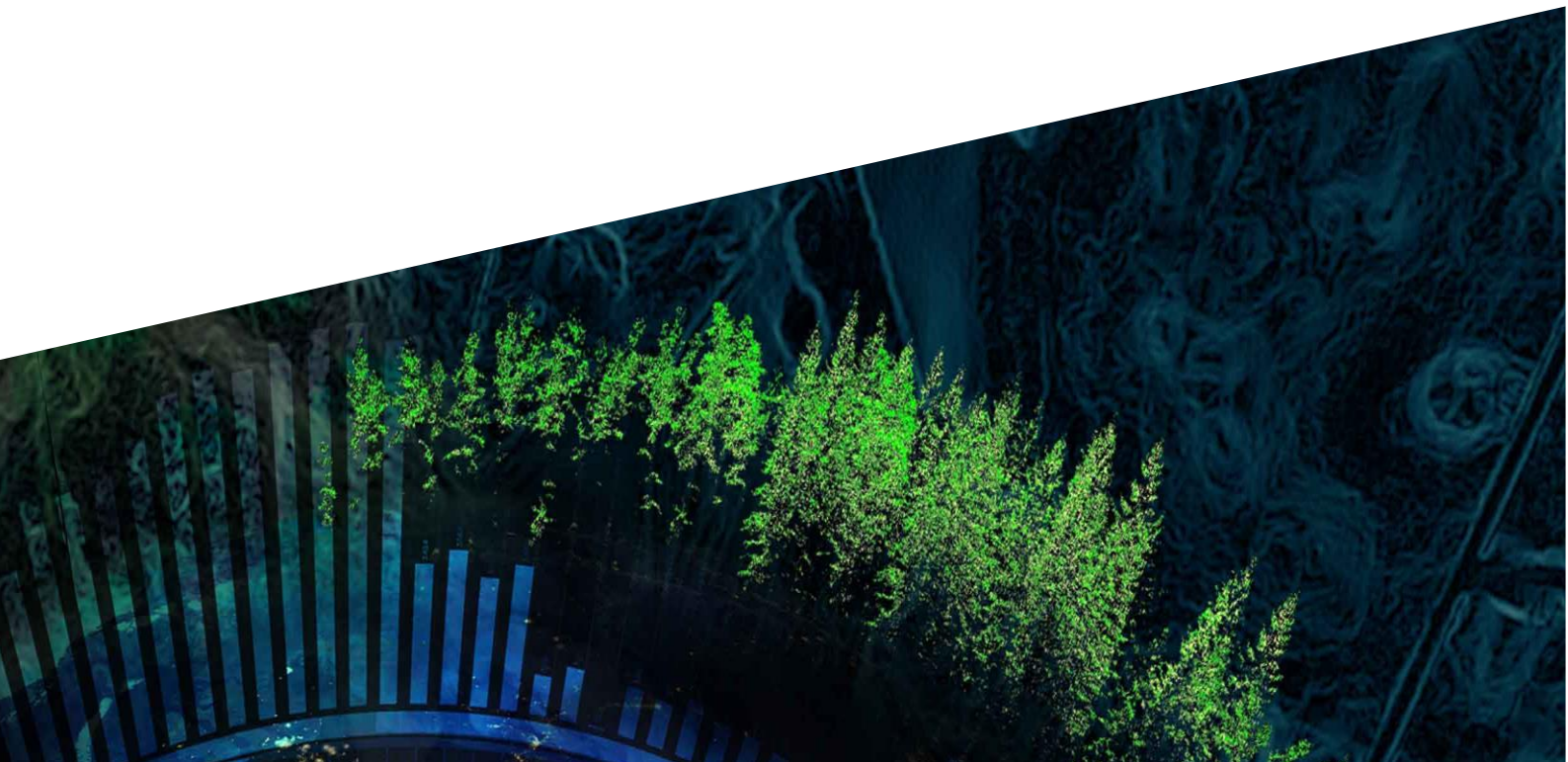
Identification and expression of genes involved in the arabinoxylan biosynthetic pathway in oat

Johan Valeur

Master thesis • 30 credits
Swedish University of Agricultural Sciences, SLU
Faculty of Landscape Architecture, Horticulture and Crop Production Science
Department of Plant Breeding
Horticultural Science Programme
Alnarp 2022



Identification and expression of genes involved in the arabinoxylan biosynthetic pathway in oat
Identifiering och uttryck av gener som är involverade I Ababinoxylans biosyntetiska bana.

Johan Valeur

Supervisor: Erik Alexandersson, SLU, Department of Plant Protection Biology
Supervisor: Jose Alfredo Zambrano, Lund University, ScanOats
Examiner: Selvaraju Kanagarajan, SLU, Department of Plant Breeding

Credits: 30 credits
Level: Second cycle, A2E, master's thesis
Course title: Independent Project in Biology, A2E
Course code: EX0856
Programme/education: Horticultural Science Programme

Course coordinating dept: Department of Plant Breeding
Place of publication: Alnarp
Year of publication: 2022
Copyright: All featured images are used with permission from the copyright owner.

Keywords: *Arabinoxylan, Oat, Gene expression analysis*

Swedish University of Agricultural Sciences
Faculty of Landscape Architecture, Horticulture and Crop Production Science
Department of Plant Breeding

Abstract

Oat is one of the most cultivated cereal crops in the world, and Arabinoxylan (AX) is a dietary fiber in oats with positive health effects. The biosynthetic pathway AX in Oat is still unknown, and which genes are active in its production. Ortholog genes have been identified in oat from other species like rice, wheat, barley, and Arabidopsis. These genes include *GT43, GT47, GT61, UAM1,* and *UAM2.* Expression analysis of three oat mutant populations was made in the early, middle, and late developmental stages. Quantifying of AX was also made in the reference oat variety Belinda in 4 different developmental stages. All genes analyzed were expressed in oat with a higher upregulation in the earlier stages except *GT61*. Accumulation of AX had a similar pattern, suggesting that AX accumulation and gene expression is higher in its first developmental stages. A phylogenetic tree was built using the protein sequences of *GT43, GT47, GT61, UAM1,* and *UAM2* in *Arabidopsis thaliana* (At), *Avena sativa* (As), *Oryza sativa* (Os), *Triticum aestivum* (Ta), and *Hordeum vulgare* (Hv) by the maximum likelihood method. The phylogenetic tree showed homology between the genes in the different species. Another phylogenetic tree was also built for the genes identified in oat and showed a clear relation between the genes.

Keywords: Arabinoxylan, Oat, Gene expression analysis

Table of contents

Abbreviations **6**

1. Introduction **7**

1.1 Arabinoxylan pathway 8

1.1.1 Aim of project 10

2. Method **11**

2.1 The mutant population of CropTailor 11

2.2 Gene expression analysis 11

2.2.1 RNA extraction 12

2.2.2 Quantification of RNA concentration 13

2.2.3 Verification of RNA integrity 13

2.2.4 DNase treatment 14

2.2.5 cDNA synthesis 14

2.2.6 RT-qPCR analysis 14

2.3 Analysis of AX content in different developmental stages in oat 15

2.3.1 Cultivation of Oat in greenhouse 15

2.3.2 Quantification of AX in Belinda 16

2.3.3 Calculations 16

2.3.4 Evolutionary map 17

3. Results and discussion **18**

3.1 Gene expression analysis 18

3.2 Bioinformatic analysis 22

3.2.1 Evolutionary map of orthologs 22

3.2.2 Evolutionary map of genes in oat 23

3.2.3 Quantification of Arabinoxylan Belinda 24

3.3 Conclusion 25

3.4 Further research 25

References **26**

Popular science summary **28**

Acknowledgements **29**

Abbreviations

AX	Arabinoxylan
CT	CropTailor
GT	Glycosyltransferase
LU	Lund University
SLU	Swedish University of Agricultural Sciences
UAM	UDP-arabinopyranose mutase

1. Introduction

Oat (*Avena sativa*) is one of the most cultivated cereal crops in the world, being ranked number six, following wheat, maize, rice, barley, and sorghum. It is an annual plant that can be sown in either fall or spring. Oat is a plant from the Poaceae family which can occur at a diploid and tetraploid level, but the primarily cultivated oat is hexaploid, with seven base chromosomes (Kole 2006).

Oat is a resilient crop adapted to moist and cool climates with acidic soil, making it fit for the Scandinavian environment. Although, because of the harsh winters in Scandinavia, Oat is often cultivated during shorter periods, such as spring (Mushtaq *et al*. 2014). Oat has been steadily rising in popularity as human food, with its protein and dietary fiber quality highlighted as two of its main perks of consumption (Mushtaq *et al*. 2014). One of these soluble dietary fibers is Beta-glucan, which has been shown to have cholesterol-lowering effects (Hughes & Grafenauer 2021).

Another dietary fiber called Arabinoxylan (AX) is hemicellulose found in the oats' primary and secondary cell walls. It is present in many kinds of cereal, like oat, wheat, and barley. AX has been shown to affect cardiovascular health positively and has possible uses as prebiotics and immunomodulators (Mendis & Simsek 2014). Research shows AX significantly improves glycaemic index in people with type 2 diabetes. Although to reach this level of effect, the dietary intake of fiber needs to be up to 50g/day, which can be challenging to implement into western diets. Therefore it can be supplemented into a healthy western diet with a purified dietary fiber supplement to make it a viable option for consumers (Lu *et al*. 2004).

The general structure of AX consists of a main chain of β-(1−4)-Xylp Units that can be monosubstituted by α-L-Ara*f* residues at the O-2 or O-3 position. Still, it can also be disubstituted at O-2 and O-3. The ratio between arabinose and xylose has been found to vary between different extracts. The positions of substituents over the xylan backbone can also vary depending on which population the AX was sourced from and which part of the plant it was extracted (Tian *et al*. 2015).

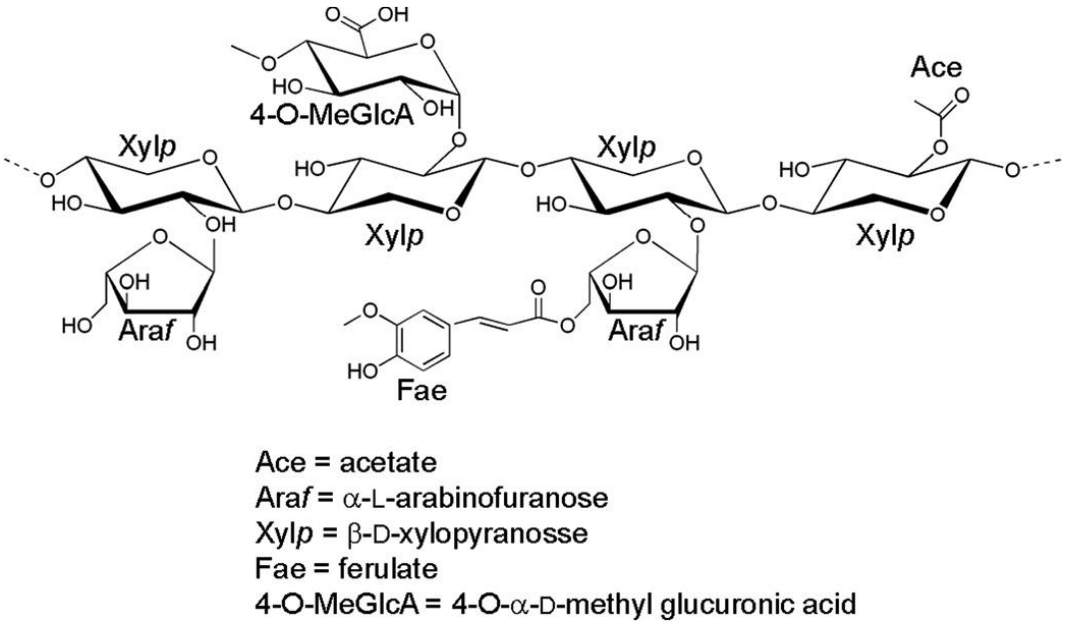


Figure 1. The structure of arabinoxylan with its substitutes is connected to the xylan backbone (Correia et al. 2011)

Arabinoxylan can be either water-soluble or non-soluble (Zhang *et al*. 2010). Stress and similar condition can affect the structure of AX. In a study where wheat was exposed to high temperatures and drought, it was possible to see an increase in water-soluble AX, although the reason was unclear (Zhang *et al*. 2010). Abiotic stress and genetic variation can impact AX content and solubility, and both have a relative impact depending on which genotype and environment are compared (Lafiandra *et al*. 2014).

Soluble AX has been shown to be functional in products like bread quality and promote digestive health by increasing water content and lowering the retention time in the intestine. The ratio of arabinose to xylose affects the structure of the molecule due to how it affects the hydrogen bonds, which play a role in the spatial arrangement of AX. Depending on the extraction method (water, alkaline, or enzymes), the structure of AX extracted can also vary (Chen *et al*. 2019).

1.1 Arabinoxylan pathway

The biosynthetic pathway of arabinoxylan has not been fully described yet. Still, several genes responsible for AX synthesis have been identified in other species like *Arabidopsis thaliana*, wheat (*Triticum aestivum*), barley (*Hordeum vulgare*), and rice (*Oryza glaberrima*). Previous research from Lund University has identified genes from the Glycosyltransferase and UDP-arabinopyranose families and suggested they may take part in synthesizing arabinoxylan.

By repressing the gene expression of the UDP-arabinopyranose family in rice by RNAi, it was shown to be responsible for forming the thermostable UDP-Arap into UDP-L-Afra*f*, which acts as a substitution to the xylan backbone by the UDP-arabinopyranose mutase (*UAM*). The transgenic plants had decreased amounts of Ada*f* and a smaller extent of substitutions of the xylan backbone in arabinoxylan. The repression of the gene *OsUAM1* was shown to have a dominant effect on the biosynthetic pathway of UDP-L-Ara*f* (Konishi *et al*. 2011).

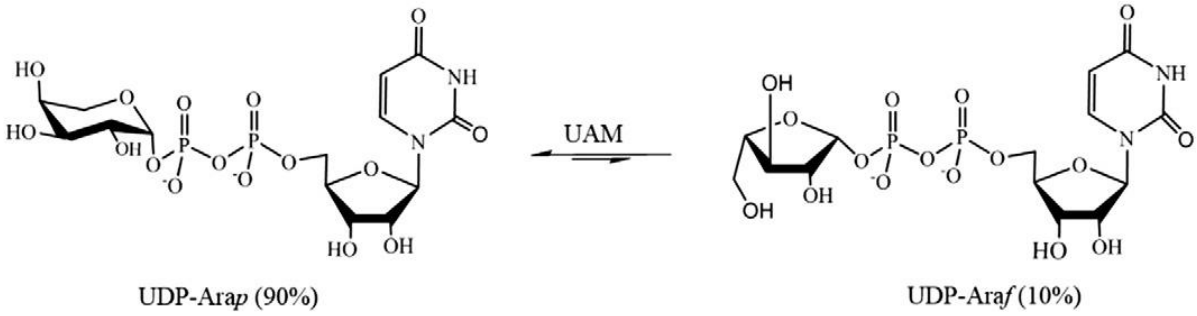


Figure 2 Reaction catalyzed by UAM enzyme (Kuttiyatveetil & Sanders 2017)

The glycosyltransferase (GT) gene family was identified in *Arabidopsis thaliana* (At), where IRX9, IRX14 from GT43, and IRX10 from GT47 were identified as candidates forming the xylan backbone. RNAi suppression of *GT43* and *GT47* in At resulted in reduced arabinoxylan. *GT61*'s diverse family members may also be responsible for several side-chain structures (Jensen *et al*. 2013). RNAi suppression of *GT61* in wheat resulted in a reduced total amount of soluble and non-soluble AX (Lafiandra *et al*. 2014). *GT61* has further been identified as a candidate to play a vital role in adding arabinosyl substituents to the xylan backbone in wheat (Anders *et al*. 2012)

A study tracked the amount of AX in different developmental stages of *Plantago ovata*. The research has shown the gene expression patterns connected to AX at different developmental stages of the seed. An accumulation pattern of xylan and arabinose was observed when the ovule reached a particular size after pollination. The gene expression analysis showed a 5-8 fold increase in the expression of UDP-arabinopyranose mutase and UDP-xylosyltransferase 2 after 15 days from pollination (Gupta *et al*. 2018).

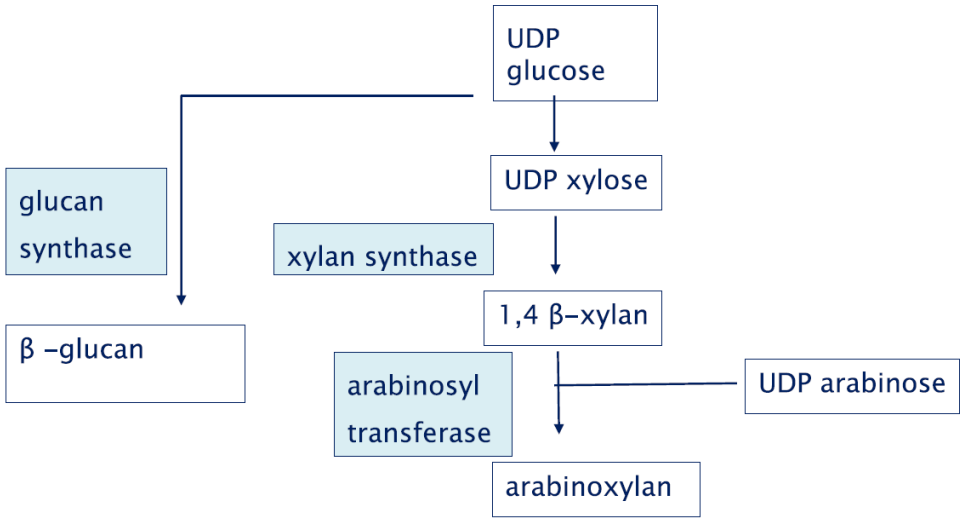


Figure 3 Proposed pathway of AX biosynthesis. Key enzymes are labeled in blue boxes (Lafiandra et al. 2014)

CropTailor AB, a commercial company in Lund, Sweden, has generated a mutagenized oat population of ca 2600 different lines by chemically induced random mutations. The company has also recently sequenced the entire Belinda oat genome at a 270 redundancy (Kamal *et al*. 2022). The biosynthetic pathway of AX in oat is currently unknown. Eight ortholog genes related to *GT43*, *GT47*, *GT61*, *UAM1*, and *UAM2* have been identified in the oat genome of rice, wheat, barley, and Arabidopsis by using the genomic database from CropTailor. Primers to these genes have been developed at Lund University to further research the connection between them and the biosynthetic pathway of AX in Oat.

1.1.1 Aim of project

Understand the biosynthesis and regulation of Arabinoxylans (AX) in oat at different developmental stages.

2. Method

2.1 The mutant population of CropTailor

CropTailor (CT) is a Swedish plant biotech company that specializes in oat breeding and have developed a sizeable mutant population of oat. The mutant population was produced by inducing random mutations in the elite cultivar Belinda by treating seeds with the mutagenic compound: ethyl methanesulfonate (EMS) (Chawade *et al*. 2010). EMS caused random mutations in the Belinda genome, making it hard to determine what causes specific changes in the new mutants' traits. Genomic mutations cause changes in regulatory regions, promoter sequences, silent mutations, knock-out genes, junk DNA and changes in the gene composition (Chawade *et al*. 2010). Chemical-induced random mutation is a technique that has been approved to modify genes for commercial use in new varieties within the EU and is not considered a GMO (Hundleby & Harwood 2019). A commercial company like CropTailor benefits from using random mutagenesis even though dealing with the uncertainty of what changes occurred in the genome that caused the new traits can be a troublesome ordeal. A commercial company may take this approach because new mutagenesis varieties can be taken to the market, unlike a type made by the more accurate gene-editing technique like CRISPR Cas9 (Hundleby & Harwood 2019).

2.2 Gene expression analysis

In order to measure the gene expression of a specific gene during a particular time, plant tissue can be gathered and flash-frozen to preserve the RNA, which is then extracted, cleaned from residue DNA, and converted into complementary DNA (cDNA). The cDNA can then be run through a real-time quantitative polymerase chain reaction (RT-qPCR) to measure which gene has the highest expression. The result from the RT-qPCR tells us which gene is expressed and how much relative to a reference gene. A reference gene has a consistent expression throughout the plant's lifetime and is always used with the genes of interest to analyze the gene expression (Bustin *et al*. 2009). This project used this technique to measure the gene

expression in the CT-oat populations Belinda (Reference), CT-high1, CT-high2, and CT-low. Oat was harvested at three developmental stages of cultivation: early, middle and late stages. In the early stage, the leaf was harvested as a sample and the kernel in the middle and late stages.
The genes chosen for analysis from the glycosyltransferase (*GT*) family were *GT61-1*, *GT61-2*, *GT47-1*, *GT47-2*, *GT43-1*, and *GT43-2*. UAM-1 and UAM-2 were chosen from the UDP-arabinopyranose family, and Tubulin-α was selected as a reference gene. Previous research at Lund university developed primers designed for each target gene; these were also used in this project.

2.2.1 RNA extraction

To extract RNA, the RNeasy Plant Mini Kit was used. The principal function of the kit is to break the cells with lysate and filter out rough plant tissue through a filter membrane. Then add ethanol to precipitate the RNA and run it through a second membrane that traps the RNA in its filter so it can be eluted with Milli-Q water from the spin column into a 2 ml tube and stored at -80°C.

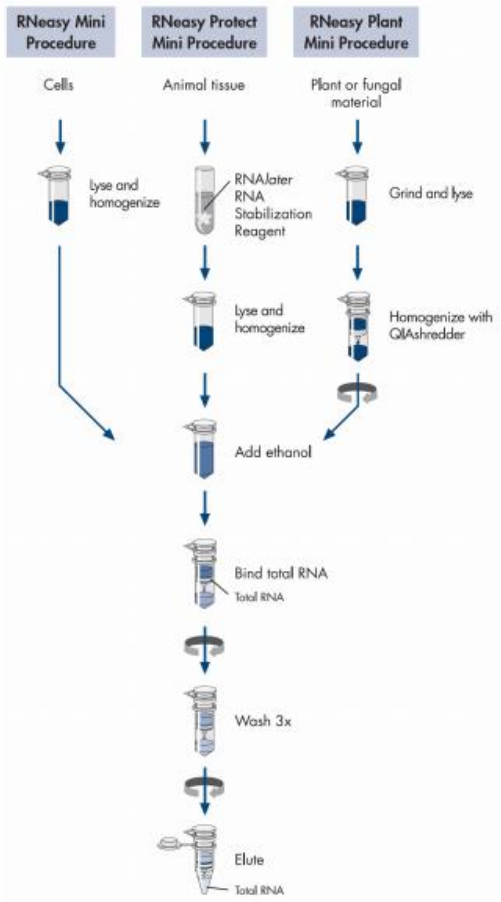


Figure 4 Visual instructions for RNA extraction (RNeasy Mini Handbook - (EN) - QIAGEN 2019)

Three samples of each stage (early, middle, and late) were extracted for every variety (Belinda, CT-high1, CT-high2, and CT-low).
100 mg of each ground sample frozen at -80°C was put in a 2 ml microcentrifuge tube, and 450 µl buffer RLT (lysis) was added and vortexed vigorously. The samples were incubated for three minutes at 56°C. The lysate was transferred to a QIAshredder spin column, placed in a 2 ml collection tube, and centrifuged at 12000 RCF for 2 min. The supernatant was carefully transferred into a new microcentrifuge 1.5 ml tube without disturbing the cell-debris pellet in the collection tube. 225 µl of 99,5% ethanol was added and mixed immediately by pipetting. 650 µl of the sample was later transferred to an RNeasy spin column and placed in a 2 ml collection tube. The RNeasy was then centrifuged for 15s at 8000 RCF. The flow-through was discarded. 700 µl of RW1 buffer was added to the RNeasy spin column and

centrifuged at 8000 RCF for 15 seconds. The flow-through was discarded. 500 µL was added to the RNeasy spin column and centrifuged at 8000 RCF for 15 seconds. The flow-through was discarded. This step was once again repeated but with centrifugation for 2 min. The RNeasy spin column was placed in a new 1.5 ml collection tube and centrifuged at 12000 RCF for 1 min. The RNeasy spin column was placed in a new 1.5 ml collection tube, and 50 µl of Milli-Q water was added to the column and left for 5 min before being centrifuged at 8000 RCF for 1 min. The RNeasy column was discarded, and the 1.5 ml collection tube with the supernatant was stored at -80°C.

2.2.2 Quantification of RNA concentration

A Nanodrop 3000 spectrophotometer was used to measure the RNA concentration in each sample. A 1:1 solution of RNA and TE buffer was used with the Nanodrop to generate accurate values with the equipment. A 1:1 H_2O and RNA solution was used as a blank to calibrate the machine.

2.2.3 Verification of RNA integrity

RNA integrity was confirmed by running the extracted RNA solution through a 1% agarose gel stained with GelRed. Although most of the visible RNA is rRNA since the mRNA of interest is too small, it is a positive sign of the general RNA integrity when clear bands are visible.

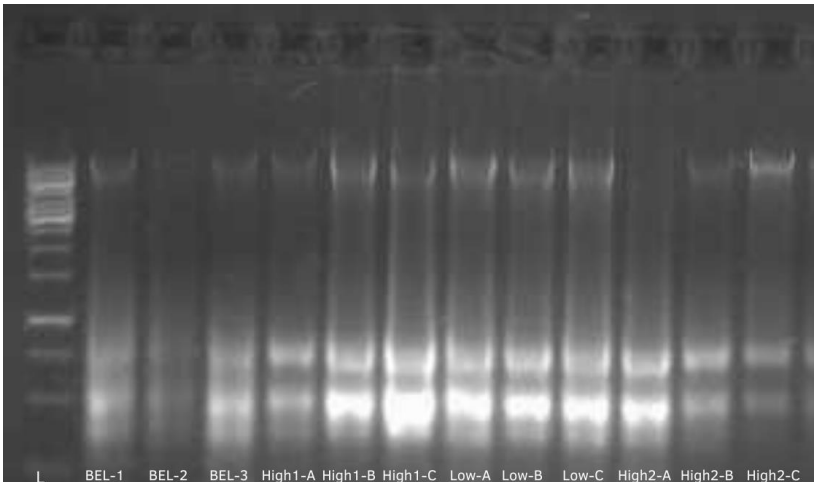


Figure 5. Clear and visible bands of separated RNA from DNA reference latter (L), Belinda (BEL), High1, High2 and Low with three biological replicates each, labeled A, B and C.

2.2.4 DNase treatment

The Ambion® DNA-free™ DNase Treatment and Removal Reagents were used in this step to destroy the remaining DNA in the samples. In order to get a concentration of 500 ng/10 µl of RNA for later steps, the volume of each solution was calculated using a simple formula of $500ng/RNAc = Volume$.
1 µl of DNase I Buffer and 1 µl DNase I was added to a 1.5 ml collection tube. In order to get a total sample volume of 10 µl, MilliQ water was added. The sample was incubated for 30 min, and 2 µl of resuspended DNase Inactivation Reagent was added. The samples were incubated for 2 min at room temperature and mixed occasionally. After this, the samples were centrifuged at 10 000 RCF for 1.5 min, and the supernatant RNA was transferred to fresh PCR tubes.

2.2.5 cDNA synthesis

The iScript™ cDNA Synthesis Kit by BIO-RAD was used in this step.
4 µl of 5x iScript Reaction Mix, 1 µl iScript Reverse Transcriptase, and 5 µl MilliQ water was added to the samples of 10 µl containing 500ng RNA in PCR tubes from the previous step. The samples were incubated in a thermal cycler using the following protocol:

Table 1. Thermocycler protocol was used for cDNA synthesis with The iScript™ cDNA Synthesis Kit by BIO-RAD. Protocol as suggested by BIO-RAD

Priming	50 min at 25°C
Reverse transcription	20 min at 46°C
RT inactivation	1 min at 95°C
Cool off	10min at 4°C

2.2.6 RT-qPCR analysis

Two brands of SYBR green were used in the qPCR: Platinum® SYBR® Green qPCR SuperMig-UDG with ROX (SYBR-UDG) and KAPA SYBR® FAST One Step for LightCycler® 480 (SYBR-FAST) due to lack of supply.

Tubulin-α was used as a reference in the qPCR. Three replicates of 20 µl reaction mix of each three samples were loaded for every nine genes in a 96 well plate.

All early, late-stage CT-low and CT-high2 samples were run with SYBR-UDG, and the remaining samples were run with SYBR-FAST.

Table 2. Reaction mix used for qPCR run. The mix contains SYBR Green from Bio-Rad or Thermo fisher, forward primer, reverse primer, MilliQ water, and cDNA.

SYBR Green reaction mix	10 µl
Forward primer	1,2 µl (10nM)
Reverse primer	1,2 µl (10nM)
MilliQ water	3,6 µl
cDNA	4 µl (diluted 1:29)
Total volume:	20 µl

Table 3. Genes found in Avena sativa (As) from previous research at Lund University. Two genes from each family were analyzed in the RT-qPCR.

AsGT61-1	*AsGT47-1*	*AsGT43-1*	*AsUAM1*
AsGT61-2	*AsGT47-2*	*AsGT43-2*	*AsUAM2*

2.3 Analysis of AX content in different developmental stages in oat

2.3.1 Cultivation of Oat in greenhouse

Seeds from the wild-type oat variety Belinda were sown with potting soil in three pots with three seeds in each in the biology greenhouse at Lund university. The plants were manually irrigated and exposed to both natural and artificial light. Harvesting of seeds was done in four stages: A (Early), E (Middle), H (Late), and J (Last), see figure 6. Seed samples harvested at stages A and E were flash-frozen immediately after harvest in liquid nitrogen and stored at -80°C. All samples were freeze-dried for 24h at a pressure of 0.05 millibar.

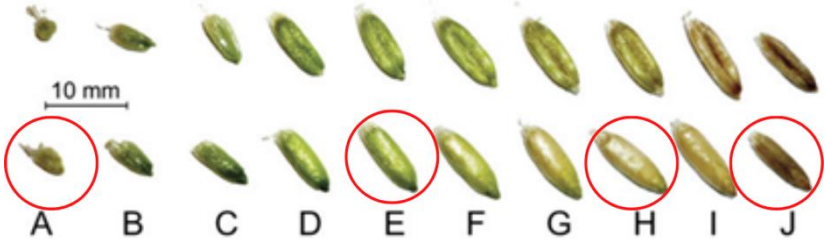


Figure 6. Different developmental stages of the Oat kernel are divided from A to J, with A, E, H, and J marked with a red circle to indicate which stage was used in the experiment.

2.3.2 Quantification of AX in Belinda

The method starts by solubilizing the fiber (polysaccharides) sample in order to break down the branched starch into mono sugars by hydrolyzation using a strong acid under high pressure. The samples are then neutralized and filtrated. AX is later quantified using High-Performance Anion-Exchange Chromatography (HPAEC).

Seeds from every developmental stage were milled in a Pulverisette for three minutes at 50 Hz. 25 mg of each sample was weight into autoclavable glass tubes. 250 µl of 72% sulfuric acid was added to each sample and incubated at 30°C in a shaker at 200 Hz for 30 min. 6,8 ml MilliQ water was added to each sample and incubated in an autoclave at 121 °C for 1 hour. All samples were vortexed vigorously, and 2 ml was transferred to 2 ml tubes and centrifuged for 10 min at 4500 rpm. 1 ml of the supernatant was transferred to 15 ml tubes and neutralized by adding 0,1 M $Ba(OH)_2$ stepwise. Gravitational pull was used to separate supernatant from the precipitate. The supernatant was transferred to 2 ml tubes.
All samples were diluted in new 2 ml tubes, and five standard dilutions with the sugars arabinose, galactose, glucose, and xylose were made, see table 4. Every sample was filtered through a 0,2 µm syringe filter into HPLC tubes and loaded into a Dionex IC-5000 using a Dionex CarboPac PA-20 analytical column coupled to a Dionex CarboPac PA-20 guard column diluted according to table 4 with standards 5 to 40.

Table 4. Dilutions of AX samples and standards from 5 to 10 with sugar solution.

Sample dilution	**AX sample**	**MilliQ H_20**
SD 1-24	50 µl	950 µl
Standards	**Sugar solution**	**MilliQ H_20**
STD 5	10 µl	1990 µl
STD 10	20 µl	19980 µl
STD 20	40 µl	1960 µl
STD 30	60 µl	1970 µl
STD 40	80 µl	1920 µl
BLANK	0 µl	2000 µl

2.3.3 Calculations

AX content was calculated as the sum of xylose, arabinose with a correcting factor of 0,88 plus galactose with a correcting factor of 0,9 due to the anhydrous sugars (Döring et al. 2016)

$$m_{xylan} = \left(m_{arabinose} + m_{xylose}\right)0.88 + m_{galactose} * 0{,}9$$

2.3.4 Evolutionary map

Protein sequences were found by searching for *GT61*, *GT47*, *GT43*, *UAM1*, and *UAM2* in the genome database Plants.ensembl.org. This database was used to find sequences in rice, wheat, and barley. The protein sequences were found in *Arabidopsis thaliana* by searching in the genome database Arabidopsis.org.

The software MEGA11 was used to produce a phylogenetic tree by aligning all protein sequences with the MUSCLE function, and the phylogenetic tree was constructed using the Maximum likelihood method.

All primer sequences from the primers used in the gene expression analysis from previous research at Lund University were also made into a phylogenetic tree, using the same method as previously stated.

3. Results and discussion

3.1 Gene expression analysis

The data generated by the expression analysis shows an expression of every tested gene. This indicates that the ortholog genes related to AX synthesis identified in barley, wheat, and rice are also present and active in oat.
The results from analyzing *GT43-2* in CT-high1 (see figure 7) show an upregulated gene expression in the early stage of development compared to Belinda. This could indicate that the gene is upregulated in the beginning to start the elongation of the xylan backbone and decrease at later stages. *GT47* has been liked with the xylan synthesis, and both of these genes have an upregulated expression at the beginning of the developing stages, see figures 7 and 8 (Brown *et al*. 2009; Wu *et al*. 2009, 2010).

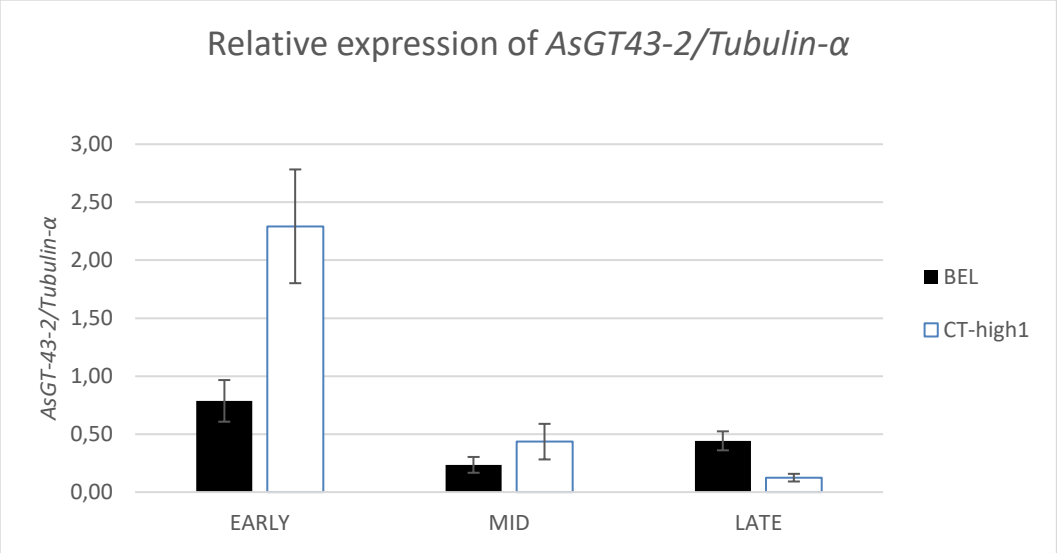


Figure 7. Relative expression of AsGT43-2/Tubulin-α. AsGT43-2 is highly upregulated at the early stage compared to Belinda and is significantly less active later during the middle and late stages. Error bars represent the standard deviation (SD) of three biological replicates.

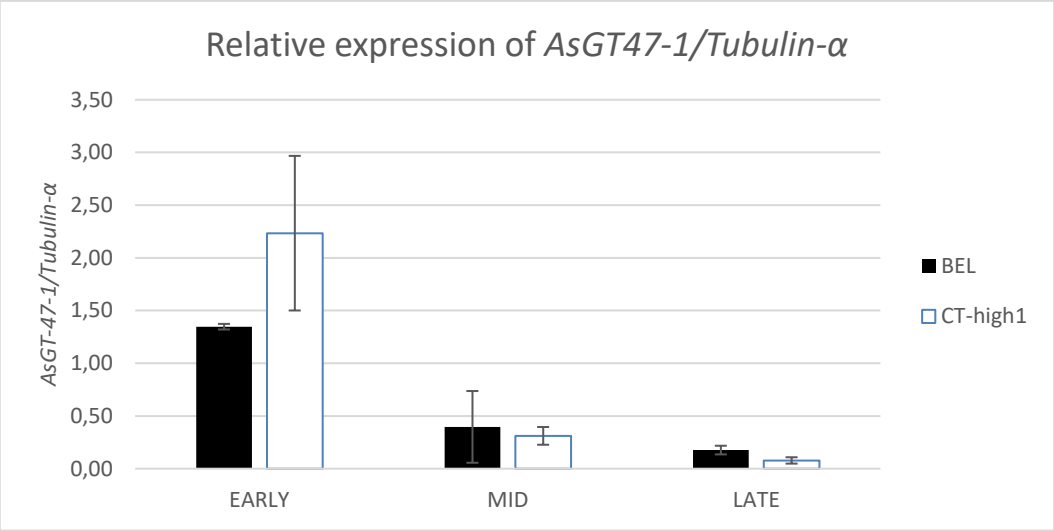


Figure 8. Relative expression of AsGT47-1/Tubulin-α at three developmental stages. AsGT47-1 is significantly upregulated at the early stage and then decreases in activation throughout the middle and late stages. Error bars represent the standard deviation (SD) of three biological replicates.

GT61-2 in CT-high1 is noticeably low in the early stage compared to Belinda and then increases in the middle and late stage, see figure 9. The *GT61* has previously been shown in other studies to be responsible for expressing enzymes to produce substituents for the xylan backbone (Jensen *et al*. 2013). In accordance with this, an upregulation of *GT61-2* in the stage after *GT43-2* and *GT47-2* have been upregulated may show a possible feedback regulation between the genes. This theory is further strengthened by the research suggesting that *GT43* and *GT47* are vital for xylan backbone production, and *GT61* plays a part in the substitution of the backbone (Jensen *et al*. 2013; Lafiandra *et al*. 2014).

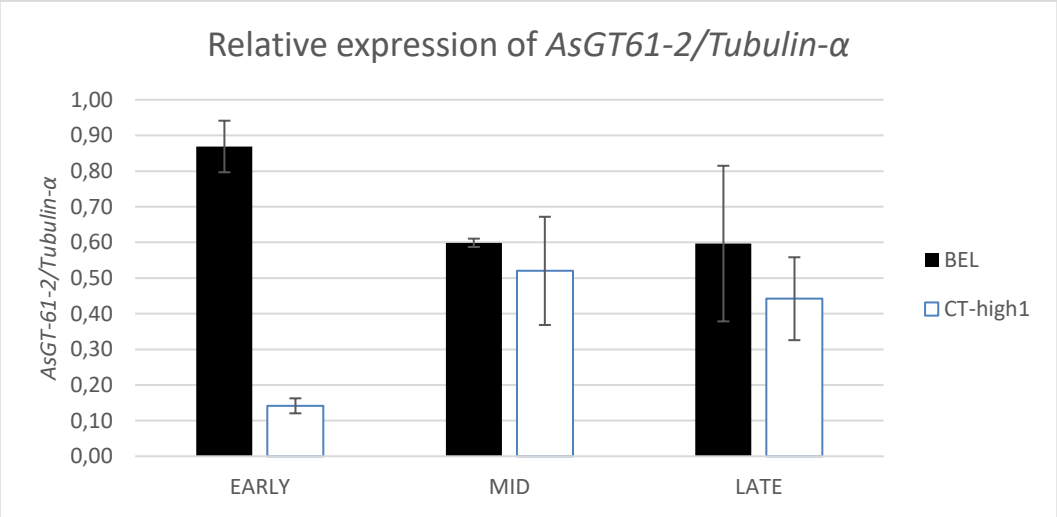


Figure 9 Relative expression of AsGT61-2 at three developmental stages. AsGT61-2 has a higher upregulation than Belinda in the middle and late-stages and significantly lower activation in the early stage. Error bars represent the standard deviation (SD) of three biological replicates.

The same pattern of early upregulation for genes related to AX synthesis can be seen in *UAM2*, see figure 10. This pattern may suggest that the genes responsible for producing arabinose residues for the xylan backbone are upregulated at the same stage as the other genes related to the xylan backbone synthesis. The substitutions produced in the early stage may play a role for *GT61* in the middle and late stages.

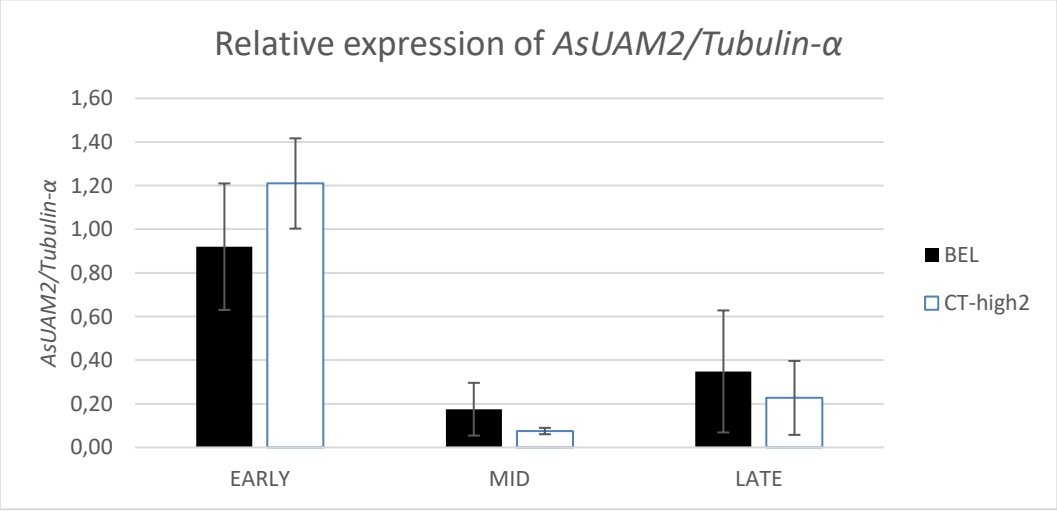


Figure 10. Relative expression of AsUAM2/Tubulin-α in three developmental stages. AsGT43-1 is more upregulated than Belinda in the early stage and later declines in the middle and late stage. Error bars represent the standard deviation (SD) of three biological replicates.

The results show varying gene regulation between the different mutants. This variation is expected since some of the mutations causing a higher or lower AX may be situated at other parts of the genome, which may affect promoter regions,

suppressors, and similar factors. Random mutagenesis needs to be considered when comparing the results from Belinda, CT-low, CT-high1, and CT-high2. The gene expression varies between all the mutant lines at different intervals. All of the mutant varieties have random mutations scattered throughout the genome, so it is impossible to draw an accurate conclusion. It was also impossible to find any coherent gene expression in the remaining gene expression analyzed, and this may be because of the random gene expression that mutants have. In order to be able to claim more extensively which genes cause what, a more exhaustive gene expression analysis with a larger sample of genes is needed. If the genetic difference from the reference variety were a knock-out mutation, it would be easier to understand better which gene regulates each process of the biosynthetic pathway of AX.

Many more genes could be analyzed in the future to understand the biosynthetic pathway of AX further, but this was not possible during this thesis due to time constraints.

3.2 Bioinformatic analysis

3.2.1 Evolutionary map of orthologs

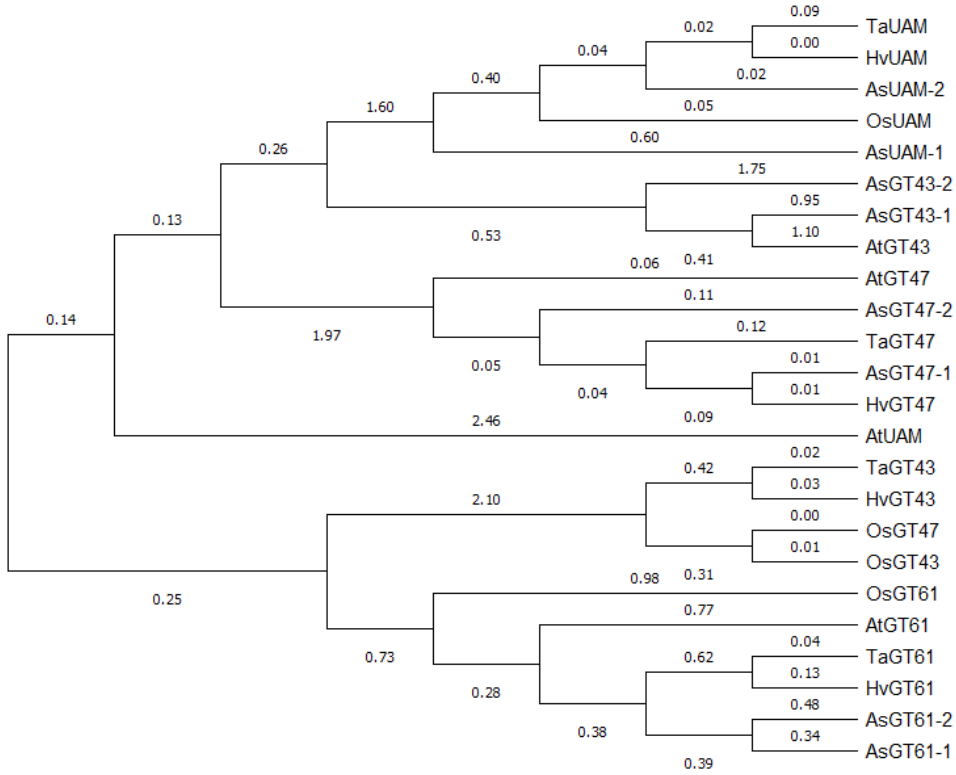


Figure 11. Evolutionary analysis of GT43, GT47, GT61, UAM1, and UAM2 in Arabidopsis thaliana (At), Avena sativa (As), Oryza sativa (Os), Triticum aestivum (Ta), and Hordeum vulgare (Hv) by the maximum likelihood method calculated by the software MEGA11.

The phylogenetic tree in figure 11 shows homology to many of the genes across different species. The protein sequences of *GT43*, *GT47*, *GT61*, *UAM1*, and *UAM2* related to the biosynthetic pathway of AX were identified in rice (*Oryza sativa*), barley (*Hordeum vulgare*), and wheat (*Triticum aestivum*). Protein sequences of *GT43-1*, *GT43-2*, *GT47-1*, *GT47-2*, *GT61-1*, *GT61-2*, *UAM1*, and *UAM2* were identified in Oat (*Avena sativa*). The genes aligned of every species were *GT43*, *GT47*, *GT61*, *UAM1*, and *UAM2*.

Although the genes across species are often related, the *GT43* and *GT47* genes have a wider variation throughout the species. This variation could indicate that the function of the genes is less predictable and could play different roles in the biosynthesis of AX across species.

3.2.2 Evolutionary map of genes in oat

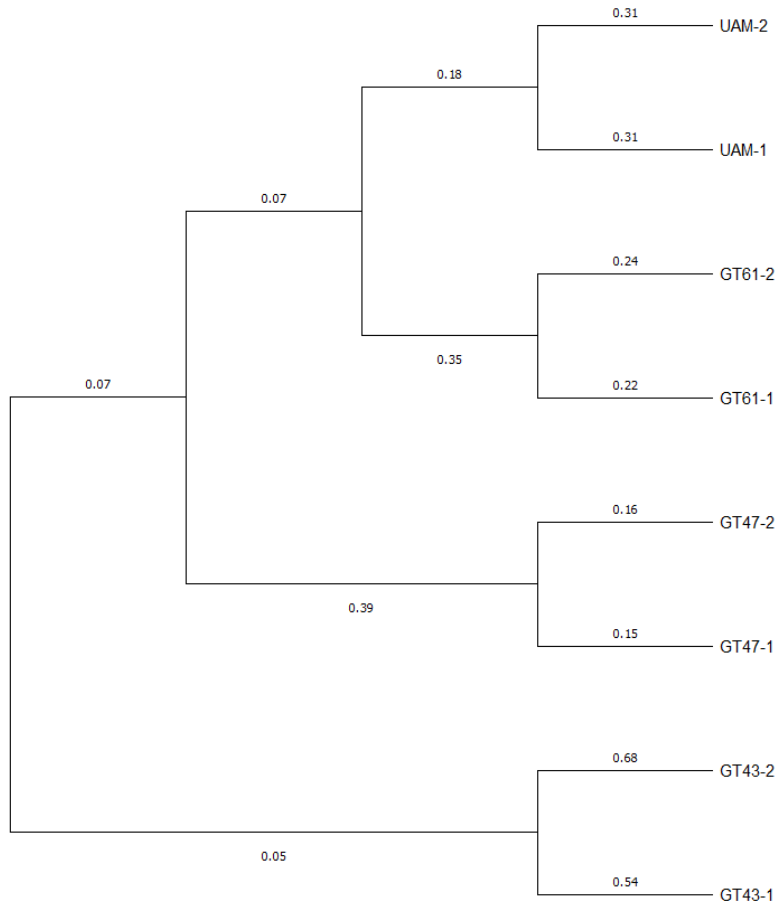


Figure 12. Phylogenetic tree of genes suspected to be related to the AX biosynthetic pathway in Hordeum vulgare (Hv). The tree was aligned and built using the maximum likelihood method in MEGA11.

The phylogenetic tree was built by aligning the eight genes in oat suspected to be involved in the biosynthetic pathway of AX: *GT43-1*, *GT43-2*, *GT47-1*, *GT47-2*, *GT61-1*, *GT61-2*, *UAM1*, and *UAM2*.
Looking at the phylogenetic tree in figure 12, we can see a close relation of each gene by the aliment made of all the genes selected in oat.

3.2.3 Quantification of Arabinoxylan Belinda

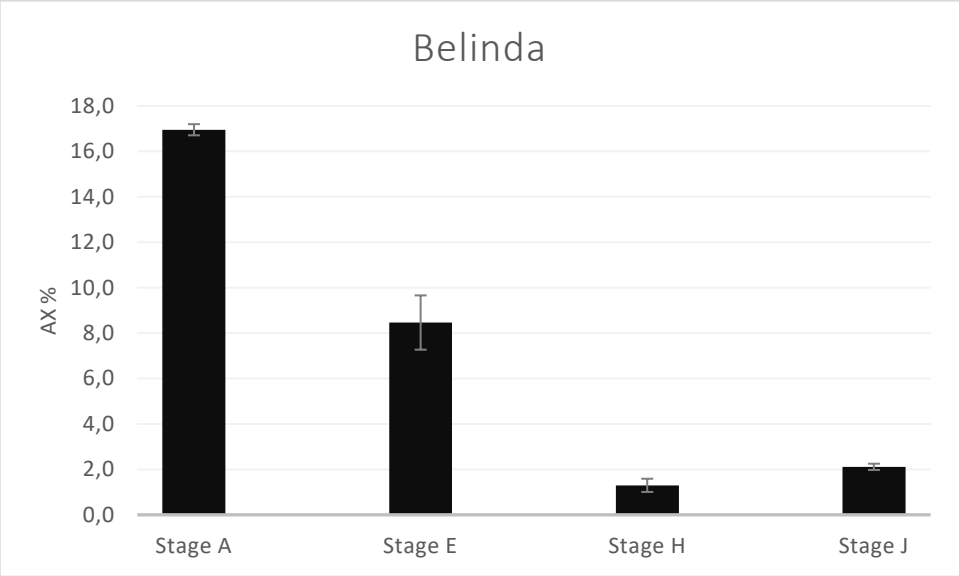


Figure 13. Percentage of AX per 25mg of milled oat kernel in four stages of maturity, going from A to J, with the highest percentage being in the early stage (A) with a steady decrease until the last stage (J). Error bars represent the standard deviation of two biological replicates.

The purpose of examining the distribution of AX in Belinda was to quantify AX in different developmental stages to understand how AX is distributed in different developmental stages and to see how this may relate to the gene expression.
Looking at the results of AX distribution in the reference Belinda throughout its development, it is possible to see a steep decline in AX through the first stage (A) with a slight increase in the last stage (J), see figure 13. These results suggest an apparent accumulation of AX in the early stage, consistent with the gene expression analysis, see Figures 7 and 8. This relation further strengthens the thesis that the genes examined are related to the biosynthetic pathway of AX in Oat. Further quantification of AX in different developmental stages needs to be done to explore the relationship between AX accumulation and gene expression. However, this was not possible during this project due to time constraints.

3.3 Conclusion

The gene expression analysis showed that *GT43, GT47, GT61, UAM1*, and *UAM2*, suspected to be involved in the AX biosynthetic pathway, were present and active in oat. The upregulation of the genes was highest in the early developmental stage, except for *GT61-2*, which had a higher upregulation in the middle and late stages. *GT61* has been shown to be a part of synthesizing substitutes for the xylan backbone in previous research, which is the last step of AX biosynthesis. The connection between the gene activation and the proposed function of *GT61* may indicate that the gene works as a feedback regulator to the other genes that had a high upregulation in the early stage.

Quantifying AX at different developmental stages showed a similar pattern to the gene activation and correlated with the gene expression analysis. This correlation may indicate that the AX is synthesized simultaneously as the examined genes are upregulated.

3.4 Further research

In order to further understand which genes are responsible for what in the AX biosynthetic pathway of oat, RNA sequencing could be a method to investigate the expression of more genes at different developmental stages.

References

Anders, N., Wilkinson, M.D., Lovegrove, A., Freeman, J., Tryfona, T., Pellny, T.K., Weimar, T., Mortimer, J.C., Stott, K., Baker, J.M., Defoin-Platel, M., Shewry, P.R., Dupree, P. & Mitchell, R.A.C. (2012). Glycosyl transferases in family 61 mediate arabinofuranosyl transfer onto xylan in grasses. *Proceedings of the National Academy of Sciences*, 109 (3), 989–993. https://doi.org/10.1073/pnas.1115858109

Brown, D.M., Zhang, Z., Stephens, E., Dupree, P. & Turner, S.R. (2009). Characterization of IRX10 and IRX10-like reveals an essential role in glucuronoxylan biosynthesis in Arabidopsis. *The Plant Journal: For Cell and Molecular Biology*, 57 (4), 732–746. https://doi.org/10.1111/j.1365-313X.2008.03729.x

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J. & Wittwer, C.T. (2009). The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clinical Chemistry*, 55 (4), 611–622. https://doi.org/10.1373/clinchem.2008.112797

Chawade, A., Sikora, P., Bräutigam, M., Larsson, M., Vivekanand, V., Nakash, M.A., Chen, T. & Olsson, O. (2010). Development and characterization of an oat TILLING-population and identification of mutations in lignin and β-glucan biosynthesis genes. *BMC Plant Biology*, 10 (1), 86. https://doi.org/10.1186/1471-2229-10-86

Chen, Z., Li, S., Fu, Y., Li, C., Chen, D. & Chen, H. (2019). Arabinoxylan structural characteristics, interaction with gut microbiota and potential health functions. *Journal of Functional Foods*, 54, 536–551. https://doi.org/10.1016/j.jff.2019.02.007

Correia, M.A.S., Mazumder, K., Brás, J.L.A., Firbank, S.J., Zhu, Y., Lewis, R.J., York, W.S., Fontes, C.M.G.A. & Gilbert, H.J. (2011). Structure and Function of an Arabinoxylan-specific Xylanase. *Journal of Biological Chemistry*, 286 (25), 22510–22520. https://doi.org/10.1074/jbc.M110.217315

Döring, C., Jekle, M. & Becker, T. (2016). Technological and Analytical Methods for Arabinoxylan Quantification from Cereals. *Critical Reviews in Food Science and Nutrition*, 56 (6), 999–1011. https://doi.org/10.1080/10408398.2012.749207

Gupta, M., Kaul, S. & Dhar, M.K. (2018). Identification and characterization of some putative genes involved in arabinoxylan biosynthesis in Plantago ovata. *3 Biotech*, 8 (6), 266. https://doi.org/10.1007/s13205-018-1289-9

Hughes, J. & Grafenauer, S. (2021). Oat and Barley in the Food Supply and Use of Beta Glucan Health Claims. *Nutrients*, 13 (8), 2556. https://doi.org/10.3390/nu13082556

Hundleby, P.A.C. & Harwood, W.A. (2019). Impacts of the EU GMO regulatory framework for plant genome editing. *Food and Energy Security*, 8 (2), e00161. https://doi.org/10.1002/fes3.161

Jensen, J.K., Johnson, N. & Wilkerson, C.G. (2013). Discovery of diversity in xylan biosynthetic genes by transcriptional profiling of a heteroxylan containing

mucilaginous tissue. *Frontiers in Plant Science*, 4. https://doi.org/10.3389/fpls.2013.00183

Kamal, N., Tsardakas Renhuldt, N., Bentzer, J., Gundlach, H., Haberer, G., Juhász, A., Lux, T., Bose, U., Tye-Din, J.A., Lang, D., van Gessel, N., Reski, R., Fu, Y.-B., Spégel, P., Ceplitis, A., Himmelbach, A., Waters, A.J., Bekele, W.A., Colgrave, M.L., Hansson, M., Stein, N., Mayer, K.F.X., Jellen, E.N., Maughan, P.J., Tinker, N.A., Mascher, M., Olsson, O., Spannagl, M. & Sirijovski, N. (2022). The mosaic oat genome gives insights into a uniquely healthy cereal crop. *Nature*, 606 (7912), 113–119. https://doi.org/10.1038/s41586-022-04732-y

Kole, C. (2006). *Cereals and millets*. Berlin: Springer. (Genome mapping and molecular breeding in plants; 1)

Konishi, T., Aohara, T., Igasaki, T., Hayashi, N., Miyazaki, Y., Takahashi, A., Hirochika, H., Iwai, H., Satoh, S. & Ishii, T. (2011). Down-regulation of UDP-arabinopyranose mutase reduces the proportion of arabinofuranose present in rice cell walls. *Phytochemistry*, 72 (16), 1962–1968. https://doi.org/10.1016/j.phytochem.2011.07.012

Lafiandra, D., Riccardi, G. & Shewry, P.R. (2014). Improving cereal grain carbohydrates for diet and health. *Journal of Cereal Science*, 59 (3), 312–326. https://doi.org/10.1016/j.jcs.2014.01.001

Lu, Z.X., Walker, K.Z., Muir, J.G. & O'Dea, K. (2004). Arabinoxylan fibre improves metabolic control in people with Type II diabetes. *European Journal of Clinical Nutrition*, 58 (4), 621–628. https://doi.org/10.1038/sj.ejcn.1601857

Mendis, M. & Simsek, S. (2014). Arabinoxylans and human health. *Food Hydrocolloids*, 42, 239–243. https://doi.org/10.1016/j.foodhyd.2013.07.022

Mushtaq, A., Gul-Zaffar, ., Z., A.D. & Mehfuza, H. (2014). A review on Oat (Avena sativa L.) as a dual-purpose crop. *Scientific Research and Essays*, 9 (4), 52–59. https://doi.org/10.5897/SRE2014.5820

Tian, L., Gruppen, H. & Schols, H.A. (2015). Characterization of (Glucurono)arabinoxylans from Oats Using Enzymatic Fingerprinting. *Journal of Agricultural and Food Chemistry*, 63 (50), 10822–10830. https://doi.org/10.1021/acs.jafc.5b04419

Wu, A.-M., Hörnblad, E., Voxeur, A., Gerber, L., Rihouey, C., Lerouge, P. & Marchant, A. (2010). Analysis of the Arabidopsis *IRX9* / *IRX9* - *L* and *IRX14* / *IRX14* - *L* Pairs of Glycosyltransferase Genes Reveals Critical Contributions to Biosynthesis of the Hemicellulose Glucuronoxylan. *Plant Physiology*, 153 (2), 542–554. https://doi.org/10.1104/pp.110.154971

Wu, A.-M., Rihouey, C., Seveno, M., Hörnblad, E., Singh, S.K., Matsunaga, T., Ishii, T., Lerouge, P. & Marchant, A. (2009). The Arabidopsis IRX10 and IRX10-LIKE glycosyltransferases are critical for glucuronoxylan biosynthesis during secondary cell wall formation. *The Plant Journal*, 57 (4), 718–731. https://doi.org/10.1111/j.1365-313X.2008.03724.x

Zhang, B., Liu, W., Chang, S.X. & Anyia, A.O. (2010). Water-deficit and high temperature affected water use efficiency and arabinoxylan concentration in spring wheat. *Journal of Cereal Science*, 52 (2), 263–269. https://doi.org/10.1016/j.jcs.2010.05.014

Popular science summary

Oat is one of the world's most cultivated cereal crops and is resilient enough to grow in the Scandinavian climate. Oat has been increasing in popularity because of its nutritional benefits and rich dietary fiber content. One of these fibers is arabinoxylan (AX). AX has been shown to have positive effects, like lowering the glycemic index of diabetes patients and having immunoregulatory effects as a pre-biotic. AX is hemicellulose and is found mainly in the primary and secondary cell walls.

The biosynthetic pathway of AX is not yet known in oat and has not been fully described in any other species. Although, ortholog genes coding for key enzymes have been found in other species like wheat, barley, rice, and Arabidopsis. These genes have been identified in oat but have not yet been analyzed.

CropTailor (CT) is a Swedish biotech company that has developed a large population of oat with random mutations in its DNA. Mutants with a low and high AX line compared to the wild-type Belinda have been identified within these populations.

In this thesis, eight ortholog genes suspected to play a role in AX synthesis have been identified in oat and were analyzed in two high and one low mutant line at three different developmental stages by quantitative reverse transcription-polymerase chain reaction (RT-qPCR). RT-qPCR is used to measure how much a gene is upregulated during different developmental stages.

AX content at four different developmental stages was also analyzed using High-Performance Anion-Exchange Chromatography (HPAEC).

This analysis showed a higher upregulation of genes in the early stages of development. The HPAEC also showed a higher AX content at the earlier stages.

This shows that the selected genes are also expressed in oat. The genes have a higher upregulation in the earlier stages which corresponds with the AX content found as a result of the HPAEC. An evolutionary analysis of the genes examined was also done in wheat, barley, oat, and *Arabidopsis*, which showed the evolutionary origin of all the genes in the different species.

Acknowledgements

Firstly, I would like to acknowledge and give a big thank you to my supervisor **Dr. Jose Alfredo Zambrano**, for guiding me in writing this thesis and being a mentor.

Secondly, I would like to thank my other supervisor **Erik Alexandersson** for helping me with the writing and making this thesis possible.

I would also like to thank **Karthik Vishaal** and **Aishwarya Thyagarajan** for teaching and helping me with the method of this thesis.

I would also like to give a big thank you to **CropTailor** for making this project possible.

Lastly, I want to thank **LTH**, **Lund University,** for providing the opportunity to write my thesis in collaboration with **SLU**.

Publishing and archiving

Approved students' theses at SLU are published electronically. As a student, you have the copyright to your own work and need to approve the electronic publishing. If you check the box for **YES**, the full text (pdf file) and metadata will be visible and searchable online. If you check the box for **NO**, only the metadata and the abstract will be visible and searchable online. Nevertheless, when the document is uploaded it will still be archived as a digital file. If you are more than one author, the checked box will be applied to all authors. Read about SLU's publishing agreement here:

- https://www.slu.se/en/subweb/library/publish-and-analyse/register-and-publish/agreement-for-publishing/.

☒ YES, I/we hereby give permission to publish the present thesis in accordance with the SLU agreement regarding the transfer of the right to publish a work.

☐ NO, I/we do not give permission to publish the present work. The work will still be archived and its metadata and abstract will be visible and searchable.