Supporting Information

for

Large variability of proanthocyanidin content and composition in sainfoin (Onobrychis viciifolia)

Carsten S. Malisch[†,#], Andreas Lüscher[*,†], Nicolas Baert[§], Marica T. Engström[§], Bruno Studer[#], Christos Fryganas[‡], Daniel Suter[†], Irene Mueller-Harvey[‡], Juha-Pekka Salminen[§]

† Forage Production and Grassland, Research Division Grassland Sciences and Agro-Ecosystems, Agroscope, Reckenholzstrasse 191, CH-8046 Zurich

Forage Crop Genetics, Institute of Agricultural Sciences, ETH Zurich, CH-8092 Zurich

§ Laboratory of Organic Chemistry and Chemical Biology, Department of Chemistry, University of Turku, FI-20500 Turku

‡ Chemistry and Biochemistry Laboratory, Food Production and Quality Division, School of Agriculture, Policy and Development, University of Reading, UK-RG6 6AT Reading

Corresponding author

*(A.L.) Phone: +41 58 468 7273. E-mail: andreas.luescher@agroscope.admin.ch

Table S1: Examined accessions, their status of cultivation, country of origin and ploidy level

Accession name	Status	Origin	Ploidy
247	Unknown	Morocco	Tetraploid
Brunner	Landrace	Austria	Unknown
Buceanskij	Unknown	Romania	Tetraploid
CPI 63750	Unknown	Turkey	Tetraploid
CPI 63764	Wild	Turkey	Unknown
CPI 63767	Cultivated	USA	Tetraploid
CPI 63780	Wild	Switzerland	Unknown
CPI 63810	Unknown	Lithuania	Tetraploid
CPI 63820	Unknown	Spain	Tetraploid
CPI 63825	Unknown	Spain	Unknown
CPI 63826	Unknown	Spain	Unknown
CPI 63854	Cultivated	Switzerland	Tetraploid
Esparsette	Cultivar	Poland	Tetraploid
Cholderton-Hampshire Common	Cultivated	UK	Tetraploid
La Rippe	Landrace	Switzerland	Unknown
NA / RCAT028437	Unknown	Hungary	Unknown
Nova	Cultivar	Canada	Tetraploid
Wiedlisbach	Ecotype	Switzerland	Unknown
Perdix	Cultivar	Switzerland	Unknown
Perly	Cultivar	Switzerland	Tetraploid
Premier	Landrace	Switzerland	Tetraploid
Rees “A”	Cultivar	UK	Tetraploid
Sarzens	Landrace	Switzerland	Unknown
Taja	Cultivar	Poland	Tetraploid
TU86-43-03	Cultivated	Turkey	Tetraploid
Visnovsky	Cultivar	Czech Republic	Tetraploid
WKT 10	Wild	Turkey	Tetraploid

Figure S1: Visual comparison of leaves of different ages, harvested at the same time from the same plant. Selection was mainly based on leaf size.

Calculation of mDP and highest average polymer sizes (maxmDP)

The mean degree of polymerization was calculated according to Engström (2014) [1], based on the ratio of extension units and terminal units for procyanidins (m/z 287 and 289 respectively) and prodelphinidins (m/z 303 and 305 respectively). The mDP was determined using these areas with the formula

$$\text{mDP} = \frac{0.37A_{287}+0.42A_{289}+2.15A_{303}+0.68A_{305}}{0.42A_{289}+0.68A_{305}} \quad \text{eqn S1}$$

Larger polymers elute at later retention times and consequently, when shifting the retention time window from the entire elution period to only a fraction at the end, the share of larger polymers is increased. For example, in Figure S2, when calculating the mDP based upon different retention time windows, a window of 0-2 minutes yields an mDP of 2, whereas the retention time window of 2-4 minutes already yields an mDP of 15 and from 4-6 minutes, the mDP is 33. Consequently, for the calculation of the on average largest polymer sizes, we chose the retention time window from 3.70 minutes, as this allowed for a sufficiently strong signal intensity while at the same time focusing on the larger polymer sizes.

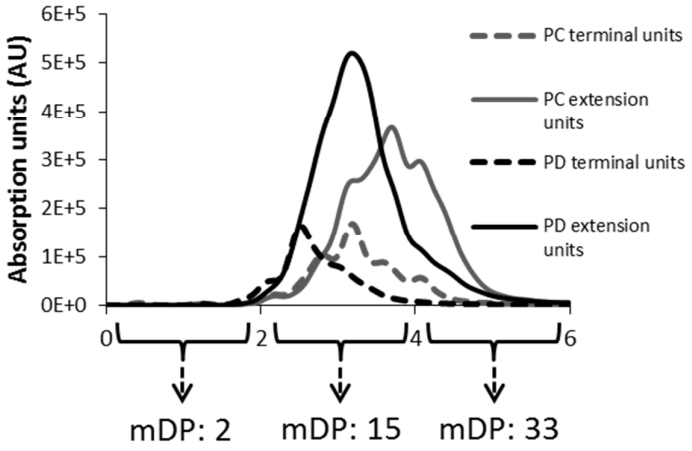

Figure S2: Fingerprint of PC and PD terminal and extension units, as compared by their elution time.

Calculation for amount of proanthocyanidins

The amount of proanthocyanidins in individual plants was calculated as

$$PA_{amount} = (([PA]_{leaves} \times lr) + ([PA]_{leaves} \times (\overline{[PA]}_{stems} / \overline{[PA]}_{leaves})) \times sr) \times pDM \quad \text{eqn S2}$$

where $[PA]$ was the concentration of proanthocyanidins of an individual plant, $\overline{[PA]}$ was the average of the proanthocyanidin concentrations in all plants, lr was the leaf ratio, sr was the stem ratio and pDM the dry matter weight of the whole plant. Leaf and stem ratio were the weight fraction of the whole plant in percent, which belonged to either leaves or stems. By dividing the average PA concentration in stems, $\overline{[PA]}_{stems}$, with the average PA concentration in leaves, $\overline{[PA]}_{leaves}$, a general coefficient was calculated. This in turn was used to estimate the PA concentration in stems of individual plants, as the stem concentration of proanthocyanidins was only measured in a subset of seven accessions. By multiplying the

concentration with the biomass formed, and adding the values for leaves and stems, the overall PA amount of the plant was determined.

Identification of the phenolic compounds

A Xevo UPLC system, coupled with a diode array and MS/MS detector was used for the identification of the phenolic compounds. For previously identified compounds [2-4], Multiple Reaction Monitoring (MRM) methods were utilized for quantification. If compounds had not been previously reported, UV chromatograms at 280 nm and 349 nm were used as a first step for detection. For final identification, a combination of *m/z* values and retention time from the MS data together with the UV-spectrum yielded further information. To increase the certainty in identification, the fractions obtained with Sephadex LH-20 purification were analyzed with tandem mass spectrometry via daughter ion scans. All identified compounds are listed in Table 2, with the main important criteria that led to their identification. Compounds which were compared against an available standard (in retention time, UV and MS/MS spectrum) and are therefore considered to be formally identified are marked as such in Table 1. All other compounds were characterized in accordance with the existing literature, based on UV and MS/MS spectrum and, where possible, the order of elution. Compounds which were quantified are also denoted with their relative standard deviation from replicate analysis. The detected compounds were then utilized to determine the variability of PAs, as well as the simple phenolics in 27 accessions of sainfoin on an individual plant level.

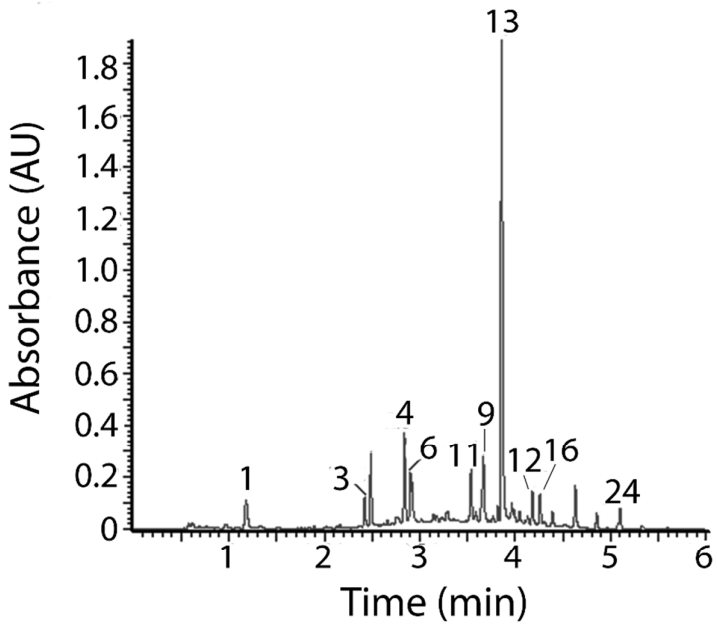

Figure S3: UPLC- chromatogram (at 280 nm) of a typical sainfoin cultivar with main compounds identified (in order of elution: 1: arbutin; 3: caffeoylquinic acid; 4: chlorogenic acid; 6: Quercetindihexoside; 11: Quercetin-3-*O*-rhamnosylrutinoside; 9: coumaric acid glucoside; 13: quercetin-3-*O*-rutinoside; 12: Kaempferol-3-O-rhamnosylrutinoside; 16: Isorhamnetin-3-*O*-rutinoside; 24: Quercetin 3-*O*-(4′′′-*O*-E-feruloyl)-α-rhamnopyranosyl-(1′′′ → 2′′)[α-rhamnopyranosyl-(1′′ ′′ → 6′′)]-β-glucopyranoside) (see also Table 1)

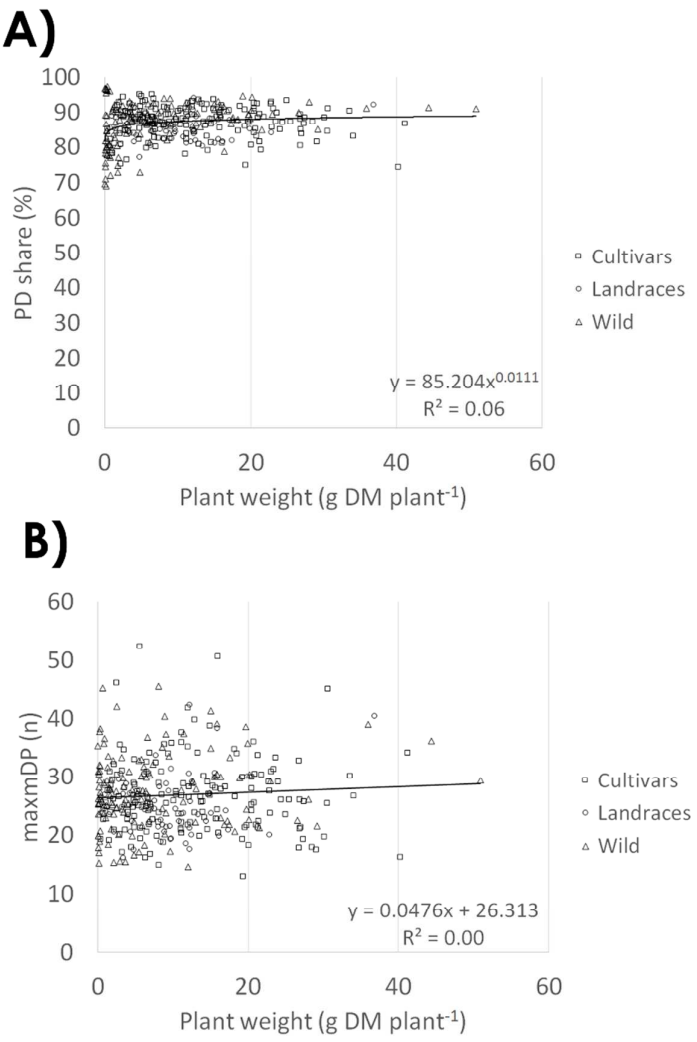

Figure S4: A) PD share in leaves compared against the plant yield, and B) maximum degree of polymerization (maxmDP) compared to plant yield. Equation for exponential trendline and its regression analysis are denoted at the bottom right

References

(1) Engström, M.T.; Pälijärvi, M.; Fryganas, C.; Grabber, J.H.; Mueller-Harvey, I.; Salminen, J.-P., Rapid Qualitative and Quantitative Analyses of Proanthocyanidin Oligomers and Polymers by UPLC-MS/MS. *Journal of Agricultural and Food Chemistry* **2014,** *62*, (15), 3390-3399.

(2) Regos, I.; Urbanella, A.; Treutter, D., Identification and Quantification of Phenolic Compounds from the Forage Legume Sainfoin (Onobrychis viciifolia). *Journal of Agricultural and Food Chemistry* **2009,** *57*, (13), 5843-5852.

(3) Veitch, N.C.; Regos, I.; Kite, G.C.; Treutter, D., Acylated flavonol glycosides from the forage legume, Onobrychis viciifolia (sainfoin). *Phytochemistry* **2011,** *72*, (4-5), 423-429.

(4) Lu, Y.R.; Sun, Y.; Foo, L.Y.; McNabb, W.C.; Molan, A.L., Phenolic glycosides of forage legume *Onobrychis viciifolia*. *Phytochemistry* **2000,** *55*, (1), 67-75.