Bacterial artificial chromosome–based physical map of *Gibberella zeae* (*Fusarium graminearum*)

Yueh-Long Chang, Seungho Cho, H. Corby Kistler, Chun-Sheng Hsieh, and Gary J. Muehlbauer

Abstract: *Fusarium graminearum* is the primary causal pathogen of Fusarium head blight of wheat and barley. To accelerate genomic analysis of *F. graminearum*, we developed a bacterial artificial chromosome (BAC)–based physical map and integrated it with the genome sequence and genetic map. One BAC library, developed in the *Hin*dIII restriction enzyme site, consists of 4608 clones with an insert size of approximately 107 kb and covers about 13.5 genome equivalents. The other library, developed in the *Bam*HI restriction enzyme site, consists of 3072 clones with an insert size of approximately 95 kb and covers about 8.0 genome equivalents. We fingerprinted 2688 clones from the *Hin*dIII library and 1536 clones from the *Bam*HI library and developed a physical map of *F. graminearum* consisting of 26 contigs covering 39.2 Mb. Comparison of our map with the *F. graminearum* genome sequence showed that the size of our physical map is equivalent to the 36.1 Mb of the genome sequence. We used 31 sequence-based genetic markers, randomly spaced throughout the genome, to integrate the physical map with the genetic map. We also end-sequenced 17 *Bam*HI BAC clones and identified 4 clones that spanned gaps in the genome sequence. Our new integrated map is highly reliable and useful for a variety of genomics studies.

Key words: Fusarium graminearum, Fusarium head blight, bacterial artificial chromosome, physical map.

Résumé : Le *Fusarium graminearum* est le principal agent pathogène causant la fusariose de l'épi chez le blé et l'orge. Afin d'accélérer l'analyse génomique du *F. graminearum*, les auteurs ont développé une carte physique à l'aide de clones BAC (chromosomes bactériens artificiels) et ont intégré celle-ci avec la séquence génomique et la carte génétique. Une banque de BAC, dont les inserts ont été clonés dans le site de restriction *Hin*dIII et mesurent environ 107 kpb, totalise 4608 clones ce qui correspond à environ 13,5 génomes. Une autre banque, dont les inserts ont été clonés dans le site de restriction *Bam*HI et mesurent environ 95 kpb, compte 3072 clones et offre une couverture équivalant à 8,0 génomes. Les auteurs ont produit des empreintes pour 2688 des clones de la banque *Hin*dIII et pour 1536 clones de la banque *Bam*HI. Ils ont ensuite assemblé une carte physique du *F. graminearum* comprenant 26 contigs totalisant 39,2 Mpb. Une comparaison de cette carte avec la séquence génomique du *F. graminearum* a montré que la taille de la carte physique est proche des 36,1 Mpb de la séquence génomique. Trente et un marqueurs génétiques de séquence connue, aléatoirement distribués sur le génome, ont été employés pour aligner les cartes physique et génétique. Les auteurs ont également séquencé les extrémités de 17 clones *Bam*HI et identifié 4 clones qui chevauchaient des discontinuités dans la séquence génomique. La nouvelle carte intégrée s'avère très fiable et utile pour toute une gamme d'études génomiques.

Mots-clés : Fusarium graminearum, fusariose de l'épi, chromosome bactérien artificiel, carte physique.

[Traduit par la Rédaction]

Introduction

Gibberella zeae (anamorph *Fusarium graminearum*) is the primary casual agent of Fusarium head blight (FHB) of wheat and barley. FHB poses a major problem for barley and wheat grown in humid and semi-humid climates (McMullen et al. 1997). For example, in the US between 1993 and 2001, FHB caused an estimated $8 billion in economic losses (Nganje et al. 2004). *Fusarium graminearum* infects spike tissue and causes premature necrosis and bleaching, resulting in a reduction in grain yield and quality (Schroeder and Christensen 1963; McMullen et al. 1997). Trichothecene mycotoxins produced by *F. graminearum*, such as deoxynivalenol, accumulate in infected tissues and reduce grain quality (Sutton 1982; Tuite et al. 1990; Bai and Shaner 1994). To date, there are no effective means to control this disease.

The development of genetic and genomics resources will greatly accelerate the understanding of this important fungal pathogen. *Fusarium graminearum* is a homothallic ascomy-

Received 13 March 2007. Accepted 29 July 2007. Published on the NRC Research Press Web site at genome.nrc.ca on 17 October 2007.

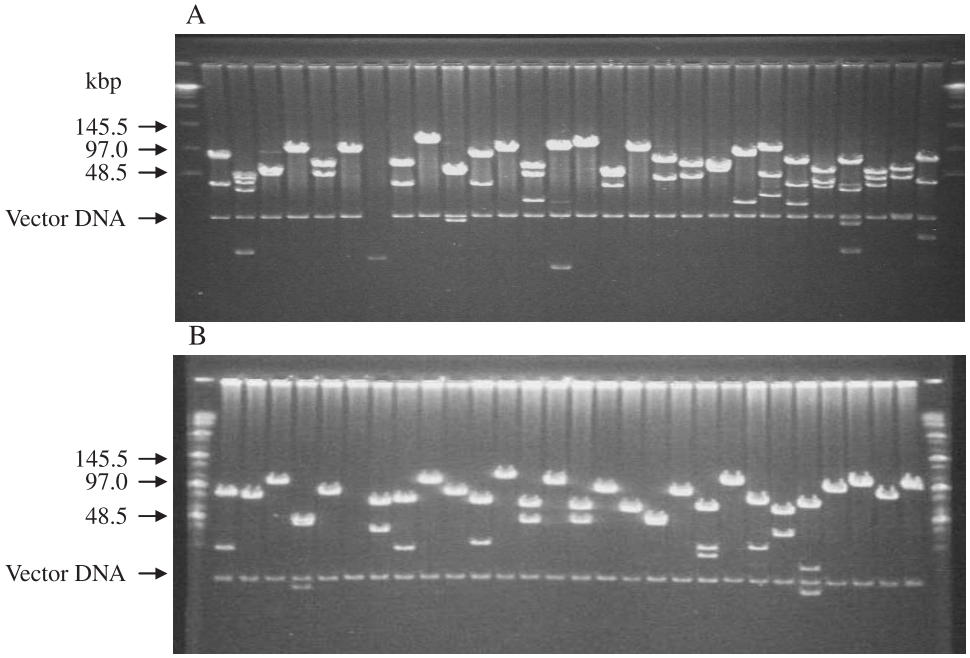
Corresponding Editor: F. Belzile.

Y.-L. Chang and C.-S. Hsieh. Institute of Agricultural Biotechnology, National Chiayi University, Chiayi 600, Taiwan.
S. Cho and G.J. Muehlbauer.[1] Department of Agronomy and Plant Genetics, University of Minnesota, St. Paul, MN 55108, USA.
H.C. Kistler. USDA-ARS, Cereal Disease Laboratory, Department of Plant Pathology, University of Minnesota, St. Paul, MN 55108, USA.

[1]Corresponding author (e-mail: muehl003@umn.edu).

 doi:10.1139/G07-079 © 2007 NRC Canada

Fig. 1. Insert sizes of randomly picked clones from the *Hin*dIII BAC library (A) and the *Bam*HI BAC library (B) of *Fusarium graminearum*.



cete (Nelson et al. 1983; Yun et al. 2000) that can be induced to cross in laboratory conditions (Bowden and Leslie 1999). The fungus is part of a species complex that can be partitioned into at least 9 phylogenetically defined lineages (O'Donnell et al. 2000; Ward et al. 2002; O'Donnell et al. 2004). In North America and Europe, the species *F. graminearum sensu stricto* (formerly known as "lineage 7") is the predominant form. *Fusarium graminearum* genomics resources consist of a full genome sequence assembly (Cuomo et al. 2007; Broad Institute 2007), approximately 12 000 expressed sequence tags (Trail et al. 2003), a genome sequence–based genetic map (Gale et al. 2005), and a Fusarium Affymetrix GeneChip (Güldener et al. 2006*a*). However, large-insert genomic DNA libraries in the form of bacterial artificial chromosomes (BACs) have not been described for *F. graminearum*. In addition, a BAC-based physical map of *F. graminearum* integrated with the genome sequence and the genetic map has not been available.

Large-insert libraries are essential tools for characterizing genomes. BAC libraries are the technology of choice because they are simple to develop, easy to use, and have a low frequency of chimeric clones. BAC libraries of *F. graminearum* also can be used for genome organization studies, for genome sequence finishing, to positionally clone genes known only by phenotype, and for comparative genome analysis within *F. graminearum* and between different *Fusarium* species or other fungal species. BAC libraries have been developed for a variety of fungal organisms including *Magnaporthe grisea* (Martin et al. 2002) and *Neurospora crassa* (Aign et al. 2001). Physical maps have been developed for a few fungal species including *Cryptococcus neoformans* (Schein et al. 2002), *M. grisea* (Thon et al. 2004), *Trichoderma reesei* (Diener et al. 2004), *Ustilago maydis* (Meksem et al. 2005), and *Penicillium chrysogenum* (Xu et al. 2005). For *Phytophthora sojae*, a BAC-based physical map has been integrated with the genome sequence (Zhang et al. 2006).

To enhance the genomics resources for *F. graminearum,* we developed 2 BAC libraries and a BAC-based physical map of *F. graminearum*, integrated the physical map with the genetic map and genome sequence, and showed that the physical map can be used to fill in gaps in the genome sequence. The BAC libraries represent the first publicly available large-insert libraries of *F. graminearum*. This integrated resource will provide the ability to efficiently access and study genes involved in the pathogenicity and virulence of the fungus, to conduct comparative genomics, and to study the ecological genetics of *F. graminearum*.

Materials and methods

Fusarium graminearum strain

Fusarium graminearum s. str. (*Gibberella zeae*) strain PH-1 (NRRL31084; Trail and Common 2000) was used for BAC library construction. PH-1, the strain previously used for genome sequencing, is highly fertile, produces high amounts of the mycotoxins deoxynivalenol and zearelanone, sporulates abundantly in pure culture, and is highly pathogenic to wheat, barley, and rice.

Fusarium graminearum large genomic DNA isolation and BAC library construction

Genomic DNA for construction of BAC libraries was isolated from fungal protoplasts prepared as previously described (Boehm et al. 1994). The libraries were constructed in the *Hin*dIII site of pBeloBAC11 (Kim et al. 1996; Zhang et al. 1996) and the *Bam*HI site of pECBAC1 (Frijters et al. 1997). These BAC libraries are permanently maintained in 384-well plates and are publicly available from the Fungal Genetics Stock Center at the University of Missouri, Kansas City (http://www.fgsc.net/).

BAC fingerprinting and physical map development

BAC DNA was isolated from *Hin*dIII clones (Nos. 1 to

© 2007 NRC Canada

2688 in the map) and *Bam*HI clones (Nos. 2689 to 4224) following standard procedures. The isolated BAC DNA was double-digested with *Hin*dIII and *Hae*III and end-labelled with [^{33}P]dATP using reverse transcriptase at 37 °C for 2 h. The labelled fragments were separated on 3.5% (*w*/*v*) polyacrylamide denaturing sequencing gels (Chang et al. 2001). Lambda DNA digested with *Sau*3AI and labelled with [^{33}P]dATP was used as a marker in the 1st and every 7th lane. The gels were dried and used to expose autoradiographic film overnight for development. The fingerprints were processed using a PowerLook 2100XL scanner and edited using Image 4.0 of the Fingerprinted Contig (FPC) package (Soderlund et al. 2000). All obvious DNA fragments from each fingerprint, distributed from the top to the bottom of the sequencing gel, were scored except for vector DNA and faint bands.

The fingerprints were used to assemble the clones into contigs using the FPC program (Soderlund et al. 2000). A 2-step process was used to construct the contigs. First, we assembled all contigs using FPC version 7.0 at a high stringency (cutoff = 10^{-12} and tolerance = 7). Next, we merged the contigs to each other and with any singletons using a lower comparison stringency (cutoff = 10^{-4} and tolerance = 7).

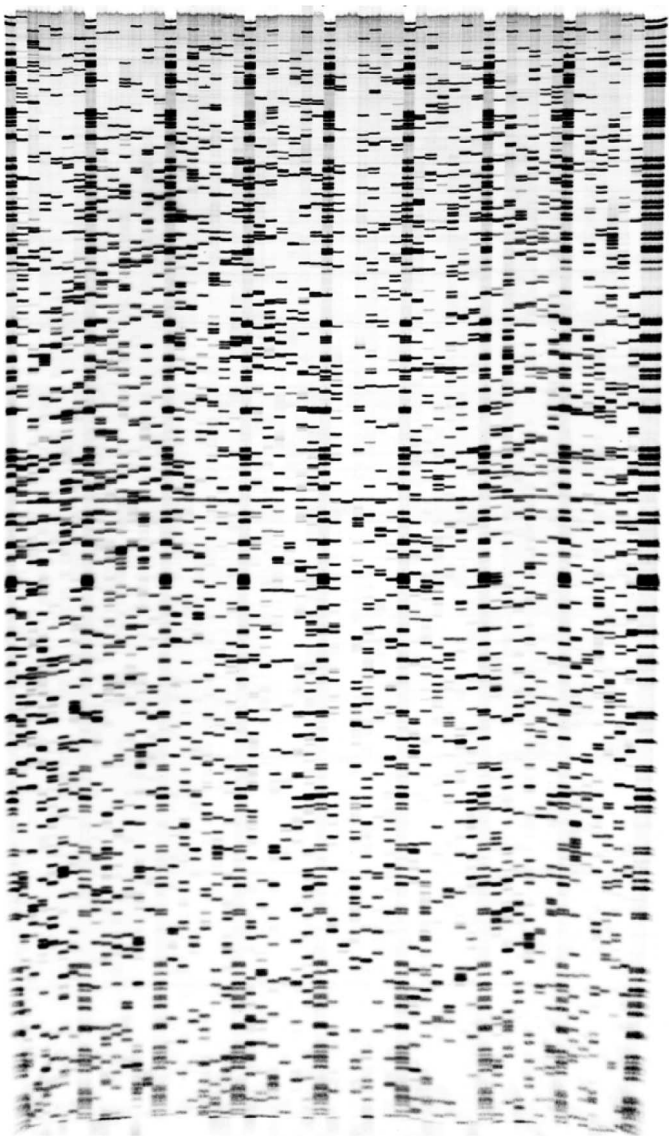
Library screening with molecular markers

Thirty-one mapped sequence-based markers that span the *F. graminearum* genome were used for PCR screening of the *Hin*dIII and *Bam*HI libraries. These markers are HK245, HK339, HK375, HK623, HK661, HK673, HK679, HK693, HK731, HK757, HK863, HK879, HK907, HK919, HK929, HK931, HK933, HK961, HK1009, HK1011, HK1015, HK1021, HK1039, HK1043, HK1045, HK1057, HK1059, HK1063, HK1097, HK1115, and HK1139. Map locations and sequences for each of the markers can be found in the MIPS *Fusarium graminearum* Genome Database, available from http://mips.gsf.de/genre/proj/fusarium/ (Güldener et al. 2006*b*; Gale et al. 2005). Each library was pooled, using 5 μL of bacteria from each BAC clone. Twelve 384-well *Hin*dIII BAC library plates were pooled to four 96-well plates by combining 12 clones located at the same position on each 384-well plate. Eight 384-well *Bam*HI library plates were pooled to four 96-well plates by combining 8 clones located at the same position on each 384-well plate. BAC pools were PCR-amplified with primers specific to each genetic marker, and amplified BAC pools were identified and deconvoluted to identify the specific amplified BAC clone. BAC clones were confirmed by conducting PCR on each individual BAC clone.

Large-scale comparison of our physical map with the genome sequence map

The *Hin*dIII BAC library was used both in end sequencing for the construction of the genome sequence map (Cuomo et al. 2007; Broad Institute 2007) and in fingerprinting for the development of our physical map. Therefore, *Hin*dIII BAC clones provide a platform for large-scale comparison between the physical map and the genome sequence map. Owing to the different naming systems used in the two maps, a conversion table was established for the large-scale comparison (Table S1[2]).

Fig. 2. Sequencing gel–based fingerprints of BAC clones. The 1st and every 7th lane are DNA markers and the others are BAC DNAs.



End sequencing of BAC clones

*Bam*HI BAC clones were selected from the borders of some contigs and grown in LB broth with the appropriate concentration of antibiotics overnight. BAC DNA was isolated using the conventional alkaline lysis method and phenol–chloroform extraction. According to the manufacturer's instructions, purified BAC DNA was used in end sequencing by the Applied Biosystems 3130 sequencer. End sequences were used to conduct BLASTN searches of the *F. graminearum* genome sequence (Cuomo et al. 2007; Broad Institute 2007).

[2] Supplementary data for this article are available on the journal Web site (http://genome.nrc.ca) or may be purchased from the Depository of Unpublished Data, Document Delivery, CISTI, National Research Council Canada, Building M-55, 1200 Montreal Road, Ottawa, ON K1A 0R6, Canada. DUD 5225. For more information on obtaining material refer to http://cisti-icist.nrc-cnrc.gc.ca/irm/unpub_e.shtml.

© 2007 NRC Canada

Table 1. Summary of completed contigs.

Contig	No. of clones	Length (kb)	Mapped marker(s)
101	71	838.77	HK929
102	737	7426.29	HK907, HK623, HK1043, HK1045, HK931, HK919, HK693, HK933
103	138	1543.95	HK961
201	132	1800.18	HK731
202	15	234.33	
203	11	170.82	
204	17	177.39	HK661
205	3	105.12	
206	8	179.58	
207	155	1817.70	HK1057, HK757
208	135	1624.98	HK1115
301	607	5893.29	HK1059, HK375, HK1039, HK1015, HK679, HK339
401	307	2895.18	HK1063, HK1097, HK1009
402	159	1500.15	HK1011
403	57	775.26	
501	62	744.60	
502	326	2838.24	HK245, HK879, HK1139
601	15	324.12	
602	303	2831.67	HK1021, HK673
701	242	2538.21	HK863
702	19	293.46	
801	153	1609.65	
901	39	602.25	
1001	5	219.00	
1	5	122.64	
2	2	137.97	

Display of the physical map

The WebFPC program (Pampanwar et al. 2005) was adopted to display our individual contigs at http://140.130.98.33/fpc/WebAGCoL/fusa/WebFPC and is useful to search any specific marker or BAC clone in the physical map. To view the map on a personal computer, the Java Runtime Environment software is required and can be downloaded from http://www.java.com/en/download/manual.jsp.

Results and discussion

Fusarium graminearum BAC libraries

Two BAC libraries were constructed, one with the *Hin*dIII restriction enzyme and the other with the *Bam*HI restriction enzyme. The *Hin*dIII library contains 4608 clones arrayed in twelve 384-well microtitre plates. *Hin*dIII restriction enzyme digestion and electrophoresis of 125 randomly selected clones from the *Hin*dIII library indicated that the insert sizes ranged from 20 to 195 kb with an average insert size of 107 kb (Fig. 1A). The *Bam*HI library contains 3072 clones arrayed in eight 384-well microtitre plates. *Bam*HI restriction enzyme digestion and electrophoresis of 54 randomly selected clones from the *Bam*HI library indicated that the insert sizes ranged from 75 to 115 kb with an average insert size of 95 kb (Fig. 1B). Combined, the average insert size for the two BAC libraries is approximately 103 kb. Based on the genome sequence, the size of the *F. graminearum* genome is approximately 36.1 Mb (Cuomo et al. 2007; Broad Institute 2007). Therefore, the *Hin*dIII and *Bam*HI libraries cover an estimated 13.5 (4608 clones × 107 kb/36.5 Mb) and 8.0 (3072 clones × 95 kb/36.5 Mb) genome equivalents, respectively, with a combined coverage of 21.5 genome equivalents.

BAC library fingerprinting and initial physical map development

To generate BAC fingerprints for developing a physical map, we used a sequencing gel–based, radioactive nucleotide method (Chang et al. 2001; see Materials and methods). We randomly picked and fingerprinted 4224 BAC clones (2688 *Hin*dIII and 1536 *Bam*HI). Based on the average of insert sizes in both libraries, these clones exhibit approximately 11.9-fold haploid genome coverage. An example of the BAC clone fingerprints is shown in Fig. 2. According to the results derived from data analysis using the FPC program, an average of 47 bands per BAC clone were obtained.

To develop contigs, each fingerprint was scanned into an image file, edited, and imported into the FPC database (Soderlund et al. 2000). After editing, 318 fingerprints were deleted owing to no BAC DNA, few bands per clone, or ambiguous patterns. Using a tolerance of 7 and a cutoff of 10^{-12}, 3906 BAC clones were assembled into 112 initial contigs and 183 singletons. Thus, the initial contigs were derived from 3723 BACs and comprised a total of 19 850 bands. According to the average insert size of approximately 103 kb from the two BAC libraries and an average of 47 bands for each clone, each band represents approximately 2.19 kb (103 kb/47 bands). The size distribution of the initial 112 contigs over a total physical distance of 43.5 Mb (19 850 bands × 2.19 kb/band) is in the range of 99 to 1643 kb. The average size of each contig was 388 kb.

© 2007 NRC Canada

Table 2. Positive fingerprinted BAC clones identified by sequence-based markers.

Marker	No. of BAC clones	Positive clones (final contig number)
HK245	9	213, 748, 1094, 2168, 2210, 2324, 2352, 3350, 4171 (ctg502)
HK339	12	556, 704, 1086, 1185, 1599, 1787, 3290, 3451, 3649, 3665, 3933, 3981 (ctg301)
HK375	14	75, 720, 876, 938, 962, 1074, 1237, 1595, 1674, 2195, 2380, 2713, 2831, 3335 (ctg301)
HK623	4	228, 689, 3046, 3757 (ctg102)
HK661	12	369, 451, 1315, 2736, 2959, 3438, 3481, 3595, 3603, 3730, 3774, 4144 (ctg204)
HK673	7	32, 221, 521, 664, 1403, 2729, 3034 (ctg602)
HK679	9	113, 374, 976, 1622, 1668, 2060, 2120, 3136, 3445 (ctg301)
HK693	8	1785, 2677, 2902, 3151, 3385, 3894, 3945, 4159 (ctg102)
HK731	6	1024, 1495, 1549, 1784, 2957, 4060 (ctg201)
HK757	9	146, 736, 784, 1898, 2407, 3293, 3321, 3503, 3610 (ctg207)
HK863	9	197, 1418, 2169, 2197, 2488, 2961, 3230, 3311, 3882 (ctg701)
HK879	14	69, 667, 1723, 1738, 1971, 1992, 2187, 2220, 2532, 2844, 2917, 2972, 3299, 3992 (ctg502)
HK907	10	281, 433, 566, 807, 1581, 1629, 2080, 3041, 4068, 4121 (ctg102)
HK919	13	306, 817, 989, 1168, 1325, 2374, 2433, 2951, 3010, 3558, 3956, 4036, 4123 (ctg102)
HK929	12	27, 613, 1399, 1424, 1978, 2983, 3045, 3075, 3172, 3188, 3390, 4223 (ctg101)
HK931	15	169, 841, 877, 903, 1271, 1353, 1868, 1967, 2096, 2549, 2617, 2847, 2933, 3265, 3808 (ctg102)
HK933	8	132, 337, 440, 500, 1135, 2344, 2567, 4053 (ctg102)
HK961	7	897, 921, 1133, 2240, 3049, 3184, 4181 (ctg103)
HK1011	11	17, 824, 1219, 1299, 1525, 1626, 1948, 2151, 2709, 3001, 3852 (ctg402)
HK1009	16	140, 310, 579, 597, 620, 640, 1088, 1277, 1432, 1728, 2390, 2969, 3281, 3339, 3863, 4195 (ctg401)
HK1015	13	50, 558, 769, 976, 1067, 1543, 1622, 1984, 2154, 2236, 2697, 3219, 3606 (ctg301)
HK1021	17	155, 862, 1072, 1229, 1435, 1527, 1530, 2090, 2111, 2635, 2834, 3274, 3491, 3845, 4034, 4117, 4145 (ctg602)
HK1039	13	121, 352, 362, 885, 1361, 1366, 1678, 2212, 2518, 2590, 3025, 3656, 4023 (ctg301)
HK1043	14	1055, 1056, 1110, 1369, 1398, 1452, 1502, 1796, 2131, 2404, 2578, 3479, 3671, 4191 (ctg102)
HK1045	14	202, 519, 742, 882, 1220, 1951, 1987, 2788, 2845, 3280, 3291, 3682, 3974, 4066 (ctg102)
HK1057	5	1812, 3295, 3797, 3830, 4096 (ctg207)
HK1059	8	120, 258, 400, 422, 1516, 3764, 3812, 3880 (ctg301)
HK1063	15	293, 317, 424, 426, 1598, 1859, 2458, 2656, 2955, 3419, 3513, 3602, 3743, 3768, 3868 (ctg401)
HK1097	11	96, 428, 518, 1178, 1491, 1753, 2540, 2553, 2919, 3760, 4097 (ctg401)
HK1115	3	9, 619, 3437 (ctg208)
HK1139	11	616, 1459, 1716, 1962, 1994, 2132, 2626, 2642, 2940, 3537, 3657 (ctg502)
Total	329	
Average	10.6	

Note: Fingerprinted *Hin*dIII clones were numbered 1 to 2688 and fingerprinted *Bam*HI clones were numbered 2689 to 4224.

Integration of the physical map with the genome sequence and refinement of the physical map

As part of the *F. graminearum* sequencing project, the 4608 clones from the *Hin*dIII library were end-sequenced and used in the assembly of the genome sequence map (Cuomo et al. 2007; Broad Institute 2007). In our study, the first 2688 *Hin*dIII clones were fingerprinted for the development of our physical map. Therefore, the *Hin*dIII BAC ends used in the sequence assembly and the sequence-based genetic markers derived from the genome sequence provide an excellent resource for the integration of the genome sequence with the physical and genetic maps. After the physical map was constructed, it was used for large-scale comparison with the genome sequence. The comparison showed that 100 of 112 initial contigs contain common *Hin*dIII BAC clones with the genome sequence and only 3 contigs (ctg42, ctg18, and ctg19) are not consistent with the sequence contigs (Table S2[2]). For example, some common BAC clones in initial contig ctg42 were found in sequence supercontigs 1 and 5. After carefully inspecting ctg42, we concluded that the false assembly was likely caused by clone 2819, which exhibited 75 bands. The large number of bands observed in clone 2819 possibly resulted in us placing it in the incorrect contig. This situation also occurred with initial contig ctg18, which contained clone 4094 with 82 bands. At higher stringency (cutoff of 10^{-13}), the initial contigs ctg42 and ctg18 were divided into 3 (rearranged ctg112, ctg113, and ctg218) and 2 (rearranged ctg409 and ctg505) contigs, respectively (Table S2). The other inconsistency occurred in ctg19, which contained several incorrectly located clones (414, 934, and 2178). This was likely due to their low number of bands resulting in their improper assembly with 2 large clones, 3196 and 4021. Therefore, the large-scale comparison confirms the integrity of the sequence assembly and indicates that the BAC contig map is also highly reliable. The comparison was also used to rearrange the 112 initial contigs into 115 rearranged contigs that are nearly equivalent to the 10 largest supercontigs of the DNA sequence, which contain 99.3% of the sequence assembly (Table S2). Furthermore, these 115 contigs were further assembled into 26 final contigs using lower analysis stringency (cutoff of 10^{-4} and tolerance of 7) (Table 1). These 26 larger contigs comprised 17 920 bands, which span 39.2 Mb (17 920 bands $\times$ 2.19 kb/band = 39.2 Mb). The size is

© 2007 NRC Canada

Table 3. *Bam*HI BAC clone end sequencing.

End-sequenced clones				BLASTN results	
Clone No.	Insert size (kb)	Contig	Sequence	Supercontig (contig: sequenced region)	*E* value
4174	105	101	Forward	1 (2: 48977–49641)	0.0
			Reverse	21 (485: 1864–2472)	0.0
3313	88	103	Forward	20 (484: 3526–3814)	e–147
			Reverse	1 (136: 133008–133461)	0.0
3586	100	208	Forward	2 (194: 108642–109340)	0.0
			Reverse	2 (194: 208554–209133)	0.0
3033	90	301	Forward	3 (196: 64845–64958)	4e–30
			Reverse	3 (196: 156815–157512)	0.0
4042	90	301	Forward	3 (258: 4956–5475)	0.0
			Reverse	3 (257: 47280–47988)	0.0
2997	92	401	Forward	4 (259: 155722–156264)	0.0
			Reverse	4 (259: 62866–63517)	0.0
4211	85	403	Forward	Sequencing failed	
			Reverse	4 (321: 158035–158303)	7e–56
4142	100	601	Forward	6 (367: 172090–172826)	0.0
			Reverse	6 (370: 33458–34154)	0.0
3862	85	702	Forward	7 (446: 130047–130624)	0.0
			Reverse	7 (446: 215674–215852)	e–73
3146	100	801	Forward	8 (460: 148748–149121)	0.0
			Reverse	8 (460: 250488–251155)	0.0
3082	110	901	Forward	9 (464: 19976–20697)	0.0
			Reverse	30 (496: 1156–1859)	0.0
3592	110	901	Forward	9 (467: 131140–131853)	0.0
			Reverse	9 (467: 17720–17856)	8e–30
4166	85	1001	Forward	10 (473: 160955–161611)	0.0
			Reverse	10 (473: 75627–76231)	0.0
3875	87	1001	Forward	10 (473: 71089–71710)	0.0
			Reverse	12 (476: 6765–7217)	0.0
4070	102	1	Forward	7 (444: 13732–14172)	0.0
			Reverse	7 (442: 19846–20386)	0.0
2871	97	2	Forward	9 (467: 146091–146807)	0.0
			Reverse	9 (470: 7183–7881)	0.0
4111	82	2	Forward	9 (467: 157179–157926)	0.0
			Reverse	9 (470: 3071–3780)	0.0

close to the total length of the sequence contigs and the estimated genome size. Because of the analysis limitations of the FPC program, the possible overlap of these large contigs was difficult to detect, but it could be identified using end-probe hybridization or end sequencing in the future.

Integration of the genetic map with the physical map

To further validate the contigs, and to anchor the genetic map to the physical map, we screened both BAC libraries with 31 markers. We used markers from a sequence-based genetic map that spans the entire genome (Gale et al. 2005). We developed pools of both libraries and conducted PCR with 31 sequence-based genetic markers (Table 1). We used 11, 8, 6, and 6 markers on *F. graminearum* chromosomes 1, 2, 3, and 4, respectively. One marker, HK339, was located on the genome sequence but was not genetically mapped. Positive BAC clones were confirmed by PCR on the individual BAC clone. For the *Hin*dIII library, we identified an average of 10.3 BAC clones (range 2 to 18) per genetic marker. For the *Bam*HI library, we identified an average of 8.5 BAC clones (range 3 to 15) per genetic marker. Based on the BAC library screening data, the *Hin*dIII library covers 10.3 genome equivalents and the *Bam*HI library covers 8.5 genome equivalents, for a total of 18.8 genome equivalents, very similar to the estimated 21.5 genome equivalents based on the number of clones and the size of the inserts in the libraries. The marker screening data were integrated with our physical map data (Tables 1 and 2). According to the screening results of the fingerprinted BAC clones, the average number of positive clones per marker is 10.6, similar to the 11.9-fold genome coverage of the fingerprinted BAC clones.

We identified 6 final contigs (ctg102, ctg207, ctg301, ctg401, ctg502, and ctg602) with more than one marker and 7 with a single marker (Table 1 and Table S2). In the case of contigs with more than one marker, the existence or order of markers in the contigs matches with that in the genetic map (Gale et al. 2005). For example, in the case of ctg102, 8 markers were observed that were consistent with the genetic map. A portion of ctg102 with markers HK623, HK1043, and HK1045 is shown in Fig. 3. These screening results confirm the reliability of our physical map.

© 2007 NRC Canada

Fig. 3. A portion of final contig 102. Positive clones (highlighted in blue in the Web version) screened by markers (HK623, HK1043, and HK1045) have been located to specific regions in the contig. The order of these markers in the contig is identical to that on the public genetic map (Gale et al. 2005).

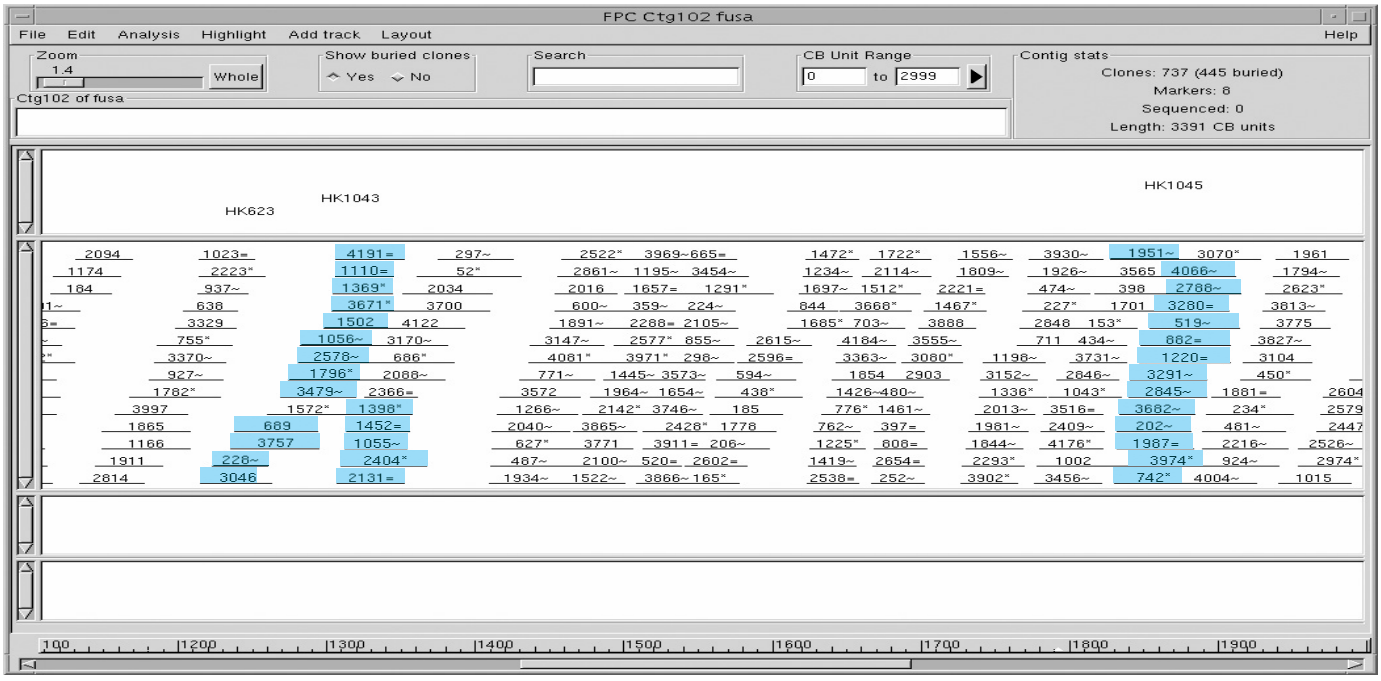
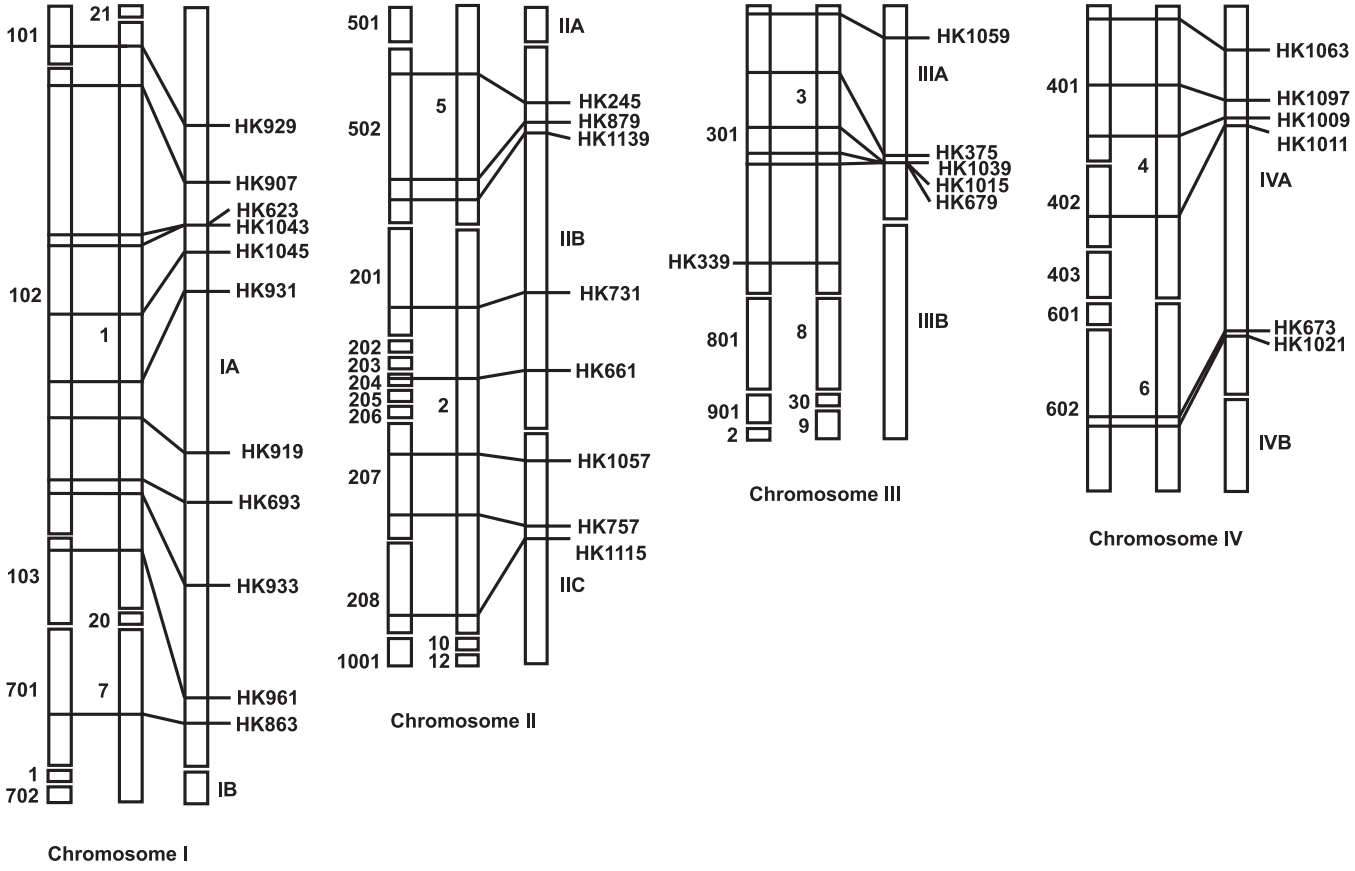


Fig. 4. Integrated BAC-based physical map, genetic map, and genome sequence of *F. graminearum*. The 4 chromosomes of *F. graminearum* are shown. Each chromosome is indicated with the BAC contigs on the left side, the genome sequence contigs in the middle, and the genetic map on the right. Except for HK339, all markers are shown on the right side of the genetic linkage groups.



© 2007 NRC Canada

Extending the genome sequence

To further validate and extend the genome sequence–based contigs, we conducted end sequencing of 17 *Bam*HI BAC clones that were on the ends of our fingerprint-based contigs. Since the *Bam*HI library was not used in the assembly of the genome sequence, our assumption was that the *Bam*HI clones could connect additional sequence-based supercontigs. Four of the BAC clones that were end-sequenced provided new links between sequencing supercontigs (Table 3). End sequences of clone 4174 at the end of BAC contig 101 provide a direct link between sequencing supercontigs 1 and 21 at the end of chromosome 1. Clone 3313, at the opposite end of sequencing supercontig 1, links this end with supercontig 20, falling between supercontigs 1 and 7 in the middle of chromosome 1. The end sequences for clone 3875 provide a link between sequencing supercontigs 10 and 12 at the end of chromosome 2. Clone 3082 provides a link between sequencing supercontigs 9 and 30 that had been previously suggested by manual assembly methods (Gale et al. 2005) between sequencing supercontigs 9 and 8 on chromosome 3. End sequences of other clones, while not providing new linkages among sequencing supercontigs, are consistent with sequences predicted based on the results of large-scale comparisons. In conclusion, with a total length of 39.2 Mb, the new fingerprinting-based map is a reliable addition to the DNA sequence map for genome research.

The integrated physical, genetic, and sequence map of *F. graminearum* and future applications

We developed a genome-wide BAC-based physical map of the plant pathogen *F. graminearum* from 4224 BAC clones randomly selected from 2 libraries. The genome coverage of the BAC clones used for the physical map assembly is estimated at 11.9-fold. The physical map consists of 26 contigs that span approximately 39.2 Mb. The physical map has been integrated with the genetic map and the genome sequence and was tested for accuracy based on end sequencing results, comparisons with the genetically mapped markers, and the genome sequence. Figure 4 shows a schematic of the integrated maps. The comparisons that we conducted showed that our BAC-based physical map is highly consistent with the genetic map and the genome sequence. Therefore, we present here the first integrated physical, genetic, and genome sequence map of the *F. graminearum* genome. The BAC-based contigs and marker screening results can be found at http://140.130.98.33/fpc/WebAGCoL/fusa/WebFPC, which is useful to search any specific marker or BAC clone in the physical map. Multiple applications of the integrated physical, genetic, and sequence map include structural and comparative genomics, finishing the genome sequence, and map-based cloning. The *Bam*HI BAC library clones will be particularly useful for filling in gaps in the genome sequence.

Acknowledgements

We thank Liane Gale for providing genetic marker information and Karen Hillburn for isolating Fusarium protoplasts for the DNA extractions. We also thank the College of Agriculture at National Chiayi University for providing funding to purchase a workstation computer for data analysis. This work was supported by grants from the University of Minnesota Rapid Agricultural Response Fund and the US Wheat and Barley Scab Initiative.

References

Aign, V., Schulte, U., and Hoheisel, J.D. 2001. Hybridization-based mapping of *Neurospora crassa* linkage groups II and V. Genetics, **157**: 1015–1020. PMID:11238391.

Bai, G.H., and Shaner, G. 1994. Scab of wheat: prospect for control. Plant Dis. **78**: 760–766.

Boehm, E.W.A., Ploetz, R.C., and Kistler, H.C. 1994. Statistical analysis of electrophoretic karyotype variation among vegetative compatibility groups of *Fusarium oxysporum* f. sp. *cubense*. Mol. Plant Microbe Interact. **7**: 196–207.

Bowden, R.L., and Leslie, J.F. 1999. Sexual recombination in *Gibberella zeae*. Phytopathology, **89**: 182–188. doi:10.1094/PHYTO.1999.89.2.182.

Broad Institute. 2007. *Fusarium graminearum* sequencing project [online]. Available from http://www.broad.mit.edu/annotation/genome/fusarium_graminearum [accessed August 2006].

Chang, Y.-L., Tao, Q., Scheuring, C., Ding, K., Mekesm, K., and Zhang, H.-B. 2001. An integrated map of *Arabidopsis thaliana* for functional analysis of its genome sequence. Genetics, **159**: 1231–1242. PMID:11729165.

Cuomo, C.A., Güldener, U., Xu, J.R., Trail, F., Turgeon, B.G., Di Pietro, A., et al. 2007. The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. Science (Washington, D.C.), **317**: 1400–1402. doi:10.1126/science.1143708.

Diener, S.E., Chellappan, M.K., Mitchell, T.K., Dunn-Coleman, N., Ward, M., and Dean, R.A. 2004. Insight into *Trichoderma reesei*'s genome content, organization and evolution revealed through BAC library characterization. Fungal Genet. Biol. **41**: 1077–1087. doi:10.1016/j.fgb.2004.08.007. PMID:15531212.

Frijters, A.C.J., Zhang, Z., Damme, M.V., Wang, G.L., and Ronald, P.C. 1997. Construction of a bacterial artificial chromosome library containing large *Eco*RI and *Hin*dIII genomic fragments of lettuce. Theor. Appl. Genet. **94**: 390–399. doi:10.1007/s001220050428.

Gale, L.R., Bryant, J.D., Calvo, S., Giese, H., Katan, T., O'Donnell, K., et al. 2005. Chromosome complement of the fungal plant pathogen *Fusarium graminearum* based on genetic and physical mapping and cytological observations. Genetics, **171**: 985–1001. doi:10.1534/genetics.105.044842. PMID:16079234.

Güldener, U., Seong, K.-Y., Boddu, J., Cho, S., Trail, F., Xu, J.-R., et al. 2006*a*. Development of a *Fusarium graminearum* Affymetrix GeneChip for profiling fungal gene expression *in vitro* and *in planta*. Fungal Genet. Biol. **43**: 316–325. doi:10.1016/j.fgb.2006.01.005. PMID:16531083.

Güldener, U., Mannhaupt, G., Münsterkötter, M., Haase, D., Oesterheld, M., Stümpflen, V., et al. 2006*b*. FGDB: a comprehensive fungal genome resource on the plant pathogen *Fusarium graminearum*. Nucleic Acids Res. **34**: D456–D458. doi:10.1093/nar/gkj026. PMID:16381910.

Kim, U.-J., Birren, B.W., Slepak, T., Mancino, V., Boysen, C., Kang, H.-L., Simon, M.I., and Shizuya, H. 1996. Construction and characterization of a human bacterial artificial chromosome library. Genomics, **34**: 213–218. doi:10.1006/geno.1996.0268. PMID:8661051.

Martin, S.L., Blackmon, B.P., Rajagopalan, R., Houfek, T.D., Sceeles, R.G., Denn, S.O., et al. 2002. MagnaportheDB: a federated solution for integrating physical and genetic map data with BAC end derived sequences for the rice blast fungus *Magna-*

© 2007 NRC Canada

porthe grisea. Nucleic Acids Res. **30**: 121–124. doi:10.1093/nar/30.1.121. PMID:11752272.

McMullen, M., Jones, R., and Gallenberg, D. 1997. Scab of wheat and barley: a re-emerging disease of devastating impact. Plant Dis. **81**: 1340–1348. doi:10.1094/PDIS.1997.81.12.1340.

Meksem, K., Shultz, J., Tebbji, F., Jamai, A., Henrich, J., Kranz, H., et al. 2005. A bacterial artificial chromosome based physical map of the *Ustilago maydis* genome. Genome, **48**: 207–216. PMID:15838542.

Nelson, P.E., Tousson, T.A., and Marasas, W.F.O. 1983. *Fusarium* species: an illustrated manual for identification. The Pennsylvania State University Press, University Park, Pa.

Nganje, W.E., Katiebie, S., Wilson, W.W., Leistritz, F.L., and Bangsund, D.A. 2004. Economic impacts of Fusarium head blight in wheat and barley: 1993–2001. North Dakota State University Agribusiness and Applied Economics Report 538.

O'Donnell, K., Kistler, H.C., Tacke, B.K., and Casper, H.H. 2000. Gene genealogies reveal global phylogeographic structure and reproductive isolation among lineages of *Fusarium graminearum*, the fungus causing wheat scab. Proc. Natl. Acad. Sci. U.S.A. **97**: 7905–7910. doi:10.1073/pnas.130193297. PMID: 10869425.

O'Donnell, K., Ward, T.J., Geiser, D.M., Kistler, H.C., and Aoki, T. 2004. Genealogical concordance between the mating type locus and seven other nuclear genes supports formal recognition of nine phylogenetically distinct species within the *Fusarium graminearum* clade. Fungal Genet. Biol. **41**: 600–623. doi:10.1016/j.fgb.2004.03.003. PMID:15121083.

Pampanwar, V., Engler, F., Hatfield, J., Blundy, S., Gupta, G., and Soderlund, C. 2005. FPC web tools for rice, maize, and distribution. Plant Physiol. **138**: 116–126. doi:10.1104/pp.104.056291. PMID:15888684.

Schein, J.E., Tangen, K.L., Chiu, R., Shin, H., Lengeler, K.B., MacDonald, W.K., et al. 2002. Physical maps of genome analysis of serotype A and D strains of the fungal pathogen *Cryptococcus neoformans*. Genome Res. **12**: 1445–1453. doi:10.1101/gr.81002. PMID:12213782.

Schroeder, H.W., and Christensen, J.J. 1963. Factors affecting resistance of wheat scab caused by *Gibberella zeae*. Phytopathology, **53**: 831–838.

Soderlund, C., Humphray, S., Dunham, A., and French, L. 2000. Contigs built with fingerprints, markers, and FPC v4.7. Genome Res. **10**: 1772–1787. doi:10.1101/gr.GR-1375R.

Sutton, J.C. 1982. Epidemiology of wheat head blight and maize ear rot caused by *Fusarium graminearum*. Trans. Br. Mycol. Soc. **70**: 187–192.

Thon, M.R., Martin, S.L., Goff, S., Wing, R.A., and Dean, R.A. 2004. BAC end sequences and a physical map reveal transposable element content and clustering patterns in the genome of *Magnaporthe grisea*. Fungal Genet. Biol. **41**: 657–666. doi:10.1016/j.fgb.2004.02.003. PMID:15275661.

Trail, F., and Common, R. 2000. Perithicial development by *Gibberella zeae*: a light microscopy study. Mycologia, **92**: 130–138. doi:10.2307/3761457.

Trail, F., Xu, J.R., San Miguel, P., Halgren, R.G., and Kistler, H.C. 2003. Analysis of expressed sequence tags from *Gibberella zeae* (anamorph *Fusarium graminearum*). Fungal Genet. Biol. **38**: 187–197. doi:10.1016/S1087-1845(02)00529-7. PMID:12620255.

Tuite, J., Shaner, G., and Everson, R.J. 1990. Wheat scab in soft red winter wheat in Indiana in 1986 and its relation to quality measurements. Plant Dis. **74**: 959–962. doi:10.1094/PD-74-0959.

Ward, T.J., Bielawski, J.P., Kistler, H.C., Sullivan, E., and O'Donnell, K. 2002. Ancestral polymorphism and adaptive evolution in the trichothecene gene cluster of phytopathogenic *Fusarium*. Proc. Natl. Acad. Sci. U.S.A. **99**: 9278–9283. doi:10.1073/pnas.142307199. PMID:12080147.

Xu, Z., Berg, M.A., Scheuring, C., Covaleda, L., Lu, H., Santos, F.A., et al. 2005. Genome physical mapping from large-insert clones by fingerprint analysis with capillary electrophoresis: a robust physical map of *Penicillum chrysogenum*. Nucleic Acids Res. **33**: e50. doi:10.1093/nar/gni037. PMID:15767275.

Yun, S.H., Arie, T., Kaneko, I., Yoder, O.C., and Turgeon, B.G. 2000. Molecular organization of mating type loci in heterothallic, homothallic, and asexual *Gibberella/Fusarium* species. Fungal Genet. Biol. **31**: 7–20. doi:10.1006/fgbi.2000.1226. PMID: 11118131.

Zhang, H.B., Choi, S., Woo, S.S., Li, Z., and Wing, R.A. 1996. Construction and characterization of two rice bacterial artificial chromosome libraries from the parents of a permanent recombinant inbred mapping population. Mol. Breed. **2**: 11–24. doi:10.1007/BF00171348.

Zhang, X., Sheuring, C., Tripathy, S., Xu, Z., Wu, C., Ko, A., et al. 2006. An integrated BAC and genome sequence physical map of *Phytophthora sojae*. Mol. Plant Microbe Interact. **19**: 1302–1310. doi:10.1094/MPMI-19-1302. PMID:17153914.

© 2007 NRC Canada