Identification of a cis-acting factor modulating the transcription of FUM1, a key fumonisin-biosynthetic gene in the fungal maize pathogen Fusarium verticillioides

Identification of a cis-acting factor modulating the transcription of FUM1, a key fumonisin-biosynthetic gene in the fungal maize pathogen Fusarium verticillioides

V. Montis, M. Pasquali, I. Visentin, , P. Karlovsky, F. Cardinale

Abstract: Fumonisins, toxic secondary metabolites produced by some *Fusarium* spp. and *Aspergillus niger*, have strong agro-economic and health impacts. The genes needed for their biosynthesis, named FUM, are clustered and co-expressed in fumonisin producers. In eukaryotes, coordination of transcription can be attained through shared transcription factors, whose specificity relies on the recognition of *cis*-regulatory elements on target promoters. A bioinformatic analysis on FUM promoters in the maize pathogens *Fusarium verticillioides* and *Aspergillus niger* identified a degenerated, over-represented motif potentially involved in the *cis*-regulation of FUM genes, and of fumonisin biosynthesis. The same motif was not found in various FUM homologues of fungi that do not produce fumonisins. Comparison of the transcriptional strength of the intact *FUM1* promoter with a synthetic version, where the motif had been mutated, was carried out *in vivo* and *in planta* for *F. verticillioides*. The results showed that the motif is important for efficient transcription of the *FUM1* gene.

Abbreviations

- dw, dry weight;
- FB1, B-series fumonisin 1;
- GFP, green fluorescent protein;
- P_{FUM1}, *FUM1* promoter;
- P_{FUM1}mod, mutated *FUM1* promoter;
- RT-qPCR, quantitative reverse-transcriptase PCR;
- TF, transcription factor;
- TFBS, transcription-factor binding site;
- TSS, transcription start-site

1. INTRODUCTION

Filamentous fungi are able to produce a multitude of secondary metabolites endowed with bioactive properties. Secondary metabolites are organic molecules that are not essential for growth, differentiation or reproduction of their producer but increase its fitness in certain ecological contexts (Fox and Howlett, 2008). In fungi, most genes needed for the synthesis of secondary metabolites are organized in clusters. Physical clustering may allow the genes to be co-regulated by chromatin modification (Keller et al., 2005), and it may also facilitate horizontal gene transfer between species (Rosewich and Kistler, 2000).

The fumonisin (FUM) gene cluster contains genes involved in the synthesis of fumonisins. Fumonisins are polyketide-derived mycotoxins produced mainly by the maize pathogens *Fusarium verticillioides* and*Fusarium proliferatum* (teleomorphs *Gibberella moniliformis* and *Gibberella intermedia*) (Munkvold, 2003), by some strains of *Fusarium oxysporum* (Rheeder et al., 2002) and by some species of *Aspergillus*, including*A*. *niger* and related species (Frisvad et al., 2007 and Varga et al., 2010). Fumonisin B_1 (FB1) is the major fumonisin produced by *F*. *verticillioides* and *F*. *proliferatum* (Bartók et al., 2006 and Nelson et al., 1993). Since fumonisins are natural contaminants of food and feed and a serious threat to animal and human health (reviewed by Voss et al., 2007 and Wan Norhasima et al., 2009), a considerable effort is being devoted to a better understanding of the regulation of their synthesis.

FUM genes are clustered in *F*. *oxysporum* as they are in *F*. *verticillioides* with the number, order and orientation of the genes being the same (Proctor et al., 2003 and Proctor et al., 2008). Two other *Fusarium*species host a full version of the FUM cluster, only missing the *FUM17* orthologue (encoding one of the two putative ceramide synthases in the cluster): *F*. *proliferatum* and *Fusarium fujikuroi* (B. Tudzynski, personal communication). Both fumonisin-producing and non-producing strains were identified in these species (Munkvold, 2003 and Proctor et al., 2004). In *A*. *niger* instead, the cluster contains only 11 homologues of the 17 FUM genes identified (*FUM1*, *FUM6–10*, *FUM13–15*, *FUM19* and *FUM21*) (Pel et al., 2007). Here too, both producing and non-producing strains were identified (Frisvad et al., 2007). The difference in gene composition and order of the *Fusarium vs*. *Aspergillus* FUM clusters likely reflects a long history of independent evolution from a common ancestor. Indeed, the FUM cluster in *A*. *niger* probably originated by horizontal gene transfer from a common ancestor prior of the divergence between *F*. *verticillioides* and *F*.*oxysporum* (Khaldi and Wolfe, 2011). Interestingly, a gene-by-gene phylogenetic analysis revealed a total of 13 putative orthologues of FUM genes (*FUM1*, *FUM6–7*, *FUM10–19*, hereby called FUM-like), scattered in the genomes of five other Euascomycetes: *Cochliobolus heterostrophus* (Kroken et al., 2003), *Fusarium graminearum*, *Neurospora crassa*, *Magnaporthe grisea*, and *Aspergillus nidulans* (Khaldi and Wolfe, 2011). Physical clustering of genes may also influence the evolution of their regulatory sequences, likely due to the possibility of tight co-regulation of expression by localized events of chromatin remodelling, and to the activity of pathway-specific regulators embedded in the cluster (Bayram and Braus, 2012 and Yu and Keller, 2005). One proposed mechanism of regulation of fumonisin synthesis is indeed at the epigenetic level (Visentin et al., 2012), as for other secondary metabolic gene clusters

(Palmer and Keller, 2010 and Reyes-Dominguez et al., 2010). However, chromatin-controlled regulation normally coexists with a shared set of transcription factors (TFs), which bind to the promoters of co-regulated genes by recognizing common *cis*-regulatory elements acting as TF-binding sites (TFBSs). Knowledge of the interplay between *trans*-factors and their cognate *cis*-sequences is crucial to understanding the molecular underpinnings of the coordinated regulation of gene expression.

A number of proteins with a demonstrated or postulated global function in signaling and chromatin remodeling were connected to fumonisin biosynthesis, together with broad-domain TFs with genes outside of the FUM cluster (reviewed in Picot et al. (2010)). At present, only one narrow-domain TF, encoded by*FUM21*, has been characterized as required for fumonisin biosynthesis. Fum21 is a predicted $Zn(II)_2Cys_6$zinc-finger protein found in all five FUM gene clusters sequenced so far (Brown et al., 2007, Pel et al., 2007 and Proctor et al., 2008; B. Tudzynski, personal communication). No *cis*-regulatory elements with TFBS function have been reported in FUM gene promoters yet. One plausibly successful approach to identify them would be based on their conservation among co-regulated FUM promoters within the same genome as well as among orthologous FUM clusters. Extension of sequence conservation to regions adjacent to TFBSs by functional constraints may facilitate identification (Doniger et al., 2005). A motif-discovery algorithm has been already successfully used for the identification of *cis*-regulatory elements in the promoters of *F. graminearum* genes induced during trichothecene synthesis (Seong et al., 2009).

In this work, our aim was to identify putative *cis*-regulatory elements involved in the coordinated transcription of FUM genes and to validate the prediction experimentally, with the broader goal of deepening our understanding of the molecular mechanisms controlling fumonisin synthesis. We detected two similar DNA sequences, over-represented within the promoters of the clustered FUM genes of the fumonisin-producing fungi *F. verticillioides* and *A. niger*, while not in the scattered FUM-like genes of *C. heterostrophus*, *F.graminearum*, *A. nidulans*, *M. grisea* and *N. crassa*, which do not produce fumonisins. These sequences are also present in the regulatory sequences of the clustered FUM genes in *F. oxysporum*, *F. proliferatum* and *F.fujikuroi*. The involvement of one of them in the transcriptional regulation of *FUM1*, encoding a key polyketide synthase (Proctor et al., 1999), was demonstrated *in vivo* in *F. verticillioides*.

2. MATERIALS AND METHODS

2.1. In silico analyses

To investigate motif over-representation in the promoters of FUM genes, 1 kb upstream of the TSS of all FUM and FUM-like genes were used as query in RSAT (http://rsat.ulb.ac.be/) (Thomas-Chollier et al., 2008 and van Helden, 2003). *FUM20* was excluded from the analysis since its function is unknown and the gene shares no significant similarity with any previously described gene in the NCBI database, although it was proven to be transcribed (Brown et al., 2005). Sequences were downloaded from the Broad Institute for all fungi

(http://www.broadinstitute.org/scientific-community/data) with the exception of sequences of the *A.niger* strain ATCC1015 (Andersen et al., 2011), which were obtained from the JGI database (http://genome.jgi-psf.org/Aspni1/Aspni1.home.html), and of the upstream sequence of the *FUM1*-like from *C.heterostrophus* (Kroken et al., 2003), from the JGI database (http://genomeportal.jgi-psf.org/CocheC5_3/CocheC5_3.home.html).

F. verticillioides and *F. oxysporum* sequences (NCBI accessions AF155773.5 and EU449979.1, respectively) are from strains 7600 and O-1890, respectively. *A. niger* sequences presented in the manuscript are from strain CBS 513.88 (Pel et al., 2007).

F. graminearum, *A. nidulans*, *M. grisea* and *N. crassa* sequences were retrieved according to a recent article (Khaldi and Wolfe, 2011). *F. fujikuroi* (IMI58289) and *F. proliferatum* (strain ET1) FUM cluster sequences were retrieved from a recent genome sequencing program (B. Tudzynski and co-workers, unpublished). For pattern discovery, the oligo-analysis (words) and the dyad-analysis (spaced pairs) tools of RSAT were used. For pattern matching, DNA-pattern (string description) and genome-scale DNA-pattern, tools of the same software were used, and feature maps generated. Background frequencies for all these analyses were calculated genome-wide on 1000 bp upstream of the TSSs for all the analyzed fungi, and for the range 100–600 or 600–1000 bp. As negative controls, three sets of 20 random sequences (1 kb upstream of TSSs) from the genomes of *F. verticillioides*, *F. graminearum* and *Saccharomyces cerevisiae* were created by the random gene-selection tool (RSAT), and a subsequent pattern-matching computation for the motif found was performed on them. To prevent methodological biases (Tompa et al., 2005), the query sequences from *F.verticillioides* and *A. niger* were also searched for conserved regulatory sequences using the stand-alone version of Weeder (among the MoD Tools; http://159.149.109.9/modtools/) (Pavesi et al., 2004 and Pavesi et al., 2006). Background frequencies for this analysis were calculated on genome-wide sequences upstream of TSSs of both *F. verticillioides* and *A. niger* by using the Frequency Maker program of Weeder. The Motif Locator and the Motif *p*-value Calculator tools were used to compute significance *P*-values and to locate instances of a given motif in a set of sequences, respectively.

The databases JASPAR Core Fungi (http://jaspar.genereg.net/), TRANSFAC (http://www.gene-regulation.com/cgi-bin/pub/databases/transfac/search.cgi), and YEASTRACT (http://www.yeastract.com/) were employed for a first characterization of the 6-mers found.

2.2. Nucleic acids methods

A mutated version (P_{FUM1}mod) of *FUM1* putative promoter sequence (P_{FUM1}) was created by replacing the main 6-mer sequence in P_{FUM1} (CGGATA) with the *Xba*I target sequence (TCTAGA). We replaced $P_{FUM1}$with this modified version within the construct pCAM-P_{FUM1}::GFP (green fluorescent protein) (Visentin et al., 2012) in several cloning steps, using a set of three modified primer pairs (Xba_seq3_for and seq3_Sma_rev; Aat_seq1_for and

seq1_Xba_rev; Xba_seq2_for and seq2_Xba_rev, see *Table* S1) to create the final construct pCAM-P_{FUM1}mod::GFP, which was fully sequenced (*Fig.* S1). For cloning purposes, PCR reactions were set up in 25 µl (final volume) of Pfu Reaction Buffer containing 1.5 mM $MgSO_4$(Promega), dNTPs 0.2 mM each, primers 0.5 µM each, DNA template 1 ng, Pfu DNA polymerase 0.375 U and GoTaq DNA polymerase 0.875 U (Promega). PCR conditions were the following: 94 °C for 2 min; then 35 cycles at 94 °C for 45 s, appropriate annealing temperature (*Table* S1) for 45 s, 72 °C for 1 min; and final extension at 72 °C for 5 min. PCR products were purified from agarose gels by NucleoSpin Extract II kit (Macherey–Nagel). All restriction enzymes were from Fermentas, while T4 DNA ligase was from Promega and DNA phosphatase was from Amersham Pharmacia Biotech.

For the extraction of genomic DNA to be used in Southern blot analyses, we used a variant of the CTAB method (Brandfass and Karlovsky, 2008) with slight modifications, and analyzed 5 µg for each strain as described previously (Visentin et al., 2012). For induction kinetics and plant infection experiments, total RNA was extracted using the NucleoSpin RNA Plant kit (Macherey–Nagel) according to the manufacturer's instructions. For all other samples, RNA was obtained from 15 mg of fungal mycelium by a slightly modified Chang's method (Chang et al., 1993). Quality and quantity of RNA were checked with the RNA 6000 Nano LabChip kit in the Agilent 2100 Bioanalyzer (Agilent Technologies). Only samples with spectrophotometric ratios of $A_{260}/A_{280} \geqslant 1.9$, $A_{260}/A_{230} \geqslant 2$, and RNA integrity number values $\geqslant 7$ were used. After extraction, RNA samples were stored at −80 °C. cDNA was synthesized starting from 10 µg of total RNA with the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems), following the manufacturer's instructions. 1 µl of each sample was analyzed by NanoDrop ND 1000 equipped with the ND-1000 V3.7 software (Thermo Fisher Scientific). cDNA samples were stored at −20 °C.

For quantitative RT-PCR experiments (RT-qPCR) analyses, absolute quantification (copy number) of target DNA was performed on the StepOnePlus Real-Time PCR System (Applied Biosystems) with the FAM dye filter set in combination with SYBR Green dye (SYBR Green PCR Master Mix, Applied Biosystems). RT-qPCR reactions were set up in 10 µl (final volume) of SYBR Green PCR Master Mix, with primers (500 nM each) and cDNA (10 ng). Transcripts from each of the two targets and one reference gene (*FUM1*, *GFP* and*TUB2*) were quantified in each of three independent biological replicates per experimental condition, in analytical triplicates. For standard curves, target sequences from each of the three genes were amplified from genomic DNA of *F. verticillioides* (primer pairs FUM1_for/_rev, GFP_for/_rev and TUB2_for/_rev in Table S1, the same used for RT-qPCR) (Flaherty et al., 2003) and cloned in pGEM-T Easy (Promega). Five-point curves were obtained by serially diluting (1:5) the obtained constructs; concentrations ranged from 25 pg to 40 fg of each plasmid for the three target sequences, corresponding to 7.30×10^6–1.17×10^4 copies/reaction for *FUM1* and to 7.04×10^6–1.13×10^4 copies/reaction for *GFP* and *TUB2*. Differences in copy number with respect to quantity are due to the difference in length of the PCR amplification products (*Table* S1). All standard samples were run in analytical triplicates. PCR conditions were the following: 95 °C for 10 min; 41 cycles at 95 °C for 15 s and 60 °C for 1 min; and 95 °C for 15 s. Differences in the amplification efficiencies of the reference (*TUB2*) and the target gene

transcripts (*FUM1*and *GFP*) in RT-qPCR experiments were considered acceptable if ⩽15% within each run and between independent experiments, with slopes of the standard curves ranging from −3.5 to −3.2, R^2 values larger than 0.98, and C_T values between 18 and 30 for target genes and about 23 for *TUB2* in all experiments. Transcripts of *FUM1* and *GFP* were normalized using *TUB2* values and ratios of *GFP* to *FUM1* transcripts were calculated.

2.3. Agrobacterium-mediated transformation of F. verticillioides and characterization of fungal transformants

F. verticillioides strain VP2 (Visentin et al., 2009) was used to create transgenic strains named Fv-P_{FUM1}mod via *Agrobacterium tumefaciens*-mediated transformation (Takken et al., 2004), with minor modifications (Visentin et al., 2012). Briefly, competent *A. tumefaciens* strain EHA 105 was prepared by $CaCl_2$ treatment (Yong et al., 2006) and transformed by heat shock with pCAM-P_{FUM1}mod::GFP plasmid DNA (Fig. S1). Rifampicin (50 μg/ml) and Kanamycin (50 μg/ml) were used to select positive transformants. Vector acquisition by *Agrobacterium* was checked by PCR. Fungal transformation was carried out as described before (Visentin et al., 2012). Hygromycin-resistant fungal colonies were tested for the presence of the Hygromycin-resistance cassette by PCR, using FTA cards according to the manufacturer's instructions (Whatman–Schleicher and Schuell). For each positive Fv-P_{FUM1}mod transformant, a single conidium was picked manually, transferred onto a plate of Czapek medium supplemented with Hygromycin (100 μg ml^{-1}) and Cefotaxime (200 μM), and incubated for 4 d at 25 °C. Monoconidial stock colonies were stored at 4 °C onto slants of Malt Extract Agar (Sigma). Southern blot and hybridization analysis were performed as follows. 5 μg of genomic DNA from wild-type *F. verticillioides* strain VP2 and 13 independent Fv-P_{FUM1}mod transformants were analyzed as in Visentin et al. (2012). Transformant Fv-P_{FUM1}mod-C was excluded from analyses because of its altered colony morphology. Single-copy, independent Fv-P_{FUM1}mod transformants (indicated by capital letters B, D E and F in Fig. S2) were selected for further analyses. Their toxigenicity was compared to the one of the wild-type parental strain by quantifying FB1 production on cracked maize kernels after 18 d of culture at 25 °C in the light (10^6 spores in 3 ml sterile water were used to inoculate 8 g of autoclaved cracked kernels, in triplicate). Concentration of FB1 was determined by HPLC/MS as described previously (Visentin et al., 2012).

2.4. Fungal cultures and media, plant infection experiments

For Southern blot analysis, *F. verticillioides* strains were grown for 7 d at 25 °C on Czapek medium (fumonisin non-inducing. In 1 l: 2 g $NaNO_3$, 1 g KH_2PO_4, 0.5 g $MgSO_4 \cdot 7H_2O$, 0.5 g KCl, 0.01 g $FeSO_4 \cdot 7H_2O$, 30 g sucrose). For quantitative analysis of P_{FUM1}*vs*. P_{FUM1}mod transcriptional activity, the transgenic control strains Fv-P_{FUM1}-1 and -3 (Visentin et al., 2012) and the Fv-P_{FUM1}mod-B -D -E and -F transformants were grown in triplicate (18 d at 25 °C in the dark and in stationary conditions) on Czapek medium or on two different inductive substrates for fumonisin production: GYAM (0.12 M glucose, 50 mM malic acid, 8 mM L-Asp,

1.7 mM NaCl, 4.4 mM K_2HPO_4, 2 mM $MgSO_4$, 8.8 mM $CaCl_2$, 0.05% w/v yeast extract, pH 3) and fructose medium (in 1 l: 0.5 g malt extract, 1 g yeast extract, 1 g peptone, 1 g KH_2PO_4, 0.3 g $MgSO_4 \cdot 7H_2O$, 0.3 g KCl, 0.05 g $ZnSO_4 \cdot 7H_2O$, 0.01 g $CuSO_4 \cdot 5H_2O$, 20 g fructose). For induction kinetics, wild-type, Fv-P_{FUM1}-1 and Fv-P_{FUM1}mod-E were grown for 7, 14 or 21 d at 25 °C in the dark and in stationary conditions in liquid fructose medium. All liquid media (40 ml per replicate, in 100 ml flasks) were inoculated with a starting concentration of 10^5 conidia ml^{-1}. Hyphal mats were harvested by filtration (0.8 μm pores, Millipore), frozen in liquid nitrogen and stored at −80 °C. For seedling infection experiments, maize kernels (Silver Queen hybrid, Syngenta Seeds) were incubated overnight in sterile distilled water containing, per ml, 10^6 conidia of transformants Fv-P_{FUM1}-1, Fv-P_{FUM1}mod-E or Fv-P_{FUM1}mod-F. Infected and control, non-infected seeds were germinated on water-agar plates (1 week at RT, 12 h dark/light cycle), then sown in sterilized soil and grown in the greenhouse for one additional week (10–25 °C and about 12 h dark/light cycle). At harvest, rootlets were quickly washed under tap water, frozen in liquid nitrogen and stored at −80 °C until RNA extraction. Three independent pools of three rootlets each were analyzed for each sample. The experiment was conducted thrice.

2.5. Statistical analysis

Fisher's tests (http://www.langsrud.com/fisher.htm) for assessing hexamers representation within the two analyzed fungal genomes were calculated on occurrences in the FUM clusters as compared to the full genome, both on 1000 bp and in 100–600 bp regions upstream of TSSs. Bonferroni correction was applied and corrected P values $\leqslant 0.05$ were considered significant. Analysis of variance between the *TUB2*-normalized *GFP*/*FUM1* ratios was performed using the LSD test, with 95% or 99% *P*-values. Data were analyzed by the SGwin software (StatGraphics Plus for Windows version 2.1, Statistical Graphics Corp.).

3. RESULTS

3.1. *In silico* analysis reveals similar over-represented sequences within promoters of clustered FUM genes in fumonisin producers F. verticillioides and A. niger

In silico pattern-discovery (by RSAT software) detected two similar 6-mers significantly over-represented in the 1000 bp upstream of transcription start-sites (TSSs) of clustered FUM genes of both fungi, as compared to background values obtained on the corresponding regions of all protein-coding genes (Tables 1 and S2). All sequences will be reported hereafter in a simplified way, i.e. only in their forward orientation, but they are meant to indicate both forward and reverse orientation. The sequences CGGATA, present in 87.5% of FUM promoters in *F. verticillioides*, and CGGCTA, present in 63.6% of FUM promoters in *A. niger*, were suggested as most significant by RSAT oligo-analysis and were the only ones in common with

the Weeder analysis output on the two FUM clusters (Table S2). They both contain the CGG triplet. Instead, running the pattern-discovery tool (RSAT) on FUM-like genes scattered in the genomes of the fumonisin non-producers*F*. *graminearum*, *M*. *grisea*, *A*. *nidulans* and *N*. *crassa* identified different over-represented hexamers than in the promoters of clustered FUM genes in *F*. *verticillioides* and *A*. *niger* (Tables 1 and S2). Notably, the CGG triplet was not present anymore within the suggested hexamers. The motif was absent also from the putative regulatory sequence of the *FUM1*-like gene of *C*. *heterostrophus* (Kroken et al., 2003). When the dyad-search algorithm was performed as implemented in RSAT, the most significant output was the CGGN(0)ATA dyad, which is equivalent to the sequence identified in *F*. *verticillioides*. In *A*. *niger* no significant dyads were found (Table S2).

Organism	6-mer	*P*	*E*	SI
Most highly-scoring sequences by oligo-analysis (RSAT)				
F. verticillioides	CGGATA|TATCCG	$1.5e^{-08}$	$3.2e^{-05}$	4.49
A. niger	CGGCTA|TAGCCG	$1.9e^{-04}$	$3.8e^{-01}$	0.42
F. graminearum	CTCCAC|GTGGAG	$1.7e^{-06}$	$3.4e^{-03}$	2.47
A. nidulans	AAAAAA|TTTTTT	$5.1e^{-12}$	$1.0e^{-08}$	7.99
N. crassa	TCTTCA|TGAAGA	$4.43e^{-04}$	$4.0e^{-01}$	0.1
M. grisea	AAAAAA|TTTTTT	$2.6e^{-15}$	$5.2e^{-12}$	11.28

Table 1 Most significant 6-mers in the FUM cluster as identified by oligo analysis (RSAT) on 1 kb sequences upstream of TSSs in FUM genes of the fumonisins producers *F*. *verticillioides* and *A*. *niger*, and in the promoters of FUM-like genes scattered in the genome of four other fungal species, which do not produce fumonisins (*F*. *graminearum*, *A*. *nidulans*, *N*. *crassa*, *M*. *grisea*). Full results of RSAT analysis are provided in Table S2. The second column shows the over-represented sequences identified, in both orientations. *P*-value is the probability to observe the same result by chance (false positive). *E*-value is the number of patterns with the same level of over-representation which would be expected by chance alone. Significance index (SI) is the $-\mathrm{Log}_{10}$(*E*-value), and the higher values are associated to the more significant patterns.

The sequences identified in *F*. *verticillioides* (CGGATA) and in *A*. *niger* (CGGCTA) were analyzed for their distribution and abundance within both fungal genomes; Fisher's Exact Test with Bonferroni correction was used to test over-representation of each of the two sequences in the genome of the other fungal species by comparing sequence occurrences in the regulatory regions of the cluster and of the full corresponding genomes. Results confirmed the significant over-representation ($P \leqslant 0.05$) of each sequence in both genomes when the analysis was limited to the region 0.1–0.6 kb upstream of the TSSs, where TFBSs are most probably located (Table S3). It is noteworthy that both sequences display the best significance in the species where they were initially discovered, while being significantly over-represented also in the FUM cluster of the other species. This latter finding is compatible with the hypothesis that both sequences contribute to *cis*-regulation of the FUM genes, possibly as TFBSs, being under selective pressure for conservation in phylogenetically distant species. Combining the main sequences as obtained from the above analyses, a degenerated motif

CGGMTA can be defined (with M = A or C). This is common to the two species and is over-represented as such in the FUM clusters of the two species ($P < 0.05$, Tables S2 and S3). Fig. S3 localizes visually the CGGMTA sequence and its orientation within the FUM promoters of *F.verticillioides* and *A. niger*, and shows a similar distribution of occurrences per gene in the two species. Motif distribution was also determined in the *A. niger* strain ATCC1015, which does not produce fumonisins (Andersen et al., 2011). No differences in position and number of occurrences of the motif were found between the two *A. niger* strains (data not shown). Localization of the motif in the regulatory regions of the clustered FUM genes of fumonisin-producing *F. oxysporum* and *F. proliferatum* strains, and of a non-producing *F. fujikuroi* isolate (the latter two sets of sequences were kindly provided by B. Tudzynski, Westfälische Wilhelms-Universität Münster, Germany) was also carried out (Fig. S3). The lack of genome sequences for these fungi did not make it possible to evaluate the significance of the motif over-representation, but it is clear that almost all FUM genes, if clustered, have at least one motif in the region 0.1–0.6 kb upstream of the TSS, both in fumonisin-producers (*F. verticillioides*, *F. oxysporum*, *F.proliferatum*, *A. niger*) and non-producers (*F. fujikuroi*).

When the 6-mers identified were searched as such within public databases specific for TFBSs (TRANSFAC, JASPAR and YEASTRACT), no matches were found. However, when binding stringency was lowered to 80% or less in the JASPAR database, the main sequence of *F. verticillioides* appeared as recognizable by several *S. cerevisiae* TFs belonging to the group of the zinc (or binuclear) cluster proteins (Table S4). This protein family, unique to the fungal kingdom, is characterized by a zinc-finger domain of the $Zn(II)_2Cys_6$ type, where two Zn atoms are coordinated by six Cys residues in the DNA-recognition domain. Binding sites for this family of TFs invariably contain a CGG triplet (MacPherson et al., 2006).

Overall, our *in silico* results suggest that the CGGMTA motif identified on the basis of its abundance within the FUM promoters might be the conserved core of a *cis*-regulatory element. Possibly, this motif is part of a TFBS specific for a zinc-cluster TF involved in fumonisin biosynthesis.

3.2. A version of *F. verticillioides* P_{FUM1} mutated in the main sequence drives significantly less transcription than native P_{FUM1}

The main sequence CGGATA is present twice in the *FUM1* promoter of *F. verticillioides*, P_{FUM1} (at −643 and −390 bp from TSS; Table 2). To validate the results obtained *in silico*, we replaced both instances with *Xba*I restriction sites (TCTAGA) to generate a modified version of P_{FUM1} (P_{FUM1}mod). This was inserted upstream of a GFP-encoding reporter gene (Fig. S1) in a construct then transferred into *F. verticillioides*. A set of 14 transformants (Fv-P_{FUM1}mod-A to -O) was generated by *A. tumefaciens*-mediated transformation in two independent experiments. Single-copy, independent transformants Fv-P_{FUM1}mod-B, -D, -E and -F (Fig. S2) with wild-type colony morphology were retained for further investigation. Toxigenicity assays on cracked maize kernels revealed that these transformants produced fumonisins at levels comparable to the parental strain (300–500 μg FB1 per g of substrate; data not shown).

Gene	Putative function	CGGATA|TATCGG OCC	Start position (relative to TSS)
FUM21	Zn_2Cys_6 transcription factor	2	−978, −131
FUM1 (*FUM5*)	Polyketide synthase	2	−643, −390
FUM6	Cytochrome P450 monooxygenase/reductase	3	−815, −116, −85
FUM7	Alcohol dehydrogenase	2	−244, −213
FUM8	Oxoamine synthase	2	−217, −150
FUM3 (*FUM9*)	Dioxygenase	4	−828, −647, −228, −110
FUM10	Fatty acyl-CoA synthetase	2	−127, −9
FUM11	Tricarboxylate transporter	–	–
FUM2 (*FUM12*)	Cytochrome P450 monooxygenase	3	−906, −827, −499
FUM13	Short-chain dehydrogenase/reductase	3	−367, −316, −299
FUM14	Peptide synthetase/condensation domains	–	–
FUM15	Cytochrome P450 monooxygenase	1	−818
FUM16	Fatty acyl-CoA synthetase	1	−152
FUM17	Ceramide synthase	1	−183
FUM18	Ceramide synthase	1	−115
FUM19	ABC transporter	1	−161

Table 2 Summary of the number of occurrences (OCCs) and positions relative to the TSSs of the main sequence discovered by oligo analysis in the clustered FUM genes of *F. verticillioides*.

GFP transcripts were quantified in one non-inducing (Czapek) and two fumonisin-inducing media (GYAM and fructose medium). To evaluate if the rate of expression driven by P_{FUM1} was changed by the mutation of the two putative TFBSs, Fv-P_{FUM1}mod transformants were compared to strains Fv-P_{FUM1}-1 and -3. These bear a cognate expression cassette with the intact P_{FUM1} sequence coupled to a GFP reporter gene (Visentin et al., 2012). After 18 d of growth in the fumonisin non-inducing medium, all transformants showed detectable amounts of *GFP* transcript, while the transcript of the endogenous *FUM1* gene was almost undetectable (Fig. 1A). *GFP* expression on this medium was higher to a statistically significant degree ($P < 0.01$) in the control transformants Fv-P_{FUM1}-1 and -3 with respect to the transformants containing the P_{FUM1}mod::GFP expression cassette (Fig. 1A). The trend was confirmed in the fumonisin-inducing media:*GFP* expression was lower in all the Fv-P_{FUM1}mod compared to the Fv-P_{FUM1} transformants (data not shown). The results are better visualized by *GFP*/*FUM1* transcript ratios, which are significantly lower in the latter set of transformants compared to the former ($P < 0.01$) (Fig. 1B and C). The time course of *GFP* and*FUM1* transcript levels in Fv-P_{FUM1}mod-E was assessed on fructose medium at 7, 14 and 21 d post-inoculation and compared to strain Fv-P_{FUM1}-1. Impairment of P_{FUM1} activity by the mutation of the putative TFBSs approximately by a factor of 3 was consistent over the

whole incubation period (Fig. 1D). These differences were also observed for strains colonizing maize rootlets, confirming that the mutation has an effect also under physiologically relevant conditions (Fig. 2). Comparison of Fv-P_{FUM1}*vs.* a Fv-P_{FUM1}mod strain in fructose medium revealed that *GFP* transcript was less abundant than transcript of the endogenous*FUM1* in the transformant containing the mutated P_{FUM1}, but more abundant than endogenous *FUM1*transcript in the transformant containing native, unmodified P_{FUM1} (Fig. S4). The difference was consistent and significant at all three time points. Summarizing, on all synthetic substrates tested as well as during plant infection, targeted mutation of the two putative *cis*-regulatory elements CGGATA from the promoter of*FUM1* significantly reduced the level of transcript generated by the promoter ectopically inserted in *F.verticillioides* genome.

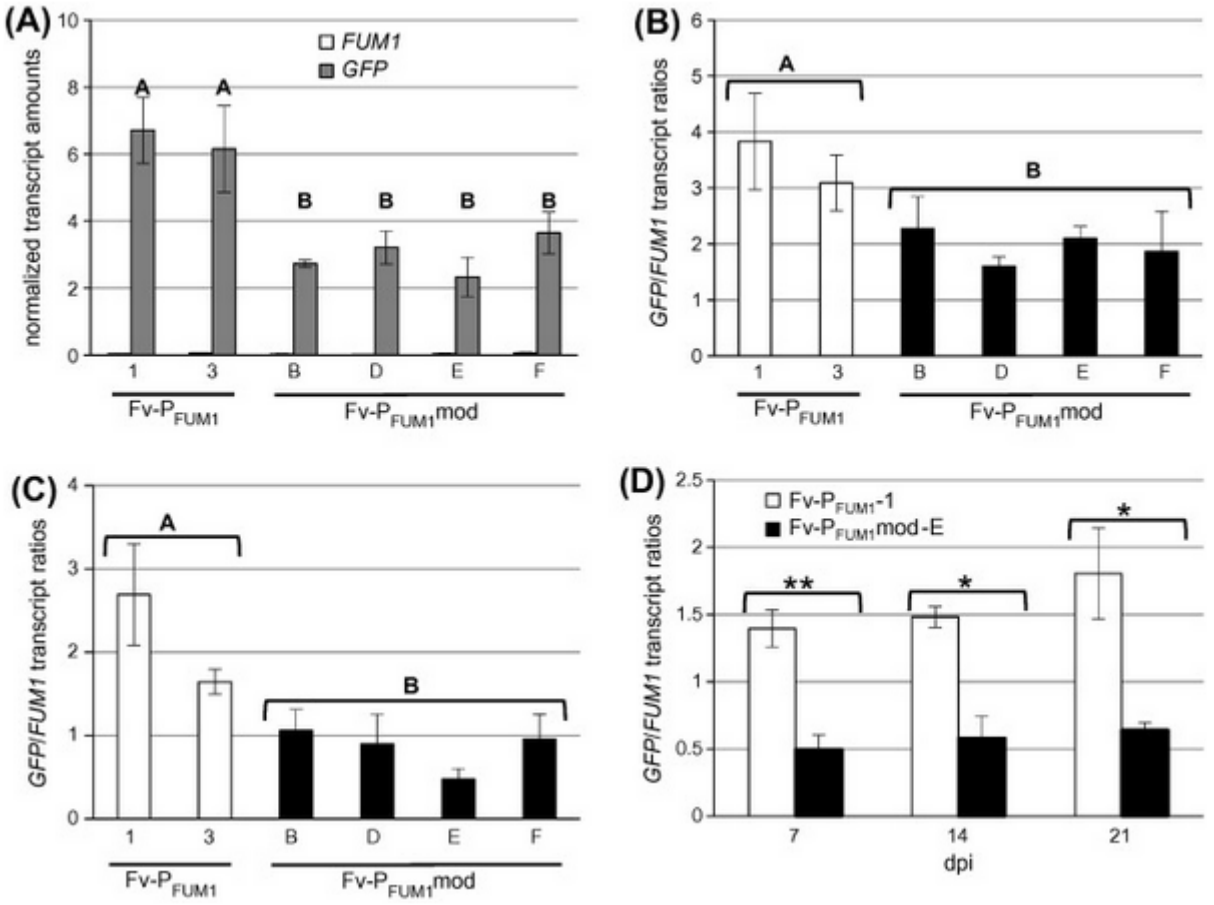

Figure 1 Mutations in the main sequence CGGATA of *F . verticillioides* impairs *FUM1* -promoter (P_{FUM1}) activity *in vivo* . Transcript abundance for *FUM1* and *GFP* from *F . verticillioides* transformants Fv-P_{FUM1}-1 and -3 and Fv-P_{FUM1}mod-B, -D, -E, -F grown for 18 days in (A) Czapek (non-inducing medium for fumonisins); (B) GYAM (fumonisin-inducing) and (C) fructose medium (fumonisin-inducing). Results for *GFP* and *FUM1* were obtained by absolute RT-qPCR and then normalized over *TUB2* transcript quantities before being divided by each other. For statistical analysis, Fv-P_{FUM1} transformants were either individually (A) or collectively (B and C) compared with Fv-P_{FUM1}mod strains. Different letters indicate a statistically significant difference (LSD test, $P < 0.01$). (D) *GFP /FUM1* transcript ratios in *F . verticillioides* transformants Fv-P_{FUM1}-1 and Fv-P_{FUM1}mod-E during a 3-point kinetics in fructose medium (fumonisin-inducing). For each time-point, significant differences are indicated by stars. $\bigstar = P < 0.05;\ \bigstar\bigstar = P < 0.01$. dpi = days post-inoculation.

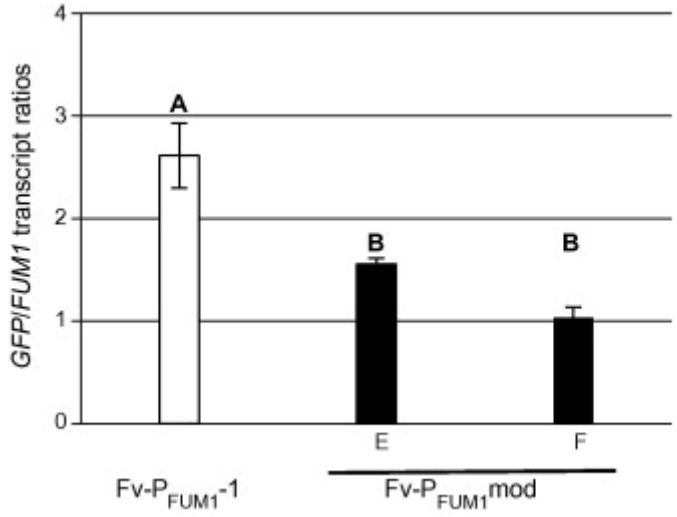

Figure 2 Mutations in the main sequence CGGATA of *F. verticillioides* impairs *FUM1*-promoter (P_{FUM1}) activity during plant infection.*GFP*/*FUM1* transcript ratios in *F. verticillioides* transformants Fv-P_{FUM1}-1, Fv-P_{FUM1}mod-E and P_{FUM1}mod-F were assessed in infected maize rootlets 1 week after sowing. Results for *GFP* and *FUM1* were obtained by absolute RT-qPCR and then normalized over *TUB2* transcript copy number, before being divided by each other. Significant differences are indicated by different letters on top of bars (LSD test, $P < 0.05$).

4. DISCUSSION

4.1. A candidate cis-regulatory element is found in promoters of clustered FUM genes, and a regulatory role for the main sequence is confirmed experimentally in *F. verticillioides*

In silico analyses revealed a set of two 6 bp-long sequences common to clustered FUM genes in the fumonisin-producing fungi *F. verticillioides* and *A. niger*, which can be described by the common motif CGGMTA. Depending on the fungus, one of the two sequences, the most abundant, can be regarded as the main sequence; while the other is less represented and can be defined as an auxiliary or sister sequence. Neither of these sequences was found to be significantly present in the promoters of the FUM-like genes scattered in the genomes of fungi that do not produce fumonisins (*F. graminearum*, *M. grisea*, *N. crassa*, *A.nidulans*, *C. heterostrophus*).

To test the involvement of the putative *cis*-regulatory element in the regulation of the FUM cluster, we focused on the contribution of the main sequence found in *F. verticillioides* to the expression of the key biosynthetic gene *FUM1*. Comparison of the intact *FUM1* promoter with a synthetic version in which the sequence was mutated revealed that under all conditions, *in vitro* as well as *in planta*, the transcriptional strength of the mutated promoter was significantly reduced. The transcript of the *GFP* gene driven by the intact P_{FUM1} promoter was more abundant than the transcript of the endogenous *FUM1* gene; as commented elsewhere, this might be due to lack of epigenetic control on the ectopic P_{FUM1} (Visentin et al., 2012)

However, the transcript level of the GFP driven by *FUM1* promoter deprived of the putative *cis*-regulatory element was significantly lower, *in vivo*, than that of the endogeneous *FUM1*, suggesting that positive control of *FUM1* via the main putative *cis*-regulatory element plays a crucial role together with epigenetically mediated control. Since the main sequence is present in nearly all FUM gene promoters, it is likely involved in the control of the FUM gene cluster as a whole in *F. verticillioides*. As in trichothecene biosynthesis (Seong et al., 2009), comparative genomics proved here successful in identifying putative TFBSs, confirming that statistical approaches for canonical motif finding perform best on sets of genes where co-regulation is strictly required (as in toxin-biosynthetic clusters).

4.2. FUM clusters in closely and distantly-related fungal species show localized conservation for the putative cis-regulatory element

The significant local over-representation of the degenerated motif CGGMTA with respect to the whole genome was confirmed in the −100 to −600 regions of clustered FUM genes for the two fumonisin producers for which whole reference genomes are available. This positional conservation suggests functional constraints on the putative TFBSs (Hare and Palumbi, 2003) and is perfectly compatible with the hypothesis of a remote horizontal gene transfer event (Khaldi and Wolfe, 2011). The sequence was also found to be similarly distributed in the regulatory regions of clustered FUM genes in a fumonisin-producing *F. oxysporum*strain, and in *F. proliferatum* and *F. fujikuroi*. The lack of whole reference genome sequences for these species and strains did not allow estimation of the statistical significance of this finding. However, the number and distribution of the instances of the putative *cis*-regulatory element is similar to those observed on *F. verticillioides* and *A. niger*, further suggesting its association with the regulatory regions of clustered FUM genes. We expect that when whole background genomes become available, it will be possible to understand if the core motif identified here should be further degenerated or changed for these fungi. Finally, the overall comparable number and distribution of the occurrences in the FUM promoters of the *F. fujikuroi*and *A. niger* ATCC1015 strains analyzed, which bear the cluster but do not produce fumonisins, may imply that an epistatic level of repression on the cluster must exist and be regulated differently than in fumonisin-producing strains.

The fact that the same motif is significantly over-represented in *F. verticillioides* and *A. niger* is surprising, since *cis*-regulatory elements are typically divergent in sequence between fungi belonging to different classes (Gasch et al., 2004). This finding is even more intriguing since the conditions inducing fumonisin synthesis in the two fungi are completely different (Frisvad et al., 2007); fumonisin production may indeed serve as a model for the diversification of metabolic regulation in spite of the retention of shared *cis*-regulatory elements and of the key TF. However, our observation also supports the hypothesis that orthologous TFs (most probably, but not necessarily, a narrow-domain TF located within the FUM cluster) may recognize these conserved *cis*-regulatory elements in the two species. Binding sites of TFs belonging to the Zinc binuclear cluster and possessing a $Zn(II)_2Cys_6$ motif, which are only found in fungi, share the CGG core contained in the motif identified in this

study (Campbell et al., 2008 and MacPherson et al., 2006). We therefore hypothesize that the TF binding to the putative *cis*-regulatory elements is a Zinc binuclear TF with a $Zn(II)_2Cys_6$ motif. Indeed, several Zinc binuclear TFs from *S. cerevisiae* may recognize the primary sequence of *F. verticillioides*, as seen when the stringency for the binding constraints was lowered to 80% in searching the JASPAR database. Fum21 belongs to this same family (Brown et al., 2007), as many other narrow-domain TFs specific for gene clusters in secondary metabolism (Bok et al., 2006, Fox et al., 2008,Wiemann et al., 2009 and Woloshuk et al., 1994). *FUM21* is present in all FUM clusters analyzed (Khaldi and Wolfe, 2011 and Pel et al., 2007). The most parsimonious hypothesis is therefore that Fum21 is the TF binding the identified motif, assuming that this is a TFBS. *fum21* Knock-out mutants of *F. verticillioides*accumulate only very limited amounts of *FUM1* transcripts under fumonisin-inducing conditions (Brown et al., 2007). In our experimental system, we observed residual activity by P_{FUM1}mod (although significantly impaired with respect to P_{FUM1}). Together, these results suggest that the effect of mutating the putative TFBSs of P_{FUM1} is weaker than that of deleting *FUM21*. We reasoned that this may be due to a incomplete inactivation of the binding site, if (as likely) the hexamer is part of a larger binding domain. If this is the case, residual binding activity may account for a certain extent of activity by P_{FUM1}mod. As an obvious alternative hypothesis, another TF may be involved, which is not absolutely required for *FUM1* transcription. However, of course, more experiments are needed to prove or disprove this hypothesis.

Acknowledgments

We wish to thank Prof. G. Tamietti for financial and D. Valentino for technical support, respectively; M. Beinhoff and K. Döll for help in Southern blot analysis and FB1 quantification, respectively; B. Amato for assistance in statistical analysis; U. Güldener and B. Tudzynski for sharing unpublished FUM regulatory sequences from *F. proliferatum* and *F. fujikuroi*; G. Munkvold for critical reading of the manuscript.

REFERENCES

Comparative genomics of citric-acid-producing Aspergillus niger ATCC 1015 versus enzyme-producing CBS 513.88, Genome Res., 21 (2011), pp. 885–897

T. Bartók, et al., Detection of new fumonisin mycotoxins and fumonisin-like compounds by reversed-phase high-performance liquid chromatography/electrospray ionization ion trap mass spectrometry, Rapid Commun. Mass Spectrom., 20 (2006), pp. 2447–2462

O. Bayram, G.H. Braus, Coordination of secondary metabolism and development in fungi: the velvet family of regulatory proteins, FEMS Microbiol. Rev., 36 (2012), pp. 1–24

J.W. Bok, et al., Secondary metabolic gene cluster silencing in Aspergillus nidulans, Mol. Microbiol., 61 (2006), pp. 1636–1645

C. Brandfass, P. Karlovsky, Upscaled CTAB-based DNA extraction and real-time PCR assays for Fusarium culmorum and F. graminearum DNA in plant material with reduced sampling error, Int. J. Mol. Sci., 9 (2008), pp. 2306–2321

D.W. Brown, et al., The Fusarium verticillioides FUM gene cluster encodes a Zn(II)2Cys6 protein that affects FUM gene expression and fumonisin production, Eukaryot. Cell, 6 (2007), pp. 1210–1218

D.W. Brown, et al., Comparative analysis of 87,000 expressed sequence tags from the fumonisin-producing fungus Fusarium verticillioides, Fungal Genet. Biol., 42 (2005), pp. 848–861

R.N. Campbell, et al, Metabolic control of transcription: paradigms and lessons from Saccharomyces cerevisiae, Biochem. J., 414 (2008), pp. 177–187

S. Chang, et al., A simple and efficient method for isolating RNA from pine trees, Plant Mol. Biol. Rep., 11 (1993), pp. 113–116

S.W. Doniger, et al., Identification of functional transcription factor binding sites using closely related Saccharomyces species, Genome Res., 15 (2005), pp. 701–709

J.E. Flaherty, et al., PAC1, a pH-regulatory gene from Fusarium verticillioides, Appl. Environ. Microbiol., 69 (2003), pp. 5222–5227

E. Fox, et al., A Zn(II)2Cys6 DNA binding protein regulates the sirodesmin PL biosynthetic gene cluster in Leptosphaeria maculans, Fungal Genet. Biol., 45 (2008), pp. 671–682

E.M. Fox, B.J. Howlett, Secondary metabolism: regulation and role in fungal biology, Curr. Opin. Microbiol., 11 (2008), pp. 481–487

J.C. Frisvad, et al., Fumonisin B production by Aspergillus niger, J. Agric. Food Chem., 55 (2007), pp. 9727–9732

A.P. Gasch, et al., Conservation and evolution of cis-regulatory systems in ascomycete fungi, PLoS Biol., 2 (2004), p. e398

M.P. Hare, S.R. Palumbi, High intron sequence conservation across three mammalian orders suggests functional constraints, Mol. Biol. Evol., 20 (2003), pp. 969–978

N.P. Keller, et al., Fungal secondary metabolism – from biochemistry to genomics, Nat. Rev. Microbiol., 3 (2005), pp. 937–947

N. Khaldi, K.H. Wolfe, Evolutionary origins of the fumonisin secondary metabolite gene cluster in Fusarium verticillioides and Aspergillus niger, J. Evol. Biol., 2011 (2011), pp. 1–7

S. Kroken, et al., Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes, Proc. Natl. Acad. Sci. USA, 100 (2003), pp. 15670–15675

S. MacPherson, et al., A fungal family of transcriptional regulators: the zinc cluster proteins, Microbiol. Mol. Biol. Rev., 70 (2006), pp. 583–604

G. Munkvold, Epidemiology of Fusarium diseases and their mycotoxins in maize ears, Eur. J. Plant Pathol., 109 (2003), pp. 705–713

P.E. Nelson, et al., Fumonisins, mycotoxins produced by Fusarium species: biology, chemistry, and significance, Annu. Rev. Phytopathol., 31 (1993), pp. 233–252

J.M. Palmer, N.P. Keller, Secondary metabolism in fungi: does chromosomal location matter?, Curr. Opin. Microbiol., 13 (2010), pp. 431–436

G. Pavesi, et al., Weeder web: discovery of transcription factor binding sites in a set of sequences from co-regulated genes, Nucleic Acids Res., 32 (2004), p. W119

G. Pavesi, et al., MoD tools: regulatory motif discovery in nucleotide sequences from co-regulated or homologous genes, Nucleic Acids Res., 34 (2006), pp. W566–W570

H.J. Pel, et al., Genome sequencing and analysis of the versatile cell factory Aspergillus niger CBS 513.88, Nat. Biotechnol., 25 (2007), pp. 221–231

A. Picot, et al., Factors of the Fusarium verticillioides-maize environment modulating fumonisin production, Crit. Rev. Microbiol., 36 (2010), pp. 221–231

R.H. Proctor, et al., Co-expression of 15 contiguous genes delineates a fumonisin biosynthetic gene cluster in Gibberella moniliformis, Fungal Genet. Biol., 38 (2003), pp. 237–249

R.H. Proctor, et al., A fumonisin biosynthetic gene cluster in Fusarium oxysporum strain O-1890 and the genetic basis for B versus C fumonisin production, Fungal Genet. Biol., 45 (2008), pp. 1016–1026

R.H. Proctor, et al., A polyketide synthase gene required for biosynthesis of fumonisin mycotoxins in Gibberella fujikuroi mating population A, Fungal Genet. Biol., 27 (1999), pp. 100–112

R.H. Proctor, et al., Discontinuous distribution of fumonisin biosynthetic genes in the Gibberella fujikuroi species complex, Mycol. Res., 108 (2004), pp. 815–822

Y. Reyes-Dominguez, et al., Heterochromatic marks are associated with the repression of secondary metabolism clusters in Aspergillus nidulans, Mol. Microbiol., 76 (2010), pp. 1376–1386

J.P. Rheeder, et al., Production of fumonisin analogs by Fusarium species,Appl. Environ. Microbiol., 68 (2002), pp. 2101–2105

U.L. Rosewich, H.C. Kistler, Role of horizontal gene transfer in the evolution of fungi, Annu. Rev. Phytopathol., 38 (2000), pp. 325–363

K.-Y. Seong, et al., Global gene regulation by Fusarium transcription factors Tri6 and Tri10 reveals adaptations for toxin biosynthesis, Mol. Microbiol., 72 (2009), pp. 354–367

F.L. Takken, et al., A one-step method to convert vectors into binary vectors suited for Agrobacterium-mediated transformation, Curr. Genet., 45 (2004), pp. 242–248

M. Thomas-Chollier, et al., RSAT: regulatory sequence analysis tools, Nucleic Acids Res., 1 (2008), pp. W119–W127

M. Tompa, et al., Assessing computational tools for the discovery of transcription factor binding sites, Nat. Biotechnol., 23 (2005), pp. 137–144

J. van Helden, Regulatory sequence analysis tools, Nucleic Acids Res., 31 (2003), pp. 3593–3596

J. Varga, et al., Fumonisin contamination and fumonisin producing black Aspergilli in dried vine fruits of different origin, Int. J. Food Microbiol., 143 (2010), pp. 143–149

I. Visentin, et al., Transcription of genes in the biosynthetic pathway for Fumonisin mycotoxins is epigenetically and differentially regulated in the fungal maize pathogen Fusarium verticillioides, Eukaryot. Cell, 11 (2012), pp. 252–259

I. Visentin, et al., The ITS region as a taxonomic discriminator between Fusarium verticillioides and Fusarium proliferatum, Mycol. Res., 113 (2009), pp. 1137–1145

K.A. Voss, et al., Fumonisins: toxicokinetics, mechanism of action and toxicity, Anim. Feed Sci. Technol., 137 (2007), pp. 299–325

W.M. Wan Norhasima, et al., The health and toxic adverse effects of Fusarium fungal mycotoxin, fumonisins, on human population, Am. J. Infect. Dis., 5 (2009), pp. 273–281

P. Wiemann, et al., Biosynthesis of the red pigment bikaverin in Fusarium fujikuroi: genes, their function and regulation, Mol. Microbiol., 72 (2009), pp. 931–946

C.P. Woloshuk, et al., Molecular characterization of AflR, a regulatory locus for aflatoxin biosynthesis, Appl. Environ. Microbiol., 60 (1994), pp. 2408–2414

W.T.L. Yong, et al., Optimization of Agrobacterium-mediated transformation parameters for Melastomataceae spp. using green fluorescent protein (GFP) as a reporter, Sci. Hortic., 109 (2006), pp. 78–85

J.H. Yu, N.P. Keller, Regulation of secondary metabolism in filamentous fungi, Annu. Rev. Phytopathol., 43 (2005), pp. 437–458