Fungal Genetics and Biology 46 (2009) 400–409

Contents lists available at ScienceDirect

Fungal Genetics and Biology

journal homepage: www.elsevier.com/locate/yfgbi

A motif within a complex promoter from the oomycete *Phytophthora infestans* determines transcription during an intermediate stage of sporulation

Qijun Xiang, Kyoung Su Kim, Sourav Roy, Howard S. Judelson *

Department of Plant Pathology and Microbiology, University of California, Riverside, CA 92521, USA

ARTICLE INFO

Article history:
Received 16 November 2008
Accepted 3 February 2009
Available online 27 February 2009

Keywords:
Developmental biology
Late blight
Oomycete
Promoter structure
Protein kinase
Transcription factor

ABSTRACT

Sporulation in *Phytophthora infestans* is associated with a major remodeling of the transcriptome. To better understand promoter structure and how sporulation-specific expression is determined in this organism, the *Pks1* gene was analyzed. *Pks1* encodes a protein kinase that is induced at an intermediate stage of sporulation, prior to sporangium maturation. Major and minor transcription start sites mapped throughout the promoter, which contains many T-rich stretches and Inr-like elements. Within the T-rich region are several motifs which bound nuclear proteins in EMSA. Tests of modified promoters in transformants implicated a CCGTTG located 110-nt upstream of the transcription start point as a major regulator of sporulation-specific transcription. The motif also bound a sporulation-specific nuclear protein complex. A bioinformatics analysis indicated that the motif is highly over-represented within co-expressed promoters, in which it predominantly resides 100–300-nt upstream of transcription start sites. Other sequences, such as a CATTTGTT motif, also bound nuclear proteins but did not play an essential role in spore-specific expression.

1. Introduction

Many eukaryotic microbes, including true fungi and oomycetes, utilize asexual spores for dissemination and infection. The differentiation of these structures is largely regulated at the level of RNA synthesis by transcription factors that recruit RNA polymerase II to promoters. These promoters are typically bipartite, containing a proximal core region which binds the transcriptional apparatus near the initiation site, and distal activation or repression domains (Lee and Young, 2000). Tissue-specific transcription is thought to be predominantly regulated by the distal sites, although contributions may also be made by core promoters (Smale, 2001). An important element of understanding an organism entails identifying the promoter motifs and transcription factors that regulate its development.

Little is known of transcriptional mechanisms that control sporulation or other developmental processes in oomycetes, a group of important plant and animal pathogens and saprophytes. Compared to well-studied model organisms, oomycetes have a unique evolutionary history as they reside in the kingdom Stramenopila (Baldauf et al., 2000). The extent to which transcriptional regulation in oomycetes resembles that of other eukaryotes remains to be determined. Data from oomycete genomes indicate that their intergenic regions are typically small (<500 nt) and may therefore lack the complexity of many metazoan and plant promoters. Many oomycete promoters contain sequences resembling the Initiator or Inr (YYANWYY), which is one of several core promoter motifs found in higher eukaryotes (Juven-Gershon et al., 2006; McLeod et al., 2004; Smale, 2001).

Phytophthora infestans, which causes the devastating late blight disease of potato, is a promising model for studying both oomycete spores and promoters (Fry, 2008; Judelson and Blanco, 2005). Copious amounts of sporangia are produced which are easily purified for molecular studies, and genome sequence and expression data are available to aid analyses of transcription during development (Judelson et al., 2008). Although its application is complicated by copy number and position effects, a transformation system has proved useful for manipulating genes and promoters associated with development. This has helped to define the binding sites for two stage-specific transcription factors, one acting during early sporulation and another during zoosporogenesis (Ah Fong et al., 2007; Tani and Judelson, 2006).

Sporulation in *P. infestans* begins when vegetative mycelia, in response to environmental and internal cues, produce aerial hyphae and then sporangiophores (Hardham and Hyde, 1997). After cytoplasm, nuclei, and other organelles move into each developing sporangium, a septum forms at its base and a papilla at its apex that is important for germination. Unlike the typically dormant fungal spores, *P. infestans* sporangia are physiologically active and capable of rapid germination. At lower temperatures and high humidity, which are the climatic factors most conducive to epidemics, the sporangium cleaves into six or more flagellated zoospores that exit through the papilla. After an interval of swimming, each zoospore forms a cyst, germ tube, and appressorium. Genes transcribed during

* Corresponding author.
E-mail address: howard.judelson@ucr.edu (H.S. Judelson).

1087-1845/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.fgb.2009.02.006

sporulation appear to participate in both early events in sporangia development and subsequent stages including germination and encystment (Clark et al., 1978; Penington et al., 1989).

To enhance our understanding of both asexual reproduction and gene expression mechanisms in *P. infestans*, this study describes the transcription pattern and promoter structure of the sporulation-induced *Pks1* protein kinase gene, which had been identified as a cDNA in an expression profiling study (Kim and Judelson, 2003). Using transformants expressing the *Pks1* promoter fused to the β-glucuronidase reporter, the gene was shown to be expressed at an intermediate stage of sporulation. Through functional dissection and bioinformatics, the promoter was shown to exhibit a complex structure containing multiple regulatory elements and binding sites for nuclear proteins. In addition, a motif necessary for activating *Pks1* during sporulation was detected which also exists within many similarly-expressed promoters. Such data helps illuminate the pathways regulating asexual reproduction, which in the long-term may lead to targets for crop protection strategies.

2. Materials and methods

2.1. Manipulations of P. infestans

Isolates were cultured on rye-sucrose media at 18 °C in the dark. Strains employed included isolate 1306 and transformants made using the protoplast method (Judelson, 1993). Non-sporulating mycelia were obtained by inoculating clarified rye-sucrose broth with sporangia, followed by 3 days incubation. Sporulating hyphae were from rye-sucrose agar cultures grown for 9–11 days. Sporangia were purified from such cultures by adding water, rubbing with a glass rod, and passing the fluid through a 50-μm mesh to remove hyphal fragments. To induce germination, sporangia were either placed in rye-sucrose broth for 6 h at 18 °C, or water at 4 °C for 1 h to induce zoospores. RNA was extracted using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA).

2.2. Cloning and sequencing of Pks1

A cDNA clone corresponding to *Pks1* was hybridized to a bacterial artificial chromosome library (Randall and Judelson, 1999). Positive colonies were subcloned and sequenced.

2.3. Gene expression analysis

Staining for β-glucuronidase (GUS) was performed as described (Judelson, 1993). For blot studies, total RNA was resolved on formaldehyde gels and hybridized with ^{32}P-labeled probes. These were generated by polymerase chain reaction using primer pairs for *Pks1* (5′-TAGGATCCTTCTCTTCCAACGTAAG and 5′-CGGCGGCCGCGA AATCTTGATCCG AACCGA), GUS (5′-AAACCCCAACCCGTGAAAT-CAAAA and 5′-ATCGGTGTGAGCGTCGCAGAA CA), or elongation factor 1-α (EF-1; 5′-ATTCTGGTGGTCGCTTCGGGTGTGG and 5′-GA CGAA ACCACGACGCAGCTCCTTG) and labeled by the random primer method. To map transcription start sites (TSSs), RACE was performed using a kit from Invitrogen (Carlsbad, CA, USA) with specific primers QX68R (5′-GCGTGTTTTGCTGCCAGAA and QX69R (5′-CCTCACTCAACACCGAGTCCTTA). Products were resolved on 2% agarose, cloned into pGEM-T Easy (Promega, Madison, WI, USA), and sequenced.

2.4. GUS plasmid construction

Promoter fragments were generated by polymerase chain reaction (PCR) using Platinum Taq DNA polymerase High-Fidelity (Invitrogen) and primers with *Apa*I sites, and cloned into pOGUS which contains a promoterless GUS gene along with neomycin phosphotransferase (*nptII)* for G418 selection. For 5′ deletion constructs, plasmids are named based on the size of the promoter as counted upstream from the TSS (+1); p877 contains 877 nt of sequences upstream of the TSS plus the 125-nt 5′ untranslated region, for example. These were amplified using reverse primer RP1 (5′-AAGGGCCC GATAGCTTCGGACTTAG) and upstream primers 5′-AAGGGCCCGGGATACTATTGATGCT for p877, 5′-AAGGGCCCAC CCGAACGACCCCAAGA for p525, 5′-AAGGGCCCAATGGTAG GTG CTCTTAG for p225, 5′-AAGGGCCCTTTGCGTTGCCGTTGTGC for p119, 5′-AAGGGC CCTTGTGCCATTTGTTTT for p107, 5′-AAGGG CCCTTTTTCCTCGTCCCA for p95, 5′-AAGG GCCCCGTCCCATTTTCTCA for p87, 5′-AAGGGCCCTTCTCATGCTGCTCGTAT for p78, or 5′-AA GGGCCCAAAAGTGTTCTCCATCG for p50. To mutate the CGTTG motifs, PCR was performed using RP1 with MR1 (5′-TTGGGCCCT TTTGACGGTCCGTTGTGCCATTTGTT), MR2 (5′-TTGGGCCCTTTT GCGTTGCATGACTGCCATTTGTTTTTCC), or MR12 (5′-TTGGGC CCTTTTGACGGTCATGACTGCCATTTGTTTT TCC) using p119 as template. To mutate the CAGTTTGTT motif, PCR used RP1 and NC (5′-TT GGGCCCTTTTGCGTTGCCGTTGTGCACGGGTGGTTTCCTCGTCCCAT) against p119. Other mutations described in the text were made by two stage PCR using outer primers matching wild-type *Pks1* and inner mutagenic primers.

2.5. Electrophoretic mobility shift assay (EMSA)

Nuclear protein isolation and EMSA were performed as described (Ah Fong et al., 2007), except that heparin agarose was not used for the extractions. Briefly, EMSA involved mixing 5 μg of nuclear protein with 1 μg poly dI-dC and 1.6 ng of ^{32}P-labeled probe for 30 min on ice, followed by electrophoresis at room temperature. For competition assays, protein was incubated with unlabeled DNA for 30 min and then the labeled probe. Double-stranded oligonucleotides listed in Results were used as probe and cold competitors. Bands were quantified using a phosphorimager.

2.6. Bioinformatics surveys

Genes induced or repressed during sporulation were identified from microarray data (Judelson et al., 2008). The corresponding promoters, and total promoters, were extracted from the *P. infestans* genome database developed by the Broad Institute of MIT and Harvard (www.broad.mit.edu). Motif searches were performed using a combination of text editors and custom perl scripts, and background frequencies were calculated based on the base composition of *P. infestans* promoters.

3. Results

3.1. Pks1 encodes a protein kinase induced upon spore maturation

Our previous macroarray study identified a cDNA clone named *pisp18* that corresponds to a gene induced during sporulation (Kim and Judelson, 2003). To identify the structure of the full-length gene, a genomic clone was obtained from a BAC library and sequenced. The gene is renamed here as *Pks1* (*P. infestans* protein kinase in sporangia 1), as its sequence matched a serine/threonine protein kinase. Subsequently, *Pks1* was defined in the *P. infestans* genome sequencing project underway at the Broad Institute of MIT and Harvard as gene PITG_10884. DNA blot analysis indicated that *Pks1* is a single-copy gene (not shown).

Pks1 encodes an intron-lacking open reading frame of 541 amino acids with similarity to the AGC group of the eukaryotic protein kinase superfamily. AGC kinases typically control essential cellular processes such as growth, cell cycle progression, differentiation,

and stress resistance (Bogre et al., 2003). The *P. infestans* protein contains a central kinase domain (Pfam00069), a protein kinase C-like terminal catalytic domain (Pfam00433), and a weak match at its N-terminus to a Pleckstrin homology domain (Pfam00169). Its best match in GenBank is to an uncharacterized protein kinase from *Paramecium* (accession CAK85399, $E = 2e{-}62$), and the best match to a functionally characterized protein is to a serum/glucocorticoid-regulated kinase from *Xenopus laevis* (NP_001083809, $E = 1e{-}57$). Based on alignments with a putative full-length cDNA and RACE experiments described below, 5′ and 3′ untranslated regions of the *Pks1* transcript are 125 and 59 nt, respectively, yielding a mRNA of about 1.9 kb.

The expression pattern of *Pks1* was assessed using RNA blot and GUS reporter methods. In blots, *Pks1* mRNA was not detected in non-sporulating mycelia but had strong signals in sporangia (Fig. 1A). The mRNA was maintained after sporangia were induced to indirectly germinate, *i.e.* release swimming zoospores. However, when sporangia were placed in rye-sucrose broth to stimulate direct germination, the mRNA disappeared. To follow expression using GUS (and obtain a functional promoter for later analysis), sequences upstream of the ATG translation start were cloned next to the promoter-lacking GUS gene in pOGUS and stable G418-resistant transformants were generated. Initial studies utilized fragments contained 877 and 525 nt of DNA upstream of the transcription start point (TSS), which yielded similar results. As in previous studies with *Phytophthora* promoters, convincing expression was detected in only a subset of transformants due to position effects (Ah Fong et al., 2007; Tani and Judelson, 2006). Consequently, transformants were first screened for expression by GUS histochemical staining and then analyzed using RNA blots. Such blots indicated that the transgene showed a pattern of expression similar to that of wild-type *Pks1* (Fig. 1C). Histochemical staining was typically observed in both immature sporangia connected to sporangiophores (Fig. 1D, panels 2–4) and mature sporangia (Fig. 1D, panel 5). Whether the staining in sporangiophores reflects expression in that compartment or diffusion of the reaction product from immature sporangia cannot be determined. Nevertheless, the timing of *Pks1* expression follows early-induced genes such as *Cdc14* that are expressed in sporangiophore and sporangia initials (Ah Fong et al., 2007), which lack *Pks1* expression as shown in Fig. 1D panel 1. Therefore, *Pks1* expression appears to normally begin at an intermediate stage. In much older cultures, about 5 days after the onset of sporulation, *Pks1* expression was sometimes observed in hyphae containing lateral bulges (Fig. 1C, panel 6). This suggests that development proceeds more slowly or with a slightly different pattern in aged cultures, if these represent sporangiophore initials.

3.2. Multiple transcription start sites from Inr-like and T-rich regions

5′ RACE was performed to locate the transcription start sites (TSSs). Typical reactions are shown in Fig. 1B, and the TSSs are mapped in Fig. 2 with bent arrows. Both a predominant and minor band were detected in sporangia, with the major TSS being 125-nt upstream of the translation start codon at a cytosine defined as +1. A few bands extended to the adenosine at −1. Both are within a region that resembles the eukaryotic Inr consensus YYANWYY (Juven-Gershon et al., 2006), except that the last base is not conserved in *Pks1* ($TCAC_{+1}TTG$). This and other Inr-like sequences are denoted by wavy lines in Fig. 2. It should be noted that the Inr has a loose consensus in most organisms, and some other *P. infestans* genes such as *ipiO1* also contain a G at the 3′ position (McLeod et al., 2004). Interestingly, while the Inr-like sequence at +1 is identical in *P. infestans* and *Phytophthora sojae*, it is very different in *Phytophthora ramorum* where the corresponding region (TGGCAAT) is tandemly repeated. The 23 bases downstream of the putative *P. infestans* Inr are nearly identical between the species; while the 5′ region of this conserved block resembles the "flanking promoter region" or FPR motif noted by McLeod et al. (2004), the rest appears to be a novel motif.

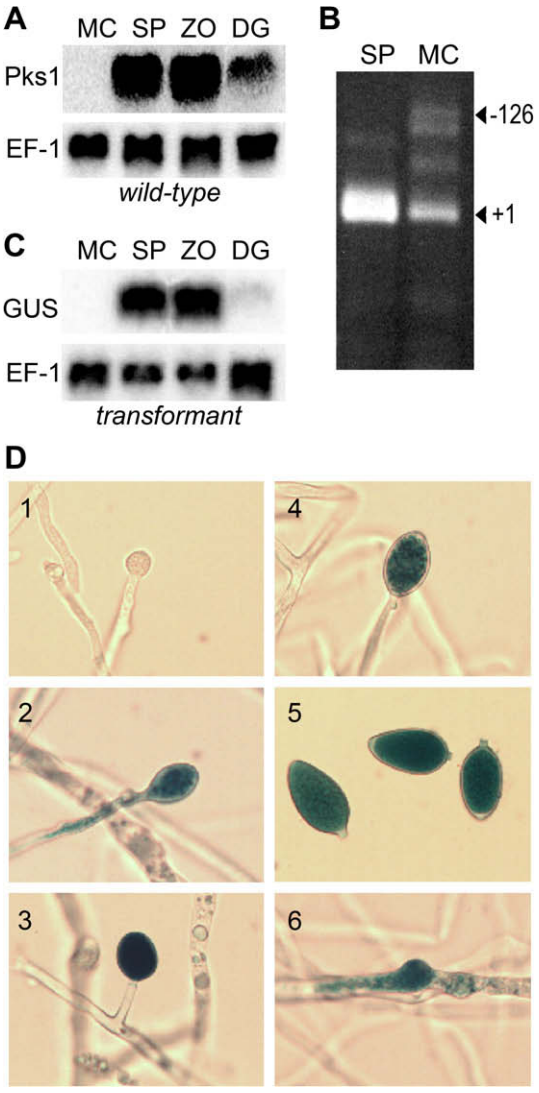

Fig. 1. Expression pattern of *Pks1*. (A) RNA blot analysis of wild-type *P. infestans*. RNA (5 μg) from non-sporulating mycelia (MC), sporangia (SP), sporangia undergoing zoosporogenesis (ZO), and sporangia directly germinating in rye-sucrose broth (DG) were hybridized with probes for *Pks1* (Pks1) or elongation factor-1α (EF-1). (B) 5′ RACE using RNA from sporangia (SP) or non-sporulating mycelia (MC). (C) RNA blot analysis of transformant expressing GUS driven by a plasmid p877, which contains 877 nt of DNA upstream of the major TSS. RNA is from the tissue types described in panel A, hybridized against GUS or EF-1. (D) GUS reporter assay using transformant grown in submerged rye-sucrose broth and then transferred onto agar media to induce sporulation. Illustrated is a young sporangial initial with no staining (1), activity in sporangiophore and developing sporangium (2), activity in a slightly older but still immature sporangium (3), sporangium approaching maturity based on the presence of an early-stage papilla (4), mature detached sporangia (5), and lobed hyphae occasionally seen in older cultures (6).

The minor RACE bands from sporangia started at −49 and −51, which are within an A-rich region that does not resemble an Inr. Due to the absence of an ATG between +1 and −51, these minor mRNAs encode the same protein as the major transcription product.

RACE products were also obtained from non-sporulating mycelia, even though no signal was detected in RNA blots (Fig. 1A and B). As no amplification occurred in the absence of reverse transcriptase (not shown), it can be concluded that a very low level of

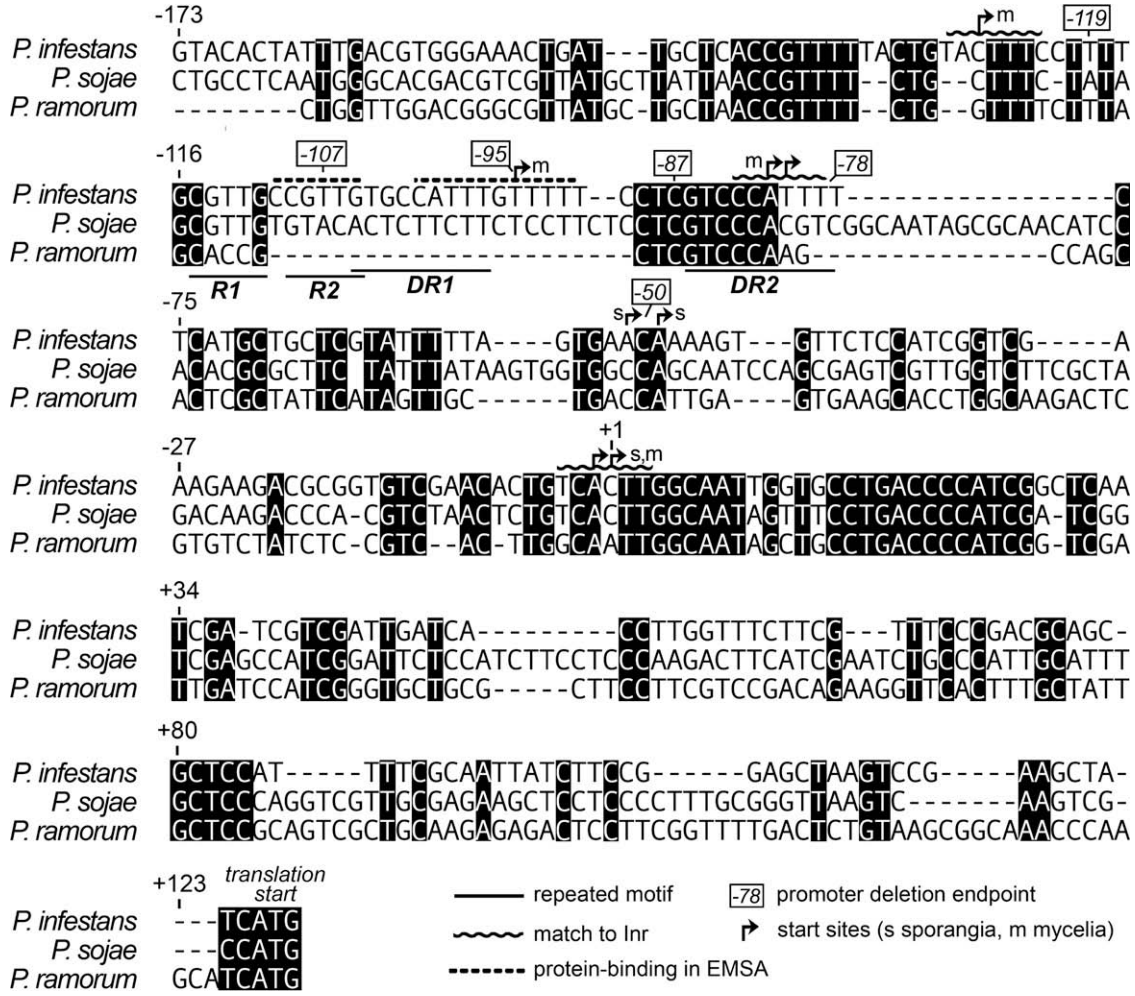

Fig. 2. Comparison of *Pks1* promoter with *P. ramorum* and *P. sojae* orthologs. Alignment was performed using CLUSTALW, with conserved bases indicated in black. No significant regions of alignment were observed upstream of nucleotide −173. Indicated are transcription start sites (TSS; bent arrow) from sporangia (s) or mycelia (m); endpoints of deletion constructs (boxed numbers), Inr-like sequences (wavy lines), repeats described in the text (R1, R2, DR1 and DR2), and regions showing specific binding in EMSA (dashed lines).

basal transcription of *Pks1* occurs in hyphae. One type of RACE product had the same 5′ terminus as the major band from sporangia, while the others mapped further upstream to positions −80 (T), −81 (A), −95 (T), and −126 (C). All of these are within an extremely T-rich region; the −58 to −137 interval, for example, contains 53% T. Potential Inr-like sequences are also within this region. For example, the minor initiation sites at −80 and −81 are within an Inr-like element, CCATTTT. The existence of the minor start sites illustrate some leakiness in the regulation of transcription, but may be insignificant to *Phytophthora* growth or development since the resulting transcripts would not be translated into a functional kinase due to an out-of-phase ATG at −73.

3.3. Deletion analysis identifies region required for sporangia-specific transcription

To search for the DNA element(s) activating *Pks1* during sporulation, the ability of truncated promoters to express GUS was measured. This involved screening sporulating cultures of stable transformants by histochemical staining, followed by RNA blot analysis of positive transformants (if any) to confirm that transcription started at the expected site within the promoter. Due to position effects, not all transformants obtained with the same plasmid expressed GUS at equal levels. Therefore, the conclusions reported represent a consensus drawn from analyses of multiple strains.

Tests of 5′ deletions indicated that a motif with a major positive effect on sporulation-specific transcription was near position −110. Promoters containing 525, 225, and 119-nt upstream of the TSS (p525, p225, and p119) conferred strong GUS staining in sporangia with no staining in hyphae. However, a plasmid with a larger deletion (p107) yielded only extremely low staining in the sporangia of some transformants and none in hyphae, suggesting that an important element was near the −107 to −119 region. Several further deletions exhibited the same phenotype (p95, p87, and p78), but when position −50 was reached by the deletions (p50) this residual staining was eliminated.

RNA blot data supported these findings. As shown in Fig. 3B and C, a strong spore-specific band of the expected size (2.1 kb) was observed in p119-containing transformants. However, transformants generated with the smaller promoters showed either no signal (p107, p95, and p78) or very weak bands (p87). An exception was one transformant generated with p107, which produced an aberrantly-sized (1.4 kb) transcript which may be due to promoter activity from a chromosomal locus flanking the insertion site. Some plasmids yielded weak GUS staining but no signal in the RNA blots (p107 and p95), indicating that expression was below the limit of detection. Also, a faint band of 1.6 kb was detected in many of the

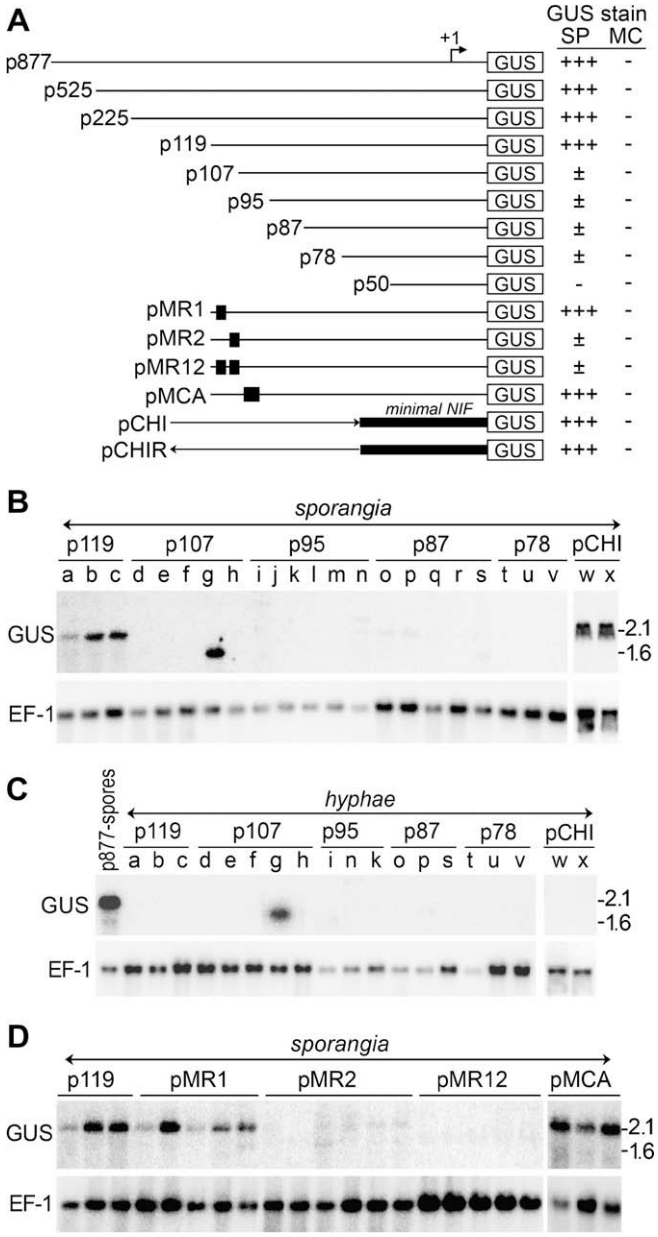

Fig. 3. Identification of functional regions by mutation analysis. (A) Phenotypes obtained with promoter-GUS fusion constructs in sporangia (SP) and mycelia (MC). The sizes of 5′ deletion plasmids correspond to their names, *i.e.* p225 contains 225-nt 5′ of the major TSS. Names with "M" contain mutagenized bases within repeat R1 (pMR1), R2 (pMR2), R1 and R2 (pMR12), or the CATTTGTT motif (pMCA). pCHI and pCHIR contain *Pks1* promoter sequences (−144 to −51) fused to minimal promoter in sense and antisense orientations, respectively. (B) RNA blot analysis of representative transformants using total RNA from sporangia and probes for GUS or EF-1. Letters a to x denote individual transformants. (C) RNA blot of same transformants as panel B but RNA is from non-sporulating hyphae, except for the first lane that shows RNA from sporangia as a control. (D) RNA blot of transformants containing p119 or versions of p119 that contain mutagenized promoter elements as shown in panel A.

transformants, which could represent weak cryptic promoter activity or non-specific hybridization.

Several conclusions can be drawn from these results. First, since no deletion caused constitutive transcription (*i.e.* activity in both hyphae and sporangia), the sporulation-specific expression of *Pks1* is determined primarily by positive and not negative regulation. Secondly, an important regulator binds between −119 and −107. Thirdly, elements with other functions may exist between −107 and −50. These may confer sporulation-specific expression independently of a sporulation-specific activator in the −119 to −107 interval; there are many precedents of autonomous transcription factors activating the same promoter (Hoshino and Fujii, 2007). Alternatively, the motif between −119 and −107 could be a specific enhancer of a sporulation-specific motif between −107 and −50.

As will be described below, additional experiments and bioinformatics analyses were used to address the scenarios listed above, leading to the conclusion that a CCGTTG motif at −110 was the critical sporulation-specific determinant. These included studying additional promoter mutations, identifying regions with protein binding activity using electrophoretic mobility binding assays (EMSA), and searching for motifs shared with other sporulation-induced genes.

3.4. A functional domain in the −119 to −107 interval spans a repeated motif

A notable feature of this region is the presence of two tandem repeats of CGTTG (R1 and R2 in Fig. 2), since repeats are often used to bind multimeric transcription factors. The role of the CGTTG repeats were tested by mutating them individually or in combination within p119, by making base changes (A to C, C to A, G to T, and T to G). A plasmid incorporating a mutated upstream repeat (pMR1) maintained expression in sporangia (Fig. 3D) and its pattern of GUS staining also indicated sporulation-specific transcription. However, when the 3′ repeat was mutated (Fig. 3D, pMR2), transcription was reduced to barely detectable levels. Mutating both repeats also blocked expression (Fig. 3D, pMR12). This indicates that the sporulation-specific transcription factor primarily binds the downstream repeat, or an overlapping sequence. The upstream repeat region apparently retains a low level of activity.

We also tested whether this region could drive transcription in a heterologous setting, by making a chimera with the 74-nt minimal *NifS* promoter (Ah Fong et al., 2007). By itself, this minimal promoter is transcriptionally inactive. However, adding a 94-nt fragment with the two CGTTG motifs (bases −144 to −51; pCHI) conferred robust sporulation-specific expression (Fig. 3B and C). Moreover, expression was observed in both forward or reverse orientations. However, smaller fragments from throughout the region did not enable GUS expression when fused to the minimal promoter, such as the 19-nt region between −117 and −99 which contained both CGTTG repeats. This was true even when random bases were added downstream of the fragments to maintain their normal spacing from the major TSS. The results suggest that multiple sites must interact to enable expression, due to a need to bind a complex of transcription factors or assemble chromatin in a specific configuration.

3.5. A sporulation-specific protein binds the CCGTTG at −110

EMSA was performed to test whether the region containing the CGTTG repeats might associate with a transcription factor, and define the precise boundaries of the binding site. When nuclear proteins from sporangia were incubated with a ^{32}P-labeled double-stranded oligonucleotide probe bearing the −117 to −100 region, three closely-spaced bands representing specific binding were observed based on their competition with unlabeled probe (Fig. 4A). For example, based on phosphorimager analysis their intensity was reduced by 39% using a 1:1 combination of radiolabeled and cold oligonucleotide, and 98% with a 1:100 ratio, which is consistent with a specific interaction. The nature of the specific bands, two high-intensity and a weaker middle band, is suggestive of a multi-protein complex or complexes. A smear of smaller and less-specific bands was also observed.

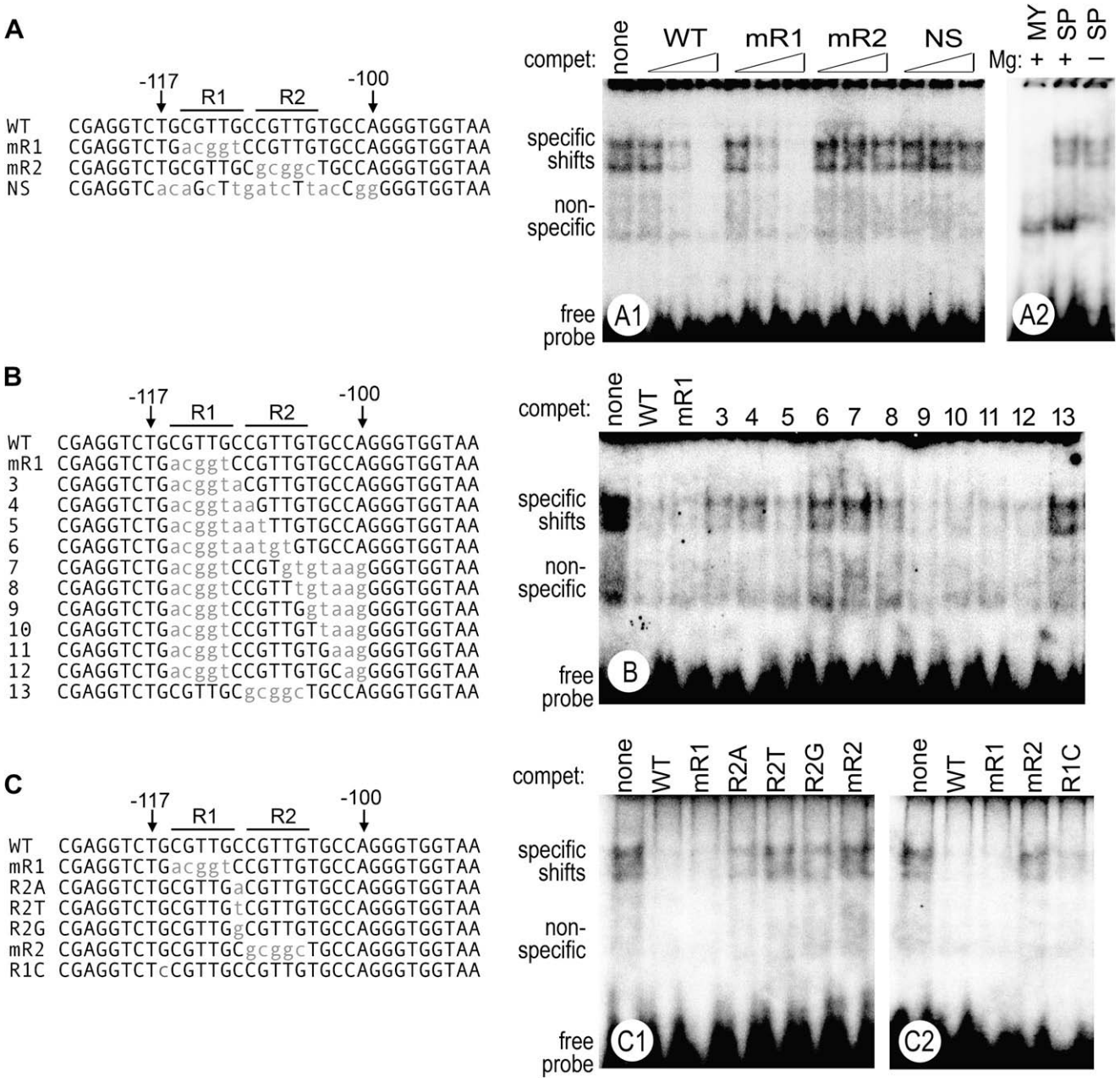

Fig. 4. Identification of CCGTTG as binding site by EMSA. Nuclear proteins were mixed with 1.6 ng of double-stranded ^{32}P-labeled oligonucleotides containing 18 nt of wild-type *Pks1* DNA flanked by random bases (WT) and cold competitors (compet). Sequences are shown left of the relevant gel images, with positions altered from wild-type in grey lower-case. The two repeats (R1 and R2) are highlighted. Unless indicated otherwise, assays used protein from sporangia at 0 mM Mg^{2+}. (A) Detection of specific binding near R2. Gel A1 indicates protein mixed with labeled WT and no competitor, wild-type competitor, competitors with mutations in R1 or R2 (mR1 and mR2), or non-specific competitor (NS). Competitors were 1×, 10×, and 100× the concentration of labeled probe. Gel A2 demonstrates that the retarding protein is sporulation-specific, based on testing extracts from non-sporulating mycelia (MY) and sporangia (SP). (B) Defining boundaries of binding site using WT probe and 25× excess of cold competitors with progressive mutations. (C) Refinement of binding site. Gel C1 tests effect of altering the C between R1 and R2. Gel C2 shows that changing the base upstream of R1 from G to C improves competition.

Only the 3′ CGTTG repeat (R2) was very important for binding, based on using cold oligonucleotides with altered sequences as competitors against the wild-type ^{32}P-labeled −117 to −100 region. For example, mutating the 5′ CGTTG repeat (R1) did not diminish competition compared to the unlabeled wild-type sequence (mR1 in Fig. 4B, C1, and C2). In contrast, oligonucleotides with a mutated R2 competed ineffectively (mR2 in Fig. 4C1 and C2; also lanes 8–10 in Fig. 4B). This is consistent with the GUS reporter data described above, which indicated that removing R1 by mutation or deletion within p119 did not eliminate sporulation-specific expression, while removing R2 drastically reduced expression.

Also consistent with the participation of repeat R2 in activating transcription during sporulation was the finding that its binding activity is sporulation-specific. As shown in Fig. 4A2, the three closely-spaced specific bands were obtained using nuclear proteins from sporangia, but not non-sporulating mycelia. Attempts to detect binding from mycelial extracts in EMSA performed under different conditions (altered salt concentration, Mg^{+2}, etc.) were unsuccessful. In contrast, the protein from the sporangial extract

showed binding under multiple conditions, as illustrated in Fig. 4A2 for EMSA performed with and without Mg^{+2}.

Specific binding was shown to be actually determined by a 6-nt motif (CCGTTG) encompassing R2, based on EMSA using competitors bearing progressive mutations (Fig. 4B). When mutations were made from the 5′ side inwards, competition was reduced when base changes reached the C residue before R2 (Fig. 4C, lane 3). When alterations moved from the 3′ side inwards, diminished competition started at the terminal G of R2 (Fig. 4C, lane 8). The site is therefore bordered by these bases, and defined as CCGTTG. This is consistent with the previous finding that R1, which is not within a CCGTTG, was unneeded for GUS expression.

To further test whether CCGTTG is the binding site, the region of the R1 CGTTG repeat (preceded by G in wild-type) was converted to CCGTTG and tested in an EMSA competition assay. This substantially increased its ability to bind the nuclear protein (Fig. 4C2, lane R1C). However, binding was slightly less than that observed against wild-type competitor (lane WT), indicating that residues flanking the CCGTTG also may affect protein interactions. That adjacent bases influence the binding of a transcription factor to its target is a long-known phenomenon (Arndt and Fink, 1986).

Interestingly, the CCGTTG is not well-conserved in other *Phytophthora* species, although somewhat related sequences are located nearby in *P. ramorum* and *P. sojae* (Fig. 2). Such variation may be explained by several factors. For example, the cognate transcription factors may bind distinct sequences, which would be consistent with rates of binding site evolution characterized within non-oomycete genera (Moses et al., 2003). Alternatively, there could be flexibility in the binding specificity of the transcription factor as is commonly seen with transcription factors in other species (Segal et al., 1999). To test this, base changes were made at the 5′ end of the R2-spanning CCGTTG motif. As shown in the EMSA competition assay in Fig. 4C1, DNA containing ACGTTG (lane R2A) and GCGTTG (lane R2G) still bound the nuclear protein although to a lower extent than CCGTTG (lane WT).

In EMSA, the three specifically retarded bands were observed consistently. This suggests that several independent proteins or complexes attach to the CCGTTG. Alternatively, a single complex may bind which partially dissociates during EMSA. To address the latter, different ionic strengths, Mg^{2+} concentrations, and binding temperatures were tested. However, in all cases the same three-banded pattern was obtained (not shown).

3.6. The −110 CCGTTG motif is over-represented in sporulation-induced promoters

As further support for the role of the CCGTTG in activating *Pks1* during sporulation, it was found to be highly enriched in co-expressed promoters. This was demonstrated by studying promoters from 388 *P. infestans* genes that like *Pks1* are induced >10-fold during sporulation (Judelson et al., 2008; Judelson, Narayan, and Ah Fong, unpublished results). For this analysis, a promoter was defined as 500 nt of DNA 5′ of the predicted open reading frame. CCGTTG was detected 128 times in the 388 sporulation-induced promoters (with matches in 104 promoters), or an average of 0.33 hits per promoter. Consistent with data from the GUS assays indicating that the sequence may function in both orientations, the motif was found in forward and reverse orientations 65 and 63 times, respectively. In contrast, only 0.14 occurrences per promoter were detected in 188 genes repressed >10-fold in sporangia versus mycelia, and 0.19 per promoter for all genes predicted from the genome sequence. This over-abundance of CCGTTG in sporangia-induced versus sporangia-repressed and total promoters were both highly significant ($P < 0.01$).

The CCGTTG motif also showed a positional bias in harmony with expectations for the binding site of a transcriptional regulator. When 1000 nt of sporulation-induced promoter DNA was examined, the sequence was most common within 400 nt of the start codon or about 300-nt upstream of the typical *Phytophthora* TSS (Fig. 5). Moreover, the most common location was about 200-nt upstream of the start codon which is the spacing seen in *Pks1*. In contrast, the motif had an unbiased distribution within all promoters and was slightly under-represented in the first 400 nt of sporulation-repressed promoters.

The subset of sporangia-induced genes that contain CCGTTG in their promoters and also have significant matches in GenBank are shown in Supplementary Table 1. The most prominent functions represented within the group included signal transduction (including nine protein kinases) and transporters (including six ion channels).

3.7. Potential elements in the T-rich region downstream of the CCGTTG motif

As noted earlier, sporulation-specific GUS expression resulted when a chimera was made between the −144 to −51 interval and the *NifS* minimal promoter, but no expression resulted using smaller fragments including the CCGTTG at −110. A likely reason is that the sporulation-specific CCGTTG-binding factor must interact with other proteins. The inability of many transcription factor sites to function autonomously or in an improper context is a common challenge when studying complex promoters (Chiang et al., 2007).

An alternative explanation is that CCGTTG (−110 to −105) is actually an enhancer of a sporulation-specific factor that binds further downstream. While this seems unlikely considering that CCGTTG is enriched in sporulation-specific promoters, this was nevertheless tested in transformants by mutagenizing blocks between −80 and −88, −89 to −96, and −97 to −104. Histochemical staining and RNA blot analysis indicated that all displayed sporulation-specific activity (not shown). This supports the premise that CCGTTG determines sporulation-specific expression, although this interpretation could be complicated if functional redundancy exists within the −80 to −104 region or elsewhere. The results also suggest that the 9-nt block between −81 and −89 that is conserved in *Phytophthora* is not the sporulation-specific determinant, although it could contribute quantitatively to expression. The 9-nt motif is not over-represented in sporulation-specific promoters.

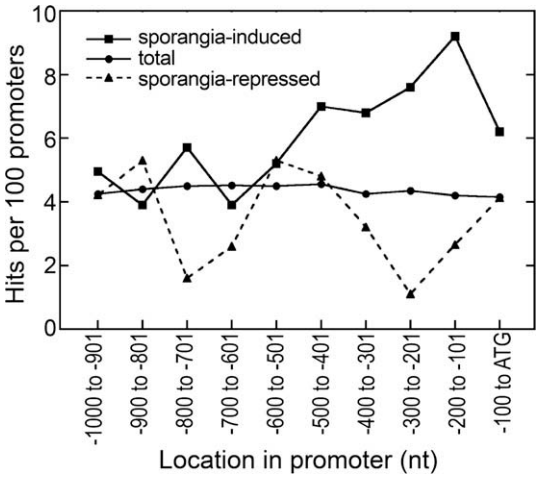

Fig. 5. Positional bias of CCGTTG motif. Results indicate frequency of detection of motif within 100-nt intervals going upstream of the ATG codon (+1). Promoter sets were from genes induced >10-fold in sporangia versus mycelia, total predicted *P. infestans* genes, and genes repressed >10-fold in sporangia versus mycelia. Searches were done in both plus and minus orientations.

3.8. The CATTTGTT at −101 associates with a nuclear protein but not sporulation-specific expression

As part of another strategy to identify functional elements 3′ of the CCGTTG, EMSA was performed using overlapping fragments spanning the region between −110 and −50. Binding activity was not detected with the 3′ half under a range of conditions. However, nuclear proteins interacted with the 5′ half (Fig. 6A). Two specific bands were detected, based on the ability to compete them away using cold probe.

Examination of the binding region revealed that it contained two similar 9-nt regions, GTGCCATTT and GTCCCATTT, which are labeled as DR1 and DR2 in Fig. 2. Their role in binding was assessed using an EMSA competition assay. DNA lacking DR1 or both repeats lost the ability to compete with wild-type probe (Fig. 6, panel A1; mDR1, mDR12) while mutating DR2 had little effect (mDR2). Thus, a protein specifically binds part or all of DR1. To define the binding consensus, progressive mutations were made inwards from the 5′ end (Fig. 6B, lanes 2–6) or 3′ end (7–13) of DR1. From the 5′ direction, competition was impaired severely when alterations reached the C at position −101, and the same occurred when mutations from the 3′ direction reached a T at −94. As mutating two additional T's at the 3′ end slightly reduced competition, the binding site is defined as CATTTGTTTT.

Several data suggested that this motif is not involved in sporulation-specific expression. First, when the site was mutated within p119 and tested in transformants, spore-specific transcription continued based on histochemical staining and RNA blot analysis (Fig. 3D, pMCA). Second, the interacting proteins detected by EMSA were present in nuclear extracts from both sporangia and non-sporulating hyphae. Third, the motif was not over-represented in sporulation-induced promoters, although it is acknowledged that not all stage-specific motifs will necessarily be over-represented.

When considering alternate roles for the CATTTGTT, the possibility of it binding the Inr complex was considered. This was because it is adjacent to a minor TSS, and would perfectly match the mammalian Inr if the G was replaced with C or T. Consequently, an EMSA competition assay was used to test if protein binding would be improved by altering it to the Inr consensus (Fig. 6, panel A2). However, competition was impaired when this base was changed to C, T, or A.

4. Discussion

Asexual sporulation in *P. infestans* is a complex process involving the expression of genes in coordinated temporal and spatial patterns. In addition to identifying putative transcription factor binding sites responsible for the regulated expression of *Pks1*, this study helps to define a specific point of the sporulation program by demonstrating that the gene is induced at a middle to late-stage, prior to formation of the basal plug. The definition of other stages were aided by our prior studies of genes such as *Cdc14* which is expressed prior to sporangiophore emergence, *PufA* which starts being transcribed in mid-stage sporangiophores, and *NifS* for which expression is typically detected coincident with maturation (Ah Fong and Judelson, 2003; Cvitanich and Judelson, 2003; Tani et al., 2005). *Pks1* therefore appears to be activated between *PufA* and *NifS*. Identifying the regulatory hierarchy that controls such genes is central to understanding this important developmental pathway. By analogy, studies of sporulation in species as diverse

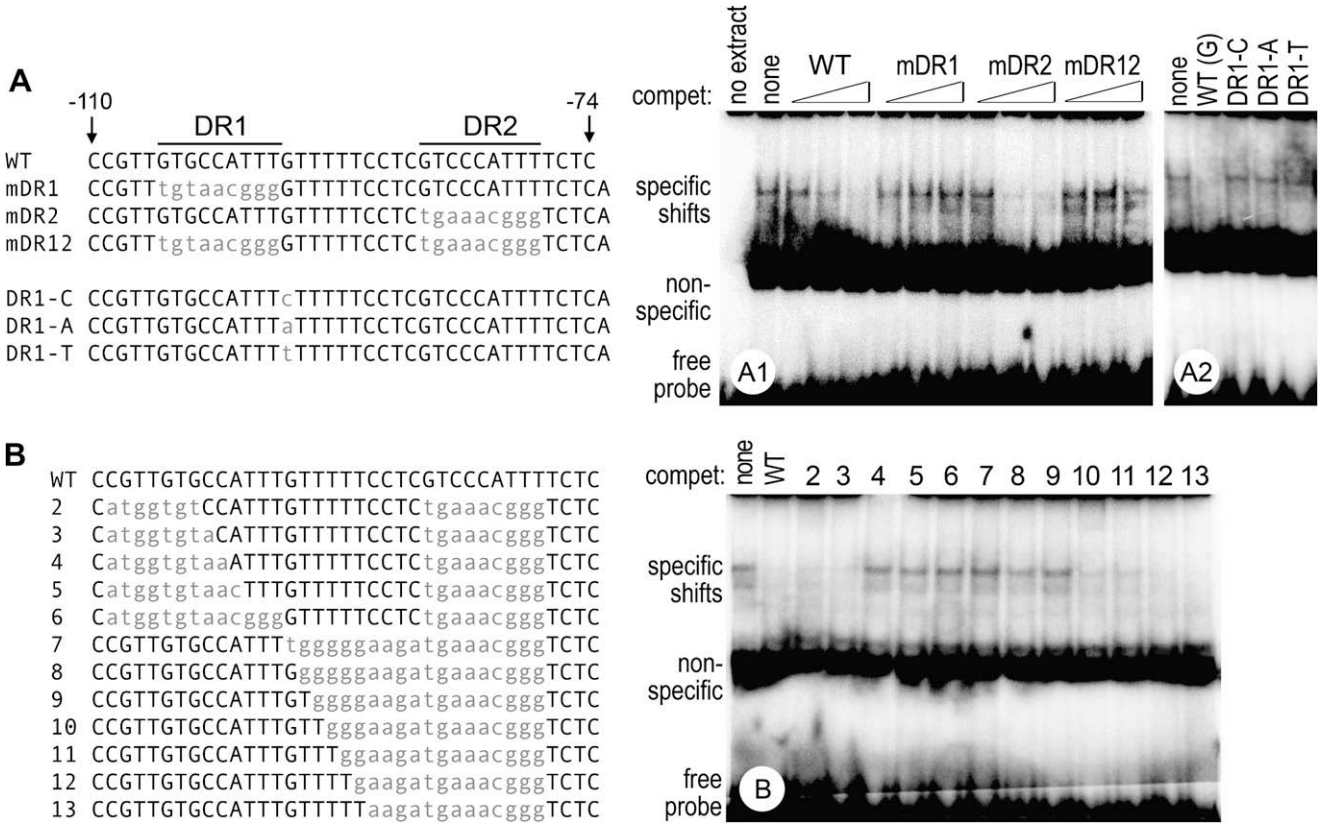

Fig. 6. Binding of nuclear proteins from sporangia to CATTTGTT region in EMSA, using labeled wild-type DNA and a 100× excess of the cold competitors shown to the left. (A) Detection of specific binding to downstream repeat 1 (DR1). Gel A1 demonstrates that proteins specifically bind DR2, based on competition with wild-type DNA (WT) or DNA bearing mutations in the 5′ or 3′ repeats (mDR1, mDR2), or both (mDR12). Gel A2 shows results from competition assays using DNA in which the base at −96 was altered. (B) Refinement of binding site by using competitors with progressive mutations.

as *Bacillus* and *Aspergillus* have revealed how they are controlled largely through the coordinated activity of transcription factors that act in sequential cascades (Adams et al., 1998; Kroos, 2007).

The −110 CCGTTG sequence from *Pks1* is only the third oomycete motif to be identified as being responsible for developmentally regulated transcription, and the second participating in sporulation-specific expression. It is distinguished in several ways from the other sporulation-specific motif, CTYAAC, which was recently identified in the *Cdc14* promoter (Ah Fong et al., 2007). While the *Cdc14* motif is not highly over-represented in other sporulation-induced promoters, the *Pks1* motif is significantly over-represented. A second difference is that the binding activity in EMSA against the −110 *Pks1* motif was more complicated than against the *Cdc14* motif, as a multi-banded phenotype was detected in the former. Moreover, while the putative transcription factor associating with the *Cdc14* motif was in both sporulating and non-sporulating tissues, the protein binding activity against the *Pks1* motif was sporulation-specific. These differences are consistent with *Cdc14* acting at a very early tier of the regulatory hierarchy of sporulation, and *Pks1* at a subsequent stage along with a consortia of other genes needed for late-stage functions such as zoospore behavior. *Cdc14* is likely activated by transcription factors preformed in hyphae, while regulators expressed coincident with *Cdc14* or soon thereafter presumably activate through transcriptional or post-translational means the transcription factor complex that turns on *Pks1*.

The identities of these putative transcription factors are unknown, although efforts to purify them are underway. DNA microarray data has revealed candidates for regulators of *Pks1*, such as several Myb family transcription factors that are induced during sporulation (Judelson et al., 2008). In other eukaryotes, many transcription factors required for development are transcribed in a stage-specific manner (Chen et al., 2002), and this would be an economical strategy for *P. infestans* since sporangia development represents only a short portion of its life cycle.

The role of the −110 CCGTTG motif in inducing *Pks1* during sporulation is very strongly supported by the promoter mutation studies, its over-representation and positional bias in sporulation-induced promoters, and its binding to a sporulation-specific protein complex in EMSA. However, we were unable to directly prove this using the chimeric promoter approach although a larger region was positive for sporulation-specific activity in that assay. This is probably since other features of the *Pks1* promoter are required for the cognate transcription factor to bind, which is a common situation in complex promoters (Chiang et al., 2007; Harbison et al., 2004). Several such features can be proposed. These include the CATTTGTT at −101 which was shown to bind a nuclear protein, and the evolutionarily conserved 9-nt sequence at −89.

An additional feature that may aid the −110 CCGTTG motif is the T-rich region between −137 and −58 which likely influences chromatin structure. In budding yeast, for example, a T-rich zone in the *HIS3* promoter establishes a nucleosome-free domain that increases transcription factor accessibility (Mai et al., 2000; Sekinger et al., 2005). The T-rich region in *Pks1* is also predicted to be devoid of nucleosomes by the Nucleosome Position Prediction tool (genie.weizmann.ac.il/pubs/nucleosomes06/segal06_prediction.html; Segal et al., 2006). The majority of *P. infestans* promoters do not contain a similarly extensive T-rich zone. Assembly of the transcriptional apparatus may also be aided by the multiple Inr-like sequences within the region. The presence of Inr-like elements combined with low nucleosome occupancy may also explain why we could detect minor transcripts initiating in the region in hyphae.

Besides containing multiple Inr-like motifs, the *Pks1* promoter contained other repeats. For example, the CCGTTG at −110 is just downstream of a similar sequence (repeats R1 and R2 in Fig. 2), and a different repeat resides just downstream (repeats DR1 and DR2). Originally it was thought that R1 and R2 might bind a dimerized transcription factor complex or have additive effects. However this does not seem likely based on the functional analyses and EMSA data. In addition, neither *P. ramorum* or *P. sojae* contain the repeats. The lack of evolutionary conservation, however, is insufficient to exclude present or past function of the repeats (or any motif for that matter) based on studies of orthologous promoters in other species, where the duplication, deletion, inversion, and other gross changes in transcription factor binding sites have been documented (Chuzhanova et al., 2000).

Although this study focused on the mechanism of *Pks1* transcription, we predict that the protein kinase it encodes plays a role in sporangia maturation or zoospores. The latter is most likely since our RNA blots indicate that *Pks1* mRNA persists through zoosporogenesis, and because zoosporogenesis and encystment are known to proceed without requiring new RNA or protein synthesis (Clark et al., 1978; Penington et al., 1989). The kinase may therefore cause post-translational modifications required for activities such as swimming, chemotaxis, or encystment. Attempts to define the role of *Pks1* through gene silencing by expressing inverted repeats of its coding sequences in transformants have so far been unsuccessful, based on examinations of over 150 transformants. However, interactions of *Pks1* with a Gβ-like WD-domain adapter protein are indicated by yeast two-hybrid and co-immunoprecipitation studies (unpublished data). *Pks1* may therefore mediate specific protein–protein interactions in maturing sporangia or zoospores, just as other regulators likely associate to form the complex that activates *Pks1* transcription during an intermediate stage of sporulation.

Acknowledgments

This work was supported by the National Research Initiative program of the United States Department of Agriculture Cooperative State Research, Education, and Extension Service. We are grateful to the Broad Institute of MIT and Harvard for providing access to the *P. infestans* sequence database.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.fgb.2009.02.006.

References

Adams, T.H., Wieser, J.K., Yu, J.-H., 1998. Asexual sporulation in *Aspergillus nidulans*. Microbiol. Mol. Biol. Rev. 62, 35–54.

Ah Fong, A., Judelson, H.S., 2003. Cell cycle regulator *Cdc14* is expressed during sporulation but not hyphal growth in the fungus-like oomycete *Phytophthora infestans*. Mol. Microbiol. 50, 487–494.

Ah Fong, A., XIang, Q., Judelson, H.S., 2007. Architecture of the sporulation-specific Cdc14 promoter from the oomycete *Phytophthora infestans*. Eukaryot. Cell 6, 2222–2230.

Arndt, K., Fink, G.R., 1986. GCN4 protein, a positive transcription factor in yeast, binds general control promoters at all 5′ TGACTC 3′ sequences. Proc. Natl. Acad. Sci. USA 83, 8516–8520.

Baldauf, S.L., Roger, A.J., Wenk-Siefert, I., Doolittle, W.F., 2000. A kingdom-level phylogeny of eukaryotes based on combined protein data. Science 290, 972–977.

Bogre, L., Okresz, L., Henriques, R., Anthony, R.G., 2003. Growth signalling pathways in *Arabidopsis* and the AGC protein kinases. Trends Plant Sci. 8, 424–431.

Chen, W. et al., 2002. Expression profile matrix of *Arabidopsis* transcription factor genes suggests their putative functions in response to environmental stresses. Plant Cell 14, 559–574.

Chiang, D.Y., Nix, D.A., Shultzaberger, R.K., Gasch, A.P., Eisen, M.B., 2007. Flexible promoter architecture requirements for coactivator recruitment. BMC Mol. Biol. 7, 16.

Chuzhanova, N.A., Krawczak, M., Nemytikova, L.A., Gusev, V.D., Cooper, D.N., 2000. Promoter shuffling has occurred during the evolution of the vertebrate growth hormone gene. Gene 254, 9–18.

Clark, M.C., Melanson, D.L., Page, O.T., 1978. Purine metabolism and differential inhibition of spore germination in *Phytophthora infestans*. Can. J. Microbiol. 24, 1032–1038.

Cvitanich, C., Judelson, H.S., 2003. A gene expressed during sexual and asexual sporulation in *Phytophthora infestans* is a member of the Puf family of translational regulators. Eukaryot. Cell 2, 465–473.

Fry, W.E., 2008. *Phytophthora infestans:* the plant (and R gene) destroyer. Mol. Plant Pathol. 9, 385–402.

Harbison, C.T. et al., 2004. Transcriptional regulatory code of a eukaryotic genome. Nature 431, 99–104.

Hardham, A.R., Hyde, G.J., 1997. Asexual sporulation in the oomycetes. Adv. Botan. Res. 24, 353–398.

Hoshino, A., Fujii, H., 2007. Redundant promoter elements mediate IL-3-induced expression of a novel cytokine-inducible gene, cyclon. FEBS Lett. 581, 975–980.

Judelson, H.S., 1993. Intermolecular ligation mediates efficient cotransformation in *Phytophthora infestans*. Mol. Genom. Genet. 239, 241–250.

Judelson, H.S. et al., 2008. Gene expression profiling during asexual development of the late blight pathogen *Phytophthora infestans* reveals a highly dynamic transcriptome. Mol. Plant–Microbe Interact. 21, 433–447.

Judelson, H.S., Blanco, F.A., 2005. The spores of *Phytophthora:* weapons of the plant destroyer. Nat. Microbiol. Rev. 3, 47–58.

Juven-Gershon, T., Hsu, J.Y., Kadonaga, J.T., 2006. Perspectives on the RNA polymerase II core promoter. Biochem. Soc. Trans. 34, 1047–1050.

Kim, K.S., Judelson, H.S., 2003. Sporangia-specific gene expression in the oomyceteous phytopathogen *Phytophthora infestans*. Eukaryot. Cell 2, 1376–1385.

Kroos, L., 2007. The *Bacillus* and *Myxococcus* developmental networks and their transcriptional regulators. Ann. Rev. Genet. 41, 13–39.

Lee, T.I., Young, R.A., 2000. Transcription of eukaryotic protein-coding genes. Ann. Rev. Genet. 34, 77–137.

Mai, X., Chou, S., Struhl, K., 2000. Preferential accessibility of the yeast his3 promoter is determined by a general property of the DNA sequence, not by specific elements. Mol. Cell Biol. 20, 6668–6676.

McLeod, A., Smart, C.D., Fry, W.E., 2004. Core promoter structure in the oomycete *Phytophthora infestans*. Eukaryot. Cell 3, 91–99.

Moses, A.M., Chiang, D.Y., Kellis, M., Lander, E.S., Eisen, M.B., 2003. Position specific variation in the rate of evolution in transcription factor binding sites. BMC Evol. Biol. 3, 19.

Penington, C.J., Iser, J.R., Grant, B.R., Gayler, K.R., 1989. Role of RNA and protein synthesis in stimulated germination of zoospores of the pathogenic fungus *Phytophthora palmivora*. Exp. Mycol. 13, 158–168.

Randall, T.A., Judelson, H.S., 1999. Construction of a bacterial artificial chromosome library of *Phytophthora infestans* and transformation of clones into *P. Infestans*. Fungal Genet. Biol. 28, 160–170.

Segal, D.J., Dreier, B., Beerli, R.R., Barbas, C.F., 1999. Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5′-GNN-3′ DNA target sequences. Proc. Natl. Acad. Sci. USA 96, 2758–2763.

Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thastrom, A., Field, Y., Moore, I.K., Wang, J.P., Widom, J., 2006. A genomic code for nucleosome positioning. Nature 442, 772–778.

Sekinger, E.A., Moqtaderi, Z., Struhl, K., 2005. Intrinsic histone–DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. Mol. Cell 18, 735–748.

Smale, S.T., 2001. Core promoters: active contributors to combinatorial gene regulation. Genes Dev. 15, 2503–2508.

Tani, S., Judelson, H.S., 2006. Activation of zoosporogenesis-specific genes in *Phytophthora infestans* involves a 7-nucleotide promoter motif and cold-induced membrane rigidity. Eukaryot. Cell 5, 745–752.

Tani, S., Kim, K.S., Judelson, H.S., 2005. A cluster of NIF transcriptional regulators with divergent patterns of spore-specific expression in *Phytophthora infestans*. Fungal Genet. Biol. 42, 42–50.