	-16;-12;-2;0	events with at least one wild type allele
	SPD-12	181017	-9;-4;-2;-1;0	
	SPD-14	181065	-36;-12;-2;-1;0	
	SPD-15	181130	-11;-8;-4;-3;-2;-1;0	

"0" represents at least one WT allele present, and "–X" and "X" represents a deletion or insertion where X is the number of bp. The events have been sorted in three groups: full knockout and out of frame indels, all alleles with mutations but at least one allele has an in-frame indel, and events with mutations but at least one allele is wild type.

default parameters (http://spuddb.uga.edu/blast.shtml) (Pham et al., 2020). Top hits above 99% identity and E-value 0.0 on chromosome 3 were identified as putative *Pho1a* loci in DM v6.1 assembly.

These putative *Pho1a* genomic sequences were aligned with previously reported *Pho1a* transcripts based on PGSC v4.03 assembly (Sharma et al., 2013; Van Harsselaar et al., 2017), *Pho1a* genomic sequence extracted from Desirée genomic DNA library (Zhao et al., 2021), and copia-like transposable element *Tst1* (CamIrand et al., 1990) using Clustal Omega (v1.2.2) at EMBL-EBI (https://www.ebi.ac.uk/Tools/msa/clustalo/) to refine exon/intron structure and deduce full length genomic sequence of *Pho1a* loci. Additionally, genomic sequences of homologous *Pho1a* loci from 24 wild potato species, representative of tuber-bearing clade of *Petota* section, two non-tuber-bearing wild potato species from the neighboring section *Etuberosum*, and two landraces of *S. tuberosum* Group Phureja (http://solomics.agis.org.cn/potato/tool/blast)(Priyam et al., 2019; Tang et al., 2022) and *S. lycopersicum* (http://spuddb.uga.edu/SollycM82_v1_download.shtml; (Alonge et al., 2022) were retrieved by blast search using reference mRNA sequence of *Pho1a*. These putative *Pho1a* genomic sequences from above assemblies were also aligned with above-mentioned complete *Pho1a* genomic sequence using Clustal Omega (v1.2.2) to detect the presence of duplication and insertion of *Tst1* in respective species. Specific genomic regions were amplified by PCR using primers listed in Table 1 and subsequently cloned in pJET1.2 using Clone Jet PCR Cloning kit (Thermo Fisher Scientific, Waltham, MA, USA). The resulting clones with inserts were subjected to Sanger sequencing (LGC Genomics, Berlin, Germany) in triplicates to confirm organization of duplicated *Pho1a* loci on chromosome 3 of *S. tuberosum* L. cv Desirée.

The duplication of *Pho1a* in cv. Desirée based on HRFA and ICE analysis was examined by copy number estimation by Real Time quantitative PCR (RT-qPCR) on QuantStudio3 thermocycler (Applied Biosystems, USA) using Maxima SYBR Green/ROX qPCR Master Mix (2X) (K0221, Thermo Fisher Scientific, Waltham, MA, USA). *StGBSS1* (GenBank accession no. A23741.1), a known single-copy gene, was used as reference gene (Andersson et al., 2017). The primer pairs for *Pho1a* were selected to match *StGBSS1* primers in terms of amplicon length and T_m (Table 1) (Andersson et al., 2006). The optimal T_m was determined based on R^2 closest to =1 from the standard curve of Desirée gDNA dilution series in duplicates (10, 1, 0.1, 0.01 and 0.001 ng) using VeriFlex (Applied Biosystems, USA), and the initial copy number in 10 ng gDNA sample was calculated using following formula:

$$X_n = X_0(1+E)_n$$

where X_n is PCR product after cycle n, X_0 is initial copy number, E is amplification efficiency, and n is cycle number.

2.3 Phenotypic characterization

Plants were photographed at three time points during greenhouse cultivation (4 weeks, 12 weeks, and maturity) to compare growth rate characteristics and individual leaf samples from greenhouse cultivated events were collected (middle and end of light phase) to evaluate transitory starch characteristics. Briefly, the leaf samples were fixed by immersion into fixative solution (3.7% formaldehyde and 0.1 M phosphate buffer, pH 6.5) for 24 h at 37°C, dehydrated and decolored (50% (v/v) ethanol for 24 h, and 96% (v/v) ethanol for 2 × 24 h; both steps at 37°C) and stored at 4°C.

Afterwards, the leaf samples were rehydrated (50% (v/v) ethanol for 30–60 min; dH_2O for 20–30 min), stained with Lugol's solution (2% KI (w/v) and 1% I_2 (w/v) for 3 min), and visualized under a light microscope (Leica DMLB, Wetzlar, Germany) equipped with an Infinity X-32 digital camera (DeltaPix, Samourn, Denmark) as described previously (Ovecka et al., 2012).

The number, shape, and total fresh weight of the tubers were recorded at the time of harvest. A subset of freshly harvested tubers from each mutational line was sliced into halves. One was immediately flash frozen in liquid nitrogen and stored at −80°C for subsequent initial free sugar content analysis. The other was freeze-dried for 48 h and utilized to determine dry matter, total starch, and starch composition. The weight of freeze-dried tuber samples was recorded, and dry matter was calculated as described previously (Zhao et al., 2021). In addition, the freshly harvested tubers were also cross-sectioned along vertical and horizontal axis, stained with Lugol's solution for 1 min, washed with dH_2O, and photographed (Cannon 450D DSLR) on a light table to record distribution of starch granules across tuber axis(s) in triplicates. These tuber samples were further sliced into 0.5-mm-thin sections using a mandolin and visualized under a light microscope after staining with Lugol's solution as previously described (Zhao et al., 2021). The size of starch granules were graded into three groups: a) small up to 25 µm, b) medium 25–50 µm, and c) large above 50 µm. The remaining tubers from each event were stored at 4°C for 3 months and used to determine free sugar levels in tubers after cold storage.

2.4 Starch phosphorylase activity staining

Tuber extracts from a subset of mutational events (based on above phenotypic characterization) and Desirée (control) were assessed for starch phosphorylase activity using a modified version of (Sweetloove et al., 1996). In total, nine mutational events and Desirée (as control) were selected for further analysis. Tuber samples from three biological replicates of each individual line were pooled and ground using a mortar and pestle in extraction buffer (100 mM HEPES, 10 mM EDTA, 5 mM DTT, 10% Glycerol, 0.1% PVPP, pH 7.5) and 1× Protease inhibitor cocktail for plant cell and tissue extracts (Sigma-Aldrich, USA), on ice. Extracts were centrifuged for 10 min at 14,000 *g*, at 4°C. Supernatant was loaded directly on a native PAGE gel (Novex, 10% Tris-Glycine, Invitrogen or 12% native PAGE cast with 0.8% glycogen). Novex gel was run for 120 min at 200 V at 4°C. The glycogen containing gel was run for 120 min at 250 V at 4°C. Subsequently, both gels were washed in an incubation buffer (100 mM Tris–HCl, 1 mM $MgCl_2$, 1 mM $CaCl_2$, 5% glycerol, pH 7) for 10 min before being moved to a substrate buffer (incubation buffer supplemented with 50 mM glucose-1-phosphate, 2.5 mM AMP) and incubated at room temperature (approximately 22°C) for 3 h. After incubation, the gel was briefly rinsed in dH_2O before staining for 5–10 min in 0.1× Lugol's solution and imaged using a photocopier (Epson Perfection V750 Pro, Seiko Epson Corporation, Suwa, Japan).

2.5 Determination of total starch content

Freeze-dried samples from five full knockout events were homogenized in a Retsch Mixer Mill MM400 (Retsch, Germany), at 30 Hz, for 30 s, and 50 mg of homogenized material was utilized to determine total starch content using Total Starch Assay Kit (K-TSTA-100A, Megazyme, Wicklow, Ireland) following the manufacturer's instructions. Briefly, total starch in the samples was converted into maltodextrins by thermostable α-amylase (100°C, 15 min), which were subsequently quantitatively hydrolyzed into D-glucose by amyloglucosidase (50°C, 30 min). The resulting D-glucose was measured in a colorimetric reaction employing glucose oxidase/peroxidase (GOPOD) reagent, and absorbance was measured at 510 nm (Multiskan GO, Thermo Fisher Scientific, USA). All analysis were performed in triplicates; total starch concentration was calculated using Mega-Calc^{TM} (Megazyme, Ireland) and reported as mean values of percentage on dry weight (DW) basis.

2.6 Determination of starch composition and *StGBSS1* expression

Amylose content was measured in starch from tubers of nine mutational events and Desirée (as control) by a colorimetric method, as described previously (Chrastil, 1987). In short, 150 mg of homogenized material from freeze-dried tuber samples was suspended in 70% EtOH. The suspension was sieved through a nylon mesh to remove debris, and total starch was pelleted by centrifugation (2,000 *g*, 20 min). The supernatant was discarded, and the pellet was dried overnight. For each sample, 8 mg of purified starch was used for analysis. Samples were suspended in water, solubilized with 5M NaOH, and incubated at room temperature with intense agitation for 1.5 h. After solubilization, samples were neutralized with 3M HCl and then buffered by addition of 50mM sodium phosphate buffer. Samples were stained using iodine dissolved in 85% Dimethyl sulfoxide (DMSO), incubated for 10 min at room temperature, and the absorbance was measured at 620 and 550 nm (Multiskan GO, Thermo Fisher Scientific, USA). Amylose content was calculated as percent of total starch content using the following formula.

$$190.2R^2 - (281.52R) + 106.6$$

where $R = A_{620nm}/A_{550nm}$ The expression level of *StGBSS1* was estimated in tuber samples of selected events by RT-qPCR on a QuantStudio3 thermocycler (Applied Biosystems, USA) using Maxima SYBR Green/ROX qPCR Master Mix (2X) (K0221, Thermo Fisher Scientific, Waltham, MA, USA), using primers listed in Table 1, and *StTUBB1* (NM_001288449.1) was used as reference gene as described above in *Section 2.2*. The samples were run in triplicates, and the fold change was calculated by $2^{-\delta\delta Ct}$ as described previously (Livak and Schmittgen, 2001).

2.7 Determination of free sugars levels

The effect of Pho1a starch phosphorylase presence on starch degradation was determined by estimation of free sugars levels, i.e., sucrose, fructose, and glucose in flash-frozen and cold-stored tuber samples using Sucrose/D-Fructose/D-Glucose Assay Kit (K-SUFRG, Megazyme, Wicklow, Ireland) following the manufacturer's instructions. A total of 50 mg of homogenized material from flash-frozen tubers samples was utilized to determine the initial free sugar content in tuber samples at the time of harvest. The free sugar analysis was repeated on 50 mg of homogenized tuber samples, from tubers stored at 4°C for 3 months. In brief, the assay involved pH-dependent conversion of D-glucose and D-fructose into glucose-6-phosphate (G-6-P) intermediates using a hexokinase/phosphoglucose isomerase/glucose-6-phosphate dehydrogenase-based reaction and subsequent stoichiometric quantification of Nicotinamide adenine dinucleotide phosphate (NADPH) by absorbance at 340 nm (light path: 1 cm; ~25°C). The sucrose level was calculated from the difference in D-glucose concentration before and after the hydrolysis by β-fructosidase. D-Glucose and sucrose standards were used to ensure accuracy of spectrophotometer measurements and effectiveness of the β-fructosidase hydrolysis reaction by comparing D-glucose to D-fructose ratio, respectively. All analyses were performed in triplicates; free sugar levels were calculated using Mega-Calc^{TM} (Megazyme, Ireland) and reported as mean values of percentage on fresh weight (FW) basis.

3 Results

3.1 *Pho1a* mutagenesis yield more than four variant alleles in tetraploid potato

Mutations were induced in exon 2 of the *Pho1a* gene in the auto-tetraploid potato cv. Desirée. In total, 140 regenerated events were selected for HRFA screening for indels, and 129 out of the 140 events had induced mutations in at least one allele (92%) (data not shown). A total of 14 events, 5 events from Group 1 (full knockouts; FKO), 5 events from Group 2 (in-frame mutation; IFM), and 4 events from Group 3 (containing wild-type alleles; WTA), were selected for greenhouse cultivation and further characterization (Table 2, Figure 2, and Supplementary Figure S3). These selected mutational events, i.e., Starch Phosphorylase mutational events in Desirée, were termed as SPD-1 to SPD-15 in sequential order. A wide variation in allelic dosage of mutations was detected, and allelic variants in individual events spanned from one to seven. The presence of up to seven mutated alleles was confirmed by ICE analysis, and copy number of the gene was estimated to be 2.23 relative to 1 for a confirmed single-copy gene, *StGBSS1* by RT-qPCR-based copy number estimation. Taken together, these results indicated that the *Pho1a* locus is duplicated in cultivated tetraploid potato cv Desirée.

3.2 The *Pho1a* gene is duplicated in potato

The duplication of *Pho1a* and its genomic organization in the current DM v6.1 genome assembly was further confirmed by *in silico* analysis and targeted sequencing. A blast search for reference mRNA sequence of *Pho1a* (X52385.1) against DM v6.1 genome assembly resulted in five hits on chromosome 3 with high sequence identity, namely, Soltu.DM.03G007710 to Soltu.DM.03G007760 (Supplementary Table S1). The cDNA sequence of Soltu.DM.03G007710 shared high sequence homology with the first part of *Pho1a* mRNA, whereas Soltu.DM.03G007720 was highly similar to the distal part of reference mRNA sequence of *Pho1a* (X52385.1) (Supplementary Figure S1). Similarly, DNA sequence of Soltu.DM.03G007740 (annotated on reverse strand) and Soltu.DM.03G007750 were highly similar to the first part of *Pho1a* mRNA, whereas Soltu.DM.03G007760 was similar to the distal part (Supplementary Figure S1). The first part would contain predicted exons 1–5, while the distal part would contain predicted exons 6–15. In addition, a previously reported copia-like transposable element *Tst1* (X52287.1) was found to align with the genomic sequence between Soltu.DM.03G007710 and Soltu.DM.03G007720, and Soltu.DM.03G007750 and Soltu.DM.03G007760 with high similarity (Supplementary Figure S1). Based on sequence homology, these two genomic segments were considered to be duplicated *Pho1a* loci and termed *Pho1.1a* (i.e., Soltu.DM.03G007710 and Soltu.DM.03G007720) and *Pho1.2a* (i.e., Soltu.DM.03G007740 to Soltu.DM.03G007760) loci. Furthermore, the genomic sequence up to 7.7 kbp upstream of *Pho1.1a* and *Pho1.2a* coding sequence was nearly identical (Figure 3). However, a further 1.7 kbp upstream genomic region consisted of a 1.3-kbp insertional segment specific to the *Pho1.1a* locus. These two highly similar genomic regions, i.e., approximately 25 kbp region (Chr03:16803346.16827878) covering *Pho1.1a* and approximately 23 kbp region (Chr03:16827879.16851055) encompassing *Pho1.2a*, were delimited by approximately 300-bp-long conserved sequence, which may have served as recombination hotspots (Figure 3). The genomic assembly of both *Pho1a* loci was confirmed by targeted PCR amplification, cloning, and Sanger sequencing (Figure 3). Cloned sequences from amplicons 1 to 3 mapped specifically to the *Pho1.1a* locus, whereas amplicon 5 sequence was specific for the *Pho1.2a* locus (Figure 3). Amplicon 4 sequence, on the other hand, mapped to the end of exon 15 of *Pho1.1a* locus into the upstream region of the *Pho1.2a* locus, thus confirming the sequential order of *Pho1.1a* and *Pho1.2a* loci in the DM v6.1 genome assembly (Figure 3).

3.3 The *Pho1a* duplication is independent of the *Tst1* insertion

The origins of the *Tst1* insertion and its possible role in duplication were investigated by *in silico* analysis of putative *Pho1a* genomic loci in recently released genome assemblies of two

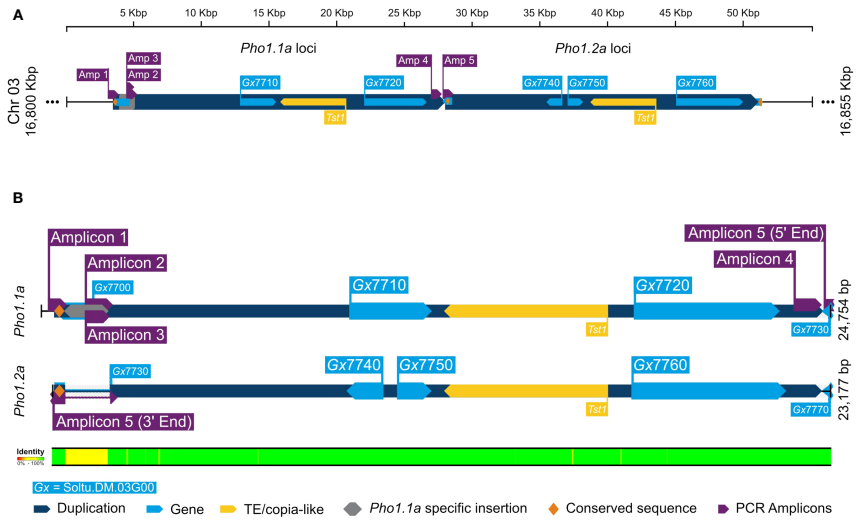

FIGURE 3
A graphical representation of alignment of duplicated *Pho1a* genomic loci in DM v6.1. **(A)** The duplicated genomic segments as dark blue blocks and gene annotations in DM v6.1 as light blue blocks. **(B)** Sequence alignment of *Pho1.1a* and *Pho1.2a* loci. The targeted cloning and sequencing amplicons are in purple, *Pho1.1a*-specific insertion sequence in gray, and conserved repeating sequence as yellow. The respective sequence IDs are annotated, and sequence identity conservation is represented as color coded legend at the bottom.

non-tuber-bearing wild potato species from *Etuberosum* section, 24 diploid wild potato species from tuber-bearing clade of *Petota* section, 2 landraces of *S. tuberosum* Group Phureja and tomato cv. M82. A blast search, using reference mRNA sequence of *Pho1a* (X52385.1) against the above-mentioned genome assemblies resulted in the identification of putative 1 or 2 *Pho1a* loci in respective assemblies (Supplementary Table S2). When compared to 13,716 bp in DM v6.1, the length of *Pho1a* loci in the above assemblies was variable, ranging from 8,440 to 16,228 bp, depending on the presence of insertion and the length of *Tst1* in intron 5 and species-specific insertions in introns 2, 9, 11, and 13 (Figure 4; Supplementary Table S2). However, the length of duplicated *Pho1a* loci, i.e., *Pho1.1a* and *Pho1.2a* were nearly identical in individual species (Figure 4; Supplementary Table S2). The duplication of *Pho1a* was found to be independent of insertion of *Tst1*, as both single and duplicated *Pho1a* loci were detected with and without a *Tst1* insertion in potato species (Figure 4; Supplementary Table S2). On the other hand, a single *Pho1a* locus without *Tst1* insertion was detected in both potato species from *Etuberosum* section and tomato, which suggests that the insertion of *Tst1* in *Pho1a* may be specific to certain clades of *Petota* section in *Solanaceae* (Figure 4; Supplementary Table S2). As a result, the above potato species could be organized into four groups, namely, 1) single *Pho1a* locus without insertion of *Tst1*, 2) single *Pho1a* locus with insertion of *Tst1*, 3) duplicated *Pho1a* loci without insertion of *Tst1*, and 4) duplicated *Pho1a* loci with insertion of *Tst1* (Figure 4).

3.4 *Pho1a* knockout results in an increased number of tubers

The general phenotypic impact of induced mutations was investigated in greenhouse cultivation. In general, the plant height and growth rate of mutated events was similar to Desirée (control) during cultivation in greenhouse, except for SPD-9, which had an aberrant phenotype (Figure 5; Supplementary Figure S4). The average total weight of harvested tubers per pot was higher in all groups as compared to control (Desriee), however, it was only significant in WTA group (Figure 6A). Most events had two- to four fold higher number of tubers as compared to Desirée, except for SPD-9, which only produced two relatively small tubers/pot (Figure 6B; Supplementary Figure S4). SPD-1, SPD-5, SPD-11, and SPD-14 had 30, 49, 52, and 64 tubers/pot on average as compared to 9 tubers/pot for Desirée (Figures 5, 6B; Supplementary Figure S4). Tubers from the majority of the events in the FKO and IFA groups were elongated, whereas tubers from the WTA group were more similar to Desirée (Figure 5; Supplementary Figure S4). SPD-5, SPD-11, and SPD-14 had a disproportionately high number of small round tubers (Figures 5, 6B; Supplementary Figure S4). Overall, in most events, the number of tubers were higher than Desirée, and a majority of them (30%–60%) were smaller in size (Figure 6B; Supplementary Figure S4). The average weight per tuber (FW) was lower in FKO and WTA groups as compared to Desirée, except for IFA group, which was similar to Desirée (Figure 6C). However, the average dry matter content of the harvested tubers was mostly similar to Desirée (Figure 6D).

10.3389/fpls.2023.1220973

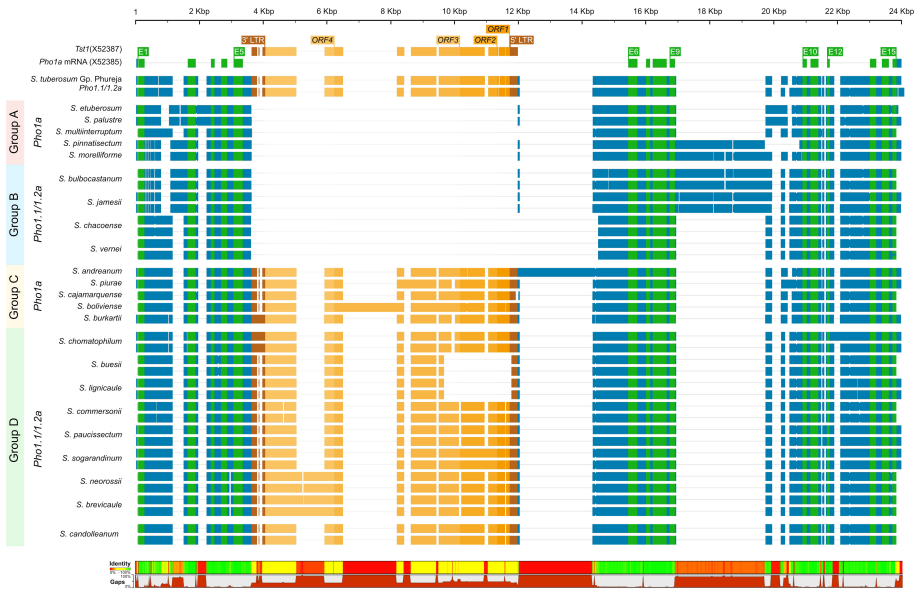

FIGURE 4
A graphical representation of alignment of *Pho1.1a* and *Pho1.2a* genomic sequences from diploid assemblies of wild and cultivated potato species from *Petota* and *Etuberosum* section. The species are divided into four groups based on presence of *Tst1* and duplication of *Pho1a* loci. Group 1: no duplication and insertion of *Tst1;* Group 2: duplicated *Pho1a* loci and insertion of *Tst1;* Group 3: No duplication and with insertion of *Tst1;* Group 2: duplicated *Pho1a* loci and with insertion of *Tst1.* The sequences are color coded as per: genomic as blue, mRNA as green, CDS/Exon as yellow, 5′ and 3′ LTRs as brown, and ORFs of *Tst1* as shades of dark yellow. The top scale represents length of single/duplicated *Pho1a* loci, and bottom bar chart represents sequence identity and gaps. The alignment was generated with Clustal Omega.

3.5 Full knockouts and in-frame mutants are deficient in Pho1a activity

To verify that the FKO events were deficient in Pho1a activity and to assess the other mutant groups, Zymograms were run on representatives of all three groups (Figure 7). Two types of zymograms were used, one glycogen containing running gel providing affinity retardation and primer (Figure 7A) and one lacking any added primer in the form of maltooligosaccharides, glycogen, or soluble starch (Figure 7B). Starch phosphorylase activity can be detected in both gels for Pho1a and the cytosolic form, Pho2. The migration of Pho2 is greatly affected by glycogen in the gel and can be found close to the well in Figure 7A but represents the fastest migrating activity band in Figure 7B. As expected, none of the events in the FKO group showed detectable activity for Pho1a. Interestingly, the same was true for all events with in-frame deletions (SPD-6 and SPD-10) and one of the mutants still carrying at least one wild-type alleles (SPD-11). This means that in-frame mutations where one or more amino acids are lost at the target site were detrimental for Pho1a activity. The only mutant with detectable Pho1a activity was SPD-15, which showed activity comparable to that of Desirée. This could indicate that all eight alleles encoding Pho1a do not have the same influence on total Pho1a activity.

3.6 Tuber amyloplasts of mutated events accumulate more starch granules

The amyloplasts from tubers of mutated events from both FKO and IFM groups contained an increased number of small granules as compared to Desirée (control) (Figure 8). The number of starch granules in WTA events were similar to Desirée (Figure 8). In general, amyloplasts from tubers of mutated events from both FKO and IFM groups predominantly contained a higher proportion of small granules (up to 80%), whereas the proportion of small granules in WTA group and Desirée varied from 30% to 60% (Figure 8). Smaller granules were found to be present in the amyloplast throughout the tubers from FKO and IFM groups, whereas for WTA and Desirée, the smaller granules were predominantly located in amyloplasts near the skin (Figure 9). The shape of the granules was more spherical in most FKO and IFM mutational events, irrespective of the size as compared to WTA events, which accumulated more oval granules similar to Desirée (Figure 8). Tubers from both FKO and IFM mutational events contained a high presence of reddish-brown stained spherical particles, in and around the vascular tissue of the tubers, which were rarely found in WTA events and Desirée (Figures 8, 9). The number of starch granules per chloroplast in stomatal guard cells from leaf samples from all groups of mutated events were found to be comparable to Desirée (Supplementary Figure 5).

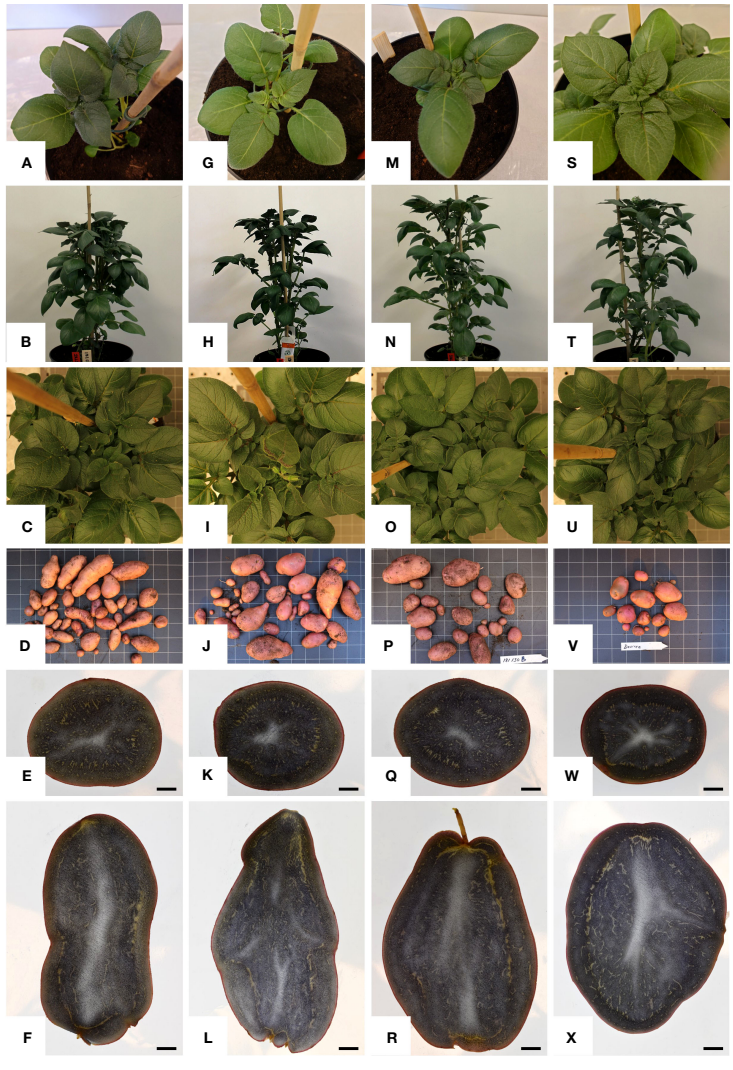

FIGURE 5
Phenotypic characterization of mutational events selected for greenhouse trial. One representative from each group, i.e., from left: column 1, **(A–F)** SPD-1: Group 1 (full knockouts; FKO); column 2, **(G–I)** SPD-6: Group 2 (in-frame knockouts; IFM); Column 3, **(M–R)** SPD-15 Group 3 (partial knockouts; WTA), and column 4, **(S–X)** Desirée (WT, control). First row: **(A, G, M, S)** represent top view of 2 weeks old plants; second row: **(B, H, N, T)** represent front view of 4-week-old plants; and third row: **(C, I, O, U)** represent top view of 4-week-old plants. Row 4: **(D, J, P, V)** represent tubers harvested from 4-month-old plants in greenhouse. Row 5: **(E, K, Q, W)**; row 6: **(F, L, R, X)** represent lateral and vertical sections of representative "medium-sized" tubers, which are stained with Lugol's solution and harvested from 4-month-old plants in greenhouse, respectively. The black bar in row 5 indicates 1 cm. For list of mutational events, see Table 2.

3.7 Full knock out of *Pho1a* affect the tuber starch content and amylose to amylopectin ratio is significantly decreased in tubers

The total starch content in four of the five FKO events was found slightly lower as compared to Desirée (control) (Figure 10A). SPD-1, 2, 3, and 5 had an average of total starch content at 16.3%, 14.4%, 14.0%, and 13.8%, respectively, as compared to 17.1% for Desirée (DW), but only the decreased content in SPD-2 and 5 was found to be significant (Figure 10A). SPD-4 had a slightly, but not significant, higher average total starch content measuring 18.0% (DW) (Figure 10A).

The amylose content was found to be consistently lower in the FKO events at 14.1%–14.6% for SPD-1, 2, 3, and 5, and 16.2% in

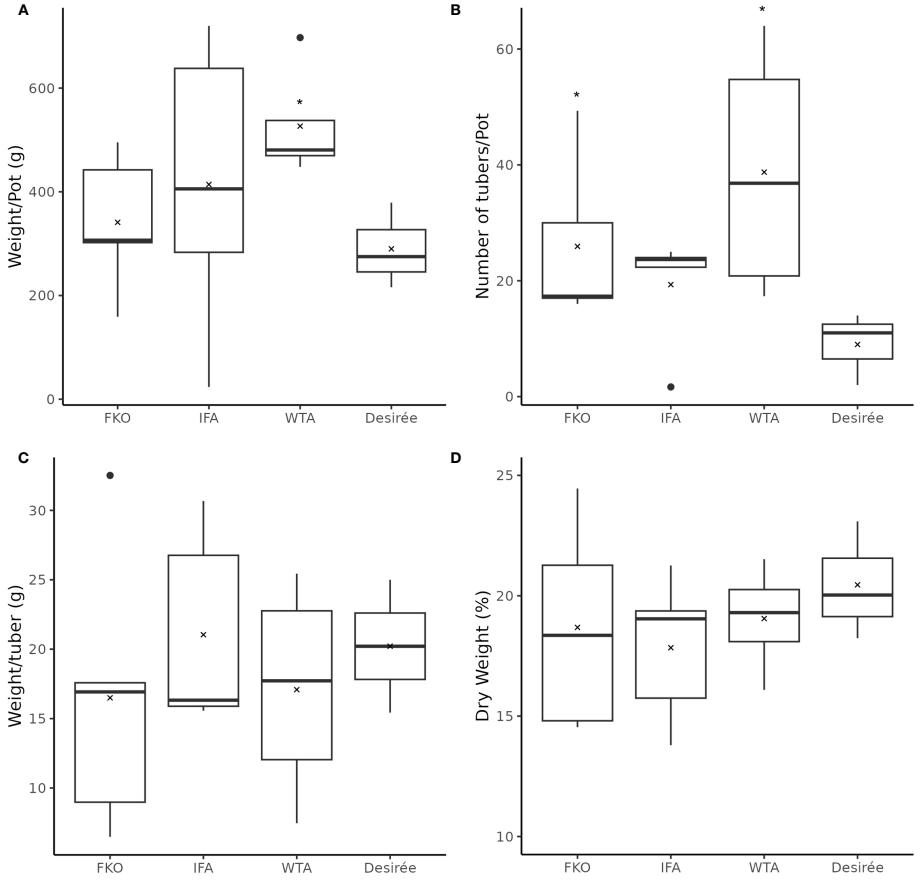

FIGURE 6
Phenotypic data of mutational events from greenhouse harvest at time of maturity. The above charts summarize four tuber characteristics of mutational events and Desirée (control). The mutational events are grouped according to groups, i.e., (from left to right) Group 1 (full knockouts; FKO), Group 2 (in-frame knockouts; IFM), and Group 3 (partial knockouts; WTA), and control. **(A)** Average total weight of tubers per plant (g), **(B)** average number of tubers per plant, **(C)** average weight per tuber (total weight of all tubers per pot/total number of tubers per pot; (g)), and **(D)** Average dry weight of tubers (percentage). All data were recorded in triplicates. Average and median values are represented as "X" and horizontal black bars. (*p<0.05, t-test—one tailed, two samples, equal variance).

SPD-4, as compared to 22.0% in Desirée (DW) (Figure 10B; Supplementary Figure S6). Similarly, amylose content in IFM and WTA group was also decreased as compared to Desirée (Supplementary Figure S6). The amylose content in SPD-10 and SPD-11 samples at 16.4% and 15.4%, respectively, was comparable to FKO group, whereas the determined amylose content in SPD-6 and SPD-15 at 21.2% and 21.6% was more comparable to Desirée (Supplementary Figure S6).

As *StGBSS1* is responsible for amylose synthesis in potato tubers, the expression of the corresponding gene was investigated in the mutational events. The expression of *StGBSS1* was found to be lower in all mutational events (Supplementary Figure S7). The decrease in expression could possibly explain the observed decrease in amylose content, although there was not a consistent pattern. Overall, the most reduction in the expression levels of *StGBSS1* was observed in WTA group, followed by IFA and least in FKO group (Supplementary Figure S7).

3.8 *Pho1a* knockout does not prevent cold sweetening of tubers

Free sugar levels, i.e., concentration of sucrose, glucose, and fructose in potato tubers were measured at harvest and after a 3-month cold storage at 4°C in potato tubers of FKO mutational events (Figures 11A–C). In general, the levels of free sugars were higher in cold stored tubers for FKO events and Desirée (control).

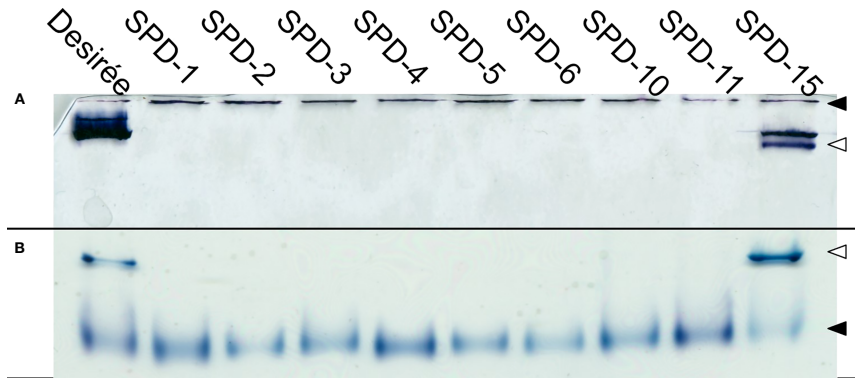

FIGURE 7
Phosphorylase activity gel. Protein extracts prepared from tubers of *Solanum tuberosum* were run on both native PAGE containing 0.8% glycogen **(A)**, and commercially available native PAGE **(B)**. Both conditions show clear activity for Pho2 (◀) for all the events, while only Desirée and SPD-15 show activity for Pho1a (◁).

The average free sucrose levels in cold stored tubers of FKO events was higher as compared to fresh tubers, differently from Desirée (control) where no significant change was detected in sucrose content between fresh and cold stored tubers (Figure 11A; Supplementary Figure S8A). The average free sucrose level of fresh tubers of the FKO group was slightly lower (0.32%) as compared to Desirée (0.43%) that increased in cold-stored FKO tubers (0.72%) as compared to no change in Desirée (0.43%) (Figure 11A; Supplementary Figure S8A). Very little to no free glucose was detected in fresh tubers of FKO mutational events (0.08%) and Desirée (0.09%) (Figure 11B; Supplementary Figure 8B), with significant increase in the cold-stored tuber samples. The average free glucose level was increased to 0.87% in tubers of the FKO group as compared to 1.50% in Desirée (Figure 11B). Similarly, the free fructose level in the fresh tuber samples of the FKO group and Desirée (control) was at the limit of detection but was significantly increased in respective cold-stored samples (Figure 11C). The average fructose content increased to 0.65% in FKO group tubers as compared to 0.77% in Desirée.

The higher average sucrose level of cold-stored FKO group tubers compared to Desirée (control), concomitant with a lower glucose level of cold-stored FKO group tubers compared to Desirée (control), could indicate that Pho1a has a role in cold sweetening of potato tubers.

4 Discussion

The role of plastid starch phosphorylase is still under investigation in plants. Even though it is one of the most highly expressed genes in potato tubers and one of the most prominent enzymatic activities in amyloplasts, its role, if any, for reserve starch synthesis is still unclear. In this manuscript, we identified that the major form, Pho1a, is encoded by a duplicated gene locus, and we further characterized its role for reserve starch composition and phenotype by observing CRISPR/Cas9 mutants of *Pho1a*.

The first indication of a duplicated *Pho1a* locus was found by the detection of up to seven differently mutated alleles in a single regenerated potato event. A plausible cause for this could be chimeric regeneration consisting of differently mutated somatic cells. However, applying RNP CRISPR/Cas9 tools to potato protoplasts generally avoids this problem from the transient nature of the application. Stable transformation for CRISPR/Cas9 mutations carries a higher risk of creating chimeric plants as mutations can occur at any time during the regenerative and vegetative phase. Copy number estimation of the *Pho1a* locus by RT-qPCR confirmed that more than one copy exists in the parental variety Desirée. A duplication of the *Pho1a* locus was further verified, analyzing the current available potato genome assembly DM v6.1. In addition, duplication of the *Pho1a* locus was also evident in multiple assembled wild potato genomes. The *Pho1a* copies on chromosome 3 were found to have highly conserved sequences in the DM v6.1 assembly. A ~7.5-kbp upstream region of both copies was also highly conserved. With the exception of a ~1.3-kbp long upstream genomic segment, ~25 kbp long genomic segment harboring *Pho1.1a* and ~23 kbp long genomic segment encompassing *Pho1.2a* were nearly identical. Both segments were delimited by ~300-bp-long conserved sequences, which were only detected in *Pho1a* loci from cultivated tetraploid potato cultivars. The presence of these conserved sequences may act as recombination hotspots during meiosis.

A reference mRNA sequence of *Pho1a* (X52385.1) spanned across 15 exons that were highly conserved in wild and cultivated potato assemblies, indicating an evolutionary constraint. However, the intronic regions were less conserved and contained species specific insertions. As a result, the length of a putative *Pho1a* locus ranged from 8,440 to 16,228 bp, depending on insertion and length of *Tst1* in intron 5 and species-specific insertions in intron 2, 9, 11, and 13.

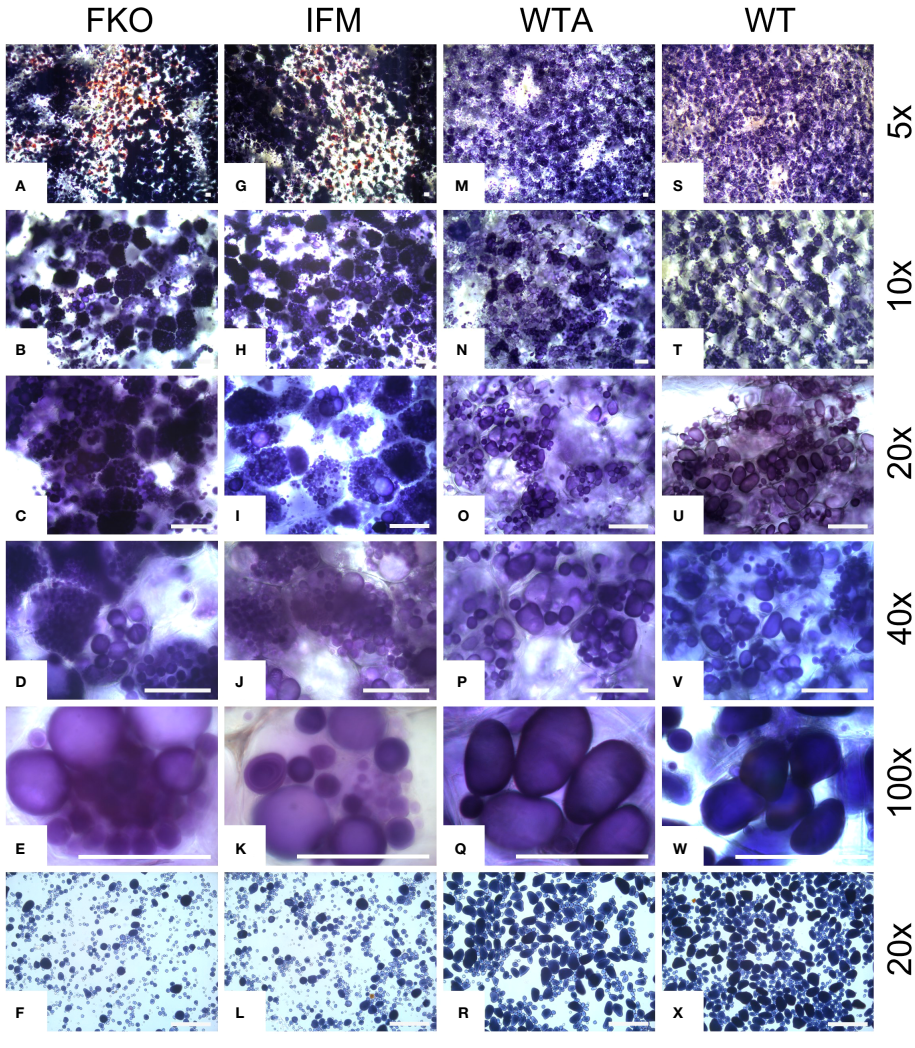

FIGURE 8
Light micrographs of starch granules stained with Lugol's solution in tubers of representative mutational events. **(A–F)** SPD-1: Group 1 (full knockouts; FKO); **(G–L)** SPD-6: Group 2 (in-frame knockouts; IFM); **(M–R)** SPD-15: Group 3 (partial knockouts; WTA) and **(S–X)** Desirée (WT, control). Rows 1–5 represent stained tuber slices visualized under a microscope in increasing order of magnification under magnification as indicated in the figure. Each row has the same magnification, and black/white bar indicates 100 µm. For list of mutational events, see Table 2.

The duplicated *Pho1a* locus is miss-annotated in the DM v6.1 assembly where *Pho1.1a* locus is annotated as Soltu.DM.03G007710 and Soltu.DM.03G007720. Both genes share a high sequence identity with exons 1–5 and exons 6–15 of reference mRNA sequence of *Pho1a*. Similarly, *Pho1.2a* locus is annotated as Soltu.DM.03G007750 and Soltu.DM.03G007760, which share high sequence identity to exons 1 and 2, 3–5 and exons 6–15 of reference mRNA sequence of *Pho1a*, respectively. A comparison of putative *Pho1a* loci among DM v6.1, DM v8.1, DM1S1, and *S. tuberosum* Group Phureja assemblies of E86-69 and E4-63 indicated probable sequencing errors leading to putative stop codons and thus resulting in mis-prediction of coding sequences of above genes. Furthermore, these sequencing errors are resolved in DM1S1, and duplicated *Pho1a* loci are annotated as Soltu.DM1S1.03G006810, i.e., *Pho1.1a* and Soltu.DM1S1.03G006830, i.e., *Pho1.2a*.

Pho1a has previously been reported to contain the transposon sequence *Tst1* in the intron between exons 5 and 6 (CamIrand et al., 1990). It could be speculated that the transposon insertion could have triggered a gene duplication. However, *Tst1* sequences were detected in both single and duplicated *Pho1a* genes across 24 wild

10.3389/fpls.2023.1220973

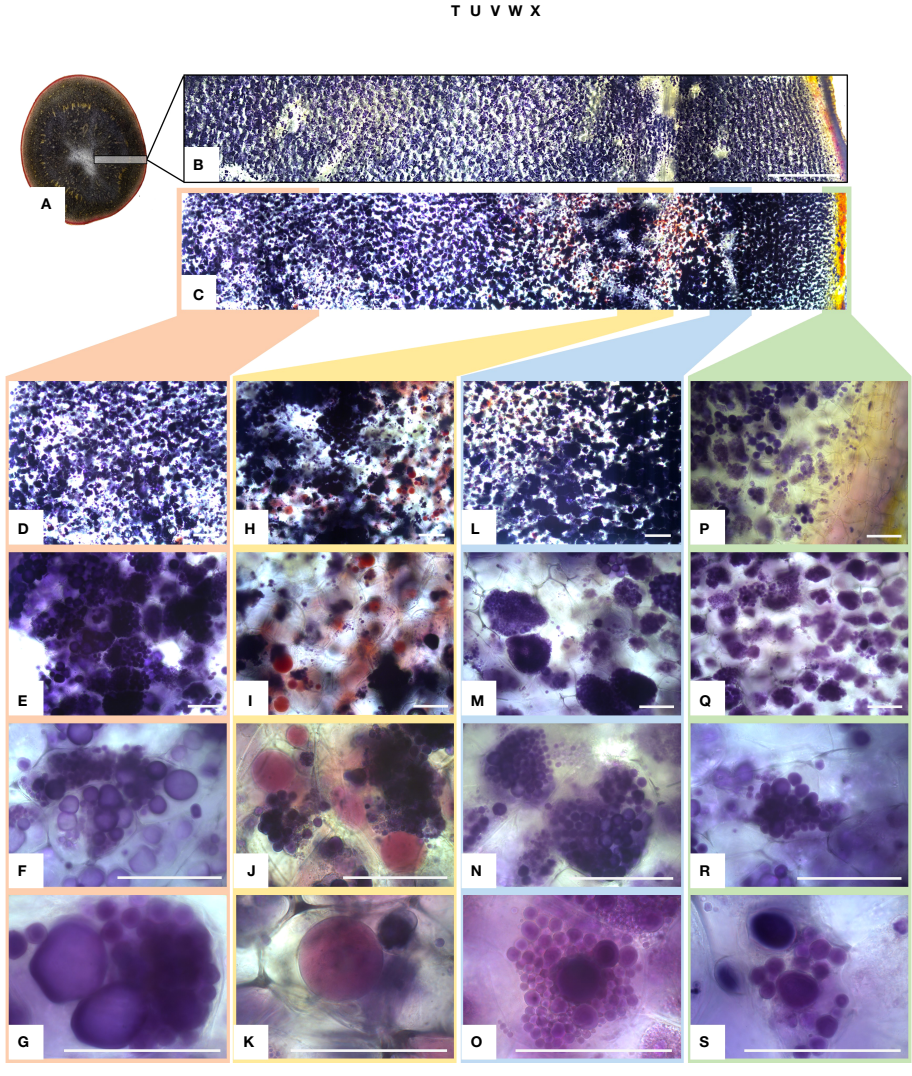

FIGURE 9
Light micrographs of starch granules stained with Lugol's solution in tubers of mutational event SPD-1 from Group 1 (full knockouts; FKO) and Desirée control. **(A)** Lateral section of tuber stained with Lugol's solution, harvested from 4-month-old plants in greenhouse. **(B, C)** Desirée and SPD-1, respectively, section of the tuber slice used for visualization under a microscope; left to right: center to periphery (skin) of tuber. Panels **(B, C)** are composite image of six individual images. Columns 1–4 (left to right: center to periphery (skin) of tuber slice) represent four areas under visualization with increasing magnification (top to bottom). **(D–G)** Center; **(H–K)** area around the vascular tissue; **(I–O):** area right to vascular tissue and **(P–S):**: periphery (skin) of tuber slice. Magnification is indicated in the figure. Each row has the same magnification, and black/white bar indicates 100 µm. For list of mutational events, see Table 2.

potato species. The *Tst1* sequences were found to be highly variable among species; however, both *Tst1* sequences of duplicated copies were nearly identical. Insertion of *Tst1* was not detected in *Pho1a* sequences from *S. palustre* and *S. etuberosum* from *Etuberosum* section or tomato, and coding sequence of *Pho1a* from tomato spanned only 14 exons. This indicates separate origins of found duplications of *Pho1a* and independence of *Tst1* presence. The duplication of *Pho1a* and insertion of *Tst1* must have originated in distinct duplication events during evolution. In addition, *Pho1b* genomic sequence was found to span 14 exons similar to *Pho1* from Etuberosum section and tomato, lacking *Tst1* insertion, which may suggest that *Pho1b* is the ancestral copy of *Pho1* gene.

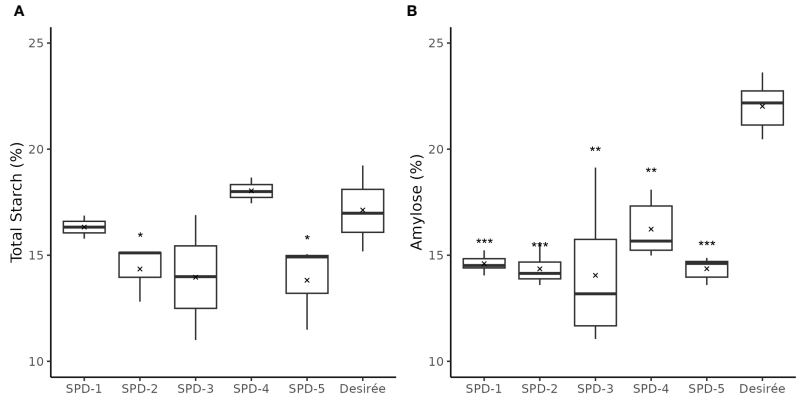

FIGURE 10
Total starch and amylose content in tubers of five full knockout events from Group 1 (FKO) and Desirée (WT, control). **(A)** Total starch measurements and **(B)** amylose fraction, represented as percentage of dry weight. All measurements were made in triplicates. Average and median values are represented as "X" and horizontal black bars. (*p<0.05, **p<0.01, ***p<0.001; *t-test*—one tailed, two samples, equal variance).

Tuber development of complete knockout events was clearly affected. Tubers were elongated, smaller, and the number per plant was higher than the control. This may suggest that Pho1a has an important role in sink development and maintenance, affecting tuber setting and growth. An altered tuber sink capacity might result in the induction of more tubers from available sucrose transport. The overall tuber weight, dry matter, and starch content were similar to the control, which shows that total sink capacity into tubers and starch accumulation is not affected. On the other hand, the amylose content of the mutational events was significantly reduced in tubers of mutational events from the FKO group, although the reduction in amylose among IFM and WTA groups was more variable. This observation may have different explanations, since it is apparent amylose content that is measured. Thus, structural changes could either be in long chains of amylopectin or in amylose. StGBSS1 (granule bound starch synthase) is the enzyme responsible for amylose synthesis, and the expression of the corresponding gene was assessed. The relative expression of *StGBSS1* was found to be reduced in all mutational events. Interestingly, the reduction in the expression was highest in mutational events from WTA, followed by IFM and FKO groups. Taken together, this suggests a complex interaction of expression regulation among starch biosynthetic genes and possible protein–protein interactions, which needs to be further elucidated in potato.

Starch phosphorylase activity of mutant events was investigated via banding patterns in native PAGE gel zymograms with and without glycogen in the separation gel. Glycogen functions as affinity retardation for starch phosphorylase proteins and a

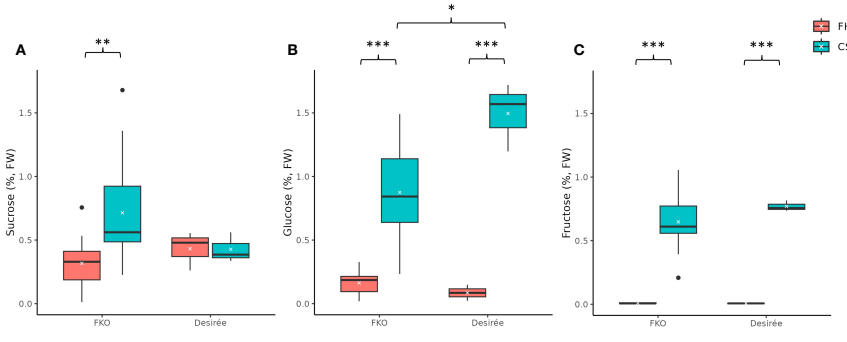

FIGURE 11
Free sugar levels in fresh and cold-stored tubers of five full knockout events from Group 1 (FKO) and Desirée (WT, control). **(A)** Sucrose, **(B)** glucose, and **(C)** fructose levels, represented as percentage of dry weight on respective y-axis. Fresh and cold-stored tuber samples are colored as orange and green, respectively. All measurements were recorded as triplicates. Statistical analyses were made on freshly harvested tubers and cold stored tubers, respectively. Average and median values are represented as "X" and horizontal black bars. (*p<0.05, **p<0.01, ***p<0.001; t-test—one tailed, two samples, equal variance).

possible primer for their biosynthetic activity. No Pho1a activity was detected in mutational events from FKO and IFM groups, which shows that full knockout and in-frame induced mutations lead to complete lack of Pho1a activity in respective events. However, Pho1a activity was only detected in SPD-15 and not in SPD-11 from WTA group. For the IFM group, this indicates that the amino acids not present in the mutant allele are crucial for detectable activity or that the particular in-frame mutant allele is of lesser importance as an allele for Pho1a activity. The results of the WTA group may suggest that there is an allelic variance among the eight alleles of *Pho1.1a* and *Pho1.2a* regarding contribution to Pho1a activity. Consistent with the literature (Albrecht et al., 1998), the Pho2 activity band migration was greatly retarded in the glycogen containing gel while having greater migration than Pho1a in a non-glycogen gel. Pho1a and Pho2 was observed to yield activity bands without a deliberately added primer in the form of maltooligosaccharides, glycogen, or starch.

While the AGPase-mediated starch biosynthesis is considered to be the absolute major pathway for reserve starch, plastid starch phosphorylase has been suggested to play a role for reserve starch accumulation, at least in some species and under some conditions (Hwang et al., 2020). A reduction in the size of the starch granules and a change in amylopectin chain length distribution due to mutations of *Pho1* has been reported in rice endosperm (Satoh et al., 2008). Similarly, the amyloplasts in the tubers from our mutated events were found to contain a large number of small starch granules. The abundance of small granules was higher at the periphery of the tuber tissue. Larger mature granules were also found, but then mostly located towards the center of the tubers. In addition, the shape of the granules was consistently rounder irrespective of the size as compared to the control. The roundish structure could be due to changed chain distribution or growth direction of starch and hence granule formation. Plastid starch phosphorylase could be of importance for the synthesis of long chains in amylopectin that might be reflected in the observed decrease in apparent amylose content. A change in chain length distribution where iodine no longer detects certain chains as amylose because of them now being shorter.

The accumulation of massive number of small granules may come from impairment in a biosynthetic and degradative direction. Pho1 has been suggested to be involved in the degradation of MOs chains to G-1-Ps in rice, where the MOs are resulting from trimming of pre-amylopectin by isoamylase-type DBEs (Hwang et al., 2010; Hwang et al., 2016a; Hwang et al., 2020). In wheat, Pho1 has been suggested to act phosphorolytically, directly at the surface of the granule (Tickle et al., 2009) and on linear glucans originating from degradation of starch granules (Tetlow and Bertoft, 2020). For both routes, this would release G-1-P, which converted to ADP-glucose, and might be utilized by SSs for transfer to other growing granules (Mérida and Fettke, 2021). In this way, Pho1a may contribute, together with other enzymatic activities, to ensure building, trimming, and structuring of the final starch granule phenotype and indirectly keep the number of granules in amyloplasts down.

A proliferation of smaller non-staining granules and red-stained granules was detected in the amyloplasts surrounding vascular tissue of the tubers of most full knockout mutational events. A similar phenotype has been reported previously in the tubers of transgenic events with antisense RNA suppression of two isoamylase genes, i.e., *Stisa1* and *Stisa2* in cv Desiree (Bustos et al., 2004). These very small granules were reported to consist of both amylose and amylopectin; however, the ratio and molecular masses of both polymers and organization were reported to be substantially different from large granules. The differences were attributed to the small surface area for amylopectin synthesis, resulting in lack of growth rings. The red-stained granules, on the other hand, were reported to have a much higher proportion of short chains than potato amylopectin and were more similar to phytoglycogen from the *sugary*1 mutant of maize. Phytoglycogen molecules result in spherical structures due to intrinsic lack of granular organization arising from a high degree of branching with short α-1,4 glucan chains. However, in our study, the isoamylase activity was not targeted, and no reports of any multi-enzyme complexes involving Pho1a and DBEs exist so far (Zhong et al., 2022). Pho1b has been shown to be specifically located to vascular tissue in potato tubers while the corresponding gene predominantly being expressed in leaf tissue (Albrecht et al., 2001). This leads to the possibility that extraction of RNA from whole tuber sample may result in the misrepresentation of expression in vascular tuber tissue. Pho1b is reported to exist as a heterodimer complex with Pho1a in potato leaves (Albrecht et al., 1998; Albrecht et al., 2001). It could be hypothesized that proliferation of tiny and red-stained granules may be the result of loss of a heterodimer complex in amyloplasts surrounding vascular tissue, although this would need further experimentation.

Pho1 has been suggested to play a role in the transitory starch degradation due to high [Pi] concentration (Kruger and ap Rees, 1983). However, downregulation of *PHS1* in *Arabidopsis* (Zeeman et al., 2004) and *Pho1b (STP-1)* in potato leaves (Sonnewald et al., 1995) did not significantly alter starch structure or diurnal starch metabolism, suggesting that Pho1 is not a major determinant in starch metabolism under normal conditions (Hwang et al., 2020). In agreement with above, knockout of *Pho1a* did not alter the number or size of starch granules in leaves of mutational events.

Pho1a has been suggested to be involved in the regulation of potato response to changes in temperature; however, effects of cold stress and cold storage of tubers have opposite effects (Slugina et al., 2020a). While potato tubers grown at low temperature had no significant effect on tuber starch content and granule size (Orawetz et al., 2016), the tubers stored at low temperature resulted in decreased tuber starch and increased free sugar accumulation along with increased Pho1a activity (Schreiber et al., 2014; Slugina et al., 2020a). The free sugar content in cold stored tubers of the fully mutated events was found to be higher than for fresh tubers. However, the increase in free sucrose, glucose, and fructose followed different patterns in comparison to Desirée (control). While the average sucrose level in cold-stored tuber samples was higher, the glucose level was significantly lower and with fructose on a similar level to cold-stored tuber samples from control. Potato tubers remain metabolically active, and starch degradation and sucrose accumulation are a perquisite for loss of post-harvest ecodormancy (Sonnewald and Sonnewald, 2014). Exposure to

cold storage is suggested to influence duration of dormancy by affecting activity levels of saccharolytic enzymes like sucrose synthase, or vacuolar acid invertase and its inhibitor (Slugina et al., 2020a). The concomitant higher sucrose and lower glucose level of the FKO group in comparison to the Desirée control indicate a role of Pho1a in cold sweetening of potato tubers.

5 Conclusion

Pho1a was shown to be tandemly duplicated in potato and thus named *Pho1.1a* and *Pho1.2a*. The gene duplication does not exist in tomato but can be found in species more related to *Solanum tuberosum*. Both genes encoding Pho1a were mutated using CRISPR/Cas9, and the impact on tuber development and starch accumulation was investigated in full knockouts. The results showed that lack of Pho1.1a and Pho1.2a have a large impact on tuber shape, size, and number per plant. Furthermore, an increase in small starch granules and more round granule shape was found, while the overall apparent amylose content was decreased. A differential, although low effect from absence of Pho1a, was observed upon cold storage of tubers that may indicate a role for Pho1a in starch degradation and cold sweetening. This study firmly establishes Pho1.1a and Pho1.2a as an important enzymatic factor in forming the starch granule phenotype and structure in reserve starch accumulation of potato tubers.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Files, further inquiries can be directed to the corresponding author/s.

Author contributions

PH conceptualized the study. PH, MA, and SS planned the study. SS, MF, MA, PV, HT, and NO performed experiments. All authors analyzed results. SS and PH wrote the manuscript. SS, PH, MA, and MF edited the manuscript. All authors approved the manuscript in its final form.

Funding

This study was funded by Formas grants 2018-01846 and 2018-01459 and Crops for the Future (C4F) a Strategic Research Area at SLU supported by the Swedish Government.

Acknowledgments

We thank Marina Kuzmenkova, Mirela Beganovic, and Ying Liu for assistance with tissue culture, greenhouse cultivation, and harvesting.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1220973/full#supplementary-material

References

Ahmed, Z., Tetlow, I. J., Ahmed, R., Morell, M. K., and Emes, M. J. (2015). Protein–protein interactions among enzymes of starch biosynthesis in high-amylose barley genotypes reveal differential roles of heteromeric enzyme complexes in the synthesis of A and B granules. *Plant Sci.* 233, 95–106. doi: 10.1016/j.plantsci.2014.12.016

Albrecht, T., Greve, B., Pusch, K., Kossmann, J., Buchner, P., Wobus, U., et al. (1998). Homodimers and heterodimers of Pho1-type phosphorylase isoforms in Solanum tuberosum L. as revealed by sequence-specific antibodies. *Eur. J. Biochem.* 251 (1–2), 343–352. doi: 10.1046/j.1432-1327.1998.2510343.x

Albrecht, T., Koch, A., Lode, A., Greve, B., Schneider-Mergener, J., and Steup, M. (2001). Plastidic (Pho1-type) phosphorylase isoforms in potato (Solanum tuberosum L.) plants: Expression analysis and immunochemical characterization. *Planta* 213 (4), 602–613. doi: 10.1007/s004250100525

Alonge, M., Lebeigle, L., Kirsche, M., Jenike, K., Ou, S., Aganezov, S., et al. (2022). Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biol.* 23 (1), 258. doi: 10.1186/s13059-022-02823-7

Andersson, M., Melander, M., Pojmark, P., Larsson, H., Bülow, L., and Hofvander, P. (2006). Targeted gene suppression by RNA interference: An efficient method for production of high-amylose potato lines. *J. Biotechnol.* 123 (2), 137–148. doi: 10.1016/j.jbiotec.2005.11.001

Andersson, M., Turesson, H., Nicolia, A., Fält, A.-S., Samuelsson, M., and Hofvander, P. (2017). Efficient targeted multiallelic mutagenesis in tetraploid potato (Solanum tuberosum) by transient CRISPR-Cas9 expression in protoplasts. *Plant Cell Rep.* 36 (1), 117–128. doi: 10.1007/s00299-016-2062-3

Brisson, N., Giroux, H., Zollinger, M., Camirand, A., and Simard, C. (1989). Maturation and subcellular compartmentation of potato starch phosphorylase. *Plant Cell* 1 (5), 559–566. doi: 10.1105/tpc.1.5.559

Bustos, R., Fahy, B., Hylton, C. M., Seale, R., Nebane, N. M., Edwards, A., et al. (2004). Starch granule initiation is controlled by a heteromultimeric isoamylase in potato tubers. *Proc. Natl. Acad. Sci.* 101 (7), 2215–2220. doi: 10.1073/pnas.0305920101

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. MGG* 224 (1), 33–39. doi: 10.1007/BF00259448

Chen, H. M., Chang, S. C., Wu, C. C., Cuo, T. S., Wu, J. S., and Juang, R. H. (2002). Regulation of the catalytic behaviour of L-form starch phosphorylase from sweet potato roots by proteolysis. *Physiol. Plantarum* 114 (4), 506–515. doi: 10.1034/j.1399-3054.2002.1140402.x

Chrastil, J. (1987). Improved colorimetric determination of amylose in starches or flours. *Carbohydr. Res.* 159 (1), 154–158. doi: 10.1016/S0008-6215(00)90013-2

Crofts, N., Nakamura, Y., and Fujita, N. (2017). Critical and speculative review of the roles of multi-protein complexes in starch biosynthesis in cereals. *Plant Sci.* 262, 1–8. doi: 10.1016/j.plantsci.2017.05.007

Cuesta-Seijo, J. A., Ruzanski, C., Krucewicz, K., Meier, S., Hägglund, P., Svensson, B., et al. (2017). Functional and structural characterization of plastidic starch phosphorylase during barley endosperm development. *PloS One* 12 (4), e0175488. doi: 10.1371/journal.pone.0175488

Fettke, J., Albrecht, T., Hejazi, M., Mahlow, S., Nakamura, Y., and Steup, M. (2010). Glucose 1-phosphate is efficiently taken up by potato (Solanum tuberosum) tuber parenchyma cells and converted to reserve starch granules. *New Phytol.* 185 (3), 663–675. doi: 10.1111/j.1469-8137.2009.03126.x

Fettke, J., Leifels, L., Brust, H., Herbst, K., and Steup, M. (2012). Two carbon fluxes to reserve starch in potato (Solanum tuberosum L.) tuber cells are closely interconnected but differently modulated by temperature. *J. Exp. Bot.* 63 (8), 3011–3029. doi: 10.1093/jxb/ers014

Flores-Castellanos, J., and Fettke, J. (2022). The plastidial glucan phosphorylase affects the maltooligosaccharide metabolism in parenchyma cells of potato (Solanum tuberosum L.) tuber discs. *Plant Cell Physiol.* 00 (00), 1–11. doi: 10.1093/pcp/pcac174

Hanes, C. S. (1940). The reversible formation of starch from glucose-1- phosphate catalysed by potato phosphorylase. *Proc. R. Soc. London. Ser. B - Biol. Sci.* 129 (855), 174–208. doi: 10.1098/rspb.1940.0035

Hwang, S. K., Nishi, A., Satoh, H., and Okita, T. W. (2010). Rice endosperm-specific plastidial α-glucan phosphorylase is important for synthesis of short-chain malto-oligosaccharides. *Arch. Biochem. Biophys.* 495 (1), 82–92. doi: 10.1016/j.abb.2009.12.023

Hwang, S. K., Koper, K., Satoh, H., and Okita, T. W. (2016a). Rice endosperm starch phosphorylase (Pho1) assembles with disproportionating enzyme (Dpe1) to form a protein complex that enhances synthesis of malto-oligosaccharides. *J. Biol. Chem.* 291 (38), 19994–20007. doi: 10.1074/jbc.M116.735449

Hwang, S. K., Koper, K., and Okita, T. W. (2020). "The plastid phosphorylase as a multiple-role player in plant metabolism," in *Plant Science* (Ireland: Elsevier Ireland Ltd). doi: 10.1016/j.plantsci.2019.110303

Hwang, S. K., Singh, S., Cakir, B., Satoh, H., and Okita, T. W. (2016b). The plastidial starch phosphorylase from rice endosperm: catalytic properties at low temperature. *Planta* 243 (4), 999–1009. doi: 10.1007/s00425-015-2461-7

Jayakody, T. B., Hamilton, J. P., Jensen, J., Sikora, S., Wood, J. C., Douches, D. S., et al. (2023). Genome Report: Genome sequence of 1S1, a transformable and highly regenerable diploid potato for use as a model for gene editing and genetic engineering. *G3 Genes|Genomes|Genetics*. 13, 1–8. doi: 10.1093/g3journal/jkad036

Kruger, N. J., and ap Rees, T. (1983). Properties of α-glucan phosphorylase from pea chloroplasts. *Phytochemistry* 22 (9), 1891–1898. doi: 10.1016/0031-9422(83)80007-7

Li, R., Zheng, W., Jiang, M., and Zhang, H. (2021). A review of starch biosynthesis in cereal crops and its potential breeding applications in rice (Oryza Sativa L.). *PeerJ* 9, e12678. doi: 10.7717/peerj.12678

Lin, Y.-C., Chang, S.-C., and Juang, R.-H. (2017). Plastidial α-glucan phosphorylase 1 complexes with disproportionating enzyme 1 in Ipomoea batatas storage roots for elevating malto-oligosaccharide metabolism. *PloS One* 12 (5), e0177115. doi: 10.1371/journal.pone.0177115

Lin, Y.-C., Chang, S.-C., and Juang, R.-H. (2012). Plastidial starch phosphorylase in sweet potato roots is proteolytically modified by protein-protein interaction with the 20S proteasome. *PloS One* 7 (4), e35336. doi: 10.1371/journal.pone.0035336

Liu, F., Chen, H.-M., Chou, I.-M., Chen, A.-N., Chen, C.-P., Young, G.-H., et al. (2009). The amylose extender mutant of maize conditions novel protein-protein interactions between starch biosynthetic enzymes in amyloplasts. *J. Exp. Bot.* 60 (15), 4423–4440. doi: 10.1093/jxb/erp297

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2–ΔΔCT method. *Methods* 25 (4), 402–408. doi: 10.1006/meth.2001.1262

Lu, Y., Steichen, J. M., Yao, J., and Sharkey, T. D. (2006). The role of cytosolic α-glucan phosphorylase in maltose metabolism and the comparison of amyloMaltase in Arabidopsis and Escherichia coli. *Plant Physiol.* 142 (3), 878–889. doi: 10.1104/pp.106.086850

Mérida, A., and Fettke, J. (2021). Starch granule initiation in Arabidopsis thaliana chloroplasts. *Plant J.* 107 (3), 688–697. doi: 10.1111/tpj.15359

Mori, H., Tanizawa, K., and Fukui, T. (1991). Potato tuber type H phosphorylase isozyme: Molecular cloning, nucleotide sequence, and expression of a full-length cDNA in Escherichia coli. *J. Biol. Chem.* 266 (28), 18446–18453. doi: 10.1016/s0021-9258(18)55082-1

Nakamura, Y., Ono, M., Utsumi, C., and Steup, M. (2012). Functional interaction between plastidial starch phosphorylase and starch branching enzymes from rice during the synthesis of branched maltodextrins. *Plant Cell Physiol.* 53 (5), 869–878. doi: 10.1093/pcp/pcs030

Nakamura, Y., Ono, M., Sawada, T., Crofts, N., Fujita, N., and Steup, M. (2017). Characterization of the functional interactions of plastidial starch phosphorylase and starch branching enzymes from rice endosperm during reserve starch biosynthesis. *Plant Sci.* 264, 83–95. doi: 10.1016/j.plantsci.2017.09.002

Nakano, K., and Fukui, T. (1986). The complete amino acid sequence of potato α-glucan phosphorylase. *J. Biol. Chem.* 261 (18), 8230–8236. doi: 10.1016/s0021-9258(19)83900-5

Nakano, K., Mori, H., and Fukui, T. (1989). Molecular Cloning of cDNA Encoding Potato Amyloplast a-Glucan Phosphorylase and the Structure of Its Transit Peptide1. *J. Biochem.* 106 (4), 691–695. doi: 10.1093/oxfordjournals.jbchem.a122918

Nicolia, A., Fält, A.-S., Hofvander, P., and Andersson, M. (2021). "Protoplast-based method for genome editing in tetraploid potato," in *Methods in molecular biology* (Humana Press Inc.), 177–186. doi: 10.1007/978-1-0716-1201-9_12

Orawetz, T., Malinova, I., Orzechowski, S., and Fettke, J. (2016). Reduction of the plastidial phosphorylase in potato (Solanum tuberosum L.) reveals impact on storage starch structure during growth at low temperature. *Plant Physiol. Biochem.* 100, 141–149. doi: 10.1016/j.plaphy.2016.01.013

Ovecka, M., Bahaji, A., Muñoz, F. J., Almagro, G., Ezquer, I., Baroja-Fernández, E., et al. (2012). A sensitive method for confocal fluorescence microscopic visualization of starch granules in iodine stained samples. *Plant Signaling Behav.* 7 (9), 1146–1150. doi: 10.4161/psb.21370

Pham, G. M., Hamilton, J. P., Wood, J. C., Burke, J. T., Zhao, H., Vaillancourt, B., et al. (2020). Construction of a chromosome-scale long-read reference genome assembly for potato. *GigaScience* 9 (9). doi: 10.1093/gigascience/giaa100

Preiss, J., and Levi, C. (1980). "Starch biosynthesis and degradation," in *Carbohydrates: structure and function* (Elsevier), 371–423. doi: 10.1016/B978-0-12-675403-2.50016-8

Priyam, A., Woodcroft, B. J., Rai, V., Moghul, I., Munagala, A., Ter, F., et al. (2019). Sequenceserver: A modern graphical user interface for custom BLAST databases. *Mol. Biol. Evol.* 36 (12), 2922–2924. doi: 10.1093/molbev/msz185

Rathore, R. S., Garg, N., Garg, S., and Kumar, A. (2009). Starch phosphorylase: Role in starch metabolism and biotechnological applications. *Crit. Rev. Biotechnol.* 29 (3), 214–224. doi: 10.1080/07388550902926063

Satoh, H., Shibahara, K., Tokunaga, T., Nishi, A., Tasaki, M., Hwang, S. K., et al. (2008). Mutation of the plastidial α-glucan phosphorylase gene in rice affects the synthesis and structure of starch in the endosperm. *Plant Cell* 20 (7), 1833–1849. doi: 10.1105/tpc.107.054007

Schönhals, E. M., Ortega, F., Barandalla, L., Aragones, A., Ruiz de Galarreta, J. I., Liao, J. C., et al. (2016). Identification and reproducibility of diagnostic DNA markers for tuber starch and yield optimization in a novel association mapping population of potato (Solanum tuberosum L.). *Theor. Appl. Genet.* 129 (4), 767–785. doi: 10.1007/s00122-016-2665-7

Schreiber, L., Nader-Nieto, A. C., Schönhals, E. M., Walkemeier, B., and Gebhardt, C. (2014). SNPs in genes functional in starch-sugar interconversion associate with natural variation of tuber starch and sugar content of potato (Solanum tuberosum L.). *G3 Genes|Genomes|Genetics* 4 (10), 1797–1811. doi: 10.1534/g3.114.012377

Schupp, N., and Ziegler, P. (2004). The relation of starch phosphorylases to starch metabolism in wheat. *Plant Cell Physiol.* 45 (10), 1471–1484. doi: 10.1093/pcp/pch170

Sharma, S. K., Bolser, D., de Boer, J., Sønderkær, M., Amoros, W., Carboni, M. F., et al. (2013). Construction of reference chromosome-scale pseudomolecules for potato: Integrating the potato genome with genetic and physical maps. *G3: Genes Genomes Genet.* 3 (11), 2031–2047. doi: 10.1534/g3.113.007153

Shoaib, N., Liu, L., Ali, A., Mughal, N., Yu, G., and Huang, Y. (2021). Molecular functions and pathways of plastidial starch phosphorylase (PHO1) in starch metabolism: current and future perspectives. *Int. J. Mol. Sci.* 22 (19), 10450. doi: 10.3390/ijms221910450

Slugina, M. A., Kochieva, E. Z., and Shchennikova, A. V. (2020a). The opposite effect of low temperature on the pho1a starch phosphorylase gene expression in solanum tuberosum L. Tubers and petota species leaves. *Am. J. Potato Res.* 97 (1), 78–87. doi: 10.1007/s12230-019-09758-z

Slugina, M. A., Kochieva, E. Z., and Shchennikova, A. V. (2020b). Polymorphism and phylogeny of the vacuolar invertase inhibitor gene INH2 homologs in solanaceae species. *Biol. Bull.* 47 (5), 474–481. doi: 10.1134/S1062359020040111

Sonnewald, U., Basner, A., Greve, B., and Steup, M. (1995). A second L-type isozyme of potato glucan phosphorylase: cloning, antisense inhibition and expression analysis. *Plant Mol. Biol.* 27 (3), 567–576. doi: 10.1007/BF00019322

Sonnewald, S., and Sonnewald, U. (2014). Regulation of potato tuber sprouting. *Planta* 239 (1), 27–38. doi: 10.1007/s00425-013-1968-z

Subasinghe, R. M. (2013) *Role and regulation of starch phosphorylase and starch synthase IV in starch biosynthesis in maize endosperm amyloplasts* (University of Guelph) (Accessed 4 April 2022).

Sweetloove, L. J., Burrell, M. M., and ap Rees, T. (1996). Characterization of transgenic potato (Solanum tuberosum) tubers with increased ADPglucose pyrophosphorylase. *Biochem. J.* 320 (2), 487–492. doi: 10.1042/bj3200487

Tang, D., Jia, Y., Zhang, J., Li, H., Cheng, L., Wang, P., et al. (2022). Genome evolution and diversity of wild and cultivated potatoes. *Nature* 606 (7914), 535–541. doi: 10.1038/s41586-022-04822-x

Tetlow, I. J., and Bertoft, E. (2020). A review of starch biosynthesis in relation to the building block-backbone model. *Int. J. Mol. Sci.* 21, 1–37. doi: 10.3390/ijms21197011

Tetlow, I. J., Beisel, K. G., Cameron, S., Makhmoudova, A., Liu, F., Bresolin, N. S., et al. (2008). Analysis of protein complexes in wheat amyloplasts reveals functional interactions among starch biosynthetic enzymes. *Plant Physiol.* 146 (4), 1878–1891. doi: 10.1104/pp.108.116244

Tickle, P., Burrell, M. M., Coates, S. A., Emes, M. J., Tetlow, I. J., and Bowsher, C. G. (2009). Characterization of plastidial starch phosphorylase in Triticum aestivum L. endosperm. *J. Plant Physiol.* 166 (14), 1465–1478. doi: 10.1016/j.jplph.2009.05.004

Tiessen, A., Nerlich, A., Faix, B., Hümmer, C., Fox, S., Trafford, K., et al. (2012). Subcellular analysis of starch metabolism in developing barley seeds using a non-aqueous fractionation method. *J. Exp. Bot.* 63 (5), 2071–2087. doi: 10.1093/jxb/err408

Van Harsselaar, J. K., Lorenz, J., Senning, M., Sonnewald, U., and Sonnewald, S. (2017). Genome-wide analysis of starch metabolism genes in potato (Solanum tuberosum L.). *BMC Genomics* 18 (1), 37. doi: 10.1186/s12864-016-3381-z

Yang, X., Zhang, L., Guo, X., Xu, J., Zhang, K., Yang, Y., et al. (2023). The gap-free potato genome assembly reveals large tandem gene clusters of agronomical importance in highly repeated genomic regions. *Mol. Plant* 16 (2), 314–317. doi: 10.1016/j.molp.2022.12.010

Young, G. H., Chen, H. M., Lin, C. T., Tseng, K. C., Wu, J. S., and Juang, R. H. (2006). Site-specific phosphorylation of L-form starch phosphorylase by the protein kinase activity from sweet potato roots. *Planta* 223 (3), 468–478. doi: 10.1007/s00425-005-0103-1

Zeeman, S. C., Thorneycroft, D., Schupp, N., Chapple, A., Weck, M., Dunstan, H., et al. (2004). Plastidial α-glucan phosphorylase is not required for starch degradation in arabidopsis leaves but has a role in the tolerance of abiotic stress. *Plant Physiol.* 135 (2), 849–858. doi: 10.1104/pp.103.032631

Zhao, X., Jayarathna, S., Turesson, H., Fält, A.-S., Nestor, G., González, M. N., et al. (2021). Amylose starch with no detectable branching developed through DNA-free CRISPR-Cas9 mediated mutagenesis of two starch branching enzymes in potato. *Sci. Rep.* 11 (1), 4311. doi: 10.1038/s41598-021-83462-z

Zhong, Y., Qu, J. Z., Liu, X., Ding, L., Liu, Y., Bertoft, E., et al. (2022). Different genetic strategies to generate high amylose starch mutants by engineering the starch biosynthetic pathways. *Carbohydr. Polymers* 287, 119327. doi: 10.1016/j.carbpol.2022.119327

Acta Universitatis agriculturae Sueciae
Doctoral Thesis No. 2024:48

As more people are switching to a more plant based diet, demand for novel plant based protein products is rising. One crop that has been highlighted as under-utilized in this context is potato. This thesis explores the possibility of using biotechnology to enhance proteins already present in the tubers, to make them accessible to the food industry. In the process, it delves in to available methods for achieving the goal of *in planta* protein engineering.

Martin Friberg received his graduate education at the Department of Plant Breeding, SLU, Alnarp, Sweden. He received his MSc in Biology at Uppsala University.

Acta Universitatis agriculturae Sueciae presents doctoral theses from the Swedish University of Agricultural Sciences (SLU).

SLU generates knowledge for the sustainable use of biological natural resources. Research, education, extension, as well as environmental monitoring and assessment are used to achieve this goal.

ISSN 1652-6880
ISBN (print version) 978-91-8046-036-1
ISBN (electronic version) 978-91-8046-037-8