UNIVERSITY OF HOHENHEIM

Institute of Crop Science
Biostatistics Unit (340c)
Prof. Dr. Hans-Peter Piepho

Improvement of breeding strategies for the trait vase life in cut carnations (*Dianthus caryophyllus L.*)

Dissertation

in fulfillment of the requirements for the degree

“Doktor der Agrarwissenschaften”

(Dr. sc. agr. / Ph. D. in Agricultural Sciences)

submitted to the

Faculty of Agricultural Sciences

Maike Boxriker (Msc.)

born in Schwäbisch Gmünd

Stuttgart – Hohenheim

2018

This thesis was accepted as a doctoral dissertation in fulfillment of the requirements for the degree “Doktor der Agrarwissenschaften” by the Faculty of Agricultural Sciences at the University of Hohenheim on 29th March 2018.

Date of oral examination: 08th October 2018

Examination Committee:

Supervisor and Reviewer: Prof. Dr. Hans-Peter Piepho

Co-reviewer: Prof. Dr. Thomas Debener

Additional examiner: PD Dr. Tobias Würschum

Head of the Committee: Prof. Dr. Jörn Bennewitz

Content

List of figures

List of tables

Abbreviations

σ_g^2	Genotypic variance
ABA	Abscisic acid
ACC	1-aminocyclopropane-1-carboxylic acid
ADP	Adenosindiphosphat
AIC	Akaike information criterion
ATP	Adenosintriphosphat
bp	Base pair
BSA	Bulked segregant analysis
BSR-Seq	Combination of RNA-Seq and bulked segregant analysis
cDNA	complementary DNA
CF	Count of flowers stem^{-1}
D1, D2	Experimental design in the first phase
DE	Differentially expressed / differential expression
DNA	Deoxyribonucleic acid
E1, E2, E3	Experiments 1, 2 and 3 conducted in three consecutive years
EST	Expressed sequence tag sequence
FD	Flower diameter
GA	Gibberellin
GO	Gene ontology
GS	Genomic selection
H^2	Heritability
JA	Jasmonic acid
L1, L2	Experimental design in the second phase
LSD	Least significant difference
MACE	Massive analysis of cDNA ends
Mn	Mini carnation
MYC2	Transcription factor

N50	Similar to a mean or median of lengths, but has greater weight given to the longer contigs
NADPH	Nicotinamidadenindinukleotidphosphat
NCBI	National Center for Biotechnology Information
NCED	9-cis-epoxycarotenoid dioxygenase
NGS	Next generation sequencing
P1	First phase of the experiment (greenhouse)
P2	Second phase of the experiment (laboratory)
PA	Polyamine
QTL	Quantitative trait locus
R	Response to selection
RCBD	Randomized complete block design
RNA	Ribonucleic acid
RNA-Seq	RNA sequencing
RSE	Relative selection efficiency
S	Selection differential
SED	Standard error of a difference between means
SL	Stem length
SNP	Single nucleotide polymorphism
St	Standard carnation
VL	Vase life

1. General introduction

1.1 The species *Dianthus caryophyllus L.*

Carnation (*Dianthus caryophyllus* L.), a member of the family *Caryophyllaceae*, is one of the major cut flower species around the world. The genus *Dianthus* includes about 300 species in the *Caryophyllaceae* family. It comprises several economically important species like *Dianthus caryophyllus*, *D. barbatus*, *D. chinensis*, *D. plumarius*, *D. superbus* and their hybrids which are used as floricultural crops (Tanase et al., 2012). *Dianthus caryophyllus* originated in the Mediterranean region has 2,000 years of assumed breeding history and are used as potted plants or cut flowers. Cut carnations are one of the ten most popular cut flowers worldwide. The company 'Klemm + Sohn' is well known for breeding of vegetative propagated ornamentals. The company is producing more than 50 million young plants per year for worldwide sales. 'Klemm + Sohn' possesses an assortment of more than 1,500 carnation varieties of hybrid nature, which can be divided into single-headed standard carnations and multi-headed mini carnations. The development of mini carnations arose out of a standard carnation with strong side shoots in the 1950ies (Köhlein, 1990). Standard carnations are characterized by a single big flower (approximately 7 cm in diameter) per stem and is produced by removal of lateral flower buds. The mini carnation type has multiple smaller flowers (2.5 to 5 cm diameter) and is produced by removing the terminal flower bud (Mc George and Hammett, 2012). Carnations are diploid with a chromosome set of n=15 and with a small genome size of 1,223 mega base pairs. Carnations are highly heterozygous, the level of heterogeneity in the seed population is high as well. The reduced seed viability and low seed germination leads to a vegetative propagation (Dole and Wilkins, 1999).

1.2 Vase life of cut carnations

Vase life (VL) is one of the important breeding objectives in carnations due to its direct influence on consumer satisfaction resulting in a repeated purchase decision (Onozaki et al., 2001). VL is defined as the duration from putting stems into vases to the loss of ornamental value due to wilting, rolling-in of petals or discoloration of the petal margins (Satoh et al., 2005). After harvesting, physiological

changes occur due to respiration, transpiration and biosynthesis. Microbial deterioration limits the VL and storability. VL is regulated by a manageable number of genes especially in the case of biosynthesis of ethylene, a gas that regulates senescence of flowers (Fu et al., 2011; Licausi et al., 2013; van Doorn and Woltering, 2008). By using transgenic approaches, it could be proven that repression of ethylene biosynthesis or blocking of ethylene receptors leads to an improved VL (Satoh, 2011). In the following, reasons for the end of VL are described.

1.2.1 Effect of water stress

VL is limited due to gradually reduced water uptake during floral development, caused by the obstruction of xylem vessels. Microbial growth, deposition of lignin, suberin or tannin in the lumen of xylem vessels, presence of air emboli in the vascular system and formation of tyloses could be the reason for reduced water uptake (Twumasi et al., 2005; van Doorn, 1996). Bacteria in the water can enter into the vascular system, but it is not possible for the bacteria to pass the pit membrane and water uptake is obstructed. The same happens to air, which enters in the plant due to vacuum that is still present after cutting the stems from the stock. To reduce the number of bacteria, biocides can be applied to the water. To reduce the amount of air emboli the cut height must be adjusted and the stems should be put into water directly. The use of cold water could be beneficial (van Meeteren, 1992).

1.2.2 Plant hormones affecting VL

The gaseous plant hormone ethylene is the simplest unsaturated hydrocarbon ($CH^2=CH^2$). The function as senescence inducing component was first described more than 100 years ago. For a long time illuminating gas was used for lighting in homes, businesses and streetlights. In 1858, George Fahnestock observed first that leaking illuminating gas was affecting plants in greenhouse. Newton and Cook (1927) demonstrated that ethylene was responsible for a higher respiration of bananas. With ^{14}C- und ^{35}S-marked methionine, the pathway of ethylene in apples was discovered (Adams and Yang, 1979). Ethylene is derived from the amino acid methionine. Via s-adenosylmethionine-synthetase, methionine is converted to s-adenosylmethionine, which serves as a precursor in many

biosynthetic pathways. The next biosynthesis step is the conversion of S-adenosylmethionine to 1-aminocyclopropane-1-carboxylic acid (ACC). The enzyme ACC-oxidase regulates the ethylene formation (Figure 1). Ethylene is not only affecting the senescence of flowers and ripening of fruits, several development stages are affected like the growth, respiration rate, abscission of flowers and leaves, color synthesis, flower formation and root development. The life of cut flowers can be divided into three major stages: growth, maturation and senescence. Senescence is defined as the time when the process is changed to degradation (Willis et al., 2007). Onozaki et al. (2018) detected that a better VL is based on a lower production of Ethylene than on a higher sensitivity against it.

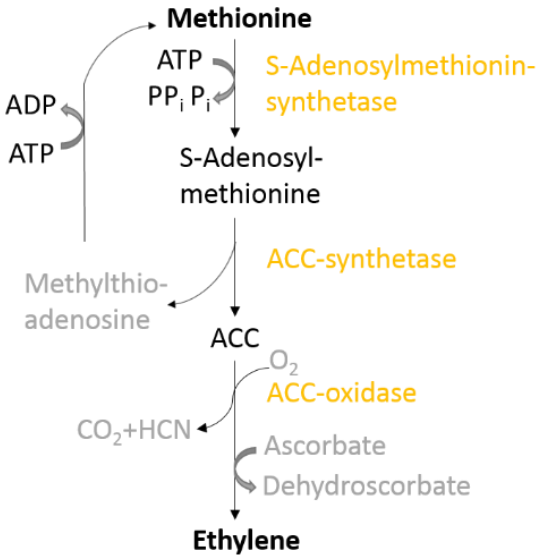


Figure 1: Ethylene biosynthetic pathway.
ATP – Adenunnucleotidtriphosphate, ADP – Adeninnucleotiddiphosphate, ACC - 1-Aminocyclopropane-1-carboxylic acid, PPi – Diphosphate, Pi – Phosphate, CO2 - Carbon dioxide, HCN - Hydrogen cyanide (Arc et al. 2013).

Some other plant hormones like abscisic acid (ABA) are as well associated with senescence for example by the regulation of stomatal closure and transpiration (Nambara and Marion-Poll, 2005). The molecular basis of ABA metabolism was discovered by genetic approaches (Xiong and Zhu, 2003). ABA is upregulated in later stages of senescence and reduces the water uptake. This suggests that a higher ABA content leads to a reduced postharvest duration (Müller et al., 1999). Gibberellins (GA) in plants have a function in stem elongation. Fletcher and Osborne (1965) first identified a biochemical relation between GA and leaf senescence. GA have the ability to delay senescence by antagonizing ABA and

have no direct influence on the initiation of senescence (Jibran et al., 2013). Auxin has the ability to inhibit the expression of genes associated with senescence and counts as negative regulator (Mueller-Roeber and Balazadeh, 2014). Jasmonic acid (JA) acts as a repressor of gene expression (Häffner et al., 2015). Exogenously supplied methyl jasmonate leads to senescence of leaves (Ueda et al., 1981). Metabolic activity like gluconeogenesis and the synthesis of sugars from complex carbohydrates is increased during the senescence process (Ho and Nichols, 1977; Sood et al., 2006). During plant senescence, transport of mineral nutrients such as aluminum, calcium and copper ions between roots and leaves is increased, stems and leaves which have a membrane protecting function against oxidative stress (Bartoli et al., 1997). Ubiquitin is a regulatory protein, signaling the degradation of proteins and it is proven that it has a role during senescence (Belknap and Garbarino, 1996). Ubiquitin and ubiquitin-like proteins affect the regulation of senescence. Several ubiquitin-dependent degradation pathways like ubiquitin ligases are important for their degradation by the 26S proteasomes (Woo et al., 2013).

The understanding of the effects that are affecting the VL is getting more important due to the long transportation times. The production of carnations for the global market occurs almost exclusively in the Southern Hemisphere and most customers are located in the Northern Hemisphere. In addition to the duration of transport, the consumer expects a VL of two weeks from carnations. This shows the importance of a long postharvest duration and a good transportability of carnations. In the past, however, breeders mainly concentrated on flower color and size, stem length and productivity, because VL and transportability are complex traits and are difficult to measure. The evaluation of VL is only possible in the late phase of the rating of carnations. The plants must be fully developed and the flower must be just before opening. The trait VL is complicated to measure because it is not possible to rate it in the greenhouse. The stems must be harvested, transported to the lab and put into vases. The rating is time consuming as well, because stems should be checked daily of senescence symptoms. In the case of such quantitative and difficult traits it is important to ensure good data quality.

1.3 Experimental designs in two-phase experiments

Good experimental design is necessary for the success of experiments in plant science (Casler, 2015; Jones et al., 2015; Piepho et al., 2013). The importance of replication is a matter of common knowledge because neither the error variance nor significance tests or confidence intervals can be computed without replication. Randomization is crucial to justify assumption that errors in a linear model are independent and to assess any differences in environmental conditions between experimental units. Blocking can be used to improve precision of the experiment. In some cases, experiments comprise two or more phases due to chemical, physical or molecular biological measurements made in a second phase in the laboratory that is distinct from the first phase where plants are raised (Brien et al., 2011; Smith et al., 2006). VL is a second-phase trait, because the experiment has a greenhouse phase and the second phase includes the VL assessment in the laboratory, as visualized in Figure 2.

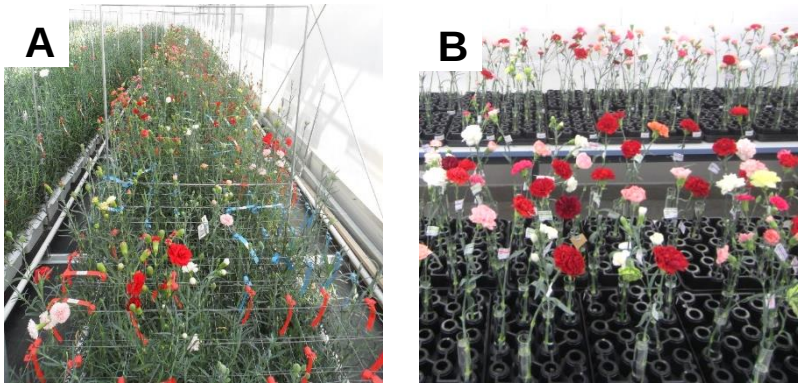


Figure 2: Visual representation of the two phases of VL trials.
A - the first phase includes the cultivation of carnations in the greenhouse, B- the second phase means the process of determination the vase life in laboratory.

To maximize precision, efficient experimental designs must be established in both phases. As noted by Wood et al. (1998) 'the distinctive feature of two-phase [or multi-phase] experiments is that each phase has its own block structure and these must be combined to form the overall variance model.' In field or greenhouse experiments, it is common to use experimental designs (Casler, 2015; Piepho et al., 2015), while in the laboratory often no particular randomization is applied. Molenaar et al. (2017) described the use of experimental designs in both phases in a rooting experiment of *Pelargonium zonale*. Other current examples

are given in Brien (2018). The benefit of experimental designs as well as of data analysis based on mixed models is so far uncommon for VL investigations. In recent research of VL, no use of experimental design in the second phase has been mentioned. However, especially when the target trait is measured in a second phase, the experimental design in the laboratory is important. The topic of two-phase experiments is getting more and more in the focus of research. Other traits relevant for breeding like FD, productivity or SL are one-phase traits, but measurements in the lab, retail tests or test of outdoor performance are conducted in a second phase like VL.

1.4 Improved selection of the trait VL

In the past, successful breeding of carnations and other cultivars was mainly based on the experience and intuition of plant breeders. Nowadays, breeders have the aim to use molecular knowledge such as marker-assisted or genomic selection for the selection of new varieties. Differences in the phenotype, like different levels in resistance, reactions to diseases or expression of quantitative traits, are caused by various types of alterations at various levels, such as nucleotide substitutions, insertion/deletion of nucleotides, differences in the copy number of repetitive elements and combinations. To detect genetic variations between different genotypes, various techniques have been established. The development of molecular markers was first published in the 1980s when isozyme markers were used (Tanksley, 1983; Tanksley and Rick, 1980). Lande and Thomson (1990) introduced the theoretical concept of marker-assisted selection (MAS) for quantitative traits. MAS is a suitable method for traits where the accuracy of selection is low. This is the issue in traits with low heritability or with few records and traits that are measured in a late stage (Meuwissen, 1998). MAS has been a useful tool for plant breeders but had less success in improving complex traits and it is difficult to find QTLs with small effects (Bernardo, 2008; Xu and Crouch, 2008). Success has been expected for mono- and oligogenic traits. Genomic selection (GS) was first described from Meuwissen et al. (2001). In contrast to MAS, in GS all available marker information to predict breeding values is used, therefore no significance tests to identify markers linked to QTL with large effects on a trait must be used. This indicates that MAS is more suited to traits controlled by a few QTL with large effects. In ornamentals, not many studies are published

with successful implementation of molecular markers and the use in practical breeding. Table 1 shows a comparison of different crops from the ornamental sector and maize (*Zea mays*) as an example from the agricultural sector presented regarding the availability of ESTs (expressed sequence tag sequences), SNPs (single nucleoid polymorphism), genome sequencing and sequenced protein at NCBI (National Center for Biotechnology Information; accessed February 2018). It is striking that in the agricultural sector much more effort has been made to introduce SNP markers. For example, while in maize breeding SNP arrays are already developed (Ganal et al., 2011), in ornamental breeding no SNP markers have been established so far.

Table 1: Publicly available ESTs, SNPs, Genome sequencing and protein sequences at NCBI of important crops for comparison.
Accessed February 2018; NCBI - National Center for Biotechnology Information, ESTs - expressed sequence tag sequences, SNP – single nucleoid polymorphism.

	ESTs	SNPs	Genome	Proteins
Gerbera	17,000	0	0	529
Rose	77,041	0	1	3,88,561
Carnation	23,222	0	1	10,934
Petunia	64,293	0	0	25,631
Pelargonium	44	0	0	11,901
Maize	2,187,011	58,916,206	10	280,263

Reasons for the delayed research in ornamental science compared to crop science are polyploidy in most cultivars, simpler breeding goals, high sales for cultivars and less public funding (Debener and Linde, 2009). In economically relevant ornamentals such as *Dianthus, Chrysanthemum, Petunia or Rose* or agricultural crops, the development of marker-assisted selection is currently underway. To improve selection of long lasting carnations, within this thesis it was tried to implement MAS in carnation by using bulked segregant analysis (BSA).

1.4.1 Bulked segregant analysis

BSA is a method for identifying SNPs and to develop molecular markers for the trait of interest. The main idea is to use a segregating population from one cross or a population of plants with diverse genetic background. With the phenotypic data two groups (bulks) are selected which display contrasting phenotypes for the trait of interest, meaning they differ only in the selected region. Molecular methods are used for analyzing the bulks instead of single lines, which reduces costs and time. The goal is to detect differences in allele frequencies between

the bulks. The key idea of BSA is that markers with differences in allele frequencies between the two contrasting bulks are closely linked to genes regulating the phenotype and selection can be done already at seedling stage, long before the phenotypic trait is expressed, which is a great advantage (Michelmore et al., 1991). In *Petunia hybrida* it was possible to establish molecular markers for double flower traits via BSA (Liu et al., 2016). Schulz et al. (2016) identified SNP-based markers for flower pigment influencing traits in *Rosa*.

1.4.2 Transcriptome sequencing

Ribonucleic acid (RNA) is synthesized from DNA by an enzyme known as RNA polymerase during a process called transcription. The transcriptome is the set of all RNA molecules expressed by an organism and describes the array of RNA transcripts produced in a particular tissue. RNA-Seq is used to analyze the transcriptome and was applied to establish a data basis of molecular data in carnations (Wang et al., 2009b). The advantage of using RNA-based sequencing methods compared to DNA-based is that statements about relevant genes are possible. Only expressed genes are being sequenced. On the other hand, it is important to decide which tissue to sample and at which time point. Some investigations were conducted in the case of cut carnation, like the whole genome sequencing of Yagi et al. (2014) or the RNA-Seq of carnations from Tanase et al. (2012b). The combination of the methods BSA and RNA-Seq is called bulked segregant RNA-Seq (BSR-Seq) (Liu et al., 2012). Massive analysis of cDNA ends (MACE) is a sequencing method where from each mRNA transcript via next generation sequencing (NGS) only a sequence of a length between 50 – 500 bp near to the poly(A) trail is generated. Each transcript represents one single read. Within the focus on the smaller area of the transcripts, a higher sequencing deep is expected. With RNA-Seq full-length transcripts are sequenced and MACE is using the 3’ ends where the most polymorphism are expected. The workflow of MACE is represented in Figure 3. The most interesting transcripts such as transcription factors or receptors are usually transcribed in very few copies, which cannot be found with RNA-Seq or Microarrays. Therefore, the advantage of MACE is that gene-based markers can be applied and indirect selection can be conducted in early stages.

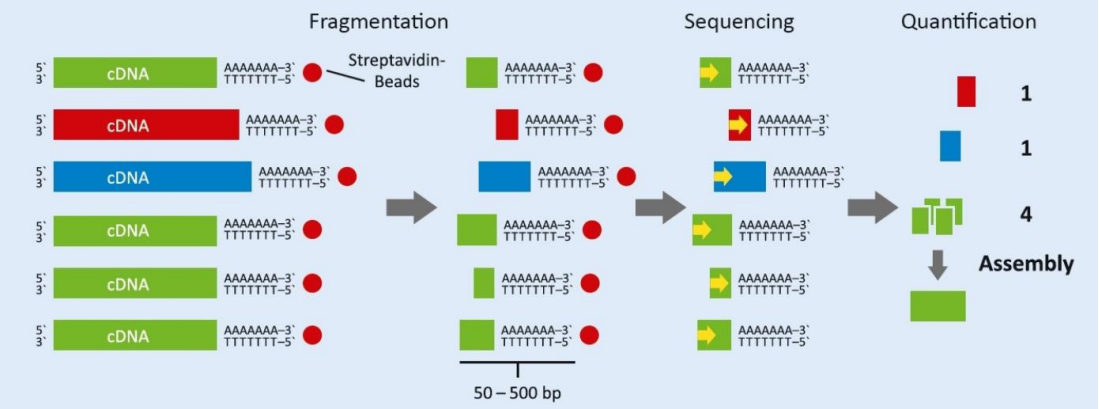


Figure 3: Workflow of MACE (massive analysis of cDNA ends).
After synthesis of cDNA from RNA, the cDNA is fragmented and sequenced from the 3′ single end to the 5’. (available at: http://genxpro.net/sequencing/transcriptome/mace-massive-analysis-of-cdna-ends/)

1.4.3 Indirect selection

One of the most popular examples for indirect selection is MAS (Bos and Caligari, 1995). Genetic correlation between two traits is used to implement indirect selection. To calculate genetic correlation via bivariate analysis the methods described in Piepho and Möhring (2011) are used. Through genetic correlation, the mean phenotypic value will be changed not only for the target trait of selection but also the other so called auxiliary trait will be influenced indirectly (Bos and Caligari, 1995). In complex traits like VL, which are hard to measure and are quantitative involving multiple genes with additive effects, indirect selection can be a helpful tool. The best results could be obtained by using a combination of indirect and direct selection (Falconer and Mackay, 1996). Due to the use as additional information, the auxiliary trait provides an indication of the performance for the breeding goal trait. Another advantage is to save time or effort to measure the target trait. The relative selection efficiency (RSE) could be calculated to see if indirect selection is the better measurement compared to direct selection.

2. Objectives

This thesis will demonstrate the development of molecular markers for the trait VL in practical cut carnation breeding. This included the implementation of experimental designs in two-phase VL experiments. The statistical analysis and proof of the high data quality to implement BSA was an important issue of this work. An important step was the definition of contrasting genotypes in the trait of interest. The genotypic data was obtained with two different sequencing methods. The goal was to find SNP-based molecular markers to improve selection efficiency of long-lasting carnations and to implement a fast genotyping-method. As alternative genotypic correlations were estimated in bivariate mixed models analysis to be used for indirect selection.

In **Chapter 3,** the use of experimental design in the greenhouse as well in the second phase, the VL assessment in the laboratory, is discussed. Response to selection was used to decide how to design VL trials in an efficient way.

In **Chapter 4,** comparative transcriptome analysis was conducted. GO terms were analyzed to find differences between in VL contrasting genotypes and between standard and mini carnations. The goal was to evaluate the conducted transcriptome data with the aim to find molecular markers.

In **Chapter 5,** genetic and phenotypic correlations for breeding relevant traits in carnations were calculated in bivariate analysis with the aim of indirect selection.

3. Efficient statistical design in two-phase experiments on vase life in carnations (*Dianthus caryophyllus* L.)

Maike Boxriker, Robert Boehm, Jens Möhring, Hans-Peter Piepho

The original publication is available at

http://www.sciencedirect.com/science/article/pii/S0925521416302022?via%3Dihub

Article history:

Received 9 August 2016
Received in revised form 30 November 2016
Accepted 12 December 2016
Available online 26 January 2017

Keywords: Vase life assessment, Carnation, Experimental design, Two-phase experiment, Evaluation of efficiency

Abstract

Good experimental design is necessary for the success of experiments. Many experiments on ornamentals comprise at least two phases, for example vase life trials in carnations. The first phase entails the cultivation in the greenhouse and the second the vase life assessment in the laboratory. In field or greenhouse experiments, it is common to use efficient experimental designs such as α-designs or row-column designs, while in the laboratory quite often no particular experimental designs are applied. We conducted vase life trials in carnations, using a randomized layout in both phases. We utilized the data from these trials to compare with competing designs based on the response to selection. This comparison revealed that the choice of design is crucial.

3.1 Introduction

Carnation (*Dianthus caryophyllus L*.), a member of the family *Caryophyllaceae*, is one of the major cut flower species around the world (Yagi et al., 2013). The genus Dianthus includes about 300 species in the *Caryophyllaceae* family. It comprises several economically important species like *Dianthus caryophyllus, D. barbatus, D. chinensis, D. plumarius, D. superbus* and their hybrids which are used as floricultural crops (Tanase et al., 2012). Improving vase life (VL) of cut flowers is one of the important breeding objectives due to its direct influence on consumer satisfaction resulting in a repeated purchase decision (Onozaki et al., 2001). VL is the duration from placement of stems in water to the loss of ornamental value, which is defined as the state of wilting, rolling-in of petals or discoloration of the petal margins (Satoh et al., 2005). In carnation, VL is regulated by a manageable number of genes especially in the case of biosynthesis of ethylene, a gas that regulates senescence of flowers (Fu et al., 2011; Licausi et al., 2013; van Doorn and Woltering, 2008). By using transgenic approaches, it was proven that repression of ethylene biosynthesis or blocking of ethylene receptors leads to an improved VL (Satoh, 2011). VL is an important criterion for the quality of carnations and required for sustainable transportation of carnation stems around the world. In the past, however, breeders mainly concentrated on flower color and size, stem length and productivity, because VL is a highly complex trait to measure.

Good experimental design is necessary for the success of experiments in plant science (Casler, 2015). In some cases, experiments comprise at least two phases due to chemical, physical or molecular biological measurements made in a second phase in the laboratory that is distinct from the first phase where plants are raised (Brien et al., 2011; Smith et al., 2006). VL trials have a greenhouse phase while the second phase includes the VL assessment in the laboratory. To maximize efficiency experimental designs have to be established in both phases. In field or greenhouse experiments, it is common to use for example α-designs or resolvable row-column designs, while in the laboratory often no particular randomization is applied. An α-design is an experimental design using incomplete blocks with only subsets of genotypes appearing in each block. But blocks can be grouped to form complete replicates (John and Williams, 1995). A row-column design utilizes double blocking with orthogonally arranged rows and columns

(John and Eccleston, 1986). The use of experimental designs as well as of data analysis based on mixed models is so far uncommon for VL investigations (Kramer et al., 2016). The aim of this study was to show how to design VL trials in a statistically efficient way by using an experimental design in the greenhouse phase, where the plants are cultivated, and in the second, the laboratory phase, where the VL assessment occurs. In order to achieve this main objective, we accounted for important sources of environmental variation in both phases and used simulation to evaluate the efficiency. The design and its analysis are illustrated using experiments conducted within a commercial breeding program for carnations.

3.2 Material and methods

3.2.1 Description of the vase life assessment

3.2.1.1 Cultivation of carnations in the greenhouse – first phase

For the cultivation of carnations, subsurface boxes (Beekenkamp, Netherlands) with nine planting holes were used (Figure 4A). The boxes have two drop inlets through which the irrigation water as well as the fertilizer were supplied. The substrate was ‘Cocopeat Premium’ (Euroveen B.V., Netherlands). Supports made of wire were used for growing the plants. The greenhouse comprised eleven parallel rows with 244 subsurface boxes each. 228 subsurface boxes in one parallel row were used for planting 500 genotypes (Figure 4B). Different gene pools were used, where the main part consisted of Klemm + Sohn (Germany) material, but also some genotypes to different breeders from Italy and Netherlands with whom Klemm + Sohn is cooperating. Not only established varieties were used, a main part was breeding strains of the years 2012 and 2013. The plants were planted in April and positioned randomly as described below in detail, because randomization can compensate edge effects (Cochran and Cox, 1957).

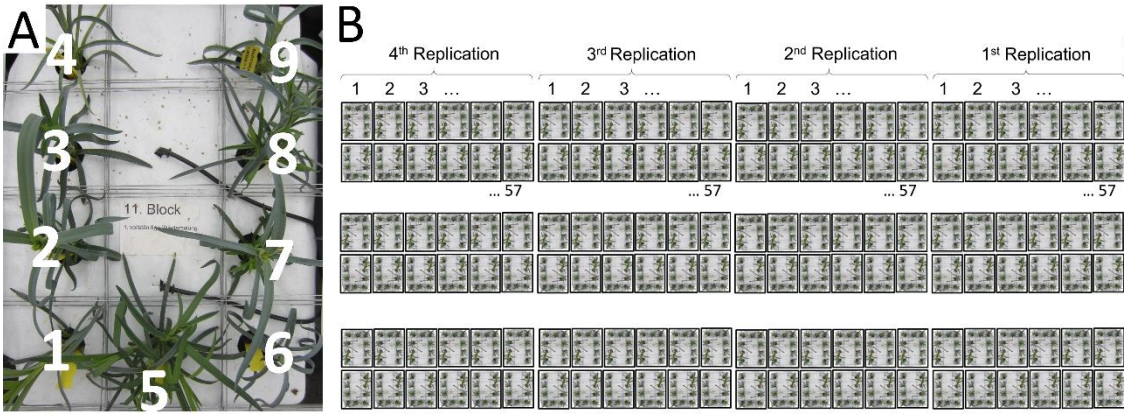


Figure 4: Experimental design in the greenhouse (first phase).

A - Arrangement of plants in subsurface boxes used as blocks and the observed positional effect, visible at position five; B – Overview of the greenhouse with four replicates, subsurface boxes and parallel rows.

3.2.1.2 Vase life assessment in the laboratory – second phase

The carnation stems were harvested in the greenhouse at flower development stage before blooming. Mini carnations were picked when more than two flowers of the same stem showed the desired development stage. For assessment of VL, two stems of every genotype in each replicate (in total eight stems per genotype) were harvested and transported to the laboratory. To simulate shipment of carnation stems, one stem of each pair was stored for two weeks at 4 °C, while the other stem was taken to the second phase immediately (Figure 5). All harvested stems from a single day, which are supposed to be stored for two weeks, were packed together in one plastic bag and put in containers to be stored. Harvesting of both stems at the same time was difficult because only one plant in each replicate was available.

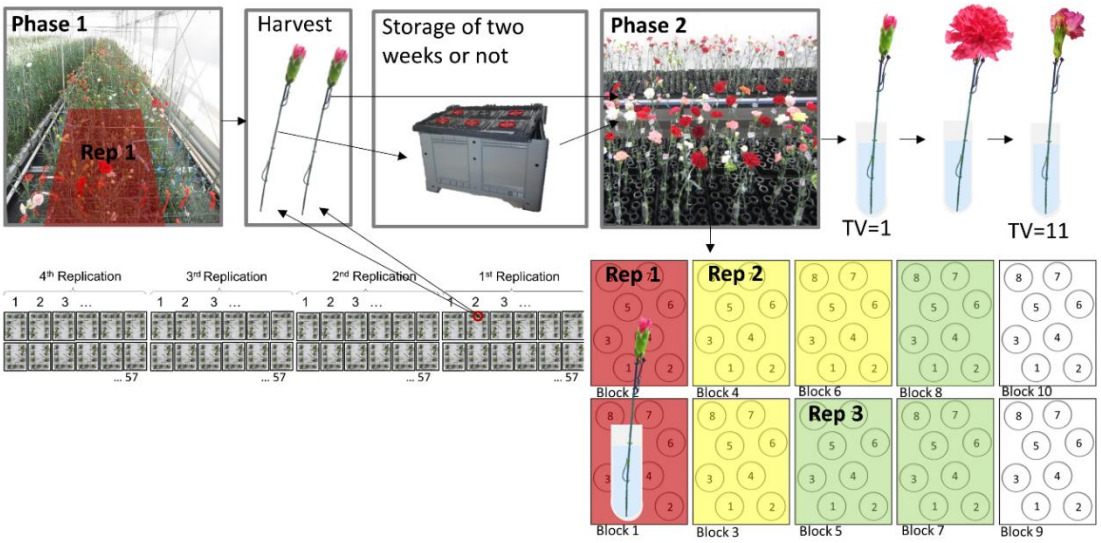


Figure 5: Experimental design in the laboratory (second phase).

Harvest of stems at right development stage and the proceeding of stems to asses for vase life (VL) of one out of four replicates (yellow)

After trimming the harvested or stored stems to 50 cm and removing of lower leaves, they were put in water in order to assess VL. Fresh water but no distilled water was used. The water was not changed but the filling level was held constant. The flowers were kept under controlled conditions: 12 h photoperiod of 300 lx in the first and 800 lx in the second year, provided by LED lamps and an average daily air temperature of 20.1 ±0.5 °C. Air was exchanged by using ventilators in the first year. In the second year, a 1.5-fold air exchange per hour was conducted by an air conditioning system. Each stem was observed daily to record senescence symptoms of the flowers like wilting, rolling-in of petals or discoloration of the petal margins or by bent necks of the stems. In the case of mini carnations, the stems were recorded as wilted when half or more of the opened flowers showed one of the named symptoms. For the second phase, an experimental design should be used as well. Difficulties occur due to the unknown point in time when the genotypes will reach the desired development stage.

3.2.2 Implementation of experimental designs

3.2.2.1 Generation of experimental designs for growing carnation in the greenhouse – first phase

In the first VL experiment (E1), two α-designs with four replicates were used. Each replicate comprised 57 incomplete blocks (subsurface box). In total 500

genotypes were tested that could be divided into 200 standard carnations (St) and 300 mini carnations (Mn). Since blocks with nine positions were used, there were 57x9 = 513 positions overall per replicate, so for thirteen randomly picked genotypes we added a further plant per replicates to make sure that all blocks were filled. The cultivation of both carnation types differs in some respects like the removal of flower buds. Therefore, in each box only one type of carnation was tested. We created two different α-designs, one for each carnation type. To integrate both carnation types into the same experiment, for every replication the incomplete blocks of a replicate were jointly randomized and renumbered. Thus, 207 standard carnations were randomly assigned to 23 subsurface boxes (blocks according to an α-design with 23 blocks per replicate) and 306 mini carnations were randomly assigned to 34 subsurface boxes, according to an α-design with 34 blocks per replicate. Each genotype was present in each replicate.

The year after (E2), a resolvable row-column design (D2) was used to test a second set of 500 genotypes. The row-column design was created similarly as the α-design in the first year. The subsurface boxes were defined as rows and the positions as columns, so there were 57 rows with nine columns in each of the four replicates. The designs were generated by using CycDesigN 4.0 or 5.1 (VSN International, United Kingdom). In both experiments, altogether 494 standard carnations and 502 mini carnations were tested.

3.2.2.2 Vase life assessment using statistical designs – second phase

It was not feasible to create a design for the laboratory phase in advance, as it was not known which genotype would reach the desired flower development stage at which time. To deal with this issue, two different approaches were established. The culture trays, which were used for placing the vases, were assigned as block units (TRAY). Each block had eight spaces (PT) for eight vases, which represented the observational units. Thus, 148 blocks, each comprising eight positions, were defined in the laboratory. Stems harvested on the same day in the same replicate in the greenhouse (first phase) were assigned to free blocks in the laboratory (second phase). The stems were randomly assigned to the free positions in the allocated blocks (L1, Figure 5). In the second year, the approach was changed and two random numbers (one for the number of the tray and the other for the position within the tray) were generated for each stem to

assign it randomly to a position in the laboratory (L2). In the case that a position was already occupied, the next free position was used and noted.

3.2.3 Data analysis

All data were statistically analyzed by linear mixed models using the procedure MIXED of the software package SAS 9.3 (SAS Institute 2002–2010). The effects included in these models will be described below.

3.2.3.1 Positional effect in subsurface boxes

In the first year after two months of cultivation, all plants showed a better development at position five in each box (Figure 4A). The reason was insufficient watering at the beginning of cultivation. The plants at position five were positioned nearer to the drop inlets and that is why the other plants suffered more from water stress. The used α-design can compensate heterogeneity between blocks, but not within the blocks. The reason is that an α-design only uses one-directional blocking. The positional effect in the blocks will result in a high residual error, which should be avoided. This was the reason why we decided to use a row-column design in the second year.

To evaluate the visible developmental differences in the subsurface boxes, the tillering was measured by counting all side shoots. In Table 2, the model set up is presented using the notation described in Piepho et al. (2013), where fixed and random effects are separated, with fixed effects appearing before the colon. The dot operator is used to define crossed effects and the residual term is underlined. The data analysis was done using a linear model for an α-design with an additional effect for the positions (POS). This effect had nine levels, one for each position within a subsurface box. The model includes a fixed effect for the four complete replicates (REP) and random effects for all incomplete blocks (BOX) in all replicates. The treatment factor (GENO) was fitted as a fixed effect. Note that this model only comprises design effects for the first phase because the trait analyzed here was measured in the first phase, meaning there was no second phase for this trait. Inspection of residual plots revealed no departures from normality or homogeneity of variance.

Table 2: Model set up for analysis of positional effect in subsurface boxes in the greenhouse.

Fixed and random effects are separated with fixed effects appearing before the colon, the dot operator is used to define crossed effects and the residual term is underlined

Design model	
REP : REP•BOX + <u>REP•BOX•POS</u>	
Treatment model	
GENO	
Full model	
Y = REP + GENO : REP•BOX + REP•POS + <u>REP•BOX•POS</u>	
Y – count of side shoots	*BOX – subsurface box*
REP – complete replicate	*POS – position in box*
GENO – genotypes	

Vase life was assessed in the second phase, so the model now comprises design effects for both phases. The VL data analysis of E1 and E2 was based on the design model set up for both phases, which was done separately for each phase. As noted by Wood et al. (1988) 'The distinctive feature of two-phase [or multi-phase] experiments is that each phase has its own block structure and these must be combined to form the overall variance model.' The design models were then combined with the model for treatment Table 3. Design factors like the blocks (BOX) and positions (POS) should be coded separately from treatment factors and are considered as random effects. In the case of the VL trials, the treatment factors were the storage of stems (STO), the genotype (GENO) and the carnation type (NA). For coding the storage of carnations stems, a factor with two levels was used. Unstored stems were coded with STO=0 and STO=1 denotes the stems which were stored for two weeks at 4 °C. The genotype was nested in carnation type (NA•GENO). The treatment factors were modelled by fixed effects, except the interaction of genotypes and storage (STO•NA•GENO), because in some cases only few stems were harvested resulting in a situation where all harvested stems of one genotype were stored for two weeks or meaning that the STO•GENO classification was incomplete. In this case, means for GENO can be obtained only when the interaction is modelled as random.

The model for the first year (E1) includes an effect for the four complete replicates and random effects for all incomplete blocks in all replicates. A fully developed carnation plant possess four to six stems. Two stems of each plant in each replicate were used, where FL identifies the harvested stem of a plant. FL is a factor

with FL=1 for the first harvested stem in one replicate and FL=2 is the second stem. An additional random effect for the positions has to be added (REP•BOX•POS). The row-column design is modeled by a fixed effect for complete replicates (REP). The row (BOX) and column (POS) effects must be fitted within replicates and are taken as random. Rows of the design correspond to boxes and columns to the nine positions within boxes. Additionally, as described for the α-design, a residual effect for the planting holes (REP•BOX•POS) was included.

In the second phase (the laboratory), culture trays were used as experimental units with 148 blocks (TRAY) and eight positions (PT) as illustrated in Figure 5. In the first year, the replicates (REP) from the greenhouse were transferred to the laboratory, allocating the harvested stems of one replicate together in free blocks. In the second year, we performed a complete randomization in the random block effect for recovery of inter-day information, because the stems were allocated to some positions twice or more often. All effects in the second phase were taken as random.

In both years, the trials lasted from August to December. In order to account for differences in temperature at the harvest of stems, temperatures were measured hourly 24 h before the harvest of each stem. The temperatures were averaged and used as a covariate (GT).

Table 3: Model set up for vase life trials in the first (E1) and second year (E2).

Fixed and random effects are separated with fixed effects appearing before the colon, the dot operator is used to define crossed effects, / is used as nesting operator and the residual term is underlined

Design model - phase 1
E1: *REP : REP•BOX+ REP•BOX•POS+ REP•BOX•POS•FL*
E2: *REP : REP•BOX+ REP•POS+ REP•BOX•POS+ REP•BOX•POS•FL*
Design model - phase 2
E1: *DV+ DV•REP•TRAY+ DV•REP•TRAY•PT*
E2: *DV+ DV•TRAY+ DV•TRAY•PT*
Treatment model
E1: *STO+ NA+ NA•GENO + STO•NA : STO•NA•GENO*
E2: *STO+ NA+ NA•GENO + STO•NA : STO•NA•GENO*
Full model [a]
E1: *VL = GT+ REP+ STO+ NA+ NA•GENO + STO•NA : DV+ DV•REP•TRAY+ STO•NA•GENO + REP•BOX+ REP•BOX•POS+ REP•BOX•POS•FL+ DV•REP•TRAY•PT*
E2: *VL = GT+ REP+ STO+ NA+ NA•GENO + STO•NA : DV+ DV•TRAY+ STO•NA•GENO + REP•BOX+ REP•POS+ REP•BOX•POS+ REP•BOX•POS•FL+ DV•TRAY•PT*

VL - vase life (d)	*DV* - start vase life assessment (categorical variable)
GT - temperature in greenhouse	*TRAY* - blocking unit in laboratory (culture trays)
STO - storage (yes or no)	*PT* - position in Trays
REP - complete replicate	*BOX* - blocking unit in greenhouse (subsurface boxes)
NA - carnation type (Mn or St)	*POS* - position in subsurface box
GENO - genotypes	*FL* - first or second harvested stems

[a] *Residual error terms of both phases were confounded so the effect REL•BOX•POS•FL was dropped from the full model.*

The residual plots of both VL trials showed a variance heterogeneity (data not shown), indicating that the variance increases with the mean. In both experiments, this was fixed by using a square root transformation. The data were back transformed by squaring the means, yielding median estimates on the original scale (Piepho, 2009).

3.2.3.2 Efficiency evaluation with statistical parameters

Heritability is an important parameter for plant breeders to quantify precision of field trials. In the narrow sense, heritability is defined as proportion of phenotypic variance among individuals that is due to additive genetic effects. The proportion

of phenotypic variance that is attributable to the whole genotype is called heritability in the broad sense. Whereas the initial definition of heritability was based on animal breeding in which the observational unit is always the individual animal, in plant breeding different types of observational units are possible which can be single plants or means of genotypes tested in different locations. Most formulae for heritability and response to selection assume balanced data and independent genotypic effects, which is often not the case in plant breeding trials. This complicates the estimating of heritability (Holland et al., 2003). The heritability (H^2) was calculated by using Formula 1 (see below) and is called ad hoc heritability (Piepho and Möhring, 2007). We illustrate the computation of the ad hoc heritability for the trait vase life in the first year (E1). Analysis of the data according to the model in Table 3 taking the genotype factor as fixed, yielded a mean variance of genotype differences of vd = 0.1012. Fitting the same model taking the genotype factor as random, we obtained a genotypic variance of σ_g^2 = 0.05664. With these two estimates, the heritability was computed as

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + {vd}/{2}} = \frac{0.05664}{0.05664 + {0.1012}/{2}} = 0.5282\ . \qquad |1|$$

The common formula for response to selection R = S x H^2 (Falconer and Mackay, 1996), where S is the difference of the mean of the selected parents and the population mean, is invalid in the case of unbalanced data and dependent genotypic effects. The linear relationship between response to selection (R) and selection differential (S) no longer exists. Response to selection was therefore calculated based on simulation as described in Piepho and Möhring (2007) and the SAS codes published at https://www.uni-hohenheim.de/bioinformatik/beratung/index.html. The simulation was implemented to show which response to selection was achieved by using the real data sets (E1 and E2) and different scenarios with alternative designs (Table 3). As criterion for evaluation, simulated response to selection for 1, 50, 100, 200 and 250 selected genotypes was calculated following the procedure described in Piepho and Möhring (2007). The calculation of response to selection was programmed in a SAS macro, which repeats the simulation 1000 times and calculates the mean of the response to selection across all simulation runs.

The Akaike Information Criterion (AIC) (Akaike, 1974; Wolfinger, 1996) was used to establish the importance of experimental design in both phases. This was done

by comparing models with different random effects omitted, but with the same set of fixed effects included. It was calculated as

$AIC = -2logL + 2p$,

where p is the number of free parameters of components of variance in the model and $logL$ is the logarithm of the residual log likelihood. AIC involves a penalty ($2p$) for the complexity of the model.

3.2.3.3 Generation of new data sets for comparisons

We simulated various scenarios, using variance components estimated for the datasets obtained in the first year of VL experiments (E1) and the second year (E2) using different designs in the first phase (D1/D2) and second phase (L1/L2). We use abbreviations for the different scenarios as described in Table 4. For the first phase, E1_D1_L1 and E2_D2_L2 denote α-design (D1) in the first and row-column design (D2) in the second year. For the second phase, L1 and L2 denote the randomization layouts used in both experiments (Figure 6).

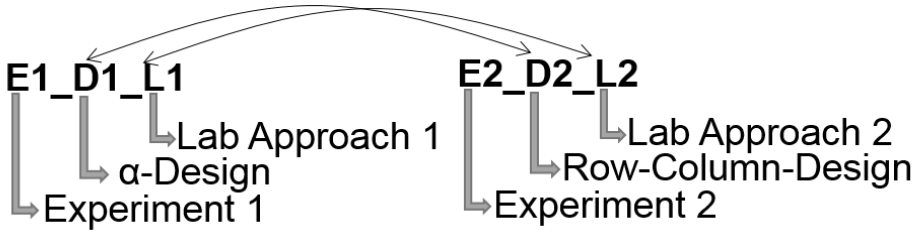


Figure 6: Usage of abbreviations to explain the simulated scenarios of the data set of the first and second year.

To denote different scenarios (Table 4), the changed design is coded with the corresponding number. For example, E1_D2_L1 denotes the analysis of a simulated row-column design (D2) for the first phase and layout L1 for the second phase, but the variance component estimates of E1 were used for the simulation of response to selection. E1_D1_L2 superimposed an α-design (D1) in phase 1 and simulated L2 in the second phase using variance components from E1. Using the same notation, the scenarios in which the experiment is analyzed as randomized are denoted as E1_D1_L1 for experiment E1 and as E2_D2_L2 for experiment E2.

Table 4: Real data sets and scenarios established by changing experimental designs for a comparison in simulation.

E1_D1_L1 – VL experiment in the first year with using α-design (D1) and laboratory approach with using replicates of the greenhouse (L1), E2_D2_L2 – VL experiment in the second year, D1 – α-design of the first year, D2 – row-column design of the second year, L1 – Laboratory approach with using replicates of the first year, L2 – complete randomization of the second year

		First Phase		**Second Phase**	
Name	**Variance components**	**D1**: α-design	**D2**: row-column design	**L1**: using REP	**L2**: complete randomization
E1_D1_L1	E1	x		x	
E1_D2_L1	E1		x	x	
E1_D1_L2	E1	x			x
E2_D2_L2	E2		x		x
E2_D1_L2	E2	x			x
E2_D2_L1	E2		x	x	

To improve comparisons of all data sets the genotype numbers were set to the same value. Data set E1 (N=502) and E2 (N=494) followed the same order and were merged by using the genotype numbers, the replicates (1–4) and FL (1 or 2), which codes for the first or the second harvested stem. The genotype number was decreased in both cases to 494 genotypes, caused by the merging of both data sets. The result was one big data set for both experiments, including the experimental designs in both phases and the vase life for each year.

3.3 Results

3.3.1 Positional effect in the first year

In the first year at the beginning of cultivation, a positional effect occurred and was measured by using the count of side shoots per plant (tillering). Plants at position five showed a highly significantly ($p<0.001$) better development compared to plants at the other positions (Figure 7). A significantly impaired development of carnations was detected at the positions 1, 4 and 9. Equivalent positions (1 and 6, 2 and 7, 3 and 8, and 4 and 9; Figure 4A) showed no significant difference. This was the reason why in the second year a row-column design was implemented.

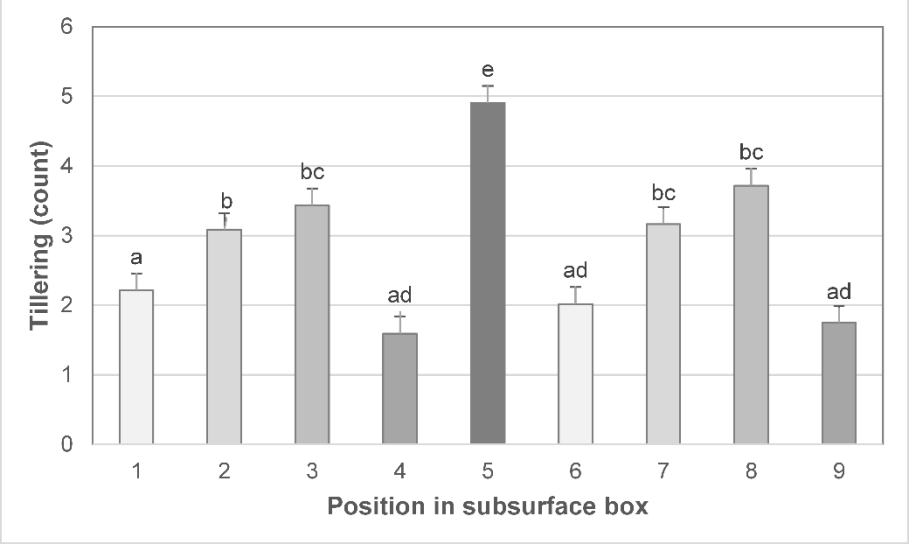


Figure 7: Adjusted means of rated tillering (count of side shoots) of all positions with standard error in the first year (E1).

Values with different letters indicate significant differences at $p \leq 0.05$ by an LSD-test

The analysis showed a significant difference among the four replicates (Table 5). Specifically, the first and second replicate showed a significantly better development compared to third and fourth replicate.

Table 5: Adjusted means of rated tillering (count of side shoots) of all replicates with standard error.

Values with different letters indicate significant differences at $p \leq 0.05$ by an LSD-test

Replicate	Mean	Standard Error	
1	3.38	0.11	a
2	3.19	0.11	a
3	2.41	0.11	b
4	2.51	0.11	b

After finishing the phenotyping of all 500 genotypes (E1) it was checked whether the positional effect was still present in subsurface boxes and measureable with VL (Figure 8). The plants at position five show a significantly inferior VL compared to other positions in subsurface boxes. In the second year (E2), no positional effect was detected (data not shown).

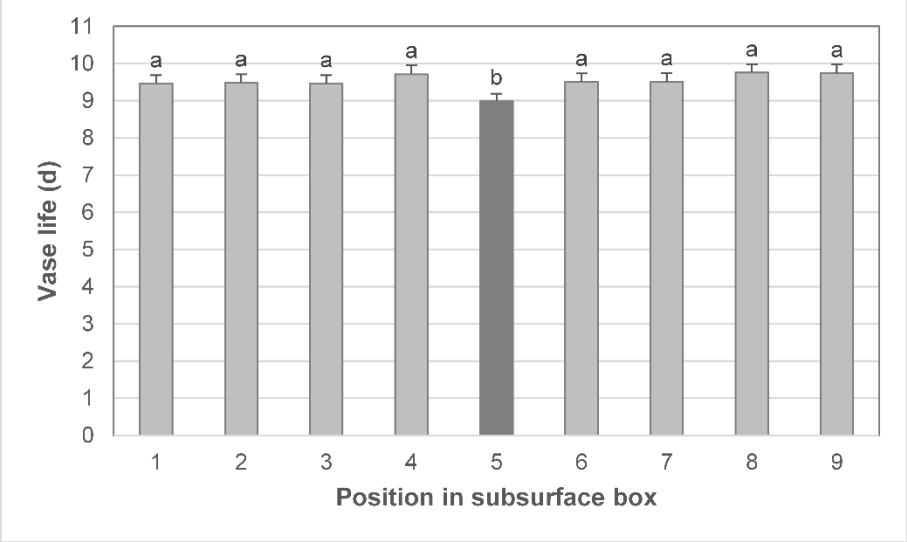


Figure 8: Adjusted means of vase life (d) of all positions with standard error in the first year (E1).
Values with different letters indicate significant differences at p≤0.05 by an LSD-test

3.3.2 Evaluation of vase life trials

3.3.2.1 Ad hoc heritability

The repeatability of the conducted VL trials was estimated with the measurement of ad hoc heritability (Piepho and Möhring, 2007). In the first year, there was a heritability of 52.82 % and in the subsequent year, heritability was improved to 59.99 % (Table 6).

Table 6: Ad hoc heritability (repeatability) of the randomized vase life trials in the first (E1) and second year (E2) in %.

	E1	**E2**
Ad hoc heritability	52.82	59.99

3.3.2.2. Efficiency evaluation of experimental designs in both phases

For the comparison of both conducted VL trials, the response to selection (R) in d was simulated for a different number of selected genotypes (1, 50, 100, 200 and 250).

The second year (E2_D2_L2) showed better results compared to the first year (E1_D1_L1) (Figure 9). In both cases, the average response to selection decreases with a higher number of selected genotypes. In the case of selecting only one genotype, a response to selection of 0.6867 d was calculated and in the year before, selection gain was 0.15 d lower by comparison. Therefore, the data of the second year (E2_D2_L2) showed a better precision and efficiency.

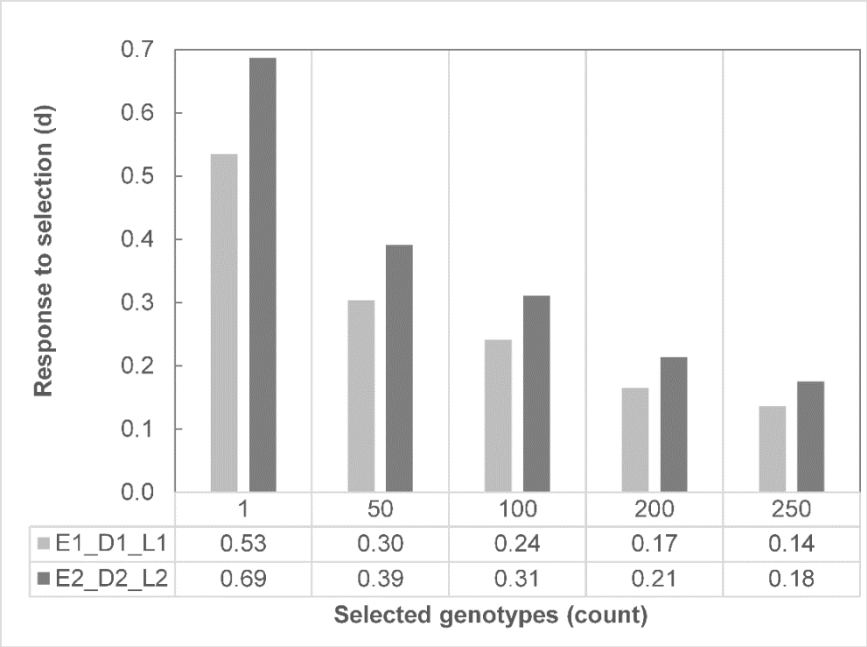


Figure 9: Response to selection of both years VL trials (E1_D1_L1 and E2_D2_L2) by a selection of 1, 50, 100, 200 or 250 genotypes.

In the first year the data E1_D1_L1 (R= 0.5349 d) was not improved by using a row-column design (E1_D2_L1, R= 0.5340 d) instead of the α-design (Table 7). By applying the completely randomized approach (L2) to the data of the first year (E1_D1_L2, R= 0.5336 d), no improvement was possible. Therefore, in the first year (E1_D1_L1) no improvement of the design was possible by changing the experimental designs.

In the second year, no positional effect was detectable in the greenhouse and the second blocking factor was not needed. The scenario E2_D1_L2 (α-design

instead of row-column design, R= 0.6877 d) achieves better results than the row-column design (E2_D2_L2, R= 0.6867 d). Nevertheless, the calculated response to selection is higher in the second year compared to the first year and by omitting the second blocking factor, i.e. by analyzing the data like an α-design (E2_D2/1_L2), the highest response to selection was attained (R= 0.7044 d). By using the L1 approach instead of L2, a higher response to selection (E2_D2_L1, R= 0.6947 d) is obtained compared to the real data set (E2_D2_L2, R= 0.6867 d), when using a complete randomization in the laboratory (L2). This shows that the approach L1 is more effective.

Table 7: Response to selection (R in d) for different scenarios by changing different approaches used in both years vase life experiments (E1_D1_L1 and E2_D2_L2) by a selection of only one genotype.

E1_D1_L1 and E2_D2_L2 are the real data set of the first and second year, D1 – α-design in first phase, D2 – row-column design in first phase, L1 – Laboratory approach using replicates in the first year (second phase), L2 – complete randomization in the second year, V1 and V2- variance estimated from E1_D1_L1 and E2_D2_L2

Name	R (d)	Description
E1_D1_L1	0.5349	real data set of E1
E1_D2_L1	0.5340	simulated row-column design in E1
E1_D1_L2	0.5336	2nd design in laboratory transferred to E1
E2_D2_L2	0.6867	real data set of E2
E2_D1_L2	0.6877	simulated α-design in E2
E2_D2/1_L2	0.7044	E2 analyzed like α-design
E2_D2_L1	0.6947	1st design in laboratory transferred to E2

3.3.3 Importance of using experimental designs by AIC comparisons

AIC was used for showing the importance of accounting for the experimental design in both phases of the VL trial. Table 8 illustrates that the full model (accounting for the design in both phases) is in both years more appropriate than models ignoring the first (P1) or second phase (P2). To not include the experimental design in the second phase results in a higher increase of AIC compared to omission of the first phase. This further emphasizes the importance of experimental design for the assessment of VL especially in the second phase.

Table 8: AIC comparison (smaller is better) of both randomized vase life trials (E1 and E2) with omitting one of the used experimental design in the first (P1) or second phase (P2).

	AIC
E1	3391.3
E1 without P1	3404.8
E1 without P2	3430.0
E2	3464.2
E2 without P1	3492.9
E2 without P2	3517.3

3.4 Discussion

3.4.1 Dealing with gradients in the greenhouse with experimental designs

Every greenhouse is characterized by gradients caused by different light conditions, aeration, humidity and CO_2 develop along the structure (Teitel et al., 2010). This can directly influence the development of plants as shown by the positional effect in subsurface boxes, which was caused by insufficient irrigation during the first weeks. In the second year, no positional effect occurred because there was an additional watering in the first weeks. In field or greenhouse experiments it is common to use randomized complete block designs (RCBD) or α-designs with one-way blocking (Jones et al., 2015). By using designs with two-dimensional blocking like a row-column design, a higher precision of treatment comparisons (Piepho et al., 2015) or improved response to selection was obtained. The four replicates exploited in the randomized trials showed a significantly different development of plants. The use of a statistical design with four different replicates can therefore be regarded as beneficial. In the laboratory, it was tried to reduce heterogeneity by using controlled conditions, but this was not entirely successful, as indicated by significant block effects, and it was not possible to control humidity.

3.4.2 Advantages and disadvantages of using experimental designs in first phase

The positive aspects of good experimental design are considered in several studies (Casler, 2015; Jones et al., 2015; Piepho et al., 2013; Smith, 1907). The row-column design showed good results in the case of a positional effect in boxes, because the second blocking factor could account for that effect. Conversely, lacking a gradient within blocks, the α-design was more effective. In the second VL trial, no effect of positions within in the subsurface boxes was detected

and the simulation showed that an α-design would have been more effective compared to the row-column design. Our results show that a row-column design is better, if some gradients in the greenhouse exists. When no positional effect is present, the row-column design can easily be analyzed like an α-design (Speed et al., 1985). On the other hand, allocating more than 2000 plants according to the generated experimental design is very time consuming, so there is an associated additional cost in terms of the logistics of the trial. In some instances, the wrong rooted cuttings were picked, set in wrong planting holes and the genotype ID mislabeled. In addition, it is more difficult to detect falsely labeled genotypes in a completely randomized arrangement compared to a design with all plants of a genotype placed in the same patch. The higher AIC when omitting first or second phase design effects indicates a less efficient data analysis. In addition, the lack of randomization entails a substantial risk of (unquantifiable) biases due to confounding of genotypic and environmental effects and a valid statistical analysis is not possible. Thus, randomized experimental designs with an efficient blocking structure are clearly preferable.

3.4.3 Randomization in the second phase is important

To create an experimental design for the second phase was difficult because it was not known which genotype would reach the desired flower development stage at which time. Therefore, it was first decided to order the harvested stems in the replicates as performed in the greenhouse. In the second year, it was tried to improve the design by doing a complete randomization. The fully randomized design was not beneficial as indicated by our simulation results. In a complete randomization, a blocking structure is missing and the error term is higher. A design handling the vases in the laboratory in the first year using the replicates from greenhouse as blocking unit was more precise. By AIC comparison, it was shown that design effects were important in both phases. When omitting the blocking units of the experimental design, the AIC increases more by ignoring the second phase, i.e., the laboratory. This indicates that the second phase has a dominant influence on the efficiency of data analysis. Further research should be conducted to increase the benefit of experimental designs in the second phase and to develop suitable designs for experiments in carnations involving two phases.

3.4.4 Improvement in the second year

The assessment of efficiency using simulated response to selection and heritability, indicated that the second year showed better results, probably due to the experience of staff in phenotyping gained in the first year. The trials differed between years in terms of the experimental design in the first phase (α-design vs. row-column design) and in terms of the randomization layout in the second phase (complete randomization in the second year). Exclusively in the first year, a positional effect was detected in the subsurface boxes. In the second year, the VL assessment occurred in a new room, which was built especially for this purpose. The light conditions as well as the aeration were improved. All other methods like the rating of senescence symptoms were the same in both years.

3.5 Conclusion

The best way to design VL trials in a statistically efficient way in practical breeding is to apply row-column designs in the greenhouse. By omitting the second blocking factor when this is found to be unimportant, the data can be analyzed like an α-design, if no gradient occurs along the column factor. In the second phase, it is also important to perform a randomization. It was shown that using replicates of the greenhouse as blocking units in the second phase improved precision compared to a design with complete randomization. The development of efficient designs for the second stage deserves further research.

4. Comparative transcriptome analysis of vase life and carnation type in *Dianthus caryophyllus L.*

Maike Boxriker, Robert Boehm, Nicolas Krezdorn, Björn Rotter, Hans-Peter Piepho

The original publication is available at

http://www.sciencedirect.com/science/article/pii/S0304423817300316?via%3Dihub

Article history:

Received 30 August 2016
Received in revised form 19 November 2016
Accepted 9 January 2017
Available online 24 January 2017

Keywords: Transcriptome analysis, Dianthus caryophyllus L., Vase life, Carnation type, Gene ontology

Abstract

Vase life (VL) and storage are the main breeding objectives in carnations due to their direct influence on consumer satisfaction and the transportability around the world. Molecular markers in carnation are available for different traits but not so far for VL, especially not SNP-based markers (single nucleotide polymorphism). Therefore, bulked segregant analysis (BSA) was introduced in carnation for the first time. 500 genotypes were phenotyped for the trait VL, contrasting bulks were generated and used for transcription sequencing with a combination of RNA-Seq and MACE (massive analysis of cDNA ends). Gene ontology (GO) terms were used for comparative analysis of annoted transcripts by comparing bulks with short and long VL. Differential expression (DE) of transcripts between the bulks with short and long VL was detected only by a higher expression of transcripts in the GO terms in the bulk with short VL. By literature review, many DE transcripts attributable to VL were identified. Little is known about the molecular differences between the carnation types standard and mini. The comparison of carnation types showed more DE transcripts as compared to the analysis of VL. Potential SNP markers were selected, but they are usable only for one carnation type, which indicates that the markers have to be established separately for the carnation types standard or mini.

4.1 Introduction

Carnation (*Dianthus caryophyllus L.)* is one of the most popular cut flowers of the world. There are two types of cut carnations available; the multi-flower mini carnation (Mn) and the one-headed standard carnation (St) (Shanan et al., 2010). The development of Mn arose out of a St with strong side shoots in the 1950ies (Köhlein, 1990). In the ornamental sector, vase life (VL) and storability are important traits due to long transportation times between producer and consumer. The production of carnation for the global market occurs almost exclusively in the Southern Hemisphere and the costumers are located in the Northern Hemisphere. In addition to the duration of transport, the consumer expects a VL of two weeks. The challenge is to slow the process of senescence to avoid high respiration rates and a high exposure of tissues to dehydration (Ebrahimzadeh et al., 2008). VL is defined as the duration from putting stems into vases to the loss of ornamental value due to senescence of the flower. Symptoms for senescence are wilting, rolling-in of petals or discoloration of the petal margins (Satoh et al., 2005). The regulation of VL in carnation is well studied and a lot of literature is available (Ebrahimzadeh et al., 2013; Rani and Singh, 2014). VL is limited due to gradually reduced water uptake during floral development, caused by the obstruction of xylem vessels. Microbial growth, deposition of lignin, suberin or tannin in the lumen of xylem vessels, presence of air emboli in the vascular system and formation of tyloses could be the reason for reduced water uptake (Twumasi et al., 2005; van Doorn, 1996).

Senescence is a biological growth process and includes some physiological processes, which are controlled by hormones. Well known is the regulation of senescence by the biosynthesis of ethylene, a volatile plant hormone that induces programmed cell death (Licausi et al., 2013; van Doorn and Woltering, 2008). The gaseous plant hormone ethylene is the simplest unsaturated hydrocarbon (CH_2=CH_2). The function as senescence inducing component was first described more than 100 years ago. Illuminating gas was used for a long time for lighting in homes, businesses and streetlights. In 1858, George Fahnestock observed first that leaking illuminating gas was affecting plants in the greenhouse. Newton and Cook (1927) demonstrated that ethylene was responsible for a higher respiration of bananas. With ^{14}C- und ^{35}S-marked Methionine, the pathway of ethylene in apples was discovered (Adams and Yang, 1979). Ethylene is derived from the

amino acid methionine. Via s-adenosylmethionine-synthetase, methionine is converted to s-adenosylmethionine, which serves as a precursor in a number of biosynthetic pathways. The next biosynthesis step is catalyzed by ACC synthase and is the conversion of s-adenosylmethionine to 1-aminocyclopropane-1-carboxylic acid (ACC). The enzyme ACC-oxidase regulates the formation of ethylene. Transgenic approaches showed that repression of ethylene biosynthesis or blocking of ethylene receptors results in improved VL (Satoh, 2011). Some other plant hormones like abscisic acid (ABA) are also associated with senescence for example by the regulation of stomatal closure and transpiration. The ABA biosynthesis is regulated by cleavage of cis-isomere from violaxanthin over neoxanthin to xanthoxin and is catalyzed by 9-cis-epoxycarotenoid dioxygenase (NCED) (Nambara and Marion-Poll, 2005). The product xanthoxin is then exported to the cytosol and converted to ABA (Xiong and Zhu, 2003). The molecular basis of ABA metabolism was discovered by genetic approaches. Non-stressed plant cells have a low ABA level. ABA is upregulated in later stages of senescence and reduces the water uptake. A higher ABA content leads to a reduced postharvest duration (Müller et al., 1999). The inhibitory effect of stress-induced ethylene on plant growth has been proven (Sharp, 2002). In some physiological processes sugar, ethylene and ABA biosynthesis have an interaction, which is not yet fully understood (Xiong and Zhu, 2003). Gibberellins (GA) in plants have a function in stem elongation, and GA-deficient mutants have dwarf phenotypes. Fletcher and Osborne (1965) first identified a biochemical relation between GA and leaf senescence. GA has the ability to delay senescence by antagonizing ABA and have no direct influence on the initiation of senescence (Jibran et al., 2013). Auxin has the ability to inhibit the expression of genes associated with senescence and counts as negative regulator (Mueller-Roeber and Balazadeh, 2014). Auxin is involved in many processes regulating the growth and development and that is why it is hard to analyze the role of auxin in developmental leaf senescence (Lim et al., 2007). Jasmonic acid (JA) acts as repressor of gene expression (Häffner et al., 2015). Exogenously supplied methyl jasmonate leads to senescence of leaves (Ueda et al., 1981). Metabolic activity like gluconeogenesis and the synthesis of sugars from complex carbohydrates is increased during the senescence process (Ho and Nichols, 1977; Sood et al., 2006). During plant senescence, transport of mineral nutrients such as aluminum,

calcium and copper ions is increased between roots, stems and leaves which have a membrane protecting function against oxidative stress (Bartoli et al., 1997). Ubiquitin is a regulatory protein, signaling the degradation of proteins and it is proven that it has a role during senescence (Belknap and Garbarino, 1996). Ubiquitin and ubiquitin-like proteins affect the regulation of senescence. Several ubiquitin-dependent degradation pathways like the E3 Ub-ligases are important for their degradation by the 26S proteasomes (Woo et al., 2013).

Senescence is an issue in practical breeding but the challenge is to measure the trait in a statistically efficient manner (In et al., 2009). Molecular methods such as marker-assisted or genomic selection for the breeding of new varieties are promising methods to improve response to selection (Dekkers, 2007). Molecular markers in carnation are already developed for variety identification (Arens et al., 2009), resistance to bacterial wilt (Yagi et al., 2013) or flower types differentiation (Yagi et al., 2014b). Bulked segregant analysis (BSA) is a method for identification of molecular markers by screening for allele frequency differences in two groups (bulks) displaying contrasting phenotypes (Liu et al., 2012). Using these methods in practical breeding, selection can be done already at seedling stage, long before the phenotypic trait is expressed (Michelmore et al., 1991). Publicly available at NCBI (National Center for Biotechnology Information, USA) for *Dianthus Caryophyllus L.* are 23,222 EST-sequences (expressed sequence tag sequences), 13 known genes, 56 sequence sets from phylogenetic and population studies, 594 protein sequences, 55,001 RNA- and DNA-sequences and one genome assembly information (Yagi et al., 2014a). For carnations, no entries of SNPs (single nucleotide polymorphisms) are public. Transcriptome sequencing of carnations was conducted by Tanase et al. (2012) and Wan et al. (2015) using RNA-Seq (Wang et al., 2009b). In this manuscript, a combination of RNA-Seq and massive analysis of cDNA ends (MACE) was used for the first time in carnation. MACE is a high-throughput Next Generation Sequencing method that sequences only the 3’-end of a transcript, where most of the polymorphisms of a transcript are present (Müller et al., 2014). Compared to RNA-Seq, where full length transcripts were sequenced, MACE requires only minimal sequencing effort. The important, but difficult-to-measure trait VL was chosen as a first target for the identification of SNP-based molecular markers. An analysis of gene ontology (GO) terms should help to identify differentially expressed (DE)

transcripts and different mechanisms between short and long VL. The analysis should also estimate genetic differences between the carnation type standard (St) and mini (Mn), to decide if SNP markers can be applied to both carnation types. The objectives of this paper were to analyze the conducted transcriptome sequencing and to identify mechanisms regulating the VL in carnations with the aim to find SNP-based markers for VL and transportability.

4.2 Material and methods

4.2.1 Experimental design and VL assessment

Details of the experimental design are described in Boxriker et al. (2017b). Briefly, for the cultivation of carnations, subsurface boxes (Beekenkamp, the Netherlands) with the substrate 'Cocopeat Premium' (Euroveen B.V., the Netherlands) were used. The boxes have two drop inlets through which the irrigation water as well as the fertilizer were supplied. An α-design with four replicates was used (John and Williams, 1995). Each replicate comprised 57 subsurface boxes, which were used as incomplete blocks with nine planting holes each. Each genotype was present in four complete replicates. 500 genotypes were tested which could be divided into 200 standard carnations (St) and 300 mini carnations (Mn). Since blocks with nine positions were used, randomly chosen genotypes were repeated eight times to fill blocks. The handling of both carnation types differs in some respects, like the removal of flower buds or senescence rating. The main part of the used gene pool consisted of Klemm + Sohn (Stuttgart) material. Some of the genotypes originate from different breeders in Italy and Netherlands. The main progeny originated from a breeding strain of 2012. The carnations were planted in April 2013.

For the assessment of VL, eight carnation stems per genotype (two per replicate) were harvested at flower development stage before blooming as shown in Figure 10. To simulate shipment of carnation stems and to test for transportability, one of the two harvested stems of each replicate was stored for two weeks at 4 °C, while the other stems were tested for VL without storage treatment. After trimming the harvested or stored stems to 50 cm and removal of lower leaves, flower stems were put in water to assess VL. Fresh, non-distilled tap water was used. The water was not changed but filling level was kept constant (1 cm under grim) and checked daily.

Figure 10: Flower development stages.

Marked in grey is the flower stage of harvest (2) for vase life investigation

The assessment of VL is a two-phase experiment (Brien et al., 2011; Smith et al., 2006), because the stems are cultivated in the greenhouse and the second phase includes the VL measurements in the lab. In the second phase, an experimental design should be used as well to improve efficiency (Boxriker et al., 2017b). Difficulties to implement experimental design in the second phase occur due to the unknown point in time when the genotypes will reach the desired development stage. To deal with this issue, the culture trays, which were used for placing the vases, were defined as block units (TRAY) and the spaces for eight vases as positions (PT) of each block. Thus, 148 blocks with eight positions each were defined in the laboratory. All harvested stems were grouped in replicates as used in the greenhouse and then the harvested stems of each replicate were randomly assigned to the blocks. The flowers were kept under controlled conditions: 12 h photoperiod of 300 lx provided by LED lamps and a mean temperature of 20.1 ±0.5 °C. The conditions were not totally homogenous because significant block effects were detected (data not shown), that is why the experimental design in the second phase is important. Air was exchanged by ventilators. Each flower was observed daily to record senescence symptoms. The stems were terminated when the flower (St) or 50 % of the flowers (Mn) lost their ornamental value by wilting, rolling-in of petals or discoloration of the petal margins or the stems showed bent necks.

VL was assessed in the second phase, so the model comprises design effects for both phases. In the first phase, an α-design with four complete replicates (REP) and random effects for all incomplete blocks (BOX) in all replicates at all positions (POS) was applied. In Table 9 the model set up is presented using the notation described in Piepho et al. (2003). A fully developed carnation plant possesses four to six stems. Two stems (FL) of each plant in each replicate were

used. In the second phase, the day of the beginning of VL assessment (DV) was used as first blocking unit in the laboratory. The replicates (REP) from the α-design were used to allocate the harvested stems to replicates in the laboratory. Culture trays (TRAY) were used as second blocking unit with eight positions each (PT). The treatment was the storage or no storage of carnation stems and for coding, a factor with two levels (STO) was used. The genotype was nested in the carnation type (NA). The hourly temperatures of each day between the duration of the trial from August to December were averaged per harvest day and used as a covariate (GT).

The residual plots of both VL showed that variance increased with the mean (data not shown). Heterogeneity of variance was removed by a square root transformation. Wald-type F-statistics, adjusted means and standard error estimates were calculated with the software package SAS 9.3 (SAS Institute 2002–2010).

Table 9: Model set up for the analysis of vase life (VL) trial.

Fixed and random effects are separated with fixed effects appearing before the colon, the dot operator is used to define crossed effects, / is used as nesting operator and the residual term is underlined

Design model: phase 1
REP : REP•BOX+ REP•BOX•POS+ REP•BOX•POS•FL
Design model: phase 2
DV+ DV•REP•TRAY+ DV•REP•TRAY•PT
Treatment model
STO+ NA+ GENO/NA+ STO•NA : STO•GENO/NA
Full model [§]
VL = GT+ REP+ STO+ NA+ GENO/NA+ STO•NA : DV+ DV•REP•FL+ STO•GENO/NA+ REP•BOX+ REP•BOX•POS+ REP•BOX•POS•FL+ DV•REP•TRAY•PT

VL - vase life in d	DV - start VL assessment
GT - temperature in greenhouse	TRAY - blocking unit in laboratory (culture trays)
STO - storage (yes or no)	PT - position in trays
REP - complete replicate in greenhouse	BOX - blocking unit in greenhouse (subsurface boxes)
NA - carnation type (Mn or St)	POS - position in subsurface box
GENO - genotypes	FL - first or second harvested stems

[§] Residual error terms of both phases were confounded so the effect REP•BOX•POS•FL was dropped from the full model.

4.2.2 RNA extraction, cDNA library and next generation sequencing

Transcriptome sequencing with a combination of two different sequencing methods was performed. RNA-Seq is a common method and reviewed in Wang et al. (2009b). MACE is a method developed by GenXPro GmbH (Germany). Here, only one sequence of the 3'-end of cDNA fragments is sequenced (Müller et al., 2014). We obtained the full length cDNA sequence from a de novo assembly of the RNA-Seq data. These were used as reference for the MACE data. To improve the transcript spectrum in RNA-Seq, many different tissues were used. For MACE more senescence related tissues were used, but no tissue of a fully developed flower was used.

4.2.2.1 Tissue samples

For RNA-Seq, two genotypes of both carnation types St and Mn with the highest contrast in VL (d) were selected (Table 10). The 22 best and 22 worst genotypes of the first year were selected based on the estimated genotype means and used for the sequencing with two different methods (RNA-Seq and MACE). A higher VL was obtained in the Mn bulk. The genotypes were cultivated in four replicates together within a set of 500 genotypes in the same greenhouse following a randomized experimental design (Boxriker et al., 2017b).

Table 10: Number of contrasting genotypes (N) of both sequencing methods for carnation types Mn and St with adjusted means of VL (d).

RNA-Seq	Mn	St	mean	MACE	Mn	St	mean
long VL	1	1	14.88	long VL	12	8	13.99
short VL	1	1	5.37	short VL	12	8	6.47
mean	10.4	9.85	10.13	mean	10.43	9.93	10.23

In the case of RNA-Seq, four contrasting genotypes were selected and ten different tissues at different developmental stages (leaves from young plants, shoots before formation of buds, roots, buds, young leaves from adult plants, old leaves from adult plants, shoot from adult plants, flower leaves, cold stressed leaves and water stressed leaves) were collected. The material was harvested from three biological replicates and pooled to one probe per single genotype (Figure 11A). Roots were harvested by cultivation in perlite (Knauf, Germany) and the stressed material was prepared by a treatment with 18 h cold (4 °C) and

hot (35 °C) storage, without watering. Out of 500 phenotyped genotypes, two contrasting bulks of 20 genotypes each contrasting in VL based on adjusted means, were selected and analyzed with MACE (Table 10). Buds and material from upper and lower stem were harvested from the greenhouse as described before and isolated separately (Figure 11B).

4.2.2.2 RNA extraction and cDNA synthesis

Total RNA was extracted from 50 mg frozen plant material of a pooled sample from three biological replicates following the procedure of RNeasy Mini Kit (Qiagen, Germany). Quality control was conducted by using LabChip GX software version 4.2.1745.0 (Perkin Elmer Inc, Swiss). The cDNA was generated by using a first-strand synthesis with hexamers combined with a ligation-based method. Reverse transcription was performed by using an anchored and biotinylated oligo(dT) primer that preferentially hybridizes to the proximal end of poly(A) tails.

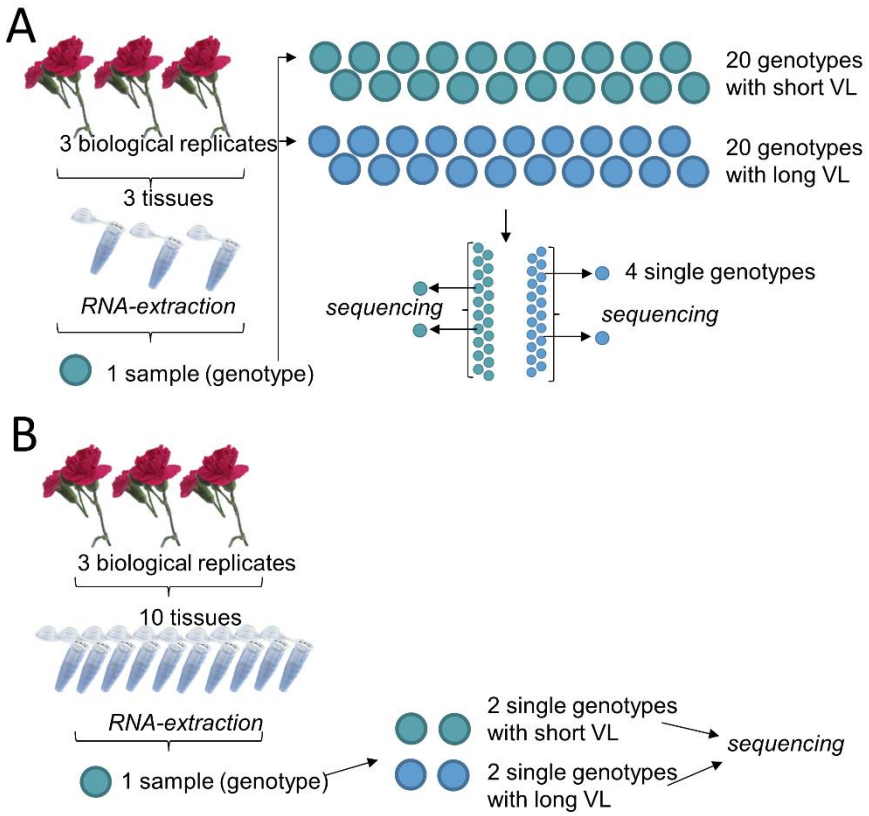


Figure 11: Sampling scheme of the transcriptome sequencing.

(**A**) MACE of bulks of genotypes contrasting in VL and (**B**) RNA-Seq of in VL contrasting genotypes; green – short VL, blue – long VL.

4.2.2.3 Next-generation sequencing

The RNA of all four contrasting genotypes was sequenced separately with the method RNA-Seq. In the case of MACE, two bulks of both carnation types were sequenced and an additional four single lines were sequenced separately (Figure 11B). The purified mRNA was fragmented in a Zn^{2+}-solution and Illumina sequencing adapters were ligated to the RNA. Each library was sequenced on an Illumina HiSeq2500 machine (Illumina, USA).

4.2.3 Sequence annotation

4.2.3.1 Raw data cleaning

The RNA-Seq and MACE reads from the Illumina sequencing were cleaned before further analysis of the data. Low-quality ends from reads, adapter and PCR primer sequences were removed using cutadapt (Martin, 2011).

4.2.3.2 De novo assembly and annotation

The resulting cleaned reads of each library were used for assembly with Trinity package version 2.01 (Grabherr et al., 2013). The annotation was conducted via BLASTX search (NCBI) against UniProt (Universal Protein Resource), SWIS-SPROT and TREMBL using a threshold (e-value of 10^{-5}) for considering the transcripts as homologs. The transcripts from the de novo assembly were mapped to the *Beta vulgaris* genome to locate their chromosomal location and hence the chromosomal location of the polymorphisms (SNPs).

4.2.4 Identification of transcripts differential expression between bulks

Following to Anders and Huber (2010), the average raw count of each transcript within a library was divided by the geometric mean of all counts and the median of the quotients were calculated per library. Each raw count was divided by library-specific median value to display the normalized values of all replicates by DEGseq (Wang et al., 2009a), and DE transcripts were identified as those transcripts with a Benjamini and Hochberg (1995) corrected *p*-value 0.05. The logarithm to the basis 2 of the fold change of the average of the normalized values was calculated. The combination of *p*-value and fold-change threshold criteria gives more evidence than using either of them alone (Chen et al., 2015). A transcript was treated as upregulated when two times higher normalized expressed in the bulks with long VL or St is present compared to the bulk associated with

short VL or the Mn. Downregulation means a higher expression of transcripts in the bulk with short VL or of Mn compared to the contrary bulk (long VL or St).

4.2.5 Analysis of MACE data

4.2.5.1 Venn diagrams

For graphical illustration of MACE data, Venn diagrams were generated using a web-based tool Venny 2.1 (Oliveros, J.C. 2007–2015).

4.2.5.2 GO Terms

The annoted hits were analyzed for their respective GO terms (Ashburner et al., 2000) and classified into GO slims using GoSlimViewer online software (AG Base version 2.0, Mississippi State University, 03/16/16). GO slims are a reduced version of GO terms. The annotation of contigs to GO terms was based on the “gene association.goa uniprot” file from Uniprot. The R-package TopGO (Alexa and Rahnenführer, 2013) was used to do the GO-enrichment analysis, which is based on the Fisher’s exact test among DE transcripts. Significant DE GO terms were identified by filtering the data by *p*-value (<0.05) and a higher ratio in the contrasting bulks than two for both conducted comparisons of the carnation type (Mn and St) and VL (short or long). A literature review identified pathways with a potential of influencing the VL of carnation like plant hormones, which are inducing programmed cell death (Ethylene, Auxin, Gibberellin, Abscisic Acid), carbohydrates metabolic pathway or the response to stress. The biological functions of the found significant DE transcripts were compared between short and long VL and both carnation types. The carnation types were also analyzed separately for short and long VL.

4.2.6 SNP identification

Sequence variants of MACE were analyzed with JointSNVMix (Roth et al., 2012). Sequence variants must have only two SNP types. DE SNPs were identified by using filter methods to get contrasting SNPs between the bulks. The goal was to find in one bulk SNPs with an allele frequency of 100 % and in the other bulk with an allele frequency of 0 %. The expression of a sequence variant as reference base (ref_base) or variable base (var_base) was chosen randomly.

4.3 Results

4.3.1 Quality of RNA-Seq

RNA-Seq data was used as reference for the analysis of MACE data. The total number of sequenced nucleotides was 71,864,367 bp. 57,120 contigs could be assembled. The longest contig was 24,374 bp long. A mean contig length of 1258 was obtained. As shown in Figure 12A, Mn showed longer contigs compared to St. N50 is similar to median of lengths, but has greater weight given to longer contigs (Miller et al., 2010). N50 is higher in both genotypes having a short VL (Figure 12B).

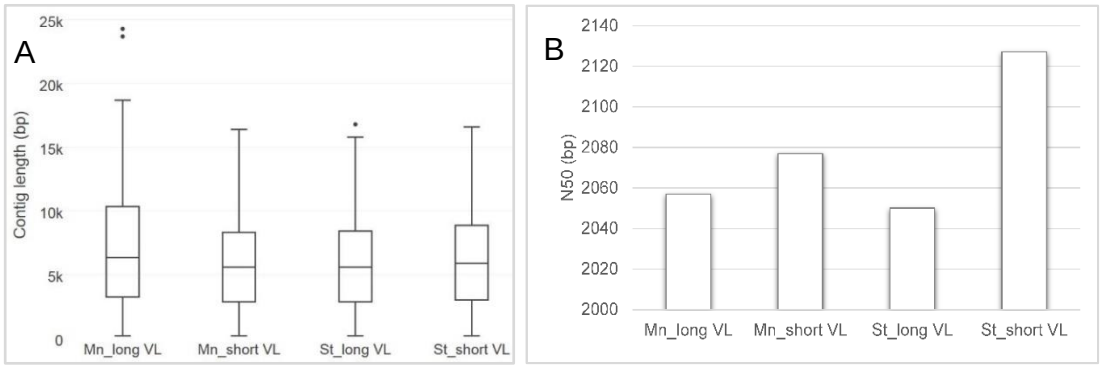


Figure 12: Quality of RNA-Seq.

A- Box-Plot of Contig length (bp) of RNA-Seq Contigs, Boxes represent the interquartile range between first and third quartiles and the line inside represents the median. Whiskers denote the lowest and highest contig length (bp). Circles represent outliers beyond the whiskers. B- N50 of all four genotypes separated in both carnation types (Mn and St) and short and long vase life (VL).

4.3.2 Analysis of MACE data

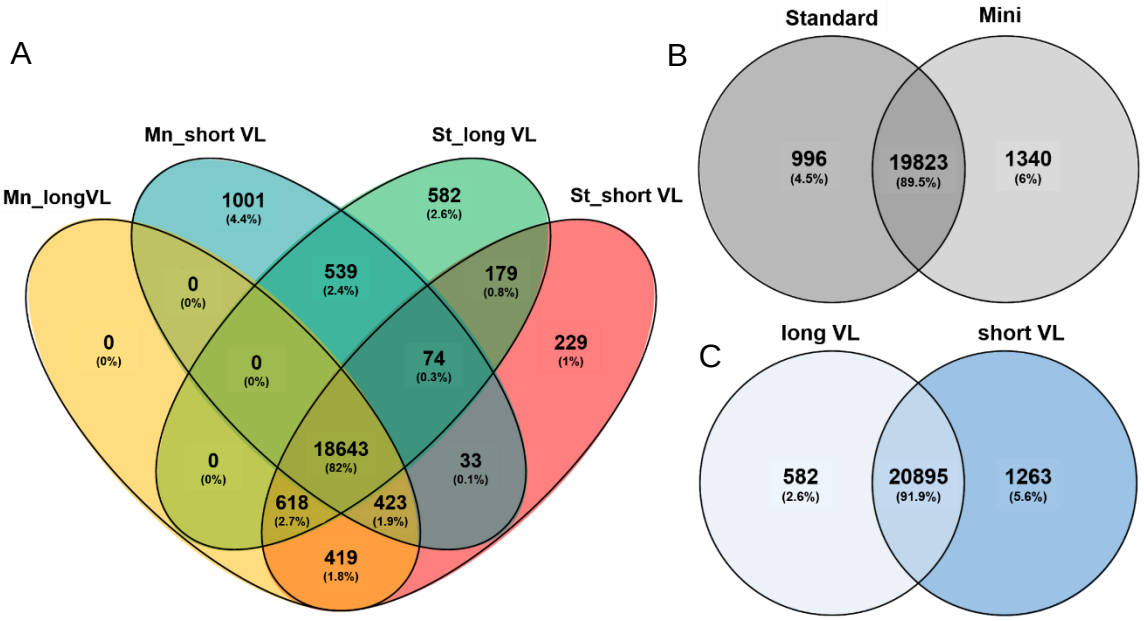
Altogether 40 genotypes of both carnation types (Mn and St) with a contrast in VL were sequenced with MACE. A difference between adjusted means of 7 d was obtained (Table 10). Additionally, four genotypes were chosen and sequenced as single genotypes for comparison (Figure 11B). Around 85 % of all contigs were annotated to existing protein sequences (Table 11). In the bulks, more reads were detected in St and more reads are available in the bulk with long VL.

Table 11: Annotation statistics of MACE separated in both carnation types (Mn and St) and in short or long VL.

	Bulks (40 genotypes)			Single genotypes (4 genotypes)		
sample	reads (bp)	annotated reads (bp)	% annotated	reads (bp)	annotated reads (bp)	% annotat
long VL (Mn)	12,157,003	10,486,487	86.26	11,995,526	10,046,162	83.75
short VL (Mn)	9,976,805	8,449,053	84.69	6,820,020	5,686,072	83.37
long VL (St)	18,555,843	15,911,350	85.75	7,545,855	6,473,937	85.79
short VL (St)	17,575,482	15,111,742	85.98	8,268,137	6,975,330	84.36

4.3.2.1 Venn diagram

For graphical illustration, Venn diagrams of different MACE bulks were drawn (Figure 13). 82 % of all contigs are available in all four bulks (St and Mn; short and long VL). 1001 contigs were exclusively present in the Mn bulk with short VL (4.4 %). All contigs of Mn bulk with long VL were shared with other bulks (Figure 13A). The St and Mn bulks have 89.5 % of the contigs in common (Figure 13B). 1340 contigs were only present in the Mn bulk, 344 more than in the St bulk. 91.1 % of all contigs were shared between the bulk with short and long VL (Figure 13C). By separating the contigs into the two carnation types and into bulks with short or long VL, a similar number of contigs and distribution are present (Figure 13D and E).



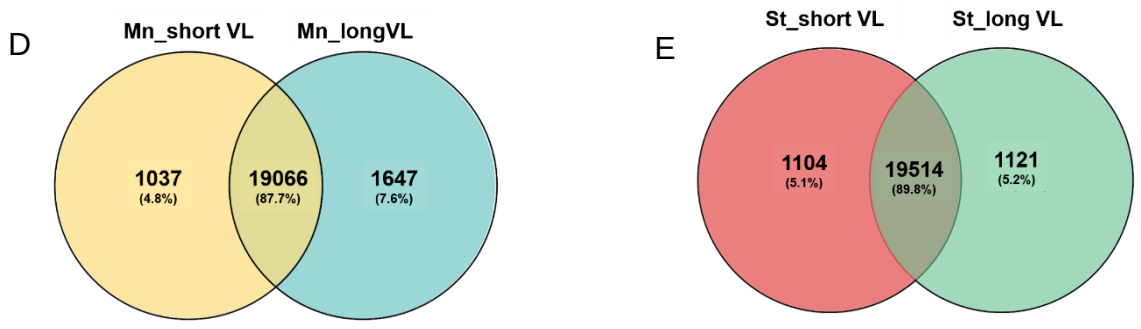


Figure 13: Venn diagram of MACE data.

A: all conditions (VL and carnation type), **B**: the carnation type standard (St) and mini (Mn), **C**: the tested VL (short and long) C) and **D**+**E**: the tested VL of both carnation types separated.

4.3.3 GO slims

The GO slim most represented in the category Biological Process (Figure 14) was “nucleobase-containing compound metabolic processes” (20.19 %) like RNA-processing, glycolytic process, DNA-dependent transcription or regulation of cellular transcription. “Transport” was one of the ten most frequent GO slims (14.97 %). Examples are the transport of metal ions, oligopeptides, water or carbohydrates. The “protein metabolic process” includes 232 unigenes referring to ubiquitin-dependent protein catabolic processes, signal peptide processing or proteolysis. “Carbohydrate metabolic processes” include the formation of molecules by the addition of a carbohydrate residue. The next six most frequent GO slims are related to “translation”, “response to stress”, “lipid and DNA metabolic process”, “cellular component organization” and “signal transduction”. The stress reactions refer to heat, desiccation, herbicides, oxidative stress and other stress response reactions.

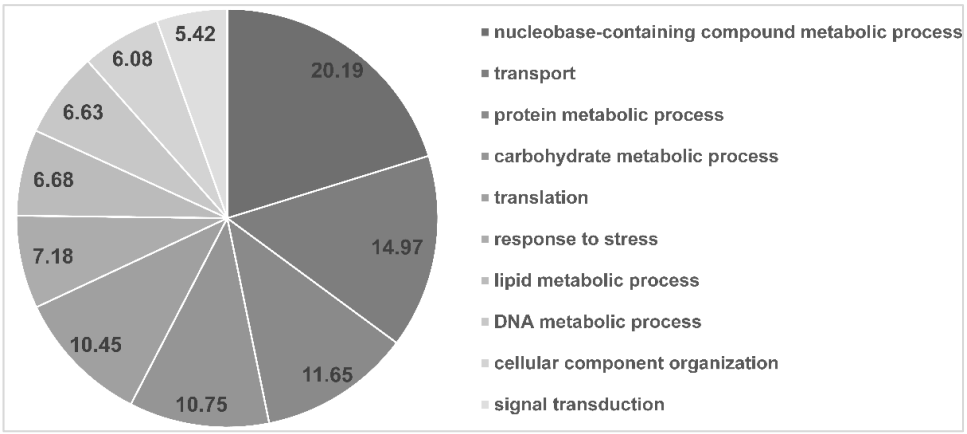


Figure 14: Percentage of *Dianthus caryophyllus L.* transcripts of the ten most redundant GO slims of the category Biological Process.

The most common GO slim in the category Cellular Component (Supplemental 1) is associated with the “membrane”, which includes all gene products and protein complexes, having at least some part of their peptide sequence embedded in the hydrophobic region of the membrane. 413 genes are part of the term “intracellular” and include different complexes like the proteasome core complex, signal peptidase complex or intracellular ribonucleoprotein complex. The biggest part of the category Molecular Function (Supplemental 1) is the term “binding” with almost 1800 unigenes and pertained for example to the binding of coenzymes, pyridoxal phosphate, metal ions and ATP.

4.3.4 GO annotation

4.3.4.1 Significant DE GO terms between short and long VL

Eight significant (p<0.05) DE GO terms with a higher ratio than two were identified between the bulks with short and long VL. In the bulks with short VL, the associated transcripts were upregulated, while in the other bulk, the transcripts were less present (Figure 15). The smallest p-value was found in “proteasome assembly”, which is the aggregation, arrangement and bonding together of a mature, active proteasome complex. Most genes in this GO term were assigned to the mitochondrial membrane respiratory chain NADH dehydrogenase and the degradation of ubiquitinated proteins. The next highly significant GO term is “response to toxic substances”, and means any response reaction to a toxic stimulus and is regulated with glutathione metabolic process, oxidoreductase activity or UDP-glycosyltransferase activity. “Response to misfolded proteins” showed a DE of transcripts between the bulks, for example in the function of hydrogen peroxide removal, import into peroxisomes, degradation of ubiquitinated proteins, ATP:ADP antiporter activity, glutamine synthetase pathway and NADH dehydrogenase pathway. Nearly the same functions could be associated with the DE GO term “response of metal ions”. A DE of transcripts in the GO term “gluconeogenesis” was obtained, which is the formation of glucose from pyruvate, amino acids or glycerol. Transcripts associated with these GO terms have the function of ubiquitin protein ligase activity, catalysis of the isomerization of citrate to isocitrate, ATP:ADP antiporter activity, production of NADPH, multiubiquitin binding protein and some transcripts are part of the phospholipid metabolism. A significant DE between short and long VL in the GO term “photorespiration” was

found. Photorespiration is the oxidative photosynthetic carbon cycle of plants. Significant DE of transcripts were obtained as well in the GO terms “toxin metabolic process”, “response to poisonous compounds” and “proteasomal protein catabolic process”, the breakdown of a protein mediated by the proteasome. These GO terms are associated with nearly the same transcripts as those already named.

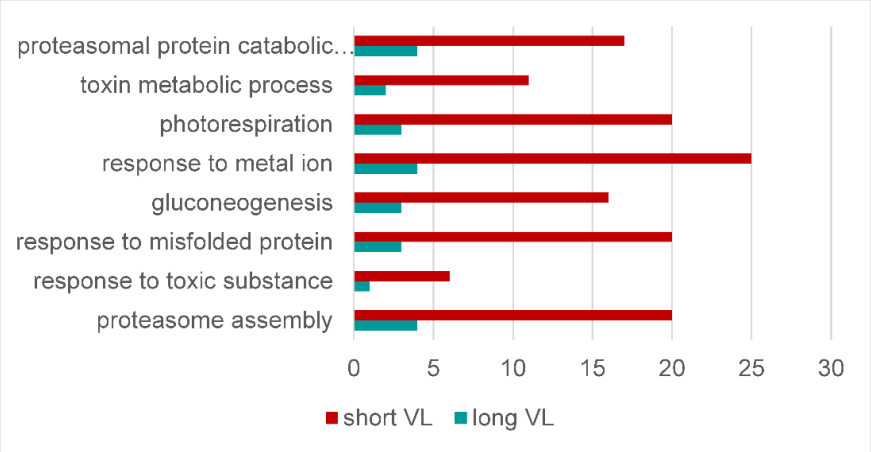


Figure 15: Significant DE GO terms (p<0.05) based on the Fisher’s exact test and a ratio of significant contigs between bulks with short (red) and long (green) VL higher than 2.

4.3.4.2 Senescence associated GO terms

GO terms associated with senescence were identified by literature review and further analyzed. The number of significant DE transcripts of interesting GO terms were counted in the comparison of the bulks with short and long VL as well between St and Mn bulks. In addition, the genotypes of both carnation types with short and long VL were compared. The different comparisons demonstrated that within the carnation types more DE transcripts were identified than between short and long VL bulks. The separation of both carnation types into short and long VL yielded fewer transcripts because fewer genotypes were part of the bulks. Mn bulks showed more DE transcripts than the St bulks (Table 12). The comparison of short and long VL and of carnation types shares only some DE transcripts with ethylene associated GO terms. For example, ACO1 (Aminocyclopropane-1-carboxylate oxidase) is only DE in the comparison of Mn bulk with St bulk. By considering both conditions (carnation type and VL) separately, no transcripts showed a significant DE in both bulks. In the comparison of St and Mn, more

genes with a direct influence on the biosynthesis of ethylene like EIN2 or ERF023 are DE.

In the following, DE transcripts with senescence associated GO terms in the comparison of short and long VL will be described. Five GO terms associated with ethylene were found (Table 12). In total, 18 significant DE transcripts ($p<0.05$) including ethylene responsive transcription factors (ERF1A, ERF8), response to stress associated transcripts like GAI (Gibberellic acid-insensitive mutant protein), DREB3 (Dehydration-responsive element binding protein) and CML37 (Calmodulin-like protein), which are involved in drought sensitivity and leaf transpiration, were detected. Nine DE transcripts associated with Gibberellins (GA) were found. These transcripts are associated with degradation of cell wall or development steps as seed germination or flowering. ABA has an inhibitory effect and induces stomatal closure and decreasing transpiration to prevent water loss. 28 DE transcription factors and genes responsible for the stomata closure and the turgor pressure were found. Interesting is MIF1 (Mini zinc finger protein 1), because it inhibits zinc finger homeodomain (ZHD) transcription factors, which are involved in the signaling pathways of GA, auxin and brassinosteroid signaling and is only DE in the St bulk. Transcripts associated with auxin have functions in cell wall thickening, xylem and phloem pattern formation, stomatal closure and some signaling pathways. In the transcriptome sequencing, 24 different DE transcripts were found which are associated with auxin. As described above, stress reactions are associated with senescence and the stress was caused by osmosis, oxidation or salt. In total, 229 DE transcripts were identified. The plants in both bulks suffered from stress and the significance of the difference in transcript abundance between the bulks with short and long VL suggests potentially different strategies in the contrasting bulks. 62 DE transcripts correspond to ubiquitin, a signaling protein of degradation. The biggest share of transcripts, which have the potential of being associated with a longer VL, are transcripts belonging to metabolism of carbohydrates with 323 significant DE transcripts in both bulks like for example Glycosyltransferase, STP4 (Sugar transport protein 4) or SS1 (Sucrose synthase 1). Polyamine (PA) has essential functions in all living cells. There are some suggestions that a lower amount of total PAs is associated with greater longevity (Ebrahimzadeh et al., 2013). It is noteworthy that no transcripts associated with PA are part of the bulk with long VL. 36 DE transcripts were found

in the GO terms associated with water stress. Some significant proteins were involved in water economy, stomatal closure or regulation of turgor pressure in cells. MYC2 is a common transcription factor of light, ABA and jasmonic acid (JA) signaling pathways. The transcription factor is only present in the bulk with short VL.

Table 12: Analysis of with senescence associated GO terms in the comparison of short and long VL as well in the comparison of carnation type (Mn and St) and combined conditions (Mn long / Mn short and St long / St short).

GO Term	Description	reads	Long VL *	Short VL *	St *	Mn *	Mn long *	Mn short *	St long *	St short *
GO:0009692	ethylene metabolic process	**22**	3	2	3	3	2	1	1	1
GO:0009693	ethylene biosynthetic process	**22**	3	2	3	3	2	1	1	1
GO:0009723	response to ethylene	**62**	8	8	8	9	4	3	3	5
GO:0009873	ethylene-activated signaling pathway	**28**	5	4	4	4	2	1	1	2
GO:0071369	cellular response to ethylene stimulus	**29**	5	4	4	4	1	2	1	2
GO:0009739	response to gibberellin	**20**	3	6	3	4	2	4	3	2
GO:0010476	gibberellin mediated signaling pathway	**14**	1	4	1	4	1	4	1	1
GO:0071370	cellular response to gibberellin stimulus	**14**	1	4	1	4	1	4	1	1
GO:0009737	response to abscisic acid	**116**	12	16	11	22	6	10	1	1
GO:0009738	abscisic acid-activated signaling pathway	**54**	9	4	7	11	4	5	5	3
GO:0071215	cellular response to abscisic acid stimulus	**59**	9	5	7	12	4	6	5	4
GO:0009733	response to auxin	**93**	12	7	9	20	4	6	10	6
GO:0009734	auxin-activated signaling pathway	**65**	10	5	5	15	3	3	9	6
GO:0071365	cellular response to auxin stimulus	**66**	10	5	5	15	3	3	9	6
GO:0006950	response to stress	**1005**	79	150	93	201	63	127	87	80
GO:0006970	response to osmotic stress	**161**	13	26	10	36	13	25	10	12
GO:0006979	response to oxidative stress	**195**	18	26	14	40	13	28	18	16
GO:0009651	response to salt stress	**142**	12	22	9	31	12	22	8	10
GO:0033554	cellular response to stress	**357**	28	47	31	68	23	32	30	28
GO:0034599	cellular response to oxidative stress	**28**	0	6	1	7	1	5	30	28

GO:0034976	response to endoplasmic reticulum stress	**69**	9	9	6	12	6	7	8	7
GO:0080134	regulation of response to stress	**117**	14	18	9	23	6	12	13	9
GO:0080135	regulation of cellular response to stress	**67**	11	9	8	15	5	6	11	5
GO:0006511	ubiquitin-dependent protein catabolic process	**158**	17	25	16	39	13	18	13	8
GO:0016567	protein ubiquitination	**164**	15	11	8	29	13	7	1	1
GO:0005975	carbohydrate metabolic process	**947**	94	143	92	195	96	105	75	115
GO:0009743	response to carbohydrate	**98**	4	10	8	15	11	8	2	10
GO:0016051	carbohydrate biosynthetic process	**317**	27	51	27	71	32	36	21	32
GO:0016052	carbohydrate catabolic process	**235**	26	48	26	58	28	31	19	31
GO:0034637	cellular carbohydrate biosynthetic process	**204**	19	27	17	46	22	18	13	18
GO:0044262	cellular carbohydrate metabolic process	**343**	35	52	31	83	42	33	1	4
GO:0044723	single-organism carbohydrate metabolic process	**621**	57	97	55	129	56	76	47	63
GO:0044724	single-organism carbohydrate catabolic process	**155**	15	38	17	41	18	27	12	19
GO:0046835	carbohydrate phosphorylation	**34**	3	7	1	13	3	5	3	8
GO:0071322	cellular response to carbohydrate stimulus	**41**	1	3	3	6	5	2	0	5
GO:1901135	carbohydrate derivative metabolic process	**556**	48	75	33	97	49	70	3	1
GO:1901136	carbohydrate derivative catabolic process	**27**	0	5	0	6	2	8	27	25
GO:1901137	carbohydrate derivative biosynthetic process	**354**	30	45	18	62	26	46	1	5
GO:0006595	polyamine metabolic process	**27**	0	9	2	5	1	6	1	3
GO:0006596	polyamine biosynthetic process	**15**	0	5	1	3	0	4	1	1
GO:0006598	polyamine catabolic process	**8**	0	3	1	2	1	2	0	1
GO:0006833	water transport	**43**	3	8	5	9	5	6	4	6
GO:0009414	response to water deprivation	**86**	9	15	9	18	4	14	6	5

4.3.4.3 Significant DE GO terms between carnation types

The smallest *p*-value for DE between GO terms up- and down-regulated in St and Mn bulks with a higher ratio than two, was obtained with the GO term called “secondary metabolite catabolic process”, which is the chemical pathway resulting in many of the chemical changes of compounds, which are not necessarily required for growth of cells (Figure 16). This includes some aspects of the ubiquitin and glutathione pathway. The “toxin catabolic process” as a reaction to toxic substances showed significant DE between both carnation types. The process is regulated via glutathione, proline, nitriles or ubiquitin. A DE of transcripts between the carnation types was detected in the GO term “negative regulation of translation”, which means any process resulting in the prevention of the formation of proteins. The GO term “dicarboxylic acid biosynthetic process” includes chemical reactions of two carboxyl (COOH) groups like the phosphorylation of serines, threonines or tyrosines and is DE between both carnation types. The transcripts in this GO term are responsible for cell wall thickening, cell wall biogenesis and stem elongation. Xylan is a group of hemicelluloses and is an important part of cell walls of plants. Interestingly, the transcripts associated with xylan biosynthetic processes are upregulated in the Mn bulk.

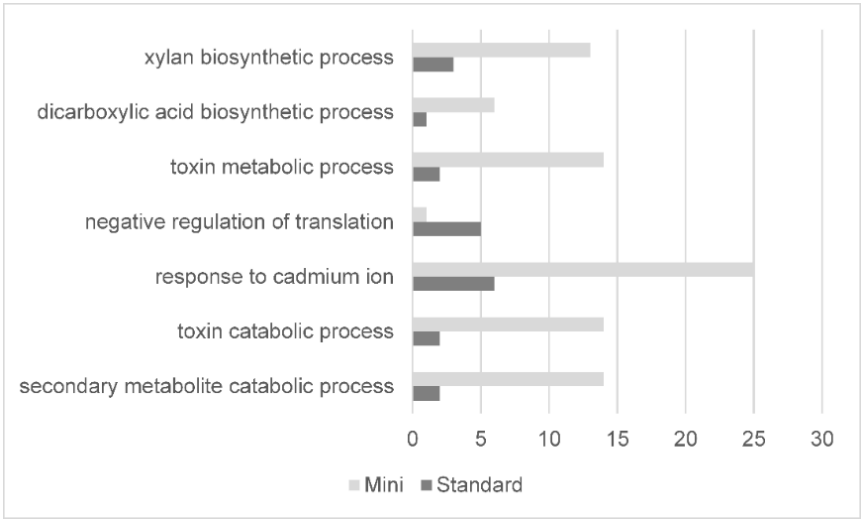


Figure 16: Significant GO terms (p<0.05) based on the Fisher’s exact test and a ratio of significant contigs between bulks of standard and mini carnation higher than 2.

4.3.5 SNP detection in bulks

SNPs with different allele frequencies were identified by using filter methods and randomly selected results are present in Table 13. SNP 1 show good results in the bulks when only considering the VL. In the bulk with long VL (allele frequency:

80 %), the reference base (ref base, for example base A) is more frequent than the variable base (var base, for example base C). The bulk with short VL shows the reverse result (allele frequency: 11 %). More reads of the reference base (A) were found in the bulk with short VL. By splitting the reads into Mn and St, the results are similar, except that the contrast in the bulk of Mn with a short VL (Mn short VL) is reduced, caused by a small count of reads. SNP 2 result in a high contrast in the bulk of both carnation types, but the additional separation of the VL contrasting bulks into St and Mn yielded no result for the St bulk (n.a.). The SNP 3, SNP 4 and SNP 7 showed good results, but only in the St bulks. In the Mn bulk, the contrast of the reference and variable base between the bulks could not be reproduced. SNP 5 showed a high contrast of sequence variants of MACE contigs between short and long VL, but in the case of only using the St bulk, the contrast seems high, but the reads are too low in the bulk with long VL. The desired contrast of SNPs between bulks was detected in SNP 6, but the contrast is not present anymore in the bulks for St.

Table 13: Selected SNPs (ref_base vs. var_base) between the bulks with short and long VL and the same SNPs separated between standard (St) and mini carnation (Mn) and allele frequencies in the bulks. No result available (n.a.)

	SNP_1	SNP_2	SNP_3	SNP_4	SNP_5	SNP_6	SNP_7
long VL ref_base	40	43	37	122	83	155	101
long VL var_base	10	1	114	4	1	41	36
Allele frequency	80	98	25	97	99	79	74
short VL ref_base	10	0	88	19	52	4	5
short VL var_base	78	29	45	240	129	48	63
Allele frequency	11	0	66	7	52	4	7
Mn_long VL ref_base	3	43	22	1	81	154	25
Mn_long VL var_base	1	0	29	11	1	15	19
Allele frequency	75	100	43	8	99	91	57
Mn_short VL ref_base	8	0	16	54	51	2	15
Mn_short VL var_base	10	29	38	6	30	17	5
Allele frequency	44	0	30	90	63	11	75
St_long VL ref_base	37	n.a.	92	3	2	1	11
St_long VL var_base	9	n.a.	8	111	0	26	82
Allele frequency	80	-	92	3	100	4	82
St_short VL ref_base	2	n.a.	29	186	1	2	48
St_short VL var_base	68	n.a.	50	13	99	31	0
Allele frequency	3	-	37	93	1	6	100

4.4 Discussion

4.4.1 Quality of sequencing

Compared to the sequencing statistics from transcriptome analysis (RNA-Seq) of carnation from Tanase et al. (2012), our sequencing showed a higher sequencing depth. Tanase et al. (2012) had a total number of 1,435,398 reads and 71,864,367 reads were sequenced. As well, the number of contigs was higher. Twenty thousand more contigs were assembled. In this manuscript, an average transcript length of 1258 bp (226–24,374) was obtained. Tanase et al. (2012) report a mean length of 605 bp (117–3850). Compared to Wan et al. (2015), 35 % more contigs were annoted to known proteins by using a combination of RNA-Seq and MACE, despite the fact that the sequencing of Wan et al. (2015) resulted in more assembled contigs. The length of the contigs is comparable. A comparison of MACE data with published data is not possible because it is the first time that the carnation transcriptome was sequenced with this sequencing method. Compared to other publications the combination of RNA-Seq and MACE show good transcriptome sequencing results. This is important for finding molecular markers for VL.

4.4.2 Differences between VL bulks smaller than expected

Taking a closer look to carnation type bulks and VL bulks, the differences were small. More than 80 % of all contigs were present in all four bulks. The phenotyping of the genotypes resulted in a VL between 6 and 14 d. All contigs of the Mn bulk with a long VL share one or more contigs of the other bulks. The analysis of RNA-Seq data showed more than 90 % common contigs in all conditions (data not shown). More transcripts are only expressed in the bulks with short VL, which indicates an upregulation of transcripts in the genotypes, which are associated with a short VL. More than 1.800 unique transcripts are available, which have the potential of influencing the VL in a negative way.

4.4.3 GO slims

The huge amount of data produced by a transcriptome analysis was managed by using GO terms, which are controlled vocabularies of defined terms representing gene product properties. The analysis of GO slims in the category Biological Process showed processes most present in the transcriptome of carnation. One of the first steps in the cascade of senescence is the mobilization of nutrients from

senescing tissues to storage tissues and a regulated degradation of proteins. It is interesting that “transport” is one of the most represented GO slims in the transcriptome of carnation. The “protein metabolic process” was associated with ubiquitin, which has the function of signaling the degradation of proteins. ”Response to stress” is induced by plant hormones like ABA, GA or Auxin and those hormones, which have a regulation part in the VL of plants. In the category Cellular Component, the GO slim “membrane” was dominant. The membrane in plants is a highly elaborated structure and is responsible for the exchange between cells, cell walls and the external environment. This includes the uptake of water and nutrients and signaling molecules. Some of these functions are senescence associated.

4.4.4 Differences between bulks with short and long VL

Among significant DE transcripts of GO terms showing differences between the bulks with short and long VL, only similar transcripts were detected. This suggests that these genes play a role in the senescence of carnation. Most frequent were genes associated with mitochondrial membrane respiratory chain NADH dehydrogenase. This points to sugar status of the plants and the significant GO term gluconeogenesis, which will be discussed later. In all significant GO terms, the degradation of ubiquitinated proteins was involved. During senescence, ubiquitin is upregulated (Belknap and Garbarino, 1996). The metabolism changes within senescent tissues lead to a general mobilization of nutrients from senescing tissues to storage tissues within the plant (Belknap and Garbarino, 1996). This mobilization includes a general loss of protein and it is interesting that glutathione-s-transferase transcription is ethylene-inducible and a correlation to senescence of plants are obvious (Itzhaki et al., 1994). These significant differences in the GO terms were caused by an expression of transcripts in the bulk with short VL. The process of senescence is already initiated in the bulk with short VL. The negative effect of polyamines on VL is controversial (Ebrahimzadeh et al., 2013). Our data suggest that no transcripts belonging to polyamines are present in bulk with long VL. MYC2 is a transcription factor, which promotes chlorophyll degradation by directly activating chlorophyll catabolic genes (Zhu et al., 2015). MYC2 is only activated in the bulk with short VL, which further proves that in this bulk the senescence is already introduced. The VL of carnation stems is highly dependent on respiration rates and a high exposure of tissues to dehydration

(Ebrahimzadeh et al., 2008). The GO terms “response to water deprivation and water transport showed both more activated transcripts in the bulk with short VL. This indicates that in these tissues a kind of water stress was already present.

4.4.5 Ethylene enrichment in the bulks with short and long VL

Some transcripts and direct members, like ethylene-responsive transcription factors (ERF1A, ERF8), were involved in the pathway of ethylene. Nevertheless, in the transcriptome sequencing with MACE only plant material from buds and stems was harvested. The direct “response to ethylene” was not yet induced. In the bulk with short VL, the accumulation of ethylene was further advanced compared to the bulk with long VL. We assume an earlier accumulation of ethylene, which results in a shorter VL.

4.4.6 Sugar metabolism and VL

The comparative analysis of the transcriptome data showed a relation between the sugar metabolism and VL. One of the most significant DE GO terms was “gluconeogenesis”. In the literature, some theories of the function of sugar in the senescence of plants are available. One theory suggests that sugar starvation induces senescence through increased ethylene sensitivity (van Doorn, 2004).

Other work demonstrated a complex molecular crosstalk between sugar and hormone signaling pathways (Dar et al., 2014). At the first stages of senescence, the sugar content is still high. One of the first steps of senescence is the degradation of polysaccharides and the mobilization of sugars and nitrogenous compounds through the plant (van Doorn, 2004). “Transport” was one of the ten most frequent GO slims and this indicates that the analyzed plants are in mobilization status. In our data, all found DE transcripts are more highly expressed in the bulk with short VL. This indicated that senescence reaction of degradation of carbohydrates and the mobilization and transport of sugar within the plant are more advanced in the genotypes with short VL. A complicated crosstalk between sugar and plant hormones is assumed. In our data, a crosstalk between sugar and GA was detectable in the bulks with short VL, because transcripts belonging to this GO term were upregulated compared to the bulk with long VL. ABA and auxin showed an upregulation in the bulk associated with a long VL.

4.4.7 Senescence as a kind of stress response

In the processes of senescence and stress response, the same hormones like ABA, GA, JA and others are involved. In total, 229 transcripts showed a DE between the bulks with short and long VL. A different process of senescence between the bulks with short and long VL was assumed. Some of the named plant hormones have a role in the early stages of senescence, like auxin, which acts as negative regulator of gene expression of senescence-associated transcripts. The MACE sequencing was based on plant material of buds and stems and no transcripts of senescence have to be included. The transcripts are significantly upregulated most of the time in the bulk with short VL and this suggests that in these genotypes the senescence already sets in at early stages. One of the earliest changes in the cell structure is breakdown of chloroplasts. More than fifty DE transcripts correspond to ubiquitin, a signaling protein of degradation. Some of the found DE transcripts have interesting functions in cell wall thickening, xylem and phloem pattern formation or stomatal closure, which are potential processes differentially regulated between genotypes with a short or long VL.

4.4.8 Differences between St and Mn

The development of Mn was initiated from a St with strong side shoots. No literature is available about differences between both carnation types. The analysis of DE transcripts of both carnation types showed a difference between the presence of xylan biosynthesis, which is responsible for cell wall thickening, cell wall biogenesis and stem elongation. A higher xylan content is associated with more stable cell walls. A higher expression of transcripts involved in the xylan biosynthesis was obtained caused by the more stable side shoot of Mn compared to St with only one main stem. The DE transcripts between St and Mn bulks of the pathway of ubiquitin, which is a degradation initiating protein for other proteins, were very frequently found. During senescence, a change of metabolism and cell structure is present, which includes the mobilization of nutrients from senescing tissues to storage tissues and a regulated degradation of proteins (Belknap and Garbarino, 1996). The second pathway is the redox reaction of glutathione, where a correlation to the ethylene pathway is assumed as already mentioned above.

4.4.9 SNP identification different in carnation type

The different counts of the reference base and variable base between the bulks of VL, which are additionally separated into St or Mn, further support a difference between both carnation types. In some SNPs, the desired contrast of sequence variants between the bulks was obtained by taking both carnation types into account. Nevertheless, in some cases, the detected SNP could be validated only in St or Mn. The separation of St and Mn is comparable to that of spring and winter wheat. Some varieties were more adapted to the cold period in winter and the breeding and selection of these varieties eventually led to the current winter wheat. The genetic variation between spring and winter wheat was proven in some regions of the genome by measuring linkage disequilibrium within populations (Chao et al., 2010). To introduce SNP-based marker-assisted selection into carnation breeding, molecular markers for both carnation types have to be established.

4.5 Conclusion

The comparative analysis of GO terms of the carnation transcriptome exposed some VL regulating mechanisms. The results of the literature review of the senescence pathway in carnation were partly confirmed by the results of the analysis of GO terms and differences between short and long VL were identified. The comparative analysis of transcriptome data showed that St and Mn have to be analyzed independently. SNP markers have to be established for both carnation types separately. The differences between St and Mn suggest that selection and breeding had a differentiating impact on each carnation type and the assessment of the extent of the differences provided valuable information for future development of molecular markers in carnation.

5. Genetic and phenotypic correlation for breeding relevant traits in *Dianthus caryophyllus L.*

Maike Boxriker, Jens Möhring and Hans-Peter Piepho

The original publication is available at

https://www.sciencedirect.com/science/article/abs/pii/S0925521418300711

Article history:

Received 20 January 2018
Received in revised form 24 March 2018
Accepted 25 March 2018
Available online 17 May 2018

Keywords: Phenotypic correlation, Genetic correlations, Indirect selection, Vase life, Dianthus caryophyllus L.

Abstract

Indirect selection is a useful tool for the selection of complex traits like vase life (VL) of cut carnations. Phenotypic and genetic correlations were calculated for breeding relevant traits like VL, flower diameter (FD) of standard carnation, stem length (SL) and count of flowers (CF) in mini carnations, phenotyped in three different experiments. The conducted experiments are peculiar in that some traits are measured in a second phase. Data were analyzed in bivariate mixed models by considering in which phase the trait was measured. Some relevant genetic correlations were found in VL and FD and SL and CF, but indirect selection would show no benefit compared to direct selection. However, the combination of direct and indirect selection could improve breeding of long-lasting cut carnations.

5.1 Introduction

Carnation (*Dianthus caryophyllus L*.), a member of the family *Caryophyllaceae*, is one of the major cut flower species worldwide (Yagi et al., 2013). As an important cut flower, carnation is cultivated in temperate regions all around the world. Carnations have the highest economic value among all species in the floricultural industry. They can be divided into two types: standard carnation (with one flower on the stem) and mini carnation (with multiple flowers per stem) (Köhlein, 1990). In the past, breeding of carnation was mainly based on floral traits such as flower colour, flower diameter (FD) and duration of flowering. Some economic traits such as productivity had a higher priority compared to vase life (VL) (Onozaki et al., 2001). Nowadays the importance of VL breeding is becoming a major focus due to long transportation times between producer and consumer. It is increasingly important to broaden genetic variability to breed new cultivars (Nimura et al., 2008). Understanding the association between traits such as VL or FD is important for direct or indirect selection (Shimelis and Hugo, 2011). Correlated traits could be used as indirect markers for quantitative traits (Ofori, 1996). VL is a highly complex trait to measure and is quantitative in nature, involving multiple genes with additive effects. In some cases, indirect selection is an interesting method to improve response to selection in a hard-to-measure trait or when the phenotyping involves high costs. Within a genetic correlation between two traits, the mean phenotypic value will be changed not only for the target trait of selection, but indirectly other so called auxiliary traits are also influenced (Bos and Caligari, 1995). The best results could be obtained by using a combination of indirect and direct selection (Falconer and Mackay, 1996). The auxiliary trait can provide an indication of the performance for the target trait or time and effort can be saved measuring the target trait (Bos and Caligari, 1995).

Bivariate mixed models are suited for the analysis of two traits measured on the same observational units for the purpose of determining correlations between the traits (Piepho and Möhring, 2011). Bivariate analysis is straightforward to implement for experiments performed in a single phase. By contrast, analysis becomes more complex for two-phase experiments. Two-phase experiments were first described by McIntyre (1955) for a trait cannot be measured directly such that a further stage of experiment has to be implemented for evaluation. Phase-specific

analyses are important in two-phase experiments like VL trials in carnation (Boxriker et al., 2017b) or rooting experiments in *Pelargonium zonale* (Molenaar et al., 2017). The objective of our study was to implement bivariate analysis in two-phase experiments in carnation with the goal to calculate genetic correlations for breeding relevant traits. The aim was to introduce indirect selection for the hard-to-measure traits.

5.2 Material and methods

5.2.1 Phenotypic data

Details of the experimental design are described in Boxriker et al. (2017b). Briefly, for the cultivation of carnations three two-phase experiments each with a phase in the greenhouse (cultivation) and a phase in the laboratory (VL assessment) were used. In each year a set of approximately 500 genotypes were cultivated using different experimental designs. In the first experiment (E1), an α-design with four replicates was used in the first phase (John and Williams, 1995). In the following experiments (E2 and E3), resolvable row-column designs were used. Each replicate comprised 57 subsurface boxes, which were used as incomplete blocks in E1 or as rows otherwise with nine planting holes each (used as positions in the blocks or columns). In experiments E1 to E3 a set of 501, 495 and 556 genotypes were tested, respectively. Each genotype was present in four complete replicates. The carnations were planted in April in the years 2013, 2014 and 2015.

For VL assessment in the second phase, eight carnation stems per genotype (two per replicate) were harvested at flower development stage before blooming. The stem length (SL) was measured in cm and the flowers stem^{-1} of mini carnations (CF) were counted (Figure 17) for all harvested stems. To simulate shipment of carnations, four stems (one of each replicate) were stored for two weeks at 4 °C, while the other stems were not stored and tested immediately. When both stems of one replicate were not harvested together, it was decided randomly if the first stem is stored or put in water directly. When the stems were ready for VL assessment, they were trimmed to 50 cm and lower leaves were removed and then put in water. Fresh tap water was used. The water was refilled to hold water level constant in the vases. The flowers were kept under controlled conditions: 12 h

photoperiod of 300 (E1) or 800 (E2+E3) lux provided by LED lamps and a constant temperature of 20.1±0.5 °C. Ventilators in E1 exchanged air. In the following experiments, a 1.5-fold air exchange per hour was effected by an air conditioning system. The diameter of open flowers was measured in cm. Each flower was checked daily for senescence symptoms like wilting, rolling-in of petals or discoloration of the petal margins to phenotype for VL of carnations. In Figure 17 the phenotyping scheme is summarized.

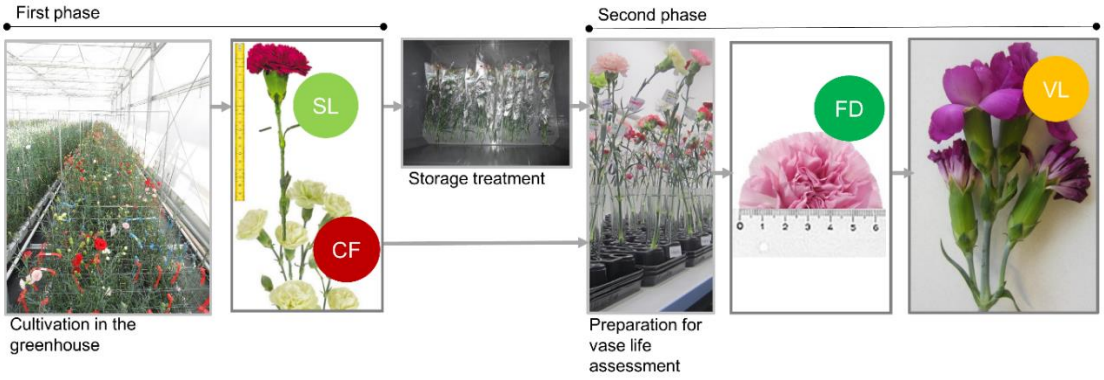


Figure 17: Phenotyping scheme of carnation stems for first-phase and second-phase traits in three experiments (E1, E2 and E3) of vase life (VL) in d.

Flower diameter (FD) in centimetre of standard carnations; stem length (SL) in centimetre and count of flowers (CF) of mini carnations

5.2.3 Data analysis

The data of each experiment were analyzed by using a model representing the two phases which is described in detail in Boxriker et al. (2017b). VL and FD are second-phase traits, means they were measured in the second phase of the experiment, because the stems were harvested before bud opening and put into vases, where the trait was measured after opening of the flower. The traits SL and CF are first-phase traits and are not affected by the storage treatment. The traits were measured in the first phase before the storage treatment and no second-phase effects have to be taken into account in analysis. Table 14 provides a summary of all traits and a description of the design for their assessment.

Table 14: Summary of the influence of storage treatment and the need of a second phase in the model.

Trait [§]	Storage	Two phases	Design for assessment of trait
VL	yes	yes	Two-phase experiment with storage treatment
CF	no	no	Single-phase experiment without storage treatment
SL	no	no	Single-phase experiment without storage treatment
FD	yes	yes	Two-phase experiment with storage treatment

[§] vase life (VL) in d; flower diameter (FD) in centimetre of standard carnations; stem length (SL) in centimetre; count of flowers (CF) of mini carnations.

Following the procedure described in Piepho et al. (2003), the mixed models for the different traits were formulated. Fixed and random effects of these models are separated by a colon with fixed effects appearing before the colon. The dot operator (•) is used to define crossed effects and the residual term is underlined.

5.2.3.1 Single-phase model without storage treatment

The simplest model was used for the first-phase traits CF and SL because no storage treatment and second phase must be considered (Table 14). The model for this case is formulated in Table 15, where Y is the trait of interest like CF or SL. CT is a factor representing the differentiation between standard and mini carnation and genotypes (GENO) were nested in carnation type (CT•GENO). The factor CT must used for the traits VL and SL, because these traits were measured in both carnation types. The α- and row-column designs require inclusion of a fixed effect for complete replicates (REP). The subsurface boxes (BOX) and the position within the subsurface box (P) are factors representing the crossed incomplete blocks and must be nested within replicates to model the corresponding effects (REP•BOX, REP•BOX•P). From each genotype in each replicate, two stems were harvested, which is why a factor for the first or second stem was added (W), which is used to define the residual error (underlined). The mixed model was fitted using the MIXED procedure of the SAS System (9.3; SAS Institute, USA). As shown in Boxriker et al. (2017b), analysis of the datasets E2 and E3 without using the second blocking effect (REP•P) of the row-column design was more effective than using the full model and data was analyzed like an α-design (REP•BOX•P).

In order to account for differences in temperature at the harvest of stems for the complete duration of the experiments, temperatures were measured hourly 24 h before the harvest of each stem. The temperatures were averaged over each day and used as a covariate in the model (GT).

Table 15: Model set up for the analysis of first-phase traits measured in the experiments E1, E2 and E3 without storage treatment.

Y = GT + REP + CT + CT•GENO : REP•BOX + REP•BOX•P + <u>REP•BOX•P•W</u>
Notes: trait of interest (Y), temperature in greenhouse (GT), complete replicate (REP), carnation type (CT), genotype (GENO), blocking unit in greenhouse (BOX), position in subsurface box (P), first or second harvested stem (W). The residual error term is underscored.

5.2.3.2 Two-phase model with storage treatment

FD was observed after the preparation for the VL assessment, because the flowers must be fully opened. The storage treatment influenced the measured second-phase traits and the analysis of variance showed a difference in VL and FD in all experiments (Table 19). The treatment storage (STO) was added as fixed effect into the model. Except for the traits CF and FD, the factor CT was used to differentiate between standard and mini carnation. The genotype was nested in CT. The treatment factors were modelled by fixed effects, except for the interaction of genotypes and storage (STO•CT•GENO), which was modelled as random, because in some cases the stems of one genotype were only treated with or without a storage treatment, meaning that analysis with fixed interaction would not have allowed computing marginal means for genotypes. FD was only measured in the case of standard carnations, so the carnation type (CT) is not considered.

In the second (laboratory) phase, culture trays were used as experimental units with 148 blocks (TRAY) and eight positions (PT). In E1, replicates (REP) from the greenhouse were transferred to the laboratory, allocating the harvested stems of one replicate together in not occupied blocks. In E2 and E3, a complete randomization was conducted in the laboratory. The day at which the stems were put in vases (DV, counted days) was used as random block effect for recovery of inter-day information, because the stems were allocated to some positions twice or more often. All effects in the second phase were taken as random. To be able to analyze two traits with the same model, when traits were measured both in the

greenhouse phase and in the laboratory phase, a dummy variable called P2 was introduced. P2=0 was used for traits with no second phase and P2=1 for second-phase traits. The residual error was defined using block factors of the first phase, but it also comprises the residual error of the second phase, with which it is confounded.

Table 16: Model set up for the analysis of second-phase traits measured in the experiments E1, E2 and E3 with storage treatment.

E1: Y = GT + REP + CT + STO + CT•GENO : STO•CT•GENO + REP•BOX + REP•BOX•P + P2•DV + P2•DV•REP•TRAY+ P2•DV•REP•TRAY•POS + <u>REP•BLOCK•P•W</u> [$]
E2 + E3: Y = GT + REP + CT + STO + CT•GENO : STO•CT•GENO + REP•BOX + REP•BOX•P + P2•DV + P2•DV•TRAY + P2•DV•TRAY•POS + <u>REP•BLOCK•P•W</u> [$]
Notes: trait of interest (Y), temperature in greenhouse (GT), complete replicate (REP), first or second harvested stem (W), carnation type (CT), storage (STO) of two weeks: yes or no, genotype (GENO), blocking unit in greenhouse (BOX), position in subsurface box (POS), start of lab phase (DV), blocking unit in lab (TRAY), position in tray (POS). The residual error term is underscored. [$] Residual error terms of both phases were defined, the effect P2•DV•(REP•)TRAY•POS was dropped from the model.

5.2.4 Heritability

The data was unbalanced and hence the common formulae for heritability are invalid (Holland et al., 2003).The *ad hoc* heritability was calculated in univariate analysis as shown in formula 1, following the procedure given in Piepho and Möhring (2007).

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \bar{v}/_2} \ , \qquad |1|$$

where σ_g^2 means the genotypic variance and $\bar{v}$ is the mean variance of a difference of two adjusted genotype means, corresponding to the best linear unbiased estimators (BLUE) (Searle et al., 1992).

5.2.5 Phenotypic correlation

Pearson's correlation coefficient was calculated between adjusted genotypic means of different traits.

5.2.6 Bivariate mixed model analysis

A univariate model for an α-design can be formulated as

$$y_{ijk} = \mu + g_i + r_j + b_{jk} + e_{ijk} \quad |2|$$

where y_{ijk} is the phenotypic observation of the *i*-th genotype in the *k*-th complete block within the *j*-th replicate, μ is the general mean, g_i is the effect of the *i*-th genotype, r_j is the effect of the *j*-th replicate, b_{jk} is the effect of the *k*-th incomplete block within the *j*-th replicate and e_{ijk} is the residual plot error associated with y_{ijk}. This model can be extended to two traits by adding a subscript *t* for the traits to each of the effects of the univariate model: $y_{ijkt} = \mu_t + g_{it} + r_{jt} + b_{jkt} + e_{ijkt}$. This model was fitted defining a new factor called TRAIT, which codes the first and second trait in the bivariate data set. Each effect of the univariate model has to be crossed with TRAIT as explained in Piepho and Möhring (2011).

As shown in Table 14, different cases need to be considered for statistical analysis depending on the phase in which a trait was assessed and how this was done. This leads to different models for the combinations of two traits. The combination with second-phase traits VL and FD in a bivariate model means that the experimental units of the second phase have to be included. The dummy variable P2 was used as on-switch (P2=1) for activating the units of the second phase for VL and FD. And conversely, it was used as off-switch (P2=0) in the case of first-phase traits SL and CF (Piepho et al., 2006). The bivariate model of the traits VL and FD do not require the variable P2, because both traits are measured in the second phase. In the case of CF and SL, the second phase plays no role, because the data were collected directly after the harvest and the second phase had no influence on these traits. The traits FD and CF were measured in only one of the carnation types. Therefore, the effect CT can be excluded from the model. Despite, VL and SL, the first-phase traits were phenotyped before the storage treatment and STO is not included in the models. With the aim to calculate the genetic correlation, the genotype (GENO) and all effects crossed with genotype should be modelled as random effects. In Table 17 the model for bivariate analysis is presented. Effects marked in grey either must be in- or excluded for each bivariate analysis, depending on the traits and their design of assessment (Table 14).

Table 17: Model set up for the bivariate analysis of first-phase and two-phase traits measured in the experiments E1, E2 and E3 with and without storage treatment, depending on the traits and their design of assessment.

E1: Y = TRAIT + TRAIT•GT + TRAIT•REP + TRAIT•CT + TRAIT•STO : TRAIT•CT•GENO + TRAIT•STO•CT•GENO + TRAIT•REP•BOX + TRAIT•REP•BOX•P + P2•TRAIT•DV + P2•TRAIT•DV•REP•TRAY + P2•TRAIT•DV•REP•TRAY•POS + TRAIT•REP•BOX•P•W $
E2 + E3: Y = TRAIT + TRAIT•GT + TRAIT•REP + TRAIT•CT + TRAIT•STO : TRAIT•CT•GENO + TRAIT•STO•CT•GENO + TRAIT•REP•BOX + TRAIT•REP•BOX•P + P2•TRAIT•DV + P2•TRAIT•DV•TRAY + P2•TRAIT•DV•TRAY•POS + TRAIT•REP•BOX•P•W $
Notes: trait of interest (Y), temperature in greenhouse (GT), distinction between both traits (TRAIT), complete replicate (REP), first or second harvested stem (W), carnation type (CT), storage (STO) of two weeks: yes or no, genotype (GENO), blocking unit in greenhouse (BOX), position in subsurface box (P), start of lab phase (DV), blocking unit in lab (TRAY), position in tray (POS), dummy variable second phase (P2). The residual error term is underscored. Marked in grey are effects, which are dependent on the traits and it is possible that they must be excluded from the model. $ Residual error terms of both phases were defined, the effect P2•TRAIT•DV•(REP•)TRAY•POS was dropped from the model.

5.2.7 Genetic correlation

The variances and covariances, needed for the calculation of the genetic correlation (Falconer and Mackay, 1996), were calculated in bivariate model in SAS 9.4 with the MIXED procedure (Piepho and Möhring, 2011). Starting values for the variance components were calculated by univariate analysis of each trait. Starting values for the covariance of two traits were computed based on a univariate analysis of the sum of both trait values using

$$var_{X+Y} = var_X + var_Y + 2cov_{XY}, \quad |3|$$

where var_{X+Y} is the variance of the sum of both traits. The variances of both traits and their sum (var_X, var_Y and var_{X+Y}) were estimated by univariate mixed model analysis. The calculated values were used as start values for bivariate analysis of different traits. The UNR structure was used to model the variance-covariance matrix using the variances of the traits and the pairwise correlations.

Some bivariate analyses could not be done in SAS, because of insufficient memory. The problem was solved by using ASReml v3.0 (2000-2017 VSN International Ltd). In a supplementary file, example codes for bivariate analyses for SAS (Supplemental 2) and ASReml (Supplemental 3) are presented. According to Bos and Caligari (1995) genetic correlation can be evaluated for the use of indirect selection via the relative selection efficiency (RSE).

$$RSE = \frac{i_Y\, r_A\, h_Y}{i_X\, h_X}. \qquad |4|$$

Assuming the same selection intensity ($i_Y = i_X$) and with the square root of the heritabilities (h_Y and h_X) given in Table 20 and the calculated genetic correlation (r_A) in Table 22, the ratio can be calculated for each bivariate analysis (Falconer and Mackay, 1996). When RSE > 1 indirect selection will be a better option. Our use of equation (5) constitutes an approximation because strictly speaking both the numerator and denominator of the equation require a variance-balanced design (Piepho and Möhring, 2007).

5.3 Results

5.3.1 Influence of carnation type and storage treatment

The difference between standard and mini carnations in different traits is shown in Table 18. A highly significant effect (p<.0001) of carnation type to VL was detected. There was a difference between 0.5 and 0.2 d. In the three experiments (E1, E2 and E3), the standard carnations showed longer stems than the mini carnations. In the second year, the difference was smaller.

Table 18: Influence of carnation type (Mn – Mini carnation, St – standard carnation) to the traits vase life (VL) in d and stem length (SL) in cm (adjusted means) with standard error in all experiments (E1, E2 and E3).

	Trait [§]	VL			SL		
	Trial	E1	E2	E3	E1	E2	E3
Mean	Mn	10.1	9.9	9.7	60.4	54.4	59.4
	St	9.6	9.7	9.4	63.6	55.6	62.8
Standard error	Mn	0.1	0.1	0.1	0.2	0.2	0.2
	St	0.1	0.1	0.1	0.3	0.3	0.2
p-value [$]		*<.0001*	*<.0001*	*<.0001*	*<.0001*	*<.0001*	*<.0001*

[§] vase life (VL) in d; stem length (SL) in centimeter;

[$] for t-test comparing means in the same column.

The influence of storage and transportation of carnations is described in Table 19. From each plant of each replicate, two stems were harvested and one of them

was stored for two weeks. Stored stems showed a decreased VL. In E1 and E3 there was a difference of about one day between the stored and unstored stems, and in the second year the decrease was two days. The storage treatment also showed a highly significant effect (p<.0001) on the FD. In the three experiments, FD of not stored stems was 0.03, 0.05 and 0.18 cm bigger, respectively, than the stored stems (two weeks, 4 °C).

Table 19: Influence of storage (two weeks at 4 °C) to the traits vase life (VL) in d and flower diameter (FD) in cm (adjusted means) with standard error in all experiments (E1, E2 and E3).

	Trait [§]	VL			FD		
	Trial	E1	E2	E3	E1	E2	E3
Mean	not stored	9.9	10.5	10.0	6.2	6.7	6.4
	stored	8.9	8.4	8.5	6.2	6.7	6.2
Standard error	not stored	0.2	0.1	0.1	0.1	0.03	0.04
	stored	0.2	0.1	0.1	0.04	0.03	0.04
p-value [$]		*<.0001*	*<.0001*	*<.0001*	*<.0001*	*<.0001*	*<.0001*

[§] vase life (VL) in d; flower diameter (FD)

[$] for t-test comparing means in the same column.

5.3.2 Characterization of measured traits in carnation

Mean response of measured traits (VL, FD, SL and CF) in E1, E2, E3 are presented in Table 20. In all experiments, the different genotypes showed a high range between the best and worst result in all traits. In the case of VL, a range of 10 d (E1), 13 d (E2) and 15 d (E3) were detected. The range of FD was between 2.5 and 2.9 cm. The mean response of the trait SL showed especially in E3 a high range between the longest (96 cm) and smallest (35.4 cm). In the case of CF, a range of 12 flowers per stem was observed in all experiments. The mean standard error of a difference (SED) can be used as a measure of precision of the experimental design. The smallest SED were found in E3. In E1, the SED for the trait CF is higher compared to the other experiments. Compared to other traits, the genotypic variance (σ_g^2) is relatively high in the traits CF and SL. The heritability (H^2) is between 0.50 and 0.89 in all traits and all experiments. The highest heritability (0.89) was calculated in SL in E3. The heritability of the trait FD was small compared to other traits.

Table 20: Mean response and rank of the ten best and ten worst carnations in all experiments (E1, E2 and E3).

Genotype	VL			FD			SL			CF		
rank §	E1	E2	E3	E1	E2	E3	E1	E2	E3	E1	E2	E3
1 good	15.60	17.68	19.79	7.82	8.05	7.68	85.05	74.00	95.96	15.53	14.59	14.65
2 good	15.34	17.04	17.62	7.42	8.01	7.51	84.96	73.94	91.05	14.75	13.39	14.28
3 good	15.03	16.33	17.24	7.41	7.96	7.48	79.61	72.48	88.51	14.64	13.31	13.94
4 good	15.01	16.15	17.22	7.34	7.95	7.45	79.28	71.70	84.28	14.40	13.23	12.85
5 good	14.87	16.09	17.06	7.31	7.81	7.44	78.78	71.64	82.50	14.03	12.79	12.65
6 good	14.48	15.94	16.54	7.22	7.76	7.32	78.46	71.60	78.92	13.96	12.53	12.64
7 good	14.22	15.93	16.25	7.17	7.75	7.31	78.43	71.49	78.72	13.71	12.51	12.54
8 good	14.16	15.80	16.25	7.09	7.75	7.15	78.06	70.74	78.67	13.45	12.45	12.38
9 good	14.15	15.61	15.66	7.09	7.74	7.06	77.14	70.50	78.45	13.38	12.24	12.35
10 good	14.07	15.39	15.55	7.08	7.70	7.02	77.12	70.23	78.18	13.04	11.75	12.20
1 bad	6.33	5.62	5.41	5.36	5.94	5.47	49.37	42.85	42.00	4.18	4.14	3.64
2 bad	6.27	5.62	5.31	5.33	5.93	5.46	48.92	42.36	41.91	4.04	4.03	3.47
3 bad	6.26	5.50	5.25	5.32	5.93	5.45	48.86	41.98	41.84	3.98	4.02	3.47
4 bad	6.24	5.19	5.16	5.31	5.87	5.41	48.69	41.95	41.56	3.83	3.35	3.35
5 bad	6.21	5.11	5.09	5.29	5.86	5.31	48.51	41.66	40.10	3.81	3.32	3.30
6 bad	6.16	5.09	5.07	5.26	5.85	5.29	48.51	40.87	40.05	3.62	3.29	3.22
7 bad	5.97	5.07	5.01	4.97	5.83	5.10	48.37	40.86	39.10	3.60	3.23	3.01
8 bad	5.90	4.50	4.83	4.97	5.74	5.08	46.77	40.84	38.78	3.57	3.15	2.68
9 bad	5.54	4.30	4.63	4.89	5.54	4.84	45.99	39.95	37.88	3.36	2.90	2.59
10 bad	5.21	3.23	4.34	4.83	5.37	4.13	44.19	38.11	35.39	3.26	2.31	1.18
N	501	495	556	199	305	193	490	495	556	294	202	355
Mean	9.69	9.65	9.20	6.23	6.68	6.39	62.07	55.19	61.48	7.81	7.60	8.06
p-value $	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
SED	0.29	0.28	0.27	0.47	0.49	0.42	4.30	4.24	4.00	1.99	1.77	1.72
σ_g^2	0.05	0.08	0.10	0.16	0.13	0.15	29.50	30.01	68.69	3.23	3.86	3.28
H^2	0.53	0.60	0.74	0.58	0.50	0.56	0.76	0.77	0.89	0.62	0.70	0.70

§ vase life (VL) in d; flower diameter (FD) in centimetre of standard carnations; stem length (SL) in centimetre; number of genotypes (N); coefficient of variation (CV); mean standard error of a difference (SED), genotypic variance (σ_g^2) and ad hoc heritability (H^2) (Piepho and Möhring, 2007).

$ F-test of the null hypothesis of no genotype effect.

5.3.3 Phenotypic correlation

Pearson's correlation coefficients of adjusted genotypic means, showing pairwise associations between traits of carnation genotypes in three experiments, were used for calculating the phenotypic correlations between traits presented in Figure 17. A highly significant (<.0001) correlation between VL and FD of 0.4 (E1) and 0.25 (E2) was detected. FD was only measured in the standard carnations and the CF was assessed only in the multi-flowered mini carnations. This is the reason that no results are available for this trait combination. In E3, an association between SL and FD of 0.21 was found. Between SL and CF of mini carnations, correlations of 0.24 and 0.31 were observed in E1 and E2.

Trait §	VL			FD			SL			CF		
Trial	E1	E2	E3	E1	E2	E3	E1	E2	E3	E1	E2	E3
VL	1.0	1.0	1.0	0.40 *<.0001*	0.25 *<.0001*	-0.04 *0.74*	0.04 *0.49*	-0.02 *0.73*	0.07 *0.10*	-0.04 *0.55*	0.03 *0.63*	0.003 *0.96*
FD				1.0	1.0	1.0	0.05 *0.49*	0.11 *0.06*	0.21 *0.0025*	n.a.	n.a.	n.a.
SL							1.0	1.0	1.0	0.09 *0.11*	0.24 *0.0006*	0.31 *<.0001*
CF				n.a.	n.a.	n.a.				1.0	1.0	1.0

Figure 18: Phenotypic correlation showing pairwise associations between adjusted trait means of carnation genotypes.

In the upper triangle are the phenotypic correlations with *p*-values in italic. Plots of the means of the trait combinations in the lower triangle.

§ vase life (VL) in d; flower diameter (FD) in centimetre of standard carnations; stem length (SL) in centimetre; count of flowers (CF) of mini carnations; not analysed (n.a.);

The traits VL and SL are influenced by the carnation type and were therefore analyzed separately (Table 21). FD was only analyzed in the standard carnations and the CF only in mini carnations. Both carnation types showed a correlation between VL and SL of 0.12 and 0.17 in E1 or E2, which was not found without separating the data into standard and mini carnations.

Table 21: Pearson correlation coefficients showing pairwise associations (VL and SL) between traits of genotypes split in mini carnations and standard carnations.

Trait	Mini carnation			Standard carnation		
combination §	E1	E2	E3	E1	E2	E3
VL + SL	0.04	0.05	0.12	0.17	-0.06	-0.02
	0.95	*0.48*	*0.03*	*0.02*	*0.29*	*0.77*

§ vase life (VL) in d; stem length (SL) in centimetre; *p*-values in italics.

5.3.4 Genetic correlation

The genetic correlation of the traits FD and VL was between 0.47 and 0.30 in E1 and E2 (Table 22). As shown in Figure 18 the phenotypic correlation of these traits in these experiments were 0.40 and 0.25. This means that the phenotypic correlation was nearly the same as the genetic correlation. The RSE reported in italics showed, however, that indirect selection would not be better than direct selection.

In E2 and E3, the traits SL and CF showed a phenotypic correlation of 0.24 and 0.31. A high genetic correlation of 0.17, 0.35 and 0.36 was found in all experiments. However, the ratio between expected and correlated response showed that indirect selection would have no benefit.

The phenotypic correlation of FD and SL of 0.21 in E3, could not be explained with genetic reasons, because the genetic correlation was low. The genetic correlations between the traits VL and CF or SL are very low and partly negative. In the case of the traits FD and SL, a small genetic correlation was found. Yet again, indirect selection would not be better than direct selection in all traits.

As shown in Table 20, the heritability of all traits was high. This indicates that the environment has low influence on the phenotypic correlation and the genetic correlation has a higher relevance (Falconer and Mackay, 1996).

Table 22: Genetic correlation calculated via bivariate analysis of the combination of all measured traits.

Trait [§]	VL			FD			SL			CF		
Trial	E1	E2	E3	E1	E2	E3	E1	E2	E3	E1	E2	E3
VL	1.0	1.0	1.0	0.47	0.30	0.02	-0.02	-0.04	0.10	-0.11	0.01	0.07
				0.49	*0.27*	*0.02*	*-0.02*	*-0.15*	*0.11*	*-0.12*	*0.01*	*0.07*
FD				1.0	1.0	1.0	0.09[$]	0.07	0.21[$]			
							0.14	*0.09*	*0.24*	n.a.	n.a.	n.a.
SL							1.0	1.0	1.0	0.17	0.30	0.36
										0.15	*0.29*	*0.32*
CF										1.0	1.0	1.0

[§] VL (vase life) in d; FD (flower diameter) in centimetre of standard carnations; SL (stem length) in centimetre; CF (count of flowers) of mini carnations; (n.a.) not available, relative selection efficiency (RSE) in italic, [$]results from SAS analysis, all other results obtained with ASREML.

5.3 Discussion

5.4.1 Importance of VL

The aim of breeding long-lasting carnations is becoming more and more important because of the current change of transportation of stems from shipment to airfreight. The transportation of carnation stems by ship requires a transportability of four weeks. The end-consumer expects a VL of more than one week, which needs to be added to the transportation time. The breeding goal VL is difficult because VL is hard to measure. Many different factors affect VL and it is not possible to measure all of them. The assessment of VL is labor-intensive. Each stem has to be grown, cultivated, harvested, maybe stored to simulate transportation, prepared and then put into water. Subsequently, plants must be scrutinized every day for senescence symptoms. In the frame of this work, we used up to eight stems per genotype. In practical breeding often fewer stems are used. The approach of indirect selection for VL is imaginable. Because of the high effort needed to assess VL, indirect selection is worthwhile.

5.4.2 Indirect selection for VL

Genetic correlation can be used for indirect selection, which means that the selection is based on another trait than the trait of interest. A genetic correlation of VL was only established with the trait FD (0.47 and 0.30). However, it was found that indirect selection would have no benefit for VL. Even so, indirect selection could be an interesting option due to the smaller effort to measure the auxiliary trait FD.

Genetic correlation is dependent on gene frequencies (Bohren et al., 1966), which could be a reason that E1, E2 and E3 showed different results. In each experiment another set of genotypes was phenotyped. A bigger flower leads to a longer VL. As opposed to this, Onozaki et al. (2018) reported a negative correlation between both traits. In cut gerbera a phenotypic correlation between VL and FD was found (Mehdikhah et al., 2016). Accurate selection is important for breeding success. For complex traits like VL, where the selection is not possible until the plant is fully developed, indirect selection is an interesting option. VL and FD are genetically correlated, so selection on FD will also result in a change in VL. FD is a very easily measurable trait. By using a ruler, the diameter can be

assessed on a metric scale and analysed easily. One disadvantage is that the FD can be measured only until the plant shows open flowers.

5.4.3 Stems of mini carnation with more flowers are longer

The genetic and phenotypic correlation between the traits CF and SL is relatively high. This indicates that genes, which are responsible for the stem growth, also influence the floweriness. It may be speculated that genotypes with longer stems with many flowers have bigger leaves as well and that this correlation is linked to genes influencing the biomass production. Swamy (2010) found phenotypic correlations of 0.85 between the number of flowers per m^2 and SL. A phenotypic correlation of 0.90 was found between SL and leaf area (cm^2) in carnations. In alfalfa (*Medicago sativa*) a high correlation between the SL and the number of inflorescences per stem was obtained (Pelikan et al., 2014).

5.4.4 Statistical analysis to calculate genetic correlations

A challenge was to formulate models for first-phase and second-phase traits, and to combine these models in a bivariate analysis. The solution was the inclusion of a dummy variable to switch-on or switch-off the second phase, depended on the trait. The traits VL and FD were measured in second phase, because the stems were harvested before bud opening. In the second phase, another structure of experimental design was used due to limited space in the lab.

Each year, we phenotyped about 500 genotypes. Some bivariate analyses could not be performed with SAS because it ran out of memory. Data volume was tried to reduce by using the mean of both stems of each genotype in each replicate, which were not influenced by the storage treatment (CF, SL and FD). However, the problem was associated with bivariate analysis of the second-phase trait FD. Both harvested stems were not placed on the same position in the second phase, because the allocation was random. The bivariate analysis of traits without a second phase (CF and SL) was done with SAS, without problems of insufficient memory and there was no need to reduce the data set. The bivariate analyses, which were not bogged down due to insufficient memory, needed more than one week on a personal computer with a 3.4-GHz processor and 32 GB of RAM. ASReml can fit complex data set due to an improved REML algorithm and use of

sparse matrix methods (Thompson et al., 2003). Comparing the results of both statistical packages, only small variations were obtained.

5.4.5 Quality of data sets

The quality of the used data sets is high due to the amount of tested individuals in different experiments (Boxriker et al., 2017b). In total, 849 single-headed standard carnation and 703 multi-headed, small-flowered mini carnation were phenotyped for the traits VL, FD, SL and CF. The smaller SED in the last experiment indicates a more precise phenotyping in all traits compared to the other experiments. The contrast between the best and worst genotypes were high in all traits. This indicates a good quality of the phenotyping. The heritability in all traits was quite high. The genetic variance is small in the case of VL and FD, maybe because these are traits where the breeding effort is high.

5.4.6 Influence of carnation type and storage treatment

As shown in Boxriker et al. (2017a) the carnation type has a high influence due to the morphological differences between these two types. Within the current manuscript, it was shown that the mini carnation stems have longer VL than the standard carnations. Standard carnations have longer stems, which could be explained by the architecture of the plants. Mini carnations are branched while the standard carnations have only one stem. By separating the carnation types into two different data sets, more relevant correlations were found.

Storage treatment decreased VL and FD. The VL shortening is due to the longer period from harvesting until the flower opening in vases. It is interesting that FD is influenced by the storage of two weeks, which is relevant due to the long transportation time as already mentioned. The selection of genotypes with big flowers, however, takes place in the greenhouse, without harvesting and storage treatment.

5.5 Conclusion

Bivariate mixed model analysis was performed the first time for a two-phase experiment. Interesting phenotypic and genotypic correlations of breeding relevant traits were found like VL and FD. Indirect selection is in our case not worthwhile, but a combination of indirect and direct selection could improve breeding of long-lasting cut carnation because the auxiliary trait provides an indication of the performance of the target trait. Moreover, time and effort can be saved when the correlated traits are easy to measure.

6. General discussion

This chapter discusses the key results of this thesis and compares the results with the relevant literature. One of the main results was that special attention should be paid to the second phase of the experiment in which VL traits are assessed. An experimental design in the second phase showed a higher impact on precision than the design in first phase, the greenhouse. Within this thesis, a high influence of the carnation type on the statistical analysis of VL data and molecular data was detected. The trait flower diameter was highlighted as a potential trait influencing VL indirectly.

6.1 VL trials are multi-phase experiments

The use of experimental designs and to analyze data with mixed models is common practice in agriculture. In the practical breeding of 'Klemm + Sohn' no statistical methods are used till now. All genotypes are replicated in multiple locations, but they are not replicated in each environment and no randomization is applied. Some experiments include a second phase caused by chemical, physical or molecular biological measurements in the laboratory, which leads to a second location with need to implement experimental designs (Brien et al., 2011; Smith et al., 2006). The term phase is defined as the duration in which the experimental units are engaged in producing their outcome (Brien et al., 2011). In the case of crop science, multi-phase experiments are common. Examples are a tobacco mosaic virus experiment (McIntyre, 1955), a four-phase experiment in apples (Wilkinson et al., 2015), rooting experiments in *Pelargonium* (Molenaar et al., 2017) or a cotton fiber test, where the cotton from the field phase was tested for strength in the laboratory phase (Brien and Bailey, 2005; Cox, 1958). Designs are needed for both phases to maximize efficiency. The research need of multi-phase design was first described by McIntyre (1955) and included a single randomization in both phases. Smith et al. (2006) described that often only a single field replicate has been taken to the next phase of the experiment and no randomization is applied anymore. The easiest way to design the second phase is to keep blocks intact across phases (Brien and Bailey, 2006). In some instances, the second phase of the experiment is more important especially when the target trait is measured in the second phase, as it is the case in VL (Chapter 3). First-phase traits like CF or SL do not involve a second phase because the traits are

measured before the stems were transferred to the second phase. Therefore, no influence of the second phase is present.

In this work, the rooting phase of vegetative cuttings was not implemented into the experimental design. In future, this should be added as the first phase of VL assessment experiments, followed by the cultivation of the rooted cuttings in the greenhouse. The third phase will be the VL assessment in the laboratory. A further aspect, which should be considered in experimental design, is the analysis in molecular laboratories, like RNA-extraction and the sequencing of the cDNA. Auer and Doerge (2010) showed that in RNA-Seq experiments, often significant lane effects are visible. These examples show that nearly all conducted experiments are multi-phase experiments and this should be considered in future.

6.2 How to conduct VL trials in a statistically efficient way

Within this thesis, more than 1.500 carnation genotypes were tested in three different experiments (E1, E2 and E3). In E1, other experimental designs were used in the first and second phase than in E2 and E3. This allowed a comparison of efficiency via response to selection as shown in Chapter 3. The result was that it is best to use row-column designs, because you can account for experimental errors. Positional effects can be compensated and when the second blocking factor turns out to be unimportant, the data can be analyzed as α-design. In ornamental breeding, it is not common to use experimental designs. Several studies describing the method of VL testing (Macnish et al., 2010; T. Onozaki et al., 2001; Satoh et al., 2005) are available but no randomization is applied neither in the greenhouse nor in the second phase. The reasons are often the limit of space in the greenhouse for replicates and in vegetative propagated crops, it is important to see the homogeneity which can be rated best in big plots. Within this work, the importance of experimental designs was highlighted and ‘Klemm + Sohn’ is now working on suitable designs for each crop.

Special attention must be given to the second phase. In the case of VL trials, the experimental design of the second phase was more important for precise data analysis as discussed in Chapter 3, which highlights that an efficient method should be used. Especially when the target trait is measured in the second phase, the experimental design is important. The best option is to apply the design of the

first phase one-to-one to the second phase (Smith et al., 2006), but good alternatives when this is not possible because block structures and sizes do not match, should be suggested.

That was the case in these experiments, because the space in the second phase was limited and the units of the experimental designs had another size. A second difficulty was that it was not known which genotypes and how many plants would arrive at the desired development stage at the same time. Therefore, it was not possible to plan the design before the start of the experiment. In E1, the harvested stems of one day and replicates of the first phase were retained, blocks were randomly assigned and the stems were located to free positions in the allocated blocks. In E2 and E3, another approach was used. For every harvested carnation stem two random numbers were generated, which define the location of the vases in the second phase. The approach of E1 was statistically more efficient as shown in Chapter 3. The reason could be that replicates of the first phase were transferred to the second phase. The replicates worked as an additional blocking structure, which improved the efficiency of the second phase compared to the complete randomization in the other experiments. The harvested stems of one day where grouped together in the second phase, meaning that the harvest day works as blocking factor (Smith et al., 2006), which can explain the more precise data analysis in E1. In the study of Smith et al. (2006) the samples per day were used as blocking factor and it was possible to use resolvable incomplete block design with days as blocks in the second phase of the experiment. In plant science, harvest days are available in most cases and are a good option to use it as blocking structure in the second phase. In the laboratory for example, the charge of probes could be used as blocking structure, because often the analysis of samples is limited due the space in the centrifuge or time to handle the method.

With hindsight, the second phase should have been planned in another way: for the laboratory phase, the first and second replicates of the conducted VL trials should have been used twice in the laboratory and the third and forth as well. Further, the 57 used blocks of each replicate could have been used two times in the laboratory, because there was space for 148 blocks. This indicates that the experimental design of the first phase would have been adjusted to the second phase. It would be interesting to see if this method can improve the precision of the second phase. Nevertheless, it could be shown that the second phase had a

higher impact on precision than the first phase because the trait VL was measured in this phase of the experiment.

6.3 Analysis of transcriptome data

Compared to other publicly available RNA-Seq data (Tanase et al., 2012; Wan et al., 2015) the data quality was evaluated as good, as discussed in Chapter 4. The MACE data could not be compared with the literature, because it was the first time that carnations were sequenced with this sequencing method. However, the data quality was declared as good because it was possible to analyze the data and reasonable results were generated as discussed in Chapter 4. The data analysis of the molecular data achieved via RNA-Seq and MACE was challenging due to its big amount. Another difficulty was the absence of a suitable reference genome where the data could be annotated. The available carnation reference genome (Yagi et al., 2014a) was not applicable and it was decided to use the *Beta vulgaris* genome (Dohm et al., 2014; Kubo et al., 2000; Mower and Palmer, 2006). Many sequenced transcripts are unknown proteins. However, due to the big amount of data, the use of GO (gene ontology) terms simplified the analysis. The annotation of genes to GO terms is an international bioinformatics initiative to standardize the vocabulary across all species (Berardini et al., 2004). However, it is hard to discover potential candidate genes for the trait of interest with this kind of approach.

Within the project, 49,212 SNPs were detected. The challenge was to select SNPs, which showed different point mutations between the bulk with long-lasting genotypes and the bulk with genotypes showing a short VL. To account for the differences in the bulks was challenging. Less literature is available about the selection of potential SNPs between bulks. Liu et al. (2012) described a selection method, but for the VL data set, it was not precise enough. A new approach was established to find different expressed SNPs between the bulks. Confidence intervals were used to find contrasts between the bulks associated with long and short VL, because it was possible to account for allele frequencies as well as for the sample size. The requirement for the selection of SNPs was that allele frequencies should be as different as possible. Pearson-Clopper intervals for binominal parameter (Reiczigel, 2003) were used and SNPs, where the confidence intervals includes 50 were not used, because otherwise the contrast between the

bulks would not be high enough. Furthermore, an index was implemented to calculate how far the boundaries of the confidence intervals are separated. SNPs for potential marker implementation were selected, when the confidence intervals were as far as possible separated from each other. For primer design, it was important to select reads with only one SNP per contig or with a distance of more than 25 bp between SNPs. However, the SNPs selection seemed to be successful, but it was not possible to design primers to generate genotypic data. As a solution, it was tried to use different sequencing methods, but it was not possible to see the contrast of phenotype also in the genotypic data of the selected loci. It was even not clear if the selected locations are involved in this kind of analyses.

The most important point, which was learned by the analysis of the transcriptome data, was that mini and standard carnations must be analyzed independently. More transcripts (around 10 %) are only expressed in the standard carnation bulk or the mini carnation bulk compared to the comparison of short and long VL (around 8 %). It could be proven that the selection of SNPs was not correct by using the bulk with both carnation types as discussed in Chapter 4. It should be noted, however, that mini carnations were developed from a standard carnation. By contrast, Nontaswatsri et al. (2002) found no difference between the carnation type by analyzing the shoot regeneration. DNA-fingerprints (Vainstein et al., 1991) and SSR-marker analysis (Kimura et al., 2009) between both carnation types showed less differences as well. However, within this thesis, differences could be obtained, and sixty years of intensive breeding and selection led to different gene pools with not only phenotypic differences.

Reasons for the differences can maybe be explained with the used methods. The experimental design in the greenhouse was planned for both carnation types. The carnation types were separated within the blocks (subsurface box), caused by a different treatment; the removal of flower buds. The carnation type was used as factor in mixed model analysis because it showed a significant effect. The analysis of VL showed that mini carnations had a longer VL compared to the standard carnations. This was shown in all three experiments. Additional to the different treatment of both carnation types, the mini carnations had another harvest and senescence decision because mini carnations have multiple flowers to rate and not like in the case of standard carnations only one flower. It was decided to use two buds or flowers to decide if the genotype is ready for harvest in the

greenhouse or lost ornamental value in VL testing. However, in the trait SL, which is not influenced by the harvest stage or evaluation of flowers, a difference was detected as well. Standard carnations showed longer stems compared to the mini carnations. By the molecular analysis it could be shown that traits were differentially expressed between the carnation types and within the selection of potential SNPs for the use as molecular markers, selected SNPs were differently expressed in either one of the carnation type (Chapter 4). This indicates that in such analyses the carnation types should be analyzed independently. With the aim to introduce molecular markers for cut carnations for each carnation type, a marker must be implemented.

Interestingly most of the differentially expressed genes or transcription factors, which are assumed to have an influence on VL, are differentially overexpressed in the bulks associated with short VL. Like for example ubiquitin, a signaling protein of degradation has a role during senescence (Belknap and Garbarino, 1996) and was overexpressed in the bulk of short VL. Furthermore, 1.800 transcripts were only expressed in the bulk with short VL and have the potential to influence the VL. Onozaki et al. (2018) discovered that breeding of long-lasting genotypes is more based on less production of ethylene and not on ethylene sensitivity. This indicates that in long-lasting genotypes less transcripts belonging to ethylene are available because no ethylene is produced and less receptors are available.

Nevertheless, it should be noted that for this kind of analysis not the right tissue was sampled. For RNA-Seq, the sampling was conducted in a suitable way. However, in the case of MACE the sampling was insufficient. Only samples from stem and buds were used. This indicates that it was the wrong tissue and too early in developmental stage of the plant to detect ethylene dependent gene expression. To analyze transcriptome data the decision of which kind of tissue and at which time point are crucial decisions. Auer and Doerge (2010) highlighted that relatively less attention is given to sampling like tissue sampling, RNA extractions or the sequencing method itself. However, it would be a better choice to use DNA-based methods because the results are easier to interpret and sampling is easier. Sampling of DNA is not dependent on tissue, time or temperature and other factors. In retrospect, the question is if BSA-Seq for MAS was the best choice of methods.

6.4 The choice of marker-enabled selection method

The main difference between MAS and GS is that in GS all available marker information to predict breeding values is used. This indicates that in GS no significant tests are necessary to find markers linked to QTLs with large effects (Meuwissen et al., 2001). As VL is a complex trait, which is not influenced by only some QTLs with relatively large effects, MAS was maybe not the best choice (Bernardo, 2008; Xu and Crouch, 2008). MAS is suggested for traits with low heritability (Meuwissen, 1998) but within this thesis, it was possible to achieve a high heritability for VL. Compared to Onozaki et al. (2018) the heritability we found was considerably higher. The basis of GS is a regression model relation phenotypic data to molecular marker information to predict breeding values. The problem so far is that no molecular markers have been developed for the trait VL. Carnations are highly heterozygous, which could be a difficulty in developing molecular markers. VL is a highly complex trait and is hard to measure. Therefore, molecular markers would have many benefits but make the implementation of it more difficult. The data quality of the conducted VL trials is good as shown in this thesis. The idea of BSA is to group the extreme phenotypes into bulks (Liu et al., 2012). It was decided to use twenty genotypes per bulk. Afterwards it was noticed that the carnation types must be separated. By keeping the bulks separated, only 12 mini carnations and 8 standard carnations were representing the bulks, which could be too less. However, in sunflowers it was possible to find molecular markers for resistance genes, even when only 16 genotypes were selected for each bulk (Livaja et al., 2013), but resistance is mainly based on some major genes (Vale et al., 2001). Within this project, it was tried to use other traits like FD, CF and SL, but no association with the genomic data was possible. This indicates that something is wrong with the molecular data, because these are less complex traits compared to VL and even then, no association was possible. In the follow-up project, more molecular data will be collected within other methods to get precise genotypic data. All conducted genotypic data was so far not precise enough. To learn from mistakes, in the follow up project, MACE sequencing was repeated with more genotypes and the sampling was changed to flower tissue and the sampling was conducted in the laboratory while VL assessment.

6.5 Importance of breeding long-lasting carnations

Flower trade has increasingly become a global business. Therefore, transportability is getting more and more important as a trait in cut flowers. The production in South America, Africa and Asia is increasing, because of better climate conditions, more available land and cheap labor. The demand on special days like Mother's Day and Valentine's Day coerce growers to adapt their cultivation. Therefore, breeding of long-lasting, storable and long-transportable cut flowers is getting more and more in the focus. Transportation of stems, especially carnations, is changed from air freight to ocean shipment. In addition to VL at home, the stems must survive the time of shipment, which is mainly two or four weeks. There are many advantages of shipping like the costs, capacity, climate control and less CO_2 emissions compared to airfreight. Ocean shipping has the lowest environmental impact for long distance transports (Psaraftis and Kontovas, 2009).

The found differences in VL of up to 15 d between the best and worst genotypes and the high heritability of VL in all experiments showed the potential for improvement of VL in cut carnation. Improved selection via molecular marker or auxiliary traits will increase VL of the population. Often it is tried to improve not only one main breeding goal but it is difficult not to reduce the performance of other traits at the same time (Falconer and Mackay, 1996). In carnation breeding, important traits despite VL are SL, FD, CF and productivity of carnation stems. The selection on multiple traits is possible, when the traits share the same genetic basis (Ghalambor et al., 2003; Griffing, 1967). Multiple-trait selection can be facilitated by using pedigree information, indirect selection via auxiliary traits or marker information (Dekkers, 2007a; Kadarmideen et al., 2003) . The implementation of a molecular marker for the trait VL and transportability will help to improve breeding effort in these complex traits.

The difference between stored (14 d at 2 °C) and not stored stems was between one and two days VL and a significant reduction of FD was detected. In other cut flowers like *Backhousia myrtifolia,* the VL is halved after storage treatment (one week at 2 °C) (Ekman et al., 2008). Carnations are more robust against a stress factor like transportation, but still there is a potential for improvement visible. Precise data analysis of VL can influence the planning of crossing combinations of

the breeder, which as well leads to an improvement of the population mean. Not only the visible documentation of the breeder is taken into account, adjusted means of repeated genotypes are more accurate. Kramer et al. (2016) summarized the missing or wrongly used statistics in several studies published in a horticultural journal. The carnation breeder at 'Klemm + Sohn' is currently not applying statistical data analysis or any experimental design neither in the greenhouse nor the laboratory where the VL assessment occurs. There is a potential to improve selection of long lasting genotypes via precise data analysis.

6.6 Indirect selection

Since no molecular markers for VL were found, correlated traits could be an alternative for indirect selection. To calculate genetic correlation, bivariate analysis was used and the first time applied to two-phase experiments with different experimental designs in both phases (Chapter 5). Bivariate analysis of two-phase traits and traits without a second phase was complicated. The reason was that the observation of first-phase trait does not have experimental units in the second phase. The use of a dummy variable to code the experimental design in the first and second phase was an efficient method to solve the problem as discussed in Chapter 5. Easier would be to use the same experimental design in each phase as suggested from Smith et al. (2006). Relevant genetic correlations were detected in VL and FD. The effort to assess for VL is compared to FD very high. The RES showed that indirect selection would have no higher response than direct selection. However, FD can provide an indication of the performance of the target trait because of the high genetic correlation. The additional information can be incorporated into breeder's decisions like for example the choice of crossing partners. Onozaki et al. (2018) found a negative phenotypic correlation of VL and FD. In this study, the genetic correlation was not calculated. Phenotypic correlations occur when the phenotype of two traits are associated explained by genetic and non-genetic causes. The most popular reasons for genetic correlation is pleiotropy or linkage disequilibrium (Jiang and Zeng, 1995). It is always better to calculate the genetic correlation as well, otherwise it is not obvious if the phenotypic correlation is influenced by the environment or by pleiotropy or linkage (Falconer and Mackay, 1996). All other measured traits showed only low genetic

correlation with the trait VL or both traits were relatively easy to measure, so direct selection will be the better option for both traits.

The big advantage of molecular markers is the early assessment stage based on young tissue. Already at germling stage or as cutting, analysis could be conducted (Collard et al., 2008). Measuring FD is possible when the first flowers are open. This means an earlier stage compared to VL, but much later compared to the approach that will be possible with a molecular marker. VL is a hard-to-measure trait. Further investigations of the implementation of molecular markers should be concentrated on the trait FD with the main goal to select for long-lasting carnations. In these experiments, FD was measured in the second phase, but this would not have been necessary. They were harvested before bud opening, because this is required to follow the protocol of VL assessment. For a better transportability, stems were harvested before bud opening. Open flowers will not survive transportation. The open flowers can be as well measured in the greenhouse to make the data collection faster and easier.

6.7 Further research

In a follow-up project, a genetic map should be generated. A genetic map is based on frequencies of gene recombination between markers during cross-over of homologous chromosomes (Singh and Singh, 2015). Yagi et al. (2013) generated a genetic map of carnations already and within this work, the map should be complemented. The advantage of a genetic map is that identified SNPs can be mapped helping to identify QTLs for traits like VL (Falconer and Mackay, 1996, Lynch and Walsh, 1998). The goal of the project is as well to expand the implemented methods to other traits like resistance to fusarium and to other crops like pot carnations. The challenge will be that the pot carnations are tetraploid and not diploid like the used cut carnations.

References

Adams, D.O., Yang, S.F., 1979. Ethylene biosynthesis: Identification of 1-aminocyclopropane-1-carboxylic acid as an intermediate in the conversion of methionine to ethylene. Proc. Natl. Acad. Sci. U. S. A. 76, 170–174.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control 19, 716–723.

Alexa, A., Rahnenführer, J., 2013. Gene set enrichment analysis with topGO., R package version 2.24.0.

Anders, S., Huber, W., 2010. Differential expression analysis for sequence count data. Genome Biol. 11, R106.

Arc, E., Sechet, J., Corbineau, F., Rajjou, L., Marion-Poll, A., 2013. ABA crosstalk with ethylene and nitric oxide in seed dormancy and germination. Frontiers in Plant Science. 4, 1-19.

Arens, P., Esselink, D., Noordijk, Y., Kodde, L., Hof, L., Wietsma, W., Vosman, B., 2009. Identification of carnation varieties using microsatellite markers. Acta Hortic. 836, 175–182.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., J.E., R., Ringwald, M., Rubin, G.M., Sherlock, G., 2000. Gene ontology: tool for the unification of biology. Nat. Genet. 25, 25–29.

Auer, P.L., Doerge, R.W., 2010. Statistical design and analysis of RNA sequencing data. Genetics 185, 405–416.

Bartoli, C.G., Simontacchi, M., Montaldi, E.R., Puntarulo, S., 1997. Oxidants and antioxidants during aging of chrysanthemum petals. Plant Sci. 129, 157–165.

Belknap, W.R., Garbarino, J.E., 1996. The role of ubiquitin in plant senescence and stress responses. Trends Plant Sci. 1, 789–793.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. 57, 289–300.

Berardini, T.Z., Mundodi, S., Reiser, L., Huala, E., Garcia-hernandez, M., Zhang, P., Mueller, L.A., Yoon, J., Doyle, A., Lander, G., Moseyko, N., Yoo, D., Xu, I., Zoeckler, B., Montoya, M., Miller, N., Weems, D., Rhee, S.Y., 2004. Functional annotation of the *Arabidopsis* genome using controlled vocabularies. Plant Physiol. 135, 745–755.

Bernardo, R., 2008. Molecular markers and selection for complex traits in plants: learning from the last 20 years. Crop Sci. 48, 1649–1664.

Bohren, B., Hill, W., Robertson, A., 1966. Some observations on asymmetrical correlated responses to selection. Genet. Res. 7, 44–57.

Bos, I., Caligari, P., 1995. Selection methods in plant breeding., Second. ed. Chapman & Hall, London.

Boxriker, M., Boehm, R., Krezdorn, N., Rotter, B., Piepho, H.P., 2017a. Comparative transcriptome analysis of vase life and carnation type in *Dianthus caryophyllus L.* Sci. Hortic. 217, 61–72.

Boxriker, M., Boehm, R., Möhring, J., Piepho, H.P., 2017b. Efficient statistical design in two-phase experiments on vase life in carnations (*Dianthus caryophyllus L.*). Postharvest Biol. Technol. 128, 161–168.

Brien, C.J., 2018. Multiphase experiments in practice: A look back. Aust. N. Z. J. Stat. 59, 327–352.

Brien, C.J., Bailey, R.A., 2006. Multiple randomizations. J. R. Stat. Soc. 68, 571–609.

Brien, C.J., Harch, B.D., Correll, R.L., Bailey, R.A., 2011. Multiphase experiments with at least one later laboratory phase. I. Orthogonal designs. J. Agric. Biol. Environ. Stat. 16, 422–450.

Casler, M.D., 2015. Fundamentals of experimental design: Guidelines for designing successful experiments. Agron. J. 107, 692–705.

Chao, S., Dubcovsky, J., Dvorak, J., Luo, M.-C., Baenziger, S.P., Matnyazov, R., Clark, D.R., Talbert, L.E., Anderson, J. a, Dreisigacker, S., Glover, K., Chen, J., Campbell, K., Bruckner, P.L., Rudd, J.C., Haley, S., Carver, B.F., Perry, S., Sorrells, M.E., Akhunov, E.D., 2010. Population- and genome-specific patterns of linkage disequilibrium and SNP variation in spring and winter

wheat (*Triticum aestivum L.*). BMC Genomics 11, 727.

Chen, Y., Mccarthy, D., Robinson, M., Smyth, G.K., 2015. edgeR: differential expression analysis of digital gene expression data. User's Guid.

Cochran, W.G., Cox, G.M., 1957. Experimental designs. John Wiley & Sons, New York.

Collard, B.C.Y., Mackill, D.J., Collard, B.C.Y., Mackill, D.J., 2008. Marker-assited selection: an approach for precision plant breeding in the twenty-first century. Phil. Trans. R. Soc. B 363, 557–572.

Cox, D.R., 1958. Planning of experiments., First. ed. Wiley, New York.

Dar, R.A., Tahir, I., Ahmad, S.S., 2014. Sugars and sugar alcohols have their say in the regulation of flower senescence in *Dianthus chinensis L.* Sci. Hortic. 174, 24–28.

Debener, T., Linde, M., 2009. Exploring complex ornamental genomes: the rose as a model plant. CRC. Crit. Rev. Plant Sci. 28, 267–280.

Dekkers, J.C.M., 2007. Prediction of response to marker assisted and genomic selection using selection index theory. J. Anim. Breed. Genet. 124, 331–341.

Dohm, J.C., Minoche, A.E., Holtgräwe, D., Capella-Gutierrez, S., Zakrzewski, F., Tafer, H., Rupp, O., Sörensen, T.R., Stracke, R., Reinhardt, Richard Goesmann, A., Kraft, T., Schulz, B., Stadler, P.F., Schmidt, T., Gabaldo, T., Lehrach, H., Weisshaar, B., Himmelbauer, H., 2014. The genome of the recently domesticated crop plant sugar beet (*Beta vulgaris*). Nature 505, 546–562.

Dole, J.M., Wilkins, H.F., 1999. Floriculture, principles and species., Second. ed. Prentice Hall Inc., New Jersey.

Ebrahimzadeh, A., Jimenez-Becker, S., Chaparro-Torres, L.A., Fernandez-Trujillo, J.P., Serrano-Mula, M., Lao-Arenas, M.T., 2013. Postharvest petal senescence of two cultivars of carnation flowers with different vase lives. Aust. J. Crop Sci. 7, 1829–1835.

Ebrahimzadeh, A., Jiménez, S., Teixeira da Silva, J.A., Satoh, S., Lao, M.T., 2008. Post-harvest physiology of cut carnation flowers. Fresh Prod. J. 2, 56–71.

Ekman, J., Eyre, J., Joyce, D., 2008. Flowers by sea. Rural Industries Research and Development Corporation, Kingston.

Falconer, D.S., Mackay, T.F.C., 1996. Introduction to quantitative genetics, Forth. ed. Longman, Harlow.

Fletcher, R.A., Osborne, D.J., 1965. Regulation of protein and nucleic acid synthesis by gibberellin during leaf senescence. Nature 207, 864–865.

Fu, Z., Wang, H., Liu, J., Zeng, H., Zhang, J., Kuang, X., Yu, Y., 2011. Molecular cloning and characterization of carnation EBF1 gene during flower senescence and upon ethylene exposure and sugar. Agric. Sci. China 10, 1872–1880.

Ganal, M.W., Durstewitz, G., Polley, A., Berard, A., Buckler, E.S., Charcosset, A., Clarke, J.D., Graner, E., Hansen, M., Joets, J., Le Paslier, M.-C., Mcmullen, M.D., Montalent, P., Rose, M., Schön, C.-C., Sun, Q., Walter, H., Martin, O.C., Falque, M., 2011. A large maize (*Zea mays L.*) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. PLoS One 6, 1–15.

Ghalambor, C.K., Walker, J.A., Reznick, D.N., 2003. Multi-trait selection, adaptation, and constraints on the evolution of burst swimming performance. Soc. Integr. Comp. Biol. 43, 431–438.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, Raktima Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2013. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat. Biotechnol. 29, 644–652.

Griffing, B., 1967. Selection in reference to biological groups. Aust. J. Biol. Sci. 20, 127–139.

Häffner, E., Konietzki, S., Diederichsen, E., 2015. Keeping control: The role of senescence and development in plant pathogenesis and defense. Plants 4, 449–488.

Ho, L.C., Nichols, R., 1977. Translocation of ^{14}C-sucrose in relation to changes in carbohydrate content in rose corollas cut at different stages of development. Ann. Bot. 41, 227–242.

Holland, J.B., Nyquist, W.E., T, C.-M.C., 2003. Estimating and interpreting heritability for plant breeding: an update. Plant Breed. Rev. 22, 9–112.

In, B.C., Inamoto, K., Doi, M., 2009. A neural network technique to develop a vase life prediction model of cut roses. Postharvest Biol. Technol. 52, 273–278.

Itzhaki, H., Maxson, J.M., Woodson, W.R., 1994. An ethylene-responsive enhancer element is involved in the senescence-related expression of the carnation glutathione-S-transferase (GST1) gene. Proc. Natl. Acad. Sci. U. S. A. 91, 8925–8929.

Jiang, C., Zeng, Z., 1995. Multiple trait analysis of genetic mapping for quantitaive trait loci. Genetics 140, 1111–1127.

Jibran, R., Hunter, D.A., Dijkwel, P.P., 2013. Hormonal regulation of leaf senescence through integration of developmental and stress signals. Plant Mol. Biol. 82, 547–561.

John, J.A., Eccleston, J.A., 1986. Row-column alpha-designs. Biometrika 73, 301–306.

John, J.A., Williams, E.R., 1995. Cyclic and computer generated designs., Second. ed. Chapman & Hall, London.

Jones, M., Woodward, R., Stoller, J., 2015. Increasing precision in agronomic field trials using latin square designs. Agron. J. 107, 20–24.

Kadarmideen, H.N., Thompson, R., Coffey, M.P., Kossaibati, M.A., 2003. Genetic parameters and evaluations from single- and multiple-trait analysis of dairy cow fertility and milk production. Livest. Prod. Sci. 81 81, 183–195.

Kimura, T., Yagi, M., Nishitani, C., Onozaki, T., Ban, Y., Yamamoto, T., 2009. Development of SSR markers in carnation (*Dianthus caryophyllus*). J. Japan. Soc. Hort. Sci. 78, 115–123.

Köhlein, F., 1990. Nelken, First. ed. Eugen Ulmer GmbH & Co., Stuttgart.

Kramer, M.H., Paparozzi, E.T., Stroup, W.W., 2016. Statistics in a Horticultural Journal: Problems and Solutions. J. Am. Soc. Hortic. Sci. 141, 400–406.

Kubo, T., Nishizawa, S., Sugawara, A., Itchoda, N., Estiati, A., Mikami, T., 2000. The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris L.*) reveals a novel gene for tRNA Cys (GCA). Nucleic Acids Res. 28, 2571–2576.

Lande, R., Thompson, R., 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124, 743–756.

Licausi, F., Ohme-Takagi, M., Perata, P., 2013. APETALA2/ethylene responsive factor (AP2/ERF) transcription factors: mediators of stress responses and developmental programs. New Phytol. 639–649.

Lim, P.O., Kim, H.J., Nam, H.G., 2007. Leaf Senescence. Annu. Rev. Plant Biol. 58, 115–136.

Liu, C., He, Y., Gou, T., Li, X., Ning, G., Bao, M., 2016. Identification of molecular markers associated with the double flower trait in Petunia hybrida. Sci. Hortic. 206, 43–50.

Liu, S., Yeh, C.T., Tang, H.M., Nettleton, D., Schnable, P.S., 2012. Gene mapping via bulked segregant RNA-Seq (BSR-Seq). PLoS One 7, 1–8.

Livaja, M., Wang, Y., Wieckhorst, S., Haseneyer, G., Seidel, M., Hahn, V., Knapp, S.J., Taudien, S., Schön, C.-C., Bauer, E., 2013. BSTA: a targeted approach combines bulked segregant analysis with next- generation sequencing and de novo transcriptome assembly for SNP discovery in sunflower. BMC Genomics 14, 628.

Lynch, M., Walsh, B., 1998. Genetics and analysis of quantitative traits., First. ed. Sinauer, Sunderland.

Macnish, A.J., Leonard, R.T., Borda, A.M., Nell, T.A., 2010. Genotypic variation in the postharvest performance and ethylene sensitivity of cut rose flowers. HortScience 45, 790–796.

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10–12.

Mc George, P., Hammett, K., 2012. Carnations and Pinks., First. ed. Firefly Books Ltd)., Ontario.

McIntyre, G.A., 1955. Design and analysis of two phase experiments. Biometrics 11, 324–334.

Mehdikhah, M., Onsinejad, R., Hashemabadi, D., 2016. Postharvest life of cut gerbera (*Gerbera jamesonii*) flowers as affected by salicylic acid, citric acid and ascorbic acid. J. Agric. Biol. Sci. 11, 170–174.

Meuwissen, T., 1998. Genomic selection: the future of marker assited selection and animal breeding., in: Marker-assisted selection: A fast track to increase genetic gain in plant and animal breeding? pp. 54–59.

Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. Genet. Soc. Am. 157, 1819–1829.

Michelmore, R.W., Paran, I., Kesseli, R. V, 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. U. S. A. 88, 9828–32.

Molenaar, H., Glawe, M., Boehm, R., Piepho, H.P., 2017. Selection for production-related traits in *Pelargonium zonale*: improved design and analysis make all the difference. Hortic. Res. 4, 1–9.

Mower, J.P., Palmer, J.D., 2006. Patterns of partial RNA editing in mitochondrial genes of *Beta vulgaris*. Mol Gen Genomics 278, 285–293.

Mueller-Roeber, B., Balazadeh, S., 2014. Auxin and its role in plant senescence. J. Plant Growth Regul. 33, 21–33.

Müller, R., Stummann, B.M., Andersen, A.S., Serek, M., 1999. Involvement of ABA in postharvest life of miniature potted roses. Plant Growth Regul. 29, 143–150.

Müller, S., Rycak, L., Afonso-Grunz, F., Winter, P., Zawada, A.M., Damrath, E., Scheider, J., Schmäh, J., Koch, I., Kahl, G., Rotter, B., 2014. APADB: A database for alternative polyadenylation and microRNA regulation events. Database 2014, 1–11.

Nambara, E., Marion-Poll, A., 2005. Abscisic acid biosynthesis and catabolism. Annu. Rev. Plant Biol. 56, 165–85.

Newton, R., Cook, W.H., 1927. The effect of ethylene on the respiration of bananas during ripening. Plant Physiol. 2, 357–360.

Nimura, M., Kato, J., Mii, M., Ohishi, K., 2008. Cross-compatibility and the polyploidy of progenies in reciprocal backcrosses between diploid carnation (*Dianthus caryophyllus L.*) and its amphidiploid with *Dianthus japonicus Thunb*. Sci. Hortic. 115, 183–189.

Nontaswatsri, C., Fukai, S., Touma, T., Goi, M., 2002. Comparison of adventitious shoot formation from node and leaf explants of various carnation (*Dianthus caryophyllus L.*) cultivars. J. Hortic. Sci. Biotechnol. 77, 520–525.

Ofori, I., 1996. Correlation and path-coefficient analysis of components of seed yield in bambara groundnut (*Vigna subterranea*). Euphytica 91, 103–107.

Onozaki, T., Ikeda, H., Yamaguchi, T., 2001. Genetic improvement of vase life of carnation flowers by crossing and selection. Sci. Hortic. 87, 107–120.

Onozaki, T., Yamada, M., Yagi, M., Tanase, K., Shibata, M., 2018. Effects of crossing and selection for seven generations based on flower vase life in carnations (*Dianthus caryophyllus L.*), and the relationship between ethylene production and flower vase life in the breeding lines. Japanese Soc. Hortic. Sci. 87, 106–114.

Pelikan, J., Vymyslicky, T., Knotova, D., Raab, S., 2014. Quantitative traits breeding for multifunctional grasslands and turf., First. ed. Springer, New York.

Piepho, H.P., 2009. Data transformation in statistical analysis of field trials with changing treatment variance. Agron. J. 101, 865–869.

Piepho, H.P., Büchse, A., Emrich, K., 2003. A hitchhiker's guide to mixed models for randomized experiments. J. Agron. Crop Sci. 189, 310–322.

Piepho, H.P., Möhring, J., 2007. Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177, 1881–8.

Piepho, H.P., Möhring, J., 2011. On estimation of genotypic correlations and their standard errors by multivariate REML using the MIXED procedure of the SAS System. Crop Sci. 51, 2449–2454.

Piepho, H.P., Möhring, J., Williams, E.R., 2013. Why randomize agricultural experiments? J. Agron. Crop Sci. 199, 374–383.

Piepho, H.P., Williams, E.R., Fleck, M., 2006. A note on the analysis of designed experiments with complex treatment structure. HortScience 41, 446–452.

Piepho, H.P., Williams, E.R., Michel, V., 2015. Beyond latin squares: A brief tour of row-column designs. Agron. J. 107, 2263–2270.

Psaraftis, H.N., Kontovas, C.A., 2009. CO_2 emission statistics for the world commercial fleet. WMU J. Marit. Aff. 8, 1–25.

Rani, P., Singh, N., 2014. Senescence and postharvest studies of cut flowers: A critical review. Pertanika J. Trop. Agric. Sci. 37, 159–201.

Reiczigel, J., 2003. Confidence intervals for the binomial parameter: some new considerations: Stat. Med. 22, 611–621.

Roth, A., Ding, J., Morin, R., Crisan, A., Ha, G., Giuliany, R., Bashashati, A., Hirst, M., Turashvili, G., Oloumi, A., Marra, M.A., Aparicio, S., Shah, S.P., 2012. JointSNVMix: A probabilistic model for accurate detection of somatic mutations in normal/tumour paired next-generation sequencing data. Bioinformatics 28, 907–913.

Satoh, S., 2011. Ethylene production and petal wilting during senescence of cut carnation (*Dianthus caryophyllus*) flowers and prolonging their vase life by genetic transformation. J. Japanese Soc. Hortic. Sci. 80, 127–135.

Satoh, S., Nukui, H., Inokuma, T., 2005. A method for determining the vase life of cut spray carnation flowers. J. Appl. Hortic. 7, 8–10.

Schulz, D.F., Schott, R.T., Voorrips, R.E., Smulders, M.J.M., Linde, M., Debener, T., 2016. Genome-wide association analysis of the anthocyanin and carotenoid contents of rose petals. Front. Plant Sci. 7, 1–15.

Searle, S.R., Casella, G., McCulloch, C.E., 1992. Variance components., First. ed. John Wiley & Sons, Inc., New Jersey.

Shanan, T.N., Emara, K.S., Barakat, S.O., 2010. Prolonging vase life of carnation flowers using natural essential oils and its impact on microbial profile of vase solutions. Aust. J. Basic Appl. Sci. 4, 3559–3574.

Sharp, R.E., 2002. Interaction with ethylene: changing views on the role of abscisic acid in root and shoot growth responses to water stress. Plant Cell Environ. 25, 211–222.

Shimelis, H., Hugo, A., 2011. Determination of selection criteria for seed yield and seed oil content in Vernonia (*Vernonia galamensis* variety <i>ethiopica). Ind. Crops Prod. 33, 436–439.

Singh, B.D., Singh, A.K., 2015. Marker-assited plant breeding: principles and practices., First. ed. Springer, New York.

Smith, A.B., Lim, P., Cullis, B.R., 2006. The design and analysis of multi-phase plant breeding experiments. J. Agric. Sci. 144, 393.

Sood, S., Vyas, D., Nagar, P.K., 2006. Physiological and biochemical studies during flower development in two rose species. Sci. Hortic. 108, 390–396.

Speed, T.P., Williams, E.R., Patterson, H.D., 1985. A note on the analysis of resolvable block designs. J. R. Stat. Soc. 47, 357–361.

Swamy, G., 2010. Studies on integrated nutrient management in carnation (*Dianthus caryophyllus L.*) Cv. dona under polyhouse. Univ. Agric. Sci. Bangalore.

Tanase, K., Nishitani, C., Hirakawa, H., Isobe, S., Tabata, S., Ohmiya, A., Onozaki, T., 2012. Transcriptome analysis of carnation (*Dianthus caryophyllus L.*) based on next-generation sequencing technology. BMC Genomics 13, 292.

Tanksley, S.D., 1983. Molecular markers in plant breeding. Plant Mol. Biol. Report. 1, 3–8.

Tanksley, S.D., Rick, C.M., 1980. Isozymic gene linkage map of the tomato: Applications in genetics and breeding. Theor. Appl. Genet. 58, 161–170.

Teitel, M., Atias, M., Barak, M., 2010. Gradients of temperature, humidity and CO_2 along a fan-ventilated greenhouse. Biosyst. Eng. 106, 166–174.

Thompson, R., Cullis, B., Smith, A., Gilmour, A., 2003. A sparse implementation of the average information algorithm for factor analytic and reduced rank variance models. Aust. N. Z. J. Stat. 45, 445–459.

Twumasi, P., Ieperen, W. van, Woltering, E., Emons, A.M., Schel, J., Snel, J., Meeteren, U. van, Marwijk, D. van, 2005. Effects of water stress during growth on xylem anatomy, xylem functioning and vase-life in three *Zinnia elegans* cultivars. Acta Hort 669, 6709.

Ueda, J., Kato, J., Yamane, H., Takahashi, N., 1981. Inhibitory effect of methyl jasmonate and its related compounds on kinetin-induced retardation of oat leaf senescence. Physiol. Plant. 52, 305–309.

Vainstein, A., Hiuef, J., Lavi, U., Tzuri, G., 1991. Assessment of genetic relatedness in carnation by DNA fingerprint analysis. Euphytica 56, 225–229.

Vale, X.F.R. Do, Parlevliet, J.E., Zambolim, L., 2001. Concepts in plant disease resistance. Fitopatol. bras. 26, 577–589.

van Doorn, W.G., 1996. Water relations of cut flowers., First. ed. John Wiley & Sons, Inc., Oxford.

van Doorn, W.G., 2004. Is petal senescence due to sugar starvation? Plant Physiol. 134, 35–42.

van Doorn, W.G., Woltering, E.J., 2008. Physiology and molecular biology of petal senescence. J. Exp. Bot. 59, 453–480.

van Meeteren, U., 1992. Role of air embolism and low water temperature in water balance of cut chrysanthemum flowers. Sci. Hortic. 51, 275–284.

Wan, X.L., Zhou, Q., Wang, Y.Y., Wang, W.E., Bao, M.Z., Zhang, J.W., 2015. Identification of heat-responsive genes in carnation (*Dianthus caryophyllus L.*) by RNA-seq. Front. Plant Sci. 6, 1–13.

Wang, L., Feng, Z., Wang, X., Wang, X., Zhang, X., 2009a. DEGseq: An R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics 26, 136–138.

Wang, Z., Gerstein, M., Snyder, M., 2009b. RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. 10, 57–63.

Wilkinson, R.I., Frisina, C., Partington, D.L., Franz, P.R., Brien, C.J., Thomson,

F., Tomkins, R.B., Faragher, J.D., Frisina, C., Partington, D.L., Franz, P.R., Brien, C.J., 2015. Effects of 1-methylcyclopropropene on firmness and flesh browning in Pink Lady apples. J. Hortic. Sci. Biotechnol. 83, 165–170.

Willis, R., McGlasson, B., Graham, D., Joyce, D., 2007. Postharvest., Fifth. ed. CABI, Oxfordshire.

Wolfinger, R.D., 1996. Heterogeneous variance: covariance structures for repeated measures. J. Agric. Biol. Environ. Stat. 1, 205–230.

Woo, H.R., Kim, H.J., Nam, H.G., Lim, P.O., 2013. Plant leaf senescence and death – regulation by multiple layers of control and implications for aging in general. J. Cell Sci. 126, 4823–4833.

Wood, J.T., Williams, E.R., Speed, T.P., 1988. Non-orthogonal block structure in two-phase designs. Aust. J. Stat. 30, 225–237.

Xiong, L., Zhu, J., 2003. Regulation of abscisic acid biosynthesis. Plant Physiol. 133, 29–36.

Xu, Y., Crouch, J.H., 2008. Marker-assisted selection in plant breeding: from publications to practice. Crop Sci. 48, 391–407.

Yagi, M., Kosugi, S., Hirakawa, H., Ohmiya, A., Tanase, K., Harada, T., Kishimoto, K., Nakayama, M., Ichimura, K., Onozaki, T., Yamaguchi, H., Sasaki, N., Miyahara, T., Nishizaki, Y., Ozeki, Y., Nakamura, N., Suzuki, T., Tanaka, Y., Sato, S., Shirasawa, K., Isobe, S., Miyamura, Y., Watanabe, A., Nakayama, S., Kishida, Y., Kohara, M., Tabata, S., 2014a. Sequence analysis of the genome of carnation (*Dianthus caryophyllus L.*). DNA Res. 21, 231–241.

Yagi, M., Yamamoto, T., Isobe, S., Hirakawa, H., Tabata, S., Tanase, K., Yamaguchi, H., Onozaki, T., 2013. Construction of a reference genetic linkage map for carnation (*Dianthus caryophyllus L.*). BMC Genomics 14, 734.

Yagi, M., Yamamoto, T., Isobe, S., Tabata, S., Hirakawa, H., Yamaguchi, H., Tanase, K., Onozaki, T., 2014b. Identification of tightly linked SSR markers for flower type in carnation (*Dianthus caryophyllus L.*). Euphytica 198, 175–183.

Zhu, X., Chen, J., Xie, Z., Gao, J., Ren, G., Gao, S., Zhou, X., Kuai, B., 2015. Jasmonic acid promotes degreening via MYC2/3/4- and ANAC019/055/072-mediated regulation of major chlorophyll catabolic genes. Plant J. 84, 597–610.

7. Summary

Carnation (*Dianthus caryophyllus* L.) is one of the ten most famous cut flowers worldwide. A single big flower characterizes standard carnations, while mini carnations possess multiple flowers per stem. Vase life (VL) is one of the most important breeding objectives in carnations due to the need of long transportation times and direct influence on the costumers. But VL is a complex trait with several effects influencing it. Two-phase traits like VL are traits where the assessment is done in a second phase, in the laboratory and the plants are cultivated in the greenhouse, the first phase. Many experiments have a two-phase character, but little research has been conducted to develop experimental designs in the second phase. To improve breeding efficiency, molecular markers and genomic selection is used in agriculture science but it is so far not common in ornamental breeding. The goal of this thesis was the implementation of SNP-based molecular markers for the trait VL to improve selection of long-lasting, transportable cut carnations.

For marker association, 1,500 carnation genotypes were screened for VL behavior in an experimental design in both phases. Response to selection was used to assess efficiency. The second-phase experimental design was more important for precise data analyses. This highlights the research need on this topic. Furthermore, it was possible to suggest row-column designs for VL trials. Row-column designs are more flexible in the case of positional effects compared with one-dimensional blocking and can be easily analyzed like an α-design. The easiest way to design the following phases are to apply the design one-to-one.

The carnation types, mini and standard, showed an influence on VL. The mini carnations last 0.5 d longer than the standard carnations. The same conclusion was drawn based on the molecular data. Transcriptome data was generated with two different sequencing methods. By independent analysis of both carnation types, different results than via the analysis of the whole data set were found. This indicates that the analysis of carnations should be done separately for each carnation type. Association of the phenotypic and genotypic data was so far not possible. As an alternative to molecular markers, genetic correlations for the use as indirect selection for the trait VL and others for breeding relevant traits was calculated. For the first time, bivariate analysis was conducted in two-phase experiments. The genotypic correlation between VL and FD was high, but indirect

selection would be less effective than direct selection. However, the information can provide an indication of the performance and the effort to measure FD is small. The calculated high heritability of VL and found differences in VL of up to 15 d between the best and worst genotypes showed the potential of improving the population mean by using improved selection strategies like marker-assisted selection or auxiliary traits and the use of statistical methods like experimental designs in all phases of the experiment.

The influence of carnation type was shown with this thesis and indicates that the implementation of molecular markers must be done independently for each carnation type. The importance of experimental designs in multi-phase experiments was highlighted and statistical analysis by mixed models and a bivariate analysis of different traits was performed. Until now, no molecular marker for VL was identified but in a further research project, this will be solved by generating more genotypic data and the construction of a genetic map.

8. Zusammenfassung

Nelken (*Dianthus caryophyllus L.)* gehören zu den zehn wichtigsten Schnittpflanzen weltweit. Standardnelken sind durch eine große Blüte charakterisiert, während Mininelken mehrere kleine Blüten je Stiel haben. Eins der wichtigsten Zuchtziele ist die Vasenhaltbarkeit (VL), da Nelken weite Transportwege haben und die Haltbarkeit die Wiederkaufbereitschaft direkt beeinflusst. Allerdings ist die VL ein komplexes Merkmal, das durch viele Bedingungen beeinflusst werden kann. Zweiphasige Merkmale wie die VL zeichnen sich dadurch aus, dass sie in einer zweiten Phase des Experiments gemessen werden, während die Pflanzen in der ersten Phase im Gewächshaus kultiviert werden. Viele Versuche sind zweiphasig, jedoch wird wenig Forschung betrieben, um effiziente experimentelle Designs für die zweite Phase zu entwerfen. Im Bereich der Agrarwissenschaften werden molekulare Marker und die genomische Selektion zur Erhöhung der Züchtungseffizienz verwendet. In der Zierpflanzenzüchtung findet dies jedoch noch kaum Anwendung. Das Ziel der vorliegenden Arbeit war die Etablierung SNP-basierter, molekularer Marker für das Merkmal VL, um die Selektionseffizienz von lang haltbaren und gut transportierbaren Schnittnelken zu verbessern.

Zur Generierung einer Datengrundlage zur Markerassoziation wurden insgesamt 1.500 Nelkengenotypen auf Vasenhaltbarkeit phänotypisiert, wobei das experimentelle Design sowohl in der ersten als auch der zweiten Phase angelegt wurde. Der Selektionserfolg wurde verwendet, um die Effizienz bewerten zu können. Es konnte gezeigt werden, dass das Design in der zweiten Phase einen höheren Einfluss auf die Effizienz hat, wodurch der Forschungsbedarf in diesem Themengebiet bestätigt wurde. Zudem konnte gezeigt werden, dass Zeilen-Spalten Pläne für VL-Versuche am besten geeignet sind. Ein Zeilen-Spalten Plan ist verglichen mit einem eindimensional geblocktem α-Design flexibler gestaltbar im Fall von Positionseffekten und kann auch einfach zu einem effizientem α-Design umgewandelt werden. Der einfachste Weg die zweite Phase zu gestalten ist die Übertragung des Designs aus der ersten Phase.

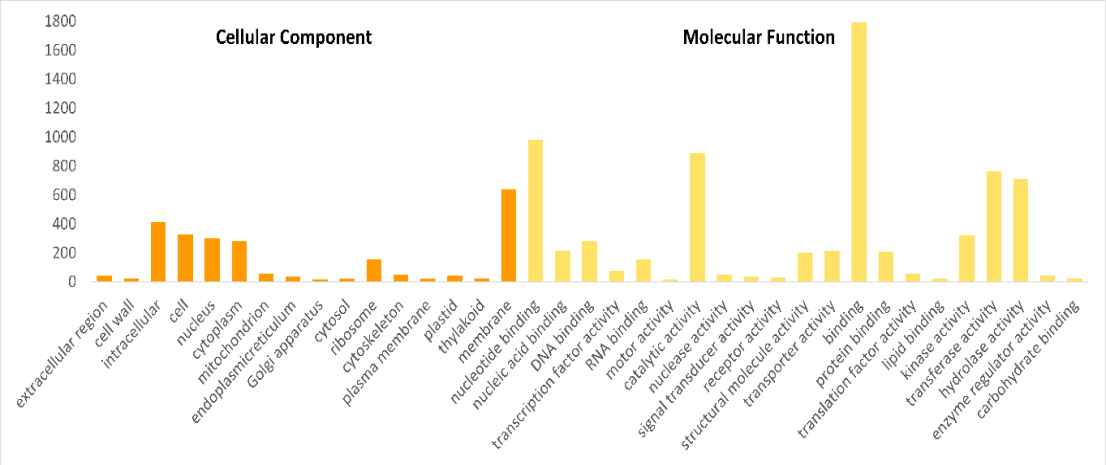
Der Nelkentyp, Mini- oder Standardnelken, zeigte einen Einfluss auf die VL: Die Mininelken waren bis zu 0,5 d länger haltbar als die Standardnelken. Bei der Analyse der molekularen Daten konnten ebenfalls Unterschiede zwischen den Nelkentypen festgestellt werden. Transkriptomdaten wurden mit zwei

verschiedenen Methoden erzeugt. Bei einer getrennten Analyse beider Nelkentypen wurden andere Ergebnisse erreicht als bei der kombinierten Analyse des gesamten Datensatzes. Daraus lässt sich folgern, dass die Analysen für jeden Nelkentyp einzeln durchgeführt werden muss. Assoziationen zwischen den phänotypischen und genotypischen Daten konnten bisher nicht gefunden werden. Als Alternative zu molekularen Markern wurden genetische Korrelationen zur Nutzung als indirekte Selektion des Merkmals VL und anderen für die Züchtung relevanten Merkmalen berechnet. Zum ersten Mal wurde eine bivariate Analyse in einem zweiphasigen Experiment angewendet. Die genetische Korrelation zwischen der VL und dem Blütendurchmesser ist relativ hoch, allerdings konnte gezeigt werden, dass eine indirekte Selektion weniger effizient wäre. Die genetische Korrelation kann jedoch als Vorinformation dienen, da der Blütendurchmesser deutlich einfacher gemessen werden kann im Vergleich zur VL. Die errechnete hohe Heritabilität und die Differenz zwischen dem besten und dem schlechtesten Genotypen von einer VL bis zu 15 d macht deutlich, dass das Populationsmittel erhöht werden kann durch verbesserte Selektionsstrategien wie die markergestützte Selektion, durch korrelierte Merkmale und die Verwendung von statistischen Methoden wie experimentellen Designs in allen Phasen des Experiments.

In dieser Arbeit konnte gezeigt werden, dass der Nelkentyp einen großen Einfluss auf die Daten hat und die Entwicklung molekularer Marker für jeden Nelkentyp einzeln durchgeführt werden sollte. Die Wichtigkeit experimenteller Designs in mehrphasigen Experimenten wurde hervorgehoben und die statistische Auswertung zweiphasiger Designs in einem gemischten Modell sowie die bivariate Analyse verschiedener Merkmale wurde durchgeführt. Bisher konnte kein molekularer Marker für die VL assoziiert werden, allerdings soll dies in einem Nachfolgeprojekt anhand weiterer genotypischen Daten und der Erstellung einer genetischen Karte erreicht werden.

9. Appendix

9.1 Supplementary material of Chapter 4



Supplemental 1: GO slim analysis of *Dianthus caryophyllus L.* transcripts of the category Cellular Component and Molecular Function. Plotted are the numbers of unigenes against the GO slims.

9.2 Supplementary material of Chapter 5

Supplemental 2: Example SAS Codes

SAS code for bivariate analysis of SL and CF in E1

```
proc mixed data=SLCF;
class REP BOX P GENO W TRAIT;
model Y = TRAIT TRAIT*GT TRAIT*REP /ddfm=residual;
random TRAIT/sub=GENO type=UNR;
random TRAIT/sub=REP*BOX type=UNR;
random TRAIT/sub=REP*BOX*P type=UNR gcorr;
repeated TRAIT /sub=REP*BLOCK*P*W type=UNR;
parms (0.2292)(2.4094)(0.4092)(0.01262)(0)(-0.01262)(0.1439)
(3.1609)(0.0265)(0.02359)(4.0885)(0.21);
run;
```

SAS code for bivariate analysis of VL and CF in E1

```
proc mixed data=VLCF;
class REP BOX P GENO TRAY POS STO DV W TRAIT;
model Y = TRAIT TRAIT*GT TRAIT*REP TRAIT*STO /ddfm=residual;
random TRAIT /sub=GENO type=UNR;
random TRAIT /sub=GENO*STO type=UNR;
random TRAIT /sub=REP*BOX type=UNR;
random TRAIT /sub=REP*BOX*P type=UNR gcorr;
random P2*DV;
random P2*DV*REP*TRAY;
repeated TRAIT /sub=REP*BOX*P*W type=UNR;
parms (2.5063)(2.4516)(-0.5128)(0.7788)(0)(0.2035)(0)(0)(0)
(0.6097)(3.1253)(0.0856)(0.2058)(0.8742)(4.6906)(4.0096)(0.6119);
run;
```

SAS code for bivariate analysis of VL and FD in E2

```
proc mixed data=VLFD;
class REP BOX P GENO TRAY POS STO DV W TRAIT;
model Y = TRAIT TRAIT*GT TRAIT*REP TRAIT*STO/ddfm=residual;
random TRAIT /sub=GENO type=UNR;
random TRAIT /sub=GENO*STO type=UNR;
random TRAIT /sub=REP*BOX type=UNR;
random TRAIT /sub=REP*BOX*P type=UNR;
random TRAIT /sub=DV type=UNR;
random TRAIT /sub=DV*rep*TRAY type=UNR gcorr;
repeated TRAIT /sub=REP*BOX*P*W type=UNR;
parms (0)(0.394)(-0.394)(0.4735)(0)(7.3719)(3.4014)(3.1144)
(-3.4479)(0.4355)(0.2023)(0.1326)(6.1588)(2.8889)(-8.094);
run;
```

Supplemental 3: Example ASReml Codes

ASReml code for bivariate analysis of CF and SL in E1

```
rep !I 4
box !I
p !I
geno !I 2500
W
GT !I
y
trait !I
nr !I 1903
D:\Desktop\ASReml\CFSL\E1.txt !SKIP 1 !EXTRA 5 !MAXIT 30 !AISINGULARIES
y ~ trait trait.GT trait.rep !r trait.geno trait.rep.box trait.rep.box.p trait.nr
1 1 3
0 0 0 !S2==0.00001
trait.geno 2
trait trait US 1 .1 1
geno geno ID !GF
trait.nr 2
trait trait US 1 .1 1
nr nr ID !GF
trait.rep.box 2
trait trait US 1 .1 1
rep.box rep.box ID !GF
trait.rep.box.p 2
trait trait US 1 .1 1
rep.box.p rep.box.p ID !GF
```

ASReml code for bivariate analysis of VL and CF in E1

```
REP !I 4
BOX !I
P !I
GENO !I 2500
STO !I
W
DV
GT !I
TRAY !I
POS !I
Y
TRAIT !I
P2
NR !I 1898
D:\ASREML\VLCF\E1.TXT !SKIP 1 !EXTRA 5 !MAXIT 30 !AISINGULARIES
Y ~ TRAIT TRAIT.GT TRAIT.STO TRAIT.REP !R TRAIT.GENO TRAIT.GENO.STO TRAIT.REP.BOX TRAIT.REP.BOX.P P2.DV
P2.DV.REP.TRAY TRAIT.NR
1 1 5
0 0 0 !E1==0.00001
TRAIT.GENO 2
TRAIT TRAIT US 1 .1 1
GENO GENO ID !GF
TRAIT.GENO.STO 2
TRAIT TRAIT US 1 .1 1
GENO.STO GENO.STO ID !GF
TRAIT.NR 2
TRAIT TRAIT US 1 .1 1
NR NR ID !GF
TRAIT.REP.BOX 2
TRAIT TRAIT US 1 .1 1
REP.BOX REP.BOX ID !GF
TRAIT.REP.BOX.P 2
TRAIT TRAIT US 1 .1 1
REP.BOX.P REP.BOX.P ID !GF
```

ASReml code for bivariate analysis of VL and FD in E2

```
rep !I 4
box !I
p !I
geno !I 2500
STO !I
W
DV
GT !I
TRAY !I
POS !I
y
trait !I
P2
nr !I 1898
D:\ASReml\VLFD\E2.txt !SKIP 1 !EXTRA 5 !MAXIT 30 !AISINGULARIES
y ~ trait trait.GT trait.STO trait.rep !r trait.geno trait.geno.STO trait.rep.box trait.rep.box.p DV DV.TRAY trait.nr
1 1 5
0 0 0 !S2==0.00001
trait.geno 2
trait trait US 1 .1 1
geno geno ID !GF
trait.geno.STO 2
trait trait US 1 .1 1
geno.STO geno.STO ID !GF
trait.nr 2
trait trait US 1 .1 1
nr nr ID !GF
trait.rep.block 2
trait trait US 1 .1 1
rep.box rep.box ID !GF
trait.rep.box.p 2
trait trait US 1 .1 1
rep.box.p rep.box.p ID !GF
```