

# A non-destructive method to quantify leaf starch content in red clover

**Lea Antonia Frey**

ETH Zurich

**Philipp Baumann**

ETH Zurich

**Helge Aasen**

ETH Zurich

**Bruno Studer**

Eidgenössische Technische Hochschule Zurich

**Roland Kölliker** (✉ [roland.koelliker@usys.ethz.ch](mailto:roland.koelliker@usys.ethz.ch))

ETH Zurich <https://orcid.org/0000-0003-1959-0402>

---

## Research

**Keywords:** Red clover, starch content, hyperspectral imaging, partial least square regression, forage quality, grassland

**Posted Date:** January 14th, 2020

**DOI:** <https://doi.org/10.21203/rs.2.20769/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

Grassland-based ruminant livestock production provides a sustainable alternative to intensive production systems relying on concentrated feeds. However, grassland-based roughage often lacks the energy content required to meet the productivity potential of modern livestock breeds. Forage legumes, such as red clover, with increased starch content could partly replace maize and cereal supplements. However, breeding for increased starch content requires efficient phenotyping methods. This study is unique in evaluating a non-destructive hyperspectral imaging approach to estimate leaf starch content in red clover for enabling efficient development of high starch red clover genotypes.

## Results

We assessed prediction performance of partial least square regression models (PLSR) and validated model performance with an independent test set. Starch content of the training set ranged from 0.1 to 120.3 mg g<sup>-1</sup> DW. The best cross-validated PLSR model explained 56% of the measured variation and yielded a root mean square error (RMSE) of 17 mg g<sup>-1</sup> DW. Model performance decreased when applied to the independent test set (RMSE = 29 mg g<sup>-1</sup> DW, R<sup>2</sup> = 0.36). Different filtering methods did not increase model performance.

**Conclusion** The non-destructive spectral method presented here, provides a tool to detect large differences in leaf starch content of red clover. Breeding material can be sampled and selected according to their starch content without destroying the plant.

## Background

Temporary and permanent grassland account for roughly 70% of agricultural land and plays a significant role in sustainable agriculture worldwide by providing roughage for ruminant livestock production. Pasture and grassland-based agroecosystems maintain carbon balances, nutrient cycles, biodiversity, and water quality [1]. However, they were gradually replaced by intensified production systems where the high feed energy content required by today's livestock breeds is largely covered through starch from cereals and maize. Starch is an important form of assimilated carbohydrates in plants, which diurnally accumulates in the leaf and is nocturnally mobilized to support growth [2–4]. The accumulation of starch and its linear degradation at night is thought to be crucial for stable growth and directly correlated to plant biomass [5, 6]. However, plant biomass and leaf starch content are not always negatively correlated in species such as birdsfoot trefoil (*Lotus japonicus* L.) or red clover (*Trifolium pratense* L.) [7, 8]. Starch accumulation and degradation varies not only among plant species but also among genotypes, seasons, and management regimes [9–12].

Red clover is one of the most important forage legumes in temperate climates [13]. Its high yield potential, high crude protein content and high digestibility make it an excellent feed, not only for cattle but

also for other livestock and poultry [14, 15]. Red clover has the potential to accumulate up to one third of its leaf dry mass as starch, and some genotypes degrade less than 50% of their starch during the night [8]. Thus, selecting for red clover plants with high starch content and low degradation rates is likely to result in high starch cultivars. These could provide an alternative, high energy feed source, which would significantly improve sustainability of ruminant livestock production.

Developing a high starch red clover requires a better understanding of the starch metabolism in red clover and an efficient method to quantify starch in leaf tissue. Starch is commonly quantified with an enzymatic method, where leaf samples are flash frozen, ground and weighted before extraction is performed [16]. This procedure is laborious, expensive, and involves destructive sampling. A non-destructive method to measure leaf starch content would enable detailed studies of starch turnover in red clover plants. Dynamic changes during plant development could be traced throughout the entire season. Specifically, different genotypes could be investigated under different management regimes and across different environments over an extended period of time. Furthermore, this method could be applied to develop high starch red clover cultivars.

Hyperspectral imaging and near infrared spectroscopy (NIRS) are routinely used to estimate biochemical compounds such as lignin, cellulose, starch, sugars and proteins in numerous crops [17–19]. These two methods have largely replaced wet chemistry as the standard analytical procedure for detection and quantification of plant biochemical compounds in the food industry [20, 21]. Infrared spectra result from the fundamental vibrational absorptions of photons in the mid-infrared region ( $500\text{--}4000\text{ cm}^{-1}$ ,  $350\text{--}25000\text{ nm}$ ) by bonds within specific functional groups of molecules. These absorptions are mirrored to the NIR region [21]. Multivariate statistics, chemometrics, or machine learning methods are then used to quantify and classify specific compounds or properties [22]. NIRS or other spectral techniques are most accurate when using dried and homogenized (i.e. milled) plant material. For example, starch has been accurately quantified on dried cotton leaves or dry forage maize using NIRS ( $R^2 > 0.9$ ) [19, 23–25]. Estimating chemical compounds with spectral measurements on fresh leaf tissue is often less reliable due to masking effects of light absorption by the cuticle or the leaf water content [26, 27]. For successful spectroscopy-based diagnostics using fresh leaf tissue, spectral pre-processing and statistical modeling are essential to at least partially correct for confounding effects [26, 27].

The following study aimed at developing a non-destructive spectroscopic method to estimate leaf starch content in fresh leaf tissue of red clover, enabling to monitor starch turnover on the same genotype over a longer period, under different management regimes, and under various environmental conditions.

## Methods

Leaf starch was determined in two independent sets (i.e., a training set and a test set) of plants grown in two separate experiments. Wet laboratory measurements were taken on exactly the same material as used for the hyperspectral measurements.

# Plant material and growth conditions

Thirty red clover plants were used for the study (Additional file 1: Table S1). All plants were clonally propagated, using cuts that contained only one shoot and one root meristem to ensure comparable physiological states of all plants. These clonally propagated plants were grown in a climate chamber for 90 days before harvesting (3:2:1 soil: peat: perlite substrate, photoperiod of 14:10-h L:D, day temperature  $20 \pm 2$  °C, night temperature  $15 \pm 2$  °C and relative humidity,  $60 \pm 10\%$ ). Samples were taken at the end of the night (EN; before lights were turned on), and at the end of the day (ED; before lights were turned off). The training set included 18 genotypes, six thereof clonally duplicated, resulting in a total of 24 plants (Additional file 1: Table S1). Starch was measured either at EN or at ED and 15 leaf cuts were measured per plant, taken on all three leaflets of each leaf. The youngest, fully emerged leaf (y), the oldest leaf (o) and three intermediate leaves (m) were sampled (Fig. 1), resulting in 360 measurements.

The test set included six plants, three harvested at ED and three at EN (Additional file 1: Table S1). Ten leaf cuts were taken per plant on mature leaves, resulting in 60 measurements.

## Leaf spectroscopy

Leaflets were cut using a round, sharpened tube with a diameter of 12 mm to standardize leaf area (Additional file 2: Fig. S1). These leaf cuts were placed on the matt black surface of the FieldSpec4 pro device (Analytical Spectral Devices, Boulder, Colorado, USA). Radiance between 350 and 2500 nm was measured. The spectrometer's contact probe was fixed on a clamp, and the sample was placed so that no light escaped through the sides. Leaf samples were referenced to a spectralon white reference every fifth recording and the radiance measurements were transformed to reflectance. Immediately after taking spectral measurements, leaf cuts were flash frozen in liquid nitrogen and freeze-dried for 48 h.

After taking spectral measurements, whole plants from the training set were cut 2 cm above ground, flash-frozen and freeze-dried for starch quantification.

## Wet lab analysis for starch quantification

Leaf cuts and whole plants were quantified as described by Ruckle et al. [8]. Two additional clones of one genotype, not included in the correlation model, were iodine stained to visualize the starch pattern within a plant. Plants were harvested either at ED or at EN, washed with tap water and placed in 80% (v / v) boiling ethanol. After two hours, when plants were transparent, they were removed and placed in Lugol's solution. After 10 minutes, the Lugol's solution was rinsed off to destain the non-target areas. The plants were photographed on a light-table.

## Statistical analyses

Statistical analyses were performed using the R statistical software version 3.6.0 [28]. Significance testing was performed using the library MASS [29]. Assumptions of homoscedasticity of variances and normality of residuals were not met, therefore an exact Wilcoxon rank sum test at  $\alpha = 5\%$  was performed.

# Pre-processing of spectral data

Spectral analysis was realized using the R package `simplerspec` [30]. The mean reflectance values of 10 measurements per sample were used. Leaf spectra were pre-processed prior to modelling. Gaps between the different detector arrays at  $\lambda = 1000$  nm and at  $\lambda = 1800$  nm were splice corrected. Spectra were smoothed with the Savitzky-Golay first derivative filter using a 3rd-order polynomial at a 21-point window (42 nm at a spectrum interval of 2 nm; R package `prospectr` [31]). Spectral pre-processing is crucial to reducing significant noise and baseline drift resulting from light scatter before establishing a correlation model. Partial least square regression (PLSR) is a substantial chemometric method, which can cope with multicollinearity in spectra and delivers robust calibration models with many predictors and few observations [32, 33].

## Model development

Leaf reflectance data from the training set was modelled by PLSR [32], using either raw or pre-processed spectra as predictors. A 5-times repeated 10-fold cross-validation scheme was used to fit the models, to determine the best number of components ( $n_{comp}$ ), and to estimate model performance of the final model. A constant random seed was set for resampling, yielding identical hold out data across all models. Model reliability was assessed by coefficient of determination ( $R^2$ ) and the root mean square error (RMSE, Eq. 1). RMSE of the held-out sets was calculated by aggregating all repeated k-fold cross-validation predictions ( $\hat{y}_i$ ) and corresponding observed values ( $y_i$ ) grouped by  $n_{comp}$ .

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (1)$$

Variable influence on projection scores (VIP) is a measure of variable importance tailored to PLS regression [34, 35]. VIP scores were calculated from the PLS regression parameters taking multicollinearity into account, which is likely to occur because of the nature of spectroscopic data. VIP scores are considered as a robust measure to identify relevant predictors, here important wavelengths. A variable with VIP above 1 contributes more than the average variable to the model prediction. The VIP value  $v_j$  was calculated for each wavelength variable  $j$  as

$$v_j = \sqrt{p \sum_{a=1}^A [SS_a (w_{aj} / \|w_{aj}\|)^2] / \sqrt{\sum_{a=1}^A (SS_a)}} \quad (2)$$

where  $w_{aj}$  are the PLS regression weights for the  $a^{\text{th}}$  component for each of the wavelength variables and  $SS_a$  is the sum of squares explained by the  $a^{\text{th}}$  component (Eq. 2). The sum of squares  $SS_a$  for the  $a^{\text{th}}$  component was calculated from the score  $q_a$  of the predicted variable  $y$  and the  $t_a$  scores of the spectral matrix  $X$  (Eq. 3):

$$SS_1 = q_a^2 t_a^T t^a \quad (3)$$

VIP scores were also used to filter important predictors. Predictors were screened with a threshold of VIP > 1 within the training set, and the identified predictors were used to re-calibrate the test set and assess performance. This separation in the VIP filtering by independent tests was needed to avoid overfitting and over-optimistic assessment that typically occurs when identifying subsets of features on the modelling data.

In addition to the VIP based filtering, two other procedures were applied for wavelength selection. First, the 50 most relevant wavelengths to estimate starch according to Pearson's correlation coefficient ( $r$ ) were taken to re-perform PLS regression. Second, four wavelengths that were assigned to starch in previous literature were taken and normalized with the reflection at the wavelength that had the smallest standard deviation across the entire wavelength range, prior to performing a multiple linear regression (MLR; [22]).

## Model evaluation using the test set

The best training model tuned by cross-validation and refitted on all training data, and the training models with three different wavelength selection (filtering) methods were tested on the independent test data set (60 samples). The predictive ability of these final models was again evaluated using  $R^2$  and RMSE on the test set. Besides these test set predictions, a PLSR model was re-calibrated using only the test data. This re-calibration allowed to determine whether the test set possibly contained different or differently weighted spectral features relevant for starch prediction, so that PLSR training relationships did not generalize to this independent test experiment.

## Results

### Wet lab analysis for starch quantification

Starch concentration of samples for the training harvested at ED ranged from 2.0 to 120.3 mg starch per g dry weight (DW), with a median of 46.3 mg g<sup>-1</sup> DW. For the samples harvested at the end of the night (EN), starch concentration ranged from 0.1 to 47.8 mg g<sup>-1</sup> DW, with a median of 9.6 mg g<sup>-1</sup> DW. Starch concentration for the test set ranged from 26.41 to 125.44 mg g<sup>-1</sup> DW for ED harvested samples, with a median of 66.18 mg g<sup>-1</sup> DW. Plants harvested at EN had lower starch concentrations, ranging from 3.66 to 79.51 mg g<sup>-1</sup> DW, with a median of 23.28 mg g<sup>-1</sup> DW. Differences between samples harvested at ED and EN were statistically significant ( $p < 0.5$ ) for both sets (Fig. 2).

The iodine stained leaves displayed analogous patterns. Plants harvested at ED showed higher starch accumulation than plants harvested at EN, indicated by a dark coloration of the leaves (Fig. 3A, B). Differences in coloration were not only visible between diurnal time points, but also across and within

leaves (Fig. 3B, C). Dark coloration indicated that old leaves accumulated more starch than young ones. Coloration varied within a plant, showing a clear pattern with young leaves accumulating less starch compared to old leaves. While starch accumulation varied within one leaf, no clear pattern was distinguishable based on iodine visualization when observing the starch accumulation within individual leaflets (Fig. 3C). Observations in the iodine staining were confirmed by starch quantification in different leaf types. Young leaves had a significantly lower ( $p < 0.05$ ) starch concentration when compared to mature leaves, old leaves or whole plants (Fig. 4). Some differences between genotypes of the training set harvested at the end of the day were statistically significant ( $p < 0.05$ ), but a high variation within genotypes was observed (Fig. 5).

## Spectral measurements and modelling

The average reflectance spectra of the training set revealed similar patterns for both harvest time points ED and EN (Fig. 6A). VNIR/SWIR (350 nm-2500 nm) spectra had three main absorption regions, around the absorption bands of 700 nm, 1400 nm, and 1900 nm. Savitzky-Golay smoothed spectra showed very similar patterns across the entire wavelength range for samples harvested at ED and those harvested at EN (Fig. 6B).

The best PLSR training model with pre-processed spectra resulted in an accurate starch prediction for the training set ( $R^2 = 0.72$ ,  $RMSE = 13 \text{ mg g}^{-1} \text{ DW}$ ), using nine PLS components (Fig. 7A). Five times repeated 10-fold cross validation performed on the same data set revealed a moderate correlation coefficient ( $R^2_{CV}$ ) of 0.56 and a  $RMSE_{CV}$  of  $17 \text{ mg g}^{-1}$  (Fig. 7B).

PLSR modelling using pre-processed spectra performed better than modelling using raw spectra as predictors (Additional file 3: Fig. S2). Separating cross-validated predictions by ED and EN resulted in lower correlation coefficients of  $R^2_{CV} = 0.39$ , and  $R^2_{CV} = 0.25$  for ED and EN, respectively (Additional file 4: Fig. S3). Including only the most relevant wavelengths for estimating starch content, first based on filtering training variable importance in the projection ( $VIP > 1$ ), second based on the 50 top starch correlated wavelengths (PCC), and third based on MLR divided by the minimum standard error of reflectance did not increase model performance substantially (Table 1).

## Model evaluation with an independent test set

Independent test set predictions ( $n = 57$ ) using the best training PLSR model calibrated with pre-processed spectra ( $n = 337$ ,  $n_{comp} = 9$ , all wavelengths) yielded a substantially lower  $R^2$  of 0.36 and larger  $RMSE$  of  $29 \text{ mg g}^{-1} \text{ DW}$  (Fig. 8). The three training models calibrated with variable selection ( $VIP > 1$ , top 50 correlations, and MLR with normalized assigned starch bands) resulted in inferior accuracy when applied to the test set (Table 1).

Table 1 Model performance using different filtering methods

Model strategy	Training set		Test set		Description
	R <sup>2</sup>	RSME	R <sup>2</sup>	RSME	
VIP > 1	0.58	16.6	0.37	38	Filtering based on importance of the training variable in the projection (VIP)
PCC	0.53	18.5	0.37	32	Filtering based on the top 50 with starch correlated wavelengths
MLR	0.26	21.8	0.18	38	Reflectance at 556 nm, 702 nm, 1300 nm, and 1960 nm divided by reflectance at 670 nm (minimum standard deviation)

Model development using different filtering methods such as variable importance in the projection (VIP), the top 50 starch correlated wavelengths (PCC), multiple linear regression (MLR) before performing partial least square regression (PLSR). Best model performance of each filtering method determined by five times repeated 10-fold-cross validation was used to estimate leaf starch content of an independent test set

## Discussion

Hyperspectral imaging on dry homogenized material is a widely used and well established technique, but applying this method on fresh plant material is not yet a standard analytical procedure [26, 27]. Consequently, the correlation of hyperspectral measurements and wet lab results for starch reported in this study were clearly lower, when compared to NIRS measurements on dried plant material, where coefficients of determination ( $R^2$ ) reached 0.99 for nitrogen and starch content of cotton leaves [19]. One problem using fresh leaf material is the water in the fresh leaves. Water is a strong absorber of the infrared radiation and predominant bands are in the region between 1300 nm and 2500 nm, where important wavelengths were present in this study [26]. It is therefore likely, that water absorption masked the absorption bands of starch molecules, impairing prediction of starch content [26, 27]. Not only water absorption can obscure the starch absorption characteristics, but also the cell structure of fresh plants scattering light as it passes through multiple air and water boundaries. Furthermore, the distribution of starch in fresh leaves is not uniform with respect to the organization of cells and organelles [26]. The problems associated with the prediction of starch content in fresh leaves might be reduced, if spectral data is preprocessed [35]. Indeed, preprocessing of the spectra considerably improved predictive accuracy compared to unprocessed reflectance spectra (Additional file 3: Fig. S2), by removing systematic variation in spectra such as light scattering and thereby increasing the signal to noise ratio [35].

Total starch concentration of the plant material in the training was between 0.2% and 12% for plants harvested at ED and ranged from 0.01% up to 5% for the plant material harvested at EN. The starch concentrations of the training set were slightly lower than the concentrations of the test set. The total

starch concentration was substantially lower than the ones published by Ruckle et al. [8], where leaf starch concentration ranged from 6% up to 35% for ED harvested plants. This difference occurred most likely due to different growing conditions, since light and temperature have a high impact on starch accumulation.

Many studies have shown, that starch contents highly depend on the diurnal cycle [3, 5, 37, 38]. An over 3-fold difference in starch content was observed between ED harvested plants, and plants harvested at EN (Fig. 4). Mean genotypic differences for the training set ranged from 31.6 to 59.7 mg g<sup>-1</sup> DW for the ED harvested plants (Fig. 5) and from 2.2 to 37.6 mg g<sup>-1</sup> DW for the EN harvested plants (data not shown), respectively, showing high variation within genotypes. The best PLSR training model explained 56% of the measured starch variation with an RSME of 17 mg g<sup>-1</sup> DW. Consequently, differences between harvest time points and differences between extreme genotypes can be predicted with our model. Nevertheless, this 56% of starch variation explained by the model is lower than found by Shorten et al. [39], using hyperspectral imaging systems (550 nm – 1700 nm) to estimate more than ten different quality compounds in perennial ryegrass (*Lolium perenne* L.). Low and high weight sugars were estimated separately and best model prediction for the high weight sugars using PLS regression resulted in an R<sup>2</sup> of 0.68 and a RMSE of 19.9 mg g<sup>-1</sup>. Splitting two-third of the data for calibration and using the remaining data for validation resulted in a slightly lower model performance (R<sup>2</sup> = 0.63 and RMSE of 21.6 mg g<sup>-1</sup>) [39]. Variable importance in the projection (VIP) did not considerably improve model performance. This is in contrast to comparable studies where selecting important wavelengths improved model accuracy and reduced the redundancy effects of wavelengths, which had low weight in the model [35, 36]. Our model results indicate that the selected regions were not sufficient to explain the data with equal effectiveness and that many features are important for prediction. The wavelengths near 550 nm, 770 nm, 850 nm and 1920 nm and from 1650 nm to 1850 nm had high model contribution for estimating starch content (VIP assessment, Additional file 5: Fig. S4). Starch absorbance in fresh leaves was previously shown to be associated with wavelengths in the regions of 556 nm, 702 nm, 1300 nm and 1960 nm [28]. These absorptions partly corresponded with the VIP patterns across wavelengths. However, many additional wavelength ranges had high model importance for starch prediction.

A model built from a single set of training observations is often not adequate to predict an independent data set [32, 33]. If a model is tested on the same data that was used to fit the model, performance is often overestimated [32]. Our study showed that cross-validated training resulted in a substantial overfitting and high starch contents were particularly underestimated (Fig. 7). The independent second dataset from a second experiment (test set) allowed us to further validate model performance, in addition to cross-validation during training. As expected, the test prediction resulted in a 1.7-fold increase in RMSE (Fig. 8). Moreover, models including only a subset of wavelengths were validated on the test set, resulting in an even lower predictability (Table 1). The VIP analysis of the two independent sets (training and test) indicated that some important wavelengths regions occurred in both sets, but with different VIP magnitudes (Additional files 5, 6: Figs. S4, S5). Further, the training model had important features between 500 nm and 750 nm, whereas the re-calibrated test model had important wavelengths below

500 nm. These differences in VIP magnitudes and the additional regions relevant for prediction partly explain the poorer prediction performance of the test set when applying the training model. Despite the fact that two of the three genotypes from the test set were included in the training set, the spectra and models had only limited generalization capacity for starch contents.

Recalibration using only test data led to a slight decrease in RMSE compared to test prediction, but this substantially reduced bias. Thus, a new calibration may be needed for each independent trial or the current red clover starch spectral library needs to be augmented with more measurements from different independent trials with both genotypic and phenotypic variance in starch. Various environmental growth conditions influence starch accumulation and can thereby mask genotypic effects [38]. Hence, we suggest follow up research to ultimately verify whether a separate spectral model is required for each independent trial to improve predictive accuracy under substantial genotype x environment interaction. Despite the relatively low prediction accuracy, performance of the best PLSR training model was sufficient to detect differences between red clover genotypes with very high or very low levels of starch content. Therefore, the method developed has the potential to contribute to the breeding of high energy red clover cultivars.

The success of breeding forage crops with increased energy content was previously demonstrated by breeding perennial ryegrass cultivars with high levels of water-soluble carbohydrates (WSC). These WSC cultivars can substantially increase animal performance and nitrogen use efficiency in pasture-based animal production systems [40]. Red clover and ryegrasses are often cultivated in mixtures, not only due to their attractive diet composition, but also due to the transfer of N between species. In addition, grass-clover mixtures require fewer pesticide and herbicide applications, and protect soils against erosion [41, 42]. Therefore, high starch red clover cultivars in mixtures with high WSC ryegrasses appear a particularly promising option, which brings us one-step closer towards an environmental sustainable feed production meeting the high-energy requirements of modern livestock production.

## Conclusion

This study is unique in developing and testing a non-destructive method to predict leaf starch content in red clover plants. The described method is suitable to differentiate between high and low starch content in red clover genotypes. Unfortunately, model performance is not sufficient to trace small changes in starch accumulation. Therefore, the method is only partially suited to monitor starch metabolism in detail or to investigate the effect of environmental influences or management regimes. We suggest follow up studies to enlarge the current red clover starch spectral library by means of additional measurements from different independent trials, covering both genotypic and phenotypic variation in starch. Currently, the level of resolution is sufficient for the method to differentiate high variance in starch and thus, can be integrated into existing breeding programs to get a first impression on starch levels of different red clover cultivars.

## Abbreviations

**ED:** at the end of the day

**EN:** at the end of the night

**DW:** dry weight

**PLSR:** partial least square regression

**VIP:** variable importance in the projection

**PCC:** top 50 starch correlated wavelengths

**MLR:** multiple linear regression

## **Declarations**

### **Ethics approval and consent to participate**

Not applicable

### **Consent for publication**

Not applicable

### **Availability of data and materials**

Raw data and scripts can be accessed via:

[https://zenodo.org/record/3598699#.XhOLvdko\\_d4](https://zenodo.org/record/3598699#.XhOLvdko_d4)

### **Competing interests**

The authors declare that they have no competing interests

### **Funding**

This research was partially supported by the Coop–Research Fellowship Program through the ETH–World Food System Center (WFSC). The funding body assisted in conceiving the study and in designing the experiments.

### **Author contribution statement**

LAF performed the research, analyzed the wet lab data and drafted the manuscript. PB analyzed the spectral data and HA assisted with data analysis. PB and HA helped to improve the manuscript. RK and BS supported the research design and helped interpreting the results and drafting the manuscript. All authors read and approved the final version of the manuscript.

## Acknowledgements

We thank Dr. Michael Ruckle for initiating the research on high starch red clover and for helping with the wet lab starch analyses, Prof. Dr. Achim Walter for the allocation of the FieldSpec4 pro device and Dr. Dániel Carrera and Ms Verena Knorst for the excellent technical support.

## References

1. Steinfeld H, Gerber P, Wassenaar T, Castel V, Rosales M, De Haan: C. Livestock's long shadow. Rome: FAO Rome; 2006
2. Stettler M, Eicke S, Mettler T, Messerli G, Hörtensteiner S, Zeeman SC. Blocking the metabolism of starch breakdown products in Arabidopsis leaves triggers chloroplast degradation. Mol Plant 2009, 2:1233-1246
3. Geiger DR, Servaites JC. Diurnal regulation of photosynthetic carbon metabolism in C3 Plants. Annu Rev Plant Physiol Plant Mol Biol. 1994, 45:235-256
4. Stitt M, Gibon Y, Lunn JE, Piques M. Multilevel genomics analysis of carbon signalling during low carbon availability: Coordinating the supply and utilisation of carbon in a fluctuating environment. Funct Plant Biol. 2007, 34:526-549
5. Graf A, Schlereth A, Stitt M, Smith AM. Circadian control of carbohydrate availability for growth in Arabidopsis plants at night. Proc Natl Acad Sci USA 2010, 107:9458-9463
6. Mugford ST, Fernandez O, Brinton J, Flis A, Krohn N, Encke B, et al. Regulatory properties of ADP glucose pyrophosphorylase are required for adjustment of leaf starch synthesis in different photoperiods. Plant Physiol 2014, 166:1733-1747
7. Vriet C, Welham T, Brachmann A, Pike M, Pike J, Perry J, et al. A suite of *Lotus japonicus* starch mutants reveals both conserved and novel features of starch metabolism. Plant Physiol 2010, 154:643-655
8. Ruckle ME, Meier MA, Frey L, Eicke S, Kölliker R, Zeeman SC, et al. Diurnal leaf starch content: An orphan trait in forage legumes. Agronomy 2017, 2:1-15
9. Liu W, Su J, Li S, Lang X, Huang X. Non-structural carbohydrates regulated by season and species in the subtropical monsoon broad-leaved evergreen forest of Yunnan Province, China. Sci Rep 2018, 8:1-10
10. Moraes MG, Chatterton NJ, Harrison PA, Filgueiras TS, Figueiredo-Ribeiro RCL. Diversity of non-structural carbohydrates in grasses (*Poaceae*) from Brazil. Grass Forage Sci 2013, 68:165-177
11. Griggs TC, MacAdam JW, Mayland HF, Burns JC. Temporal and vertical distribution of nonstructural carbohydrate, fiber, protein, and digestibility levels in orchardgrass swards. Agron J 2007, 99:755-763
12. Pelletier S, Tremblay GF, Bélanger G, Bertrand A, Castonguay Y, Pageau D, et al. Forage nonstructural carbohydrates and nutritive value as affected by time of cutting and species. Agron J 2010, 102:1388-1398

13. Taylor NL. A century of clover breeding developments in the United States. *Crop Sci* 2008, 48:1-13
14. Halling MA, Hopkins A, Nissinen O, Paul C, Tuori M, et al. Forage legumes – productivity and composition Magnus. *Landbauforsch Voelkenrode, Sonderh 234* 2001, 8-9
15. Broderick GA. Desirable characteristics of forage legume for improving protein utilisation in ruminants. *J Anim Sci.* 1995;73:2760–73
16. Hostettler C, Kölling K, Santelia D, Streb S, Kötting O, Zeeman SC. Chloroplast research in *Arabidopsis*. *Chloroplast Res Arab Methods Protoc* 2011, 775:387-410
17. Goetz AFH, Gao BC, Wessman CA and Bowman WD. By spectrum matching techniques. *Brows Conf , Geosci Remote Sensing* 1990, 971-974
18. Yoder BJ, Pettigrew-Crosby RE. Predicting nitrogen and chlorophyll content and concentrations from reflectance spectra (400-2500 nm) at leaf and canopy scales. *Remote Sens Environ* 1995, 53:199–211
19. Hattey JA, Sabbe WE, Baten GD, Blakeney AB. Nitrogen and starch analysis of cotton leaves using near infrared reflectance spectroscopy (NIRS). *Commun Soil Sci Plant Anal* 1994, 25:9-10
20. Barton FE. New methods for the structural and compositional analysis of cell walls for quality determinations. *Anim Feed Sci Technol* 1991, 32:1–11
21. Card DH, Peterson DL, Matson PA, Aber JD. Prediction of leaf chemistry by the use of visible and near infrared reflectance spectroscopy. *Remote Sens Environ* 1988, 26:123-147
22. Kumar L, Schmidt K, Dury S, Skidmore A. Imaging spectrometry and vegetation science. In *Imaging Spectrometry. Volume 4*. Edited by van der Meer F and de Jong SM. Dordrecht: Springer; 2002:111-154
23. Hetta M, Mussadiq Z, Wallsten J, Halling M, Swensson C, Geladi P. Prediction of nutritive values, morphology and agronomic characteristics in forage maize using two applications of NIRS spectrometry. *Soil Plant Sci* 2017, 67:326-333
24. Lu X, Sun J, Mao H, Wu X, Gao H. Quantitative determination of rice starch based on hyperspectral imaging technology. *Int J Food Prop* 2017, 00:1-8
25. Kjær A, Nielsen G, Stærke S, Clausen MR, Edelenbos M, Jørgensen B. Prediction of starch, soluble sugars and amino acids in potatoes (*Solanum tuberosum* L.) using hyperspectral imaging, Dielectric and LF-NMR Methodologies. *Potato Res* 2016, 59:357-374
26. Fourty T, Baret F. On spectral estimates of fresh leaf biochemistry. *Int J Remote Sens* 1998, 19:1283-1297
27. Curran PJ, Dungan JL, Macler BA, Plummer SE, Peterson DL. Reflectance spectroscopy of fresh whole leaves for the estimation of chemical concentration. *Remote Sens Environ* 1992, 39:153-166
28. R Core Team. R: A language and environment for statistical computing. Vienne, Austria: R Foundation for Statistical Computing, 2017
29. Venables WN and Ripley BD. *Modern Applied Statistics with S*. 4<sup>th</sup> Ed. Springer, New York, 2002

30. Baumann P. SimplerSpec. 2016. <https://github.com/philipp-baumann/leaf-starch-spc>. Accessed 21 Aug 2019
31. Stevens A, Ramirez Lopez L. An introduction to the prospectr package. 2014:1-22
32. Kuhn M, Johnson K: Measuring performance in regression models. In Applied Predictive Modeling. Volume 5. New York: Springer; 2013: 59-100
33. Naes T, Martens H. Comparison of prediction methods for multicollinear data. Communications in Stat 1985, 14:545-576
34. Zhao N, Wu ZS, Zhang Q, Shi XY, Ma Q, Qiao YJ. Optimization of Parameter Selection for Partial Least Squares Model Development. Sci Rep 2015, 5:1-10
35. Wold S, Jonsson J, Sjöström M, Sandberg M, Rännar S: DNA and peptide sequences and chemical processes multivariately modelled by principal component analysis and partial least-squares projections to latent structures. ACACAM 1993, 277: 239-253
36. Chong IG, Jun CH. Performance of some variable selection methods when multicollinearity is present. Chemom Intell Lab Syst 2005, 78:103-112
37. Claessens A, Castonguay Y, Bertrand A, Bélanger G, Tremblay G.F: Breeding for improved nonstructural carbohydrates in alfalfa. In Breeding in a World of Scarcity: 13-17 September 2015: Gent. Edited by Roldán-Ruiz I, Baert J, Reheul D. Springer: Cham; 2016: 231-235
38. Holt DA, Hilst AR. Daily variation in carbohydrate content of selected forage crops. Agron J 1969, 61:239-242
39. Shorten PR, Leath SR, Schmidt J, Ghamkhar K. Predicting the quality of ryegrass using hyperspectral imaging. Plant Methods 2019, 15:63
40. Rasmussen S, Parsons AJ, Xue H, Newman JA. High sugar grasses – harnessing the benefits of new cultivars through growth management. Proc New Zeal Grassl Assoc 2009, 71:167-75
41. McKenna P, Cannon N, Conway J, Dooley J. The use of red clover (*Trifolium pratense*) in soil fertility-building: A Review. F Crop Res 2018;221:38-49
42. Dhamala NR, Rasmussen J, Carlsson G, Søgaard K, Eriksen J. Nitrogen fixation in red clover grown in multi-species mixtures with ryegrass, chicory, plantain and caraway. Mult roles Grassl Eur 2016;21:576-578.

## Additional Files

### Additional file 1

Table S1 Name and origin of the used red clover genotypes

### Additional file 2

Figure S1 A) Round sharpened tube, used to cut out leaflets, B) leaf cuts C) and FieldSpec4 pro device opened and closed

### Additional file 3

Figure S2 PLS regression of raw spectra and best model performance of the cross-validation (ncomp = 9; n=337). Different colors and shapes indicate the age of the leaves, m for matures leaves (red, circular), o for the oldest leaf (green, rectangle) and y for the youngest fully emerged leaf (blue, square)

### Additional file 4

Figure S3 PLS regression of the training set separated for the two different harvest times at the end of the day (ED) and at the end of the night (EN). Different colors and shapes indicate the age of the leaves, m for matures leaves (red, circular), o for the oldest leaf (green, rectangle) and y for the youngest fully emerged leaf (blue, square)

### Additional file 5

Figure S4 Reflectance spectra (top panel), pre-processed reflectance (Savitzky-Golay preprocessed; middle panel) and VIP filtering (bottom panel) for the training set

### Additional file 6

Figure S5 Reflectance spectra (top panel), pre-processed reflectance (Savitzky-golay preprocessed; middle panel) and VIP filtering (bottom panel) for the test set

## Figures

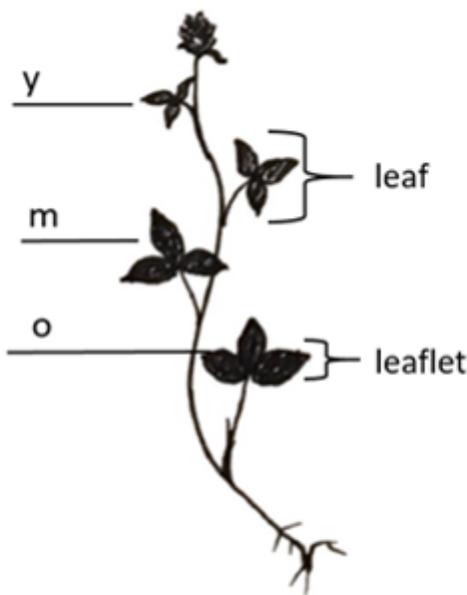


Figure 1

Sampling on red clover plants was performed on the youngest leaf, where three leaflets were fully emerged (y), on the three leaflets of the oldest leaf (o), and on three different intermediate leaves, in total nine leaflets (m)

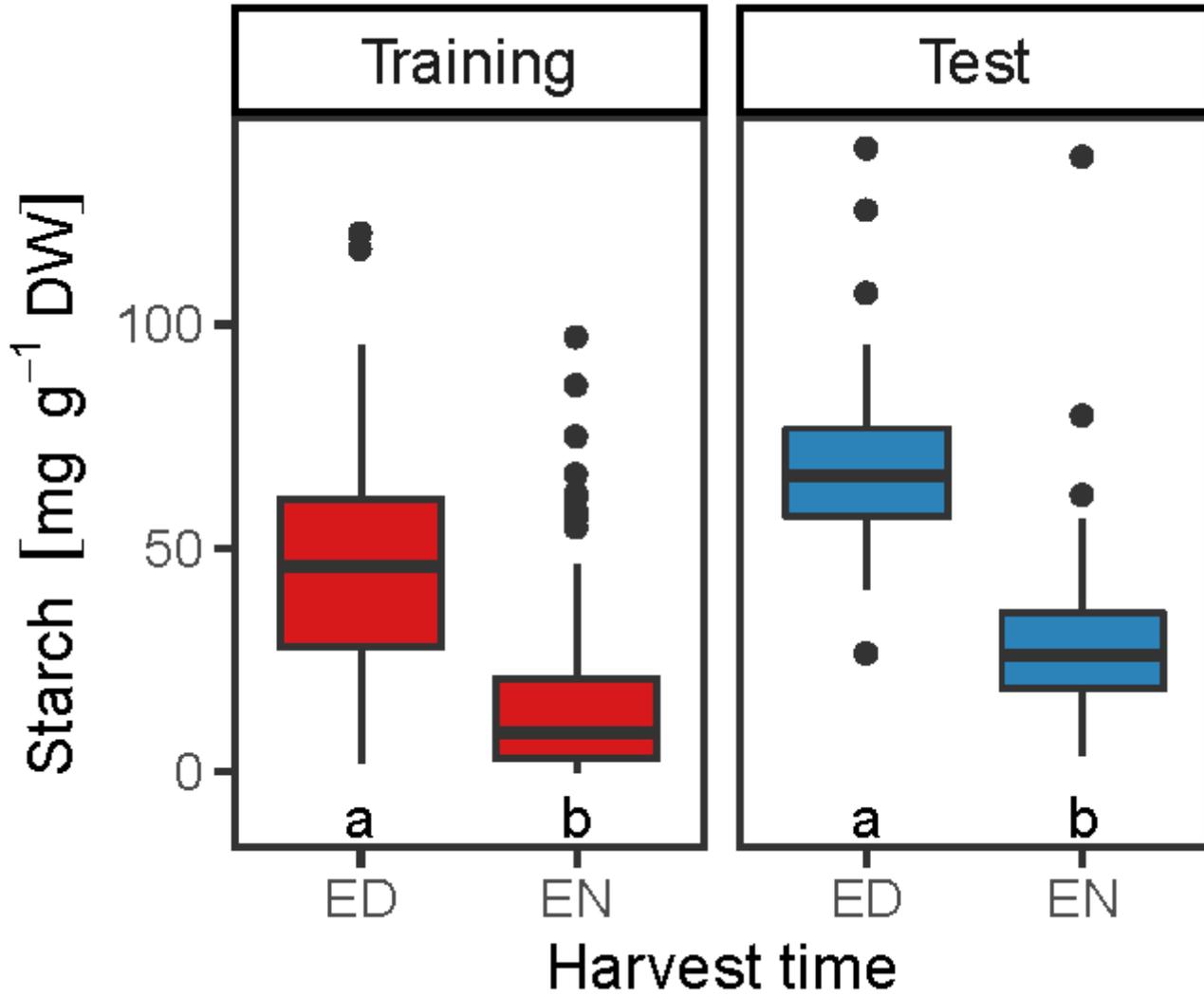


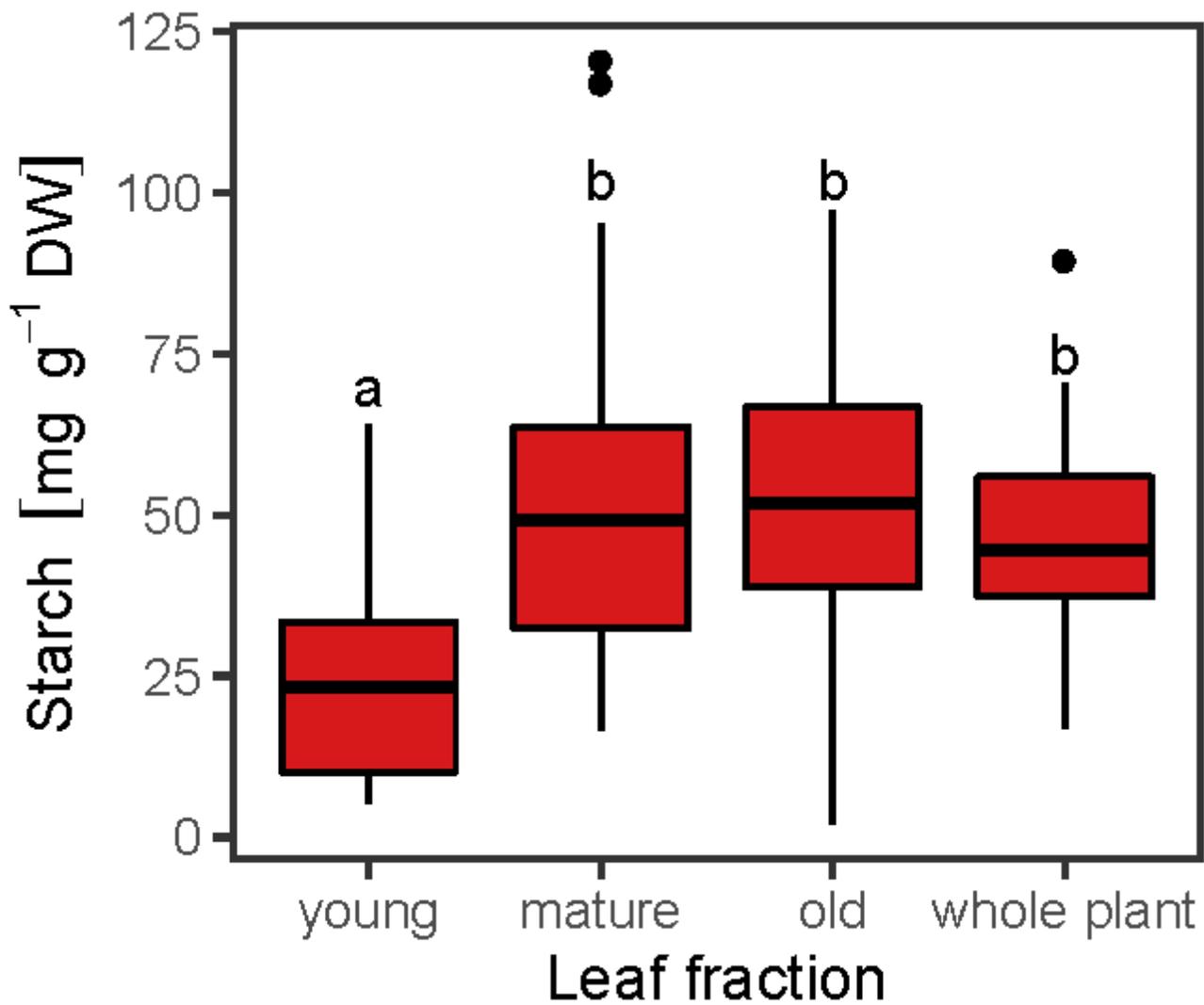
Figure 2

Starch content in mg g<sup>-1</sup> DW in red clover plants of the training (n=449) and the test set (n=60) at the end of the day (ED) and at the end of the night (EN). Medians within the two sets with no letter in common are significantly different (Wilcoxon signed-rank test;  $\alpha = 5\%$ )



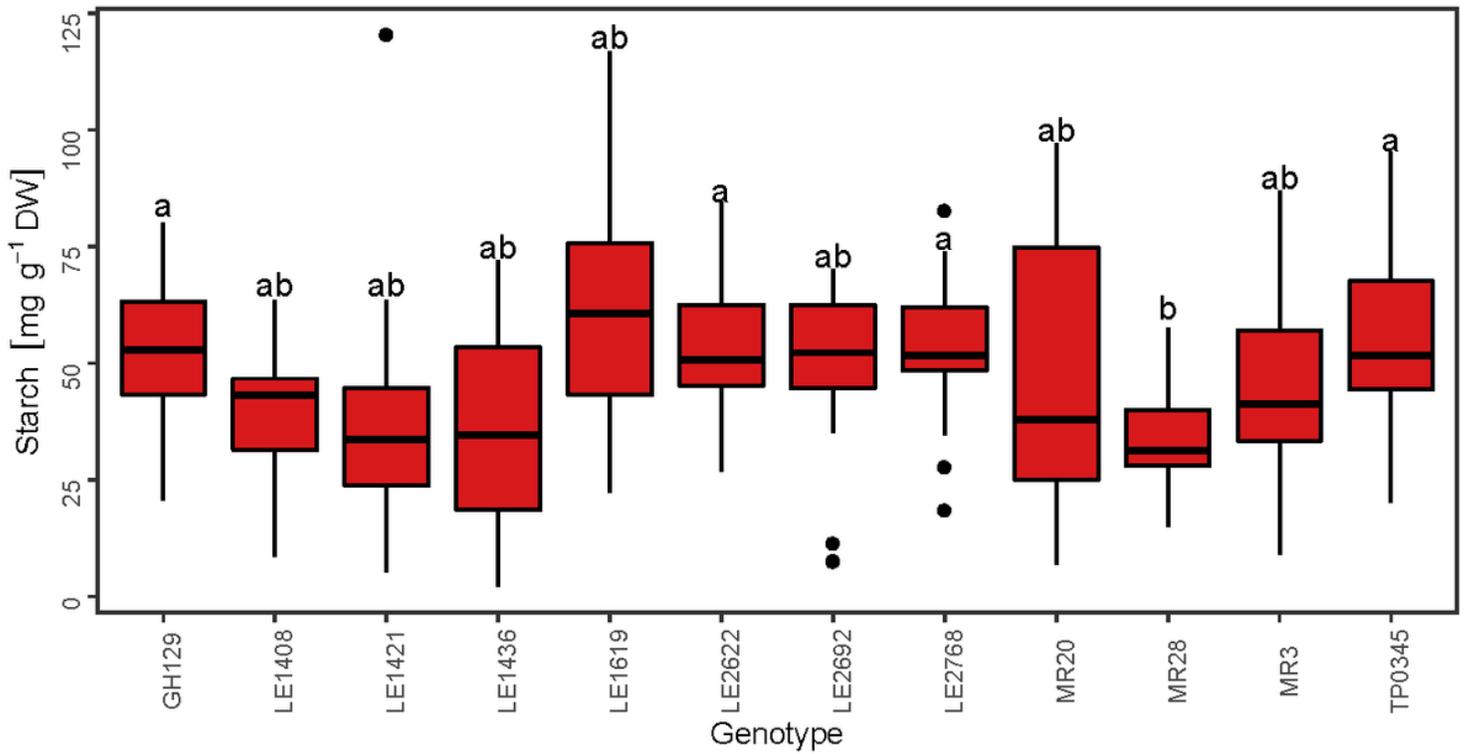
**Figure 3**

Iodine stained red clover plants (clones from the same genotype) visualizing starch distribution (dark coloration). There are clear differences in starch accumulation between harvest times (a) EN (b) ED. Differences in starch accumulation are also visible within one plant (b) and within one leaf (c)



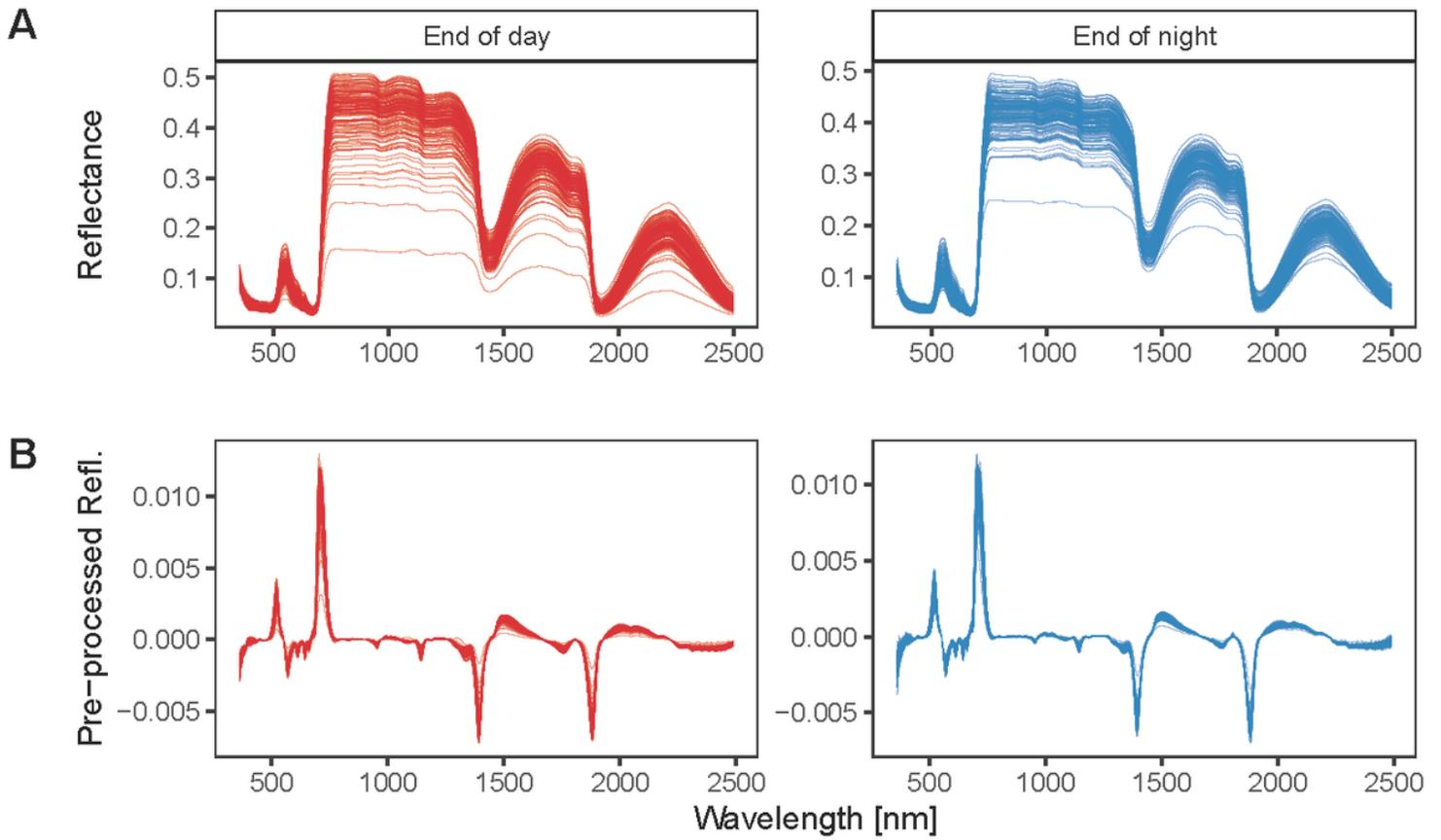
**Figure 4**

Starch concentrations in different leaf fractions (young [n = 35], mature [n = 108], old [n= 36]) and whole plants (n = 48) of red clover genotypes harvested at the end of the day. Horizontal bars represent medians. Medians with no letter in common are significantly different (Wilcoxon- signed-rank test;  $\alpha = 5\%$ )



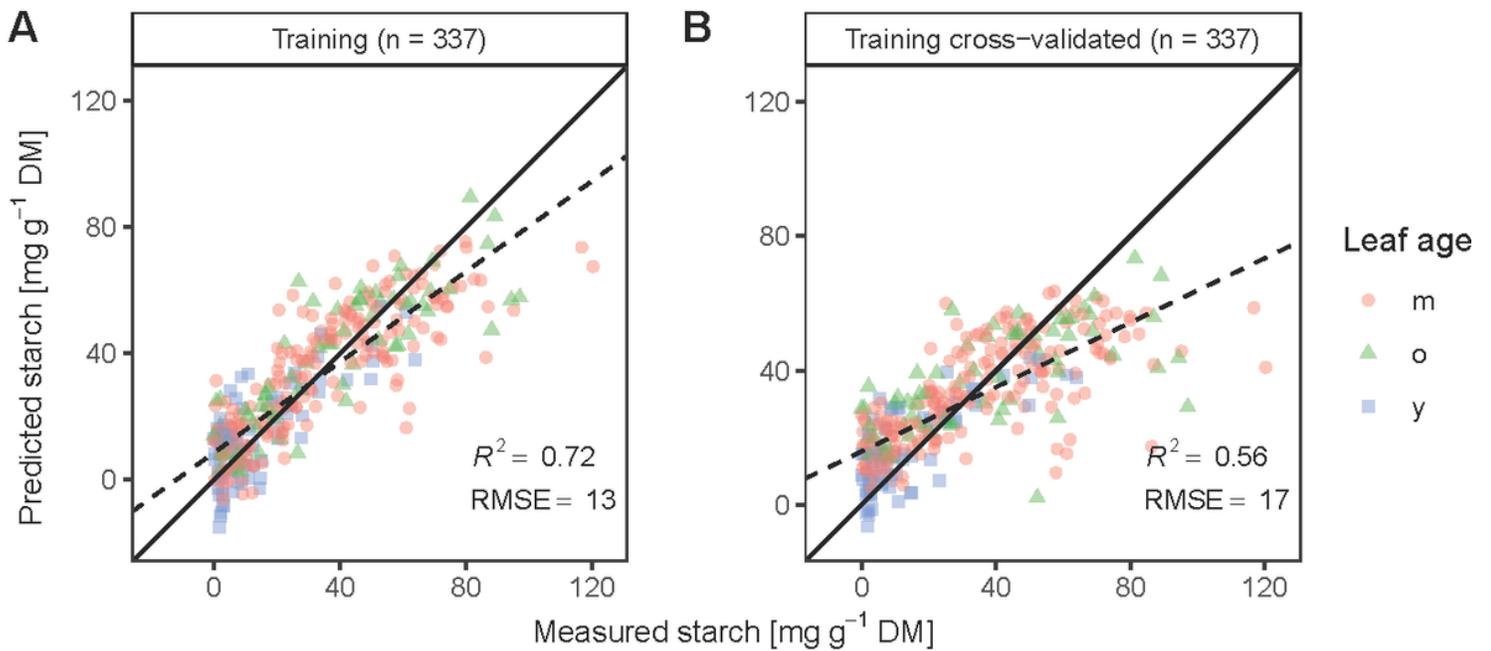
**Figure 5**

Starch content in mg g<sup>-1</sup> DW for the red clover genotypes of the training set harvested at the end of the day (n=226). Medians with no letter in common are significantly different (Wilcoxon signed-rank test;  $\alpha = 5\%$ )



**Figure 6**

Raw (A) and Savitzky-Golay pre-processed reflectance spectra (B) of red clover samples of the training set ( $n = 338$ ) harvested at the end of day (red) or end of night (blue)



**Figure 7**

PLS regression of the training set and best model performance of the cross-validation on the same set (ncomp = 9; n=337, different colors and shape indicate the age of the leaves, m for matures leaves, o for the oldest leaf, and y for the youngest fully emerged leaf)

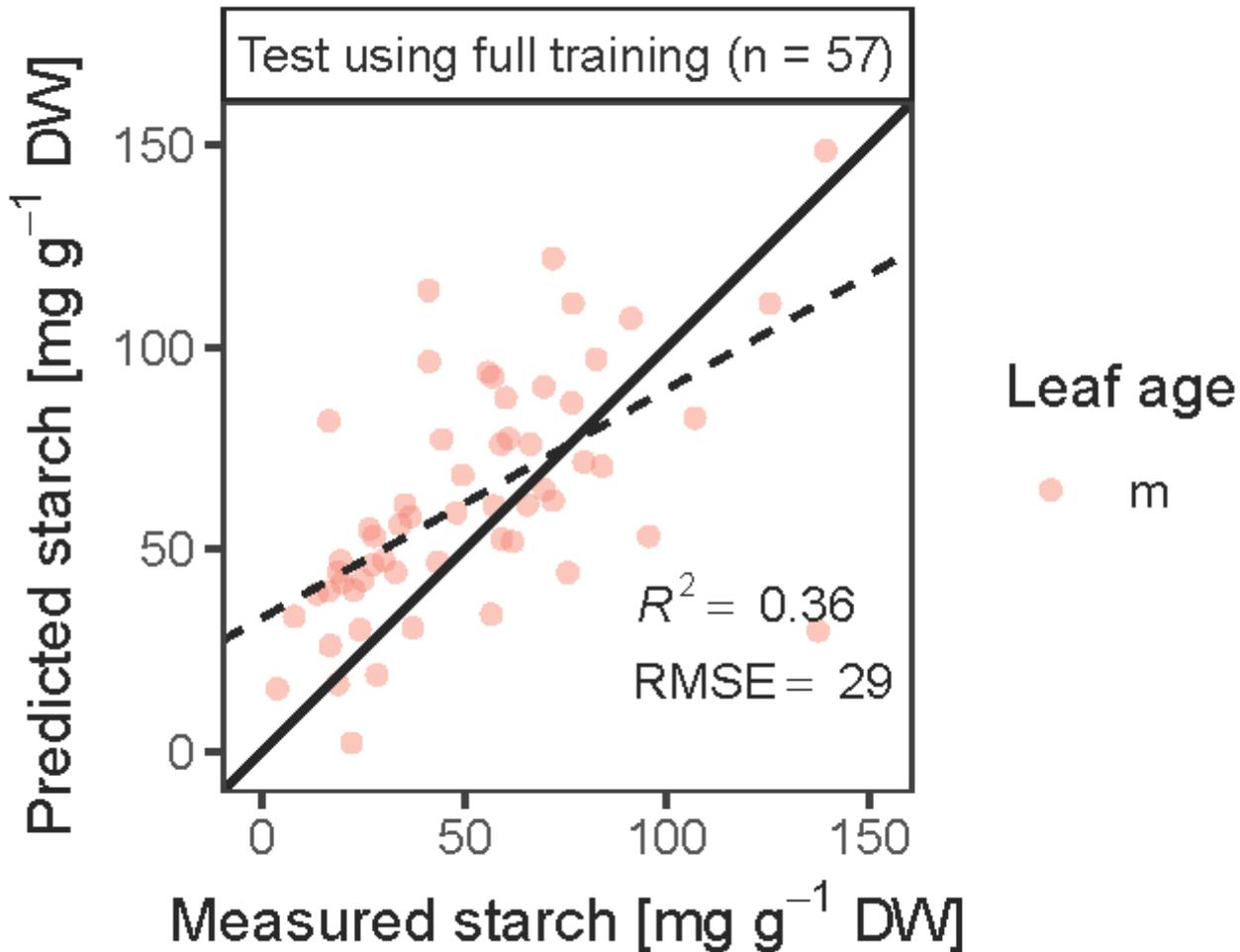


Figure 8

Best performing model from the training set was used to predict leaf starch content in an independent test set (n = 57)

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile6.pdf](#)
- [Additionalfile4.pdf](#)
- [Additionalfile3.pdf](#)
- [Additionalfile5.pdf](#)
- [AdditionalFile1.docx](#)
- [Additionalfile2.png](#)