Published as: *Science*. 2013 August 16; 341(6147): 783–786.

Identification of Wheat Gene *Sr35* that Confers Resistance to Ug99 Stem Rust Race Group

Cyrille Saintenac[#1], **Wenjun Zhang**[#2], **Andres Salcedo**[1], **Matthew N. Rouse**[3], **Harold N. Trick**[1], **Eduard Akhunov**[#1,*], and **Jorge Dubcovsky**[#2,4,*]

[1] Kansas State University, Dept. Plant Pathology, Manhattan, KS 66506

[2] University of California, Davis Dept. Plant Sciences, Davis, CA 95616

[3] USDA-ARS Cereal Disease Laboratory, St. Paul, MN 55108

[4] Howard Hughes Medical Institute, Chevy Chase, MD 20815

[#] These authors contributed equally to this work.

Abstract

Wheat stem rust, caused by *Puccinia graminis* f. sp. *tritici* (*Pgt*), is a devastating disease that can cause severe yield losses. A new *Pgt* race designated Ug99 has overcome most of the widely used resistance genes and is threatening major wheat production areas. Here we demonstrate that the *Sr35* gene from *Triticum monococcum* is a coiled coil-nucleotide binding-leucine rich repeat gene that confers near-immunity to Ug99 and related races. This gene is absent in the A-genome diploid donor and in polyploid wheat, but is effective when transferred from *T. monococcum* to polyploid wheat. The cloning of *Sr35* opens the door to the use of biotechnological approaches to control this devastating disease and to the analyses of the molecular interactions that define the wheat-rust pathosystem.

Puccinia graminis f. sp. *tritici* (henceforth *Pgt*) is the causal agent of wheat stem rust, a devastating disease responsible for major outbreaks and large losses of wheat yields in the past. The deployment of *Pgt* resistance genes, combined with the eradication of the alternative host (barberry) provided an effective control of this disease for the last fifty years (*1*). However, the widely deployed *Pgt* resistance gene *Sr31* was overcome by a new race of *Pgt* identified in Uganda in 1999 designated Ug99 (or TTKSK according to the North American system for *Pgt* race nomenclature) (*2*). A decade later, six new Ug99-related *Pgt*

[*] Correspondence to: Eduard Akhunov, eakhunov@ksu.edu, Dept. Plant Pathology, Kansas State University, Manhattan, KS 66506, phone (785)-532-1342; and Jorge Dubcovsky, jdubcovsky@ucdavis.edu, University of California, Davis Dept. Plant Sciences, Davis, CA95616, phone (530) 752-5159. .

One Sentence Summary

The wheat stem rust resistance gene *Sr35* is effective against the new *Puccinia graminis* f. sp. *tritici* race Ug99.

Supplementary Materials

Materials and Methods

Figures S1-S7

Tables S1-S10

References (25-47)

Author contributions

Author contributions are listed in the Supplemental Online Materials.

races, some showing a broader virulence spectrum, have been detected and have spread to the wheat growing regions of Africa, Yemen and Iran (*3*). Roughly 90% of the wheat varieties grown worldwide are susceptible to Ug99 and related races, which represents a serious threat to global food security (*3*). The Borlaug Global Rust Initiative was launched in 2005 to coordinate international efforts to fight Ug99 (http://www.globalrust.org). The identification and characterization of Ug99 resistance genes *Sr35* in this study and *Sr33* in a companion paper (*4*) are part of these efforts.

The stem rust resistance gene *Sr35* was identified in previous screens for resistance to *Pgt* in the diploid wheat species *Triticum monococcum* (*5, 6*). The genome of *T. monococcum*, designated A^m, is closely related to the genome of *T. urartu*, the diploid donor of the A genome in tetraploid (*T. turgidum*, pasta wheat) and hexaploid wheat (*T. aestivum*, bread wheat) (*7*). *Sr35* was prioritized for cloning because it confers near-immunity against Ug99, Ug99-realated races, and the TRTTF group of races from Africa, Yemen and Pakistan, which has a broad but different virulence profile from the Ug99 race group (*3, 8*). *Sr35* was also selected because previous studies have confirmed that this gene is effective against the same virulent races when it is transferred to hexaploid wheat by crossing and recombination (*5, 8*).

Sr35 was previously mapped on the long arm of chromosome $3A^m$ in *Triticum monococcum* (*8*). In this study, we used 4,575 recombinant gametes and seven molecular markers derived from the colinear region in *Brachypodium distachyon* (Fig. 1A) to map *Sr35* between markers *AK331487* (0.02 cM) and *AK332451* (0.98 cM, Fig. 1B). We then used the closest proximal markers *AK331487* and *SFGH* (*S-formylglutathione hydrolase-like*) to screen a *T. monococcum* BAC library of the *Sr35*-resistant accession DV92 (*9*). The 23 selected BAC clones were assembled by fingerprinting into a single contig that spanned the *Sr35* locus (Fig. 1C-D, Table S1, Fig. S1).

We sequenced three overlapping BACs covering the *Sr35* region (*10*) and annotated the 307,519 bp sequence (KC573058). This sequence includes a cluster of coiled coil-nucleotide binding-leucine rich repeat (henceforth *CNL*) disease resistance genes including five intact genes (*CNL1*, *CNL2*, *CNL4*, *CNL6* and *CNL9*), two pseudogenes (*pCNL3* and *pCNL10*) and three small gene fragments (*pCNL5*, *pCNL7* and *pCNL8*) (Fig. 1E). A phylogenetic tree of the complete *CNL* genes showed that *CNL4* and *CNL9* are the most closely related members of this cluster (Fig. S2). The annotated sequence also includes two unrelated genes (*SFGH* and *APGG1*) and two pseudogenes (*pABC* and *pAP2*) (Fig. 1E). Additional markers developed from this sequence were used to delimit the *Sr35* candidate region to a 213 kb segment including candidate genes *APGG1*, *CNL4*, *CNL6* and *CNL9* (Table S1, Fig. 1E).

We sequenced these four candidate genes in a *T. monococcum* collection including 24 Ug99-resistant accessions carrying *Sr35* and 25 susceptible accessions without *Sr35*. We identified two resistant (R1 and R2) and six susceptible haplotypes (S1-S6, Table S2, primers in Tables S3-S5). The two resistant haplotypes differ in a short *CNL4* region with a 6-bp deletion and four single nucleotide polymorphisms (SNPs) but show no differences in *APGG1*, *CNL6* and *CNL9*. All susceptible accessions have mutations in *CNL9* and among them five have mutations only in *CNL9*, which suggests that this gene is necessary to confer

resistance to Ug99. Among these five susceptible accessions, three share three close SNPs that result in amino acid changes at positions 854, 856 and 858 (RLWFT to HLRFS) in the C-terminal region of the leucine-rich repeat domain (Fig. 2A). The same three SNPs are present in the closely related *CNL4* gene suggesting a conversion event.

To validate the previous results, we mutagenized the *Sr35*-resistant accession G2919 with ethyl-methanesulfonate (*10*). Out of 1,087 M_2 mutant families screened with race RKQQC we identified two mutant families segregating for susceptibility, which were validated with races Ug99 and TRTTF (Fig. 2B). Sequencing of the four candidate genes in these susceptible plants confirmed the presence of mutations only in *CNL9*. The first mutant (*cnl9*^{1296}) contained a G to A mutation that resulted in a premature stop codon at position 856 (Fig. 2A) and truncated the last 64 amino acids. In the progeny of a cross between *cnl9*^{1296} and the resistant parental line G2919 (33 F_2 plants), homozygocity for the mutation co-segregated with susceptibility to Ug99.

The second susceptible mutant (*cnl9*^{1120}) showed the same three SNPs detected in accession PI428167-2 (RLWFT to HLRFS, Table S2). To test if this was the result of seed contamination or crosspollination, we used genotyping-by-sequencing (*11*) to estimate the level of polymorphisms among *cnl9*^{1296}, *cnl9*^{1120} and the non-mutagenized line G2919 (Table S6 and supplemental text). We show that *cnl9*^{1120} has the level of mutations and the ratio of homozygous to heterozygous loci expected from a mutagenized plant. Therefore, a spontaneous gene conversion between *CNL4* and *CNL9* is the most parsimonious explanation for the three linked mutations in *cnl9*^{1120}. Two of these amino acid positions (856 and 858) overlap with 15 amino acids located in the C-terminal half of the leucine-rich repeat domain of CNL9 that show evidence of positive selection (Fig. S3A, S3B and Table S7). In summary, mutants *cnl9*^{1296} and *cnl9*^{1120} confirmed that *CNL9* is necessary for the *Sr35*-mediated resistance and that the distal region of the leucine-rich repeat domain is critical for *Sr35* function.

To determine if *CNL9* is sufficient to confer resistance to Ug99, we generated transgenic hexaploid wheat plants expressing the *CNL9* gene under the control of its native promoter (*10*). Out of four putative T_0 transgenic plants only one, designated #1123, showed consistent expression of the transgene (Fig. S4) and co-segregation between the presence of the transgene and resistance to Ug99 and RKQQC in the T_1 and T_2 progeny (Fig. 2C, Fig. S5, Table S8). In contrast, all #1123 T_1 and T_2 plants were susceptible to the *Sr35*-virulent race QTHJC, regardless of the presence or absence of the transgene (Fig. 2C, Fig. S5). This result suggests that the *CNL9* transgene has the same race specificity as *Sr35*. Taken together; the natural variation, mutant and transgenic results demonstrate that *CNL9* is *Sr35*.

Using rapid amplification of cDNA ends (*10*) we found that the *CNL9* transcripts have a 196 bp 5'UTR and a 1,526 bp 3'UTR that includes three introns (Fig. S6A). The three introns in the 3' UTR were also detected in all the *T. urartu*, *T. turgidum* cv. *durum*, and *T. aestivum* related *CNL* genes for which we were able to obtain both genomic and transcript data (Table S9). Both *CNL* homologs from *Brachypodium distachyon* (Table S9) also have two introns in the 3'UTR, which indicates that this structural feature is conserved in this disease

resistance cluster. Exons 3 and 4 from the *B. distachyon CNL* genes correspond to exons 4 and 5 from the *T. monococcum CNL9* homolog.

Transcript levels of *CNL9* in leaves from G2919 plants inoculated with *Pgt* race RKQQC (*10*) were 40-, 81- and 411-fold higher than those of candidate genes *APGG1*, *CNL4*, and *CNL6*, respectively (Fig. 2D); but not significantly different from mock-inoculated plants at different time points (Fig. 2E). Using isoform specific primers (Table S5, Fig. S6B), we found that roughly 8% of the *T. monococcum CNL9* transcripts were represented by an alternative splicing variant that retained the second intron in the 3'UTR (Fig. 2E). We also detected transcripts with and without the same intron in *T. turgidum* (Table S9). The ratio between the two *CNL9* transcript isoforms did not show changes in *T. monococcum* G2919 plants mock-inoculated and inoculated with *Pgt* race RKQQC (Fig. 2E). This suggests that the relative proportion of the two alternative splice forms is not affected by the presence of the pathogen. Previously reported alternative splicing events in *CNL* genes do not involve introns in the 3' UTR (*12-18*), which might be a distinctive feature of this particular group of *CNL* genes.

Sr35 has not been reported so far in *T. urartu* or polyploid wheat species. To understand better the reasons for this absence we performed a comparative analysis of the *T. monococcum* (KC573058) and *T. urartu* (KC816724) colinear regions, which diverged less than one million years ago (*19*). The *T. monococcum* region encompassing genes *CNL6* and *CNL9* and pseudogenes *pCNL5, pCNL8, pCNL10* and *pABC* is absent in *T. urartu* (Fig. 1F). Conversely, the *T. urartu* region including *TuCNL-D* and pseudogene *pCNL-E* is missing in *T. monococcum*. Large insertions and deletions have been found in other colinear intergenic regions of the *T. monococcum* and *T. urartu* genomes (*19, 20*). The large and repetitive genomes of wheat show higher rates of insertion and deletions than the human genome (*19*).

A screen of 41 *T. urartu* accessions and 19 wild tetraploid wheat *T. turgidum* ssp. *dicoccoides* accessions (Table S10) revealed no orthologues of *TmCNL9* (Fig. S7). Gene *TuCNL-H* from *T. urartu* accession G1545 from Iran encoded the same RWT amino acids found in CNL9 at positions 854, 856 and 858, but the rest of the sequence was different and clustered with a separate set of *CNL* genes (Fig. S7). Since *T. urartu* is the donor of the A genome to the polyploid wheat species (*7*), it is not surprising that *CNL9* homologues have not been detected in the genomic sequence of *T. aestivum* (http://www.wheatgenome.org/) or in the transcriptome of *T. turgidum* (http://wheat.pw.usda.gov/GG2/WheatTranscriptome/) (Fig. S2).

The absence of *Sr35* in the tested pasta and bread wheat varieties highlights the value of wheat landraces and wild relatives as a reservoir of novel resistance specificities. It also suggests that *Sr35* has the potential to improve stem rust resistance in a wide range of wheat germplasm. Our transgenic experiments also indicate that the transfer of *CNL9-Sr35* to hexaploid wheat is sufficient to confer effective levels of resistance to Ug99. In contrast, some *CNL* genes, as for example wheat leaf rust resistance gene *Lr10*, require the presence of additional *CNL* genes to provide resistance (*21*).

CNL proteins mediate recognition of pathogen-derived effector molecules as well as host protein altered by the pathogen and subsequently activate host defenses. These proteins have an ancient origin and are encoded by one of the largest, most variable multigene families in plants (*22*). Members of this family confer resistance to a wide range of pathogens and pests. Remaining challenges are to identify which genes are responsible for resistance to a specific pathogen and to understand the signal transduction pathways involved in the plant resistance response. This information is particularly important in the case of Ug99, which now threatens the major wheat producing areas in Asia (*3*).

The identification of *Sr35* and of *Sr33* in a companion paper (*4*) opens the door to transgenic approaches to control this devastating pathogen. *Sr35* shows a strong hypersensitive reaction to the TTKSK and TRTTF race groups when introgressed into hexaploid what, but is susceptible to some *Pgt* races and, therefore, should not be deployed alone. In contrast, *Sr33* is resistant to all races tested so far (*23*, *24*) but confers only moderate resistance to the Ug99 race group when introgressed alone in hexaploid wheat. Based on these complementary characteristics it might be beneficial to combine these two genes either by crossing and recombination or by transforming wheat with a cassette including both genes. The insertion of multiple resistance genes in a single locus can accelerate breeding efforts to pyramid multiple sources of resistance, which is a reasonable strategy to increase the durability of available resistance genes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This project was supported by the National Research Initiative Competitive Grants 2011-68002-30029 (Triticeae-CAP) and 2012-67013-19401 from the USDA National Institute of Food and Agriculture, by the Borlaug Global Rust Initiative and by support to JD from the Howard Hughes Medical Institute and the Gordon and Betty Moore Foundation. We thank Jayaveeramuthu Nirmala, Sunshi Sehgal, Mariana Padilla, Shiaoman Chao, Katherine Jordan, Hyeonju Lee and Dehlia Burdan for excellent technical support, Bob Bowden, Jan Dvorak and Mike Pumphrey for providing critical materials, Ksenia Krasileva, Alina Akhunova, and Chengxia Li for valuable suggestions, and the UC-Davis and K-state Genomic facilities. Sequences have been deposited in GenBank under accession numbers KC573058, KC816724, KF113354-KF113357 and KC876115-KC876121.

References and Notes

1. McIntosh, RA.; Wellings, CR.; Park, RF. Wheat rusts, an atlas of resistance genes. Jean, K., editor. CSIRO; Melbourne, Australia: 1995. p. 200
2. Pretorius ZA, Singh RP, Wagoire WW, Payne TS. Plant Dis. 2000; 84:203.
3. Singh RP, et al. Annu. Rev. Phytopathol. 2011; 49:465. [PubMed: 21568701]
4. Periyannan S, et al. Science. 2013 To be completed by Science.
5. McIntosh RA, Dyck PL, The TT, Cusick J, Milne DL. Plant Breeding. 1984; 92:1.
6. Rouse MN, Jin Y. Plant Dis. 2011; 95:941.
7. Dvorak J, McGuire PE, Cassidy B. Genome. 1988; 30:680.
8. Zhang W, et al. Crop Sci. 2010; 50:2464.
9. Lijavetzky D, et al. Genome. 1999; 42:1176. [PubMed: 10659785]
10. SOM, Materials and methods are available as supplementary material on Science Online
11. Saintenac C, Zhang D, Wang S, Akhunov E. G3 (Bethesda). 2013 doi:10.1534/g3.113.005819.

12. Costanzo S, Jia YL. Plant Sci. 2009; 177:468.
13. Ferrier-Cana E, et al. Theor. Appl. Genet. 2005; 110:895. [PubMed: 15660237]
14. Tan XP, et al. BMC Plant Biol. 2007; 7:56. [PubMed: 17956627]
15. Sela H, et al. Mol. Plant. Pathol. 2012; 13:276. [PubMed: 21952112]
16. Halterman DA, Wei FS, Wise RP. Plant Physiol. 2003; 131:558. [PubMed: 12586880]
17. Halterman DA, Wise RP. Mol. Plant Pathol. 2006; 7:167. [PubMed: 20507437]
18. Gassmann W. Nuclear pre-mRNA processing in plants. 2008; 326:219–233.
19. Dubcovsky J, Dvorak J. Science. 2007; 316:1862. [PubMed: 17600208]
20. Wicker T, et al. The Plant Cell. 2003; 15:1186. [PubMed: 12724543]
21. Loutre C, et al. Plant J. 2009; 60:1043. [PubMed: 19769576]
22. Yue JX, Meyers BC, Chen JQ, Tian DC, Yang SH. New Phytol. 2012; 193:1049. [PubMed: 22212278]
23. Huerta-Espino, J. University of Minnesota; 1992.
24. Rouse MN, Olson EL, Gill BS, Pumphrey MO, Jin Y. Crop Sci. 2011; 51:2074.
25. Wilkinson PA, et al. BMC Bioinformatics. 2012; 13:219. [PubMed: 22943283]
26. Luo MC, et al. Genomics. 2003; 82:378. [PubMed: 12906862]
27. Soderlund C, Longden I, Mott R. Comput. Appl. Biosci. 1997; 13:523. [PubMed: 9367125]
28. Akhunov ED, Akhunova AR, Dvorak J. Theor. Appl. Genet. 2005; 111:1617. [PubMed: 16177898]
29. Dubcovsky J, et al. Genetics. 1996; 143:983. [PubMed: 8725244]
30. Carver TJ, et al. Bioinformatics. 2005; 21:3422. [PubMed: 15976072]
31. Tamura K, et al. Mol. Biol. Evol. 2011; 28:2731. [PubMed: 21546353]
32. Fu LM, Niu BF, Zhu ZW, Wu ST, Li WZ. Bioinformatics. 2012; 28:3150. [PubMed: 23060610]
33. Langmead B, Trapnell C, Pop M, Salzberg SL. Genome Biol. 2009; 10:R25. [PubMed: 19261174]
34. Li H, et al. Bioinformatics. 2009; 25:2078. [PubMed: 19505943]
35. Kosakovsky Pond SL, Frost SDW, Muse SV. Bioinformatics. 2005; 21:676. [PubMed: 15509596]
36. Scheffler K, Martin DP, Seoighe C. Bioinformatics. 2006; 22:2493. [PubMed: 16895925]
37. Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SDW. Mol. Biol. Evol. 2006; 23:1891. [PubMed: 16818476]
38. Christensen AH, Quail PH. Transgenic Res. 1996; 5:213. [PubMed: 8673150]
39. Ayella AK, Trick HN, Wang WQ. Mol. Nutr. Food Res. 2007; 51:1518. [PubMed: 18030664]
40. Klosterman SJ. Methods Mol. Biol. 2012; 835:121. [PubMed: 22183651]
41. Long XY, et al. Plant Mol. Biol. 2010; 74:307. [PubMed: 20658259]
42. Chen A, Dubcovsky J. PLOS Genet. 2012; 8:e1003134. [PubMed: 23271982]
43. Rawat N, et al. BMC Plant Biol. 2012; 12:205. [PubMed: 23134614]
44. Kilian B, et al. Mol Biol Evol. 2007; 24:2657. [PubMed: 17898361]
45. Takken FLW, Goverse A. Curr. Opin. Plant Biol. 2012; 15:375. [PubMed: 22658703]
46. Kelley LA, Sternberg MJE. Nat. Protoc. 2009; 4:363. [PubMed: 19247286]
47. Stakman EC, Steward DM, Loegering WQ. Identification of physiologic races of Puccinia graminis var. tritici. 1962

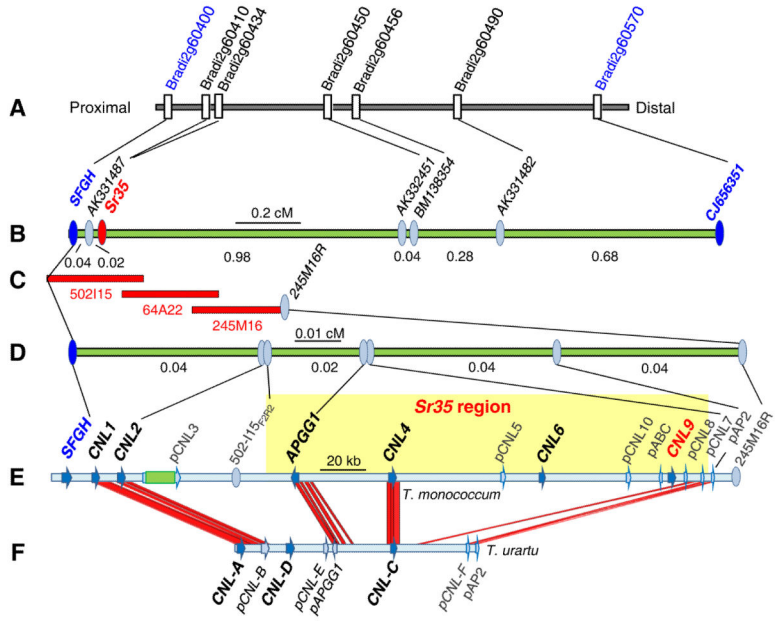

Fig. 1.

Genetic and physical maps of *Sr35*. (**A**) 174-kb collinear region of *Brachypodium* (*8*). Only genes for which a wheat orthologous gene was found in databases are represented. (**B**) Genetic map of *Sr35* locus. (**C**) Screening the DV92 BAC library with proximal markers *SFGH* and *AK331487* (only BACs from the minimum tilling path are shown). (**D**) High density map. (**E**) Graphical representation of the *T. monococcum* annotated sequences (KC573058). *CNL*= coiled coil-nucleotide binding-leucine rich repeat genes, 'p' before gene name = pseudogene (*pCNL3* has an inserted retroelement). The *Sr35* candidate gene region is highlighted in yellow. (**F**) Comparison of *T. monococcum* DV92 and *T. urartu* G1812 (KC816724) orthologous regions (92% identity threshold).

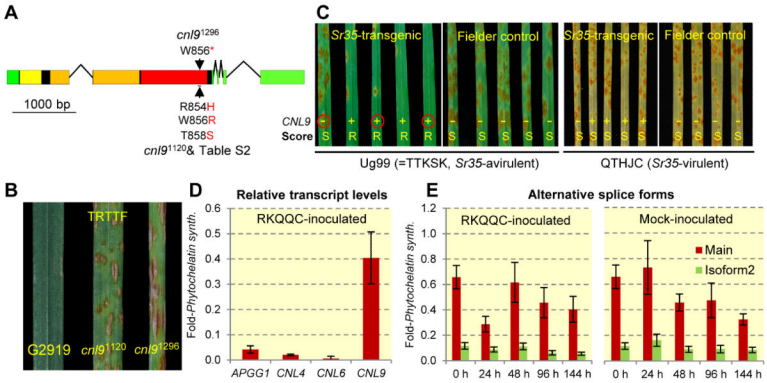

Fig. 2.

(A) *CNL9* gene structure. Green=UTR, black= coding exons, yellow= coiled-coil domain, orange= nucleotide-binding domain, red= leucine-rich repeat domain, arrows= amino acid changes in susceptible induced mutants *cnl9*^{1296} (W856*) and *cnl9*^{1120} or natural mutants (Table S2, RLWFT to HLRFS) (**B**) Infection types produced on *T. monococcum* G2919 and *CNL9* mutants *cnl9*^{1120} and *cnl9*^{1296} inoculated with *Pgt* race TRTTF. TRTTF is *Sr35*-avirulent and *Sr21*-virulent, which is required because of the presence of *Sr21* in these lines. (**C)** Infection types on seedlings of T_1 lines from event #1123 segregating for the *CNL9* transgene. Plants carrying the *CNL9* transgene (+) were resistant to Ug99 (R) and plants without the transgene (−) were susceptible (S) (Table S8). When inoculated with *Sr35*-virulent race QTHJC all plants were susceptible suggesting similar race specificity between the transgenic and the natural *Sr35*. Red circles indicate available progeny tests in Figure S5. (**D**) Relative transcript levels of candidate genes *APPG1*, *CNL4*, *CNL6*, and *CNL9* (main isoform) in G2919 six days after inoculation with race RKQQC (**E**) Transcript levels of *CNL9* main isoform (red) and isoform two (green, retained intron) in mock- or race RKQQC inoculated G2919 plants. Leaves were collected at 0, 24, 48, 96 and 144 h after inoculation. Transcript levels are expressed relative to the *Phytochelatin synthase* internal control using the $2^{-\Delta Ct}$ method. Bars are standard errors of the means based on six biological and two technical replicates.