Investigating genotype by environment and QTL by environment interactions for developmental traits in potato

Paula Ximena Hurtado Lopez

Thesis committee

Promotors

Prof. dr. R.G.F. Visser
Professor of Plant Breeding
Wageningen University

Prof. dr. F.A. van Eeuwijk
Professor of Applied Statistics
Wageningen University

Co-promotor

Dr. ir. C.A. Maliepaard
Assistant professor, Laboratory of Plant Breeding
Wageningen University

Other members
Prof. dr. I. Romagosa, University of Lleida, Spain
Prof. dr. ir. P.C. Struik, Wageningen University
Prof. dr. ir. R.F. Veerkamp, Wageningen University
Prof. dr. ir. A.J. Haverkort, Plant Research International (PRI), Wageningen

This research was conducted under the auspices of the C.T de Wit Graduate School of Production Ecology and Resource Conservation.

Investigating genotype by environment and QTL by environment interactions for developmental traits in potato

Paula Ximena Hurtado Lopez

Thesis

Submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Wednesday 12 December 2012
at 4:00 p.m in the Aula.

Paula Ximena Hurtado Lopez
Investigating genotype by environment and QTL by environment interactions for developmental traits in potato
160 pages.

PhD thesis, Wageningen University, Wageningen, The Netherlands, 2012
With references, with summaries in Dutch and English

ISBN 978-94-6173-438-9

CONTENTS

Chapter 1

General introduction

Potato – Origin and genetic background

Potato *(Solanum tuberosum* L) is a tuberous and starchy crop species in the Solanaceae family. According to the Food and Agricultural Organization (FAO) it is the world's third-largest food crop, following rice and wheat. Potato originally comes from the Andean region of South America, more specifically southern Peru and north-western Bolivia, where it was domesticated 7,000–10,000 years ago (Innovation 1989; Spooner et al. 2005a). The first edible potatoes are considered to have been cultivated about 4000 years ago in Peru. The South American Indians were able to select alkaloid-free potato varieties and after centuries of selective breeding there are now over a thousand different types of potato (Innovation 1989; Pena-Cortes et al. 1989). Among the 5000 cultivated varieties worldwide, 3000 are found in the Andes alone, mainly in Peru, Bolivia, Ecuador, Chile, and Colombia. Apart from the cultivated species, about 200 wild species and subspecies are found throughout the Americas, from the United States to southern Chile (Hijmans and Spooner 2001).

Although most of the cultivated potato varieties are tetraploids, the first cultivated potato species were diploid and some of them are still cultivated in South America (Innovation 1989; Raker and Spooner 2002). The development of modern varieties was related to the spontaneous occurrence of tetraploid species with superior yield. The major species grown worldwide is *S. tuberosum* (tetraploid with 48 chromosomes), and modern varieties of this species are the most widely cultivated. The ploidy level has been one of the most important taxonomic characters to recognize cultivated potato species, containing diploid (2n = 2x = 24), triploid (2n = 3x = 36), tetraploid (2n = 4x = 48), and pentaploid (2n = 5x = 60) cultivars (Spooner et al. 2010). The majority of diploid species are self-incompatible and tetraploids are self-compatible allopolyploids that exhibit disomic inheritance (Hawkes 1990). *S. tuberosum* an is autotetraploid with tetrasomic inheritance and two major subspecies have been reported: andigena, or Andean; and tuberosum, or Chilean (Raker and Spooner 2002). The Andean potato is adapted to the short-day conditions prevalent in the mountainous equatorial and tropical regions from where it originated. The Chilean potato, native to the Chiloé Archipelago, is adapted to the long-day conditions prevalent in the higher latitude region of southern Chile. It has been suggested that subsp. andigena originated from two diploid cultivated species (Matsubayashi 1991), while subsp. tuberosum evolved from

subsp. andigena after transport to Chile, parallel with the evolution of subsp. andigena in Europe (Hawkes 1990).

"European" potato - origin and distribution worldwide

The first reported presence of cultivated potato in Europe was on the Canary Islands in 1567 (Hawkes and Franciscoortega 1993; Rios et al. 2007). The large-scale cultivation of the crop began only at the beginning of the 19th century. Initially, the crop was used as a medicinal plant and grown by pharmacists in Spain, it was later introduced to other parts of Europe by merchants and king (Ames and Spooner 2008). Potato was slowly adopted as a major food crop and in some European countries it was rejected as an acceptable food crop well into the late 1700s. By 1800s, potato was established as a field crop in England, it was widely grown in the rest of Europe and became an important staple crop due to its high yielding capacity and highly nutritious qualities (Salaman 1949). From Europe, potato spread to the rest of the world during the 18th and 19th centuries and it has been referred to as the "European" or "Irish" potato (Ames and Spooner 2008). It rapidly migrated to the tropics and subtropics (Center 1984; Simmonds 1971) and it is now a worldwide crop grown under many climatic conditions. The successful use of the crop did not only require changes in the dietary habits of the people, but also biological adaptation of the crop to other day length conditions.

Until 2007, two controversial hypotheses about the origin of the European potato existed. The first one proposed a Chilean origin due to similarities between Chilean landraces and modern European cultivars regarding morphology and tuberization under long days (Juzepczuk and Bukasov 1929). The second hypothesis has been widely accepted for the last 60 years. It suggested that potato came from the Andes and tuberization evolved from Andean landraces to a Chilean-type. These potatoes persisted until the late blight epidemics in the UK in 1845, after which they were replaced with Chilean germplasm through new introductions and breeding efforts (Salaman 1937). The Andean origin has been questioned in the last years due to Indian landraces showing to be of Chilean origin (Spooner et al. 2005b). A recent DNA study has resolved the controversy about the origin of the European potato. It has been shown that in fact the first potatoes came from the Andes and persisted until 1982 while the Chilean potato first appeared in Europe in 1811, long before the late

blight epidemics(Ames and Spooner 2008). It refutes the idea that Chilean germplasm replaced the Andean types after the epidemics of 1845. These new evidences showed that the European potato germplasm was derived from a breeding process with high latitude Chilean landraces beginning in 1845 before the late blight epidemics. It partially explains the adaptation of the European potato to long photoperiod. However, the selection of cultivars against the obligatory demand of short photoperiod for tuber initiation and growth also played a role. The long days and moderate temperatures of the temperate regions of Europe allowed longer periods of photosynthesis, efficient translocation of assimilate from haulm to tubers and low transpiration rates during the cool nights.

In the 20th century the need for new resistance genes against pest and diseases promoted the use of South American wild potato species for introgression of such genes in the European varieties. Therefore, modern European varieties are the result of intensive breeding programs improving traits like disease resistance, tolerance to environmental factors, and others.

Worldwide production

In the last 20 years the world potato production has been undergoing major changes. Until the early 1990s potato was mainly grown and consumed in developed countries with a major production in Europe. The Food and Agriculture Organization of the United Nations (FAO) has made available the updated worldwide potato production data until 2010 (FAO 2012). Production and demand has shown a dramatic increase in Asia, where output rose from less than 65 million tonnes in the early 1990s to more than 153 million tonnes in 2010 (Figure 1). Since 2007 Asian production has exceeded that of Europe; in 2010 almost 50% of the world production was harvested in Asia.

Although Asia and Europe account for more than 80 percent of the world production, the highest average yields are obtained in regions with a moderate climate, such as the northern United States and north-western Europe, where potato is grown with modern cultivation methods under long days and moderate temperatures. North America has been the leader in yields with more than 40 tonnes per hectare since 2004. Europeans are still the major consumers but per capita consumption is increasing in Africa and Latin America.

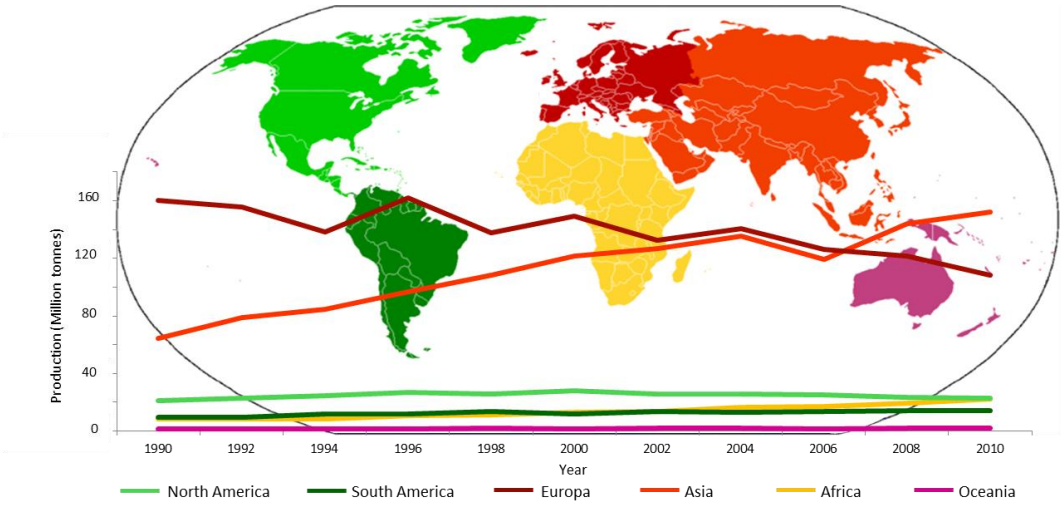


Figure 1. Worldwide potato production per continent between 1990 and 2010 according to FAOSTAT

Developmental events

A potato plant experiences a number of developmental events that can be divided in two main groups: A) above ground events related to canopy development, and B) below ground events related to tuber formation and maturation. These developmental events occur in several stages. Above ground events are divided in five main stages: 1) emergence, 2) canopy growth 3) flowering, 4) berry production and ripening of the fruits and 5) senescence. On the other hand, below ground events occur in five major stages: 1) sprout development from eyes on the seed tuber, 2) root and stolon development, 3) tuber initiation at stolon tips, 4) tuber bulking and 5) maturation (maximum tuber dry matter content).

The most interesting and striking fact is the apparent lack of synchrony between developmental stages above and below ground. During a large part of the plant cycle, it is possible to observe newly formed stolons, swollen stolon tips, tuber incipient and growing tubers of different sizes, while above ground, the plant produces leaves, flowers and berries and may even be senescing (Celis-Gamboa et al. 2003). Figure 2 illustrates the above and below ground events as well as the different developmental stages in each one. The duration of each developmental stage depends among others on cultivar and environmental factors, such as temperature, soil type, water availability and day length.

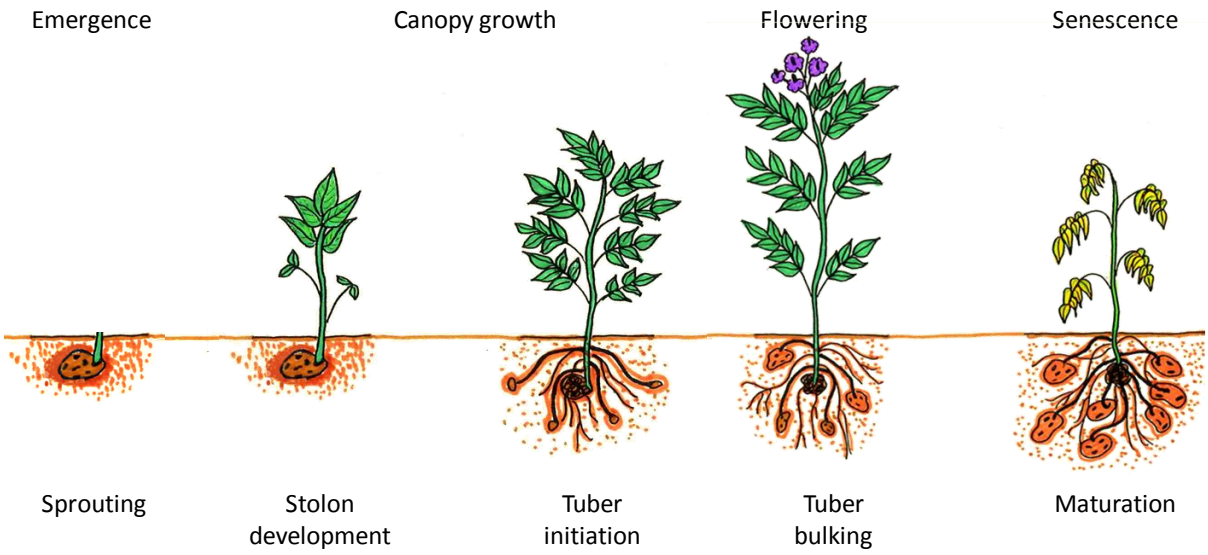


Figure 2. Potato developmental stages.

Effect of photoperiod and temperature on potato development

Previous studies have shown temperature and photoperiod to be the major environmental factors controlling development in potato (Ewing and Struik 1992; Levy and Veilleux 2007). Above- and below-ground parts of the plant as well as the different stages of the potato developmental cycle are differentially affected by both factors. For instance, it is known that photoperiod has little or no effect on the duration of developmental phases after tuber initiation (Haverkort and Kooman 1997; Kooman et al. 1996b), but these phases are determined mainly by temperature (Kooman et al. 1996b). The photoperiod signal is perceived in the leaves, so that potato plants are able to respond to it immediately after emergence (Ewing and Struik 1992). Short photoperiods (10–12 h) hasten tuber initiation compared to long photoperiods (14–18 h) (Kooman et al. 1996a; VanDam et al. 1996). On the other hand, temperature has a big impact on tuber initiation and senescence. Temperatures higher than 20°C delay tuber initiation and higher than 30°C shorten crop senescence(Midmore 1984). Thus, both factors have a major influence on the total length of the growth cycle.

The effect of both factors on physiological processes has been extensively studied (Borah and Ilthorpe 1962; Haverkort 1990; Kooman et al. 1996b; Levy and Veilleux 2007; Manrique 1992; Midmore 1984; Struik and Ewing 1995; Struik et al. 1988; Turner and Ewing 1988; VanDam et al. 1996) and different effects have been observed on below and above ground

development. However, the changes on physiological traits in response to them are also influenced by other factors such as cultivar, N fertilization and physiological conditions of the seed tuber (Struik and Ewing 1995).

Above ground: After emergence, photoperiod and temperature affect the development of the individual main stem by influencing the number of stems per shoot, the number of leaves per stem, the life span of the shoot and the leaf characteristics. However it has been reported that the effects on number of stems are also dependent on the cultivar (Struik and Ewing 1995). Under short day conditions the development of the haulm is restricted, while tuber formation and maturation are accelerated (Maris 1964). When photoperiod is longer than a critical day-length, haulm development is greatly promoted while tuber formation is restricted due to the fact that the plants are not physiologically old. Life span of the entire shoot is increased by long days. It has been observed that under long days, high temperatures tend to increase the life span of the potato plant and under short days high temperatures may decrease it (Struik and Ewing 1995). Earlier crop senescence, delayed flowering and reduced vegetative growth have been reported under short days (Haverkort 1990; VanDam et al. 1996).

Below ground: Cultivars of *S. tuberosum* usually grown in Europe and other temperate regions do not require a strong induction to form stolons. However during the early phases of plant development stolon initiation and the switch from stolon initiation to tuberization are stimulated by short days. Therefore, when tuberization is delayed by long days, more and longer stolons and increased number of potential tuber sites are observed (Struik et al. 1988). Similar to the effects of long days, high temperatures also stimulate stolon initiation and stolon growth by delaying tuberization (Struik et al. 1989). Although some reports have shown that long days or high temperatures increase the number of tubers (Struik et al. 1989; Struik et al. 1988), other studies indicated a low optimum temperature for number of tubers, especially under long days (Wheeler et al. 1986) or strong genotype by environment interactions (VanDam et al. 1996). Final yield depends on the effects of photoperiod and temperature on the duration of tuber bulking (Struik and Ewing 1995). However it is known that yield is lower in tropical highlands (short day length) which is mainly due to the short season (Haverkort 1990). The short day sensitivity of potato implies that varieties that make

use of the whole growing season and produce well in northern Europe (5-6 months), may mature too early and experience a short senescence period which reduced yield. In the same way, varieties that perform well at low latitudes start tuberizing in 3 to 4 months and mature too late under long days. Despite the rather large effect of temperature and photoperiod discussed previously, the changes on physiological traits in response to these factor are also influenced by other biotic and abiotic factors such as cultivar, N fertilization and physiological conditions of the seed tuber (Struik and Ewing 1995).

Optimal conditions for potato growth in different agro-climatic zones

In tropical highlands where environmental conditions are similar to those in the centre of origin, potato is a staple food. Considering that two crops per years can be grown, potential yields are very low. On the other hand, in the tropical lowlands potato production is only possible during the cooler dry period of the year. Yields there are higher than in the tropical highlands basically because of the absence of late blight (Haverkort 1990).

In subtropical highlands the low night temperatures assured adaptation of the crop, while the long cloudless days assure maximal radiation and thus, high dry matter production. Although the absence of rain avoids the need for late blight control the major constrains are the lack of irrigation and the killing frost. In these areas, potato is also a staple food.

In the subtropical lowlands two distinct growing seasons are observed (spring and autumn). Spring is the highest yielding season, it starts in the rainy season and its foliar development coincides with increased solar radiation (Fahem and Haverkort 1988). In autumn, the initial high temperatures favour foliar development and the shortening days favour subsequent tuberization (Haverkort 1990).

In the temperate zones, the largest fields and highest yield are encountered in the lowlands of the northern latitudes. Centuries of selection produced genotypes adapted to the long days; adequate rainfall, the use of healthy seed and high levels of inputs assure high yields. In these regions yields mainly depend on solar radiation and rainfall (Haverkort 1990).

Despite the differences in temperature, growing seasons, agronomic practices and biotic/abiotic factors in each agro-climatic zone, the range of optimum temperatures for potato production has been estimated between 18 and 24°C (Kooman and Haverkort 1995).

The estimation has considered some of the optimal temperatures reported in the literature (Borah and Ilthorpe 1962; Ingram and McCloud 1984; Manrique 1992; Midmore 1984; Sands and Regel 1983); a wider range of optimum temperatures has been observed in long days than in short days. This wider optimum temperature range for growth of the crop in longer days explains the broad adaptation of potato to a wide variety of climates. Besides its short day and low temperature origin, the crop also performs well in regions with higher temperatures such as the Mediterranean region and temperate regions during summer because of the long days during the growing season (Haverkort and MacKerron 1994; Kooman and Haverkort 1995).

Complex data analysis in potato

Potato experiments conducted in the field and greenhouse are a generous source of multiple types of data. In addition to phenotypic data measured only once during the experiments, there are also traits where data are collected at several instances during the growing season in order to closely monitor the development of the plants. Therefore, complex data analysis considering multiple environments, multiple traits and multiple observations during the growing season can provide the basis to understand plant development as a complex and dynamic process influenced by different factors. Such studies can also benefit from molecular data to identify pleiotropic genetic regions and networks of genes driving developmental and physiological processes.

1.Multi-environment studies

In breeding programs data collected in multi-environment trials (METs) provide information to predict cultivar performance used for breeding programs. Cultivar means estimated from METs are the best predictors of future performance but they are normally estimated with error due to the low number of replicates, locations or years included in the studies. Although this error can be minimized by increasing the experimental sizes, in term of replicates and environments, METs are expensive and time consuming.

In potato extensive studies simulating different day lengths and different temperature regimes have been performed to identify the effects of temperature and photoperiod on physiological processes (Borah and Ilthorpe 1962; Haverkort 1990; Kooman et al. 1996b; Levy and Veilleux 2007; Manrique 1992; Midmore 1984; Struik and Ewing 1995; Struik et al.

1988; Turner and Ewing 1988; VanDam et al. 1996). These artificial conditions avoid the confounding effects of photoperiod and/or temperature with other environmental factors such as rainfall, abiotic/biotic stresses, etc. Therefore the qualitative effects of photoperiod and temperature on traits such as haulm size, productivity, onset, duration and rate of tuber bulking are relatively well-known (Struik et al. 1989; Struik and Ewing 1995; Struik et al. 1988; Turner and Ewing 1988; VanDam et al. 1996; Wheeler et al. 1986). However, quantitative information on the effects of photoperiod and temperature on the different aspects of tuber formation, development and growth and their interactions are limited. In addition, a major drawback of most studies is the number of cultivars included in the experiments. For instance, genotype by environment interactions have been reported for a few traits but using only a limited number of cultivars (VanDam et al. 1996). MET studies for a wide range of cultivars or mapping populations are scarce.

Despite the efforts to study the physiological response of potato to different environmental conditions and to select cultivars with good performance in specific locations, a genetic component is still missing in most of the studies. Understanding the genetic basis of development and other complex traits can support previous physiological studies and facilitate breeding strategies. Complex traits are the result of physiological processes and environmental influences during the crop cycle which involve interactions of many gene pathways and networks. Therefore the detection of quantitative trait loci (QTL) for these types of traits in environments with contrasting conditions can be done using information generated from multi-environment trials (METs). The inclusion of populations of genetically related individuals in such trials, facilitate the understanding of the genetic control of adaptive traits by identifying associations with QTLs (Boer et al. 2007; Malosetti et al. 2008). Hence, the combined use of genetic and phenotypic information from METs allows the detection of genetic factors controlling development and complex traits and also to study QTL by environment interactions.

2.Multi-trait studies

Although many QTL studies considered multiple traits, usually those traits have been analysed separately. An integrated analysis combining traits related to developmental processes simultaneously is required to get a better understanding of the genetic and

environmental forces driving plant development. QTL analysis combining data of multiple traits related to plant development will not only increase the power of QTL detection, it will also improve the understanding of the genetic control of developmental processes. As a consequence, a multi-trait QTL analysis of a single population allows for the detection of closely linked chromosomal regions affecting several traits simultaneously (Jiang and Zeng 1995).

In potato a first attempt to estimate the optimal set of consensus QTLs for several traits simultaneously was done through a QTL meta-analysis (Danan et al. 2011). It allowed the co-localization of late blight resistance and plant maturity traits by projecting individual QTLs onto a consensus map. However, there are no reports of such integrative analyses for developmental traits in potato. So far, data on traits related to plant development in potato have not been integrated into a single study to get insight in the genetic architecture of crop development and the presence of putative pleiotropic-QTL related to developmental processes.

3.Multiple time points during the growing season (time series data)

The ability to understand the genetic basis of plant development depends on an accurate description of developmental morphology, yet such descriptions are often lacking and conclusions are drawn based on the observations of fully grown phenotypes (Kellogg 2004). This means that similarities between structures or developmental phases are often superficial. Therefore, a proper characterization of development over time, taking into account the continuous nature of development, is needed to describe each part of the process. Therefore, growth models have to be implemented that are flexible enough for the usually non-linear trait responses over time. Examples of commonly used models in biology are the exponential growth models and the family of s-shaped curves (Schnute 1981). These models have the advantage of describing the development of the trait in terms of curve parameters with a biological interpretation. The differences in growth trajectories between genotypes are reflected by genotype-specific curve parameters. Growth models and QTL analysis can be combined by modelling growth curve parameters in terms of QTLs in one-step (Ma et al. 2002b; Malosetti et al. 2006) or two-step model approaches (Reymond et al. 2003; Yin et al. 1995b). In the former, logistic growth curves and QTL mapping are combined

modelling growth parameters as a function of the QTL genotypes described by molecular marker scores. In the latter, first genotype-specific parameters are estimated from the growth curves, which are then used as phenotypic traits in a conventional QTL analysis. In both approaches, the accurate estimation of the parameters describing the curves plays an important role when meaningful results have to be found.

In potato, previous studies have incorporated well characterized time series data into growth models and QTL analysis, allowing for instance the genetic description of senescence in terms of parameters at different aging stages (Hurtado et al. 2012b; Malosetti et al. 2006). Further studies in potato incorporating the time-dependent nature of plant development into QTL analysis have not been reported; therefore the genetic control of different developmental stages is still poorly understood.

Motivation and purpose of this thesis

Wageningen UR Plant Breeding has carried out studies on different developmental processes of potato making use of a segregating population evaluated under field conditions in different day length regimes. The field trials included the evaluation of below and above ground traits in different stages of the growing season. Extensive molecular data is available on this experimental potato population as well. Thus, this population provided an excellent opportunity for genetic studies based on information previously generated in research projects making use of statistical methods and models to integrate the different data sets and developmental stages. Therefore, based on previous studies we identified four major areas of potato research on which we have focused the present study.

First, a genetic component is still missing in most of the physiological studies. Physiological responses under different climatic conditions have been extensively studied in the last 30 years (Borah and Ilthorpe 1962; Haverkort 1990; Kooman et al. 1996b; Levy and Veilleux 2007; Manrique 1992; Midmore 1984; Struik and Ewing 1995; Struik et al. 1988; Turner and Ewing 1988; VanDam et al. 1996). Most studies have been performed under highly controlled conditions, simulating different day lengths and different temperature regimes. Despite the efforts to study the physiological response of potato to different environmental conditions and to select cultivars with good performance in specific climates, a genetic

component is still missing in most of these studies. Understanding the genetic basis of development and other complex traits can support previous physiological studies and facilitate breeding strategies.

Secondly, the presence of closely linked or pleiotropic chromosomal regions affecting developmental and agronomic traits has not been investigated. Potato studies including molecular information normally present separate QTL analyses for each trait. So far, data on traits related to plant development in potato have not been integrated in a single study to get insight in the genetic architecture of crop development and the presence of putative pleiotropic-QTL related to developmental and agronomical processes.

Thirdly, quantitative information on the effects of photoperiod and temperature (major environmental factors controlling development) on the different aspects of tuber formation, development and growth and their interactions are limited. In addition, a major drawback of most studies is the number of cultivars included in the experiments. Studies of a wide range of cultivars or mapping populations are scarce. Therefore genotype by environment interactions (GEI), as the major complication in potato breeding, are not fully understood for most of the agronomic and developmental traits.

Fourthly, potato studies incorporating the time-dependent nature of plant development into QTL analysis are scarce. Previous studies have incorporated well characterized time series data into growth models and QTL analysis, allowing for instance the genetic description of senescence in terms of parameters at different aging stages (Hurtado et al. 2012b; Malosetti et al. 2006). Further studies in potato incorporating the time-dependent nature of plant development into QTL analysis have not been reported.

Scope and outline of the thesis

In this thesis we make use of data from multi-environment trials (METs) where the CxE diploid potato population has been evaluated. This population was previously developed for research purposes (Jacobs et al. 1995) based on the different genetic background of *S.*

phureja and S.tuberosum and their phenotypic contrast. The most relevant characteristics of *S. phureja* are the lack of a tuber dormancy period, early maturity and short day tuberization induction (Ochoa 1990). In contrast, *S.tuberosum* varieties are grown in temperate regions, have long tuber dormancy, variable maturity and long-day tuberization induction (Celis-Gamboa 2002; Hawkes 1990). These characteristics make the CxE population suitable for the study of phenological processes along the life cycle and also to study performance and adaptation to different day length conditions. Thus, in our study the CxE population was used to investigate 1) the use of flexible methods to model potato growth and development in a genetic context, 2) the genetic factors underlying dynamic developmental and growth processes, 3) the presence of pleiotropic genetic regions controlling developmental and agronomical traits simultaneously, 4) the presence of genotype by environment interactions, GEI, and QTL by environment interactions, QEI, for developmental and agronomical traits.

In Chapter 2 we propose methods to model ordinal and continuous data over time and show their use in a biological context. The methods provide a flexible framework which can accommodate the behaviour of all genotypes in a test population. Haulm senescence and plant height were used as example traits to model development and growth processes. The characterization of the fitted curves enabled us to zoom in on certain physiological stages during the growing season. The characteristics derived from the curves were used to describe different phenological phases.

In Chapter 3, the genetic factors underlying phenological phases of haulm senescence were investigated in two consecutive years (2004 and 2005), under field conditions in Finland using a flexible method to model ordinal data. The available time series data were used in a smoothed generalized linear model to characterize curves describing the senescence development in terms of its onset, mean and maximum progression rate and inflection point. These characteristics together with the individual time points were used in a QTL analysis to identify among others time-dependent QTLs.

In Chapter 4, the aim was to identify the genetic basis of plant developmental processes in potato by means of a multi-trait QTL analysis combining several traits describing plant

development and agronomic characteristics measured under short day length. Parameters derived from fitted curves for flowering, senescence and plant height were simultaneously analysed with agronomic traits in a multi-trait QTL analysis to investigate the presence of pleiotropic genetic regions controlling those traits.

In Chapter 5, multi-environment data collected under 3 contrasting day lengths, in Ethiopia and Venezuela (short), The Netherlands (long) and Finland (very long) was used to study genotype by environment and QTL by environment interactions. Flowering, haulm senescence and plant height were evaluated as a time series during the growing season, while some important agronomic traits were measured at harvest to get a better understanding of potato development and adaptation under the contrasting environments.

In Chapter 6, we highlight relevant facts related to multi-environment trials (METs) that have been identified during our research and we discuss the implications of complex data analysis, including multiple time points, multiple traits and multiple locations for potato breeding. We close this chapter with the main conclusion and outlook of this thesis.

Chapter 2

Flexible tools for analysing phenotypic data collected over time

Sabine K. Schnabel[1,2], Paula X. Hurtado-Lopez[1,3,5], Paul H.C. Eilers[1,4], Chris Maliepaard[2,3], Fred A. van Eeuwijk[1,2]

[1] *Biometris, Wageningen UR, P.O. Box 100, 6700 AC Wageningen, The Netherlands,*
[2]*Centre for BioSystems Genomics, P.O. Box 98, 6700 AB Wageningen, The Netherlands*
[3]*Plant Breeding, Wageningen UR, P.O. Box 386, 6700 AJ Wageningen, The Netherlands*
[4]*Department of Biostatistics, Erasmus Medical Center, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands*
[5]*C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC), Wageningen University, P.O. Box 386, 6708 PB Wageningen, The Netherlands*

To be submitted

Abstract

Field and greenhouse experiments are a valuable source of phenotypic data of plants. In addition to data collected only once during the experiments, there are also a number of traits for which data are collected at several instances during the growing season to closely monitor the development of the plants. By modelling a developmental trait over time we can incorporate the dynamic nature of the process described by a fitted curve. The continuous nature of growth processes, such as plant height as well as ordinal scales used to indicate developmental stages of senescence or flowering can be modelled through flexible methods. We present here flexible curve modelling techniques for ordinal and continuous data (the most common data types) collected over time. The methods provide a framework that can accommodate the behaviour of all genotypes in a test population. They are also flexible enough in terms of the length of data series to be analyzed. The characterization of the fitted curve enables us to zoom in on certain time periods in the growing season. Therefore characteristics derived from these curves can be used to describe different phenological phases. This improves the comparison across environments since the curve characteristics can be evaluated instead of individual time points, which would not have the same interpretation in different environments. Therefore developmental processes can be compared across different studies. These derived characteristics also provide a good basis for subsequent QTL analysis. Thus the use of curve characteristics in QTL analysis also favours the study of time-dependent QTLs and QTL by environment interactions. To illustrate the methodologies we are making use of plant height and haulm senescence time series data collected from a diploid potato population under contrasting field conditions.

Key words: Curve modelling, phenotypic traits, plant development, statistical methods, time series data, QTL mapping

1. Motivation and introduction

Plant experiments in the field and greenhouse are a valuable source of multiple types of data. In addition to phenotypic data measured only once during the experiments, there are also traits where data is collected at several instances during the growing season in order to closely monitor the development of the plants. Previous studies have made used of these dynamic phenotypic data in genetic studies (Ma et al. 2002b; Malosetti et al. 2006; Wu and

Lin 2006b; Wu et al. 2004). However, many researchers present results on data from single time points without making use of the knowledge of the process that lies behind it. In the following we propose methods to pre-process and analyze dynamic phenotypic data in a flexible framework. Thorough data preparation is also a key to appropriate and successful statistical analysis and useful results. Therefore we also comment on possible problems with data collection and preparation. Ultimately we want to use our results to perform a QTL analysis with characteristics that can be associated to the developmental process rather than to phenotypic traits measured at single time points.

The multitude of available data requires flexible methods. The proposed techniques are flexible enough to analyze different types of data such as ordinal and continuous data. The methods provide a framework that can accommodate the behaviour of all genotypes in a test population. When it comes to traits such as senescence, flowering, plant height etc. not all cultivars complete the development trajectory within the experimental time frame. This leads to data series that are not easily described by the same parametric model. The family of S-shaped curves where growth converges to a maximum has been used to described biological processes (Schnute 1981). A classic example for a parametric model for growth and development is the (two parameter) logistic model (Verhulst 1845). Especially late genotypes with an incomplete growth curve at harvest cannot be modelled by a logistic curve. The use of a semi-parametric model overcomes this problem and all genotypes can be captured inside the same framework and described by similar characteristics. As we have seen in our analyses, our methods are also flexible enough in terms of the length of data series to be analyzed. Beyond a certain minimal number of data points the models give equally good results for longer and for shorter data series. This is of special value as it helps to analyze existing datasets that might not have been suited for analysis with other methods. Describing a developmental trait through a curve enables us to zoom in on certain time periods in the growing season. We can define characteristics for these parts and search for QTLs only associated with a period of interest. Therefore we can also investigate properties such as progression rate for a period between defined developmental stages. Characteristics of the curve can be easily related to the biological processes behind it. Thus they provide a good basis for a successful subsequent QTL analysis. Plant breeders in

research as well as in companies can profit from making better use of the available data and ultimately arrive at more productive QTL analyses.

In Section 2 we describe the types of data as well as their handling for the intended use with the curve fitting techniques presented in Section 3. This section also includes an extensive example using a diploid potato population. We show the results in terms of the fitted curves as well as the resulting QTL analysis for plant height in two different environments. The paper closes with a short conclusion and outlook on future work.

2. Data

2.1. General description and motivation

As described in the introduction, field and greenhouse experiments produce a range of phenotypic data. In the following we are proposing methods for so-called dynamic traits measured over time during the growing season. Growth characteristics such as plant height, and developmental traits, such as senescence or flowering can be expressed in different ways. Developmental processes are usually scored on an ordinal scale to indicate which stage the plant in question has reached. We will take haulm senescence as an example. In this case the potato plants were scored on a discrete scale from “green plant” (1), “upper leaves with first signs of yellowing” (2) etc. to “dead plant” (7) at several instances during the growing season (Hurtado et al. 2012b). For these data we propose to fit a smooth curve over time as described in Section 3.1.1.

Plant height is typically a continuous variable. In the case considered here, potato plants were measured by using the longest stem of each plant as the distance between the ground and the main apex. For these data we estimate a smooth frontier curve as presented in Section 3.1.2. Working with phenotypic data collected over time in field trials offers a lot of possibilities, but also bears some problems. Complications can arise when analyzing data from the same population but from different experiments (e.g. in different years, locations, under different irrigation levels or fertilizer treatments). This can be due to differences in field designs: in order to have sufficient data for fitting a curve over time, measurements from several blocks (in time or space) might have to be combined. Different harvest dates for entire or partial blocks result in data series of varying length. In addition, poorly

documented scales for ordinal scores might cause problems. E.g. only the first and the last level of a scale are well described while the intermediate levels are not explained in detail or not at all. Especially when working with historical data (data from not so recent experiments) often the person in charge of data collection is hard to come by in order to overcome these problems.

Better phenotyping protocols for frequently used research populations –such as the CxE potato population described below– can help to improve future data quality and consistency across locations and years of experimentation. Other means to improve the data quality are discussed in the conclusion of this paper.

2.2. Data description

For the analysis in the following we are using data from two field experiments of the CxE potato population(Jacobs et al. 1995). This is a diploid backcross population resulting from the cross of two diploid potato clones. The parents involved are characterized by diverging dormancy, maturity type and different tuberization properties (Celis-Gamboa 2002). Therefore the resulting CxE population is especially suited for studying developmental traits. This population was planted and evaluated in a series of trials at different locations and years. The phenotypic data from the 2004 experiment in Finland (Zaban et al. 2006) and the 2010 experiment in Ethiopia (Hurtado et al. 2012a) will give the examples in our analysis. Table 1 summarizes the basic information about the locations and the experiment.

Table 1: Summary of local conditions and general information of the field experiments

	Location	
Local conditions	**Ruukki, Finland**	**Holetta, Ethiopia**
Coordinates	64°24'N, 25°00'E	09°04'N, 38°01'E
Altitude	48m	2400m
Average daily air temperature (°C)	12.7	14.26
Temperature range (°C)	2.6-20.5	11.3-17.5
Planting date	June 1, 2004	July 16, 2010
Number of genotypes	197	169
Observation period senescence (days)	52	53
Observation period plant height (days)	49	48

Although the two environments are rather different in terms of geographic location and altitude, the observation period both for senescence and plant height are similar. In Figure 1 an overview of the data is given. Here, we collapsed the measurements of all genotypes and indicate overlapping observations by darker shades of grey. Due to the data collection on special days the graphs show a stripe patterns. More observations in each time point for senescence and plant height were obtained from the Ethiopian experiment. Data collection was done every week and the experiment included more replications per genotype. On the time axis instead of the classical days after planting (DAP) we used thermal days adjusted for day length and temperature (photo beta thermal time units – PBTT). This takes into account the temperature as well as the exposure to daylight. These two factors are known to have a major influence on plant growth and development (Ewing and Struik 1992; Levy and Veilleux 2007). Details on the transformation including the cardinal temperatures of potato (Khan et al. 2012) and the daily air temperature in Finland are presented in a previous study (Hurtado et al. 2012b).

2.3. Data handling

The methods in this paper are developed for phenotypic data from field and greenhouse experiments. For the traits under investigation measurements are taken at different times during the growing season. There are several factors that can already influence the recording of the data. For haulm senescence one problem is missing data at the beginning of the growing season. Typically the first category “green plant” (1) is under-represented in the data. Although a (potato) plant has been in this stage since its emergence from the ground often it is not mentioned in the data and only from first signs of yellow on the leaves (for at least one plant in the entire field) the data collection really started. Therefore, we imputed additional observations of the first level of the scale at the beginning of the growing season when it was clear from the data that the plants emerged and had not shown any signs of ageing. By doing so, we introduce a lower asymptote at level 1. The last level of the ordinal scale can be seen as an upper asymptote. Haulm senescence is a monotone increasing process. Thus, we did the necessary data cleaning to remove inconsistencies. In practice, we checked the data series per single plant whether an observation was preceded by an

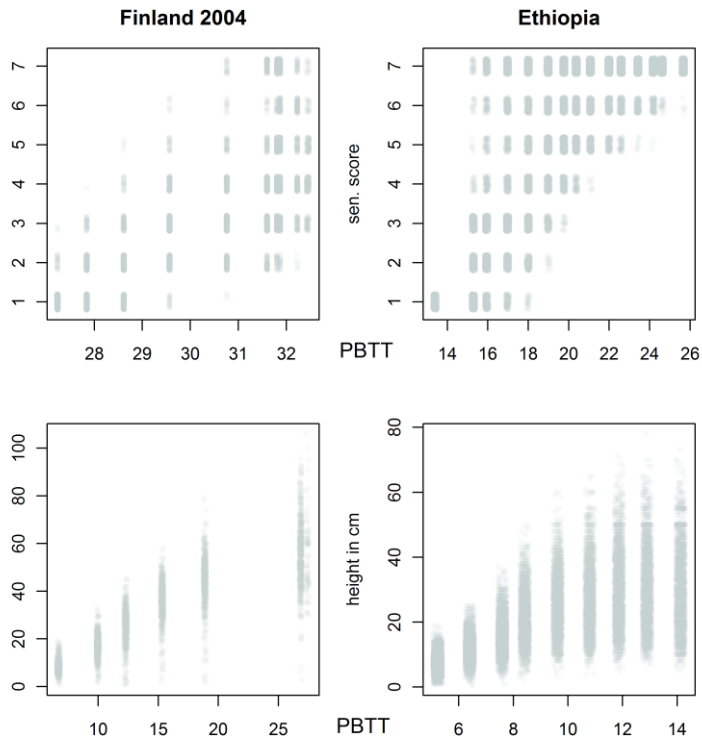


Figure 1: Original data for haulm senescence (top row) and plant height (bottom row) from Finland 2004 (first column) and Ethiopia (second column). All data is plotted in one figure. For better visualization positions on the *x*−axis and *y*−axis have been jittered. Darker shades indicated observations that overlap. The darker the area, the more observations are identical between genotypes.

observation of the same or lower score. In the latter case, we removed the higher observation in order to have monotone increasing data. This is in line with the definition of haulm senescence as the haulm can only stay in the same stage or progress on the scale.

For continuous data as plant height there are similar problems. Obviously plant height is only recorded once the plant is growing. In order to have comparable curves as a result of our modelling activity, we decided to impute pseudo-observations of 0 cm before the date of emergence. For comparing the same characteristics on all developmental traits we also assumed that plant height is a monotone increasing process. We had to do some data cleaning following the same philosophy as for haulm senescence. Due to the measurement of plant height in our study (see Section 2.1. above) plant height can also only stay the same or show an increase. We also imputed pseudo-observations at the end of the measurement period by carrying forward the highest measured height.

For all traits we pooled the data for the different replicates and estimated curves per genotype instead of curves per plant. As the data used in the example originates only from one complete block of the experiment or from a complete block design, no additional adjustments are necessary for accounting for possible block effects. However, in general for data from more complicated experimental designs we advise to first correct for experimental design effects such as block effects.

3. Methods for modelling dynamic phenotypic data

3.1. Curve modelling for ordinal and continuous data

3.1.1. A modelling technique for ordinal data over time

In this section we propose a method to fit a smooth curve on time series of ordinal monotone increasing data. In what follows we are describing the approach for the data at hand, but this technique can be generalized easily to other data situations e.g. flowering or other development related traits (Hurtado et al. 2012a).

Haulm senescence is measured on an ordinal scale $y \in [1,n]$ with $n=7$ in our case. From the field experiment we obtained a discrete data series of senescence scores. Figure 2A shows an example for raw data from one selected genotype in our study. Possible data imputations are explained in Section 2.3.

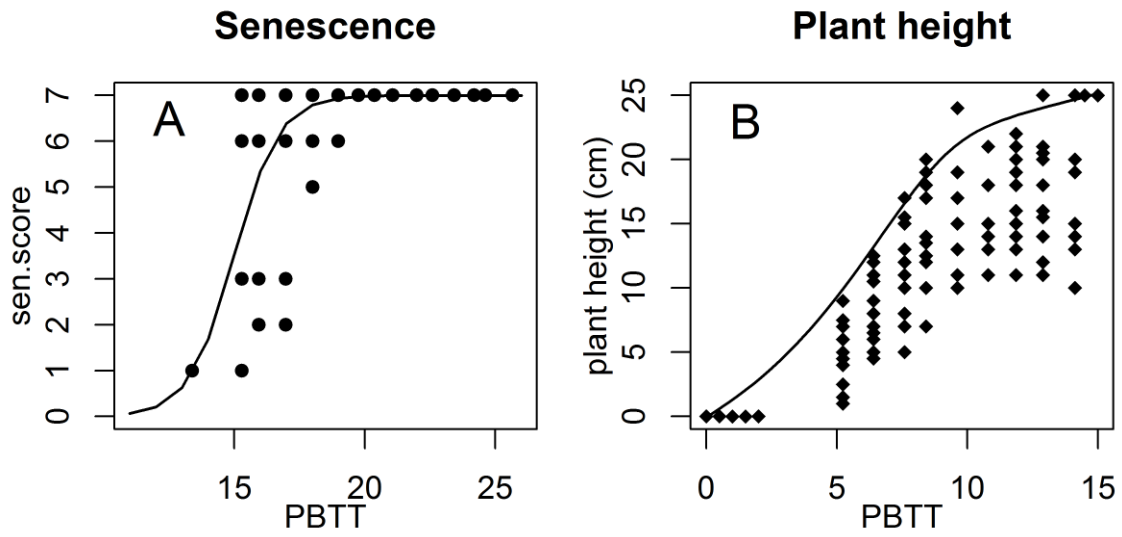


Figure 2. Raw data and fitted curves for haulm senescence (A) and plant height (B) of one selected genotype evaluated in Ethiopia.

We re-interpret the senescence score *y* as having *y* successes in *n* binomial trials. A smooth curve is then fitted for the success probability *p*, 0<*p*<1. We use a logit transformation of the probabilities as the linear predictor, i.e.

$$\eta = log\left(\frac{p}{1-p}\right) \in R. \qquad (1)$$

The estimation uses the analogue between the parametric technique of logistic regression and its non-parametric counterpart the generalized linear model (GLM) (Nelder and Wedderbu.Rw 1972) with a logit link. We want to model the development of senescence over time with a smooth curve. We make use of *P*–splines (Eilers and Marx 1996) as the functional form for our smooth curves. An adaptation of the original form is applied in the context of modelling senescence scores over time (Hurtado et al. 2012b).

Our goal is to estimate a smooth curve for haulm senescence over time per genotype. Results can be seen in Figure 3. As the plant population in the example originates from a cross of two contrasting parents in terms of plant maturity, we expect to see a range of different shapes of the senescence curves. This cannot easily be captured in the classical logistic model. However with splines a flexible framework is provided to capture all types of curves in the same framework. This is an additional advantage of the proposed technique.

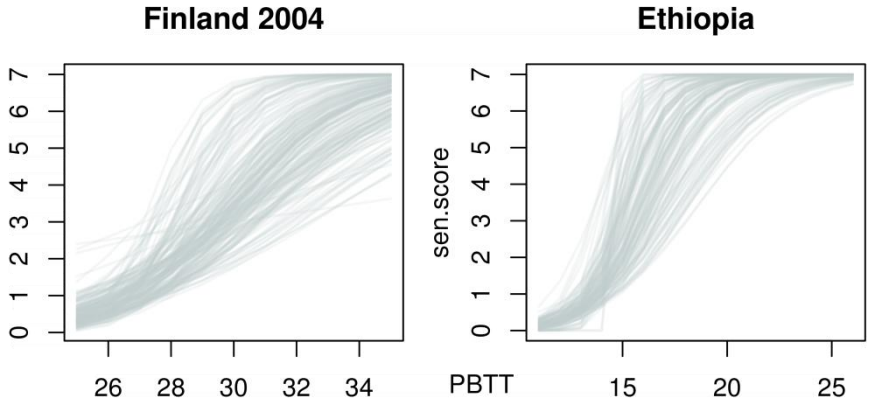


Figure 3: Fitted curves per genotypes for haulm senescence. All curves shown in one figure. Darker shades indicate overlap of the curves.

3.1.2 Modelling a continuous phenotypic trait over time

In this section we present a new approach to estimate a smooth curve for measurements of a continuous variable over time. As we are interested in the best performance of a genotype –e.g. in light of selection for future breeding programmes– we aim at modelling an upper frontier instead of the mean trend for the relationship between the trait and time. In the plant breeding context an upper frontier can be interpreted as an upper potential. Therefore we choose a frontier estimation technique. The methodology presented can be used for other data situations and is not limited to monotone time series of data. An example for the application to plant height in the context of a multi-trait analysis can be found in a previous study (Hurtado et al. 2012a).

From the field experiment we obtain data in form of a set of continuous plant height measurements per genotype. An example for the observed data –including some imputed observations as explained in Section 2.3- can be found in Figure 2B. As we are interested in modelling the upper potential of a genotype we propose to use expectiles (Newey and Powell 1987) to estimate a curve at the upper border of the data cloud. Expectile curves are an asymmetric generalization of the classical least squares approach and provide a suitable model for estimating frontier curves. Estimation includes a weight p, $0<p<1$ for observations located above the estimated curve and a weight $1-p$ for points located under the curve. Ordinary least squares is a special case for p=0.5. Formally we are minimizing

$$S=\sum_i w_i(p)(y_i-\mu(x_i,\alpha,p))^2 \qquad (2)$$

with weights

$$w_i(p) = \begin{cases} p & \text{if } y_i > \mu(x_i; \alpha; p) \\ 1\text{-}p & \text{if } y_i \leq \mu(x_i; \alpha; p) \end{cases} \qquad (3)$$

Here, y_i is the response variable. μ is the estimated expectile based on the explanatory variable x_i, parameters α of the functional description of the curve and the selected asymmetry p. For the expectile curve μ we choose P–splines (Eilers and Marx 1996) as a flexible functional form and α are the spline coefficients. This combination has been previously described in detail (Schnabel and Eilers 2009). In the context of plant height we opted for an asymmetry parameter of p=0.99.

Our goal is to estimate a smooth curve for plant height over time per genotype. Results can be seen in Figure 4. Using the flexible functional form different types of curves can be captured in the same framework.

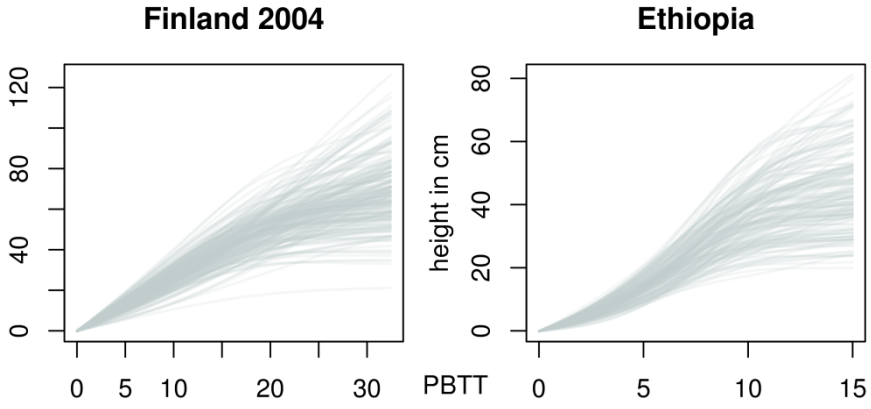


Figure 4: Plant height fitted expectiles curve per genotype in the diploid potato population. All curves shown in one figure. Darker shades indicate overlap of the curves.

All estimations and visualizations have been done within the R software environment (CoreTeam 2011). For modelling continuous variables we used the package expectreg (Sobotka et al. 2012).

3.2. New traits for QTL mapping from fitted curves

The above described procedures result in smooth curves over time for developmental and growth traits. Once the curves are fitted per genotype we use the first and second differences on the curves to deduce characteristics with a biological interpretation. These are the empirical approximations of the first and second derivatives of the curves. The point where the 2nd derivative of the curve is maximal is used as a trait in genetic analysis. This can be interpreted as the onset of the developmental process. If applicable in the context of the experiment, the point where the 2nd derivative of the curve is minimal can be seen as the end of the developmental process. Furthermore we derive other characteristics from the fitted curves. These include the mean slope (as a general indication of the speed of the process), the maximum slope during onset (indicating the maximum progression rate of the initial phase of growth) and the maximum and mean of the continuously measured phenotypic traits such as plant height.

Standard parameters derived from logistic models are also estimated: The maximum slope of the curve (maximum rate of growth/development) as well as the time point where it occurs. By definition this is the inflection point of the curve. It often coincides with the point in time when half of the process is reached.

Phenotypic traits that are measured at several instances over time are usually used as single observations and a QTL analysis per time point is done. Depending on the number of measurements this approach can amount to a larger number of analyses and results in potentially different QTLs that can be only associated with a single point in time. By modelling a developmental trait over time we can study a process that is best described by a fitted curve. Characteristics of this curve can be used to describe different phenological phases and they can be used for QTL mapping afterwards.

3.2.1. Results of the QTL analysis

A direct application of the parameters derived from fitted curves is shown in the multi-trait/multi-environment QTL analysis shown in Figure 5. The analysis was performed using some of the characteristics describing the fitted curves for plant height in Ethiopia and Finland. We have made use of parameters derived from the fitted curves that a general logistic regression does not estimate. The parameters presented in the figure are: 1) mean slope, indicating the average progression rate during the growth process, 2) onset of plant height defined in terms of PBTT, indicating the beginning of the exponential part of the growth process and 3) maximum slope during onset, indicating the maximum progression rate of the initial phase of growth.

Significant QTL by environment interactions were found for all QTLs. We identified two time-dependent and environment-specific QTLs on C4 and E12 associated only to the initial part of plant growth in Ethiopia (Figure 5). The first one was related to the maximum progression rate during onset and the second one was associated to the time when the plant started the exponential phase of the growth process. One QTL on E5 was detected during the whole growth process in Ethiopia and Finland with a major effect. The genetic region where this QTL was located has been previously related to senescence and plant maturity (Celis-Gamboa 2002; Hurtado et al. 2012b; Malosetti et al. 2006). Thus, our results

not only support previous findings but provide new insights into the genetics of growth in contrasting environments.

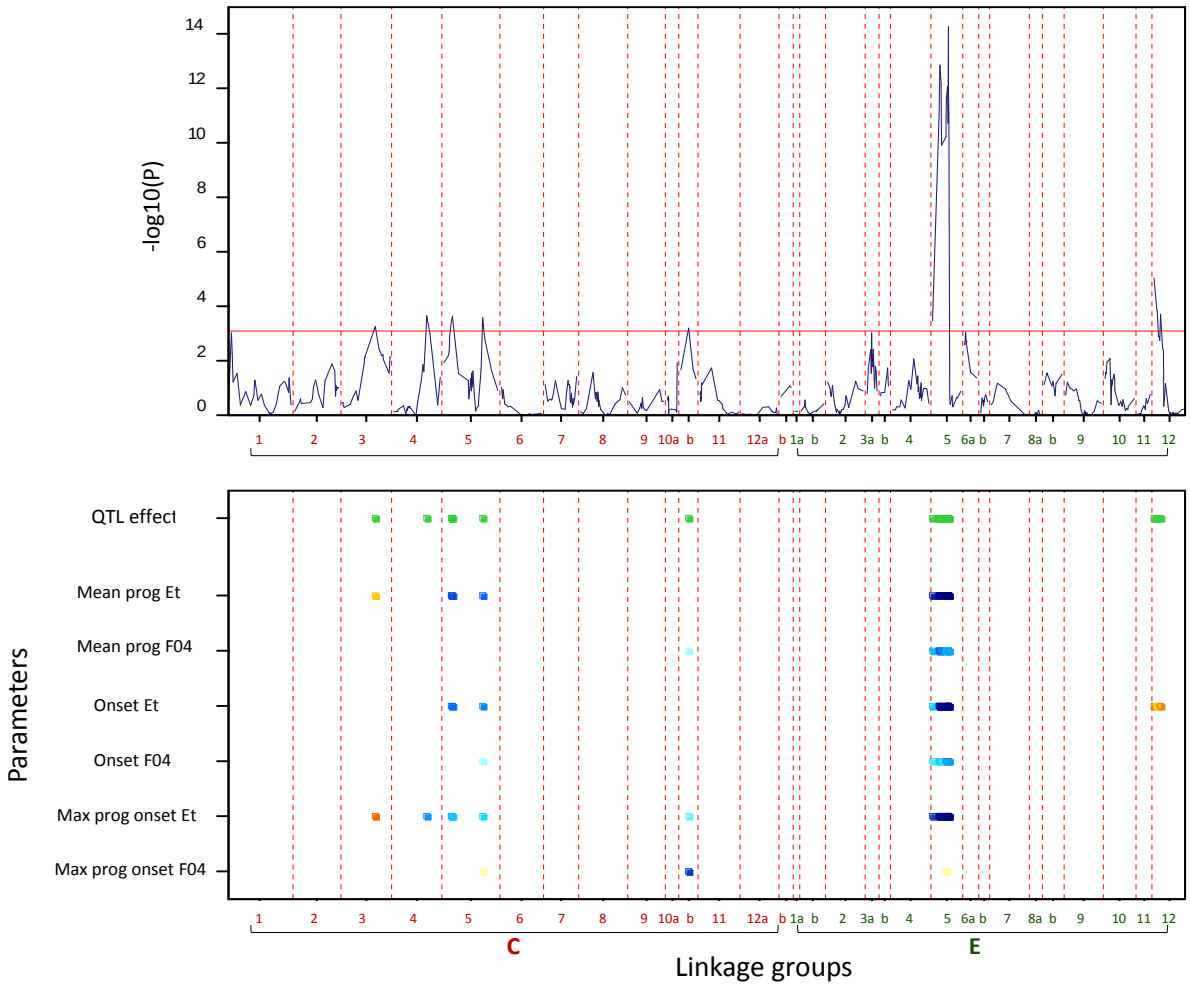


Figure 5: Multi-trait QTL linkage analysis combining parameters estimated for plant height fitted curves in Ethiopia (Et) and Finland 2004 (F04). The linkage groups of each parental map (C and E), which have been previously described (Hurtado et al. 2012a) are indicated on the x-axis . The upper plot shows the significance of QTLs (-$\log_{10}$ scale for the associated probability value). The lower plot shows the positive (red) and negative (blue) allele substitution effects at positions of a significant QTL. The intensity of the colour is proportional to the QTL effect size (the darker the colour, the larger the effect). Three parameters were analysed in each environment: mean progression rate, onset and maximum progression rate for onset.

Here we only show the results for plant height, but we performed a similar analysis for haulm senescence. Details on the QTLs found for the Finnish data set as well as a comparison of the results for the derived traits and the analysis per time point are presented in (Hurtado et al. 2012b). A comprehensive treatment of haulm senescence, flowering and plant height for Ethiopia is given in (Hurtado et al. 2012a).

4. Conclusions

In addition to some recommendations on data handling we presented two curve modelling techniques to analyze data series of phenotypic traits measured over time in field experiments. Ordinal and continuous data were analyzed here which include the most common data types. Although the presented data was monotone (increasing), we can also adapt our methods to accommodate non-monotone series of observations. This is one of the advantages of using a flexible functional form for the curves. Another advantage is the curve fitting of genotypes with an incomplete growth curve at harvest (e.g. late genotypes) which cannot be modelled by a logistic curve. The use of a semi-parametric model overcomes this problem and all genotypes can be captured inside the same framework and described by similar characteristics. The model described here also gives us the opportunity to zoom in on different phases of the development in time and investigate more in detail. Additional characteristics can be defined according to the need of the research study and no additional data collection is needed. In our example we made use of this property by including the maximum progression rate for the onset of plant growth as a trait into the QTL analysis. This new trait leads to the discovery of three new QTLs.

Even in different environments the respective curve characteristics have the same biological meaning. Therefore developmental phases can be compared in different studies. Comparing results from different environments using data from individual time points is less straightforward. In this case the researcher has to make sure that he or she is comparing the time points corresponding to the same stage of plant development. Thus, the use of curve characteristics in QTL analysis also favours the study of time-dependent QTLs and QTL by environment interactions.

5. Outlook

We suggest to optimally scale the response variable when working with data from experiments at different locations or years. Scales might be poorly documented and a comparison is difficult. This can be the case with ordinal scales where only the first and the last category is well described while the intermediate levels are not well explained. Optimal scaling is a technique from psychometrics. Questionnaires often use scales where the order of the different responses is known, but the difference between two subsequent categories

is unknown. In expressing the scale in integer number it is assumed that neighbouring categories are equidistant. This might not relate to the actual situation. In optimal scaling, the response is rescaled to make the relation between the transformed response and the explanatory variable linear. This linear summary is the simplest form to express the relationship and its parameters can be used in QTL mapping. An introductory description of optimal scaling -the response-and an extension to optimally scaling -the time axis- have been recently presented (Schnabel et al. 2011a; Schnabel et al. 2011b). A more comprehensive treatment of optimal scaling in plant breeding will be reported elsewhere.

Automatic phenotyping is a very recent field of research. With these facilities more and much larger phenotypic data sets will become available and they will also include more often developmental and growth traits measured in time. Analyzing these data per individual time point is time consuming, computationally intensive and will require multiple testing corrections as there will be correlations among consecutive time points. Curve modelling is well suited to summarize longer series of repeated measurements into new traits with a biological meaning. Statistical modelling can also provide valuable methods for analyzing the image data that is produced in high throughput phenotyping. Often pictures are also taken over periods of time which brings us back to the situation that motivated the research presented in this paper. We will explore more applications of our methodology to denser data sets and observations originating from automatic phenotyping.

Acknowlegdments

Part of this project was (co)financed by the Centre for BioSystems Genomics (CBSG) which is part of the Netherlands Genomics Initiative/Netherlands Organisation for Scientific Research.

Chapter 3

Dynamics of senescence-related QTLs in Potato

Paula Hurtado López[1,2,5], Sabine Schnabel[2,6], Alon Zaban[3], Merja Veteläinen[3], Elina Virtanen[3], Paul Eilers[2,4], Fred van Eeuwijk[2,6], Richard GF Visser[1,6], Chris Maliepaard[1]

[1.] Wageningen UR Plant Breeding. Wageningen University. P.O Box 386, 6700 AJ Wageningen, The Netherlands
[2.] Biometris – Applied Statistics. Wageningen University. P.O Box 100, 6700 AC Wageningen, The Netherlands
[3] AgriFood Research Finland (MTT), North Ostrobothnia Research Station in Ruukki, Finland
[4] Department of Biostatistics. Erasmus Medical Center. PO Box 2040 3000 CA. Rotterdam, The Netherlands
[5] C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC). Wageningen University. P.O Box 386, 6708 PB Wageningen, The Netherlands
[6] Centre for BioSystems Genomics, PO Box 98, 6700 AB Wageningen, The Netherlands

Published in Euphytica (2012) 183:289–302

Abstract

The study of quantitative traits expression over time helps to understand developmental processes which occur in the course of the growing season. Temperature and other environmental factors play an important role. The dynamics of haulm senescence was observed in a diploid potato mapping population in two consecutive years (2004 and 2005) under field conditions in Finland. The available time series data were used in a smoothed generalized linear model to characterize curves describing the senescence development in terms of its onset, mean and maximum progression rate and inflection point. These characteristics together with the individual time points were used in a Quantitative Trait Loci (QTL) analysis. Although QTLs occurring early in the senescence process coincided with QTLs for onset of senescence, the analysis of the time points made it difficult to study senescence as a continuous trait. Characteristics estimated from the senescence curve allowed us to study it as a developmental process and provide a meaningful biological interpretation to the results. Stable QTLs in the two experimental years were identified for progression rate and year-specific QTLs were detected for onset of senescence and inflection point. Specific interactions between loci controlling senescence development were also found. Epistatic interaction between QTLs on chromosomes 4, 5 and 7 were detected in 2004 and pleiotopic effects of QTLs on chromosomes 3 and 4 were observed in 2005.

Keywords

Beta thermal time; epistasis; functional QTL mapping; smoothing; time series; GLM.

Introduction:

The study of developmental processes in plants requires the evaluation of traits over time taking into account the continuous nature of development. Quantitative traits normally involved in these processes require assessments at several time points during the life cycle. However, conventional experiments include evaluations at a single fixed time point during the growing season. Quantitative trait loci (QTL) analysis using the data collected from those experiments give just an impression of the loci affecting the trait at a particular developmental stage. The understanding of the genetic basis controlling quantitative traits improves when the evolution of the trait during the life cycle and the developmental

pattern are considered. In addition, environmental factors affecting crop development, such as temperature and photoperiod, must be also taken into account. Therefore, growth models flexible enough to be adapted to the usually non-linear trait responses over time have to be implemented. Examples of commonly used models in biology are the exponential growth models and the family of s-shaped curves (Schnute 1981). These models have the advantage of describing the development of the trait in terms of curve parameters with a biological interpretation. The differences in growth trajectories between genotypes are reflected by genotype-specific curve parameters. Growth models and QTL analysis can be combined by modelling growth curve parameters in terms of QTLs in one-step (Ma et al. 2002a; Malosetti et al. 2006) or two-step model approaches (Reymond et al. 2003; Yin et al. 1995a). In the former, logistic growth curves and QTL mapping are combined modelling growth parameters as a function of the QTL genotypes described by molecular marker scores. In the latter, first genotype-specific parameters are estimated from the growth curves, thereafter they are used as phenotypic traits in a conventional QTL analysis. In both approaches, the accurate estimation of the curve parameters plays an important role to find meaningful results.

In our study, haulm senescence was assessed at several time points during the growing season on a diploid potato mapping population evaluated for two consecutive years in Finland (2004 and 2005). We are presenting here an alternative two-step approach: modelling first in a flexible way the curve trajectories with a smoothed generalized linear model (GLM) procedure. We used the collected time series data to model the development of senescence in terms of time and temperature. To allow a better comparability of the trait across the years, we converted the calendar days after planting into thermal days after planting (TAP) using a non-linear temperature effect function described by Khan et al. 2012. Once the senescence trajectory was defined, the characteristics of the curve -such as onset, maximum and average progression rate and inflection point- were estimated from the first and second derivative of the curve. These characteristics were used as phenotypic traits describing senescence development in the QTL analysis. In this way, the characterization of the smoothed GLM curves allowed us to map the genetic basis of the senescence process giving a meaningful biological interpretation to the results.

Materials and methods:

Description of the CxE potato population

The evaluation of haulm senescence was done on a diploid backcross potato population (CxE) composed of 250 genotypes. The CxE population is the result of a cross between two diploid potato clones. The female parent, the C clone US-W5337.3 (Hanneman and Peloquin 1967) is a hybrid between *Solanum phureja* (PI225696) and *S. tuberosum* dihaploid US-W42. The male parent, the E clone (Jacobsen 1980), is a hybrid between $VH^3$4211 (a *S. vernei* - *S. tuberosum* backcross) and the C clone itself. *S. phureja* is characterized by a lack of dormancy, early maturity and short day tuberization induction. In contrast, *S. tuberosum* varieties have long dormancy, variable maturity and long day tuberization induction (Ewing and Struik 1992; Hawkes 1990). The different day length requirements for tuberization of *S. phureja* and *S. tuberosum* have provided a source of genetic variation in the CxE population making it suitable for the study of developmental processes under different photoperiods.

Experimental design and description of the haulm senescence evaluation

The very long day conditions of northern Finland were used to conduct an experiment in two consecutive years (2004 and 2005) in the experimental field of AgriFood Research Finland (MTT) at the North Ostrobothnia Research Station in Ruukki (64°42'N,25°00'E). The experiment performed in 2004 was previously described by Zaban et al. 2006. Sets of 197 and 222 genotypes of the CxE population were planted in plots of three plants per genotype, where the genotypes were randomized, on June 1st 2004 and May 16th 2005, respectively. We acknowledge that in these trials there was no real replication of genotypes, but only pseudo-replications. Each plant was evaluated at several time points spaced in intervals of 3-7 days. In 2004 the observation period was 52 days with evaluations at 77, 81, 86, 91, 95, 100, 105, 109, 116, 123, 129 days after planting (DAP). In the second year the observation period was shorter (31 days) and haulm senescence was evaluated at 84, 88, 93, 98, 101, 106, 109, 112 and 115 DAP. The process of senescence was defined as the period between the last observation at which the plant was entirely green and the first date at which the plant was dead (Celis-Gamboa 2002). The progress of the trait was measured using the scale described by Celis-Gamboa et al. 2003 in which 1= green plant; 2= upper leaves with the first signs of yellowing (light green); 3= yellow leaves; 4= 25% of haulm

tissue brown; 5= 50% of haulm tissue brown; 6=more than 75% of haulm tissue brown; 7= dead plant.

Description of Beta thermal time estimation

Crop development is mainly affected by temperature and can be modified by other factors such as photoperiod (Hodges 1990). Previous studies have shown that potato growth is influenced by temperature, where warmer temperatures favour vegetative growth accelerating reproductive and vegetative development (Benoit et al. 1986; Haun 1975; Struik and Ewing 1995). Whereas, lower temperatures facilitate tuber growth (Marinus and Bodlaender 1975). The effect of temperature on crop development rate is often described by using a thermal time concept. Various non-linear models have been developed to describe temperature response of developmental processes in plants (Gao et al. 1992a; Johnson and Thornley 1985; Yin et al. 1995a). In our study, the daily contribution of temperature to plant development in 2004 was different from 2005 due to the fluctuations in daily air temperature. A summary of the meteorological conditions reported by the station in Ruukki (Finland) during the growing season in the two experimental years is presented in Table 1.

Table 1. Summary of meteorological conditions reported in Ruukki, Finland during the observation period in 2004 and 2005.

Metorological variables	**2004**	**2005**
Average daily air temperature (°C)	12.7	14.5
Temperature range (°C)	2.6-20.5	4.8-21.9
Amount of rainfall (mm)	514.0	359.0
Average daily rainfall (mm)	2.1	1.5
Average daily radiation (kJ/cm2)	142.9	155.1
Radiation range (kJ/cm2)	19.2-297.2	35.0-280.7

**Observation period: 52 and 31 days in 2004 and 2005 respectively*

The growing season defined as the period between the last killing frost of spring and the first killing frost of autumn (Allaby 1998) was defined in 2004 from April 16 to November 10 and in 2005 from May 3 to November 14. To allow for a better comparability of senescence

across the years, we converted the calendar DAP into beta thermal days after planting (TAP) using the non-linear temperature effect function *g(T)* described by Khan et al. 2012:

$$g(T) = \left[\left(\frac{T_c - T}{T_c - T_o} \right) \left(\frac{T - T_b}{T_o - T_b} \right)^{\frac{T_o - T_b}{T_c - T_o}} \right]^{c_t} \qquad (1)$$

where the three main temperatures, base, optimum and ceiling temperature for phenological development of potato, were defined as T_b = 5.5 °C, T_o = 23.4 °C and T_c = 34.6 °C, respectively. The temperature response curvature coefficient was estimated as c_t = 1.7 according to Khan et al. 2012. Because the function *g(T)* is non-linear and temperature fluctuates daily, *g(T)* was estimated using the average daily air temperature to obtain the daily value. The cumulative TAP, combining temperature and time (beta thermal time, BTT), was the scale of the x-axis used to compare senescence development in the two experimental years.

Description of the smoothed generalized linear model procedure

To estimate the senescence curve non-parametrically, a variant of penalized B-splines, P-splines (Eilers and Marx 1996), was used. Then we interpreted a senescence value *y* as having observed *y* "successes" in *n* binomial trials and we estimated a smooth curve for the probability *p* of a "success". To guarantee that 0<p<1, we worked with the linear predictor. In this case it is the logit, η = log(*p*/(1 – *p*)), which has no restrictions on its range. This is standard practice in logistic regression and generalized linear models (Nelder and Wedderburn 1972). Smooth logistic regression with *P*-splines was explained by Eilers and Marx 1996. They also showed that automatic interpolation was obtained. We exploited this property by working with zero degree *B*-splines including a knot on each beta thermal day (BTT units) of the domain we study. The penalized log-likelihood in this situation is

$$l^* = \sum_i w_i [y_i \log p_i + (n - y_i) \log(1 - p_i)] - \lambda \sum_i (\Delta^2 \eta_i)^2 / 2 - \kappa \sum_i \eta_i^2 / 2, \qquad (2)$$

with $p_i = 1/(1+e^{-\eta i})$. The weight is determined by $w_i = np_i(1-p_i)$ following GLM methodology for our situation. The second term in the equation above is the penalty on the linear

predictor η. The parameter λ tunes its weight: the larger λ, the smoother the result. We used the operator notation to indicate repeated differences: $\Delta^2\eta_i=\eta_i-2\eta_{i-1}+\eta_{i-2}=(\eta_i-\eta_{i-1})-(\eta_{i-1}-\eta_{i-2})$

One can show that for very large λ the estimated curve for *p* approaches the logistic curve. Furthermore we used a ridge penalty on η tuned by the parameter κ to avoid numerical instabilities, when η becomes too large. To estimate η, the iterative weighted linear regression algorithm (Nelder and Wedderburn 1972) was modified, to account for the penalties.

The fitting procedure resulted in a smooth curve for the development of haulm senescence over time for each genotype. This flexible functional form using splines allowed to model different shapes of these curves (e.g. for early and late genotypes) within the same framework.

Characteristics estimated from the senescence curves

Once the senescence curves were fitted, some characteristics of the curves describing the ageing process were estimated to allow the study of senescence as a continuous trait changing in time. The characteristics estimated from the curve have a meaningful biological interpretation for the senescence development and facilitate the understanding of the QTL mapping results. We calculated the first differences on the curve values (first derivative) as a proxy of the slope of the curve. The mean slope or mean progression rate (*mprate*) is reflecting the average rate of change of the senescence curve during the whole observation period giving an idea of how fast a genotype is experiencing the senescence process. The maximum slope (*prate*) explains the maximum rate of change of the senescence curve when the plant has completed half of the senescence process (between 4 and 5 in the senescence scale). The higher the value, the faster full senescence is reached during the observation period. The inflection point (*ipoint*) reflects the point in time when half of the senescence process has been reached and the trajectory curve change from convex to concave shape indicating the beginning of the final stage. Additional characteristics were deduced from the second differences of the curve values (second derivative), including the maximum and

minimum change of the slope, which are interpreted as the *onset* and the end of the senescence process. The *onset* is indicating the beginning of the senescence process in terms of time or more accurately in our case in terms of BTT. The lower the value, the earlier the senescence process starts.

Repeatability estimation

Data from the three plants within a plot for individual genotypes were used for repeatability estimation, calculated for each year. The repeatability was defined by the ratio of genetic (genotypic) to phenotypic variance. Phenotypic variance was equal to the sum of the genetic variance and a third times the between-plants-within-plot variance.

GxE interactions

The estimation of genotype by year interactions included 186 genotypes evaluated in the two consecutive years using a two-way analysis of variance with genotype and environment fixed in the model. The curve characteristics (*mprate*, *prate*, *ipoint* and *onset*) were used as response variables, *y*, and the genotype (G), year (E) and Genotype by Year (GxE) interaction as the explanatory variables in the model, which also includes an error term (based on the variation between the three plants per genotype):

$$y = G + E + GxE + error \qquad (3)$$

Description of the genetic map and molecular data

The CxE population was genotyped using amplified fragment length polymorphism, AFLP (Celis-Gamboa 2002), simple sequence repeat, SSR, and cleavage amplified polymorphism, CAPS (Werij et al. 2007). Our study included a subset of dominant and co-dominant markers with the expected segregation ratios 1:1 and 1:1:1:1, respectively. JoinMap 4 (van Ooijen 2006) was used to construct the C and E parental maps using 164 and 198 markers respectively, each with 12 linkage groups (LG) previously reported by Celis-Gamboa 2002. The C map consisted of 135 markers spanning 917.4 cM. Four of the 12 LG were split in two sub-groups and one was split in 3 sub-groups due to the large distances between adjacent markers (more than 30 cM). The E map consisted of 132 markers spanning 629.8 cM and 2 of the 12 LG were split in 2 sub-groups. Since the maternal (C) and paternal (E) maps were not integrated due to the expected differences in the recombination frequencies between

the two parents (with genetic background from two different *Solanum* species), the separated parental maps were used to perform the QTL analysis. Co-dominant markers present in both parental maps (C and E) were used to identify the same LG in the two maps. The assignment of linkage groups was done using as a reference the maps of Celis-Gamboa 2002, each LG is preceded by the letter C or E according to the parental map, followed by the LG number.

QTL analysis using curve characteristics and single time points

The detection of QTLs, estimation of QTL main effects and their interactions was done in two steps. In the first step, the detection of QTLs was performed separately for each type of phenotypic data available. In the analysis per time point the senescence scores, from 1 to 7, were used in the QTL mapping. In the analysis of the characteristics estimated from the curves (*mprate*, *prate*, *ipoint* and *onset*), each characteristic was used as a quantitative trait. The use of model parameters in QTL mapping has been introduced in functional mapping for the analysis of time-series traits. It incorporates biological principles (defined with mathematical functions) into the framework for QTL mapping (Wu and Lin 2006a; Wu et al. 2003).

A nonparametric QTL analysis was performed for the time point data using the rank sum test of Kruskal-Wallis (KW). The curve characteristics were analyzed using both KW and interval mapping (IM). Both procedures are available in MapQTL 6 (van Ooijen 2009). For the curve characteristics, QTLs detected by KW and IM were the same. For transparent comparison of QTL analysis on time points and curve characteristics, in this paper we present only results from KW.

QTL mapping was performed separately on the parental maps and the criterion for detecting QTLs was set at a significance level of $p \leq 0.005$ (van Ooijen 2009). To identify the detected QTLs, the origin of the map (C and E parent) and the linkage group are indicated in the QTL name using italic letters. The QTLs were named according to the parental map and the linkage group in which they were identified. As an example, the most significant QTL on linkage group 4 in the C parent was called *C4* and in the case of common markers detected

in both parents, like *E5* and *C5*, the QTL was called *Ch5* (corresponding to a QTL on chromosome 5).

QTLxQTL interactions

IM was used for detecting main effect QTLs for the curve characteristics *mprate*, *prate*, *ipoint* and *onset*. The presence of epistatic interactions was investigated per year by fitting linear models (Genstat release 13.2, (Payne et al. 2010)) containing pairwise interactions between those QTLs for whom earlier main effects were detected for any of the four curve characteristics. A full model was defined with each curve characteristic as response variable, *y*, and the main effect QTLs together with all possible QTLxQTL interactions as fixed explanatory variables.

$$y = QTLs + QTLxQTL\ interactions + error \qquad (4)$$

From the full model (4), for each curve characteristic a subset of model terms was created. This list contained the significant main effects QTLs, all significant epistatic interactions, plus non-significant main effect QTLs whenever these QTLs were involved in significant epistatic interactions. A final model was fitted for each curve characteristic including only the earlier selected terms.

Results

Curve fitting

The transformation of normal calendar days after planting into TAP and their cumulative values generated a new scale (BTT) to measure the senescence development in terms of time and accumulated temperature. BTT went from 33.3 to 41.85 and from 37.1 to 50.8 during the observation period in 2004 (77 to 129 DAP) and 2005 (84 to 115 DAP) respectively. This scale was used as the x-axis to fit the curves and visualize the progression of haulm senescence scored from 1 to 7. The use of the smoothed GLM procedure in our study allowed flexible modeling of different curve shapes. Different types of curves were observed in the CxE population according to the maturity type of each genotype (Fig.1a).

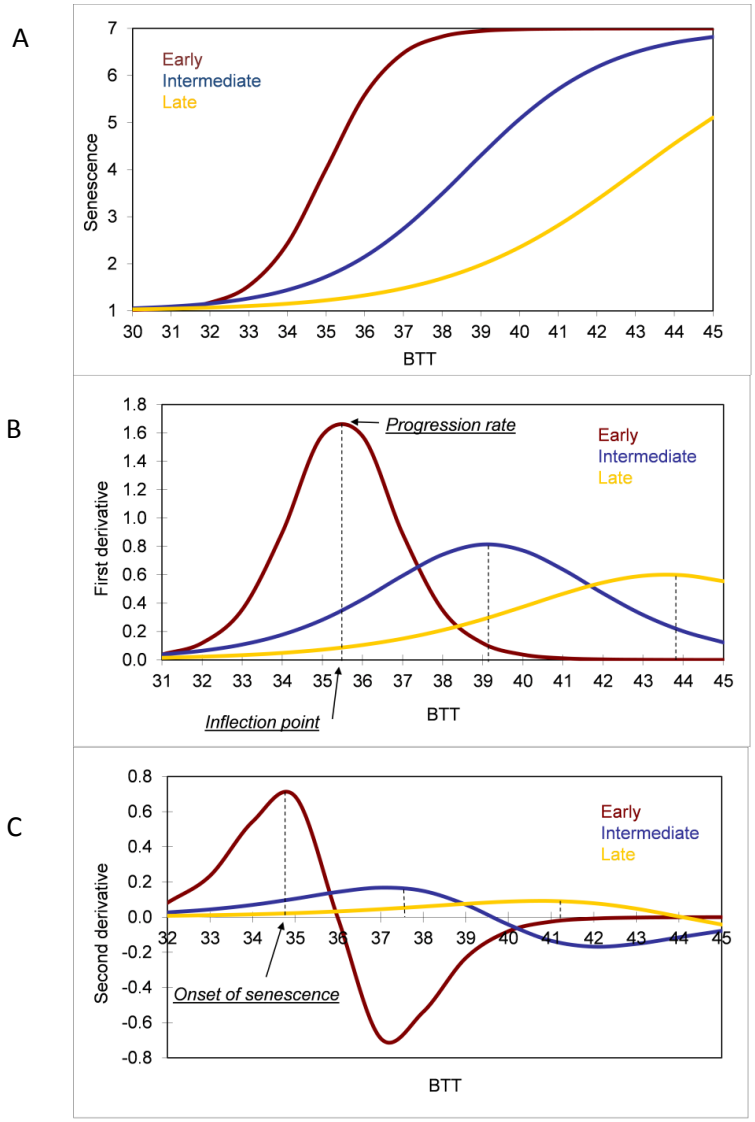


Figure 1. Characterization of senescence development in the CxE population represented by three genotypes with different maturity type: CE78 (early: red), CE47 (intermediate: blue) and CE697 (late: yellow). Fig.1a: Fitted curves showing the senescence development. Fig.1b: First derivative of the curves in which the maximum slope represents the progression rate and the dashed lines represent the time when half of the senescence process has been completed (inflection point) in each genotype. Fig.1c: Second derivative of the curves in which the maximum second difference is interpreted as the onset of senescence and it is indicated with a dashed line in each genotype. BTT: beta thermal time representing cumulative beta thermal days after planting.

The description of the five categories (very early, early, intermediate, late and very late) according to the duration of the plant cycle was reported by Celis-Gamboa 2002. In early genotypes we observed a logistic S-shape in the curve trajectory while in the late genotypes we had only the exponential part of the curves often corresponding to completion of only about half of the senescence process.

After fitting the curves for each genotype, several characteristics describing the ageing process were estimated to be used in the QTL analysis. The first differences on the curves allowed the estimation of *mprate, prate* and *ipoint* (Fig.1b) and the second derivative was used to determine the *onset* of the senescence process (Fig.1c). The minimum second derivative was calculated for the early and intermediate genotypes but it was not accurately estimated for the late genotypes, therefore it was not considered in this study. However, it can be interpreted as the end of the senescence process and experiments with a longer observation period could make use of this characteristic as well to describe the final stage of senescence.

Repeatability estimation

Once the curve characteristics were calculated, the repeatability of each one was estimated for each year. High values were observed for all of them in the two years with values between 0.89 and 0.98. Repeatability was also estimated in each time point and values higher than 0.92 were obtained during the observation period in each year. The high repeatability values observed in the curve characteristics and in the individual time points are due to the experimental design.

GxE interactions

A set of 186 genotypes evaluated in 2004 and 2005 was used to estimate the year effect, the genotype main effect and GxE interaction for *prate*, *mprate*, *ipoint* and *onset* according to model (3). There was a clear year effect for all three traits ($p<0.001$). We observed in 2005 a senescence period lasting longer than in 2004 with a slow progression rate, probably due to the fact that the daily air temperature was higher. In 2004, the average *mprate* and *prate* were higher (0.36 and 0.86) than in 2005 (0.24 and 0.55), whereas the average *onset* and *ipoint* were higher in 2005 (45.97 and 49.90) than in 2004 (37.28 and 39.65). The observation period in 2005 was not long enough for late genotypes to complete the senescence process. Figure 2 shows the senescence fitted curves of three genotypes with different maturity type (CE78: early, CE47: intermediate and CE697: late) in 2004 and 2005 to exemplify the effect of genotype and year in the CxE population.

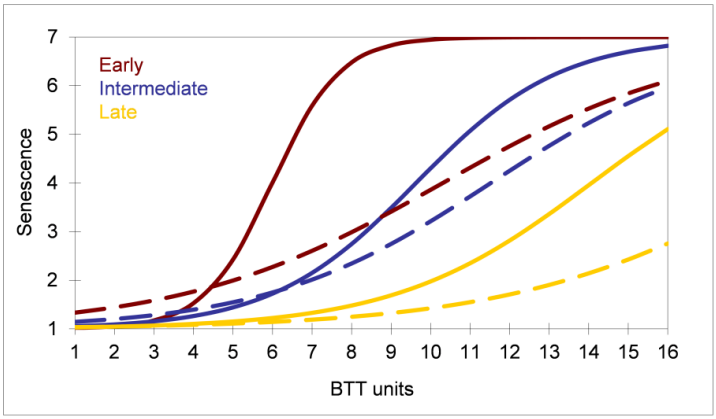


Figure 2. Comparison of senescence development in Finland during 2004 (solid line) and 2005 (dashed line) represented by three genotypes with different maturity type: CE78 (early: red), CE47 (intermediate: blue) and CE697 (late: yellow). The x-axis corresponds to BTT units adjusted for comparison of the two years.

Significant interactions between genotypes and year (p-value<0.001) were also observed for each curve characteristic. Larger differences in *mprate*, *ipoint* and *onset* were observed between early and late genotypes in 2005, whereas the differences between early and late genotypes were larger for *prate* in 2004. Early genotypes had a fast senescence development and late genotypes became very late in 2005.

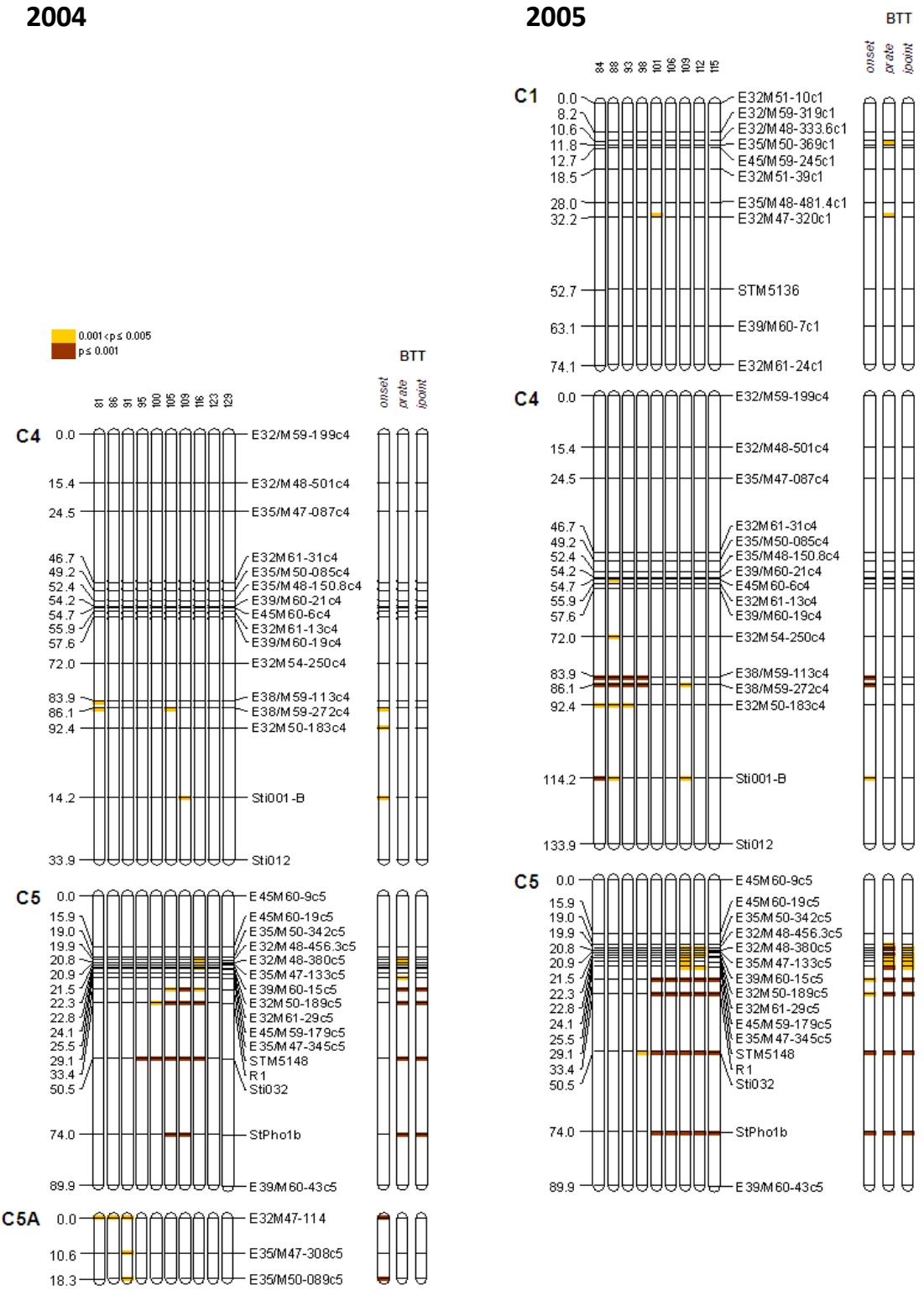
QTL analysis: characteristics of the curves and single time points

QTLs were detected in each parental map using the individual time points and the characteristics estimated from the curves in each year, considering as a significance threshold a p-value from the Kruskal-Wallis test lower than 0.005. Only the LGs in which QTLs were identified are shown in Fig.3 and two levels of significance are highlighted ($0.001<p\leq0.005$, $p\leq0.001$). In the analysis per time point, 10 and 9 time points were used in the QTL analysis in 2004 and 2005 respectively. In the first year, the first time point was excluded from the analysis because all the genotypes were still green. In the analysis using the curve characteristics, although four of them were used for the QTL mapping using KW and IM, results are presented only for *onset*, *prate* and *ipoint* with Kruskal-Wallis. It is due to the fact that high correlations were observed between *mprate* and *prate* in the two consecutive years (0.85 and 0.80 respectively) and the same QTLs were detected for each curve characteristic using both statistical methods.

Comparing the results using time points and curve characteristics, we observed that QTLs detected in early time points, coincided with QTLs on LG *C4, C5A* and *E3* associated with

onset of senescence. QTLs occurring in more advanced stages of the senescence development coincided with QTLs for *prate* and *ipoint* as shown in *C5* and *E5* (they correspond to the same LG in both parents and hence are termed *Ch5*).

A



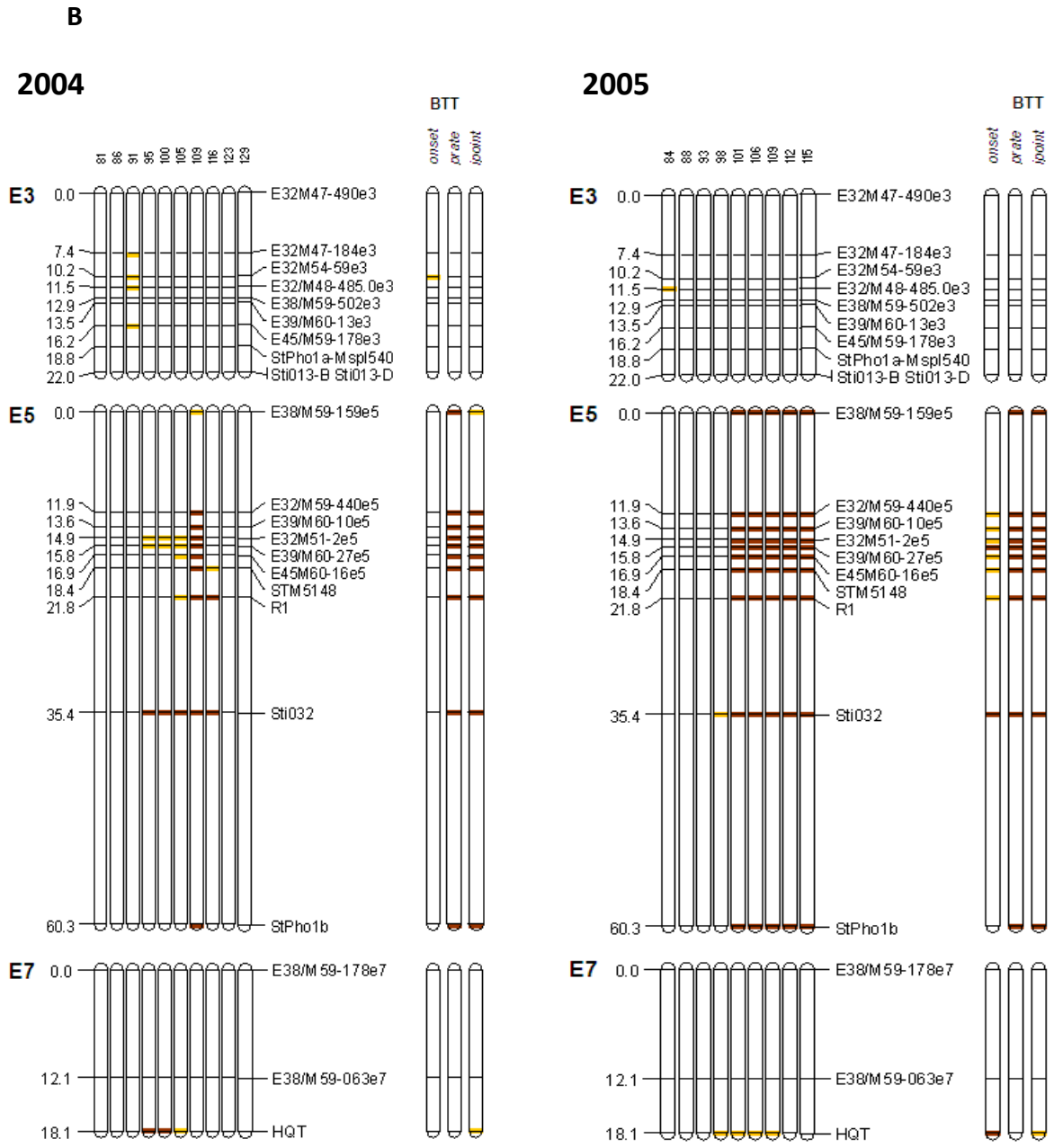


Figure 3. Graphical representation of the QTLs detected in the C (Fig.3a) and E (Fig.3b) parental maps. Results from the QTL analysis using data from the field experiments in 2004 and 2005 are presented in the left and right panel respectively. Only the linkage groups in which significant QTLs were detected using individual time points (in 2004: 81,86,91,95,100,105,109,116,123 and 129DAP; in 2005:84,88,93,98,101,106,109,112 and 115 DAP) and curve characteristics (*onset, ipoint, prate*) are presented. The colors in the linkage groups indicate the significance levels (white: p>0.005, yellow: 0.001<p≤0.005, brown: p≤0.001).

Environment (year)-specific QTLs were found on *C1* and *C5A* associated with *prate* in 2005 and *onset* in 2004, respectively. QTLs associated with *ipoint* were also related with *onset* in 2005 as it is shown in *E7* and *Ch5*. The pleiotropic QTLs detected for *ipoint* and *onset* on both LG explained the high correlation (0.8) observed between the two traits in that year. On the other hand, QTLs consistently found in *C4* in the two experimental years were associated with *onset* and the more significant p-value was observed in 2005 (p<0.0001).

Table 2. Main effect QTLs and QTLxQTL interactions included in the full models for *prate*, *ipoint* and *onset* of senescence in 2004 and 2005. Significant p-values (p<0.05) are highlighted in bold.

			2004			2005		
Variable	**Segregation type[1]**	**Possible genotypic classes[2]**	***prate***	***ipoint***	***onset***	***prate***	***ipoint***	***onset***
E3	"nnxnp"	*nn,np*	0.860	**0.032**	**0.003**	0.856	**0.006**	**0.011**
Ch5	"efxeg"	*ee,ef,eg,fg*	**<0.001**	**<0.001**	**0.007**	**<0.001**	**<0.001**	**<0.001**
E7	"ijxik"	*ii,ij,ik,jk*	0.886	**0.007**	**0.021**	0.223	0.056	**0.004**
C1	"abxaa"	*aa,ab*	0.841	0.614	0.719	**0.030**	**0.034**	0.969
C4	"abxaa"	*aa,ab*	0.689	0.159	**0.007**	**0.010**	0.475	**0.006**
C5a	"abxaa"	*aa,ab*	0.081	0.428	**<0.001**	0.763	0.076	0.154
E3xCh5	"nnxnp"x"efxeg"	*nn ee, nn ef, nn eg, nn fg, np ee, np ef, np eg, np fg*	0.838	0.117	0.293	0.829	0.486	0.563
E3xE7	"nnxnp"x"ijxik"	*nn ii, nn ij, nn ik, nn jk, np ii, np ij, np ik, np jk*	0.440	0.943	0.427	0.264	0.610	0.471
Ch5xE7	"efxeg"x"ijxik"	*ee ii, ee ij, ee ik, ee jk, ef ii, ef ij, ef ik, ef jk, eg ii, eg ij, eg ik, eg jk fg ii, fg ij, fg ik, fg jk*	0.658	0.121	0.161	0.399	0.561	0.515
C1xC4	"abxaa"x"abxaa"	*aa aa, aa ab, ab aa, ab ab*	0.930	0.178	0.139	0.566	0.398	0.925
C1xC5a	"abxaa"x"abxaa"	*aa aa, aa ab, ab aa, ab ab*	0.267	0.789	0.269	0.115	0.891	0.744
C4xC5a	"abxaa"x"abxaa"	*aa aa, aa ab, ab aa, ab ab*	0.302	0.701	0.264	0.426	0.075	0.325
E3xC1	"nnxnp"x"abxaa"	*nn aa, nn ab, np aa, np ab*	0.392	0.330	0.620	0.396	0.750	0.906
E3xC4	"nnxnp"x"abxaa"	*nn aa, nn ab, np aa, np ab*	0.373	0.701	0.895	0.909	0.784	0.574
E3xC5a	"nnxnp"x"abxaa"	*nn aa, nn ab, np aa, np ab*	**0.019**	0.657	**0.037**	0.500	0.366	0.922
Ch5xC1	"efxeg"x"abxaa"	*aa ee, aa ef, aa eg, aa fg, ab ee, ab ef, ab eg, ab fg*	0.989	0.805	0.793	0.704	0.397	0.385
Ch5xC4	"efxeg"x"abxaa"	*aa ee, aa ef, aa eg, aa fg, ab ee, ab ef, ab eg, ab fg*	**0.043**	**0.039**	0.470	**0.022**	0.736	0.977
Ch5xC5a	"efxeg"x"abxaa"	*aa ee, aa ef, aa eg, aa fg, ab ee, ab ef, ab eg, ab fg*	0.247	0.479	0.553	0.962	0.203	0.669
E7xC1	"ijxik"x"abxaa"	*aa ii, aa ij, aa ik, aa jk, ab ii, ab ij, ab ik, ab jk*	0.414	0.566	0.959	0.904	0.570	0.948
E7xC4	"ijxik"x"abxaa"	*aa ii, aa ij, aa ik, aa jk, ab ii, ab ij, ab ik, ab jk*	0.601	0.105	**0.020**	0.636	0.080	0.204
E7xC5a	"ijxik"x"abxaa"	*aa ii, aa ij, aa ik, aa jk, ab ii, ab ij, ab ik, ab jk*	0.294	0.475	0.880	0.691	0.265	0.708
Adjusted R^2			31.5	47.7	29.4	13.9	49.5	30.2

[1]Segregation type codes for a CP population according to MapQTL 6 (van Ooijen 2009). [2]Genotype codes for a CP population, depending on the locus segregation type and the QTLxQTL interaction.

QTLxQTL interactions

The most significant QTLs for *prate*, *ipoint* and *onset* in each linkage group and all the pairwise interactions between them were used in a linear model to investigate epistatic interactions, within a year. The full model was based on the joined set of main effect QTLs as detected for any of the three curve characteristics. Therefore the full model was the same for *prate*, *ipoint* and *onset* and it was set according to model (4) including only two-way interactions between QTLs:

y = E3 + Ch5 + E7 + C1 + C4 + C5A+ E3xCh5 + E3xE7 + Ch5xE7 + C1xC4 + C1xC5A + C4xC5A + E3xC1 + E3xC4 + E3xC5A + Ch5xC1 + Ch5xC4 + Ch5xC5A + E7xC1 + E7xC4 + E7xC5A + error

QTL main effects, QTLxQTL interactions and their corresponding genotypic classes for *prate*, *ipoint* and *onset* are presented in Table 2. To go from the full models to the final models, only significant terms in the full models were retained ($p<0.05$, highlighted in bold). The p-values for the retained terms and adjusted R^2 in the final models for each curve characteristic are presented in Table 3. For instance, after fitting the full model for *prate*, 6 terms were retained in the final model (Table 2): two epistatic interactions (*E3xC5a* and *Ch5xC4*), one main effect QTL (*Ch5*) and three non-significant main effect QTLs involved in the epistatic interactions (*E3,C4* and *C5a*). In the final model, only the interaction *Ch5xC4* showed to be significant and the main effect of *Ch5* and *C5a* (Table 3).

In 2004 and 2005, a main effect of *Ch5* on *prate*, *ipoint* and *onset* was observed. It is related with a QTL on the same region previously reported as associated with plant maturity type (Celis-Gamboa 2002), which is a trait reflecting the general development of the plant.

In 2005 no epistatic interactions were found between the QTLs but pleiotropic effects of *E3* on *ipoint* and *onset*, and *C4* on *prate* and *onset* (Table 3).

In 2004, *C4* associated with *onset* of senescence interacted epistatically with *Ch5* associated with *prate* and *ipoint* (Table 3). In the interaction *C4xCh5*, eight genotypic classes are present (‘aa ee’, ‘aa ef’, ‘aa eg’, ‘aa fg’, ‘ab ee’, ‘ab ef’, ‘ab eg’, ‘ab fg’) as shown in Table 2 and the average senescence curves per class are presented in Fig.4.

Table 3. Main effect QTLs and QTLxQTL interactions retained in the final models for *prate*, *ipoint* and *onset* of senescence in 2004 and 2005. Significant p-values (p<0.05) are highlighted in red. Non-significant main effect QTLs are still included in the final model when the QTLs are involved in a significant interaction.

		2004			2005		
Variable	**Possible genotypic classes**	***prate***	***ipoint***	***onset***	***prate***	***ipoint***	***onset***
E3	*nn,np*	0.887	**0.028**	**0.040**		**0.007**	**0.028**
Ch5	*ee,ef,eg,fg*	**<0.001**	**<0.001**	**0.011**	**<0.001**	**<0.001**	**<0.001**
E7	*ii,ij,ik,jk*		**0.019**	**0.013**			**0.019**
C1	*aa,ab*				0.067	**0.016**	
C4	*aa,ab*	0.738	0.157	**0.005**	**0.013**		**0.002**
C5a	*aa,ab*	**0.026**		**<0.001**			
E3x C5a	*nn aa, nn ab, np aa, np ab*	0.066		0.267			
Ch5xC4	*aa ee, aa ef, aa eg, aa fg, ab ee, ab ef, ab eg, ab fg*	**<0.001**	0.109		0.086		
E7xC4	*aa ii, aa ij, aa ik, aa jk, ab ii, ab ij, ab ik, ab jk*			**0.005**			
Adjusted R^2		37.0	44.0	29.9	13.9	43.0	31.7

The genotypic classes including 'ab' (fitted curves with dashed line) had a faster *onset* and senescence development than the classes in which 'aa' (fitted curves with solid line) was present. The curve 'ab ee' showed a higher *prate* than 'aa ee' and the genotypes in this class had in general a faster *onset* of senescence and died earlier. The classes 'aa fg' and 'ab fg' were associated with late senescence development and the *prate* and *ipoint* of these two classes were the lowest. Genotypes in these two classes were at almost half of the maximum senescence when the observation period ended. Genotypes in the class 'aa fg' showed the latest senescence development among all the genotypic classes. Interestingly the class 'ab fg' showed the earliest *onset* but this was followed by a delayed senescence and the genotypes in this group had the latest senescence.

Another epistatic interaction was observed in 2004 between *C4* associated with *onset* and *E7* related with *ipoint* (Table 3). Fig.5 shows the average senescence curves observed for the interaction *C4xE7* with eight genotypic classes ('aa ii', 'aa ij', 'aa ik', 'aa jk', 'ab ii', 'ab ij', 'ab ik', 'ab jk') as shown in Table 2. The classes including 'ab' (fitted curves with dotted line) had in general, a faster senescence development than classes in which 'aa' (solid line) was present except for the class 'aa ik'. In this class earlier *onset* of the senescence process was observed with a faster *prate* and *ipoint* while in the group 'ab ik' a delayed *onset* with a

slower *prate* is present. Interestingly, big differences were also observed between the classes 'aa jk' and 'ab jk'. The class 'aa jk' showed the same trajectory than the class 'aa ii' and both of them have late *onset* and *ipoint* making the senescence process slower than in the other classes. On the other hand, the class 'ab jk' showed the earliest *onset* with a delayed senescence that ended up simultaneously with the classes exhibiting the slowest *prate* ('aa ii' and 'aa jk').

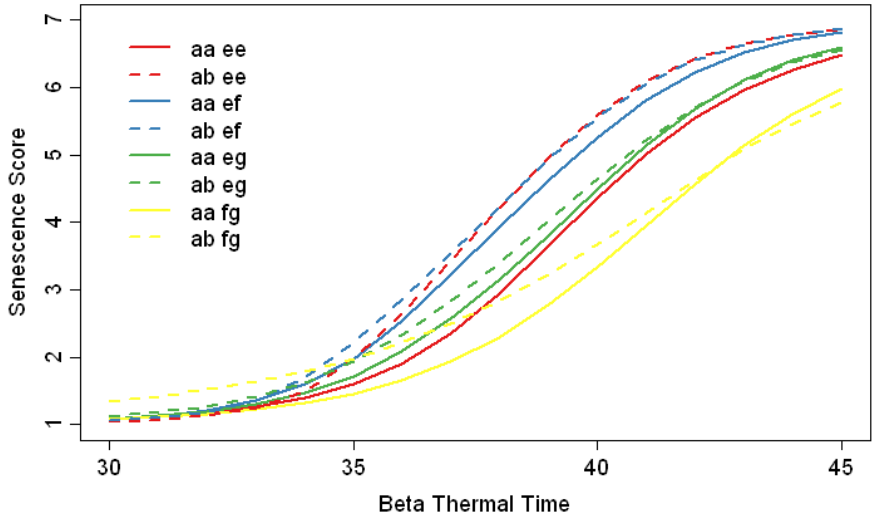


Figure 4. Average senescence curves for each genotypic class in the epistatic interaction between QTLs *C4* and *Ch5* observed in 2004.

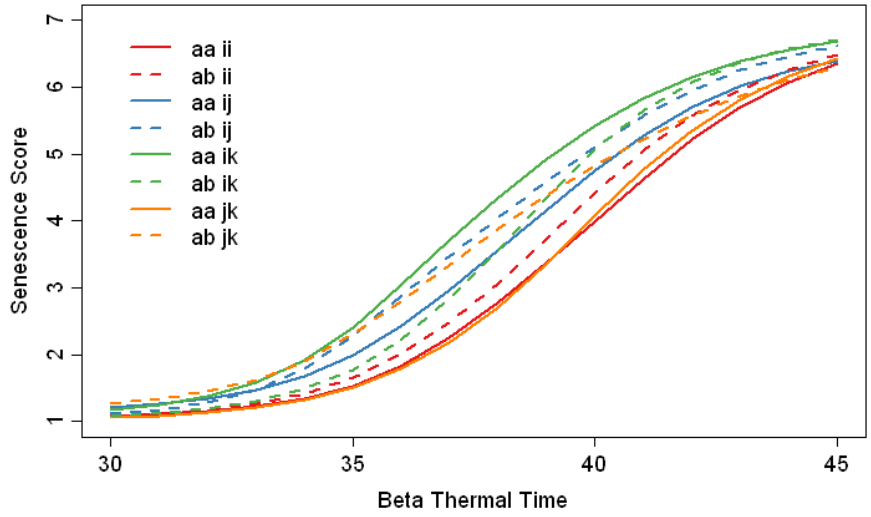


Figure 5. Average senescence curves for each genotypic class in the epistatic interaction between QTLs *C4* and *E7* observed in 2004.

Discussion

The smoothed GLM procedure in our study allows flexible modelling of different senescence curves in the mapping population. Genotypes with early development showed an s-shape senescence curve, whereas genotypes with slow development exhibited an exponential curve during the observation period. In both cases the curve trajectories would eventually reach the maximum of 7, which is the highest senescence score. In the case of late genotypes, the observation period was not long enough to complete the senescence process. Four characteristics estimated from the curves (*onset, prate, mprate* and *ipoint*) allowed us to explain senescence in terms of development, whereas individual time points only gave us an impression of the trait at selected different moments during the growing season. In the first case sensible biological interpretations could be made and the results were compared with previous studies describing the senescence process in the same population evaluated in the Netherlands (Celis-Gamboa 2002).

The weather and the general conditions of the experiment had a clear effect on the performance of the population in the two years of field experimentation in Finland. Clearly the senescence period lasted longer in 2005 and showed a slower progression rate. This was probably due to the higher daily air temperature observed that year. Marinus and Bodlaender, 1975 reported that high temperatures accelerate the reproductive and vegetative development of potato plants and delay the senescence process. Our results are therefore in line with their findings.

Comparing the QTL analysis using the individual time points and the curve characteristics, QTLs detected at early time points were associated with *onset* of senescence. QTLs occurring in intermediate to late time points coincided with QTLs associated with *prate* and *ipoint.* Although the results are comparable with both phenotypic data, the analysis using the characteristics of fitted curves offers a biological interpretation of the results in terms of senescence development. In our study, four curve characteristics described the whole process and allowed us to study the progression of the trait. When the CxE population was evaluated in the Netherlands, QTLs on chromosome 5 were identified for different traits related with senescence and life cycle. Malosetti et al. 2006 reported two QTLs on chromosome 5 associated with mid-senescence and rate of senescence. Celis-Gamboa et al.

2003 identified three QTLs on the same chromosome associated with the duration of the plant cycle and one of them was found associated to plant maturity type explaining 40% of the phenotypic variance. In this study, we detected QTLs associated with senescence development on chromosome 5. These are in the same region where the latter QTL was reported by Celis-Gamboa 2002, suggesting a pleiotropic effect of this region for *prate*, *ipoint, onset* and plant maturity.

An interesting QTL associated with *onset* was found on *C4*, whereas no QTLs have been previously reported for developmental traits on this chromosome. A wide range of functional genes involved in resistance against various pathogens have been reported on chromosome 4, among others a major late blight resistance gene cluster (Celebi-Toprak et al. 2002; Li et al. 1998; Park et al. 2005; Zimnoch-Guzowska et al. 2000).

Epistatic interactions were observed in 2004 between *C4* and *Ch5* (Fig.4) and between *C4* and *E7*(Fig.5). Interestingly, in both epistatic interactions, the genotypic class 'ab' in *C4* was associated with early development of the senescence process. In the interaction *C4*x*Ch5*, the genotypic classes 'ab ee' and 'ab ef' associated with early senescence showed delayed *onset* but fast *prate* and *ipoint* whereas the class 'ab fg" showed early *onset* but a senescence period lasting longer. In 2005 no epistatic interactions were observed but main effects of *C1* and pleiotropic effects of *E3* and *C4* on *ipoint/onset* and *prate/onset* respectively.

In summary, by modelling haulm senescence in the CxE potato population using a smoothed GLM procedure we were able to characterize the curve in terms of developmental traits and to identify QTLs associated with the senescence process. Pleiotropic effects and epistatic interactions between QTLs were detected when two-way interactions were studied. Delayed senescence was associated with particular genotypic classes (*C4xCh5*: 'ab fg' and *C4xE7*: 'aa ii' 'aa jk') and some classes showing early *onset* turned to have delayed senescence development (*C4xCh5*:'ab fg' and *C4xE7*: 'ab jk').

Different regulatory genes have been reported for onset and progression rate of senescence in *Arabidopsis thaliana* (Gepstein et al. 2003) and it will be interesting to compare the genes

involve in the same process in different crops. The complexity of QTLxE and QTLxQTLxE as well as the performance of genotypes under different environments will be considered in a further study providing insights for the better understanding of adaptation and developmental processes in potato.

Acknowledgements

We thank Prof. Paul Struik (Centre for Crop Systems Analysis, Wageningen University) for the discussion on the use of beta thermal time in this study.

Chapter 4

Understanding the genetic basis of potato development using a multi-trait QTL analysis

P.X Hurtado-Lopez[1,2,5], B. Bekele-Tessema[1,6], S.K. Schnabel[2,3], C. Maliepaard[1], C.G. Van der Linden[1], P.H.C. Eilers[2,4], J. Jansen[2,3], F.A van Eeuwijk[2,3] & R.G.F. Visser[1,3]

[1]Wageningen UR Plant Breeding, Wageningen University & Research Centre, Wageningen, The Netherlands
[2]Biometris, Wageningen University & Research Centre, Wageningen, The Netherlands
[3]Centre for BioSystems Genomics, Wageningen, The Netherlands.
[4]Erasmus University Medical Center. Rotterdam, The Netherlands
[5]C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC). Wageningen University, Wageningen, The Netherlands
[6]The graduate school Experimental Plant Sciences (EPS). Wageningen University, Wageningen, The Netherlands

To be submitted

Abstract

Background

The ability to understand the genetic basis of plant development depends on a proper characterization of developmental morphology over time. Developmental processes can be described in terms of parameters related to different aging stages. When time-dependent parameters are incorporated in QTL mapping procedures the genetic control of the different stages is better understood. Because development involves related events, it is necessary to study developmental traits simultaneously. QTL analyses combining these traits not only increase the power of QTL detection, but also improve the precision of QTL map position especially for co-localized QTLs related to different traits. Multi-trait QTL analysis makes it possible to study relationship between traits and to detect the presence of pleiotropic regions affecting several processes simultaneously. Multi-trait QTL analysis is of major importance for the development of breeding strategies targeting stable pleiotropic QTLs affecting complex traits.

Results

The growth and development parameters used in our study facilitated the understanding of complex developmental traits as a continuous and dynamic process in time. The combined use of developmental parameters and agronomic characteristics in a multi-trait QTL analysis provided insight in (1) the genetic architecture of complex traits in potato, (2) the presence of time-dependent QTLs and (3) the existence of pleiotropic regions related to above and below ground traits. The genetic control of developmental traits such as flowering, senescence and plant height was explained by 9, 10 and 12 QTLs, respectively. Some of these were expressed permanently throughout the entire development process, while others were time-dependent and associated with specific development stages. Emergence, number of main stems, number of tubers and yield were explained by 9, 5, 4 and 6 QTLs, respectively. These traits were measured only once during the growing season, therefore time-dependent QTLs could not be detected. A genetic relationship between above and below ground traits in potato was observed through 14 pleiotropic QTLs associated with developmental and agronomic traits.

Conclusions

In our study we found time-dependent QTLs linked to different developmental stages of potato. The genetic relationship between above and below ground traits was observed through 14 pleiotropic QTLs associated with plant development and agronomic traits. Further research will help to confirm the stability of the pleiotropic QTLs found in our study across environments. Some of our results suggest the presence of QTLxE interactions. Therefore additional studies comparing development under different photoperiods are necessary to investigate the plasticity of the crop. The anchoring of stable pleiotropic QTLs to the annotated potato genome sequence will provide target genes for marker assisted breeding and candidate gene approaches.

Keywords

Multi-trait QTL analysis, Plant development, Pleiotropy, Potato

Background

The development of plants is a complex, dynamic process controlled by networks of genes and environmental factors. As a consequence, QTL analysis of traits related to plant development requires the use of advanced statistical-genetic models and methods (Atchley 1984; Wolf et al. 2001). Conventional QTL mapping strategies neglect the fact that traits related to plant development are changing in time. For example, in potato plants height and tuber size change in time, and their development is influenced by changing environmental factors. Therefore, such traits should be represented as functions of time and/or optional variables describing the major changes in environmental factors over time. This requires an approach that is able to detect time-dependent genetic effects.

In *Arabidopsis* molecular markers have been associated with phenotypes observed at different development stages and the differences between these stages have been compared (Mauricio 2005). In the same model plant, simulated time series data have been used to infer growth curves in order to study the quantitative nature of plant development (Mundermann et al. 2005). A more general strategy to study the genetic architecture of complex, dynamic traits, so-called functional mapping, has been proposed to integrate time-dependent traits into QTL mapping (Lin and Wu 2006; Wu and Lin 2006a; Wu et al. 2003).

Dissecting the genetic basis of plant development requires an accurate description of developmental morphology. Such descriptions are often lacking and conclusions are drawn based on observations of fully grown plants (Kellogg 2004). This means that comparisons between developmental phases are often superficial. Therefore, a proper characterization of development over time is needed to describe each part of the process.

In potato, previous studies have incorporated well characterised time series data into growth models and QTL analysis. This approach allowed for a genetic description of senescence in terms of parameters related to different aging stages (Hurtado et al. 2012b; Malosetti et al. 2006). To our knowledge, studies in potato including the time-dependent nature of plant development into QTL analysis have not been reported. Therefore, the genetic control of different developmental stages is still poorly understood.

Although many QTL studies considered multiple traits, usually those traits were analysed separately. An integrated analysis combining traits related to developmental processes simultaneously is required to get a better understanding of the genetic and environmental forces driving plant development. QTL analysis combining data from multiple traits related to plant development will not only increase the power of QTL detection, it will also improve the understanding of the genetic control of developmental processes. As a consequence, a multi-trait QTL analysis of a single population allows the detection of closely linked chromosomal regions affecting several traits simultaneously (Jiang and Zeng 1995). Although different methodologies have been proposed not only to map multiple trait simultaneously(Jiang and Zeng 1995; Knott and Haley 2000; Malosetti et al. 2008) but also to differentiate between close linkage and pleiotropy of coincident QTLs (Jiang and Zeng 1995; Knott and Haley 2000; Lebreton et al. 1998; Liu et al. 2007), the identification of pleiotropic genes requires additional genomic information such as high density maps and genome sequence information.

A first attempt to estimate the optimal set of consensus QTLs for several traits simultaneously in potato was done through a QTL meta-analysis (Danan et al. 2011). It permitted the co-localization of late blight resistance and plant maturity traits by projecting individual QTLs onto a consensus map. However, there are no reports of such integrative

analysis for developmental traits in potato. So far, data on traits related to plant development in potato have not been integrated in a single study to get insight into the genetic architecture of crop development and the presence of putative pleiotropic QTLs related to plant development.

The aim of this study was to identify the genetic basis of plant developmental processes in potato by means of a multi-trait QTL analysis combining several traits describing plant development in time. A total of 23 traits related to plant developmental and agronomic value were incorporated in the multi-trait QTL analysis. For this purpose, a diploid potato mapping population was evaluated under field conditions. The time dependent traits plant height, flowering and senescence were assessed on a weekly basis. The agronomic traits yield, number of main stems and number of tubers were measured at harvest. We were interested in the presence and genetic positions of putative pleiotropic regions associated with plant development and traits of agronomic value. 14 Pleiotropic QTLs were detected in our study, providing insights of the genetic architecture of developmental processes and the genetic relationship between above and below ground traits in potato. The anchoring of putative pleiotropic QTLs to the annotated potato genome sequence (Potato Genome Sequencing et al. 2011) will provide target genes for marker assisted breeding and candidate gene approaches.

Materials and methods

Plant materials

Potato development was assessed in the diploid backcross population CxE. It was obtained from the cross between clone C (US-W5337.3 (Hanneman and Peloquin 1967)) which is a hybrid between *Solanum phureja* (PI225696) and a dihaploid *S. tuberosum* (US-W42) and clone E which is a hybrid between $VH^3$4211 (a *S. vernei* - *S. tuberosum* backcross) and clone C. The CxE population was developed for research purposes (Jacobs et al. 1995) based on the genetic background of the parents. It is known for its segregation of agronomic and quality traits (Celis-Gamboa 2002; Kloosterman et al. 2010). *S.tuberosum* and *S. phureja* have different day length requirements for tuberization making the CxE population suitable for the study of developmental processes influenced by photoperiod and other environmental conditions. 169 genotypes of the CxE population, parents C and E and a selected group of 9

European (Astarte, Bintje, Gloria, Granola, Karnico, Mondial, Première, Saturna and Desiree) and 10 Ethiopian cultivars (Awash, Belete, Bulle, Gera, Gorebella, Guassa, Gudene, Jalene, Shenkolla, and Zengena) cultivars were used in the experiment.

Experimental setup

The CxE population was planted in light clay soil under rain fed conditions on July16 2010 at Holetta Agricultural Research Center, Ethiopia (9.07'N, 38.03'E in west Ethiopia at an altitude of 2400m). Planting was done by hand, with a spacing of 75cm between and 30 cm within rows. Fertilizer (165kg UREA and 196 kg diammonium phosphate per hectare) was applied during planting and a fungicide (RidomilGold) was sprayed against late blight. Ridging of the field was done three times throughout the experiment and weeding was done by hand whenever necessary. The experiment was laid out in a randomized complete block design with three replicated blocks laid against the slope of the field. In each block, the CxE progeny, the European and Ethiopian varieties as well as the two parents were randomized over 190 plots, with 4 plants per plot. The observation period of the developmental traits was 5 months (between July and December 2010) and meteorological data were obtained during this period from the meteorological service present at the research station. The air temperature was recorded daily, every three hours, day and night. Over the whole observation period, the temperature fluctuated between 4 and 23°C between 6:00am to 6:00pm and during the night between 2 and 20°C. During the experiment the day length was 12 hours.

Agronomic traits

During the growing period for each plant the development was assessed by measuring above and below ground traits. Above ground, the date of emergence and the number of main stems were assessed once, while plant height, flowering and senescence were measured over time at regular intervals. Below ground, number of tubers and yield were assessed after the final harvest.

The evaluation of flowering and senescence was done using a scale from 0 to 7 and 1 to 7 respectively, as it has been previously described (Celis-Gamboa et al. 2003). Flowering was recorded 17 times with intervals of 2-6 days at 38, 40, 42, 45, 47, 49, 52, 54, 56, 59, 61, 63,

66, 68, 70, 74, 80, 83, 87, 89 and 95 days after planting, DAP. Senescence was assessed 16 times with intervals of 3-7 days at (80, 83, 87, 91, 95, 99, 103, 107, 111, 115, 119, 123, 129 and 136 DAP. Plant height was measured using the longest stem of each plant as the distance from ground level to main apex. The assessment was done at nine occasions with intervals of 6 days (26, 32, 38, 44, 50, 56, 62, 68 and 74 DAP). All plots were harvested at 138 DAP and the tubers of each plant were counted and weighed.

Conversion of days after planting into thermal days

Average daily air temperature was used in a non-linear temperature effect function (Yin et al. 1995b) to transform DAP into beta-thermal time (BTT) units using the base, optimum and ceiling temperature for phenological development of potato (Hurtado et al. 2012b). Day length was incorporated into the function as a constant anticipating on a later comparison of the performance of the CxE population under different day length conditions. The new thermal unit is the cumulative beta-thermal days after planting combining, temperature, time and photoperiod (photo-beta thermal time, PBTT). This scale was used as the x-axis to analyse the time series data of plant height, flowering and senescence.

Curve fitting and characterization of the curves

Curve fitting of plant height, flowering and senescence was done using PBTT units on the x-axis. For modelling flowering and senescence we used a methodology previously described to fit senescence data in potato (Hurtado et al. 2012b). A smooth generalized linear model was used to estimate smooth curves for the development of flowering and senescence over time. The estimation was done using the R software environment (CoreTeam 2011). A different approach was used to model plant height. In contrast to flowering and senescence, plant height was measured as a continuous variable (in cm). Up to twelve observations per genotype were available per time point. We pooled the 12 observations per genotype in each time point and fitted a curve to the relationship between plant height and time. A smooth expectile curve was well suited for this purpose and the expectiles were estimated using least asymmetrically weighted squares (Schnabel and Eilers 2009). They were combined with P-splines to provide a flexible functional form(Schnabel et al. 2012). This modelling procedure resulted in a smooth frontier curve to describe the development of

plant height over time. For the calculations we used the package “expectreg” in R (Sobotka et al. 2012).

Once the developmental curves were fitted, parameters describing the aging process were estimated. Those parameters facilitated the study of development as a continuous time-dependent process by breaking down the complex traits into components related to the different developmental stages. The first and second derivative of the fitted curves have been used to characterise senescence processes under long day length conditions (Hurtado et al. 2012b). The parameters used to characterised senescence were also used in our study to describe plant height, flowering and senescence under short photoperiod (Figure 1). Those parameters are mean and maximum progression rates (average and maximum speed of the process), inflection point or the turning point at which the processes enter into the final phase, onset and end. We also considered additional traits describing growth and development, such as maximum and mean plant height, duration of flowering and maximum progression rate for onset of plant height (maximum speed of the process between emergence and the first observation of plant height). Note that the parameters have different units and their interpretation is different. For instance, small values of progression rate indicate slow flowering, senescence or plant height processes, mainly associated to late genotypes; while small values of inflection point, onset or end are related to early genotypes.

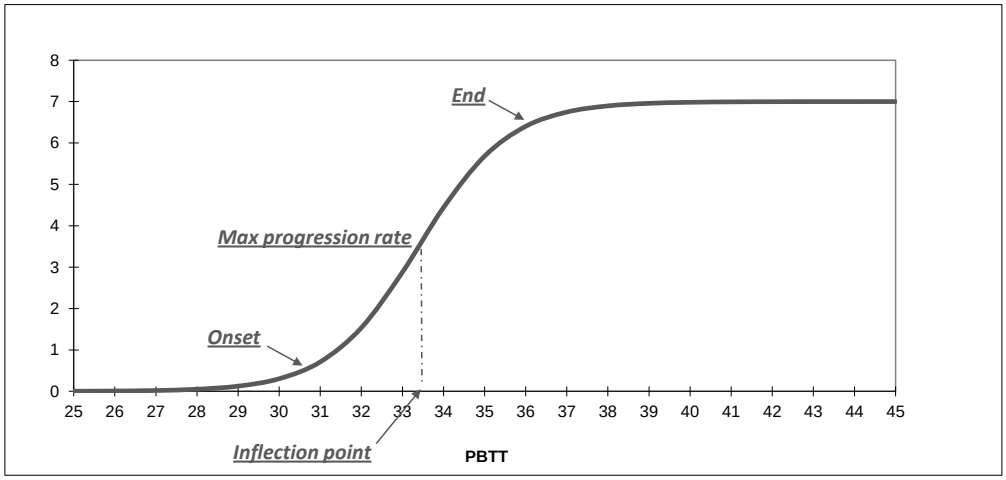


Figure 1. Fitted curve for flowering development of a random genotype of the CxE population. It is used as example to show the parameters describing flowering, senescence and plant height. On the x-axis: photo-beta thermal time (PBTT), on the y-axis: flowering on a scale from 0 to 7.

Genetic maps and molecular data

Single nucleotide polymorphism (SNP) markers scored in a core set of the CxE population (Anithakumari et al. 2010) were added to the separate CxE parental maps described in Hurtado et al, 2012 (Hurtado et al. 2012b). Together with the SNP markers, AFLP, SSR and CAPS with the expected segregation ratios for dominant and co-dominant markers (1:1 and 1:1:1:1, respectively) were used to construct more saturated maternal (C) and paternal (E) maps (additional file 1). JoinMap 4 (van Ooijen 2009) was used to map 521 and 560 markers in the C and E maps, respectively, with 12 linkage groups (LG) each as previously reported (Celis-Gamboa 2002).

Considering the expected differences in the recombination frequencies between the two parents (due to their genetic background from two different *Solanum* species), the C and E maps were not integrated. The data were analysed for each parental meiosis separately in an integrated two-way pseudo-testcross analysis (Grattapaglia and Sederoff 1994). Dominant and codominant markers segregating 1:1 and 1:1:1:1, respectively, were used in the QTL analysis and the latter ones were converted into 1:1 type to separate the meiosis of the two parents.

QTL analysis

The parental maps were combined in a single dataset where the map of the second parent was appended to the map of the first one. It allowed the use of cofactor markers of one parent while searching for QTLs in the other parent and vice versa gaining power to detect QTLs.

Two types of phenotypic traits were considered in our study (Table 1): parameters derived from the fitted curves and characteristics measured once during the growing season. The curve parameters were treated as phenotypic traits with a single value per genotype. The data were unbalanced for the traits measured once and the best linear unbiased estimators (BLUEs) were obtained from a mixed model analysis of the field trial, in which genotype was fitted as a fixed term. The BLUEs, having a single value per genotype, were the input for the QTL analysis. The QTL library of Genstat 14 (VSN 2011) was used for the multi-trait QTL analysis.

Table 1. Phenotypic traits included in the multi-trait QTL analysis, trait units and described developmental processes

	Description		
Trait type	**Traits**	**Units**	**Developmental processes**
Parameters derived from fitted curves	Onset	Thermal days	Flowering, senescence, plant height
	Maximum progression rate		Flowering, senescence, plant height
	Inflection point	Thermal days	Flowering, senescence, plant height
	End	Thermal days	Flowering, senescence, plant height
	Mean progression rate		Flowering, senescence, plant height
	Maximum progression rate in onset		Plant height
Characteristics measured once	Duration of flowering	Days after planting	Flowering
	Maximum height	cm	Plant height
	Mean height	cm	Plant height
	Emergence	Days after planting	
	Number of main stems	Number	
	Total number of tubers	Number	
	Yield	Kg	

Multi-trait linkage analysis

The different stages of potato development were studied using the fitted curves for flowering, senescence and plant height. Each developmental trait was broken down in time-related parameters derived from the curves and they were considered as new phenotypic traits. Those parameters together with agronomic traits were used in a multi-trait QTL analysis including 23 traits: 5 common traits for the three developmental processes (onset, maximum progression rate, inflection point, end and mean progression rate), one additional trait describing flowering (duration of flowering), three additional traits related to plant height (maximum progression of onset, maximum and mean height) and four agronomic traits (emergence, number of main stems, total number of tubers and yield). All the traits were auto-scaled (subtracting the average and dividing by the standard deviation) to make traits with different scales and units better comparable for the multi-trait analysis. The best variance-covariance model was selected based on the Bayesian Information Criterion (BIC) and it was used in the multi-trait QTL analysis for the selection and fitting of the QTL model (Malosetti et al. 2008), (Boer et al. 2007). An initial genome-wide scan to identify candidate QTL positions was done by SIM and one round of CIM was performed to scan for QTL effects in the presence of co-factors genome-wide. A REML (residual maximum likelihood) procedure was used iteratively to fit the final QTL model at each chromosome position, storing the Wald statistic for hypothesis testing. The associated probability value (on a minus log10 scale) was plotted to visualize the results along the chromosomes. The

threshold for QTL detection was calculated using the Li and Ji method (Li and Ji 2005) with a genome-wide significance level of 0.05. A final QTL model was fitted to estimate QTL effects, assuming traits and QTLs as fixed terms and genotypes as random term.

Results

Curve fitting and characteristics of the curves

Curves describing development over time were fitted to the data of the individuals of the CxE population, their parents and the control varieties. Differences in curve trajectories were observed between early and late genotypes for flowering, senescence and plant height (Figure 2). The maturity type of the CxE population was previously assessed under field conditions (Celis-Gamboa 2002) and it is used as reference in the present study. Early genotypes completed their life cycle faster and a complete s-shaped curve could be observed. Late genotypes showed slow progression of the developmental traits and some of them did not even complete the flowering and aging processes during the observation period. In that case, only the first part of the s shape could be observed.

In the CxE population there was a direct relationship between growth and maturity. Most of the late genotypes were tall and the early genotypes were short. However there was not a complete relationship between plant height and maturity respect to the Dutch cultivars. For instance Dutch varieties, irrespective of their maturity type, showed fast progression of senescence and all of them were shorter than the Ethiopian cultivars. It indicates that under short day conditions maturation was accelerated while growth was restricted in these varieties. In addition, flowering curves could not be fitted for the Dutch varieties due to the absence of flowers or flower abortion observed in all of them. Thus, the reduction in photoperiod affected dramatically the Dutch varieties which are adapted to long day length. Suppressed flower development was also observed in previous potato studies in growth chambers when the irradiance was reduced (Clarke and Lombard 1939; Turner and Ewing 1988). In all the CxE genotypes flowering and senescence curves presented parallel trajectories and they overlapped in early genotypes at the final stage of both processes. Although some common genetic factors could be involved in both processes, genes controlling individual developmental stages are also expected. It has been reported that

some QTLs are expressed at early developmental stages and they are switched off after a particular age (Wu and Lin 2006a). Time-dependent QTLs have been observed in potato, controlling for instance onset and progression rate of senescence under long day length conditions (Hurtado et al. 2012b).

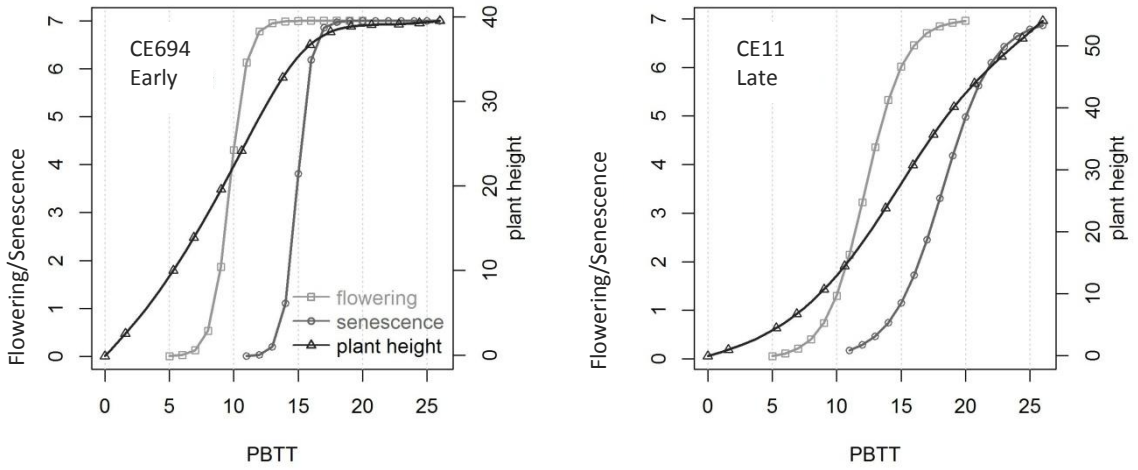


Figure 2. Fitted curves for plant height, flowering and senescence of two genotypes representing early and late maturing groups. On the x-axis: PBTT (Photo-beta thermal time) units combining average daily air temperature and photoperiod. On the y-axis: flowering and senescence scales from 0 to 7 (left side) and plant height in cm on a continuous scale (right side)

Genetics of complex traits

The genetic architecture of complex developmental traits in potato was studied using the parameters derived from the fitted curves for flowering, senescence and plant height. Together with the agronomic traits they were included in a multi-trait QTL analysis and the QTLs detected with the maternal and paternal maps could be observed in Figure 2. Although our study was mainly focused on the presence and positions of QTLs (upper plot of Figure 3) rather than on the allelic effects (lower plot), the QTL effects (positive: red; negative: blue) related to different values of the phenotypic traits, are also reported for the 23 traits on each QTL position. The size of QTL effects, indicated by the intensity of the colour (the darker the larger the effect), is also shown in Figure 2 and the explained variance for each trait is provided in Table 2. It is important to notice that the allele effects in each parent are related to the different sets of allele pairs of the grandparents. Furthermore, the order of the grandparental alleles on each parent may be switched between linkage groups and therefore the colours can be arbitrary. Only within the same linkage group the allele effect can be attributed to the same grandparental difference between the alleles. In that sense,

within a pleiotropic region the allele effects can be compared and they can go in opposite directions as it was observed in all linkage groups were a QTL was detected except for linkage groups E6 and E11. Opposite effects within a pleiotropic region are expected for a QTL related to negatively correlated traits. For instance, progression of flowering is negatively correlated to end of flowering (Additional file 2) and opposite QTL effect on C5 and E5 were observed for both traits. Plants with fast flowering development (high values for progression rate) are expected to have an early end of the flowering process (small values for end of flowering).

Complex traits

For each complex trait multiple QTLs were identified as it was expected due to the quantitative inheritance of developmental and agronomic traits (Figure 3). Some of the QTLs were time-dependent and they were associated to different stages of development. We checked the position of the QTLs on the parental maps and the QTLs detected on a particular linkage group were different from the QTLs detected on the homologous linkage group in the other parent. Only 1 QTL on the same genetic region on C5 and E5 was detected (second QTL observed on C5 in the right side of the linkage group). This was a major QTL associated to all developmental and agronomic traits (except emergence) and the E parent was showing a main effect on all of them with a –Log10(p) going up to 50. It also showed the highest explained variance in most of the traits included in our study (Table 2) going up to 60% for onset of senescence. This finding is in agreement with previous reports indicating a major effect of a QTL in the same chromosomal region, which was associated to plant maturity and could have a pleiotropic effect on many developmental traits (Celis-Gamboa 2002; Hurtado et al. 2012b; Malosetti et al. 2006). According to our results there is not a major contribution of this QTL to the agronomic traits as it is shown by the low explained variances. Since our study is focused on new QTLs contributing to the understanding of the genetic architecture of complex traits, we have limited our discussion and main conclusions to those QTLs.

Flowering

In our study the genetic control of flowering was driven by 9 QTLs, one found throughout the entire process (E5) and the others found according to the development stage. QTLs on

C2, E1, E3 and E8 were expressed at early developmental stages, including up to half way of the process (inflection point) and the maximum speed reached at that point. Those QTLs seemed to be switched off at later stages when other QTLs were expressed. The late QTLs (C10 and the first QTL on C5), silent during early stages, were expressed in the final parts of flowering and they were associated with the length of the flowering period.

Senescence

Ten QTLs controlled the aging process in our study. QTLs on E1, E8 and E12 were related to early onset of senescence and they seem to be switched off when half of the process was reached. At that point, late QTLs on C3, C4 and E6 seem to be switched on and these QTLs were associated with the end of senescence.

Plant height

We found 12 QTLs related to plant height. QTLs permanently expressed during the growing process were identified on C2, first half of C5, E5 and E12. Early QTLs on C1, C3 and C4 were expressed between onset and half of the growing process and they were also associated with the average and maximum plant height. The presence of common QTLs for those traits could also be explained by the high phenotypic correlation between them (Additional file 2).

Agronomic traits

Emergence, number of main stems, total number of tubers and yield were explained by 9, 5, 4 and 6 QTLs respectively. Those traits were measured once during the growing season, therefore time-dependent QTLs could not be detected for any of them. Some QTLs have been reported for yield on Chromosomes 1 and 6 in a tetraploid potato full-sib family (Bradshaw et al. 2008). In our study QTLs on C1 and E1 explained 11% of the yield phenotypic variance suggesting the presence of a stable genomic region on chromosome 1 for yield in potato. Although there was an effect of chromosome 5 on the agronomic traits, it was smaller compared with the effect on developmental traits, except for yield (Table 2). These results suggest that plant maturity does not have a central role on the agronomic traits considered in our study.

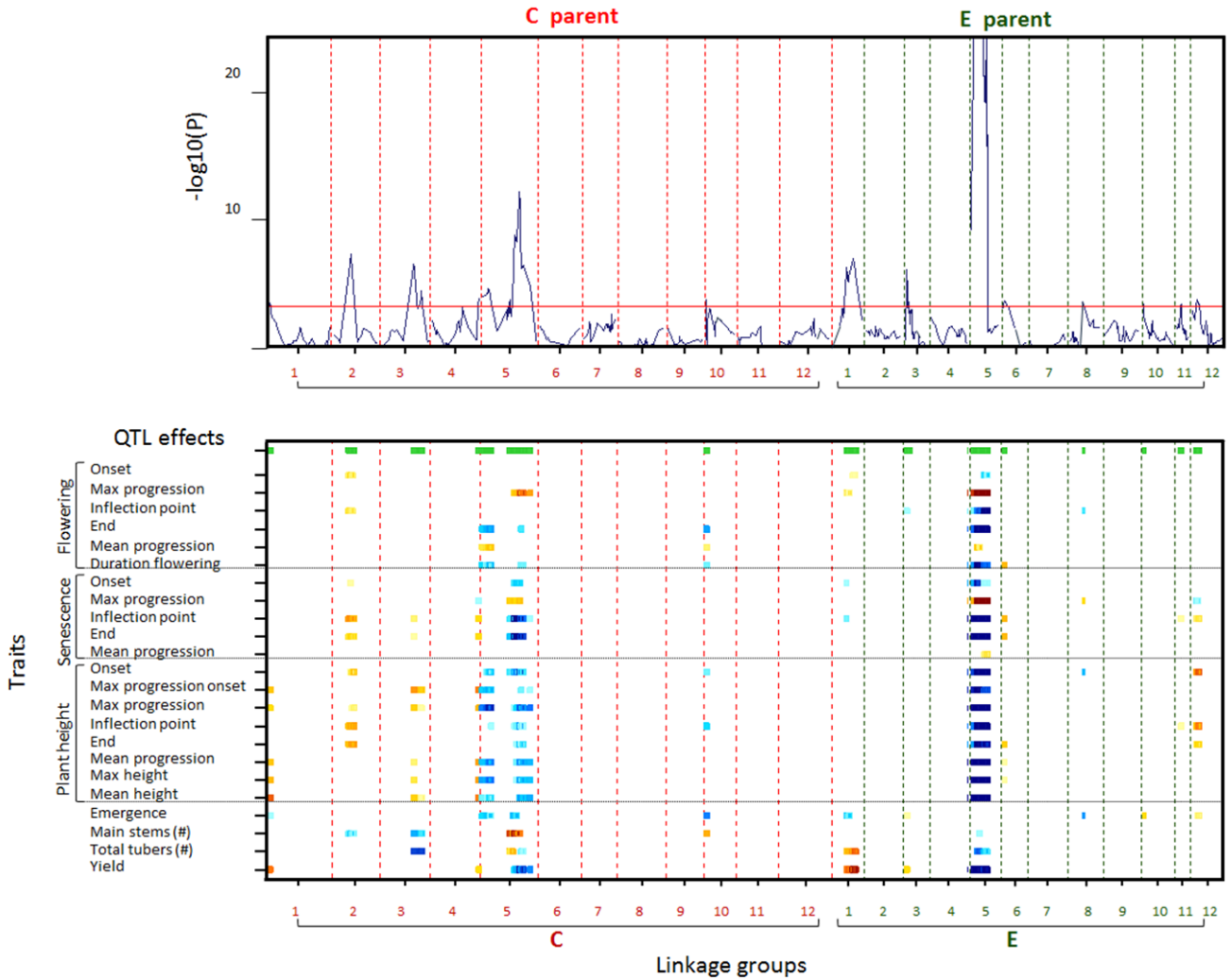


Figure 3. Multi-trait QTL linkage analysis. The upper plot shows the significance of QTLs (-log10 scale for the associated probability value). The lower plot shows the positive (red) and negative (blue) allele substitution effects at positions where there was a significant QTL. The intensity of the colour is proportional to the QTL effect size (the darker the larger the effect). Only –log10(p) values lower than 25 are presented in the figure.

Table 2. Phenotypic variance explained by each QTL associated to developmental or agronomic traits in the multi-trait QTL analysis. The explained variance is given in percentage.

		Linkage group (cM)													
Trait	**Parameters**	**C1: 7.7**	**C2: 45.1**	**C3: 84.8**	**C4: 121.0**	**C5: 15.3**	**C5: 97.3**	**C10: 2.5**	**E1: 35.1**	**E3: 3.1**	**E5: 18.8**	**E6: 4.3**	**E8: 0.0**	**E10: 0.0**	**E11: 13.0**	**E12: 13.5**
Flowering	Onset		4.6						2.4		1					
	Max progression						7.7		1		44.6					
	Inflection point		3.4							3	8.6		3.6			
	End					5.1	1.3	5.6			28.1					
	Mean progression					10.2		8.7			13.7					
	Duration flowering					3.8	1	3.8			39.7	7.4				
Senescence	Onset		9.3				18.8		2		60.2					
	Max progression				3.2		5.3				27.5		3.3			2.5
	Inflection point		4.5	1.5	2		13.6		1		51	2.7			2.7	1.8
	End		2	1.2	2.1		16.3				53.1	2.4				
	Mean progression										1					
Plant height	Onset		1.4			3.3	6.9	1.7			31.6		3.2			5.2
	Max progression onset	2.8		5.2	5.2	4.2	1.7				22.3					
	Max progression	2.1	1	2.3	1.9	6.4	5.4				42.6					
	Inflection point		4			2	4.5	1.9			32				2.6	4.9
	End		4.9				5.9				20.7	4.2				5
	Mean progression	3.3		1.6	2.5	3.6	4.3				49.2	1.5				
	Max height	3.7		2	2.7	4.2	3.5				47.1	1.3				
	Mean height	4.2		2.1	3.6	1.7	3.1				49					
Agronomic traits	Emergence	1.4				5.1	1	6.7	1	2	1.4		4.9	2.3		2.8
	Main stems		3.4	5.9			6.7	4.6			1.6					
	Total tubers			9.9			1		6.9		8.3					
	Yield	3.1			2.1		10.1		8.1	5	29.6					

Pleiotropic regions

The multi-trait QTL analysis combining developmental and agronomic traits not only increased the power of QTL detection, compared with single trait linkage analysis (Additional file 3), but it also helped us to detect pleiotropic regions controlling above and below ground traits in potato.

14 pleiotropic QTLs associated with developmental and agronomic traits were identified in our study and different QTL effects were observed in each one. In the C parent, 7 pleiotropic QTLs were identified and all of them showed QTL effects going in opposite directions. For instance, the QTL on C2 was related to genotypes with late onset of plant height, flowering and senescence, fast progression of the three processes and few main stems. In C3 the QTL was related to tall plants with fast growth and low number of tubers and main stems. In fact, previous studies have shown that tuber formation could be reduced when the development of the haulm is accelerated (Maris 1964). A positive correlation between number of main stems and number of tubers has also been reported (Lemaga and Caesar 1990) but the genetic control of these traits was not yet clear. Here, we are reporting a QTL on C3 for both traits explaining 6 and 10% of the phenotypic variance main stems and total number of tubers, respectively. The QTL on C10 was associated with early emergence, early onset of growth, short duration of flowering and more than 4 main stems per plant. It could be interesting to investigate if the presence of this QTL is also related to yield in other environments because a positive correlation between number of main stems and yield has been reported in a previous study (Lemaga and Caesar 1990). This QTL could facilitate the selection of high yielding varieties with fast growth and a short flowering period.

In the E parent, we detected 1 QTL associated to late emergence on E10, 5 pleiotropic QTLs on E1, E3, E5, E8 and E12 with opposite effects on the associated traits and 2 QTLs with allele effects going in the same direction on E6 and E11. The QTL on E1 was associated with genotypes having early emergence, early onset of senescence, high number of tubers and high yield, showing the highest explained variance for yield and number of tubers (8.1 and 6.9%, respectively). The QTL on E8 was associated with early emergence, early onset of growth and fast senescence, while the QTL on E12 show opposite effect on the same traits. On E6 and E11 we detected QTLs with similar effect on senescence and plant height. It was

associated with long growth and senescence processes but without any effect on agronomic traits. Further research will help to confirm the stability of the pleiotropic regions associated with developmental traits found in our study and to investigate the presence of one or more genes in those regions.

Discussion

The curve fitting approaches followed in our study provided an effective characterization of the developmental processes that occur during the potato life cycle under short day length conditions. The parameters derived from the curves characterise different stages of the development of the above ground parts of the plant. Plant height, flowering and senescence are described by five parameters: onset, end, progression rate (average and maximum speed of the process) and inflection point (time point when half of the developmental process has been reached) These parameters can also be used to characterise other processes in which growth curves are fitted using discrete or continuous data collected as a time series. For some traits additional characteristics were taken into account, such as duration of flowering or maximum plant height and they were directly calculated from the data. We also considered an additional trait for plant height (progression rate between emergence and the first observation of plant height) that was estimated from the fitted curves. It shows that the methodology we used for curve fitting permits not only the characterization of the processes with the conventional parameters, but also the estimation of new characteristics according to the needs of the study.

Differences in trajectories were observed when comparing the fitted developmental curves according to earliness. In the case of flowering and senescence, early genotypes showed a complete s-shape curve while late genotypes show slow progression and only the first part of the s-shape was observed in most of the genotypes. As it is known, the genomic region on chromosome 5 controlling maturity has a pleiotropic effect on developmental traits (Celis-Gamboa 2002; Hurtado et al. 2012b; Malosetti et al. 2006) and it can explain the curve's trajectories defined according to earliness. On the other hand, there was not a clear relation between plant height and maturity as it was also observed in a previous study (Maris 1964). Photoperiod played a role in both development and agronomic performance of the plants and it was specially observed in the Dutch varieties used as controls in the experiment. They

were shorter compared with their height in the Netherlands and all of them showed fast senescence development indicating that under short day length, growth was restricted and maturation was accelerated. Another indication of the photoperiod effect on development was the flower abortion suffered by the Dutch varieties. It is known that reduction in day length can suppress flower development (Turner and Ewing 1988).

To understand the genetic basis of the complex traits included in our study, developmental traits were treated as continuous and dynamic processes instead of looking at particular single moments of the life cycle. During the curve fitting all the time points were analysed together, a proper characterisation of different developmental stages was done and then the genetic factors underlying the processes were identified. A more efficient QTL analysis was performed using the estimated developmental parameters instead of searching for QTLs per single time point. In addition, the number of QTL analyses was reduced. For instance, flowering was assessed in the field 17 times and we analysed only 6 parameters describing this trait. In the multi-trait QTL analysis presented here, all the parameters were analyzed simultaneously and the presence of pleiotropic QTLs was also investigated.
On the other hand, the combined use of development parameters and agronomic traits in the multi-trait analysis provided insight into (1) the genetic architecture of development and complex traits in potato, (2) the presence of time-dependent QTLs and (3) the genetic link between above and below ground traits as discussed below.

For each complex trait multiple QTLs were identified explaining the genetic bases of such quantitative traits. Some QTLs were permanently expressed during the whole developmental process and some others were associated with specific developmental stages. These time-dependent QTLs were detected for flowering, senescence and plant height. They showed a very low explained variance compared with the QTLs expressed during the whole process (e.g. QTLs related to mean progression rate). Some of the time-dependent QTLs were expressed at early development stages (QTLs associated with onset) and they seem to be switched off after half developmental process was completed. Some others were silent during early stages but were expressed in the final part of the process (QTLs associated with inflection point and end of the processes).

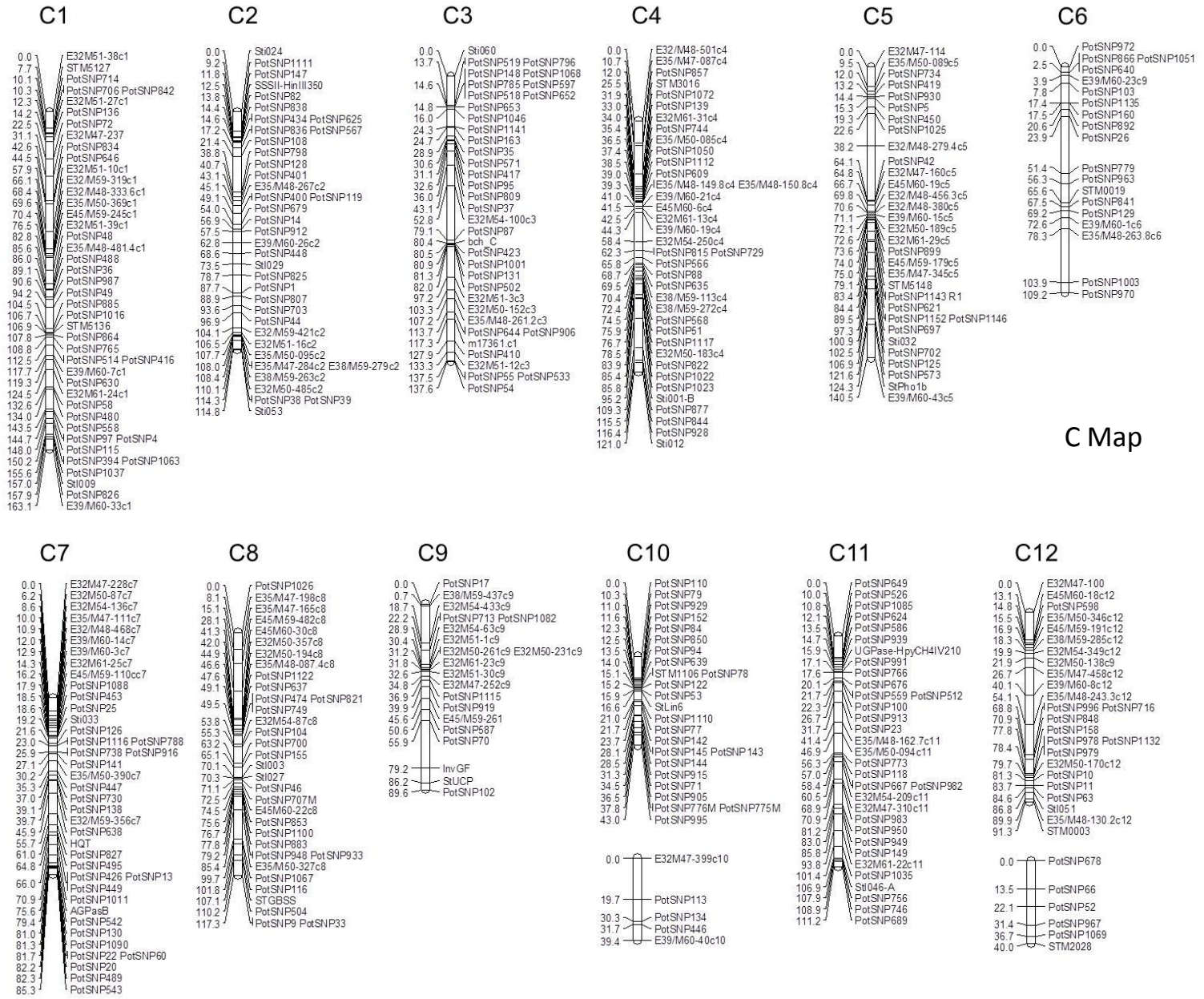
The genetic relation between above and below ground traits in potato was observed in 14 genetic regions associated with development and agronomic traits. It is important to notice that the traits have different scales and the interpretation of the allele effects depends on the grandparental origin of the alleles. Only within the same linkage group the allele effect can be attributed to the same grandparental difference between the alleles. In that sense, within a pleiotropic region the allele effects can be compared and they can even go in opposite directions as it was observed in all linkage groups where a QTL was detected except for linkage groups E6 and E11.

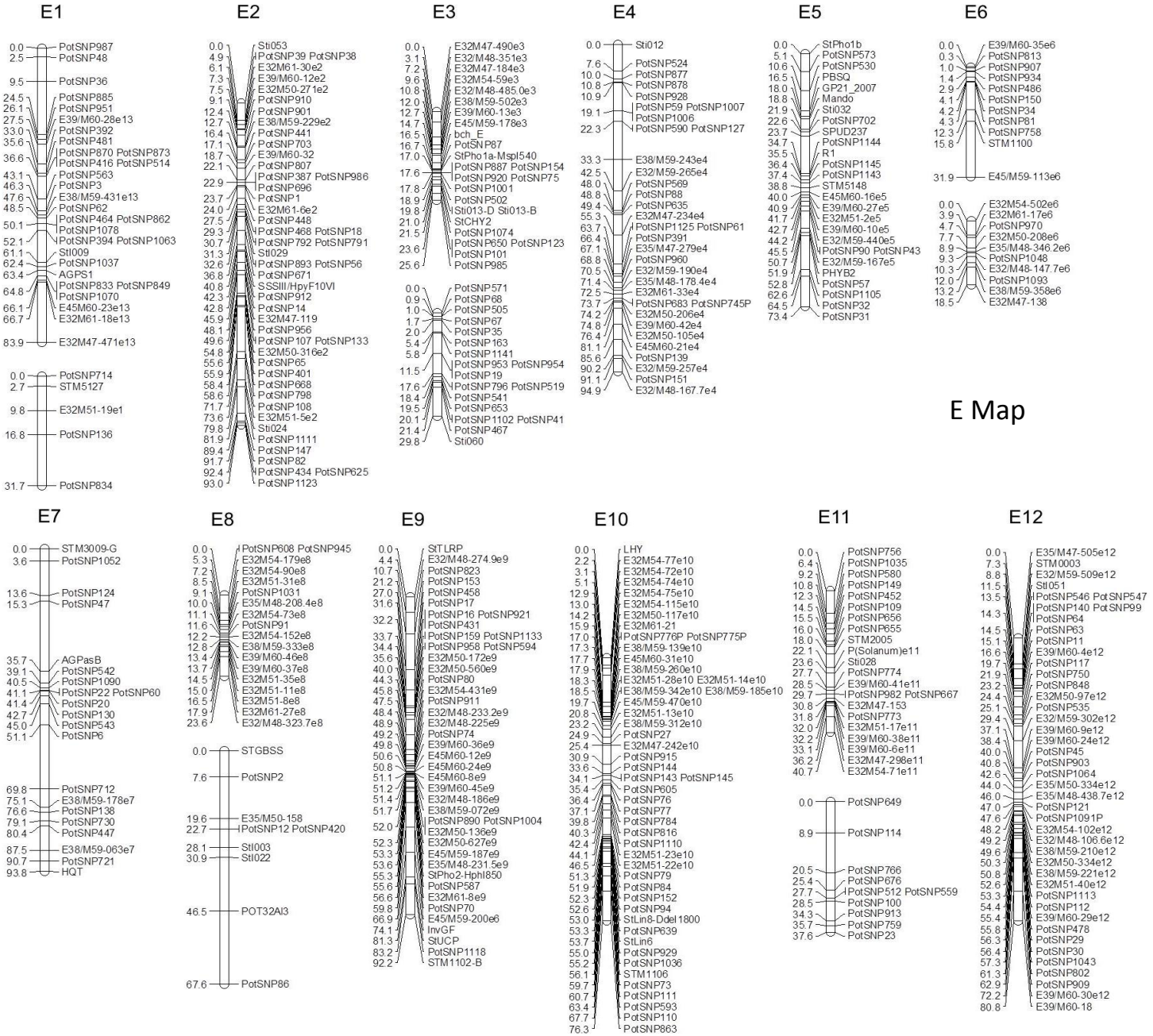
Further research will help 1) to confirm the stability of the pleiotropic regions found in our study across environments, 2) to check the consistency of the allele effects, which can vary according to the environmental setup where they are expressed (Clark 2000) and 3) to investigate the presence of one or more genes in those regions when evidences of linkage are existing.

Some of our results suggest the presence of QTLxE interactions but additional studies comparing development under different photoperiods are still necessary to take advantage of the plasticity of the crop. Multi-environment experiments will allow us to better quantify the effect of the different photoperiod on traits, such as the ones presented in this study. The anchoring of stable pleiotropic QTLs to the annotated potato genome sequence (Potato Genome Sequencing et al. 2011), will also provide target genes for marker assisted breeding and candidate gene approaches.

Additional files

Additional file 1. C and E linkage maps. The C map consists of 399 markers spanning 1403.3 cM with an average distance between adjacent markers of 3.5 cM. Two of the 12 LG (C10 and C12) were split in two sub-groups due to the large distance between adjacent markers (more than 30cM). The E map consists of 424 markers spanning 995.1 cM with average distance between adjacent markers of 2.3 cM. Five of the 12 LG (E1, E3, E6, E8, E11) were split in two subgroups due to the large distance between adjacent markers. The assignment of linkage groups was done according to Celis-Gamboa, 2002 (Celis-Gamboa 2002) and each LG is preceded by the letter C or E according to the parental map.


C1
0.0 E32M51-38c1
7.7 STM5127
10.1 PotSNP714
10.3 PotSNP706 PotSNP842
12.3 E32M51-27c1
14.2 PotSNP136
22.5 PotSNP72
31.1 E32M47-237
42.6 PotSNP834
44.5 PotSNP646
57.9 E32M51-10c1
66.1 E32/M59-319c1
68.4 E32/M48-333.6c1
69.6 E35/M50-369c1
70.4 E45/M59-245c1
76.5 E32M51-39c1
82.8 PotSNP48
85.6 E35/M48-481.4c1
86.0 PotSNP488
89.1 PotSNP36
90.6 PotSNP987
94.2 PotSNP49
104.5 PotSNP886
106.7 PotSNP1016
106.9 STM5136
107.8 PotSNP864
108.8 PotSNP765
112.5 PotSNP514 PotSNP416
117.7 E39/M60-7c1
119.3 PotSNP630
124.5 E32M61-24c1
132.6 PotSNP58
134.0 PotSNP480
143.5 PotSNP558
144.7 PotSNP97 PotSNP4
148.0 PotSNP115
150.2 PotSNP394 PotSNP1063
155.6 PotSNP1037
157.0 Sti009
157.9 PotSNP826
163.1 E39/M60-33c1
C2
0.0 Sti024
9.2 PotSNP1111
11.8 PotSNP147
12.5 SSSII-HinIII350
13.8 PotSNP82
14.4 PotSNP838
14.6 PotSNP434 PotSNP625
17.2 PotSNP836 PotSNP567
21.4 PotSNP108
38.8 PotSNP798
40.7 PotSNP128
43.1 PotSNP401
45.1 E35/M48-267c2
49.1 PotSNP400 PotSNP119
54.0 PotSNP679
56.9 PotSNP14
57.5 PotSNP912
62.8 E39/M60-26c2
68.6 PotSNP448
73.5 Sti029
78.7 PotSNP825
87.7 PotSNP1
88.9 PotSNP807
93.6 PotSNP703
96.9 PotSNP44
104.1 E32/M59-421c2
106.5 E32M51-16c2
107.7 E35/M50-095c2
108.0 E35/M47-284c2 E38/M59-279c2
108.4 E38/M59-263c2
110.1 E32M50-485c2
114.3 PotSNP38 PotSNP39
114.8 Sti053
C3
0.0 Sti060
13.7 PotSNP519 PotSNP796
PotSNP148 PotSNP1068
14.6 PotSNP785 PotSNP597
PotSNP518 PotSNP652
14.8 PotSNP653
16.0 PotSNP1046
24.3 PotSNP1141
24.7 PotSNP163
28.9 PotSNP35
30.6 PotSNP571
31.1 PotSNP417
32.6 PotSNP95
36.0 PotSNP809
43.1 PotSNP37
52.8 E32M54-100c3
79.1 PotSNP87
80.4 bch_C
80.5 PotSNP423
80.9 PotSNP1001
81.3 PotSNP131
82.0 PotSNP502
97.2 E32M51-3c3
103.3 E32M50-152c3
107.2 E35/M48-261.2c3
113.7 PotSNP644 PotSNP906
117.3 m17361.c1
127.9 PotSNP410
133.3 E32M51-12c3
137.5 PotSNP55 PotSNP533
137.6 PotSNP54
C4
0.0 E32/M48-501c4
10.7 E35/M47-087c4
12.0 PotSNP857
25.5 STM3016
31.9 PotSNP1072
33.0 PotSNP139
34.0 E32M61-31c4
35.4 PotSNP744
36.5 E35/M50-085c4
37.4 PotSNP1050
38.5 PotSNP1112
39.0 PotSNP609
39.3 E35/M48-149.8c4 E35/M48-150.8c4
41.0 E39/M60-21c4
41.5 E45M60-6c4
42.5 E32M61-13c4
44.3 E39/M60-19c4
58.4 E32M54-250c4
62.3 PotSNP815 PotSNP729
65.8 PotSNP566
68.7 PotSNP88
69.5 PotSNP635
70.4 E38/M59-113c4
72.4 E38/M59-272c4
74.5 PotSNP568
75.9 PotSNP51
76.7 PotSNP1117
78.5 E32M50-183c4
83.9 PotSNP822
85.4 PotSNP1022
85.8 PotSNP1023
95.2 Sti001-B
109.3 PotSNP877
115.5 PotSNP844
116.4 PotSNP928
121.0 Sti012
C5
0.0 E32M47-114
9.5 E35/M50-089c5
12.0 PotSNP734
13.2 PotSNP419
14.4 PotSNP930
15.3 PotSNP5
19.3 PotSNP450
22.6 PotSNP1025
38.2 E32/M48-279.4c5
64.1 PotSNP42
64.8 E32M47-160c5
66.7 E45M60-19c5
69.8 E32/M48-456.3c5
70.6 E32/M48-380c5
71.1 E39/M60-15c5
72.1 E32M50-189c5
72.6 E32M61-29c5
73.6 PotSNP899
74.0 E45/M59-179c5
75.0 E35/M47-345c5
79.1 STM5148
83.4 PotSNP1143 R1
84.4 PotSNP621
89.5 PotSNP1152 PotSNP1146
97.3 PotSNP697
100.9 Sti032
102.5 PotSNP702
106.9 PotSNP125
121.6 PotSNP573
124.3 StPho1b
140.5 E39/M60-43c5
C6
0.0 PotSNP972
2.5 PotSNP866 PotSNP1051
PotSNP640
3.9 E39/M60-23c9
7.8 PotSNP103
17.4 PotSNP1135
17.5 PotSNP160
20.6 PotSNP892
23.9 PotSNP26
51.4 PotSNP779
56.3 PotSNP963
65.6 STM0019
67.5 PotSNP841
69.2 PotSNP129
72.6 E39/M60-1c6
78.3 E35/M48-263.8c6
103.9 PotSNP1003
109.2 PotSNP970
C Map
C7
0.0 E32M47-228c7
6.2 E32M50-87c7
8.6 E32M54-136c7
10.0 E35/M47-111c7
10.9 E32/M48-468c7
12.0 E39/M60-14c7
12.9 E39/M60-3c7
14.3 E32M61-25c7
16.2 E45/M59-110cc7
17.9 PotSNP1088
18.5 PotSNP453
18.6 PotSNP25
19.2 Sti033
21.6 PotSNP126
23.0 PotSNP1116 PotSNP788
25.9 PotSNP738 PotSNP916
27.1 PotSNP141
30.2 E35/M50-390c7
35.3 PotSNP447
37.0 PotSNP730
39.1 PotSNP138
39.7 E32/M59-356c7
45.9 PotSNP638
55.7 HQT
61.0 PotSNP827
64.8 PotSNP495
66.0 PotSNP426 PotSNP13
70.9 PotSNP449
75.6 PotSNP1011
79.4 AGPasB
81.0 PotSNP542
81.3 PotSNP130
81.7 PotSNP1090
82.2 PotSNP22 PotSNP60
82.3 PotSNP20
85.3 PotSNP489
PotSNP543
C8
0.0 PotSNP1026
8.1 E35/M47-198c8
15.1 E35/M47-165c8
28.1 E45/M59-482c8
41.3 E45M60-30c8
42.0 E32M50-357c8
44.9 E32M50-194c8
46.6 E35/M48-087.4c8
47.6 PotSNP1122
49.1 PotSNP637
49.5 PotSNP474 PotSNP821
53.8 PotSNP749
55.3 E32M54-87c8
63.2 PotSNP104
65.1 PotSNP700
70.1 PotSNP155
70.3 Sti003
71.1 Sti027
72.5 PotSNP46
74.5 PotSNP707M
75.6 E45M60-22c8
76.7 PotSNP853
77.8 PotSNP1100
79.2 PotSNP883
85.4 PotSNP948 PotSNP933
99.7 E35/M50-327c8
101.8 PotSNP1067
107.1 PotSNP116
110.2 STGBSS
117.3 PotSNP504
PotSNP9 PotSNP33
C9
0.0 PotSNP17
0.7 E38/M59-437c9
18.7 E32M54-433c9
22.2 PotSNP713 PotSNP1082
28.9 E32M54-63c9
30.4 E32M51-1c9
31.2 E32M50-261c9 E32M50-231c9
31.8 E32M61-23c9
32.6 E32M51-30c9
34.8 E32M47-252c9
36.9 PotSNP1115
39.9 PotSNP919
45.6 E45/M59-261
50.6 PotSNP587
55.9 PotSNP70
79.2 InvGF
86.2 StUCP
89.6 PotSNP102
C10
0.0 PotSNP110
10.3 PotSNP79
11.0 PotSNP929
11.6 PotSNP152
12.3 PotSNP84
12.5 PotSNP850
13.5 PotSNP94
14.0 PotSNP639
15.1 STM1106 PotSNP78
15.2 PotSNP122
15.9 PotSNP53
16.6 StLin6
21.0 PotSNP1110
21.7 PotSNP77
23.7 PotSNP142
28.1 PotSNP145 PotSNP143
28.5 PotSNP144
31.3 PotSNP915
34.5 PotSNP71
36.5 PotSNP905
37.8 PotSNP776M PotSNP775M
43.0 PotSNP995
0.0 E32M47-399c10
19.7 PotSNP113
30.3 PotSNP134
31.7 PotSNP446
39.4 E39/M60-40c10
C11
0.0 PotSNP649
10.0 PotSNP526
10.8 PotSNP1085
12.1 PotSNP624
13.5 PotSNP586
14.7 PotSNP939
15.9 UGPase-HpyCH4IV210
17.1 PotSNP991
17.6 PotSNP766
20.1 PotSNP676
21.7 PotSNP559 PotSNP512
22.3 PotSNP100
26.7 PotSNP913
31.7 PotSNP23
41.4 E35/M48-162.7c11
46.9 E35/M50-094c11
56.3 PotSNP773
57.0 PotSNP118
58.4 PotSNP667 PotSNP982
60.5 E32M54-209c11
68.9 E32M47-310c11
70.9 PotSNP983
81.2 PotSNP950
83.0 PotSNP949
85.8 PotSNP149
93.8 E32M61-22c11
101.4 PotSNP1035
106.9 Sti046-A
107.9 PotSNP756
108.9 PotSNP746
111.2 PotSNP689
C12
0.0 E32M47-100
13.1 E45M60-18c12
14.8 PotSNP598
15.5 E35/M50-346c12
16.9 E45/M59-191c12
18.3 E38/M59-285c12
19.9 E32M54-349c12
21.9 E32M50-138c9
26.7 E35/M47-458c12
40.1 E39/M60-8c12
54.1 E35/M48-243.3c12
68.8 PotSNP996 PotSNP716
70.9 PotSNP848
77.8 PotSNP158
78.4 PotSNP978 PotSNP1132
PotSNP979
79.7 E32M50-170c12
81.3 PotSNP10
83.7 PotSNP11
84.6 PotSNP63
86.8 Sti051
89.9 E35/M48-130.2c12
91.3 STM0003
0.0 PotSNP678
13.5 PotSNP66
22.1 PotSNP52
31.4 PotSNP967
36.7 PotSNP1069
40.0 STM2028


E1
0.0 PotSNP987
2.5 PotSNP48
9.5 PotSNP36
24.5 PotSNP885
26.1 PotSNP951
27.5 E39/M60-28e13
33.0 PotSNP392
35.6 PotSNP481
36.6 PotSNP870 PotSNP873
PotSNP416 PotSNP514
43.1 PotSNP563
46.3 PotSNP3
47.6 E38/M59-431e13
48.5 PotSNP62
50.1 PotSNP464 PotSNP862
PotSNP1078
52.1 PotSNP394 PotSNP1063
61.1 StI009
62.4 PotSNP1037
63.4 AGPS1
64.8 PotSNP833 PotSNP849
PotSNP1070
66.1 E45M60-23e13
66.7 E32M61-18e13
83.9 E32M47-471e13
0.0 PotSNP714
2.7 STM5127
9.8 E32M51-19e1
16.8 PotSNP136
31.7 PotSNP834
E2
0.0 StI053
4.9 PotSNP39 PotSNP38
6.1 E32M61-30e2
7.3 E39/M60-12e2
7.5 E32M50-271e2
9.1 PotSNP910
12.4 PotSNP901
12.7 E38/M59-229e2
16.4 PotSNP441
17.1 PotSNP703
18.7 E39/M60-32
22.1 PotSNP807
22.9 PotSNP387 PotSNP986
PotSNP696
23.7 PotSNP1
24.0 E32M61-6e2
27.5 PotSNP448
29.3 PotSNP468 PotSNP18
30.7 PotSNP792 PotSNP791
31.3 StI029
32.6 PotSNP893 PotSNP56
36.8 PotSNP671
40.8 SSSIII/HpyF10VI
42.3 PotSNP912
42.8 PotSNP14
45.9 E32M47-119
48.1 PotSNP956
49.6 PotSNP107 PotSNP133
54.8 E32M50-316e2
55.6 PotSNP65
55.9 PotSNP401
58.4 PotSNP668
58.6 PotSNP798
71.7 PotSNP108
73.6 E32M51-5e2
79.8 StI024
81.9 PotSNP1111
89.4 PotSNP147
91.7 PotSNP82
92.4 PotSNP434 PotSNP625
93.0 PotSNP1123
E3
0.0 E32M47-490e3
3.1 E32/M48-351e3
7.2 E32M47-184e3
9.6 E32M54-59e3
10.8 E32/M48-485.0e3
12.0 E38/M59-502e3
12.7 E39/M60-13e3
14.7 E45/M59-178e3
16.5 bch_E
16.7 PotSNP87
17.0 StPho1a-MspI540
17.6 PotSNP887 PotSNP154
PotSNP920 PotSNP75
17.8 PotSNP1001
18.9 PotSNP502
19.8 StI013-D StI013-B
21.0 StCHY2
21.5 PotSNP1074
23.6 PotSNP650 PotSNP123
25.6 PotSNP101
PotSNP985
0.0 PotSNP571
0.9 PotSNP68
1.0 PotSNP505
1.7 PotSNP67
2.0 PotSNP35
5.4 PotSNP163
5.8 PotSNP1141
11.5 PotSNP953 PotSNP954
PotSNP19
17.6 PotSNP796 PotSNP519
18.4 PotSNP541
19.5 PotSNP653
20.1 PotSNP1102 PotSNP41
21.4 PotSNP467
29.8 StI060
E4
0.0 StI012
7.6 PotSNP524
10.0 PotSNP877
10.8 PotSNP878
10.9 PotSNP928
19.1 PotSNP59 PotSNP1007
PotSNP1006
22.3 PotSNP590 PotSNP127
33.3 E38/M59-243e4
42.5 E32/M59-265e4
48.0 PotSNP569
48.8 PotSNP88
49.4 PotSNP635
55.3 E32M47-234e4
63.7 PotSNP1125 PotSNP61
66.4 PotSNP391
67.1 E35/M47-279e4
68.8 PotSNP960
70.5 E32/M59-190e4
71.4 E35/M48-178.4e4
72.5 E32M61-33e4
73.7 PotSNP683 PotSNP745P
74.2 E32M50-206e4
74.8 E39/M60-42e4
76.4 E32M50-105e4
81.1 E45M60-21e4
85.6 PotSNP139
90.2 E32/M59-257e4
91.1 PotSNP151
94.9 E32/M48-167.7e4
E5
0.0 StPho1b
5.1 PotSNP573
10.6 PotSNP530
16.5 PBSQ
18.0 GP21_2007
18.8 Mando
21.9 StI032
22.6 PotSNP702
23.7 SPUD237
34.7 PotSNP1144
35.5 R1
36.4 PotSNP1145
37.4 PotSNP1143
38.8 STM5148
40.0 E45M60-16e5
40.9 E39/M60-27e5
41.7 E32M51-2e5
42.7 E39/M60-10e5
44.2 E32/M59-440e5
45.5 PotSNP90 PotSNP43
50.7 E32/M59-167e5
51.9 PHYB2
52.8 PotSNP57
62.6 PotSNP1105
64.5 PotSNP32
73.4 PotSNP31
E6
0.0 E39/M60-35e6
0.3 PotSNP813
1.0 PotSNP907
1.4 PotSNP934
2.9 PotSNP486
4.1 PotSNP150
4.2 PotSNP34
4.3 PotSNP81
12.3 PotSNP758
15.8 STM1100
31.9 E45/M59-113e6
0.0 E32M54-502e6
3.9 E32M61-17e6
4.7 PotSNP970
7.7 E32M50-208e6
8.9 E35/M48-346.2e6
9.3 PotSNP1048
10.3 E32/M48-147.7e6
12.0 PotSNP1093
13.2 E38/M59-358e6
18.5 E32M47-138
E Map
E7
0.0 STM3009-G
3.6 PotSNP1052
13.6 PotSNP124
15.3 PotSNP47
35.7 AGPasB
39.1 PotSNP542
40.5 PotSNP1090
41.1 PotSNP22 PotSNP60
41.4 PotSNP20
42.7 PotSNP130
45.0 PotSNP543
51.1 PotSNP6
69.8 PotSNP712
75.1 E38/M59-178e7
76.6 PotSNP138
79.1 PotSNP730
80.4 PotSNP447
87.5 E38/M59-063e7
90.7 PotSNP721
93.8 HQT
E8
0.0 PotSNP608 PotSNP945
5.3 E32M54-179e8
7.2 E32M54-90e8
8.5 E32M51-31e8
9.1 PotSNP1031
10.0 E35/M48-208.4e8
11.1 E32M54-73e8
11.6 PotSNP91
12.2 E32M54-152e8
12.8 E38/M59-333e8
13.4 E39/M60-46e8
13.7 E39/M60-37e8
14.5 E32M51-35e8
15.0 E32M51-11e8
16.5 E32M51-8e8
17.9 E32M61-27e8
23.6 E32/M48-323.7e8
0.0 STGBSS
7.6 PotSNP2
19.6 E35/M50-158
22.7 PotSNP12 PotSNP420
28.1 StI003
30.9 StI022
46.5 POT32AI3
67.6 PotSNP86
E9
0.0 StTLRP
4.4 E32/M48-274.9e9
10.7 PotSNP823
21.2 PotSNP153
27.0 PotSNP458
31.6 PotSNP17
32.2 PotSNP16 PotSNP921
PotSNP431
33.7 PotSNP159 PotSNP1133
34.4 PotSNP958 PotSNP594
35.6 E32M50-172e9
40.0 E32M50-560e9
44.3 PotSNP80
45.8 E32M54-431e9
47.5 PotSNP911
48.4 E32/M48-233.2e9
48.9 E32/M48-225e9
49.2 PotSNP74
49.8 E39/M60-36e9
50.6 E45M60-12e9
50.8 E45M60-24e9
51.1 E45M60-8e9
51.2 E39/M60-45e9
51.4 E32/M48-186e9
51.7 E38/M59-072e9
52.0 PotSNP890 PotSNP1004
52.3 E32M50-136e9
53.3 E32M50-627e9
53.6 E45/M59-187e9
55.3 E35/M48-231.5e9
55.6 StPho2-HphI850
56.6 PotSNP587
59.8 E32M61-8e9
66.9 PotSNP70
74.1 E45/M59-200e6
81.3 InvGF
83.2 StUCP
92.2 PotSNP1118
STM1102-B
E10
0.0 LHY
2.2 E32M54-77e10
3.1 E32M54-72e10
5.1 E32M54-74e10
12.9 E32M54-75e10
13.0 E32M54-115e10
14.2 E32M50-117e10
15.9 E32M61-21
17.0 PotSNP776P PotSNP775P
17.3 E38/M59-139e10
17.7 E45M60-31e10
17.9 E38/M59-260e10
18.3 E32M51-28e10 E32M51-14e10
18.5 E38/M59-342e10 E38/M59-185e10
19.7 E45/M59-470e10
20.8 E32M51-13e10
23.2 E38/M59-312e10
24.9 PotSNP27
25.4 E32M47-242e10
30.9 PotSNP915
33.6 PotSNP144
34.1 PotSNP143 PotSNP145
35.4 PotSNP605
36.4 PotSNP76
37.1 PotSNP77
39.8 PotSNP784
40.3 PotSNP816
42.4 PotSNP1110
44.1 E32M51-23e10
46.5 E32M51-22e10
51.3 PotSNP79
51.9 PotSNP84
52.3 PotSNP152
52.6 PotSNP94
53.0 StLin8-DdeI 1800
53.3 PotSNP639
53.7 StLin6
55.0 PotSNP929
55.2 PotSNP1036
56.1 STM1106
59.7 PotSNP73
60.7 PotSNP111
63.4 PotSNP593
67.7 PotSNP110
76.3 PotSNP863
E11
0.0 PotSNP756
6.4 PotSNP1035
9.2 PotSNP580
10.8 PotSNP149
12.3 PotSNP452
14.5 PotSNP109
15.5 PotSNP656
16.0 PotSNP655
18.0 STM2005
22.1 P(Solanum)e11
23.6 StI028
27.7 PotSNP774
28.5 E39/M60-41e11
29.7 PotSNP982 PotSNP667
30.8 E32M47-153
31.8 PotSNP773
32.0 E32M51-17e11
32.2 E39/M60-38e11
33.1 E39/M60-6e11
36.2 E32M47-298e11
40.7 E32M54-71e11
0.0 PotSNP649
8.9 PotSNP114
20.5 PotSNP766
25.4 PotSNP676
27.7 PotSNP512 PotSNP559
28.5 PotSNP100
34.3 PotSNP913
35.7 PotSNP759
37.6 PotSNP23
E12
0.0 E35/M47-505e12
7.3 STM0003
8.8 E32/M59-509e12
11.5 StI051
13.5 PotSNP546 PotSNP547
14.3 PotSNP140 PotSNP99
PotSNP64
14.5 PotSNP63
15.1 PotSNP11
16.6 E39/M60-4e12
19.7 PotSNP117
21.9 PotSNP750
23.2 PotSNP848
24.4 E32M50-97e12
25.1 PotSNP535
29.4 E32/M59-302e12
37.1 E39/M60-9e12
38.4 E39/M60-24e12
40.0 PotSNP45
40.8 PotSNP903
42.6 PotSNP1064
44.0 E35/M50-334e12
46.0 E35/M48-438.7e12
47.0 PotSNP121
47.6 PotSNP1091P
48.2 E32M54-102e12
49.2 E32/M48-106.6e12
49.6 E38/M59-210e12
50.3 E32M50-334e12
50.8 E38/M59-221e12
52.6 E32M51-40e12
53.3 PotSNP1113
54.4 PotSNP112
55.4 E39/M60-29e12
55.8 PotSNP478
56.3 PotSNP29
56.4 PotSNP30
57.3 PotSNP1043
61.3 PotSNP802
62.9 PotSNP909
72.2 E39/M60-30e12
80.8 E39/M60-18

Additional file 2. Correlation between residuals of the traits characterizing plant height, flowering, senescence and agronomic traits (emergence, main stems, total number of tubers and yield). The effect of chromosome 5 has been removed from all the traits.

FonsresM	1	-																							
FmaxresM	2	0.108	-																						
FipresM	3	0.589	-0.204	-																					
FendresM	4	0.433	-0.515	0.651	-																				
FmeanresM	5	-0.057	0.301	-0.298	-0.454	-																			
FdurresM	6	0.020	-0.620	0.438	0.752	-0.668	-																		
SonsresM	7	0.054	-0.247	0.178	0.244	-0.238	0.271	-																	
SmaxresM	8	-0.192	0.002	-0.244	-0.197	0.083	-0.137	-0.222	-																
SipresM	9	0.192	-0.245	0.327	0.438	-0.226	0.363	0.707	-0.526	-															
SendresM	10	0.151	-0.202	0.315	0.429	-0.263	0.366	0.653	-0.645	0.888	-														
SmeanresM	11	-0.185	0.085	-0.311	-0.340	0.318	-0.251	-0.117	0.438	-0.303	-0.480	-													
PhonsresM	12	0.246	-0.122	0.328	0.419	-0.144	0.259	0.205	-0.315	0.373	0.398	-0.167	-												
PhmaxonsresM	13	-0.082	-0.174	-0.080	0.009	0.083	0.015	-0.099	-0.145	-0.003	0.049	0.008	-0.077	-											
PhmaxresM	14	0.154	-0.325	0.255	0.453	-0.194	0.361	0.145	-0.300	0.330	0.381	-0.179	0.393	0.745	-										
PhipresM	15	0.179	-0.138	0.338	0.404	-0.184	0.311	0.242	-0.282	0.360	0.363	-0.113	0.674	-0.274	0.229	-									
PhendresM	16	0.099	-0.074	0.242	0.295	-0.133	0.230	0.272	-0.154	0.411	0.380	-0.091	0.580	-0.221	0.219	0.480	-								
PhmeanresM	17	0.112	-0.352	0.313	0.515	-0.228	0.445	0.226	-0.344	0.433	0.484	-0.238	0.441	0.603	0.915	0.355	0.381	-							
PhmaxheresM	18	0.091	-0.336	0.268	0.456	-0.196	0.395	0.172	-0.328	0.374	0.430	-0.202	0.402	0.689	0.937	0.291	0.318	0.988	-						
PhmeanheresM	19	-0.045	-0.254	0.111	0.252	-0.072	0.275	0.127	-0.263	0.295	0.359	-0.142	0.107	0.804	0.768	-0.020	0.052	0.790	0.806	-					
emeresM	20	0.404	0.103	0.312	0.212	0.043	-0.029	0.033	-0.066	0.064	0.002	-0.058	0.313	-0.291	-0.085	0.375	0.059	-0.110	-0.135	-0.315	-				
mainstemsresM	21	-0.125	0.003	-0.107	-0.124	0.062	-0.028	-0.097	0.085	-0.167	-0.142	0.120	-0.186	0.187	0.012	-0.240	-0.083	0.004	0.033	0.138	-0.329	-			
totaltibersresM	22	0.055	0.017	0.097	0.162	-0.120	0.125	0.068	-0.140	0.142	0.203	-0.093	0.135	0.048	0.157	0.053	0.184	0.217	0.205	0.218	-0.245	0.647	-		
totalweightresM	23	-0.044	-0.121	-0.003	0.097	-0.078	0.188	0.146	-0.301	0.276	0.406	-0.243	0.073	0.459	0.437	-0.042	0.102	0.486	0.486	0.650	-0.348	0.338	0.496	-	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

Flowering

Senescence

Plant height

Additional file 3. QTL results of single trait linkage analysis using the characteristics describing Plant height (A), flowering (B) and senescence (C) and QTLs associated to the agronomic traits measured at harvest (D)

A. Plant height

		QTL description					
Trait	**Marker**	**LG**	**cM**	**(-log10P)**	**%Expl**	**Add eff**	**s.e**
Onset	E35/M47-345c5	C5	75.03	3.071	4.535	-0.332	0.098
	E39/M60-27e5	E5	40.94	11.401	21.961	-0.73	0.097
	E39/M60-4e12	E12	16.57	4.731	7.691	0.432	0.098
maxons (slope)	Sti012m	C4	121.04	3.014	5.328	0.063	0.019
	PotSNP450	C5	19.3	3.162	6.375	-0.069	0.02
	Sti032f	E5	21.89	10.192	21.852	-0.128	0.018
Max slope	PotSNP450	C5	19.3	5.308	8.818	-0.136	0.029
	Sti032m	C5	100.88	4.12	5.538	-0.108	0.027
	Sti032f	E5	21.89	20.686	38.935	-0.286	0.026
Ipoint	Sti032f	E5	21.89	10.845	22.445	-0.632	0.087
	E39/M60-4e12	E12	16.57	3.591	6.067	0.329	0.088
End	GP21_2007	E5	17.97	6.8	15.602	-0.511	0.093
Mean slope	PotSNP706	C1	10.29	3.131	3.682	0.051	0.015
	PotSNP450	C5	19.3	3.838	5.551	-0.063	0.016
	Sti032m	C5	100.88	3.023	3.563	-0.051	0.015
	Mando	E5	18.79	23.461	44.482	-0.179	0.015
Max height	PotSNP706	C1	10.29	2.729	4.117	2.416	0.764
	PotSNP450	C5	19.3	2.84	4.888	-2.632	0.812
	E39/M60-27e5	E5	40.94	13.585	27.368	-6.229	0.746
meanheight	STM5127m	C1	7.66	3.775	4.678	1.253	0.325
	StPho1bm	C5	124.25	4.009	4.976	-1.292	0.324
	Mando	E5	18.79	21.439	40.675	-3.695	0.328

B. Flowering

		QTL description					
Trait	**Marker**	**LG**	**cM**	**(-log10P)**	**%Expl**	**Add eff**	**s.e**
onset	-						
max slope	Sti032m	C5	100.88	6.238	10.745	0.261	0.05
	SPUD237	E5	23.67	17.576	41.061	0.51	0.051
Ipoint	PotSNP1145	E5	36.43	6.396	16.113	-0.421	0.079
End	PotSNP450	C5	19.3	2.454	4.677	-0.242	0.082
	Sti032m	C5	100.88	2.777	4.719	-0.243	0.076
	SPUD237	E5	23.67	14.62	35.705	-0.67	0.076
mean slope	-						
Duration Flow	Sti032m	C5	100.88	3.815	6.896	-0.339	0.087
	Mando	E5	18.79	13.435	33.378	-0.746	0.089

C. Senescence

		QTL description					
trait	**marker**	**LG**	**cM**	**(-log10P)**	**%Expl**	**Add eff**	**s.e**
onset	R1f	E5	35.47	4.987	11.823	-0.389	0.085
max slope	E39/M60-27e5	E5	40.94	9.712	22.992	0.388	0.057
ipoint	Sti032m	C5	100.88	8.005	10.711	-0.397	0.066
	Sti032f	E5	21.89	24.714	46.025	-0.823	0.065
	PotSNP81	E6	4.27	3.654	4.474	0.257	0.068
	PotSNP91	E8	11.57	2.8	4.114	0.246	0.077
end	PotSNP125	C5	106.9	9.833	15.146	-0.683	0.099
	Mando	E5	18.79	25.252	47.066	-1.204	0.094
	PotSNP486	E6	2.94	3.162	3.541	0.33	0.095
mean slope	-						

D. Traits measured at harvest

			QTL description				
Trait	**h2**	**Marker**	**LG**	**cM**	**LOD**	**%Expl. Var**	**Add effect**
emergence	0.8093	PotSNP142	C10	11.13	5.308	13.548	0.647
	StI022f	E8A	30.87	3.944	9.404	-0.539	
total tubers	0.8777	PotSNP95	C3	104.97	3.835	11.355	-1.891
	Sti032f	E5	21.89	4.703	9.761	-1.753	
total weight	0.8324	Sti032m	C5	100.88	5.248	8.43	-26.857
	StI009f	E1A	35.07	3.889	5.793	22.263	
	Sti032f	E5	21.89	13.98	27.294	-48.324	
# main stems	0.8015	PotSNP621	C5	84.41	3.742	8.427	0.379

Chapter 5

Genetic basis of adaptation in potato under different day-length conditions

P.X. Hurtado-Lopez[1,2,5], S.K. Schnabel[2,3], J. Jansen[2,3], C. Maliepaard[1], P.H.C. Eilers[2,4], R.G.F. Visser[1,3], F.A van Eeuwijk[2,3]

[1]Wageningen UR Plant Breeding, Wageningen University & Research Centre, Wageningen, The Netherlands
[2]Biometris, Wageningen University & Research Centre, Wageningen, The Netherlands
[3]Centre for BioSystems Genomics, Wageningen, The Netherlands.
[4]Erasmus University Medical Centre. Rotterdam, The Netherlands
[5]C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC). Wageningen University, Wageningen, The Netherlands

To be submitted

Abstract

The presence of genotype by environment and QTL by environment interactions play an important role in the expression of complex traits involved in plant development under field conditions. To understand the genetic basis of developmental processes in potato, experiments have been carried out at latitudes with different day lengths. In this study, field trials including ~200 genotypes from a diploid backcross mapping population were planted under 3 contrasting day lengths, in Ethiopia and Venezuela (short), The Netherlands (long) and Finland (very long). Flowering, haulm senescence and plant height were evaluated as time series during the growing season, while some important agronomic traits were measured at harvest to get a better general understanding of potato development and adaptation under the contrasting environments.

The analysis of the multi-environment trials (METs) provided the bases for understanding 1) the effect of temperature and photoperiod on development, growth and agronomic traits evaluated under field conditions, 2) genotype by environment interactions that play an important role in the phenological plasticity of the crop and 3) associations between developmental and agronomic traits. Furthermore, the multi-environment QTL analyses combining MET data with genetic information allowed the identification of 4) QTL by environment interactions, 5) pleiotropic genetic regions, 6) QTLs which are stable across environments and 7) time dependent QTLs for development and growth.

The successful identification of pleiotropic regions, photoperiod-independent QTLs and time-dependent QTLs based on the multi-environment field trials opens the gate to further research to anchor the identified genetic regions to annotated functional genes which can be implemented in new breeding strategies.

Introduction

Potato (*Solanum tuberosum*), originating from the Andes in South America, is adapted to full sun light, moderate day time temperature and cool nights. However, it has been considered as a long day crop with optimal day length for development depending on temperature and cultivar (Manrique 1992). Previous studies have shown temperature and photoperiod to be the major environmental factors controlling development in potato (Ewing and Struik 1992; Levy and Veilleux 2007). Above- and below-ground parts of the plant as well as the different stages of the potato developmental cycle are differentially affected by both factors. For

instance, it is known that photoperiod has little or no effect on the duration of developmental phases after tuber initiation (Haverkort and Kooman 1997; Kooman et al. 1996b), but these phases are determined mainly by temperature (Kooman et al. 1996b). The effect of both factors on physiological processes in potato has also been extensively studied (Borah and Ilthorpe 1962; Haverkort 1990; Kooman et al. 1996b; Levy and Veilleux 2007; Manrique 1992; Midmore 1984; Struik and Ewing 1995; Struik et al. 1988; Turner and Ewing 1988; VanDam et al. 1996). It has been reported that they affect among others the rate of accumulation of biomass, the partitioning of biomass to leaves, stems, roots, stolons and tubers, the number and size of tubers and the rates of canopy development and crop senescence (Struik and Ewing 1995). However, most of the information from previous studies has been obtained under highly controlled conditions simulating different day lengths and different temperature regimes. These artificial conditions avoid the confounding effects of photoperiod and/or temperature with other environmental factors such as rainfall, abiotic/biotic stresses, etc. Therefore the qualitative effects of photoperiod and temperature on traits such as size of the haulm, productivity, onset, duration and rate of tuber bulking are relatively well-known (Struik et al. 1989; Struik and Ewing 1995; Struik et al. 1988; Turner and Ewing 1988; VanDam et al. 1996; Wheeler et al. 1986). However, quantitative information on the effects of photoperiod and temperature on the different aspects of tuber formation, development and growth and their interactions are limited. In addition, a major drawback of most studies is the number of cultivars included in the experiments. For instance, genotype by environment interactions have been reported for a few traits but using only a limited number of cultivars (VanDam et al. 1996). Studies for a wide range of cultivars or mapping populations are scarce.

Due to the lack of information about development and yield for large numbers of cultivars, different models have been used to simulate tuber dry matter production in specific environments and to select cultivars for such environments (Kooman and Haverkort 1995). Although a range of optimum temperatures for potato production including some of the optimal temperatures reported in literature (Borah and Ilthorpe 1962; Ingram and McCloud 1984; Manrique 1992; Midmore 1984; Sands and Regel 1983) have been estimated between 18 and 24°C (Kooman and Haverkort 1995), a wider span of optimal temperatures has been observed in long days than in short days. This wider optimal temperature range for growth

of the crop in longer days explains the broad adaptation of potato to a wide range of climates. Besides its short day and low temperature origin, the crop also performs well in regions with higher temperatures such as the Mediterranean region and temperate regions during summer because of the long days during the growing season (Kooman and Haverkort 1995).

Despite the efforts to study the physiological response of potato to different environmental conditions and to select cultivars with good performance in specific locations, a genetic component is still missing in most of the studies. Understanding the genetic basis of development and other complex traits can support previous physiological studies and facilitate breeding strategies. Complex traits are the result of physiological processes and environmental influences during the crop cycle which involve interactions of many gene pathways and networks. The inclusion of populations of genetically related individuals in multi-environment trials (METs), facilitates the understanding of the genetic control of adaptive traits by identifying associations with QTLs (Boer et al. 2007; Malosetti et al. 2008). Hence, the use of a methodology combining genetic and phenotypic information from METs allows the detection of genetic factors controlling development and complex traits and also to study the presence of QTL by environment interactions.

In our study we made use of a diploid potato population (~200 genotypes) evaluated under contrasting day lengths. The field trials were carried out in Ethiopia, Venezuela, Finland and The Netherlands. Development and growth were periodically assessed during the growing season, while agronomic traits were mainly measured at harvest. The aims of our study were 1) to investigate the effects of temperature and photoperiod on development, growth and agronomic traits evaluated under field conditions, 2) to study the role of genotype-environment interaction in phenological and agronomic traits 3) to investigate the relationship between developmental and agronomic traits in contrasting environments and 4) to identify the genetic factors controlling development, growth and agronomic traits under different day length conditions.

The use of a multi-environment QTL analysis allowed us to successfully identify pleiotropic genetic regions, photoperiod-independent QTLs and time-dependent QTLs for development

and growth processes. Our results allow for further research to be initiated making use of the recently sequenced potato genome (Potato Genome Sequencing et al. 2011) to anchor the identified genetic regions to annotated functional genes which could be implemented in new breeding strategies.

Materials and methods

Plant materials and genetic background of the potato population

A diploid backcross potato population was obtained from a cross between the diploid parental clones C (USW5337.3 (Bonhomme 2000; Hanneman and Peloquin 1967)) and E (77.2102.37 (Jacobsen 1980)). The female parent, clone C, is a hybrid between *Solanum phureja* (PI225696.1) and the *S. tuberosum* dihaploid (USW42). The E clone was obtained from a cross between clone C and the *S. vernei* - *S. tuberosum* backcross clone $VH^3$4211. The CxE population was developed for research purposes (Jacobs et al. 1995) based on the different genetic background of *S. phureja and S.tuberosum* and their phenotypic contrast. The most relevant characteristics of *S. phureja* are lack of tuber dormancy period, early maturity and short day tuberization induction (Ochoa 1990). In contrast, *S.tuberosum* varieties are grown in temperate regions, have long tuber dormancy, variable maturity and long-day tuberization induction (Celis-Gamboa 2002; Hawkes 1990). These characteristics make the CxE population suitable for the study of phenological processes taking place along the life cycle and also to study performance and adaptation to different day length conditions. This population is also known for its segregation of agronomic and quality traits (Celis-Gamboa 2002; Kloosterman et al. 2010). Therefore the relationships between relevant agronomic and developmental traits can also be studied.

The parental clones C and E and up to 244 of their offspring (referred to as CE clones) were evaluated under 3 day lengths (short, long and very long days). Common European cultivars and regional varieties (local standards) were planted in each location as controls.

Multi-environment trials (MET)

Field trial data of the CxE population was obtained from 3 contrasting day lengths (short, long and very long days). Ethiopia and Venezuela represented short day length, the Netherlands long days and two experimental years in Finland (2004 and 2005) represented

very long days with up to 21 hours of light per day. The local conditions in each of the 5 environments during the observation period are described in Table 1. The description includes year, growing season planting date, observation period (days after planting, DAP), range of temperatures recorded in the weather stations closest to each experiment, range of hours of light per day, altitude and the exact location of the experiments.

The experimental designs used in Finland 2004, The Netherlands and Ethiopia have been previously described (Celis-Gamboa et al. 2003; Hurtado et al. 2012a; Zaban et al. 2006). In Venezuela, the experiment included in our study was conducted in La Fresa (2200 meters above sea level), located in Mérida, one of the three Andean states of Venezuela. The parental clones (C and E), 238 CE genotypes and the control varieties were planted in plots with 3 plants per genotype in a randomized complete block design In Finland 2005, the field trial was carried out in the same location as the experiment in 2004 (North Ostrobothnia Research Station in Ruukki). In 2005 the parental clones, 222 CE genotypes and the same control varieties included in Finland 2004 were planted in plots of three plants per genotypes. The genotypes /plots were randomized within 4 blocks that were harvested 43, 57, 71 and 127 DAP respectively.

In The Netherlands and Finland the experimental designs included between 4 and 14 blocks where individual blocks were harvested at particular times during the growing season. The experiments were planned in such a way that intermediate harvest could also be done to measure some tuberization traits along the life cycle. Therefore, most of the assessments were done before or during harvesting of each block along the observation period.
None of the 5 environments had real replication of genotypes, but only pseudo-replications (multiple plants of a genotype per plot) within each block.

Agronomic and developmental traits

The traits included in this study were divided in two groups: characteristics measured once during the life cycle (referred to as agronomic traits) and traits evaluated in time series (referred to as developmental or growth traits). As it is expected in METs, different sets of genotypes were evaluated in each environment and different groups of traits were assessed

in each location. Table 1 shows how many genotypes and which traits were included in each field trial.

Considering that the first field trial of the CxE population was planted in the Netherlands and the experiments in the other locations were inspired on it, we are taking the description of the traits evaluated there as a reference for the other environments. However, some modifications were done to particular traits or the description was incomplete for few traits in other environments (Additional file 1). Special data pre-processing was needed to allow for the comparability between environments.

Table 1. Description of local condition in each field trial, number of CxE genotypes and traits assessed in each experiment

	Environment				
Local conditions	**Ethiopia**	**Venezuela**	**Finland (2004)**	**Finland (2005)**	**Netherlands**
Year	2010	2003	2004	2005	1999
Growing season [a]	July-December	June-November	16 April-10 Nov	3 May-14 Nov	12 Mar-14Nov
Planting date	Jul-16	Jun-24	Jun-01	May-16	May-12
Obs. period (DAP)	137	85	129	115	185
Temperature (°C)	11.3-17.5	14-21.5	2.6-20.5	4.8-21.9	3.0-23.6
Hours of light per day	11-Dec	11-Dec	Oct-21	13-21	Sep-17
Altitude (meters)	2400	2200	48	48	11
Location	09°07'N, 38°03'E	08°36' N, 71°08'W	64°42' N, 25°00'E	64°42' N, 25°00'E	51°58' N, 05°38'E
Number of CE genotypes	169	238	197	222	244
Flowering (# tpoints)*/scale	17 / 0-7	6**	13 / 0-7	15/01/2007	14 / 0-7
Observation period (DAP)	38-89	39-71	50-123	42-115	42-155
Senescence(# tpoints)*/scale	16/01/2007		11/01/2007	09/01/2007	14/01/2007
Observation period (DAP)	83-136	1**	77-129	84-115	75-185
Plant height (# tpoints)*	9		7	3**	
Observation period (DAP)	26-74	1	30-79	43-71	1
Total tuber number	X	X	X	X	X
No of tubers < 20mm	X	N.A	X	N.A	X
No of tubers > 20mm	X	N.A	X	N.A	X
Total tuber Weight	X	X	X	X	X
Weight of tubers < 20mm	X	N.A	X	N.A	N.A
Weight of tubers > 20mm	X	N.A	X	N.A	N.A
Tuber size	N.A	X	X	N.A	X
Date of emergence (DAP	X	N.A	N.A	X	N.A
No of main stems	X	X	X	X	X

[a]Growing season: defined as the period between the last frost of spring and the first severe frost in autumn (Allaby 1998). In Ethiopia and Venezuela (tropical conditions), it depended on the precipitation pattern.

**Traits measured in time series with the number of time points and the scale used in each environment*

***Data not used in this study*

x: Data included in this study

N.A: Data not analysed

Data pre-processing

Data pre-processing was done for all the traits before proceeding with the MET analysis.

Agronomic traits: A linear mixed model was used to analyse the data from each field trial and the best linear unbiased estimators (BLUEs) were obtained for each trait. In the mixed models genotype was fitted as fixed term and the random term depended on the experimental design of the field trial. In Ethiopia block was fitted as random term, while in The Netherlands, Finland and Venezuela no additional terms were fitted in the models. The BLUEs, having a single value per genotype, were the new phenotypic values for each trait.

For tuber size and total tuber weight, the BLUEs were normalized (subtracting the average and dividing by the standard deviation) in each environment due to the big differences between the measurements. This environment-based transformation has been proposed to investigate the relationship among environments based on the way how these environments discriminate among genotypes grown in each one (DeLacy et al. 1996).

Development and growth traits: Before the curve fitting was done for flowering, senescence and plant height, the scales used in each environment were shifted, standardized and in some cases data imputation for the beginning or the end of the curves was done if necessary to make the scales comparable (Schnabel et al. 2012). For instance, to allow the comparability between the developmental and growth curves, the maximum height reached by each plant was imputed in the final time points of the observation period. This data imputation was done because at the end of the growth process plant height was not recorded anymore for some genotypes in which hardly any changes were observed (mainly early genotypes). The imputation was also done for some genotypes when the final observations were smaller than the maximum height probably due to measurements done on a different stem. Such data imputation allowed the growth curves to have an approximate s-shape with an upper asymptote comparable to the flowering and senescence curves. Therefore the parameters derived from developmental and growth curves had a comparable meaning.

Conversion of DAP into PBTT

The effect of temperature on the crop development rate is often described by using a thermal time approach (Bonhomme 2000). Various non-linear models have been developed to describe temperature response of developmental processes in plants (Gao et al. 1992b;

Johnson and Thornley 1985; Yin et al. 1995b). In our study, the daily contribution of temperature to plant development was different between the environments due to the fluctuations in daily air temperature during the observation period (Figure 1A). Therefore, daily air temperature was incorporated in a non-linear temperature effect function to transform calendar days (days after planting, DAP) into beta thermal time, BTT, as previously described (Hurtado et al. 2012b). We also incorporated day length into the model considering the effect of photoperiod on growth and development in the different environments. The new thermal unit is the cumulative beta thermal days after planting. This unit combines, temperature, time and photoperiod: photo-beta thermal time, PBTT (Figure 1B). These PBTT units are used in the x-axis in the analysis of the time series data of plant height, flowering and senescence to be able to compare these traits at the different locations considered in this study.

Curve fitting of development traits

The curve fitting for flowering, senescence and plant height was done using PBTT units on the x-axis and the corresponding scale for each trait in the y-axis. For modelling flowering and senescence we used a methodology previously described to fit senescence data (Hurtado et al. 2012b). A smoothed generalized linear model was used to estimate smooth curves for the development of flowering and senescence over time, pooling all the observations per genotype, per time point. The estimation was done using the R software environment (CoreTeam 2011).

A different approach was used to model plant height. In contrast to flowering and senescence, plant height was measured as a continuous variable (in cm). Up to twelve observations per genotype were available per time point. We pooled the 12 observations per genotype, per time point and fitted a curve to the relationship between plant height and PBTT. A smooth expectile curve was well suited for this purpose and the expectiles were estimated using least asymmetrically weighted squares (Schnabel and Eilers 2009). They were combined with P-splines to provide a flexible functional form(Schnabel et al. 2012). This modelling procedure resulted in a smooth frontier curve to describe the upper potential of plant height over time. For the calculations we used the package “expectreg” in R (Sobotka et al. 2012).

Curve parameters: New phenotypic traits

Once the curves were fitted, parameters describing the aging and growth processes were estimated. We used the first and second derivative of the fitted curves to characterise the processes as it was previously done to characterise senescence in potato under long day length conditions (Hurtado et al. 2012b). Those parameters facilitated the study of development and growth as continuous time-dependent processes by breaking down the complexity of such traits into components related to the different phenological stages. The mean slope reflected the average rate of change of each trait during the observation period and explained how fast a genotype experienced each process. In the maximum slope the higher the value, the faster the process was completed. The inflection point is the time point where the growth rate of the curve changes from increasing to decreasing growth rate. It reflects the turning point at which development or growth entered into the final phase. The onset and end indicated the beginning and the final part of the processes in terms of PBTT. The lower the values, the earlier the processes started or ended. We also considered two additional descriptive traits describing growth, maximum and mean plant height. The curve parameters were treated as phenotypic traits with a single value per genotype per environment for subsequent QTL analysis.

Phenotypic relations between traits according to day length and between environments

We estimated the phenotypic correlations between traits in each environment and a principal component biplot was used to visualize these relationships (Gabriel 1971). The data were normalized and only genotypes with complete information for all the traits were included in the analysis.

The relationship between environments was investigated for each parameter derived from the curves and for the agronomic traits. Correlations between environments were estimated for each trait, GGE biplots were used to visualise such correlations and the presence of Genotype by Environment Interactions (GEI). Low correlation between environments indicated GEI and negative correlations indicated even strong GEI.

Highly discriminant environments were also identified for each trait based on the standard deviation (sd) estimated for each trait using BLUEs or single values per genotype for the parameters. The higher the sd the most discriminant the environment.

Multi-environment QTL analysis

The multi-environment data analysis consisted of three steps carried out using the QTL library of Genstat 14 (VSN 2011) and was based on the use of REML (residual maximum likelihood) procedures. In the first step, we fitted a phenotypic mixed model to genotype-by environment data and we selected the best variance–covariance (VCOV) model for each trait based on the Schwarz (Bayes) Information Criterion (SIC). The best model was selected among seven VCOV models allowing heterogeneity of genetic variances across individual environments and heterogeneity of genetic correlations between pairs of environments (Boer et al. 2007). In the second step, the best model was used in a multi-environment QTL analysis per trait including repeated genome scans for the detection of QTL effects (Boer et al. 2007; Malosetti et al. 2008). The QTL analysis was done using the parental linkage maps previously described and used in a multi-trait QTL analysis performed on the CxE population (Hurtado et al. 2012a). An initial genome-wide scan to identify candidate QTL positions was done by simple interval mapping (SIM) followed by one round of multi-environment composite interval mapping (CIM) to scan for QTL effects in the presence of co-factors genome-wide. In the third step backward selection of markers representing candidate QTLs was performed iteratively to fit the final QTL model at each chromosome position. The procedure was checking whether in the combined model all QTLs were significant and, if not, the QTL with the lowest Wald test statistic was dropped from the model. This process was repeated until all the selected loci in the remaining model were significant. The procedure then tested whether the remaining QTLs show significant “QTL-by-Environment” Interaction (QEI), by breaking down the QTL effects into QTL main effects and QEI effects. If the QEI term was not significant, only a main effect was retained in the model for the corresponding QTL. The associated significance p-values (on a -log10 scale) were plotted to visualize the results along the chromosomes. The threshold for QTL detection was calculated using the Li & Ji method (Li and Ji 2005) with a genome-wide significance level of 0.05.

Results

Effects of temperature and photoperiod on development and growth

The differences in daily air temperature between the five experiments (Figure 1A) were reduced by incorporating it into the temperature effect function. However the cumulative

PBTT showed differences between the patterns according to day length (Figure 1B). Under short photoperiod, the fluctuations in temperatures were relatively small during the observation period which is reflected in the nearly linear cumulative PBTT. However under very long days the temperatures depended on the season starting and ending with low values, with an increase in values in the middle of the observation period. These increases and decreases in daily air temperature are reflected in a nearly S-shaped curve for the relation between PBTT and DAP with the highest profit of heat accumulation in the middle of the growing season.

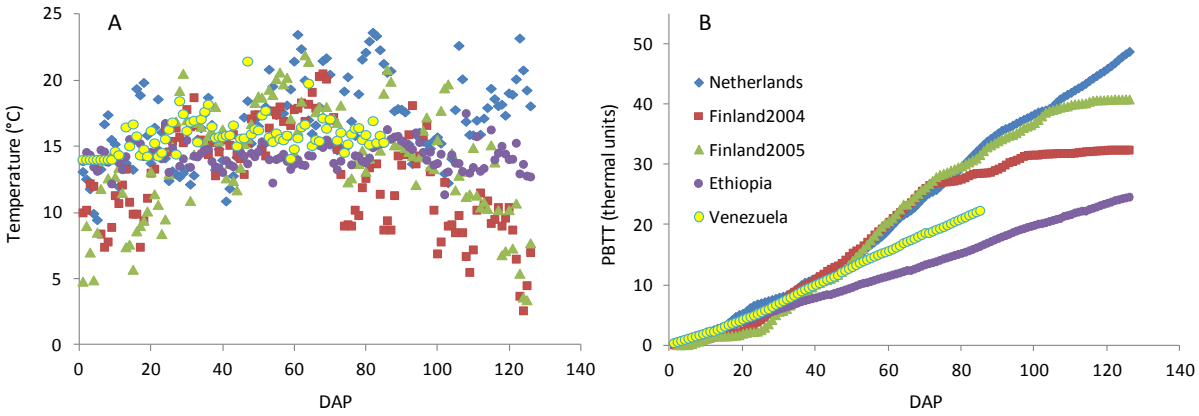


Figure 1. Fluctuations of daily air temperature (A) and cumulative PBTT (B) during the observation period (DAP) of multi-environment potato field trials in Ethiopia, Venezuela, The Netherlands, and Finland (2004 and 2005).

The general effects of temperature and photoperiod on development and growth of the CxE population can be observed in Figure 2. Flowering, senescence and plant height were differentially affected by both environmental factors during the growing season as we will explain below.

In our MET study the ageing process was extended and the life span was increased under long days with high temperatures (Figure 2A). In the Netherlands a longer period of high temperature during the growing season was observed (Figure 1a) and it was the environment with the highest accumulation of BPTT (Figure 1b). In this environment, the senescence period was extended in the whole population and even late genotypes completed only half of the process. A highly negative correlation was observed between onset of senescence and the slope of the senescence curves (Additional file 2), meaning a late start of the yellowing process with an extended ageing period.

On the other hand, short photoperiod accelerated development, as can be observed in the progression of flowering (Figure 2B) and senescence in Ethiopia. All genotypes completed both processes within the observation period and the curves in the whole population were closer to each other compared with the other environments. The flowering process in the Netherlands was similar to the process in Ethiopia probably due to the high temperatures in the Netherlands right in the middle of the growing season promoting development.

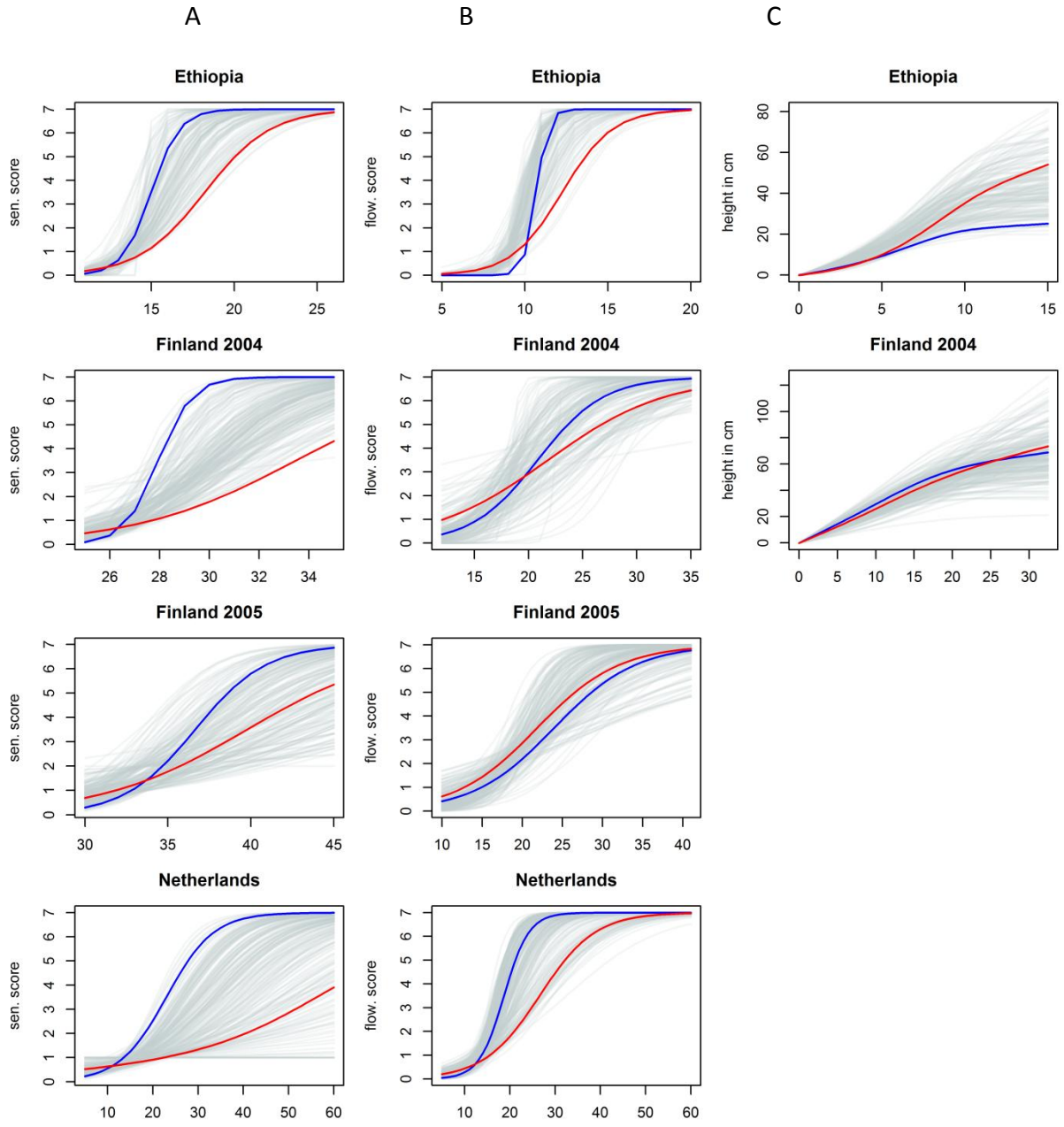


Figure 2. Senescence (A), flowering (B) and plant height (C) fitted curves for 200 genotypes evaluated in Ethiopia, Finland (2004 and 2005) and The Netherlands. Two genotypes were randomly selected from the early (blue) and late (red) maturity groups according to the Netherlands trial (1999) and their fitted curves are they are highlighted in the figure.

It is interesting to notice that early onset of flowering in Ethiopia and the Netherlands was related to fast progression of the flowering process (negative correlations in additional file 2). Whereas early onset of flowering in Finland was related to small values of the slope, so the flowering process was extended even when the process started early.

Plant growth was stimulated under long days of Finland and taller plants were observed than in Ethiopia (Figure 2C), where short days seemed to restrict growth. The relationship between maturity and plant height was not very clear, as is shown with the two highlighted genotypes in Figure 2C. However it was clear that steep curves (fast growth) were related to tall plants with high positive correlation between the slope and maximum plant height in both environments (Additional file 2).

Relationship between traits

Development and growth

Developmental processes are defined in our study in terms of biological parameters describing the progression of flowering and senescence. The results showed that in each environment both processes were strongly related and the parameters describing each developmental phase were positively correlated (Additional file 2). For instance, slow progression of flowering was related to a slow ageing process irrespective of day length. An example of such a relationship is shown in the plot based on traits measured in Ethiopia (Figure 3). The trait vectors Fmax (maximum progression rate of flowering) and Smax (maximum progression rate of senescence) have a small angle, which indicates a strong positive correlation. The strongest correlation between Fmax and Smax was observed in the Netherlands (Additional file 2).

On the other hand, a negative correlation between progression rate of growth and progression rate of development was observed in Ethiopia. It indicates that slow growth, leading to short plants, was related to fast flowering and senescence. In Figure 3, the trait vectors Fmax and Smax are going in opposite directions compared to Phmax (progression rate of plant height), indicating a strong negative correlation between them.

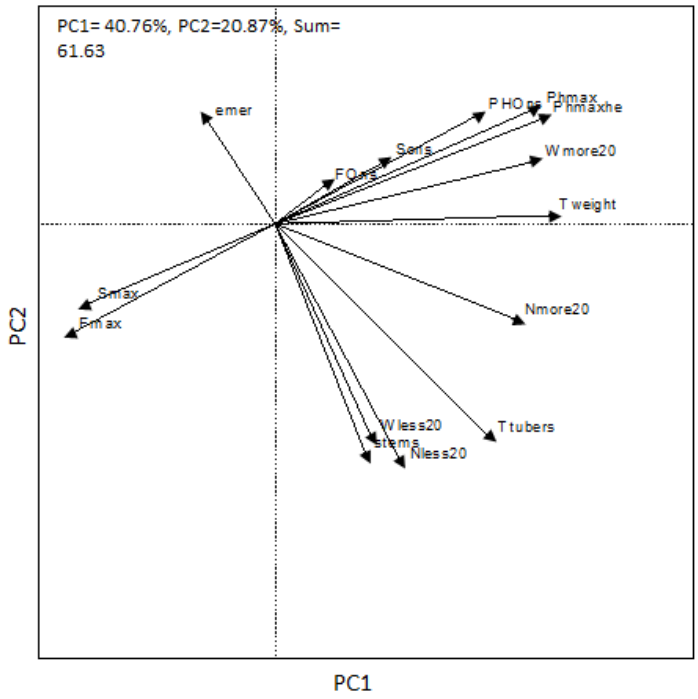


Figure 3. Principal components plot showing the relationship between agronomic, developmental and growth traits evaluated under field conditions in Ethiopia. Each arrow represents a trait vector. Agronomic traits: Tweight=total tuber weight; Wmore20/Wless20=weight of tubers bigger/smaller than 20mm; Ttubers=Total tuber number; Nmore20/Nless20=number of tubers bigger/smaller than 20mm; stems=number of main stems; emer=emergence in DAP. Developmental traits: Smax/Fmax=maximum progression rate of senescence/flowering; Sons/Fons=Onset of senescence/flowering. Growth traits: PHons=onset of plant height; PHmax= maximum progression rate of plant height; PHmaxhe=maximum plant height.

Agronomic and development traits

Some important associations between agronomic, developmental and growth traits were identified across environments and other relations were found dependent on day length.

Slow progression of senescence was associated to high number of tubers in all the environments. In Ethiopia and Finland 2004, growth also showed an important association to senescence and total number of tubers. Tall plants with fast growth but slow senescence progression showed higher total number of tubers (especially tubers bigger than 20mm) at harvest. The number of tubers also increased when more main stems were present per genotype in Ethiopia and Finland 2005. An example of this relationship is shown in Figure 3 in which the vectors Ttubers and stems have a small angle. Early emergence was related to early onset of flowering in Ethiopia and Finland 2005. The contrasting environments showed almost the same correlation between the traits.

The differences in assessment of tuber weight and tuber size estimates between environments made it difficult to compare across day lengths. However, the correlation of normalized traits per environment allowed us to make some conclusions about the relationship between tuber weight and tuber size and the other agronomic traits and then linked the conclusions across environments. Under long and very long days, fast progression of flowering with an early end of this developmental process was associated to higher tuber weight, whereas in Ethiopia a long flowering (and also senescence) process was required to have higher tuber weight (especially for tubers bigger than 20mm). In the Netherlands and Finland 2005 fast senescence was associated to high tuber weight. On the other hand, photoperiod had an influence on the relationship between tuber size and total tuber weight. In Venezuela bigger tubers were related to higher tuber weight, while in Finland 2004 bigger tubers were associated to relative lower tuber weight.

Relationship between environments (Genotype by Environment Interactions)

Each trait was studied separately to identify the relationships between environments (Table 2) and GGE biplots were used to visualize these relationships (two examples are shown in Figure 4). For each trait the most discriminant environment was identified based on the standard deviation per environment (underlined values in Table 2).

In Table 2 the correlations between environments are highlighted in grey. Although low correlation indicates GEI, negative correlations suggest even stronger GEI and they are highlighted in bold. GEI was identified in some agronomic traits and in only one of the developmental parameters. The strongest GEI was observed in total tuber weight, followed by weight of tubers smaller than 20mm. Total tuber weight was measured in all environments but as was mentioned earlier large differences were observed between the values due to differences in measurement units per environment. The GGE biplot for total tuber weight (Figure 4A) was done using the normalized data per environment and it showed the Netherlands as the most discriminant environment with the longest vector given by the high standard deviation (Table 2, highest sd per trait are underlined).

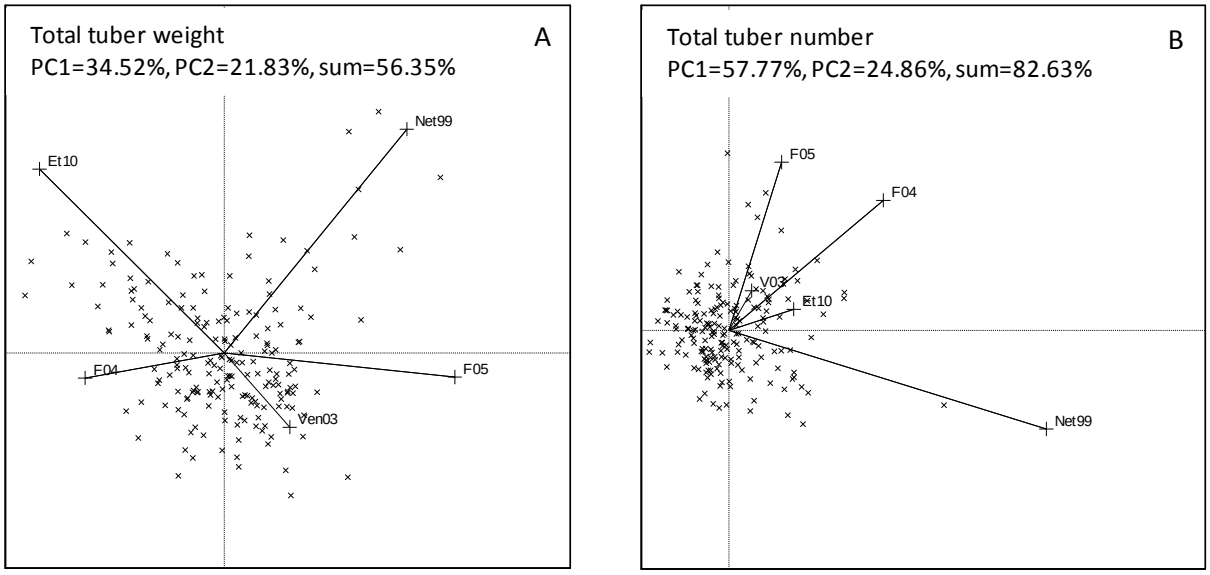


Figure 4. GGE biplot for total tuber weight (A) and total number of tubers (B). The genotypes are represented by x and each line represents an environment vector (Et10: Ethiopia 2010, V03: Venezuela 2003, Net99: Netherlands 1999, F04: Finland 2004 and F05: Finland 2005).

Small negative correlations, indicating strong GEI, were found between the environments for the traits tuber size, number of main stems, total number of tubers (Figure 4B) and onset of flowering. Note that these traits were not measured in all environments. No evidence of GEI was observed for the parameters describing senescence and plant height, except for onset of senescence with low correlations between Finland 2004 and Ethiopia (Table 2).

Genetic bases of agronomic, developmental and growth traits in MET

Our study was mainly focused on the presence and positions of QTLs identified in a multi-environment QTL setup. The QTL effects related to different values of the phenotypic traits were estimated for each environment on each QTL position. Figure 5 is an example of the multi-environment composite interval mapping performed on each trait. In this case, we show the QTL results for total number of tubers and the QTL effects on each environment. QTL by Environment Interactions, QEI, can be observed (e.g. QTL on E5), as well as stable-QTLs across environments (e.g. QTL on C3) and environment-specific QTLs (e.g. QTL on C8).

Table 2. Description of agronomic, developmental and growth traits evaluated in each environment. The description includes mean, standard deviation (sd) and range of values per environment. Correlations between environments are highlighted in gray. The highest sd per trait is underlined and it represents the most discriminant environment. Negative correlations are in bold and they reflect strong genotype by environment interactions.

		Environments				
Trait		**Et10**	**F04**	**F05**	**Net99**	**V03**
Total tuber number	Mean	13.53	28.38	18.15	40.74	16.03
	stdev	5.59	11.60	10.42	<u>17.11</u>	8.54
	range	3-36	9-63	1-52	9-135	3-45
	F04	0.686				
	F05	0.300	0.460			
	Net99	0.514	0.486	**-0.004**		
	V03	0.246	0.244	0.051	0.032	
Number of tubers < 20mm	Mean	5.52	7.03		21.18	
	stdev	3.78	5.94		<u>10.69</u>	
	range	1-27	1-29		4-86	
	Net99	0.47	0.46			
	Et10		0.42			
Number of tubers > 20mm	Mean	8.70	9.70		19.19	
	stdev	2.87	6.45		<u>8.12</u>	
	range	3-18	1-34		1-49	
	Net99	0.20	0.22			
	Et10		0.43			
Total tuber Weight*	Mean	248.46	2672.67	186.68	1523.98	121.22
	stdev	92.44	843.47	120.67	<u>1403.10</u>	108.43
	range	46-567	1726-6449	10-471	11-6590	8-975
	F04	0.356				
	F05	**-0.151**	**-0.255**			
	Net99	**-0.037**	**-0.164**	0.376		
	V03	**-0.121**	**-0.041**	0.024	0.029	
Weight of tubers < 20mm	Mean	28.30	220.86			
	stdev	20.65	<u>122.31</u>			
	range	8-132	20-682			
	F04	**-0.280**				

		Environments				
Trait		**Et10**	**F04**	**F05**	**Net99**	**V03**
Weight of tubers > 20mm	mean	229.51	53.24			
	stdev	<u>88.08</u>	62.70			
	range	85-537	10-281			
	F04	0.195				
Tuber size*	Mean		30.02		2.53	1.91
	stdev		<u>4.21</u>		0.51	0.53
	range		20-41		0.5-4	6-36
	F04				**-0.06**	0.29
	Net99					**-0.02**
Date of emergence (DAP)	Mean	15.59		29.40		
	stdev	1.79		<u>4.37</u>		
	range	11-21		23-46		
	Et10			0.231		
Plant height (cm)	Mean			45.00	62.54	32.24
	stdev			10.82	<u>35.14</u>	6.81
	range			18-72	28-193	18-48
	F05				0.03	0.21
	Net99					0.22
Number of main stems	mean	4.73	3.37	3.87	4.15	2.61
	stdev	1.30	1.28	<u>1.47</u>	1.19	1.42
	range	2-9	1-8	1-11	1-7	1-8
	F04	0.324				
	F05	0.276	0.217			
	Net99	0.420	0.213	0.130		
	V03	0.098	0.131	0.065	**-0.063**	

Trait			Environments			
			Et10	F04	F05	Net99
Flowering	Mean	Mean	0.32	0.28	0.20	0.12
		s.d.	0.002	0.018	0.028	0.003
		range	0.28-0.32	0.23-0.30	0.11-0.22	0.11-0.13
		Fin04	0.407			
		Fin05	0.594	0.475		
		Net99	0.459	0.405	0.647	
	Max	Mean	1.56	0.62	0.46	0.52
		s.d.	0.51	0.36	0.21	0.19
		range	0.4-3.7	0.16-2.6	0.11-1.03	0.19-0.93
		Fin04	0.372			
		Fin05	0.675	0.419		
		Net99	0.702	0.488	0.685	
	ipoint (PBTT)	Mean	15.49	21.18	22.69	21.06
		s.d.	1.04	2.42	2.34	3.47
		range	14-20	17-29	18-31	16-30
		Fin04	0.676			
		Fin05	0.517	0.559		
		Net99	0.639	0.703	0.536	
	onset (PBTT)	Mean	14.28	17.38	17.06	16.28
		s.d.	0.75	2.30	2.24	1.62
		range	12-17	14-27	12-24	13-22
		Fin04	0.456			
		Fin05	0.016	0.159		
		Net99	0.331	0.294	**-0.081**	
	end (PBTT)	Mean	17.76	26.42	29.51	26.85
		s.d.	1.59	3.80	5.07	5.73
		range	15-23	20-35	21-41	21-43
		Fin04	0.734			
		Fin05	0.709	0.783		
		Net99	0.712	0.764	0.757	

Trait			Environments			
			Et10	Fin04	Fin05	Net99
senescence	mean	Mean	0.28	0.59	0.31	0.10
		s.d.	0.01	0.09	0.10	0.04
		range	26-0.29	0.12-0.7	0.08-0.45	0.003-0.14
		Fin04	0.474			
		Fin05	0.467	0.541		
		Net99	0.617	0.577	0.678	
	max	Mean	0.95	1.03	0.45	0.17
		s.d.	0.38	0.43	0.21	0.08
		range	0.37-2.8	0.12-2.3	0.04-0.96	0.004-0.34
		Fin04	0.477			
		Fin05	0.607	0.581		
		Net99	0.585	0.593	0.689	
	ipoint (PBTT)	Mean	24.72	30.56	39.90	43.29
		s.d.	1.94	1.25	3.24	11.32
		range	17-29	28-34	34-45	25-60
		Fin04	0.500			
		Fin05	0.524	0.538		
		Net99	0.703	0.545	0.696	
	onset (PBTT)	Mean	22.53	28.44	34.34	29.70
		s.d.	1.16	0.80	1.77	8.61
		range	19-25	27-30	32-40	16-49
		Fin04	0.036			
		Fin05	0.124	0.343		
		Net99	0.563	0.086	0.196	

Trait			Environments	
			Et10	F04
Plant height	mean	Mean	0.92	1.44
		stdev	0.27	0.34
		range	0.4-1.6	0.7-2.4
		F04	0.475	
	max	Mean	1.53	1.98
		stdev	0.46	0.32
		range	0.6-2.8	1.3-3.1
		F04	0.202	
	ipoint (PBTT)	Mean	7.25	10.65
		stdev	1.20	5.48
		range	3.9-15	2-24
		F04	0.506	
	onset (PBTT)	Mean	4.20	5.81
		stdev	1.56	3.32
		range	2.1-7.6	1.3-15.2
		F04	0.420	
	end (PBTT)	Mean	10.21	20.52
		stdev	1.29	4.44
		range	7-13.4	9.9-29
		F04	0.371	
	max height (cm)	Mean	43.70	61.55
		stdev	11.91	13.45
		range	20-78	31-100
		F04	0.452	
	mean height (cm)	Mean	22.12	33.41
		stdev	5.79	8.05
		range	11-36	17-79
		F04	0.428	

Environments: Ethiopia 2010 (Et10), Finland 2004 (F04), Finland 2005 (F05), Netherlands 1999 (Net99), Venezuela 2003 (V03).
**: Total tuber weight and Tuber size were assessed with different scales in each environment.*

The multi-environment QTL analyses allowed the identification of QEI (e.g. QTL on E5, Figure 5), pleiotropic genetic regions, stable-QTLs across environments (e.g. QTL on C3, Figure 5), environment-specific QTLs (e.g. QTL on C8, Figure 5), and time dependent QTLs for development and growth processes.

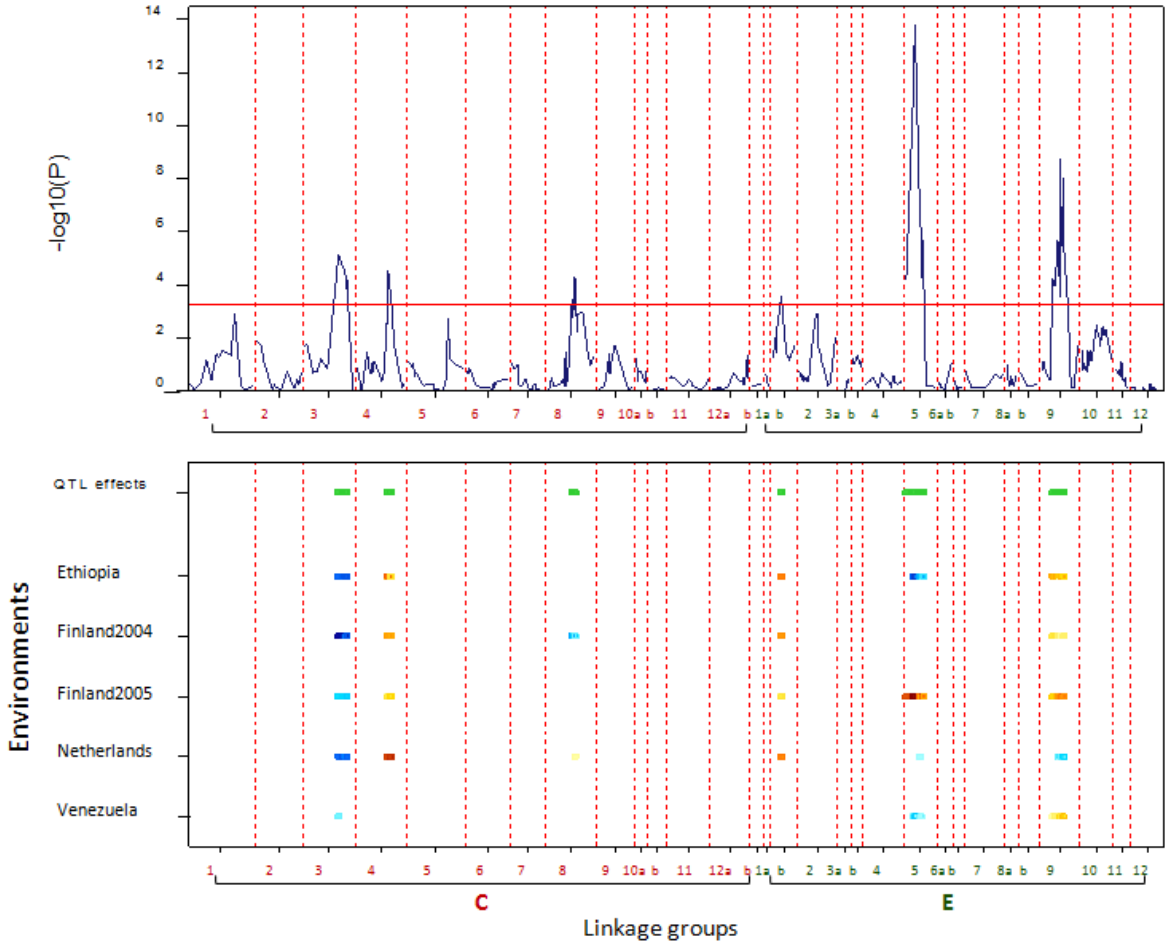


Figure 5. Graphical display of the results of multi-environment composite interval mapping for total number of tubers. The upper plot shows the significance of the QTLs (-log10 scale of the associated probability value). The lower plot shows the positive (red) and negative (blue) allele substitution effects at positions where there was a significant QTL. The intensity of the colour is proportional to the QTL effect size (the darker the larger the effect).

Table 3 contains a summary of the results of the multi-environment QTL analysis performed for all the traits. Genetic regions where QTLs were detected were defined to span maximum 20cM. Significant QTLs in each environment are highlighted in bold. The signs of the QTL effects are differentiated with colours and they can be compared across traits (along the rows). Positive effects indicate higher values of the traits, while a negative sign is related to low values.

The QTLs were checked to determine whether they were equal to QTLs detected on the homologous linkage group of the other parent. Only one QTL was detected in the same genetic region in both parents, C5 and E5 located at 105 and 21.9 cM respectively (Hurtado et al. 2012a). Be aware that the molecular markers on C5 have inverted order as compared to E5.

After identification of the QTLs, the presence of QEI was determined. Cells highlighted in grey indicate the presence of significant QEI and the effects of the QTLs were estimated for individual environments. QTL effects for different environments can go in the same or opposite direction (e.g. QTLs on C3 and E9 respectively, Figure 5), indicated by different font colour within a grey cell (Table 3). We notice that the presence of GEI (negative correlations between environments shown in Table 2) points to QTL effects with opposite directions when there was significant QEI. As an example the total number of tubers will be considered: the small negative correlation between Finland 2005 and The Netherlands (Table 2) was an indication of the presence of GEI. Five out of six QTLs for this trait showed QEI and the effects of the QTLs on linkage groups E5 and E9 had opposite signs in Netherlands and Finland 2005 as indicated by different colours (Table 3 and Figure 5).

Quantitative nature of development, growth and agronomic traits

The number of QTLs associated with the parameters describing senescence and plant height progression showed the quantitative nature of these traits (Table 3C and 3D). Apart from the inflection point and the end of plant height, all the parameters gave rise to at least two QTLs located in different genetic regions.

As expected, complex agronomic traits were also explained by at least two QTLs. The QTL on chromosome 5 related to flowering was associated with all traits except for emergence and number and weight of tubers smaller than 20mm. However, it was not the only QTL with high explained variance per trait. For instance, the QTLs on C3 and E6 explained 13.4 and 7.8% of the variance in Finland 2004 for number of main stems and total number of tubers, respectively.

Our results showed QEI in all the traits except for onset of plant height (PHons) and average height (PHmehe). The strongest QEI was detected for onset of flowering, total number of tubers and total tuber weight. Finland 2005 and the Netherlands were the most contrasting environment for onset of flowering. In Finland 2005 late onset of flowering was related to higher number of tubers and higher tuber weight. In contrast, in The Netherlands early onset of flowering was related to short plants producing more tubers bigger than 20mm and a higher total tuber weight.

Pleiotropic regions

The main pleiotropic region was detected on E5 between 18 - 25cM and it was associated with all flowering, senescence and plant height parameters (except for onset of senescence) and to some of the agronomic traits (total number of tubers and tuber weight, number and weight of tubers bigger than 20mm). This region showed a major effect on development, growth and agronomic traits with explained variances going up to 58% (inflection point of senescence).

Other pleiotropic regions associated to developmental or growth traits and agronomic traits were observed on C4, C5 and C10. The pleiotropic region on C4 between 72 – 78cM showed an effect on onset of senescence and total number of tubers (small and big tubers). This QTL associated to late onset of senescence in Finland (2004 and 2005) was also related to higher number of tubers in these environments. The pleiotropic region on C5 (66-69cM) had an effect on onset of senescence and weight of tubers smaller than 20mm. In Ethiopia early onset of senescence was related to higher weight of tuber smaller than 20mm. In this case, probably short photoperiod stimulated the beginning of the yellowing process and the bulking of the small tubers. The pleiotropic region on C10 between 12 – 25cM was associated with early emergence, fast growth and more tubers bigger than 20mm in Ethiopia.

Photoperiod-independent QTLs

For some traits consistent QTLs across environments were detected. Considering that not all the traits were evaluated in all the environments, the most interesting QTLs are the ones in which there were no significant differences between QTL effects in contrasting

environments. For instance, if we compare plant height in Ethiopia and Finland (2004), the most interesting parameter was onset of plant height because QEI was no detected, neither at E5 nor at E12. E5 was related to early onset of growth in both environments and it showed the highest explained variance in Ethiopia (Table 3D).

The QTL effects of C2, C5 (105-125cM) and E8 in short and long days were almost the same for inflection point and maximum progression rate of senescence. C2 and C5 were related to fast senescence development in all the environments, while E8 was associated to slow ageing process (Table 3C).

For each agronomic trait, at least one of the multiple QTLs detected did not show QEI except for plant height and tuber size showing significant QEI for all the detected QTLs.

Environment-specific QTLs

A higher number of environment-specific QTLs was detected under short photoperiod than under long days for agronomic and growth traits. QTLs with significant effects in Ethiopia were detected for emergence (C2,C10), total tuber weight (E1A), number of tubers bigger than 20mm (E5: 18-25cM, E10), weight of tubers smaller than 20mm (C5: 66.6cM), onset of plant height (E12) and mean height (C1) with explained variances going up to 14%. In Venezuela QTLs with significant effect on tuber size were detected on C11 and E5 (18-25cM).

Time dependent QTLs for development and growth processes.

Time-dependent QTLs were identified in the growth process. For instance, QTLs on E12, C10 and E6 were related to onset, maximum progression rate and maximum plant height respectively (Table 3D). They pop up under contrasting day lengths (Ethiopia and Finland 2004) only in specific growth phases.

The senescence process also showed time-dependent QTLs for onset and inflection point (Table 3C). Although onset of senescence was the developmental phase with the higher number of QTLs, there were only significant effects under long days. For instance, early onset of senescence in The Netherlands and Finland 2004 was related to a second QTL on C5 (69.77cM) and a QTL on C9, respectively.

Regarding inflection point of senescence, a QTL on C2 was related to early entry into the final phase of senescence under contrasting day lengths (Ethiopia and Finland 2004), while a QTL on E4 was related to the same process only under long days (Netherlands and Finland 2004).

Trait-specific QTLs

The QTLs associated to only one trait showed significant effects in only one or two environments. The most interesting QTLs, because of the significant effects on contrasting environments, were detected on C1, E6 and E9 associated to tuber size, number of main stems and weight of tubers smaller than 20mm respectively (Table 3A)

Discussion

The analyses of multi-environment trials provides the basis to understand 1) the effect of temperature and photoperiod on development, growth and agronomic traits evaluated under field conditions, 2) GEI playing an important role in the phenological plasticity of the crop and 3) relationships between developmental and agronomic traits.

Temperature and photoperiod are considered major environmental factors controlling development in potato (Ewing and Struik 1992; Levy and Veilleux 2007). In our study developmental processes were defined in terms of biological parameters describing the progression of flowering and senescence. An influence of both environmental factors was observed on the developmental parameters in each environment. Controlled experiments using potato cultivars have shown that high temperatures combined with long days increase the life span of the entire shoot (Struik and Ewing 1995). High temperatures combined with long days in the Netherlands accelerated the flowering process but delayed onset of senescence thereby increasing the life span of the entire shoot. Finland with lower temperatures showed longer periods of flowering and senescence and the same increase in life span. On the other hand, short days promoted early onset and also fast progression of both developmental processes reducing the life span with almost 20 PBTT units compared with long days.

Table 3. Summary of the results from multi-environment QTL analysis performed separately on agronomic (A), developmental (B,C) and growth traits (D). QTL positions (in cM, span maximum 20cM), QTL effects, standard errors and explained variances (%var) are presented for each environment.

A

		Stems			Ttubers			Tweight			emergence			plant height			tuber size			# of tubers <20			# of tubers >20			weigth <20			weight >20		
	Va-covar	CS			FA			FA			HCS			HCS			DIA			FA			HCS			UNS			HCS		
QTL	Env	Effect	s.e	%var	Effect	s.e	%var	Effect	s.e	%var	Effect	s.e	%var	Effect	s.e	%var	Effect	s.e	%var	Effect	s.e	%var	Effect	s.e	%var	Effect	s.e	%var	Effect	s.e	%var
C1: 157.9	F04																-0.05	0.38	0.0												
	Net99																-0.14	0.04	7.2												
	V03																0.19	0.05	12.6												
C2: 0.0	Et10										-0.44	0.13	6.1																		
	F05										-0.44	0.13	1.0																		
C3: 84 - 102	Et10	-0.25	0.06	3.6	-1.68	0.36	9.0													-1.66	0.28	19.1	-0.35	0.22	1.5	-7.12	1.69	11.9	21.47	4.64	5.9
	F04	-0.25	0.06	3.7	-4.25	0.81	13.4										1.31	0.33	9.7	-1.66	0.28	7.8	-2.21	0.57	11.8	43.92	10.96	12.9	21.47	4.64	11.7
	F05	-0.25	0.06	2.8	-2.05	0.72	3.9																								
	Net	-0.25	0.06	4.3	-5.38	1.14	9.9										-0.05	0.04	1.1	-1.66	0.28	2.0	-1.94	0.66	5.7						
	Ven	-0.25	0.06	3.0	-1.74	0.77	4.1										0.06	0.04	1.4												
C4: 78.53	Et10				1.35	0.37	5.9													0.96	0.25	6.5	0.41	0.19	2.1						
	F04				2.60	0.82	5.0													0.96	0.25	2.6	1.25	0.48	3.8						
	F05				1.63	0.74	2.4																								
	Net99				4.88	1.17	8.1													0.96	0.25	0.7	2.64	0.57	10.6						
	V03				-0.22	0.77	0.1																								
C5: 66.68	Et10																									6.63	1.35	10.3			
	F04																									6.63	1.35	0.3			
C5: 101	Et10	0.33	0.10	6.6				-27.34	5.61	8.7																			-25.15	5.61	8.2
	F04	0.40	0.10	9.8				-56.85	69.39	0.5																			1.46	4.88	0.1
	F05	0.36	0.10	5.8				24.62	8.00	4.2				-0.44	0.77	0.2															
	Net	0.27	0.10	5.0				-92.12	103.60	0.4				-13.53	2.15	14.8															
	Ven	-0.12	0.11	0.7				6.45	10.42	0.4				-0.48	0.57	0.5															
C8: 68.29	Et10				0.54	0.36	0.9																								
	F04				-2.29	0.81	3.9																								
	F05				-1.10	0.73	1.1																								
	Net99				2.16	1.17	1.6																								
	V03				-0.39	0.75	0.2																								
C10: 2 - 12	Et10										0.63	0.14	12.4										0.69	0.18	5.8						
	F04																						0.69	0.18	1.1						
	F05										0.63	0.14	2.1																		
	Net99																						0.69	0.18	0.7						
	V03																														
C11: 0.0	F04																0.28	0.35	0.4												
	Net99																-0.01	0.04	0.0												
	V03																-0.19	0.04	12.1												
E1A: 22 - 36	Et10				1.16	0.32	4.3	16.65	4.13	3.2																					
	F04				1.16	0.32	1.0	16.65	4.13	0.0																					
	F05				1.16	0.32	1.2	16.65	4.13	1.9																					
	Net99				1.16	0.32	0.5	16.65	4.13	0.0																					
	V03				1.16	0.32	1.9	16.65	4.13	2.4																					
E5: 0.0	Et10																			0.29	0.25	0.6									
	F04																			-0.20	0.45	0.1									
	Net99																			-2.77	0.82	5.6									
E5: 18 - 25	Et10				-1.59	0.35	8.1	-50.36	5.65	29.7													-1.08	0.18	14.1				-47.36	5.65	28.9
	F04				-0.37	0.80	0.1	-118.32	70.44	2.0							0.57	0.33	1.8				-0.36	0.48	0.3				-6.23	4.96	1.0
	F05				3.94	0.71	14.3	71.18	8.01	34.8				-4.09	0.77	14.3															
	Net99				-1.89	1.13	1.2	313.43	103.90	5.0				-18.35	2.15	27.3	0.00	0.04	0.0				0.57	0.55	0.5						
	V03				-1.82	0.74	4.5	16.65	10.54	2.4				-1.91	0.58	7.8	0.25	0.04	22.4												
E6: 31.9	Et10	-0.07	0.10	0.3																											
	F04	0.36	0.11	7.8																											
	F05	0.19	0.10	1.6																											
	Net99	0.12	0.10	1.1																											
	V03	0.42	0.11	8.7																											
E9: 48 - 56	Et10				1.40	0.36	6.2													0.95	0.26	6.3	0.31	0.19	1.2						
	F04				2.13	0.82	3.4													1.02	0.46	2.9	1.40	0.49	4.7						
	F05				3.10	0.72	8.9																								
	Net99				-3.06	1.16	3.2													-1.25	0.83	1.1	-1.66	0.57	4.2						
	V03				2.44	0.76	8.2																								
E9: 83.17	Et10																									2.92	1.47	2.0			
	F04																									-33.14	9.48	7.3			
E10: 36.43	Et10																						-0.92	0.20	10.2						
	F04																						-0.92	0.20	2.0						
	Net99																						-0.92	0.20	1.3						

B		Flowering														
		Fmean			Fmax			Fip			Fons			Fend		
	Var-covar	FA			FA			FA			HCS			FA		
QTL	Env	Effect	s.e	%var	Effect	s.e	%var	Effect	s.e	%var	Effect	s.e	%var	Effect	s.e	%var
C5: 105 - 125	Et10	**0.001**	**0.001**	**3.3**	**0.175**	**0.031**	**11.9**	**-0.353**	**0.116**	**5.5**	-0.184	0.116	1.7	**-0.499**	**0.132**	**7.2**
	F04	**0.006**	**0.001**	**10.8**	**0.058**	**0.026**	**2.5**	**-0.871**	**0.155**	**13.0**	**-0.426**	**0.183**	**3.4**	**-1.316**	**0.206**	**12.0**
	F05	**0.012**	**0.002**	**19.2**	**0.069**	**0.012**	**10.6**	**-0.724**	**0.141**	**9.6**	**0.516**	**0.158**	**5.3**	**-2.142**	**0.252**	**17.8**
	Net99	**0.001**	**0.000**	**20.4**	**0.074**	**0.009**	**15.4**	**-1.330**	**0.181**	**14.7**	-0.187	0.109	1.3	**-2.503**	**0.287**	**19.1**
E5: 18 - 25	Et10	**0.000**	**0.001**	**0.5**	**0.332**	**0.032**	**42.6**	**-0.419**	**0.118**	**7.8**	**-0.136**	**0.116**	**0.9**	**-0.863**	**0.135**	**21.4**
	F04	**0.009**	**0.001**	**26.4**	**0.196**	**0.026**	**28.7**	**-1.301**	**0.157**	**28.9**	**-0.230**	**0.182**	**1.0**	**-2.621**	**0.208**	**47.6**
	F05	**0.017**	**0.002**	**34.9**	**0.143**	**0.012**	**44.6**	**-1.363**	**0.141**	**33.9**	**0.762**	**0.156**	**11.6**	**-3.526**	**0.253**	**48.3**
	Net99	**0.001**	**0.000**	**26.4**	**0.132**	**0.009**	**49.1**	**-2.268**	**0.182**	**42.8**	**-0.778**	**0.108**	**23.0**	**-3.787**	**0.288**	**43.7**

C		Senescence											
		Smean			Smax			Sip			Sons		
	Var-covar	FA			FA			FA			HCS		
QTL	Env	Effect	s.e	%var	Effect	s.e	%var	Effect	s.e	%var	Effect	s.e	%var
C2: 45.08	Et10							**0.303**	**0.071**	**2.5**			
	F04							**0.303**	**0.071**	**5.8**			
	F05							0.303	0.071	0.9			
	Net99							0.303	0.071	0.1			
C4: 72.43	Et10										0.151	0.057	0.7
	F04										**0.151**	**0.057**	**3.6**
	F05										**0.151**	**0.057**	**0.7**
	Net99										0.151	0.057	0.0
C5: 0.0 - 20.0	Et10										-0.288	0.056	2.5
	F04										**-0.288**	**0.056**	**13.1**
	F05										**-0.288**	**0.056**	**2.6**
	Net99										-0.288	0.056	0.1
C5: 69.77	Et10										-0.090	0.137	0.3
	F04										0.138	0.057	3.0
	F05										0.056	0.128	0.1
	Net99										**-2.259**	**0.430**	**6.9**
C5: 105 - 125	Et10	0.000	0.002	0.0	**0.029**	**0.003**	**0.6**	**-0.670**	**0.106**	**12.0**			
	F04	**0.032**	**0.006**	**11.8**	**0.029**	**0.003**	**0.4**	**-0.205**	**0.088**	**2.7**			
	F05	**0.026**	**0.006**	**7.2**	**0.029**	**0.003**	**1.9**	**-0.819**	**0.196**	**6.4**			
	Net99	**0.017**	**0.002**	**18.1**	**0.029**	**0.003**	**14.4**	**-3.969**	**0.503**	**12.3**			
C9: 44.88	Et10										0.312	0.140	3.0
	F04										**-0.183**	**0.058**	**5.3**
	F05										-0.121	0.130	0.5
	Net99										0.181	0.438	0.0
E4: 81.11	Et10							-0.219	0.105	1.3			
	F04							**0.326**	**0.087**	**6.8**			
	F05							0.130	0.194	0.2			
	Net99							**1.002**	**0.497**	**0.8**			
E5: 18 - 25	Et10	0.001	0.002	0.1	**0.207**	**0.026**	**28.9**	**-1.264**	**0.107**	**42.6**	**-0.461**	**0.140**	**6.5**
	F04	**0.050**	**0.006**	**28.6**	**0.232**	**0.029**	**28.5**	**-0.570**	**0.089**	**20.7**	0.018	0.058	0.1
	F05	**0.061**	**0.006**	**39.2**	**0.136**	**0.013**	**41.1**	**-1.956**	**0.196**	**36.5**	**-0.366**	**0.130**	**4.3**
	Net99	**0.028**	**0.002**	**45.8**	**0.053**	**0.003**	**49.4**	**-8.639**	**0.502**	**58.2**	**-5.944**	**0.436**	**47.6**
E6A: 0.0	Et10				0.025	0.026	0.4						
	F04				**-0.116**	**0.029**	**7.2**						
	F05				-0.008	0.013	0.1						
	Net99				**-0.009**	**0.004**	**1.5**						
E8: 10 - 15	Et10	0.001	0.002	0.1	**-0.018**	**0.004**	**0.2**	**0.412**	**0.124**	**4.5**			
	F04	**-0.012**	**0.007**	**1.6**	-0.018	0.004	0.2	0.034	0.102	0.1			
	F05	0.002	0.007	0.0	-0.018	0.004	0.7	-0.124	0.228	0.1			
	Net99	**-0.010**	**0.002**	**6.4**	**-0.018**	**0.004**	**5.5**	**1.961**	**0.574**	**3.0**			

D		Plant height																		
		Phmean			Phmax			Phip			Phons			Phend			Phmehe			Phmaxhe		
	Var-covar	HCS			HCS			UNS			UNS			UNS			UNS			CS		
QTL	Env	Effect	s.e	%var	Effect	s.e	%var	Effect	s.e	%var	Effect	s.e	%var	Effect	s.e	%var	Effect	s.e	%var	Effect	s.e	%var
C1 : 7.66	Et10																**1.053**	**0.312**	**3.3**			
	F04																1.053	0.312	1.7			
C5: 0.0 - 20.0	Et10	**-0.068**	**0.015**	**6.4**	**-0.139**	**0.029**	**9.1**															
	F04	-0.068	0.015	4.1	**-0.034**	**0.026**	**1.1**															
C5: 105 - 125	Et10				**-0.111**	**0.026**	**5.8**										**-1.281**	**0.310**	**4.9**			
	F04				0.019	0.024	0.3										**-1.281**	**0.310**	**2.5**			
C10: 25.3	Et10				0.105	0.023	5.2															
	F04				**0.105**	**0.023**	**10.5**															
E5: 18 - 25	Et10	**-0.181**	**0.015**	**45.7**	**-0.304**	**0.026**	**44.0**	**-0.651**	**0.090**	**23.8**	**-0.774**	**0.099**	**24.7**	**-0.511**	**0.093**	**15.6**	**-3.507**	**0.314**	**36.7**	**-7.640**	**0.841**	**41.2**
	F04	**-0.122**	**0.023**	**13.3**	-0.037	0.023	1.3	**-1.589**	**0.395**	**8.4**	**-0.774**	**0.099**	**5.4**	**-1.824**	**0.309**	**16.9**	**-3.507**	**0.314**	**19.0**	**-4.619**	**0.822**	**11.8**
E6: 2.94	Et10	**0.067**	**0.015**	**6.3**																**3.405**	**0.694**	**8.2**
	F04	**0.067**	**0.015**	**4.0**																**3.405**	**0.694**	**6.4**
E12: 16.6	Et10										**0.392**	**0.100**	**6.3**									
	F04										0.392	0.100	1.4									

QTL effects are differentiated by colours (red=positive, blue=negative) and main effects in each environment are highlighted in bold. Cells in grey indicate significant QTL by environment interaction. Var-covar: Variance-covariance models: CS= compound symmetry, FA= Factor Analytic Order 1, HCS= Uniform co-variances with unequal variances, DIA= Diagonal, UNS= Unstructured

The senescence scale was used in the Netherlands to classify the CE genotypes in groups according to plant maturity (Celis-Gamboa 2002). Two genotypes randomly selected from the early and late maturity groups were used to compare flowering and senescence processes in the different environments. The plant maturity of the selected genotypes was stable across environments. The early genotype showed faster progression and earlier end of flowering and senescence than the late one in Ethiopia, Finland 2004 and the Netherlands. This suggests that there is not much GEI in terms of maturity and only small shifts are observed across environments. For instance early flowering genotypes tend to be also earlier under long photoperiod and/or high temperatures.

Plant growth seemed to be restricted by short days and smaller plants were observed in Ethiopia than in Finland. Controlled experiments have shown that the development of the haulm is restricted under short day length (Maris 1964). Despite the differences between plant heights in both environments, the growth process under long and short days showed similar progression (high positive correlation between the average progression rates). It means parallel growth curves per genotype under contrasting environments without much GEI. The absence of GEI for the different stages of development and growth indicates good adaptation of the CE genotypes to different environments.

Few indications of GEI were detected in our study. The strongest GEI was found for total tuber weight, which is one of the most important yield determinants (Lemaga and Caesar 1990). GEI is the main constraint in potato breeding, which makes it difficult to use varieties worldwide. In our study we observed positive correlations between total tuber weight in the Netherlands and Finland 2005 indicating similar performance of the CE genotypes under long photoperiod with high temperatures during the growing season. Similar tuber weight was expected comparing Ethiopia and Venezuela but strong GEI was detected between them. This is maybe due to the short observation period in Venezuela. The final harvest was done very early and probably tuber bulking was just starting and low tuber weights were recorded for all the CE genotypes.

Strong GEI was found for tuber size, number of main stems and total number of tubers. A positive correlation between number of main stems and number of tubers has been reported under controlled conditions (Lemaga and Caesar 1990). In our study this correlation was observed across all the environments.

The small GEI observed in some agronomic traits and onset of flowering and the absence of GEI in senescence and plant growth give insights about the adaptation of the CE genotypes to different environments. Although a wider range of optimal temperatures for potato growth has been reported in long than in short days (Borah and Ilthorpe 1962; Ingram and McCloud 1984; Manrique 1992; Midmore 1984; Sands and Regel 1983), optimum crop growth has been estimated between 18 and 24°C independently of day length using the LINTUL-POTATO model (Kooman and Haverkort 1995). The wide range of optimal temperatures explains the broad adaptation of potato to a wide range of climates and ensures the good performance of the crop not only under short days and low temperatures, but also in temperate regions during summer as is shown in our study.

Relationships between traits were found based on phenotypic correlations between them in each environment. Although some of the relationships have been already described, we are showing consistency across environments or indications of trait by environment interactions. The identified pleiotropic genetic regions associated to multiple traits can help to understand the correlations between traits. For instance, negative correlation between

progression rate of growth and progression rate of development was observed in all the environments indicating that fast growth, leading to tall plants, was related to slow flowering and senescence processes.

Important association between agronomic, developmental and growth traits were identified under contrasting environments. In Ethiopia and Finland 2004, tall plants with fast growth but slow senescence progression showed higher total number of tubers (especially tubers bigger than 20mm) at harvest. Although the effect of day length on the relationship between number of tubers and other traits, such as average tuber weight and tuber yield, has been previously reported (Lemaga and Caesar 1990; Maris 1964), there were no reports showing the relationship between growth, development and total number of tubers across photoperiods in potato. An interesting relationship under contrasting environments was also found between emergence and onset of flowering. Early emergence was related to early onset of flowering consistently in Ethiopia and Finland 2005.

Trait by environment interactions were also detected. With long and very long days, fast progression of flowering was associated with higher tuber weight. In Ethiopia (with short days) a long flowering (and also senescence) process was required to obtain higher tuber weights (especially in tubers bigger than 20mm). The relationship between tuber size and total tuber weight was probably affected by photoperiod as well. In Venezuela bigger tubers were related to higher tuber weight, while in Finland 2004 bigger tubers were associated to low tuber weight. However the same relationship could not be detected between Venezuela and The Netherlands suggesting that other environmental factors are also playing a role.

QTL by environment interactions: The multi-environment QTL analyses combining MET data with genetic information allowed the identification of 1) QTL by environment interactions, 2) pleiotropic genetic regions, 6) stable-QTLs across environments, 3) environment-specific QTLs and 4) time-dependent QTLs for development and growth processes.

Although QEI was detected for most of the QTLs in our study, only few could be matched to GEI. We noticed that most of the QEI was due to significant differences between the QTL effects going in the same direction and in that case GEI was not detected. On the other

hand, the presence of strong GEI with negative correlations between environments was related to QTL effects with crossover interactions.

The polygenic inheritance of the complex traits included in our study was successfully identified for most of the traits with at least two QTLs associated to each trait in the different environments. On the other hand, all the development stages of flowering were associated with only one QTL on chromosome 5 (C5: 105cM and E5: 21.9 cM) with the highest explained variance given by the E parent. Other QTLs with small effects have been detected in the flowering process using a multi-trait QTL analysis where correlations between traits is also taken into account (Hurtado et al. 2012a). Since we are using a multi-environment QTL approach for each trait, such correlations were not included in our analysis and possibly we are lacking extra power across correlated traits to detect these QTLs with smaller effects.

Pleiotropic genetic regions were identified in different linkage groups and they allowed to make a link between agronomic and development/growth traits. The main pleiotropic region was detected on E5 between 18 - 25cM and it was associated to all flowering, senescence and plant height parameters (except for onset of senescence) and to some of the agronomic traits. QEI was also detected for most of the traits. On the same genetic region a QTL with main effect on parameters describing senescence processes in Finland was previously reported (Celis-Gamboa 2002; Hurtado et al. 2012b). However, no evidence of QEI has been reported in previous studies. In our study the strongest QEI was detected for onset of flowering, total number of tubers and total tuber weight.

For some traits *photoperiod-independent QTLs* were detected and the most interesting QTLs are the ones in which there were no significant differences between QTL effects in contrasting environments. The QTLs for onset of plant height, inflection point and maximum progression rate of senescence did not show QEI suggesting a very similar effect of the QTLs across all the environments. On the other hand, a higher number of *environment-specific QTLs* were detected under short photoperiod than under long days for agronomic and growth traits. Interestingly most of the environment-specific QTL were detected in Ethiopia.

It was probably due to the combination of short days and low temperatures promoting the phenotypic differences between CE genotypes.

Time dependent QTLs were identified for growth and development. In the growth process, QTLs pop up under contrasting day lengths (Ethiopia and Finland 2004) only in specific growth phases. The senescence process showed time-dependent QTLs for onset and inflection point but they were differentially expressed under contrasting environments. Although onset of senescence was the developmental phase with the highest number of QTLs, the QTLs only had significant effects under long days.

Further research

Considering that we were able to successfully identify pleiotropic regions, photoperiod-independent QTLs and time-depending QTLs based on multi-environment field trials, further research could help to 1) investigate whether QTLs in pleiotropic regions are really pleiotropic or just in closely linkage identifying one or more genes controlling development/growth and agronomic traits and 2) anchor the identified genetic regions to the annotated potato genome sequence (Potato Genome Sequencing et al. 2011) to find candidate genes involved in specific pathways.

Additional files

Additional file 1. Description of agronomic, development and growth traits included in the MET analysis

Agronomic traits

In our study we only included the observations done in the last block (last harvest) of the experiments in the Netherlands (block 14) and Finland (2004: block 8 and 2005: block 4) in which the plants were fully developed.

Tuber number and tuber weight were available for 2 groups of tubers (smaller and bigger than 20mm) in the datasets of Ethiopia, Finland 2004 and The Netherlands.

a. Total **Tuber number** and **tuber weight**: In Ethiopia, Finland 2004 and The Netherlands the total tuber number and total tuber weight was the sum of tuber number and tuber weight in each of the two classes.

   In Finland 2005 and Venezuela only total number of tubers and total weight per plant was available.

b. **Tuber size**: In Venezuela and Finland 2004 the average diameter per plant was given in mm. In the Netherlands, values of a scale between 0.5 and 4 were given to each plant.

c. **Number of main stems**: These are the stems emerging directly from the seed tuber.

d. **Date of emergence**: Days from planting to emergence (days after planting, DAP).

Development and growth traits (time series traits)

Flowering and senescence are refer as developmental traits and they were measured along the life cycle. To complement the study of above ground development, plant height was also measured during the observation period. The number of time points, the scale used to measure each trait and the evaluation period for each one are described in the second part of Table 1.

Time series data from the Venezuelan data set were not considered. Although we have time series for flowering, the description of the scale included bud development and flower abortion and these descriptions were not part of the scales used in the other environments.

a. **Flowering**: In The Netherlands and Ethiopia the progress of this trait was monitored using a scale previously described for potato (Celis-Gamboa et al. 2003) with values

between 0 (no open flowers) and 7 (end of flowering: no developing buds or open flowers). In Finland the description of the scale was basically the same (1= first open flower; 4= peak of flowering; 7= end of flowering).

Data collection: In The Netherlands and Finland the assessments were done per plant before harvesting each block. Therefore independent data per genotype were collected on plants from different blocks along the observation period. While in Ethiopia dependent data per genotype were collected on the same plants along the observation period.

b. **Senescence**: In The Netherlands and Ethiopia the progress of this trait was measured using a scale previously described for potato (Celis-Gamboa et al. 2003) with values between 1 (green plant) and 7 (dead plant).

In Finland the scale was shifted to the right (1= first yellow leaves and 7= no green tissue left).

Data collection: In The Netherlands independent data per genotype were collected from the last 9 blocks of the experiment. Whereas in Finland, dependent data per genotype were the result of senescence assessments on the same plants from the last block of each experiment (2004 and 2005).

c. **Plant height**: the length of the highest stem was measured from the ground level to the main apex (cm)

Data collection: In Finland 2004 and Ethiopia data were collected on each plant of all the blocks included in the experiments.

Additional file 2. Correlations between traits measured in Venezuela, Ethiopia, The Netherlands, Finland 2004 and Finland 2005

Venezuela								
Pheight	1	-						
Sen	2	-0.06	-					
Ttubers	3	0.02	-0.1	-				
Tweight	4	-0.19	-0.01	0.15	-			
Tsize	5	-0.14	0.08	**-0.32**	**0.57**	-		
Stems	6	**0.26**	0.11	0.06	0.03	-0.04	-	
		1	2	3	4	5	6	

Ethiopia

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
FOns	1	-																					
FMe	2	**-0.14**	-																				
FMax	3	-0.16	**0.62**	-																			
FIp	4	**0.64**	**-0.5**	**-0.53**	-																		
FEnd	5	**0.57**	**-0.73**	**-0.73**	**0.81**	-																	
Fdur	6	**0.25**	**-0.85**	**-0.79**	**0.66**	**0.85**	-																
PHons	7	**0.39**	**-0.36**	**-0.41**	**0.55**	**0.61**	**0.48**	-															
Phme	8	**0.3**	**-0.5**	**-0.62**	**0.59**	**0.72**	**0.67**	**0.62**	-														
PHmax	9	**0.33**	**-0.45**	**-0.6**	**0.55**	**0.68**	**0.61**	**0.59**	**0.95**	-													
PHip	10	**0.41**	**-0.44**	**-0.46**	**0.62**	**0.66**	**0.56**	**0.84**	**0.66**	**0.58**	-												
PHend	11	**0.21**	**-0.3**	**-0.3**	**0.4**	**0.44**	**0.38**	**0.64**	**0.47**	**0.34**	**0.73**	-											
PHmaxhe	12	**0.28**	**-0.46**	**-0.6**	**0.55**	**0.68**	**0.63**	**0.59**	**0.99**	**0.96**	**0.61**	**0.42**	-										
PHmehe	13	**0.17**	**-0.39**	**-0.54**	**0.44**	**0.55**	**0.55**	**0.38**	**0.87**	**0.85**	**0.38**	**0.21**	**0.88**	-									
SOns	14	**0.21**	**-0.39**	**-0.49**	**0.41**	**0.47**	**0.46**	**0.4**	**0.46**	**0.4**	**0.51**	**0.4**	**0.42**	**0.38**	-								
SMe	15	**-0.33**	**0.44**	**0.43**	**-0.54**	**-0.62**	**-0.51**	**-0.46**	**-0.54**	**-0.49**	**-0.49**	**-0.31**	**-0.51**	**-0.46**	**-0.28**	-							
SMax	16	**-0.34**	**0.3**	**0.39**	**-0.5**	**-0.5**	**-0.43**	**-0.51**	**-0.57**	**-0.53**	**-0.57**	**-0.32**	**-0.55**	**-0.5**	**-0.43**	**0.67**	-						
SIp	17	**0.38**	**-0.48**	**-0.59**	**0.61**	**0.68**	**0.61**	**0.57**	**0.66**	**0.59**	**0.66**	**0.51**	**0.62**	**0.57**	**0.79**	**-0.61**	**-0.7**	-					
Send	18	**0.36**	**-0.51**	**-0.59**	**0.61**	**0.69**	**0.62**	**0.59**	**0.7**	**0.63**	**0.68**	**0.49**	**0.65**	**0.61**	**0.75**	**-0.73**	**-0.78**	**0.94**	-				
Ttubers	19	0.13	**-0.32**	**-0.19**	**0.24**	**0.32**	**0.3**	**0.27**	**0.35**	**0.28**	**0.31**	**0.26**	**0.33**	**0.33**	**0.21**	**-0.26**	**-0.29**	**0.3**	**0.34**	-			
Tweight	20	0.14	**-0.4**	**-0.44**	**0.32**	**0.42**	**0.47**	**0.32**	**0.64**	**0.59**	**0.33**	**0.24**	**0.63**	**0.73**	**0.37**	**-0.52**	**-0.5**	**0.52**	**0.61**	**0.54**	-		
Emer	21	**0.41**	0.05	0.08	**0.24**	0.15	-0.04	**0.25**	-0.08	-0.04	**0.22**	0.09	-0.1	**-0.25**	0.03	-0.06	-0.04	0.05	0.01	**-0.22**	**-0.27**	-	
Stems	22	-0.13	-0.01	-0.02	-0.06	-0.06	0.04	-0.11	0.04	0.02	-0.09	-0.06	0.07	0.13	-0.05	0.07	0	-0.09	-0.07	**0.63**	**0.27**	**-0.33**	-

Netherlands

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
FOns	1	-															
FMe	2	**-0.15**	-														
FMax	3	**-0.49**	**0.83**	-													
FIp	4	**0.77**	**-0.73**	**-0.88**	-												
FEnd	5	**0.65**	**-0.83**	**-0.92**	**0.98**	-											
Pheight	6	**0.47**	**-0.63**	**-0.67**	**0.73**	**0.75**	-										
SOns	7	**0.5**	**-0.69**	**-0.79**	**0.79**	**0.81**	**0.67**	-									
SMe	8	**-0.55**	**0.7**	**0.77**	**-0.83**	**-0.85**	**-0.73**	**-0.82**	-								
SMax	9	**-0.52**	**0.64**	**0.73**	**-0.77**	**-0.78**	**-0.67**	**-0.81**	**0.96**	-							
SIp	10	**0.54**	**-0.71**	**-0.81**	**0.83**	**0.85**	**0.69**	**0.94**	**-0.92**	**-0.93**	-						
Ttubers	11	0.08	-0.07	-0.11	0.09	0.1	**0.18**	**0.25**	-0.1	**-0.15**	**0.2**	-					
Nless20	12	**0.24**	-0.14	**-0.24**	**0.24**	**0.24**	**0.3**	**0.32**	**-0.24**	**-0.26**	**0.3**	**0.91**	-				
Nmore20	13	**-0.17**	0.05	0.11	**-0.15**	-0.13	-0.06	0.07	0.12	0.04	0	**0.81**	**0.5**	-			
Tweight	14	-0.12	**0.2**	**0.23**	**-0.22**	**-0.23**	**-0.22**	-0.1	**0.19**	0.1	**-0.11**	**0.26**	0.09	**0.42**	-		
tusize	15	0.06	0.01	-0.1	0.06	0.05	0.05	0.03	-0.1	-0.11	**0.05**	-0.04	0.02	-0.11	-0.1	-	
Stems	16	**-0.15**	**0.19**	**0.21**	**-0.24**	**-0.23**	**-0.19**	-0.15	**0.19**	**0.16**	**-0.17**	**0.33**	**0.24**	**0.35**	0.13	-0.01	-

Finland 2004

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
FOns	1	-																							
FMe	2	**0.3**	-																						
FMax	3	0.1	**0.68**	-																					
FIp	4	**0.79**	**-0.29**	**-0.38**	-																				
FEnd	5	**0.55**	**-0.59**	**-0.61**	**0.93**	-																			
PHons	6	**0.43**	-0.17	-0.15	**0.52**	**0.49**	-																		
Phme	7	0.13	**-0.28**	**-0.3**	**0.31**	**0.37**	**0.65**	-																	
PHmax	8	**-0.19**	-0.08	-0.07	-0.12	-0.07	**0.28**	**0.7**	-																
PHip	9	**0.44**	**-0.21**	**-0.22**	**0.56**	**0.55**	**0.9**	**0.7**	**0.25**	-															
PHend	10	**0.39**	**-0.25**	**-0.21**	**0.53**	**0.52**	**0.68**	**0.57**	0.07	**0.79**	-														
PHmaxhe	11	0.09	**-0.28**	**-0.31**	**0.27**	**0.34**	**0.59**	**0.99**	**0.71**	**0.64**	**0.5**	-													
PHmehe	12	-0.05	**-0.28**	**-0.27**	0.14	**0.2**	0.1	**0.53**	**0.41**	0.1	**0.21**	**0.57**	-												
SOns	13	0.07	-0.03	0.05	0.03	0	0.05	0.15	0.04	0.11	0.11	0.16	0.11	-											
SMe	14	**-0.45**	**0.35**	**0.31**	**-0.65**	**-0.68**	**-0.35**	**-0.22**	0.11	**-0.37**	**-0.33**	**-0.19**	-0.12	**0.14**	-										
SMax	15	**-0.33**	**0.32**	**0.3**	**-0.52**	**-0.57**	**-0.32**	**-0.29**	-0	**-0.34**	**-0.36**	**-0.27**	**-0.18**	-0.07	**0.78**	-									
SIp	16	**0.41**	**-0.27**	-0.16	**0.54**	**0.54**	**0.29**	**0.23**	-0.07	**0.33**	**0.29**	**0.23**	0.15	**0.5**	**-0.66**	**-0.76**	-								
Ttubers	17	0.16	0.1	0.04	0.14	0.1	**0.2**	**0.24**	0.17	**0.21**	**0.21**	**0.23**	**0.22**	-0.05	-0.1	-0.1	0.08	-							
Nless20	18	**0.22**	0.04	0.03	**0.23**	**0.17**	**0.3**	**0.2**	0.09	**0.24**	**0.27**	**0.18**	**0.18**	-0.1	**-0.19**	-0.15	0.1	**0.84**	-						
Nmore20	19	**0.22**	0.01	-0.02	**0.26**	**0.22**	**0.23**	**0.25**	0.13	**0.24**	**0.21**	**0.23**	**0.22**	-0.06	**-0.18**	-0.09	0.11	**0.91**	**0.73**	-					
Tweight	20	**0.33**	-0.02	-0.05	**0.36**	**0.32**	0.17	0.06	-0.12	**0.21**	**0.27**	0.03	0.02	**-0.2**	**-0.3**	**-0.19**	0.11	**0.45**	**0.51**	**0.46**	-				
Wless20	21	**-0.33**	-0.14	-0.06	**-0.29**	**-0.21**	**-0.21**	-0.03	0.08	**-0.19**	-0.14	-0.01	-0.03	0.15	0.14	-0.03	0.03	**-0.63**	**-0.68**	**-0.68**	**-0.56**	-			
Wmore20	22	-0.09	**-0.17**	-0.12	-0.02	0.05	-0.1	0.02	-0.02	-0.04	0.08	0.02	-0.01	0	-0.11	**-0.2**	0.13	**-0.34**	**-0.34**	**-0.39**	**0.24**	**0.67**	-		
stems	23	-0.02	0.12	0.03	-0.08	-0.09	0.06	0.15	**0.17**	0.05	0.02	0.15	0.1	-0.03	-0.09	-0.07	0.03	**0.51**	**0.44**	**0.49**	**0.19**	**-0.3**	**-0.18**	-	
tusize	24	**-0.36**	-0.11	-0.05	**-0.32**	**-0.24**	**-0.27**	-0.08	0.06	**-0.23**	**-0.2**	-0.06	-0.07	0.14	**0.18**	0.01	-0.01	**-0.68**	**-0.77**	**-0.72**	**-0.65**	**0.97**	**0.56**	**-0.34**	-

Finland 2005

FOns	1	-													
FMe	2	**0.74**	-												
FMax	3	**0.49**	**0.74**	-											
FIp	4	**0.18**	**-0.44**	**-0.62**	-										
FEnd	5	**-0.37**	**-0.86**	**-0.85**	**0.81**	-									
Pheight	6	**-0.19**	**-0.25**	**-0.46**	**0.22**	**0.31**	-								
SOns	7	0.06	-0.17	**-0.27**	**0.34**	**0.34**	0.06	-							
SMe	8	**0.27**	**0.58**	**0.61**	**-0.56**	**-0.67**	**-0.31**	**-0.19**	-						
SMax	9	**0.25**	**0.51**	**0.63**	**-0.52**	**-0.63**	**-0.34**	**-0.15**	**0.94**	-					
SIp	10	**-0.18**	**-0.55**	**-0.62**	**0.62**	**0.71**	**0.28**	**0.57**	**-0.88**	**-0.82**	-				
Ttubers	11	**0.17**	**0.46**	**0.44**	**-0.46**	**-0.54**	0.04	**-0.23**	**0.42**	**0.36**	**-0.45**	-			
Tweight	12	0.05	**0.47**	**0.53**	**-0.64**	**-0.65**	-0.03	**-0.25**	**0.49**	**0.46**	**-0.53**	**0.76**	-		
Emer	13	**0.4**	0.03	-0.08	**0.44**	**0.21**	**-0.31**	**0.21**	-0.13	-0.12	**0.21**	**-0.3**	**-0.46**	-	
Stems	14	0.04	**0.19**	**0.21**	**-0.25**	**-0.25**	0.04	**-0.15**	0.13	0.07	-0.15	**0.57**	**0.33**	-0.09	-
		1	2	3	4	5	6	7	8	9	10	11	12	13	14

Chapter 6

General discussion

1.Introduction

The Laboratory of Plant Breeding of Wageningen University UR has used a diploid potato population to study development and agronomic traits under different field conditions for the last 15 years. The exploration of the extensive phenotypic data generated in previous studies offers a lot of opportunities. However, it was also a major challenge to use this valuable information to draw conclusions about relationships between traits under different day lengths, genetic architecture of plant development and effects of the two major environmental factors (temperature and day length) on potato adaptation.

This chapter is highlighting relevant facts related to multi-environment trials (METs) that have been identified during our research and the implications for potato breeding of handling and analysing complex data involving multiple locations, multiple traits and multiple observations during the growing season

2. Important facts related to METs

As shown in our research, METs are a generous source of multiple types of data and have great potential. However special attention should be given to three main factors identified during our research: the importance of using standard phenotyping protocols in all the experiments, the value of historical data as a source of phenotypic information in breeding programs and the relevance of multidisciplinary research combining different disciplines.

2.1. Standard phenotyping protocols

Phenotyping is the most important step of a breeding process. Therefore accurate and reliable protocols must be followed in a consequent and consistent way. The accurate detection of genetic regions controlling a trait depends on high quality phenotypes. However, phenotyping is not straightforward due to the strong interaction between genetic and environmental factors, the subjectiveness and bias in phenotypic observations and the expression of a phenotype depending on complex interactions between the plant and external factors (biotic, abiotic). For instance, evaluation of senescence under field conditions can be biased due to the presence of late blight symptoms that can be confounded with some stages of the ageing process.

Poorly documented scales for ordinal scores may also cause problems. In some cases only the first and the last level of a scale are well described while the intermediate levels are not explained in detail, or even not at all, and sometimes the scores are not entirely ordinal because of some more descriptive classes being used as well.

Conventional, non-invasive phenotyping methods currently in use depend on manual workflows that require extensive plant handling as well as manual measurements with human intervention. Protocols are often characterized by low precision and possible bias in the results caused by plant handling or subjective interpretation. A major bias occurs when a single observation is done at the end of the life cycle and the impact of the early stages of development (emergence, sprouting, etc) are ignored. Thus, inaccuracy and inadequacy of phenotyping make marker assisted selection (MAS) for complex traits one of the biggest challenges faced by breeders (Araus et al. 2003). Considering that breeding progress depends on accurate selection of rare genotypes that possess new or improved attributes due to superior combinations of alleles at multiple loci, precise phenotyping must be one of the pillars in plant breeding.

Despite the low accuracy of conventional phenotyping methods, they have been used extensively in research and breeding programs due to low cost, easy implementation and sometimes also due to the lack of a better alternative. Nowadays the main emphasis of phenotyping procedures is on new methods to obtain objective, robust, non-destructive and automated ways of collecting data. Some of the currently used automated phenotyping techniques are spectral reflectance (Montes et al. 2007), spectral imaging (Ruckelshausen 2007), chlorophyll fluorescence (Chaerle et al. 2007), thermal imaging(Chaerle et al. 2007; Moeller et al. 2007) and image analysis (Hartmann et al. 2011). These phenotyping methods allow the screening of genotypes for several traits, in multi-location field trials. They also help to reduce the main limitation of phenotyping large mapping populations for dynamic traits at short regular time intervals during the growing season. Therefore one of the direct effects of automated phenotyping is an increase in heritability due to the repeated observations that can be obtained for the same trait during different development stages. However the best moment to take data may differ between traits according to the needs of

the study. Indeed a more powerful genetic analysis is achieved by phenotyping at the right moment(s).

Although automated phenotyping allows high throughput evaluations under field conditions it also has some drawbacks. The high complexity of the data generated makes the biological interpretation and the posterior analysis difficult. Automated phenotyping of large field trials is also expensive and requires further practical and technical optimization. Therefore the successful implementation of automated phenotyping technologies for large field trials is only possible in developed countries, where labour costs are also higher. The adoption of these technologies in developing countries is limited by the shortage of well-trained personnel, lack of suitable field infrastructure, lack of information systems or adapted analytical tools or simply by limited resources in the breeding programs (Ribaut et al. 2010). Thus, conventional phenotyping techniques depending on manual measurements will still be used worldwide for some time. Therefore, breeding programs must promote the use of standard and accurate phenotyping protocols which facilitate the comparison of traits across experiments, the understanding of genotype by environment interactions, GEI, and ultimately the detection of the genetic basis of complex traits. Meanwhile, improving the phenotyping infrastructure in developing countries must thus be a priority.

In our research we have made used of traits measured with non-automated methods. As expected, we encountered difficulties to compare traits across locations and years due to the different scales or units reported and sometimes also due to incomplete description of the scales. Therefore some time was devoted to make the traits comparable across environments and it was also one of the motivations of **Chapter 2** describing flexible methods to handle and analyse ordinal and continuous data.

2.2. Historical data as source of phenotypic information

Historical data, defined as data generated in previous experiments, are of great value in breeding programs, in research and gene banks. As standard phenotyping protocols, historical data must be clearly documented and properly stored to be able to benefit and profit from them.

A centralised data base, well curated and storing the information generated in previous experiments is a key part of a breeding program if historical data is to be used. A comprehensive description (in words and images) of the experiments, the scales used to measure different traits and the genotypes considered in each trial are as important as a proper data recording. It is more effective if the documentation of the trial is generated during, rather than after, the data recording. It results in a better quality data collection, as well as a more accurate description of the overall experiment. The process of documenting data generated during project execution will also help to refine research questions which in general will benefit the quality of the research.

Good documentation of historical data is essential for their long term use. Without it, data are not suitable for further studies and their usefulness is lost. In our research we have made use of historical data generated in the last 15 years to investigate GEI and QTL by environment interactions, QEI, as shown in **Chapters 3 and 5** and we were able to successfully draw conclusions. However we have identified three important elements that should be described for a proper documentation of historical data used in MET studies: content, origin/source and structure.

The description of the content should provide sufficient details to allow a potential user to judge whether the data are appropriate for their needs. The description of the experiments must include the geographic position, the list of traits measured and the corresponding scale or units, the time period, periodicity of data collection and list of genotypes names/codes. The original name of the genotypes should be kept to avoid inconsistencies. Multiple names that refer to the same genotype are source for misleading conclusions. If the genotypes included in the experiments are to be used over time and multiplication rounds are needed, the identity of genotypes should be checked regularly.

The description of the data's origin must provide enough information to understand the framework in which the data were generated and the steps that were taken during data collection. Therefore the most important details to be provided are a brief description of the project which gave rise to data collection, references indicating how the data have been

used (including publications which have made use of the data) and the reference number or name used to identify a particular experiment.

The structure, form and organization of the data must be comprehensive enough to enable others to explore it and put it in a format suitable for new studies. A full description of the files must include among others the number of records, the exact experimental design as laid out in the field (including location and orientation), the field length and the type of data. When means have been stored for genotypes instead of plot data, it is important to know how these means were obtained, which statistical model was used to summarize the data, if the genotypes were interpreted as representing a fixed or a random term and consequently if the genotypic means are best linear unbiased estimates or predictions (BLUEs or BLUPs).

2.3. Multidisciplinary research combining different expertise

Given the complexity of our data, involving multiple locations, multiple traits and multiple observations during the growing season, a multi-disciplinary approach is of great help to answer complex research questions.

A simple definition of multidisciplinary research is “research that involves the interaction among two or more different disciplines”. This may range from sharing ideas to full integration of concepts, methodology, procedures, theory, terminology, data and organization of research and training in a fairly large field (www.nserc-crsng.gc.ca 2012). Multidisciplinary research draws on knowledge from different disciplines but stays within the boundaries of those fields.

In our research we have greatly benefited from the interaction between biology, mathematics and applied statistics. Statistical models were used to describe biological processes as shown in **Chapter 2**. Flexible methods for modelling dynamic phenotypic data were proposed taking plant height and senescence as example traits. The methods proposed to analyse ordinal and continuous data taken at different times during the growing season can be extended to model other time-dependent traits. Therefore the multidisciplinary approach was not only beneficial to this particular research but will also provide a useful tool

for other researchers in different fields. These methods could also be used and implemented in the context of high-throughput automated phenotyping over time.

This research has also benefitted from the close interaction between biology and applied statistics making use of newly developed tools for the analysis of multi-environment and multi trait data as shown in **Chapters 4 and 5**. Working in close collaboration with the developers of the QTL-library of GenStat (VSN 2011) we have been able to use advanced statistical tools to analyse the complex data sets of METs.

3. Analysis of complex data

The potato experiments considered in our research provided multiple types of data including phenotypic data measured only once during the experiments and traits where data were collected at several instances during the growing season. Given the complexity of our data, complex data analyses were required and they provided the bases to understand development and agronomic traits of potato influenced by different factors. In addition, our research benefited from molecular data incorporated into the analysis to understand the genetic architecture of crop development, to identify time-dependent QTLs controlling different processes, to detect putative pleiotropic-QTLs related to development and agronomic traits and to find QTL by Environment Interactions, QEI, affecting different traits.
Extending our results to a broader context, this study can be used as an example of complex analysis using historical data. The benefits of such kind of analysis were illustrated and evidence was provided of the valuable information existing in breeding programs: their data just need to be explored.

We expect historical data in breeding companies and in research to be used more often as it has become much easier and cheaper to also analyze molecular markers/sequence data in high-throughput fashion. For applications such as genome wide association studies, GWAS, and genomic selection it might be useful to have a very large phenotypic data set and also large sets of genotypes. Thus, this could create a demand to re-use pedigreed historical data as well.

3.1. Multiple time points analysis

Quantitative information on plant development characters (e.g. leaf area, biomass and development stage) is rare and usually presented for a small number of genotypes. Experiments involving a large number of genotypes often include only qualitative information about plant phenotypes (Kuromori et al. 2006), which, although helpful for detecting improved phenotypes, are less informative for studying the genetic basis of development processes (Arvidsson et al. 2011). For that we require data collected on large populations at several instances during the growing season. The use of flexible and accurate methods for modelling this type of data is crucial to be able to monitor plant development over time. In our research we make use of semi-parametric models, flexibleenough to analyse ordinal and continuous data as shown in **Chapters 2 and 3**. The methods provide a framework that can accommodate the behaviour of all genotypes in a test population because they are less restrictive than parametric models in terms of symmetry. When it comes to traits such as senescence, flowering, plant height and so on, not all cultivars complete the development trajectory within the experimental time frame. This leads to incomplete data series that are not easily described by parametric models. A classic example of a parametric model for growth and development is the (two parameter) logistic model (Verhulst 1845). Especially late genotypes with an incomplete growth curve at harvest cannot be modelled by a logistic curve. The use of semi-parametric models overcomes this problem and all genotypes can be captured in the same framework and be described by the same set of parameters.

As we have observed in our analyses, our methods are also flexible enough in terms of the length of the data series to be analyzed. Beyond a certain minimal number of data points the models give equally good results for longer and for shorter data series. This is of special value as it helps to analyse existing datasets that might not have been suited for analysis with other methods.

Additionally, describing a development trait by means of a curve enables us to zoom in on certain periods in the growing season. We can define parameters for specific parts of the curve and search for QTLs associated with the different stages of development and growth. The use of curve characteristics also facilitates the modelling of development traits within a

particular biparental population where comparisons of traits at single time points is always confounded with development differences. Interactions between loci controlling different stages of developmental processes can also be studied as shown in **Chapter 3** where epistatic interactions were detected.

Additional characteristics can be defined according to the need of the research study and no additional data collection is needed. In **Chapter 2** we made use of this property by including the maximum progression rate for the onset of plant growth as a trait into the QTL analysis. This new trait leaded to the discovery of three new QTLs. This illustrates how characteristics of the curves can be related easily to the biological process behind it and how estimating such characteristics provides a basis for a successful, subsequent QTL analysis. As a consequence, breeding programs and researchers can profit from making a better use of available data and ultimately arrive at more productive QTL analyses.

When development processes are evaluated in different locations and years, the use of curve characteristics to compare development stages is appropriated considering that even in different environments the characteristics have the same biological meaning. In genetic studies the use of curve characteristics also favours the identification of QEI and QTLs associated to particular phases of growth and/or development (defined as phase-dependent QTLs). For instance, when different environments were compared in our study phase-dependent QTLs were identified for onset and inflection point of senescence and they were also differentially expressed under contrasting environments.

3.2. Multi-trait QTL analysis

Although many QTL studies have considered multiple traits, usually those traits are analysed separately. An integrated analysis combining traits related to development processes simultaneously is required to get a better understanding of the genetic forces driving plant development. Combining data from multiple traits related to plant development in QTL analysis not only increases the power of QTL detection, it also improves the understanding of the genetic control of development processes. As a consequence, a multi-trait QTL analysis of a single population allows detection of closely linked chromosomal regions affecting several traits simultaneously (Jiang and Zeng 1995). A first attempt to estimate the

optimal set of consensus QTLs for several traits simultaneously in potato was done through a QTL meta-analysis (Danan et al. 2011). It allowed the co-localization of late blight resistance and plant maturity traits by projecting individual QTLs onto a consensus map. However, there are no reports of such integrative analysis for development traits in potato. So far, data on traits related to plant development in potato have not been integrated in a single study to get insight in the genetic architecture of crop development and the presence of putative pleiotropic QTL related to plant development.

In our research we made use of data collected under short day length for multiple traits (**Chapter 4**). The combined use of characteristics describing development and growth as well as agronomic traits in a multi-trait QTL analysis provided insights into (1) the genetic architecture of development and complex traits in potato, (2) the presence of phase-dependent QTLs associated to particular development/gowth stages and (3) the genetic link between above and below ground traits.

For each complex trait multiple QTLs were providing a genetic basis for such quantitative traits. Some QTLs were permanently expressed during the whole development process and some of them were associated with specific development stages. These phase-dependent QTLs were detected for flowering, senescence and plant height. Some of them were expressed at early development stages and they seem to be switched off after half the development process was completed. Others were silent during early stages but they were expressed in the final part of the process.

Fourteen pleiotropic QTLs explained the relationships between above and below ground traits under short day length. The relationships between some traits have already been reported in physiological studies but the genetic component behind them was still missing. To illustrate the results from our multi-trait QTL analysis we take the pleiotropic QTL detected on linkage group C3 as example. It was related to tall plants, with fast growth, showing few main stems and few tubers. The relationship between number of tubers and growth is consistent with previous studies showing that tuber formation could be reduced when the development of the haulm is accelerated (Maris 1964). Our results add a genetic

component to such relationship. Thus, we show the potential of multi-trait analysis incorporating molecular information to find the genetic bases of physiological relationships.

3.3. Multi-environment QTL analysis

Understanding the genetic basis of development and other complex traits can support previous physiological studies and facilitate breeding strategies. The detection of QTLs for these types of traits in environments with contrasting conditions can be done using information from multi-environment trials (METs). The inclusion of populations of genetically related individuals in such trials, facilitates the understanding of the genetic control of adaptive traits by identifying associations with QTLs (Boer et al. 2007; Malosetti et al. 2008). Hence, the combined use of genetic and phenotypic information from METs allows the detection of genetic factors controlling development and complex traits and also to study the presence of QEI.

The applications and usefulness of multi-environment QTL is presented in **Chapter 5**. In our research the analysis of the METs provided the bases for understanding 1) the effect of temperature and photoperiod on development, growth and agronomic traits evaluated under field conditions, 2) genotype by environment interactions that play an important role in the phenological plasticity of the crop and 3) associations between development and agronomic traits. Furthermore, the multi-environment QTL analyses combining MET data with genetic information allowed the identification of 4) QTL by environment interactions, 5) pleiotropic genetic regions, 6) QTLs which are stable across environments and 7) time dependent QTLs for development and growth.

Controlled experiments using potato cultivars have shown that high temperatures combined with long days increase the life span of the entire shoot (Struik and Ewing 1995). In our research we observed that day length had a stronger effect than temperature on the life span. High temperatures combined with long days in the Netherlands accelerated the flowering process but delayed onset of senescence thereby increasing the life span of the entire shoot in the whole population. Finland with lower temperatures showed longer periods of flowering and senescence and the same increase in life span. On the other hand, short days, as in Ethiopia, promoted early onset and also fast progression of flowering and

senescence thereby reducing the life span with almost 20 PBTT units compared with long days. PBTT (photo-beta thermal time) is the cumulative thermal unit, combining temperature and day length, considered in our study for the comparison between environments. Despite the differences in development under contrasting day lengths, no much GEI was observed and only small shifts were observed across environments. For instance early flowering genotypes tend to be even earlier under long photoperiod and high temperatures. Plant growth in Finland was compared to that in Ethiopia where growth was restricted by short days and smaller plants were observed. Despite the differences in plant heights, the growth process under long and short days showed similar progression (high positive correlation between the average progression rates). It means that we can observe parallel growth curves of genotypes across contrasting environments without much GEI. The absence of GEI for the different stages of development and growth indicates good adaptation of the CE genotypes to different environments.

Regarding the agronomic traits, few indications of GEI were detected in our study. However, as it was expected, the strongest GEI was found for total tuber weight, which is one of the most important yield determinants (Lemaga and Caesar 1990). GEI is the main constraint in potato breeding, which makes it difficult to use varieties worldwide when good yield is a pre-requisite. Our results show strong GEI for total tuber weight with cross-over interactions between some environments. A negative correlation was observed between the two experimental years in Finland. Although the experiments were performed in the same location, the growing season started earlier in 2004 and lower temperatures were recorded during the observation period. Our interpretation is that the strong GEI observed between 2004 and 2005 was due to these differences in temperature over the two seasons. On the other hand, positive correlations were observed between total tuber weight in the Netherlands and Finland 2005 indicating similar performance of the CE genotypes under long photoperiod with high temperatures during the growing season. In addition, positive correlation was also observed between Ethiopia and Finland 2004 where similar temperatures were observed in the middle of the growing season. These results suggest that the main factor controlling GEI for total tuber weight is temperature and there is not a major role of photoperiod.

Associations between agronomic, development and growth traits were found under contrasting environments. Photoperiod seem to have an effect on flowering and tuber weight, thus under long photoperiod, fast progression of flowering was associated with higher tuber weight. Whereas, in Ethiopia (with short days) long flowering and senescence processes were associated with high tuber weights (especially in tubers bigger than 20mm). The relationship between tuber size and total tuber weight was probably affected by photoperiod as well. In Venezuela bigger tubers were related to higher tuber weight, while in Finland 2004 bigger tubers correspond with low tuber weight. However, the same relationship could not be detected between Venezuela and The Netherlands suggesting that other environmental factors also play a role.

The polygenic inheritance of the complex traits included in our study was successfully identified for most of the traits with at least two QTLs associated with each trait in the different environments.

On the other hand, all the development stages of flowering were associated with only one QTL on chromosome 5 (E5: 21.9 cM and C5: 105cM with inverted order of the molecular markers(Hurtado et al. 2012a)) with the highest explained variance given by contrasting alleles of the E parent. Other QTLs with small effects have been detected that influence the flowering process using a multi-trait QTL analysis where correlations between traits were also taken into account (Hurtado et al. 2012a). Since we are using a multi-environment QTL approach for each trait, such correlations were not included in our analysis and possibly we are lacking the extra power across correlated traits required to detect these QTLs with smaller effects.

The combined use of MET data and genetic information also allowed the detection of pleiotropic genetic regions associated with multiple traits and helped to understand the relationships between traits under contrasting environments. For instance, a negative correlation between progression rate of growth and progression rate of development was observed in all environments indicating that fast growth, leading to tall plants, was related to slow flowering and senescence processes; it was explained by a QTL on chromosome 5. In Ethiopia and Finland 2004, this QTL was associated with tall plants, with fast growth but slow

senescence progression and higher total number of tubers (especially tubers bigger than 20mm) at harvest. Previous reports indicated the effect of day length on the relationship between number of tubers and other traits, such as average tuber weight and tuber yield (Lemaga and Caesar 1990; Maris 1964). However to our knowledge there are no reports showing the relationship between growth, development and total number of tubers across photoperiods as we presented here.

The main pleiotropic region was detected on E5 between 18 - 25cM and it was associated with all flowering, senescence and plant height parameters (except onset of senescence) and to some of the agronomic traits. In the same genetic region a QTL with main effect on parameters describing senescence processes in Finland was previously reported (Celis-Gamboa 2002; Hurtado et al. 2012b). However, no evidence of QEI was reported using the CxE population in previous studies. In our study the strongest QEI was detected for onset of flowering, total number of tubers and total tuber weight.

QEI was detected for most of the QTLs in our study. We noticed that most of the QEI was due to significant differences between the QTL effects going in the same direction and in that case low positive correlation was found between the environments. On the other hand, the presence of QTL effects with crossover interactions was associated with strong GEI with negative correlations between environments.

For some traits photoperiod-independent QTLs were detected and the most interesting QTLs were the ones in which there were no significant differences between QTL effects in contrasting environments. The QTLs for onset of plant height, inflection point and maximum progression rate of senescence did not show QEI suggesting a very similar effect of the QTLs across all the environments. On the other hand, a higher number of environment-specific QTLs were detected under short photoperiod than under long days for agronomic and growth traits. Interestingly most of the environment-specific QTLs were detected in Ethiopia and multiple reasons could support our finding. Probably the combination of short days and low temperatures promoted phenotypic differences between CE genotypes. It could be also due to the extra power given by the larger number of observations across time points compared with the other environments. In addition it could also be due to more accurate

observations because of better defined phenotyping protocols based on previous experiments in other locations.

Conclusion and outlook

Our research involves data generated in multiple experiments during the last 15 years and shows the potential of historical data to elucidate physiological and genetic relationships, as well as the influence of the main environmental factors on crop performance. Although we are making use of a diploid mapping population, the complex analysis used in our research can be extended to sets of cultivars for association studies and to polyploidy cross populations. We stress the importance of standard phenotyping protocols when data from multiple experiments have to be combined. Furthermore, we show the benefits of multidisciplinary research integrating biology, mathematics and applied statistics to make full use of all available information. The methods presented here to model ordinal or continuous data over time could also be used and implemented in the context of high throughput automated phenotyping over time. These methods are flexible enough in terms of the length of the data series to be analysed, they also deal with the behaviour of all genotypes in a test population because they are less restrictive than parametric models in terms of symmetry and they can be automated to make them even more useful in the context of high throughput phenotyping.

The results presented in this thesis open the door to further research to investigate whether QTLs in pleiotropic regions are really pleiotropic or just in close linkage and to find candidate genes involved in specific pathways through the anchoring of the identified genetic regions to the annotated potato genome sequence (Potato Genome Sequencing et al. 2011).

Literature cited

Allaby, M. "A Dictionary of Plant Sciences." http://www.encyclopedia.com. : Encylopedia.com, 1998.

Ames, M., and D. M. Spooner. "DNA from Herbarium Specimens Settles a Controversy About Origins of the European Potato." *American Journal of Botany* 95, no. 2 (Feb 2008): 252-57.

Anithakumari, A. M., J. Tang, H. J. van Eck, R. G. Visser, J. A. Leunissen, B. Vosman, and C. G. van der Linden. "A Pipeline for High Throughput Detection and Mapping of Snps from Est Databases." *Molecular Breeding* 26, no. 1 (Jun 2010): 65-75.

Araus, J. L., J. Bort, P. Steduto, D. Villegas, and C. Royo. "Breeding Cereals for Mediterranean Conditions: Ecophysiological Clues for Biotechnology Application." *Annals of Applied Biology* 142, no. 2 (2003 2003): 129-41.

Arvidsson, S., P. Perez-Rodriguez, and B. Mueller-Roeber. "A Growth Phenotyping Pipeline for Arabidopsis Thaliana Integrating Image Analysis and Rosette Area Modeling for Robust Quantification of Genotype Effects." *New Phytologist* 191, no. 3 (2011): 895-907.

Atchley, William R. "Ontogeny, Timing of Development, and Genetic Variance–Covariance Structure." *The American Naturalist* 123, no. 4 (1984): 519-40.

Benoit, G. R., W. J. Grant, and O. J. Devine. "Potato Top Growth as Influenced by Day-Night Temperature Differences 1." *Agronomy Journal* 78, no. 2 (1986 1986): 264-69.

Boer, M. P., D. Wright, L. Feng, D. W. Podlich, L. Luo, M. Cooper, and F. A. van Eeuwijk. "A Mixed-Model Quantitative Trait Loci (Qtl) Analysis for Multiple-Environment Trial Data Using Environmental Covariables for Qtl-by-Environment Interactions, with an Example in Maize." *Genetics* 177, no. 3 (Nov 2007): 1801-13.

Bonhomme, Raymond. "Bases and Limits to Using 'Degree.Day' Units." *European Journal of Agronomy* 13, no. 1 (2000): 1-10.

Borah, M.N. , and F.L. Ilthorpe. "Growth of the Potato as Influenced by Temperature." *Indian Journal of Plant Physiology* 5 (1962): 53-72.

Bradshaw, J. E., C. A. Hackett, B. Pande, R. Waugh, and G. J. Bryan. "Qtl Mapping of Yield, Agronomic and Quality Traits in Tetraploid Potato (Solanum Tuberosum Subsp. Tuberosum)." *Theoretical and applied genetics.* 116, no. 2 (Jan 2008): 193-211.

Celebi-Toprak, F., S. A. Slack, and M. M. Jahn. "A New Gene, Ny Tbr , for Hypersensitivity to Potato Virus Y from Solanum Tuberosum Maps to Chromosome Iv." *Theoretical and Applied Genetics* 104, no. 4 (2002/03/01 2002): 669-74.

Celis-Gamboa, C. "The Life Cycle of the Potato (Solanum Tuberosum L.): From Crop Physiology to Genetics." PhD thesis, Wageningen University, Plant breeding department, 2002.

Celis-Gamboa, C., P. C. Struik, E. Jacobsen, and R. G. F. Visser. "Temporal Dynamics of Tuber Formation and Related Processes in a Crossing Population of Potato (Solanum Tuberosum)." *Annals of Applied Biology* 143, no. 2 (2003): 175-86.

Chaerle, L., I. Leinonen, H. G. Jones, and D. Van der Straeten. "Monitoring and Screening Plant Populations with Combined Thermal and Chlorophyll Fluorescence Imaging." *Journal of experimental botany* 58, no. 4 (2007): 773-84.

CIP. *Potatoes for the Developing World*. Lima, Peru1984.

Clark, Melody. *Comparative Genomics*. 1 ed.: Springer, 2000.

Clarke, A.E, and P.M Lombard. "Relation of Length of Day to Flower and Seed Production in Potato Varieties.". *American Potato Journal* 16 (1939): 236-44.

R: A Language and Environment for Statistical Computing, Vienna, Austria.

Danan, S., J. B. Veyrieras, and V. Lefebvre. "Construction of a Potato Consensus Map and Qtl Meta-Analysis Offer New Insights into the Genetic Architecture of Late Blight Resistance and Plant Maturity Traits." *BMC Plant Biology* 11 (2011): 16.

DeLacy, I.H., K.E. Basford, M. Cooper, J.K. Bull, and C.G. McLaren. *Analysis of Multi-Environment Trials - an Historical Perspective*. Plant Adaptation and Crop Improvement. 1996.

Eilers, P. H. C., and B. D. Marx. "Flexible Smoothing with B-Splines and Penalties." *Statistical Science* 11, no. 2 (May 1996): 89-102.

Ewing, E. E., and P. C. Struik. "Tuber Formation in Potato: Induction, Initiation and Growth." *Horticultural reviews* 14 (1992): 89-198.

Fahem, M., and A. J. Haverkort. "Comparison of the Growth of Potato Crops Grown in Autum and Spring in North-Africa." *Potato Research* 31, no. 3 (Sep 1988): 557-68.

FAO, [Food and Agriculture Organization]. "Faostat. Fao Statistics Division 2012. ." http://faostat.fao.org/.

Gabriel, K. R. "The Biplot Graphic Display of Matrices with Application to Principal Component Analysis." *Biometrika* 58, no. 3 (December 1, 1971 1971): 453-67.

Gao, L. Z., Z. Q. Jin, Y. Huang, and L. Z. Zhang. "Rice Clock Model - a Computer-Model to Simulate Rice Development." *Agricultural and Forest Meteorology* 60, no. 1-2 (Aug 1992): 1-16.

Gao, Liangzhi, Zhiqing Jin, Yao Huang, and Lizhong Zhang. "Rice Clock Model—a Computer Model to Simulate Rice Development." *Agricultural and Forest Meteorology* 60, no. 1–2 (1992): 1-16.

Gepstein, Shimon, Gazalah Sabehi, Marie-Jeanne Carp, Taleb Hajouj, Mizied Falah Orna Nesher, Inbal Yariv, Chen Dor, and Michal Bassani. "Large-Scale Identification of Leaf Senescence-Associated Genes." *The Plant Journal* 36, no. 5 (2003): 629-42.

Grattapaglia, D., and R. Sederoff. "Genetic Linkage Maps of *Eucalyptus Grandis* and *Eucalyptus Urophylla* Using a Pseudo-Testcross: Mapping Strategy and Rapd Markers." *Genetics* 137, no. 4 (Aug 1994): 1121-37.

Hanneman, R., and S. Peloquin. "Crossability of 24-Chromosome Potato Hybrids with 48-Chromosome Cultivars." *Potato Research* 10, no. 1 (1967): 62-73.

Hartmann, Anja, Tobias Czauderna, Roberto Hoffmann, Nils Stein, and Falk Schreiber. "Htpheno: An Image Analysis Pipeline for High-Throughput Plant Phenotyping." *BMC Bioinformatics* 12 (May 12 2011).

Haun, J. R. "Potato Growth-Environment Relationships." *Agricultural Meteorology* 15, no. 3 (1975): 325-32.

Haverkort, A. J. "Ecology of Potato Cropping Systmens in Relation to Latitude and Altitude." *Agricultural Systems* 32, no. 3 (1990): 251-72.

Haverkort, A. J., and P. L. Kooman. "The Use of Systems Analysis and Modelling of Growth and Development in Potato Ideotyping under Conditions Affecting Yields." *Euphytica* 94, no. 2 (1997): 191-200.

Haverkort, A.J. , and D.K.L MacKerron. "Potato Ecology and Modelling of Crops under Conditions Limiting Growth." Paper presented at the Proceedings of the second international potato modelling conference, Wageningen, 1994.

Hawkes, J. G., and J. Franciscoortega. "The Early History of the Potato in Europe." *Euphytica* 70, no. 1-2 (1993): 1-7.

Hawkes, J.G. *The Potato: Evolution, Biodiversity, and Genetic Resources*. London: Belhaven Press, 1990.

Hijmans, R. J., and D. M. Spooner. "Geographic Distribution of Wild Potato Species." *American Journal of Botany* 88, no. 11 (Nov 2001): 2101-12.

Hodges, T. "Temperature and Water Stress Effects on Phenology." In *Predicting Crop Phenology*. 7-13. Boca Raton, FL: Taylor & Francis, 1990.

Hurtado, P. X., S. K. Schnabel, A. Zaban, M. Vetelainen, E. Virtanen, P. H. C. Eilers, F. A. van Eeuwijk, R. G. F. Visser, and C. Maliepaard. "Dynamics of Senescence-Related Qtls in Potato." *Euphytica* 183, no. 3 (Feb 2012): 289-302.

Hurtado, P.X., B. Bekele-Tessema, S.K. Schnabel, C. Maliepaard, G. van der Linden, P.H.C. Eilers, J. Jansen, F.A. van Eeuwijk, and R.G.F. Visser. "Understanding the Genetic Basis of Potato Development Using a Multi-Trait Qtl Analysis ". *In preparation* (2012).

Ingram, K. T., and D. E. McCloud. "Simulation of Potato Crop Growth and Development." *Crop Science* 24, no. 1 (1984): 21-27.

Innovation, Ad Hoc Panel of the Advisory Committee on Technology, Board on Science, Technology for International Development, and National Research Council. *Lost Crops of the Incas: Little-Known Plants of the Andes with Promise for Worldwide Cultivation*. The National Academies Press, 1989.

Jacobs, J. M. E., H. J. Van Eck, P. Arens, B. Verkerkbakker, B. T. L. Hekkert, H. J. M. Bastiaanssen, A. Elkharbotly*, et al.* "A Genetic Map of Potato (*Solanum Tuberosum*) Integrating Molecular Markers, Including Transposons, and Classical Markers." *Theoretical and Applied Genetics* 91, no. 2 (Jul 1995): 289-300.

Jacobsen, E. "Increase of Diplandroid Formation and Seed Set in 4x X 2x-Crosses in Potatoes by Genetical Manipulation of Dihaploids and Some Theoretical Consequences." *Zeitschrift für Pflanzenzüchtung* 85, no. 2 (1980): 110-21.

Jiang, C., and Z. B. Zeng. "Multiple Trait Analysis of Genetic Mapping for Quantitative Trait Loci." *Genetics* 140, no. 3 (Jul 1995): 1111-27.

Johnson, I. R., and J. H. M. Thornley. "Temperature-Dependence of Plant and Crop Processes." *Annals of Botany* 55, no. 1 (1985): 1-24.

Juzepczuk, S.W., and S.M. Bukasov. "A Contribution to the Question of the Origin of the Potato." Paper presented at the Proceedings of the USSR Congress of Genetics (In Russian). Plant Animal Breed. 3:592–611 Rusia, 1929.

Kellogg, Elizabeth A. "Evolution of Developmental Traits." *Current Opinion in Plant Biology* 7, no. 1 (2004): 92-98.

Khan, M.S., P.C. Struik, P.E.L. van der Putten, H.J. van Eck, F.A. van Eeuwijk, and X. Yin. "An Ecophysiological Modelling Approach to Analyze Variation in Canopy Growth Using a Segregating Potato Population." *Under review* (2012).

Kloosterman, B., M. Oortwijn, J. Uitdewilligen, T. America, R. de Vos, R. G. F. Visser, and C. W. B. Bachem. "From Qtl to Candidate Gene: Genetical Genomics of Simple and Complex Traits in Potato Using a Pooling Strategy." *BMC Genomics* 11 (Mar 2010).

Knott, S. A., and C. S. Haley. "Multitrait Least Squares for Quantitative Trait Loci Detection." *Genetics* 156, no. 2 (Oct 2000): 899-911.

Kooman, P. L., M. Fahem, P. Tegera, and A. J.(a). Haverkort. "Effects of Climate on Different Potato Genotypes 1. Radiation Interception, Total and Tuber Dry Matter Production." *European Journal of Agronomy* 5, no. 3–4 (1996): 193-205.

Kooman, P. L., M. Fahem, P. Tegera, and A. J.(b). Haverkort. "Effects of Climate on Different Potato Genotypes 2. Dry Matter Allocation and Duration of the Growth Cycle." *European Journal of Agronomy* 5, no. 3–4 (1996): 207-17.

Kooman, P. L., and A. J. Haverkort. *Modelling Development and Growth of the Potato Crop Influenced by Temperature and Daylength: Lintul-Potato*. Potato Ecology and Modelling of Crops under Conditions Limiting Growth. edited by A. J. Haverkort and D. K. L. MacKerron. Vol. 3,1995.

Kuromori, Takashi, Takuji Wada, Asako Kamiya, Masahiro Yuguchi, Takuro Yokouchi, Yuko Imura, Hiroko Takabe*, et al.* "A Trial of Phenome Analysis Using 4000 Ds-Insertional Mutants in Gene-Coding Regions of Arabidopsis." *Plant Journal* 47, no. 4 (Aug 2006): 640-51.

Lebreton, C. H., P. M. Visscher, C. S. Haley, A. Semikhodskii, and S. A. Quarrie. "A Nonparametric Bootstrap Method for Testing Close Linkage Vs. Pleiotrophy of Coincident Quantitative Trait Loci." *Genetics* 150, no. 2 (Oct 1998): 931-43.

Lemaga, B., and K. Caesar. "Relationships between Numbers of Main Stems and Yield Components of Potato (*Solanum Tuberosum* L. Cv Erntestolz) as Influenced by Different Daylengths." *Potato Research* 33, no. 2 (Jun 1990): 257-67.

Levy, D., and R. E. Veilleux. "Adaptation of Potato to High Temperatures and Salinity - a Review." *American Journal of Potato Research* 84, no. 6 (Nov-Dec 2007): 487-506.

Li, J., and L. Ji. "Adjusting Multiple Testing in Multilocus Analyses Using the Eigenvalues of a Correlation Matrix." *Heredity (Edinb)* 95, no. 3 (Sep 2005): 221-27.

Li, X., H. J. van Eck, J. N. A. M. Rouppe van der Voort, D. J. Huigen, P. Stam, and E. Jacobsen. "Autotetraploids and Genetic Mapping Using Common Aflp Markers: The R2 Allele Conferring Resistance to Phytophthora Infestans Mapped on Potato Chromosome 4." *Theoretical and Applied Genetics* 96, no. 8 (1998/06/01 1998): 1121-28.

Lin, M., and R. Wu. "A Joint Model for Nonparametric Functional Mapping of Longitudinal Trajectory and Time-to-Event." *BMC Bioinformatics* 7 (2006): 138.

Liu, J. F., Y. J. Liu, X. G. Liu, and H. W. Deng. "Bayesian Mapping of Quantitative Trait Loci for Multiple Complex Traits with the Use of Variance Components." *American Journal of Human Genetics* 81, no. 2 (Aug 2007): 304-20.

Ma, C. X., G. Casella, and R. L. Wu. "Functional Mapping of Quantitative Trait Loci Underlying the Character Process: A Theoretical Framework." *Genetics* 161, no. 4 (Aug 2002): 1751-62.

Ma, Chang-Xing, George Casella, and Rongling Wu. "Functional Mapping of Quantitative Trait Loci Underlying the Character Process: A Theoretical Framework." *Genetics* 161, no. 4 (August 1, 2002 2002): 1751-62.

Malosetti, M., J. M. Ribaut, M. Vargas, J. Crossa, and F. A. van Eeuwijk. "A Multi-Trait Multi-Environment Qtl Mixed Model with an Application to Drought and Nitrogen Stress Trials in Maize (Zea Mays L.)." *Euphytica* 161, no. 1-2 (May 2008): 241-57.

Malosetti, M., R. G. F. Visser, C. Celis-Gamboa, and F. A. van Eeuwijk. "Qtl Methodology for Response Curves on the Basis of Non-Linear Mixed Models, with an Illustration to Senescence in Potato." *Theoretical and Applied Genetics* 113, no. 2 (Jul 2006): 288-300.

Manrique, L. A. "Potato Production in the Tropics - Crop Requirements." *Journal of Plant Nutrition* 15, no. 12 (1992): 2679-726.

Marinus, J, and K.B.A Bodlaender. "Response of Some Potato Varieties to Temperature." *Potato Research* 18 (1975): 189-204.

Maris, B. "Studies Concerning the Relationship between Plant Height of Potatoes in the Seedling Year and Maturity in the Clonal Generations." *Euphytica* 13, no. 2 (1964): 130.

Matsubayashi, M. . "Phylogenetic Relationships in the Potato and Its Related Species." In *Chromosome Engineering in Plants: Genetics, Breeding, Evolution. Part B.*, edited by T.;Gupta Tsuchiya, P. K. . 1991.

Mauricio, Rodney. "Ontogenetics of Qtl: The Genetic Architecture of Trichome Density over Time in Arabidopsis Thaliana." *Genetica* 123 (2005): 75-85.

Midmore, D. J. "Potato (*Solanum* Spp) in the Hot Tropics. 1. Soil - Temperature Effects on Emergence, Plant Development and Yield." *Field Crops Research* 8, no. 4 (1984): 255-71.

Moeller, M., V. Alchanatis, Y. Cohen, M. Meron, J. Tsipris, A. Naor, V. Ostrovsky, M. Sprintsin, and S. Cohen. "Use of Thermal and Visible Imagery for Estimating Crop Water Status of Irrigated Grapevine." *Journal of experimental botany* 58, no. 4 (Mar 2007): 827-38.

Montes, Juan M., Albrecht E. Melchinger, and Jochen C. Reif. "Novel Throughput Phenotyping Platforms in Plant Genetic Studies." *Trends in Plant Sciencie* 12, no. 10 (Oct 2007): 433-36.

Mundermann, L., Y. Erasmus, B. Lane, E. Coen, and P. Prusinkiewicz. "Quantitative Modeling of Arabidopsis Development." *Plant Physiology* 139, no. 2 (Oct 2005): 960-8.

Nelder, J. A., and Wedderbu.Rw. "Generalized Linear Models." *Journal of the Royal Statistical Society Series a-General* 135, no. 3 (1972): 370-&.

Nelder, J.A., and R.W.M. Wedderburn. "Generalized Linear Models." *Journal of the Royal Statistical Society. Series A (General)* (1972): 370-84.

Newey, W. K., and J. L. Powell. "Asymetric Least-Squares Estimation and Testing." *Econometrica* 55, no. 4 (Jul 1987): 819-47.

Ochoa, C.M. *The Potatoes of South America: Bolivia*. Cambridge University press, 1990.

Park, Tae-Ho, Jack Gros, Anne Sikkema, Vivianne G. A. A. Vleeshouwers, Marielle Muskens, Sjefke Allefs, Evert Jacobsen, Richard G. F. Visser, and Edwin A. G. van der Vossen. "The Late Blight Resistance Locus Rpi-Blb3 from Solanum Bulbocastanum Belongs to a Major Late Blight R Gene Cluster on Chromosome 4 of Potato." *Molecular Plant-Microbe Interactions* 18, no. 7 (2005/07/01 2005): 722-29.

Genstat for Windows (13th Edition). Hemel Hempstead, UK.

Pena-Cortes, H., J. J. Sanchez-Serrano, R. Mertens, L. Willmitzer, and S. Prat. "Abscisic Acid Is Involved in the Wound-Induced Expression of the Proteinase Inhibitor Ii Gene in Potato and Tomato." *Proceedings of the National Academy of Science, USA* 86, no. 24 (Dec 1989): 9851-5.

Potato Genome Sequencing, Consortium, X. Xu, S. Pan, S. Cheng, B. Zhang, D. Mu, P. Ni*, et al.* "Genome Sequence and Analysis of the Tuber Crop Potato." *Nature* 475, no. 7355 (Jul 14 2011): 189-95.

Raker, C. M., and D. M. Spooner. "Chilean Tetraploid Cultivated Potato, Solanum Tuberosum, Is Distinct from the Andean Populations: Microsatellite Data." *Crop Science* 42, no. 5 (Sep-Oct 2002): 1451-58.

Reymond, M., B. Muller, A. Leonardi, A. Charcosset, and F. Tardieu. "Combining Quantitative Trait Loci Analysis and an Ecophysiological Model to Analyze the Genetic Variability of the Responses of Maize Leaf Growth to Temperature and Water Deficit." *Plant Physiology* 131, no. 2 (Feb 2003): 664-75.

Ribaut, J. M., M. C. de Vicente, and X. Delannay. "Molecular Breeding in Developing Countries: Challenges and Perspectives." *Current Opinion in Plant Biology* 13, no. 2 (Apr 2010): 213-18.

Rios, D., M. Ghislain, F. Rodriguez, and D. M. Spooner. "What Is the Origin of the European Potato? Evidence from Canary Island Landraces." *Crop Science* 47, no. 3 (May-Jun 2007): 1271-80.

Ruckelshausen, A. "Autonomous Robots in Agricultural Field Trials." Paper presented at the Proceedings of the International Symposium "Agricultural Field Experiments—Today and Tomorrow" 2007.

Salaman, R.N.(a). *The History and Social Influence of the Potato*. London: Cambridge University Press, 1949.

Salaman, R.N.(b). "The Potato in Its Early Home and Its Introduction into Europe." *Journal of the Royal Horticultural Society* 62 (1937).

Sands, P. J., and P. A. Regel. "A Model of the Development and Bulking of Potatoes (*Solanum Tuberosum*). 5. A Simple Model for Predicting Graded Yields." *Field Crops Research* 6, no. 1 (1983): 25-40.

Schnabel, S. K., and P. H. C. Eilers. "Optimal Expectile Smoothing." *Computational Statistics & Data Analysis* 53, no. 12 (Oct 2009): 4168-77.

Schnabel, S. K., P. X. Hurtado, P. H. C. Eilers, C. Maliepaard, and F. Van Eeuwijk. "Statistical Tools for Pre-Processing Phenotypic Data." *In progress* (2012).

Schnabel, S.K., P.H.C. Eilers, and F.A. van Eeuwijk. "Optimal Time Scaling for Plant Growth Analysis. Proceedings." Paper presented at the 26th International Workshop on Statistical Modelling, Universitat de Valencia, Spain, 2011.

Schnabel, S.K., P. Hurtado Lopez, P.H.C. Eilers, and F.A. van Eeuwijk. "Optimal Scaling for Plant Growth Analysis. Poster." Paper presented at the 3rd IBS Channel Network Conference of the International Biometric Society Bordeaux, France, 2011.

Schnute, J. "A Versatile Growth-Model with Statistically Stable Parameters." *Canadian Journal of Fisheries and Aquatic Sciences* 38, no. 9 (1981): 1128-40.

Simmonds, N. W. "Potential of Potatoes in Tropics." *Tropical Agriculture* 48, no. 4 (1971 1971): 291-&.

Expectreg: Expectile and Quantile Regression Version 0.35.

Spooner, D. M., T. Gavrilenko, S. H. Jansky, A. Ovchinnikova, E. Krylova, S. Knapp, and R. Simon. "Ecogeography of Ploidy Variation in Cultivated Potato (*Solanum* Section *Petota*)." *American Journal of Botany* 97, no. 12 (Dec 2010): 2049-60.

Spooner, D. M., K. McLean, G. Ramsay, R. Waugh, and G. J. Bryan. "A Single Domestication for Potato Based on Multilocus Amplified Fragment Length Polymorphism Genotyping." *Proceedings of the National Academy of Science, U.S.A* 102, no. 41 (Oct 2005): 14694-99.

Spooner, D. M., J. Nunez, F. Rodriguez, P. S. Naik, and M. Ghislain. "Nuclear and Chloroplast DNA Reassessment of the Origin of Indian Potato Varieties and Its Implications for the Origin of the Early European Potato." *Theoretical and Applied Genetics* 110, no. 6 (Apr 2005): 1020-26.

Struik, P. C., and E. E. Ewing. *Crop Physiology of Potato (Solanum Tuberosum): Responses to Photoperiod and Temperature Relevant to Crop Modelling*. Potato Ecology and Modelling of Crops under Conditions Limiting Growth. edited by A. J. Haverkort and D. K. L. MacKerron. Vol. 3,1995.

Struik, P. C., E. Van Heusden, and K. Burger-Meijer. "Effects of Short Periods of Long Days on the Development Yield and Size Distribution of Potato Tubers." *Netherlands Journal of Agricultural Science* 36 (1988): 11-22.

Struik, Paul, Jan Geertsema, and Constan Custers. "Effects of Shoot, Root and Stolon Temperature on the Development of the Potato (*Solanum Tuberosum* L.) Plant. 3: Development of Tubers." *Potato Research* 32, no. 2 (1989): 151-58.

Turner, A. D., and E. E. Ewing. "Effects of Photoperiod, Night Temperature, and Irradiance on Flower Production in the Potato." *Potato Research* 31, no. 2 (Jun 1988): 257-68.

Joinmap 4. Software for the Calculation of Genetic Linkage Maps in Experimental Populations., Wageningen, The Netherlands.

Mapqtl 6. Software for the Mapping of Quantitative Trait Loci in Experimental Populations of Diploid Species, Wageningen, The Netherlands.

VanDam, J., P. L. Kooman, and P. C. Struik. "Effects of Temperature and Photoperiod on Early Growth and Final Number of Tubers in Potato (Solanum Tuberosum L)." *Potato Research* 39, no. 1 (1996): 51-62.

Verhulst, P.F. "Recherches Mathématiques Sur La Loi D'accroissement De La Population.", edited by Nouv. mém. de l'Academie Royale des Sci. et Belles-Lettres de Bruxelles, 1-41, 1845.

Genstat for Windows 14th Edition. VSN International, Heel Hempstead, UK.

Werij, JeroenS, Bjorn Kloosterman, Carolina Celis-Gamboa, C. H. Ric de Vos, Twan America, RichardG F. Visser, and ChristianW B. Bachem. "Unravelling Enzymatic Discoloration in Potato through a Combined Approach of Candidate Genes, Qtl, and Expression Analysis." *Theoretical and Applied Genetics* 115, no. 2 (2007/07/01 2007): 245-52.

Wheeler, R. M., K. L. Steffen, T. W. Tibbitts, and J. P. Palta. "Utilization of Potatoes for Life-Support-Systems. 2. The Effects of Temperature under 24-H and 12-H Photoperiods." *American Potato Journal* 63, no. 11 (Nov 1986): 639-47.

Wolf, Jason B., W Anthony. Frankino, Aneil F. Agrawal, Edmund D. 3rd Brodie, and Allen J. Moore. "Developmental Interactions and the Constituents of Quantitative Variation." *Evolution* 55, no. 2 (2001): 232-45.

Wu, R. L., and M. Lin. "Opinion - Functional Mapping - How to Map and Study the Genetic Architecture of Dynamic Complex Traits." *Nature Reviews Genetics* 7, no. 3 (Mar 2006): 229-37.

Wu, R. L., C. X. Ma, M. Lin, and G. Casella. "A General Framework for Analyzing the Genetic Architecture of Developmental Characteristics." *Genetics* 166, no. 3 (Mar 2004): 1541-51.

Wu, R., and M. Lin. "Functional Mapping - How to Map and Study the Genetic Architecture of Dynamic Complex Traits." *Nature Reviews Genetics* 7, no. 3 (Mar 2006): 229-37.

Wu, R., C. X. Ma, W. Zhao, and G. Casella. "Functional Mapping for Quantitative Trait Loci Governing Growth Rates: A Parametric Model." *Physiological Genomics* 14, no. 3 (Aug 15 2003): 241-9.

www.nserc-crsng.gc.ca. "Guidelines for the Preparation and Review of Applications in Interdisciplinary Research." http://www.nserc.ca/pubs/agir/AGIR_e_annex4.pdf.

Yin, X. Y., M. J. Kropff, G. McLaren, and R. M. Visperas. "A Nonlinear Model for Crop Development as a Function of Temperature." *Agricultural and Forest Meteorology* 77, no. 1-2 (Nov 1995): 1-16.

Yin, Xinyou, Martin J. Kropff, Graham McLaren, and Romeo M. Visperas. "A Nonlinear Model for Crop Development as a Function of Temperature." *Agricultural and Forest Meteorology* 77, no. 1–2 (1995): 1-16.

Zaban, A., M. Veteläinen, C. Celis-Gamboa, R. van Berloo, H. Häggman, and R.G.F. Visser. "Physiological and Genetic Aspects of a Diploid Potato Population in the Netherlands and Northern Finland.". *Maataloustieteen Päivät* (2006): 1-7.

Zimnoch-Guzowska, E., W. Marczewski, R. Lebecka, B. Flis, R. Scha¨fer-Pregl, F. Salamini, and C. Gebhardt. "Qtl Analysis of New Sources of Resistance to Erwinia Carotovora Ssp. Atroseptica in Potato Done by Aflp, Rflp, and Resistance-Gene-Like Markers." *Crop Science* 40, no. 4 (/7 2000): 1156-67.

Summary

Potato experiments conducted under field conditions are a tremendous source of multiple types of data in which environmental factors play a major role. In addition to phenotypic data measured only once during the experiments, there are also traits for which data are collected at several instances during the growing season in order to monitor the development of the plants. Therefore, complex analysis involving multiple environments, multiple traits and multiple observations over time can provide the basis for understanding plant development as a complex and dynamic process influenced by different genetic and non-genetic factors. Such studies can also profit from molecular data to identify pleiotropic genetic regions and networks of genes driving developmental and physiological processes. In the last 15 years Wageningen UR Plant Breeding has used a diploid potato population to extensively study developmental, quality and agronomic traits under different field conditions. We had the opportunity to explore the extensive phenotypic data generated in previous studies using this population and we incorporated molecular information to gain insights into the genetic factors controlling development, growth and agronomic traits.

In this thesis we benefited from a multidisciplinary approach combining biology, mathematics and applied statistics to analyse phenotypic and molecular data. We first proposed methods for modelling ordinal and continuous data over time, we described parameters characterizing different development/growth stages and we made use of them to identify QTLs related to these stages and to compare development across environments. Afterwards we combined agronomic and development/growth parameters describing flowering, senescence and plant height in a multi-trait QTL analysis to identify pleiotropic regions explaining relationships between above and below ground traits. Lastly, we performed a multi-environment QTL analysis using data collected under different day lengths (Ethiopia, Venezuela, the Netherlands and Finland). We included development/growth parameters and agronomic traits, such as total number of tubers and total tuber weight among others. We detected QTL by environment interactions (QEI) for almost all the traits, with the strongest QEI observed for total tuber weight, we also identified stable-QTLs across environments, we found the genetic bases of physiological

relationships between traits and we identified pleiotropic regions conserved across environments.

We also present in this thesis relevant issues related to multi-environment trials (METs) and the implications for potato breeding. Despite the great potential of MET data, special attention should be given to three main factors identified during our research: the importance of using standard phenotyping protocols in all the experiments, the value of historical data as a source of phenotypic information in breeding programs and the relevance of multidisciplinary research combining different expertises.

Phenotyping is the most important step of a breeding process and it is also often the limiting step in detection of QTLs in research. Considering that progress in breeding depends on accurate selection of rare genotypes that possess new or improved attributes due to superior combinations of alleles at multiple loci, precise phenotyping is one of the pillars of plant breeding. The use of standard and accurate phenotyping protocols must be promoted within breeding programs to facilitate the comparison of traits across experiments, the understanding of genotype by environment interactions, GEI, and ultimately the detection of the genetic basis of complex traits. Historical data, understood as data generated in previous experiments, are of great value in breeding programs, research and genebanks. These data are expected to be used more in future, in combination with pedigree information, sequence data and data collected in public and breeding databases. Therefore good documentation and a proper storage of historical data is essential for their long term use.

Extending our results to a broader context, this study can be used as an example of complex analyses using historical data. The benefits of such analyses were illustrated and evidence was provided of the valuable information existing in breeding programs: their data just need to be explored. We expect that in the future, especially in automated high-throughput phenotyping experiments across time, our methods will prove to be of great benefit and will become included in the design of breeding better varieties..

Samenvatting

Veldproeven met aardappel zijn een geweldige bron van verschillende soorten gegevens waarin omgevingsfactoren een grote rol spelen. Behalve phenotypische gegevens die slechts één keer zijn gemeten gedurende de proeven, zijn er ook eigenschappen waarvoor op verschillende tijdstippen tijdens het groeiseizoen gegevens verzameld worden om zo de ontwikkeling van de planten de observeren. Om de plant-ontwikkeling te begrijpen als een dynamisch proces beïnvloed door genetische en niet-genetische factoren, zijn complexe analyses nodig over de verschillende omgevingen, de gemeten kenmerken en de observaties in de tijd. Dergelijke studies kunnen ook gebruik maken van moleculaire gegevens bijvoorbeeld om pleiotrope genetische loci en gennetwerken op te sporen die fysiologische ontwikkelingsprocessen sturen. In de afgelopen 15 jaar is er door Wageningen UR Plant Breeding in veldproeven onder verschillende omstandigheden aan een diploïde aardappelpopulatie uitgebreid studie gedaan naar ontwikkelings-, agronomische- en kwaliteits-eigenschappen. Dit gaf ons de mogelijkheid om een groot aantal phenotypische gegevens die in het verleden verzameld zijn te onderzoeken, en we hebben deze gegevens gecombineerd met moleculaire merkers om inzicht te krijgen in de genetische factoren die een rol spelen bij ontwikkeling, groei en ook agronomische kenmerken zoals opbrengst.

In dit proefschrift hebben we een multidisciplinaire aanpak gebruikt waarin biologie, wiskunde en toegepaste statistiek werden gecombineerd om de phenotypische en moleculaire gegevens te analyseren. Om te beginnen hebben we methoden ontwikkeld om ordinale en continue gegevens over de tijd te modelleren en parameters afgeleid die de verschillende stadia in groei en ontwikkeling karakteriseren; deze parameters zijn gebruikt om QTLs (‘Quantitative Trait Loci’) gerelateerd aan die stadia te identificeren en om plantontwikkeling in verschillende omgevingen te vergelijken. Daarna hebben we de agronomische en ontwikkelings/groei-parameters voor bloei, afrijping en planthoogte gecombineerd voor een multi-kenmerk QTL-analyse om pleiotrope regio’s op het aardappel-genoom te identificeren die verbanden tussen bovengrondse en ondergrondse groei verklaren. Tenslotte hebben we een QTL-analyse over meerdere omgevingen gedaan voor gegevens verzameld in locaties met verschillende daglengten: Ethiopië, Venezuela, Nederland en Finland. We hebben daarbij gekeken naar ontwikkelings- en groei-kenmerken alsook naar agronomische eigenschappen zoals onder andere het totaal aantal knollen en

totaal knolgewicht. We vonden QTL maal omgevings-interactie (QEI) voor bijna alle kenmerken, met de grootste interactie voor totaal knolgewicht; we hebben ook stabiele QTLs over omgevingen gevonden, en de genetische basis voor fysiologische verbanden tussen eigenschappen, alsook pleiotrope genoom-regio's die consistent waren voor de verschillende omgevingen.

In dit proefschrift benoemen we ook kwesties die relevant zijn met betrekking tot het uitvoeren van proeven in meerdere omgevingen en de implicaties daarvan voor de aardappelveredeling. Bij het gebruik van gegevens uit meerdere omgevingen zou er speciale aandacht geschonken moeten worden aan drie belangrijke factoren die we tijdens ons onderzoek hebben vastgesteld: het belang van het gebruik van standaardprotocollen voor phenotypering in alle omgevingen, de waarde van historische gegevens als een bron van informatie in veredelingsprogramma's en de relevantie van multidisciplinair onderzoek waarin gebruik gemaakt wordt van verschillende expertises.

Phenotypering is de belangrijkste stap in de veredeling en vaak ook de beperkende stap voor het vinden van QTLs bij onderzoek. Gegeven dat vooruitgang in de veredeling afhangt van nauwkeurige selectie van zeldzame genotypen met nieuwe of verbeterde kenmerken dankzij superieure combinaties van allelen van verschillende loci, is nauwkeurige phenotypering één van de dragers van de veredeling. Het gebruik van gestandaardiseerde en nauwkeurige phenotyperings-protocollen binnen veredelingsprogramma's zou bevorderd moeten worden om de vergelijking van kenmerken over proeven, het begrijpen van genotype maal milieu-interacties en uiteindelijk het vinden van de genetische basis van complexe eigenschappen, te vergemakkelijken. Historische gegevens, gegevens verzameld in eerdere experimenten, zijn van grote waarde in veredelingsprogramma's, in onderzoek en voor genenbanken. Deze gegevens zullen naar verwachting meer gebruikt gaan worden in de toekomst, in combinatie met informatie over de afstamming van accessies, sequentie-informatie van gewassen, en gegevens uit publieke en eigen databanken. Daarom is een goede beschrijving en bewaring van historische gegevens van essentieel belang voor het gebruik ervan op de lange termijn.

Als we onze resultaten bezien in een bredere context, kan deze studie gezien worden als een voorbeeld van complexe analyses van historische data. De voordelen van dergelijke analyses zijn hier toegelicht en we geven aan hoe waardevol de bestaande informatie binnen

veredelingsprogramma's kan zijn: deze moet echter goed onderzocht worden. We verwachten dat in de toekomst, vooral in geautomatiseerde phenotyperings-experimenten in de tijd over grote aantallen planten, onze methoden erg waardevol zullen blijken te zijn en onderdeel zullen worden van experimenten ten behoeve van de ontwikkeling van betere rassen.

Acknowledgements

This part of my thesis takes me directly to Colombia, my beloved country, where this journey started. I could have never thought about doing a PhD without the encouragement, motivation and great inspiration that I got from Martin Fregene, Carmen de Vicente and Myriam Cristina Duque in the International Center for Tropical Agriculture, CIAT. **Martin** and **Carmen** you drove me to professional and personal challenges that made life very dynamic and full of excitement. **Myriam** you were the first one to show me that statistics is the most useful tool a biologist could have to understand genetics.

One of the biggest challenges working at CIAT was being part of the Generation Challenge Program, GCP, and representing the cassava genetics team at international level. The participation in GCP gave me the opportunity to meet one if my promotors, Fred van Eeuwijk. **Fred** I will be always grateful with you for taking me to the next step in my career and introducing me to Richard Visser (my other promotor). **Richard** and **Fred** thanks a lot for this great opportunity of doing my PhD combining the expertise of plant breeding and biometris. You took me to the biggest challenge of my life and without knowing you helped me to develop hidden personal and professional skills, thanks a lot!

This PhD would have never been successful without the great support of the experts beside me, Chris Maliepard, Paul Eilers and Hans Janssen. I always felt honoured to have your help and support. **Chris**, I got support from you from the first day when you took me in the back of your bike. You not only kept the door of your office open every time that I was coming with a question, but you also open the door of your home for me and my family, I am really grateful for that! **Paul**, thanks for all the great ideas, for our discussions and for joining every monthly meeting. **Hans**, Thanks for your patience, for all the time you spend with me trying to answer my questions. I have learned a lot from you and I am very happy to know that we can continue working together now under different circumstances.

I was also lucky to have always the support from Biometris staff, it was making life much more easy during my PhD. M**arcos**, **Joao** and **Martin**, thanks for our discussions and for your support with GenStat issues. Marcos and Joao, I will always remember how inspiring was the GCP workshop in Zaragoza where I met you. Thanks for sharing not only work experiences but also parenthood stories.

I could have never reached this stage of my PhD without Bas Engel and Sabine Schnabel during the first year of my PhD. **Bas**, thanks to your inspiring way of teaching statistics I decided to continue with this challenge, you will always be my favourite teacher. **Sabine**, my dear friend and colleague, thanks for all your support at personal and professional level, you made team work the most enjoyable experience during the last three years. Thanks for being always there for me!

I am also very grateful with the secretaries of Plant breeding. **Annie** thanks for your support, I will never forget the first shoulder in which I cried. **Letty**, how to forget our chats and your charm hearth. **Mariame**, **Janneke** and **Nicole** thanks for making things smooth, easy to organize and for your great smile. **Nicole**, your daughter shoes are giving big steps every day with my little girl.

I am also very grateful with my colleagues and friends from the university, who have made my life in The Netherlands a great experience. I would like to start with my office mates who shared with me special moments of this journey: Songlin, Krisana, Anoma, Rene, Cristal, Paul Bijman, Nick, Luis, Suxian and Rafael. **Rene**, we really appreciate all your advices about cars, thanks for helping us to get our first car in The Netherlands. **Krisana** and **Anoma**, it was very useful to get all your pregnancy tips. **Anoma**, thanks for your friendship and for sharing important moments with me. You, Teun and Maelya are part of my big family in The Netherlands. Cristal, thanks for your advices, for all the talks we had in the office and for your friendship. **Paul** and **Songlin**, thanks for giving me insights about your beautiful flowers.

Thanks to all the great friends who made the first year of my PhD a wonderful time despite the difficulties, **Hulya**, **Shiva**, **Leila**, **Arwa**, **Anitha**, **Nasim**, and **Animesh**. You were the first pillar of my new life in Wageningen. **Hulya**, **Serdal** and **Efe**, you are our family in Turkey; **Animesh**, **Sharmi** and **Annirudho** you are also an extension of our family no matter where you are.

Thanks to all my friends in Radix for sharing little bits of your life with me. **Ram**, **Yusuf**, **Sabas**, **Myluska**, **Andres**, **Stathis**, **Gullia**, **Stelle**, **Pierre-Yves**, **Yolanda**, **Pauline**, **Erica**, **Catherijn**, **Caucasela**, **Mr.Ali**, **Mr.Khan**, **Gisella**, **Thomas O**, **Daniella-Marcelo**, **Alba**, **Patricia**, **Priscilla**, **Caroline**, **Willem**, **Ana-Carolina** and all the others who shared coffee breaks**,** lunches or dinners in Forum with me. **Ningwen** and **Collete**, thanks for sharing your

motherhood experiences. **Helen F**, thanks for always trying to teach me Dutch. **Willem** Stol, thanks for your enthusiasm to learn Spanish and for always caring about me even if we did not see each other very often.

How to forget the kindness of **Diana** Londoño, who picked me up in Schiphol when I arrived in the Netherlands; **Enrique** and **Claudia**, who took care of me the first two nights in Bornsesteeg; **Natalia** Carreño and **Luis** Montes who made my first birthday in the Netherlands a great celebration and have been a great emotional support in the last four years. **Luis**, it seems that life always wants to bring us together wherever we go, so I am sure we will continue sharing special moments.

Beside the university life, one of the more valuable things in Wageningen is the opportunity to meet people from all over the world and also to strength the ties with my Colombian and Latin-American fellows. The almost weekly gatherings have made my life in Wageningen full of fun. Thanks to all my dear friends who have been part of this wonderful journey in The Netherlands. I fell blessed with all the great friends God have put in my way, you are my angels: **Carito H** and **Roland**, **Bjorn** and **Naty**, **Natica** Piñeros, **Naty** Patiño, **Madhu** and **Manoj**, **Marce** and **Rafa**, **Miguel** and **Maria**, **Caro U** and **Robert**, **Indira** and **Chris**, **Jose** and **Nanine**, **William** van Dijk, **Eric** and **Henriet**, **Chris** and **Joost**, **Esperanza** and **John**, **Nancy**, **Suly** and **Kees**, **Diana** Z and **Luigi.** Thanks to all of you for opening the doors of your house, you have made us part of your family and this is priceless.

I will never forget the invigorating hugs that I got from **Eric** van de Weg and **Gerard** van der Linden the last weeks of my writing phase. You have no idea how much I appreciate such hugs, thanks for caring about me and my family. **Eric**, thanks for taking me to a new dancing experience, for always passing by and caring about my personal issues, thanks for making us part of your thoughts.

How to forget my housemate (for two years), my brother and friend: Mahmoud Haidar. **Maha**, thanks for your friendship, for being part of our family and for making our life so enjoyable. **Neven**, you and your little princess (**Nathaly)** are another extension of our family in the Netherlands. A big thank for my other housmates who came to our place for short periods but were actively involved in our daily activities. **Nara**, you and your mom were two

more angels present in the most difficult part of this amazing journey, thanks for all your love!

I feel very proud of reaching this stage in my life and also to have been able to combine motherhood with my PhD. Of course it would have never been possible without the great help from all my family. Papitos ustedes han sido siempre mi motor, mi inspiracion y sin ustedes nunca habria llegado hasta aqui. Gracias por el apoyo incondicional, por el gran ejemplo que me han dado y por estar presentes en los momentos mas importantes de mi vida, los amo con todo mi Corazon!

Doña Lucia y Don Jesus gracias por su apoyo, su amor y su paciencia. Ustedes son mis segundos papitos, el gran ejemplo de bondad y lucha. Don Jesus, se que desde el cielo nos sigue cuidando y sus enseñanzas y consejos estaran siempre presentes. Que placer fue compartir una parte de mi vida con usted. A mis hermanitos Juan Felipe y Natalia, gracias por su presencia en la distancia. A mis cuñados Diana, Laura, Yovany y Shanna gracias por todo su cariño.

Last but not least all my gratitude, love and respect to the person who have been building with me this life path for the last 10 years. Mi Cesi, gracias por todo tu amor, tu ayuda, tus sacrificios y por recorrer conmigo este camino, te amo mi preciosito. Thanks for encouraging me to assume new challenges and for being my grounding when my optimism makes me drift away from reality. Thanks for sharing with me the most amazing experience one can have in life, parenthood. We are enjoying together the most precious gift from God, our little angel Maria Antonieta. The last words for you my little angel, you are my everything.

Paula Ximena Hurtado Lopez

Wageningen, December 2012

About the author

Paula Ximena Hurtado López was born on July 10, 1979 in Bogota, Colombia. She is the first child of a traditional Colombian family consisting of parents (Juan Pablo and Maria Helena), two girls (Paula and Natalia) and one boy (Juan Felipe). In 2001 she got her Bachelor degree in Microbiology from Universidad de los Andes, Colombia after doing an internship in Bioremediation at CIMIC (Microbiological research center). On January 15th 2002 she moved to Palmira, Colombia to carry out her master thesis in the International Center for Tropical Agriculture (CIAT). Her thesis was part of a collaborative project between Cassava pathology and Cassava Genetics and her research was awarded with the National phytopathology prize "Rafael Obregon". Her MSc thesis was laureate and she got her master degree with honors in Molecular Biology from Universidad de los Andes in 2004. In 2003 she started working as research assistant in the Cassava Genetics group under the leadership of Dr. Martin Fregene. Between 2004 and 2008 she was coordinating the genotyping activities of several Cassava projects under the Generation Challenge program (GCP). Her participation in international projects took her to India, Brasil, Australia, Tanzania and Spain representing the Cassava team. One of the international projects puts her in contact with the Biometris department of Wageningen University to get statistical support in the analysis of data generated under GCP. In 2008 she got a fellowship from Plant Breeding and Biometris to carry out her PhD in Wageningen University. She moved to the Netherlands on June 22nd 2008 and she started working on a project investigating genotype by environment and QTL by environment interactions in Potato. Through the analysis and interpretation of potato data generated in the last 15 years she developed skills in data analysis, quantitative genetics and statistics. After 4 years and 3 months she has successfully finished her PhD and the results of her research are described in this thesis. On September 17th, 2012 she joined Rijk Zwaan Breeding B.V as Researcher in Quantitative Genetics and she is enjoying her dream job in The Netherlands next to her husband (Cesar) and her daughter (Maria Antonieta).

PE&RC PhD Education Certificate

With the educational activities listed below the PhD candidate has complied with the educational requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)

Review of literature (5.6 ECTS)

- Biometris colloquium (2009)
- Plant Breeding colloquium (2009)

Writing of project proposal (4.5 ECTS)

- Investigating GxE and QTLxE interactions for developmental traits in potato (2009)

Post-graduate courses (4.8 ECTS)

- Natural variation in plants; EPS (2008)
- Systems biology course: statistical analysis of ~omics data; EPS (2008)
- Introduction to R for statistical analysis; PE&RC (2009)
- Multivariate analysis; PE&RC (2009)

Laboratory training and working visits (0.3 ECTS)

- Bioinformatics, biostatistics and phytopathology; Scottish Crop Research Institute (SCRI) (2009)

Invited review of (unpublished) journal manuscript (2 ECTS)

- Heredity: functional mapping of biomass growth trajectories in soybean (2010)
- Revista cientifica UDO Agricola: detection of SSR markers associated with resistance to Curvularia in Maize (2011)

Deficiency, refresh, brush-up courses (3 ECTS)

- Plant breeding (2008)
- Advanced statistics (2008)
- Modern statistics for the life sciences (2009)

Competence strengthening / skills courses (3.3 ECTS)

- Competence assessment; WUR (2009)
- Academic writing II; Language Centre (2009)
- Working with EndNote; WUR-Library (2010)
- Techniques for writing and presenting a scientific paper; Graduate schools (2011)

PE&RC Annual meetings, seminars and the PE&RC weekend (2.4 ECTS)

- PE&RC Days (2010, 2011)
- PE&RC Weekends (2010, 2011)

Discussion groups / local seminars / other scientific meetings (5.4 ECTS)

- Biometris colloquium – Statistical Genetics (2008-2012)
- Plant breeding colloquium (2008-2012)
- Literature discussions Plant Breeding (2009-2011)
- Dies Natalis symposium “The world in transition” (2010)
- Technology transfer in the plant sciences (2010)

International sysmposia, workshops and conferences (9.2 ECTS)

- XIV Meeting of the Eucarpia meeting Biometris in Plant breeding (2009)
- QTL-MAS Workshop (2009, 2010)
- EAPR-Eucarpia Congress “Potato breeding after completion of the DNA sequence of the potato genome” (2010)
- EPS Lunteren (2010)
- XVIII Triennial conference of the EAPR (2011)
- XV Meeting of the Eucarpia section Biometris in Plant breeding (2012)

Lecturing / supervision of practical ‘s / tutorials (1.5 ECTS)

- Plant breeding (2009, 2010)
- Breeding for resistance and quality (2009, 2010, 2012)

Cover design:
Cesar Ospina

Cover Art:
Cesar Ospina
Natalia Moreno
Sabine Schnabel

Printed by: CPI Wöhrmann Print Service, Zutphen, The Netherlands