

# **Towards automated phenotyping in plant tissue culture**

Von der Naturwissenschaftlichen Fakultät der  
Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades  
Doktor der Gartenbauwissenschaften (Dr. rer. hort.)

genehmigte Dissertation

von

Hans Lukas Bethge, M. Sc.

2023

Referentin: Prof. Dr. rer. hort. habil. Traud Winkelmann

Korreferent: Prof. Dr. rer. hort. habil. Thomas Rath

Tag der Promotion: 18.08.2023

## **Abstract**

Plant in vitro culture techniques comprise important fundamental methods of modern plant research, propagation and breeding. Innovative scientific approaches to further develop the cultivation process, therefore, have the potential of far-reaching impact on many different areas. In particular, automation can increase efficiency of in vitro propagation, a domain currently constrained by intensive manual labor. Automated phenotyping of plant in vitro culture bears the potential to extend the evaluation of in vitro plants from manual destructive endpoint measurements to continuous and objective digital quantification of plant traits. Consequently, this can lead to a better understanding of crucial developmental processes and will help to clarify the emergence of physiological disorders of plant in vitro cultures.

The aim of this dissertation was to investigate and exemplify the potential of optical sensing methods and machine learning in plant in vitro culture from an interdisciplinary point of view. A novel robotic phenotyping system for automated, non-destructive, multi-dimensional in situ detection of plant traits based on low-cost sensor technology was conceptualized, developed and tested. Various sensor technologies, including an RGB camera, a laser distance sensor, a micro spectrometer, and a thermal camera, were applied partly for the first time under these challenging conditions and evaluated with respect to the resulting data quality and feasibility. In addition to the development of new dynamic, semi-automated data processing pipelines, the automatic acquisition of multisensory data across an entire subculture passage of plant in vitro cultures was demonstrated. This allowed novel time series images of different developmental processes of plant in vitro cultures and the emergence of physiological disorders to be captured in situ for the first time. The digital determination of relevant parameters such as projected plant area, average canopy height, and maximum plant height, was demonstrated, which can be used as critical descriptors of plant growth performance in vitro. In addition, a novel method of non-destructive quantification of media volume by depth data was developed which may allow monitoring of water uptake by plants and evaporation from the culture medium.

The phenotyping system was used to investigate the etiology of the physiological growth anomaly hyperhydricity. Therefore, digital monitoring of the morphology and along with spectroscopic studies of reflectance behavior over time were conducted. The new optical characteristics identified by classical spectral analysis, such as reduced reflectance and major absorption peaks of hyperhydricity in the SWIR region could be validated to be the main discriminating features by a trained support vector machine with a balanced accuracy of 84% on test set, demonstrating the feasibility of a spectral detection of hyperhydricity. In addition, an RGB image dataset was used for automated detection of hyperhydricity using deep neural networks. The high-performance metrics with precision of 83.8% and recall of 95.7% on test images underscore the presence of for detection sufficient number of discriminating features within the spatial RGB data, thus a second approach is proposed for automatic detection of hyperhydricity based on RGB images.

The resulting multimodal sensor data sets of the robotic phenotyping system were tested as a supporting tool of an e-learning module in higher education to increase the digital skills in the field of sensing, data processing and data analysis, and evaluated by means of a student survey. This proof-of-concept study revealed an overall high level of acceptance and advocacy by students with 70% good to very good rating. However, with increased complexity of the learning task, students experienced excessive demands and rated the respective session lower.

In summary, this study is expected to pave the way for increased use of automated sensor-based phenotyping in conjunction with machine learning in plant research and commercial micropropagation in the future.

Keywords: In vitro culture, phenotyping, imaging, hyperhydricity, sensors, machine learning, teaching

### Zusammenfassung

Die pflanzliche In-vitro-Kultur umfasst wichtige grundlegende Methoden der modernen Pflanzenforschung, -vermehrung und -züchtung. Innovative wissenschaftliche Ansätze zur Weiterentwicklung des Kultivierungsprozess können daher weitreichenden Einfluss auf viele unterschiedliche Bereiche haben. Insbesondere die Automatisierung kann die Effizienz der In-vitro- Vermehrung steigern, die derzeit durch die intensive manuelle Arbeit beschränkt wird. Automatisierte Phänotypisierung von In-vitro-Kulturen ermöglicht es, die Erfassung von manuellen destruktiven Endpunktmessungen auf eine kontinuierliche, objektive und digitale Quantifizierung der Pflanzenmerkmale auszuweiten. Dies kann zu einem besseren Verständnis entscheidender Entwicklungsprozesse führen und die Entstehung physiologischer Störungen zu klären.

Ziel dieser Dissertation war es, das Potential optischer Erfassungsmethoden und des maschinellen Lernens für die pflanzliche In-vitro-Kultur unter interdisziplinären Gesichtspunkten zu untersuchen und exemplarisch aufzuzeigen. Ein neuartiger Phänotypisierungsroboter zur automatisierten, zerstörungsfreien, mehrdimensionalen In-situ-Erfassung von Pflanzenmerkmalen wurde auf Basis kostengünstiger Sensortechnik entwickelt. Unterschiedliche Sensortechnologien, darunter eine RGB-Kamera, ein Laser-Distanzsensor, ein Mikrospektrometer und eine Wärmebildkamera, wurden teils zum ersten Mal unter diesen schwierigen Bedingungen eingesetzt und im Hinblick auf die resultierende Datenqualität und Realisierbarkeit bewertet. Neben der Entwicklung dynamischer, halbautomatischer Datenverarbeitungspipelines, wurde die automatische Erfassung multisensorischer Daten über eine gesamte Subkulturpassage der In-vitro-Kulturen demonstriert. Dadurch konnte erstmals Zeitrafferaufnahmen verschiedener Entwicklungsprozesse von pflanzlichen In-vitro-Kulturen und das Auftreten von physiologischen Störungen in situ erfasst werden. Die digitale Bestimmung relevanter Kenngrößen wie der projizierten Pflanzenfläche, der durchschnittlichen Bestandhöhe und der maximalen Pflanzenhöhe wurde demonstriert, die als wichtige Deskriptoren für das pflanzliche Wachstum dienen können. Darüber hinaus konnte eine neue Methode für die Pflanzenwissenschaften entwickelt werden, um die Wasseraufnahme von Pflanzen und die Verdunstung von Kulturmedien auf der Grundlage einer zerstörungsfreien Quantifizierung des Medienvolumens zu überwachen.

Der Phänotypisierungsroboter wurde zur Untersuchung der Entstehung der Wachstumsanomalie Hyperhydrizität eingesetzt. Hierfür wurden ein digitales Monitoring der Morphologie der Explantate mit begleitenden spektroskopischen Untersuchungen des Reflexionsverhaltens im Zeitverlauf durchgeführt. Die durch Spektralanalyse identifizierten optischen Merkmale, wie den reduzierter Reflexionsgrad und die Hauptabsorptionspeaks der Hyperhydrizität in der SWIR-Region, konnten als die wichtigsten Unterscheidungsmerkmale durch ein Support-Vektor-Maschine-Model mit einer Genauigkeit von 84% auf dem Testsatz validiert werden und damit Machbarkeit der spektrale Identifizierung von Hyperhydrizität aufzeigen. Darüber wurde für die automatische Detektion der Hyperhydrizität auf Basis von RGB-Bildern ein neuronales Netz trainiert. Die hohen Kennzahlen im Testdatensatz wie die Präzision von 83,8 % und einem Recall von 95,7 % unterstreichen das Vorhandensein einer für die Erkennung ausreichenden Anzahl von Unterscheidungsmerkmalen innerhalb der räumlichen RGB-Daten. Somit konnte ein zweiter Ansatz der automatischen Detektion von Hyperhydrizität durch RGB-Bilder präsentiert werden.

Die resultierenden Sensordatensätze des Phänotypisierungsroboters wurden als unterstützendes Werkzeug eines E-Learning Moduls zur Steigerung digitaler Kompetenzen im Bereich Sensortechnik, Datenverarbeitung und -auswertung in der Hochschulausbildung erprobt und anhand der Befragung von Studierenden evaluiert. Diese Machbarkeitsstudie ergab eine insgesamt hohe Akzeptanz durch die Studierenden mit 70% guten bis sehr guten Bewertungen. Mit zunehmender Komplexität der Lernaufgabe fühlten sich die Studierenden jedoch überfordert und bewerteten die jeweilige Session schlechter.

Zusammenfassend zielt diese Arbeit darauf ab den Weg für einen verstärkten Einsatz der automatisierten, sensorbasierten Phänotypisierung in Kombination mit den Techniken des maschinellen Lernens der Forschung und der kommerziellen Mikrovermehrung zukünftig zu ebnet.

Schlüsselwörter: In-vitro-Kultur, Phänotypisierung, Bildgebung, Hyperhydrizität, Sensoren, maschinelles Lernen, Lehre

## **Abbreviations**

ANN	Artificial neural networks
CNN	Convolutional neural networks
DNN	Deep neural networks
HH	Hyperhydricity
HTP	High-throughput plant phenotyping
LDA	Linear discriminant analysis
LIDAR	Light detection and ranging
ML	Machine learning
MOOC	Massive open online courses
NIR	Near-infrared
PCA	Principal component analysis
PGRs	Plant growth regulators
PLS	Partial least square
RF	Random forest
SVM	Supported vector machines
SWIR	Shortwave infrared
ToF	Time-of-flight
UV	Ultraviolet
VIS	Visible

## Table of Contents

<b>Abstract</b> .....	<b>I</b>
<b>Zusammenfassung</b> .....	<b>II</b>
<b>Abbreviations</b> .....	<b>III</b>
<b>Table of Contents</b> .....	<b>IV</b>
<b>1 General Introduction</b> .....	<b>1</b>
1.1 Hyperhydricity - a major bottleneck in plant in vitro culture.....	1
1.2 Automated phenotyping in plant in vitro culture.....	3
1.3 Challenges of optical sensing methods in plant in vitro culture.....	7
1.4 Machine learning in plant in vitro culture .....	9
1.5 Digital qualification in higher education in plant science .....	12
1.6 Thesis objectives.....	14
<b>2 Manuscripts</b> .....	<b>15</b>
2.1 Low-cost and automated phenotyping system “Phenomenon” for multi-sensor in situ monitoring in plant in vitro culture .....	15
2.2 Towards automated detection of hyperhydricity in plant in vitro culture .....	49
2.3 Remote plant sensing and phenotyping – an e-learning tool in higher education.....	81
<b>3 Summarizing discussion and further perspective</b> .....	<b>94</b>
3.1 Automated phenotyping in commercial micropropagation .....	94
3.2 Automated detection of morpho-physiological disorders in plant in vitro culture....	98
3.3 Application of automated phenotyping robot in higher education .....	101
3.4 Summary with critical review of the state of fulfillment.....	103
<b>4 Outlook</b> .....	<b>105</b>
<b>5 References</b> .....	<b>107</b>
<b>6 Curriculum vitae</b> .....	<b>112</b>
<b>7 List of publications</b> .....	<b>114</b>
7.1 Peer reviewed publications .....	114
7.2 Non reviewed publications .....	114
7.3 Conference contributions.....	114
<b>8 Acknowledgement</b> .....	<b>116</b>

# 1 General Introduction

## 1.1 Hyperhydricity - a major bottleneck in plant in vitro culture

Plant in vitro culture techniques form the basis of most biotechnological methods for plant breeding and propagation. Plant in vitro culture enables clonal mass propagation of valuable plants, supports plant breeding by regeneration from single cells, preserves endangered plant species in gene banks and allows efficient production of secondary metabolites by medical plants, among others (Birnbaum & Alvarado 2008, George et al. 2008). Although data on in vitro plant production on a worldwide scale are lacking — probably due to the withholding of production data by global players in commercial micropropagation, and being composed by different horticultural sectors such as ornamental, forestry, fruit and aquatic plants — for Germany, these data are available and revealed the main micropropagated plant genera to be *Phalaenopsis spp.*, *Rubus spp.* and *Helleborus spp.* (Winkelmann et al. 2006, Hutter & Schneider 2019). Other globally important micropropagated plant genera are *Musa spp.*, *Solanum spp.*, *Fragaria spp.*, *Vaccinium spp.* and others (Gamborg 2002, Podwyszyńska et al. 2022).

Plant in vitro propagation can reach high multiplication factors in short time and allows to produce genetically uniform and disease-free explants in many plant species. However, due to the special conditions of explants cultivated aseptically in closed containers, under a high relative humidity, with a reduced gas exchange and further stress caused by plant growth regulators (PGRs), physiological and morphological malformations may occur (Hazarika et al. 2006). In addition to recalcitrant and habitual behaviour (Gaspar et al. 2000, Abdalla et al. 2022) of some plant species in response to PGRs and other disorders such as shoot-tip necroses, fasciation and somaclonal variation, hyperhydricity (HH) is one of the major bottlenecks of efficient micropropagation in several plant species (Hazarika & Bora 2008, Ruffoni & Savona 2013).

This phenomenon also known as vitrification, glassiness or translucency, is characterized by a water-soaked, translucent, curled and fragile appearance. HH negatively affects plant quality, propagation rate, adventitious root and shoot formation and ex vitro survival rate (Debergh et al. 1992, Gribble 1999, Cardoso et al. 2018). Despite these drawbacks and the fact that at least 150 plant species can be affected seriously by HH (Kemat et al. 2020), the causes and underlying physiological mechanism of HH are still not fully understood. However, Rojas-Martínez et al. (2010) and van den Dries et al. (2013) provided strong evidence that in HH etiology the flooding of the apoplast, resulting in hypoxia and oxidative stress plays a major role. High water availability was identified as one of the key triggers of HH in several studies (Smith & Spomer 1995, Casanova et al. 2008, van den Dries et al. 2013, De Klerk et al. 2015, Kemat et al. 2021). This fact drastically restricts the usage of liquid culture and bioreactor systems, although they would allow very high multiplication rates and saving of cost-intensive gelling agents.

Paques et al. (1985) refer to HH as an inducible and reversible phenomenon and demonstrated that *Malus sp.* 'M26' plantlets could return to non-hyperhydric state if the induction phase in liquid culture did not exceed five days or the symptoms of HH are not too severe. In commercial in vitro laboratories, visual monitoring for contaminations and disorders is part of the routine work and therefore a costly and time-consuming repetitive matter (Mestre et al. 2017). Automation of processes offers great economic potential for micropropagation laboratories since 60 - 70% of total costs of a micropropagated plant is due to manual labor (Chen 2016).



## **1.2 Automated phenotyping in plant in vitro culture**

Plant phenotyping refers to a systematic and quantitative determination of the plant's anatomical, ontogenetic, physiological and biochemical properties (Guo & Zhu 2006). High-throughput plant phenotyping (HTP) is often realized by optical sensor-based approaches allowing a non-invasive, fast and objective quantification of plant traits while replacing time-consuming manual evaluation (Walter et al. 2015). Here, automation of data acquisition or processing meets the needs of plant breeding for measuring large populations of plants to select superior individuals (Dhondt et al. 2013). Recent technological advances in imaging sensors such as LIDAR (light detection and ranging) 3D sensors and hyperspectral cameras and in combination with rapid progress in machine learning turns automated HTP into a promising tool for plant research and production.

Applications of this discipline range from phenotyping of plant canopies with field-based platforms (Busemeyer et al. 2013) down to the phenotyping of tissue, organ and cell cultures in laboratory systems (Dhondt et al. 2014). Relevant performance metrics of plant in vitro cultures such as multiplication rate, plant quality and biomass are usually accessed manually and limited to single endpoint measurements of subculture passages. So far, very limited research using automated sensors in plant tissue culture has been reported and most are restricted to “plant to sensor” approaches (Table 1).

## General Introduction

**Table 1:** Summary table of selected literature of phenotyping of plant in vitro cultures, based on internet research using a combination of the keywords: plant in vitro culture, image analysis, phenotyping, monitoring

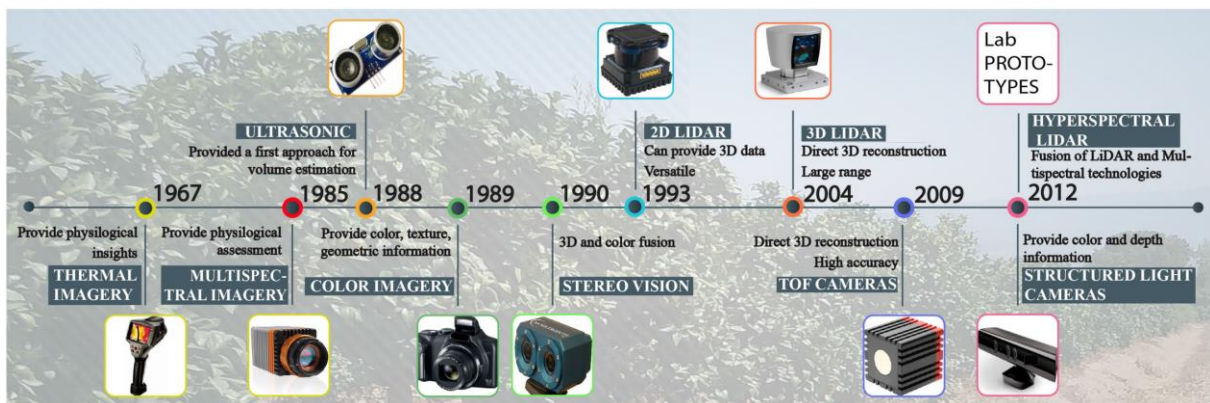
Automation/ Authors	Imaging setup	Dimension- ality	Digital parameters	Plant species	Culture phase	Aim
<b>Plant to sensor</b>						
Smith et al. 1989	Side view camera	2D- Mono- chrome	Plant height, plant area	<i>Acer sp.</i> <i>Daphne sp.</i> <i>Malus sp.</i> <i>i.a.</i>	Shoot culture	Correlating pro- jected plant area to plant height and fresh mass
Aynalem et al. 2006	Top view camera	2D-RGB	Vegetation indices	<i>Pyrus sp.</i>	Shoot culture	Monitoring plant quality of in vitro conserved plants
Ibaraki & Gupta 2011	Handheld camera	2D- Thermal	Leaf temper- ature	<i>Solanum sp.</i>	Shoot culture	Prediction of wilting after acclimatization
Mansouri et al. 2016	Top view camera	2D-RGB	Plant area, shape analy- sis	<i>Cuminum sp.</i>	Callus culture	Prediction of fresh mass and volume
Gupta & Karmakar 2017	Top view camera	2D-RGB	Plant area, shape analy- sis	<i>Swertia sp.</i>	Shoot culture	Computer vision- assisted evaluation
Mestre et al. 2017	Top view camera	2D- RGB & NIR	Plant area, vegetation indices	<i>Nandina sp.</i>	Shoot culture	Monitoring plant quality
Faragó et al. 2018	Top view camera	2D-RGB	Plant area, shape analy- sis	<i>Arabidopsis sp.</i>	In vitro germi- nation	Image processing software “PlantSize”
<b>Sensor to plant<sup>[1]</sup></b>						
Dhondt et al. 2014	Top view 2D camera with rotary carousel for petri dishes “IGIS”	2D-NIR	Plant area, shape analy- sis	<i>Arabidopsis sp.</i>	In vitro germi- nation	Live-monitoring of growth curves during cultivation
Barbez et al. 2017	Side view 2D camera with rotary carousel for petri dishes	2D-RGB	n/a	<i>Arabidopsis sp.</i>	In vitro germi- nation	Hardware setup for live-monitoring during cultivation
Lube et al. 2022	Side view camera with rotary carousel for petri dishes	2.5D-RGB by stereo vision	Root area and length, plant area, shape analy- sis	<i>Arabidopsis sp.</i>	In vitro germi- nation	High resolution live-monitoring of root growth during cultivation

<sup>[1]</sup>Note: Selected literature with closest proximity to “Sensor to plant” approach, albeit with restricted transferability for “automated live-monitoring during cultivation”.

Automated phenotyping approaches with minimal invasiveness allowing a direct live-monitoring plant during cultivation are heavily focused on fundamental research in *Arabidopsis thaliana*. Although the importance of *A. thaliana* as a model plant for research is undisputed, *A. thaliana* is a poor model plant for commercial micropropagation due to

the small explant size and height, uniformly structured growth habitus (rosette), presence of radicle and lack of adventitious shoots (in vitro seed germination).

Previous research on automated phenotyping of in vitro plants has primarily concentrated on RGB imaging, determining 2D parameters such as projected plant area and morphological description of the plant objects. However, over the past two decades, several other sensor technologies have gained interest for plant research in field phenotyping or HTP under controlled conditions, whether due to cost reductions, advances in sensor technology, or the launch of the first commercial products (Fig. 1; Roitsch et al. 2019). The following sensor technologies are worth mentioning due to their great potential for plant research and production, while focusing on small portable sensor technology which may have potential application for plant in vitro cultures.



**Fig. 1: Evolution in the use of range and artificial vision sensors** for morphological characterization and fruit/plant detection. The years report the first use of these sensing systems for agricultural purposes (Narvaez et al. 2017).

Biomass quantification and shoot length estimation in field or controlled environment phenotyping systems are enabled by **2.5/3D-imaging** sensors (Li et al. 2014). Light detection and ranging (LIDAR) sensors operate by the Time-of-Flight (ToF) principle where the distance is calculated by measuring the time of a specific laser pulse required to reach an object and be reflected back to the sensor's detector. In contrast to laser distance sensors, LIDAR systems differ in their dimensionality of acquiring multiple spatial

information at once. Laser distance sensors determine the distance of a single point, 1D-LIDAR system can be seen a line scanning approach, while 2D-LIDAR systems measure a plane by capturing multiple lines. The key features of the technology are the high accuracy, robustness and the fast speed while acquiring 2.5D/3D point clouds. In addition, the reflectance of the object can be determined by the amount backscattered light intensity and used for segmentation of the point cloud (Lin 2015). In contrast to the LIDAR system as a sequential scanning approach, ToF-cameras simultaneously illuminate the whole image scene at once and detect the backscattered light by an imaging sensor, therefore offering real-time 3D sensing but have generally lower spatial resolution and lower accuracy. Other principles to be mentioned for generating depth data are i) stereo vision, where distance is calculated by triangulating two different views of an object, and ii) structured light, where depth data is calculated by deforming a grid of light through the object (Narvaez et al. 2017). To date, there are no reports on the application of these 2.5D/3D-imaging techniques to the quantification of biomass of plant in vitro cultures.

Plant sensing by **optical spectroscopy** has demonstrated significant potential for stress detection (Lichtenthaler & Rinderle, 1988, Lichtenthaler et al. 1998, Buschmann et al. 2000), quantification of plant metabolites (Schulz & Baranska 2007), and plant classification (Zomer et al. 2009), as evidenced by numerous publications (reviewed by Cavaco et al. 2022). The sub-disciplines of optical spectroscopy can be differentiated based on the type of the detected interaction between light and plant tissue, e.g., transmission spectroscopy, reflectance spectroscopy, and fluorescence spectroscopy, or by the spectral region in which the detector operates, such as ultraviolet (UV), visible (VIS), near-infrared (NIR), shortwave infrared (SWIR), or thermal. Here, each spectral region highlights different biochemical plant compounds, based on their specific absorption characteristics.

In particular **imaging spectroscopic sensors** based on reflectance such as multi-spectral cameras — imaging sensors with more than three or different to those of a RGB camera— and hyperspectral cameras — imaging sensors with more than 100 spectral channels — have gained interest due to their feasibility, wide applicability and huge content of information used for disease detection and plant monitoring ([Araus et al. 2018](#), [Roitsch et al. 2019](#)). Fluorescence imaging sensor are most commonly used to estimate the photosynthetic performances of plants based on the detection of chlorophyll fluorescence kinetics during the transition from dark to light, also known as Kautsky effect ([Kautsky & Hirsch 1931](#)). On the other hand, multispectral or hyperspectral fluorescence imaging spectroscopy, e.g., excited by UV radiation, is less common, although it holds great potential by providing physiological information transported within the autofluorescence signals of chlorophyll and phenolic compounds ([Pérez-Bueno et al. 2016](#)). Other promising non-destructive phenotyping sensor technologies such as Raman spectroscopy, tomographic imaging techniques, were neglected so far due to the lack of portable systems.

Finally, the emergence of innovative products with **sensor fusion** approaches such as Intel®RealSense™ (fusion of RGB with depth imaging), or PlantEye ([Hummel 2012](#), fusion of multispectral imaging with structured light). Furthermore, there are several research activities to fuse the strengths of different sensor technologies such as hyperspectral LIDAR ([Hakala et al. 2012](#), [Chen et al. 2018](#)) with no commercial product launched so far.

### **1.3 Challenges of optical sensing methods in plant in vitro culture**

For live-monitoring of the dynamic growth processes of plant in vitro cultures over longer periods of time it is essential to maintain the **aseptic conditions** of the explants — ensured by closed culture containers. Otherwise, bacterial and fungal contaminations would establish on the sugar containing culture media. Therefore, most imaging

approaches following two different strategies, either image acquisition takes place under a laminar flow hood (“plant-to-sensor approach”, [Table 1](#)) — where explants can be kept sterile even if the culture container is opened. However, during image acquisition explants are exposed to completely different environmental conditions (active air movement, lower relative humidity, different atmospheric composition). Alternatively, image acquisition during cultivation has to happen through the closed culture vessels facing numerous challenges, such as specular lighting ([Mestre et al. 2017](#)) and water condensation ([Faragó et. al 2018](#)).

Plant in vitro cultures are commonly cultivated in multi-layered shelf systems with tubular fluorescent lamps or LEDs. A typical distance between the cultivation area and the illumination is around 400 mm in each layer. Due to this **limitation in space** only small sensor systems are suitable, and in addition, selection of sensor systems is further restricted as their sensor technology has to perform in a close-range setup. This is of major concern for depth sensing sensor e.g., based on stereo vision ([Kazmi et al. 2014](#)).

The rel. humidity inside the culture vessels in general is higher than 95% ([Kozai 1991](#)). **Water condensation** occurs predominantly at the coldest surface, which is often the lid of the culture vessel due to the heat dissipation of the tubular fluorescent lamps below the cultivation area. The complex interaction of light and these condense water droplets, such as scattering, diffraction, refraction, dispersion and absorption complicate the imaging of the plant in vitro cultures. [Dhondt et al. \(2014\)](#) prevented the condensation of water by a heated air stream that shifted the dew point toward the culture medium. Based on the same principle, a so called “bottom-cooling” is well established in plant in vitro culture rooms and prevents water condensation ([Vanderschaeghe & Debergh 1987](#)).

Depending on plant species, established laboratory processes and in vitro culture phase, various types of culture containers are used in commercial micropropagation and plant research. While Petri dishes provide a clear view of the cultured tissue, certain types of plastic lids for culture vessels exhibit a **degree of opacity** despite a high degree of light transmission. The decreased visibility interferes with the accurate observation of plant tissues by imaging techniques, especially under certain lighting conditions where **total internal reflection** can occur. In addition, the **spectral transmittance** of the culture vessels, which is mainly determined by the plastic material, has to be considered for in vitro imaging approaches beyond the VIS region e.g., UV-excited fluorescence imaging.

In plant in vitro culture high humidity, hetero/mixotrophic growth, low light condition, impaired gaseous exchange lead to various physiological and anatomical variations compared to ex vitro cultivated plants. Among them are obvious macroscopic changes such as smaller and differently shaped leaves, as well as anatomically such as very thin or absent cuticle, permanently open stomata, irregular structured spongy parenchyma with larger intercellular air-spaces and others ([George et al. 2008](#)). The associated multiple scattering processes and differences in light path length call into question the transferability of ex vitro results from optical plant spectroscopy to spectroscopic monitoring of plants in vitro.

### 1.4 Machine learning in plant in vitro culture

Machine learning (ML) techniques as a generic term for statistical learning models are established as a state-of-the-art for processing and analyzing data of high dimensionality (less observations than describing features), complex interaction and non-linear relationships. ML-models applied to in plant tissue cultures were already reviewed by [Prasad & Gupta \(2008\)](#) and [Hesami & Jones \(2020\)](#) including, plantlets clustering ([Mahendra et al. 2004](#)), classification of somatic embryos ([Zhang et al. 1999](#)),

estimation of shoot length ([Honda et al. 1997](#)) and optimizing culture medium ([Nezami-Alanagh et al. 2019](#), [Hameg et al. 2020](#)). ML-models for classification can be grouped into two main types: unsupervised and supervised models. Supervised models aim to discriminate between classes based on known class memberships, while unsupervised methods attempt to identify underlying patterns or structures in the data without prior knowledge of class labels.

Principal component analysis (PCA, [Pearson 1901](#)) is a unsupervised dimension reduction method. Reduction of dimensionality is performed by the projection of the data into a hyperspace (linear combinations of original variables named principal components) of lower dimension, while maximizing the variance of the whole data set ([Ringnér 2008](#)). The principal components are uncorrelated and sorted with decreasing captured variance, thus the first principal component captures the largest variance (information) of the data set. The principal components can be used for outlier detection, noise filtering, data visualization, identification of correlated factors or as input with lower dimensionality for classification models like linear discriminant analysis.

Linear discriminant analysis (LDA) as a generalization of Fisher's linear discriminant ([Fisher 1936](#)) is a classification method for data of high-dimensionality based on supervised extraction of discriminating features by transformation. Reduction of dimensionality is performed by the projection of the data into a hyperspace of lower dimension, while minimizing intra-class variance and maximizing inter-class variance ([Du & Wang 2011](#)).

Partial least square (PLS) is a supervised dimension reduction method ([Wold 1984](#)). Reduction of dimensionality is performed by the projection of the data into a hyperspace (linear combinations of original variables named latent variables) of lower



dimension, while maximizing the variance of the whole data set and the covariance between the dependent variable (labels/class membership) and the latent variable (Lee et al. 2018).

Random Forest (RF) is an ensemble learning method of supervised tree-based models such as decision trees (Breiman 2001). With a random sub-selection of the data set, multiple tree-shaped models based on sequential decisions are constructed each on their subset of the data. The final random forest model aggregates the prediction of the decision trees and thus is less sensitive to overfitting.

Support vector machine (SVM) is supervised classification method (Cortes & Vapnik 1995). Following the transformation into a hyperspace, SVMs select the best class-separating hyperplane, while maximizing the distance of the two classes. The distance in SVM is the width of the margin of the hyperplane and reflects the space between the support vectors. These support vectors are the data points closest to the hyperplane and restricting the margin width.

Neural Networks, also known as Artificial Neural Networks (ANNs), are a family of highly flexible models (McCulloch & Pitts 1943). They are typically composed of different layers, including the input layer, hidden layer(s), and the output layer. In each layer, the output signal of a unit is calculated using the input value, weights, and bias, which are then passed on to all connected units in the next layer via an activation function, such as the sigmoid function. The model is trained to optimize the weights of each unit in order to predict the output with the lowest possible error (e.g. with gradient descent). This is achieved by repeatedly presenting the training data to the network, which adjusts its parameters through a process called backpropagation.

Networks with a certain depth of hidden layers are commonly called deep neural networks (DNN, [LeCun et al. 2015](#)). Convolutional neural networks (CNN) are DNN allowing grid-shaped inputs such as images. In a convolutional layer, additional features are extracted from the inputs by applying several different image processing filters, like edge detection by a sliding window (kernel). In the pooling layer the spatial information is reduced by pooling filters. A CNN has several convolutional layers and pooling layers generating features for a final fully connected layer, which can be seen as traditional neural network estimating a prediction.

To ensure strategic use of the potential of digital technologies such as machine learning and robotics, 26 European countries signed the 2019 Declaration for "a smart and sustainable digital future for European agriculture and rural areas" ([European Commission 2019](#)), emphasizing the need for strong digital skills in the future.

### **1.5 Digital qualification in higher education in plant science**

Along with the increased trend for digitalization and automation of plant production processes and in order to be able to understand and use the potential of modern machine learning techniques, a need arose to equip the students of "tomorrow" with a repertoire of digital competences. This is exemplified by the emergence of courses of study such as "precision farming/agriculture" (University of Applied Science "Ostwestfalen-Lippe", Germany) and "Agriculture Informatics" (Shobhit University, India; [Paul et al. 2020](#)) which incorporate teaching units on topics such as informatics, automation, robotics, and sensor technology as a common base in addition to biological topics. Furthermore, high ranked universities such as the University of Wageningen have also introduced a Massive Open Online Courses (MOOCs) "Smart agriculture" to provide a globally accessible tool to strengthen participants' digital skills in this highly interdisciplinary field. The COVID19 pandemic has caused further acceleration in the use

of digital formats such as e-learning, blended learning approaches or MOOCs in higher education (Bedenlier et al. 2021). However, digital formats carry the risk that the sustainability of knowledge transfer will decrease due to the increasing abstractness and the lack of practical application experience. The question therefore arises how the strengths of digital formats such as accessibility, flexibility, adaptivity, can be combined with other formats that establish personal knowledge and practical relevance in order to maintain a sustainable knowledge transfer in higher education. This is especially true for areas that are partly still perceived as abstract, such as “machine learning” and “sensor technology”.

## **1.6 Thesis objectives**

In order to investigate the potential and the required digital competences for the use of automation and optical sensor technology in horticulture using the labor-intensive plant in vitro culture as an example, this thesis was subjected to the following objectives:

- I. To develop and establish a low-cost phenotyping robot system suitable for direct live-monitoring of plant in vitro cultures during cultivation in an established multi-layered shelf system, scalable for high-throughput use in commercial laboratories and capable of monitoring a wide range of plant species and various different in vitro culture techniques. For this, low-cost sensor systems should be evaluated in regards to their feasibility and potential to quantify key growth parameters of plant in vitro cultures.
- II. To investigate the spectral fingerprints of hyperhydricity and to identify specific absorption features of hyperhydric tissues that are sufficient for discrimination by ML techniques. The morphological characteristics of hyperhydric explants in time-series image data should be described by in situ monitoring with the developed novel phenotyping system. In addition, optical technologies towards automated detection of hyperhydricity should be identified.
- III. To establish and explore the use of an interactive e-learning approach supported by low-cost sensor sets and phenotyping data in order to provide students in higher education with in-depth digital competence in the field of sensor technology in horticulture.

## 2 Manuscripts

### 2.1 Low-cost and automated phenotyping system “Phenomenon” for multi-sensor in situ monitoring in plant in vitro culture

Hans Bethge<sup>1,2</sup>, Traud Winkelmann<sup>2</sup>, Patrick Lüdeke<sup>3</sup>, Thomas Rath<sup>1</sup>

<sup>1</sup>Laboratory for Biosystems Engineering, Faculty of Agricultural Sciences and Landscape Architecture, Osnabrück University of Applied Sciences, 49090 Osnabrück, Germany

<sup>2</sup>Institute of Horticultural Production Systems, Section of Woody Plant and Propagation Physiology, Leibniz Universität Hannover, Herrenhäuser Str. 2, 30419, Hannover, Germany

<sup>3</sup>private person

Type of authorship:	First author
Type of article:	Research article
Status of article:	Published
Contribution to the article:	Developed the hardware and software of the system, designed and performed the experiments and analysed the data. Prepared the figures and wrote the manuscript.
Journal:	Plant Methods
Impact factor:	5.827 (2022-2023)
DOI:	10.1186/s13007-023-01018-w
Acknowledgment:	Reproduced with permission from Springer Nature

## RESEARCH

## Open Access



# Low-cost and automated phenotyping system “Phenomenon” for multi-sensor in situ monitoring in plant in vitro culture

Hans Bethge<sup>1\*</sup>, Traud Winkelmann<sup>2</sup>, Patrick Lüdeke<sup>3</sup> and Thomas Rath<sup>1</sup>**Abstract**

**Background** The current development of sensor technologies towards ever more cost-effective and powerful systems is steadily increasing the application of low-cost sensors in different horticultural sectors. In plant in vitro culture, as a fundamental technique for plant breeding and plant propagation, the majority of evaluation methods to describe the performance of these cultures are based on destructive approaches, limiting data to unique endpoint measurements. Therefore, a non-destructive phenotyping system capable of automated, continuous and objective quantification of in vitro plant traits is desirable.

**Results** An automated low-cost multi-sensor system acquiring phenotypic data of plant in vitro cultures was developed and evaluated. Unique hardware and software components were selected to construct a xyz-scanning system with an adequate accuracy for consistent data acquisition. Relevant plant growth predictors, such as projected area of explants and average canopy height were determined employing multi-sensory imaging and various developmental processes could be monitored and documented. The validation of the RGB image segmentation pipeline using a random forest classifier revealed very strong correlation with manual pixel annotation. Depth imaging by a laser distance sensor of plant in vitro cultures enabled the description of the dynamic behavior of the average canopy height, the maximum plant height, but also the culture media height and volume. Projected plant area in depth data by RANSAC (random sample consensus) segmentation approach well matched the projected plant area by RGB image processing pipeline. In addition, a successful proof of concept for in situ spectral fluorescence monitoring was achieved and challenges of thermal imaging were documented. Potential use cases for the digital quantification of key performance parameters in research and commercial application are discussed.

**Conclusion** The technical realization of “Phenomenon” allows phenotyping of plant in vitro cultures under highly challenging conditions and enables multi-sensory monitoring through closed vessels, ensuring the aseptic status of the cultures. Automated sensor application in plant tissue culture promises great potential for a non-destructive growth analysis enhancing commercial propagation as well as enabling research with novel digital parameters recorded over time.

**Keywords** Chlorophyll fluorescence, Image analysis, Laser distance sensor, Non-destructive growth analysis, Plant tissue culture, RGB imaging, Spectrometer, Thermal sensor

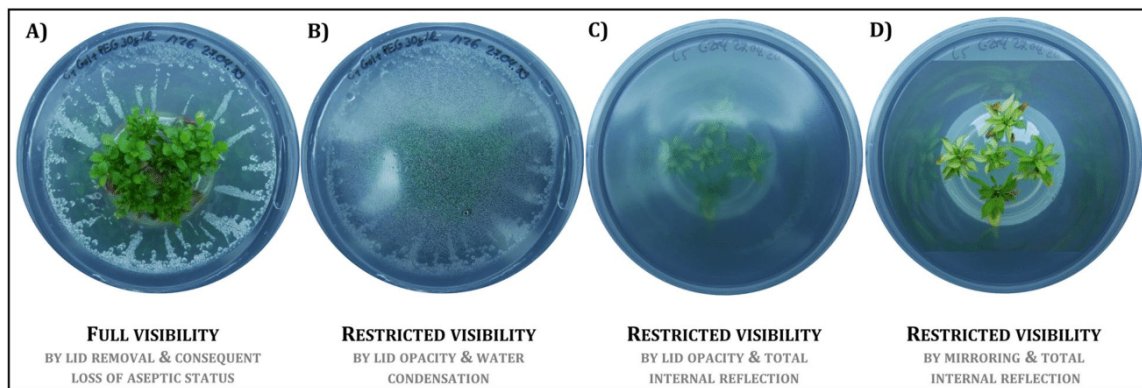
\*Correspondence:

Hans Bethge  
bethge@baum.uni-hannover.de

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.



**Fig. 1** Highly challenging imaging situation of plant in vitro cultures. **A** Culture vessel lid removal offers a proper imaging, but results in the loss of aseptic status of the cultures. Problems for optical monitoring arise from **B** water condensation, **C** opacity of culture containers and total internal reflection of ambient or detection light **D** as well as mirroring of plantlets inside the culture vessel. RGB images were taken from two shoot cultures of *Malus* spp. The image pairs **A–B** and **C–D** each show one image scene **B, C** with and **A, D** without lid, respectively

## Background

A bottleneck of the promising discipline “phenomics”, which combines high-throughput phenotyping with genome and transcriptome analyses, is the automated acquisition of phenotypic data [1]. Applications of digital phenotyping range from monitoring individual plant cells in controlled environments to satellite-based remote sensing at the plant canopy level using various ground-based and mobile platforms such as gantries, agricultural vehicles, drones, and various sensor technologies such as LIDAR, RGB camera and spectral devices [1]. Although plant in vitro culture is the basis of most biotechnological methods for breeding and propagation of disease-free plants, very limited research using automated sensors in plant tissue culture has been reported, mainly using “plant to sensor” approaches [2–7] and thus involved a significant degree of invasiveness. So far, only few sensor technologies were used, including monochromatic imaging sensors [2], RGB cameras [3, 5–7], modified RGB camera setups with a near infrared (NIR) channel [4, 8] and thermal imaging sensors [9]. Therefore, most studies (reviewed by Gupta and Karmakar [5]) focused on image analysis to estimate parameters like biomass of callus [10], classification of somatic embryos and regenerated shoots [11, 12], as well as chlorophyll determination [13] and growth of embryogenic suspension cultures [14]. A fully automated image acquisition customized for in vitro cultured plantlets was demonstrated by Dhondt et al. [4]. The “in vitro growth imaging system” (IGIS) consisted of a rotating metal platform (carousel) to capture top-down images of *A. thaliana* rosettes cultivated in Petri dishes. However practical usage of the setup is limited in terms of scalability and it is not suited for

phenotyping of cultures of commercially important micropropagated species like *Phalaenopsis* spp., *Rubus* spp. and *Helleborus* spp. [15] due to their larger explant size and height.

Visual monitoring of the cultures is a costly and time-consuming repetitive task [8]—typically once a week in research and depending on the plant species every 2 to 10 weeks in commercial propagation—to assess the plant quality, the occurrence of contaminations, the outgrowth of endophytes, and morphophysiological disorders in research laboratories and commercial micropropagation laboratories. Quantitative assessments, such as biomass increase or multiplication rate, are up to now limited to single point measurements at the end of a subculture. Automation offers great potential for increasing efficiency of micropropagation laboratories since 60–70% of total costs of a micropropagated explant is due to manual labor [16]. According to Cardoso et al. [17], the high cost of labor for skilled workers is the most common reason for plant tissue laboratories to switch from manual to automated processes. However, the switch is currently often hindered by the high initial cost of automation, which increases the interest in low-cost monitoring systems for commercial use.

Due to the specific in vitro culture conditions in closed vessels, optical monitoring approaches face a number of challenges such as water condensation on the lid, opacity and total reflection of plastic lids or media surfaces (Fig. 1) [4, 6, 8]. Therefore, most plant evaluation methods were destructive and non-real-time methods, while digital phenotyping of in vitro plants allows objective and continuous quantification of plant characteristics over time. Important biological

key parameters for the performance of micropropagated plants include biomass, multiplication rate, shoot length, plant quality, and the absence of malformations, contaminations, and outgrowing endophytes.

Here, we describe the development of a low-cost phenotyping platform (named “Phenomenon”) suitable for direct monitoring of plant *in vitro* cultures while cultivation in an established multi-layered shelf system. In addition, the “Phenomenon” system is scalable for high-throughput use in commercial laboratories and capable of monitoring a wide range of plant species and various different *in vitro* culture techniques. In the present study, we aimed (i) to describe in detail the hard- and software components of the established phenotyping system, (ii) to validate the four sensor systems and (iii) to demonstrate the performance of the system for the quantification of growth parameters, such as projected plant area, average canopy and maximum plant height.

## Results

### Phenotyping system concept

The phenotyping system was designed as a scanning imaging system (xyz-gantry) for an autonomously operating acquisition of multi-sensor data, including RGB, thermal, depth and spectral data with specifically developed illumination (Fig. 2). Essential steps of continuous data acquisition with non-imaging and imaging sensor technologies were developed (Fig. 3). We could experimentally determine the technical repeatability for xy-axis with a  $MAE_x$  of 0.23 mm and a  $MAE_y$  of 0.08 mm of the repositioning over 16 days via RGB image analysis of a reference object (described in detail “Methods” section). For the z-axis, a technical repeatability with a  $MAE_z$  of 0.09 mm was obtained by using the calibrated laser distance sensor. For data segmentation a RGB image processing pipeline based on a random forest classifier and a depth image processing pipeline based on RANSAC [18] were newly established (Fig. 4).

### Optical properties of culture vessels

In order to ensure high quality data acquisition for the four sensors and their respective spectral working ranges, spectral transmittance measurements were conducted from the ultraviolet (UV) to the long wavelength infrared (LWIR) region of three possible culture vessels and lids (Fig. 5). The polypropylene lid and the polystyrene Petri dish represented the standard culture vessels, while the polyvinyl chloride foil was included as an alternative sealing.

While all of the three tested sealings had high transmittance (> 91%) in the visible spectrum (VIS) (Fig. 5A), the tested materials differed strongly in the proportion of transmitted diffuse light (Table 1). The ratio of both is

described by the Haze index, according to standard test method ASTM D1003 [19], thus representing an indicator for light scattering effects and visual perception by camera chips. Haze index should be kept to a minimum in imaging situations to maintain sharpness and clarity of the monitored object. The high Haze index of 34.2% excluded the standard polypropylene lid for being used in the phenotyping approach, while the polystyrene Petri dish and the PVC foil provided a clear VIS transmittance indicated by much lower Haze indices of 0.5% and 1.4%, respectively. In addition, a low to medium mean transmittance in the thermal range of 1.9% for the Petri dish and 50.6% for the polypropylene lid was determined (Fig. 5B). However, the foil still perceived a mean transmittance of 78.4% in thermal region. Thus the PVC foil was most suitable as a sealing system for imaging approaches for plant tissue culture, neglecting other not tested physical properties.

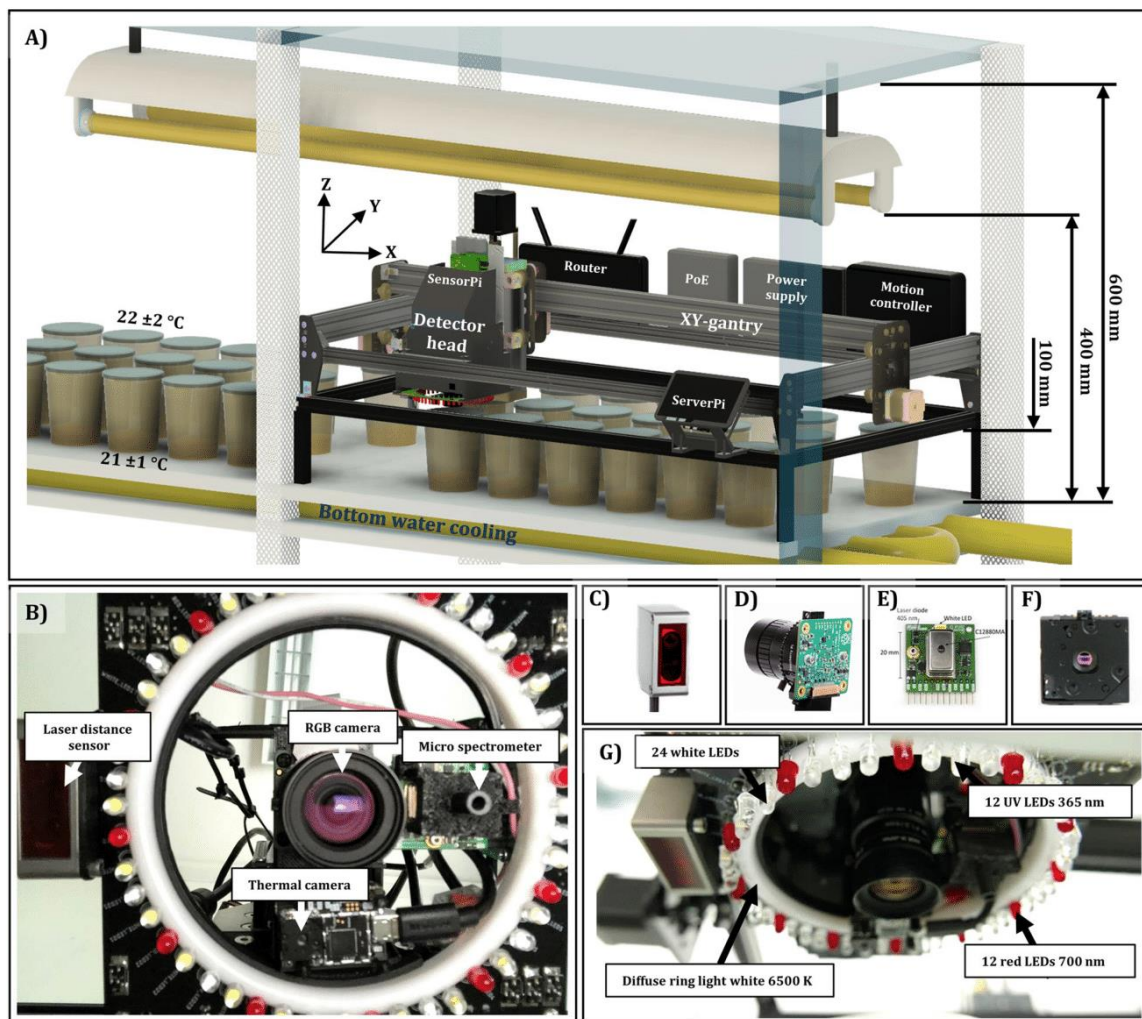
### Collection of representative phenotypic data of plant *in vitro* cultures

The following results derived from automated data acquisition by the phenotyping system “Phenomenon” according to Fig. 3, which included an automated sequential approach of culture vessel positions and acquisition of multi-sensory data over weeks. Exemplary data analysis were conducted by automated data processing pipelines presented in Fig. 4, where automated segmentation of RGB and depth data were performed.

### RGB data—Exemplary data analysis and validation of RGB image processing pipeline

Several *in vitro* phenotyping approaches were conducted with the “Phenomenon” system to demonstrate its full potential, including different plant species (*Arabidopsis thaliana*, *Nicotiana tabacum* and *Malus domestica*—data not shown) and developmental phases (*in vitro* germination, shoot and root regeneration and shoot multiplication). Figure 6 demonstrates the regeneration of adventitious shoots of *N. tabacum* from leaf explants monitored (6 images per day) over 32 days after treatment (DAT) and the output of the automated RGB processing pipeline of Fig. 4. This experiment clearly illustrated the segmentation challenges for image analysis such as similar color appearance of developing cell and organ types such as callus or roots and the cultivation medium, medium adhering to the plant cluster and camera specific changes in color balance. Additional file 1 contains a complete time-lapse video of one of the culture vessels. Regardless of the challenges mentioned, this video demonstrates the great potential of “Phenomenon” in terms of time series observations.



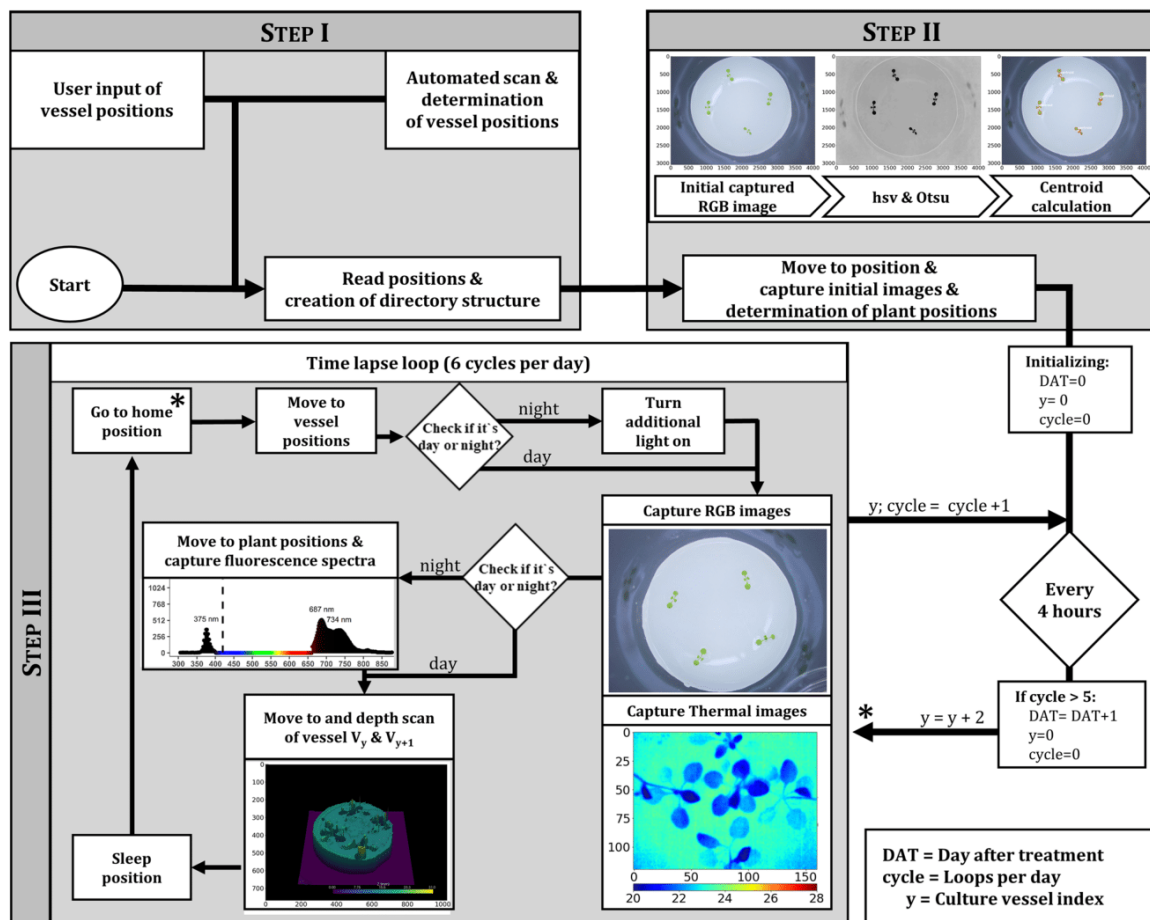


**Fig. 2** Experimental setup of the phenotyping system designed for direct monitoring of plantlets and explants cultured in vitro. **A** 3D representation of the designed robot platform inside a multi-layered shelf system with bottom water cooling. **B** Closeup of sensor arrangement of the developed multi-sensor detector head. Four different sensors, including **C** a laser distance sensor, **D** RGB camera, **E** a micro spectrometer and **F** a thermal camera defined the multi-sensor detector head. Furthermore, **G** a ring-light printed circuit board, including UV, white and red LEDs was added to a purchasable diffuse ring light to meet the highly specific illumination situation of monitoring plant in vitro cultures. Detailed description in “Methods” section

As second demonstration of the functions of our phenotyping system, Fig. 7 shows the whole life-cycle monitoring of *A. thaliana* in vitro (seedling to flowering plant) and the calculation of growth performance metrics. Time-lapse videos of *A. thaliana* monitored over 16 days are provided in Additional files 2, 3.

The validation of the projected plant area obtained as one output from the RGB image processing pipeline by relating it to the projected plant area determined by manual annotation of plant pixels (ground truth) indicated a

high  $R^2$  of  $>0.99$  (Fig. 8A). The automated classification approach overestimated the plant area with an average error of 7591 px. The relative error of the different acquisition time points (Fig. 8B) indicated a slightly higher overestimation at day time images, while an underestimation occurred for night time images (with highest error at 23 o'clock), resulting in a mean relative error (MRE) of 0.37% overestimation. To quantify the classification performance, confusions statistics of 221,834,880 pixel pairs were conducted and disclosed a classification accuracy of



**Fig. 3** Flow chart of the three main steps of the automated phenotypic data acquisition (indicated in gray). In step I, the position of the culture vessel is determined, while in step II the initial images and the calculated plant positions are acquired. To determine plant positions, the original image was transformed to hsv-colorspace and the h-channel was segmented with Otsu-Method [20]. Four largest objects were selected as plant positions. Step III includes the actual time-lapse loop (start indicated by asterisk), where data of the four sensors are recorded. Detailed description in "Methods" section

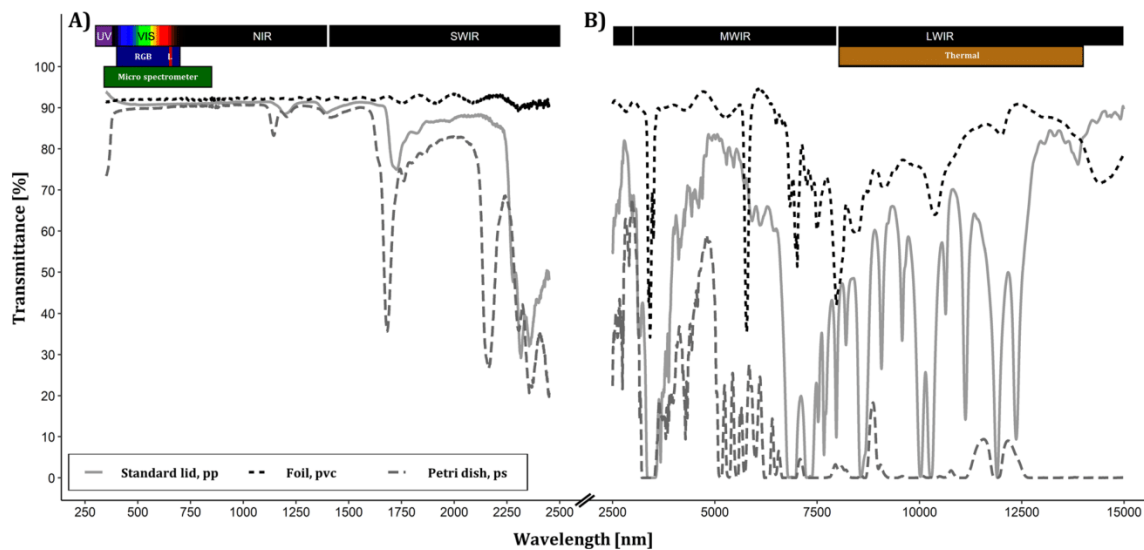
97.7%, a sensitivity of 97.7%, a specificity of 96.9% and a precision of 99.9% for the segmentation of the proposed RGB image processing pipeline.

**Depth data—Exemplary data analysis and validation of depth image processing pipeline**

The first report of depth data acquisition and analysis in plant in vitro culture is illustrated in Fig. 9A using

(See figure on next page.)

**Fig. 4** Overview of main data processing steps and used software packages to process the different types of acquired data. **A** A trainable Ilastik [21] classification model was trained to robustly cover the diversity of background (yellow labels) due to changing background and media color and diversity of foreground (blue labels) such as different plant species appearance and explant color changes during cultivation. **B** RGB image processing pipeline was developed in Python [22] with OpenCv [23] and PlantCv [24] for batch processing and including the ilastik classification model headless for segmentation. Upper row: RGB image processing workflow included an automated brightness and contrast adjustment by histogram stretching, down-scaling of image resolution from 4054 px x 3040 px to 1014 px x 760 px. Lower row: The trained classifier predicted binary mask of plant pixels rescaled to the original image resolution and applied to the original image for background removal. Exemplary images from monitoring of *A. thaliana* (Trial A). **C** For depth data processing, Python with Open3D [25] was used as an essential component to perform RANSAC [18]-based segmentation. Depth data of in vitro grown *A. thaliana* seedlings (Trial A). Upper row: Day 0 (Media with 10 day old, small seedlings), Hough Transform circle detection [26] and edge-removed depth image. Lower row: Pseudo 3D visualization of depth data of Day 11, estimated RANSAC plane and plant height surface corrected by estimated RANSAC plane at Day 11. Detailed description in "Methods" section



**Fig. 5** Spectral transmittance of culture vessel closures. Transmittance was measured with **A** an UV/VIS/NIR spectrometer and **B** with a FT-IR Spectrometer. Three independent replicates were measured and mean spectra per lid material are shown. Colored rectangles indicate waveband regions and spectral sensitivity of the sensors (according to the manufacturer’s specifications) installed in the phenotyping platform (blue, RGB camera with Sony IMX 477 sensitivity: 400 to 700 nm; green, micro spectrometer sensitivity: 340 to 850 nm; red, L, Laser distance sensor emission wavelength: 655 nm; brown, thermal camera sensitivity: 8000 to 14,000 nm)

**Table 1** Optical characteristics of in vitro culture vessel sealings

Culture vessel sealing	Replicates	Total transmittance [%]		Diffuse transmittance [%]		Haze index [%]		Thick-ness [µm]
		Average	SD	Average	SD	Average	SD	
Standard lid, PP	3	91.3	0.3	3.3	0.9	34.2	0.9	200
Foil, PVC	3	92.6	0.3	1.4	0.2	1.4	0.2	20
Petri dish, PS	1	91.2		0.6		0.5		900

PP polypropylene, PVC polyvinyl chloride, PS polystyrene

Measured with UV/VIS/NIR Spectrometer (PerkinElmer Lambda 1050) in 5 nm intervals from 380 to 780 nm according to standard test method ASTM D1003[19]

Haze index [%] = ((Diffuse transmittance/total transmittance) – rel. scattered transmittance by the system) × 100

*A. thaliana* as an example. It included the calculation of important biological parameters from the depth data set to monitor culture medium height, culture medium volume, mean canopy height, maximum plant height and degree of coverage (Fig. 9B). Figure 9 clearly demonstrates a height and volume reduction of culture media, while in plant growth related parameters a height increase was notable. Corresponding RGB images revealed first signs of flower induction of the *A. thaliana* seedlings at DAT 12, and an associated increase in maximum plant height was to be seen in depth data at DAT 16. In addition, also the variance in average canopy height increased at DAT 16.

Comparing the projected plant area obtained from the RGB processing pipeline (assumed as the ground truth) with the projected plant area as output from the depth processing pipeline, a high correlation, expressed in an  $R^2$  of 0.93 was observed with an average underestimation of 59.7 mm<sup>2</sup> (MAE) of plant area determined by depth data (Fig. 10). The mean relative error (MRE) revealed that the depth data processing pipeline projected the plant area by 65% compared to RGB processing pipeline, meaning 35% of plant pixels were systematically not detected by the sensor or have been removed due to segmentation. This can be considered as a rough estimator of how accurately the different sensor technologies (RGB camera vs. scanning laser distance sensor) detect plant pixels,

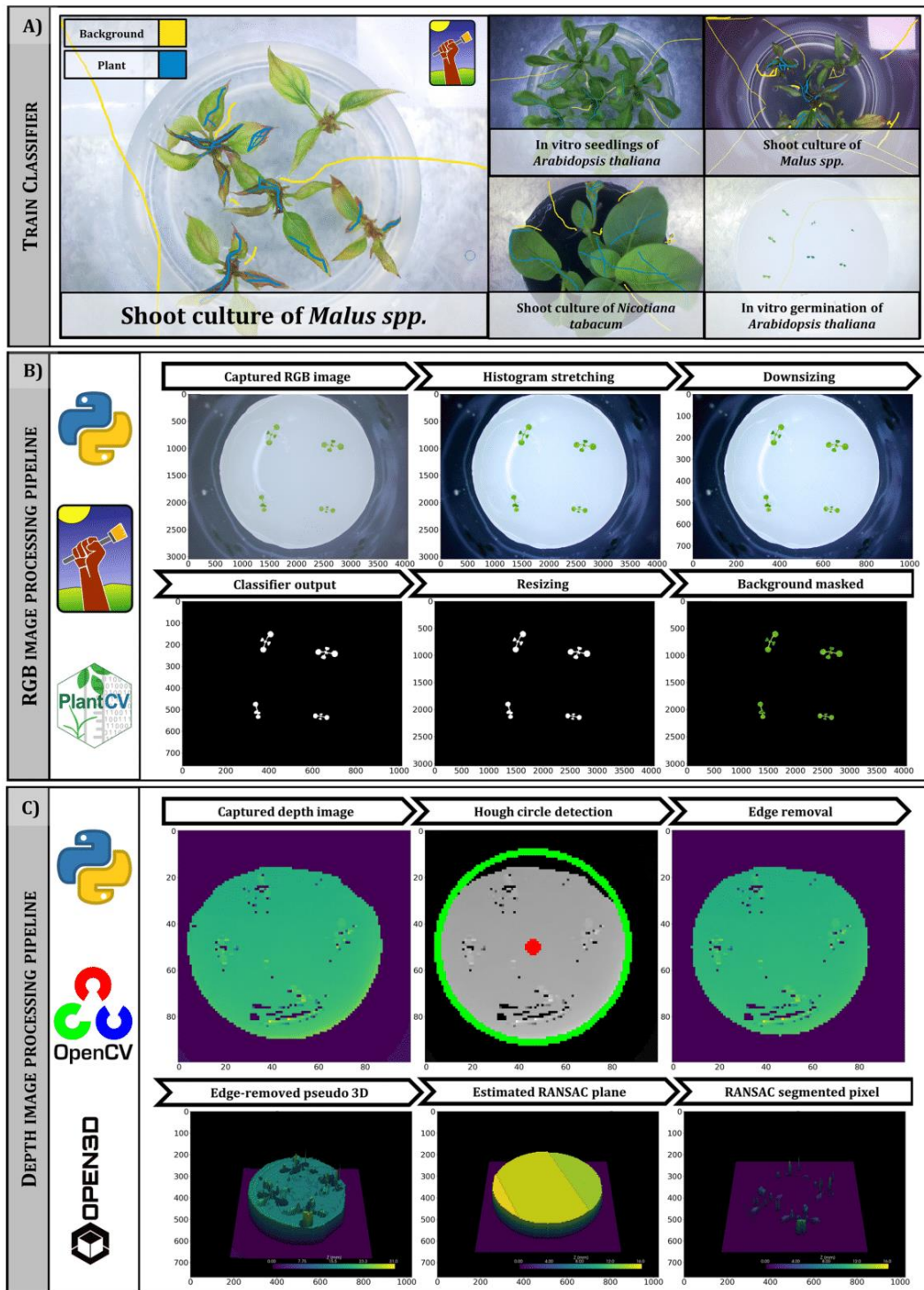
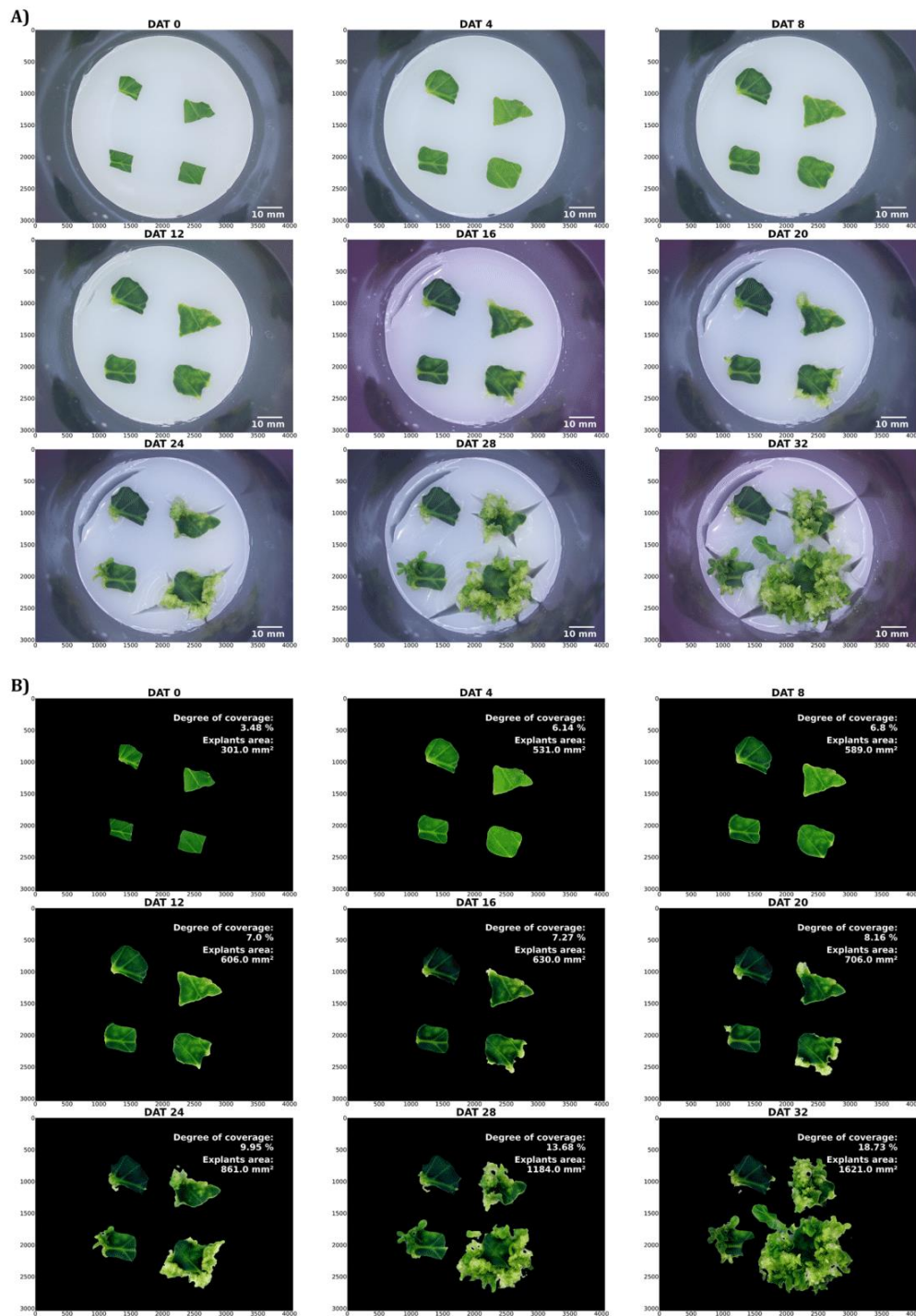
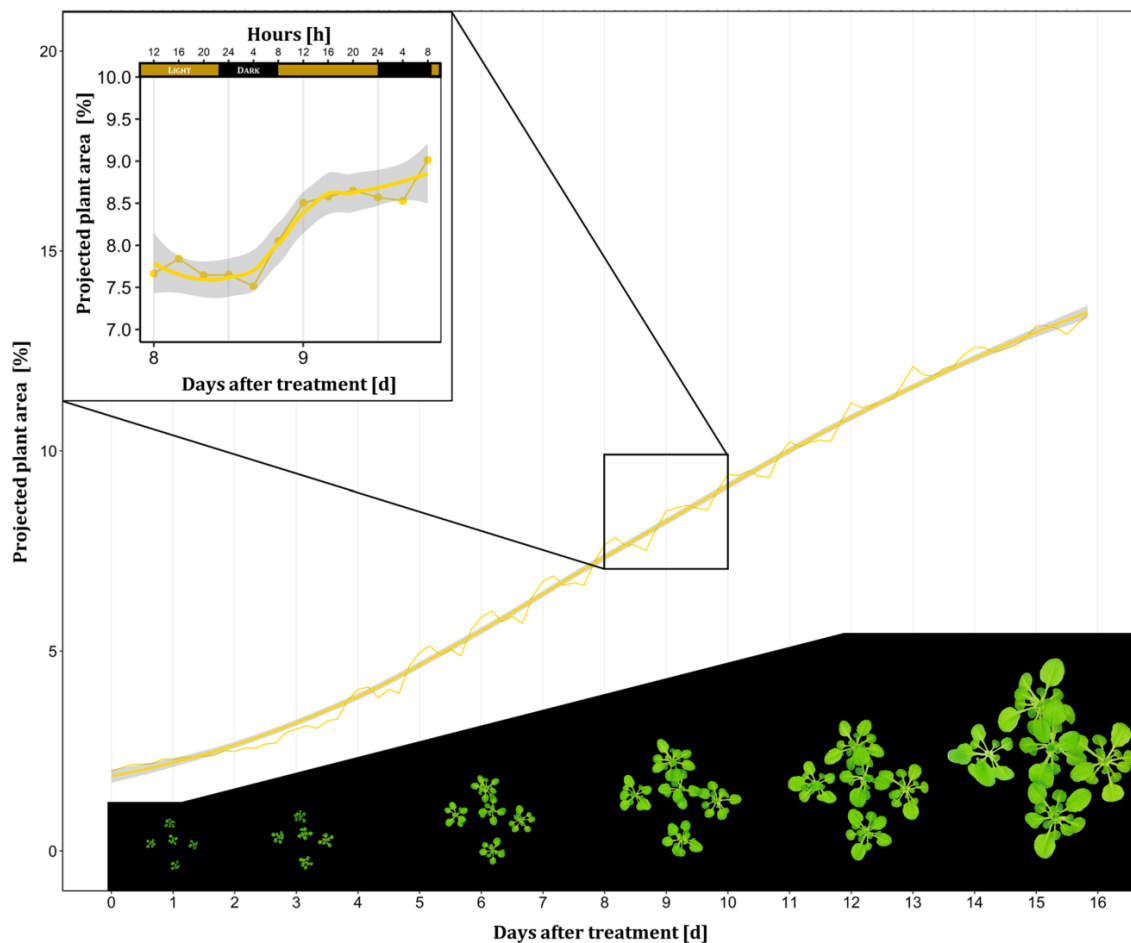


Fig. 4 (See legend on previous page.)



**Fig. 6** Monitoring of shoot regeneration of *N. tabacum* leaf explants. **A** RGB raw images and **B** processed images with the RGB imaging processing pipeline. *N. tabacum* leaf explants were placed on regeneration medium and developing adventitious shoot clusters were recorded over 32 days after treatment (DAT). Degree of coverage was calculated as the sum of plant pixels divided by total number of pixels within an image. For determination of explants area, sum of plant pixels was multiplied by pixel-metric-conversion factor. Time lapse video of *N. tabacum* regeneration is provided in Additional file 1



**Fig. 7** Exemplary growth curves of one culture vessel containing five *A. thaliana* seedlings (Trial B) expressed as projected plant area. Projected plant area was calculated as the sum of plant pixels divided by total number of pixels within an image. Yellow smoothed line plot, method=logistic regression, gray indicates confidence interval borders  $\alpha=0.95$ . Six images per day over 16 days resulted in a total number of 96 images per growth curve. Left corner highlights a closeup showing the diurnal rhythm of plant growth. The bottom part contains segmented images of 0, 3, 6, 9, 12, 15 DAT (days after treatment). Time-lapse video of *A. thaliana* (Trial B) is provided in Additional files 2, 3

including the errors derived from segmentation and differences in object area representation by the two sensor technologies.

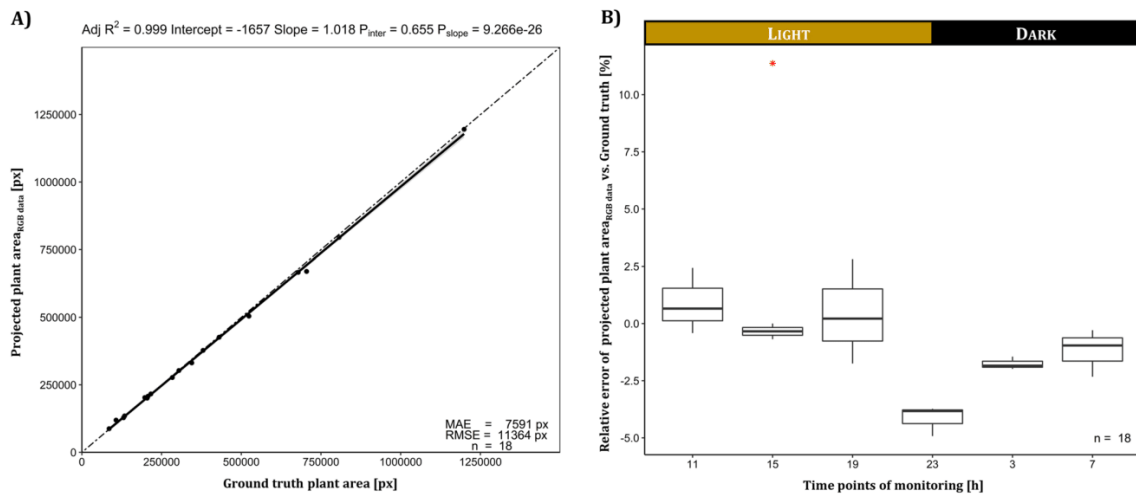
**Spectral data—Exemplary data analysis and validation of detection spot size**

An automated and dynamic monitoring of the chlorophyll fluorescence signature of an *A. thaliana* seedling cultured in vitro over 21 days is illustrated in Fig. 11. Excitation light emission maximum at 375 nm as well as a sequential increase in the fluorescence signal depending on the plant growth were evident. Typical emission maxima derived from the reaction centers of the photosystem

(PS): mainly PSII (F690) and PSII and PSI (F730) were detected. Furthermore, we have determined the diameter of the detection spot of the modified spectrometer to be 23 mm (detailed description in “Methods” section).

**Thermal data—Exemplary data analysis and validation**

Thermal imaging of in vitro cultivated *A. thaliana* seedlings was attempted, but faced the challenges of the special imaging situation (Fig. 12). When captured without



**Fig. 8** Characterization of the segmentation performance of the RGB image processing pipeline from 18 randomly selected images of the *A. thaliana* Trial A dataset. **A** Linear regression of projected plant area vs. ground truth plant area. The regression line is colored black, while the angle bisector line is drawn two-dashed. Gray indicates confidence interval limits at  $\alpha = 0.95$ . Adj  $R^2$  denotes the coefficient of determination adjusted according to Yin and Fan [27], while  $P_{slope}$  and  $P_{inter}$  represent p-values of the coefficients for the intercept and slope determined by simple T-test. MAE and RMSE indicate the mean absolute error and the root mean square error of the projected plant area. **B** Relative error of plant area projection for different acquisition time points (23, 3 and 7 o'clock represented night conditions) calculated from 18 randomly selected images of the *A. thaliana* Trial A dataset. First manual annotated image was identified as an outlier marked as red asterisk

the sealing foil, the thermal images (Fig. 12A) of culture vessels at a room temperature of 25 °C (day) and placed on a bottom-cooled shelf surface with a temperature of 21.5 °C, revealed reasonable absolute values of thermal data, as indicated in the corresponding histogram. However, thermal images taken through the sealing foil (Fig. 12B) only allowed a weak separation between plant and background pixels and the corresponding histogram indicated an increase in radiometric data.

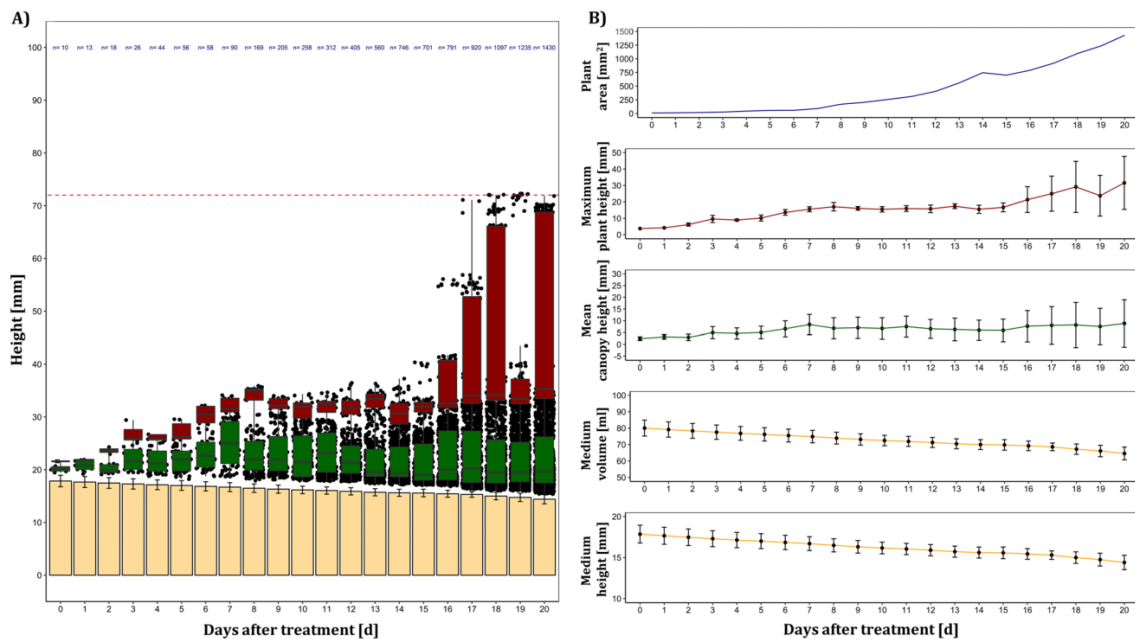
### Discussion

To the best of our knowledge, this is the first report of a multi sensor phenotyping system, based on an xyz-gantry that is capable of autonomous acquisition of relevant sensor data of plant in vitro cultures. The selection of exclusively low-cost hardware (Table 2) and open-source software components accessible via the GitHub repository [28] enables other researchers to rebuild the “Phenomenon” system and to benefit from it in science, education and commercial micropropagation.

As proposed by Dhondt et al. [29] phenotyping systems can be defined by system properties like throughput, resolution and dimensionality. With the current setup, we reached a throughput of multi-dimensional data (RGB, depth, spectral, thermal) at a macroscopic resolution for ten culture vessels per day. Therein, the main limiting factor was the time-consuming process of depth data scans (45 min per vessel; compared to RGB and thermal

image and spectral point measurements with only a few seconds per vessel) and system dimensions restricting the working area. Low cost imaging depth sensors based on “time-of-flight” principle (ToF) such as Pieye Nimbus 3D or Onion tau could reduce substantially the acquisition time of depth images. However respective sensors need to be tested how they perform under the highly challenging imaging conditions (Fig. 1) of plant in vitro cultures. Nevertheless, a large-scale application can be achieved with minimal effort and costs if the robot system working area is scaled up to a whole shelf.

We aimed at monitoring plant in vitro cultures with minimal invasiveness, consequently phenotyping took place dynamically within in the cultivation of in vitro cultures, instead of monitoring open culture containers under laminar flow to ensure aseptic conditions. Non-destructive phenotyping approaches where optical sensing happens through the vessel encounter a challenging imaging situation (Fig. 1) and could be solved in parts by the technical design of “Phenomenon”. However, three modifications were necessary to increase sensor data quality: (i) The culture vessels were placed on a bottom-cooled surface to avoid condense water formation. Bottom cooling systems are widely applied in tissue culture to reduce the relative humidity in the vessels and thereby increasing plant quality, but in case of rose roots also slowed down the growth of cultures due to the lower



**Fig. 9** Exemplary depth data analysis of a culture vessel with four *A. thaliana* seedlings grown in vitro for 21 days (Trial A). **A** Yellow bar plot displays calculated media height (Mean  $\pm$  SD), while black dots indicate sensor values of plant pixels after segmentation and colored boxplots indicate values for the calculation of the two digital parameters mean canopy height (green) and maximum plant height as mean of the upper 10 percentile (red). Red dashed line represents the maximum height of the sensor reliable distance ( $\leq 72$  mm) and the amount of plant pixels after segmentation was colorized in blue. **B** Individual plots of the dynamic behavior of the calculated digital parameters with equal color code and depicted as means  $\pm$  SD. The calculation of medium volume and all other parameters is described in “Methods” section. 10,000 data points for each date were processed from depth scans of an area of 100 mm  $\times$  100 mm with a scan pattern of 1 mm  $\times$  1 mm. Depending on the necessary segmentation for the calculation of the individual parameters, a corresponding proportion of the 10,000 data points was included in the analysis

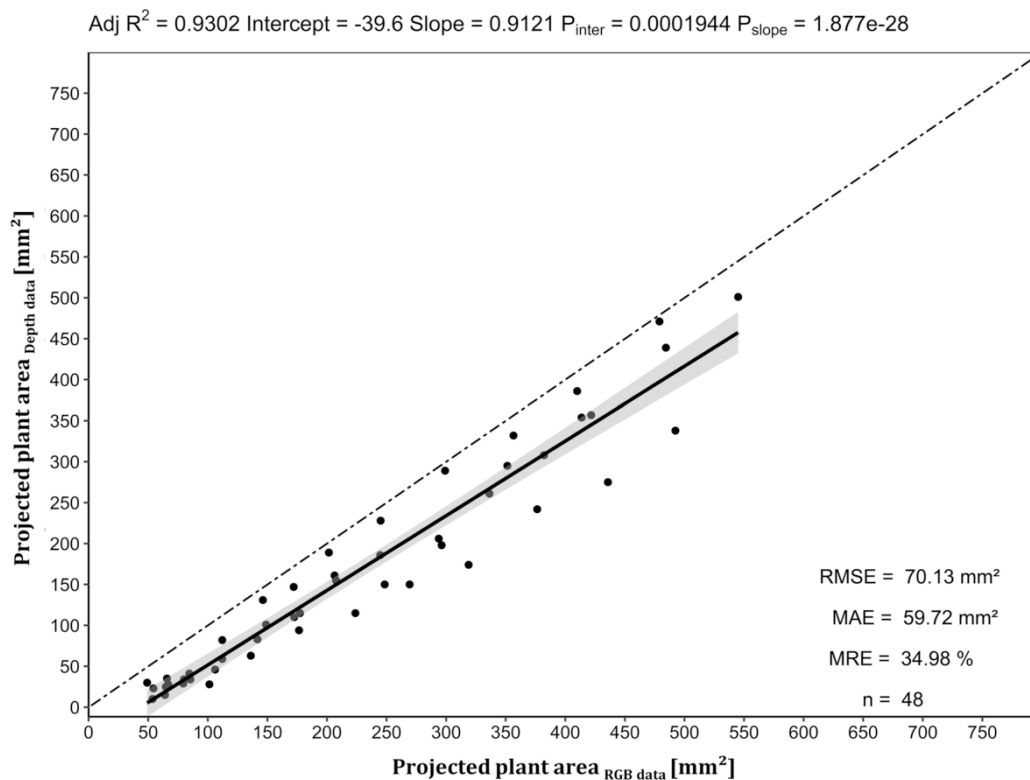
temperature [30]. (ii) The plastic lid was substituted by a PVC foil to maintain a clear and undistorted view, evidenced by the Haze index (Table 1) and to increase the spectral transmittance in the thermal region (Fig. 5). This also affected the gaseous exchange capacity of the culture containers, which was notable by increased evaporation of water from the culture media. The use of the foil also prevented condense water formation, when the bottom was not cooled. Thus, this intervention might be sufficient. Nevertheless, future research should address an optimization of the culture vessels/lids to enable proper imaging sensor application in vitro. (iii) The supplementation of the culture media with TiO<sub>2</sub> allowed a detection of the surface of the normally semitransparent media with the laser distance sensor (detailed description in “Methods” section). TiO<sub>2</sub> had already been used in plant in vitro culture due to an antimicrobial activity induced by UV excitation [31]. However, beneficial or cytotoxic effects of TiO<sub>2</sub> nanoparticles (NPs) in particular, are currently under research and most likely will be depending on the dose and UV exposure time [32]. TiO<sub>2</sub> NPs had

no negative effect on the growth of soybean seedlings in vitro at concentrations of 10 and 100 mg L<sup>-1</sup> TiO<sub>2</sub> NPs, but slightly reduced fresh mass and root growth at 1000 mg L<sup>-1</sup> TiO<sub>2</sub> NPs [33], suggesting a reduction of the TiO<sub>2</sub> concentration in the culture media for upcoming experiments.

The **automated scanning imaging system “Phenomenon”** based on a belt-driven xy-gantry and screw-driven z-axis was specified by the manufacturer to provide an accuracy of 0.1 to 0.2 mm for the xy-axes and 0.05 to 0.1 mm for the z-axis. Experimentally, we determined the technical repeatability for MAE<sub>X</sub> of 0.23 mm, MAE<sub>Y</sub> of 0.08 mm and MAE<sub>Z</sub> of 0.09 mm. Considering the fact that an exclusively low-cost phenotyping system was intended, a sufficient technical repeatability was achieved for consistent data acquisition.

RGB data acquisition was conducted with a low-cost **RGB sensor** equipped with a low distortion lens to minimize the error of projection. This error resulted in a distortion of the projected plant area at the edges of the image compared to the midpoint. Furthermore, the





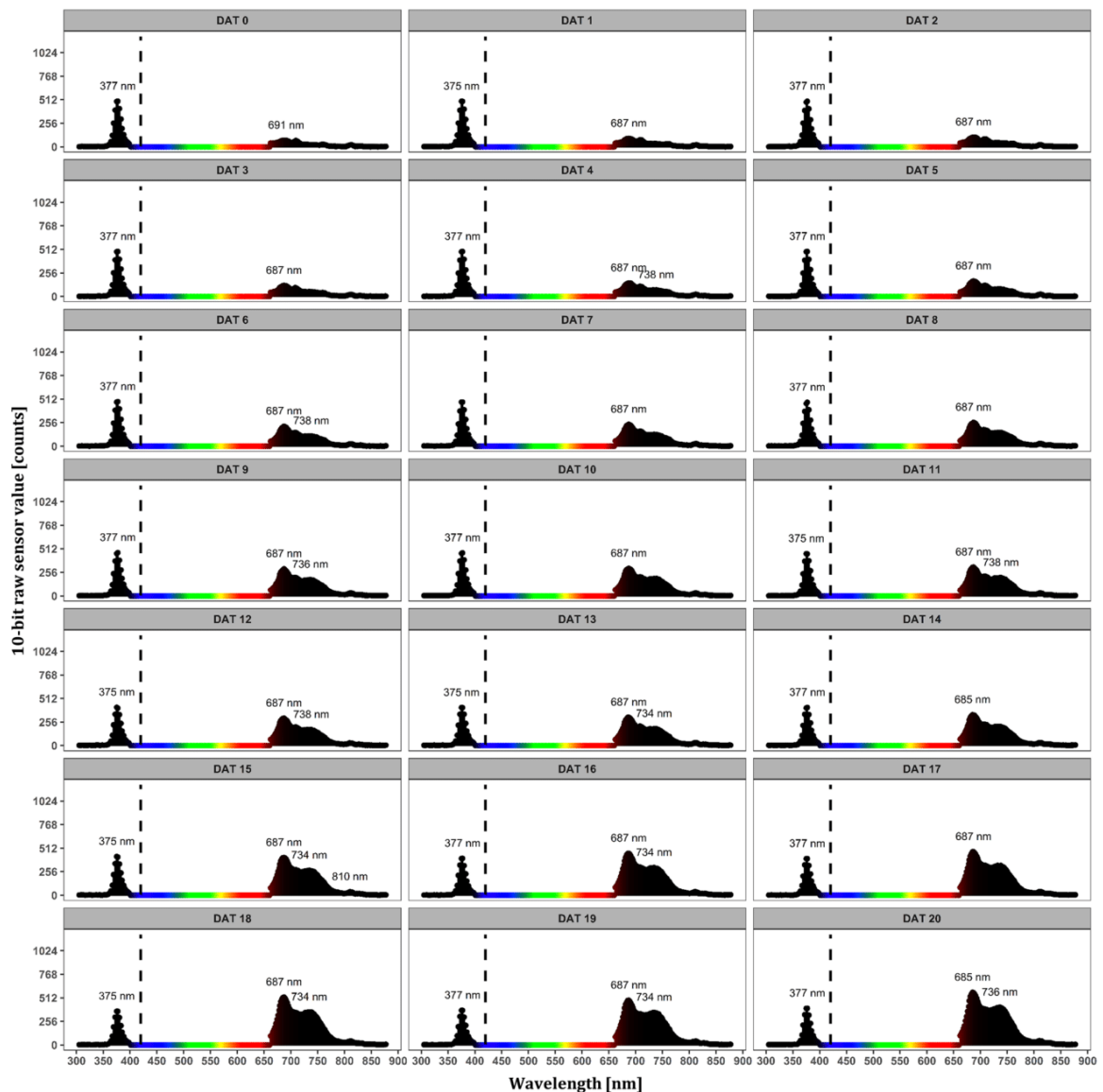
**Fig. 10** Linear regression of projected plant area as output of the depth data pipeline vs. projected plant area obtained from the RGB data pipeline (ground truth). The regression line is colored black, while the angle bisector line is drawn two-dashed. Gray indicates confidence interval limits at  $\alpha=0.95$ . Adj  $R^2$  denotes the coefficient of determination adjusted according to Yin and Fan [27], while  $P_{slope}$  and  $P_{inter}$  represent p-values of the coefficients for the intercept and slope determined by simple T-test. MAE, MRE and RMSE indicate the mean absolute error, mean relative error and the root mean square error of projected plant area. Sampling (n) was formed out of 12 images from four different culture containers and 12 time points respectively (DAT 0–DAT 11)

estimation of plant area, i.e. of a 3D object, with a 2D sensor without a telecentric lens can be put into question. However, plant cultivation in multi-shelf systems (Fig. 2) with a distance between the cultivation area and the illumination of 400 mm, limited not only the application of optics greater in size but also the selection of other sensor technology by their optical specifications (minimum working distance;  $MWD < 150$  mm).

In this study, we demonstrated a successful implementation of a scanning laser distance sensor resulting in a **depth image** of plant in vitro cultures for the first time. Novel relevant traits of micropropagated cultures like medium height and deduced from this medium volume, average canopy height and maximum plant height could be quantified and will be validated in upcoming experiments. We showed a reliable application of this technology (Figs. 9, 10, Additional files 4, 5 and 6), but the reflection-based time-of-flight sensor failed, if the reflection surface was tilt with a higher angle (upwards

growing leaves) and at the upper part of depth images, where the emission beam was inside and the detector side of the sensor still outside of the culture vessel (Fig. 4: missing part of detected Hough circle). In addition, it is worth mentioning that the detection error of the reflection-based sensor could be due to the fact that the emission wavelength of the laser distance sensor hits the absorption of the plant pigments at 655 nm. Therefore, depth sensors with spectral detection range in near infrared might be superior due to the higher reflectance signal derived from the red edge shape of the plant spectra.

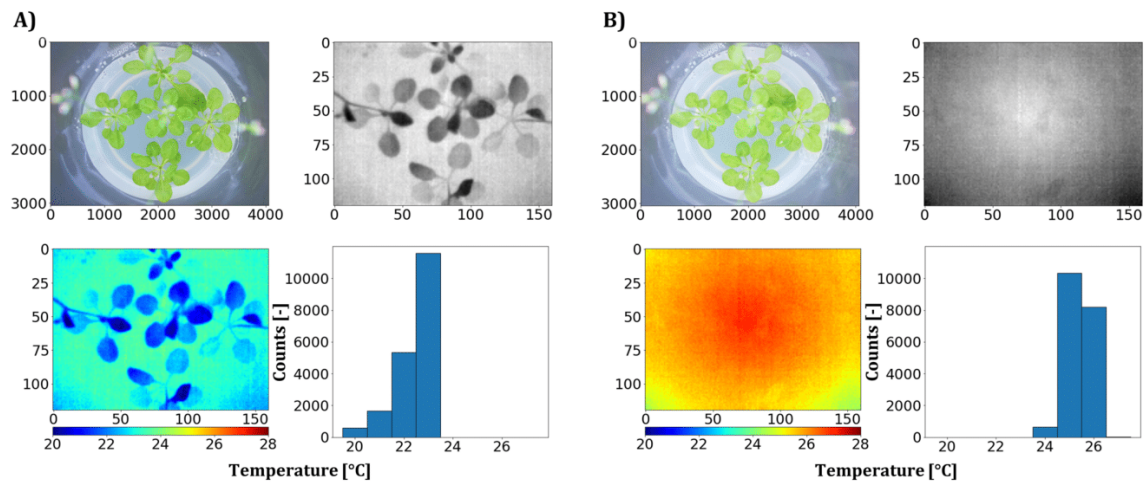
The second novelty was the proof-of-concept for applying a low-cost **micro spectrometer** to determine spectral signatures, offering great potential for monitoring the stress status of in vitro cultivated plantlets. The point measuring device was limited in spatial resolution due to the detection spot size of around 23 mm (Additional file 7). Reflection-based measurements were therefore not exclusively-plant-specific. However, the fluorescence



**Fig. 11** Exemplary determination of dynamic chlorophyll fluorescence monitoring of one of four *A. thaliana* seedlings grown in vitro for 21 days (Trial A). The first peak at 375 nm can be assigned to the excitation light provided by UV LEDs imperfectly blocked by the long pass filter at 420 nm (black long dashed line). The region from 400 to 660 nm has been masked for simplified representation. Emission peaks in the region from 660 to 780 nm indicated the two typical maxima of the chlorophyll fluorescence, derived from PSII (F690) and PSII and PSI (F730). Micro spectrometer integration time was set to 300 ms

signature reflected plant specific peaks (Fig. 11). Known stress indices, like F690/F740 as a chlorophyll content estimator [34, 35]—can be calculated from the fluorescence spectra on an explant base and their potential use to detect early stress responses opens new ways in in vitro stress screenings, for example.

Leaf temperature quantification of micropropagated plants by **thermal imaging** approach was already investigated by Ibaraki and Gupta [9], but so far only after their transfer to ex vitro conditions. Thermal imaging of plants is widely used to estimate evapotranspiration-based parameters like water loss, water stress indices or stomatal conductance [9, 36, 37]. We could disprove the



**Fig. 12** Exemplary thermal imaging of in vitro cultivated *A. thaliana* seedlings (Trial A). **A** Left side demonstrates imaging without the foil that was used to seal the culture vessels, while **B** the right side shows acquired sensor data through the foil. The respective RGB images are shown and thermal data are presented as grayscale and false color images and corresponding histograms

**Table 2** Main system components and costs

Description	Quantity	Hardware	Price
G420 Long pass filter	1	Dielectric coated long pass filter	40 €
PCB manufacturing cost	1	Ring light PCB and a Circuit PCB	40 €
Various LEDs	48	Standard 5 mm LEDs (375 nm, 6500 K, 700 nm)	40 €
RGB camera	1	Raspberry Pi Camera High quality	50 €
3D filament, cable chain, limit switch	1	Small mechatronic parts	100 €
Network communication	1	Router & PoE-Switch	100 €
Single-board computer	2	Raspberry Pi 4B & PoE-Shield	120 €
Z-axis with Nema 23 Stepper motor	1	OpenBuilds Linear Actuator	160 €
Thermal camera	1	PureThermal 2 & Lepton 3.5	250 €
Micro spectrometer	1	Mini-Spectrometer C12880MA	350 €
Low distortion lens	1	Edmund Optics 6 mm lens	400 €
Xy-gantry with 24 V power supply and 3xNema 17 Stepper motor	1	OpenBuilds ACRO 1515 60" x 60"	410 €
Laser distance sensor	1	OD-Mini OB1-B100	1000 €
Total			3060 €

assumption that thermal imaging of in vitro cultures is impossible, even if data quality was limited in terms of contrast (Fig. 12). By using the PVC foil, we improved the average transmittance in the thermal waveband up to 78.4% (Fig. 5), but still absorption and reflection occurred and reduced the quality of the sensor data (Fig. 12). The increased mean temperature of explants imaged through the foil might be due to sensor self-reflection compared to the imaging without foil. Whether temperature differences between plants due to evapotranspiration can be quantified by thermal imaging of high humidity culture

vessels (93 to 97% RH with bottom cooling [38]) remains to be answered.

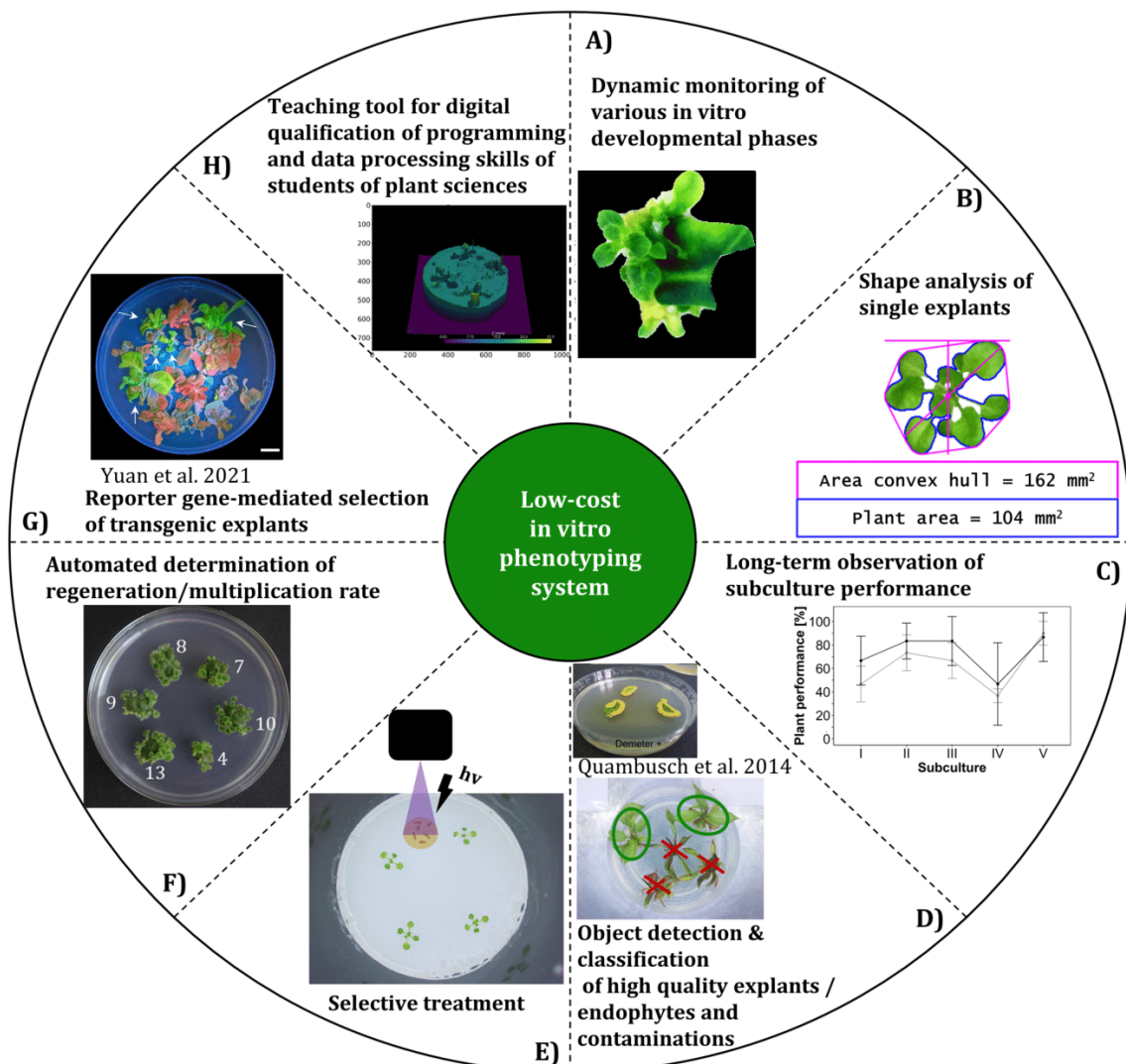
The validation of the **RGB image processing pipeline** demonstrated the power of digital image analysis accomplished through successful segmentation. Figure 7 could demonstrate the potential for researchers to compare treatments, such as different media compositions, or to track small leaf movement like the diurnal growth rhythm. A robust and specific segmentation covering the required range of the imaging situation was only possible by a trainable segmentation model. Despite the

acceptance of evoking errors by the use of reduced resolution images as an input of the segmentation model, a nearly perfect segmentation was achieved as indicated by the high coefficient of determination of  $R^2$  of  $>0.99$  referred to manual annotation of plant pixels. Confusions statistics revealed an even higher accuracy of 97.7% of classification compared to the study of Mestre et al. [8] reaching an accuracy of 96.9% although they used multidimensional data (RGB, NIR) as input for a random forest classifier to segment in vitro grown *Nandina domestica* explants. Main classification errors originated from overestimation of leaf borders by the automated RGB image processing pipeline compared to the ground truth segmentation and from false-positive classification due to root greening. The time point-dependent performance of the segmentation during the day (Fig. 8B) can likely be attributed to insufficient illumination of tiny plant structures such as leaf petioles, where the average light intensity captured by the camera revealed minimum residual light at the time point of the greatest underestimation (23 o'clock). Interestingly, a correlation coefficient of 0.75 indicated a strong correlation between the absolute classification error and the average mean intensity of the RGB images, explaining the difference at night time points where variations in residual light intensity occurred due to the switch timing of the tube fluorescent lamps (data not shown). Projected plant area can be used as good estimator for biomass as a key performance parameter of plant in vitro cultures as shown by Faragó et al. [6] who identified a coefficient of determination of  $R^2=0.99$  between *A. thaliana* digital rosette size and fresh mass. As a common issue of image analysis, it might be questionable whether the probabilistic based random forest model or the human labeled classification better reflect the real ground truth of in particular imperfect-focused leaf borders. It has to be stressed, that for other growth habits, such as upright growth with several layers of overlapping leaves, it will be more difficult to correlate projected plant area and biomass, but in these cases additional information from depth data may be used to define additional covariates.

The **depth image processing pipeline** showed that segmentation of plant in vitro depth data over time requires a dynamic approach to accommodate changing processes like plant growth or culture media shrinkage via evaporation. The separation of background, culture medium and plant pixels was the main objective of the segmentation for calculating relative plant height, culture medium height, and accounting for tilts of the cultivation surface or the medium surface. An image registration approach of RGB (where a good segmentation was already achieved) and depth images was not satisfying due to the too different representation of objects by the

two sensor technologies. A Random sample consensus (RANSAC [18]) algorithm fulfilled the requirements of the task and was able to dynamically and robustly detect the culture medium surface planes within the one-dimensional and therefore difficult to segment data. RANSAC is a robust method for an iterative determination of outliers from a mathematical model in an over-determined data set. The RANSAC approach is commonly used in depth data segmentation of plants [39–41] and allowed the determination of relevant and novel digital features of plant in vitro cultures like culture medium height, mean canopy height, maximum plant height and plant area by depth data. However, limitations will arise when the culture medium surface is fully covered by plants and therefore, no longer represents the largest plane. The determination of the height of each explant inside the culture vessel is to be aimed, but requires connected compounds after segmentation. To estimate the quality of the representation of plants in depth images we selected the projected plant area as a basis of comparison between RGB and depth sensor data and corresponding pipelines, respectively. The high mean relative error (Fig. 10) demonstrated the limitations of the scanning laser distance sensor, as only two thirds of the projected plant area were represented in the depth data after segmentation. This error could be attributed to detection errors, unfavorable reflection due to inclined surfaces (e.g., leaves growing upwards) or water drops on plants, the low spatial resolution of the scan pattern (1 mm×1 mm) or errors caused by segmentation. Nevertheless, depth data of plant in vitro cultures could be used to estimate plant biomass, especially when combined with projected plant area by RGB images. Furthermore, the determination of culture medium volume opens the possibility to collect new data of plant water uptake and evaporation from culture medium.

We have designed, constructed and tested a novel multi-sensor robot platform for phenotyping in plant in vitro cultures offering great potential for automatization of specific tasks in commercial micropropagation, but also offering new possibilities in research (Fig. 13A–C). The “Phenomenon” phenotyping system differentiates from existing in vitro monitoring approaches that focused primarily on shape analysis and the application of which was limited to *A. thaliana*, mainly. The system allows phenotyping of different species and different developmental phases in in vitro culture due to its customized and specific hardware design (Fig. 2). Repeated monitoring of individual cultures regarding their growth performance over several culture passages will reveal new insights into phenomena such as the habituation against phytohormones or seasonal variation of growth. Tracing back the development of individual



**Fig. 13** Potential applications of the automated low-cost phenotyping system in plant vitro culture. While the requirements for the use cases from **A–C, E** and **H** have already been met, further research is required for the use cases **D, F** to **G** (External images from Yuan et al. [42] and Quambusch et al. [44])

explants over time pave the way for improvements in cultivation. Furthermore, the system could be used to optimize culture medium compositions like amount of plant growths regulators via an objective quantification of the plant phenotypic characteristics. A future perspective of sensor application in plant in vitro culture by automated imaging robots includes the construction of multi-sensor data sets to benefit from ever easier access to the power of artificial intelligence such as artificial neural networks as reviewed in Prasad and Gupta [12] for complex classification or regression tasks such

as the detection of endophytes or the calculation of multiplication rates of plant in vitro cultures (Fig. 13D and F). Furthermore, robots may offer the ability to identify and treat explants which exhibit a low or high stress level after certain treatments or to early detect potential contaminations (Fig. 13E). A marker gene-free early selection of transgenic plant material as proposed by Yuan et al. [42] with the usage of new reporter genes such as eYGFpuv could be automated by the presented low-cost phenotyping system (Fig. 13G). Finally, we suggest using the system in teaching to promote

digital skills of plant science students. Since it features low-cost, stand-alone and portable characteristics, it may provide students with handling and processing of multi-sensory phenotypic data (Fig. 13H) [43].

### Conclusions

We developed a novel low-cost multi-sensor automated phenotyping system for application in plant in vitro cultures. The unique hard- and software concept is characterized by using exclusively low-cost compounds and open-source-based software components. This allows remote and programming language-independent access to its functionalities, enabling plant scientists to benefit from the capabilities with minimal financial investment. Various sensor technologies were applied for the first time under these challenging culture conditions and were evaluated with respect to resulting data quality and feasibility with proposed data processing pipelines. We demonstrated the digital determination of relevant parameters such as projected plant area, average canopy height, and maximum plant height, which can be used as critical descriptors of plant growth performance in vitro. The initial exemplary demonstration of resulting data promises great potential. The technical realization of “Phenomenon” enabled phenotyping of plant in vitro cultures under highly challenging conditions and will lead to increased sensor application approaches for research and commercial propagation in upcoming years.

### Methods

#### Adventitious shoot regeneration from *N. tabacum* leaf explants

From in vitro grown *Nicotiana tabacum* ‘Samsun’ shoot stock cultures, leaf explants (5 to 6 mm edge length) were prepared and four each were placed in four 500 mL polypropylene containers containing 80 mL MS medium [45] supplemented with 3% (w/v) sucrose, 0.75% Plant agar (w/v) (Duchefa, Harlem, The Netherlands), 4.44  $\mu\text{M}$  6-benzylaminopurine (BAP) and 1 g L<sup>-1</sup> titanium dioxide. The pH of the medium was adjusted to 5.8 prior to autoclaving at 121 °C for 15 min.

#### Seedling growth of *A. thaliana*

*Arabidopsis thaliana* Col-0 seeds stored since 2018 at 4 °C were surface-disinfected using 70% (v/v) isopropanol for 30 s, followed by 2% (v/v) sodium hypochlorite plus Tween 20 for 5 min and rinsing three times in water. The

seeds were germinated for 10 days at 24 °C in Petri dishes containing plant growth regulator-free B5 medium [46] with 1.5% (w/v) sucrose and 0.8% (w/v) Plant agar at pH 5.8. Ten days old uniform seedling were transferred to the same medium but supplemented with 0.1% (w/v) titanium dioxide (food dye; Ruth GmbH & Co.KG, Bochum, Germany) to achieve an opaque white colored appearance which simplified the detection with optical sensors. Titanium dioxide is commonly used in food production [47]. For this cultivation step, ten 500-mL polypropylene containers containing approximately 80 mL of medium were used, in each of which four seedlings were placed for Trial A and five seedlings for Trial B.

#### Culture conditions

A polyvinyl chloride foil (PVC system foil; Klarsichtpackung GmbH, Hofheim, Germany) sealed each vessel as a substitution of the lid to provide a fully transparent view while ensuring the aseptic condition of the cultures for both experiments. The cultures were incubated for either 21 days (Trial A) or 16 days (Trial B) for *A. thaliana* and 32 days for *N. tabacum* at 25 °C with a 16 h photoperiod (7 am till 11 pm) and a PPFD (Photosynthetic Photon Flux Density) of 35 to 40  $\mu\text{mol m}^{-2} \text{s}^{-1}$ , provided by two tubular fluorescent lamps (Philips MASTER TL-D 58W/865). The lab’s bottom-cooling system—provided by water-cooled plastic tubes below the shelf—prevented water condensation due a local shift of dew point (Fig. 2). Room temperature ranged from 19 °C (night) to 25 °C (day) with an average of 22 °C, while the average surface temperature of the cooled cultivation area ranged from 19 °C (night) to 24 °C (day) with an average of 21 °C.

#### Optical properties of culture vessel

Spectral transmittance was measured with an UV/VIS/NIR spectrometer (PerkinElmer Lambda 1050) equipped with 150 mm indium gallium arsenide (InGaAs) integrating sphere in a 1 nm wavelength interval from 250 to 2500 nm and with a FT-IR Spectrometer (PerkinElmer Spectrum Two) in a 3.75 nm wavelength interval from 2500 to 15,000 nm. Three independent replicates were measured for transmittance curves (Fig. 5). Additionally Haze index was measured with an UV/VIS/NIR spectrometer (PerkinElmer Lambda 1050) in 5 nm intervals and in a wavelength interval from 380 to 780 nm, according to standard test method ASTM D1003 [18]. Thus, the Haze index was calculated by the following equation:

$$\text{Haze index}[\%] = \left( \frac{\text{Diffuse transmittance}}{\text{Total transmittance}} - \text{Rel. scattered transmittance by the system} \right) \times 100 \quad (1)$$

## Development of the automated phenotyping system

### *Environmental conditions of the application area*

Plant in vitro cultures are usually cultivated in multi-layered shelf systems equipped with tubular fluorescent lamps (TFL) as a light source with a photoperiod of 16/8 h in a temperature-controlled culture room (Fig. 2). Plant in vitro culture techniques are characterized among others by the potential of cultivating high numbers of plantlets at minimum space—up to 50 culture vessels can be placed at a cultivation area of  $\sim 0.6 \text{ m}^2$  (1000 mm  $\times$  600 mm) containing multiple explants. The distances between the different levels of the multi-layer shelf systems are mainly determined by the heat dissipation of the fluorescent tubes, which limits the available space of potential sensor application to 400 mm between cultivation area and TFL. For the purpose of automated phenotyping of explants cultured under common in vitro conditions, we therefore developed a low-cost multi sensor system at minimum space.

### *Phenotyping platform hardware setup*

As backbone of the phenotyping system, a commercially available belt-driven xy-gantry was chosen (ACRO system; OpenBuilds, Zephyrhills, USA), that allows direct control of movement via a G-code sent to the native motion controller. The xy-gantry was specified with an accuracy of 0.1 to 0.2 mm by manufacturer. The dimensions of the xy-gantry were reduced to 1000 mm  $\times$  600 mm (X, Y) to match the dimensions of the shelf used in the culture room of Leibniz University Hannover (Fig. 2A). To fulfill the specific demands of monitoring in vitro cultures, several hardware components were added to the gantry. In order to control the height of the multi-sensor detector head (Fig. 2), and in particular to accommodate the variable needs of dynamically monitoring different plant species, we installed an additional screw-driven z-axis (C-Beam Linear Actuator, OpenBuilds, Zephyrhills, USA; modified to a stroke length of 60 mm) and connected it to a motion controller. The linear actuator for the z-axis was specified with an accuracy of 0.05 to 0.1 mm by manufacturer. The cable management was ensured by various 3D-printed parts and common cable chains (GitHub repository [28]). Network connection and power supply of the two single-board computers (Raspberry Pi 4 Model B), controlling either the sensors of the detector head or the serial communication of the G-code to the motion controller, were provided by a router and a Power-over-Ethernet switch (Table 2).

### *Detector head hardware setup*

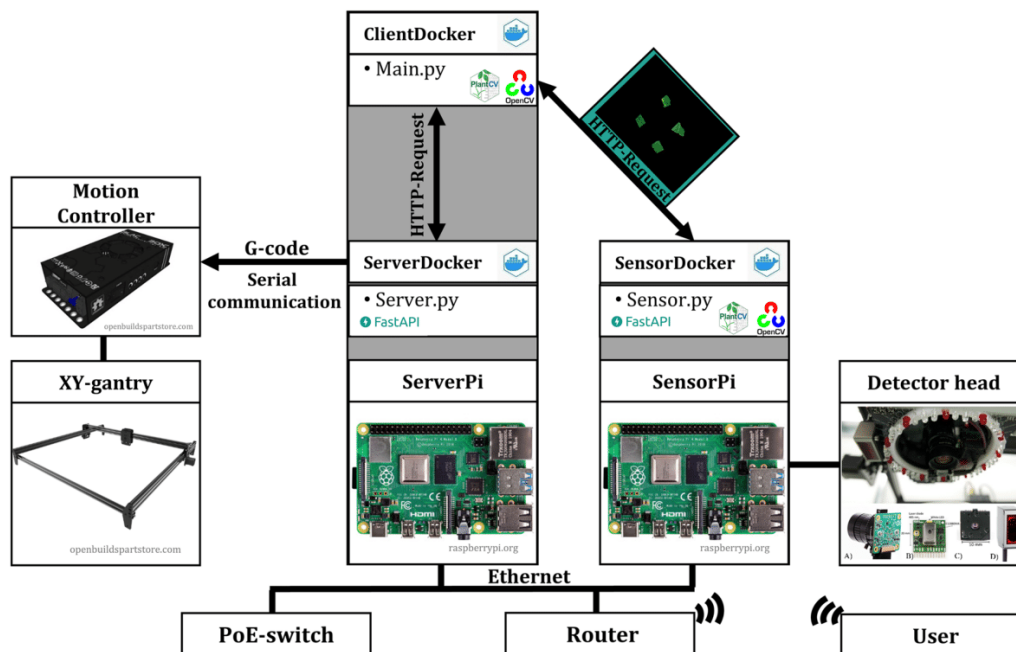
The detector head installed at the z-axis of the system consists of four different sensors (Fig. 2B–F) and diverse

LEDs for the illumination (Fig. 2G), including a laser distance sensor (Fig. 2C), a low cost RGB camera (Fig. 2D), a micro spectrometer breakout board (Fig. 2E) and a thermal camera board (Fig. 2F).

The laser distance sensor (OD-Mini OB1-B100, Sick AG, Waldkirch, Germany) used in this setup was specified by the manufacturer with a power consumption of  $< 1.92 \text{ W}$ , laser emission wavelength of 655 nm, max. output of 390  $\mu\text{W}$  (laser class 1), a measuring range of 50 to 150 mm and a linearity of  $\pm 100 \mu\text{m}$  as well as spot size of 700  $\mu\text{m} \times 600 \mu\text{m}$  at a measuring distance of 100 mm. The analog output of the laser distance sensor (10 V) was connected via a small voltage divider circuit to a high precision 16-bit A/D-converter (ADS 1115), which communicated via Inter-Integral Circuit (I<sup>2</sup>C) with a microcontroller board (Wemos D1 Mini). The A/D-converter gain was set to 2/3 to read a voltage range of  $\pm 6.144 \text{ V}$  and therefore, cover the analog output range of 0 to 5 V. Each distance measurement consisted of a up to ten single readouts and averaging (excluding default sensor values), to achieve a robust and low noise measurement. The microcontroller was powered and read out via USB by the Raspberry Pi of the detector head (Fig. 14: SensorPi).

The 12.3-megapixel RGB camera (Raspberry Pi Camera HQ, Raspberry Pi Foundation, Cambridge, UK) was installed in the center of the ring light PCB to capture top-down images of the in vitro culture vessels (Fig. 2B). The device was electrically connected to a Raspberry Pi via CSI (Camera Serial Interface). The RGB camera was equipped with a 6 mm fixed focal length low-distortion lens (Table 2: Edmund Optics: 6 mm wide angle lens, f/1.2, high resolution = 120 lp/mm, low distortion  $< 0.5\%$ ) to achieve a field of view (FOV) of  $> 100 \text{ mm} \times 100 \text{ mm}$  at a minimum working distance (MWD) of  $\sim 100 \text{ mm}$ , mainly determined by the height of the culture vessels used (500 mL transparent polypropylene containers with a height of 104 mm). Images of in vitro cultures were captured with the following camera parameters: resolution = 4054 px  $\times$  3040 px, shutter speed = 2000 ms, iso = 100, autowhite-balance = off and a fixed gain of 3.3, 1.5 (red, blue).

The micro spectrometer board (micro spectrometer and Breakout Board v2, GroupGets, Reno, USA) allows an easy application of the ultra-compact Hamamatsu CMOS image sensor (C12880MA, Hamamatsu Photonics K.K., Hamamatsu, Japan), which has 288 channels with a spectral range of 340 to 850 nm and a spectral resolution of 15 nm. The sensor's pixel index was converted to wavelength with the device-specific factory calibration coefficients and resulting wavelengths were round to integers. The micro spectrometer board was powered



**Fig. 14** Software design and network communication of the phenotyping system components. Two Raspberry Pis hosting Docker containers executing scripts for the two main tasks of motion control and sensor control and providing the access over HTTP request via the Python framework FastAPI by the main script (semi-autonomous mode/local communication) or the user (reliant mode/ wireless communication). Gray areas represent physically co-located software elements, arrows indicate the direction of data transfer, and black lines mark physically connected hardware components

and readout by a microcontroller (Wemos D1 Mini) connected via USB to the Raspberry Pi of the detector head (SensorPi). The analog values of the micro spectrometer were digitized by the 10-bit internal A/D converter of the microcontroller. The micro spectrometer was equipped with a 3D printed tubular aperture ( $\varnothing$  5 mm), which was used to integrate a long pass filter (Table 2: Edmund Optics: G420; OD > 5; transmission > 90%) and to limit the detection spot size of the spectrometer. Due to the limited signal, an integration time of 300 ms was specified for fluorescence detection. The dielectric long pass filter with a cut-on wavelength of  $420 \pm 5$  nm was used to separate excitation light of the UV-LEDs (Fig. 2G) from the chlorophyll fluorescence signal measured in dark condition (night).

The thermal camera board (PureThermal 2, GroupGets, Reno, USA) was equipped with the FLIR Lepton 3.5 thermal camera (Lepton 3.5, Teledyne FLIR LLC, Wilsonville, USA). This low-cost device is a radiometrically calibrated thermal camera, sensitive to longwave infrared radiation from 8 to 14  $\mu$ m, with a spatial resolution of 160 px  $\times$  120 px, a horizontal field of view (FOV) of 57°, a radiometric accuracy of up to  $\pm 5$  °C and a thermal sensitivity of 0.05 °C. Power supply and data readout was

ensured by a USB connection to one of the Raspberry Pis (SensorPi). The internal flat field calibration (dark current correction with closed shutter) was set to be performed every 90 s.

As an essential requirement of the image acquisition of plant in vitro culture, the illumination of the detector head (Fig. 2G) had to limit total reflection, occurring at the culture media surface and the lid of the culture container, to a minimum. Therefore, the illumination setup mainly included diffuse and non-direct lateral illumination. To enable various illumination options and to provide appropriate signal for the spectral measurements, we designed a ring light printed circuit-board (GitHub repository [28]), which consists of 24 white standard LEDs (ROHM Semiconductor: SLA560WBC7T3), 12 UV standard LEDs with a peak maximum of 375 nm (Nichia: NSPU510CS) and 12 red standard LEDs (Lumex: SSL-LX5093HD) with a peak maximum of 700 nm. LEDs were controlled by a Mosfet circuit connected to one of the microcontrollers (Wemos D1 mini) and powered by eight 20 mA micro constant current power supplies. Additionally, a 24 V diffuse ring light with white LEDs and a color temperature of 6500 K was added as the main illumination source for image acquisition.



### Phenotyping platform software setup

The requirements of the software setup comprised (i) remote and programming language independent access and control, and (ii) automatic and robust data transfer from the phenotyping system sensors over weeks. The software design is mainly based on the Python [22] programming language and includes open-source components like Docker [48], FastAPI [49], OpenCV [23] and PlantCv [24] (Fig. 14). To ensure software reproducibility and flexibility—independent of framework and operating system versions—we decided to containerize the applications with Docker (ServerDocker, SensorDocker and ClientDocker) which contain the Python-based main scripts, according to respective task (sensor control, motion control and the fusion of the two tasks). Network communication between the different containers is ensured by a Python framework “FastAPI,” which allows the execution of Python functions, provides the network addressing and the access of sensor data via HTTP requests of the different physically separated network components, resulting in control of the system via HTTP independently of programming languages. The system specific Python library containing all self-defined functions is accessible at our GitHub repository [28]. The hardware and software setup yielded a portable and stand-alone system, allowing a semi-automated sensor data acquisition.

### Automated data acquisition

#### Step I: Start of the system

Step I included the start of the system and the determination of culture vessel position (Fig. 3). To run the phenotyping system in an autonomous way over weeks of monitoring, the positions of the to be monitored culture containers as a single user input had to be initially set in the Python main script. Alternatively, a vessel detection algorithm based on Circle Hough Transform implemented in our Python library can be used—if some input constants are adjusted to the respective imaging situation. To run the monitoring experiment the started ClientDocker executed the main Python script (Main.py) and thus the library is included with all necessary functions and system constants.

#### Step II: Data structure and capture initial images for determination of plant positions

Once the system is started, the output directory (256 GB USB drive connected to ServerPi) is checked for already existing experiments, then a new experiment folder and subfolders for each culture vessel are created. The culture vessel positions are sequentially approached, capturing initial images and directly determining plant positions

of the four largest objects, found by image color space transformation in HSV (hue, saturation and value) and thresholding of the hue channel with Otsu’s method [20] (Fig. 3). Plant positions are calculated by deviation of the centroid of the found objects—converted from pixel to mm—and the known position of image midpoint (xy-position of the motion controller).

#### Step III: Time lapse data acquisition

After the initial steps I and II, the actual time lapse data acquisition is continuously looped over the time of the experiment (Fig. 3). In our experiments, RGB image acquisition was performed sequentially for each culture vessel at the midpoint every 4 h. Thermal images were captured simultaneously with the RGB images, with the thermal sensor shifted to the center point in the xy direction. To determine whether additional illumination is required for RGB night shots, the average pixel intensity of an RGB image previously captured without system illumination was calculated. Once the system recognized a night image situation, the estimated plant positions were sequentially approached to capture fluorescence spectral information with the micro spectrometer at the centroid of the found objects and UV excitation lights turned on (Fig. 3). After that, consequently the acquisition of depth data with the point-measuring laser distance sensor was obtained via spatial scan by sequential readout of the sensor point measurements while shifting the detector head in xy direction, according to the scan pattern (e.g., 100 mm×100 mm; with a resolution of 1 mm×1 mm) with a speed around 0.27 s per point×10,000 points per vessel (~45 min), which limited the measurement of depth data to two culture vessels per 4-h cycle. For the experiments conducted in this paper, a depth measurement for each culture container once per day was ensured.

### Data processing

#### RGB data processing

Classical image processing approaches—applying thresholds to certain color space channels—failed in different previously conducted experiments due to a high variability and diversity in the obtained image data sets, for instance due to changing illumination situations during the day or due to changes in leaf pigment composition (Fig. 4A). To obtain a robust image classification, we therefore trained a pixel-wise random forest classifier with Ilastik [21]. Ilastik is an open-source toolkit offering machine learning (ML) based image processing for pixel and object classification and tracking. 50 random RGB images of the *A. thaliana* dataset were selected and partially labeled pixel-wise in either background or plant pixels. Features selection was limited to a number

of 14 features to reduce computation time (Additional file 7). After verification of the classifier, the model was exported and used in the RGB image processing pipeline in Python script.

PlantCv [24] was used to set up the image processing pipeline allowing a uniform batch-processing of hundreds of images. The trained classifier was executed in the headless mode to obtain the segmentation binary image containing only plant pixels (Fig. 4B). Image processing included an automated brightness and contrast adjustment by histogram stretching and a temporary reduction in resolution to reduce computation time from 4054 px × 3040 px to 1014 px × 760 px while using the pixel-wise classifier. After obtaining the binary plant masks, the connected components analysis was carried out, mainly with established PlantCv functions. For single plant analysis, the following parameters could be calculated: projected plant area, perimeter, convex hull area to calculate solidity/compactness and stockiness (data not shown). Additionally, the cumulative projected plant area of all explants could be determined by the sum of non-zero pixels in the segmented binary plant mask.

#### Depth data processing

The level of zero depth was calculated separately for each culture vessel as the mean of the raw sensor data from four quadrilaterals (10 px × 10 px) of each corner of the scan area (100 mm × 100 mm)—where values only derived from the cultivation surface and not from the media or plant. To obtain depth data from raw sensor values of the laser distance sensor, the calibration curve of a reference object for data conversion was used (Additional file 5). Circle Hough Transform [26] was employed to detect the culture medium in the depth data from Day 0 (Fig. 4C). With the radius ( $r_1$ ) of the detected circle, a circular binary mask was created with  $r_{\text{new}} = r_1 - 3$  px, which allowed the removal of disturbing edges of the culture medium for further determination of plant height parameters. RANSAC [18]-based plane detection was therefore applied to the edge-removed point cloud to dynamically identify the eventually tilted medium surface (Fig. 4C). Here, the following parameters were set to detect the planes (distance threshold = 1.5, sample size = 3, iterations = 10,000). The obtained RANSAC plane of the medium was subtracted from the processed point cloud, resulting in height correction and segmentation of the plant depth data. Sum of non-zero pixels of segmented depth pixels (Background: 0, plant: 1) allowed the calculation of plant area by depth data.

With the processed depth data, the following parameters were calculated:

- Medium height (mm): Mean of estimated RANSAC plane
- Medium volume ( $\text{mm}^3$ ): Assuming a circular conical frustum ( $V = \frac{1}{3}\pi h (r_1^2 + r_1r_2 + r_2^2)$ ;  $h$  = medium height,  $r_2 = 37$  mm; radius of the bottom surface of the culture vessel (constant))
- Average canopy height (mm): Mean of the height corrected plant depth data (output)
- Maximum plant height (mm): Mean of the upper 10 percentile of the corrected plant depth data
- Projected plant area<sub>depth data</sub> ( $\text{mm}^2$ ): Count of non-zero pixels of segmented and height corrected plant depth data

#### Spectral data processing

For spectral data processing, the analysis focused on the fluorescence measurement, since, in contrast to the spectral reflection, the fluorescence spectra were derived almost exclusively from plant tissue. First, the dark current noise was calculated (spectrometer readout at night, with no excitation light on) as the mean of all dark current measurements. From all fluorescence spectra this mean dark spectrum was then subtracted. For a simplified visualization, the region between 400–660 nm was masked. The masked region contained signals of residual light of the culture room and excitation light (UV) due to imperfect blocking properties of the used longwave filter.

#### Thermal data processing

The image situation in the wavelength region of the spectral sensitivity of the thermal camera was challenging due to the optical properties of the culture vessels. A successful and robust implementation of thermal data acquisition allowed the readout of the 14-bit raw grayscale image by the use of Python library *Flirpy*. Thermal data processing included a conversion of 14 bit grayscale values to °C by manufacturer-specified conversion ( $y_{\text{Celsius}} = y_{\text{raw}} / 100 - 273.15$ ).

#### Calibration of the “Phenomenon” phenotyping system

##### Xyz-gantry movement calibration

The motion controller of xy-gantry has been set up with manufacturer-specificized GRBL settings for each axis respectively (GitHub repository [28]), that allow the stepper motor motion to be translated into steps in metric units.

##### RGB sensor calibration

A relation between pixel and metric units was established to express the projected plant area in square millimeters by counting pixels of a graph paper image at the average media height of 20 mm (1 mm = 37.7 px).

**Laser distance sensor calibration**

Laser distance raw sensor data were technically calibrated by measuring a staircase shaped reference object. Therefore, z-axis was set to the same value as used in later experiments (z-axis = -40 mm = detector head height ~ 130 mm). Reference heights were obtained by a caliper for 6 different heights and 119 raw sensor values were used for calibration. Thus, the zero plane for specific sensor Z-height (z-axis<sub>zero</sub> = 19430) as well as the maximal valid height could be determined (Additional file 5, maximum height < = 72 mm). The obtained linear regression function determined the metric conversion of raw sensor data in all conducted experiments.

**Spectrometer detection spot size determination**

Two approaches were used to estimate the measuring spot diameter of the micro spectrometer that had to be modified with a 3D-printed aperture tube (GitHub repository [28]) to reduce the size of the measurement spot: a graphical estimation and an experimental determination. Additional file 8 contains a schematic sketch for the graphical estimation of the detection spot size diameter of 23.5 mm. The experimental determination included a sequential spectrometer readout every 1 mm, while linear movement in x-axis over a grid with black background and white squares of decreasing size and a side length ranging from 30 to 21 mm. Spectrometer channel readouts with the highest signal were picked from the array and plotted over the x axis. We assumed that if the detection spot size diameter is smaller than the side length of the square a constant plateau is found in the respective peak. The first square where a sharp maximum was identifiable, or in particular its side length of 23 mm, determined the detection spot size diameter of this approach.

**Validation of the “Phenomenon” phenotyping system**

**Validation of xy and z-axis repositioning accuracy of the “Phenomenon” phenotyping system**

Determination of technical repeatability of xy-axis repositioning over time was conducted by measuring the midpoint deviation by RGB images of a reference object with a flat surface and a height of 41 mm over 16 days for a certain timepoint (12 o'clock), under the settings that were used in all experiments. The initial midpoint (Day 0) of the largest found object in Otsu-binarized L-Channel of CIELAB colorspace was set as the reference for calculation of the mean absolute error (MAE) for x- and y-axis (Table 3). The daily measurement procedure included an initial zeroing through limit switches, repositioning and RGB data acquisition. Reference object surface area of 50 mm x 50 mm and founded counts of px were related

**Table 3** Technical repeatability of xy-gantry repositioning via RGB image analysis

Day [d]	MAE <sub>x</sub> [px]	MAE <sub>x</sub> [mm]	MAE <sub>y</sub> [px]	MAE <sub>y</sub> [mm]
1	8	0.17	3	0.06
2	10	0.21	4	0.09
3	7	0.15	2	0.04
4	8	0.17	7	0.15
5	9	0.19	5	0.11
6	9	0.19	4	0.09
7	10	0.21	3	0.06
8	11	0.24	4	0.09
9	11	0.24	6	0.13
10	12	0.26	6	0.13
11	12	0.26	4	0.09
12	14	0.3	4	0.09
13	15	0.32	6	0.13
14	14	0.3	0	0
15	14	0.3	0	0
Total	10.9	0.23	3.9	0.08

to convert the midpoint deviation of the reference object in px to metric units at a height of 41 mm of the reference object (1 mm = 46.7 px).

Determination of technical repeatability of z-axis over time was conducted by setting five different z-axis values by the motion controller. Each Z step (0 mm, -6 mm, -20 mm, -40 mm, -50 mm) was approached five times with initial zeroing through limit switches each time. Actual height changes were recorded by the calibrated laser distance values. Linear regression analysis revealed an R<sup>2</sup> > 0.99, a MAE<sub>z</sub> of 0.09 mm and a RMSE of 0.11 mm (Additional file 4).

**Validation of the RGB image processing pipeline**

The performance of the RGB image processing pipeline, in particular the image segmentation part was checked by manual plant pixel labeling with the annotation software “LabelMe” [50]. 18 randomly selected images from the *A. thaliana* Trial A dataset were used with 3 images per time point and including images of 9 different culture vessels. The 18 binary masks from manual segmentation, thus forming the ground truth dataset, were matched against respective binary masks derived from our RGB image processing pipeline. Plant area was calculated by the sum of non-zero pixels in binary images (Background: 0, plant: 1), while for confusion statistics a full comparison between the two data sets were necessary, revealing 221,834,880 pixel pairs where plant pixels reflected the true positive class and background pixels represented the true negative class.

### Validation of depth data processing

To estimate the quality of the representation of plants in depth images we selected the projected plant area as a basis of comparison between RGB and depth sensor data and corresponding pipelines, respectively. Therefore, we converted projected plant area by the RGB image processing pipeline from px to square millimeters ( $37.7 \text{ px} \times 37.7 \text{ px} = 1 \text{ mm}^2$ ), which allowed comparison with projected plant area by depth data processing pipeline. Projected plant area from RGB and depth data of 4 culture containers at 12 time points ( $n = 48$ ) were submitted to a linear regression analysis, assuming the RGB segmentation as the ground truth data.

### Software environment for data acquisition, processing, analysis and visualization

Data acquisition was done mainly with *Python v3.8.8* [22], using in particular the libraries *FastAPI* [49], *OpenCV v3.4.9* [23], *NumPy v1.20.2* [51], *Serial v3.4* [52], *Picamera v1.13* [53], *Flirpy v0.3.0* [54] and with *Arduino IDE 1.8.19* [55] with the following libraries: *arduino-microspec* [56], *SerialCommand* [57] and *Adafruit\_ADS1015* [58]

RGB Image processing and analysis was conducted with *Python v3.8.8* [22] in the *Jupyter Notebook v6.3.0* [59] environment using the following packages: *PlantCv v3.11.0* [24], *OpenCV v3.4.9* [23], *NumPy v1.20.2* [51], *Matplotlib v3.4.1* [60], *scikit-image v0.18.1* [61] and the Software toolkit *Ilastik v1.3.3* [26] headless integrated in the Python script.

Depth data analysis included subsequent additional *Python* libraries: *Pandas v1.4.2* [62], *Open3D v0.15.1* [25], *Pyvista v0.34.0* [63].

For data visualization, spectral data analysis and statistical analysis, where statistical test assumptions were proofed graphically, we used *R v4.1.2* [64] and the R-packages *dplyr v1.0.8* [65], *ggplot2 v 3.3.5* [66], *kable-Extra v1.3.4* [67], *purrr v0.3.4* [68], *readr v2.1.2* [69], *tidyverse v1.3.1* [70], *hyperSpec v0.100.0* [71] and *photobiology* [72].

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13007-023-01018-w>.

**Additional file 1.** Time lapse video of shoot regeneration of *N. tabacum* in vitro. Leaf explants were cultivated at MS medium supplemented 4.44  $\mu\text{M}$ . Shoot development was monitored over 32 days of cultivation. Images were segmented with a trained classifier. Uncompressed video is available from the corresponding author on reasonable request.

**Additional file 2.** Time lapse video with original images of *A. thaliana* growth in vitro. 10 days old seedlings were cultivated on modified B5 medium and monitored for 16 days. Uncompressed video is available from the corresponding author on reasonable request.

**Additional file 3.** Time lapse video with segmented images of *A. thaliana* growth in vitro. 10 days old seedlings were cultivated on modified B5 medium and monitored for 16 days. Images were segmented with a trained classifier. Uncompressed video is available from the corresponding author on reasonable request.

**Additional file 4.** Technical repeatability of Z-axis repositioning. Five different z-axis values were set to the motion controller and approached five times with initial zeroing through limit switches each time. Actual height changes were recorded by the calibrated laser distance values.

**Additional file 5.** Calibration of laser distance sensor. Linear regression of raw sensor values of the laser distance sensor. The reference height was determined with a caliper of a staircase-shaped object (RGB and depth image in bottom left corner). The regression line is colored black, while the linear regression extrapolation is drawn dashed. Gray indicates confidence interval limits at  $\alpha = 0.95$ .  $\text{Adj } R^2$  denotes the coefficient of determination adjusted according to Yin and Fan [27], while  $P_{\text{slope}}$  and  $P_{\text{intercept}}$  represent p-values of the coefficients for the intercept and slope determined by simple T-test. MAE and RMSE indicate the mean absolute error and the root mean square error of calibration.  $n = 119$ .

**Additional file 6.** Technical repeatability of spatial scanning with laser distance sensor over time. Determination of technical repeatability over time was conducted by measuring a reference object with a flat surface and a height of 41 mm once per day over 6 days, under the settings that were used in all experiments. The initial depth measurement of an area of 50 mm  $\times$  50 mm was set as the reference for calculation of the mean absolute error and the root mean square error. The daily measurement procedure included an initial zeroing through limit switches, repositioning and depth data acquisition by spatial scan.

**Additional file 7.** Random forest classification model features for segmentation of *A. thaliana* Trail A.

**Additional file 8.** Experimental and graphical determination of modified spectrometer detection spot size. Image of the modified spectrometer are shown in upper right corner. A) Experimental determination of spectrometer detection spot size by a sequential spectrometer readout every 1 mm, while linear movement in x-axis over a grid with black background and white squares of decreasing size and a side length ranging from 30 to 21 mm. Spectrometer channel readouts with the highest signal were picked from the array and plotted over the x-axis. We assumed that if the detection spot size diameter is smaller than the side length of the square a constant plateau is found in the respective peak. The first square where a sharp maximum was identifiable, or in particular its side length of 23 mm determined the spot size diameter. B) Graphical estimation by drawing at a 1:1 scale. Graphical determination found a spot size diameter of 23.5 mm.

### Acknowledgements

We thank Ms. Zahra Mohammadi Nakhjiri and Mr. Dr. Philipp Bethge for their assistance in lecturing the manuscript. We also thank the technical assistants of the department of Woody Plant and Propagation Physiology for their lab work.

### Author contributions

HB and PL developed the hardware and software of the system. HB designed and performed the experiments and analysed the data. HB wrote the manuscript and HB, TW, LP, TR revised the manuscript. All the authors discussed the results and collectively edited the manuscript. All authors read and approved the final manuscript.

### Funding

Open Access funding enabled and organized by Projekt DEAL. This project took place within the research project "Experimentierfeld Agro-Nordwest", which is funded by the Federal Ministry of Food and Agriculture (BMEL, Grant No.: 28DE103F18) via the Federal Agency for Agriculture and Food (BLE).

**Availability of data and materials**

The dataset supporting the conclusions of this article (Hard- and Software of "Phenomenon" phenotyping system) are available in an open-access Github repository, <https://github.com/halube/Phenomenon>. Supporting information contain beside time lapse videos of conducted experiments, mainly "Methods" section supporting figures and tables. The images data sets of the biological experiments are available from the corresponding author on reasonable request.

**Declarations****Ethics approval and consent to participate**

Not applicable.

**Competing interests**

The authors declare no competing interests.

**Author details**

<sup>1</sup>Laboratory for Biosystems Engineering, Faculty of Agricultural Sciences and Landscape Architecture, Osnabrück University of Applied Sciences, Oldenburger Landstraße 24, 49090 Osnabrück, Germany. <sup>2</sup>Institute of Horticultural Production Systems, Section of Woody Plant and Propagation Physiology, Leibniz Universität Hannover, Herrenhäuser Str. 2, 30419 Hannover, Germany. <sup>3</sup>Hannover, Germany.

Received: 12 September 2022 Accepted: 14 April 2023

Published online: 02 May 2023

**References**

- Araus JL, Kefauver SC, Zaman-Allah M, Olsen MS, Cairns JE. Translating high-throughput phenotyping into genetic gain. *Trends Plant Sci.* 2018;23(5):451–66.
- Smith MA, Spomer L, Meyer MJ, McClelland MT. Non-invasive image analysis evaluation of growth during plant micropropagation. *Plant Cell, Tissue Organ Cult.* 1989;19(2):91–102.
- Aynalem HM, Righetti TL, Reed BM. Non-destructive evaluation of in vitro stored plants: a comparison of visual and image analysis. *In Vitro Cell Dev Biol-Plant.* 2006;42(6):562–7.
- Dhondt S, Gonzalez N, Blomme J, De Milde L, Van Daele T, Van Akoleyen D, Storme V, Coppens F, Beemster GTS, Inzé D. High-resolution time-resolved imaging of in vitro *Arabidopsis* rosette growth. *Plant J.* 2014;80(1):172–84.
- Gupta SD, Karmakar A. Machine vision based evaluation of impact of light emitting diodes (LEDs) on shoot regeneration and the effect of spectral quality on phenolic content and antioxidant capacity in *Swertia chirata*. *J Photochem Photobiol, B.* 2017;174:162–72.
- Faragó D, Sass L, Valkai I, Andrási N, Szabados L. PlantSize offers an affordable, non-destructive method to measure plant size and color in vitro. *Front Plant Sci.* 2018;9:219.
- Niazian M, Sadat-Noori SA, Abdipour M, Tohidfar M, Mortazavian SM. Image processing and artificial neural network-based models to measure and predict physical properties of embryogenic callus and number of somatic embryos in ajowan (*Trachyspermum ammi* (L.) Sprague). *In Vitro Cell Dev Biol-Plant.* 2018;54(1):54–68.
- Mestre D, Fonseca JM, Mora A. Monitoring of in-vitro plant cultures using digital image processing and random forests. 2017.
- Ibaraki Y, Gupta SD. Thermal imaging of micropropagated plantlets for evaluation of possible wilting. *Environ Control Biol.* 2011;49(3):141–8.
- Mansouri A, Fadavi A, Mortazavian SM. An artificial intelligence approach for modeling volume and fresh weight of callus—a case study of cumin (*Cuminum cyminum* L.). *J Theor Biol.* 2016;397:199–205.
- Zhang C, Timmis R, Hu WS. A neural network based pattern recognition system for somatic embryos of Douglas fir. *Plant Cell, Tissue Organ Cult.* 1999;56(1):25–35.
- Prasad VS, Gupta SD. Applications and potentials of artificial neural networks in plant tissue culture. In: *Plant tissue culture engineering.* 2008:47–67.
- Dutta Gupta S, Ibaraki Y, Pattanayak AK. Development of a digital image analysis method for real-time estimation of chlorophyll content in micro-propagated potato plants. *Plant Biotechnol Rep.* 2013;7(1):91–7.
- Ibaraki Y, Kenji K. Application of image analysis to plant cell suspension cultures. *Comput Electron Agric.* 2001;30(1–3):193–203.
- Winkelman T, Geier T, Preil W. Commercial in vitro plant production in Germany in 1985–2004. *Plant Cell, Tissue Organ Cult.* 2006;86(3):319–27.
- Chen C. Cost analysis of plant micropropagation of *Phalaenopsis*. *Plant Cell, Tissue Organ Cult.* 2016;126(1):167–75.
- Cardoso JC, Sheng Gerald LT, Teixeira da Silva JA. Micropropagation in the twenty-first century. In: *Plant cell culture protocols.* 2018:17–46.
- Fischler MA, Bolles RC. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun ACM.* 1981;24(6):381–95.
- American Society for Testing and Materials. Standard test method for haze and luminous transmittance of transparent plastics. ASTM D 1003. 2003
- Otsu N. A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern.* 1979;9(1):62–6.
- Sommer C, Straehle C, Koethe U, Hamprecht FA. Ilastik: Interactive learning and segmentation toolkit. In: 2011 IEEE international symposium on biomedical imaging: from nano to macro 2011 Mar 30. IEEE. pp. 230–3.
- Van Rossum G, Drake FL. Python 3 reference manual. CreateSpace; 2009.
- Bradski G. The openCV library. Dr Dobb's J Softw Tools Prof Program. 2000;25(11):120–3.
- Gehan MA, Fahlgren N, Abbasi A, Berry JC, Callen ST, Chavez L, Doust AN, Feldman MJ, Gilbert KB, Hodge JG, Hoyer JS. PlantCV v2: image analysis software for high-throughput plant phenotyping. *PeerJ.* 2017;5:e4088.
- Zhou QY, Park J, Koltun V. Open3D: a modern library for 3D data processing. arXiv preprint [arXiv:1801.09847](https://arxiv.org/abs/1801.09847). 2018.
- Duda RO, Hart PE. Use of the Hough transformation to detect lines and curves in pictures. *Commun ACM.* 1972;15(1):11–5.
- Yin P, Fan X. Estimating R<sup>2</sup> shrinkage in multiple regression: a comparison of different analytical methods. *J Exp Educ.* 2001;69(2):203–24.
- Bethge H. Phenomenon—low-cost and multi-sensor system for automated phenotyping of plant in vitro culture. 2022. <https://github.com/halube/Phenomenon>. Accessed 07 Sept 2022.
- Dhondt S, Wuyts N, Inzé D. Cell to whole-plant phenotyping: the best is yet to come. *Trends Plant Sci.* 2013;18(8):428–39.
- Ghashghaie J, Brenckmann F, Saugier B. Water relations and growth of rose plants cultured in vitro under various relative humidities. *Plant Cell, Tissue Organ Cult.* 1992;30(1):51–7.
- Safavi K. Effect of titanium dioxide nanoparticles in plant tissue culture media for enhance resistance to bacterial activity. *Bull Environ Pharmacol Life Sci.* 2014;3:163–6.
- Cox A, Venkatachalam P, Sahi S, Sharma N. Silver and titanium dioxide nanoparticle toxicity in plants: a review of current research. *Plant Physiol Biochem.* 2016;107:147–63.
- Clapa D, Borsai O, Leopold L, Coman C, Toma A, Oprea I, Hârta M. The effect of TiO<sub>2</sub> and ZnO<sub>2</sub> nanoparticles upon some biometrical characteristics in soybean (*Glycine max* L. Merrill) in vitro cultures. *Sci Bull Ser F Biotechnol.* 2020;24:31–6.
- Lichtenthaler HK, Hak R, Rinderle U. The chlorophyll fluorescence ratio F690/F730 in leaves of different chlorophyll content. *Photosynth Res.* 1990;25(3):295–8.
- Buschmann C, Langsdorf G, Lichtenthaler HK. Imaging of the blue, green, and red fluorescence emission of plants: an overview. *Photosynthetica.* 2000;38(4):483–91.
- Merlot S, Mustilli AC, Genty B, North H, Lefebvre V, Sotta B, Vavasseur A, Giraudat J. Use of infrared thermal imaging to isolate *Arabidopsis* mutants defective in stomatal regulation. *Plant J.* 2002;30(5):601–9.
- Grant OM, Davies MJ, James CM, Johnson AW, Leinonen I, Simpson DW. Thermal imaging and carbon isotope composition indicate variation amongst strawberry (*Fragaria x ananassa*) cultivars in stomatal conductance and water use efficiency. *Environ Exp Bot.* 2012;76:7–15.
- Honjo T, Takakura T. Effects of CO<sub>2</sub> concentration, light intensity and liquid medium composition for the growth of *Cymbidium* PLB in vitro. *J Agric Meteorol.* 1987;43(3):223–7.
- Miao Y, Wang L, Peng C, Li H, Li X, Zhang M. Banana plant counting and morphological parameters measurement based on terrestrial laser scanning. *Plant Methods.* 2022;18(1):1–6.

40. Zheng J, Liu G, Liu X. Phenotypic parameter extraction system for crops based on supervoxel segmentation. In: Proceedings of the 2nd international conference on graphics and signal processing. 2018 Oct 6. pp. 89–94.
41. Vázquez-Arellano M, Reiser D, Paraforos DS, Garrido-Izard M, Burce ME, Griepentrog HW. 3-D reconstruction of maize plants using a time-of-flight camera. *Comput Electron Agric.* 2018;1(145):235–47.
42. Yuan G, Lu H, Tang D, Hassan MM, Li Y, Chen JG, Tuskan GA, Yang X. Expanding the application of a UV-visible reporter for transient gene expression and stable transformation in plants. *Hortic Res.* 2021;8:234.
43. Bethge H, Mählmann T, Winkelmann T, Rath T. Remote plant sensing and phenotyping—an e-learning tool in higher education. 43. GIL-Jahrestagung, Resiliente Agri-Food-Systeme. 2023.
44. Quambusch M, Pirttilä AM, Tejesvi MV, Winkelmann T, Bartsch M. Endophytic bacteria in plant tissue culture: differences between easy-and difficult-to-propagate *Prunus avium* genotypes. *Tree Physiol.* 2014;34(5):524–33.
45. Murashige T, Skoog F. A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol plant.* 1962;15:473–97.
46. Gamborg OL, Miller R, Ojima K. Nutrient requirements of suspension cultures of soybean root cells. *Exp Cell Res.* 1968;50(1):151–8.
47. Gázquez MJ, Bolívar JP, García-Tenorio García-Balmaseda R, Vaca F. A review of the production cycle of titanium dioxide pigment. 2014.
48. Merkel D. Docker: lightweight linux containers for consistent development and deployment. *Linux J.* 2014;239(2):2.
49. Ramírez S. FastAPI. 2018. <https://github.com/tiangolo/fastapi>. Accessed 05 June 2022.
50. Russell BC, Torralba A, Murphy KP, Freeman WT. LabelMe: a database and web-based tool for image annotation. *Int J Comput Vis.* 2008;77(1):157–73.
51. Van Der Walt S, Colbert SC, Varoquaux G. The NumPy array: a structure for efficient numerical computation. *Comput Sci Eng.* 2011;13(2):22–30.
52. Liechti C. PySerial v3.4. 2020. <https://github.com/pyserial/pyserial>. Accessed 05 June 2022.
53. Jones D. Picamera v1.13. 2018. <https://picamera.readthedocs.io/en/release-1.13/>. Accessed 05 June 2022.
54. Veitch-Michaelis J. Flirpy. 2018. <https://github.com/LJMUAstroecology/flirpy/tree/v0.3.0>. Accessed 05 June 2022.
55. Banzi M. Arduino IDE v1.8.19. 2021. <https://github.com/arduino/Arduino>. Accessed 05 June 2022.
56. Woodworth P. Arduino-microspec. 2016. <https://github.com/open-eio/arduino-microspec/>. Accessed 05 June 2022.
57. Cogswell S. Arduino-SerialCommand. 2011. <http://github.com/p-v-o-s/Arduino-SerialCommand>. Accessed 05 June 2022.
58. Neson C. Adafruit\_ADS1X15. 2015. [https://github.com/adafruit/Adafruit\\_ADS1X15](https://github.com/adafruit/Adafruit_ADS1X15). Accessed 05 June 2022.
59. Kluyver T, Ragan-Kelley B, Pérez F, Granger BE, Bussonnier M, Frederic J, Kelley K, Hamrick JB, Grout J, Corlay S, Ivanov P. Jupyter Notebooks—a publishing format for reproducible computational workflows. 2016.
60. Hunter JD. Matplotlib: a 2D graphics environment. *Comput Sci Eng.* 2007;9(03):90–5.
61. Van der Walt S, Schönberger JL, Nunez-Iglesias J, Boulogne F, Warner JD, Yager N, Gouillart E, Yu T. scikit-image: image processing in Python. *PeerJ.* 2014;19(2):e453.
62. McKinney W. pandas: a foundational Python library for data analysis and statistics. *Python High Perform Sci Comput.* 2011;14(9):1–9.
63. Sullivan C, Kaszynski A. PyVista: 3D plotting and mesh analysis through a streamlined interface for the Visualization Toolkit (VTK). *J Open Source Softw.* 2019;4(37):1450.
64. Team R. RStudio: integrated development for R. Boston: RStudio, Inc; 2015;42(14):84. <http://www.rstudio.com>.
65. Wickham H, François R, Henry L, Müller K. dplyr: a grammar of data manipulation. R package version 0.4. 2015;3:p156.
66. Wickham H, Chang W, Wickham MH. Package 'ggplot2'. Create elegant data visualisations using the grammar of graphics. Version. 2016;2(1):1–89.
67. Zhu H. KableExtra: construct complex table with 'kable' and pipe syntax. R package version. 2019;1(0).
68. Henry L, Wickham H. Purrr: functional programming tools. R package version 0.3. 2020;4.
69. Wickham H, Hester J, Francois R, Jylänki J, Jørgensen M. readr: read rectangular text data. R package version 1.1.1. R Foundation for Statistical Computing. 2017.
70. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M. Welcome to the Tidyverse. *J Open Source Softw.* 2019;4(43):1686.
71. Beleites C, Sergio V. Chemometric analysis of spectroscopic data in R: hyperSpec. In: 7th Workshop FT-IR Spectroscopy in Microbiological and Medical Diagnostic Robert Koch-Institute. 2009.
72. Aphalo PJ. The r4photobiology suite: spectral irradiance. *UV4Plants Bull.* 2015;2015(1):21–9.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

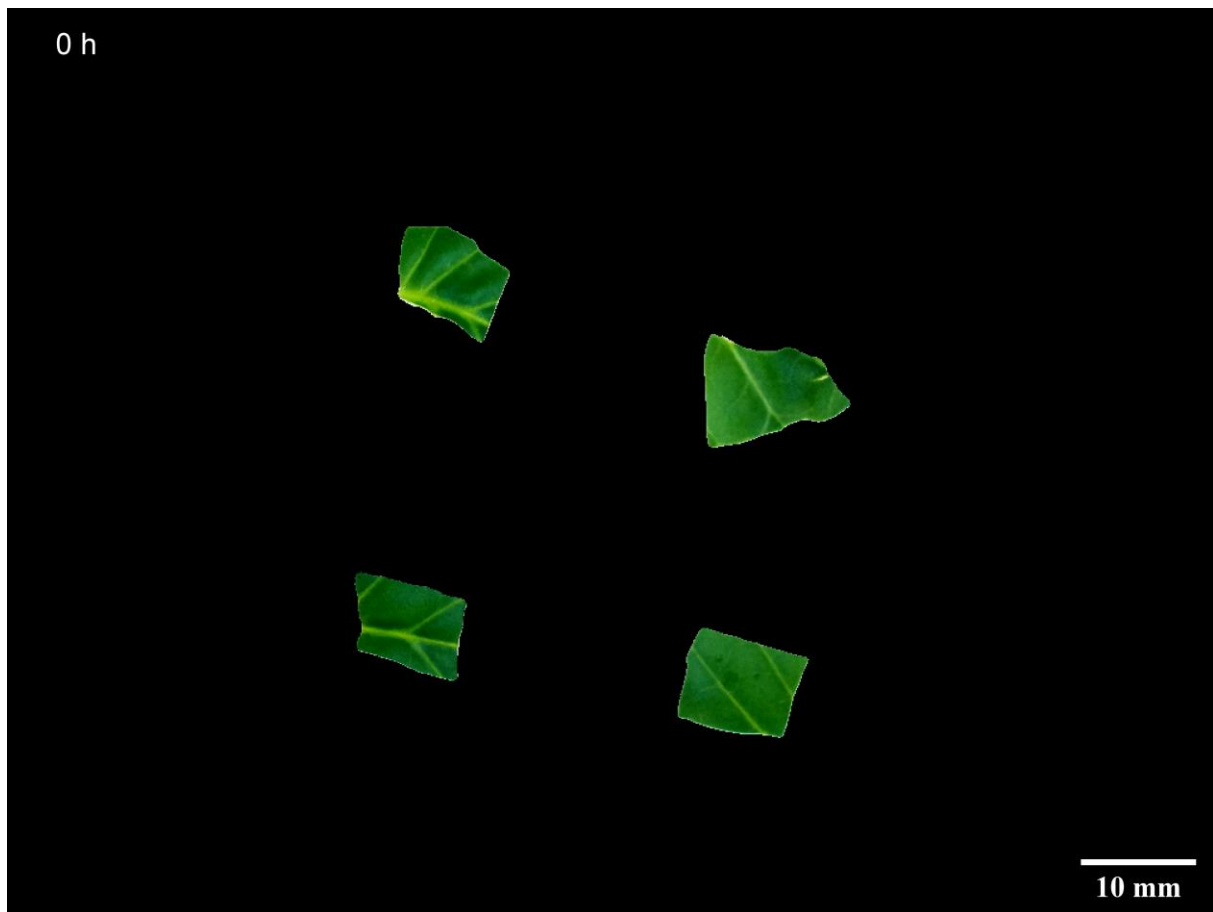
- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

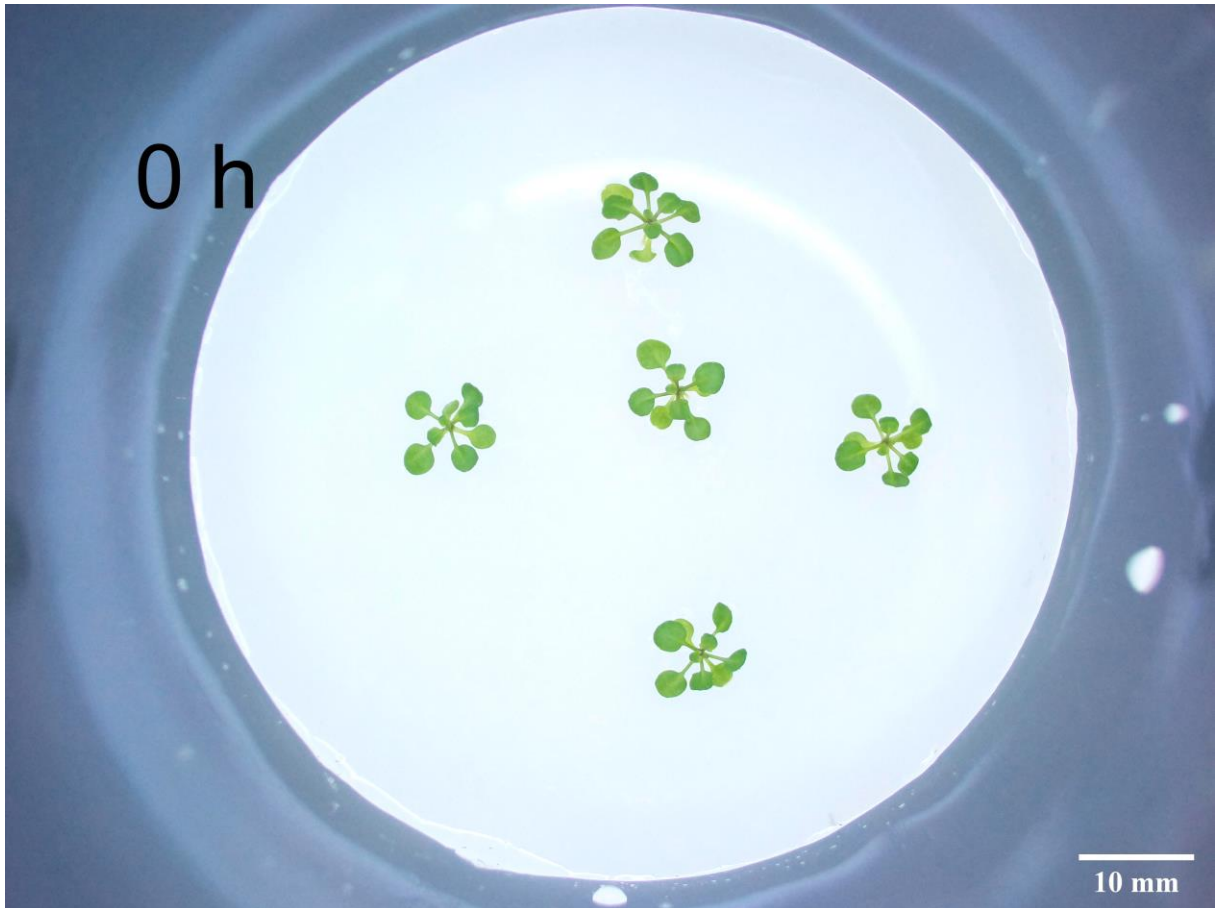
Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)



## Supplementary Information

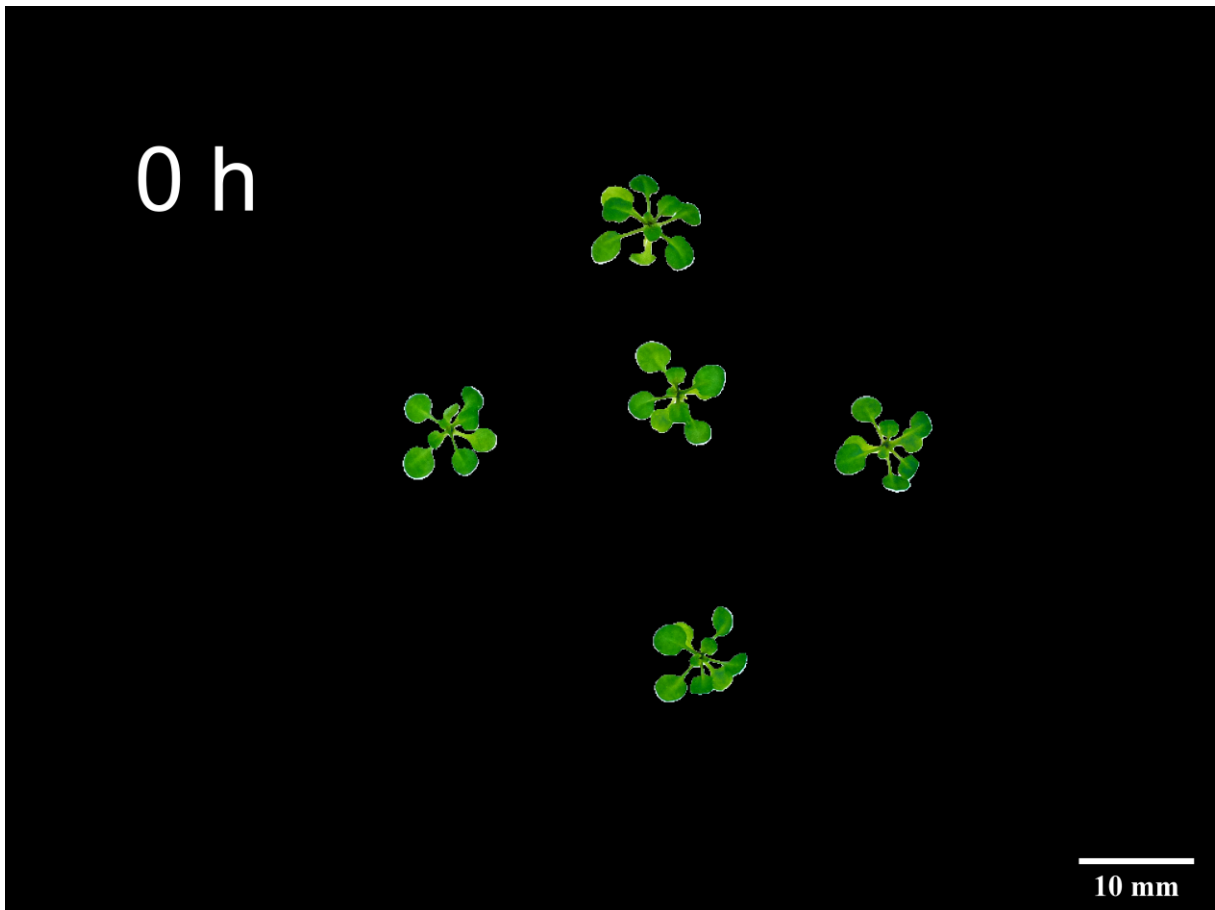


SI. 1: Time lapse video of shoot regeneration of *N. tabacum* in vitro. Leaf explants were cultivated at MS medium supplemented 4.44  $\mu\text{M}$  (BAP). Shoot development were monitored over 32 Days of cultivation. Images were segmented with a trained classifier.



SI. 2: Time lapse video with original images of *A. thaliana* growth in vitro. 10 days old seedling were cultivated on modified B5 medium (see Methods) and monitored for 16 days.

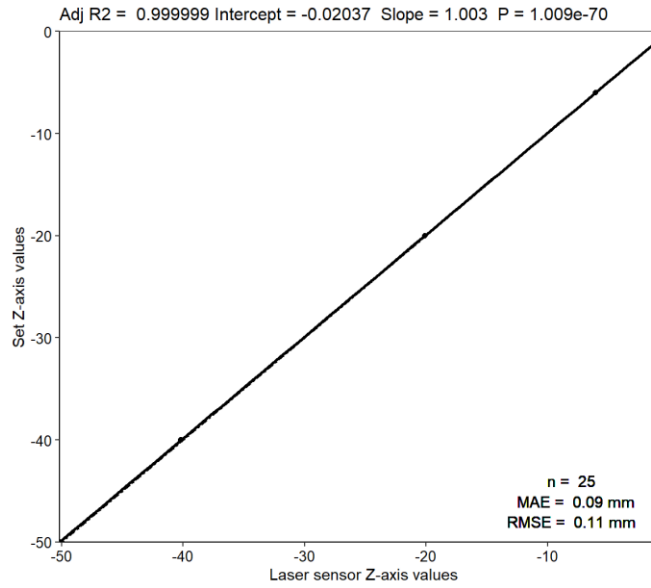




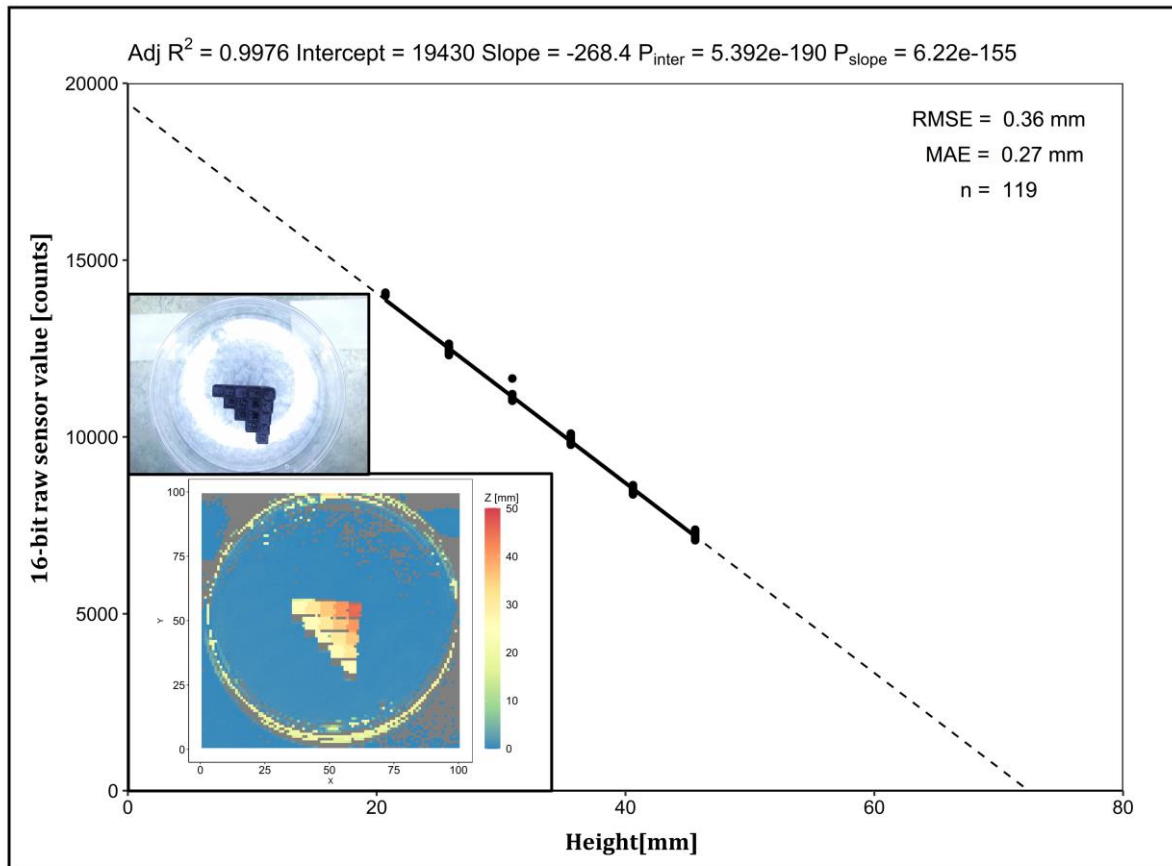
SI. 3: Time lapse video with segmented images of *A. thaliana* growth in vitro. 10 days old seedling were cultivated on modified B5 medium (see Methods) and monitored for 16 days. Images were segmented with a trained classifier.

SI. 4 - Table 1: Technical Repeatability of Z-axis repositioning

Specified Z-axis value [mm]	Calibrated laser distance sensor readout [mm]	MAE [mm]
0	0.000	0.000
0	-0.041	-0.0401
0	-0.011	-0.0111
0	-0.037	-0.0373
0	-0.034	-0.0335
-6	-6.028	0.0283
-6	-6.025	0.0246
-6	-6.021	0.0209
-6	-6.058	0.0581
-6	-6.017	0.0171
-20	-20.078	0.0782
-20	-20.082	0.0820
-20	-20.082	0.0820
-20	-20.097	0.0969
-20	-20.086	0.0857
-40	-40.142	0.1416
-40	-40.138	0.1379
-40	-40.142	0.1416
-40	-40.179	0.1788
-40	-40.183	0.1826
-50	-50.156	0.1565
-50	-50.171	0.1714
-50	-50.171	0.1714
-50	-50.130	0.1304
-50	-50.201	0.2012
<b>Total</b>		<b>0.0923</b>



SI. 4: Technical Repeatability of Z-axis repositioning. Five different Z-Axis values were set to the motion controller and approached five times with initial zeroing through limit switches each time. Actual height changes were recorded by the calibrated laser distance values.



SI. 5: Calibration of laser distance sensor. Linear regression of raw sensor values of the laser distance sensor. The reference height was determined with a caliper of a staircase shaped object (RGB and depth image in bottom left corner). The regression line is colored black, while the linear regression extrapolation is drawn dashed. Gray indicates confidence interval limits at  $\alpha = 0.95$ . Adj  $R^2$  denotes the coefficient of determination adjusted according to Yin and Fan (2001), while  $P_{slope}$  and  $P_{inter}$  represent p-values of the coefficients for the intercept and slope determined by simple T-test. MAE and RMSE indicate the mean absolute error and the root mean square error of calibration. n = 119

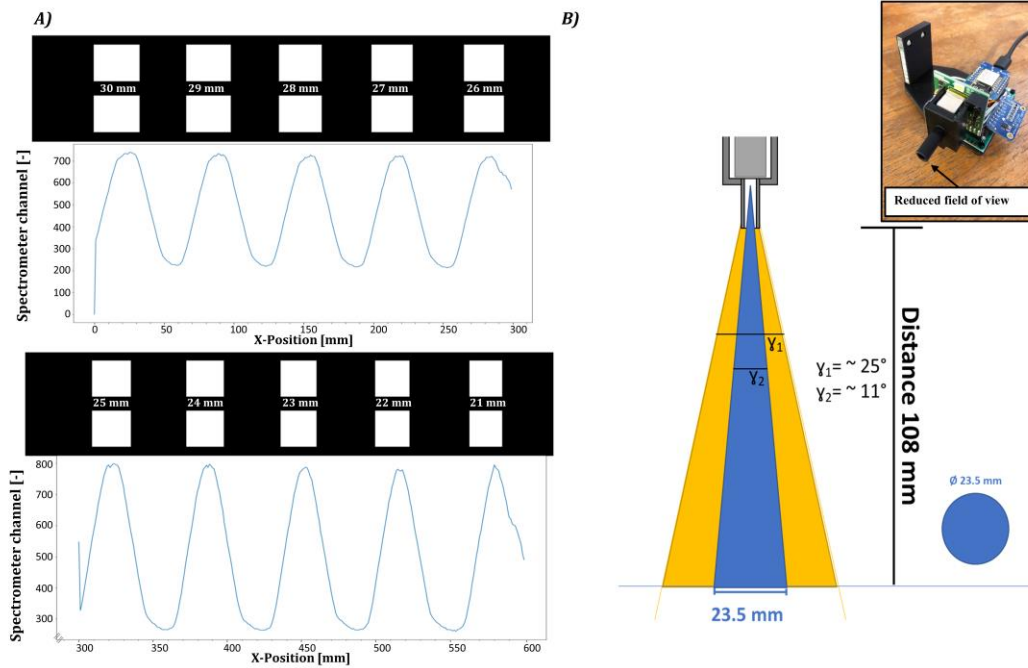
**SI. 6 - Table 2:** *Technical Repeatability of spatial scanning with laser distance sensor over time*

Day [d]	MAE [mm]	RMSE [mm]	Sample number [-]
1	0.31	0.32	2500
2	0.49	0.51	2500
3	0.48	0.5	2500
4	0.47	0.49	2500
5	0.37	0.4	2500
<b>Total</b>	<b>0.42</b>	<b>0.44</b>	<b>12500</b>

SI. 6: Technical Repeatability of spatial scanning with laser distance sensor over time. Determination of technical repeatability over time was conducted by measuring a reference object with a flat surface and a height of 41 mm once per day over 6 days, under the settings that were used in all experiments (Table 3). The initial depth measurement (Day 0) of an area of 50 mm × 50 mm was set as the reference for calculation of the mean absolute error (MAE) and the root mean square error (RMSE). The daily measurement procedure included an initial zeroing through limit switches, repositioning and depth data acquisition by spatial scan.

```
Model:
Parallel Random Forest (VIGRIA)
Selected Features:
Gaussian Smoothing ( $\sigma=0.3$ )
Gaussian Smoothing ( $\sigma=0.7$ )
Gaussian Smoothing ( $\sigma=1.0$ )
Gaussian Smoothing ( $\sigma=1.6$ )
Gaussian Smoothing ( $\sigma=3.5$ )
Gaussian Smoothing ( $\sigma=5.0$ )
Gaussian Smoothing ( $\sigma=10.0$ )
Gaussian Smoothing ( $\sigma=20.0$ )
Laplacian of Gaussian ( $\sigma=10.0$ )
Laplacian of Gaussian ( $\sigma=20.0$ )
Gaussian Gradient Magnitude ( $\sigma=10.0$ )
Gaussian Gradient Magnitude ( $\sigma=20.0$ )
Difference of Gaussians ( $\sigma=10.0$ )
Difference of Gaussians ( $\sigma=20.0$ )
```

SI. 7: Random forest classification model features for segmentation of *A. thaliana* trail A.



SI. 8: Experimental and graphical determination of modified spectrometer detection spot size. Image of modified spectrometer are shown in upper right corner. A) Experimental determination of spectrometer detection spot size by a sequential spectrometer readout every 1 mm, while linear movement in X-axis over a grid with black background and white squares of decreasing size and a side length ranging from 30 mm to 21 mm. Spectrometer channel readouts with the highest signal were picked from the array and plotted over the x axis. We assumed that if the detection spot size diameter is smaller than the side length of the square a constant plateau is found in the respective peak. The first square where a sharp maximum was identifiable, or in particular its side length of 23 mm determined the spot size diameter. B) Graphical estimation by drawing at a 1:1 scale. Graphical determination found a spot size diameter of 23.5 mm.

## 2.2 Towards automated detection of hyperhydricity in plant in vitro culture

Hans Bethge<sup>1,2</sup>, Zahra Mohammadi Nakhjiri<sup>2</sup>, Thomas Rath<sup>1</sup>, Traud Winkelmann<sup>2</sup>

<sup>1</sup>Laboratory for Biosystems Engineering, Faculty of Agricultural Sciences and Landscape Architecture, Osnabrück University of Applied Sciences, 49090 Osnabrück, Germany

<sup>2</sup>Institute of Horticultural Production Systems, Section of Woody Plant and Propagation Physiology, Leibniz Universität Hannover, Herrenhäuser Str. 2, 30419, Hannover, Germany

Type of authorship:	First author
Type of article:	Research article
Status of article:	Accepted/In press
Contribution to the article:	Designed and performed the experiments and analysed the data. Prepared the figures and wrote the manuscript.
Journal:	Plant Cell, Tissue and Organ Culture (PCTOC)
Impact factor:	2.726 (2022-2023)
DOI:	10.1007/s11240-023-02528-0
Acknowledgment:	Reproduced with permission from Springer Nature



# Towards automated detection of hyperhydricity in plant in vitro culture

Hans Bethge<sup>1,2</sup> · Zahra Mohammadi Nakhjiri<sup>2</sup> · Thomas Rath<sup>1</sup> · Traud Winkelmann<sup>2</sup> Received: 24 March 2023 / Accepted: 12 May 2023  
© The Author(s) 2023

## Abstract

Hyperhydricity (HH) is one of the most important physiological disorders that negatively affects various plant tissue culture techniques. The objective of this study was to characterize optical features to allow an automated detection of HH. For this purpose, HH was induced in two plant species, apple and *Arabidopsis thaliana*, and the severity was quantified based on visual scoring and determination of apoplastic liquid volume. The comparison between the HH score and the apoplastic liquid volume revealed a significant correlation, but different response dynamics. Corresponding leaf reflectance spectra were collected and different approaches of spectral analyses were evaluated for their ability to identify HH-specific wavelengths. Statistical analysis of raw spectra showed significantly lower reflection of hyperhydric leaves in the VIS, NIR and SWIR region. Application of the continuum removal hull method to raw spectra identified HH-specific absorption features over time and major absorption peaks at 980 nm, 1150 nm, 1400 nm, 1520 nm, 1780 nm and 1930 nm for the various conducted experiments. Machine learning (ML) model spot checking specified the support vector machine to be most suited for classification of hyperhydric explants, with a test accuracy of 85% outperforming traditional classification via vegetation index with 63% test accuracy and the other ML models tested. Investigations on the predictor importance revealed 1950 nm, 1445 nm in SWIR region and 415 nm in the VIS region to be most important for classification. The validity of the developed spectral classifier was tested on an available hyperspectral image acquisition in the SWIR-region.

## Key message

This study provides an approach that paves the way to automatic detection of hyperhydricity by identifying the key spectral features of this phenomenon.

**Keywords** Hyperhydricity · Spectral analysis · Phenotyping · Machine learning · Automated object detection

## Abbreviations

HH Hyperhydricity  
ML Machine learning  
UV Ultra violet

VIS Visible radiation  
NIR Near infrared radiation  
SWIR Shortwave infrared radiation  
MWIR Mid-wave infrared radiation  
LWIR Longwave infrared radiation  
DAT Days after treatment/transfer  
CV Cross validation  
CNN Convolutional neuronal network  
HSI Hyperspectral imaging

Communicated by Victor M. Jimenez.

✉ Hans Bethge  
bethge@baum.uni-hannover.de

<sup>1</sup> Laboratory for Biosystems Engineering, Faculty of Agricultural Sciences and Landscape Architecture, Osnabrück University of Applied Sciences, Oldenburger Landstraße 24, 49090 Osnabrück, Germany

<sup>2</sup> Institute of Horticultural Production Systems, Section of Woody Plant and Propagation Physiology, Leibniz Universität Hannover, Herrenhäuser Str. 2, 30419 Hannover, Germany

## Introduction

Hyperhydricity (HH) represents one of the major challenges for increasing the efficiency of plant in vitro propagation as it limits plant quality, adventitious root formation and ex vitro survival rate, in particular when



using liquid culture or bioreactor systems (Cardoso et al. 2018; Debergh et al. 1992; Gribble 1999). According to Kemat et al. (2020), at least 200 species are sensitive to HH and around 150 species can be affected seriously by HH emphasizing the relevance for commercial micropropagation. HH not only restricts the propagation of in vitro plants, but also affects the efficiency of genetic transformation mediated by *Agrobacterium* (van Altvorst et al. 1996) and the conservation of important species in germplasm banks (Lizárraga et al. 2017).

HH is a physiological disorder occurring under the specific conditions of plant tissue culture such as high humidity, high supplementation of sucrose, impaired gaseous exchange capacity and consequently low photosynthetic activity (George et al. 2008; Ziv 1991). The work of van den Dries et al. (2013) and Rojas-Martínez et al. (2010) provided strong evidence that the underlying mechanism of the HH etiology is the flooding of the apoplast, resulting in hypoxia and causing oxidative stress. This in turn leads to the macroscopic symptoms of water-soaked, wrinkled, curled, brittle and translucent tissue. The occurrence of HH was shown to be increased when the water availability for the in vitro explant was increased (Smith and Spomer 1995), e.g., by reduced concentration of the gelling agent (Ivanova and Van Staden 2011), or by the type of the gelling agent used (Pasqualetto et al. 1988; Tsay et al. 2006). The gelling agent gelrite induced HH in a wide range of plant genera (e.g., *Arabidopsis* sp., van den Dries et al. 2013, *Malus* sp. Pasqualetto et al. 1988, *Prunus* sp. Franck et al. 1998), even though the same gel strength as agar was used.

In addition to the anatomical changes of hyperhydric tissue which include larger intercellular spaces in the mesophyll and a drastically reduced number of palisade cells (Vieitez et al. 1985), several biochemical changes of hyperhydric tissue such as decreased chlorophyll contents (Phan and Letouze 1983; Franck et al. 1998), hypolignification (Kevers et al. 1987; Kemat et al. 2021), and high apoplastic water volume (Dries et al. 2013; Tian et al. 2015; de Klerk and Pramanik 2017) were reported. Paques et al. (1985) refer to HH as an inducible and reversible phenomenon and demonstrated that *Malus* sp. 'M26' plantlets could return to non-hyperhydric state if the induction phase in liquid culture did not exceed five days or if the symptoms of HH were not too severe. Recently, there were reports that hyperhydricity can be reversed by supplementation of agents to media such as silver nitrate and trichloroacetate (Gao et al. 2017; de Klerk and Pramanik 2017) or by controlling the environmental conditions in addition to media optimization (Mohamed et al. 2023), but no general countermeasure has been derived up to now. In commercial in vitro laboratories visual monitoring for contaminations and disorders are part of the routine work and therefore a costly and time-consuming repetitive matter (Mestre et al. 2017).

Nowadays, digitalization enters the horticultural sector, driven by digital solutions to increasingly complex work processes achieved through technological advances in sensors, automation and robotization, as well as data analysis through classical and advanced machine learning (ML) techniques. Automation of processes offers great economic potential for micropropagation laboratories since 60–70% of total costs of a micropropagated plant is due to manual labor (Chen 2016). An increasing number of reports on automating micropropagation processes such as explant cutting (Huang and Lee 2010), the commercial laser-based robotic cut and transplanting system RoBo@Cut (Bock Biosciences GmbH 2018), monitoring of cultures (Dhondt et al. 2014, Bethge et al. 2023) and transplanting of explants (Lee et al. 2019) were published within the recent years. In addition, there are several studies on the application of computer vision to micropropagation (Smith et al. 1989; Aynalem et al. 2006; Dhondt et al. 2014; Gupta and Karmakar 2017; Mestre et al. 2017) with imaging sensors being the crucial technology. Imaging sensors used in horticulture consist of affordable RGB cameras, multispectral cameras, thermal cameras, expensive hyperspectral imaging (HSI) systems, ToF (Time of Flight), LIDAR systems (Light Detection and Ranging) and more. The different sensor systems can be discriminated by their operating spectral range (UV, VIS, NIR, SWIR, MWIR, LWIR/Thermal-IR), spectral resolution from one (monochrome) to > 100 (hyperspectral) channels and cost of purchase. For example, the price of silicon (Si)-based hyperspectral cameras rise by a factor of 2 to 20 when switching the operating spectral range from VIS/NIR (400–1000 nm) to SWIR (900–1700 nm) with an Indium-Gallium-Arsenide (InGaAs) camera chip (Tisserand 2021). This needs to be considered, when selecting the appropriate spectral range and corresponding imaging technology. While computer vision coupled with ML offers already great potential to solve complex detection task in agriculture (reviewed in Patrício and Rieder 2018), for application in plant tissue culture only few reports are available up to now (reviewed in Prasad and Gupta 2008; Hesami and Jones 2020). However, these are limited in terms of live-monitoring, since they followed the “object to sensor” approach for plantlet clustering (Mahendra et al. 2004), classification of somatic embryos (Zhang et al. 1999) and estimation of shoot length (Honda et al. 1997).

The visual appearance of plants, and in particular leaf pigments, can be estimated by spectroscopic approaches based on their interaction with electromagnetic radiation. Single biochemical plant metabolites can be associated with specific wavelengths based on their major absorption peaks (Table 1).

Univariate data analysis, e.g., spectral indices or multivariate data analyses like partial least square (PLS), allows the prediction of leaf pigments' concentrations and can be

**Table 1** Selected reported symptoms of hyperhydric tissues (HH) and corresponding expected major changes in optical absorbance features

Reference	Plant species	Observation	Deduced optical absorbance features in VIS-SWIR [nm]*
Phan and Letouze (1983)	<i>P. avium</i>	Lower chlorophyll content in HH	<b>430</b> , 460, <b>640</b> , 660
Van den Dries et al. (2013)	<i>A. thaliana</i>	Higher apoplastic water volume in HH	<b>970</b> , <b>1200</b> , 1400, <b>1450</b> , <b>1940</b>
Phan and Letouze (1983)	<i>P. avium</i>	Less protein content in HH	910, 1020, <b>1510</b> , 1940, 1980
Kemat et al. (2021)	<i>A. thaliana</i>	Hypolignification in HH	<b>1200</b> , <b>1420</b> , 1450, <b>1690</b> , 1940
Saher et al. (2005)	<i>D. caryophyllus</i>	Higher sugar content in HH	<b>1450</b> , 1490, 1580, <b>1780</b> , <b>1960</b>
Van den Dries et al. (2013)	<i>A. thaliana</i>	Anthocyanins accumulation in HH	<b>550</b>

\*Absorptions peaks according to Curran (1989) in a wavelength range of 400 to 2000 nm. Bold wavelength indicating stronger absorption of the respective chemical compound

used for classification. These techniques also enable the discrimination of different plant species or the identification of growth anomalies by specific spectral features (Shaw and Kelley 2005). According to Hesami and Jones (2020), ML techniques applied to plant tissue culture problems will help in future to solve classification and regression problems and can be employed for automation and mechanization of in vitro propagation, genetic engineering and genome editing technologies. In addition, Nezami-Alanagh et al. (2019) demonstrated the positive impact of ML models in optimizing culture media in terms of time, cost and the occurrence of physiological disorders in the propagation of pistachio rootstocks. Prasad and Gupta (2008) proposed that an automated decision-making system based on computer vision coupled with ML models and combined with a robotic system will result in the mechanization of commercial mass propagation and help in evaluating various aspects of plant quality such as HH status, which might be difficult to determine by human visual inspection. To our knowledge, the spectral properties of HH have not yet been studied or used as a distinguishing feature for ML classification of in vitro cultured explants.

The objective of this study was to investigate the spectral fingerprints of hyperhydric tissue in two different plant species (*Malus* sp. and *Arabidopsis thaliana*) after forced induction of the growth anomaly and subsequent spectral analysis of the explants. Here, we selected *Malus* as a representative of classical in vitro shoot cultures and *Arabidopsis* as a model plant for the underlying mechanism of HH. A novel phenotyping system was tested to monitor the morphological characteristics of hyperhydric explants in time-series image data. Furthermore, we aimed at identifying specific absorption features of hyperhydric tissues that are sufficient for discrimination by ML techniques and to locate them within in the electromagnetic radiation spectrum. Putative discriminating models should be validated and discussed in terms of their feasibility in plant tissue culture. The findings of this study should pave the way for an automatic detection of HH by live-monitoring of in vitro cultures.

## Material and methods

### Plant material and experimental setup

#### Morphological characteristics of hyperhydricity

From in vitro apple shoot cultures (*Malus* sp. ‘G214’) uniform shoots of 10–15 mm length were prepared and cultivated on modified MS medium (Murashige and Skoog 1962) containing 2.2 µM 6-benzylaminopurine (BAP), 0.5 µM indole-3-butyric acid (IBA), 3% (w/v) sucrose and solidified with either 0.8% (w/v) agar (Plant agar, Duchefa, Haarlem, The Netherlands) for the control variant (“MS + agar”) or with 0.25% (w/v) gelrite (Duchefa, Haarlem, The Netherlands) for the HH induction variant (“MS + gelrite”). The pH of the medium was adjusted to 5.8 prior to autoclaving at 121 °C for 15 min.

*Arabidopsis thaliana* ‘Col-0’ seeds which had been stored at 4 °C, were surface-disinfected using 70% (v/v) isopropanol for 30 s, followed by 2% (v/v) sodium hypochlorite plus Tween 20 for 5 min and then rinsed thoroughly three times using sterile deionized water. The seeds were germinated for 10 days at 24 °C in 9 cm-Petri dishes (polystyrene) on modified plant growth regulator-free B5 medium (Gamborg et al. 1968), containing 1.5% (w/v) sucrose with 0.8% (w/v) Plant agar and pH 5.8. Uniform 10 day-old seedlings were selected and five seedlings per 500 mL-vessel were transferred to modified plant growth regulator-free B5 medium (Gamborg et al. 1968), containing 1.5% (w/v) sucrose and either 0.8% (w/v) Plant agar for the control variant (“B5 + agar”) or 0.25% (w/v) gelrite (“B5 + gelrite”) to induce HH. The pH of the medium was adjusted to 5.8 prior to autoclaving at 121 °C for 15 min.

Ten 500 mL polypropylene vessels were prepared for Experiment I (Table 2) and Experiment II, each with four plantlets and containing ~ 80 mL of one of the two different media (“B5/MS + agar”/“B5/MS + gelrite”) supplemented with 1 g L<sup>-1</sup> titanium dioxide. Titanium dioxide (food dye; Ruth GmbH & Co.KG, Bochum, Germany)

**Table 2** Overview of conducted experiments and measurements ten

Experiment	Plant species	Time series [day]	Evaluations	Determination/Device
I	<i>Arabidopsis thaliana</i>	0–20	RGB growth curve	RGB image sensor of Phenomenon <sup>a</sup>
			RGB shape analysis	RGB image sensor of Phenomenon <sup>a</sup>
			Depth mean canopy height	Laser distance sensor of Phenomenon <sup>a</sup>
			Depth maximum plant height	Laser distance sensor of Phenomenon <sup>a</sup>
II	<i>Malus sp.</i>	0–27	RGB growth curve	RGB image sensor of Phenomenon <sup>a</sup>
			RGB shape analysis	RGB image sensor of Phenomenon <sup>a</sup>
			RGB image data set	RGB image sensor of Phenomenon <sup>a</sup>
			Depth mean canopy height	Laser distance sensor of Phenomenon <sup>a</sup>
			Depth maximum plant height	Laser distance sensor of Phenomenon <sup>a</sup>
III	<i>Malus sp.</i> , <i>Arabidopsis thaliana</i>	0, 5, 10, 15, 20	HH Score	Visual scoring
			Apoplastic liquid volume	Apoplastic liquid volume
			Reflection spectra	UV–VIS Spectrometer Perkin-Elmer
IV	<i>Malus sp.</i>	14, 21, 28	HH Score	Visual scoring
			Apoplastic liquid volume	Apoplastic liquid volume
			Reflection spectra	UV–VIS Spectrometer Perkin-Elmer
V	<i>Malus sp.</i>	0, 4, 8, 12, 16	HH Score	Visual scoring
			Apoplastic liquid volume	Apoplastic liquid volume
			Reflection spectra	UV–VIS Spectrometer Perkin-Elmer

<sup>a</sup>The multisensory robot system “Phenomenon” developed by Bethge et al. (2023), consisting of 4 sensors (RGB camera, laser distance sensor, thermal camera and a microspectrometer), was used to enable in-situ measurement of the morphology through the lid of the culture vessels

was used to add a white color to the medium, because this enabled the height measurements of the robot system due to increased reflection of the culture media. A plastic film (PVC system foil; Klarsichtpackung GmbH, Hofheim, Germany) sealed the containers as a substitution of the lid of the containers to provide a fully transparent view while ensuring the aseptic condition of the cultures. These cultures were cultivated at 22 °C with a 16 h photoperiod and under a PPF (Photosynthetic Photon Flux Density) of 35–40 μmol m<sup>-2</sup> s<sup>-1</sup>, provided by two tubular fluorescent lamps (Philips MASTER TL-D 58W/865). The lab’s bottom-cooling system—provided by water-cooled plastic tubes below the shelf—prevented water condensation due to a local shift of dew point. Room temperature ranged from 19 (night) to 25 °C (day) with an average of 22 °C over 24 h, while the average surface temperature of the cooled cultivation area ranged from 19 (night) to 24 °C (day) with an average of 21 °C over 24 h. In addition to the non-destructive monitoring approach (Exp. I & II), three experiments (Exp. III, Exp. IV, Exp. V; Table 2) were conducted with different evaluation time points. The evaluation time points were chosen based on the key events in the dynamic etiology of hyperhydricity during a culture passage (~4–5 weeks for *Malus*). Important morphological changes were observed during the first two weeks, so Exp. III and V covered this time span, while measurements in Exp. IV were undertaken to cover the second half of the culture passage.

### Hyperhydricity induction

For *Malus* shoot cultures, 500 mL polypropylene containers containing 80 mL of the two different media were used and each container was inoculated with five shoots. Cultivation took place for 20 days (Table 2; Exp. III), 28 days (Exp. IV) and 16 days (Exp. V) at 22 °C (room temperature ranged from 19 (night) to 25 °C (day) with an average of 22 °C over 24 h) with a 16 h photoperiod and under a PPF (Photosynthetic Photon Flux Density) of 35–40 μmol m<sup>-2</sup> s<sup>-1</sup>, provided by tubular fluorescent lamps (Philips MASTER TL-D 58W/865). *Arabidopsis* plantlets were cultivated as described above for 20 days (Exp. III).

### Evaluations

#### Morphological characteristics of hyperhydricity via image analysis

For visualization of the etiology of HH, the multisensory robot system “Phenomenon” (Bethge et al. 2023) was used. RGB images were captured in Exp. I and Exp. II every 4 h with a 12.3-megapixel RGB camera (Raspberry Pi Camera HQ, Raspberry Pi Foundation, Cambridge, UK) equipped with a 6 mm fixed focal length low-distortion lens (Edmund Optics: 6 mm wide angle lens, f/1.2, high resolution = 120 lp mm<sup>-1</sup> (lp = line pairs), low distortion < 0.5%) and with the following camera parameters: resolution = 4054 px × 3040

px, shutter speed = 2000 ms, iso = 100, autowhite-balance = off and a fixed gain of 3.3, 1.5 (red, blue).

Sensor data from the multisensory robot system “Phenomenon” (Bethge et al. 2023) were processed, segmented, and various parameters were calculated. RGB image analysis was performed in Python (Van Rossum and Drake 2009), using the following packages: *OpenCV* v3.4.9 (Bradski 2000), *NumPy* v1.20.2 (Van Der Walt et al. 2011) and *PlantCv* v3.11.0 (Gehan et al. 2017) and the Software toolkit *Ilastik* v1.3.3 (Sommer et al. 2011) headless integrated in the Python script. RGB image analysis included a histogram stretching for normalization, segmentation via a trained random forest classifier, normalization to the day 0 plant area and calculation of projected plant area (37.7 px = 1 mm). Shape analyses were performed on the four largest objects by area and limited to the first nine days to avoid errors from overlapping explants. We used the installed shape function of *PlantCv* to calculate solidity (measure of density as the ratio between object area and area of the convex hull of the object) and eccentricity (measure of deviation of an ellipse to a circle (eccentricity = 0) as the ratio between major and minor axis).

Depth data were acquired once per day for each culture container with the point-measuring laser distance sensor as a spatial scan by sequential readout of the sensor while shifting the detector head of the “Phenomenon” robot system in xy direction, according to the scan pattern (100 mm × 100 mm; with a resolution of 1 mm × 1 mm). The laser distance sensor (OD-Mini OB1-B100, Sick AG, Waldkirch, Germany) used in this setup was specified by the manufacturer with a power consumption of < 1.92 W, laser emission wavelength of 655 nm, max. output of 390 μW (laser class 1), a measuring range of 50 to 150 mm and a linearity of ± 100 μm as well as spot size of 700 μm × 600 μm at a measuring distance of 100 mm. The analog output of the laser distance sensor (10 V) was connected via a small voltage divider circuit to a high precision 16-bit A/D-converter (ADS 1115), which communicated via Inter-Integral Circuit (I<sup>2</sup>C) with a microcontroller board (Wemos D1 Mini). Each distance measurement consisted of a up to ten single readouts and averaging (excluding default sensor values), to achieve a robust and low-noise measurement. A detailed description of the robot system “Phenomenon” can be found in Bethge et al. (2023).

Depth data of explants were obtained by measuring 10,000 data points of each culture vessel once a day with a scanning laser distance sensor. The depth data processing pipeline included the segmentation of culture media by a RANSAC (random sample consensus, Fischler and Bolles 1981) segmentation approach, subtraction of RANSAC plane, normalization to the day 0 plant height with the *Python* libraries: *Open3D* v0.15.1 (Zhou et al. 2018) and *Pyvista* v0.34.0 (Sullivan et al. 2019). Pipelines construction is described in detail in Bethge et al. (2023).

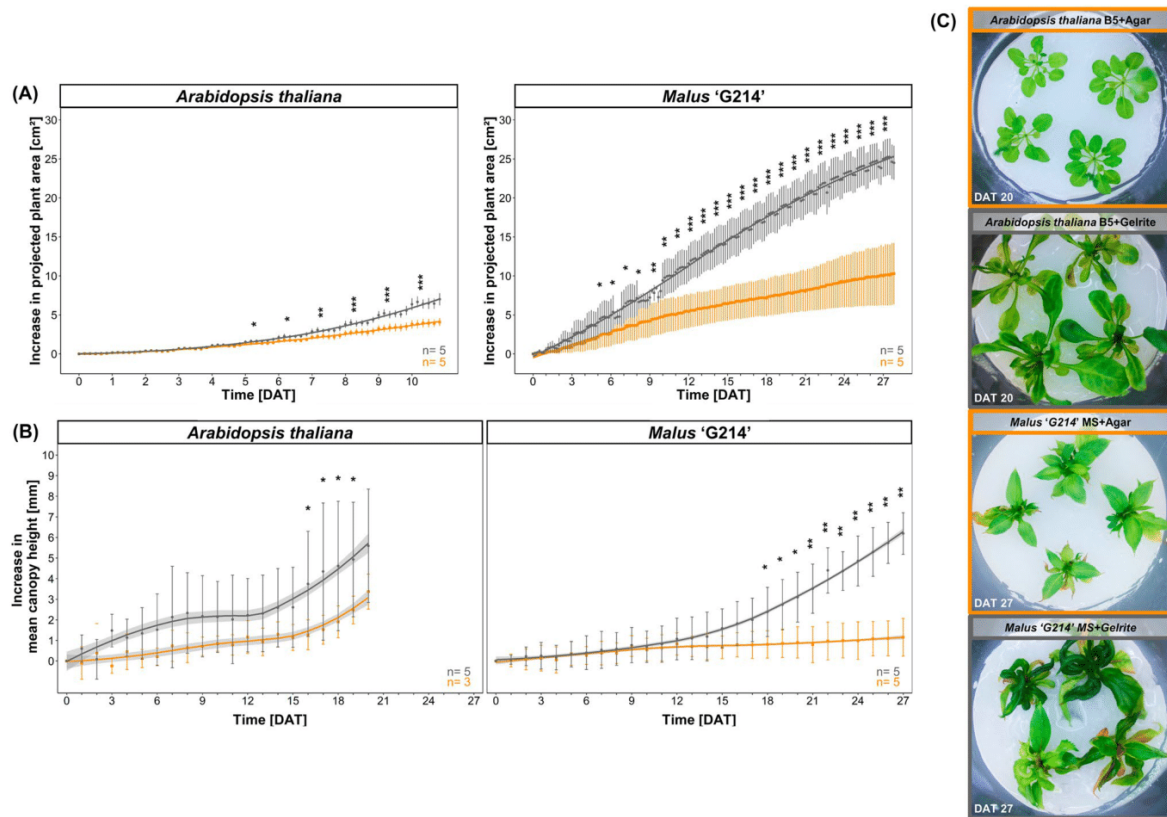
Statistical analysis of repeated measures data was performed using R software. Data were transformed, if necessary, with the R package *bestNormalize* v1.8.2 (Peterson and Peterson 2020). Different linear mixed-effect models from *nlme* v3.1-153 package (Pinheiro et al. 2017) were fitted to the data with different covariance structures: scaled identity, first-order autoregressive, first-order heterogeneous autoregressive, compound symmetry, Toeplitz and heterogeneous Toeplitz. The mean model consisted of the fixed effects treatment/medium type and time and their interaction terms. An extra random effect was included in the model to account for the dependencies between measurements from the same culture container or in SI. 1 for shape analysis from the same explant (as nested random effect). We also included linear models with random intercept (CulturecontainerID) and random slope (Time). The respective best model (Fig. 1A: linear mixed model with scaled identity covariance structure and random slope; Fig. 1B and SI. 1A: linear mixed model with heterogeneous Toeplitz covariance structure and random slope; SI. 1B: linear mixed model with heterogeneous Toeplitz covariance structure and random intercept) was selected based on the Akaike Information Criterion (AIC, Sakamoto et al. 1986) values and residual analysis (QQ-plot). Pairwise comparisons using Tukey’s HSD test at  $p < 0.05$  was performed and show significant differences between treatments within a time point.

#### Visual scoring of hyperhydricity severity level

In the Experiments III to V, the severity of HH was assessed for each explant and at every time point (Exp. III: 0, 5, 10, 15, 20 days; Exp. IV: 14, 21, 28 days; Exp. V: 0, 4, 8, 12, 16 days) according to Tian et al. (2015) with minor modifications (Table 3). The starting plant material cultivated on control media represented the samples of 0 days after transfer (DAT 0).

#### Determination of apoplastic liquid volume

Per time point at least 10 samples per treatment were collected for the determination of apoplastic liquid volume, with DAT 0 samples representing the starting material. Apoplastic liquid was extracted from leaf tissue by mild centrifugation according to van den Dries et al. (2013) and Terry and Bonner (1980): Leaves (50–150 mg FM) from a single explant were excised, weighed, and placed into a 2 mL tube microcentrifuge filter without membrane (Clear-Line®; Kisker Biotech GmbH & Co, Steinfurt, Germany). Samples were centrifuged at 3000 g for 20 min at 4 °C. Immediately after centrifugation, the leaves were reweighed to determine the apoplastic liquid volume ( $V_{AL}$ ) in  $\mu\text{L g}^{-1}$  fresh mass (FM) using the Eq. 1.



**Fig. 1** Morphological differences in growth patterns of explants of *A. thaliana* Col-0 and *Malus 'G214'* cultivated on either agar or gelrite solidified media (Mean ± SD). **A** The curve for the increase in the projected plant area was calculated from the analysis of the segmented RGB images normalized to the plant area of day 0 and presented as projected plant area [cm<sup>2</sup>]. Since flower initiation started at later time points for *A. thaliana* and thus an error in the estimation of projected plant area might occur, the analysis of growth curves was limited to the first ten days. **B** The relative increase in mean canopy height resulted from analysis of segmented depth data collected with a scanning laser distance sensor and normalized to day 0 plant height.

Yellow lines indicate cultivation on standard media formulation on Gamborg-B5 (*A. thaliana*) and MS-Medium (*Malus*) solidified with 0.8% agar (w/v), while dark gray lines display the cultivation on induction media containing 0.25% (w/v) gelrite, inducing HH. **C** Representative images at the endpoint of the experiments. Sample number (n) indicates the individual culture containers. Significance stars indicate comparisons of treatments within a time point (day) with \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001. RGB and depth data were acquired with the multisensory robot system “Phenomenon” (Bethge et al. 2023). (Color figure online)

**Table 3** Scoring of hyperhydricity by visual observation (Tian et al. 2015, with minor modifications)

Hyperhydricity score	Symptoms
0	No visual symptoms
1	≤ 50% curled leaves
2	> 50% curled leaves
3	> 50% curled and thickened leaves
4	Curled, thickened, translucent, fragile leaves

$$V_{AL} = \frac{(FM - M_{ac}) \cdot \rho_{H_2O}}{FM} \quad (1)$$

where FM = fresh mass of leaves in mg, M<sub>ac</sub> = mass of leaves after centrifugation and ρ<sub>H<sub>2</sub>O</sub> = water density (the water density was taken as equal to 1 g mL<sup>-1</sup> assuming the apoplastic liquid is mainly water and has a temperature of 4 °C).

### Spectral data acquisition and analysis

Prior to the quantification of apoplastic liquid volume, one fully expanded leaf per explant under study was collected. The leaf was then placed in a 3D printed sample holder (SI. 2) in an adaxial position that allowed for flat clamping

without exerting too much pressure on the leaf (with a cavity of 1 mm). The curled hyperhydric leaves were handled with care to obtain reflection spectra from a planar surface. The leaf reflectance spectra were examined with a Perkin-Elmer Lambda 900 UV–VIS–NIR–SWIR spectrometer (Perkin-Elmer Instruments, Norwalk, USA) equipped with 150 mm Indium–Gallium–Arsenide (InGaAs) integrating sphere. The reflectance intensity was measured in steps of 1 nm in the wavelength range between 200 and 2000 nm, and the reflectance was calculated using the reflection spectrum of the white reference standard Spectralon®. Raw spectra were pre-processed in R *v4.1.2* using Rstudio (RStudio Team 2015) with the *hsdar v1.0.4* package (Lehnert et al. 2018) allowing the cleaning of device errors, trimming to spectral range of 400 nm to 2000 nm and smoothing with the Savitzky-Golay filter at a window size of 25 data points of third-degree polynomials to remove noise from data.

Spectra of leaves obtained from three experiments (Exp. III: 147, Exp. IV: 39 spectra, Exp. V: 51) were divided into two groups based on the significance level of the apoplastic liquid volume and the HH score of the whole explant was assessed by visual observation. Here, the explants with a HH score of 0 and 1 were classified as normal explants while the explants with a HH score of 2 to 4 represented hyperhydric explants. This resulted in 100 and 137 spectra of normal and hyperhydric leaves, respectively, covering the two plant species *Malus* ‘G214’ (187 spectra) and *A. thaliana* (50 spectra). For visualization and isolation of the HH-specific absorption features, leaf spectra were further processed with the segmented upper hull continuum removal method described in detail in Lehnert et al. (2018). This normalization method allowed a comparison of individual absorption features on a common baseline formed by a segmented upper hull of local maxima and resulted in absorption features spectra. In addition, difference spectra of absorption feature spectra were calculated by subtracting normal leaf spectra from hyperhydric leaf spectra.

In addition, we defined three spectral ranges based on the sensitivity of the state-of-the-art sensor technologies such as standard RGB camera systems with silicon sensor chips (3 channels: B: 400 nm to 500 nm, G: 500 nm to 600 nm and R: 600 nm to 700 nm), multispectral camera systems with silicon sensor chips (4 channels: B: 400 nm to 500 nm, G: 500 nm to 600 nm, R: 600 nm to 700 nm and NIR: 750 nm to 850 nm) and SWIR-HSI camera systems with Indium–Gallium–Arsenide (InGaAs) sensor chips (SWIR: 900 nm to 1700 nm). This division was made as a decision support for assessing the potential of the candidate detection systems to detect HH based on their spectral sensitivity range and considering their affordability.

### Identification of hyperhydricity-specific absorption features

Different ML models were trained with the *caret v6.0-90* package (Kuhn 2008) in the R software to identify the key absorption features that discriminate between normal and hyperhydric explant leaf spectra. Here, pre-processed spectral data sets (237) were centered and scaled and divided into a training set (178 spectra; *Malus*: 143, *A. thaliana*: 35, with 103 normal and 75 hyperhydric explants, in total) and a test set (59 spectra; *Malus*: 44, *A. thaliana*: 15, with a total of 34 normal and 25 hyperhydric explants). All classification models were trained with the same resampling procedure consisting of a 10 times tenfold repeated cross validation (CV). The tenfold repeated CV divides the training data into 10 equal parts (10 subsamples with a size of 178/10). These parts are iterated 10 times, during each iteration, 9 of the 10 parts serve as training data, and the remaining 10th part as the validation set to calculate model performance metrics. In 10 times repeated tenfold CV this process is repeated 10 times; therefore, performance of training was validated on 100 validation subsamples consisting of 17–18 individual spectra.

In the confusion metrics, correctly classified normal and hyperhydric leaves formed the true-positive (TP) and the true-negative (TN) class, while false classified ones constituted the false-positive (FP) and false-negative (FN) class, respectively. For evaluation of model validation performance, the sensitivity (Eq. 2; TPR: true positive rate) and the specificity (Eq. 3; TNR: true negative rate) were calculated with normal explants as the positive class and the area under the curve (AUC) of the receiver-operator-characteristics (Eq. 4;  $AUC_{ROC}$ ), while for evaluation of model test performance, the accuracy (Eq. 5) was determined. Here, misclassifications are described by the false negative rate (FNR) and false positive rate (FPR). Balanced accuracy (Eq. 6) and  $F_1$  score (Eq. 7) were calculated to account for putative class imbalances. To find the best suitable model for discriminating between normal and hyperhydric leaf spectra, different ML model structures were tested, including a neuronal net with the maximum allowable number of weights set to 2000 (“nnet” from *nnet v7.3-16* package; Ripley et al. 2016), a linear discriminate analysis (“lda” from *caret* package), a supported vector machine (“svmLinear” from *caret* package), a random forest (“rf” from *caret* package), a high dimensional discriminate analysis (“hdda” from *caret* package) as well as a linear discriminate analysis (“lda” from *caret* package with PCA-preprocessed data set) with an upstream principal component analysis (PCA). Based on their resampled performance metrics, the best model was selected to identify its most relevant features/wavelengths on the basis of the underlying variable importance in the model.

$$\text{Sensitivity} = \text{TPR} = 1 - \text{FNR} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{Specificity} = \text{TNR} = 1 - \text{FPR} = \frac{TN}{TN + FP} \quad (3)$$

$$AUC_{ROC} = \int TPR(FPR) d(FPR) \quad (4)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (5)$$

$$\text{Balanced accuracy} = \frac{\text{TPR} + \text{TNR}}{2} \quad (6)$$

$$F_1 \text{ score} = 2 \times \frac{\left(\frac{TP}{TP+FP}\right) \times TPR}{\left(\frac{TP}{TP+FP}\right) + TPR} \quad (7)$$

### Automated hyperhydrycity detection

To test the validity of the developed spectral classifier, an HSI-system operating in the shortwave infrared (SWIR) region was used to acquire a single HSI data cube from a culture vessel containing a HH-sensitive apple genotype (*Malus* ‘Selection 4’). The imaging system that was developed and described by Thiel (2018) consisted of an EVK Helios Core NIR Line-scan camera (240 px × 1 px and 252 spectral channels in the wavelength region of 900 nm to 1700 nm), two 65 W halogen spot lights and a conveyer-belt system to move the sample. Image acquisition was performed in closed polypropylene culture vessels, so that sterile conditions could be maintained inside the vessel and water condensation was prevented by heat radiation from the halogen lamps. Since only one HSI data cube could be acquired, these results were considered to be an exemplary and preliminary validation test.

The developed spectral classifier was retrained with a reduced number of features to match the spectral channels of the imaging system (Features/wavelengths: 252 channels in the range between 900 and 1700 nm). Due to the binary classification output of the classifier, most of the background pixels were removed by creation of a binary mask with simple thresholding of the image slice at a wavelength of 1000 nm. Then each pixel of the segmented hyperspectral data cube was inserted as an input to the spectral classifier and class membership was predicted.

As a more affordable approach and as a proof of concept, an object detection model based on annotated RGB images acquired by the robot system was trained. Therefore, 250 images were randomly selected and annotated with the

graphical user interface Roboflow<sup>®</sup> (Dwyer et al. 2022). The image data set consisted of 200 annotated images of eight culture containers from Experiment II and 50 images from a comparable experiment to increase variance in the number of explants, background colour, and colour of culture media. A total of 504 normal explants and 545 hyperhydryc explants were included. The image data set was divided into 175 images as training set, 50 images as validation set and 25 images as test set. Data augmentation of annotated bounding boxes increased the training set to 1800 images and included: horizontal and vertical flip, rotation by 90° (clockwise, counter-clockwise, upside down), rotation by ± 5°, brightness by ± 10%, exposure by ± 7%, blur with 2px and noise with 2% of pixels. The data set (Bethge 2023) is publicly accessible via Roboflow<sup>®</sup> universe. Time series images of two culture vessels from Experiment II were retained and used to visualize the trained model. Object detection models perform attempts to identify and locate objects in images while assigning them to the appropriate classes. We selected YOLOv8 (Jocher et al. 2023) architecture as the latest versions of the YOLO (“You only look once”, Redmond et al. 2016) family. YOLO is a single-stage object detector, consisting of three parts in its architecture: backbone, neck and head. The backbone is defined by several convolutional layers which extract key features from the images, the neck uses the features and forms the feature pyramid by fully connected layers and the head is the final output layer for prediction of bounding boxes and classification. The training process was performed in the Google Colaboratory (Colab/Colab Pro) environment on a NVIDIA A100-SXM4-40 GB graphical processing unit (GPU) serviced by Google. In addition, the model was trained with the following parameters: epochs = 250 (early stopping occurred after 188 epochs), batch size = 16 images, image size = 640 px, patience = 100 epochs, learning rate = 0.01, momentum = 0.94, intersection over union (IoU) = 0.7. We let Roboflow train two object detection models, one from scratch and one with weights from a previously trained model (additional 125 images from the same experiment) to see the full potential of the dataset with the optimized pipeline. Evaluation of model performance was based on precision (Eq. 8), recall (Eq. 9), average precision (AP; Eq. 10) and mean average precision (mAP; Eq. 11) of the validation set. Here true positive (TP) indicate a correct detection and classification, false negative (FN) describes cases where the prediction missed the detection contained in the ground truth data, while in a false positive (FP) case a bounding box was predicted on a location not contained in the ground truth data. Thereby, AP represents the area under the precision-recall-curve across a range of probability confidence threshold values from 0 to 1. The mAP is the sum of AP of each class (k) divided by the number of classes (n) at a given intersection over union (IoU) threshold of 0.5. Intersection

over union is defined as ratio between the overlap area to the united area of the predicted and ground truth bounding box. After the training process predications were obtained using the Python library *roboflow v0.2.25* (Dwyer et al. 2021) with IoU threshold and confidence threshold set to 0.5.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (8)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (9)$$

$$AP = \int_0^1 \text{Precision}(\text{Recall}) d(\text{Recall}) \quad (10)$$

$$\text{mAP} = \frac{1}{n} \sum_{k=1}^{k=n} AP(k) \quad (11)$$

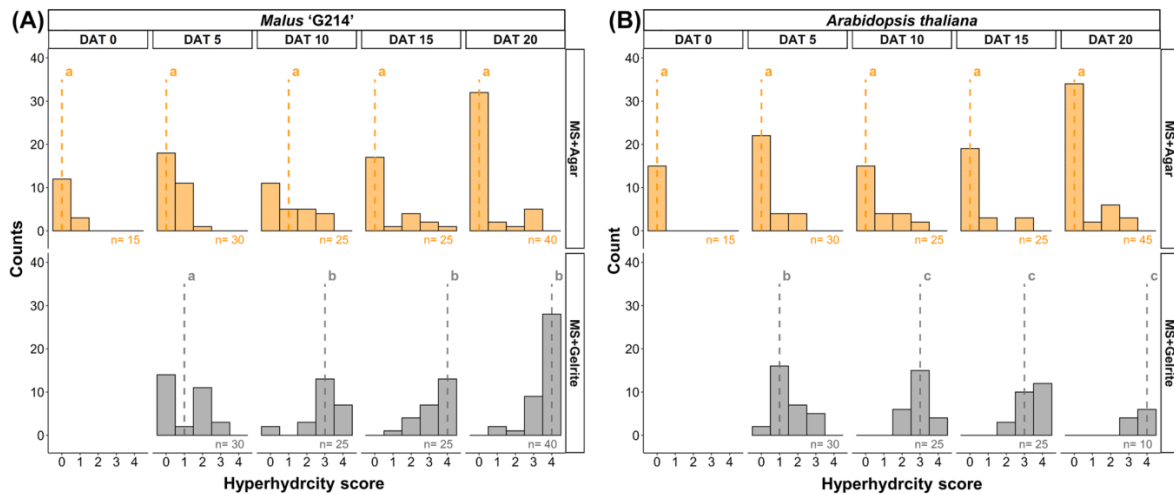
### Morphological characteristics of hyperhydricity via image analysis

Studying the morphology of the shoots of the two treatments revealed major differences in horizontal and vertical growth. Significantly stronger growth, quantified as projected plant area, was observed for the gelrite treatment at early time

points (5 days) for both plant species (Fig. 1A). After 4 weeks of cultivation, shoots of *Malus* in culture vessels with gelrite medium had with 24.5 cm<sup>2</sup> a 2.4 times greater increase in projected plant area than shoots in vessels with agar medium with 10.3 cm<sup>2</sup>. Here, 65% of the explants of *Malus* had a HH score > 2 in the gelrite treatment compared to 0% for agar treatment. For *A. thaliana* we evaluated the projected plant area only until day 10 to avoid distorting effects on projected plant area due to flower initiation starting at day 12. Shape analysis of single explants showed significant differences in solidity at day 3 and in eccentricity at day 6 for *A. thaliana*, whereas the shape differences of *Malus* explants were not significant (SI. 1). Vertical growth analysis, quantified as mean canopy height (Fig. 1B) and maximum shoot height as mean of upper 10<sup>th</sup> percentile (SI. 1), showed a significantly higher mean canopy height of *Malus* for the gelrite treatment at day 18 and of *A. thaliana* at day 16. An even earlier distinction was recorded for the maximum shoot height, i.e. at day 11 and 14 for *A. thaliana* and *Malus*, respectively.

### Hyperhydricity induction

Visual scoring of HH revealed the dynamics of HH induction using gelrite in the two plant species under investigation. Anthocyanin accumulation was noted within the first 4 days in both treatments for *Malus*. However, it persisted



**Fig. 2** Visual scoring of hyperhydricity of **A** *Malus* ‘G214’ and **B** *A. thaliana* Col-0 in vitro cultures over 20 days (DAT, Days After Treatment). Samples from 0 days after transfer (DAT 0) represent the starting plant material cultured on control media. Yellow bars indicate cultivation on standard media formulation **A** MS-Medium, **B** Gamborg-B5 solidified with 0.8% (w/v) agar, while gray bars display the cultivation on induction media containing 0.25% (w/v) gelrite. Dashed lines represent the medians of each histogram. Sample

number (n) indicates the individual explants. The different sample numbers result from the combined evaluation with different methods (apoplastic liquid evaluation, reflection spectroscopy) of the same samples. Different letters resulting from Kruskal–Wallis test followed by Fisher’s LSD ( $p < 0.05$ ) indicate significant differences between histograms. Kruskal–Wallis effect size could be determined to be very strong with **A**  $\eta^2 = 0.62$  and **B**  $\eta^2 = 0.65$ . (Color figure online)



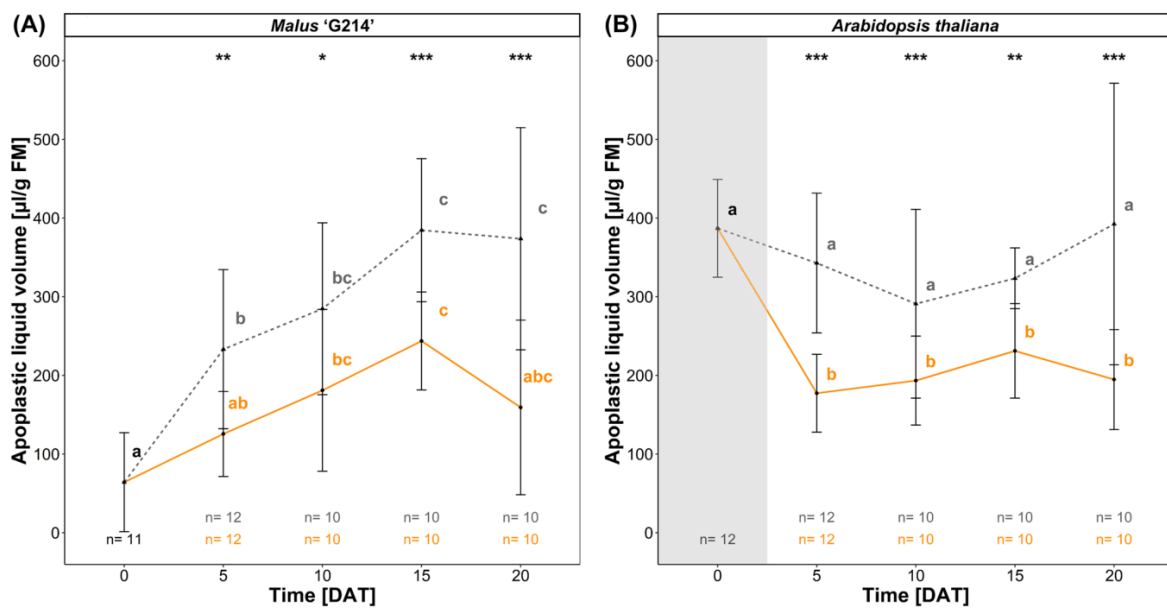
only in the gelrite treatment until the end of the experiment in most explants. In *Malus*, severe symptoms of HH were induced even on the agar control medium in 12.5% of the shoots (Fig. 2A). In two experiments, significant differences in the HH score and the occurrence of severe symptoms (curled, thickened and translucent leaves = level 4 of the HH score) between the agar control and the gelrite induction treatment were identified 10 days (Fig. 2A) and 8 days (SI. 3) after transfer. When performing this experiment under a novel phenotyping system, time-lapse videos were taken. They confirmed these observations and visualized the temporal development of HH in the two plant species (*Malus*: SI. 4 and *A. thaliana*: SI. 5). Three of four apple shoots turned into a hyperhydric status and formed first hyperhydric leaves (SI. 4: arrows) on gelrite at 5 DAT (SI. 4: 100 h). They started to curl at 8 DAT and also became much larger than those on agar. After 27 days, severe symptoms appeared on dark green to reddish explants that exhibited compact growth with curled, epinastic, and brittle leaves.

For *A. thaliana*, there was already a significant increase in the HH score after 5 days of treatment (Fig. 2B). Furthermore, decolorization of leaves was the predominating symptom of HH in *A. thaliana* on gelrite induction medium. For *A. thaliana* seedlings, first signs of HH (SI. 5: arrows)

became visible at 5 DAT (SI. 5: 100 h) on gelrite-solidified medium and shoots developed longer petioles and much larger leaves with severe HH symptoms.

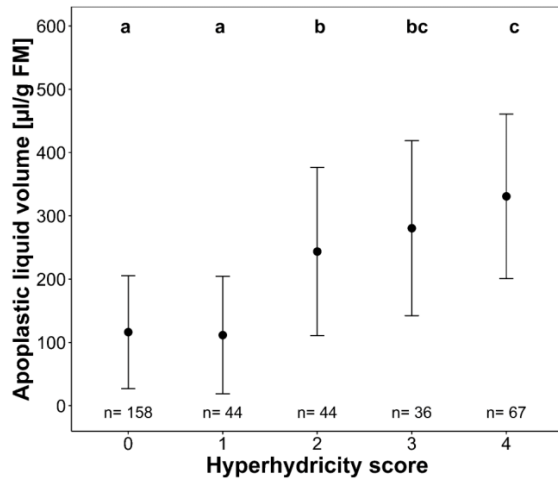
The apoplastic liquid volume increased steadily for *Malus* (Fig. 3A) until 15 DAT, while Experiment IV (SI. 3) demonstrated a decrease at later time points: 21 days and 28 days. A significant difference in apoplastic liquid content in both plant species was detected at the earliest time point: 4 DAT (SI. 3) and 5 DAT (Fig. 3A and B), where the apoplastic liquid volumes of explants on gelrite induction media were already twice as high as those of explants on agar control media. In *Malus*, three independent experiments (Fig. 3A and SI. 3) allowed us to confine the time of peak in apoplastic liquid volume at 12 to 16 DAT. Apparently, up to this timepoint, quantification of apoplastic liquid volume reflected the HH score well—even the occurrence of some hyperhydric explants on the agar control medium was also reflected in the increase in apoplastic water volume (Fig. 3A vs. SI. 3). However, at later time points, the severity of HH symptoms steadily increased, while apoplastic liquid volume stayed constant.

To prove the relation between the objective quantification of apoplastic liquid volume and the HH score determined by visual scoring, data pairs of a total of 349 measurements



**Fig. 3** Apoplastic liquid volume of **A** *Malus* 'G214' and **B** *A. thaliana* Col-0 in vitro cultures over time (DAT, Days After Treatment). Samples from 0 days after transfer (DAT 0) represent the starting plant material cultured on control media. Yellow lines indicate cultures on standard media **A** MS-Medium, **B** Gamborg-B5 solidified with 0.8% (w/v) agar, while gray dashed lines display the cultures on HH induction media containing 0.25% (w/v) gelrite (Mean ± SD). The values of *A. thaliana* at DAT 0 **B** were masked in gray to indicate

the authors' uncertainty, because the plants were very small when the apoplastic water volume was determined at this time, and therefore a large influence of adhering water could not be excluded. Replicate number (n) indicates the individual explants. Different letters resulted from Tukey's HSD test at  $p < 0.05$  and indicate significant differences when comparing time points within one treatment, while asterisks indicate comparisons of treatments within a time point with \* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ . (Color figure online)



**Fig. 4** Relation between visual scoring of hyperhydricity and apoplastic liquid volume of *Malus* 'G214' (Mean  $\pm$  SD). Data obtained from three different induction experiments (Exp. III-V) covering time points from 0 to 28 DAT. Replicate number (n) indicates the individual explants. Different letters resulted from Tukey's HSD test at  $p < 0.05$  and show significant differences between score levels

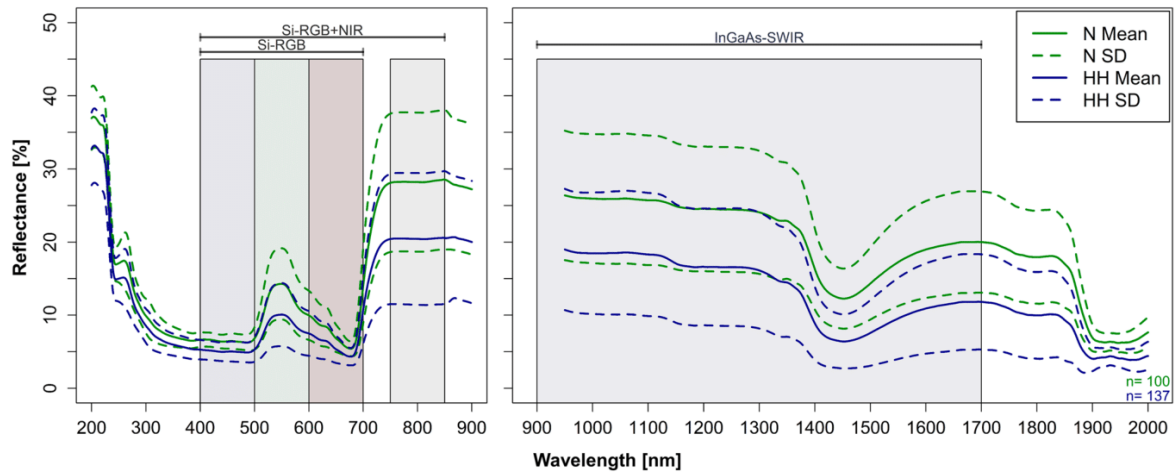
from both treatments of *Malus* were used (Fig. 4). We found the highest correlation between HH score and apoplastic volume to be  $\rho(18) = 0.83$  ( $p < 0.001$ ) 12 DAT using Spearman's rank correlation for all acquired time points for *Malus*. Interestingly, only three groups could be distinguished

significantly by apoplastic liquid volume. Explants with a HH score of two had more than  $> 50\%$  curled leaves and in average a double amount of apoplastic liquid volume. We therefore restricted the three significant groups to two classes (HH score 0–1: normal explants and HH score 2–4: hyperhydric explants) in further analysis, regardless of the treatment in order to exclude treatment-dependent effects on the spectral analysis.

### Spectral analysis of hyperhydricity

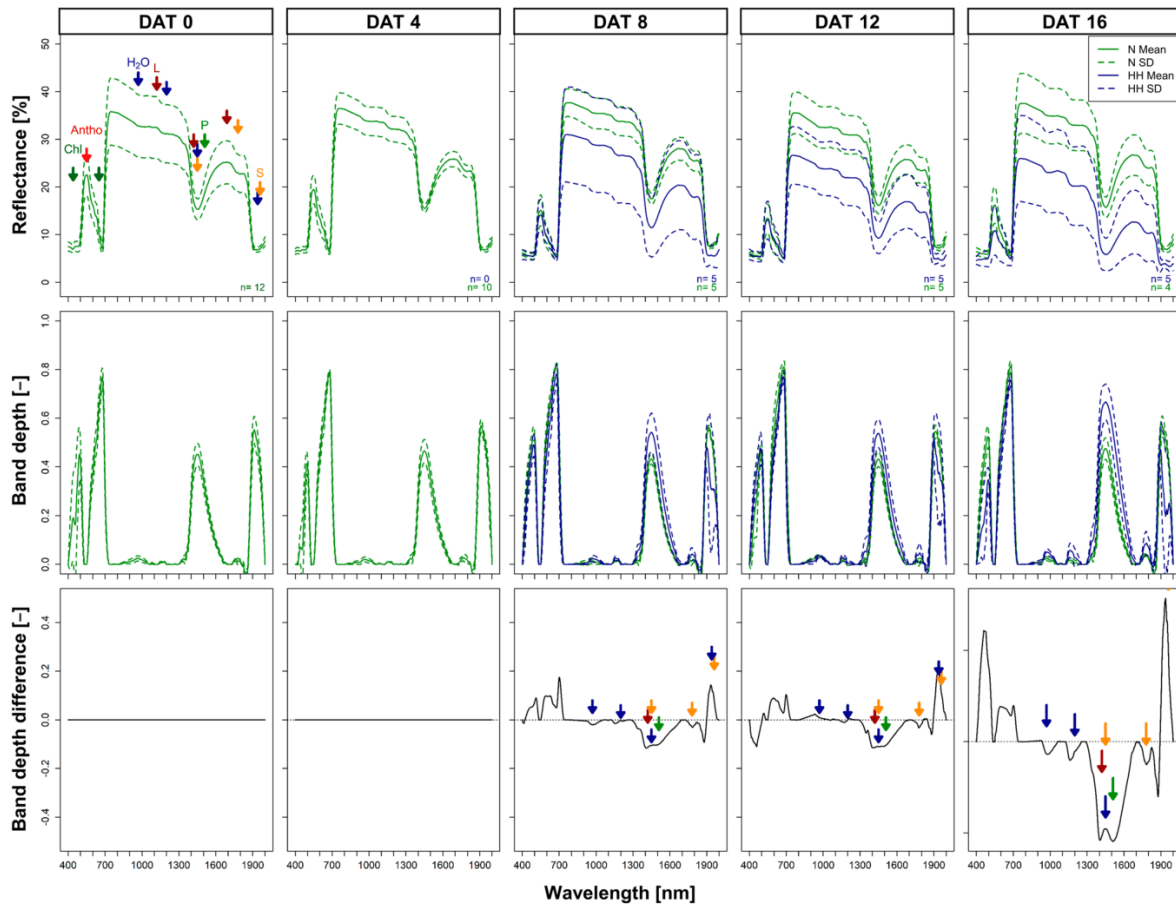
The evaluation of leaf explants via UV–VIS–NIR–SWIR spectroscopy (Fig. 5) revealed the first major difference in significantly reduced reflectance in the RGB (400 nm to 700 nm) region of  $6.5 \pm 3.2\%$  for the hyperhydric explants compared to  $8.7 \pm 3.9\%$  for normal explants. The largest difference in reflectance was recorded for NIR (750 nm to 850 nm) region with a reflectance of  $20.5 \pm 9.0\%$  hyperhydric explants and normal explants with  $28.3 \pm 9.5\%$ . Also, for the SWIR (950 nm to 1700 nm) region the overall reflectance was lower in hyperhydric explants ( $13.8 \pm 8.2\%$  for hyperhydric and  $21.4 \pm 8.7\%$  for normal explants). Differences in average reflectance were most significant in the blue ( $p < 2.2 \cdot 10^{-16}$ ) region followed by SWIR ( $p < 4.3 \cdot 10^{-14}$ ), green ( $p < 1.1 \cdot 10^{-12}$ ), red ( $p < 2.5 \cdot 10^{-11}$ ) and the NIR region ( $p < 2.0 \cdot 10^{-9}$ ) according to the results of a Mann–Whitney test.

The emergence of HH-specific absorption features over time was recorded applying the continuum removal method to pre-processed spectra of Experiment V and the



**Fig. 5** Raw reflectance spectra of *Malus* 'G214' and *A. thaliana* Col-0 in vitro leaves. Mean (solid)  $\pm$  SD (dashed) spectra of normal leaves (N, in green) and hyperhydric leaves (HH, in blue). The distinction was based on visual scoring of HH (N: 0–1 HH score; HH: 2–4 HH score). Wavebands represent different spectral regions, defined by the sensitivity of silicon-based (Si) cameras, such as

affordable RGB and RGB–NIR multispectral cameras, and a more expensive Indium–Gallium–Arsenide-based (InGaAs) detection sensor. Reflectance spectra were measured with an UV–VIS–NIR–SWIR spectrometer (PerkinElmer Lambda 950) in a wavelength range of 200 nm to 2000 nm and at a resolution of 1 nm. (Color figure online)



**Fig. 6** Spectral contrasting reflectance of *Malus* 'G214' normal and hyperhydraulic explants over time (DAT, Days After Treatment). The spectral data used originate from Experiment V. Upper row: Raw reflection spectra; middle row: extracted absorption features after segmented convex-hull removal of raw spectra; bottom row: difference spectrum of the absorption peaks, where the absorption features spectra of normal explants were subtracted from absorption features spectra of hyperhydraulic explants. Mean (solid)  $\pm$  SD (dashed) spectra of normal explant leaves (N, in green) and hyperhydraulic explant leaves (HH, in blue). Distinction of N and HH was based on visual scoring

of HH (N: 0–1 HH score; HH: 2–4 HH score). Arrows indicate putative major biochemical compounds absorbing in the given wavelength region, according to Curran (1989). Colored arrows represent: "Chl" = chlorophyll (dark green), "Antho" = anthocyanin (red), "H<sub>2</sub>O" = water (dark blue), "L" = lignin (dark red), "P" = protein (green), "S" = sugar (yellow). Reflection spectra were measured with an UV–VIS–NIR–SWIR spectrometer (PerkinElmer Lambda 950) in a wavelength range of 200 nm to 2000 and at a resolution of 1 nm. Absorption features spectra of the other conducted experiments, showing similar results, can be found in SI. 6. (Color figure online)

formation of difference spectra of the isolated band depth spectra, where the absorption features spectra of normal explants were subtracted from absorption features spectra of hyperhydraulic explants (Fig. 6). Greater absorption of hyperhydraulic explants was observed as early as 8 DAT, with a maximum at 1402 nm and a full width at half maximum of 157 nm. At later time points the difference in absorbance at around 980 nm, 1150 nm, 1400 nm, 1520 nm and 1780 nm increased negatively, while at around 1930 nm the difference positively increased. In the VIS region, two further local maxima arose at 460 nm and 695 nm at DAT 10, which were also detected at the later time points. However, these peaks

can be considered as artefacts of the reduced reflection in the green region due to the continuum removal method based on connection of local maxima. In addition, a consistent positive peak indicating less absorption or higher reflection of hyperhydraulic explants was found with a maximum around 1930 nm, besides the two local minima at 1400 nm and 1520 nm. When combining data from all experiments (SI. 6), including different time points and the two different plant species, we identified reliable minima (arrows) at 980 nm, 1150 nm, 1400 nm, 1520 nm, and 1780 nm, indicating stronger absorption of the hyperhydraulic explants, and a reliable maximum at 1930 nm.

Plant Cell, Tissue and Organ Culture (PCTOC)

**Table 4** Performance metrics of machine learning (ML)-based spectral classifiers. Bold letters indicate the value for the best performing model in each column

ML model	Training				
	Data set [No. of spectra]	Train time [s]	AUC <sub>ROC</sub> <sup>a</sup> [Mean ± SD]	Sensitivity <sup>a</sup> [Mean ± SD]	Specificity <sup>a</sup> [Mean ± SD]
NNET	178	761.7	0.93 ± 0.07	0.89 ± 0.10	<b>0.83 ± 0.13</b>
LDA	178	115.2	0.87 ± 0.09	0.82 ± 0.12	0.74 ± 0.16
SVM	178	42.2	0.93 ± 0.07	<b>0.91 ± 0.08</b>	0.79 ± 0.15
RF	178	819.5	0.92 ± 0.06	0.84 ± 0.11	0.80 ± 0.15
PLS	178	5.6	<b>0.94 ± 0.06</b>	0.90 ± 0.09	0.82 ± 0.13
HD.DA	178	38.0	0.83 ± 0.09	0.83 ± 0.12	0.78 ± 0.14
PCA.LD*	178	<b>1.4</b>	<b>0.94 ± 0.06</b>	0.89 ± 0.09	0.82 ± 0.13
NDRI <sup>b</sup>		–	–	–	–

ML model	Test						
	Data set [No. of spectra]	Accuracy	Accuracy [95% CI]	Balanced accuracy	Sensitivity <sup>a</sup>	F <sub>1</sub> score	Specificity <sup>a</sup>
NNET	59	0.81	0.69–0.90	0.81	0.85	0.84	0.76
LDA	59	0.73	0.59–0.84	0.72	0.79	0.77	0.64
SVM	59	<b>0.85</b>	<b>0.82–0.93</b>	<b>0.84</b>	<b>0.91</b>	<b>0.87</b>	0.76
RF	59	0.71	0.58–0.82	0.72	0.68	0.73	0.76
PLS	59	0.69	0.56–0.81	0.70	0.68	0.72	0.72
HD.DA	59	0.61	0.47–0.73	0.62	0.53	0.61	0.72
PCA.LD*	59	0.69	0.56–0.81	0.70	0.68	0.72	0.72
NDRI <sup>b</sup>	59	0.63	0.51–0.77	0.66	0.47	0.59	<b>0.84</b>

\*Principal component analysis (PCA) was performed prior to linear discriminate analysis (LD), therefore the training time should be considered slightly higher

<sup>a</sup>Note: AUC<sub>ROC</sub>, sensitivity and specificity were calculated with normal explants as the positive class

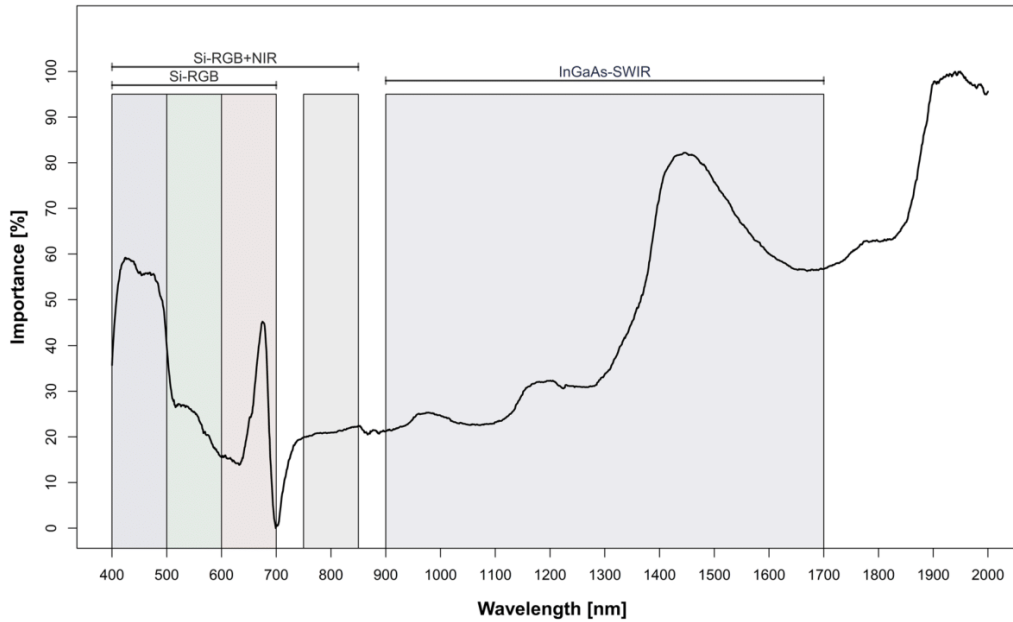
<sup>b</sup>Note: Normalized difference ratio index with a threshold of 0.35

To demonstrate whether the observed differences in reflectance spectra are sufficient to reliably discriminate between hyperhydric and non-hyperhydric explants, while generalizing plant species and time points, we performed a model spot checking for several ML models with whole spectral data sets as input (Table 4). The models spot check based on AUC<sub>ROC</sub> metrics identified partial least square (PLS) and linear discriminate analysis with upstream principal component analysis (PCA.LD) with 0.94 to be superior in the training step when classifying the explants against the other models, while supported vector machine (SVM), neutral net (NNET) with 0.93 and random forest model (RF) with 0.92 performed only slightly worse. High dimensional linear discriminate analysis (HD.LD) showed the lowest performance and was therefore excluded. Furthermore, SVM was best in classifying normal explants as expressed in the sensitivity metrics with 0.91 ± 0.08, while NNET reached with 0.93 ± 0.13 the highest specificity indicating the best performance in identifying hyperhydric explants. On the test set consisting of 59 unseen spectra, SVM outperformed the other models with the highest accuracy with 0.85, the

highest balanced accuracy with 0.84, the highest sensitivity with 0.91 and the highest F1 score of 0.87 and was therefore selected as final model, besides for its low training time and its better human interpretability. As a reference of a classical approach, we checked the classification performance of a two-band normalized difference ratio index using a threshold of 0.35, which resulted in a low accuracy of 0.63.

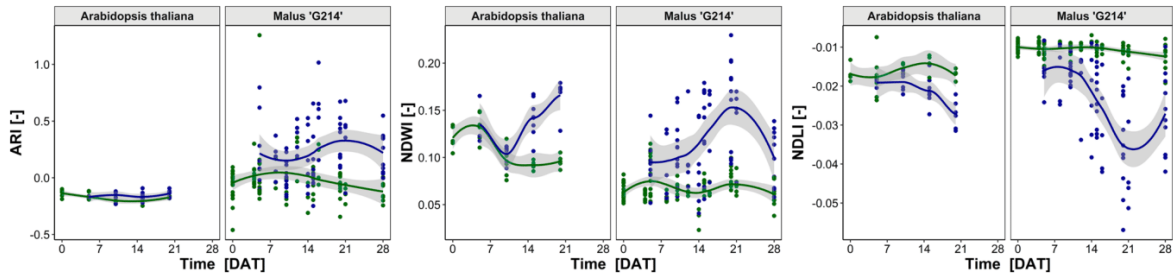
The evaluation of the predictor importance based on ROC-curve importance of SVM revealed the most important wavelength for classification (Fig. 7). The most relevant wavelength for classification was found at 1949 nm, followed by the peak at 1445 nm in the SWIR region, 424 nm in the blue region and 676 nm in the red region. The wavelength region from 700 to 900 nm, including the NIR region, contained the least essential information for the classification. In the green region, 500 nm was most important, while in the SWIR region two further peaks were identified at 975 nm and 1202 nm.

The 237 acquired spectra of the two species were further used to simulate three in literature stated HH-affected leaf compounds over time (anthocyanin, water, lignin) via



**Fig. 7** Variable importance of spectral classification of hyperhydricity using a support vector machine approach. Classification classes consisted of reflection spectra from either normal or hyperhydric leaves based on visual scoring of HH (N: 0–1 HH score; HH: 2–4 HH

score). Wavebands representing different spectral regions, defined by the sensitivity of silicon-based (Si) cameras, such as affordable RGB and RGB-NIR multispectral cameras, and more expensive Indium-Gallium-Arsenide-based (InGaAs) sensors as candidates for detection



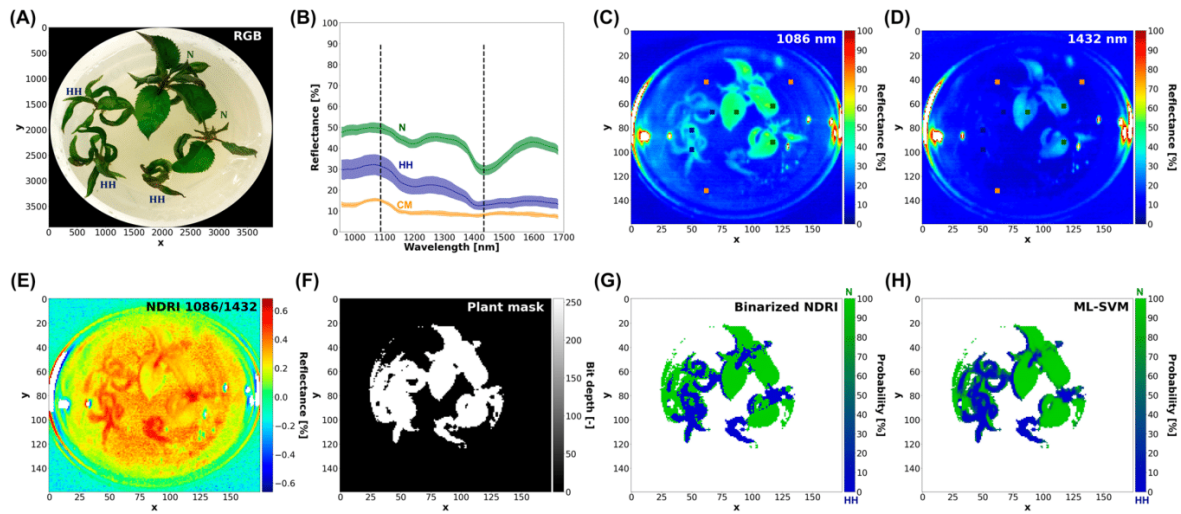
**Fig. 8** Selection of contrasting vegetation indices to hyperhydricity inducing cultivation of *Malus* 'G214' and *A. thaliana* 'Col-0'. Vegetation indices were calculated from spectra from three different experiments (Exp. III-V). Data points from normal explants are indicated in green (N: 0–1 HH score), while the blue color represents data from

hyperhydric tissue (HH: 2–4 HH score). Estimated 95% confidence interval was colored in light gray, while lines illustrate the locally weighted data trend by 2<sup>nd</sup> order polynomial regression. ARI/mARI defined according to Gitelson et al. (2006), NDWI from Gao (1996) and NDLI according to Serrano et al. (2002). (Color figure online)

described vegetation indices (Fig. 8; mARI, Gitelson et al. 2006; NDWI, Gao 1996; NDLI, Serrano et al. 2002). Hyperhydric explants of *A. thaliana* showed a relatively small increase in ARI, high increase in NDWI and notable reduction in NDLI compared to normal explants. All three vegetation indices simulated using spectra of *Malus* "G214" classified as hyperhydric, revealed a strong change over time compared to normal spectra.

**Automated detection of hyperhydricity**

To test the validity of the SWIR region of the trained spectral classifier as spectral region with high importance for discrimination of HH and to see the generalization to a new domain, the classifier was applied on a previously acquired SWIR-HSI data set from culture vessels containing normal (N, green) and hyperhydric explants (HH, blue)



**Fig. 9** Validation test of major absorption features by hyperspectral imaging. **A** Reference RGB image of a *Malus* ‘Selection 4’ vessel with normal (N, green) and hyperhydric (HH, blue) explants used for hyperspectral imaging of the SWIR region with the EVK Helios Core NIR Line-scan camera (240 px × 1 px and 252 spectral channels in the wavelength region of 900 nm to 1700 nm, according to Thiel 2018). **B** SWIR reflectance spectra of normal leaf, hyperhydric leaf and culture media (CM) pixels (green, blue and orange) with the dashed vertical lines indicating spectral locations of selected wavelengths. **C** and **D** False-color images of selected wavelengths. **E** Normalized difference ratio of selected wavelengths used to illustrate classical discriminating approach via **F** segmentation of plant pixels and **G** binarization by thresholding. Proposed discriminating approach by application of **H** ML model “Support vector machine” (ML-SVM) on segmented plant pixels of the SWIR-Hyperspectral-Image-Cube (SWIR-HSI-Cube). ML-SVM was laboratory-trained with single leaf reflection spectra and is presented as the predicted probability images of plant pixels. Hyperspectral imaging was performed through the lid of the culture vessel. (Color figure online)

at 1086 nm and 1432 nm. **E** Normalized difference ratio of selected wavelengths used to illustrate classical discriminating approach via **F** segmentation of plant pixels and **G** binarization by thresholding. Proposed discriminating approach by application of **H** ML model “Support vector machine” (ML-SVM) on segmented plant pixels of the SWIR-Hyperspectral-Image-Cube (SWIR-HSI-Cube). ML-SVM was laboratory-trained with single leaf reflection spectra and is presented as the predicted probability images of plant pixels. Hyperspectral imaging was performed through the lid of the culture vessel. (Color figure online)

of *Malus* ‘Selection 4’ (Fig. 9A). From the spectral signatures (Fig. 9B), a normalized difference ratio index (NDRI, Fig. 9C–E) as a two-band index with a HH-insensitive wavelength at 1086 nm (Fig. 9C) and a HH-responsive wavelength at 1432 nm (Fig. 9D) was derived. Hyperhydric explants became almost invisible due to their high absorption/ reduced reflection (R) at 1432 nm (Fig. 9D and SI. 7). Based on the acquired spectral signature a normalized difference ratio index (NDRI) could be derived (Eq. 12), which is formed by two wavelengths, an HH-insensitive correction wavelength at 1086 nm and a HH-sensitive at 1432 nm.

$$\text{NDRI} = \frac{(R_{1086\text{nm}} - R_{1432\text{nm}})}{(R_{1086\text{nm}} + R_{1432\text{nm}})} \quad (12)$$

The NDRI image (Fig. 9E) was segmented with a mask for plant pixels (Fig. 9F) and a threshold was applied to produce the classification image (Fig. 9G). For the ML approach that included the application of the spectral classifier (Fig. 9H), some modifications were made to the trained spectral classifier, such as spectral resampling to fit the spectral sensor channels and segmentation to limit the task to a two-class problem (see Materials and Methods section).

As a more affordable approach of HH detection—SWIR camera systems can cost hundred to thousand times more than an RGB camera system—three different object detection models were trained based on RGB image time series data sets to determine if the information contained in the three spectral channels of the RGB images (in addition to the observed morphological differences in the shape of the explants) was sufficient to correctly classify the hyperhydric explants. With all three trained models (Table 5) a high mAP of > 88% was observed for the validation set, indicating a high accuracy in localization and correct classification of the explants in the images. Highest precision of 86.8% in validation set was reached with the model PCTOC\_V2. For this model, we used the Roboflow Train option to train an object model from scratch. The model PCTOC\_V3 performed best in terms of the recall metric with 95.7% and mAP with 95.6% in the validation set. In an unseen test data set PCTOC\_V3 outperformed the other models in mAP with 97.0% and highest recall 89.0% and was therefore selected to visualize its performance on a selection of test set images (Fig. 10) and on unseen time-series data from two culture vessels of the same experiment (SI. 8).

The PCTOC\_V3 model identified multiple objects on the selection of test set images (Fig. 10) with only slightly greater predicted bounding boxes compared to ground truth.

**Table 5** Performance metrics of object detection models trained on RGB images. Bold letters indicate the value for the best performing model in each column

Name	Model architecture	Training		Validation				Test			
		Data set [No. of images]	Description	Data set [No. of images]	mAP [%]	Precision [%]	Recall [%]	Data set [No. of images]	mAP [%]	Precision <sup>c</sup> [%]	Recall <sup>c</sup> [%]
PCTOC_V1	YOLOv8	250	Colab with weights from scratch	50	88.4	83.0	82.1	25	NYI <sup>a</sup>	<b>94.4</b>	49.5
PCTOC_V2	Roboflow 2.0 OD	250	Weights from scratch	50	93.5	<b>86.8</b>	86.4	25	95.0	90.4	87.6
PCTOC_V3	Roboflow 2.0 OD	375 <sup>b</sup>	Weights from PCTOC_V2	50	<b>95.6</b>	83.8	<b>95.7</b>	25	<b>97.0</b>	93.7	<b>89.0</b>

<sup>a</sup>Note: Not yet implemented in Ultralytics YOLOv8.0.20

<sup>b</sup>Note: Trained with weights from PCTOC\_V2 (based 250 images) and additionally 125 images

<sup>c</sup>Note: Precision and recall on test set were calculated with IoU and confidence threshold of 0.5

A supposedly perfect classification could be reached with the prediction settings used. However, severely hyperhydric explants (Fig. 10B1) received a lower class membership probability than explants with developing HH symptoms (Fig. 10B2). Class membership probability of normal explants was generally high on the test set (Fig. 10B3) and seemed to be stable even on the time-series RGB image data set (SI. 8, left image). For hyperhydric explants, prediction confidence increased until day 10 and decreased at day 16 in the time-series RGB image data set (SI. 8, right image).

## Discussion

### Time-lapse videos enable insights into early phases of HH development

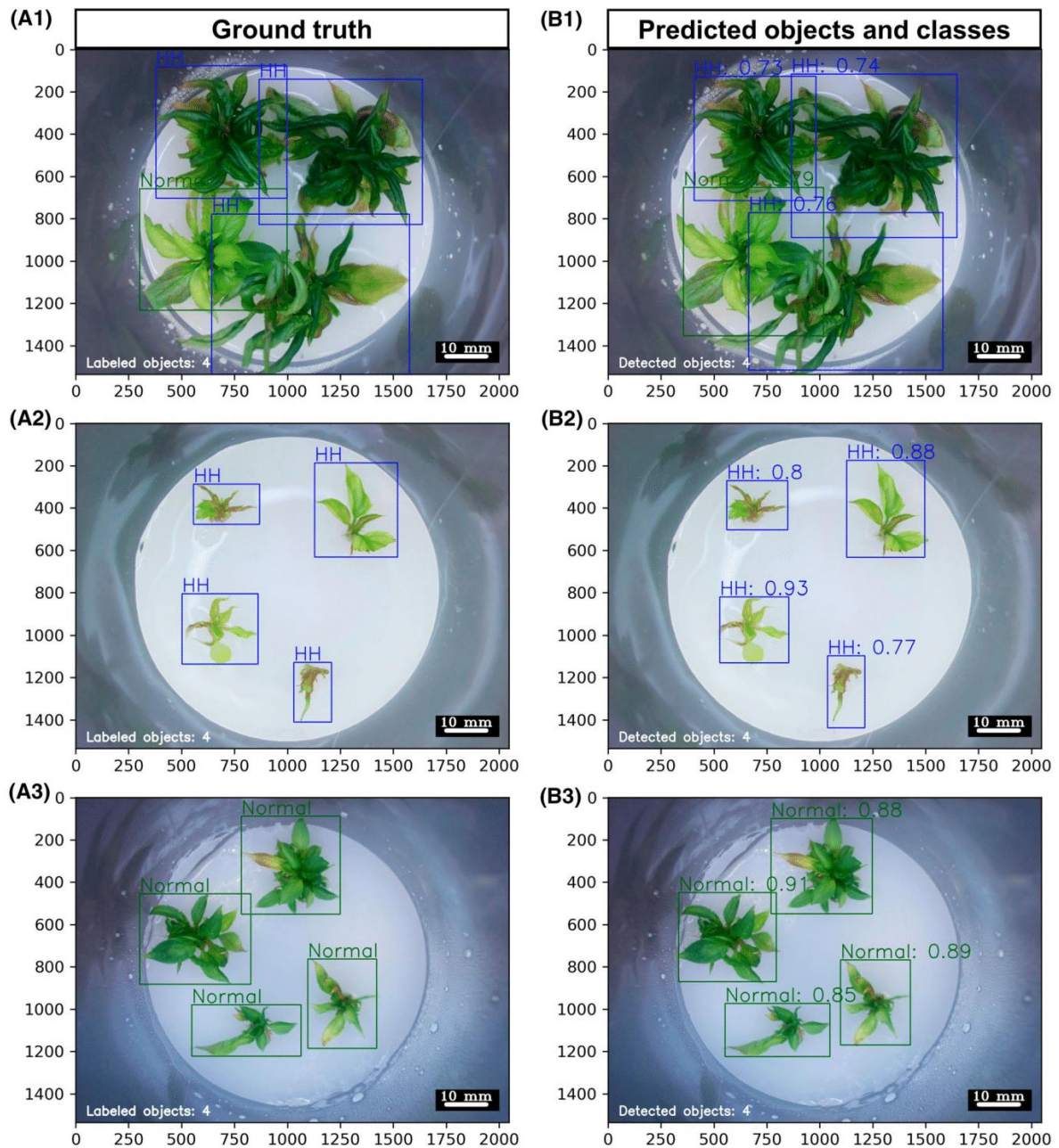
HH is a serious limitation of plant tissue propagation affecting multiple phases of in vitro cultivation. The use of the novel monitoring system “Phenomenon” capturing time series image data (SI. 4 and SI. 5) identified (i) the first visual symptoms of HH to occur 5 DAT and (ii) an accelerated and higher growth of shoots of the gelrite treatment (Fig. 1A). Thereby, significant differences in the projected plant area between the two treatments were found already 5 days after transferring to the culture media in both species. As discussed previously by Kevers et al. (1984), HH may be considered as morphological response to waterlogging, which in turn induces ethylene synthesis. For *A. thaliana*, we observed a higher vertical growth (Fig. 1B) with hyponasty (SI. 5), which was described as ethylene-triggered strategy of ex vitro plants in waterlogging conditions to re-establish contact with air and restore successful gas exchange (Voeselek and Blom 1989). Furthermore, Vreeburg et al. (2005) described a flooding-induced petiole elongation in a two-stage process, starting with acidification of the apoplast

followed by cell wall expansion. This is in agreement with our observation of a significantly higher eccentricity (deviation of the ellipse to circle) and significantly less solidity (density of the object) for explants in the gelrite treatment (SI. 1B). A more pronounced curling of the leaves was observed in *Malus* (SI. 4) which also resulted in epinastic leaf growth. In addition, a significant higher mean canopy height (Fig. 1B) and maximum shoot height of *Malus* shoots on gelrite medium (SI. 1A) indicated a more pronounced vertical orientation of growth.

### Hyperhydricity induction by increased water availability

Although HH symptoms vary between different plant species and cultivars, and several factors have been described to trigger HH, a putative common underlying mechanism of apoplast flooding has been described (van den Dries et al. 2013). Several studies showed that increasing the water availability by decreasing the concentration of gelling agent, changing the type of gelling agent or the cultivation in liquid media induced HH in a large set of plant species (*Dianthus sp.*, Casanova et al. 2008; *Aloe sp.*, Ivanova and Van Staden 2011, *Malus sp.* Chakrabarty et al. 2003).

In our study, we demonstrated the HH-inducing effects of gelrite for *Malus* and *Arabidopsis* indicated by the overall increase in HH scores (Fig. 2 and SI. 3) and apoplastic liquid volume (Fig. 3 and SI. 3) over time. Gelrite differs from agar in terms of consistency and purity and resulted in a superior growth of explants at a comparable gel strength (Scherer 1987; Tsay et al. 2006; Pasqualetto et al. 1988). However, gelrite induced HH in several species (*Arabidopsis sp.*, van den Dries et al. (2013), *Malus sp.* Pasqualetto et al. 1988, *Prunus sp.* Franck et al. 1998) limiting the use of this gelling agent. Scherer et al. (1988)



**Fig. 10** Object detection performance of the PCTOC\_V3 model on an image selection of the test set. (A) Ground truth RGB image of the *Malus* ‘G214’ test set annotated with normal (Normal, green) and hyperhydric (HH, blue) explants (A1–A3). (B)

(B1–B3) and class membership probability (0 to 1 corresponds 0 to 100%). Prediction was performed with confidence threshold and intersection of union threshold of 0.5. (Color figure online)

could show that there is no difference in the osmotic and water potential of gelrite compared to agar. Van den Dries et al. (2013) suspected therefore a local dissolution of the culture medium due to the excretion of chelators by the

explants and thus a higher water availability and water uptake. This higher water availability in gelrite-solidified media most likely explains HH-induction and accelerated growth, but other putative factors like differences in



uptake of nutrients or plant hormones were also found: Higher contents of magnesium (Mg) and a higher ratio of potassium (K) to sodium (Na) were detected in the leaves of walnut explants grown on gelrite medium compared to agar, which can affect stomatal function (Barbes et al. 1993). Furthermore, Arthur et al. (2004) found a lower concentration of IAA-like compounds in gelrite than in different types of agar powder.

With the collected data and the time-lapse videos, we could narrow down crucial key points within the development of HH of the two species in time. First visual identifiable symptoms (SI. 4 and SI. 5) and significant increases in apoplastic liquid volume were observed already after 5 days of cultivation on gelrite media in both species. Time series dynamics of apoplastic liquid volume confirmed previous data for *A. thaliana* (van den Dries et al. 2013)—in both studies hyperhydric explants of *A. thaliana* had an apoplastic liquid volume of around 300  $\mu\text{L g}^{-1}$  FM 15 days after treatment, but were carried out for the first time for *Malus*. Quantification of apoplastic liquid volume of *A. thaliana* seedlings at very early time points was limited by the very small amounts of apoplastic liquid and the distorting effect of adhering water (Fig. 3B). For *Malus*, the highest increase in apoplastic liquid volume for the gelrite treatment was detected within the first 4–5 days in two independent experiments (SI. 3, Fig. 3B). Furthermore, a different behavior of the HH score and the apoplastic liquid volume was found after 21 days of cultivation in *Malus*: While the severity of HH symptoms steadily increased over time, the apoplastic liquid volume seemed to reach saturation at later time points (SI. 3, Fig. 4). Therefore, we suggest the HH score to be useful to determine the symptoms of HH, whereas the quantification of apoplastic liquid volume better reflects the physiological state of the explants.

### Identification of HH-specific spectral absorption features

Despite the fact that clear visible symptoms (Table 1) of HH were reported and still are the major distinguishing parameter for classification, spectroscopic analysis of HH is limited. Only Marques et al. (2021) using Fourier-transform infrared spectroscopy in attenuated total reflectance mode (FTIR-ATR), evaluated chemical properties of prepared cell walls of hyperhydric *Arbutus unedo*. Assuming HH as a consequence of flooding of the air-filled apoplast by water, UV–VIS–NIR–SWIR reflection spectroscopy was expected to detect these physiological changes due to higher light absorption of water compared to air. Therefore, we applied this technique to identify specific absorption features of HH essential for designing an automated detection system. However, we excluded the UV region ( $< 400$  nm) from further analysis due to the low penetration depth of UV light in plant

tissue (Qi et al. 2010), since most reflection signals can only be attributed to anatomical and biochemical properties of cuticle, trichomes and the upper epidermis.

The observed overall reduction in reflectance of hyperhydric explants (Fig. 5) compared to normal ones is consistent with the visual appearance of the observed darkening of the affected explants (SI. 4 and SI. 5). The visualization of isolated absorption features over the time course of the development of HH in *Malus* (Fig. 6) should give insights whether there is at least a trend in the time course of the presumed absorption characteristics. We used the continuum removal method to exclude the observed overall absolute reduction in reflection and to compare all spectra on a common base. This allowed an automated extraction of absorption peaks for the SWIR region with predominant absorptions valleys, however, produced artefacts in the VIS region. In the SWIR region, a consistent difference between absorption of normal and hyperhydric leaves was observed for the wavelengths 980 nm, 1150 nm, 1400 nm, 1520 nm, 1780 nm and 1930 nm, both over time (Fig. 6) and in the different experiments (SI. 6). Most likely, the absorption of water in the plant tissue is most responsible for the wavelengths 980 nm, 1150 nm, 1400 nm. Curran et al. (1989) described the intense absorption of liquid water at 970 nm, 1200 nm, 1400 nm and 1450 nm due to the fundamental O–H bending vibrations of the first overtone. Thus, the tendency of an increase in water absorption (970 nm, 1200 nm, 1400 nm, 1450 nm) within time is in accordance with the increase in apoplastic liquid volume over time. However, absorption bands of other compounds like proteins, lignins and sugars are located within the peak between 1300 to 1600 nm and contribute to the total absorption in this region. Curran et al. (1989) associated the absorption at 1780 nm to cellulose, sugars and starch. Since for this wavelength a higher absorption in hyperhydric leaves was observed in our study, this is in line with the detection of a higher sugar content (sucrose, glucose and fructose) in hyperhydric explants of *Dianthus* (Saher et al. 2005), but contradicting Kevers et al. (1987) who reported less lignin and cellulose in hyperhydric *Dianthus*.

Simulation of vegetation indices (Fig. 8) demonstrated traceable trends, that closely match the dynamics of the physiological reference data (Fig. 3 & SI. 3) and support the observation of time-series data (SI. 4 & SI. 5). Overall, the vegetation indices from normal explants exhibited low variance, although they were derived from different experiments. The high variance of hyperhydric explants indicated by the confidence interval can be explained by different physiological states of explants with different degrees of hyperhydricity. The simulation of a modified anthocyanin index, indicated a higher anthocyanin content in hyperhydric leaves of *Malus*, but not *Arabidopsis*, supporting our RGB image time series. The normalized difference water index

(NDWI) displayed higher water contents for hyperhydric explants of both species and supported our observation that apoplastic liquid volume did not increase any more after 4 weeks of cultivation. The normalized difference lignin index (NDLI) showed in both species less lignin for hyperhydric leaves. However, the trend of the NDLI curves in both species followed inversely that of the NDWI indicating a putative dependency on plant water content. Marques et al. (2021) found no significant difference in the lignin content per dry weight of hyperhydric and normal leaves of *Arbutus*, whereas Kevers et al. (1987) reported a lower lignin content per fresh weight of hyperhydric tissue. It remains to be clarified, whether these divergent results are due to different species or to the fact that the fresh mass of hyperhydric explants is much higher.

### Automated detection of HH by machine learning

In order to evaluate the performance of the spectral data in the classification of hyperhydric and normal leaves, we trained different ML models (Table 4), investigated the most important wavelengths of the best model (Fig. 7) and compared them against a novel vegetation index as the classical approach (Fig. 9). The ML models differed in their architecture, complexity, performance, prediction time and interpretability (Singh et al. 2016; Liakos et al. 2018; and Hesami and Jones 2020). All ML models reached a high  $AUC_{ROC} > 0.83$  in training, however only SVM and NNET had a high accuracy  $> 0.80$  on test data. Both ML models outperformed with an accuracy of 0.81 for NNET (balanced accuracy of 0.81) and 0.85 for SVM (balanced accuracy of 0.84) the univariate vegetation index approach with a lower accuracy of 0.63 (balanced accuracy of 0.66). Furthermore, SVM was best in classifying normal spectra indicated by highest sensitivity of 0.91 on the test set. The two-band vegetation index NDRI reached the highest specificity of 0.84 in the test data, followed by SVM with 0.76, meaning highest ratio in the identification of hyperhydric tissue, however, low sensitivity of 0.47, low accuracy of 0.63 and low  $F_1$  score of 0.59 indicated a conservative behavior of classification towards hyperhydric explants. SVM was selected due to its high performance on training and testing datasets, low training data volume requirements, performance on high-dimensional datasets, low risk of overfitting, good generalization ability, and its advantages over NNET in terms of training time, simplified structure, and interpretability (Singh et al. 2016; Liakos et al. 2018). The evaluation of the feature importance of SVM for classification (Fig. 7) supported our findings that bands (peaks with maxima at 1949 nm, 1445 nm, 1202 nm and 975 nm) associated with water absorption were crucial to distinguish between hyperhydric and normal leaves. However, the method indicated essential features importance in the VIS region with maxima

at 424 nm and 676 nm. In regard of an automated HH detection system, we further evaluated two different approaches (i) HH detection based on an HSI-SWIR camera system (Fig. 9) and (ii) HH detection based on RGB camera system coupled with a deep neuronal network (DNN) to provide two putative solutions for commercial plant propagation based on our findings (Table 5, Fig. 10, SI. 8).

Following the HSI-SWIR camera system approach, we could only test the validity of our spectral classifier as a proof of concept because we only had a single HSI acquisition (SI. 7), so these results should be interpreted with caution. In addition, our analysis followed a two-class classification problem, but under the assumption that an automated HH detection system monitors explants on culture media during cultivation, culture media spectra could presumably interfere with the other classes within classification. Therefore, for further studies, we propose to include the acquisition of reflectance spectra of the culture media in the dataset. Nevertheless, we could test our ML-SVM classifier, trained on spectra from *Malus* 'G214' and *A. thaliana*, on the single SWIR-HSI acquisition of *Malus* 'Selection 4' segmented plant pixels, indicating the generalization ability of the classifier with respect to experimental setup and plant species/genotype. The NDRI and ML-SVM both classified most pixels correctly, however, ML-SVM segmented the borders of different classes much sharper. These preliminary results demonstrated that classification of HH is possible during in vitro cultivation and through the lid of the vessel with either an expensive SWIR-HSI system classifying with our novel ML-SVM classifier or more cost-effectively with a two-channel SWIR camera system using a novel vegetation index.

Alternatively, an RGB camera setup coupled with convolutional neural network (CNN) can be the most cost-effective solution for an automated HH-detection. Since we had identified feature importance also in the VIS region, a proof-of-concept study was conducted to demonstrate object detection via CNN. Therefore, we used the Roboflow<sup>®</sup> pipeline, which allowed an easy access to these tools and provided an interface for data annotation, pre-processing, data augmentation, training, data availability and deployment of the trained models. Comparing a self-trained YOLOv8 with the unknown object detection algorithms of Roboflow Train (Table 5), we did not reach the performance of their optimized model, which was particularly evident in the performance on test set, where PCTOC\_V1 reached the highest precision with 94.4%, but with low recall of 49.5%—indicating only half of all explants could be detected. The best trained model PCTOC\_V3, however had a precision of 83.8% on validation and of 97.0% on test set, indicating that prediction was mostly correct (Table 5, Fig. 10, SI. 8). In addition the explants were reliably detected (recall of 95.7% on

validation set and 89.0% on test set). By using the relatively new Python library roboflow, we encountered some unsolved issues as seen in SI. 8 where non-maximum-suppression only works so far within one class, resulting in multiple predictions per object. Considering the properties of the dataset, the low amount of data (250 to 375 images), resulting from time series images (1049 explants) of only 32 individual explants, we could see already good performance on the test set and the time-series set (Fig. 10 & SI. 8).

## Conclusions

To our knowledge this study is the first report of (i) identifying discriminating wavelengths in the VIS–NIR–SWIR region for the detection of HH, (ii) application of short wave infrared hyperspectral imaging to detect growth anomalies in vitro, (iii) proposing a spectral classifier for hyperhydricity. Wavelength bands (around 1940 nm, 1450 nm, 1200 nm and 970 nm) associated with absorption of water are the most distinguishable between hyperhydric and normal leaves within the analyzed spectral data set (400 nm to 2000 nm). In addition, minor important wavelengths were found in the RGB region (around 430 nm and 680 nm), whereas the NIR region seemed to be less important. Furthermore, RGB images of hyperhydric explants contain sufficient morphological and spectral features to allow a reliable detection of HH in an affordable manner via convolutional neuronal networks. However, this needs to be proven in an in-depth study. Nonetheless, these results can serve as a proof-of-concept for CNN-assisted live monitoring of plant tissue cultures and pave the way for increased use of CNN to estimate other key parameters such as multiplication rate, nutrient deficiency, and contamination.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11240-023-02528-0>.

**Acknowledgements** We thank the technical assistants Ewa Schneider and Bärbel Ernst of the department of Woody Plant and Propagation Physiology, Institute of Horticultural Production Systems, Leibniz Universität Hannover for their excellent support in the lab. Furthermore, we thank Matthias Igelbrink and Prof. Dr. Arno Ruckelshausen at University of Applied Science Osnabrück in their support in recording the SWIR-HSI data. In addition, we are grateful for the scholarship for the completion of a dissertation of the University of Applied Science Osnabrück.

**Author contributions** HB and TW designed the experiments, HB and ZM performed the experiments and analysed the data. HB wrote the manuscript and HB, ZM, TR, TW revised the manuscript. All the authors discussed the results and collectively edited the manuscript. All authors read and approved the final manuscript.

**Funding** Open Access funding enabled and organized by Projekt DEAL. This project took place within the research project “Experimentierfeld Agro-Nordwest”, which is funded by the Federal Ministry of Food and Agriculture (BMEL, Grant No.: 28DE103F18) via the Federal Agency for Agriculture and Food (BLE).

**Data availability** The datasets generated during the current study are available from the corresponding author on reasonable request. RGB image dataset analysed during the current study available in the Bethge (2023) repository, [<https://universe.roboflow.com/hains/hh-detection-in-vitro/dataset/8>].

## Declarations

**Competing interests** The authors declare no competing interests.

**Ethical approval** The research work was carried out in compliance with the ethical standards that do not involve the use of humans.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Arthur GD, Stirk WA, Van Staden J, Thomas TH (2004) Screening of aqueous extracts from gelling agents (Agar and Gelrite) for root-stimulating activity. *S Afr J Bot* 70(4):595–601. [https://doi.org/10.1016/S0254-6299\(15\)30197-6](https://doi.org/10.1016/S0254-6299(15)30197-6)
- Aynalem HM, Righetti TL, Reed BM (2006) Non-destructive evaluation of in vitro-stored plants: a comparison of visual and image analysis. *In Vitro Cell Dev Biol-Plant* 42(6):562–567. <https://doi.org/10.1079/IVP2006816>
- Barbas E, Jay-Allemand C, Doumas P, Chaillou S, Cornu D (1993) Effects of gelling agents on growth, mineral composition and naphthoquinone content of in vitro explants of hybrid walnut tree (*Juglans regia* × *Juglans nigra*). *Annales Des Sci* for 50(2):177–186. <https://doi.org/10.1051/forest:19930205>
- Bethge H (2023) HH Detection in vitro Image Dataset. <https://universe.roboflow.com/hains/hh-detection-in-vitro/dataset/8>. Accessed 10 Feb 2023
- Bethge H, Winkelmann T, Lüdeke P (2023) Rath T (2023) Low-cost and automated phenotyping system “Phenomenon” for multi-sensor in situ monitoring in plant in vitro culture. *Plant Methods* 19(1):1–25. <https://doi.org/10.1186/s13007-023-01018-w>
- Bock Biosciences GmbH (2018) RoBo®Cut. <https://www.robotec-ptic.com/>. Accessed 14 Feb 2023
- Bradski G (2000) The openCV library. *Dr. Dobb’s J Softw Tools Prof Progr* 25(11):120–123
- Cardoso JC, Sheng Gerald LT, Teixeira da Silva JA (2018) Micro-propagation in the twenty-first century. In: Loyola-Vargas VM, Ochoa-Alejo N (eds) *Plant cell culture protocols*. Springer, Dordrecht, pp 17–46

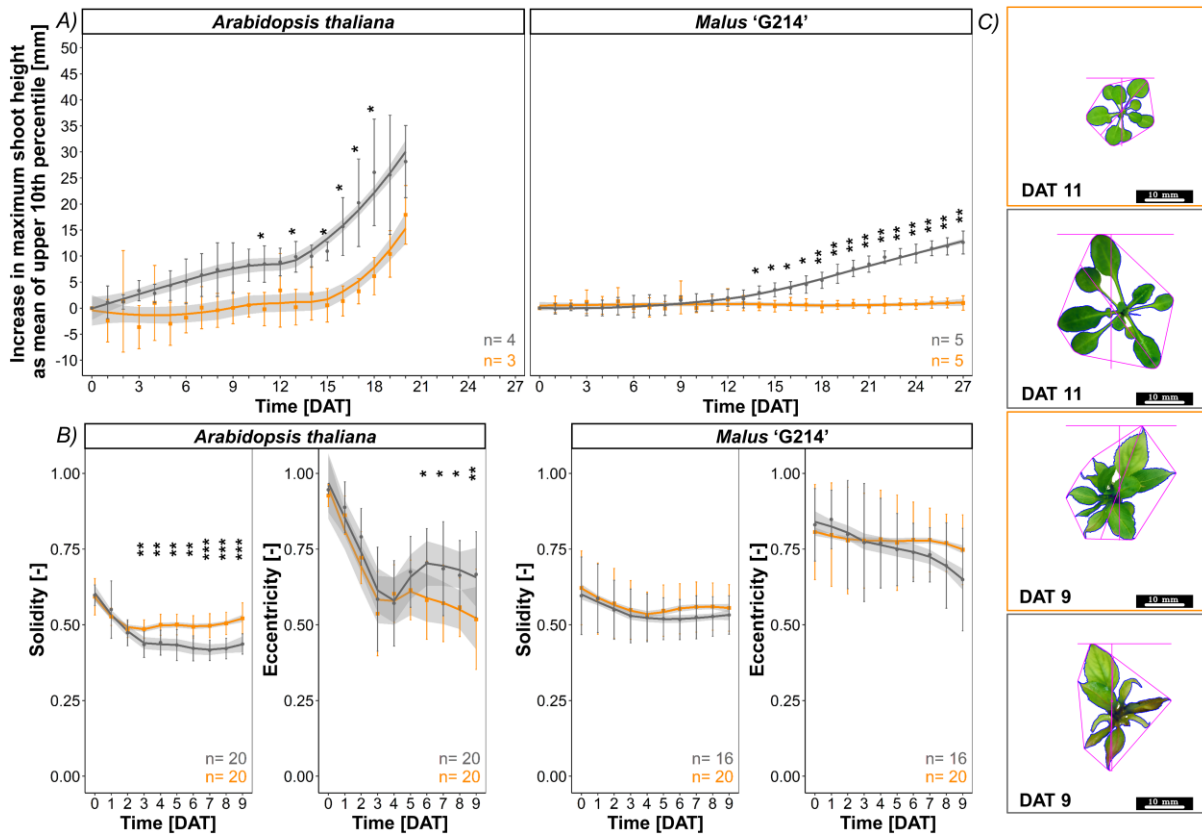
- Casanova E, Moysset L, Trillas MI (2008) Effects of agar concentration and vessel closure on the organogenesis and hyperhydricity of adventitious carnation shoots. *Biol Plant* 52:1–8. <https://doi.org/10.1007/s10535-008-0001-z>
- Chakrabarty D, Hahn EJ, Yoon YJ, Paek KY (2003) Micropropagation of apple rootstock M. 9 EMLA using bioreactor. *J Hort Sci Biotechnol* 78(5):605–609. <https://doi.org/10.1080/14620316.2003.11511671>
- Chen C (2016) Cost analysis of plant micropropagation of *Phalaenopsis*. *Plant Cell, Tissue Organ Cult* 126(1):167–175. <https://doi.org/10.1007/s11240-016-0987-4>
- Curran PJ (1989) Remote sensing of foliar chemistry. *Remote Sens Environ* 30(3):271–278. [https://doi.org/10.1016/0034-4257\(89\)90069-2](https://doi.org/10.1016/0034-4257(89)90069-2)
- de Klerk GJ, Pramanik D (2017) Trichloroacetate, an inhibitor of wax biosynthesis, prevents the development of hyperhydricity in *Arabidopsis* seedlings. *Plant Cell, Tissue Organ Cult* 131(1):89–95. <https://doi.org/10.1007/s11240-017-1264-x>
- Debergh P, Aitken-Christie J, Cohen D, Grout B, Von Arnold S, Zimmermann R, Ziv M (1992) Reconsideration of the term ‘vitrification’ as used in micropropagation. *Plant Cell, Tissue Organ Cult* 30(2):135–140. <https://doi.org/10.1007/BF00034307>
- Dhondt S, Gonzalez N, Blomme J, De Milde L, Van Daele T, Van Akoleyn D, Storme V, Coppens F, Beemster TS, Inzé D (2014) High-resolution time-resolved imaging of in vitro *Arabidopsis* rosette growth. *Plant J* 80(1):172–184. <https://doi.org/10.1111/tj.12610>
- Dwyer B, Nelson J, Solawetz J (2021) Roboflow python package. <https://github.com/roboflow/roboflow-python>. Accessed 10 Feb 2023
- Dwyer B, Nelson J, Solawetz J (2022) Roboflow (v1.0). <https://roboflow.com>. Accessed 14 Feb 2023
- Fischler MA, Bolles RC (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun ACM* 24(6):381–395. <https://doi.org/10.1145/358669.358692>
- Franck T, Crèvecoeur M, Wuest J, Greppin H, Gaspar T (1998) Cytological comparison of leaves and stems of *Prunus avium* L. shoots cultured on a solid medium with agar or gelrite. *Bio-technic Histochem* 73(1):32–43. <https://doi.org/10.3109/10520299809140504>
- Gamborg OL, Miller R, Ojima K (1968) Nutrient requirements of suspension cultures of soybean root cells. *Exp Cell Res* 50(1):151–158. [https://doi.org/10.1016/0014-4827\(68\)90403-5](https://doi.org/10.1016/0014-4827(68)90403-5)
- Gao BC (1996) NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens Environ* 58(3):257–266. [https://doi.org/10.1016/S0034-4257\(96\)00067-3](https://doi.org/10.1016/S0034-4257(96)00067-3)
- Gao H, Xia X, An L, Xin X, Liang Y (2017) Reversion of hyperhydricity in pink (*Dianthus chinensis* L.) plantlets by AgNO<sub>3</sub> and its associated mechanism during in vitro culture. *Plant Sci* 254:1–1. <https://doi.org/10.1016/j.plantsci.2016.10.008>
- Gehan MA, Fahlgren N, Abbasi A, Berry JC, Callen ST, Chavez L, Doust AN, Feldman MJ, Gilbert KB, Hodge JG, Hoyer JS (2017) PlantCV v2: image analysis software for high-throughput plant phenotyping. *PeerJ* 5:e4088. <https://doi.org/10.7717/peerj.4088>
- George EF, Hall MA, De Klerk GJ (2008) Plant propagation by tissue culture. In: George EF, Hall MA, De Klerk G-J (eds) Volume I. The background. *Plant propagation by tissue culture*. Springer, Dordrecht
- Gitelson AA, Keydan GP, Merzlyak MN (2006) Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves. *Geophys Res Lett.* <https://doi.org/10.1029/2006GL026457>
- Gribble K (1999) The influence of relative humidity on vitrification, growth and morphology of *Gypsophila paniculata* L. *Plant Growth Regul* 27(3):181–190. <https://doi.org/10.1023/A:1006235229848>
- Gupta SD, Karmakar A (2017) Machine vision based evaluation of impact of light emitting diodes (LEDs) on shoot regeneration and the effect of spectral quality on phenolic content and antioxidant capacity in *Swertia chirata*. *J Photochem Photobiol, B* 174:162–172. <https://doi.org/10.1016/j.jphotobiol.2017.07.029>
- Hesami M, Jones AM (2020) Application of artificial intelligence models and optimization algorithms in plant cell and tissue culture. *Appl Microbiol Biotechnol* 104(22):9449–9485. <https://doi.org/10.1007/s00253-020-10888-2>
- Honda H, Takikawa N, Noguchi H, Hanai T, Kobayashi T (1997) Image analysis associated with a fuzzy neural network and estimation of shoot length of regenerated rice callus. *J Ferment Bioeng* 84(4):342–347. [https://doi.org/10.1016/S0922-338X\(97\)89256-2](https://doi.org/10.1016/S0922-338X(97)89256-2)
- Huang YJ, Lee FF (2010) An automatic machine vision-guided grasping system for *Phalaenopsis* tissue culture plantlets. *Comput Electron Agric* 70(1):42–51. <https://doi.org/10.1016/j.compag.2009.08.011>
- Ivanova M, Van Staden J (2011) Influence of gelling agent and cytokinins on the control of hyperhydricity in *Aloe polyphylla*. *Plant Cell, Tissue Organ Cult* 104(1):13–21. <https://doi.org/10.1007/s11240-010-9794-5>
- Jocher G, Chaurasia, A, Qiu J (2023) YOLO by Ultralytics (Version 8.0.0). <https://github.com/ultralytics/ultralytics>. Accessed 14 Feb 2023
- Kemat N (2020) Improving the quality of tissue-cultured plants by fixing the problems related to an inadequate water balance, hyperhydricity. Doctoral dissertation, Wageningen University and Research. <https://doi.org/10.18174/517434>
- Kemat N, Visser RG, Krens FA (2021) Hypolignification: a decisive factor in the development of hyperhydricity. *Plants* 10(12):2625. <https://doi.org/10.3390/plants10122625>
- Kevers C, Coumans M, Coumans-Gillès MF, Caspar TH (1984) Physiological and biochemical events leading to vitrification of plants cultured in vitro. *Physiol Plant* 61(1):69–74. <https://doi.org/10.1111/j.1399-3054.1984.tb06102.x>
- Kevers C, Prat R, Gaspar T (1987) Vitrification of carnation in vitro: changes in cell wall mechanical properties, cellulose and lignin content. *Plant Growth Regul* 5(1):59–66. <https://doi.org/10.1007/BF00035020>
- Kuhn M (2008) Building predictive models in R using the caret package. *J Stat Softw* 28:1–26. <https://doi.org/10.18637/jss.v028.i05>
- Lee TJ, Zobayed SM, Firmani F, Park EJ (2019) A novel automated transplanting system for plant tissue culture. *Biosys Eng* 181:63–72. <https://doi.org/10.1016/j.biosystemseng.2019.02.012>
- Lehnert LW, Meyer H, Obermeier WA, Silva B, Regeling B, Bendix J (2018) Hyperspectral data analysis in R: the hsdar package. *arXiv Preprint*. <https://doi.org/10.48550/arXiv.1805.05090>
- Liakos KG, Busato P, Moshou D, Pearson S, Bochtis D (2018) Machine learning in agriculture: a review. *Sensors* 18(8):2674. <https://doi.org/10.3390/s18082674>
- Lizárraga A, Fraga M, Ascasíbar J, González ML (2017) In vitro propagation and recovery of eight apple and two pear cultivars held in a germplasm bank. *Am J Plant Sci* 8(9):2238–2254. <https://doi.org/10.4236/ajps.2017.89150>
- Mahendra PVS, Gupta SD (2004) Trichromatic sorting of in vitro regenerated plants of gladiolus using adaptive resonance theory. *Curr Sci* 10:348–353
- Marques MP, Martins J, de Carvalho LA, Zuzarte MR, da Costa RM, Canhoto J (2021) Study of physiological and biochemical events leading to vitrification of *Arbutus unedo* L. cultured in vitro. *Trees* 35:241–253. <https://doi.org/10.1007/s00468-020-02036-0>

- Mestre D, Fonseca JM, Mora A (2017) Monitoring of in-vitro plant cultures using digital image processing and random forests. 8th International Conference on Pattern Recognition Systems. <https://doi.org/10.1049/cp.2017.0137>
- Mohamed SM, El-Mahrouk ME, El-Banna AN, Hafez YM, El-Ramady H, Abdalla N, Dobránszki J (2023) Optimizing medium composition and environmental culture condition enhances antioxidant enzymes, recovers *Gypsophila paniculata* L. hyperhydric shoots and improves rooting in vitro. *Plants* 12(2):306. <https://doi.org/10.3390/plants12020306>
- Murashige T, Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol Plant* 15(3):473–497. <https://doi.org/10.1111/j.1399-3054.1962.tb08052.x>
- Nezami-Alanagh E, Garoosi GA, Landín M, Gallego PP (2019) Computer-based tools provide new insight into the key factors that cause physiological disorders of pistachio rootstocks cultured in vitro. *Sci Rep* 9(1):1–5. <https://doi.org/10.1038/s41598-019-46155-2>
- Paques M, Boxus P, Dulos M (1985) “Vitrification”: an inducible and reversible phenomenon. In: symposium on in vitro problems related to mass propagation of horticultural plants 212. pp 253–258. <https://doi.org/10.17660/ActaHortic.1987.212.38>
- Pasqualetto PL, Zimmerman RH, Fordham I (1988) The influence of cation and gelling agent concentrations on vitrification of apple cultivars in vitro. *Plant Cell, Tissue Organ Cult* 14(1):31–40. <https://doi.org/10.1007/BF00029573>
- Patrício DI, Rieder R (2018) Computer vision and artificial intelligence in precision agriculture for grain crops: a systematic review. *Comput Electron Agric* 153:69–81. <https://doi.org/10.1016/j.compag.2018.08.001>
- Peterson RA, Peterson MR (2020) Package ‘bestNormalize’. Normalizing transformation functions. R package version
- Phan CT, Letouze R (1983) A comparative study of chlorophyll, phenolic and protein contents, and of hydroxycinnamate: CoA ligase activity of normal and ‘vitreous’ plants (*Prunus avium* L.) obtained in vitro. *Plant Sci Lett* 31(2–3):323–327. [https://doi.org/10.1016/0304-4211\(83\)90071-8](https://doi.org/10.1016/0304-4211(83)90071-8)
- Pinheiro J, Bates D, DebRoy S, Sarkar D, Heisterkamp S, Van Willigen B, Maintainer R (2017) Package ‘nlme.’ Linear Nonlinear Mixed Eff Models Vers 3(1):274
- Prasad VS, Gupta SD (2008) Applications and potentials of artificial neural networks in plant tissue culture. In: Gupta SD, Ibaraki Y (eds) *Plant tissue culture engineering*. Springer, Dordrecht, pp 47–67. [https://doi.org/10.1007/978-1-4020-3694-1\\_3](https://doi.org/10.1007/978-1-4020-3694-1_3)
- Qi Y, Heisler GM, Gao W, Vogelmann TC, Bai S (2010) Characteristics of UV-B radiation tolerance in broadleaf trees in southern USA. In: Gao W, Slusser JR, Schmoldt DL (eds) *UV radiation in global climate change*. Springer, Berlin, Heidelberg
- Redmon J, Divvala S, Girshick R, Farhadi A (2016). You only look once: Unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779–788. <https://doi.org/10.1109/CVPR.2016.91>
- Ripley B, Venables W, Ripley MB (2016) Package ‘nnet’. R Package version. 2;7(3–12):700
- Rojas-Martínez L, Visser RG, De Klerk GJ (2010) The hyperhydricity syndrome: waterlogging of plant tissues as a major cause. *Propag Ornament Plants* 10(4):169–175
- Rstudio Team (2015) RStudio: integrated development for R. RStudio, Inc., Boston, p 879
- Saher S, Fernández-García N, Piqueras A, Hellín E, Olmos E (2005) Reducing properties, energy efficiency and carbohydrate metabolism in hyperhydric and normal carnation shoots cultured in vitro: a hypoxia stress? *Plant Physiol Biochem* 43(6):573–582. <https://doi.org/10.1016/j.plaphy.2005.05.006>
- Sakamoto Y, Ishiguro M, Kitagawa G (1986) Akaike information criterion statistics. D. Reidel, Dordrecht. <https://doi.org/10.2307/2983028>
- Scherer PA (1987) Standardization of plant micropropagation by usage of a liquid medium with polyurethane foam plugs or a solidified medium with the gellan gum gelrite instead of agar. In: *International Symposium on Propagation of Ornamental Plants* 226. pp. 107–114. <https://doi.org/10.17660/ActaHortic.1988.226.10>
- Scherer PA, Müller E, Lippert H, Wolff G (1988) Multielement analysis of agar and gelrite impurities investigated by inductively coupled plasma emission spectrometry as well as physical properties of tissue culture media prepared with agar or the gellan gum gelrite. In: *International Symposium on Propagation of Ornamental Plants* 226, pp 655–658. <https://doi.org/10.17660/ActaHortic.1988.226.91>
- Serrano L, Penuelas J, Ustin SL (2002) Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: decomposing biochemical from structural signals. *Remote Sens Environ* 81(2–3):355–364. [https://doi.org/10.1016/S0034-4257\(02\)00011-1](https://doi.org/10.1016/S0034-4257(02)00011-1)
- Shaw DR, Kelley FS (2005) Evaluating remote sensing for determining and classifying soybean anomalies. *Precision Agric* 6(5):421–429. <https://doi.org/10.1007/s11119-005-3681-9>
- Singh A, Thakur N, Sharma A (2016) A review of supervised machine learning algorithms. In: 2016 3rd international conference on computing for sustainable global development (INDIACom). pp. 1310–1315. <https://doi.org/10.35940/ijscce.E3583.1112522>
- Smith MA, Spomer L (1995) Vessels, gels, liquid media, and support systems. In: Aitken-Christie J, Kozai T, Smith MAL (eds) *Automation and environmental control in plant tissue culture*. Springer, Dordrecht, pp 371–404. [https://doi.org/10.1007/978-94-015-8461-6\\_16](https://doi.org/10.1007/978-94-015-8461-6_16)
- Smith MA, Spomer L, Meyer MJ, McClelland MT (1989) Non-invasive image analysis evaluation of growth during plant micropropagation. *Plant Cell, Tissue Organ Cult* 19(2):91–102. <https://doi.org/10.1007/BF00035809>
- Sommer C, Straehle C, Koethe U, Hamprecht FA (2011) Ilastik: Interactive learning and segmentation toolkit. In: 2011 IEEE international symposium on biomedical imaging: From nano to macro. pp. 230–233. <https://doi.org/10.1109/ISBI.2011.5872394>
- Sullivan C, Kaszynski A (2019) PyVista: 3D plotting and mesh analysis through a streamlined interface for the Visualization Toolkit (VTK). *J Open Sour Softw* 4(37):1450. <https://doi.org/10.21105/joss.01450>
- Terry ME, Bonner BA (1980) An examination of centrifugation as a method of extracting an extracellular solution from peas, and its use for the study of indoleacetic acid-induced growth. *Plant Physiol* 66(2):321–325. <https://doi.org/10.1104/pp.66.2.321>
- Thiel M (2018) Bildgebende NIR-Hyperspektral-Technologie zur in-situ Erfassung des Blattwassergehalts. Doctoral dissertation, Leibniz Universität Hannover. <https://doi.org/10.15488/3882>
- Tian J, Jiang F, Wu Z (2015) The apoplastic oxidative burst as a key factor of hyperhydricity in garlic plantlet in vitro. *Plant Cell Tiss Organ Cult* 120(2):571–584. <https://doi.org/10.1007/s11240-014-0623-0>
- Tisserand S (2021) Vis-NIR hyperspectral cameras. *Photoniques* 110:58–64
- Tsay HS, Lee CY, Agrawal DC, Basker S (2006) Influence of ventilation closure, gelling agent and explant type on shoot bud proliferation and hyperhydricity in *Scrophularia yoshimurae*—a medicinal plant. In *Vitro Cell Dev Biol-Plant* 42(5):445–449. <https://doi.org/10.1079/IVP2006791>
- van Altvorst AC, Koehorst H, de Jong J, Dons HJ (1996) Transgenic carnation plants obtained by *Agrobacterium tumefaciens*-mediated

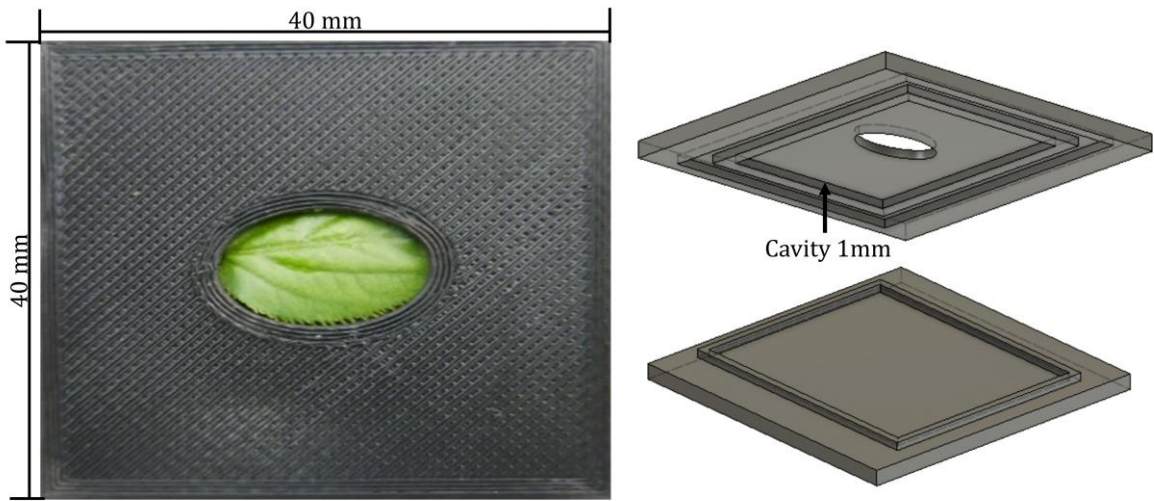
- transformation of petal explants. *Plant Cell, Tissue Organ Cult* 45(2):169–173. <https://doi.org/10.1007/BF00048762>
- van den Dries N, Gianni S, Czerednik A, Krens FA, de Klerk GJ (2013) Flooding of the apoplast is a key factor in the development of hyperhydricity. *J Exp Bot* 64(16):5221–5230. <https://doi.org/10.1093/jxb/eru497>
- Van Der Walt S, Colbert SC, Varoquaux G (2011) The NumPy array: a structure for efficient numerical computation. *Comput Sci Eng* 13(2):22–30. <https://doi.org/10.48550/arXiv.1102.1523>
- Van Rossum G, Drake FL (2009) Python 3 Reference manual: python documentation manual part 2. Scotts Valley, CA: CreateSpace
- Vieitez AM, Ballester A, San-José MC, Vieitez E (1985) Anatomical and chemical studies of vitrified shoots of chestnut regenerated in vitro. *Physiol Plant* 65(2):177–184. <https://doi.org/10.17660/ActaHortic.1987.212.34>
- Voeselek LA, Blom CW (1989) Growth responses of *Rumex* species in relation to submergence and ethylene. *Plant Cell Environ* 12(4):433–439. <https://doi.org/10.1111/j.1365-3040.1989.tb0159.x>
- Vreeburg RA, Benschop JJ, Peeters AJ, Colmer TD, Ammerlaan AH, Staal M, Elzena TM, Staals RH, Darley CP, McQueen-Mason SJ, Voeselek LA (2005) Ethylene regulates fast apoplastic acidification and expansin A transcription during submergence-induced petiole elongation in *Rumex palustris*. *Plant J* 43(4):597–610. <https://doi.org/10.1111/j.1365-3113X.2005.02477.x>
- Zhang C, Timmis R, Hu WS (1999) A neural network based pattern recognition system for somatic embryos of Douglas fir. *Plant Cell, Tissue Organ Cult* 56:25–35. <https://doi.org/10.1023/A:1006287917534>
- Zhou QY, Park J, Koltun V (2018) Open3D: a modern library for 3D data processing. arXiv Preprint. <https://doi.org/10.48550/arXiv.1801.09847>
- Ziv M (1991) Vitrification: morphological and physiological disorders of in vitro plants. In: Debergh PC, Zimmerman RH (eds) Micropropagation. Springer, Dordrecht, pp 45–69. [https://doi.org/10.1007/978-94-009-2075-0\\_4](https://doi.org/10.1007/978-94-009-2075-0_4)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Supplementary Information

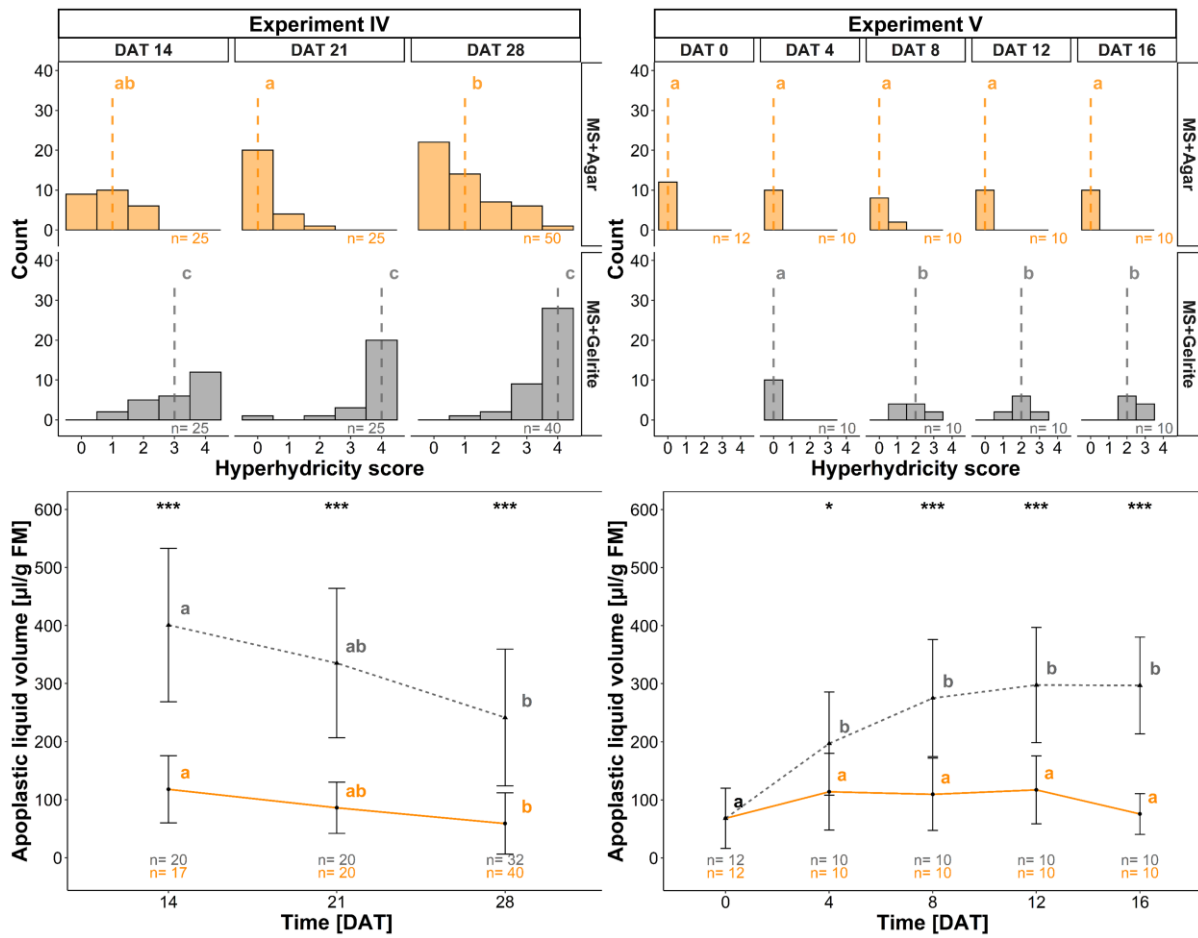


**SI. 1: Maximum shoot height and shape analysis** of explants of *A. thaliana* Col-0 and *Malus* 'G214' (Mean  $\pm$ SD). A) The relative increase in mean maximum shoot height of upper 10<sup>th</sup> percentile resulted from analysis of segmented depth data collected with a scanning laser distance sensor and normalized to day 0 maximum plant height. B) Shape analysis parameter of single explants. Yellow lines indicates cultivation on standard media formulation on Gamborg-B5 (*A. thaliana*) and MS-Medium (*Malus*) solidified with 0.8% agar (w/v), while blue lines display the cultivation on induction media containing 0.25% (w/v) gelrite, inducing HH. Estimated 95<sup>th</sup> confidence interval was illustrated in light gray. C) Representative images of segmented single explants of the different plant species and treatments. Convex hull, longest path, centroid is colored in pink, while object area and perimeter are outlined in blue. Sample number (n) indicates A) individual culture containers and B) individual explants. Significance stars indicate comparisons of treatments within a time point (day) with \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . Depth data were acquired with the multisensory robot system "Phenomenon" (Bethge et al. 2023).



**SI. 2: Sample holder for reflectance measurement** with Perkin-Elmer Lambda 900 UV-VIS-NIR-SWIR spectrometer. Leaf samples were placed with adaxial orientation inside the 3D-printed sample holder.

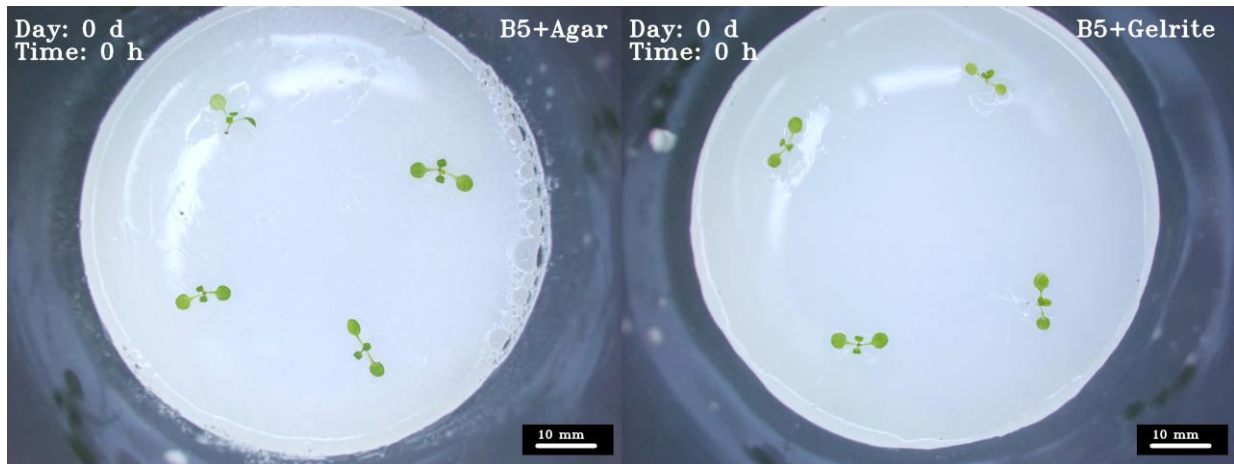




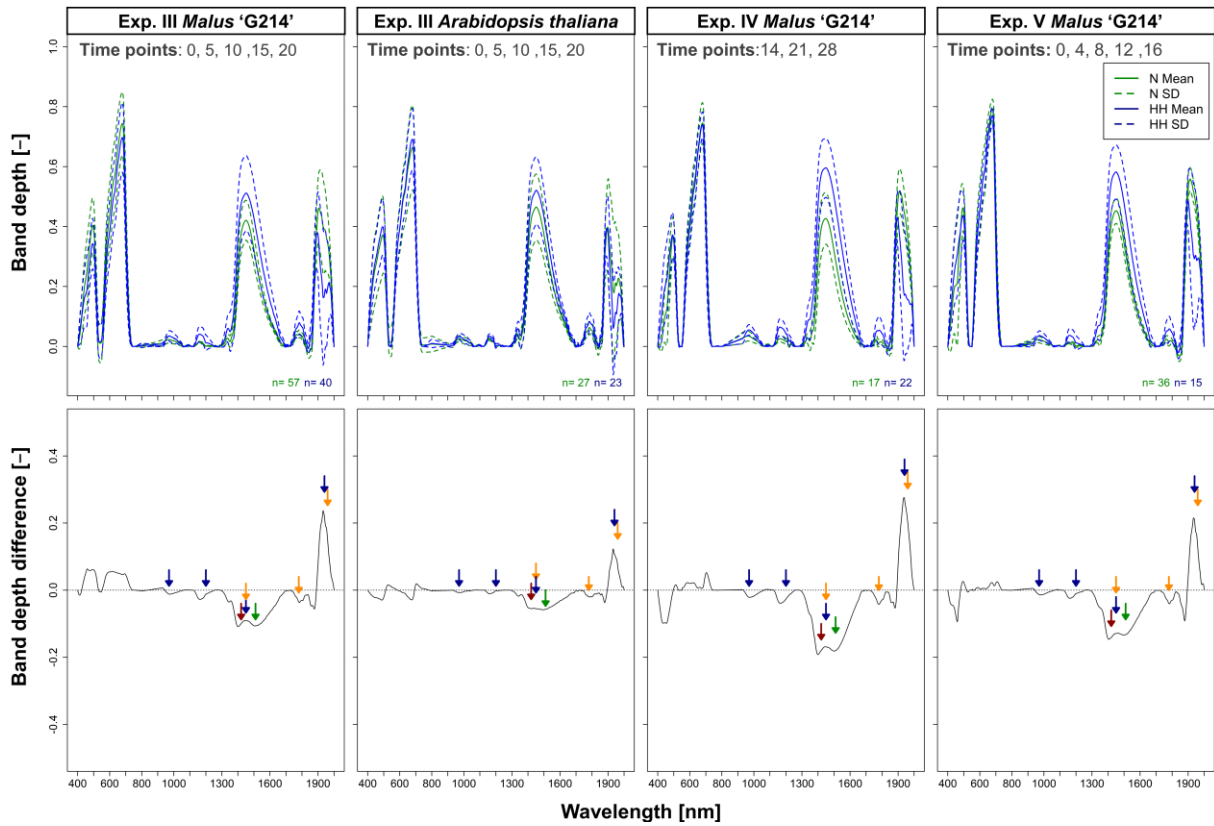
**SI. 3: Visual scoring and apoplastic liquid volume of *Malus* 'G214' Experiment IV & V.** Samples from 0 days after transfer (DAT 0) represent the starting plant material cultured on control media. Yellows bars indicate cultivation on standard media formulation MS-Medium with 0.8% (w/v) agar, while gray bars display the cultivation on induction media containing 0.25% (w/v) gelrite. Dashed lines represent the medians of each histogram. Sample number (n) indicates the individual explants. Different letters resulting from Kruskal-Wallis-Test followed by Fisher's LSD ( $P < 0.05$ ) indicate significant differences between histograms. Different letters resulted from Tukey's HSD test at  $P < 0.05$  and show significantly different time points within one treatment, while asterisks indicate comparisons of treatments within a time point with \* =  $P < 0.05$ , \*\* =  $P < 0.01$ , \*\*\* =  $P < 0.001$ .



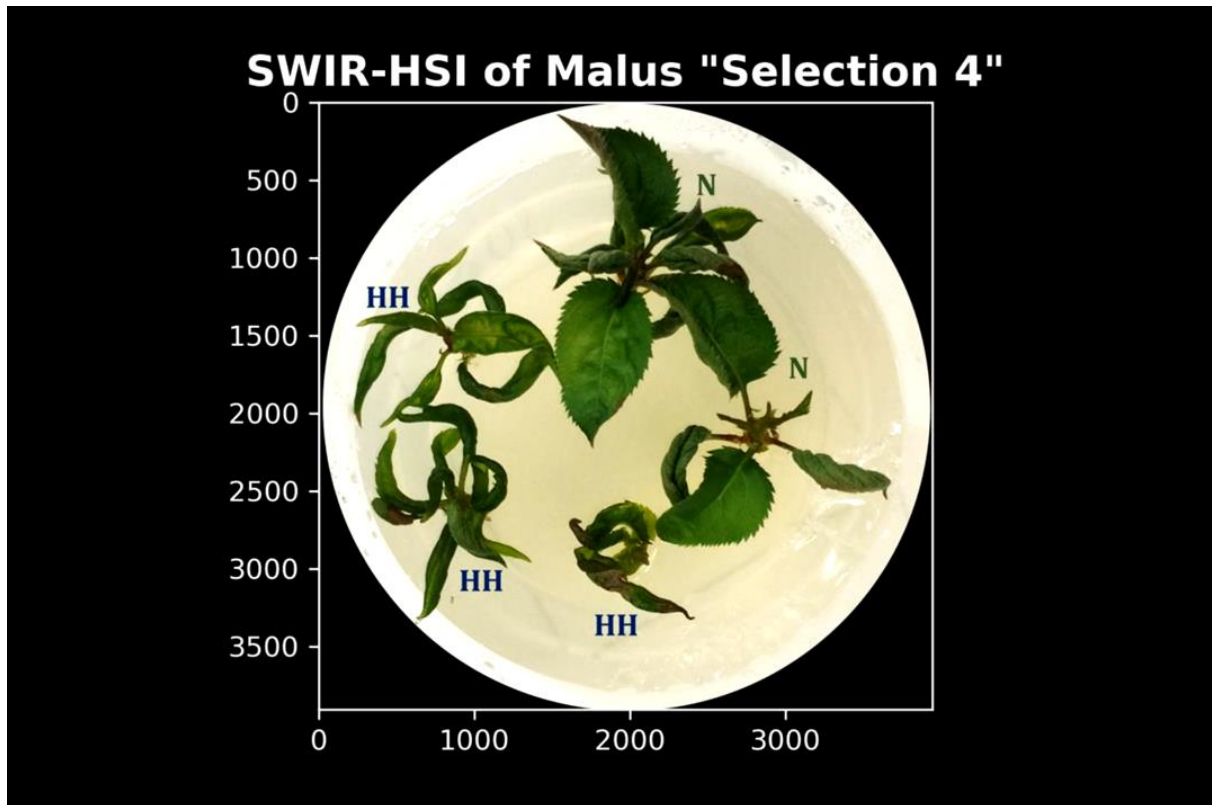
**SI. 4: Hyperhydricity induction time-lapse video in Malus 'G214'.** Left side represents control MS-Medium solidified with 0.8% agar (w/v), while right image displays the cultivation on induction media containing 0.25% (w/v) gelrite. Both media were supplemented with  $1 \text{ g L}^{-1}$  titanium dioxide. RGB Images were acquired with the multisensory robot system "Phenomenon" (Bethge et al. 2023).



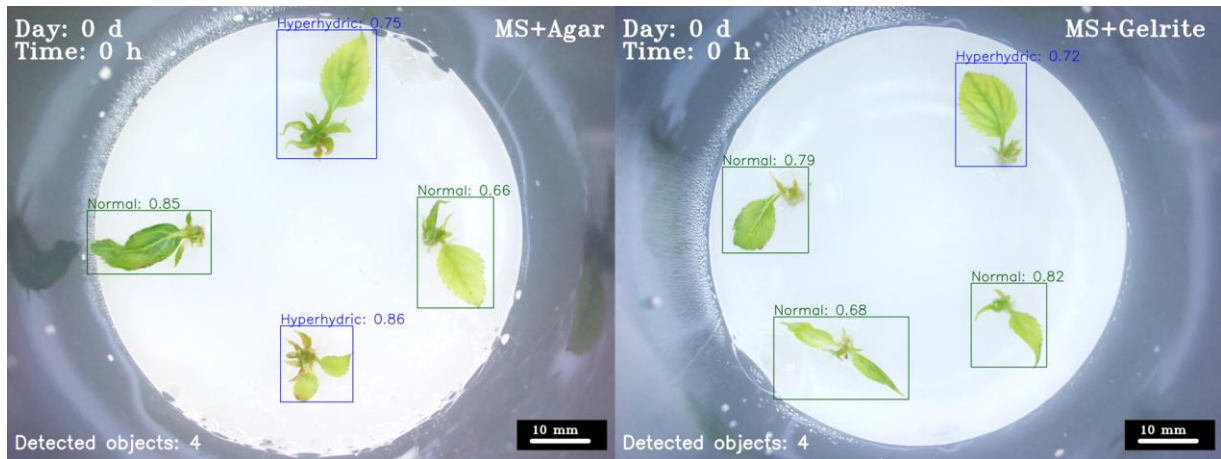
**SI. 5: Hyperhydricity induction time-lapse video in Arabidopsis.** Left side represents control B5-Medium solidified with 0.8% agar (w/v), while right image displays the cultivation on induction media containing 0.25% (w/v) gelrite. Both media were supplemented with  $1 \text{ g L}^{-1}$  titanium dioxide. RGB Images were acquired with the multisensory robot system “Phenomenon” (Bethge et al. 2023).



**SI. 6: Similarities in spectral reflectance of *Malus sp.* “G214” and *A. thaliana* “Col-0”.** Upper row represents extracted absorption features after segmented convex-hull removal of raw spectra and bottom row demonstrates difference spectrum of the absorption peaks. Normal (NN, colored in green) and hyperhydric (HH, colored in blue) spectra membership were based on visual scoring of HH (NN, 0-1; HH 2-4). Colorized arrows indicating potential major biochemical leaf compounds absorbing in the given wavelength region, according to Curran, 1989. Reflection spectra were measured with UV/VIS/NIR Spectrometer (PerkinElmer Lambda 950) in wavelength range of 200 – 2000 and a resolution of 1 nm.



**SI. 7: SWIR-Hyperspectral image stack/video of Malus 'Selection 4'.** Image data set were acquired with the imaging system developed and described by [Thiel \(2018\)](#). This system consisted of an EVK Helios Core NIR Line-scan camera (240 px × 1 px and 252 spectral channels in the wavelength region of 900 nm to 1700 nm), two 65 W halogen spot lights and a conveyer-belt system to move the sample. Spectral reflectance was recorded top-down through the lids of the culture containers.



**SI. 8: Automated detection of Hyperhydricity via CNN, while the induction of HH in *Malus* 'G214' in time-lapse video.** Left side represents control MS-Medium solidified with 0.8% agar (w/v), while right image displays the cultivation on induction media containing 0.25% (w/v) gelrite. RGB images were acquired with the multisensory robot system "Phenomenon" (Bethge et al. 2023). Class membership predictions resulted from the trained PCTOC\_V3 object detection model, available under Bethge (2023).

## 2.3 Remote plant sensing and phenotyping – an e-learning tool in higher education

Hans Bethge<sup>1,2</sup>, Thomas Mählmann<sup>1</sup>, Traud Winkelmann<sup>2</sup>, Thomas Rath<sup>1</sup>

<sup>1</sup>Laboratory for Biosystems Engineering, Faculty of Agricultural Sciences and Landscape Architecture, Osnabrück University of Applied Sciences, 49090 Osnabrück, Germany

<sup>2</sup>Institute of Horticultural Production Systems, Section of Woody Plant and Propagation Physiology, Leibniz Universität Hannover, Herrenhäuser Str. 2, 30419, Hannover, Germany

Type of authorship:	First author
Type of article:	Long Paper Proceeding
Status of article:	Published
Contribution to the article:	Analysed the data. Prepared the figures and wrote the manuscript.
Journal:	Lecture Notes in Informatics (LNI), Proceedings - Series of the Gesellschaft für Informatik in der Landwirtschaft (GIL)
Impact factor:	0.19 (LNI-GI)
Acknowledgment:	Reproduced with permission from “Gesellschaft für Informatik in der Land-Forst- und Ernährungswirtschaft e.V.”

## Remote plant sensing and phenotyping – an e-learning tool in higher education

Hans Bethge <sup>1</sup>, Thomas Mählmann<sup>2</sup>, Traud Winkelmann <sup>3</sup> and Thomas Rath <sup>4</sup>


**Abstract:** Within the consortium “Experimentation Field Agro-Nordwest”, a practical concept for knowledge and technology transfer of digital competence in agriculture was created. For this purpose, the web-based e-learning system “SensX” was set up, consisting of videos, presentations and instructions. In addition, the classical e-learning concept was extended by data sets, student experiments and sensor data of plants acquired by a remote phenotyping robot. This resulted in a massive open online course (MOOC), which was tested with agricultural and biotechnology students in higher education at the University of Applied Sciences Osnabrück over two years. The evaluation process of “SensX” included an empirical survey, qualitative interviews of the participating students by an external institution and an evaluation of the concept by the lecturers.

**Keywords:** agriculture, digital competence, e-learning concepts, remote experiments, sensors in teaching


### 1 Introduction


In higher education, the number of teaching modules based on e-learning systems, blended learning systems (traditional teaching combined with e-learning) or MOOCs (massive open online courses) is steadily increasing, both nationally (Germany) and internationally [Lü20; A118]. The majority of these (approx. 30%) are offered in the computer sciences, but almost 6% are part of the agricultural and life sciences curricula, worldwide [A118]. The COVID19 pandemic has caused further acceleration in the use of e-learning, blended learning approaches or MOOCs in higher education [Be21]. Yet the terms used to describe the use of computer technologies in education (here e-learning, blended learning, MOOCs) vary widely and are not coherent. In many publications and reports, the methods are

---

<sup>1</sup>University of Applied Sciences Osnabrueck, Laboratory for Biosystems Engineering (BLab), Oldenburger Landstr. 24, D-49090 Osnabrueck, h.bethge@hs-osnabrueck.de,  <https://orcid.org/0000-0002-3487-9725>

<sup>2</sup>University of Applied Sciences Osnabrueck, Laboratory for Biosystems Engineering (BLab), Oldenburger Landstr. 24, D-49090 Osnabrueck, thomas.maehlmann@hs-osnabrueck.de

<sup>3</sup>Leibniz Universität Hannover, Institute of Horticultural Production Systems, Woody Plant and Propagation Physiology Section, Herrenhaeuser Str. 2, D-30419 Hannover, traud.winkelmann@zier.uni-hannover.de,  <https://orcid.org/0000-0002-2509-1418>

<sup>4</sup>University of Applied Sciences Osnabrueck, Laboratory for Biosystems Engineering (BLab), Oldenburger Landstr. 24, D-49090 Osnabrueck, t.rath@hs-osnabrueck.de,  <https://orcid.org/0000-0001-7277-7335>



grouped under terms such as "digital education" or "e-learning". Defined competences should be specified and consolidated. In contrast to the original intention of the Bologna process to focus academic education very strongly on vocational training, digital education in particular is based on four (or even more) areas of competence acquisition: (i) self-competence: ability to act on one's own responsibility, (ii) subject-matter competence: the ability to be able to make judgments and take action in specific areas, (iii) time competence: the ability to plan and carry out actions and intellectual achievements in a chronological sequence, (iv) social competence: ability to make judgements and act in a complex society [Ar18]. When evaluating digital education with regard to the criteria and competences listed above, digital education is sometimes seen as having a disruptive character [Ki19], since universal access makes it possible to support any university. Thus knowledge and methods are, at least in theory, available to many institutions and locations. On the other hand, difficulties are also seen that ultimately cause high dropout rates, e.g. with MOOCs [Ki19]. Empirical studies do not show a uniform picture of the educational success of digital higher education. In an analytical-theoretical study, it was demonstrated that there is only a slight correlation between digital and classic teaching methods in higher education with regard to learning success [Sc20]. In contrast, specific digital evidence-based and tested teaching concepts in the agricultural sector showed the superiority of digital teaching concepts to classical teaching in terms of learning success [We22; Ke16].

Of course, it should also be noted that the use and success is largely dependent on the users, i.e. the learners themselves. Thus, Kahan et al. differentiated users of MOOCs into five categories: (i) tasters, (ii) downloaders, (iii) disengagers, (iv) offline engagers and (v) online engagers [Ka17]. Each of these groups handles digital education differently and presumably, this leads to strong dispersion in learning success or in other evaluation parameters of digital teaching. On closer inspection, the teaching concepts and methods used in digital education or e-learning are also complex, and the terms used for them are multifaceted, not clearly demarcated, and only partially defined (see the comprehensive table of terms in [Lü20]). [Ca20] distinguished only three teaching approaches: (i) e-learning by distributing, (ii) e-learning by interacting and (iii) e-learning by collaboration. They were able to show that e-learning in higher education is still largely dominated by e-learning by distribution (uploading texts, graphics, PDFs, etc.), while more far-reaching approaches such as e-learning by interacting or e-learning by collaboration often remain largely unexplored [Ca20]. Approaches that go beyond this, such as e-learning supported by self-performing experiments, do not appear at all in the considerations and therefore, seem to be so far unconsidered in digital education and literature. The need to develop educational concepts that go beyond the three approaches of [Ca20] in order to provide efficient and successful academic training in as many areas of competence as possible is becoming increasingly evident.

In agriculture, the field of sensor technologies is very suitable for this purpose, as it is a subject with constantly evolving contents. Additionally, it is gaining increasing importance in all areas of agribusiness, and already plays a dominant role in practice, research, and development (see [Ha19] and [Yi21]). Moreover, students of agricultural sciences usually have little affinity for sensor technologies prior to their studies. So for

many of them, a deeper engagement in sensor topics with new e-learning methods can lead to new knowledge, skills and competences. Therefore, the aim of our project is to establish and explore the use of sensor technology in agriculture using an interactive e-learning approach in order to provide students with in-depth digital competence.

## 2 SensX

### 2.1 Concept and categories

Within the consortium, “Experimentation Field Agro-Nordwest” a proof-of-concept project called “SensX” was initiated to promote and establish the use of sensor technology in the plant sector at university and college level. For that purpose, a MOOC e-learning system was developed that extends classical e-learning concepts by using sensor data, sensor kits, and, as a future perspective, fully remote and collaborative teaching with a robotic sensor demonstrator (Fig. 1). We have structured the features of the e-learning system into four categories, as described in detail in the following subchapters.

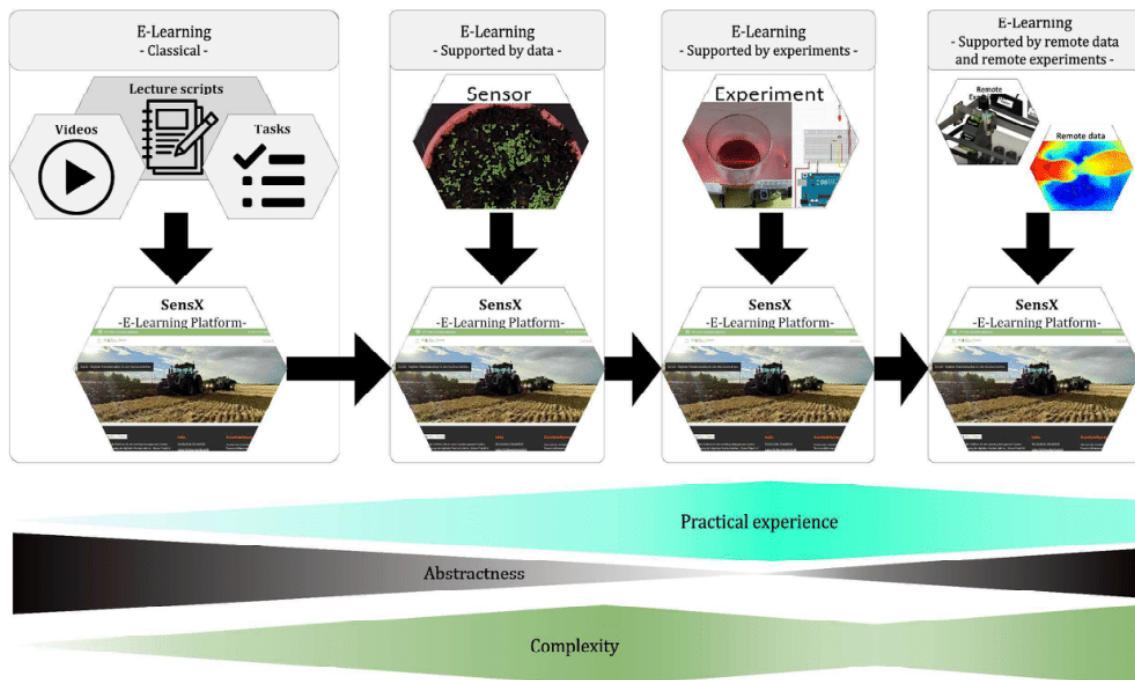


Fig. 1: Supported e-learning concepts of the MOOC platform “SensX”

### Classical e-learning

Up to now, SensX contains 12 sessions (see Tab. 1 at the end of this chapter), four of which were structured according to basic/classical e-learning concepts following e-learning by distribution. The transfer of knowledge took place through the exchange of information by uploading lecture scripts, explanatory videos and exercises about

agricultural engineering topics, for instance operating principles of specific sensors used in crop production and research, such as ultrasonic and spectral sensors, RGB, NIR and IR cameras, fluorescence spectroscopy, LIDAR (light detection and ranging), soil moisture, temperature and humidity sensors.

### **E-learning supported by sensor data**

Two of the 12 SensX sessions (see Tab. 1) were conceptualized according to e-learning supported by RGB sensor data sets to demonstrate the measurement of spatially resolved data and their processing with current approaches of artificial intelligence. Based on this, image-processing algorithms were developed to evaluate the generated feature vectors with a neural network.

### **E-learning supported by sensor experiments**

Since SensX is a hands-on online course, each participant received a sensor kit (worth about € 30) for the session categorized as e-learning supported by experiments at the beginning of the module. The kit consisted of a microcontroller platform, electronic components and various sensors. This enabled the participants to carry out their own experiments at home and to collect data of plants and other real objects. In addition, for specific tasks, participants carried out exercises with smartphone-based sensors.

### **E-learning supported by remote sensor data and remote sensor experiments**

In order to also enable multisensory applications at a high academic level, the novel low-cost demonstrator “Phenomenon” was developed [Be22], by which different sensory information from real plants was obtained (Fig. 2). It consists of exclusively low-cost hardware and open-source software components, which were selected to construct a xyz-scanning system with an adequate accuracy for consistent data acquisition and total costs of around € 3000. The developed device allows remote control via HTTP of all the functions such as motion control, data acquisition and access of sensor data due to its unique software design. We have installed four different sensors inside the robot, (RGB and thermal camera as imaging sensors, laser-based depth sensor and spectrometer as point measuring sensor) that correspond to the current sensor technologies used in modern agriculture. The resulting data sets provide students with multisensory data acquired remotely from real phenotypic experiments that they learned to handle and process. We already included the resulting data sets of the system in two modules of SensX.

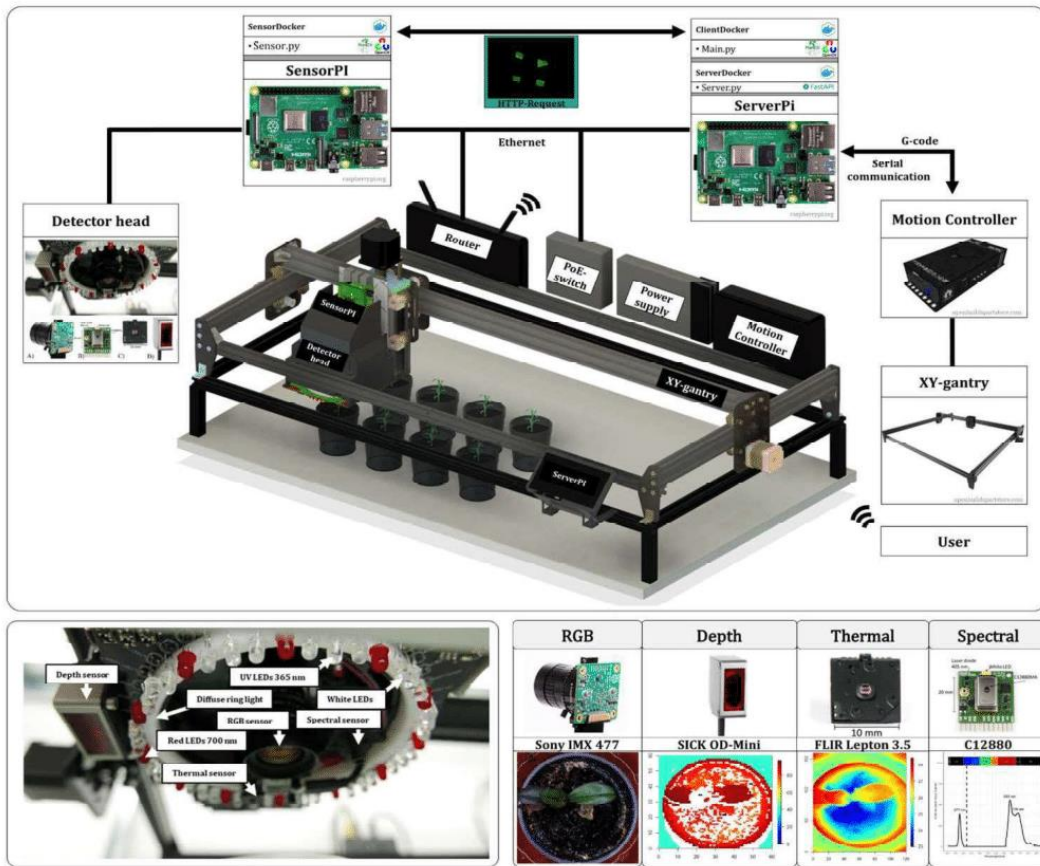


Fig. 2: Demonstrator “Phenomenon” for multisensor technology

Agrotechnical content	Category	Hardware
Analog sensor data acquisitions	Basic/Classical	PC
Thermodynamics in greenhouses	Basic/Classical	PC
CO <sub>2</sub> -tracer gas method	Basic/Classical	PC
Computer vision (CV) and machine-learning	Basic + data sets	PC + RGB data
Plant classification, CV and neural networks	Basic + data sets	PC + RGB data
Optoacoustic signals	Basic + experiments	PC + Microcontroller
Radiometry & spectroscopy	Basic + experiments	PC + Smartphone
Temperature and spectral data acquisition, analysis and visualization with R/Rstudio	Basic + (rem.) exp. + (remote) data sets	PC + Demonstrator + MCont. + sensor data

Tab. 1a: Sessions and categories in SensX for *Sensor control and analysis*

Agritechnical content	Category	Hardware
Randomization of experiments in controlled environment	Basic/Classical	PC
Moisture sensor data acquisition	Basic + experiments	PC + Microcontroller
Determination of crop performance traits	Basic + experiments	PC + Smartphone
Spectroscopic Methods and data visualization	Basic + (remote) experiments + (remote) data sets	PC + Demonstrator and Microcontroller sensor data

Tab. 1b: Sessions and categories in SensX for *Applied technology in crop experimentation*

## 2.2 Evaluation methods and statistics

The 12 application sessions of SensX were evaluated internally with surveys within two university lectures (“Sensor control and analysis” with 18 participants and “Applied technology in crop experimentation” with 16 participants) addressing agriculture and biotechnology students over two semesters (winter 20/21 and summer 21) at the University of Applied Sciences Osnabrück. Within the surveys, the students have to rank a) the different sessions in general (scale: very good, good, average, bad, very bad) and b) the difficulty level of the sessions (scale: very easy, easy, average, heavy, extreme heavy). Because of different semesters, different survey years, different studies and different prior knowledge of the students, the results were not statistically condensed. Additionally, the students were asked to rank every session, which was integrated into their course, with a German grading scale from 1 (excellent) to 6 (very poor). These rankings were combined with categories (see above) and statistically analysed with Kruskal-Wallis rank sum test with Fisher’s Least Significant Difference using R.

In the following year (summer semester 22, “Sensor control and analysis” with 9 participants), the courses were evaluated externally as qualitative interview by the Institute for Futures Studies and Technology Assessment IZT, Berlin.

## 3 SensX evaluation

### 3.1 Internal empirical survey

The internal empirical survey with the participating students revealed that over 90% rated SensX as very good, good or average. However, only 65-70% of the students were satisfied with the clarity and design of the Moodle-based web interface of SensX (data not shown). Participants were highly satisfied with subject-specific content of SensX in general, regardless of the teaching concept (Fig. 3) and there was no significant difference in the overall rating of SensX in terms of the defined teaching concepts (Tab. 2).

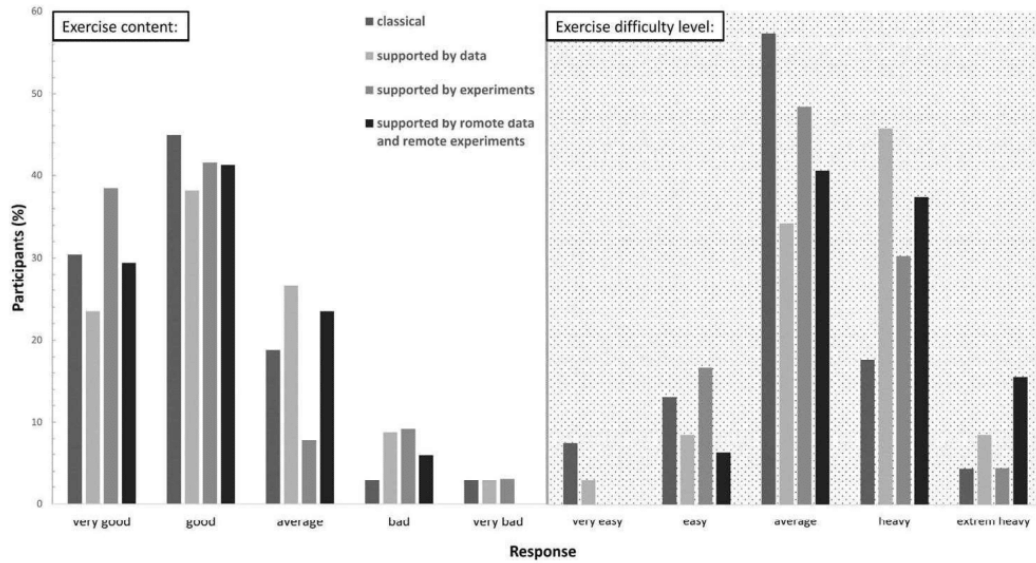


Fig. 3: Internal evaluation of SensX by participants

Concepts	Mean±sd	Median	Count	Rank
Classical	2.13±0.95a	2	70	94.6
Supported by data	2.34±1.19a	2	35	103.1
Supported by experiments	2.27±1.13a	2	66	99.4
Sup. by rem. data and rem. experiments	2.68±1.25a	2	31	119.6

(German grading scale, Kruskal-Wallis rank sum test with Fisher's Least Significant Difference,  $\alpha = 0.05$ )

Tab. 2: Statistical summary of participants' overall rating of SensX by in the internal evaluation

Further analyses by the students regarding the developed sessions showed that there were no difficulties with regard to the usability and the clarity of the presentation of the individual subject contents. This was of great importance, as it was the basic prerequisite for quantitatively evaluating the students' further statements regarding the sessions. In terms of difficulty level, the optimum was exceeded for some participants, in particular when complexity and abstractness (Fig. 1) were increased due to the teaching method. Nevertheless, that also indicated the need for specific teaching tools addressing those skills. This was particularly evident in the sessions with remote data of the “Phenomenon” robot, where 70% of the participants rated the content of the sessions as very good to good (Fig. 3), but at the same time 70% of the participants rated the difficulty level as heavy to very heavy.

The motivation of the students to engage with the sessions of SensX as well as the results achieved in the exams were rated as very good by the lecturers. These evaluations showed that especially the e-learning sessions, which were based on the analysis of data and simple experiments, received very good ratings from the students. The more complex the applications became, the more critically the students rated the individual sessions (see Tab. 2). These evaluations are understandable, as individual modules with a high degree of complexity demanded a lot from the students, especially methodological and procedural expertise and competence orientation.

### **3.2 External qualitative interview by IZT**

The post hoc analysis of the interviews, conducted and reported by the IZT, revealed that participants rated SensX an average of 7.9 in terms of overall satisfaction on a scale from 1 (low satisfaction) to 10 (high satisfaction). It should be noted that this survey was conducted a year later with other participants of the module and slightly optimized contents of the sessions. The students surveyed the gain in their own digital competence particularly positive being achieved through the independently performed experiments with microcontrollers and sessions regarding computer vision. However, the surveys also indicated that there is a strong desire for more collaboration.

Representative statement from students:

- "I would definitely also say that I was able to take away a lot of digital skills in this module and that will continue."
- "What disappointed me a little bit: I thought we would be standing together more in the greenhouse looking at plants."

### **3.3 Evaluation by lecturers**

In the first six columns, Figure 4 shows the main characteristics and criteria of the developed sessions in relation to the four e-learning categories (lines) that were considered and implemented in the module design.

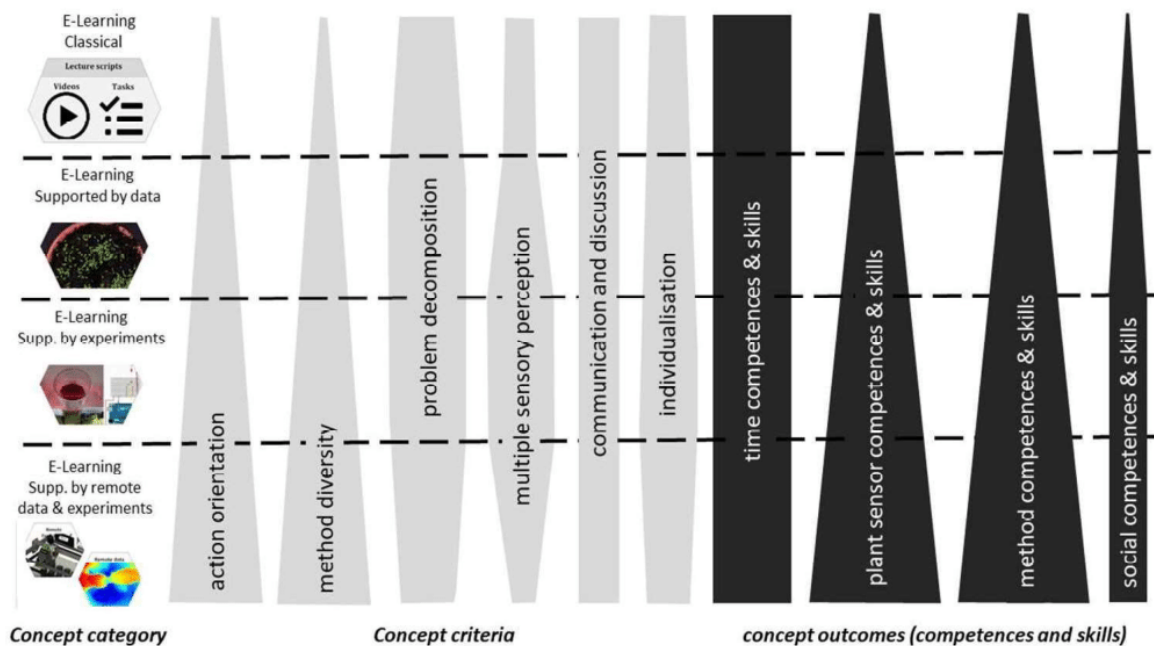


Fig. 4: Criteria of SensX sessions based on experiences of module creators and lecturers and the competence output profile within two years (four semesters) of application

It should be emphasized that the information on the module properties in Figure 4 refers exclusively to the sessions created and the concepts behind them. Nevertheless, Figure 4, columns 1 to 6 provide important information about the general concept of SensX. In contrast, the last four columns represent the actual competences achieved within the sessions based on the evaluations (see chapters 3.1 & 3.2), but also based on the oral and written examination results and the statements made by the students. A summary of the different evaluations indicated that the new e-learning concept supported by experiments and data analysis mainly raised plant sensor expertise and methodological competence and skills of the participants.

### 3.4 General discussion

The SensX system presented in this study goes beyond the **classical e-learning concepts** [see Ca20] and uses technical tools that meet today's digital potential and requirements for higher education. Our teaching concept successfully demonstrated the proof-of-concept that an extension of classical e-learning systems by action-oriented methods in the field of agricultural engineering is possible and reasonable.

However, the extension of e-learning requires considerable additional technical and financial effort for higher education institutions. The hardware used (microcontrollers, sensors) must be provided as a kit to each student for home work and should be in the low-cost range so that students can **experiment** with it freely and without prior knowledge.



The additional more complex systems required, such as the “Phenomenon” robot in our case, which enables remote experiments at the educational institution, must also meet low-cost conditions. Our “Phenomenon” system fulfils this criterion, costing less than € 3000 [Be22]. It should be noted, however, that remote access to “Phenomenon” requires an open IT structure of the educational institution that is freely accessible from outside. Usually this is not the case, or only to a limited extent. Also in our case, we had to provide students with the data collected on site without direct students’ access to the sensor system. Although this limitation can be reduced by video presentations (in our case) or similar documentation, a free **remote access** from outside is actually necessary, or at least desirable, for the complete implementation of our teaching concept. We were therefore only able to show the perspective of our remote experimental concept, where students will be taught how to remotely program, control and use a remote robot to phenotype a small biological experiment through a collaborative online programming workshop. We are convinced that this will address the identified students’ demand for collaborative work.

At this point, we would like to once again highlight the different e-learning types of students [Ka17]. Possibly, here lies a problem of the created system: due to the high technology employment, for individual students, e.g. the group designated as downloaders, the study within a technical unknown field (developing circuits, calling sensor data with programmed scripts etc.) may need too much time and self-initiative so that no new knowledge or new method competences can be gained. On the other hand, the complex sessions forced students to interact with each other, which is otherwise considered difficult and critical in e-learning systems [Ca20]. Despite the very high content-related, methodical and self-organizational requirements, the students rated our new e-learning concept as efficient and useful in the overall evaluation. Understandably, sessions with their own (easy) experiments tended to score best. Analysis of **complex remote data sets** was rated as very difficult and was obviously at the learning limit of the students involved. Nevertheless, these sessions were generally not devalued significantly, supporting that an appropriate methodological teaching approach was chosen. A direct comparison of the final grades of SensX participants with the final grades of students from the previous five years, which had the same content but no e-learning components, showed a better final grade on average (data not shown). In the evaluations of our overall concept, it is essential to take into account that the newly designed teaching units have so far only been used in the period affected by the COVID19 pandemic, where the boundary conditions were difficult due to the temporary closure of the university and were uncharted territory for all involved. The clarification of possible correlations is still pending and can only be answered after several runs of SensX in different study programs and at different educational institutions.

## 4 Conclusions

In this article, we have demonstrated the feasibility of a new e-learning approach that addresses the need of modern agriculture for high digital competences in higher education. We propose that a deep understanding of sensor technologies and methods of digital data processing can only be obtained when higher education, which has been dominated by e-

learning concepts in the last 2 years due to pandemic circumstances, includes hands-on sessions and is supported by self-performing experiments. Further expansion of SensX will accommodate even more collaborative opportunities for participants and forces to prepare tomorrow's plant sciences students for the challenges of digitized modern agriculture.

**Acknowledgements:** We would like to thank the participating students of University of Applied Sciences Osnabrück and the Federal Ministry of Food and Agricultural (BMEL, Grant No: 28DE103F18).

## Bibliography

- [Ar18] Arnold, P.; Kilian, L.; Thilloßen, A.; Zimmer, G.M.: Handbuch e-learning: Lehren und lernen mit digitalen Medien. Vol. 4965. UTB, 2018.
- [Al18] Al-Rahmi, W.; Aldraiweesh, A.; Yahaya, N.; Bin Kamin, Y.: "Massive open online courses (MOOCs): Systematic literature review in Malaysian higher education." *International Journal of Engineering & Technology* 7, no. 4 (2018): 2197-2202.
- [Be21] Bedenlier, S.; Händel, M., Kammerl, R; Gläser-Zikuda, M.; Kopp, B.; Ziegler, A.: „Akademische Mediennutzung Studierender im Corona-Semester 2020: Digitalisierungsschub oder weiter wie bisher?." *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung* 40 (2021): 229-252.
- [Be22] Bethge, H; Winkelmann, T.; Lüdecke, P.; Rath, T.: Low-cost and automated phenotyping system "Phenomenon" for multi-sensor in situ monitoring in plant in vitro culture. Manuscript submitted, September 2022 in *Plant Methods*.
- [Ca20] Cammann, F.; Hansmeier, E., Gottfried K.: Möglichkeiten und Szenarien einer durch digitale Medien gestützten Lehre – zentrale Tendenzen des aktuellen E-Learning-Einsatzes im Hochschulsektor. In: *Vom E-Learning zur Digitalisierung - Mythen, Realitäten, Perspektiven. Medien in der Wissenschaft, Band 76*, Waxmann Verlag, Münster <https://doi.org/10.31244/9783830991090>; 2020.
- [Ha19] Halachmi, I.; Guarino, M.; Bewley, J.; Pastell, M.: "Smart animal agriculture: application of real-time sensors to improve animal well-being and production." *Annu. Rev. Anim. Biosci* 7, no. 1: 403-425, 2019.
- [Ka17] Kahan, T.; Soffer, T.; Nachmias, R.: "Types of participant behavior in a massive open online course." *International Review of Research in Open and Distributed Learning: IRRODL* 18, no. 6: 1-18, 2017.

- [Ke16] Kersebaum, A.: „Bewertung von hochschulübergreifendem eLearning unter Berücksichtigung von statischer Lerner-Adaptivität und Lernstilen.“ PhD diss., Hannover: Gottfried Wilhelm Leibniz Universität Hannover, 2016.
- [Ki19] Kirchner, K.; Lemke, C.: „MOOCs als disruptive Innovation für die akademische Bildung.“ In *Hochschulen in Zeiten der Digitalisierung*, pp. 239-263. Springer Vieweg, Wiesbaden, 2019.
- [Lü20] Lübecke M., Wannemacher K.: Digitalisierung ohne Wandel? Der hochschuldidaktische Diskurs in Schlüsseljournals. In: Bauer R., Hafer J., Hofhues S., Schiefner-Rohs M., Thillosen A., Volk B., Wannemacher K. (Hrsg.) 2020: *Vom E-Learning zur Digitalisierung - Mythen, Realitäten, Perspektiven. Medien in der Wissenschaft*, Band 76, Waxmann Verlag, Münster <https://doi.org/10.31244/9783830991090>.
- [Sc20] Schaper, N.: „Entwicklung und Validierung eines Modells zur E-Lehrkompetenz.“ *MedienPädagogik: Zeitschrift für Theorie und Praxis Der Medienbildung* 37: 313-342, 2020.
- [We22] Wernecke, A.: „E-Learning: Das Schwein in Tiermedizin und Landwirtschaft – Patient und Schnittstelle zwischen tierärztlicher Praxis und landwirtschaftlichem Betrieb.“ PhD diss., LMU, 2022.
- [Yi21] Yin H., Cao Y., Marelli B., Zeng X., Mason A. J., Cao C.: Soil Sensors and Plant Wearables for Smart and Precision Agriculture. *Advanced Materials*, 33, 20; 2021; <https://doi.org/10.1002/adma.202007764>.

### **3 Summarizing discussion and further perspective**

Plant phenotyping, performed by satellites, drones, agricultural vehicles, robots or ground-based platforms, is a rapidly expanding research area forming a pivotal discipline on the journey towards the digitalization of the agricultural and horticultural sector (Araus et al. 2018, Roitsch et al. 2019). However, phenotyping of plant in vitro cultures by sensor systems is an underrepresented research area with limited reports so far (Table 1). Thus, the potential of automated in situ phenotyping of plant in vitro cultures has also not been explored in depth. The special conditions of plant in vitro cultures such as imaging through closed vessels, specular lighting of vessel material, limited space in multi-layer shelf systems and formation of water condensation place highly challenging demands on phenotyping in this field.

#### **3.1 Automated phenotyping in commercial micropropagation**

Considering these specific claims, we have developed a novel robotic system named “Phenomenon” enabling multi-sensory in situ phenotyping of plant in vitro cultures (Bethge et al. 2023a). This study covered the design and realization of the mechatronic backbone, the implementation of the software architecture of the robotic system, as well as the creation of automated data acquisition and appropriate data processing pipelines. Furthermore, besides the description of the optical properties of culture containers, the feasibility of four tested low-cost sensors was investigated and unprecedented sensor data from plant in vitro cultures were demonstrated. Here, we could show the potential of the RGB-imaging in monitoring the horizontal growth of explants inside the culture vessels as well as the application of a laser-distance sensor for tracking the vertical growth and the feasibility of a micro spectrometer for sensing of chlorophyll fluorescence.

Most of the approaches to phenotyping in vitro cultures reported to date (Table 1) have either the disadvantage that the "plant-to-sensor" approach limits the automation and monitoring of cultivation (Smith et al. 1989, Aynalem et al. 2006, Ibaraki & Gupta 2011, Mansouri et al. 2016, Gupta & Karmakar 2017, Mestre et al. 2017, Faragó et al. 2018), or they focused heavily on *Arabidopsis* research (Dhondt et al. 2014, Barbez et al. 2017, Lube et al. 2022). The roboticized Phenomenon phenotyping system has several unique features such as the multi-sensory approach, direct in situ monitoring during cultivation, and by proposing an approach for solving the challenging imaging situation such as water condensation, specular lighting and mirroring that set it apart from other previously reported systems.

We claimed that the system is of practical use for plant research and can be scaled up **from a prototype to a commercial micropropagation system**. However, the multi-sensory approach of the "Phenomenon" system limited the throughput by an increased size of the detection module and by the temporal and spatial boundaries, when acquiring multi-dimensional data of each object/explant. Furthermore, the system dimensions limited the operation area to 0.6 m<sup>2</sup>. This reduced the system throughput to 6 data acquisition cycles of 10 culture vessels per day in a multiple sensor mode or up to 50 culture vessels per day in a single sensor mode. However, an application of phenotyping prototype system to commercial micropropagation with thousands to millions of vessels in a culture room would drastically change the system demands. Therefore, it is expected that a **higher throughput**, lower spatial and temporal resolution would be required in the application to commercial micropropagation. In addition, a complete transition to full automation of already partially automated data processing and analysis pipelines would be required.

Three out of four tested sensors (RGB camera, laser-distance sensor in shifted scan mode and micro spectrometer) produced novel, and reasonable data of plant in vitro cultures. Whereas with RGB and depth data, important information about plant growth only emerges through their spatiality, chlorophyll fluorescence monitoring with a micro spectrometer holds the potential of high throughput and high information density through a single point measurement that allows quantification of biomass-correlated signals (review in Tremblay et al. 2012, Malapascua et al. 2014) and simultaneous information of physiological stress status (Lichtenthaler et al. 1998, Buschmann et al. 2000). Here, most studies to date focused either on ex vitro plants (Lichtenthaler et al. 1998, Buschmann et al. 2000) or microalgal cell culture (Malapascua et al. 2014), hardly comparable in size, habitus and physiological state to explants in micropropagation. Even if not investigated in detail, we expect that a spectral differentiating chlorophyll fluorescence sensor such as “Closed FluorCam” (Photon System Instruments) would meet the demands of high throughput, fast and accurate measurements of biomass and stress indicators. Thus, the “Phenomenon” system is suggested to be used for further investigations in this topic, in particular if biomass quantification by ground fluorescence signals or stress estimation by fluorescence signature can be generalized across multiple influencing factors such as plant species, chlorophyll content and explant height. Here, the multi-sensory approach of the system could use to fuse the data of projected plant area by RGB data, explant height by depth data and fluorescence intensity to robustly predict biomass in regression models. Furthermore, there is significant scientific interest in investigating the relationship between alterations in spectral fluorescence signatures and various stressors (Lichtenthaler & Rinderle 1988) encountered in vitro such as imbalanced culture media or emergence of disorders.

In addition, a scaled-up version of the “Phenomenon” system that achieves high-throughput data collection could be perfectly used for media optimization of new in vitro cultures, which are still time-consuming and a major problem in plant tissue culture of some species (Nezami-Alanagh et al. 2019, Hameg et al. 2020, Hesami & Jones 2020). Hesami & Jones (2020) pointed out that Hildebrandt (1946) needed 108 different media, while Murashige and Skoog (1962) required five years and 81 different combinations to determine the balance between macro- and micronutrients in the basal medium. Here, projected plant area determined by RGB imaging and average canopy height determined by depth imaging could be used as decision variables that are maximized or minimized by an optimization algorithm to predict the optimal media composition in terms of nutrients and plant growth regulators.

Furthermore, an industrial application would require implementation in a partially existing **database with a unique identifier** for each culture vessel — realized, for example by automated capturing of barcodes, radio-frequency identification chips (RFID chips) or quick response codes (QR code), in agreement with the study of Seelye et al. (2014) for in vitro propagation of kiwifruit. Such a traceability would enable not only analyses of the tracing of historical performance of plant in vitro cultures, but also a deeper investigation on multi-factorial phenomena such as the habitation (Gaspar et al. 2000) by ensuring sufficient phenotyped replicates across multiple subcultures.

We were able to demonstrate the successful monitoring of three plant species and different stages of plant in vitro culture using the presented method. However, the substitution of the vessel lid by PVC foil and the use of a bottom-cooling system affected the growth speed of in vitro cultures by differences in gaseous exchange capacity and water vapor permeability. In parts, this can be solved if the Haze index is considered in addition

to high transparency and good gaseous exchange capacity when **developing “easy to phenotype” culture vessels in future**. This creates new requirements for culture vessel design, such as complete clarity and high transmittance in different spectral regions and in addition to those already existing such as sterility, high gaseous exchange capacity, chemical inertness, high stability for transport and others (Chen 2003, Prakash et al. 2004). The condensation of water results from the thermal load of the cultivation surface due to the high heat dissipation of the subjacent tubular fluorescent lamps in the multi-layer shelf systems. Water condensation was avoided in our study by a bottom

cooling system — the use of which was often hindered in commercial production by a reduction of the growth rate of some plant species— might be solved differently in future. On the one hand the whole horticultural sector undergoes a transition from traditional light sources to semi-conductor-based light emitting diodes (LEDs), and these are expected to become mainstream in commercial micropropagation as well, thus producing less absolute heat emission due to higher efficiency. On the other hand, with an expected increase in energy costs which are mainly caused by the necessity to cool the culture rooms, the use of bottom cooling systems could gain new relevance, since excess heat can be removed through the water pipes directly from culture rooms to a heat exchanger.

### **3.2 Automated detection of morpho-physiological disorders in plant in vitro culture**

Plant in vitro culture covers a highly diverse field of research that involves the cultivation of different plant species as cell, protoplast, tissue, or organ cultures using various pathways to generate plants undergoing different stages of development (Birnbaum & Alvarado 2008, George et al. 2008). Considering the major challenges such as shoot-tip necrosis, shoot fasciation, habituation, recalcitrance, albinism, somaclonal



variation ([Hazarika & Bora 2008](#), [Ruffoni & Savona 2013](#), [Nezami-Alanagh et al. 2019](#), [Abdalla et al. 2022](#)) in plant in vitro culture in terms of plant physiological and genetic variation, we focused on physiological and morphological malformations in the context of hyperhydricity in our study ([Bethge et al. 2023b](#)). We were able to report new optical characteristics, as such reduced reflectance and major absorption peaks of hyperhydricity at 980 nm, 1150 nm, 1200 nm, 1400 nm, 1450 nm, 1520 nm, 1780 nm and higher reflection at 1930 nm of this disorder, provided new insights into the relationship between two established methods for assessing HH, and proposed the first two competing approaches for automatic detection of HH using ML techniques.

While our findings allowed us to temporally narrow down the physiological events within the etiology of HH, it is important to note that our **understanding of the underlying mechanism of hyperhydricity** remains limited. Other aspects have been highlighted in our study, but in particular the reason for the flooding of the apoplast is not yet clear. Similar to many other studies ([Pasqualetto et al. 1988](#), [Franck et al. 1998](#), [van den Dries et al. 2013](#)), our investigation primarily focused on comparing non-induced explants cultivated under conditions of low water availability such as high plant agar concentrations with HH-induced explants cultivated in media offering high water availability (low plant agar concentration, gelrite, liquid culture). However, it is worth noting that the same non-induced explants can also develop HH if exposed to conditions of high water availability. Thus, the non-induced explants are already in an “inherent physiological state of latent HH”, in other words, the limited water availability prevents the expression of the disorder. In order to identify the fundamental cause, it is suggested to compare the metabolomic profiles (focused on osmotically active molecules within in apoplastic liquid) of multiple non-susceptible to serious affected genotypes of one or more plant species in their non-induced state to visualize differences in their physiological

preconfiguration responsible for their susceptibility. This hypothesis is currently being investigated in parts as a side project examining the osmotic potential of susceptible and non-susceptible *Malus* genotypes in their latent HH state ([Mohammadi Nakhjiri et al., unpublished](#)).

We have presented, to the best of our knowledge, the first application of an automated live-monitoring and detection based on **computer vision combined with a convolutional neural network** in plant in vitro culture. Certainly, there is room for debate regarding the optimal integration of imaging in the production process of commercial micropropagation, whether live-monitoring during cultivation or single endpoint imaging on a sterile conveyer belt system after a subculture is most efficient and provides sufficient information for the task at hand. As in other areas of horticulture, a tremendous potential of machine learning-assisted monitoring in plant in vitro culture for various tasks such as determination of optimal transfer time point, media composition, environmental conditions for cultivation, early detection of contaminations and endophytes, nutrient deficiencies, malformations, estimation of multiplication rate, and estimation of transferable explants. Although we have obtained promising results in detecting hyperhydricity, more complex tasks such as multi-classification approaches (for example, specific and early identification between different malformations or culture problems) would demand larger data sets covering different domains, imaging systems, plant species, media and genotypes. To accelerate technological and scientific progress, image datasets should be used in the future to create a common online dataset for benchmarking the different tasks.

Therefore, special attention was paid to the publication of new scientific findings under the aspects of unrestricted **accessibility**, **sustainable research** and easy **applicability** in accordance to FAIR (findability, accessibility, interoperability,

reusability) guiding principles (Wilkinson et al. 2016). Image data sets as well as written software modules were uploaded to free online repositories (Bethge 2022, Bethge 2023) that allow a permanent usage and all scientific articles were published under open access. In addition, the weights of trained object detection models are now available online (Bethge 2023) and can be reused both, as initial weights to improve similar approaches in this area and to directly test model performance as an out-of-the-box feature of the hosting platform.

### **3.3 Application of automated phenotyping robot in higher education**

At latest since the outbreak of the COVID19 pandemic at the beginning of this thesis project, universities worldwide have been forced to critically evaluate their digital learning offerings due to the partial or even complete elimination of face-to-face teaching. In order to compensate for the lack of practical courses and application relevance in digital teaching formats, we have created a concept for knowledge transfer for digital competencies in agriculture as a web-based e-learning system "SensX" (Bethge et al. 2023c). This approach focused primarily on integrating research data sets, small home experimental sensor kits sent to students, and Phenomenon system sensor data (Bethge et al. 2023a) into the digital teaching sessions of "SensX". The evaluation of this concept was carried out by an empirical survey, qualitative interviews of the participating students by an external institution and the evaluation by the lecturers over a two-years period with agricultural and biotechnology students at the University of Applied Sciences Osnabrück. This study revealed an overall high level of acceptance and advocacy of the concept by students with 70% good to very good rating. The student survey showed that as complexity of the learning task increased, students experienced excessive demands and rated the respective session lower. It therefore remains a fine line to provide student with ambitious hands-on exercises that promote a profound and

enduring understanding of digital skills, while avoiding overwhelming them with excessive demands.

While we could demonstrate a successful proof-of-concept of the applied teaching concept, in future we are aiming for the implementation of **a remote phenotyping workshop** on “How to remotely program and control a phenotyping robot” in the SensX e-learning platform. For this purpose, we draw on the software properties of Phenomenon phenotyping robot ([Bethge et al. 2023a](#)), where a full remote control of all capabilities of the system was realized via programming language-independent online access. The goal is to i) design a biological experiment, ii) collaboratively program the robot with interactive computing notebooks (jupyter notebook or google colab) and iii) design suitable data processing and analysis pipelines with a small group of students over the course of a semester. In that case, the conceptualized experiment would then be physically set up by the instructor. In this way, students will be provided with deep insights into sensing, data processing, statistical analysis of biological experiments, programming, and skills at the intersection of digital technology and plant biology.

### 3.4 Summary with critical review of the state of fulfillment

As a critical review of the thesis objectives (Table 2), the state of fulfillment is evaluated by the author.

**Table 2:** Personal assessment of the thesis objectives

Nr.	Thesis objectives	Personal Assessment <sup>[1]</sup>	Comments	References
I.I	Development of a low-cost phenotyping robot system suitable for live-monitoring of in vitro cultures during cultivation.	++	Creation of Phenomenon phenotyping system. However, a certain degree of invasiveness is present due to use bottom-cooling system, PVC film and TiO <sub>2</sub>	<a href="#">Bethge et al. 2023a</a>
I.II	Scalable for high-throughput use in commercial laboratories.	++	Scalability provided Throughput limited	<a href="#">Bethge et al. 2023a</a>
I.III	Capable of monitoring a wide range of plant species and various developmental phases.	+++	Monitoring of in vitro germination, shoot regeneration and shoot multiplication in <i>Arabidopsis</i> , <i>Nicotiana</i> and <i>Malus</i> demonstrated	<a href="#">Bethge et al. 2023a</a>
I.IV	Low-cost sensor systems should be evaluated in regards to their feasibility and potential to quantify key growth parameters of plant in vitro cultures.	+++	Four tested sensors were investigated and feasibility were described	<a href="#">Bethge et al. 2023a</a>
II.I	Investigation of spectral fingerprints of hyperhydricity and identification of specific absorption features that are sufficient for discrimination by ML techniques.	+++	Studies were conducted and major discriminating spectral features were identified: 980 nm, 1150 nm, 1450 nm, 1520 nm, 1780 nm and 1930 nm	<a href="#">Bethge et al. 2023b</a>
II.II	Morphological characterization of hyperhydric explants in time-series image data by Phenomenon phenotyping system.	+++	Quantitative description and qualitative time-lapse video of HH etiology were created	<a href="#">Bethge et al. 2023b</a>
II.III	Identification optical technologies towards automated detection of hyperhydricity should be identified.	++	A SWIR spectral classifier and an RGB object recognition model were proposed for HH detection. Validation were promising, but also preliminary.	<a href="#">Bethge et al. 2023b</a>
III	Establish an interactive e-learning approach supported by phenotyping data to provide students in higher education with in-depth digital competence in the field of sensor technology in horticulture.	+	Proof-of-concept study was promising, but still needs deeper evaluation, fine-tuning, extension and more participants	<a href="#">Bethge et al. 2023c</a>

<sup>[1]</sup>Note: Personal assement ranged from +++ completely fulfilled to — not fulfilled with 4 levels.

In summary, the majority of the goals were either partially or fully reached. In this thesis a novel phenotyping robot for automated multi-sensory in-situ monitoring of plant in vitro culture was developed. Although a certain degree of invasiveness in cultivation could not be avoided and high throughput has not yet been achieved, a successful imaging of a wide range of plant species and developmental phase has been demonstrated. The platform was used to investigate the growth anomaly hyperhydricity, and optical characteristics of the growth anomaly — mainly located in the SWIR region — were

identified and the first approaches based on a spectral classifier and an object detection on RGB images for an automated detection of hyperhydricity were developed. However, the presented detection approaches require further validation to test their generalizability. A concept to use the phenotyping robot to strengthen digital skills in plant sciences in higher education was established.

## **4 Outlook**

Automated phenotyping of developmental processes of plant in vitro cultures is an intriguing and relatively unexplored research domain. In commercial micropropagation, the application of automated phenotyping could help provide continuous real-time data on growth performance for efficient use of production capacities. In order to advance the field, it is crucial to conduct a comprehensive survey on the necessity and implementation of digital phenotyping in commercial propagation and to promote new research projects in collaboration with micropropagation companies enhancing the throughput of phenotyping systems. By harnessing computer vision, sorting tasks aimed at generating uniform, high-quality plant material can be effectively addressed, thereby enhancing standardization while simultaneously reducing labor costs. Moreover, computer vision coupled with statistical learning methods could facilitate the early detection of contaminations, endophytic outgrowths, malformations, and morphological variations. Estimating multiplication rates, identifying transferable explants, and quantifying biomass are particularly valuable for commercial propagation practices. The integration of a high-throughput plant phenotyping system with advanced statistical learning methods offers the potential to extend the applications of phenomics to plant in vitro culture. This combination enables the comprehensive analysis of complex plant traits, facilitating large-scale genome-wide association studies and uncovering novel insights into the genetic basis of propagation rates and other phenotypic traits.

Further research is needed to address the new requirements of optical monitoring of culture vessels. A temporal and spatial high resolution imaging of the physiological processes of plant in vitro cultures will support plant science in overcoming the challenges of micropropagation and lead to a better understanding of the development of growth anomalies. Here, applied research is needed to evaluate the use of advanced

## *Outlook*

---

imaging technology such as hyperspectral cameras and 3D-sensing sensors in plant in vitro culture. The use of machine learning, advanced imaging technologies and sensor fusion approaches in automated phenotyping can further enhance accuracy in detection and the prediction of complex phenotypic traits.



## 5 References

- Abdalla N, El-Ramady H, Seliem MK, El-Mahrouk ME, Taha N, Bayoumi Y, Shalaby TA, Dobránszki J (2022) An academic and technical overview on plant micropropagation challenges. *Horticulturae*. Jul 25;8(8):677.
- Araus JL, Kefauver SC, Zaman-Allah M, Olsen MS, Cairns JE (2018) Translating high-throughput phenotyping into genetic gain. *Trends in plant science*. May 1;23(5):451-66.
- Aynalem HM, Righetti TL, Reed BM (2006) Non-destructive evaluation of in vitro-stored plants: a comparison of visual and image analysis. *In Vitro Cellular & Developmental Biology-Plant*. Nov;42(6):562-7.
- Barbez F, Kleine-Vehn J and Barbez E (2017) Low-Cost Microprocessor-Controlled Rotating Stage for Medium-Throughput Time-Lapse Plant Phenotyping. *Plant Hormones: Methods and Protocols*, pp.37-45.
- Bedenlier, Svenja, Marion Händel, Rudolf Kammerl, Michaela Gläser-Zikuda, Bärbel Kopp, and Albert Ziegler (2021) "Akademische Mediennutzung Studierender im Corona-Semester 2020: Digitalisierungsschub oder weiter wie bisher?." *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung* 40. 229-252.
- Bethge H (2022) Phenomenon—low-cost and multi-sensor system for automated phenotyping of plant in vitro culture. <https://github.com/halube/Phenomenon>. Accessed 07 Sept 2022.
- Bethge H (2023) HH Detection in vitro image Dataset. <https://universe.roboflow.com/hains/hh-detection-in-vitro/dataset/8>. Accessed 10 February 2023
- Bethge H, Mählmann T, Winkelmann T, Rath T (2023c) Remote plant sensing and phenotyping—an e-learning tool in higher education. 43. GIL-Jahrestagung, Resiliente Agri-Food-Systeme.
- Bethge H, Mohammadi Nakhjiri Z, Rath T, Winkelmann T (2023b) Towards auto-mated detection of hyperhydricity in plant in vitro culture. Manuscript accepted, May 2023 in *Plant Cell, Tissue and Organ Culture (PCTOC)*.
- Bethge H, Winkelmann T, Lüdeke P, Rath T (2023a) Low-cost and automated phenotyping system “Phenomenon” for multi-sensor in situ monitoring in plant in vitro culture. *Plant Methods*. 2023 Dec;19(1):1-25. <https://doi.org/10.1186/s13007-023-01018-w>
- Birnbaum KD, Alvarado AS (2008) Slicing across kingdoms: regeneration in plants and animals. *Cell*. Feb 22;132(4):697-710.
- Breiman L (2001) Random forests. *Machine learning*. Oct;45:5-32.
- Buschmann C, Langsdorf G, Lichtenthaler HK (2000) Imaging of the blue, green, and red fluorescence emission of plants: an overview. *Photosynthetica*. Aug;38(4):483-91.
- Busemeyer L, Mentrup D, Möller K, Wunder E, Alheit K, Hahn V, Maurer HP, Reif JC, Würschum T, Müller J, Rahe F (2013) BreedVision—A multi-sensor platform for non-destructive field-based phenotyping in plant breeding. *Sensors*. Feb 27;13(3):2830-47.
- Cardoso JC, Sheng Gerald LT, Teixeira da Silva JA (2018) Micropropagation in the twenty-first century. *Plant cell culture protocols*. 17-46.
- Casanova E, Moysset L, Trillas MI (2008) Effects of agar concentration and vessel closure on the organogenesis and hyperhydricity of adventitious carnation shoots. *Biologia Plantarum*. Mar;52:1-8.
- Cavaco AM, Utkin AB, Marques da Silva J, Guerra R (2022) Making Sense of Light: The Use of Optical Spectroscopy Techniques in Plant Sciences and Agriculture. *Applied Sciences*. Jan 19;12(3):997.

## References

---

- Chen C (2003) Development of a heat transfer model for plant tissue culture vessels. *Biosystems Engineering*. May 1;85(1):67-77.
- Chen C (2016) Cost analysis of plant micropropagation of *Phalaenopsis*. *Plant Cell, Tissue and Organ Culture (PCTOC)*. 126(1):167-75.
- Chen Y, Jiang C, Hyyppä J, Qiu S, Wang Z, Tian M, Li W, Puttonen E, Zhou H, Feng Z, Bo Y (2018) Feasibility study of ore classification using active hyperspectral LiDAR. *IEEE Geoscience and Remote Sensing Letters*. Jul 31;15(11):1785-9.
- Cortes C, Vapnik V (1995) Support-vector networks. *Machine Learning*. Sep;20:273-97.
- De Klerk GJ, Van Den Dries N, Krens FA (2015) Hyperhydricity: underlying mechanisms. In: VI International Symposium on Production and Establishment of Micropropagated Plants 1155 Apr 19 (pp. 269-276).
- Debergh P, Aitken-Christie J, Cohen D, Grout B, Von Arnold S, Zimmerman R, Ziv M (1992) Reconsideration of the term 'vitrification' as used in micropropagation. *Plant cell, Tissue and organ culture*. Aug;30:135-40.
- Dhondt S, Gonzalez N, Blomme J, De Milde L, Van Daele T, Van Akoleyen D, Storme V, Coppens F, TS Beemster G, Inzé D (2014) High-resolution time-resolved imaging of in vitro *Arabidopsis* rosette growth. *The Plant Journal*. Oct;80(1):172-84.
- Dhondt S, Wuyts N, Inzé D (2013) Cell to whole-plant phenotyping: the best is yet to come. *Trends in plant science*. Aug 1;18(8):428-39.
- Du M, Wang X (2011) Linear discriminant analysis and its application in plant classification. In 2011 Fourth International Conference on Information and Computing Apr 25 (pp. 548-551). IEEE.
- European Commission (2019) A Smart and Sustainable Digital Future for European Agriculture and Rural Areas. <https://digital-strategy.ec.europa.eu/en/news/eu-member-states-join-forces-digitalisation-european-agriculture-and-rural-areas>.
- Faragó D, Sass L, Valkai I, Andrási N, Szabados L (2018) PlantSize offers an affordable, non-destructive method to measure plant size and color in vitro. *Frontiers in Plant Science*. Feb 22;9:219.
- Fisher RA (1936) The use of multiple measurements in taxonomic problems. *Annals of Eugenics*. Sep;7(2):179-88.
- Franck T, Crèvecoeur M, Wuest J, Greppin H, Gaspar T (1998) Cytological comparison of leaves and stems of *Prunus avium* L. shoots cultured on a solid medium with agar or gelrite. *Biotechnic & histochemistry*. 73(1):32-43.
- Gamborg OL (2002) Plant tissue culture. biotechnology. milestones. *In Vitro Cellular & Developmental Biology-Plant*. Mar;38:84-92.
- Gaspar T, Kevers C, Bisbis B, Franck T, Crèvecoeur M, Greppin H, Dommes J (2000) Loss of plant organogenic totipotency in the course of in vitro neoplastic progression. *In Vitro Cellular & Developmental Biology. Plant*. May 1:171-81.
- George EF, Hall MA, De Klerk GJ (2008) Plant propagation by tissue culture. Volume I. The background. *Plant Propagation by Tissue Culture*. Springer, Dordrecht.
- Gribble K (1999) The influence of relative humidity on vitrification, growth and morphology of *Gypsophila paniculata* L. *Plant growth regulation*. Mar;27:181-90.
- Guo Q, Zhu Z. Phenotyping of plants (2006) *Encyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation*. Sep 15:1-5.
- Gupta SD, Karmakar A (2017) Machine vision based evaluation of impact of light emitting diodes (LEDs) on shoot regeneration and the effect of spectral quality on phenolic content and anti-oxidant capacity in *Swertia chirata*. *Journal of Photochemistry and Photobiology B: Biology*. Sep 1;174:162-72.
- Hakala T, Suomalainen J, Kaasalainen S, Chen Y (2012) Full waveform hyperspectral LiDAR for terrestrial laser scanning. *Optics express*. Mar 26;20(7):7119-27.

- Hameg R, Arteta TA, Landin M, Gallego PP, Barreal ME (2020) Modeling and optimizing culture medium mineral composition for in vitro propagation of *Actinidia arguta*. *Frontiers in Plant Science*. Dec 23;11:554905.
- Hazarika BN, Bora A (2008) Hyperhydricity-A bottleneck to micropropagation of plants. In *IV International Symposium on Acclimatization and Establishment of Micropropagated Plants*. 865 Dec 8 (pp. 95-101).
- Hazarika BN, da Silva JT, Talukdar A (2006) Effective acclimatization of in vitro cultured plants: methods, physiology and genetics. *Floriculture, ornamental and plant biotechnology: advances and topical issues*. Dec;2:427-38.
- Hesami M, Jones AM (2020) Application of artificial intelligence models and optimization algorithms in plant cell and tissue culture. *Applied Microbiology and Biotechnology*. Nov;104:9449-85.
- Hildebrandt AC, Riker AJ, Duggar BM (1946) The influence of the composition of the medium on growth in vitro of excised tobacco and sunflower tissue cultures. *American Journal of Botany*. Jul 1:591-7.
- Honda H, Takikawa N, Noguchi H, Hanai T, Kobayashi T (1997) Image analysis associated with a fuzzy neural network and estimation of shoot length of regenerated rice callus. *Journal of Fermentation and Bioengineering*. 84(4):342-7.
- Hummel (2012) *PlantEye*—A novel 3D sensor platform for automated determination of plant growth dynamics
- Hutter I, Schneider C (2019) Commercial micropropagation in Germany. *Journal of Applied Botany and Food Quality*. 92:226-31.
- Ibaraki Y, Gupta SD (2011) Thermal imaging of micropropagated plantlets for evaluation of possible wilting. *Environmental Control in Biology*. 49(3):141-8.
- Kautsky H, Hirsch A (1931) Neue Versuche zur Kohlensäureassimilation. *Naturwissenschaften*. 1931 Nov;19(48):964.
- Kazmi W, Foix S, Alenyà G, Andersen HJ (2014) Indoor and outdoor depth imaging of leaves with time-of-flight and stereo vision sensors: Analysis and comparison. *ISPRS journal of photogrammetry and remote sensing*. Feb 1;88:128-46.
- Kemat N (2020) Improving the quality of tissue-cultured plants by fixing the problems related to an inadequate water balance, hyperhydricity (Doctoral dissertation, Wageningen University and Research).
- Kemat N, Visser RG, Krens FA (2021) Hypolignification: A decisive factor in the development of hyperhydricity. *Plants*. Nov 29;10(12):2625.
- Kozai T (1991) Micropropagation under photoautotrophic conditions. *Micropropagation: technology and application*. 447-69.
- LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature*. May 28;521(7553):436-44.
- Lee LC, Liong CY, Jemain AA (2018) Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. *Analyst*. 143(15):3526-39.
- Li L, Zhang Q, Huang D (2014) A review of imaging techniques for plant phenotyping. *Sensors*. Oct 24;14(11):20078-111.
- Lichtenthaler HK, Rinderle U (1988) The role of chlorophyll fluorescence in the detection of stress conditions in plants. *CRC Critical Reviews in Analytical Chemistry*. Jan 1;19(sup1):S29-85.
- Lichtenthaler HK, Wenzel O, Buschmann C, Gitelson A (1998) Plant stress detection by reflectance and fluorescence. *Annals of the New York Academy of Sciences*. 851:271-85.
- Lin Y (2015) LiDAR: An important tool for next-generation phenotyping technology of high potential for plant phenomics?. *Computers and electronics in Agriculture*. Nov 1;119:61-73.

## References

---

- Lube V, Noyan MA, Przybysz A, Salama K, Blilou I (2022) MultipleXLab: A high-throughput portable live-imaging root phenotyping platform using deep learning and computer vision. *Plant Methods*. Mar 27;18(1):38.
- Mahendra, Prasad VS, Gupta SD (2004) Trichromatic sorting of in vitro regenerated plants of gladiolus using adaptive resonance theory. *Current Science*. 10:348-53.
- Malapascua JR, Jerez CG, Sergejevová M, Figueroa FL, Masojídek J (2014) Photosynthesis monitoring to optimize growth of microalgal mass cultures: application of chlorophyll fluorescence techniques. *Aquatic biology*. Nov 20;22:123-40.
- Mansouri A, Fadavi A, Mortazavian SM (2016) An artificial intelligence approach for modeling volume and fresh weight of callus—A case study of cumin (*Cuminum cyminum* L.). *Journal of Theoretical Biology*. May 21;397:199-205.
- McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*. Dec;5:115-33.
- Mestre D, Fonseca JM, Mora A (2017) Monitoring of in-vitro plant cultures using digital image processing and random forests. 8th International Conference on Pattern Recognition Systems.
- Murashige T, Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiologia plantarum*. 15(3):473-97.
- Narvaez FY, Reina G, Torres-Torriti M, Kantor G, Cheein FA (2017) A survey of ranging and imaging techniques for precision agriculture phenotyping. *IEEE/ASME Transactions on Mechatronics*. Oct 9;22(6):2428-39.
- Nezami-Alanagh E, Garoosi GA, Landín M, Gallego PP (2019) Computer-based tools provide new insight into the key factors that cause physiological disorders of pistachio rootstocks cultured in vitro. *Scientific Reports*. Jul 5;9(1):9740.
- Paques M, Boxus P, Dulos M (1985) " Vitrification": an induceable and reversible phenomenon. In: *Symposium on In Vitro Problems Related to Mass Propagation of Horticultural Plants* 212. pp 253-258.
- Pasqualetto PL, Zimmerman RH, Fordham I (1988) The influence of cation and gelling agent concentrations on vitrification of apple cultivars in vitro. *Plant Cell, Tissue and Organ Culture*. 14(1):31-40.
- Paul PK, Sinha RR, Baby P, Shivraj KS, Aremu B, Mewada S (2020) Agricultural Informatics as a Branch of Study in Information Sciences and Technology Domain-A Proposal towards Digital Agriculture. *International Journal of World Policy and Development Studies*.;6(6):56-65.
- Pearson K. LIII (1901) On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*. Nov 1;2(11):559-72.
- Pérez-Bueno ML, Pineda M, Cabeza FM, Barón M (2016) Multicolor fluorescence imaging as a candidate for disease detection in plant phenotyping. *Frontiers in Plant Science*. Dec 2;7:1790.
- Podwyszyńska M, Orlikowska T, Trojak-Goluch A, Wojtania A (2022) Application and Improvement of In Vitro Culture Systems for Commercial Production of Ornamental, Fruit, and Industrial Plants in Poland. *Acta Societatis Botanicorum Poloniae*. Jan 1;91.
- Prakash S, Hoque MI, Brinks T (2004) Culture media and containers. Low cost options for tissue culture technology in developing countries, *FAO/IAEA Division of Nuclear Techniques in Food and Agriculture*, Vienna. Feb 1:29-40.
- Prasad VS, Gupta SD (2008) Applications and potentials of artificial neural networks in plant tissue culture. *Plant tissue culture engineering*. pp. 47-67.
- Ringnér M (2008) What is principal component analysis?. *Nature Biotechnology*. Mar;26(3):303-4.

## References

---

- Roitsch T, Cabrera-Bosquet L, Fournier A, Ghamkhar K, Jiménez-Berni J, Pinto F, Ober ES (2019) New sensors and data-driven approaches—A path to next generation phenomics. *Plant Science*. May 1;282:2-10.
- Rojas-Martínez L, Visser RG, De Klerk GJ (2010) The hyperhydricity syndrome: waterlogging of plant tissues as a major cause. *Propag Ornament Plants*. Dec 1;10(4):169-75.
- Ruffoni B, Savona M (2013) Physiological and biochemical analysis of growth abnormalities associated with plant tissue culture. *Horticulture, Environment, and Biotechnology*. Jun;54(3):191-205.
- Schulz H, Baranska M (2007) Identification and quantification of valuable plant substances by IR and Raman spectroscopy. *Vibrational Spectroscopy*. Jan 16;43(1):13-25.
- Seelye JF, Corpe S, Debenham MC (2014) A management system for tracking high health in vitro kiwifruit germplasm. In XXIX International Horticultural Congress on Horticulture: Sustaining Lives, Livelihoods and Landscapes (IHC2014): 1113 Aug 17 (pp. 113-118).
- Smith MA, Spomer LA (1995) Vessels, gels, liquid media, and support systems. *Automation and environmental control in plant tissue culture*. 371-404.
- Smith MA, Spomer LA, Meyer MJ, McClelland MT (1989) Non-invasive image analysis evaluation of growth during plant micropropagation. *Plant Cell, Tissue and Organ Culture*. Nov;19:91-102.
- Tremblay N, Wang Z, Cerovic ZG (2012) Sensing crop nitrogen status with fluorescence indicators. A review. *Agronomy for sustainable development*. Apr;32:451-64.
- van den Dries N, Gianni S, Czerednik A, Krens FA, de Klerk GJ (2013) Flooding of the apoplast is a key factor in the development of hyperhydricity. *Journal of experimental botany*. Nov 1;64(16):5221-30.
- Vanderschaeghe AM, Debergh PC (1987) Technical aspects of the control of the relative humidity in tissue culture containers. *Mededelingen van de Faculteit landbouwwetenschappen. Rijksuniversiteit Gent*. 52(4a):1429-37.
- Walter A, Liebisch F, Hund A (2015) Plant phenotyping: from bean weighing to image analysis. *Plant methods*. Dec;11(1):1-1.
- Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Botten JW, da Silva Santos LB, Bourne PE, Bouwman J (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*. Mar 15;3(1):1-9.
- Winkelmann T, Geier T, Preil W (2006) Commercial in vitro plant production in Germany in 1985–2004. *Plant cell, tissue and organ culture*. Sep;86(3):319-27.
- Wold S, Ruhe A, Wold H, Dunn, Iii WJ (1984) The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM Journal on Scientific and Statistical Computing*. Sep;5(3):735-43.
- Zhang C, Timmis R, Hu WS (1999) A neural network based pattern recognition system for somatic embryos of Douglas fir. *Plant cell, tissue and organ culture*. 56:25-35.
- Zomer RJ, Trabucco A, Ustin SL (2009) Building spectral libraries for wetlands land cover classification and hyperspectral remote sensing. *Journal of environmental management*. 90(7):2170-7.

## 6 Curriculum vitae

### **Hans Bethge**

06.10.1989

Bredenbecker Str. 6

30449 Hannover

---

#### **Education**

---

- |             |  |
|-------------|--|
| 2019 — 2023 | <b>Ph.D. student</b><br>Institute of Horticultural Production Systems.<br>Department Woody Plant and Propagation Physiology.<br>Cooperative: University of Applied Sciences Osnabrück,<br>Leibniz University Hannover. <ul style="list-style-type: none"><li>• „Towards automated phenotyping in plant tissue culture“</li></ul> |
| 2016 — 2018 | <b>Master studies Horticultural Science M.Sc.</b><br>Institute of Horticultural Production Systems.<br>Department Woody Plant and Propagation Physiology.<br>Leibniz University Hannover. <ul style="list-style-type: none"><li>• „Development of a LED-Research-module for plant in vitro culture“</li></ul>                    |
| 2012 — 2016 | <b>Bachelor studies Horticultural Science B.Sc.</b><br>Wissenschaftszentrum Weihenstephan (WZW).<br>Technische Universität München (TUM). <ul style="list-style-type: none"><li>• „Quality optimization of sports turf through different light quality“</li></ul>  |
| 2010 — 2012 | <b>Bachelor studies Brewing and beverage technology</b><br>Wissenschaftszentrum Weihenstephan (WZW).<br>Technische Universität München (TUM).  |
| 2009 — 2010 | <b>Civilian service</b><br>Haus Tobias, Integrativer Kindergarten Freiburg.  |
| 2009        | <b>University Entrance Degree (Abitur)</b><br>Bertha-von-Suttner-Schule<br>Agrarwissenschaftliches Gymnasium Ettlingen.  |

---

#### **Work experience**

---

- |      |  |
|------|--|
| 2023 | <b>Scientific project assistant</b><br>Leibniz University Hannover. Department Phytophotonic.<br>MWK-funded project: “HyperKorr” — Korrelative optische Erfassung pflanzlicher Krankheitssymptome zur Steigerung der Vorhersagegüte hyperspektraler Messsysteme in der Landwirtschaft. |
|------|--|

- 2019-2022      **Scientific project assistant**  
University of Applied Sciences Osnabrück  
BMEL-funded project: “Agro-Nordwest” — Experimentation Field for the Digital Transformation in Crop Farming.
- 2019            **Scientific assistant for writing of a research proposal**  
University of Applied Sciences Osnabrück.  
Research proposal: “AgriGain” — Agricultural Gateway for Artificial Intelligence”.
- 2019            **Scientific assistant**  
University of Applied Sciences Osnabrück  
Proof study: “UV imaging for plant phenotyping”  
Laboratory for Micro- and Optoelectronics.

---

### **Scholarships**

---

- 2022-2023      **Completion of the doctoral thesis**  
University of Applied Sciences Osnabrück.

---

### **Awards & Research activity**

---

- 2022            **Best Poster Award** — ISHS Young Minds Award 2022  
“Phenotyping robot for plant in vitro culture”.  
International horticulture congress. IHC, Anger 2022.
- 2020-2022      **Scientific review activity**  
ORCID: 0000-0002-3487-9725.  
Plant Cell, Tissue and Organ Culture (PCTOC) Journal  
Springer Nature.
- 2019            **TASPO-Award** — Junge Wissenschaft 2019  
„Monitoring von Nährstoffmangelerscheinungen,  
Wachstumsanomalien und Kontaminationen  
ausgewählter pflanzlicher In-vitro-Kulturen“.  
Haymarket Media GmbH & Co. KG, Berlin 2019.
- 2018            **Best Poster Award** — ICULTA 2018  
“Design and fabrication of a microcontroller based wireless LED-research module for application in in vitro culture labs”. International Conference on UV LEDs Technology and Applications. ICULTA, Berlin 2018.

---

### **Selected certified workshops**

---

- 2022            **Winter School** — Image Analysis for Plant Phenotyping.  
5-days online workshop. Wageningen University & Research. International Plant Phenotyping Network (IPPN).
- 2021            **Hackaton** — Plant phenotyping with minicomputers and low-cost cameras. 3-days online workshop. Anger, France. International Plant Phenotyping Network (IPPN).

## **7 List of publications**

### **7.1 Peer reviewed publications**

Steger S, Steinbacher F, Bethge H, Dreier C (2016) Quality optimization of sport lawns by LED lighting technology. *European Journal of Turfgrass Science*, 47(4), 76-80.

Assou J, Bethge H, Wamhoff D, Winkelmann T (2022) Effect of cytokinins and light quality on adventitious shoot regeneration from leaflet explants of peanut (*Arachis hypogaea*). *The Journal of Horticultural Science and Biotechnology*, 1-18. <https://doi.org/10.1080/14620316.2022.2160382>

Bethge H, Winkelmann T, Lüdeke P, Rath T (2023) Low-cost and automated phenotyping system “Phenomenon” for multi-sensor in situ monitoring in plant in vitro culture. *Plant Methods* 19, 42. <https://doi.org/10.1186/s13007-023-01018-w>

Bethge H, Mohammadi Nakhjiri Z, Rath T, Winkelmann T (2023) Towards automated detection of hyperhydricity in plant in vitro culture. Manuscript accepted/In press, May 2023 in *Plant Cell, Tissue and Organ Culture* (PCTOC). <https://10.1007/s11240-023-02528-0>

### **7.2 Non reviewed publications**

Bethge H, Rath T, Winkelmann T (2018) Entwicklung eines Mikrocontroller-basierten LED-Belichtungsmoduls für die pflanzliche In-vitro-Kultur. *Versuche im deutschen Gartenbau- Zierpflanzen*. Hortigate.

### **7.3 Conference contributions**

#### ***Conference manuscripts***

Bethge H, Mählmann T, Winkelmann T, Rath T (2023) Remote plant sensing and phenotyping — an e-learning tool in higher education. *Lecture Notes in informatics (LNI)*, 43. GIL-Jahrestagung, Gesellschaft für Informatik eV (GIL), 2023

#### ***Oral presentations***

Gesellschaft für Pflanzenzüchtung e.V. GPZ-AG 18 (Zierpflanzen)  
September, 2021, Hann. Münden (Germany)

Bethge H, Rath T, Winkelmann T: “Chancen und Herausforderungen bei der Entwicklung eines In-vitro-Phänotypisierungssystems“

7th International Plant Phenotyping Symposium (IPPS 2022)  
September, 2022, Wageningen (Netherlands)

Bethge H, Winkelmann T, Lüdeke P, Rath T: “In situ monitoring and phenotyping of plant in vitro cultures”



***Poster presentations***

International horticulture congress (IHC 2022).

August, 2022, Angers (France)

Bethge H, Winkelmann T, Lüdeke P, Rath T: "Phenotyping robot for plant in vitro culture".

7th International Plant Phenotyping Symposium (IPPS 2022)

September, 2022, Wageningen (Netherlands)

Bethge H, Winkelmann T, Lüdeke P, Rath T: "Phenotyping robot for everyone".

## **8 Acknowledgement**

First of all, I would like to thank Prof. Dr. Traud Winkelmann, who managed to get me fascinated to plant in vitro culture and guided me from study to doctorate. Your systematic and meticulous scientific work is remarkable and every student gains self-confidence from your unique human mentoring, despite the emotional valleys that must be passed on the way of the PhD. I am very thankful that you offered me your high quality supervision, even though I was a horticulturist with somehow exotic research interest.

I would like to express my special appreciation and thanks to Prof. Dr. Thomas Rath, who strengthened my interest in photonics and managed to find the financial backbone of my doctoral studies. I always could identify myself with your unconventional point of view with simultaneous huge interest in science. I am very grateful for your support and tough discussions about technical and methodical details.

My special thanks go to Prof. Dr. Dag Heinemann, who has granted me the urgently needed time to finish my final manuscript and that he has agreed to chair the examination committee for the disputation. I look forward to working with you in the new research group in the future.

I would also like to mention Prof. Dr. Arno Ruckelshausen, who gave me the chance to gain insights into sensor technology and robotics in the agricultural sector. I always enjoyed being a temporary member of your working group.

Furthermore, I especially thank Patrick Lüdeke without whom I wouldn't be able to solve all the software issues within the project. Thank you for being a good friend and for taking your time.

Many thanks to David Wamhoff in whom I have found an allied friend, supporting me with lots of jokes in the daily struggle of the PhD.

I would also like to thank Zahra Mohammadi Nakhjiri for being an excellent master student, for proof reading and the support in writing the manuscript.

Finally, I like to express my special thanks to the whole research group of the department of Woody Plant and Propagation Physiology providing a wonderful working environment. In particular, coffee breaks with Bärbel Ernst and Ewa Schneider has

## *Acknowledgement*

---

proven to be essential for me. Thank you for your open ears to subject and non-subject related questions, for your daily support in the lab and the time with you.