Linköping University | Department of Physics, Chemistry and Biology
Bachelor thesis, 16 hp | Biology programme: Physics, Chemistry and Biology
Spring term 2018 | LITH-IFM-G-EX--18/3515--SE

Is there genetic variation in *VicJ*, which can be associated with protein content variation in pea (*Pisum sativum* L.)?

Erika Andersson

Examinator, Dominic Wright, IFM Biologi, Linköpings universitet
Tutor, Jenny Hagenblad, IFM Biologi, Linköpings universitet

Department of Physics, Chemistry and Biology
Linköping University

Datum
Date
180527

Språk
Language

☐ Svenska/Swedish
☒ Engelska/English

☐ _____

Rapporttyp
Report category

☐ Licentiatavhandling
☒ Examensarbete
☐ C-uppsats
☐ D-uppsats
☐ Övrig rapport

☐ _____

ISBN

ISRN: LITH-IFM-G-EX--18/3515--SE

Serietitel och serienummer **ISSN**
Title of series, numbering _____

URL för elektronisk version

Title
Is there genetic variation in VicJ, which can be associated with protein content variation in pea (*Pisum sativum* L.)?

Author
Erika Andersson

Sammanfattning
Abstract

Today, the livestock sector accounts for 18 % of greenhouse gas emissions. To prevent negative environmental effects, dietary changes are required. Locally cultivated legumes with high protein content can be used in order to produce plant-based protein, which can replace animal-based protein. In Sweden, pea (*Pisum sativum* L.) has been cultivated for centuries and been a valuable protein source for both human consumption and animal feed. *VicJ*, a gene in pea, has previously been associated with variation in protein content. In the present study, a primarily Swedish material of 31 accessions from different improvement stages were analysed for differences in protein content. It was also tested if genetic variation of *VicJ* was associated with variation in protein content. The result showed no differences in protein content between various improvement stages, which indicated that selection on the trait has not occurred. No genetic variation associated with variation in protein content in *VicJ* was detected either. However a stop codon in *VicJ*, known to be associated with reduced protein content was missing in the material, suggesting that the accessions studied may be suitable for breeding to increase protein content in pea.

Keywords
genetic variation, Pea, *Pisum sativum* L, protein content, VicJ

Table of content

1 Abstract

Today, the livestock sector accounts for 18 % of greenhouse gas emissions. To prevent negative environmental effects, dietary changes are required. Locally cultivated legumes with high protein content can be used in order to produce plant-based protein, which can replace animal-based protein. In Sweden, pea (*Pisum sativum* L.) has been cultivated for centuries and been a valuable protein source for both human consumption and animal feed. *VicJ*, a gene in pea, has previously been associated with variation in protein content. In the present study, the Swedish material of 31 accessions from different improvement stages were analysed for differences in protein content. It was also tested if genetic variation of *VicJ* was associated with variation in protein content. The result showed no differences in protein content between various improvement stages, which indicated that selection on the trait has not occurred. No genetic variation associated with variation in protein content in *VicJ* was detected either. However a stop codon in *VicJ*, known to be associated with reduced protein content was missing in the material, suggesting that the accessions studied may be suitable for breeding to increase protein content in pea.

2 Introduction

Higher income and urbanization have resulted in a global increase of meat consumption (Tilman and Clark 2014). This has a large environmental cost. Today, the livestock sector cause 18 % of greenhouse gas emissions (Food and Agricultural Organisation 2006). If this trend continues is has been calculated that by 2050 meat, milk and egg production will be responsible for 39 % of greenhouse gas emissions (Pelletier and Tyedmers 2010).

To prevent the negative environmental effects of livestock production dietary changes are required (Pelletier and Tyedmers 2010). Plant-based protein cause less emission of greenhouse gases than animal-based protein (Tilman and Clark 2014). Legumes are suitable as substitute to meat since they contain high amounts of protein. In addition they provide energy, dietary fibre, minerals and vitamins (Boye, Zare and Pletch 2010). Besides being good nutritional sources legumes are nitrogen-fixing crops. Cultivation could therefore reduce the usage of fertilizer, which in turn decreases emission of nitric oxide (Peoples Herridge and Ladha 1995).

Soybean is the dominating crop on an industrial scale and it is used for both human consumption and animal feed (Barack et al 2015). The

protein concentration varies between 330 to 490 g per kilo (Hymowitz et al 1971). Although soybean is appropriate as food and feed protein, its usage contributes to global warming. Since cultivation of soybean is not suitable in temperate climate countries like Sweden, these countries are dependent on import, which results in emission of greenhouse gases (Fearnside 2001, Sly 2017). In 2011 the European Union imported 70 % of its total use of protein crops, which accounted for 14 % of the soybean production in the world. To limit the reliance of imported soybean the European Union exhorts the usage of locally produced crops with high protein content (European Parliament 2013).

Pea (*Pisum sativum* L.) is a protein legume widely cultivated in Europe (Olle et al 2015). In Sweden, pea has been cultivated for centuries and is a valuable protein source for both food and feed. However, the cultivation in Sweden has declined. Today pea constitutes 1 % of Sweden's arable land while in the beginning of the 1800-century it was 3% (Leino and Nygårds 2008). The average protein content in pea is 230 g per kilo dry weight but it can vary depending on genetic factors and cultivation conditions. Selection of genetic variants with elevated protein concentration has been of interest in breeding programmes but it has proven to be difficult (Olle et al 2015). Identification of genes, which are associated with variation in protein content, would therefore be useful to increase the amount of protein in pea.

VicJ is a gene that has been associated with variation of protein content in pea. Chinoy et al (2011) found an allele of *VicJ* with a premature stop codon. The allele, mapped to the vc-2 locus, was present in some but not all accessions tested. The allele showed to be non-functional and encoded a truncated protein. Accessions with the non-functional *VicJ* allele showed lower nitrogen/protein concentration (Chinoy et al 2011). Apart from the study by Chinoy et al (2011) genetic variation regarding protein content in *VicJ* seems not to have been studied.

The present study investigated if various improvement stages differ in protein content in Swedish pea material. It was also investigated if there is genetic variation in *VicJ*, which can explain variation in protein content between different pea cultivars.

3 Material and methods

3.1 Material

A total of 31 accessions from different improvement stages were analysed (table 1). Seeds were provided by Nordic Genetic Resource

Center (NordGen) and John Innes Centre. Data of protein content in different improvement stages of pea was available by measurements performed by Agrilab. The measurements were conducted on protein content, seed weight, and ripening. In addition sequences for 25 accessions were already available at Linköping University. For the 6 remaining accessions without complete sequences, DNA was also available at Linköping University.

All accessions could be considered to be pure lines due to propagation by single seed decent. As a consequence all accessions were homozygotes and the offspring had the same genotype. This means that measurements of protein content on each accession were performed on the same genotype hence the average for protein content, seed weight and ripening could be used for statistical analyses.

Table 1. Data of analysed pea accessions containing their name, accession number, improvement stage and average seed weight, protein content and ripening with standard deviation in parenthesis. For the accessions marked with an asterisk sequencing was performed since no sequences were available.

Accession	Name (location)	Improvement stage	Protein content (g kg^{-1} dry weight)	Seed weight (g)	Ripening (Days)
JIC1031	Tyskland	European landrace	235 (55.29)	9.27 (3.4)	32.4 (1.14)
JIC1525*	Prasia	European landrace	295 (61.98)	6.86 (1.95)	33.4 (3.29)
JIC1778	Culdris'	European landrace	292 (16.02)	10.79 (4.68)	33.8 (2.17)
NGB102814	WBH2814	European landrace	308 (9.01)	9.71 (1.7)	24.8 (2.49)
NGB17871	Baltikum	European landrace	191 (47.25)	4.21 (2.68)	30.2 (1.10)
NGB17883*	Papardes	European landrace	246 (75.87)	9.55 (1.35)	31.4 (1.67)
NGB17884	Vidzermes Tirgus	European landrace	232 (38.06)	5.23 (0.84)	38.8 (1.30)
NGB20117	Lollandske Rosiner'	European landrace	269 (24.83)	5.08 (1.3)	35.6 (1.67)
NGB20123*	Errindler	European landrace	200 (52.16)	9.04 (2.68)	32.2 (1.48)
NGB101819	WBH1819 (Norra Rörum)	Swedish landrace	214 (54.80)	8.04 (1.57)	30.2 (1.64)
NGB103517*	Jämtländsk grå'	Swedish landrace	273 (66.0)	9.19 (2.03)	28.6 (1.82)
NGB103518	WBH3518 (Tollestorp)	Swedish landrace	224 (31.21)	8.72 (3.47)	30.8 (2.05)
NGB103590	WBH3590 (Skararp)	Swedish landrace	277 (14.25)	2.91 (2.82	33.6 (1.52)
NGB13469	Gråär from Laholm, 'Stäme'	Swedish landrace	260 (63.25)	8.16 (2.66)	38.2 (5.54)
NGB13487	Östgöta gulärt'	Swedish landrace	299 (36.40)	11.32 (4.6)	26.2 (6.30)
NGB14153	Solberga	Swedish landrace	277 (61.98)	9.87 (2.3)	33.2 (1.30)
NGB14154	Maglaby	Swedish landrace	301 (24.24)	10.63 (4.94)	27.4 (1.14)
NGB14155	Skånsk gråärt	Swedish landrace	298 (21.56)	11.77 (5.08)	34.6 (1.14)
NGB14639	Orust	Swedish landrace	267 (21.37)	10.84 (4.13)	27.8 (3.50)
NGB14642	Lit	Swedish landrace	268 (58.09)	8.14 (2.41)	27.8 (1.10)
NGB17868	Väse	Swedish landrace	253 (31.58)	6.71 (1.37)	31.2 (1.79)
NGB17873	Puggor from Glimåkra	Swedish landrace	272 (14.86)	12.62 (4.45)	32.2 (2.17)
NGB17881	Rättviks gråärt	Swedish landrace	298 (21.54)	12.4 (2.5)	33.8 (9.68)
NGB102027*	WBH2027	Wild	287 (12.49)	8.49 (6.52)	28.2 (1.64)
NGB102123	WBH2123	Wild	262 (7.77)	5.87 (0.92)	32.3 (1.53)
NGB103567*	WBH3567	Wild	338 (88.68)	4.6 (2.47)	32.0 (3.61)
NGB101997	Brioärt	Cultivar	254 (28.68)	9.92 (2.9)	33.6 (1.52)
NGB103071	Rosakrone	Cultivar	227 (24.88)	12.52 (1.29)	34.0 (0.82)
NGB10660	Capella	Cultivar	307 (25.89)	4.0 (1.7)	34.0 (1.00)
NGB13138	Odalett	Cultivar	212 (31.59)	6.87 (2.27)	32.6 (1.67)
NGB4018	Timo	Cultivar	248 (66.92)	10.36 (0.94)	32.4 (1.14)

3.2 DNA extraction and PCR

JIC1525 was grown and DNA was extracted in order to optimize PCR conditions. Five seeds were grown in plastic pots (12x12 cm) on standard cultivation soil. Two leaves were harvested (1x1 cm) and placed in eppendorf tubes with silica gel to dry. After JIC1525 had been harvested the dry leaves were homogenized with Tissuelyser II (Qiagen®) and DNA was extracted using Qiagens DNeasy Plant miniprep kit according to the manufacturer´s instructions. DNA concentration for JIC1525 was measured with NanoDrop™ instrument.

JIC1525 DNA was used to optimize PCR conditions for two different primer combinations (table 2). All amplifications of DNA were conducted using a S1000™ Thermal Cycler (Bio-Rad Laboratories). When PCR conditions had been optimized amplification was performed on the remaining accessions (table 1). The annealing temperatures were 57 and 58°C for PVicJ1F + PVicJ1R and PVicJ2F + PVicJ2R, respectively. The volume of each PCR reaction was 20 µl and contained 10X DreamTaq-buffert (Thermo Scientific), 1µM of each primer, 5 U μl^{-1} DreamTaq DNA polymerase (Thermo Scientific), 0.25 mM dNTP and ddH_2O. Reaction runs consisted of 94°C during 2.30 minutes, 34 cycles where each cycle consisted of 94°C during 15 seconds, 57/58°C during 40 seconds and 72°C during 40 seconds. The final step was set at 72°C during 10 minutes. Amplification products yielding a distinct band upon electrophoresis were used for sequencing.

3.3 ExoSap

The amplification products were cleaned from primers and nucleotides that were not incorporated using ExoSap. Each reaction had a total volume of 6 µl and contained 20 U μl^{-1} exonuclease (Fermentas), 1 U μl^{-1} FastAP (Thermo Scientific) and ddH_2O. The reactions ran for 30 minutes at 37°C followed by 5 minutes at 95°C.

3.4 Sequencing

After the samples had been treated with ExoSap 5 µl amplification product and 5 µl of relevant the primer (5µM) were sent to Macrogen for Sanger sequencing with fluorescently labelled termination nucleotides.

VicJ was amplified in three overlapping DNA products. DNA products for the last primer pair (PVicJ3F + PVicJ3R) were already available for all accessions. The primer sequences are shown in table 2.

Table 2. Sequences of all primers used in the experiment.

Primer	Sequence
VicJ1 Forward	ACAGTGCCTGTCTTTATGTTTCTCT
VicJ1 Reverse	TCGGCATCGGTGTATTGTGGAAGA
VicJ2 Forward	ACGGTCACATTCGTCTTCTCCA
VicJ2 Reverse	CTGTGTTGTGGCTCTTGTTCCTGTTG
VicJ3 Forward	TGTAAGCATGCAACACACACATCC
VicJ3 Reverse	ACACCTGGAGACAACTTAGCTCT

Existing sequences were edited upon visual inspection in Geneious 11.1.4. The same procedure was applied on sequences returning from Macrogen. Geneious was also used for identifying different haplotypes by visual inspection in accessions with good sequence quality.

3.5 Statistical analysis

IBM SPSS statistics version 24.0 was used for statistical analyses regarding differences in nutrient content between various improvement stages and differences in protein content between haplotypes. The software used for investigating selection on *VicJ* was DNAsp5 where Tajima´s D was calculated. The same software was also used to calculate nucleotide diversity (Pi-value) of sequences. The selected significance probability was 5 % for all tests.

The data of protein content, seed weight and ripening were tested for normal distribution by Shapiro-Wilk´s test. One-way ANOVA was conducted to test if various improvement stages differed in protein content. Pearson correlation test was performed to test correlation between seed weight and protein content and ripening and protein content. To test if the different haplotypes that were found could be associated with variation in average protein content one-way ANOVA or independent T-test was conducted. To detect selection on *VicJ* Tajima´s D was calculated on accessions for different improvement stages. The test could not be performed for wild accessions where only one sequence was available, as four or more sequences are required to calculate Tajima´s D. To establish the nucleotide diversity of sequences from different improvement stages the Pi-value was calculated. Nucleotide diversity for wild accessions could not be tested since only one accession was available.

4 Results

4.1 DNA extraction

In order to confirm that DNA extraction of JIC1525 was successful DNA concentration was measured. The concentration was 56.6 ng μl^{-1}, which was a sufficient amount of DNA for continuing with amplification. The 260/280 ratio was 2.10 which was deemed sufficiently close to the 1.8 expected for pure DNA.

4.2 Protein content for different improvement stages

Protein content, seed weight and ripening in different improvement stages of pea were normally distributed (Shapiro-Wilk´s test, N = 31, p > 0.05). The average protein content was 264 g kg^{-1} dry weight and protein content of all accessions varied between 191-338 g kg^{-1} dry weight (Table 1, figure 1).

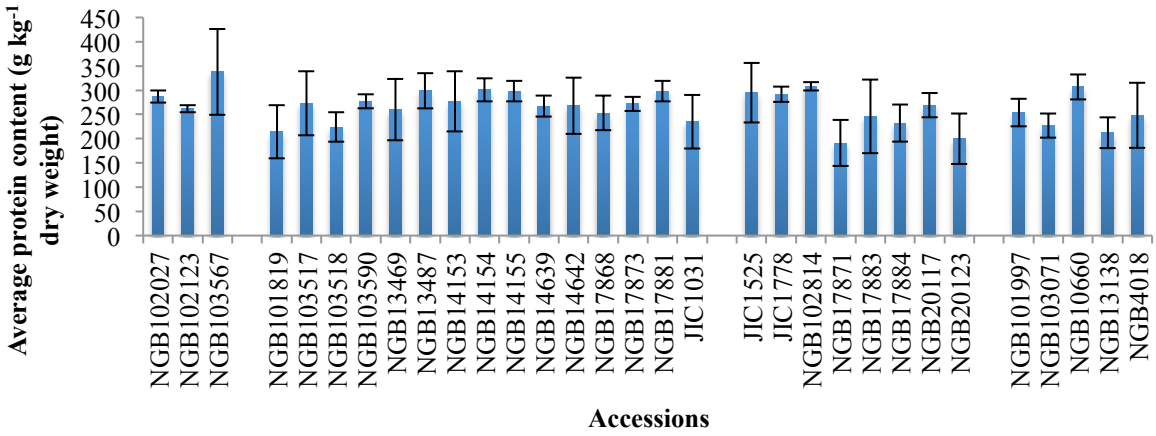

Figure 1. Each accession´s average protein content and standard deviation divided into different improvement stages. From left: wild, Swedish landrace, European landrace, and cultivar.

No significant differences in protein content were observed between the different improvements stages European landrace (N = 9), Swedish landrace, (N = 14), wild (N = 3) and cultivar (N = 5) (ANOVA, N = 31, F = 1.68, p = 0.20) (figure 2). When the outlier (NGB10660) was excluded the result still showed no significant differences in protein content between various improvement stages (ANOVA, N = 30, F =2.56, p = 0.08) (figure 2). Significant differences between Swedish landraces and cultivars was not detected (N = 18, F = 3.23 p = 0.09) (figure 2).

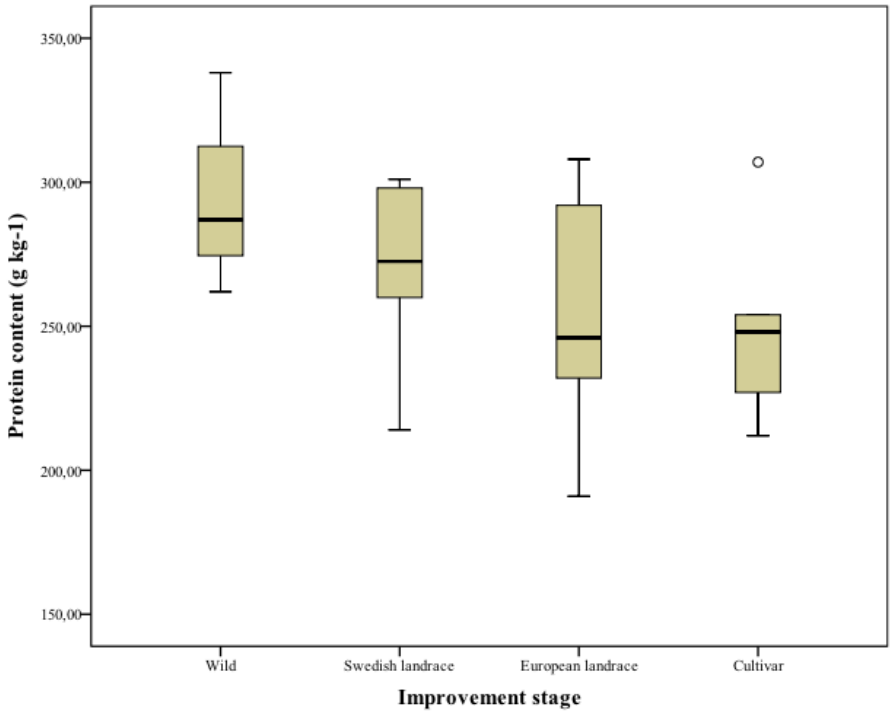

Figure 2. Box-plots of protein content of various improvement stages. The boxes report median, minimum, maximum and the first respectively third quartile. Mean values (g kg^{-1}): wild = 296, Swedish landrace =270, European landrace =252 and cultivar = 250.

No significant correlation between protein content and seed weight was detected (Pearson correlation test, N = 31, p = 0.63, r = 0.089) (figure 3A). Significant correlation between protein content and ripening could not be observed either (Pearson correlation test, N = 31, p = 0.40, r = -0.15) (figure 3B).

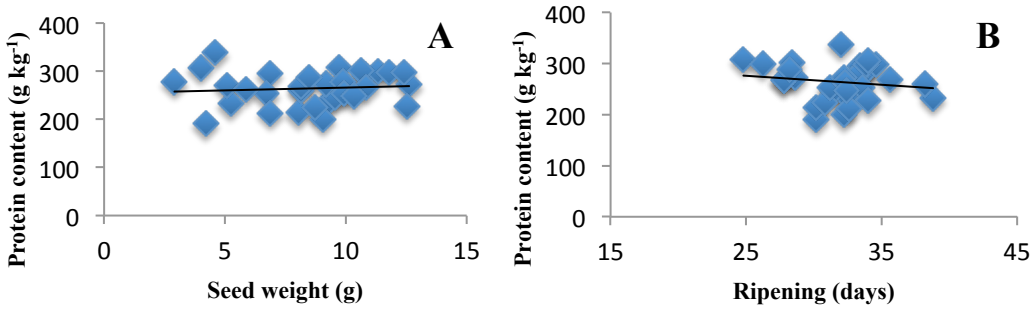

Figure 3. Correlation between protein content and seed weight (3A) and protein content and ripening (3B)

4.3 Sequence analysis

After visual inspection of all sequences JIC1031, JIC1525, NGB17883, NGB20123, NGB103517 NGB102027, NGB102123, and NGB4018

were excluded from the sequence analysis due to poor quality. 23 accessions had good sequence quality and were used for sequence analyses. After excluding singletons four different haplotypes were identified among 18 accessions (appendix table 1). There was no significant difference in protein content between haplotypes (ANOVA, N = 18, F = 1.10, p =0.38) (figure 4). Comparing only the two most common haplotypes no significant differences in protein could be determined either (Independent T-test, t = -0.17, p = 0.09). Neither did Swedish accessions and Swedish landraces differ in protein content between haplotypes (ANOVA, N = 13, F = 2.03, p = 0.18).

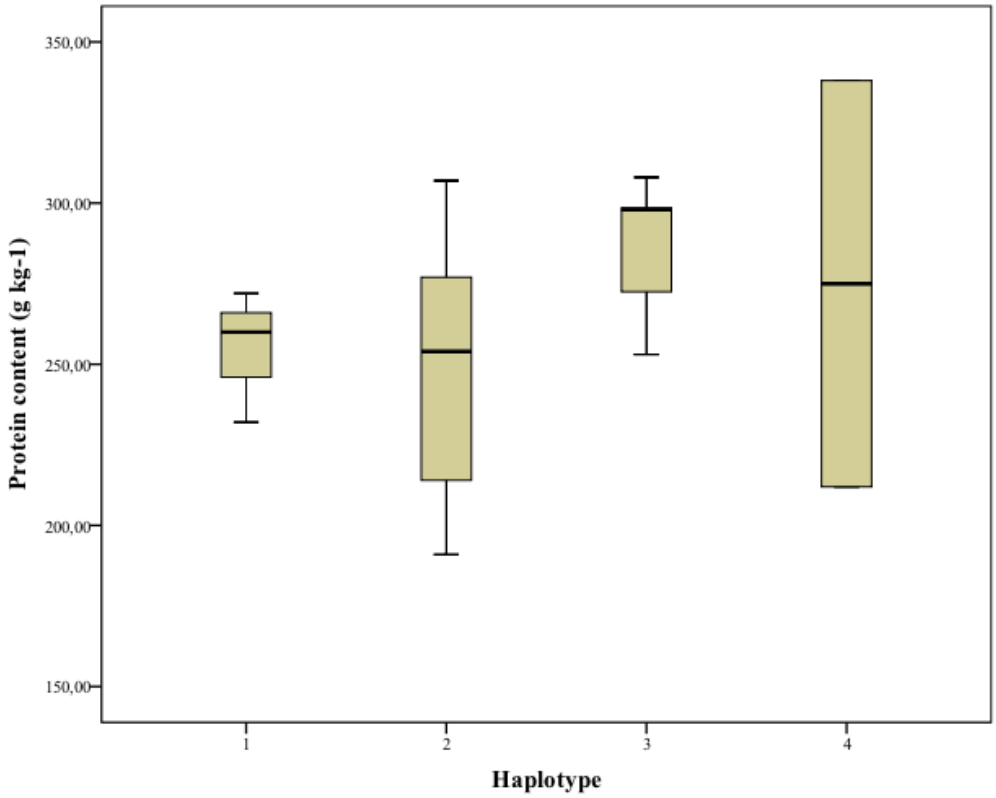

Figure 4. Box-plots of each haplotypes´ protein content. The boxes report median, minimum, maximum and the first respectively third quartile. Mean values (g kg^{-1}): 1 = 255, 2 =249, 3 =286 and 4 =273.

4.4 Selection of *VicJ*

To detect selection on *VicJ* Tajima´s D was calculated on accessions for different improvement stages but no statistical significance could be observed (table 3). The nucleotide diversity was lowest for cultivars (table 3).

Table 3. Tajima´s D, Pi-value, number of accessions tested, p-value and statistical significance for different improvement stages of pea.

Improvement stage	Tajima´s D	Pi-value	N	P-value	Statistical significance
All improvement stages	1.31	0.00497	23	p > 0.05	Not significant
European landrace	-1.07	0.0043	5	p > 0.05	Not significant
Swedish landrace	1.77	0.00545	13	p > 0.05	Not significant
Cultivar	-0.8	0.00116	4	p > 0.05	Not significant

5 Discussion

5.1 Protein content in various improvement stages

In Sweden, breeding on pea started in the beginning of the 19th-century. The starting material was Swedish landraces and breeding was mainly performed on yield, maturation and thousand grain weight (Sjödin 1997). When comparing protein content between different improvement stages no significant differences could be observed. This suggests that selection on protein content has not been targeted for improvement. Similar results were shown in a previous study where unimproved pea varieties and elite cultivars did not significantly differ in protein content (Santalla et al 2001). Tajima´s test confirmed that no selection on VicJ has occurred, which suggest that the gene has not contributed to pea fitness. The result from Tajima´s test agrees with the fact that there is no variation in protein content between different improvement stages. If differences in protein content had been detected it would be reasonable that selection for the trait had occurred. However, the nucleotide diversity of *VicJ* decreased from landraces to cultivar, possibly due to the occurrence of a bottleneck in conjunction with breeding.

A previous study on wheat showed negative correlation between protein content and yield in cereal grains (Simmonds 1995). The explanation for this was that large seeds, which are associated with high yield, cause a dilution effect on nutrient concentrations (Fan et al 2008, Singh et al 1990). However, no significant correlation between protein content and seed weight could be detected in this study. No correlation between ripening and protein content was detected, which is consistent with a previous study performed on soybean (Tanteeratarm and Steinberg 1989).

5.2 Genetic variation

The results showed no significant difference regarding protein content between haplotypes in the material studied. Genetic variation in *VicJ* was

found but it could not be linked to differences in protein content in pea. Four haplotypes of VicJ were identified as candidates for variation in protein content but no significant difference between the haplotypes could be observed. Furthermore, no significant differences could be determined between the two most frequent haplotypes. However, eight accessions were excluded from the sequence analysis due to poor sequences quality. The number of sequences may be to few to determine the statistical significance regarding differences in protein content between different haplotypes.

Studies regarding identification of genes associated with protein content in pea are few. However, a handful of quantitative trait loci (QTL) for protein content in pea have been found. All these QTLs have been located in linkage group V, which is the region to where the vc-2 locus and *VicJ* has been mapped (Izrykowska and Wolko 2004, Chinoy et al 2011, Krajewski et al 2012). The results in this study contrasts with earlier findings made by Chinoy et al (2011). They found a premature stop codon in *VicJ*, which was associated with reduced protein content in pea (Chinoy et al 2011). When the *VicJ* coding amino acid sequence from the Chinoy et al study was aligned with the coding amino acid sequences from present study no such stop codon could be detected, which could explain the differences between this study and that of Chinoy et al (2011). The average protein content of the material was 264 g kg^{-1} dry weight, which is higher than the general content of 230 g kg^{-1} dry weight (Olle et al 2015). The lack of premature stop codon in the accessions studied can be an explanation for this phenomenon. The fact that accessions in the present study had higher protein content than average makes them suitable for breeding to increase the amount of protein in pea.

Since the results of this study indicate that there is no genetic variation in *VicJ* that can be associated with protein content variations other aspects regarding protein content needs to be considered. Future prospects would be to study if accessions from Chinoy et al (2011) and accessions from this study differ in protein content when grown in a common garden and if the potential difference could be linked to genetic variation in the Swedish material. Genome-wide association study (GWAS) could be an alternative method to apply since it identifies single-nucleotide polymorphism linked to a specific trait. This has successfully been done in chickpea where seven alleles governing protein content was detected by using GWAS (Upadhyaya et al 2016). A previous study investigated environmental effects on protein concentration in pea cultivated in Spain. The results showed significant differences for the trait between different environments (Santalla et al 2001). The present study excluded

environmental effects on variation in protein content but for investigations of accessions from Chinoy et al (2011) and the present study the material should be cultured under same conditions to eliminate environmental effects.

5.3 Conclusion

No differences in protein content between various improvement stages were observed. It can therefore be assumed that protein content has not been a priority for improvement.

VicJ showed no genetic variation associated with variation in protein content in different improvement stages of pea. One explanation could be that *VicJ* is not involved in variation in protein content in the material studied. However a stop codon in *VicJ*, known to be associated with reduced protein content was missing in the material, which could be an explanation for why accessions in the present study showed higher average protein content than general. The results suggest that the accessions in this study could be suitable for breeding to produce pea with high protein content.

5.4 Societal and ethical aspects

The discovery of genes, which generates elevated protein content, enables production of legumes with high protein content for local usage. This could lead to better utilization of vegetable protein for food and feed. The consequence would be positive environmental effects since it would result in reduced emission of green house gases.

Knowledge of nutrient genes can lead to increased production of genetically modified organisms (GMO). GMO has been a controversial debate for many years and several studies have been performed to investigate positive and negative effects (Klumper and Qaim, 2014). One argument for GMO is that nutrient genes could prevent malnutrition since food with higher nutrient content could be produced. Arguments against GMO are e.g. that it can reduce biodiversity and those companies, which owns the patent of seeds controls the market and shift the balance of economic power.

6 Acknowledgement

I would like to thank my supervisor Jenny Hagenblad for her valuable knowledge and support. Also, I would like to thank Maria Lundström for the advices given during sequence analyses. Finally, I would like to thank

my co-worker Louise Carlsson, my opponents Liselotte Neumann and Anna Lotsander and my examiner Dominic Wright.

7 References

Barack M, Pesic M, Stanojevic S, Kostic A, Cabrilo S (2015) Techno-functional properties of pea (*Pisum sativum*) protein isolates – a review. APTEFF. 46, 1-269

Boye J, Zare F, Pletch A (2010) Pulse proteins: Processing, characterization, functional properties and application in food and feed. Food Research International. 43, 414-431

Chinoy C, Welham T, Turner L, Moreau C, Domoney C (2011) The genetic control of seed quality traits: effects of allelic variation at the *Tri* and Vc-2 loci in *Pisum sativum* L. Euphytica. 180, 107-122

European Parliament (2013) The Environmental Role of Protein Crops in the New Common Agricultural Policy. http://www.europarl.europa.eu/studies (accessed 10 April 2018)

Fan MS, Zhao FJ, Fairweather-Trait SJ, Poulton PR, Dunaham SJ, McGrath SP (2008) Evidence of decreasing mineral density in wheat grain over the last 160 years. Journal of Trace Elements in Medicine nd Biology. 22, 315-324

Fearnside P. (2001) Soybean cultivation as a threat to the environment in Brazil. Environmental Conservation. 28, 23-38

Food and Agriculture Organization of the United States. (2006) Livestock´s long shadow: environmental issues and options. FAO, Rome, Italy.

Hymowitz T, Collins FI, Panczner J, Walker WM (1971) Relationship Between the Content of Oil, Protein, and Sugar in Soybean Seed. Agronomy Journal. 64, 613 – 616

Irzykowska L, Wolko B (2004) Interval mapping of QTLs controlling yield-related traits and seed protein content in *Pisum sativum*. J. Appl. Genet. 45, 297-306

Klumper W, Qaim M (2014) A Meta-Analysis of the Impacts of Genetically Modified Crops. PLoS ONE. 9(11):e111629 https://doi.org/10.1371/journal.pone.0111629

Krajewski P, Bocianowski J, Gawlowska M, Kaczmarek Z, Pniewski T, Swiecicki, Wolko B (2012) QTL for yield components and protein content: a multienvironment study of two pea (*Pisum sativum* L.) populations. Euphytica. 183, 323-336

Leino MW, Nygårds L. (2008) Puggor och pelusker - svenska lokalsorter av ärt. Svensk botanisk tidskrift. 103, 3-4

Olle M, Narits L,Williams IH (2015). The influence of variety on the yeild and content of protein and nutrients of peas (*PISUM SATIVUM*). Research for rural development. 1, 45-50.

Pelletier N, Tyedmers P (2010) Forecasting potential global environmental costs of livestock production 2000-2050. PNAS. 107, 18371-18374

Peoples MB, Herridge DF, Ladha JK (1995) Biological nitrogen fixation: an efficient source of nitrogen for sustainable agricultural production? Plant and Soil. 174, 3-28

Santalla M, Amurrio JM, De Ron AM (2001) Food and feed potential breeding value of green, dry and vegetable pea germplasm. Can. J. Plant Sci. 81, 601–610

Simmonds NW (1995) The Relation Between Yield and Protein in Cereal Grain. J Sci Food Agic. 67, 309-315

Singh KB, Bejiga B, Malhotra RS (1990) Association of some characters with seed yield n chickpea collections. Euphytica. 49, 83-88

Sjödin J (1997) Trindsäd. Den svenska västförädlingens historia. Skogs- och Lantbrukshistoriska meddelanden. 20, 215-222

Sly M (2017) The Argentine portion of the soybean community chain. *Palgrave communications*. 3. https://doi.org.10.1057/palcomms.2017.95

Tanteeratarm K, Wei LS, Steinberg MP (1989) Effect of Soybean Maturity on Storage Stability and Process Quality. Journal of Food Science. https://doi.org/10.1111/j.1365-2621.1989.tb04660.x

Tilman D, Clark M (2014) Global diets link environmental sustainability and human health. Nature. 515, 518-522

Upadhyaya H, Bajaj D, Narnoliya L, Das S, Kumar V, Gowda CLL, Sharma S, Tyagi AK, Parida SK (2016) Genome-wide Scans or

Delinetion of Candidate Genes Regulating Seed-Protein Content in Chickpea. Front. Plant. Sci. https://doi.org/10.3389/fpls.2016.00302

8 Appendix

Table 1. The haplotypes found among 18 accessions.

Haplotype	Accession
Haplotype 1	NGB17884, NGB13469, NGB17873
Haplotype 2	NGB14153, NGB101819, NGB17871, NGB101997, NGB10660
Haplotype 3	NGB14155, NGB103590, NGB17868, NGB17881, NGB13487, NGB14642, NGB102814
Haplotype 4	NGB13138, NGB103567, NGB20117