Histone H3K27 methylation perturbs transcriptional robustness and underpins dispensability of highly conserved genes in fungi

Sabina Moser Tralamazza^{1,2}, Leen Nanchira Abraham¹, Benedito Corrêa², Daniel Croll^{1*}

¹ Laboratory of Evolutionary Genetics, Institute of Biology, University of Neuchatel, Switzerland

² Department of Microbiology, Institute of Biomedical Sciences, University of Sao Paulo, Brazil

* corresponding author: Daniel Croll, daniel.croll@unine.ch.

Abstract

Epigenetic modifications are key regulators of gene expression and underpin genome integrity. Yet, how epigenetic changes affect the evolution and transcriptional robustness of genes remains largely unknown. Here, we show how the repressive histone mark H3K27me3 influences the trajectory of highly conserved genes in fungi. We first performed transcriptomic profiling on closely related species of the plant pathogen *Fusarium graminearum* species complex. We determined transcriptional responsiveness of genes across environmental conditions to determine expression robustness. To infer evolutionary conservation of coding sequences, we used a comparative genomics framework of 23 species across the *Fusarium* genus. We integrated histone methylation data from three *Fusarium* species across the phylogenetic breadth of the genus. Gene expression variation is negatively correlated with gene conservation confirming that highly conserved genes show higher expression robustness. Furthermore, we show that highly conserved genes marked by H3K27me3 deviate from the typical housekeeping gene archetype. Compared to the genomic background, H3K27me3 marked genes encode smaller proteins, exhibit lower GC content, weaker codon usage bias, higher levels of hydrophobicity and are enriched for functions related to regulation and membrane transport. The evolutionary age of conserved genes with H3K27me3 histone marks falls typically within the origins of the *Fusarium* genus. We show that highly conserved genes marked by H3K27me3 are more likely to be dispensable for survival. Lastly, we show that conserved genes exposed to repressive H3K27me3 marks across distantly related fungi predict transcriptional perturbation at the microevolutionary scale in *Fusarium* fungi. In conclusion, we establish how repressive histone marks determine the evolutionary fate of highly conserved genes across evolutionary timescales.

Introduction

Highly conserved genes are ensuring housekeeping functions essential for the survival of the organism (She et al. 2009). Beyond sequence conservation, highly conserved genes are often constitutively expressed. Changes in gene expression is a major factor associated with the evolutionary trajectory of genes. In yeasts, gene expression levels are negatively correlated with the evolutionary rate of gene sequences (Pál et al. 2001; Drummond et al. 2006). The correlation, also termed E-R correlation, was later confirmed across the tree of life including for bacteria (Rocha and Danchin 2004; Drummond et al. 2006) Metazoa (Krylov et al. 2003) and plants (Ingvarsson 2006). Both empirical and theoretical approaches support that the E-R correlation stems from the fact that highly expressed genes show stronger functional constraints. These include purifying selection acting against deleterious mutations affecting translation, protein folding and maladaptive protein interactions (Zhang and Yang 2015). Despite the broad evidence that gene expression levels are strongly associated with the rate of protein evolution across species, how the association arises through genetic and epigenetic modifications over shorter evolutionary time spans remains largely unknown.

Gene expression in eukaryotes is governed largely by higher-order chromatin structure (Woodcock and Ghosh 2010). Chromatin is determined by the nucleosome, which contains a histone octamer attached to a stretch of nuclear DNA (Luger et al. 1997). Histones and DNA are constantly being modified by various proteins catalyzing enzymatic activities, such as phosphorylation, acetylation, methylation, ubiquitination, and O-GlcNAcylation. The joint effect of these post-translational modifications regulates and stabilize gene expression during the cell cycle and development while facilitating responses to environmental stimuli (Youn 2017). For example, cytosine methylation is a covalent DNA modification associated with transcriptional responses to biotic stress (Downen et al. 2012) and plays a key role in genomic defenses against selfish genetic elements in plants (Cui et al. 2013), mammals (Smith and Meissner 2013) and fungi (Zhou et al. 2001). Histone modifications also dynamically remodel chromatin structure to activate or repress gene expression following intrinsic and extrinsic signals (Flavahan et al. 2017). Specific histone modifications such as the methylation of H3K9 residues are frequently associated with transcriptionally silent heterochromatin, while acetylation and methylation of H3K4 and H3K36 residues are hallmarks of transcriptionally active euchromatin (Gates et al. 2017). The H3K27me3 modification plays a particularly important role by underpinning facultative heterochromatin allowing for responsive gene expression regulation (Trojer and Reinberg 2007).

The H3K27 residue is methylated by the protein Polycomb repressive complex 2 (PRC2) (Qian et al. 2006). The canonical formation of the complex constitutes three core domains KMT6/EZH2, EED, SUZ12 which are conserved in multiple taxa such as vertebrates, insects, plants, and fungi (Ridenour et al. 2020). H3K27 marks are often associated with transcriptional repression in gene-rich chromosomal

regions (Jamieson et al. 2013), control of developmental stages in higher organisms such as plants and flies (Schwartz and Pirrotta 2007) and genome instability in fungi (Janevska et al. 2018; Möller et al. 2019). Due its importance for both gene regulation and genome integrity, H3K27me3 is one of the best-studied post-translational marks in fungi. H3K27me3 marks show a wide range of abundance from ~7% in the *Neurospora crassa* genome to ~30% in *Fusarium graminearum* (Freitag 2017). In the budding yeast *Saccharomyces cerevisiae* and fission yeast *Schizosaccharomyces pombe*, H3K27 methylation is absent due to the loss of the Polycomb complex (Connolly et al. 2013). Experimentally induced loss of H3K27 methylation has highly variable impacts. Inactivation of the Polycomb complex (*i.e.* PRC2) in *Fusarium* fungi causes growth and developmental defects, as well as sterility (Freitag 2017). However, the inactivation has no overt effect in *N. crassa*. Yet deletion of the methyltransferase *set-1* underlying H3K4me3 causes promiscuous spread of H3K27me3 across the genome and severe fitness costs (Jamieson et al. 2013; Kronholm et al. 2016).

Across fungi, H3K27me3 regulates the expression of rapidly evolving genes encoding virulence factors and biosynthetic pathways of specialized metabolites (Connolly et al. 2013; Jamieson et al. 2013). How facultative heterochromatin such as governed by H3K27 methylation impacts the evolution and transcriptional robustness of more conserved genes remains largely unknown. Studies focused on deep evolutionary timescales (Fair et al. 2020) and a small number of highly conserved orthologues across taxa. However, identifying causal factors underlying among species transcriptional variability is challenging because of environmental heterogeneity and niche differentiation (Dötsch et al. 2015). Furthermore, technical noise arising from different experimental setups renders comparisons difficult (De Jong et al. 2019) and uncertainty about orthology inference and functional divergence add additional uncertainty. In contrast, closely related fungi provide ideal models to analyze how repressive histone marks underpin the robustness of expression and sequence conservation across species. The ability to access transcriptional variability in identical experimental settings in species sharing colinear genomes, similar habitats and life cycles is a key advantage to make robust inferences about causal factors.

Fungi of the genus *Fusarium* are important pathogens of crops causing a wide range of diseases (Summerell 2019). Originally described as a single species (O'Donnell et al. 2004), the *Fusarium graminearum* species complex (FGSC) comprises 16 recognized species (O'Donnell et al. 2008). All species can cause Fusarium head blight in cereals and hence have strongly overlapping host ranges (van der Lee et al. 2015). FGSC are all closely related (Walkowiak et al. 2016; Tralamazza et al. 2019) but encode a vast and variable repertoire of specialized metabolites (Tralamazza et al. 2019). The genome of *F. graminearum* has one of the most strongly affected gene bodies in terms of repressive H3K27me3 marks among fungi (46% of genes) (Freitag 2017) making FGSC an ideal model to analyze epigenetic control of gene expression.

To investigate the impact of repressive histone marks on gene expression robustness and protein conservation across species, we performed genome-wide expression analyses of five FGSC members including the reference genome strain of *F. graminearum* (PH-1) (Cuomo et al. 2007). We analyzed transcriptional responsiveness across two environmental conditions including the infection of the wheat host and a nutrient-rich growth medium. To infer evolutionary conservation of coding sequences, we used a comparative genomics framework of 23 species across the genus of *Fusarium* spanning approximately 110-420 million years of evolution (Summerell et al. 2010). We integrated histone methylation data from three *Fusarium* species across the phylogenetic breadth of the genus. We found that gene expression variation is negatively correlated with gene conservation across the genus confirming that highly conserved genes show higher gene expression robustness. Genes silenced through H3K27me3 histone marks in any of the species showed lower expression levels and higher gene expression variation among species. Marked genes encoded for on average smaller proteins, showed lower GC content, codon usage bias and were enriched for functions related to regulation and membrane transport. Estimates of the evolutionary age of conserved genes showed that genes with H3K27me3 histone marks are of much more recent origin than unaffected genes. Lastly, we show that highly conserved genes marked by H3K27me3 are more likely to be dispensable.

Material and Methods

Infection Assay

RNA-seq analyses were performed on fungal mycelium grown in culture and *in planta* for five species of the FGSC (*F. graminearum*, *F. meridionale*, *F. cortaderiae*, *F. asiaticum* and *F. austroamericanum*) (Figure 1A). Each species was grown in a Petri dish containing V8 agar medium for four days at 25°C. Following the culturing, the mycelium was transferred to a 100 ml mung bean liquid medium and agitated at 170 rpm for five days at 25°C (Zhang et al. 2012). Next, the medium was filtered, and spores counted using a haemocytometer. A 10^6 per ml spore solution was prepared for the following infection procedure. For *in vitro* assays, cultures were inoculated from spore solutions in yeast sucrose agar medium for 72h at 25°C until RNA extraction. For *in planta* assays, wheat coleoptiles with intermediate susceptibility to *Fusarium* head blight (cultivar CH Combin, harvest 2018-2019) were infected with each species individually according to (Zhang et al. 2012). Briefly, wheat seeds were soaked in sterilized water for germination in a culture chamber at 25°C with a 12h white light cycle and 93% humidity. After 3 days of germination, the tip of the coleoptile was cut and 10µl of the spore solution was used for inoculation. Coleoptiles were collected 72h after inoculation (approximately 0.4 mm lesion size)

and processed for RNA extraction. All *in planta* assays were performed in triplicates and each replicate was composed of a pool of 25 infected seedlings to obtain sufficient material and homogenize infection conditions.

RNA extraction and sequencing

RNA extraction was performed using the NucleoSpin RNA Plant and Fungi kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany) according to the manufacture's recommendation. The RNA quality was assessed using a Qubit (Thermo Fisher Scientific, Waltham, USA) and an Agilent Bioanalyzer 2100 (Agilent Technologies, CA, USA). The NEB Next® Ultra™ RNA Library Prep (NEB, CA, USA) kit based on the polyA method was used for RNA library preparation. Samples were sequenced on a NovaSeq 6000 system (Illumina Inc.), CA, USA and 150 bp paired-end reads were generated. Library preparation and sequencing was performed by Novogene Co., Ltd, Beijing, China.

RNA-seq read alignment and transcript quantification

Illumina raw reads were trimmed and filtered for adapter contamination using Trimmomatic v. 0.32 (parameters: ILLUMINACLIP:TruSeq3_PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36) (Bolger et al. 2014). Filtered reads were aligned using Hisat2 v. 2.0.4 with default parameters (Kim et al. 2019) to the *Fusarium graminearum* PH-1 reference genome (King et al. 2017). Mapped transcripts were quantified using HTSeq-count (Anders et al. 2015). Read counts were normalized based on the trimmed mean of M-values (TMM) method using the calcNormFactors option. To account for gene length, we calculated reads per kilobase per million mapped reads (RPKM) values using the R package edgeR (Robinson et al. 2010). Gene expression robustness (*i.e.* gene expression variation) among the five species of the FGSC was calculated using the coefficient of variation (CV) based on the ratio of the standard deviation to the mean (Figure 1B). To compare gene expression levels among the more distant species, we normalized and ranked gene expression values using the *percent_rank* function of the *dplyr* R package (Wickham et al. 2019). To compare expression congruity (*i.e.* expression range within a group), we calculated the interquartile range distribution, which reports the difference between the 25th and 75th percentiles (Figure 1B).

Orthology and sequence conservation analyses

Gene orthology within the *Fusarium* genus was predicted using Orthofinder v2.4.0 with default parameters (Emms and Kelly 2019). The analysis included eight distinct species (for 11 genomes in total) for the FGSC and 15 additional species of the genus. Protein conservation was assessed as the mean of the pair-wise amino acid sequence identity based on the top BLASTP (Camacho et al. 2009) hit between

protein sequences of *F. graminearum* (PH-1) against the 25 other *Fusarium* genomes. We classified the underlying gene as highly conserved if an ortholog of *F. graminearum* reference genome was conserved in at least 25 of the 26 genomes and the mean of the amino acid identity was >70% among orthologs. Only single copy genes in the FGSC were retained. Conversely, orthologs failing the above criteria were classified as variable. To infer orthology and duplication events among more distantly related fungi, we analyzed the following set of genomes using Orthofinder with default parameters: *F. graminearum*, *F. oxysporum*, *Neurospora crassa*, *Magnaporthe oryzae*, *Zymoseptoria tritici* and *Saccharomyces cerevisiae*. The outline of the datasets used in the study based on orthology conservation is available in Figure 1B. To infer the age of highly conserved genes in *Fusarium*, we used the *F. oxysporum* gene set available from the Geneorigin database covering 565 species from Ensembl and Ensembl Genomes databases following the protein family-based pipeline based on the Wagner parsimony algorithm (Tong et al. 2020).

Chromatin immunoprecipitation (ChIP-seq)

F. graminearum PH-1 strain ChIP-seq and RNA-seq datasets of the histone modifications H3K27me3, H3K36me3, H3K4m3 and H3K4m2 from fungal mycelium grown in culture were retrieved from the NCBI SRA database (accession PRJNA221153) (Connolly et al. 2013). Chip-seq raw reads were trimmed with Trimmomatic v. 032 (Bolger et al. 2014) (parameters: ILLUMINACLIP:Truseq3_SE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:20) and mapped to the *F. graminearum* reference genome using Bowtie2 v. 2.4.0 (Langmead and Salzberg 2012). Alignment bam files were converted using BEDtools v.2.30.0 (Quinlan 2014) and peak calling was performed using the makeTagDirectory and findPeak (parameters: -style histone -region -size 1000 -minDist 1000 -C 0) programs included in Homer v.4.11 (Heinz et al. 2010). Peak calls from replicates were merged with BEDtools intersect. Peak coverage was calculated with BEDtools coverage. Similarly, gene coverage was analyzed with BEDtools intersect based on the *F. graminearum* reference genome annotation. For RNA-seq analyses of *F. graminearum* (Connolly, et al. 2013) and *F. oxysporum* (Fokkens et al. 2018), raw reads were analyzed as described above. Similarly, ChIP-seq datasets for different species were analyzed following the same protocol as described above using the matching reference genome (Supplementary Table 1). For *F. fujikuroi*, we used gene expression data based on a *F. fujikuroi* specific NimbleGen microarray analysis provided by the authors (Studt et al. 2016). Further information on the datasets used are available in Supplementary Table 1.

Functional enrichment analyses and codon usage

Gene ontology (GO) term enrichment analyses were performed using the Fisher's exact test based on gene counts with the *topGO* R package (Alexa and Rahnenfuhrer, 2019) and plotted using the *GOpilot* R package (Walter et al. 2015). Codon adaptation usage (CAI) between marked and unmarked genes were calculated using the CAI software of EMBOSS (<http://www.ch.embnet.org/EMBOSS/>) package. We used the codon usage table for *F. graminearum* (PH-1) available from the Kazusa DNA Research Institute (Nakamura et al. 2000). Hydrophobicity (GRAVY) and aromatic scores were calculated using the CodonW software v.1.4.4 (Peden, 2000). The GRAVY score represents the average hydrophobic index across all amino acids of a predicted protein. The aromatic score was calculated as the proportion of aromatic amino acids.

Repeat-induced point mutations and data analyses

To detect signatures of repeat-induced point mutations (RIP), we used the software The RIPper (Van Wyk et al. 2019) on the *F. graminearum* (PH-1) genome with a window size of 1000 bp and a 500 bp step size. Kendall tau- β rank correlation tests were visualized using the *corrplot* R package (Wei et al. 2017). Genome-wide expression heatmaps were generated using the *pheatmap* R package (Kolde 2017). Upset diagrams were created with the *UpSetR* R package (Conway et al. 2019). Other figures were produced using the *ggplot2* R package (Wickham 2011).

Data availability

All access information for data generated in this study or retrieved from public databases is described in Supplementary Table 1. RNA-Seq raw reads have been deposited at NCBI SRA database under the accession number PRJNA542165.

Results

Robustness of transcriptomic responses among closely related species

We performed transcriptomic profiling on closely related fungi belonging to the *Fusarium graminearum* species complex (FGSC). The group of fungi include crop pathogens infecting mainly wheat. The highly overlapping habitat and lifestyle together with the recent history of speciation make FGSC highly suitable to assess the impact of repressive histone modifications on gene conservation and expression robustness. We first assessed the degree of protein sequence conservation within the complex using orthology and find that 82.5% of genes are shared among all the members analyzed (Figure 1C). Then, we expanded the orthology analysis to 23 species covering the entire genus *Fusarium* and find

4332 highly conserved genes (orthologs of *F. graminearum* conserved in at least 25/26 *Fusarium* genomes and with > 70% mean amino acid identity) and 9813 variable genes (genes which do not meet the former criteria) across the *Fusarium* species. To assess gene expression robustness under standardized environmental conditions, we generated RNA-seq data for five FGSC members during infection of wheat and growth on nutrient-rich medium (Figure 1A). The transcriptomic data showed high reproducibility (Supplementary Figure 1). We detected reliable expression in ~78% of all genes ($n = 11,136$ out of 14,145). Most genes (90.7%) were transcribed in all five species during wheat infection and on growth medium (84.8%; Supplementary Table S2). Highly conserved genes showed higher gene expression robustness compared to variable genes (Figure 1A).

Fig.1

Figure 1. The *Fusarium graminearum* species complex (FGSC). A) Phylogenomic tree of *Fusarium* species was inferred from a coalescence-based analysis of 4192 single-copy orthologues (Tralamazza et al. 2019). Values indicate support from 100 bootstrap replicates. *Trichoderma reesei* was used as an outgroup. Colored circles indicate data type utilized in the current study. Host and culture medium describe the two environmental conditions. B) Outline of the datasets used in the study. Colored dashed arrows refer to the evolutionary fate inquiry across taxon. Boxplot refers to the expression robustness and expression congruity defined in the study. Green solid arrow indicates comparison within the same group (e.g. group a), grey dotted arrow indicate comparison between groups (e.g. a vs. b). C) Percentage of orthologues shared among FGSC members. The green circle indicates RNA-seq analysis.

Chromatin landscape of the *F. graminearum* genome and expression robustness

Following the association of gene conservation and transcriptional robustness within FGSC, we analyzed links between chromatin modification marks and transcription profiles of individual genes (Figure 2). For this, we assessed the occupancy of histone modifications in gene body regions of the reference genome *F. graminearum* (Connolly et al. 2013). As expected, euchromatin marks (i.e. H3k4me2/me3) are dominating gene body regions. Facultative heterochromatin marks (i.e. H3K27me3) are enriched in variable genes compared to highly conserved genes (Supplementary Figure 2A). Interestingly, the distribution of gene body H3K27me3 marks is strongly bimodal (Supplementary figure 2B). Hence, we categorized genes having either >50% covered by H3K27me3 marks or less. Transcriptional profiling of genes within the species complex revealed strong genome-wide associations between gene expression levels, expression robustness, protein sequence conservation and H3K27me3 gene marks (Figure 2A). As expected, we found that highly conserved unmarked genes were the most highly expressed genes. Remarkably, we found that highly conserved marked genes were less repressed (median 11.42 RPKM, with 11.6% showing no transcription) than less conserved marked genes (median of 0.83 RPKM, with 42.3% showing no transcription; Figure 2B). Furthermore, we found that genes marked by H3K27me3 showed much lower gene expression robustness (expression variation) compared to unmarked genes independent of protein sequence conservation (Figure 2B).

Fig.2

Figure 2. Genome-wide expression and expression variation profile of the *Fusarium graminearum* species complex (FGSC). A) Transcriptome analysis of the FGSC based on hierarchical clustering. Host and culture medium describe the two environmental conditions. Green and orange colors indicate the presence or absence of H3K27me3 in the gene body, as well as 1 kb upstream and downstream of genes. B) Mean expression and gene expression variation of highly conserved and variable genes during host infection in the FGSC. Fgram- *F. graminearum*, Faus – *F. austroamericanum*, Fcor – *F. cortaderiae*, Fmer- *F. meridionale*, Fasi- *F. asiaticum*.

To identify general patterns how H3K27me3 marks impact sequence conservation and gene expression robustness, we expanded the analyses to 4332 highly conserved genes of the *Fusarium* genus. Protein sequences of *F. graminearum* share a 98.8% mean amino acid identity with orthologs in the genus. We find support for a general E-R correlation (Figure 3A). Protein sequence conservation shows a significant positive correlation (p -value < 0.0001) with gene expression levels within species (*F. graminearum* on host; Kendall tau = 0.27), or between species (among FGSC members on host; Kendall tau = 0.29). We found that gene expression variation was negatively correlated (p -value < 0.0001) with protein sequence conservation (Kendall tau = -0.24 and -0.19 of FGSC members on host and in culture, respectively; Figure 3A). We found a strong positive correlation with histone H3K27me3 marks and gene expression variation (*i.e.* lower expression robustness; Kendall tau = 0.36 during infection and culture condition; p -value < 0.0001; Figure 3A).

Given the likely impact of H3K27me3 marks on expression robustness, we focused on gene expression variation among members of the FGSC. Higher gene expression is strongly associated with higher protein sequence conservation both on the host and under culture conditions (Figure 3B; Supplementary Figure 3A). Highly conserved genes (*i.e.* shared among 25/26 species and >70% identity) with H3K27me3 marks ($n = 673$) show broad transcriptional repression (Figure 3B). However, highly conserved genes show a similar range of expression (*i.e.* congruity) between different conservation rates as unmarked genes ($n = 3659$; Supplementary Figure 3B). Highly conserved genes marked by H3k27me3 are on average more divergent than unmarked genes (90.2% and 93.6% mean protein identity, respectively; Supplementary Figure 3D). Next, we analyzed gene expression robustness among members of the FGSC. Higher protein conservation was associated with reduced gene expression variation (*i.e.* higher robustness) both on the host and in culture condition (Figure 3C; Supplementary Figure 3B). We also found a positive association between protein conservation and expression congruity (*i.e.* expression range within the same category; Supplementary Figure 3C). Which supports the idea that highly conserved genes are under selection to retain transcriptional activity largely independent of the environment (*i.e.* robustness). Next, we examined the impact of H3K27me3 on robustness and found that marked genes showed significantly higher levels of gene expression variation (Figure 3C) and reduced constraints on gene expression congruity (Supplementary Figure 3B). Marked genes also show no

311 significant association with protein conservation (Kendall tau = -0.05; p -value = 0.05) in contrast to
312 unmarked genes (Kendall tau = -0.20; p -value < 2.2e-16).

Fig.3

313
314 **Figure 3.** Analyses of gene expression and gene body histone methylation marks among *Fusarium graminearum*
315 species complex (FGSC) A) Correlation plot between marks and conditions. Test was performed between paired

samples, using Kendall's *tau* method to estimate a rank-based measure of association. Circles indicate *p*-values < 0.0001. Circle size and color indicate degrees of correlation. Culture and host describe growth conditions. Fgram – *F. graminearum*. FGSC – mean value of the *F. graminearum* species complex. FGSC variation – coefficient of variation based on each species-specific RPKM values. B) Mean gene expression of highly conserved genes in the FGSC during host infection based on protein sequence conservation. C) Mean expression variation of highly conserved genes during host infection based on protein sequence conservation. Orange and green colors indicate the presence and absence of H3K27me3 in the gene body, respectively. D) Gene expression analysis of the *F. graminearum* wild type strain (wt) and mutant lacking the methyltransferase gene *kmt6* ($\Delta kmt6$) assessed under culture conditions (Connolly et al. 2013). Grey and light pink refer to the degree of gene conservation.

To investigate if the loss of expression is a direct consequence of H3K27me3 marks, we analyzed RNA-seq and ChIP-seq dataset of a *F. graminearum kmt6* mutant strain (Connolly et al. 2013). The mutant lacks the methyltransferase enzyme KMT6 responsible for H3K27me3 marks in fungi (Freitag 2017). We found that highly conserved genes marked by H3K27me3 in the wild type background had expression levels similar to unmarked H3K27me3 genes in the *kmt6* mutant background. The highly conserved genes marked by H3K27me3 in the wild type background also showed higher expression congruity in the mutant background (Figure 3D; Supplementary Figure 3C). Less conserved genes displayed similar patterns of de-repression in the *kmt6* mutant background but overall lower levels of transcription. This is consistent with highly conserved genes being under selection for gene expression robustness and the H3K27me3 marks acting to disrupt gene regulatory mechanisms in place.

Fig.4

A

B

Feature enrichment of H3K27me3-marked genes

We investigated protein functions encoded by the H3K27me3 marked gene body. We found that H3K27me3 marked genes have shorter transcripts than unmarked genes (median 1452 vs. 1841 bp; p -value $< 2.2\text{e-}16$) and have a lower GC content (51% vs. 53%; p -value $< 2.2\text{e-}16$; Figure 4A). We found that highly conserved genes covered by H3K27me3 marks have significantly less codon usage bias than unmarked genes (CAI median of 0.778 vs. 0.785; p -value = 0.0002). The encoded proteins differ also in amino acid composition with an aromatic codon score median of 0.092 and 0.074 for marked and unmarked genes, respectively. We evaluated thermodynamic properties related to protein interaction and folding. Marked genes have higher amino acid hydrophobicity compared to unmarked genes (median of -0.264 vs -0.426; Figure 4A). Conserved genes with H3K27me3 marks are enriched for functions related to oxidation-reduction, transmembrane functions, and transcriptional regulation (p -value < 0.0001 ; Figure 4B, Supplementary Figure 4A). Highly conserved unmarked genes are enriched for housekeeping functions such as cellular metabolic and macromolecular processes (Supplementary Figure 4B).

Evolutionary conservation of H3K27me3 marks among *Fusarium* fungi

To understand the evolutionary longevity of epigenetic effects on gene expression robustness, we analyzed H3K27me3 marks across the genomes of distantly related *Fusarium* species. We examined 3727 (*Fusarium* conserved) highly conserved single copy orthologs shared by *F. graminearum*, *F. oxysporum* and *F. fujikuroi*. As expected, the large majority of the highly conserved genes are not covered by histone H3K27me3 marks ($n = 3253$; Figure 5A). Genes marked by H3K27me3 majoritarily shared the mark among orthologs (59.8%; $n = 297/473$). This suggests that H3K27me3 is often retained over significant evolutionary timescales (Figure 5A). Next, we asked if H3K27me3 mark conservation in *Fusarium* is correlated with transcriptional variation. Using ranked gene expression, we compared gene transcription levels among *F. oxysporum*, *F. fujikuroi* and *F. graminearum*, as well as members of the FGSC during host infection and in culture condition. H3K27me3 marked genes in *Fusarium* show strong and consistent transcriptional repression across species and conditions (Figure 5B). Interestingly, genes with conserved H3K27me3 marks show also conserved patterns of up-regulation during infection. This is despite the fact that *F. graminearum* and *F. oxysporum* assays have been carried out on different hosts (wheat and tomato, respectively). Conversely, in culture conditions H3K27me3 marked genes of *F. oxysporum* and *F. fujikuroi* show stronger repression compared to *F. graminearum*. Overall, gene expression of H3K27me3 marked genes was consistently lower in culture conditions compared to host infection highlighting the de-repression caused by infection stress (Figure 5B).

Fig.5
A

Figure 5. Conservation of H3K27me3 gene body occupancy in highly conserved genes across different species in the *Fusarium* genus. A) Co-occurrence of H3K27me3 marks in *F. graminearum*, *F. oxysporum* and *F. fujikuroi*. B) Relative gene expression (ranked) of *Fusarium* species under different growth conditions with or without H3K27me3 marks. Host and culture medium describe the two environmental conditions. C) Age inference of the highly conserved gene set. D) Highly conserved orthologues among ascomycetes. The phylogenetic tree is based on single-copy orthologues. Orange and green colors indicate the presence and absence of H3K27me3 in the gene body, respectively.

Evolutionary origins of highly conserved genes

To investigate the evolutionary origins of highly conserved genes within FGSC ($n = 4332$), we expanded the orthology analyses to a set of representative ascomycetes. We found a striking difference in the evolutionary age of H3K27me3 marked genes compared to unmarked genes (Figure 5D). The majority of the unmarked genes are shared among ascomycetes. Conversely, H3K27me3 marked genes were largely restricted to the *Fusarium* genus. Based on gene age calibrations for *F. oxysporum*, we find that unmarked highly conserved genes have a median age of ~439 million years. H3K27me3 marked genes have a median age of ~394 million years (Figure 5C) matching the predicted age of the *Fusarium* genus (Summerell et al. 2010). Genes marked by H3K27me3 showed a much higher duplication rate compared to unmarked genes (Figure 5D). Gene duplication events can trigger genomic defense mechanisms such as repeat-induced point mutations (RIP). Hence, we analyzed evidence for RIP-like mutations in the gene sets of *F. graminearum* but found only a small portion of the genome to be affected (<1.1%). We found no difference between the extent of RIP-like mutations in H3K27me3 marked versus unmarked genes (Supplementary Figure 5).

Fig.6

Figure 6. Phenotypic effect analyses of transcription factor deletion mutants in the *F. graminearum* reference isolate PH1 background. Normal refers to mutants without any overt phenotypic change compared to the wild type. Green and orange colors indicate the presence or absence of H3K27me3 in the gene body. The presented data is a subset of the dataset by Son et al. 2011 filtered for highly conserved genes. DON – deoxynivalenol, ZEA- zearalenone.

Impact of H3K27me3 marks on gene dispensability

Repressive histone marks can have an impact on gene dispensability. We first analyzed a dataset of 657 transcription factor deletion mutant lines screened across 17 phenotypic traits using the *F. graminearum* PH1 strain as a genetic background (Son et al 2011). Out of the 657 transcription factors, 238 are considered highly conserved within FGSC based on our classification. Highly conserved H3K27me3 marked genes encoding transcription factors show lower transcription levels, lower expression congruity and lower protein sequence conservation compared to marked genes

(Supplementary Figure 6B). The screened phenotypes include sexual development, mycelial growth, secondary metabolite production, virulence, and stress response. As previously reported by Son et al. (2011), only 26% of the deletion mutants show detectable phenotypic effects. Mutants with phenotypic effects related to sexual development carried mostly (94.4%) a knock-out of an unmarked gene. Whereas altered phenotypes associated with virulence and stress response were only found in unmarked gene mutants (Figure 6). Expanding to the full data set (highly conserved + variable genes; n=657 genes), the same pattern was retained (Supplementary figure 6A). Next, we analyzed a dataset of 101 *F. oxysporum* f. *sp. lycopersici* mutants with reduced pathogenicity compared to the wild type (Michielse et al. 2008). Similarly, out of all genes with phenotypic effects only 7% (7/101) were marked with H3K27me3. Among the highly conserved genes, only one gene (1/36) was marked by H3K27me3 (Supplementary Figure 6B). Hence, our analysis shows that H3K27me3 marks are linked to gene dispensability in the *Fusarium* genus.

Fig.7

Figure 7. H3K27me3 histone methylation conservation among ascomycetes. A) Single-copy orthologues shared between *Neurospora crassa*, *Magnaporthe oryzae* and *Zyoseptoria tritici* with overlapping H3K27me3 gene body occupancy between species. B) Orthologues found in the FGSC according to *Fusarium* genus conservation

categories (highly conserved vs. variable). Subsequent green arrow and bar charts refer to a subset of the data including genes marked by H3K27me3 ($n = 111$) with shared orthologues in the *Fusarium* genus. C) Gene expression analyses for host condition, expression robustness and protein sequence conservation rate of highly conserved genes in the FGSC. H3K27me3 mark conservation was assessed among distant ascomycetes.

Linking H3K27me3 marks to protein conservation in filamentous fungi

To investigate repressive histone marks and protein conservation over deep evolutionary timespans, we analyzed single-copy orthologues ($n = 4034$) shared between the distantly related model fungi *N. crassa*, *M. oryzae* and *Z. tritici*. Only a single ortholog had conserved H3K27me3 marks among the species (Figure 7A). Overall, we found 111 genes with H3K27me3 marks in at least one of the species (Figure 7B). The majority (55.4%) of all orthologues were also highly conserved in *Fusarium* and 38% of H3K27me3 marked genes in any of the distant fungi show also H3K27me3 marks in *F. graminearum* (Figure 7B). However, only 44/111 of the H3K27me3 marked genes were highly conserved in *Fusarium*. Hence, protein conservation is reduced in *Fusarium* for genes with repressive histone marks. We also investigated links between H3K27me3 marked genes among distant fungi and transcriptional responses in different environments (host and culture condition) among members of the FGSC (Figure 7C). We selected genes, which were both conserved among distant fungi and in *Fusarium* ($n = 2213/4034$). Genes with shared H3K27me3 marks among distant fungi and *Fusarium* showed lower gene expression and high expression robustness across environments (Figure 7C, Supplementary Figure 7). Interestingly, highly conserved genes with H3K27me3 marks in distant fungi but not in *Fusarium* share similar sequence conservation levels but have lower expression robustness in FGSC (Figure 7C). Together, our findings show that highly conserved genes with repressive H3K27me3 marks in distantly related fungi underpin transcriptional perturbation at microevolutionary scale in the species complex.

Discussion

We analyzed a major group of plant pathogenic fungi to understand factors associated with gene expression robustness over deeper evolutionary timescales. We performed transcriptome-wide analyses to assess expression patterns across environments and species. Integrating histone marks data across the genus and additional ascomycetes, we established predictors for expression robustness, protein conservation and gene dispensability.

Repressive histone marks predict gene expression and dispensability

Highly conserved genes tend to be constitutively expressed to maintain basic cellular functions (She et al. 2009). We show that highly conserved genes marked by H3K27me3 deviate from these archetypical housekeeping properties. Compared to unmarked genes, marked genes exhibit shorter transcripts, lower GC content, and weaker codon usage bias, which are features associated with fast evolving genes (Lipman et al. 2002; Zhang and Yang 2015). Functions encoded by marked genes were also enriched for transcriptional regulation and binding functions. This is consistent with findings in the *Arabidopsis* genus with an enrichment for similar functions (Ha et al. 2011). The higher levels of aromatic codons and hydrophobicity of proteins encoded by marked genes suggest weakened selection against misfolding and aggregation (Bastolla et al. 2004). In contrast, highly conserved genes are typically under strong purifying selection to encode key functions in molecular and cellular processes. Similarly, highly expressed genes are under strong selection for expression robustness to reduce translational errors, protein misfolding and to avoid promiscuous protein-protein interactions (Zhang and Yang 2015; Agozzino and Dill 2018). Destabilized proteins can cause protein-protein aggregation with deleterious consequences for cellular functions including membrane integrity (Stefani and Dobson 2003). Our analyses show that highly conserved genes marked by H3K27me3 encode distinct functions compared to similarly conserved genes without heterochromatin marks. In conjunction, our analyses show that H3K27me3 marked genes are under relaxed selection pressure to avoid protein cytotoxicity.

We show that transcription factor deletion mutants affecting highly conserved genes marked by H3K27me3 are less likely to generate an overt phenotypic defect compared to unmarked but similarly conserved genes. A lack of apparent phenotypic consequences could stem from genetic redundancy. However, our analysis focused only on single copy orthologs making gene duplications an unlikely explanation (*i.e.* by creating redundancy). Alternatively, genes marked by H3K27me3 could have recently lost essentiality either through changes in the environmental interactions of the organism or mutations in the genetic background creating redundancy (Bergmiller et al. 2012). Interestingly, essential genes marked by H3K27me3 can become dispensable through epistatic interactions, monogenic suppressors or similar mechanisms (Li et al. 2019; Rousset et al. 2021). Genomic regions with repressive histone marks may coincide with higher densities of repetitive elements. In some fungi, such regions are targeted by genomic defense mechanisms including RIP (Baez-Ortega et al.) causing loss-of-function mutations (Möller et al. 2021). We found no evidence for higher rates of RIP-like mutations in marked versus unmarked genes. Hence, gene dispensability most likely evolved through other mechanisms.

H3K27me3 perturbs gene expression robustness and congruity of highly expressed genes

Highly conserved genes in the *Fusarium* genus show strong positive associations between gene expression and degree of conservation. Hence, our results confirm the E-R correlation observed broadly

among eukaryotes stipulating that purifying selection acts on translation efficiency (Zhang and Yang 2015). Beyond levels of expression, we found that expression robustness is strongly associated with the degree of protein conservation independent of the environment. The association likely reflects selection acting on regulatory elements to retain expression homeostasis under external and internal perturbations (Kitano 2004). H3K27me3 plays a significant role in gene expression robustness among closely related species. Less conserved genes covered by H3K27me3 are largely transcriptionally silent confirming well-documented patterns in fungi (Freitag 2017) and other eukaryotes (Zhang and Yang 2015). We find that highly conserved genes covered by H3K27me3 suffer mostly partial transcriptional repression and rarely complete silencing. Hence, the genes retained translation proficiency despite the repressive nature of the histone modifications. Robustness typically evolves in order to tolerate environmental perturbations and facilitates the evolvability of complex systems (Kitano 2004; Liu et al. 2020). We found that highly conserved genes with repressive histone marks showed a much higher level of gene expression variation and lower expression congruity compared to unmarked genes. Associations of histone modifications and gene expression across species was found previously in primates (Zhou et al. 2014). However, much more extensive work was done on cytosine DNA methylation. Interestingly, conservation of gene body methylation in coding sequences is positively correlated with both gene expression levels and slower protein sequence evolution, but negatively correlated with gene expression robustness (Chuang and Chiang 2014; Seymour and Gaut 2020).

We found that genes with repressive histone marks showed a weaker association between protein conservation and transcriptional robustness compared to unmarked genes. The loss in expression variation could be a consequence of polymorphism in H3K27me3 marks within the species complex. However, our phylogenetic analyses showed that gene body H3K27me3 marks of highly conserved genes are largely maintained. Interestingly, the deletion of the methyltransferase KMT6 responsible for deposition of H3K27me3 (Connolly et al. 2013) restored gene expression of marked genes to similar levels as highly conserved unmarked genes. Hence, there is an additional layer of transcriptional regulation and maintenance of translation efficiency in highly conserved genes. After depletion of H3K27me3 in the KMT6 mutant, highly conserved genes marked in the wild type showed a strong increase in gene expression congruity similar to unmarked genes. Hence, H3K27me3 seems to be the causal factor disturbing otherwise similar expression congruity of highly conserved genes. Our analyses also indicate that the sensitivity to transcriptional perturbation by H3K27me3 marks within the species complex is likely shared among distantly related ascomycetes.

Our study establishes evolutionary links between expression robustness and protein evolution mediated by repressive histone modifications. Highly conserved genes marked by H3K27me3 are transcriptionally active despite the repressive nature of the mark but suffer from perturbed expression

robustness compared to unmarked genes. Highly conserved marked genes show enrichment in environmental stress related functions, carry hallmarks of fast evolving genes and, hence, do not follow the housekeeping gene archetype. We show that H3K27me3 marks can blur the general E-R correlation of sequence conservation and expression levels. Hence, histone modifications are a key link between protein evolvability and gene essentiality.

Acknowledgements

We thank Dirk Balmer at Syngenta Inc. and the CBS-KNAW culture collection for providing fungal material. Ursula Oggenfus and Vinciane Mossion provided helpful comments on the manuscript. The research was supported by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) grant process 2019/16045-0. DC received support from the Swiss National Science Foundation (grant 31003A_173265).

References

- Alexa A, Rahnenführer J. 2009. Gene set enrichment analysis with topGO. *Bioconductor Improv*, 27, 1-26.
- Agozzino L, Dill KA. 2018. Protein evolution speed depends on its stability and abundance and on chaperone concentrations. *Proc Natl Acad Sci U S A*. 115(37):9092–9097. doi:10.1073/pnas.1810194115.
- Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 31(2):166–169. doi:10.1093/bioinformatics/btu638.
- Baez-Ortega B-O, A. A, Gori G, K. K, Strakova S, A. A, Allen A, L. J, Allum A, M. K, et al. Somatic evolution and global expansion of an ancient transmissible cancer lineage. *Science* (80-). doi:10.1126/science.aau9923.
- Bastolla U, Moya A, Viguera E, Van Ham RCHJ. 2004. Genomic determinants of protein folding thermodynamics in prokaryotic organisms. *J Mol Biol*. 343(5):1451–1466. doi:10.1016/j.jmb.2004.08.086.
- Bergmiller T, Ackermann M, Silander OK. 2012. Patterns of evolutionary conservation of essential genes correlate with their compensability. *PLoS Genet*. 8(6). doi:10.1371/journal.pgen.1002803.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30(15):2114–2120. doi:10.1093/bioinformatics/btu170.
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics*. 10(1):421. doi:10.1186/1471-2105-10-421.

Chuang TJ, Chiang TW. 2014. Impacts of pretranscriptional DNA methylation, transcriptional transcription factor, and posttranscriptional microRNA regulations on protein evolutionary rate. *Genome Biol Evol.* 6(6):1530–1541. doi:10.1093/gbe/evu124.

Connolly LR, Smith KM, Freitag M. 2013. The *Fusarium graminearum* histone H3 K27 methyltransferase KMT6 regulates development and expression of secondary metabolite gene clusters. *PLoS Genet.* 9(10). doi:10.1371/journal.pgen.1003916.

Conway JR, Lex A, Gehlenborg N. 2019. UpSetR: an R package for the visualization of intersecting sets and their properties. doi:10.1093/bioinformatics/btx364.

Cui Xiekui, Jin P, Cui Xia, Gu L, Lu Z, Xue Y, Wei L, Qi J, Song X, Luo M, et al. 2013. Control of transposon activity by a histone H3K4 demethylase in rice. *Proc Natl Acad Sci U S A.* 110(5):1953–1958. doi:10.1073/pnas.1217020110.

Cuomo CA, Güldener U, Xu J, Trail F, Turgeon BG, Pietro A Di, Walton JD, Ma L, Baker SE, Rep M, et al. 2007. The *Fusarium graminearum* Genome. *Science* (80). 317:1400–1403. doi:10.1126/science.1143708.

Dötsch A, Schniederjans M, Khaledi A, Hornischer K, Schulz S, Bielecka A, Eckweiler D, Pohl S, Häussler S. 2015. The *Pseudomonas aeruginosa* transcriptional landscape is shaped by environmental heterogeneity and genetic variation. *MBio.* 6(4). doi:10.1128/mBio.00749-15.

Dowen RH, Pelizzola M, Schmitz RJ, Lister R, Dowen JM, Nery JR, Dixon JE, Ecker JR. 2012. Widespread dynamic DNA methylation in response to biotic stress. *Proc Natl Acad Sci U S A.* 109(32):E2183–E2191. doi:10.1073/pnas.1209329109.

Drummond DA, Raval A, Wilke CO. 2006. A single determinant dominates the rate of yeast protein evolution. *Mol Biol Evol.* 23(2):327–337. doi:10.1093/molbev/msj038.

Emms DM, Kelly S. 2019. OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20(1):1–14. doi:10.1186/s13059-019-1832-y.

Fair BJ, Blake LE, Sarkar A, Pavlovic BJ, Cuevas C, Gilad Y. 2020. Gene expression variability in human and chimpanzee populations share common determinants. *Elife.* 9:1–26. doi:10.7554/eLife.59929.

Flavahan WA, Gaskell E, Bernstein BE. 2017. Epigenetic plasticity and the hallmarks of cancer. *Science* (80-). 357(6348). doi:10.1126/science.aal2380.

Fokkens L, Shahi S, Connolly LR, Stam R, Schmidt SM, Smith KM, Freitag M, Rep M. 2018. The multi-speed genome of *Fusarium oxysporum* reveals association of histone modifications with sequence divergence and footprints of past horizontal chromosome transfer events. *bioRxiv.*:465070. doi:10.1101/465070.

Freitag M. 2017. Histone Methylation by SET Domain Proteins in Fungi. *Annu Rev Microbiol.*

71(1):413–439. doi:10.1146/annurev-micro-102215-095757.

Gates LA, Foulds CE, Malley BWO. 2017. Histone Marks in the ‘ Driver ’ s Seat ’ □ : Functional Roles in Steering the Transcription Cycle. Trends Biochem Sci. 42(12):977–989. doi:10.1016/j.tibs.2017.10.004.

Ha M, Ng DWK, Li WH, Chen ZJ. 2011. Coordinated histone modifications are associated with gene expression variation within and between species. Genome Res. 21(4):590–598. doi:10.1101/gr.116467.110.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Mol Cell. 38(4):576–589. doi:10.1016/j.molcel.2010.05.004.

Ingvarsson PK. 2006. Gene Expression and Protein Length Influence Codon Usage and Rates of Sequence Evolution in *Populus tremula*. Mol Biol Evol. 24(3):836–844. doi:10.1093/molbev/msl212.

Jamieson K, Rountree MR, Lewis ZA, Stajich JE, Selker EU. 2013. Regional control of histone H3 lysine 27 methylation in *Neurospora*. Proc Natl Acad Sci U S A. 110(15):6027–6032. doi:10.1073/pnas.1303750110.

Janevska S, Güldener U, Sulyok M, Tudzynski B, Studt L. 2018. Set1 and Kdm5 are antagonists for H3K4 methylation and regulators of the major conidiation-specific transcription factor gene *ABAI* in *Fusarium fujikuroi*. Environ Microbiol. 20(9):3343–3362. doi:10.1111/1462-2920.14339.

De Jong T V., Moshkin YM, Guryev V. 2019. Gene expression variability: The other dimension in transcriptome analysis. Physiol Genomics. 51(5):145–158. doi:10.1152/physiolgenomics.00128.2018.

Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 37(8):907–915. doi:10.1038/s41587-019-0201-4.

King R, Urban M, Hammond-Kosack KE. 2017. Annotation of *Fusarium graminearum* (PH-1) version 5.0. Genome Announc. 5(2). doi:10.1128/genomeA.01479-16.

Kitano H. 2004. Biological robustness. Nat Rev Genet. 5(11):826–837. doi:10.1038/nrg1471.

Kolde R, Kolde MR. 2015. Package ‘pheatmap’. R package, 1(7), 790.

Kronholm I, Johannesson H, Ketola T. 2016. Epigenetic control of phenotypic plasticity in the filamentous fungus *Neurospora crassa*. G3 Genes, Genomes, Genet. 6(12):4009–4022. doi:10.1534/g3.116.033860.

Krylov DM, Wolf YI, Rogozin IB, Koonin E V. 2003. Gene loss, protein sequence divergence, gene dispensability, expression level, and interactivity are correlated in eukaryotic evolution. Genome Res. 13(10):2229–2235. doi:10.1101/gr.1589103.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods. 9(4):357–359.

doi:10.1038/nmeth.1923.

van der Lee T, Zhang H, van Diepeningen A, Waalwijk C. 2015. Biogeography of *Fusarium* *graminearum* species complex and chemotypes: a review. Food Addit Contam Part A. 32(4):453–460. doi:10.1080/19440049.2014.984244.

Li J, Wang HT, Wang WT, Zhang XR, Suo F, Ren JY, Bi Y, Xue YX, Hu W, Dong MQ, et al. 2019. Systematic analysis reveals the prevalence and principles of bypassable gene essentiality. Nat Commun. 10(1):1–15. doi:10.1038/s41467-019-08928-1.

Lipman DJ, Souvorov A, Koonin E V., Panchenko AR, Tatusova TA. 2002. The relationship of protein conservation and sequence length. BMC Evol Biol. 2(1):1–10. doi:10.1186/1471-2148-2-20.

Liu X, Maiorino E, Halu A, Glass K, Prasad RB, Loscalzo J, Gao J, Sharma A. 2020. Robustness and lethality in multilayer biological molecular networks. Nat Commun. 11(1):1–12. doi:10.1038/s41467-020-19841-3.

Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature. 389(6648):251–260. doi:10.1038/38444.

Michielse CB, van Wijk R, Reijnen L, Cornelissen BJC, Rep M. 2008. Insight into the molecular requirements for pathogenicity of *Fusarium oxysporum* f. sp. *lycopersici* through large-scale insertional mutagenesis. Genome Biol. 10(1):1–18. doi:10.1186/gb-2009-10-1-r4.

Möller M, Habig M, Lorrain C, Feurtey A, Haueisen J, Fagundes WC, Alizadeh A, Freitag M, Stukenbrock EH. 2021. Recent loss of the Dim2 DNA methyltransferase decreases mutation rate in repeats and changes evolutionary trajectory in a fungal pathogen. PLOS Genet. 17(3):e1009448. doi:10.1371/journal.pgen.1009448.

Möller M, Schotanus K, Soyer JL, Haueisen J, Happ K, Stralucke M, Happel P, Smith KM, Connolly LR, Freitag M, et al. 2019. Destabilization of chromosome structure by histone H3 lysine 27 methylation. PLoS Genet. 15(4):2–7. doi:10.1371/journal.pgen.1008093.

Nakamura Y, Gojobori T, Ikemura T. 2000. Codon usage tabulated from international DNA sequence databases: Status for the year 2000. Nucleic Acids Res. 28(1):292. doi:10.1093/nar/28.1.292.

O'Donnell K, Ward TJ, Aberra D, Kistler HC, Aoki T, Orwig N, Kimura M, Bjørnstad Å, Klemsdal SS, Bjørnstad S, et al. 2008. Multilocus genotyping and molecular phylogenetics resolve a novel head blight pathogen within the *Fusarium graminearum* species complex from Ethiopia. Fungal Genet Biol. 45(11):1514–22. doi:10.1016/j.fgb.2008.09.002.

O'Donnell K, Ward TJ, Geiser DM, Kistler HC, Aoki T. 2004. Genealogical concordance between the mating type locus and seven other nuclear genes supports formal recognition of nine phylogenetically distinct species within the *Fusarium graminearum* clade. Fungal Genet Biol. 41(6):600–623. doi:10.1016/j.fgb.2004.03.003.

Pál C, Papp B, Hurst LD. 2001. Highly expressed genes in yeast evolve slowly. Genetics. 158(2):927–

931. doi:10.1093/genetics/158.2.927.

Peden JF. 2000. Analysis of codon usage.73-74.

Qian C, Wang X, Manzur K, Sachchidanand, Farooq A, Zeng L, Wang R, Zhou MM. 2006. Structural Insights of the Specificity and Catalysis of a Viral Histone H3 Lysine 27 Methyltransferase. J Mol Biol. 359(1):86–96. doi:10.1016/j.jmb.2006.03.006.

Quinlan AR. 2014. BEDTools: The Swiss Army Tool for Genome Feature Analysis. Curr Protoc Bioinforma. 47(1):11.12.1-11.12.34. doi:10.1002/0471250953.bi1112s47.

Ridenour JB, Möller M, Freitag M. 2020. Polycomb Repression without Bristles: Facultative Heterochromatin and Genome Stability in Fungi. Genes (Basel). 11(6):638. doi:10.3390/genes11060638.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 26(1):139–140. doi:10.1093/bioinformatics/btp616.

Rocha EPC, Danchin A. 2004. An Analysis of Determinants of Amino Acids Substitution Rates in Bacterial Proteins. Mol Biol Evol. 21(1):108–116. doi:10.1093/molbev/msh004.

Rousset F, Cabezas-Caballero J, Piastra-Facon F, Fernández-Rodríguez J, Clermont O, Denamur E, Rocha EPC, Bikard D. 2021. The impact of genetic diversity on gene essentiality within the *Escherichia coli* species. Nat Microbiol. 6(3):301–312. doi:10.1038/s41564-020-00839-y.

Schwartz YB, Pirrotta V. 2007. Polycomb silencing mechanisms and the management of genomic programmes. Nat Rev Genet. 8(1):9–22. doi:10.1038/nrg1981.

Seymour DK, Gaut BS. 2020. Phylogenetic Shifts in Gene Body Methylation Correlate with Gene Expression and Reflect Trait Conservation. Mol Biol Evol. 37(1):31–43. doi:10.1093/molbev/msz195.

She X, Rohl CA, Castle JC, Kulkarni A V., Johnson JM, Chen R. 2009. Definition, conservation and epigenetics of housekeeping and tissue-enriched genes. BMC Genomics. 10:1–12. doi:10.1186/1471-2164-10-269.

Smith ZD, Meissner A. 2013. DNA methylation: Roles in mammalian development. Nat Rev Genet. 14(3):204–220. doi:10.1038/nrg3354.

Son H, Seo YS, Min K, Park AR, Lee J, Jin JM, Lin Y, Cao P, Hong SY, Kim EK, et al. 2011. A phenome-based functional analysis of transcription factors in the cereal head blight fungus, *Fusarium graminearum*. PLoS Pathog. 7(10). doi:10.1371/journal.ppat.1002310.

Stefani M, Dobson CM. 2003. Protein aggregation and aggregate toxicity: New insights into protein folding, misfolding diseases and biological evolution. J Mol Med. 81(11):678–699. doi:10.1007/s00109-003-0464-5.

Studt L, Rösler SM, Burkhardt I, Arndt B, Freitag M, Humpf HU, Dickschat JS, Tudzynski B. 2016.

Knock-down of the methyltransferase Kmt6 relieves H3K27me3 and results in induction of cryptic and otherwise silent secondary metabolite gene clusters in *Fusarium fujikuroi*. Environ Microbiol. 18(11):4037–4054. doi:10.1111/1462-2920.13427.

Summerell BA. 2019. Resolving *Fusarium*: Current Status of the Genus. Annu Rev Phytopathol. 57:323–339. doi:10.1146/annurev-phyto-082718-100204.

Summerell BA, Laurence MH, Liew ECY, Leslie JF. 2010. Biogeography and phylogeography of *Fusarium*: a review. Fungal Divers. 44(1):3–13. doi:10.1007/s13225-010-0060-2.

Tong Y-B, Shi M-W, Qian SH, Chen Y-J, Luo Z-H, Tu Y-X, Chen C, Chen Z-X. GenOrigin: A Comprehensive Protein-coding Gene Origination Database on the Evolutionary Timescale of Life.

Tralamazza SM, Rocha LO, Oggenfuss U, Corrêa B, Croll D, Rose L. 2019. Complex Evolutionary Origins of Specialized Metabolite Gene Cluster Diversity among the Plant Pathogenic Fungi of the *Fusarium graminearum* Species Complex. Genome Biol Evol. 11(11):3106–3122. doi:10.1093/gbe/evz225.

Trojer P, Reinberg D. 2007. Facultative Heterochromatin: Is There a Distinctive Molecular Signature? Mol Cell. 28(1):1–13. doi:10.1016/j.molcel.2007.09.011.

Walkowiak S, Rowland O, Rodrigue N, Subramaniam R. 2016. Whole genome sequencing and comparative genomics of closely related Fusarium Head Blight fungi: *Fusarium graminearum*, *F. meridionale* and *F. asiaticum*. BMC Genomics. 17(1):1014. doi:10.1186/s12864-016-3371-1.

Walter W, Sánchez-Cabo F, Ricote M. 2015. GOpilot: an R package for visually combining expression data with functional analysis: Fig. 1. Bioinformatics. 31(17):2912–2914. doi:10.1093/bioinformatics/btv300.

Wei T, Simko V, Levy M, Xie Y, Jin Y, Zemla J. 2017. Package ‘corrplot’. Statistician, 56(316), e24.

Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Golemund G, Hayes A, Henry L, Hester J, et al. 2019. Welcome to the Tidyverse. J Open Source Softw. 4(43):1686.

Wickham H. 2011. ggplot2. Wiley Interdiscip Rev Comput Stat. 3(2):180–185. doi:10.1002/wics.147.

Woodcock CL, Ghosh RP. 2010. Chromatin higher-order structure and dynamics. Cold Spring Harb Perspect Biol. 2(5):a000596. doi:10.1101/cshperspect.a000596.

Van Wyk S, Harrison CH, Wingfield BD, De Vos L, Van Der Merwe NA, Steenkamp ET. 2019. The RIPper, a web-based tool for genome-wide quantification of Repeat-Induced Point (RIP) mutations. PeerJ. 2019(8):e7447. doi:10.7717/peerj.7447.

Youn HD. 2017. Methylation and demethylation of DNA and histones in chromatin: The most complicated epigenetic marker. Exp Mol Med. 49(4):e321–e321. doi:10.1038/emmm.2017.38.

Zhang J, Yang JR. 2015. Determinants of the rate of protein sequence evolution. Nat Rev Genet. 16(7):409–420. doi:10.1038/nrg3950.

734 Zhang X-W, Jia L-J, Zhang Y, Jiang G, Li X, Zhang D, Tang W-H. 2012. In Planta Stage-Specific Fungal
735 Gene Profiling Elucidates the Molecular Strategies of *Fusarium graminearum* Growing inside Wheat
736 Coleoptiles. Plant Cell. 24(12):5159–5176. doi:10.1105/tpc.112.105957.

737 Zhou X, Cain CE, Myrthil M, Lewellen N, Michelini K, Davenport ER, Stephens M, Pritchard JK, Gilad
738 Y. 2014. Epigenetic modifications are associated with inter-species gene expression variation in
739 primates. Genome Biol. 15(12):1–19. doi:10.1186/s13059-014-0547-3.

740 Zhou Y, Cambareri E, Kinsey J. 2001. DNA methylation inhibits expression and transposition of the
741 *Neurospora* Tad retrotransposon. Mol Genet Genomics. 265(4):748–754.
742 doi:10.1007/s004380100472.

743

Fig.1

Fig.2

A bioRxiv preprint doi: <https://doi.org/10.1101/2021.05.21.445187>; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

B

Fig.3

bioRxiv preprint doi: <https://doi.org/10.1101/2021.05.21.445187>; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A

B

C

D

Fig.4

bioRxiv preprint doi: <https://doi.org/10.1101/2021.05.21.445187>; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A

B

Fig.5

A

bioRxiv preprint doi: <https://doi.org/10.1101/2021.05.21.445187>; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

B

C

D

Fig.6

bioRxiv preprint doi: <https://doi.org/10.1101/2021.05.21.445187>; this version posted May 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig.7