An experimental study on morphological diversity and genetic erosion in *Phaseolus vulgaris* L.

Micha Groenewegen
Student number: 920822-282-120

MSc Thesis Plant Breeding
(PBR-80436)

1st Supervisor: Edith Lammerts van Bueren
2nd Supervisor: Chris Kik

April 2016

Preamble

I would like to thank for Jan Velema for this research opportunity as well as the many lessons in plant breeding. Furthermore, I would like to thank Edith Lammerts van Bueren and Chris Kik for their critical questions and comments during the research and the writing of the thesis. I would also like to thank Ravi van Marissingen for his help during the cultivation and the very nice lunches. Furthermore, I would like to thank Ravi's parents, his brother as well as Andy, Christa, and Maartje for their help during the planting, sowing, and weeding. Finally, I would like to thank my parents for their support and the nice food they cooked that got me through the last few days of writing.

Abstract

Phaseolus vulgaris L. is one of the most consumed legumes in the world. However, in the Netherlands, legumes have decreased in popularity and due to this the number of bean varieties and number of hectares cultivated with beans has decreased. To analyze the total variation present in the bean varieties considered and the change in morphological diversity, that might have occurred between the 1920s and the 1990s, a group of 188 varieties were analyzed for 24 traits. This was done via a field experiment using a completely randomized block design. The beans were sown twice, in pots, on the 21st of May and on the 10th of June, and planted on the 4th of June and the 24th of June, respectively. ANOVAs were performed, and PCAs and Nei's variation index were calculated for the data, which were divided in a total set encompassing all shared variables, and three subsets with variables for pole beans, bush beans, and French beans. Furthermore, the data were analyzed based on the following three extra subgroups: growth and consumption type, release type: before the 1920s and after, and seed colour type group, to understand the potential loss of diversity. The statistical tests did show that the varieties released between the 1920s and the 1990s compared to those from before the 1920s or unregistered, had a reduced diversity in number of seed colors, darker leaf colour and rounder pods. However, the analysis of varieties grouped per decade over time did not yield any significant patterns suggesting the influence of breeding on the diversity has been small in this period.

Key words: *Phaseolus vulgaris*, diversity, genetic erosion

Introduction

Beans (*Phaseolus vulgaris* L.*)* belong to the *Fabaceae* family and are part of the genus *Phaseolus* (Gepts, Singh & DeBouck; 1990). Beans originate from Southern America and were discovered and brought to the Old World by Columbus (Bitochhi *et al.*, 2012). *Phaseolus vulgaris* L. is the most consumed leguminous pulse in the world, and is a major dietary staple and source of protein in Africa and Latin America (Mamidi *et al.*, 2013). Beans are leguminous pulses including chickpeas, lentils and peas. Beans themselves can be separated into dry beans and green beans. Green beans are immature pods of the bean plant, whereas dry beans are the mature seeds of the bean plant. Both are important staples in the diets of many people, in the past and the present. The possible genetic erosion of *Phaseolus vulgaris* L., in the field, is the focus of this MSc research. The first part of the thesis describes the origin and worldwide significance of *Phaseolus vulgaris* L. to provide the background for the problem statement that follows. After the problem statement, the methods of measuring diversity are briefly discussed after which the research objectives and questions are presented. In the materials and methods section, the setup of the experiment, the actions at different stages of the experiment and the situation on the farm are explained. Furthermore, the data collection and methods of analysis are described. The results section, presents the data based on the tests performed, and the discussion and conclusions section discusses the results and their implications on the basis of the research questions.

Origin

Recent research shows that *Phaseolus vulgaris* L., the common bean, was already cultivated a long time before the discovery of the New World in 1492 by Columbus. He observed the bean three weeks after his arrival in the New World (Uilenburg & Steenbergen, 1950-53). Some researchers hypothesize that the bean has been domesticated twice in the New World, on separate occasions (Mamidi *et al.*, 2011, Bitochhi *et al.*, 2012). This led to the development of two major gene pools, the Mesoamerican and the Andean (Zeven, 1999). Both derived from a common ancestor which existed 110,000 years ago, after which both pools suffered a bottleneck event which was most severe for the Andean gene pool (Mamidi *et al.*, 2013). The Andean genepool became the origin of most cultivated and improved *P. vulgaris* varieties, although it has less genetic diversity than the Mesoamerican genepool (Bitochhi *et al.*, 2012). There is evidence that *P. vulgaris* is in the process of speciation: crosses between the Andean and the Mesoamerican germplasm are leading to F1 lethality (Gepts & Bliss, 1985). On the European Continent *P. vulgaris* was embraced after its introduction in the 16th century and has led to the development of many different varieties; many home gardeners and small farmers have developed and maintained their own bean varieties (Zeven, 1999). However, in recent years many of these beans have not been cultivated anymore and are only maintained in gene banks.

Significance of beans

In 2013 23 million tons of dry beans were produced worldwide. The average yield worldwide was 800 kilograms per hectare. Of the total dry bean production 46.0 % were produced in Asia, 30.6% in the Americas, 21.0% in Africa and only 2.2% in Europe. Myanmar and India are the two largest dry bean producers. Between 2010 and 2013 20 million tons of green beans were produced worldwide and the yield in 2012 was 13 tons per hectare. Green beans are mostly produced in China, which produces 91.1% of all green beans (FAOSTAT, 2015).

In Western Europe the production and consumption of dry beans and green beans has decreased over the past 40 to 50 years. This is often correlated with the rise of readymade meals. The Netherlands can be taken as a case study: the production has fluctuated over the years, but has decreased from 15,000 tonnes in the 1990s to less than 5000 tonnes in 2013. The production of green beans has decreased from 95,000 tonnes in the 1990s to 36,000 tonnes in 2013 (FAOSTAT, 2015). However, at the same time the population has grown more than two fold from 7.8 million in the 1930 to 16.7 million in 2013 (CBS, 2015).

It is to be expected that bean consumption will increase again. An increasing number of studies are published on the health benefits of beans and beans are becoming more popular in urban areas than ever before. Beans are high in fibers (Messina, 2014; Siddiq *et al.*, 2011), something which is lacking from the diets of many people in Western Europe (Messina, 2014). Moreover, in 2015 the Dutch health council recommended Dutch people to increase their intake of legumes (Gezondheidsraad, 2015). Also, beans contain high levels of protein, making them good replacements for animal based food. Furthermore, beans contain many minerals and vitamins (FAO, 2016). With increasing interest in their nutritional properties during the UN 2016 year of the pulses (FAO, 2016), it is to be expected that interest for legumes such as dry beans and green beans will continue to grow. However, difficulties remain for increasing the popularity of pulses, as for example the time needed to prepare beans is considerable, especially with dry beans (Arruda *et al.*, 2012).

Problem Statement

In the Netherlands many different landraces and varieties of *Phaseolus vulgaris* L. used to be cultivated (Zeven, 1997). However, currently only a few of the many varieties once actively cultivated, are cultivated today. For example, for dry brown bean production in the Netherlands only two cultivars are used: Berna and Narda (Timmer, 2012). Motivated home gardeners and a few small farms maintain some of the landraces. Many varieties have been stored in gene banks: the collection of the Dutch national gene bank, the CGN (Centrum voor Genetische Bronnen, Nederland), which contained over 235 Dutch varieties, together with 729 other varieties, from Europe and the USA, was sent to the CIAT (Centro Internacional de Agriculturo Tropical) between 1976 and 1980 (Drijfhout, 1987a). Many of the traditional Dutch varieties that are still cultivated on a commercial scale are old, as indicated by Timmer (2012). For example, the two only commercially used dry bean cultivars, Berna and Narda, in the Netherlands were released in 1960 and 1976. The total cultivated area of brown beans, 1000 – 1500 hectares, is sown with these two varieties. Furthermore, farmers harvest their own seeds for next year's planting since there is close to no professional propagation of these varieties, leading to many off types. Furthermore, the breeding work in beans, both dry and green for non-industry purposes has completely stopped in the Netherlands (Velema, 2015). It could thus be hypothesized that *Phaseolus vulgaris* L. might suffer from genetic erosion in the Netherlands, if one compares the diversity from before the 1900s with the current available genetic diversity. Genetic erosion is defined by Luquet *et al*. (2012) and others as: "the loss of genetic diversity within a crop". Genetic erosion has occurred in many other crops aside from the common bean. The process repeats itself: traditional varieties and landraces are replaced by modern varieties, furthermore, modern breeding, by focusing mostly on yield improvement, results in new varieties which are lacking the diversity found in traditional[1] varieties and landraces (Bennett *et al.*, 1999). Therefore, one would expect that without control this would lead to genetic erosion *on farm*.

In this context Van der Wouw *et al.* (2010) showed that for varieties released in the 20th century an interesting phenomenon occurred. They found, on the basis of a meta-analysis encompassing many cereal crops, that the amount of genetic variation in subsequent decades, using as a proxy Nei gene diversity index (Nei 1973) was identical. The only exception was the 1960's decade in which the variation observed was less and in many cases significantly less compared to other decades. Van de Wouw and colleagues suggested that this difference might be due to the strong influence of breeding on genetic variation in varieties during the green revolution (1960-1970) resulting in varieties with reduced genetic variation. However due to the improved access to genetic variation, because of the establishment of genebanks starting in the 1960's and the use of wild relatives in breeding programmes, the genetic variation in varieties increased again after this period, resulting in similar levels of genetic variation as before the 1960's. It would be interesting to investigate if these trends are also present in the bean varieties released in the 20th century in the Netherlands.

In the case of *P. vulgaris* L. all major breeding done in the Netherlands by Pop Vriend seeds (Rogers, 2012; Dodde, 2014) and Holland Select (Holland Select, 2016). Both companies focus on varieties for industrial processing. With the specific requirements of the industry the genetic diversity generated by breeder's new varieties will be low and with the current challenges posed by climate change such as drought, nutrient and salinity stress, genetic diversity is important as this might lead to a better adaption to (a)biotics stresses (Graham & Vance 2003; McClean *et al.*, 2011).

Before traditional bean varieties can be cultivated again, and thus maintained *on farm*, proper descriptions are necessary to provide information for interested growers. The description of old bean varieties is not comprehensive. For example, the Dutch bean germplasm that has been transferred to the CIAT, has been described incompletely. The observed characteristics at the CIAT in Columbia for *Phaseolus vulgaris* L. are: growth habit, seed color, seed shape, seed brightness, 100 seed weight, days to flowering date of the first harvest and the last harvest were measured. This is a relatively small selection of characteristics that was described. For example, for farmers it might be necessary to know the

[1] Varieties not officially registered and originating from before 1920

suitability for machine harvest, which could be indicated by the inflorescence position, which has not been described by current genebank descriptions. The International Union for the Protection of new Varieties of Plants (UPOV) guidelines are the internationally most used ones for DUS (Distinctness – Uniformity – Stability) evaluation. The Community Plant Variety Office (CPVO) protocol is the standard within the European Union concerning description guidelines. UPOV is the most descriptive one allowing for more distinction between the various traits. This is not surprising when one knows that the CPVO protocol has been derived from the UPOV guidelines. Uilenburg & Steenbergen (n.d.) did one of the last systematic bean characterizations of the available old Dutch varieties. The characteristics noted by Uilenburg & Steenbergen (n.d.) focus mostly on the morphology of the fruit and the plant. The shape, presence of string, color, length, number of seeds, form of the cross section, presence of inedible fiber of the pod are all noted. The disease susceptibility, development (vigor), and flower color are noted as well. There is a large overlap between the UPOV characteristics and the characteristics noted by Uilenburg & Steenbergen. However, Uilenburg and Steenbergen had a stronger focus on pseudo qualitative characteristics such as vigor. UPOV guidelines focus more on quantitative and qualitative characteristics.

Hypothesis

In this thesis, it is hypothesized that the rise of producing crops for industrial processing corresponds with a decline in morphological and genetic diversity in commercially released *Phaseolus vulgaris* L. varieties. It is expected that this decline has happened due to demand from the industry to provide uniform varieties meeting a few very strict set of demands, amongst those demands two are the most important ones: green pods and white seeds. It is hypothesized that because of this focus on these criteria breeding companies have started to release less diverse varieties which limits the available morphological diversity.

Research objective

- Describe patterns of morphological variation in bean landraces and varieties cultivated in the Netherlands and processes involved in shaping this variation.

Research questions

- What is the morphological variation in a set of 188 varieties of *Phaseolus vulgaris* L. cultivated in the Netherlands during the last century in a common garden experiment?
- Have any changes in traits occurred for varieties, grouped in decades, and cultivated between the 1920s and the 2000s in the Netherlands?
- Is there a significant difference between undated varieties and varieties cultivated in the Netherlands from the 1920s onwards?

Material and methods

The experiment

One hundred eighty-eight accessions were selected by Jan Velema director of *De Zaderij*. The selection criteria for including accessions were a combination of origin and availability. All of the varieties originate from the Netherlands. The majority of the seeds were obtained from small scale seed savers, the *IPK Gatersleben* gene bank, and various seed companies such as *Vreeken's Zaden* and *Zaadhandel van der Wal*. A few varieties were multiplied before the experiment by Jan Velema at his own home farm to provide sufficient sowing material. Other varieties, mostly those retrieved from the gene banks, had been stored for long times (see Appendix I for a complete overview of all the selected varieties, their source of origin and type of bean). The accessions were divided in various groups as defined by other authors such as Uilenburg & Steenbergen (n.d.) to gain a better insight into:

a. the variation present in Dutch germplasm,
b. the difference between traditional accessions originating from before 1920 and varieties released between the 1920s and the 1990s, and
c. the changes over time within the group of varieties released between the 1920s and the 1990s.

Firstly, the division based on consumption (French or dry bean types) and growth (bush or pole) type is used since it represents the way that beans are categorized most often. The second group is defined as the seed color group. The division is based on Uilenburg & Steenbergen (n.d.) method of accessing bean varieties which starts from seed color. The seed color group has been divided into the following colors: white, even colored, hilum spotted, marmered, striped and spotted, and lastly bicolored. Zeven (1997) mentioned that research has consistently shown a difference between white colored beans and other colors for vigor, this can thus be investigated. The third group is defined as the decade group. This group is based on the official Dutch variety catalogues released between the 1920s and the 1990s. All varieties have been sorted into the decade group during which they were released. The last group is the release type group. This group distinguishes between those varieties which were released commercially between the 1920s and the 1990s, and those accessions which have been maintained but never have been commercially or officially released from before the 1920s. Table 1 shows an overview of the varieties as defined according to the consumption/growth group. Table 2 provides an overview of the tested beans based on the seed colour group. Table 3 provides an overview of both the decade group as well as the seed release group.

Table 1: An overview of the selection of beans ordered by consumption and growth type

Type			number of accessions
Pole beans			*53*
	French beans		42
		Round Podded	35
		Flat Podded	7
	Dry beans		11
Bush beans			*135*
	French beans		47
		Round Podded	43
		Flat Podded	4
	Dry beans		88
Total			*188*

Table 2: An overview of the selection of beans ordered by seed color group

Seed type		number of accessions
Bicolour		4
Even colour		64
	black	11
	brown	38
	red	3
	yellow	12
Hilum spot		6
Marmered		1
Striped		37
White		76
Total		188

Table 3: An overview of the varieties ordered by release type and decade group

Group		number of accessions
Not released		*150*
1920-1990s		*38*
	1920-1930	2
	1940-1950	12
	1950-1960	4
	1960-1970	5
	1970-1980	6
	1980-1990	5
	1990-2000	4
Total		*188*

Seedling raising

The first block of beans was sown on May the 21st and was planted June the 4th. The second block was sown on June the 10th and planted on June the 24th. The plants were raised inside a greenhouse where they were sown into small 6×7×6 cm pots positioned on growing benches inside an insect proof cage. This was to prevent early transmission of viruses from the plants through aphids. Plants with virus were carefully removed by hand to prevent early spread of virus. As seedling soil, potting soil for lettuce from Jiffy with 50% compost was used. The second block was sown in non-fertilized potting soil from the same brand and covered with the first type of soil, the fertilized one. This was due to a shortage of non-fertilized soil to cover them with it.

Number of plants

There were 20 plants per plot for the majority of the accessions. The UPOV guidelines state that for *P. vulgaris* all observations should be made on 20 plants, whilst the CPVO guidelines state that 30 is the ideal number. The present research deviates from this for 17 accessions, namely 6720, 6751, 6735, 6757, 6744, 8848, 6749, 6752, 8834, 6717, 8833, 8881, 8891, 6723, 6765, 6762 and 8827, for practical reasons as not enough seed material was available for these numbers of plants. Because the beans will be sown twice, there will be a total of 376 plots with Dutch bean varieties.

Soil type

The soil type at the test location at *De Zaderij is* an *Enkeerd* type. This is a sandy type of soil rich in humus, of which the top soil layer is at least 50 cm deep and contains loam. The soil type has been created by humans who have cultivated this type of soil since 6000-3500 BC. The organic material yearly brought to these soils developed the soils into fertile soils as it is now. The soil is rich of nutrient and allows for good plant growth (Jongmans & Peek, n.d.).

Fertilization

The complete field experiment was conducted at one test location: D*e Zaderij* in Voorst. Voorst is located in the province of Gelderland in the Netherlands. The experiment as described in this paragraph took place in the summer 2015 season. The soil was not especially fertilized for this experiment; however, the entire location had been fertilized with compost at the beginning of 2015. In 2014 the crops at the location of the experiment were potatoes and pumpkins, the years before the soil had not been used for intensive vegetable production.

Weed management

The weeds were managed by a combination of hand weeding, a wheel hoe, and a Dutch hoe. Hand weeding was done in the first stages; later a Dutch hoe was used in the beds, and a wheel hoe to keep the paths clear.

Layout of the experiment

The beans were planted in a completely randomized block design. Due to the fact that there were two sowing dates there were two blocks for the bush beans, and two blocks for the pole beans. A graphical representation of the experiment (Figure 1) illustrates the experiment.

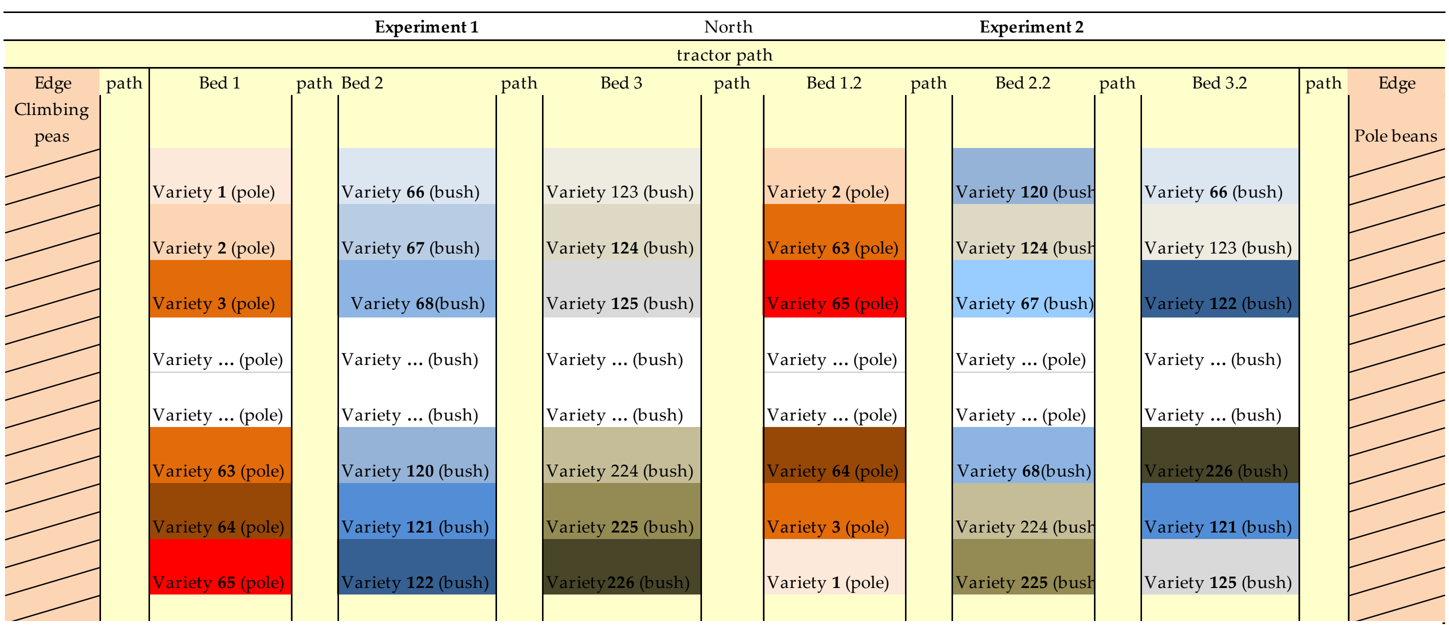


Figure 1: A graphical representation of the experiment

In both repetitions, first the pole beans were planted. After that, the bush beans were planted in the same bed and direction, starting 0,20 meters from the last planted pole bean. In the graphical representation of the experiment it is illustrated how the planting started at the beginning (the top) of the first bed (bed 1), continues in bed 2 from the bottom and then continues in bed 3 downwards again. For the second repetition the planting started 1 meter from the first repetition in the last bed, and started again with the pole beans. For both repetitions the varieties were divided in bush and pole beans, however, in the first repetition the plants are planted in order of consumption type, starting with French beans and ending with dry beans. The varieties in the second repetition were also separated based on growth type but were randomized compared to the first one for consumption type. The randomization was done using excel randomize function for the field numbers. Lastly, the experiments were flanked on both sides with tall growing crops; Experiment 1 had a boarder row with climbing peas and Experiment 2 had a boarder row with pole beans. However, there was no planting on the other two sides where the tractor path was located. For a list of all the varieties and the exact field numbers see Appendix I.

Beds at *De Zaderij* had a net length of 40 meters and a net width of 1.2 meters. Bush bean varieties were planted in rows of approximately ten plants, depending on their availability. The plants were planted with a plant distance of ten centimeters and with an in between row distance of 0.75 meters, and 0.20 meters between varieties. The pole varieties were grown along bamboo poles. At the base of each pole two to three beans were planted depending on availability. The distance between the poles was 0.6 meter. For the entire test and both sowings 15 beds were reserved.

Data collection

For the description of most of the characteristics the UPOV guidelines were used. The main two reasons for this were that UPOV is one of the two universally used guidelines in Europe for identifying and describing varieties of plants. Also, UPOV is more detailed than the CPVO. UPOV lists 52 characteristics related to *Phaseolus vulgaris* L. The last three relate to disease resistance, which could not be tested for practical reasons. The 49 remaining traits were too many to monitor in the short period of time available. Therefore, we decided to use the minimum requirements by the UPOV for beans and 18 traits are in the UPOV minimum set. Without those related to disease tests this left 15 traits for observation. Another nine traits were added due to their importance for growers. All the traits are mentioned in Table 4 including when and how these were measured. The nine added traits are marked with an asterisk and are explained in further detail below the table.

Table 4: Measured traits (24) and the number, type of observations, and how the variable has been treated in the analysis

Measured trait	When	How	Number of observations per plot	Type of variable
Leaf color	During plant growth and plant maturity	Scored on a scale, measuring darkness of the leaves, from one (lightest) to seven (darkness). Neckargold was used as a reference variety scored at six.	One	Continuous
Color of flower	During flowering	Scored on a scale, measuring purpleness ranging from, from one (white) to five (dark purple). Trebona was used as a reference variety scored at one.	One	Continuous
Length of pod	At pod maturity	Measured in centimeters	One overall first sowing	Continuous
Width of pod at maximum points	At pod maturity	Measured in centimeters	One overall first sowing	Continuous
Shape of cross section pod	At pod maturity	Divided into five categories: circular (one), cordate circular, circular to elliptic, elliptic to narrow, narrow elliptic (five). For statistical purposes	One overall first sowing	Continuous

		recoded based on length width ratio.		
Number of colors seed	Before sowing	Counted the number of colors	One overall first sowing	Continuous
Days to flowering	During season	Flowering defined as the time at which 50% of the plants in a plot are flowering.	One overall	Continuous
***Inflorescence position**	During season	Graded on a score from one to five measuring height of flowers: one: low, three: medium, five: high.	One overall	Continuous
Seed shape	After seed harvest	Crosssection and longitudinal section judged on a scale from one to five ranging from circular through elliptic to kidney shaped.	One overall	Continuous
***Earliness and harvest period**	Throughout the season	Judged on a scale from one to five beginning with the earliest ripening plants to those ripening at the end of the season.	One overall per plot	Continuous
***Lodging**	Beginning of first harvest	% of lodged plants transformed into five classes	One overall per plot	Continuous
***Seedling vigour**	One week after 50% sprouting	Disease occurrence scored on a scale from one to nine, based on	One overall per plot	Continuous

		average performance of the sprouted seeds.		
*** Performance**	Twice in the growing season	Scored on a basis from one to five based on the average of the plot per variety.	One overall per plot	Continuous
***% Seedling virus**	1st week after sprouting	Seedlings with virus/ total seedlings *100%	One overall per plot	Continuous
***% Germination**	1st week after sprouting	Sprouted seeds / total sown * 100%	One overall per plot	Continuous
***Height**		Based on height of the pole covered measured when the majority of the varieties reached the point defined as medium. The pole was divided into five equal parts.	One overall per plot	Continuous
***100 seed weight**	Before planting	Measured 20 seeds, multiplied by five or more if less seeds were available	One	Continuous

The eight characteristics which are not mentioned in the UPOV list, but were added for our observations due to their potential importance for the grower, are described in further detail below.

1. **Lodging**

Lodging is of importance in the Netherlands due to fungal infections that can easily occur if pods lie on the ground. Certain type of beans are more susceptible to lodging than others (Smak, pers. comm, 2015, Beattie *et al.*, 2003). The lodging will be measured as the percentage of plants lodging (Beattie *et al.* 2003). Later these lodging percentages can be translated into qualitative categories to increase comprehension which could range from severe through intermediate to none.

2. **Earliness and harvest period**

Commercial cultivars often have a very short or single harvest period. Knowing when this period begins and ends, can help to plan the sowing dates for farmers.

3. **Inflorescence position**

For dwarf beans it is important to know the inflorescence position. Inflorescence positioned above the crop reduces losses and increases ease of harvest. This trait is mentioned in the complete UPOV guidelines for beans and will be used in the same manner here and divided in three categories: predominantly in the foliage (low), intermediate (medium), predominantly above the foliage (high).

4. **Seed Shape**

Seed shape is a characteristic mentioned in the UPOV guidelines as well. It was decided to include this trait as well in order to find small differences between various accessions with the same name but with different characteristics.

5. **Height**

The height of the pole beans was measured in order to give an indication of the speed and ease with which they start to climb up poles. This can help to make a better planning of available growing space. For example, in 1950s a combination of two bean varieties was released, one which focused its growth more on the lower part of the pole or rope, and the second one had most of its leaf mass and fruiting in the top. This combination was sold as a mixture variety.

6. **% germination**

Percentage germination could give an indication of the sprouting speed and the sensitivity for rot. However, in this research, since we tried to minimize seed rot by not watering too much and by the usage of the non-fertilized potting soil in the experiment. The germination percentage should be taken as an indication of the vigor and quality of the seed at the start of the experiment.

7. **100 seed weight**

The weight of the seeds was measured per 20, multiplied by 5, or more, if less seeds were available.

8. **Seedling vigor**

Seedling vigor was rated on a scale from one to nine, with nine representing highest vigour. It was measured by comparing the heights one week after sowing.

9. Seedling virus

Seedling virus is mostly an indication of the quality of the basic seed. Because it was counted whilst cleaning up the seedlings it was also included as a trait.

Figure 2: Example of diversity in leaf colour, scores for these leaves would be from right to left: 7,5,3 and 2

Figure 3: Example of diversity in vigor and flowering ***, the most left variety would have a score of 1 for vigour whereas the right variety would be scored as 3.***

Statistical methods applied

The variables can be divided into groups of discrete and continuous variables. Table 4 shows the measured traits, whether it has been analyzed as discrete or not, and the motivation.

Principal component analysis (PCA) will be used to gain a better insight into the data. By clustering the varieties by traits that provide the most variation, PCA helps to explore the data more easily. The PCA analysis, can indicate clustering based on the divisions of growth and consumption type, seed color type, and release type. The PCA can indicate whether some divisions are more useful than others in finding patterns in the data. Thus, the PCA can suggest what is a better division for ordering the data. For analyzing trends, per individual characteristic, mean values were calculated and grouped by the introduction date of the material (Diederichsen, Solberg, Jeppson, 2013). However, only the introduction date is known of a small selection of the test material. Therefore, trends will only be calculated on material that was introduced between 1920 and 1990 as a result of professional breeding programs. This information was derived from a list created by Jan Velema on the basis of old variety release booklets published yearly between 1900 and the 1990s by the agricultural research service.

Nei's variation index as described by Hennink & Zeven (1990) indicates the degree to which certain characteristics are evenly represented in the population. H is the measure of Nei's variation index. H is obtained through h ($h = \sum X_i^2$), where X represents the frequency at which a certain characteristic is represented in the total population. H can be obtained through h; H= 1-h. This is a measure of the variation that is present in the population for a certain trait. A maximum is possible for the division of the characteristics over the groups, hence Hennink & Zeven (1990) suggested H'. In order to obtain H', H^{max} first needs to be obtained. $H^{max}$is the maximum possible diversity if a certain characteristic is divided over the number Y of observed groups. Thus $H^{max} = 1 - \frac{1}{Y}$ through this formula a relative index can be calculated: H'=H /H^{max} which indicates on a scale from 0 to 1, with 1 representing maximum variation, how much of the possible variation exists within in a population.

Particularities of these experiments

The 188 bean accessions differed considerably, both between different varieties and between the same variety in the different blocks. Literature indicates that sowing has a very large influence on many characteristics for *Phaseolus vulgaris* L. (Beattie *et al.*, 2003), which was definitely reflected in our results. Also climate, weather conditions, and soil have been described to have significant influences on the performance of beans (Buishand, 1959).

Due to time constraints, it was not possible to measure all characteristics for all varieties. It was decided to measure only the most relevant traits. Thus pod length, width, pod colours, and presence of string was not measured for dry beans, which are cultivated for their seeds. Furthermore, for some varieties certain traits were not relevant or not possible to measure, for example, lodging percentage is useful for bush beans but irrelevant for pole beans.

The growing season of 2015 was characterized by many extremes: extreme temperatures with a heat wave of 38 degrees, the worst June storm ever recorded and in the first two weeks of June frost was measured at soil level around 60 kilometers from the experiment trial (Willemsen, 2015). Moreover, the year also seemed to have extremes in virus pressure, with virtually all plants affected with golden bean mosaic virus, especially in later stages complicating the assessment of leaf colour and other traits. However, it also made virus scoring very difficult and unreliable. For these reasons it was decided to treat the different repetitions as different experiments.

Results

The results have been divided into two parts. In the first part, general patterns of diversity are shown. This means that the Nei's variation indices are described for all characteristics, and their averages and standard deviations are given. The descriptive statistics for the seed color group and the growth type group are given and further discussed. After this PCAs performed on a common set of shared variables and graphs showing the three main groups: growth type, seed color group and release type are shown to provide further insight into diversity of Dutch germplasm. Table 5 shows which traits have been measured for which consumption and growth type subgroup, and have been analyzed. Table 6 shows the variables which have been measured but which were either qualitative or for information purposes and are thus not further analyzed.

Table 5 Overview of the traits that have been measured and were used for statistical analysis divided on growth type group

	French bush bean	Dry bush bean	Pole french bean	Pole dry bean
100 seed weight	x	x	x	x
Days to flower	x	x	x	x
Earliness	x		x	
Flower colour	x	x	x	x
Height			x	x
Inflorescence position	x	x		
Leaf Colour	x	x	x	x
Lodging	x	x		
Performance	x	x		
Pod Crosssection	x		x	
Pod Length	x		x	
Pod Width	x		x	
Seed Crosssection	x	x	x	x
Seed longitudinal Section	x	x	x	x
Seed number of colours	x	x	x	x
Seedling Vigour	x	x	x	x

Table 6 Overview of the measured variables that have been measured for qualitative description of the varieties, indicated with one asterisk, and the variables measured for the research to gain a better understanding of the seed quality at the beginning

	French bush bean	Dry bush bean	Pole french bean	Pole dry bean
100 seed weight	x	x	x	x
Days to flower	x	x	x	x
Earliness	x		x	
Flower colour	x	x	x	x
Height			x	x
Inflorescence position	x	x		
Leaf Colour	x	x	x	x
Lodging	x	x		
Performance	x	x		
Pod Crosssection	x		x	
Pod Length	x		x	
Pod Width	x		x	
Seed Crosssection	x	x	x	x
Seed longitudinal Section	x	x	x	x
Seed number of colours	x	x	x	x
Seedling Vigour	x	x	x	x
% germination	x	x	x	x
% seedling virus	x	x	x	x

General patterns of diversity

Table 7 shows the Nei's variation index, average, standard deviation and characters per class for all variables measured in each repetition. Table 8 shows the Nei's variation index average, standard deviation and characters per class for all variables measured once. In both cases the characters observed per class refers to the different scores during observing, for example flower color scored from one to five counts as five classes. For all variables which were measured without classification the range was divided into six classes, for example, lodging was divided into six classes of equal size.

Table 7 Nei's variation index (H'), the averages, standard deviations, number of observations and number of observations per class for all variables which were measured once for each repetition

Variable	observed characters per class														N		H'		Average		Standard deviation	
Character	1		2		3		4		5		6		7									
Repetition	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
Days to flowering	29	65	87	38	48	32	7	0	9	2	8	0	*	*	188	137	0,839	0,771	51,436	46,891	6,106	5,780
Earliness	9	6	14	19	17	9	*	*	*	*	*	*	*	*	40	34	0,969	0,879	2,76	2,89	1,71	1,45
Flower colour	99	114	32	33	9	12	6	7	4	2	*	*	*	*	150	168	0,641	0,617	2,207	2,141	6,214	5,851
Height	5	4	18	4	7	13	15	7	8	8	*	*	*	*	53	36	0,921	0,909	3,096	3,306	1,257	1,261
Inflorescensce position	29	28	5	0	43	53	20	10	32	27	*	*	*	*	129	118	0,939	0,853	3,185	3,084	1,456	1,406
Leaf colour	5	1	25	9	41	27	56	63	47	32	9	21	1	3	184	156	0,904	0,910	3,726	4,165	1,338	1,261
Lodging	11	3	4	5	11	18	23	27	23	26	34	30	*	*	106	109	0,936	0,931	0,675	0,681	0,280	0,226
Performance	16	18	49	34	49	36	18	37	2	9	*	*	*	*	134	134	0,875	0,955	2,758	3,092	0,886	1,144
Seedling vigour	9	4	10	5	24	10	33	33	62	55	28	37	9	6	175	150	0,920	0,875	5,411	5,737	1,490	1,306

Table 8 Nei's variation index (H'), the averages, standard deviations, number of observations and number of observations per class for all variables which were measured once

Variable	observed characters per class						N	H'	Average	Standard deviation
Character	1	2	3	4	5	6				
100 seed weight	16	41	53	38	24	4	176	0,938	38,05	18,19
Pod length (cm)	3	39	22	17	2	7	90	0,865	14,16	3,18
Pod width (mm)	9	46	11	6	6	1	79	0,617	11,22	3,40
Pod crosssection	18	9	20	14	18	*	70	0,985	3,06	1,46
Seed Crossection	27	50	87	20	2	*	186	0,845	2,54	0,94
Seed longitudinal section	25	45	93	20	3	*	186	0,827	2,62	0,92
Seed number of colours	145	43	0	*	*	*	188	0,523	1,23	0,55

From the table it becomes clear that number of seed colors and flower colors have the lowest Nei's variation index. This means that the number of seed colors was more often only one color compared to even distribution over the three classes. Flower color was more often white and light purple than that it was dark purple. Earliness, pod cross section and performance have the highest score for Nei's variation index. This means that the scores for these traits were most evenly distributed over the total group.

Table 9 and Table 10 show the means and standard deviations for the seed color group and the growth type group. Interesting to notice is that the data show the differences between the colour groups, for example, bicoloured seeds had the lowest 100 seed weight, and the longest number of days to flowering. Even coloured varieties had the earliest ripe pods (earliness) and the longest pods. White seeded varieties and varieties with hilum spot had the lowest scores for seedling vigour. Differences between growth and consumption type also exist, as for example, bush beans had shorter days to flower than pole beans. Between consumption types it is difficult to indicate consistent differences.

Table 9 Means and standard deviations for all variables used in the analysis based on the seed color group, for all variables for which there was no variation to calculate the mean or the standard deviation an asterisk has been placed

Seed colour group	White		Even coloured		Bicoloured		Striped		Hilum Spot	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
100 seed weight	43,47	14,14	39,91	12,73	20,85	8,31	44,00	16,50	23,15	7,47
Days to flowering	49,76	4,55	48,71	4,60	54,20	7,18	49,49	4,63	47,38	1,38
Earliness	1,55	1,49	0,28	0,75	*	*	3,36	1,36	*	*
Flower colour	1,08	0,45	2,16	1,23	*	*	1,74	0,95	*	*
Height	3,51	0,99	2,86	1,05	2,75	1,06	3,00	1,03	*	*
inflorescence position	3,06	1,19	3,44	1,08	*	*	2,68	1,05	2,80	1,10
Leaf colour	4,46	1,16	3,82	1,04	3,25	0,96	3.75	0,81	4,00	0,41
Lodging	0,52	0,36	0,55	0,35	*	*	0,52	0,31	0,52	0,30
Number of seed colours	*	*	*	*	2,25	0,50	1,94	0,33	*	*
Performance	2,76	0,84	3,01	0,91	4,13	0,18	2,86	0,88	3,15	0,74
pod crossection	3,25	1,41	2,80	1,47	*	*	2,50	1,72	*	*
Pod length (cm)	14,57	3,93	15,10	3,52	*	*	13,50	1,78	*	*
Pod width (mm)	11,32	3,89	9,75	2,10	*	*	12,08	2,43	*	*
Seed crosssection	2,80	0,77	2,78	0,92	2,50	1,73	2,69	0,82	3,25	0,50
Seedling vigour	4,92	1,40	5,95	1,03	7,17	1,04	6,01	0,99	4,41	1,39

Table 10 Means and standard deviations for all variables used in the analysis based on the growth type group, for all variables for which there was no variation to calculate the mean and the standard deviation an asterisk has been placed

Growth type	Bush beans				Pole beans			
	Dry		French		Dry		French	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
100 seed weight	34,59	17,23	45,30	14,85	44,00	17,93	37,73	20,17
Days to flowering	48,44	3,20	46,78	4,65	62,80	10,00	54,73	5,34
Earliness	2,67	1,17	*	*	*	*	*	*
Flower colour	1,63	0,96	1,43	0,78	1,40	0,89	1,39	0,94
Height	*	*	*	*	2,45	0,79	3,32	0,96
inflorescence position	3,37	1,14	2,70	1,07	*	*	*	*
Leaf colour	3,76	0,82	3,89	1,51	2,95	1,21	4,43	1,34
Lodging	0,75	0,15	0,52	0,25	*	*	*	*
Number of seed colours	1,23	0,47	1,26	0,44	1,64	0,50	0,95	0,31
Performance	3,06	0,75	2,62	0,84	*	*	*	*
pod crossection	*	*	2,90	1,59	*	*	3,24	1,30
Pod length (cm)	*	*	12,80	2,16	*	*	15,63	3,46
Pod width (mm)	*	*	10,61	3,04	*	*	11,87	3,69
Seed crosssection	2,42	0,89	2,78	0,79	1,82	0,60	2,87	0,98
Seedling vigour	5,65	1,10	5,38	1,26	5,59	1,63	5,32	1,70

To further investigate and understand the differences between the groups a PCA was performed. The results of the PCA for the first three components is shown below in Table 11. In the first component seed cross section and seed longitudinal section are determining this PCA, in the second component it is seedling vigour and the last component is determined by the number of seed colours.

Table 11 Results of the ANOVA performed on the common set of variables measured for all growth and consumption types, indicating the first three components, the variance they explain, their eigenvalue and the value for the variables per component

	Component 1	Component 2	Component 3
Variance explained (%)	23,996	19,738	16,515
Eigen value	1,68	1,382	1,156
Seedling vigour	-,164	,721	-,160
Flower colour	-,009	,609	-,197
Days to flowering	-,175	-,317	,600
Leaf colour	,313	-,482	-,257
Number of seed colours	-,239	,320	,741
Seed cross section	,851	,159	,263
Seed longitudinal section	,871	,183	,127

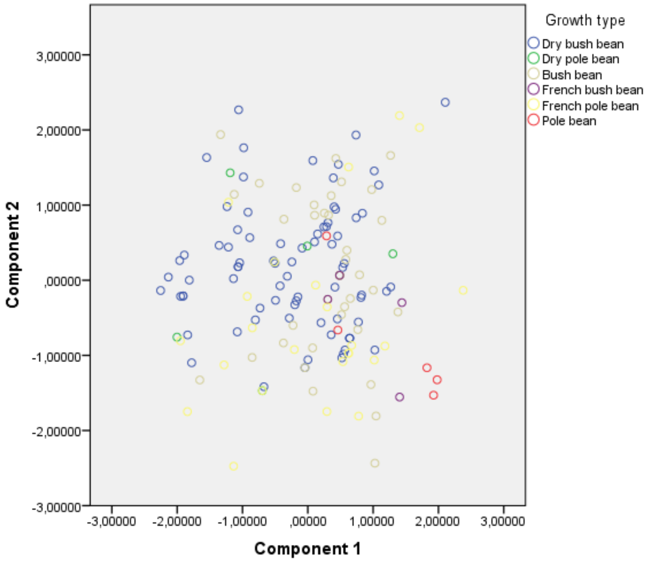


Figure 4a Results of the PCA analysis graphed for the first two components colored on growth type

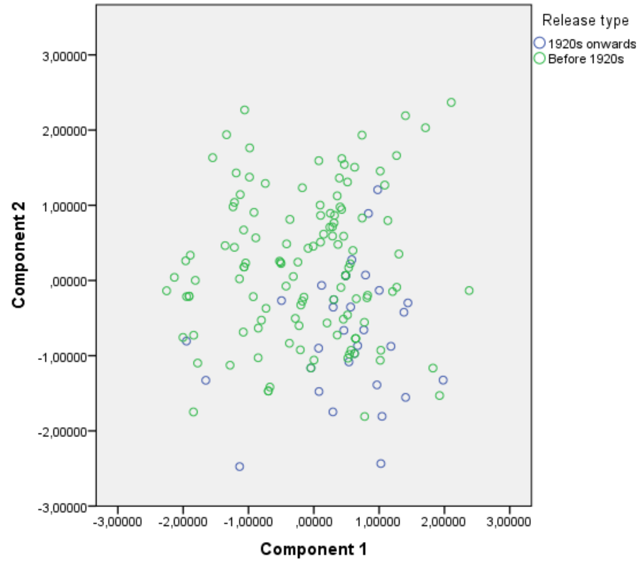


Figure 6a Results of the PCA analysis graphed for the first two components colored on release type

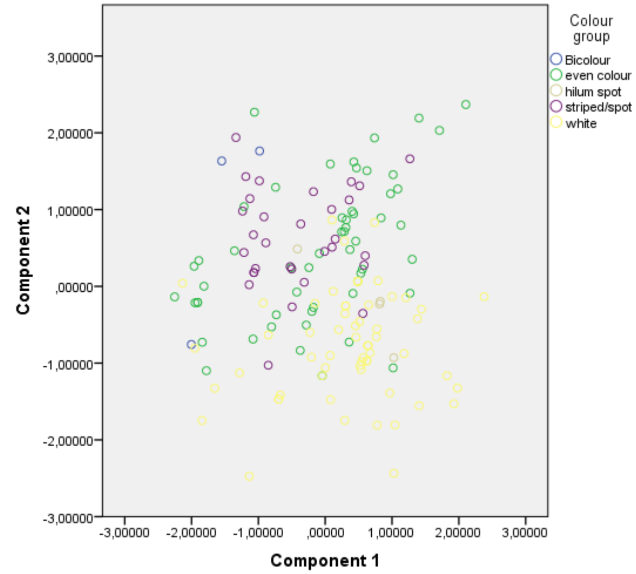


Figure 7a Results of the PCA analysis graphed for the first two components colored on color group

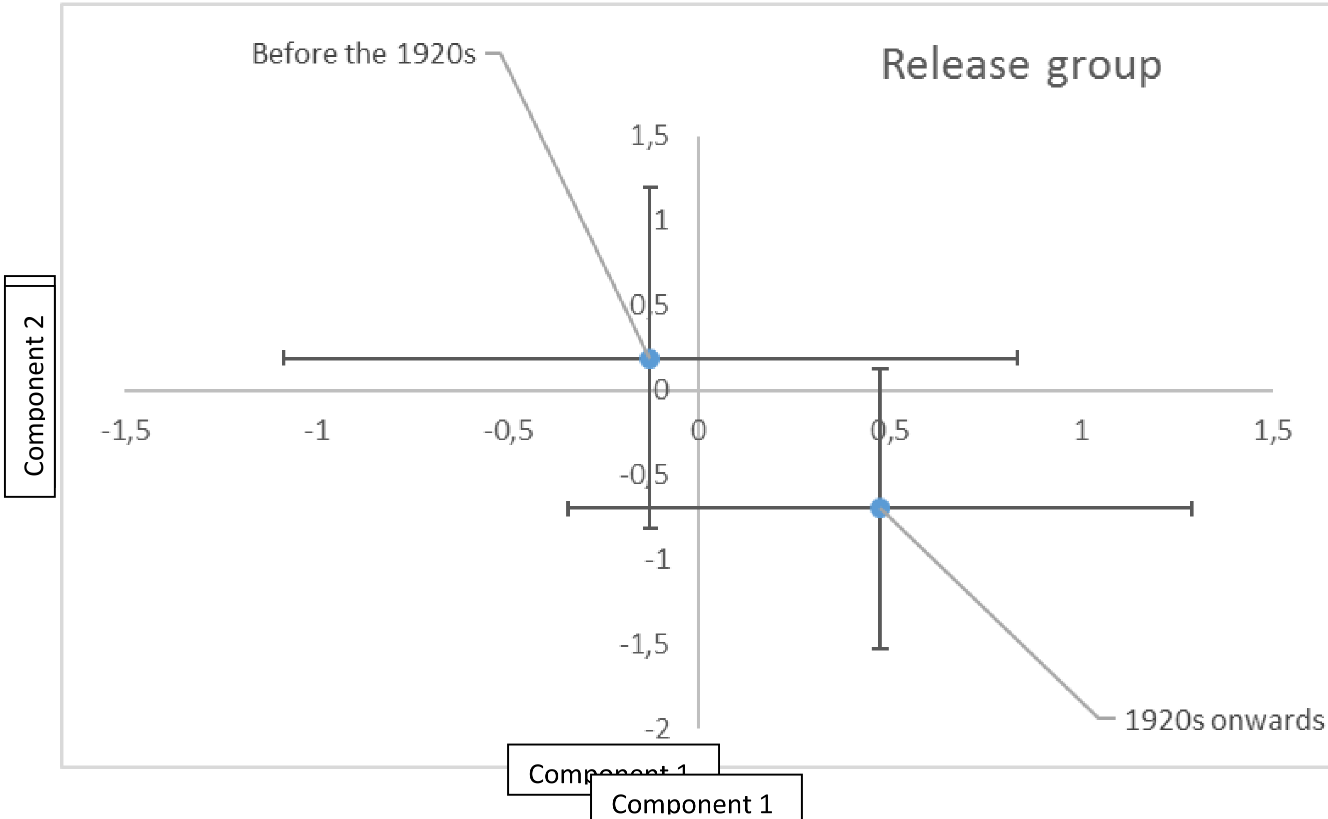


Figure 5b Results of the PCA analysis graphed for the first two components, indicating the mean component scores and the standard deviation range for the growth type

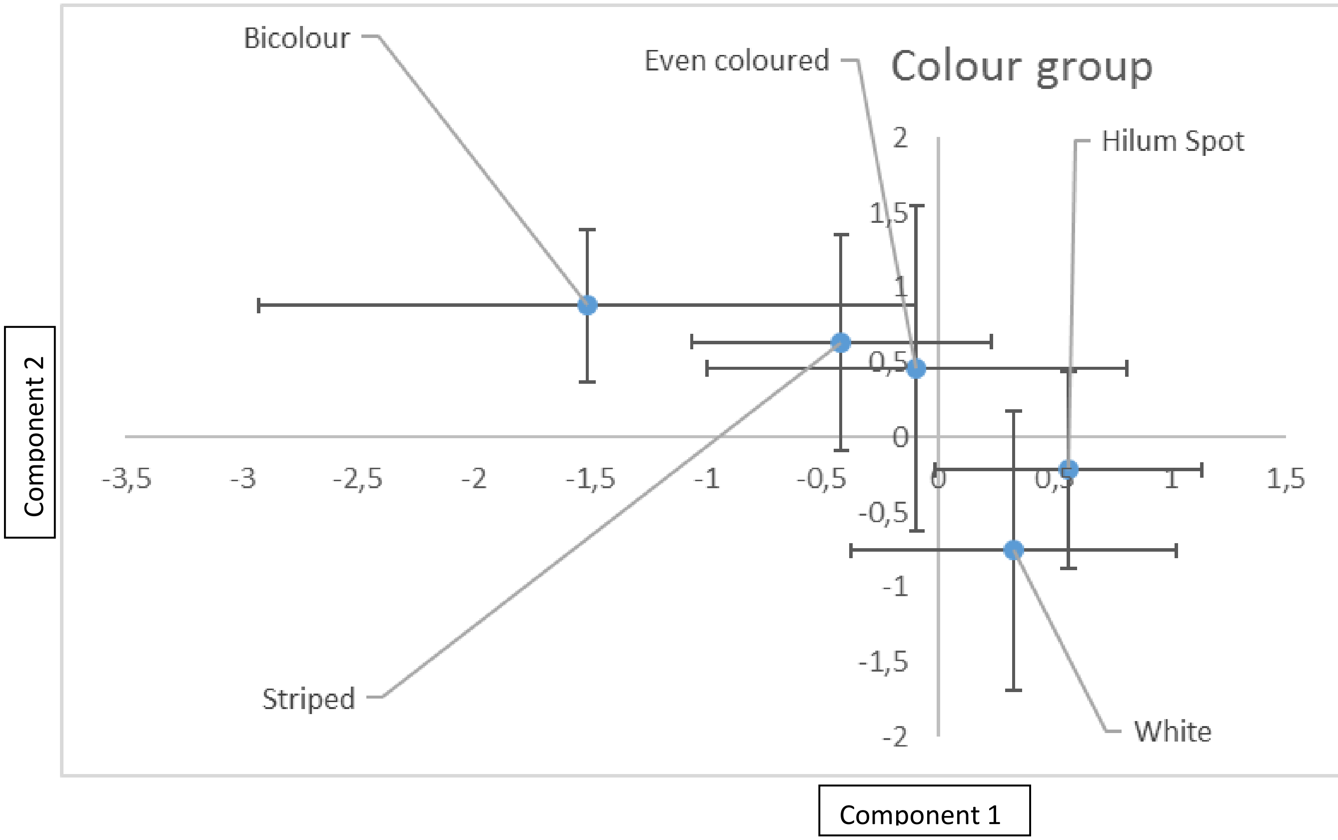


Figure 7b Results of the PCA analysis graphed for the first two components, indicating the mean component scores and the standard deviation range for the colour group

Figures 5a,b, 6a,b, 7a and b show the PCA analysis in graphs, colored based on the different groups used in the analysis. The, a, figures show the separate data points for all varieties, with colour the relevant group was indicated. The b figures show the mean component score and the standard deviation range, this helps to distinguish the different groups better by not showing all individual points. It is interesting to notice that the color group more clearly shows differences between the first two components for the various colors. This was already expected since the averages for the various colors differed for various traits. Results of PCAs on different release type clearly show how the varieties released after the 1920s are a selective part of the diversity present in the before the 1920s group. Growth type and consumption type show little distinction between the various groups in this research.

To confirm whether the difference in seedling vigour for the various colour groups was significant a one-way ANOVA was performed. The difference was indeed significant for both white seeded beans and those with a hilum spot.

Table 12 Results of one-way ANOVA, indicated by the P-value and F-value and significant differences obtained through SNK test for the different colour groups.

Colour group	P-value	F-Value	White	Even colour	Striped	Hilum Spot	Bicoloured
Seedling vigour	0,000	10,343	ab	bc	bc	a	b

Trends in diversity

To investigate the trends in diversity differences in means were analyzed using ANOVA. Table 13 shows the means and standard deviations for the two release type groups. To determine if the differences found between the variables are significant a two-way ANOVA was performed, see Table 14.

Table 13 Means and standard deviations for the release type group: the 1920s onwards and varieties from before the 1920s

Release type group				
	1920s onwards		Before 1920s	
	Mean	Std. Dev.	Mean	Std. Dev.
100 seed weight	40,30	15,76	38,45	17,57
Days to flowering	49,05	4,60	50,65	6,57
Earliness	2,43	1,13	0,80	1,44
Flower colour	1,06	0,23	1,61	0,94
Height	3,53	0,85	3,12	1,06
inflorescence position	2,66	1,02	3,25	1,16
Leaf colour	4,74	1,06	3,75	1,15
Lodging	0,61	0,50	0,88	0,33
Number of seed colours	1,03	0,28	1,25	0,50
Performance	2,72	0,80	2,96	0,89
pod crossection	2,79	1,51	3,27	1,41
Pod length (cm)	14,28	3,62	14,44	3,14
Pod width (mm)	11,08	0,38	11,30	3,11
Seed crosssection	2,94	0,47	2,74	0,90
Seedling vigour	5,04	1,39	5,62	1,29

Table 14 shows the results for the two-way ANOVA performed on the release type group. Significant results were obtained for the variables flower color, inflorescence position, leaf color and lodging. Lodging showed also significant interaction between the blocks and the group. For the other significantly different variables a one-way ANOVA was performed to confirm the results.

Table 14 Two-way ANOVA performed on the variables which were measured for each repetition based on the release type group, significant results have been put in bold. The table indicates the significance and the F-value for the release type group, block and their interaction

	Release type		Block		Interaction	
	P	F	P	F	P	F
Days to flowering	0,072	3,27	**0,000**	31,27	0,628	0,24
Earliness	0,637	0,22	0,739	0,11	0,819	0,05
Flower colour	**0,000**	22,87	0,887	0,02	0,887	0,02
Height	0,181	1,82	**0,000**	0,99	0,135	2,28
Inflorescensce position	**0,016**	5,85	0,340	0,92	0,436	0,61
Leaf colour	**0,000**	41,95	**0,001**	10,86	0,494	0,47
Lodging	**0,000**	22,11	**0,006**	7,63	**0,000**	15,14
Performance	0,151	2,07	**0,012**	6,45	0,150	2,08
Seedling vigour	0,016	5,82	**0,035**	4,49	0,462	0,54

Further analyses were performed within the after 1920s group to investigate if changes occurred over time in the professionally released varieties. Table 15 shows the means and standard deviation for all measured variables per decade. In some decades no variation for certain traits existed, these have been marked with an asterisk.

Table 15 Means and standard deviations for the decade groups, decades in which there was not enough variation to calculate mean or standard deviation there has been placed an asteriks.

Decade group	1940-1950		1950-1960		1960-1970		1970-1980		1980-1990		1990-2000	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
100 seed weight	38,99	12,84	50,89	11,82	37,31	6,48	47,43	11,16	43,53	10,93	43,46	18,08
Days to flowering	48,67	4,61	51,10	3,63	46,50	5,70	45,90	4,31	50,13	2,95	48,67	4,65
Earliness	1,80	1,10	*	*	2,00	1,15	3,00	1,58	2,50	0,87	2,75	0,35
Flower colour	1,42	0,90	*	*	*	*	*	*	*	*	*	*
Height	3,83	1,21	3,50	0,50	*	*	*	*	3,25	0,35	*	*
inflorescence position	2,17	0,75	2,50	0,71	2,63	1,70	2,80	1,04	3,17	1,04	3,25	0,35
Leaf colour	4,21	1,21	4,80	0,45	5,40	0,89	4,90	0,65	5,00	1,35	5,00	0,87
Lodging	0,59	0,32	*	*	0,82	0,22	0,38	0,41	0,47	0,41	.291	0,31
Number of seed colours	1,08	0,29	*	*	*	*	*	*	*	*	*	*
Performance	2,04	0,188	3,37	0,53	2,187	0,625	3,3	0,778	2,75	0,66	3,625	0,53
pod crossection	2,70	1,70	2,75	1,26	3,80	1,10	3,00	1,79	2,40	0,89	*	*
Pod length (cm)	15,73	4,10	13,80	3,49	14,60	4,78	14,17	4,12	13,00	1,87	12,00	1,00
Pod width (mm)	12,73	4,84	10,60	0,89	11,80	4,09	10,67	4,23	9,00	2,12	9,00	3,00
Seed crosssection	3,00	0,60	2,80	1,10	2,60	0,89	3,20	0,45	3,25	0,50	*	*
Seedling vigour	4,87	1,29	5,00	1,78	5,70	1,25	5,33	1,47	5,20	0,76	3,50	1,80

To investigate the differences between the various means a two-way ANOVA was performed. The ANOVA results are shown in Table 16. From Table 16 it becomes clear that there was a significant difference between the lodging and performance in the various decades.

Table 16. Two-way ANOVA as performed on the decade group, on the variables which have been measured for both repetitions, significant results are in bold. The table indicates the significance and the F-value for the release type group, block and their interaction.

	Decade		Block		Interaction	
	P	F	P	F	P	F
Days to flowering	0,692	0,65	**0,001**	11,91	0,621	0,71
Earliness	0,310	1,26	0,811	0,06	0,066	2,39
Flower colour	0,250	1,35	1,000	0,00	1,000	0,00
Height	0,811	0,48	0,427	0,68	**0,060**	2,90
Inflorescensce position	0,672	0,64	0,544	0,38	0,677	0,63
Leaf colour	0,077	2,04	0,007	7,93	0,258	1,35
Lodging	**0,024**	3,12	**0,001**	13,63	0,797	0,47
Performance	**0,000**	9,02	0,152	2,16	0,710	0,59
Seedling vigour	0,851	4,09	0,851	0,44	0,260	1,35

Table 17 shows the results of the SNK test performed on lodging and performance. The letters indicate the statistical significant difference between the decade groups based on the ANOVA performed, if letters are shared it mean there is no statistical difference between those decades. According to the hypothesis a gradual but significant difference would be found between the decades. However, all the decades retain variation for the two investigated traits. The results are insufficient to point towards a trend in the decade groups.

Table 17 Results of the SNK test performed on lodging and performance, indicating the significant differences between the different decades

Decade group							
	1920-1930	1940-1950	1950-1960	1960-1970	1970-1980	1980-1990	1990-2000
Lodging	a	ab	b	b	ab	ab	ab
Performance	*	a	ab	a	b	ab	b

Discussion and conclusions

The discussion is divided into two parts. Part one concerns the analysis of the general patterns of diversity in which the research question entitled 'Analyze the morphological variation in a set of 188 varieties of *Phaseolus vulgaris* L. cultivated in the Netherlands during the last century in a common garden experiment' will be discussed. In part 2 (Trends in diversity) the other two research questions will be discussed namely 'Analyze if changes in traits occurred for varieties, grouped in decades, and cultivated between the 1920s and the 2000s in the Netherlands' and 'Is there a significant difference between undated varieties and varieties cultivated in the Netherlands from the 1920s onwards?'

For both parts it is important to consider the particularities of the experiment set out in the Materials and Methods section. The extreme temperature, virus pressure, and differences in seedling raising are perhaps the most important to keep in mind. Border rows were present at the alongside of both the experiments, with both border rows being quite similar in height: climbing peas and pole beans. However, there were no border rows present at the ends of the beds as the tractor paths were located there, the experiment started directly from these paths. Therefore, it is possible that the border rows have been influenced by an edge effect. Furthermore, sampling size might have influenced the results, including more varieties in the different groups, especially in the decades could have resulted in a different and better understanding of the diversity and its development. Lastly, the experiment was a completely randomized block design, but organized on growth type making it not completely randomized with a possibility that the pole beans had an influence on the bush beans due to sheltering, shade or otherwise. Furthermore, due to not all variables having been measured for all growth types the analysis became more fragmented. All the aforementioned has to be taken into account when assessing the results.

General Patterns of Diversity

The research question" *What is the morphological variation in a set of 188 varieties of Phaseolus vulgaris L. cultivated in the Netherlands during the last century in a common garden experiment?" was answered through using Nei's diversity index.* On the basis of Nei's variation indexes it could be concluded that for most traits there was an even distribution of the characters over the classes. Most of the Nei's variation indexes were above 0,850 suggesting even distribution of the variation over the various classes. For pod width, seed number of colours and flower colour the Nei's index scores were lower, this suggests that the selected group had lower variation for these traits. The width of the pod was often narrow, and most accessions had a low number of seed colours and a white to light purple flower colour.

For further insight into the division of the traits PCAs were performed. The PCAs managed to explain more variance than expected, around 59% was explained in most cases. This is more than

Hennink & Zeven (1999), and Grahic *et al.* (2013), and Freitas *et al.* (2011). The performed PCAs brought more clarity to the dataset.

The first component explained seed shape. For the seed colour group little distinction between groups was to be expected for PCA I as both groups contain different shapes of seeds. The second component related seedling vigour and flower colour. For the seed colour group it was possible to identify the white seeded varieties as they scored mostly negative for this component. In the dataset all white seeded bean varieties had white flowers, Zeven (1997) mentioned that research had consistently shown white seeded varieties to be less vigorous and suffer more from rot during the sprouting phase due to their genetic susceptibility to fungi. This is clearly indicated in graphs 7a and 7b where the white seeded varieties had lower scores for the second component. A one-way ANOVA performed on seedling vigour and colour group clearly indicated that there was a statistical significant difference between the white and other coloured beans. However, beans with a hilum spot had a significantly lower seedling vigour as well. No satisfactory explanation is given by Zeven (1997) or otherwise available. Graphs 6a and 6b clearly indicated that varieties released from the 1920s onwards were more positive for the first component and more negative for the second component. This can be explained because the varieties released from the 1920s onwards were mostly white and less round seeded than the varieties from before the 1920s.

Trends in Diversity

Previous research in Dutch bean germplasm (Hennink & Zeven, 1999) had focused on diversity; however, it did not focus on the change in diversity over time. In the present research this was carried out and in this context to questions were analyzed.

Firstly: *Analyze if changes in traits occurred for varieties, grouped in decades, and cultivated between the 1920s and the 2000s in the Netherland*s

ANOVA analysis comparing the mean values per decade group showed only significant interaction for lodging and performance. The SNK test performed showed that this difference was due to fluctuations between the year groups. However, it was not possible to identify any trends or consistent changes over time. Research on Dutch germplasm and breeding was not able to provide further answers to this question. Zeven (1997) described very shortly for which traits there was natural selection in the natural environment of the bean, and for which traits man has selected. Apart from yield no traits were added by professional breeders as unique traits. However, as Drijfhout (1978b) described, professional breeders have been more rigorous in selecting for traits such as disease resistance. However, literature on breeding techniques from the 1920 to the 2000s is sparse, which makes it difficult to find more patterns in the data.

Secondly: *Is there a significant difference between undated varieties (landraces?) and varieties cultivated in the Netherlands from the 1920s onwards?*

To answer this question an ANOVA was performed which compared the means of the varieties released after 1920s to the means of the varieties from before the 1920s. Significant results were obtained for lodging and performance. Flower color was significantly lighter for the varieties released after the 1920s. Interestingly, also inflorescence position was lower for varieties after the 1920s. The results were contradicting the expectations; it was expected that varieties from after the 1920s would bear their flowers higher to facilitate easier harvesting. Leaf color was significantly darker for the varieties released from the 1920s onwards, which was also influenced by the blocks. This can be explained with the newer varieties being less susceptible to mosaic virus and thus also receiving a higher score for leaf color. Lodging was significantly influenced per release type as well as the block, the two factors also interacted with each other. This can be explained by varieties from the 1920s onwards having a more upright and compact growth to facilitate machine harvest. The influence of the blocks comes from the big summer storm that was experienced. Overall the experiment has showed small differences between the two groups, however, more research would be necessary to indicate if these changes

Conclusions

This research has shown that Dutch bean germplasm lacks clear boundaries when considering morphological diversity: a conclusion which was also reached by Hennink & Zeven (1999). Apart from this it ties in with similar research performed on the Iberian Peninsula, Greece, Bosnia Herzegovina and Italy on morphological diversity of bean varieties and their origin. The majority of the European common bean germplasm originates from the Andean gene pool. This is illustrated by literature on Italian cultivars (Sicard *et al.*, 2005; Freitas *et al.*, 2011; Lioi *et al.*, 2012; Lioi *et al.* 2005), Greek and cultivars from the Iberian Peninsula (Motley *et al.*, 2013) and Bosnian Herzegovinan cultivars (Grahic *et al.*, 2013). Furthermore, literature has confirmed extensive hybridization between the Meso American and the Andean gene pool in Europe (Lioi *et al.*, 2012).

Many researchers have failed to establish groups in common bean collections using morphological data (Hennink & Zeven, 1999; Grahic *et al.*, 2013; Freitas *et al.* 2011), this was also shown in this research. However, the grouping based on seed color, and a one-way ANOVA confirmed statements by Zeven (1997) that white seeded varieties are lower in vigour. However, it did also show that beans with a hilum spot were scored lower for seedling vigour, for which no apparent explanation was available. Furthermore, the PCA analysis suggested that the varieties released after the 1920s might be a subset of the varieties from before the 1920s (figure 6a and 6b).

Recommendations for further research

For further research it is important to first, raise accessions which are suspected to have virus, inside a greenhouse with insect proof caging. The first seedling raising can thus be guaranteed to be virus free when sown outside the field. In this experiment the plants were raised inside an insect proof cage in a greenhouse, however, as soon as they were planted out a few plants started showing signs of virus that were missed when cleaning out the seedlings. This virus spread rapidly and infected most of the plants. This would have been less of a problem if we would have started with virus free seed.

For optimal germination it is important to sow the bean varieties in unfertilized potting soil. Our experience has shown that fertilized potting soil keeps the moisture longer increasing the rotting risk of the seeds.

Furthermore, future research in Dutch bean germplasm will need to answer why beans with hilum spots are as likely or even more likely than white seeded beans to a getting lower scores for seedling vigour. Genomic research might also provide extra answers through clarifying the origin of the different bean varieties as performed by Zeven *et al.* (1999). Lastly, future research into developments over time should include more varieties over time to make a more secure investigation into trends. If combined with yield it might provide a more accurate picture of the work done by Dutch breeders on the germplasm.

References

Arruda, B., Guidolin, A.F., Coimbra, J.L.M., Battilana, J. (2012). Environment is crucial to the cooking time of beans. *Cienca e Tecnologia de Alimentos, Campinas* 32(2): 537-578

Beattie, A.D., Michaels, T.E., Pauls, K.P. (2003). Predicting progeny performance in common bean (Phaseolus vulgaris L.) using molecular marker based cluster analysis. *Genome* 46: 259-267

Bennett, S., Cocks, P.S. & Summerfield, R.J. (ed.) (1999). Genetic resources of Mediterranean pasture and forage legumes. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Bitochhi E., Nanni, L., Bellucci, E., Rossi, M., Giardini, A., Zeuli, P.S., Logozzo, G., Stougaard, J., Mcclean, P., Attene, G. Papa, R. (2012). Mesoamerican origin of the common bean (*Phaseolus vulgaris* L.) is revealed by sequence data. *PNAS* E788 – E796.

Buishand, J.T. (1959). Rassenonderzoek bij bonen. *Proefstation voor de groenteteelt in de volle grond.*

CBS (2015). Population, households and population dynamics; from 1899. Retrieved from: http://statline.cbs.nl/Statweb/publication/?VW=T&DM=SLEN&PA=37556eng&D1=0-44&D2=1,11,21,31,41,51,61,71,81,91,101,l&HD=150129-1350&LA=EN&HDR=G1&STB=T Retrieved on 04-09-2015

Diederichsen, A., Solberg, S.Ø., Jeppson, S. (2013). Morphological changes in Nordic spring wheat (*Triticum aestivum L.*) landraces and cultivars released from 1892 to 1994. *Genetic Resources and Crop Evolution*. 60(2): 569-585.

Dodde, H. (2014). Zaadfirma zoekt zesde gewas. *De Nieuwe Oogst.* Retrieved January 9 from: http://www.popvriendseeds.com/images/64116_PDF.pdf

Drijfhout, E. (1978a). Ten years collaborative bean research IVT - CIAT in IVT : activities and results. *IVT rapport* 242

Drijfhout, E. (1978b). Genetic interaction between Phaseolus vulgaris and bean common mosaic virus with implications for strain identification and breeding for resistance. *Verslagen van landbouwkundige onderzoekingen der rijkslandbouwproefstations* 872

FAOSTAT (2015). Food and Agriculture Organization of the United Nations, FAOSTAT database Retrieved from: http://faostat3.fao.org/home/E

FAO (2016). International Year of Pulses. www.fao.org. Retrieved January 9 from: http://www.fao.org/pulses-2016/en/

Freitas, G., Gananca, J.F.T., Nobrega, H., Nunes, E., Costa, G. Slaski, J.J., Pinheiro de Carvalho, M.A.A. (2011). Morphological evaluation of common bean diversity on the Island of Madeira. *Genetic Resources and Crop Evolution* 58(6): 861-874

Gezondheidsraad (2015). Richtlijnen goede voeding 2015. Retrieved January 9 from: http://www.gezondheidsraad.nl/sites/default/files/201524_richtlijnen_goede_voeding_2015.pdf

Govindaraj, M., Vetriventhan, M., Srinivasan, M. (2015). Importance of Genetic Diversity Assessment in Crop Plants and Its Recent Advances: An Overview of Its Analytical Perspectives. *Genetics Research International* 1- 14.

Graham, P.H., Vance, C.P. (2003). Legumes: Importance and Constraints to Greater Use. *Plant Physiology* 131(3): 872 -877

Grahic, J., Gasi, F., Kurtovic, M., Karic, L., Dikic, M., Gedzo, D. (2013). Morphological evaluation of common bean diversity in Bosnia and Herzegovina using the discriminant analysis of principal components (DAPC) multivariate method. *Genetika* 45 (3) 963-977.

Hennink, S., Zeven, A.C. (1991). The interpretation of Nei and Shannon-Weaver within population variation indices. *Euphytica* 51: 235-240.

Holland Select (2016). About us. History. *Holland Select*. Retrieved January 9 from: http://www.holland-select.com/about-us/history

Lioi, L., Piergiovanni, A.R., Pignonne, D., Puglisi, S., Santantionio, M., Sonnante, G. (2005). Genetic diversity of some surviving on-farm Italian common bean (*Phaseolus vulgaris* L.) landraces. *Plant Breeding* 124: 576-581

Lioi, L., Nuzzi, A., Campion, B., Piergiovanni, A.R. (2012). Assessment of genetic variation in common bean (*Phaseolus vulgaris* L.) from Nebrodi mountains (Sicily, Italy). *Genetic Resources and Crop Evolution* 59(3): 455-464

Luquet, E., Garner, T.W.J., Léna, J.P., Bruel, C., Joly, P., Lengagne, Grolet, O., Plénet, S. (2012). Genetic Erosion in Wild Populations Makes Resistance to a Pathogen More Costly. *Evolution* 66(6): 1942-1952

Messina, V. (2014). Nutritional and health benefits of dried beans. *American Journal of Clinical Nutrition* 100(1): 437S – 442S

Mamidi, S., Rossi, M., Moghaddam, S.M., Annam, D., Lee, R., Papa, R., McClean, P.E.(2013). Demographic factors shaped diversity in the two gene pools of wild common bean *Phaseolus vulgaris* L. *Heredity* 110(3): 267 - 276

Mamidi, S., Rossi, M., Annam, D., Moghaddam, S.M., Lee, R., Papa, D., McClean, P. (2011). Investigation of the domestication of common bean (*Phaseolus vulgaris*) using multilocus sequence data. *Functional Plant Biology* 38 (12): 953 - 967

McClean, P.E. (2011). Crop improvement in the era of climate change: an integrated, multi-disciplinary approach for common bean (*Phaseolus vulgaris)*. *Functional Plant Biology* 38 (12): 927 - 933

Motley, T.J., Zerega, N., Cross, H. (2006). Darwin's Harvest: New Approaches to the origins, Evolution and Conservation of Crops. *Columbia University Press*. New York.

Rogers, J. (2012). No title. *Special Uitgangsmateriaal*. 38 (3) 23-24. Retrieved January 9 from: http://www.popvriendseeds.com/images/documents/bb_maart_2012_special_pop_vriends.pdf

Sicard, D. Nanni,L. Porfiri, O., Bulfon, D., Papa, R. (2005). Genetic diversity of *Phaseolus vulgaris* L. and *P. coccineus* L. landraces in central Italy. *Plant Breeding* 124: 464-472.

Siddiq, M., Butt, M.S., Sultan, M.T. (2011). Dry Beans: Production, Processing, and Nutrition. *Chapter 27 Handbook of Vegetables and Vegetable Processing.* Sinha, N. (ed.) Singapore: Blackwell Publishing 2011. 771. 1st edition.

Singh, S.P., Gepts, P., Debouck, D.G. (1991). Races of Common Bean (*Phaseolus vulgaris*, Fabaceae). *Economic Botany* 45(3): 379-396

Timmer, R.D. (2012). Verbetering ketenresultaat door beter uitgangsmateriaal bruine bonen. Stichting Dienst Landbouwkundig Onderzoek, PPO rapport 487, pp. 1-12

UPOV (2016). What is UPOV? *UPOV* Retrieved January 9 from: http://www.upov.int/overview/en/upov.html

Velema, J. (2015). [untitled overview of release dates of commercial germplasm]. Unpublished data.

Willemsen, J. (2015). Juli 2015: van hittegolf tot zomerstorm. *Meteovista*. Retrieved January 9 from: http://nieuws.weeronline.nl/29-7-2015-juli-2015-van-hittegolf-tot-zomerstorm/

Wouw, M. van den., Hintum, T., Chris, K., Treuren, R., Visser, B. (2010). Genetic diversity trends in twentieth century crop cultivars: a meta-analysis. *Theoretical and Applied Genetics* 120: 1241 -1252

Zeven, A.C. (1997). The Introduction of the Common Bean (*Phaseolus vulgaris* L.) into the Western Europe and the Phenotypic Variation of Dry Beans Collected in the Netherlands in 1946. *Euphytica* 94: 319-328

Zeven, A.C., Waninge, J., van Hintum, Th, Singh, S.P. (1999). Phenotypic variation in a core collection of common bean (*Phaseolus vulgaris* L.) in the Netherlands. *Euphytica* 109: 93-106.

Appendix I: Overview of all varieties, growth and consumption type, seed colour group & release date.

Experiment 1	Experiment 2	Variety	Pod group		Colour group	Experiment 1	Experiment 2	Variety	Pod group		Colour group
6,701	9,538	Bernardo	French pole bean	1991	white	6,753	9,507	Veense Snijboon Venray	Pole bean	prior to the 1920s	white
6,702	9,546	Cobra	French pole bean	prior to the 1920s	even colour	6,754	9,520	Helda	Pole bean	prior to the 1920s	white
6,703	9,515	Trebona (kas)	French pole bean	prior to the 1920s	white	6,755	9,526	Wonder van Venetie z dr	Pole bean	prior to the 1920s	even colour
6,704	9,568	Farba	French pole bean	1983	white	6,757	9,533	Gritse Kievitsboon HS	Dry pole bean	prior to the 1920s	striped/spotted
6,705	9,519	Mentor	French pole bean	1948	white	6,758	9,516	Stokkievitsboon Aalsmeer	Dry pole bean	prior to the 1920s	striped/spotted
6,706	9,514	Aromata	French pole bean	1953	white	6,759	9,567	Stokkievitsboon Hoogmade	Dry pole bean	prior to the 1920s	striped/spotted
6,707	9,545	Rakker	French pole bean	1979		6,761	9,524	Odriscal HS	Dry pole bean	prior to the 1920s	striped/spotted
6,708	9,560	Rinox	French pole bean	1983	white	6,762	9,542	Patrijzenoogjes	Dry pole bean	prior to the 1920s	even colour
6,709	9,506	Koning der Belgen	French pole bean	prior to the 1920s	even colour	6,763	9,564	Rieser boon	Dry pole bean	prior to the 1920s	white
6,710	9,508	Mechelse Markt	French pole bean	1951	white	6,764	9,556	Stok Kievitsboon (Borlotti)	Dry pole bean	prior to the 1920s	striped/spotted
6,711	9,547	Mechelse Tros z dr (syn. M.markt)	French pole bean	prior to the 1920s	white	6,765	9,505	Citroenboontje	Dry pole bean	prior to the 1920s	even colour
6,712	9,530	Non Plus Ultra	French pole bean	1926	white	6,766	9,511	Rattekeutel Barneveld	Dry pole bean	prior to the 1920s	even colour
6,713	9,562	Overvloed Rentegevers	French pole bean	prior to the 1920s	white	6,767	9,518	Zwols Heilig Stokboontje	Dry pole bean	prior to the 1920s	hilum spot
6,714	9,529	Rentegevers z dr	French pole bean	prior to the 1920s	white	6,768	9,509	Ying Yang HS	Dry pole bean	prior to the 1920s	Bicolour
6,715	9,521	Precores	French pole bean	1969	white	8,801	9,858	Arosa	Bush bean	unknown	white
6,716	9,557	Bolts stokslaboon	French pole bean	prior to the 1920s	white	8,802	9,948	Cantare	Bush bean	1991	white
6,717	9,555	Dreuge Boon Hinrichs Reuzen Groningen	French pole bean	prior to the 1920s	white	8,803	9,812	Flair	Bush bean	1965	white
6,718	9,528	Groninger Dreuge Boon	French pole bean	prior to the 1920s	white	8,804	9,880	Forum	Bush bean	1991	white
6,719	9,534	Oldambster whitete	French pole bean	prior to the 1920s	white	8,805	9,833	Groffy	Bush bean	1976	white
6,720	9,535	Korte Trosjes 1989	French pole bean	prior to the 1920s	white	8,806	9,888	Koudumer boontje (Impala)	Bush bean	1968	white
6,721	9,531	Koudumer Stokboon	French pole bean	prior to the 1920s	white	8,807	9,893	Montano	Bush bean	1985	white
6,722	9,517	Lange Stokslaboon 1989	French pole bean	prior to the 1920s	even colour	8,808	9,807	Prelubel	Bush bean	1967	white
6,723	9,551	Limburgse Vit	French pole bean	prior to the 1920s	white	8,809	9,901	Prelude	Bush bean	1955	white
6,724	9,544	Knobbelboontje	French pole bean	prior to the 1920s	white	8,810	9,922	Sonate	Bush bean	1992	white
6,725	9,550	Ohnegleichen	French pole bean	prior to the 1920s	even colour	8,811	9,867	Verbeterde Dubbele, Centrum	Bush bean	1968	white
6,726	9,532	Westlandia	French pole bean	1941	white	8,812	9,818	Widusa	Bush bean	1955	white
6,727	9,549	Zwolse Paarse Stokboon	French pole bean	prior to the 1920s	even colour	8,813	9,925	Bovensmildese Dubbele whitete 1973	Bush bean	1973	white
6,728	9,504	Epser Spekboon	French pole bean	prior to the 1920s	white	8,814	9,824	Dubbele whitete z dr	Bush bean	1943	white
6,731	9,539	Phenomeen	French pole bean	1943	white	8,815	9,906	Chicobel	Bush bean	1971	white
6,732	9,552	Markant	French pole bean	prior to the 1920s	white	8,816	9,829	Chicobel	Bush bean	1971	white
6,733	9,512	Mombacher Speck	French pole bean	1943	white	8,817	9,940	Conserva Konserva	Bush bean	1925	white
6,734	9,501	Boterboon whitez. 1988	French pole bean	prior to the 1920s	white	8,818	9,932	Dubresco	Bush bean	1986	white
6,735	9,559	Zwarte Stokslaboon 2010	French pole bean	prior to the 1920s	even colour	8,819	9,838	Knobbelboon Roggel	Bush bean	prior to the 1920s	white
6,736	9,566	Friese Gele boon Sytske	French pole bean	prior to the 1920s	even colour	8,820	9,945	Rijstkorreltjes	Bush bean	prior to the 1920s	white
6,737	9,563	Gele Posthoorn	French pole bean	prior to the 1920s	even colour	8,821	9,830	white-Roodboon Barneveld	Bush bean	prior to the 1920s	striped/spotted
6,749	9,565	Raadsheer	Pole bean	prior to the 1920s	white	8,822	9,854	whitete Stamboon Avenhoorn	Bush bean	prior to the 1920s	white
6,750	9,503	Raadsheer (nabouw Veense)	Pole bean	1943	white	8,823	9,878	Drentse Rode Weekschil	Bush bean	prior to the 1920s	striped/spotted
6,751	9,561	Stoksnijboon 1988	Pole bean	prior to the 1920s	white	8,824	9,885	Gieterse Slaboon Gieten	Bush bean	prior to the 1920s	striped/spotted
6,752	9,513	Veense	Pole bean	1943	white	8,825	9,875	Giethoornse weekschil	Bush bean	prior to the 1920s	striped/spotted

Experiment 1	Experiment 2	Variety	Pod group		Colour group	Experiment 1	Experiment 2	Variety	Pod group		Colour group
8,826	9,802	Giethoornse weekschil	Bush bean	prior to the 1920s	striped/spotted	8,871	9,910	Beka	Dry bush bean	prior to the 1920s	even colour
8,827	9,924	Giethoornse Weekschil	Bush bean	prior to the 1920s	striped/spotted	8,872	9,834	Berna IvP	Dry bush bean	prior to the 1920s	even colour
8,828	9,868	Groninger Weekschil	Bush bean	1943	striped/spotted	8,873	9,810	Berna van Namen	Dry bush bean	prior to the 1920s	even colour
8,829	9,913	Groninger Weekschil	Bush bean	1943	striped/spotted	8,874	9,920	Bovensmildese Bruine Boon 1973 HS	Dry bush bean	prior to the 1920s	even colour
8,830	9,846	Hinrich Riesen	Bush bean	prior to the 1920s	striped/spotted	8,875	9,826	Bruine Boon Zeeland	Dry bush bean	prior to the 1920s	even colour
8,831	9,891	Padleggers	Bush bean	prior to the 1920s	white	8,876	9,851	Bruine Rotjes Breezand	Dry bush bean	prior to the 1920s	even colour
8,832	9,939	Rôner Weekschil 2011	Bush bean	prior to the 1920s	striped/spotted	8,877	9,923	Bruine Soepboon Venray	Dry bush bean	prior to the 1920s	even colour
8,833	9,947	Rôner Weekschil 2014	Bush bean	prior to the 1920s	striped/spotted	8,878	9,831	Bruine Woudboon	Dry bush bean	prior to the 1920s	even colour
8,834	9,929	Roner Weekschil	Bush bean	prior to the 1920s	striped/spotted	8,879	9,886	Ceka IVT	Dry bush bean	prior to the 1920s	even colour
8,835	9,876	Gijsje Teunissen Lunteren	Bush bean	prior to the 1920s	even colour	8,880	9,921	Dikpenskes Baflo	Dry bush bean	prior to the 1920s	even colour
8,836	9,936	Goudhoorntje	Bush bean	prior to the 1920s	even colour	8,881	9,949	Enkhuizer Kleine 2013	Dry bush bean	prior to the 1920s	even colour
8,837	9,904	Honderd voor Een (Sans Rival)	Bush bean	prior to the 1920s	even colour	8,882	9,856	Goudkorrel	Dry bush bean	prior to the 1920s	even colour
8,838	9,953	Limburgische	Bush bean	prior to the 1920s	even colour	8,883	9,907	Goudkorrel (kogelboon)	Dry bush bean	prior to the 1920s	even colour
8,839	9,938	Saxa- Verbeterde Wagenaars z dr	Bush bean	1943	even colour	8,884	9,890	Graddus Reuzen (Friese Gele)	Dry bush bean	prior to the 1920s	even colour
8,840	9,848	Saxa, Verb. Wagenaar	Bush bean	prior to the 1920s	even colour	8,885	9,895	Koffiebruine boon Vlaardingen	Dry bush bean	prior to the 1920s	even colour
8,842	9,835	Epser Wasboon	Bush bean	prior to the 1920s	white	8,886	9,950	Kogelronde Bruine Boon Wageningen	Dry bush bean	prior to the 1920s	even colour
8,843	9,862	Gele Boterboon	Bush bean	prior to the 1920s	even colour	8,887	9,918	Leverkleurige boon	Dry bush bean	prior to the 1920s	even colour
8,846	9,866	Oktoberli	Bush bean	prior to the 1920s	striped/spotted	8,888	9,903	Noordhollandse Bruine	Dry bush bean	prior to the 1920s	even colour
8,847	9,827	Admires	Bush bean	1970	white	8,889	9,828	Noordhollandse Bruine	Dry bush bean	prior to the 1920s	even colour
8,848	9,811	Antwerpse Noordster	Bush bean	prior to the 1920s	white	8,890	9,905	Noordhollandse Bruine	Dry bush bean	prior to the 1920s	even colour
8,849	9,801	Koudumer Stamsnijboon	Bush bean	prior to the 1920s	white	8,891	9,915	Paaseitjes (kogel)	Dry bush bean	prior to the 1920s	even colour
8,850	9,842	Minidor	Bush bean	1987	white	8,892	9,809	Rosse Boon (kogel)	Dry bush bean	prior to the 1920s	even colour
8,851	9,944	Aka whitezadig	Dry bush bean	prior to the 1920s	white	8,893	9,819	Walcherse Bruine Kogelboon	Dry bush bean	prior to the 1920s	even colour
8,852	9,931	Dikkontjes s'Gravenhage	Dry bush bean	prior to the 1920s	white	8,894	9,822	Walcherse Bruine Kogelboon	Dry bush bean	prior to the 1920s	even colour
8,853	9,911	Eiboon Zevenbergschehoek	Dry bush bean	prior to the 1920s	white	8,895	9,803	Zeeuwse Grootzadige Bruine boon	Dry bush bean	prior to the 1920s	even colour
8,854	9,927	Kaboon IPO	Dry bush bean	prior to the 1920s	white	8,896	9,916	Bruine Kogelboon	Dry bush bean	prior to the 1920s	even colour
8,855	9,850	Vroege Wagenaar	Dry bush bean	1943	even colour	8,899	9,852	Blokkerder Boon HS	Dry bush bean	prior to the 1920s	even colour
8,856	9,853	Walcherse whitete	Dry bush bean	prior to the 1920s	white	8,900	9,882	Friese Woudboon	Dry bush bean	prior to the 1920s	even colour
8,858	9,836	Epser	Dry bush bean	prior to the 1920s	white	8,901	9,884	Gele Krobbe	Dry bush bean	prior to the 1920s	even colour
8,859	9,902	Laagpeulige Krombek	Dry bush bean	prior to the 1920s	white	8,902	9,909	Giele Waldsbeantsje	Dry bush bean	prior to the 1920s	even colour
8,860	9,820	whitete Krombek	Dry bush bean	prior to the 1920s	white	8,903	9,843	Groninger Strogele	Dry bush bean	prior to the 1920s	even colour
8,861	9,926	whitete Krombek	Dry bush bean	prior to the 1920s	white	8,904	9,814	Groninger Strogele	Dry bush bean	prior to the 1920s	even colour
8,862	9,946	whitete Krombek	Dry bush bean	prior to the 1920s	white	8,905	9,908	Groninger Strogele	Dry bush bean	prior to the 1920s	even colour
8,863	9,933	whitete Krombek	Dry bush bean	prior to the 1920s	white	8,906	9,825	Strogele boon	Dry bush bean	prior to the 1920s	even colour
8,864	9,804	whitete Krombek	Dry bush bean	prior to the 1920s	white	8,907	9,934	Citroenboon Texel	Dry bush bean	prior to the 1920s	even colour
8,865	9,917	whitete Krombek	Dry bush bean	prior to the 1920s	white	8,908	9,823	Citroenboon van Namen	Dry bush bean	prior to the 1920s	even colour
8,866	9,877	whitete Krombek	Dry bush bean	prior to the 1920s	white	8,909	9,870	Citroenboon Venray	Dry bush bean	prior to the 1920s	even colour
8,867	9,857	whitete Krombek	Dry bush bean	prior to the 1920s	white	8,910	9,952	Gele Citroen	Dry bush bean	prior to the 1920s	even colour
8,870	9,879	Aka LU-LPT	Dry bush bean	prior to the 1920s	even colour	8,911	9,860	Reade Krobbe Friesland	Dry bush bean	prior to the 1920s	striped/spotted

Experiment 1	Experiment 2	Variety	Pod group		Colour group
8,912	9,937	Reade Krobbe HS	Dry bush bean	prior to the 1920s	striped/spotted
8,913	9,855	Reade Krobbe HS	Dry bush bean	prior to the 1920s	striped/spotted
8,914	9,859	Bonties WS (platte bonte ranker)	Dry bush bean	prior to the 1920s	striped/spotted
8,915	9,849	Breezandse Boon	Dry bush bean	prior to the 1920s	striped/spotted
8,916	9,871	Drentse Kievitsboon HS	Dry bush bean	prior to the 1920s	striped/spotted
8,917	9,806	Grootzadige Kievitsboon Wouda	Dry bush bean	prior to the 1920s	striped/spotted
8,918	9,894	Grootzadige Paarse Kievitsboon Tubergen	Dry bush bean	prior to the 1920s	striped/spotted
8,919	9,881	Kievistboon Batenburg	Dry bush bean	prior to the 1920s	striped/spotted
8,920	9,935	Kievitsbohne CH-VD	Dry bush bean	prior to the 1920s	striped/spotted
8,921	9,840	Kievitsboon Druten	Dry bush bean	prior to the 1920s	striped/spotted
8,922	9,943	Kievitsboon Texel	Dry bush bean	prior to the 1920s	striped/spotted
8,923	9,951	Musseneitjes	Dry bush bean	prior to the 1920s	striped/spotted
8,924	9,942	Paaseitjes	Dry bush bean	prior to the 1920s	striped/spotted
8,925	9,930	Patrijzenoogjes	Dry bush bean	prior to the 1920s	striped/spotted
8,926	9,844	Renka	Dry bush bean	prior to the 1920s	striped/spotted
8,927	9,919	Renka LU-LPT	Dry bush bean	prior to the 1920s	striped/spotted
8,928	9,896	Kievitsboon	Dry bush bean	prior to the 1920s	striped/spotted
8,929	9,805	Transvaalse Bonte m dr LU-LPT	Dry bush bean	1943	striped/spotted

Appendix II:

Table 18: PCA results on the pole bean subset

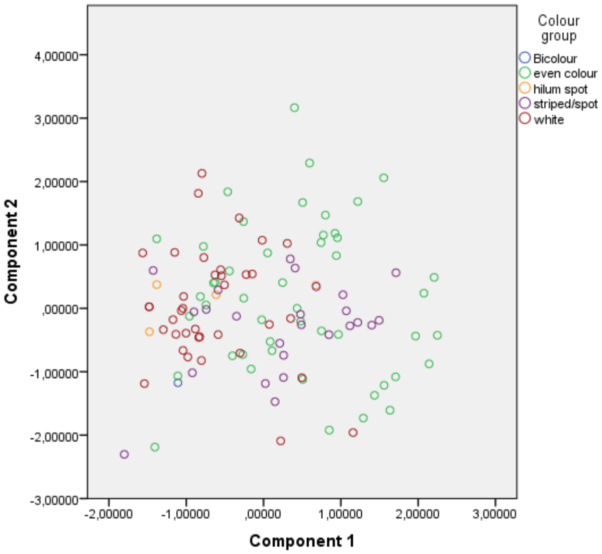
	Component 1	Component 2	Component 3
Variance explained (%)	24,819	21,069	15,509
Eigen value	1,985	1,686	1,241
Seedling vigour	-,097	,788	-,070
Flower colour	,253	-,671	-,368
Days to flowering	-,708	-,054	,359
Leaf colour	,192	,012	-,534
Number of seed colours	-,679	,404	,358
Seed cross section	,569	,390	,594
Seed longitudinal section	,761	-,038	,450
Height	0,097	-,544	0,058

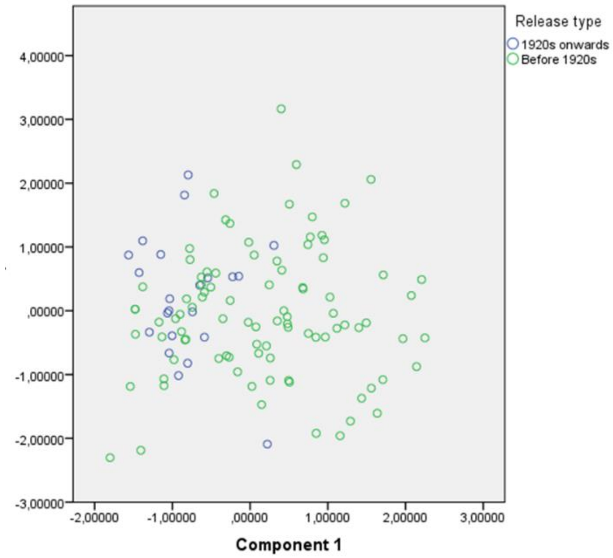
Table 1915: PCA results on the subset of bush beans

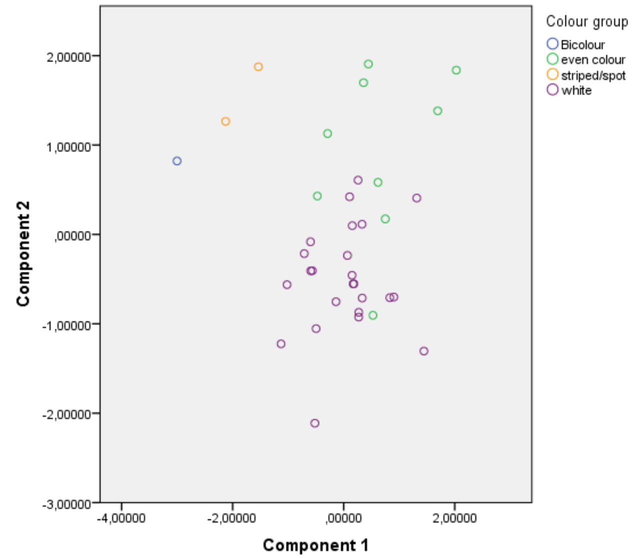
	Component 1	Component 2	Component 3
Variance explained (%)	20,779	17,265	14,006
Eigen value	2,078	1,727	1,401
Seedling vigour	,490	,486	-,384
Flower colour	,539	,324	-,308
Days to flowering	,234	-,449	,455
Leaf colour	-,148	,224	,700
Number of seed colours	,037	-,191	-,061
Seed cross section	-,562	,576	,210
Seed longitudinal section	-,578	,599	-,022
Performance 1	,636	,522	,079
Performance 2	,470	,319	,617
Lodging (%)	,442	-,179	,228

Table 160: PCA results French beans subset

	Component 1	Component 2	Component 3
Variance explained (%)	21,696	20,394	14,611
Eigen value	2,604	2,447	1,753
Seedling vigour	,620	,252	-,431
Flower colour	,668	-,091	-,035
Days to flowering	-,232	,080	,840
Leaf colour	-,704	,066	-,402
Number of seed colours	,625	,077	,533
Seed cross section	-,507	,695	,006
Seed longitudinal section	-,110	,598	-,261
Performance 1	,204	,570	,404
Performance 2	-,520	,320	,371
Lodging (%)	-,134	-,376	-,057
Pod width (cm)	,071	,774	-,187
Pod length (cm)	,479	,592	-,085


Colour group
Bicolour
even colour
hilum spot
striped/spot
white
Component 2
Component 1


Release type
1920s onwards
Before 1920s
Component 1


Colour group
Bicolour
even colour
striped/spot
white
Component 2
Component 1

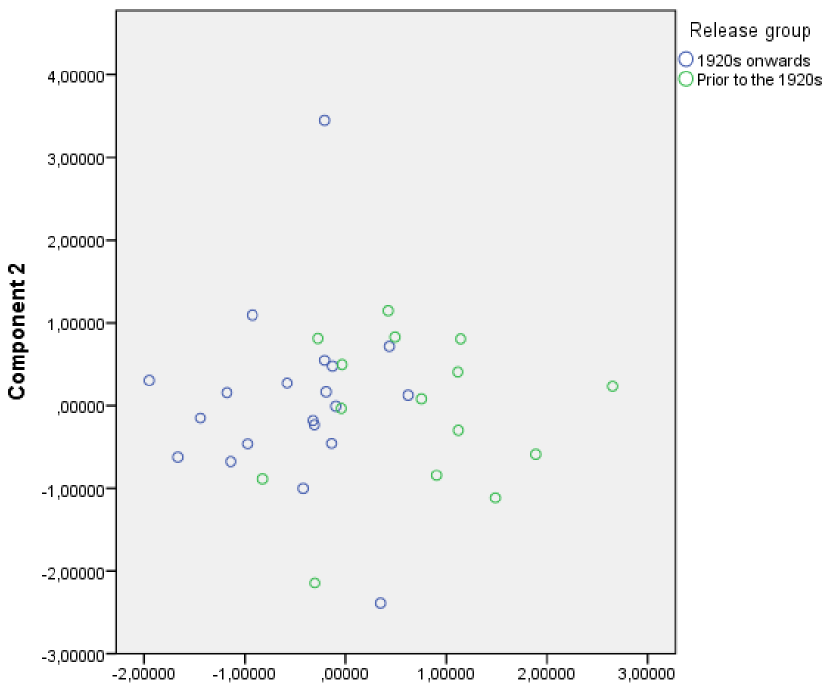


Figure 11: PCA results for the French bean data subset coloured on release group

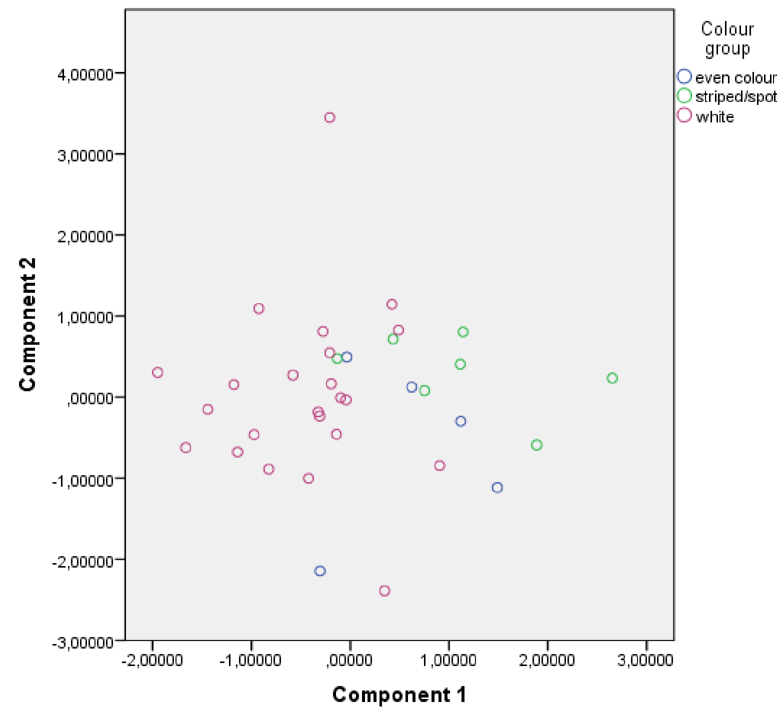


Figure 12: PCA results for the French bean data subset coloured on colour group