



**Common Sequence Variants in the LOXL1 Gene
Confer Susceptibility to Exfoliation Glaucoma**

Gudmar Thorleifsson, *et al.*

Science **317**, 1397 (2007);

DOI: 10.1126/science.1146554

**The following resources related to this article are available online at
www.sciencemag.org (this information is current as of September 21, 2007):**

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/317/5843/1397>

Supporting Online Material can be found at:

<http://www.sciencemag.org/cgi/content/full/1146554/DC1>

This article **cites 22 articles**, 7 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/317/5843/1397#otherarticles>

This article appears in the following **subject collections**:

Genetics

<http://www.sciencemag.org/cgi/collection/genetics>

Information about obtaining **reprints** of this article or about obtaining **permission to reproduce this article** in whole or in part can be found at:

<http://www.sciencemag.org/about/permissions.dtl>

11. A. T. Saurin, H. Neubert, J. P. Brennan, P. Eaton, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 17982 (2004).
12. J. P. Brennan *et al.*, *J. Biol. Chem.* **279**, 41352 (2004).
13. J. P. Brennan *et al.*, *J. Biol. Chem.* **281**, 21827 (2006).
14. F. Hofmann, R. Feil, T. Kleppisch, J. Schlossmann, *Physiol. Rev.* **86**, 1 (2006).
15. R. Feil, S. M. Lohmann, H. de Jonge, U. Walter, F. Hofmann, *Circ. Res.* **93**, 907 (2003).
16. J. R. Schnell, G. P. Zhou, M. Zweckstetter, A. C. Rigby, J. J. Chou, *Protein Sci.* **14**, 2421 (2005).
17. C. E. Monken, G. N. Gill, *J. Biol. Chem.* **255**, 7067 (1980).
18. Material and methods are available as supporting material on *Science Online*.
19. L. J. Ignarro, G. M. Buga, K. S. Wood, R. E. Byrns, G. Chaudhuri, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 9265 (1987).
20. L. J. Ignarro, R. G. Harbison, K. S. Wood, P. J. Kadowitz, *J. Pharmacol. Exp. Ther.* **237**, 893 (1986).
21. R. M. Palmer, A. G. Ferrige, S. Moncada, *Nature* **327**, 524 (1987).
22. H. Miura *et al.*, *Circ. Res.* **92**, e31 (2003).
23. D. Hasdai *et al.*, *Hypertension* **32**, 228 (1998).
24. R. S. Walikonis *et al.*, *J. Neurosci.* **21**, 423 (2001).
25. W. Landgraf, S. Regulla, H. E. Meyer, F. Hofmann, *J. Biol. Chem.* **266**, 16305 (1991).
26. This research was supported by grants from the British Heart Foundation, The Biotechnology and Biological

Sciences Research Council and the Wellcome Trust. We also thank M. Marber and M. Shattock for helpful comments.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1144318/DC1
Materials and Methods
Figs. S1 to S5
References

26 April 2007; accepted 16 July 2007
Published online 23 August 2007;
10.1126/science.1144318
Include this information when citing this paper.

Common Sequence Variants in the *LOXL1* Gene Confer Susceptibility to Exfoliation Glaucoma

Gudmar Thorleifsson,^{1*} Kristinn P. Magnusson,^{1*} Patrick Sulem,^{1*} G. Bragi Walters,¹ Daniel F. Gudbjartsson,¹ Hreinn Stefansson,¹ Thorlakur Jonsson,¹ Adalbjorg Jonasdottir,¹ Aslaug Jonasdottir,¹ Gerdur Stefansdottir,¹ Gisli Masson,¹ Gudmundur A. Hardarson,¹ Hjorvar Petursson,¹ Arsaell Arnarsson,² Mehdi Motallebipour,³ Ola Wallerman,³ Claes Wadelius,³ Jeffrey R. Gulcher,¹ Unnur Thorsteinsdottir,¹ Augustine Kong,¹ Fridbert Jonasson,^{2,4†} Kari Stefansson^{1†}

Glaucoma is a leading cause of irreversible blindness. A genome-wide search yielded multiple single-nucleotide polymorphisms (SNPs) in the 15q24.1 region associated with glaucoma. Further investigation revealed that the association is confined to exfoliation glaucoma (XFG). Two nonsynonymous SNPs in exon 1 of the gene *LOXL1* explain the association, and the data suggest that they confer risk of XFG mainly through exfoliation syndrome (XFS). About 25% of the general population is homozygous for the highest-risk haplotype, and their risk of suffering from XFG is more than 100 times that of individuals carrying only low-risk haplotypes. The population-attributable risk is more than 99%. The product of *LOXL1* catalyzes the formation of elastin fibers found to be a major component of the lesions in XFG.

Glaucoma is the second most common cause of blindness worldwide (1). Its pathophysiology is poorly understood, and there is a compelling need for improved risk assessment and better treatment.

Glaucoma is a heterogeneous group of disorders that share a distinct optic nerve damage. In most populations, open-angle glaucoma (OAG), characterized by painless loss of vision, constitutes the majority of glaucoma cases and is defined as a progressive loss of neuroretinal rim tissue within the optic disk and consequent excavation of the optic disk with corresponding loss of visual field (2, 3). OAG may be divided into primary open-angle glaucoma (POAG) and secondary glaucoma. POAG is without an identifiable cause of aqueous outflow resistance, whereas in secondary glaucoma the outflow re-

sistance is of a known cause and in exfoliation glaucoma (XFG) it is considered to be due to the exfoliative material from which the syndrome derives its name. Exfoliation syndrome (XFS) is characterized by accumulation of abnormal microfibrillar deposits that line the aqueous bathed surfaces of the anterior segment of the eye. The prevalence of XFS increases with age, and a number of studies have pointed to a geographical clustering of XFS, although this condition is found worldwide; reported prevalence rates average about 10 to 20% of the general population over age 60 (4). In the Reykjavik Eye Study (3), 40% of individuals 80 years and older were found to have XFS. XFS is the most common identifiable cause of secondary glaucoma in most populations. A recent study (5) found the 15-year risk of XFS conversion to XFG to be about 60%, which is similar to results of some previous studies. XFG is characterized by rapid progression, high resistance to medical therapy, and a worse prognosis than in POAG (6).

Family history is an important risk factor for both POAG and XFS which, together with ethnic differences in prevalence of POAG, points

to a role of genetic factors in the risk of suffering from these conditions (7). Three genes, *MYOC* (8), *OPTN* (9), and *WDR36* (10), have been found to be mutated among POAG patients. However, mutations in these genes are of moderate frequency and thus explain only a small fraction of the POAG cases (7).

To identify sequence variants that confer risk of glaucoma, we conducted a genome-wide association study on Icelandic patients with glaucoma, using the Illumina Hap300 chip. After quality filtering, 304,250 single-nucleotide polymorphisms (SNPs) were tested for association to glaucoma in a sample of 195 cases and 14,474 population controls [see (11) for a description of study groups]. The results were adjusted for relatedness between individuals and potential population stratification by the method of genomic control (12). Specifically, the chi-square statistics were divided by an adjustment factor of 1.055 [see (11) for quality-control and statistical analysis].

Overall, three SNPs achieved genome-wide significance ($P < 1.6 \times 10^{-7}$, fig. S1) and are all located within a small region in strong linkage disequilibrium on chromosome 15q24.1 (fig. S2). The strongest association with glaucoma was observed with allele T of rs2165241 (Table 1) with an odds ratio (OR) of 2.28 ($P = 2.0 \times 10^{-14}$). Also achieving genome-wide significance are allele C of rs2304719 (OR = 2.07, $P = 1.2 \times 10^{-8}$) and allele A of rs893817 (OR = 1.85, $P = 1.4 \times 10^{-7}$), but they are both substantially correlated with rs2165241 and are no longer significant ($P > 0.05$) after adjusting for the effect of rs2165241.

The 195 glaucoma cases included 90 cases classified as POAG, 75 known XFG cases, and 30 cases without a precise classification. Further analysis showed that the estimated effect of rs2165241 was weak and only marginally significant for POAG (OR = 1.36, $P = 0.040$), but very strong for XFG (OR = 3.40, $P = 4.3 \times 10^{-12}$) (Table 1). To replicate the observed association, we genotyped rs2165241 in Swedish samples including 200 POAG cases, 199 XFG cases, and 198 controls. No association was seen with POAG (OR = 0.83, $P = 0.18$), but association similar to that in the Icelandic samples was observed for XFG (OR = 3.78, $P = 3.1 \times 10^{-17}$). Combining the results from the two sample sets

¹deCODE genetics Inc, 101 Reykjavik, Iceland. ²Medical Faculty, University of Iceland, 101 Reykjavik, Iceland. ³Department of Genetics and Pathology, Uppsala University, Rudbeck Laboratory, Uppsala, Sweden. ⁴Department of Ophthalmology, National University Hospital, 101 Reykjavik, Iceland.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: fridbert@landspitali.is (F.J.); kstefans@decode.is (K.S.)

for XFG using a Mantel-Haenszel model (13) gave an OR of 3.62 ($P = 1.0 \times 10^{-27}$) (Table 1).

To further explore the impact of the variant, we genotyped an additional 55 Icelandic XFS cases without glaucoma. Compared to the controls, the OR is 3.18 ($P = 1.9 \times 10^{-8}$), and the frequency of rs2165241 T in XFS cases without glaucoma is similar to that in XFS cases with glaucoma ($P > 0.5$). These results indicate that the susceptibility variant tagged by rs2165241 T is a major susceptibility variant for XFS and support the notion that the variant confers risk of glaucoma mainly through XFS.

SNP rs2165241 is located in the first intron of the lysyl oxidase-like protein 1 (*LOXLI*) gene. To refine the observed association signal, we identified SNPs that are substantially correlated with rs2165241 ($r^2 > 0.2$) on the basis of the HapMap CEPH Utah (CEU) data and are not part of the Illumina Hap300 chip (table S2). Eight of those SNPs, in addition to the three best SNPs from the genome-wide scan, were successfully genotyped in all the Icelandic and Swedish XFG cases, in all the Swedish controls, and in 647 of the Icelandic controls. Also genotyped were two known nonsynonymous SNPs, rs1048661 (Arg¹⁴¹→Leu, R141L) and rs3825942 (Gly¹⁵³→Asp, G153D), both located in the first exon of *LOXLI*. rs1048661 was identified through the dbSNP database and rs3825942 is a HapMap SNP. Both nonsynonymous SNPs showed strong association with XFG (combining Iceland and Sweden, OR = 2.46, $P = 2.3 \times 10^{-12}$ for allele G of rs1048661, and OR = 20.10, $P = 3.0 \times 10^{-21}$ for allele G of rs3825942) (Table 1). Further analysis revealed that, although rs2165241 ($P = 1.0 \times 10^{-27}$) was more significant than rs1048661 and rs3825942

individually, it was no longer significant ($P = 0.71$) after adjusting for both nonsynonymous SNPs simultaneously (tables S3 to S5); the latter was also true for the other SNPs that we typed. Results from investigating the joint effect of two nonsynonymous SNPs rs1048661 and rs3825942 are summarized in Fig. 1. The two SNPs are in substantial linkage disequilibrium ($D' = 1$), and only three of the four possible haplotypes were detected in our samples. Among the three observed haplotypes, (G, A) had the lowest estimated risk. Combining results from Iceland and Sweden, relative to (G, A), the (G, G) haplotype had an OR of 27.05 ($P = 4.0 \times 10^{-27}$) and the (T, G) haplotype had an OR of 8.90 ($P = 1.6 \times 10^{-8}$). Allele T of the intronic SNP rs2165241 was strongly associated with XFG because it effectively tagged the high-risk haplotype (G,G) ($r^2 = 0.9$). On the basis of a multiplicative model for the risks of the two risk alleles, allele G of rs1048661 has a relative risk of 3.04 = 27.05/8.90 compared to allele T, and allele G of rs3825942 has a relative risk of 27.05 compared to allele A. Notably, the haplotype (T, A) that was not seen in our samples would be predicted to have an even lower risk than (G, A). The three observed haplotypes did not show deviation from Hardy-Weinberg equilibrium in either the cases or the controls, which is consistent with the model that the risks of the two haplotypes carried by an individual multiply. Under this model, the risk of individuals carrying two copies of the high-risk haplotype (G, G) would be about 700 times the risk of those carrying two copies of (G, A) and about 2.47 times the population average risk. If the risk of the two higher-risk haplotypes, (G, G) and (T, G), could be reduced to

that of (G, A), it would eliminate more than 99% of the XFG cases. Hence, the population attributable risk of the two higher-risk haplotypes is more than 99%. Sequencing of the seven exons of *LOXLI* did not identify further variants associated with the disease (table S7).

To determine if the nonsynonymous risk variants could affect the mRNA expression of *LOXLI*, we analyzed *LOXLI* expression in adipose tissue from 659 individuals with genotype data for rs1048661 and rs3825942 [microarray expression data (11)]. *LOXLI* expression was reduced by an estimated 7.7% with each copy carried of the risk G allele of rs1048661 ($P = 8.3 \times 10^{-7}$); this effect was significant for both sexes and did not change if the expression was adjusted for the weight of the individuals (Fig. 2). In contrast, weak positive correlation was observed between the risk G allele of rs3825942 and expression of *LOXLI* ($P = 0.034$), and this effect disappeared completely when the correlation was adjusted for the effect of rs1048661 ($P = 0.55$). The result from the microarray expression data was confirmed with real-time polymerase chain reaction for a subset of 564 of the 659 individuals (fig. S3).

The *LOXLI* gene is a member of the lysyl oxidase family of proteins that catalyzes oxidative deamination of lysine residues of tropoelastin, which leads to their spontaneous cross-linking with consequential formation of elastin polymer fibers (14, 15). Elastogenesis also requires fibrillin-containing microfibrils that act as scaffolds that guide the cross-linking process and deposition of elastine (16). The lysyl oxidase family has five members, and these encode the prototypic LOX protein and LOX-like proteins LOXL1 to LOXL4. All five LOX family members have a

Table 1. Association between POAG, XFG, and XFS and rs2165241, rs1048661, and rs3825942. The association of the risk alleles of the SNP rs2165241, located in the first intron of *LOXLI*, and of the two nonsynonymous SNPs rs1048661 (R141L) and rs3825942 (G153D) with glaucoma in the Icelandic discovery case-control group, the Swedish replication case-control group, and the two groups combined. Results are shown for all glaucoma cases

and for POAG cases, XFG, and exfoliation without glaucoma separately. Study population includes the number of individuals (*n*). The results include the OR, 95% confidence intervals (CI), and *P* values assuming the multiplicative model. For the Icelandic case-control group, the *P* values and CI were adjusted for relatedness as described in the methods (11). For the combined group, we calculated OR and *P* values using a Mantel-Haenszel model.

Study population (<i>n</i>)	rs2165241 T			rs1048661 (R141L) G			rs3825942 (G153D) G		
	Frq.	OR (95% CI)	<i>P</i>	Frq.	OR (95% CI)	<i>P</i>	Frq.	OR (95% CI)	<i>P</i>
Iceland									
Controls (14,474)	0.473			0.651			0.847		
Glaucoma combined (195)	0.672	2.28 (1.85–2.82)	2.0×10^{-14}	0.777	1.87 (1.49–2.35)	7.4×10^{-8}	0.936	2.66 (1.86–3.80)	7.9×10^{-8}
POAG (90)	0.550	1.36 (1.01–1.83)	0.04	0.711	1.32 (0.96–1.82)	0.085	0.872	1.25 (0.81–1.91)	0.32
XFG (75)	0.753	3.40 (2.41–4.81)	4.3×10^{-12}	0.827	2.56 (1.74–3.77)	1.8×10^{-6}	0.987	13.23 (5.59–31.29)	4.1×10^{-9}
XFS no glaucoma (55)	0.740	3.18 (2.12–4.76)	1.9×10^{-8}	0.789	2.02 (1.32–3.09)	1.3×10^{-3}	0.982	10.10 (4.02–25.36)	8.5×10^{-7}
Sweden									
Controls (198)	0.535			0.682			0.879		
Glaucoma combined (399)	0.649	1.61 (1.26–2.05)	0.00016	0.737	1.31 (1.00–1.70)	0.048	0.929	1.79 (1.19–2.70)	0.0052
POAG (200)	0.488	0.83 (0.63–1.09)	0.18	0.638	0.82 (0.61–1.10)	0.19	0.863	0.87 (0.57–1.31)	0.49
XFG (199)	0.813	3.78 (2.77–5.14)	3.1×10^{-17}	0.834	2.39 (1.72–3.34)	2.7×10^{-7}	0.995	27.28 (11.44–65.07)	9.1×10^{-14}
Combined									
Controls (14,672)									
Glaucoma combined (594)		1.96 (1.67–2.29)	1.3×10^{-16}		1.59 (1.35–1.89)	7.5×10^{-8}		2.20 (1.69–2.85)	3.4×10^{-9}
POAG (290)		1.04 (0.85–1.28)	0.67		1.02 (0.83–1.27)	0.83		1.04 (0.78–1.39)	0.81
XFG (274)		3.62 (2.87–4.55)	1.0×10^{-27}		2.46 (1.91–3.16)	2.3×10^{-12}		20.10 (10.80–37.41)	3.0×10^{-21}

similar exon structure consisting of seven exons, five of which (exons 2 to 6) exhibit strong homology and encode the C-terminal catalytic domain of these proteins. The sequence difference between the *LOX* genes resides mainly in exon 1, which encodes pro-peptide that is, after the attachment of *LOXLI* to the scaffolding structure, cleaved off for catalytic activation of the enzyme. Several studies have demonstrated that the *LOXLI* pro-peptide binds to both tropoelastine and fibulin-5 and that these interactions are essential for directing the deposition of the enzyme onto elastic fibers (14, 16).

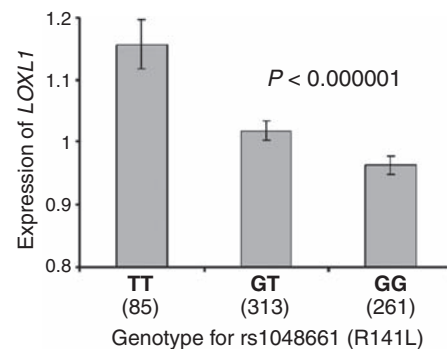
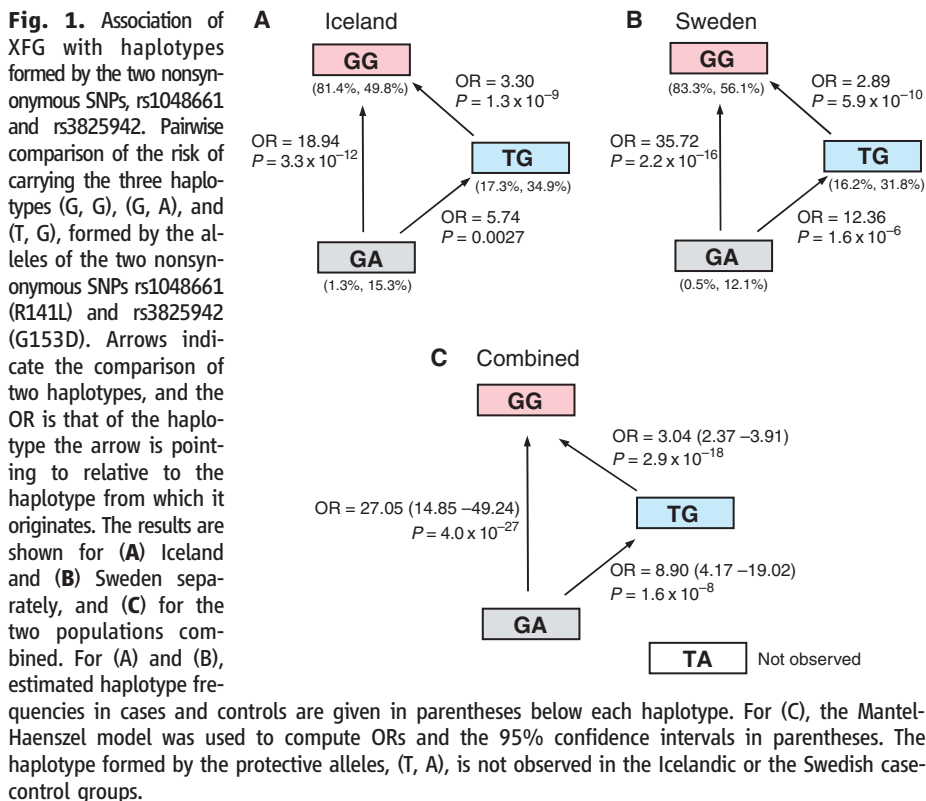
The pathology of XFS is characterized by chronic accumulation of abnormal fibrillar material in the anterior segment of the eye, leading to numerous clinical complications apart from secondary glaucoma development. From analysis of the XFS material, it has been proposed that XFS arises from abnormal aggregation of elastin microfibrillar components (elastic microfibrilopathy) produced by various intraocular cell types (6, 17). Although a role for *LOXLI* in the formation of the extracellular matrix of the eye has not been documented, *LOXLI* expression is detected in ocular tissues such as lamina cribrosa, lens epithelium, cornea, ciliary muscle, and trabecular meshwork, all of which may be involved in extracellular matrix formation (18–20) [data accessible at NCBI GEO database (11)]. We demonstrate here the association of two coding SNPs, rs1048661 and rs3825942, with XFG that leads to an amino acid change at position 141 (Arg→Leu) and 153 (Gly→Asp), respectively, both of which are located in the

N-terminal pro-peptide. Based on the functional role of the pro-peptide, these alterations could affect both the catalytic activity of the protein through modifications of pro-peptide cleavage and the binding to substrates like tropoelastine and fibulin-5. In addition, we demonstrate that the risk allele of rs1048661 associates with lower expression levels of the *LOXLI* mRNA in adipose tissue. This effect could be mediated through its linkage disequilibrium to noncoding regulatory elements or through its own effect on mRNA stability or processing, as previously documented for both synonymous and nonsynonymous coding mutations in genes such as *DRD1*, *MDRI*, and *OPRM1* (21–23). Assuming a similar regulatory network for *LOXLI* expression in adipose and ocular tissues, these data suggest that low levels of *LOXLI* expression could predispose to XFS.

Ocular tissue was not available to us to study the effect of the risk alleles on the expression of *LOXLI*, and we considered it unlikely that we could obtain such tissue from a sufficiently large number of individuals to do a meaningful study. Our assumption was that it would be difficult to predict what tissue would best reflect ocular tissue in this respect and that any tissue expressing the gene in an easily detectable amount would serve our purpose as well as other such tissues. *LOXLI* is expressed at very low levels in the blood, and thus we were unable to determine whether the risk variants affect its expression. Hence, we analyzed RNA from adipose tissue because samples from several hundred individuals were available to us. In adipose tissue, the

expression of *LOXLI* is decreased by 7.7% per risk allele of rs1048661 (R141L), which is a small change and not necessarily biologically meaningful, although in a late-onset disease it could be relevant. It is, however, notable that the risk allele of rs3825942 (G153D), the variant that confers the greater risk, has no effect on *LOXLI* expression.

In summary, we have shown in two independent study groups that two nonsynonymous changes in exon 1 of the *LOXLI* gene on chromosome 15q24.1 confer risk to XFG, possibly through XFS. In Iceland and Sweden, the high-risk haplotype is very common with an average frequency of about 50% in the general population. About 25% of individuals in the general population are homozygous for the haplotype with the highest risk, and their risk of suffering from XFG is estimated to be about 700 times that of individuals carrying only the low-risk haplotype, or about 2.47 times that of the population average. Jointly, the two nonsynonymous changes account for more than 99% of all XFG cases. The product of the *LOXLI* gene modifies elastin fibers that are a major constituent of the intraocular lesions in XFG. As to other forms of glaucoma, after removing the SNPs in the *LOXLI* region, the genome-scan Q-Q plots for POAG and glau-



coma overall cannot be distinguished from that resulting from random noise (fig. S1b), which suggests that POAG may be a more complex disease than XFG.

References and Notes

1. S. Resnikoff *et al.*, *Bull. World Health Org.* **82**, 844 (2004).
2. P. J. Foster, R. Buhrmann, H. A. Quigley, G. J. Johnson, *Br. J. Ophthalmol.* **86**, 238 (2002).
3. F. Jonasson *et al.*, *Eye* **17**, 747 (2003).
4. A. Ringvold, *Acta Ophthalmol. Scand.* **77**, 371 (1999).
5. S. M. Jeng *et al.*, *J. Glaucoma* **16**, 117 (2007).
6. U. Schlotzer-Schrehardt, G. O. Naumann, *Am. J. Ophthalmol.* **141**, 921 (2006).
7. A. W. Hewitt, J. E. Craig, D. A. Mackey, *Clin. Exp. Ophthalmol.* **34**, 472 (2006).
8. E. M. Stone *et al.*, *Science* **275**, 668 (1997).
9. T. Rezaie *et al.*, *Science* **295**, 1077 (2002).
10. S. Monemi *et al.*, *Hum. Mol. Genet.* **14**, 725 (2005).
11. Materials and methods are available as supporting material on Science Online.
12. B. Devlin, K. Roeder, *Biometrics* **55**, 997 (1999).
13. N. Mantel, W. Haenszel, *J. Natl. Cancer Inst.* **22**, 719 (1959).
14. X. Liu *et al.*, *Nat. Genet.* **36**, 178 (2004).
15. H. A. Lucero, H. M. Kagan, *Cell. Mol. Life Sci.* **63**, 2304 (2006).
16. L. Thomassin *et al.*, *J. Biol. Chem.* **280**, 42848 (2005).
17. R. Ritch, U. Schlotzer-Schrehardt, A. G. Konstas, *Prog. Retin. Eye Res.* **22**, 253 (2003).
18. R. P. Kirwan *et al.*, *Mol. Vis.* **11**, 798 (2005).
19. P. A. Netland, H. Ye, B. W. Streeten, M. R. Hernandez, *Ophthalmology* **102**, 878 (1995).
20. J. D. Pena *et al.*, *Exp. Eye Res.* **67**, 517 (1998).
21. J. Duan *et al.*, *Hum. Mol. Genet.* **12**, 205 (2003).
22. D. Wang, A. D. Johnson, A. C. Papp, D. L. Kroetz, W. Sadée, *Pharmacogenet. Genomics* **15**, 693 (2005).
23. H. Zhang *et al.*, *Hum. Mol. Genet.* **15**, 807 (2006).
24. We thank the participants whose contribution made this study possible. We also thank the nurses at Noatun (deCODE's sample recruitment center) and personnel at the deCODE core facilities for their hardwork and enthusiasm. C. W. was supported by the Swedish Research Council. The authors from deCODE genetics declare competing financial interests.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1146554/DC1

Materials and Methods

Figs. S1 to S3

Tables S1 to S7

References

15 June 2007, accepted 18 July 2007

Published online 9 August 2007;

10.1126/science.1146554

Include this information when citing this paper.

The *Fusarium graminearum* Genome Reveals a Link Between Localized Polymorphism and Pathogen Specialization

Christina A. Cuomo,¹ Ulrich Güldener,^{2,3} Jin-Rong Xu,⁴ Frances Trail,⁵ B. Gillian Turgeon,⁶ Antonio Di Pietro,⁷ Jonathan D. Walton,⁵ Li-Jun Ma,¹ Scott E. Baker,⁸ Martijn Rep,⁹ Gerhard Adam,¹⁰ John Antoniw,¹¹ Thomas Baldwin,¹¹ Sarah Calvo,¹ Yueh-Long Chang,¹² David DeCaprio,¹ Liane R. Gale,¹² Sante Gnerre,¹ Rubella S. Goswami,¹² Kim Hammond-Kosack,¹¹ Linda J. Harris,¹³ Karen Hilburn,¹⁴ John C. Kennell,¹⁵ Scott Kroken,¹⁶ Jon K. Magnuson,⁸ Gertrud Mannhaupt,³ Evan Mauceli,¹ Hans-Werner Mewes,^{2,3} Rudolf Mitterbauer,¹⁰ Gary Muehlbauer,¹² Martin Münsterkötter,³ David Nelson,¹⁷ Kerry O'Donnell,¹⁸ Thérèse Ouellet,¹³ Weihong Qi,⁵ Hadi Quesneville,¹⁹ M. Isabel G. Roncero,⁷ Kye-Yong Seong,¹² Igor V. Tetko,^{3,21} Martin Urban,¹¹ Cees Waalwijk,²⁰ Todd J. Ward,¹⁸ Jiqiang Yao,⁴ Bruce W. Birren,¹ H. Corby Kistler^{12,14*}

We sequenced and annotated the genome of the filamentous fungus *Fusarium graminearum*, a major pathogen of cultivated cereals. Very few repetitive sequences were detected, and the process of repeat-induced point mutation, in which duplicated sequences are subject to extensive mutation, may partially account for the reduced repeat content and apparent low number of paralogous (ancestrally duplicated) genes. A second strain of *F. graminearum* contained more than 10,000 single-nucleotide polymorphisms, which were frequently located near telomeres and within other discrete chromosomal segments. Many highly polymorphic regions contained sets of genes implicated in plant-fungus interactions and were unusually divergent, with higher rates of recombination. These regions of genome innovation may result from selection due to interactions of *F. graminearum* with its plant hosts.

Fusarium, a genus of plant pathogenic fungi, causes diseases that affect most species of cultivated plants, including root and stem rots, blights, and wilts (*1*). *F. graminearum*, which causes Fusarium head blight (FHB) disease on wheat and barley, is a leading cause of economic loss in these crops (*2*). In addition to reducing seed mass and quality, the fungus contaminates grain with toxic metabolites that are a threat to human health (*3*). *Fusarium* species also can directly infect humans, causing localized necrotic diseases (*4*) and invasive infection, especially in immunocompromised individuals (*5*).

The *F. graminearum* genome was whole-genome shotgun sequenced by paired-end se-

quencing of plasmid, Fosmid, and bacterial artificial chromosome (BAC) clones. The resulting assembly totals 36.1 Mb and displays high sequence quality and continuity. Nearly all (99.8%) of the assembly was anchored to the four chromosomes by genetically mapping markers derived from the genome sequence (*6*), and an initial set of 11,640 genes was predicted (table S1 and SOM text). Functional categories for the predicted genes were inferred by the presence of conserved InterPro domains (*7*) and were compared with those found in genomes of the related fungi, *Neurospora crassa*, *Magnaporthe grisea* and *Aspergillus nidulans*. The *F. graminearum* genome has greater numbers of genes for several protein categories, including predicted transcrip-

tion factors, hydrolytic enzymes, and transmembrane transporters (Fig. 1 and table S2).

The *F. graminearum* genome has few high-identity duplicated sequences, fewer by at least a factor of 15 than other related fungi, including *Saccharomyces cerevisiae* (table S3 and SOM text). Only a few gene pairs originated from recent duplications (fig. S1 and table S4), and we identified only two small families of transposons (table S5 and SOM text). *F. graminearum* differs from other filamentous fungi because it is homothallic (self-fertile) and rarely out-crosses, which limits the opportunity to acquire new repeats (*2*). In some ascomycetous fungi, including *F. graminearum*, the lack of repetitive sequence is due to a genome-wide defense system known as repeat-induced point mutation (RIP) (*8*). RIP identifies duplicated sequences (*9*) and introduces C:G to T:A transition mutations in both copies during the sexual cycle; this mutational bias was observed in *F. graminearum* transposons (tables S6 and S7 and SOM text).

¹Broad Institute of the Massachusetts Institute of Technology and Harvard, Cambridge, MA 02142, USA. ²Technische Universität München, Freising-Weihenstephan, Germany. ³Institute for Bioinformatics, GSF National Research Center for Environment and Health, Neuherberg, Germany. ⁴Purdue University, West Lafayette, IN 47907, USA. ⁵Michigan State University, East Lansing, MI 48824, USA. ⁶Cornell University, Ithaca, NY 14853, USA. ⁷Universidad de Córdoba, Córdoba, Spain. ⁸Pacific Northwest National Laboratory, Richland, WA 99352, USA. ⁹University of Amsterdam, Netherlands. ¹⁰BOU, University of Natural Resources and Applied Life Sciences, Vienna, Austria. ¹¹Rothamsted Research, Harpenden, UK. ¹²University of Minnesota, St. Paul, MN 55108, USA. ¹³Agriculture and Agri-Food Canada and University of Ottawa, Ottawa, ON, Canada. ¹⁴U.S. Department of Agriculture (USDA) Agricultural Research Service, Cereal Disease Laboratory, St. Paul, MN 55108, USA. ¹⁵St. Louis University, St. Louis, MO 63103, USA. ¹⁶University of Arizona, Tucson, AZ 85721, USA. ¹⁷University of Tennessee, Memphis, TN 38163, USA. ¹⁸USDA ARS, National Center for Agricultural Utilization Research, Peoria, IL 61604, USA. ¹⁹Institut Jacques Monod, Paris, France. ²⁰Plant Research International, Wageningen, Netherlands. ²¹Institute of Bioorganic Chemistry and Photochemistry, National Ukrainian Academy of Sciences, Kiev, Ukraine.

*To whom correspondence should be addressed. E-mail: hckist@umn.edu