CrossMark

ORIGINAL ARTICLE

Variability of the phenolic profiles in the fruits from old, recent and new apple cultivars cultivated in Belgium

Domien De Paepe · Dirk Valkenborg · Bart Noten ·
Kelly Servaes · Ludo Diels · Marc De Loose ·
Bart Van Droogenbroeck · Stefan Voorspoels

Abstract The objective of this study was to explore the utility of the phenolic pattern (1) to differentiate or associate between cultivars with a different or same genetic background respectively and (2) to assess some important quality traits (color, disease resistance, and browning) in apple fruit. A unique set of cultivars/candivars occurring in Belgium were selected for this survey. It was possible to segregate the studied apple cultivars/candivars into three clusters: red-flesh apple candivars, classic/new cultivars and candivars and heritage apple cultivars. It was found that the heritage apple varieties selected for their disease resistance and cultivated under organic agricultural practice demonstrate a remarkable different phenolic profile. Furthermore, it was found that the phenolic makeup of the peel could suggest the genetic relationships among the apple cultivars. In conclusion, the knowledge of the phenolic profile could be a useful instrument in breeding programs to identify apples with better quality traits, processing characteristics and health-benefiting effects.

Keywords Phenolic compounds · Domesticated apple · Heritage cultivars · Genetic diversity · Quality traits

D. De Paepe · D. Valkenborg · B. Noten · K. Servaes ·
L. Diels · S. Voorspoels
Business Unit Separation and Conversion Technology (SCT),
Flemish Institute for Technological Research (VITO),
Boeretang 200, 2400 Mol, Belgium

D. De Paepe (✉) · M. De Loose · B. Van Droogenbroeck
Technology and Food Sciences Unit (T&V), Product Quality and
Innovation (PI), Institute for Agricultural and Fisheries Research
(ILVO), Burgemeester Van Gansberghelaan 115/1,
9820 Merelbeke, Belgium
e-mail: depaepedomien@gmail.com;
domien.depaepe@ilvo.vlaanderen.be

D. Valkenborg
Business Unit Quality of the Environment (MRG), Flemish
Institute for Technological Research (VITO), Boeretang 200,
2400 Mol, Belgium

D. Valkenborg
Interuniversity Institute for Biostatistics and Statistical
Bioinformatics, Hasselt University, Agoralaan 1,
3590 Diepenbeek, Belgium

1 Introduction

The domesticated apple (*Malus domestica* Borkh., family *Rosaceae*, tribe *Pyreae*), one of the world's most ancient and most widely cultivated temperate fruit, may have originated in western Asia from natural hybridization between several species, including the wild Central Asian species *Malus sieversii* and the wild European crabapple *Malus sylvestris* (Cornille et al. 2012). In the past, most important commercial apples were selected by chance seedlings and produced on small scales (Son et al. 2012). This changed from the latter half of the twentieth century by the introduction of intensive apple orchard systems (Lammerts van Bueren and Myers 2011). In this process, only the cultivars that were adapted to the new practices were conserved and the old traditional cultivars (cultivated before 1940) were abandoned (Van Treuren et al. 2010). The current world production of about 70 million metric tons/year is based on only a small number of cultivars: 'Golden Delicious', 'Delicious', 'Cox's Orange Pippin', 'Rome Beauty', 'Granny Smith', 'McIntosh', 'Jonathan', 'Braeburn', 'Fuji', 'Gala' and 'Jonagold' (van Nocker et al. 2012). This unbalance result in a reduction of diversity: of

the 7,098 apple varieties documented as having been cultivated between 1804 and 1904 in the United States, about 86 % has been lost (Richards et al. 2009).

In Europe, a rich heritage of apple fruit genetic resources is still present, either in traditional orchards or in genebank collections. Several programs on conservation and evaluation of indigenous apple tree genetic resources have been conducted (van Treuren et al. 2010). Only a fraction of these collections is used in a limited way in breeding programs, in which often only a narrow genetic pool of standard commercial cultivars is represented (Kellerhals et al. 2012). Such focused selection programs result in a decrease of the apple genetic background diversity (Cornille et al. 2012). However, it has been recognized that conserving the genetic diversity in commercial fruit species reveals a relevant and not substitutable access to genetic resources encoding for fruit quality and disease resistance traits, useful as raw material for breeding (Iacopini et al. 2010).

In Belgium, a wide collection of 1,600 old apple cultivars, mainly from the 18th and 19th centuries, was established progressively at the National Orchard Foundation and Agricultural Research Centre of Gembloux (Kellerhals et al. 2012). Until now, studying the phenotypic variation of this germplasm was restricted to a systematic evaluation of the resistance to most important fungal diseases (scab, canker and mildew), and other characteristics of agronomic interest (fruit quality, yield and storage ability). In this way, the so called 'R.G.F. varieties' were selected, ancient varieties which posses fairly resistant to several apple diseases allowing organic cultivation (Lateur and Populer 1994). With the exception of vitamin C, only little attention has so far been given to the exploration of metabolite diversity (Cuthbertson et al. 2012). However, exploring the metabolite makeup is essential in evaluating the potential beneficial health effects and processing properties of apple fruit. Furthermore, it could provide insight in the susceptibility/tolerance against various apple diseases. In this context, phenolic compounds, a large group of secondary plant metabolites, have drawn attention in the last decades (Kellerhals et al. 2012). Indeed, the presence of specific phenolic compounds could be the underlying factor of low susceptibility against the most important apple diseases, such as scab (black spot), caused by *Venturia inaequalis* (Petkovsek et al. 2010) and powdery mildew (rusty spot) caused by *Podosphaera leucotricha* (Singh et al. 2010). In addition, disease pressure, for example by the lack of pesticides in organic cultivation, can lead to the production and accumulation of an array of defence related phenolic compounds, particularly in the skin which acts as a protective barrier between the plant tissue and its environment (Petkovsek et al. 2010; Zhao and Rajashekar 2006). In this way, the type of agricultural practice (conventional vs.
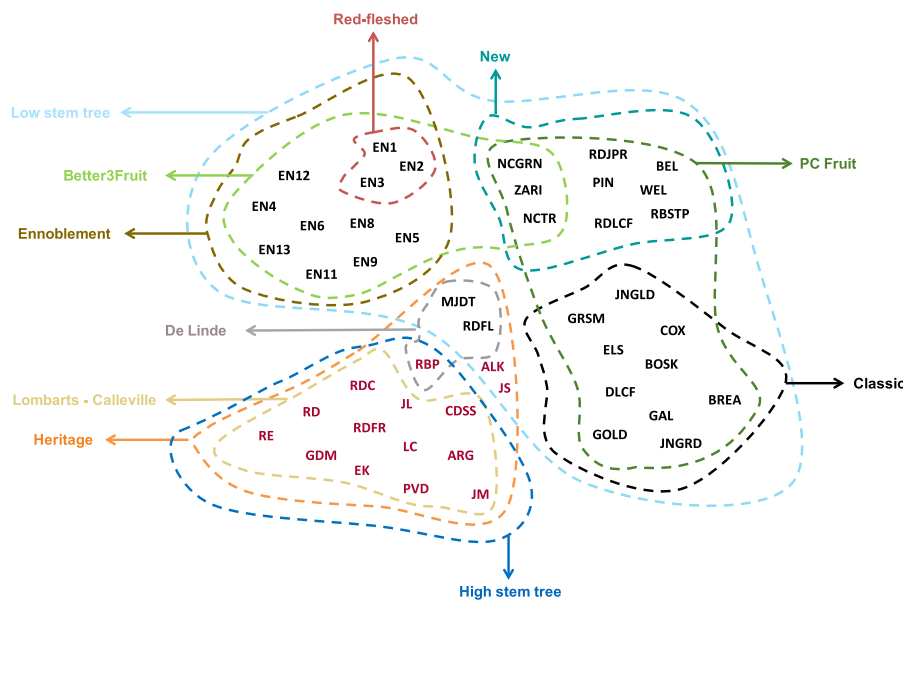
organic) could affect the phenolic content of apple fruit. However, the influence of this factor is still controversial in literature for apple fruit and other species (Vanzo et al. 2013; Petkovsek et al. 2010; Zhao and Rajashekar 2006). Other factors which could affect the phenolic profile of apple fruit are: geographic origin, weather conditions during harvesting season, rootstock, fruit position within the canopy, crop load, fruit maturity stage, fruit part, disease status, and most importantly, genetic variation (Ceymann et al. 2012).

Therefore, the aim of this study was to explore the utility of the phenolic pattern (1) to differentiate or associate between cultivars with a different or same genetic background respectively and (2) to assess some important quality traits (color, disease resistance, and browning) in apple fruit. A representative set of 47 different apple cultivars/candivars occurring in Belgium, comprising heritage cultivars, commercial cultivars and novel selections was studied. A variability study was performed to ensure that the observed variation in phenolic profile could be substantially linked to the variation in genotype (cultivar).

2 Materials and methods

2.1 Plant material

An Euler representation of the selected apple cultivars and candivars is given in Fig. 1. See Supplementary Data 1 for more details. The 'classical' varieties chosen for this survey (besides the international references 'Granny Smith', 'Gala', 'Breaburn'), represent about 70 % of the total production of the dessert apple market in Belgium and are selected because of their relevance to the commercial cultivation. Several recently introduced cultivars like 'Nicoter' (Kanzi®, 2002), 'Nicogreen' (Greenstar®, 2002) and 'Zari' (2006) are selected for reasons of the growing importance in the Belgian and European market and the limited information about their phenolic content. Also some candidate varieties from breeding programs (further called 'candivars') are includes in this survey to evaluate if there are differences in their phenolic profile to what is currently on the market. Heritage cultivars, whose production is gradually eroding in Belgium, are included because their potential to be used in breeding programs. A case example is the recent introduced cultivar 'Belgica', which is a cross between the cultivars 'Vlijtingse kraker' and 'Bellefleur de Hermée'. This 'heritage varieties' includes the so called 'R.G.F. varieties' as well as the 'E.N.R. varieties' which were tested over years for the most important apple diseases (scab, cancer and mildew) by Walloon Agricultural Research Centre (CRA-W, Wallon, Belgium) and Regional Natural Space (Region Nord/

Fig. 1 Euler representation of the properties of the selected apple cultivars and candivars. The cultivars colored in *red* have been cultivated under organic agricultural practice while the cultivars/candivars colored in *black* have been cultivated under conventional agricultural practice. For further details, see Supplementary Data 1 (Color figure online)

Pas-de-Calais, France) respectively, and which were found to be fairly resistant allowing organic cultivation. To the best of our knowledge, the phenolic profile of the following cultivars was not yet reported: 'Belgica', 'Rubinstep', 'Red Jonaprince', 'Red Delcorf', 'Wellant', 'Zari', 'Delcorf', 'Alkmene', 'Argilière', 'Reinette Descarde', 'Reinette de France', 'Jan Steen', 'Calville de Saint-Sauveur', 'Gueule de Mouton', 'Sabot d'Eysden', 'Jaque Lebel', 'Lombarts-Calville', 'Joseph Mush', 'Président Henri Van Dievoet', 'Reinette Étoilée', 'Reinette des Capucins', 'Marie-Joseph d'Othée', 'Reinette Bakker Parmentier', and 'Reinette de Flandre'.

The 47 different apple cultivars analyzed were harvested at optimal maturity in 2010 from different locations in Belgium (Supplementary Data 1). All classical apple cultivars and the recent cultivars 'Appache', 'Wellant', 'Belgica', and 'Rubinstep' were collected from the experimental Orchard of the Department of Pomology of the Belgian Research Station Fruit ('PC Fruit', Kerkom, Sint-Truiden, 50°46′N 5°08′E, altitude 78 m, loamy soil). The trees were grafted on M.9 rootstocks (dwarf trees, 2–2.5 m high) and grown under conventional agricultural practices, expected 'Rubinstep', which was grown under organic practices. All cultivars from breeding programs were harvested at the experimental orchard of a Belgian professional breeder ('Better3Fruit', Rillaar, Aarschot, 50°58′N 4°53′E, altitude 21 m, sandy soil). The trees were grafted on M.9 rootstocks ('Jaune de Metz') with interstem of cv. 'Golden Delicious' (dwarf trees, 1.8–2.5 m high) and grown under conventional agricultural practices. The heritage cultivars included in this study were grown at a private orchard ('Lombarts-Calville', Voormezele, Ieper, 50°49′N 2°52′E, altitude 31 m, sandy loam soil), except from 'Reinette Bakker', 'Parmentier', 'Reinette de Flandre', 'Marie-Joseph d'Othée' which were harvested at a private orchard of a tree nursery ("De Linde", Kemmel, Heuvelland, 50°47′N 2°49′ E, altitude 32 m, sandy soil). The heritage varieties were all originating from standard trees (*Malus communis* 'Bittenfelder' rootstock, interstem 'Gueule de Mouton', +4.5 m high, organic agricultural practices), expected 'Alkmene' (MM.106 rootstock, half-standard tree, 2.5–4.5 m high), 'Jan Steen' (M.26 rootstock, half-standard tree, 2.5-3 m), 'Marie-Joseph d'Othée' (M26 rootstock, interstem of cv. 'Golden Delicious', half-standard tree, 2.5–3 m), and 'Reinette de Flandre' (MM.106 rootstock, without interstem, half-standard tree, 2.5–4.5 m high) which have other constellations. Only the heritage cultivars 'Marie-Joseph d'Othée', 'Reinette

Bakker Parmentier' and 'Reinette de Flandre' were subjected to conventional agricultural practices.

2.2 Sampling and storage

For each apple cultivar, 12 visual healthy fruits were sampled. For all classical cultivars, these 12 fruits were picked out of three trees on the same orchard, growing on a equidistant (13 trees × 1.5 m intertree distance) position in a single row system (3.5 m interrow distance × 1.5 m intertree distance) of a non-randomized block design, excepted cv. 'Pinova' and cv. 'Delcorf' which were picked out of three neighboring trees, in the same single row system. The cultivars 'Wellant', 'Rubinstep' and 'Nicoter' (from 'PC Fruit') were harvested out of three trees growing in on a equidistant (5 trees × 1.5 m intertree distance) position in one row of a monocultivar orchard (3.5 m interrow × 1.5 m intertree distance). For the cultivars 'Belgica', 'Red Delcorf', 'Nicogreen', 'Pinova', 'Zari' and 'Red Jonaprince', 12 fruits were randomly chosen out of the whole production harvest. All other cultivars were harvested from one tree (one biological replicate). Where possible, equal amounts of fruit from the sun and shadow side of the trees were harvested from the outer periphery of the designated trees, avoiding the tops and bottoms of the trees. Three recent cultivars namely, 'Nicoter', 'Nicogreen' and 'Zari' were harvested on two locations ('PC Fruit', 'Better3Fruit'), allowing us to estimate the variation caused by the differences in environmental effects (geographic replicates: climatic conditions during growth and soil). In order to roughly estimate the variation that arises from differences between trees growing under the same environmental conditions, samples of 'Nicoter' were separately collected from 6 different trees.

Fruits were loosely packed inside conventional modular bulk containers and stored at 0 °C for 4 days. The air was exchanged with fans to remove ethylene. Prior to freeze-drying, all 12 fruits from each cultivars were peeled and cored using a manual apple peeler (Préparation & Plaisir C5550, JD Diffusion, France) for a reproducible peel thickness of 3 mm. The samples (peel and flesh) were immediately frozen in liquid nitrogen to avoid enzymatic browning. Directly from the liquid nitrogen, the samples were transferred into a freeze dryer with heated shelves at 25 °C (GAMMA 1-16 LSC, Martin Christ, Osterode am Harz, Germany). Following the freeze-drying, the apple peel and flesh were grounded in a commercial blender (DP705 LA Moulinette, Group SEB, Fleurus, Belgium) and consequently stored in an amber-coloured flask at −25 °C. An inert atmosphere of N_2 gas was established in each flask, avoiding rehydration, biological contamination, and compound degradation.

2.3 Measurement method

An ultrasound assisted consecutive extraction with MeOH/H_2O (20/80, v/v) and 100 % MeOH on 0.5 g freeze-dried material was carried out (De Paepe et al. 2013). The liquid chromatographic (LC) system consisted of an Accela^TM quaternary solvent manager, a 'Hot Pocket' column oven (Thermo Fisher Scientific, Bremen, Germany) and a CTC PAL^TM autosampler (CTC Analytics, Zwingen, Switzerland). A reversed phase separation was performed on a Waters Acquity UPLC^TM BEH SHIELD RP18 column, with dimensions 3.0 × 150 mm, 1.7 μm (Waters, Milford, MA). To protect the UHPLC column, an Acquity BEH RP18 VanGuard pre-column, with dimensions 1.7 lm, 2.1 × 5 mm (Waters, Milford, MA) was coupled with the analytical column. Elution was completed in 23 min with a sequence of linear water-acetonitrile and a flow rate of 500 μL min^{-1}. The UV spectra of all selected phenolic compounds were recorded in the range of 220-400 nm using an Accela^TM photo diode array (PDA) detector. The UHPLC system was coupled to an Orbitrap mass spectrometer operating with an Ion Max^TM ESI source (Thermo Fisher Scientific, Bremen, Germany) in negative ionization mode (ESI-). The software Xcalibur^TM version 2.2.1 (Thermo Fisher Scientific, Bremen, Germany) was used for integration and calculations based on external standards. Samples were analyzed in triplicate, and in random order. Quality control (QC) samples were analysed after every six unknown samples.

2.4 Data preprocessing

Identification and screening workflows were carried out with ToxID 2.1.1.56 (Thermo Fisher Scientific, Bremen, Germany) and Exact Finder Beta version (Thermo Fisher Scientific, Bremen, Germany). Identification of the phenolic compounds indicated with "MSI-level 1" in Table 1 was conducted by accurate mass retention time tags, UV/Vis spectra and MS data (accurate mass, isotopic distribution and fragmentation pattern with in-source fragments) in negative-ion mode with those of the analytical standards that were available (Phytolab, Vestenbergsgreuth, Germany). Identification of the phenolic compounds indicated with "MSI-level 2" in Table 1 was based on the MS data (accurate mass, isotopic distribution and fragmentation pattern with in-source fragments). Peak shape similarities between putative corresponding ions (isotopes, in-source fragments) were manually checked by means of the Qual Browser module of the Xcalibur 2.2.1 acquisition software (Thermo Fisher Scientific, Bremen, Germany) during the identification process. Doing so, ions with both the same apex location and similar peak shape are clustered into one molecular entity. A summary of the parameters used to

Table 1 Identification characteristics of the selected phenolic compound in apple fruit. **For detailed identification characteristics, see Supplementary Data 2**

Phenolic class	Name	Common name	Code	MSI level	ChemSpider ID	Molecular formula	RT (min)	[M-H]⁻ m/z	Wavelength[a] (nm)
Hydroxybenzoic acids	Protocatechuic acid		PA	1	71	$C_7H_6O_4$	7.1	153.01914	259, 294
Hydroxycinnamic acids	*p*-coumaric acid		pCA	1	553148	$C_9H_8O_3$	12.7	163.03980	310
	p-coumaroyl quinic acid		pCQA	2	4945466	$C_{16}H_{18}O_8$	10.98	337.09289	312
	Ferulic acid		FA	1	393368	$C_{10}H_{10}O_4$	12.85	193.05029	325
	Sinapic acid		SINA	1	553361	$C_{11}H_{12}O_5$	12.7	223.06116	324, 225
	Caffeic acid		CA	1	600426	$C_9H_8O_4$	10.36	179.03481	325
	5-*O*-caffeoylquinic acid	Chlorogenic acid	CHLA	1	1405788	$C_{16}H_{18}O_9$	9.6	353.0879	326
Flavones	Apigenin-7-*O*-glucoside	Apigetrin	APIGLU	1	4444290	$C_{21}H_{24}O_9$	13.57	431.09854	267, 336
	Luteolin-7-*O*-glucoside	Cynaroside	LUTGLU	1	4444241	$C_{21}H_{20}O_{11}$	12.62	447.09364	348, 253
Flavonols	Kaempferol-3-*O*-glucoside	Astragalin	KAEGLU	1	4445311	$C_{21}H_{20}O_{11}$	13.64	447.09361	347, 265
	Quercetin		QUE	1	4444051	$C_{15}H_{10}O_7$	16.67	301.03542	371, 255
	Quercetin-3-*O*-glucoside	Isoquercitrin	QUEGLU	1	4444361	$C_{21}H_{20}O_{12}$	12.99	463.08832	354, 255
	quercetin-3-*O*-galactoside	Hyperin	QUEGAL	1	9926320	$C_{21}H_{20}O_{12}$	12.93	463.08853	354, 255
	quercetin-3-*O*-rutinoside	Rutin	QUERUT	1	4444362	$C_{27}H_{30}O_{16}$	12.5	609.14621	355, 255
	Quercetin-3-*O*-arabinoside	Avicularin	QUEARA	1	16736526	$C_{20}H_{18}O_{11}$	13.6	433.07726	353, 254
	Quercetin-3-*O*-rhamnoside	Quercitrin	QUERHA	1	4444112	$C_{21}H_{20}O_{11}$	13.86	447.09315	348, 255
	Isorhamnetin	Isorhamnetin	ISORHA	1	4444973	$C_{16}H_{12}O_7$	18.03	315.05088	371, 253
Flavanols	(+)-catechin		CAT	1	8711	$C_{15}H_{14}O_6$	9.40	289.0717	279
	(−)-epicatechin		eCAT	1	65230	$C_{15}H_{14}O_6$	10.10	289.0717	278
Anthocyanidins	Cyanidin-3-*O*-glucoside[b]	Kuromanin	CYAGLU	1	170681	$[C_{21}H_{21}O_{11}]^+$	8.28	447.09344	512, 279
	Cyanidin-3-*O*-galactoside[b]	Ideain	CYAGAL	1	390306	$[C_{21}H_{21}O_{11}]^+$	8.02	447.09353	512, 279
Dihydrochalcones	Phloretin-2′-*O*-glucoside	Phlorizin	PHLGLU	1	16498836	$C_{21}H_{24}O_{10}$	13.56	435.12936	284
	Phloretin-2′-*O*-xylosylglucoside		PHXYGLU	2	–	$C_{26}H_{32}O_{14}$	12.79	567.17193	284
	3-hydroxyphloretin-2′-*O*-glucoside	3-hydroxyphloridzin	HPHLGLU	2	–	$C_{21}H_{24}O_{11}$	12.66	451.12458	281
	3-hydroxyphloretin-2′-*O*-xylosylglucoside		HPHXYGLU	2	–	$C_{26}H_{32}O_{15}$	11.87	583.16684	281
Procyanidins	(−)-epicatechin-(4β → 8)-(−)-epicatechin	Procyanidin B2	eCATB2	2	4478723	$C_{30}H_{26}O_{12}$	10.21	577.13477	279
	(epi)catechin trimer (unknown)		eCAT3	2	–	$C_{45}H_{38}O_{18}$	10.59	865.19854	280
	(epi)catechin tetramer (unknown)		eCAT4	2	–	$C_{60}H_{50}O_{24}$	10.59	1153.26083	278
	(epi)catechin pentamer 1 (unknown)		eCAT51	2	–	$C_{75}H_{62}O_{30}$	11.25	1441.32531	280
	(epi)catechin pentamer 2 (unknown)		eCAT52	2	–	$C_{75}H_{62}O_{30}$	11.34	1441.32531	279

RT Retention time, *[M-H]*⁻ deprotonated molecule

[a] wavelength of maximum absorption, measured in MeOH/water, 60/40 (v/v)

[b] measured as carbinol pseudobase

identify all phenolic compounds included in this study is depicted in Table 1. Detailed identification characteristics are given as supplementary data (Supplementary Data 2).

MS based quantitative workflows were carried out using Xcalibur 2.2.1 (Thermo Fisher Scientific, Bremen, Germany) software, by using extracted ion chromatograms from Full-Scan Mass Spectra. Quantification was performed by reporting the measured integration areas in the calibration equation of the corresponding standard if available (indicated with "MSI-level 1" in Table 1). If no analytical standard was available, the calibration equation of a similar compound was utilized for quantification of the corresponding phenolic compound (indicated with "MSI-level 2" in Table 1). p-Coumaronyl quinic acid was quantified by p-coumaric acid, phloretin-2′-O-xylosylglucoside, 3-hydroxyphloretin-2′-O-glucoside and 3-hydroxyphloretin-2′-O-xylosyl glucoside were quantified by phloretin-2′-O-glucoside, and (epi)catechin trimer, (epi)catechin tetramer, and two (epi)catechin pentamers were quantified by procyanidin B2. Compound identity was confirmed by close inspection of the mass spectrometric profiles. Spectra that did not meet the quality criteria were excluded from the data. A custom report macro was generated to export raw data into a Microsoft Excel compatible format. By using Microsoft Excel (Microsoft, Redmond, USA), both 'raw' data sets (for peel and flesh) were adjusted for dilution and original sample weight. Concentrations below the corresponding limit of quantification (LOQ) were replaced by the LOQ of the relevant compound. As a result two 'clean' multivariate data sets (for peel and flesh) were obtained, both consisting of a 50×30 matrix, in which rows represented the apple cultivars/candivars analyzed (50 objects: 47 different cultivars/candivars, 3 geographic replicates), and columns the concentrations of 30 individual phenolic compounds.

2.5 Sources of variation

The 'environmental variance' (variation arising from differences in environmental effects: weather and soil) was estimated as the difference between the cultivars, 'Nicoter', 'Nicogreen' and 'Zari' from two locations 'PC Fruit' and 'Better3Fruit'. The 'tree-to-tree variance' which arises from differences between apples with the same genetic background from different trees subject to the same environmental conditions was estimated from the separated analysis on 'Nicoter' samples collected from 6 different trees on the same location ('PC Fruit'). The 'technical variance' or the sum of the 'analytical variance' (error arise from differences during instrumental analysis) and the 'sample work-up variance' (error arise from differences during sample work-up) was estimated from triplicate measurements of the candivars 'Ennoblement 1', 'Ennoblement 2', 'Ennoblement 3',

'Ennoblement 4', 'Ennoblement 6'. Due to the composition of the studied 'variety pool', the possible variance that arises from the difference in agricultural practice could not be estimated in this study. To get insight into the contribution of the studied sources of noise, all technical and biological repeats were included in both peel and flesh datasets.

2.6 Data pretreatment

Data pretreatment was performed before conducting an exploratory data analysis. In order to avoid a heteroscedastic error structure in the data sets, several transformation methods were tested for their effectiveness: Briggsian log, square root, and Anscombe transformation (van den Berg et al. 2006). With the aim to make phenolic compounds comparable to each other when studying the variation between the cultivars/candivars by means of principal component analysis, large differences in offset between the mean concentrations of high and low abundant phenolic compounds in both peel and flesh data sets were removed by means of column-wise arithmetic mean centering. To make the variation study more independent from the abundance of the phenolic compounds, several column-wise normalization (scaling) methods were tested (van der Werf et al. 2005). Due the fact that some phenolic compounds were quantified with the sensitivity factor of a similar compound, only scaling methods that results in dimensionless datasets were considered: autoscaling, range scaling, vast scaling and level scaling (van den Berg et al. 2006; van der Werf et al. 2005). The effectiveness of the scaling methods with regard to the removal of the dependency of the variable value (not proportional to their biological relevance) to the corresponding variance contribution was studied by the correlation between the rank of a phenolic compound based on the corresponding mean concentration and fold change on the one hand, and the rank based on the cumulative contributions of the loadings on the first three principal components on the other hand (van den Berg et al. 2006). To avoid that phenolic compounds with large values contribute more to the distance measure than variables with small values, column-wise autoscaling was conducted before hierarchical clustering was applied on both Briggsian log transformed peel and flesh datasets. In order to allow direct visual representation of the abundance of each phenolic compound throughout the samples, a column-wise normalization by scaling between 0 and 1 was conducted on the Briggsian log transformed datasets. In the perspective to permit a visual comparison of samples with a similar phenolic profile, a row-wise normalization by scaling between 0 and 1 was applied on the Briggsian log transformed datasets. All data pretreatments were conducted by using R

statistical computing software version 2.15.2 (R Foundation, Auckland, New Zealand).

2.7 Exploratory data analysis

Exploratory data analysis was used to search for natural groupings of samples and to obtain an easy-to-understand form, maximizing insight into the datasets (Cuthbertson et al. 2012). In order to provide a partial visualization of data structure in a reduced dimension while retaining the maximum amount of variability present in data, principal component analysis (PCA) was applied on Briggsian log transformed and column-wise scaled datasets (Nyamundanda et al. 2010). To discover structures in data, cluster analysis was used to assign the different samples into groups such that the degree of association between them was maximal if they belong to the same group and minimal otherwise (Tikunov et al. 2012). In this study, k-means clustering (kMC) based on the squared Euclidean distance was performed on the scores of the observations to the first two principal components after PCA for both peel and flesh datasets. The number of clusters was determined based on the 'Elbow' method wherein the percentage of explained variance is the ratio of the between-group variance to the total variance (F test, $\alpha = 0.05$). Furthermore, hierarchical clustering (HCA) based on maximum dissimilarity ('complete linkage') was applied on the Briggsian log transformed peel and flesh datasets. In order to further detect underlying interrelationships, HCA based on the squared Euclidean distance and maximum dissimilarity was applied on column-wise autoscaled and Briggsian log transformed datasets. To graphically visualize the abundance of the different phenolic compounds throughout the samples, heatmaps of the Briggsian log transformed and column-wise normalized peel and flesh dataset were employed. With the aim to visualize similarities in phenolic profile between the studied samples, heatmaps of the Briggsian log transformed and row-wise normalized peel and flesh datasets were generated. All data exploration methods were performed by using R statistical computing software version 3.0.1 (R Foundation, Auckland, New Zealand) with follow packages: 'Bioconductor' version 2.12 (PCA, kMC, HCA) and 'pheatmap' version 0.7.4 (heatmaps).

3 Results

Concentration of considered phenolic compounds in the peel and flesh of the studied apple cultivars/candivars are given in Supplementary Data 3 and Supplementary Data 4 respectively.

3.1 Data pretreatment

Due to nearly constant relative standard deviations of technical error, it was possible to largely remove the heteroscedasticity in both peel and flesh datasets by applying a Briggsian log transformation, which was not the case for the square root and Anscombe transformation (Supplementary Data 5 and Supplementary Data 6). Applying different column-wise scaling methods on the Briggsian log-transformed data sets demonstrates that range scaling and autoscaling were able to remove the dependency (expressed as Pearson's correlation coefficient) of the rank of the phenolic compounds, on both average concentration and magnitude of the fold changes. This result was not achieved for vast scaling and level scaling, which is in conformity with the findings of van den Berg et al. (2006). Furthermore, the variance explained by the first two principal components in the PCA analysis is also affected by the method of scaling: autoscaling and range scaling result in similar explained variances (58.3 and 57.0 % respectively), while vast scaling and level scaling result in higher levels of explained variances (74.1 and 86.4 % respectively). However, conducting k-means clustering on the scores of the performed PCA on both datasets results in only one cluster when using vast scaling and level scaling, while autoscaling and range scaling delivers 3 separate clusters. It was decided to perform auto-scaling before each PCA and HCA.

3.2 Sources of variation

The result of the PCA, conducted on both peel and flesh datasets including the technical and biological repeats, are displayed in Figs. 2 and 3. The figures indicate the sample scores and variable loading by means of a bidimensional plot defined by the two first principal components. This plot provides insight into the various sources of variation that contribute to the total phenolic diversity in the evaluated set of apple cultivars and candivars. The technical replicates (indicated with "M1, M2, …" in Figs. 2 and 3) of the candivars 'EN1', 'EN2', 'EN3', 'EN4', 'EN6' are tightly grouped, indicating a very low amount of technical variance. The replicates are also tightly linked in the corresponding dendrogram (Figs. 4a and 5a), which illustrates that the variance among technical replicates is smaller than differences between the biological variants. Furthermore, the tree-to-tree replicates (indicated with "T1, T2, …" in Figs. 2 and 3) of 'NCTR' also group together. It should be noted that additional tree-to-tree variation is present besides the technical variation, leading to slightly more dispersed data. This artifact indicates that there is few heterogeneity between trees of the same genotype growing under the same conditions (weather, soil, applied
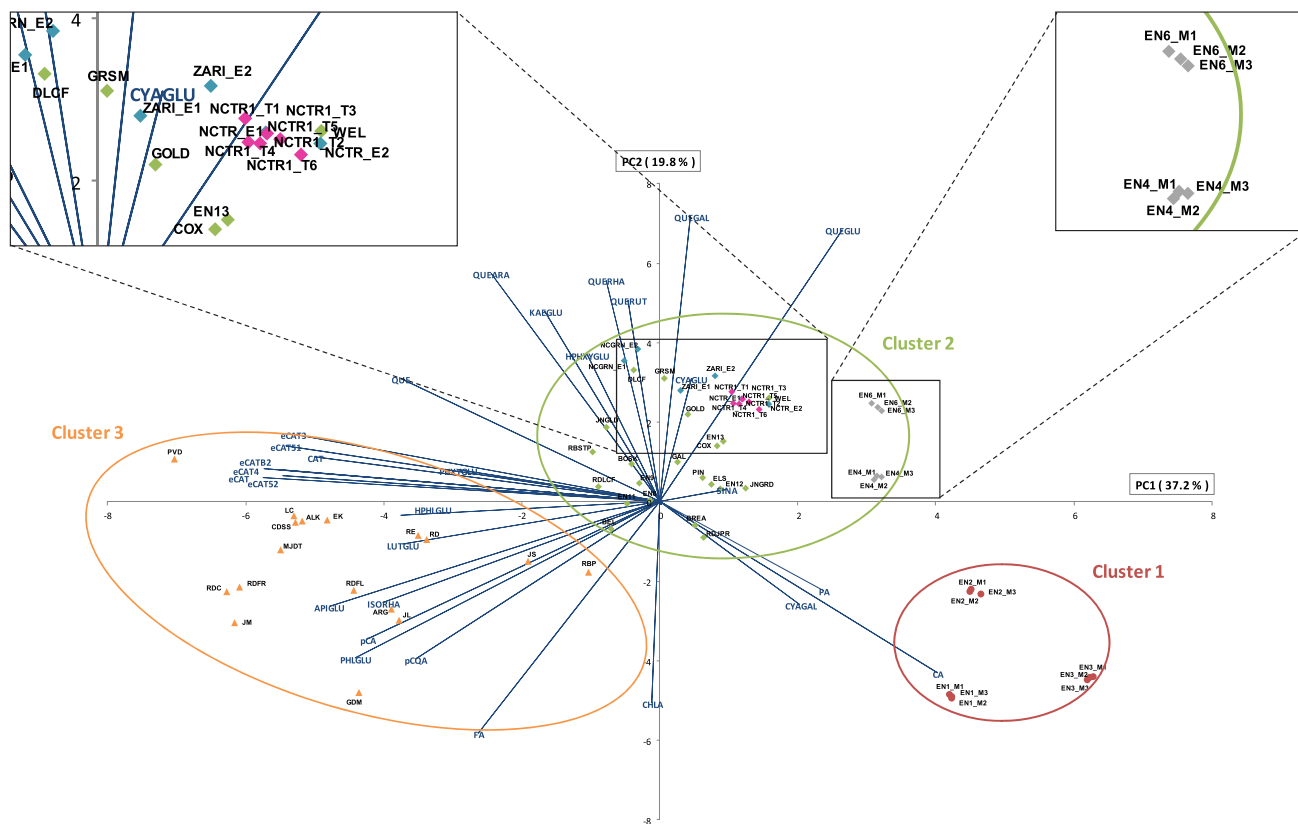
Fig. 2 *k*-mean clustering conducted on the score plot of the first two principal components (PC1 × PC2) and of peel dataset including 47 different apple cultivars/candivars and 30 phenolic compounds. *Red filled circle* cluster 1-*red* flesh, *green filled diamond* cluster 2-ennoblement, new and classic, *pink filled diamond* between tree replicates in cluster 2, *blue filled diamond* environmental replicates in cluster 2, *orange filled traingle* cluster 3-heritage cultivars. For explanation of abbreviations used, see Table 1 (Color figure online)

agricultural practice) and supports our hypothesis to disregards the effect of the tree on the phenolic profile. Finally, a grouping between the environmental replicates of respectively 'NCTR', 'NCGRN' and 'ZARI' (indicated with "E1, E2, …" in Figs. 2 and 3) can be observed, although the dispersal between the environmental replicates for each genotype is the result of the environmental variation and the technical variation. This observation suggests a relative low impact of environmental conditions (weather and soil) on the whole phenolic profile, with the constraint of possible small differences in growing conditions between the two harvest places (locations 'PC Fruit' and 'Better3Fruit'). A slightly increased heterogeneity can be seen between the environmental replicates. The possible influence of the type agricultural practice was not taken into account in this variability study.

3.3 Distinction based on phenolic profile of the peel data

Conducting *k*-means clustering on the scores of the performed PCA (Supplementary Data 7) on the peel dataset (Briggsian log transformation and column-wise auto-

scaled) results in the segregation of the apple cultivars/candivars into three clusters (explained variation 57.0 %) (Fig. 2). A first cluster was formed by the red-flesh apple candivars included in this study. Furthermore, it was possible to make a distinction between the classic/new cultivars and candivars (cluster 2) on the one hand and the heritage apple cultivars (cluster 3) on the other hand. Large difference in phenolic profile could be observed between the mentioned clusters. Performing HCA on the columns of the pretreated data delivers almost the same segregation with more nuances (Fig. 4a).

The phenolic composition of the peel of the red-fleshed apple candivars 'EN1', 'EN2', 'EN3', which all have a red to dark-red skin, are typified by high amounts of chlorogenic acid and caffeic acid and the lack of oligomeric procyanidins and flavanols (Figs. 2 and Fig. 4b). In contrast with all studied classical/new cultivars and candivars, the most abundant phenolic compounds in the peel of red-fleshed apple candivars were not quercetin-glycosides, but chlorogenic acid and caffeic acid (Fig. 4c). As for each studied red colored apple cultivar/candivar, the main anthocyanin identified in apple skin is cyanidin-3-*O*-galactoside, while cyanidin-3-*O*-glucoside levels were very
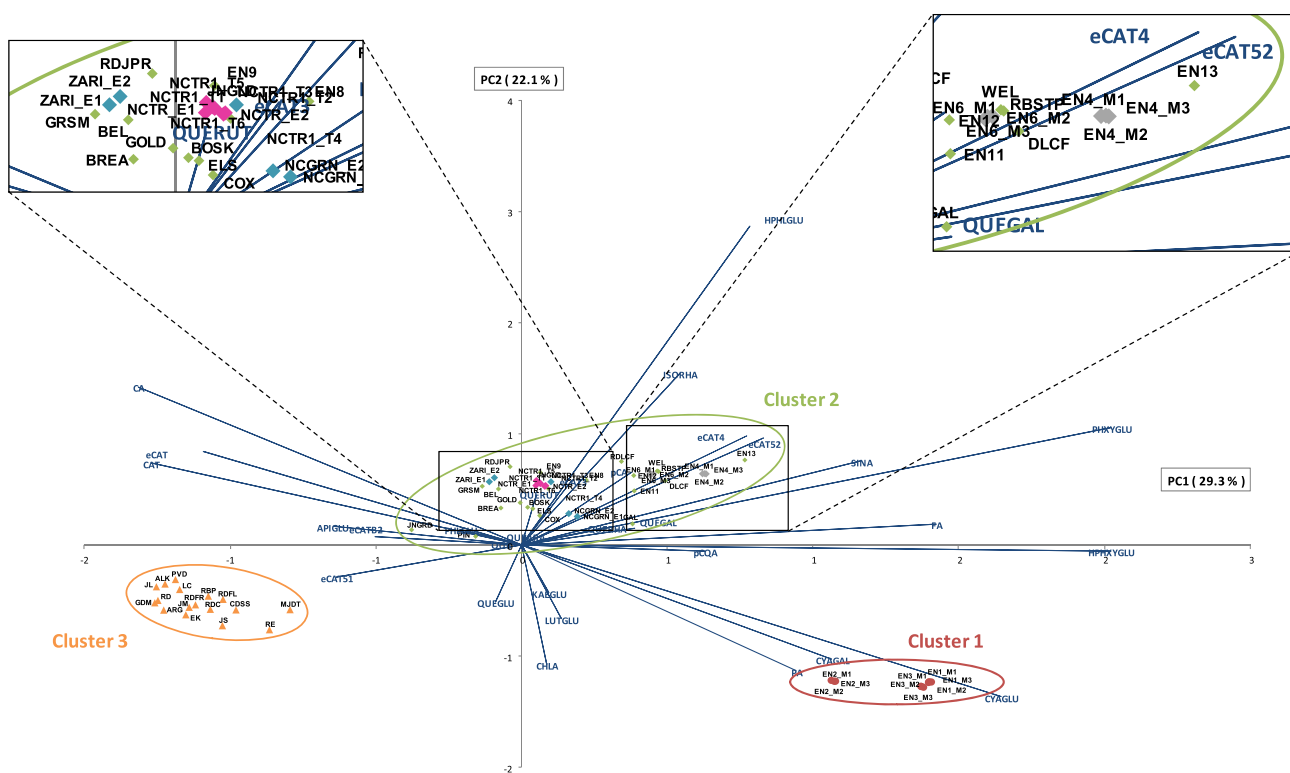
Fig. 3 *k*-mean clustering conducted on the score plot of the first two principal components (PC1 × PC2) and of flesh dataset including 47 different apple cultivars/candivars and 30 phenolic compounds. *red filled circle* cluster 1-*red* flesh, *green filled diamond* cluster 2-ennoblement, new and classic, *pink filled diamond* between tree replicates in cluster 2, *blue filled diamond* environmental replicates in cluster 2, *orange filled traingle* cluster 3-heritage cultivars. For explanation of abbreviations used, see Fig. 1 (Color figure online)

low (Fig. 4c). Based on the peel phenolic profile, the candivars 'EN6' (intense red), 'EN9' (dark-red), and 'EN4' (purple) cluster together with the red-fleshed apple candivars (Fig. 4a).

The phenolic composition of the peel of most classic/new cultivars and candivars was typified by the flavonol subclass, especially quercetin glycosides (Figs. 2 and 4b). The most abundant flavonol in all studied genotypes of cluster 2 (Fig. 2) is quercetin-3-*O*-galactoside, excluding 'EN8' where quercetin-3-*O*-rhamnoside dominates the flavonol spectrum. For each cultivar/candivar of this group, the content of quercetin-3-*O*-glucoside was larger than quercetin-3-*O*-arabinoside. This observation is in contrast with the heritage varieties group where the opposite sequence was found (Fig. 4c). Noteworthy is the clustering of 'ZARI', 'NCGRN', 'RDLCF' and 'EN11' with their corresponding ancestors 'ELS', 'GOLD', 'GRSM' and 'DLCF' (Fig. 4a). Furthermore, following our approach, JNGLD' could be combined in a cluster with its corresponding color mutants 'JNGRD' and 'RDJPR' (Fig. 4).

The phenolic composition of the peel of heritage apple cultivars is mainly characterized by phenolic compounds out of the flavonoid subclasses flavan-*3*-ols, dihydrochalcones and procyanidins, and two hydroxycinnamic acids

p-coumaronyl quinic acid and chlorogenic acid (Fig. 4b). The most abundant dihydrochalcones found in heritage cultivars were phloretin-2′-*O*-glucoside and 3-hydroxyphloretin-2′-*O*-glucoside, while the classical and new cultivars and candivars are more typified by phloretin-2′-*O*-xylosylglucoside and 3-hydroxyphloretin-2′-*O*-xylosylglucoside (Fig. 4b). Based on this segregating phenolic profile, it was possible to cluster all studied heritage cultivars with exception of 'RBP' by means of HCA (Fig. 4).

3.4 Distinction based on phenolic profile of the flesh data

Applying *k*-means clustering on the scores of the performed PCA (Supplementary Data 8) on the flesh dataset results in the same segregation of the apple cultivars/candivars as derived from the peel datasets (explained variation 51.4 %) (Fig. 3). However, due to the difference in phenolic profile between peel and flesh, other phenolic compounds are responsible for these assignments (Fig. 5). Both phenolic profiles of heritage cultivars as classic/new cultivars and candivars are dominated by flavanols, flavanols and dihydrochalcones, although, based on the individual representatives, differences could be observed.
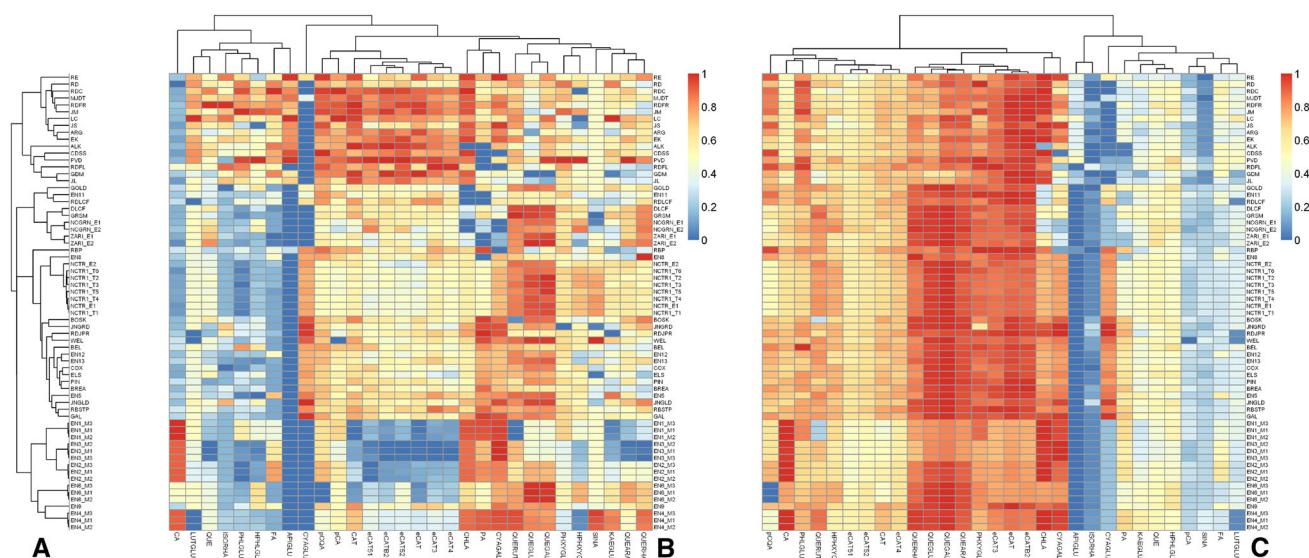
Fig. 4 Hierarchical clustering (**a**), heatmap on the column scaled (**b**), and heatmap of the row scaled (**c**) peel dataset. For explanation of abbreviations used, see Table 1 and Fig. 1
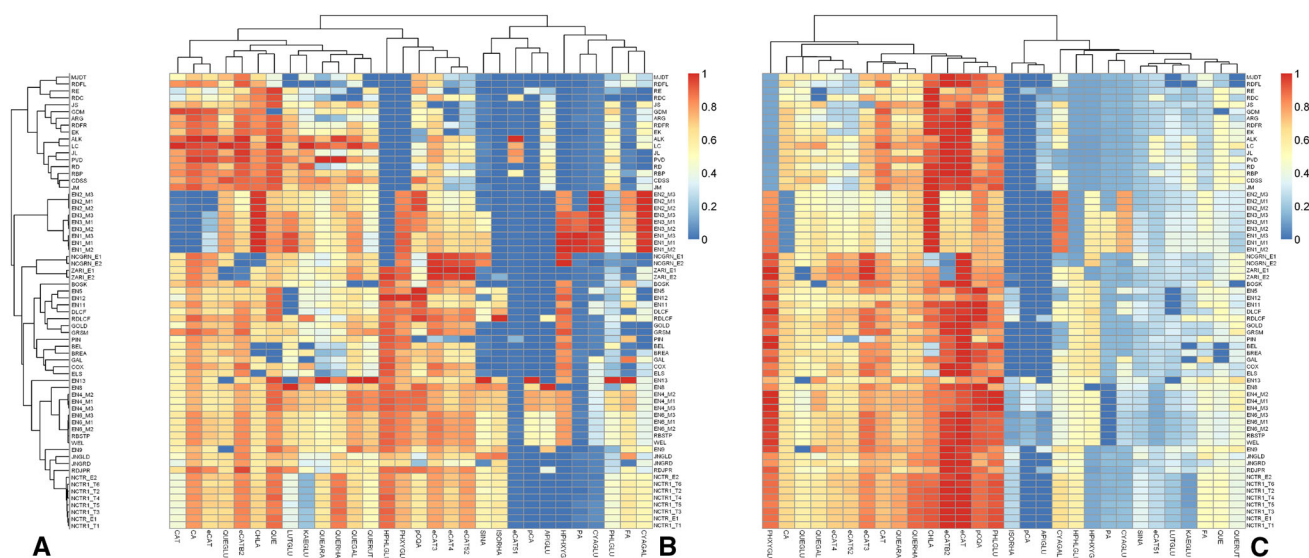

Fig. 5 Hierarchical clustering (**a**), heatmap on the column scaled (**b**), and heatmap of the row scaled (**c**) flesh dataset. For explanation of abbreviations used, see Table 1

Based on HCA, it was possible to sort all studied heritage cultivars into one group (Fig. 5a).

The phenolic profile of the flesh of heritage cultivars was dominated by (−)-epicatechin, (+)-catechin and caffeic acid, while the phenolic profile of the classic/new cultivars and candivars excelled by the presence of 3-hydroxyphloretin-2′-*O*-glucoside, phloretin-2′-*O*-xylosylglucoside and 3-hydroxyphloretin-2′-*O*-xylosylglucoside (Fig. 5b). The phenolic profile of the flesh of red-fleshed apple candivars is characterized by anthocyanidins cyanidin-3-*O*-glucoside, cyanidin-3-*O*-galactoside

and protocatechuic acid which allows them to segregate in a separated cluster (Fig. 3).

4 Discussion

4.1 Sources of variation

Within the selected 'variety pool', great variation was observed in the distribution and composition of the measured phenolic compounds. This variation arises from the

dependency of phenolic composition on the interactions between genetic makeup, the developmental stage of the fruit, and environmental factors (Treutter 2005).

In apple fruit, hydroxybenzoic acids are derived from chorismic acid (salicylic acid) and 3-hydroshikimic acid (protocatechuic acid and gallic acid), both intermediates from the shikimate pathway which have L-phenylalanine as an end product. All other phenolic compounds are synthesized from L-phenylalanine through the phenylpropanoid pathway and flavonoid pathway. Chalcones, flavanones, flavanonols, and flavanols can be considered as intermediate products of phenylpropanoid pathway. Anthocyanins, procyanidins, and flavonols are synthesized via the flavonoid pathway, a branch of the phenylpropanoid pathway (Takos et al. 2006). These complex biosynthetic pathways involve a series of enzymes, which have been, as well as the encoding genes, already functionally characterized in apple fruit (Khan et al. 2012). The formation and activity of these enzymes are determined by the maturity stage, as well as internal (derived from genome differences) and several external (stress conditions existing before and after harvest) factors (Slatnar et al. 2012). Therefore, studying the variation in the phenolic profiles engendered by genetic background apart from other influences is not feasible. However, based on low contributions of the studied sources of noise (tree-to-tree and environmental variation) and the harvest of all cultivars/candivars at optimal maturity, which establish a sound basis for equivalent comparison, it is possible to suggest that the observed variability in both peel and flesh datasets could be substantially linked to the variation in genetic background.

4.2 Distinction based on phenolic profile of the peel data

Apple peel color is one of the most important factors to determine apple market acceptance. In general, red cultivars are the most preferred, and within a cultivar more highly colored fruits are favored (Telias et al. 2011). In red-fleshed apple genotypes, which exhibit an intense dark red skin color, cyanidin-3-O-galactoside represents the major flavonoid but occurs in all red peel colored cultivars/candivars. This finding is in accordance with other studies (Rupasinghe et al. 2009). Generally, the intensity and pattern of anthocyanin accumulation in apple peels can be affected by UV-B irradiation and low-temperature treatment, but is connected to a large genetic component (Telias et al. 2011). Several genes required for anthocyanin synthesis in the skin of apple fruit were coordinately transcribed in response to light exposure by means of the transcription factor genes MdMYB1 and MdMYBA (Takos et al. 2006). However, the transcripts of some required structural genes were found to be barely detectable in the

non-red skin cultivars but were, to a greater or lesser extent, abundant in the red skin cultivars suggesting that expression of these genes are controlled by a common regulator that is defective in non-red skin cultivars (Honda et al. 2002). This regulator is responsible for anthocyanin in the peel of the studied genotypes including the red-fleshed candivars.

Beside cyanidin-3-O-galactoside, the hydroxcinnamic acids chlorogenic acid and caffeic acid dominates the phenolic profile of the studied red-fleshed candivars (Balázs et al. 2012). It was found that caffeic acid and chlorogenic acid exhibit intermolecular copigmentation with cyanidins. Indeed, they meet the main prerequisite as the copigment, containing a planar π electron rich ring, which allows a large π-π overlap with the pyrilium ring of the flavylium ion. The complexation of a copigment with an anthocyanin causes a hyperchromic effect (increase in colour intensity) and a bathochromic shift (red shift) (Sun et al. 2010). This conjecture could be a molecular explanation for the different shades of red in apple skin of the studied red-fleshed candivars. The high levels of cyanidin-3-O-galactoside, caffeic acid, and chlorogenic acid in the candivars 'EN4', and 'EN6', which all have an intense red and dark-red blush respectively, results in the clustering of these candivars together with the other red-fleshed apple cultivars. The remarkably lower content of cyanidin-3-O-glucoside, chlorogenic acid, and caffeic acid in 'EN9' could be a possible indication for its purple-like blush.

The phenolic profile of the group of classic/new cultivars and candivars is typified by higher concentrations of flavonols and lower concentrations of dihydrochalcones, proanthocyanidins and chlorogenic acid. Interestingly, the cluster analysis of the row-wise normalized peel dataset correctly reflected the genetic relationships among the apple cultivars 'ZARI', 'NCGRN', 'RDLCF', 'EN11' with their corresponding ancestors 'ELS', 'GOLD', 'GRSM' and 'DLCF' (Supplementary Data 9). 'GD' belongs to the founding clones. The ancestry of 'ZARI' is quite complicated. A cross of 'GOLD' and 'Ingrid Marie' (not included in this analysis) generated 'ELS'. Furthermore, a cross of 'Jon Grimes' (not included in this study) and 'GOLD' generated 'DLCF'. 'ZARI' emerged from a further cross of 'ELS' with 'DLCF'. As a consequence, 'GOLD' shows up twice in the pedigree of 'ZARI'. 'NCGRN' is a cross of 'DCF' and 'GS' and has 'GOLD' in his pedigree. One of the ancestors of 'EN11' is 'DLCF'. 'GRSM' was propagated from a chance seedling of unknown identity, but it has been suggested that *Malus sylvestris* may have been a significant contributor to its gene pool (Harrison and Harrison 2011). 'RDCLF' is a color mutant of 'DLCF'. All constituents of the pedigree of 'ZARI' and 'NCGRN' have a green to green/yellow background color, in some cases with a red blush (case for 'ZARI', 'EST', 'RDLCF'). As

expected, 'ZARI' and 'NCGRN' are positioned near their ancestors in both for the HCA on peel and flesh datasets (Figs. 4 and 5). Especially for the peel dataset, a complete subcluster with all constituents of the pedigree within the cluster of classical and new cultivars and candivars could be identified. Also a tight clustering of 'JS' with one of his ancestors 'RE' was found (Supplementary Data 10). The pedigree of 'Alkmene', 'Pinova', 'Jonagold', and 'Jonagored' are given in Supplementary Data 11.

This metabolomic data analysis illustrates the consequences of selection and breeding techniques used in modern apple breeding. The repeated use of a limited number of cultivars with good characteristics as breeding parents, results in the narrowing of the genetic variation encoding for different traits, as illustrated in this study for diversity in the phenolic profiles. Furthermore, the utility of metabolomics towards a tool for assessing the quality of traits in apple fruit is confirmed (Cuthbertson et al. 2012).

The studied heritage cultivars (except 'JS', 'CDSS', 'LC' and 'MJDT') have shown resistance over years to the most important apple diseases (scab, cancer, mildew). Given the fact that phenolic compounds are involved in disease defense, it could be expected that this group of cultivars would exhibit a deviating phenolic profile, particularly in the skin which acts as a protective barrier between the plant tissue and its environment. The cluster of heritage cultivars was indeed typified by an elevated concentration of flavanols, procyanidins, dihydrochalcones and hydroxycinnamic acids. There is evidence that, flavan-3-ols and chlorogenic acid content are the highlighted criteria affecting the resistance of apple fruits against scab disease (Khanizadeh et al. 2006). Moreover, these compounds are also inversely related with the powdery mildew incidence in apple orchards (Singh et al. 2010). Furthermore, phloridzin, which represents the major dihydrochalcone in heritage apples, has often been linked to resistance to numerous diseases (Slatnar et al. 2010). Studies on apple have shown that resistance to *V. inaequalis* responsible for scab was related to the presence of this compound (Hamauzu 2006). The fungus degrades phloridzin to its toxic aglycone, phloretin. So, it could be suggested that the typical phenolic composition of the heritage apple varieties included in this study forms a partially explanation for resistance to scab and powdery mildew. Indeed, most of the heritage apple varieties (except 'JS', 'CDSS', 'LC' and 'MJDT') sampled from the collection orchards of Walloon Agricultural Research Centre (CRA-W, Wallon, Belgium) and Regional Natural Space (Region Nord/Pas-de-Calais, France) are maintained for their excellent disease characteristics.

It should be noted that the most (except 'MJDT' and 'RDFL') of the heritage cultivars were cultivated under organic agricultural practices, while most of the classic/new cultivars and candivars were cultivated under conventional agricultural practice. This could be also a possible explanation for the deviating phenolic profile of the heritage cultivars. Namely, phenolic compounds are involved in plant defence against biotic stress factors, leading to the production and accumulation of an array of defence related phenolic compounds during disease pressure (Petkovsek et al. 2010; Zhao and Rajashekar 2006; Mayr et al. 1997). On this manner, the observed elevated phenolic content some phenolic constituents in the heritage cultivars could be the underlying factor of their low susceptibility against the most important apple diseases, or the result of disease pressure by the lack of pesticides use. In this study, no conclusion can be placed in this context, by the fact that no comparison is made between heritage cultivars cultivated under conventional and organic practice. Moreover, the comparison between conventional and organic agricultural practice for heritage cultivars is hampered by the fact that the most heritage cultivars mostly concern high stem trees, whereby the application of conventional agricultural practice is impossible.

Several studies have been conducted to gain information regarding the impact of the production method, organic versus conventional (or under integrated agricultural methods with controlled fertilizer and pesticides use), on the phenolic content in apples and other fruit species (Stracke et al. 2010). However, the results are inconsistent in literature by the fact that the difference between both agricultural practices seems to be dependent on the cultivar, the harvest season, and the identity of individual phenolic compounds taken into account (Valavanidis et al. 2009; Petkovsek et al. 2010). However it has been found that the production method appears to have only a small impact on the phenolic content in comparison with cultivar differences (Ceyman et al. 2013).

4.3 Distinction based on phenolic profile of the flesh data

Whilst the majority of apples have white or cream flesh, a small number of varieties have red flesh or pink flesh. In contrast with the skin color where the transcription factor genes MdMYB1 and MdMYBA are involved, flesh color is regulated by means of the transcriptor factor gene MYB10 (Telias et al. 2011). The transcription of some required structural genes for anthocyanin synthesis in the flesh of white-fleshed apple fruit is limited by levels of MdMYB10 (van Nocker et al. 2012). However, in red-fleshed apple fruit, the structural genes involved in producing anthocyanins were more highly expressed, forming the genetic basis for the phenotypical difference between red-fleshed and white-fleshed genotypes (Espley et al. 2007).

Within the group of the current commercial varieties and classical varieties, 'NCGRN' exhibits a distinct phenolic profile. 'NCGRN' is characterized by an extreme low chlorogenic acid, while the *p*-coumaronyl quinic acid content is higher. As a result, a small chlorogenic acid to *p*-coumaronyl quinic acid ratio is obtained which is an important parameter in browning susceptibility. Namely, chlorogenic acid is considered to be a preferential natural substrate of the catecholase activity of polyphenoloxidase, whereas *p*-coumaronyl quinic acid seems to be a competitive inhibitor of this enzymatic activity. Furthermore, the high content of high-molecular weight procyanidins in 'NCGRN' flesh could provide an additional anti-browning effect (Le Bourvellec et al. 2004). It is therefore not surprising that 'NCGRN' is known for his good resistance to enzymatic browning after cutting, making this cultivar suitable for processing fruit salads and bakery products.

5 Conclusion

In this study, a liquid chromatography—high-resolution accurate mass spectrometry approach was employed to explore the utility of the phenolic pattern to differentiate between cultivars and to assess some important quality traits (color, disease resistance, and browning) within a unique variety pool of 47 apple cultivars. It was possible to segregate the studied apple cultivars/candivars into three clusters: the red-flesh apple candivars, classic/new cultivars and candivars and the heritage apple cultivars. It was also shown that heritage apple varieties selected for their disease resistance have a remarkable different phenolic profile than the classic/new cultivars. This is a promising result that must be confirmed with heritage cultivars cultivated under the same conditions as the other studied cultivars/candivars. Furthermore, several linkages could be made between quality traits like color and resistance against diseases and browning on one hand and the phenolic makeup on the other hand. Moreover, it was found that the phenolic makeup of the peel could suggest the genetic relationships among the apple cultivars. The phenolic profiles from the different apple cultivars and candivars evaluated in the present study could help in (1) selecting cultivars with a typical phenolic profile for quality traits or disease resistance related breeding programs, (2) selection of apple cultivars with better processing characteristics, (3) selecting cultivars for studies designed to measure the physiological effects of phenolic compounds. By the fact that this study just use one biological replicate, the outcomes of this study needs further validation by choosing representative cultivars based on the results presented in this manuscript.

References

Balázs, A., Tóth, M., Blazics, B., Héthelyi, É., Szarka, S., Ficsor, E., et al. (2012). Investigation of dietary important components in selected red fleshed apples by GC-MS and LC-MS. *Fitoterapia, 83*, 1356–1363.

Ceymann, M., Arrigoni, E., Schärer, H., Nising, A. B., & Hurrell, R. F. (2012). Identification of apples rich in health-promoting flavan-3-ols and phenolic acids by measuring the polyphenol profile. *Journal of Food Composition and Analysis, 26*, 128–135.

Cornille, A., Gladieux, P., Smulders, M. J. M., Roldán-Ruiz, I., Laurens, F., Le Cam, B., et al. (2012). New insight into the history of domesticated apple: Secondary contribution of the European wild apple to the genome of cultivated varieties. *PLoS Genetics, 8*, 1–13.

Cuthbertson, D., Andrews, P. K., Reganold, J. P., Davies, N. M., & Lange, B. M. (2012). Utility of metabolomics toward assessing the metabolic basis of quality traits in apple fruit with an emphasis on antioxidants. *Journal of Agricultural and Food Chemistry, 60*, 8552–8560.

De Paepe, D., Servaes, K., Noten, B., Diels, L., De Loose, M., Van Droogenbroeck, B., et al. (2013). An improved mass spectrometric method for identification and quantification of phenolic compounds in apple fruits. *Food Chemistry, 136*, 368–375.

Espley, R. V., Hellens, R. P., Putterill, J., Stevenson, D. E., Kutty-Amma, S., & Allan, A. C. (2007). Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. *The Plant Journal, 49*, 414–427.

Hamauzu, Y. (2006). Role and evolution of fruit phenolic compounds during ripening. *Stewart Postharvest Review, 2*, 1–7.

Harrison, N., & Harrison, J. R. (2011). On the evolutionary history of the domesticated apple. *Nature Genetics, 43*, 1043–1044.

Honda, C., Kotoda, N., Wada, M., Kondo, S., Kobayashi, S., Soejima, J., et al. (2002). Anthocyanin biosynthetic genes are coordinately expressed during red coloration in apple skin. *Plant Physiology and Biochemistry, 40*, 955–962.

Iacopini, P., Camangi, F., Stefani, A., & Sebastiani, L. (2010). Antiradical potential of ancient Italian apple varieties of Malus x domestica Borkh. in a peroxynitrite-induced oxidative process. *Journal of Food Composition and Analysis, 23*, 518–524.

Kellerhals, M., Szalatnay, D., Hunziker, K., Duffy, B., Nybom, H., Ahmadi-Afzadi, M., et al. (2012). European pome fruit genetic resources evaluated for disease resistance. *Trees, 26*, 179–189.

Khan, S. A., Chibon, P. Y., de Vos, R. C. H., Schipper, B. A., Walraven, E., Beekwilder, J., et al. (2012). Genetic analysis of metabolites in apple fruits indicates an mQTL hotspot for phenolic compounds on linkage group 16. *Journal of Experimental Botany, 63*, 2895–2908.

Khanizadeh, S., Groleau, Y., Levasseur, A., Carisse, O., & Rekika, D. (2006). 'SuperMac' apple. *HortScience, 41*, 1159–1161.

Lammerts van Bueren, E. T., & Myers, J. R. (2011). *Organic crop breeding*. Hoboken: Wiley, Inc.

Lateur, M., & Populer, C. (1994). Screening fruit tree genetic resources in Belgium for disease resistance and other desirable characters. In H. Schmidt & M. Kellerhals (Eds.), *Progress in temperate fruit breeding* (pp. 425–431). Dordrecht: Kluwer Academic Publishers.

Le Bourvellec, C., Le Quéré, J.-M., Sanoner, P., Drilleau, J.-F., & Guyot, S. (2004). Inhibition of apple polyphenol oxidase activity by procyanidins and polyphenol oxidation products. *Journal of Agricultural and Food Chemistry, 52*, 122–130.

Mayr, U., Michalek, S., Treutter, D., & Feucht, W. (1997). Phenolic compounds of apple and their relationship to scab resistance. *Journal of Phytopathology, 145*, 69–75.

Nyamundanda, G., Brennan, L., & Gormley, I. C. (2010). Probabilistic principal component analysis for metabolomic data. *BMC Bioinformatics, 11*, 571–580.

Petkovsek, M., Slatnar, A., Stampar, F., & Veberic, R. (2010). The influence of organic/integrated production on the content of phenolic compounds in apple leaves and fruits in four different varieties over a 2-year period. *Journal of the Science of Food and Agriculture, 90*, 2366–2378.

Richards, C. M., Gayle, M. V., Reilley, A. A., Henk, A. D., Lockwood, D. R., Reeves, P. A., et al. (2009). Genetic diversity and population structure in *Malus sieversii*, a wild progenitor species of domesticated apple. *Tree Genetics & Genomes, 5*, 333–347.

Rupasinghe, H. P. V., Wang, L., Pitts, N. L., & Astatkie, T. (2009). Baking and sensory characteristics of muffins incorporated with apple skin powder. *Journal of Food Quality, 32*, 685–694.

Singh, A., Singh, K. P., Rajwar, G. S., & Singh, U. P. (2010). Phenolic acid content—a criterion for selection of resistant apple cultivars against *Podosphaera leucotricha* (Ell. and Ev.) Salmon. *Phytopathology And Plant Protection, 43*, 1138–1143.

Slatnar, A., Petkovsek, M. M., Halbwirth, H., Stampar, F., Stich, K., & Veberic, R. (2010). Response of the phenylpropanoid pathway to Venturia inaequalis infection in maturing fruit of 'Braeburn' apple. *Journal of Horticultural Science & Biotechnology, 85*, 465–472.

Slatnar, A., Petkovsek, M. M., Halbwirth, H., Stampar, F., Stich, K., & Veberic, R. (2012). Polyphenol metabolism of developing apple skin of a scab resistant and a susceptible apple cultivar. *Trees, 26*, 109–119.

Son, K. M., Kwon, S. I., & Choi, C. (2012). Inbreeding, coancestry, and founding clones of apple cultivars released from Korea. *Horticulture, Environment, and Biotechnology, 53*, 404–409.

Stracke, B. A., Rüfer, C. E., Bub, A., Seifert, S., Weibel, F. P., Kunz, C., et al. (2010). No effect of the farming system (organic/conventional) on the bioavailability of apple (Malus domestica Bork., cultivar Golden Delicious) polyphenols in healthy men: a comparative study. *European Journal of Nutrition, 49*, 301–310.

Sun, J., Cao, X., Bai, W., Liao, X., & Hu, X. (2010). Comparative analyses of copigmentation of cyanidin 3-glucoside and cyanidin 3-sophoroside from red raspberry fruits. *Food Chemistry, 120*, 1131–1137.

Takos, A. M., Jaffe, F. W., Jacob, S. R., Bogs, J., Robinson, S. P., & Walker, A. R. (2006). Light-induced expression of a MYB gene regulates anthocyanin biosynthesis in red apples. *Plant Physiology, 142*, 1216–1232.

Telias, A., Lin-Wang, K., Stevenson, D. E., Cooney, J. M., Hellens, R. P., Allan, A. C., et al. (2011). Apple skin patterning is associated with differential expression of MYB10. *BMC Plant Biology, 11*, 93–106.

Tikunov, Y. M., Laptenok, S., Hall, R. D., Bovy, A., & de Vos, R. C. H. (2012). MSClust: a tool for unsupervised mass spectra extraction of chromatography-mass spectrometry ion-wise aligned data. *Metabolomics, 8*, 714–718.

Treutter, D. (2005). Significance of flavonoids in plant resistance and enhancement of their biosynthesis. *Plant Biology, 7*, 581–591.

Valavanidis, A., Vlachogianni, T., Psomas, A., Zovoili, A., & Siatis, V. (2009). Polyphenolic profile and antioxidant activity of five apple cultivars grown under organic and conventional agricultural practice. *International Journal of Food Science & Technology, 44*, 1167–1175.

van den Berg, R. A., Hoefsloot, H. J. C., Westerhuis, J. A., Smilde, A. K., & van der Werf, M. J. (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics, 7*, 142–156.

van der Werf, M. J., Jellema, R. H., & Hankemeier, T. (2005). Microbial metabolomics: Replacing trial-and-error by the unbiased selection and ranking of targets. *Journal of Industrial Microbiology and Biotechnology, 32*, 234–252.

van Nocker, S., Berry, G., Najdowski, J., Michelutti, R., Luffman, M., Forsline, P., et al. (2012). Genetic diversity of red-fleshed apples (Malus). *Euphytica, 185*, 281–293.

van Treuren, R., Kemp, H., Ernsting, G., Jongejans, B., Houtman, H., & Visser, L. (2010). Microsatellite genotyping of apple (*Malus × domestica Borkh.*) genetic resources in the Netherlands: application in collection management and variety identification. *Genetic Resources and Crop Evolution, 57*, 853–865.

Vanzo, A., Jenko, M., Vrhovsek, U., & Stopar, M. (2013). Metabolomic profiling and sensorial quality of 'Golden Delicious', 'Liberty', 'Santana', and 'Topaz' apples grown using organic and integrated production systems. *Journal of Agricultural and Food Chemistry, 61*, 6580–6587.

Zhao, X., & Rajashekar, C. B. (2006). Does Organic Production Enhance Phytochemical Content of Fruit and Vegetables? Current Knowledge and Prospects for Research. *HortTechnology, 16*, 449–456.