Analysis of chlorophyll synthesis mutants *xantha-m* and *xantha-j* of *Hordeum vulgare L.*

Fredrik Klinghammer

Supervisor: Mats Hansson

2020-06-02 VT

fr8560ni-s@student.lu.se

BIOK01

Abstract

The xantha phenotype of barley is observable as a yellow plant and caused by an inability to correctly synthesize chlorophyll. Chlorophyll is a tetrapyrrolic pigment containing magnesium, and is integral for photosynthesis in plants.

The phenotype can be caused by a number of mutations, and through methods like linkage mapping and sequencing it is possible to find the causative genotype. In this project xantha mutants *xan-m* and *xan-j* were investigated. The *xan-m* mutants were thought to possess mutations in the uncategorized gene HORVU3Hr1G050010 and *xan-j* a mutation in the chlorophyll synthase gene. The candidate gene of xan-m mutants is very large, and so only its transcribed mRNA was investigated through mRNA-selective RT-PCR and sequencing of the gene in eight fragments. A previously known mutation in the candidate gene of mutant *xan-m.53* was in this project confirmed through sequencing. However, no other sequenced *xan-m* mutants possessed mutations in the candidate gene.This suggests that the candidate gene mutation is not the one causing the xantha phenotype. *xan-j* investigation through western blotting was during the project unsuccessful, due to ordered antibodies not being of sufficient concentration.

Introduction

Barley (*Hordeum vulgare L.*), is an annual member of the Poaceae grass family. The crop grows 60 to 120 cm tall and features few leaves along the erect stem. Barley comes in two variants, plants with two rows of seeds opposite each other, and plants with six rows. Two-row barley is known to have a lower protein content along with a higher sugar content than six-row barley. Thus they are suited for different purposes, where the high protein six-row barley is preferred as animal feed and two-row preferred for fermentation in beer production. (Harwood, 2019).

Barley is a cornerstone of modern agriculture being the fourth most produced grain globally (Ullrich, 2011), and grown in high quantity during spring in Sweden (Jordbruksverket, 2019). It is also a useful model organism, as it is both diploid and self-pollinating. Today, barley is primarily used as animal feed and in the brewing industry. However, historically and in some parts of the world it is consumed in bread and porridges. It is noted to be a good source of beta-glucan. Barley is thought to have been domesticated around 10.000 years ago from a wild relative, *Hordeum spontaneum* (Harwood, 2019). Its genome is large, at 5.1 gigabases with around 36400 high confidence genes having assigned proteins (Liu et al. 2019). The genome is divided into 7 chromosome pairs, as is typical for the genus *Hordeum* (Stein & Muehlbauer).

Chlorophyll is a pigment and a tetrapyrrole molecule integral for all green plant life. It is vital for the process of photosynthesis, allowing plants to make use of solar energy. The chlorophyll molecules are situated in the thylakoid membranes of the chloroplasts, where they are part of photosystems I and II. These systems use the energy absorbed to create high-energy electrons that are pushed along a so-called electron transport chain. These electrons are then used by the plant to produce biomass (Brooker et al, 2014, pp158). Two very similar types of chlorophyll are present in plants, the lighter green type a and darker green type b. An important part of the chlorophyll molecule is the central magnesium ion to which four nitrogen atoms are coupled. Another integral part is the 20-carbon alcohol phytol that is coupled to a carboxylic acid side chain of the molecule (Berg J. M, Gatto G. J Jr., Stryer L., Tymoczko J. L., 2015).

Chlorophyll is synthesized in multiple steps. The first is synthesis of 5-aminolevulinic acid (ALA) from glutamyl-tRNA. This is a regulatory step down-regulated by both the presence of heme (another tetrapyrrolic molecule with biological significance) and by the regulatory protein FLU in the dark. The ALA molecule is then condensed with another ALA, forming a pyrrole. Four pyrroles are then polymerized into a linear pyrrole which is cyclidized into a precursor molecule of chlorophyll, protoporphyrin IX. Protoporphyrin IX is the substrate for the enzyme Mg-chelatase which will insert Mg^{2+} into the centre of the molecule. It is also substrate for Ferrochelatase, which instead inserts Fe^{2+} into the molecule centre. The steps including and following insertion of Mg^{2+} are called the Mg-branch. The protoporphyrin IX will then go through a number of oxidations and reductions through the help of multiple enzymes, as well as branch into the formation pathways of chlorophyll types a and b. At the final step of synthesis, chlorophyll synthase will couple the longer phytol chain to the tetrapyrrole via an ester bond. (Matsuda, 2008)

In this study I investigated chlorophyll synthesis mutants of Hordeum vulgare *L., xantha-j* (*xan-j*) and *xantha-m* (*xan-m*). These mutants are unable to synthetize chlorophyll correctly, leading to yellow plants.

xan-j mutants *xan-j.19*, *xan-j.59* and *xan-j.64* were investigated regarding the amount of expression of chlorophyll synthase in three xan-j mutants. This was done through western blotting using rabbit antibodies against chlorophyll synthase ordered from Agrisera. It was predicted that mutant *xan-j59*, as its allele contains an early stop codon, should have no chlorophyll synthase present.

xan-m mutants *xan-m.3*, *xan-m.48*, *xan-m.53*, *xan-m.72* and *xan-m.73* were investigated to see if they possess mutations in the candidate gene HORVU3Hr1G050010 (Appendix). This gene is uncharacterized but is of high confidence and is assigned a protein which may be related to cell division and stomata formation. They were investigated through growing of mutants and harvesting and extracting RNA. The RNA was converted into cDNA by the viral enzyme reverse transcriptase. The conversion selected for mRNA, ignoring other types of RNA by the usage of $Oligo(dT)_{18}$ primers conjugating to poly-A tails only present on mRNA. The cDNA is then amplified using PCR with primer pairs already designed by David Stuart and then sent in for sequencing.

As the HORVU3Hr1G050010 candidate gene is very large (30 kb) (IPK Leibniz Institute, n/d) we only examined exons transcribed into mRNA.

The gene was chosen as candidate because a whole genome sequencing experiment identified it as the only gene with a mutation on chromosome 3H in the region where the *Xantha-m* gene should be. The sequence library was created by TruSeq DNA PCR-free kit, fragmenting the whole genome into 350bp long pieces. The pieces were then sequenced by 150bp paired-end Illumina NovaSeq covering the whole genome 25-30x. This process sequences 150bp long strings and matches them up to form a complete genome sequence.

It was found during previous whole genome sequencing that *xan-m* mutant *xan-m.53* contains a base substitution mutation in the candidate gene, removing a splice site. The normal removal of introns through RNA splicing primarily occurs at certain splice junctions, situated on the border between exon and intron. These junctions are characterized by the bases GT at the 5' junction and AG at the 3' junction (Burset et al, 2001). The base substitution is G8728A, removing a 5' splice site.

In *Arabidopsis thaliana* the HORVU3Hr1G050010 homolog SCD1 has been studied. SCD1 loss of function mutants were found to be unable to correctly form stomata, resulting in binucleate guard cells without the complete ventral cell wall needed for cytokinesis. *xan-m* mutants were therefore investigated by microscope, observing stomata structure. SCD1 was also indicated in affecting seedling growth, root elongation and flower morphogenesis. Mutants of SCD1-deficient mutants grown on high-nutrient media were noted to be a darker green (Falbel T.G et al.).

Methods and Materials

Planting of seeds:

Seeds were planted in vermiculite into 0.1 l planting pots. Water was added until sufficiently moist. Plants were regularly watered until harvest.

RNA extraction from Xantha-m mutants and WT cultivars:

Xantha-m mutants *Xan-m.3*, *Xan-m.48*, *Xan-m.53*, *Xan-m.72*, *Xan-m.73* along with wild-type (WT) cultivars *Bonus* and *Gull* were harvested for 0.1 g of leaf material. At the point of harvest the plants were around 2 weeks old, standing at a height of 10-20 cm. Leaf material was collected into screw cap tubes containing two 4mm glass beads and were initially stored at -80 °C for extraction at later date. RNA extraction was done using the Nucleospin RNA Plant Kit (Macherey-Nagel)

RT-PCR:

2,5 µg mRNA extracted from xantha-m mutants and WT cultivars was converted into cDNA through the use of the viral enzyme reverse transcriptase (Thermo Fisher #EP0441). The Oligo$(dT)_{18}$ primer were chosen as to only bind mRNA expressed and exclude rRNA. Oligo$(dT)_{18}$ primer binds 3' poly-A tails only present on mRNA. One PCR tube with a total volume of 20 µl contained 10 µl REDExtract-N-Amp PCR Readymix (Sigma #R4775), along with 2 µl cDNA, 4 µl 1:1 mix of extraction and dilution buffers, 2 µl dH_2O aswell as 1 µl of both forward and reverse primers. 8 segments were amplified using primer pairs 1 through 8. Samples were after PCR analysed on a 2% agarose 50 ml gel electrophoresis to ensure amplified sequences were of correct length. GeneRuler 1 kb Plus DNA ladder (Thermo Fisher) was here used for size comparison.
cDNA of *xan-m* mutans were amplified by following PCR program:
94°C for 3 minutes followed by 35 cycles of 94 °C for 30 seconds - 60 °C for 30 seconds - 72 °C for 1 minute and then ended with 72 °C for 5 minutes. The PCR machine was set to cool the samples to 4 °C after completing the program.

Sample preparation for sequencing:

After observing correct sequence sizes in agarose gel 6µl of the PCR product was added to PCR tubes along with 2.4 µl ExoProStar 1 Step (GE Healthcare). PCR tubes were then

incubated in the PCR machine at 37 °C for 15 minutes, followed by 80 °C for another 15 minutes. After incubation, 8 µl of the contents were added to sequencing tubes together with 2 µl of corresponding forward primer and 7 µl of dH_2O for a total volume of 17 µl. Samples were then sent in for sequencing at eurofins genomics.

Protein extraction from *xantha-j* mutants and WT cultivars:

xantha-j mutants *xan-j.9*, *xan-j.59*, *xan-j.64* along with WT cultivars *Bonus* and *Foma* were harvested for 0.2 - 0.4 g of leaf material. At the point of harvest the plants were around 2 weeks old, standing at a height of 10-20 cm. Leaf material was collected into screw cap tubes containing two 4 mm glass beads and 10 µl 2.5% PMSF in isopropanol and were initially stored at -80°C for extraction at later date. The leaf material was grinded while frozen in a FastPrep 24 benchtop homogenizer. After homogenization, 1 ml of prepared protein extraction buffer (Appendix) was added to the tubes and they were let thaw to 60°C in water bath. Samples were then grinded two times while kept in water bath between grinds, after which they were incubated in heat block at 99 °C whilst shaking (14000 rpm). They were then centrifuged at 16000 g for 15 minutes and supernatant was transferred to new 1.5 ml eppendorf tubes and centrifuged again for 30 minutes. ~800 µl of the second supernatant was transferred to new 1.5 ml eppendorf tubes containing one part 5x sample buffer (Appendix) to four parts supernatant. The supernatant mix was boiled at 99 °C for 10 minutes. Solution not to be immediately used was aliquoted into parts of 100 µl and stored at -20 °C.

Western Blotting:

Precast gels (Mini-PROTEAN TGX Precast Gels) were loaded with 5 µl of Pageruler Prestained Protein Ladder (Thermo Scientific #26616) and 15 µl of sample and let run for 30 minutes at 200 V. Resulting gel was then blotted over to membrane by the use of Turboblot (Transblot Turbo Transfer Pack). Unspecific protein binding sites on the membrane were blocked by incubation for 1 hour in 5% milk in phosphate-buffered saline (PBS). Membranes were then washed for 5 minutes 3 times with PBS + 0.1% Tween. Membranes were then incubated for 1 hour in 20 ml 3% milk in PBS along with 10ul of primary antibody against chlorophyll synthase, bought from Agrisera. They were then washed for 5 minutes 3 times with PBS + 0.1% Tween. After the second wash they were let incubate in 20ml 3% milk in PBS containing the secondary antibody (Goat Anti-Rabbit IgG (H+L)-HRP Conjugate). After

incubation they were again washed 3 times with PBS + 0.1% Tween. Membranes were developed with (reagent) and photographed in (maskin namn).

SDS-PAGE:

A vertical gel (märke) were loaded with 5 µl ladder and 15 µl of sample and let run for 30 minutes at 200 V. After running, gel was stained with (product) according to instructions. Gels were photographed in (maskin namn). An addition gel was ran in this manner as a loading control for the western blotting.

Observation of stomata by microscope:

Samples were prepared through carefully scraping off cell layers ofthe upper epidermis of the leaf by scalpel. Remaining leaf was then turned around and fixed on slide and coverslip by water. Prepared slide was observed and photographed in microscope at 100x magnification. (10x ocular, 10x objective).

Results

xan-m mutant identification:

xan-m53 segments were aligned into a complete sequence and a base substitution of G to A was found in *xan-m53* HORVU3Hr1G050010 at point 1818 of the transcribed candidate gene (8728 of gDNA, Appendix) (Fig. 1.) . The mutation causes the removal of a splice site. As the number of added intron bases are a multiple of 3 the insertion will cause no shift in reading frame. The insertion is translated into 7 additional amino acids (Fig. 2.), possibly affecting the function of the translated protein.

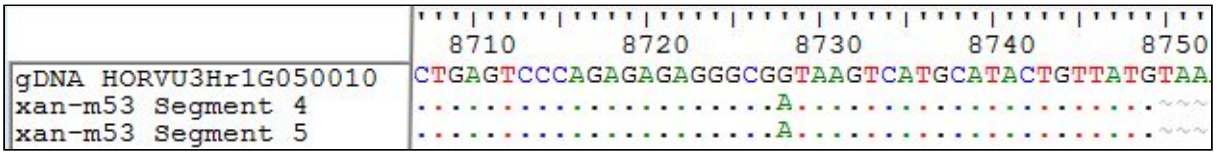

Fig. 1. *Alignment showing xan-m53 base substitution G8728A removing a 5' GT splice site, leading to 21 intron bases upstream from 8728 being transcribed.*

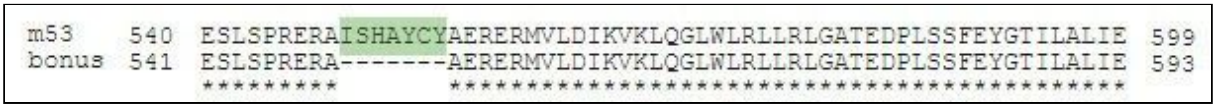

Fig. 2. *Protein sequence alignment of mutant xan-m53 versus WT bonus showing a 7 amino acid insertion at positions 550 to 557.*

Fig. 3. *HORVU3Hr1G050010 heterozygous and homozygous wild-type xan-m53 mutants (both green), along with homozygous mutant yellow phenotype xan-m53 mutants. Xantha phenotype plants exhibit severely stunted growth due to their inability to use photosynthesis for energy generation. They die before they reach a mature state.*

The *xan-m53* mutation matches the one previously found during whole-genome sequencing, thus confirming its location. The *xan-m* mutants *xan-m48* and *xan-m72* were successfully sequenced but no mutations were found. Mutants *xan-m3* and *xan-m73* were mostly

sequenced, but one of eight fragments was unable to be sequenced for each mutant. No mutations were found in the sequenced fragments of *xan-m3* and *xan-m73*. This observation suggests that the candidate gene mutation is not the mutation causing the phenotype (Fig. 3.).

Wild types cultivars *Bonus* and *Gull* were also successfully sequenced, as expected showing no mutations in the candidate gene outside those already known as differences from the Morex reference genome.

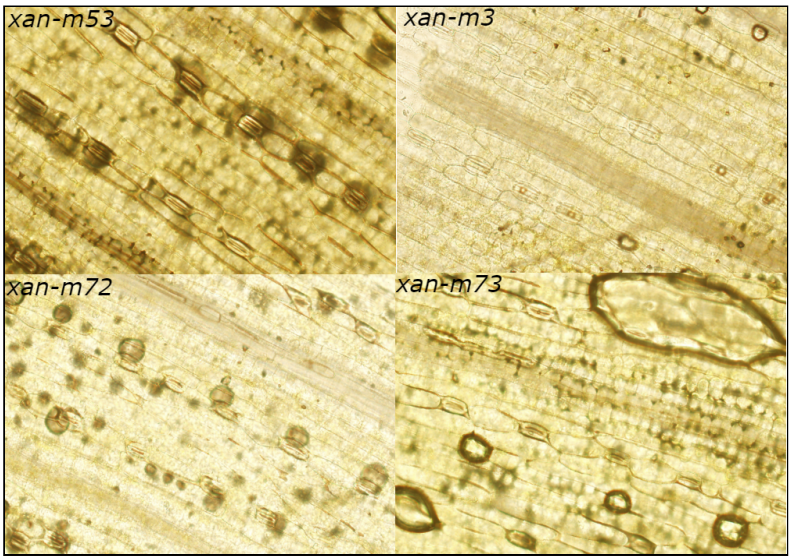

Fig. 4. *100x microscopic view of leaf epidermis with stomata. No difference in stomata structure is visible between mutants. xan-m.53 mutation in possibly stomata-related gene seems to have no impact on stomata integrity. Note that water in microscope-slide seems to gather above stomata, interfering with observation.*

Investigating stomata structure of xan-m mutants by microscope (Fig. 4.) shows no difference between mutants. Stomata of *xan-m.53* with confirmed mutation are compared with sequenced mutant *xan-m.72* - confirmed to contain no mutation in the candidate gene. The *xan-m.53* mutation seems to have no impact on stomata phenotype.

***xan-j* mutant chlorophyll synthase content:**

Due to a low primary antibody concentration, no results could be seen on the western blot. The blot was repeated with 5x and 10x the original antibody amount added but still no binding was visible (Fig. 5.).

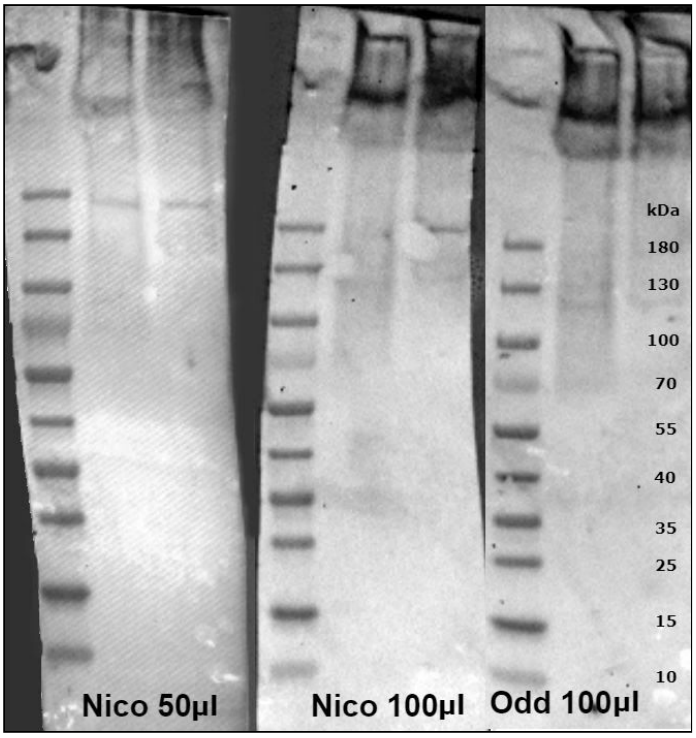

Fig. 5. *Western blot showing mutant xan-j.59 and WT bonus using 5x and 10x concentrations (50µl and 100µl) of the two antibody serums from rabbits Odd and Nico. No bands indicating chlorophyll synthase (42kDa) are visible. Wells are in the order of ladder - mutant - WT.*

Discussion

The previously found *xan-m.53* mutation was successfully sequenced, further confirming its location. The mutation is a base substitution, G1818A, causing an in-frame insertion of 21 base pairs through the removal of a 5' splicing site. The insertion is translated into 7 additional amino acids. A short insertion like this can result in a number of effects on the protein, like loss of function, a decrease in stability or very miniscule effects, having no impact on fitness (Lin et al. 2017).

Observation by microscope showed no difference between *xan-m.53* and *xan-m.3* stomata. This is notable as HORVU3Hr1G050010 homolog SCD1 loss of function mutants were inhibited in their ability to form intact stomata. Note however that the homolgy is only 70%, and that the protein could have different effects in different organisms. Along with the sequences of mutatants *xan-m.3* and *xan-m48* showing no mutations in the candidate gene, this strongly implies that the candidate gene mutation is unrelated to the phenotype.

The additional intron amino acids may also have a negligible effect on stomata structure while still causing the yellow phenotype.

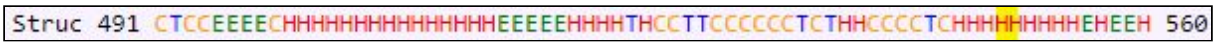

Fig. 6. *CFSSP secondary structure prediction of Bonus WT protein. The 7 amino acid insertion of ISHAYCY between residues A550 and E551 (yellow highlight) is likely to be situated in a helix region. H: Helix. E: Sheet. T: Turn. C: Coil.*

Short amino acid insertions into helices are known to sometimes be accommodated by the protein without major loss of function through two different mechanisms. The **register shift mechanism** preserves the geometry of the structure by rotating all residues on one side of the insertion by 100°. The other mechanism is called the **bulge mechanism** and distorts the local geometry but does not affect other residues to the same extent. Both mechanisms result in a possibly negligible effect on protein function, as long as the insert region is not one of high importance, for example an enzyme active site. However, they may both result in a substantial reduction in protein stability, with the register shift mechanisms resulting in a generally less severe destabilization than the bulge mechanism (Vetter et al, 1996) (Kavanaugh et al., 1993). Protein destabilization in turn leads to a larger fraction of incorrectly folded proteins unable to perform their function as well as being more likely to aggregate or be degraded by the cell (Redler et al. 2015).

There may be a number of reasons why a mutation corresponding to the phenotype was not found during the whole genome sequencing. The corresponding mutation may have been situated in an area with a high frequency of repeats, this leads to very similar short reads and unreliable sequences. Another issue is the GC content of the fragment sequenced. A too high or too low segment GC content causes disturbances during sequencing, resulting in unreliable sequences (Benjamini & Speed, 2012). The *xan-m.53* mutants was generated by x-rays, thus also leading to plenty of mutations in locations other than the one causing the phenotype (Breimer, 1988). Additionally, the analysis pipeline used previously for the whole genome sequencing data would not have been able to detect larger genomic rearrangements such as deletions, which can also be induced by x-ray mutagenesis.

Finding the mutation responsible for the xantha phenotype will require further research. A new mapping population of mutant *xan-m.48* is to be created. This mutant has the xantha phenotype but was generated not by x-ray but spontaneously, leading it to possess fewer mutations unrelated to the phenotype of interest. Finally, HORVU3Hr1G050010 may not be the gene of interest, but through further analysis it will be found. Perhaps the mutation is situated in an area unable to be correctly sequenced or is the cause of a deletion large enough to be overlooked. It is important to remember that negative results are still results and in this case they are a stepping stone towards valuable answers about the synthesis of chlorophyll in barley.

Chlorophyll is one of the most important compounds on earth, being integral to plant function. The importance of barley for the global market is also not to be understated. Discovering the connections between genotype and phenotype in a study like this is a step towards understanding the synthesis of the molecule. A greater understanding may then elucidate strategies of crop improvement and of finding beneficial mutations.

References

Benjamini Y., Speed T. P. (2012). Summarizing and correcting the GC content bias in high-throughput sequencing. *Nucleic Acids Research*, 40(10), 72. https://doi.org/10.1093/nar/gks001

Berg J. M, Gatto G. J Jr., Stryer L., Tymoczko J. L. (2015). *Biochemistry* (8th ed.) Kate Ahr Parker.

Breimer L. H. (1988). Ionizing radiation-induced mutagenesis. *Br J Cancer*. 57(1), 6-18. https://doi.org/10.1038/bjc.1988.2

Brooker R. J, Graham L. E, Stiling P. D, Widmaier E. P. (2014). *Biology* (4th ed.). McGraw-Hill Education.

Falbel T. G., Koch L. M., Nadeau J. A., Segui-Simarro J. M., Sack F. D., Bednarek S. Y. (2003) SCD1 is required for cytokinesis and polarized cell expansion in Arabidopsis thaliana. *Developement*, 130, 4011-4024. https://doi.org/10.1242/dev.00619

Harwood W. A.. (2019). An Introduction to Barley: The Crop and the Model. *Methods in Molecular Biology, vol 1900.* https://doi-org.ludwig.lub.lu.se/10.1007/978-1-4939-8944-7_1

IPK Leibniz Institute. (n/d). HORVU3Hr1G050010. Barlex: The Barley Genome Explorer. https://apex.ipk-gatersleben.de/apex/f?p=284:45:::NO::P45_GENE_NAME:HORVU3Hr1G050010.11

Jordbruksverket. (2019). *Skördestatistik spannmål*. Jordbruksverket. http://www.jordbruksverket.se/amnesomraden/odling/jordbruksgrodor/vete/skordestatistik.4.32b12c7f12940112a7c800020268.html?fbclid=IwAR0SO6fMMTSfUFxcEwGlEARq88wBEMiI9D-vZUQFIeDO-jUsYmnsc7boAt4

Kavanaugh J. S., Moo-Penn W. F., Arnone A. (1993). Accomodation of Insertions in Helices: The Mutation in Hemoglobin Catonsville (Pro 37α-Glu-Thr 38α) Generates a $3_{10} \rightarrow \alpha$ Bulge. *Biochemistry*, 32, 2509-2513.

Liu, M., Li, Y., Ma, Y., Zhao, Q., Stiller, J., Feng, Q., Tian, Q., Liu, D., Han, B., & Liu, C. (2020). The draft genome of a wild barley genotype reveals its enrichment in genes related to biotic and abiotic stresses compared to cultivated barley. *Plant biotechnology journal*, 18(2), 443–456. https://doi.org/10.1111/pbi.13210

Macherey-Nagel. (2020.) *NucleoSpin RNA Plant, Mini kit for RNA from plant.* Macherey-Nagel. https://www.mn-net.com/nucleospin-rna-plant-mini-kit-for-rna-from-plant-740949.50

Matsuda T. (2008). Recent overview of the Mg branch of the tetrapyrrole biosynthesis leading to chlorophylls. Springer Link. https://doi.org/10.1007/s11120-008-9291-4
Burset M., Seledtsov I. A., Solovyev V. V. (2001). SpliceDB: database of canonical and non-canonical mammalian splice sites. *Nucleic Acids Research*, 1, 255-259. https://doi.org/10.1093/nar/29.1.255

Redler R. L., Das J., Diaz J. R., Dokholyan N. V. (2016). Protein Destabilization as a Common Factor in Diverse Inherited Disorders. *J Mol Evol*, 82(1), 11-16. https://doi.org/10.1007/s00239-015-9717-5

Stein, N., Muehlbauer, G. J. (2018). *Compendium of Plant Genomes: The Barley Genome*. Springer International Publishing AG. https://doi.org/10.1007/978-3-319-92528-8

Ullrich S. E. (2010) *Barley: Production, Improvement, and Uses.* Wiley-Blackwell.

Vetter I. R., Baase W. A., Heinz D. W., Xiong J., Snow S., Matthews B. W. (1996). Protein structural plasticity exemplified by insertion and deletion mutants in T4 lysozyme. *Protein Science*. 5. 2399-2415.

Appendix

Sample Buffer

Compound	**Conc. in 5x**
Tris pH 6.8	150mM
Glycerol	50%
SDS	2%
2-mercaptoethanol	5%
Bromophenol blue (BFB)	0.1%

Protein Extraction Buffer

Compound	**Conc. in 1x**
Urea	12M
SDS	2%
DTT	100mM
Tris-HCl pH 8.0	50mM
EDTA pH 8.0	10mM

HORVU3Hr1G050010.11 notation, IPK Leibniz Institute, *Barlex: The Barley Genome Explorer*. https://apex.ipk-gatersleben.de/apex/f?p=284:45:::NO::P45_GENE_NAME:HORVU3Hr1G050010.11