Establishing a biomarker set for the diagnosis of apple replant disease using qPCR

Luisa Baader[1], Susan Schröpfer[1], Stefanie Reim[1], Traud Winkelmann[2], Nils Orth[2], Jiem Krueger[3], Georg Guggenberger[3], Jens Boy[3], Jacqueline Kaldun[4], Eva Lehndorff[4], Nele Meyer[5], Kristin Hauschild[6], Henryk Flachowsky[1]

[1]Julius Kühn Institute (JKI) – Federal Research Centre for Cultivated Plants, Institute for Breeding Research on Fruit Crops, Dresden-Pillnitz, Germany

[2]Leibniz University Hannover, Institute of Plant Genetics, Hannover, Germany

[3]Leibniz University Hannover, Institute of Earth System Sciences, Section Soil Science, Hannover, Germany

[4]University Bayreuth, Bayreuth Centre for Ecology and Environmental Research, Bayreuth, Germany

[5]Goethe University Frankfurt, Institute of Physical Geography, Frankfurt, Germany

[6]Julius Kühn Institute (JKI) – Federal Research Centre for Cultivated Plants, Institute for Epidemiology and Pathogen Diagnostics, Braunschweig, Germany

Author email addresses

Luisa Baader: luisa.baader@julius-kuehn.de
Susan Schröpfer: susan.schroepfer@julius-kuehn.de
Stefanie Reim: stefanie.reim@julius-kuehn.de
Traud Winkelmann: traud.winkelmann@zier.uni-hannover.de
Nils Orth: orth@baum.uni-hannover.de
Jiem Krueger: krueger@ifbk.uni-hannover.de
Georg Guggenberger: guggenberger@ifbk.uni-hannover.de
Jens Boy: boy@ifbk.uni-hannover.de
Jacqueline Kaldun: jacqueline.kaldun@uni-bayreuth.de
Eva Lehndorff: eva.lehndorff@uni-bayreuth.de
Nele Meyer: nele.meyer@em.uni-frankfurt.de
Kristin Hauschild: kristin.hauschild@ufz.de
Henryk Flachowsky: henryk.flachowsky@julius-kuehn.de

Abstract

Apple replant disease (ARD) is a soil-borne disease that arises from replanting apple trees on land previously used for apple cultivation. There is interest in biomarkers that can reliably assess the severity of ARD by quantifying how strongly apple plants react to the disease in soils of different agro-environments. Thus far, transcriptomic studies of ARD-affected plants have examined only a few soils at a time, revealing that expression patterns vary among different soils. Here, we analyzed the expression of 90 candidate genes in the roots of apple plants (rootstock genotype 'M.26') grown in ARD-affected soils from 151 sites across Germany to test whether a consistent pattern of gene expression under ARD-conditions exists. Additionally, the expression of the candidate genes was analyzed in the leaves of apple plants grown in 18 different ARD-affected soils. Most of the genes (72) showed significantly upregulated expression in roots under ARD conditions, while only 11 showed significantly upregulated expression in leaves, suggesting that these genes play a significant role in the ARD reaction in roots but only a limited or no role in leaves. The candidate genes were evaluated for their potential as ARD biomarkers, defined by their consistently increased expression under ARD conditions across different soils and correlation with ARD severity. The accordingly selected ARD biomarker genes in roots include genes involved in phytoalexin biosynthesis, lignin metabolism, ethylene metabolism, cyanogenesis, detoxification, programmed-cell-death, and plant defense. These biomarkers have the potential to assess the severity of ARD and open up new possibilities for disease diagnosis.

Keywords: apple replant disease (ARD), gene expression, high-throughput RT-qPCR, *Malus domestica*, soil borne disease, biotic stress response, biomarkers

Introduction

Apple replant disease (ARD) is a soil-borne disease that arises from replanting apple trees on land previously used for apple cultivation and can persist in the soil for decades [1]. Affected plants typically exhibit stunted shoot growth and root damage as well as reduced fruit yield and quality [2-4]. As a result, ARD poses a significant economic threat to orchards and tree nurseries worldwide, with estimated yield losses ranging from 20 to 50% [2,3,5].

Previous studies have shown that ARD is associated with harmful shifts in the soil biota community composition and it is suspected that root exudates and the decomposition products of dead apple plant material induce these changes in the soil biome [1,6,7]. Many microorganisms have been described as being associated with ARD, including members of *Pythium* [2], *Rhizoctonia* [2], *Fusarium* [8], *Actinomycetes* [9], *Nectriaceae* [4,9,10], and *Streptomyces* [11]. At the same time, a decrease in taxa with presumable beneficial effects was recorded [12,13]. Abiotic factors, such as soil texture, influence ARD severity, which often varies significantly, even within the same field [3,14]. However, the replanting history of an orchard appears to play a bigger role than the soil characteristics [15].

Currently, there are few sustainable and effective countermeasures against ARD and the most common approaches, soil disinfection and crop rotation, are often impractical due to cost or environmental concerns [1]. Planting ARD-tolerant rootstocks appears to be the most reliable long-term method of mitigating the impact of the disease, however the performance of commercially available tolerant rootstocks varies depending on the site [7]. Other promising ARD countermeasures, such as microbial inoculants, intercropping with *Tagetes*, and biofumigation with brassica seed meal, have produced inconsistent results as well [1,3,16]. Developing strategies for managing ARD requires

a deeper understanding of the disease's underlying causes and the plant's response mechanisms [1,17].

Transcriptomic studies of apple plants grown in ARD soil revealed the increased expression of genes associated with the biotic stress response [13,17-22]. Among these genes, those involved in phytoalexin biosynthesis have been found to be especially overexpressed under ARD conditions. Recent studies have shown that biphenyl and dibenzofuran phytoalexins are induced in response to ARD and inhibit pathogenic as well as commensal and beneficial microorganisms [13,15,17,18,20,23-25]. Other genes that have been shown to be differentially expressed under ARD conditions or after infection with ARD-associated pathogens are involved in the metabolism of phytohormones, flavonoids, lignin, and other secondary metabolites, as well as having functions related to signaling, detoxification, and pathogen defense processes [18,19,21,22,26-29]. However, these expression patterns vary between different soils and apple genotypes [15,18,20,30,31].

Fruit growers are interested to know whether and to what extent the soil on their site is affected by ARD, as this information could help them plan new orchards. Biomarkers that can detect and quantify how apple trees react to ARD could help evaluate ARD severity, management options and rootstock tolerance [20]. To reliably assess ARD severity, an evaluation system is needed that can be applied universally to soils from different agro-environments [17]. Thus far, transcriptomic studies of ARD-affected plants have examined only a few soils at a time [30]. To identify robust biomarker genes, experiments involving a large number of different ARD-affected soils are needed.

As a part of the joint BonaRes ORDIAmur project [32], 151 apple-replanted soils from different apple growing regions in Germany were utilized in an ARD biotest experiment as established by Yim *et al.* [12]. For this, 'M.26' apple plants sensitive to ARD were cultivated in untreated, potentially ARD-affected soils as well as in gamma-irradiated control samples of each soil. Experimental data from the various analyses such as soil characteristcs, plant growing parameters, as well as gene expression profiles were collected by the project partners and are available in the BonaRes Repository [33-35].

The aim of this study is to identify universal biomarker genes for assessing the severity of ARD in different soil types. Based on previous studies, 90 candidate genes (CGs) will be selected that were either differentially expressed in apple plants under ARD conditions or shown to be involved in pathways induced by ARD-mediated stress. The expression of these CGs will be analyzed in roots and leaves derived from the large-scale experiment described above. Considering qPCR performance, differential gene expression, and their correlations with plant growth parameters, the large sample size will enable the evaluation of the CGs as ARD biomarkers with high statistical power.

Results

Selection of candidate biomarker genes

A total of 90 CGs were selected as potential biomarkers for ARD (Tab. S1). Eighty six of these genes have previously been found to be differentially expressed in apple plants grown under ARD conditions [13,17-22] or in apple plants infected by pathogens associated with ARD [26-29]. In addition, four putative phytoalexin biosynthesis genes, for which sequence information was provided by a collaboration partner (Liu *et al.,* unpublished), were included in the analysis. The main selection criterion for the CGs was the reported upregulation of their expression under ARD conditions. Additionally, genes showing higher upregulation in the ARD-tolerant genotype 'Mxr.5' compared to the ARD-susceptible genotype 'M.9' were preferably selected, as these genes are presumed to play a role in ARD tolerance [19]. Furthermore, genes with functions related to biotic stress response were preferentially chosen as well.

Based on the literature, UniProt [36] and KEGG [37] the CGs were classified into functional groups (Fig. 1; Tab. S1). Seven of the selected genes play a role in the biosynthesis of phytoalexins, five genes in the metabolism of lignin, and eight genes in the metabolism of other secondary metabolites. These include the synthesis of phenylpropanoids, flavonoids, sesquiterpenoids and triterpenoids, cyanides, and glucosinolates. Twelve CGs have functions related to plant-pathogen interactions, including genes involved in the recognition and response to pathogens. Out of the eleven genes playing a role in plant hormone metabolism and signaling, eight are related to ethylene, two to jasmonate and one to auxin. Twelve CGs have functions related to signaling, and eight each to transport and detoxification. The 13 genes classified as playing a role in other stress responses include genes involved in programmed cell death as well as genes related to biotic and abiotic stress response. The remaining six genes have other functions, such as being involved in the metabolism of lipids or protein processing.

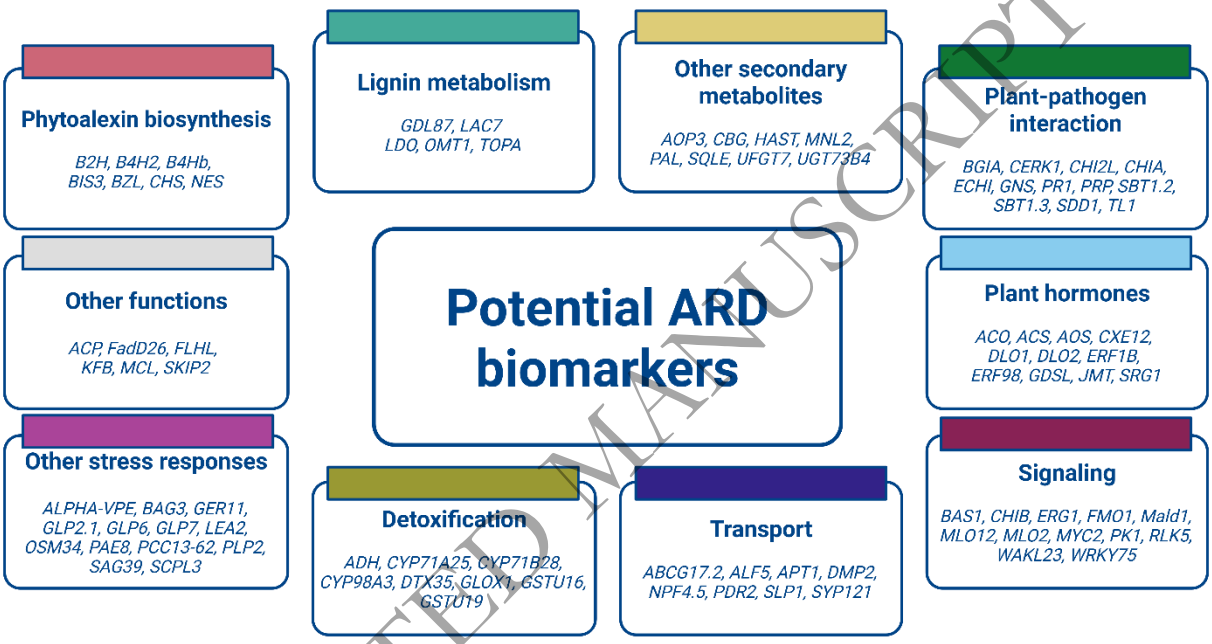

Figure 1: Functional classification of the selected potential ARD biomarkers.
Based on literature, 90 CGs were selected that were shown to be differentially expressed in apple under ARD conditions, in apple plants infected by pathogens associated with ARD, or genes involved in pathways, which were described to be induced by ARD-mediated stress. These genes were classified into functional groups according to literature, UniProt and KEGG and their full names are listed in Table S1. Figure 1 was created with BioRender.com.

Establishing gene specific primers for high-throughput real-time PCR

For 21 of the CGs, qPCR primer sequences were already available, while for the other 69 genes primers had to be established (Tab. S1). A total of 79 primer pairs were designed with NCBI Primerblast and tested for specificity *in silico* by blasting against the *Malus domestica* transcriptome (taxid:3750). In order to rule out qualitative errors in advance, the new primer pairs were tested in an establishing qPCR prior to the high-throughput qPCR. Only the primers that showed the expected amplification product size in the gel electrophoresis, exhibited one specific melt-curve peak and demonstrated sufficient amplification efficiency were selected. Any primer pairs that failed the quality check were discarded and redesigned.

Quality parameters for expression of reference and CGs

To analyze the CGs' expression in response to ARD conditions in root and leaf tissue, plants of the 'M.26' rootstock were grown for six weeks in 151 different ARD-affected soil samples (Fig. 2). In accordance with the biotest established by Yim *et al.* [12], samples of each soil that were disinfected through gamma-irradiation served as ARD-unaffected controls. The origins and characteristics of the soils and the growth difference of the plants grown in ARD and control soil are available in the BonaRes Repository [34,35]. The shoot fresh mass reduction in ARD soils compared to their respective control soil ranged from -74.5% to 77.9% and was significantly reduced in 103 soils. The shoot dry mass reduction (-76.3% to +72.3%) was significant in 100 ARD soils. As an example, Figure S1 illustrates the growth reduction of plants grown in the untreated and disinfected variants of one soil.

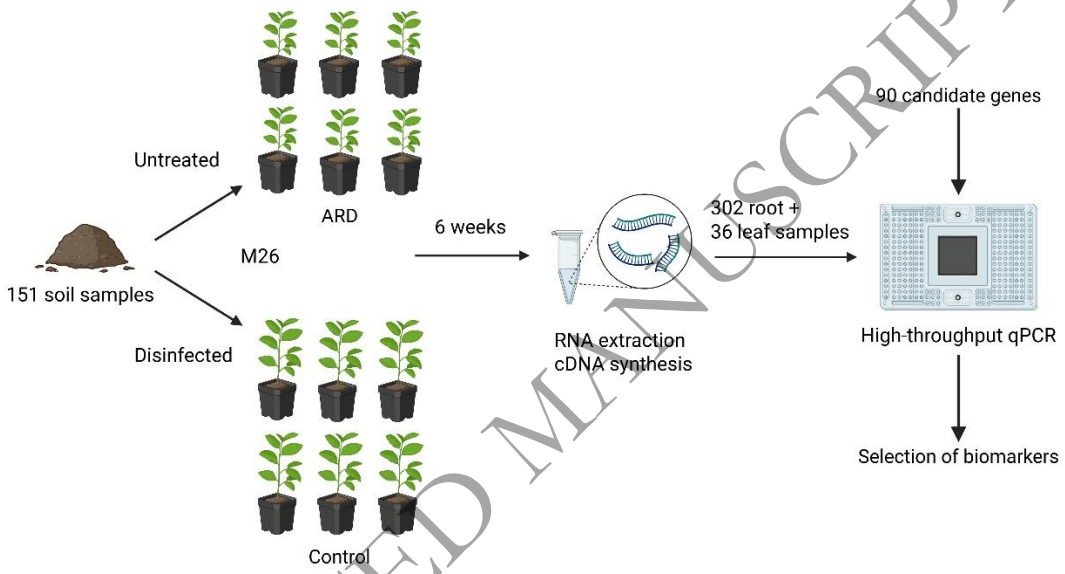

Figure 2: Experimental design.
A total of 151 soil samples were collected from different apple growing regions across Germany that were affected by ARD. 'M.26' apple plants were grown in a greenhouse biotest for six weeks under ARD conditions using an untreated soil sample. A soil sample that was disinfected using gamma-irradiation served as corresponding control. Samples were taken from the roots of six plants from both the ARD and the control variants in each soil type. In addition, leaves from six plants in the ARD and control variants were sampled for 18 soils. Per soil and treatment, the samples were pooled, and total RNA was extracted and reverse-transcribed into cDNA. Gene expression analysis of the 90 CGs was performed by qPCR using the BioMark HD high-throughput system and based on the results, the most promising potential biomarker genes were selected. Figure 2 was created with BioRender.com.

The expression analysis of the 90 CGs (Tab. S1) and five reference genes (Tab. S2) was performed by qPCR using the BioMark HD high-throughput system. The stability of the reference genes was analyzed with RefFinder, a tool for evaluating the stability of reference genes that combines the programs BestKeeper, the comparative ΔCt method, geNorm, and NormFinder [38]. The comprehensive stability ranking was as follows, from most to least stable: *EF1b, ACT7, EF1a, TUBB, UBE210* (Tab. S3). Because *UBE210* consistently ranked last, it was excluded from further analysis and only *ACT7, EF1a, EF1b* and *TUBB* were used as reference genes.

The qPCR showed efficiencies between 98 and 100% for 70 primer pairs and efficiencies above 90% for 15 further primer pairs (Tab. S4). The qPCR with the *SRG1* primers exhibited the lowest efficiency

at 86% and for the primer pairs *FMO1, GER11, GLP7* and *JMT* no amplification efficiencies could be calculated, as the standard samples failed the quality check of the Fluidigm software.

Expression of the CGs in response to ARD in roots

A total of 58 740 qPCR reactions were performed using the high-throughput system. Of these, 5 239 were excluded from analysis due to technical issues. In order to calculate the relative expression (fold change values) of each gene per soil sample, both the ARD sample and the corresponding control sample had to pass the technical quality check. Therefore, the number of samples included in the relative expression analysis differs among the genes (Tab. S4). For the majority of the genes (58), samples from plants grown on at least 140 soils were used in the analysis. The analysis used samples from over 100 soils for 21 of the genes and over 40 soils samples for 10 genes. For the *GLP7* gene the relative expression could be calculated for only three samples.

Nearly all CGs have been documented in the literature as being upregulated under ARD conditions. However, a wide variation in their differential expression was observed in this study (Fig. 3), with the fold change values of the individual samples ranging from 0.0001 to 1 064 865 (log_2 fold-change (log_2FC): −12.8 to 20.0). The gene expression data for each soil sample is available in the BonaRes Repository [33]. Seven genes showed lower expression on average under ARD conditions compared to control conditions (*CYP71A25, ERF1B, ERG1, Mald1, MLO12, SKIP2, UFGT7*), while the expression of 11 genes was not significantly changed on ARD soil compared to the control (Tab. S4). The remaining 72 genes were, on average of all soils, significantly upregulated under ARD conditions ($p < 0.05$). Of these, 57 genes exhibited average fold change values above 4.5 (log_2FC: 2.2), seven genes exhibited average fold changes between 100 and 1 000 (log_2FC: 6.6 to 10.0), and 17 genes exhibited average fold changes above 1 000. Figure S2 shows the relative expression values of three example genes for each soil sample, illustrating how the relative expression of each gene varies between the soils.

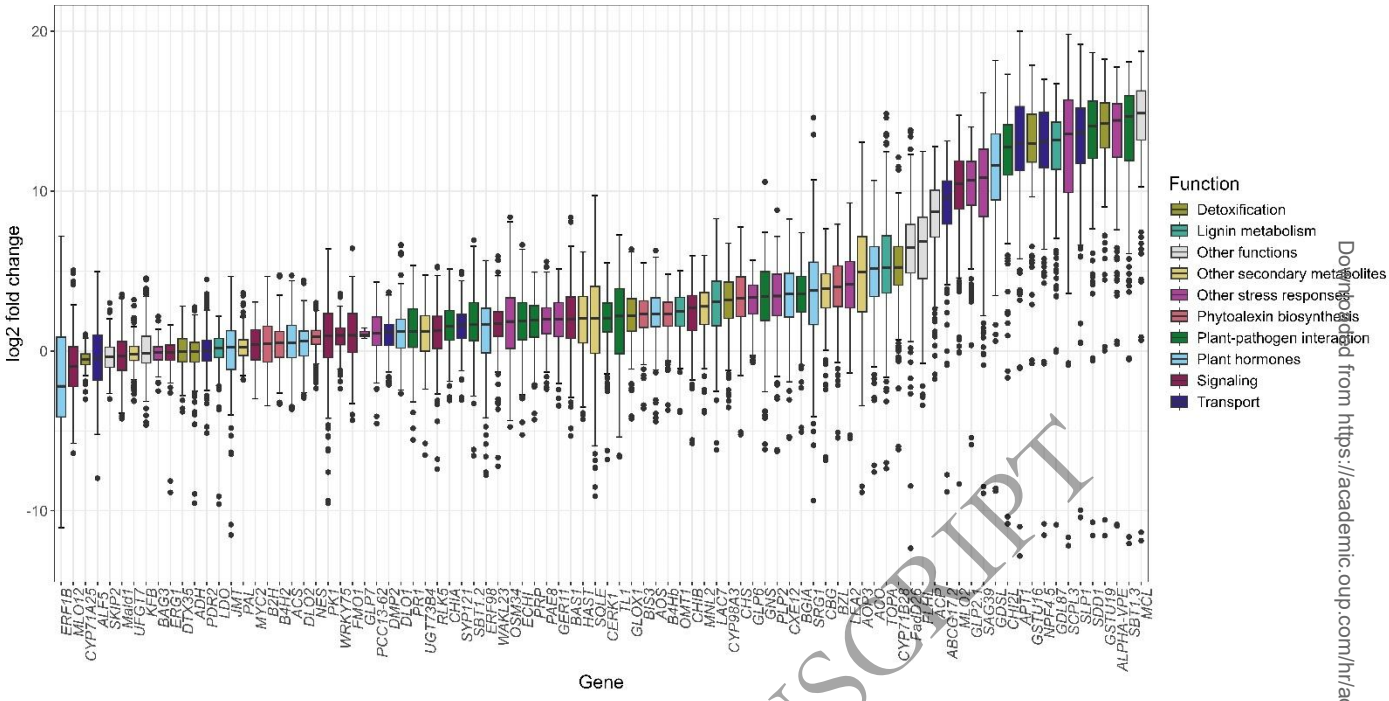

Figure 3: Differential gene expression in apple roots under ARD conditions compared to control conditions. The diagram shows the differential expression of 90 CGs (Tab. S1) in the roots of apple plants in response to ARD conditions. The $\log_2$ fold change (ARD-affected soil/disinfected control soil) is depicted as a boxplot, calculated with data obtained from up to 151 different soils. The genes are color-coded according to their functional classification.

Expression of novel postulated phytoalexin biosynthesis genes was upregulated under ARD conditions

For the first time, this study analyzed the expression of the genes *B2H*, *B4H2*, *BLZ,* and *NES*, which are assumed to be involved in phytoalexin biosynthesis (Liu *et al.,* unpublished), in apple roots under ARD conditions. The expression of the *B2H*, *B4H2* and *BLZ* genes were found to be significantly upregulated in ARD-affected roots ($p < 0.05$; Tab. S4). The *NES* gene expression also showed significant upregulation; however, the qPCR with the *NES* primers exhibited non-specific amplification in several samples.

Evaluation of the CGs for their suitability as biomarkers for ARD in roots

The present study primarily used root samples for the evaluation of the CGs as biomarkers for ARD, because the roots are the organs directly exposed to ARD-affected soil and the previous transcription analyses have mainly focused on root tissue [13,17-22,26-29]. For the evaluation, several factors were taken into account and each gene was assessed individually. The results of the evaluation are recorded in Table S6.

The first selection criterion was the qPCR performance of the gene-specific primer pair, which encompasses both amplification efficiency and specificity (Tab. S6, PCR quality), as well as the number of soil samples that passed the Fluidigm software quality check (Tab. S6, n soils). For instance, the PCR performed with *GLP7*-specific primers was disregarded because only three samples passed the qPCR quality check and *ALF5* was excluded because the melt curve analysis indicated non-specific

amplification. Additionally, the strength of expression was taken into consideration, as higher expression levels make detection by qPCR more robust (Tab. S6, Ct). A total of 24 genes were deselected due to their poor qPCR performance (Tab. S6, Assessment: PCR issues).

The second selection criterion was the differential expression of each gene under ARD conditions. For this, the mean fold change value for all soil samples was evaluated (Tab. S6, FC, log_2FC) and the significance of the upregulation under ARD conditions was calculated based on the calibrated normalized relative quantification of gene expression (Tab. S6, Difference ARD/Control). Sixteen genes were deselected due to a lack of upregulated expression, and an additional eight genes were deselected due to low fold changes (Tab. S6, Assessment: low fold change). Furthermore, the consistency of differential expression across the different soils was assessed for each gene, as illustrated exemplarily in Figure S3 for two genes. Both genes exhibited significantly upregulated expression under ARD conditions and similar mean fold changes (Tab. S6); however, the expression of *BGIA* was induced much more consistently across the different soils than that of *OSM34* (Fig. S3). A total of 17 genes were deselected due to inconsistent upregulation of expression (Tab. S6, Assessment: inconsistent upregulation).

The third criterion was the correlation of the differential gene expression with the growth difference of plants grown in ARD and control soil (Tab. S7). These growth parameters encompass the shoot dry mass reduction, shoot fresh mass reduction and root fresh mass reduction, and serve as indicators of ARD severity [34,39]. Overall, genes with moderate fold change values showed stronger correlations than those with very high values. A total of 13 genes were deselected due to lack of correlation with the growth parameters (Tab. S6, Assessment: lack of growth correlation).

Based on the evaluation, the qPCR results for 12 genes that were rated positively in all criteria were designated as the most promising potential ARD biomarkers (category I, Tab. 1). For 35 genes, the evaluation was positive for the majority of the criteria, resulting in the classification as potential ARD biomarkers (category II, Tab. S6). Lastly, 43 genes were deemed not suited as ARD biomarkers based on the results of the performed qPCRs (category III, Tab. S6). Category I includes four phytoalexin biosynthesis genes: *B4Hb, BIS3, BZL* and *CHS*. *CHS* is also involved in flavonoid biosynthesis. Three genes involved in lignin metabolism, *LAC7*, *OMT1* and *TOPA*, were also classified as category I, as were the cyanogenesis related gene *MNL2*, the ethylene biosynthesis gene *ACO*, the plant defense related gene *BGIA*, the detoxification related gene *CYP98A3*, and the cell death related gene *PLP2*. Figure 4A shows that the normalized relative expression of the category I genes was significantly higher ($p < 0.001$) in the roots of ARD-affected plants compared to roots of plants grown in the control soils.

Table 1: CGs rated most promising for their suitability as biomarkers for ARD in roots.
The CGs were evaluated for their potential to serve as biomarkers for ARD. This table shows the most promising potential ARD biomarkers (category I), the evaluation of the other genes can be found in Table S6. For each gene the table shows the mean Ct value of the ARD samples (Ct), the mean fold change (FC) and log_2 fold change (log_2 FC), the p-value and significance level of a paired Wilcoxon signed rank test of difference in gene expression between the ARD and control samples (difference ARD/Control), the qPCR performance of the gene specific primer pair (PCR quality), the number of successfully analyzed soil samples (n soils), and the Spearman correlation coefficient and significance level of gene expression with shoot dry mass reduction (SDM_red). Significance level: *** $p < 0.001$.

Gene	Ct	FC	Log_2 FC	Difference ARD/Control		PCR quality	n soils	Correlation SDM_red	
ACO	10.08	76.98	4.70	0.00E+00	***	good	148	0.310	***
B4Hb	5.54	6.14	2.21	0.00E+00	***	good	129	0.339	***
BGIA	5.99	19.25	3.47	0.00E+00	***	good	144	0.496	***
BIS3	4.24	6.29	2.18	0.00E+00	***	good	130	0.324	***
BZL	10.54	31.36	3.89	0.00E+00	***	good	149	0.205	***
CHS	7.63	18.70	3.24	5.43E-16	***	good	149	0.225	***
CYP98A3	9.99	15.22	2.98	0.00E+00	***	good	149	0.386	***
LAC7	14.47	20.80	2.92	0.00E+00	***	good	149	0.256	***
MNL2	6.09	10.06	2.59	0.00E+00	***	good	138	0.347	***
OMT1	5.74	7.55	2.35	0.00E+00	***	good	138	0.414	***
PLP2	13.32	25.28	3.44	0.00E+00	***	good	148	0.411	***
TOPA	10.48	612.81	5.40	0.00E+00	***	good	147	0.244	***

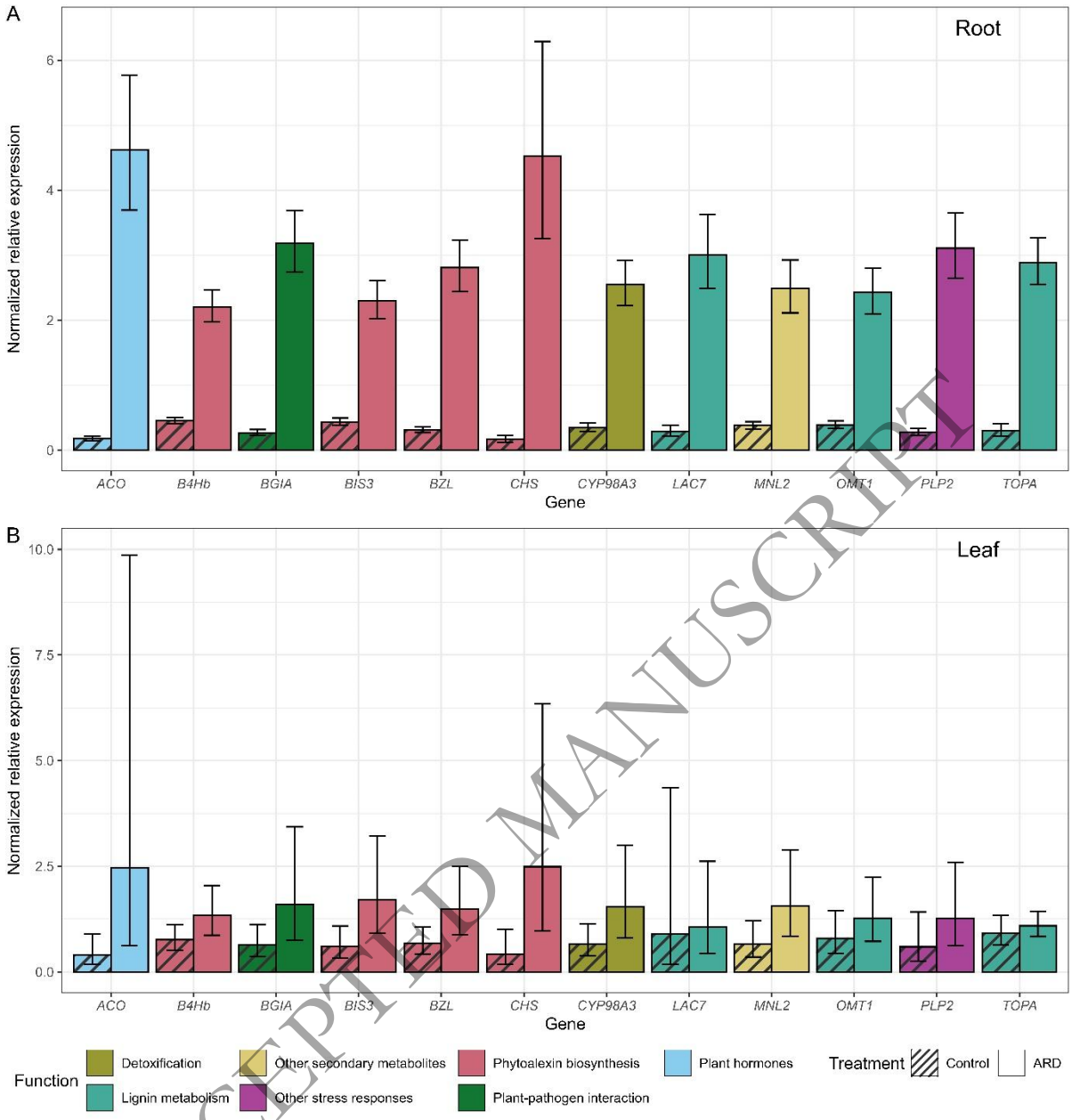

Figure 4: Normalized relative expression of the suggested ARD biomarkers.
The expression of the 12 genes classified as the most promising ARD biomarkers was analyzed by qPCR in the roots (**A**) and leaves (**B**) of apple plants grown on up to 151 (roots) or 18 (leaves) ARD-affected soils and disinfected control samples of each soil. The normalized relative expression values are the averages for all soils, the 95% confidence interval is represented as error bar, and the genes are color-coded according to their functional classification. **A)** The expression of the genes was significantly higher in the roots of ARD-affected plants compared to the control ($p < 0.001$). **B)** Only the expression of *CHS* was significantly higher in the leaves of ARD-affected plants compared to the control ($p < 0.05$).

Expression of the CGs in leaves in response to ARD

Since taking leaf samples is much more convenient and feasible than taking root samples, it would be preferable to use leaf material to analyze biomarker gene expression. Thus, the expression of the 90 CGs was analyzed in the leaves of apple plants grown in a subset of 18 different ARD-affected soils and disinfected control samples of each soil (Fig. 5). A total of 6 840 qPCR reactions were performed

using the high-throughput system. Of those, 1 718 were excluded from the analysis due to technical issues, proportionally three times more than recorded for the root samples (see above). The expression of 13 genes in leaves was too low to obtain usable data (Tab. S5). For 50 genes, samples from plants grown in 18 soils were used in the analysis, for 14 genes, samples from over 10 soils were used and for 13 genes samples from fewer than 10 soils were used. The differential expression of the CGs under ARD conditions exhibited lower fold changes in leaves than in roots, as the fold change values for individual samples ranged from 0.015 to 171.43 (log_2FC -6.02 to 10.19). The gene expression data for each soil sample is available in the BonaRes Repository [33].

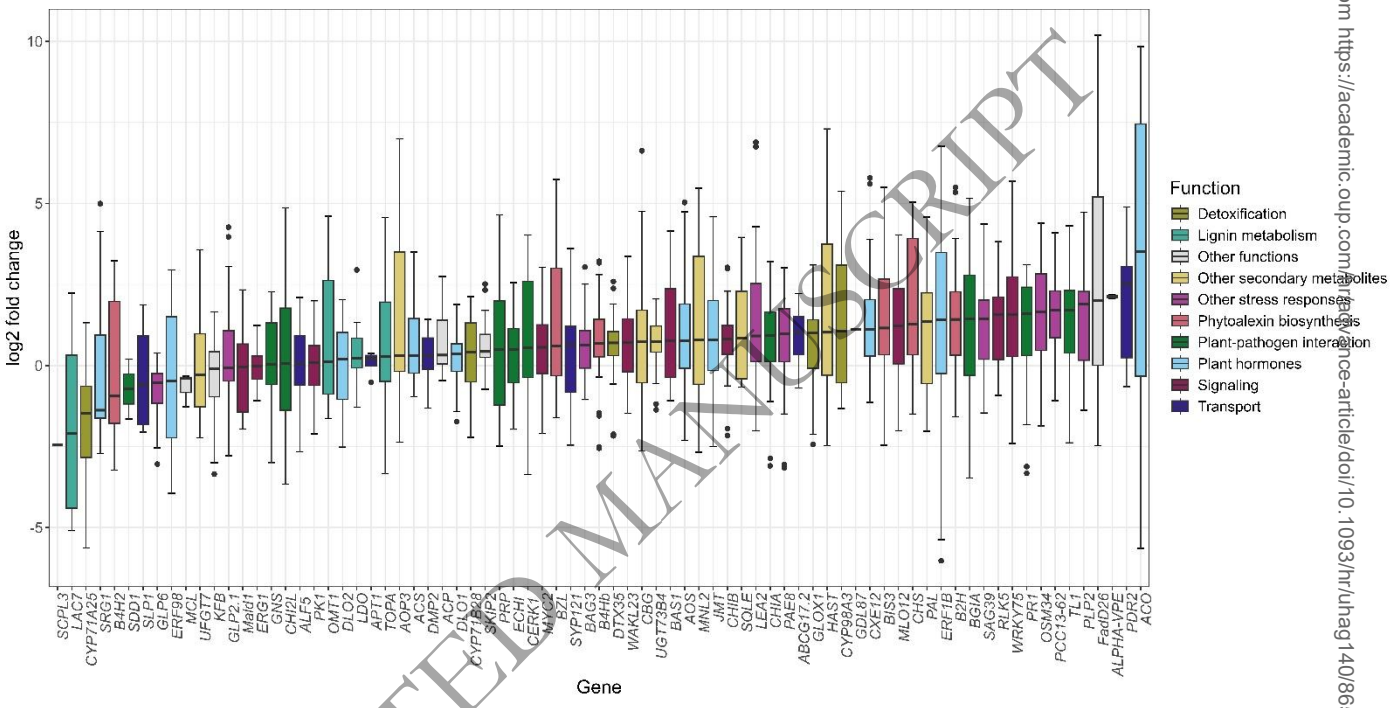

Figure 5: Differential gene expression under ARD conditions in leaves.
The diagram shows the differential expression (log_2 fold change values) of 77 CGs (Tab S1) in the leaves of apple plants in response to ARD conditions. The log_2 fold change values (ARD-affected soils/disinfected control soils) are depicted as a boxplot, calculated with data obtained from up to 18 different soils. The genes are color-coded according to their functional classification. The expression of the remaining 13 CGs in leaves was too low to obtain usable data.

Overall, the significance test of the differential gene expression showed higher p-values in the leaf samples compared to the root samples (Tab. S4 and S5). A total of 11 genes exhibited significantly upregulated expression in leaves under ARD conditions, with six genes exhibiting average fold change values greater than 4.5 (log_2FC: 2.2; Tab. S5). Of the 12 genes identified as the most promising biomarkers in roots, only the *CHS* gene was found to be significantly overexpressed ($p < 0.05$) in leaves as well (Fig. 4B). Interestingly, the *CYP71A25* and *GLP6* genes showed, on average, higher levels of expression in the control compared to the ARD samples. The remaining 64 genes did not show a significant change in expression in ARD soil in comparison with the control.

Evaluation of the CGs for their suitability as biomarkers for ARD in leaves

As with the gene expression in roots, the CGs were evaluated for their potential of being used as biomarkers in leaves and classified into three categories based on their qPCR performance, their differential expression and their correlation with the growth parameters (Tab. S8). Based on the evaluation, four genes were classified as the most promising potential ARD biomarkers (category I, Tab. 2), 17 genes were classified as potentially suitable ARD biomarkers (category II, Tab. S8), and the qPCR results revealed 69 genes to be not suited as ARD biomarkers (category III, Tab. S8). The category I genes include the phytoalexin and flavonoid biosynthesis gene *CHS*, which was identified as one of the most promising potential biomarker genes in roots as well. A second putative phytoalexin gene, *B2H*, was also classified as category I gene, as were the stress response related gene *PCC13-62* and the pathogen defense response gene *OSM34.* Figure 6 shows that the normalized relative expression of the category I genes was significantly higher ($p < 0.05$) in ARD-affected plants compared to the control.

Table 2: CGs rated most promising for their suitability as biomarkers for ARD in leaves.
The CGs were evaluated for their potential to serve as biomarkers for ARD. This table shows the most promising potential ARD biomarkers (category I). The evaluation of the other genes can be found in Table S8. For each gene the table shows the mean Ct value of the ARD samples (Ct), the mean fold change (FC) and log_2 fold change (log_2 FC), the result of a paired Wilcoxon signed rank test of difference in gene expression between the ARD and control samples (difference ARD/Control), the qPCR performance of the gene specific primer pair (PCR quality), the number of successfully analyzed soil samples (n soils), and the Spearman correlation coefficient and significance level of gene expression with shoot dry mass reduction (SDM_red). Significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Gene	Ct	FC	Log_2 FC	Difference ARD/Control		PCR quality	n soils	Correlation SDM_red	
B2H	11.01	5.97	1.43	3.94E-02	*	good	18	0.534	**
CHS	12.56	7.20	1.73	1.97E-02	*	good	18	0.296	*
OSM34	8.49	5.10	1.59	2.18E-02	*	good	18	0.532	**
PCC13-62	15.86	4.49	1.62	9.83E-03	**	good	18	0.686	***

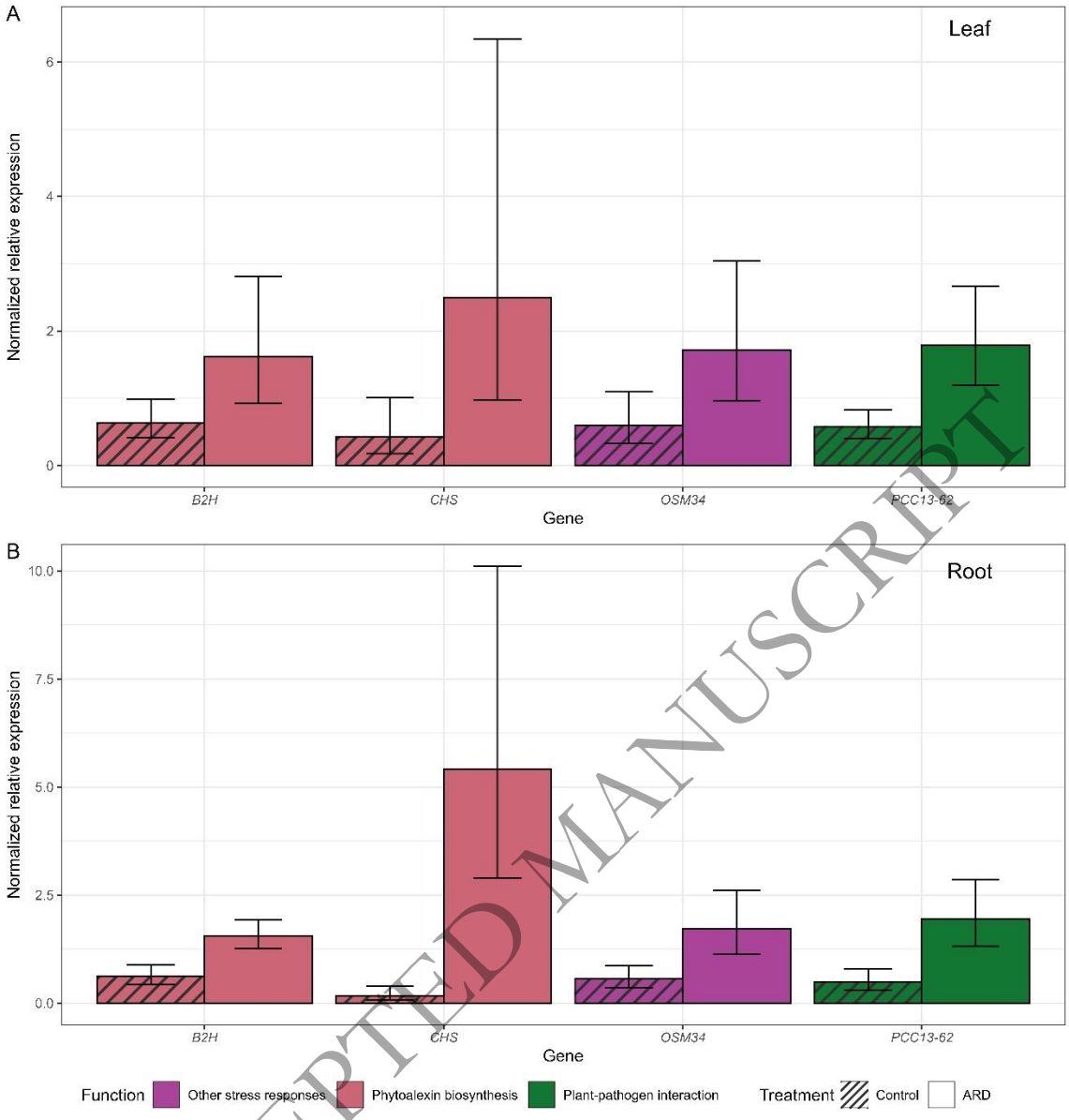

Figure 6: Normalized relative expression of most promising potential ARD biomarkers in leaves. The expression of the four genes classified as the most promising ARD biomarkers in leaves was analyzed by qPCR in the leaves **(A)** and roots **(B)** of apple plants grown on up to 18 ARD-affected soils and disinfected control samples of each soil. The normalized relative expression values are the averages for all soils, the 95% confidence interval is represented as error bar, and the genes are color-coded according to their functional classification. The expression of the genes was significantly higher in ARD-affected plants compared to the control ($p < 0.05$).

Discussion

The majority of CGs was upregulated in roots under ARD conditions

In order to reliably assess the severity of ARD in a soil, there is interest in biomarkers that can universally detect and quantify the ARD response of plants, regardless of the soil's agro-environmental origin [20,21]. This study analyzed the expression of 90 CGs in the roots of 'M.26' plants grown in ARD-affected soils from 151 different sites across Germany. To compare apple plants grown in ARD soils to those grown in non-ARD soils with the same properties, half of each soil sample was left untreated, while the other half was disinfected using gamma-irradiation. This method has been used in many

previous studies [12,15,17-22,30,40] and has the advantage of reversing ARD through disinfection while preserving most of the chemical and physical soil properties.

Overall, a wide variation in the degree of upregulation of gene expression was observed (Fig. 3), with some genes exhibiting very high fold changes of over 1 000 (log_2FC: 10.0). These very large fold change values were caused by extremely low expression in the control samples and higher expression in the ARD samples. The expression of the majority of the genes (72) was significantly upregulated under ARD conditions (Tab. S4). However, the remaining 18 genes were not significantly upregulated in roots from ARD-affected soils compared to roots from control soils. This finding was unexpected, as nearly all CGs (86) were selected for their reported overexpression under ARD conditions. For example, *ERF1B,* an ethylene related-gene previously recommended as an ARD biomarker [18,20,21], showed a wide variance in differential expression among different soils in this study. *ERF1B* fold changes ranged from 0.0004 to 114, and *ERF1B* expression was upregulated in 39 soils but downregulated in 74 soils [33]. These results align with those of Orth *et al.* [15], who found that *ERF1B* expression was only upregulated in one of the five tested ARD-affected soils. This demonstrates the significant influence that soil has on the plants' molecular reactions to ARD, due to the different abiotic and biotic soil characteristics as well as the replant histories [15,18,20]. Therefore, it is crucial for the identification of robust biomarker genes to analyze their expression in a large number of different ARD-affected soils.

Another reason for the lack of agreement with the previous literature of the expression data for some of the genes could be that the genes were analyzed in different apple genotypes [18,19,30,31]. Additionally, some genes were shown to be upregulated in previous RNA-sequencing analyses [19]; however, their expression may be too low for robust detection by the performed qPCR assays (e.g., *FMO1, GER11,* and *GLP7*). Further, a technical consideration has to be taken into account, because high-throughput qPCR using the BioMark HD system was susceptible to technical problems, such as air bubbles in the reaction chambers. The strict quality control measures excluded samples with these errors. Finally, different primers may produce better results for some genes and could be designed and tested in future studies.

In this study, the expression of nine genes was described as being upregulated in apple plants grown in ARD affected soil for the first time. These include the putative phytoalexin biosynthesis genes *B2H*, *B4H2*, *BZL* and *NES*, confirming the important role of phytoalexins in the ARD plant response. Previous studies have shown the induction of phytoalexin-related genes, as well as the correlation of their expression with phytoalexin production and ARD symptom severity [15,18,20,21,24]. Furthermore, the expression of CGs, which were previously only described after infection with ARD-associated pathogens, was shown to be upregulated in in this study. These comprise *ACS*, *CBG*, *CERK1*, *SQLE* [27,28] and *ERF98* [29].

Promising potential ARD biomarkers in roots have various functions in the ARD response

The twelve genes evaluated as the most promising potential ARD biomarkers in roots (category I, Tab. 1) were significantly and consistently upregulated under ARD conditions compared to control conditions across different soils. Their expression correlated with ARD severity (Tab. S6), which was measured by the reduction of plant growth [39].

The potential ARD biomarkers include *BIS3* and *B4Hb*, which have been previously recommended as biomarkers as well [13,17-21,30]. These genes are involved in biphenyl formation in the phytoalexin biosynthetic pathway [41,42] and their expression has been shown to correlate with phytoalexin formation and ARD symptoms [15,18,20,30]. Furthermore, additional genes from category I, *BZL* and *CHS*, are also related to phytoalexin formation. Lastly, the *ACO* gene was also classified as category I

and is involved in ethylene biosynthesis, which can induce phytoalexins of the phenylpropanoid pathway [43]. This underpins the importance of the phytoalexin pathways during the ARD response of the plant. It is known that apple roots exude phytoalexins into the soil in a compound-specific percentage, where they play a role in shaping the soil microbiome through their antimicrobial activities, which might influence ARD severity [15,44]. However, the overall impact of phytoalexins on ARD development is not yet fully understood, as they can cause a decrease of both pathogenic and beneficial microorganisms in the soil. The induction of phytoalexins in response to ARD is genotype-specific, and the production of individual phytoalexins may influence the ARD susceptibility of different genotypes [25]. Weiß *et al.* [17] hypothesized that the phytoalexin defense response of the ARD-susceptible genotype 'M.26', which was used in this study, is ineffective against ARD and that the roots might even be damaged by the high concentration of phytoalexins.

Further genes classified in category I include the lignin biosynthesis genes *LAC7*, *OMT1*, and *TOPA*. Reim *et al.* [19] proposed that increased cell wall stabilization through lignin biosynthesis reduces the negative effects of ARD on plants, which were also evident in term of disruptions of the outer root cell layers [12]. In this study, all analyzed lignin metabolism genes were significantly upregulated under ARD conditions, which emphasizes their role in the ARD response. Interestingly, *LAC7* and other laccase genes have shown increased expression after infection with the ARD-associated pathogen *Pythium ultimum* [28]. Zhu *et al.* [45] discovered that *P. ultimum* infection leads to the downregulation of a laccase-targeting miRNA, particularly in tolerant genotypes. The downregulation of this miRNA could lead to faster and stronger cell wall stabilization through increased laccase activity and thereby contribute to tolerance against the pathogen. Corroborating this finding, a QTL for *P. ultimum* tolerance has been identified near the putative location of this miRNA [46].

Another gene in category I is *CYP98A3*, which belongs to the cytochrome P450 superfamily. This family is involved in the detoxification of reactive oxygen species and bioactive compounds [47]. The category I gene *MNL2* has its function in cyanogenesis, which can play a role in the defense against fungi [48,49]. The expression of other cyanogenesis genes has been found to be induced after infection with *P. ultimum* [28], and the cyanogenic gene *CBG* was found upregulated in this study as well. Another gene in category I is *PLP2*, which plays a role in programmed cell death and may restrict pathogen growth [50]. Lastly, the gene *BGIA* was classified in category I. In *Arabidopsis,* a *BGIA* homolog was found to serve a function in plant defense and response to wounding [51].

All things considered, we suggest that the expression response of category I genes could be used as a diagnostic tool for detecting and quantifying ARD across a wide range of soils under controlled test conditions. However, other soil-independent factors may influence the expression of the proposed ARD biomarker genes. These factors should be controlled for or investigated in future studies, e.g. analyzing the expression under both abiotic and biotic stress conditions. Thus far, only *BIS3* and *B4Hb* have been examined in apple plants under abiotic stress conditions, where neither showed an upregulation response [20]. Furthermore, the ARD biomarkers' expression should also be investigated in other rootstock genotypes and at earlier time points after planting.

The ARD response gene expression pattern differs between roots and leaves

Since leaf samples are more convenient to take than root samples, it would be preferable to use leaf material to analyze biomarker gene expression. Thus, the expression of the CGs was also analyzed in the leaves (Fig. 5). Overall, the genes showed less significant differential expression in the leaf samples than in the root samples (Tab. S5), as expected, since most of the genes have previously only been described in roots. Several of the genes were barely expressed in the leaves. These results align with

462 those of Weiß & Winkelmann [22] and Reim *et al.* [18], who also found lower gene expression and
463 upregulation in leaves. However, the genes that were previously identified as being upregulated in
464 leaves upon growth in ARD-affected soils were not significantly upregulated in this study, which may
465 be due to the use of different soils.

466 Of the genes identified as the most promising biomarkers in roots, only the phytoalexin and flavonoid
467 biosynthesis gene *CHS* was also classified as category I in leaves. A second putative phytoalexin gene,
468 *B2H*, was also classified as a category I gene. Weiß & Winkelmann [22] also found the induced
469 expression of many phytoalexin genes in leaves as well, suggesting a delayed systemic response to
470 ARD. Other category I genes in leaves include the stress-responsive antifungal gene *OSM23,* which has
471 been found to be overexpressed in *Arabidopsis* infected with a fungal pathogen [52]. The *PCC13-6* gene
472 was classified as category I and plays a role in the abiotic water stress response [53], pointing to
473 potential water stress in the leaves of ARD-affected plants due to a damaged root system.

474 Given the smaller expression changes observed in leaves, it is questionable whether the sensitivity of
475 the proposed biomarker genes is sufficient for a reliable assessment of the ARD response, or whether
476 only strong ARD contamination can be diagnosed. This issue should be addressed in subsequent
477 studies.

478

479 **Conclusion**

480 In conclusion, this study identified genes that were consistently upregulated under ARD conditions
481 across many different soils. The results indicate that these genes play a significant role in the ARD
482 reaction in roots, but only to a limited extent in leaves. Potential ARD biomarkers that could reliably
483 assess ARD severity were selected. However, future studies are needed to analyze the selected
484 genes' specificity to ARD and their performance in other apple genotypes.
485

486 Materials and methods

487

488 **Soil and plant material**

489 Surface soil (0-30 cm) was collected from 151 different sites with apple cultivation history in Germany.
490 Half of each soil was disinfected by gamma-irradiation (>10 kGy). Successful disinfection was validated
491 via plating of soil suspensions (1 g in 9 mL 0.85% NaCl) onto R2A agar (Carl Roth, Germany) with 100 mg
492 L^{-1} cycloheximide or 1/3 potato dextrose agar (PDA, Carl Roth, Germany) with 100 mg L^{-1} penicillin,
493 10 mg L^{-1} tetracycline, 50 mg L^{-1} streptomycin. Colony forming units (CFUs) were counted after
494 inoculation for 48 h (bacteria, R2A) or 7 days (fungi, PDA) at room temperature in the dark.

495 *In vitro* propagated and rooted young plantlets of the rootstock 'M.26' were acclimatized for four
496 weeks in a peat substrate as described by Rohr *et al.* [40]. The biotest was carried out in nine sets of
497 up to 18 soils between 13.07.2022 and 20.08.2023. Each set included one variant with peat substrate
498 to normalize possible growth differences among the sets. Each of the 151 soils used as untreated ARD-
499 affected soil (ARD) and gamma-irradiated soil (Control) was first supplemented with 2 g L^{-1} of the slow-
500 release fertilizer Osmocote exact 3-4 M (N-P_2O_5-K_2O-MgO: 16-9-12-2, ICL Deutschland, Germany) and
501 per soil variant eight pots (600 mL each) were prepared with one 'M.26' plantlet per pot. The pots
502 were placed randomly onto four tables in a climate chamber and cultivated for six weeks (19.9 ± 2.6°C,
503 61.9 ± 7% relative humidity, 16/8 h day/night cycle with SON-T AGRO 400 W lamps (Philips,
504 Netherlands)). Irrigation with tap water was done manually on a daily basis. For evaluation of plant
505 growth and sampling, the two plants with the smallest and largest shoot length were excluded and the

remaining six plants per soil variant were used to collect growth data and samples. On the day of sampling, the roots were carefully washed with tap water. Per soil and treatment, 100 mg roots and the first fully expanded leaf of each of the six selected plants were pooled (roots and leaves separately) and immediately frozen in liquid nitrogen. The samples were stored at -80°C until RNA isolation.

RNA isolation and cDNA synthesis

Root and leaf pooled samples were homogenized with a Mixer Mill MM400 (Retsch, Germany) at 25 Hz for 2 min while cooled by liquid nitrogen. With the InviTrap Spin Plant RNA Mini Kit (Invitek Molecular, Germany), total RNA was isolated according to the manufacturer's instructions with RP lysis buffer and the RNA was eluded in 60 µl elution buffer R. Genomic DNA was removed with DNase I (Thermo Fisher Scientific, USA) following the manufacturer's instructions. RNA concentrations were measured using the Nanodrop One (Thermo Fisher Scientific, USA). The cDNA synthesis was performed with the RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, USA) using oligo dT primers and 1 µg RNA as template. The cDNA synthesis was verified in a PCR with *EF1* primers (5'-ATTGTGGTCATTGGYCAYGT-3'/5'-CCAATCTTGTAVACATCCTG-3') using 1 µl cDNA and genomic DNA as control [54]. The amplified PCR products from the cDNA and the genomic DNA have different fragment sizes (707bp/905bp) and no product would be generated with only RNA as template. Verified cDNA samples were stored at -20°C until further use.

Primer selection and qPCR validation

A total of 90 CGs were selected as potential biomarkers for ARD (Tab. S1). For 61 of these genes, new qPCR primers were designed with NCBI Primer Blast and tested *in silico* for specificity by blasting against the *Malus domestica* transcriptome (taxid:3750). The primer pairs were then tested for efficiency and specificity by qPCR using the BioRad CFX96 Real-Time System (Bio-Rad Laboratories, USA). The qPCR was performed with the Maxima SYBR Green master mix (Thermo Fisher Scientific, USA) using 1 µl 'M.26' apple cDNA diluted 1:10, 1:100, and 1:1000 as template and a final primer concentration of 75 nM. The reaction was conducted with an initial denaturation step at 95°C for a duration of 5 minutes, followed by 39 cycles of denaturation at 95°C for 10 s, annealing at 60°C for 30 s, and extension at 72°C for 30 s. The PCR products were analyzed by melt-curve analysis, ranging from 65°C to 95°C with an increment of 0.5°C for a duration of 5 s at each step. The amplification efficiencies were calculated with the CFX manager software. Afterwards, the amplification products were analyzed by gel electrophoresis. Only primer pairs that showed the expected amplification product size in the gel electrophoresis, one specific melt-curve peak, and sufficient amplification efficiency (85-115%) were selected for the following analysis.

Gene expression analysis

The expression of the 90 CGs was analyzed using the BioMark HD high-throughput qPCR system (Standard BioTools, USA). The genes *ACT7*, *EF1a*, *EF1b*, *TUBB*, and *UBE210* (Tab. S2) were used as reference genes according to Weiß *et al.* [21]. Equal amounts of 57 cDNA samples isolated from 'M.26' roots grown in ARD affected soils were mixed and diluted 1:10, 1:100, and 1:1000 in series to be used as standards for the calculation of amplification efficiencies. A total of 338 cDNA samples with two technical replicates each were analyzed with the 96.96 Dynamic Array™ IFC (Standard BioTools, USA) following the manufacturer's instructions. First, 1.25 µl of the cDNA were pre-amplified with the 5× PreAmp Master Mix (Standard BioTools, USA) and 500 nM pooled primers filled up to 5 µl with

nuclease-free water. The PCR was performed with an initial denaturation at 95°C for 10 min, followed by 14 cycles of 95°C for 15 s and 60°C for 4 min. Following this, the pre-amplified cDNAs were purified with exonuclease I (New England Biolabs GmbH, Germany), diluted 1:5 with DNA suspension buffer pH 8.0 (Alpha Teknova, USA) and stored at -20°C until use. For the sample-mix, 2.7 µl of the pre-amplified and diluted cDNA was mixed with 3 µl 2x SsoFast EvaGreen supermix with low ROX (Bio-Rad Laboratories, USA) and 0.3 µl 20x DNA binding dye sample loading reagent (Standard BioTools, USA). The assay-mix consisted of 0.5 mM primer mix, 2.5 µl 2x assay loading reagent (Standard BioTools, USA), and 2.25 µl DNA suspension buffer pH 8.0 (Alpha Teknova, USA). The qPCR was performed with 5 µl sample-mix and assay-mix per inlet in 96.96 Dynamic Array™ IFCs (Standard BioTools, USA). The cycling program was as follows: 95°C for 1 min, 30 cycles of 96°C for 5 s and 60°C for 20 s plus melting curve analysis from 60°C to 95°C with an increment of 1°C for a duration of 3 s at each step.

Data analysis

The normalization of the growth data of each biotest set was done by multiplying it by a normalization factor derived from the growth data of the plants grown in peat substrate, using the following formula:

$$\text{Normalisation Factor (Biotest x)} = \frac{\text{Mean (all Biotests)}}{\text{Mean (Biotest x)}}$$

The normalized growth reduction for each growth parameter (SDM, SFM, and RFM) was calculated as a percentage of the reduction in the ARD variant of a soil compared to the control variant of the same soil.

The initial gene expression data analysis was performed with the Fluidigm Real-Time PCR Analysis program version 4.8.1 (Standard BioTools, USA). To ensure the quality of the measurements, each amplification curve was scored based on how well it compared to an ideal curve and all samples that scored below the recommended threshold set by the Fluidigm program were excluded from further analysis. Samples showing unspecific amplification in the melt curve analysis were excluded as well. Additionally, the Fluidigm program calculated the amplification efficiencies of each primer pair. The stability of the reference genes was analyzed with RefFinder [38] and based on the results, the gene *UBE210* was excluded.

Next, the Fluidigm program calculated the ΔCt, ΔΔCt and fold change values for every sample according to the following formulas:

$$\Delta\text{Ct} = \text{Ct (candidate gene)} - \text{average Ct (reference genes)}$$

$$\Delta\Delta\text{Ct} = \Delta\text{Ct (ARD sample)} - \Delta\text{Ct (corresponding control sample)}$$

$$\text{Fold change} = 2^{-\Delta\Delta\text{Ct}}$$

Afterwards, statistical analysis was performed using R version 4.4.1 [55] and the $\log_2$FC was calculated for each sample. The correlation between the $\log_2$FC and the plant growth data (SDM_red, SFM_red, RFM_red) for each gene was analyzed by Spearman correlation.

Additionally, the calibrated normalized relative quantification of gene expression was determined with the qBase+ software version 3.4 (CellCarta, Canada). A paired Wilcoxon signed rank test was used to calculate the significance of the difference in gene expression between the ARD and control samples.

Acknowledgments and funding

The authors would like to thank Ines Hiller, Eric Fritzsche, Bärbel Ernst, and Eva Schneider for their technical assistance. We are grateful to Dr. Franco Röckel (Julius Kühn Institute (JKI) - Federal Research Centre for Cultivated Plants, Institute for Grapevine Breeding, Siebeldingen, Germany) for enabling us to conduct gene expression analyses using the BioMark HD system. We thank Dr. Benye Liu (Technical University Braunschweig, Institute of Pharmaceutical Biology, Braunschweig, Germany) for providing the sequences of the putative phytoalexin genes.

This work was part of the BonaRes-ORDIAmur project, funded by the German Federal Ministry of Research and Education as part of the BonaRes program (grant no. 031B1070B).

Contributions

HF, SR, TW, GG, and NM conceptualized the project. The experiment and sample procedure were designed by TW, EL, NM, GG, and JB. JKr and JKa conducted the soil sampling. NO, KH, and JKr conducted the growth experiment. LB, SR and SS designed and LB performed the gene expression analyses. LB and SS contributed to data analysis. The manuscript was written by LB and SS and edited by HF, NO, SR, TW, KH, NM, JKr, EL, GG and JB. All authors contributed to the article and approved the submitted version.

Data availability statement

The datasets presented in this study can be found in the supplementary material or are available in the BonaRes Repository [33-35].

Conflict of interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary information

Figure S1: Growth reduction of plants grown in ARD-affected soil.
'M.26' apple plants were grown in a greenhouse biotest for six weeks under ARD conditions using an untreated ARD-affected soil sample (**A**). A sample of the same soil that was disinfected using gamma-irradiation served as ARD-unaffected control (**B**). The plants shown in this figure were grown in the soil with the ID number 3659 [34].

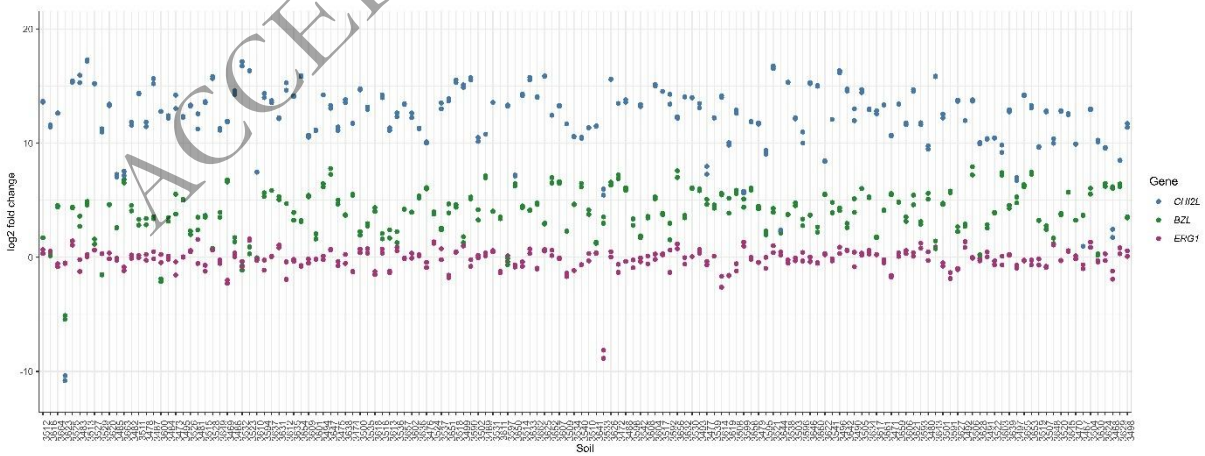

Figure S2: Relative expression values of three example genes for each soil sample.
The differential expression (log_2 fold change values) in roots under ARD conditions is shown for three exemplary genes for each soil sample: the non-upregulated *ERG1*, the significantly upregulated *BZL*, and the significantly and very strongly upregulated *CHI2L*. The log_2 fold change values are displayed separately for the two technical replicates. The soils are listed by their ID number [33].

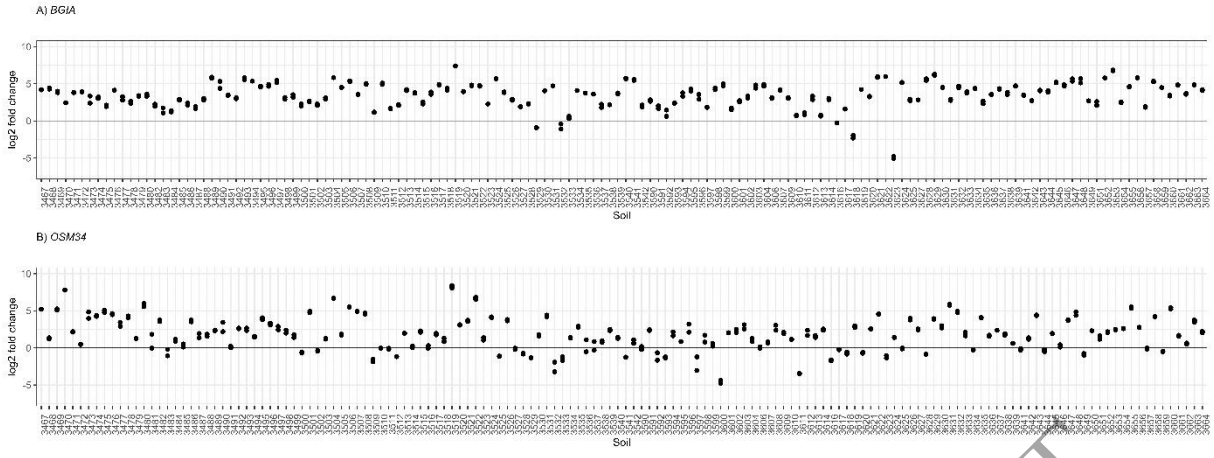

Figure S3: Consistency of expression upregulation across the different soils of two example genes. Differential expression (log_2 fold change values) under ARD conditions of two example genes for each soil. Both genes exhibit significant upregulation under ARD conditions and similar mean fold changes (Tab. S6); however, *BGIA* (**A**) is upregulated much more consistently across the different soils than *OSM34* (**B**). The log_2 fold change values (ARD-affected soils/Disinfected control soils) are displayed separately for the two technical replicates of each soil. The soils are listed by their ID number [33].

Table S1: Candidate genes.
List of 90 CGs that were selected from literature as potential ARD biomarkers. The table includes their apple gene identification V1.0 (MDP) and V1.1 (MD), genebank accession mRNA and protein number from NCBI, protein name and functions as described in literature, the functional classification used in this work, references, primer sequences, PCR product size, annealing temperature of the primer pair, the reference of the primers, and alternative gene names used in other works.

Table S2: Reference genes.
The table lists the reference genes with their MDP ID (Apple Gene Identification V1.0), protein name, primer sequences, size of the PCR product, primer annealing temperature, PCR amplification efficiency in the BioMark HD high-throughput qPCR, and primer reference.

Table S3: Reference gene stability analysis.
The stability of the reference genes was analyzed with RefFinder. This table lists the individual stability rankings of the programs BestKeeper, the comparative ΔCt method, geNorm, and NormFinder, as well as the comprehensive RefFinder ranking.

Table S4: Summary of the qPCR data in roots.
The gene expression analysis of the CGs in the roots of apple plants grown in in 151 different ARD-affected soils and disinfected control samples of each soil was performed by qPCR using the BioMark HD high-throughput system. This table shows the qPCR data for each gene in all soil samples, including the mean and median fold change value with standard deviation, the mean and median log_2 fold change value with standard deviation, the p-value of a paired Wilcoxon signed rank test of difference in gene expression between the ARD and control samples, the number of included soils and samples, the PCR amplification efficiency, and the mean melt peak temperature of the PCR product with standard deviation. The gene expression data for every individual sample is available in the BonaRes Repository [33].

Table S5: Summary of the qPCR data in leaves.
The gene expression analysis of the CGs in the leaves of apple plants grown in 18 different ARD-affected soils and disinfected control samples of each soil was performed by qPCR using the BioMark HD high-throughput system. This table shows the qPCR data for each gene for all soil samples, including the mean and median fold change value with standard deviation, the mean and median log_2 fold change value with standard deviation, the p-value of a paired Wilcoxon signed rank test of difference in gene expression between the ARD and control samples, the number of included soils and samples, the PCR amplification efficiency, and the mean melt peak temperature of the PCR product with standard deviation. The gene expression data for every individual sample is available in the BonaRes Repository [33].

Table S6: Evaluation of the CGs for their suitability as biomarkers for ARD in roots.
The CGs were evaluated for their potential to serve as biomarkers for ARD and classified into three different categories. Category I: most promising potential ARD biomarkers, category II: potentially suitable ARD biomarkers, category III: not suited as ARD biomarkers. For each gene the table shows the mean Ct value of the ARD samples (Ct), the mean fold change (FC) and log_2 fold change (log_2 FC), the result of a paired Wilcoxon signed rank test of difference in gene expression between the ARD and control samples (difference ARD/Control), the qPCR performance of the gene specific primer pair (PCR quality), the number of successfully analyzed soil samples (n soils), the Spearman correlation coefficient and significance level of gene expression with the growth parameters (SDM_red, SFM_red, RFM_red), and a comment for deselection or selection as potential biomarker (assessment). Significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, n.s. $p > 0.05$.

Table S7: Correlation of gene expression with growth data.
The table shows the Spearman correlation coefficient and p-value of the differential gene expression in root and leaf samples (log_2 FC) with the growth difference of apple plants grown in ARD and control soil. The growth parameters encompass the shoot dry mass reduction (SDM_red), shoot fresh mass reduction (SFM_red), and root fresh mass reduction (RFM_red).

Table S8: Evaluation of the CGs for their suitability as biomarkers for ARD in leaves.
The CGs were evaluated for their potential to serve as biomarkers for ARD and classified into three different categories. Category I: most promising potential ARD biomarkers, category II: potentially suitable ARD biomarkers, category III: not suited as ARD biomarkers. For each gene the table shows the mean Ct value of the ARD samples (Ct), the mean fold change (FC) and log_2 fold change (log_2 FC), the result of a paired Wilcoxon signed rank test of difference in gene expression between the ARD and control samples (difference ARD/Control), the qPCR performance of the gene specific primer pair (PCR quality), the number of successfully analyzed soil samples (n soils), the Spearman correlation coefficient and significance level of gene expression with the growth parameters (SDM_red, SFM_red, RFM_red), and a comment for deselection or selection as potential biomarker (assessment). Significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, n.s. $p > 0.05$, NaN = missing data.

References

1. Winkelmann T, Smalla K, Amelung W, et al. Apple Replant Disease: Causes and Mitigation Strategies. *Current Issues in Molecular Biology*. 2019 2019:89-106. doi:10.21775/cimb.030.089
2. Mazzola M. Elucidation of the microbial complex having a causal role in the development of apple replant disease in washington. *Phytopathology*. 1998/09// 1998;88(9):930-938. doi:10.1094/PHYTO.1998.88.9.930

3. Mazzola M, Manici LM. Apple replant disease: role of microbial ecology in cause and control. *Annual Review of Phytopathology*. 2012 2012;50:45-65. doi:10.1146/annurev-phyto-081211-173005
4. Grunewaldt-Stöcker G, Mahnkopp F, Popp C, Maiss E, Winkelmann T. Diagnosis of apple replant disease (ARD): Microscopic evidence of early symptoms in fine roots of different apple rootstock genotypes. *Scientia Horticulturae*. 2019/01/03/ 2019;243:583-594. doi:https://doi.org/10.1016/j.scienta.2018.09.014
5. George Geldart H. THE IMPACT OF REPLANT PROBLEM ON THE ECONOMICS OF HIGH DENSITY APPLE PLANTINGS. International Society for Horticultural Science (ISHS), Leuven, Belgium; 1994:11-18.
6. Yim B, Smalla K, Winkelmann T. Evaluation of apple replant problems based on different soil disinfection treatments—links to soil microbial community structure? *Plant and Soil*. 2013/05/01 2013;366(1):617-631. doi:10.1007/s11104-012-1454-6
7. Somera TS, Mazzola M. Toward a holistic view of orchard ecosystem dynamics: A comprehensive review of the multiple factors governing development or suppression of apple replant disease. *Frontiers in Microbiology*. 2022/07/25/ 2022;13:949404. doi:10.3389/fmicb.2022.949404
8. Wang G, Yin C, Pan F, et al. Analysis of the Fungal Community in Apple Replanted Soil Around Bohai Gulf. *Horticultural Plant Journal*. 2018/09/01/ 2018;4(5):175-181. doi:https://doi.org/10.1016/j.hpj.2018.05.003
9. Popp C, Grunewaldt-Stöcker G, Maiss E. A soil-free method for assessing pathogenicity of fungal isolates from apple roots. *Journal of Plant Diseases and Protection*. 2019;126(4):329-341.
10. Manici LM, Caboni E, Caputo F, Frattarelli A, Lucioli S. Phytotoxins from Dactylonectria torresensis involved in replant disease of fruit trees. *Rhizosphere*. 2021/03/01/ 2021;17:100300. doi:https://doi.org/10.1016/j.rhisph.2020.100300
11. Mahnkopp-Dirks F, Radl V, Kublik S, Gschwendtner S, Schloter M, Winkelmann T. Molecular Barcoding Reveals the Genus Streptomyces as Associated Root Endophytes of Apple (Malus domestica) Plants Grown in Soils Affected by Apple Replant Disease. *Phytobiomes Journal*. 2021/01/01 2020;5(2):177-189. doi:10.1094/PBIOMES-07-20-0053-R
12. Yim B, Winkelmann T, Ding G-C, Smalla K. Different bacterial communities in heat and gamma irradiation treated replant disease soils revealed by 16S rRNA gene analysis – contribution to improved aboveground apple plant growth? Original Research. *Frontiers in Microbiology*. 2015-November-06 2015;Volume 6 - 2015doi:10.3389/fmicb.2015.01224
13. Balbín-Suárez A, Jacquiod S, Rohr A-D, et al. Root exposure to apple replant disease soil triggers local defense response and rhizoplane microbiome dysbiosis. *FEMS Microbiology Ecology*. 2021/04/01/ 2021;97(4):fiab031. doi:10.1093/femsec/fiab031
14. Simon M, Lehndorff E, Wrede A, Amelung W. In-field heterogeneity of apple replant disease: Relations to abiotic soil properties. *Scientia Horticulturae*. 2020/01/03/ 2020;259:108809. doi:https://doi.org/10.1016/j.scienta.2019.108809
15. Orth N, Krueger J, Liu B, et al. Linking soil characteristics, rhizosphere microbiome composition, and plant defence reaction to apple replant disease severity. *Plant and Soil*. 2024/12/12 2024;doi:10.1007/s11104-024-07091-x
16. Hauschild K, Giongo A, Liu B, et al. Catch crop amendments and microbial inoculants differently modulate apple rhizosphere microbiomes and plant responses. *FEMS Microbiology Ecology*. 2025;101(6):fiaf055. doi:10.1093/femsec/fiaf055
17. Weiß S, Liu B, Reckwell D, Beerhues L, Winkelmann T. Impaired defense reactions in apple replant disease-affected roots of Malus domestica 'M26'. *Tree Physiology*. 2017/12/01/ 2017;37(12):1672-1685. doi:10.1093/treephys/tpx108
18. Reim S, Rohr A-D, Winkelmann T, et al. Genes Involved in Stress Response and Especially in Phytoalexin Biosynthesis Are Upregulated in Four Malus Genotypes in Response to Apple Replant Disease. *Frontiers in Plant Science*. 2020 2020;10

19. Reim S, Winkelmann T, Cestaro A, Rohr A-D, Flachowsky H. Identification of Candidate Genes Associated With Tolerance to Apple Replant Disease by Genome-Wide Transcriptome Analysis. *Frontiers in Microbiology*. 2022 2022;13:888908. doi:10.3389/fmicb.2022.888908
20. Rohr A-D, Schimmel J, Liu B, Beerhues L, Guggenberger G, Winkelmann T. Identification and validation of early genetic biomarkers for apple replant disease. *PLOS ONE*. 2020/09/24/ 2020;15(9):e0238876. doi:10.1371/journal.pone.0238876
21. Weiß S, Bartsch M, Winkelmann T. Transcriptomic analysis of molecular responses in Malus domestica 'M26' roots affected by apple replant disease. *Plant Molecular Biology*. 2017/06/01/ 2017;94(3):303-318. doi:10.1007/s11103-017-0608-6
22. Weiß S, Winkelmann T. Transcriptome profiling in leaves representing aboveground parts of apple replant disease affected Malus domestica 'M26' plants. *Scientia Horticulturae*. 2017/08/19/ 2017;222:111-125. doi:10.1016/j.scienta.2017.05.012
23. Busnena BA, Beerhues L, Liu B. Biphenyl and Dibenzofuran Phytoalexins Differentially Inhibit Root-Associated Microbiota in Apple, Including Fungal and Oomycetal Replant Disease Pathogens. *Phytopathology®*. 2025;115(2):181-191. doi:10.1094/phyto-03-24-0088-r
24. Busnena BA, Beuerle T, Mahnkopp-Dirks F, Winkelmann T, Beerhues L, Liu B. Formation and exudation of biphenyl and dibenzofuran phytoalexins by roots of the apple rootstock M26 grown in apple replant disease soil. *Phytochemistry*. 2021/12/01/ 2021;192:112972. doi:https://doi.org/10.1016/j.phytochem.2021.112972
25. Siefen N, Staudt J, Busnena BA, et al. Differential accumulation of phenolics and phytoalexins in seven Malus genotypes cultivated in apple replant disease-affected soil. *Scientia Horticulturae*. 2024/03/15/ 2024;328:112902. doi:10.1016/j.scienta.2024.112902
26. Zhu Y, Shin S, Mazzola M. Genotype responses of two apple rootstocks to infection by Pythium ultimum causing apple replant disease. *Canadian Journal of Plant Pathology*. 2016/10/01/ 2016;38(4):483-491. doi:10.1080/07060661.2016.1260640
27. Shin S, Lv J, Fazio G, Mazzola M, Zhu Y. Transcriptional regulation of ethylene and jasmonate mediated defense response in apple (Malus domestica) root during Pythium ultimum infection. *Horticulture Research*. 2014/10/29/ 2014;1(1):1-10. doi:10.1038/hortres.2014.53
28. Shin S, Zheng P, Fazio G, Mazzola M, Main D, Zhu Y. Transcriptome changes specifically associated with apple (Malus domestica) root defense response during Pythium ultimum infection. *Physiological and Molecular Plant Pathology*. 2016/04/01/ 2016;94:16-26. doi:10.1016/j.pmpp.2016.03.003
29. Xiang L, Wang M, Pan F, et al. Transcriptome analysis Malus domestica 'M9T337' root molecular responses to Fusarium solani infection. *Physiological and Molecular Plant Pathology*. 2021/01/01/ 2021;113:101567. doi:10.1016/j.pmpp.2020.101567
30. Hauschild K, Orth N, Giongo A, et al. Apple Genotypes and Soil Properties Shape the Soil Microbiome and Influence the Plant Response to Replant Disease-Affected Soil. *Phytobiomes Journal*. 2025/01/23/ 2025;doi:10.1094/PBIOMES-08-24-0082-R
31. Wang L, Somera TS, Hargarten H, Honaas L, Mazzola M. Comparative Analysis of the Apple Root Transcriptome as Affected by Rootstock Genotype and Brassicaceae Seed Meal Soil Amendment: Implications for Plant Health. *Microorganisms*. Apr 6 2021;9(4)doi:10.3390/microorganisms9040763
32. ORDIAmur. BonaRes (Modul A): ORDIAmur. Accessed 14.08.25, 2025. www.ordiamur.de
33. Baader L, Schröpfer S, Reim S, et al. Data from: ORDIAmur Central Experiment 12 – gene expression data. 2026. *BonaRes Repository*. doi:10.20387/BONARES-HKD9-YW18
34. Winkelmann T, Orth N. Data from: ORDIAmur Central Experiment 12 (CE12) Growth Data. 2026. *BonaRes Repository*. doi:10.20387/BONARES-NSXD-TJ79
35. Kaldun J, Lehndorff E, Krüger J, Guggenberger G, Boy J. Data from: ORDIAmur Central Experiment 12 – Soil parameters for 151 ARD soils. *BonaRes Repository*. 2026. doi:10.20387/BONARES-N3HM-W883
36. Consortium TU. UniProt: the Universal Protein Knowledgebase in 2025. *Nucleic Acids Research*. 2025;53(D1):D609-D617. doi:10.1093/nar/gkae1010

37. Kanehisa M, Furumichi M, Sato Y, Matsuura Y, Ishiguro-Watanabe M. KEGG: biological systems database as a model of the real world. *Nucleic Acids Res*. Jan 6 2025;53(D1):D672-d677. doi:10.1093/nar/gkae909

38. Xie F, Wang J, Zhang B. RefFinder: a web-based tool for comprehensively analyzing and identifying reference genes. *Funct Integr Genomics*. 2023/04/15/ 2023;23(2):125. doi:10.1007/s10142-023-01055-7

39. Mahnkopp F, Simon M, Lehndorff E, Pätzold S, Wrede A, Winkelmann T. Induction and diagnosis of apple replant disease (ARD): a matter of heterogeneous soil properties? *Scientia Horticulturae*. 2018/11/18/ 2018;241:167-177. doi:https://doi.org/10.1016/j.scienta.2018.06.076

40. Rohr A-D, Staudt J, Cziborra K, Fritz A, Schmitz M, Winkelmann T. Split-root approach reveals localized root responses towards apple replant disease (ARD) in terms of ARD biomarker gene expression and content of phenolic compounds. *Scientia Horticulturae*. 2021/08/25/ 2021;286:110117. doi:https://doi.org/10.1016/j.scienta.2021.110117

41. Busnena BA, Beerhues L, Liu B. Biphenyls and dibenzofurans of the rosaceous subtribe Malinae and their role as phytoalexins. *Planta*. Sep 9 2023;258(4):78. doi:10.1007/s00425-023-04228-7

42. Khalil MNA, Brandt W, Beuerle T, et al. O-Methyltransferases involved in biphenyl and dibenzofuran biosynthesis. *The Plant Journal*. 2015/07/01 2015;83(2):263-276. doi:https://doi.org/10.1111/tpj.12885

43. Ishigaki E, Asamizu T, Arisawa M, Kurosaki F. Cloning and Expression of Calmodulin Genes Regulating Phytoalexin Production in Carrot Cells. *Biological and Pharmaceutical Bulletin*. 2004;27(8):1308-1311. doi:10.1248/bpb.27.1308

44. Busnena BA, Beerhues L, Liu B. Biphenyl and Dibenzofuran Phytoalexins Differentially Inhibit Root-Associated Microbiota in Apple, Including Fungal and Oomycetal Replant Disease Pathogens. *Phytopathology®*. 2025/02/01 2024;115(2):181-191. doi:10.1094/PHYTO-03-24-0088-R

45. Zhu Y, Li G, Singh J, et al. Laccase Directed Lignification Is One of the Major Processes Associated With the Defense Response Against Pythium ultimum Infection in Apple Roots. Original Research. *Frontiers in Plant Science*. 2021-September-07 2021;Volume 12 - 2021doi:10.3389/fpls.2021.629776

46. Gennaro Fazio MM, Yanmin Zhu. Genetic Analysis of Resistance to Pythium Ultimum a Major Component of Replant Disease in Apple Rootstocks. *Journal of the American Pomological Society*. %01/%01 2023;77(1):28-38. doi:10.71318/apom.2023.77.1.28

47. Pandian BA, Sathishraj R, Djanaguiraman M, Prasad PVV, Jugulam M. Role of Cytochrome P450 Enzymes in Plant Stress Response. *Antioxidants*. 2020;9(5):454.

48. Nahrstedt A. Cyanogenic compounds as protecting agents for organisms. *Plant Systematics and Evolution*. 1985/03/01 1985;150(1):35-47. doi:10.1007/BF00985566

49. Wu HC, Poulton JE. Immunocytochemical Localization of Mandelonitrile Lyase in Mature Black Cherry (Prunus serotina Ehrh.) Seeds. *Plant Physiol*. Aug 1991;96(4):1329-37. doi:10.1104/pp.96.4.1329

50. La Camera S, Balagué C, Göbel C, et al. The Arabidopsis patatin-like protein 2 (PLP2) plays an essential role in cell death execution and differentially affects biosynthesis of oxylipins and resistance to pathogens. *Mol Plant Microbe Interact*. Apr 2009;22(4):469-81. doi:10.1094/mpmi-22-4-0469

51. Chassot C, Nawrath C, Métraux JP. Cuticular defects lead to full immunity to a major plant pathogen. *Plant J*. Mar 2007;49(6):972-80. doi:10.1111/j.1365-313X.2006.03017.x

52. Mukherjee AK, Carp M-J, Zuchman R, Ziv T, Horwitz BA, Gepstein S. Proteomics of the response of Arabidopsis thaliana to infection with Alternaria brassicicola. *Journal of Proteomics*. 2010/02/10/ 2010;73(4):709-720. doi:https://doi.org/10.1016/j.jprot.2009.10.005

53. Piatkowski D, Schneider K, Salamini F, Bartels D. Characterization of Five Abscisic Acid-Responsive cDNA Clones Isolated from the Desiccation-Tolerant Plant Craterostigma plantagineum and Their Relationship to Other Water-Stress Genes. *Plant Physiol*. Dec 1990;94(4):1682-8. doi:10.1104/pp.94.4.1682

54. Boudichevskaia A, Flachowsky H, Dunemann F. Identification and molecular analysis of candidate genes homologous to HcrVf genes for scab resistance in apple. *Plant Breeding*. 2009 2009;128(1):84-91. doi:10.1111/j.1439-0523.2008.01537.x
55. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2024. https://www.R-project.org/