Composition and evolution of oligomeric proanthocyanidin–malvidin glycoside adducts in commercial red wines

Juuso Erik LAITILA

Natural Chemistry Research Group, Department of Chemistry, University of Turku, FI-20014 Turku, Finland

E-mail: juerlai@utu.fi, Tel.: +358504329070

Abstract:

Studies regarding composition and evolution of oligomeric proanthocyanidin–anthocyanin adducts in red wines have often focused only on a limited number of small dimers. Here, a group-specific liquid chromatography–tandem mass spectrometry method was utilized to measure two-dimensional (2D) chromatographic fingerprints of three different types of oligomeric adducts in commercial red wines. A new protocol was developed to visualize and summarize the chromatographic data. The 2D fingerprints showed how the compositions of the oligomeric adducts had typically only minor differences between wine varieties in young wines, excluding the Pinot Noir wines. Major quantitative differences were found between the wine varieties despite the lack of major compositional differences. The evolution of the concentrations differed between the three structural sub-groups, while similar general patterns were observed in the compositional evolution. Via statistical modelling, several characteristics in the polyphenolic starting material composition were tentatively suggested to affect the formation of the oligomeric adducts.

Keywords:

Anthocyanins, chromatographic fingerprints, pigments, proanthocyanidins, tannins

1. Introduction

Pigment-wise, red wines are complex polyphenolic mixtures. Anthocyanins, which are pigments extracted to wine from grape skins, are still relatively few in number in red wines (He et al., 2012) but during wine making and aging the complexity of the overall pigment composition increases rapidly. The anthocyanins react with other wine components to yield pyranoanthocyanins, pinotins and other first generation adducts (de Freitas & Mateus, 2011; He et al., 2012). Some of the first generation adducts can react to form second generation adducts, such as portisins and oxovitisins (He, Oliveira, Silva, Mateus & de Freitas, 2010; Mateus, Oliveira, Santos-Buelga, Silva & de Freitas, 2004), which further increase the complexity of the pigment composition. Yet, in comprehensive screening studies dozens of individual anthocyanins and monomeric anthocyanin adducts can be readily and efficiently detected and characterized with liquid chromatography–mass spectrometry (LC–MS; Alberts, Stander & de Villiers, 2012; Alcalde-Eon, Escribano-Bailón, Santos-Buelga & Rivas-Gonzalo, 2006; Blanco-Vega, Gómez-Alonso & Hermosín-Gutiérrez, 2014; Boido, Alcalde-Eon, Carraus, Dellacassa & Rivas-Gonzalo, 2006; Willemse, Stander, Vestner, Tredoux & de Villiers, 2015). The analytical issues and the true complexity are caused by the oligomeric and polymeric adducts consisting of proanthocyanidins (PA) and anthocyanins. Malvidin glycosides (Mv) are the most common anthocyanins in red wines (He et al., 2012) and the PAs and malvidin glycosides can be combined in different ways to form oligomeric adducts. For instance, the malvidin unit can be the terminal unit (PA–Mv^+), first extension unit with the terminal unit being a flavan-3-ol monomer (FL; PA–Mv–FL) or the PA and malvidin units can be linked via a methylmethine bridge (PA–methylmethine–Mv^+; Fig. S1; Somers, 1971; Timberlake & Bridle, 1976; Zeng, Teissedre, & Jourdes, 2016).

Just like the monomeric anthocyanin adducts, small dimers of the oligomeric adducts can be detected and quantified efficiently and specifically with non-targeted or compound-specific LC–MS and LC–MS/MS methods (Alberts et al., 2012; Alcalde-Eon et al., 2006; Blanco-Vega et al., 2014; Boido et al., 2006; Willemse et al., 2015). In these studies, only the most abundant dimers have

been detected and larger oligomers have completely eluded themselves from the analytical methods. Some methods, such as the high-performance gel permeation chromatography method of Kennedy and Taylor (2003), or the combination of SO_2-bleaching and BSA precipitation by Harbertson, Picciotto & Adams (2003), enable the detection and semi-quantification of larger colored oligomeric adducts. However, the different structural sub-groups cannot be detected separately. Therefore, when oligomeric and polymeric pigments are discussed in studies utilizing these methods, they are often referred to as polymeric pigments rather than specifying the pigment groups more precisely (Bindon, Kassara, Hayasaka, Schulkin & Smith, 2014; Casassa, Huff & Steele, 2019).

Recently, a group-specific ultra-performance liquid chromatography–tandem mass spectrometry (UPLC–MS/MS) method was published, which utilizes multiple reaction monitoring (MRM) in the detection and semi-quantification of anthocyanins and anthocyanin adducts in red wine (Laitila, Suvanto & Salminen, 2019). In addition to many monomeric pigment groups, the method can detect oligomeric PA–Mv^+, PA–methylmethine–Mv^+ and PA–Mv–FL type adducts separately (Fig. S1). By the nomenclature proposed by Laitila et al., the method detects small oligomeric adducts (SOAs), medium-sized oligomeric adducts (MOAs) and large oligomeric adducts (LOAs) separately. Importantly, the typically detected small dimers comprised a smaller part of the total responses compared to the larger oligomers (Laitila et al., 2019). Therefore, the method provides more comprehensive picture of the composition of the oligomeric adducts than other methods, which only detect individual small dimers. Moreover, information about the initial composition is retained in the form of two dimensional (2D) chromatographic fingerprints, which enables comparison of the compositions between wines made of different grape varieties, for instance.

Here, the group-specific MRM method was utilized to study how the concentrations and average sizes of the oligomeric adducts differed in commercial red wines. The 2D chromatographic fingerprints were visually inspected to find out whether they would reveal completely new aspects

about red wines and their polyphenolic composition. To this end, a new protocol was developed to form quantile 2D chromatographic fingerprints to visualize and summarize the vast chromatographic data from multiple samples simultaneously. The red wines were studied in two phases. First (**i**) by concentrating only on the young wines to study the differences in wine varieties, and then (**ii**) by studying the evolution of the oligomeric compound groups by focusing on wines over a wider age range. Finally, several factors related to the polyphenolic starting material composition of the oligomeric adducts were analyzed from the 1-year-old wines to find out which factors contribute the most to the formation of the oligomeric adducts. The commercial red wines utilized in this study were part of a bigger wine set consisting of 317 red wines and the whole wine set was previously utilized to establish general connections between the pigment composition and color intensity of red wines (Laitila & Salminen, 2020). In the present paper, several sub-sets of the whole wine set (Tables S1–S3) were utilized to answer the questions above.

2. Materials and methods

2.1. Chemicals and red wines

The monomeric flavan-3-ols, i.e., (+)-Catechin, (−)-epicatechin, (−)-gallocatechin and (−)-epigallocatechin, were purchased from Extrasynthese (Lyon, France). Alko Inc., a finish national alcoholic beverage retailing company, provided part of the whole red wine set (n=272) and some were collected by the Natural Chemistry Research Group of University of Turku (n=45). In this paper, specific sub-sets of the whole wine set were utilized (Tables S1–S3). All wines were commercial red wines and aliquots of the wines were stored in Eppendorf tubes at –80 °C prior to the analyzes.

2.2 Quantitative and semi-quantitative analyzes

Concentrations of the malvidin glycosides and the oligomeric PA–Mv^+, PA–Mv–FL and PA–methylmethine–Mv^+ adducts were semi-quantified with the UPLC–MS/MS method of Laitila et al. (2019) as described by Laitila and Salminen (2020). The UPLC–MS/MS system consisted of a Waters Acquity UPLC system and of a Xevo triple quadrupole mass spectrometer. Shortly, the

compound groups were detected with the quantitative transitions of the group-specific MRM methods. The chromatogram areas were transformed to proportional concentrations with calibration curves, which were prepared from a dilution series of a single reference wine, a JP Chenet Merlot 2015. In other words, the concentrations were reported as percentage values compared to the reference wine and a concentration of 70%, for instance, meant that the concentration was 70% of the concentration in the reference wine. This approach was utilized solely to take into account non-linear responses in some of the compound groups (Laitila et al., 2019). Total concentrations of the oligomeric adducts were calculated from the summed total chromatogram areas of the SOAs, MOAs and LOAs and the corresponding SOA-%, MOA-% and LOA-% were defined and calculated as the proportions of different-sized adducts of the total summed chromatogram area. All instrumentational and methodological details and analytical conditions are described in detail in Laitila et al. (2019) and Laitila and Salminen (2020).

PA concentrations and mean degree of polymerization (mDP) of the PAs were acquired with the Engström method (Engström, Pälijärvi, Fryganas, Grabber, Mueller-Harvey & Salminen, 2014) as described by e.g., Malisch et al. (2015). Exceptions were that the procyanidin (PC) and prodelphinidin (PD) sub-units were detected using three different cone voltages, which were 75 V, 85 V and 140 V for the PCs, and 55 V, 80 V and 130 V for the PDs, and that the gradient was the same as utilized by Laitila et al. (2019). The UPLC–MS/MS system, eluent gradients and other analytical conditions were the same as utilized in quantifying the oligomeric adducts. The PC and PD units were quantified separately with calibration curves in the range of 0.1875–1.50 $mg \times mL^{-1}$ and 0.25–2.00 $mg \times mL^{-1}$, respectively and the calibration was achieved by using two PA mixtures with known PC and PD contents. The calibration curve for the mDP was similarly obtained by using six different Sephadex LH-20 fractions with known PC, PD and mDP contents. The Engström method contains an optimized selected reaction monitoring channel for the detection of monomeric flavan-3-ols (Engström et al., 2014), and this channel was used to detect the monomers. (+)-catechin, (–)-epicatechin, (+)-gallocatechin and (–)-epigallocatechin were quantified compound

specifically using the corresponding standards. A stock solution of catechin (1 μg mL^{-1}) was used as an external standard to monitor the quantitative performance of the instrument and five replicate analyzes were performed after every 10 samples or approximately once in every two hours with a catechin specific selected reaction monitoring method.

2.3. Formation of the two-dimensional quantile fingerprints

Comparison of the qualitative adduct composition was done by calculating 2D quantile fingerprints from multiple individual 2D fingerprints (Fig. 1). To compare the compositions of the oligomeric adducts in the young wines (1–2 years old wines), ten red wines were randomly sampled from young Merlot, Cabernet Sauvignon and Shiraz wines and from the Pinot Noir wines of the Pfalz region in Germany. The Touriga Nacional and Pinot Noir wines from the Beaune region in France had only eight samples, which were all used to calculate the quantile 2D fingerprints. Fifteen red wines were randomly sampled from the 1-year-old and 3-year-old Shiraz, Cabernet Sauvignon and Merlot wines to present the evolutionary patterns of the adduct composition. All 15 4–6 years old Shiraz, Cabernet Sauvignon and Shiraz wines were used.

Briefly, median 2D fingerprints of the oligomeric adducts and 10^{th} and 90^{th} percentiles were calculated to visualize and summarize the adduct composition and their variations within various groups (Fig. 3, 5). Shortly, the SOA, MOA and LOA fingerprints were summed to form a single total fingerprint and the total fingerprints were scaled to the highest intensity. Then, multiple total fingerprints were aligned with a parametric time warping algorithm (Bloemberg et al., 2010; Eilers, 2004) and the quantile 2D fingerprints were calculated from multiple aligned total fingerprints. The workflow is visually demonstrated in Figure 1.

In detail, raw chromatograms were first smoothed in MassLynx software (version 4.2, Waters Corporation) and then extracted in a list format. Mean smoothing algorithm was applied twice with the smoothing width of four scans for the PA–methylmethine–Mv^+ and PA–Mv–FL adducts and four times with the same width for the PA–Mv^+ adducts and malvidin glycosides. Summation of the SOA, MOA and LOA fingerprints and calculation of the quantile fingerprints

were done in Microsoft Excel 2016. Chromatogram alignment was done with R package "*ptw*" (Bloemberg et al., 2010; Eilers, 2004). Best reference chromatogram in each group (wine type or and age group) was first selected using the bestref-function using weighted cross-correlation criterion without smoothing and with a triangle width of five. Then, remaining chromatograms were aligned to the selected refence with quadratic warping functions using the same parameters as in the bestref-function. Finally, the quantile fingerprints were plotted in Origin 2016 (OriginLab Corporation, Northampton, MA, USA).

2.4. Statistical analyzes

All statistical analyzes and the chromatogram alignment were performed with R (version 3.5.3; R Core Team, 2019) in Rstudio integrated development environment (version 1.1.463; Rstudio Team, 2018). The differences between groups in the quantitative data (Fig. 2 and 4) were tested with non-parametric Kruskal–Wallis tests and the subsequent multiple comparisons were done with Holm-corrected Mann–Whitney U-tests. Partial least squares regression (PLSR) models were utilized two study the connections between the compositions of the oligomeric adducts (responses) and the PA, FL and Mv concentrations, the mDP of the PAs and the concentration ratios of the FLs and PAs (FL/PA; predictors) in the 1-year-old wines. Separate model was built for each response (concentrations and LOA-%). The responses, FL concentrations and FL/PA ratios were log-transformed to meet the assumption of linear correlations. All variables were centered and scaled to zero means and to unit variances. The PLSR models were fitted with R package "*plsdepot*" (Sanchez, 2012), all models consisted of two latent variables and the models were cross-validated to avoid over-fitting.

3. Results and discussion

3.1. Quantitative composition of the oligomeric adducts in young wines

First, the focus was on the young wines, which were defined to be the 1–2 years old wines. The primary aim was to see how (**i**) the quantitative composition, i.e., concentrations and sizes of the oligomeric adducts, and (**ii**) the 2D chromatographic fingerprints varied between different wine

types. The secondary objective was to find out which wine types could be pooled to study the evolution of the oligomeric adducts. Such wines were selected from the young wines which had at least eight replicates based on the principal grape variety (n=92; Table S1). These grape varieties were Pinot Noir, Shiraz, Merlot, Cabernet Sauvignon and Touriga Nacional. The wines were mainly single-varietal wines except the Touriga Nacional wines, all of which contained Touriga Franca and Tinta Roriz grapes as well. Principal component analysis and hierarchical clustering with the principal component scores were done to group the data but the best way turned out to be to simply group the wines based on the principal grape varieties used in the wine making. However, the Pinot Noir wines from the Pfalz (Germany) and Beaune (France) regions were considered separately because the unsupervised multivariate methods showed them to form two separated groups.

In the young wines, all the three groups of oligomeric adducts had differences in their concentrations between the wine types (Fig. 2A–C; $p<0.001$; refer to Tables S4–S6 for statistical pairwise comparisons). The variation in the concentrations was the largest for the PA–Mv^+ adducts where nearly a five-fold difference was observed between the concentrations in the Shiraz, Cabernet Sauvignon and Touriga Nacional wines (average 79%) and the Pinot Noir wines from the Beaune region (average 16%, Fig. 2A). Approximately, a two-fold difference in the concentrations of the PA–methylmethine–Mv^+ adduct was observed between the Merlot, Shiraz and Cabernet Sauvignon wines (average 79%) and the Touriga Nacional and Pinot Noir wines from the Pfalz region (average 43%, Fig. 2B). The largest statistically significant differences in the concentrations of the PA–Mv–FL adducts were between Shiraz and Cabernet Sauvignon wines (average 111%) and the Pinot Noir wines of the Pfalz region (average 74%, Fig. 2C) where the difference was only one and a half fold. The Shiraz and the Cabernet Sauvignon wines had consistently the highest concentration of all three types of oligomeric adducts, but in the other groups the relative order of the concentrations varied between the adduct groups. For instance, the Touriga Nacional wines had as high a concentration of PA–Mv^+ and PA–Mv–FL adducts as Shiraz and Cabernet Sauvignon wines, while the concentration of PA–methylmethine–Mv^+ adducts was lower and similar to the levels in Pinot Noir wines from

the Beaune and Pfalz regions (Fig. 2A–C). Additionally, the concentrations of the PA–methylmethine–Mv^+ adducts in the Merlot wines were similar to the concentrations in Shiraz and Cabernet Sauvignon wines, while the concentrations of PA–Mv^+ adducts in Merlot wines were similar to the Cabernet Sauvignon wines but lower than in Shiraz wines (Fig. 2A–B). Between the Pinot Noir wines from the Beaune and Pfalz regions, the differences in the concentrations were significant only in the case of the PA–Mv^+ adducts and, overall, the concentrations of all three groups of the oligomeric adducts were in the lower end of the concentration ranges in the Pinot Noir wines.

The average sizes of the oligomeric adducts had differences between the groups as well (Fig. 2D–F). The SOA-% and LOA-% of all three oligomeric adduct groups had significant differences between the wine types ($p<0.001$, refer to Tables S7–S12 for statistical pairwise comparisons), while the MOA-% of all three adduct groups were found to be equal ($p\geq0.066$). For the PA–Mv^+ and PA–methylmethine–Mv^+ adducts, these differences were solely caused by the Pinot Noir wines from the Pfalz region, which had higher SOA-% and lower LOA-% than the other groups (Fig. 2D–F). Therefore, the Pinot Noir wines from the Pfalz region had the lowest average sizes of the PA–Mv^+ and PA–methylmethine–Mv^+ adducts, while the average sizes of these adducts were similar in the Merlot, Shiraz, Cabernet Sauvignon, Touriga Nacional and Pinot Noir wines from the Beaune region. Most variation in the adduct sizes was observed with the PA–Mv–FL adducts (Fig. 2D–F). The Touriga Nacional and Shiraz wines had the lowest SOA-% (average 23%) and the highest LOA-% (average 40%) meaning that the average sizes of the PA–Mv–FL adducts were the highest in these two wine groups. Based on the LOA-% and SOA-% values, the average sizes of the PA–Mv–FL adducts were the second highest in the Merlot and Cabernet Sauvignon wines, followed by the Pinot Noir wines from the Beaune region. Again, the average adduct size was the lowest in the Pinot Noir wines from the Pfalz region.

3.2. Qualitative composition of the oligomeric adducts in young wines

The workflow described in section 2.3., and demonstrated in Figure 1, made it possible to visualize and summarize the chromatographic information from 504 individual 2D chromatographic fingerprints in only 18 quantile fingerprints (Fig. 3). Quantile-based approach was chosen rather than calculating average chromatograms and standard errors or confidence intervals because the quantiles were more flexible in visualizing the shapes of the fingerprints, and especially when it came to visualizing the variation within different groups. Standard error (or related errors) are symmetrical around the average, which can lead to problems with chromatographic data because there is a certain baseline under which the intensities should never go in a well working method. Moreover, outliers are always expected to be positive. This can lead to situations where typical error measures could cover negative intensities because of the symmetry (Fig. S2), which is not chemically sensible to allow. This sort of a problem is avoided altogether with quantiles because distances between median and 10^{th} and 90^{th} percentiles are not (necessarily) equal (e.g., Fig. 3O and Fig. S2). Overall, the quantile 2D chromatographic fingerprints made it possible to summarize both within-group variation (e.g., Fig. 3C) and between-group variation (e.g., Fig. 3C–F) effectively, and in an intuitive and visual manner.

Compositionally, the Pinot Noir wines from the Pfalz and Beaune regions differed notably from the other groups. First, the 2D fingerprints of the PA–Mv^+ adducts in Pinot Noir wines lacked the peaks of many individual compounds that were present in the fingerprints of other wine groups (e.g., dimeric adducts containing malvidin-3-*O*-acetylglucoside and malvidin-3-*O*-coumaroylglucoside units; Fig. 3A–F). Second, the PA–methylmethine–Mv^+ fingerprints had much more within-group variation than the other groups, and the fingerprints contained major individual peaks, namely the dimeric adducts containing malvidin-3-*O*-glucoside unit (Fig. 3G–L). Third, the chromatographic hump of the PA–Mv–FL adducts was less abundant than in the other groups (Fig. 3M–R). Regarding the first point, the lack of multiple individual PA–Mv^+-type adducts was presumably explained by the malvidin glycoside compositions of the wines (Fig. S3). Pinot Noir grapes and wines are known to differ from other common grape and wine varieties in their

anthocyanin composition as they do not contain acetylated anthocyanins (Dimitrovska, Bocevska, Dimitrovski & Murkovic, 2011). The malvidin glycoside fingerprints replicated this result in Pinot Noir wines from the Beaune region, whereas the wines from the Pfalz region contained small amounts of acetylated malvidin glycosides (Fig. S3). The malvidin glycoside composition might explain the qualitative differences in the PA–methylmethine–Mv^+ fingerprints as well. In the Pinot Noir wines, a catechin–methylmethine–malvidin-3-*O*-glucoside dimer was the main individual compound in the median fingerprints (Fig. 3G, H). In the other wine groups, the same dimers were only visible in the 90th percentile fingerprint (Fig. 3I and K), if they were visible at all. When there is only one major precursor for the adduct formation (i.e., malvidin glucoside), the adducts containing the said precursor should be more emphasized in the resulting composition (Fig. 3G, H) compared to a situation when there are multiple potential and competing pre-cursors (Fig. 3I–L). Alternatively, the qualitative differences in the PA–methylmethine–Mv^+ compositions could be related to the compositions of the other precursor, i.e., the PAs.

Interestingly, the fingerprints of the PA–Mv–FL adducts had fewer qualitative differences between the Pinot Noir wines and the other groups (Fig. 3M–R) than with the other two types of oligomeric adducts. The major difference was the chromatographic hump, which was less abundant than in Pinot Noir wines, and which supported the earlier observation about the lower average sizes of the PA–Mv–FL adducts in the Pinot Noir wines. The individually distinguishable peaks appeared to be mostly the same in all wine groups (Fig. 3M–R), despite the differing malvidin glycoside composition (Fig. S3). Previously, it was noted that the dimeric adducts of catechin and acylated malvidin glucosides were less abundant in PA–Mv–FL fingerprints than the corresponding dimers were in the PA–Mv^+ fingerprints (Laitila et al., 2019). The reason for this could arise from the mechanism of the adduct formation reaction. When the PA–Mv–FL adducts are formed, the reactive moiety of anthocyanins (in their flavylium cation form) is the electrophilic C4 carbon of the C-ring (Dangles & Fenger, 2018; Salas, Fulcrand, Meudec & Cheynier, 2003). In this case, the glycosidic substituent at the C3-position might produce a steric hindrance, which in turn would be

proportional to the size of the substituent. On the contrary, the glycosidic substituent probably does not affect the formation of the PA–Mv^+ and PA–methylmethine–Mv^+ adducts, as the reactive moieties are the C6 and C8 carbons of the A-ring. Overall, the dimeric adducts appeared to be the major individual compounds of the PA–Mv^+ and PA–Mv–FL adducts (Fig. 3A–F, M–R) and of PA–methylmethine–Mv^+ adducts in the Pinot Noir wines from the Beaune and Pfalz regions (Fig. 3G–H). The fingerprints of the PA–methylmethine–Mv^+ adducts of the other wine groups were dominated by the chromatographic humps (Fig. 3I–L). Subsequently, the PA–methylmethine–Mv^+ adducts formed compositionally the most complex compound group.

If the Pinot Noir wines were excluded, then in most cases there was remarkably little of either within-group or between-group compositional variation in the red wines (Fig. 3). Overall, the qualitative differences in the young wines were mainly limited to differences in abundances of individual peaks. For instance, the abundance of PA–Mv^+-type catechin–malvidin-3-*O*-acetylglucoside dimer was lower in the Touriga Nacional wines than in Merlot, Shiraz and Cabernet Sauvignon wines (Fig. 3C–F). Differences in the locations or in the shapes of the underlying chromatographic humps are often signs of major differences in the chemical compositions. Such drastic differences were not observed in the PA–Mv^+ and PA–methylmethine–Mv^+ adducts (Fig. 3C–F, I–L) in the Merlot, Shiraz, Cabernet Sauvignon and Touriga Nacional wines. On the contrary, the fingerprints of the PA–Mv–FL adducts displayed similarity with the conclusions made of the SOA-%, MOA-% and LOA-% parameters. These parameters suggested that the average size of the PA–Mv–FL adducts were higher in the Shiraz and Touriga Nacional wines than in the Merlot and Cabernet Sauvignon wines (Fig. 2E). This was visible in the fingerprints, as the underlying chromatographic humps were more predominant in the Shiraz and Touriga Nacional wines (Fig. 3O–R) than in Merlot and Cabernet Sauvignon wines (Fig. 3O, Q). Typically, the fingerprints of naturally occurring PAs can have major differences between plant species (Salminen, 2018). The same did not appear to be true for the three types of oligomeric adducts in different *Vitis vinifera*

red wines, especially when the malvidin glycoside composition was similar between the wine groups (Fig. S3).

To summarize the results regarding the young wines (sections 3.1. and 3.2.), there were only few differences in the composition of the oligomeric adducts between Shiraz, Merlot, Cabernet Sauvignon and Touriga Nacional wine varieties (Fig. 3). Moreover, the average sizes of the oligomeric adducts in these wines were mainly similar based on the quantitative size-distribution parameters (Fig. 2D–F) and the chromatographic fingerprints. The PA–Mv–FL adducts were an exception, as there were small differences in the average adduct sizes between the wine types. However, the biggest differences between these wines were found to be quantitative (Fig. 2A–C). The Pinot Noir wines stood out from the other wine groups both by their typically lower concentrations and by their compositions (Fig. 2 and 3). Additionally, the Pinot Noir wines of the Pfalz region had systematically the lowest average sizes of the oligomeric adducts (Fig. 2D–F).

3.3. Quantitative evolution of the oligomeric adducts

To test how the age of the wines affected the adduct composition, wine types with similar chemical composition (based on Fig. 3) were pooled together. These were the wines where Merlot, Shiraz or Cabernet Sauvignon grapes were used as primary grapes (n=95, Table S2). The concentrations of the PA–methylmethine–Mv^+ adducts ($p<0.001$) and PA–Mv^+ adducts ($p<0.001$) were found to decrease, whereas the concentrations of PA–Mv–FL adducts remained unchanged (Fig. 4A–C; $p=0.632$; refer to Tables S13 and S14 for statistical multiple comparisons). Age had the biggest impact on the concentration of the PA–methylmethine–Mv^+ adducts, as the concentration decreased significantly year by year from the initial average of 90.0% to 21.7%. The concentration of the PA–Mv^+ adducts halved, as the concentration decreased from the initial average of 80.5% to 41.5%. This decrease was not as extreme as with the PA–methylmethine–Mv^+ adducts, whose concentrations decreased to a fourth of the initial concentration. All three size distribution parameters of the PA–Mv^+ and PA–methylmethine–Mv^+ adducts were affected by aging ($p<0.001$), while the size distribution parameters of the PA–Mv–FL adducts remained unchanged ($p\geq0.056$;

Fig. 4D–F, refer to Tables S15–S20 for statistical multiple comparisons). Again, the PA–methylmethine–Mv^+ adducts turned out to be the most heavily affected by aging as their SOA-% decreased from the initial 25.0% to 17.8%, the MOA-% decreased from 38.0% to 35.4% and the LOA-% increased from 37.1% to 46.8%. In the oldest analyzed wines, the LOAs produced nearly a half of the areas of the total fingerprints, which was markedly more than in the 1-year-old wines. Recently, the LOA-% was shown to have a positive effect on the color intensity in a wine set with a wide age range, but not in the subset of 1-year-old wines (Laitila et al., 2019). It was proposed that there could be and an evolutionary trend in the composition of the PA–methylmethine–Mv^+ adducts, which would become relevant for the wine color as the wines aged. Now, a strong evolutionary trend in the sizes of these adducts was discovered, which might be relevant for the wine color. The SOA-% of the PA–Mv^+ adducts decreased only from 26.8% to 24.2%, the MOA-% decreased from 35.9% to 34.2% and the LOA-% increased from 37.3% to 41.6%. While these changes were statistically significant, the changes in the average sizes of the PA–Mv^+ adducts appeared to be less substantial compared to the PA–methylmethine–Mv^+ adducts. Based on the quantitative metrics, the PA–Mv–FL adducts were the most stable oligomeric adducts.

The previous knowledge about the quantitative evolution of the oligomeric adducts has been obtained by measuring, almost exclusively, the concentrations of individual dimeric adducts (e.g., Alcalde-Eon et al., 2006; Arapitsas, Perenzoni, Nicolini & Mattivi, 2012; Blanco-Vega et al., 2014; Boido et al., 2006). These studies showed that the FL–methylmethine–anthocyanin dimers began to diminish in concentration already during maturation stage, which continued to the aging stage. The concentrations of the directly-linked FL–anthocyanin dimers increased during aging but began to decrease during bottled storage. Moreover, the degradation of acetaldehyde-mediated dimers was more rapid than the degradation of their directly linked counterparts. Now, it was confirmed that these evolutionary trends were valid for the whole compound groups of PA–Mv^+ and PA–methylmethine–Mv^+ adducts as well (Fig. 4A, D). It was also confirmed that the degradation of the PA–methylmethine–Mv^+ adducts was faster than the degradation of the PA–Mv^+ adducts, which

was in line with the observations made of the dimeric adducts. It is important to note the difference between the previous studies and the present study. To this date, the conclusions made about the evolution of the oligomeric adducts have often been based on monitoring the evolution of dimeric and trimeric adducts, whereas the present results were based on monitoring the evolution of much larger number of compounds. Additionally, the present results showed how there was an evolutionary trend in the size-distribution of the PA–Mv^+ and PA–methylmethine–Mv^+ adducts, which is something that would not be possible to detect by only monitoring individual dimers.

3.4. Qualitative evolution of the oligomeric adducts

For the comparison of the qualitative evolutionary trends of the oligomeric adducts, the quantile 2D fingerprints of the Shiraz, Merlot and Cabernet Sauvignon wines of different ages were compared with a similar approach as the fingerprints of the young wines. This approach made it possible to visualize and summarize the chromatographic information from 405 individual 2D fingerprints into only nine fingerprints in Figure 5.

The relative abundance of some individual PA–Mv^+-type adducts, e.g., catechin–malvidin-3-*O*-acetylglucoside and catechin–malvidin-3-*O*-coumaroylglucoside, were observed to decrease as the wines aged (Fig. 5A–C). At the same time, the underlying chromatographic hump appeared to become slightly more prevalent, which together suggested that the average size of the PA–Mv^+ adducts was larger in the older wines. Therefore, the qualitative observations made of the 2D fingerprints and quantitative size-distribution parameters revealed the same trend. In the 2D fingerprints of the PA–Mv–FL adducts, an increase in the relative abundance of trimeric adducts was clearly observed (Fig. 5G–I). Additionally, a slight shift of the chromatographic hump could be seen, as the middle point of the hump moved slightly to a lower retention time (Fig. 5G–I). Trimeric PA–Mv–FL adducts have lower retention times than dimers when analyzed with reversed-phase liquid chromatography systems (Laitila et al., 2019; Willemse et al., 2015), meaning that the change in the chromatographic hump might indicate that larger adducts were formed. These observations did not match with the conclusion made of the size-distribution parameters, which suggested that

there was no change in the average sizes of the PA–Mv–FL adducts. However, in this case the 2D chromatographic fingerprints provided more compelling data for the evolutionary patterns compared to the quantitative data, because the fingerprints were purely qualitative. The size-distribution parameters of the PA–Mv–FL adducts in the young wines (Fig. 2E) had statistically significant differences between the three wine varieties (Shiraz, Merlot and Cabernet Sauvignon). This might disturb the investigation of the quantitative evolutionary pattern if the evolutionary trend is only modest in comparison to stronger trends, such as the one observed for the PA–methylmethine–Mv^+ adducts. Interestingly, even though the PA–methylmethine–Mv^+ adducts were the most influenced by the aging based on the quantitative measures (Fig. 4C, F), their 2D fingerprints were relatively alike in the three age categories (Fig. 5D–F). This could be due to the PA–methylmethine–Mv^+ composition itself, which was already initially so complex that the 2D fingerprints were dominated by the chromatographic hump. The complexity, in turn, likely hid any major visual changes in the composition.

Even in the oldest wines, the dimeric adducts remained clearly present as the main individual compounds in the fingerprints of the PA–Mv^+ and PA–Mv–FL adducts, even though the average sizes of the adducts increased. The within-group variation in the 2D fingerprints decreased when the wines aged, which was especially apparent for the PA–methylmethine–Mv^+ adducts (Fig. 5D–F). Chemically, this could occur if the individual 1-year-old wines still had compositional differences, which all then converged to similar composition after years of storage. When the observations from the quantitative metrics and 2D chromatographic fingerprints were combined, it could be concluded that the PA–methylmethine–Mv^+ adducts were the most unstable and reactive compounds, as both their concentration and composition changed drastically. The PA–Mv^+ adducts were the second most unstable oligomeric adducts followed by the PA–Mv–FL adducts, which remained stable based on the quantitative metrics but faced some compositional changes according to the 2D fingerprints.

At least two mechanisms could have contributed to the compositional evolution in the size-distributions of the oligomeric adducts. First, (**i**) the adducts degraded regardless of their size as the wines aged but the smallest adducts were more susceptible to the degradation than the larger ones, leading to the increased average size of the adducts and in increase in the LOA-% and decrease in the SOA-%. Second, **(ii)** the smaller adducts oligomerized to larger compounds, possibly by reactions with carbocationic cleavage products of PAs. The prerequisite for mechanism **ii**, i.e., the production of carbocationic PA degradation products, has been shown to occur in mildly acidic red wine model solutions (Dallas, Ricardo-da-Silva & Laureano, 1996; Vidal, Cartalade, Souquet, Fulcrand & Cheynier, 2002). Third possibility was that both **i** and **ii** occurred at the same time. In the case of the PA–Mv–FL adducts, it was likely that **ii** occurred because the changes in the concentrations were minimal to non-existent, while the changes in the 2D fingerprints suggested that higher oligomers were formed. On the contrary, the degradation of PA–methylmethine–Mv^+ and PA–Mv^+ adducts was confirmed by the diminishing concentrations meaning that mechanism **i** was a more probable cause for the changes in the size distribution. However, mechanism **ii** could not be ruled out and it might occur as well. Finally, it needs to be noted that while the utilized method showed that there were compositional evolutionary trends, the methodology itself cannot reveal what exactly caused the observed changes. Therefore, the explanation offered in this section are hypothetical and tentative.

3.5. Factors regulating the formation of the oligomeric adducts

The red wine data set made it possible to investigate what factors possibly regulated the formation of the oligomeric adducts. To this end, the PA concentration, mDP of the PAs, concentration of FLs and the concentration of malvidin glycosides were measured from the 1-year-old wines (Table S3). The FL concentration was measured because the monomers share the same nucleophilic properties as the PAs and they can participate as such to the adduct formation of the PA–Mv–FL and PA–methylmethine–Mv^+ adducts (Dangles & Fenger, 2018; Salas et al., 2003). FLs cannot participate to the formation of PA–Mv^+ adducts as such but they can compete with the malvidin glycosides over

the carbocationic PA degradation products, which in turn are needed in the formation of the PA–Mv^+ adducts. Therefore, a ratio of the FL and PA concentrations (FL/PA) was calculated to test for such a scenario. The data was analyzed by fitting PLSR models where the above-mentioned variables were used as predictors for the concentrations and proportions of LOAs of the oligomeric adducts. The LOA-% was used as a metric for the degree of oligomerization of the oligomeric adducts, just as it was in Laitila and Salminen (2020). Only the youngest of wines were studied to minimize the effect of the age. In this setup, it was assumed that the values of the predictors reflected their initial values at the beginning of the fermentation because, though they were inaccessible information in this study, they would have been more suitable predictors. For this reason, the following results should be considered as tentative ones. The main results of the regression analyzes are presented in Figure 6 and Table S21.

The models explained large portions of the variation in the PA–Mv^+ (R^2=0.725, Q^2=0.703) and PA–methylmethine–Mv^+ concentrations (R^2=0.589, Q^2=0.513) but the PA–Mv–FL concentrations could not be accurately modelled (R^2=0.111, Q^2=0.017). The Mv concentration and the mDP of the PAs were the most important predictors for the PA–Mv^+ concentration (Fig. 6A). The Mv concentration and the FL/PA ratio were the most important predictors for the PA–methylmethine–Mv^+ concentration (Fig. 6B) and only the Mv concentration was important for the PA–Mv–FL concentration (Fig. 6C). Based on these results, the Mv concentration was a more important limiting factor than the PA concentration was as such. Interestingly, the mDP of the PAs had a positive effect on the concentrations of the PA–Mv^+ adducts. This observation could be rationalized by considering the adduct formation mechanism. The PA–Mv^+ adducts can only be formed when malvidin glycosides in their hemiacetal forms get to react with carbocationic degradation products of PAs (Salas et al., 2003). The depolymerization of PAs happens through degradation of the interflavan C–C linkages and as the number of these linkages increases in PAs (i.e., the mDP increases), so does the number of possible sites for the degradation to happen. The formation mechanism could also explain the importance of the FL/PA-ratio. The monomeric FLs

cannot produce the reactive carbocations, but the FLs can react with them leading to a competition with malvidin glycosides. Now, a low FL/PA ratio means that there are more PAs to produce the carbocationic species and less monomeric FLs to react with them (repolymerization) compared to a high FL/PA ratio. The large regression coefficient of the FL/PA ratio was negative, suggesting that low FL/PA ratio was favorable for the formation of the PA–Mv^+ adducts. The FL/PA ratio was an important predictor for the PA–methylmethine–Mv^+ concentration as well (with a negative regression coefficient; Fig. 6B). In this case, both the monomeric FLs and oligomeric PAs can participate to the adduct formation reaction as such. Here, the chemical importance of the FL/PA ratio was not as apparent as for the PA–Mv^+ adducts but, again, low FL/PA ratio was associated with high PA–methylmethine–Mv^+ concentration. Possibly, a high PA content in relation to the FL content simply helps the formation of the adducts because PAs have more reactive sites available for the adduct formation.

Only 11.1% of the variation in the PA–Mv–FL concentrations could be explained with the measured starting material composition and even that model was over-fitted (Q^2=0.017; Fig. 6C). This implied that some other factor must have a dominant role in regulating the formation of the PA–Mv–FL adducts. The most appealing of these factors was the pH values of the wines. The formation of the Mv–FL dimers involve a malvidin glycoside reacting as an electrophile with an FL unit (Dueñas, Fulcrand & Cheynier, 2006). Only the flavylium cation form of the malvidin glycosides has the necessary electrophilic properties but this form is a minor constituent in the anthocyanin equilibrium in mildly acidic pH conditions in red wines compared to the nucleophilic hemiacetals (Nave, Petrov, Pina, Teixeira, Mateus & de Freitas, 2010; Pina, 1998). Consequently, the anthocyanin equilibrium and the abundance of the flavylium cationic malvidin glycosides could be more important in regulating the formation of Mv–FL adducts than the polyphenolic starting material composition. The PA–Mv–FL adducts probably form by further oligomerization reactions between Mv–FL dimers and carbocationic PA degradation products (Fig. S4, mechanism 1). Alternatively, it could be possible for the PA–Mv–FL adducts to be formed from the PA–Mv^+

adducts by reactions with FLs (Fig. S4, mechanism 2). The concentrations of the PA–Mv^+ and PA–Mv–FL adducts were indeed found to correlate in the 1-year-old wines (r=0.68). However, in the case of mechanism 2 (Fig. S4), it would be expected that the concentrations of the PA–Mv–FL adducts would be better explained by the same predictors as the PA–Mv^+ adducts but this was not case in this study.

In this study, the focus was entirely on the polyphenolic components in red wines and there are other non-polyphenolic predictors, such as pH and acetaldehyde content, that can potentially regulate the adduct formation. Despite lacking this information, the models explained large proportions of the variation in the PA–Mv^+ and PA–methylmethine–Mv^+ concentrations, thereby revealing what factors in the polyphenolic starting material composition could be important for their formation.

Similar regression models were constructed for the LOA-% of the oligomeric adducts to find out what factors favor the formation of larger adducts (Fig. 6D–F, Table S21). The FL/PA ratio was a significant predictor for each adduct type (with negative regression coefficient) and, additionally, the malvidin glycoside concentration was significant predictor for the PA–methylmethine–Mv^+ adducts. Concerning all three oligomeric adduct groups, the mechanism described for the formation of the PA–Mv^+ adducts might explain the importance of the negative connection between the FL/PA ratio and the LOA-% of the adducts. The lower FL/PA ratio is, the less there are FLs to compete with the oligomeric adducts in further oligomerization reactions. Regarding the methylmethine-linked adducts, FLs and PAs can both participate to the adduct formation as such and, certainly, a reaction with a PA leads to a larger adduct. Additionally, a PA moiety would have more remaining nucleophilic sites left in a newly formed adduct for further oligomerization reactions compared to a corresponding adduct containing an FL unit. The same rationalization could also explain the negative connection between the malvidin glycoside concentration and the LOA-% of the PA–methylmethine–Mv^+ adducts. While a higher malvidin concentration yields more PA–methylmethine–Mv^+ adducts (Fig. 6B), the higher concentration

favors acetaldehyde-mediated dimerization of malvidin glycosides as well, resulting in larger abundance of small methylmethine-linked malvidin glycoside dimers. Mv–methylmethine–Mv dimers have been shown to produce the same vinylmalvidin fragment (Bindon et al., 2014), which was used to detect the PA–methylmethine–Mv^+ adducts. Oligomeric methylmethine-linked malvidin glycosides are therefore likely detected as PA–methylmethine–Mv^+ adducts by the utilized UPLC–MS/MS method, even though such oligomers have not been separately detected or characterized in the PA–methylmethine–Mv^+ fingerprints so far.

The results of this section are important because recently it was shown that the high concentration and high LOA-% of the PA–Mv^+ and PA–methylmethine–Mv^+ adducts were associated with strong color intensity (Laitila & Salminen, 2020). In this context, the present results imply that if a strong color is wanted, the PAs in the initial polyphenolic starting material composition should have a high mDP, Mv concentration should be high, and the FL concentration should be low compared to the PA concentration. The role of the malvidin glycosides is interesting because only the pinotin-type pigments contributed less to the color intensity than malvidin glycosides (Laitila & Salminen, 2020). However, it does seem that the malvidin glycoside concentration is one of the main limiting factors for the formation of the oligomeric pigments and, thereby, malvidin glycoside concentration has an important impact on the color intensity. As mentioned in the beginning of this section, the results of this section rely on certain assumptions and, hence, the results should be considered as tentative.

4. Conclusions

The quantile 2D fingerprints revealed how the young Pinot Noir wines were compositionally distinctively different from Shiraz, Cabernet Sauvignon, Merlot and Touriga Nacional wines, which in turn were relatively similar with one another. Besides the lack of major between-group compositional variation between the latter four wine types, there was also only little within-group compositional variation. However, there were significant quantitative differences between wine varieties, which had only minor compositional differences. Regarding the quantitative evolution of

544 the oligomeric adducts, the PA–methylmethine–Mv^+ adducts were the most unstable adducts, as
545 their concentration in the 5–6-year-old wines was only one fourth of the initial concentration in the
546 1-year-old wines. The 2D quantile fingerprints, and the SOA-%, MOA-% and LOA-% parameters,
547 showed how the compositions of the oligomeric adducts evolved during storage. Generally, the
548 average sizes of all three groups oligomeric adducts increased towards the older wines and the
549 changes were again most notable for the PA–methylmethine–Mv^+ adducts. Finally, tentative results
550 were obtained about what factors in the starting material composition of the oligomeric adducts
551 might have an impact on their formation. The most important implications of this paper come from
552 the evolutionary patterns and of the tentative results regarding formation of the oligomeric adducts.
553 The evolutionary trends of the oligomeric adducts, and other pigments in red wines, have been
554 extensively studied during the last decade (e.g., references in the introduction). This study takes
555 another profound step forward in this field because the quantitative and qualitative results were not
556 obtained by monitoring individual small oligomers, but rather a much larger number of compounds
557 using novel group-specific methodology. On the other hand, the tentative results about the
558 formation of the oligomeric adducts might turn out to be useful if one wants to optimize the
559 accumulation of color intensity by optimizing the polyphenolic starting material composition of the
560 initial wine must, or simply understand why two different polyphenolic compositions lead to
561 different color intensities. In this study, only the oligomeric derivatives of malvidin glycosides were
562 considered but the results should be indicative of the behavior and properties of other analogous
563 anthocyanin adducts as well.
564

Abbreviations

FL, flavan-3-ol monomer; FL/PA ratio; ratio of the concentrations of the monomeric flavan-3-ols and proanthocyanidins; LC, liquid chromatography; LOA, large oligomeric adduct; MOA, medium-sized oligomeric adduct; MRM, multiple reaction monitoring; MS, mass spectrometry; MS/MS, tandem mass spectrometry; Mv, malvidin glycoside; MvGlc, malvidin glucoside; MvAcGlc, malvidin acetylglucoside, MvCouGlc, malvidin coumaroylglucoside; PA, proanthocyanidin; PA–methylmethine–Mv^+, methylmethine-linked proanthocyanidin–malvidin glycoside adducts; PA–Mv^+, directly-linked proanthocyanidin–malvidin glycoside adduct where the malvidin unit is the terminal unit; PA–Mv–FL, proanthocyanin–malvidin glycoside–flavan-3-ol adduct where the malvidin is the first extension unit; SOAs, small oligomeric adduct; UPLC, ultra-performance liquid chromatography

Acknowledgements

Juha-Pekka Salminen contributed to conceptualization and resources (CRediT) in this study. His helpful and insightful comments during the study, and his help in improving this manuscript, are acknowledged and much appreciated. Anne Koivuniemi, Milla Leppä and Jorma Kim are acknowledged for maintenance of the UPLC–MS/MS instrument. Jussi Suvanto is acknowledged for preparing calibration curves for the monomeric flavan-3-ols. Maarit Karonen kindly helped to improve an early version of this manuscript. Alko Inc. provided a red wine set consisting of 272 commercial red wines and part of this wine set was utilized in this study. Alko Inc. is gratefully acknowledged for their contribution, which greatly improved the quality of the present paper. The whole of Natural Chemistry Research Group is thanked for general help and discussion.

Funding

This study was supported by the University of Turku Graduate School (UTUGS).

Declaration of interest

None

Appendix A

Supplementary figures (Fig. S1–S4) and tables (Tables S1–S21).

References:

Alberts, P., Stander, M. A., & De Villiers, A. (2012). Advanced ultra high pressure liquid chromatography–tandem mass spectrometric methods for the screening of red wine anthocyanins and derived pigments. *Journal of Chromatography A*, *1235*, 92–102. https://doi.org/10.1016/j.chroma.2012.02.058

Alcalde-Eon, C., Escribano-Bailón, M. T., Santos Buelga, C., & Rivas-Gonzalo, J. C. (2006). Changes in the detailed pigment composition of red wine during maturity and ageing: A comprehensive study. *Analytica Chimica Acta*, *563*(1–2), 238–254. https://doi.org/10.1016/j.aca.2005.11.028

Arapitsas, P., Perenzoni, D., Nicolini, G., & Mattivi, F. (2012). Study of sangiovese wines pigment profile by UHPLC-MS/MS. *Journal of Agricultural and Food Chemistry*, *60*(42), 10461–10471. https://doi.org/10.1021/jf302617e

Bindon, K. A., Kassara, S., Hayasaka, Y., Schulkin, A., & Smith, P. A. (2014). Properties of wine polymeric pigments formed from anthocyanin and tannins differing in size distribution and subunit composition. *Journal of Agricultural and Food Chemistry*, *62*(47), 11582–11593. https://doi.org/10.1021/jf503922h

Blanco-Vega, D., Gómez-Alonso, S., & Hermosín-Gutiérrez, I. (2014). Identification, content and distribution of anthocyanins and low molecular weight anthocyanin-derived pigments in Spanish commercial red wines. *Food Chemistry*, *158*, 449–458. https://doi.org/10.1016/j.foodchem.2014.02.154

Bloemberg, T. G., Gerretzen, J., Wouters, H. J. P., Gloerich, J., van Dael, M., Wessels, H. J. C. T., van den Heuvel, L. P., Eilers, P. H. C., Buydens, L. M. C., & Wehrens, R. (2010). Improved parametric time warping for proteomics. *Chemometrics and Intelligent Laboratory Systems*, *104*(1), 65–74. https://doi.org/10.1016/j.chemolab.2010.04.008

Boido, E., Alcalde-Eon, C., Carrau, F., Dellacassa, E., & Rivas-Gonzalo, J. C. (2006). Aging effect on the pigment composition and color of *Vitis vinifera* L. Cv. Tannat wines. Contribution of

the main pigment families to wine color. *Journal of Agricultural and Food Chemistry*, *54*(18), 6692–6704. https://doi.org/10.1021/jf061240m

Casassa, L. F., Huff, R., & Steele, N. B. (2019). Chemical consequences of extended maceration and post-fermentation additions of grape pomace in Pinot noir and Zinfandel wines from the Central Coast of California (USA). *Food Chemistry*, *300*, 125147. https://doi.org/10.1016/j.foodchem.2019.125147

Dallas, C., Ricardo-da-Silva, J. M., & Laureano, O. (1996). Products formed in model wine solutions involving anthocyanins, procyanidin B2, and acetaldehyde. *Journal of Agricultural and Food Chemistry*, *44*(8), 2402–2407. https://doi.org/10.1021/jf940433j

Dangles, O., & Fenger, J. A. (2018). The chemical reactivity of anthocyanins and its consequences in food science and nutrition. *Molecules*, *23*, 1970. https://doi.org/10.3390/molecules23081970

De Freitas, V., & Mateus, N. (2011). Formation of pyranoanthocyanins in red wines: A new and diverse class of anthocyanin derivatives. *Analytical and Bioanalytical Chemistry*, *401*(5), 1467–1477. https://doi.org/10.1007/s00216-010-4479-9

Dimitrovska, M., Bocevska, M., Dimitrovski, D., & Murkovic, M. (2011). Anthocyanin composition of Vranec, Cabernet Sauvignon, Merlot and Pinot Noir grapes as indicator of their varietal differentiation. *European Food Research and Technology*, *232*(4), 591–600. https://doi.org/10.1007/s00217-011-1425-9

Dueñas, M., Fulcrand, H., & Cheynier, V. (2006). Formation of anthocyanin-flavanol adducts in model solutions. *Analytica Chimica Acta*, *563*(1–2), 15–25. https://doi.org/10.1016/j.aca.2005.10.062

Eilers, P. H. C. (2004). Parametric Time Warping. *Analytical Chemistry*, *76*(2), 404–411. https://doi.org/10.1021/ac034800e

Engström, M. T., Pälijärvi, M., Fryganas, C., Grabber, J. H., Mueller-Harvey, I., & Salminen, J.-P. (2014). Rapid qualitative and quantitative analyses of proanthocyanidin oligomers and polymers by UPLC-MS/MS. *Journal of Agricultural and Food Chemistry*, *62*(15), 3390–3399.

https://doi.org/10.1021/jf500745y

Harbertson, J. F., Picciotto, E. A., & Adams, D. O. (2003). Measurement of polymeric pigments in grape berry extract sand wines using a protein precipitation assay combined with bisulfite bleaching. *American Journal of Enology and Viticulture*, *54*(4), 301–306.

He, F., Liang, N. N., Mu, L., Pan, Q. H., Wang, J., Reeves, M. J., & Duan, C. Q. (2012). Anthocyanins and their variation in red wines I. Monomeric anthocyanins and their color expression. *Molecules*, *17*(2), 1571–1601. https://doi.org/10.3390/molecules17021571

He, J., Oliveira, J., Silva, A. M. S., Mateus, N., & De Freitas, V. (2010). Oxovitisins: A new class of neutral pyranone-anthocyanin derivatives in red wines. *Journal of Agricultural and Food Chemistry*, *58*(15), 8814–8819. https://doi.org/10.1021/jf101408q

Kennedy, J. A., & Taylor, A. W. (2003). Analysis of proanthocyanidins by high-performance gel permeation chromatography. *Journal of Chromatography A*, *995*, 99–107. https://doi.org/10.1016/S0021-9673(03)00420-5

Laitila, J. E., & Salminen, J.-P. (2020). Relevance of the concentrations and sizes of oligomeric red wine pigments to the color intensity of commercial red wines. *Journal of Agricultural and Food Chemistry*, *68*(11), 3576–3584. https://doi.org/10.1021/acs.jafc.9b07941

Laitila, J. E., Suvanto, J., & Salminen, J.-P. (2019). Liquid chromatography–tandem mass spectrometry reveals detailed chromatographic fingerprints of anthocyanins and anthocyanin adducts in red wine. *Food Chemistry*, *294*, 138–151. https://doi.org/10.1016/j.foodchem.2019.02.136

Malisch, C. S., Lüscher, A., Baert, N., Engström, M. T., Studer, B., Fryganas, C., Suter, D., Mueller-Harvey, I., & Salminen, J.-P. (2015). Large variability of proanthocyanidin content and composition in sainfoin (*Onobrychis viciifolia*). *Journal of Agricultural and Food Chemistry*, *63*(47), 10234–10242. https://doi.org/10.1021/acs.jafc.5b04946

Mateus, N., Oliveira, J., Santos Buelga, C., Silva, A. M. S., & De Freitas, V. (2004). NMR structure characterization of a new vinylpyranoanthocyanin-catechin pigment (a portisin). *Tetrahedron*

Letters, *45*(17), 3455–3457. https://doi.org/10.1016/j.tetlet.2004.03.007

Nave, F., Petrov, V., Pina, F., Teixeira, N., Mateus, N., & De Freitas, V. (2010). Thermodynamic and kinetic properties of a red wine pigment: catechin-(4,8)-malvidin-3-*O*-glucoside. *Journal of Physical Chemistry B*, *114*(42), 13487–13496. https://doi.org/10.1021/jp104749f

Pina, F. (1998). Thermodynamics and kinetics of flavylium salts Malvin revisited. *Journal of the Chemical Society, Faraday Transactions*, *94*(15), 2109–2116. https://doi.org/10.1039/a802602e

Rstudio Team. (2018). RStudio: Integrated Development Environment for R (1.1.463). RStudio, PBC, Boston, MA. http://www.rstudio.com/

Salas, E., Fulcrand, H., Meudec, E., & Cheynier, V. (2003). Reactions of anthocyanins and tannins in model solutions. *Journal of Agricultural and Food Chemistry*, *51*(27), 7951–7961. https://doi.org/10.1021/jf0345402

Salminen, J.-P. (2018). Two-dimensional tannin fingerprints by liquid chromatography tandem mass spectrometry offer a new dimension to plant tannin analyses and help to visualize the tannin diversity in plants. *Journal of Agricultural and Food Chemistry*, *66*(35), 9162–9171. https://doi.org/10.1021/acs.jafc.8b02115

Sanchez, G. (2012). *plsdepot: Partial Least Squares (PLS) Data Analysis Methods* (R package version 0.1.17). https://cran.r-project.org/package=plsdepot

Somers, T. C. (1971). The polymeric nature of wine pigments. *Phytochemistry*, *10*(9), 2175–2186. https://doi.org/10.1016/S0031-9422(00)97215-7

R Core Team. (2019). R: A language and environment for statistical computing (3.5.3). R Foundation for Statistical Computing, Vienna, Austria. http://www.r-project.org

Timberlake, C. F., & Bridle, P. (1976). Interactions between anthocyanins, phenolic compounds, and acetaldehyde and their significange in red wines. *American Journal of Enology and Viticulture*, *27*(3), 97–106.

Vidal, S., Cartalade, D., Souquet, J. M., Fulcrand, H., & Cheynier, V. (2002). Changes in

proanthocyanidin chain length in winelike model solutions. *Journal of Agricultural and Food Chemistry*, *50*(8), 2261–2266. https://doi.org/10.1021/jf011180e

Willemse, C. M., Stander, M. A., Vestner, J., Tredoux, A. G. J., & De Villiers, A. (2015). Comprehensive two-dimensional hydrophilic interaction chromatography (HILIC) × reversed-phase liquid chromatography coupled to high-resolution mass spectrometry (RP-LC-UV-MS) analysis of anthocyanins and derived pigments in red wine. *Analytical Chemistry*, *87*(24), 12006–12015. https://doi.org/10.1021/acs.analchem.5b03615

Zeng, L., Teissédre, P.-L., & Jourdes, M. (2016). Structures of polymeric pigments in red wine and their derived quantification markers revealed by high-resolution quadrupole time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, *30*(1), 81–88. https://doi.org/10.1002/rcm.7416

Figure captions:

Figure 1. Protocol for the formation of the two-dimensional (2D) quantile chromatographic fingerprints. In the lowest panel, the solid black line is the median and the grey area is the interval between 10^{th} and 90^{th} percentiles. The alignment of the chromatograms by parametric time warping algorithm (Bloemberg et al., 2010; Eilers, 2004) was done to correct small fluctuations in the retention times between individual wines.

Figure 2. Concentrations of the oligomeric PA–Mv^+, PA–methylmethine–Mv^+ and PA–Mv–FL adducts (A–C), and the relative proportions of small, medium-sized and large oligomeric adducts (SOA-%, MOA-% and LOA-%; D–F) in young, 1–2 years old wines (n=92). The wines were grouped based on the principal grape variety, except for the Pinot Noir wines, which were grouped based on their region (Pfalz, Germany; Beaune, France). Most wines were single-varietal wines except for the Touriga Nacional wines, all of which were made from Touriga Nacional, Touriga Franca and Tinta Roriz grapes. Refer to the supplementary table S1 for a summary of the wines and to tables S4–S12 for the pairwise statistical significance tests. Abbreviations: PA, proanthocyanidin; FL, flavan-3-ol monomer; Mv, malvidin, SOA, small oligomeric adduct; MOA, medium-sized oligomeric adduct; LOA, large oligomeric adduct.

Figure 3. Two-dimensional (2D) quantile chromatographic fingerprints of the oligomeric PA–Mv^+, PA–methylmethine–Mv^+ and PA–Mv–FL adducts in young, 1–2 years old wines. The groups Pfalz (Germany) and Beaune (France) consisted of Pinot Noir wine from the corresponding regions. The solid black lines are the medians and the grey areas are the intervals between 10^{th} and 90^{th} percentiles. The quantile fingerprints were formed using the protocol in Figure 1 to visualize and summarize the chromatographic information from 504 individual chromatograms. The individual compounds were identified based on their retention times and the compound were characterized in Laitila et al. (2019). Abbreviations: PA, proanthocyanidin; FL, flavan-3-ol monomer; Mv, malvidin glycoside, MvGlc, malvidin glucoside; MvAcGlc, malvidin acetylglucoside, MvCouGlc, malvidin coumaroylglucoside; Cat, catechin.

Figure 4. Evolution of the concentrations of the oligomeric PA–Mv^+, PA–Mv–FL and PA–methylmethine–Mv^+ adducts (A–C), and the evolution of the relative proportions of the small, medium-sized and larger oligomeric adducts (SOA-%, MOA-% and LOA-%; D–F) in Merlot, Shiraz and Cabernet Sauvignon red wines (n=95). Most wines were single-varietal wines with the Cabernet Sauvignon being the most used secondary grape in some Merlot wines. Refer to the supplementary table S2 for a summary of the wines and to tables S13–S20 for the pairwise statistical significance tests. Abbreviations: PA, proanthocyanidin; FL, flavan-3-ol monomer; Mv, malvidin, SOA, small oligomeric adduct; MOA, medium-sized oligomeric adduct; LOA, large oligomeric adduct.

Figure 5. Evolution of the two-dimensional (2D) quantile chromatographic fingerprints of the oligomeric PA–Mv^+, PA–methylmethine–Mv^+ and PA–Mv–FL adducts. The solid black lines are the medians and the grey areas are the intervals between 10^{th} and 90^{th} percentiles. The quantile fingerprints were formed using the protocol in Figure 1 to summarize the chromatographic information from 405 individual chromatograms. The individual compounds were identified based on their retention times and the compounds were identified in Laitila et al. (Laitila et al., 2019). Abbreviations: PA, proanthocyanidin; FL, flavan-3-ol monomer; Mv, malvidin glycoside, MvGlc, malvidin glucoside; MvAcGlc, malvidin acetylglucoside, MvCouGlc, malvidin coumaroylglucoside; Cat, catechin.

Figure 6. Results of the partial least squares regression models explaining the concentrations and the relative proportions of the large oligomeric adducts (LOA-%) of the PA–Mv^+ (A, D), PA–methylmethine–Mv^+ (B, E) and PA–Mv–FL adducts (C, F) in the 1-year-old wines (n=78). Refer to the supplementary table S3 for a summary of the 1-year-old wines. Abbreviations: PA, proanthocyanidin; FL, flavan-3-ol monomer; FL/PA, ratio of the concentrations of the monomeric flavan-3-ols and proanthocyanidins; Mv, malvidin glycoside; LOA-%, proportion of large oligomeric adducts; mDP, mean degree of polymerization of proanthocyanidins.