

Untargeted metabolomics analysis reveals the biochemical variations of polyphenols in a diverse apple population

Jun Song^{1*}, Beatrice Amyotte¹, Cindy H.J. Yu², Leslie Campbell-Palmer¹, Melinda Vinqvist-Tymchuk¹ and H.P. Vasantha Rupasinghe²

¹ Agriculture and Agri-Food Canada, Kentville Research and Development Centre, Kentville, Nova Scotia, B4N 1J5, Canada

² Department of Plant, Food and Environmental Science, Faculty of Agriculture, Dalhousie University, Truro, NS, B2N 5E3, Canada

* Corresponding author, E-mail: Jun.Song@agr.gc.ca

Abstract

Apples are an important dietary source of polyphenols as antioxidants that contribute to human health. Based on total phenolic content and total antioxidant capacity data gathered from a large apple biodiversity collection, we applied untargeted metabolomics analyses to further understand the phenolic composition in select accessions with very low and very high phenolic contents, as well as in commercial cultivars. The liquid chromatography-mass spectrometry (LC-MS) analysis provided both qualitative and quantitative information, with 2,946 features detected using the positive mode of electrospray ionization. We found significant variation in total polyphenols, and strong correlations between total phenolic content and total antioxidant capacity. Polyphenolic compounds significantly associated with each group of apples were quantified, and new features were putatively identified. This study provides new knowledge regarding the specific polyphenols that contribute to the variation in total phenolic content and antioxidant capacity among apples, and new insights into the biochemical regulations of polyphenol biosynthesis and composition.

Citation: Song J, Amyotte B, Yu CHJ, Campbell-Palmer L, Vinqvist-Tymchuk M, et al. 2023. Untargeted metabolomics analysis reveals the biochemical variations of polyphenols in a diverse apple population. *Fruit Research* 3:29 <https://doi.org/10.48130/FruRes-2023-0029>

Introduction

Fruits are an important dietary source of vitamins, minerals, antioxidants, and other biological compounds. Consumption of fresh fruit has been increasing as consumers become more aware of their nutritional value and their role in disease prevention. Apples are one of the most highly produced and consumed fruits in the world, with a gross production value of 430 billion USD in 2021 (www.fao.org/about/who-we-are/departments/statistics-division/en/). It is recognized as a rich source of phytochemicals that provide a wide range of nutrient benefits for human health and wellbeing with reduced risks of some cancers, cardiovascular disease, asthma, and diabetes^[1].

Polyphenolic compounds are major contributors to the phytochemical composition of apples^[2]. Apples and their polyphenols have been found to have very strong antioxidant activity, inhibit cancer cell proliferation, decrease lipid oxidation, and lower cholesterol with anti-inflammatory and anti-neurodegenerative properties^[1,3]. Recent research has expanded our understanding of polyphenols as bio-active compounds in apple fruits. There are various assays for measuring total phenolic content, composition, flavonoids, and antioxidants. Among them, the measurement of total phenolic content by the Folin Ciocalteu method and of antioxidant capacity through ferric reducing antioxidant power have been widely applied^[4].

In general, total phenolic content has been shown to correlate with antioxidant activity^[5]. Five major polyphenolic groups including 16 unique compounds were found in apples^[2,6]. It was reported that dihydrocinnamic acid esters, phloretin

glycosides, and flavan-3-ols were predominant in both flesh and peel tissues, whereas quercetin glycosides were present only in the peels, and cyanidin-3-galactoside was found almost exclusively in red apple peels. Generally, the predominant group of polyphenols was the procyanidins, followed by quercetin glycosides in the peel, and hydroxycinnamic acid esters in the flesh^[2]. More recent studies have revealed additional phenolic compounds uniquely present in apples, such as groups of dihydrochalcones, phloretin, and phloridzin-related compounds^[7,8].

Significant differences in phenolic composition and abundance have been detected among apple cultivars. These differences suggest strong genetic control of apple phenolic composition^[9]. Furthermore, differences in phenolic composition have been found between older and newer varieties^[10], and a recent study suggests that apple breeding has resulted in a significant decline in phenolic content over the past 200 years^[11]. Quantitative trait locus (QTL) mapping and genome-wide association studies (GWAS) have revealed genetic loci of significant effect underlying the phenolic diversity in apple fruit^[12–14]. These genetic mapping efforts align well with recent progress that has been made to identify metabolic control points in apples, for example, in the anthocyanin biosynthesis pathway^[15], and the metabolism of procyanidin^[16]. In addition, the role of UGT-glycosyltransferase in flavonoid metabolism in apples was also characterized^[17]. In contrast to the evidence for strong genetic effects, differences in growing conditions have had marginal to no impact on apple phenolic compounds^[18]. Developmental stages, harvest maturities, postharvest treatments, and storage

conditions also seem to have negligible effect on phenolic composition in comparison with apple genotype^[19].

Different polyphenols may have varied biological activities, including antioxidant activity. In order to better understand the complement of phenolic compounds, high-performance liquid chromatography (HPLC) has routinely been used to characterize total and individual compounds with reference standards^[2]. For better identification and quantitation, liquid chromatography-mass spectrometry (LC-MS) is one of the common investigation tools for metabolomics analysis because of its sensitivity, speed, and broad-range coverage of significant numbers of metabolites^[20]. LC-MS has been successfully applied in apples^[21].

A new approach of data-independent acquisition of MS, an untargeted metabolomics approach, has recently become popular for biological research^[22]. The advantages of untargeted metabolomics include the unbiased and improved detection of both molecular ions and fragment ions, and specifically the detection of low abundance molecular ions. Applying an untargeted metabolomics approach, a group of A-type procyanidin dimers was detected and quantified in yellow raspberries, which would not have been possible in the data-dependent mode^[23]. An untargeted metabolomics approach revealed that knockdown-based modulation of a steroidal glycoalkaloid biosynthetic enzyme in potato leaf could result in metabolic changes in response to the Colorado potato beetle^[24]. In apples, LC-MS based untargeted metabolomics uncovered metabolomic changes in apple fruit during the development of superficial scald, and in response to scald inhibitors^[25].

Apples contain a large number of phenolic compounds, but it is necessary to conduct in-depth compound characterization to understand the relationship of chemical composition to overall antioxidant capacity. With access to a large and diverse population of apple germplasm, it becomes possible to evaluate the range of phenolic compounds in apple fruit. Untargeted metabolomics provides the opportunity to identify and quantify the diverse phenolic compounds present in apples with either very high or very low antioxidant capacities. This approach will help to advance our understanding of the contribution of specific individual polyphenols to overall total phenolic profile in apple, and their associations to the health benefits thereof.

Materials and methods

The apple biodiversity collection

Apple fruits were harvested from the Canadian Apple Biodiversity Collection (ABC), which is located at the Kentville Research and Development Centre, Agriculture and Agri-Food Canada (AAFC) in Nova Scotia, Canada. Detailed information about the ABC was reported previously^[11]. Briefly, the ABC consists total of 1,100 apple accessions. The accessions include diverse *Malus × domestica* and *Malus sieversii* material from the United States Department of Agriculture (USDA) Plant Genetic Resources Unit apple germplasm collection in Geneva, New York, USA; commercial cultivars from the Nova Scotia Fruit Growers' Association Cultivar Evaluation Trial; and advanced breeding material from the AAFC breeding programs^[11]. In 2016, 476 accessions were harvested and evaluated.

Fruit materials

During the 2016 fruiting season, apples included in this study were harvested at full maturity based on a multi-factor evaluation of fruit ripeness, which was determined by changes in ground color, firmness assessed by touch, sweetness assessed by taste, browning of seeds, and starch-iodine content^[11]. For each accession, 15 fruits were harvested and stored at 3 °C with relative humidity > 95% for 30 d before samples were taken and frozen in liquid nitrogen, then ground into powder and stored at -80 °C. Each frozen sample represented a ground mixture of 10–15 chopped fruits with flesh and peel.

Chemicals

The chemicals chlorogenic acid, catechin, epicatechin, cyanidin 3-glucoside, quercetin, quercitrin, phloridzin and rutin, as well as the reagents for Folin-Ciocalteu reagent, were purchased from Sigma/Aldrich Chemical Co. (St. Louis, MO, USA). Procyanidins B₁, B₂, C₁ were purchased from Indofine Chemical Co. (Hillsborough, NJ, USA). Isoquercetin was from Fluka Chemie GmbH (Buchs, Switzerland). All solvents were HPLC grade and purchased from Fisher Scientific (Georgetown, ON, USA).

Determination of total phenolic content and antioxidant capacity

Both the Folin-Ciocalteu assay and a Ferric Reducing Antioxidant Power (FRAP) assay were performed to estimate the total phenolic content (TPC) and total antioxidant capacity (TAC), respectively, with methods as previously described^[3,26]. Each apple sample was extracted and quantified in duplicate. TPC of apple extracts was reported in micromolar of gallic acid equivalents (GAE) per gram fresh weight. The ferric reducing antioxidant powers (FRAP) of apple extracts were reported in micromolar of Trolox equivalents (TE) per gram fresh weight^[27].

TPC and TAC values for the 476 apple accessions were previously reported by Watts et al.^[11]. In this study, a subset of 30 apples was chosen for further analysis based on those TPC and TAC values. The first group included the ten apples with the highest TPC and TAC values, the second group included the ten apples with the lowest TPC and TAC values, and the third group included ten commercial cultivars ([Supplemental Data File S1](#)). The methods described below, and results reported thereafter, correspond to data collected from these 30 apple accessions.

Metabolomics analysis

Metabolites extraction and analysis

Apple tissue (0.5 g) was suspended twice with 0.7 mL 80 % methanol (80 : 20 methanol : water, V/V, with 0.1% formic acid), then vortexed, sonicated for 20 min and centrifuged at 10,000 × g for 10 min (Thermal ICE Microlite, Fisher Scientific Company, Ottawa, Ontario, Canada). The supernatants from the two extractions were then combined and recovered to a new clean tube and dried in a vacuum centrifuge (Thermo Fisher). The pellet was re-suspended in 1 mL solution (10% methanol with 0.1% formic acid) and mixed and sonicated for 15 s and vortexed for 10 s followed by another centrifugation at 10,000 × g for 10 min. The supernatants were transferred to HPLC vials for injection^[25,28]. Two extractions were conducted on each accession to serve as technical replicates.

Data acquisition - mass spectrometry

Chromatography separation was performed on a NanoAcquity UPLC system (Waters Corporation, Milford MA, USA)

Biochemical variations of polyphenols in apples

equipped with a BEH C₁₈ 1.7 μm 1.0 × 100 mm column^[25]. Briefly, binary mobile phases were employed with mobile phase A as 99.9% water with 0.1% (v/v) formic acid and mobile phase B as 99.9% acetonitrile with 0.1% (v/v) formic acid. The gradient was carried out with initial mobile phase A at 95%, decreased to 50% for 13.5 min, decreased to 45% for 1 min, maintained for 3.5 min, and then returned to the initial stage at 95% for another 5.5 min. The total run time was 25 min at flow rate of 45 μL·min⁻¹. Each sample was injected with 1.0 μL using a partial loop mode. During the entire study, the temperatures for the column and sample holder were maintained at 35 °C and 4 °C respectively.

Analysis of apple phenolic compounds was performed on a Synapt XS HDMS mass spectrometer (Waters Corporation, Milford MA, USA) equipped with an electrospray ionization source (ESI) ^[26]. Mass acquisition was set using MS^e (data-independent acquisition) mode in the continuum mode using Masslynx (version 4.2, Waters Corporation, Milford MA, USA). The MS^e acquisition parameters were set as follows: 75–1,000 amu; capillary voltage was 2.0 eV; and a scan speed of 0.4 s including an inter scan delay of 0.02 s. The mass detector was operated in both positive or negative modes at high resolution of 40,000 FWHM. The collision energy was set at 6 eV for the low energy and ramp was 15–35 eV for high energy. The cone voltage was 20 V. Argon was used as collision gas. A lock mass solution (Leucine enkephalin, [M + H]⁺ m/z: 556.2771) was applied in 30-s intervals during the sample acquisition. The mass detector was calibrated with sodium formate over the mass range of 100–1,500 amu prior to analysis. A standard mixture of 12 reference compounds (listed under Chemicals in Materials and methods) was run as quality control and to ensure the consistency of retention time, mass accuracy, and mass fragmentation^[25].

Data processing and metabolites identification

Data processing and analysis

Untargeted LC-MS data was processed using Progenesis QI (Version 3.0, Nonlinear Dynamics, Waters Corporation, Milford MA, USA). Raw files were processed through alignment of the low energy and high energy data, as well as the detection of mass peak (i.e. chromatographic peaks). Results were generated with detected mass, retention time, and integrated intensity values. Data were normalized based on all analyzed features using default parameters in Progenesis QI ^[26].

Metabolites identification

The obtained mass features were set to search against public databases such as Chemspider and FoodDB (<http://foodb.ca/spectra/ms/search>) using Progenesis QI. For identifications, mass tolerance for mass and MS/MS were set as ± 5.0 ppm, respectively. MS/MS fragmentation spectra, mass, and isotope pattern were manually inspected. Only features with a Progenesis QI score > 39 with fragment score and mass error less than ± 5.0 ppm were accepted as putative identifications^[26]. Reference standards mixtures with known compounds (listed under Chemicals in Materials and methods) were injected to verify identified mass features and retention times. The annotation of identified features followed the Metabolomics Standard Initiative guidelines^[29].

Statistical analysis

Mean TPC and TAC were estimated for each group of ten apples through a linear model variance analysis, followed by a

Kruskal-Wallis test with a Bonferroni correction at $\alpha = 0.05$. Means comparisons were performed using the 'Kruskal' function of the R package 'agricolae'^[30]. The effects of apple group on the normalized intensity of each metabolite were evaluated through ANOVA using the Progenesis QI software (Version 3.0). A False Discovery Rate correction (q) was applied at $\alpha = 0.05$ to correct for multiple testing across metabolites. Progenesis QI results for metabolites with significant group effects were exported to excel spreadsheets and further analyzed. A principal components analysis (PCA) was performed using the PCA function of the R package 'FactoMineR' with scaled variables^[31]. Pearson correlations between metabolites, TAC and TPC values were estimated using the 'corr' procedure of the R package 'Hmisc'^[32] with a Bonferroni correction applied at $\alpha = 0.05$.

Results

Characterization of total phenolic content (TPC) and antioxidant capacity (TAC)

Significant variations in the total phenolic content (TPC) and total antioxidant capacity (TAC) were found among the three groups of 10 apples: those with the highest TPC and TAC (top group), those with the lowest TPC and TAC (bottom group), and the 10 commercial cultivars (commercial group) (Fig. 1a, Supplemental Data File S1). Across these groups, TPC was significantly positively correlated with TAC ($r = 0.98$, $p < 0.05$) (Fig. 1b); this correlation was also found in a previous study of 476 apples^[11,33]. The top group had average TPC and TAC values of 16.4 GAE·g⁻¹·FW and 117.5 μmol·TE·g⁻¹·FW respectively, while the low group had averages of 0.6 GAE·g⁻¹·FW and 1.2 μmol·TE·g⁻¹·FW, respectively (Supplemental Data File S1). The 10 commercial genotypes included were: 'Empire', 'Elstar', 'Sweetango', 'Ambrosia', 'Honeycrisp', 'Jonagold', 'Gala', 'Red Delicious', 'McIntosh', and 'Reinette Russet', and they were found to have an average TPC value of 3.04 GAE·g⁻¹·FW and TAC value of 3.39 μmol·TE·g⁻¹·FW (Supplemental Data File S1). These results demonstrate the substantial range that exists in TPC and TAC in diverse apples, and demonstrates that most popular commercial cultivars have relatively low total phenolic content and antioxidant capacity. For example, when the 'Honeycrisp' cultivar was compared with others in terms of TPC and TAC, it showed 8- and 12-fold lower values compared with the top group, respectively (Fig. 1b). The top TPC and TAC group contained a combination of specialty European cultivars and wild accessions from Kazakhstan (Supplemental Data File S1).

Identification of phenolic compounds in apple accession groups

Taking advantage of untargeted LC-MS analysis, many known and unknown features were detected, and the corresponding abundances based on the normalized MS detector response were quantified. A total of 2,946 masses/features were detected under the positive mode (Supplemental Data File S2). The relative abundances of these features were compared among accessions from the three groups of 10 accessions (top TPC, bottom TPC, and commercial cultivars), and a total of 1,849 features were found to be significantly different between apple groups ($q < 0.05$) (data not shown). Among the features with significant differences, 29 were putatively identified through the use of reference standards and the analysis of

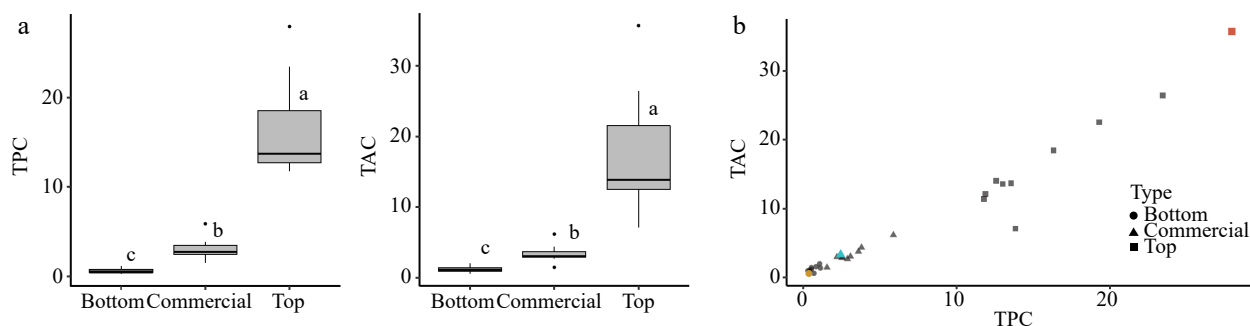


Fig. 1 Comparison of high total phenolic content (TPC) and high total antioxidant capacity (TAC) in the 10 apples with the highest (Top) vs. lowest (Bottom) TPC and TAC values, alongside 10 commercial apple cultivars (Commercial). (a) Boxplots for TPC and TAC of each category showing significant differences between category means according to a Kruskal-Wallis test ($\alpha = 0.05$). (b) Scatterplot showing significant correlation between TPC and TAC ($r = 0.98$, $p < 0.05$). Cultivars 'Zestar' (yellow dot), 'Honeycrisp' (blue triangle) and 'Marechal' (red square) are highlighted to illustrate the range of responses.

MS/MS fragmentation spectra, mass and isotope pattern (Supplemental Data File S1). A principal component analysis of peak area responses for these 29 putative compounds showed no clear pattern of clustering or separation based on the apple groups (Supplemental Data File S1).

Quantitative differences in phenolic compounds between accession groups

Twenty-two putative phenolic compounds were found to have significantly higher peak area responses in the apples from the top TPC group than those from the other groups ($q < 0.05$) (Fig. 2, Supplemental Data File S1). The average peak area responses for this group ranged from 52 (methylarbutin) to 41,900 (procyanidin B₂). The most highly expressed putative compounds were phloretin like (peak area response 10,400), procyanidin C₂ (10,500), procyanidin B₁ (12,500), epicatechin (17,900), procyanidin C₁ (28,363), and procyanidin B₂ (41,900). Among these, epicatechin and procyanidins B₁, C₁, and C₂ were all found to be at least ten-fold or more abundant in the top TPC group than the bottom TPC group (Fig. 2, Supplemental Data File S1).

In contrast, there were only four putative compounds found to have significantly higher peak area responses in the bottom TPC group than in the other groups ($q < 0.05$) (Fig. 3, Supplemental Data File S1). These included cyanidin-3-galactoside, chlorogenic acid, feruloyl glucose, and quercetin. The most highly expressed putative compounds from the bottom TPC group were cyanidin-3-galactoside (peak area response 11,622) and chlorogenic acid (13,279), however neither were substantially more abundant in the bottom TPC group than in the other groups. All four compounds were detected in some amount in the other two groups; the minimum abundance was for feruloyl glucose in the top TPC group (peak area response 85) (Fig. 3, Supplemental Data File S1).

Lastly, there were three putative compounds with significantly higher peak area responses in the commercial apple group than in the other two groups (Fig. 4, Supplemental Data File S1). These were a feruloyl-quinic acid, a quercetin glucoside, and reynoutrin. Reynoutrin had the largest abundance (peak area response 3,660), while the other two were below 1,000. Reynoutrin was found to be 1.5-fold more abundant in the commercial group than in the bottom TPC group, and 1.2-fold more abundant than in the top TPC group. Once again, all three compounds were also detected in the other two groups (Fig. 4, Supplemental Data File S1).

A correlation analysis was performed for the 29 putative compounds that had significant peak area response differences between apple group (Fig. 5, Supplemental Data File S1). Strong positive correlations ($r = 0.80$ to 1.00) were observed between procyanidins B₁, B₂, C₁ and C₂, catechin, epicatechin, cinchonain la, arecatannin B₁, hydroxynaringenin, and shikimic acid. Procyanidin A₂ was moderately correlated ($r = 0.65$ to 0.80) with several compounds in this grouping, and strongly correlated with catechin gallate. Hesperitin and 4-coumaroyl-shikimic acid were strongly correlated with each other, and were moderately correlated with methylarbutin, cinchonain la, hydroxynaringenin, arecatannin B₁, and procyanidin C₂. Phloretin and phloridzin were moderately correlated. There were no significant correlations among any of the putative compounds associated with the bottom TPC or commercial apple groups. In addition, none of the individual compounds from the bottom TPC or commercial groups were found to be significantly correlated with TPC or TAC (Fig. 5, Supplemental Data File S1).

Discussion

Apples have gained significant attention as nutritious fruit rich in dietary antioxidants, which include polyphenols and other bioactive compounds. Apple polyphenols play an important role in physiological functions related to human health^[1]. Total phenolic compounds have been summarized as total phenolics, total hydroxycinnamates, total flavonols, total procyanidins, as well as total anthocyanins^[6]. Understanding the polyphenolic composition of apple biodiversity collections provides a pathway toward the genetic improvement of apples by targeting bioactive compounds^[13].

Significant research has been conducted to characterize the polyphenols as well as antioxidant capacity in apples with various chemical methods and enzymatic assays^[4]. It has been suggested that multi-faceted approaches such as measuring TPC and TAC, should be applied to effectively evaluate natural products as dietary sources of antioxidants^[5]. The majority of the research on phenolic compounds in apples has been conducted on selected cultivars with a limited numbers of samples. These types of studies are useful to compare the effects of certain growing conditions on one or more apple cultivars^[2,34] but do not necessarily investigate the variability across diverse germplasm. However, a recent study investigated the relative abundances of several dihydrochalcones in

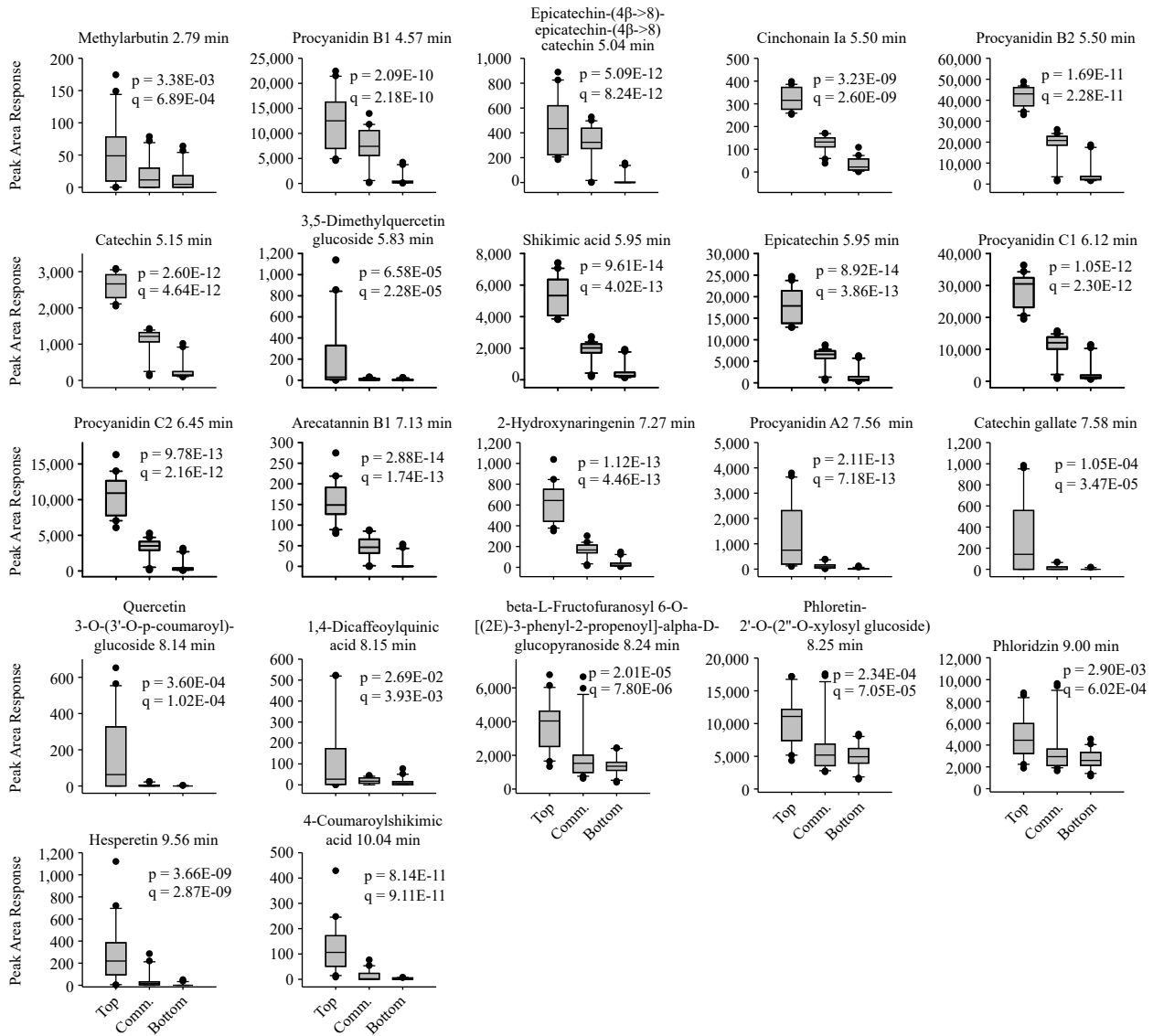


Fig. 2 Boxplots of apple phenolic compounds significantly more abundant in the top TPC and TAC group as compared with the bottom and commercial groups. The statistics p and q represent the unadjusted and multiple means corrected significance values, respectively, for a variance analysis comparing the effects of apple group on phenolic compound content.

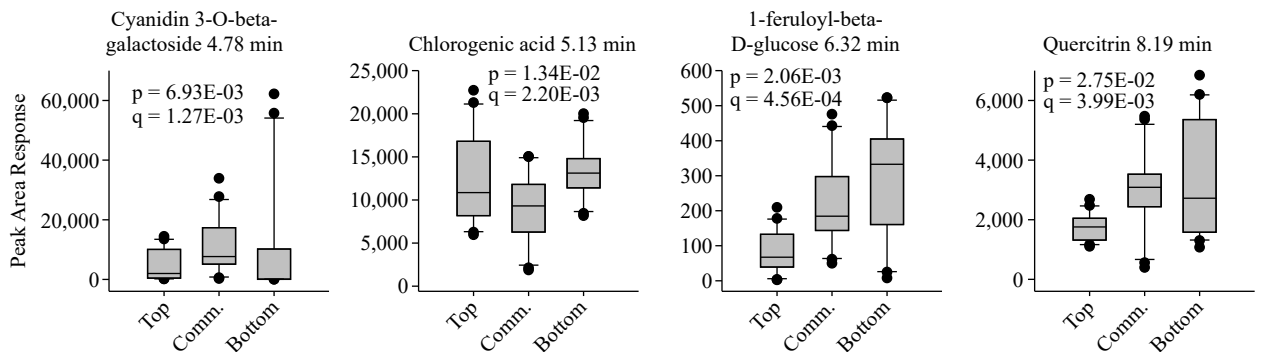


Fig. 3 Boxplots of apple phenolic compounds significantly more abundant in the bottom TPC and TAC group as compared with the top and commercial groups. The statistics p and q represent the unadjusted and multiple means corrected significance values, respectively, for a variance analysis comparing the effects of apple group on phenolic compound content.

140 cultivars^[7]. It was subsequently recognized that phenolic content is highly heritable in apple, and that genetic diversity seems to be the most significant factor determining the

phenolic content across apples^[13]. This research led the way toward exploring the variability in a larger apple collections, such as Canada's Apple Biodiversity Collection (ABC).

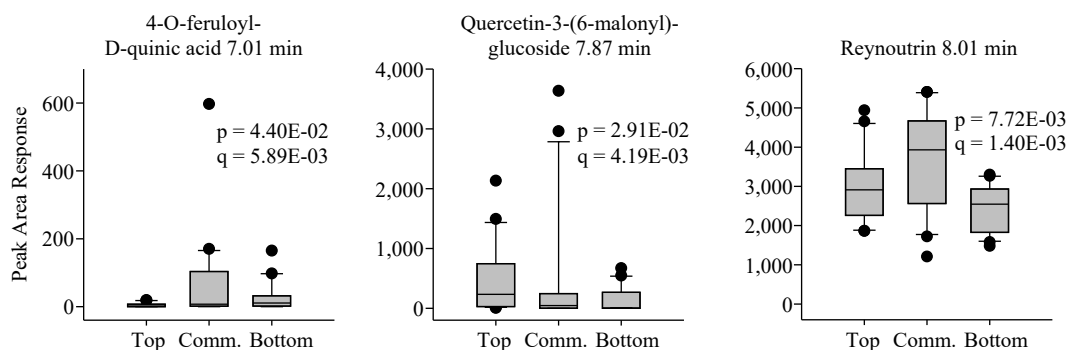


Fig. 4 Boxplots of apple phenolic compounds significantly more abundant in the commercial TPC and TAC group as compared with the top and bottom groups. The statistics p and q represent the unadjusted and multiple means corrected significance values, respectively, for a variance analysis comparing the effects of apple group on phenolic compound content.

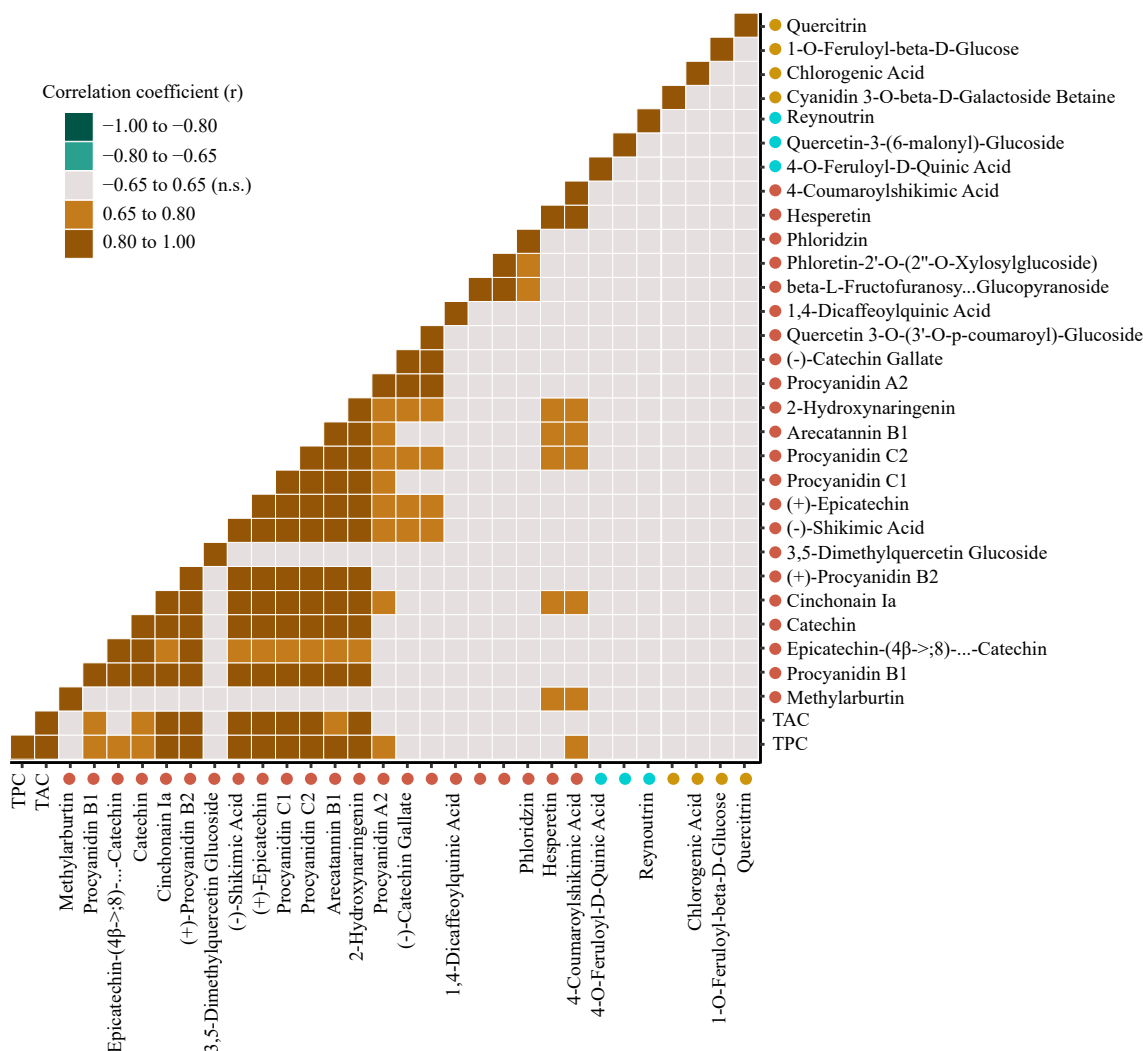


Fig. 5 Pearson correlations between the phenolic compounds significantly associated with top total TPC and TAC apples (red dots), bottom TPC and TAC apples (yellow dots), or commercial apple cultivars (blue dots). Pearson correlation coefficients significant at $\alpha = 0.05$ are shown.

We previously took advantage of the ABC to assess 476 accessions, determining the total phenolic content (TPC) using the Folin-Ciocalteu assay, and measuring the total antioxidant capacity (TAC) by the FRAP assay^[11]. Significant differences in TPC and TAC were reported across the whole population^[11], which allowed us to identify the 20 accessions with the highest

and lowest TPC and TAC values, as well as 10 commercial cultivars with a range of TPC and TAC values. Overall, the commercial cultivars in this study had 80% less TPC than the accessions with the highest TPC values. Among the tested commercial cultivars, 'Reinette Russet' and 'Red Delicious' had the highest concentrations of total phenolic compounds, while 'Empire'

Biochemical variations of polyphenols in apples

had the lowest. Our results confirmed the previous report that 'Empire' had lower TPC than 'Red Delicious'^[2].

We assessed the distribution of maturity traits including harvest date, days to harvest, soluble solids content, and acidity at harvest for the 30 apples included in this study, and found that the traits were generally evenly distributed without any major skewedness which would suggest an incorrect assessment of ripeness (data not shown). We also performed a correlation analysis of the 30 apples, including harvest date and the phenolic compounds. No significant correlation was found either on flowering date, days to ripen, harvest date, fruit soluble solids content, and nor fruit acidity ($p > 0.05$, data not shown). Similarly, TPC and TAC were not significantly correlated with any of these traits. These results corresponded with published results from the larger Apple Biodiversity Collection in which TPC was not correlated with the phenology or harvest traits^[11]. These results also suggest that the regulation of phenolic biosynthesis is cultivar-specific, and is not controlled by a universal phenology-based ripening signal in apple.

The significant biodiversity of apples provides an invaluable genetic resource to further characterize the chemical contributors to TPC and TAC. We applied untargeted LC-MS metabolomic analysis to identify and quantify compounds in apples with very high and very low TPC and TAC values, as well as in commercial cultivars. Significant differences in composition and content were found across the different groups. Among the three groups, groups of phenolic compounds such as catechin, epicatechin, and their gallate derivatives, procyanidins, as well as a group of dihydrochalcones, including phloretin and phloridzin, were all significantly higher in the top TPC group. These results imply that these classes of compounds contribute positively to an overall high value of TPC and TAC in apples.

Catechin, epicatechin as well as procyanidins were previously found to be the major components of the phenolic profile in apples^[10]. Catechin and epicatechin are also precursors of procyanidins, through leucoanthocyanidin reductase (LAR1) for catechin and anthocyanidin reductase (ANR) for epicatechin^[35]. Genetic studies have identified a major QTL around LAR1 on chromosome 16 that likely controls the accumulation of flavanols and procyanidins in commercial apples^[13,36], as well as in cider apples^[14]. Therefore, from total antioxidant and phenolic perspectives, high abundances of catechin, epicatechin and related procyanidins contribute to high antioxidant activity.

It is worth noting that russeted accessions, such as 'Reinette Russet', were outliers with respect to their high concentrations of phloridzin and phloretin, which coincides with previous reports^[8,13]. We also found high abundances of dihydrochalcones in the top TPC group from the present study. Recently, a 3-hydroxylase that convert the phloretin to 3-hydroxyphloretin has been identified in some *Malus* species^[37]. These result indicate that incorporating selections with high abundance of phloretin-related compounds in an apple breeding program may result in an improvement in TPC as an element of nutritional quality.

Interestingly, despite having been identified through their low TPC content and low TAC, the bottom apple group did contain significantly higher amounts of a few compounds in comparison with the top and commercial groups. These included chlorogenic acid and quercitrin. Chlorogenic acid is one of the most prevalent phenolic compounds in apples and belongs to the hydroxycinnamic acids class of phenolic

compounds^[9]. It has been identified in apples in both flesh and peel tissues, however, it showed lower antioxidant capacity than cyanidin-3-galactoside and procyanidin B₁ and B₂^[38].

The biosynthesis of chlorogenic acid is controlled by shikimate O-hydroxycinnamoyl transferase and quinate O-hydroxycinnamoyl transferase (HCT/HQT) (EC:2.3.1.133, KEGG pathways), which play a critical role in phenylpropanoid biosynthesis^[39]. At the genetic level, the HCT/HQT enzyme pair was identified as a good candidate target for controlling the chlorogenic acid in cider apples^[14]. Our previous research indicated the possible connection between HCT/HQT and chlorogenic acid at a locus of interest on chromosome 17^[13]. New research recently reported that chlorogenic acid in apples is primarily synthesized via the caffeoyl-CoA and quinic acid route, with a positive correlation to phenylalanine ammonia-lyase 3 (PAL3) and HQT^[40]. The same group also reported that cultivated apples have higher chlorogenic acid than wild apples.

Another interesting finding in our study was that apples in the commercial group showed significantly lower TPC and TAC than the top TPC group, but contained relatively higher amounts of quercetin glycosides. Quercetin and its derivative compounds are in the flavonol class of phenolic compounds; they are formed from dihydrokaempferol and dihydroquercetin through flavonol synthase (FLS) to kaempferol and quercetin, respectively (KEGG pathways). Quercetin can be converted to quercetin glycosides such as reynoutrin via UDP-glucose flavonoid-3-O-glucosyl transferase (UFGT) (KEGG pathways). A significantly higher amount of quercetin derivatives were found in skin tissue of 'Hetlian' and 'Devonshire Quarrenden' compared to 'Royal Gala'; the two heritage apples also showed a corresponding increase in expression of UFGT and flavonol synthase genes^[16]. Scab-resistant cultivars had a significantly higher concentration of quercitrin compared to scab susceptible cultivars, and a significant GWAS hit for quercitrin occurred 94 kb upstream of a UDP-glycosyltransferase gene on chromosome 1^[13]. The high abundances of quercetin derivatives that we found in newer cultivars such as 'Gala', 'Honeycrisp', and 'Sweetango' may in part be the consequence of commercial breeding efforts over the years aimed at disease resistance.

This study found that while select phenolic compounds were overexpressed in apples with low TPC and in commercial apples, apples with high TPC values had overall high levels of polyphenols across phenolic compound classes (Fig. 6). It therefore appears that the biosynthesis of phenolic compounds could be influenced at a number of critical points in the pathway. For example, apples with high TPC values were abundant in dihydrochalcones as well as in complex flavonoids, whereas apples with low TPC values were limited to accumulating chlorogenic acid or quercetin derivatives. The action of select enzymes acting at downstream stages in the dihydrochalcone and flavonoid biosynthesis pathways could have significant influence of total phenolic content and total antioxidant capacity in apple.

The biosynthesis of polyphenols in the flavonoid class depends on the action of a flavonoid 3'-hydroxylase (F3'H) acting on precursors derived from 4-coumaroyl CoA or caffeoyl CoA (Fig. 6; KEGG pathways). It was reported that apples lack a functional flavonoid 3'5' hydroxylase, and therefore are unable to hydroxylate positions 3 and 5 of the b-ring to form delphinidin-based compounds^[41]. Another group subsequently

detected the expression of an F3'H which was associated with flavonol, procyanidin, and anthocyanin biosynthesis in apple. This enzyme is proposed to control the synthesis of dihydroquercetin and quercetin from eriodictoyl or dihydrokaempferol, and kaempferol, respectively. It therefore likely plays an important role in the biosynthesis and metabolism of catechins, epicatechins, and procyanidins, together with leucoanthocyanidin reductase (LAR), anthocyanidin reductase (ANR), anthocyanidin synthase (ANS)^[42].

Downstream of F3'H and upstream of LAR, ANS, and ANR, the genetic control for flavonol synthase (FLS) and flavanone-4-reductase (F4R) may represent a branch point and regulating steps in the flavonoid synthesis pathways of high versus low TPC and TAC apple varieties (Fig. 6). Our data suggest that F4R activity may be required for apples to accumulate high levels of catechins, epicatechins, and procyanidins, which are associated with high antioxidant capacity. Our data also support the

potential importance of dihydrochalcone 2-O-glucosyltransferase (D2'GT) in driving the accumulation of phloretin glycosides, which were similarly abundant in apples with high TPC and TAC (Fig. 6). Although suggestive of potential biological regulation, these findings are putative associations and require additional research to support the hypothesis.

The use of HPLC and LC-MS to measure phenolic compounds and their relation to antioxidant capacity has widely been reported^[2,43]. In the present study, we applied LC-MS based untargeted metabolomics to identify and quantify the compounds contributing to antioxidant capacity in apple. This untargeted metabolomics approach (so called data-independent acquisitions, DIA) showed advantages for unbiased data collection and acquisition, since features with low abundance could be detected and quantified. Both primary metabolites (e.g. sugars, fatty acids, and amino acids) and secondary metabolites (flavonols and phenolic acids) can be identified

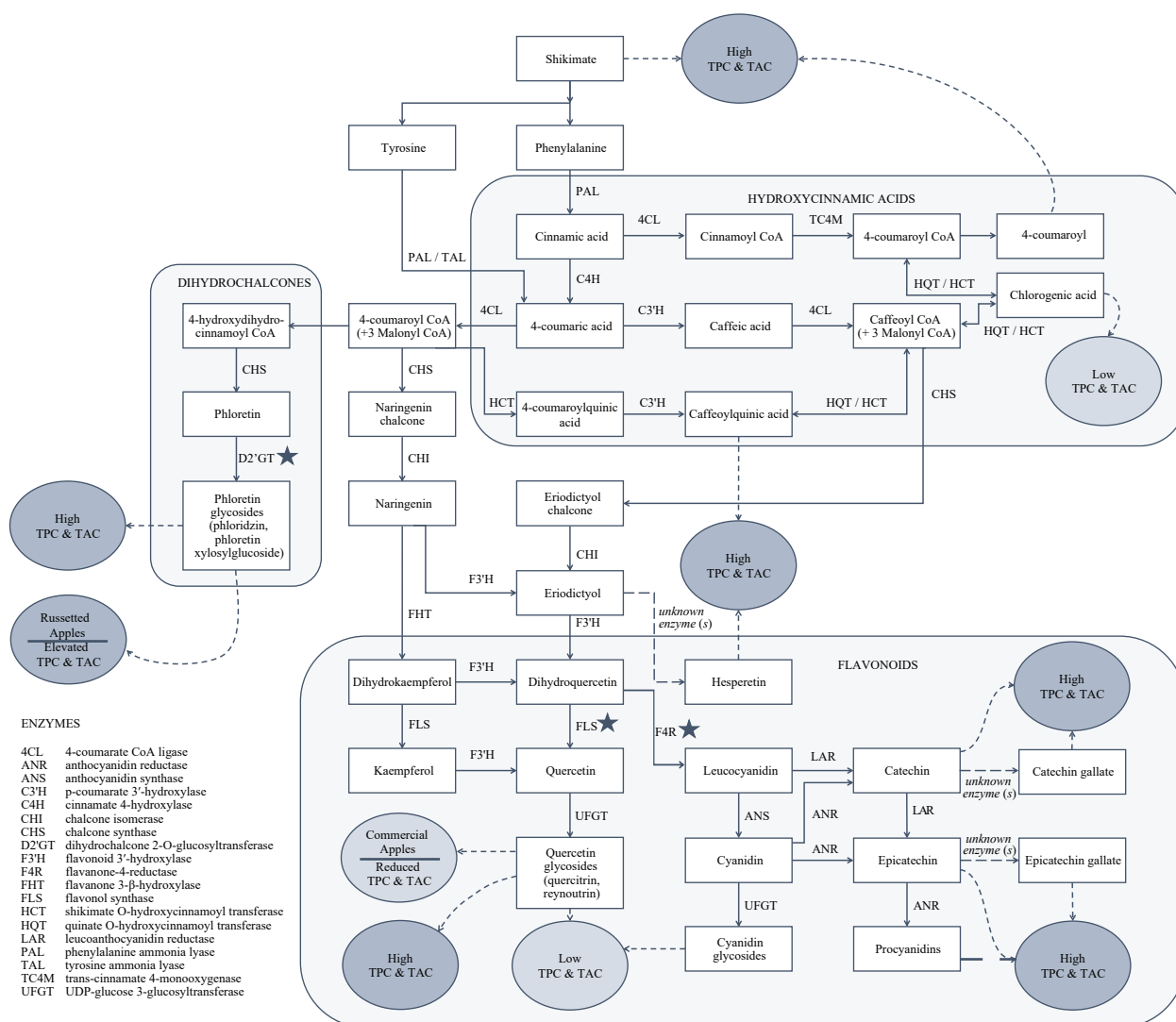


Fig. 6 Diagram of the main phenolic compounds and classes detected in apples with high total phenolic content (TPC) and high total antioxidant capacity (TAC) vs apples with low TPC and TAC. Biosynthetic pathways are briefly depicted where known according to previous reports, including putative enzymes. The main phenolic compounds found primarily in commercial apple cultivars and in russetted apples are also listed. Starred (★) enzymes may represent key branch points between high vs low TPC and TAC apples. The diagram is modified from Verdu et al. (diagram structure) ^[14], Liao et al. (LAR) ^[35], Henry-Kirk et al. (LAR and ANR) ^[16], Liao et al. ^[40] and Hoffmann et al. ^[39] (HCT / HQT), KEGG pathways^[47].

Biochemical variations of polyphenols in apples

and quantified using this method^[44]. Its application has been reported in *Brassicaceae*^[28], tomatoes^[45], and wine grapes^[46]. However, the challenge that remains is that not all features can be successfully identified due to limitations in available databases for species-specific phenolic compounds. We found this challenge to be substantial, considering that we were only able to confidently identify 29 out of the 1849 metabolomic features (1.5%) that were significantly different between the apple groups we studied. Despite these limitations, our results provide a further understanding of the polyphenolic composition and diversity in apples, and highlight the associations of select compounds with overall TPC and TAC. Specific compounds such as phloridzin and other phloretin glycosides could serve as targets for breeding apples with improved health value, either through crosses with some of the high TPC accessions evaluated here, or through gene editing tools targeting the putative enzymes controlling their biosynthesis in commercial cultivars. To gain more fundamental knowledge on genetic control of apple phenolic compounds, a larger study of the association between phenolic compounds and genome-wide variation is underway and will be published separately in the near future.

Conclusions

Apples are a widely consumed fruit and a rich source of polyphenolic compounds, which play an important role in physiological functions related to human health and well-being. In this study, employing an untargeted LC-MS approach to phenotype a diverse apple germplasm collection, we identified many putative phytochemicals in apple fruits and probed the relationships between 29 individual phenolic compounds, total phenolic content, and antioxidant capacity. We found significant variation between genotypes, with more than ten-fold differences in compound abundances between the apples with the highest total phenolic contents and those with the lowest. Our work also shows that while popular commercial cultivars have relatively lower phenolic contents, they still contain high concentrations of some well-studied health-associated compounds. Together these results suggest complex dynamics in apple polyphenol biosynthesis and antioxidant activity that merit further study. Overall, this research reveals the role and contribution of specific polyphenols and compound classes toward the total phenolic content and antioxidant capacity of apple. The results provide new insights into the possible genetic control and regulation of polyphenols in apple, and offer potential targets for breeding or gene editing.

Acknowledgments

This research was supported by A-Base funding (LOI-1767) from Agriculture and Agri-Food Canada (JS), and by the National Sciences and Engineering Research Council (NSERC) of Canada (HPVR). The authors gratefully acknowledge the Farm Services team at AAFC-Kentville for establishing and maintaining the orchards studied. The authors would like to thank Dr. Sean Myles for critical review of the manuscript and Sophie Watts for providing information on ABC accessions.

Conflict of interest

The authors declare that they have no conflict of interest.

Supplementary Information accompanies this paper at (<https://www.maxapress.com/article/doi/10.48130/FruRes-2023-0029>)

Dates

Received 6 July 2023; Accepted 17 August 2023; Published online 3 November 2023

References

- Boyer J, Liu RH. 2004. Apple phytochemicals and their health benefits. *Nutrition Journal* 3:5
- Tsao R, Yang R, Young JC, Zhu H. 2003. Polyphenolic profiles in eight apple cultivars using high-performance liquid chromatography (HPLC). *Journal of Agricultural and Food Chemistry* 51:6347–53
- Huber GM, Rupasinghe HPV. 2009. Phenolic profiles and antioxidant properties of apple skin extracts. *Journal of Food Science* 74:C693–C700
- Sethi S, Joshi A, Arora B, Bhowmik A, Sharma RR, et al. 2020. Significance of FRAP, DPPH, and CUPRAC assays for antioxidant activity determination in apple fruit extracts. *European Food Research and Technology* 246:591–98
- López-Alarcón C, Denicola A. 2013. Evaluating the antioxidant capacity of natural products: a review on chemical and cellular-based assays. *Analytica Chimica Acta* 763:1–10
- Tsao R, Yang R. 2003. Optimization of a new mobile phase to know the complex and real polyphenolic composition: towards a total phenolic index using high-performance liquid chromatography. *Journal of Chromatography A* 1018:29–40
- Zhou K, Hu L, Liu B, Li Y, Gong X, et al. 2018. Identification of apple fruits rich in health-promoting dihydrochalcones by comparative assessment of cultivated and wild accessions. *Scientia Horticulturae* 233:38–46
- Gutierrez BL, Zhong GY, Brown SK. 2018. Genetic diversity of dihydrochalcone content in *Malus* germplasm. *Genetic Resources and Crop Evolution* 65:1485–502
- Awad MA, de Jager A, van Westing LM. 2000. Flavonoid and chlorogenic acid levels in apple fruit: characterisation of variation. *Scientia Horticulturae* 83:249–63
- Wojdyło A, Oszmiański J, Laskowski P. 2008. Polyphenolic compounds and antioxidant activity of new and old apple varieties. *Journal of Agricultural and Food Chemistry* 56:6520–30
- Watts S, Migicovsky Z, McClure KA, Yu CHJ, Amyotte B, et al. 2021. Quantifying apple diversity: a phenomic characterization of Canada's Apple Biodiversity Collection. *PLANTS, PEOPLE, PLANET* 3:747–60
- Ali Khan S, Chibon PY, de Vos RCH, Schipper BA, Walraven E, 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* 63:2895–908
- McClure KA, Gong Y, Song J, Vinqvist-Tymchuk M, Campbell Palmer L, et al. 2019. Genome-wide association studies in apple reveal loci of large effect controlling apple polyphenols. *Horticulture Research* 6:107
- Verdu CF, Guyot S, Childebrand N, Bahut M, Celton JM, et al. 2014. QTL analysis and candidate gene mapping for the polyphenol content in cider apple. *PLoS ONE* 9:e107103
- Takos MA, Robinson PS, Walker RA. 2006. Transcriptional regulation of the flavonoid pathway in the skin of dark-grown 'Cripps' Red' apples in response to sunlight. *The Journal of Horticultural Science and Biotechnology* 81:735–44

16. Henry-Kirk RA, McGhie TK, Andre CM, Hellens RP, Allan AC. 2012. Transcriptional analysis of apple fruit proanthocyanidin biosynthesis. *Journal of Experimental Botany* 63:5437–50
17. Elejalde-Palmett C, Billet K, Lanoue A, De Craene JO, Glévarec G, et al. 2019. Genome-wide identification and biochemical characterization of the UGT88F subfamily in *Malus x domestica* Borkh. *Phytochemistry* 157:135–44
18. Yuri JA, Maldonado FJ, Razmilic I, Neira A, Quilodran Á, et al. 2012. Concentrations of total phenols and antioxidant activity in apple do not differ between conventional and organic orchard management. *Journal of Food, Agriculture & Environment* 10:207–16
19. Kevers C, Pincemail J, Tabart J, Defraigne JO, Dommès J. 2011. Influence of cultivar, harvest time, storage conditions, and peeling on the antioxidant capacity and phenolic and ascorbic acid contents of apples and pears. *Journal of Agricultural and Food Chemistry* 59:6165–71
20. Alseekh S, Fernie AR. 2018. Metabolomics 20 years on: what have we learned and what hurdles remain? *The Plant Journal* 94:933–42
21. Bars-Cortina D, Macià A, Iglesias I, Romero MP, Motilva MJ. 2017. Phytochemical profiles of new red-fleshed apple varieties compared with traditional and new white-fleshed varieties. *Journal of Agricultural and Food Chemistry* 65:1684–96
22. Schrimpe-Rutledge AC, Codreanu SG, Sherrod SD, McLean JA. 2016. Untargeted metabolomics strategies—challenges and emerging directions. *Journal of the American Society for Mass Spectrometry* 27:1897–905
23. Carvalho E, Franceschi P, Feller A, Herrera L, Palmieri L, et al. 2016. Discovery of A-type procyanidin dimers in yellow raspberries by untargeted metabolomics and correlation based data analysis. *Metabolomics* 12:144
24. Paudel JR, Davidson C, Song J, Maxim I, Aharoni A, et al. 2017. Pathogen and pest responses are altered due to RNAi-mediated knockdown of *GLYCOALKALOID METABOLISM 4* in *Solanum tuberosum*. *MPMI* 30:876–85
25. Gong Y, Song J, Campbell Palmer L, Vinqvist-Tymchuk M, Fillmore S, et al. 2021. Tracking the development of the superficial scald disorder and effects of treatments with diphenylamine and 1-MCP using an untargeted metabolomic approach in apple fruit. *Food Chemistry: Molecular Sciences* 2:100022
26. Singleton VL, Orthofer R, Lamuela-Raventós RM. 1999. Analysis of total phenols and other oxidation substrates and antioxidants by means of folin-ciocalteu reagent. *Methods in Enzymology* 299:152–78
27. Benzie IFF, Strasin JJ. 1999. Ferric reducing/antioxidant power assay: direct measure of total antioxidant activity of biological fluids and modified version for simultaneous measurement of total antioxidant power and ascorbic acid concentration. *Methods in Enzymology* 299:15–27
28. De Vos RCH, Schipper B, Hall RD. 2012. High-performance liquid chromatography–mass spectrometry analysis of plant metabolites in *Brassicaceae*. In *Plant Metabolomics: Methods and Protocols*, eds. Hardy NW, Hall RD. Totowa, NJ: Humana Press. 111–28. https://doi.org/10.1007/978-1-61779-594-7_8
29. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, et al. 2007. Proposed minimum reporting standards for chemical analysis. *Metabolomics* 3:211–21
30. De Mendiburu F. 2021. Package 'agricolae' version 1.3-5. <https://cran.r-project.org/web/packages/agricolae/index.html>.
31. Lê S, Josse J, Husson F. 2008. FactoMineR: an R Package for multivariate analysis. *Journal of Statistical Software* 25:1–18
32. Harrell FE, Dupont C. 2021. Harrell Miscellaneous 'Hmisc' R Package v. 4.6. <https://rdrr.io/github/harrelfe/Hmisc/>
33. Yu CHJ, Migicovsky Z, Song J, Rupasinghe HPV. 2023. (Poly)phenols of apples contribute to in vitro antidiabetic properties: assessment of Canada's Apple Biodiversity Collection. *PLANTS, PEOPLE, PLANET* 5:225–40
34. Thompson-Witrick KA, Goodrich KM, Neilson AP, Hurley EK, Peck GM, et al. 2014. Characterization of the polyphenol composition of 20 cultivars of cider, processing, and dessert apples (*Malus x domestica* Borkh.) grown in Virginia. *Journal of Agricultural and Food Chemistry* 62:10181–91
35. Liao L, Vimolmangkang S, Wei G, Zhou H, Korban SS, et al. 2015. Molecular characterization of genes encoding leucoanthocyanidin reductase involved in proanthocyanidin biosynthesis in apple. *Frontiers in Plant Science* 6:243
36. Chagné D, Krieger C, Rassam M, Sullivan M, Fraser J, et al. 2012. QTL and candidate gene mapping for polyphenolic composition in apple fruit. *BMC Plant Biology* 12:12
37. Miranda S, Lagrèze J, Knoll AS, Angeli A, Espley RV, et al. 2022. De novo transcriptome assembly and functional analysis reveal a dihydrochalcone 3-hydroxylase (DHC3H) of wild *Malus* species that produces sieboldin *in vivo*. *Frontiers in Plant Science* 13:1072765
38. Tsao R, Yang R, Xie S, Sockovie E, Khanzadeh S. 2005. Which polyphenolic compounds contribute to the total antioxidant activities of apple? *Journal of Agricultural and Food Chemistry* 53:4989–95
39. Hoffmann L, Besseau S, Geoffroy P, Ritzenthaler C, Meyer D, et al. 2004. Silencing of hydroxycinnamoyl-coenzyme A shikimate/quinate hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis. *The Plant Cell* 16:1446–65
40. Liao L, Zhang W, Zhang B, Cai Y, Gao L, et al. 2021. Evaluation of chlorogenic acid accumulation in cultivated and wild apples. *Journal of Food Composition and Analysis* 104:104156
41. Forkmann G. 1991. Flavonoids as flower pigments: the formation of the natural spectrum and its extension by genetic engineering. *Plant Breeding* 106:1–26
42. Han Y, Vimolmangkang S, Soria-Guerra RE, Rosales-Mendoza S, Zheng D, et al. 2010. Ectopic expression of apple *F3'H* genes contributes to anthocyanin accumulation in the Arabidopsis *tt7* mutant grown under nitrogen stress. *Plant Physiology* 153:806–20
43. Petkovska A, Gjamovski V, Stanoeva JP, Stefova M. 2017. Characterization of the polyphenolic profiles of peel, flesh and leaves of *Malus domestica* cultivars using UHPLC-DAD-HESI-MSⁿ. *Natural Product Communications* 12:35–42
44. Chaleckis R, Meister I, Zhang P, Wheelock CE. 2019. Challenges, progress and promises of metabolite annotation for LC–MS-based metabolomics. *Current Opinion in Biotechnology* 55:44–50
45. Rogachev I, Aharoni A. 2012. UPLC-MS-based metabolite analysis in tomato. In *Plant Metabolomics. Methods in Molecular Biology*, eds. Hardy N, Hall R. Totowa, NJ: Humana Press. 860: 129–44. https://doi.org/10.1007/978-1-61779-594-7_9
46. Arbulu M, Sampedro MC, Gómez-Caballero A, Goicolea MA, Barrio RJ. 2015. Untargeted metabolomic analysis using liquid chromatography quadrupole time-of-flight mass spectrometry for non-volatile profiling of wines. *Analytica Chimica Acta* 858:32–41
47. Kanehisa M, Goto S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research* 28:27–30



Copyright: © 2023 by the author(s). Published by Maximum Academic Press, Fayetteville, GA. This article is an open access article distributed under Creative Commons Attribution License (CC BY 4.0), visit <https://creativecommons.org/licenses/by/4.0/>.