

Association of single nucleotide polymorphic sites in candidate genes with aggressiveness and deoxynivalenol production in *Fusarium graminearum* causing wheat head blight

Talas *et al.*

RESEARCH ARTICLE

Open Access

# Association of single nucleotide polymorphic sites in candidate genes with aggressiveness and deoxynivalenol production in *Fusarium graminearum* causing wheat head blight

Firas Talas<sup>1,3</sup>, Tobias Würschum<sup>1</sup>, Jochen C Reif<sup>1</sup>, Heiko K Parzies<sup>2</sup> and Thomas Miedaner<sup>1\*</sup>

## Abstract

**Background:** *Fusarium graminearum* sensu stricto (s.s.) is an ubiquitous pathogen of cereals. The economic impact of Fusarium head blight (FHB) is characterized by crop losses and mycotoxin contamination. Our objective was to associate SNP diversity within candidate genes with phenotypic traits. A total of 77 *F. graminearum* s.s. isolates was tested for severity of fungal infection (= aggressiveness) and deoxynivalenol (DON) production in an inoculated field experiment at two locations in each of two years. For seven genes known to control fungal growth (*MetAP1*, *Erf2*) or DON production (*TRI1*, *TRI5*, *TRI6*, *TRI10* and *TRI14*) single nucleotide polymorphic sites (SNPs) were determined and evaluated for the extent of linkage disequilibrium (LD). Associations of SNPs with both phenotypic traits were tested using linear mixed models.

**Results:** Decay of LD was in most instances fast. Two neighboring SNPs in *MetAP1* and one SNP in *Erf2* were significantly ( $P < 0.05$ ) associated with aggressiveness explaining proportions of genotypic variance ( $\rho_G$ ) of 25.6%, 0.5%, and 13.1%, respectively. One SNP in *TRI1* was significantly associated with DON production ( $\rho_G = 4.4$ ).

**Conclusions:** We argue that using the published sequence information of *Fusarium graminearum* as a template to amplify comparative sequence parts of candidate genes is an effective method to detect quantitative trait loci. Our findings underline the potential of candidate gene association mapping approaches to identify functional SNPs underlying aggressiveness and DON production for *F. graminearum* s.s. populations.

**Keywords:** Aggressiveness, Association mapping, DON, *Fusarium graminearum*, FHB, Linkage disequilibrium, QTL, *Triticum aestivum*, SNP

## Background

Fusarium head blight (FHB) is a destructive disease to cereals including wheat and barley. *Fusarium graminearum* (teleomorph *Gibberella zeae*) is considered to be the main causal agent of this disease in addition to other species such as *F. culmorum* [1]. *F. graminearum* complex has been subdivided into several cryptic species [2], the main member in Germany is *F. graminearum* sensu stricto (s.s.) [3]. FHB disease leads to prematurely bleached spikes in infected plants [1] with considerable

yield losses and contamination by mycotoxins such as deoxynivalenol (DON) [4]. DON is the most common food and feed contaminant in Europe and the maximum permissible level in unprocessed wheat used for food is 1.25 mg kg<sup>-1</sup> [5].

The quantitative ability of an isolate to cause disease on a susceptible host plant in a non-race specific pathosystem is defined as aggressiveness [6]. Aggressiveness is an important factor determining the potential ability of an isolate to cause yield losses. Large genetic variation of aggressiveness, type of mycotoxin, and DON production was found among isolates sampled in the same country or even the same field [7-11]. The molecular causes of this tremendous variation in aggressiveness

\* Correspondence: miedaner@uni-hohenheim.de

<sup>1</sup>Universitaet Hohenheim, State Plant Breeding Institute (720), Fruwirthstr. 21, 70593 Stuttgart, Germany

Full list of author information is available at the end of the article

and DON production in *F. graminearum* are still unclear. To date, only one QTL (quantitative trait locus) study from a single cross was reported detecting at least two QTL for aggressiveness [12]. It was shown before, that a large amount of molecular variation was found within field populations of *F. graminearum* s.s. using SSR (simple sequence repeat) markers. Comparing 12 individual field populations each consisting of 30 isolates, Talas et al. [3] reported that 71% of the molecular variation was assigned within populations and 29% between populations. A cause for this high diversity might be the mating system. *F. graminearum* is a haploid, homothallic fungus and propagates asexually by conidia and sexually by ascospores allowing both selfing and outcrossing [7]. With the availability of the full genomic sequence of *F. graminearum* [13], it is now possible to deeper analyze this variation by a candidate gene approach. An array of candidate genes, including components of transcription, signal transduction, host-specific nutrition, host infection/colonization and trichothecene biosynthesis is available in the version 3.2 of the Pathogen-Host Interactions database [14]. The candidate genes we were aiming for are known to control DON biosynthesis and/or aggressiveness (*TR11*, *TR15*, *TR16*, *TR110*, *TR114*) [15,16] or are expected to have a link to aggressiveness but are yet uncharacterized for *Fusarium* spp. (*MetAP1*, *Erf2*) [17,18]. Candidate gene association mapping is a sensitive tool if the mapping resolution is high enough [19]. The resolution of association mapping depends on the extent of linkage disequilibrium (LD), i.e. the non-random association of alleles present in a species [20]. Low LD was reported in *F. graminearum* estimated by different types of genetic makers such as VNTR, AFLP, RFLP [7,11,21], but no data on LD within and among genes is available. Because LD is species and population specific, it should consequently be determined before conducting an association mapping study.

The specific objectives of our study were to: (i) investigate the nucleotide diversity on gene level, (ii) investigate the extent of LD between single nucleotide polymorphic sites (SNPs) within and among candidate genes, (iii) identify SNPs of candidate genes *Erf2* and *MetAP1* for testing whether they affect the quantitative variation for aggressiveness, and (iv) evaluate associations of SNPs in *TR11* and other four genes of the *TR15* cluster with variation in DON production.

## Results

### Phenotypic analysis

Phenotypic data were analyzed in detail in a companion study [9]. All 77 *F. graminearum* s.s. isolates produced symptoms of FHB disease in all environments. Briefly, genotypic proportion of phenotypic variance of the

isolates was significant ( $P < 0.01$ ) for both traits: Aggressiveness measured as FHB rating on scale from 0 to 100 and DON production measured as DON concentration in wheat kernels in  $\text{mg kg}^{-1}$ . Isolate  $\times$  environment interaction variance was significant ( $P < 0.01$ ) only for mean FHB rating but not for DON production. Entry-mean heritability was moderate for mean FHB rating (0.55) and DON production (0.62).

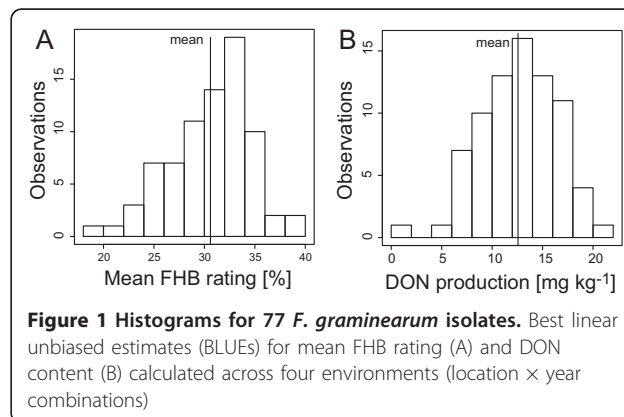
Histograms based on best linear unbiased estimators (BLUEs) followed a normal distribution for both traits (Figure 1). The mean of FHB infection among all isolates was 30.6% ranging from 18.4 to 38.9%. Average DON production was  $12.6 \text{ mg kg}^{-1}$  ranging from 3.5 to  $21.6 \text{ mg kg}^{-1}$ . One isolate (WET24) produced a very low concentration of DON ( $0.3 \text{ mg kg}^{-1}$ ) and was found to be a nivalenol producer by chemotype-specific primers.

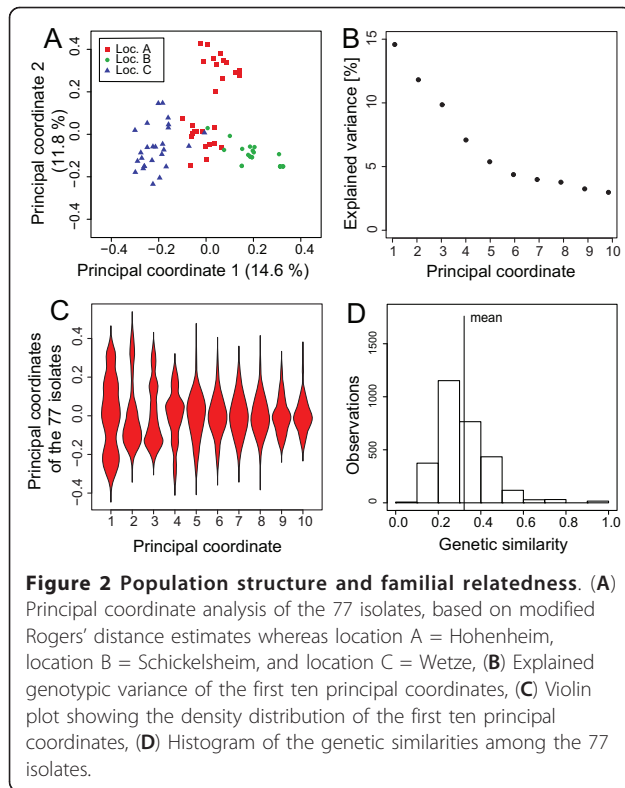
### Analysis of population structure and diversity

Principal coordinate analysis (PCoA) based on modified Rogers' distances between all isolates did not show a distinct separation of the isolates sampled from different locations (Figure 2A). Explained variance gradually decreased according to the first ten principal coordinates (Figure 2B). The violin plot (Figure 2C) had a continuous density of distribution over all ten principal coordinates without any division within the principal coordinates. Genetic similarity ranged from 0.057 to 1.0 with a mean value of 0.31 (Figure 2D).

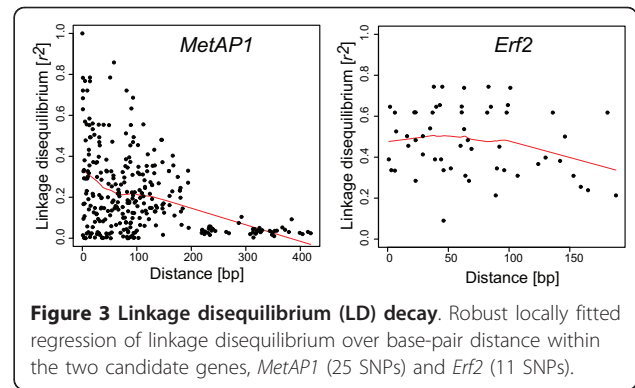
### Estimation of nucleotide diversity and linkage disequilibrium

Percentage of polymorphic sites per total sequenced region of each gene (without singletons) varied from 0.9% (5/513) on *TR110b* to 8.8% (65/734) on *MetAP1* indicating a high nucleotide diversity in most tested genes (Table 1). LD of SNPs within the gene *MetAP1* decayed rapidly within 200 bp of physical distance, i.e., the robust locally fitted regression of  $r^2$  values has a trend to decay from  $r^2 = 0.35$  to  $r^2 < 0.1$ , whereas LD





within the gene *Erf2* had  $r^2$  values ranging from 0.8 to 0.2 and the regression of  $r^2$  trends to decay already after 150 bp (Figure 3). LD with  $r^2$  values higher than 0.1 were detected between all allele combinations within the tested genes followed by a rapid decay negatively correlated with the physical distance in base pair (Figure 4). In line with the rapid decay, LD between genes located



on the same chromosome (i.e., *TR11* and *MetAP1*) is low ( $r^2 < 0.1$ ). Interestingly, 48%, 19%, and 45% of the SNP pairs between the genes *TR110/MetAP1*, *MetAP1/TR15* and *MetAP1/Erf2*, respectively, have higher values of  $r^2$  than 0.1, although they are located on different chromosomes. Low  $r^2$  values ( $< 0.1$ ) were observed for SNPs of gene pairs *TR110/Erf2* and *TR15/Erf2*, whereas  $r^2$  values of 0.2 were detected between SNP pairs of *TR110b* and *TR15*, followed by a rapid decay of LD.

**Association analysis for aggressiveness and mycotoxin**

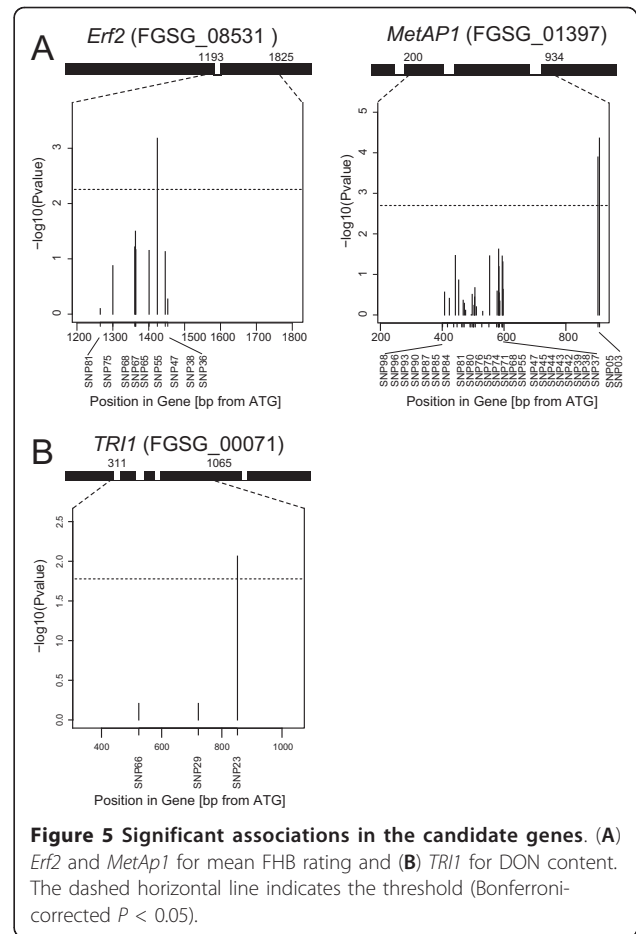
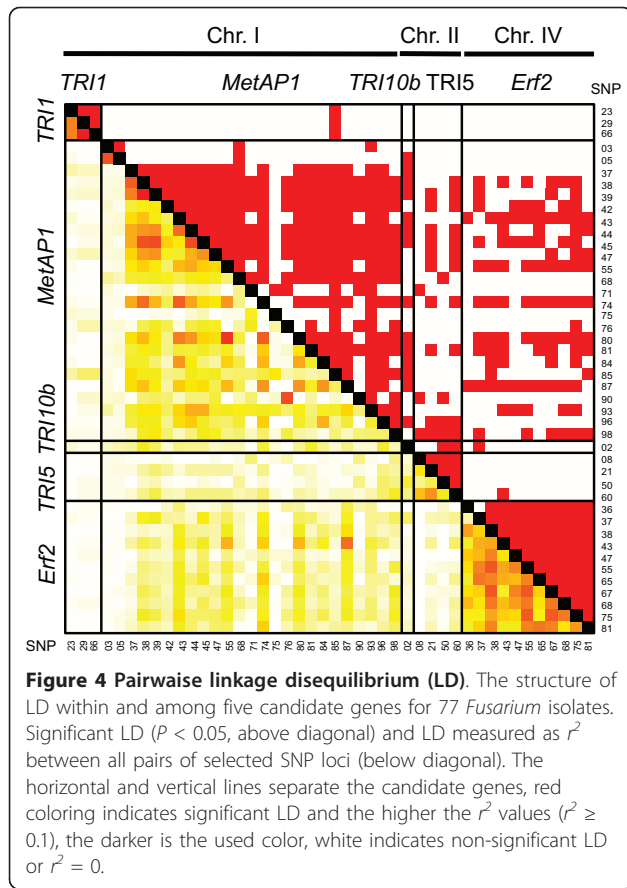
Two adjacent SNPs of the gene *MetAP1* were significantly associated ( $P < 0.05$  using Bonferroni-Holm correction) [22] with mean FHB rating with an explained genotypic proportion of variance  $p_G = 25.6$  and 0.5, respectively (Table 2, Figure 5A). Additionally, one SNP in the gene *Erf2* was significantly associated with this trait showing a  $p_G$  of 13.1% (Table 2, Figure 5A). A single SNP significantly associated with DON production was identified in the *TR11* gene explaining 4.4% of the

**Table 1 Sequenced regions of candidate genes and its related nucleotide polymorphism**

| Gene ID <sup>a</sup>           | Sequenced region relative to the ATG | No. of SNPs detected | No. of singletons | Nucleotide diversity (%) <sup>b</sup> | No. of SNPs with allele frequency > 0.1 |
|--------------------------------|--------------------------------------|----------------------|-------------------|---------------------------------------|---|
| <i>TR11</i><br>(FGSG_00071)    | 311-1065                             | 73                   | 54                | 2.5                                   | 3                                       |
| <i>TR16</i><br>(FGSG_16251)    | 35-558                               | 36                   | 29                | 1.3                                   | 0                                       |
| <i>TR15</i><br>(FGSG_03537)    | 479-1080                             | 70                   | 38                | 5.3                                   | 4                                       |
| <i>TR110 a</i><br>(FGSG_03538) | 31-678                               | 123                  | 80                | 6.6                                   | 0                                       |
| <i>TR110 b</i><br>(FGSG_03538) | 760-1273                             | 34                   | 29                | 0.9                                   | 1                                       |
| <i>TR114</i><br>(FGSG_03543)   | 277-976                              | 43                   | 23                | 2.9                                   | 0                                       |
| <i>MetAP1</i><br>(FGSG_01397)  | 200-934                              | 106                  | 41                | 8.8                                   | 25                                      |
| <i>Erf2</i><br>(FGSG_08531)    | 1193-1825                            | 80                   | 42                | 6.0                                   | 11                                      |

<sup>a</sup> The given gene ID is the entry number of the MIPS *F. graminearum* genome database (FGDB; Wong et al. 2011)

<sup>b</sup> Nucleotide diversity is the frequency of SNPs (without singletons) relative to the total length of the sequenced gene region



genotypic variance of DON production (Table 2, Figure 5B). All detected SNPs that were associated to the mentioned phenotypic traits were non-synonymously substituted (Table 2). Applying a haplotype analysis across genes, we could confirm a significant association between mean FHB rating and *MetAP1* and between DON production and *TRI1* (Additional file1).

### Discussion

Association mapping based on candidate genes is a promising tool for high-resolution mapping of genes contributing to quantitative traits [23]. Nevertheless, it has not

yet applied to investigate the basis of quantitative variation in aggressiveness and/or DON production in *F. graminearum*. This is a totally different approach than using knock-out mutants [24], because we are aiming for the analysis of quantitative differences produced by single nucleotide changes of the respective genes in a set of 77 isolates. The high number of singletons may, besides sequencing noise, be caused by frequent sexual recombination of the fungus. Therefore, we used a strict threshold and were analyzing only SNPs with allele frequencies  $> 0.1$ .

**Table 2** SNPs in the candidate genes significantly associated with mean FHB rating or DON content

| Trait/Candidate gene                  | SNP #  | Position <sup>a</sup> | Polymorphism | Changes in amino acids <sup>b</sup> | P-value               | $p_G$ (%) <sup>c</sup> |
|---------------------------------------|--------|-----------------------|--------------|-------------------------------------|-----------------------|------------------------|
| Mean FHB rating (%)                   |        |                       |              |                                     |                       |                        |
| <i>MetAP1</i>                         | SNP 05 | 904                   | A, G         | H, R                                | $1.28 \times 10^{-4}$ | 25.6                   |
|                                       | SNP 03 | 909                   | A, C         | T, P                                | $1.94 \times 10^{-4}$ | 0.5                    |
| <i>Erf2</i>                           | SNP 47 | 1424                  | C, T         | Q, Stop                             | $6.61 \times 10^{-4}$ | 13.1                   |
| DON production (mg kg <sup>-1</sup> ) |        |                       |              |                                     |                       |                        |
| <i>TRI1</i>                           | SNP 23 | 851                   | A, G         | R, Q                                | $8.62 \times 10^{-3}$ | 4.4                    |

<sup>a</sup> The position is relative to the start codon (ATG)

<sup>b</sup> Different amino acids coded by SNPs located in exon regions (non-synonymous SNPs)

<sup>c</sup> The proportion of explained genotypic variance in percent ( $p_G$ ) after Bonferroni-Holm procedure

### Choice of candidate genes

The genes for deoxynivalenol (DON) biosynthesis reside primarily in a 25 kb cluster (*TRI5* cluster), *TRI1* gene belong to a second smaller cluster that has a different chromosomal localization [16]. *TRI1*, *TRI5*, *TRI14* have a direct effect on the production of DON or acetylated DON, *TRI6* and *TRI10* are regulating this pathway. *TRI14* export the deoxynivalenol outside the mycelia in addition to its major role in aggressiveness [15].

The function of *MetAP1* is still not described in *Fusarium* spp., a deletion, however, reduces growth in *Saccharomyces cerevisiae* [18,25,26]. In both, prokaryotes and eukaryotes the N-terminal methionine is often cleaved by methionine aminopeptidase encoded by the gene *MetAP1* [18]. *Erf2* gene is a component of the *RAS* protein subcellular localization pathway in yeast [17]. *RAS2* gene is known to affect the pathogenesis of *F. graminearum* by regulation of hyphal growth and expression of hydrolytic enzymes [27]. Failing of palmitoylation mainly controlled by the *Erf2* gene in *S. cerevisiae* reduces amount of *RAS* protein at the plasma membrane and lead to poor growth of yeast [17]. Both genes are not indispensable for fungal growth because their deletion does not prevent any growth at least in *S. cerevisiae*. And the 77 isolates analyzed in this study, all had a similar growth *in vitro*, produced enough spores for inoculation and were able to infect the host in the field. Significant, quantitative differences for aggressiveness and DON production, however, were found illustrating those genes with regulatory functions might be involved. Implementing associations between new candidate genes and important phenotypic traits may provide a useful tool to rapidly check a possible contribution of these genes to QTL of DON production or aggressiveness in *F. graminearum* as shown previously in mammal research [23]. If an association is found, a more detailed analysis of the function of the respective genes should, of course, follow.

### Molecular diversity of the 77 isolates

All isolates were checked by respective primers for their designation to *F. graminearum* s.s. Nucleotide diversity within candidate genes of this study ranged from 1 to 9 per 100 bp being high compared to 0 to 2 SNPs per 100 bp reported by Cuomo et al. [13]. This was not unexpected since Cuomo et al. [13] used just two isolates (PH-1 and GZ3639) while in this study 77 isolates were analyzed. Many genes in *F. graminearum* are localized in regions of high SNP density, i.e. highly variable regions, especially those expressed during plant infection. Moreover, it is known that the density of SNPs is biased in *F. graminearum* in a way that 50% of the SNPs are within 13% of the genome sequence, the highest SNP density was reported in some regions of

chromosome II, where most of the *TRI* genes are located [13]. The high nucleotide diversity of these groups of genes suggests that the fungus has a great capacity for adaptability and genetic change during its interaction with the host plant [13]. Despite this, we found much more SNPs with allele frequencies > 0.1 in the *MetAP1* and *Erf2* genes than in the *TRI* genes analyzed. This was not the case, when just the number of detected SNPs is regarded, indicating that nucleotide diversity may be similar, but more rare haplotypes occur in the *TRI* genes. Moreover, high SNPs density was detected in one or two large interstitial regions on chromosomes I, II, and IV [13].

### Population structure and consequences for association mapping

Correcting for population and/or family structure is essential for association mapping to decrease the number of false positive QTL [28]. An appropriate statistical model should provide an excellent compromise between correcting for population stratification to decrease the probability of detecting false positive SNP-trait associations but still retaining enough information within the SNPs for QTL detection [29]. Although some grouping according to the sampling location might be seen in the PCoA, the violin plots did not show any separated knots on the first ten principle coordinates (Figure 2). Thus, no distinct subpopulations were identified and correction for familial relatedness between isolates should be sufficient. Therefore, trait-SNP association was investigated in detail with the K model, which incorporated estimates of kinship coefficients based on SSR data.

### Extent of linkage disequilibrium and resolution of association mapping

Rapid decay of regression parameter  $r^2$  in a short physical distance of 200 bp within two genes was presented in this study (Figure 3), providing a high resolution of association mapping. Generally, low LD was reported in *F. graminearum* populations using selection-neutral markers (VNTR, AFLP, or RFLP) especially in the region that includes the *TRI5* gene cluster [11]. Moreover, the weak correlation between  $r^2$  and physical distance in addition to the large proportion of unlinked SNP with significant LD, such as SNPs between gene pairs *TRI5/MetAP1*, *TRI10/MetAP1*, or *MetAP1/Erf2* (Figure 4), suggest the presence of other forces generating LD between unlinked SNPs. The values of  $r^2$  between most pairs of SNPs in *TRI5/TRI10b* are > 0.1, presenting these two genes in LD, that might refer to the controlling role of *TRI10* on *TRI5* gene cluster [24,30,31]. That *TRI10* and *TRI5* are located close to each other on the same chromosome play a minor role hence the detected LD decay is occurring within 200 bp.

Selection acting on oligo- or polygenic traits such as aggressiveness and DON production might be responsible for this [32]; an alternative explanation would be the involvement in the same trait network with other physiologically important genes.

#### Association mapping of genes underlying aggressiveness and DON production

We identified three SNPs related to aggressiveness and one SNP related to DON production using the fairly conservative Bonferroni-Holm correction. Setting the threshold for detection of trait associations to an allele frequency of 0.1 might underestimate the number of QTL detected but is due to the restricted population size. Association between SNPs of candidate gene *Erf2* and phenotypic data revealed a significant association to mean FHB rating, but not to DON content. This might be explained by the fact that the *Erf2* gene is involved in *RAS2* processing or trafficking that precedes palmitoylation of *RAS2* genes [17]. *RAS2* is known to regulate the aggressiveness through affecting fungal growth and regulating other pathogenicity genes, e.g., *Gmpk1*, which controls the induction of extracellular enzymes required for pathogenesis [27]. Further study is needed to understand how *Erf2* affects differences in aggressiveness of *F. graminearum* isolates.

Two SNPs in *MetAP1* gene significantly associated with aggressiveness were located adjacent to each other forming a collinearity pattern. The role of *MetAP1* was reported in *S. cerevisiae* as reducing cell growth by N-terminal protein modification. A similar role might be expected for *MetAP1* in *F. graminearum*, hence significant positive correlations between aggressiveness and fungal biomass ( $r = 0.7$ ,  $P = 0.01$ ), and fungal biomass and DON content ( $r = 0.8$ ,  $P = 0.01$ ) were reported among 50 *F. graminearum* isolates [33]. Nevertheless, analyzed isolates were sampled from visually diseased spikelets, thus all were able to infect wheat ears and induce symptoms, non-aggressive isolates were not included in our *F. graminearum* population sample.

A single SNP detected in *TR11* was associated with DON production in *F. graminearum*. *TR11* encodes a cytochrome P450 monooxygenase that catalyzes hydroxylation of C-8 position during trichothecene biosynthesis [34]. This confirmation change obviously does not affect aggressiveness, because the respective SNP is just associated with DON content but not with aggressiveness.

#### Conclusions

This is the first candidate gene association mapping study provided insights on some genes involved in aggressiveness and DON concentration of *F. graminearum* s.s. The described associations should be validated using a larger number of isolates and different

environments. The validated genes are an important starting point for further functional analyses.

#### Methods

##### Fungal material

Ears of winter wheat (*Triticum aestivum* L.) showing symptoms of FHB were sampled from three commercial fields in Germany to establish a fungal population of *F. graminearum* s.s.: Stuttgart-Hohenheim in southwest Germany in 2008, Wetze and Schickelsheim in Lower Saxony in 2006 and 2007, respectively. From each infected head, one isolate was recovered, transferred onto a fresh SNA plate and placed under permanent UV light for induction of sporulation as described recently [3]. From each isolate, one single spore was picked out under the microscope and transferred onto a fresh SNA plate to establish a single-spore culture. The single-spore isolates were checked morphologically and analyzed for their species specificity and chemotype by different PCR-based assays as described recently in detail [3]. In total, 77 single-spore isolates of *F. graminearum* s.s. (former lineage 7) were inoculated on the moderately resistant German spring wheat cultivar Taifun (KWS LOCHOW GMBH, Bergen, Germany) for analyzing aggressiveness and DON production in the field.

##### Design of field studies

Field experiments were planted at each of two locations in 2009 and 2010: Hohenheim (HOH, longitude 9° 12' 58", latitude 48° 42' 50", altitude 400 m) and Oberer Lindenhof (OLI, longitude 9° 18' 12", latitude 48° 28' 26", altitude 700 m), resulting in four environments (location × year combinations). Mean annual temperatures at HOH and OLI were 10.1°C and 9.0°C, respectively, mean annual precipitations were 644 mm and 723 mm, respectively, across 2009 and 2010. Plants were grown in two-rowed micro-plots of 1 m length and 0.42 m width. Plots were arranged in a chessboard-like design, i.e., each plot with a wheat entry was bordered by four plots of similar size that were planted with a long-strawed spring triticale cultivar Nilex (NORDSAAT GmbH, Halberstadt, Germany) to reduce inter-plot interference caused by drifting of inoculum during spraying or secondary distribution of spores. The experiment was arranged in a split-plot design with two replications. The main plot factor was the *F. graminearum* s.s. population, the subplot factor the isolate. Both factors were assigned to a randomized complete block design. Eight plots per replication included in the subplot factor were not inoculated to estimate the degree of natural infection.

Inoculum was sprayed with a concentration of  $2 \times 10^5$  spores ml<sup>-1</sup> onto the wheat heads of each plot.

Inoculation was performed at full flowering time of the wheat cultivar to ensure maximum susceptibility of wheat to *F. graminearum*. Fusarium head blight (FHB) aggressiveness was rated visually four times as the percentage of infected spikelets per plot (0-100%). This rating included both the number of infected spikes per plot and the number of infected spikelets per spike. To compare FHB reactions, the arithmetic means of four ratings were used and assigned as mean FHB rating throughout the paper. All plots were harvested by hand, threshed, and the grain analyzed to quantify the amount of DON by a commercially available immunotest (R-biopharm AG, Darmstadt, Germany) as previously described in detail [9]. This test cannot differentiate between DON and 3-ADON, so the results include both mycotoxins. Generally, however, the amount of 3-ADON is only 2-3% of total DON content [35]. In each of the four environments, the natural infection rate was very low ranging from 0 to 3% FHB rating and from 0 to 0.34 mg kg<sup>-1</sup> DON concentration. Prediction of NIV chemotype was done using the primers N11, 15D11, 3D11, and 11R in a multiplex polymerase chain reaction (PCR) as designed by Zhang et al. [36].

#### Detection of population structure and gene sequencing

In order to analyze the population structure, all isolates were fingerprinted with 19 simple sequence repeat (SSR) markers dispersed throughout the whole genome following standard protocols [3].

Five candidate genes were chosen according to their role in trichothecene biosynthesis and/or aggressiveness (*TR11*, *TR15*, *TR16*, *TR110*, and *TR114*). Two other, yet uncharacterized candidate genes in *Fusarium* spp. (*Erf2*, *MetAP1*) were tested for a possible association with aggressiveness. Nucleotide sequences of these genes were imported from *Fusarium graminearum* database FGDB [37]. The selected genes were (Table 3): *TR11* (FGSG\_00071), *TR15* (FGSG\_03538), *TR16* (FGSG\_16251), *TR110* (FGSG\_03538), *TR114* (FGSG\_03543), *MetAP1* (FGSG\_01397), and *Erf2* (FGSG\_08531). Specific primers were designed (see Additional file 2) to amplify parts of these genes using the software Primer Premier 4.0 (Premier Biosoft International, CA, U.S.A.).

Polymerase chain reaction (PCR) was performed using the designed primers related to each gene separately following standard protocols, however, with a different annealing temperature for each gene (Table 3). Purification of PCR products was performed by precipitating the DNA with 10% (v/v) of 3 M Sodium Acetate and 25% (v/v) of absolute cold ethanol overnight. Precipitated DNA was cleaned twice with 70% ethanol and finally diluted in 10 µl ddH<sub>2</sub>O. Purified PCR products were sequenced (QIAGEN® Sequencing Services,

Hilden, Germany). The sequence was performed once for each isolate, hence the alignment was mostly identical (the sequence was repeated in case it fail to align to the other sequence or in case of high noise in the sequence). Expected sizes of PCR products were obtained for all tested genes. All sequences of *TR16* and *TR114* were located in the coding region of the genes (exon, see Additional file 1). Sequences of the genes *TR15*, *MetAP1*, *Erf2*, and *TR110* (two parts) were stretched over two exons. Sequences of genes were obtained from right to left sides of PCR products, just genes *MetAP1* and *Erf2* were obtained from left to right side. Sequence of *TR11* was located over four exons according to the reference sequence. Number of identified SNPs was high over all sequenced regions of candidate genes (Table 1). The sequences were aligned (according to the initial matrix published in FGDB) using CLC sequence viewer 6.3 (CLC-bio, Aarhus, Denmark) to identify single nucleotide polymorphisms (SNPs) among the 77 isolates.

#### Phenotypic data analyses

The following linear mixed model was used to estimate variance components:  $y_{ijn} = \mu + Iso_i + Env_j + (Iso \times Env)_{ij} + Rep_n + e_{ijn}$ , where  $\mu$  is the population mean,  $Iso_i$  the genetic effect of the  $i$ th isolate,  $Env_j$  the effect of the  $j$ th environment,  $Iso \times Env$  the isolate times environment interaction effect,  $Rep_n$  the effect of the  $n$ th replication, and  $e_{ijn}$  the residual error. Variance components were determined by the restricted maximum likelihood (REML) method using the software ASReml 2.0 (VSN International Ltd, Hemel Hempstead, U.K.). Significance for variance component estimates was tested by model comparison with likelihood ratio tests where the halved  $P$  values were used as an approximation [38]. Heritability ( $h^2$ ) on an entry-mean basis was estimated as the ratio of genotypic to phenotypic variance according to Melchinger et al. [39]. Furthermore, genotypes were regarded as fixed effects and best linear unbiased estimates (BLUEs) were determined for all isolates and traits.

Genetic relatedness among the 77 isolates was determined by applying principal coordinate analysis (PCoA) [40] based on the modified Rogers' distances of the isolates [41]. Linkage disequilibrium (LD) between the selected SNPs was assessed by the LD measure  $r^2$  [42] and significance of LD was tested with Fisher's exact tests [43]. LD and PCoA computations were performed with the software package Plabsoft [44].

#### Association analysis

A two-step association approach was applied in this study and the BLUEs per environment were used as input for the association analysis. The linear mixed



**Table 3 Name of tested genes, chromosomal localization, primer sequence and expected amplified DNA products**

| No. | Gene ID <sup>a</sup>        | Chromosome |   | Sequences of the primers | Ta (°C) | Expected products (bp) | Gene size (bp) |
|-----|-----------------------------|------------|---|--------------------------|---------|------------------------|----------------|
| 1   | <i>TRI1</i> (FGSG_00071)    | I          | F | CACCAGTTGCAGGATGT        | 53.2    | 850                    | 1980           |
|     |                             |            | R | AATGGGAGTGATTAGTTCCG     |         |                        |                |
| 2   | <i>TRI5</i> (FGSG_03537)    | II         | F | TGGCGGATCTATCTATTCA      | 57.1    | 750                    | 1129           |
|     |                             |            | R | CTTCTGGCGTCCTCTGT        |         |                        |                |
| 3   | <i>TRI6</i> (FGSG_1625)     | II         | F | ATGGAGGCCGAATCTCAC       | 54.4    | 600                    | 673            |
|     |                             |            | R | CCACCCCTGCTAAAGACCCT     |         |                        |                |
| 4   | <i>TRI10 a</i> (FGSG_03538) | II         | F | ATTTCCCAAAGCCTAGACAA     | 54.2    | 700                    | 1264           |
|     |                             |            | R | GGCCGTAATCTTCAAATGGT     |         |                        |                |
| 5   | <i>TRI10 b</i> (FGSG_03538) | II         | F | GCACCATTTGAAGATTACGG     | 54.2    | 550                    |                |
|     |                             |            | R | CGTCAAGTCTTCCCATCTCA     |         |                        |                |
| 6   | <i>TRI14</i> (FGSG_03543)   | II         | F | AGCCCGAACCCGCTACAT       | 57.3    | 750                    | 1116           |
|     |                             |            | R | CGAACCTGCTGCTCTTAC       |         |                        |                |
| 7   | <i>MetAPI</i> (FGSG_01397)  | I          | F | AACGATGCCGGATCGTCA       | 55.3    | 800                    | 1403           |
|     |                             |            | R | TTGCCAAACTCTCGGATCAA     |         |                        |                |
| 8   | <i>Erf2</i> (FGSG_08531)    | IV         | F | AGGCATCTTTGTTGTTGTG      | 52.4    | 750                    | 2040           |
|     |                             |            | R | TTGTAATACGTGGGTTGT       |         |                        |                |

<sup>a</sup> The given gene ID is the entry number of the MIPS *F. graminearum* genome database (FGDB; Wong et al. 2011)

model for the association approach was:  $y_{ijn} = \mu + a_p + Iso_i + Env_j + e_{ijp}$ , where  $a_p$  is the effect of allele  $p$ . The allele effect  $a_p$  was modeled as fixed effect whereas  $Iso_i$  and  $Env_j$  were regarded as random effects. We assumed that the variance of the random genetic effect was  $Var(g) = 2K\sigma_g^2$ , where  $\sigma_g^2$  refers to the genetic variance estimated by REML and  $K$  was a  $77 \times 77$  matrix of kinship coefficients that define the degree of genetic covariance between all pairs of entries. We followed the suggestion of Bernardo [45] and calculated the kinship coefficient  $K_{ij}$  between isolates  $i$  and  $j$  on the basis of the SSR marker data as  $K_{ij} = 1 + (S_{ij} - 1) / (1 - T_{ij})$ , where  $S_{ij}$  is the proportion of marker loci with shared variants between isolates  $i$  and  $j$ , and  $T_{ij}$  is the average probability that a variant from one isolate  $i$  and a variant from one isolate  $j$  are alike in state, given that they are not identical by descent. The coefficient  $T_{ij}$  was estimated separately for each gene and trait using a REML method setting negative kinship values between isolates to zero. SNPs with allele frequencies  $< 0.1$  were not considered in the association analysis or LD estimation. The obtained optimum T values were for DON content: *TRI1* (0.025), *TRI5* (0.575), *TRI10* (0.375), *MetAPI* (0.300), and *Erf2* (0.400). For mean FHB rating: *TRI1* (0.600), *TRI5* (0.150), *TRI10* (0.200), *MetAPI* (0.075), and *Erf2* (0.275). For the detection of main effects of the candidate gene SNPs, these were fitted as fixed effects in the mixed model and their significance was tested by a Wald  $F$  test. Based on the Wald  $F$  statistic, we performed tests for the presence of marker-phenotype associations with a significant ( $P < 0.05$ ) effect on DON content and mean FHB rating applying the Bonferroni-Holm procedure [22] to correct for multiple testing.

The proportion of genotypic variance ( $p_G$ ) explained by the detected SNP was calculated by fitting each SNP in a linear model to obtain  $R^2_{adj}$ . The ratio  $p_G = R^2_{adj} / h^2$  yielded the proportion of genotypic variance [46]. In the case of *MetAPI*, where two SNPs were detected for mean FHB rating, both were simultaneously fitted in the linear model in the order of their  $P$  values to correct for collinearity. Haplotype analyses were performed with the same procedure to test the association with both phenotypic traits (Additional file 2: Table S1). We applied the same threshold level like for SNP analyses. All mixed model calculations were performed using the software ASReml 2.0 (VSN International Ltd, Hemel Hempstead, U.K.).

## Additional material

**Additional file 1: Genomic organizations of the seven candidate genes with exons in black color.** The sequenced region is shown as a green bar above each gene.

**Additional file 2: Table S1.** Haplotype analysis of the genes *MetAPI* and *TRI1*.

## Acknowledgements

This work was funded by the German Academic Exchange Service (DAAD), Bonn, Germany and by the State Plant Breeding Institute (720), Universitaet Hohenheim, Germany. The authors like to thank Mrs. B. Lieberherr for excellent assistance in field work.

## Author details

<sup>1</sup>Universitaet Hohenheim, State Plant Breeding Institute (720), Fruwirthstr. 21, 70593 Stuttgart, Germany. <sup>2</sup>Universitaet Hohenheim (350), Institute of Plant Breeding, Seed Science & Population Genetics, Fruwirthstr. 21, 70593 Stuttgart, Germany. <sup>3</sup>National Commission of Biotechnology (NCBT), P. O. Box. 31902, Damascus, Syria.

#### Authors' contributions

FT carried out the phenotypic and molecular analyses, performed parts of the statistical analyses and drafted the manuscript. TW performed parts of the statistical analyses and helped to draft the manuscript, JCR edited the manuscript, HKP† participated in the design of the study and supported the technical realization. TM participated in the design of the field study and helped to draft the manuscript. All authors read and approved the final manuscript.

#### Competing interests

The authors declare that they have no competing interests.

Received: 23 September 2011 Accepted: 12 March 2012

Published: 12 March 2012

#### References

1. Agrios GN: *Plant Pathology*. 5 edition. San Diego: Elsevier Press; 2004.
2. O'Donnell K, Ward TJ, Geiser DM, Kistler HC, Aoki T: **Genealogical concordance between the mating type locus and seven other nuclear genes supports formal recognition of nine phylogenetically distinct species within the *Fusarium graminearum* clade.** *Fungal Genet Biol* 2004, **41**(6):600-623.
3. Talas F, Parzies HK, Miedaner T: **Diversity in genetic structure and chemotype composition of *Fusarium graminearum* sensu stricto populations causing wheat head blight in individual fields in Germany.** *Eur J Plant Pathol* 2011, **131**(1):39-48.
4. Miedaner T, Cumagun CJR, Chakraborty S: **Population genetics of three important head blight pathogens *Fusarium graminearum*, *F. pseudograminearum* and *F. culmorum*.** *J Phytopathol* 2008, **156**(3):129-139.
5. European Commission: *Commission Regulation (EC) No. 1126/2007 of 28 September 2007 amending Regulation (EC) No. 1881/2006 setting maximum levels for certain contaminants in foodstuffs as regards Fusarium toxins in maize and maize products 2007*, Official Journal of the European Union L255, 14-17. Published online.
6. Vanderplank J: *Disease Resistance of Plants* New York: Academic Press; 1968.
7. Gale LR, Chen LF, Hernick CA, Takamura K, Kistler HC: **Population analysis of *Fusarium graminearum* from wheat fields in Eastern China.** *Phytopathology* 2002, **92**(12):1315-1322.
8. Miedaner T, Schilling AG, Geiger HH: **Molecular genetic diversity and variation for aggressiveness in populations of *Fusarium graminearum* and *Fusarium culmorum* sampled from wheat fields in different countries.** *J Phytopathol* 2001, **149**:641-648.
9. Talas F, Kalih R, Miedaner T: **Within-field variation of *Fusarium graminearum* sensu stricto isolates for aggressiveness and deoxynivalenol production in wheat head blight.** *Phytopathology* 2012, **102**(1):128-134.
10. Ward TJ, Clear RM, Rooney AP, O'Donnell K, Gaba D, Patrick S, Starkey DE, Gilbert J, Geiser DM, Nowicki TW: **An adaptive evolutionary shift in *Fusarium* head blight pathogen populations is driving the rapid spread of more toxigenic *Fusarium graminearum* in North America.** *Fungal Genet Biol* 2008, **45**(4):473-484.
11. Zeller KA, Bowden RL, Leslie JF: **Population differentiation and recombination in wheat scab populations of *Gibberella zeae* from the United States.** *Mol Ecol* 2004, **13**(3):563-571.
12. Cumagun CJR, Bowden RL, Jurgenson JE, Leslie JF, Miedaner T: **Genetic mapping of pathogenicity and aggressiveness of *Gibberella zeae* (*Fusarium graminearum*) towards wheat.** *Phytopathology* 2004, **94**(5):520-526.
13. Cuomo CA, Guldener U, Xu JR, Trail F, Turgeon BG, Di Pietro A, Walton JD, Ma LJ, Baker SE, Rep M, Adam G, Antoniw J, Baldwin T, Calvo S, Chang YL, DeCaprio D, Gale LR, Gnerre S, Goswami RS, Hammond-Kosack K, Harris LJ, Hilburn K, Kennell JC, Kroken S, Magnuson JK, Mannhaupt G, Mauceli E, Mewes HW, Mitterbauer R, Muehlbauer G, et al: **The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization.** *Science* 2007, **317**(5843):1400-1402.
14. Winnenburg R, Urban M, Beacham A, Baldwin TK, Holland S, Lindeberg M, Hansen H, Rawlings C, Hammond-Kosack KE, Kohler J: **PHI-base update: additions to the pathogen host interaction database.** *Nucleic Acids Res* 2008, **36**:572-576.
15. Dyer RB, Plattner RD, Kendra DF, Brown DW: ***Fusarium graminearum* *TRI1* is Required for high virulence and DON production on wheat but not for DON synthesis *in vitro*.** *J Agric Food Chem* 2005, **53**:9281-9287.
16. Kimura M, Tokai T, Takahashi-Ando N, Ohsato S, Fujimura M: **Molecular and genetic studies of *Fusarium trichothecene* biosynthesis: pathway genes and evolution.** *Biosci Biotechnol Biochem* 2007, **71**(9):2105-2123.
17. Bartels DJ, Mitchell DA, Dong X, Deschenes RJ: ***Erf* a novel gene product that affects the localization and palmitoylation of *RAS* in *Saccharomyces cerevisiae*.** *Mol Cell Biol* 1999, **19**(10):6775-6787.
18. Chen S, Vetro JA, Chang YH: **The specificity *in vivo* of two distinct methionine aminopeptidases in *Saccharomyces cerevisiae*.** *Arch Biochem Biophys* 2002, **408**(1):87-93.
19. Yu J, Buckler ES: **Genetic association mapping and genome organization of maize.** *Curr Opin Biotechnol* 2006, **17**(2):155-160.
20. Wilson LM, Whitt SR, Ibanez AM, Rocheford TR, Goodman MM, Buckler ES: **Dissection of maize kernel composition and starch production by candidate gene association.** *Plant Cell* 2004, **16**(10):2719-2733.
21. Karugia GW, Suga H, Gale LR, Nakajima T, Tomimura K, Hyakumachi M: **Population structure of *Fusarium graminearum* species complex from a single Japanese wheat field sampled in two consecutive years.** *Plant Dis* 2009, **93**(2):170-174.
22. Holm S: **A simple sequentially rejective multiple test procedure.** *Scand J Stat* 1979, **6**:65-70.
23. Magee DA, Sikora KM, Berkowicz EW, Berry DP, Howard DJ, Mullen MP, Evans RD, Spillane C, MacHugh DE: **DNA sequence polymorphism in a panel of eight candidate bovine imprinted genes and their association with performance traits in Irish Holstein-Friesian cattle.** *BMC Genet* 2010, **11**:93.
24. Proctor RH, Hohn TM, McCormick SP, Desjardins AE: ***TRI* encodes an unusual zinc finger protein involved in regulation of trichothecene biosynthesis in *Fusarium sporotrichioides*.** *Appl Environ Microbiol* 1995, **61**(5):1923-1930.
25. Dummitt B, Micka WS, Chang YH: **Yeast glutamine fructose 6-phosphate aminotransferase (*Gfa*) requires methionine aminopeptidase activity for proper function.** *J Biol Chem* 2005, **280**(14):14356-14360.
26. Li X, Chang YH: **Amino terminal protein processing in *Saccharomyces cerevisiae* is an essential function that requires two distinct methionine aminopeptidases.** *Proc Natl Acad Sci* 1995, **92**(26):12357-12361.
27. Bluhm BH, Zhao X, Flaherty JE, Xu JR, Dunkle LD: ***RAS* regulates growth and pathogenesis in *Fusarium graminearum*.** *Mol Plant Microbe Interact* 2007, **20**(6):627-636.
28. Reif JC, Maurer HP, Korzun V, Ebmeyer E, Miedaner T, Würschum T: **Genetic architecture of grain yield and heading time in soft winter wheat.** *Theor Appl Genet* 2011, **123**(2):283-292.
29. Würschum T, Maurer HP, Schulz B, Möhring J, Reif JC: **Genome-wide association mapping reveals epistasis and genetic interaction networks in sugar beet.** *Theor Appl Genet* 2011, **123**(1):109-118.
30. Peplow AW, Tag AG, Garifullina GF, Beremand MN: **Identification of new genes positively regulated by *TRI1* and a regulatory network for trichothecene mycotoxin production.** *Appl Environ Microbiol* 2003, **69**(5):2731-2736.
31. Seong KY, Pasquali M, Zhou X, Song J, Hilburn K, McCormick S, Don Y, Xu JR, Kistler HC: **Global gene regulation by *Fusarium* transcription factors *TRI6* and *TRI10* reveals adaptations for toxin biosynthesis.** *Mol Microbiol* 2009, **72**(2):354-367.
32. Stich B, Melchinger AE, Frisch M, Maurer HP, Heckenberger M, Reif JC: **Linkage disequilibrium in European elite maize germplasm investigated with SSRs.** *Theor Appl Genet* 2005, **111**(4):723-730.
33. Cumagun CJR, Rabenstein F, Miedaner T: **Genetic variation and covariation for aggressiveness deoxynivalenol production and fungal colonization among progeny of *Gibberella zeae* in wheat.** *Plant Pathol* 2004, **53**(4):446-453.
34. Meek IB, Peplow AW, Ake CJR, Phillips TD, Beremand MN: ***TRI* encodes the cytochrome P450 monooxygenase for C-8 hydroxylation during trichothecene biosynthesis in *Fusarium sporotrichioides* and resides upstream of another new *TR* gene.** *Appl Environ Microbiol* 2003, **69**(3):1607-1613.
35. Von der Ohe C, Gauthier V, Tamburic-Illincic L, Brule-Babel A, Fernando WGD, Clear R, Ward TJ, Miedaner T: **A comparison of aggressiveness and deoxynivalenol production between Canadian *Fusarium graminearum* isolates with 3-acetyl and 15-**

- acetyldeoxynivalenol chemotypes in field-grown spring wheat. *Eur J Plant Pathol* 2010, **127**:407-417.
36. Zhang H, Zhang Z, Van der Lee T, Xu J, Yang L, Yu D, Waalwijk C, Feng J: Population genetic analyses of *Fusarium asiaticum* populations from barley suggest a recent shift favoring 3ADON producers in southern China. *Phytopathology* 2010, **100**(4):328-336.
  37. Wong P, Walter M, Lee W, Mannhaupt G, Münsterkötter M, Mewes HW, Adam G, Güldener U: FGDB: Revisiting the genome annotation of the plant pathogen *Fusarium graminearum*. *Nucleic Acids Res* 2011, **39**(4):637-639.
  38. Stram DO, Lee JW: Variance components testing in longitudinal mixed effects model. *Biometrics* 1994, **50**(4):1171-1177.
  39. Melchinger AE, Utz HF, Schön CC: Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and larger bias in estimates of QTL effects. *Genetics* 1998, **149**(1):383-403.
  40. Gower JC: Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* 1966, **53**(3):325-338.
  41. Wright S: *Evolution and Genetics of Populations, Variability within and among Natural Populations*. 4 edition. Chicago: The University of Chicago Press; 1978.
  42. Weir BS: *Genetic data analysis II*. 2 edition. Sunderland: Sinauer Associates; 1996.
  43. Hill WG, Robertson A: Linkage disequilibrium in finite populations. *Theor Appl Genet* 1968, **38**(2):226-231.
  44. Maurer HP, Melchinger AE, Frisch M: Population genetic simulation and data analysis with Plabsoft. *Euphytica* 2008, **161**(1):133-139.
  45. Bernardo R: Estimation of coefficient of coancestry using molecular markers in maize. *Theor Appl Genet* 1993, **85**(8):1055-1062.
  46. Utz HF, Melchinger AE, Schön CC: Bias and sampling error of the estimated proportion of genotypic variance explained by quantitative trait loci determined from experimental data in maize using cross validation and validation with independent samples. *Genetics* 2000, **154**(4):1839-1849.

doi:10.1186/1471-2156-13-14

**Cite this article as:** Talas et al.: Association of single nucleotide polymorphic sites in candidate genes with aggressiveness and deoxynivalenol production in *Fusarium graminearum* causing wheat head blight. *BMC Genetics* 2012 **13**:14.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

