Genome-wide association studies reveal a rapidly evolving candidate avirulence effector in the Cercospora leaf spot pathogen *Cercospora beticola*

Chen Chen, Harald Keunecke, Felix Bemm, Gabor Gyetvai, Enzo Neu, Friedrich J. Kopisch-Obuch, Bruce A. McDonald, Jessica Stapley.

Plant Pathology Group, Institute of Integrative Biology, ETH Zurich, 8092 Zürich, Switzerland
KWS SAAT SE & Co. KGaA, Grimsehlstr. 31, 37574 Einbeck, Germany
Corresponding author: Jessica Stapley, Jessica.stapley@usys.ethz.ch

Abstract

The major resistance gene *BvCR4* recently bred into sugar beets provides a high level of resistance to Cercospora leaf spot caused by the fungal pathogen *Cercospora beticola*. The occurrence of pathogen strains virulent to *BvCR4* was studied using field trials in Switzerland and Germany. Virulence of a subset of these strains was tested in a field trial conducted under elevated artificial disease pressure. We created a new *C. beticola* reference genome and mapped whole genome sequences of 256 field-collected isolates and combined this with the virulence phenotypes to conduct three separate GWAS to identify candidate avirulence genes. GWAS analyses identified a locus associated with avirulence containing a single candidate avirulence effector gene named *AvrCR4*. All virulent isolates either lacked *AvrCR4* or had non-synonymous mutations within the gene. *AvrCR4* was present in all 74 unique isolates obtained from non-*BvCR4* hybrids, whereas 33 of 89 unique isolates obtained from *BvCR4* hybrids carried this deletion. We also mapped genomic data from 190 publicly available isolates from the U.S to our new reference genome. The *AvrCR4* deletion was found in only one of 95 unique isolates from non-*BvCR4* hybrids in the U.S. *AvrCR4* presents a unique example of an avirulence effector in which virulent alleles have only recently emerged. Most likely these were selected out of the standing genetic variation after deployment of a new major resistance gene. Identification of *AvrCR4* will enable real-time screening of *C. beticola* populations for the emergence and spread of virulent isolates as well as long-term studies of effector evolution.

Keywords: avirulence effector, GWAS, pathogen evolution, R-gene durability

34 **Introduction**

35 Plant pathogens pose significant risks to crop production and food security around the
36 world (Savary et al., 2019). Disease resistance genes (R-genes) provide economical and
37 ecologically-friendly approaches to disease control and make important contributions to
38 production of many important crops. Plant breeders often select for R-genes with major
39 effects that form a critical part of the plant's immune system. Major R-genes typically
40 encode immune receptors, e.g. receptor-like kinases (RLKs) or nucleotide-binding leucine-
41 rich repeat proteins (NB-LRRs) that detect pathogen proteins and trigger an immune
42 response (Deng et al., 2020; Petit-Houdenot and Fudal, 2017). Thus, R-genes play a major
43 role in disease control. In sugar beet (*Beta vulgaris ssp. vulgaris*), R-genes have provided the
44 only effective control for beet cyst nematode (*Heterodera schachtii*) (Stevanato et al., 2015)
45 and rhizomania (Beet necrotic yellow vein virus) (Capistrano-Gossmann et al., 2017).
46
47 The effectiveness of R-genes can be reduced when virulent strains of the pathogen evolve.
48 This evolution of virulent strains can happen over various periods of time, with well-
49 documented cases including 2-7 years for *Leptosphaeria maculans* on oilseed rape (Rouxel
50 and Balesdent, 2017; Van De Wouw et al., 2010); 3 years for *Zymoseptoria tritici* on wheat
51 (Cowger et al., 2000), and 2-3 years for *Pyricularia oryzae* on rice (Kiyosawa, 1982).
52 However, even when the population contains virulent strains, major genes can still prove
53 useful, e.g. the *Rlm7* gene encoding resistance to *L. maculans* remains useful more than 15
54 years after commercial introduction despite the identification of virulent isolates (Alnajar et
55 al., 2022; Winter and Koopmann, 2016). Understanding the mechanisms governing the
56 emergence of virulence in fungal pathogens can be useful for developing more effective
57 strategies to mitigate the impact of pathogens in agro-ecosystems.
58
59 Plant-pathogen coevolution in natural ecosystems created the reservoir of R-genes that can
60 be extracted from wild host populations and integrated into domesticated crops via plant
61 breeding programs. These R-genes typically recognise pathogen molecules, often called
62 effectors and they function to increase susceptibility of the host to pathogen infection
63 (Kamoun, 2007). If the recognition of an effector by an R-protein leads to a major immune
64 response that severely limits pathogen infection, it is called an avirulence effector. The
65 genes encoding avirulence effectors typically encode small, secreted, cysteine-rich proteins

with limited homology to other characterized proteins (Rouxel and Balesdent, 2017; Sonah et al., 2016). Many avirulence effector genes have now been identified and cloned in several important plant pathosystems, including *L. maculans* on oilseed rape (Rouxel and Balesdent, 2017), *P. oryzae* on rice (Fernandez, 2023), and *Z. tritici* on wheat (Meile et al., 2018; Zhong et al., 2017).

Following deployment of a major R-gene, the corresponding avirulence effector gene is under strong directional selection and typically evolve rapidly to evade host recognition, creating virulent pathogen strains that can infect plants carrying the major R-gene. This evolution can involve gene deletions, non-synonymous mutations, and alterations in gene transcription that enable evasion of the R-gene recognition. The genes encoding fungal effectors are often located in dynamic, TE-rich regions of the genome that are thought to facilitate an accelerated evolution for these genes (Lo Presti et al., 2015). After years of selection imposed by a widely deployed R-gene, the gene encoding the corresponding avirulence effector typically displays very high levels of genetic diversity, while the original wild-type sequence of the effector gene may have disappeared entirely or be present at only very low frequencies (Brunner et al., 2007; Brunner and McDonald, 2018).

Sugar beet is an important crop that provides around one-fifth of the global sucrose production every year(Rangel et al., 2020). Cercospora leaf spot (CLS), caused by the fungus *Cercospora beticola*, is the most common and destructive disease of sugar beet. Sugar beet cultivars with improved resistance to CLS have been under development for thirty years (Rangel et al., 2020) . Wild sea beet (*Beta vulgaris* subsp. *maritima*) is the most important source of resistant genes and at least four quantitative trait loci (QTL) were identified for resistance to CLS . The CLS-resistant cultivars generally produce lower yields compared with susceptible cultivars, and it has proved difficult to combine high resistance and high yield into the same cultivar (Setiawan et al., 2000; Taguchi et al., 2011). To date, the most widely planted sugar beet cultivars combine high yield and moderate CLS resistance (Windels et al., 1998), relying mainly on fungicides to limit the losses associated with CLS. Recently, Törjék et al (2022; 2020) identified an NBS-LRR resistance gene named *BvCR4* that provides new opportunities in CLS resistance breeding. This gene showed a high level of CLS resistance when bred into sugar beets, enabling a unique combination of high CLS resistance and high

98 yield in the same variety. However, *BvCR4* might be vulnerable to the evolution of virulent
99 *C. beticola* strains if it were deployed as a monogenic resistance source without any
100 supporting quantitative background resistance (Pilet-Nayel et al., 2017) and without
101 supplemental integrated plant protection measures.
102
103 The evolutionary potential of *C. beticola* is high because it has a mixed reproductive system,
104 exhibits extensive gene flow, and undergoes several cycles of reproduction during each
105 sugar beet growing season. Field populations contain high levels of genetic diversity as
106 expected for pathogen populations that maintain a high effective population size (Vaghefi et
107 al., 2017). *C. beticola* has already shown the capacity to rapidly evolve resistance to several
108 fungicides representing different modes of action, including triazole, benzimidazole,
109 difenoconazole, organotin, sterol demethylation inhibitor (DMI), and quinone outside
110 inhibitor (QoI) (Bolton et al., 2012; Karaoglanidis et al., 2000). This evolution of fungicide
111 resistance underlines the need to make more effective use of CLS resistance, but also
112 illustrates the potential for pathogen adaptation following the deployment of major R-
113 genes.
114
115 We conducted a series of field, greenhouse, and lab experiments to assess the risks
116 associated with deployment of *BvCR4* in sugar beets. We monitored CLS epidemics in
117 experimental hybrids carrying *BvCR4* and isolated putative virulent strains of *C. beticola*
118 from infected leaves. We confirmed the virulence of many of these strains on *BvCR4* hybrids
119 growing under field conditions. We created a new reference genome sequence for *C.*
120 *beticola*, analysed genome sequences for 256 isolates and conducted three GWAS analyses
121 to search for candidate avirulence effectors. Each GWAS identified the same candidate
122 avirulence effector (*AvrCR4*) on chromosome 1. The *AvrCR4* gene contained several
123 different mutations encoding virulence and was deleted in many isolates coming from
124 *BvCR4* hybrids. We discuss strategies that should be considered to extend the life
125 expectancy of *BvCR4* in sugar beet agriculture.
126
127 **Methods**
128 **Collections of leaves infected by CLS**

129 In 2019 a field in Switzerland (47°45'15.4"N, 9°03'70.6"E) was planted with two
130 experimental sugar beet hybrids; Hybrid_R1 carrying *BvCR4* in a genetic background with
131 low quantitative resistance, and Hybrid_S1 conferring only a low level of quantitative
132 resistance. Both hybrids were examined in 2019 for evidence of CLS infections. Infected
133 leaves were collected from clusters of plants exhibiting CLS symptoms and stored at ETH
134 until isolations were made as described below.
135
136 In 2020, four sugar beet experimental hybrids with different levels of resistance to CLS were
137 planted at two field sites in Switzerland separated by 60 km (Rudolfingen, Trüllikon
138 47°38'18.6"N, 8°40'46.1"E and Hendschiken, Othmarsingen 47°23'40.4"N, 8°13'41.1"E);
139 Hybrid_R2 with *BvCr4* embedded in a background with low quantitative resistance,
140 Hybrid_R3 with *BvCr4* embedded in a background with moderate quantitative resistance,
141 Hybrid_S2 with only low quantitative resistance and Hybrid_S3 with moderate quantitative
142 resistance. In Hendschiken, 12 rows of each hybrid were planted in adjacent plots and in
143 Rudolfingen, 18 rows of each hybrid were planted in adjacent plots (Figure S1). Sugar beets
144 were sown in March and harvested in early November. A similar experimental plot was
145 planted at two field sites in Germany. The same four experimental hybrids were planted in
146 fields located near Tabertshausen (48°43'22.42"N, 12°52'25.25"E) and Friesendorf
147 (48°48'1.02"N, 12°46'52.48"E). No fungicides were applied at any of the field sites.
148
149 We used two sampling strategies, one temporal (Swiss sites only) and the other spatial
150 (Swiss and German sites), to obtain leaves infected by *C. beticola* (Figure S1). The temporal
151 sampling strategy was designed to identify the infected plants in each plot and determine
152 how a CLS infection spreads to neighboring plants to form a cluster of CLS-infected plants.
153 We identified every newly infected plant in each plot at weekly intervals from July 21st to
154 September 24th 2020. We stopped sampling in a plot when it was no longer possible to find
155 new solitary infected plants because the clusters of infected plants began to overlap. Every
156 infected plant was given an alphanumeric code and its position was marked with a labeled
157 flag to avoid resampling. One young leaf with obvious CLS lesions was taken from each
158 infected plant. The excised leaves were transported back to the laboratory for *C. beticola*
159 isolation (details below). In the early stages of the epidemic we often observed single newly
160 infected plants surrounded by healthy plants; we refer to this pattern as a solitary infection.

161 We observed that the disease often spread from a solitary infected plant to neighboring
162 plants to form a cluster of infected plants, referred to hereafter as a hotspot. While
163 sampling, we determined whether a newly infected plant was a solitary infection or a part
164 of a hotspot. A GPS location was recorded for the first infected plant in each hotspot. A
165 maximum of five plants was sampled from each hotspot.
166
167 Transect sampling was conducted in all plots in Switzerland and Germany at the end of the
168 2020 season (13th and 15th October in Switzerland and on September 17th in Germany). Two
169 transect sites were selected within each hybrid field plot. Each transect area was at least
170 two rows away from another hybrid and at least 3 m from the plot edge. An individual
171 transect ran perpendicular to the direction of the planting rows, and a diseased leaf was
172 collected every third row, for a total of 3 leaves per transect line. A total of 7 transect lines
173 each 3 m apart were made for each transect site, giving a total of 21 leaves per transect site
174 (See Figure S1). Two transect sites were sampled for each hybrid, yielding 42 leaf samples
175 total for each hybrid in each plot at both sites in Switzerland and Germany.
176
177 In the field, the sampled leaves were stored flat and transported in an ice box back to the
178 lab for isolation. Each flattened leaf was placed on a labeled piece of paper with a dry paper
179 towel placed on top. Multiple leaves were stacked together and then pressed between rigid
180 cardboard sheets using strong rubber bands before being placed in an ice box filled with ice.
181 After transport to the lab, the leaves were air dried at room temperature in a drying
182 cupboard. Dried leaves were then stacked and placed into sealed plastic bags with
183 anhydrous silica gel and stored at 4°C until pathogen isolation.
184
185 **Pathogen isolation and DNA extraction**
186 Leaf pieces bearing CLS lesions were cut from each leaf and surface-sterilized by placing
187 them in a 2% sodium hypochlorite solution for 3 min and then rinsing them twice in sterile
188 water. After removing excess water, the leaf pieces were placed in glass petri dishes on a
189 sterile plastic screen laid on top of rubber bands on a piece of moist filter paper. The petri
190 dishes were sealed with Parafilm and incubated at 24°C for 48-72 hours to induce
191 sporulation. Lesions were checked under a dissecting microscope (Olympus Schweiz AG) and
192 spores from single sporulating lesions were gathered using hand-made fine glass needles

193 and transferred to fresh potato dextrose agar (PDA, Difco). The plates were cultured under
194 constant darkness at 24°C for 7-10 days. When a characteristic *C. beticola* colony had
195 formed, a 5 mm agar plug was excised from the edge of each colony and transferred to a
196 new PDA plate for 14 days to form a pure culture. Any plates contaminated by other fungi
197 were discarded. Long term storage of each isolate was made by cutting 4-6 agar plugs from
198 the edges of 2-week-old pure cultures and putting the plugs into cryotubes in duplicate -
199 with and without anhydrous silica gel, and then stored at 4°C.
200
201 To obtain fresh mycelia for DNA extraction, we transferred hyphae from pure culture plates
202 to liquid culture, by scraping small pieces of hyphae from the edge of each pure culture and
203 placing it into 100 mL flasks with 50 mL potato dextrose broth (PDB, Difco). Two methods
204 were used to extract DNA. A Kingfisher Flex robot (Thermo Fisher Scientific, U.S.A.) was
205 used for 103 samples following the recommended protocol for automated DNA extraction
206 from fungal samples (Fungus Sbeadex, LGC Genomics GmbH). For the remaining 153
207 isolates, DNeasy Plant Mini Kits (Qiagen) were used with a modified protocol where 3.5 µL
208 of proteinase K was added to each sample at the lysis step to enhance protein degradation.
209 DNA was quantified using the Qubit 2.0 Fluorometer (Themo Fisher Scientific, U.S.A.) with a
210 Qubit dsDNA HS Assay kit (Thermo Fisher Scientific, U.S.A.) and Nanodrop (ThermoFisher
211 Scientific, U.S.A.). Species identity was confirmed using PCR (Knight and Pethybridge, 2020).
212
213 For the reference genome, *C. beticola* strain "Hevensen" was grown in liquid media by
214 transferring mycelium from a V8 agar petri dish (500 ml of V8 juice, 500 ml distilled water, 3
215 g of $CaCO_3$ and 15 g of agar) to 225 ml of CB medium (1 g/l casein, 1 g/l yeast extract, 10 g/l
216 D-Glucose, 10 mM NaCl, 0.02 mM $FeCl_3x6H_2O$, 0.1 mM $ZnSO_4x7H_2O$, 7.9 µM $CuSO_4x5H_2O$,
217 0.4 mM $MnSO_4xH_2O$, 0.9 mM H_3BO_3, 4.2 mM $Ca(NO_3)_2x4H_2O$, 1.5 mM KH_2PO_4, 1 mM
218 $MgSO_4x7H_2O$). The culture was incubated for 6 to 8 days at 24°C at 100 rpm. Mycelium was
219 harvested in a sieve, dried on paper towels and ground to a fine powder in liquid nitrogen.
220 DNA was extracted using the method described by (Huang et al., 2018) with minor
221 modifications. DNA was repaired and end-prepped using NEBNext FFPE DNA repair mix
222 (NEB) and the NEBNext Ultra II End Repair/dA-Tailing Module (NEB), respectively. Small DNA
223 fragments were removed from each sample using a Short Read Eliminator reagent
224 (Circulomics).

225

226 **Genome data acquisition and assembly**

227 For the Hevensen reference isolate, Oxford Nanopore Technology (ONT) sequencing
228 libraries were prepared using the SQK-LSK109 library kit (ONT). Each library was loaded onto
229 three FLO-FLG001 (R9.4.1 pore) flongle cells (ONT) and run for 24 hours on MinION or
230 GridIONx5 instruments (ONT). Raw data was base called using Guppy (version 4.0.15,
231 default parameters) and assembled with miniasm (version r179; default parameters). The
232 resulting draft assembly was polished with Racon (version 1.4.13, input long reads, default
233 parameters), Medaka (version 1.2.0, input long reads, default parameters) and Pilon
234 (version 1.23, input short reads, default parameters). Polished sequences were size-ordered
235 and named in consecutive order as chromosomes 1 through 11. Two short contigs were
236 annotated via Uniprot Blast and contained rDNA (CB_CbHevensen-1_hq-v2.rDNA) and
237 mtDNA (CB_CbHevensen-1_hq-v2.mitochondrial). Final sequences were structurally
238 annotated with Funannotate (version 1.8.3, RNA-seq evidence – NCBI PRJNA294383, protein
239 evidence - NCBI RefSeq Fungi 2021-05-06, otherwise default fungal parameters) and
240 CodingQuarry-PM (run_CQ-PM_mine.sh, version 2.0).

241

242 DNA from the 256 field-collected isolates was sequenced using Illumina paired end (150 bp)
243 sequencing (sequence provider: Novogene). In addition, raw reads from 199 isolates were
244 downloaded from the SRA (https://www.ncbi.nlm.nih.gov/sra) using SRAtoolkit (v2.0) for
245 downstream analyses. We performed multiple analyses to check the quality of the isolate
246 sequence data (see below for details) and removed a total of 33 isolates (24 sequenced by
247 us, 9 from SRA). After mapping, quality checking and removal of sequence duplicates, we
248 retained 422 isolate genomes (232 sequenced by us, 190 from SRA) for further analyses. See
249 Table S1 for information about each isolate, including why they were removed from further
250 analyses as described below, noting that samples could be removed for multiple reasons.

251

252 We used k-mer based approaches to identify potential contaminants or mixed cultures. To
253 verify that isolations made in the lab were composed of a single clone, we first counted k-
254 mers with KMC (https://github.com/refresh-bio/KMC) (Kokot et al., 2017) and then
255 estimated heterozygosity and sequence coverage using Smudgeplot
256 (https://github.com/KamilSJaron/smudgeplot). Four isolates had high heterozygosity (>0.1)

257 and five had a k-mer based genome coverage lower than the expected coverage based on
258 the number of reads (number of reads * read length)/haploid genome size < 0.5). Next, we
259 investigated genetic similarity to identify isolates that were genetically distant, representing
260 potentially different species, subspecies or contaminated cultures. We used the program
261 mashtree (Katz et al., 2019) to draw a k-mer-based neighbor joining tree. Thirteen isolates
262 that were highly divergent were excluded from further analyses.
263
264 Sequence data processing and variant discovery are described in detail at
265 https://github.com/jessstapley/CercosporaGWAS. In brief, reads were trimmed of adapters
266 and low-quality sequence using trimmomatic (v0.35). The reads were mapped to the
267 Hevensen reference genome using bwa-mem (v0.7.17). A total of 13 samples did not map
268 well to the reference genome (< 80% properly paired reads).
269
270 Variant calling was done using the GATK (v4.2.4.1) Germline Short Variant Discovery
271 pipeline, following their Best Practices recommendations
272 (https://gatk.broadinstitute.org/hc/en-us/sections/360007226651-Best-Practices-
273 Workflows). GATK default filters are: MQ=20.0, ReadPosRankSum_lower=-2.0,
274 ReadPosRankSum_upper=2.0, MQRankSum_lower=-2.0, MQRankSum_upper=2.0,
275 BaseQRankSum_lower=-2.0, BaseQRankSum_upper=2.0. In addition to the GATK default
276 filters we excluded rDNA and mtDNA. We also removed indels and loci with >50% missing
277 genotypes, as well as loci that did not meet the following criteria: depth <20000, mapping
278 quality (MQ) >50, quality depth (QD) >2.0, Symmetric Odds Ratio (SOR) <3, and Phred-
279 scaled p-value using Fisher's exact test to detect strand bias (FS) <60. For the GWAS analyses
280 we used only biallelic SNPS and we removed markers with minor allele frequency <0.05.
281 Principal component analysis was performed on LD-pruned SNP datasets (see
282 https://github.com/jessstapley/CercosporaGWAS for details). The SNP error rate based on
283 isolates sequenced twice (sequence duplicates, n=29) was 0.00125.
284
285 Clones were identified using an IBS distance matrix from LD pruned data (13,696 SNPS). To
286 determine the appropriate genetic similarity threshold, we used SNP data from the known
287 sequence replicates. We calculated the genetic similarity of SNP data known to come from
288 the same genotype and found it ranged between 0.9968-0.9990. We considered three

289 thresholds (0.990, 0.995, 0.999) to identify clones. A threshold of 0.999 was too strict, with
290 only one pair of known sequence replicates correctly recorded as having the same
291 genotype. The thresholds of 0.990 and 0.995 gave identical results and correctly assigned all
292 sequence replicates as the same genotype. Thus, isolates with a genetic similarity >0.990
293 were considered a single genotype or clone. Following clone correction, we retained 163
294 unique isolates from Switzerland and Germany, and 95 unique isolates from the publicly
295 available genomes of isolates coming from the USA.
296
297 **Field phenotyping of virulence for strains used in the GWAS**
298 Fifty-three isolates phenotyped for virulence were unique genotypes originating from CLS
299 hotspots found in Swiss plots planted to *BvCR4* hybrids (20 isolates from Hybrid_R2, 33
300 isolates from Hybrid_R3). Three isolates were from the same field but collected on the
301 susceptible hybrids (2 from Hybrid_S3, 1 from Hybrid_S4). One putatively virulent clone that
302 was widespread in one of the experimental fields was replicated five times, hence only 52
303 unique genotypes were included in the field phenotyping for virulence. Four additional
304 reference inocula were used as controls with known virulence properties, but were not
305 sequenced nor included in the GWAS analyses.
306
307 Inoculum of the *C. beticola* strains was multiplied at KWS in Einbeck, Germany between
308 March and June 2022. The strains were cultured on Petri plates containing V8 agar. A 5 mm
309 plug of mycelium for each isolate was transferred onto a petri plate and incubated at 20°C
310 under 12-hour darkness:12 hour fluorescent light until the plates were at least 50%
311 colonized, requiring approx. two weeks. Mycelium plugs from these starter plates were then
312 transferred onto fresh V8 plates (3 plates per isolate) and incubated as described above for
313 another 2 weeks. Fully colonized plates were blended in 50 ml of distilled water to make a
314 suspension. Approximately 5 ml of this suspension was poured into 25 new V8 agar plates,
315 which were then rotated to evenly spread the liquid on the agar surface. The plates were
316 then placed into a cardboard box and incubated in the dark at 20°C for 2 weeks. Inoculum
317 consisting of mycelial fragments and spores was harvested by adding 5 ml of distilled water
318 per plate followed by gently scraping off the fungal cells growing on the agar surface. The
319 resulting fungal mass, containing both hyphal fragments and spores, was filtered through
320 cheesecloth and then adjusted to a concentration of 25,000 hyphal fragments/spores per

321 ml. The suspension from each isolate was amended with 0.1% Tween 20 (Sigma-Aldrich)
322 prior to inoculation.
323
324 Hybrid_R2, an experimental hybrid shown in earlier experiments to be susceptible to *C.*
325 *beticola* under greenhouse conditions, was planted in trays filled with peat soil mix
326 (HAWITA). After 2 weeks, seedlings were transplanted into new trays and kept at 16°C for 6
327 weeks under diffuse light. Plants were then transferred into plastic pots (11x11x10 cm) filled
328 with peat soil mix and arranged in groups of 30 plants in a greenhouse chamber set at 20°C
329 with 16h fluorescent light per day for four weeks. Plants were fertilized weekly with 0.8%
330 Kamasol green (10N-4P-7K). The distance between each group of 30 plants was 0.5 m to
331 minimize the potential for cross-contamination of the inoculum produced for each strain.
332
333 Each group of thirty 12-week-old plants was inoculated with one of the 56 isolates used in
334 the GWAS. Suspensions of spores and hyphal fragments were sprayed onto each group of
335 plants until runoff with a hand-held atomizer. Inoculated plants were then covered with
336 plastic sheets draped over metal frames and incubated for 5 days without supplemental
337 lighting at 25°C and kept at 100% RH to facilitate infection. After the incubation, plants were
338 kept under a 16 h photoperiod at 25°C/20°C (day/night) with RH maintained between 60-
339 80%. The first symptoms were observed at around 14 dpi for most isolates. Inoculum for
340 each isolate was harvested between 28 and 42 dpi by removing leaves that were fully
341 covered with CLS lesions. The removed leaves were placed in paper envelopes and dried at
342 35°C for 3 days. The dried leaves were then milled in a kitchen blender for 30 sec,
343 thoroughly mixed with semolina flour (40 g of milled leaf inoculum added to 200 g of flour)
344 and then transferred into labelled plastic boxes.
345
346 Virulence assessments were conducted in 2022 on a field site near Northeim, Germany
347 (51°40'52.1"N, 9°57'09.2"E). To minimize the risk of infection by local strains of *C. beticola*,
348 the field had not been planted to sugar beet for about 20 years (corn-wheat rotation), the
349 nearest sugar beet field was 800 m away, and the experimental field site was surrounded by
350 a 50 m border of maize. Seeds treated with the fungicide Tachigaren 70WP (Hymexazole)
351 and the insecticide Force 20CS (Tefluthrine) from four sugar beet hybrids with varying levels
352 of CLS resistance (Table 1) were planted on April 19th, 2022 in a randomized block design

353 with 6-row plots that were 0.45 m wide and 5 m long. Each isolate was inoculated onto a
354 block of plots that contained all 4 hybrids. Weeds were controlled with three herbicide
355 applications and aphids were controlled with two insecticide applications during the
356 growing season.
357
358 Plots were inoculated after row closure on July 27th, 2022. A total of 56 isolates and 4
359 control inoculants were inoculated separately onto blocks of plots containing all four
360 hybrids. Inoculation was performed by manually spreading 60 g of leaf inoculum mix (10 g
361 milled leaf inoculum + 50 g semolina flour) onto one row (3rd or 4th row, 5 m length, approx.
362 25 plants) of each 6-row plot. Inoculation rows were sprayed prior to inoculation with water
363 containing an adjuvant (Silwet 0.05%) so that most of the inoculum mix would stick to the
364 leaf canopy. Following the inoculation, all plots were watered twice a week for 6 weeks to
365 facilitate infection and disease progress. Approximately 4 l water/m^2 was applied each week
366 with a tractor-mounted crop protection sprayer. Disease severity data were collected from
367 the inoculated rows across five time points in August and September at 32, 40, 47, 56 and
368 64 days post inoculation (dpi), using the KWS rating scale (Table S3). We investigated how
369 disease rating changed at different time points using linear models and we used least
370 significant differences to make post hoc comparisons (Bonferroni adjusted) between mean
371 rating at each time point. Disease rating was bimodal at 40 dpi on the resistant cultivar
372 Hybrid_R2 (Figure S2), providing a clear distinction between virulent (rating >4 on KWS
373 rating scale) and avirulent (rating ≤ 4) strains and enabling us to code isolates as either
374 virulent or avirulent for some of the GWAS analyses.
375
376 **GWAS analyses**
377 SNP-based GWAS were performed using GAPIT (http://zzlab.net/GAPIT). We performed two
378 GWAS analyses using different phenotype data. The first GWAS was based on the 52 unique
379 isolates that were included in the phenotyping field trial, using the quantitative virulence
380 rating on the resistant hybrid (Hybrid_R2) at 40 dpi. This analysis is hereafter referred to as
381 the 'quantitative GWAS'. The second GWAS was done by coding the phenotype data as
382 either virulent (Hybrid_R2 rating at 40 dpi > 4) or avirulent (Hybrid_R2 rating at 40 dpi ≤ 4)
383 and included 38 additional isolates that were sampled from the susceptible hybrid
384 Hybrid_S2 and scored as avirulent. This created an expanded dataset with improved

385 statistical power and is referred to hereafter as 'binary GWAS'. We used a Mixed Linear
386 Model (MLM), which accounts for relatedness of individuals, and we included three
387 principal components into the model as covariates (random effects). The model fit was
388 assessed by inspecting the Q-Q plots (Figure S3). We repeated the analysis using
389 quantitative virulence ratings at other time points (Figure S4). A quantitative GWAS using
390 the Hybrid_R2 virulence rating at 47 dpi identified the same peak that was found using the
391 rating at 40 dpi.
392
393 We conducted a k-mer GWAS as these can provide more power to detect associations (see
394 https://github.com/jessstapley/CercosporaGWAS for details) as illustrated in a recent
395 analysis conducted in *Z. tritici* (Dutta et al., 2021). We used recommendations and scripts
396 from https://github.com/voichek/kmersGWAS, but used GAPIT to run the GWAS. In brief,
397 KMC was used to count k-mers with a length of 31 bp. The k-mer plink output file (.bed) was
398 imported into plink (v1.9, https://www.cog-genomics.org/plink/) and converted to a vcf file
399 (--recode vcf). The vcf file was split into 99 smaller files (containing 20,000 k-mers each),
400 then converted to a hapmap file using R::vcfR (https://github.com/knausb/vcfR). GAPIT
401 GWAS was then run on each of these smaller files. The k-mers were mapped to the
402 reference genome using bowtie2 (v2.3) and the results were concatenated into
403 chromosome files for plotting. The k-mer GWAS was performed using only 47 of the
404 phenotyped strains because we could not count k-mers in all 52 strains.
405
406 **Annotation of the GWAS interval**
407 Genes within the GWAS interval were identified using the reference genome annotation file
408 (gff3). The function of the putative effector was confirmed using EffectorP (v3.0)
409 (https://effectorp.csiro.au/ (Sperschneider and Dodds, 2022)). Additional annotation of the
410 GWAS interval was performed to identify possible transposable elements (TEs) using Repeat
411 Masker (v4.0.6) and One Code to Find them all (Bailly-Bechet et al., 2014). Signatures of RIP
412 in the GWAS interval were investigated using RIPPER (van Wyk et al., 2019).
413
414 SNP genotypes across the GWAS interval were obtained for every sequenced isolate using
415 bcftools (samtools v1.9). Individuals with missing genotype calls spanning the entire
416 candidate gene were recorded as 'gene deleted' for the candidate gene. Individuals with

417 SNP genotype calls within the gene were recorded as 'gene present'. We compared the
418 frequency of gene presence or absence between resistant and susceptible hybrids in
419 Germany and Switzerland using a binomial exact test.
420
421 **Results**
422 Swiss field plots planted to the resistant hybrids carrying *BvCR4* showed a later onset of CLS
423 symptoms and a slower epidemic progress compared to plots planted to the susceptible
424 sugar beet hybrids (unpublished data). The total levels of CLS infection in the resistant
425 hybrids never reached the levels found in the susceptible hybrids. Leaf death and
426 subsequent regrowth was evident across the whole plot in the susceptible hybrids, but was
427 less prevalent in the resistant hybrids carrying *BvCR4*. From field isolates we chose 52
428 unique isolates for phenotyping. In choosing strains for the field phenotyping, we aimed to
429 maximize genotypic diversity among a sample of putatively virulent isolates sampled mainly
430 from hotspots of infected plants in plots planted to resistant hybrids carrying *BvCR4*.
431 Interestingly, one of those putatively virulent genotypes was more frequent and widely
432 distributed at the Rudolfingen field site. We included 5 separate isolations of this genotype
433 in the field phenotyping conducted in 2022 in Northeim, Germany to test reproducibility of
434 the phenotypes.
435
436 In this validation trial, the disease ratings increased over the growing season, peaking at 56
437 dpi (Figure 1). The disease ratings at 32-56 dpi were highly correlated while the rating at 64
438 dpi was less correlated with the other time points (Figure S5). Using the 40-dpi rating on
439 Hybrid_R2, isolates were scored as virulent (>4) or avirulent (≤4), resulting in 25 of the 52
440 unique field-collected isolates scored as virulent and 27 scored as avirulent. The three
441 isolates collected from susceptible hybrids were all avirulent.
442
443 We considered how dpi, hybrid and virulence score (0/1) explained variation in the disease
444 rating and found significant pairwise interactions but no three-way interaction (all statistics
445 are provided in Table S4). Disease ratings increased over time on all hybrids, but different
446 hybrids showing different patterns of disease development (Figure 1). When isolates were
447 separated into two categories (virulent or avirulent) based on a threshold rating of 4 on
448 hybrid_R2 at 40 dpi, isolates scored as virulent caused more disease at all timepoints on

449 Hybrid_R3 carrying *BvCR4* compared to the avirulent isolates (it was not statistically
450 appropriate to do this analysis on Hybrid_R2 because this was used to define
451 virulent/avirulent). When tested on hybrids without *BvCR4* (hybrid_S3 and hybrid_S4),
452 virulent and avirulent isolates caused similar levels of disease.
453
454 **Genome-wide association scans (GWAS) identified a single effector gene candidate**
455 A high-quality reference genome was created using Nanopore long read sequencing. The
456 genome was assembled into 11 continuous sequences (hereafter called chromosomes). The
457 genome showed a completeness of 98.4% according to the BUSCO benchmark against
458 Ascomycota single copy orthologs. Structural annotation resulted in 13,869 genes being
459 present.
460
461 The quantitative GWAS included 52 phenotyped individuals genotyped at 358,365 SNPs. The
462 resulting Manhattan plots showed a single distinctive, but not statistically significant, LOD
463 peak on chromosome 1 (Figure 2a). The binary GWAS was performed on 90 individuals
464 genotyped at 358,365 SNPS. A significant LOD peak was observed on chromosome 1, with
465 the significant interval spanning 6730 bp (4087634-4094364 bp) (Figure 2b). This peak was
466 in the same position as the peak observed in the quantitative GWAS analysis. The GWAS
467 interval contained a single gene (ID=Hevensen_hq-v2.G01679) that is 340 bp long (Figure 3).
468 The gene is predicted to encode an 89 amino acid (aa) apoplastic/cytoplasmic effector
469 according to EffectorP (v3.0) that includes a 21 aa signal peptide (predicted by TOPCONS,
470 Tsirigos et al., 2015). We did not find a significant match using a BLAST search against the
471 publicly available reference genome, but a partial hit was found on Chr 7. A copy of the gene
472 with 88% identity was found in the *Cercospora sojina* (strain RACE15) chromosome XII:
473 4346561-4346898.
474
475 The k-mer GWAS was performed using 47 of the phenotyped individuals. Several significant
476 peaks were identified. A peak on Chromosome 1 was found, with the significant k-mers
477 located within the SNP-based GWAS interval identified in the other two GWAS analyses.
478 Other significant associations containing more than one k-mer were found on Chromosomes
479 2 and 7. In both cases the k-mers with significant associations were not tightly clustered,
480 unlike the peak on Chr 1. So, we consider these to be most likely false positives. Significant

481 singleton k-mers found on Chromosomes 1 (1644617bp) and 8 (2223033bp) were also
482 considered to be false positives.
483
484 The effector gene candidate was located in a gene-sparse, repeat-rich region of the genome
485 that exhibits signatures of Repeat Induced Point mutations (RIP) (Figure 3). Among the 258
486 unique sequenced genotypes, 34 had the candidate effector deletion. When considering
487 isolates collected from Switzerland and Germany, the gene was absent in over 40% of the
488 isolates collected from the resistant hybrids carrying *BvCR4* (mean = 43.5%) and was present
489 in all of the isolates collected from the susceptible hybrids (Table 2). In the publicly available
490 genome sequences from the USA, the gene deletion was observed only once out of 95
491 unique genotypes.
492
493 **Sequence variation in the candidate effector gene**
494 Among the 224 isolates with the candidate effector gene present, 195 (87%) had no
495 variation in the gene (158 from susceptible hybrids, 37 from resistant hybrids). Among the
496 remaining 29 isolates, we identified 24 SNPs within the gene. No SNPs were found within
497 the intron. Haplotype network analysis revealed higher diversity in the isolates collected on
498 the resistant hybrids, compared to the susceptible hybrids (Figure 4a). The sequence
499 diversity (pi) was 10X higher in the isolates originating from the resistant hybrids (pi=0.0037)
500 compared to isolates coming from susceptible hybrids (pi = 0.00030) in Switzerland.
501 Translation of the two exons identified 12 different amino acid sequences, with most of the
502 variation in the amino acid sequence resulting in virulence at 40 dpi on Hybrid_R2 (Figure
503 4b).
504
505 To assess the potential for a fitness cost associated with loss of function of the putative
506 effector, we tested if the loss or inactivation of the effector influenced the virulence of an
507 isolate infecting resistant and susceptible hybrids. We grouped isolates into two categories:
508 those encoding the reference wild type isoform of the putative effector, and those encoding
509 either deletions or alternative effector isoforms. We found no difference in disease rating
510 between isolates with the deleted isoform or with alternative isoforms (Table S5). Across
511 phenotyped isolates, we found that dpi, hybrid and effector gene isoform explained
512 variation in disease rating, with significant interactions between all pairs of factors (dpi and

513 hybrid, hybrid and isoform, and dpi and isoform) but no three-way interaction (Table S6).
514 Considering each hybrid and dpi separately, we found that the effector isoform had little
515 effect on an isolate's virulence rating when infecting the susceptible cultivars, but had a
516 significant effect on an isolate's virulence when infecting the resistant cultivars carrying
517 *BvCR4* (Figure 5). Thus, there was no evidence for a reduction in fitness in strains lacking
518 *AvrCR4* or encoding alternative isoforms of AvrCR4. Rather it appears that strains carrying
519 deletions or alternative isoforms exhibit higher virulence on the susceptible cultivars (Figure
520 5, p-values ranging from 0.02-0.09, See Table S6).
521
522 **Discussion**
523 The identification and use of *BvCR4* represents an important advance in breeding for
524 resistance to CLS, as it provides a high level of resistance to CLS without a corresponding
525 reduction in sugar yield. Given the long history of boom-and-bust cycles associated with the
526 release of major-effect R-genes, there was concern that *C. beticola* would be able to evolve
527 virulence against resistant hybrids carrying *BvCR4*, the first reported major R-gene deployed
528 against this pathogen.
529
530 The first evidence that virulent isolates could evolve was manifested by the emergence of
531 hotspots of CLS disease in experimental fields planted with experimental *BvCR4* hybrids in
532 Switzerland. These hotspots did not emerge until late in the growing season, but their
533 occurrence suggested that hotspots could reflect the selection and local spread of *C.*
534 *beticola* clones that were virulent to *BvCR4*. Isolations of *C. beticola* made from these
535 hotspots in Switzerland in 2019 provided the first strains that were shown in 2021 to be
536 virulent on sugar beet hybrids carrying *BvCR4*. The discovery of these infection hotspots
537 triggered field experiments designed to assess the emergence and spread of virulent strains,
538 including a search for evidence of fitness costs associated with a gain in virulence.
539
540 We used a GWAS approach to identify genes associated with higher virulence on *BvCR4*
541 hybrids, with collections of *C. beticola* made from plots planted with resistant *BvCR4* hybrids
542 and susceptible hybrids lacking *BvCR4* growing in the same field. *C. beticola* isolates coming
543 from hotspots identified in plots planted with *BvCR4* hybrids were phenotyped under field
544 conditions to test their degree of virulence. Only 51% of these isolates were confirmed to be

545 virulent on *BvCR4*, illustrating that *BvCR4* hybrids can be colonized by *C. beticola* strains that
546 are not specifically adapted to *BvCR4*. Though we did not yet conduct experiments to
547 determine how the avirulent strains were able to infect plants carrying *BvCR4*, we speculate
548 that abiotic or biotic stresses other than CLS could play a role. We observed several
549 instances of plants infected by *Rhizoctonia solani* in Rudolfingen in 2020 that might have
550 compromised CLS resistance. Furthermore, experiments with mixed infections in *Z. tritici*
551 showed that virulent strains were able to suppress the immune system of plants carrying
552 the *Stb6* R-gene, enabling avirulent strains to infect the resistant plants (Bernasconi et al.,
553 2023). We hypothesize that similar mechanisms could be operating in the CLS pathosystem.
554
555 A quantitative GWAS that included only the 52 isolates with field-based phenotypes
556 identified a region on Chromosome 1 associated with virulence, but the LOD scores did not
557 exceed the significance threshold. To increase the power to detect genes associated with
558 virulence on *BvCR4*, we assumed that any allelic variants encoding virulence to *BvCR4* would
559 be present at a very low frequency (e.g., at mutation-drift equilibrium) prior to the
560 deployment of hybrids carrying *BvCR4*. Hence, we scored 38 *C. beticola* isolates sampled
561 from susceptible hybrids growing in the same field as avirulent, even though we lacked
562 phenotypes for most of these strains. None of the phenotyped isolates sampled from
563 susceptible hybrids showed the virulent phenotype. This binary phenotype approach
564 identified the same Chromosome 1 region, but with a significant LOD score of 9.3. A third
565 GWAS based on k-mers but including only the phenotyped strains also identified the same
566 Chromosome 1 region, with a very high LOD score of 21.3, as well as small intervals on
567 Chromosomes 2 (LOD = 14.5) and 7 (LOD = 11.4).
568
569 A single predicted gene was found within the Chromosome 1 interval that was identified in
570 all three GWAS analyses. This gene shares many of the properties that are found in
571 avirulence effectors that have been functionally validated in other plant pathogenic fungi,
572 including: a) it encodes a small, secreted, cysteine-rich protein; b) it displays a presence-
573 absence polymorphism in natural populations, with the absence allele associated with a
574 gain in virulence occurring at a much higher frequency on *BvCR4* hybrids compared to
575 susceptible hybrids; c) its exons encode non-synonymous substitutions that can be
576 associated with a gain in virulence, but nucleotide diversity is lower in the introns and the

577 signal peptide sequence; d) it is located in a genomic region that is rich in transposable
578 elements and associated with RIP. Based on these properties, we named this candidate
579 avirulence effector gene *AvrCR4*.
580
581 The main difference between *AvrCR4* and other cloned fungal avirulence effectors is that
582 the overall level of DNA and amino acid diversity found in this study is lower than what has
583 been reported for other avirulence effectors (Brunner et al., 2007; Brunner and McDonald,
584 2018; Schürch et al., 2004). This lower diversity in *AvrCR4* may reflect the very short time
585 period over which selection has been operating on this gene. In Germany, the first
586 commercial cultivation of *BvCR4* sugar beet hybrids occurred in 2021 and in Switzerland in
587 2022. Our interpretation is that *AvrCR4* has not yet been placed under strong diversifying
588 selection that would favor mutations that enable it to avoid recognition encoded by the
589 newly deployed *BvCR4* R-gene. It can be expected that *AvrCR4* will begin evolving in all
590 regions planted to *BvCR4* sugar beets. The rate of evolution may be moderated if the *BvCR4*
591 resistance is embedded in genetic backgrounds providing at least moderate quantitative
592 resistance and if fields planted to *BvCR4* sugar beets are supported by other agronomic
593 measures to reduce selection pressure as discussed below. Combining *BvCR4* with an
594 adequate level of quantitative resistance is the current breeding approach used to mitigate
595 the risk that *C. beticola* will evolve to overcome the overall CLS resistance.
596
597 Our current data support the hypothesis that the *AvrCR4* mutations creating virulent
598 phenotypes were selected from already existing, standing genetic variation in *C. beticola*
599 populations. As recently described for the wheat pathogen *Zymoseptoria tritici* (McDonald
600 et al., 2022), populations of fungal plant pathogens can generate and maintain vast
601 reservoirs of functional genetic diversity that enable them to adapt quickly to changes in
602 their local environments, including applications of fungicides and deployments of resistant
603 cultivars. *C. beticola* shares many properties with *Z. tritici*, including a mixed reproduction
604 system and high gene flow (Rangel et al., 2020; Vaghefi et al., 2017), suggesting that its
605 populations will also maintain high levels of standing genetic variation.
606
607 We believe that *AvrCR4* is the first avirulence effector to be discovered at the onset of its
608 evolution that will be driven by deployment of a novel major R-gene. Hence *AvrCR4*

609 provides a unique opportunity to monitor in real time the changes in effector amino acid
610 sequence that enable an escape from recognition by the cognate NBS-LRR protein. We
611 expect that the most fit AvrCR4 protein isoforms will rapidly increase in frequency in each
612 region where the *BvCR4* R-gene is deployed. The diversity found in other cloned avirulence
613 genes is thought to have developed over decades of selection imposed by their
614 corresponding R-genes. Support for the hypothesis that *AvrCR4* has just begun to evolve is
615 as follows: a) the same wild-type *AvrCR4* sequence is found at a high frequency on both
616 resistant and susceptible sugar beets growing on different continents (Figure 4a); b) *AvrCR4*
617 is found in nearly all *C. beticola* isolates sampled from locations where *BvCR4* has not yet
618 been deployed, but is often missing in isolates sampled from hybrids carrying *BvCR4*; c)
619 there is little nucleotide diversity in *AvrCR4* in *C. beticola* isolates sampled from locations
620 where *BvCR4* has not yet been deployed; d) a relatively small number of non-synonymous
621 mutations are found in the *AvrCR4* alleles associated with virulence coming from isolates
622 sampled from *BvCR4* hybrids. As a reference comparison, we consider diversity for the
623 *AvrStb6* gene found in the wheat pathogen *Z. tritici*. The *Stb6* R-gene has been deployed in
624 European wheat cultivars for at least 100 years (Dutta et al., 2021). The *AvrStb6* gene found
625 in every *Z. tritici* isolate analyzed to date (no deletions have been detected) encodes a
626 protein containing 81 or 82 amino acids. In a global sample of 142 *Z. tritici* strains, there was
627 no predominant wild-type allele. Instead, 71 polymorphic sites produced 41 distinct DNA
628 haplotypes encoding 30 AvrStb6 protein isoforms, with significant positive diversifying
629 selection operating on nine of the 82 codons (Brunner and McDonald, 2018) and with
630 several polymorphisms found in both the introns and the signal peptide. In our clone-
631 corrected sample of 258 European and North American genomes of *C. beticola*, 34 isolates
632 were missing the gene (almost all originating from a resistant hybrid carrying *BvCR4*) and
633 only 29 strains carried *AvrCR4* alleles that differed from the predominating wild type
634 sequence found in the reference strain. Among those 29 strains, we identified 11 distinct
635 DNA haplotypes carrying 24 polymorphic sites encoding 13 AvrCR4 protein isoforms.
636 Monitoring changes in the sequence of *AvrCR4* over the next decade will allow us to
637 determine which mutations enable evasion of *BvCR4* resistance. We also consider this to be
638 a unique opportunity to search for compensatory mutations that could ameliorate any
639 reductions in pathogen fitness caused by changing the function of the encoded wild-type
640 avirulence effector protein.

641

642 Strains carrying deletions for *AvrCR4* showed a similar disease development on sugar beet
643 hybrids lacking *BvCR4* compared to strains carrying the wild type *AvrCR4* allele. This lack of
644 an obvious fitness cost associated with the total loss of function due to deletion or mutation
645 of the avirulence gene has been reported before (Zhan and McDonald, 2013). While our first
646 experiments did not detect any fitness costs in the field environment, we are currently
647 conducting additional lab-based experiments to search for fitness costs associated with loss
648 of the *AvrCR4* function.

649

650 **Outlook**

651 The identification of *AvrCR4* will enable us to closely monitor the frequency of virulent
652 mutants in CLS populations. Monitoring may prove useful to support a region-specific
653 variety choice and to apply regional management strategies aiming to reduce both the
654 overall CLS population and the frequency of virulent strains. These management strategies
655 will likely utilize an integrated plant protection approach that combines precise timing of
656 fungicide applications with agronomic practices such as crop rotation and tillage that reduce
657 an inoculum build up in the soil (Harveson et al., 2009; Khan et al., 2008; Khan and Khan,
658 2010; Secor et al., 2010) . Varieties that combine *BvCR4* with other resistance sources can
659 provide a strong CLS background resistance that will further contribute to both a reduction
660 of the CLS population and the durability of *BvCR4*. This latter approach was shown to be
661 effective in oilseed rape, where the adaptation time of *L. maculans* to the resistance gene
662 *Rlm6* was doubled by protecting this resistance with additional polygenic background
663 resistance (Brun et al., 2010; Delourme et al., 2014; Pilet-Nayel et al., 2017).

664

665

666

667

668 **References**

669 Alnajar, D., Von Tiedemann, A. & Koopmann, B. (2022) Efficacy of blackleg major resistance
670 genes in *B. napus* in Germany. *Pathogens*, 11, 461.
671 https://doi.org/10.3390/pathogens11040461.

672 Bailly-Bechet, M., Haudry, A. & Lerat, E. (2014) "One code to find them all": a perl tool to
673 conveniently parse RepeatMasker output files. *Mobile DNA*, 5, 13.
674 https://doi.org/10.1186/1759-8753-5-13.

675 Bernasconi, A., Lorrain,C, Flury, P, Alassimone, J, McDonald, B A, & Sánchez-Vallet, A (2023)
676 *Zymoseptoria tritici* suppresses the host immune response and facilitates the success
677 of avirulent strains in mixed infections.
678 http://biorxiv.org/lookup/doi/10.1101/2023.05.13.540507.

679 Bolton, M.D., Birla, K., Rivera-Varas, V., Rudolph, K.D. & Secor, G.A. (2012) Characterization
680 of *CbCyp51* from field isolates of *Cercospora beticola*. *Phytopathology*, 102, 298–305.
681 https://doi.org/10.1094/PHYTO-07-11-0212.

682 Brun, H., Chèvre, A., Fitt, B.D., Powers, S., Besnard, A., Ermel, M., et al. (2010) Quantitative
683 resistance increases the durability of qualitative resistance to *Leptosphaeria*
684 *maculans* in *Brassica napus*. *New Phytologist*, 185, 285–299.
685 https://doi.org/10.1111/j.1469-8137.2009.03049.x.

686 Brunner, P.C. & McDonald, B.A. (2018) Evolutionary analyses of the avirulence effector
687 AvrStb6 in global populations of *Zymoseptoria tritici* identify candidate amino acids
688 involved in recognition: Evolution of *AvrStb6*. *Molecular Plant Pathology*, 19, 1836–
689 1846. https://doi.org/10.1111/mpp.12662.

690 Brunner, P.C., Schürch, S. & McDonald, B.A. (2007) The origin and colonization history of the
691 barley scald pathogen *Rhynchosporium secalis*. *Journal of Evolutionary Biology*, 20,
692 1311–1321. https://doi.org/10.1111/j.1420-9101.2007.01347.x.

693 Capistrano-Gossmann, G.G., Ries, D., Holtgräwe, D., Minoche, A., Kraft, T., Frerichmann,
694 S.L.M., et al. (2017) Crop wild relative populations of *Beta vulgaris* allow direct
695 mapping of agronomically important genes. *Nature Communications*, 8, 15708.
696 https://doi.org/10.1038/ncomms15708.

697 Cowger, C., Hoffer, M.E. & Mundt, C.C. (2000) Specific adaptation by *Mycosphaerella*
698 *graminicola* to a resistant wheat cultivar. *Plant Pathology*, 49, 445–451.
699 https://doi.org/10.1046/j.1365-3059.2000.00472.x.

700 Delourme, R., Bousset, L., Ermel, M., Duffé, P., Besnard, A.L., Marquer, B., et al. (2014)
701 Quantitative resistance affects the speed of frequency increase but not the diversity
702 of the virulence alleles overcoming a major resistance gene to *Leptosphaeria*
703 *maculans* in oilseed rape. *Infection, Genetics and Evolution*, 27, 490–499.
704 https://doi.org/10.1016/j.meegid.2013.12.019.

705 Deng, Y., Ning, Y., Yang, D.-L., Zhai, K., Wang, G.-L. & He, Z. (2020) Molecular basis of disease
706 resistance and perspectives on breeding strategies for resistance improvement in
707 crops. *Molecular Plant*, 13, 1402–1419. https://doi.org/10.1016/j.molp.2020.09.018.

708 Dutta, A., Croll, D., McDonald, B.A. & Krattinger, S.G. (2021) Genome-wide association study
709 for septoria tritici blotch resistance reveals the occurrence and distribution of *Stb6* in
710 a historic Swiss landrace collection. *Euphytica*, 217, 108.
711 https://doi.org/10.1007/s10681-021-02843-7.

712 Fernandez, J. (2023) The phantom menace: latest findings on effector biology in the rice
713 blast fungus. *aBIOTECH*. https://doi.org/10.1007/s42994-023-00099-4.

714 Harveson, R.M., Hanson, L.E. & Hein, G.L. eds. (2009) Compendium of Beet Diseases and
715 Pests, Second Edition. The American Phytopathological Society.

716 Huang, X., Duan, N., Xu, H., Xie, T.N., Xue, Y.-R. & Liu, C.-H. (2018) CTAB-PEG DNA extraction
717 from fungi with high contents of polysaccharides. *Molecular Biology*, 52, 621–628.
718 https://doi.org/10.1134/S0026893318040088.

719 Kamoun, S. (2007) Groovy times: filamentous pathogen effectors revealed. *Current Opinion*
720 *in Plant Biology*, 10, 358–365. https://doi.org/10.1016/j.pbi.2007.04.017.

721 Karaoglanidis, G.S., Ioannidis, P.M. & Thanassoulopoulos, C.C. (2000) Reduced sensitivity of
722 *Cercospora beticola* isolates to sterol-demethylation-inhibiting fungicides. *Plant*
723 *Pathology*, 49, 567–572. https://doi.org/10.1046/j.1365-3059.2000.00488.x.

724 Katz, L., Griswold, T., Morrison, S., Caravas, J., Zhang, S., Bakker, H., et al. (2019) Mashtree: a
725 rapid comparison of whole genome sequence files. *Journal of Open Source Software*,
726 4, 1762. https://doi.org/10.21105/joss.01762.

727 Khan, J., Río, L.E. del, Nelson, R.S., Rivera-Varas, V.V., Secor, G.A. & Khan, M.F.R. (2008)
728 Survival, dispersal, and primary infection site for *Cercospora beticola* in sugar beet.
729 *Plant Disease*, 92 5, 741–745.

730 Khan, M.F.R. & Khan, J. (2010) Survival, spore trapping, dispersal, and primary infection site
731 for *Cercospora beticola* in sugar beet. In: Lartey, R.T., Weiland, L., and Panella, P.W.
732 (Eds.) Cercospora Leaf Spot of Sugar Beet and Related Species. St Paul, MN:
733 American Phytopathological Society, pp. 67–75.

734 Kiyosawa, S. (1982) Genetics and epidemiological modeling of breakdown of plant disease
735 resistance. *Annual Review of Phytopathology*, 20, 93–117.
736 https://doi.org/10.1146/annurev.py.20.090182.000521.

737 Knight, N.L. & Pethybridge, S.J. (2020) An improved PCR assay for species-specific detection
738 and quantification of *Cercospora beticola*. *Canadian Journal of Plant Pathology*, 42,
739 72–83. https://doi.org/10.1080/07060661.2019.1621380.

740 Kokot, M., Długosz, M. & Deorowicz, S. (2017) KMC 3: counting and manipulating k-mer
741 statistics Berger, B. (Ed.). *Bioinformatics*, 33, 2759–2761.
742 https://doi.org/10.1093/bioinformatics/btx304.

743 Lo Presti, L., Lanver, D., Schweizer, G., Tanaka, S., Liang, L., Tollot, M., et al. (2015) Fungal
744 effectors and plant susceptibility. *Annual Review of Plant Biology*, 66, 513–545.
745 https://doi.org/10.1146/annurev-arplant-043014-114623.

746 McDonald, B.A., Suffert, F., Bernasconi, A. & Mikaberidze, A. (2022) How large and diverse
747 are field populations of fungal plant pathogens? The case of *Zymoseptoria tritici*.
748 *Evolutionary Applications*, 15, 1360–1373. https://doi.org/10.1111/eva.13434.

749 Meile, L., Croll, D., Brunner, P.C., Plissonneau, C., Hartmann, F.E., McDonald, B.A., et al.
750 (2018) A fungal avirulence factor encoded in a highly plastic genomic region triggers
751 partial resistance to septoria tritici blotch. *New Phytologist*, 219, 1048–1061.
752 https://doi.org/10.1111/nph.15180.

753 Petit-Houdenot, Y. & Fudal, I. (2017) Complex interactions between fungal avirulence genes
754 and their corresponding plant resistance genes and consequences for disease
755 resistance management. *Frontiers in Plant Science*, 8, 1072.
756 https://doi.org/10.3389/fpls.2017.01072.

757 Pilet-Nayel, M.-L., Moury, B., Caffier, V., Montarry, J., Kerlan, M.-C., Fournet, S., et al. (2017)
758 Quantitative resistance to plant pathogens in pyramiding strategies for durable crop
759 protection. *Frontiers in Plant Science*, 8, 1838.
760 https://doi.org/10.3389/fpls.2017.01838.

761 Rangel, L.I., Spanner, R.E., Ebert, M.K., Pethybridge, S.J., Stukenbrock, E.H., Jonge, R., et al.
762 (2020) *Cercospora beticola*: the intoxicating lifestyle of the leaf spot pathogen of
763 sugar beet. *Molecular Plant Pathology*, 21, 1020–1041.
764 https://doi.org/10.1111/mpp.12962.

765 Rouxel, T. & Balesdent, M. (2017) Life, death and rebirth of avirulence effectors in a fungal
766 pathogen of Brassica crops, *Leptosphaeria maculans* . *New Phytologist*, 214, 526–
767 532. https://doi.org/10.1111/nph.14411.

768 Savary, S., Willocquet, L., Pethybridge, S.J., Esker, P., McRoberts, N. & Nelson, A. (2019) The
769 global burden of pathogens and pests on major food crops. *Nature Ecology &*
770 *Evolution*, 3, 430–439. https://doi.org/10.1038/s41559-018-0793-y.

771 Schürch, S., Linde, C.C., Knogge, W., Jackson, L.F. & McDonald, B.A. (2004) Molecular
772 population genetic analysis differentiates two virulence mechanisms of the fungal
773 avirulence gene *NIP1*. *Molecular Plant-Microbe Interactions*, 17, 1114–1125.
774 https://doi.org/10.1094/MPMI.2004.17.10.1114.

775 Secor, G.A., Rivera, V.V., Khan, M.F.R. & Gudmestad, N.C. (2010) Monitoring fungicide
776 sensitivity of *Cercospora beticola* of sugar beet for disease management decisions.
777 *Plant Disease*, 94, 1272–1282. https://doi.org/10.1094/PDIS-07-09-0471.

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

782 Sonah, H., Deshmukh, R.K. & Bélanger, R.R. (2016) Computational prediction of effector
783 proteins in fungi: opportunities and challenges. *Frontiers in Plant Science*, 7.
784 https://doi.org/10.3389/fpls.2016.00126.

785 Sperschneider, J. & Dodds, P.N. (2022) EffectorP 3.0: prediction of apoplastic and
786 cytoplasmic effectors in fungi and oomycetes. *Molecular Plant-Microbe Interactions*,
787 35, 146–156. https://doi.org/10.1094/MPMI-08-21-0201-R.

788 Taguchi, K., Kubo, T., Takahashi, H. & Abe, H. (2011) Identification and precise mapping of
789 resistant QTLs of Cercospora leaf spot resistance in sugar beet (*Beta vulgaris* L.). *G3*
790 *Genes|Genomes|Genetics*, 1, 283–291. https://doi.org/10.1534/g3.111.000513.

791 Törjék, O., Borchardt, D., Mechelke, W, Schulz, B, & Lein, J C (2022) Plant resistance gene
792 and means for its identification.

793 Törjék, O., Borchardt, D., Rekoskem, M., Mechelke, W., Schulz, B. & Lein, J.C. (2020) Gene
794 conferring resistance to *Cercospora beticola* in beets.

795 Tsirigos, K.D., Peters, C., Shu, N., Käll, L. & Elofsson, A. (2015) The TOPCONS web server for
796 consensus prediction of membrane protein topology and signal peptides. *Nucleic*
797 *Acids Research*, 43, W401–W407. https://doi.org/10.1093/nar/gkv485.

798 Vaghefi, N., Kikkert, J.R., Bolton, M.D., Hanson, L.E., Secor, G.A., Nelson, S.C., et al. (2017)
799 Global genotype flow in *Cercospora beticola* populations confirmed through
800 genotyping-by-sequencing Chiang, T.-Y. (Ed.). *PLOS ONE*, 12, e0186488.
801 https://doi.org/10.1371/journal.pone.0186488.

802 Van De Wouw, A.P., Cozijnsen, A.J., Hane, J.K., Brunner, P.C., McDonald, B.A., Oliver, R.P., et
803 al. (2010) Evolution of linked avirulence effectors in *Leptosphaeria maculans* is
804 affected by genomic environment and exposure to resistance genes in host plants
805 Sil, A. (Ed.). *PLoS Pathogens*, 6, e1001180.
806 https://doi.org/10.1371/journal.ppat.1001180.

807 Windels, C.E., Lamey, H.A., Hilde, D., Widner, J. & Knudsen, T. (1998) A Cercospora leaf spot
808 model for sugar beet: In practice by an industry. *Plant Disease*, 82, 716–726.

809 Winter, M. & Koopmann, B. (2016) Race spectra of *Leptosphaeria maculans*, the causal
810 agent of blackleg disease of oilseed rape, in different geographic regions in northern
811 Germany. *European Journal of Plant Pathology*, 145, 629–641.
812 https://doi.org/10.1007/s10658-016-0932-8.

813 Wyk, S. van, Harrison, C.H., Wingfield, B.D., De Vos, L., Merwe, N.A. van der & Steenkamp,
814 E.T. (2019) The RIPper, a web-based tool for genome-wide quantification of Repeat-
815 Induced Point (RIP) mutations. *PeerJ*, 7, e7447. https://doi.org/10.7717/peerj.7447.

816 Zhan, J. & McDonald, B.A. (2013) Experimental measures of pathogen competition and
817 relative fitness. *Annual Review of Phytopathology*, 51, 131–153.
818 https://doi.org/10.1146/annurev-phyto-082712-102302.

819 Zhong, Z., Marcel, T.C., Hartmann, F.E., Ma, X., Plissonneau, C., Zala, M., et al. (2017) A small
820 secreted protein in *Zymoseptoria tritici* is responsible for avirulence on wheat
821 cultivars carrying the *Stb6* resistance gene. *New Phytologist*, 214, 619–631.
822 https://doi.org/10.1111/nph.14434.

823

824

825 **Table 1:** Sugar beet hybrids used for the virulence assessments.

Experimental Hybrids	***Cercospora* resistance level**
Hybrid_S4	low quantitative resistance
Hybrid_S3	moderate quantitative resistance
Hybrid_R2	*BvCR4* with low quantitative background resistance
Hybrid_R3	*BvCR4* with moderate quantitative background resistance

826

827

828 **Table 2.** Frequency of unique (clone corrected) isolates carrying a deletion of the candidate
829 avirulence effector gene. Frequency is the percentage of the total number of individuals
830 analyzed. The number of individuals with deletions and the total are shown in parentheses.
831

Country	Sampling	Year	Resistant: Deletion freq. (n)	Susceptible: Deletion freq. (n)	Clone Corrected (Total)
Switzerland	Pilot	2019	25% (1/4)	NA	4 (7)
Switzerland	Temporal	2020	32% (22/68)	0% (0/42)	110 (147)
Switzerland	Transect	2020	57% (4/7)	0% (0/15)	22 (38)
Germany	Transect	2020	60% (6/10)	0% (0/17)	27 (40)
USA	Public	2016-17	NA	1.0% (1/95)	95 (190)

832

833

834 **Figure 1.** Disease rating of isolates at five different time points in the field trial. Each panel
835 shows the rating on one of the four different hybrids (2 susceptible and 2 resistant). Red
836 points are isolates scored as virulent (rating >4 on Hybrid_R2 at 40 dpi) and black are
837 avirulent (rating ≤4 on Hybrid_R2 at 40 dpi), lines are loess-smoothed lines. Black letters at
838 the top of the plots indicate least significant differences between mean rating of all isolates
839 at each time point. Red text indicate difference in disease rating between virulent and
840 avirulent for that hybrid and time point (ns=nonsignificant, *p<0.05, **p<0.01, ***
841 p<0.001). No analysis was performed on Hybrid_K2 because data from this cultivar was used
842 to score virulence/avirulence.
843
844 **Figure 2.** Manhattan plots of GWAS. a) Manhattan plot from a genome-wide association
845 (GWAS) scan using 52 unique isolates with quantitative phenotypes at 40 dpi on Hybrid_R2.
846 b) Manhattan plot from GWAS using 90 unique isolates, with 52 isolates scored with binary
847 phenotypes (virulent or avirulent) and 38 isolates collected from Hybrid_S4 scored as
848 avirulent. c) Manhattan plot from a k-mer GWAS using 47 unique isolates with quantitative
849 phenotypes. Horizontal line is the genome-wide significance level. All SNPs and k-mers with
850 p-value >1.22e-01 were downsampled for plotting.
851
852 **Figure 3.** Manhattan plot of the significant GWAS interval (± 25 kb of flanking sequence) on
853 chromosome 1:4087634-4094364 bp (shaded green). Genes are shaded in blue, full length
854 transposable elements are shaded in grey and large RIP affected regions (LRAR) are shaded
855 in yellow. Inset: Gene and repeat density around the GWAS interval (including 300 kb of
856 flanking sequence). Gene (solid line) and repeat (dashed line) density shown in 10 kb
857 windows. The GWAS interval is shaded in green and the LRAR is shaded in yellow.
858
859 **Figure 4.** DNA and amino acid (aa) sequence variation in the candidate effector gene. a)
860 Haplotype networks for isolates collected in Switzerland from resistant and susceptible
861 hosts and in Renville USA. Ref = Identical to reference genome, 'aa' numbers correspond to
862 the translated effector sequence variation shown in (b). b) Shows the 13 (incl. Ref) different
863 amino acid sequences and the aa sequence for *Cercospora sojina*. In parentheses are the
864 numbers and their host or virulence status (if phenotyped): sus = collected on susceptible
865 cultivar, res = collected on resistant hybrid, virulence rating based on rating at 40 dpi on

866 Hybrid_R2 vir >4 virulent, avr ≤4 avirulent. c) Effect of amino acid sequence variation (seen
867 in b) and gene deletion on virulence rating (40 dpi Hybrid_R2).
868
869 **Figure 5.** Disease rating of isolates at five different time points in the field trial coloured
870 according to the effector gene copy (blue=same as reference, orange=gene absent or
871 sequence codes for different amino acids (inactivated)). Each panel shows the rating on one
872 of the four different hybrids (2 susceptible and 2 resistant). Stars at the top of the plots
873 indicate significant differences between the gene variants at each time point
874 (ns=nonsignificant, *p<0.05, **p<0.01, *** p<0.001).
875

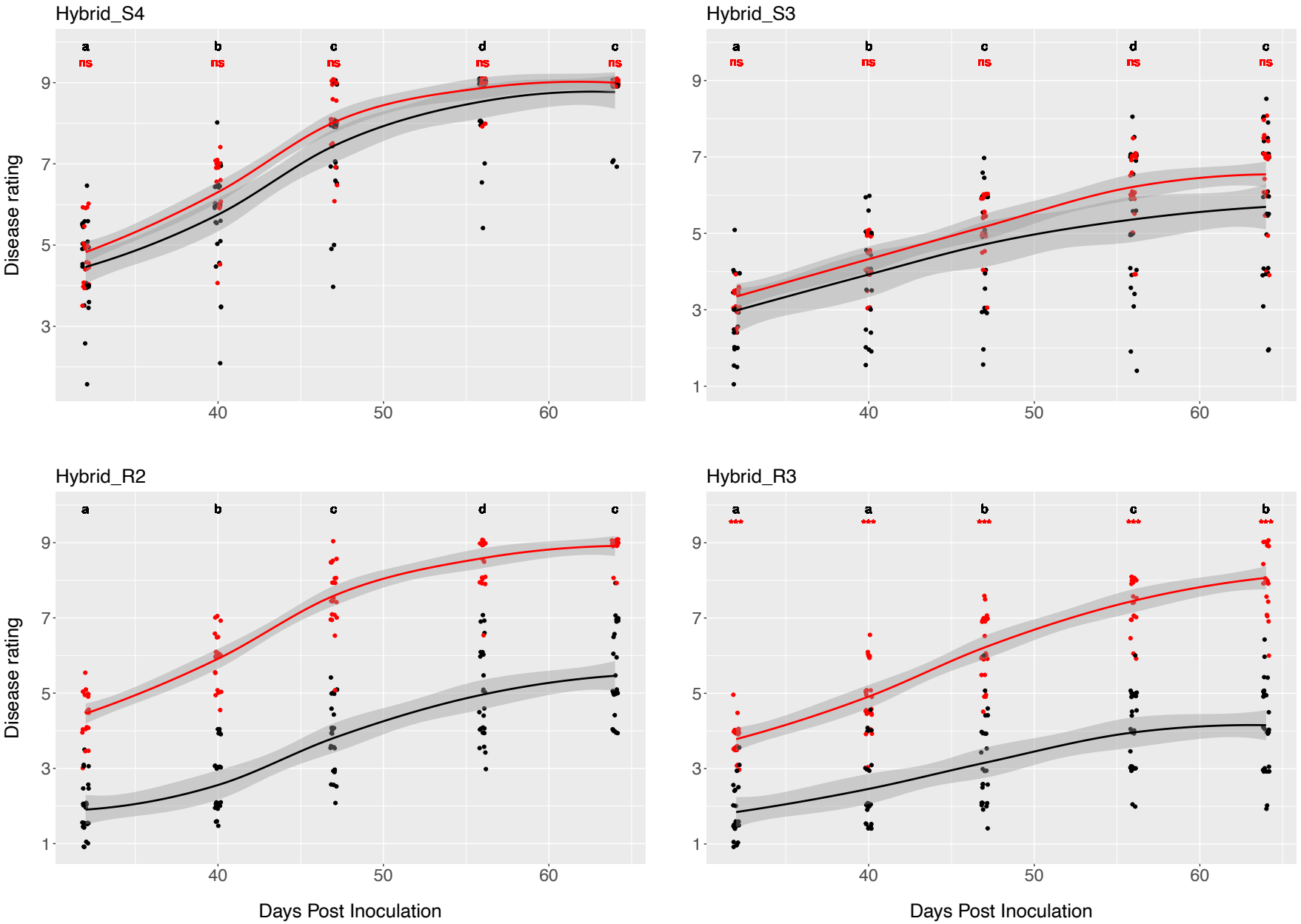

Figure 1. Disease rating of isolates at five different time points in the field trial. Each panel shows the rating on one of the four different hybrids (2 susceptible and 2 resistant). Red points are isolates scored as virulent (rating >4 on Hybrid_R2 at 40dpi) and black are avirulent (rating ≤4 on Hybrid_R2 at 40dpi), lines are loess-smoothed lines. Black letters at the top of the plots indicate least significant differences between mean rating of all isolates at each time point. Red text indicate difference in disease rating between virulent and avirulent for that hybrid and time point (ns=nonsignificant, *p<0.05, **p<0.01, *** p<0.001). No analysis was performed on Hybrid_K2 because data from this cultivar was used to score virulence/avirulence.

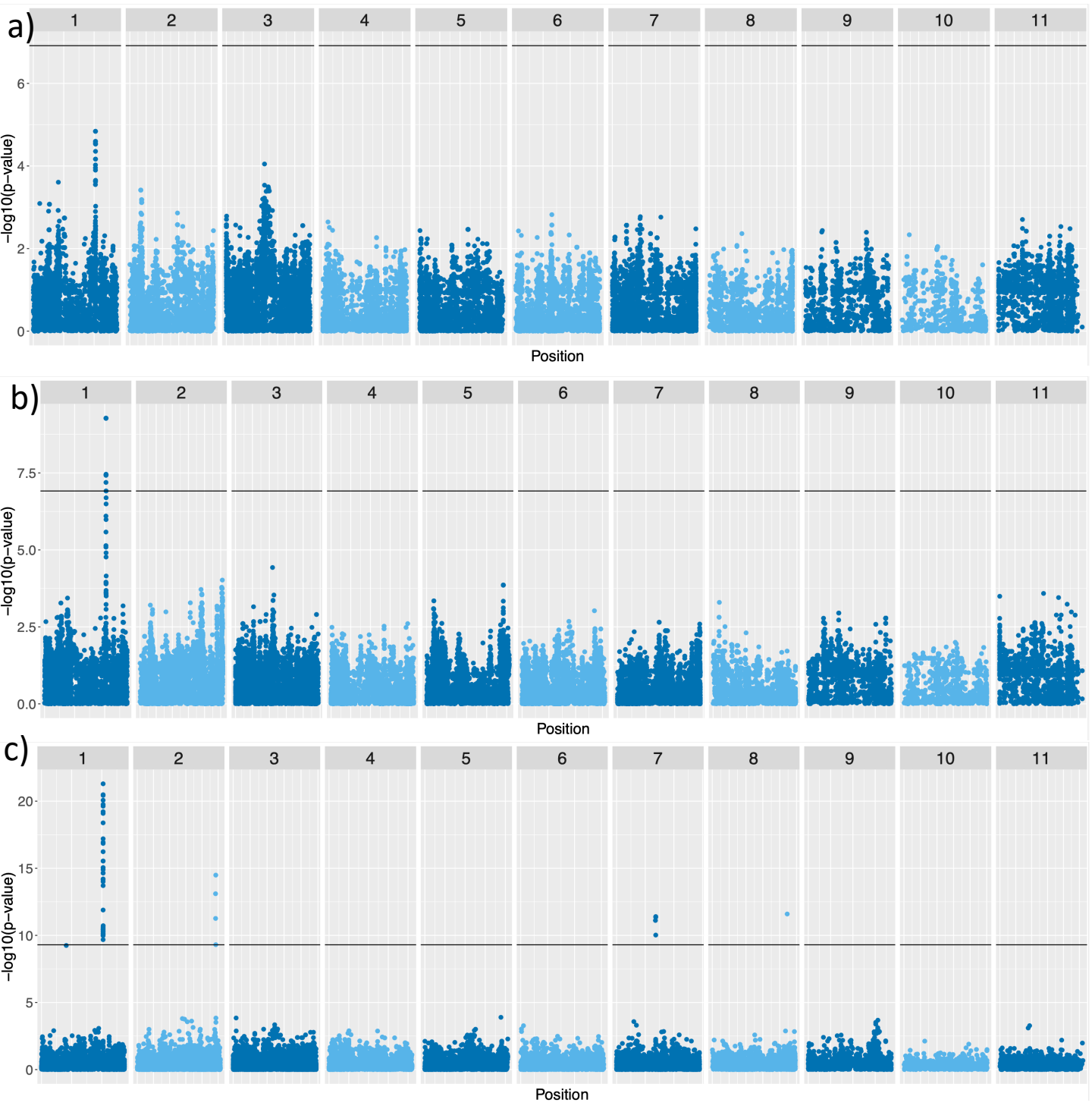

Figure 2. Manhattan plots of GWAS. a) Manhattan plot from a genome-wide association (GWAS) scan using 52 unique isolates with quantitative phenotypes at 40 dpi on Hybrid_R2. b) Manhattan plot from GWAS using 90 unique isolates, with 52 isolates scored with binary phenotypes (virulent or avirulent) and 38 isolates collected from Hybrid_S4 scored as avirulent. c) Manhattan plot from a k-mer GWAS using 47 unique isolates with quantitative phenotypes. Horizontal line is the genome-wide significance level. All SNPs and k-mers with p-value >1.22e-01 were downsampled for plotting.

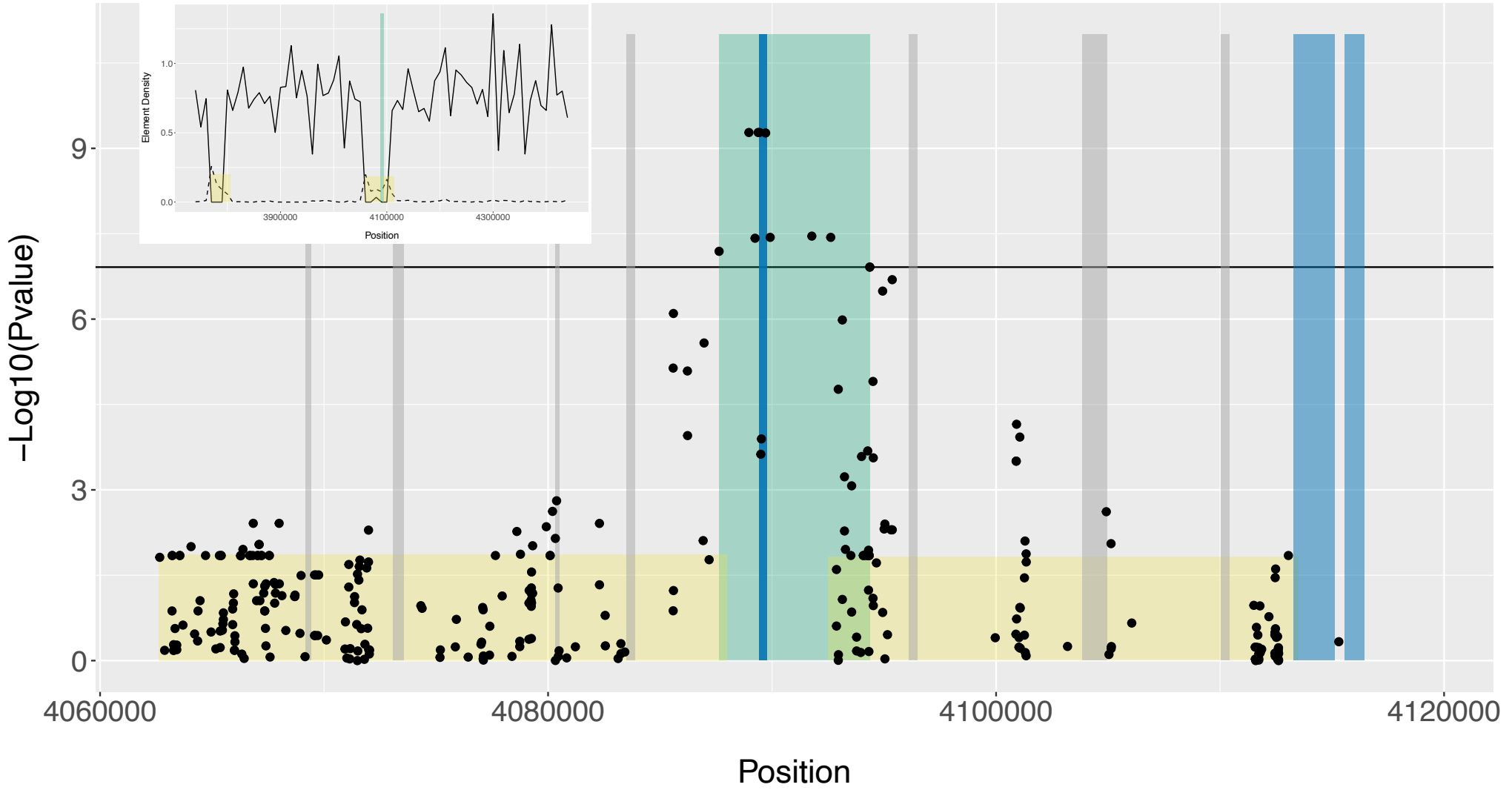

Figure 3. Manhattan plot of the significant GWAS interval (± 25 kb of flanking sequence) on chromosome 1:4087634-4094364 bp (shaded green). Genes are shaded in blue, full length transposable elements are shaded in grey and large RIP affected regions (LRAR) are shaded in yellow. **Inset:** Gene and repeat density around the GWAS interval (including 300 kb of flanking sequence). Gene (solid line) and repeat (dashed line) density shown in 10 kb windows. The GWAS interval is shaded in green and the LRAR is shaded in yellow.

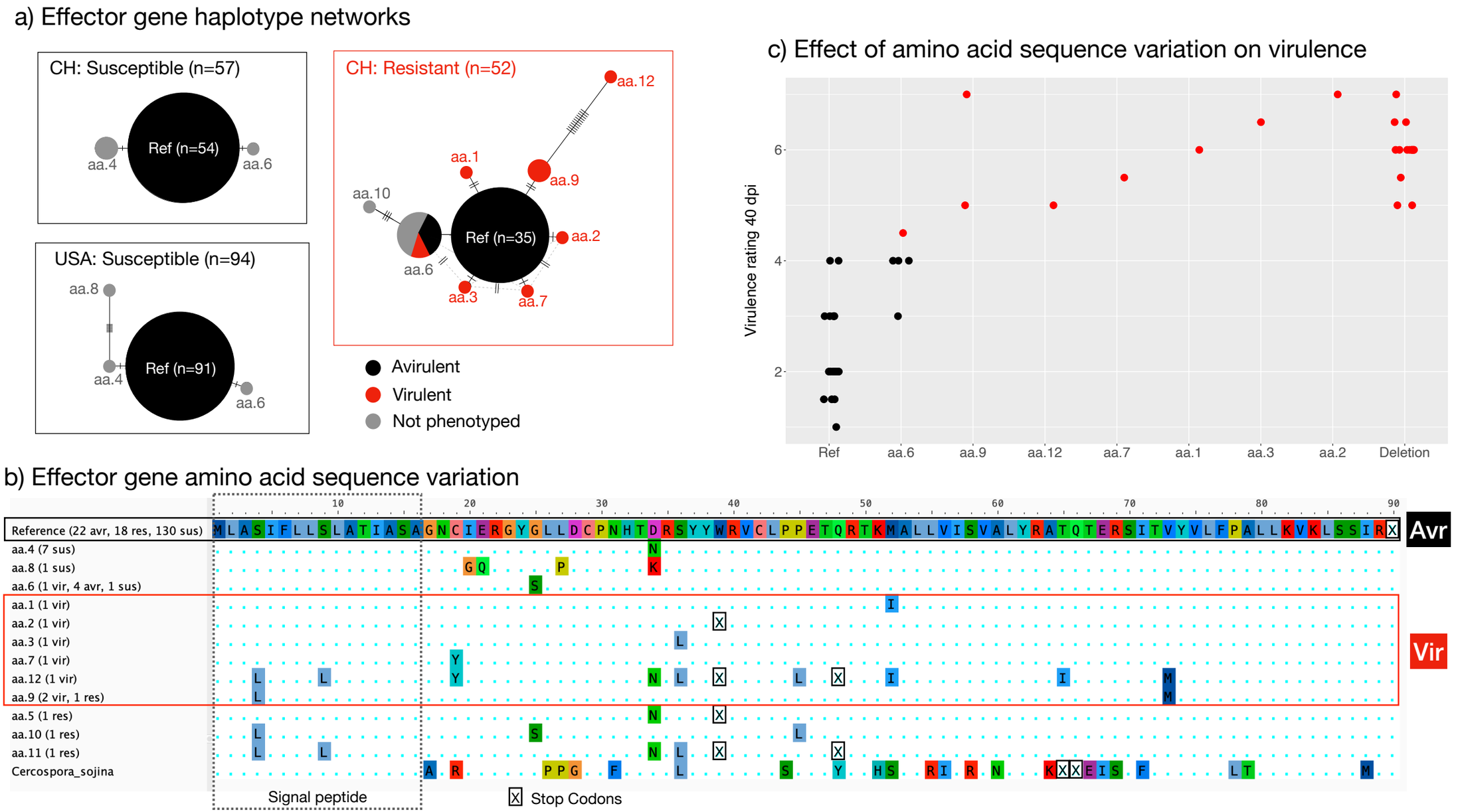

Figure 4. DNA and amino acid (aa) sequence variation in the candidate effector gene. a) Haplotype networks for isolates collected in Switzerland from resistant and susceptible hosts and in Renville USA. Ref = Identical to reference genome, 'aa' numbers correspond to the translated effector sequence variation shown in (b). b) Shows the 13 (incl. Ref) different amino acid sequences and the aa sequence for *Cercospora sojina*. In parentheses are the numbers and their host or virulence status (if phenotyped): sus = collected on susceptible hybrid, res = collected on resistant hybrid, virulence rating based on rating at 40 dpi on Hybrid_R2 (vir >4 virulent, avr ≤4 avirulent). c) Effect of amino acid sequence variation (seen in b) and gene deletion on virulence rating (40 dpi Hybrid_R2).

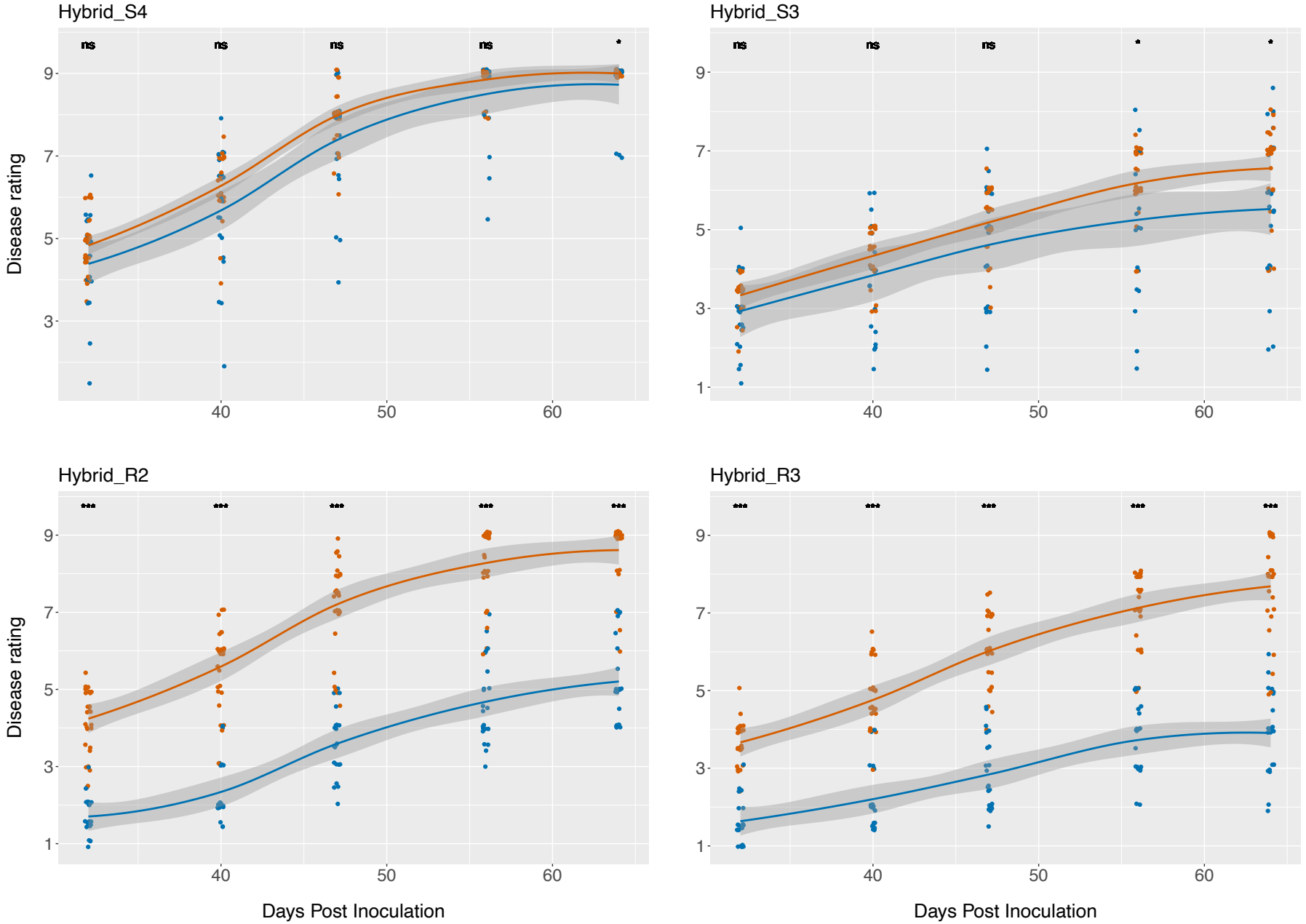

Figure 5. Disease rating of isolates at five different time points in the field trial coloured according to the effector gene copy (blue=same as reference, orange=gene absent or sequence codes for different amino acids (inactivated)). Each panel shows the rating on one of the four different hybrids (2 susceptible and 2 resistant). Stars at the top of the plots indicate significant differences between the gene variants at each time point (ns=nonsignificant, *p<0.05, **p<0.01, *** p<0.001).