# Molecular Genetic Study on Seed Coat Cracking in Soybean

**January 2017**

**Auchithya Chathurani DISSANAYAKA**

# Molecular Genetic Study on Seed Coat Cracking in Soybean

A Dissertation submitted to

the Graduate School of Life and Environmental Sciences,

the University of Tsukuba

in Partial Fulfilment of the Requirements

for the Degree of Doctor in Philosophy in Agricultural Science

(Doctoral Program in Advanced Agricultural Technology and Sciences)

**Auchithya Chathurani DISSANAYAKA**

# Table of contents

List of abbreviations ..... iii

List of tables ..... iv

List of figures ..... vi

**Chapter 1. Introduction ..... 1**

1.1 Uses of soybean and its quality requirement ..... 1

1.2 Seed coat quality in soybean ..... 3

1.3 Seed coat cracking in soybean ..... 5

1.4 Study objectives ..... 9

**Chapter 2. QTL mapping of maturity gene *E5* affecting low-temperature induced seed coat cracking ..... 11**

2.1 Background ..... 11

2.2 Materials and methods ..... 14

2.3 Results ..... 19

2.4 Discussion ..... 24

**Chapter 3. Molecular genetic studies on net-like cracking ..... 50**

3.1 Background ..... 50

3.2 Materials and methods ..... 52

3.3 Results ..... 55

3.4 Discussion ..... 58

**Chapter 4. Molecular genetic studies on stress induced seed coat cracking.....71**

4.1 Background.....71

4.2 Materials and methods.....73

4.3 Results .....77

4.4 Discussion.....80

**Chapter 5. General discussion.....96**

**Summary .....102**

**Acknowledgements.....104**

**References.....105**

# List of abbreviations

| | |
|---|---|
| ABA | Abscisic acid |
| ACI | Average cracking index |
| ALP | Amplicon length polymorphism |
| BC | Back-crossing |
| cDNA | Complementary DNA |
| CTAB | Cetyl trimethylammonium bromide |
| dCAPs | Derived cleaved amplified polymorphic sequence |
| DNA | Deoxyribonucleic acid |
| LOD | Logarithm of the odds |
| MLG | Molecular linkage group |
| NILs | Near isogenic lines |
| PCR | Polymerase chain reaction |
| QTL | Quantitative trait loci |
| RHLs | Residual heterozygous lines |
| RILs | Recombinant inbred lines |
| RNA | Ribonucleic acid |
| SSR | Simple sequence repeat |
| USDA | United States Department of Agriculture |

# List of Tables

1. Frequency of browned or cracked seeds in Harosoy and its NILs under control and chilling treatments at Tsukuba, Japan in 1998 .....29

2. Soybean material used (Previous study).....30

3. $F_2$ populations used (Previous study) .....31

4. Soybean material used in developing mapping populations .....32

5. $F_2$ populations used in the study.....33

6. Sequences of DNA markers used for identification of alleles at loci *E1-E4* .....34

7. Linkage groups obtained from $F_2$ populations .....35

8. QTLs observed in two $F_2$ populations.....36

9. Alleles at maturity loci *E1-E4* in parental lines.....37

10. QTLs observed in $F_2$ populations with diagnostic markers.....38

11. Genotype of Harosoy-*E5* at SSR markers compared to Harosoy and PI 80837 .39

12. QTLs responsible for average cracking index in $F_2$ population (Previous study) .....60

13. QTLs associated with ACI in RIL $F_6$ and $F_7$ populations .....61

14. Genotype and ACI of the progeny of RHLs of RIL 10-2.....62

15. List of genes located between BARCSOYSSR_04_1292 and BARCSOYSSR_04_1308 .....64

16. QTLs responsible for average cracking index (*cr1* and *cr2*) (Previous study) ...82

17. ACI and other characters evaluated in $F_8$ NILs of *cr1* (Previous study).....83

18. Mean ACI recorded in 2013 (field) and 2014 (pot) experiments with parents and NILs of *cr1* .....84

19. Genotypes and ACI obtained in the fine mapping population .....85

20. List of genes located between BARCSOYSSR_02_0277 and BARCSOYSSR_02_0310 .....87

21. Selected genes that over-expressed in the seed coats of susceptible NIL 86-26-22-10-18 .....89

22. Selected genes that over-expressed in the seed coats of tolerant NIL 86-26-3-4-2 .....90

23. Genes differentially expressed between tolerant and susceptible NILs that located between BARCSOYSSR_02_0277 and BARCSOYSSR_02_0310 .....91

# List of Figures

1. Linkage groups containing QTLs for days to flowering and maturity (Previous study) ..... 40

2. Frequency distribution of days to flowering in Harosoy-*E5Dt2* x Clark-*e2* $F_2$ population ..... 41

3. Frequency distribution of days to flowering and maturity in Harosoy x PI 80837 $F_2$ population ..... 42

4. Linkage map of Harosoy-*E5Dt2* x Clark-*e2* $F_2$ population ..... 43

5. Linkage map of Harosoy x PI 80837 $F_2$ population ..... 44

6. MLG B1 (chr 11) containing QTL qDF 03 for days to flowering in $F_2$ population of Harosoy-*E5Dt2* x Clark-*e2* ..... 45

7. MLG C2 (chr 6) containing QTLs qDF 04 and qDM 03 for days to flowering and maturity in $F_2$ population of Harosoy x PI 80837 ..... 46

8. MLG L (chr 19) containing QTLs qDF 05 for days to flowering in $F_2$ population of Harosoy x PI 80837 ..... 47

9. Results of ALP/dCAPs analysis used to identify alleles at the *E1*, *E2*, *E3* and *E4* loci in parental lines ..... 48

10. Linkage groups containing QTLs for days to flowering and maturity with diagnostic markers ..... 49

11. Linkage group containing QTLs responsible for average cracking index in $F_2$ population (Previous study) ..... 66

12. Cracking scale for scoring intensity of net-like cracking ..... 67

13. MLG C1 (chr 4) containing QTLs responsible for ACI in RIL $F_6/F_7$ populations .....68

14. Comparison of ACI and genotype at BARCSOYSSR_04_1296 in RIL 10-2 and RIL 10-6.....69

15. Genotypes at each loci with higher and lower ACI.....70

16. Linkage groups containing QTLs responsible for ACI (Previous study).....92

17. Cracking scale for scoring intensity of pod removal cracking .....93

18. Steps followed in developing the mapping population .....94

19. Genotypes at each loci and corresponding ACI .....95

20. Soybean MLGs with genes/QTLs useful in developing tolerance to seed coat cracking.....100

21. Implications on physiological mechanisms leading to seed coat cracking .....101

# Chapter 1

# Introduction

## 1.1. Uses of soybean and its quality requirements

### World scenario

Soybean (*Glycine max* (L.) Merr) plays an important role in global economy as the world's leading oilseed and legume crop. It is the primary source of protein feed supplement for livestock. Soybean is renowned as "King of Beans" due to its higher protein content covering a more complete range of essential amino acids compared to other food legumes.

Around 85% of the world's soybean production is used to produce soy meal and oil (Brown-Lima *et al*, https://www.nature.org/ourinitiatives/regions/southamerica/brazil/explore/brazil-china-soybean-trade.pdf). Soybean oil is wide spread in consumption as an edible oil due to its neutral flavour and stability in both dehydrogenated and hydrogenated forms. It accounts for about 25% of total global oil and fat consumption (Thoenes, 2006). Soy meal accounts for about 60% of total world meal production (Thoenes, 2006).

The traditional soy foods derived from fresh beans play an important role in the diet of Far East. However, only about 6% of the world soybean production is used for human consumption (Brown-Lima *et al*, https://www.nature.org/ourinitiatives/regions/southamerica/brazil/explore/brazil-china-soybean-trade.pdf). The soy foods are considered very nutritious and they are rich in protein, fat, carbohydrates, dietary fibre, minerals and isoflavones. The recent experiments have revealed the health benefits of

isoflavones in reducing the risk of heart diseases, osteoporosis and hormone dependent cancers (Thoenes, 2006). Therefore in recent years, soy food consumption has increased in non-traditional countries. Furthermore, soy foods are a healthy food substitute for vegetarians and important in food safety of developing countries.

A small fraction of soybeans is used in manufacturing industry to produce biodiesel, solvents, lubricants *etc*.

**Types of soybeans based on the end use**

Based on the end use of soybeans, two types have been identified. The two types are oil beans and food beans. In soybean grading and marketing, different specifications are used for oil beans and food beans.

Minimum test weight per bushel, maximum percent limit of damaged kernels, foreign material, splits and seed coat colours other than yellow and maximum count limits of other material are the main factors established by USDA in grading oil beans (Islas-Rubio and Higuera-Ciapara, 2002).

The quality requirements for food beans depend on the type of soy food. However, food beans require superior seed grades and aesthetic appearance plays a major role. Lighter seed coat, clear hilum and higher protein content with lower oil content are some of the common quality requirements of food beans (Islas-Rubio and Higuera-Ciapara, 2002).

**Use of food soybean in Japan**

In contrast to the world scenario, in Japan soybeans are mainly used for human consumption. In 2011, 66% of soybeans were used as edible oils and followed by food beans (30%) and then animal feed (4%) (U.S.A. Soybean export council, 2012). The food beans are either used as whole seeds or processed into soy foods such as tofu, miso, soymilk *etc*.

Japan demands higher quality food beans and all the food beans used in the market are non-GMO. The quality of traditional soy foods is significantly affected by not only the nutritional and chemical composition of seed but also by the properties and appearance of seed coat. The acceptability of the traditional soy foods is determined by the aesthetic standards such as size of the seed, seed coat colour, texture and uniformity of seed coat. Because of the demand for higher quality beans, Japanese soy food processers prefer domestic soybeans and beans from USA (U.S.A. Soybean export council, 2012). Therefore, Japanese soybean research focuses more on studying the quality related traits important for soy food industry.

## 1.2 Seed coat quality in soybean

**Anatomy of soybean seed coat**

Soybean seed coat derives largely from the integument of maternal tissues after fertilization. It has specialized areas as hilum, micropyle, and raphe. Soybean seed coat consists of three distinct layers as epidermis, hypodermis and inner parenchyma layer.

The outer epidermal layer consists of closely packed thick walled palisade cells. The hypodermal layer consists of a single layer of cells having unevenly thickened cell walls resulting considerable intercellular space. The inner parenchyma layer is composed of 6-8 layers of thin-walled flattened cells (Carlson and Lestern, 2004).

Soybean seed coats vary in colour and colour is genetically controlled. While the most common colour is yellow; green, brown and black soybeans can be seen. Sometimes, there are bicolour and variegated germplasm.

**Factors affecting seed coat quality of soybean**

There are several internal and external factors that contribute to poor seed coat quality resulting poor marketability.

Seed coat cracking is one of the common quality deteriorating factors in soybean. Scientists have identified different types of seed coat cracking occurred in soybean that is controlled genetically. Certain types are induced by the environment and other stress factors. These types of seed coat cracking have been extensively studied over the years. However, the actual mechanisms of the genetic control are yet to be identified.

Growth marks also contribute to the poor appearance of seed coats. It is believed that this condition is a result of faster development of embryo compared to seed coat.

Among external factors, fungi are one of the common constraints resulting mouldy seeds and purple stain seeds. Seed coat mottling or hilum bleeding is another biotic constraint caused by viruses. Soybean seeds are viable to mechanical damages during

harvesting, handling and storage resulting chipped, cracked and broken seeds. Environmental factors such as early frost can cause green colouring in seed coats.

### 1.3 Seed coat cracking in soybean

Seed coat cracks have many adverse effects on commercial value of soybean. It reduces the aesthetic appeal of soybean and reduces the commercial value. Furthermore, under severe cracking conditions soybean lose their usability as food beans. Seed coat cracking reduces the longevity in storage and also increases the susceptibility to pathogen infections like fungi.

Seed coat cracking is a result of the separation of epidermal and hypodermal tissues from the inner layers of seed coat (Wolf *et al*, 1981; Yalkichi and Barla-Szabo, 1993). It can be caused by low temperature at flowering (Takahashi, 1997), adverse environmental stresses such as hot and dry weather during maturity, alternate wetting and drying of mature seeds, mechanical impact during harvesting and handling and genetics of the cultivars (Wolf *et al*, 1981).

Light and electron microscopy studies done by Wolf *et al* (1981) and Yalkichi and Barla-Szabo (1993) have revealed that the initial separation occurs along a plane perpendicular to the seed coat and further lateral movement occurs between hourglass cells of hypodermis and the underlying cellular layers. Furthermore, Yalkichi and Barla-Szabo (1993) observed that in many instances, the initial minute cracks began parallel to the long axis of the seed coat around hilum region. They found that cracking begins at approximately R7 growth stage after seeds achieved their maximum weight and just before the physiological maturity.

## Types of seed coat cracking found in soybean

Scientists have conducted studies to find out the causes and factors affecting seed coat cracking from the first quarter of last century and have been able to identify different types of seed coat cracking that control genetically. Some of these types are affected by environment and show varietal differences in tolerance.

## Genetically-controlled seed coat cracking that occurs regardless of environments

There are two types of seed coat cracking occur in some cultivars regardless of the environment they are growing. They are Type I with irregular cracks and Type II with net-like cracks (Liu, 1949). *I* and *T* loci which are responsible for seed coat colour and pubescence colour control Type I cracking (Stewart and Wentz, 1930). Soybeans with double recessive alleles at *I* and *T* loci produce severe cracking while other allele combinations show no cracking (Nicholas *et al*, 1993). However, Type II or net-like cracking is not influenced by genotypes at *I* and *T* loci (Liu, 1949). Oyoo *et al* (2010b) identified two quantitative trait loci (QTLs) at molecular linkage group (MLG) C1 associated with net-like cracking and further studies are needed to identify the genes associated with net-like cracking.

## Seed coat browning and cracking induced by chilling temperatures during flowering

Soybeans grown in temperate zone are liable to chilling stress ($10^{\circ}$C-$18^{\circ}$C) at different stages of their life cycle resulting in significant damage to the crop. Some of the damages depend on the growth stage of the crop at the time of exposure and sometimes

the damage is cultivar specific. Cultivars with yellow hilum and grey pubescence have shown browning around the hilum region and seed coat cracking, when the plants exposed to chilling temperatures during flowering (Sunada and Ito, 1982; Takahashi and Asanuma, 1996; Takahashi, 1997; Takahashi and Abe, 1999). Takahashi (1997) observed that the degree of browning and cracking is dependent on developmental stages of flowers and increases with the age of flowers at the time of chilling stress. Furthermore, Takahashi and Abe (1994) identified one or two major genes are involved in browning and cracking and one of the genes is linked with a dominant gene for late maturity. In an evaluation of the effect of *T* and *I* genes, it is revealed that the dominant *T* allele completely suppress browning and the dominant *I* allele partly suppress seed coat cracking under chilling stress (Takahashi and Asanuma, 1996; Takahashi, 1997).

**Association of soybean maturity genes with low-temperature induced seed coat browning and cracking**

In soybean *E* series of maturity genes; *E1* to *E9* and locus *J* control time of flowering and maturity (Bernard, 1971; Buzzell, 1971; Buzzell and Voldeng, 1980; McBlain and Bernard, 1987; Cober and Voldeng, 2001; Cober *et al,* 2010 and Kong *et al*, 2014). As mentioned earlier, Takahashi and Abe (1994) found that a maturity gene is associated with low-temperature induced browning and cracking. Takahashi and Abe (1999) conducted a study to evaluate the effect of maturity genes *E1* to *E5* on intensity of seed coat browning and cracking under chilling temperatures using cultivar Harosoy and its NILs for *E1* to *E5*. They revealed that intensity of seed coat browning was not affected by *e3*, slightly reduced by *E2* and *e4* and effectively reduced by dominant alleles of *E1*

and *E5*. Furthermore, they found that degree of cracking was slightly increased by *e3* and effectively reduced by *e4*, *E1* and *E5*. They concluded dominant alleles *E1* and *E5* are most effective in suppressing both browning and cracking. Therefore, these two genes are useful in developing tolerance to seed coat browning and cracking under chilling stress.

Molecular basis of maturity gene *E1* has already been uncovered. It is located on MLG C2 (Song *et al*, 2004). However, the molecular basis of *E5* locus is yet to be uncovered and the research on identification of the location of *E5* locus on molecular linkage maps will be useful to uncover the molecular basis and function of *E5* locus and thereby will help in improving seed coat quality under chilling stress.

**Seed coat cracking observed under field condition without chilling stress**

Researchers have frequently observed seed coat cracking under field conditions even at the absence of chilling stress. So far, the causes of this cracking have not been identified. However, genotypic differences have been observed with respect to the tolerance to this type of seed coat cracking. In Japan, different treatments are being used in research to reproduce the genotypic differences observed under field conditions. Those treatments are pod removal (Maruyama and Mikoshiba, 1976; Sasaki and Nakamura, 1981), drying of imbibed seeds (Murata *et al*, 1991) and application of ethylene generating reagents (Figaron, Nissan Chemical Industry Inc.,Tokyo) (Okabe, 1996). Out of these three methods pod removal treatment have shown reproducible results over the years and high heritability estimates of seed coat cracking (Okabe, 1996). Oyoo *et al* (2010a) identified two QTLs *cr1* (MLG D1b) and *cr2* (MLG M) associated with seed coat

cracking under field condition using pod removal technique. However, further studies are needed to identify the genes affecting stress induced cracking.

The knowledge on genes or QTLs controlling seed coat cracking is important in breeding soybeans tolerant to seed coat cracking. Furthermore, such knowledge helps to identify the physiological mechanisms leading to seed coat cracking. QTL mapping and fine mapping techniques are useful tools in identifying the location of genes or QTLs in MLGs and subsequent marker assisted breeding. This study aims to identify the genes or QTLs associated with low-temperature induced cracking, net-like cracking and pod removal cracking using QTL mapping and fine mapping.

### 1.4 Study objectives

This research consists of three experiments which aim the identification of molecular basis of three factors affecting seed coat cracking.

1. In Japan, Hokkaido is the major soybean producing area. However, due to its high latitude, chilling temperatures frequently prevail in Hokkaido during soybean growing season. As previously explained, chilling temperatures during flowering induce seed coat cracking and browning in yellow hilum cultivars. Yellow hilum cultivars are preferred over brown hilum cultivars as a food bean in Japan. As discussed earlier in this chapter, researchers have found genotypic variations among yellow hilum cultivars with respect to tolerance to browning and cracking and this tolerance is associated with maturity loci *E1* and *E5*. Although the location and functions of *E1* is already known, very little information is available on maturity locus *E5*. Therefore, the first objective of the current research is

quantitative trait locus (QTL) mapping of maturity locus *E5* to identify its location in MLGs and assist breeding programmes for tolerance to low-temperature induced browning and cracking.

2. Type II or net-like cracking is a type of seed coat cracking that occurs in some cultivars regardless of the environment they are growing (Liu, 1949). Although, the genes controlling type I cracking has been identified, the genetic control of net-like cracking is yet to be identified. Oyoo *et al* (2010 b) identified two QTLs affecting net-like seed coat cracking in MLG C1 (*ncr1* and *ncr2*). Therefore, the second research objective is confirmation of QTLs affecting net-like cracking and narrow-down the QTL region by fine mapping.

3. Pod removal is a technique used to reproduce the genotypic differences among cultivars to seed coat cracking occurs under field condition. This method has been widely used in evaluating tolerant cultivars. Two QTLs associated with this type of seed coat cracking have been identified using pod removal technique (Oyoo *et al*, 2010a). One QTL was identified in MLG D1b (*cr1*) and the other in MLG M (*cr2*). In this study the association of QTL *cr1* with seed coat cracking will be confirmed using pod removal technique and the QTL region will be narrowed-down by fine mapping.

# Chapter 2

# QTL mapping of soybean maturity gene *E5* affecting low-temperature induced seed coat cracking

## 2.1 Background

Soybeans growing in high latitude and altitude areas are sensitive to chilling stress. Sunada and Ito (1982) reported that chilling stress during flowering induce seed coat browning and cracking in soybean. Furthermore, they reported that these defects were found only in yellow hilum cultivars.

However, yellow hilum cultivars are superior in protein content and in aesthetic appearance over brown hilum cultivars. Therefore, yellow hilum cultivars are preferred as food beans in Japan and majority of the cultivars used in Japan are yellow hilum cultivars (Takahashi and Asanuma, 1996). Therefore, several researches have been carried out to identify the association between low-temperature induced seed coat browning and cracking, and the locus *T* and *I* which control pubescence and hilum colour (Takahashi and Asanuma, 1996; Takahashi, 1997). Takahashi and Asanuma (1996) and Takahashi (1997) found that the locus *T* completely suppressed seed coat browning and partly suppressed seed coat cracking independent of the genotypes at *I* locus. Furthermore, they observed that under the genotype *t/t*, dominant *I* allele also suppressed seed coat browning and cracking but the effect was comparatively less. However, the dominant *T* alleles results tawny pubescence and in Japan the tawny pubescence and yellow hilum is not a preferred combination because, allelic combination *II TT* darkens the entire seed

coat. Therefore, it is important to identify the genetic variation in tolerance within yellow hilum and grey pubescence cultivars.

Takahashi and Abe (1994) observed genetic variation in tolerance to browning and cracking among cultivars with yellow hilum and grey pubescence. Furthermore, they found that the tolerance was closely associated with a dominant gene for late maturity.

Eight loci belong to *E* series of maturity genes control flowering in soybean. Takahashi and Abe (1999) carried out an experiment to study the relationship between maturity genes and low-temperature induced seed coat browning and cracking. They used Harosoy isolines for maturity genes *E1* to *E5* for analysis and all those lines used had yellow hilum and grey pubescence. This study revealed that dominant alleles *E1* and *E5* were effective in suppressing both seed coat pigmentation and cracking compared to other alleles tested (Table 01). Therefore, they suggested the use of maturity loci *E1* and *E5* in developing yellow hilum cultivars with tolerance to low-temperature induced seed coat browning and cracking.

Maturity locus *E1* has been assigned in MLG C2 (Cregan *et al,* 1999). *E1* locus has four alleles; dominant *E1*, *e1-as* with a single nucleotide polymorphism leading to missense mutation, *e1-fs* with a single base deletion leading to premature stop codon and *e1-nl* null allele where entire *E1* gene was deleted (Xia *et al*, 2012). Molecular characterization of *E1* locus has revealed *E1* protein contains a putative bipartite nuclear localization signal and a region distantly related to B3 domain suggesting a probable role as a transcription factor (Xia *et al*, 2012). Furthermore, Xia *et al* (2012) observed a negative correlation of transcript abundance of *E1* with *GmFT2a* and *GmFT5a*

homologues of *FLOWERING LOCUS T* demonstrating the key role of *E1* in repressing flowering and delaying maturity.

Maturity locus *E5* is known to scientists since 1987 (McBlaine and Bernard, 1987). It was first identified in a backcrossed population of Harosoy (6) x PI 80837 which was made to transfer dense pubescence (*Pd1*) into Harosoy (McBlaine and Bernard, 1987). When they observed segregation for late maturity in the above population, a uniformly late BC5F4 plant progeny (named as L64-4830) had been selected for further studies on late maturity. The $F_2$ population made with L64-4830 x Harosoy had shown segregation for a single allele for lateness. Test crosses of L64-4830 with Harosoy maturity allele isolines that have single maturity allele substitutions of *E1*, *E2* and *e3* have shown that the observed maturity allele was not any of those loci. Because of the facts that *e4* conditions photoperiod insensitivity and both Harosoy and L64-4830 were photoperiod sensitive they presumed both Harosoy and L64-4830 have the *E4* allele. Therefore, they designated the new locus as *E5*; the allele in L64-4830 as *E5* and the allele in Harosoy as *e5*. However, the molecular identity of *E5* locus is unknown. The knowledge on molecular identity of *E5* locus is important in developing cultivars tolerant to low-temperature induced browning and cracking. Therefore, objective of this study is to identify the quantitative trait loci (QTLs) correspond to the maturity locus *E5* using molecular markers.

**Previous work on QTL mapping of *E*5**

Two $F_2$ populations expected to segregate for *E5* have been used to identify QTLs correspond to *E5* (Rodriguez Torrico, 2014). The first population had been derived from

a cross between Harosoy-*E5* and Clark-*e2*. The second population had been derived from a cross between Harosoy and Clark-*e2E5*. Details of the plant material used and the details of $F_2$ populations are shown in Table 02 and Table 03.

The linkage mapping and QTL analysis of the $F_2$ population made with Harosoy-*E5* x Clark-*e2* have revealed two QTLs with larger effects; qDF 01 for days to flowering and qDM 01 for days to maturity. Both these QTLs have been found in the MLG O (chr 10) between molecular markers Satt581 and Satt153 (Figure 01) where the maturity gene *E2* is assigned.

In Harasoy x Clark-*e2E5* population a QTL with a LOD score of 5.71 has been found (qDF 02) for days to flowering in the MLG D1a (chr 1) between Satt198 and Satt077. Furthermore, a QTL for days to maturity (qDM 02) has been found between Satt509 and Sat_247 in the MLG B1 (chr 11) (Rodriguez Torrico, 2014).

These two populations were expected to segregate for the same locus however; the four QTLs found were not consistence. Therefore, no candidate QTL correspond to *E5* was found with these two populations.

## 2.2 Materials and methods

### QTL mapping using $F_2$ populations

### Plant materials

Canadian cultivar Harosoy, a NIL with genetic background of Harosoy (Harosoy-*E5Dt2)*, a NIL with genetic background of US cultivar Clark and Japanese landrace PI

80837 were used to develop two $F_2$ populations expected to segregate for maturity gene *E*5. The first $F_2$ population was developed by crossing Harosoy-*E5Dt2* with pollen from Clark-*e2*. The second population was developed by crossing Harosoy with pollen from PI 80837. Harosoy x PI 80837 cross is identical with the cross in which *E*5 was originally identified (McBlain and Bernard, 1987). Seeds of the plant material were obtained from the USDA Soybean Germplasm Collections. The Harosoy and Clark NILs have been developed by crossing the cultivars with lines having the respective alleles and backcrossing the progeny up to BC5 (Bernard *et al*, 1991). The genotype at maturity genes of each plant material is shown in Table 04. Flowers of Harosoy and Harosoy-*E*5*Dt*2 were emasculated one day before flower opening and then were pollinated either with pollen of Clark-*e2* or PI 80837. As shown in Table 04, Harosoy-*E5Dt2* had recessive allele (*t*) for pubescence colour while Clark-*e2* had dominant allele (*T*) for pubescence colour. Therefore in Harosoy-*E5Dt2* x Clark-*e2* population, hybridity of $F_1$ plants was assessed using pubescence colour. Harosoy had recessive allele (*pd1*) for pubescence density while PI 80837 had dominant allele (*Pd1*) for pubescence density. Therefore in Harosoy x PI 80837 population, hybridity of $F_1$ plants was assessed using pubescence density.

**Plant cultivation and phenotypic data collection**

Harosoy-*E*5*Dt*2 x Clark-*e2* population was planted in the experimental field at Hokkaido University, Sapporo, Japan (43°25′N, 143°32′E) from May to October, 2011 (Table 05). Seeds were sown on paper pots and 10 day old seedlings were transplanted into the field. Plants were spaced 25 cm within row and 60 cm between rows. Fertilizers

were applied at N, P, K ratio of 3.0, 4.4 and 8.3 $gm^{-2}$. Number of days from planting to opening the first flower (R1) was recorded for individual plants.

Harosoy x PI 80837 population was grown in an unheated vinyl plastic green house at the Institute of Crop Science, NARO, Tsukuba, Japan (36°06′N, 140°05′E) (Table 05). Pots (12.5 cm diameter) were prepared by filling 2.5 kg soil (low-humic andosols) supplemented with 0.8 g ammonium sulphate, 1.6 g monocalcium phosphate, 3.2 g fused magnesium phosphate and 0.8 g potassium sulphate. Three $F_2$ seeds and 5 seeds from each parental line were sown per pot. One week after the seedling emergence, seedlings were thinned to one plant per pot. Pots were positioned randomly inside the green house and repositioned twice a week. Number of days from planting to opening the first flower (R1) and days from planting to maturity (R8:95% of pods showing mature colour) were recorded for individual plant.

**DNA extraction and SSR analysis**

Total DNA was extracted individually from trifoliate leaves of the parents and $F_2$ plants based on the CTAB method (Murray and Thompson, 1980). Since the PCR reaction plates and the electrophoresis apparatus were designed for multiples of 96 samples, 94 $F_2$ plants from each population were selected randomly together with parents for SSR analysis. SSR markers developed by USDA (Song *et al*, 2004) and Kazusa DNA Research Institute (Hisano *et al*, 2007) were used in the analysis. The total volume of PCR mixture was 5 µl containing 20 ng of genomic DNA, 2.25 pmol of primer, 625 pmol of nucleotides and 0.125 units of ExTaq in 1 x ExTaq buffer supplied by the manufacturer (Takara Bio, Ohtsu, Japan). PCR was performed in an Applied Biosystems 9700 thermal cycler

(Applied Biosystems, Foster City, CA). An initial 4 minutes denaturation at 95°C was followed by 35 cycles of 1 minute denaturation at 95°C, 1 minute annealing at 49°C and 1 minute extension at 68°C. PCR products were separated in 8% non-denaturing polyacrylamide gels and the fragments were visualized by staining with ethidium bromide.

**Linkage mapping and QTL analysis**

The linkage maps were constructed with the maximum likelihood function (threshold LOD score=3) using MAPMAKER/EXP. Ver. 3.0 (Lander *et al*, 1987). Designation of MLGs followed Cregan *et al* (1999). QTL analysis was performed by composite interval mapping (Zeng, 1993) using QTL Cartographer version 2.5 (Wang *et al*, 2007). The threshold score was determined by permutation test with 1000 repetitions corresponding to genomic-wide 5% level of significance.

**Genotyping parental lines for maturity loci *E1* to *E4***

Total DNA was extracted individually from trifoliate leaves of the parental lines listed in Table 02 and Table 04 and from reference cultivars having specific maturity alleles. Reference cultivars used were Bay for *E1* and *E2-dl* alleles, Tachinagaha for *e3-tr* allele, Moshidou Gong 503 for *e3-Mo* allele and Tokei 780 for *e4-SORE-1* allele (Liu *et al*, 2008; Tsubokura *et al*, 2014). Alleles at each locus, *E1* to *E4* were determined by ALP or dCAPS analyses following previous reports (Liu *et al*, 2008; Tsubokura *et al*, 2014).

Alleles at locus *E1* (*E1* from *e1-as*) were determined by *Taqα*I restriction digestion of the PCR product generated with specific primers G33snpTaqcutF and G33snpTaqcutR (Marker E1_*Taqα*I). The allele *e2-ns* was discriminated from other alleles at *E2* locus by *Dra*I digestion of the PCR product obtained with PCR primers SoyGI_dCAPaMs19300FW and SoyGI_dCAPa19440RV (Marker E2_*Dra*I). *E2-in* and *E2-dl* alleles at *E2* locus were discriminated based on the PCR amplicon size obtained using PCR primers E2_15345FW and E2_15856RV (marker E2_InDel). Allele *e3-tr* was differentiated from other alleles at locus *E3* by the PCR amplicon size using mixed primers E3_08557FW, E3_09908RV, E3Ha_1000RV and e3tr_0716RV (marker E3_Mix). Alleles *E3-Ha* and *e3-Mo* at locus *E3* were determined by *Mse*I digestion of the PCR product of primers E3_08094FW and E3_08417RV (marker E3_*Mse*I). Alleles *E*4 and *e*4-*SORE-1* at locus *E4* were discriminated using the PCR amplicon size obtained with primers PhyA2-for, PhyA2-Rev/E4 and PhyA2-Rev/e4 (Marker E4_Mix). Primer sequence information are shown in Table 06.

PCR products of E2_InDel and E4_Mix were separated in agarose gels of 2% and 1% respectively. The other products were separated in 8% non-denaturing polyacrylamide gels.

**Genotyping $F_2$ populations using diagnostic markers**

$F_2$ population Harosoy-*E5* x Clark-*e2* (used in a previous QTL analysis) and Harosoy x PI 80837 $F_2$ populations were genotyped with diagnostic markers for *E2* and *E1/E3* respectively. Harosoy-*E5* x Clark-*e2* population was genotyped with marker E2_*Dra*I while Harosoy x PI 80837 population was genotyped with E1_*Taqα*I and

E3_Mix (Liu *et al*, 2008; Tsubokura *et al*, 2014). Linkage mapping and QTL analysis were repeated including genotypes with diagnostic markers.

**Evaluation of the out-crossing possibility in *E5* gene**

Five SSR markers were selected from each MLG and were used to genotype Harosoy, Harosoy-*E5* and PI 80837. The genotype of Harosoy-*E5* was compared with its recurrent parent Harosoy and donor parent PI 80837.

## 2-3 Results

**Distribution of days to flowering and maturity**

The frequency distribution of days to flowering (R1) of Harosoy-*E5Dt2* x Clark-*e2* $F_2$ population together with their parents is shown in Figure 02. The parent, Harosoy-*E5Dt2* took 65 days to initiate flowering at Sapporo while Clark-*e2* took only 61 days. The time taken by $F_2$ population to initiate flowering ranged from 57 days to 68 days with an average of 63 days. One $F_2$ plant flowered 4 days earlier than Clark-*e2* and three plants flowered 3 days later than Harosoy-*E5Dt2*. Therefore, transgressive segregation was observed in Harosoy-*E5Dt2* x Clark-*e2* $F_2$ population.

The frequency distribution of days to flowering (R1) and days to maturity (R8) of Harosoy x PI 80837 $F_2$ population together with their parents are shown in Figure 03. The parent, Harosoy initiated flowering 31 days after sowing and matured 94 days after sowing. PI 80837 initiated flowering 35 days after sowing and matured 99 days after

sowing the seeds. With respect to $F_2$ population, the time taken for flowering ranged from 25 days to 43 days and time taken for maturity ranged from 79 days to 116 days. One $F_2$ plant was observed to flower 6 days earlier than Harosoy and another plant flowered 8 days later than PI 80837. Furthermore, two $F_2$ plants matured 15 days earlier than Harosoy and 17 days later than PI 80837. These observations were clear evidences of transgressive segregation.

In both these populations, it was difficult to categorize the $F_2$ plants either into early or late categories for days to flowering or days to maturity. Therefore, linkage mapping and QTL analysis were performed for days to flowering and maturity.

**Linkage mapping**

In Harosoy-*E5Dt2* x Clark-*e2* population, 185 SSR markers showed polymorphism between parents and distinctly segregated in the population. Out of these 185 markers, 179 markers were linked to form 27 linkage groups. The total map length covered was 1982cM. In Harosoy x PI 80837 population, the number of markers showed polymorphism between parents and distinct segregation was 207. Out of these 207 markers 197 markers linked to compose 35 linkage groups (Table 07). Linkage maps obtained with two populations are shown in Figure 04 and 05 respectively.

**QTL analysis for days to flowering and days to maturity**

In Harosoy-*E5Dt2* x Clark-*e2* population a QTL for days to flowering (qDF 03) was found between Sat_149 and Sat_348 in the MLG B1 (chr 11). This locus had a LOD

score of 6.46 accounting for 25.5% of phenotypic variation. The allele from Harosoy-*E5Dt2* had increased the days to flowering at the QTL (Figure 06 and Table 08).

In Harosoy x PI 80837 population, QTLs with larger LOD scores were observed for both traits (days to flowering and days to maturity) between Sat_402 and Satt365 in MLG C2 (chr 6) where *E1* was repeatedly reported to locate. The QTL for flowering qDF 04 at MLG C2 had a LOD score of 29.9 accounting for 73% of phenotypic variation whereas, the QTL for maturity had a LOD score of 5.25 accounting for 46.9% of phenotypic variation. The alleles from PI 80837 had increased both the days to flowering and maturity at the QTLs (Figure 07 and Table 08). Another QTL for days to flowering qDF 05 was found in MLG L (chr 19) between Sat_286 and Satt513 closer to the location reported for locus *E3*. This QTL had a LOD score of 11.75 accounting for 16.6% of phenotypic variation. At this QTL allele from PI 80837 had decreased the days to flowering (Figure 08 and Table 08).

**Genotyping of parental lines for maturity loci *E*1 to *E*4**

The results are shown in Figure 09 and summarized in Table 09. PCR with G33snpTaqcutF and G33snpTaqcutR1 produced a fragment of 444 bp in all parental material and in reference cultivar Bay. The subsequent *Taqα*I digestion, resulted a 412 bp fragment in Harosoy, Harosoy-*E5*, Harosoy-*E5Dt2*, Clark, Clark-*e2* and Clark-*e2E5*. However, PCR products of PI 80837 and Bay remained undigested suggesting that PI 80837 and Bay had the dominant allele at *E1* locus while other parental lines had the recessive allele *e1-as*.

PCR primers SoyGI_dCAPaMs19300FW and SoyGI_dCAPa19440RV generated a 142 bp fragment in all the material used and subsequent *Dra*I digestion produced a 115 bp fragment in Harosoy, Harosoy-*E5Dt2*, Clark-*e2*, Clark-*e2E5*, and in PI 80837 whereas the products of Harosoy-*E5*, Clark and Bay were undigested. According to this observation, at *E2* locus Harosoy-*E5* had the dominant allele similar to Clark and reference cultivar Bay while other parental lines had the *e2-ns* allele. PCR with E2_15345FW and E2_15856RV (E2_InDel) produced a 512 bp fragment in all the lines except in Clark. In Clark a 544 bp fragment was observed suggesting *E2-in* allele at *E2* locus. These results suggest that in Harosoy-*E5* and Bay the alleles at *E2* locus were *E2-dl.*

PCR with primers of E3_Mix produces a 274bp fragment in PI 80837 and in reference cultivar Tachinagaha whereas a 558 bp fragment in other lines. Therefore, PI 80837 and Tachinagaha had *e3-tr allele* at locus *E3*. PCR with E3_08094FW and E3_08417RV produced a fragment of 324 bp in all material tested. However, with the subsequent *Mse*I digestion the digested 223 bp fragment was observed only in reference cultivar Moshidou Gong 503. Products of other material remained undigested suggesting only Moshidou Gong had *e3-Mo* allele at *E3* locus while Harosoy, Clark and their NILs had *E3-Ha* allele.

PCR with E4-Mix primers produced 1229 bp fragment in all parental material whereas they produced an 837 bp fragment in reference cultivar Tokei. Since Tokei has the allele *e4-SORE1* at locus *E4*, the results suggest that all the parental lines had the dominant allele *E4.*

**Genotyping $F_2$ populations using diagnostic markers**

Harosoy-*E5* x Clark-*e2* $F_2$ population was genotyped with primers of marker E2_*Dra*I and followed by *Dra*I restriction digestion. The genotypes obtained were used to repeat the linkage mapping and QTL analysis. In linkage mapping, 142 markers linked to form 36 linkage groups spanning 1524 cM map length. The repeated QTL analysis revealed that the two larger QTLs found, qDF 01 (for flowering) and qDM 01 (for maturity) located in the vicinity of marker E2_*Dar*I in MLG O (chr 10) (Table 10 and Figure 10). The locus qDF 01 recorded a LOD score of 26.44 accounting for 51.2% of phenotypic variation. The qDM 01 locus had a LOD score of 19.51 accounting for 29.5% of phenotypic variation. The alleles from Harosoy-*E5* had increased the days to flowering and days to maturity at the QTLs.

In the Harosoy x PI 80837 population, with the genotypes of E1_*Taqα*I and E3_Mix markers the same number of linkage groups (35) as previously were obtained with 199 linked markers spanning of 2555 cM map length.

The QTL analysis revealed that the two larger QTLs found, qDF 04 (for days to flowering) and qDM 03 (for days to maturity) located in the close proximity of marker E1_*Taqα*I in the MLG C2 (chr 06) (Table 10 and Figure 10). The qDF 04 locus had a LOD score of 37.15 accounting for 69% of phenotypic variation. The qDM 03 had a LOD score of 6.76 accounting for 56.2% of phenotypic variation. The alleles from PI 80837 had increased the days to flowering and days to maturity at the QTL. It was found that the QTL qDF 05 (for flowering) found in MLG L (chr 19) located in the vicinity of marker E3_Mix. Furthermore, a QTL for days to maturity qDM 04 was found in the same vicinity. The qDF 05 locus had a LOD score of 15.87 accounting for 19.5% of phenotypic variation

whereas qDM 04 had a LOD score of 4.47 accounting for 14.3% of phenotypic variation. The alleles from PI 80837 had decreased the days to flowering and days to maturity (Table 10 and Figure 10).

**Evaluation of the out-crossing possibility in *E5* gene**

Out of 100 markers used to compare the genotype of Harosoy-*E5* with its parents Harosoy and PI 80837, at 95 markers Harosoy-*E5* showed the same genotype as Harosoy. At marker Satt148 in MLG I (chr 20) Harosoy-*E5* showed the same genotype as PI 80837. However, at four markers (Sat_183 in MLG D1b, Satt049 in MLG I, Satt406 in MLG J and Satt592 in MLG O) Harosoy-*E5* showed genotypes different from both its parents Harosoy and PI 80837 suggesting an out-crossing occurred during the development of Harosoy-*E5* (Table 11).

## 2.4 Discussion

Two QTLs for flowering (qDF 01 and qDF 02) and two QTLs for maturity (qDM 01 and qDM 02) have been found in previous studies with Harosoy and Clark NILs (Rodriguez Torrico, 2014). However, the QTLs found were not consistent. Furthermore, the QTLs found with current $F_2$ populations were also not consistent. We repeated the QTL analysis for Harosoy-*E5* x Clark-*e2* with a diagnostic marker for maturity locus *E2* and found that qDF 01 and qDM 01 located in the close proximity of *E2* locus at MLG O. Furthermore, ALP/dCAPS analysis revealed that Harosoy-*E5* had *E2-dl* allele at locus *E2* whereas both of its parents Harosoy and PI 80837 had *e2-ns* allele. Analysis was

repeated with a new batch of seed samples from USDA but, the results were same. These results suggested that the QTLs we identified in Harosoy-*E5* x Clark-*e2* population might be corresponding to locus *E2*.

As mentioned earlier in this chapter, locus *E5* was first identified in a cross between Harosoy x PI 80837. Therefore, if a unique *E5* locus exists it should segregate in the Harosoy x PI 80837 $F_2$ population. However, the QTLs found in this population were located in close vicinity of loci *E1* and *E3*. ALP/dCAPs analysis revealed the probable genotypes of Harosoy and PI 80837 at maturity loci *E1-E4* as *e1 e2 E3 E4* and *E1 e2 e3 E4* respectively. Therefore, Harosoy x PI 80837 population can be segregated for both *E1* and *E3* loci. These results suggest that the QTLs qDF 04 and qDM 03 might be corresponding to maturity locus *E1*. This finding is further strengthened by higher LOD score of qDF 04 since; *E1* has recorded to have larger effect on flowering and maturity (Takahashi and Abe, 1999). The other two QTLs found in this population qDF 05 and qDM 04 might be corresponding to *E3*. Appearance of plants with double recessive and double dominant alleles of *E1* and *E3* might be the reason for higher transgressive segregation observed in $F_2$ plants. Therefore, no candidate QTL corresponding to maturity locus *E5* was found in Harosoy x PI 80837 $F_2$ population. Since, a candidate QTL was not found with Harosoy x PI 80837 population, QTLs found with other populations, qDF 02, qDF 03, qDM 02 may not be corresponding to locus *E5*. In addition, the LOD scores of these QTLs were small to consider as QTLs corresponding to *E5*.

The current study revealed that the genotypes of Harosoy-*E5* at maturity loci *E1* to *E4* is *e1-as E2-dl E3-Ha E4* whereas its parents Harosoy and PI 80837 is *e1-as e2-ns E3-Ha E4 and E1 e2-ns e3-tr E4* respectively. Therefore, Harosoy-*E5* may have been

generated by unexpected out-crossing with pollens having *E2-dl* allele. In the genotype comparison of Harosoy-*E5* with Harosoy and PI 80837, four out of hundred SSR markers showed different genotypes from Harosoy and PI 80837 revealing the possibility of out-crossing. Therefore, late maturity found in Harosoy-*E5* may be attributed to *E2-dl* allele. During the identification of *E5*, McBlain and Bernard (1987) conducted test crosses between Harosoy-*E5* (L64–4830) and Harosoy-*E2* (L74-21 or L74-27) to exclude the possibility of newly identified gene to be *E2*. However, the donor parent of L74-21 and L74-27 was Clark which has *E2-in* allele. Therefore, the transgressive late segregation observed by McBlaine and Bernard (1987) might be caused by the appearance of plants with heterozygous *E2-in E2-dl* genotypes in the progeny of test crosses.

Scientists have observed different responses of Harosoy-*E2* and Harosoy-*E5*. Takahashi and Abe (1999) have reported that in Tsukuba Harosoy-*E5* flowered three days later and matured nine days later than Harosoy-*E2.* The different effects of Harosoy-*E5* and Harosoy-*E2* are possibly caused by the allelic difference between *E2-in* and *E2-dl*. In addition to these two alleles, *e2-ns* a nonsense mutation is also found at locus *E2.* Tsubokura *et al* (2014) identified the *E2-dl* allele and they observed an insertion and a deletion of 36 bases in the eighth intron at *E2* locus in *E2-in* and *E2-dl* with each specific sequence in the 5'upstream region, exon, intron and 3' downstream region. Watanabe *et al* (2011) identified *E2* as an ortholog of *GIGANTEA* in *Arabidopsis.* GIGANTEA (GI) is a plant specific nuclear protein involves in various plant processes (Mishra and Panigrahi, 2015). It regulates circadian rhythm and flowering in plants and functions upstream of *CONSTANS (CO)* and *FLOWERING LOCUS (FT)* which encodes flowering promotion signal florigen (Samach *et al*, 2000; Abe *et al*, 2005).

Takahashi and Abe (1999) observed significantly lower incidence of low-temperature induced seed coat cracking and browning in Harosoy-*E5* compared to Harosoy. As shown in Table 01, the percentage reduction of cracked seeds and pigmented seeds in Harosoy-*E5* compared to Harosoy were approximately 40% and 60% respectively (Takahashi and Abe, 1999). Harosoy-*E5* and Harosoy have *E2-dl* and *e2-ns* alleles respectively at maturity locus *E2* (Table 09). Furthermore, according to the results of Takahashi and Abe (1999), Harosoy-*E5* is also effective in suppressing low temperature induced seed coat cracking and browning compared to Harosoy-*E2* which has *E2-in* allele at maturity locus *E2.* The percentage reduction of pigmentation and seed coat cracking in Harosoy-*E5* compared to Harosoy-*E2* were approximately 25% and 30% respectively (Table 01). Therefore, we concluded that the maturity allele that associated with low-temperature induced seed coat browning and cracking might be *E2-dl* allele at maturity locus *E2.* The *Arabidpsis* orthologues of locus *E2 GIGANTEA* have also been reported to function not only in flowering time regulation but also in diverse physiological processes including cold tolerance (Mishra and Panigrahi, 2015).

In ALP/dCAPs analysis, we found that Clark-*e2E5* and Harosoy-*E5Dt2* has *e2-ns* allele. Therefore, the differences we observed with regard to days to flowering and maturity between parents in Harosoy x Clark-*e2E5* population and Harosoy-*E5Dt2* x Clark- *e2* population might be a result of introduction of other genomic regions affecting time of flowering and maturity during the development of Harosoy-*E5Dt2* and Clark-*e2E5.*

Present study concludes the absence of a unique *E5* gene. Furthermore, we revealed that the late flowering Harosoy-*E5* has *E2-dl* allele at locus *E2* and may have been generated by unexpected out-crossing with pollen having *E2-dl* allele. In Harosoy-

*E5*, we observed a genotype different from its parents at Satt592 which is linked with *E2* in MLG O with a distance of 22.3 cM. Therefore, Harosoy-*E5* may have received a chromosomal fragment containing *E2-dl* and Satt592 from an unknown donor. A previous study revealed that Harosoy-*E5* is tolerant to low-temperature induced seed coat browning and cracking compared to Harosoy-*E2*. Hence, *E2-dl* allele may have attributed to higher tolerance of Harosoy-*E5* on low-temperature induced seed coat browning and cracking.

**Table 01. Frequency of browned or cracked seeds in Harosoy and its NILs under control and chilling treatments at Tsukuba, Japan in 1998**

(Takahashi and Abe, 1999)

| Line | Frequency of browned seeds (%) | | Frequency of cracked seeds (%) | |
|---|---|---|---|---|
| | Control | 15 $^0$C | Control | 15 $^0$C |
| Harosoy (*e1e2E3E4e5*) | 0.0 | 92.2 | 3.0 | 77.9 |
| Harosoy-*E1* | 0.0 | 42.6 | 0.0 | 17.3 |
| Harosoy-*E2* | 0.0 | 55.9 | 0.0 | 69.2 |
| Harosoy-*e3* | 0.0 | 86.6 | 5.7 | 92.1 |
| Harosoy-*e4* | 0.0 | 82.0 | 0.4 | 48.1 |
| Harosoy-*E5* | 0.0 | 29.4 | 0.4 | 38.9 |

The pair-wise mean comparison using Tukey's HSD test; the mean of pigmentation index differed significantly among NILs except between Harosoy and Harosoy-*e3* and Harosoy-*E2* and Harosoy-*e4.* Means of cracking index differed significantly among NILs except Harosoy and Harosoy-*E2* and Harosoy-*e4* and Harosoy-*E5*.

**Table 02. Soybean material used (Previous study)**

| Line name | Line designation | Genotype | Pedigree |
|---|---|---|---|
| **Harosoy** | - | *e1 e2 E3 E4 e5 E7 dt2 t pd1* | |
| **Harosoy-*E5*** | L64-4830 | *e1 e2 E3 E4 E5 E7 dt2 t pd1* | Harosoy (6) x PI80837 |
| **Clark-*e2*** | L62-1392 | *e1 e2 E3 E4 e5 E7 dt2 T pd1* | Clark (6) x PI86024 |
| **Clark-*e2E5*** | L94-1110 | *e1 e2 E3 E4 E5 E7 dt2 T pd1* | L63-3117 (Clark-*e2*) (6) x L644830 |

**Table 03. $F_2$ populations used (Previous study)**

| Crossing combination | Year of crossing | Date of planting | Number of parents | Number of $F_2$ plants |
|---|---|---|---|---|
| Harosoy-*E*5 x Clark-*e*2 | 2007 | June 20, 2008 | 4 | 98 |
| Harosoy x Clark-*e*2*E*5 | 2010 | June 16, 2011 | 8 | 110 |

**Table 04. Soybean material used in developing mapping populations**

| Line name | Line designation | Genotype | Pedigree |
|---|---|---|---|
| **Harosoy** | - | *e*1 *e*2 *E*3 *E*4 *e*5 *E*7 *dt*2 *t pd*1 | - |
| **Harosoy-*E*5*Dt*2** | L62-812 | *e*1 *e*2 *E*3 *E*4 *E*5 *E*7 *Dt*2 *t pd*1 | Harosoy (6) x PI80837 |
| **Clark-*e*2** | L62-1392 | *e*1 *e*2 *E*3 *E*4 *e*5 *E*7 *dt*2 *T pd*1 | Clark (6) x PI86024(*e*2) |
| **PI 80837** | - | *E*5 *Dt*2 *t Pd*1 | - |

**Table 05. $F_2$ populations used in the study**

| Crossing combination | Year of Crossing | Date of planting | Cultivation method | Location | Number of Parents | Number of $F_2$ plants |
|---|---|---|---|---|---|---|
| Harosoy-*E*5*Dt*2 x Clark-*e*2 | 2010 | May 25, 2011 | Field experiment | Sapporo | 4 | 119 |
| Harosoy x PI 80837 | 2012 | June 27, 2013 | Pot experiment inside vinyl plastic house | Tsukuba | 8 | 104 |

**Table 06. Sequences of DNA markers used for identification of alleles at loci *E1-E4***
(Liu *et al*, 2008 and Tsubokura *et al*, 2014)

| Marker | Primer | Sequence (5′–3′) |
|---|---|---|
| E1_TaqαI | G33snpTaqcutF | TCAGATGAAAGGGAGCAGTGTCAAAAGAAGT |
| | G33snpTaqcutR1 | TCCGATCTCATCACCTTTCC |
| E2_DraI | SoyGI_dCAPaMs19300FW | GAAGCCCATCAGAGGCATGTCTTATT |
| | SoyGI_dCAPa19440RV | GAGGCAGAGCCAAAGCCTAT |
| E2_InDel | E2_15345FW | TGTTGATATTACATGCACATGCAT |
| | E2_15856RV | GGCAGTTTCACCTTCTTAGC |
| E3_Mix | E3_08557FW | TGGAGGGTATTGGATGATGC |
| | E3_09908RV | CTAAGTCCGCCTCTGGTTTCAG |
| | E3Ha_1000RV | CGGTCAAGAGCCAACATGAG |
| | e3tr_0716RV | GTCCTATACAATTCTTTACGACG |
| E3_MseI | E3_08094FW | TTGCATGAAGTTTTGGTTGC |
| | E3_08417RV | CAACTGAACTGAAGACCCACAA |
| E4_Mix | PhyA2-for | AGACGTAGTGCTAGGGCTAT |
| | PhyA2-Rev/E4 | GCATCTCGCATCACCAGATCA |
| | PhyA2-Rev/e4 | GCTCATCCCTTCGAATTCAG |

**Table 07. Linkage groups obtained from $F_2$ populations**

| $F_2$ population | Number of polymorphic markers | Number of linked markers | Number of linkage groups | Total map length (cM) |
|---|---|---|---|---|
| **Harosoy-*E5Dt2* x Clark-*e2*** | 185 | 179 | 27 | 1982 |
| **Harosoy x PI 80837** | 207 | 197 | 35 | 2522 |

**Table 08. QTLs observed in two $F_2$ populations**

| $F_2$ population | Phenotypic trait | Name of the QTL | Linkage group | Proximal marker | Position (cM)* | LOD score | Additive effect ⁑ | Dominance effect | Variance explained (%) |
|---|---|---|---|---|---|---|---|---|---|
| Harosoy-*E5Dt2* x Clark-*e2* | Days to flowering | qDF 03 | B1 (chr 11) | Sat_149 | 73.4 | 6.46 | 1.59 | -0.18 | 25.5 |
| Harosoy x PI 80837 | Days to flowering | qDF 04 | C2 (chr 6) | Satt365 | 74.4 | 29.9 | -4.45 | 2.27 | 73.0 |
| | | qDF 05 | L (chr 19) | Sat_286 | 40.7 | 11.75 | 2.86 | 0.53 | 16.6 |
| | Days to maturity | qDM 03 | C2 (chr 6) | Satt365 | 69.3 | 5.25 | -6.49 | 3.33 | 46.9 |

* Distance from top of linkage group.

⁑ Additive effect of each QTL are those of Harosoy or Harosoy-NIL allele in contrast to pollen parent.

**Table 09. Alleles at maturity loci *E1-E4* in parental lines**

| Parental line | *E1* | *E2* | *E3* | *E4* |
|---|---|---|---|---|
| **Harosoy** | *e1-as* | *e2-ns* | *E3-Ha* | *E4* |
| **Harosoy-*E5*** | *e1-as* | *E2-dl* | *E3-Ha* | *E4* |
| **Harosoy-*E5Dt2*** | *e1-as* | *e2-ns* | *E3-Ha* | *E4* |
| **Clark** | *e1-as* | *E2-in* | *E3-Ha* | *E4* |
| **Clark-*e2*** | *e1-as* | *e2-ns* | *E3-Ha* | *E4* |
| **Clark-*e2E5*** | *e1-as* | *e2-ns* | *E3-Ha* | *E4* |
| **PI 80837** | *E1* | *e2-ns* | *e3-tr* | *E4* |

**Table 10. QTLs observed in two $F_2$ populations with diagnostic markers**

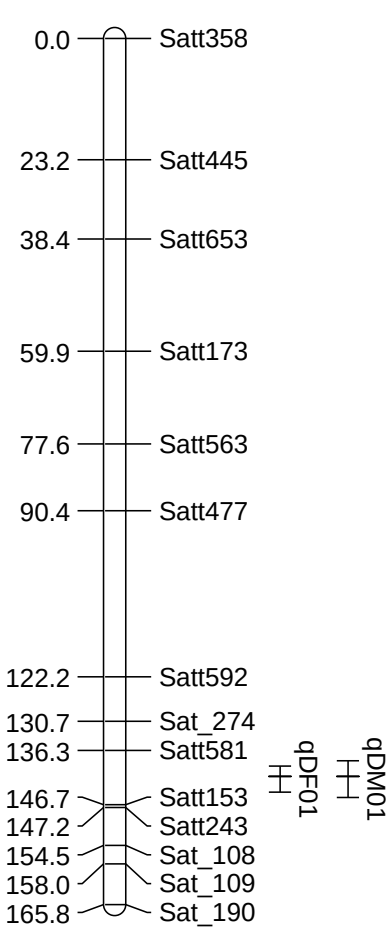
| $F_2$ population | Phenotypic trait | Name of the QTL | Linkage group | Proximal marker | Position (cM)* | LOD score | Additive effect ⁂ | Dominance effect | Variance explained (%) |
|---|---|---|---|---|---|---|---|---|---|
| Harosoy-*E5* x Clark-*e2* | Days to flowering | qDF 01 | O (chr 10) | E2_*Dra*I | 146.5 | 26.44 | 2.38 | 0.43 | 51.2 |
| | Days to maturity | qDM 01 | O (chr 10) | E2_*Dra*I | 145.5 | 19.51 | 4.74 | 1.64 | 29.5 |
| Harosoy x PI 80837 | Days to flowering | qDF 04 | C2 (chr 6) | E1_*Taqα*I | 72.2 | 37.15 | -4.48 | 2.07 | 69.5 |
| | | qDF 05 | L (chr 19) | E3_Mix | 24.3 | 15.87 | 2.90 | 0.27 | 19.5 |
| | Days to maturity | qDM 03 | C2 (chr 6) | E1_*Taqα*I | 69.3 | 6.76 | -5.71 | 4.45 | 56.2 |
| | | qDM 04 | L (chr 19) | E3_Mix | 25.3 | 4.47 | 5.14 | 1.14 | 14.3 |

∗ Distance from top of linkage group.

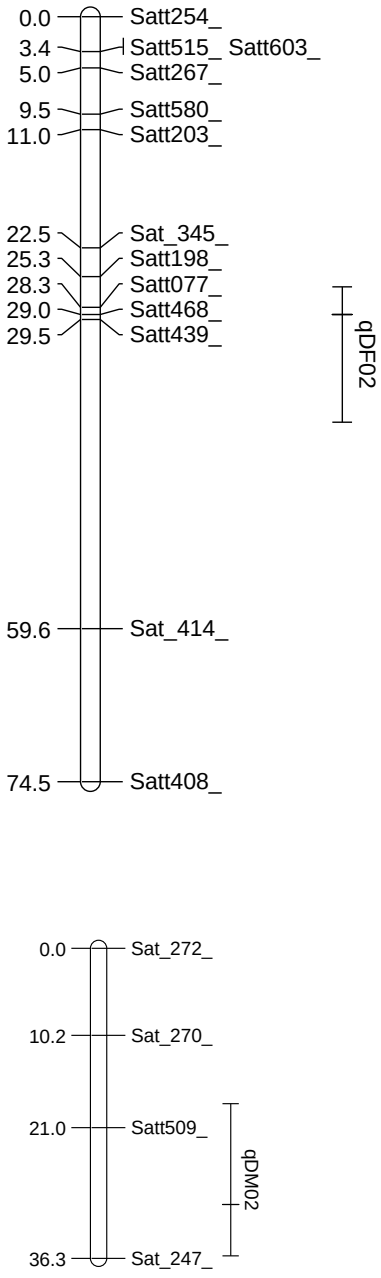
⁂ Additive effect of each QTL are those of Harosoy or Harosoy-NIL allele in contrast to pollen parent.

**Table 11. Genotype of Harosoy-*E5* at SSR markers compared to Harosoy and PI 80837**

| Linkage group | Marker | Genotype | Linkage group | Marker | Genotype | Linkage group | Marker | Genotype |
|---|---|---|---|---|---|---|---|---|
| A1 | Satt165 | Harosoy-type | D1b | Sat_279 | Harosoy-type | J | Satt249 | Harosoy-type |
| | Satt382 | Harosoy-type | | Sat_135 | Harosoy-type | | Satt285 | Harosoy-type |
| | Satt300 | Harosoy-type | | Satt537 | Harosoy-type | | Satt406 | different-type |
| | Sat_171 | Harosoy-type | | Sat_183 | different-type | | Satt621 | Harosoy-type |
| | Sat_267 | Harosoy-type | | Satg001 | Harosoy-type | | Satt431 | Harosoy-type |
| A2 | Sat_383 | Harosoy-type | D2 | Satt135 | Harosoy-type | K | Satt715 | Harosoy-type |
| | Sat_212 | Harosoy-type | | Satt372 | Harosoy-type | | Satt055 | Harosoy-type |
| | Sat_233 | Harosoy-type | | Satt311 | Harosoy-type | | Satt247 | Harosoy-type |
| | Satt329 | Harosoy-type | | Satt301 | Harosoy-type | | Satt499 | Harosoy-type |
| | Sat_294 | Harosoy-type | | Satt386 | Harosoy-type | | Sat_352 | Harosoy-type |
| B1 | Sat_272 | Harosoy-type | E | Sat_112 | Harosoy-type | L | Satt495 | Harosoy-type |
| | Satt509 | Harosoy-type | | Satt651 | Harosoy-type | | Satt182 | Harosoy-type |
| | Sat_247 | Harosoy-type | | Satt699 | Harosoy-type | | Satt278 | Harosoy-type |
| | Satt597 | Harosoy-type | | Satt706 | Harosoy-type | | Satt076 | Harosoy-type |
| | Satt359 | Harosoy-type | | Satt685 | Harosoy-type | | Sat_245 | Harosoy-type |
| B2 | Sat_264 | Harosoy-type | F | Satt193 | Harosoy-type | M | Sat_389 | Harosoy-type |
| | Sat_182 | Harosoy-type | | Satt149 | Harosoy-type | | Satt245 | Harosoy-type |
| | Satt474 | Harosoy-type | | Satt374 | Harosoy-type | | Satt175 | Harosoy-type |
| | Satt726 | Harosoy-type | | Satt490 | Harosoy-type | | Sat_422 | Harosoy-type |
| | Satt687 | Harosoy-type | | Sat_074 | Harosoy-type | | Sat_330 | Harosoy-type |
| C1 | Satt396 | Harosoy-type | G | Satt163 | Harosoy-type | N | Satt009 | Harosoy-type |
| | Sat_140 | Harosoy-type | | Satt688 | Harosoy-type | | Sat_280 | Harosoy-type |
| | Satt338 | Harosoy-type | | Sat_131 | Harosoy-type | | Satt521 | Harosoy-type |
| | Sat_311 | Harosoy-type | | Sct_199 | Harosoy-type | | Satt257 | Harosoy-type |
| | Satt164 | Harosoy-type | | Sat_372 | Harosoy-type | | Satt022 | Harosoy-type |
| C2 | AW734043 | Harosoy-type | H | Sat_200 | Harosoy-type | O | BF08905 | Harosoy-type |
| | Satt291 | Harosoy-type | | Satt442 | Harosoy-type | | Satt679 | Harosoy-type |
| | Satt643 | Harosoy-type | | Satt637 | Harosoy-type | | Sat_193 | Harosoy-type |
| | Sat_312 | Harosoy-type | | Satt181 | Harosoy-type | | Satt477 | Harosoy-type |
| | Satt371 | Harosoy-type | | Sat_401 | Harosoy-type | | Satt592 | different-type |
| D1a | Satt184 | Harosoy-type | I | Satt571 | Harosoy-type | | | |
| | Satt179 | Harosoy-type | | Satt700 | Harosoy-type | | | |
| | Satt283 | Harosoy-type | | Satt049 | different-type | | | |
| | Satt077 | Harosoy-type | | Satt162 | Harosoy-type | | | |
| | Satt071 | Harosoy-type | | Satt148 | PI 80837-type | | | |




Harosoy-*E5* x Clark-*e2* $F_2$ population




Harosoy x Clark-*e2E5* $F_2$ population

**Figure 01. Linkage groups containing QTL for days to flowering and maturity (Previous study)**

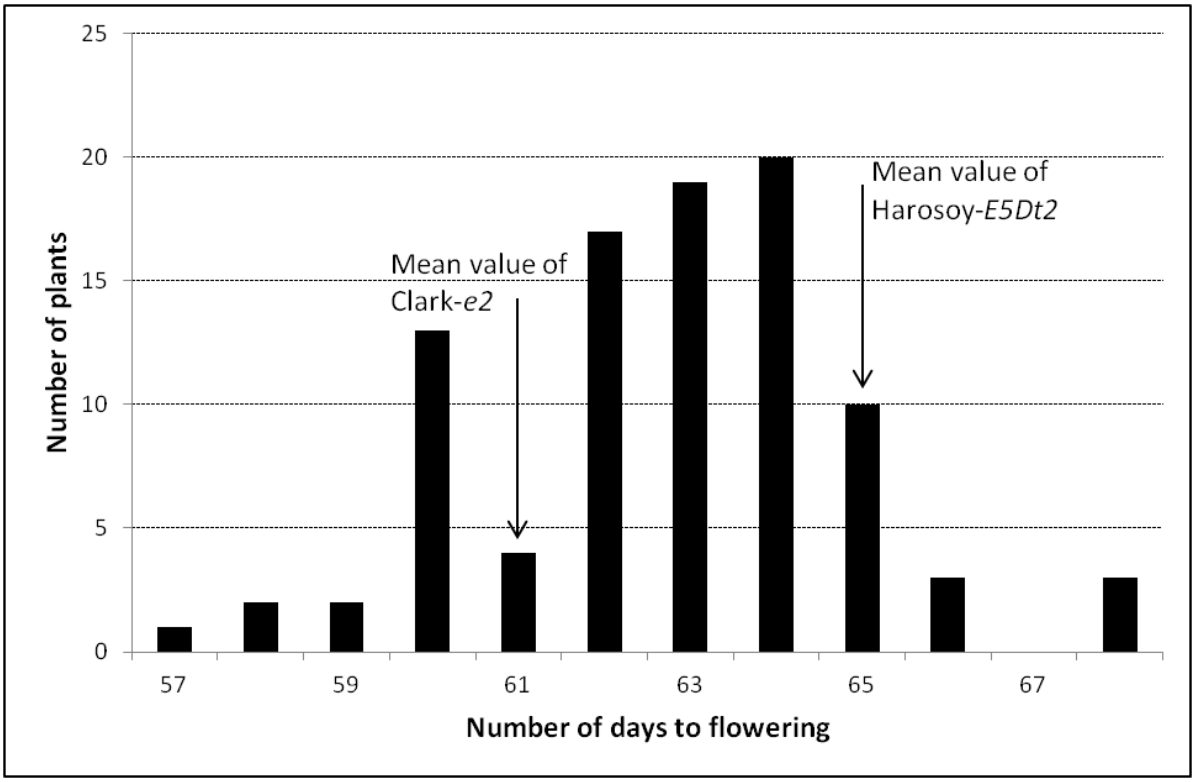
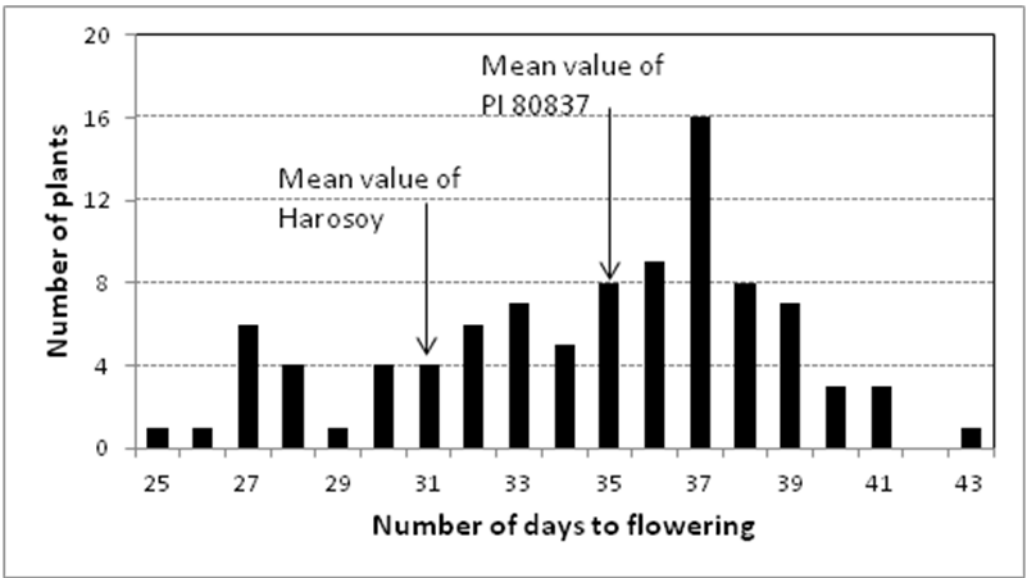



**Figure 02. Frequency distribution of days to flowering in Harosoy-*E5Dt2* x Clark-*e2* $F_2$ population**



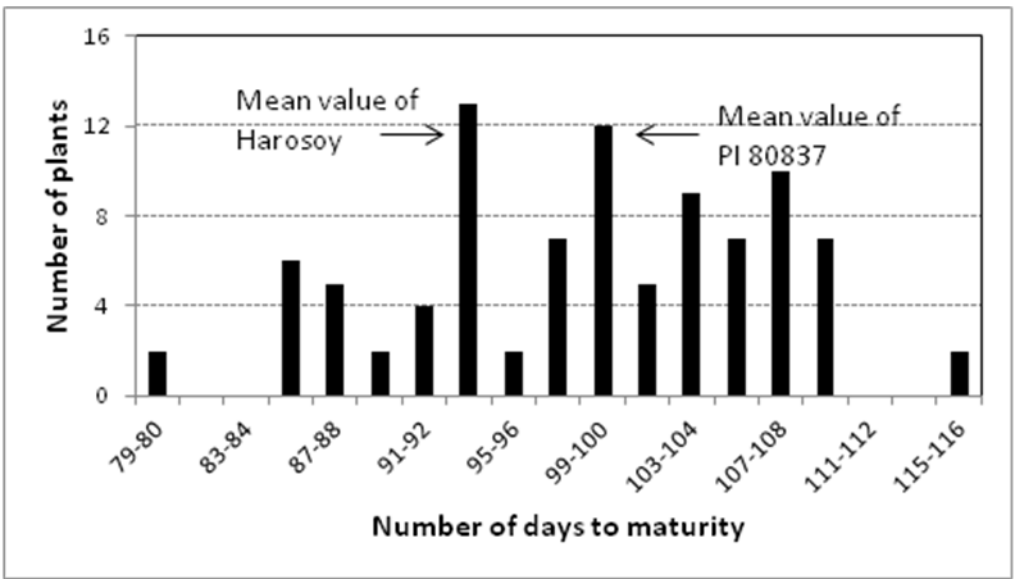




**Figure 03. Frequency distribution of days to flowering and maturity in Harosoy x PI 80837 $F_2$ population**

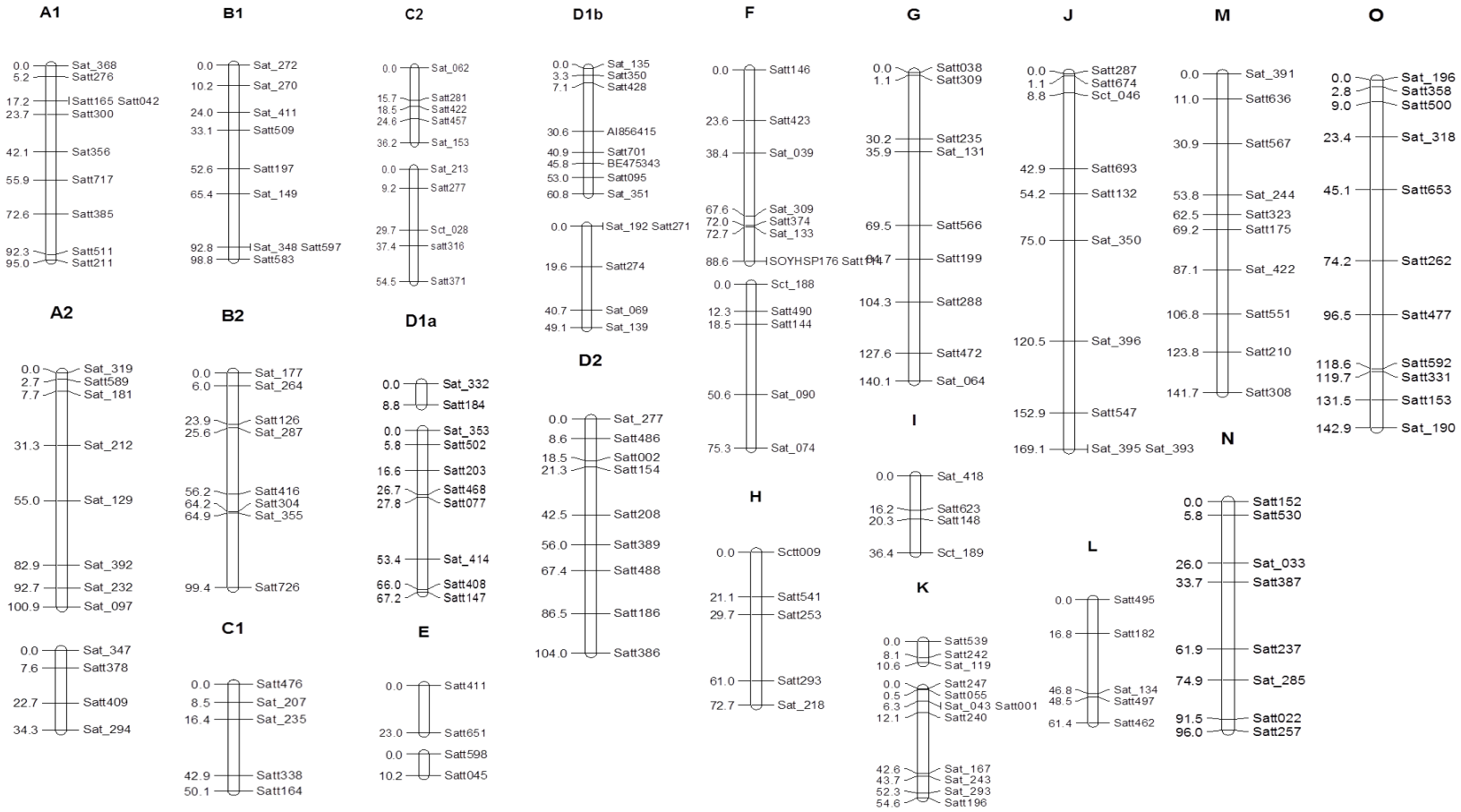



**Figure 04. Linkage map of Harosoy-*E5Dt2* x Clark-*e2* $F_2$ population**

The numbers on the left side of each linkage group indicate the genetic distance (cM) from the top. The name of the linkage group is indicated at the top.

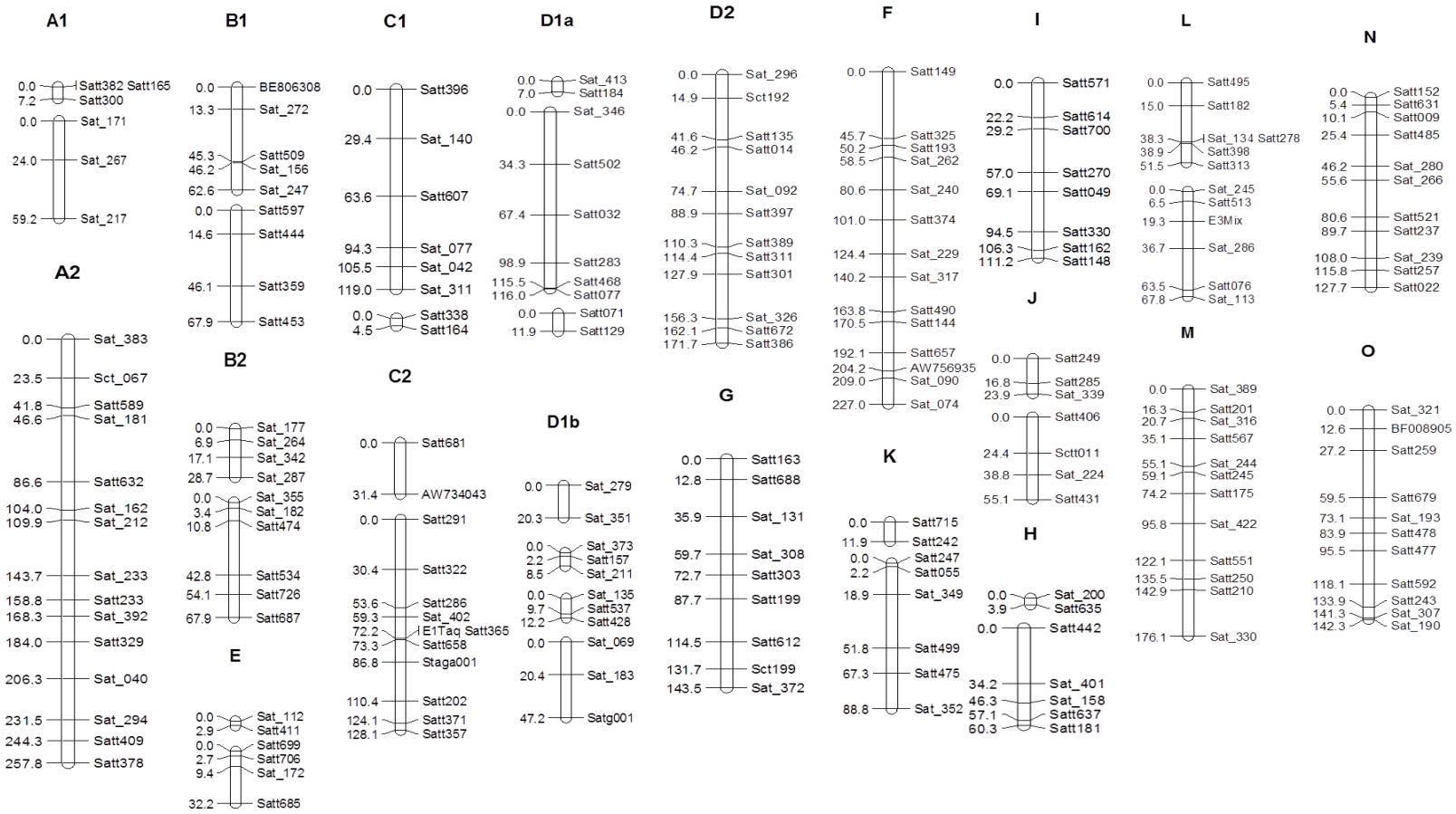



**Figure 05. Linkage map of Harosoy x PI 80837 $F_2$ population**

The numbers on the left side of each linkage group indicate the genetic distance (cM) from the top. The name of the linkage group is indicated at the top.

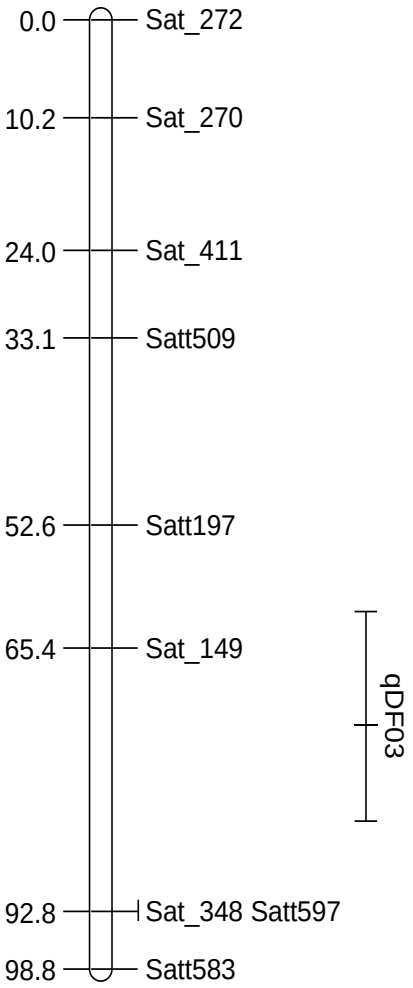



**Figure 06. MLG B1 (chr 11) containing QTL qDF 03 for days to flowering in $F_2$ population of Harosoy-*E5Dt2* x Clark-*e2***

The distance of markers (cM) from the top of each linkage group is shown on the left. Length of vertical bar (of QTL) is equal to the one-LOD likelihood confidence interval. Horizontal line at the centre of the bar indicate the position of the QTL peak.

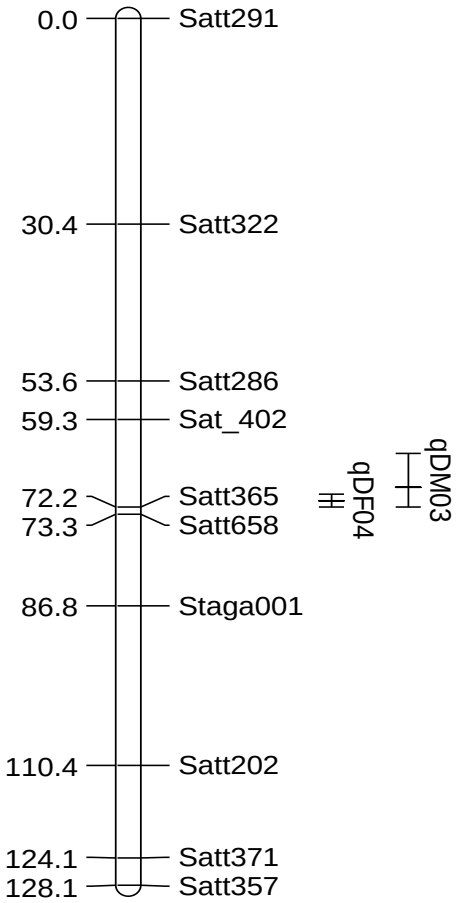



**Figure 07. MLG C2 (chr 6) containing QTLs qDF 04 and qDM 03 for days to flowering and maturity in $F_2$ population of Harosoy x PI 80837**

The distance of markers (cM) from the top of each linkage group is shown on the left.

Length of vertical bar (of QTL) is equal to the one-LOD likelihood confidence interval.

Horizontal line at the centre of the bar indicate the position of the QTL peak.

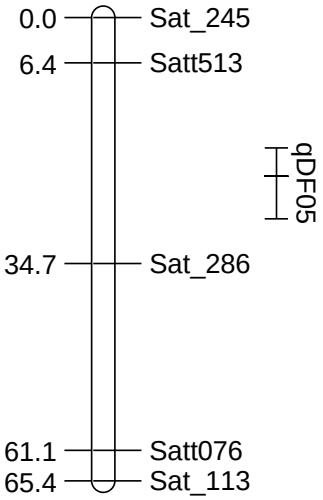



**Figure 08. MLG L (chr 19) containing QTL qDF 05 for days to flowering in $F_2$ population of Harosoy x PI 80837**

The distance of markers (cM) from the top of each linkage group is shown on the left.

Length of vertical bar (of QTL) is equal to the one-LOD likelihood confidence interval.

Horizontal line at the centre of the bar indicate the position of the QTL peak.

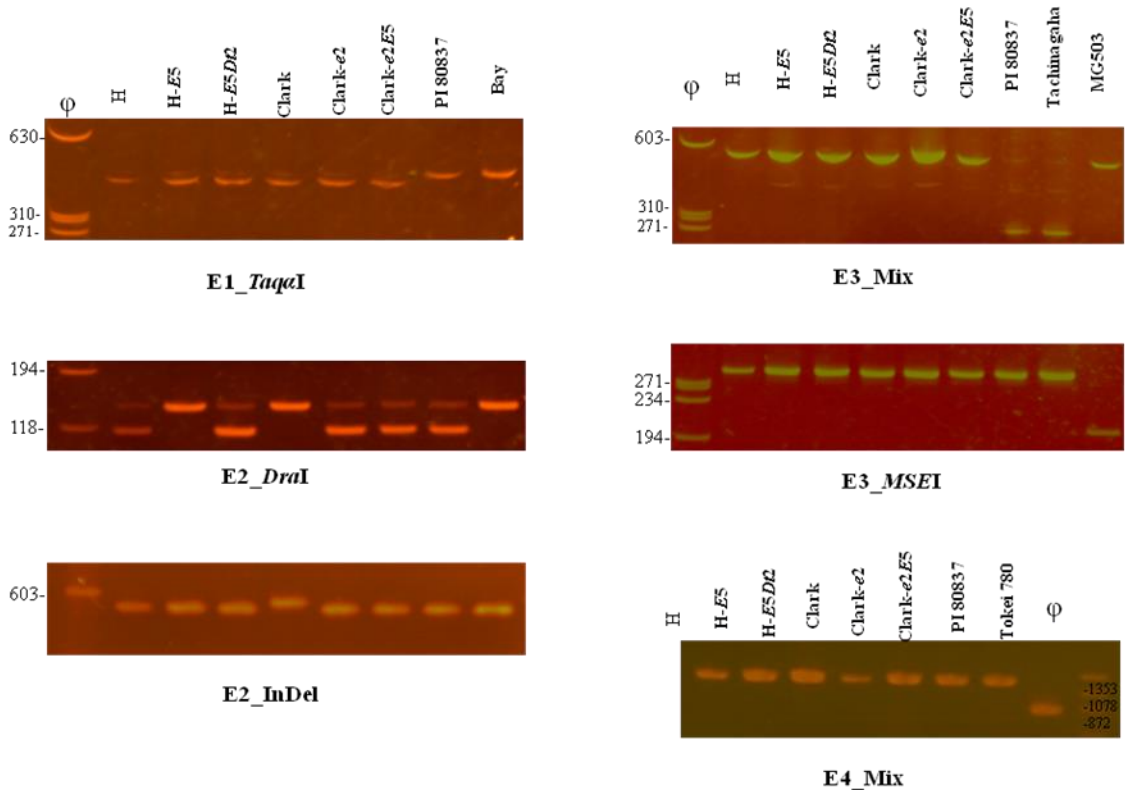
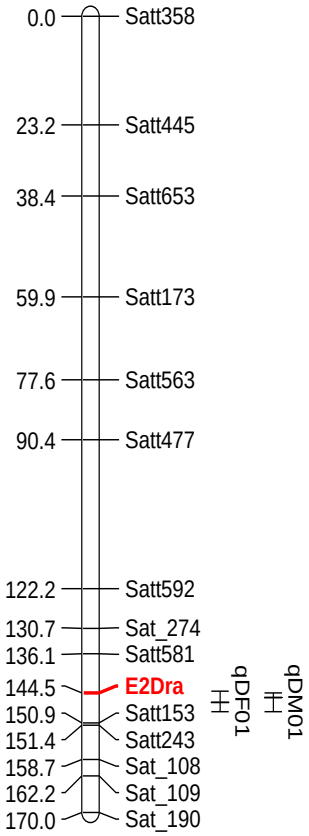



**Figure 09. Results of ALP/dCAPS analysis used to identify alleles at the *E1, E2, E3* and *E4* loci in parental lines**

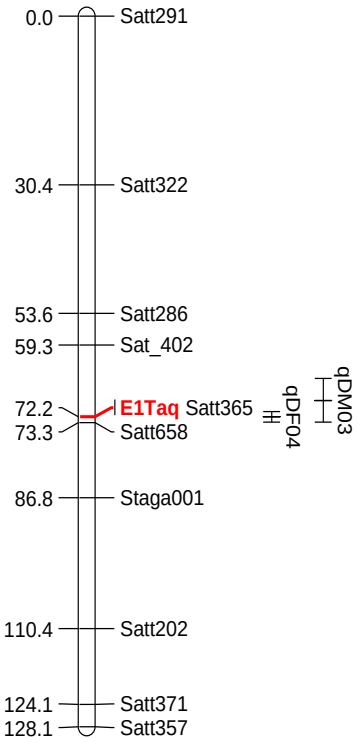
Marker designations are indicated below of each gel picture. The migration of size marker (bp) is shown to the left of each gel picture.

ϕ, molecular marker ϕx174/*Hae*III; H, Harosoy; H-*E5*, Harosoy-*E5*;

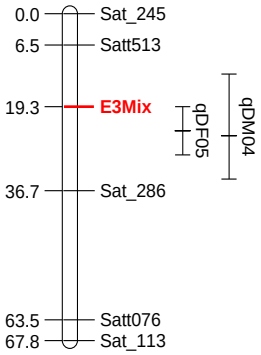
H-*E5Dt2*, Harosoy- *E5Dt2*.




MLG O (chr 10)
Harosoy-*E5* x Clark-*e2* $F_2$ population




MLG C2 (chr 06)
Harosoy x PI 80837 F2 population




MLG L (chr 19)
Harosoy x PI 80837 F2 population

**Figure 10. Linkage groups containing QTLs for days to flowering and maturity with diagnostic markers**

## Acknowledgements

My first sincere gratitude goes to my academic advisor Prof. Ryoji Takahashi for his continuous support during my PhD study. His guidance, vast knowledge and experience made this study a success.

Next, I would like to thank Prof. Chikako Otobe, Assoc. Prof. Junichi Tanaka, and Prof. Sachio Maruyama for their valuable advices.

I am so much grateful to Japanese government, MEXT scholarship program and to all the tax payers of Japan for providing me an opportunity to study in Japan.

I also express my gratitude to University of Tsukuba and Institute of Crop Science, NARO for all the research facilities provided and for comfortable accommodation. I like to extend my gratitude to the staff of Tsukuba International House, NARO.

I like to remind my amazing lab mates Dr. Felipe Rojas, Dr. Tito Rodriguez, Dr. Yan Fan and Dr. Shaokang Di. I thank them for their invaluable support I received in conducting research and for creating a friendly atmosphere at the lab. I also thank Mrs. Iizumi and Mrs. Tulia Lopez for their friendship and support.

I thank the Board of directors, Coconut Research Institute of Sri Lanka for granting me study leave to carry out this PhD study.

I remember the support and friendship I received from Sri Lankan community living in Tsukuba and the Tsukuba MEXT scholars, 2013 group. At Last but not least, I would like to thank my family in Sri Lanka; my mother, brother, sister in-law and kids for their continuous encouragement and love. Finally, I thank my late father for being the strength of my life even in his absence.

# List of References

Abe, M., Y. Kobayashi, S. Yamamoto, Y. Daimon, A. Yamaguchi, Y. Ikeda, H. Ichinoki, M. Notaguchi, K. Goto and T. Araki (2005) FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot Science 309: 1052-1056

Adachi, T., K. Tabuchi and A. Kikuchi (1994) Screening of soybean gene resources for seed coat cracking resistance and estimation of soybean seed coat resistance genes by maximum likelihood method. (In Japanese) Tohoku Agric. Res. 47: 163-164

Agarwal, K. and K. Menon (1974) Lignin content and seed coat thickness in relation to seed coat cracking in soybean. Seed Res. 2: 64–66.

Bernard, R. L. (1971) Two major genes for time of flowering and maturity in soybeans. Crop Sci. 11:242-244

Benitez, E.R., H. Funatsuki, Y. Kaneko, Y. Matsuzawa, S.W. Bang and R. Takahashi (2004) Soybean maturity gene effects on seed coat pigmentation and cracking in response to low temperatures. Crop Sci. 44: 2038-2042

Bernard, R., R. Nelson and C. Cremeens (1991) USDA soybean genetic collections: isoline collection. Soyb. Genet. Newsl. 18: 27-57

Brown-Lima, C. M. Cooney and D. Cleary (n.d.) An overview of the Brazil-China soybean trade and its strategic implications for conservation. The nature conservancy. Latin American region. (http://www.nature.org/ourinitiatives/regions/southamerica/brazil/explore/brazil-china-soybean-trade.pdf). accessed on 18/09/2016

Buzzell, R.I (1971) Inheritance of a soybean flowering response to fluorescent-day length conditions. Can. J. Genet. Cytol. 13: 703–707

Buzzell, R.I. and H.G. Voldeng (1980) Inheritance of insensitivity to day length. Soyb. Genet. Newsl. 7: 26-29

Carlson, J.B. and N.R. Lerstern (2004) Reproduction morphology. In: Boerma H.R. and E. Specht James (Ed). Soybeans: Improvement, production and uses. 3rd.ed. Madison, WI. 59-93pp

Chang S., J. Puryear and J. Cairney (1993) A Simple and efficient method for isolating RNA from pine trees. Plant Mol. Biol. Rep. 11: 113-116

Cober, E.R. and H.D. Voldeng (2001) A new soybean maturity and photoperiod-sensitivity locus linked to *E1* and *T*. Crop Sci. 41:698-701

Cober, E.R., S.J. Molnar, M. Charette, and H.D. Voldeng (2010) A new locus for early maturity in soybean. Crop Sci. 50:524-527

Cregan, P.B., T. Jarvik, A.L. Bush, R.C. Shoemaker, K.G. Lark, A.L. Kahler, N. Kaya, T.T. vanToai, D.G. Lohnes, J. Chung, and J.E. Specht (1999) An integrated genetic linkage map of the soybean genome. Crop Sci. 39:1464–1490

Hisano, H., S. Sato, S. Isobe, S. Sasamoto, T. Wada, A. Matsuno, T. Fujishiro, M. Yamada, S. Nakayama, Y. Nakamura, S. Watanabe, K. Harada and S. Tabata (2007) Characterization of the soybean genome using EST-derived microsatellite markers. DNA Res. 14:271–281

Hong, J.C., R.T. Nagao and J.L Key (1989) Characterization and sequence analysis of a developmentally regulated putative cell wall protein gene isolated from soybean. J. Biol. Chem. 262: 8367-8376

Islas-Rubio, A.R. and I. Higuera-Ciapara (2002) Soybeans: Post-harvest operations. Post harvest compendium. Food and Agricultural Organization of United Nations. (http://www.fao.org/3/a-ax444e.pdf). accessed on 13/05/2016)

Kamiya, M. and T. Kiguchi (2003) Rapid DNA extraction method from soybean seeds. Breed. Sci. 53:277-279

Kong, F.J., H.Y. Nan, D. Cao, Y. Li, F.F. Wu, J.L. Wang, S.J. Lu, X.H. Yuan, E.R. Cober and J.Abe (2014) A new dominant gene *E9* conditions early flowering and maturity in soybean. Crop Sci. 54: 2529-2535

Kour, A., A. M. Boone and L.O. Vodkin (2014) RNA-Seq Profiling of a defective seed coat mutation in *Glycine max* reveals differential expression of proline-rich and other cell Wall protein transcripts. PLoS ONE 9 (5): e96342. doi:10.1371/journal.pone.0096342

Lander, E.S., P. Green, J. Abrahamson, A. Barlow, M.J. Daly. S.E. Lincoln and L. Newburg (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage map of experimental and natural populations. Genomics 1: 174-181

Lehti-Shiu, M.D. and S. Shiu (2012) Diversity, classification and function of the plant protein kinase superfamily. Phil. Trans. R. Soc. B, 367: 2619–2639

Liu, B., A. Kanazawa, H. Matsumura, R. Takahashi, K. Harada and J. Abe (2008) Genetic redundancy in soybean photoresponses associated with duplication of the phytochromeA gene. Genetics 180: 995-1007

Liu, H. L. (1949) Inheritance of defective seed coat in soybeans. J. Hered. 40: 317-322

Maruyama, N. and K. Mikoshiba (1976) Seed coat cracking in soybeans (in Japanese). Proc. Crop Sci. Soc. Jpn. (Suppl. 2) 45: 45-46

McBlain, B.A and R.L. Bernard (1987) A new gene affecting the time of flowering and maturity in soybeans. J. Hered. 78: 160-162

Mishra, P. and K.C. Panigrahi (2015). GIGANTEA–an emerging story. Front. Plant Sci. 6:8

Murata, K., A. Kikuchi and S. Sakai (1991) Soybean seed coat cracking test by soak and drying method (in Japanese). Rep. Tohoku Br. Crop Sci. Soc. Jpn. 34: 57-58

Murray, M.G. and W.F Thompson (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. 8: 4321-4325

Nagai, I. (1926) [On the genetics of the soybean.] part 2 (In Japanese). Agriculture and Horticulture 1: 107–118

Nakamura, T., D. Yang, S. Kalaiselvi, Y. Uematsu and R. Takahashi (2003) Genetic analysis of net-like cracking in soybean seed coats. Euphytica 133: 179–184

Nicholas, C.D., J.T. Lindstrom and L.O. Vodkin (1993) Variation of proline rich cell wall proteins in soybean lines with anthocyanin mutations. Plant Mol. Biol. 21: 145-156

Okabe, A. (1996) Inheritance of seed coat cracking and effective selection method for the resistance in soybean. JARQ 30: 15-20

Oyoo, M. E., E.R. Benitez, H. Matsumura and R. Takahashi (2010a) QTL analysis of seed coat cracking in soybean. Crop Sci. 50: 1230-1235

Oyoo, M.E., S.M. Githiri, E.R. Benitez and R. Takahashi (2010b) QTL analysis of net-like cracking in soybean seed coats. Breeding Science 60: 28–33

Oyoo, M.E., E.R. Benitez, H. Kurosaki, S. Ohnishi, T. Miyoshi, C. Kiribuchi-Otobe, A. Horigane and R. Takahashi (2011) QTL analysis of soybean seed coat discolouration associated with *II TT* genotype. Crop Sci. 51: 464-469

Percy, J.D., R. Philip and L.O. Vodkin (1999) A defective seed coat pattern (Net) is correlated with the post-transcriptional abundance of soluble proline-rich cell wall proteins. Plant Mol Biol. 40: 603–613

Piper, C.V. and W.J. Morse (1923) The Soybean. New York: McGraw-Hill Book Co. 329

Rodriguez, T.O., F.R. Rodas, M.E. Oyoo, M. Senda and R. Takahashi (2013) Inverted repeats of chalcone synthase 3 pseudogene is associated with seed coat discolouration in soybean. Crop Sci. 53: 518-523

Rodriguez Torrico, T. (2014) Molecular genetics studies on abiotic stress tolerance in soybean. Doctoral thesis. University of Tsukuba. Japan

Samach, A., H. Onouchi, S.E. Gold, G.S. Ditta, Z. Schwarz-Sommer, M.F. Yanofsky and G. Coupland (2000) Distinct roles of *CONSTANS* target genes in reproductive development of *Arabidopsis*. Science. 288: 1613-1616

Sasaki, K. and S. Nakamura, (1981) Test for resistance to seed coat cracking by pod removal in soybeans (In Japanese). Tohoku Agric. Res. 29: 107-108

Song, Q. J., L.F. Marek, R.C. Shoemaker, K.G. Lark, V.C. Concibido, X. Delannay, J.E. Specht and P.B. Cregan (2004) A new integrated genetic linkage map of the soybean. Theor. Appl. Genet. 109: 122-128

Song, Q.J., G.F. Jia, Y.L Zhu, D. Grant, R.T. Nelson, E.Y Hwang, D.L. Hyten and P.B. Cregan (2010) Abundance of SSR Motifs and Development of Candidate Polymorphic SSR Markers (BARCSOYSSR_1.0) in Soybean, Crop Sci. 50: 1950-1960

Stewart, R.T. and J.B. Wentz, (1930) A defective seed coat character in soybeans. J. Am. Soc. Agron. 22: 658-662

Sunada, K. and T. Ito (1982) Soybean grain quality as affected by low temperature treatments in plants (colour of hilum, seed coat cracking) (in Japanese). Rep. Hokkaido Branch, Crop Sci. Soc. Jpn and Hokkaido Branch. Jpn. Soc. Breeding 22: 34

Suzuki, M., E. Takahashi, H. Miyakawa (1979) Characteristics of seed coat cracking of soybean and its difference by climatic conditions (in Japanese). Tohoku Agric. Res. 25:59-60

Takahashi R. (1997) Association of soybean genes *I* and *T* with low temperature induced seed coat deterioration. Crop Sci. 37: 1755-1759

Takahashi, R and J. Abe (1994) Genetic and linkage analysis of low temperature-induced browning in soybean seed coats. J. Hered. 85: 447-450

Takahashi, R and J. Abe (1999) Soybean maturity genes associated with seed coat pigmentation and cracking in response to low temperatures. Crop Sci. 39: 1657-1662

Takahashi, R. and S. Asanuma (1996) Association of *T* gene with chilling tolerance in soybean. Crop Sci. 36: 559-562

Thoenes, P. (2006) Soybean international commodity profile: Background paper for the competitive commercial agriculture in sub–Saharan Africa (CCAA) study. Food and Agriculture Organization of the United Nations.

Tsubokura, Y., S. Watanabe, Z. Xia, H. Kanamori, H. Yamagata, A. Kaga, Y. Katayose, J. Abe, M. Ishimoto and K. Harada (2014) Natural variation in the genes responsible for maturity loci *E1, E2, E3* and *E4* in soybean. Ann Bot. 113: 429-441

U.SA. Soybean export council (2012) Japanese soybean market intelligence: Special Report. (http://www.ussec.org/wp-content/uploads/2012/10/ASA-IM-Special-Report.pdf). accessed on 17/05/2016

Wang, S.C., J. Basten and Z.B. Zeng (2007) Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC. (http://statgen.ncsu.edu/qtlcart/WQTLCart.htm)

Watanabe, S., Z. Xia, R. Hideshima, Y. Tsubokura, S. Sato, N. Yamanaka, R. Takahashi, T. Anai, S. Tabata, K. Kitamura and K. Harada (2011) Map-based cloning strategy employing a residual heterozygous line reveals that the *GIGANTEA* gene is involved in soybean maturity and flowering. Genetics 188: 395–407

Weber, H., L. Borisjuk and U. Wobus (2005) Molecular physiology of legume seed development. Annu.Rev.Plant Biol. 56:253-279

Weiss, M.G. (1970) Genetic linkage in soybeans: Linkage group I. Crop Sci. 10:69-72

Wolf, W.J., F.L. Baker and R.L. Bernard (1981) Soybean seed coat structural features: pits, deposits and cracks. Scanning Electron Microscopy III: 531-544 Chicago, IL

Woodworth, C.M. and L.F. Williams (1938) Recent studies on the genetics of the soybean. J. Am. Soc. Agron. 30: 125–129

Xia, Z., S. Watanabe, T. Yamada, Y. Tsubokura, H. Nakashima, H. Zhai, T. Anai, S. Sato, T. Yamazaki, S. Lü, H. Wu, S. Tabata, and K. Harada (2012) Positional cloning and characterization reveal the molecular basis for soybean maturity locus *E1* that regulates photoperiodic flowering. Proc.Natl. Acad. Sci. USA 109: E2155–E2164

Yalkichi, R.W. and G. Barla-Szabo (1993) Seed coat cracking in soybean. Crop Sci. 33:1016-1019

Zeng, Z.B. (1993) Theoritical basis for separation of multiple linked gene effects in mapping quantitative trait loci. Proc.Natl. Acad. Sci. USA 90:10972-10976

Zhang, X., M. Wang, T. Wu, C. Wu, B. Jiang, C. Guo, T. Han (2016) Physiological and molecular studies of stay-green caused by pod removal and seed injury in soybean. The Crop Journal, 4: 435-443