

# Phenotypic Analyses of Multi-Environment Data for Two Diverse Tetraploid Potato Collections: Comparing an Academic Panel with an Industrial Panel

Björn B. D'hoop · Maria João Paulo ·  
Richard G. F. Visser · Herman J. van Eck ·  
Fred A. van Eeuwijk



Received: 14 April 2010 / Accepted: 31 January 2011 /

Published online: 8 April 2011

© The Author(s) 2011. This article is published with open access at Springerlink.com

**Abstract** Phenotypic analyses of two different association panels of tetraploid potato cultivars are presented. Association panels are sets of variously related genotypes assembled for association analysis purposes. The aims of this research were to inspect, analyse and compare two phenotypic data sets, a first step in association mapping analysis. A first panel of 205 contemporary and historical cultivars, selected to represent the commercial potato germplasm pool, was evaluated in two trials in 2006, one on sandy soil and the other on clay soil, both with two replications. It was called the academic panel. Data for the second panel with 299 genotypes were compiled from contributions from five breeding companies and included 66 locations and 18 years. Each of the participating breeding companies contributed data from their clonal selection programmes for 38

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s11540-011-9186-1) contains supplementary material, which is available to authorized users.

B. B. D'hoop · M. J. Paulo · R. G. F. Visser · H. J. van Eck  
Laboratory of Plant Breeding, Wageningen University, 6708 PB Wageningen, Netherlands

B. B. D'hoop · R. G. F. Visser · H. J. van Eck  
Graduate school of Experimental Plant Sciences, 6708 PB Wageningen, Netherlands

B. B. D'hoop · M. J. Paulo · R. G. F. Visser · H. J. van Eck (✉) · F. A. van Eeuwijk  
CBSG, Centre for BioSystems Genomics, 6708 PB Wageningen, Netherlands  
e-mail: herman.vaneck@wur.nl

B. B. D'hoop  
e-mail: bjorn.dhoop@wur.nl

M. J. Paulo · F. A. van Eeuwijk  
Biometris, Wageningen University & Research Centre, 6708 PB Wageningen, Netherlands

*Present Address:*

B. B. D'hoop  
Rijk Zwaan, 2678 KX De Lier, Netherlands

advanced breeding clones and a series of standard cultivars. It was called the industrial panel. Variance components for genotypic main effects and genotype-by-environment interactions were calculated, and estimates for the random genotypic main effects were produced. The genotypic main effects for 19 agro-morphological and quality traits were used to study trait by trait correlations within each panel. In addition, for the genotypes shared by both panels, the correlation of genetic main effects between the panels was investigated. The heritability of all traits was high and no large differences were observed between panels. Coefficients of trait variation were highly correlated ( $r=0.9$ ) for both panels and trait by trait correlations in both panels showed highly similar patterns. These results demonstrate that a single-year balanced field trial as well as using breeders' records yields robust phenotypic information that can be used in a genome-wide association study. Issues related to data management and definition of traits are discussed.

**Keywords** Genotype-by-environment interaction · Heritabilities · Mixed models · Potato

## Introduction

Performance of cultivars across environments can be captured in the form of cultivar main effects. Good estimates of these effects are of great value for compiling national lists of crop species to produce reliable recommendations. Also for potato, such estimates are valuable. However, the production of such estimates requires caution. Potato cultivars are maintained as vegetatively propagated clones. The origin and physiological status of the planting material (seed tubers) may already have a profound impact on the plant (Asiedu et al. 2003). Therefore, it is required to plant trial fields with seed tubers which have been propagated, stored and sprouted under identical conditions for at least one propagation cycle. Potato cultivars may serve different purposes and market niches. Since the 1970s and 1980s, potato breeding companies have specialised and consequently their gene pools have increasingly diversified. Currently, cultivars are developed for specific markets such as the processing industry (pre-baked frozen fries, crisps and flakes), fresh consumption (domestic and export markets) or the starch industry, each with their own specific requirements (Kirkman 2007).

Studies on the inheritance of complex traits are nowadays no longer limited by the number of genetic data points. High multiplex SNP assays and other DNA marker techniques have enabled the construction of marker-dense linkage maps (van Os et al. 2006), as well as marker-dense genome-wide association scans (Zhu et al. 2008). Accurate phenotypic information of biological samples has become the limiting step in both linkage and association mapping studies. In plants, large  $F_2$ , backcross or recombinant inbred mapping populations can be generated easily, which in animal genetics is not so straightforward. Phenotypic information always has intrinsic inaccuracies due to the methods of data collection. Replication of observations through installing extra field trial years or locations will have a diminishing return in the end, while including the number of genotypes will overcome limitations in phenotyping accuracy. In human genetics, controlled crosses and replications are imaginary only, and therefore, association genetics with thousands of patients is habitually used by necessity. In plants, association studies have gained popularity,

mostly because association mapping allows historical phenotypic data to be used for quantitative trait locus (QTL) mapping purposes (Flint-Garcia et al. 2003).

It is generally accepted that association studies (studies aimed at identifying QTL by statistically associating trait and marker data using collections of variously related germplasm) using well-known cultivars (both contemporary and historical) have two major attractive properties compared to QTL studies. First, in a QTL mapping population, the number of segregating alleles is limited to the specific alleles of the crossing parents, which is far from a comprehensive reflection of the genetic and phenotypic trait diversity of the entire gene pool. Second, phenotypic data may already be available for these cultivars. In potato, for example, the ECPGR Potato Working Group (a continuation of EU project RESGEN-CT95-34/45) aims to collect phenotypic data ([www.europotato.org](http://www.europotato.org)). Unfortunately, it appeared that these data emphasise disease resistance, are highly incomplete with respect to quality traits and are highly incomplete with respect to cultivars released in the recent years.

In the framework of a public–private research effort ([www.cbsg.nl](http://www.cbsg.nl)), two association panels (sets of variously related germplasm assembled for association studies) have been composed. One panel comprises 205 contemporary and historical cultivars (the academic panel), and the other panel (the industrial panel) contains 299 cultivars including recent breeds from five Dutch breeding companies next to a series of standard cultivars. The academic panel, in distinction to the industrial panel, was chosen to represent the commercial potato gene pool (excluding native potato land races from Latin America) of the previous century with emphasis on the oldest “founding fathers”, the most widely used in cultivation or the most widely used as crossing parents (major contributing ancestors, see Love 1999), as well as progenitor clones, representing widely used donors of disease resistance genes. The industrial panel covers mostly advanced clones just before granting Plant Breeders’ Rights. Obviously, breeders are hesitant to release this material for centralised phenotypic analysis, but in this project breeders gladly shared all existing multi-year–multi-location data obtained during the clonal selection process.

In this paper, we describe, analyse and compare two phenotypic data sets (“the academic panel” and “the industrial panel”). The industrial panel is based on existing multi-year–multi-location observations of recent breeds and standard cultivars at five breeding companies. The academic panel is based on a single-year field experiment with contemporary and historical potato cultivars. Both data sets are intended for use in a genome-wide association study, and this paper will describe the approach used to generate adjusted mean genotypic values. Furthermore, the strengths and weaknesses of both panels will be analysed. Finally, broader aspects related to phenotypic data sets and their analyses are discussed.

## Materials and Methods

### Plant Material

Tubers of 205 potato cultivars, listed in Online Resource 1, have been collected from five Dutch breeding companies (Agrico Research, Averis seeds, C. Meijer, HZPC Holland BV and Van Rijn-KWS) or were kindly provided by gene banks [Agriculture and Agri-Food (Canada), Arche Noah (Austria), IPK Gatersleben

(Germany), INRA (France), SASA (Scotland), Teagasc (Ireland) and USDA (USA)]. Uniform planting material was generated by propagation of seed tubers in 2005 to perform the 2006 field trials, with a few exceptions for in-vitro-grown cultivars donated by overseas gene banks.

### Field Trial Design Academic Panel

For the academic panel, it appeared problematic to use existing data for many reasons. Existing data were highly incomplete and were obtained in periods with outmoded agronomic practices (e.g. obsolete cultivars), and traits were evaluated with different methods or recorded on scales that do not relate to recent practices in varietal assessment. Therefore, it initially appeared necessary to collect new phenotypic values on all contemporary quality traits for the academic panel.

In 2006, the 205 contemporary and historical potato cultivars were grown on trial fields at two locations near Wageningen (Netherlands), one with sandy soil and one with river clay. Both locations harboured two replicates. The cultivars were arranged into maturity classes according to their maturity type (early maturing–main crop–late maturing) to avoid competition as much as possible. Plots were four-plant units. Plants were spaced at 40 cm within rows and at 75 cm between rows. Plots were randomised within maturity classes, while maturity classes were randomised within replicates. Tubers were planted on April 26 and harvested from October 9 to 11 at the river clay location and planted on May 3 and harvested from September 28 to October 2 at the sand location. Three weeks before harvest, vines were killed. No irrigation was applied. The fertility regime followed standard commercial potato production field practices.

### Phenotypic Observations

Agro-morphological traits were evaluated in a similar way for both panels. For the academic panel, these traits were recorded during the growing season in the field and using the harvested material of these fields. Phenotypic values of quality traits for both the academic and industrial panels were assessed according to standard protocols of the breeding companies. In this study, we focused on the non-enzymatic discolouration due to oxidation of an iron–chlorogenic acid complex (McKenzie et al. 2005), which is observed after baking of French fries (after baking darkening, ABD) or after boiling potatoes (after cooking darkening, ACD). We also paid attention to the Maillard discolouration of French fries due to reducing sugars (frying colour, FryingCol). The traits ABD and FryingCol were assessed after varying months of storage of the harvested material. Another form of discolouration, observed on peeled or grated tubers, is caused by the enzymatic oxidation of phenolic compounds by *polyphenol-oxidase* (enzymatic browning, EnzBrow), see Lærke et al. (2002). Tuber flesh colour of potato is usually recorded as an ordinal trait, where values below 5.5 are regarded as white fleshed, while yellow-fleshed potato is ranked six and higher depending on the intensity of the yellow colour. The white–yellow contrast, controlled by the well-known *Y*-locus on chromosome 3 encoding the gene *beta-carotene hydroxylase* (Brown et al. 2006), was not regarded as a relevant quality trait. In this study, only the intensity of the yellow pigmentation was recorded, which means that a truncated scale was used starting from light yellow

six to orange nine, in an attempt to study the genetic control of tuber flesh yellowness, both in raw tuber flesh (FleshColY) and in cooked tuber tissue (CookColY). Furthermore, we recorded cooking type, maturity, tuber shape, tuber size and starch concentration (underwater weight). An overview with descriptions and scales of all the traits measured is presented in Table 1.

## Data Analysis

To get an indication of the feasibility of genome-wide association mapping with our phenotypic data sets (academic and industrial), we studied the variance components for the genotypic main effect and genotype-by-environment interaction effects, as well as the heritabilities. Furthermore, we looked at the genetic correlations between traits within both the academic and industrial data sets. Finally, we looked at genetic correlations between both data sets for the genotypes that were common to both association panels. The genetic correlations were calculated from the estimates for the random genotypic main effects produced with appropriate mixed models for the individual traits in both data sets. For the analysis, we used the mixed model facilities of GenStat, 11th edition (VSN International Ltd., Oxford, UK).

## Academic Panel

The academic panel contained phenotypic values obtained during a 1-year–two-locations field trial with replicates. The genotypes in the academic panel were representative for the commercial potato germplasm because cultivars were selected based on (1) the acreage for seed potato production in the Netherlands, (2) the country of origin, (3) the year of commercial introduction, (4) the market niche, (5) phenotypic diversity for quality traits and (6) availability of the cultivars, an important criterion for older genotypes which formed about one third of the academic panel.

For the analysis of the academic panel, the following model was used:

$$\begin{aligned}
 \text{Response} &= \text{Location} + \text{Maturity class} \\
 &+ (\text{Location} - \text{by} - \text{maturity class interaction}) \\
 &+ (\text{Replicate within location}) + (\text{SecGr within replicate}) \\
 &+ (\text{Green within replicate}) + \underline{\text{Genotype}} \\
 &+ (\underline{\text{Genotype} - \text{by} - \text{location interaction}}) + \underline{\text{Error}}
 \end{aligned}$$

Most terms in the model are self-explanatory. Random terms are underlined. Maturity class was a factor with three levels (early, main crop and late). SecGr and Green were two covariates representing secondary growth and tuber greening, respectively. Both disturbances were caused by delayed and excessive rain in the 2006 growing season.

Heritabilities for cultivar means across locations and replicates were calculated according to:  $h^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_{gl}^2/n_l + \sigma_e^2/n_e n_r)$ , with  $\sigma_g^2$  the variance component for cultivar main effects,  $\sigma_{gl}^2$  the variance component for genotype-by-location interaction and  $\sigma_e^2$  the error variance. The number of environments,  $n_e$ , and the

**Table 1** Abbreviated names of phenotypic traits along with a short description of the protocol, the data type and the scale

Trait	Trait description	Description protocol	Type	Scale
ABD_Nov8c	After baking darkening	Shades of grey scored 1 h after frying in November after storage at 8°C	Ordinal	3 (very dark) to 8 (very light) with intermediates
ABD_Feb8c	After baking darkening	As above after storage until February at 8°C	Ordinal	3 (very dark) to 8 (very light) with intermediates
ABD_Apr8c	After baking darkening	As above after storage until April at 8°C	Ordinal	3 (very dark) to 8 (very light) with intermediates
ABD_Apr4c	After baking darkening	As above after storage until April at 4°C	Ordinal	3 (very dark) to 8 (very light) with intermediates
ABD_24hNov8c	After baking darkening	Shades of grey scored 24 h after frying in November after storage at 8°C	Ordinal	3 (very dark) to 8 (very light) with intermediates
ACD_1h	After cooking darkening	Shades of grey scored 1 h after cooking	Ordinal	3 (very dark) to 8 (very light) with intermediates
ACD_24h	After cooking darkening	Shades of grey scored 24 h after cooking	Ordinal	3 (very dark) to 8 (very light) with intermediates
CookCoLY	Intensity of yellow flesh in cooked tissue	Scored 1 h after 30 min of steam cooking	Ordinal	6 light yellow, 7 yellow, 8 dark yellow, 9 orange
Cooking type	Cooking type	Scored 1 h after 30 min of steam cooking	Ordinal	1 (A, firm) to 7 (D, very mealy) with intermediates
EnzBrow_30min	Enzymatic browning	Scored 30 min after exposure of grated raw tuber tissue to air	Ordinal	3 (very dark) to 8 (very light) with intermediates
FleshCoLY	Intensity of yellow flesh in fresh tissue	Scored immediately after cutting tubers	Ordinal	6 light yellow, 7 yellow, 8 dark yellow, 9 orange
FryingCol_Nov8c	Maillard discolouration of French fries	Scored 1 h after 3 min frying at 180°C using an IBVL chart in November after storage at 8°C	Ordinal	3 (very dark) to 8 (very light) with intermediates
FryingCol_Feb8c	Maillard discolouration of French fries	As above after storage until February at 8°C	Ordinal	3 (very dark) to 8 (very light) with intermediates
FryingCol_Apr8c	Maillard discolouration of French fries	As above after storage until April at 8°C	Ordinal	3 (very dark) to 8 (very light) with intermediates
FryingCol_Apr4c	Maillard discolouration of French fries	As above after storage at 4°C until April	Ordinal	3 (very dark) to 8 (very light) with intermediates
Maturity	Maturity type	Visual field observations end of July, half September	Ordinal	3 (extremely late) to 9 (extremely early) with intermediates
Shape	Tuber shape	Visual observation on tubers from 4 plants	Ordinal	1 long, 3 oval, 5 round with intermediates for academic panel. The industrial data were recorded in 9 classes

Size	Visual impression of the proportion of the tubers with a tuber diameter exceeding 55 mm	Visual observation on tubers from 4 plants	Ordinal	3 (predominantly undersized tubers) to 9 (predominantly large tubers) with intermediates
Underwater weight	Dry matter content; in particular starch composition	Measurements on tubers from 4 plants	Continuous	Gram per 5 kg tubers

number of locations,  $n_l$ , were two. The number of replicates within locations,  $n_r$ , was two as well. The variance components  $\sigma_g^2$  and  $\sigma_{gl}^2$  were estimated for each maturity class separately. The error variance,  $\sigma_e^2$ , was estimated across maturity classes. Coefficients of variation (%CV) were calculated according to:  $\%CV = 100 * \sqrt{(\sigma_e^2) / \text{mean}}$ .

Furthermore, heritabilities across the three maturity classes were calculated by a weighted average with weights proportional to the number of genotypes in each class. Likewise, overall coefficients of variation were calculated.

Although installing potato cultivars according to their maturity class in the field is a common practice for potato breeding trial fields, variance component and heritability estimation will likely be influenced due to the subsequent inclusion of a maturity class term and its interaction with the environment in the phenotypic analysis model. To appraise the importance of this effect, a simplified model, in which all maturity class-related terms were omitted, was tested. The principal findings related to removal of maturity class from the model are presented in the “[Discussion](#)” section.

Most traits were measured on an ordinal scale. Nevertheless, we analysed these traits without transformation as we found that the diagnostic plots did not show serious violations of the standard assumptions for the application of mixed models. This was true for both the academic and industrial panels.

### Industrial Panel

The industrial data set was composed of phenotypic values of 190 advanced breeders’ clones (38 per breeding company), supplemented with additional observations on a variable series of standard potato cultivars, leading to a total of 299 genotypes. The 109 standard cultivars are well-known and widely grown cultivars used by each of the participating breeding companies. A majority of the standards (101 out of 109) was also included in the academic panel. Obviously, the companies use different sets of standard cultivars, in view of their differences in market niches. Data of the industrial data set were assembled throughout a variable number of successive years of clonal selection and commercial testing of the 190 recent breeds. Standard commercial potato production field practices were applied to all of the included trials in the industrial panel. [Table 2](#) presents an overview of the structure of the industrial data set. The number of tested genotypes is highly variable, depending on the trait, simply because cultivars for the starch industry are habitually not evaluated for cooking and processing traits. Conversely, observations on processing and starch properties (not part of this study) are poorly known for cultivars used for fresh consumption. [Table 2](#) also shows the number of locations, the number of years and the number of trials in which a particular trait was measured. As illustrated in [Table 2](#), the phenotypic observations on agromorphological traits were abundant, whereas data on processing traits were scarce. Plot sizes were in general 16 plants per genotype, in contrast to the academic panel, where plot sizes were 4 plants.

The model for the trait values in the industrial data set was:

$$\text{Response} = \text{trial} + \left( \text{repeatwithintrial} \right) + \text{genotype} + \left( \text{genotype} - \text{by} - \text{year interaction} \right) \\ + \left( \text{genotype} - \text{by} - \text{trial} - \text{within} - \text{year interaction} \right) + \text{error}$$



**Table 2** Illustration of the structure of the industrial panel

Traits	No. of genotypes	No. of locations	No. of years	No. of trials
ABD_Nov8c	151	9	8	39
ABD_Feb8c	130	5	8	21
ABD_Apr8c	113	3	3	9
ABD_Apr4c	77	3	7	15
ABD_24hNov8c	29	14	2	16
ACD_1h	256	33	8	103
ACD_24h	116	26	5	54
CookCoIY	151	5	8	37
Cooking type	258	33	8	103
EnzBrow_30min	98	1	3	3
FleshCoIY	267	50	18	219
FryingCol_Nov8c	180	25	8	66
FryingCol_Feb8c	161	8	8	34
FryingCol_Apr8c	133	5	5	19
FryingCol_Apr4c	108	6	7	29
Maturity	275	21	15	117
Shape	295	66	18	249
Size	252	64	8	173
Underwater weight	292	58	18	255
Total data set	299	85	18	342

For each phenotypic trait, the number of genotypes indicates on how many genotypes the trait was evaluated. The number of locations and the number of years are indicative for how many locations and years were involved in trait evaluation. The number of trials should be interpreted as in how many combinations of year and location a specific trait has been tested

Trial represents the factorial combination of location and year. “Repeat within trial” is a kind of replicate within trial effect. We use the term “repeat” instead of “replicate”. “Replicate”, as used for the academic panel, means fully duplicated for all traits. “Repeat” in case of the industrial data set indicates that a trait has been assessed in a variable number of replicates, depending on the trait and the breeding company. Again, random terms are underlined, from which it becomes evident that for these data we used a fully random model. The large number of trials made it statistically attractive to use a purely random model.

Heritabilities were calculated as  $h^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_{gy}^2 + \sigma_{gty}^2 + \sigma_e^2)$ , with  $\sigma_g^2$  the variance component for the genotypic main effect,  $\sigma_{gy}^2$  the variance component for the genotype-by-year interaction,  $\sigma_{gty}^2$  the variance component for the genotype-by-trial-within-year interaction and  $\sigma_e^2$  the error variance. As the numbers of repeats varied strongly among genotypes, we first calculated heritabilities on a per plot basis instead of on a per genotype-mean basis.

To compare heritabilities between the academic and industrial data sets, we calculated adjusted heritabilities according to:  $h^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_{gty}^2/2 + \sigma_e^2/4)$ , where the variance component for genotype-by-year interaction was left out of the

formula as this component could not be estimated in the academic data set. Analogous to the heritability calculation for the academic data set, we divided the genotype-by-trial-within-year variance component by two and the error component by four. Also for this data set, we calculated coefficients of variation according to:  $\%CV = 100 * \sqrt{(\sigma_e^2)}/\text{mean}$ .

### Comparative Analysis of Both Association Panels

The genotypic main effects obtained with the above mixed models were used to inspect correlation structures within the academic and industrial data sets. In addition, for individual traits, we inspected the correlation ( $r$ ) between the academic and industrial data sets. As we calculated the correlations based on the estimates for the random genotypic main effects, we were effectively investigating the structure of the genetic correlations within and between the two data sets. To visualize the within-set genetic correlations, we produced biplots based on a principal component analysis of standardized genetic effects. For the biplots, we used the software SC-Biplot (Cosinus Computing BV, Waalwijk, Netherlands).

## Results

More than 15,000 data points were collected during the academic field trial. Nineteen phenotypic traits were evaluated for 205 potato genotypes on two locations with two replicates ( $205 * 2 * 2 * 19 = 15,580$ ). The industrial data set was even larger, with 73,968 observations. This number, however, represented only about 4% of the possible data points that would have been obtained from the evaluation of all 299 genotypes in all 342 trials for all 19 traits (Table 2).

Estimates for variance components and their standard errors for the academic panel are shown in Table 3. An analogous overview is presented for the industrial panel in Table 4. Academic heritabilities are plotted versus industrial heritabilities in Fig. 1.

In general, high heritabilities were obtained. For the academic field trial, the heritabilities ranged from 0.53 for ABD\_Apr4c to 0.95 for CookCoLY. For the industrial data set, the heritabilities ranged from 0.33 for ABD\_Feb8c to 0.87 for underwater weight. However, when we compared the academic heritabilities with the comparably scaled industrial heritabilities, there was for the majority of traits a clear trend for the industrial heritabilities to be higher (see Fig. 1).

The coefficients of variation for the academic and industrial panels are plotted in Fig. 2, from which it can be deduced that there was consistency between both data sets. Traits that had a higher coefficient of variation in one data set also showed a higher coefficient of variation in the other data set, likewise for the traits with a lower coefficient of variation.

Figure 3 presents a biplot showing the traits following a principal components analysis on standardized genotypic effects (Gabriel 1971, 1978). Figure 3 thus gives an impression of the genetic correlations within both data sets, with the academic field trial in Fig. 3a (46% of the original variation) and the industrial data set in Fig. 3b (40% of the original variation). Acute angles between trait directions

**Table 3** Results of the phenotypic analysis of the academic field trial

Traits	Maturity	$V_g$ ( $\pm$ SE)	$V_{gl}$ ( $\pm$ SE)	Error ( $\pm$ SE)	$h^2$	Predicted means	%CV	%present
Shape	Early	1.75 ( $\pm$ 0.48)	0.06 ( $\pm$ 0.14)	0.32 ( $\pm$ 0.02)	0.90	2.61	19.7	100
	Mid	2.32 ( $\pm$ 0.40)	0.38 ( $\pm$ 0.14)			2.88		
	Late	1.43 ( $\pm$ 0.44)	0.25 ( $\pm$ 0.20)			3.27		
Size	Early	0.61 ( $\pm$ 0.23)	0.08 ( $\pm$ 0.15)	0.31 ( $\pm$ 0.02)	0.79	6.43	8.4	100
	Mid	1.04 ( $\pm$ 0.25)	0.67 ( $\pm$ 0.19)			6.69		
	Late	1.70 ( $\pm$ 0.48)	0.04 ( $\pm$ 0.15)			6.48		
Maturity	Early	0.26 ( $\pm$ 0.13)	0	0.39 ( $\pm$ 0.03)	0.85	8.04	9.8	100
	Mid	1.83 ( $\pm$ 0.33)	0.30 ( $\pm$ 0.13)			6.76		
	Late	4.17 ( $\pm$ 1.09)	0.57 ( $\pm$ 0.28)			4.69		
Underwater weight	Early	479 ( $\pm$ 179)	232 ( $\pm$ 123)	505 ( $\pm$ 38)	0.76	324	6.5	100
	Mid	1207 ( $\pm$ 224)	431 ( $\pm$ 107)			341		
	Late	2716 ( $\pm$ 786)	949 ( $\pm$ 320)			377		
Cooking type	Early	0.51 ( $\pm$ 0.20)	0.02 ( $\pm$ 0.13)	0.42 ( $\pm$ 0.03)	0.85	1.90	30.3	100
	Mid	0.68 ( $\pm$ 0.14)	0.05 ( $\pm$ 0.09)			2.11		
	Late	1.60 ( $\pm$ 0.49)	0.23 ( $\pm$ 0.20)			2.48		
ACD_1h	Early	0.26 ( $\pm$ 0.13)	0	0.22 ( $\pm$ 0.02)	0.88	6.90	7.0	100
	Mid	0.75 ( $\pm$ 0.15)	0.06 ( $\pm$ 0.09)			6.63		
	Late	1.19 ( $\pm$ 0.37)	0.13 ( $\pm$ 0.17)			6.22		
ACD_24h	Early	0.17 ( $\pm$ 0.11)	0	0.85 ( $\pm$ 0.06)	0.69	6.39	16.2	100
	Mid	0.90 ( $\pm$ 0.19)	0.16 ( $\pm$ 0.10)			5.72		
	Late	0.67 ( $\pm$ 0.23)	0			5.09		
ABD_Nov8c	Early	0.98 ( $\pm$ 0.33)	0.23 ( $\pm$ 0.18)	0.86 ( $\pm$ 0.06)	0.68	6.18	14.8	100
	Mid	1.02 ( $\pm$ 0.19)	0			6.20		
	Late	0.10 ( $\pm$ 0.20)	0.34 ( $\pm$ 0.24)			6.43		
ABD_24hNov8c	Early	0.89 ( $\pm$ 0.36)	0.55 ( $\pm$ 0.26)	0.88 ( $\pm$ 0.07)	0.58	5.93	16.2	100
	Mid	1.11 ( $\pm$ 0.22)	0.27 ( $\pm$ 0.12)			5.77		
	Late	0	0.84 ( $\pm$ 0.25)			5.70		
ABD_Feb8c	Early	0.90 ( $\pm$ 0.40)	0.26 ( $\pm$ 0.29)	0.41 ( $\pm$ 0.04)	0.77	6.25	9.9	80
	Mid	0.76 ( $\pm$ 0.20)	0.35 ( $\pm$ 0.15)			6.51		
	Late	0.69 ( $\pm$ 0.26)	0.08 ( $\pm$ 0.17)			6.56		
ABD_Apr8c	Early	1.04 ( $\pm$ 0.42)	0	0.38 ( $\pm$ 0.04)	0.71	6.45	9.3	69
	Mid	0.71 ( $\pm$ 0.21)	0.37 ( $\pm$ 0.18)			6.62		
	Late	0.35 ( $\pm$ 0.29)	0.54 ( $\pm$ 0.32)			6.65		
ABD_Apr4c	Early	0.36 ( $\pm$ 0.28)	0	0.49 ( $\pm$ 0.06)	0.53	6.51	10.5	51
	Mid	0.29 ( $\pm$ 0.14)	0.09 ( $\pm$ 0.15)			6.64		
	Late	0	0.57 ( $\pm$ 0.26)			6.70		
CookColY	Early	1.32 ( $\pm$ 0.41)	0.11 ( $\pm$ 0.16)	0.16 ( $\pm$ 0.01)	0.95	6.88	6.0	86
	Mid	1.47 ( $\pm$ 0.29)	0.09 ( $\pm$ 0.10)			6.77		
	Late	1.23 ( $\pm$ 0.41)	0			6.53		
FleshColY	Early	1.03 ( $\pm$ 0.33)	0.07 ( $\pm$ 0.15)	0.15 ( $\pm$ 0.01)	0.92	6.51	5.9	77
	Mid	0.99 ( $\pm$ 0.22)	0.16 ( $\pm$ 0.12)			6.61		
	Late	0.83 ( $\pm$ 0.35)	0			6.49		
FryingCol_Nov8c	Early	1.07 ( $\pm$ 0.39)	0.48 ( $\pm$ 0.25)	0.67 ( $\pm$ 0.05)	0.75	5.97	13.6	100
	Mid	1.08 ( $\pm$ 0.21)	0.12 ( $\pm$ 0.10)			5.98		
	Late	0.35 ( $\pm$ 0.22)	0.22 ( $\pm$ 0.20)			6.09		

**Table 3** (continued)

Traits	Maturity	$V_g$ ( $\pm$ SE)	$V_{gl}$ ( $\pm$ SE)	Error ( $\pm$ SE)	$h^2$	Predicted means	%CV	%present
FryingCol_Feb8c	Early	2.84 ( $\pm$ 0.74)	0.20 ( $\pm$ 0.18)	0.38 ( $\pm$ 0.03)	0.88	5.23	10.4	99
	Mid	2.79 ( $\pm$ 0.46)	0.24 ( $\pm$ 0.12)			5.99		
	Late	1.04 ( $\pm$ 0.46)	0.82 ( $\pm$ 0.35)			6.58		
FryingCol_Apr8c	Early	2.22 ( $\pm$ 0.89)	1.00 ( $\pm$ 0.55)	0.34 ( $\pm$ 0.03)	0.80	5.63	9.7	79
	Mid	1.36 ( $\pm$ 0.30)	0.40 ( $\pm$ 0.17)			6.03		
	Late	0.98 ( $\pm$ 0.45)	0.61 ( $\pm$ 0.33)			6.39		
FryingCol_Apr4c	Early	0.91 ( $\pm$ 0.34)	0	0.89 ( $\pm$ 0.07)	0.74	4.77	17.7	79
	Mid	0.82 ( $\pm$ 0.18)	0			5.42		
	Late	0.59 ( $\pm$ 0.33)	0.56 ( $\pm$ 0.29)			5.79		
EnzBrow_30min	Early	0.80 ( $\pm$ 0.33)	0.40 ( $\pm$ 0.23)	0.39 ( $\pm$ 0.03)	0.80	6.57	9.8	100
	Mid	0.77 ( $\pm$ 0.15)	0			6.30		
	Late	0.32 ( $\pm$ 0.19)	0.21 ( $\pm$ 0.19)			6.10		

$V_g$  represents the genetic variance component per maturity class and  $V_{gl}$  contains the estimates for the variance component for genotype-by-location interaction. %CV tabulates the coefficient of variation and %present is indicative of the percentage of genotypes for which data were available

represent positively correlated traits, obtuse angles represent negative correlations and orthogonal directions represent independent traits. Distance from the origin indicates how well individual traits are represented in the biplot projection plane. For traits close to the origin, one should not try to reconstruct the original correlation from the biplot.

In Fig. 3b, which we consider first because genetic correlations are visually better observed, FryingCol\_Apr8c, FryingCol\_Nov8c, FryingCol\_Feb8c and FryingCol\_Apr4c appeared highly correlated. The same observation holds for ABD\_Nov8c, ABD\_Feb8c, ABD\_Apr8c and ABD\_Apr4c. These formed expected clusters of related traits because they only differed in their time point of evaluation or in their storage temperature before evaluation. The protocols to assess the traits were exactly the same. The relationships between cooking colour and flesh colour and between cooking type and underwater weight were equally predictable. The first two measured principally the same phenotype, whereas the latter two were correlated because cooking type is influenced by dry matter concentration: higher dry matter concentration usually results in a looser cooking type (van Dijk et al. 2002). The negative correlation between maturity and underwater weight (and indirectly cooking type) was due to the development of a higher dry matter concentration by later maturing genotypes.

In Fig. 3a, the same cluster of correlated frying colour traits existed, although the correlations were less pronounced, i.e. the angles were less sharp. The cluster of after baking darkening traits became less obvious as well, only ABD\_Apr8c and ABD\_Feb8c appeared highly correlated, whereas ABD\_Nov8c and ABD\_Apr4c shifted away from the cluster. The relationship between ACD\_1h and ACD\_24h became more apparent in this single-year data set whereas the relationships between cooking colour and flesh colour and between cooking type and underwater weight became less obvious. The negative correlation between maturity and underwater weight also existed in this data set.

**Table 4** Overview of the estimates of variance components and their respective standard errors for the industrial data set

Traits	$V_g$ ( $\pm$ SE)	$V_{gy}$ ( $\pm$ SE)	$V_{gy}$ ( $\pm$ SE)	Error ( $\pm$ SE)	$h^2$	$h^2_{adj}$	Predicted mean ( $\pm$ SE)	%CV	No. of years	No. of locations
Shape	2.55 ( $\pm$ 0.22)	0.07 ( $\pm$ 0.01)	0.26 ( $\pm$ 0.02)	0.57 ( $\pm$ 0.01)	0.74	0.90	5.15 ( $\pm$ 0.02)	14.7	18	66
Size	0.23 ( $\pm$ 0.02)	0.01 ( $\pm$ 0.00)	0.04 ( $\pm$ 0.00)	0.16 ( $\pm$ 0.00)	0.52	0.79	7.76 ( $\pm$ 0.01)	5.2	8	64
Maturity	1.41 ( $\pm$ 0.13)	0.02 ( $\pm$ 0.00)	0.09 ( $\pm$ 0.01)	0.14 ( $\pm$ 0.00)	0.85	0.95	6.38 ( $\pm$ 0.02)	5.8	15	21
Underwater weight	2837 ( $\pm$ 245)	78 ( $\pm$ 8)	132 ( $\pm$ 8)	209 ( $\pm$ 6)	0.87	0.96	406 ( $\pm$ 1)	3.6	18	58
Cooking Type	0.99 ( $\pm$ 0.10)	0.03 ( $\pm$ 0.01)	0.11 ( $\pm$ 0.01)	0.65 ( $\pm$ 0.02)	0.56	0.82	2.99 ( $\pm$ 0.02)	27.0	8	33
ACD_1h	0.35 ( $\pm$ 0.04)	0.03 ( $\pm$ 0.01)	0.06 ( $\pm$ 0.01)	0.23 ( $\pm$ 0.01)	0.52	0.80	6.67 ( $\pm$ 0.01)	7.2	8	33
ACD_24h	0.43 ( $\pm$ 0.08)	0.05 ( $\pm$ 0.02)	0.05 ( $\pm$ 0.02)	0.31 ( $\pm$ 0.04)	0.51	0.81	6.55 ( $\pm$ 0.02)	8.5	5	26
ABD_Nov8c	0.29 ( $\pm$ 0.04)	0.03 ( $\pm$ 0.01)	0.06 ( $\pm$ 0.01)	0.24 ( $\pm$ 0.01)	0.48	0.77	6.88 ( $\pm$ 0.02)	7.1	8	9
ABD_Feb8c	0.13 ( $\pm$ 0.02)	0.01 ( $\pm$ 0.01)	0.03 ( $\pm$ 0.00)	0.23 ( $\pm$ 0.01)	0.33	0.64	6.84 ( $\pm$ 0.02)	7.0	8	5
ABD_Apr8c	0.16 ( $\pm$ 0.03)	0	0.07 ( $\pm$ 0.02)	0.19 ( $\pm$ 0.02)	0.38	0.66	6.82 ( $\pm$ 0.03)	6.4	3	3
ABD_Apr4c	0.31 ( $\pm$ 0.07)	0.16 ( $\pm$ 0.04)	0.01 ( $\pm$ 0.03)	0.22 ( $\pm$ 0.03)	0.45	0.84	6.54 ( $\pm$ 0.04)	7.1	7	3
ABD_24hNov8c	0.96 ( $\pm$ 0.23)	0.04 ( $\pm$ 0.06)	0.01 ( $\pm$ 0.06)	0.42 ( $\pm$ 0.08)	0.68	0.90	6.99 ( $\pm$ 0.08)	9.2	2	14
CookColY	0.26 ( $\pm$ 0.03)	0.01 ( $\pm$ 0.00)	0.01 ( $\pm$ 0.00)	0.13 ( $\pm$ 0.00)	0.65	0.88	6.57 ( $\pm$ 0.01)	5.5	8	5
FleshColY	0.29 ( $\pm$ 0.03)	0.01 ( $\pm$ 0.00)	0.01 ( $\pm$ 0.00)	0.15 ( $\pm$ 0.00)	0.63	0.87	6.52 ( $\pm$ 0.01)	5.9	18	50
FryingCol_Nov8c	0.92 ( $\pm$ 0.11)	0.03 ( $\pm$ 0.01)	0.10 ( $\pm$ 0.01)	0.22 ( $\pm$ 0.01)	0.72	0.90	6.59 ( $\pm$ 0.02)	7.2	8	25
FryingCol_Feb8c	1.23 ( $\pm$ 0.15)	0.04 ( $\pm$ 0.01)	0.08 ( $\pm$ 0.01)	0.22 ( $\pm$ 0.01)	0.78	0.93	6.13 ( $\pm$ 0.03)	7.7	8	8
FryingCol_Apr8c	1.07 ( $\pm$ 0.15)	0.05 ( $\pm$ 0.02)	0.11 ( $\pm$ 0.03)	0.31 ( $\pm$ 0.02)	0.70	0.89	5.97 ( $\pm$ 0.04)	9.3	5	5
FryingCol_Apr4c	1.13 ( $\pm$ 0.18)	0.00 ( $\pm$ 0.02)	0.21 ( $\pm$ 0.04)	0.32 ( $\pm$ 0.03)	0.68	0.86	5.09 ( $\pm$ 0.04)	11.0	7	6
EnzBrow_30min <sup>a</sup>	0.59 ( $\pm$ 0.11)			0.43 ( $\pm$ 0.05)	0.58	0.85	6.59 ( $\pm$ 0.07)	9.9	3	1

$V_g$  represents the genetic variance component,  $V_{gy}$  the variance component for genotype-by-year interaction and  $V_{gry}$  the variance component for genotype-by-trial-within-year interaction.  $h^2$  is the actual heritability for this data set and  $h^2_{adj}$  is the adjusted heritability calculated on a comparable scale to allow a fair comparison with the academic data set. %CV presents the coefficient of variation, no. of years the number of years in which the trait was tested and no. of locations the number of locations where the trait was evaluated

<sup>a</sup> For enzymatic browning, a more simple model could be used for analysis due to the trait being analysed on one location only

**Fig. 1** Comparison between heritability estimates of the academic and industrial panels for 19 phenotypic traits. For the industrial panel, the adjusted heritabilities on a comparable scale were used

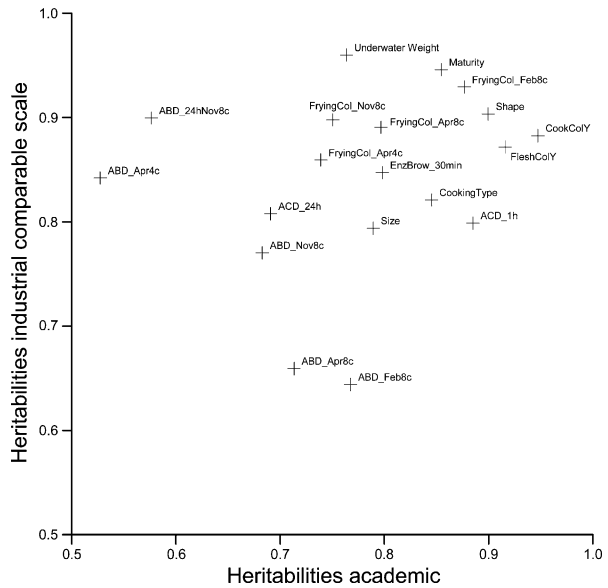
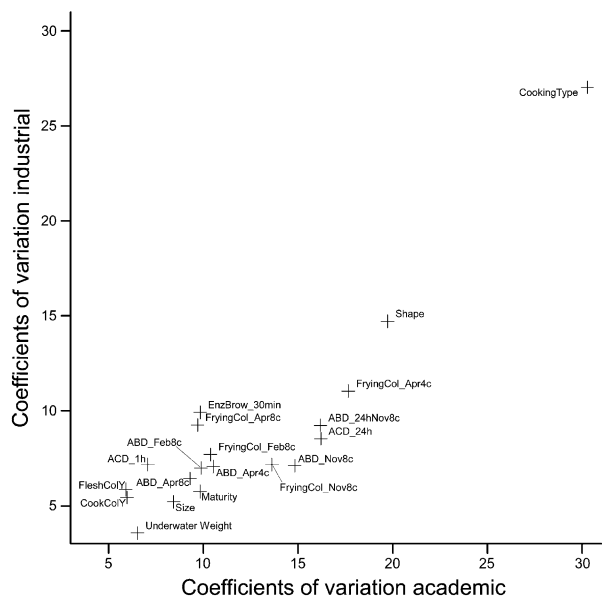
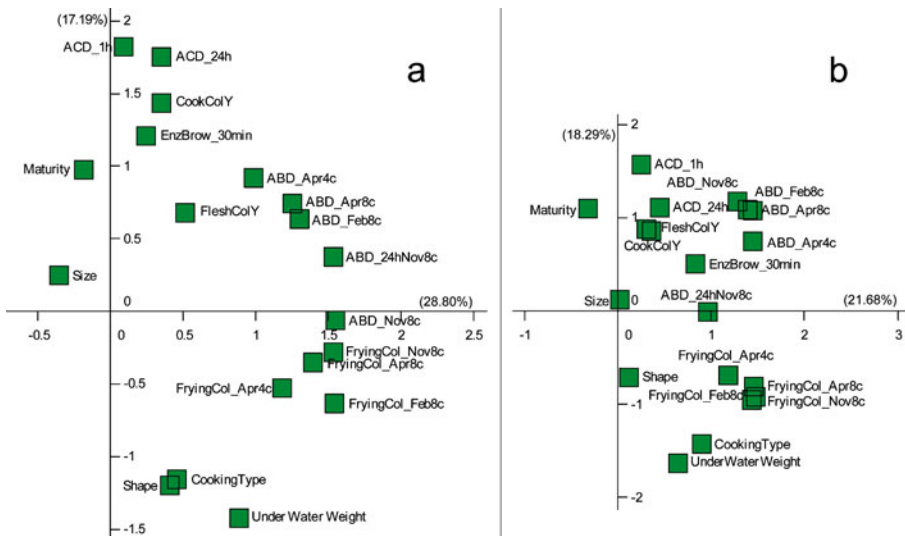


Figure 4 shows scatter plots of the estimated genetic effects in both panels, illustrates for individual traits the relationships between the two data sets and also contains information about the magnitude of variances for specific traits present in both data sets. When the points are concentrated along the diagonal, this is an indication of a similar behaviour of genotypes in both data sets and high correlation. By comparing the range of the genotypic main effects, it can be deduced whether both data sets hold the same amount of information, i.e. variance, for a particular

**Fig. 2** Comparison of the coefficients of variation of 19 agro-morphological and quality traits between the academic and industrial panels





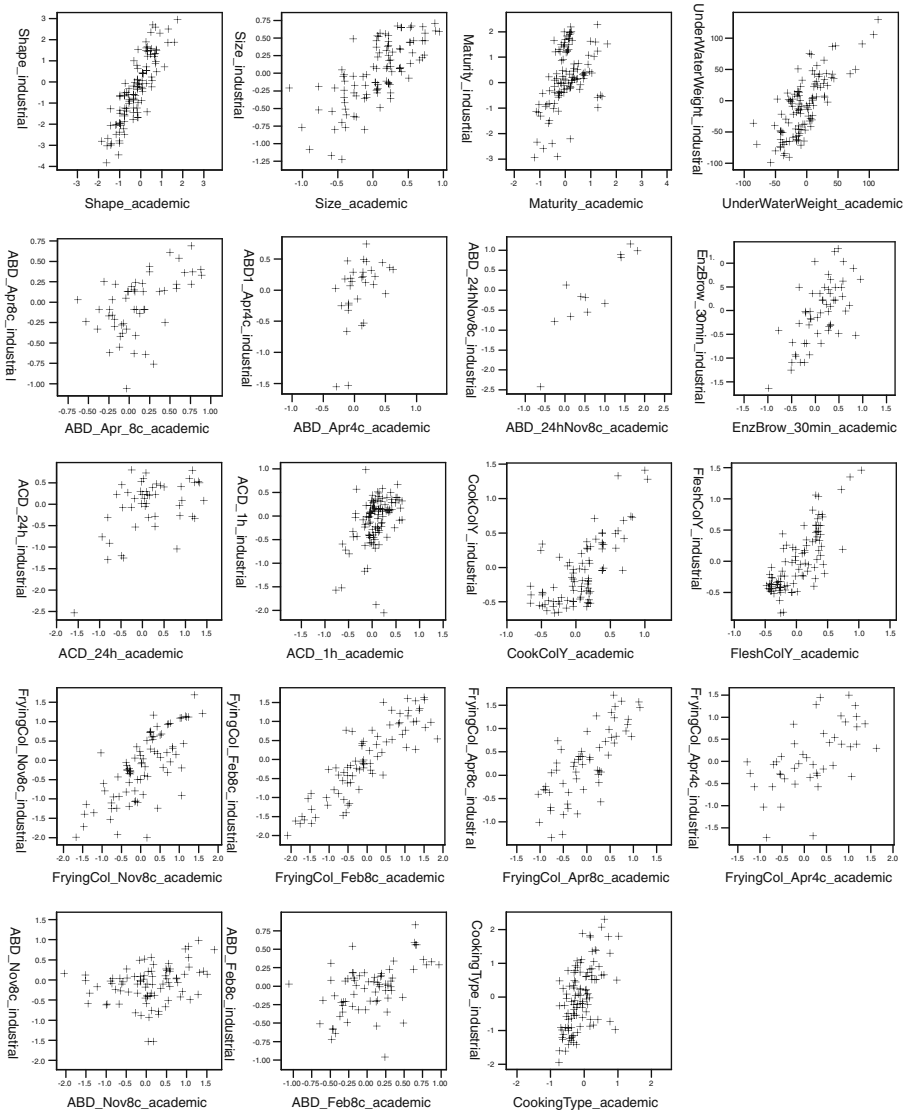
**Fig. 3** Biplot showing genetic correlations. Clustered traits (*squares*), sharing approximately the same direction (a sharp angle), represent highly correlated traits. *Squares* at opposite side of the origin (obtuse angles) represent negatively correlated traits. *Squares* with orthogonal directions represent independent traits. **a** Trait correlations resulting from the academic panel. **b** Trait correlations deduced from the industrial panel

trait. For example, tuber shape, tuber size, underwater weight, maturity, cooking type, ACD\_1h, flesh colour, cooking colour, enzymatic browning and the frying colour traits all have their points more or less along the diagonal, an indication that these traits correlate well. Tuber size, underwater weight and the frying colour traits all have comparable ranges for genotypic main effects in both data sets, indicating that equal variance is present in both data sets for these traits. Tuber shape, maturity, cooking type, ACD\_1h, flesh colour, cooking colour and EnzBrow\_30min have a slightly higher range for their genotypic main effects in the industrial data set, which suggests that the industrial panel has a slightly higher information content for these traits. These abovementioned observations illustrate what could be interpreted from the correlations presented in Table 5 as well, namely that, at least for the traits with a high correlation, both data sets are equally valuable.

**Discussion**

Data Management is Necessary

A first complication while assembling data from different companies, as we did for the industrial data set, was to match traits with each other. Trait names can be different while the same phenotype is measured, or the trait names can be the same while a different phenotype is measured. The scales that are used can be different in order, in increments or in range. In addition, the protocols to measure the same phenotype can differ. These problems were solved through consulting the trait evaluators to obtain detailed descriptions allowing rescaling of the trait values. Inspection of unprocessed trait data also resulted in the detection of systematic differences between companies



**Fig. 4** Comparison of genotypic effects for 19 traits between the academic and industrial panels. The scatter plots display the variability of the traits across genotypes as observed in the academic and industrial data sets, as well as the pair-wise correlation of the traits between both panels

for trait values of reference cultivars. Upon discussions and adaptation of scales to score traits, we could harmonize the data obtained from five different breeding companies. Table 1 offers a clear trait ontology, analogous to the plant structure ontology for anatomy and morphology of flowering plants (Ilic et al. 2007), with accurate definitions of the trait, the evaluation protocol and the scaling of the trait values. A continued effort towards a comprehensive trait ontology could be very helpful in the case of potato. Examples of trait ontologies exist in cereal genomics (Jaiswal et al. 2002) and in tomato (Brewer et al. 2006). Also, to increase robustness and power to detect QTL, it is advisable to develop or apply objective measurement



**Table 5** Correlations between genotypic effects of the academic and industrial panels

	ABD_Nov8c_academic	ABD_Apr8c_academic	ABD_Apr4c_academic	ABD_Feb8c_academic	ABD_24hNov8c_academic	ACD_1h_academic
ABD_Nov8c_industrial	0.20	0.32	0.34	0.37	0.29	0.40
ABD_Apr8c_industrial	0.31	0.31	0.20	0.31	0.41	0.32
ABD_Apr4c_industrial	0.23	0.22	0.29	0.23	0.31	0.18
ABD_Feb8c_industrial	0.17	0.28	0.31	0.35	0.27	0.27
ABD_24hNov8c_industrial	0.24	0.16	0.20	0.21	0.25	0.13
ACD_1h_industrial		0.21	0.18	0.23		0.39
ACD_24h_industrial		0.18	0.23	0.22	0.16	0.23
CookCoLY_industrial	0.19	0.12	0.26	0.17	0.26	
CookingType_industrial	0.39	0.23		0.29	0.36	
EnzBrow_30min_industrial	0.20		0.22	0.22	0.20	
FleshCoLY_industrial	0.19	0.28	0.20	0.29	0.27	
FryingCol_Nov8c_industrial	0.59	0.42	0.34	0.41	0.55	
FryingCol_Feb8c_industrial	0.65	0.40	0.28	0.43	0.61	
FryingCol_Apr8c_industrial	0.59	0.35	0.33	0.36	0.54	
FryingCol_Apr4c_industrial	0.34	0.27	0.33	0.31	0.30	
Maturity_industrial						
Shape_industrial	0.15				0.10	
Size_industrial						
UnderWaterWeight_industrial	0.39	0.18	0.12	0.24	0.31	
ACD_24h_academic						
CookCoLY_academic						
CookingType_academic						
EnzBrow_30min_academic						
FleshCoLY_academic						
FryingCol_Nov8c_academic						
FryingCol_Feb8c_academic						
0.37		0.12			0.16	0.15
0.41	0.12				0.32	0.17
0.21		0.15			0.22	0.11
0.39		0.13			0.16	0.11
0.20			0.10		0.24	0.21
0.41	0.14					
0.35	0.20					
0.11	0.17			0.71	0.22	0.47
	0.57				0.38	0.11
0.10		0.48			0.14	

**Table 5** (continued)

ACD_24h_academic	CookCo1Y_academic	CookingType_academic	EnzBrow_30min_academic	FleshCo1Y_academic	FryingCol_Nov8c_academic	FryingCol_Feb8c_academic
0.17	0.69		0.17	0.75	0.18	
0.11		0.18		0.13	0.59	0.70
0.12		0.19		0.13	0.65	0.73
0.11		0.14		0.13	0.57	0.62
0.15					0.36	0.40
		0.25			0.15	0.25
		0.51	0.14		0.40	0.49

FryingCol_Apr8c_academic	FryingCol_Apr4c_academic	Maturity_academic	Shape_academic	Size_academic
0.18				0.21
0.24	0.16			
0.14	0.14		0.14	
0.29	0.17	0.13		0.15
0.31	0.11		0.28	
	0.31			
0.13	0.12			
0.63	0.51			
0.57	0.49		0.10	
0.47	0.38			
0.38	0.38			
0.12	0.24	0.44	0.86	0.70
0.42	0.46		0.33	

On the diagonal, the correlations between the two assessments for the same trait

tools. Such objective measurements could replace more subjective trait evaluations based on the breeder's eye.

Patterns of variation and correlations in two data sets, one resulting from a designed field trial (the academic panel) and one containing data from multi-year–multi-location clonal selection trials (the industrial panel), have been compared. Heritabilities, coefficients of variation, trait relationships within and between data sets and phenotypic variance attributable to genotypes have been estimated. In Tables 3 and 4, an overview of the analyses of both data sets is presented. Especially the columns with the genotypic variance component and heritability are interesting because they give an impression of the feasibility of genetic improvement and possibilities for QTL detection. Figure 1 shows that the industrial heritabilities are generally higher than the ones from the academic panel. Figure 2 demonstrates that traits in both panels behaved consistently when their coefficients of variation are considered, indicating that both data sets contain data of comparable quality. Trait relationships within each panel are depicted through biplots in Fig. 3, while correlations between traits in both data sets are listed in Table 5. Basically, the same trait relationships existed within the two data sets, but the strength of the relationships varied between them. Agromorphological traits tend to have higher pair-wise correlations between the two data sets than quality traits, except for frying colour. This trend is visually expressed in Fig. 4, where the individual pair-wise trait correlations are depicted in scatter plots.

### The Importance of Maturity

As mentioned in the “[Materials and Methods](#)” section, we also tested a simplified model without maturity class-related terms for the analysis of the academic panel, to quantify the influence of maturity class on the resulting estimates. Based on the results of this simplified model (see [Online Resource 2 to 8](#)), it can be deduced that the presence or absence of maturity terms in the analysis model did not alter the observed trends for the majority of traits when comparing the academic and the industrial panels. For example, removal of maturity terms did not influence the coefficients of variation ([Online Resource 5](#)), nor the within data set genetic correlations ([Online Resource 6](#)). However, for maturity and underwater weight, there were some changes apparent, as anticipated. [Online Resource 3 and 4](#) show the effect of removal of maturity terms on heritabilities. The effect seemed rather unpredictable for most traits, except for maturity and underwater weight where higher heritabilities were expected after removal of maturity terms because in that case the between-maturity-class trait variation was added to the within-maturity-class trait variation in the estimation of the genetic variance component. The range of the genetic effects for the trait “maturity” obtained with the simplified model did become perfectly similar with the range in the industrial panel ([Online Resource 7](#)). This shift was also visible in the improved correlation for maturity between genotypic main effects for the industrial panel on the one hand and the academic panel on the other hand: from  $r=0.44$  to  $r=0.87$  where the latter was obtained with the simplified model ([Table 5](#) and [Online Resource 8](#)).

## High Heritabilities Were Obtained

Heritabilities obtained were high, which is not abnormal for highly variable traits of clonally propagated crops. The high heritabilities (both for the academic and industrial panels) for tuber shape and maturity confirm that these traits are very stable across locations and years. The high heritability for underwater weight can be interpreted as the absence of strong genotype-by-environment ( $G \times E$ ) interaction. The same interpretation holds for frying colour because the  $G \times E$  variance component for this set of traits was low in comparison with the genetic variance component, as illustrated in Tables 3 and 4 and Online Resource 2. Notwithstanding the truncated scales of flesh colour and cooking colour, now reflecting yellow colour intensity only, the heritabilities of these traits were still high. Tuber size in the industrial data set had a low genetic variance component, which was reflected in the lower heritability for this trait and was probably the result of variation introduced by the size and physiological age of the seed tuber, as these factors are known to have a profound effect on size grading variability within a genotype. The genetic variance components for discolouration phenotypes (ABD and ACD) and cooking type were lower when compared to their respective error variance components for both panels, which resulted in lower heritabilities for discolouration traits and cooking type. This could be because the trait values describe composite effects, i.e. it is difficult to observe pure phenotypes. ABD observations, for instance, are influenced by the dark frying colours, caused by the Maillard discolouration of French fries due to reducing sugars, and cooking type scores are largely dependent on starch concentration.

The higher heritabilities for the academic compared to the industrial data set (here we consider the true heritability, not the comparably scaled “adjusted” heritability) can be explained in two ways. Firstly, the industrial data set included more environments. The more environments for which trait data are available, the lower heritability estimates tend to be, due to a larger contribution of  $G \times E$ . Secondly, the choice for diversity among the cultivars in the academic panel will have inflated the heritabilities. About one third of the cultivars was first registered on a national list in Europe before 1930, and in those days for most traits, more deviating values were accepted that are not observed anymore in contemporary cultivars.

Phenotypic values of traits with high heritabilities were expected to correlate well between the academic and industrial data sets. This was confirmed in Table 5 where the correlations between genotypic main effects range from 0.20 for ABD\_Nov8c to 0.86 for tuber shape. Indeed, tuber shape, underwater weight, flesh colour and cooking colour did not only have higher heritability estimates, but were also more highly correlated between the two data sets than quality traits like after baking darkening and after cooking darkening. Although the data structure of both panels was quite different, most of the traits correlated reasonably well. As an important consequence, existing data (industrial panel) may serve equally well for association studies as a designed trial (academic panel) for particular traits.

## Expected Genetic Correlations Appeared

Figure 3 and Online Resource 6 show genetic correlations between traits within data sets. The frying colour and after baking darkening trait clusters detected in both

panels were expected as these traits measure exactly the same phenotypes. They only differed in the temperature at which the tubers have been stored and the duration of storage prior to evaluation. In theory, there are no genes involved in the discolouration process itself: frying colour is the result of the Maillard reaction, whereas after baking darkening is the result of oxidation of an iron–chlorogenic acid complex (McKenzie et al. 2005). However, genetic background does influence the amount of reaction products available for the discolouration processes. Olsson et al. (2004) provided evidence that storage time and temperature do not influence the ranking of genotypes with respect to frying colour considerably, which could explain the strong relationships between the frying colour traits (Fig. 3 and Online Resource 6). Likewise, Wang-Pruski and Nowak (2004) state that for after baking darkening, storage time and temperature only influence the severity of the discolouration; nothing is mentioned about changes in genotype ranking. The positive correlation between cooking type and underwater weight in both data sets is equally predictable. Underwater weight is a measure for dry matter concentration of potato tubers and indirectly also for starch concentration; the higher the underwater weight, the higher the starch concentration of a cultivar. Starch granules create swelling pressure while being cooked and therefore can contribute to disintegration of the cooked tuber structure, which creates a looser cooking type (van Marle et al. 1997). The relationship between flesh colour and cooking colour is even more predictable: the mere difference between both phenotypes is that flesh colour is evaluated on raw tubers and cooking colour on cooked tubers. The negative correlation between maturity and underwater weight can easily be explained because later maturing cultivars reside longer in the field and thus have the opportunity to assemble more sucrose as their photosynthesis remains active for a longer time period and, so, can stack larger amounts of starch in their tubers which increases their underwater weight. The unexpected relationship between ACD\_24h, cooking colour and flesh colour (visible in Fig. 3b) is possibly an artefact due to the difficulty for breeders to score after cooking darkening objectively after a 24-h period in naturally darker coloured potato genotypes.

#### Captured Genetic Variation Comparable with Literature

To get an impression of the genetic variation in our data sets, we compared obtained heritabilities with previously reported heritabilities, where the remark has to be made that in our study, a ratio of variance components was calculated that strictly speaking is not really a measure of heritability, but rather repeatability. Where possible, we clearly indicated how heritabilities were estimated in previous studies. However, in some cases, this was not well described. Maris (1969) reports heritabilities pooled over 12 progenies of potato. For maturity and underwater weight, he obtained values of 0.85 and 0.76 respectively, similar to our estimates (see also Tables 3 and 4 and Online Resource 2). Tai and Young (1984) reported a somewhat lower heritability of 0.54 for maturity and 0.65 for specific gravity, a trait directly related to underwater weight. Yildirim and Çalışkan (1985) obtained 0.77 as a heritability estimate on a clone mean basis for tuber shape evaluated in a field trial performed over 2 years and three locations, likewise comparable to what we computed. Love et al. (1997) reported heritabilities on a clone mean basis of 27 traits obtained from a study

involving three clonal generations with 200 clones. For shape, they calculated a heritability of 0.79; we obtained 0.74 for the industrial data set and 0.90 for the academic data set (Tables 3 and 4). Thompson et al. (1983) estimated tuber size heritability to be 0.79 for a *Solanum andigena* population propagated by true potato seed. We obtained a lower value: 0.52 within the industrial data set, but the same value was obtained with the academic panel. Neele et al. (1991) estimated heritabilities for maturity, shape and underwater weight to be 0.77, 0.61 and 0.68, respectively. Bradshaw et al. (2008) reported heritabilities on a clone mean basis for maturity of 0.92, dry matter concentration of 0.81, tuber size of 0.87, tuber shape of 0.85, after cooking darkening of 0.60, frying colour at 4 °C 0.79 and frying colour at 10 °C of 0.78. These estimates are slightly higher than or comparable to the values we obtained (see Tables 3 and 4 and Online Resource 2).

### Field Trial Design

Multi-year–multi-location (MYML) data sets will always be preferred over single-year–few-location data sets because they allow investigation of  $G \times E$ . Often dilemmas regarding field trial design have to be solved before the start of new projects: one trial year with several locations and replicates versus more years with several locations and less replicates. The results of our phenotypic analyses show that the heritabilities obtained with the two types of data sets (Tables 3 and 4) were different. However, when trait correlations within or between data sets are concerned, both data sets showed similar patterns. Both MYML and single-year–few-location approaches have their advantages and disadvantages. When choosing just 1 year of field trials, heritabilities are likely overestimated due to confounding of the genotypic main effect variance and the genotype-by-year interaction variance (and genotype-by-year-by-location interaction variance). MYML trials guarantee robust phenotypic observations, heritabilities will be more conservative and probably more realistic and more variable (environment-sensitive) traits will become better assessable. However, with MYML data, some aspects can change in the course of time, for instance reference cultivars can change as new standards are set by breeding companies or trait scoring methodology can alter due to new techniques or insights. Phenotypic analyses will become increasingly complex as more trials need to be included in the analysis and unbalancedness of data sets will increase.

Plot size is another important factor in field trial design. Some researchers claim that small plot size, e.g. four plants per genotype, can suffice to evaluate most of the traits (personal communication), whereas potato breeding companies generally use a plot size of a minimum of 16 plants per genotype to get a reliable impression of the performance of a genotype. Considering the phenotypic analysis of the academic field trial—given acceptable environmental conditions and good experimental design—for most simple traits such as maturity, tuber shape or flesh colour, small plot sizes indeed suffice to obtain good quality phenotypic data. However, when more variable traits like discolouration have to be assessed, a larger plot size can significantly improve the outcome of a field experiment by providing more sample material and offering a buffer to extreme within-plot environmental conditions.

## Cultivar Performance Is Better Illustrated Using Adjusted Genotypic Means

Marketing brochures and national lists usually provide information regarding a cultivar's performance to allow growers to make well-informed decisions. However, in order to be really informative about a cultivar's performance for specific traits, it is necessary to present reliable genotypic means. Therefore, inclusion of adjusted trait mean values, corrected for genotype-by-environment interaction is advisable, in contrast to straightforwardly averaged trait values. In our case, the genotypic effects which are directly related to these adjusted means by means of the phenotypic analysis model offered robust data that could be reliably compared throughout data sets resulting from different field trials (Fig. 4).

### Conclusion

With two differently structured association panels, high heritabilities have been obtained, which also appeared comparable to previous studies. Both panels were consistent in their data quality. Anticipated trait correlations were found within both panels and the majority of the traits correlated well between both panels. The industrial panel had a somewhat higher information content for some of the traits. Maturity class separation in potato field trials was shown to have consequences for the estimation of the genotypic main effects of maturity and underwater weight, which will likely carry over to association results since association studies make direct use of the estimated genotypic main effects. Based on the results presented in this manuscript, we can conclude that two equally valuable association panels have been assembled.

The phenotypic analyses in this article represent the first step of a genome-wide two-step association mapping study. We intend to integrate a high amount of genotypic marker information to detect associations for the traits mentioned. The outcome of the current analyses produced adjusted genotypic means that subsequently will serve as entry values for further association analysis, a similar approach as described in D'hoop et al. (2008) and Li et al. (2008).

**Acknowledgements** We thank Mariëlle Muskens, Sjefke Allefs, Nico Kers and stagiairs (Agrico Research), Nick de Vetten (Averis seeds), Jeroen van Soesbergen (Van Rijn-KWS), Jacqueline Verdijck-Lamers, Robert Graveland and Kees Zijlstra (HZPC Holland BV) and Guus Heselmans (C. Meijer) for sharing their phenotypic data and for their support and help in the 2006 field trial. We deeply acknowledge the help of Martin Keijbets and Leo van Marion (VAVI) with scoring frying quality traits. We are grateful to the "2006 field trial" team of the Laboratory of Plant Breeding for their contribution to the academic field trial. This research project was financed by the Centre for BioSystems Genomics (CBSG), which is part of the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Asiedu SK, Astatkie T, Yiridoe EK (2003) The effect of seed-tuber physiological age and cultivar on early potato production. *J Agron Crop Sci* 189:176–184
- Bradshaw J, Hackett C, Pande B, Waugh R, Bryan G (2008) QTL mapping of yield, agronomic and quality traits in tetraploid potato (*Solanum tuberosum* subsp. *tuberosum*). *Theor Appl Genet* 116:193–211
- Brewer MT, Lang L, Fujimura K, Dujmovic N, Gray S, van der Knaap E (2006) Development of a controlled vocabulary and software application to analyze fruit shape variation in tomato and other plant species. *Plant Physiol* 141:15–25
- Brown C, Kim T, Ganga Z, Haynes K, De Jong D, Jahn M, Paran I, De Jong W (2006) Segregation of total carotenoid in high level potato germplasm and its relationship to *beta-carotene hydroxylase* polymorphism. *Am J Potato Res* 83:365–372
- D'hoop B, Paulo M, Mank R, van Eck H, van Eeuwijk F (2008) Association mapping of quality traits in potato (*Solanum tuberosum* L.). *Euphytica* 161:47–60
- Flint-Garcia SA, Thornsberry JM, Buckler ES (2003) Structure of linkage disequilibrium in plants. *Annu Rev Plant Biol* 54:357–374
- Gabriel KR (1971) Biplot graphic display of matrices with application to principal component analysis. *Biometrika* 58:453–467
- Gabriel KR (1978) Biplot as a tool for diagnosing models. *Biometrics* 34:156–157
- Ilic K, Kellogg EA, Jaiswal P, Zapata F, Stevens PF, Vincent LP, Avraham S, Reiser L, Pujar A, Sachs MM, Whitman NT, McCouch SR, Schaeffer ML, Ware DH, Stein LD, Rhee SY (2007) The plant structure ontology, a unified vocabulary of anatomy and morphology of a flowering plant. *Plant Physiol* 143:587–599
- Jaiswal P, Ware D, Ni J, Chang K, Zhao W, Schmidt S, Pan X, Clark K, Teytelman L, Cartinhour S, Stein L, McCouch S (2002) Gramene: development and integration of trait and gene ontologies for rice. *Comp Funct Genom* 3:132–136
- Kirkman MA (2007) Global markets for processed potato products. In: Vreugdenhil D, Bradshaw J, Gebhardt C, Govers F, MacKerron DKL, Taylor MA, Ross HA (eds) *Potato biology and biotechnology: advances and perspectives*. Elsevier, Amsterdam, pp 27–44
- Lærke PE, Christiansen J, Veierskov B (2002) Colour of blackspot bruises in potato tubers during growth and storage compared to their discolouration potential. *Postharvest Biol Technol* 26:99–111
- Li L, Paulo M-J, Strahwald J, Lübeck J, Hofferbert H-R, Tacke E, Junghans H, Wunder J, Draffehn A, van Eeuwijk F, Gebhardt C (2008) Natural DNA variation at candidate loci is associated with potato chip color, tuber starch content, yield and starch yield. *Theor Appl Genet* 116:1167–1181
- Love SL (1999) Founding clones, major contributing ancestors, and exotic progenitors of prominent North American potato cultivars. *Am J Potato Res* 76:263–272
- Love S, Werner B, Pavek J (1997) Selection for individual traits in the early generations of a potato breeding program dedicated to producing cultivars with tubers having long shape and russet skin. *Am J Potato Res* 74:199–213
- Maris B (1969) Studies on maturity, yield, underwater weight and some other characters of potato progenies. *Euphytica* 18:297–319
- McKenzie MJ, Sowokinos JR, Shea IM, Gupta SK, Lindlauf RR, Anderson JAD (2005) Investigations on the role of *acid invertase* and *UDP-glucose pyrophosphorylase* in potato clones with varying resistance to cold-induced sweetening. *Am J Potato Res* 82:231–239
- Neele A, Nab H, Louwes K (1991) Components of visual selection in early clonal generations of a potato breeding programme. *Plant Breed* 106:89–98
- Olsson K, Svensson R, Roslund C-A (2004) Tuber components affecting acrylamide formation and colour in fried potato: variation by variety, year, storage temperature and storage time. *J Sci Food Agric* 84:447–458
- Tai G, Young D (1984) Early generation selection for important agronomic characteristics in a potato breeding population. *Am J Potato Res* 61:419–434
- Thompson P, Mendoza H, Plaisted R (1983) Estimation of genetic parameters for characters related to potato propagation by true seed (TPS) in an *andigena* population. *Am J Potato Res* 60:393–401
- van Dijk C, Fischer M, Holm J, Beekhuizen J-G, Stolle-Smits T, Boeriu C (2002) Texture of cooked potatoes (*Solanum tuberosum*). 1. Relationships between dry matter content, sensory-perceived texture, and near-infrared spectroscopy. *J Agr Food Chem* 50:5082–5088
- van Marle J, de Vries R, Wilkinson E, Yuksel D (1997) Sensory evaluation of the texture of steam-cooked table potatoes. *Potato Res* 40:79–90



- van Os H, Andrzejewski S, Bakker E, Barrena I, Bryan GJ, Caromel B, Ghareeb B, Isidore E, de Jong W, van Koert P, Lefebvre V, Milbourne D, Ritter E, Rouppe van der Voort JNAM, Roussele-Bourgeois F, van Vliet J, Waugh R, Visser RGF, Bakker J, van Eck HJ (2006) Construction of a 10,000-marker ultradense genetic recombination map of potato: providing a framework for accelerated gene isolation and a genomewide physical map. *Genetics* 173:1075–1087
- Wang-Pruski G, Nowak J (2004) Potato after-cooking darkening. *Am J Potato Res* 81:7–16
- Yildirim M, Çalışkan C (1985) Genotype X environment interactions in potato (*Solanum tuberosum* L.). *Am J Potato Res* 62:371–375
- Zhu C, Gore M, Buckler ES, Yu J (2008) Status and prospects of association mapping in plants. *Plant Genome* 1:5–20