

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Improving Seed Quality of Soybean Suitable for Growing in Europe

Aleksandra Sudarić, Maja Matoša Kočar,
Tomislav Duvnjak, Zvonimir Zdunić and
Antonela Markulj Kulundžić

Additional information is available at the end of the chapter

Abstract

The potential of soybean for food, feed, and pharmaceutical industry arises from the composition of its seed. Since European countries import 95% of the annual demand for soybean grains, meal, and oil, causing an enormous trade deficit, the governments in Europe had started to introduce additional incentives to stimulate soybean cropping. To rebalance the sources of soybean supply in the future, production must be followed by continuous research to create varieties that would make European soybean more appealing to the processing industry and profitable enough to satisfy European farmers. This chapter is giving an overview of the European soybean seed quality research and an insight into soybean seed quality progress made at the Agricultural Institute Osijek, Croatia. The studies presented are mainly considering maturity groups suitable for growing in almost all European regions. The most important traits of soybean seed quality discussed are protein content and amino acid composition, oil content and fatty acid composition, soluble sugars, and isoflavones. Defining quality traits facilitates the parental selection in breeding programs aiming to improve the added value properties of final soybean products and enables the exchange of materials between different breeding and research institutions to introduce diversity, which is a prerequisite for genetic advance.

Keywords: soybean, seed quality, breeding, variability, Europe, chemical composition

1. Introduction

Soybean (*Glycine max* (L.) Merr.) is the main oilseed crop of the world [1], a staple crop for protein-rich food and feed as well as a significant source of nutraceutical compounds

with many different medical benefits [2]. One of the major health-promoting traits of soybean seed is the proposed ability to reduce the risk of metabolic disorders, cardiovascular diseases and cancers [3, 4]. Even though the majority of soybean in the European Union (EU) is used as poultry and pork feed, and to a less extent for feeding dairy cows [5], the high nutritional value of the seed and its health-promoting traits are making soybean more and more appealing for human consumption. In the last decades (1998–2017), most of the world's soybean seed was produced in the Americas (85.6%), whereas Europe produced only 1.6% [1]. Moreover, European countries import 95% of the annual demand for soybean grains, meal and oil from overseas, causing an enormous trade deficit, so the governments in Europe had started to introduce additional incentives aiming to stimulate soybean cropping [6, 7]. As a result, the harvested area in Europe has been continuously increasing over the past few years [1]. Furthermore, EU has a relatively high demand for non-genetically modified (non-GM) soybean in comparison to other parts of the world, and food/feed ingredients containing more than 0.9% of GM material must be labelled. This generated a system of segregation and identity preservation (IP) that ensures the identity of non-GM soybean is preserved through the entire supply chain. If we consider that EU imports about 2.7 million tons of non-GM IP soybean meal equivalent yearly and that the premium for non-GM IP is 20–30% of the price of non-segregated soybean while more than 80% of soybean area worldwide is planted with GM soybean [5], inciting the production of non-GM soybean in Europe becomes a matter of high importance. In order to keep the positive trends and rebalance the sources of soybean supply in the future, growing soybean production must be followed by continuous and intensive research in order to create improved varieties which would make European soybean more appealing to the processing industry as well as profitable enough to satisfy European farmers. Furthermore, because the frequency of adverse weather events has been increasing over the last 20 years [8], continuous and intensive research is necessary to create stable European varieties with high seed quality which would become an integral part of conservation agriculture.

Regardless of the final goal, the prerequisite of successful crop improvement is genetic variability of the traits of interest. Assessment of genetic diversity is necessary for germplasm characterisation, conservation, utilisation and the establishment of breeding programmes [9]. If sources of genetic variability are available, the genetic advance can be achieved not only with genetic engineering but also with continuous breeding efforts using conventional hybridization and selection methods together with modern chemical, biochemical and genetic analyses [10]. Conventional soybean improvement starts with the creation of a large recombinant inbred line population of significant diversity by hybridization of chosen parental components carrying the traits of interest. According to Burton [11], genetically distant elite parental lines have the greatest chance of producing superior progeny. Genetic characterisation and evaluation of divergence of parental material prior to hybridization can be made based on the pedigree [12], by measuring and analysing the variability of qualitative and quantitative morphological properties [13], on the basis of biochemical properties [14, 15] and with the help of molecular markers [16, 17]. All phenotype determination

methods must consider the interaction of genotype and environment, whereas the efficacy and reliability of molecular markers are based on the ability of direct and rapid determination of genotype divergence excluding the environmental influences. Even if molecular methods of determining the differences in genetic constitutions are not available, creating a diverse population is enabled by the positive correlation between phenotypic variability and genetic divergence [18]. Therefore, phenotype evaluation is crucial for successful crop improvement.

The process that negatively affects crop improvement is the loss of genetic diversity as a consequence of human activity and the influence of the environment. This process, which Harlan [19] called genetic erosion, appears because of the replacement of diverse indigenous populations with modern, new, uniform cultivars and hybrids, and it causes a considerable threat to the production of food and hence the survival of humans. A narrow genetic base has been identified in most soybean germplasm studies [20, 21]. However, Hahn and Würschum [22] noted that the genetic base of middle European genotypes was not as narrow as expected because of the unsystematic phenotype selection. If narrowing of the genetic base is not stopped, it can result in complete crop destruction due to the lack of tolerance to adverse abiotic or biotic factors. This is why it is important to take all the measures necessary for preserving genetic diversity not only as it is crucial for crop improvement [23] but also to preserve biodiversity which ensures natural sustainability for all life forms. Furthermore, it is crucial to ensure the availability of this biodiversity to breeders, researchers and producers in order to ensure the genetic advance of cultivars in the future. As in other plant species, genetic diversity for soybean is preserved by creating germplasm collections on a local level and gene banks on a global level. Creating large and diverse germplasm collections as genetic resources necessary for the production of high-yielding and high-quality, commercially important cultivars is only possible if there is a continuous research. Since the commercial significance of soybean emanates mainly from the chemical composition of its seed, further improving the genetic basis of soybean seed quality should nowadays be as important as increasing the seed yield. Higher seed yield means more food produced on the same land area, but when 2 billion people worldwide are known to suffer from malnutrition [24], the quality of that food is of concern as well. This is why crops with high nutritional value, such as soybean, can greatly contribute to improving human and animal diets and health, as well as provide quality stock for pharmaceutical and functional food industries.

This chapter is giving insight into soybean seed quality research beneficial for defining the quality traits in available germplasm and choosing parental components in the soybean breeding programmes aiming to improve the added value properties of final soybean products. The studies presented here are mainly focused on maturity groups (MGs) 00 to II commonly sown in Central and South-eastern Europe but suitable for growing in almost all European regions [7]. The most important traits of soybean seed quality discussed here are protein content and amino acid composition; oil content and fatty acid composition; content and composition of soluble sugars, especially oligosaccharides; and content and composition of isoflavones.

2. Protein content and amino acid composition

Soybean protein, taking up 40% of the dry seed weight (DW) on average [25], is highly valued for food and feed because of its amino acid composition and a high digestibility [26]. Although it is thought to be the equivalent of animal protein [27], soybean proteins are deficient in sulphur-containing amino acids of which methionine and cysteine are considered the most limiting in animal feed [28]. Nevertheless, after appropriate heat treatment to reduce protease inhibitor activity, soybean proteins are considered to be superior to the proteins of other legumes in their growth-promoting properties [29]. Soybean seed proteins are made up from four major fractions: 2S, 7S, 11S and 15S, with 7S (glycinin) and 11S (β -conglycinin) being the most abundant [30]. The ratio between 7S and 11S subunits has an important role in determining the functional properties of food made from soybean [31] and protein quality in soybean as glycinin has more sulphur-containing amino acids (cysteine and methionine) than β -conglycinin [32]. Furthermore, glycinin is considered to lower the cholesterol levels in human serum [33], and it is important for tofu gel formation [34], whereas α subunit of β -conglycinin is identified as one of the major allergenic proteins in soybean [32]. The nutritional value, utilization and digestibility of soybean proteins are further affected by bioactive compounds with toxic and/or anti-nutritional properties such as lipoxygenases, lectins, urease, the Kunitz trypsin inhibitor and the Bowman-Birk trypsin inhibitor [35], causing digestive and metabolic diseases in animals [36]. Some of these anti-nutritional factors can be destroyed or inactivated by heat treatment and some by supplemental enzymes, but others are unaffected by the methods applied commercially [37]. Nevertheless, in 2018, soybean represented 70% of protein meal used worldwide [38]. Defatted soybean meal has the highest level of crude protein among plant-based protein sources [26], and it is the main source of protein in commercial feed mixtures for poultry, livestock and fish farms [39]. For the production of these nutritionally valuable proteins, soybean needs less land area and less N than wheat and other small grains, maize, rice, fruits and vegetables. The emergence of soybean as a dominant global crop resulted in a partial offsetting of the dramatic increases in global soil nutrient withdrawals occurring as a result of the expansion of agriculture over newly cultivated land and rising yields in the recent decades while, at the same time, slightly increasing the overall protein content of the global harvest [40]. In the last several decades, the dominance of the soybean on the world crop market was further enhanced by the increase in meat consumption and livestock production, which caused the increase of the demand for high-protein materials for livestock feed in Europe as well [41]. It was estimated that global meat and milk demand would have an increase of 57 and 48% respectively, between 2005 and 2050, due to the fast-growing population and rising incomes in developing countries [42], while livestock production was estimated to increase by 21% between 2010 and 2025 [43]. Plant protein production in the EU, on the other hand, does not follow this increase, so almost 70% of domestic needs are covered by imported plant protein [44]. Because importing 1 kg of dried soybean meal to the European Union from South America is associated with 11.65 kg CO₂ equivalent emissions [45] and because the price of proteins on the world market is continuously rising [46], this imposes major environmental, economic and social problems. Therefore, increasing local plant protein production and superseding the import of feed protein would not only increase

Europe's protein self-sufficiency but also decrease the negative environmental footprint of animal production [47]. Furthermore, including legumes such as soybean in crop rotation dominated mainly by cereals and non-legume oilseeds, which is the case in Europe [44], can have positive effects on soil quality and agrobiodiversity and can contribute to reducing nitrous oxide emissions and nitrate-N leaching [48]. These environmental benefits, together with benefits arising from the better balance of EU agriculture and trade, are the main reasons the provisions of the new Common Agriculture Policy (CAP) include the promotion of protein crops in Europe as a priority [49].

In general, increasing plant protein production can be done by increasing the area under high-protein crops or by increasing the protein content by breeding. Even some management practices can influence protein content. For example, minimum tillage and herbicide application can decrease the protein content in comparison to conventional tillage and no herbicides [50], and irrigation at the beginning of pod formation and during seed-filling can result in protein content increase [51–53]. Furthermore, Foroud et al. [54] and Bouniols et al. [55] noted that protein content was highest when irrigation was applied after the flowering, whereas continuous water supply decreased it. According to Afza et al. [56], nitrogen (N) fertilization during the seed-filling stage increased the protein content, and earlier N application, e.g. in flowering, did not affect the protein content [57], whereas, sulphur-containing amino acid composition fluctuated depending on the nitrogen source [58] and on the availability of reduced forms of sulphur as well [59]. Zimmer et al. [60] reported that protein content of non-inoculated soybeans was significantly lower than protein content of the inoculated soybeans, and Vollmann et al. [61] reported that N fertilization at the flowering stage was superior to both the control and rhizobium inoculation in increasing seed protein content, but as symbiotic N fixation is a highly complex phenomenon, the difference in protein content could have been the result of many different factors and their interactions which demand further investigation. Although correct management practices can increase the protein yield of soybean crops, improving the genetic base of protein content is the right approach for providing high-quality soybean cultivars. Furthermore, because the agroclimatic conditions in Europe are not ideal for widespread cultivation of protein crops like soybeans [62], increasing protein yield per area of arable land by plant breeding is the only efficient way of decreasing the plant protein deficit in Europe in which genetic diversity and population structure have a key role [22].

The soybean protein content is a quantitative trait inherited polygenetically, and it ranges between 30 and 50% on a dry weight basis [10]. Besides earlier mentioned management practices (M), protein content is influenced by genotype (G), environment (E) and interactions $G \times E \times M$, but the contribution of each is still not well established [50, 63–67]. The G and E interaction is one of the main problems in genotype selection as well as in the recommendation of cultivars. Because of this, estimating the broad sense trait heritability is essential for successful genetic advance and the larger the estimated value of this parameter is, the greater will be the chance of success with selection [68]. Broad-sense heritability assessments for soybean vary from middle, which suggests variability due to the combination of genetic and environmental factors, to high, where genetic factors are more pronounced in determining the protein phenotypes [67, 69, 70]. The influence of the environment on protein content has been researched by many authors. According to Popović et al. [71] and Josipović et al.

[66], protein content is higher in years with lower average temperatures and more precipitation during pod formation and seed-filling stages. On the other hand, Hurburgh [72] and Nian et al. [73] reported reduced protein content in northern regions of soybean cultivation with lower temperatures and higher amounts of precipitation, which can be the result of reduced symbiotic N fixation and synthesis of proteins due to the low root-zone temperatures [74]. Gibson and Mullen [75] reported that high daily temperatures in combination with high night temperatures during the period from seed-filling to maturity can cause positive linear relationship of average temperatures and protein content. Furthermore, Dornbos and Mullen [76] reported that high temperatures in combination with drought in seed-filling period can result in significant protein content rise, whereas Vollmann et al. [61] and Matoša Kočar et al. [77] noticed that seed protein content was highest for soybean crops grown under moderately dry conditions and high temperature during the seed-filling stage. Such contradictory reports are a major concern in plant breeding since they complicate decision-making process during selection. Breeding for increased protein content is further complicated by negative correlation between protein content and yield reported by many authors [70, 78, 79], although there are studies indicating no significant correlation between those two parameters [80, 81]. Furthermore, Vollmann et al. [61] suggested that even moderately negative correlation between seed protein content and seed yield, means the selection of breeding lines with both improved protein content and acceptable yield level should be possible. On the other hand, the relationship between protein content and oil content is almost always significantly negative [61, 70, 81] which is caused by either pleiotropic effect or linkage [82] where every 2% protein content increase usually decreases oil content by 1% [83]. Apart from protein content, the content of sulphur-containing amino acids should be considered as it affects the nutritional value of soybean meal, so increasing the content of both would be beneficial. However, Wilcox and Shibles [84] determined a weak and inverse correlation between protein content and methionine and cysteine levels which means protein quality decreases with the increase in protein content, so Paek et al. [58] concluded that it would be beneficial to focus on improving protein quality over protein content. Nevertheless, Burton et al. [85] noticed no significant changes in methionine content among cycles of recurrent selection for high protein. Furthermore, Krober and Cartter [86] reported a positive relationship between methionine content and protein content suggesting the possibility of combining high-yield, high-quality protein soybean lines.

As mentioned earlier, improving any trait of interest should start with screening the available materials. In Europe, soybean seed content was researched by many authors over the years. Vollmann et al. [61] determined a considerable variation in protein content of early maturing genotypes studied during 6 years in Austria. In their study, genotype \times treatment and genotype \times year interactions were of a much lower magnitude than the effect of genotype on protein content [61]. Sudarić et al. [63] researched 14 MG 0-I soybean genotypes at three locations in eastern Croatia during 5 years and reported highly significant effects of genotype, year, location, genotype \times year, genotype \times location interactions as well as a highly significant effect of the interaction between genotype, year and location with year effects being larger than location effects. Protein content range in the research by Popović et al. [71] was 36.65–37.66%, and average value was 37.15%. Fogelberg and Recknagel [87] determined the protein content

range of early matured, well-adapted soybean cultivars evaluated at three sites in Germany during 3 years to be 37–43%, while Pannecouque et al. [81] reported a protein content range of 35.5–43.3% in 14 early maturing soybean varieties tested in two consecutive years in two locations in Belgium. In a study by Kurasch et al. [70] two types of European soybean varieties that maximise protein yield were distinguished: one that produces high grain yields per hectare with protein contents around 39–42% and the other with slightly lower protein yields than the first type but characterised by very high-protein contents (above 42%). Soybean breeding at the Agricultural Institute Osijek (Croatia) resulted in an increase of protein content over time, and this progress can be followed through three studies. The first one [88], conducted in the 1990s, researched 22 soybean cultivars (11 standard cultivars and 11 promising lines) whose protein content range was determined to be 35.9–38.4%. Afterwards, in a 3-year study (2001–2003), Vratarić et al. [89] tested 132 elite breeding lines and standard cultivars grouped according to their MG (00, 0, I) and reported that for all MGs, elite breeding lines had significantly higher values for protein content than their respective standards. Ranges determined were 35.2–38.7%, 36.2–39.6% and 38.7–41.2% for MG 00, 0 and I, respectively [89]. In the third research, Matoša Kočar et al. [77] determined significant variation for soybean seed protein content, influenced by genotype and year in a 3-year (2010–2012) screening study with 21 soybean genotypes (MG 00 to I) created at the Agricultural Institute Osijek (Croatia). The lack of significant difference in average protein content between elite breeding lines created at the Agricultural Institute Osijek and standard cultivars (**Table 1**) can be explained with the significant progress already achieved in previous selection cycles [89]. Nevertheless, wide ranges of variation for protein content in genotypes from the Agricultural Institute Osijek indicate there is still a room for improvement [77]. Although, according to Sato et al. [90], most of the world is focused towards increasing soybean seed oil content and oil yield, average protein content is at the world level higher than protein content reported for Central European soybean varieties [61]. For example, Bueno et al. [64] reported that average protein content value in 18 Brazilian soybean genotypes was 42.44% (40.20–44.49%), Sharma et al. [91] reported average protein content values being 41.4% (39.40–44.40%) in eight genotypes in India, while Ramteke et al. [92] reported somewhat lower average protein content (40.23%) determined in 92 soybean varieties in India with values ranging from 37.69 to 42.74%. Furthermore, protein content determined in the whole US Department of Agriculture (USDA) soybean germplasm

MG	Elite breeding lines		Standard cultivars		t	P
	Range	Mean	Range	Mean		
	Protein (% DW)					
00	35.86–43.82	39.87	35.52–43.55	39.44	0.37 ^{ns}	P > 0.05
0	35.83–44.33	40.42	34.14–42.48	39.08	1.28 ^{ns}	P > 0.05
I	36.02–44.04	40.19	36.34–42.70	39.07	1.40 ^{ns}	P > 0.05

The means were tested with Student's t-test with ns indicating no significant difference.

Table 1. The 3-year (2010–2012) protein content range and mean for different maturity group (MG) elite breeding lines created at the agricultural institute Osijek in comparison to standard cultivars.

collection ranged from 34.1 to 56.8% [93]. Higher average protein contents in non-European studies are maybe due to the fact that soybean breeding is still at a relatively low level in Europe compared to the USA or Asia because most of the soybean demand in Europe has been relying on import [44], so developing high-quality genotypes was not of great economic importance. Therefore, European breeding programmes would benefit from introducing foreign accessions with high-protein contents to be used as parental components in crosses, providing these accessions are not transgenic, since that is not socially acceptable in Europe [70].

If sources of variation for high-protein content are available, the next step is choosing the most efficient breeding method. Sato et al. [90] reported that crossing adapted elite genotypes with advanced high-protein donors exhibiting a seed protein content of approximately 470 g/kg protein improved protein content in progeny more than crossing adapted genotypes and food-grade cultivars. Increasing the protein content from 46.3% in the initial parental population to 48.4% was achieved after six cycles of selection, without significantly reducing yield [78]. Furthermore, Thorne and Fehr [94] concluded that three-way crosses were superior to two-way crosses for creating high-protein, high-yielding lines. Wehrmann et al. [95] reported that selection for high protein between two backcross generations increased protein content while maintaining seed yield. Wilcox and Cavins [96] were able to create high-protein BC₃ line with the yield at the same level as the recurrent parents, but it took them 20 years because it was necessary to select for high protein between each backcross. Cober and Voldeng [80] evaluated single-cross and backcross breeding methods of achieving high-yield, high-protein soybean genotypes and found no significant differences between them. Selecting appropriate genotypes during the breeding process can be made easier and more efficient with marker-assisted selection (MAS), especially since quantitative trait loci (QTL) for proteins have been mapped to locations on all chromosomes. QTL mapping and genome-wide association studies (GWAS) have identified 252 QTL associated with soybean protein and distributed on all 20 soybean chromosomes [97]. Brummer et al. [98] evaluated eight soybean populations from the Midwestern USA for genetic markers linked to seed protein content detecting a significant association between markers and traits and identifying environmentally stable and sensitive QTL. Sebolt et al. [99] successfully integrated *G. soja* alleles for high-protein content into *G. max* background by backcrossing. Genetic marker alleles linked to the QTL allele from *G. soja* on linkage group (LG) I were significantly associated not only with higher protein content but with lower oil content, reduced yield, smaller seeds, taller plants and earlier maturity. Hwang et al. [100] used GWAS to identify QTL controlling seed protein content in 298 soybean germplasm accessions exhibiting a wide range of protein content and identified 40 single-nucleotide polymorphisms (SNPs) in 17 different genomic regions significantly associated with seed protein. GWAS also resulted in narrowing the genomic region previously reported to contain protein content QTL [101] which is important because narrower GWAS-defined genome regions will allow more precise marker-assisted allele selection and will expedite positional cloning of the causal gene(s) [100]. Bandillo et al. [102] studied 12,000 accessions from the USDA soybean germplasm collection by using GWAS and had identified SNPs for protein and oil content with strong signals on chromosomes 20 and 15 with chromosome 20 region, previously reported to be important for protein and oil content, further narrowed so it contained only three plausible

candidate genes. Besides genomic studies, identifying markers for protein content can be done using metabolomics, a non-targeted approach monitoring hundreds of metabolites, but there is still little knowledge about the association between metabolites and protein content in soybean seeds, especially in similar genetic backgrounds [103]. Nevertheless, Wang et al. [103] evaluated metabolic diversity in a soybean near-isogenic line (NIL) population derived from parents with contrasting seed oil contents comparing seed primary metabolites of high-protein/low oil lines, low-protein/high oil lines and their parents. Results indicated that metabolic profiles of all progeny lines could be discriminated based on protein and oil contents. All such molecular and genomic investigations help to reveal the genetic architecture of complex traits, protein content being one of them, and enable understanding of the genetic basis of trait variation which is crucial for improving seed quality as well as other important agronomic traits. Nonetheless, local cultivar testing is particularly important for the development of the crop, despite the wide adaptation of soybean, since there could be substantial site effects within regions [87] which mean it would be beneficial for the European soybean breeders to develop their own, local varieties with increased protein content, rather than relying on foreign introductions.

The next step after increasing soybean seed protein content could be breeding for more favourable amino acid composition and improved 11S/7S protein fraction ratio. For example, while the manipulation of other aspects of seed composition and processing may improve amino acid assimilation, increasing the relative proportion of methionine, lysine and threonine has become a goal in soybean breeding [104]. Almost 20 years ago, the economic benefit of improved essential amino acid content was estimated to be ~\$5 per ton per 10% increase of any of the above-mentioned amino acids [105], and today the benefit could be even bigger. The knowledge of the molecular mechanism of amino acid biosynthesis in soybean was mainly limited to genetic mapping [106, 107], but recent studies have enhanced our understating of amino acid genetic architecture in soybean. For example, Vaughn et al. [108] conducted GWAS for methionine, threonine, cysteine and lysine and identified multiple loci associated with multiple amino acids across different populations not reported previously. Furthermore, Qiu et al. [109] predicted 12 candidate genes based on the synteny between 113 genes of sulphur-containing amino acid synthases and 33 related QTLs in soybean and bioinformatic analyses, and many QTL related to the 7S (β -conglycinin) and 11S (glycinin) fractions of soybean storage proteins have been identified on chromosomes 1, 3, 4, 6, 10, 13, 16, 17, 19 and 20 [110–112]. The 11S/7S ratio of soybean seed can vary greatly due to both genetic and environmental differences [113, 114], which means there is room for improvement. In a study by Žilić et al. [32], the ratio of 11S/7S proteins varied from 2.43 to 3.29 among the seven soybean varieties adapted to South-Eastern European conditions. Murphy and Resurreccion [113] reported the 11S/7S protein ratio to vary from 2.1 to 3.4 in 12 soybean varieties, while Mujoo et al. [114] determined this ratio to be 1.63–2.05 in 7 soybean varieties. Since 7S protein fraction is more favourable for food and feed, lowering 11S/7S ratio would be beneficial, but at the moment it is still far from a priority for European breeding programmes. Furthermore, according to Zhang et al. [115], breeding efforts for the improvement of soybean seed amino acids have rarely been reported, much less developed on a commercial scale, because methods used for quantifying amino acids in soybean, such as HPLC, are time-consuming and not

cost-effective, therefore not suitable for high-throughput screening of a large number of lines [116].

3. Oil content and fatty acid composition

Soybeans represent 59% of the world's oilseed production and 29% of the total vegetable oil consumption in the world [117]. Oil from soybean seed can be found in margarine, salad dressings and cooking oils and in industrial products such as plastics and biodiesel fuel as well. The content of oil in soybean seed can range from 12 to 24% DW, while most commercial cultivars contain between 19 and 23% [110]. Soybean oil is composed of triacylglycerols (TAGs), which are products of fatty acid esterification [118]. Soybean oil fatty acids are responsible for nutritional value, stability and taste of soybean oil, and commodity soybean oil is in average made up from 13% palmitic acid (16:0), 4% stearic acid (18:0), 20% oleic acid (18:1), 55% linoleic acid (18:2) and 8% linolenic acid (18:3) [119]. Fatty acids in soybean can be saturated (palmitic and stearic acid) or unsaturated (oleic, linoleic and linolenic acid). Soybean oil intended for frying should, for example, contain 7% saturates, 60% oleic acid, 31% linoleic acid and 2% linolenic acid, whereas desired fatty acid composition of oils intended for industrial use is 11% saturates, 12% oleic acid, 55% linoleic acid and 22% linolenic acid. Nevertheless, the most desirable fatty acid phenotypes for soybean oil are those having the most applications for food, feed and industry, i.e. phenotypes with saturates reduced to less than 7%, linolenic acid reduced to less than 3% and oleic acid increased to more than 55% [93]. Lowering saturates makes oil more suitable for the food industry and consumers concerned about dietary health issues such as high cholesterol and increased risk of coronary heart disease usually associated with diets high in saturated fats [120]. However, as stearic acid is thought to be heart neutral [121], increasing its content could satisfy the market demand for healthier foods without compromising desired functional qualities contributed by saturated fats [122]. Unsaturated fatty acids can be monounsaturated (MUFA), such as oleic acid (n-9, omega-9), or polyunsaturated (PUFA), such as linoleic (n-6, omega-6) and linolenic (n-3, omega-3) acids. Oleic acid is not considered to be an essential fatty acid (EFA) since the human body can synthesise it in small amounts, but it is significant as a precursor of some EFAs [123]. On the other hand, polyunsaturated linoleic and linolenic acids are considered essential, supporting the cardiovascular, reproductive, immune and nervous systems, crucial for manufacturing and repairing cell membranes [123]. However, these PUFAs are susceptible to oxidation, so they reduce the shelf life of the oil, causing low stability at high cooking temperatures as well as off-flavours [124]. Oxidative stability of soybean oil generally decreases with the increasing degree of unsaturation [125] and can be assessed by the ratio between MUFA and PUFA contents (MUFA/PUFA) [126]. Soybean oil has poor oxidative stability compared to other vegetable oils, and its MUFA/PUFA is about 0.5 on average [126]. The average value for MUFA/PUFA in the research by Matoša Kočar et al. [127] was 0.4 suggesting these genotypes would give less stable final product, i.e. edible oil with shorter shelf life and lower stability for cooking at high temperatures [126]. Oxidative stability of soybean oil can be improved in two ways: by trans-isomer producing catalytic hydrogenation or by breeding for higher content of oleic acid (MUFA), which is also known to reduce cholesterol and reduce the risk of

arteriosclerosis and heart disease [93, 124, 126]. On the other hand, diets rich in trans-fatty acids increase the risk of cardiovascular diseases [128], so increasing oil stability by hydrogenation is becoming more and more unpopular. Soybean oil with high-oleic acid (>70%) and low-linolenic acid content (<3%) is desirable for its oxidative stability and health benefits, with the linolenic acid content of <3% being the current market target to increase soybean oil functionality [129]. As oxidative stability of soybean oil can be assessed by MUFA/PUFA, nutritional quality can be assessed by the linoleic and linolenic acid ratio (LA/ALA) with an optimal value between 5 and 10:1 [126]. Three-year average LA/ALA reported in the study by Matoša Kočar et al. [127] was 7.36, which falls within the recommended range for vegetable oils suggested by Rani et al. [126], meaning that tested genotypes can produce nutritionally acceptable oil. On the other hand, very high LA/ALA is considered detrimental for human health, and so lowering it by breeding can help with the prevention against degenerative pathologies [130]. Even though the demand for vegetable oil is continuously increasing and this increase is expected to continue in the period from 2018 to 2027, as the production of oilseeds is mainly concentrated in few regions of the world, weather fluctuations and extreme weather having a negative effect on the yield impact oilseed market more than other major crop markets, causing market uncertainties and price volatility [131]. This, together with evident changes in consumer preferences concerning dietary value and functional properties of oil, is causing the world oilseed market to become more competitive, which encourages breeding programmes to focus not only on increasing oil yield but also on increasing oil quality to meet the demands of industry and end-users alike [132]. Furthermore, because European consumers still prefer GM-free oilseed end products, while the area planted with soybean, the main oilseed crop of the world [1], is more than 80% planted with GM soybean [5], European soybean breeders should focus on creating high-oil GM-free varieties with conventional breeding methods so that they are able to provide high-quality and healthy end products to European consumers. Seed oil content is a complex quantitative trait, controlled by multiple genes and affected by environmental factors [124, 133]. The broad-sense heritability of soybean oil content varied considerably in the range between 0.13 and 0.99 [7, 134–136] suggesting that trait inheritance is a function of G, E, M and $G \times E \times M$. For example, oil content can be in positive correlation with temperature [137, 138], but negative linear relationships [139] and quadratic relationships [140] were also reported. Water stress negatively impacts oil content [141], but irrigation at the beginning of pod formation and during seed filling has resulted in oil content decline [51–53]. Some other crop management factors (no-till, seed treatments, foliar N, fungicide and insecticide applications and rotation) had an overall positive effect on oil content because they were improving crop growing conditions through conserving or supplying water and nutrients or improving soil physical and chemical properties and protecting the crop from diseases [142]. On the other hand, a decline in oil content with increased N application [141] and lack of response [143] were reported as well. While monocropping negatively affected soybean oil composition in comparison to crop rotation with maize [144], early planting benefited oil [145–147]. Furthermore, Assefa et al. [142] reported that planting date had larger impacts at northern latitudes (40–45 N) than southern latitudes (30–35 N). Although the influence of M for oil content was found to be significant [144–147], in a comprehensive analysis of G, M and E factors influencing soybean yield quantity and quality across the US Corn Belt, Assefa et al. [142] concluded that E is a dominant factor for the significant variability in seed composition

and yield. According to Josipović et al. [66] and Popović et al. [71], the amount of oil in soybean seed was significantly higher in years with less precipitation and higher air temperatures at the time of pod formation and dry matter accumulation. This was also true for the average oil content determined by Matoša Kočar et al. [127], noting that hot and extremely humid conditions resulted in the lowest average amount of oil in soybean seed, whereas hot and dry conditions resulted in higher average amounts of oil. On the other hand, Vollmann et al. [61] reported that low precipitation rates and high temperatures did not favour oil synthesis and that, in genotypes of later maturity, oil synthesis was enhanced by late water availability. The significance of genotype and environmental factors for fatty acid content was determined by many authors as well [7, 61, 65, 67, 127, 146, 148]. Bellaloui et al. [146] noted that, among fatty acids, oleic acid was the most sensitive to environmental factors and palmitic and stearic acids were the least sensitive in clay soil, whereas stearic, linoleic and linolenic acids were the least sensitive in sandy loam soil. According to the same authors, cooler temperatures favour the synthesis of linolenic acid, whereas the synthesis of oleic acid is negatively affected as a result of the inverse relationship between them [146]. The same was true in the research done by Xue et al. [149], where increasing air temperature during pod fill significantly increased oleic acid and significantly decreased linoleic and linolenic acid contents which was later confirmed by Matoša Kočar et al. [129].

Together with understanding the sources of variation, knowing the strength and direction of relationships among different traits is of great value for breeding programmes as it enables the enhancement of more than one trait at the same time. Although significantly negative relationship between oil and protein content was reported in most researches [61, 70, 81, 84], in the research analysing variations from 21 studies conducted over 15 years (2002–2017), Assefa et al. [142] concluded that, when pooling data across all studies, there was no significant relationship between oil and protein content, but a tendency for a negative relationship was observed when plotting data separately for each of the studies evaluated in the database. Furthermore, oil content increased slowly with yield increase suggesting a positive relationship, but when relationships were investigated by study, 63% of studies supported a positive relationship, and the other 37% displayed a slightly negative relationship between seed yield and oil content [142]. Wilcox and Shibles [84] reported that oil increased by 1.9 g kg^{-1} for every 100 kg ha^{-1} increase in seed yield, whereas protein decreased by 15.6 g kg^{-1} for each 10 g kg^{-1} increase in oil, which is in accordance with determined significant positive correlation between oil content and yield and significant negative correlation between oil and protein contents. The analysis of the correlation between oil content and different fatty acid contents in eight early maturing soybean genotypes revealed a positive correlation between oil and saturated fatty acids but negative correlation between oil and unsaturated fatty acids [115]. Matoša Kočar et al. [127] reported a highly significant and positive correlation of oil with stearic acid and significant positive correlation with oleic acid but significant negative correlation with linoleic and linolenic acid. The results were somewhat different in the research conducted by Rani et al. [126] where oil was in highly significant positive correlation with stearic acid and linoleic acid but in negative correlation with oleic acid. The correlation between favourable oleic acid and linoleic acid was highly significant and negative in the research by Matoša Kočar et al. [127], the same as in some previous studies [126, 150], while

the correlation between oleic acid and unfavourable linolenic acid was also highly significant and negative. These relationships indicate that it could be possible to create varieties with high oil amount, a high amount of oleic acid and low amount of linolenic acid, which is favourable for edible oil.

Whether the improvements in soybean oil yield and its quality are achieved by conventional breeding or genetic engineering [2, 119], having a favourable gene pool concerning the trait of interest is crucial. In Europe, the variability of oil content was studied by many authors [61, 70, 77, 88, 89, 127]. At the Agricultural Institute Osijek, breeding for soybean seed oil over time is presented by three studies. In the first one [88], the range for oil content determined in 22 soybean cultivars tested from 1993 to 1995 was 18.9–20.5%. The second one [89] tested 132 elite breeding lines and standard cultivars grouped according to their MG (00, 0, I) during 3 years (2001–2003). All three elite breeding line groups had significantly higher average oil contents than their respective standards. The range for MG 00 was 22.1–23.8%, for MG 0 22.5–23.4% and for MG I 21.9–22.7% [89]. The third research from the Agricultural Institute Osijek, conducted between 2010 and 2012 with eight MG 0 advanced soybean breeding lines, reported oil content ranges from 22.09 to 24.08% [77]. As it is evident from **Table 2**, there were no significant differences determined between the average oil content of elite breeding lines created at the Agricultural Institute Osijek in the last decade in comparison to standard cultivars. This is expected because significant progress has been achieved in earlier selection cycles and oil content in genotypes from Osijek is already at a relatively high level. Furthermore, as

MG	Elite breeding lines		Standard cultivars		t	P
	Range	Mean	Range	Mean		
	Oil (% DW)					
00	20.59–25.48	23.29	21.01–23.81	22.54	1.49 ^{ns}	P > 0.05
0	20.24–25.84	22.85	20.06–23.85	22.15	1.45 ^{ns}	P > 0.05
I	20.93–26.40	23.24	20.66–24.40	22.40	1.76 ^{ns}	P > 0.05
	Oleic acid (% oil)					
00	20.58–31.72	25.72	21.37–27.51	23.92	1.67 ^{ns}	P > 0.05
0	19.51–28.91	23.59	18.74–24.84	21.38	2.68 ^{**}	P < 0.01
I	20.27–28.05	24.78	19.48–26.83	23.54	1.30 ^{ns}	P > 0.05
	Linolenic acid (% oil)					
00	4.41–7.04	5.99	5.36–8.27	6.93	2.23 [*]	P < 0.05
0	4.99–8.65	7.15	6.97–7.99	7.59	2.54 [*]	P < 0.05
I	5.10–7.98	6.59	5.37–8.73	7.17	1.37 ^{ns}	P > 0.05

The means were tested with Student's t-test with * and ** indicating significant differences at $P \leq 0.05$ and $P \leq 0.01$, respectively, and ns indicating no significant difference.

Table 2. The 3-year (2010–2012) oil, oleic and linolenic acid content range and mean for different maturity group elite breeding lines created at the agricultural institute Osijek in comparison to standard cultivars.

soybean seed is mostly used for animal feed in Croatia; further increasing oil content is not a priority. However, there were genotypes among the newer elite breeding lines that accumulated significantly higher amounts of oil than the rest, indicating their suitability for being used as parental components [77]. Somewhat lower oil content range values (18.55–18.73%) were reported by Kurasch et al. [70] researching 1008 $F_{5,8}$ recombinant inbred lines MG (MG 000–00) that showed good agronomic performance in Central Europe. Pannecouque et al. [81] determined the oil content range being 20.3–24.1% in 14 very early (MG 0000–00) soybean genotypes selected from the European plant variety catalogue. Outside Europe, Bueno et al. [64] reported 22.1% average oil content determined in 18 Brazilian soybean genotypes with values ranging from 20.72 to 22.81%, and Sharma et al. [91] noted 16.3% average oil content for 8 genotypes in India ranging from 14 to 18.7%, whereas oil content determined in the whole US Department of Agriculture soybean germplasm collection ranged from 8.1 to 27.9% [93]. Variability of soybean seed fatty acid content in European commercial breeding programmes has seldom been researched, but in other parts of the world where soybean is considered the main crop feeding oil industry, many studies have been conducted. Significant variability of the fatty acid content was determined in eight advanced soybean breeding lines (MG 0) developed at the Agricultural Institute Osijek, Croatia [127]. The contents were 10.81, 5.78, 26.41, 49.75 and 5.97% for palmitic, stearic, oleic, linoleic and linolenic acid, respectively [127], which mostly coincided with fatty acid content in commercial soybean [151]. The positive outcomes of initial breeding efforts for increasing oleic and reducing linolenic acid contents at the Agricultural Institute Osijek can be seen from the significant difference between the average oleic acid content of MG 0 elite breeding lines in comparison with the MG 0 standard cultivars and significant difference between the average linolenic acid content of MG 00 and 0 elite breeding lines in comparison with the MG 00 and 0 standard cultivars, respectively (**Table 2**). Nevertheless, according to Matoša Kočar et al. [127], none of the newly developed genotypes from the Agricultural Institute Osijek exhibited desirable values for oleic or linolenic acid contents; the breeding programme would benefit from the introduction of high-oleic and low-linolenic soybean germplasm. Average year values determined in a 3-year research with RILs exhibiting a broad variation for fatty acid content ranged between 10.3 and 11.1%, 3.7 and 3.9%, 22.3 and 26.2%, 51.4 and 55% and 7.5 and 8.6% for palmitic, stearic, oleic, linoleic and linolenic acid, respectively [148]. Average fatty acid contents determined in 96 diverse accessions originating from different regions of the world and evaluated for 2 years in Brazil were 10.59, 3.21, 24.54, 52.81 and 6.36% for palmitic, stearic, oleic, linoleic and linolenic acids, respectively [152]. In Indian soybean germplasm, the average fatty acid contents were 11.15, 3.45, 25.6, 52.56 and 6.91% [126], while in Chinese they amounted to 11, 3.63, 21.47, 53.99 and 9.93% [153] for palmitic, stearic, oleic, linoleic and linolenic acids, respectively. As it can be seen from the data presented, none of the genotypes exhibited target values for oleic acid (45–60%) and linolenic acid (<3%) contents desirable for avoiding partial hydrogenation [93] which emphasises the need for further research and screening studies.

Nowadays, breeding for increased oil content can be facilitated by genetic markers or by monitoring metabolites, but although Wang et al. [103] reported that metabolic profiles of all progeny lines could be discriminated based on oil contents, there is still lack of sufficient information about the association between metabolites and oil content in soybean seeds.

Nevertheless, QTL mapping and GWAS have identified 322 QTLs associated with soybean oil distributed over all 20 soybean chromosomes but mainly on chromosomes 5, 15 and 20 [97]. Fan et al. [148] identified 35 additive QTLs underlying individual fatty acid contents in a single environment and 17 additive QTLs across multiple environments or underlying multiple fatty acids. Priolli et al. [152] discovered 19 single-nucleotide polymorphism loci on 10 different chromosomes significantly associated with palmitic acid, oleic acid and total oil contents with loci and specific alleles that contributed to lower palmitic and higher oleic acid contents. Thapa et al. [154] identified independent mutations in the FAD3A associated with a reduced level of linolenic acid. Combs and Bilyeu [129] reported that seed stearic acid increased to 10–11% in lines containing combinations of FAD2-1A and FAD2-1B mutant alleles plus the SACPD-C missense mutant alleles, but this increase was associated with a decrease in the oleic acid content and did not meet the target of at least 20% stearic acid in the seed oil. Of all the reported QTLs associated with oil traits in soybean, only two (cqPro/oil-15, cqPro/oil-20) have been officially confirmed and repeatedly detected in several different populations, both associated with protein and oil contents and each showing opposite additive effect directions for the two traits, which is why identifying an environmentally stable major QTL regulating seed oil content is of crucial importance [100, 102, 103, 155]. Moreover, QTLs directly related to seed oil accumulation in soybean have not been cloned, so the underlying mechanism has not been thoroughly elucidated to date [133]. Although advance in oil content has been achieved over time, further increasing oil content by traditional breeding based on genetic crossing and phenotypic selection may be difficult and inefficient, because of the polygenic nature of oil regulation and the majority of oil-related loci having varying additive, epistatic or QTL \times E effects [156–158]. Some studies suggest that increasing soybean oil could be achieved by genetic engineering of transcription factors involved in oil accumulation [159, 160]. As further increasing oil content is deemed challenging without genetic engineering [156–158] and taking into consideration that European oil industry is dominated by sunflower and rapeseed [1], it is evident that improving oil content and oil quality in Europe is not economically as important as in regions of the world where soybean is the main stock for oil production. Nevertheless, research and genotype screening for determining germplasm with favourable QTLs controlling oil traits are of great importance, especially since fatty acid phenotypes that meet the breeding objectives of providing improved oil quality are still lacking, not just in European soybean germplasm but in genotypes from all regions of the world.

4. Soluble sugars

Totally dry soybean seed contains 33% DW of carbohydrates on average, of which 16.6% DW is soluble sugars [25]. The significance of the soluble sugar profile is in its effect on quality, digestibility and nutritional value of soybean for food and feed. Five main soluble sugars in soybean seed are glucose, fructose, sucrose, raffinose and stachyose with sucrose and stachyose being the predominant ones [161]. Monosaccharides glucose and fructose and disaccharide sucrose can be easily digested and give soybean food products their characteristic sweet taste, whereas indigestible galactooligosaccharides (stachyose and raffinose) limit soybean food and feed nutritional

value, reduce metabolizable energy and cause irritation of the gastrointestinal tract in humans and animals [162, 163]. In consequence, the use of soybean for food and feed is limited, but as the demand for animal feed in Europe rises [41], lowering galactooligosaccharides would increase the share of soybean seed used for feeding livestock, thus increasing the market for seed production. Furthermore, due to the increasing awareness of the health benefits connected with soy food consumption, more favourable saccharide profile would make soybean more amenable to human consumption. However, oligosaccharides are not necessarily undesirable because they serve as transport carbohydrates in the phloem, as cryoprotectants, and reportedly play a positive role in desiccation tolerance during seed maturation [164]. Furthermore, in human intestines, galactooligosaccharides ferment into low-chain fatty acids with prebiotic qualities, which make soybean seed interesting for pharmaceutical as well as functional food industry [165].

Broad-sense heritability is considered to be moderate to high for monosaccharides and high for oligosaccharides [166–168]. Geater et al. [169] determined that the differences in total sugar (TS) content among tested genotypes were constant in all environments, but significant year effect was determined for raffinose. Taira [170] found G to be the larger source of variation than E for raffinose and stachyose, whereas the opposite was true for TS and sucrose. Both G and E were statistically significant sources of variation for sucrose in the research by Maughan et al. [171]. Matoša Kočar et al. [172] reported that TS content was mostly under genetic control, whereas individual soluble sugars, as well as total oligosaccharides (TO), were significantly influenced by the year which means breeding for TS content should be much more predictable than altering sugar profiles in that particular set of genotypes. In mentioned research, glucose, fructose and raffinose contents were higher in years with higher temperatures, whereas the opposite was true for TO, sucrose and stachyose contents [172]. Wolf et al. [173] concluded that sucrose and stachyose concentrations decreased with the rise of average year temperature; however, the temperature had no effect on glucose, fructose and raffinose. The often occurring absence of G × E interaction [169], together with the determined maternal and additive effects [171–174], facilitates the selection and breeding for improved saccharide contents in soybean. To lower the expenses and reduce the time invested in analyses, plant breeders rely on favourable relationships between important traits which enable effective indirect selection. Significant positive correlation between raffinose and stachyose [161, 175] and significant and very strong negative relationship between sucrose and raffinose and/or stachyose [175, 176] enable breeders to increase the contents of favourable soluble sugars while decreasing the contents of unfavourable oligosaccharides simultaneously. Furthermore, Akond et al. [177] noted that the lack of strong correlations among tested soluble sugars indicated it should be possible to obtain lines that are high in sucrose but low in raffinose and stachyose. No significant correlation between raffinose, stachyose and sucrose was noticed by Geater et al. [169] as well. However, Hartwig et al. [178] reported a significant moderate positive correlation between raffinose and sucrose, while Hou et al. [161] determined the relationship between sucrose and raffinose to be significant, positive and strong. Another unfavourable relationship is the negative correlation determined between protein content and TS and protein content and sucrose [84, 178–179], whereas the negative correlation determined between protein content and raffinose [178] and between stachyose plus raffinose and protein content [84] is considered desirable. The difference in correlation reports from different researchers emphasises the importance of determining trait relationships in each breeding population.

Significant difference in sugar contents, i.e. variability among soybean genotypes that justifies selection, was determined in many studies [25, 161, 172, 176, 180–182]. Geater and Fehr [181] reported TS determined in 23 soybean cultivars tested at eight Iowa locations to range between 2.19 and 18.4%. The range for TS content determined by Matoša Kočar et al. [172] in a 3-year trial (2010–2012) with 22 soybean genotypes (MG 00-II) created at the Agricultural Institute Osijek (Croatia) was 5.69–7.68% with the average value being 6.62%. In the same study, glucose content ranged 0.16–0.29%, fructose 0.11–0.28%, sucrose 2.32–3.46%, raffinose 0.66–1.10% and stachyose 1.64–2.41% [173]. Breeding for improved soluble sugar content is at its beginnings at the Agricultural Institute Osijek, but some progress has been achieved in elite breeding lines, in which higher glucose, fructose, sucrose and TS contents and a lower TO content have been determined than the standard cultivars of

MG	Elite breeding lines		Standard cultivars		t	P
	Range	Mean	Range	Mean		
	Glucose (% DW)					
00	0.17–0.33	0.25	0.12–0.27	0.19	3.05**	P < 0.01
0	0.11–0.35	0.23	0.13–0.19	0.16	6.58**	P < 0.01
I	0.13–0.42	0.24	0.12–0.23	0.19	2.84**	P < 0.01
	Fructose (% DW)					
00	0.04–0.30	0.18	0.05–0.25	0.16	0.65 ^{ns}	P > 0.05
0	0.02–0.38	0.17	0.02–0.18	0.11	2.87**	P < 0.01
I	0.02–0.43	0.21	0.02–0.19	0.13	3.11**	P < 0.01
	Sucrose (% DW)					
00	2.56–3.91	3.18	2.20–2.83	2.62	4.19**	P < 0.01
0	1.77–3.96	3.01	1.89–2.71	2.32	5.74**	P < 0.01
I	1.13–3.78	3.03	2.34–2.95	2.63	4.55**	P < 0.01
	Total oligosaccharides (% DW)					
00	2.47–3.24	2.76	2.66–3.29	3.08	3.95**	P < 0.01
0	1.81–3.53	2.75	3.02–3.34	3.18	6.83**	P < 0.01
I	1.05–3.38	2.66	2.47–3.38	3.04	2.94**	P < 0.01
	Total soluble sugars (% DW)					
00	5.23–7.67	6.97	5.56–6.95	6.09	3.99**	P < 0.01
0	4.96–9.19	6.67	5.01–6.63	5.73	4.17**	P < 0.01
I	5.04–9.21	6.68	2.95–7.25	5.69	1.75 ^{ns}	P > 0.05

The means were tested with Student's t-test with * and ** indicating significant differences at $P \leq 0.05$ and $P \leq 0.01$, respectively, and ns indicating no significant difference.

Table 3. The 3-year (2010–2012) soluble sugar content range and mean for different maturity group elite breeding lines created at the Agricultural Institute Osijek in comparison to standard cultivars.

the respective MG (**Table 3**). Hou et al. [161] investigated worldwide soybean germplasm collections with 241 genotypes and reported average contents for TS, glucose, fructose, sucrose, raffinose and stachyose to be 96.4, 5.4, 4.4, 46.8, 8.3 and 31.7 mg g⁻¹, respectively. Values varied in wide ranges with genotypes from Peru having the least TS (72.6 mg g⁻¹) but the lowest contents of raffinose (0.5 mg g⁻¹) and stachyose (11.8 mg g⁻¹) as well [161]. The highest contents for favourable glucose (23.9 mg g⁻¹) and fructose (25.2 mg g⁻¹) were determined in genotypes from Peru as well, while the highest sucrose content (83.1 mg g⁻¹) was determined in genotypes from Angola [161]. If considering only genotypes from European countries, TS contents varied between 91.5 (Bulgaria) and 111.9 mg g⁻¹ (Poland) [161]. The lowest glucose and fructose contents (0.3 mg g⁻¹) together with the highest sucrose (49.2 mg g⁻¹) and stachyose (39.2 mg g⁻¹) contents were found in a genotype from Germany, whereas the highest glucose (15.4 mg g⁻¹) and fructose (14.9 mg g⁻¹) but the lowest sucrose (16.5 mg g⁻¹), raffinose (5.2 mg g⁻¹) and stachyose (26.7 mg g⁻¹) contents were determined in three genotypes from Turkey [161]. Geater et al. [169] determined 6.2–7.1% range for sucrose, 0.49–0.58% range for raffinose and 4.7–4.9% range for stachyose in 16 small-seeded soybean cultivars. Wilcox and Shibles [84] reported combined stachyose and raffinose content in 43 random breeding lines varied in seed protein content and grown in three environments to be 38.5–42.2 g kg⁻¹ and sucrose content to be 43.3–56.7 kg⁻¹. Wide content ranges determined for soybean sugar traits [25, 161, 178] imply there is sufficient variability for altering TS or individual sugar contents. Furthermore, increasing the digestibility and taste of European soybean meal or food products necessitates the introduction of cultivars with specifically tailored sugar profiles intended for human consumption to be used as parent donors in European soybean breeding programmes, which earlier mostly have not put their focus on soluble sugars.

The introduction of germplasm with favourable sugar profiles into European breeding programmes and selection of favourable phenotypes after backcrossing can be facilitated by marker-assisted selection. The first set of markers associated with sucrose content in soybean was discovered in the F₂ population of an interspecific cross between *Glycine max* and *Glycine soja* detecting six QTL on linkage groups (LGs), A, E, F, I, L and M [171]. A novel allele of the putative soybean raffinose synthase gene (RS2) associated with the low raffinose and stachyose and elevated levels of sucrose content was discovered in PI200508 in Korea [183] which was used as a donor parent in crosses with three popular Korean soybean cultivars. The effects of RS2 allele were confirmed in the F₂ progenies through the use of allele-specific molecular markers proving that breeding and selection for low raffinose-type soybean can be done efficiently [183]. Akond et al. [177] identified and mapped 14 significant QTLs for sucrose, raffinose and stachyose contents on eight different LGs and chromosomes with only two QTLs underlying seed sucrose and stachyose contents being previously mapped. Although previous studies identified QTL associated with seed sucrose, raffinose and stachyose contents in different populations [166, 167, 171, 173, 177], only 37 QTLs for sucrose and no QTL for raffinose and stachyose are found in SoyBase [97] to date, emphasising the necessity for more researches. Nevertheless, a high-sucrose, low-raffinose and low-stachyose germplasm was developed [168] which can

be introduced to European soybean breeding programmes to be used as a parent for food-grade cultivars.

5. Isoflavones

Soybean is considered to be the most abundant natural source of isoflavones in the human and animal diet [184]. Isoflavones are the main components of flavonoids and the most common form of phytoestrogens, i.e. non-steroidal compounds with oestrogen-like biological properties [185]. They are considered nutraceuticals as it is claimed they have potential benefits in preventing the development of cardiovascular diseases and cancers, menopausal symptoms and osteoporosis, as well as antifungal and antioxidant properties [186–188]. Their role in plants is to encourage infection and nodulation by nitrogen-fixing bacteria [189] and to help in abiotic and biotic stress resistance [190]. The most studied isoflavones are genistein, daidzein and glycitein [191]. As a quantitative trait controlled by many minor genes [192], isoflavone content in soybean seed largely depends on the environment [185, 193–195], to the extent that even the smallest changes in the microclimate can cause significant changes in the isoflavone contents. For example, it is reported that lower temperatures during the seed fill period increase the isoflavone content, whereas environments which were warmer and drier resulted in lower isoflavone contents [196–199]. In 76 environments during 12 years, Carrera and Dardanelli [200] found a 91% decrease in total isoflavone content with mean temperatures during the seed development rising from 14.1 to 26.7°C. Nevertheless, Morrison et al. [201] evaluated 14 cultivars across 12 years at one location determining that high mean temperature during seed development did not result in isoflavone concentration decrease which might be due to the fact that only one location provides a somewhat narrower range of temperature variation [200]. The negative impact of water deficit on soybean isoflavone concentration was confirmed in many earlier studies [192, 198, 202, 203] as well. Furthermore, synthesis of isoflavones in soybean seed depends on the geographical origin, seed size, seed colour, maturity, disease tolerance and resistance to insects as well [185, 187, 204–206]. Nevertheless, isoflavone broad-sense heritability was estimated to range from moderate [207] to high [192], indicating that genotype effects were high enough to enable efficient improvement of isoflavone contents [194, 208]. The predominance of genotype effects over environmental was confirmed by Gutierrez-Gonzalez et al. [192], Murphy et al. [199] and Hoeck et al. [206]. Successful high-isoflavone cultivar development necessitates investigating the relationships between isoflavones and other important traits. For example, total isoflavone (TI) content was reported to be in a positive [204, 205, 209, 210] and negative [211] correlation with seed yield, but the lack of interrelationship has also been reported [199, 212], which suggests that the development of high-yield, high-isoflavone cultivars would be possible. The relationship between isoflavone content and each oil and protein content was reported to be negative [194, 199, 205, 209, 2011], but no correlation between isoflavones and proteins was reported as well [212]. Conflicting reports on the relationships between total seed isoflavone content and other traits necessitate the need for studies including a wide range of genetic material grown across a range of environments to provide additional insight into the associations [199].

Because of the many abiotic and biotic factors influencing the isoflavone content, it is expected to vary in wide ranges. Wang and Murphy [213] found TI content to vary from 1176 to 3309 $\mu\text{g/g}$ within a single cultivar of soybean. Gutierrez-Gonzales et al. [192] reported of extremely wide content range (12.4–317.4 $\text{mg } 100^{-1} \text{ g DM}$) for TI determined in the population of RILs in the USA. Murphy et al. [199] determined 160–370 $\text{mg } 100^{-1} \text{ g}$ range for TI content in RIL population tested in 2-year, multilocation trials in Canada, while Adie et al. [194] determined 14.97–39.85 $\text{mg } 100^{-1} \text{ g}$ content range for 10 soybean lines tested during 1-year trial on eight locations in Indonesia. In Europe, Cvejić et al. [208] found TI content in 20 F1 soybean progenies ranged from 156 to 366 $\text{mg } 100^{-1} \text{ g}$, while Bursać et al. [214] determined the TI content range of 211 to 524 $\text{mg } 100 \text{ g}^{-1}$ for different seed coat-coloured offsprings derived from the single cross between commercial variety and germplasm collection genotype with black seed coat. Matoša Kočar et al. [195] investigated 22 MG 00 to II soybean genotypes from the Agricultural Institute Osijek (Croatia) during 3 years (2010–2012) and determined average TI content varied from 124.06 to 286.2 $\text{mg } 100^{-1} \text{ g}$, while average daidzein, glycitein and genistein contents were 49.27 $\text{mg } 100^{-1} \text{ g}$, 20.90 $\text{mg } 100^{-1} \text{ g}$ and 94.41 $\text{mg } 100^{-1} \text{ g DM}$, respectively. Although breeding of isoflavone content at the Agricultural Institute Osijek is at its beginnings, some shifts in the right direction have been made which can be seen from the higher average isoflavone contents in elite breeding lines than in standard cultivars of the same MG (**Table 4**). In Matoša Kočar et al. [195] study, genistein was the most abundant isoflavone component, followed by daidzein and glycitein. The same order of abundance was determined in the research by Lozovaya et al. [198] for two French and three US cultivars, whereas Sumardi et al. [215] researched 34 black soybeans in Indonesia and found that daidzein content was higher in 31 genotypes. In the research done by Cvejić et al. [208], Bursać et al. [214] and Tepavčević et al. [216], total daidzein was the highest followed by total genistein and total glycitein. Gutierrez-Gonzalez et al. [192] found that glycitein was the most abundant isoflavone, followed by genistein and daidzein. Although all three isoflavones are considered to have health benefits, the order of abundance is important because genistein is reported to have approximately 10 times higher biological activity compared to daidzein and glycitein [217], so genotypes with high genistein content should be favoured in selection processes aiming to create soybean genotypes suitable for food and dietetic supplement industries.

The large influence the environmental factors have on the isoflavone content emphasises the necessity for molecular research and identification of molecular markers associated with favourable isoflavone profiles. To date, almost 90 QTLs associated with seed isoflavone content are registered in the SoyBase [97]. Akond et al. [193] identified 16 QTLs for seed isoflavones content on 12 different chromosomes (Chr) or linkage groups. Wang et al. [218] identified 33 expression QTLs on 15 soybean chromosomes, with 5 of them overlapping with phenotype QTL. In the overlapping region, 11 candidate genes underlying the accumulation of isoflavones were discovered which could be beneficial for the development of marker-assisted selection to breed soybean cultivars with high-isoflavone contents [218]. Furthermore, Wang et al. [219] identified 23 new isoflavone content QTLs and 34 QTLs in total in 130 RILs derived from the cross between high-isoflavone and low-isoflavone cultivars of which 6, 7, 10 and 11 QTLs were associated with daidzein, glycitein, genistein and TI, respectively. Akond et al. [193] identified three QTLs on three different linkage groups, one controlling daidzein

MG	Elite breeding lines		Standard cultivars		t	P
	Range	Mean	Range	Mean		
Daidzein (mg 100 g⁻¹)						
00	10.65–106.48	65.07	20.34–63.99	45.02	2.11*	P < 0.05
0	11.08–93.39	50.54	19.02–48.53	36.71	3.28**	P < 0.01
I	25.03–91.35	51.26	9.81–53.51	36.01	2.70**	P < 0.01
Glycitein (mg 100 g⁻¹)						
00	10.72–40.31	24.26	19.08–27.24	22.77	0.68 ^{ns}	P > 0.05
0	13.09–41.76	21.78	12.97–18.62	15.54	6.46**	P < 0.01
I	10.09–47.38	19.91	9.70–18.94	13.41	4.71**	P < 0.01
Genistein (mg 100 g⁻¹)						
00	43.35–355.31	184.85	57.59–128.23	93.89	4.08**	P < 0.01
0	52.32–112.25	80.49	54.82–86.38	67.44	2.86**	P < 0.01
I	48.19–142.22	92.24	54.48–104.39	83.80	1.12 ^{ns}	P > 0.05
Total isoflavones (mg 100 g⁻¹)						
00	70.49–501.79	273.18	97.22–206.06	161.68	3.50**	P < 0.01
0	87.61–212.99	150.05	115.12–148.60	136.25	0.76 ^{ns}	P > 0.05
I	91.34–239.75	160.75	71.14–179.53	129.94	2.34*	P < 0.05

The means were tested with Student's t-test with * and ** indicating significant differences at $P \leq 0.05$ and $P \leq 0.01$, respectively, and ns indicating no significant difference.

Table 4. The 3-year (2010–2012) isoflavone content range and mean for different maturity group elite breeding lines created at the agricultural institute Osijek in comparison to standard cultivars.

content and two controlling glycitein content. Furthermore, QTL epistatic interactions are also thought to contribute to isoflavone variation [205]. Although many QTL have been discovered for isoflavone content, these are mainly minor effect often influenced by the environment, so discovering sufficient loci associated with isoflavone content stable in different environments that are going to significantly aid in selection is yet to be accomplished [219].

6. Conclusion

Although genetic enhancement of soybean seed quality contributes to advances in processing industries and improves the added value properties of final soybean products, soybean as a commodity is still mostly being paid for by weight and not by composition. Considering the fact that in Europe soybean is predominantly used as animal feed, i.e. as a protein source rather than oilseed crop, it is expected of European breeders to mainly focus on increasing the seed yield and protein content, since this would be most profitable for soybean

producers. As soybean is becoming more appealing for human consumption in Europe, because of high nutritional value and health-promoting traits, breeding for improved amino acid, fatty acid and soluble sugar compositions as well as for increased isoflavone content is gaining importance but still not on a larger scale. Nevertheless, all efforts in describing the variability of important traits through research are invaluable not only for creating superior progeny but for germplasm preservation to oppose the narrowing of the genetic base. Although soybean is widely adaptable, significant environment effects for different seed quality traits, as well as the increased frequency of adverse weather events, emphasise the need for the development of local cultivars with improved performance and stability. The need for creating European soybean cultivars emanates from the demand for non-GM soybean as well, which motivates European breeders to use conventional breeding methods focusing on phenotype selection and MAS. Positive outcomes of breeding for improved seed quality recorded at the Agricultural Institute Osijek (Croatia) and other European soybean breeding programmes indicate that progress can be achieved even without genetic engineering. As this paper mainly focuses on MG 00 to II, commonly sown in Central and South-eastern Europe but suitable for growing in almost all European regions, information presented should be useful for soybean breeders and researches all over Europe and could promote the exchange of germplasm for introducing diversity, which is a prerequisite for any genetic advance.

Author details

Aleksandra Sudarić^{1,2}, Maja Matoša Kočar^{1*}, Tomislav Duvnjak¹, Zvonimir Zdunić^{1,2} and Antonela Markulj Kulundžić¹

*Address all correspondence to: maja.matoso@poljinis.hr

1 Agricultural Institute Osijek, Osijek, Republic of Croatia

2 Centre of Excellence for Biodiversity and Molecular Plant Breeding, Faculty of Agriculture, University of Zagreb, Zagreb, Republic of Croatia

References

- [1] FAOSTAT Database [Internet]. 2019. Available from: <http://faostat3.fao.org/home/E> [Accessed: July 24, 2019]
- [2] Cober ER, Cianzio SR, Pantalone VR, Rajcan I. Soybean. In: Vollmann J, Rajcan I, editors. Oil Crops. Handbook of Plant Breeding. Vol. 4. New York, USA: Springer Nature Switzerland AG; 2009. pp. 57-90
- [3] Tsuda T, Osawa T, Ohshima K, Kawakishi S. Antioxidative pigments isolated from the seeds of *Phaseolus vulgaris* L. Journal of Agricultural and Food Chemistry. 1994;**42**:248-251

- [4] Zou P. Traditional Chinese medicine, food therapy, and hypertension control: A narrative review of Chinese literature. *The American Journal of Chinese Medicine*. 2016;**44**:1-16. DOI: 10.1142/S0192415X16500889
- [5] Pascal T, Rodriguez Cerezo E. Markets for Non-Genetically Modified, Identity-Preserved Soybean in the EU. JRC Science and Policy Report. Luxembourg: Publications Office of the European Union; 2015. DOI: 10.2791/949110
- [6] Dima DC. Soybean demonstration platforms: The bond between breeding, technology and farming in central and Eastern Europe. *Agriculture and Agricultural Science Procedia*. 2016;**10**:10-17. DOI: 10.1016/j.aaspro.2016.09.003
- [7] Kurasch AK, Hahn V, Leiser WL, Vollmann J, Schori A, Béatrix CA, et al. Identification of mega-environments in Europe and effect of allelic variation at maturity *E. loci* on adaptation of European soybean. *Plant, Cell & Environment*. 2017;**40**:765-778
- [8] Jug D, Jug I, Brozović B, Vukadinović V, Stipešević B, Đurđević B. The role of conservation agriculture in mitigation and adaptation to climate change. *Poljoprivreda*. 2018;**24**(1):35-44. DOI: 10.18047/poljo.24.1.5
- [9] Li G, Ra WH, Park JW, Kwon SW, Lee JH, Park CB, et al. Developing EST-SSR markers to study molecular diversity in *Liriope* and *Ophiopogon*. *Biochemical Systematics and Ecology*. 2011;**39**:241-252. DOI: 10.1016/j.bse.2011.08.012
- [10] Vratarić M, Sudarić A. Soja. Osijek: Poljoprivredni Institut Osijek; 2008. pp. 1-459
- [11] Burton JW. Soybean (*Glycine max* (L.) Merr.). *Field Crops Research*. 1997;**53**:171-186
- [12] Gizlice Z, Carter TE Jr, Gerig TM, Burton JW. Genetic diversity patterns in North American public soybean cultivars based on coefficient of parentage. *Crop Science*. 1996;**33**:753-765. DOI: 10.1590/S1415-47572006000400019
- [13] Mudibu J, Nkongolo KKC, Kalonji-Mbuyi A. Morphovariability and agronomic characteristics of soybean accessions from the Democratic Republic of Congo (DR-Congo) gene pool. *Journal of Plant Breeding and Crop Science*. 2011;**3**:660-668
- [14] Malik MFA, Qureshi A, Ashraf M, Khan M, Javed A. Evaluation of genetic diversity in soybean (*Glycine max*) lines using seed protein electrophoresis. *Australian Journal of Crop Science*. 2009;**3**:107-112
- [15] Salimi S. Relationships of some soybean genotypes based on morphological and biochemical markers. *International Agronomy and Plant Production*. 2013;**4**:2237-2243
- [16] Mulato BM, Möller M, Zucchi MI, Quecini V, Pinheiro JB. Genetic diversity in soybean germplasm identified by SSR and EST-SSR markers. *Pesquisa Agropecuária Brasileira*. 2010;**45**:276-283
- [17] Tantasawat P, Trongchuen J, Prajongjai T, Jenweerawat S, Chaowiset W. SSR analysis of soybean (*Glycine max* (L.) Merr.) genetic relationship and variety identification in Thailand. *Australian Journal of Crop Science*. 2011;**5**:283-290

- [18] Moose SP, Mumm RH. Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiology*. 2008;**147**:969-977. DOI: 10.1104/pp.108.118232
- [19] Harlan JR. Genetics of disaster. *Journal of Environmental Quality*. 1972;**1**:212-215
- [20] Zhang G, Xu S, Mao W, Hu Q, Gong Y. Determination of the genetic diversity of vegetable soybean (*Glycine max* (L.) Merr.) using EST-SSR markers. *Journal of Zhejiang University – Science B (Biomedicine & Biotechnology)*. 2013;**14**:279-288. DOI: 10.1631/jzus.B1200243
- [21] Perić V, Nikolić A, Babić V, Sudarić A, Srebrić M, Đorđević V, et al. Genetic relatedness of soybean genotypes. *Genetika-Belgrade*. 2014;**46**:839-854. DOI: 10.2298/GENSR1403839P
- [22] Hahn V, Würschum T. Molecular genetic characterization of Central European soybean breeding germplasm. *Plant Breeding*. 2014;**133**:748-755. DOI: 10.1111/pbr.12212
- [23] Grainger CM, Rajcan I. Characterization of the genetic changes in a multi-generational pedigree of an elite Canadian soybean cultivar. *Theoretical and Applied Genetics*. 2014;**127**:211-229. DOI: 10.1007/s00122-013-2211-9
- [24] Varshney RK, Thudi M, Pandey MK, Tardieu F, Ojiewo C, Vadez V, et al. Accelerating genetic gains in legumes for the development of prosperous smallholder agriculture: Integrating genomics, phenotyping, systems modelling and agronomy. *Journal of Experimental Botany*. 2018;**69**:3293-3312. DOI: 10.1093/jxb/ery088
- [25] Hymowitz T, Collins FI. Variability of sugar content of seed of *Glycine max* (L.) Merr. and *G. soja* Serb. and Zucco. *Agronomy Journal*. 1974;**66**:239-240. DOI: 10.2134/agronj1974.00021962006600020017x
- [26] Cromwell GL. Soybean Meal: An Exceptional Protein Source. Ankeny, IA: Soybean Meal InfoCenter [Internet]. 2012. Available from: https://www.soymeal.org/wpcontent/uploads/2018/04/soybean_meal_an_exceptional_protein_source.pdf [Accessed: July 25, 2019]
- [27] Hasler CM. The cardiovascular effects of soy products. *Journal of Cardiovascular Nursing*. 2002;**16**:50-63
- [28] Zarkadas CG, Voldeng HD, Yu ZR, Choi V. Assessment of the protein quality of nine northern adapted yellow and brown seed coated soybean cultivars by amino acid analysis. *Journal of Agricultural and Food Chemistry*. 1999;**47**:5009-5018. DOI: 10.1021/jf981381r
- [29] Everson G, Heckert A. The biological value of some leguminous sources of protein. *Journal of the American Dietetic Association*. 1944;**20**:81-82
- [30] Wagner JR, Sorgentini DA, Anon MC. Thermal and electrophoretic behavior, hydrophobicity, and some functional properties of acid treated soy isolates. *Journal of Agricultural and Food Chemistry*. 1996;**44**:1881-1889. DOI: 10.1021/jf950444s
- [31] Cai T, Chang KC. Processing effect on soybean storage proteins and their relationship with tofu quality. *Journal of Agricultural and Food Chemistry*. 1999;**47**:720-727. DOI: 10.1021/jf980571z

- [32] Žilić SM, Barać MB, Pešić MB, Mladenović Drinić SD, Ignjatović-Micić DD, Srebrić MB. Characterization of proteins from kernel of different soybean varieties. *Journal of the Science of Food and Agriculture*. 2010;**91**:60-67. DOI: 10.1002/jsfa.4148
- [33] Kito M, Moriyama T, Kimura Y, Kambara H. Changes in plasma lipid levels in young healthy volunteers by adding an extruder-cooked soy protein to conventional meals. *Bioscience, Biotechnology, and Biochemistry*. 1993;**57**:354-355. DOI: 10.1271/bbb.57.354
- [34] Poysa V, Woodrow L, Yu K. Effect of soy protein subunit composition on tofu quality. *Food Research International*. 2006;**39**:309-317. DOI: 10.1016/j.foodres.2005.08.003
- [35] Liener IE. Implications of antinutritional components in soybean foods. *Critical Reviews in Food Science and Nutrition*. 1994;**34**:31-67. DOI: 10.1080/10408399409527649
- [36] Yin Y, Fatufe AA, Blachier F. Soya bean meal and its extensive use in livestock feeding and nutrition, soybean and nutrition. In: El-Shemy H, editor. *Soybean and Nutrition*. Rijeka, Croatia: InTech; 2011. pp. 369-384. DOI: 10.5772/18823
- [37] Yasothai R. Antinutritional factors in soybean meal and its deactivation. *International Journal of Science, Environment and Technology*. 2016;**5**:3793-3797
- [38] SoyStats [Internet]. 2019. Available from: <http://soystats.com/international-world-protein-meal-consumption/> [Accessed: July 30, 2019]
- [39] Gatrell S, Lum K, Kim J, Lei XG. Nonruminant nutrition symposium: Potential of defatted microalgae from the biofuel industry as an ingredient to replace corn and soybean meal in swine and poultry diets. *Journal of Animal Science*. 2014;**92**:1306-1314. DOI: 10.2527/jas.2013-7250
- [40] Jobbágy EG, Sala OE. The imprint of crop choice on global nutrient needs. *Environmental Research Letters*. 2014;**9**:084014. DOI: 10.1088/1748-9326/9/8/084014
- [41] Bues A, Preißel S, Reckling M, Kuhlman T, Topp K, Watson C, et al. The Environmental Role of Protein Crops in the New Common Agricultural Policy [Internet]. 2013. Available from: [http://www.europarl.europa.eu/RegData/etudes/etudes/JOIN/2013/495856/IPOL-AGRI_ET\(2013\)495856_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/etudes/JOIN/2013/495856/IPOL-AGRI_ET(2013)495856_EN.pdf) [Accessed: July 14, 2019]
- [42] Alexandratos N, Bruinsma J. World agriculture towards 2030/2050: The 2012 revision. *ESA Working Paper No. 12-03*. Rome, FAO [Internet]. 2012. Available from: <http://www.fao.org/3/a-ap106e.pdf> [Accessed: July 20, 2019]
- [43] Mottet A, de Haan C, Falcucci A, Tempio G, Opio C, Gerber P. Livestock: On our plates or eating at our table? A new analysis of the feed/food debate. *Global Food Security*. 2017;**14**:1-8
- [44] Watson CA, Reckling M, Preissel S, Bachinger J, Bergkvist G, Kuhlman T, et al. Grain legume production and use in European agricultural systems. *Advances in Agronomy*. 2017;**144**:235-303. DOI: 10.1016/bs.agron.2017.03.003
- [45] Carbon Trust. Carbon-Footprinting Software: Footprint Expert. London: Carbon Trust [Internet]. 2013. Available from: <http://www.carbontrust.com/software> [Accessed: July 16, 2019]

- [46] Chéreau D, Videcoq P, Ruffieux C, Pichon L, Motte JC, Belaid S, et al. Combination of existing and alternative technologies to promote oilseeds and pulses proteins in food applications. *Oilseeds and Fats, Crops and Lipids*. 2016;**23**:D406. DOI: 10.1051/ocl/2016020
- [47] Flachowsky G, Meyer U, Südekum KH. Land use for edible protein of animal origin— A review. *Animals*. 2017;**7**:25. DOI: 10.3390/ani7030025.
- [48] Zander P, Amjath-Babu TS, Preissel S, Reckling M, Bues A, Schläfke N, et al. Grain legume decline and potential recovery in European agriculture: A review. *Agronomy for Sustainable Development*. 2016;**36**:26. DOI: 10.1007/s13593-016-0365-y
- [49] Roman GV, Epure LI, Toader M, Lombardi AR. Grain legumes— Main source of vegetal proteins for European consumption. *AgroLife Scientific Journal*. 2016;**1**:178-183
- [50] Chetana C, Rusub T, Chetana F, Simona A. Influence of soil tillage systems and weed control treatments on root nodules, production and qualitative indicators of soybean. *Procedia Technology*. 2016;**22**:457-464. DOI: /10.1016/j.protcy.2016.01.088
- [51] Latifi N. Yield and morphological response of soybean to time of irrigation and sowing rate. *Dissertation Abstracts International*. 1989;**40**:5088-5098
- [52] Kirnak H, Dogan E, Turkoglu H. Effect of drip irrigation intensity on soybean seed yield and quality in the semi-arid Harran plain, Turkey. *Spanish Journal of Agricultural Research*. 2010;**8**:1208-1217. DOI: 10.5424/sjar/2010084-1239
- [53] Ghassemi-Golezani K, Farshbaf-Jafari S. Influence of water deficit on oil and protein accumulation in soybean grains. *International Journal of Plant, Animal and Environmental Sciences*. 2012;**2**:46-52
- [54] Foroud N, MuÈndel HH, Saindon G, Entz T. Effect of level and timing of moisture stress on soybean yield, protein, and oil responses. *Field Crops Research*. 1993;**31**:195-209. DOI: 10.1016/0378-4290(93)90062-R
- [55] Bouniols A, Texier V, Mondies M, Piva G. Soybean seed quality among genotypes and crop management: Field experiment and simulation. *Eurosoya*. 1997;**11**:87-99
- [56] Afza R, Hardarson G, Zapata F, Danso SKA. Effects of delayed soil and foliar N fertilisation on yield and N₂ fixation of soybean. *Plant and Soil*. 1987;**97**:361-368
- [57] Isfan D. Fertilizer nitrogen uptake by soybean as related to cultivars and time of application using ¹⁵N technique. *Journal of Plant Nutrition*. 1991;**14**:1369-1380. DOI: 10.1080/01904169109364292
- [58] Paek NC, Imsande J, Shoemaker RC, Shibles R. Nutritional control of soybean seed storage protein. *Crop Science*. 1997;**37**:498-503. DOI: 10.2135/cropsci1997.0011183X003700020031x
- [59] Grabau LJ, Blevins DG, Minor HC. Stem infusions enhanced methionine content of soybean storage protein. *Plant Physiology*. 1986;**82**:1013-1018. DOI: 10.1104/pp.82.4.1013

- [60] Zimmer S, Messmer M, Haase T, Piepho HP, Mindermann A, Schulz H, et al. Effects of soybean variety and Bradyrhizobium strains on yield, protein content and biological nitrogen fixation under cool growing conditions in Germany. *European Journal of Agronomy*. 2016;**72**:38-46. DOI: 10.1016/j.eja.2015.09.008
- [61] Vollmann J, Fritz CN, Wagentristsl H, Ruckebauer P. Environmental and genetic variation of soybean seed protein content under central European growing conditions. *Journal of the Science of Food and Agriculture*. 2000;**80**:1300-1306. DOI: 10.1002/1097-0010(200007)80:9<1300:AID-JSFA640>3.3.CO;2-9
- [62] Kim SW, Less JF, Wang L, Yan T, Kiron V, Kaushik SJ, et al. Meeting global feed protein demand: Challenge, opportunity, and strategy. *Annual Review of Animal Biosciences*. 2019;**7**:221-243. DOI: 10.1146/annurev-animal-030117-014838
- [63] Sudarić A, Šimić D, Vratarić M. Characterization of genotype by environment interactions in soybean breeding programmes of Southeast Europe. *Plant Breeding*. 2006;**125**: 191-194. DOI: 10.1111/j.1439-0523.2006.01185.x
- [64] Bueno RD, Borges LL, Arruda KMA, Bhering LL, de Barros EG, Moreira MA. Genetic parameters and genotype x environment interaction for productivity, oil and protein content in soybean. *African Journal of Agricultural Research*. 2013;**8**:4853-4859. DOI: 10.5897/AJAR2013.6924
- [65] Ghodrati G. Study of genetic variation and broad sense heritability for some qualitative and quantitative traits in soybean (*Glycine max* L.) genotype. *Current Opinion in Agriculture*. 2013;**2**:31-35
- [66] Josipović M, Sudarić A, Sudar R, Plavšić H, Marković M, Jug D, Stojić B. Influence of irrigation and variety on the soybean grain yield and quality in the no nitrogen fertilisation soil condition. In: *Soil and Crop Management: Adaptation and Mitigation of Climate Change*; 26-28 September 2013; Osijek. Osijek: Grafika d.o.o; 2013. pp. 237-245
- [67] Rodrigues JID, Arruda KMA, Cruz CD, Piovesan ND, de Barros EG, Moreira MA. Biometric analysis of protein and oil contents of soybean genotypes in different environments. *Pesquisa Agropecuária Brasileira*. 2014;**49**:475-482. DOI: 10.1590/S0100-204X2014000600009
- [68] Gravois KA, Bernhardt JL. Heritability x environment interactions for discoloured rice kernels. *Crop Science*. 2000;**40**:314-318. DOI: 10.2135/cropsci2000.402314x
- [69] Malik MFA, Ashraf M, Qureshi AS, Ghafoor A. Utilisation of diverse germplasm for soybean yield improvement. *Asian Journal of Plant Sciences*. 2006;**5**:663-667. DOI: 10.3923/ajps.2006.663.667
- [70] Kurasch AK, Hahn V, Leiser WL, Starck N, Würschum T. Phenotypic analysis of major agronomic traits in 1008 RILs from a diallel of early European soybean varieties. *Crop Science*. 2017;**57**:726-738. DOI: 10.2135/cropsci2016.05.0318
- [71] Popović V, Vidić M, Tatić M, Jakšić S, Kostić M. Uticaj sorte i godine na prinos i komponente kvaliteta soje. *Ratarstvo i Povrtarstvo*. 2012;**49**:132-139. DOI: 10.5937/ratpov49-1140

- [72] Hurburgh CR Jr. Long-term soybean composition patterns and their effect on processing. *Journal of the American Oil Chemists' Society*. 1994;**71**:1425-1427
- [73] Nian H, Wang JL, Yang QK. Effects of ecological conditions of the northeast of China on protein, oil and protein+oil content in soybean seeds. *Journal of Northeast Agricultural University (English Edition)*. 1996;**3**:1-6
- [74] Zhang F, Lynch DH, Smith DL. Impact of low root temperatures in soybean [*Glycine max* (L.) Merr.] on nodulation and nitrogen fixation. *Environmental and Experimental Botany*. 1995;**35**:279-285. DOI: 10.1016/0098-8472(95)00017-7
- [75] Gibson LR, Mullen RE. Soybean seed composition under high day and night growth temperatures. *Journal of the American Oil Chemists' Society*. 1996;**73**:733-737
- [76] Dornbos DL Jr, Mullen RE. Soybean seed protein and oil contents and fatty acid composition adjustments by drought and temperature. *Journal of the American Oil Chemists' Society*. 1992;**69**:228-231
- [77] Matoša Kočar M, Sudarić A, Vila S, Petrović S, Rebečić A, Josipović A, et al. Varijabilnost fenotipske ekspresije svojstava kvalitete zrna elitnih linija soje. *Poljoprivreda*. 2017;**23**:40-48. DOI: 10.18047/poljo.23.1.7
- [78] Brim CA, Burton JW. Recurrent selection in soybeans. II. Selection for increased percent protein in seeds. *Crop Science*. 1979;**19**:494-498
- [79] Leffel RC, Rhodes WK. Agronomic performance and economic value of high-seed protein soybean. *Journal of Production Agriculture*. 1993;**6**:365-368
- [80] Cober ER, Voldeng HD. Developing high-protein, high-yield soybean populations and lines. *Crop Science*. 2000;**40**:39-42. DOI: 10.2135/cropsci2000.40139x
- [81] Pannecoucq J, Goormachtigh S, Heungens K, Vleugels T, Ceusters J, Van Waes C, et al. Screening for soybean varieties suited to Belgian growing conditions based on maturity, yield components and resistance to *Sclerotinia sclerotiorum* and *Rhizoctonia solani* anastomosis group 2-IIIB. *The Journal of Agricultural Science*. 2018;**156**:342-349. DOI: 10.1017/S0021859618000333
- [82] Chung J, Babka HL, Graef GL, Staswick PE, Lee DJ, Cregan PB, et al. The seed protein, oil, and yield QTL on soybean linkage group I. *Crop Science*. 2003;**43**:1053-1067. DOI: 10.2135/cropsci2003.1053
- [83] Clemente TE, Cahoon EB. Soybean oil: Genetic approaches for modification of functionality and total content. *Plant Physiology*. 2009;**151**:1030-1040. DOI: 10.1104/pp.109.146282
- [84] Wilcox JR, Shibles RM. Interrelationships among seed quality attributes in soybean. *Crop Science*. 2001;**41**:11-14. DOI: 10.2135/cropsci2001.41111x
- [85] Burton J, Purcell A, Walter W. Methionine concentration in soybean protein from populations selected for increased percent protein. *Crop Science*. 1982;**22**:430-432. DOI: 10.2135/cropsci1982.0011183X002200020054x

- [86] Krober OA, Cartter JL. Relation of methionine to protein levels in soybeans. *Cereal Chemistry*. 1966;**43**:320-325
- [87] Fogelberg F, Recknagel J. Developing soy production in central and northern Europe. In: Murphy-Bokern D, Stoddard FL, Watson CA, editors. *Legumes in Cropping Systems*. CAB International; 2017. pp. 109-124. DOI: 10.1079/9781780644981.0109
- [88] Sudarić A, Vratarić M. Variability and interrelationships of grain quantity and quality characteristics in soybean. *Die Bodenkultur*. 2002;**53**:137-142
- [89] Vratarić M, Sudarić A, Sudar R, Duvnjak T, Jurković D, Jurković Z. Genetic advance in quantitative traits of soybean lines within different maturity groups. *Poljoprivreda*. 2005;**11**:5-10
- [90] Sato T, van Schoote M, Wagentristsl H, Vollmann J. Effects of divergent selection for seed protein content in high-protein vs. food-grade populations of early maturity soybean. *Plant Breeding*. 2014;**133**:74-79. DOI: 10.1111/pbr.12138
- [91] Sharma S, Kaur M, Goyal R, Gill BS. Physical characteristics and nutritional composition of some new soybean (*Glycine max* (L.) Merrill) genotypes. *Journal of Food Science and Technology*. 2014;**51**:551-557. DOI: 10.1007/s13197-011-0517-7
- [92] Ramteke R, Kumar V, Muralidharan P, Agarwal DK. Study on genetic variability and traits interrelationship among released soybean varieties of India (*Glycine max* (L.) Merrill). *Electronic Journal of Plant Breeding*. 2010;**1**:1483-1487
- [93] Wilson RF. Seed composition. In: Boerma HR, Specht JE, editors. *Soybeans: Improvement, Production and Uses*, Ed 3. Madison, Wisconsin, USA: American Society of Agronomy, Crop Science Society of America, Soil Science Society of America; 2004. pp. 621-677
- [94] Thorne JC, Fehr WR. Incorporation of high-protein, exotic germplasm into soybean populations by 2- and 3-way crosses. *Crop Science*. 1970;**10**:652-655
- [95] Wehrmann VK, Fehr WR, Cianzio SR, Cavins JF. Transfer of high seed protein to high-yielding soybean cultivars. *Crop Science*. 1987;**27**:927-931. DOI: 10.2135/cropsci1987.0011183X002700050020x
- [96] Wilcox JR, Cavins JF. Backcrossing high seed protein to a soybean cultivar. *Crop Science*. 1995;**35**:1036-1041. DOI: 10.2135/cropsci1995.0011183X003500040019x
- [97] SoyBase [Internet]. 2019. Available from: <https://www.soybase.org/> [Accessed: July 23, 2019]
- [98] Brummer EC, Graef GL, Orf J, Wilcox JR, Shoemaker R. Mapping QTLs for seed protein and oil content in eight soy-bean populations. *Crop Science*. 1997;**37**:370-378
- [99] Sebolt AM, Shoemaker RC, Diers BW. Analysis of a quantitative trait locus allele from wild soybean that increases seed protein concentration in soybean. *Crop Science*. 2000;**40**:1438-1444. DOI: 10.2135/cropsci2000.4051438x

- [100] Hwang EY, Song QJ, Jia GF, Specht JE, Hyten DL, Costa J, et al. A genome-wide association study of seed protein and oil content in soybean. *BMC Genomics*. 2014;**15**:1-12. DOI: 10.1186/1471-2164-15-1
- [101] Nichols DM, Glover KD, Carlson SR, Specht JE, Diers BW. Fine mapping of a seed protein QTL on soybean linkage group I and its correlated effects on agronomic traits. *Crop Science*. 2006;**46**:834-839. DOI: 10.2135/cropsci2005.05-0168
- [102] Bandillo N, Jarquin D, Song Q, Nelson R, Cregan P, Specht J, et al. A population structure and genome-wide association analysis on the USDA soybean germplasm collection. *Plant Genome*. 2015;**8**:1-13. DOI: 10.3835/plantgenome2015.04.0024
- [103] Wang J, Zhou P, Shi X, Yang N, Yan L, Zhao Q, et al. Primary metabolite contents are correlated with seed protein and oil traits in near-isogenic lines of soybean. *The Crop Journal*. DOI: 10.1016/j.cj.2019.04.002
- [104] Durham D. The United Soybean Board's better bean initiative: Building United States soybean competitiveness from the inside out. *The Journal of Agrobiotechnology Management & Economics*. 2003;**6**:23-26
- [105] Clarke E, Wiseman J. Developments in plant breeding for improved nutritional quality of soya beans. *The Journal of Agricultural Science*. 2000;**134**:111-124. DOI: 10.1017/S0021859699007431
- [106] Fallen BD, Hatcher CN, Allen FL, Kopsell DA, Saxton AM, Chen P, et al. Soybean seed amino acid content QTL detected using the universal soy linkage panel 1.0 with 1,536 SNPs. *Journal of Plant Genome Sciences*. 2013;**1**:68-79. DOI: 10.5147/jpgs.2013.0089
- [107] Wang XZ, Jiang GL, Song QJ, Cregan PB, Scott RA, Zhang JP, et al. Quantitative trait locus analysis of seed sulfur-containing amino acids in two recombinant inbred line populations of soybean. *Euphytica*. 2015;**201**:293-305. DOI: 10.1007/s10681-014-1223-0
- [108] Vaughn JN, Nelson RL, Song Q, Cregan PB, Li Z, Vaughn JN, et al. The genetic architecture of seed composition in soybean is refined by genome-wide association scans across multiple populations. *G3 (Bethesda)*. 2014;**4**:2283-2294. DOI: 10.1534/g3.114.013433
- [109] Qiu H, Hao W, Gao S, Ma X, Zheng Y, Meng F, et al. Gene mining of sulfur-containing amino acid metabolic enzymes in soybean. *Yi Chuan = Hereditas*. 2014;**36**:934-942. DOI: 10.3724/SP.J.1005.2014.0934
- [110] Panthee D, Kwanyuen P, Sams C, West D, Saxton A, Pantalone V. Quantitative trait loci for β -conglycinin (7S) and glycinin (11S) fractions of soybean storage protein. *Journal of the American Oil Chemists' Society*. 2004;**81**:1005-1012
- [111] Ma Y, Kan G, Zhang X, Wang Y, Zhang W, Du H, et al. Quantitative trait loci (QTL) mapping for glycinin and β -conglycinin contents in soybean (*Glycine max* L. Merr.). *Journal of Agricultural and Food Chemistry*. 2016;**64**:3473-3483. DOI: 10.1021/acs.jafc.6b00167
- [112] Boehm JD Jr, Nguyen V, Tashiro RM, Anderson D, Shi C, Wu X, et al. Genetic mapping and validation of the loci controlling 7S a' and 11S A-type storage protein subunits in

- soybean [*Glycine max* (L.) Merr.]. *Theoretical and Applied Genetics*. 2018;**131**:659-671. DOI: 10.1007/s00122-017-3027-9
- [113] Murphy PA, Resurreccion AP. Varietal and environmental differences in soybean glycinin and β -conglycinin content. *Journal of Agricultural and Food Chemistry*. 1984;**32**:911-915
- [114] Mujoo R, Trinh DT, Perry KW. Characterization of storage proteins in different soybean varieties and their relationship to tofu yield and texture. *Food Chemistry*. 2003;**82**:265-273. DOI: 10.1016/S0308-8146(02)00547-2
- [115] Zhang D, Lü H, Chu S, Zhang H, Zhang H, Yang Y, et al. The genetic architecture of water-soluble protein content and its genetic relationship to total protein content in soybean. *Scientific Reports*. 2017;**7**:5053. DOI: 10.1038/s41598-017-04685-7
- [116] Krishnana HB, Jezb JM. Review: The promise and limits for enhancing sulfur-containing amino acid content of soybean seed. *Plant Science*. 2018;**272**:14-21. DOI: 10.1016/j.plantsci.2018.03.030
- [117] SoyStats [Internet]. 2018. Available from: <http://soystats.com/> [Accessed: July 26, 2019]
- [118] Snyder CL, Yurchenko OP, Siloto RM, Chen X, Liu Q, Mietkiewska E, et al. Acyltransferase action in the modification of seed oil biosynthesis. *New Biotechnology*. 2009;**26**:11-16. DOI: 10.1016/j.nbt.2009.05.005
- [119] Fehr WR. Breeding for modified fatty acid composition in soybean. *Crop Science*. 2007;**47**:S-72-S-87. DOI: 10.2135/cropsci2007.04.0004IPBS
- [120] Baum SJ, Kris-Etherton PM, Willett WC, Lichtenstein AH, Rudel LL, Maki KC, et al. Fatty acids in cardiovascular health and disease: A comprehensive update. *Journal of Clinical Lipidology*. 2012;**6**:216-234. DOI: 10.1016/j.jacl.2012.04.077
- [121] Crupkin M, Zambelli A. Detrimental impact of trans fats on human health: Stearic acid-rich fats as possible substitutes. *Comprehensive Reviews in Food Science and Food Safety*. 2008;**7**:271-279. DOI: 10.1111/j.1541-4337.2008.00045.x
- [122] Jeong JE, Krishnanand KP, Chang JH, Ha BK, Kang ST, Bilyeu K, et al. A novel allele of GmSACPD-C associated with high seed stearic acid concentration in an EMS-induced mutant PE980 in soybean. *Crop Science*. 2018;**58**:192-203. DOI: 10.2135/cropsci2017.05.0313
- [123] Johnson S, Saikia N. *Fatty Acids Profile of Edible Oils and Fats in India*. New Delhi, India: Centre for Science and Environment; 2009
- [124] Lee GJ, Wu X, Shannon JG, Sleper DA, Nguyen HT. Soybean. In: Kole C, editor. *Genome Mapping and Molecular Breeding in Plants*. Berlin: Springer; 2007. pp. 1-3
- [125] Duh PD, Yen WJ, Yen GC. Oxidative stability of polyunsaturated fatty acids and soybean oil in an aqueous solution with emulsifiers. *Journal of the American Oil Chemists' Society*. 1999;**76**:201-204

- [126] Rani A, Kumar V, Verma SK, Shakya AK, Hussain SM, Chauhan GS. Interrelationship between oil content and fatty acid composition in Indian soybean (*Glycine max*) cultivars. *The Indian Journal of Agricultural Sciences*. 2007;**77**:59-62
- [127] Matoša Kočar M, Sudarić A, Sudar R, Duvnjak T, Zdunić Z. Screening of early maturing soybean genotypes for production of high quality edible oil. *Zemdirbyste-Agriculture*. 2018;**105**:55-62. DOI: 10.13080/z-a.2018.105.008
- [128] Brouwer IA, Wanders AJ, Katan MB. Effect of animal and industrial trans fatty acids on HDL and LDL cholesterol levels in humans—A quantitative review. *PLoS One*. 2010;**5**:e9434. DOI: 10.1371/journal.pone.0009434
- [129] Combs R, Bilyeu K. Novel alleles of FAD2-1A induce high levels of oleic acid in soybean oil. *Molecular Breeding*. 2019;**39**:79. DOI: 10.1007/s11032-019-0972-9
- [130] Williams CD, Whitley BM, Hoyo C, Grant DJ, Irraggi JD, Newman KA, et al. Freedland SJ. A high ratio of dietary n-6/n-3 polyunsaturated fatty acids is associated with increased risk of prostate cancer. *Nutrition Research*. 2011;**31**:1-8. DOI: 10.1016/j.nutres.2011.01.002
- [131] Chapter 4: Oilseeds and Oilseed Products. *Agricultural Outlook 2018-2027*. OECD-FAO [Internet]. 2018. pp. 127-274. Available from: <http://www.agri-outlook.org/commodities/Agricultural-Outlook-2018-Oilseeds.pdf> [Accessed: July 29, 2019]
- [132] Hemingway J, Eskandari M, Rajcan I. Genetic and environmental effects on fatty acid composition in soybeans with potential use in automotive industry. *Crop Science*. 2015;**55**:1-11. DOI: 10.2135/cropsci2014.06.0425
- [133] Zhang D, Zhang H, Hu Z, Chu S, Yu K, Lv L, et al. Artificial selection on GmOLEO1 contributes to the increase in seed oil during soybean domestication. *PLOS Genetics: A Peer-Reviewed Open-Access Journal*. 2019;**15**:7. DOI: 10.1371/journal.pgen.1008267
- [134] Panthee D, Pantalone V, West D, Saxton A, Sams C. Quantitative trait loci for seed protein and oil concentration, and seed size in soybean. *Crop Science*. 2005;**45**:2015-2022. DOI: 10.2135/cropsci2004.0720
- [135] Mao T, Jiang Z, Han Y, Teng W, Zhao X, Li W. Identification of quantitative trait loci underlying seed protein and oil contents of soybean across multi-genetic backgrounds and environments. *Plant Breeding*. 2013;**132**:630-641. DOI: 10.1111/pbr.12091
- [136] Lee S, Van K, Sung M, Nelson R, La Mantia J, McHale LK, et al. Genome-wide association study of seed protein, oil and amino acid contents in soybean from maturity groups I to IV. *Theoretical and Applied Genetics*. 2019;**132**(6):1639-1659. DOI: 10.1007/s00122-019-03304-5
- [137] Ren C, Bilyeu KD, Beuselinck P. Composition, vigor, and proteome of mature soybean seeds developed under high temperature. *Crop Science*. 2009;**49**:1010-1022. DOI: 10.2135/cropsci2008.05.0247
- [138] Mourtzinis S, Gaspar AP, Naeve SL, Conley SP. Planting date, maturity, and temperature effects on soybean seed yield and composition. *Agronomy Journal*. 2017;**109**:2040-2049. DOI: 10.2134/agronj2017.05.0247

- [139] Kumar V, Rani A, Solanki S, Hussain SM. Influence of growing environment on the biochemical composition and physical characteristics of soybean seed. *Journal of Food Composition and Analysis*. 2006;**19**:188-195. DOI: 10.1016/j.jfca.2005.06.005
- [140] Pipolo AE, Sinclair TR, Camara GMS. Effects of temperature on oil and protein concentration in soybean seed cultured in vitro. *Annals of Applied Biology*. 2004;**144**:71-76. DOI: 10.1111/j.1744-7348.2004.tb00318.x
- [141] Rotundo JL, Westgate ME. Meta-analysis of environmental effects on soybean seed composition. *Field Crops Research*. 2009;**110**:147-156. DOI: 10.1016/j.fcr.2008.07.012
- [142] Assefa Y, Purcell LC, Salmeron M, Naeve S, Casteel SN, Kovács P, et al. Assessing variation in US soybean seed composition (protein and oil). *Frontiers in Plant Science*. 2019;**10**:298. DOI: 10.3389/fpls.2019.00298
- [143] Wood CW, Torbert HA, Weaver DB. Nitrogen fertilizer effects on soybean growth, yield, and seed composition. *Journal of Production Agriculture*. 1993;**6**:354-360. DOI: 10.2134/jpa1993.0354
- [144] Bellaloui N, Bruns HA, Gillen AM, Abbas HK, Zablutowicz RM, Mengistu A, et al. Soybean seed protein, oil, fatty acids, and mineral composition as influenced by soybean-corn rotation. *Agricultural Sciences*. 2010;**1**:102-109. DOI: 10.4236/as.2010.13013
- [145] Robinson AP, Conley SP, Volenec JJ, Santini JB. Analysis of high yielding, early-planted soybean in Indiana. *Agronomy Journal*. 2009;**101**:131-139. DOI: 10.2134/agronj2008.0014x
- [146] Bellaloui N, Bruns HA, Abbas HK, Mengistu A, Fisher DK, Reddy KN. Agricultural practices altered soybean seed protein, oil, fatty acids, sugars, and minerals in the Midsouth USA. *Frontiers in Plant Science*. 2015;**6**:31. DOI: 10.3389/fpls.2015.00031
- [147] Jaureguy LM, Rodriguez FL, Zhang L, Chen P, Brye K, Oosterhuis D, et al. Planting date and delayed harvest effects on soybean seed composition. *Crop Science*. 2013;**53**:2162-2175. DOI: 10.2135/cropsci2012.12.0683
- [148] Fan S, Li B, Yu F, Han F, Yan S, Wang L, et al. Analysis of additive and epistatic quantitative trait loci underlying fatty acid concentrations in soybean seeds across multiple environments. *Euphytica*. 2015;**206**:689-700
- [149] Xue HQ, Upchurch RG, Kwanyuen P. Relationships between oleic and linoleic acid content and seed colonization by *Cercospora kikuchii* and *Diaporthe phaseolorum*. *Plant Disease*. 2008;**92**:1038-1042. DOI: 10.1094/PDIS-92-7-1038
- [150] Kumar V, Rani A, Joshi OP. Fatty acid profile released cultivars of Indian soybean with special reference to identification of high comparatively low linolenic and high oleic acid cultivars. *The Indian Journal of Agricultural Sciences*. 2004;**74**:388-391
- [151] Fehr WR, Curtiss CF. Breeding for fatty acid composition of soybean oil. In: 7th World Soybean Research Conference, 4th International Soybean Processing and Utilization Conference, 3rd Brazilian Soybean Congress. 29-5 March; Foz Do Iguacu, Brazil; 2004. pp. 815-821

- [152] Priolli RHG, Carvalho CRL, Bajay MM, Pinheiro JB, Vello NA. Genome analysis to identify SNPs associated with oil content and fatty acid components in soybean. *Euphytica*. 2019;**215**:54. DOI: 10.1007/s10681-019-2378-5
- [153] Xia C. Study on oil content and fatty acid composition of soybean cultivars cultivated in Heliongjiang province in China. *Soybean Genetics Newsletter*. 1995;**22**:99-102
- [154] Thapa R, Carrero-Colón M, Addo-Quaye C, Held J, Dilkes B, Hudson KA. New alleles of FAD3A lower the linolenic acid content of soybean seeds. *Crop Science*. 2018;**58**:1-6. DOI: 10.2135/cropsci2017.08.0490
- [155] Zhang J, Wang X, Lu Y, Bhusal SJ, Song Q, Cregan PB, et al. Genome-wide scan for seed composition provides insights into soybean quality improvement and the impacts of domestication and breeding. *Molecular Plant - Cell Press*. 2018;**11**:460-472. DOI: 10.1016/j.molp.2017.12.016
- [156] Cao Y, Li S, Wang Z, Chang F, Kong J, Gai J, et al. Identification of major quantitative trait loci for seed oil content in soybeans by combining linkage and genome-wide association mapping. *Frontiers in Plant Science*. 2017;**8**:1222. DOI: 10.3389/fpls.2017.01222
- [157] Van K, McHale LK. Meta-analyses of QTLs associated with protein and oil contents and compositions in soybean [*Glycine max* (L.) Merr.] seed. *International Journal of Molecular Sciences*. 2017;**18**:pii: E1180. DOI: 10.3390/ijms18061180
- [158] Li YH, Reif JC, Hong HL, Li HH, Liu ZX, Ma YS, et al. Genome-wide association mapping of QTL underlying seed oil and protein contents of a diverse panel of soybean accessions. *Plant Science: An International Journal of Experimental Plant Biology*. 2018;**266**:95-101. DOI: 10.1016/j.plantsci.2017.04.013
- [159] Roesler K, Shen B, Bermudez E, Li CJ, Hunt J, Damude HG, et al. An improved variant of soybean type 1 diacylglycerol acyltransferase increases the oil content and decreases the soluble carbohydrate content of soybeans. *Plant Physiology*. 2016;**171**:878-893. DOI: 10.1104/pp.16.00315
- [160] Li QT, Lu X, Song QX, Chen HW, Wei W, Tao JJ, et al. Selection for a zinc-finger protein contributes to seed oil increase during soybean domestication. *Plant Physiology*. 2017;**173**:2208-2224. DOI: 10.1104/pp.16.01610
- [161] Hou A, Chen P, Alloatti J, Li D, Mozzoni L, Zhang B, et al. Genetic variability of seed sugar content in worldwide soybean germplasm collections. *Crop Science*. 2009;**49**:903-912. DOI: 10.2135/cropsci2008.05.0256
- [162] Gulewicz P, Martinez-Villaluenga C, Kasprowicz-Potocka M, Frias J. Non-nutritive compounds in Fabaceae family seeds and the improvement of their nutritional quality by traditional processing—A review. *Polish Journal of Food and Nutritional Sciences*. 2014;**64**:75-89. DOI: 10.2478/v10222-012-0098-9
- [163] Qiu D, Vuong T, Valliyodan B, Shi H, Guo B, Shannon JG, et al. Identification and characterization of a stachyose synthase gene controlling reduced stachyose content in soybean. *Theoretical and Applied Genetics*. 2015;**128**:2167-2176. DOI: 10.1007/s00122-015-2575-0

- [164] Jones DA, DuPont MS, Ambrose MJ, Frias J, Hedley CL. The discovery of compositional variation for the raffinose family of oligosaccharides in pea seeds. *Seed Science Research*. 1999;**9**:305-310. DOI: 10.1017/S0960258599000318
- [165] Zheng R, Yang L, Zhou X, Zhu C, Shu X, Wu X, et al. Effect of soybean oligosaccharides on immunity and TLR2-NF- κ B signal pathway response for weanling pigs. *Journal of Food, Agriculture and Environment*. 2012;**10**:273-279
- [166] Kim HK, Kang ST, Cho JH, Choung MG, Suh DY. Quantitative trait loci associated with oligosaccharides and sucrose contents in soybean (*Glycine max* L.). *Journal of Plant Biology*. 2005;**48**:106-112
- [167] Kim HK, Kang ST, Oh KW. Mapping of putative quantitative trait loci controlling the total oligosaccharide and sucrose content of *Glycine max* seeds. *Journal of Plant Research*. 2016;**119**:533-538. DOI: 10.1007/s10265-006-0004-9
- [168] Mozzoni L, Chen P. Correlations of yield and quality traits between immature and mature seed stages of edamame soybean. *Journal of Crop Improvement*. 2019;**33**:67-82. DOI: 10.1080/15427528.2018.1542366
- [169] Geater CW, Fehr WR, Wilson LA. Association of soybean seed traits with physical properties of natto. *Crop Science*. 2000;**40**:1529-1534. DOI: 10.2135/cropsci2000.4061529x
- [170] Taira H. Quality of soybeans for processed foods in Japan. *Japan Agricultural Research Quarterly*. 1990;**24**:224-230
- [171] Maughan PJ, Saghai Maroof MA, Buss GR. Identification and quantitative trait loci controlling sucrose content in soybean (*Glycine max*). *Molecular Breeding*. 2000;**6**:105-111. DOI: 10.1023/A:1009628614988
- [172] Matoša Kočar M, Vila S, Petrović S, Rebekić A, Sudarić A, Josipović A, et al. Assessment of phenotypic variability of saccharides in soybean genotypes suitable for growing in Europe. *Journal of Central European Agriculture*. (accepted manuscript)
- [173] Wolf RB, Cavins JF, Kleiman R, Black LT. Effect of temperature on soybean seed constituents: Oil, protein, moisture, fatty acids, amino acids and sugars. *Journal of the American Oil Chemists' Society*. 1982;**59**:230-232. DOI: 10.1007/BF02582182
- [174] Openshaw SJ, Hadley HH. Maternal effects on sugar content in soybean seeds. *Crop Science*. 1978;**18**:581-584. DOI: 10.2135/cropsci1978.0011183X001800040014x
- [175] Neus JD, Fehr WR, Schnebly SR. Agronomic and seed characteristics of soybean with reduced raffinose and stachyose. *Crop Science*. 2005;**45**:589-592. DOI: 10.2135/cropsci2005.0589
- [176] Mozzoni L, Shi A, Chen P. Genetic analysis of high sucrose, low raffinose, and low stachyose content in v99-5089 soybean seeds. *Journal of Crop Improvement*. 2013;**27**:606-616. DOI: 10.1080/15427528.2013.812998
- [177] Akond M, Liu S, Kantartzi SK, Meksem K, Bellaloui N, Lightfoot DA, et al. Quantitative trait loci underlying seed sugars content in "MD96-5722" by "Spencer" recombinant

- inbred line population of soybean. *Food and Nutrition Sciences*. 2015;**6**:964-973. DOI: 10.4236/fns.2015.611100
- [178] Hartwig EE, Kuo TM, Kenty MM. Seed protein and its relationship to soluble sugars in soybean. *Crop Science*. 1997;**37**:770-773. DOI: 10.2135/cropsci1994.0011183X003400050001x
- [179] Krober OA, Carter JL. Quantitative interrelationships of protein and non-protein constituents of soybeans. *Crop Science*. 1962;**2**:171-172
- [180] Openshaw SJ, Hadley HH. Selection to modify sugar content of soybean seeds. *Crop Science*. 1981;**21**:805-808. DOI: 10.2135/cropsci1981.0011183X002100060002x
- [181] Geater CW, Fehr WR. Association of total sugar content with other seed traits of diverse soybean cultivars. *Crop Science*. 2000;**40**:1552-1555. DOI: 10.2135/cropsci2000.4061552x
- [182] Hollung K, Øverland M, Hrustić M, Sekulić P, Miladinović J, Martens H, et al. Evaluation of nonstarch polysaccharides and oligosaccharide content of different soybean varieties (*Glycine max*) by near-infrared spectroscopy and proteomics. *Journal of Agriculture and Food Chemistry*. 2005;**53**:9112-9121. DOI: 10.1021/jf051438r
- [183] Yang K, Ko JM, Joung Ha T, Lee JH, Baek IY, Yang TY, et al. Development of molecular markers for low raffinose and stachyose in Korean soybean cultivars. *Plant Breeding and Biotechnology*. 2014;**2**:151-157. DOI: 10.9787/PBB.2014.2.2.151
- [184] Bhagwat S, Haytowitz DB, Holden JM. USDA Database for the Isoflavone Content of Selected Foods, Release 2.0. Beltsville, MD, USA. Department of Agriculture, Agricultural Research Service, Nutrient Data Laboratory. DOI: 10.15482/USDA.ADC/1209670. Available from: <http://www.ars.usda.gov/nutrientdata/isoflav>
- [185] Kim EH, Ro HM, Kim SL, Kim HS, Chung IM. Analysis of isoflavone, phenolic, soyasapogenol, and tocopherol compounds in soybean (*Glycine max* (L.) Merrill) germplasm of different seed weights and origins. *Journal of Agricultural and Food Chemistry*. 2012;**60**:6045-6055. DOI: 10.1021/jf300463f
- [186] Jeng TL, Shih YJ, Wu MT, Sung JM. Comparisons of flavonoids and anti-oxidative activities in seed coat, embryonic axis and cotyledon of black soybeans. *Food Chemistry*. 2010;**123**:1112-1116. DOI: 10.1016/j.foodchem.2010.05.070
- [187] Lee SJ, Seguin P, Kim JJ, Moon HI, Ro HM, Kim EH, et al. Isoflavones in Korean soybeans differing in seed coat and cotyledon color. *Journal of Food Composition and Analysis*. 2010;**23**:160-165. DOI: 10.1016/j.jfca.2009.08.005
- [188] Li Y, Kong D, Bao B, Ahmad A, Sarkar FH. Induction of cancer cell death by isoflavone: The role of multiple signalling pathways. *Nutrition*. 2011;**3**:877-896. DOI: 10.3390/nu3100877
- [189] Pregelj L, McLanders JR, Gresshoff PM, Schenk PM. Transcription profiling of the isoflavone phenylpropanoid pathway in soybean in response to *Bradyrhizobium japonicum* inoculation. *Functional Plant Biology*. 2011;**38**:13-24. DOI: 10.1071/FP10093

- [190] Meng FL, Han YP, Teng WL, Li YG, Li WB. QTL underlying the resistance to soybean aphid (*Aphis glycines* Matsumura) through isoflavone-mediated antibiosis in soybean cultivar 'Zhongdou 270'. *Theoretical and Applied Genetics*. 2011;**123**:1459-1465. DOI: 10.1007/s00122-011-1680-y
- [191] Messina MJ. Legumes and soybeans: Overview of their nutritional profiles and health effects. *The American Journal of Clinical Nutrition*. 1999;**70**:439-450. DOI: 10.1093/ajcn/70.3.439s
- [192] Gutierrez-Gonzalez JJ, Vuong TD, Zhong R, Yu O, Lee JD, Shannon G, et al. Major locus and other novel additive and epistatic loci involved in modulation of isoflavone concentration in soybean seeds. *Theoretical and Applied Genetics*. 2011;**123**:1375-1385. DOI: 10.1007/s00122-011-1673-x
- [193] Akond AGM, Richard B, Ragin B, Herrera H, Kaodi U, Akbay C, et al. Additional quantitative trait loci and candidate genes for seed isoflavone content in soybean. *The Journal of Agricultural Science*. 2013;**5**:20-33. DOI: 10.5539/jas.v5n11p2
- [194] Adie MM, Krisnawati A, Harnowo D. Agronomic characteristic and nutrient content from several soybean promising lines with high isoflavones. *Procedia Food Science*. 2015;**3**:348-354. DOI: 10.1016/j.profoo.2015.01.038
- [195] Matoša Kočar M, Vila S, Petrović S, Rebekić A, Sudarić A, Duvnjak T, et al. Isoflavone concentrations in soybeans suitable for growing in Europe. *Genetika-Belgrade*. 2019;**51**(1):47-59. DOI: 10.2298/GENSR1901047M
- [196] Tsukamoto C, Shimada S, Igita K, Kudou S, Kokubun M, Okubo K, et al. Factors affecting isoflavone content in soybean seeds: Changes in isoflavones, saponins, and composition of fatty acids at different temperatures during seed development. *Journal of Agricultural and Food Chemistry*. 1995;**43**:1184-1192. DOI: 10.1021/jf00053a012
- [197] Caldwell CR, Britz SJ, Mirecki RM. Effect of temperature, elevated carbon dioxide, and drought during seed development on the isoflavone content of dwarf soybean (*Glycine max* (L.) Merrill) grown in controlled environments. *Journal of Agricultural and Food Chemistry*. 2005;**53**:1125-1129. DOI: 10.1021/jf0355351
- [198] Lozovaya VV, Lygin AV, Ulanov AV, Nelson RL, Dayde J, Widholm JM. Effect of temperature and soil moisture status during seed development on soybean seed isoflavone concentration and composition. *Crop Science*. 2005;**45**:1934-1940. DOI: 10.2135/cropsci2004.0567
- [199] Murphy SE, Lee EA, Woodrow L, Seguin P, Kumar J, Rajcan I, et al. Association of seed and agronomic traits with isoflavone levels in soybean. *Canadian Journal of Plant Science*. 2009;**89**:477-484. DOI: 10.4141/CJPS08148
- [200] Carrera CS, Dardanelli JL. Changes in the relationship between temperature during the seed-filling period and soya bean seed isoflavones under water-deficit conditions. *Journal of Agronomy and Crop Science*. 2015;**202**:421-432. DOI: 10.1111/jac.12147

- [201] Morrison MJ, Cover ER, Saleem MF, McLaughlin NB, Frégeau-Reid J, Ma BL, et al. Seasonal changes in temperature and precipitation influence isoflavone concentration in short-season soybean. *Field Crops Research*. 2010;**117**:113-121. DOI: 10.1016/j.fcr.2010.02.005
- [202] Rasolohery CA, Berger M, Lygin AV, Lozovaya VV, Nelson RL, Daydé J. Effect of temperature and water availability during late maturation of the soybean seed on germ and cotyledon isoflavone content and composition. *Journal of the Science of Food and Agriculture*. 2008;**88**:218-228. DOI: 10.1002/jsfa.3075
- [203] Hasanah Y, Chairun Nisa T, Armidin H, Hanum H. Isoflavone content of soybean [*Glycine max* (L). Merr.] cultivars with different nitrogen sources and growing season under dry land conditions. *Journal of Agriculture and Environment for International Development*. 2015;**109**:5-17. DOI: 10.12895/jaeid.20151.216
- [204] Wang C, Sherrard M, Pagadala S, Wixon R, Scott RA. Isoflavone content among maturity group 0 to II soybeans. *Journal of the American Oil Chemists' Society*. 2000;**77**:483-487. DOI: 10.1007/s11746-000-0077-6
- [205] Primomo VS, Poysa V, Ablett GR, Jackson CJ, Rajcan I. Agronomic performance of recombinant inbred line populations segregating for isoflavone content in soybean seeds. *Crop Science*. 2005;**45**:2203-2211. DOI: 10.2135/cropsci2004.0610
- [206] Hoeck JA, Fehr WR, Murphy PA, Welke GA. Influence of genotype and environment on isoflavone contents of soybean. *Crop Science*. 2000;**40**:48-51. DOI: 10.2135/cropsci2000.40148x
- [207] Akond AGM, Liu S, Kantartzi SK, Meksem K, Bellaloui N, Lightfoot DA, et al. Quantitative trait loci for seed isoflavone contents in 'MD96-5722' by 'Spencer' recombinant inbred lines of soybean. *Journal of Agricultural and Food Chemistry*. 2014;**62**:1464-1468. DOI: 10.1021/jf4040173
- [208] Cvejić J, Tepavčević V, Bursać M, Miladinović J, Malenčić Đ. Isoflavone composition in F1 soybean progenies. *Food Research International*. 2011;**44**:2698-2702. DOI: 10.1016/j.foodres.2011.05.024
- [209] Seguin P, Zheng W, Smith DL, Deng W. Isoflavone content of soybean cultivars grown in eastern Canada. *Journal of the Science of Food and Agriculture*. 2004;**84**:1327-1332. DOI: 10.1002/jsfa.1825
- [210] Yin X, Vyn TJ. Relationships of isoflavone, oil, and protein in seed with yield of soybean. *Agronomy Journal*. 2005;**97**:1314-1321. DOI: 10.2134/agronj2004.0316
- [211] Al-Tawaha AM, Seguin P. Seeding date, row spacing, and weed effects on soybean isoflavone concentrations and other seed characteristics. *Canadian Journal of Plant Science*. 2006;**86**:1079-1087. DOI: 10.4141/P06-043
- [212] Charron CS, Allen FL, Johnson RD, Pantalone VR, Sams CE. Correlations of oil and protein with isoflavone concentration in soybean (*Glycine max* (L.) Merr.). *Journal of Agricultural and Food Chemistry*. 2005;**53**:7128-7135. DOI: 10.1021/jf050610o

- [213] Wang HJ, Murphy PA. Isoflavone composition of American and Japanese soybeans in Iowa: Effects of variety, crop year, and location. *Journal of Agricultural and Food Chemistry*. 1994;**42**:1674-1677. DOI: 10.1021/jf00044a017
- [214] Bursać M, Atanacković Krstonošić M, Miladinović J, Malenčić Đ, Gvozdenović LJ, Hogervorst J, et al. Isoflavone composition, total phenolic content and antioxidant capacity of soybeans with colored seed coat. *Natural Product Communications*. 2017;**12**:475-640
- [215] Sumardi D, Pancoro A, Yulia E, Musfiroh I, Prasetyono J, Karuniawan A, et al. Potential of local black soybean as a source of the isoflavones daidzein and genistein. *International Food Research Journal*. 2017;**24**:2140-2145
- [216] Tepavčević V, Atanacković M, Miladinović J, Malenčić D, Popović J, Cvejić J. Isoflavone composition, total polyphenolic content, and antioxidant activity in soybeans of different origin. *Journal of Medicinal Food*. 2010;**13**:657-664. DOI: 10.1089/jmf.2009.0050
- [217] Morito K, Hirose T, Kinjo J, Hirakawa T, Okawa M, Nohara T, et al. Interaction of phytoestrogens with estrogen receptors alpha and beta. *Biological & Pharmaceutical Bulletin*. 2001;**24**:351-356. DOI: 10.1248/bpb.24.351
- [218] Wang Y, Han Y, Teng W, Zhao X, Li Y, Wu L, et al. Expression quantitative trait loci infer the regulation of isoflavone accumulation in soybean (*Glycine max* L. Merr.) seed. *BMC Genomics*. 2014;**15**:680. DOI: 10.1186/1471-2164-15-680
- [219] Wang Y, Han Y, Zhao X, Li Y, Teng W, Li D, et al. Mapping isoflavone QTL with main, epistatic and QTL × environment effects in recombinant inbred lines of soybean. *PLoS One*. 2015;**10**(3):e0118447. DOI: 10.1371/journal.pone.0118447

