Tri6 Is a Global Transcription Regulator in the Phytopathogen Fusarium graminearum

Charles G. Nasmith1,*, Sean Walkowiak1,*, Li Wang1,*, Winnie W. Y. Leung1, Yunchen Gong2, Anne Johnston1, Linda J. Harris1, David S. Gutman2, Rajagopal Subramaniam1,*/2

1 Agriculture and Agri-Food Canada, Ottawa, Ontario, Canada, 2 CAGEF, University of Toronto, Toronto, Ontario, Canada

Abstract

In F. graminearum, the transcriptional regulator Tri6 is encoded within the trichothecene gene cluster and regulates genes involved in the biosynthesis of the secondary metabolite deoxynivalenol (DON). The Tri6 protein with its Cys2His2 zinc-finger may also conform to the class of global transcription regulators. This class of global transcriptional regulators mediate various environmental cues and generally responds to the demands of cellular metabolism. To address this issue directly, we sought to find gene targets of Tri6 in F. graminearum grown in optimal nutrient conditions. Chromatin immunoprecipitation followed by Illumina sequencing (ChiP-Seq) revealed that in addition to identifying six genes within the trichothecene gene cluster, Tri1, Tri3, Tri6, Tri7, Tri12 and Tri14, the ChiP-Seq also identified 192 additional targets potentially regulated by Tri6. Functional classification revealed that, among the annotated genes, ~40% are associated with cellular metabolism and transport and the rest of the target genes fall into the category of signal transduction and gene expression regulation. ChiP-Seq data also revealed Tri6 has the highest affinity toward its own promoter, suggesting that this gene could be subject to self-regulation. Electro mobility shift assays (EMSA) performed on the promoter of Tri6 with purified Tri6 protein identified a minimum binding motif of GTGA repeats as a consensus sequence. Finally, expression profiling of F. graminearum grown under nitrogen-limiting conditions revealed that 49 out of 198 target genes are differentially regulated by Tri6. The identification of potential new targets together with deciphering novel binding sites for Tri6, casts new light into the role of this transcriptional regulator in the overall growth and development of F. graminearum.

Introduction

Fusarium graminearum Schwabe [teleomorph Gibberella zeae (Schwein.) Petch] is the causal agent of Fusarium head blight (FHB), one of the most destructive crop diseases in temperate climates throughout the world. In addition to yield reduction, FHB is often associated with the accumulation of the secondary metabolite DON in grain [1]. DON belongs to the trichothecene family of secondary metabolites; it binds to the peptidyltransferase of ribosomes thereby inhibiting protein synthesis [2]. DON accumulates in infected plant tissues, is phytotoxic, and poses considerable health risk to consumers [3]. Although considerable effort has been expended to both detect and regulate the amount of this mycotoxin in the infected cereals, there is less information with regard to the regulation of DON biosynthesis.

Our knowledge of mechanisms involved in the biosynthesis of DON and other secondary metabolites comes largely from in vitro culture studies. A considerable amount of evidence gathered over many years suggested that the physiological status of the fungus and the availability of nutrients are the main contributors for secondary metabolite production [4], [5]. Other growth conditions such as pH have also been shown to influence the production of secondary metabolite in numerous fungi including F. graminearum [6], [7]. Recent studies have assessed the role of various carbon and nitrogen sources in the induction of DON [5], [8]. For example, while the products of the polyamine biosynthesis pathway such as agmatine and putrescine strongly influenced DON biosynthesis, the study also revealed negative effects of other nitrogenous compounds on the induction of DON [5]. Other studies have also demonstrated the importance of carbon sources in the regulation of DON biosynthesis [8]. Growth conditions modified by the addition of salt solutions, hydrogen peroxide, and various phytochemicals and fungicides have also been shown to influence DON production [9].

In addition to the physiological conditions, factors affecting fungal developmental also impact the synthesis of secondary metabolites [4]. In Aspergillus species, the canonical heterotrimeric G protein/cyclic AMP/protein kinase A signalling pathway involved in diverse cellular responses including cell division, morphogenesis and pathogenic development affect the production of the secondary metabolites penicillin and sterigmatocystin. For example, mutations in fadA, a gene encoding for the Gα subunit of the heterotrimeric G protein, negatively impacts aflatoxin biosynthesis [10], [11]. Conversely, a dominant active fadA mutant inhibited expression of the transcription factor AflR, resulting in the blockage of sterigmatocystin synthesis. Interestingly, introduc-
Tri6 Is a Global Transcription Regulator

Author Summary

Our knowledge of mechanisms involved in the activation and biosynthesis of DON comes largely from in vitro culture studies. Cumulated knowledge suggests that the physiological status of the fungus and the availability of nutrients are the main determining factors for DON production. Integration of various environmental cues to coordinate expression of secondary metabolic genes is thought to be mediated by a combination of global and pathway-specific transcription factors. While the global transcriptional factors respond to broad range of environmental cues such as the availability of carbon and nitrogen, the pathway-specific transcriptional factors regulate genes within a gene cluster. In Aspergillus, the transcriptional regulator Tri6 is encoded within the trichothecene gene cluster and regulates genes involved in the synthesis and transport of DON. In this report, we utilized ChiP-Seq to demonstrate that Tri6 can potentially bind to promoters and regulate genes not involved in the synthesis of DON and furthermore, many of these non-trichothecene genes are involved in various aspects of cellular metabolism, including transport and energy. Expression profiling revealed that many of the target genes are differentially regulated by Tri6, thus validating our hypothesis that Tri6 is a global regulator involved in cellular metabolism.

Results

Targets of Tri6 include structural and regulatory genes involved in metabolism

Tri6 has been defined as a pathway-specific transcription factor that regulates genes of the trichothecene gene cluster under nitrogen-deprived conditions. However, to characterize Tri6 as a global transcriptional regulator, we sought to identify targets of Tri6 by performing a genome wide ChiP-Seq in F. graminearum grown in nutrient-rich conditions. The ChiP-Seq was performed in the Tri6-HA complemented strain and was compared to the Tri6Δ strain. It should be noted that the addition of the HA epitope to the C-terminus of Tri6 did not compromise its function [21]. Moreover, the presence of Tri6 protein in the Tri6-HA complemented strain was confirmed by immunoblot analysis using HA antibodies (Fig. S1A). ChiP DNA samples from the Tri6-HA and the Tri6Δ strains were sequenced by Illumina Genome Analyzer as 38 base tags. The software Novaalign was used to map the tags/reads to the reference genome (F. graminearum, PH-1; NRRL 31084) and the software Site Identification from Short Sequence Reads (SISRRs) was used to identify potential binding sites [22]. A browser shot of the output from the SISRRs analysis is displayed in Fig. 1. Among the 1491 enriched binding sites in the Tri6-HA complemented strain, we identified a total of 198 protein-coding genes with at least one binding site 1 Kb upstream of the ORF as potential targets of Tri6, distributed in all four chromosomes (Fig. 1). The binding site was defined by a high stringent criterion with a minimum of 120 tags and the number of tags per given target is proportional to the affinity of Tri6 to its target genes [22]. This is highlighted by the substantial enrichment of region in chromosome 2, where Tri6 is located (dotted box, Fig. 1). The 198 target genes with their tags are listed in Table S1.

The analysis of the 198 target genes by the MIPS F. graminearum FunCat database (http://mips.helmholtz-muenchen.de/genre/proj/DB/SS/S/Catalogs/searchCatFirstFun.html) and Kyoto encyclopedia of genes and genomes (http://www.genome.jp/kegg/)
**Tri6 Is a Global Transcription Regulator**

In summary, the genome-wide ChiP-Seq performed in nutrient-rich conditions identified new targets involved in various aspects of metabolism. The targets encompassed not only regulatory genes, but also genes involved in primary and secondary metabolism, energy, and transport. The analysis also identified genes of the trichothecene gene cluster. This was particularly interesting given the fact that these genes are activated only in nutrient-deprived conditions. This suggested that that under nutrient-rich conditions, Tri6 could potentially exert transcriptional control over itself and other Tri genes.

**Tri6 auto-regulates its own expression in nutrient-rich conditions**

The ChiP-Seq data suggested that Tri6 had high affinity to its own promoter so we were interested to know if Tri6 would bind to its own promoter and self-regulate its expression. To demonstrate that Tri6 protein binds to its own promoter, the DNA-Tri6 complex was immunoprecipitated with HA antibodies from both the Tri6Δ and the Tri6-HA complemented strains grown in nutrient-rich conditions. This suggested that under nutrient-rich conditions, Tri6 could potentially exert transcriptional control over itself and other Tri genes.

---

**Figure 1. Browser shot of Tri6 binding sites in all four chromosome of *F. graminearum*.** The binding sites were identified by the site identification software SISSR [22]. The black bars in each window refer to the number of tags or reads for a given site. The corresponding P-value is displayed in red. The dotted box in chromosome 2 indicates the site of the Tri6 promoter region.

doi:10.1371/journal.ppat.1002266.g001
monitor Tri6 transcripts. The first primer set (Tri6-ORF F/R) was designed in the coding region of Tri6 (vertical open box, Fig. 3A) which allowed us to monitor Tri6 transcripts originating only in the wildtype and the tri6Δ strains. As shown in the Fig. 3B and Fig. 3C, a significant increase of Tri6 expression (Tri6-5 UTR, 4.4 ± 0.4, Fig. 3B and Tri6-5 UTR, 7 ± 0.9, Fig. 3C) was observed in the tri6Δ strain, compared to the wildtype strain. However, over expression of Tri6 in the tri6Δ strain (tri6ΔTri6) resulted in decreased Tri6 expression (Tri6-5 UTR, 0.72 ± 0.1, Fig. 3B and Tri6-5 UTR, ± 0.06, Fig. 3C). This suggested that Tri6 acts as a repressor, regulating its own expression in nutrient-rich conditions. We did not observe any significant change in the expression of other Tri genes in the tri6ΔTri6 strain (tri6ΔTri6, Fig. 3B and 3C), suggesting that additional factors are required for the expression of these genes.

**Identification and characterization of the Tri6 binding site in the Tri6 promoter**

A recent study performed in *F. graminearum* suggested TNAGGCC as a DNA binding site for Tri6 protein [17]. A recent study that examined the promoters of genes differentially regulated by Tri6 during the *F. graminearum* infection process also suggested a similar DNA binding motif [16]. The results described here confirmed that Tri6 is able to bind its own promoter in vivo (Fig. 2) and regulate its own expression (Fig. 3). Since the promoter of Tri6 (~836 to ~768) harbours two RNaGGCC (where R = G or A) binding sites (Fig. 4A), we employed EMSA analyses to further delineate the binding site for Tri6. All the EMSA assays were performed with purified recombinant Tri6 protein (Fig. S1B). First, we tested the probe which contained two of the RNaGGCC motifs and as the results indicated, Tri6 did not bind to this probe (Tri6-1, Fig. 4B). This prompted us to examine the region that surrounds this motif. The probe Tri6-2 which included sequences proximal to the binding sites (Fig. 4A) also did not bind the Tri6p (Fig. 4B). However, a probe (Tri6-3) which included sequences distal to the binding sites was able to bind Tri6 protein. The specificity of the binding to the Tri6-3 probe was confirmed by the addition of 10-fold excess non-labelled Tri6-3 probe in the EMSA assays (Tri6-3, competitor, *+*, Fig. 4B). In addition, we also used a probe which did not contain this consensus sequence as a negative control (Tri6-NS, Fig. 4B). These results suggested that RNaGGCC motif is not involved in Tri6 binding *in vitro*. To further confirm this, nucleotides GGGC within the motif were mutated in the Tri6-3 probe and the results indicated that the mutations did not abolish Tri6 binding (Tri6-3-1, Fig. 2). These results confirmed that the RNaGGCC motif is not required for binding of Tri6, but suggested that nucleotides outside of this motif

![Figure 2. Tri6 binds to its cognate promoter in vivo.](image)

**Table 1. Tri6 targets summarized by functional categories.**

| Functional Category | Number of genes |
|---------------------|-----------------|
| 01 Metabolism       | 25              |
| 02 Energy           | 3               |
| 10 Cell cycle and DNA processing | 9               |
| 11 Transcription    | 3               |
| 12 Translation      | 12              |
| 14 Protein fate     | 23              |
| 16 Protein with binding function | 3               |
| 18 Regulation of metabolism and protein function | 24               |
| 20 Cellular transport, transport facilities and routes | 4               |
| 20.01.10 Protein transport | 5               |
| 20.01.27 Drug/toxin transport | 8               |
| 20.03 Transport routes | 13              |
| 30 Signal transduction | 6               |
| 30.01.05.05 G-protein mediated | 3               |
| 30.01.05.05.01 GTPase mediated | 3                |
| 32 Cell rescue, defense and virulence | 15             |
| 32.05 Disease, virulence and defense | 7               |
| 32.05.03 Defense related proteins | 4               |
| 32.07 Detoxification | 6               |
| 34 Interaction with the environment | 6               |
| 40 Cell fate        | 4               |
| 42 Biogenesis of cellular components | 10              |
| 43 Cell type differentiation | 6               |
| 99 Unclassified proteins | 114             |

doi:10.1371/journal.ppat.1002266.t001

**Table 2. Enrichment of Tri6 binding sites in the promoters of trichothecene genes in nutrient-rich conditions.**

| Tags | Trichothecene genes |
|------|---------------------|
| 88725 | FGSG_03336, Tri6     |
| 6882  | FGSG_00071, Tri1     |
| 5426  | FGSG_03543, Tri14    |
| 362   | FGSG_03343, Tri3     |
| 260   | FGSG_03541, Tri12    |
| 206   | FGSG_03333, Tri7     |

ChIP-Seq was performed with Tri6-HA complemented and the tri6Δ strains. The Tags represent binding of Tri6 protein to the promoters of Tri genes in the Tri6-HA strain in nutrient-rich conditions.

doi:10.1371/journal.ppat.1002266.t002
in the Tri6-3 probe are involved in Tri6 binding. Outside of the RNAGGCC motif, three domains in the Tri6-3 probe were recognized that could potentially bind Tri6 (Fig. 5A). We designated CTGA sequence which partially overlaps the AGGCC site as Domain I and the two GTGA repeats separated by six nucleotides as Domain II and III, respectively (Fig. 5A). As the EMSA results indicated, individual nucleotide mutations within Domain I did not have any noticeable effect on Tri6 binding (Fig. 5B). However, individual mutations in Domain II and Domain III of the GTGA sequences, respectively, dramatically decreased Tri6 binding (Fig. 5C). Furthermore, combined mutations in both Domain II and III completely abolished the Tri6 binding (Fig. 5D). These results suggested that either a single GTGA sequence or GTGA repeats in the promoter of Tri6 are required for Tri6 binding.

**Tri6 targets are differentially regulated under nitrogen limiting conditions**

The 198 Tri6 targets identified in nutrient-rich conditions included several genes of the trichothecene gene cluster (Table 2). Since the Tri genes are expressed only under nitrogen-limiting conditions, we were interested to know how many of the non-Tri genes targets are co-regulated with the Tri genes. To address this, we performed a genome-wide expression analysis of wildtype and tri6Δstrain grown in nutrient-limiting conditions. RT-qPCR reactions were performed in triplicates using Applied Biosystems Power SYBR Green kit and the Applied Biosystems Step One Plus Real-Time PCR System. A list of qPCR primers for all the Tri genes is listed in the Table S2. The β-tubulin gene (FGSG_09530) was used as the internal control and the data was imported and Relative quantity (RQ) was derived by the Relative standard method included in the StepOne 2.1 software. (C) Identical to B, except the internal control used was Gapdh (FGSG_06257). The figures are representative of two independent biological replicates.

doi:10.1371/journal.ppat.1002266.g003

**Discussion**

Much of our understanding regarding the transcription regulator Tri6 stems from earlier studies that identified its control of the mycotoxin pathway leading to trichothecene biosynthesis and designated it as a pathway-specific transcriptional regulator [17], [18]. Since then, evidence from numerous laboratories have challenged this perception and thus prompted us to hypothesize that Tri6 is a global transcriptional regulator whose influence extends beyond the trichothecene gene cluster [16], [20]. Accordingly, the ChiP-Seq experiments performed in nutrient-rich conditions demonstrated that Tri6 could bind to promoters of both structural and regulatory genes associated with many aspects of nitrogen, carbon and lipid metabolisms (Table S2). One of the intriguing revelations was that even under these conditions Tri6...
was bound to the promoters of genes involved in the biosynthesis of the secondary metabolite, trichothecene 15-ADON (Table 2). Since this cluster is only activated in Fusarium grown under nitrogen-deficient conditions, binding of Tri6 to the promoters of Tri genes suggested that Tri6 could potentially act as a negative regulator, suppressing these genes under nutrient-rich conditions. The results from the RT-qPCR clearly demonstrated that with the addition of Tri6 protein (Tri6p), the probe and the unlabelled probe (competitor) is indicated by ‘+’ sign and the omission is indicated by the ‘−’ sign. The capital bold letters indicate the consensus binding sequence and the bold non-capital letters represent mutations within the sequence YNAGGCC. Arrow indicates the migration of the free probe. This is representative of at least three independent experiments. doi:10.1371/journal.ppat.1002266.g004

Tri6 is a Global Transcription Regulator

**Tri6 and other global transcriptional regulators**

Nitrogen catabolite repression, or NCR, refers to genes that are repressed when a preferred source of nitrogen such as ammonia or glutamine are present in the environment [24]. NCR is mediated by the action of global nitrogen regulators which regulate the expression of structural genes required to metabolize alternate nitrogen sources. AreA is a prototypic NCR gene regulator and has been identified in many fungi, including *F. graminearum* and *F. fujikuroi* [30–33]. Under nutrient-rich conditions, expression of AreA is repressed by the binding of the repressor protein NmrA to the promoter of *AreA* [31], [32]. However, under nitrogen-starving conditions, *AreA* repression is relieved and target genes involved in the utilization of non-preferred sources of nitrogen are expressed [31], [32]. Coincidentally, some of the known *AreA* target genes such as the branched amino acid transferase (*FGSG_05696*; −2.2-fold, Table S4) and the general amino acid permease (*FGSG_05574*; −4-fold, Table S4) that are normally expressed in the wildtype *Fusarium* strain under nitrogen-deprived conditions are repressed in the *tri6* strain [34]. This leads one to speculate that *Tri6*, similar to AreA, is involved in NCR. Studies in *F. fujikuroi* have also shown that *AreA*, in addition to regulating those genes involved in NCR, also regulates expression of genes that synthesize secondary metabolites like gibberellins and the pigment bikaverin [35], [36]. An additional feature that is shared between *Tri6* and *AreA* is that they are both subject to auto-regulation [31]. Taken together, our evidence suggests one possible scenario where *Tri6* regulates functions of *AreA* through *NmrA* in *F. graminearum*. In support, we identified two genes (*FGSG_05942* and *FGSG_10774*, Table S2) with Rossman-fold domains as targets of *Tri6*. Rossman-folds are distinguishing features of NmrA proteins [26]. A further link between *NmrA* and virulence was recently established in *F. oxysporum* where the deletion of the transcription factor *MeaB*, which regulates *Nmr1*, an orthologue of *NmrA* resulted in the repression of the virulence gene *Six1* [37].

**DNA binding properties of Tri6**

The EMSA studies demonstrated that mutations in either one of the GTGA elements in the promoter of *Tri6* reduced the binding of Tri6. However, combined mutations in both the GTGA elements completely abolished this binding (Fig. 5D). This suggested that the presence of two GTGA elements in close proximity to each other in the *Tri6* promoter likely contributes to the affinity of Tri6 to this site. Studies with other global regulators, specifically those responding to nitrogen, such as *AreA* in *A. nidulans* or its counterpart Nit2 in *N. crassa* showed that binding sites located within 30bp of each other in either orientation

![Figure 4. The Tri6 protein does not recognize the YNAGGCC motif by EMSA analyses. A) Sequence of probes Tri6-1, Tri6-2 and Tri6-3 in the Tri6 promoter (-836 to -767) used in the EMSA analyses. The bold letters indicate the putative consensus binding sequence YNAGGCC for Tri6p. B) EMSA analyses were performed using the probes Tri6-1, Tri6-2, Tri6-3, a non-specific probe (NS) and a probe with mutations within the binding motif (Tri6-3-1). The addition of Tri6 protein (Tri6p), the probe and the unlabelled probe (competitor) is indicated by ‘+’ sign and the omission is indicated by the ‘−’ sign. The capital bold letters indicate the consensus binding sequence and the bold non-capital letters represent mutations within the sequence YNAGGCC. Arrow indicates the migration of the free probe. This is representative of at least three independent experiments. doi:10.1371/journal.ppat.1002266.g004](image-url)
enhances their DNA binding affinity [38]. When these criteria were applied to the analysis of Tri6 target genes, 109 of the 198 target genes harboured multiple GTGA/TCAC binding sites in their promoters, separated by eight nucleotides or less, suggesting a common mechanism of binding and regulation by Tri6 (Fig. 7). It should be noted that GTGA/TCAC motifs separated by more than eight nucleotides are represented in the promoters of all the Tri6 target genes (Table S1). The GTGA/TCAC motif described in this study is distinct from the previously reported YNAGGCC for Tri6 binding [17]. Unlike this study, Hohn et al. did not use purified Tri6 protein and furthermore, their EMSA analysis with regions of the Tri3 promoter which contained the putative YNAGGCC consensus binding site did not result in a shift [17]. This led them to speculate an alternate binding site for Tri6. It is noteworthy to indicate that Tri3 is one of the genes identified as targets of Tri6 with the GTGA repeats in its promoter.

Studies with global regulators have outlined several mechanisms that influence both binding affinity and specificity [31]. For example, proteins that bind to direct repeats most likely bind as dimers. As an example, Da130 protein in S. cerevisiae, involved in nitrogen catabolism, binds to two closely spaced GATA elements as a dimer in either a tail-to-tail or head-to-head orientation [39]. Interactions between global and pathway-specific regulators also influence both DNA binding specificity and affinity. For example, some of the genes involved in nitrate assimilation are regulated by the interaction between the global regulators AreA and NirA in A. nidulans, while other genes are regulated by the global regulator Nit2 and the pathway-specific regulator Nit4 in N. crassa [24]. Since we observed closely spaced GTGA repeats in the Tri6 promoter of F. graminearum, we could envisage a scenario, where Tri6 could function as a dimer and repress the expression of Tri6 in nutrient-rich conditions. However, under nutrient-deprived conditions, where the expression of Tri6 increases, Tri6 could interact with other known regulators such as Tri10 or PacC to provide specificity and regulate expression in condition-specific manner [14], [19]. This interaction may also provide a platform for the Tri6-complexes to bind to an alternate site, such as the YNAGGCC proposed in several studies. Such a scenario could account for the differences between the genes regulated by Tri6 in this study from those seen in a recent study of F. graminearum infection in planta [16]. It is noteworthy to mention that the authors in that study also observed stark differences between the genes regulated by Tri6 in culture and in planta, albeit the comparison was made between two different Fusarium strains. One of the intriguing findings from the in planta study was that the promoters of genes, specifically the genes of the isoprenoid pathway essential for trichothecene production in both in planta and culture, showed enrichment of the YNAGGCC motif only in F. graminearum and

![Figure 5. Tri6 binds to GTGA repeats in the Tri6 promoter by EMSA analyses.](https://www.plospathogens.org/article/10.1371/journal.ppat.1002266.g005)
Figure 6. *Tri6* differentially regulates 26 targets under nitrogen-deprived conditions. Expression profiling was performed with wildtype and *Tri6* mutant (*tri6*Δstrain grown in nitrogen-limiting conditions). Eighteen targets that are down-regulated and eight targets that are up-regulated in the *tri6*Δ strain are represented. The top five genes in each category were verified by RT-qPCR analysis represented by relative quantity (RQ). Two internal standards (*β*-tubulin and Gapdh) were used. This is a representative of three independent experiments.

doi:10.1371/journal.ppat.1002266.g006

Figure 7. GTGA/TCAC repeats are present in the promoters of *Tri6* target genes. One kilobase region upstream of the open reading frame of each of the *Tri6* target genes were analyzed for GTGA/TCAC repeats separated by 0 to 8 nucleotides (N). This is a representative of three independent experiments.

doi:10.1371/journal.ppat.1002266.g007
not in other Fusarium species unable to make trichothecenes or in other related fungi [16]. It was proposed that co-regulation of the isopenoid genes and the trichothecene genes represent an evolutionary adaptation specific to F. graminearum. A logical extension of this argument is that the source of nutrition, determined by the environment such as culture conditions or the type of host that this fungus infects, is the driving force behind this adaptation.

In conclusion, this study has identified new or additional role for the transcriptional regulator, Tri6. The findings from this study together with earlier studies have led us to conceive a new mode of action for Tri6. In this model, Tri6 recognizes and binds the GTGA/TCAC elements under one condition such as the nutrient-rich condition where it regulates itself and other genes involved in various aspects of metabolism. However, when the environmental conditions change, be it in culture or in planta, activation of other regulators including Tri10 and PacC may change the dynamics of Tri6-complex and initiate new interactions with the DNA, which could involve recognition of an alternate site such as the YNAGGCC motif. Although, we have no evidence to show that Tri6 directly interacts with either Tri10 or PacC, isolation and identification of Tri6 and Tri10 protein complexes in vivo will shed more light on both the dynamics and the links between these regulators.

Materials and Methods

Strains and growing conditions

Fusarium graminearum wildtype strain GZ3639 (NRRL 38155) was provided by C. Babcock of the Canadian Collection of Fungal Cultures (CCFC), Agriculture and Agri-Food Canada, Ottawa. F. graminearum and transformants were grown on PDA (Sigma C for 30 s for annealing, 68°C for 1 min extension of this argument is that the source of nutrition, determined by the environment such as culture conditions or the type of host that this fungus infects, is the driving force behind this adaptation.

In conclusion, this study has identified new or additional role for the transcriptional regulator, Tri6. The findings from this study together with earlier studies have led us to conceive a new mode of action for Tri6. In this model, Tri6 recognizes and binds the GTGA/TCAC elements under one condition such as the nutrient-rich condition where it regulates itself and other genes involved in various aspects of metabolism. However, when the environmental conditions change, be it in culture or in planta, activation of other regulators including Tri10 and PacC may change the dynamics of Tri6-complex and initiate new interactions with the DNA, which could involve recognition of an alternate site such as the YNAGGCC motif. Although, we have no evidence to show that Tri6 directly interacts with either Tri10 or PacC, isolation and identification of Tri6 and Tri10 protein complexes in vivo will shed more light on both the dynamics and the links between these regulators.

Materials and Methods

Strains and growing conditions

Fusarium graminearum wildtype strain GZ3639 (NRRL 38155) was provided by C. Babcock of the Canadian Collection of Fungal Cultures (CCFC), Agriculture and Agri-Food Canada, Ottawa. F. graminearum and transformants were grown on PDA (Sigma Chemical Co. USA) plates.

Construction of Tri6Δdeletion, Tri6-HA complement and tri6Δ/Tri6 over expression strains

Constructs for the tri6Δand Tri6-HA complement strains have been described (Fig. S1; http://apsjournals.apsnet.org/doi/suppl/10.1094/MPMI-09-10-0210). To construct the Tri6 over expression vector, we first PCR amplified Tri6 ORF with the primer set Tri6-GUE-F/ Tri6-GUE-R (Table S2) and F. graminearum genomic DNA as template. The PCR fragment was cloned into the vector pSW-GU. The backbone of pSW-GU vector is identical to pRF-HUE [40] except that the selection marker Hygromycin gene was replaced with Geneticin, encoded by the gene neomycin phosphotransferase [40]. Transformation of the tri6Δstrain with the Tri6 over expression vector was performed according to Rasmus et al [40]. The transgenic Fusarium strains were verified by PCR (Fig. S2). PCR conditions: 200 µM Primers, 200 µM NTP’s, 1 unit of Expand long polymerase (Roche), 95°C for 30 s, 52°C for 30 s for annealing, 68°C for extension for 37 cycles. All PCR products were purified using Qiagen’s PCR purification kit.

Trichothecene induction and detection

To induce trichothecene production in liquid culture, a two stage media protocol, modified from Miller and Blackwell was employed [41]. 20,000 spores of wildtype, tri6Δand tri6Δ/Tri6 over expression strains were inoculated into 4 mL of first stage growth media and incubated in Falcon Multitwell 6-well culture trays (Fig. S2). The culture trays were affixed to an orbital shaker and the spores were grown for 24 hr at 28°C in the dark with constant shaking at 170 rpm. Following 24 hr growth, the mycelial solids were washed with water and resuspended in 4 mL of second stage media (pH 4.0) [41] and then transferred to the 6-well culture trays. The mycelia was grown in second stage media under the same conditions as described previously. The supernatant was collected after 24 hr for trichothecene analyses. Trichothecenes were analysed on an AKTA 10 purifier (GE Healthcare, Canada) with direct injection of 100 µL of the culture filtrate into a 150×4.6 mm, 5 µm Hypersil ODS column (Thermo-Electron Corp.), using a methanol: water gradient from 15:85 to 60:40 over 25 min at a flow rate of 1 mL/min. Under these conditions, 15-ADON elutes at a retention time ~10 min, monitored by UV 220nm.

Chromatin Immunoprecipitation

Spores from the tri6Δand the complemented (Tri6-HA) strains were used to inoculate first stage media and grown for 19 hr as described before. The mycelia were washed with water and filtered (sterile 1MM; Whatman). The ChiP-enriched DNA was prepared according to Saleh et al. [42] with modifications. The mycelial pellet was incubated in the cross-linking buffer for 30 min with continuous shaking. After thorough washing, the pellet was ground in liquid N2 and resuspended in the lysis buffer (250 mM, HEPES pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton, 0.1% DeoxyCholate, 10 mM DTT and Protease inhibitor (Roche complete Mini) and incubated for 1 hr on ice. To obtain uniform fragments of cross-linked DNA, the suspension was sonicated 6x for 15 sec (60% amplitude, Misonix 4500 Sonicator). A small aliquot was used to confirm a dominant diffused band between 500 and 1000 bp. Approximately 750 µg of protein was incubated with 30 µL HA-magnetic beads (Miltenyi Biotech) and left overnight at 4°C. Immunoprecipitates were washed six times with lysis buffer and eluted with 250 µL of freshly prepared elution buffer (0.5% SDS and 0.1 M NaHCO3) with incubation at room temperature (RT) for 15 min with gentle agitation. The elution step was repeated once with 30 min of incubation at RT. To reverse the cross-linking, 20 µL of 5 M NaCl was added to the 500 µL elutes and incubated overnight at 65°C. The samples were further incubated for 1.5 hr at 45°C with 1 µL of Proteinase K (20 mg/mL). 30 mM Tris·HCl pH 6.5 and 10 mM EDTA to digest the proteins. DNA from these samples was precipitated according to Saleh et al. [42].

Preparation of DGE and ChiP libraries and Illumina sequencing

ChiP-enriched DNA was prepared for sequencing using the Illumina ChiP sample prep protocol with minor modifications. In brief, 2 µg of ChiP-enriched DNA was end repaired, and an adenosine overhang added to the 3’ ends. Standard Illumina adapters were ligated to the DNA and 240 bp fragments size selected from a 1.5% agarose gel. The eluted products were enriched with 18 cycles of amplification with Illumina PCR primers. Libraries were validated on the Bioanalyzer 2100 (Agilent) and Qubit fluorometer (Invitrogen, Carlsbad, CA). Each sample was run on two lanes of the Illumina GAII sequencer at the Centre for the Analysis of Genome Evolution and Function (CAGEF, University of Toronto). For each sample, 6 pmol was loaded in one lane of a standard flow cell, and 8 pmol was loaded in a second lane. The GAII was run for 38 cycles.

Statistical analysis of ChiP data

In total, 24,201,991 out of 29,933,903 tags from Tri6-HA complemented strain and 15,756,703 out of 22,010,411 tags from tri6Δstrain were normalized and uniquely mapped to the genome by Novoalign [http://www.novocraft.com] [43], [44]. The rest of the tags were either of low quality or mapped to multiple locations.
Of the mapped tags, ~50% were mapped to sense strand and another 50% were mapped to anti-sense strand. The mapping data were analyzed by the Site Identification from Short Sequence Reads (SISSRs) software for identification of the binding sites with tags originating from the tri6Δ strain as control [22]. The following parameters were used to run the program: 36,986,000 base pairs for the genome size, 0.8 for the fraction of genome mappable by reads, 1 for the E-value (minimum number of directional tags required on each side of the inferred binding site), 0.1 for the p-value (for fold enrichment of ChIP tags at a binding site location compared to that at the same location in the control data), 2 for the scanning window size, 240 base pairs for the average DNA fragment length, and 350 base pairs for the upper bound DNA fragment length. The switch -u was turned on to allow for identification of the binding sites supported only by reads mapped to one strand. Under these conditions, we were able to identify 1,491 potential binding sites with at least 120 tags (Tag#, the sum of the tags mapped to positive strand of the left half of the binding site and the tags mapped to negative strand of the right half of the binding site) and at least 2.15 fold more tags in treatment vs control. We then compiled a list of 198 protein-coding genes with at least one binding site in 1 kb upstream of the start site. The sum of the Tag# indicates the relative affinity of Tri6 to the binding sites in the promoter of the gene.

RNA isolation and Quantitative RT-PCR Analysis

The two-stage media was employed to grow F. graminearum as described before. Mycelia grown for 5 hr in the second stage (nitrogen-limiting) media were filtered and ground to a powder in liquid nitrogen. The RNA was isolated from 0.25 g of ground mycelia with 1 mL of Trizol (Invitrogen, USA) according to manufacturer’s instructions. The RNA was purified further with the Invitrap Spin Cell RNA mini kit (Invitek, Germany) according to the manufacturer’s instructions. cDNA for the quantitative RT-PCR Analysis (RT-qPCR) experiments was synthesized from 1 μg total RNA using random hexamers using the high-capacity cDNA reverse transcription kit (Applied Biosystems, USA). RT-qPCR was performed using the Applied Biosystems StepOne Plus Real Time PCR system (ABI, Foster City, USA). Standard curves were created for the housekeeping genes (β-actin, α-tubulin; Fgsg_05530 and Gap1p; Fgsg_06257) and the genes of interest according to Relative standard curve method outlined in StepOne 2.1 software. Primers used in the RT-qPCR are listed in supplementary Table S2. Using the relative standard curve, the relative quantity (RQ) was determined by comparing target quantity in each sample to the reference sample. P values <0.05 were considered to be statistically significant.

Purification of Tri6 protein and Electromobility shift assays

Tri6 was cloned into pDEST17 vector (Invitrogen, USA) and expressed in E. coli BL21 and the Tri6 protein was purified by the His Trap FF affinity column (GE Health care, Sweden). Double stranded DNA probes (100 ng/mL) were labelled using T4 Polynucleotide Kinase (Fermentas MBI) and 20 μCi of γ-32P ATP (Perkin Elmer, USA). The labelled probes were purified using the Quick spin column (Roche, USA) and specific activity varied from 4–7 x 10^6 cpm/μg. Binding reactions for the EMSA were performed in 15 μL volume with 2 μg purified Tri6-HA recombinant protein, labelled probes (20,000 cpm; 2ng) for 30 minutes at room temperature in a buffer containing 20 mM HEPES pH 7.9, 2 mM DTT, 5% (V/V) Glycerol and 1 μg poly[dIdT]. The reaction was run on 6% polyacrylamide gels (acrylamide/bis ratio of 19:1 w/w) in 50 mM Tris-HCl, pH 8.0, 50 mM sodium borate and 1 mM EDTA. The competition experiments were performed by adding unlabelled oligonucleotides to the reaction at 10X molar excess (20 ng) of the labelled probe. The shift obtained by the Tri6 was scanned by the Storm 840 densitometer (GE healthcare, USA) and quantified by Image Quant TL 7.0 software (GE healthcare, USA). All the DNA probes used in EMSA were synthesized by Sigma Genosys (SigmaAldrich, Canada).

DNA microarray and statistical analysis

Wildtype and the tri6Δ strains were grown in the second stage media for 5 hrs and the RNA was extracted as described before. The integrity of RNA was confirmed by an Agilent 2100 Bioanalyzer (Agilent, Canada). RNA was then converted into cDNA using the Agilent Quick Amp Labeling Kit and converted back into labeled RNA using T7 RNA polymerase and cyanine 3-labeled CTP or 5-labeled CTP from the Agilent two-color RNA Spike-in kit. The cRNA was then hybridized to a custom F. graminearum 4X44K oligomer microarray (Agilent Technologies, CA, USA), using an Agilent gene expression hybridization kit. Both dye combinations were performed for each biological replicate. Each array consisted of 1,417 spike-in and negative controls and up to three 60-mer oligos designed for each of 13,918 predicted F. graminearum genes (NCBI GEO, Platform Accession #GPL11046). The hybridizations were scanned using the GenePix Professional 4200A scanner, and the signals quantified using GenePix Pro 6. The microarray data was transferred into Acuity 4.0 and the data was normalized using Lowess Normalization. Data points with low intensities were removed, and the dye swapped replicates were combined and expressed in log2 ratio. ANOVA was used to determine the consistency between samples for each hybridization (P<0.05). The data points were then averaged first between three biological replicates, then within the three average hybridizations. The data was then back transformed from the log scale as fold-expression. Separate datasets were then generated that contained genes that were positively identified in the chromatin immunoprecipitation data, as well as the genes that are up and down regulated in the entire genome by either 1.5-fold or greater and 2.0-fold or greater. Raw data can be accessed at NCBI GEO, Accession # GSE30892.

Supporting Information

Figure S1 A) Detection of Tri6 protein in the Tri6 complemented strains (Tri6-HA) by immunoblot analysis. 100 μg of total protein from both Tri6 mutant (tri6Δmutant Tri6 complemented strains (Tri6-HA) was separated by SDS-PAGE and detected by HA-antibodies. The arrow indicates the migration of the Tri6 protein and the * indicates non-specific cross reacting with HA antibodies and serves as internal loading control. B) Purification of Tri6 protein expressed in bacteria. Tri6 gene was His tagged at the C-terminus and expressed in E.coli BL21Lys E. coli and purified over a Nickel affinity column. Tri6 protein eluted from the Nickel affinity column was detected by Coomassie Blue G-250. (TIF)

Figure S2 A) Characterization of Tri6 over expression strains by PCR. Genomic DNA was isolated from wildtype (Wt), Tri6 mutant (tri6Δstrains and the Tri6 over expressor transgenic strain (tri6ΔTri6). PCR was performed as outlined in the methods section with the primer set Tri6-ORF-F/Tri6-ORF-R to detect Tri6 (lane 1), the primer set HygF/HygR to detect the selection marker Hygromycin (lane 2), the primer set GenF/GenR to detect the selection marker Geneticin (lane 3) and the primer set
Tri6 GUE F/Tri6 GUE R to detect the entire Tri6 over expression construct (lane 4). B) HPLC analysis of the production of 15-ADON from all three strains grown in DON-inducing media. The strains grown in six well culture plates and induced for 15-ADON as described in Methods. The retention time for the elution 15-ADON was ~10 mins and the elution from each strain is indicated. There is no production of 15-ADON from the tri6Δstrain (Blue line). The quantity of 15-ADON is indicated by arbitrary absorbance units measured at 220 nm (AU220).

Table S1 Complete list of Tri6 targets identified in ChIP-Seq experiment. The Tri6 gene targets are presented with gene description and the number of associated tags in the ChIP experiment. Number of tags is proportional to the binding affinity of Tri6 to their cognate promoter DNA sequences. * refers to the number of GTGA and TCAC sequence present in the one kb promoter of each gene.

Table S2 Tri6 targets organized by functional category as determined by MIPS.

Table S3 List of primers used in this study.

Table S4 Complete list of Fusarium genes differentially expressed under nitrogen-deprived conditions. Genes expressed in the tri6Δstrain relative to expression in the wildtype Fusarium strain and values represent expression over 2-fold or higher.

Table S5 Differential expression of Tri6 targets with cut-off of 1.5-fold or higher.

Acknowledgments

The authors are grateful to George Fedak and Andre Kalikililo from Agriculture & Agri-Food Canada for wheat seed sources and members of CAGEF at the University of Toronto. The authors would like to thank Drs. D. Desveaux (U of Toronto) for critical reading of the manuscript.

Author Contributions

Conceived and designed the experiments: CGN LW AJY WL RS. Performed the experiments: CGN LW AJ WWYL RS. Analyzed the data: CGN SW AJ YG DSG. Contributed reagents/materials/analysis tools: YG DSG LJH RS. Wrote the paper: RS.

References

1. Kang Z, Buchnerauer H (2002) Studies on the infection process of Fusarium culmorum in wheat spikes: Degradation of host cell wall components and localization of trichothecene toxins in infected tissue. Eur J Plant Pathol 108: 633–660.
2. Bennet JW, Kich M (2003) Mycotoxins. Clinical Micro Rev 16: 497–516.
3. Peska JJ, Szenolinski AT (2005) Deoxynivalenol: toxicology and potential effects on humans. J Toxicol Environ Health B Critical Rev 9: 39–69.
4. Calvo AM, Wilson RA, Bok JW, Keller NP (2002) Relationship between secondary metabolism and fungal development. Microbiol and Mol Bio Reviews 66: 447–459.
5. Gardiner DM, Kazan K, Manners JM (2009) Nutrient profiling reveals potent inducers of trichothecene biosynthesis in fusarium graminearum. Fungal Genet Biol 46: 604–613.
6. Yu J-H, Keller N (2005) Regulation of secondary metabolism in filamentous fungi. Annu Rev Phytopathol 43: 437–58.
7. Gardiner DM, Osborne S, Kazan K, Manners JM (2009) Low pH regulates the production of deoxynivalenol by Fusarium graminearum. Microbiology 155: 3149–3156.
8. Jiao F, Kawakami A, Nakajima T (2008) Effects of different carbon sources on trichothecene production and Tri10 gene expression by Fusarium graminearum in liquid culture. FEMS Microbiol Lett 285: 212–219.
9. Ochiai N, Tokai T, Takahashi-Ando N, Fujimura M, Kimura M (2007) Regulation of secondary metabolism and fungal development. Microbiol and Mol Bio Reviews 71: 224–235.
10. Hicks J, Yu H-J, Keller N, Adams TH (1997) ApeGelder sporation and mycotoxin production both require inactivation of FdA G p protein-dependent signaling pathway. EMBO J 16: 4916–4923.
11. Rhee LV, Beaudry KM, Keller NP, Linz JE (2004) Regulation of aflatoxin synthesis by FdA/GMP:protein kinase A signaling in ApeGelder parasiticus. Mycopathologia 158: 219–232.
12. Tag A, Hicks J, Garifullina G, Ake jr C, Phillips TD, et al. (2000) G-protein signaling mediates differential production of toxic secondary metabolites. Mol Microbiol 38: 659–665.
13. Petalva MA, Arest HN (2004) Recent advances in the characterization of ambient pH regulation of gene expression in filamentous fungi and yeasts. Annu Rev Microbiol 58: 423–451.
14. Jerelj J, Richard-Ferget F, Barreau C (2011) The pH regulatory factor Pacl regulates Tri10 gene expression and trichothecene production in Fusarium graminearum. Fungal Genet Biol 48: 275–284.
15. Fernandes M, Keller N, Adams TH (1998) Sequence-specific binding by ApeGelder nubulin ADD, a C6 zinc cluster protein regulating mycotoxin biosynthesis. Mol Microbiol 30: 1355–1366.
16. Seong Kye-Y, Pasquali M, Hilburn KL, McCormick SP, Xia J, et al. (2009) Global gene regulation by Fusarium transcription factors Tri6 and Tri10 reveals adaptations for toxin biosynthesis. Mol Microbiol 72: 354–367.
17. Hohn TM, Krishna R, Proctor RH (1999) Characterization of a transcriptional activator controlling trichothecene toxin biosynthesis. Fungal Genet Biol 26: 224–235.
18. Proctor RH, Hohn TM, McCormick SP, Dejardins AE (1995) Tri6 encodes an unusual zinc finger protein involved in regulation of trichothecene biosynthesis in Fusarium sporotrichioides. Appl Environ Microbiol 61: 1923–1930.
19. Perplow AW, Tag AG, Garifullina GF, Beremand MN (2005) Identification of new genes positively regulated by Tri10 and a regulatory network for trichothecene mycotoxin production. Appl Environ Microbiol 69: 2731–2736.
20. Gardiner DM, Kazan K, Manners JM (2009) Novel genes of Fusarium graminearum that negatively regulate deoxynivalenol production and virulence. Mol Plant-Microbe Interact 12: 1569–1600.
21. Schreiber K, Naemicht C, Allard G, Singh J, Subramaniam R, et al. (2011) Found in translation: High-throughput chemical screening in Arabidopsis thaliana identifies small molecules that reduce Fusarium head blight disease in wheat. Mol Plant-Microbe Interact 24: 640–648.
22. Jothi R, Cudدادap S, Barski A, Ciu K, Zhao K (2008) Genome-wide identification of in vivo protein-DNA binding sites from ChIP-seq data. Nucleic Acids Res 36: 5221–5231.
23. Slaughter JC (1980) Nitrogen metabolism. In: Berry DR, ed. Physiology of Industrial Fungi. Oxford: Blackwell Scientific Publications. pp 58–76.
24. Marzluf GA (1997) Genetic regulation of nitrogen metabolism in the fungi. Microbiol Mol Biol Rev 16: 17–32.
25. Bopoulous A, Mrozowicz Z, Thivenian F, Le Dall M-T, Hapala I, et al. (2006) Mastering lipid accumulation in the yeast Torulaspora delbrueckii. Appl Environ Microbiol 74: 7779–7789.
26. Stammers DK, Ren J, Leslie K, Nichols CE, Lamb HK, et al. (2001) The structure of the negative transcriptional regulator NimR reveals a structural superfamily which includes the short-chain dehydrogenase/reductases. The EMBO Journal 20: 6619–6626.
27. Andrianopoulos A, Kouranbas S, Sharp JA, Davis MA, Hynes MJ (1998) Characterization of the ApeGelder nubulin add gene involved in nitrogen metabolism expression. J Bacteriol 180: 1973–1977.
28. Blühm BH, Zhou X, Flaherty JE, Xu J, Dunkle LD (2007) Ras2 regulates growth and pathogenesis in Fusarium graminearum. Mol Plant Microbe Interact 20: 627–636.
29. Hofmann CS (2005) Glucose sensing via the protein kinase A pathway in Schizaosacharomyces pombe. Biochem Soc Trans 33: 257–260.
30. Froeliger EH, Carpenter BE (1996) NMR1, a major nitrogen regulatory gene in nitrogen metabolism repression. J Bacteriol 180: 627–636.
31. Mihlan M, Hemann V, Lai T-WD, Tuzdyski B (2003) AREA directly mediates nitrogen regulation of gibberellin biosynthesis in Gibberella fujikuroi, but its activity is not affected by NMR. Mol Microbiol 47: 917–925.
32. Wilson RA, Artrr Jr. RN (1996) Mutational analysis of AREA, a transcriptional activator mediating nitrogen metabolism repression in ApeGelder nubulin and a member of the ‘streetwise’ GATA family. Microbiol Mol Biol Reviews 62: 586–596.
33. Schöning B, Brown DW, Oester B, Tuzdyski B (2008) Cross-species hybridization with Fusarium verticillioides microarrays reveals new insights into...
Fusarium fujikuroi nitrogen regulation and the role of AreA and NMR. Eukaryotic Cell 7: 1831–1846.

35. Teichert S, Wottawa M, Schöning B, Tudzynski B (2006) Role of the Fusarium fujikuroi TOR kinase in nitrogen regulation and secondary metabolism. Eukaryotic Cell 5: 1007–1019.

36. Tudzynski B (1999) Biosynthesis of gibberellins in Gibberella fujikuroi biomolecular aspects. Appl Microbiol Biotechnol 52: 298–310.

37. López-Berges MS, Risapail N, Prados-Rosales RC, Di Pietro A (2010) A nitrogen response pathway regulates virulence functions in Fusarium oxysporum via the protein kinase TOR and the bZIP protein MeaB. The Plant Cell 22: 2459–2475.

38. Chiang TY, Marzluf GA (1994) DNA recognition by NIT2 nitrogen regulatory protein: importance of number, spacing, and orientation of GATA core elements and their flanking sequences upon NIT2 binding. Biochemistry 33: 576–582.

39. Cunningham TS, Cooper TG (1993) The Saccharomyces cerevisiae DAL80 repressor protein binds to multiple copies of GATAA-containing sequences (URS GATA). J Bacteriol 175: 5851–5861.

40. Frandsen EJN, Kristensen MB, Andersson JA, Giese H (2008) Efficient four fragment cloning for the construction of vectors for targeted gene replacement in filamentous fungi. BMC Mol Biol 9: 70.

41. Miller JD, Blackwell BA (1986) Biosynthesis of 3-acetyldeoxynivalenol and other metabolites by Fusarium culmorum HLX 1503 in a stirred jar fermenter. Can J Bot 64: 1–5.

42. Saleh A, Alvarez-Venegas R, Avramova Z (2008) An efficient chromatin immunoprecipitation (ChIP) protocol for studying histone modifications in Arabidopsis plants. Nat Protocols 3: 1018–1025.

43. Li H, Homer N (2010) A survey of sequence alignment algorithms for next-generation sequencing. Briefings in Bioinformatics 11: 473–483.

44. Krawitz P, Rödelisperger C, Jäger M, Jostins L, Bauer S, et al. (2010) Microindel detection in short-read sequence data. Bioinformatics 26: 722–729.