Journal of Functional Food and Nutraceutical

SOYBEAN (*GLYCINE MAX* (L.) MERR.) AND WILD SOYBEAN (*G. SOJA* SIEB. ET ZUCC) GENETIC DIVERSITY FOR FUNCTIONAL FOOD

Maria Stefanie Dwiyanti

Research Faculty of Agriculture, Hokkaido University, Sapporo 060-8589, Japan

ABSTRACT

Soybean has been a major protein source for people in East and Southeast Asia. It contains numerous secondary metabolites that are useful for human and animal health, such as isoflavones, polyphenols, tocopherols, lutein, carotenoids, and saponins. Isoflavones are the most popular nutraceutical compounds of soybean. Several soybean varieties having high isoflavones have been developed and utilized as functional food resources. Tocopherols, lutein, and saponins are also antioxidants present in soybean seeds. Tocopherols are known as vitamin E, whereas lutein has potential to prevent age-related macular degeneration. DDMP saponins in soybean is reported to inhibit HIV infection in vitro and prevent colon cancer. Screening of soybean germplasm showed that soybean holds a large genetic diversity of content and composition of tocopherols, lutein and saponins that can be utilized to breed new soybean cultivars. In addition, the wild counterpart of soybean (wild soybean) can also be source of genetic variations useful for soybean breeding. This review will summarize studies on genetic diversity of tocopherols, lutein, and saponins found in soybean and wild soybean.

Keywords: *DDMP saponin, lutein, soybean, vitamin E, wild soybean*

ABSTRAK

Kedelai telah menjadi sumber protein utama bagi masyarakat di Asia Timur dan Tenggara. Kedelai mengandung banyak metabolit sekunder yang bermanfaat untuk kesehatan manusia dan hewan, seperti isoflavon, polifenol, tokoferol, lutein, karotenoid, dan saponin. Hasil skrining plasma nutfah kedelai menunjukkan bahwa kedelai memiliki keragaman kandungan metabolit sekunder yang besar dan dapat dimanfaatkan untuk pembiakan kultivar kedelai baru. Selain itu, kerabat kedelai atau kedelai liar (wild soybean) juga dapat menjadi sumber variasi genetik yang berguna untuk pemuliaan kedelai. Review ini merangkum studi tentang keragaman genetik zat gizi yang ditemukan pada kedelai dan kedelai liar, dengan fokus pada tokoferol, lutein, dan saponin.

Kata kunci: *Kedelai, kedelai liar, lutein, saponin DDMP, vitamin E*

Article Information

Article Type: Review Paper
Journal Type: Open Access
Volume: 2 Issue 1

Manuscript ID
v2n1400-1

Received Date
12 July 2020

Accepted Date
21 August 2020

Published Date
31 August 2020

DOI:
10.33555/jffn.v2i1.53

Corresponding author:
Maria Stefanie Dwiyanti
Sapporo, Japan
Email:
dwiyanti@abs.agr.hokudai.ac.jp

Citation:
Dwiyanti, M.S., 2020. Soybean (*Glycine Max* (L.) Merr.) and Wild Soybean (*G. Soja* Sieb. Et Zucc) Genetic Diversity for Functional Food. J. Functional Food & Nutraceutical, 2(1), pp.29-42.

I. SOYBEAN AND WILD SOYBEAN

Soybean (*Glycine max* (L.) Merrill) has been a major protein source for human consumption, particularly in East and Southeast Asia. Traditional soy-based food products can be found in East Asia and Southeast Asian countries. Worldwide, soybean is grown in a large scale in Brazil, US, Argentine, China (http://soystats.com/international-world-soybean-production/, accessed July 12th, 2020), mainly for oil and feed. Soybean seeds vary in shape, size, seed coat color, and cotyledon color (Figure 1), as well as in oil content, fatty acid composition, protein, and sugar content.

The ancestor of soybean is wild soybean or *G. soja* (Siebold and Zucc.). It can be found naturally in wide region of China, Japan, Korean Peninsula, and Russia. Wild soybean has small, ellipse black seeds, and the plants form vine (Figure 1). The genetic diversity analyses showed that wild soybean has larger genetic diversity analysis than that of soybean (Kofsky et al. 2018). Moreover, wild soybean also has unique genes that are not present in cultivated soybeans (Xie et al. 2019). Alleles related to soybean cyst nematode resistance; aphid resistance; salt, alkalinity and drought tolerance; yield, seed oil, protein and unsaturated fatty acid have been identified from wild soybean (Xie et al. 2019). Crossing between wild soybean and soybean produces fertile progenies, thus wild soybean has a large potential to be a donor parent for soybean breeding.

The Japan National Institute of Agrobiological Sciences (NIAS) Genebank holds a collection of approximately 11,300 accessions consisting of Japan and overseas landraces, improved varieties and breeding lines developed by agriculture research centers, and also wild soybean accessions (Kaga et al. 2012). In addition, U.S. Departement of Agriculture (USDA) genebank holds a collection of nearly 22,000 accessions of soybean, wild soybean, and perennial soybeans (Bandillo et al. 2015). After filtering out duplicates, there were 14,430 unique accessions originated from 85 countries. Of these, 14,000 soybean and wild soybean accessions have been genotyped using 50,000 single nucleotide polymorphism (SNP) chip and they were evaluated for their oil and protein content (Bandillo et al. 2015).

There are growing interests in developing soybean as functional food. The most popular compounds of soybean that have been utilized in functional food, supplements, or cosmetics are isoflavones. However, soybean and wild soybean also contain many other secondary metabolites that can be developed further for new products. Here, three compounds selected: tocopherols, lutein, and saponins. This review will discuss about the genetic diversity that found in soybean and wild soybean, the impact of crop management and the environmental condition to the content and composition of the three traits.

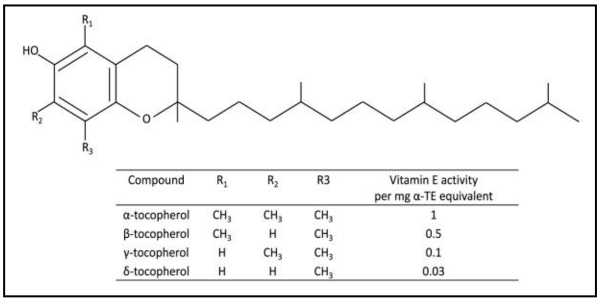
Figure 1. Seed and plant types of soybean and wild soybean. left: example of seeds of soybean and wild soybean. right: example of plants of wild soybean and soybean.

II. STUDIES ON GENETIC DIVERSITY OF NUTRITIONAL COMPOUNDS

2.1. Tocopherols

2.1.1. Tocopherol in soybean and its biosynthesis regulation

Tocopherols are lipophilic antioxidants belong to tocochromanol group, which is known to possess vitamin E activity (Bramley et al. 2000). The structure consists of a chromanol head and a phytyl tail. Based on the number and position of methyl group on the chromanol head, tocopherols are divided to four isoforms (α-, β-, γ-, and δ-tocopherol) (Figure 2). Natural isoforms of tocopherols have *RRR*-forms, thus tocopherols from natural sources are designated as *RRR*-α-, β-, γ-, and δ-tocopherol, respectively. Among four isoforms, α-tocopherol has the highest vitamin E activity because it is preferentially absorbed by α-tocopherol transfer protein in liver and it is retained at high level in human blood plasma. When *RRR*-α-tocopherol vitamin E activity is 1 mg α-tocopherol equivalents (α-TE mg-1 compound), the vitamin E activities of *RRR*-β-, *RRR*-γ-, and *RRR*-δ-tocopherol are 0.5, 0.1, and 0.03 mg, respectively (Figure 2; Bramley et al. 2000). Besides tocopherols, there are tocotrienols which differ from tocopherols by the presence of unsaturated hydrogen bonds in the phytyl tail. Rice bran contains high level of tocotrienols, however, soybean contains no tocotrienols.



Compound	R_1	R_2	R3	Vitamin E activity per mg α-TE equivalent
α-tocopherol	CH_3	CH_3	CH_3	1
β-tocopherol	CH_3	H	CH_3	0.5
γ-tocopherol	H	CH_3	CH_3	0.1
δ-tocopherol	H	H	CH_3	0.03

Figure 2. Tocopherol isoforms and its vitamin E activity

Tocopherols are commonly added in food and cosmetics to prevent oil oxidation. Soybean oil is also one of the sources of vitamin E supplement sold in markets. Soybean oil contains about 1,000 mg tocopherols per kg edible portion (Bramley et al. 2000). This value is higher compared with oils sourced from canola (253.4 mg tocopherols per kg edible portion), sunflower (546 mg tocopherols per kg edible portion), or palm (386 mg tocopherols per kg edible portion) (Bramley et al. 2000; Sheppard et al. 1993). However, the vitamin E activity as α-TE (mg) of soybean oil is lower than that of canola, sunflower, or palm, because soybean seeds contain α-tocopherol only less than 10% of total tocopherol content. The predominant form of tocopherol in soybean seeds is γ-tocopherol (60-70% of total tocopherol content), followed by δ-tocopherol (20-30% of total tocopherol content). The remaining isoform, β-tocopherol is usually present at negligible amount (1-2% of total tocopherol content).

It has been great interest to increase vitamin E content in soybean seeds by increasing its α-tocopherol ratio. Tocopherol biosynthesis pathway is already elucidated (Figure 3), and genes involved in the tocopherol biosynthesis have been characterized in model plant *Arabidopsis thaliana*, and several oil crops such as soybean, maize, rapeseed, and sunflower (Fritsche et al. 2017). The first step in tocopherol biosynthesis is the creation of 2-methyl-6-phytyl-1,4-benzoquinol (MPBQ) by joining homogentisic acid (HGA) and phytyl-diphosphate (PDP). This step is catalyzed by homogentisic phytyltransferase (HPT). MPBQ-methyltransferase enzyme (MPBQ-MT) adds a methyl group to MPBQ, creating 2,3-dimethyl-5-phytyl-1,4-benzoquinone (DMPBQ). The chromanol heads of MPBQ and DMPBQ were cyclized to create δ-tocopherol and γ-tocopherol, respectively. This step is catalyzed by tocopherol cyclase (TC). The last step of tocopherol biosynthesis is the addition of a methyl group to δ-tocopherol and γ-tocopherol, creating β-tocopherol and α-tocopherol, respectively. This step is catalyzed by γ-tocopherol methyltransferase (γ-TMT). In *A.thaliana*, *VTE3* and *VTE4* genes encode MPBQ-MT and γ-TMT, respectively. Overexpressing *VTE4* in soybean increased α-tocopherol ratio to 50~75% of total tocopherol content, whereas overexpressing both *VTE3* and *VTE4* in soybean increased α-tocopherol ratio to 60~91% of total tocopherol content (Van Eenennaam et al. 2003).

2.1.2. Genetic diversity of tocopherol composition and content in soybean germplasm

Soybean has a considerable diversity of tocopherol composition and tocopherol content. Studies from India (Rani et al. 2007), Brazil (Carrão-Panizzi and Erhan 2007), China (Zhan et al. 2020), worldwide germplasm (Ujiie et al. 2005), wild soybean (Dwiyanti et al. 2016) will be introduced here. It is important to take note that the measurement differs between studies. Rani et al. (2007) used μg/g of oil, Carrão-Panizzi and Erhan (2007) used ppm, Ujiie et al. (2005) used mg/ 100 g seed meal whereas

Dwiyanti et al. (2016) used percentage of total tocopherol content.

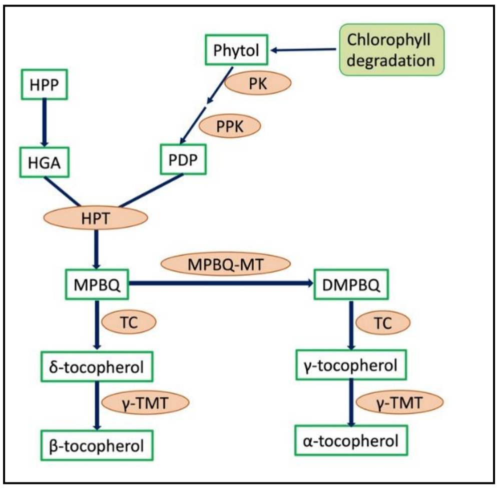


Figure 3. Tocopherol biosynthesis pathway. Substrates are shown in green boxes: HPP: *p*-hydroxyphenylpyruvic acid, PDP: phytyl diphosphate, HGA: homogentisic acid, MPBQ: 2-methyl-6-phytyl-1,4-benzo-quinol, DMPBQ: 2,3-dimethyl-6-phytyl-1,4-benzoquinone. Enzymes are shown in brown circles. PK: phytol kinase, PPK: phytyl phosphate *kinase*, HPT: homogentisate phytyltransferase, MPBQ-MT: MPBQ methyltransferase, TC: tocopherol cyclase, γ-TMT: γ-tocopherol methyltransferase.

Screening of 1,109 worldwide soybean and wild soybean accessions, Ujiie et al. (2005) identified three soybean accessions having high α-tocopherol ratio Keszthelyi Aproszemu Sarga (KAS), Dobrogeance, and Dobrudza 14 Pancevo). Two soybean cultivars used as control (Toyokomachi and Ichihime) have total tocopherol content 10 - 12 mg/ 100 g seed meal. Initial screening using seeds from genebank showed that KAS had higher α-tocopherol content, but significantly lower total tocopherol content compared with other accessions. Its α-tocopherol content was about 4.26 mg / 100 g seed meal of total tocopherol content (8 mg / 100 g seed meal), thus the α-tocopherol ratio was more than 50% (Ujiie et al. 2005). The other high α-tocopherol accessions Dobrogeance and Dobrudza had α-tocopherol content about 3.71 mg / 100 g seed meal and 3.26 mg / 100 g seed meal, but their total tocopherol content was similar to that of Toyokomachi and Ichihime (Ujiie et al. 2005). It is important to note that the screening used seeds coming from plants grown at different locations. KAS was grown in National Shikoku Agricultural Experiment Station, Kagawa; Dobrogeance and var. Dobrudza 14 Pancevo were grown at the National Agricultural Research Center for Tohoku Region, Akita; and the control cultivars were grown at Hokkaido University Agricultural Farm (Ujiie et al. 2005). Among these, Kagawa is the southernmost location. When grown in Hokkaido, the α-tocopherol ratio of KAS will be about 20% of total tocopherol content (Dwiyanti et al. 2007, Dwiyanti et al. 2011). It is known that the α-tocopherol ratio is largely affected by temperature during seed filling (Britz and Kremer 2002, Dwiyanti et al. 2016), thus the 50% α-tocopherol ratio of KAS in initial screening may be resulted from the temperature effect. In addition, screening of 600 wild soybean accessions originated from Japan and South Korea (Dwiyanti et al. 2016) identified 16 wild soybean accessions having α-tocopherol ratio more than 20%. Cultivation of these accessions in controlled temperature showed that the α-tocopherol ratio was more than 20% for 15 accessions, and 14% for 1 accession (Dwiyanti et al. 2016), indicating that the high α-tocopherol ratio in these accessions was genetically inherited. Genetic analysis showed that the genetic variations in *γ-TMT3*, one of the genes encoding γ-tocopherol methyltransferase were associated with high α-tocopherol ratio in soybean KAS and wild soybean B04009.

Analysis of 66 Indian soybean cultivars (Rani et al. 2007) showed large variability for all tocopherol isoforms as well as total tocopherol content. The α-tocopherol content of 66 Indian soybean cultivars ranged from 58 µg/g to 794 µg/g, with an overall mean value of 269 g/g of oil, whereas the total tocopherol content ranged from 422 µg/g to 3,311 µg/g of soybean oil. If the value is converted to ratio, the α/total tocopherol ratio ranged between 11 to 26%, with an average value 18%. This ratio is higher than the values reported in Ujiie et al. (2005). On the other hand, analysis of 89 Brazil soybean accessions showed the range from 11 ppm to 191 ppm for α-tocopherol, and from 561 ppm to

1,983 ppm for total tocopherols (Carrão-Panizzi and Erhan 2007). The α-tocopherol ratio range was between 1.8% to 11.7%, which is lower than the values obtained from India or Japan study. Recently, Zhan et al. (2020) investigated the tocopherol content of 180 soybean accessions from China. The populations were grown at two locations for 3 years. The range of α-tocopherol and total tocopherol content was different for each year × location but the range was between 1.4 μg/g to 98.0 μg/g for α-tocopherol content, and 155.0 μg/g to 399.30 μg/g for total tocopherol content.

2.1.3. Effect of growing condition to soybean tocopherol composition and content

High temperature and drought stress during seed filling affected α-tocopherol content (Britz and Kremer 2002). A 5°C increase in temperature from 23°C to 28°C during seed filling increase the α-tocopherol up to twice of the usual content (Britz and Kremer 2002). The same phenomenon was also observed in wild soybean. Growing 16 wild soybean genotypes having high α-tocopherol ratio in 20°C, 25°C, and 30°C during seed filling period showed that α-tocopherol ratio increase following the increment in temperature (Dwiyanti et al. 2016). Interestingly, the increase rate differed among genotypes. For example, the α-tocopherol ratio of genotype B09092 matured in 30°C was twice of the α-tocopherol ratio of the same genotype matured in 25°C. On the contrary, the α-tocopherol ratio of genotype B00092 was not significantly different between seed matured in 25°C or in 30°C (Dwiyanti et al. 2016). Difference in temperature response may be resulted from genetic variations in genes involved in tocopherol biosynthesis pathway. Park et al. (2019) investigated the expression level of three soybean genes (*γ-TMT1*, *γ-TMT2*, and *γ-TMT3*) encoding γ-tocopherol methyltransferase of developing seeds of TK780 (soybean, low α-tocopherol ratio) and B04009 (wild soybean, high α-tocopherol ratio) grown in 20°C and 30°C during seed filling. Previous studies showed that genetic variations in *γ-TMT3* are associated with high α-tocopherol ratio in KAS (Dwiyanti et al. 2011) and B04009 (Park et al. 2019). Interestingly, the gene expression analysis showed that only *γ-TMT2* expression level increased in both genotypes when temperature increased to 30°C. In contrast to expectation, *γ-TMT3* expression decreased in both genotypes when temperature increased to 30°C (Park et al. 2019). Whether this phenomenon is common to all soybean genotypes is remained to be elucidated.

Other environment effects such as seeding rate, seeding date, row spacing, and P+K fertilization were also investigated in Canada environment (Seguin et al. 2010). Among these, early seeding date resulted to higher α-tocopherol ratio up to 45% compared to that of later seeding date. It is suggested that higher temperature during seed filling in early seeding date regime affected the α-tocopherol ratio. Other treatments impacted the α-tocopherol ratio, but the results were not consistent across environments.

It is important to note that increasing only α-tocopherol in soybean seeds may not be beneficial. It has been suggested that γ-tocopherol plays important role in seed desiccation tolerance, which may be the reason of γ-tocopherol being the major isoform of tocopherols in seeds of most plants (Fritsche et al. 2017). If present as tocopherol mixture in oil, α-tocopherol will be oxidized first followed by other tocopherol isoforms (Seppanen et al. 2010). In addition, the antioxidant effect of 50% α-tocopherol : 50% γ-tocopherol mixture and natural tocopherols mixture from soybean (13% α-tocopherol, 64% γ-tocopherol and 21% δ-tocopherol) were 250 and 500 ppm, respectively (Huang et al. 1995), indicating that α-tocopherol over certain ratio can act as prooxidant.

2.2. Lutein

2.2.1. Lutein in soybean and biosynthesis regulation

Lutein is known as pro-vitamin A carotenoids and is able to prevent early, intermediate, and advanced AMD (age-related macular degeneration) (Ranard et al. 2017). Major sources of lutein are egg yolk, green leafy vegetables, and orange/yellow fruits. Lutein is often present together with zeaxanthin, another type of carotenoid, and they are both present in the macula, a part of retina where photoreceptor cells are highly concentrated. Daily intake about 3-5 mg lutein per day may reduce the risk of all stages of AMD (Ranard et al. 2017). The maximum dose recommended for dietary intake is yet to be determined, however it is suggested that the intake up to 20 mg/day for lutein is considered safe (Ranard et al. 2017, Toti et al. 2018).

Lutein is also an important component in poultry feed. Since animals cannot produce carotenoids by their own, lutein supplementation into the poultry feed helps enrichment of lutein content in eggs. Purified lutein supplements up to 1,000 mg/kg, increased lutein concentrations by up to 10-fold in eggs (Pitargue et al. 2019). Lutein supplements for feed are mainly sourced from corn co-products, alfalfa, marigold extract, or algae (Pitargue et al. 2019). Globally, soybean meal contributed up to 69% of all protein sources for animal feeds. In the U.S., it reached up to approximately 92% of animal feed, and 48% of which was used for poultry feed (Cromwell, 2017). It has been an interest in high-lutein soybean in order to save cost used in lutein supplementation from other sources.

Lutein is the most abundant carotenoids in soybean seeds (Monma et al. 1994). Carotenoid biosynthesis pathway has been elucidated (Figure 4). Carotenoid biosynthesis starts by combining two geranylgeranyl pyrophosphate (GGPP) to form phytoene. The process is catalyzed by phytoene synthase (PSY). Four reaction steps catalyzed by phytoene desaturase, ζ-carotene desaturase, ζ-carotene isomerase and carotenoid isomerase convert phytoene to lycopene. Two processes catalyzed by lycopene ε-cyclase and lycopene β-cyclase respectively convert lycopene to α-carotene. Finally, α-carotene is converted to lutein through two pathways, zeinoxanthin or α-cryptoxanthin. Generally, immature soybean seeds contain lutein. However, these compounds degrade during seed filling, as the result, lutein is present at low concentration in mature yellow soybean seeds (Monma et al. 1994).

2.2.2. Genetic diversity of lutein content in soybean and wild soybean

Genetic diversity of lutein content in soybean was reported by Kanamaru et al. (2006) who evaluated 490 soybean and 610 wild soybean genotypes. The lutein content of soybean genotypes ranged between 0.16 to 1.48 mg/100 g meal (Kanamaru et al. 2006). Wild soybean showed higher range of lutein content, between 0.58 to 3.28 mg/100 g meal (Kanamaru et al. 2006). Of these, seven wild soybean genotypes had lutein content more than 2.1 mg/100 g meal, and the values were stable across two years planting experiment. High lutein wild soybean can be distinguished by deep yellow cotyledon. High lutein wild soybean also had high β-carotene and chlorophyll content but these compounds were found in seed coat whereas lutein was produced in cotyledon. Crossing high lutein wild soybean to low lutein soybean cultivar produced progenies with segregating lutein content (Kanamaru et al. 2008). Comparing lutein content of high lutein progenies and low lutein progenies showed that there was no difference in lutein content at immature seed state (Kanamaru et al. 2008). However, lutein content significantly decreased toward maturation in low lutein progenies compared to high lutein progenies, indicating that genetic variations in gene(s) involved in lutein degradation may be associated with high lutein content.

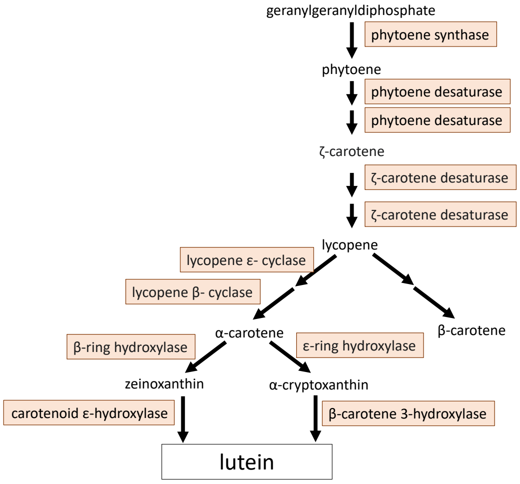


Figure 4. Lutein biosynthesis pathway (redrawn based on kyoto encyclopedia of genes and genomes (KEGG) database: http: //www.genome. jp/keggbin/show_pathway?gmx00906; last accessed July 8th, 2020).

2.2.3. Effect on crop management and growth environment on soybean lutein content

Unlike tocopherols, soybean's lutein content is more determined by the genotype and it is less affected by crop management practice and growth environment (Kanamaru et al. 2006, Lee et al. 2008, Seguin et al. 2011). However, more studies are needed to confirm the effect of planting date on lutein content, particularly in the tropical regions where there is little difference in daylight length and temperature. A genotype × environment interaction study conducted on 7 high lutein wild soybean genotypes showed that there was no large variation in between two years planting (2004 and 2005; Sapporo, latitude 43°04'N) (Kanamaru et al. 2006). The trend was confirmed by another study showed that when 15 soybean genotypes grown on farms in Portageville, US (latitude 36°44'N) (Lee et al. 2008). Lee et al. (2008) also showed that planting date affected lutein content but the response between soybean genotypes varied. However, a study conducted in Quebec, Canada showed that the planting date affected on lutein content (Seguin et al. 2011). This study only tested the planting date effect on two soybean genotypes, AC Proteina and OAC Vision, grown in two locations Ste-Anne-de-Bellevue (45°24'N) and Normandin (48°49'N). Later planting date in Ste-Anne-de-Bellevue significantly increased lutein content but the same treatment in Normandin significantly decreased lutein content. Whether temperature during seed filling affects lutein content is remained to be investigated.

Other crop management practices were assessed such as seeding rate, row spacing, P and K fertilization and there was no clear trend of between treatments and the differences were small. It was less than 8% for row spacing and P treatment, as well as there was no difference in K fertilization (Seguin et al. 2011).

2.3. Saponins

2.3.1. Saponins in soybean and biosynthesis regulation

Soyasaponins are oleanane-type triterpenoids, widely distributed among Leguminosae species. Based on its aglycone structure, soyasaponins are categorized into soyasapogenol A and soyasapogenol B (Figure 5, Takada et al. 2013, Krishnamurthy et al. 2019). Soyasapogenol A and soyasapogenol B can be distinguished by the presence of hydroxy group on the C-21 position. In plants, sugar chain is attached to the C-3 position of soyasapogenol A, creating group A saponins. In soybean seeds, soyasapogenol B group is present as DDMP saponin. In addition to attachment of sugar chain at C-3 position, a DDMP moiety is attached to C-22 position, creating DDMP saponin. Based on type of sugars attached to C-3 or C-22 position, there are more than 30 saponins exist in soybean seeds (Table 1). When soybean seeds are crushed, extracted, or processed in cooking, DDMP saponin degrades to B saponin and E saponin. Several studies report the saponin composition or content of soybean seeds as group A and group B saponins, because when DDMP saponin is extracted from soybean seeds, it is rapidly degraded to B saponins during the processing or during analysis using high performance liquid chromatography (HPLC) (Takahashi et al. 2017).

The acetylated group A saponins cause bitter and astringent taste of soybean (Sundaramoorthy et al. 2018), however recently it is also reported that group A saponins may have anti-obesity effect (Yang et al. 2015). On the other hand, group DDMP saponins and its derivative B saponins are reported to be less bitter and possess health benefits such as radical oxygen scavenger, anti-mutagenic activities (Berhow et al. 2002), inhibitor of HIV infection in vitro (Nakashima et al. 1989), and colon cancer prevention (Tsai et al. 2010). It has been an interest to increase DDMP saponin and to decrease the amount of A saponin in soybean seeds. Group A saponins are present mainly in hypocotyl, whereas DDMP saponins are found in cotyledon (Sundaramoorthy et al. 2018).

Saponin biosynthesis pathway in soybean has not been fully elucidated (Figure 5). The first step of saponin biosynthesis is the creation of β-amyrin. The next step is the hydroxylation of C-24 position of β-amyrin to 24-hydroxyl β-amyrin. The conversion may be in several steps, but it is not elucidated yet. Hydroxylation of C-22 position creates soyasapogenol B, and hydroxylation of C-

Table 1. Soyasaponin types based on sugar-chain at C-3 and C-22 positions

	C-22 position						
C-3 position (R1)	Soyasapogenol A				Soyasapogenol B		
	Aa-series	Ab-series	A0-series	A-series	DDMP	Group B	Group E
	acetylXyl-	acetylGlc-	H-	*	DDMP	H-	O=
β-**D**-glucosyl-(1→2)-*O*-β-**D**-Gal-†	Aa	Ab	A0-αg	A-αg	αg	Ba	Bd
β-**L**-rhamnosyl-(1→2)-*O*-β-**D**-Gal-	Au	Ac	A0-βg	A-βg	βg	Bb	Be
β-**D**-Gal-	Ae	Af	A0-γg	A-γg	γg	Bb’	Be’
β-**D**-glucosyl-(1→2)-*O*-α-**L**-Ara-‡	Ax	Ad	A0-αa	A-αa	αa	Bx	Bf
β-**L**-rhamnosyl-(1→2)-*O*-α-**L**-Ara-	Ay	Az	A0-βa	A-βa	βa	Bc	Bg
α-**L**-Ara-	Ag	Ah	A0-γa	A-γa	γa	Bc’	Bg’
(δ-series) H-	Aa-δ	Ab-δ	A0-δ	A-δ	DDMP-δ	B-δ	E-δ

*A-series saponins do not contain arabinose at the C-22 position. †: galactose; ‡: arabinose. The categorization was redrawn based on Khrisnamurthy et al. (2019).

21 of soyasapogenol B creates soyasapogenol A. Additions of sugar chains and DDMP create DDMP saponins, whereas the additions of sugar chains at C-3 and C-22 positions of soyasapogenol A create group A saponins.

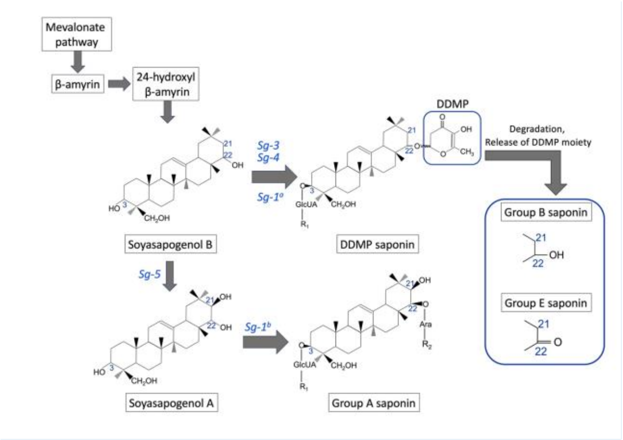


Figure 5. Putative saponin biosynthesis pathway in soybean (redrawn based on krishnamurthy et al. 2019). *Sg-1*^{a}, *Sg-1*^{b}, *Sg-3*, *Sg-4*, and *Sg-5* are loci where genes involved in saponin biosynthesis pathway are located. number 3, 21, 22 correspond to the hydroxyl position in the soyasapogenol.

There are several types of *Sg-1* locus that lead to structural diversity of sugar at C-22 position of group A-saponins (Sayama et al. 2012). *Sg-1*^{a} and *Sg-1*^{b} encode UDP-sugar–dependent glycosyltransferases, UGT73F4 and UGT73F2, which add xylose and glucose, respectively, to the arabinose residue at the C-22 position. A rare *sg-1*^{0} mutant does not have the ability to add sugar chain, and this is resulted in A0-αg type soybean. The *sg-1*^{0} mutant has been used in breeding program, led to development of ‘Kinusayaka’, a soybean cultivar having less beany and astringent flavor, and suitable for soymilk and tofu (Kato et al. 2007).

Sg-3 locus contains a gene *Glyma.10G104700* encoding a putative glycosyltransferase (UGT91H9) that adds glucose as third moiety atsugar chain at C-3 position of saponins (Yano et al. 2018). Soybean cultivar ‘Mikuriya-ao’ is a *sg-3* natural mutant that lost *Glyma.10G104700 gene, resulting in the absence of* Ab and αg saponins and accumulation of Af saponin. *Sg-4 locus contains gene Glyma.01G046300 encoding a UDP-arabinose glycosyltransferase that adds second arabinose moiety at the C-3 position of soyasaponins (Takagi et al. 2018). Sg-5 locus contains gene Glyma.15G243300 encoding* the cytochrome P450 72A69 enzyme and is responsible for conversion of soyasapogenol B to soyasapogenol A (Sundaramoorthy et al. 2018). A mutant lacking sg-5 gene and has no group A-saponins has developed into variant ‘Tohoku 152’ (Sakai et al. 2002).

Recently, rare *Sg-6* saponin types were identified from wild soybean germplasms in Japan and South Korea (Krishnamurthy et al. 2014), of which there are additional moieties to C-29 position and hydroxy group at C-22 position, creating new type of soyasapogenols, designated as soyasapogenol-H, soyasapogenol-I, and soyasapogenol-J. Based on type of sugar chain at C-29 position, soyasapogenol-H, soyasapogenol-I, and soyasapogenol-J have $-CH_2OH$, -COOH, and malonyl-CH2- moiety, respectively (Krishnamurthy et al. 2014).

2.3.2. Genetic diversity of saponin content and composition in soybean and wild soybean

Three major studies analyze saponin content and composition will be introduced here. A study conducted on 3,720 Korean wild soybean accessions discovered 7 common saponin allele combination (Panneerselvam et al. 2013). The 7 combinations are *Sg-1*^{a}*/sg-4/sg-6* (phenotype: Aa), *Sg-1*^{b}*/sg-4/sg-6* (Ab), *Sg-1*^{a}*/Sg-4/sg-6* (AaBc), *Sg-1*^{b}*/Sg-4/sg-6* (AbBc), *Sg-1*^{a}*/sg-4/Sg-6* (Aa+α), *Sg-1*^{a}*/Sg-4/Sg-6* (AaBc+α) and *Sg-1*^{b}*/Sg-4/Sg-6* (AbBc+α). Interestingly, the composition of allele combinations differs between regions. Aa phenotype is the most common in east region of Korea (50%), followed by AaBc (42%). From west region, the composition is reversed with AaBc phenotype is the most dominant (60-65%) followed by Aa (30%). The wild soybean originated from South region of Korea has similar composition with the west region, but the percentage of Aa is lower (4-16%), and the third genotype AaBc+α is 2 to 16% of total genotypes. Another study on soyasaponin diversity in the Chinese wild soybean germplasm was conducted on 3,795 accessions (Takahashi et al. 2017). These wild soybeans contain combination of several saponins, with Aa and Ae being two dominant group A-saponins. 78% and 78.9% of total samples have Aa and Ae saponins, respectively. Of DDMP saponins, almost all samples have αg and βg saponins (99% and 100%, respectively). Five accessions carrying rare A0-αg type similar to Kinusayaka, and three accessions carrying both A0-αg and A0-αa saponins. The genetic basis of these components is unknown, but 7 accessions were collected from northeastern China and 1 accession from Henan Province, central China. Interestingly, an accession No.3137 showed Bd accumulation. Bd is derived from soyasapogenol E, which is thought as degradation product of DDMP saponin. Takahashi et al. (2017) claimed that Bd accumulation observed from No. 3137 was inherited through three generations. The last study came from the screening of high-density soybean mutant library (Panneerselvam et al. 2019). The mutant library was developed by treating cultivar 'Enrei' with ethyl methanosulfonate to induce genetic mutations. Thirty-five mutants having unusual saponin phenotypes were identified. Among these, six mutants deficient in total saponins, 11 mutants deficient in DDMP saponins, 3 mutants accumulating only group A saponins, and 5 mutants showing A0-type saponins will be useful for further study of saponin biosynthesis pathway (Panneerselvam et al. 2019). Interestingly, mutants lacking total saponins did not show any abnormalities in the progenies. Mutants lacking in total saponins will be also useful as donor parent in breeding new soybean cultivars.

2.3.3. Saponin content as response to environment

There are only few reports on growing environment effect on saponin content in soybean seeds (Tsukamoto et al. 1995, Seguin et al. 2014). Tsukamoto et al. (1995) particularly investigated the effect of temperature during seed development to DDMP saponin content in three soybean cultivars (Koganedaizu, Suzuyutaka, and Higomusume). There were tendency of higher αg, βg-saponin in higher temperature regime, and lower βa-saponin in three cultivars, although the result was not statistically significant. There was no difference in total DDMP saponin content between high and low temperature. Similar trend was obtained from study in Canada (Seguin et al. 2014). Seguin et al. (2014) investigated the effect of seeding dates and high temperature on soyasaponin B content of 20 early maturing soybean genotypes grown at four locations in Canada. Early seeding date was associated with higher soyasaponin B content (Seguin et al. 2014). In Canada, early seeding date is usually associated with higher temperature during seed filling. Interestingly, test on high temperature during seed filling using two soybean genotypes OAC Champion and AC Proteina showed that the effect depends on soybean genotypes. Exposure to high temperature (33/25

°C, day/night temperatures) exclusively during seed filling, total soyasaponin B content in AC Proteina increased by 28% compared to that of the control condition (23/15 °C), but not in OAC Champion. However, high temperature during all growing stages decreased the total soyasaponin B content in both cultivars. The decrease may be resulted from the heat stress experienced by the plants; therefore, this factor should be considered when growing soybean for high soyasaponin B content purpose in the tropical regions.

Another interesting report was the investigation on the effect of artificially damaged goldenrod volatiles (ADGV) to saponin content in black soybean seeds (Shiojiri et al. 2020). The idea came from plant-plant communication concept; when a plant is damaged by insects or herbivores, it will release volatiles that are received by other plants. When the other plants received such volatiles, they also release volatiles as defense response. In the study, two soybean cultivars: black soybean cultivar 'Hyokei Kuro-3' and yellow soybean cultivar 'Yumesayou' were grown together with goldenrod (*Solidago altissima*). The Ab, βa, and Bc saponin content of 'Hyokei Kuro-3' increased when exposed to ADGV, compared to non-exposed plants. For 'Yumesayou', the increase was observed for Af, Ab, αg, βg, γg, compared to non-exposed plants. It is important to note that the ADGV effect might differ depending on soybean genotypes.

III. UTILIZATION OF THE RESOURCES INTO SOYBEAN BREEDING

Most of accessions having high nutraceutical content here are landraces or wild soybean. Thus, these traits can be introduced to current leading cultivars through breeding. For efficient breeding, it is important to introgress only the genetic variations associated with high nutritional content. To do this, we need to elucidate the genetic variations associated with the beneficial traits and develop genetic markers. The availability of whole genome sequence information of soybean (Schmutz et al. 2010) and wild soybean (Xie et al. 2019) accelerates characterization of genetic variations exist in soybean and wild soybean germplasm. It is also important to further study the genotype × environment effect to the traits. Most of examples shown here based on studies in subtropics, where the daylength change greatly throughout the growing period and temperature is much lower compared to tropical countries such as Indonesia. Particularly, temperature effect should be investigated in case of α-tocopherol. Study cases shown here used temperature maximum at 28°C. Temperature higher than that may increase α-tocopherol content, but it also induces heat stress in plants. Lastly, as breeders may want to stack several beneficial traits into one cultivar, it is important to investigate whether it is possible to increase the content of α-tocopherol, lutein or DDMP saponin without decreasing the content of oil, protein, or other carotenoids that share same biosynthesis pathway.

REFERENCES

Bandillo, N., Jarquin, D., Song, Q., Nelson, R., Cregan, P., Specht, J. and Lorenz, A. (2015), A Population Structure and Genome-Wide Association Analysis on the USDA Soybean Germplasm Collection. The Plant Genome. 8, pp.1-13.

Berhow, M.A., Cantrell, C.L., Duval, S.M., Dobbins, T.A., Maynes, J., and Vaughn, S.F. 2002. Analysis and quantitative determination of group B saponins in processed soybean products. *Phytochemical Analysis*. 13, pp.343-348.

Bramley, P., Elmadfa, I., Kafatos, A., Kelly, F., Manios, Y., Roxborough, H., Schuch, W., Sheehy, P. and Wagner, K.H. 2000. Vitamin E. *Journal of the Science of Food and Agriculture*. 80, pp.913-938.

Britz, S.J., and Kremer, D.F. 2002. Warm temperatures or drought during seed maturation increase free α-tocopherol in seeds of soybean (*Glycine max* [L.] Merr.). *Journal of Agriculture and Food Chemistry*. 50(21), pp.6058-6063.

Carrão-Panizzi, M., and Erhan, S. 2007. Environmental and genetic variation of soybean tocopherol content under Brazilian growing conditions. *Journal of the American Oil Chemists' Society*. 84(10), pp. 921-928.

Cromwell, G.L. 2017. Soybean meal - An exceptional protein source. www.soymeal.org, accessed July 5th, 2020.

Dwiyanti, M.S., Maruyama, S., Hirono, M., Sato, M., Park, E., Yoon, S.H., Yamada, T., and Abe, J. 2016. Natural diversity of seed α-tocopherol ratio in wild soybean (*Glycine soja*) germplasm collection. *Breeding Science*. 66, pp.653-657.

Dwiyanti, M.S., Ujiie, A., Thuy, L.T.B., Yamada, T., and Kitamura, K. 2007. Genetic analysis of high α-tocopherol content in soybean [*Glycine max*] seeds. *Breeding Science*. 57, pp.23-28.

Fritsche, S., Wang, X., Jung, C. 2017. Recent advances in our understanding of tocopherol biosynthesis in plants: An overview of key genes, functions, and breeding of vitamin E improved crops. *Antioxidants*. 6(4), pp.99.

Huang, S.W., Frankel, E.N., and German, J.B. 1995. Effects of individual tocopherols and tocopherol mixtures on the oxidative stability of corn oil. *Journal of Agriculture and Food Chemistry*. 43, pp. 2345-2350.

Kaga, A., Shimizu, T., Watanabe, S., Tsubokura, Y., Katayose, Y., Harada, K., Vaughan, D. A., and Tomooka, N. 2012. Evaluation of soybean germplasm conserved in NIAS genebank and development of mini core collections. *Breeding Science*. 61(5), pp.566–592.

Kanamaru, K., Wang, S., Abe, J., Yamada, T., Kitamura, K. 2006. Identification and characterization of wild soybean (*Glycine soja* Sieb. et Zecc.) strains with high lutein content. *Breeding Science*. 56(3), pp.231-234.

Kanamaru, K., Yamada, T., and Kitamura, K. 2008. Biochemical characterization of high lutein trait of wild soybean (*Glycine soja* Sieb. et Zucc.) and expression analysis of gene encoding enzymes involved in lutein biosynthesis for the breeding of lutein rich soybean. *Soy Protein Research, Japan* 11, pp.44-50.

Kato, S., Yumoto, S., Takada, Y., Kono, Y., Shimada, S., Sakai, T., Shimada, H., Takahashi, K., Adachi, T., Tabuchi, K., and Kikuchi, A. 2007. A new soybean cultivar “Kinusayaka” lacking three lipoxygenase isozymes and group A acetyl saponin. *Bulletin* of the National Agriculture Research Center *Tohoku Region*. 107, pp.29-42.（in Japanese with English summary）

Kofsky, J., Zhang, H., and Song, B.H. 2018. The untapped genetic reservoir: The past, current, and future applications of the wild soybean (*Glycine soja*). *Frontiers in Plant Science*. 9, pp.949.

Krishnamurthy, P., Fujisawa, Y., Takahashi, Y., Abe, H., Yamane, K., Mukaiyama, K., Son, H.R., Hiraga, S., Kaga, A., Anai, T., Tsukamoto, C., and Ishimoto, M. 2019. High-throughput screening and characterization of a high-density soybean mutant library elucidate the biosynthesis pathway of triterpenoid saponins, *Plant and Cell Physiology*. 60(5), pp.1082-1097.

Krishnamurthy, P., Lee, J.M., Tsukamoto, C. Takahashi, Y., Singh, R.J., Lee, J.D., and Chung, G. 2014. Evaluation of genetic structure of Korean wild soybean (*Glycine soja*) based on saponin allele polymorphism. *Genetic Resources and Crop Evolution*. 61, pp.1121–1130.

Lee, J.D., Shannon, J., So, Y.S., Sleper, D., Randall, N., Lee, J.H., and Choung, M.G. 2008. Environmental effects on lutein content and relationship of lutein and other seed components in soybean. *Plant Breeding*. 128, pp.97-100.

Monma, M., Ito, M., Saito, M., and Chikuni, K. 1994. Carotenoid components in soybean seeds varying with seed color and maturation stage. *Bioscience, Biotechnology, and Biochemistry*. 58(5), pp.926-930.

Nakashima, H., Okubo, K., Honda, Y., Tamura, T., Matsuda, S., and Yamamoto, N. 1989. Inhibitory effect of glycosides like saponin from soybean on the infectivity of HIV *in vitro*. *AIDS*. 3, pp.655-658.

Panneerselvam, K., Tsukamoto, C., Honda, N., Kikuchi, A., Lee, J.-D., Yang, S.-H. and Chung, G. 2013. Saponin polymorphism in the Korean wild soybean (*Glycine soja* Sieb. and Zucc.). *Plant Breeding*. 132, pp.121-126.

Park, C., Dwiyanti, M.S., Nagano, A.J., Liu, B., Yamada, T., and Abe, J. 2019. Identification of quantitative trait loci for increased α-tocopherol biosynthesis in wild soybean using a high-density genetic map. *BMC Plant Biology*. 19(1), pp.510.

Pitargue, F., Kang, H., and Kil, D. 2019. Lutein-enriched egg production for laying hens. *World's Poultry Science Journal*, 75(4), pp. 633-645.

Ranard, K.M., Jeon, S., Mohn, E.S., Griffiths, J.C., Johnson, E.J., and Erdman Jr., J.W. 2017. Dietary guidance for lutein: consideration for intake recommendations is scientifically supported. *European Journal of Nutrition*. 56, pp.37-42.

Rani, A, Kumar, V., Verman, S.K., Shakya, A.K., and Chauhan, G.S. 2007. Tocopherol content and profile of soybean: genotypic variability and correlation studies. *Journal of the American Oil Chemists' Society*. 84, pp. 377-383.

Sakai, T., Kikuchi, A., Takada, Y., Kono, Y., Kunishi, I., Tezuka, M., Asao, H., and Shimada, S. 2002. Characteristics of soybean lines ‘‘Tohoku 151” and “Tohoku 152” which are lacking all lipoxygenases and modified saponin composition in seeds. *Tohoku Agricultural Research*. 55, pp. 59-60.

Sayama, T., Ono, E., Takagi, K., Yoshitake Takada, Y., Horikawa, M., Nakamoto, Y., Hirose, A., Sasama, H., Ohashi, M., Hasegawa, H., Terakawa, T., Kikuchi, A., Kato, S., Tatsuzaki, N., Tsukamoto, C., and Ishimoto, M. 2012. The *Sg-1* glycosyltransferase locus regulates structural diversity of triterpenoid saponins of soybean. *The Plant Cell*. 24(5), pp.2123-2138.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Jay J. Thelen, J.J, Cheng, J., Xu, D., Hellsten, U., May, G.D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M.K., Sandhu, D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S., Goodstein, D., Barry, K., Futrell-Griggs, M., Abernathy, B., Du, J., Tian, X., Zhu, L., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X, Shinozaki, K., Nguyen, H.T., Wing, R.A., Cregan, P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Randy C. Shoemaker, R.C. and Scott A. Jackson, S.A. 2010. Genome sequence of the palaeopolyploid soybean. *Nature*. 463, pp.178–183.

Seguin, P., Tremblay, G., Pageau, D., and Liu, W. 2010. Soybean tocopherol concentrations are affected by crop management. *Journal of Agricultural and Food Chemistry*. 58, pp. 5495–5501.

Seguin, P., Tremblay, G., Pageau, D., Liu, W. and Turcotte, P. 2011. Soybean lutein concentration: impact of crop management and genotypes. *Crop Science*. 51, pp.1151-1160.

Seguin, P., Chennupati, P., Tremblay, G. and Liu, W. 2014. Crop management, genotypes, and environmental factors affect soyasaponin B concentration in soybean. *Journal of Agricultural and Food Chemistry*. 62, pp.7160–7165.

Seppanen, C.M., Song, Q., and Csallany, A.S. 2010. The antioxidant functions of tocopherol and tocotrienol homologues in oils, fats, and food systems. *Journal of the American Oil Chemists' Society*. 87, pp.469-481.

Shao, A., and Hathcock, J.N. 2006. Risk assessment for the carotenoids, lutein and lycopene. *Regulatory Toxicology Pharmacology*. 45, pp.289-298.

Sheppard, A.J., Pennington, J.A.T., and Weihrauch, J.L. 1993. Analysis and distribution of vitamin E in vegetable oils and foods. In: Packer, L. and Fuchs, J. (eds) *Vitamin E in health and disease*. New York: Marcel Dekker Inc., pp.9-31.

Shiojiri, K., Ozawa, R., Yamashita, K., Uefune, M., Matsui, K., Tsukamoto, C., and Takabayashi, T. 2020. Exposure to artificially damaged goldenrod volatiles increases saponins in seeds of field-grown soybean plants. *Phytochemistry Letters*. 36, pp.7-10.

Sundaramoorthy, J., Park, G., Mukaiyama, K., Tsukamoto, C., Chang, J.H., Jeong-Dong, L., Kim, J.H., Seo, H., and Song, J. 2018. Molecular elucidation of a new allelic variation at the *Sg-5* gene associated with the absence of group A saponins in wild soybean. *PLOS ONE*. 13, p. e0192150.

Takada, Y., Sasama, H., Sayama, T., Kikuchi, A., Kato, S., Ishimoto, M., and Tsukamoto, C. 2013. Genetic and chemical analysis of a key biosynthetic step for soyasapogenol A, an aglycone of group A saponins that influence soymilk flavour. *Theoretical and Applied Genet*ics. 126, pp. 721-731.

Takahashi, Y., Li, X.H., Tsukamoto, C., and Wang, K.J. 2017. Categories and components of soyasaponin in the Chinese wild soybean (*Glycine soja*) genetic resource collection. *Genetic Resources and Crop Evolution*. 64, pp.2161-2171.

Takagi K, Yano R, Tochigi S Fujisawa, Y., Tsuchinaga, H., Takahashi, Y., Takada, Y., Kaga, A., Anai, T., Tsukamoto, C., Seki, H., Muranaka, T. and Ishimoto, M. 2018. Genetic and functional characterization of *Sg-4* glycosyltransferase involved in the formation of sugar chain structure at the C-3 position of soybean saponins. *Phytochemistry*. 156, pp.96-105.

Toti, E., Chen, C.O., Palmery, M., Villaño Valencia, D., and Peluso, I. 2018. Non-provitamin A and provitamin A carotenoids as immunomodulators: Recommended dietary allowance, therapeutic index, or personalized nutrition? *Oxidative Medicine and Cellular Longevity*. 2018, pp.4637861.

Tsai, C.Y., Chen, Y.H., Chien, Y.W., Huang, W.H., and Lin, S.H. 2010. Effect of soy saponin on the growth of human colon cancer cells. *World Journal of Gastroenterology*. 16(27), pp.3371–3376.

Tsukamoto, C., Shimada, S., Igita, K., Kudou, S., Kokubun, M., Okubo, K., and Kitamura, K. 1995. Factors affecting isoflavone content in soybean seeds: Changes in isoflavones, saponins, and composition of fatty acids at different temperatures during seed development. *Journal of Agriculture and Food Chemistry*. 43(5), pp.1184-1192.

Ujiie, A., Yamada, T., Fujimoto, K., Yasushi Endo, Y., and Kitamura, K. 2005. Identification of soybean varieties with high α-tocopherol content. *Breeding Science*. 55(2), pp.123-125.

Van Eenennaam, A. L., Lincoln, K., Durrett, T. P., Valentin, H. E., Shewmaker, C. K., Thorne, G. M., Jiang, J., Baszis, S. R., Levering, C. K., Aasen, E. D., Hao, M., Stein, J. C., Norris, S. R., and Last, R. L. 2003. Engineering vitamin E content: from Arabidopsis mutant to soy oil. *The Plant Cell*. 15(12), pp.3007–3019.

Xie, M., Chung, C.Y., Li, M., Wong, F., Wang, X., Liu, A., Wang, Z., Leung, A.K., Wong, T., Tong, S., Xiao, Z., Fan, K., Ng, Ming., Qi, X., Yang, L., Deng, T., He, L., Chen, L., Fu, A., Ding, Q., He, J., Chung, G., Isobe, S., Tanabata, T., Valliyodan, B., Nguyen, H.T., Cannon, S.B., Foyer, C.H., Chan, T., and Lam, H. 2019. A reference-grade wild soybean genome. *Nature Communications.*10, pp.1216.

Yang, S.H., Ahn, E.K., Lee, J.A., Shin, T.S., Tsukamoto, C., Suh, W., Itabashi, M., and Chung, G. 2015. Soyasaponins Aa and Ab exert an anti-obesity effect in 3T3-L1 adipocytes through downregulation of PPARc. *Phytotherapy Research*. 29, pp.281-287.

Yano, R., Kyoko Takagi, K., Tochigi, S., Fujisawa, Y., Nomura, Y., Tsuchinaga, H., Takahashi, Y., Takada, Y., Kaga, A., Anai, T., Tsukamoto, C., Seki, H., Muranaka, T., and Ishimoto, M. 2018. Isolation and characterization of the soybean *Sg-3* gene that is involved in genetic variation in sugar chain composition at the C-3 position in soyasaponins. *Plant and Cell Physiology*. 59(4), pp.797-810.

Zhan, Y., Li, H., Sui, M., Zhao, X., Jing, Y., Luo, J., Teng, W., Qiu, L., Zheng, H., Li, W., Yang, D., Han, Y. 2020. Genome wide association mapping for tocopherol concentration in soybean seeds across multiple environments. *Industrial Crops & Products*. 154, pp. 112674.