We detected common genomic hallmarks in the powdery mildews associated with obligate biotrophy. These include gene losses and extensive gene reshuffling correlated with expansion in (retro-)transposon number and genome size. Together, these hallmarks may represent a tradeoff between advantages of increased genetic variation independent of sexual recombination and irreversible deletion of genes dispensable for biotrophy. Hence, their evolution provides a notable example of Dollo's law of evolutionary irreversibility (*15*). This may explain why powdery mildews and possibly other biotrophic parasites became obligate.

References and Notes

1. D. A. Glawe, *Annu. Rev. Phytopathol.* **46**, 27 (2008).
2. Materials and methods are available as supporting material on *Science* Online.
3. J. R. Miller *et al*., *Bioinformatics* **24**, 2818 (2008).
4. S. Takamatsu, *Mycoscience* **45**, 147 (2004).
5. N. Möbius, C. Hertweck, *Curr. Opin. Plant Biol.* **12**, 390 (2009).
6. J. Collemare, A. Billard, H. U. Böhnert, M. H. Lebrun, *Mycol. Res.* **112**, 207 (2008).
7. J. Glazebrook, *Annu. Rev. Phytopathol.* **43**, 205 (2005).
8. B. L. Cantarel *et al*., *Nucleic Acids Res.* **37** (database issue), D233 (2009).
9. H. A. B. Wösten, *Annu. Rev. Microbiol.* **55**, 625 (2001).
10. R. Debuchy, B. Turgeon, in *The Mycota I: Growth, Differentiation and Sexuality*, U. Kües, R. Fischer, Eds. (Springer, Berlin, 2006), vol. I, pp. 293–323.
11. S. Sacristán *et al*., *PLoS ONE* **4**, e7463 (2009).
12. R. Panstruga, P. N. Dodds, *Science* **324**, 748 (2009).
13. D. Godfrey *et al*., *BMC Genomics* **11**, 317 (2010).
14. A.-M. Catanzariti, P. N. Dodds, G. J. Lawrence, M. A. Ayliffe, J. G. Ellis, *Plant Cell* **18**, 243 (2006).
15. C. R. Marshall, E. C. Raff, R. A. Raff, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 12283 (1994).

16. We thank B. Thomas, T. Carver, and D. Grunewald for providing scanning electron micrographs. This work was supported by funds from the Biotechnology and Biological Sciences Research Council [grant BB/E0009831/1], INRA, the BioExploit European Union Framework 6 project, Deutsche Forschungsgemeinschaft (priority program SPP1212), the Leverhulme Trust, and the Max Planck Society. GenBank project identification numbers are as follows: *Blumeria*, 28821; *E. pisi*, 50315; and *G. orontii*, 50317.

Supporting Online Material

www.sciencemag.org/cgi/content/full/330/6010/1543/DC1
Materials and Methods
SOM Text
Figs. S1 to S9
Tables S1 to S6
References

2 July 2010; accepted 1 November 2010
10.1126/science.1194573

Pathogenicity Determinants in Smut Fungi Revealed by Genome Comparison

Jan Schirawski,[1,2]* Gertrud Mannhaupt,[1,3] Karin Münch,[1] Thomas Brefort,[1]† Kerstin Schipper,[1] Gunther Doehlemann,[1] Maurizio Di Stasio,[1] Nicole Rössel,[1] Artemio Mendoza-Mendoza,[1]‡ Doris Pester,[1]§ Olaf Müller,[1]|| Britta Winterberg,[1]¶ Elmar Meyer,[1] Hassan Ghareeb,[1]* Theresa Wollenberg,[1]* Martin Münsterkötter,[3] Philip Wong,[3] Mathias Walter,[3] Eva Stukenbrock,[1] Ulrich Güldener,[3] Regine Kahmann[1]#

Biotrophic pathogens, such as the related maize pathogenic fungi *Ustilago maydis* and *Sporisorium reilianum*, establish an intimate relationship with their hosts by secreting protein effectors. Because secreted effectors interacting with plant proteins should rapidly evolve, we identified variable genomic regions by sequencing the genome of *S. reilianum* and comparing it with the *U. maydis* genome. We detected 43 regions of low sequence conservation in otherwise well-conserved syntenic genomes. These regions primarily encode secreted effectors and include previously identified virulence clusters. By deletion analysis in *U. maydis*, we demonstrate a role in virulence for four previously unknown diversity regions. This highlights the power of comparative genomics of closely related species for identification of virulence determinants.

Smut fungi are biotrophic pathogens causing disease in a number of agriculturally important crop plants. *Ustilago maydis* and the related fungus *Sporisorium reilianum* both parasitize maize (*1*, *2*). Their life cycle leading to the infectious form is similar (*2*, *3*); however, shortly after infection *U. maydis* locally induces tumors on all aerial parts of the plant, whereas *S. reilianum* spreads systemically and causes symptoms in male and female inflorescences only (Fig. 1). Both *S. reilianum* and *U. maydis* establish an intimate communication with their host through secreted protein effectors that enable biotrophic development (*3*, *4*). Effector proteins like *U. maydis* Pep1 can suppress plant defense responses (*5*). Additional effector genes were identified in the genome as genes encoding *U. maydis*–specific secreted proteins, most of which are up-regulated during host colonization (*3*). Many of these effector genes are clustered, and deletion of five of these clusters affected virulence in seedlings (*3*). Some cluster genes are induced in specific plant organs, and respective cluster mutants show altered virulence depending on the host tissue infected (*6*). In plant parasitic oomycetes, genes for effector proteins are under diversifying selection and occur in highly flexible genomic regions (*7*). In accordance with this emerging picture of plant-pathogen communication via rapidly evolving effector proteins, we hypothesized that virulence-associated *U. maydis* genes might be identified as genomic regions with high sequence variability in closely related smut species.

To identify regions of high diversity in the *U. maydis* genome, we sequenced the genome of *S. reilianum* strain SRZ2 (*8*). The *S. reilianum* genome assembly covers 97% of the 18.7-Mb genome (*9*). As in *U. maydis* (*3*), the genome is organized in 23 chromosomes, to which 6648 gene models could be assigned after manual annotation. The genomes of *U. maydis* and *S. reilianum* exhibit a remarkable degree of synteny (Fig. 2A) (*10*) despite an average amino acid identity of predicted proteins of only 74.2% (Fig. 2B). Interestingly, some chromosome ends are extended by up to 20 genes in *U. maydis* compared with ends in *S. reilianum*. About 90% of these chromosome end–associated genes do not carry any functional annotation and no enrichment for secreted effectors is evident (fig. S1), whereas the others likely encode enzymes for secondary metabolism (table S1). Because orthologs of these genes are lacking in *S. reilianum*, their presence is likely dispensable for virulence. Compared with an average amino acid identity of 76% for nonsecreted proteins, secreted proteins in both organisms display an average identity of only 62% and are enriched among the weakly conserved proteins (Fig. 2B). This suggests that genes coding for secreted proteins are subject to more rapid evolution.

Manual sequence comparisons of predicted gene models of *S. reilianum* and *U. maydis* led to a reannotation of more than 300 gene models of *U. maydis* (table S2). The *S. reilianum* genome has a 5.7% higher GC content than the *U. maydis* genome and a 5% higher coding potential (table S2). More than 99% of all InterPro (www.ebi.ac.uk/interpro/) domains are equally or close to equally represented in the two genomes, suggesting that the biosynthetic repertoire of both species is comparable. However, *S. reilianum* contains three putative RNA-dependent RNA polymerase genes (table S3). A search for other components (*11*, *12*) of a putative RNA interference (RNAi) machinery in *S. reilianum* identified homologs of *dicer* and *argonaute* (table S3). These genes all lie in highly syntenic regions (*10*); however, the corresponding intergenic regions in *U. maydis* lack traces of the

[1]Department of Organismic Interactions, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch Straße 10, 35043 Marburg, Germany. [2]Department for Molecular Biology of Plant-Microbe Interactions, Albrecht-von-Haller Institute for Plant Sciences, Georg-August-Universität Göttingen, Untere Karspüle 2, 37073 Göttingen, Germany. [3]Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstädter Landstraße 1, 85764 Neuherberg, Germany.

*Present address: Department for Molecular Biology of Plant-Microbe Interactions, Albrecht-von-Haller Institute for Plant Sciences, Georg-August-Universität Göttingen, Untere Karspüle 2, 37073 Göttingen, Germany.
†Present address: Febit Biomed GmbH, Im Neuenheimer Feld 519, 69120 Heidelberg, Germany.
‡Present address: Bio-Protection Research Centre, Post Office Box 84, Lincoln University, Lincoln 7647, New Zealand.
§Present address: Institute for Plant Health, Austrian Agency for Health and Food Safety (AGES), Spargelfeldstraße 191, 1226 Wien, Austria.
||Present address: Roy J. Carver Center for Comparative Genomics, University of Iowa, Department of Biological Sciences, 303 Biology Building, Iowa City, IA 52242, USA.
¶Present address: Research School of Biology, Building 41, Linnaeus Way, Australian National University, Canberra, ACT 0200, Australia.
#To whom correspondence should be addressed. E-mail: Kahmann@mpi-marburg.mpg.de

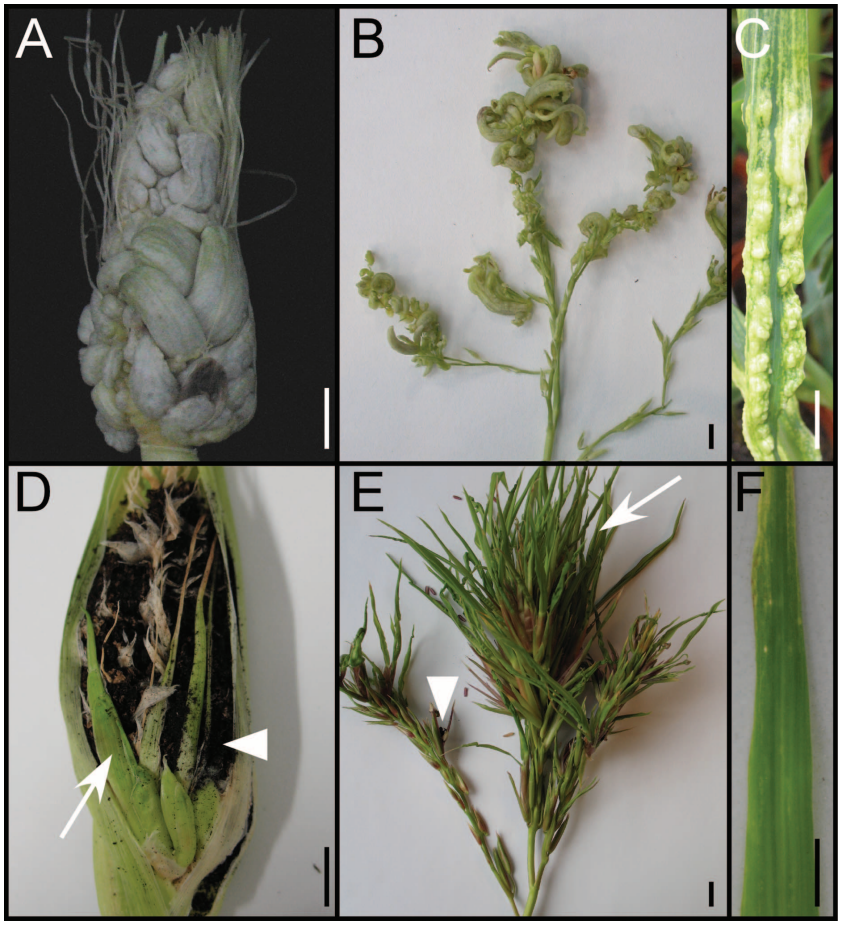

Fig. 1. Infection symptoms of *U. maydis* and *S. reilianum*. Tumor formation on ear (**A**), tassel (**B**), and leaf (**C**) after inflorescence infection [(A) and (B)] or seedling infection (C) by *U. maydis*. Maize seedling infection with *S. reilianum* leads to formation of spores (arrowheads) and leaflike structures (arrows) in ear (**D**) and tassel (**E**), whereas inoculated leaves show mild chlorosis but no tumors (**F**). Scale bars indicate 1 cm.

respective genes. Instead, we detected between one and four variants of a conserved 10–base pair (bp) sequence highly overrepresented in intergenic regions of *U. maydis* (*3*) that is largely absent from the genome of *S. reilianum* (table S4). We speculate that the genes encoding components of the RNAi machinery were lost in *U. maydis* by rare homologous recombination events involving the 10-bp motifs. To investigate whether the generation of small regulatory RNAs could explain the differences in symptoms of *U. maydis* and *S. reilianum*, we deleted the putative dicer gene *sr16838* in a solopathogenic strain of *S. reilianum* (*8*). Infection experiments revealed that *sr16838* deletion mutants were affected in neither virulence nor symptom development (fig. S2).

To detect regions of high sequence divergence, we compared the genomes of *U. maydis* and *S. reilianum* gene by gene (*8*, *10*) and identified regions encoding genes with low sequence conservation ("divergence clusters") in a conserved genomic context (*8*). This analysis revealed the presence of 43 divergence clusters (table S5) (*10*). Seventy-one percent of the genes in divergence clusters occurred in both organisms, whereas 19% were *S. reilianum*–specific and 10% were *U. maydis*–specific. Sixty-one percent of the genes in divergence clusters are predicted to encode secreted proteins. In contrast, of all *S. reilianum* and *U. maydis* genes less than 12% encode potentially secreted proteins. Ninety-four of the genes in divergence clusters code for proteins without functional annotation (table S5). It is notable that 442 of the 494 putative effectors detected in *U. maydis* are conserved in *S. reilianum*, with amino acid identities ranging from 20 to 98% (*10*). In addition, 445 genes are present only in *U. maydis*, whereas 372 exist only in *S. reilianum*, and of these about 15% encode secreted proteins.

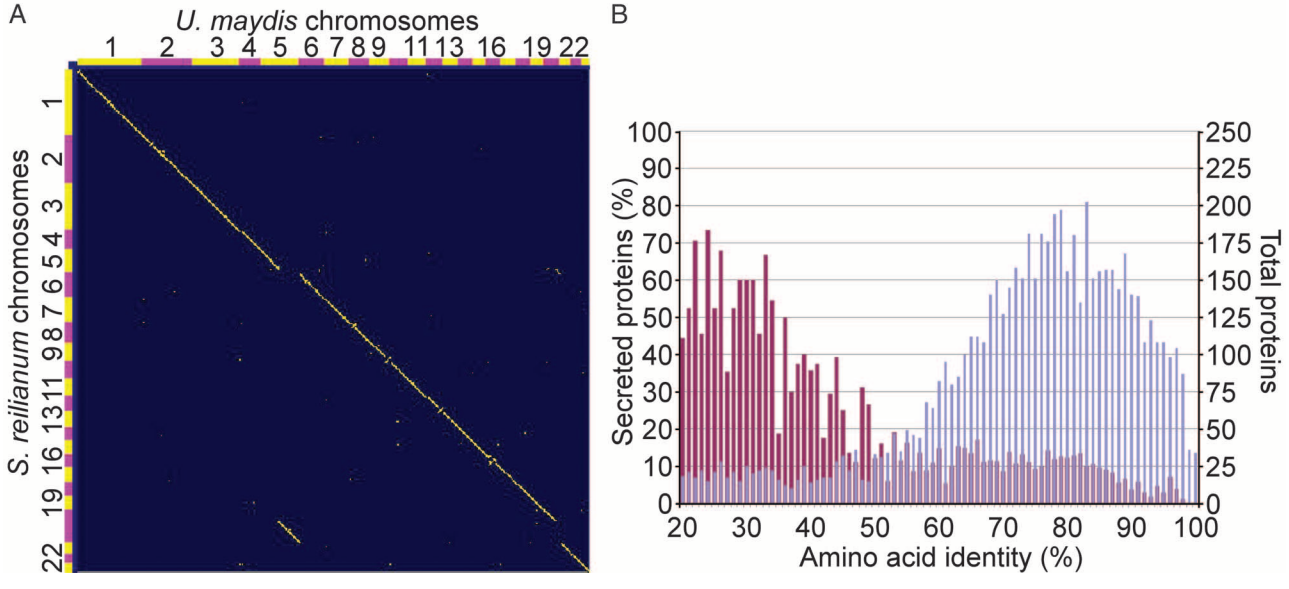

Fig. 2. Comparison of *U. maydis* and *S. reilianum* genomes. (**A**) Synteny (diagonal lines) of protein-encoding genes on the 23 chromosomes of *U. maydis* compared to those on the 23 chromosomes of *S. reilianum*. (**B**) Distribution of amino acid identities of all protein-encoding genes occurring in both genomes (right axis, blue bars). The percentage of proteins with a predicted secretion signal is given for each amino acid identity value (left axis, red bars).

Among the divergence gene clusters identified in this study were 7 of the 12 previously described *U. maydis* effector gene clusters (*3*), and these included all clusters whose deletion affected virulence in seedling infections (*3*) (fig. S3). In contrast, four of the gene clusters of secreted proteins without a deletion phenotype in *U. maydis* seedling infections (*3*) did not classify as divergence clusters (fig. S4). To test whether any of the newly identified divergence gene clusters also harbor virulence factors, we individually deleted six randomly chosen divergence clusters in *U. maydis* (Fig. 3). In seedling infection assays, loss of three divergence gene clusters (15-12, 5-21, and 20-15) attenuated virulence, whereas deletion of two clusters (5-18 and 11-16) did not affect virulence (Fig. 3 and fig. S5). The absence of a virulence phenotype likely reflects redundancy because in these cases potential paralogs exist elsewhere in the genome (table S6). In cluster 8-12, we detected the *mig1* gene that is highly induced during plant colonization and encodes an effector with similarity to apoplastic fungal avirulence proteins (*13*, *14*). Cluster 8-12 contains three additional genes, of which two encode proteins with similarity to Mig1. In *S. reilianum*, the *mig1* gene family is expanded to eight members residing in a single cluster (Fig. 3 and fig. S6). Whereas the deletion of only *mig1* did not affect virulence (*13*), cluster 8-12 deletion caused hypervirulence (fig. S5). Hypervirulence could result from an active attenuation of fungal proliferation by respective effectors (*3*). However, given the conserved features between avirulence proteins and Mig1 effectors, we now propose that genes whose deletion leads to hypervirulence encode weak avirulence proteins that trigger defense responses in plants expressing a cognate resistance protein, resulting in an attenuation of fungal growth.

Although most of the deleted genes in the four divergence clusters with an effect on virulence encode putatively secreted proteins, cluster 15-12 encodes only proteins without identifiable secretion signals, suggesting that additional mechanisms for virulence modulation exist in *U. maydis*. With respect to the origin of these clusters, we do not detect hallmarks for horizontal gene transfer like an altered GC content or an association with repetitive elements (fig. S7). Therefore, because many divergence regions contain members of small gene families (table S4) we propose that the majority of divergence clusters have been generated by local gene duplications followed by strong natural selection resulting from interaction with different host molecules.

Our studies demonstrate the power of comparative genomics of closely related species for the identification of new virulence genes. The *U. maydis* and *S. reilianum* pathosystems present a unique example of differentiation of two closely related pathogens parasitizing the same host. We have recognized that the *U. maydis* and *S. reilianum* genomes comprise conserved effector genes as expected for pathogens infecting the same host. However, although the two pathogens are both recognized and challenged by the maize immune system, they also possess strongly differentiated effectors, suggesting that they are targeting different host molecules. We speculate that their different infection strategies lead to the interaction between different host-pathogen molecules and thereby the evolution of differentiated sets of effector proteins, although we cannot exclude contributions of the species-specific genes. The assertion that closely related pathogens interact with and affect different host targets suggests a high variety in pathogen targets within the same host. It also suggests a temporal and spatial difference in the composition of different host proteins, which can drive the evolution of different sets of effectors in pathogens with different infection strategies.

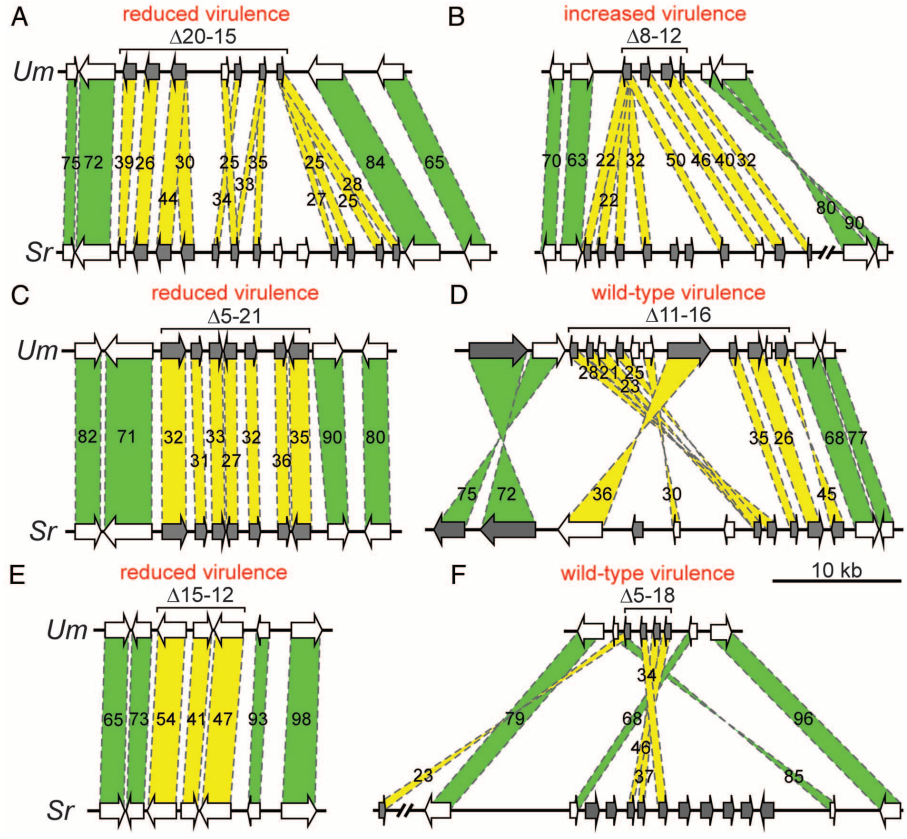

Fig. 3. Gene-by-gene comparisons of divergence clusters between *U. maydis* (*Um*) and *S. reilianum* (*Sr*) deleted in this study. (**A**) Cluster 15-12, (**B**) cluster 20-15, (**C**) cluster 11-16, (**D**) cluster 8-12, (**E**) cluster 5-21, and (**F**) cluster 5-18. Genes encoding putatively secreted proteins are shaded in gray. Bars connecting syntenic homologs are color-coded (green, high; yellow, weak conservation), and numbers give amino acid identities. Brackets denote regions deleted in *U. maydis* mutants. Virulence phenotypes of the respective mutants are indicated; the corresponding scores are found in fig. S5.

References and Notes

1. M. Stoll, D. Begerow, F. Oberwinkler, *Mycol. Res.* **109**, 342 (2005).
2. C. Martinez, C. Roux, R. Dargent, *Phytopathology* **89**, 247 (1999).
3. J. Kämper *et al.*, *Nature* **444**, 97 (2006).
4. T. Brefort *et al.*, *Annu. Rev. Phytopathol.* **47**, 423 (2009).
5. G. Doehlemann *et al.*, *PLoS Pathog.* **5**, e1000290 (2009).
6. D. S. Skibbe, G. Doehlemann, J. Fernandes, V. Walbot, *Science* **328**, 89 (2010).
7. B. J. Haas *et al.*, *Nature* **461**, 393 (2009).
8. Materials and methods are available as supporting material on *Science* Online.
9. MIPS, *Sporisorium reilianum* data base (MSRDB), http://mips.helmholtz-muenchen.de/genre/proj/sporisorium.
10. http://mips.helmholtz-muenchen.de/genre/export/sites/default/sporisorium/Download/Genome_comparison.xls
11. E. Bernstein, A. A. Caudy, S. M. Hammond, G. J. Hannon, *Nature* **409**, 363 (2001).
12. C. Catalanotto, G. Azzalin, G. Macino, C. Cogoni, *Nature* **404**, 245 (2000).
13. C. W. Basse, S. Stumpferl, R. Kahmann, *Mol. Cell. Biol.* **20**, 329 (2000).
14. I. Stergiopoulos, P. J. de Wit, *Annu. Rev. Phytopathol.* **47**, 233 (2009).

15. We are grateful to the Max Planck Society for funding genome sequencing and annotation. T.B. was supported through the Deutsche Forschungsgemeinschaft via SFB593. A.M.M. received a Humboldt fellowship. H.G. was funded by the International Max Planck Research School for Environmental, Cellular, and Molecular Microbiology. R.K. acknowledges support through the LOEWE program of the State of Hesse. We thank U. Bonas and G. Gottschalk for critical comments on the manuscript. The sequence was submitted to European Molecular Biology Laboratory database under accession numbers FQ311430 to FQ311474.

Supporting Online Material

www.sciencemag.org/cgi/content/full/330/6010/1546/DC1
Materials and Methods
Figs. S1 to S7
Tables S1 to S6
References

19 July 2010; accepted 25 October 2010
10.1126/science.1195330

Pathogenicity Determinants in Smut Fungi Revealed by Genome Comparison

Jan Schirawski, Gertrud Mannhaupt, Karin Münch, Thomas Brefort, Kerstin Schipper, Gunther Doehlemann, Maurizio Di Stasio, Nicole Rössel, Artemio Mendoza-Mendoza, Doris Pester, Olaf Müller, Britta Winterberg, Elmar Meyer, Hassan Ghareeb, Theresa Wollenberg, Martin Münsterkötter, Philip Wong, Mathias Walter, Eva Stukenbrock, Ulrich Güldener, and Regine Kahmann

Science **330** (6010), . DOI: 10.1126/science.1195330

From Blight to Powdery Mildew

Pathogenic effects of microbes on plants have widespread consequences. Witness, for example, the cultural upheavals driven by potato blight in the 1800s. A variety of microbial pathogens continue to afflict crop plants today, driving both loss of yield and incurring the increased costs of control mechanisms. Now, four reports analyze microbial genomes in order to understand better how plant pathogens function (see the Perspective by **Dodds**). **Raffaele *et al.*** (p. 1540) describe how the genome of the potato blight pathogen accommodates transfer to different hosts. **Spanu *et al.*** (p. 1543) analyze what it takes to be an obligate biotroph in barley powdery mildew, and **Baxter *et al.*** (p. 1549) ask a similar question for a natural pathogen of *Arabidopsis*. **Schirawski *et al.*** (p. 1546) compared genomes of maize pathogens to identify virulence determinants. Better knowledge of what in a genome makes a pathogen efficient and deadly is likely to be useful for improving agricultural crop management and breeding.