DISS. ETH NO. 25042

Starch: from metabolism to utilization in orphan crops

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

Presented by
Wuyan Wang

M.Sc. in Horticultural Science,
Technical University of Munich, Germany

Born on 26.12.1988
Citizen of China

Accepted on the recommendation of

Prof. Dr. Samuel C. Zeeman (examiner)
Dr. Zerihun Tadele (co-examinar)
Prof. Dr. Wilhelm Gruissem (co-examinar)
Prof. Dr. Herve Vanderschuren (co-examinar)

2018

Table of Contents

Acknowledgements I

Summary II

Sommario IV

Abbreviations VI

Chapter 1 1

General Introduction 1

1.1 Starch 1

1.1.1 Starch metabolism pathways 1

1.1.1.1 Starch synthesis 1

1.1.1.2 Starch degradation 2

1.1.2 Starch structure, composition and functional properties 5

1.1.2.1 Structure 5

1.1.2.2 Composition 7

1.1.2.3 Physicochemical and functional properties 8

1.1.2.4 Starch applicability 10

1.2 Orphan crops: Cassava (*Manihot esculenta* Crantz) and tef (*Eragrostis tef* (Zucc.) Trotter) 11

1.2.1 Cassava 11

1.2.1.1 Utilization 11

1.2.1.2 Biotechnology application 12

1.2.2 Tef 13

1.2.2.1 Genetic diversity 13

1.2.2.2 Tef grain composition 13

1.2.2.3 Utilization 14

1.3 Starch genetic modification in orphan crops 14

1.4 Aim of the PhD thesis 15

1.5 References 16

Chapter 2 28

Modification of cassava root starch phosphorylation enhances starch functional properties. 28

Chapter 3 68

Improving starch diversity in the orphan crop- tef *(Eragrostis tef)* **68**
3.1 Abstract **69**
3.2 Introduction **70**
3.3 Materials and methods **73**
3.4 Results **76**
3.4.1 Identification of starch related genes in tef 76
3.4.2 Expression profile of starch related genes in tef developing seeds 83
3.4.3 TILLING 88
3.4.3.1 Mutagenesis stage 89
3.4.3.2 Mutation detection stage 89
3.5 Discussion **109**
3.6 References **112**

Chapter 4 ..... **124**
General Discussion **124**
4.1 Investigation of starch metabolism in storage organs **124**
4.2 Biotechnology for improving orphan crops **125**
4.3 Potential application of modified starch **127**
4.4 Importance of increasing starch research in orphan crops **128**
4.5 References **129**

Acknowledgements

I would like to express my appreciation to my supervisor Prof. Dr. Samuel C. Zeeman for his support of my PhD study and research, for his patience, encouragement, immense knowledge, valuable guidance and supervision. I gratefully thank Dr. Zerihun Tadele for his significant cooperation, resource sharing and suggestions. I also gratefully thank Prof. Dr. Wilhelm Gruissem and Prof. Dr. Herve Vanderschuren for their valuable suggestions in committee meetings.

I sincerely thank all members in Plant Biochemistry for the work, the help, the discussion and the enjoyable trips together. Special thanks to our "cassava harvesting team" with the core members: Andrea Ruckle, Gabriel Deslandes and Laure David for their helps of harvesting my cassava materials. Thanks to Federica Assenza for her translation of my thesis summary into Italian. Thanks to Elisabeth Truernit and Michaela Fischer-Stettler for their contribution of time to discuss about my thesis. A great thank to Carmen Hostettler who initiated the cassava project leading to the available materials I could work during my PhD.

I kindly thank all members in Dr. Zerihun Tadele's group in University of Bern for their super support in the tef project. I enjoyed doing experiments in their lab and visiting them in Bern. I would like to specially thank Regula Bloesch, Sonia Plaza Wüthrich and Gina Cannarozzi, who gave me valuable helps for establishing the tef project at the beginning.

I also thank the great help from Matthias Hirsch-Hoffmann, Emily Jane Mc Callum, Devang Mehta and other helpers in Plant Biotechnology group ETH Zurich.

Furthermore, I would like to thank my PhD funding program-IDPBRIDGES, where I learnt a lot about the linkages between science and policy and also impressed by our secondment in Ethiopia. Thank all IDP fellows and our program coordinators: Andrea Pfisterer and Luisa Last.

Last but not least, I would like to thank my parents, my brother and my boyfriend who give me measureless support and love.

Summary

Starch is the most widespread and abundant storage carbohydrate produced in plants and valued for its use in food, feed, fuel and industry. Advance of genetics, biochemistry and biotechnology enabled detailed investigation of how starch is synthesized and degraded in the leaves of the model plant *Arabidopsis thaliana*. Our understanding of these processes in seeds, root and other organs of starch crops has also increased. Starch properties are influenced by many factors. These include the relative amounts of the two glucan polymers amylopectin and amylose, the branched structure of amylopectin, starch granule size and the presence of covalent modifications. Based on the current knowledge about starch, biotechnological applications have arisen that led to a number of transgenic crops with the altered starch content or improved starch quality, such as in rice (*Oryza sativa*), wheat (*Triticum aestivum*), maize (*Zea mays*) and potato (*Solanum tuberosum*).

Unlike those major crops, orphan crops such as cassava (*Manihot esculenta* Cranz.), tef (*Eragrostis tef* (Zucc.) Trotter), yam (*Dioscorea sp*. L.) and finger millet (*Eleusine coracana* (L.) Gaertn.) are under-researched. Programs to improve their performance have received little public or private investment, even though these crops play particularly important roles in agro-ecology and social-economy in many developing countries. Implementation of modern technologies to boost the productivity of such orphan crops are needed. Here in my research I focused on cassava (*Manihot esculenta* Crantz) and tef (*Eragrostis tef* (Zucc.) Trottter).

Cassava, as a root crop, is used as a foodstuff and as a starch source in industry. The aim of this part of my work was to explore the potential to alter starch phosphorylation in the cassava storage root. Transgenic plants were generated in which the potato glucan, water dikinase gene (*StGWD*) was overexpressed, or in which each of the endogenous phosphoglucan phospahatse genes (*MeSEX4* and *MeLSF2*) were silenced by RNAi. Overexpressing the potato *GWD* gene (*StGWD*), which specifically phosphorylates the glucosyl residues of amylopectin at the C6 position, increased the total starch-bound phosphate content at both the C3 and the C6 positions. Silencing of the *MeLSF2* gene, the protein product of which specifically dephosphorylates the C3 position of the glucosyl residues of amylopectin, increased the ratio of C3:C6 phosphorylation without altering total phosphate content. In both cases, amylopectin structure, amylose content and starch granule size were unaltered. I tested several physicochemical properties of the starch extracted from transgenic and control lines. Starch swelling power and paste clarity were both influenced by total phosphate content. However, phosphate position did not significantly influence the functional properties. In conclusion, I found that manipulating the expression of starch phosphorylating/dephosphorylating enzymes can be used to engineer starch phosphate levels in cassava storage root, potentially increasing the value of starch for the food and non-food industries.

Tef (*Eragrostis tef* (Zucc.) Trottter) is the most importance cereal in Ethiopia. It is highly appreciated for its nutritional properties and has the advantage of being naturally gluten-free.

Yet, it still receives very little attention from the global research network. Starch comprises more than 70% of tef grain and is the primary source of calories served in Ethiopian diets. Tef is also used in other ways in the food and beverage industry. Unlike rice, maize, barley and other major crops, the starch metabolism in endosperm of tef has not been researched and the existing tef cultivars lack diversity in starch traits. The aim of this part of my research was to investigate the starch metabolism and to improve tef by diversifying its starch properties. I first identified tef genes involved in starch synthesis in the endosperm using comparative genomic analysis. I probed the tetraploid tag genome with sequences from the closely-related diploid model species, rice (*Oryza sativa*) and sorghum (*Sorghum bicolor*). I then confirmed that the targeted genes were expressed in developing tef seeds through RNA-sequencing analysis. Based on the conventional reverse genetics approach Targeting Induced Local Lesions IN Genomes (TILLING), I established a new TILLING-by-sequencing method, using high throughput sequencing and bioinformatics analysis. I used this method with the aim of generating tef plants with mutations in identified starch genes. A database containing mutations for each starch-related gene has been created. It will now be possible in the longer-term to conduct the screening of mutants, followed by the cultivation and crossing of plants in order to achieve useful null-mutant plants. I expect in the future to see new tef varieties with improved starch properties to be brought into breeding programs for the benefit of Ethiopian farmers and use in both food and non-food industries.

My research has highlighted the potential of starch bioengineering in orphan crops, resulting in modified starches with numerous applications. This success should stimulate further starch research in orphan crops using modern technologies like those I employed.

Sommario

L'amido è il carboidrato di riserva più diffuso e abbondante prodotto dalle piante ed è apprezzato per il suo impiego nell'alimentazione umana ed animale, come combustibile e in svariate applicazioni industriali. Il progresso nel campo della genetica, della biochimica e delle biotecnologie ha consentito la dettagliata indagine di come l'amido è sintetizzato e degradato nelle foglie della specie modello *Arabidopsis thaliana*. È migliorata anche la nostra comprensione di tali processi in semi, radici e altri organi di specie coltivate (*crop*) ad alta produzione di amido. Le proprietà dell'amido sono influenzate da molti fattori, tra cui la quantità relativa dei polimeri glucidici amilopectina e amilosio, la struttura ramificata dell'amilopectina, le dimensioni dei granuli di amido e la presenza di modificazioni covalenti. Sono sorte una serie di applicazioni biotecnologiche, basate sulle attuali conoscenze sull'amido, che hanno portato alla generazione di varianti transgeniche di *crop* caratterizzate da quantità di amido alterate o dalla sua migliore qualità, tra queste riso (*Oryza sativa*), frumento tenero (*Triticum aestivum*), mais (*Zea mays*) e patata (*Solanum tuberosum*).

Al contrario delle *crop* principali succitate, le *crop* orfane, tra cui cassava (*Manihot esculenta* Cranz.), tef (*Eragrostis tef* (Zucc.) Trotter), yam (*Dioscorea sp*. L.) e miglio indiano (*Eleusine coracana* (L.) Gaertn.) sono poco studiate. Programmi per il miglioramento delle loro prestazioni hanno ricevuto scarsi investimenti pubblici o privati, sebbene queste specie giochino un ruolo di particolare rilievo nell'agroecologia e nell'economia-sociale di molti paesi in via di sviluppo. L'utilizzo di tecnologie moderne per incrementare la produttività di tali specie orfane è necessario. Con la mia ricerca mi sono concentrata su cassava (*Manihot esculenta* Crantz) e tef *(Eragrostis tef* (Zucc.) Trottter).

La cassava, una specie da radice, è utilizzata come alimento e come fonte di amido per l'industria. L'obiettivo di questa parte del mio lavoro è stato quello di esplorare il potenziale di alterazione della fosforilazione dell'amido nella radice tuberizzata di cassava. Sono state generate piante transgeniche in cui il gene di patata codificante la glucano acuqa dichinase (glucan, water dikinase, *StGWD*) è stato over-espresso o in cui ciascuno dei geni endogeni della fosfoglucano fosfatasi (phosphoglucan phospahatse, *MeSEX4* and *MeLSF2*) è stato silenziato tramite RNAi. Over-espressione del gene di patata GWD (*StGWD*), che fosforila specificamente il residuo glucidico dell'amilopectina in posizione C6, ha portato all'aumento del contenuto totale di fosfato legato all'amido, sia in posizione C3 che in posizione C6. Il silenziamento del gene *MeLSF2*, il cui prodotto proteico defosforila specificamente la posizione C3 dei residui glucidici dell'amilopectina, ha determinato l'aumento del rapporto C3:C6 senza alterare il contenuto totale in fosfato. In entrambi i casi la struttura dell'amilopectina, il contenuto in amilosio e la dimensione dei granuli di amido sono rimasti inalterati. Ho testato numerose proprietà fisico-chimiche dell'amido estratto da linee transgeniche e linee controllo. Il potenziale di rigonfiamento (swelling power) e la limpidezza della pasta (paste clarity) sono risultate influenzate dal contenuto di fosfato. Tuttavia, la posizione del fosfato non ha determinato una differenza significativa delle proprietà funzionali.

In conclusione, ho riscontrato che manipolare l'espressione degli enzimi di fosforilazione/defosforilazione dell'amido è un approccio valido per ingegnerizzare i livelli di fosfato dell'amido della radice tuberizzata di cassava, con il potenziale di accrescere il valore di questo polimero per l'industria alimentare e non alimentare.

Tef (*Eragrostis tef* (Zucc.) Trottter) è il più importante cereale in Etiopia. È altamente apprezzato per le sue proprietà nutrizionali e per il vantaggio di non contenere glutine. Tuttavia riceve ancora poca attenzione da parte della rete globale della ricerca. L'amido costituisce più del 70% del chicco di tef ed è la fonte primaria di calorie nella dieta etiope. Il tef ha anche altri utilizzi nell'industria alimentare e delle bevande. Al contrario di riso, mais, orzo e altre *crop* principali, sul metabolismo dell'amido nell'endosperma di tef non sono state ancora condotte ricerche e le cultivar esistenti sono carenti nella diversità dei caratteri relativi all'amido. L'obiettivo di questa parte della mia ricerca è stato quello di investigare il metabolismo dell'amido e di migliorare il tef diversificando le proprietà del suo amido. Come prima cosa, tramite analisi di genomica comparativa, ho identificato i geni di tef convolti nella sintesi dell'amido nell'endosperma. Ho sondato il genoma tetraploide di tef utilizzando sequenze geniche delle specie imparentate riso (*Oryza sativa*) e sorgo (*Sorghum bicolor*). Successivamente, attraverso analisi di sequenziamento dell'RNA ho confermato che i geni target fossero espresso nei semi di tef in via di sviluppo. Sulla base dell'approccio convenzionale di *reverse genetics* detto Targeting Induced Local Lesions IN Genomes (TILLING), ho messo a punto un nuovo metodo di TILLING-by-sequencing, usando high throughput sequencing e analisi bioinformatiche. Ho utilizzato questo metodo con l'obiettivo di ottenere piante di tef con mutazioni in specifici geni convolti nel metabolismo dell'amido. È inoltre stato creato un database di mutazioni in ciascun gene connesso all'amido. Da ora, e nel lungo periodo, sarà possibile effettuare lo screening di mutanti, seguito dalla loro coltivazione e incrocio così da ottenere piante con alleli nulli utili. In futuro mi aspetto di vedere accedere nuove varietà di tef con migliorate proprietà dell'amido a programmi di breeding a beneficio degli agricoltori etiopi e per l'utilizzo nell'industria alimentare e non.

La mia ricerca ha messo in evidenza il potenziale della bio-ingegnerizzazione dell'amido in *crop* orfane allo scopo di ottenere amidi modificati con numerose applicazioni. Tale successo dovrebbe stimolare l'ulteriore ricerca nell'ambito dell'amido in *crop* orfane attraverso l'impiego di tecnologie moderne come quelle da me utilizzate.

Abbreviations

AGPase	ADPglucose pyrophosphorylase
AMY	Alpha-amylase
BAM	Beta-amylase
CaMV	Cauliflower mosaic virus
CBB	Cassava bacterial blight disease
CBM	Carbohydrate-binding module
CBSV	Cassava brown streak virus
CLD	Chain-length distribution
CMV	Cassava mosaic virus
DBEs	Debranching enzymes
DPE1	Disproportionating enzyme 1
DSC	Differential scanning calorimetry
EMS	Ethylmethansulfonat
ESV1	Early starvation1
FEC	Friable embryogenic calli
GATK	Genome analysis tool kit
GBSS	Granule-bound starch synthases
Glc1P	D-glucose-1-phosphate
GWD	Glucan, water dikinase
HPAEC-PAD	High performance anion exchange chromatography with pulsed amperometric detection
ISA	Isoamylase
LDA	Limit dextrinase
LESV	Like ESV1
LSF2	Like sex four 2
MEX1	Maltose excess protein 1
pGlcT	Plastidic glucose translocator transporter
PHS1	Plastidial α-glucan phosphorylase 1
PTST	Protein targeting to starch
PWD	Phosphoglucan, water dikinase
RSR1	Rice starch regulator 1
RVA	Rapid visco analysis
SBEs	Starch-branching enzymes
SE	Somatic embryogenesis
SEX4	Starch excess 4
SSs	Starch synthases
TILLING	Targeting induced local lesions in genomes

Chapter 1
General Introduction

1.1 Starch
1.1.1 Starch metabolism pathways
1.1.1.1 Starch synthesis

Starch constitutes a major energy supply for humans and is produced as a reserve carbohydrate in plants, being stored leaves (transitory starch), in seeds, fruits, roots or other organs. Starch (both transitory and storage starch) is composed of the branched polymer, amylopectin and the near-linear polymer, amylose. The synthesis of amylopectin and amylose are well researched in chloroplasts of model plant *Arabidopsis thaliana*. Amylopectin synthesis involves three major classes of enzymes: soluble starch synthases (SSs), which elongating non-reducing ends of glucose chains using substrate adenosine 5′-diphosphate-glucose (ADPglucose), starch branching enzymes (SBEs), which generating branches from existing chains via glucanotransferase reactions, and starch debranching enzymes (DBEs) which selectively debranch mis-placed chains (Pfister and Zeeman 2016a).

Amylopectin chain elongation is catalyzed by soluble starch synthases (SSs). SSI and SSII preferentially elongate short single-cluster-filling chains (i.e., the A- and B1-chains), while SSIII elongates longer cluster-spanning B chains (Tetlow and Emes 2011). SSIV is less involved in the amylopectin structure, but function in starch granule initiation and granule morphology (Crumpton-Taylor et al. 2013). The complex of SSs is formed and the contribution of each enzyme varies in different plants and tissues (Hennen-Bierwagen et al. 2008; Hennen-Bierwagen et al. 2009).

Amylopectin synthesis is continued with branching, catalyzed by starch branching enzymes (SBEs), cutting an existing α-1,4-linked chain and transferring the cut segment to another linear chain to create a new α-1,6 linked branch. There are two subclasses of SBEs, SBEI (family B) and SBEII (family A). SBEIs preferentially transfer longer chains than SBEIIs. SBEI enzymes generally occur as single isoform, with the exception of Arabidopsis, which contains no SBEI (Dumez et al. 2006). SBEII has single isoform in potato and pea, but two isoforms (SBEIIa and SBEIIb) in the cereals. The expression of SBEIIb is restricted to the grain (such as in maize, rice and barley), while SBEIIa is found in all tissues but with the lower expression levels (Gao et al. 1996; Pfister and Zeeman 2016a). Reduction of SBEII activity altered starch content, structure and properties, while a lack of SBEI resulted only in minor differences compared to wild type, e.g. in potato and maize plants. Reducing the activity of SBEIIs results in a starch phenotype called *amylose-extender (ae)*, in which the starch has amylopectin with longer external and internal chains and less branching, and the amylose content is increased in some cases. *ae* starch has changed physio-chemical properties, such as swelling, gelatinization and digestibility (Pfister and Zeeman 2016a).

Amylose is synthesized by granule-bound starch synthase (GBSS), one of the starch synthases (SSs), which as its name suggests becomes exclusively bound within the granule matrix formed by amylopectin (Tatge et al. 1999; Pfister and Zeeman 2016a). Two GBSS isoforms are currently identified in cereal endosperm, GBSSI and GBSSII. The role of GBSSI is mostly confined to storage tissues, whereas GBSSII also functions in non-storage plant tissues like leaves, where transitory starch accumulates (Vrinten and Nakamura 2000). A novel protein, Protein Targeting To Starch (PTST), has been identified recently that influence amylose synthesis through affecting the delivery of GBSS to starch granule. PTST is a plastidial protein with a family 48 carbohydrate binding module (CBM48) and coiled-coil motifs for protein-protein interactions. It binds to GBSS through coiled-coil mediated interaction and this interaction enables GBSS to be targeted to starch granules via the CBM (Seung et al. 2015).

The complete amylopectin synthesis also requires debranching enzymes (DBEs). Plants lacking a particular type of DBEs have serious defects in starch biosynthesis. Two classes of DBEs, isoamylases (ISA) and limit-dextrinases (LDA; also called pullulanase), hydrolyze α-1,6 branch points. Each enzyme has a preference for different substrates (Zeeman et al. 2010b). LDA has a preference for substrates with very short branches such as in the yeast polysaccharide pullulan. ISA cannot degrade pullulan, but are more active on amylopectin-like substrates. ISA has three classes, ISA1, ISA2 and ISA3. ISA1 can form a heteromultimeric enzyme together with ISA2 and, in the cereals, can also form homomultimers by itself. The heteromultimeric complex is involved in debranching amylopectin during its synthesis, while ISA3 and LDA are specialized for branch point hydrolysis during starch breakdown (Streb and Zeeman 2012).

In recent years, several novel proteins have been identified that influence starch synthesis. One designated Early Starvation1 (ESV1) is a plastidial starch-binding protein. Mutants of *ESV1* show premature depletion of transitory starch at night and overexpressing the protein caused a starch-excess phenotype. While the precise molecular function is not known, it was proposed to help glucan molecules align correctly within the granular matrix. (Feike et al. 2016). Like ESV1 (LESV), was identified as a homolog of ESV1. LESV overexpression lead to a low-starch phenotype, similar to that of *esv1* mutants (Feike et al. 2016). Rice Starch Regulator1 (RSR1), an APETALA2/ethylene-responsive element binding protein family transcription factor, was shown to regulate the expression of a series of starch-biosynthetic genes in rice grains, thereby influencing starch structure and granule packing (Fu and Xue 2010).

1.1.1.2 Starch degradation

Starch degradation is best understood in leaves, where the transitory starch is degraded at night. The pathways of starch degradation in chloroplast is via a network of reactions rather than a linear pathway. In general, reversible glucan phosphorylation initially solubilizes starch granule surface, enables the hydrolases access to the glucan chains to perform degradation and the major degraded products, maltose and glucose, are then exported to the cytosol (Figure 1).

Starch phosphorylation is catalyzed by two dikinases: α-glucan water, dikinase (GWD) and Phosphoglucan, water dikinase (PWD). GWD phosphorylates glucosyl residues of amylopectin at the C6 position, while PWD specifically phosphorylate a different glucosyl residue at the C3 position after the pre-phosphorylation by GWD (Lorberth et al. 1998; Yu et al. 2001; Baunsgaard et al. 2005; Kötting 2005). The *gwd1* mutant of Arabidopsis (also called *sex1*) has a strong starch excess (*sex*) phenotype while *pwd* mutants have more mild *sex* phenotypes (Lorberth et al., 1998; Yu et al., 2001; Weise et al. 2012). Phosphate groups are proposed to locally disrupt the packing of amylopectin double helices. Phosphates are removed from glucan chains by two phosphoglucan phosphatases, Starch EXcess 4 (SEX4) and Like Sex Four 2 (LSF2). This is proposed to take place concurrently with glucan hydrolysis by amylolytic enzymes. SEX4 releases phosphate bound either to the C6 or the C3 positions, with a preference for C6, whereas LSF2 has a very strong preference for phosphate bound to the C3 position. A reduced rate of starch degradation, the accumulation of excess leaf starch and the appearance of soluble phospho-oligosaccharides are all phenotypic aspects of *sex4* mutants. In contrast, *lsf2* mutants have little or no *sex* phenotype and nor do they accumulate phospho-oligosaccharides. However, the starch of *lfs2* mutants has a marked increase in C3-bound phosphate. (Zeeman et al. 1998; Kötting et al. 2009; Santelia et al. 2011).

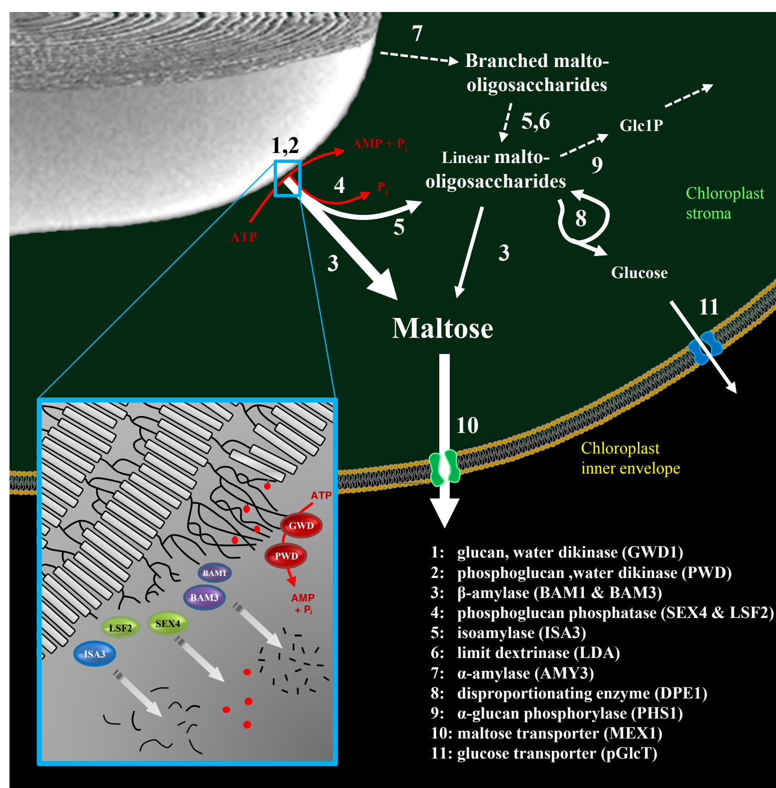


Figure 1. The pathway of starch degradation in chloroplasts. (Figure from Streb and Zeeman 2012)
Maltose and malto-oligosaccharides are released from the surface of the starch granule during degradation. Malto-oligosaccharides are metabolized in the stroma. Maltose and glucose are exported to the cytosol. Inset is a model depicting the role of phosphorylation by GWD1 and PWD in disrupting the packing of amylopectin double helices (grey boxes). This allows the release of maltose and malto-oligosaccharides (black lines) by β-amylases (BAMs) and DBE (ISA3). Phosphate (red dots) is concomitantly released by SEX4 and LSF2 to allow complete degradation.

Glucan degradation is performed by the hydrolytic enzymes to degrade the α-1,4- and α-1,6-bonds. β-Amylase (BAMs) is one of the key enzymes of starch degradation, exo-acting on linear chains to release maltose from non-reducing end. In Arabidopsis, there are nine genes encoding β-amylase-like proteins, at least four of which (BAM1 to BAM4) are targeted to the chloroplast (Fulton et al. 2008). The α-1,6 branch points are further hydrolyzed by DBEs, described earlier. Of the four genes encoding DBE-like proteins in Arabidopsis, ISA3 and LDA are critical for starch degradation (Delatte et al. 2006). The well-known enzyme α-amylases (AMY) is also present in the chloroplast. It acts on internal α-1,4-bonds to release a variety of linear and branched oligosaccharides and plays a key role in the degradation of storage starch (e.g. in the endosperms of germinated cereal seeds) but it appears to have a less important role in leaves (Yu et al. 2005).

The combined actions of β-amylase, α-amylase and DBEs will result in the release of maltose and longer malto-oligosaccharides into the chloroplast stroma. The malto-oligosaccharides can be further metabolised by β-amylases to yield maltose or maltotriose or by α-glucan phosphorylase to yield Glc1P and maltotetraose. Short malto-oligosaccharides like maltotriose are also metabolized by the glucanotransferase disproportionating enzyme (DPE1) (Critchley et al. 2001). DPE1 activity, together with that of β-amylase and phosphorylase will result in the production of maltose, glucose and Glc1P. The plastidial α-glucan phosphorylase PHS1 (PHO1 in other species) actually catalyses a reversible reaction and can also release orthophosphate (P_i) through reversible addition of the glucosyl moiety of Glc1P to the non-reducing end of an acceptor glucan. Its precise function in starch metabolism is unclear (Rathore et al. 2009). Glucose and maltose are exported to the cytosol by maltose transporter (MEX1) and glucose transporter (pGlcT) (Weber et al. 2000; Niittylä et al. 2004).

During seed germination, tuber sprouting, or fruit ripening, starch degradation occurs and the derived metabolites are used as source energy for regrowth. Unlike in Arabidopsis chloroplasts, the starch breakdown pathways in non-photosynthetic tissues are distinct. For instance, during cereal seed germination, it is generally accepted that α-amylase plays a central role in endosperm starch degradation (Sun and Henson 1991). The α-amylase protein is synthesized in the aleurone layers and scutellum, followed by secretion into the non-living starchy endosperm. Its endoamylolytic action releases a mixture of soluble linear and branched oligosaccharides. In the case of β-amylase, it is laid down within the endosperm during seed development and converted to active forms during germination by the action of proteases released from the aleurone (Kaneko et al. 2002; Henson 2014). Together with the function of debranching enzymes and maltases (α-glucosidase), the end product of starch degradation in the endosperm is glucose (Sun and Henson 1991). The produced glucose is taken up by living cells in the embryo adjacent to the endosperm (the scutellum). The total process of starch hydrolysis in storage organs is still not clear and more research is required.

1.1.2 Starch structure, composition and functional properties

1.1.2.1 Structure

Wild-type starches consisting of 75-90% amylopectin and 10-25% amylose. Both have chains of α-1,4-linked glucose residues that are branched with α-1,6 linkages. Amylopectin is extensively branched and has comparatively short chains whereas amylose is mostly linear or lightly branched and has long chains with several hundreds or even thousands of glucosyl units (Figure 2B) (Manners 1989). The external chains of amylopectin that carry no branches themselves are called A-chains. The chains carry one or more branches are B-chains. The single B chains with a reducing end is called the C-chain. The neighboring linear chain segments

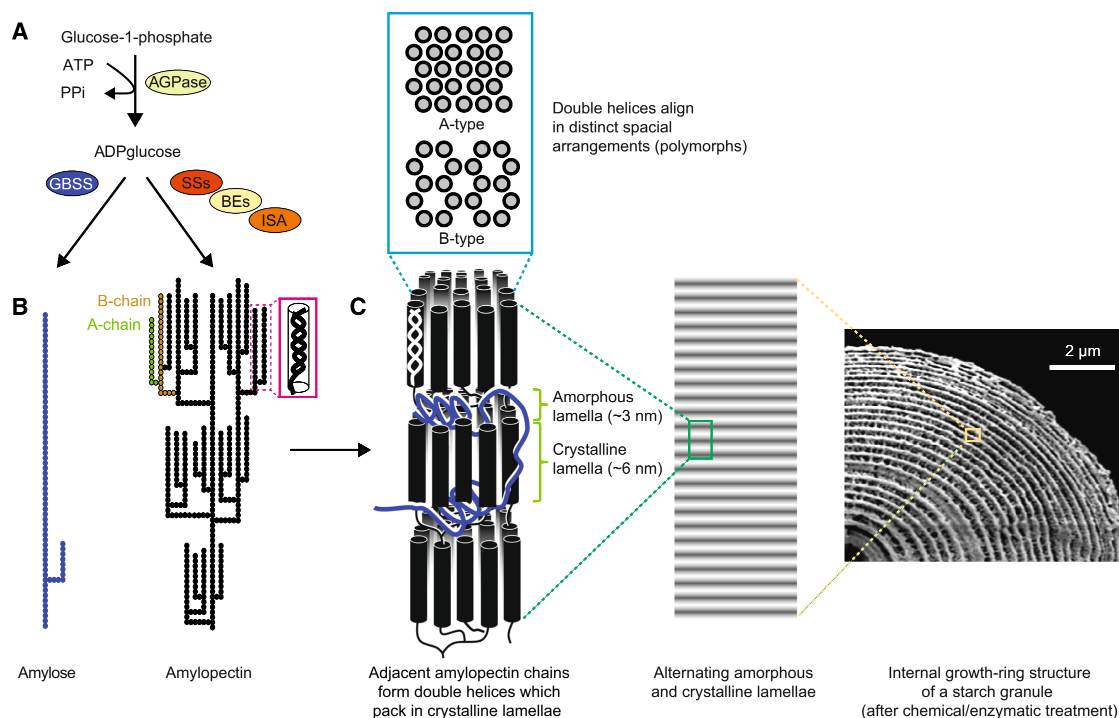


Figure 2 The structure and biosynthesis of starch. Figure from (Pfister and Zeeman 2016a) The structure and biosynthesis of starch. a Overview of the core starch biosynthesis pathway. ADPglucose pyrophosphorylase (AGPase) produces ADPglucose, the substrate of starch synthases (SSs). Granule-bound starch synthase (GBSS) synthesizes amylose, while soluble SSs, branching enzymes (BEs) and isoamylase-type debranching enzyme (ISA) collectively synthesize amylopectin. b Molecular structure of amylose and amylopectin (according to the cluster model), showing its branching pattern and formation of secondary structures. Filled, joined circles represent individual glucosyl residues. c High-order alignment of amylopectin double helices. Each growth ring (right) has a thickness of ca. 200–400 nm and contains a semi-crystalline region and an amorphous region. The semi-crystalline region consists of alternating crystalline lamellae (containing the linear parts of the chains) and amorphous lamellae (containing most of the branch points) which stack with a periodicity of ~9–10.5 nm (middle). Depending on the exact architecture of the amylopectin giving rise to the clusters, the double helices either arrange as densely packed A-type polymorph or less dense hexagonal B-type polymorph (top). A mixture of A and B is also possible and named C-type polymorph (not shown).

entwine to form clusters of parallel double helices, enabled by the branching frequency and pattern, and the distribution of chain lengths (Figure 2B). Clusters stack in crystalline lamellae with distinct spacial arrangements (polymorphs); either the dense A-type polymorphs, the less

dense B- polymorphs or the C-type polymorphs (mix of A and B) (Figure 2C) (Imberty et al. 1988; Imberty and Perez 1988). In between crystalline lamellae are amorphous lamellae containing most of the branch points and this is thought to be where amylose is formed (Figure 2C). The alternating crystalline and amorphous lamellae stack with a periodicity of ~9–10.5 nm constituting concentric semi-crystalline regions. These alternate with amorphous regions with lower degrees of order, yielding a repeating layered structure of semi-crystalline and amorphous regions, which appear as an internal growth-ring structure (Figure 2C) (Gallant et al. 1997; Pilling and Smith 2003; Pfister and Zeeman 2016a).

Starch granules have a board variation in size (0.5–100 µm in diameter), shape (round, lenticular, polygonal), size distribution (uni- or bi-modal) and association as individual (simple) or granule clusters (compound), depending on the species and tissues from which they derive. Starch granules from wild-type Arabidopsis leaves are irregularly discoid in shape and approximately 1-2 µm in diameter at the end of normal photoperiod (Zeeman 2002). A comparison of the size and morphology of starch granules from tef and cassava (the focus of my studies) with those of other well-studied crops, including rice, wheat, potato and Arabidopsis are shown in Figure 3. Starch from tef grain is most similar to that of rice, forming compound granules comprised of many sub-granules polygonal in shape (2–6 µm in diameter)(Bultosa et al. 2002). This is not the same as wheat starch, which has a bimodal distribution of large ‘A’ and small ‘B’ granules (Figure 3). Cassava root starch granules differed from those of potato tuber, being considerably more irregular in shape with oval, round and flat-ended granules, with a mean size of about 15 µm (Figure 3). Interestingly, it was reported that cassava storage root starch had both unimodal and bimodal granule size distribution and the bimodal distribution was more obvious when harvested very late during the growing season (Sriroth et al. 1999; Zhu 2015).

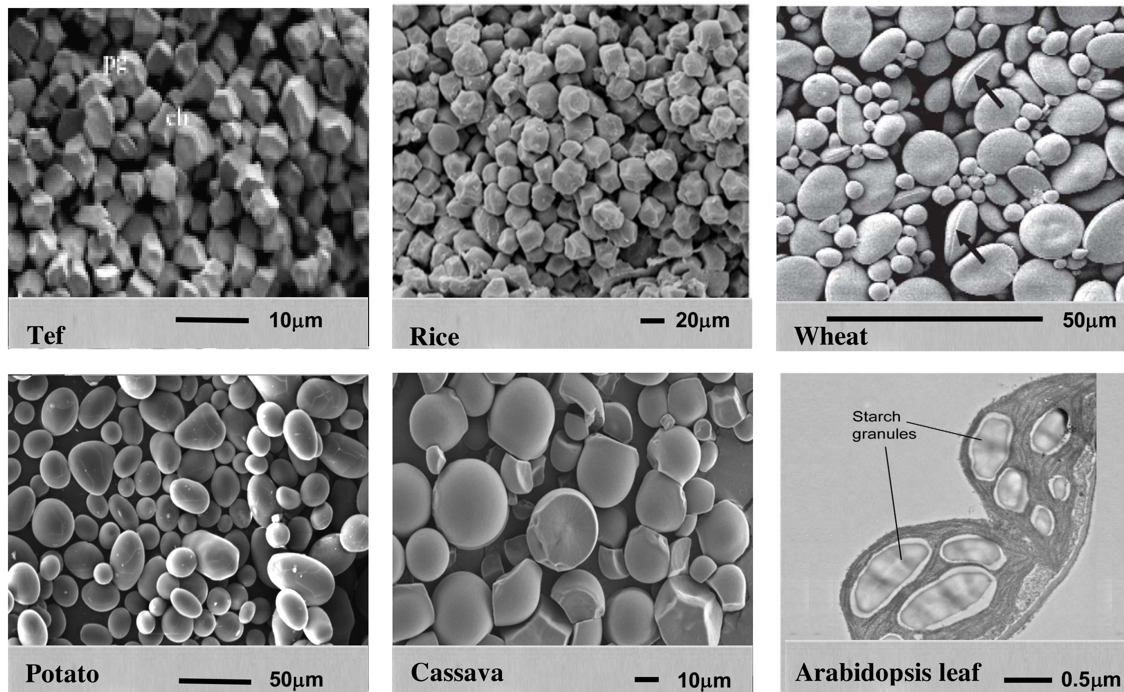


Figure 3 Morphology of tef, rice, wheat, potato, cassava and Arabidopsis leaf starch granules
Scanning electron microscopy of tef starch granule (Bultosa et al. 2002), of rice starch granule (Wang et al. 2015b), of wheat starch granule (Smith 2010), of potato starch granule (Szymońska et al. 2009). Transmission electron microscopy of Arabidopsis leaf starch (Smith 2010).

1.1.2.2 Composition

Starch granules are synthesized within plastids (chloroplasts in leaves and amyloplasts in non-green tissues). In addition to glucans, starch granule also contains protein, lipid and phosphate groups (Zhu 2015; Zhu 2017).

Protein

Starch contains protein (typically 0.1–0.7 %), some of which are embedded within the matrix of granules and some of which occur on the granule surface. Proteins located in the interior part of starch granules include a specific subset of plastid proteins and have different functions, such as starch biosynthetic enzymes (e.g., GBSSI, starch synthases, starch branching enzymes) and starch degradative enzymes (e.g., β-amylase) (Grimaud et al. 2008). Surface proteins of wheat and maize starches may also include storage proteins and they are thought to have the effect of restricting the swelling of starch granules during pasting (Tester et al. 2004).

Lipids

Lipids are also present in normal cereal starches as a minor component (<1.5%) in the form of phospholipids and free fatty acids. Root and tuber contain little or no lipid (Lim et al. 1994). Lipids have been reported to form inclusion complex with amylose in native starches, and the lipid content is positively correlated with the amylose content in cereals, with

amylose-free (or 'waxy') starches having very low contents of lipids (Morrison et al. 1984; Morrison et al. 1993).

Phosphorus

Phosphorus is also found in starch in three major forms: phosphate monoesters covalently linked to glucans, the aforementioned phospholipids and inorganic phosphate. The phosphate monoesters are selectively bound at the C6 or C3 position of glucose residues within the amylopectin molecules (Hizukuri et al. 1970; Lim et al. 1994). The starch phosphate level varies among species (Lim et al. 1994; Singh et al. 2003). In Arabidopsis leaf starch, the phosphate level is around 0.05 % (i.e., around one per 2000 glucose units is phosphorylated), while in tuber starches it can be many times higher (e.g. ~0.5 % in potato). In cereal starches, the phosphate level is extremely low. Phosphorylation is the only known naturally occurring covalent modification of starch. Phosphorylation has an important role in starch metabolism (see above) but also strongly influences the physical and functional properties of starch (Singh et al. 2003). A high phosphate content is an industrially relevant trait as it is associated with an increased granule hydration and lowered crystallinity, yielding starch pastes with higher transparency, viscosity and freeze–thaw stability (Alcázar-Alay and Meireles 2015).

1.1.2.3 Physicochemical and functional properties

When starch is heated in the presence of sufficient water, granules start to swell and the crystalline regions are disrupted. This is known as gelatinization. When the gelatinized starch is cooled down, the aggregation of starch molecules occurs and when stored at cool temperature for a longer time, recrystallization happens (Hari et al. 1989; Tako et al. 2014; Wang et al. 2015a). This is known as retrogradation. It was demonstrated that the thermal properties of starch during those processes are significantly influenced by the factors mentioned above (i.e. molecular structure, granule architecture and the composition of starch) (Hoover 2001; Singh et al. 2007). The related physicochemical and functional properties of starch are listed as follows:

Swelling power

During the inanition of heating, starch granules start to absorb water and swell, and the key functional properties at this phase are hydration, swelling, and solubility. Swelling is determined by the calculation of a ratio between the weight of the swollen granules and its initial dry weight, named as swelling power. Swelling power indicates the water holding capacity of starch, which is important to the quality and texture of some food products (Kaur et al. 2011). Amylose/amylopectin ratio, granule morphology, amylopectin crystalline structure and non-carbohydrate constituents such as proteins, lipids and phosphate monoesters have been shown to influence the swelling pattern of starch granule (Tester and Morrison 1990; Srichuwong et al. 2005; Karim et al. 2007; Vamadevan and Bertoft 2015).

Gelatinization

Gelatinization is an endothermic reaction where the starch granule loses its crystalline structure and become amorphous after heating in water (Ratnayake and Jackson 2008). Differential scanning calorimetry (DSC), one of the techniques, has been used to monitor this phase transition. DSC measures gelatinization temperatures (onset [T_o], melting [T_m], and conclusion [T_c]), and the enthalpy (ΔH) of gelatinization). T_m is generally used as a measure of crystalline perfection and ΔH represents the rupture of the H-bonds between glucan strands (loss of double helical order). A high gelatinization temperature and a narrow endothermic peak suggest a higher molecular order or more stable crystals (Lund 1984). Gelatinization is related to the structural type of the amylopectin component in the starch granule. Amylose content and lipid content have also been shown to affect the gelatinization parameters (Fredriksson et al. 1998; Liu et al. 2006).

Pasting

When the gelatinized starch is continuously heated in excess water with constant stirring, starch molecules eventually disperse in the aqueous medium. The viscosity development in this process is known as the pasting property of starch, which plays an important role in estimating the cooking, eating, and processing quality (Yan et al. 2011). Rapid visco analyser (RVA) is the instrument most commonly employed for measuring the pasting properties of starch. Pasting curves are generated to show the change in viscosity as a function of temperature and time (pasting temperature [PT], peak viscosity [PV], breakdown [BD], trough viscosity [TV], setback [SB], and final viscosity [FV]). Peak viscosity is the highest viscosity level reached during the heating cycle, normally used to indicate the pasting characters from different starch sources (Deffenbaugh and Walker 1989). Studies have shown that the peak viscosity of starch is influenced by its amylose content, the proportion of long amylopectin branch chains, the granular morphology, the friction between the swollen granules, the phosphate monoester content, and sugars (Kim and Walker 1992; Liang and King 2003; Kurakake et al. 2009; Polnaya et al. 2013).

Gelling

On cooling, starch paste starts to form gel. Strength of a starch gel can be determined using a texture analyzer or a rheometer. Normally cereal starches such as maize, wheat, rice, and pea, are known to develop strong gels, whereas, *waxy* starch and tuber starch such as from potato fail to form gels or only form weak gels at the same starch concentration. The strength of a starch gel generally increases with the amylose content of starch because the linear amylose molecules can form long and strong junction zones in the gel network. Therefore, starch having a relatively large concentration of amylose such as indica rice is chosen for the production of strong gels (Donald 2004).

Recrystallization and retrogradation

During storage for an extended period of time, the formed gel then recrystallizes gradually into semi-crystalline aggregates that differ in form from the native granules. This is called

retrogradation, the process where the glucan molecules dispersed in the paste begin to reassociate through H-bonding. Retrogradation is commonly observed in the staling of bread and the syneresis of soups and sauces with starch as the thickener. Starch retrogradation is enhanced by increasing the concentration of starch in the paste, increasing amylose content and increasing the branch chain length of amylopectin (Donald 2004; Vamadevan and Bertoft 2015).

1.1.2.4 Starch applicability

Various botanical starch sources provide diverse properties on solubility, swelling, water absorption, syneresis and rheological behavior of pastes and gels, which determine their different uses (Ahamed et al. 1996; Chan et al. 2009; Joshi et al. 2013). In general, starch is commonly used as a thickener, gelling agent, stabilizer, emulsifier, fat replacer, encapsulating agent, anticaking agent and resistant starch in food industry. It also has wide applications in textile, pharmaceutical, paper and other non-food industries (Alcázar-Alay and Meireles 2015). As a renewable resource, starch recently has been used to make synthetic polymers for the development of environment-friendly degradable and biodegradable composites (Sangwan et al. 2014). The versatile applications in industry are mainly derived from the conventional commercial starch sources, which are obtained from cereals, such as corn, wheat and rice, and from tubers, such as potato (Alcázar-Alay and Meireles 2015; Waterschoot et al. 2015). Due to the finite variation in starch sources and the limitations of the native starch properties, modification of starch is commonly carried out post-extraction through physical, chemical, or enzymatic methods to enhance the starch attributes. Such modified starches are widely used at the moment. The application and development of genetic modification approaches offers the advantages of stability, cost reduction and pollution reduction (Chiu and Solarek 2009; Kaur et al. 2012; Ashogbon and Akintayo 2014). Recently, it has been demanded in industry to have new starches (e.g. starches from orphan crops) as alternative raw materials with different properties for commercial applications (Santana et al. 2014).

1.2 Orphan crops: Cassava (*Manihot esculenta* Crantz) and tef (*Eragrostis tef* (Zucc.) Trotter)

Orphan crops is a term used to encompass a group of crops that are locally rather than globally important, and which receive little public or private investment for their development (Naylor et al. 2004). Unlike maize (*Zea mays L.*), rice (*Oryza sativa L.*), wheat (*Triticum aestivum L. em. Thell.*) and other well-known global crops, orphan crops are not extensively traded and are under-researched. They include key staple crops in many low-income countries. Orphan crops are normally nutritious, valued culturally, adapted to harsh environments, and diverse in terms of their genetic, agroclimatic, and economic niches. Good examples include cassava (*Manihot esculenta Cranz.*), tef (*Eragrostis tef* (Zucc.) Trotter), yam (*Dioscorea sp. L.*), finger millet (*Eleusine coracana (L.) Gaertn.*), cowpeas (*Vigna unguiculata (L.) Walp*) and several tree crops (Naylor et al. 2004; Nelson et al. 2004).

Orphan crops play particular roles in agro-ecology and social-economy in developing countries, specially Africa. The major bottlenecks of orphan crops are low yield (e.g. tef and finger millet), poor in nutrition (e.g. cassava) and production of toxic substances (e.g. grass pea and cassava). To improve orphan crops, agriculture research is required. In few cases of modern research on orphan crops, molecular breeding approaches have been used for crop improvement (Varshney et al. 2012). Even though modern improvement techniques could be readily applied in orphan crops, conventional breeding is still the main strategy for most orphan crops improvement. Implementation of modern technologies to boost the productivity of orphan crops is needed (Tadele 2009).

1.2.1 Cassava

1.2.1.1 Utilization

Cassava (*Manihot esculenta* Crantz, *Euphorbiaceae*, 2n = 36) is a perennial shrub widely grown in tropical and subtropical countries such as Africa, Asia and Latin America. The crop is tolerant to poor soil and harsh climatic conditions and is cultivated over 13 million hectors mainly for starchy roots. The harvested storage roots serve as one of the most important staple foods for over 500 million people (Alves 2002; El-Sharkawy 2004). Notwithstanding the increasing growth and production, cassava remains referred to as orphan crop, mainly because it is not extensively traded and receive less attention from researchers compared to the main crops.

Cassava is processed into a wide range of products for food, for animal feed and for industry. For the diverse food utilization, it can be basically classified into fermented and non-fermented products. Fermented products include cassava breads, drinks and some traditional regional food such as fufu and farina, whereas the unfermented products include tapioca, cassava chips and pellets (Falade and Akingbala 2011). In industry, cassava starch is commonly used as a food additive with special emphasis on its lack of flavor contribution. Cassava dextrin is preferred in remoistening gums for stamps and envelope flaps because of its adhesive properties (Balagopalan 2002; Breuninger et al. 2009). In addition, cassava starch has also been

chemically or physically modified, and both native and modified cassava starch have important applications in corrugated cardboard manufacture, the paper industry and the textile industry. This is due to its good water holding capacities, stable viscosity and its qualities to form strong film and clear paste (Gunorubon and Kekpugile 2012). Recently, cassava starch has also been chosen as an alternative renewable resource for bio-ethanol production (Ukwur and Egbonu 2013).

Cassava is not a perfect crop, even though it plays an increasing role in industrial utilization and in the tropical agriculture. Cassava is essential to food security, yet it contains highly toxic cyanogenic compounds and anti-nutrients (Montagnac et al. 2009a). Cassava roots are also deficient in proteins and amino acids (Montagnac et al. 2009b). The improvement of food safety and nutritional value is crucial. Cassava also faces disease and pest problems, such as cassava mosaic virus (CMV) and cassava bacterial blight (CBB), which can be improved through better agronomic practices, superior varieties and pest and disease management. The research and development of cassava has lagged behind other major staples, particularly at the levels of understanding the basic biological processes in its storage roots, and of developing technique and tools to solve the constraints limiting cassava production (Taylor et al. 2012).

1.2.1.2 Biotechnology application

Plant biotechnology has been applied to improve cassava crops (Chavarriaga-Aguirre et al. 2016). The development of transgenic cassava plants through Agrobacterium-mediated transformation was first developed at the end of 1980s and has been improved to establish robust transformation platforms (Bull et al. 2009; Chetty et al. 2013). The transgenic cassava cultivars which express selectable marker genes, as well as genes of agronomic interest have been introduced with potential value for cassava farmers. The important cassava traits targeted using biotechnology include the development of new cassava varieties with a) increased resistance to viruses (e.g. cassava brown streak disease [CBSD] and cassava mosaic disease [CMD]), b) resistance to bacterial diseases such as cassava bacterial blight (CBB), c) resistance to the vectors that transmit diseases (e.g. whiteflies) d) nutritional improvement (carotenoid content, iron content and vitamin B6) and e) and mass production by tissue culture and synthetic seeds (Naconsie and Zhang 2016).

A wide genetic diversity in chemical composition and properties of natural cassava starches has been observed. Amylose content, as a major quality attribute to determine the starch property, is documented to range from 0 to 30.3% (Zhu 2015). Cassava with low/nil amylose content has been achieved through natural selection of mutant with the absence of granule-bound starch synthase (GBSS) enzyme (Ceballos et al. 2007), and genetic modification of suppressing the expression of GBSS1(Raemakers et al. 2005; Zhao et al. 2011). To enhance the understanding of carbon metabolism in cassava root, more studies at genome, transcriptome and proteome level have been conducted: comparative genomic approach to reconstruct starch biosynthesis pathway (Rongsirikul et al. 2010; Saithong et al. 2013), transcriptomic data integration inferring the dominance of starch biosynthesis in carbon utilization (Siriwat et al.

2012) and comparative proteomics to reveal carbohydrate metabolic enzymes (Wang et al. 2016).

1.2.2 Tef

1.2.2.1 Genetic diversity

Tef (*Eragrostis tef* [Zuccagni] Trotter) is a tropical cereal crop that belongs to the family of *Poaceae*, subfamily *Eragrostoidae*, tribe *Eragrosteae*, and genus *Eragrostis*. Tef is a C4 self-pollinated allotetraploid with a chromosome number of 2n = 4x = 40. Tef is indigenous to Ethiopia. It can adapt to harsh environment and is very resistant to insect pests compared to other cereals in Ethiopia (Pa and Asseng 2018).

Genetic diversity in morphology, agronomic traits, natural populations, and agro-ecological traits have been recorded (Assefa et al. 2015). Tef cultivars for commercial use have been recognized and described based on the colour of the grains (from milky-white to almost dark brown), ramification of the inflorescences (panicle type from loose to compact) and the size of plants (50–120 cm for most varieties). The Ethiopian Institute of Biodiversity has conserved over 5000 accessions and 10000 tef genotypes to prevent the loss of genetic diversity and for use in varietal improvement. The exact number of domesticated cultivars is unclear. Cultivars such as DZ-Cr-387 RIL 355 (*Quncho*), DZ-01-974 (*Dukem*), DZ-01-196 (*Magna*), DZ-Cr-37 (*Tsedey*), DZ-01-354 (*Enatite*) and DZ-01-99 (*Asgori*) are the most widely grown in Ethiopia (Assefa et al. 2015; Bultosa 2016).

1.2.2.2 Tef grain composition

Tef grain is oval-shaped with a very small size of 0.7–1.0 mm in diameter. The individual grain mass is generally ≤ 2 mg, about 0.6–0.8% that of wheat (Bultosa et al. 2002). Starch is the major component of tef grain (over 70% of the dry weight) and individual starch granules are also very small (2–6 μm in diameter). Starch granule shape is polygonal, smooth with no surface pores, similar to rice (Figure 3). Similar to the starches of other cereals, tef starch has 25-30% amylose content (Bultosa et al. 2002).

The physicochemical properties of tef starch have been studied. X-ray diffraction revealed that tef has A-type starch granule with about 37% crystallinity (Bultosa and Taylor 2003). The starch has a high gelatinization temperature and pasting viscosities (peak, breakdown, and setback) which are lower than maize starch (Bultosa et al. 2002). The paste clarity of tef starch is opaque and the gel texture is short and smooth (Bultosa 2007).

The protein content of tef appears to be similar to that of other common cereals ranging from 9-13% and the major amino acids in tef grain are glutamic acid, alanine, proline, aspartic acid, leucine, and valine (Gebremariam et al. 2012). Tef protein is essentially free of the type of gluten found in wheat, which is important for consumers with celiac disease (Spaenij-Dekking et al. 2005). Fiber, minerals, fat, vitamins and other components have been measured in tef as well and compared with cereals (Hager et al. 2012).

1.2.2.3 Utilization

The production of tef in Ethiopia reached 4.2 million tons in 2015. Tef flour is mainly used for staple food in Ethiopia (e.g. a pancake-like local bread *injera*) and its natural gluten free property has great potential to be formulated into malting, brewing and manufacturing of gluten-free foods and beverages (Gebremariam et al. 2012). Tef flour has also been used in a range of food products, such as pasta, cookies and cakes. Recently, it is also credited as a source of bioactive compounds such as polyphenols (Zhu 2018).

1.3 Starch genetic modification in orphan crops

The application of biotechnology to alter cassava starch structure and properties has not been developed as well as in rice, maize and other staple crops. In industry, cassava native starch is usually either chemically, physically, or enzymatically modified before use for diverse applications (Zhu 2015). Genetic modification has so far been used to generate low-amylose or amylose-free cassava cultivars (Salehuzzaman et al. 1993; Zhao et al. 2011). The resultant *waxy* cassava starch has been shown to have excellent freeze thaw stability, to form clear gels and to have increased viscosity, low synerisis and therefore, is highly suitable for industrial applications like frozen food and paper making (Krishna Radhika et al. 2014). Starch production in cassava storage roots has also been increased substantially by enhancing ADP-glucose pyrophosphorylase (AGPase) activity (Ihemere et al. 2006). Besides, cDNAs encoding branching enzyme (BEI, BEII) involved in the starch biosynthetic pathway have also been cloned from cassava, which will find applications in genetic modification of cassava for enhanced starch quality (Opabode et al. 2012).

No research into starch modification at the molecular level has been reported for tef. Although conventional breeding is still the main technique used for improving tef, *in-vitro* regeneration, mutant breeding (TILLING) and other molecular approaches are being developed (Yu et al. 2007; Assefa et al. 2011; Cannarozzi et al. 2018). Recently, the genome of tef has been sequenced and the data made accessible by Cannarozzi et al (2014). The draft genome contains 672 Mbp which represents 87% of the genome size. It will be of great value for future genetic improvements of tef agricultural traits and food processing quality (Cannarozzi et al. 2014).

1.4 Aim of the PhD thesis

The overall aim of this work was to apply the knowledge of starch metabolism achieved from Arabidopsis and other model species to improve orphan crops. I expected in the future, our bio-engineered starches could potentially increase the value of orphan crops and expend their utilizations for the food and non-food industries.

The cassava part of my research aimed to alter starch phosphorylation in cassava storage roots by manipulating the expression of the phosphorylation-related enzymes. Transgenic cassava plants were generated that either overexpressed the potato glucan, water dikinase gene (StGWD), or silenced each of the endogenous phosphoglucan phospahatse genes (MeSEX4 and MeLSF2) in the root. Several phosphate-related physicochemical properties of the starch extracted from transgenic and control lines were then tested.

The tef part of my research aimed to generate tef plants with modified starch properties. This included several steps: 1) Identification of starch-related genes involved in starch synthesis in the tef endosperm using comparative genomic analysis; 2) Confirmation of the expression of the targeted genes in developing tef seeds through RNA-sequencing analysis; 3) Establishment of a new TILLING-by-sequencing method for screening a mutated tef plant population (using high throughput sequencing and bioinformatics analysis); 4) Identification of the useful null-mutant plants with improved starch properties by cultivating and crossing of plants.

1.5 References

Abebe W, Collar C, Ronda F (2015) Impact of variety type and particle size distribution on starch enzymatic hydrolysis and functional properties of tef flours. Carbohydr Polym 115:260–268. doi: 10.1016/j.carbpol.2014.08.080

Agarwal P, Kapoor S, Tyagi AK (2011) Transcription factors regulating the progression of monocot and dicot seed development. BioEssays 33:189–202.

Ahamed NT, Singhal RS, Kulkarni PR, Pal M (1996) Physicochemical and functional properties of Chenopodium quinoa starch. Carbohydr Polym 31:99–103. doi: 10.1016/S0144-8617(96)00034-3

Alcázar-Alay SC, Meireles MAA (2015) Physicochemical properties, modifications and applications of starches from different botanical sources. Food Sci Technol 35:215–236. doi: 10.1590/1678-457X.6749

Alexandersson M, Cawley S, Pachter L (2003) SLAM: Cross-species gene finding and alignment with a generalized pair hidden Markov model. Genome Res 13:496–502. doi: 10.1101/gr.424203

Allem AC, Genéticos R (2002) The Origins and Taxonomy of Cassava. Cassava Biol Prod Util 1–16. doi: 10.1080/21622965.2014.897905

Altschul SF, Gish W, Miller W, et al (1990) Basic local alignment search tool. J Mol Biol 215:403–10. doi: 10.1016/S0022-2836(05)80360-2

Alves a. a. C (2002) Cassava botany and physiology. Cassava Biol Prod Util 67–89. doi: 10.1079/9780851995243.0000

Anand S, Mangano E, Barizzone N, et al (2016) Next generation sequencing of pooled samples: Guideline for variants' filtering. Sci Rep. doi: 10.1038/srep33735

Armstead I, Huang L, Ravagnani A, et al (2009) Bioinformatics in the orphan crops. Brief. Bioinform. 10:645–653.

Ashogbon AO, Akintayo ET (2014) Recent trend in the physical and chemical modification of starches from different botanical sources: A review. Starch/Staerke 66:41–57.

Aslam U, Cheema HMN, Ahmad S, et al (2016) COTIP: Cotton TILLING Platform, a Resource for Plant Improvement and Reverse Genetic Studies. Front Plant Sci. doi: 10.3389/fpls.2016.01863

Assefa K, Chanyalew S, Tadele Z (2017) Tef, Eragrostis tef (Zucc.) Trotter. In: Millets and Sorghum: Biology and Genetic Improvement. pp 226–266

Assefa K, Yu J-K, Zeid M, et al (2011) Breeding tef [Eragrostis tef (Zucc.) trotter]: conventional and molecular approaches. Plant Breed 130:1–9. doi: 10.1111/j.1439-0523.2010.01782.x

Avni A, Blázquez MA (2011) Can plant biotechnology help in solving our food and energy shortage in the future? Curr. Opin. Biotechnol. 22:220–223.

Balagopalan C (2002) Cassava Utilization in Food, Feed and Industry. Cassava Biol Prod Util 301–318. doi: 10.1079/9780851995243.0000

Barkley N, Wang M (2008) Application of TILLING and EcoTILLING as Reverse Genetic Approaches to Elucidate the Function of Genes in Plants and Animals. Curr Genomics 9:212–226. doi: 10.2174/138920208784533656

Baunsgaard L, Lütken H, Mikkelsen R, et al (2005) A novel isoform of glucan, water dikinase phosphorylates pre-phosphorylated α-glucans and is involved in starch degradation in Arabidopsis. Plant J 41:595–605. doi: 10.1111/j.1365-313X.2004.02322.x

Blagbrough IS, Bayoumi SAL, Rowan MG, Beeching JR (2010) Cassava: An appraisal of its phytochemistry and its biotechnological prospects. Phytochemistry 71:1940–1951. doi:

10.1016/j.phytochem.2010.09.001
Blazek J, Copeland L (2008) Pasting and swelling properties of wheat flour and starch in relation to amylose content. Carbohydr Polym 71:380–387. doi: 10.1016/j.carbpol.2007.06.010
Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254. doi: 10.1016/0003-2697(76)90527-3
Breuninger WF, Piyachomkwan K, Sriroth K (2009) Tapioca/Cassava Starch: Production and Use. In: Starch. pp 541–568
Bull SE, Owiti JA, Niklaus M, et al (2009) Agrobacterium-mediated transformation of friable embryogenic calli and regeneration of transgenic cassava. Nat Protoc 4:1845–1854. doi: 10.1038/nprot.2009.208
Bultosa G (2016) Teff: Overview, 2nd edn. Elsevier Ltd.
Bultosa G (2007) Physicochemical characteristics of grain and flour in 13 tef [Eragrostis tef (Zucc .) Trotter] grain varieties. J Appl Sci Res 3:2042–2050.
Bultosa G, Hall AN, Taylor JRN (2002) Physico-chemical characterization of grain tef [Eragrostis tef (Zucc.) trotter] starch. Starch/Staerke 54:461–468. doi: 10.1002/1521-379X(200210)54:10<461::AID-STAR461>3.0.CO;2-U
Bultosa G, Taylor JRN (2003) Chemical and physical characterisation of grain tef [Eragrostis tef (Zucc.) Trotter] starch granule composition. Starch/Staerke 55:304–312. doi: 10.1002/star.200390065
Bultosa G, Taylor JRN (2004) Paste and gel properties and in vitro digestibility of Tef [Eragrostis tef (Zucc). Trotter] starch. Starch/Staerke 56:20–28. doi: 10.1002/star.200200191
Burton RA, Jenner H, Carrangis L, et al (2002) Starch granule initiation and growth are altered in barley mutants that lack isoamylase activity. Plant J 31:97–112. doi: 10.1046/j.1365-313X.2002.01339.x
Cannarozzi G, Chanyalew S, Assefa K, et al (2018) Technology generation to dissemination: lessons learned from the tef improvement project. Euphytica 214:31. doi: 10.1007/s10681-018-2115-5
Cannarozzi G, Plaza-Wüthrich S, Esfeld K, et al (2014) Genome and transcriptome sequencing identifies breeding targets in the orphan crop tef (Eragrostis tef). BMC Genomics 15:581. doi: 10.1186/1471-2164-15-581
Carciofi M, Shaik SS, Jensen SL, et al (2011) Hyperphosphorylation of cereal starch. J Cereal Sci 54:339–346. doi: 10.1016/j.jcs.2011.06.013
Carpenter MA, Joyce NI, Genet RA, et al (2015) Starch phosphorylation in potato tubers is influenced by allelic variation in the genes encoding glucan water dikinase, starch branching enzymes I and II, and starch synthase III. Front Plant Sci 6:143. doi: 10.3389/fpls.2015.00143
Caspar T, Lin TP, Kakefuda G, et al (1991) Mutants of Arabidopsis with altered regulation of starch degradation. Plant Physiol 95:1181–1188. doi: 10.1104/pp.95.4.1181
Ceballos H, Sánchez T, Morante N, et al (2007) Discovery of an amylose-free starch mutant in cassava (Manihot esculenta Crantz). J Agric Food Chem 55:7469–7476. doi: 10.1021/jf070633y
Chan HT, Bhat R, Karim AA (2009) Physicochemical and functional properties of ozone-oxidized starch. J Agric Food Chem 57:5965–5970. doi: 10.1021/jf9008789
Chavarriaga-Aguirre P, Brand A, Medina A, et al (2016) The potential of using biotechnology to improve cassava: a review. Vitr. Cell. Dev. Biol. - Plant 52:461–478.
Chetty CC, Rossin CB, Gruissem W, et al (2013) Empowering biotechnology in southern Africa: Establishment of a robust transformation platform for the production of

transgenic industry-preferred cassava. N Biotechnol 30:136–143. doi: 10.1016/j.nbt.2012.04.006
Chiu C wai, Solarek D (2009) Modification of Starches. In: Starch. pp 629–655
Colbert T (2001) High-Throughput Screening for Induced Point Mutations. PLANT Physiol 126:480–484. doi: 10.1104/pp.126.2.480
Craig SAS, Maningat CC, Seib PA, Hoseney RC (1989) Starch paste clarity. Cereal Chem 66:173–182.
Critchley JH, Zeeman SC, Takaha T, et al (2001) A critical role for disproportionating enzyme in starch breakdown is revealed by a knock-out mutation in Arabidopsis. Plant J 26:89–100. doi: 10.1046/j.1365-313X.2001.01012.x
Crumpton-Taylor M, Pike M, Lu KJ, et al (2013) Starch synthase 4 is essential for coordination of starch granule formation with chloroplast division during Arabidopsis leaf expansion. New Phytol 200:1064–1075. doi: 10.1111/nph.12455
Deffenbaugh LB, Walker CE (1989) Use of the Rapid-Visco-Analyzer to Measure Starch Pasting Properties. Part I: Effect of Sugars. Starch - Stärke 41:461–467. doi: 10.1002/star.19890411205
Delatte T, Umhang M, Trevisan M, et al (2006) Evidence for distinct mechanisms of starch granule breakdown in plants. J Biol Chem 281:12050–12059. doi: 10.1074/jbc.M513661200
Donald AM (2004) Understanding starch structure and functionality. In: Starch in Food: Structure, Function and Applications. pp 156–184
Dumez S, Wattebled F, Dauvillee D, et al (2006) Mutants of Arabidopsis Lacking Starch Branching Enzyme II Substitute Plastidial Starch Synthesis by Cytoplasmic Maltose Accumulation. PLANT CELL ONLINE 18:2694–2709. doi: 10.1105/tpc.105.037671
Edner C, Li J, Albrecht T, et al (2007) Glucan, Water Dikinase Activity Stimulates Breakdown of Starch Granules by Plastidial β-Amylases. Plant Physiol 145:17–28. doi: 10.1104/pp.107.104224
El-Sharkawy MA (2004) Cassava biology and physiology. Plant Mol Biol 56:481–501. doi: 10.1007/s11103-005-2270-7
Eliasson AC (2004) Starch in Food: Structure, Function and Applications.
Esfeld K, Uauy C, Tadele Z (2012) Application of TILLING for orphan crop improvement. In: Biotechnology of Neglected and Underutilized Crops. pp 83–113
Falade KO, Akingbala JO (2011) Utilization of Cassava for food. Food Rev Int 27:51–83. doi: 10.1080/87559129.2010.518296
Feike D, Seung D, Graf A, et al (2016) A novel starch-granule-associated protein required for the control of starch degradation in Arabidopsis thaliana leaves. Plant Cell 28:1472–1489. doi: 10.1105/tpc.16.00011
Feltus F a, Singh HP, Lohithaswa HC, et al (2006) A comparative genomics strategy for targeted discovery of single-nucleotide polymorphisms and conserved-noncoding sequences in orphan crops. Plant Physiol 140:1183–1191. doi: 10.1104/pp.105.074203
Fredriksson H, Silverio J, Andersson R, et al (1998) The influence of amylose and amylopectin characteristics on gelatinization and retrogradation properties of different starches. Carbohydr Polym 35:119–134. doi: 10.1016/S0144-8617(97)00247-6
Fu F-F, Xue H-W (2010) Coexpression Analysis Identifies Rice Starch Regulator1, a Rice AP2/EREBP Family Transcription Factor, as a Novel Rice Starch Biosynthesis Regulator. PLANT Physiol 154:927–938. doi: 10.1104/pp.110.159517
Fujita N, Yoshida M, Asakura N, et al (2006) Function and Characterization of Starch Synthase I Using Mutants in Rice. PLANT Physiol 140:1070–1084. doi: 10.1104/pp.105.071845
Fulton DC, Stettler M, Mettler T, et al (2008) -AMYLASE4, a Noncatalytic Protein

Required for Starch Breakdown, Acts Upstream of Three Active -Amylases in Arabidopsis Chloroplasts. PLANT CELL ONLINE 20:1040–1058. doi: 10.1105/tpc.107.056507
Gallant DJ, Bouchet B, Baldwin PM (1997) Microscopy of starch: evidence of a new level of granule organization. Carbohydr Polym 32:177–191. doi: 10.1016/S0144-8617(97)00008-8
Gao M, Fisher DK, Kim K-N, et al (1996) Evolutionary conservation and expression patterns of maize starch branching enzyme I and IIb genes suggests isoform specialization. Plant Mol Biol 30:1223–1232. doi: 10.1007/BF00019554
Garrison E, Marth G (2012) Haplotype-based variant detection from short-read sequencing. 1–9. doi: arXiv:1207.3907 [q-bio.GN]
Gebremariam MM, Zarnkow M, Becker T (2012) Teff (Eragrostis tef) as a raw material for malting, brewing and manufacturing of gluten-free foods and beverages: a review. J. Food Sci. Technol. 51:2881–2895.
Grimaud F, Rogniaux H, James MG, et al (2008) Proteome and phosphoproteome analysis of starch granule-associated proteins from normal maize and mutants affected in starch biosynthesis. J Exp Bot 59:3395–3406. doi: 10.1093/jxb/ern198
Gunorubon AJ, Kekpugile DK (2012) Modification of Cassava Starch for Industrial Uses. Int J Eng Technol 2:913–919.
Hager AS, Wolter A, Jacob F, et al (2012) Nutritional properties and ultra-structure of commercial gluten free flours from different botanical sources compared to wheat flours. J Cereal Sci 56:239–247. doi: 10.1016/j.jcs.2012.06.005
Hannah LC, James M (2008) The complexities of starch biosynthesis in cereal endosperms. Curr Opin Biotechnol 19:160–165. doi: 10.1016/j.copbio.2008.02.013
Hari PK, Garg S, Garg SK (1989) Gelatinization of Starch and Modified Starch. Starch ??? St??rke 41:88–91. doi: 10.1002/star.19890410304
Hennen-Bierwagen TA, Lin Q, Grimaud F, et al (2009) Proteins from Multiple Metabolic Pathways Associate with Starch Biosynthetic Enzymes in High Molecular Weight Complexes: A Model for Regulation of Carbon Allocation in Maize Amyloplasts. PLANT Physiol 149:1541–1559. doi: 10.1104/pp.109.135293
Hennen-Bierwagen TA, Liu F, Marsh RS, et al (2008) Starch Biosynthetic Enzymes from Developing Maize Endosperm Associate in Multisubunit Complexes. PLANT Physiol 146:1892–1908. doi: 10.1104/pp.108.116285
Henry IM, Nagalakshmi U, Lieberman MC, et al (2014) Efficient Genome-Wide Detection and Cataloging of EMS-Induced Mutations Using Exome Capture and Next-Generation Sequencing. Plant Cell 26:1382–1397. doi: 10.1105/tpc.113.121590
Henson (2014) A Comparison of Barley Malt Amylolytic Enzyme Thermostabilities and Wort Sugars Produced During Mashing. J Am Soc Brew Chem. doi: 10.1094/ASBCJ-2014-0130-01
Hizukuri S, Tabata S, Kagoshima, Nikuni Z (1970) Studies on Starch Phosphate Part 1. Estimation of glucose-6-phosphate residues in starch and the presence of other bound phosphate(s). Starch - Stärke 22:338–343. doi: 10.1002/star.19700221004
Hoover R (2001) Composition, molecular structure, and physicochemical properties of tuber and root starches: A review. Carbohydr Polym 45:253–267. doi: 10.1016/S0144-8617(00)00260-5
Hovenkamp-Hermelink JHM, Jacobsen E, Ponstein AS, et al (1987) Isolation of an amylose-free starch mutant of the potato (Solanum tuberosum L.). Theor Appl Genet 75:217–221. doi: 10.1007/BF00249167
Hwang S, Kim E, Lee I, Marcotte EM (2015) Systematic comparison of variant calling pipelines using gold standard personal exome variants. Sci Rep. doi: 10.1038/srep17875

Ihemere U, Arias-Garzon D, Lawrence S, Sayre R (2006) Genetic modification of cassava for enhanced starch production. Plant Biotechnol J 4:453–465. doi: 10.1111/j.1467-7652.2006.00195.x
Imberty A, Chanzy H, Pérez S, et al (1988) The double-helical nature of the crystalline part of A-starch. J Mol Biol 201:365–378. doi: 10.1016/0022-2836(88)90144-1
Imberty A, Perez S (1988) A revisit to the three-dimensional structure of B-type starch. Biopolymers 27:1205–1221. doi: 10.1002/bip.360270803
Jaiswal S, Baìšga M, Ahuja G, et al (2014) Development of barley (Hordeum Vulgare L.) lines with altered starch granule size distribution. J Agric Food Chem 62:2289–2296. doi: 10.1021/jf405424x
James MG, Denyer K, Myers AM (2003) Starch synthesis in the cereal endosperm. Curr. Opin. Plant Biol. 6:215–222.
Jeon JS, Ryoo N, Hahn TR, et al (2010) Starch biosynthesis in cereal endosperm. Plant Physiol. Biochem. 48:383–392.
Jobling S (2004) Improving starch for food and industrial applications. Curr Opin Plant Biol 7:210–218. doi: 10.1016/j.pbi.2003.12.001
Joshi M, Aldred P, McKnight S, et al (2013) Physicochemical and functional characteristics of lentil starch. Carbohydr Polym 92:1484–1496. doi: 10.1016/j.carbpol.2012.10.035
Kamei CLA, Severing EI, Dechesne A, et al (2016) Orphan Crops Browser: a bridge between model and orphan crops. Mol Breed 36:1–18. doi: 10.1007/s11032-015-0430-2
Kammerer B, Fischer K, Hilpert B, et al (1998) Molecular Characterization of a Carbon Transporter in Plastids from Heterotrophic Tissues: The Glucose 6-Phosphate/Phosphate Antiporter. Plant Cell 10:105–117. doi: 10.2307/3870632
Kampfenkel K, Möhlmann T, Batz O, et al (1995) Molecular characterization of an Arabidopsis thaliana cDNA encoding a novel putative adenylate translocator of higher plants. FEBS Lett 374:351–5. doi: 10.1016/0014-5793(95)01143-3
Kaneko M, Itoh H, Ueguchi-Tanaka M, et al (2002) The alpha -Amylase Induction in Endosperm during Rice Seed Germination Is Caused by Gibberellin Synthesized in Epithelium. Plant Physiol 128:1264–1270. doi: 10.1104/pp.010785
Karim AA, Toon LC, Lee VPL, et al (2007) Effects of phosphorus contents on the gelatinization and retrogradation of potato starch. J Food Sci 72:132–138. doi: 10.1111/j.1750-3841.2006.00251.x
Karlström A, Calle F, Salazar S, et al (2016) Biological Implications in Cassava for the Production of Amylose-Free Starch: Impact on Root Yield and Related Traits. Front Plant Sci 7:604. doi: 10.3389/fpls.2016.00604
Kaur B, Ariffin F, Bhat R, Karim AA (2012) Progress in starch modification in the last decade. Food Hydrocoll 26:398–404. doi: 10.1016/j.foodhyd.2011.02.016
Kaur M, Oberoi DPS, Sogi DS, Gill BS (2011) Physicochemical, morphological and pasting properties of acid treated starches from different botanical sources. J Food Sci Technol 48:460–465. doi: 10.1007/s13197-010-0126-x
KIM CS, WALKER CE (1992) Changes in Starch Pasting Properties Due to Sugars and Emulsifiers as Determined by Viscosity Measurement. J Food Sci 57:1009–1013. doi: 10.1111/j.1365-2621.1992.tb14344.x
Knoll JE, Ramos ML, Zeng Y, et al (2011) TILLING for allergen reduction and improvement of quality traits in peanut (Arachis hypogaea L.). BMC Plant Biol. doi: 10.1186/1471-2229-11-81
Kotting O (2005) Identification of a Novel Enzyme Required for Starch Metabolism in Arabidopsis Leaves. The Phosphoglucan, Water Dikinase. PLANT Physiol 137:242–252. doi: 10.1104/pp.104.055954
Kötting O, Santelia D, Edner C, et al (2009a) STARCH-EXCESS4 is a laforin-like

Phosphoglucan phosphatase required for starch degradation in Arabidopsis thaliana. Plant Cell 21:334–46. doi: 10.1105/tpc.108.064360

Kötting O, Santelia D, Edner C, et al (2009b) STARCH-EXCESS4 is a laforin-like Phosphoglucan phosphatase required for starch degradation in Arabidopsis thaliana. Plant Cell 21:334–46. doi: 10.1105/tpc.108.064360

Krishna Radhika N, Sheela MN, Asha Devi A, et al (2014) Genetic modification for designer starch from cassava. J. Trop. Agric. 52:1–6.

Kurakake M, Akiyama Y, Hagiwara H, Komaki T (2009) Effects of cross-linking and low molecular amylose on pasting characteristics of waxy corn starch. Food Chem 116:66–70. doi: 10.1016/j.foodchem.2009.02.006

Kusumayanti H, Handayani NA, Santosa H (2015) Swelling Power and Water Solubility of Cassava and Sweet Potatoes Flour. Procedia Environ Sci 23:164–167. doi: 10.1016/j.proenv.2015.01.025

Lee MR, Swanson BR, Baik BK (2001) Influence of amylose content on properties of wheat starch and breadmaking quality of starch and gluten blends. Cereal Chem 78:701–706. doi: 10.1094/CCHEM.2001.78.6.701

Li H, Handsaker B, Wysoker A, et al (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079. doi: 10.1093/bioinformatics/btp352

Liang X, King JM (2003) Pasting and crystalline property differences of commercial and isolated rice starch with added amino acids. J Food Sci 68:832–838. doi: 10.1111/j.1365-2621.2003.tb08251.x

Lim S-T, Kasemsuwan T, Jane J-L (1994) Characterization of Phosphorus in Starch by 31P-Nuclear Magnetic Resonance Spectroscopy. Cereal Chem 71:488–493.

Lin Q, Huang B, Zhang M, et al (2012) Functional Interactions between Starch Synthase III and Isoamylase-Type Starch-Debranching Enzyme in Maize Endosperm. PLANT Physiol 158:679–692. doi: 10.1104/pp.111.189704

Liu H, Yu L, Xie F, Chen L (2006) Gelatinization of cornstarch with different amylose/amylopectin content. Carbohydr Polym 65:357–363. doi: 10.1016/j.carbpol.2006.01.026

Lloyd JR, Kossmann J (2015) Transitory and storage starch metabolism: Two sides of the same coin? Curr. Opin. Biotechnol. 32:143–148.

Lorberth R, Ritte G, Willmitzer L, Kossmann J (1998) Inhibition of a starch-granule-bound protein leads to modified starch and repression of cold sweetening. Nat Biotechnol 16:473–477. doi: 10.1038/nbt0598-473

Lund D (1984) Influence of time, temperature, moisture, ingredients, and processing conditions on starch gelatinization. Crit Rev Food Sci Nutr 20:249–73. doi: 10.1080/10408398409527391

Mahlow S, Orzechowski S, Fettke J (2016) Starch phosphorylation: insights and perspectives. Cell Mol Life Sci 73:2753–2764. doi: 10.1007/s00018-016-2248-4

Manners DJ (1989) Recent developments in our understanding of amylopectin structure. Carbohydr. Polym. 11:87–112.

Martin C, Smith a M (1995) Starch biosynthesis. Plant Cell 7:971–985. doi: 10.1039/qr9692300057

Maurer T, Kalbitzer HR (1996) Indirect referencing of31P and19F NMR spectra. J Magn Reson - Ser B 113:177–178. doi: 10.1006/jmrb.1996.0172

Menlove KJ, Clement M, Crandall KA (2009) Similarity searching using BLAST. Methods Mol. Biol. 537:1–22.

Mikkelsen R, Mutenda KE, Mant A, et al (2005) Alpha-glucan, water dikinase (GWD): a plastidic enzyme with redox-regulated and coordinated catalytic activity and binding affinity. Proc Natl Acad Sci U S A 102:1785–1790.

Minoia S, Petrozza A, D'Onofrio O, et al (2010) A new mutant genetic resource for tomato crop improvement by TILLING technology. BMC Res Notes. doi: 10.1186/1756-0500-3-69

Montagnac JA, Davis CR, Tanumihardjo SA (2009a) Processing techniques to reduce toxicity and antinutrients of Cassava for use as a staple food. Compr Rev Food Sci Food Saf 8:17–27. doi: 10.1111/j.1541-4337.2008.00064.x

Montagnac JA, Davis CR, Tanumihardjo SA (2009b) Nutritional value of cassava for use as a staple food and recent advances for improvement. Compr Rev Food Sci Food Saf 8:181–194. doi: 10.1111/j.1541-4337.2009.00077.x

Montero WR (2003) Cassava; Biology, Production and Utilization.

Moorthy SN (2002) Physicochemical and functional properties of tropical tuber starches: A review. Starch/Staerke 54:559–592.

Morrison WR, Law R V., Snape CE (1993) Evidence for inclusion complexes of lipids with v-amylose in maize, rice and oat starches. J Cereal Sci 18:107–109. doi: 10.1006/jcrs.1993.1039

Morrison WR, Milligan TP, Azudin MN (1984) A relationship between the amylose and lipid contents of starches from diploid cereals. J Cereal Sci 2:257–271. doi: 10.1016/S0733-5210(84)80014-4

Mun JH, Kwon SJ, Yang TJ, et al (2009) Genome-wide comparative analysis of the Brassica rapa gene space reveals genome shrinkage and differential loss of duplicated genes after whole genome triplication. Genome Biol. doi: 10.1186/gb-2009-10-10-r111

Muth J, Hartje S, Twyman RM, et al (2008) Precision breeding for novel starch variants in potato. Plant Biotechnol J 6:576–584. doi: 10.1111/j.1467-7652.2008.00340.x

NACONSIE M, ZHANG P (2016) Transgenic technologies in cassava for nutritional improvement and viral disease resistance: a key strategy for food security in Africa. Front Agric Sci Eng 3:285. doi: 10.15302/J-FASE-2016119

Nakamura T, Yamamori M, Hirano H, et al (1995) Production of waxy (amylose-free) wheats. MGG Mol Gen Genet 248:253–259. doi: 10.1007/BF02191591

Naylor RL, Falcon WP, Goodman RM, et al (2004) Biotechnology in the developing world: A case for increased investments in orphan crops. Food Policy 29:15–44. doi: 10.1016/j.foodpol.2004.01.002

Nelson RJ, Naylor RL, Jahn MM (2004) The role of genomics research in improvement of 'orphan' crops. In: Crop Science. pp 1901–1904

Niittylä T, Messerli G, Trevisan M, et al (2004) A Previously Unknown Maltose Transporter Essential for Starch Degradation in Leaves. Science (80-) 303:87–89. doi: 10.1126/science.1091811

Nyakabau T, Wokadala OC, Emmambux MN (2013) Effect of steeping additives on tef starch extptraction and its quality. Starch/Staerke 65:738–746. doi: 10.1002/star.201200241

Oleykowski C a, Bronson Mullins CR, Godwin a K, Yeung a T (1998) Mutation detection using a novel plant endonuclease. Nucleic Acids Res 26:4597–4602. doi: 10.1093/nar/26.20.4597

Opabode JT, Oyelakin OO, Akinyemiju OA, Ingelbrecht IL (2012) Starch branching enzyme I from cassava (Manihot esculenta Crantz): Genomic organization , intron identification and phylogenetic analysis. Glob J Biotechnol Biocemistry 7:79–85. doi: 10.5829/idosi.gjbb.2012.7.3.1103

Pa K, Asseng S (2018) A review of tef physiology for developing a tef crop model. 94:54–66. doi: 10.1016/j.eja.2018.01.008

Pfister B, Zeeman SC (2016a) Formation of starch in plant cells. Cell. Mol. Life Sci. 73:2781–2807.

Pfister B, Zeeman SC (2016b) Formation of starch in plant cells. Cell Mol Life Sci 73:2781–2807. doi: 10.1007/s00018-016-2250-x
Pilling E, Smith a M (2003) Growth ring formation in the starch granules of potato tubers. Plant Physiol 132:365–371. doi: 10.1104/pp.102.018044
Polnaya FJ, Haryadi, Marseno DW, Cahyanto MN (2013) Effects of phosphorylation and cross-linking on the pasting properties and molecular structure of sago starch. Int Food Res J 20:1609–1615.
Puounti-Kaerlas J (2002) Cassava Biotechnology. Genet Eng Rev 15:329–364. doi: 10.1079/9780851995243.0000
Radchuk V V., Borisjuk L, Sreenivasulu N, et al (2009) Spatiotemporal Profiling of Starch Biosynthesis and Degradation in the Developing Barley Grain. PLANT Physiol 150:190–204. doi: 10.1104/pp.108.133520
Raemakers K, Schreuder M, Suurs L, et al (2005) Improved cassava starch by antisense inhibition of granule-bound starch synthase I. Mol Breed 16:163–172. doi: 10.1007/s11032-005-7874-8
Ran FA, Hsu PD, Wright J, et al (2013) Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8:2281–2308. doi: 10.1038/nprot.2013.143
Raphael M, Yona B, Stephen K, et al (2011) Amylopectin molecular structure and functional properties of starch from three Ugandan cassava varieties. J Plant Breed Crop Sci 3:195–202.
Rathore RS, Garg N, Garg S, Kumar A (2009) Starch phosphorylase: Role in starch metabolism and biotechnological applications. Crit. Rev. Biotechnol. 29:214–224.
Ratnayake WS, Jackson DS (2008) Starch Gelatinization. In: Advances in Food and Nutrition Research. pp 221–268
Regina A, Bird A, Topping D, et al (2006) High-amylose wheat generated by RNA interference improves indices of large-bowel health in rats. Proc Natl Acad Sci 103:3546–3551. doi: 10.1073/pnas.0510737103
Rhoads A, Au KF (2015) PacBio Sequencing and Its Applications. Genomics, Proteomics Bioinforma. 13:278–289.
Ritte G, Heydenreich M, Mahlow S, et al (2006a) Phosphorylation of C6- and C3-positions of glucosyl residues in starch is catalysed by distinct dikinases. FEBS Lett 580:4872–4876. doi: 10.1016/j.febslet.2006.07.085
Ritte G, Heydenreich M, Mahlow S, et al (2006b) Phosphorylation of C6- and C3-positions of glucosyl residues in starch is catalysed by distinct dikinases. FEBS Lett 580:4872–4876. doi: 10.1016/j.febslet.2006.07.085
Rongsirikul O, Saithong T, Kalapanulak S, et al (2010) Reconstruction of starch biosynthesis pathway in cassava using comparative genomic approach. In: Communications in Computer and Information Science. pp 118–129
Saithong T, Rongsirikul O, Kalapanulak S, et al (2013) Starch biosynthesis in cassava: A genome-based pathway reconstruction and its exploitation in data integration. BMC Syst Biol. doi: 10.1186/1752-0509-7-75
Salehuzzaman SNIM, Jacobsen E, Visser RGF (1993) Isolation and characterization of a cDNA encoding granule-bound starch synthase in cassava (Manihot esculenta Crantz) and its antisense expression in potato. Plant Mol Biol 23:947–962. doi: 10.1007/BF00021811
Sangwan P, Petinakis E, Dean K (2014) Starch Polymers.
Sano Y (1984) Differential regulation of waxy gene expression in rice endosperm. Theor Appl Genet 68:467–473. doi: 10.1007/BF00254822
Santana ÁL, Angela M, Meireles A (2014) New Starches are the Trend for Industry Applications: A Review. Food Public Heal 4:229–241. doi: 10.5923/j.fph.20140405.04

Santelia D, Kötting O, Seung D, et al (2011) The phosphoglucan phosphatase Like Sex Four2 dephosphorylates starch at the C3-position in Arabidopsis. Plant Cell 23:4096–4111. doi: 10.1105/tpc.111.092155
Santelia D, Zeeman SC (2011a) Progress in Arabidopsis starch research and potential biotechnological applications. Curr Opin Biotechnol 22:271–280. doi: 10.1016/j.copbio.2010.11.014
Santelia D, Zeeman SC (2011b) Progress in Arabidopsis starch research and potential biotechnological applications. Curr. Opin. Biotechnol. 22:271–280.
Satoh H, Nishi A, Yamashita K, et al (2003) Starch-branching enzyme I-deficient mutation specifically affects the structure and properties of starch in rice endosperm. Plant Physiol 133:1111–1121. doi: 10.1104/pp.103.021527
Scheben A, Wolter F, Batley J, et al (2017) Towards CRISPR/Cas crops - bringing together genomics and genome editing. New Phytol.
Seung D, Soyk S, Coiro M, et al (2015) PROTEIN TARGETING TO STARCH Is Required for Localising GRANULE-BOUND STARCH SYNTHASE to Starch Granules and for Normal Amylose Synthesis in Arabidopsis. PLoS Biol. doi: 10.1371/journal.pbio.1002080
Shumoy H, Raes K (2017) In vitro starch hydrolysis and estimated glycemic index of tef porridge and injera. Food Chem 229:381–387. doi: 10.1016/j.foodchem.2017.02.060
Silver DM, Kötting O, Moorhead GBG (2014) Phosphoglucan phosphatase function sheds light on starch degradation. Trends Plant Sci 19:471–478. doi: 10.1016/j.tplants.2014.01.008
Singh J, Kaur L, McCarthy OJ (2007) Factors influencing the physico-chemical, morphological, thermal and rheological properties of some chemically modified starches for food applications-A review. Food Hydrocoll. 21:1–22.
Singh N, Singh J, Kaur L, et al (2003) Morphological, thermal and rheological properties of starches from different botanical sources. Food Chem. 81:219–231.
Siriwat W, Kalapanulak S, Suksangpanomrung M, et al (2012) Transcriptomic data integration inferring the dominance of starch biosynthesis in carbon utilization of developing cassava roots. In: Procedia Computer Science. pp 96–106
Slade AJ, Fuerstenberg SI, Loeffler D, et al (2005) A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. Nat Biotechnol 23:75–81. doi: 10.1038/nbt1043
Smith AM (2010) Starch and Starch Granules. In: Encyclopedia of Life Sciences.
Smith AM, Zeeman SC, Smith SM (2005) Starch degradation. Annu Rev Plant Biol 56:73–98. doi: 10.1146/annurev.arplant.56.032604.144257
Sonnewald U, Kossmann J (2013) Starches-from current models to genetic engineering. Plant Biotechnol. J. 11:223–232.
Spaenij-Dekking L, Kooy-Winkelaar Y, Koning F (2005) The Ethiopian cereal tef in celiac disease. N Engl J Med 353:1748–1749. doi: 10.1056/NEJMc051492
Sreenivasulu N, Wobus U (2013) Seed-Development Programs: A Systems Biology–Based Comparison Between Dicots and Monocots. Annu Rev Plant Biol 64:189–217. doi: 10.1146/annurev-arplant-050312-120215
Srichuwong S, Sunarti TC, Mishima T, et al (2005) Starches from different botanical sources II: Contribution of starch structure to swelling and pasting properties. Carbohydr Polym 62:25–34. doi: 10.1016/j.carbpol.2005.07.003
Sriroth K, Santisopasri V, Petchalanuwat C, et al (1999) Cassava starch granule structure-function properties: Influence of time and conditions at harvest on four cultivars of cassava starch. Carbohydr Polym 38:161–170. doi: 10.1016/S0144-8617(98)00117-9
Streb S, Delatte T, Umhang M, et al (2008) Starch Granule Biosynthesis in Arabidopsis Is

Abolished by Removal of All Debranching Enzymes but Restored by the Subsequent Removal of an Endoamylase. PLANT CELL ONLINE 20:3448–3466. doi: 10.1105/tpc.108.063487
Streb S, Zeeman SC (2012) Starch metabolism in Arabidopsis. Arab B. doi: 10.1199/tab.0160
Sun Y, Jiao G, Liu Z, et al (2017) Generation of High-Amylose Rice through CRISPR/Cas9-Mediated Targeted Mutagenesis of Starch Branching Enzymes. Front Plant Sci. doi: 10.3389/fpls.2017.00298
Sun ZT, Henson CA (1991) A quantitative assessment of the importance of barley seed alpha-amylase, beta-amylase, debranching enzyme, and alpha-glucosidase in starch degradation. Arch Biochem Biophys 284:298–305.
Svihus B, Uhlen AK, Harstad OM (2005) Effect of starch granule structure, associated components and processing on nutritive value of cereal starch: A review. Anim. Feed Sci. Technol. 122:303–320.
Szymońska J, Targosz-Korecka M, Krok F (2009) Characterization of starch nanoparticles. In: Journal of Physics: Conference Series.
Tadele, Z. and Esfeld K (2009) The Improvement of African Orphan Crops through TILLING. Atdf J 6:44–49.
Tadele Z (2009) Role of Orphan Crops in Enhancing and Diversifying Food Production in Africa. African Technol y Dev Forum J 9–15.
Tako M, Tamaki Y, Teruya T, Takeda Y (2014) The Principles of Starch Gelatinization and Retrogradation. Food Nutr Sci 5:280–291. doi: 10.4236/fns.2014.53035
Tatge H, Marshall J, Martin C, et al (1999) Evidence that amylose synthesis occurs within the matrix of the starch granule in potato tubers. Plant, Cell Environ 22:543–550. doi: 10.1046/j.1365-3040.1999.00437.x
Taylor NJ, Fauquet CM, Tohme J (2012) Overview of Cassava Special Issue. Trop. Plant Biol. 5:1–3.
Taylor S, Scott R, Kurtz R, et al (2009) A practical guide to High Resolution Melt Analysis genotyping.
Tester RF, Karkalas J, Qi X (2004) Starch - Composition, fine structure and architecture. J Cereal Sci 39:151–165. doi: 10.1016/j.jcs.2003.12.001
Tester RF, Morrison WR (1990) Swelling and gelatinization of cereal starches. I. Effects of amylopectin, amylose, and lipids. Cereal Chem 67:551–557.
Tetlow I, Emes M (2017) Starch Biosynthesis in the Developing Endosperms of Grasses and Cereals. Agronomy 7:81. doi: 10.3390/agronomy7040081
Tetlow IJ (2011) Starch biosynthesis in developing seeds. Seed Sci. Res. 21:5–32.
Tetlow IJ, Emes MJ (2011) Starch Biosynthesis in Higher Plants: The Enzymes of Starch Synthesis. In: Comprehensive Biotechnology, Second Edition. pp 47–65
Tian Z, Qian Q, Liu Q, et al (2009) Allelic diversities in rice starch biosynthesis lead to a diverse array of rice eating and cooking qualities. Proc Natl Acad Sci 106:21760–21765. doi: 10.1073/pnas.0912396106
Till BJ, Reynolds SH, Weil C, et al (2004) Discovery of induced point mutations in maize genes by TILLING. BMC Plant Biol. doi: 10.1186/1471-2229-4-12
Tomlinson K, Denyer K (2003) Starch synthesis in cereal grains. Adv. Bot. Res. 40:1–61.
Uauy C, Paraiso F, Colasuonno P, et al (2009) A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat. BMC Plant Biol 9:115. doi: 10.1186/1471-2229-9-115
Ukwuru MU and Egbonu SE (2013) Recent development in cassava-based products research. Acad J Food Res 1:1–13. doi: 10.15413/ajfr.2012.0105
Vamadevan V, Bertoft E (2015) Structure-function relationships of starch components. Starch/Staerke 67:55–68.

Van der Auwera GA, Carneiro MO, Hartl C, et al (2013) From fastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. Curr Protoc Bioinforma. doi: 10.1002/0471250953.bi1110s43

Van Harsselaar JK, Lorenz J, Senning M, et al (2017) Genome-wide analysis of starch metabolism genes in potato (Solanum tuberosum L.). BMC Genomics. doi: 10.1186/s12864-016-3381-z

Varshney RK, Ribaut JM, Buckler ES, et al (2012) Can genomics boost productivity of orphan crops? Nat. Biotechnol. 30:1172–1176.

Vasconcelos LM, Brito AC, Carmo CD, Oliveira EJ (2016) Polymorphism of starch pathway genes in cassava. 1:1–15.

Vrinten PL, Nakamura T (2000) Wheat granule-bound starch synthase I and II are encoded by separate genes that are expressed in different tissues. Plant Physiol 122:255–264. doi: 10.1104/pp.122.1.255

Wang S, Li C, Copeland L, et al (2015a) Starch Retrogradation: A Comprehensive Review. Compr Rev Food Sci Food Saf 14:568–585. doi: 10.1111/1541-4337.12143

Wang S, Li S, Liu Q, et al (2015b) The OsSPL16-GW7 regulatory module determines grain shape and simultaneously improves rice yield and grain quality. Nat Genet 47:949–954. doi: 10.1038/ng.3352

Wang X, Chang L, Tong Z, et al (2016) Proteomics Profiling Reveals Carbohydrate Metabolic Enzymes and 14-3-3 Proteins Play Important Roles for Starch Accumulation during Cassava Root Tuberization. Sci Rep. doi: 10.1038/srep19643

Waterschoot J, Gomand S V., Fierens E, Delcour JA (2015) Production, structure, physicochemical and functional properties of maize, cassava, wheat, potato and rice starches. Starch/Staerke 67:14–29.

Weber A, Servaites JC, Geiger DR, et al (2000) Identification, Purification, and Molecular Cloning of a Putative Plastidic Glucose Translocator. Plant Cell 12:787–801. doi: 10.1105/tpc.12.5.787

Weise SE, Aung K, Jarou ZJ, et al (2012) Engineering starch accumulation by manipulation of phosphate metabolism of starch. Plant Biotechnol J 10:545–554. doi: 10.1111/j.1467-7652.2012.00684.x

Wilm A, Aw PPK, Bertrand D, et al (2012) LoFreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res 40:11189–11201. doi: 10.1093/nar/gks918

Wolter A, Hager AS, Zannini E, Arendt EK (2013) In vitro starch digestibility and predicted glycaemic indexes of buckwheat, oat, quinoa, sorghum, teff and commercial gluten-free bread. J Cereal Sci 58:431–436. doi: 10.1016/j.jcs.2013.09.003

Wong KM, Suchard MA, Huelsenbeck JP (2008) Alignment uncertainty and genomic analysis. Science (80-) 319:473–476. doi: 10.1126/science.1151532

Xin Z, Li Wang M, Barkley NA, et al (2008) Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population. BMC Plant Biol. doi: 10.1186/1471-2229-8-103

Xiong JS, Ding J, Li Y (2015) Genome-editing technologies and their potential application in horticultural crop breeding. Hortic. Res. 2:

Xu X, Huang XF, Visser RGF, Trindade LM (2017) Engineering potato starch with a higher phosphate content. PLoS One. doi: 10.1371/journal.pone.0169610

Xu X, Visser RGF, Trindade LM (2014) Starch Modification by Biotechnology: State of Art and Perspectives. In: Starch Polymers: From Genetic Engineering to Green Applications. pp 79–102

Yan CJ, Tian ZX, Fang YW, et al (2011) Genetic analysis of starch paste viscosity

parameters in glutinous rice (Oryza sativa L.). Theor Appl Genet 122:63–76. doi: 10.1007/s00122-010-1423-5
Yu JK, Graznak E, Breseghello F, et al (2007) QTL mapping of agronomic traits in tef Eragrostis tef (Zucc). BMC Plant Biol. doi: 10.1186/1471-2229-7-30
Yu S, Ma Y, Sun DW (2009) Impact of amylose content on starch retrogradation and texture of cooked milled rice during storage. J Cereal Sci 50:139–144. doi: 10.1016/j.jcs.2009.04.003
Yu TS, Kofler H, Häusler RE, et al (2001) The Arabidopsis sex1 mutant is defective in the R1 protein, a general regulator of starch degradation in plants, and not in the chloroplast hexose transporter. Plant Cell 13:1907–1918. doi: 10.1105/TPC.010091
Yu TS, Zeeman SC, Thorneycroft D, et al (2005) ??-amylase is not required for breakdown of transitory starch in Arabidopsis Leaves. J Biol Chem 280:9773–9779. doi: 10.1074/jbc.M413638200
Zeeman SC (2002) Starch Synthesis in Arabidopsis. Granule Synthesis, Composition, and Structure. PLANT Physiol 129:516–529. doi: 10.1104/pp.003756
Zeeman SC, Kossmann J, Smith AM (2010a) Starch: its metabolism, evolution, and biotechnological modification in plants. Annu Rev Plant Biol 61:209–234. doi: 10.1146/annurev-arplant-042809-112301
Zeeman SC, Kossmann J, Smith AM (2010b) Starch: its metabolism, evolution, and biotechnological modification in plants. Annu Rev plant … 61:209–234. doi: 10.1146/annurev-arplant-042809-112301
Zeeman SC, Northrop F, Smith AM, Ap Rees T (1998) A starch-accumulating mutant of Arabidopsis thaliana deficient in a chloroplastic starch-hydrolysing enzyme. Plant J 15:357–365. doi: 10.1046/j.1365-313X.1998.00213.x
Zhang L, Häusler RE, Greiten C, et al (2008) Overriding the co-limiting import of carbon and energy into tuber amyloplasts increases the starch content and yield of transgenic potato plants. Plant Biotechnol J 6:453–464. doi: 10.1111/j.1467-7652.2008.00332.x
Zhao SS, Dufour D, Sánchez T, et al (2011) Development of waxy cassava with different Biological and physico-chemical characteristics of starches for industrial applications. Biotechnol Bioeng 108:1925–1935. doi: 10.1002/bit.23120
Zhou W, Yang J, Hong Y, et al (2015) Impact of amylose content on starch physicochemical properties in transgenic sweet potato. Carbohydr Polym 122:417–427. doi: 10.1016/j.carbpol.2014.11.003
Zhu F (2015) Composition, structure, physicochemical properties, and modifications of cassava starch. Carbohydr. Polym. 122:456–480.
Zhu F (2017) Barley Starch: Composition, Structure, Properties, and Modifications. Compr Rev Food Sci Food Saf 16:558–579. doi: 10.1111/1541-4337.12265
Zhu F (2018) Chemical composition and food uses of teff (Eragrostis tef). Food Chem. 239:402–415.
Zhu T, Jackson DS, Wehling RL, Geera B (2008) Comparison of amylose determination methods and the development of a dual wavelength iodine binding technique. Cereal Chem 85:51–58. doi: 10.1094/CCHEM-85-1-0051
Zhu W, Buell CR (2007) Improvement of whole-genome annotation of cereals through comparative analyses. Genome Res 17:299–310. doi: 10.1101/gr.5881807

Chapter 2

Modification of cassava root starch phosphorylation enhances starch functional properties.

Wuyan Wang[1], Carmen E. Hostettler[1], Fred F. Damberger[2], Jens Kossmann[3], James R. Lloyd[3], Samuel C. Zeeman[1]*

[1]Institute of Molecular Plant Biology, Department of Biology, ETH Zurich, CH-8092, Switzerland

[2]Institute of Molecular Biology and Biophysics, Department of Biology, ETH Zurich, CH-8093 Zürich, Switzerland.

[3]Institute for Plant Biotechnology, Department of Genetics, University of Stellenbosch, Private Bag X1, Matieland 7602, Stellenbosch, South Africa

***Corresponding Author:**

Samuel C. Zeeman

Department of Biology, ETH Zurich

Universitaetstrasse 2, 8092 Zurich, Switzerland

Tel.: +41 44 632 8275

Fax.: +41 44 632 1664

E-mail: szeeman@ethz.ch

Running Title: Cassava starch phosphorylation and properties

Keywords: *Manihot esculenta* Crantz, Glucan Water Dikinase (GWD), Starch EXcess 4 (SEX4), Like Sex Four 2 (LSF2), gene overexpression, RNAi silencing, starch phosphorylation, starch properties.

Authorship:

In this work, I performed most of the experiments except for the generation of *St*GWD, *MeSEX4* and *MeLSF2* cassava transgenic lines, which was done by Carmen E. Hostettler, the 31P NMR analysis, which was done by Fred F. Damberger, and the RVA analysis, which was conducted by James R. Lloyd from the group of Jens Kossmann.

Summary

Cassava (*Manihot esculenta* Crantz) is a root crop used as a foodstuff and as a starch source in industry. Starch properties are influenced by many factors including the relative amounts of the two glucan polymers amylopectin and amylose, the branched structure of amylopectin, starch granule size and the presence of covalent modifications. Starch phosphorylation is a naturally-occurring modification, whereby phosphate esters are linked either to the C3 or C6 carbon atoms of amylopectin glucosyl residues. This work aimed to explore the potential to alter starch phosphorylation in cassava storage roots. Transgenic plants were generated in which we either overexpressed the potato glucan, water dikinase gene (*StGWD*), or silenced each of the endogenous phosphoglucan phospahatse genes (*MeSEX4* and *MeLSF2*). Overexpressing the potato GWD gene (*StGWD*), which specifically phosphorylates the C6 position, increased the total starch-bound phosphate content at both the C3 and the C6 positions. Silencing endogenous LSF2 gene (*MeLSF2*), which specifically dephosphorylates the C3 position, increased the ratio of C3:C6 phosphorylation without altering total phosphate content. In both cases, amylopectin structure, amylose content and starch granule size were unaltered. We tested several physicochemical properties of the starch extracted from transgenic and control lines. Starch swelling power and paste clarity were both influenced by total phosphate content. However, phosphate position did not significantly influence the functional properties. In conclusion, manipulating the expression of starch phosphorylating/dephosphorylating enzymes can be used to engineer starch phosphate levels in cassava storage root and potentially increased the value of starch for the food and non-food industries.

Introduction

Cassava (*Manihot esculenta* Crantz), a perennial shrub in Euphorbiaceae family, is commercially cultivated as a staple crop in tropical and subtropical regions (Allem and Genéticos 2002; Alves 2002; Puounti-Kaerlas 2002; El-Sharkawy 2004). Its swollen storage roots are rich in starch, and represent an important food source for hundreds of millions of people. In South and South-East Asia, 40% of cassava production is used to produce extracted starch for industrial use. In 2014, the international trade of cassava starch and flour was estimated to be approximately 8.5 million tons (Karlström et al. 2016). Starch is composed of two glucose polymers - amylose and amylopectin. Amylopectin is the major polymer constituting 70% or more of the starch. Its glucosyl units are α-1,4-linked to form chains that are connected by α-1,6-bonds yielding a tree-like or racemose structure. Its branches are clustered and interact to form double helices that pack into semi-crystalline lamellae, resulting in highly-ordered, insoluble starch granules (Zeeman et al. 2010b; Pfister and Zeeman 2016b).

Besides its nutritional value, starch is extensively used as a texturizer in the food industry. Cassava starch offers certain advantages over other starches. First, it is inexpensive compared with corn starch. Second, it has a low gelatinization temperature and produces relatively clear, high-viscosity pastes. Third, its bland taste makes it preferable as additive for processed food with mild flavor (Raphael et al. 2011; Vasconcelos et al. 2016). Starch is also a feedstock for non-food industries: it is used in fuel-ethanol production, in paper and textile production, and in the pharmaceutical industry as an inert carrier (Balagopalan 2002; Montero 2003). In order to fulfil these different functional uses, starch and starch derivatives need to have diverse physicochemical properties (Santelia and Zeeman, 2011; Alcázar-Alay and Meireles 2015). This diversity is obtained partly by using starches from different botanical sources and partly through chemical, physical and enzymatic treatments performed after extraction. These processes typically modify the starch polymers by fragmenting them, by crosslink them, or by covalently adding positively- or negatively-charged functional groups. Phosphorylation is one such post-extraction modification. It increases the hydration capacity of starch pastes, thereby influencing peak viscosity, gel-forming capacity, swelling power and paste stability (Jobling 2004; Carpenter et al. 2015).

Phosphorylation is also a naturally-occurring starch modification. In wild-type starches phosphate groups are bound primarily to the C6 position of amylopectin glucosyl residues with

smaller amounts bound to the C3 position. In the past two decades, the enzymes responsible for the reversible phosphorylation of starch in plants were discovered and this process was shown to play an important role in starch metabolism. Phosphorylation is mediated by two dikinases, namely Glucan, Water Dikinase (GWD) and Phosphoglucan, Water Dikinase (PWD) (Lorberth et al., 1998; Yu et al., 2001; Kötting et al. 2005; Baunsgaard et al., 2005). GWD and PWD both transfer the β-phosphate group of ATP to amylopectin glucosyl residues, phosphorylating the C6 and C3 positions respectively. The γ-phosphate group of ATP is released as orthophosphate (Ritte et al. 2006a). For PWD to function, it requires the preceding phosphorylation of amylopectin by GWD (Kötting et al., 2005; Baunsgaard et al., 2005). Thus, the starch from the *gwd* mutant is phosphate free, whereas *pwd* mutant starch lacks phosphate at the C3 position. Two phosphoglucan phosphatases responsible for amylopectin dephosphorylation have also been identified, namely Starch EXcess 4 (SEX4) and Like-SEX Four 2 (LSF2), mainly through work on transitory leaf starch in Arabidopsis. SEX4 releases phosphate from both the C6 and C3 positions of amylopectin, with the preference for C6. In contrast, LSF2 is specific for C3-bound phosphate (Gentry et al., 2007; Kötting et al., 2009; Santelia et al. 2011; Silver et al. 2014).

Although starch phosphorylation proceeds concurrently with starch synthesis, the biological role of phosphorylation is thought to be in initiating the starch degradation process in living cells. The lack of GWD or PWD results in *starch-excess* phenotypes in leaves and retardation in growth, which is particularly severe in case of *gwd* mutants (Lorberth et al., 1998; Yu et al., 2001; Nashelivitz et al., 2009; Weise et al., 2012). Phosphorylation is proposed to disrupt the semi-crystalline packing of amylopectin, making the glucan chains at the granule surface more accessible to glucan degrading enzymes such as amylases (Edner et al. 2007; Hejazi et al., 2009). Phosphoglucan phosphatases are also required for normal rates of starch degradation. The loss of SEX4 also causes a *starch-excess* phenotype in leaves and decreases plant growth (Zeeman et al., 1998; Kötting et al. 2009), which is exacerbated by the additional loss of LSF2. This is explained by the fact that phosphate groups interfere with the full degradation of starch and, after disrupting the semi-crystalline parts of starch, the phosphate groups need to be removed again concurrently with amylolysis (Kötting et al., 2009; Hejazi et al., 2010). Consequently, phosphoglucan phosphatase-deficient mutants accumulate phosphorylated intermediates of starch degradation or have increased levels of starch-bound phosphate (Kötting et al., 2009; Santelia et al. 2011). Both GWD and SEX4 have been reported to be

redox-regulated enzymes, based on the formation of an intramolecular disulfide bonds between two cysteines (Mikkelsen et al., 2005; Silver et al., 2013). When reduced, the enzymes are active whereas they are inactive in the oxidized state. In the case of potato GWD, substitution of one of the cysteines with a serine resulted in an active, redox-insensitive enzyme (Mikkelsen et al. 2005).

Various phosphate levels have been reported for starches from different species and different plant organs (Blennow A, 1998). Storage starch from cereal seeds has very low phosphate (less than 1 nMol Glc-6-P/mg starch) compared to other starches. This is probably due to the special circumstances of starch degradation in the germinated grains, which is essentially extracellular, mediated by secreted enzymes, and is unlikely to involve reversible glucan phosphorylation. In contrast, potato tuber starch has particularly high levels of starch-bound phosphate (8-33 nMol Glc-6-P/mg starch), while the levels in cassava root starch is rather low (2.5 nMol Glc-6-P/mg starch). The discovery of the starch phosphorylating/dephosphorylating enzymes has opened the door to engineering starch phosphate levels and starch content in crops through manipulating their expression. This has been applied in potato (*Solanum tuberosum*), barley (*Hordeum vulgare*), maize (*Zea mays*), wheat (*Triticum aestivum*), and rice (*Oryza sativa*) (Lorberth R et al. 1998; Carciofi et al. 2011; Lanahan Michael B et al. 2005; Schewe G et al. 2003; Cheng YL et al. 2017). This has been particularly effective in species with low starch-bound phosphate; an almost tenfold increase in grain starch-bound phosphate was achieved by heterologous overexpression of *StGWD* in barley endosperm. This high phosphate level was accompanied by altered starch granule morphology and thermal properties, enhancing economically important traits (Carciofi et al., 2011). Such biotechnological approaches are potentially valuable in both agriculture and industry, potentially increasing the value of starch and decreasing the need for post-extraction modifications.

To explore the potential to alter starch phosphorylation in the storage root starch of cassava, transgenic plants were generated in which we either overexpressed the potato GWD gene (*StGWD*; both the wild-type protein and a redox-insensitive version of it), or silenced each of the endogenous phosphoglucan phosphatase genes (*MeSEX4* and *MeLSF2*). Both approaches altered starch-phosphorylation in the predicted way resulting in starches with distinct properties.

Results

Generation of cassava transgenic plants either overexpressing the *StGWD* gene, or repressing the endogenous *SEX4* or *LSF2* genes.

To increase the rate of starch phosphorylation we generated constructs to overexpress in cassava either the wild-type *St*GWD protein or a mutant version of it (*St*GWDm) that is insensitive to redox regulation (Mikkelsen et al. 2005) (Figure S1a). Both genes were placed under the control of the constitutive cauliflower mosaic virus 35s promoter (CaMV 35S), and were transformed into the cassava variety TMS60444 through Agrobacterium-mediated friable embryonic callus (FEC) transformation (Bull et al. 2009; Talyor et al. 2004). Putative transformants displaying both shoot and root regeneration under antibiotic selection were analysed for the presence of the transgenes by PCR. Thirteen independent 35S::*StGWD* lines and 35 35S::*StGWDm* lines were selected. The *St*GWD protein expression in leaves of *in-vitro* grown plantlets was detected by immunoblotting using a polyclonal antibody raised against the recombinant *St*GWD protein (Eurogentec, Seraing, Belgium). The *St*GWD protein levels varied from low to high among the transgenic lines (Figure S1b). Of the 35S::*StGWD* lines (hereafter referred to as 'G' lines), G18, G2 (high expression), G16 (lower expression), and G21 (no *St*GWD protein detected) were selected for further analysis. From the 35S::*StGWDm* lines (hereafter referred to as 'R' lines), R9, R41, R46 (high expression), and R49 (no detectable expression) were selected. These plants were transferred to soil and grown for 8 months. In storage roots, the expression level was in some cases similar to that in leaves, but not in all. Of the G-lines, only G18 had detectable levels of *St*GWD protein while of the R-lines, both R9 and R46 had detectable levels of protein (Figure 1a).

In the greenhouse, the G- and R-lines generally grew similarly to the wild type, albeit with slightly shorter stems (Figure 1b, Figure S2a). The fresh weight of the storage roots in the transgenic cassava tended to be lower than in the wild type (Figure S2b). Storage root yield varied from one generation to another, presumably due to year-on-year variation in greenhouse conditions. Neither plant height nor storage root yield correlated with expression of the *St*GWD proteins, suggesting that there was a more general influence of cassava transformation on growth relative to the wild-type plants.

To decrease the rate of starch dephosphorylation, we generated RNAi hairpin constructs against either *MeSEX4* or *MeLSF2* gene and expressed them under the control of the *Solanum tuberosum* patatin B33 promoter (Figure S3). The constructs were transformed into TMS60444, again via Agrobacterium-mediated transformation. Antibiotic selection and screening for the presence of the transgene by PCR on genomic DNA resulted in 20 independent *MeSEX4* RNAi lines and 30 independent *MeLSF2* RNAi lines. The *MeSEX4* lines (hereafter referred to as 'S' lines) displayed severe growth phenotypes, being dwarfed and slow-growing. Only four lines (S31, S84, S107 and S132) grew sufficiently well to be stably propagated in soil (Figure 1C, Table S1) and analysed further. In contrast, the *MeLSF2* lines (hereafter referred to as 'L' lines) grew well (Figure 1c). The expression level of the targeted gene in the storage roots of each line was analysed, since the patatin promoter is known to drive expression in the cassava storage root (Naumkina et al. 2007). Immunoblotting using an antibody raised against the *Arabidopsis thaliana* SEX4 protein recognised a protein of the correct molecular weight in cassava root extracts. The blots indicated efficient silencing of *Me*SEX4 protein in storage root of all four S-lines (Figure 1d). Since an antibody recognising the *Me*LSF2 protein was not available, RT-PCR was used to assess *MeLSF2* gene expression. This revealed efficient transcriptional repression in the storage roots of most of the transgenic lines (Figure 1e). Gene expression in leaves of selected S- and L-lines were also checked by RT-PCR, revealing either similar expression levels to the wild type for the S-lines (Figure S4a), and either a similar or a small degree of repression in L-lines (Figure S4b).

Starch phosphate content and distribution is significantly altered in storage root of transgenic cassava lines

The total phosphate content of starch purified from storage roots was determined using the malachite green assay. Significant increases in starch-bound phosphate were observed in the G- and R-lines, where *St*GWD protein was detected in the storage root (Figures 1a and 2a). The highest level was observed in line R9, which had around twice as much starch-bound phosphate as the wild type in all three harvested generations (Figure 2a). R46 had increases of above 50%, and G18 had a significant increase in 2016 but not in the preceding 2 years. In the low-expressing lines G2, G16, R41 and R49, there was no significant increase in phosphate content (Figure 2a). To investigate whether the distribution of starch-bound phosphate between the C3 and the C6 positions, ^{31}P NMR was performed. All selected G- and R-lines showed

similar ratios of C3:C6 phosphate as in the wild type (Figure 2b). This is interesting since GWD specifically phosphorylates at the C6 position, while PWD phosphorylates the C3 position.

For the S-lines, the total starch-bound phosphate was slightly increased in three of the four cases over two generations, whereas in the L-lines, the level was similar to the wild type (Figure 3a). In the S-lines, no alteration in the ratio of C3:C6 phosphate was detected whereas in the L-lines, there was a significant increase in the proportion of C3-bound phosphate (from 27% in the wild-type to 35% in the *MeLSF2* RNAi lines; Figures 3b and 3c). Analysis of additional L-lines (L2, L23, L98, L101, L129 and L132) further confirmed the relative increase in C3-bound phosphate established that it was inversely correlated with the expression of the *MeLSF2* gene ($r^2 = 0.71$, $P = 0.004$; Figure 3d).

Starch structure and granule morphology are unaffected by altered phosphate

The starch granules from the storage roots of our transgenic cassava, viewed by scanning electron microscopy were irregularly-shaped, with ovoid, polygonal and round granules observed. The starch granule surfaces were predominantly smooth with some concave pits (Figure 4a). There were no obvious differences in the appearance of the starch granules among the transgenic and the wild-type samples. The micrographs showed that the granule size ranged from 3 to 30 µm. Further quantification of the particle size distributions within this range by laser diffraction indicated that the modal particle size was around 12.4 µm (Figure 4a, Table 1). Similar granule sizes were detected for the transgenic lines (modal values ranging from 10.3 to 13.6 µm; Table 1).

Amylose content - an important factor influencing starch properties - was determined in the storage root starch of our transgenic lines: Amylose content ranged from 16.9 to 22.1%, but there was no statistically significant differences amongst the transgenic lines (except line S31, the amylose content of which was slightly lower than that of wild-type plants; Table 1). We also analysed the chain length distribution (CLD) of the starches to provide insight into branched amylopectin architecture. After enzymatic debranching, the resultant linear glucan chains were analysed by high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD). The distribution of chains ranging in length from DP (degree of polymerization) 3 to 60 revealed a modal value of DP 13 (Figure 4b). No difference in distribution was observed among the transgenic lines compared to the wild type (Figure 4b).

Thus, the starch characteristics we tested here (granule size, granule surface morphology, amylose content and chain length distribution) were generally not influenced by the alteration of phosphate content.

Cassava starch physioco-chemical properties and functionality is influenced by phosphate content but not phosphate position

Alteration of the number and/or distribution of negatively-charged phosphate groups might influence starch functional properties. To investigate whether the variation in the amount of starch-bound phosphate in our cassava lines changed starch swelling power, we heated samples of starch (from the G- and R-lines) that contained different amounts of phosphate in water at 55°C. The weight of wet sediment formed during heating was determined and the swelling power calculated based on the increased weight. The line with the highest phosphate, R9, had a significantly higher swelling power (19.4 g/g) compared to the wild type (12.9 g/g; Figure 5a). Starches from lines R46 and G18, which also had high starch-bound phosphate, similarly showed increased swelling powers (17.8 and 15.6 g/g, respectively). Starches from the other transgenic lines, where there was no increase in starch-bound phosphate (G2, G16, R41 and R49) had swelling power similar to wild-type starch (Figure 5a).

Next we analyzed paste clarity - another important starch property. After gelatinization and initial measurement of clarity (as light transmittance), the starch was stored at 4°C and the clarity re-measured over the subsequent four days. Light transmittance gradually decreased due to retrogradation, reaching a stable level after 48 h. While this pattern was observed in all our cassava lines (Figure 5b), the pastes of the high-phosphate starches (R9, R46 and G18) displayed an improved initial and/or final light transmittance (Figure 5b). This suggests that the charged phosphate groups help to keep starch molecules hydrated and reduce the extent to which they re-crystallize.

Differential scanning calorimetry (DSC) was used to determine starch thermal properties during gelatinization. No significant correlation was observed between starch phosphate content and gelatinization temperature or enthalpy (Table S2). Among all selected lines, DSC peak temperatures ranged from 59.8 to 62.8°C and the gelatinization enthalpies were ranged

from 15.1 to 19.2 Jg^{-1} (Table S2). There was a negative correlation between DSC gelatinization temperature and enthalpy amongst the samples.

An overview of the significantly correlated physico-chemical properties (swelling power, paste clarity and DSC) and compositional characteristics (phosphate content, granule size and amylose content) of starch strongly suggest that the phosphate content of cassava starch influences swelling power ($r = 1.0$, $P < 0.001$) and paste clarity ($r = 0.9$, $P < 0.01$), but not gelatinization temperature, granule size or amylose content (Figure 6). We also performed rapid viscometric analyses (RVA) of selected G-and R-line starches. While there was some variation in the viscosity profiles, this did not correlate with phosphate content (Figure S5).

To investigate whether the variation of the ratio of C3-bound to C6-bound phosphate results in any changes in starch properties, the *MeLSF2* RNAi lines with different levels of C3-bound phosphate were analysed. However, no major changes in light transmittance were detected among the lines, except for a slight decrease in L99 (Figure 7). The other starch physico-chemical properties (swelling power and DSC-based gelatinization) of L-lines were also tested, but no significant correlations were observed (data not shown). Thus, unlike phosphate amount, there is no significant change in cassava starch properties resulting from the distribution of the phosphate between C6 and C3 positions.

Discussion

This work demonstrates that, using transgenic methods, it is possible to increase the phosphate content of cassava storage root starch. This leads to important changes in starch properties that can be valuable for industries that use starch as a raw material. In addition, this work shows that using genes discovered and studied in Arabidopsis leaves it is possible to change the relative abundance of phosphate at the C3 and C6 positions of the glucosyl residues of cassava starch amylopectin. This demonstrates a strong degree of conserved function between species and plant tissues.

Control of starch phosphorylation in the storage root of cassava

The particularly high levels of covalently bound phosphate in potato starch (reportedly 4-10 times higher than cassava root starch) led us to overexpress the *St*GWD gene based on the hypothesis that it may be fully or partly responsible for this potato starch trait. Further, we also over-expressed a mutant version of the *St*GWD protein to circumvent potential redox-regulation in the storage root amyloplasts. In several cassava lines, expression of either proteoform successfully increased the amount of starch-bound phosphate. This is consistent with the findings of previous studied in other crops, such as barley (Carciofi et al., 2011) and rice (Cheng YL et al. 2017), where the starch phosphate levels were increased by overexpressing *St*GWD – in these cases in the seed starchy endosperm.

In our cassava transformants, starch-bound phosphate in the storage roots was elevated upto twice the level seen in wild-type cassava (Figure 2a). This is still not as high as the amounts seen in potato tuber starch suggesting that factors other than the *St*GWD itself contribute to the high levels. Such other factors could include the concentration of ATP, which may be higher in potato amyloplasts than in cassava root amyloplasts, differences in the structure of the starch itself between potato and cassava. It is also possible that the other enzymes known to phosphorylate and dephosphorylate starch play a role in determining the overall level. Our data suggest that this latter point is at least part of the explanation. For example, in our *St*GWD-overexpressing lines, the ratio of phosphate bound to both the C6 and C3 position of the glucosyl residues of amylopectin remained the same as in the wild type (Figure 2b). This means that an increase of phosphate at the C3 position was observed when expressing the *St*GWD protein. It is known that C6- and C3-phosphorylation is selectively catalyzed by GWD and

PWD, respectively (Ritte G. et al., 2006), and that PWD is strictly dependent on the pre-phosphorylation by GWD (Kötting et al., 2005; Baunsgaard et al., 2005). Thus, it appears that the activity of endogenous *Me*PWD is increased in *St*GWD overexpression lines leading to the higher phosphorylation at C3 position.

The phosphoglucan phosphatases SEX4 and LSF2 have been shown to influence both the amount and distribution of starch-bound phosphate in Arabidopsis. Our work shows that some of these findings can be transposed to cassava. Repression of *MeLSF2* expression increased the ratio of C3-bound to C6-bound phosphate. Multiple, stably propagated lines with different degrees of *MeLSF2* repression were obtained, in which a correlated increase in the proportion of C3-bound phosphate was evident (Figure 3d). These results are consistent with the findings in Arabidopsis, where LSF2 was shown to specifically remove the C3-bound phosphate. In that system, the increase in C3:C6 ratio in the *lsf2* null mutant was greater than was observed here and it was accompanied by an increase in the total starch bound phosphate. It seems likely that these differences can be explained by that fact that the repression of the *MeLSF2* gene was incomplete.

Our results with *MeSEX4* RNAi lines were mixed and not easily interpretable. In contrast to all of the other transgenic lines, these plants grew very poorly during regeneration *in vitro* and very few survived the transfer to soil. The four successfully transferred lines continued to grow poorly on soil. The growth of Arabidopsis *sex4* mutants was also slower than the wild type (Zeeman et al., 1998a), and was accompanied by a strong starch-excess phenotype in the leaves. In that case, the incomplete mobilization of stored reserves is thought to limit vegetative growth (Stitt and Zeeman, 2012). However, in our transgenic cassava lines, *MeSEX4* expression was decreased only in the storage roots and not in the leaves (Figure S4a), and no starch-excess phenotype was observed in leaves either by qualitative iodine staining or quantitative measurements (not shown). Nevertheless, the phosphate levels in the storage root starch were slightly increased, as might be expected for the loss of a phosphoglucan phosphatase, but the strength of this effect was somewhat dependent on the year harvest (Figure 3a). The C3:C6 phosphate ratio was similar to the wild type (Figure 3b and 3c). Presently, the cause of the poor growth phenotype remains unclear. If it is direct effect of the repression of the *MeSEX4* gene, more work will be required to identify the affected tissue and reason for growth retardation.

Alternatively, it may be an off-target effect of the silencing construct used. If so, it could be circumvented using an alternative method for silencing of the *MeSEX4* gene.

Impact of starch phosphorylation on cassava starch properties

Phosphate esters are hydrophilic and can enhance starch hydration. Indeed, it is thought that the major biological function of starch phosphorylation is the solubilization of the semi-crystalline structures of starch to assist and control its degradation. Thus, it is not surprising that the presence of phosphate groups leads to the changes in the valuable physico-chemical properties of starch (Jobling et al., 2004). Our study provides a clear picture of this influence since the elevation of starch-bound phosphate in storage root starch (e.g. line R9) occurred without affecting other features of starch. This is important since amylopectin structure, amylose content, and starch granule morphology are all well known to significantly influence its functional properties. Furthermore, it is quite common for genetic perturbations to affect multiple starch traits at once. For example, in rice and barley, impressive, ten-fold increases in starch-phosphate were achieved by GWD overexpression in the grain. This altered starch physico-chemical properties such as gelatinisation enthalpy but also changed features such as starch granule morphology and amylose content (Carciofi et al., 2011; Chen et al. 2017).

In our study, the two-fold increase in starch-bound phosphate was more modest (partly because cassava starch naturally has more bound phosphate than that of rice or barley; Blennow et al., 1998; Carciofi et al., 2011). Our analysis of amylopectin structure, amylose content, and starch granule morphology revealed them essentially unchanged, allowing us to attribute the altered physico-chemical properties directly to the phosphate content. The swelling power of starch and the clarity of the starch pastes - two desirable traits in both the food and non-food industries (Alcazar-Alay SC et al., 2015; Waterschoot J et al., 2014; Singh N et al., 2003), were both strongly and positively correlated with starch phosphate content (Figures 5 and 6), while other, gelatinisation-related properties were unaffected.

Our study also allows us to propose that it is the amount of starch-bound phosphate, rather than its distribution between C3 and C6 positions of the glucosyl residues, that is most important for its functional properties (Figure 7). On one hand, this is surprising since the C3-bound phosphate is proposed to have a greater disruptive impact on the semi-crystalline packing of

native starch (Hansen et al., 2009). On the other hand, it is perhaps unsurprising that a trait like paste clarity – determined by physical processes that occur after starch gelatinization (where the natural semi-crystalline starch structure is lost) is insensitive to the location of the phosphate.

Potential for improved cassava starch through transgenic and non-transgenic means

Our data demonstrate that it is feasible to alter the phosphate content of cassava root starch and thereby its functional properties through biotechnological means. This work was performed on the cultivar (cv) 60444, which is amenable to genetic transformation, but not widely grown in agriculture. Future approaches using similar RNAi-based methodologies in other, farmer-preferred cultivars could realize this opportunity to create a higher value cassava crops. It is also plausible that alternative CRISPR/Cas9-based strategies could be used, either for mutation (e.g. of the LSF2 gene) or for targeted genome editing (e.g. to alter the expression and/or regulation of the *MeGWD* and *MePWD* genes), to create similarly enhanced, transgene-free cassava lines. Furthermore, while we have studied each gene target one at a time, a multi-target approach to simultaneously increase phosphorylation and decreased dephosphorylation may be even more successful in elevating starch bound phosphate. This may allow levels seen in potato or even higher to be achieved, gaining even greater value and broader applications for cassava in the starch industry.

Experimental procedures

Generation of transgenic lines

Cassava (*Manihot esculenta* cv 60444) was used as the wild type. For overexpression of GWD, two constructs harbouring coding sequence of either the wild-type potato GWD (pCAMBIA2300::*StGWD*) or a redox-insensitive version of it (pCAMBIA2300::$StGWD_m$) were obtained from Mikkel Glaring (University of Copenhagen, Copenhagen, Denmark). The redox-insensitive version contains a nucleotide modification leading to an amino acid substitution of cysteine 1084 to a serine (C1084S) at the peptide level. For RNA interference lines (RNAi), hairpin constructs targeting either the *MeSEX4* or the *MeLSF2* genes were designed. The protein sequences of *At*SEX4 and *At*LSF2 were used as a query to identify orthologous proteins in cassava (Manes.10G053500.1 and Manes.10G005000.1, respectively; www.phytozome.net). Hairpin constructs contained an antisense fragment, a loop, and a sense fragment designed for *MeSEX4* or *MeLSF2*, were subcloned into a modified pCAMBIA1301 vector containing the *Solanum tuberosum* patatin B33 promoter. Transgenic cassava lines were obtained via Agrobacterium-mediated friable embryonic callus (FEC) transformation as described in (Bull et al., 2009; Niklaus et al., 2011). Transgenic plantlets selected by antibiotic resistance and confirmed for the presence of the transgene by PCR on genomic DNA were transferred to soil.

Plant materials and growth condition

Cassava plants grown in a greenhouse with a minimum of 14 h of light, 60% humidity, day/night temperatures of 24 and 17°C, respectively. Greenhouse grown plants were propagated via stem cuttings from the mother plants. The stem cuttings contained at least 2 buds and were planted first in Jifi pots for rooting. Within 2 months, the rooted cuttings were transferred to 15-cm round pots containing 40% Klasmann Substrate 2, 10% Perlite, 50% Ricoter lawn soil, fertilised with Scotts Osmocote.

Quantitative RT-PCR

Eighty milligrams of cassava storage roots or 20 mg of leaves were powdered in liquid N_2 and vortexed with 600 µL RNA extraction buffer (150 mM Tris, pH 7.5, 2 % [w/v] SDS, 50 mM EDTA). Ethanol absolute (150 µL), potassium acetate (5M, 66 µL) and chloroform: isoamyl alcohol (24:1; 750 µL) were then sequentially added to the mix, vortexed and separated by

centrifugation at 13,000 g for 3 min. The supernatant (600 µL) was taken and further cleaned by vortexing with 600 µL phenol:chloroform:isoamyl alcohol (25:24:1). Ethanol was added to the supernatant and the mixture was incubated at -80 °C for 30 min to precipitate nucleotides. The precipitate was washed with 80% (v/v) ethanol, dissolved in 75 µL DEPC water and RNA precipitated by addition of 25 µL 8 M LiCl in DEPC water for 16 h at -20 °C. The RNA pellet was washed with 80% (v/v) ethanol, dissolved in 30 µL DEPC water and DNase I treated. Complementary DNA was synthesized using the SuperScript III kit (Invitrogen) and oligo(dT) primers. Real-time quantitative PCR was carried out using the SYBR Green Supermix (Eurogentec) with Applied Biosystem 7500/7500 Fast. PP2A was used as a reference gene. Gene-specific transcripts were normalized to PP2A and quantified by the ΔCt method (Ct of gene of interest – Ct of PP2A gene). Primer sequences (5' to 3') were as follows: *MeLSF2*, TGAGGAACCCATATGAGTACCA and GCTGCAAATTTAGAATGTAGGCC (150-bp amplicon); *MeSEX4*, ATTCAGCATCTACGTGCAGA and CTCTTCCTAGCCCAGCAGTG (145-bp amplicon); *MePP2A*, TGCAAGGCTCACACTTTCATC and CTGAGCGTAAAGCAGGGAAG (150-bp amplicon).

Protein Extraction and immunoblotting detection

Powdered frozen root material (0.3 g) was homogenized in 1 mL ice-cold extraction medium (100 mM MOPS, pH 7.2, 1 mM EDTA, 10% [v/v] ethanediol, 1% [w/v] PVPP, 1 mM DTT, 1 x Complete proteinase inhibitor [Roche]) using a pre-chilled glass homogenizer. After centrifugation (10 min, 16,000 *g*, 4°C), protein in the supernatant was determined using the Bradford method (Bradford 1976). Equal amounts of protein were separated by SDS-PAGE, electroblotted onto polyvinylidene fluoride membranes and detected with antibodies raised against recombinant *St*GWD protein (Eurogentec, Seraing, Belgium), or against the *Arabidopsis thaliana* SEX4 (Niittylä et al., 2006).

Iodine staining

Ten cassava leaves (counting from top) were harvested at the end of the day or end of the night from each plant, cleared in 80% (v/v) ethanol, rinsed in water and iodine-stained with Lugol's solution (Sigma-Aldrich).

Starch quantification

The method of Smith and Zeeman (2007) was used. Briefly, powdered frozen material (leaf or root) was homogenized in 1.12M perchloric acid. After centrifugation (3000 *g*, 10 min, 4°C) the insoluble material was washed four times in 80% (v/v) ethanol, resuspended in water and used for starch quantification. Starch in the insoluble fraction was gelatinized at 95°C for 15 min and digested to glucose at 37°C using α-amylase and amyloglucosidase (Roche). Starch content (in glucose equivalents) was determined by quantifying the released glucose with a hexokinase/glucose-6-phosphate dehydrogenase-based spectrophotometric assay.

Starch purification

Cassava storage roots were harvested, chopped, immediately frozen in liquid N_2 and ground to powder using a Geno Grinder (producer). Cassava powder (10 ml) was homogenized in a Waring Blender in 100 ml of 50 mM Tris-HCl, pH 8.0, 0.2 mM EDTA, 0.5% (v/v) Triton X-100. Insoluble material was collected by centrifugation (15 min, 3000 g), resuspended in the same medium and sequentially filtered through 100 μm and 60 μm nylon meshes. Starch granules were purified by sedimenting through a 95% (v/v) Percoll cushion at 2,500 *g* for 15 min. The starch pellet was washed at least three times in 0.5% (w/v) SDS, and the SDS then washed away by at least three water washes. Purified starch was then dried under vacuum for 48 h.

Total starch-bound phosphate content

Purified starch granules (5 mg) were suspended in water and acid-hydrolyzed in 50 μL 2 M HCl for 2 h at 95°C. The solution was neutralized with 1 M NaOH. Fifty microliters of supernatant was then treated with 15 units of Antarctic Phosphatase (2 h, 37°C). Released phosphate was determined using the malachite green reagent with absorbance at 660 nm determined spectrophotometrically.

^{31}P NMR Assay

Samples for ^{31}P NMR analysis were prepared according to Santelia et al. (Santelia et al. 2011). Briefly, starch (50 mg) was suspend in 3 mM NaCl, 1 mM $CaCl_2$, pH 6, and digested with α-amylase and amyloglucosidase (Roche). After adjusting the sample pH to 6 and adding 5% (v/v) D_2O, ^{31}P NMR spectra were measured on an Avance III 500 spectrometer equipped with a CPQCI CryoProbes with cryo-cooled ^{31}P preamplifier and a Z axis-pulsed field gradient unit (Bruker) at 303K. ^{31}P spectra of cassava starch were obtained for 5-10k transients (3.8 s recycle

delay, 1 s acquisition, 10 Hz line-broadening). All spectra were processed, analysed and plotted with Topspin 3.2 (Bruker). ^{31}P spectra were calibrated indirectly with the DSS signal as external reference with Ξ=0.404807356 (Maurer and Kalbitzer 1996). Wild-type Arabidopsis starch NMR was used as a reference for peak identification. Percentages of 3P and 6P peaks were obtained from integrals of baseline-corrected spectra. Errors were estimated from technical triplicates and verified for the wild type and some transgenic lines by performing analyses on replicate samples.

Scanning electron microscopy (SEM)

A starch-water mix was loaded to double-sided adhesive carbon tape mounted on SEM stub, air dried and then sputter-coated with osmium (2 nm thickness). Samples were visualized in Hitachi SU5000 microscope at an acceleration voltage of 5.0kV and a working distance of 6-7 mm.

Chain length distribution

Purified starch (0.2 mg) was suspended in 450 µL water. Samples were heated to 100 °C for 15 min and cooled to 20°C. Sodium acetate (10 mM, pH 4.8) was added together with 1 unit pullulanase and 0.04 units of isoamylase (Megazyme) to debranched the samples (37°C, 2.5 h). The reaction was stopped by heating (99°C, 10 min) and clarified by centrifugation (5 min, 16,000 *g*). The supernatant was applied to sequential 1.5mL columns of DOWEX50 and DOWEX1 to remove the charged compounds. Neutral compounds were eluted in 4 mL water and freeze-dried. After re-dissolving in water by heating (99°C, 5 min) linear chains were separated and detected by HPAEC-PAD (ICS-5000, Dionex) using a CarboPack PA-200 column as described previously (Streb et al. 2008).

Granule size distribution

Six milligrams of purified starch granules were suspended in 2 mL water, sonicated for 1 min disaggregate the granules. Granule size distribution was analysed with a LS 13 320 Laser diffraction particle size analyzer (Beckman Coulter). The starch granule suspension was further diluted in water to attain an obscuration level about 10% and continuously stirred to prevent sedimentation.

Amylose content assay

Purified starch in water (5 μg μL^{-1}) was used for determining amylose content essentially according to (Zhu et al. 2008). Samples were incubated at 98°C for 2 h with intermittent vortexing. Five microliters of gelatinized sample was stained with 10% Lugol solution. The absorbance was immediately measured at 510 and 620 nm and amylose content calculated based on potato standards. A standard curve between amylose content and absorbance difference (620 nm – 510nm) was firstly generated. Potato starches (Sigma) with different amylose contents (0, 10%, 20%, 30%, 50% and 80%) were used. The resulted equation of standard curve was: absorbance difference (y) = 0.0105 amylose content (x) – 0.2887 (R^2 for the standard curve = 0.99692).

Swelling power assay

Swelling powered was measurement by using the method from Kusumayanti et al. (2015) with modifications. The starch dispersion (10 mg starch per milliliter water) was heated at 55 °C for 1 hour with constant agitation followed by centrifugation at 371 g for 15 min. Swelling power was calculated according to the equation: Swelling power = weight of sediment paste /dry weight of the sample (g/g).

Paste clarity

The paste clarity was determined according to Craig et al. (1989). Pastes were produced by suspending 10 mg starch in 1 mL water and heating to 90°C for 30 min with constant agitation. After cooling to 30°C for 1 h, light transmittance of the paste was measured at 650 nm against water. Samples were stored at 4°C and the transmittance was remeasured every 24 h.

Differential scanning calorimetry (DSC) Assay

The thermal parameters of melting temperature (DSC-Tp) and the heat of melting (DSC-H) were determined using Mettler-Toledo DSC with STARe evaluation software. Three milligrams of starch was added to 30 uL water at 20°C for 24 h. Eleven milligrams of the water and starch mixture was sealed into an aluminum pan for measurement. DSC was carried out in an N_2 atmosphere at heating rate of 5°C/min from 30 to 100°C.

RVA

Purified starch was evaluated for paste viscosity using a RVA4500 Rapid Viscoanalyzer (Perten Instruments). Gelling properties were recorded in a by placing a 16% (w/v) starch/water

slurry in the instrument and stirring for 10 s at 960 rpm and then at 160 rpm for the remaining time. The temperature profile was as follows: 50°C for 1 min, followed by a linear increase to 95°C over 3 min 42 s, hold at 95°C for 2 min 30 s, cool to 50°C over 3 min 48 s and hold at 50°C for 2 min. Viscosity was recorded continuously.

Statistical analysis

Two tailed unpaired t-tests, used for determining significant difference between modified starches and control samples in C3 phosphate percentages and in swelling power, were conducted with Prism software, as were Dunnett's multiple comparison tests for amylose content and starch granule size. The least significant values were calculated at 5% probability. Relationships between starch characters and starch physico-chemical properties for GWD-expressing lines were analyzed by the means of Pearson correlation and displayed as heat map with the relevant P-value indicated.

Accession numbers

The gene NCBI and Phytozome accession numbers, respectively, for the following genes are: *StGWD*, XM_006357557.2 and PGSC0003DMT400019845; *MeLSF2*, XM_021769917.1and Manes.10G005000.1; *MeSEX4* XM_021770800.1 and Manes.10G053500.1; *MePP2A*, XM_021767530.1 and Manes.09G039900.1

Acknowledgments

We thank Wilhelm Gruissem, Herve Vanderschuren and Simon Bull for valuable discussions throughout this work. We thank Kim Schlegel for advive on cassava transformation and Andrea Ruckle, Laure David and Gabriel Deslandes for assistance in plant culture and harvesting. We thank Antoni Sánchez-Ferrer for advice with DSC measurements, and Anton Hochmuth, Daniel Breitenstein and Simona Rodighiero for help with the starch granule sizing and imaging.

Conflict of interest

The authors declare that they have no conflict of interests.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Figure S1. Generation of transgenic cassava expressing *Solanum tuberosum* GWD.
Figure S2. Plant height and storage root fresh weight of StGWD and StGWDm over-expression lines over three generations.
Figure S3. Generation of *MeSEX4* RNAi and *MeLSF2* RNAi transgenic cassava.
Figure S4. *MeSEX4* and *MeLSF2* expression in the leaves of transgenic cassava lines.
Figure S5. Rapid viscometric analysis of starch extracted from storage roots of wild-type cassava and transgenic lines over-expressing GWD.
Table S1. Plant height and storage root fresh weigh of *MeSEX4* and *MeLSF2* RNAi cassava lines
Table S2. Differential scanning calorimetry analysis of starch slurries from wild-type cassava and transgenic lines expressing StGWD.

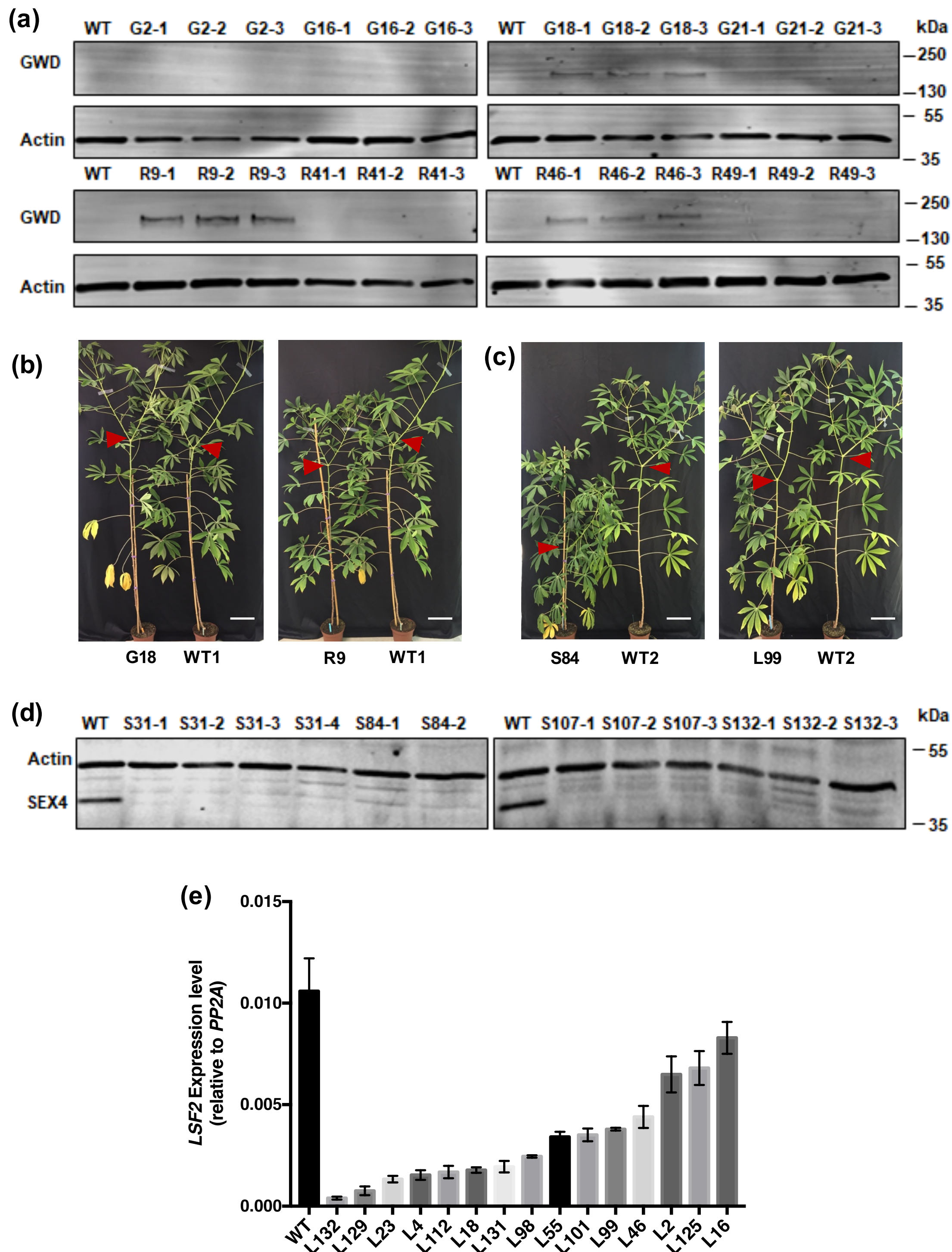


Figure 1. Analysis of gene expression in transgenic cassava lines.
(a) Immunoblot detection of GWD proteins (G, wild type *St*GWD protein; R, redox insentitive version) in soluble protein extracts from cassava storage root (5 µg protein loaded). Wild-type (WT) control and three technical replicates of pooled material (from 3 individuals of each transgenic line) were analyzed. Actin (loading control) was detected on the same membrane (b) Growth comparison of representative G- and R-lines (c) Growth comparison of representative *LSF2* RNAi lines (L-lines) and *SEX4* RNAi lines (S-lines). Red arrows indicate shoot branching. Bar = 15 cm. (d) Immunoblot detection of the *Me*SEX4 protein in soluble protein extracts (5 µg protein loaded). Wild-type control and 2-4 technical replicates of pooled material (from three individuals of each transgenic line) were analyzed. Actin was used as a loading control. (e) Expression of the *LSF2* gene in storage roots relative to the *PP2A* reference gene determined by RT-PCR. Mean values (± SE) of 3 replicate analyses of pooled material (from 3 individuals of each transgenic line).

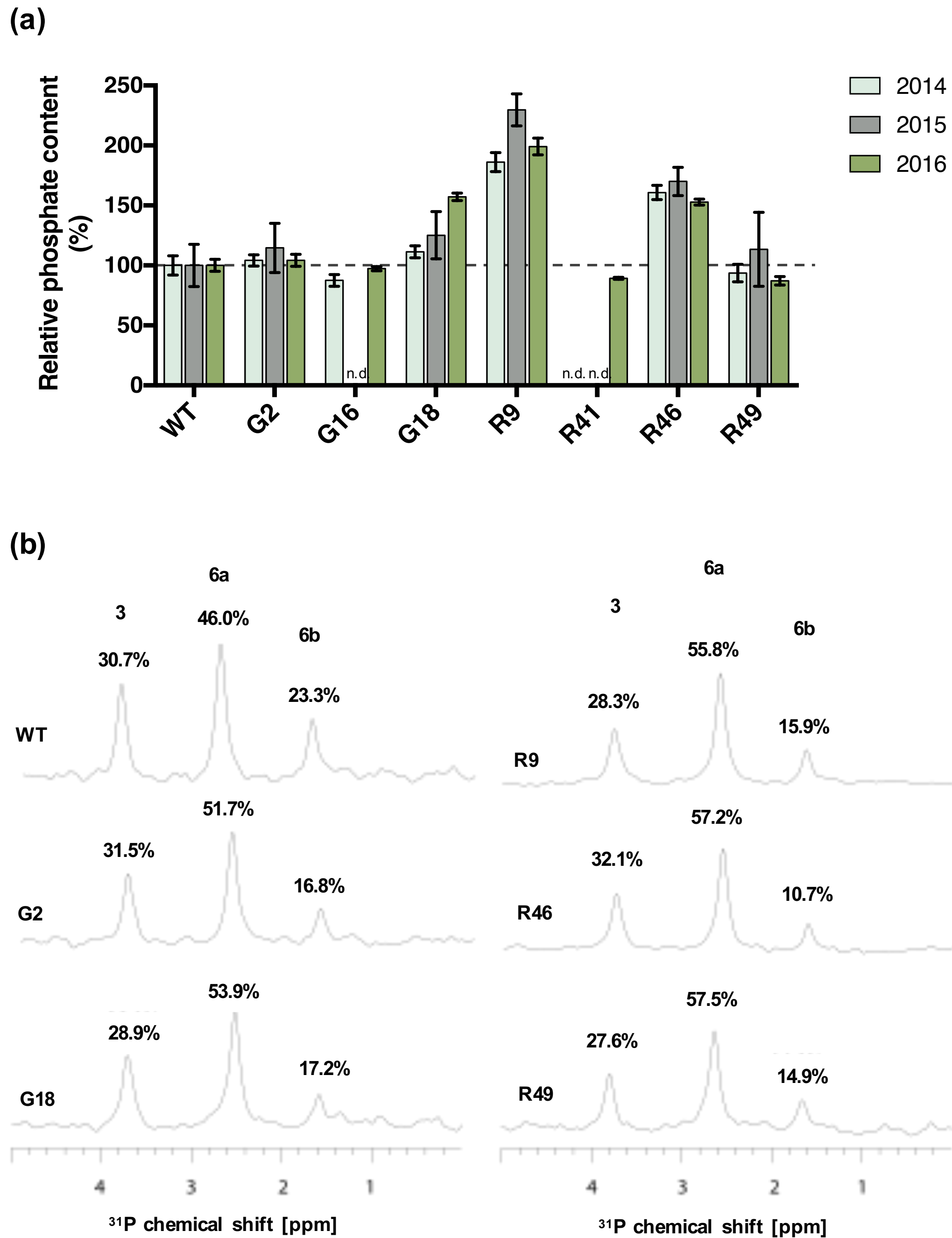


Figure 2. Total phosphate content and C3:C6 phosphate distribution of GWD overexpression lines.

(a) Starch was purified from storage roots of given lines in three consecutive harvests. Total starch bound phosphate was measured with the malachite green assay. The phosphate content relative to the wild type (WT) is shown. The mean phosphate content of WT was 0.62, 0.31, and 0.38 nmol P/µmol Glc in 2014, 2015 and 2016, respectively. Three technical replicates were measured for pooled starch derived from equal mounts of storage roots from three biological replicates of each line. n.d., not determined. (b) ^{31}P-NMR one-dimensional spectra of hydrolyzed starch was performed to detect C3-bound (3) and C6-bound (6a and 6b) phosphates. Peak areas are proportional to the relative amount of phosphate and are given as a percentage above each peak.

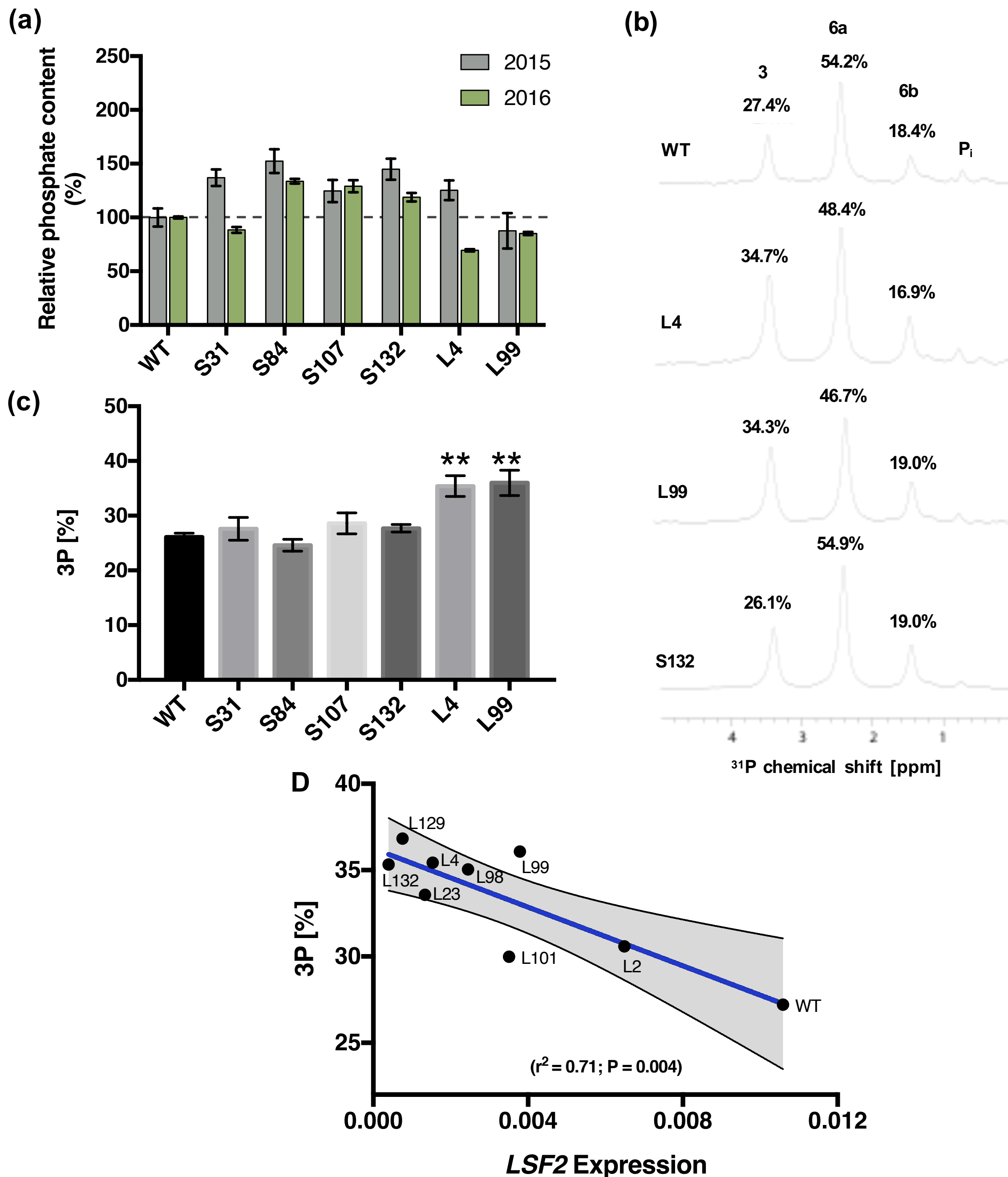


Figure 3. Total phosphate content and C3:C6 phosphate distribution of *MeSEX4* RNAi and *MeLSF2* RNAi cassava lines

(a) Starch was purified from storage root of representative S and L lines in 2015 and 2016 and dried. Total starch-bound phosphate was measured and is expressed as a percentage relative to the wild-type value for each year. Values are means ± SE of three technical replicates. (b) Example ^{31}P-NMR one-dimensional spectra of hydrolyzed starch, performed to detect C3-bound (3) and C6-bound (6a and 6b) phosphates. Peak areas are proportional to the relative amount of phosphate and are given as a percentage above each peak. Pooled samples from three biological replicates were used. (c) Percentage of C3-bound phosphate of selected lines, measured through ^{31}P-NMR, as above. Values are the means ± SE of three replicate samples. Student's t-tests were performed; ** : P < 0.01. (d) Correlation between the percentage C3-bound phosphate and *LSF2* gene expression in indicated L-lines. Blue line indicates the linear correlation, with 95% confidence intervals. Error bands (dash lines) and error area (gray) are shown.

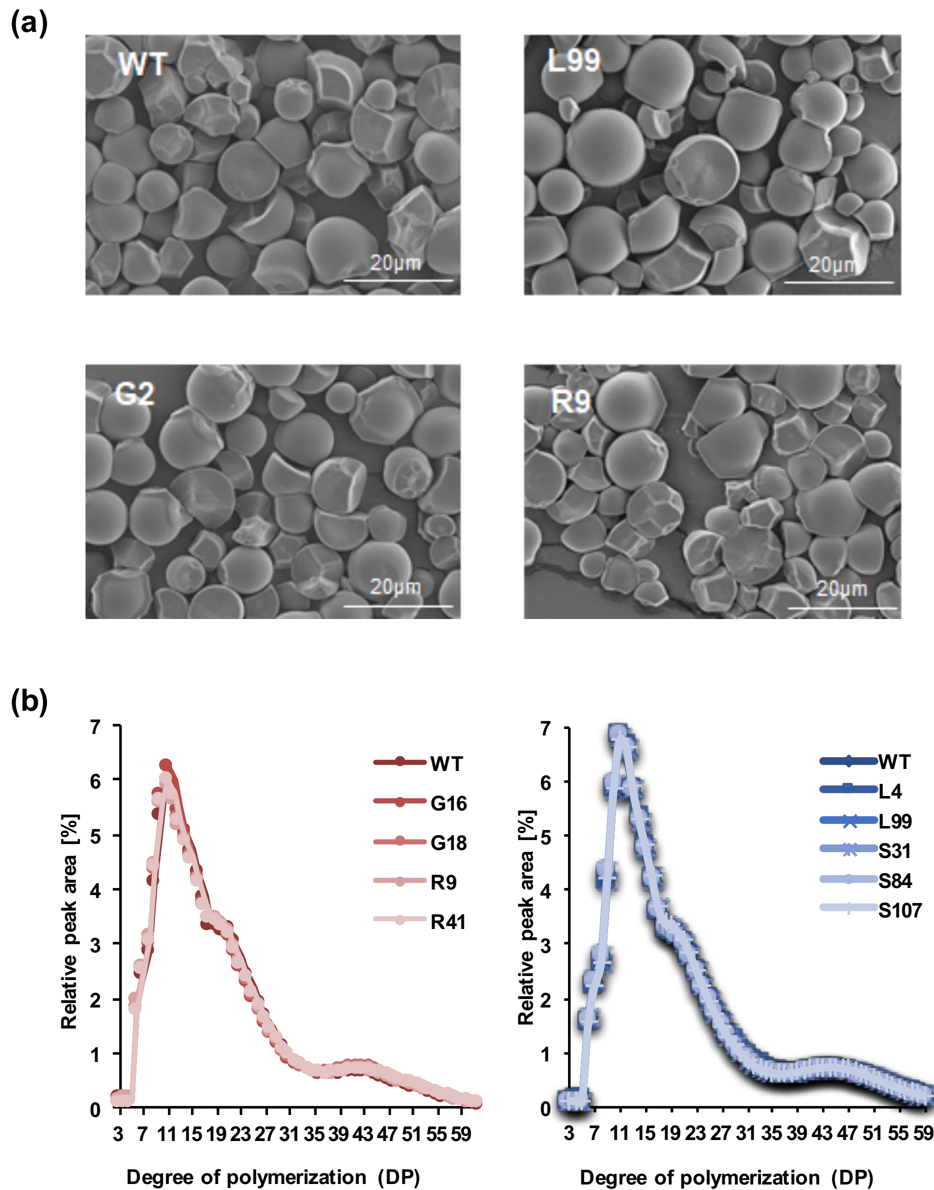


Figure 4. Starch structure in transgenic cassava roots

(a) Scanning electron micrographs of purified and dried root starch granules. Line names and scale bars are indicated. (b) Comparison of chain length distributions of starch in different transgenic lines. Left is the CLDs of G- and R-line starches from samples in 2016. Right is the CLDs of S and L line starches from samples in 2015. Values are means of 4 technical replicates performed on pooled starch samples form 3 biological replicate plants..

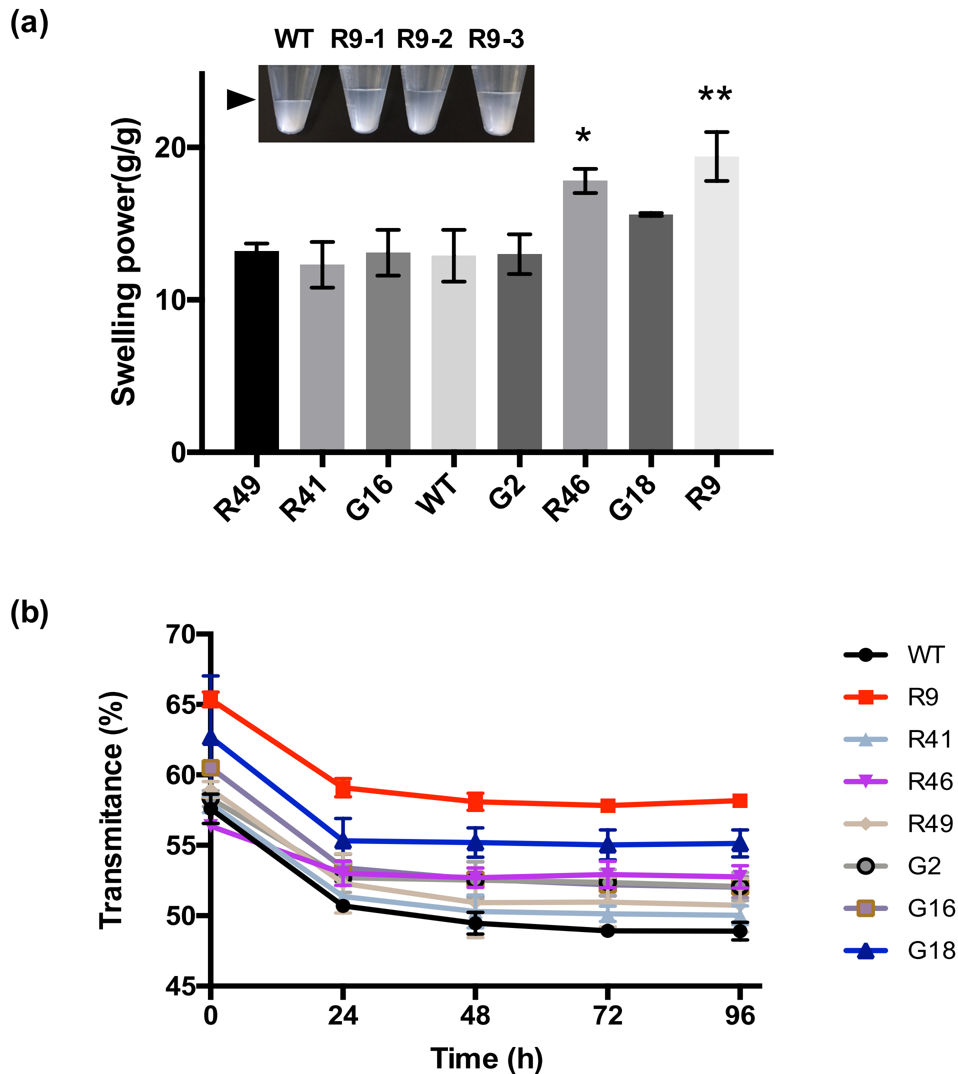


Figure 5. Swelling power and starch paste clarity of GWD over-expression lines

(a) Swelling power of starch from GWD over-expression lines. Lines R46 and R9 had significantly increased swelling (t-test; ∗: $P < 0.05$, ∗∗: $P < 0.01$). Inset: picture of R9 line with three replicates showing the level of swollen starch (arrowhead). Starches were purified from samples harvested in 2016. (b) Starch paste clarity was determined by light transmittance at 650nm of gelatinized starch stored at 4℃ for the given 24 h periods. Starches were purified from samples harvested in 2016. Values are the means ± SE of three technical replicates performed on pooled starch from 3 biological replicate plants.

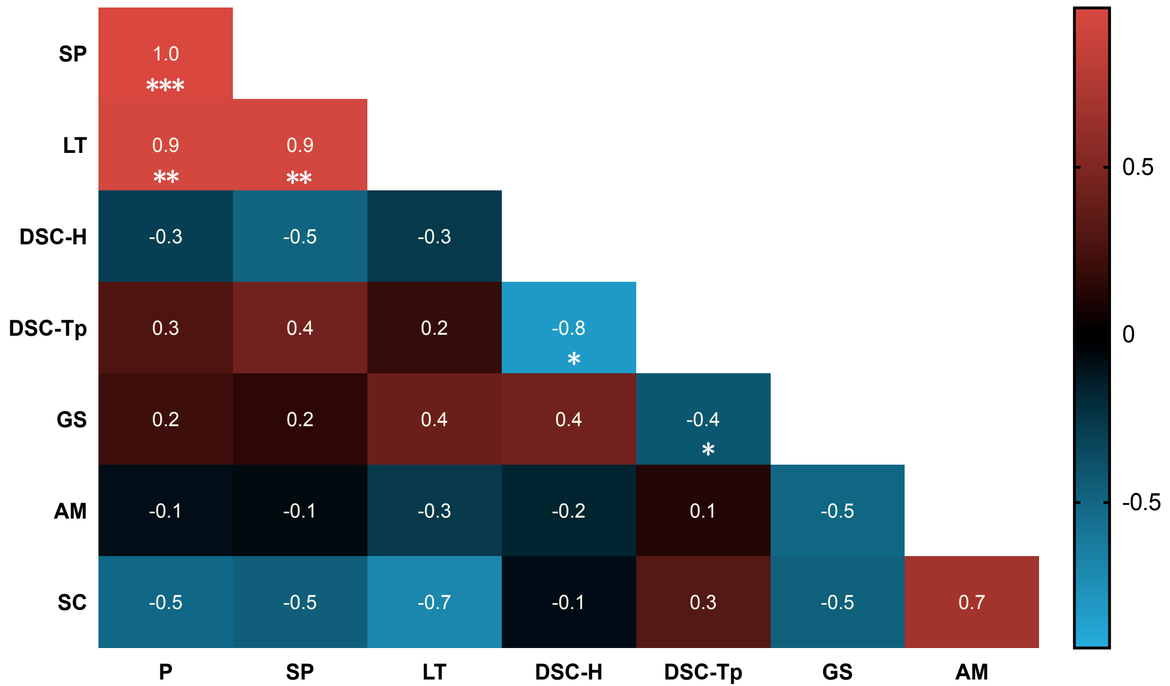


Figure 6. Heatmap displaying the Pearson correlation between starch characters and starch physico-chemical properties.

Each correlation value is indicated and the relevant P value is shown as significance (*: P < 0.05, **: P < 0.01, *** < 0.001). Red colors show positive correlations and blue colors show negative correlations. P, phosphate content; SP, swelling power; LT, light transmittance (after 96 h storage); DSC-H, DSC gelatinization enthalpy; DSC-Tp, DSC gelatinization peak temperature; GS, granule size; AM, amylose content; SC, starch content.

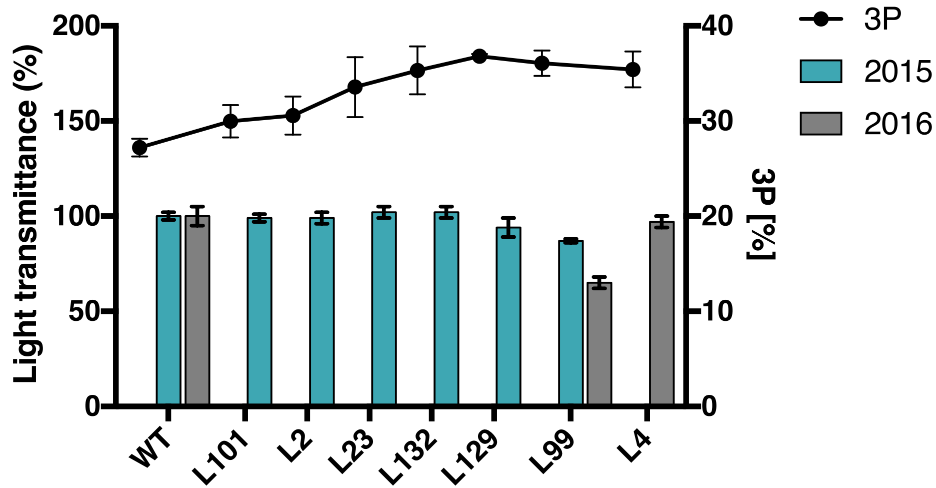


Figure 7. Starch paste clarity and C3-bound phosphate percentages of selected *MeLSF2* RNAi cassava lines

C3-bound phosphate (3P) percentage (right Y axis) was measured through ^{31}P-NMR (same values as presented in Figure 3d). The mean value (±SE) of three replicate analyses of pooled starch were included. Light transmittance of gelatinized starch (measured after 96 h – as in Figure 5) from root samples harvested in 2015 and 2016. Values are expressed relative to the wild-type (WT) transmittance (left Y axis) and are the means ± SE of three replicate analyses of starch pooled from three biological replicate plants.

Table 1. Starch compositional characters in transgenic cassava roots

Construct	Line	Starch content (mg.g-1 FW)	Amylose content[1] (%)	Starch granule size[1 2] Mode (μm)
StGWD &*StGWDm* OE	**WT**	269.5 (±9.85)	19.0 (±2.60)	12.4 (±1.38)
	G2	199.5 (±4.94)	16.9 (±0.42)	13.6 (±1.42)
	G16	239.5 (±12.41)	18.1 (±0.28)	13.6 (±1.43)
	G18	210.9 (±10.34)	18.8 (±0.92)	13.6 (±1.38)
	R9	215.4 (±5.30)	18.2 (±0.78)	12.4 (±1.37)
	R41	247.1 (±8.16)	19.0 (±0.44)	12.4 (±1.40)
MeSEX4 &*MeLSF2* RNAi	**WT**	277.8 (±8.65)	20.3 (±0.45)	12.4 (±1.38)
	S31	241.2 (±10.18)	17.9 (±0.65)*	10.3 (±1.37)
	S84	193.7 (±11.06)	18.9 (±0.68)	11.3 (±1.40)
	S107	211.4 (±7.93)	18.8 (±1.03)	11.3 (±1.42)
	L4	327.2 (±13.22)	22.1 (±0.73)	13.6 (±1.33)
	L99	166.5 (±5.63)	20.2 (±1.03)	10.3 (±1.32)

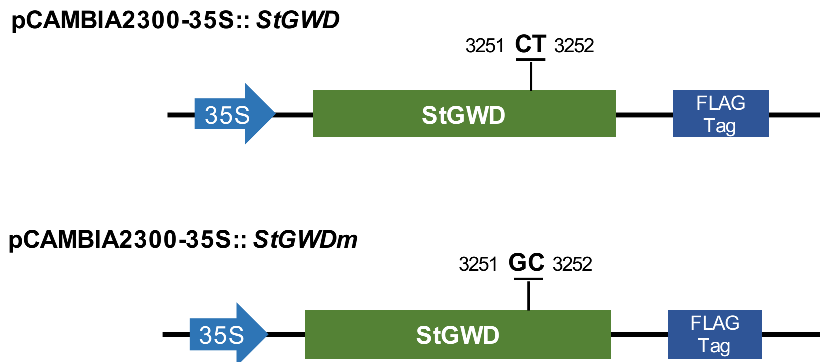
Mean values (± standard deviation of three technical replicates)

[1] Dunnett's multiple comparisons test at 95% CI (each line to WT) was performed for amylose content (*: $P < 0.05$) and starch granule size (no significant difference).

[2] The mean of granule size was calculated ranging between 3 to 30μm.

Wang et al., Supplementary data

(a)



(b)

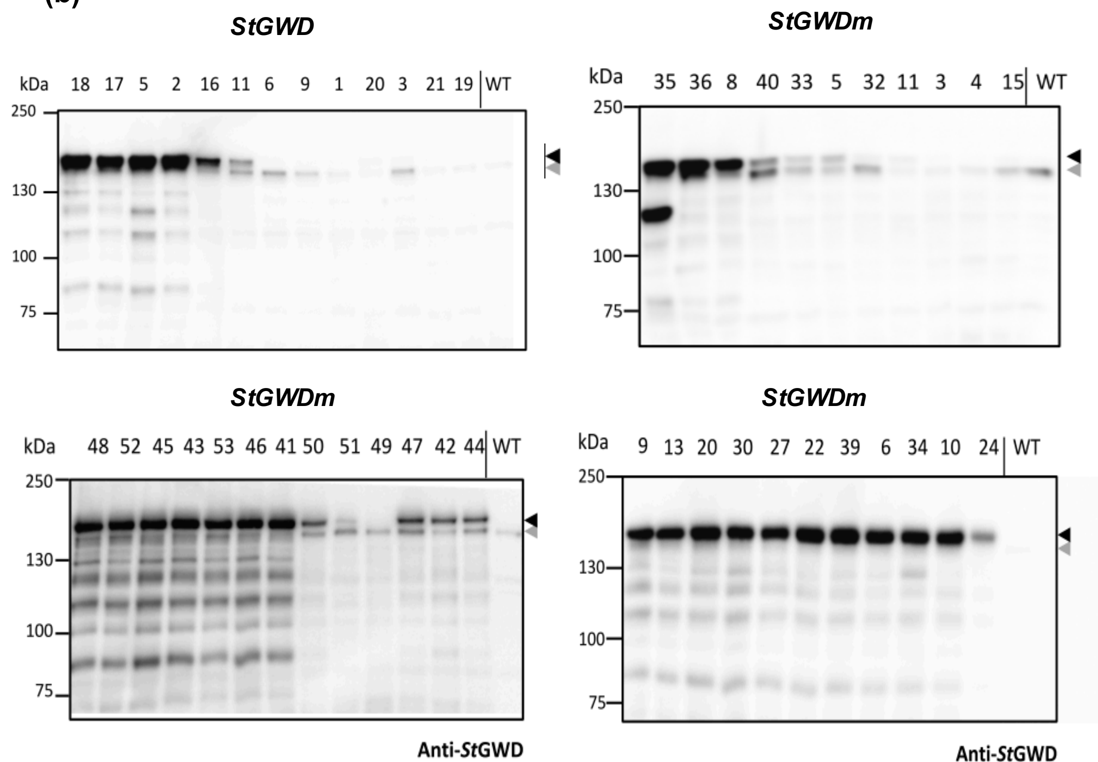


Figure S1. Generation of transgenic cassava expressing *Solanum tuberosum* GWD. (a) Two pCAMBIA2300 plasmid constructs harboring either the wild type (pCAMBIA2300::*StGWD*) or redox-insensitive (pCAMBIA2300::*StGWDm*) coding sequence of *S. tuberosum* GWD were placed under the control of 35S promoter. Nucleotide modifications at location 3251 and 3252 of *StGWDm* are indicated. (b) Thirty micrograms of total leaf protein from individual lines were subjected to immunoblot analysis using anti-StGWD antibody. The black arrow indicates the *St*GWD protein. The grey arrow represents endogenous *Me*GWD protein.

Wang et al., Supplementary data

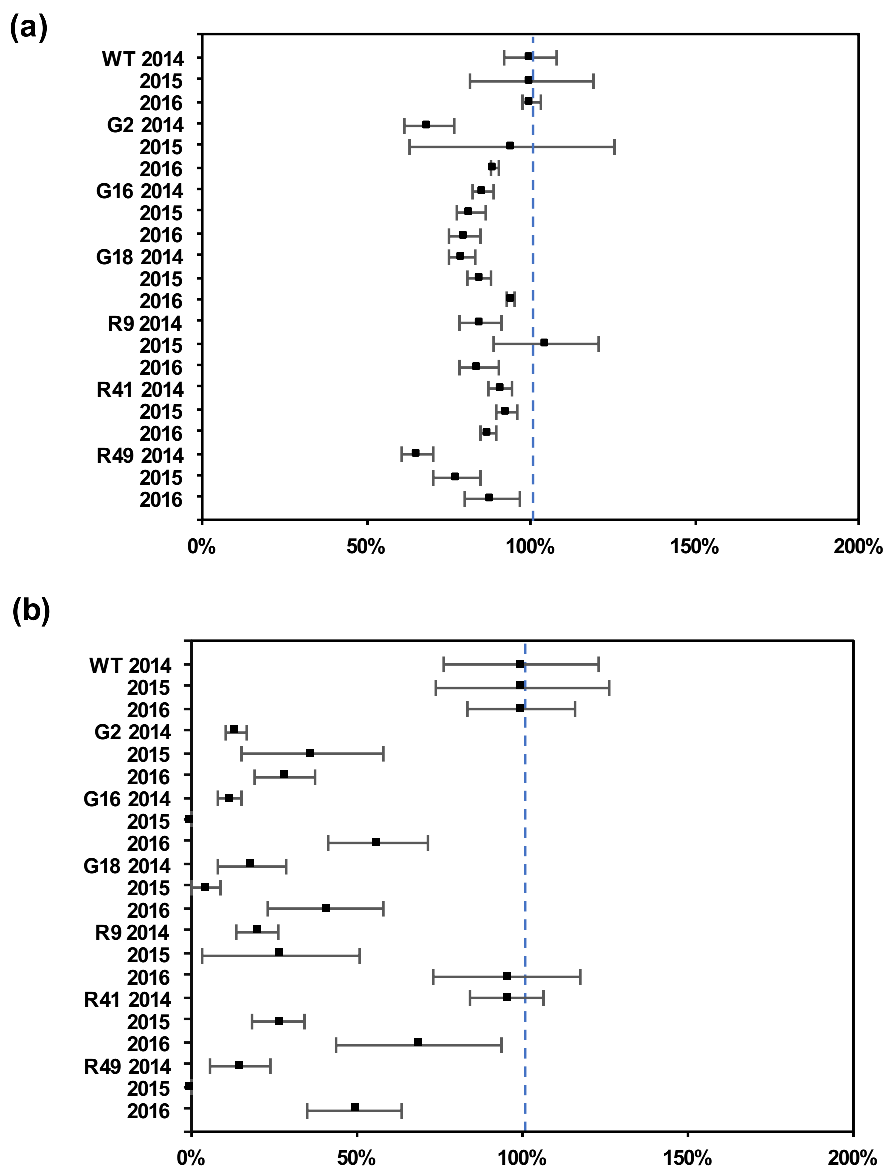
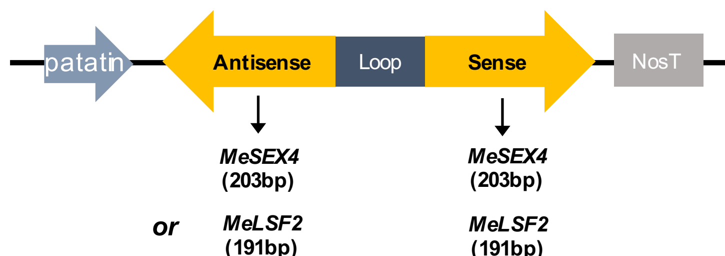


Figure S2. Plant height and storage root fresh weight of *StGWD* and *StGWDm* over-expression lines over three generations.

(a) Forest plots were generated for plant height and (b) storage root fresh weight of cassava in years 2014, 2015 and 2016. Points represent the mean value for each line (± SE, 2014 n=4-9, 2015, n=3-5, 2016 n=2-5). The wild type in each year is calculated to 100% (blue dashed line). The value of each transgenic line is given as a year-specific relative percentage. Plant height was recorded after 7 months growth. Storage roots were measured after 10 months growth.

Wang et al., Supplementary data

(a) pCAMBIA1301-patatin:: *MeSEX4* hairpin
(or MeLSF2)



(b) cassava4.1_009735m 203bp CDS for *MeSEX4* hairpin construct

GGAGGACAAGGGAAAGTCTGAGATATACAGTCATAATATGACAGAAGCTATGGGTGC
TGTTTTGACCTATAGGCATGAACTAGGAATGAACTTCAACTTCATTTGTCCAGATTTGA
TTGTAGGATCATGCCTACAGACTCCTGAAGATGTTGACAAGCTTCGAGAAATAGGAGT
GAAAACAATATTCTGCTTGCAACAGGACC

(c) cassava4.1_013314m 191bp CDS for *MeLSF2* hairpin construct

TCTCATACCCGAATGAAACTGAAGTGTTCTTCACAAGAAATAAATCTACATGGAAATT
CTTGCTTTCGAAGGATTGCTTCAAAATGGGTAGGATCAATTGTAAGCTATCAGATAGT
GGAATCGAGAAAAAACCCACTGGAAAAGATGTGTCATTGAGCTCAACGAACAGGATG
GAAGAGTACAATACAGCC

Figure S3. Generation of *MeSEX4* RNAi and *MeLSF2* RNAi transgenic cassava. (a) Construct design for *MeSEX4* and *MeLSF2* RNAi cassava lines. The hairpin construct contains an antisense and a sense region (203bp of coding sequence of the *MeSEX4* gene or 191bp of the coding sequence of *MeLSF2*) with a loop in the middle. The hairpin sequence was sub-cloned to vector pCAMBIA1301 and expression was driven by the *S. tuberosum* class 1 patatin B33 promoter. (b) The sequence targeting *MeSEX4* gene. (c) The sequence targeting *MeLSF2* gene.

Wang et al., Supplementary data

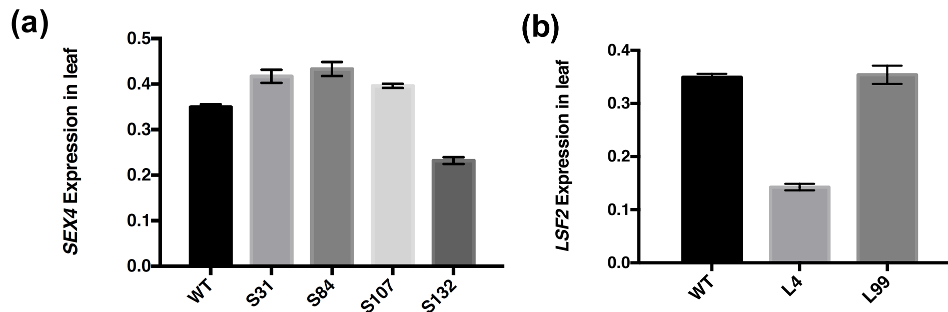


Figure S4. *MeSEX4* and *MeLSF2* expression in the leaves of transgenic cassava lines.
(a) RT-PCR was performed using RNA extracted from *in-vitro* grown plantlet leaves of the wild type (WT) and selected *MeSEX4* RNAi-lines. Mean *SEX4* expression levels (± SE, n=3) are given relative to the housekeeping gene *PP2A*. (b) *LSF2* expression levels in the wild type (WT) and selected *MeLSF2* RNAi-lines are given relative to *PP2A*, as described in (a).

Wang et al., Supplementary data

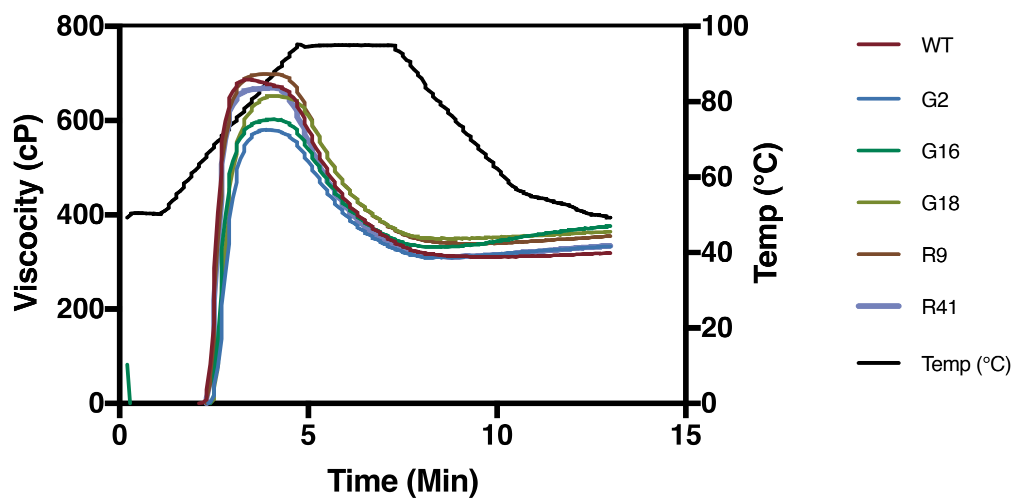


Figure S5. Rapid viscometric analysis of starch extracted from storage roots of wild-type cassava and transgenic lines over-expressing GWD.
Viscosity profiles of starch/water slurries during heating and cooling in a rapid viscoanalyser. Each profile is the mean of three replicate measurement's made for starch pooled from three biological replicates for each line.

Wang et al., Supplementary data

Table S1. Plant height and storage root fresh weigh of *MeSEX4* and *MeLSF2* RNAi cassava lines
The height of 6-month-old plants and the storage root fresh weight of 8-month-old plants of each of the given lines was measured.

Construct	Line	Plant height (cm)	SR fresh weight (g)
MeSEX4	**WT**	150 (±4)	83.77 (±18.5)
RNAi	**S31**	134.3 (±16.9)	30.2 (±15.2)
	S84	101.6 (±3.7)	10.3 (±1.9)
	S107	138.6 (±15.8)	22.1 (±21.3)
	S132	100.5 (±10.4)	27.7 (±17.1)
MeLSF2	**WT**	150 (±4)	83.77 (±18.5)
RNAi	**L4**	126.6 (±8.1)	71.0 (±39.6)
	L99	133.3 (±6.4)	10.8 (±7.04)

Table S2. Differential scanning calorimetry analysis of starch slurries from wild-type cassava and transgenic lines expressing *St*GWD.
For the DSC assay, starch from three biological replicate plants was pooled, and three replicates samples were taken. Peak temperature and enthalpy were determined using STAR software.

DSC Samples	WT	R9	R41	R46	R49	G2	G16	G18
Peak Temp. (℃)	61.7 ± 0.07	62.8 ± 0.13	61.2 ± 0.21	59.8 ± 0.14	61.4 ± 0.07	60.6 ± 0.17	61.0 ± 0.17	59.8 ± 0.05
Enthalpy (Jg-1)	17.4 ± 4.4	15.1 ± 1.3	18.6 ± 1.9	17.6 ± 2.7	15.4 ± 0.4	18.7 ± 1.9	18.2 ± 3.5	19.2 ± 1.6

References

Alcázar-Alay SC, Meireles MAA (2015) Physicochemical properties, modifications and applications of starches from different botanical sources. Food Sci Technol 35:215–236. doi: 10.1590/1678-457X.6749

Allem AC, Genéticos R (2002) The Origins and Taxonomy of Cassava. Cassava Biol Prod Util 1–16. doi: 10.1080/21622965.2014.897905

Alves a. a. C (2002) Cassava botany and physiology. Cassava Biol Prod Util 67–89. doi: 10.1079/9780851995243.0000

Balagopalan C (2002) Cassava Utilization in Food, Feed and Industry. Cassava Biol Prod Util 301–318. doi: 10.1079/9780851995243.0000

Baunsgaard L, Lütken H, Mikkelsen R, et al (2005) A novel isoform of glucan, water dikinase phosphorylates pre-phosphorylated α-glucans and is involved in starch degradation in Arabidopsis. Plant J 41:595–605. doi: 10.1111/j.1365-313X.2004.02322.x

Blennow A, Bay-Smidt AM, Wischmann B, et al (1998) The degree of starch phosphorylation is related to the chain length distribution of the neutral and the phosphorylated chains of amylopectin. Carbohydr Res 307:45–54. doi: 10.1016/S0008-6215(98)00015-9

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254. doi: 10.1016/0003-2697(76)90527-3

Bull SE, Owiti JA, Niklaus M, et al (2009) Agrobacterium-mediated transformation of friable embryogenic calli and regeneration of transgenic cassava. Nat Protoc 4:1845–1854. doi: 10.1038/nprot.2009.208

Carciofi M, Shaik SS, Jensen SL, et al (2011) Hyperphosphorylation of cereal starch. J Cereal Sci 54:339–346. doi: 10.1016/j.jcs.2011.06.013

Carpenter MA, Joyce NI, Genet RA, et al (2015) Starch phosphorylation in potato tubers is influenced by allelic variation in the genes encoding glucan water dikinase, starch branching enzymes I and II, and starch synthase III. Front Plant Sci 6:143. doi: 10.3389/fpls.2015.00143

Caspar T, Lin TP, Kakefuda G, et al (1991) Mutants of Arabidopsis with altered regulation of starch degradation. Plant Physiol 95:1181–1188. doi: 10.1104/pp.95.4.1181

Chen Y, Sun X, Zhou X, et al (2017) Highly phosphorylated functionalized rice starch produced by transgenic rice expressing the potato GWD1 gene. Sci Rep. doi: 10.1038/s41598-017-03637-5

Craig SAS, Maningat CC, Seib PA, Hoseney RC (1989) Starch paste clarity. Cereal Chem 66:173–182.

Edner C, Li J, Albrecht T, et al (2007) Glucan, Water Dikinase Activity Stimulates Breakdown of Starch Granules by Plastidial β-Amylases. Plant Physiol 145:17–28. doi: 10.1104/pp.107.104224

El-Sharkawy MA (2004) Cassava biology and physiology. Plant Mol Biol 56:481–501. doi: 10.1007/s11103-005-2270-7

Gentry MS, Dowen RH, Worby CA, et al (2007) The phosphatase laforin crosses evolutionary boundaries and links carbohydrate metabolism to neuronal disease. J Cell Biol 178:477–488. doi: 10.1083/jcb.200704094

Hansen PI, Spraul M, Dvortsak P, et al (2009) Starch phosphorylation-maltosidic restrains upon 3'- and 6'-phosphorylation investigated by chemical synthesis, molecular dynamics and NMR spectroscopy. Biopolymers 91:179–93. doi: 10.1002/bip.21111

Hejazi M, Fettke J, Paris O, Steup M (2009) The Two Plastidial Starch-Related Dikinases Sequentially Phosphorylate Glucosyl Residues at the Surface of Both the A- and B-Type Allomorphs of Crystallized Maltodextrins But the Mode of Action Differs. PLANT Physiol 150:962–976. doi: 10.1104/pp.109.138750

Hejazi M, Fettke J, Kotting O, et al (2010) The Laforin-Like Dual-Specificity Phosphatase SEX4 from Arabidopsis Hydrolyzes Both C6- and C3-Phosphate Esters Introduced by Starch-Related Dikinases and Thereby Affects Phase Transition of -Glucans. PLANT Physiol 152:711–722. doi: 10.1104/pp.109.149914

Jobling S (2004) Improving starch for food and industrial applications. Curr Opin Plant Biol 7:210–218. doi: 10.1016/j.pbi.2003.12.001

Karlström A, Calle F, Salazar S, et al (2016) Biological Implications in Cassava for the Production of Amylose-Free Starch: Impact on Root Yield and Related Traits. Front Plant Sci 7:604. doi: 10.3389/fpls.2016.00604

Kötting O (2005) Identification of a Novel Enzyme Required for Starch Metabolism in Arabidopsis Leaves. The Phosphoglucan, Water Dikinase. PLANT Physiol 137:242–252. doi: 10.1104/pp.104.055954

Kötting O, Santelia D, Edner C, et al (2009) STARCH-EXCESS4 is a laforin-like Phosphoglucan phosphatase required for starch degradation in Arabidopsis thaliana. Plant Cell 21:334–46. doi: 10.1105/tpc.108.064360

Kusumayanti H, Handayani NA, Santosa H (2015) Swelling Power and Water Solubility of Cassava and Sweet Potatoes Flour. Procedia Environ Sci 23:164–167. doi: 10.1016/j.proenv.2015.01.025

Lanahan MB, Basu SS. 2005. Modified starch, uses, methods for production thereof. International Patent Application no. WO/2005/002359

Lorberth R, Ritte G, Willmitzer L, Kossmann J (1998) Inhibition of a starch-granule-bound protein leads to modified starch and repression of cold sweetening. Nat Biotechnol 16:473–477. doi: 10.1038/nbt0598-473

Maurer T, Kalbitzer HR (1996) Indirect referencing of31P and19F NMR spectra. J Magn Reson - Ser B 113:177–178. doi: 10.1006/jmrb.1996.0172

Mikkelsen R, Mutenda KE, Mant A, et al (2005) Alpha-glucan, water dikinase (GWD): a plastidic enzyme with redox-regulated and coordinated catalytic activity and binding affinity. Proc Natl Acad Sci U S A 102:1785–1790.

Pfister B, Zeeman SC (2016) Formation of starch in plant cells. Cell Mol Life Sci 73:2781–2807. doi: 10.1007/s00018-016-2250-x

Nashilevitz S, Melamed-Bessudo C, Aharoni A, et al (2009) The legwd mutant uncovers the role of starch phosphorylation in pollen development and germination in tomato. Plant J 57:1–13. doi: 10.1111/j.1365-313X.2008.03664.x

Naumkina EM, Bolyakina YP, Romanov GA (2007) Organ-specificity and inducibility of patatin class I promoter from potato in transgenic arabidopsis plants. Russ J Plant Physiol 54:350–359. doi: 10.1134/S1021443707030090

Niittylä T, Comparot-Moss S, Lue WL, et al (2006) Similar protein phosphatases control starch metabolism in plants and glycogen metabolism in mammals. J Biol Chem 281:11815–11818. doi: 10.1074/jbc.M600519200

Niklaus M, Gruissem W, Vanderschuren H (2011) Efficient transformation and regeneration of transgenic cassava using the neomycin phosphotransferase gene as aminoglycoside resistance marker gene. GM Crops 2:193–200. doi: 10.4161/gmcr.2.3.18866

Puounti-Kaerlas J (2002) Cassava Biotechnology. Genet Eng Rev 15:329–364. doi: 10.1079/9780851995243.0000

Raphael M, Yona B, Stephen K, et al (2011) Amylopectin molecular structure and functional properties of starch from three Ugandan cassava varieties. J Plant Breed Crop Sci 3:195–202.

Ritte G, Heydenreich M, Mahlow S, et al (2006) Phosphorylation of C6- and C3-positions of glucosyl residues in starch is catalysed by distinct dikinases. FEBS Lett 580:4872–4876. doi: 10.1016/j.febslet.2006.07.085

Santelia D, Kötting O, Seung D, et al (2011) The phosphoglucan phosphatase Like Sex Four2 dephosphorylates starch at the C3-position in Arabidopsis. Plant Cell 23:4096–4111. doi: 10.1105/tpc.111.092155

Santelia D, Zeeman SC (2011) Progress in Arabidopsis starch research and potential biotechnological applications. Curr Opin Biotechnol 22:271–280. doi: 10.1016/j.copbio.2010.11.014

Schewe G, Knies P, Amati SF, Lorz H, Becker D, et al. 2003 Monocotyledon plant cells and plants which synthesize modified starch. US Patent application no. US 2003/0110534 A1

Silver DM, Kötting O, Moorhead GBG (2014) Phosphoglucan phosphatase function sheds light on starch degradation. Trends Plant Sci 19:471–478. doi: 10.1016/j.tplants.2014.01.008

Silver DM, Silva LP, Issakidis-Bourguet E, et al (2013) Insight into the redox regulation of the phosphoglucan phosphatase SEX4 involved in starch degradation. FEBS J. 280:538–548.

Singh N, Singh J, Kaur L, et al (2003) Morphological, thermal and rheological properties of starches from different botanical sources. Food Chem. 81:219–231.

Smith AM, Zeeman SC (2006) Quantification of starch in plant tissues. Nat Protoc 1:1342–1345. doi: 10.1038/nprot.2006.232

Stitt M, Zeeman SC (2012) Starch turnover: Pathways, regulation and role in growth. Curr Opin Plant Biol 15:282–292. doi: 10.1016/j.pbi.2012.03.016

Streb S, Delatte T, Umhang M, et al (2008) Starch Granule Biosynthesis in Arabidopsis Is Abolished by Removal of All Debranching Enzymes but Restored by the Subsequent

Removal of an Endoamylase. PLANT CELL ONLINE 20:3448–3466. doi: 10.1105/tpc.108.063487

Taylor N, Chavarriaga P, Raemakers K, et al (2004) Development and application of transgenic technologies in cassava. Plant Mol. Biol. 56:671–688.

Vasconcelos LM, Brito AC, Carmo CD, Oliveira EJ (2016) Polymorphism of starch pathway genes in cassava. 1:1–15.

Waterschoot J, Gomand S V., Willebrords JK, et al (2014) Pasting properties of blends of potato, rice and maize starches. Food Hydrocoll 41:298–308. doi: 10.1016/j.foodhyd.2014.04.033

Weise SE, Aung K, Jarou ZJ, et al (2012) Engineering starch accumulation by manipulation of phosphate metabolism of starch. Plant Biotechnol J 10:545–554. doi: 10.1111/j.1467-7652.2012.00684.x

Yu TS, Kofler H, Häusler RE, et al (2001) The Arabidopsis sex1 mutant is defective in the R1 protein, a general regulator of starch degradation in plants, and not in the chloroplast hexose transporter. Plant Cell 13:1907–1918. doi: 10.1105/TPC.010091

Zeeman SC, Kossmann J, Smith AM (2010) Starch: its metabolism, evolution, and biotechnological modification in plants. Annu Rev plant … 61:209–234. doi: 10.1146/annurev-arplant-042809-112301

Zeeman SC, Northrop F, Smith AM, Ap Rees T (1998) A starch-accumulating mutant of Arabidopsis thaliana deficient in a chloroplastic starch-hydrolysing enzyme. Plant J 15:357–365. doi: 10.1046/j.1365-313X.1998.00213.x

Zhu T, Jackson DS, Wehling RL, Geera B (2008) Comparison of amylose determination methods and the development of a dual wavelength iodine binding technique. Cereal Chem 85:51–58. doi: 10.1094/CCHEM-85-1-0051

Chapter 3

Improving starch diversity in the orphan crop- tef (Eragrostis tef)

Wuyan Wang[1], Matthias Hirsch-Hoffmann[1], Niklaus Zemp[2], Gina M. Cannarozzi[3], Regula Bloesch[3], Zerihun Tadele[3] and Samuel C. Zeeman[1]*

[1]Institute of Molecular Plant Biology, Department of Biology, ETH Zurich, CH-8092 Zürich, Switzerland

[2]Genetic Diversity Centre (GDC), ETH Zurich, CH-8092 Zürich, Switzerland

[3] Institute of Plant Sciences, University of Bern, CH-3013 Bern, Switzerland

***Corresponding Author:**
Samuel C. Zeeman
Department of Biology, ETH Zurich
Universitaetstrasse 2, 8092 Zurich, Switzerland
Tel.: +41 44 632 8275
Fax.: +41 44 632 1664
E-mail: szeeman@ethz.ch

Authorship:
In this work, I performed most of the experiments except for the database management platform, which was generated by Matthias Hirsch-Hoffmann, variant calling, which was conducted by Niklaus Zemp, genome annotation, which was done together with Gina M. Cannarozzi and LICOR gel running, which was done together Regula Bloesch from Zerihun Tadele's group.

3.1 Abstract

Tef (*Eragrostis tef* (Zucc.) Trottter) is the most importance cereal in Ethiopia. It is highly appreciated for its nutritional properties and has the advantage of being naturally gluten-free. Yet, it still receives very little attention from the global research network. Starch comprises more than 70% of tef grain and is the primary source of calories served in Ethiopian diets. Tef is also used in other ways in the food and beverage industry. Unlike rice, maize, barley and other major crops, the starch metabolism in endosperm of tef has not been researched and the existing tef cultivars lack diversity in starch traits. The aim of this part of my research was to investigate the starch metabolism and to improve tef by diversifying its starch properties. I first identified tef genes involved in starch synthesis in the endosperm using comparative genomic analysis. I probed the tetraploid tag genome with sequences from the closely-related diploid model species, rice (*Oryza sativa*) and sorghum (*Sorghum bicolor*). I then confirmed that the targeted genes were expressed in developing tef seeds through RNA-sequencing analysis. Based on the conventional reverse genetics approach Targeting Induced Local Lesions IN Genomes (TILLING), I established a new TILLING-by-sequencing method, using high throughput sequencing and bioinformatics analysis. I used this method with the aim of generating tef plants with mutations in identified starch genes. A database containing mutations for each starch-related gene has been created. It will now be possible in the longer-term to conduct the screening of mutants, followed by the cultivation and crossing of plants in order to achieve useful null-mutant plants. I expect in the future to see new tef varieties with improved starch properties to be brought into breeding programs for the benefit of Ethiopian farmers and use in both food and non-food industries.

3.2 Introduction

Starch consists of amylopectin (typically 75%) and amylose (typically 25%), which together form semi-crystalline, insoluble granules with an internal lamellar structure. Starch is synthesized exclusively in plants and used for long-term storage of carbohydrate in non-photosynthetic tissues, such as the seed endosperms of the most important cereals (rice, maize, wheat, barley, sorghum, tef and others), and the tubers, roots or stems of other crops. Starch is the primary source of calories in the human diet and also serves as energy source for numerous important food and non-food industrial applications (Moorthy 2002; Jobling 2004; Svihus et al. 2005; Zeeman et al. 2010).

Biochemical and genetic analyses of mutants of different species have resulted in the identification of a significant number of enzymes and enzyme isoforms involved in starch biosynthesis. These findings have resulted in a model for starch synthesis in which the various enzymes have defines roles (Martin and Smith 1995; James et al. 2003; Tomlinson and Denyer 2003; Jeon et al. 2010; Zeeman et al. 2010; Tetlow 2011). Briefly, amylopectin and amylose synthesis both start from the substrate ADPglucose, produced by ADPglucose pyrophosphorylase (AGPase). Amylopectin chains are generated by soluble starch synthases (SSs) and each starch isoform (SSI, SSII, SSIII, and SSIV) has a distinct role to play. Amylopectin branching is controlled by the actions of starch branching enzymes (SBEs that are classified as BEI and BEII in higher plants) and debranching enzymes (DBEs, ISA and LDA in plants). Amylose is synthesized by granule-bound starch synthase (GBSS) which becomes encapsulated within the growing starch granule.

As for starch degradation, this pathway is best understood in leaves, where the transitory starch is degraded at night by suite of complementary enzymes. First, the surface of the starch granules is solubilized through glucan phosphorylation by two dikinases, namely Glucan, Water Dikinase (GWD) and Phosphoglucan, Water Dikinase (PWD) (Lorberth et al. 1998; Yu et al. 2001; Kötting et al. 2005; Baunsgaard et al. 2005). GWD and PWD both transfer the β-phosphate group of ATP to amylopectin glucosyl residues, phosphorylating the C6 and C3 positions respectively. This allows hydrolases (BAMs, AMY3, DBEs, ISA and LDA) to access to the glucan chains and hydrolyse them to maltose and glucose (Smith et al. 2005; Zeeman et al. 2010). Concurrently, the glucans are dephsphorylated again by the phosphoglucan phosphatses, Starch EXcess 4 (SEX4) and Like-SEX Four 2 (LSF2). SEX4 releases phosphate from both the C6 and C3 positions of amylopectin, with the preference for C6. In contrast, LSF2 is specific for C3-bound phosphate (Gentry et al. 2007; Kötting et al. 2009; Santelia et al. 2011; Silver et al. 2014).

Mutants lacking or having reduced activities of key starch metabolic enzymes have been generated in rice, wheat, maize, barley, potato and other crops (Burton et al. 2002; Fujita et al. 2006; Regina et al. 2006; Muth et al. 2008; Lin et al. 2012). The resulting characteristics and functional properties of the starches from these different crops enable them to be utilized for a

range of purposes (Lee et al. 2001; Sun et al. 2017; Eliasson 2004). Amylose content, for example, is one of the most important factors influencing starch quality. No or low amylose starch lines (also called *waxy*, after the appearance of amylose free maize kernals), have been generated through null or partially reduced GBSS activity in wheat, rice, potato and many other crops (Sano 1984; Hovenkamp-Hermelink et al. 1987; Nakamura et al. 1995). The *waxy* starch lines have altered on pasting properties, swelling power, solubility, and gel stability (Lee et al. 2001; Yu et al. 2009; Zhou et al. 2015). High amylose starch, for instance (generated by the silencing of starch branching enzymes and consequently resulting in the suppression of amylopectin synthesis) has different physicochemical properties (Sun et al. 2017). In addition, high amylose starch is used as a resistant starch (RS) product which benefits us in the diet (Sun et al. 2017). Amylopectin structure is also an important determinant of starch properties and can be affected independently of or together with changes in amylose content. For example, in the rice sbe1 mutant, endosperm starch has an altered amylopectin fine structure, a lower onset concentration for urea gelatinization as well as a lower onset temperature for thermo-gelatinization (Satoh et al. 2003). In addition to amylose content and amylopectin structure, other starch characteristics (e.g. phosphate content, granule size) have also been modified through modification of starch related enzymes in major crops (Burton et al. 2002 ; Carciofi et al. 2011; Xu et al. 2017; Jaiswal et al. 2014).

Tef (*Eragrostis tef* (Zucc.) Trottter), an allotetraploid species, is not considered as a major crop globally, but is instead classified as an orphan crop. It is the most important cereal in Ethiopia, where it is cultivated by over 6 million small-hold farmers and constitutes the staple food for about 50 million people. Tef is a gluten-free crop, rich in essential amino acids (alanine, methione, threonine and tyrosine) and mineral content (iron and calcium), compared with major cereals like wheat, rice, barley and millets. Tef is also relatively tolerant to harsh climatic conditions including both low and high moisture stress (Assefa et al. 2017). Due to its wide use as a staple and preferred food source in Ethiopia, as well as its increasing use as a naturally gluten-free material for processes such as malting and brewing (Gebremariam et al. 2012), research into tef starch properties could be valuable for optimizing the crop and for its downstream applications. The starch physico-chemical and functional properties of tef starch, including aspects such as paste clarity, gel texture, retrogradation, water absorption, water solubility indexes and others, were evaluated and compared with other starches (Bultosa and Taylor 2004; Abebe et al. 2015). Relevant techniques, such as steeping additives, have been shown to affect the quality of isolated tef starch (Nyakabau et al. 2013). In vitro starch digestibility of the tef bread (*injera*) and porridge was investigated and the estimated glycemic index showed that tef contains the more resistant and slowly digestible starch type than crops such as buckwheat, quinoa and wheat (Wolter et al. 2013; Shumoy and Raes 2017). However, there is very little research at the molecular level to support tef crop improvement, either for its starch properties or other agronomic aspects.

Among the 35 varieties most used by farmers, none has been reported to have any diversity in starch properties. Thus, to increase the understanding of starch metabolism in tef specifically,

and to enrich the crop for its starch diversity, we aimed to identify and alter the activity of single or multiple starch biosynthetic enzymes. This can be achieved by applying a Targeting Induced Local Lesions IN Genomes TILLING strategy (Tian et al. 2009). TILLING, is a dominant reverse genetic approach, based on traditional chemical mutagenesis followed by high throughput screening for point mutations. It has been developed for and widely applied in crop species (Till et al. 2004; Slade et al. 2005; Xin et al. 2008; Uauy et al. 2009; Minoia et al. 2010). In traditional TILLING, enzymes like S1 nuclease and T4 endonuclease VII have been used for mismatch-specific cleavage to detect heterozygous polymorphisms. CEL1 (belonging to S1 nuclease family) is a particularly efficient mismatch endonuclease commonly used to perform a single-strand cut on the 3′-side of a mismatch site. CEL1 was reported as a plant-specific extracellular glycoprotein from celery, alfalfa sprout, asparagus, and tomato (Oleykowski et al. 1998; Colbert 2001). TILLING by CEL1 has been used successfully to improve traits through identification of beneficial alleles in crops (Knoll et al. 2011; Aslam et al. 2016). During the past decade of technology development, TILLING has been improved. More recently, TILLING by next-generation sequencing (NGS), as the latest mutation detection method, has been reported as a robust tool for identifying allelic series of mutations in crops, such as potato (Muth et al. 2008). For TILLING in tef, knowledge of the genome is critical. Tef is closely related to rice and sorghum in the Poaceae family, which are two well-characterized small-genome diploid species. Cannarozzi et al. (Cannarozzi et al. 2014) performed genome and transcriptome sequencing of tef to help with the identification of breeding targets and publicly released a draft genome (672Mbp and 38000 transcripts). Thus, with the help of both this draft genome as a reference and the rice and sorghum genomes, TILLING by sequencing is now an alternative strategy to traditional CEL1-based screening to identify induced variation in starch related genes in tef.

In our study, we firstly performed the identification of starch related genes in tef, followed by the investigation of their expression in tef developing seeds. To create tef varieties with altered starch properties, we applied both conventional TILLNG using the CEL1 enzyme based system and worked to establish a more efficient TILLING-by-sequencing approach.

3.3 Materials and methods

Bioinformatics for the identification of starch related genes

The coding sequence of each starch-related gene from either rice or sorghum was used as the query for BLAST search against the tef genome database in CoGe platform with the default parameters and settings (Tef draft genome: Group of Dr. Zerihun Tadele: (id 22790) v1.1.2 unmasked 607,680,327nt; CoGe platform: Cogehttps://genomevolution.org/CoGe/CoGeBlast.pl).

Sequence alignments were conducted with default setting in CLC Genomics Workbench (CLC bio, Denmark). Based on the aligned sequences of genes within each class the relatedness phylogenetic trees were constructed as neighbor-joining trees with Jukes-Cantor nucleotide distance measures. Bootstrap values were derived from 100 replicates.

GEvo analysis was performed via the web-based platform, where the default setting of sequence comparison algorithms were used (https://genomevolution.org/CoGe/GEvo.pl). The output of gene comparison was visualized and genomic regions with high sequence similarity denoted with a colored overlay.

RNAseq library sequencing, mapping, and statistical analysis

Seeds were collected from four 4-month-old tef plants (DZ-Cr-37, *Tsedey*) under a binocular microscopy and immediately frozen in liquid N_2. Total RNA from each seed pool was extracted as follows: around 80 mg seeds were powdered while frozen. RNA extraction buffer (600 µL of 150 mM Tris, pH 7.5, 2% (w/v) SDS, 50 mM EDTA) was added and the mixture vortexed. Ethanol absolute (150 µL), potassium acetate (5M, 66 µL) and chloroform: isoamyl alcohol (24:1; 750 µL) were then sequentially added to the mix, vortexed and separated by centrifugation at 13,000 g for 3 min. The supernatant (600 µL) was taken and further cleaned by vortexing with 600 µL phenol:chloroform:isoamyl alcohol (25:24:1). Ethanol was added to the supernatant and the mixture was incubated at -80 °C for 30 min to precipitate nucleotides. The precipitate was washed with 80% (v/v) ethanol, dissolved in 75 µL DEPC water and RNA precipitated by addition of 25 µL 8 M LiCl in DEPC water for 16 h at -20 °C. The RNA pellet was washed with 80% (v/v) ethanol, dissolved in 30 µL DEPC water.

About 0.5-1 µg of the extracted total RNA was used for RNA-seq library preparation using TruSeq™ RNA Sample Prep Kit v2 (Illumina, Inc.). RNAseq library quality was assessed using a Bio-analyzer and sequenced on an Illumina HiSeq2500 at Fastaris SA (Plan-les-Ouates, Switzerland). The sequence reads from each sample were aligned to the tef genome database using CLC Genomics Workbench (CLC bio, Denmark) and the RNA-seq data were analyzed using the EdgeR soft-ware in the CLC Genomics Workbench. This generated expression values as RPKM (reads per kilobase of transcript per million mapped reads), fold change values, p-value, and FDR (false discovery rate corrected p-value).

The quality of the RNA-seq data was assessed using principal component analysis (PCA) with the original expression values. The number of transcripts with the mean RPKM >1 in each developing stage was classified as expressed genes, and displayed in a Venn diagram using Venny 2.0. The number of transcripts at each stage differentially expressed at above 1.5-fold change between any pair of stages (FDR <0.05) were displayed in a Venn diagram conducted in CLC. Vocano plots were created to display the distribution of differentially expressed genes. PCA and Vocano plots were conducted in CLC as well.

TILLING by CEL1

Primers

Primer3 was used to design primers with following parameters: primer length, 18-25 bp; T_m, 55-62°C; GC content, 40-60%. The Metabion biocaclulator program was used to check the oligonucleotide properties and self-complementarity. BLAST searches of the oligonucleotide sequences against the tef genome database were performed to determine uniqueness of the target sequences.

Common primers for the amplification of both homoeologous sequences for each gene were from the tetrapoid genome designed based on conserved regions deduced from their alignments in CLC Workbench. Specific primers were designed using the sequence variation of two copies (e.g. in intron regions) with genome-specific SNP at the 3'end of the primers. For fluorescent primers detected in via the LI-COR 4300 DNA Analyser, specific primers (forward) were labeled with IR Dye 700 (681) and common primers (reverse) were labeled with IR Dye 800 (781).

Cloning

Cloning was performed using pGEM®-T easy vector system (Promega). PCR products were inserted in pGEM-T vectors and transformed into DH5-alpha competent cells. The strains were selected by ampicillin and blue/white colour screening method. Colonies were picked and success of the cloning was confirmed by PCR and Sanger sequencing.

PCR amplification and heteroduplex formation

Primers were mixed according to a volume ratio of 50 : 70 : 120 : 0 of primer stock (IRDye 700 primer : unlabelled forward primer : IRDye 800 primer : unlabelled reverse primer). Each PCR reaction contained 1 µL of pooled genomic DNA and was run using the following program: 98 °C for 1 min, 6 cycles (98 °C for 10 s, 64 °C for 30 s, 72 °C for 50 s), 38 cycles (98 °C for 10 s, 60 °C for 30 s, 72 °C for 50 s), 72 °C for 10 min, followed by 99 °C for 10 min, loop for 23 cycles (70 ºC for 20 s, reduce temperature 0.9 ºC per cycle) and finally hold at 8 °C.

CEL1 treatment and clean-up

CEL1 extract (0.2 µL, produced in house at the University of Bern) was added to each PCR reaction product and incubated at 45 °C for 15 min. The reaction was stopped by adding 0.15 M EDTA (pH 8.0). Afterwards, CEL1 digestion products were purified through gel filtration medium Sephadex® G-50 (Sigma). The total volume was reduced at 90 °C to approximately 1.5 µL (this takes about 45 min) and the formamide load dye was added to each sample.

Electrophoresis and gel image analysis

After reducing the volume by heating, aliquots were transferred to a 100-tooth membrane comb, which was then inserted into the well of a slab gel (6.5% [w/v] acrylamide, 7 M urea) and the DNA separated by denaturing gel electrophoresis (3 h 45 min at 1500 V, 40 mA, 40 W, 50 ºC). The fluorescent dyes were detected in two separate channels on the Li-COR 4300 DNA Analyser yielding IR700 and IR800 images. Mutation detection was performed using GelBuddy software.

TILLING by sequencing

Primers and PCR amplification

Multiple primers were initially designed for amplification of each complete gene. The process was similar as mentioned above, with using the Primer3, the Metabion biocaclulator program and the BLAST tool. The PCR was carried out in a 20 µL final volume containing 20 ng DNA, 0.4 U iProof high fidelity DNA polymerase in 1 × PCR buffer (Bio-Rad, Hercules, CA), 0.2 µM each of dNTP, and 0.5 µM each forward and reverse primers. The PCR conditions were as follows: 98 °C for 1 min, 6 cycles (98 °C for 10 s, 64 °C for 30 s, 72 °C for 50 s), 38 cycles (98 °C for 10 s, 60 °C for 30 s, 72 °C for 50 s), 72 °C for 10 min and hold at 4 °C. The PCR products were checked for the correct sizes and purity by agarose gel electrophoresis.

DNA sequencing data analysis

Data analysis was performed at the ETH High Performance cluster Euler (Erweiterbarer, Umweltfreundlicher, Leistungsfähiger ETH-Rechner). Bioinformatics tools used were Samtools, Bowtie2, IGV, Freebayse and CLC Genomics Workbench.

High-resolution melt (HRM) analysis

HRM experiments were performed by PCR amplification of a specific ~100 bp fragment. HRM curve acquisition and analysis were performed on ABI 7500 Fast RT-PCR System. MeltDoctor HRM master mix (Applied Biosystems, 4415440) was used for HRM analysis. A reaction volume of 20 µL reaction was prepared containing 20 ng genomic DNA. PCR cycling parameters for real-time PCR and HRM curve acquisition were followed by: one cycle of initial denaturation at 95 °C for 10 min, 6 cycles (95 °C for 10 s, 64 °C for 30 s, 72 °C for 30 s), and 34 cycles (95 °C for 10 s, 60 °C for 30 s, 72 °C for 30 s). The melt step of HRM analysis was performed as follows; heat at 95 °C for 10 s and then anneal at 60 °C for 1 min, afterwards heat at 95 °C for 30 s and then anneal at 60 °C for 15 s. The normalized HRM data were subjected

to gene scanning analysis by ABI HRM software to identify the changes in melting curves, which indicated the presence of variation in the amplified target sequence.

3.4 Results

3.4.1 Identification of starch related genes in tef

To create tef varieties with low or no amylose, genes encoding enzyme isoforms related with amylose synthesis were targeted, including GBSSI, GBSSII and PTST. Creation of tef varieties with higher amylose content could be achieved by altering genes encoding enzymes that contribute to amylopectin synthesis, particularly the SBEs. Unlike the well-known species rice and sorghum, the biosynthetic genes involved in endosperm starch biosynthesis in tef remain unknown. The starch related genes in the tef genome were identified through sequence similarity search in the closely related species rice and sorghum. The coding sequence of each gene either from rice or sorghum was used to search the tef draft genome. Gene candidates encoding each enzyme were chosen and are listed in Table 1. Using the enzyme GBSSI as an example, augustus_masked-scaffold6196-abinit-gene-0.5 and snap_masked-scaffold12082-abinit-gene-0.25 are the two highest hit genes after BLAST searches. Due to the allotetraploidy of tef, these two genes presumably represent the two homeologs and we named them as EtGBSSIa and EtGBSSIb. The same rule was applied to name other genes encoding putative starch biosynthetic enzymes (Table 1).

For most of the starch related genes identified the situation was similar as for GBSSI, and there were two tef orthologues, designated as “a” and “b”, for each rice or sorghum gene. These included genes for the enzymes GBSSI, GBSSII, SBEIIb, SSI, SSIIc, PTST, ESV1, LESV and RSRI (Table 1). To further confirm and compare the “a” and “b” genes, sequence alignments were performed. Using GBSSI as an example, we aligned the genomic and coding sequences of EtGBSSIa, EtGBSSIb, OSGBSS*I* and SBGBSS*I* genes (Figure 1). As a result, we observed that the borders between exons and introns of tef genes were identical to those in rice and sorghum and the sequence similarity of the conserved exon segments was high among all three species (Figure 1). This supports the conclusion that EtGBSSIa and EtGBSSIb were the two homeologous genes coding GBSSI enzyme in tef. The same was also applied for GBSSII, SBEIIb, SSI, SSIIc, PTST, ESVI, LESV and RSRI and the alignment based results were confirmed for those genes (not shown).

In a few cases, the situation was different. Either more than two genes for each enzyme were identified. Alternatively, some homeologous genes appeared to be missing or were only partially represented. We used an alternative strategy, GEvo (Genome Evolution Analysis tools), for syntenic comparisons to further analyze those genes. GEvo was initially designed to quickly identify patterns of genome evolution by comparing multiple genomic regions from any number of organisms using a variety of different sequence comparison algorithms. For

SBE1 enzyme, *EtSBEI* obtained three gene candidates with high blast scores, named as *EtSBEIa*, *EtSBEIb* and *EtSBEIc*.

Table 1 Starch related genes identification in tef

Enzymes isoforms	Gene accessions in sorghum	Gene accessions in rice	Blast result in tef database	Gene name
GBSSI	Sb10g002140	LOC_Os06g04200	augustus_masked-scaffold6196-abinit-gene-0.5	*EtGBSSIa*
			snap_masked-scaffold12082-abinit-gene-0.25	*EtGBSSIb*
GBSSII	Sb002g116000	LOC_Os07g22930	maker-scaffold6672-augustus-gene-0.28	*EtGBSSIIa*
			maker-scaffold66-augustus-gene-0.13	*EtGBSSIIb*
SBEI	Sb10g030776	LOC_Os06g51084	maker-scaffold686-snap-gene-0.57	*EtSBEIa*
			maker-scaffold8859-augustus-gene-0.24	*EtSBEIb*
			maker-scaffold10020-augustus-gene-0.19	*EtSBEIc*
SBEIIb	Sb10g273800	LOC_Os02g32660	maker-scaffold6829-snap-gene-0.4	*EtSBEIIa*
			maker-scaffold8367-snap-gene-0.35	*EtSBEIIb*
SSI	Sb010g04770	LOC_Os06g06560	maker-scaffold5436-augustus-gene-1.30	*EtSSIa*
			snap_masked-scaffold8478-abinit-gene-0.37	*EtSSIb*
SSIIa	Sb10g008200	LOC_Os06g12450	maker-scaffold6057-snap-gene-1.57	*EtSSIIa*
SSIIb	Sb04g028060	LOC_Os02g51070	maker-scaffold10799-snap-gene-0.15	*EtSSIIb*
SSIIc	Sb001g239500	LOC_Os10g30156	maker-scaffold8547-snap-gene-0.27	*EtSSIIc*
			maker-scaffold373-snap-gene-1.58	*EtSSIId*
SSIVa	Sb09g026570	LOC_Os01g52250	maker-scaffold5048-snap-gene-0.17	*EtSSIVa*
			maker-scaffold5595-augustus-gene-0.11	*EtSSIVb*

SSIVb	Sb09g026570	LOC_Os05g45720	maker-scaffold941-snap-gene-1.37	*EtSSIVc*
PTST	Sb04g002930	LOC_Os02g04330	maker-scaffold8269-augustus-gene-0.47	*EtPTSTa*
			maker-scaffold2368-snap-gene-1.52	*EtPTSTb*
ESVI	Sb001g515600	LOC_Os03g04100	maker-scaffold2737-augustus-gene-0.39	*EtESVIa*
			maker-scaffold12818-augustus-gene-0.29	*EtESVIb*
LESV	Sb005g167800	LOC_Os11g37560	maker-scaffold13941-augustus-gene-0.23	*EtLESVa*
			maker-scaffold2434-snap-gene-0.35	*EtLESVb*
RSRI	Sb09g002080	LOC_Os05g03040	snap_masked-scaffold306-abinit-gene-2.11	*EtRSRIa*
			maker-scaffold252-snap-gene-1.16	*EtRSRIb*

Together with the rice *SBEI* gene, a syntenic comparison was conducted. The result was displayed such that the regions of similar sequence were colored the same and connected (Figure 2). *EtSBEIb* was fully covered *OSSBEI* gene region with high similarity displayed. Both *EtSBEIa* and *EtSBEIc* were partially overlaid *OSSBEI* gene region. *EtSBEIa* and *EtSBEIc* together were able to cover the entire region. This might indicate *EtSBEIb* tends to be the tef *SBE1* gene (Figure 2). For SSII enzyme*,* only one tef gene was found for SSIIa (*EtSSIIa*) through BLAST searches, and the same for SSIIb (*EtSSIIb*) (Table 1). For SSIVa and SSIVb, one gene (*EtSSIVc)* was found to have high sequence similarity and coverage to both *OSSSIVa* and *OSSSIVb*. This might be due to the similarity between *OSSSIVa* and *OSSSIVb*. Besides *EtSSIVc*, two other candidate genes *EtSSIVa* and *EtSSIVb* also scored highly in the BLAST searches, while *EtSSIVa* had only matched part of the rice gene. *EtSSIVb* had a degree of similarity simply falls below the threshold required for defining a syntenic region. (Figure 3).

To further explore the relatedness of the identified starch biosysthesis genes, phylogenetic trees including tef genes together with the genes from rice, sorghum and, in some cases Arabidopsis, were constructed (Figure 4). In general, genes encoding same enzyme isoform from all species clustered together, and the two tef homeologs were closely related to each other within the sub-clusters (e.g. GBSSII, SBEI, SBEIIb, SSI, SSIIc and RSRI). For the cases which had uncertain homeologs (SBE1, SSII and SSIV) phylogenetic relatedness was consistent with the results obrtained from the syntenic comparisons. *EtSBEI*a, *EtSBEIb* and *EtSBEIc* were related in *SBE1*cluster. The formed sub-clusters of four *EtSSII* genes clearly separated *EtSSIIa* and *EtSSIIb* and they were more distant from *EtSSIIc* and *EtSSIId*. For SSIV, *EtSSIVa* and *EtSSIVc* were more closely related to *OSSSIV* and *SBSSIV,* and *EtSSIVb* was less related.

All in all, our search for starch metabolism genes in tef was successful in identifying a comparable number of genes as in other species, and to define likely homeologs. To investigate whether these genes were functional in tef, a transcriptome analysis of developing tef seeds was performed.

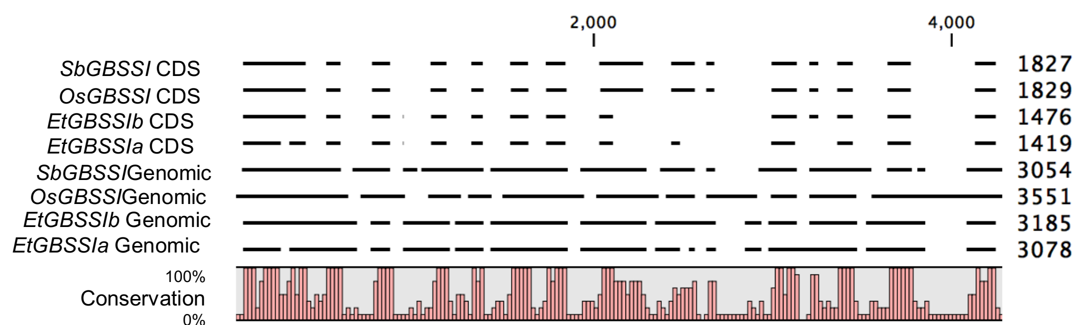


Figure 1. Alignment of genomic and coding sequence of *GBSSI* genes in tef, rice and sorghum
The name of each sequence is indicated on the left and length of sequence indicated on the right. An overview of the degree of conservation after alignment is shown below.

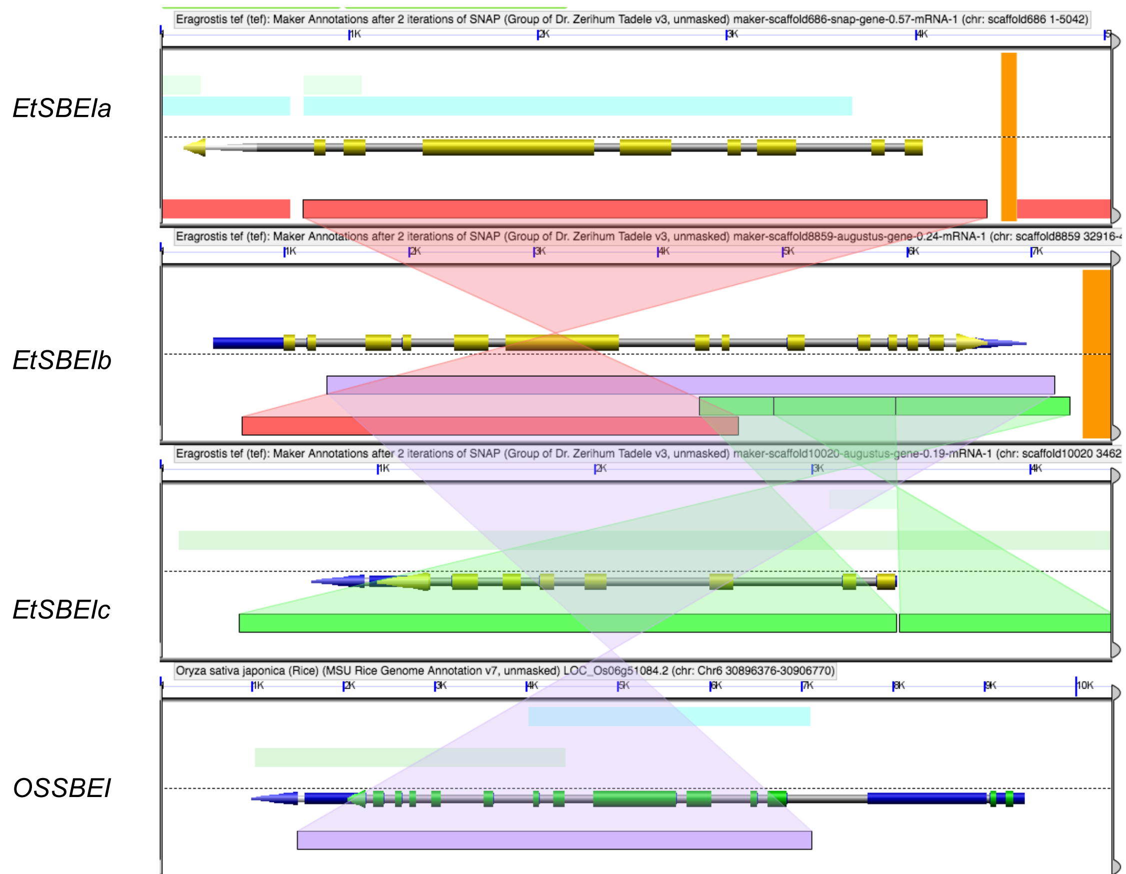


Figure 2. Syntenic comparison of tef and rice *SBEI* genes using GEvo
Gene names are indicated at the left side. Similarly coloured bars below each gene represent the similarity and the coverage region. In gene *EtSBEIb,* purple colour means the identity matched with *OSSBEI*. Red colour (*EtSBEIa*) and green colour (*EtSBEIc*) bars are matched individually for the 5' and 3' region of the *EtSBEIb* gene*.* Identity scores for all matches were above 70%.

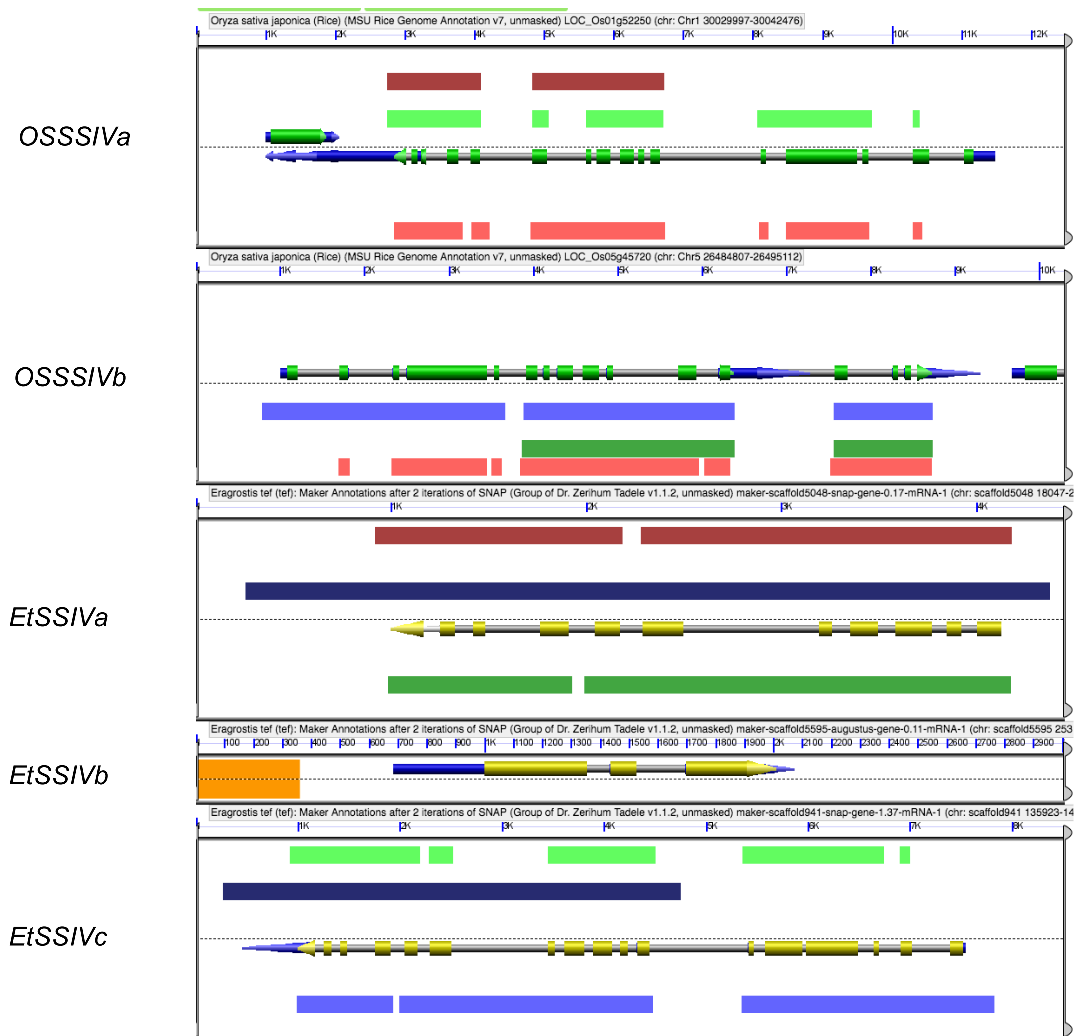
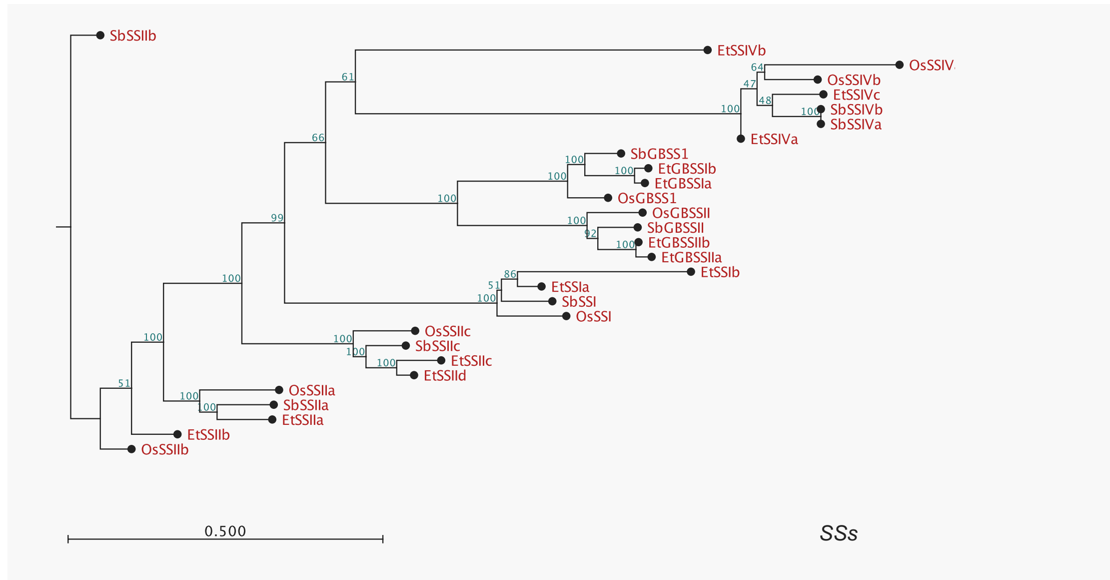
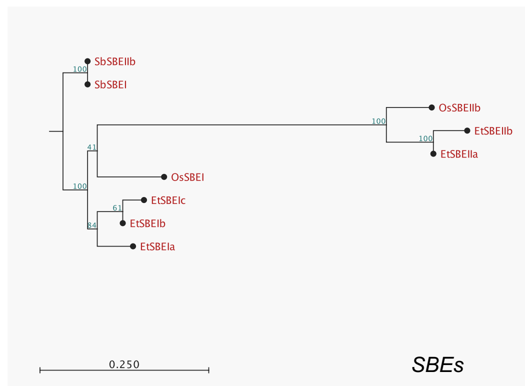


Figure 3. Syntenic comparison of tef and rice *SSIV* genes using GEvo
Gene names are indicated on the left. *EtSSIVc* has high similarity and coverage to both *OSSSIVa* (light green) and *OSSSIVb* (light blue). *EtSSIVa* partially overlaps both *OSSSIVa* (dark red) and *OSSSIVb* (green). Identity scores for all matches were above 70%. *EtSSIVb* has lower similarity to the other genes.

A



B



C

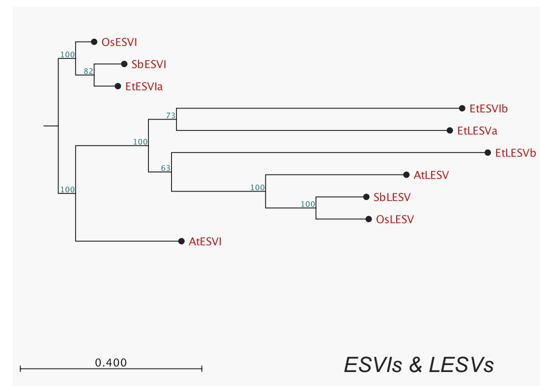


Figure 4. Relatedness of identified starch related genes
Phylogenetic trees were made for each group of genes, using rice (Os), tef (Et), sorghum (Sb) and Arabidopsis (At) sequences. Bootstrap values (%) are indicated in the figures. A) Starch synthases (SSs) group B) Starch branching enzymes (SBEs) group C) ESV1s and LESVs group.

Figure 5. Tef developing seeds
Developing seeds were collected from a four-month-old tef plant and sorted into three developmental stages (early, middle and late), according to their size and appearance.

3.4.2 Expression profile of starch related genes in tef developing seeds

To investigate whether our identified starch metabolic genes are expressed in developing tef seeds during the period of starch deposition, total RNA was extracted from seeds and sequenced by Illumina HiSeq 2500. Each tef spikelet has seeds at different developmental stages. For our sample preparation, we classified developing seeds from an individual plant into three pools; early (E), middle (M) or late (L) (Figure 5). Early developing seeds were small and soft. Late developing seeds were hard and ripe. Middle developing seeds were intermediate between early and late stages in terms of size, colour and hardness. Four tef plants were included as biological replicates and RNA was individually extracted from 12 samples (named as 1-1, 1-2, 1-3, 2-1, 2-2, 2-3, 3-1, 3-2, 3-3, 4-1, 4-2 and 4-3, representing “plant name – developing stage (1, 2 and 3 representing E, M and L)”).

After RNAseq, the sequencing quality was checked. In total, 300 million sequences reads were generated from the 12 samples, with each sample contributing between 17.7 and 31.9 million reads. The reads from each sample were aligned to reference genome. We used here the tef draft genome (id 22790 Group of Dr. Zerihum Tadele unmasked v1.1.2). On average, 75.1% of reads were uniquely mapped, 17.4% were mapped nonspecifically and 7.5% were unmapped. The uniquely mapped reads were then counted and the RPKM value (Reads Per Kilobase of transcript per Million mapped reads) was calculated. RPKM values were used for subsequent gene expression analysis.

As mentioned above, the sampling of seeds into three pools was based on seed size, colour and texture. To check the sampling quality and to potentially exclude any outlying samples from further RNAseq analysis, we performed principal component analysis (PCA). Using the RPKM value for each gene in each sample, the PCA on the 12 samples (Figure 6A) revealed that the biological replicates clustered together and that the different developing stages were separated. This indicated that the factor influencing sample distribution in PCA plot was the developing stage, validating our sampling strategy. No replicates were excluded.

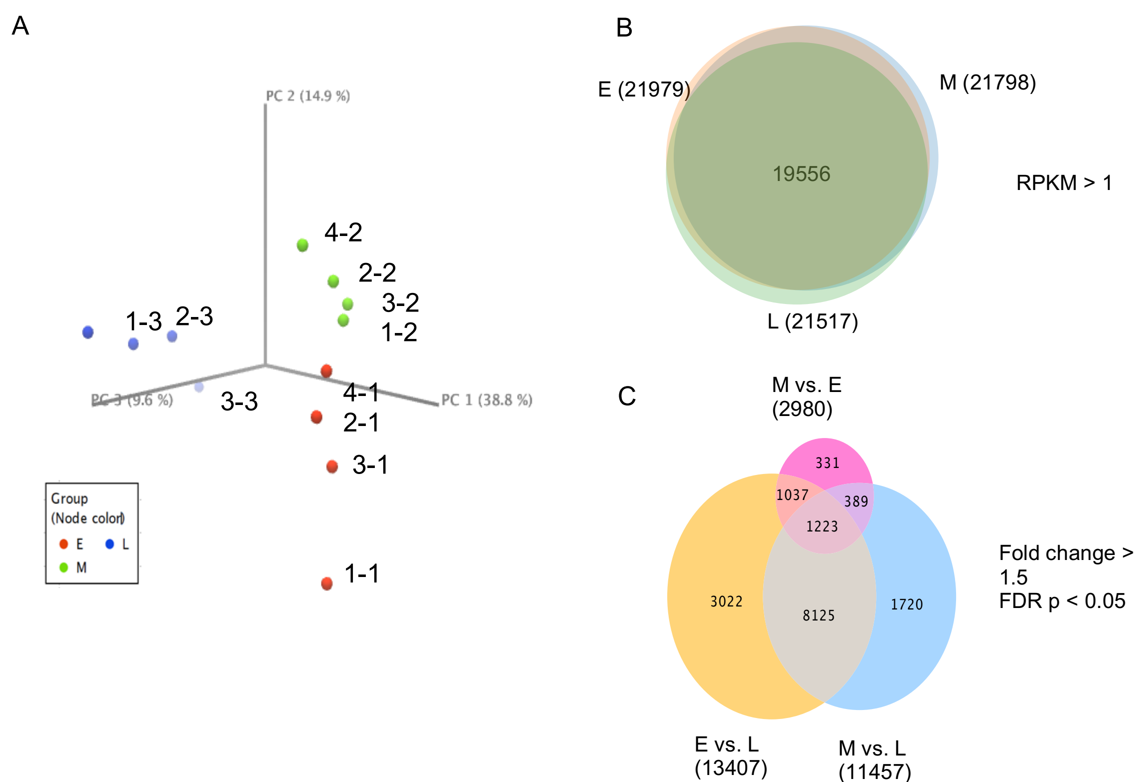


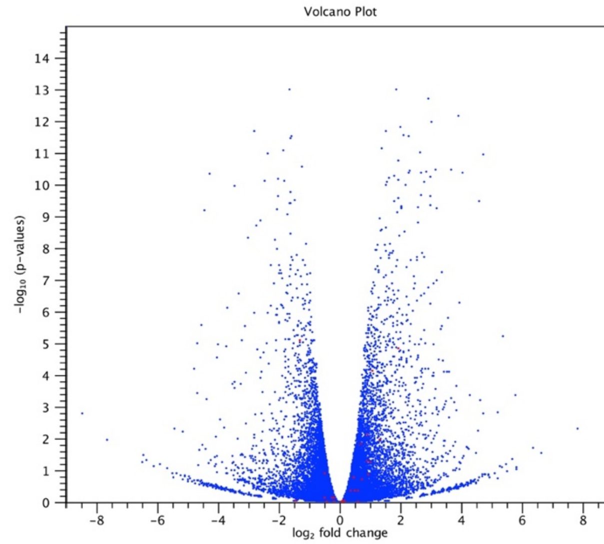
Figure 6. RNAseq for early, middle and late stage seeds
A) Principle component analysis (PCA) for all samples B) Venn diagram of genes expressed in early (E) (orange), middle (M)(blue) and late (L) (green) stages. Threshold RPKM>1. C) Venn diagram of differentially expressed genes. E vs. L (yellow), M vs. L (light blue) and M vs. E (pink).

To have an overview of genes in E, M and L stages, the mean expression value of all genes was created from the replicates samples (RPKM cutoff at >1.00). No significant difference was observed in terms of the total number of genes expressed among three stages, with 21979, 21798 and 21517 transcripts detected in E, M and L stages, respectively. The majority of transcripts (19556) were expressed in all stages (Figure 6B).

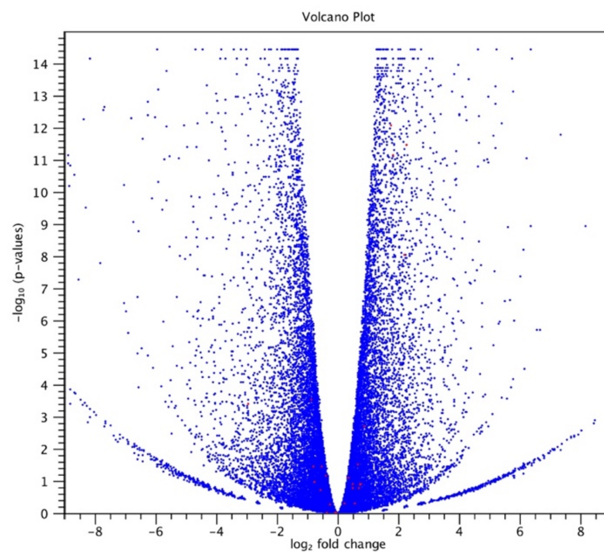
For the differentially expressed genes, pairwise comparisons of transcript abundances at different stages were conducted (Figure 6C). The late stage was very different to the middle and early stages. There were 13407 differentially expressed transcripts in E vs. L and 11457 differential transcripts in M vs. L (cutoff at fold change > 1.5 and FDR <0.05), while the number of differentially expressed genes between E vs. M was much smaller (2980) (Figure 6C). This was also reflected in our PCA plot, where the blue L cluster was more distant to the red E and green M clusters (Figure 6A).

Besides of the number of differentially expressed genes, their distribution can be displayed through Vocano plot analysis (Figure 7). The distributions of differentially expressed genes in E vs. M, E vs. L and M vs. L were different in each comparison. As expected, the E and M samples were much more similar to each other than to the L samples. there were fewer points at the top left or top right of plot, which indicate large, highly significant changes in expression. Regarding the starch related genes (Table 1), their distribution did not reveal huge differences among three stages: most of them had log2 fold-change of less than 2 in all three Volcano plots (Figure 7, red dots).

E VS M



M VS L



E VS L

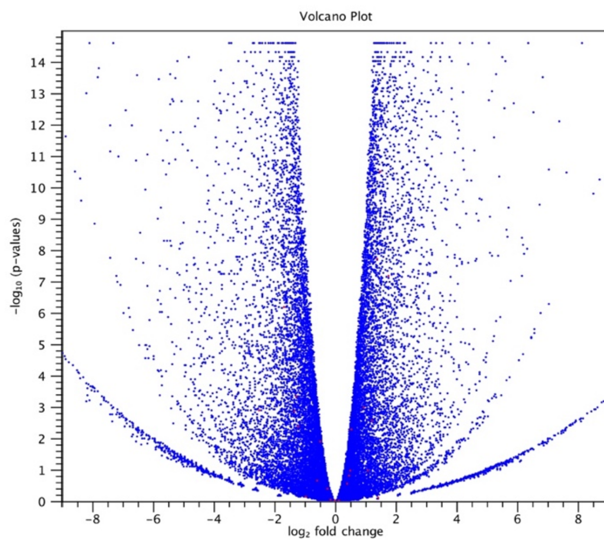


Figure 7. Volcano plots of differential expressed genes
Three comparisons of early (E) vs. middle (M) stage, middle vs. late (L) stage and early vs. late stage. Red dots are targeted starch related genes (see Table 1).

To investigate the gene expression levels of starch related genes (Table 1) and their changes in early, middle and late stages, a heatmap was created where the expression value of each gene (the mean RPKM of biological replicates) ranged from 0.6 to 208 (blue to red) (Figure 8). Among them, EtGBSSI, *EtSBEII, EtSBEI,* and EtSSI were more highly expressed than the other genes. This was very similar to the expression of the orthologous genes in rice seed development (RiceXPro). In tef, the highest expression was observed at the middle developing stage (Figure 8), while in rice, the expression trend was slightly different with a slowly increasing or similar expression during seed development (RiceXPro). However, it is difficult to compare exactly the developmental stages between the two cereals.

To conclude from the RNAseq analysis, a subset of the genes we annotated were highly expressed during seed development, with a peak in the middle stage. This analysis is important since it identifies unambiguously the genes that need to be targeted by TILLING in order to modify endosperm starch composition and structure.

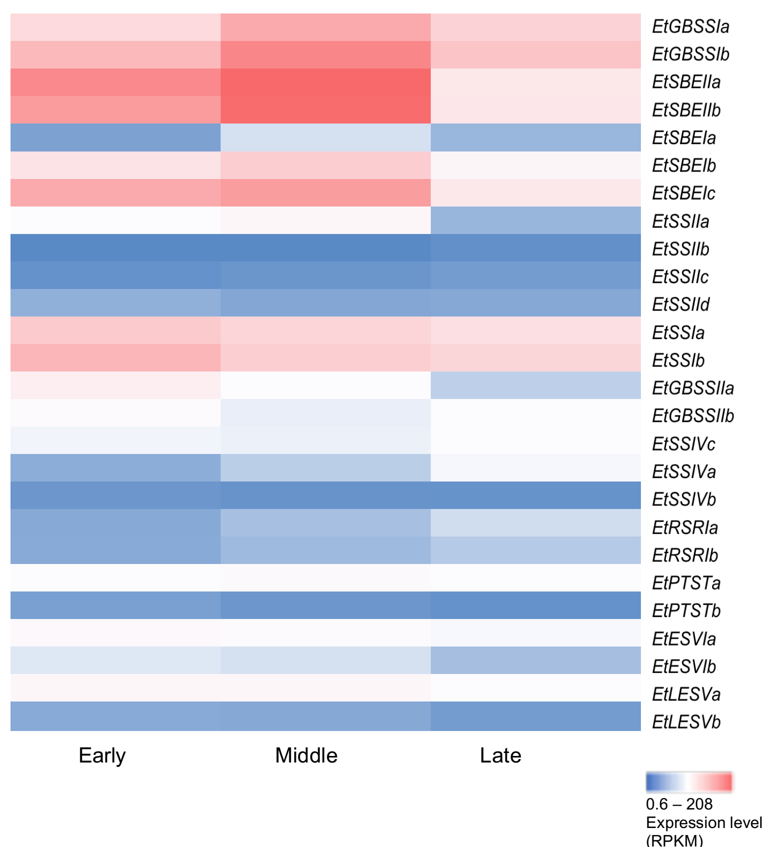


Figure 8. Heat map of expression of targeted starch related genes in three stages
All identified copies of starch metabolic genes in tef are listed. Expression level of each gene (RPKM value) was used to generate a heatmap. Red represents high expression levels and blue represents low expression levels.

3.4.3 TILLING

To create tef varieties with modified starch properties we want to screen a mutant population for useful alleles in starch-related genes. TILLING, as a reverse genetic approach, was applied since it is suitable for crops with polyploidy and large genome sizes (Barkley and Wang 2008). TILLING uses traditional chemical mutagenesis methods to create libraries of mutagenized individuals that are later subjected to high throughput screens for mutation discovery. The workflow of TILLING by use of the CEL1 enzyme is displayed in Figure 9, which summarizes the mutagenesis and mutation detection steps. TILLING by sequencing is an alternative method and differs to TILLNG by CEL1 in the mutation detection step, where the high throughput sequencing is used to achieve a greater coverage of potential mutants, saving time and cost in the long run.

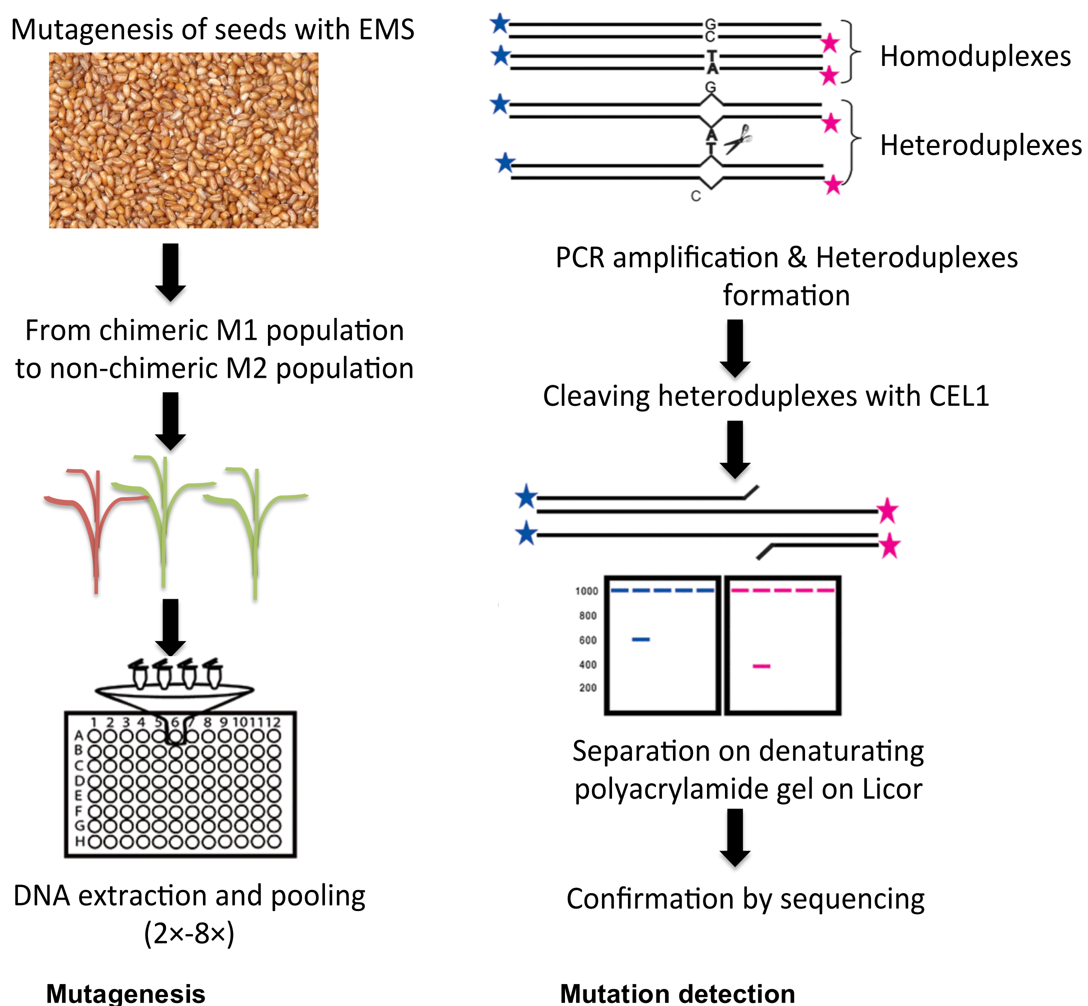


Figure 9. TILLING workflow (Targeted Induced Local Lesions IN Genomes)
Adapted from the figure generated in the group of Dr. Zerihun Tadele at the University of Bern (Esfeld et al. 2012). For a detailed description of the method, see the chapter text.

3.4.3.1 Mutagenesis stage

The mutagenesis of tef seeds (Dz-Cr-37) was performed by Dr. Zerihun Tadele's group at the University of Bern. Tef seeds were mutagenized by treatment with 2% (v/v) ethyl methanesulfonate (EMS). Plants were self-fertilized and the individuals of the M2 generation were used to extract DNA samples for mutational screening to avoid potential mosaicism in M1 generation (Figure 9). A population of genomic DNA from around 4000 individual plants were stored (Tadele, Z. and Esfeld 2009).

3.4.3.2 Mutation detection stage

TILLING by CEL1 enzyme

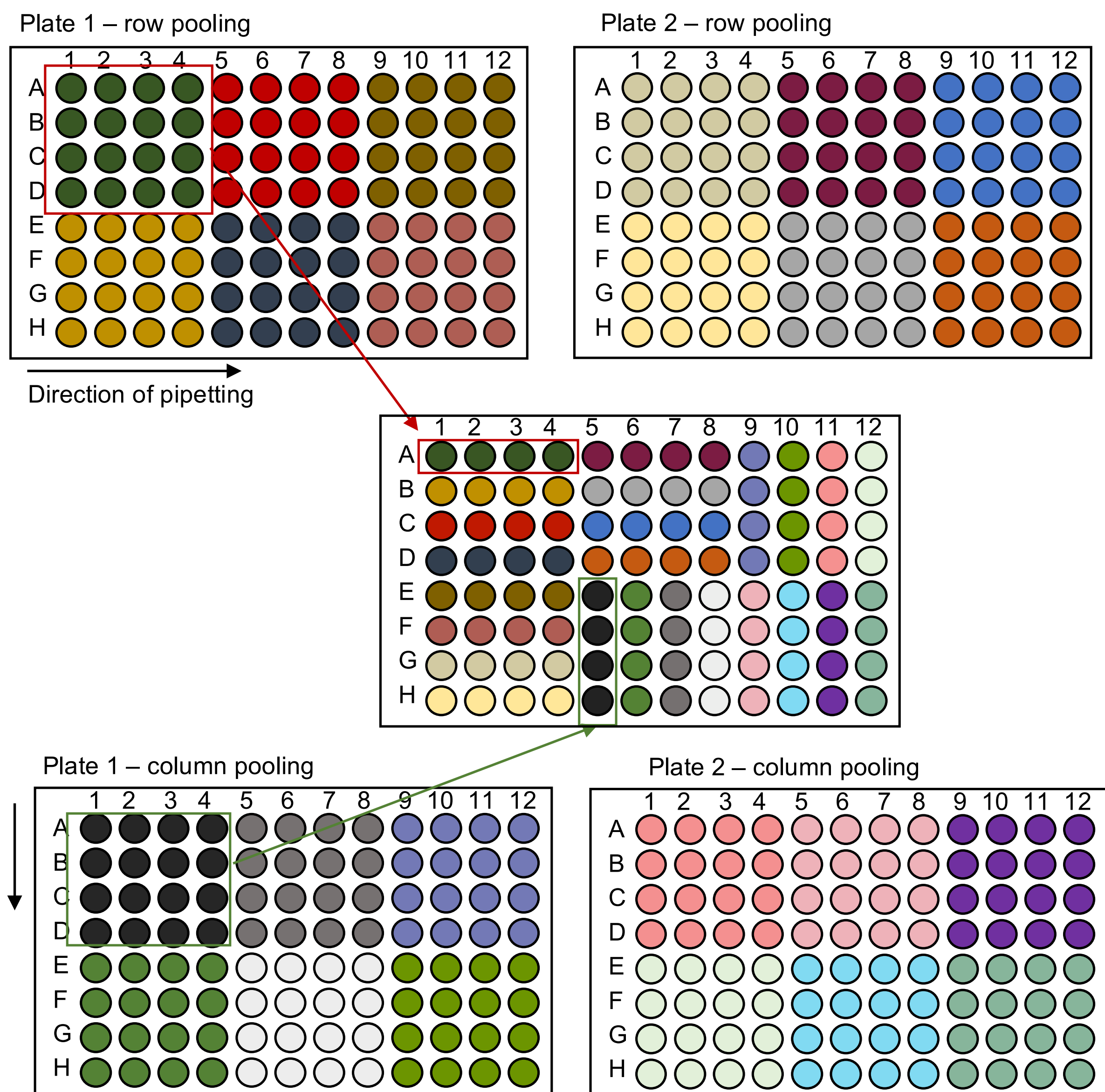


Figure 10. Two-dimension pooling for TILLING by CEL1. For a detailed description, see the chapter text.

TILLING by CEL1 for all starch related genes

Two-dimensional pooling was conducted after normalizing TILLING population DNA concentration. DNA from four thousand individuals were stored in 96 well plates. Each plate was labelled from 1 through to 44. By pooling with horizontal and vertical pipetting directions from each plate, as indicated in Figure 10, DNA from four individuals were combined into one pool and every individual was, at the end, represented in two distinct pools. For instance, DNA from plate1 A1 was also located in A1 and E5 in the final pooled plate (Figure 10). One pool plate contained individual's DNA from two plates. The resultant 22 plates generated by this two-dimensional, 4-plant pooling was used to perform mutation discovery by CEL1-based TILLING (Table 2).

Table 2, List of DNA population pools for TILLING by CEL1

Pool	Plate 1	Plate 2
1	2-7-1	2-7-113
2	2-7-244	2-7-337
3	2-7-446	2-7-465
4	2-7-561	2-7-673
5	2-7-786	2-7-948
6	2-7-1033	2-7-1136
7	2-7-1233	2-7-1335
8	2-7-1435	2-7-1539
9	2-7-1573	2-7-1686
10	2-7-1887	2-7-1927
11	2-7-2043	2-7-2275
12	2-7-2390	2-7-970
13	2-7-2587	2-7-2700
14	2-7-3002	2-7-3113
15	2-7-3226	2-7-3329
16	2-7-3432	2-7-3536
17	2-7-3640	2-7-3743
18	2-7-3852	2-7-3962
19	2-7-4055	2-7-4155
20	2-7-4457	2-7-4553
21	2-7-4649	2-7-4747
22	2-7-4853	2-7-4952

Mutation detection in the TILLING population using the CEL1-based method starts from PCR amplification of a 1-1.5 kb region of the targeted gene using fluorescent end-labeled primers. The forward-strand primers are labelled with infrared (IR) 700 dye and the reverse-strand primers are labelled with infrared (IR) 800 dye. In case of a mutation in the PCR-amplified regions in the DNA pool, heteroduplexes will formation in the PCR products. CEL1 enzymatic

cleavage at mismatch sites is visualized in denaturing polyacrylamide gel electrophoresis using a LI-COR DNA analyzer (Figure 9).

Due to the tetraploidy of tef, genome specific primers are essential for use in a successful TILLING screen. The specific primer design strategy for all starch related genes is summarized in Figure 11 and Figure 12 illustrates specific steps in creating specific primers for TILLING from two sub-genomes using GBSSI as an example. BLAST searches, as mentioned above, were conducted to identify the homeologous starch-related genes in tef (e.g. EtGBSSIa and EtGBSSIb for GBSSI gene). The optimal region within the gene for TILLING (i.e. the region which, when mutated has the highest possibility of producing nonsense and missense mutations) was chosen using the web-based program CODDLE (Codons Optimized to Discover Deleterious Lesions; Figure 12A). Non-fluorescent primers were first designed to simultaneously amplify the optimal region of both homeologous genes from wild-type tef genomic DNA. The PCR amplicons of both homeologs were inserted into pGEM-T vectors and transformed into *E. coli*. Positive single colonies were selected and the cloned fragments were sequenced. The amplicon sequences of about 16 randomly selected colonies were aligned and fell into two different groups, reflecting the homeologous genes (Figure 12B). These showed variation, particularly in intronic sequences. On the basis of these sequences fluorescent homeolog-specific primers were designed. Because of inevitably sequence similarity between the homeologous genes, primer specificity was further confirmed by PCR using plasmids containing each homeolog as templates and primers at different combinations (Figure 12C).

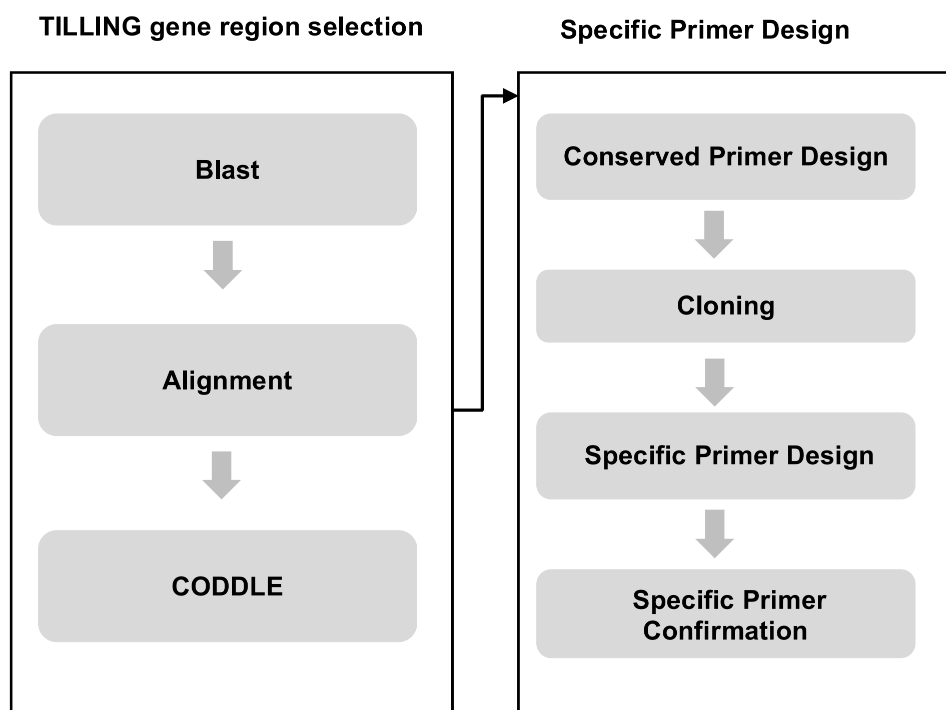
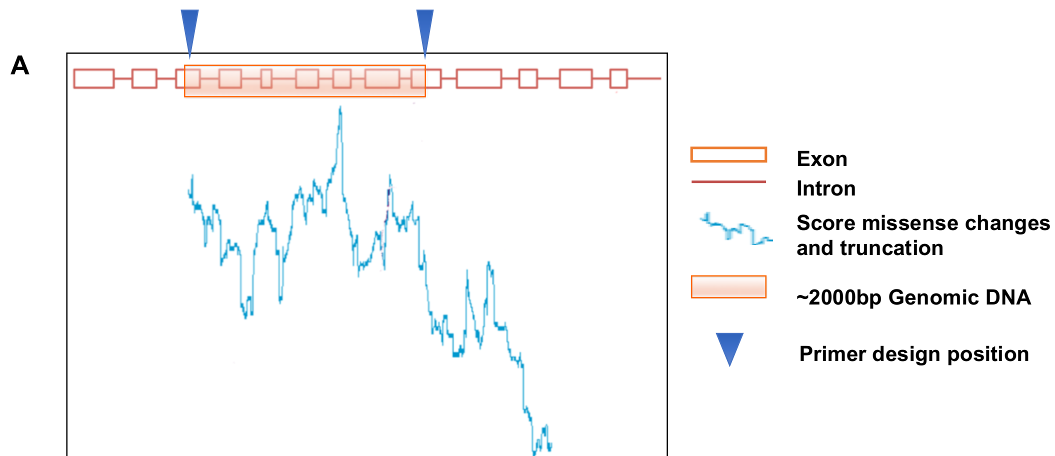
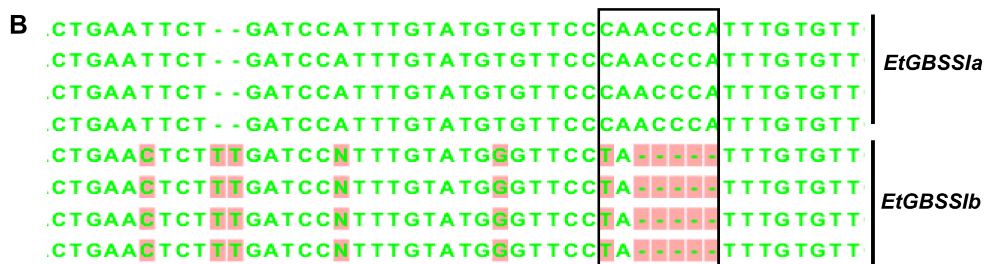


Figure 11. Specific primer design process for TILLING by CEL1





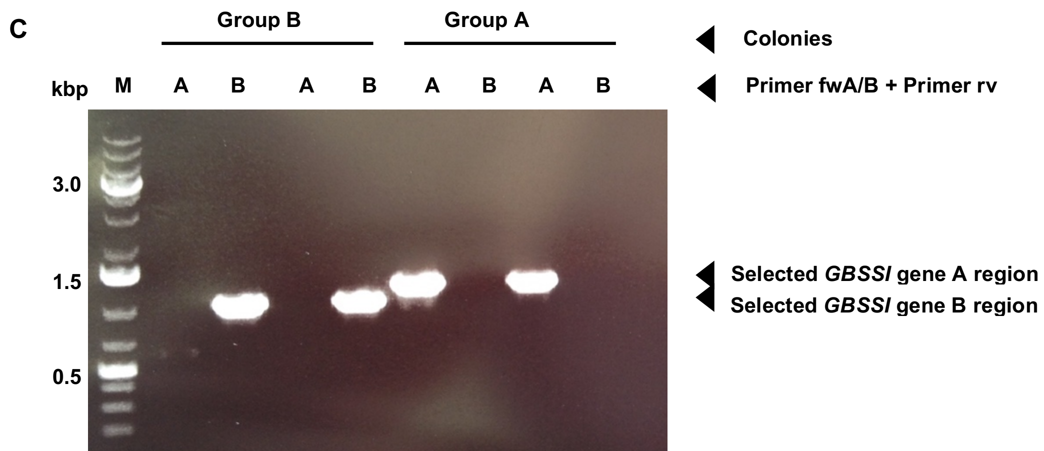


Figure 12. Specific primer design for *EtGBSSIa* and *EtGBSSIb*

A) CODDLE analysis for optimum TILLING region selection using sorghum *GBSSI*. The box denotes a 2-kb exon-rich region with a high CODDLE score

B) Alignment region of *EtGBSSIa* and *EtGBSSIb* for specific primer design from each homeologous gene

C) Confirmation of primer specificity against the intended target sequence

TILLING by CEL1 for* GBSSI *mutants

CODDLE analysis using well-annotated sorghum reference gene Sb10g002140 predicted the most suitable regions for the CEL1 TILLING of GBSS1 (Figure 12A). Amplification of a 2-kb fragment using primers designed to conserved regions allowed unambiguous identification of EtGBSSIa and EtGBSSIb as the two tef GBSSI homeologs (Figure 12B). Variation in the non-conserved intronic region, allowed the design of gene-specific primers to amplify ~1200bp and of ~1000bp for EtGBSSIa and for EtGBSSIb, respectively for the TILLING screen. Primers for one homeolog would not amplify the other despite use of a common reverse primer (e.g. plasmid DNA containing EtGBSSIa (Group A) gave no PCR product using the primer specific to EtGBSSIb, but did when using the primer specific to EtGBSSIa). For TILLING, forward specific primers were then labeled with IRDye 700 and the common reverse primer was labeled with dye IRDye 800. Both conserved and specific primers used for TILLING of GBSSI are listed in Table 3.

Table 3. List of primers for *GBSSI* in TILLING

Name	Sequence (5' - 3')
Conserved primers	
EtGBSSI FW	GAT CGC GTG TTC ATC GAC CA
EtGBSSI RV	CTT CCT TGT TGA GCG CCT TC
Specific primers	
EtGBSSIa FW	GTT GCT GAG TTG AGT ACT ACT G
EtGBSSIb FW	CCA TTT GTA TGT GTT CCC AAC CC

To screen the mutated tef population for mutations in GBSS, PCR amplification of the target fragments was done using the homeolog-specific GBSSI primers and the pooled genomic DNA samples (described above). After PCR amplification, heteroduplex formation between mutant and wild type DNA strands was performed by denaturing and annealing, followed by incubation with CEL1 to digest mismatched base pairs. DNA was then purified, sample volume reduced, and each sample loaded on a denaturing polyacrylamide gel and run on the LI-COR DNA Sequence Analyzer. Two fluorescence images from 700nm and 800nm channels were produced for per gel run.

In Figure 13, an example of mutation discovery of *EtGBSSIa* in the DNAs in pool 2 is shown. Two novel bands of ~750bp resulting from endonucleolytic cleavages are present in lane 34 and 64 of 700 nm channel. Corresponding novel bands of ~500bp in same lanes from the 800nm channel were also found (Fig. 13). The sum of the band sizes (750bp and 500bp) was equal to the full-length product visible at the top of the image, and was distinguished from background patterns caused by miss-priming. Thanks to our pooling strategy, the duplicate appearance of novel bands (Figure 13) allowed us to confirm the positive mutation over false positive signals. Furthermore, the band positions allow deconvolution of the pools to provide the identity of the individual containing the mutation.

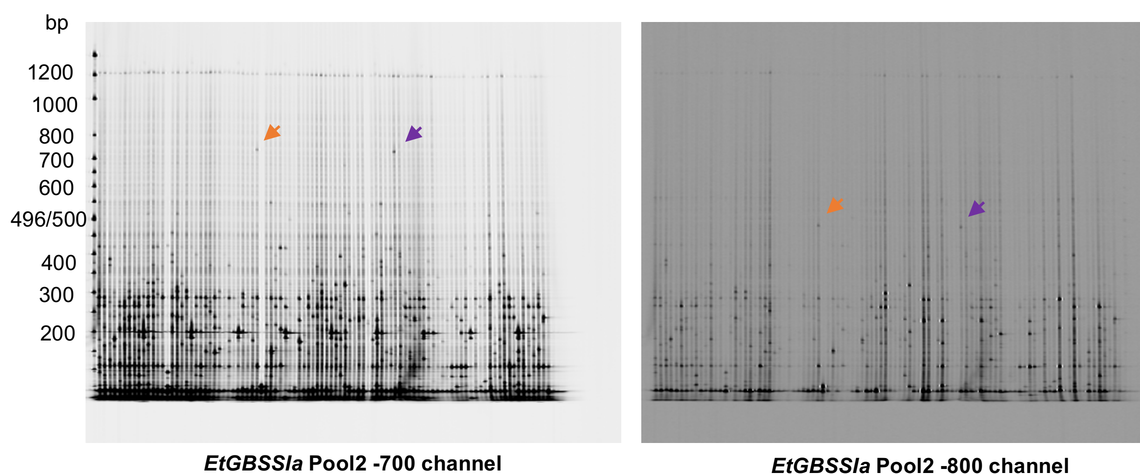


Figure 13. Mutations detected using the LI-COR 4300 DNA Sequence Analyzer
Example of individual DNA samples of 192 mutated tef plants (M_2 population) were pooled 4-fold in a two dimensional range to pool2. TILLING by CEL1 for *GBSS1* mutants was performed. Two images from 700 and 800 channels of LI-COR analyzer were displayed. From the images, each lane contains the PCR products from 4 individuals and each individual is repeated in two lanes.The orange and purple arrows (in lanes 34 and 64) indicate cleaved PCR products and the molecular weight of the representative cleaved fragments, sum up to the molecular weight of the full length PCR product size of *EtGBSSIa* (1200bp).

In conclusion, the screening of ~2112 mutant lines for the EtGBSSIa gene identified 5 mutants (2-7-737, 2-7-1134, 2-7-122, 2-7-1168 and 2-7-1387; Table 4). In the case of EtGBSSIb gene, ~2880 mutant lines were screened and three mutations were detected (2-7-1409, 2-7-142 and 2-4-791; Table 4). The exact nucleotide change in each case was determined by sequencing.

Table 4. *EtGBSSIa* and *EtGBSSIb* mutation detection using CEL1-based TILLING on a LI-COR DNA Sequence Analyser

Gene	Mutation	DNA Pool	Band-700 channel	Band-800 channel	Mutant individual	Seed name of mutant Individual	Base change
EtGBSSIa	1	Pool4	Lane 8	Lane 57	Plate 673 - 8A	2-7-737	C >> T
	2	Pool2	Lane 34	Lane 64	Plate 244 - 12G	2-7-1134	C >> T
	3	Pool1	Lane 23	Lane 74	Plate 113 - 2B	2-7-122	n.d.
	4	Pool6	Lane 5	Lane 57	Plate 1136 - 5A	2-7-1168	n.d.
	5	Pool7	Lane 7	Lane 81	Plate 1335 - 7C	2-7-1387	n.d.
EtGBSSIb	1	Pool7	Lane 41	Lane 96	Plate 1335 - 9H	2-7-1409	n.d.
	2	Pool7	Lane 31	Lane 95	Plate 1335 -11D	2-7-1421	C >> N
	3	Pool5	Lane 8	Lane 57	Plate 948 - 8A	2-4-791	C >> Y

Note:
n.d. represents that the base change has not been checked by sequencing.
Gray highlighted mutations were used as controls in TILLING by sequencing strategy below.

TILLING by sequencing

As indicated at the previous section, eight *GBSSI* TILLING mutant candidates were identified after screening around 4000 plants. Among the four sequence confirmed mutations, no stop codon was detected (Table 4). Mutation individual, named as Plate 948 - 8A, was located in intron. The nucleotide changes in mutations, Plate 673 - 8A and Plate 1335 -11D, were predicted to result no amino acid changes from protein prediction analysis. Mutant individual, named as Plate 244 - 12G, was predicted to have an amino acid change from Serine to Phenylalanine. This took several months to achieve. To boost the mutation detection for *GBSSI* and the other starch related genes (Table 1), high-throughput sequencing technologies were applied using the same DNA populations. Our strategy was to PCR-amplify each starch related gene (Table 1) from super-pooled DNA and deep sequence the mixed amplicons (Figure 14). Variant calling and other bioinformatics tools were then used to identify single nucleotide polymorphisms (SNPs) in the genes of interest (Figure 16). Based on our pooling strategy, the possible range of individuals carrying a SNP was reduced to 96 plants. HRM and Sanger sequencing were then used for individual mutant identification.

Amplicon libraries preparation and High-throughput sequencing

A pooling strategy of genomic DNA was used in preparation for the PCR and high-throughput sequencing strategy (Figure 14). Genomic DNA from two 96-well plates (i.e. 192 individuals in plates 1 and 2) were taken as one pool. To ensure each 96-well plate was represented twice, pooling was conducted as described in Figure 14, by co-sampling adjacent plates. In total, 44 genomic DNA pools were generated.

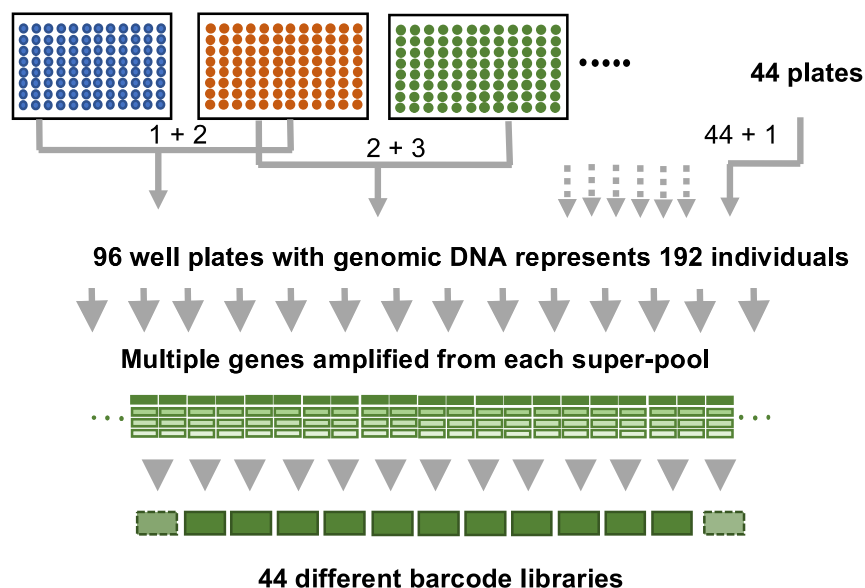


Figure 14. TILLING-by-sequencing amplicon library preparation
Two neighbor 96-well plates were pooled together to be one superpool with having one plate duplicated. Each superpool contains 192 individuals and used as template for PCR amplification of the targets. PCR products for all starch related genes were amplified from each superpool and combined to make 44 different Illumina barcode libraries.

Multiple pairs of primers were then designed to amplify each of the starch related genes. Primer specificity was first tested by PCR on genomic DNA. When tested against genomic DNA from three randomly chosen individuals, three pairs of primers for EtGBSSIa and two pairs of primers for EtGBSSIb yielded a single dominant band of the expected size, confirming primer specificity (Figure 15). Similar experiments were performed for primers against the other starch-related genes. Then, using the 44 genomic DNA super pools, PCR amplifications were individually conducted with specific primers for each of the genes (Table 5). PCR products for all starch related genes from each superpool were combined to be one amplicon library, in total 44 libraries were prepared for Illumina sequencing.

Sequencing was performed by Illumina Hiseq_4000. Thirty-seven libraries yielded between 3 and 15 million reads. Seven libraries yielded between 0.8 to 3 million reads. The average read length was 126bp. The base sequence quality scores (Q score) for all pools were at the range of Q28 to Q40 (Q30 means that the base call accuracy is 99.9%, which is considered a benchmark for quality in next-generation sequencing).

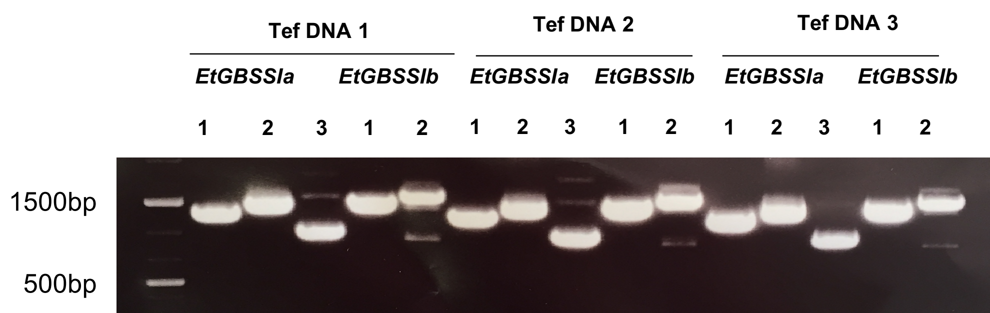


Figure 15. Amplicons of *EtGBSSIa* and *EtGBSSIb* for TILLING by sequencing libraries preparation
Random three superpools (mentioned above) were used as template for PCR amplification, named as Tef DNA1, Tef DNA2 and Tef DNA3. Three pairs of the primers for amplifying the whole *EtGBSSIa* gene (*EtGBSSIa* 1, 2 and 3) were used. Two pairs of the primers for amplifying *EtGBSSIb* gene (*EtGBSSIb* 1 and 2) were used.

Table 5. Primer list for amplification of starch related genes

Gene	Genome (bp)	Primer				Amplicon size (bp)
		Forward	Sequence (5'-3')	Reverse	Sequence (5'-3')	
EtGBSSIa	*3078*	EtGBSSIa FW1	AGCTGTGAGCGCTTCTTCA	EtGBSSIa RV1	TGAAATACGACGCTGAAACG	1276
		EtGBSSIa FW2	CACACCTTGTCGGGGTACTT	EtGBSSIa RV2	GAGGACAACCAGATGCGTTT	1338
		EtGBSSIa FW3	AAACGCATCTGGTTGTCCTC	EtGBSSIa RV3	ACACAGTGCGCTCGTCTTC	943
EtGBSSIb	3183	EtGBSSIb FW1	GCTGCGGTCACACTCTCTTA	EtGBSSIb RV1	CTCGATGGCCTGAAAATCAT	1478
		EtGBSSIb FW2	ATGATTTTCAGGCCATCGAG	EtGBSSIb RV2	GCCAAACCGAATCGTATCTC	1725
EtGBSSIIa	4491	EtGBSSIIa FW1	GTCTTTGCCATGCCCTTAGA	EtGBSSIIa RV1	AATTGCGATGGTTGGGGATC	1398
		EtGBSSIIa FW2	GATCCCCAACCATCGCAATT	EtGBSSIIa RV2	TAACAGGGATGCTCGGATCC	1868
		EtGBSSIIa FW3	CAAGCTGAAGTTGGATTGCC	EtGBSSIIa RV3	CTAGTAACCCTGGCCATGCT	1434
EtGBSSIIb	5761	EtGBSSIIb FW1	ACTGTTGAATGTATACCCGACTT	EtGBSSIIb RV1	AGCAAATGACTGGCACACTG	1890
		EtGBSSIIb FW2	CAGTGTGCCAGTCATTTGCT	EtGBSSIIb RV2	TGTGGATGCTAAGGTATTCTCTT	2008
		EtGBSSIIb FW3	AGAGAATACCTTAGCATCCACA	EtGBSSIIb RV3	ACTACAAACTACTCTCAGCCCC	712
		EtGBSSIIb FW4	GGGGCTGAGAGTAGTTTGTAGT	EtGBSSIIb RV4	TGCCATGCCCTTAGAACCAT	1305
EtSBEIa	3912	EtSBEIa FW1	CCACCAAGGGCCATTGTGAT	EtSBEIa RV1	GGGGAATGAGGTAAGATCTTGC	774
		EtSBEIa FW2	AGTAACACCAGGGTCCATGG	EtSBEIa RV2	GAGCTCCCTACGATGGTGTT	1733
		EtSBEIa FW3	AACACCATCGTAGGGAGCTC	EtSBEIa RV3	TGATCGCTACGTTGTGCTTC	978
EtSBEIb	5651	EtSBEIb FW1	GGCGCTTTAGACATTTTGCA	EtSBEIb RV1	GACACCATCGTAGGGAGCTC	1115
		EtSBEIb FW2	GAGCTCCCTACGATGGTGTC	EtSBEIb RV2	GGCCATCACCTCACATGTTT	1513
		EtSBEIb FW3	AAACATGTGAGGTGATGGCC	EtSBEIb RV3	AATGGACCAGTTACCCCTCA	1846
		EtSBEIb FW4	TGAGGGGTAACTGGTCCATT	EtSBEIb RV4	ACGAGCCAGTCCCTATTGAG	768
EtSBEIc	2380	EtSBEIc FW1	CCATCAGTCCTCTACGAGCT	EtSBEIc RV1	TGCTACCGTGTTACTTGCAC	1813
		EtSBEIc FW2	GTGCAAGTAACACGGTAGCA	EtSBEIc RV2	CACCTGGCTCTCTCTCTCTG	1610
EtSBEIIa	8166	EtSBEIIa FW1	AAGCTTTCCATACGAATTGC	EtSBEIIa RV1	AGTTCGAGGCCAGCAAGAAT	1508
		EtSBEIIa FW2	TCTTGCTGGCCTCGAACTAT	EtSBEIIa RV2	GTATGCTCCAATCACCTCACA	1745
		EtSBEIIa FW3	GTGAGGTGATTGGAGCATACC	EtSBEIIa RV3	AGATAGTCTGCGTCACCCTG	1657
		EtSBEIIa FW4	CCAGGGTGACGCAGACTATC	EtSBEIIa RV4	CACCAGGTTGTATTCGCCAA	1701
EtSBEIIb	10587	EtSBEIIb FW1	GCTAGGAGCCTGACAACGAT	EtSBEIIb RV1	TTCACGGATAGCAAAGCACC	1428
		EtSBEIIb FW2	GGTGCTTTGCTATCCGTGAA	EtSBEIIb RV2	TTCCGGTACTGAGAAGTCCC	1207
		EtSBEIIb FW3	ACCGGAACCCTGATTGTCTA	EtSBEIIb RV3	ACACACTTCTCTAGCCACCT	1914
		EtSBEIIb FW4	AGGTGGCTAGAGAAGTGTGT	EtSBEIIb RV4	CAATAGTGTACCCAACGGCT	1935
		EtSBEIIb FW5	AGTTGAGAGCCAGAAGACTCA	EtSBEIIb RV5	CGCCAAACTTACGGTAGTGA	1298
		EtSBEIIb FW6	CTCAGACGCTGGACTCTTTG	EtSBEIIb RV6	AAGTTGACCGCACATGCTTA	605

Table 5. (Continued) Primer list for amplification of starch related genes

Gene	Genome (bp)	Primer				Amplicon size (bp)
		Forward	Sequence (5'-3')	Reverse	Sequence (5'-3')	
EtSSIa	5755	EtSSIa FW1	AACACTCGTCTTTACCACCCT	EtSSIa RV1	CCCACCCAATCAACGGAATC	1436
		EtSSIa FW2	GATTCCGTTGATTGGGTGGG	EtSSIa RV2	ATGGGTGAACTTTCTTGCCA	1553
		EtSSIaFW3	TTGGCAAGAAAGTTCACCCA	EtSSIa RV3	ACTTACCCTGTACCCTGCTC	1865
		EtSSIa FW4	GAGCAGGGTACAGGGTAAGT	EtSSIa RV4	TCAAAGAGGTGCCACATGTG	671
EtSSIb	8214	EtSSIb FW1	TGTGGCTAGTGAGGACGAAG	EtSSIb RV1	ACTCTCTGACATATCCTAAGGCA	1612
		EtSSIb FW2	TGCCTTAGGATATGTCAGAGAGT	EtSSIb RV2	CTGCATGAGATCCGGGATGA	1310
		EtSSIb FW3	CCCGGATCTCATGCAGGAT	EtSSIb RV3	ACTCTCTGACATATCCTAAGGCA	1404
		EtSSIb FW4	CCCGGATCTCATGCAGGAT	EtSSIb RV4	AGTCTTTTGTCATTCCCCGC	1743
EtSSIIa	5230	EtSSIIa FW1	CGACACGGCTGACAATGTAG	EtSSIIa RV1	TGCAGGAAAATGGGGAATCAC	1239
		EtSSIIa FW2	GTGATTCCCCATTTTCCTGCA	EtSSIIa RV2	CTCCAGCTCTCCCCGTAGTT	1739
		EtSSIIa FW3	GCTACACCAACTACTCCCTG	EtSSIIa RV3	ACAAGCTGGTTATGTCCGC	1658
		EtSSIIa FW4	GCGGACATAACCAGCTTGT	EtSSIIa RV4	AGTACTGGAATTCACGAGGCA	1199
EtSSIIb	4140	EtSSIIb FW1	CCTAAAATTCCCTCGCGCG	EtSSIIb RV1	CGCTTCAACACATCCTGCAT	1864
		EtSSIIb FW2	ATGCAGGATGTGTTGAAGCG	EtSSIIb RV2	TTGCCCGTGTCCAGTGTCT	1741
		EtSSIIb FW3	AGACACTGGACACGGGCAA	EtSSIIb RV3	GGGGTTTCACTCGACTCGA	1168
EtSSIIc	9383	EtSSIIc FW1	GACTGTCCCTACTCCCCTTC	EtSSIIc RV1	CTTGCTCGTTATGGATAGAACTG	1629
		EtSSIIc FW2	CAGTTCTATCCATAACGAGCAAG	EtSSIIc RV2	GCCTGCTAAGTGAGAAAAGGG	1834
		EtSSIIc FW3	CCCTTTTCTCACTTAGCAGGC	EtSSIIc RV3	ACAGCCGCTTTGCACAAC	1959
		EtSSIIc FW4	AAGCGGCTGTAGAGGTAATC	EtSSIIc RV4	TGAAGTCGTCTATTGGGCCA	2105
		EtSSIIc FW5	TGGCCCAATAGACGACTTCA	EtSSIIc RV5	CCATGAGATCTGAAGCAGCC	1045
EtSSIVa	3127	EtSSIVa FW1	CGTATTGTGAAGCGGACGAAT	EtSSIVa RV1	TTTTCAGTTAGGGTGAGGCTT	1830
		EtSSIVa FW2	AAGCCTCACCCTAACTGAAAA	EtSSIVa RV2	ACATCTAACAGGCGTAACGTTG	1530
EtSSIVb	1009	EtSSIVb FW1	CACTCTTCATCCGCCAGTTC	EtSSIVb RV1	TGCTCCAGTGATTGTTTCTCA	1956
		EtSSIVb FW2	TGAGAAACAATCACTGGAGCA	EtSSIVb RV2	GCAACTTTCTGGGTGATAGCA	2133
EtSSIVc	6516	EtSSIVc FW1	CAACTGCCTGCGTTCATG	EtSSIVc RV1	CATGTGGTATTCGAGCTGGTC	1713
		EtSSIVc FW2	GGGACCAGCTCGAATACCAC	EtSSIVc RV2	GAGTCTTGTCAGCTAAGTAGTGT	1862
		EtSSIVc FW3	ACACTACTTAGCTGACAAGACTC	EtSSIVc RV3	GTCAGATGCAGAGACCGGT	1859
		EtSSIVc FW4	ACCGGTCTCTGCATCTGAC	EtSSIVc RV4	GTGAGTTTGGCCTTCCTCCA	1379
EtPTSTa	3918	EtPTSTa FW1	CTTCGGTGAGGTTTTAGGCC	EtPTSTa RV1	GTTTGTCCCACAGTAACCCA	1082
		EtPTSTa FW2	TCGATATGGGTTACTGTGGGA	EtPTSTa RV2	GCTATGGGTGTCTCAGTTCG	1616
		EtPTSTa FW3	CGAACTGAGACACCCATAGC	EtPTSTa RV3	GAGCACAACTTCGCATGTC	1020
EtPTSTb	1305	EtPTSTb FW	CATTGATGTCCACAGGCTCT	EtPTSTb RV	GCCAACTGAGGGGTGTATTTC	1039

EtESVIa	4350	EtESVIa FW1	AGTTGGGTATACTCCAGGTGG	EtESVIa RV1	TGCGAGAACTTGGCGTATAC	1917
		EtESVIa FW2	GTATACGCCAAGTTCTCGCA	EtESVIa RV2	TCTGTTACCTCTACAACGTGC	1347
		EtESVIa FW3	GTAATGCCACGCTTGTTTAGAT	EtESVIa RV3	TGCCTGGCCATCCCTTTA	1411
EtESVIb	6592	EtESVIb FW1	ACAAATTGACGCAGGAATGC	EtESVIb RV1	GGTTAAACGAGCAGTCCTGG	1744
		EtESVIb FW2	CCAGGACTGCTCGTTTAACC	EtESVIb RV2	ACAGGATTTCTACCCGCGG	2282
		EtESVIb FW3	TGGACTAGCTTCTGTCTAAAGC	EtESVIb RV3	CCAGTGCTAGCCCCAATC	3542
EtLESVa	2714	EtLESVa FW	CTAAACCAGTGGAAGGCAGT	EtLESVa RV	TTGGAACTGACCGGAGCT	1930
EtLESVb	10856	EtLESVb FW1	AAGTTTATGCTCTGACATCGG	EtLESVb RV1	GCAGTGATAGAACCCGAAGT	2038
		EtLESVb FW2	ACTTCGGGTTCTATCACTGC	EtLESVb RV2	TTGATCATGTGGGCAGGG	1533
		EtLESVb FW3	CCCTGCCCACATGATCAA	EtLESVb RV3	TGTAGGAGCTGGTGATAGGG	1246

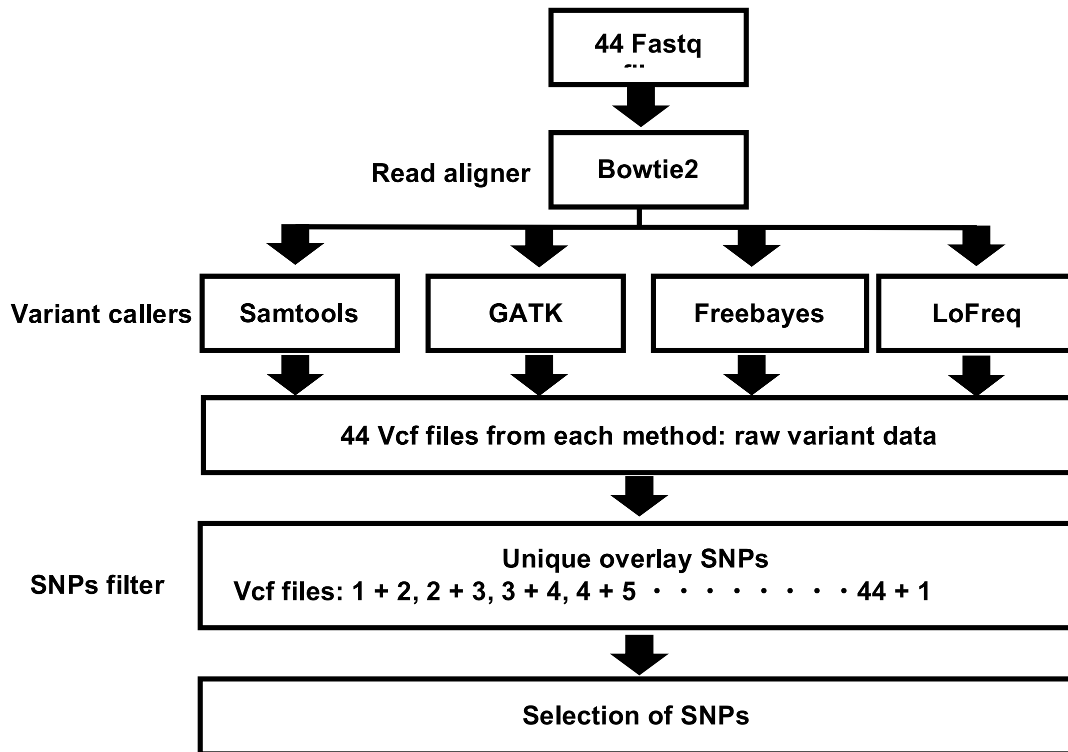


Figure 16. TILLING by sequencing bioinformatics with comparison of variant calling pipelines

Bioinformatics analysis for variant calling

To identify the single nucleotide polymorphisms (SNPs) from sequencing pools, an optimum variant caller for our data analysis was required. Thus, we compared recent, popular variant callers: Samtools (Li et al. 2009), GATK (Van der Auwera et al. 2013), Freebayes (Garrison and Marth 2012) and LoFreq (Wilm et al. 2012)(Figure 16). Using the 44 fastq files generated from sequencing, short reads from each pool were aligned to the reference (containing the genome sequence of all starch related tef genes) by Bowtie2 aligner. Between 83% and 90% of total reads were mapped. A large number of SNPs were generated from each pipeline, which might contain amount of false positive SNPs induced by different tool errors or sequencing errors. To compare the accuracy and efficiency of the four variant callers, the known mutations in GBSSI obtained from CEL1-based TILLING were used as controls. Freebayes and Samtools pipelines identified the control mutants with similar allele frequency (AF), but they were not called correctly by GATK and LoFreq variant callers (Table 6). Samtools called a larger number of mutations overall, suggesting that it may be more prone to false positive identification. Thus, of the four methods, Freebayes seems both sensitive enough to detect non-reference nucleotides and reliable enough to distinguish them from sequencing errors (Table 6). Freebayes called a total of 29736 variants in our data out of which 27529 were single nucleotide polymorphisms (SNPs). We then used a bottom up approach to filter the SNPs. This

means that our filtering criteria in Freebayes (including mean depth, mapping quality, read direction and allele frequency) was set based on the known SNPs from TILLING by CEL1 based method. The same filtering criteria for control SNPs was then applied to all 27529 SNPs.

Table 6. Allele frequency (AF) of unique overlay SNPs in variant calling pipelines

Control SNPs (CEL1 method-mutation No.)	**Control SNPs** (Barcode library-Gene-Mutant position)	**Samtools**	**GATK**	**Freebayes**	**LoFreq**
EtGBSSIa-2	2-*EtGBSSIa*-1375	C>T, 0.84%	C>T, 2.5%	C>T, 0.89%	N
EtGBSSIa-2	3-*EtGBSSIa*-1375	N	N	L	N
EtGBSSIa-1	7-*EtGBSSIa*-1726	C>T, 0.56%	C>T, 2.3%	C>T, 0.60%	N
EtGBSSIa-1	8-*EtGBSSIa*-1726	C>T, 0.48%	N	C>T, 0.40%	N
EtGBSSIb-3	9-*EtGBSSIb*-910	C>T, 0.23%	N	C>T, 0.23%	N
EtGBSSIb-3	10-*EtGBSSIb*-910	C>T, 0.33%	N	C>T, 0.33%	N
EtGBSSIb-2	13-*EtGBSSIb*-1367	C>T, 0.09%	N	L	N
EtGBSSIb-2	14-*EtGBSSIb*-1367	C>T, 0.37%	N	C>T, 0.31%	N

Note: “N” represents no detectable SNP; “L” represents SNP with low frequency

This filtering was needed to distinguish between false SNPs coming from sequencing and PCR errors compared to those from real variant alleles. Based on our pooling strategy, a true SNP from EMS mutagenesis will be present in two adjacent sequencing libraries. Therefore, SNPs were filtered with the criteria of appearance in adjacent libraries and not others. This allowed the removal of 90% of the SNPs as likely false positives. The 10% of the SNPs that remained (named as the unique overlay SNPs) were retained if they had an AF between 0.02% to 1%. In theory, if the amplicons were mixed in equal amounts, a single copy gene should have an AF of 0.26% (1/384 × 100%) for heterozygous mutants and 0.52% (1/192 × 100%) for homozygous mutants. In reality, these SNP frequencies were close to sequencing noise (e.g. machine errors) and there was deviation from the theoretical value, presumably because of uneven amplicon pooling. This made true SNP identification difficult, based on the theoretical AF. Future approaches should use a lower pooling, or a sequencing method with the lowest possible error rate.

To facilitate the management and analysis of the large-scale dataset, an interactive database platform was created, where the multiple analysis functions could be flexibly applied (Figure 17). We included the option of selection from 44 sequencing pools, the option of selection from all genes, the indications of SNP frequency, location and nucleotide change and the option of applying different AF thresholds. The information of sequencing depth and the gene exon-

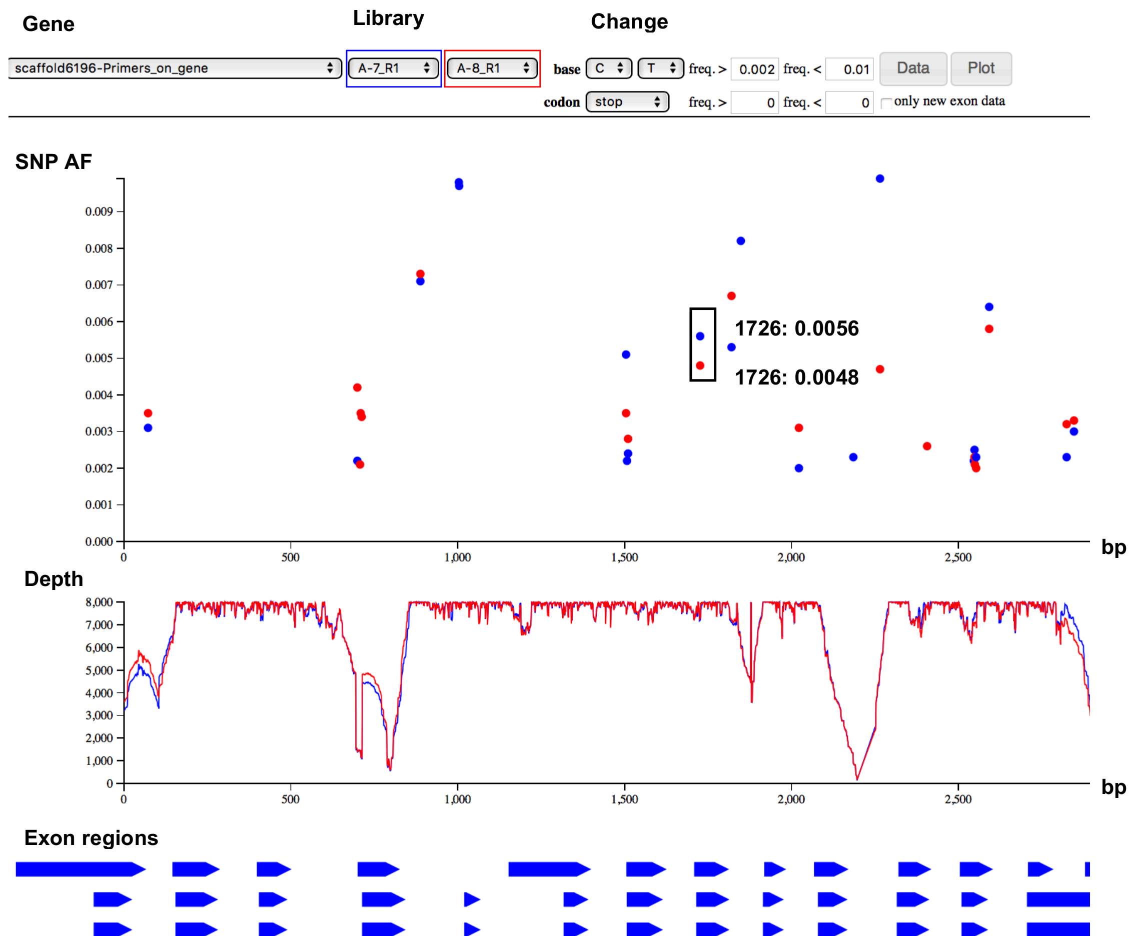


Figure 17. Visualized window of the interactive SNP database management platform

The SNPs with AF are visualized as colored dots in the middle window, where the blue dots are from sequencing library 7 and the red dots are from sequencing library 8, with the x-axes indicating the length of the gene and the y-axes indicating the AF. Two SNPs marked in black square are at the gene position 1726bp and have AFs from each library. The window below is shown the sequencing depth. At the bottom of the window, the coding region are displayed and three channels are listed due to the different annotation possibilities. At the top of the window, multiple analysis functions are available. These includes selection of the gene, the different sequencing library, the base change, the AF and the stop codon function.

intron structure are displayed. An example of the output of our database for the *EtGBSSIa* gene is shown in Figure 17. The barcode libraries 7 (blue) and 8 (red) (both containing DNA from the original DNA plate 2-7-673) are selected and C > T SNPs are displayed as colored filled

dots with AF range between 0.02% and 1%. Using the sequencing depth and coding region information shown below as a guide, certain SNPs were selected for further analysis. For instance, SNPs at position 1726bp are the positive SNPs. An additional function implemented in our interactive database is the prediction of potential stop codons were also possible in our platform. With this tool it is easy to assess whether a promising SNP introduces a stop codon. The SNPs displayed in Figure 17 were generated from Samtools pipeline, indicating the amounts of false positive background SNPs.

For each gene, SNP-calling through Freebayes pipeline were listed, with SNP site, nucleotide change, exon/intron and overlay plate. Overall, more than 500 high-confidence SNPs were detected. An example is shown here only for *GBSSI* gene.

HRM and Sanger sequencing for individual mutant identification

After variant calling, it is necessary to identify the individual mutant from amongst the 96 plants represented in one overlay plate. High-resolution melt (HRM) analysis and Sanger sequencing were used. HRM analysis is becoming a popular method for scanning large numbers of samples for genetic variants. It is a PCR based method which monitors the double strand DNA-binding of a fluorescent dye. Mutations in HRM analysis are detected as homozygotes or heterozygotes. Homozygous variation is characterized by the melting temperature shift and heterozygous variation is characterized by changes in melting curve shape (Taylor et al. 2009). Realistically, the HRM melting curves of SNPs which are homozygous mutants were more difficult than heterozygous to be distinguished from the curves of wild types. In this case, Sanger sequencing was used. The identification of individual mutants started from the *GBSS1* mutants (Figure 18) and this part is still going on. Besides, the work will be continued for identifying individuals on other starch related genes.

For instance, one of the control SNP (C > T) at position 910 on *EtGBSSIb* identified via TILLING-by-sequencing was located in plate 2-7-948. The HRM curves using A8 individual DNA (plant name 2-4-791) differed to the curves using other DNAs on plate 2-7-948 (Figure 18). This indicated A8 DNA from plate 2-7-948 carried the SNP and it was a heterozygous mutation according to the shape change of the melting curve. This mutation was confirmed by Sanger sequencing.

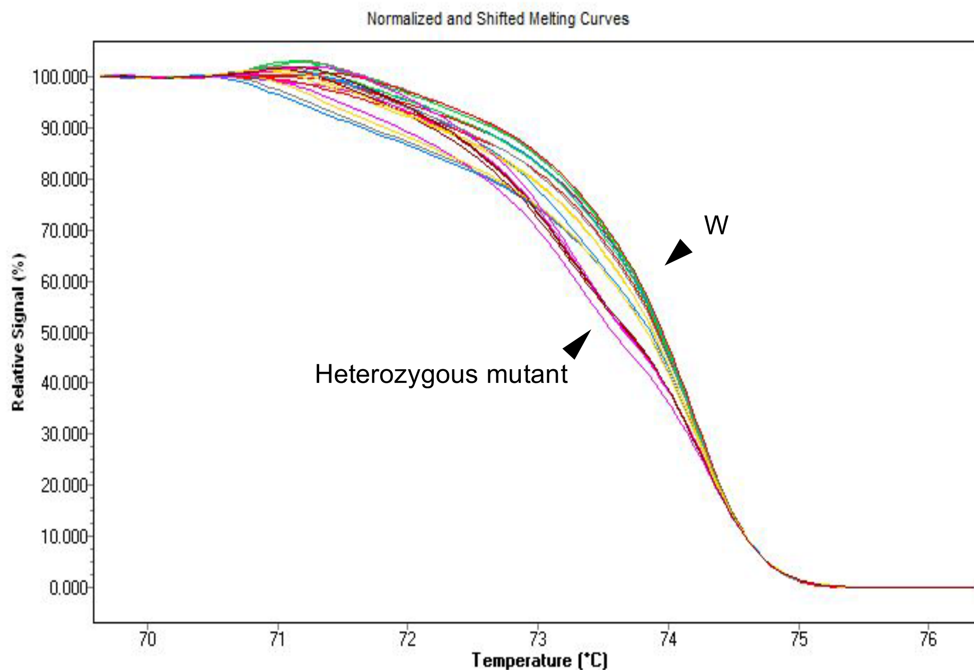


Figure 18. HRM curve of mutated SNPs and WT
The aligned melt curves are plotted as relative signal (fluorescence) vs. temperature within a pre- and post- melt region (70°C to 76°C). The plot demonstrates the sharp decrease in fluorescence when the double-stranded DNA melts into its single-stranded form. Primers were designed to amplify target regions of 100 bp. Curve colors represent different samples. The amount of 20 ng and 10 ng plate 2-7-948 A8 DNA (carried SNP identified from TILLING by CEL1 method) were used as templates. The amount of 20 ng and 20 ng four random DNAs in plate 2-7-948 were used as well. The shape of melting curves changed in heterozygous mutation samples.

Bioinformatics analysis for functional prediction

The 500 SNPs identified through the Freebayes pipeline were filtered for exon-located SNPs, which might potentially lead to protein functional change, using the annotated draft genome (Cannarozzi et al. 2014). The number of the exon-located SNPs for each gene varied, ranging from one in *EtRSRIb* to twenty-four in *EtSSIb* gene (Table 7)

To explore whether these exon-located SNPs would result in potentially useful mutations, informatics prediction tools were used to identify the most promising SNPs. The potential most valuable SNPs for *GBSS1* (containing either significant deleterious change or stop codon) are listed in Table 8. To obtain functional null mutants, mutation in both homeologous genes are needed (e.g. green color highlighted two SNPs). Of the mutations within exons of *EtGBSSIb*,

a G/A transition at position 2401 of the sequence, results in an A417>T change. This substitution is predicted to decrease protein function by I-mutant analysis. This was one of the mutants detected from both TILLING methods. For other mutants, one mutant with a C to A transition at bp position 2502 was silent, causing no change in the protein. One mutant with C to T transition at bp position 2074 caused a premature stop codon at position 378. The other five mutants were predicted to induce amino acid changes: G176>S, I432>F, Q439>H, I475>S and A484>S. The predicted effect of these substitutions on protein function by I-mutant and PROVEN differed (software tools which predict whether an amino acid substitution or indel has an impact on the biological function of a protein). For the mutants in *EtGBSSIa*, two were predicted to have neutral change and one resulted in a missense change (L 89 H).

The functional predictions for the SNPs in the other starch related genes is ongoing and not yet complete. Table 9 shows the stop codons for some of the starch related genes I have obtained so far that should be analyzed further.

Table 7. Prediction of SNPs located at exon region of each gene

Tef gene	Number of SNPs
EtGBSSIa	3
EtGBSSIb	8
EtGBSSIIa	7
EtGBSSIIb	7
EtSBEIa	4
EtSBEIb	12
EtSBEIc	3
EtSBEIIa	11
EtSBEIIb	11
EtSSIa	12
EtSSIb	24
EtSSIIa	24
EtSSIIb	15
EtSSIIc	6
EtSSIId	4
EtSSIVa	1
EtSSIVb	n.d.
EtSSIVc	9
EtPTSTa	2
EtPTSTb	n.d.
EtESVIa	4
EtESVIb	n.d.
EtLESVa	n.d.
EtLESVb	n.d.
EtRSRIa	n.d.
EtRSRIb	1

Note: n.d. means not annotated.

Table 8. Prediction of functional change in *EtGBSSIa* and *EtGBSSIb* mutants

Gene	Position	Change	Overlay plate	Prediction for functional change		
				Amino acid change	I-mutant score	PROVEN score
ErteGBSSIa	128	g to t	2-7-1686	72 E TO D	decrease -0.35, RI 3	-1.368 neutral
	178	t to a	2-7-4853	89 L TO H	decrease -1.62, RI7	-3.624 deleterious
	1175	g to t	2-7-4457	236 R TO M	decrease -0.4, RI 5	0.103 neutral
ErteGBSSIb	414	g to a	2-7-3113	176 G TO S	decrease -1.23, RI 6	-0.2764 deleterious
	2074	c to t	2-7-1	378 R TO Stop		
	2401	g to a	2-7-0113	417 A TO T	decrease -0.44, RI 2	-0.951 neutral
	2446	a to t	2-7-4649	432 I TO F	decrease -1.26, RI 8	-1.207 neutral
	2469	g to t	2-7-3432	439 Q TO H	decrease -0.85, RI 7	-0.714 neutral
	2502	c to a	2-7-4952	no change		
	2692	t to g	2-7-2587	475 I TO S	decrease -1.57, RI 7	-0.839 neutral
	2718	g to t	2-7-2700	484 A TO S	decrease -0.75, RI 8	-0.920 neutral

Note: Green highlighted mutations were used for achieving potential null mutant (listed in Table 9).

Table 9. SNPs for achieving potential null mutants

Gene	Position	SNP base change	Overlay plate	Functional change prediction (PROVEN score or stop codon)
EtGBSSIa	178	t to a	2-7-4853	-3.624 deleterious
EtGBSSIb	2074	c to t	2-7-1	378 R to stop
EtSBEIIa	582	c to a	2-7-4155	-6.947 Deleterious
EtSBEIb	1871	g to t	2-7-948	G486STOP
EtSSIIa	2689	g to t	2-7-1033	E483STOP
EtSSIb	2482	g to t	2-7-337	G156STOP
EtSSIb	2700	a to t	2-7-3536	K186STOP
EtSSIb	4834	c to t	2-7-1	R354STOP
EtSSIa	1458	g to t	2-7-786	E87STOP
EtSSIIc	2904	c to a	2-7-244	S127STOP
EtGBSSIIb	5663	a to t	2-7-4649	C33STOP
EtSSIVc	2318	c to a	2-7-337	E613STOP

3.5 Discussion

Tef, as an orphan allotetraploid crop, has a relative paucity of genomic and bioinformatics resources, which brings difficulties in identifying our starch related genes of interested. In previous studies, cross-species identification approach has been utilized to assist non-model plant research, and also for understanding phylogenetic relationships, annotating genes, and detecting functional regulatory elements (Altschul et al. 1990; Alexandersson et al. 2003; Zhu and Buell 2007; Menlove et al. 2009; Mun et al. 2009). Thus, by using a comparative sequence analysis strategy based on the two closely related species rice and sorghum we successfully identified starch-related genes encoding enzyme the isoforms GBSSI, GBSSII, SBEI, SBEIIb, SSI, SSIIa, SSIIb, SSIIc, SSIVa, SSIVb, PTST, ESVI, LESV and RSRI in tef. In most cases, we were able to identify homologous copies for each gene. However, genes encoding for enzymes such as SBEI, SSII and SSIV had equivocal homeologs (Table 1). This might be due to the uncertainty in genome assembly and sequence alignment (Wong et al. 2008) or might reflect true genetic differences in each species (e.g. gene loss or mutation in tef). To increase the gene identification accuracy, phylogenetic trees of orthologs and genome syntenic comparisons were applied and the genes most homologous to those of rice and sorghum stood out (*EtSBEIb* for *SBEI* gene and *EtSSIVc* for *SSIV* gene; Figures 2 and 3).

The role of the identified genes in endosperm starch biosynthesis was further supported by transcriptome analysis, where the expression profile of each gene during seed development was determined, revealing the predominant isoforms. *EtGBSSIa, EtGBSSIb, EtSBEIb, EtSBEIc, EtSBEIIa, EtSBEIIb*, *EtSSIa* and *EtSSIb* were all highly expressed during the whole tef seed development (Figure 8). This is highly congruent with the findings in rice, where genes encoding the isoforms of GBSS, SBE and SSs, together with AGPase, DBE, PHO and DPE have been highlighted to be the significant genes in starch biosynthesis during seed development through transcriptome sequencing studies (Agarwal et al. 2011; Sreenivasulu and Wobus 2013). Seed development is a vitally important process for cereal yield and it undergoes complex molecular and tissue reprogramming. The regulatory network controlling starch accumulation in cereals is far from understood and in tef it has not been studied yet. It is likely that there is a common mechanism in the different cereals. From our result, the importance of our targeted starch genes in rapid starch filling stage is reflected by their highest expression (peak RPKM value) in the middle of development, and decreasing expression when reaching seed maturation and desiccation (Figure 8).

EtGBSSIa and *EtGBSSIb* have been identified as two obvious genes in tef encoding GBSSI isoforms and both are confirmed to be equally present through transcriptome analysis (Figure 8). These were prioritized as gene targets since, through modifying the activity of one or both isozymes, amylose content (one of the most important functional properties of starch) is predicted be to altered (Tian et al. 2009). In wheat, a *waxy* cultivar with almost no amylose was produced through TILLING (Slade et al. 2005). Our initial use of CEL1-based TILLING to find tef mutants with alterations in either *EtGBSSIa* or *EtGBSSIb* yielded five mutants for

EtGBSSIa and three for EtGBSSIb after screening ~2112 and ~2880 mutant lines, respectively. However, only one of them has predicated amino acid change on EtGBSSIa, which is not enough to achieve any null mutant for GBSS1 gene. So far, we targeted the gene regions having high mutation possibility for CEL1-based TILLING and at the rest of the whole gene, the valuable mutations were not detected via this method. A better system was demanded to include the whole region of a gene, to target more genes and to analyze the large population. Hence, we tried a new TILLING approach based on high trough put sequencing and genome data analysis.

An original TILLING-by-sequencing approach was devised to use the existing resources developed at the University of Bern. Via this approach, we developed and a database with a large number of SNPs in all targeted starch genes. The success of our TILLING-by-sequencing method was confirmed by the identification of SNPs previously identified by the conventional tilling method (Table 6). With a large number of SNP, it is necessary to prioritise and the prediction of functional changes using bioinformatics approaches, is a good place to start. The amino acid changes and predicted effects are listed in Table 7. No phenotypic change, amylose content and GBSSI protein expression difference have been detected in our tef lines as yet (not shown). However, this is expected because a) some mutations are heterozygous, b) some SNPs cause silent mutations, c) most SNPs cause missense mutations, most of which are unlikely to abolish protein function, and d) there is very likely redundancy by other wild-type copies of homeologous genes due to the tetraploid of tef. The work to screening for GBSSI null mutants and mutants in other genes using TILLING-by-sequencing is ongoing.

As with other population genetic studies, our pooled sequencing (pool-seq) creates complexity in data analysis. Sequencing noise and potential false-positive SNPs are caused by sequencing and/or PCR errors. This is compounded by differences in AF caused by the fact that the concentration of each individual's DNA in the pools might not always be equimolar. This variation in true AF and background noise observed in our pool-seq, lead to the difficulty for accurate estimation of rare variants. Many false positive SNPs were identified from our experiment, and in hindsight, a lower degree of pooling would have facilitated SNP-calling, even though it would have increased the sequencing costs.

Low frequency SNPs identification (rare disease) are frequently studied in clinical contexts. Several related tools such as Kolmogorov-Smirnov and CRISP filters, based on different statistical tests have been developed to remove spurious variants (Anand et al. 2016). Studies have been conducted to compare different variant calling pipelines. The substantial disagreement among variant calls made by different pipelines suggesting a need for more cautious interpretation of called variants in genomic medicine (Hwang et al. 2015). The same is true for our project on tef; an accurate variant estimation pipeline is also needed since the next steps in mutation validation are both time- and resource-consuming. By starting the variant calling pipelines and filtering tools developed in clinical medicine, we modified and applied them to the data from our tef TILLING population. We found that

Freebayes was the most suitable variant caller for SNP identification in all targeted starch genes (Table 6).

Another issue that we encountered was that the relatively short-read lengths generated by the Illumina sequencing technologies present a particular problem for larger, polyploid genomes with highly repetitive and homologous genomic regions. Thus, as well as optimal bioinformatics approaches, other method for sequence data acquisition (e.g. Pacbio longer-read sequencing; Rhoads and Au 2015) could be beneficial, as could alternatives to PCR amplification of target sequences (e.g. Exome capture; Henry et al. 2014). Finally, at the same time that this work was conducted, genome editing tools have emerged (i.e. CRISPR-Cas9; Ran et al. 2013) that hold great promise for simultaneously targeting multiple copies of the same gene in polyploidy species.

In conclusion, we have successfully identified key starch related genes in tef and determined their expression during the seed development. We performed conventional TILLING using CEL1 and developed a novel TILLING-by-sequencing method for population-wide mutation detection. Hopefully in the future, useful mutants identified here can be used to create new tef varieties with improved starch properties, opening the door for the alterative used in food and non-food industry, and a higher value for the crop.

3.6 References

Abebe W, Collar C, Ronda F (2015) Impact of variety type and particle size distribution on starch enzymatic hydrolysis and functional properties of tef flours. Carbohydr Polym 115:260–268. doi: 10.1016/j.carbpol.2014.08.080

Agarwal P, Kapoor S, Tyagi AK (2011) Transcription factors regulating the progression of monocot and dicot seed development. BioEssays 33:189–202.

Ahamed NT, Singhal RS, Kulkarni PR, Pal M (1996) Physicochemical and functional properties of Chenopodium quinoa starch. Carbohydr Polym 31:99–103. doi: 10.1016/S0144-8617(96)00034-3

Alcázar-Alay SC, Meireles MAA (2015) Physicochemical properties, modifications and applications of starches from different botanical sources. Food Sci Technol 35:215–236. doi: 10.1590/1678-457X.6749

Alexandersson M, Cawley S, Pachter L (2003) SLAM: Cross-species gene finding and alignment with a generalized pair hidden Markov model. Genome Res 13:496–502. doi: 10.1101/gr.424203

Allem AC, Genéticos R (2002) The Origins and Taxonomy of Cassava. Cassava Biol Prod Util 1–16. doi: 10.1080/21622965.2014.897905

Altschul SF, Gish W, Miller W, et al (1990) Basic local alignment search tool. J Mol Biol 215:403–10. doi: 10.1016/S0022-2836(05)80360-2

Alves a. a. C (2002) Cassava botany and physiology. Cassava Biol Prod Util 67–89. doi: 10.1079/9780851995243.0000

Anand S, Mangano E, Barizzone N, et al (2016) Next generation sequencing of pooled samples: Guideline for variants' filtering. Sci Rep. doi: 10.1038/srep33735

Armstead I, Huang L, Ravagnani A, et al (2009) Bioinformatics in the orphan crops. Brief. Bioinform. 10:645–653.

Ashogbon AO, Akintayo ET (2014) Recent trend in the physical and chemical modification of starches from different botanical sources: A review. Starch/Staerke 66:41–57.

Aslam U, Cheema HMN, Ahmad S, et al (2016) COTIP: Cotton TILLING Platform, a Resource for Plant Improvement and Reverse Genetic Studies. Front Plant Sci. doi: 10.3389/fpls.2016.01863

Assefa K, Chanyalew S, Tadele Z (2017) Tef, Eragrostis tef (Zucc.) Trotter. In: Millets and Sorghum: Biology and Genetic Improvement. pp 226–266

Assefa K, Yu J-K, Zeid M, et al (2011) Breeding tef [Eragrostis tef (Zucc.) trotter]: conventional and molecular approaches. Plant Breed 130:1–9. doi: 10.1111/j.1439-0523.2010.01782.x

Avni A, Blázquez MA (2011) Can plant biotechnology help in solving our food and energy shortage in the future? Curr. Opin. Biotechnol. 22:220–223.

Balagopalan C (2002) Cassava Utilization in Food, Feed and Industry. Cassava Biol Prod Util 301–318. doi: 10.1079/9780851995243.0000

Barkley N, Wang M (2008) Application of TILLING and EcoTILLING as Reverse Genetic Approaches to Elucidate the Function of Genes in Plants and Animals. Curr Genomics 9:212–226. doi: 10.2174/138920208784533656

Baunsgaard L, Lütken H, Mikkelsen R, et al (2005) A novel isoform of glucan, water dikinase phosphorylates pre-phosphorylated α-glucans and is involved in starch degradation in Arabidopsis. Plant J 41:595–605. doi: 10.1111/j.1365-313X.2004.02322.x

Blagbrough IS, Bayoumi SAL, Rowan MG, Beeching JR (2010) Cassava: An appraisal of its phytochemistry and its biotechnological prospects. Phytochemistry 71:1940–1951. doi:

10.1016/j.phytochem.2010.09.001
Blazek J, Copeland L (2008) Pasting and swelling properties of wheat flour and starch in relation to amylose content. Carbohydr Polym 71:380–387. doi: 10.1016/j.carbpol.2007.06.010
Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254. doi: 10.1016/0003-2697(76)90527-3
Breuninger WF, Piyachomkwan K, Sriroth K (2009) Tapioca/Cassava Starch: Production and Use. In: Starch. pp 541–568
Bull SE, Owiti JA, Niklaus M, et al (2009) Agrobacterium-mediated transformation of friable embryogenic calli and regeneration of transgenic cassava. Nat Protoc 4:1845–1854. doi: 10.1038/nprot.2009.208
Bultosa G (2016) Teff: Overview, 2nd edn. Elsevier Ltd.
Bultosa G (2007) Physicochemical characteristics of grain and flour in 13 tef [Eragrostis tef (Zucc .) Trotter] grain varieties. J Appl Sci Res 3:2042–2050.
Bultosa G, Hall AN, Taylor JRN (2002) Physico-chemical characterization of grain tef [Eragrostis tef (Zucc.) trotter] starch. Starch/Staerke 54:461–468. doi: 10.1002/1521-379X(200210)54:10<461::AID-STAR461>3.0.CO;2-U
Bultosa G, Taylor JRN (2003) Chemical and physical characterisation of grain tef [Eragrostis tef (Zucc.) Trotter] starch granule composition. Starch/Staerke 55:304–312. doi: 10.1002/star.200390065
Bultosa G, Taylor JRN (2004) Paste and gel properties and in vitro digestibility of Tef [Eragrostis tef (Zucc). Trotter] starch. Starch/Staerke 56:20–28. doi: 10.1002/star.200200191
Burton RA, Jenner H, Carrangis L, et al (2002) Starch granule initiation and growth are altered in barley mutants that lack isoamylase activity. Plant J 31:97–112. doi: 10.1046/j.1365-313X.2002.01339.x
Cannarozzi G, Chanyalew S, Assefa K, et al (2018) Technology generation to dissemination: lessons learned from the tef improvement project. Euphytica 214:31. doi: 10.1007/s10681-018-2115-5
Cannarozzi G, Plaza-Wüthrich S, Esfeld K, et al (2014) Genome and transcriptome sequencing identifies breeding targets in the orphan crop tef (Eragrostis tef). BMC Genomics 15:581. doi: 10.1186/1471-2164-15-581
Carciofi M, Shaik SS, Jensen SL, et al (2011) Hyperphosphorylation of cereal starch. J Cereal Sci 54:339–346. doi: 10.1016/j.jcs.2011.06.013
Carpenter MA, Joyce NI, Genet RA, et al (2015) Starch phosphorylation in potato tubers is influenced by allelic variation in the genes encoding glucan water dikinase, starch branching enzymes I and II, and starch synthase III. Front Plant Sci 6:143. doi: 10.3389/fpls.2015.00143
Caspar T, Lin TP, Kakefuda G, et al (1991) Mutants of Arabidopsis with altered regulation of starch degradation. Plant Physiol 95:1181–1188. doi: 10.1104/pp.95.4.1181
Ceballos H, Sánchez T, Morante N, et al (2007) Discovery of an amylose-free starch mutant in cassava (Manihot esculenta Crantz). J Agric Food Chem 55:7469–7476. doi: 10.1021/jf070633y
Chan HT, Bhat R, Karim AA (2009) Physicochemical and functional properties of ozone-oxidized starch. J Agric Food Chem 57:5965–5970. doi: 10.1021/jf9008789
Chavarriaga-Aguirre P, Brand A, Medina A, et al (2016) The potential of using biotechnology to improve cassava: a review. Vitr. Cell. Dev. Biol. - Plant 52:461–478.
Chetty CC, Rossin CB, Gruissem W, et al (2013) Empowering biotechnology in southern Africa: Establishment of a robust transformation platform for the production of

transgenic industry-preferred cassava. N Biotechnol 30:136–143. doi: 10.1016/j.nbt.2012.04.006
Chiu C wai, Solarek D (2009) Modification of Starches. In: Starch. pp 629–655
Colbert T (2001) High-Throughput Screening for Induced Point Mutations. PLANT Physiol 126:480–484. doi: 10.1104/pp.126.2.480
Craig SAS, Maningat CC, Seib PA, Hoseney RC (1989) Starch paste clarity. Cereal Chem 66:173–182.
Critchley JH, Zeeman SC, Takaha T, et al (2001) A critical role for disproportionating enzyme in starch breakdown is revealed by a knock-out mutation in Arabidopsis. Plant J 26:89–100. doi: 10.1046/j.1365-313X.2001.01012.x
Crumpton-Taylor M, Pike M, Lu KJ, et al (2013) Starch synthase 4 is essential for coordination of starch granule formation with chloroplast division during Arabidopsis leaf expansion. New Phytol 200:1064–1075. doi: 10.1111/nph.12455
Deffenbaugh LB, Walker CE (1989) Use of the Rapid-Visco-Analyzer to Measure Starch Pasting Properties. Part I: Effect of Sugars. Starch - Stärke 41:461–467. doi: 10.1002/star.19890411205
Delatte T, Umhang M, Trevisan M, et al (2006) Evidence for distinct mechanisms of starch granule breakdown in plants. J Biol Chem 281:12050–12059. doi: 10.1074/jbc.M513661200
Donald AM (2004) Understanding starch structure and functionality. In: Starch in Food: Structure, Function and Applications. pp 156–184
Dumez S, Wattebled F, Dauvillee D, et al (2006) Mutants of Arabidopsis Lacking Starch Branching Enzyme II Substitute Plastidial Starch Synthesis by Cytoplasmic Maltose Accumulation. PLANT CELL ONLINE 18:2694–2709. doi: 10.1105/tpc.105.037671
Edner C, Li J, Albrecht T, et al (2007) Glucan, Water Dikinase Activity Stimulates Breakdown of Starch Granules by Plastidial β-Amylases. Plant Physiol 145:17–28. doi: 10.1104/pp.107.104224
El-Sharkawy MA (2004) Cassava biology and physiology. Plant Mol Biol 56:481–501. doi: 10.1007/s11103-005-2270-7
Eliasson AC (2004) Starch in Food: Structure, Function and Applications.
Esfeld K, Uauy C, Tadele Z (2012) Application of TILLING for orphan crop improvement. In: Biotechnology of Neglected and Underutilized Crops. pp 83–113
Falade KO, Akingbala JO (2011) Utilization of Cassava for food. Food Rev Int 27:51–83. doi: 10.1080/87559129.2010.518296
Feike D, Seung D, Graf A, et al (2016) A novel starch-granule-associated protein required for the control of starch degradation in Arabidopsis thaliana leaves. Plant Cell 28:1472–1489. doi: 10.1105/tpc.16.00011
Feltus F a, Singh HP, Lohithaswa HC, et al (2006) A comparative genomics strategy for targeted discovery of single-nucleotide polymorphisms and conserved-noncoding sequences in orphan crops. Plant Physiol 140:1183–1191. doi: 10.1104/pp.105.074203
Fredriksson H, Silverio J, Andersson R, et al (1998) The influence of amylose and amylopectin characteristics on gelatinization and retrogradation properties of different starches. Carbohydr Polym 35:119–134. doi: 10.1016/S0144-8617(97)00247-6
Fu F-F, Xue H-W (2010) Coexpression Analysis Identifies Rice Starch Regulator1, a Rice AP2/EREBP Family Transcription Factor, as a Novel Rice Starch Biosynthesis Regulator. PLANT Physiol 154:927–938. doi: 10.1104/pp.110.159517
Fujita N, Yoshida M, Asakura N, et al (2006) Function and Characterization of Starch Synthase I Using Mutants in Rice. PLANT Physiol 140:1070–1084. doi: 10.1104/pp.105.071845
Fulton DC, Stettler M, Mettler T, et al (2008) -AMYLASE4, a Noncatalytic Protein

Required for Starch Breakdown, Acts Upstream of Three Active -Amylases in Arabidopsis Chloroplasts. PLANT CELL ONLINE 20:1040–1058. doi: 10.1105/tpc.107.056507
Gallant DJ, Bouchet B, Baldwin PM (1997) Microscopy of starch: evidence of a new level of granule organization. Carbohydr Polym 32:177–191. doi: 10.1016/S0144-8617(97)00008-8
Gao M, Fisher DK, Kim K-N, et al (1996) Evolutionary conservation and expression patterns of maize starch branching enzyme I and IIb genes suggests isoform specialization. Plant Mol Biol 30:1223–1232. doi: 10.1007/BF00019554
Garrison E, Marth G (2012) Haplotype-based variant detection from short-read sequencing. 1–9. doi: arXiv:1207.3907 [q-bio.GN]
Gebremariam MM, Zarnkow M, Becker T (2012) Teff (Eragrostis tef) as a raw material for malting, brewing and manufacturing of gluten-free foods and beverages: a review. J. Food Sci. Technol. 51:2881–2895.
Grimaud F, Rogniaux H, James MG, et al (2008) Proteome and phosphoproteome analysis of starch granule-associated proteins from normal maize and mutants affected in starch biosynthesis. J Exp Bot 59:3395–3406. doi: 10.1093/jxb/ern198
Gunorubon AJ, Kekpugile DK (2012) Modification of Cassava Starch for Industrial Uses. Int J Eng Technol 2:913–919.
Hager AS, Wolter A, Jacob F, et al (2012) Nutritional properties and ultra-structure of commercial gluten free flours from different botanical sources compared to wheat flours. J Cereal Sci 56:239–247. doi: 10.1016/j.jcs.2012.06.005
Hannah LC, James M (2008) The complexities of starch biosynthesis in cereal endosperms. Curr Opin Biotechnol 19:160–165. doi: 10.1016/j.copbio.2008.02.013
Hari PK, Garg S, Garg SK (1989) Gelatinization of Starch and Modified Starch. Starch ??? St??rke 41:88–91. doi: 10.1002/star.19890410304
Hennen-Bierwagen TA, Lin Q, Grimaud F, et al (2009) Proteins from Multiple Metabolic Pathways Associate with Starch Biosynthetic Enzymes in High Molecular Weight Complexes: A Model for Regulation of Carbon Allocation in Maize Amyloplasts. PLANT Physiol 149:1541–1559. doi: 10.1104/pp.109.135293
Hennen-Bierwagen TA, Liu F, Marsh RS, et al (2008) Starch Biosynthetic Enzymes from Developing Maize Endosperm Associate in Multisubunit Complexes. PLANT Physiol 146:1892–1908. doi: 10.1104/pp.108.116285
Henry IM, Nagalakshmi U, Lieberman MC, et al (2014) Efficient Genome-Wide Detection and Cataloging of EMS-Induced Mutations Using Exome Capture and Next-Generation Sequencing. Plant Cell 26:1382–1397. doi: 10.1105/tpc.113.121590
Henson (2014) A Comparison of Barley Malt Amylolytic Enzyme Thermostabilities and Wort Sugars Produced During Mashing. J Am Soc Brew Chem. doi: 10.1094/ASBCJ-2014-0130-01
Hizukuri S, Tabata S, Kagoshima, Nikuni Z (1970) Studies on Starch Phosphate Part 1. Estimation of glucose-6-phosphate residues in starch and the presence of other bound phosphate(s). Starch - Stärke 22:338–343. doi: 10.1002/star.19700221004
Hoover R (2001) Composition, molecular structure, and physicochemical properties of tuber and root starches: A review. Carbohydr Polym 45:253–267. doi: 10.1016/S0144-8617(00)00260-5
Hovenkamp-Hermelink JHM, Jacobsen E, Ponstein AS, et al (1987) Isolation of an amylose-free starch mutant of the potato (Solanum tuberosum L.). Theor Appl Genet 75:217–221. doi: 10.1007/BF00249167
Hwang S, Kim E, Lee I, Marcotte EM (2015) Systematic comparison of variant calling pipelines using gold standard personal exome variants. Sci Rep. doi: 10.1038/srep17875

Ihemere U, Arias-Garzon D, Lawrence S, Sayre R (2006) Genetic modification of cassava for enhanced starch production. Plant Biotechnol J 4:453–465. doi: 10.1111/j.1467-7652.2006.00195.x
Imberty A, Chanzy H, Pérez S, et al (1988) The double-helical nature of the crystalline part of A-starch. J Mol Biol 201:365–378. doi: 10.1016/0022-2836(88)90144-1
Imberty A, Perez S (1988) A revisit to the three-dimensional structure of B-type starch. Biopolymers 27:1205–1221. doi: 10.1002/bip.360270803
Jaiswal S, Baïšga M, Ahuja G, et al (2014) Development of barley (Hordeum Vulgare L.) lines with altered starch granule size distribution. J Agric Food Chem 62:2289–2296. doi: 10.1021/jf405424x
James MG, Denyer K, Myers AM (2003) Starch synthesis in the cereal endosperm. Curr. Opin. Plant Biol. 6:215–222.
Jeon JS, Ryoo N, Hahn TR, et al (2010) Starch biosynthesis in cereal endosperm. Plant Physiol. Biochem. 48:383–392.
Jobling S (2004) Improving starch for food and industrial applications. Curr Opin Plant Biol 7:210–218. doi: 10.1016/j.pbi.2003.12.001
Joshi M, Aldred P, McKnight S, et al (2013) Physicochemical and functional characteristics of lentil starch. Carbohydr Polym 92:1484–1496. doi: 10.1016/j.carbpol.2012.10.035
Kamei CLA, Severing EI, Dechesne A, et al (2016) Orphan Crops Browser: a bridge between model and orphan crops. Mol Breed 36:1–18. doi: 10.1007/s11032-015-0430-2
Kammerer B, Fischer K, Hilpert B, et al (1998) Molecular Characterization of a Carbon Transporter in Plastids from Heterotrophic Tissues: The Glucose 6-Phosphate/Phosphate Antiporter. Plant Cell 10:105–117. doi: 10.2307/3870632
Kampfenkel K, Möhlmann T, Batz O, et al (1995) Molecular characterization of an Arabidopsis thaliana cDNA encoding a novel putative adenylate translocator of higher plants. FEBS Lett 374:351–5. doi: 10.1016/0014-5793(95)01143-3
Kaneko M, Itoh H, Ueguchi-Tanaka M, et al (2002) The alpha -Amylase Induction in Endosperm during Rice Seed Germination Is Caused by Gibberellin Synthesized in Epithelium. Plant Physiol 128:1264–1270. doi: 10.1104/pp.010785
Karim AA, Toon LC, Lee VPL, et al (2007) Effects of phosphorus contents on the gelatinization and retrogradation of potato starch. J Food Sci 72:132–138. doi: 10.1111/j.1750-3841.2006.00251.x
Karlström A, Calle F, Salazar S, et al (2016) Biological Implications in Cassava for the Production of Amylose-Free Starch: Impact on Root Yield and Related Traits. Front Plant Sci 7:604. doi: 10.3389/fpls.2016.00604
Kaur B, Ariffin F, Bhat R, Karim AA (2012) Progress in starch modification in the last decade. Food Hydrocoll 26:398–404. doi: 10.1016/j.foodhyd.2011.02.016
Kaur M, Oberoi DPS, Sogi DS, Gill BS (2011) Physicochemical, morphological and pasting properties of acid treated starches from different botanical sources. J Food Sci Technol 48:460–465. doi: 10.1007/s13197-010-0126-x
KIM CS, WALKER CE (1992) Changes in Starch Pasting Properties Due to Sugars and Emulsifiers as Determined by Viscosity Measurement. J Food Sci 57:1009–1013. doi: 10.1111/j.1365-2621.1992.tb14344.x
Knoll JE, Ramos ML, Zeng Y, et al (2011) TILLING for allergen reduction and improvement of quality traits in peanut (Arachis hypogaea L.). BMC Plant Biol. doi: 10.1186/1471-2229-11-81
Kotting O (2005) Identification of a Novel Enzyme Required for Starch Metabolism in Arabidopsis Leaves. The Phosphoglucan, Water Dikinase. PLANT Physiol 137:242–252. doi: 10.1104/pp.104.055954
Kötting O, Santelia D, Edner C, et al (2009a) STARCH-EXCESS4 is a laforin-like

Phosphoglucan phosphatase required for starch degradation in Arabidopsis thaliana. Plant Cell 21:334–46. doi: 10.1105/tpc.108.064360

Kötting O, Santelia D, Edner C, et al (2009b) STARCH-EXCESS4 is a laforin-like Phosphoglucan phosphatase required for starch degradation in Arabidopsis thaliana. Plant Cell 21:334–46. doi: 10.1105/tpc.108.064360

Krishna Radhika N, Sheela MN, Asha Devi A, et al (2014) Genetic modification for designer starch from cassava. J. Trop. Agric. 52:1–6.

Kurakake M, Akiyama Y, Hagiwara H, Komaki T (2009) Effects of cross-linking and low molecular amylose on pasting characteristics of waxy corn starch. Food Chem 116:66–70. doi: 10.1016/j.foodchem.2009.02.006

Kusumayanti H, Handayani NA, Santosa H (2015) Swelling Power and Water Solubility of Cassava and Sweet Potatoes Flour. Procedia Environ Sci 23:164–167. doi: 10.1016/j.proenv.2015.01.025

Lee MR, Swanson BR, Baik BK (2001) Influence of amylose content on properties of wheat starch and breadmaking quality of starch and gluten blends. Cereal Chem 78:701–706. doi: 10.1094/CCHEM.2001.78.6.701

Li H, Handsaker B, Wysoker A, et al (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079. doi: 10.1093/bioinformatics/btp352

Liang X, King JM (2003) Pasting and crystalline property differences of commercial and isolated rice starch with added amino acids. J Food Sci 68:832–838. doi: 10.1111/j.1365-2621.2003.tb08251.x

Lim S-T, Kasemsuwan T, Jane J-L (1994) Characterization of Phosphorus in Starch by 31P-Nuclear Magnetic Resonance Spectroscopy. Cereal Chem 71:488–493.

Lin Q, Huang B, Zhang M, et al (2012) Functional Interactions between Starch Synthase III and Isoamylase-Type Starch-Debranching Enzyme in Maize Endosperm. PLANT Physiol 158:679–692. doi: 10.1104/pp.111.189704

Liu H, Yu L, Xie F, Chen L (2006) Gelatinization of cornstarch with different amylose/amylopectin content. Carbohydr Polym 65:357–363. doi: 10.1016/j.carbpol.2006.01.026

Lloyd JR, Kossmann J (2015) Transitory and storage starch metabolism: Two sides of the same coin? Curr. Opin. Biotechnol. 32:143–148.

Lorberth R, Ritte G, Willmitzer L, Kossmann J (1998) Inhibition of a starch-granule-bound protein leads to modified starch and repression of cold sweetening. Nat Biotechnol 16:473–477. doi: 10.1038/nbt0598-473

Lund D (1984) Influence of time, temperature, moisture, ingredients, and processing conditions on starch gelatinization. Crit Rev Food Sci Nutr 20:249–73. doi: 10.1080/10408398409527391

Mahlow S, Orzechowski S, Fettke J (2016) Starch phosphorylation: insights and perspectives. Cell Mol Life Sci 73:2753–2764. doi: 10.1007/s00018-016-2248-4

Manners DJ (1989) Recent developments in our understanding of amylopectin structure. Carbohydr. Polym. 11:87–112.

Martin C, Smith a M (1995) Starch biosynthesis. Plant Cell 7:971–985. doi: 10.1039/qr9692300057

Maurer T, Kalbitzer HR (1996) Indirect referencing of31P and19F NMR spectra. J Magn Reson - Ser B 113:177–178. doi: 10.1006/jmrb.1996.0172

Menlove KJ, Clement M, Crandall KA (2009) Similarity searching using BLAST. Methods Mol. Biol. 537:1–22.

Mikkelsen R, Mutenda KE, Mant A, et al (2005) Alpha-glucan, water dikinase (GWD): a plastidic enzyme with redox-regulated and coordinated catalytic activity and binding affinity. Proc Natl Acad Sci U S A 102:1785–1790.

Minoia S, Petrozza A, D'Onofrio O, et al (2010) A new mutant genetic resource for tomato crop improvement by TILLING technology. BMC Res Notes. doi: 10.1186/1756-0500-3-69

Montagnac JA, Davis CR, Tanumihardjo SA (2009a) Processing techniques to reduce toxicity and antinutrients of Cassava for use as a staple food. Compr Rev Food Sci Food Saf 8:17–27. doi: 10.1111/j.1541-4337.2008.00064.x

Montagnac JA, Davis CR, Tanumihardjo SA (2009b) Nutritional value of cassava for use as a staple food and recent advances for improvement. Compr Rev Food Sci Food Saf 8:181–194. doi: 10.1111/j.1541-4337.2009.00077.x

Montero WR (2003) Cassava; Biology, Production and Utilization.

Moorthy SN (2002) Physicochemical and functional properties of tropical tuber starches: A review. Starch/Staerke 54:559–592.

Morrison WR, Law R V., Snape CE (1993) Evidence for inclusion complexes of lipids with v-amylose in maize, rice and oat starches. J Cereal Sci 18:107–109. doi: 10.1006/jcrs.1993.1039

Morrison WR, Milligan TP, Azudin MN (1984) A relationship between the amylose and lipid contents of starches from diploid cereals. J Cereal Sci 2:257–271. doi: 10.1016/S0733-5210(84)80014-4

Mun JH, Kwon SJ, Yang TJ, et al (2009) Genome-wide comparative analysis of the Brassica rapa gene space reveals genome shrinkage and differential loss of duplicated genes after whole genome triplication. Genome Biol. doi: 10.1186/gb-2009-10-10-r111

Muth J, Hartje S, Twyman RM, et al (2008) Precision breeding for novel starch variants in potato. Plant Biotechnol J 6:576–584. doi: 10.1111/j.1467-7652.2008.00340.x

NACONSIE M, ZHANG P (2016) Transgenic technologies in cassava for nutritional improvement and viral disease resistance: a key strategy for food security in Africa. Front Agric Sci Eng 3:285. doi: 10.15302/J-FASE-2016119

Nakamura T, Yamamori M, Hirano H, et al (1995) Production of waxy (amylose-free) wheats. MGG Mol Gen Genet 248:253–259. doi: 10.1007/BF02191591

Naylor RL, Falcon WP, Goodman RM, et al (2004) Biotechnology in the developing world: A case for increased investments in orphan crops. Food Policy 29:15–44. doi: 10.1016/j.foodpol.2004.01.002

Nelson RJ, Naylor RL, Jahn MM (2004) The role of genomics research in improvement of 'orphan' crops. In: Crop Science. pp 1901–1904

Niittylä T, Messerli G, Trevisan M, et al (2004) A Previously Unknown Maltose Transporter Essential for Starch Degradation in Leaves. Science (80-) 303:87–89. doi: 10.1126/science.1091811

Nyakabau T, Wokadala OC, Emmambux MN (2013) Effect of steeping additives on tef starch extptraction and its quality. Starch/Staerke 65:738–746. doi: 10.1002/star.201200241

Oleykowski C a, Bronson Mullins CR, Godwin a K, Yeung a T (1998) Mutation detection using a novel plant endonuclease. Nucleic Acids Res 26:4597–4602. doi: 10.1093/nar/26.20.4597

Opabode JT, Oyelakin OO, Akinyemiju OA, Ingelbrecht IL (2012) Starch branching enzyme I from cassava (Manihot esculenta Crantz): Genomic organization , intron identification and phylogenetic analysis. Glob J Biotechnol Biocemistry 7:79–85. doi: 10.5829/idosi.gjbb.2012.7.3.1103

Pa K, Asseng S (2018) A review of tef physiology for developing a tef crop model. 94:54–66. doi: 10.1016/j.eja.2018.01.008

Pfister B, Zeeman SC (2016a) Formation of starch in plant cells. Cell. Mol. Life Sci. 73:2781–2807.

Pfister B, Zeeman SC (2016b) Formation of starch in plant cells. Cell Mol Life Sci 73:2781–2807. doi: 10.1007/s00018-016-2250-x
Pilling E, Smith a M (2003) Growth ring formation in the starch granules of potato tubers. Plant Physiol 132:365–371. doi: 10.1104/pp.102.018044
Polnaya FJ, Haryadi, Marseno DW, Cahyanto MN (2013) Effects of phosphorylation and cross-linking on the pasting properties and molecular structure of sago starch. Int Food Res J 20:1609–1615.
Puounti-Kaerlas J (2002) Cassava Biotechnology. Genet Eng Rev 15:329–364. doi: 10.1079/9780851995243.0000
Radchuk V V., Borisjuk L, Sreenivasulu N, et al (2009) Spatiotemporal Profiling of Starch Biosynthesis and Degradation in the Developing Barley Grain. PLANT Physiol 150:190–204. doi: 10.1104/pp.108.133520
Raemakers K, Schreuder M, Suurs L, et al (2005) Improved cassava starch by antisense inhibition of granule-bound starch synthase I. Mol Breed 16:163–172. doi: 10.1007/s11032-005-7874-8
Ran FA, Hsu PD, Wright J, et al (2013) Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8:2281–2308. doi: 10.1038/nprot.2013.143
Raphael M, Yona B, Stephen K, et al (2011) Amylopectin molecular structure and functional properties of starch from three Ugandan cassava varieties. J Plant Breed Crop Sci 3:195–202.
Rathore RS, Garg N, Garg S, Kumar A (2009) Starch phosphorylase: Role in starch metabolism and biotechnological applications. Crit. Rev. Biotechnol. 29:214–224.
Ratnayake WS, Jackson DS (2008) Starch Gelatinization. In: Advances in Food and Nutrition Research. pp 221–268
Regina A, Bird A, Topping D, et al (2006) High-amylose wheat generated by RNA interference improves indices of large-bowel health in rats. Proc Natl Acad Sci 103:3546–3551. doi: 10.1073/pnas.0510737103
Rhoads A, Au KF (2015) PacBio Sequencing and Its Applications. Genomics, Proteomics Bioinforma. 13:278–289.
Ritte G, Heydenreich M, Mahlow S, et al (2006a) Phosphorylation of C6- and C3-positions of glucosyl residues in starch is catalysed by distinct dikinases. FEBS Lett 580:4872–4876. doi: 10.1016/j.febslet.2006.07.085
Ritte G, Heydenreich M, Mahlow S, et al (2006b) Phosphorylation of C6- and C3-positions of glucosyl residues in starch is catalysed by distinct dikinases. FEBS Lett 580:4872–4876. doi: 10.1016/j.febslet.2006.07.085
Rongsirikul O, Saithong T, Kalapanulak S, et al (2010) Reconstruction of starch biosynthesis pathway in cassava using comparative genomic approach. In: Communications in Computer and Information Science. pp 118–129
Saithong T, Rongsirikul O, Kalapanulak S, et al (2013) Starch biosynthesis in cassava: A genome-based pathway reconstruction and its exploitation in data integration. BMC Syst Biol. doi: 10.1186/1752-0509-7-75
Salehuzzaman SNIM, Jacobsen E, Visser RGF (1993) Isolation and characterization of a cDNA encoding granule-bound starch synthase in cassava (Manihot esculenta Crantz) and its antisense expression in potato. Plant Mol Biol 23:947–962. doi: 10.1007/BF00021811
Sangwan P, Petinakis E, Dean K (2014) Starch Polymers.
Sano Y (1984) Differential regulation of waxy gene expression in rice endosperm. Theor Appl Genet 68:467–473. doi: 10.1007/BF00254822
Santana ÁL, Angela M, Meireles A (2014) New Starches are the Trend for Industry Applications: A Review. Food Public Heal 4:229–241. doi: 10.5923/j.fph.20140405.04

Santelia D, Kötting O, Seung D, et al (2011) The phosphoglucan phosphatase Like Sex Four2 dephosphorylates starch at the C3-position in Arabidopsis. Plant Cell 23:4096–4111. doi: 10.1105/tpc.111.092155
Santelia D, Zeeman SC (2011a) Progress in Arabidopsis starch research and potential biotechnological applications. Curr Opin Biotechnol 22:271–280. doi: 10.1016/j.copbio.2010.11.014
Santelia D, Zeeman SC (2011b) Progress in Arabidopsis starch research and potential biotechnological applications. Curr. Opin. Biotechnol. 22:271–280.
Satoh H, Nishi A, Yamashita K, et al (2003) Starch-branching enzyme I-deficient mutation specifically affects the structure and properties of starch in rice endosperm. Plant Physiol 133:1111–1121. doi: 10.1104/pp.103.021527
Scheben A, Wolter F, Batley J, et al (2017) Towards CRISPR/Cas crops - bringing together genomics and genome editing. New Phytol.
Seung D, Soyk S, Coiro M, et al (2015) PROTEIN TARGETING TO STARCH Is Required for Localising GRANULE-BOUND STARCH SYNTHASE to Starch Granules and for Normal Amylose Synthesis in Arabidopsis. PLoS Biol. doi: 10.1371/journal.pbio.1002080
Shumoy H, Raes K (2017) In vitro starch hydrolysis and estimated glycemic index of tef porridge and injera. Food Chem 229:381–387. doi: 10.1016/j.foodchem.2017.02.060
Silver DM, Kötting O, Moorhead GBG (2014) Phosphoglucan phosphatase function sheds light on starch degradation. Trends Plant Sci 19:471–478. doi: 10.1016/j.tplants.2014.01.008
Singh J, Kaur L, McCarthy OJ (2007) Factors influencing the physico-chemical, morphological, thermal and rheological properties of some chemically modified starches for food applications-A review. Food Hydrocoll. 21:1–22.
Singh N, Singh J, Kaur L, et al (2003) Morphological, thermal and rheological properties of starches from different botanical sources. Food Chem. 81:219–231.
Siriwat W, Kalapanulak S, Suksangpanomrung M, et al (2012) Transcriptomic data integration inferring the dominance of starch biosynthesis in carbon utilization of developing cassava roots. In: Procedia Computer Science. pp 96–106
Slade AJ, Fuerstenberg SI, Loeffler D, et al (2005) A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. Nat Biotechnol 23:75–81. doi: 10.1038/nbt1043
Smith AM (2010) Starch and Starch Granules. In: Encyclopedia of Life Sciences.
Smith AM, Zeeman SC, Smith SM (2005) Starch degradation. Annu Rev Plant Biol 56:73–98. doi: 10.1146/annurev.arplant.56.032604.144257
Sonnewald U, Kossmann J (2013) Starches-from current models to genetic engineering. Plant Biotechnol. J. 11:223–232.
Spaenij-Dekking L, Kooy-Winkelaar Y, Koning F (2005) The Ethiopian cereal tef in celiac disease. N Engl J Med 353:1748–1749. doi: 10.1056/NEJMc051492
Sreenivasulu N, Wobus U (2013) Seed-Development Programs: A Systems Biology–Based Comparison Between Dicots and Monocots. Annu Rev Plant Biol 64:189–217. doi: 10.1146/annurev-arplant-050312-120215
Srichuwong S, Sunarti TC, Mishima T, et al (2005) Starches from different botanical sources II: Contribution of starch structure to swelling and pasting properties. Carbohydr Polym 62:25–34. doi: 10.1016/j.carbpol.2005.07.003
Sriroth K, Santisopasri V, Petchalanuwat C, et al (1999) Cassava starch granule structure-function properties: Influence of time and conditions at harvest on four cultivars of cassava starch. Carbohydr Polym 38:161–170. doi: 10.1016/S0144-8617(98)00117-9
Streb S, Delatte T, Umhang M, et al (2008) Starch Granule Biosynthesis in Arabidopsis Is

Abolished by Removal of All Debranching Enzymes but Restored by the Subsequent Removal of an Endoamylase. PLANT CELL ONLINE 20:3448–3466. doi: 10.1105/tpc.108.063487
Streb S, Zeeman SC (2012) Starch metabolism in Arabidopsis. Arab B. doi: 10.1199/tab.0160
Sun Y, Jiao G, Liu Z, et al (2017) Generation of High-Amylose Rice through CRISPR/Cas9-Mediated Targeted Mutagenesis of Starch Branching Enzymes. Front Plant Sci. doi: 10.3389/fpls.2017.00298
Sun ZT, Henson CA (1991) A quantitative assessment of the importance of barley seed alpha-amylase, beta-amylase, debranching enzyme, and alpha-glucosidase in starch degradation. Arch Biochem Biophys 284:298–305.
Svihus B, Uhlen AK, Harstad OM (2005) Effect of starch granule structure, associated components and processing on nutritive value of cereal starch: A review. Anim. Feed Sci. Technol. 122:303–320.
Szymońska J, Targosz-Korecka M, Krok F (2009) Characterization of starch nanoparticles. In: Journal of Physics: Conference Series.
Tadele, Z. and Esfeld K (2009) The Improvement of African Orphan Crops through TILLING. Atdf J 6:44–49.
Tadele Z (2009) Role of Orphan Crops in Enhancing and Diversifying Food Production in Africa. African Technol y Dev Forum J 9–15.
Tako M, Tamaki Y, Teruya T, Takeda Y (2014) The Principles of Starch Gelatinization and Retrogradation. Food Nutr Sci 5:280–291. doi: 10.4236/fns.2014.53035
Tatge H, Marshall J, Martin C, et al (1999) Evidence that amylose synthesis occurs within the matrix of the starch granule in potato tubers. Plant, Cell Environ 22:543–550. doi: 10.1046/j.1365-3040.1999.00437.x
Taylor N, Chavarriaga P, Raemakers K, et al (2004) Development and application of transgenic technologies in cassava. Plant Mol. Biol. 56:671–688.
Taylor NJ, Fauquet CM, Tohme J (2012) Overview of Cassava Special Issue. Trop. Plant Biol. 5:1–3.
Taylor S, Scott R, Kurtz R, et al (2009) A practical guide to High Resolution Melt Analysis genotyping.
Tester RF, Karkalas J, Qi X (2004) Starch - Composition, fine structure and architecture. J Cereal Sci 39:151–165. doi: 10.1016/j.jcs.2003.12.001
Tester RF, Morrison WR (1990) Swelling and gelatinization of cereal starches. I. Effects of amylopectin, amylose, and lipids. Cereal Chem 67:551–557.
Tetlow I, Emes M (2017) Starch Biosynthesis in the Developing Endosperms of Grasses and Cereals. Agronomy 7:81. doi: 10.3390/agronomy7040081
Tetlow IJ (2011) Starch biosynthesis in developing seeds. Seed Sci. Res. 21:5–32.
Tetlow IJ, Emes MJ (2011) Starch Biosynthesis in Higher Plants: The Enzymes of Starch Synthesis. In: Comprehensive Biotechnology, Second Edition. pp 47–65
Tian Z, Qian Q, Liu Q, et al (2009) Allelic diversities in rice starch biosynthesis lead to a diverse array of rice eating and cooking qualities. Proc Natl Acad Sci 106:21760–21765. doi: 10.1073/pnas.0912396106
Till BJ, Reynolds SH, Weil C, et al (2004) Discovery of induced point mutations in maize genes by TILLING. BMC Plant Biol. doi: 10.1186/1471-2229-4-12
Tomlinson K, Denyer K (2003) Starch synthesis in cereal grains. Adv. Bot. Res. 40:1–61.
Uauy C, Paraiso F, Colasuonno P, et al (2009) A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat. BMC Plant Biol 9:115. doi: 10.1186/1471-2229-9-115
Ukwuru MU and Egbonu SE (2013) Recent development in cassava-based products research. Acad J Food Res 1:1–13. doi: 10.15413/ajfr.2012.0105

Vamadevan V, Bertoft E (2015) Structure-function relationships of starch components. Starch/Staerke 67:55–68.

Van der Auwera GA, Carneiro MO, Hartl C, et al (2013) From fastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. Curr Protoc Bioinforma. doi: 10.1002/0471250953.bi1110s43

Van Harsselaar JK, Lorenz J, Senning M, et al (2017) Genome-wide analysis of starch metabolism genes in potato (Solanum tuberosum L.). BMC Genomics. doi: 10.1186/s12864-016-3381-z

Varshney RK, Ribaut JM, Buckler ES, et al (2012) Can genomics boost productivity of orphan crops? Nat. Biotechnol. 30:1172–1176.

Vasconcelos LM, Brito AC, Carmo CD, Oliveira EJ (2016) Polymorphism of starch pathway genes in cassava. 1:1–15.

Vrinten PL, Nakamura T (2000) Wheat granule-bound starch synthase I and II are encoded by separate genes that are expressed in different tissues. Plant Physiol 122:255–264. doi: 10.1104/pp.122.1.255

Wang S, Li C, Copeland L, et al (2015a) Starch Retrogradation: A Comprehensive Review. Compr Rev Food Sci Food Saf 14:568–585. doi: 10.1111/1541-4337.12143

Wang S, Li S, Liu Q, et al (2015b) The OsSPL16-GW7 regulatory module determines grain shape and simultaneously improves rice yield and grain quality. Nat Genet 47:949–954. doi: 10.1038/ng.3352

Wang X, Chang L, Tong Z, et al (2016) Proteomics Profiling Reveals Carbohydrate Metabolic Enzymes and 14-3-3 Proteins Play Important Roles for Starch Accumulation during Cassava Root Tuberization. Sci Rep. doi: 10.1038/srep19643

Waterschoot J, Gomand S V., Fierens E, Delcour JA (2015) Production, structure, physicochemical and functional properties of maize, cassava, wheat, potato and rice starches. Starch/Staerke 67:14–29.

Weber A, Servaites JC, Geiger DR, et al (2000) Identification, Purification, and Molecular Cloning of a Putative Plastidic Glucose Translocator. Plant Cell 12:787–801. doi: 10.1105/tpc.12.5.787

Weise SE, Aung K, Jarou ZJ, et al (2012) Engineering starch accumulation by manipulation of phosphate metabolism of starch. Plant Biotechnol J 10:545–554. doi: 10.1111/j.1467-7652.2012.00684.x

Wilm A, Aw PPK, Bertrand D, et al (2012) LoFreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res 40:11189–11201. doi: 10.1093/nar/gks918

Wolter A, Hager AS, Zannini E, Arendt EK (2013) In vitro starch digestibility and predicted glycaemic indexes of buckwheat, oat, quinoa, sorghum, teff and commercial gluten-free bread. J Cereal Sci 58:431–436. doi: 10.1016/j.jcs.2013.09.003

Wong KM, Suchard MA, Huelsenbeck JP (2008) Alignment uncertainty and genomic analysis. Science (80-) 319:473–476. doi: 10.1126/science.1151532

Xin Z, Li Wang M, Barkley NA, et al (2008) Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population. BMC Plant Biol. doi: 10.1186/1471-2229-8-103

Xiong JS, Ding J, Li Y (2015) Genome-editing technologies and their potential application in horticultural crop breeding. Hortic. Res. 2:

Xu X, Huang XF, Visser RGF, Trindade LM (2017) Engineering potato starch with a higher phosphate content. PLoS One. doi: 10.1371/journal.pone.0169610

Xu X, Visser RGF, Trindade LM (2014) Starch Modification by Biotechnology: State of Art and Perspectives. In: Starch Polymers: From Genetic Engineering to Green

Applications. pp 79–102
Yan CJ, Tian ZX, Fang YW, et al (2011) Genetic analysis of starch paste viscosity parameters in glutinous rice (Oryza sativa L.). Theor Appl Genet 122:63–76. doi: 10.1007/s00122-010-1423-5
Yu JK, Graznak E, Breseghello F, et al (2007) QTL mapping of agronomic traits in tef Eragrostis tef (Zucc). BMC Plant Biol. doi: 10.1186/1471-2229-7-30
Yu S, Ma Y, Sun DW (2009) Impact of amylose content on starch retrogradation and texture of cooked milled rice during storage. J Cereal Sci 50:139–144. doi: 10.1016/j.jcs.2009.04.003
Yu TS, Kofler H, Häusler RE, et al (2001) The Arabidopsis sex1 mutant is defective in the R1 protein, a general regulator of starch degradation in plants, and not in the chloroplast hexose transporter. Plant Cell 13:1907–1918. doi: 10.1105/TPC.010091
Yu TS, Zeeman SC, Thorneycroft D, et al (2005) ??-amylase is not required for breakdown of transitory starch in Arabidopsis Leaves. J Biol Chem 280:9773–9779. doi: 10.1074/jbc.M413638200
Zeeman SC (2002) Starch Synthesis in Arabidopsis. Granule Synthesis, Composition, and Structure. PLANT Physiol 129:516–529. doi: 10.1104/pp.003756
Zeeman SC, Kossmann J, Smith AM (2010a) Starch: its metabolism, evolution, and biotechnological modification in plants. Annu Rev Plant Biol 61:209–234. doi: 10.1146/annurev-arplant-042809-112301
Zeeman SC, Kossmann J, Smith AM (2010b) Starch: its metabolism, evolution, and biotechnological modification in plants. Annu Rev plant … 61:209–234. doi: 10.1146/annurev-arplant-042809-112301
Zeeman SC, Northrop F, Smith AM, Ap Rees T (1998) A starch-accumulating mutant of Arabidopsis thaliana deficient in a chloroplastic starch-hydrolysing enzyme. Plant J 15:357–365. doi: 10.1046/j.1365-313X.1998.00213.x
Zhang L, Häusler RE, Greiten C, et al (2008) Overriding the co-limiting import of carbon and energy into tuber amyloplasts increases the starch content and yield of transgenic potato plants. Plant Biotechnol J 6:453–464. doi: 10.1111/j.1467-7652.2008.00332.x
Zhao SS, Dufour D, Sánchez T, et al (2011) Development of waxy cassava with different Biological and physico-chemical characteristics of starches for industrial applications. Biotechnol Bioeng 108:1925–1935. doi: 10.1002/bit.23120
Zhou W, Yang J, Hong Y, et al (2015) Impact of amylose content on starch physicochemical properties in transgenic sweet potato. Carbohydr Polym 122:417–427. doi: 10.1016/j.carbpol.2014.11.003
Zhu F (2015) Composition, structure, physicochemical properties, and modifications of cassava starch. Carbohydr. Polym. 122:456–480.
Zhu F (2017) Barley Starch: Composition, Structure, Properties, and Modifications. Compr Rev Food Sci Food Saf 16:558–579. doi: 10.1111/1541-4337.12265
Zhu F (2018) Chemical composition and food uses of teff (Eragrostis tef). Food Chem. 239:402–415.
Zhu T, Jackson DS, Wehling RL, Geera B (2008) Comparison of amylose determination methods and the development of a dual wavelength iodine binding technique. Cereal Chem 85:51–58. doi: 10.1094/CCHEM-85-1-0051
Zhu W, Buell CR (2007) Improvement of whole-genome annotation of cereals through comparative analyses. Genome Res 17:299–310. doi: 10.1101/gr.5881807

Chapter 4
General Discussion

4.1 Investigation of starch metabolism in storage organs

Transitory starch synthesis and degradation in the Arabidopsis chloroplast is relatively well understood thanks to the wealth of genomic information and the associated resources that enable rapid research in this model species (Streb and Zeeman 2012; Pfister and Zeeman 2016a). In non-photosynthetic tissues of different crops, the synthesis and degradation of storage starch is often less well studied. It is clear that in some cases it there are distinct aspects to its metabolism, while in other cases, enzymes discovered in leaves share common functionalities in tissues line developing seeds and tubers (James et al. 2003; Hannah and James 2008; Radchuk et al. 2009; Tetlow and Emes 2017; Van Harsselaar et al. 2017). While research into model species like Arabidopsis is very valuable for gene discovery, for crop improvement, research into storage starch metabolism needs to be increased. We need to clarify and complete our understanding of the starch metabolic pathways in each tissues and hopefully improve it through breeding and biotechnological means (Zeeman et al. 2010b; Lloyd and Kossmann 2015). By utilizing tef as a cereal crop and cassava as a root crop, we explored the possibilities to enhance storage starch metabolism in both the starchy endosperm of seeds and in storage roots.

Glucan phosphorylation is good example of an important process in starch metabolism, which also has an impact on starch functionality, but which is not fully understood. We know from pioneering work in potato and Arabidopsis that reversible glucan phosphorylation occurs via the actions of dikinases (e.g. GWD and PWD) and phosphatases (e.g. SEX4 and LSF2)(Lorberth et al. 1998; Kötting 2005; Ritte et al. 2006; Kötting et al. 2009; Santelia and Zeeman 2011). The current model, mainly from work in Arabidopsis leaves, is that phosphorylation initiates starch degradation through solubilizing the semi-crystalline structure of starch at the granules surface (Streb and Zeeman 2012). There is evidence that phosphorylation is similarly important for the degradation of starch in other non-photosynthetic tissues such as Arabidopsis roots (Caspar et al, 1991), potato tubers (Lorberth et al., 1998) and tomato pollen (Nashilevitz et al., 2009). The rate of glucan phosphorylation is probably but its highest during starch degradation but, as it is transient, it is not readily measurable. The amount of phosphate that is measurable bound to starch, which is important from an applied aspect, is added during starch biosynthesis and probably determined by the balance between the background rates of phosphorylation and dephosphorylation. We actually do not fully understand what controls these rates in any system as yet.

In the cereals rice, wheat, maize and others, many studies of starch biosynthesis in the seed endosperm have been carried out. Indeed, many fundamental discoveries were made and concepts developed using these systems before the emergence of model systems like Arabidopsis. However, the functions of newly discovered genes in Arabidopsis, such as PTST(PTST1), ESV and LESV, in regulating endosperm starch metabolism has not yet been

studied. The advantage of high-through put approaches such as the TILLING-by-sequencing described in Chapter 3 is that it offers the potential to identify mutant alleles within a cereal population to allow such investigations of their functions in the endosperm. Of course, tef, being tetraploid, might not represent the ideal system in which to do this since the issue of redundancy between homeologs increases the amount of work needed.

4.2 Biotechnology for improving orphan crops

Based on our knowledge of starch biosynthesis and regulation achieved in both model and non-model species, biotechnological applications have led to development of a number of transgenic crops with elevated starch content or improved starch quality (Avni and Blázquez 2011; Santelia and Zeeman 2011; Sonnewald and Kossmann 2013). The biotechnological approaches used for producing starch with novel functionalities can be classified as either manipulation of endogenous genes involved in starch metabolism or expression of heterologous genes encoding glucan biosynthetic or modifying enzymes (Xu et al. 2014).

Technology to generate transgenic cassava cultivars expressing marker genes and/or genes of agronomic interest has been developed for cassava production (Chavarriaga-Aguirre et al. 2016). We applied this method and successfully generated transgenic cassava overexpressing *St*GWD, or RNAi lines repressing *Me*SEX4 or *Me*LSF2 individually. Our cassava work confirms the importance of GWD in regulating the degree of phosphorylation of storage starch during biosynthesis. The total starch phosphate level increased by as much as double in cassava roots overexpressing *St*GWD (i.e. R9 line). Nevertheless, as mentioned in Chapter 2, the starch phosphate content was still far from that in potato, implying that GWD is not the sole point of control. In this context, we also confirmed the functionality of the LSF2 enzyme in catalyzing the removal of C-3 bound phosphates in storage root starch. Through repressing the *MeLSF2* gene by RNAi, the proportion of C-3 bound phosphate rose from 28% seen in the wild type to 35%. Although this was a relatively modest shift, and there was no measurable increase in total starch-bound phosphate, it shows that the enzyme is performing the same role as it is in Arabidopsis, where it was discovered (Santelia et al., 2011). We predict that more significant changes, similar to what was seen in Arabidopsis, would be achieved with a *Melsf2* null mutant. Furthermore, we believe that, by targeting multiple genes simultaneously, much larger changes in starch-bound phosphate could be achieved. A line like R9 which has good overexpression of GWD could be combined with an *lsf2* null mutant. Other changes that could also be introduced to boost both starch levels and starch phosphate could include simultaneous upregulation of plastid envelope transporters such as the GPT hexose phosphate transporter that delivers G6P into the amyloplast, and the NTT adenylate transporter that provides ATP. These strategies, when combined, have been shown to increase starch production in potato tubers and could be equally effective in cassava storage roots (Kampfenkel et al. 1995; Kammerer et al. 1998; Zhang et al. 2008) .

However, it is important to note that not all of our RNAi approaches were equally successful. Repressing *MeSEX4* in cassava led to a severe growth phenotype among almost all transgenic lines. More work would be required to understand the cause of this phenotype. This in turn could help define the most suitable technology with which to modify the expression of the *MeSEX4* gene.

Realistically, it is complex to proceed directly from studies of metabolism in model systems to improvements in different crops. Compared with cassava, tef is even less well researched. Biotechnological applications are obstructed in tef due to limited understanding of starch metabolism, lack of a completed draft genome, increased genomic complexity due to polyploidy and the large size of the tef genome. It is nevertheless reasonable to assume that closely-related and better studied cereals like sorghum and rice can be used to guide improvement strategies. This was borne out by our bioinformatic analyses, which showed a similar gene complement in tef, and by our gene expression analyses, which showed similar isoform-specific expression patterns. For improving the starch characteristics of tef, we used a reverse genetic mutant breeding approach thanks to the existence of a EMS-mutagenized population established at the University of Bern. To create tef varieties with altered starch properties requires a TILLING approach. This is because the most obvious gene targets like GBSSI and SBE (to control amylose content), require null mutants and, due to the tetraploid genome of tef, the chances of identifying mutants in both homeologs through phenotypic screening is essentially zero. Conventional TILLING is effective, but time consuming. Furthermore, it focusses only on part of the target genes and may miss valuable mutant alleles in the population. These facts prompted us to develop a high-throughput sequencing-based strategy to optimize mutation detection stage and, in the long run to become more efficient than with standard TILLING methods. Many polymorphisms in targeted starch-related genes in the tef population were successfully identified, although we emphasize that the initial conventional TILLING approach was critical for optimizing the mutation detection procedure. These known mutations allowed us to assess the performance of the bioinformatics tools developed in the clinic research fields to see which could identify low frequency SNPs in our sequence dataset derived tetraploid tef. Important lessons were learned that will improve the future filtering of positive candidates from background noise, reducing both the number of false positives and the risk of accidentally filtering out genuinely useful alleles. Our approach has the capacity to deal with much larger mutant populations than used here or in many previous studies, and would allow investigation of many genes at once.

Some of the induced missense mutants identified are now currently on trial. Unfortunately, it was not possible within the timeframe of my project to both develop the tools, apply them and do the necessary downstream mutant plant identification and breeding to create lines with null mutations in both homeologs of given target genes.

The magnitude of breeding efforts and the adoption of modern technologies for orphan crops varies in different countries. Molecular approaches for improving orphan crops are demanding

and require relatively sophisticated research environments, but the output from these efforts can be integrated with conventional breeding programs in the countries where the orphan crops are grown. The advancement of genetics and genomics will undoubtedly benefit orphan crop research (Naylor et al. 2004; Nelson et al. 2004; Varshney et al. 2012). Modern DNA sequencing technologies allow rapid expansion of genomic analysis in such species. Studies with related well-characterized model plant species can then be exploited to assist orphan crop research (Feltus et al. 2006; Armstead et al. 2009; Kamei et al. 2016). Varshney et al. (2012) described the potential gain and challenges in implementing genomics-assisted breeding in orphan crops. The argued that with the use of next-generation sequencing and high-throughput molecular biology tools, orphan crops could be converted into genomic resource–rich crops for which the approaches such as association mapping or genome-wide association studies (GWAS), marker-assisted recurrent selection (MARS) and genomic selection (GS) could be readily applied to improve properties such as grain yield for developing countries. This strategy, through prediction of plant phenotype from genotype, could increase the rate of genetic gain across target environments in less time and at lower costs compared with conventional selection based exclusively on phenotype.

Recently, the power of genome-editing technologies to facilitate efficient plant breeding for crop improvement has been demonstrated (Xiong et al. 2015). Of these technologies, CRISPR/Cas9 (clustered regularly interspaced short palindromic repeat/CRISPR-associated protein 9) and associated technologies hold the most promise, allowing targeted modification of almost any crop genome sequence to generate novel variation and accelerate breeding efforts (Scheben et al. 2017). These methods will certainly find application in orphan crop improvement.

4.3 Potential application of modified starch

Here we demonstrated that starch with increased phosphate content results in stable paste properties and transparent gels. Our high phosphate cassava starch with altered physicochemical properties (swelling power and light transmittance) has the potential to be transferred to farmer-preferred varieties. The starch traits could be further enhanced by implementing genomics-assisted breeding or genome-editing technologies as mentioned above.

Through the tef TILLING approach, it should be possible to generate amylose-free tef starch composed only of amylopectin molecules if the mutations identified in the *GBSS1* genes prove to be null mutants. As for other cereal waxy starches, the mutant starch should have improved uniformity, stability and texture, and could be used not only in food products but also in the diverse industries due to its positive attributes (e.g. gluten free and nutrition value). Likewise, mutation of *SBEs* in tef, if nulls, will result in *ae* (*amylose extender*) type starch, which contains high amylose content. High-amylose starch has the ability to form strong gels and films and can be used in textiles, candies and the adhesives industry. It is also commonly known as resistant starch (RS), which is well suited for food application. Tef varieties with altered

amylopectin structure may also be achieved by modulation of other starch-related genes, which could lead to more diversity in starch properties for broader utilization. For instance, modified starch granule size and number will affect starch digestibility properties, food quality and potentially crop production. Through my project, I have created the data and the tools needed for such approaches, although as mentioned in Chapter 3 for the example of GBSSI, considerable manual curation of the gene annotation data, and care in interpretation of the sequence and mutation data and is still needed for the analysis of each gene.

4.4 Importance of increasing starch research in orphan crops

The term "orphan crop" refers to any crop to which little attention has been given to date. It is a label that applies to a large number of cereals, legumes, vegetables, root crops, fodder crops, oil crops, fiber crops and medicinal plants. Besides of their obvious local significance, their potential value in nutrition, food security, environmental stress tolerance, medicinal use and other unexplored characteristics should be recognized and further developed as a part of global food resources. Maintaining genetic diversity is increasingly recognized as critical for sustainable environmental management and agricultural production. Starch bioengineering in major crops has been well developed and the modified starches achieved through these efforts display altered properties which are utilized in diverse food and industries. This has led the way for similarly enhancing the potential of orphan crop varieties by altering their starch content and quality. Starch research in orphan crops is still limited, however a better understanding would be a great step toward the synthesis of tailored starches for different bio-based applications.

4.5 References

Abebe W, Collar C, Ronda F (2015) Impact of variety type and particle size distribution on starch enzymatic hydrolysis and functional properties of tef flours. Carbohydr Polym 115:260–268. doi: 10.1016/j.carbpol.2014.08.080

Agarwal P, Kapoor S, Tyagi AK (2011) Transcription factors regulating the progression of monocot and dicot seed development. BioEssays 33:189–202.

Ahamed NT, Singhal RS, Kulkarni PR, Pal M (1996) Physicochemical and functional properties of Chenopodium quinoa starch. Carbohydr Polym 31:99–103. doi: 10.1016/S0144-8617(96)00034-3

Alcázar-Alay SC, Meireles MAA (2015) Physicochemical properties, modifications and applications of starches from different botanical sources. Food Sci Technol 35:215–236. doi: 10.1590/1678-457X.6749

Alexandersson M, Cawley S, Pachter L (2003) SLAM: Cross-species gene finding and alignment with a generalized pair hidden Markov model. Genome Res 13:496–502. doi: 10.1101/gr.424203

Allem AC, Genéticos R (2002) The Origins and Taxonomy of Cassava. Cassava Biol Prod Util 1–16. doi: 10.1080/21622965.2014.897905

Altschul SF, Gish W, Miller W, et al (1990) Basic local alignment search tool. J Mol Biol 215:403–10. doi: 10.1016/S0022-2836(05)80360-2

Alves a. a. C (2002) Cassava botany and physiology. Cassava Biol Prod Util 67–89. doi: 10.1079/9780851995243.0000

Anand S, Mangano E, Barizzone N, et al (2016) Next generation sequencing of pooled samples: Guideline for variants' filtering. Sci Rep. doi: 10.1038/srep33735

Armstead I, Huang L, Ravagnani A, et al (2009) Bioinformatics in the orphan crops. Brief. Bioinform. 10:645–653.

Ashogbon AO, Akintayo ET (2014) Recent trend in the physical and chemical modification of starches from different botanical sources: A review. Starch/Staerke 66:41–57.

Aslam U, Cheema HMN, Ahmad S, et al (2016) COTIP: Cotton TILLING Platform, a Resource for Plant Improvement and Reverse Genetic Studies. Front Plant Sci. doi: 10.3389/fpls.2016.01863

Assefa K, Chanyalew S, Tadele Z (2017) Tef, Eragrostis tef (Zucc.) Trotter. In: Millets and Sorghum: Biology and Genetic Improvement. pp 226–266

Assefa K, Yu J-K, Zeid M, et al (2011) Breeding tef [Eragrostis tef (Zucc.) trotter]: conventional and molecular approaches. Plant Breed 130:1–9. doi: 10.1111/j.1439-0523.2010.01782.x

Avni A, Blázquez MA (2011) Can plant biotechnology help in solving our food and energy shortage in the future? Curr. Opin. Biotechnol. 22:220–223.

Balagopalan C (2002) Cassava Utilization in Food, Feed and Industry. Cassava Biol Prod Util 301–318. doi: 10.1079/9780851995243.0000

Barkley N, Wang M (2008) Application of TILLING and EcoTILLING as Reverse Genetic Approaches to Elucidate the Function of Genes in Plants and Animals. Curr Genomics 9:212–226. doi: 10.2174/138920208784533656

Baunsgaard L, Lütken H, Mikkelsen R, et al (2005) A novel isoform of glucan, water dikinase phosphorylates pre-phosphorylated α-glucans and is involved in starch degradation in Arabidopsis. Plant J 41:595–605. doi: 10.1111/j.1365-313X.2004.02322.x

Blagbrough IS, Bayoumi SAL, Rowan MG, Beeching JR (2010) Cassava: An appraisal of its phytochemistry and its biotechnological prospects. Phytochemistry 71:1940–1951. doi:

10.1016/j.phytochem.2010.09.001
Blazek J, Copeland L (2008) Pasting and swelling properties of wheat flour and starch in relation to amylose content. Carbohydr Polym 71:380–387. doi: 10.1016/j.carbpol.2007.06.010
Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254. doi: 10.1016/0003-2697(76)90527-3
Breuninger WF, Piyachomkwan K, Sriroth K (2009) Tapioca/Cassava Starch: Production and Use. In: Starch. pp 541–568
Bull SE, Owiti JA, Niklaus M, et al (2009) Agrobacterium-mediated transformation of friable embryogenic calli and regeneration of transgenic cassava. Nat Protoc 4:1845–1854. doi: 10.1038/nprot.2009.208
Bultosa G (2016) Teff: Overview, 2nd edn. Elsevier Ltd.
Bultosa G (2007) Physicochemical characteristics of grain and flour in 13 tef [Eragrostis tef (Zucc .) Trotter] grain varieties. J Appl Sci Res 3:2042–2050.
Bultosa G, Hall AN, Taylor JRN (2002) Physico-chemical characterization of grain tef [Eragrostis tef (Zucc.) trotter] starch. Starch/Staerke 54:461–468. doi: 10.1002/1521-379X(200210)54:10<461::AID-STAR461>3.0.CO;2-U
Bultosa G, Taylor JRN (2003) Chemical and physical characterisation of grain tef [Eragrostis tef (Zucc.) Trotter] starch granule composition. Starch/Staerke 55:304–312. doi: 10.1002/star.200390065
Bultosa G, Taylor JRN (2004) Paste and gel properties and in vitro digestibility of Tef [Eragrostis tef (Zucc). Trotter] starch. Starch/Staerke 56:20–28. doi: 10.1002/star.200200191
Burton RA, Jenner H, Carrangis L, et al (2002) Starch granule initiation and growth are altered in barley mutants that lack isoamylase activity. Plant J 31:97–112. doi: 10.1046/j.1365-313X.2002.01339.x
Cannarozzi G, Chanyalew S, Assefa K, et al (2018) Technology generation to dissemination: lessons learned from the tef improvement project. Euphytica 214:31. doi: 10.1007/s10681-018-2115-5
Cannarozzi G, Plaza-Wüthrich S, Esfeld K, et al (2014) Genome and transcriptome sequencing identifies breeding targets in the orphan crop tef (Eragrostis tef). BMC Genomics 15:581. doi: 10.1186/1471-2164-15-581
Carciofi M, Shaik SS, Jensen SL, et al (2011) Hyperphosphorylation of cereal starch. J Cereal Sci 54:339–346. doi: 10.1016/j.jcs.2011.06.013
Carpenter MA, Joyce NI, Genet RA, et al (2015) Starch phosphorylation in potato tubers is influenced by allelic variation in the genes encoding glucan water dikinase, starch branching enzymes I and II, and starch synthase III. Front Plant Sci 6:143. doi: 10.3389/fpls.2015.00143
Caspar T, Lin TP, Kakefuda G, et al (1991) Mutants of Arabidopsis with altered regulation of starch degradation. Plant Physiol 95:1181–1188. doi: 10.1104/pp.95.4.1181
Ceballos H, Sánchez T, Morante N, et al (2007) Discovery of an amylose-free starch mutant in cassava (Manihot esculenta Crantz). J Agric Food Chem 55:7469–7476. doi: 10.1021/jf070633y
Chan HT, Bhat R, Karim AA (2009) Physicochemical and functional properties of ozone-oxidized starch. J Agric Food Chem 57:5965–5970. doi: 10.1021/jf9008789
Chavarriaga-Aguirre P, Brand A, Medina A, et al (2016) The potential of using biotechnology to improve cassava: a review. Vitr. Cell. Dev. Biol. - Plant 52:461–478.
Chetty CC, Rossin CB, Gruissem W, et al (2013) Empowering biotechnology in southern Africa: Establishment of a robust transformation platform for the production of

transgenic industry-preferred cassava. N Biotechnol 30:136–143. doi: 10.1016/j.nbt.2012.04.006
Chiu C wai, Solarek D (2009) Modification of Starches. In: Starch. pp 629–655
Colbert T (2001) High-Throughput Screening for Induced Point Mutations. PLANT Physiol 126:480–484. doi: 10.1104/pp.126.2.480
Craig SAS, Maningat CC, Seib PA, Hoseney RC (1989) Starch paste clarity. Cereal Chem 66:173–182.
Critchley JH, Zeeman SC, Takaha T, et al (2001) A critical role for disproportionating enzyme in starch breakdown is revealed by a knock-out mutation in Arabidopsis. Plant J 26:89–100. doi: 10.1046/j.1365-313X.2001.01012.x
Crumpton-Taylor M, Pike M, Lu KJ, et al (2013) Starch synthase 4 is essential for coordination of starch granule formation with chloroplast division during Arabidopsis leaf expansion. New Phytol 200:1064–1075. doi: 10.1111/nph.12455
Deffenbaugh LB, Walker CE (1989) Use of the Rapid-Visco-Analyzer to Measure Starch Pasting Properties. Part I: Effect of Sugars. Starch - Stärke 41:461–467. doi: 10.1002/star.19890411205
Delatte T, Umhang M, Trevisan M, et al (2006) Evidence for distinct mechanisms of starch granule breakdown in plants. J Biol Chem 281:12050–12059. doi: 10.1074/jbc.M513661200
Donald AM (2004) Understanding starch structure and functionality. In: Starch in Food: Structure, Function and Applications. pp 156–184
Dumez S, Wattebled F, Dauvillee D, et al (2006) Mutants of Arabidopsis Lacking Starch Branching Enzyme II Substitute Plastidial Starch Synthesis by Cytoplasmic Maltose Accumulation. PLANT CELL ONLINE 18:2694–2709. doi: 10.1105/tpc.105.037671
Edner C, Li J, Albrecht T, et al (2007) Glucan, Water Dikinase Activity Stimulates Breakdown of Starch Granules by Plastidial β-Amylases. Plant Physiol 145:17–28. doi: 10.1104/pp.107.104224
El-Sharkawy MA (2004) Cassava biology and physiology. Plant Mol Biol 56:481–501. doi: 10.1007/s11103-005-2270-7
Eliasson AC (2004) Starch in Food: Structure, Function and Applications.
Esfeld K, Uauy C, Tadele Z (2012) Application of TILLING for orphan crop improvement. In: Biotechnology of Neglected and Underutilized Crops. pp 83–113
Falade KO, Akingbala JO (2011) Utilization of Cassava for food. Food Rev Int 27:51–83. doi: 10.1080/87559129.2010.518296
Feike D, Seung D, Graf A, et al (2016) A novel starch-granule-associated protein required for the control of starch degradation in Arabidopsis thaliana leaves. Plant Cell 28:1472–1489. doi: 10.1105/tpc.16.00011
Feltus F a, Singh HP, Lohithaswa HC, et al (2006) A comparative genomics strategy for targeted discovery of single-nucleotide polymorphisms and conserved-noncoding sequences in orphan crops. Plant Physiol 140:1183–1191. doi: 10.1104/pp.105.074203
Fredriksson H, Silverio J, Andersson R, et al (1998) The influence of amylose and amylopectin characteristics on gelatinization and retrogradation properties of different starches. Carbohydr Polym 35:119–134. doi: 10.1016/S0144-8617(97)00247-6
Fu F-F, Xue H-W (2010) Coexpression Analysis Identifies Rice Starch Regulator1, a Rice AP2/EREBP Family Transcription Factor, as a Novel Rice Starch Biosynthesis Regulator. PLANT Physiol 154:927–938. doi: 10.1104/pp.110.159517
Fujita N, Yoshida M, Asakura N, et al (2006) Function and Characterization of Starch Synthase I Using Mutants in Rice. PLANT Physiol 140:1070–1084. doi: 10.1104/pp.105.071845
Fulton DC, Stettler M, Mettler T, et al (2008) -AMYLASE4, a Noncatalytic Protein

Required for Starch Breakdown, Acts Upstream of Three Active -Amylases in Arabidopsis Chloroplasts. PLANT CELL ONLINE 20:1040–1058. doi: 10.1105/tpc.107.056507
Gallant DJ, Bouchet B, Baldwin PM (1997) Microscopy of starch: evidence of a new level of granule organization. Carbohydr Polym 32:177–191. doi: 10.1016/S0144-8617(97)00008-8
Gao M, Fisher DK, Kim K-N, et al (1996) Evolutionary conservation and expression patterns of maize starch branching enzyme I and IIb genes suggests isoform specialization. Plant Mol Biol 30:1223–1232. doi: 10.1007/BF00019554
Garrison E, Marth G (2012) Haplotype-based variant detection from short-read sequencing. 1–9. doi: arXiv:1207.3907 [q-bio.GN]
Gebremariam MM, Zarnkow M, Becker T (2012) Teff (Eragrostis tef) as a raw material for malting, brewing and manufacturing of gluten-free foods and beverages: a review. J. Food Sci. Technol. 51:2881–2895.
Grimaud F, Rogniaux H, James MG, et al (2008) Proteome and phosphoproteome analysis of starch granule-associated proteins from normal maize and mutants affected in starch biosynthesis. J Exp Bot 59:3395–3406. doi: 10.1093/jxb/ern198
Gunorubon AJ, Kekpugile DK (2012) Modification of Cassava Starch for Industrial Uses. Int J Eng Technol 2:913–919.
Hager AS, Wolter A, Jacob F, et al (2012) Nutritional properties and ultra-structure of commercial gluten free flours from different botanical sources compared to wheat flours. J Cereal Sci 56:239–247. doi: 10.1016/j.jcs.2012.06.005
Hannah LC, James M (2008) The complexities of starch biosynthesis in cereal endosperms. Curr Opin Biotechnol 19:160–165. doi: 10.1016/j.copbio.2008.02.013
Hari PK, Garg S, Garg SK (1989) Gelatinization of Starch and Modified Starch. Starch ??? St??rke 41:88–91. doi: 10.1002/star.19890410304
Hennen-Bierwagen TA, Lin Q, Grimaud F, et al (2009) Proteins from Multiple Metabolic Pathways Associate with Starch Biosynthetic Enzymes in High Molecular Weight Complexes: A Model for Regulation of Carbon Allocation in Maize Amyloplasts. PLANT Physiol 149:1541–1559. doi: 10.1104/pp.109.135293
Hennen-Bierwagen TA, Liu F, Marsh RS, et al (2008) Starch Biosynthetic Enzymes from Developing Maize Endosperm Associate in Multisubunit Complexes. PLANT Physiol 146:1892–1908. doi: 10.1104/pp.108.116285
Henry IM, Nagalakshmi U, Lieberman MC, et al (2014) Efficient Genome-Wide Detection and Cataloging of EMS-Induced Mutations Using Exome Capture and Next-Generation Sequencing. Plant Cell 26:1382–1397. doi: 10.1105/tpc.113.121590
Henson (2014) A Comparison of Barley Malt Amylolytic Enzyme Thermostabilities and Wort Sugars Produced During Mashing. J Am Soc Brew Chem. doi: 10.1094/ASBCJ-2014-0130-01
Hizukuri S, Tabata S, Kagoshima, Nikuni Z (1970) Studies on Starch Phosphate Part 1. Estimation of glucose-6-phosphate residues in starch and the presence of other bound phosphate(s). Starch - Stärke 22:338–343. doi: 10.1002/star.19700221004
Hoover R (2001) Composition, molecular structure, and physicochemical properties of tuber and root starches: A review. Carbohydr Polym 45:253–267. doi: 10.1016/S0144-8617(00)00260-5
Hovenkamp-Hermelink JHM, Jacobsen E, Ponstein AS, et al (1987) Isolation of an amylose-free starch mutant of the potato (Solanum tuberosum L.). Theor Appl Genet 75:217–221. doi: 10.1007/BF00249167
Hwang S, Kim E, Lee I, Marcotte EM (2015) Systematic comparison of variant calling pipelines using gold standard personal exome variants. Sci Rep. doi: 10.1038/srep17875

Ihemere U, Arias-Garzon D, Lawrence S, Sayre R (2006) Genetic modification of cassava for enhanced starch production. Plant Biotechnol J 4:453–465. doi: 10.1111/j.1467-7652.2006.00195.x
Imberty A, Chanzy H, Pérez S, et al (1988) The double-helical nature of the crystalline part of A-starch. J Mol Biol 201:365–378. doi: 10.1016/0022-2836(88)90144-1
Imberty A, Perez S (1988) A revisit to the three-dimensional structure of B-type starch. Biopolymers 27:1205–1221. doi: 10.1002/bip.360270803
Jaiswal S, Baìšga M, Ahuja G, et al (2014) Development of barley (Hordeum Vulgare L.) lines with altered starch granule size distribution. J Agric Food Chem 62:2289–2296. doi: 10.1021/jf405424x
James MG, Denyer K, Myers AM (2003) Starch synthesis in the cereal endosperm. Curr. Opin. Plant Biol. 6:215–222.
Jeon JS, Ryoo N, Hahn TR, et al (2010) Starch biosynthesis in cereal endosperm. Plant Physiol. Biochem. 48:383–392.
Jobling S (2004) Improving starch for food and industrial applications. Curr Opin Plant Biol 7:210–218. doi: 10.1016/j.pbi.2003.12.001
Joshi M, Aldred P, McKnight S, et al (2013) Physicochemical and functional characteristics of lentil starch. Carbohydr Polym 92:1484–1496. doi: 10.1016/j.carbpol.2012.10.035
Kamei CLA, Severing EI, Dechesne A, et al (2016) Orphan Crops Browser: a bridge between model and orphan crops. Mol Breed 36:1–18. doi: 10.1007/s11032-015-0430-2
Kammerer B, Fischer K, Hilpert B, et al (1998) Molecular Characterization of a Carbon Transporter in Plastids from Heterotrophic Tissues: The Glucose 6-Phosphate/Phosphate Antiporter. Plant Cell 10:105–117. doi: 10.2307/3870632
Kampfenkel K, Möhlmann T, Batz O, et al (1995) Molecular characterization of an Arabidopsis thaliana cDNA encoding a novel putative adenylate translocator of higher plants. FEBS Lett 374:351–5. doi: 10.1016/0014-5793(95)01143-3
Kaneko M, Itoh H, Ueguchi-Tanaka M, et al (2002) The alpha -Amylase Induction in Endosperm during Rice Seed Germination Is Caused by Gibberellin Synthesized in Epithelium. Plant Physiol 128:1264–1270. doi: 10.1104/pp.010785
Karim AA, Toon LC, Lee VPL, et al (2007) Effects of phosphorus contents on the gelatinization and retrogradation of potato starch. J Food Sci 72:132–138. doi: 10.1111/j.1750-3841.2006.00251.x
Karlström A, Calle F, Salazar S, et al (2016) Biological Implications in Cassava for the Production of Amylose-Free Starch: Impact on Root Yield and Related Traits. Front Plant Sci 7:604. doi: 10.3389/fpls.2016.00604
Kaur B, Ariffin F, Bhat R, Karim AA (2012) Progress in starch modification in the last decade. Food Hydrocoll 26:398–404. doi: 10.1016/j.foodhyd.2011.02.016
Kaur M, Oberoi DPS, Sogi DS, Gill BS (2011) Physicochemical, morphological and pasting properties of acid treated starches from different botanical sources. J Food Sci Technol 48:460–465. doi: 10.1007/s13197-010-0126-x
KIM CS, WALKER CE (1992) Changes in Starch Pasting Properties Due to Sugars and Emulsifiers as Determined by Viscosity Measurement. J Food Sci 57:1009–1013. doi: 10.1111/j.1365-2621.1992.tb14344.x
Knoll JE, Ramos ML, Zeng Y, et al (2011) TILLING for allergen reduction and improvement of quality traits in peanut (Arachis hypogaea L.). BMC Plant Biol. doi: 10.1186/1471-2229-11-81
Kotting O (2005) Identification of a Novel Enzyme Required for Starch Metabolism in Arabidopsis Leaves. The Phosphoglucan, Water Dikinase. PLANT Physiol 137:242–252. doi: 10.1104/pp.104.055954
Kötting O, Santelia D, Edner C, et al (2009a) STARCH-EXCESS4 is a laforin-like

Phosphoglucan phosphatase required for starch degradation in Arabidopsis thaliana. Plant Cell 21:334–46. doi: 10.1105/tpc.108.064360
Kötting O, Santelia D, Edner C, et al (2009b) STARCH-EXCESS4 is a laforin-like Phosphoglucan phosphatase required for starch degradation in Arabidopsis thaliana. Plant Cell 21:334–46. doi: 10.1105/tpc.108.064360
Krishna Radhika N, Sheela MN, Asha Devi A, et al (2014) Genetic modification for designer starch from cassava. J. Trop. Agric. 52:1–6.
Kurakake M, Akiyama Y, Hagiwara H, Komaki T (2009) Effects of cross-linking and low molecular amylose on pasting characteristics of waxy corn starch. Food Chem 116:66–70. doi: 10.1016/j.foodchem.2009.02.006
Kusumayanti H, Handayani NA, Santosa H (2015) Swelling Power and Water Solubility of Cassava and Sweet Potatoes Flour. Procedia Environ Sci 23:164–167. doi: 10.1016/j.proenv.2015.01.025
Lee MR, Swanson BR, Baik BK (2001) Influence of amylose content on properties of wheat starch and breadmaking quality of starch and gluten blends. Cereal Chem 78:701–706. doi: 10.1094/CCHEM.2001.78.6.701
Li H, Handsaker B, Wysoker A, et al (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079. doi: 10.1093/bioinformatics/btp352
Liang X, King JM (2003) Pasting and crystalline property differences of commercial and isolated rice starch with added amino acids. J Food Sci 68:832–838. doi: 10.1111/j.1365-2621.2003.tb08251.x
Lim S-T, Kasemsuwan T, Jane J-L (1994) Characterization of Phosphorus in Starch by 31P-Nuclear Magnetic Resonance Spectroscopy. Cereal Chem 71:488–493.
Lin Q, Huang B, Zhang M, et al (2012) Functional Interactions between Starch Synthase III and Isoamylase-Type Starch-Debranching Enzyme in Maize Endosperm. PLANT Physiol 158:679–692. doi: 10.1104/pp.111.189704
Liu H, Yu L, Xie F, Chen L (2006) Gelatinization of cornstarch with different amylose/amylopectin content. Carbohydr Polym 65:357–363. doi: 10.1016/j.carbpol.2006.01.026
Lloyd JR, Kossmann J (2015) Transitory and storage starch metabolism: Two sides of the same coin? Curr. Opin. Biotechnol. 32:143–148.
Lorberth R, Ritte G, Willmitzer L, Kossmann J (1998) Inhibition of a starch-granule-bound protein leads to modified starch and repression of cold sweetening. Nat Biotechnol 16:473–477. doi: 10.1038/nbt0598-473
Lund D (1984) Influence of time, temperature, moisture, ingredients, and processing conditions on starch gelatinization. Crit Rev Food Sci Nutr 20:249–73. doi: 10.1080/10408398409527391
Mahlow S, Orzechowski S, Fettke J (2016) Starch phosphorylation: insights and perspectives. Cell Mol Life Sci 73:2753–2764. doi: 10.1007/s00018-016-2248-4
Manners DJ (1989) Recent developments in our understanding of amylopectin structure. Carbohydr. Polym. 11:87–112.
Martin C, Smith a M (1995) Starch biosynthesis. Plant Cell 7:971–985. doi: 10.1039/qr9692300057
Maurer T, Kalbitzer HR (1996) Indirect referencing of31P and19F NMR spectra. J Magn Reson - Ser B 113:177–178. doi: 10.1006/jmrb.1996.0172
Menlove KJ, Clement M, Crandall KA (2009) Similarity searching using BLAST. Methods Mol. Biol. 537:1–22.
Mikkelsen R, Mutenda KE, Mant A, et al (2005) Alpha-glucan, water dikinase (GWD): a plastidic enzyme with redox-regulated and coordinated catalytic activity and binding affinity. Proc Natl Acad Sci U S A 102:1785–1790.

Minoia S, Petrozza A, D'Onofrio O, et al (2010) A new mutant genetic resource for tomato crop improvement by TILLING technology. BMC Res Notes. doi: 10.1186/1756-0500-3-69
Montagnac JA, Davis CR, Tanumihardjo SA (2009a) Processing techniques to reduce toxicity and antinutrients of Cassava for use as a staple food. Compr Rev Food Sci Food Saf 8:17–27. doi: 10.1111/j.1541-4337.2008.00064.x
Montagnac JA, Davis CR, Tanumihardjo SA (2009b) Nutritional value of cassava for use as a staple food and recent advances for improvement. Compr Rev Food Sci Food Saf 8:181–194. doi: 10.1111/j.1541-4337.2009.00077.x
Montero WR (2003) Cassava; Biology, Production and Utilization.
Moorthy SN (2002) Physicochemical and functional properties of tropical tuber starches: A review. Starch/Staerke 54:559–592.
Morrison WR, Law R V., Snape CE (1993) Evidence for inclusion complexes of lipids with v-amylose in maize, rice and oat starches. J Cereal Sci 18:107–109. doi: 10.1006/jcrs.1993.1039
Morrison WR, Milligan TP, Azudin MN (1984) A relationship between the amylose and lipid contents of starches from diploid cereals. J Cereal Sci 2:257–271. doi: 10.1016/S0733-5210(84)80014-4
Mun JH, Kwon SJ, Yang TJ, et al (2009) Genome-wide comparative analysis of the Brassica rapa gene space reveals genome shrinkage and differential loss of duplicated genes after whole genome triplication. Genome Biol. doi: 10.1186/gb-2009-10-10-r111
Muth J, Hartje S, Twyman RM, et al (2008) Precision breeding for novel starch variants in potato. Plant Biotechnol J 6:576–584. doi: 10.1111/j.1467-7652.2008.00340.x
NACONSIE M, ZHANG P (2016) Transgenic technologies in cassava for nutritional improvement and viral disease resistance: a key strategy for food security in Africa. Front Agric Sci Eng 3:285. doi: 10.15302/J-FASE-2016119
Nakamura T, Yamamori M, Hirano H, et al (1995) Production of waxy (amylose-free) wheats. MGG Mol Gen Genet 248:253–259. doi: 10.1007/BF02191591
Naylor RL, Falcon WP, Goodman RM, et al (2004) Biotechnology in the developing world: A case for increased investments in orphan crops. Food Policy 29:15–44. doi: 10.1016/j.foodpol.2004.01.002
Nelson RJ, Naylor RL, Jahn MM (2004) The role of genomics research in improvement of 'orphan' crops. In: Crop Science. pp 1901–1904
Niittylä T, Messerli G, Trevisan M, et al (2004) A Previously Unknown Maltose Transporter Essential for Starch Degradation in Leaves. Science (80-) 303:87–89. doi: 10.1126/science.1091811
Nyakabau T, Wokadala OC, Emmambux MN (2013) Effect of steeping additives on tef starch extptraction and its quality. Starch/Staerke 65:738–746. doi: 10.1002/star.201200241
Oleykowski C a, Bronson Mullins CR, Godwin a K, Yeung a T (1998) Mutation detection using a novel plant endonuclease. Nucleic Acids Res 26:4597–4602. doi: 10.1093/nar/26.20.4597
Opabode JT, Oyelakin OO, Akinyemiju OA, Ingelbrecht IL (2012) Starch branching enzyme I from cassava (Manihot esculenta Crantz): Genomic organization , intron identification and phylogenetic analysis. Glob J Biotechnol Biocemistry 7:79–85. doi: 10.5829/idosi.gjbb.2012.7.3.1103
Pa K, Asseng S (2018) A review of tef physiology for developing a tef crop model. 94:54–66. doi: 10.1016/j.eja.2018.01.008
Pfister B, Zeeman SC (2016a) Formation of starch in plant cells. Cell. Mol. Life Sci. 73:2781–2807.

Pfister B, Zeeman SC (2016b) Formation of starch in plant cells. Cell Mol Life Sci 73:2781–2807. doi: 10.1007/s00018-016-2250-x
Pilling E, Smith a M (2003) Growth ring formation in the starch granules of potato tubers. Plant Physiol 132:365–371. doi: 10.1104/pp.102.018044
Polnaya FJ, Haryadi, Marseno DW, Cahyanto MN (2013) Effects of phosphorylation and cross-linking on the pasting properties and molecular structure of sago starch. Int Food Res J 20:1609–1615.
Puounti-Kaerlas J (2002) Cassava Biotechnology. Genet Eng Rev 15:329–364. doi: 10.1079/9780851995243.0000
Radchuk V V., Borisjuk L, Sreenivasulu N, et al (2009) Spatiotemporal Profiling of Starch Biosynthesis and Degradation in the Developing Barley Grain. PLANT Physiol 150:190–204. doi: 10.1104/pp.108.133520
Raemakers K, Schreuder M, Suurs L, et al (2005) Improved cassava starch by antisense inhibition of granule-bound starch synthase I. Mol Breed 16:163–172. doi: 10.1007/s11032-005-7874-8
Ran FA, Hsu PD, Wright J, et al (2013) Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8:2281–2308. doi: 10.1038/nprot.2013.143
Raphael M, Yona B, Stephen K, et al (2011) Amylopectin molecular structure and functional properties of starch from three Ugandan cassava varieties. J Plant Breed Crop Sci 3:195–202.
Rathore RS, Garg N, Garg S, Kumar A (2009) Starch phosphorylase: Role in starch metabolism and biotechnological applications. Crit. Rev. Biotechnol. 29:214–224.
Ratnayake WS, Jackson DS (2008) Starch Gelatinization. In: Advances in Food and Nutrition Research. pp 221–268
Regina A, Bird A, Topping D, et al (2006) High-amylose wheat generated by RNA interference improves indices of large-bowel health in rats. Proc Natl Acad Sci 103:3546–3551. doi: 10.1073/pnas.0510737103
Rhoads A, Au KF (2015) PacBio Sequencing and Its Applications. Genomics, Proteomics Bioinforma. 13:278–289.
Ritte G, Heydenreich M, Mahlow S, et al (2006a) Phosphorylation of C6- and C3-positions of glucosyl residues in starch is catalysed by distinct dikinases. FEBS Lett 580:4872–4876. doi: 10.1016/j.febslet.2006.07.085
Ritte G, Heydenreich M, Mahlow S, et al (2006b) Phosphorylation of C6- and C3-positions of glucosyl residues in starch is catalysed by distinct dikinases. FEBS Lett 580:4872–4876. doi: 10.1016/j.febslet.2006.07.085
Rongsirikul O, Saithong T, Kalapanulak S, et al (2010) Reconstruction of starch biosynthesis pathway in cassava using comparative genomic approach. In: Communications in Computer and Information Science. pp 118–129
Saithong T, Rongsirikul O, Kalapanulak S, et al (2013) Starch biosynthesis in cassava: A genome-based pathway reconstruction and its exploitation in data integration. BMC Syst Biol. doi: 10.1186/1752-0509-7-75
Salehuzzaman SNIM, Jacobsen E, Visser RGF (1993) Isolation and characterization of a cDNA encoding granule-bound starch synthase in cassava (Manihot esculenta Crantz) and its antisense expression in potato. Plant Mol Biol 23:947–962. doi: 10.1007/BF00021811
Sangwan P, Petinakis E, Dean K (2014) Starch Polymers.
Sano Y (1984) Differential regulation of waxy gene expression in rice endosperm. Theor Appl Genet 68:467–473. doi: 10.1007/BF00254822
Santana ÁL, Angela M, Meireles A (2014) New Starches are the Trend for Industry Applications: A Review. Food Public Heal 4:229–241. doi: 10.5923/j.fph.20140405.04

Santelia D, Kötting O, Seung D, et al (2011) The phosphoglucan phosphatase Like Sex Four2 dephosphorylates starch at the C3-position in Arabidopsis. Plant Cell 23:4096–4111. doi: 10.1105/tpc.111.092155
Santelia D, Zeeman SC (2011a) Progress in Arabidopsis starch research and potential biotechnological applications. Curr Opin Biotechnol 22:271–280. doi: 10.1016/j.copbio.2010.11.014
Santelia D, Zeeman SC (2011b) Progress in Arabidopsis starch research and potential biotechnological applications. Curr. Opin. Biotechnol. 22:271–280.
Satoh H, Nishi A, Yamashita K, et al (2003) Starch-branching enzyme I-deficient mutation specifically affects the structure and properties of starch in rice endosperm. Plant Physiol 133:1111–1121. doi: 10.1104/pp.103.021527
Scheben A, Wolter F, Batley J, et al (2017) Towards CRISPR/Cas crops - bringing together genomics and genome editing. New Phytol.
Seung D, Soyk S, Coiro M, et al (2015) PROTEIN TARGETING TO STARCH Is Required for Localising GRANULE-BOUND STARCH SYNTHASE to Starch Granules and for Normal Amylose Synthesis in Arabidopsis. PLoS Biol. doi: 10.1371/journal.pbio.1002080
Shumoy H, Raes K (2017) In vitro starch hydrolysis and estimated glycemic index of tef porridge and injera. Food Chem 229:381–387. doi: 10.1016/j.foodchem.2017.02.060
Silver DM, Kötting O, Moorhead GBG (2014) Phosphoglucan phosphatase function sheds light on starch degradation. Trends Plant Sci 19:471–478. doi: 10.1016/j.tplants.2014.01.008
Singh J, Kaur L, McCarthy OJ (2007) Factors influencing the physico-chemical, morphological, thermal and rheological properties of some chemically modified starches for food applications-A review. Food Hydrocoll. 21:1–22.
Singh N, Singh J, Kaur L, et al (2003) Morphological, thermal and rheological properties of starches from different botanical sources. Food Chem. 81:219–231.
Siriwat W, Kalapanulak S, Suksangpanomrung M, et al (2012) Transcriptomic data integration inferring the dominance of starch biosynthesis in carbon utilization of developing cassava roots. In: Procedia Computer Science. pp 96–106
Slade AJ, Fuerstenberg SI, Loeffler D, et al (2005) A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. Nat Biotechnol 23:75–81. doi: 10.1038/nbt1043
Smith AM (2010) Starch and Starch Granules. In: Encyclopedia of Life Sciences.
Smith AM, Zeeman SC, Smith SM (2005) Starch degradation. Annu Rev Plant Biol 56:73–98. doi: 10.1146/annurev.arplant.56.032604.144257
Sonnewald U, Kossmann J (2013) Starches-from current models to genetic engineering. Plant Biotechnol. J. 11:223–232.
Spaenij-Dekking L, Kooy-Winkelaar Y, Koning F (2005) The Ethiopian cereal tef in celiac disease. N Engl J Med 353:1748–1749. doi: 10.1056/NEJMc051492
Sreenivasulu N, Wobus U (2013) Seed-Development Programs: A Systems Biology–Based Comparison Between Dicots and Monocots. Annu Rev Plant Biol 64:189–217. doi: 10.1146/annurev-arplant-050312-120215
Srichuwong S, Sunarti TC, Mishima T, et al (2005) Starches from different botanical sources II: Contribution of starch structure to swelling and pasting properties. Carbohydr Polym 62:25–34. doi: 10.1016/j.carbpol.2005.07.003
Sriroth K, Santisopasri V, Petchalanuwat C, et al (1999) Cassava starch granule structure-function properties: Influence of time and conditions at harvest on four cultivars of cassava starch. Carbohydr Polym 38:161–170. doi: 10.1016/S0144-8617(98)00117-9
Streb S, Delatte T, Umhang M, et al (2008) Starch Granule Biosynthesis in Arabidopsis Is

Abolished by Removal of All Debranching Enzymes but Restored by the Subsequent Removal of an Endoamylase. PLANT CELL ONLINE 20:3448–3466. doi: 10.1105/tpc.108.063487
Streb S, Zeeman SC (2012) Starch metabolism in Arabidopsis. Arab B. doi: 10.1199/tab.0160
Sun Y, Jiao G, Liu Z, et al (2017) Generation of High-Amylose Rice through CRISPR/Cas9-Mediated Targeted Mutagenesis of Starch Branching Enzymes. Front Plant Sci. doi: 10.3389/fpls.2017.00298
Sun ZT, Henson CA (1991) A quantitative assessment of the importance of barley seed alpha-amylase, beta-amylase, debranching enzyme, and alpha-glucosidase in starch degradation. Arch Biochem Biophys 284:298–305.
Svihus B, Uhlen AK, Harstad OM (2005) Effect of starch granule structure, associated components and processing on nutritive value of cereal starch: A review. Anim. Feed Sci. Technol. 122:303–320.
Szymońska J, Targosz-Korecka M, Krok F (2009) Characterization of starch nanoparticles. In: Journal of Physics: Conference Series.
Tadele, Z. and Esfeld K (2009) The Improvement of African Orphan Crops through TILLING. Atdf J 6:44–49.
Tadele Z (2009) Role of Orphan Crops in Enhancing and Diversifying Food Production in Africa. African Technol y Dev Forum J 9–15.
Tako M, Tamaki Y, Teruya T, Takeda Y (2014) The Principles of Starch Gelatinization and Retrogradation. Food Nutr Sci 5:280–291. doi: 10.4236/fns.2014.53035
Tatge H, Marshall J, Martin C, et al (1999) Evidence that amylose synthesis occurs within the matrix of the starch granule in potato tubers. Plant, Cell Environ 22:543–550. doi: 10.1046/j.1365-3040.1999.00437.x
Taylor N, Chavarriaga P, Raemakers K, et al (2004) Development and application of transgenic technologies in cassava. Plant Mol. Biol. 56:671–688.
Taylor NJ, Fauquet CM, Tohme J (2012) Overview of Cassava Special Issue. Trop. Plant Biol. 5:1–3.
Taylor S, Scott R, Kurtz R, et al (2009) A practical guide to High Resolution Melt Analysis genotyping.
Tester RF, Karkalas J, Qi X (2004) Starch - Composition, fine structure and architecture. J Cereal Sci 39:151–165. doi: 10.1016/j.jcs.2003.12.001
Tester RF, Morrison WR (1990) Swelling and gelatinization of cereal starches. I. Effects of amylopectin, amylose, and lipids. Cereal Chem 67:551–557.
Tetlow I, Emes M (2017) Starch Biosynthesis in the Developing Endosperms of Grasses and Cereals. Agronomy 7:81. doi: 10.3390/agronomy7040081
Tetlow IJ (2011) Starch biosynthesis in developing seeds. Seed Sci. Res. 21:5–32.
Tetlow IJ, Emes MJ (2011) Starch Biosynthesis in Higher Plants: The Enzymes of Starch Synthesis. In: Comprehensive Biotechnology, Second Edition. pp 47–65
Tian Z, Qian Q, Liu Q, et al (2009) Allelic diversities in rice starch biosynthesis lead to a diverse array of rice eating and cooking qualities. Proc Natl Acad Sci 106:21760–21765. doi: 10.1073/pnas.0912396106
Till BJ, Reynolds SH, Weil C, et al (2004) Discovery of induced point mutations in maize genes by TILLING. BMC Plant Biol. doi: 10.1186/1471-2229-4-12
Tomlinson K, Denyer K (2003) Starch synthesis in cereal grains. Adv. Bot. Res. 40:1–61.
Uauy C, Paraiso F, Colasuonno P, et al (2009) A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat. BMC Plant Biol 9:115. doi: 10.1186/1471-2229-9-115
Ukwuru MU and Egbonu SE (2013) Recent development in cassava-based products research. Acad J Food Res 1:1–13. doi: 10.15413/ajfr.2012.0105

Vamadevan V, Bertoft E (2015) Structure-function relationships of starch components. Starch/Staerke 67:55–68.
Van der Auwera GA, Carneiro MO, Hartl C, et al (2013) From fastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. Curr Protoc Bioinforma. doi: 10.1002/0471250953.bi1110s43
Van Harsselaar JK, Lorenz J, Senning M, et al (2017) Genome-wide analysis of starch metabolism genes in potato (Solanum tuberosum L.). BMC Genomics. doi: 10.1186/s12864-016-3381-z
Varshney RK, Ribaut JM, Buckler ES, et al (2012) Can genomics boost productivity of orphan crops? Nat. Biotechnol. 30:1172–1176.
Vasconcelos LM, Brito AC, Carmo CD, Oliveira EJ (2016) Polymorphism of starch pathway genes in cassava. 1:1–15.
Vrinten PL, Nakamura T (2000) Wheat granule-bound starch synthase I and II are encoded by separate genes that are expressed in different tissues. Plant Physiol 122:255–264. doi: 10.1104/pp.122.1.255
Wang S, Li C, Copeland L, et al (2015a) Starch Retrogradation: A Comprehensive Review. Compr Rev Food Sci Food Saf 14:568–585. doi: 10.1111/1541-4337.12143
Wang S, Li S, Liu Q, et al (2015b) The OsSPL16-GW7 regulatory module determines grain shape and simultaneously improves rice yield and grain quality. Nat Genet 47:949–954. doi: 10.1038/ng.3352
Wang X, Chang L, Tong Z, et al (2016) Proteomics Profiling Reveals Carbohydrate Metabolic Enzymes and 14-3-3 Proteins Play Important Roles for Starch Accumulation during Cassava Root Tuberization. Sci Rep. doi: 10.1038/srep19643
Waterschoot J, Gomand S V., Fierens E, Delcour JA (2015) Production, structure, physicochemical and functional properties of maize, cassava, wheat, potato and rice starches. Starch/Staerke 67:14–29.
Weber A, Servaites JC, Geiger DR, et al (2000) Identification, Purification, and Molecular Cloning of a Putative Plastidic Glucose Translocator. Plant Cell 12:787–801. doi: 10.1105/tpc.12.5.787
Weise SE, Aung K, Jarou ZJ, et al (2012) Engineering starch accumulation by manipulation of phosphate metabolism of starch. Plant Biotechnol J 10:545–554. doi: 10.1111/j.1467-7652.2012.00684.x
Wilm A, Aw PPK, Bertrand D, et al (2012) LoFreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res 40:11189–11201. doi: 10.1093/nar/gks918
Wolter A, Hager AS, Zannini E, Arendt EK (2013) In vitro starch digestibility and predicted glycaemic indexes of buckwheat, oat, quinoa, sorghum, teff and commercial gluten-free bread. J Cereal Sci 58:431–436. doi: 10.1016/j.jcs.2013.09.003
Wong KM, Suchard MA, Huelsenbeck JP (2008) Alignment uncertainty and genomic analysis. Science (80-) 319:473–476. doi: 10.1126/science.1151532
Xin Z, Li Wang M, Barkley NA, et al (2008) Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population. BMC Plant Biol. doi: 10.1186/1471-2229-8-103
Xiong JS, Ding J, Li Y (2015) Genome-editing technologies and their potential application in horticultural crop breeding. Hortic. Res. 2:
Xu X, Huang XF, Visser RGF, Trindade LM (2017) Engineering potato starch with a higher phosphate content. PLoS One. doi: 10.1371/journal.pone.0169610
Xu X, Visser RGF, Trindade LM (2014) Starch Modification by Biotechnology: State of Art and Perspectives. In: Starch Polymers: From Genetic Engineering to Green

Applications. pp 79–102
Yan CJ, Tian ZX, Fang YW, et al (2011) Genetic analysis of starch paste viscosity parameters in glutinous rice (Oryza sativa L.). Theor Appl Genet 122:63–76. doi: 10.1007/s00122-010-1423-5
Yu JK, Graznak E, Breseghello F, et al (2007) QTL mapping of agronomic traits in tef Eragrostis tef (Zucc). BMC Plant Biol. doi: 10.1186/1471-2229-7-30
Yu S, Ma Y, Sun DW (2009) Impact of amylose content on starch retrogradation and texture of cooked milled rice during storage. J Cereal Sci 50:139–144. doi: 10.1016/j.jcs.2009.04.003
Yu TS, Kofler H, Häusler RE, et al (2001) The Arabidopsis sex1 mutant is defective in the R1 protein, a general regulator of starch degradation in plants, and not in the chloroplast hexose transporter. Plant Cell 13:1907–1918. doi: 10.1105/TPC.010091
Yu TS, Zeeman SC, Thorneycroft D, et al (2005) ??-amylase is not required for breakdown of transitory starch in Arabidopsis Leaves. J Biol Chem 280:9773–9779. doi: 10.1074/jbc.M413638200
Zeeman SC (2002) Starch Synthesis in Arabidopsis. Granule Synthesis, Composition, and Structure. PLANT Physiol 129:516–529. doi: 10.1104/pp.003756
Zeeman SC, Kossmann J, Smith AM (2010a) Starch: its metabolism, evolution, and biotechnological modification in plants. Annu Rev Plant Biol 61:209–234. doi: 10.1146/annurev-arplant-042809-112301
Zeeman SC, Kossmann J, Smith AM (2010b) Starch: its metabolism, evolution, and biotechnological modification in plants. Annu Rev plant … 61:209–234. doi: 10.1146/annurev-arplant-042809-112301
Zeeman SC, Northrop F, Smith AM, Ap Rees T (1998) A starch-accumulating mutant of Arabidopsis thaliana deficient in a chloroplastic starch-hydrolysing enzyme. Plant J 15:357–365. doi: 10.1046/j.1365-313X.1998.00213.x
Zhang L, Häusler RE, Greiten C, et al (2008) Overriding the co-limiting import of carbon and energy into tuber amyloplasts increases the starch content and yield of transgenic potato plants. Plant Biotechnol J 6:453–464. doi: 10.1111/j.1467-7652.2008.00332.x
Zhao SS, Dufour D, Sánchez T, et al (2011) Development of waxy cassava with different Biological and physico-chemical characteristics of starches for industrial applications. Biotechnol Bioeng 108:1925–1935. doi: 10.1002/bit.23120
Zhou W, Yang J, Hong Y, et al (2015) Impact of amylose content on starch physicochemical properties in transgenic sweet potato. Carbohydr Polym 122:417–427. doi: 10.1016/j.carbpol.2014.11.003
Zhu F (2015) Composition, structure, physicochemical properties, and modifications of cassava starch. Carbohydr. Polym. 122:456–480.
Zhu F (2017) Barley Starch: Composition, Structure, Properties, and Modifications. Compr Rev Food Sci Food Saf 16:558–579. doi: 10.1111/1541-4337.12265
Zhu F (2018) Chemical composition and food uses of teff (Eragrostis tef). Food Chem. 239:402–415.
Zhu T, Jackson DS, Wehling RL, Geera B (2008) Comparison of amylose determination methods and the development of a dual wavelength iodine binding technique. Cereal Chem 85:51–58. doi: 10.1094/CCHEM-85-1-0051
Zhu W, Buell CR (2007) Improvement of whole-genome annotation of cereals through comparative analyses. Genome Res 17:299–310. doi: 10.1101/gr.5881807