CSIRO PUBLISHING

Functional Plant Biology
http://dx.doi.org/10.1071/FP16121

Hyperspectral imaging reveals the effect of sugar beet quantitative trait loci on Cercospora leaf spot resistance

Marlene Leucker[A,D], *Mirwaes Wahabzada*[A], *Kristian Kersting*[B], *Madlaina Peter*[C], *Werner Beyer*[C], *Ulrike Steiner*[A], *Anne-Katrin Mahlein*[A] *and Erich-Christian Oerke*[A]

[A]Institute for Crop Science and Resource Conservation (INRES) – Phytomedicine, University of Bonn, Meckenheimer Allee 166a, 53115 Bonn, Germany.
[B]Department of Computer Science, TU Dortmund University, Otto-Hahn-Str. 14, 44227 Dortmund, Germany.
[C]KWS SAAT SE, Grimsehlstrasse 31, 37555 Einbeck, Germany.
[D]Corresponding author. Email: mleucker@uni-bonn.de

Abstract. The quantitative resistance of sugar beet (*Beta vulgaris* L.) against Cercospora leaf spot (CLS) caused by *Cercospora beticola* (Sacc.) was characterised by hyperspectral imaging. Two closely related inbred lines, differing in two quantitative trait loci (QTL), which made a difference in disease severity of 1.1–1.7 on the standard scoring scale (1–9), were investigated under controlled conditions. The temporal and spatial development of CLS lesions on the two genotypes were monitored using a hyperspectral microscope. The lesion development on the QTL-carrying, resistant genotype was characterised by a fast and abrupt change in spectral reflectance, whereas it was slower and ultimately more severe on the genotype lacking the QTL. An efficient approach for clustering of hyperspectral signatures was adapted in order to reveal resistance characteristics automatically. The presented method allowed a fast and reliable differentiation of CLS dynamics and lesion composition providing a promising tool to improve resistance breeding by objective and precise plant phenotyping.

Additional keywords: hyperspectral reflectance, proximal sensing, QTL, quantitative disease resistance, spherical k-means clustering, spectral phenotyping, spectral vegetation indices.

Received 31 March 2016, accepted 2 August 2016, published online dd mmm yyyy

Introduction

The understanding of the genetics of crops, especially of more complex traits such as disease resistance, has increased with advanced molecular techniques. The optimisation of crop cultivation and breeding for high yield as well as for resistance to biotic or abiotic stress benefits from this knowledge. Mapping of quantitative trait loci (QTL) is an effective and widely used method to study genetically complex traits such as quantitative disease resistance which is not as well understood as qualitative (monogenic) resistance (Setiawan *et al*. 2000). Molecular markers, closely linked to QTL regulating quantitative disease resistance, can be used for marker-assisted breeding to introduce these valuable traits into new varieties.

The breeding of Cercospora leaf spot resistant varieties with high yield potential also under non-disease conditions is a major challenge in sugar beet breeding. Disease resistance was introgressed into the domesticated *Beta vulgaris* L. from its wild relative *B. vulgaris* ssp. *maritima* (L.) Arcang. (Munerati *et al*. 1913). According to Smith and Gaskill (1970), 4–5 major genes confer CLS resistance in sugar beet. However, recent studies indicated that more genes are involved (Weltmeier *et al*. 2011). An exception of race-specific, monogenic resistance has been reported where the dominant *Cb* resistance (R) gene confers resistance to the C2 isolate of *Cercospora beticola* (Sacc.), but not to isolate C1 (Lewellen and Whitney 1976). The resistance reaction is characterised by small, fleck-like lesions and a rapid induction of signalling genes 1 day after inoculation (dai) and pathogenesis-related (PR) genes 3 dai respectively (Weltmeier *et al*. 2011). The C2 isolate is also more aggressive on susceptible cultivars lacking the *Cb* gene. Quantitative resistance, which is of more importance, has been described as rate-reducing and is characterised by an increase in the incubation time and a reduction in the number and size of lesions as well as conidia (Rossi *et al*. 1999, 2000; Leucker *et al*. 2016). It was also observed that resistant sugar beets typically have lesions with smaller centres compared with highly susceptible genotypes (Leucker *et al*. 2016).

Several approaches of QTL mapping have been conducted to facilitate the challenging CLS resistance breeding. QTL have been mapped on four (Setiawan *et al*. 2000) and six (Koch *et al*. 2000) of the nine sugar beet chromosomes. To ensure a successful genetic evaluation and to unravel the number as well as the positions of genetic resistance, it is essential to have a precise and reliable screening method (Setiawan *et al*. 2000).

The KWS 1–9 scoring scale is generally adopted for the visual assessment of *C. beticola* resistant genotypes (Shane and Teng 1992). As phenotypic differences between QTL lines can be very small and the effect may be highly sensitive to environmental factors, visual rating scales may fail to reveal these subtle differences (Geiger and Heun 1989). Furthermore, quantitative effects on the disease epidemic may become visible only at the end of the growing season when several disease cycles have been repeated. Therefore, a more precise and objective phenotyping practice is needed to improve the breeding for quantitative disease resistance.

The application of sensor-based phenotyping techniques may contribute to improve the selection process, because sensors provide considerable advantages (Mahlein 2016). Sensors are objective, sensitive and non-invasive. Furthermore, they are suitable for standardised high-throughput screenings (Montes *et al*. 2007). Optical spectral reflectance measurements are widely used for the non-invasive assessment of the physiological status of vegetation (e.g. pigment content, leaf area) and to detect stress such as drought stress, nutrition deficiencies or diseases of plants (Bock *et al*. 2010). The chlorophyll content and changes in pigment compositions are typically used as stress indicators and can be assessed by spectral vegetation indices (SVIs) (Blackburn 1998). SVIs are calculated from specific wavelengths of the spectral reflectance and were successfully used for early detection and differentiation of fungal sugar beet diseases (Mahlein *et al*. 2010, 2013). Providing information not only on the pigment content, but also on the structure, the physiology and the water status of plant tissue, hyperspectral reflectance can characterise phenotypic differences in host-pathogen compatibility (Mutka and Bart 2015). Kuska *et al*. (2015) distinguished barley genotypes differing in susceptibility to powdery mildew based on hyperspectral dynamics of the plant-pathogen interaction. This phenotyping approach included an automated, data-driven analysis which is of great interest for the applicability in screening systems (Wahabzada *et al*. 2016).

Within this context efficient processing and handling of phenotypic data is crucial but also challenging (Wahabzada *et al*. 2015): (i) datasets can easily become massive, in particular when several samples are measured over time; (ii) although the labels of each image may be known, it is difficult – if not impossible – to assign the labels to individual hyperspectral signatures at pixel level. To deal with the massive data, its complexity can be reduced using a hierarchical decomposition of spectral signatures within the image as described by Kersting *et al*. (2010). Based on the reduced data amount, standard clustering techniques such as 'k-means' clustering can be applied to learn a suitable segmentation of hyperspectral images. The k-means algorithm is an unsupervised approach for clustering based on the geometrical distribution of the data points into a given number of clusters (Bishop 2006). For CLS, it is challenging to identify the composition of individual disease symptoms by the naked eye, and thus, a satisfactory number of classes for the characterisation of different resistances. Consequently, out-of-the-box classification techniques do not work and an unsupervised data analysis method with an automated classification of spectral information at pixel level is required.

The motivation of the present study was to characterise small differences in quantitative CLS resistance by hyperspectral imaging. Lesion development on two closely related inbred lines differing in two QTL were investigated using a combination of manual analyses and unsupervised clustering of hyperspectral images.

Materials and methods

Plant material

Two inbreed lines of *Beat vulgaris* L. ssp. *vulgaris* with different numbers of QTL conferring resistance to *C. beticola* were chosen for the experiments (KWS SAAT SE, Einbeck, Germany). Both inbred lines Bv1–2– and Bv1+2+ are monogerm inbred lines which have been developed out of crosses of a highly homozygous KWS elite line with the polygenic CLS resistant donor line K31 (Weltmeier *et al*. 2011). An F3 mapping population had been developed and used to identify QTL associated with the resistant phenotype (M. Peter and W. Beyer, unpubl. data). Lines Bv1–2– and Bv1+2+ have been developed by repeated backcrossing of single progenies of the mapping population with the respective non-resistant elite parent. The backcross procedure was accompanied by marker-assisted selection with QTL flanking markers for absence (Bv1–2–) and presence (Bv1+2+) of the QTL-flanking markers, leading to the lines Bv1–2– and Bv1+2+ with a proportion of >80% of the elite parent, respectively. While Bv1+2+ carries the alleles of the resistant donor parent K31 in the two QTL regions, Bv1–2– shows the susceptible marker alleles like the elite parent used as backcross parent. Both QTL regions are located on different chromosomes and are inherited independently. The QTL-flanking markers cover windows of 12 and 35 cM, respectively, on the individual chromosomes. Susceptibility to CLS was assessed in field trails in Italy in the growth season 2011. Both QTL act in additive manner (M. Peter and W. Beyer, unpubl. data).

For greenhouse experiments, seeds were pre-grown in small pots until growth stage BBCH 12 (GS, Meier *et al*. 2001). Seedlings were transferred into commercial substrate (Einheitserde ED73, Klasmann-Deilmann, Geeste, Germany) in plastic pots (Ø 17 cm). Plants were cultivated under controlled conditions at 22/20°C day/night temperature, 60 ± 10% RH and a photoperiod of 16 h per day. Six weeks after transplanting, the sugar beets were fertilised weekly with a 0.2% solution of Poly Crescal (Aglukon GmbH). Plants with six fully developed leaves (GS 16) were used for the experiments.

Pathogen inoculation and disease assessment

Sugar beet lines were grown in a trial field in Italy in the growth season 2011 with 90 plant per plot and four repetitions. In order to increase the infection pressure, dried *Cercospora beticola* (Sacc.) infected plant material was wetted and applied into the field for inoculation. Disease severity was assessed visually and repeated weekly starting with disease outbreak in mid-July until the end of August using a 1–9 disease rating scale (Shane and Teng 1992).

Inoculum of *C. beticola* for greenhouse experiments was harvested from wetted infected sugar beet leaves incubated at

100% RH for 48 h. Spores were washed off with a 0.01% solution of Tween 20 (Merck Millipore) and a spore suspension with 4×10^4 conidia mL^{-1} was sprayed onto sugar beet leaves (Leucker *et al*. 2016). After inoculation, plants were incubated at 100% RH and 25/20°C for 48 h and afterwards placed back to 60% RH. CLS development was visually assessed by calculating the number of lesions per leaf area.

Hyperspectral imaging

Hyperspectral reflection was measured by a hyperspectral line scanning camera (spectral camera PFD V10E, Specim) as previously described by Leucker *et al*. (2016). Samples were 4-fold magnified. Since reflectance spectra were noisy at the extremes, only reflectance values from 460 to 850 nm were considered for data analysis. Hyperspectral measurements of inoculated leaves started when first lesions appeared (9 dpi) and were repeated every other day until 19 dpi.

Data processing and calculation of spectral vegetation indices

Hyperspectral images were normalised as described by Leucker *et al*. (2016). The Savitzky-Golay filter (Savitzky and Golay 1964) was applied to smooth the spectral signals with seven supporting points to the left and right, each, and a fifth degree polynomial. Spectral signatures were manually extracted from pre-processed images as region of interest (ROI) and the average reflectance of 10 CLS lesions was calculated. Spectral vegetation indices were determined for the same regions: (1) the pigment specific simple ratio (PSSRa; Blackburn 1998) to estimate the chlorophyll a content was calculated as R_{800}/R_{680} where R denotes the waveband used, (2) the photochemical reflectance index (PRI, Gamon *et al*. 1992) as measure for the radiation use efficiency (Garbulsky *et al*. 2011) was calculated as $(R_{531}-R_{570})/(R_{531}+R_{570})$.

Fast clustering of hyperspectral data

Unsupervised techniques are applied to identify a suitable grouping of hyperspectral signatures, where no information about the groups or classes, such as groups of training data, is available. This is in contrast to supervised classification settings, e.g. the spectral angle mapper algorithm, which is a common method for the analysis of remote sensing data. In this process, the similarity of an unknown spectral signature to known reference spectra is calculated based on the angular difference (Luc *et al*. 2005).

Considering that approach, spherical k-means (Dhillon and Modha 2001) was adapted for clustering spectral signatures of CLS dynamics. Unlike the standard *k*-means algorithm, where the Euclidean distance is used to measure the distance between the vectors, spherical k-means clusters the data on the sphere. Therefore, one can employ the cosine similarity, which is based on the angles between vectors or spectral signatures. It is defined by Eqn 1 where $<x,y>$ denotes the scalar product between the vectors x and y and $||x||$ denotes the Euclidean norm:

$$d(x,y) = 1 - \cos(x,y) = 1 - \frac{<x,y>}{\| x \| \| y \|}. \quad (1)$$

Equivalently, the vectors might be normalised to have unit Euclidean norm, hence, they can be thought of as points on a high dimensional unit sphere. This captures the direction of the vectors and mitigates the effects of varying reflectance values that can result from, for example, specular reflectance or the uneven surface of leaves. Furthermore, signatures with similar shapes may have lower values, allowing a more robust clustering. In the case of normalised signatures, the Euclidean distance can be applied to compute the similarity of the signatures. However, as the mean vectors $\mu_1,\ldots,\mu_k$, k being the number of clusters, may not have unit length, they were normalised in each iteration, such that:

$$c_i = \frac{\mu_i}{\| \mu_i \|}. \quad (2)$$

The concept vector c_i is the vector closest to all vectors in the cluster based on cosine similarity. See Dhillon and Modha (2001) for more details.

In order to derive an efficient approach for clustering of hyperspectral images, data complexity was reduced by first computing a hierarchical decomposition in form of a tree (Kersting *et al*. 2010). Starting with an empty tree, a 1D projection of the data was computed repeatedly and searched for a splitting criterion minimising the weighted variance as much as possible in each resulting subtree. At the end of the hierarchical decomposition, the representatives for each leaf of the tree were determined as computed mean signature. Finally, the derived representatives – a kind of prototypes – were used for the clustering, thus, reducing the complexity of the dataset.

As the hyperspectral data lie on a smooth, structured low dimensional manifold, its intrinsic structure was captured without learning this structure explicitly (Dasgupta and Freund 2009). Overall, the approach for computing a clustering of hyperspectral data includes four steps: (1) project the data on the unit sphere by normalising each signature with its Euclidean norm; (2) compute a hierarchical decomposition in form of a tree and determine the representatives y_i for each leaf i on the tree from its normalised mean signature; (3) compute k clusters from the representatives $y_1,\ldots,y_l$; (4) obtain the cluster membership of all hyperspectral signatures using the cluster means and membership of $y_1,\ldots,y_l$ computed in the previous step.

The cluster mean for every cluster k, the prototypes, could also be used as reference spectra to classify new data. In order to do this, the distances of each unknown test vector to the reference spectra are computed and the class/cluster with the lowest distance or highest similarity is returned. Furthermore, the approach is readily parallelisable: simply divide the data into several subsets, compute the hierarchical decomposition on each subset separately, collect the results and run the clustering algorithm on the merged results. As noted before, the efficiency is of great importance due to the absence of prior information. Additionally, data are often acquired under different conditions. In turn, it requires a large number of repeated experiments with different parameters to get optimal results. In this study, 120 hyperspectral images with a total size of

more than 25 GB where analysed and, due to the hierarchical decomposition and parallelisation, it took less than 10 min to compute the clustering of the data on a Dual-CPU Xeon Workstation with total number of 20 cores.

Statistical analysis

Statistical analyses were performed using SPSS 22.0 (SPSS Inc.). Data were tested for normal distribution and equity of variances. The means between the two genotypes were compared using the *t*-test in case of a normal distribution, otherwise by the Mann–Whitney U-test with a significance level of $P = 0.05$ confidence.

Results

Disease assessment

Severity of CLS on sugar beet genotypes Bv1–2– (absence of QTL) and Bv1+2+ (presence of QTL) was assessed visually in a field trial in Italy in 2011 (Fig. 1*a*). Significant differences were detected two weeks after the disease outbreak (mid-July), when average disease severity was 1.3 scores higher for Bv1–2– than for Bv1+2+. The repeatability of the resistance trait was 0.79 at the time of best differentiation (mid-July) and 0.64 in average over seven scoring dates. Based on this assessment, the two genotypes were selected to analyse fine differences in disease resistance, and appearance of CLS lesions over time was investigated under optimal disease conditions in the greenhouse (Fig. 1*b*). Remarkably, the incubation time and the number of lesions per leaf area were very similar for both genotypes.

Spectral characterisation of lesion development

The development of CLS lesions on the sugar beet lines was analysed by hyperspectral imaging in order to identify the differences in resistance based on the two QTL. On both genotypes first lesions appeared 9 dai. Initial lesions on the less resistant genotype, lacking the resistance QTL, appeared as depressions of the leaf surface, which still was green. During further development, the tissue became necrotic brown with a white to grayish centre. Over time the spectral reflectance of CLS lesions on Bv1–2– increased continuously in the visible (VIS) range, especially from 600 to 650 nm (Fig. 2*a*). The near-infrared (NIR) reflectance decreased compared with the reflectance of healthy tissue directly at the day of lesion appearance (9 dai). The lesions on the genotype carrying the QTL were characterised by a faster and more distinct increase of reflectance in the VIS range right from 9 dai (Fig. 2*a*). The decrease in the NIR range was also stronger than on Bv1–2– and only marginal changes were detected over the further experimental time. On the RGB depiction of the hyperspectral images, the initial lesions already appeared as brown, necrotic depressions and showed little changes in size and colour (Fig. 2*b*).

In order to characterise the physiological effect of the QTL on CLS resistance, spectral vegetation indices related to pigment content and photosynthetic light use efficiency were calculated for the diseased tissue. The values of PSSRa and the PRI of healthy control leaves were similar for both genotypes (Fig. 3). After lesion appearance, the PSSRa decreased considerably from ~11 to 5 on Bv1–2–. The initial decline was significantly lower on Bv1+2+ and the PSSRa value remained at this level for the rest of the experimental period. The PSSRa of Bv1–2– further decreased over time and was significantly lower than for Bv1+2+ 19 dai. Before inoculation, the PRI of the two genotypes was approximately –0.1 and decreased during lesion development with significantly lower values for Bv1+2+. We note that the difference between the genotypes became smaller.

Clustering of hyperspectral images of CLS dynamics

The manual extraction of the spectral signatures and the vegetation indices revealed clear differences between the CLS lesions on the two breeding lines. For an implementation of hyperspectral imaging into phenotyping processes, however, a more efficient analysis is required. Hence, an approach for fast clustering was developed to identify differences in CLS phenotypes of the two sugar beet genotypes in an automated way. The spectral signatures of all image pixels were grouped into clusters without any prior information or labelling. For this clustering approach, the number of clusters has to be preset. To

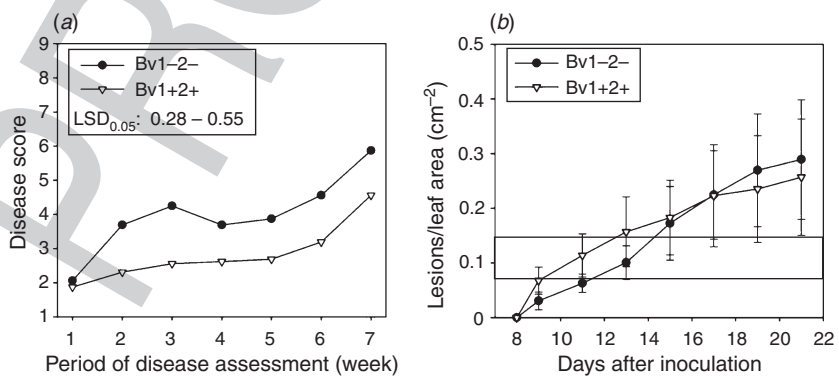

Fig. 1. Assessment of Cercospora leaf spot on sugar beet genotypes Bv1–2– (absence of QTL) and Bv1+2+ (presence of QTL) in (*a*) a field trial and (*b*) under controlled conditions. Using a 1–9 scale, significant differences were detected from the second week of disease outbreak (least significant difference (l.s.d.$_{0.05}$) >0.5). The number of lesions per leaf area was not significantly different (*b*).

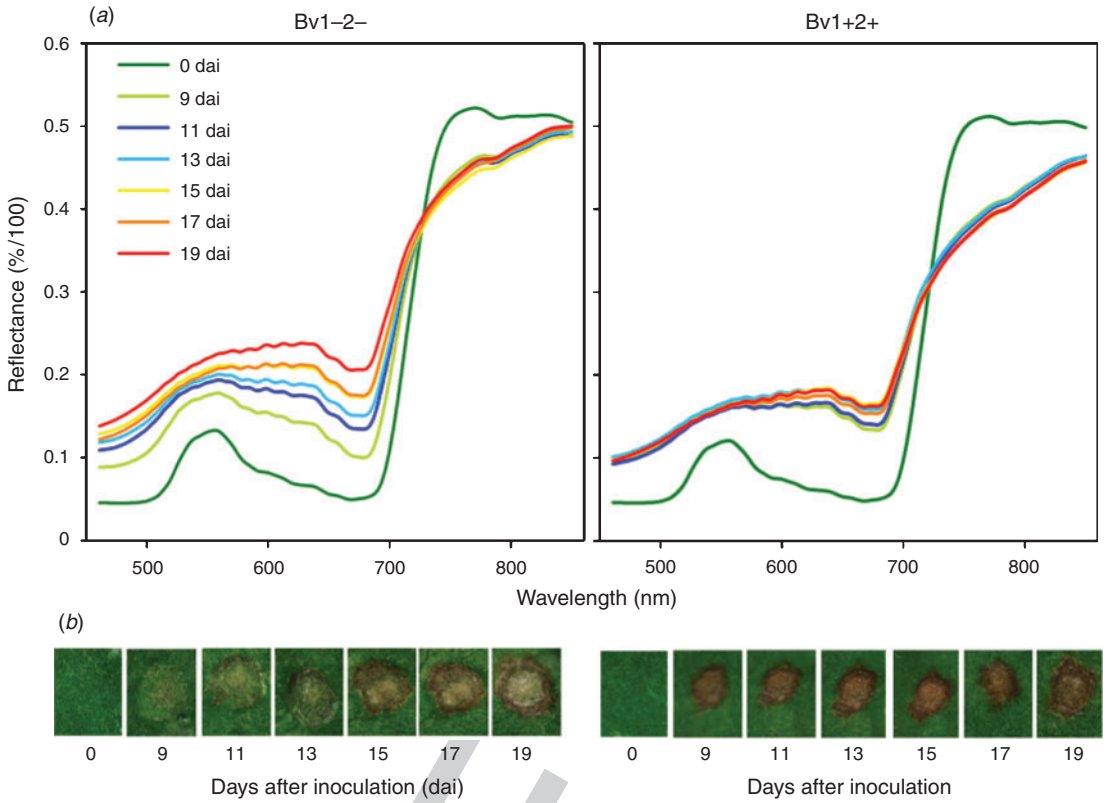

Fig. 2. (*a*) Spectral signatures and (*b*) RGB images of sugar beet genotypes Bv1–2– and Bv1+2+ non-inoculated (0 days after inoculation) and 9 to 19 days after inoculation with *Cercospora beticola*. Reflectance spectra represent the average of $n = 10$ lesions with at least 8000 pixels each.

estimate the best number of clusters, different numbers were tested on mature CLS lesions 19 dai (Fig. 4). A good separation of healthy from diseased tissue was possible with two or three clusters; however, leaf veins and specular reflections could not be identified. By increasing the number of classes the clustering within lesions improved considerably and differences between the genotypes became most evident with seven clusters. The dimension of the lesion centre on Bv1–2– was more pronounced and the tissue surrounding the lesions seemed to be more affected than on Bv1+2+. Classification with 10 or more clusters gave no further valuable segmentation and even resulted in less clear distributions within the lesions. Seven was considered to be a suitable number of clusters and was, therefore, used for analysing the full dataset.

The spectral signatures of the seven cluster means are given in Fig. 5*a*. Each cluster was assigned to one colour and the clustering results throughout the assessment period are visualised by false-colour images in Fig. 5*b*. The clusters 1 to 3 represented the green, visibly non-infected tissue (compare Fig. 2). The reflectance of the healthy tissue varied due to the inclination of the leaf surface or leaf veins leading to three different clusters. Cluster 4 was associated with the early stage of CLS, where the leaf surface is sunken and still green (Fig. 2*b*). Clusters 5 and 6 represented the outer and inner area of lesions, respectively. The seventh cluster was mainly represented in the lesion centre on the less resistant genotype Bv1–2– (Fig. 5*b*). The quantification of the pixels in the respective clusters associated with the lesions (4–7) was chosen to assess the temporal and – moreover – the spatial differences on the two genotypes. The amount of pixels per lesion on Bv1–2– increased considerably from 7000 to 42 000. Cluster 4 was also represented more frequently in lesions on Bv1–2– than on Bv1+2+, even though the proportion of cluster 4 decreased with lesion development (Fig. 5*c*). Moreover, the absolute as well as the relative amount of clusters 6 and 7 increased over time. In contrast, lesions on Bv1+2+ had ~13 000 to 15 000 pixels on average over the whole time. Not only the size, but also the composition of the clusters remained almost constant over time.

Discussion

This study demonstrated the potential of hyperspectral imaging to reveal the effect of sugar beet QTL on CLS resistance. In field trials, the genotype Bv1+2+ carrying two QTL of a highly resistant donor line performed ~15% better in CLS resistance than the inbred line Bv1–2–, which carried the susceptible marker alleles. Under controlled and disease promoting conditions, differences in disease severity were not detectable by visual assessment after one infection cycle. However, the dynamics of spectral signatures of CLS were considerably different on the two sugar beet lines. On both genotypes, the reflectance of the red shoulder declined at the time of lesion appearance and remained on this level. This decrease in reflectance was much stronger on the resistant genotype

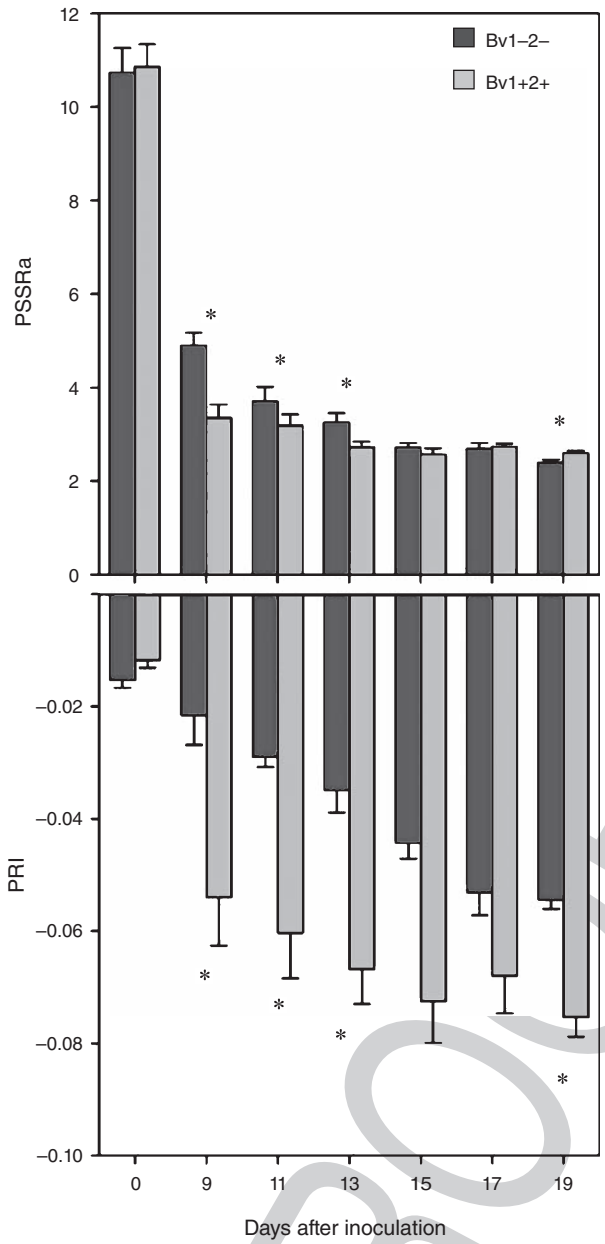

Fig. 3. Effect of Cercospora leaf spot development on pigment specific simple ratio (chlorophyll *a*, PSSRa) and photochemical reflectance index (PRI) of sugar beet genotypes Bv1–2– and Bv1+2+ 0 to 19 days after inoculation (dai). Significant differences between genotypes are indicated: *, $P < 0.05$ according to *t*-tests for PSSRa on 0, 9, 11, 17, 19 dai and PRI on 9, 15 to 19 dai or Mann–Whitney U-tests for PSSRa on 13, 15, 19 dai and PRI on 0, 11, 13 dai ($n = 10$).

Bv1+2+ carrying the QTL. Absorption in the far-red region is associated with brown pigments and phenolic compounds (Peñuelas and Filella 1998). Changes in the NIR range are also related to structural discontinuities, therefore indicating a higher degree of tissue damage on the resistant genotype. In the early stage of lesion development, the leaf tissue collapsed in a limited area, presumably because of a decreased turgor pressure and the loss of membrane integrity possibly induced by cercosporin and other toxins (Daub and Ehrenshaft 2000; Goudet *et al*. 2000). The appearance of the early lesions differed on the two genotypes. Whilst the collapsed tissue was still green on Bv1–2– until 13 dai, the lesions on Bv1+2+ had become brownish associated with significantly lower PSSRa and PRI values. This and the initial higher increase of reflectance in the VIS range on the resistant genotype indicate a stronger reaction of the affected cells compared with the continuous increase on the genotype lacking the resistance QTL.

The increased reflectance in the VIS light is a consequence of reduced absorption by photosynthetic pigments related to pigment degradation and decreased photosynthetic activity (Merzlyak *et al*. 1999). A slower reduction in photosynthetic activity of the susceptible genotype may contribute to a better and persistent assimilate supply for the pathogen compared with the resistance interaction. In barley, even an increased photosynthesis rate of leaf tissue infected with the biotroph rust fungus *Puccinia hordei* was reported (Scholes and Farrar 1986). The abrupt increase of reflectance in the VIS range on the resistant genotype may be based on an enhanced sensitivity to cercosporin or a faster and stronger induction of defence mechanisms like lignification or phytoalexin production (Martin 1977; Daub and Ehrenshaft 2000). Feindt *et al*. (1981) showed that the reduction of hyphal growth and sporulation on resistant sugar beets coincides with the accumulation of extracellular material in lesions, but also in the margin of lesions on susceptible plants.

Many parameters play a role in the rate-reducing resistance against *C. beticola* (Rossi *et al*. 1999, 2000). As incubation time and the number of lesions per leaf area were not affected by the QTL, the size of lesions and the tissue colonisation of *C. beticola* had to be investigated to identify the parameter(s) responsible for the resistance differences assessed in the field. With the efficient approach for clustering all hyperspectral images at once it was possible to quantify the CLS lesions and assess their individual spatial composition. The results of the clustering approach demonstrated the growth of the lesions and the change in composition on the susceptible genotype as previously described by Leucker *et al*. (2016) and Mahlein *et al*. (2012). Mature lesions were characterised by a large lesion centre, also well visualised by the false-colour images, known to be an indicator of susceptibility (Leucker *et al*. 2016). The resistant genotype seemed to prevent an extensive colonisation of leaf tissue by *C. beticola* reflected by the stable size and composition of lesions.

Plants have evolved mechanisms to defend themselves against attacking pathogens. Weltmeier *et al*. (2011) demonstrated a higher induction of defence-related genes in the stage when necrotic lesions appeared for the *C. beticola* resistant line K31 compared with a susceptible genotype. In order to inhibit pathogen growth, plants produce and accumulate phytoalexins which are synthesised *de novo* after pathogen infection. Geigert *et al*. (1973) isolated the isoflavone betavulgarin from sugar beet leaves infected with *C. beticola*. Betavulgarin inhibited mycelial growth drastically and resistant cultivars generally produced higher concentrations of betavulgarin in CLS lesions than more susceptible cultivars (Johnson and Taylor 1976; Martin 1977). An accelerated phytoalexin accumulation may be involved in the limited lesion development on the resistant

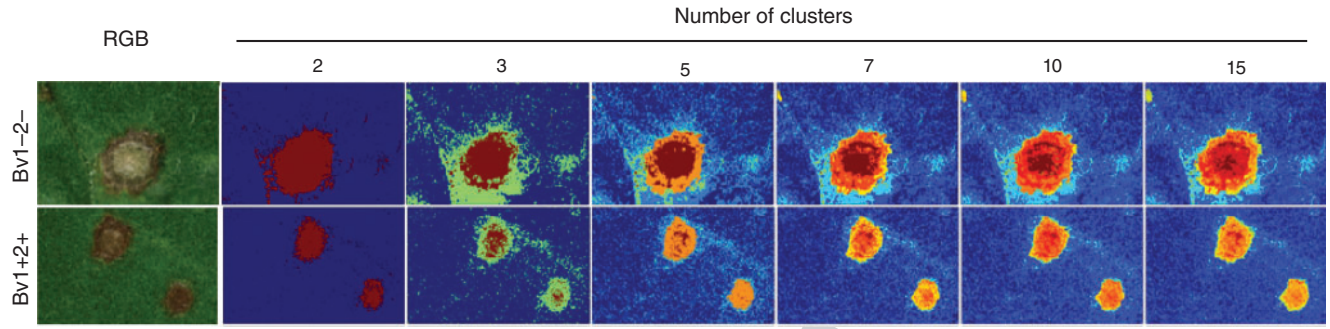

Fig. 4. Results of clustering hyperspectral images of representative Cercospora leaf spot lesions on sugar beet genotypes Bv1–2– and Bv1+2+ depending on the number of clusters.

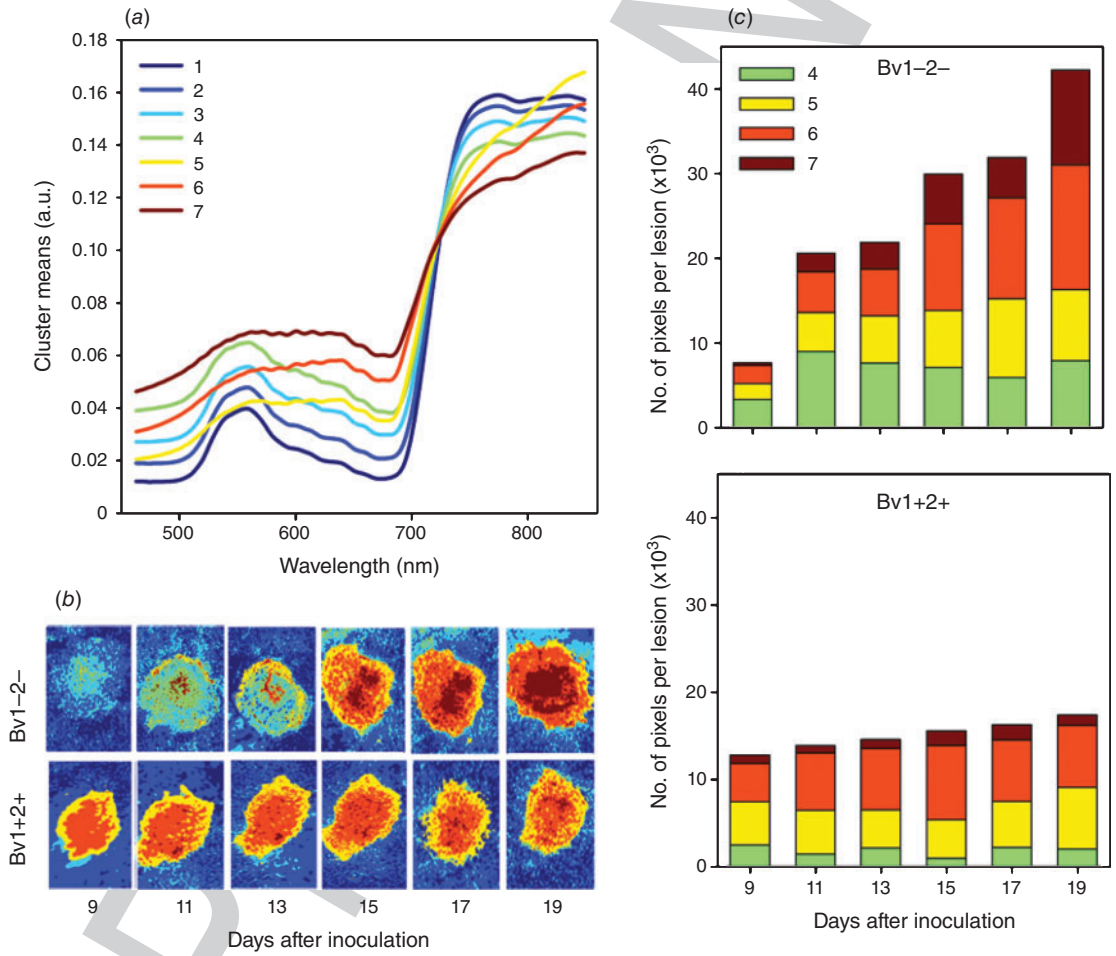

Fig. 5. Results of clustering hyperspectral images of sugar beet genotypes Bv1–2– and Bv1+2+ 9 to 19 days after inoculation (dai) of *Cercospora beticola*. (*a*) Cluster means with unit length in arbitrary units (a.u.), (*b*) false-colour images of representative Cercospora leaf spot lesions (corresponding RGB images in Fig. 2*b*) and (*c*) quantified number of pixels per lesion in cluster 4–7 (n = 10). Number of clusters was preset to seven and clusters are represented by the respective colour shown in legend (*a*).

genotype. In contrast, pathogens try to circumvent or inhibit plant defence. Schmidt and colleagues assumed that *C. beticola* suppresses the plant defence by ABA accumulation which led to a reduction of phenylalanin ammonium lyase (PAL) expression and activity (Schmidt *et al*. 2004, 2008). Moreover, ABA treatment has been shown to inhibit phytoalexin accumulation in potato (Henfling *et al*. 1980). Further investigations are needed to reveal details of the mechanisms of CLS resistance which are more difficult to investigate and to understand because of the long asymptomatic (latent) period of CLS

compared with e.g. downy and powdery mildews (Oerke *et al.* 2006; Mahlein *et al.* 2012; Wahabzada *et al.* 2015). For instance, the resistance of barley to *Blumeria graminis* f.sp. *hordei* causing powdery mildew largely depends on the tissue reaction within 48 h after inoculation (Hückelhoven and Panstruga 2011), thus hyperspectral imaging approaches enable an identification of resistant genotypes within the first days after inoculation (Kuska *et al.* 2015).

Recently, it was shown that the resistance level of a sugar beet genotype influences the centre-to-margin ratio of CLS lesions (Leucker *et al.* 2016). With a supervised classification algorithm, the lesions could be differentiated into subareas and categorised by their composition. For this approach, a detailed and elaborate extraction of reference spectra is necessary, which needs background knowledge and expertise. In phenotyping procedures unsupervised data analysis is more suitable (Wahabzada *et al.* 2015). They offer the possibility to detect new resistance phenotypes and also quantitative effects. Hyperspectral signatures can comprise several hundred reflectance values of the VIS, NIR and short wave-infrared range, providing physiological and structural information and various possibilities for interpretation (Wahabzada *et al.* 2016). In addition to the high dimensionality, data are often recorded over time which leads to big datasets. Data mining and machine learning methods are able to handle large data amounts and have been successfully used for disease detection (reviewed by Singh *et al.* 2016). For a distinction between diseased and healthy samples, a simple segmentation into two clusters is sufficient. For example, Bergsträsser *et al.* (2015) used k-means clustering to distinguish CLS-infected and non-infected leaf tissue. In order to identify phenotypic differences in disease resistance a finer differentiation is required. Therefore, several numbers of clusters were tested on one representative time point, before clustering the full dataset to analyse the CLS dynamics. Four of the seven clusters could be assigned to characteristic areas of the CLS lesions (early stage of CLS, outer and inner area of the lesion, lesion centre on susceptible genotype) and revealed the abrupt and limited lesion formation on Bv1+2+ and the slower, but more intensive colonisation on Bv1–2–. The reflectance of the non-infected tissue was also varying depending, for example, on the inclination of the leaf surface or leaf veins. This heterogeneity was reflected in the first three clusters.

Hyperspectral imaging provided very accurate analysis of disease symptoms and host plant (resistance) reaction to *C. beticola*. The resistance QTL seemed to enhance the reaction of the QTL-carrying genotype resulting in a limited CLS lesion development. The spectral characterisation of the breeding lines and a robust differentiation of finer differences in quantitative resistance may help to enhance the selection process in resistance breeding due to its sensitivity and potential for automation. This approach can be integrated in a pre-selection of breeding material before field trials or for a more detailed analysis of promising candidate lines.

Acknowledgements

The project is supported by funds of the German Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support program. Funding of Mirwaes Wahabzada and Anne-Katrin Mahlein was provided by the German Federal Ministry of Education and Research (BMBF) within the scope of the competitive grants program 'Networks of excellence in agricultural and nutrition research – CROP.SENSe.net' (Funding code: 0315529), junior research group 'Hyperspectral phenotyping of resistance reactions of barley'.

References

Bergsträsser S, Fanourakis D, Schmittgen S, Cendrero-Mateo MP, Jansen M, Scharr H, Rascher U (2015) HyperART: non-invasive quantification of leaf traits using hyperspectral absorption-reflectance-transmittance imaging. *Plant Methods* **11**, 1–17. doi:10.1186/s13007-015-0043-0

Bishop CM (2006) 'Pattern recognition and machine learning.' (Ed. M Jordan) (Springer Science + Business Media: New York)

Blackburn GA (1998) Spectral indices for estimating photosynthetic pigment concentrations : a test using senescent tree leaves. *International Journal of Remote Sensing* **19**, 657–675. doi:10.1080/014311698215919

Bock CH, Poole GH, Parker PE, Gottwald TR (2010) Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. *Critical Reviews in Plant Sciences* **29**, 59–107. doi:10.1080/07352681003617285

Dasgupta S, Freund Y (2009) Random projection trees for vector quantization. *IEEE Transactions on Information Theory* **55**, 3229–3242. doi:10.1109/TIT.2009.2021326

Daub ME, Ehrenshaft M (2000) The photoactivated *Cercospora* toxin cercosporin: contributions to plant disease and fundamental biology. *Annual Review of Phytopathology* **38**, 461–490. doi:10.1146/annurev.phyto.38.1.461

Dhillon IS, Modha DS (2001) Concept decompositions for large sparse text data using clustering. *Machine Learning* **42**, 143–175. doi:10.1023/A:1007612920971

Feindt F, Mendgen K, Heitefuss R (1981) Feinstruktur unterschiedlicher Zellwandreaktionen im Blattparenchym anfaelliger und resistenter Rueben (*Beta vulgaris* L.) nach Infektion durch *Cercospora beticola* Sacc. *Phytopathologische Zeitschrift* **101**, 248–264. doi:10.1111/j.1439-0434.1981.tb03346.x

Gamon JA, Peñuelas J, Field CB (1992) A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing of Environment* **41**, 35–44. doi:10.1016/0034-4257(92)90059-S

Geiger H, Heun M (1989) Genetics of quantitative resistance to fungal diseases. *Annual Review of Phytopathology* **27**, 317–341. doi:10.1146/annurev.py.27.090189.001533

Geigert J, Schmitz FR, Johnson G, Maag DD, Johnson DK (1973) Two phytoalexins from sugar beet (*Beta vulgaris*) leaves. *Tetrahedron* **29**, 2703–2706. doi:10.1016/S0040-4020(01)93389-7

Goudet C, Milat ML, Sentenac H, Thibaud JB (2000) Beticolins, nonpeptidic, polycyclic molecules produced by the phytopathogenic fungus *Cercospora beticola*, as a new family of ion channel-forming toxins. *Molecular Plant-Microbe Interactions* **13**, 203–209. doi:10.1094/MPMI.2000.13.2.203

Henfling JWDM, Bostock R, Kuc J (1980) Effect of abscisic acid on rishitin and lubimin accumulation and resistance to *Phytophthora infestans* and *Cladosporium cucumerinum* in potato tuber tissue slices. *Phytopathology* **70**, 1074–1078. doi:10.1094/Phyto-70-1074

Hückelhoven R, Panstruga R (2011) Cell biology of the plant-powdery mildew interaction. *Current Opinion in Plant Biology* **14**, 738–746. doi:10.1016/j.pbi.2011.08.002

Kersting K, Wahabzada M, Thurau C, Bauckhage C (2010) Hierarchical convex NMF for clustering massive data. In '2nd Asian conference machine learning'. pp. 253–268. (JMLR Workshop and Conference Proceedings: Tokyo, Japan)

Johnson R, Taylor AJ (1976) Spore yield of pathogens in investigations of the race – specificity of host resistance. *Annual Review of Phytopathology* **14**, 97–119. doi:10.1146/annurev.py.14.090176.000525

Koch G, Jung C, Asher MJC, Holtschulte B, Molard MR, Rosso F, Steinrücken G, Beckers R (2000) Genetic localization of *Cercospora* resistance genes. '*Cercospora beticola* Sacc. biology, agronomic influence and control measures in sugar beet.' (Eds MJC Asher, B Holtschulte, MR Molard, F Rosso, G Steinrücken, R Beckers) pp. 197–209. (International Institute for Beet Research: Brussels)

Kuska M, Wahabzada M, Leucker M, Dehne H-W, Kersting K, Oerke E-C, Steiner U, Mahlein A-K (2015) Hyperspectral phenotyping on the microscopic scale: towards automated characterization of plant-pathogen interactions. *Plant Methods* **11**, 28. doi:10.1186/s13007-015-0073-7

Leucker M, Mahlein A-K, Steiner U, Oerke E-C (2016) Improvement of lesion phenotyping in *Cercospora beticola* – sugar beet interaction by hyperspectral imaging. *Phytopathology* **106**, 177–184. doi:10.1094/PHYTO-04-15-0100-R

Lewellen RT, Whitney ED (1976) Inheritance of resistance to race C2 of *Cercospora beticola* in sugarbeet. *Crop Science* **16**, 558–561. doi:10.2135/cropsci1976.0011183X001600040032x

Luc B, Deronde B, Kempeneers P, Debruyn W, Provoost S, Sensing R, Observation E (2005) Optimized spectral angle mapper classification of spatially heterogeneous dynamic dune vegetation, a case study along the Belgian coastline. In 'The 9th international symposium on physical measurements and signatures in remote sensing'. (ISPMSRS: Beijing)

Mahlein A-K (2016) Plant disease detection by imaging sensors – parallels and specific demands for precision agriculture and plant phenotyping. *Plant Disease* **100**, 241–251. doi:10.1094/PDIS-03-15-0340-FE

Mahlein A-K, Steiner U, Dehne H-W, Oerke E-C (2010) Spectral signatures of sugar beet leaves for the detection and differentiation of diseases. *Precision Agriculture* **11**, 413–431. doi:10.1007/s11119-010-9180-7

Mahlein A-K, Steiner U, Hillnhütter C, Dehne H-W, Oerke E-C (2012) Hyperspectral imaging for small-scale analysis of symptoms caused by different sugar beet diseases. *Plant Methods* **8**, 3. doi:10.1186/1746-4811-8-3

Mahlein A-K, Rumpf T, Welke P, Dehne H-W, Plümer L, Steiner U, Oerke E-C (2013) Development of spectral indices for detecting and identifying plant diseases. *Remote Sensing of Environment* **128**, 21–30. doi:10.1016/j.rse.2012.09.019

Martin SS (1977) Accumulation of the flavonoids betagarin and betavulgarin in Beta vulgaris infected by the fungus Cercospora beticola. *Physiological Plant Pathology* **11**, 297–303. doi:10.1016/0048-4059(77)90072-8

Meier U (2001) Entwicklungsstadien mono- und dikotyler Pflanzen. *Biologische Bundesanstalt für Land und Forstwirtschaft* **2**, 1–165.

Merzlyak MN, Gitelson AA, Chivkunova OB, Rakitin VY (1999) Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiologia Plantarum* **106**, 135–141. doi:10.1034/j.1399-3054.1999.106119.x

Montes JM, Melchinger AE, Reif JC (2007) Novel throughput phenotyping platforms in plant genetic studies. *Trends in Plant Science* **12**, 433–436. doi:10.1016/j.tplants.2007.08.006

Munerati O, Mezzadroli G, Zapparoli T (1913) Osservazioni sulla *Beta maritima* L., nel triennio 1910–1912. *Le Stazioni Sperimentali Agrarie Italiane* **46**, 415–444.

Mutka AM, Bart RS (2015) Image-based phenotyping of plant disease symptoms. *Frontiers in Plant Science* **5**, 734. doi:10.3389/fpls.2014.00734

Oerke E-C, Steiner U, Dehne HW, Lindenthal M (2006) Thermal imaging of cucumber leaves affected by downy mildew and environmental conditions. *Journal of Experimental Botany* **57**, 2121–2132. doi:10.1093/jxb/erj170

Peñuelas J, Filella I (1998) Visible and near-infrared reflectance techniques for diagnosing plant physiological status. *Trends in Plant Science* **3**, 151–156. doi:10.1016/S1360-1385(98)01213-8

Rossi V, Battilani P, Chiusa G, Giosuè S, Languasco L, Racca P (1999) Components of rate-reducing resistance to Cercospora leaf spot in sugar beet: incubation length, infection efficiency, lesion size. *Journal of Plant Pathology* **81**, 25–35.

Rossi V, Battilani P, Chiusa G, Giosuè S, Languasco L, Racca P (2000) Components of rate-reducing resistance to Cercospora leaf spot in sugar beet: conidiation length, spore yield. *Journal of Plant Pathology* **82**, 125–131.

Savitzky A, Golay MJE (1964) Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry* **36**, 1627–1639. doi:10.1021/ac60214a047

Schmidt K, Heberle B, Kurrasch J, Nehls R, Stahl DJ (2004) Suppression of phenylalanine ammonia lyase expression in sugar beet by the fungal pathogen *Cercospora beticola* is mediated at the core promoter of the gene. *Plant Molecular Biology* **55**, 835–852. doi:10.1007/s11103-005-2141-2

Schmidt K, Pflugmacher M, Klages S, Mäser A, Mock A, Stahl DJ (2008) Accumulation of the hormone abscisic acid (ABA) at the infection site of the fungus *Cercospora beticola* supports the role of ABA as a repressor of plant defence in sugar beet. *Molecular Plant Pathology* **9**, 661–673. doi:10.1111/j.1364-3703.2008.00491.x

Scholes JD, Farrar JF (1986) Increased rates of photosynthesis in localized regions of a barley leaf infected with brown rust. *New Phytologist* **104**, 601–612. doi:10.1111/j.1469-8137.1986.tb00660.x

Setiawan A, Koch G, Barnes SR, Jung C (2000) Mapping quantitative trait loci (QTLs) for resistance to Cercospora leaf spot disease (*Cercospora beticola* Sacc.) in sugar beet (*Beta vulgaris* L.). *Theoretical and Applied Genetics* **100**, 1176–1182. doi:10.1007/s001220051421

Shane WW, Teng PS (1992) Impact of *Cercospora leaf* spot on root weight, sugar yield, and purity of *Beta vulgaris*. *Plant Disease* **76**, 812–820. doi:10.1094/PD-76-0812

Singh A, Ganapathysubramanian B, Singh AK, Sarkar S (2016) Machine learning for high-throughput stress phenotyping in plants. *Trends in Plant Science* **21**, 110–124. doi:10.1016/j.tplants.2015.10.015

Smith GA, Gaskill JO (1970) Inheritance of resistance to Cercospora leaf spot in sugarbeet. *Journal of the American Society of Sugar Beet Technologists* **16**, 172–180. doi:10.5274/jsbr.16.2.172

Wahabzada M, Mahlein A-K, Bauckhage C, Steiner U, Oerke E-C, Kersting K (2015) Metro maps of plant disease dynamics-automated mining of differences using hyperspectral images. *PLoS One* **10**, e0116902. doi:10.1371/journal.pone.0116902

Wahabzada M, Mahlein A-K, Bauckhage C, Steiner U, Oerke E-C, Kersting K (2016) Plant phenotyping using probabilistic topic models: uncovering the hyperspectral language of plants. *Scientific Reports* **6**, 22482. doi:10.1038/srep22482

Weltmeier F, Mäser A, Menze A, Hennig S, Schad M, Breuer F, Schulz B, Holtschulte B, Nehls R, Stahl DJ (2011) Transcript profiles in sugar beet genotypes uncover timing and strength of defense reactions to *Cercospora beticola* infection. *Molecular Plant-Microbe Interactions* **24**, 758–772. doi:10.1094/MPMI-08-10-0189