

This is an electronic reprint of the original article.

This reprint *may differ* from the original in pagination and typographic detail.

Author(s): Mirko Pavicic, Kirk Overmyer, Attiq ur Rehman, Piet Jones, Daniel Jacobson and Kristiina Himanen

Title: Image-Based Methods to Score Fungal Pathogen Symptom Progression and Severity in Excised Arabidopsis Leaves

Year: 2021

Version: Published version

Copyright: The Author(s) 2021

Rights: CC BY 4.0

Rights url: <http://creativecommons.org/licenses/by/4.0/>

Please cite the original version:

Pavicic, M., Overmyer, K., ur Rehman, A., Jones, P., Jacobson, D., Himanen, K. (2021). Image-Based Methods to Score Fungal Pathogen Symptom Progression and Severity in Excised Arabidopsis Leaves}. Plants 10(1), 158. <https://doi.org/10.3390/plants10010158>.

All material supplied via *Jukuri* is protected by copyright and other intellectual property rights. Duplication or sale, in electronic or print form, of any part of the repository collections is prohibited. Making electronic or print copies of the material is permitted only for your own personal use or for educational purposes. For other purposes, this article may be used in accordance with the publisher's terms. There may be differences between this version and the publisher's version. You are advised to cite the publisher's version.

Article

Image-Based Methods to Score Fungal Pathogen Symptom Progression and Severity in Excised *Arabidopsis* Leaves

Mirko Pavicic ^{1,2,3} , Kirk Overmyer ⁴ , Attiq ur Rehman ^{2,3,5} , Piet Jones ^{1,6} , Daniel Jacobson ^{1,6}  and Kristiina Himanen ^{2,3,4,*} 

¹ Oak Ridge National Laboratory, 1 Bethel Valley Rd, Oak Ridge, TN 37830, USA; pavicicenvmv@ornl.gov (M.P.); jonespc@ornl.gov (P.J.); jacobsonda@ornl.gov (D.J.)

² Department of Agricultural Sciences, Viikki Plant Science Centre, Faculty of Agriculture and Forestry, University of Helsinki, Latokartanonkaari 7, 00790 Helsinki, Finland; attiq.rehman@luke.fi

³ National Plant Phenotyping Infrastructure, HiLIFE, University of Helsinki, Latokartanonkaari 7, 00790 Helsinki, Finland

⁴ Organismal and Evolutionary Biology Research Program, Viikki Plant Science Centre, Faculty of Biological and Environmental Sciences, Viikinkaari 1, University of Helsinki, 00790 Helsinki, Finland; kirk.overmyer@helsinki.fi

⁵ Production Systems Unit, Horticulture Technologies, Natural Resources Institute (Luke), Toivonlinnanti 518, 21500 Puumala, Finland

⁶ Bredesen Center for Interdisciplinary Research and Graduate Education, University of Tennessee Knoxville, Knoxville, TN 37996, USA

* Correspondence: kristiina.himanen@helsinki.fi

Abstract: Image-based symptom scoring of plant diseases is a powerful tool for associating disease resistance with plant genotypes. Advancements in technology have enabled new imaging and image processing strategies for statistical analysis of time-course experiments. There are several tools available for analyzing symptoms on leaves and fruits of crop plants, but only a few are available for the model plant *Arabidopsis thaliana* (*Arabidopsis*). *Arabidopsis* and the model fungus *Botrytis cinerea* (*Botrytis*) comprise a potent model pathosystem for the identification of signaling pathways conferring immunity against this broad host-range necrotrophic fungus. Here, we present two strategies to assess severity and symptom progression of *Botrytis* infection over time in *Arabidopsis* leaves. Thus, a pixel classification strategy using color hue values from red-green-blue (RGB) images and a random forest algorithm was used to establish necrotic, chlorotic, and healthy leaf areas. Secondly, using chlorophyll fluorescence (ChlFl) imaging, the maximum quantum yield of photosystem II (F_v/F_m) was determined to define diseased areas and their proportion per total leaf area. Both RGB and ChlFl imaging strategies were employed to track disease progression over time. This has provided a robust and sensitive method for detecting sensitive or resistant genetic backgrounds. A full methodological workflow, from plant culture to data analysis, is described.

Keywords: *Arabidopsis*; high-throughput; plant phenotyping; imaging sensors; *Botrytis*; disease symptom; chlorophyll fluorescence



Citation: Pavicic, M.; Overmyer, K.; Rehman, A.U.; Jones, P.; Jacobson, D.; Himanen, K. Image-Based Methods to Score Fungal Pathogen Symptom Progression and Severity in Excised *Arabidopsis* Leaves. *Plants* **2021**, *10*, 158. <https://doi.org/10.3390/plants10010158>

Received: 15 December 2020

Accepted: 13 January 2021

Published: 15 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A visual assessment of plant disease symptoms is an age-old practice used for the identification of phytopathogen resistant and susceptible crops. Visual assessment is based on observations of changes in color and morphology of plants upon disease progression, rendering the method prone to variations and human error. Disease phenotyping achieved by destructive end-point assays is based on observations of pathogen colonization and quantification under microscopy, which make the analysis of disease progression cumbersome and does not allow for the following of a time series of the same event. To overcome these challenges, efforts have been made towards developing high-throughput non-invasive solutions for plant analysis and disease scoring [1,2]. These solutions are based on digitized

image processing in order to provide unbiased symptom scoring. The utilization of images to evaluate diseases of plants has been practiced for over three decades [3]. Most of the available solutions use red-green-blue (RGB) imagery and chlorophyll fluorescence (ChlFl) images to quantify symptomatic areas in the whole plant or its parts [3]. As the cost of imaging technologies decreases and computing power increases, the use of both ChlFl and RGB imaging for symptom scoring has become even more common [4–8]. The majority of RGB and ChlFl approaches have been implemented for plants with greater agricultural importance and comparatively large organ sizes, such as tobacco, wheat, bean, and soybean [4–8]. The model plant *Arabidopsis thaliana* (*Arabidopsis*) has been used to study the identification of disease resistance genes, however, there are only a few examples found, in the literature, that use *Arabidopsis* with high-throughput (HTP) potential [9–12]. The grey mold, i.e., *Botrytis cinerea* (*Botrytis*) is the second most important necrotrophic fungus impacting agricultural production worldwide [13]. It has a host range of more than 200 crops, being especially harmful to cucumber, tomato, strawberry, raspberry, grapes, ornamental flowers, and even pine and spruce seedling production [13–15]. *Botrytis* initially tends to decrease chlorophyll content causing chlorotic lesions by yellowing/chlorosis of leaf and, being a necrotrophic fungus, it continuously produces toxic compounds that eventually cause cell death (necrosis), then, the fungus feeds on the dead tissue, hence, resulting in visible necrotic lesions [15]. *Botrytis* can also effectively infect *Arabidopsis*, rendering these two model organisms an optimal system for studying plant pathogen interactions [13,15,16]. Previously, Berger et al. [11] studied the interaction between *Pseudomonas syringae* and *Arabidopsis* by combining different irradiance levels and a non-biased analysis and found that using the conventional F_v/F_m parameter provided a robust physiological interpretation of the plant–microbe interactions. Practically, F_m is the maximum fluorescence of a dark-adapted leaf, while F_v is the difference between F_m and the minimum fluorescence (F_o) from a dark-adapted leaf [17,18]. Thus, the ratio F_v/F_m provides an estimation of the maximum quantum efficiency of photosystem II, which usually tends to decrease during the onset of a pathogenic infection [6,10,17–20]. The benefits of chlorophyll fluorescence imaging include detecting disease in terms of F_v/F_m and its progression over time, and it is also able to offer physiological information regarding overall plant health [21]. Therefore, by utilizing image-based sensors at the National Plant Phenotyping Infrastructure (NaPPI) located at the University of Helsinki, Finland, we developed a high-throughput plant disease scoring method with automated image analysis, focusing on a *Botrytis*-*Arabidopsis* model. The method was based on digital imaging which allowed simultaneous screening of several hundreds of plant organs over a time series for the analysis of disease progression. RGB and ChlFl imaging were used to develop this method, which could simultaneously screen color and physiological changes over time. Here, we provide a spatial and temporal analysis of fungal infection symptom severity during disease onset and progression.

2. Results

2.1. *Botrytis* Infection Symptom Screening Assay Development

To establish a sensitive disease scoring method, we used well-characterized *Arabidopsis* lines reported as susceptible and resistant to *Botrytis* [22,23]. As a susceptible control, we deployed the double knockout line *cyp79 b2/b3* (*cytochrome p450, family 79, subfamily b polypeptide 2 and 3*). This line is unable to produce both the indole alkaloid phytoalexin camalexin and indole glucosinolates, which are the most important antimicrobial compounds in *Arabidopsis* [22]. The mutant line *lacs2-3* (*long-chain acyl-coa synthase 2*) possesses a permeable leaf cuticle layer and is associated with a strong immunity against *Botrytis*, thus, it was selected as a resistant control [23]. In this way, *cyp79 b2/b3*, *lacs2-3*, and the wild type Columbia-0 (Col-0) accession represented *Botrytis* response controls for assessing *Botrytis*-induced symptom progression.

Using these three lines, an infection assay protocol was established that lasted for four weeks in total. In the first step, the seeds were sown at high density in soil, stratified in cold, and left to grow for one week. At the beginning of the second week, five seedlings

per genotype were transferred to individual pots (Figure 1A). At this same time, potato carrot tomato agar (PCTA) plates were inoculated with Botrytis conidia and left to grow in darkness at room temperature (~25 °C) for two weeks in order to produce inoculum. At the beginning of the fourth week, leaves five, six, and seven from four plants per genotype (twelve in total) were excised at the base of the petiole and placed in six-well plates containing agar (Figure 1B). The leaf petioles were embedded in the agar to prevent leaf desiccation. Leaves were subsequently inoculated with a conidia suspension and the first image data were collected. The disease development was followed up to 96 h post infection utilizing both RGB and ChlFl imaging in order to record and track the appearance of symptoms (Figure 1C). Images were taken from trays containing eight six-well plates with a total of 48 leaves per tray. The workflow is summarized in Figure 1A–C.

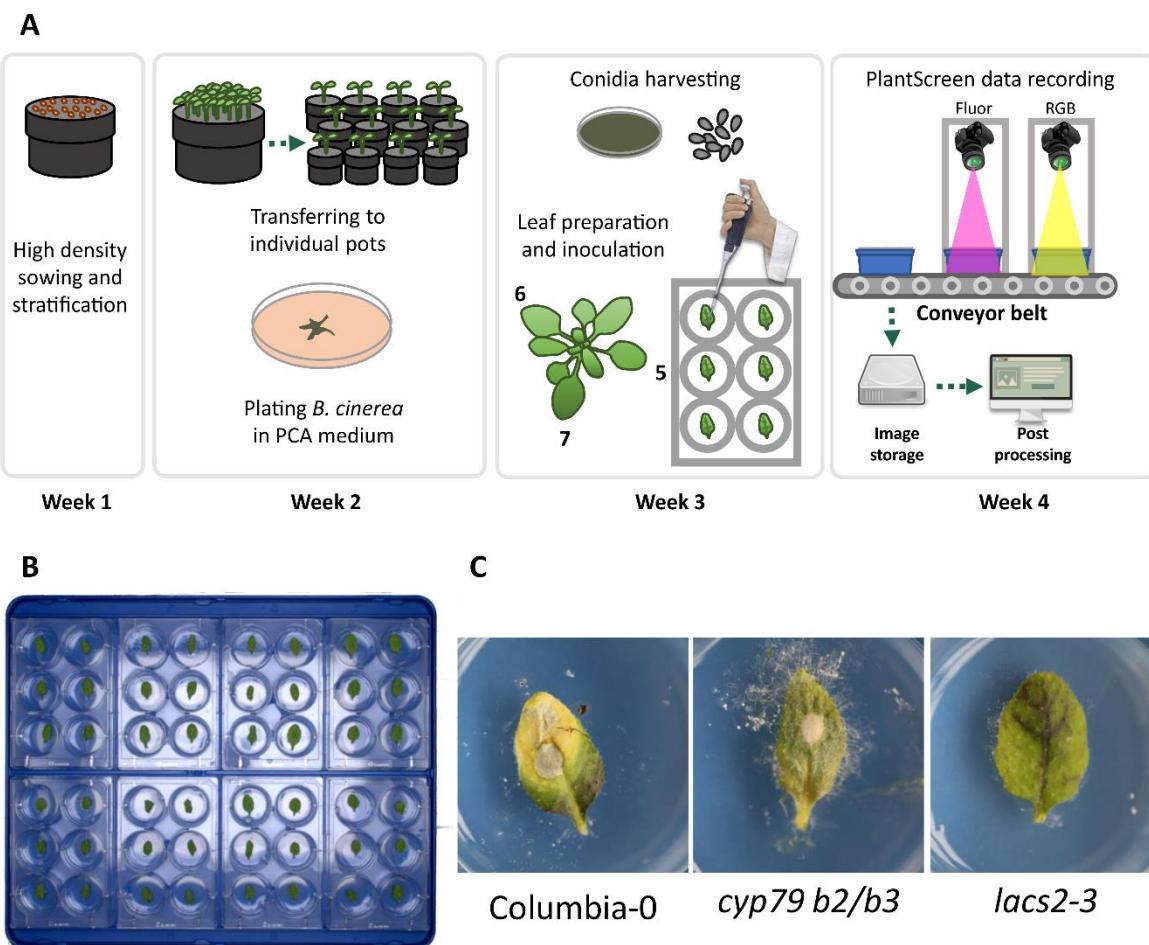


Figure 1. Arabidopsis leaf infection workflow. (A) The workflow of the infection assay used in this study starting from high density sowing of Arabidopsis seeds, potting of seedlings into individual pots, Botrytis conidia culture, inoculation of excised Arabidopsis leaves, imaging by red-green-blue (RGB) and chlorophyll fluorescence (ChlFl) sensors, online image storage and post processing, and data analysis; (B) Eight six-well plates arranged on the imaging tray; (C) Control lines Col-0, *cyp79 b2/b3*, and *lacs2-3* used in this study, after 96 h post inoculation. PCTA medium, potato carrot tomato agar medium; ChlFl, chlorophyll fluorescence imaging; RGB imaging, red-green-blue imaging; Col-0, wild type Columbia-0 accession; *cyp79 b2/b3*, the Botrytis-susceptible cytochrome p450 79-b2 and -b3 (*cyp79 b2/b3*) double mutant; *lacs2-3*, the Botrytis-resistant long-chain acyl-coa synthase2 (*lacs2-3*) mutant.

2.2. RGB Image Processing Strategies

Two methods were used to measure the diseased areas in inoculated Arabidopsis leaves, namely pixel classification for RGB color hues and pixel thresholding for ChlFl imaging derived from F_v/F_m values. The most obvious visible symptoms caused by

Botrytis infection in *Arabidopsis* leaves were the tissue color changes from healthy (green) to chlorosis (yellow) and necrosis (brown). These pixel regions were quantified from the RGB images to assess the disease severity, and then followed throughout the time series in order to record the disease progression in the excised leaves (Figures 1C and 2). For this purpose, an ImageJ (FIJI) plugin called “Trainable Weka Segmentation” was used [24]. This plugin allowed sampling images for different plant features to entrain a classifier to decide which category each leaf pixel belonged to (Figure 2). In this case, four categories were created, i.e., background, healthy, chlorotic, and necrotic. The categories are depicted in Figure 2 with false colors as follows: yellow (background), red (healthy), green (chlorotic), and purple (necrotic). The plugin used the pixels to train a random forest algorithm and create a classifier file that could be applied to batch images without human input. Then, this classifier was used to analyze the images of wild type, *cyp79 b2/b3*, and *lacs2-3* leaves infected with Botrytis (Figure 3A and Supplementary Figure S1). This step created images with pixel counts for Categories 0, 1, 2, and 3 for the background, healthy, chlorotic, and necrotic, respectively. A grid was drawn for each leaf position in the image and the pixel values for each were stored on a spreadsheet. These results were processed in R studio where the symptomatic area was calculated by counting pixels of each category per leaf per day.

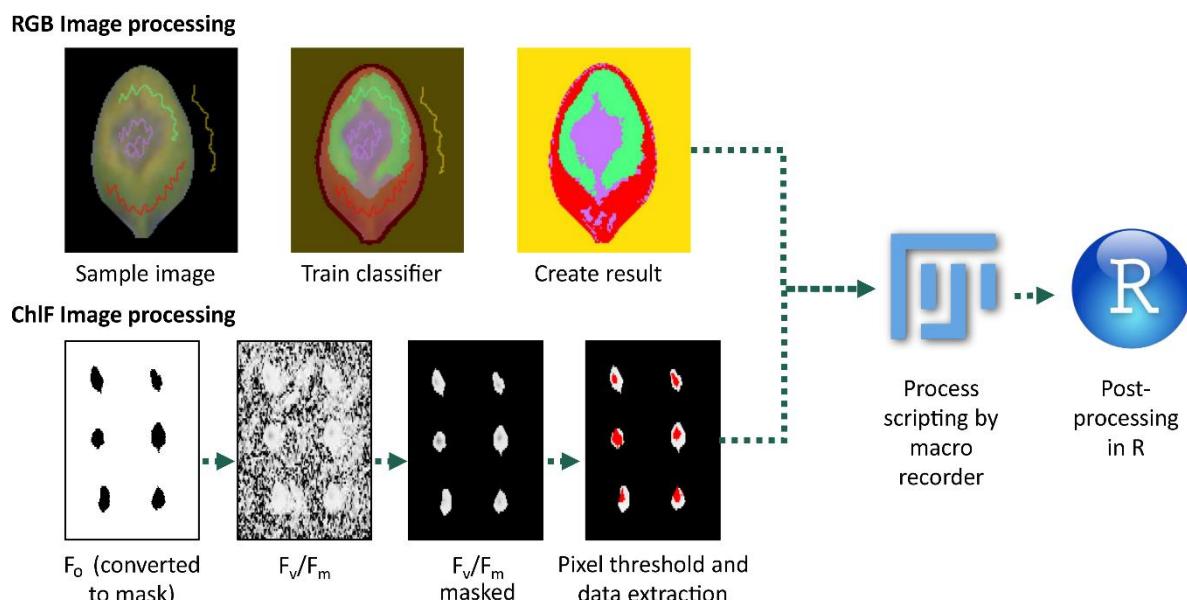


Figure 2. Image processing workflow for RGB (red-green-blue) and ChlFl (chlorophyll fluorescence) images. For RGB images, four categories were depicted, i.e., background, healthy, chlorotic, and necrotic, and are indicated by yellow, red, green, and purple colors, respectively. For ChlFl images, an F_o masked image was used to detect the leaf F_v/F_m to record the photosynthetic capacity on the leaf and the two together to assign pixel thresholds for disease. Both image derived datasets were processed in FIJI and analyzed by R scripts. F_o , minimum fluorescence; F_v , variable fluorescence; F_m , maximum fluorescence; F_v/F_m , maximum quantum yield of the photosystem II.

According to the color hue analysis, the wild type Col-0 plants showed some degree of all symptoms, with chlorotic being the most prevalent in the late stages of infection (Figure 3A). The hyper-susceptible *cyp79 b2/b3* mutant quickly developed necrosis, which spread to cover most of the leaf area. Little chlorosis was observed in *cyp79 b2/b3*, likely due to the rapid expansion of necrosis (Figure 3B,C). In the Botrytis-tolerant *lacs2-3* line, most of the leaf area remained classified as healthy (Figure 3C). Some necrosis and chlorosis were observed in *lacs2-3*, however, these were mostly restricted to the inoculation site and limited as compared with wild type Col-0 leaves.

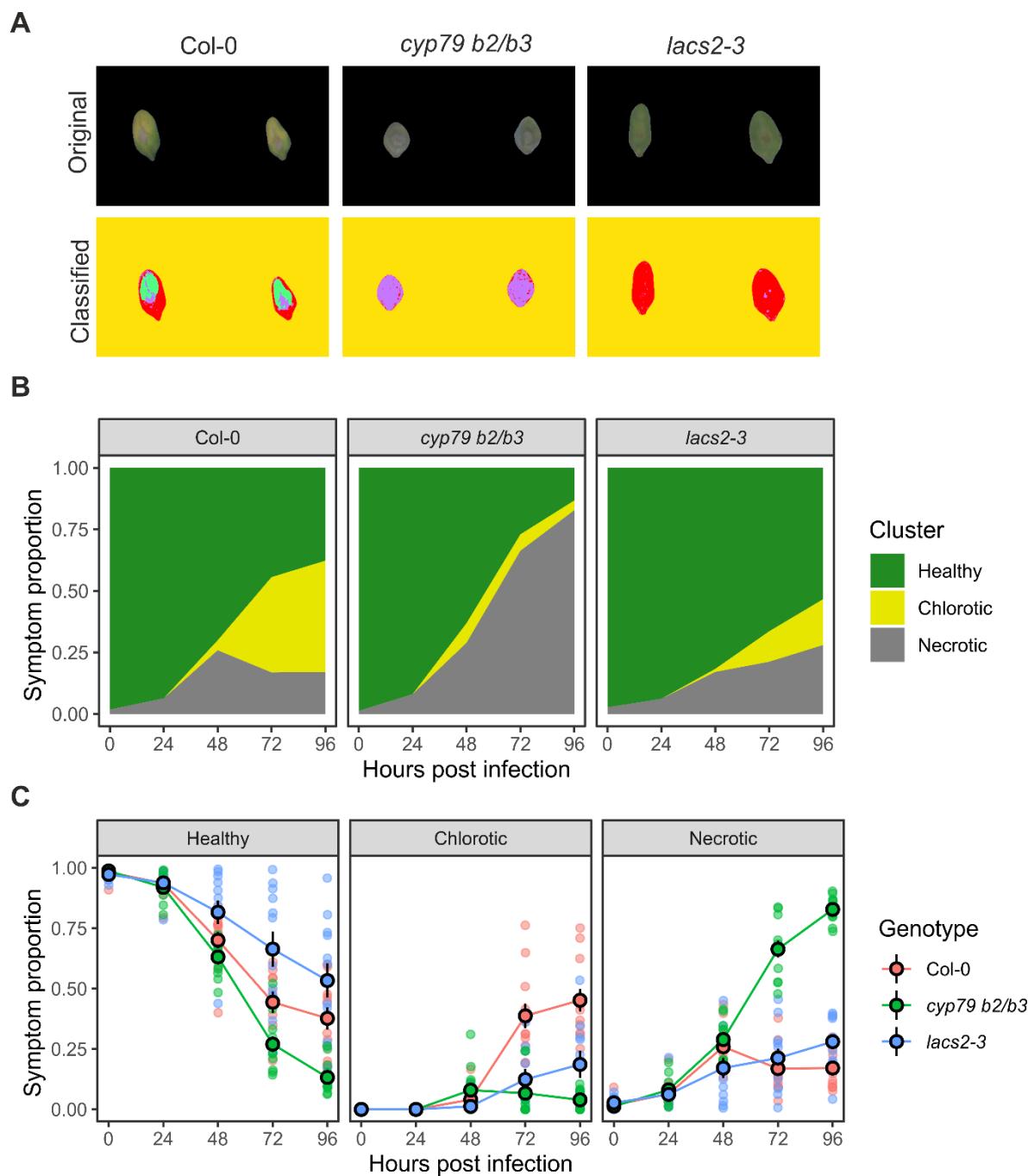


Figure 3. Validation of the RGB pixel classification strategy for Botrytis disease scoring. (A) Images of original RGB images (above) and pixel classification results with color codes (below); (B) Stacked color hue plots of the diseased area progression for the symptom categories healthy, chlorotic, and necrotic; (C) Disease progression area for healthy, chlorotic, and necrotic categories. Circles, mean; error bars, standard error of the mean; Col-0, wild type Columbia-0 accession; *cyp79 b2/b3*, the Botrytis-susceptible cytochrome p450 79-b2 and -b3 double mutant; *lacs2-3*, the Botrytis-resistant long-chain acyl-coa synthase2 mutant.

2.3. ChlFl Image Processing

The ChlFl imaging creates monochromatic images from fluorescence coming from chlorophyll in the plant leaves. In a basic ChlFl imaging protocol, plants are dark adapted so that photosynthesis stops completely [25,26]. Plants are illuminated with red flashes that trigger a minimum fluorescence (F_0) without starting photosynthesis and an image

is acquired. Another image is recorded with a saturating pulse of light that induces a maximum peak of fluorescence (F_m). The F_o pixel values are subtracted from F_m pixel values to create an image termed F_v or variable fluorescence. The F_v pixel values divided by F_m pixel values give F_v/F_m or the maximum quantum yield of the photosystem II if all reaction centers are open. The F_v/F_m in healthy plants has a value of ~0.83 and decays rapidly when plants are stressed [26]. In summary, the image output of a basic ChlFl imaging protocol is F_o , F_m , F_v , and F_v/F_m images, of which the decay of F_v/F_m can be used to record the disease progression.

The calculation of F_v/F_m results in a very noisy image background (Figure 2). To identify only leaf pixels in the F_v/F_m images, a background masking was required. The F_o images, where the leaves were clearly distinguishable from the background, allowed for the creation of the masks (Figure 2).

The F_v/F_m values at 72 h post infection of wild type, *cyp79 b2/b3*, and *lacs2-3* lines, are shown in Figure 4A. Wild type Col-0 plants showed strong F_v/F_m decay at the inoculation site, which expanded to some extent to the neighboring pixels, while a good portion of the leaves remained close to the healthy value of 0.8. In the susceptible *cyp79 b2/b3* mutant, the F_v/F_m values decayed strongly, covering nearly the whole leaf with only a small portion of some leaves remaining healthy. Conversely, the *lacs2-3* F_v/F_m pixel values remained healthy, decaying slightly only at the inoculation point. A density plot of pixel value distribution of F_v/F_m images showed that most of the leaf pixels stayed above 0.75 in the *lacs2-3* leaves (Figure 4B). Using this information, we defined the pixel value threshold of 0.75 to be the cutoff for symptomatic pixels; values below 0.75 were considered to be symptomatic. In this way, two parameters were used to track Botrytis infection in F_v/F_m images, namely the increase in size of the symptomatic area, i.e., a count of pixels with value below 0.75, and the severity of the infection, calculated as total disease pixel area per whole leaf area.

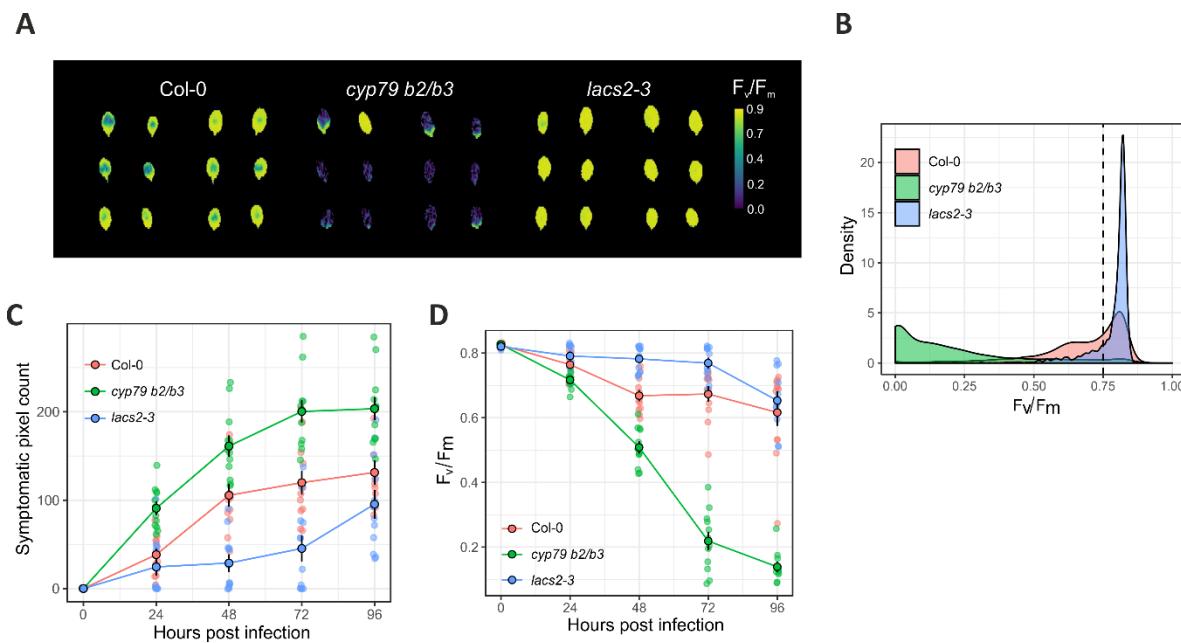


Figure 4. Chlorophyll fluorescence threshold pixel count and mean pixel value strategies for Botrytis disease scoring. (A) False color F_v/F_m image. Yellow pixels represent healthy leaf areas, while green and blue represent symptomatic areas; (B) Density plot for the distributions of pixel intensities for wild type Col-0, *cyp79 b2/b3*, and *lacs2-3*, at 72 h post infection, with an arbitrary pixel threshold (≤ 0.75) to consider a pixel as symptomatic (dashed line); (C) Symptomatic pixel count and (D) disease severity (mean pixel value over the leaf) differences among Col-0, *cyp79 b2/b3*, and *lacs2-3*. Circles represent the mean, and the error bars the standard deviation from the mean. Transparent points in the background represent the actual individual leaves measured with a soft jittering to prevent point overlapping. Col-0, wild type Columbia-0 accession; *cyp79 b2/b3*, the Botrytis-susceptible cytochrome p450 79-b2 and -b3 double mutant; *lacs2-3*, the Botrytis-resistant long-chain acyl-coa synthase2 mutant.

Figure 4C shows the progression of the disease symptomatic pixel count from ChlFl and Figure 4D shows the severity of the infection by F_v/F_m decay in the three control lines. The line *cyp79 b2/b3* showed a larger symptomatic pixel count and stronger infection severity than wild type plants. The opposite effect was observed in *lacs2-3* leaves, thus validating the method.

It should be noted that these datasets presented a nonlinear behavior and an increased dispersion along the time course. The RGB data and symptomatic F_v/F_m area severity are proportional data with values between zero and one, while the symptomatic pixel area is count data. Furthermore, these are not independent measurements since we measured the same individuals several times, therefore, the measurements are not independent from each other and this should be included in the model. To cope with these data features, a generalized additive mixed model (GAMM) using Poisson probability distribution was used for symptomatic pixel counts and a beta probability distribution was used for the other parameters. The differences among mutant lines and wild type were established by differences in the goodness of fit using Akaike information criterion (AIC) of a GAMM with and without the genotype term.

The AIC difference establishes that one model has a better fit than another, where the lower the AIC, the better the fit. For example, the healthy pixel proportion of Col-0 and *lacs2-3* are shown in Figure 5. In this case, a base model was fit without indicating their genotype (Figure 5A). A second model (full model) was fit to the same data, and indicated that there were two genotypes, i.e., Col-0 and *lacs2-3* (Figure 5B). The AIC values of both models were calculated, and their AIC difference was nine units suggesting that the model fit in Figure 5B was the best model, and therefore leaf genotype has a significant impact on healthy pixel proportion decay (Figure 5). For statistical inference, AIC difference ≤ 2 show support for the model without the genotype term, $4 \leq \text{AIC difference} \leq 7$ show considerably less support, and AIC difference > 10 show no support, and therefore the genotype term should be included in the model [27].

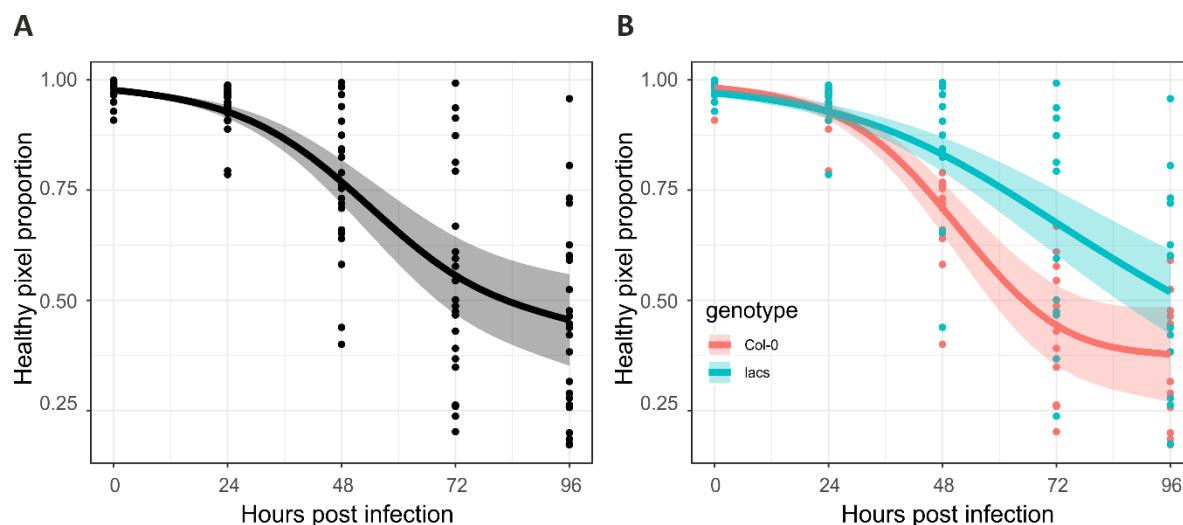


Figure 5. Modeling strategy for inference. (A) Base model without indicating leaves genotype; (B) Full model including leaves genotype term. Col-0, wild type Columbia-0 accession; *lacs2-3*, the Botrytis-resistant long-chain acyl-coa synthase2 (*lacs2-3*) mutant.

This process was repeated for each parameter for both *lacs2-3* and *cyp79 b2/b3*. For most of the parameters analyzed, *lacs2-3* and *cyp79 b2/b3* were significantly different from Col-0 leaves (Table 1). Only *cyp79 b2/b3* healthy pixel proportion and *lacs2-3* necrotic pixel proportion were not significant. Nonetheless, the symptom trend showed that *cyp79 b2/b3* had a lower proportion of healthy pixels and *lacs2-3* had a lower proportion of necrotic pixels than Col-0 (Supplementary Figure S2).

Table 1. Statistical output of the relative likelihood test performed in this study.

Line	Imaging	Parameter	AIC Base Model	AIC Full Model	AIC Diff.
<i>cyp79 b2/b3</i>	RGB	Healthy	−410.2	−412.5	2.2
		Chlorotic	−1302.1	−1325.4	23.3
		Necrotic	−322.8	−364.7	41.9
	ChlFl	Pixel count	1492.0	1305.5	−186.5
		Severity	−341.8	−382.9	41.1
<i>lacs2-3</i>	RGB	Healthy	−373.3	−383.0	9.7
		Chlorotic	−1449.7	−1448.5	1.3
	ChlFl	Necrotic	−334.3	−339.6	5.3
		Pixel count	2174.9	1799.3	−375.6
		Severity	−413.9	−458.5	44.5

Col-0, wild type Columbia-0 accession; *cyp79 b2/b3*, the Botrytis-susceptible *cytochrome p450 79-b2* and *-b3* double mutant; *lacs2-3*, the Botrytis-resistant *long-chain acyl-coa synthase2* mutant; RGB, red-green-blue; ChlFl, chlorophyll fluorescence; AIC, Akaike information criterion; Diff, difference.

3. Discussion

Botrytis, also known as grey mold, is the second most important fungal pathogen in agriculture worldwide with a wide host range of more than 200 crops [13–15]. There are several transcriptome studies available and many genes have been identified to be activated in plants upon contact with Botrytis, but it remains unclear if they are responsible for plant immunity and/or posterior disease development. Utilization of genetically engineered model plants has significantly advanced our understanding of plant pathogenesis and unraveled the role of many genes involved in diseases. In contrast to biotrophic pathogens, which have simple gene-for-gene resistance involving a single *Resistance (R)* gene in the host, resistance against necrotrophs such as Botrytis, is complex, quantitative, and polygenic. For this reason, despite recent advances, our understanding of immune signaling mechanisms against necrotrophs lags behind that of biotrophs [22,28,29]. Thus, there is a need for further studies into the genetics of plant immunity against Botrytis. The ability of Botrytis to infect Arabidopsis offers an excellent infection system to explore gene function, due to the enormous amount of information and genetic tools available for Arabidopsis. Genes identified through this system could be used for breeding disease resistance in crop plants [30]. Pathogen infection phenotyping is, however, currently time-consuming, and subjective or end-point assays make the analysis of disease progression non-viable [3]. To overcome this, many digital image-based tools have been developed with automation potential, providing the opportunity to develop an unbiased and high-throughput analysis.

Pathogens induce many physiological changes in plants that can be detected by image-based analysis [11]. Stressed plants exhibit visible symptoms such as chlorosis and necrosis at the infection sites. An analysis of these symptoms is traditionally based on a visual assessment of the color of the infection site and measurements of the size of the infection spot. Both these measurements are laborious and can be biased by human error. Automated image analysis that is based on pixel classification provides a novel alternative to these manual assays. In addition to visible responses, plants have many invisible changes that occur during infections. These changes can be detected by chlorophyll fluorescence measurements that represent the photosynthetic activity in the plant. Thus, these measurements can detect physiological responses early, for example, as early as 6 h after bacterial inoculation [11]. Furthermore, early, and late responses can be differentiated by specific chlorophyll fluorescence signatures.

In this study, RGB and ChlFl cameras were used to record Botrytis disease progression in detached leaves of Arabidopsis. The sensors are mounted on a conveyor belt system that allows automatic tray delivery and imaging (Figure 1A). The RGB images record the reflected light of red, green, and blue colors that are stored as three channels in the image. Then, this information is used to create a full color image using the different levels of red, green, and blue in the picture. In this way, each pixel in the image stores information for

these three values (red, green, and blue) that act as coordinates for any image software to compose an image on the screen. This information can also be used to classify each pixel into different groups. There are several classification algorithms available such as k-means, k-nearest neighbor, logistic regression, decision trees, random forests, etc. In this study, a random forest algorithm was trained to create a classification model. Random forests are based on collections of decision trees, where many trees vote on which category a pixel should be classified. This has the advantage of correcting the tendency of individual decision trees to overfit the training dataset. Every classification method has some degree of inaccuracy where some pixels are misclassified, especially those values close to the category borders. For instance, in Supplementary Figure S1, some chlorotic pixels were identified by manual segmentation, however, the trained model was unable to recognize those, showing room for further training and improvement of the model. Therefore, it is important to use proper controls and expert evaluation that allows validation of the classifications. Classification models ease the post processing data extraction, since it can be automatically applied to batches of images. The RGB analysis in this study utilizes completely free tools, and therefore can be applied to any picture of infected leaves taken in a systematic manner in any laboratory.

ChlFl imaging is classified as active imaging since it requires its own light source. ChlFl imaging is more sensitive than RGB imaging and can detect symptoms earlier than RGB imaging. Due to the monochromatic nature of ChlFl images, the symptoms are scored by averaging all pixel values in the leaf to estimate the infection severity or by classifying them as symptomatic or not symptomatic using pixel thresholding to estimate the symptomatic area of the leaf. Because ChlFl imaging depends on chlorophyll, it is difficult to detect signals when the infection is very advanced and little chlorophyll remains in the leaf. This event can create an underestimation of the diseased area. Therefore, ChlFl is a sensitive technique that should be used with care with aggressive pathogens.

Data can sometimes be analyzed by linear regression with a Gaussian probability distribution. In order to properly apply this technique, the data must fulfil the following assumptions: (i) measurements should be normally distributed, (ii) each measurement should be independent, and (iii) the variance of the measurements should be homogeneous across treatments (homoscedasticity). In this study, the data violated all these assumptions making Gaussian linear regression inappropriate for these datasets. For instance, symptomatic pixel counts cannot take values below zero and symptomatic pixel proportions and F_v/F_m can only take values between zero and one. To solve this problem, a Poisson probability distribution was used for symptomatic pixel counts and a beta probability distribution was used for proportional data. All parameters analyzed exhibited nonlinear patterns and a tendency for over dispersion along the time course, violating the homoscedasticity assumption of Gaussian models. Thus, GAM models were used to account for nonlinear patterns and overdispersion. Finally, to deal with the dependency created by repeated measurements, each leaf was labeled with an identification (id) that was included as random effect in the GAM.

4. Conclusions

In this study, we have proposed a complete pipeline for the analysis of *Botrytis* symptom progression in *Arabidopsis*. This pipeline is based on RGB and ChlFl images, where the full image and data processing are scripted, opening the possibility to use this method in HTP applications. The RGB image analysis was performed using freeware and could be implemented in any lab with any standard RGB camera. Here, we also show that phenotyping data may violate many modeling assumptions, and therefore it is of utmost importance to use the proper statistical tools to analyze different data types, otherwise, incorrect inferences may be drawn. In addition, many of the methodological steps regarding sample collection and inoculation still requires manual work, in turn affecting the throughput of the study and can serve as a bottleneck for genetic screening of large populations. Nonetheless, this study could serve as a useful tool for screening

Arabidopsis germplasm in various infection assays and could be extrapolated to other plant species with modifications in our open-source analysis pipeline using freely available tools such as FIJI and R.

5. Materials and Methods

5.1. Plant Material and Growth Conditions

To develop a disease scoring method, the cuticle permeable and Botrytis-resistant *long-chain acyl-coa synthase 2* (*lacs2-3*) mutant and camalexin and indole glucosinolate deficient and Botrytis-susceptible *cytochrome p450 79-b2* and *-b3* (*cyp79 b2/b3*) double mutant were used as controls (Bessire et al., 2007 and Buxdorf et al., 2013). The wild type Columbia-0 (Col-0) accession was used as a reference control. Seeds were sown in water saturated vermiculite-peat (Type B2 peat, Kekkilä, Vantaa, Finland, www.kekkila.fi) substrate (1:1 ratio) in 8 × 8 cm pots at high density (~100 seeds per pot). Then, pots were put in mini greenhouses and stratified at 4 °C for 72 h, in darkness, to ensure even germination. Mini greenhouses were placed in a growth room (PhytoScope, PSI, Drasov, Czech Republic) and kept covered to maintain humidity during germination. On day four, the lid was removed, and seedlings were manually thinned using forceps leaving 30 seedlings per pot. Growth conditions in the Arabidopsis growth chamber were 12 h light/12 h darkness and 22 °C and 60% relative air humidity. White LED was used as light source with a photosynthetically active radiation (PAR) of 130 $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$, controlled using an LI-190R Quantum Sensor coupled to an LI-250A light meter (LI-COR, Bad Homburg, Germany, www.licor.com).

5.2. Botrytis Cinerea Culture Conditions

Botrytis strain B05.10 was cultured on potato carrot tomato agar (PCTA) medium containing potato carrot tomato extract (see Supplementary Materials for details), 1% dextrose (Sigma-Aldrich, Munich, Germany, www.sigmaaldrich.com), 0.15% yeast extract (Biokar Diagnostics, Allonne, France), and 0.7% agar (Sigma-Aldrich). A detailed description of the PCTA medium preparation is provided in Appendix A. Botrytis inoculum was cultured for two weeks in PCTA plates in darkness at room temperature (~23 °C). Botrytis conidia were collected using 2/3 strength potato dextrose broth (Sigma-Aldrich) and filtered using miracloth (Merk Millipore, Darmstadt, Germany, www.merckmillipore.com) to remove mycelia. Conidia concentration was adjusted to $1 \times 10^6/\text{mL}$ using a Fuchs-Rosenthal counting chamber (Assistent, Germany, www.hecht-assistent.de).

5.3. Dissected Leaf Infection Assay

For leaf infections, one-week-old Arabidopsis seedlings were transferred to individual pots (6 cm²) and allowed to grow for two weeks more (Figure 1A). Whenever possible, leaves five, six, and seven were harvested from three-week-old plants and placed in six-well plates (Sarstedt, Nümbrecht, Germany, www.sarstedt.com) containing 0.7% agar gel (Duchefa Biochemie, Haarlem, The Netherlands, www.duchefa-biochemie.com) and distributed in trays holding eight plates each (Figure 1B). Twelve leaves per line from four individual plants were inoculated with 10 µL of the conidia suspension. The infection development was recorded daily for a period of 96 h post infection.

5.4. Phenotyping Infrastructure

The disease scoring method was developed utilizing the imaging sensors available at the NaPPI facility at the University of Helsinki (<https://www.helsinki.fi/en/infrastructures/national-plant-phenotyping>). This unit has a controlled environment PhytoScope Walk-In chamber with a PlantScreen™ Compact System and 12 LED-illuminated shelves as an integrated phenotyping platform (Photon Systems Instruments, PSI, Drasov, Czech Republic, www.psi.cz). The PlantScreen™ Compact System can hold up to 18 trays and the growth shelves another 36 trays of 20 plants each, making a potential total of 2592 leaves per experiment. Within this system, trays were transported on conveyor belts

between the light-isolated visible RGB imaging cabinet, chlorophyll fluorescence (ChlFl) imaging cabinet, the weighing and watering station, and the dark/light acclimation chamber with regulated photosynthetically active radiation (PAR) from 0–700 $\mu\text{mol m}^{-2} \text{s}^{-1}$. The ChlFl measurements were acquired using an enhanced version of the FluorCam FC-800MF pulse amplitude modulated (PAM) chlorophyll fluorometer (PSI, Czech Republic). The ChlFl imaging unit has been described by Awlia et al. (2016) and features a 1/2" monochromatic sensor of 720×560 pixels resolution and a lens type Lensagon CY0314. The ChlFl illumination panel (FluorCam SN-FC800-195) has a pulse-modulated short duration red-orange flashes (620 nm), a red-orange actinic light (620 nm) with maximum photosynthetic photon flux density (PPFD) of $300 \mu\text{mol m}^{-2} \text{s}^{-1}$, a cool white actinic light with maximum PPFD of $500 \mu\text{mol m}^{-2} \text{s}^{-1}$, and a saturating light pulse with a maximum PPFD of $3000 \mu\text{mol m}^{-2} \text{s}^{-1}$. The RGB imaging unit is equipped with three RGB cameras (one top and two side views) mounted on robotic arms, each supplemented with an LED-based light source to ensure homogeneous illumination of the imaged object. The RGB images (resolution 2560×1920 pixels) of individual trays were captured using the GigE uEye UI-5580SEC/M 5 Mpx Camera (IDS, Obersulm, Germany) with SV-0814H lens to assess plant growth and morphological traits. Light conditions, plant position, and camera settings were fixed throughout the experiments.

5.5. Imaging Recording and Processing

Following inoculation, chlorophyll fluorescence and RGB imaging was performed daily, for five days. The RGB images were processed using Morpho Analysis software for background removal (version 1.0.7, PSI). The FIJI software was subsequently used with the Weka segmentation plugin to train a random forest algorithm for classification of leaf pixels into the following four classes: healthy, chlorotic, necrotic, and background. Thus, using the GUI, pixels representing each class were sampled manually to train the model. The training features selected to create the model were Gaussian blur, Sobel filter, Hessian, difference of Gaussian, and membrane projections. Once trained, the classifier was applied to all images in batch mode using FIJI scripting tool according to the ImageJ wiki page “Scripting the Trainable Weka Segmentation” (https://imagej.github.io/Scripting_the_Trainable_Weka_Segmentation). Afterwards, a grid was drawn with squares matching each leaf position on the classified images and the pixel information was extracted using the FIJI function “save XY coordinates”. Pixel values were further processed using R studio (www.rstudio.com), as described below.

The ChlFl imaging-based detection of the infection site was done based on reduction of the chlorophyll fluorescence emitted from the infected leaf. Image capture was done using the principal F_v/F_m imaging protocol of a FluorCam system (PSI) that generated parameter images of minimum fluorescence (F_o) and maximum fluorescence (F_m) yields [26,31]. To score the infection severity, a common plant stress indicator, i.e., the quantum yield of photosystem II (F_v/F_m), was utilized [11,32]. The ChlFl raw images of .fimg format were collected for five days and stored in the central database. These images were processed to extract pixel information using the FIJI software [33,34]. Due to the high number of images produced, image processing was scripted, and run in batch mode, in FIJI. The script steps were the following:

- Open F_o images;
- Select background pixels by pixel thresholding;
- Convert to binary images and save them to be used as masks;
- Open masks and select background pixels;
- Open F_v/F_m images;
- Transfer background selection from masks to F_v/F_m images;
- Set F_v/F_m background pixels value to -100 ;
- Draw a rectangle for each plant position;
- Save XY coordinates to export each pixel value within the rectangle as comma separated values (CSV) file.

The pixel threshold was arbitrarily selected for each image in order to select as much leaf area as possible. In advanced stages of the infection, pixel values in the infected area were too low to be selected by normal thresholding and the masks corresponding to these time points were manually modified to obtain a better estimation of the symptomatic area.

The output of this script generated one CSV file per leaf per time point that were imported and further analyzed using R studio software.

5.6. Data Processing in R Studio

Due to fluctuation in the leaf size, the pixel count in each category of RGB images was normalized against the total leaf size and results were displayed as stacked color plots. For ChlFl results, the CSV files were merged into a single file and leaf genotype information was added to this dataset. Background pixels were filtered out by removing all values equal to –100. The infection size was calculated by counting the number of pixels with values below 0.75 and the infection severity was calculated by averaging all leaf pixels. This data processing and analysis was done using the Tidyverse R group packages (www.tidyverse.org/packages). Statistical analysis was done using R package mgcv and generalized additive models with different probability distributions depending on the data type [35]. Then, Akaike information criterion (AIC) was calculated for model inference [27]. All R and ImageJ script generated to process are available at https://github.com/mipavici/MDPI_leaf_infection.

Supplementary Materials: The following are available online at <https://www.mdpi.com/2223-7747/10/1/158/s1>, Figure S1: Comparison between manually and Weka segmented leaves, Figure S2: Model fit for all parameters used in this study.

Author Contributions: Conceptualization, K.H., K.O., and M.P.; methodology, M.P., K.O., A.u.R., and P.J.; software, M.P. and A.u.R.; validation, M.P., P.J., and D.J.; formal analysis, M.P. and P.J.; investigation, M.P.; resources, K.O. and K.H.; data curation, M.P. and A.u.R.; writing—original draft preparation, M.P.; writing—review and editing, M.P., K.O., K.H., A.u.R., P.J., and D.J.; visualization, M.P.; supervision, K.H.; project administration, K.H.; funding acquisition, K.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Academy of Finland (Suomen Akatemia) grant numbers #283138, #256094, and #250972.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available at https://github.com/mipavici/MDPI_leaf_infection.

Acknowledgments: This project was supported by the Academy of Finland (Suomen Akatemia #283138, #256094, and #250972) and Becas Chile from the Chilean National Agency for Research and Development (ANID). This research was also supported by the Plant-Microbe Interfaces Scientific Focus Area in the Genomic Science Program of the Office of Biological and Environmental Research (BER) in the U.S. Department of Energy Office of Science under award number DE-AC05-00OR22725. This manuscript has been co-authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the U.S. Department of Energy (DOE). The U.S. government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for U.S. government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Supplementary Table of PCTA (potato carrot tomato agar) medium

1. Prepare:
 - 300 g potatoes, unpeeled, sliced and chopped fine;
 - 25 g carrots, peeled and chopped fine;
 - (Optional, one tomato, chopped);
 - Add to 800–900 mL water;
 - Bring to a boil;
 - Simmer for 20 min;
 - Filter through miracloth (or equivalent) supported in a sieve into a beaker.
2. Add:
 - 10 g dextrose;
 - 1.5 g yeast extract.
3. Adjust volume to 1 L with water.
4. Add 3.75 g agar to each of 4500 mL bottles.
5. Add 250 mL PCA media to each bottle.
6. Autoclave sterile 121 °C × 15 min.
7. Store at 4 °C.

References

1. Fahlgren, N.; Gehan, M.A.; Baxter, I. Lights, camera, action: High-throughput plant phenotyping is ready for a close-up. *Curr. Opin. Plant Biol.* **2015**, *24*, 93–99. [[CrossRef](#)] [[PubMed](#)]
2. Furbank, R.T.; Tester, M. Phenomics-technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* **2011**, *16*, 635–644. [[CrossRef](#)] [[PubMed](#)]
3. Bock, C.H.; Poole, G.H.; Parker, P.E.; Gottwald, T.R. Plant Disease Severity Estimated Visually, by Digital Photography and Image Analysis, and by Hyperspectral Imaging. *Crit. Rev. Plant Sci.* **2010**, *29*, 59–107. [[CrossRef](#)]
4. Chaerle, L.; Hagenbeek, D.; De Bruyne, E.; Valcke, R.; Van Der Straeten, D. Thermal and Chlorophyll-Fluorescence Imaging Distinguish Plant-Pathogen Interactions at an Early Stage. *Plant Cell Physiol.* **2004**, *45*, 887–896. [[CrossRef](#)]
5. Rolfe, S.A.; Scholes, J.D. Chlorophyll fluorescence imaging of plant-pathogen interactions. *Protoplasma* **2010**, *247*, 163–175. [[CrossRef](#)]
6. Rousseau, C.; Belin, E.; Bove, E.; Rousseau, D.; Fabre, F.; Berruyer, R.; Guillaumès, J.; Manceau, C.; Jacques, M.A.; Boureau, T. High throughput quantitative phenotyping of plant resistance using chlorophyll fluorescence image analysis. *Plant Methods* **2013**, *9*, 17. [[CrossRef](#)]
7. Simko, I.; Jimenez-Berni, J.A.; Sirault, X.R.R. Phenomic Approaches and Tools for Phytopathologists. *Phytopathology* **2017**, *107*, 6–17. [[CrossRef](#)]
8. Singh, V.; Misra, A.K.K. Detection of plant leaf diseases using image segmentation and soft computing techniques. *Inf. Process. Agric.* **2017**, *4*, 41–49. [[CrossRef](#)]
9. Matouš, K.; Benediktyová, Z.; Berger, S.; Roitsch, T.; Nedbal, L. Case study of combinatorial imaging: What protocol and what chlorophyll fluorescence image to use when visualizing infection of *Arabidopsis thaliana* by *Pseudomonas syringae*? *Photosynth. Res.* **2007**, *90*, 243–253. [[CrossRef](#)]
10. Bonfig, K.B.; Schreiber, U.; Gabler, A.; Roitsch, T.; Berger, S. Infection with virulent and avirulent *P. syringae* strains differentially affects photosynthesis and sink metabolism in *Arabidopsis* leaves. *Planta* **2006**, *225*, 1–12. [[CrossRef](#)]
11. Berger, S.; Benediktyová, Z.; Matouš, K.; Bonfig, K.; Mueller, M.J.; Nedbal, L.; Roitsch, T. Visualization of dynamics of plant-pathogen interaction by novel combination of chlorophyll fluorescence imaging and statistical analysis: Differential effects of virulent and avirulent strains of *P. syringae* and of oxylipins on *A. thaliana*. *J. Exp. Bot.* **2007**, *58*, 797–806. [[CrossRef](#)]
12. Fordyce, R.F.; Soltis, N.E.; Caseys, C.; Gwinner, R.; Corwin, J.A.; Atwell, S.; Copeland, D.; Feusier, J.; Subedy, A.; Eshbaugh, R.; et al. Digital Imaging Combined with Genome-Wide Association Mapping Links Loci to Plant-Pathogen Interaction Traits. *Plant Physiol.* **2018**, *178*, 1406–1422. [[CrossRef](#)]
13. Dean, R.; Van Kan, J.L.; Pretorius, Z.A.; Hammond-Kosack, K.E.; Di Pietro, A.; Spanu, P.D.; Rudd, J.J.; Dickman, M.; Kahmann, R.; Ellis, J.; et al. The Top 10 fungal pathogens in molecular plant pathology. *Mol. Plant Pathol.* **2012**, *13*, 414–430. [[CrossRef](#)]
14. Petäistö, R.L. *Botrytis cinerea* and Norway spruce seedlings in cold storage. *Balt. For.* **2006**, *11*, 24–33.
15. Williamson, B.; Tudzynski, B. *Botrytis cinerea*: The cause of grey mould disease. *Mol. Plant* **2007**, *8*, 561–580. [[CrossRef](#)]
16. Luo, H.; Laluk, K.; Lai, Z.; Veronese, P.; Song, F.; Mengiste, T. The *Arabidopsis* *Botrytis Susceptible1* Interactor defines a subclass of RING E3 ligases that regulate pathogen and stress responses. *Plant Physiol.* **2010**, *154*, 1766–1782. [[CrossRef](#)]
17. Kalaji, H.M.; Schansker, G.; Ladle, R.J.; Goltsev, V.; Bosa, K.; Allakhverdiev, S.I.; Breštic, M.; Bussotti, F.; Calatayud, A.; Dąbrowski, P.; et al. Frequently asked questions about *in vivo* chlorophyll fluorescence: Practical issues. *Photosynth. Res.* **2014**, *122*, 121–158. [[CrossRef](#)]
18. Kalaji, H.M.; Schansker, G.; Breštic, M.; Bussotti, F.; Calatayud, A.; Ferroni, L.; Goltsev, V.; Guidi, L.; Jajoo, A.; Li, P.; et al. Frequently asked questions about chlorophyll fluorescence, the sequel. *Photosynth. Res.* **2017**, *132*, 13–66. [[CrossRef](#)]

19. Baker, N.R. Chlorophyll Fluorescence: A Probe of Photosynthesis In Vivo. *Annu. Rev. Plant Biol.* **2008**, *59*, 89–113. [[CrossRef](#)]
20. Bauriegel, E.; Giebel, A.; Geyer, M.; Schmidt, U.; Herppich, W.B. Early detection of Fusarium infection in wheat using hyperspectral imaging. *Comput. Electron. Agric.* **2011**, *75*, 304–312. [[CrossRef](#)]
21. Mutka, A.M.; Bart, R.S. Image-based phenotyping of plant disease symptoms. *Front. Plant Sci.* **2015**, *5*. [[CrossRef](#)]
22. Kliebenstein, D.J.; Rowe, H.C.; Denby, K.J. Secondary metabolites influence *Arabidopsis/Botrytis* interactions: Variation in host production and pathogen sensitivity. *Plant J.* **2005**, *44*, 25–36. [[CrossRef](#)]
23. Bessire, M.; Chassot, C.; Jacquat, A.C.; Humphry, M.; Borel, S.; Petétot, J.M.C.; Métraux, J.P.; Nawrath, C.; Line Chassot, C.; Jacquat, A.C.; et al. A permeable cuticle in *Arabidopsis* leads to a strong resistance to *Botrytis cinerea*. *EMBO J.* **2007**, *26*, 2158–2168. [[CrossRef](#)]
24. Arganda-Carreras, I.; Kaynig, V.; Rueden, C.; Eliceiri, K.W.; Schindelin, J.; Cardona, A.; Sebastian Seung, H. Trainable Weka Segmentation: A machine learning tool for microscopy pixel classification. *Bioinformatics* **2017**, *33*, 2424–2426. [[CrossRef](#)]
25. Maxwell, K.; Johnson, G.N. Chlorophyll fluorescence—A practical guide. *J. Exp. Bot.* **2000**, *51*, 659–668. [[CrossRef](#)]
26. Murchie, E.H.; Lawson, T. Chlorophyll fluorescence analysis: A guide to good practice and understanding some new applications. *J. Exp. Bot.* **2013**, *64*, 3983–3998. [[CrossRef](#)]
27. Burnham, K.P.; Anderson, D.R. Multimodel Inference. *Sociol. Methods Res.* **2004**, *33*, 261–304. [[CrossRef](#)]
28. Laluk, K.; Mengiste, T. Necrotroph Attacks on Plants: Wanton Destruction or Covert Extortion? *Arab. Book* **2010**, *8*, e0136. [[CrossRef](#)]
29. Xu, J.; Meng, J.; Meng, X.; Zhao, Y.; Liu, J.; Sun, T.; Liu, Y.; Wang, Q.; Zhang, S. Pathogen-Responsive MPK3 and MPK6 Reprogram the Biosynthesis of Indole Glucosinolates and Their Derivatives in *Arabidopsis* Immunity. *Plant Cell* **2016**. [[CrossRef](#)]
30. Schwessinger, B.; Bart, R.; Krasileva, K.V.; Coaker, G. Focus issue on plant immunity: From model systems to crop species. *Front. Plant Sci.* **2015**, *6*, 195. [[CrossRef](#)]
31. Tschiersch, H.; Junker, A.; Meyer, R.C.; Altmann, T. Establishment of integrated protocols for automated high throughput kinetic chlorophyll fluorescence analyses. *Plant Methods* **2017**, *13*. [[CrossRef](#)] [[PubMed](#)]
32. Barbagallo, R.P.; Oxborough, K.; Pallett, K.E.; Baker, N.R. Rapid, noninvasive screening for perturbations of metabolism and plant growth using chlorophyll fluorescence imaging. *Plant Physiol.* **2003**, *132*, 485–493. [[CrossRef](#)] [[PubMed](#)]
33. Rueden, C.T.; Schindelin, J.; Hiner, M.C.; DeZonia, B.E.; Walter, A.E.; Arena, E.T.; Eliceiri, K.W. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinform.* **2017**, *18*, 529. [[CrossRef](#)] [[PubMed](#)]
34. Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B.; et al. Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **2012**, *9*, 676–682. [[CrossRef](#)] [[PubMed](#)]
35. Wood, S.N. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc. Ser. B* **2011**, *73*, 3–36. [[CrossRef](#)]