Project acronym: PROTEIN2FOOD

Project No.: 635727

H2020-SFS-2014-2015/H2020-SFS-2014-2

Start date of project: March 2015

Duration: 5 years

Deliverable reference number and title

D1.4 Protein quality and quantity transcriptomes available for target crops for further use in developing SNPs.

January 2018

Organisation name of lead for this deliverable:

Swedish University of Agricultural Sciences (SLU)

Project co-funded by the European Commission within the Horizon 2020 Programme		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 635727.

0. Summary

Quinoa (*Chenopodium quinoa*) and faba bean (*Vicia faba*) were chosen as target protein crops for transcriptome analysis to be used for the development of molecular markers for higher protein content. Several genotypes of each species were grown under equal environment in controlled growth chambers. Seeds were sampled during development and used for characterization of deposition of major storage compounds, seed characteristics and for sequencing of their respective transcriptomes. Transcriptome data (from RNA-sequencing) is now available for three different quinoa genotypes that show interesting seed characteristics and will now be bio-informatically processed to identify genetic markers associated with protein content.

1. Introduction and Objectives

Among different crops interesting for use as plant-based protein sources, faba bean seeds are already relatively high in protein content (around 30%) while quinoa seeds are lower (around 15-20%) even though the latter has a very well balanced protein quality (amino acid composition) for human nutrition. From an economical perspective, from both the farmer/grower and the food industry, breeding to develop cultivars with as high protein content as possible should be of high importance to make protein crops as competitive as possible. One aim of WP1 is to achieve a better understanding of the genetic mechanisms driving the protein synthesis and accumulation in seeds from target protein crops. This knowledge can help us to identify mechanisms and genes responsible for partitioning a higher share of seed storage compounds into protein synthesis. By bioinformatic analysis of seed RNA-sequence data (transcriptomes) we will understand which genes are expressed (i.e. 'used) in the seeds. If there are interesting differences in gene expression patterns that can be correlated to protein content in different genotypes of a target crop, this information can be used to develop genetic markers for protein content for further use in plant breeding.

2. Activities for solving the task(s)

Total protein content was first screened in different genotypes of the target crops quinoa and faba bean to identify genotypes with contrasting protein content. The most interesting among these (three genotypes of quinoa and six of faba bean) were further grown under controlled and equal growth conditions in climate chambers. Faba bean plants were grown under a light: dark cycle of 16 h:8 h at 21°C/ 18°C, and quinoa plants at 12 h: 12 h at 20°C/ 17°C, both at 70% humidity. Plants were tagged at anthesis and seeds were sampled during development and up to maturity into liquid nitrogen to keep RNA intact. Seed samples were ground in liquid nitrogen in steel containers into a powder saved in freezer. Protein content in sampled seed material was estimated by total Nx6.25, where totalN was determined using elementar anlayzer (THERMO Fisher Scientific) to characterise genotype differences in protein storage accumulation. Starch was also analysed in the sampled seed material using enzymatic methods for total starch determination (Megazyme, Wicklow, Ireland). Embryo ratios of quinoa seeds were determined using X-ray MicroCT imaging (see separate deliverable for details). Altogether, these characterisations were evaluated to identify which genotypes and developmental stages to extract RNA from, to be sent for sequencing. Total RNA was extracted from samples seed material using PureLink Plant RNA Reagent (Ambion, USA). Total RNA was DNase treated to get rid of genomic DNA (TurboDNase, Ambion, USA).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 635727.

Integrity and concentration of DNased RNA was measured using Experion RNA StdSens analysis kit (BioRad, Hercules, USA) and sent for poly-A selection based cDNA library preparation and further paired-end Illumina sequencing to a depth of minimum 30 million reads per sample (GATC-Biotech, Germany).

3. Results

Two of the three selected quinoa genotypes showed similar developmental rate (Figure 1) after anthesis (*i.e.* Titicaca and Regalona). However, one genotype (Pasankalla rosada) is of altiplano ecotype and had a much slower rate of seed development under the day/night regime used (12h light/12hdark, selected to make both day length sensitive and insensitive genotypes develop flowers in the same chamber). For those three different genotypes to be comparable, developmental stages of Pasankalla rosada seeds could not be based on days post anthesis but instead on visual inspection of seeds. Due to that dry matter content of seeds from all three quinoa genotypes turned out to be roughly similar at each developmental stage (data not shown), it was assumed that sampled stages are in fact comparable between genotypes.

Storage compound analysis of quinoa seeds identified a significant genotype difference in both starch and protein content at later developmental stages (Figure 2). Interestingly, the genotype Pasankalla rosada had highest protein content in mature seeds but the lowest content of starch, indicating a potential difference in carbon allocation between different seed storage products. However, oil content will also be analysed to get a more complete picture of carbon and storage balances of seeds. There were no significant genotype differences in protein quality (Figure 3) which means that this trait cannot be taken into account when later on evaluating the transcriptome data for potential markers.

As part of the characterisation of quinoa seeds that will be used during interpretation of transcriptome data, embryo ratios of quinoa seeds were determined from seed density analysis using a medium-high resolution of X-ray scanning (an example picture from X-ray scanning 3D movies can be seen in Figure 4, together with results for different genotypes, done by Giacomo Mele at ISAFoM-CNR). From this analysis it was determined that two genotypes (Pasankalla rosada and Altiplano) had significantly higher embryo ratios as compared to the two other genotypes analysed (Titicaca and Regalona). At the same time, seed weights (mg/seed) did not differ. A high embryo ratio is an interesting seed phenotype since the majority of proteins in quinoa seeds are deposited in the embryo, and not in the perisperm which instead is starch rich. Therefore, the information of a higher embryo ratio of the high seed protein genotype Pasankalla rosada will potentially be interesting to take into account when evaluating genotype differences in transcriptome data.

RNA was extracted from three developmental stages and quality was analysed by capillary gel electrophoresis to confirm intact and high quality index of RNA (Figure 5). RNA was further DNased treated to get rid of genomic DNA contamination that can disturb sequencing results. RNA prepared from two developmental stages were selected to be sent to RNA-sequencing facility (Illumina HiSeq, yielding approximately 30 million reads per sample). Transcriptome data is now available for three different genotypes of quinoa at two developmental stages (Table 1, including three biological replicates to ensure reliability of data).

Faba been genotypes analysed did not show any significant differences in seed protein content that was more or less stable around 30% at maturity (Figure 6). This suggests that the previously observed genotype differences in seed protein content from the screening of available gene bank seed material were most probably caused by environmental factors. Thus,

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 635727.

it is not possible to use the available faba bean material for the identification of genetic markers associated to protein content through comparative transcriptome analyses.

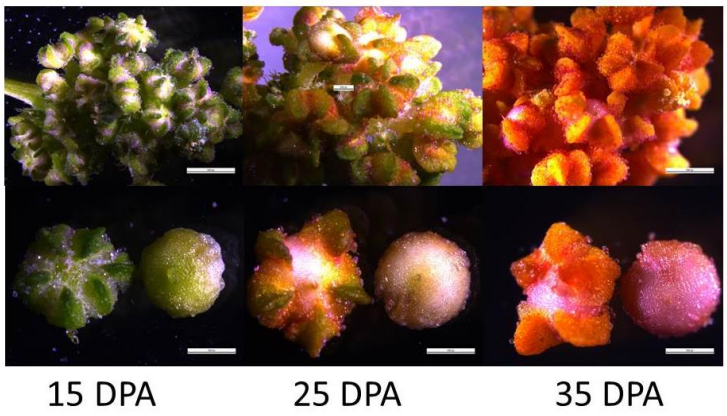

Figure 1. *Quinoa seeds cv. Titicaca during different developmental stages. DPA; days post anthesis. (Microscope photo: Å. Grimberg, SLU).*

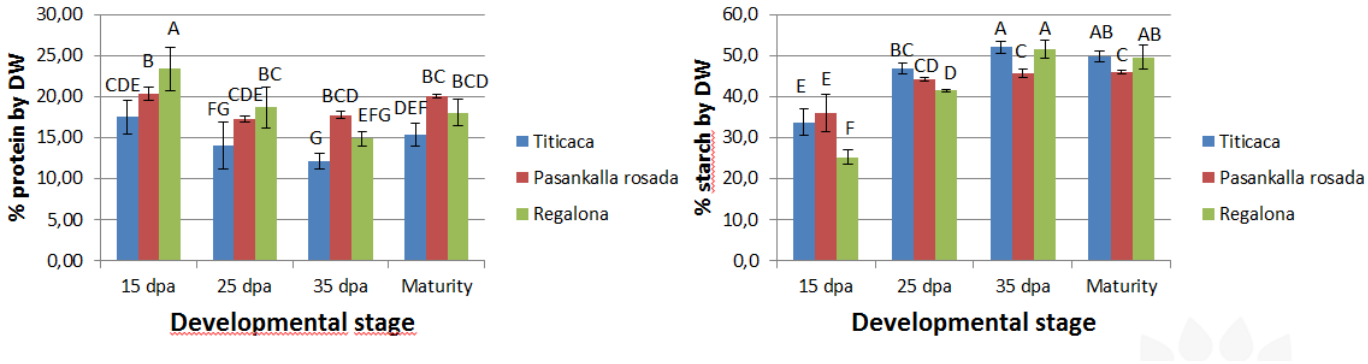

Figure 2. *Protein (left) and starch (right) content in developing seeds from three different quinoa genotypes grown under equal conditions in controlled growth chambers. Protein content was estimated by total Nx6.25, where totalN was determined using elementar analyzer. Starch content was determined by enzymatic analysis. Results are the mean±stdev from three biological replicates. DW; dry weight, dpa; days post anthesis. Bars that do not share letters are significantly different according to Fisher's test at significance level P≤0.05.*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 635727.

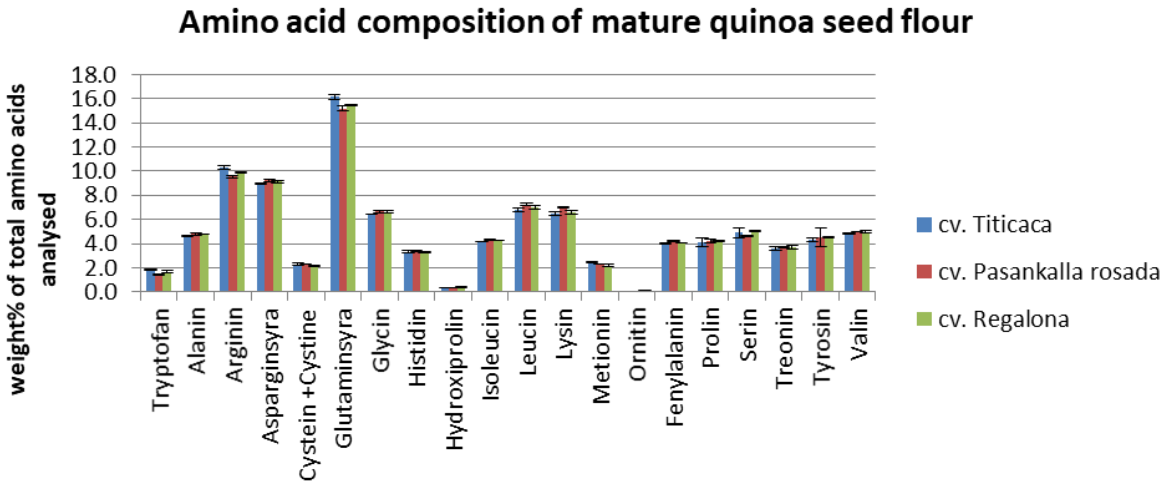

Figure 3. *Amino acid quality (weight% of total amino acids) in flour of mature seeds of three different genotypes of quinoa. Results are mean±stdev from two biological replicates.*

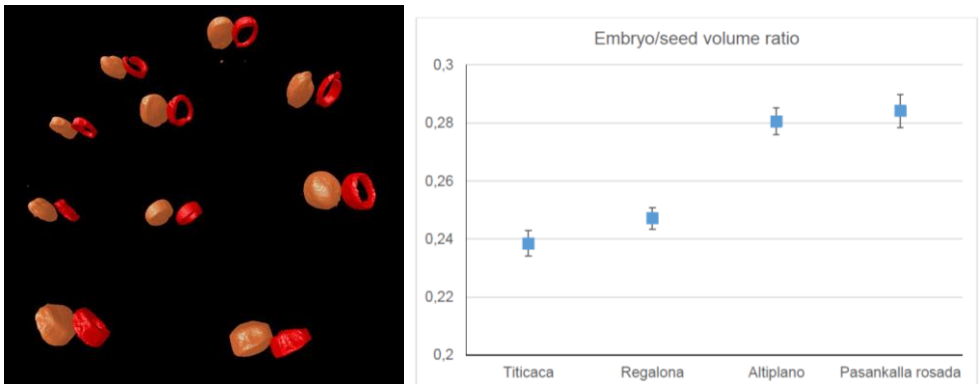

Figure 4. *Embryo/seed ratios of mature quinoa seeds as determined by X-ray MicroCT imaging. Left; picture visualising virtually separated embryos (red) and perisperm (brown) of a ten seed sample of quinoa. Right; embryo/seed ratios in four different genotypes of quinoa (results are mean±stdev from three biological replicate á 10 seeds each).*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 635727.

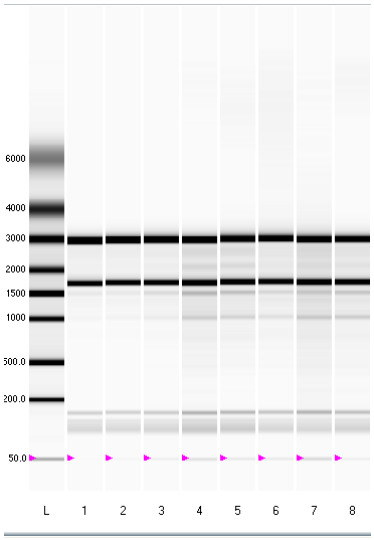

Figure 5. *Capillary gel electrophoresis visualising examples of size separation and confirmation of intact extracted total RNA from developing quinoa seeds.*

Table 1. *Transcriptome data from RNA-sequencing (Illumina HiSeq) of developing seeds of three different genotypes of quinoa, now available for bioinformatics analyses and interpretation and further for marker development.*

Sample	***Million reads (300 bp)***
Titicaca 15 dpa-1	*55*
Titicaca 15 dpa-2	*40*
Titicaca 15 dpa-3	*65*
Titicaca 25 dpa-1	*32*
Titicaca 25 dpa-2	*34*
Titicaca 25 dpa-3	*40*
Regalona 15 dpa-1	*47*
Regalona 15 dpa-2	*50*
Regalona 15 dpa-3	*32*
Regalona 25 dpa-1	*55*
Regalona 25 dpa-2	*23*
Regalona 25 dpa-3	*41*
Pasankalla rosada Early-1	*59*
Pasankalla rosada Early-2	*58*
Pasankalla rosada Early-3	*60*
Pasankalla rosada Mid-1	*71*
Pasankalla rosada Mid-2	*70*
Pasankalla rosada Mid-3	*67*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 635727.

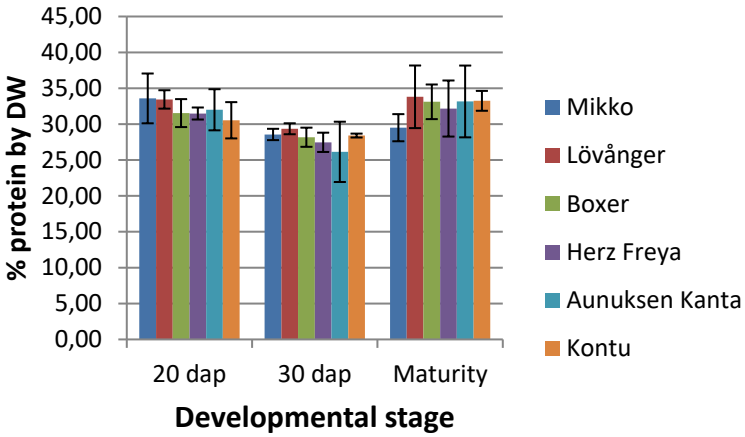

Figure 6. Protein content in developing seeds from six different faba bean genotypes grown under equal conditions in controlled growth chambers. Results are the mean±stdev from three biological replicates. DW; dry weight, dap; days after pollination.

4. Conclusion and next steps

Transcriptomes from seeds of three different genotypes of quinoa that have been carefully characterised during development is now available for bioinformatics analysis. Next step is to do comparative transcriptome analysis between genotypes to identify genes that are differentially expressed, as well as identify gene variants of those or other genes that can be determinants of protein content in quinoa.

5. Delays and difficulties

Analysis of the available faba bean material grown under equal climate conditions did not show any significant genotype differences in protein content. Therefore transcriptomes will not be produced in purpose to target protein content from this faba bean material. However, other target traits that could be interesting for agronomic performance of faba bean such as earliness are currently discussed to be included in activities to achieve D1.5.

6. Impact and dissemination activities

Making seed transcriptomes of quinoa genotypes with contrasting protein content available for further bioinformatics analysis will i) significantly increase the current sequence information available for this species which will be useful for the plant scientific community ii) increase the knowledge about genetic factors that determines protein accumulation and content in quinoa, and iii) potentially lead to development of genetic markers for protein content for use in future breeding of new quinoa cultivars with increased protein content. The results will be summarized and disseminated through open-access scientific publication and stakeholder interaction.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 635727.