Nucleosome patterns in four plant pathogenic fungi with contrasted genome structures

Colin Clairet[1], Nicolas Lapalu[1], Adeline Simon[1], Jessica L. Soyer[1,2], Muriel Viaud[1], Enric Zehraoui[3], Berengère Dalmais[1], Isabelle Fudal[1,*] and Nadia Ponts[3,*]

[1]UMR1290 BIOGER, Université Paris-Saclay, INRAE, AgroParisTech, Avenue Lucien Brétignières, BP 01, F-78850 Thiverval-Grignon, France

[2]Max Planck Institute for Evolutionary Biology, August-Thienemann-Str. 2, 24306 Plön, and Christian-Albrechts University of Kiel, Am Botanischen Garten 1-9, 24118 Kiel, Germany

[3]INRAE, MycSA, F-33882 Villenave d'Ornon, 33140, France

* To whom correspondence should be addressed.

Nadia Ponts Tel: +33557122482; Email: nadia.ponts@inrae.fr;

Isabelle Fudal Tel: +33130814590; Email: isabelle.fudal@inrae.fr;

ABSTRACT

Fungal pathogens represent a serious threat towards agriculture, health, and environment. Control of fungal diseases on crops necessitates a global understanding of fungal pathogenicity determinants and their expression during infection. Genomes of phytopathogenic fungi are often compartmentalized: the core genome contains housekeeping genes whereas the fast-evolving genome mainly contains transposable elements and species-specific genes. In this study, we analysed nucleosome landscapes of four phytopathogenic fungi with contrasted lifestyles and genome organisations to describe and compare nucleosome repartition patterns in relation with genome structure and gene expression level. We combined MNase-seq and RNA-seq analyses to concomitantly map nucleosome-rich and transcriptionally active regions during fungal growth in axenic culture; we developed the tool MSTS to analyse and visualise data obtained from MNase-seq experiments in combination with other genomic data and notably RNA-seq expression data. We observed different characteristics of nucleosome profiles between species, as well as between genomic regions within the same species. We further linked nucleosome repartition and gene expression. Our findings support that nucleosome positioning and occupancies are subjected to evolution, in relation with underlying genome sequence modifications. Understanding genomic organisation and its role in expression regulation is the next gear to understand complex cellular mechanisms and their evolution.

INTRODUCTION

Fungi account for a huge part of the Earth biodiversity; they rank second in terms of species number, behind the insects, with a current estimate of 2.2 to 3.8 million species (1). Fungi are organisms of major

environmental importance as they develop beneficial symbiotic associations with plants and are able to decay dead organic matter (2). Unfortunately, fungi are also very efficient pathogens causing important damages in agriculture, human health, and environment (3). Control of fungal diseases on crops necessitates a global understanding of fungal pathogenicity determinants and the control of their expression during infection. Among these pathogenicity determinants, fungi secrete an arsenal of molecules known as effectors, key elements of pathogenesis which modulate innate immunity of the plant and facilitate infection, corresponding to small proteins, secondary metabolites and small RNAs (4–6). Upon plant infection, fungi undergo a tightly controlled transcriptional reprogramming, and different sets of effectors are expressed at specific stages of pathogen development and host colonization. Plant-associated fungi generally show contrasted genomic landscapes including 'plastic' regions with a high prevalence of transposable elements (TE). These genomes either show an overall large proportion of TE evenly distributed throughout the genome, or TE clustered in specific regions such as long TE-rich blocks, accessory chromosomes or subtelomeric areas (7). Effector genes are over-represented in these TE-rich regions. TE-rich compartments have heterochromatin properties contrary to TE-poor regions which have euchromatin properties. The location of effector genes in regions enriched in TEs has been shown to provide a tight control of their expression through chromatin remodelling. Indeed, several recent studies pointed out the potential role of chromatin remodelling in the regulation of effector-encoding genes and the control of secondary metabolism (reviewed in (8, 9)).

Eukaryotic chromatin is packaged into nucleosomes, each composed of DNA wrapped around a histone octamer associated with various other proteins, and separated by linker DNA (10). These histone proteins are composed of histone core where the DNA is wrapped and histone tails which can be chemically modified by specific enzymes changing the chromatin 3D-structure and DNA accessibility to polymerases and transcription factors (TF). Positioning of nucleosomes throughout the genome and post-translational modifications of histones have a significant regulatory function by modifying availability of binding sites to TF and to polymerases, affecting DNA-dependent processes such as transcription, DNA repair, replication and recombination (11, 12). Nucleosome positioning and occupancy are determined by a combination of DNA sequence features, TF, chromatin remodelers and histones modifiers. Genome-wide maps of nucleosome occupancy and positioning are still sparse in fungi and have only been developed in a few Hemiascomycota yeast species, including *Saccharomyces cerevisiae* (13, 14), in the ascomycete *Aspergillus fumigatus* (15) and the basidiomycete *Mixia osmundae* (16). The studies revealed that promoter, enhancer and terminator regions were depleted in nucleosomes, allowing access to TF, and that the nucleosomal DNA length distribution was similar in *M. osmundae* and *A. fumigatus* but differed from that of hemiascomycetous yeasts. In contrast, no comparative genome-wide analyses of nucleosome positioning have been performed in ascomycetes and notably not in plant pathogenic fungi.

In the present study, we investigated genome-wide nucleosome localization in four different plant pathogenic fungi showing different host ranges, lifestyles and genomic organizations: i) *Leptosphaeria maculans* 'brassicae' (Lmb), a hemibiotrophic pathogen of *Brassica* species, including oilseed rape; ii) the most closely related species of Lmb, *Leptosphaeria maculans* 'lepidii' (Lml), a pathogen of *Lepidium*

spp.; iii) *Fusarium graminearum*, a hemibiotrophic pathogen of cereals and iv) *Botrytis cinerea*, a polyphagous necrotrophic pathogen causing grey mould on more than 1,400 plant species. The genome of Lmb has been invaded by TE (which represent more than 30 % of its genome) and is composed of alternating compartments: gene-rich GC-equilibrated and TE-rich AT-rich genomic regions (17, 18). In contrast, the Lml genome presents only 4 % of repeats which are evenly distributed throughout the genome (19). Genomes of *F. graminearum* and *B. cinerea* have a very low to low TE-content (20–23). The genome of the reference strain of *B. cinerea*, B05.10, contains 4 % of TE, which are localized essentially in the telomeric and centromeric regions of the core chromosomes, or on the two dispensable chromosomes (20, 24). The genome of *F. graminearum* contains very little TE identified to date (0.3 % (21–23)).

In this study, we compare nucleosome repartition patterns in relation with genome structure and gene expression level in these four phytopathogenic Ascomycota. To gain insight into the role of nucleosome positioning and occupancy in regulating fungal pathogen transcription, we applied micrococcal nuclease digestion of mono-nucleosomes (MNase-assisted isolation of nucleosomes, MAINE-seq or MNase-seq) with regards to mRNA abundance to concomitantly map nucleosome-rich regions and transcriptionally active regions during fungal growth in contrasted media.

MATERIAL AND METHODS

Strains and culture conditions

Leptosphaeria maculans 'brassicae' v23.1.3 and *Leptosphaeria maculans* 'lepidii' IBCN84 mycelia were inoculated into 100 mL of Fries liquid medium (1 g/L NH_4NO_3, 5 g/L $C_4H_{12}N_2O_6$, 1 g/L KH_2PO_4, 0.5 mg/L $MgSO_4$ $7H_2O$, 130 mg/L $CaCl_2$, 100 mg/L NaCl, 30 g/L $C_{12}H_{22}O_{11}$ and 5 g/L Yeast extract). Tissues were harvested after growing for seven days in the dark at 25°C. *Botrytis cinerea* strain B05.10 (10^6 spores/mL) was grown for two days on solid Malt Medium (MM, 20 g/L malt extract, 5 g/L yeast extract and 15 g/L agar) covered with a cellophane layer (25, 26). The plates were incubated in a growth chamber (Sanyo MLR-350H) at 23°C with an alternation of 14 h of white light and 10 h of darkness. After two days of culture, mycelia were ground in liquid nitrogen and stored to -80°C until further processing. *Fusarium graminearum* strain CBS185.32 (Centraal Bureau voor Schimmelcultures, Utrecht, the Netherlands) was grown for three days in modified liquid MS (glucose was substituted with sucrose) as previously described (27). All cultures were done in three biological replicates.

Preparation of nucleosomal DNA

Fungal material was harvested and treated with microccocal nuclease (MNase, cat. #MS0247S, New England BioLabs). For Lmb and Lml, ~300 mg of mycelium were digested with 5 µL of MNase for 10 min at 37°C (28), directly followed by DNA purification as previously described (29). For *F. graminearum*, mycelia were harvested by filtering and immediately homogenized for 1 min at 30 Hz using a TissueLyzer (Qiagen). Then, 100 mg of ground mycelium was digested for 10 min at 37°C with

15 µL of MNase in 600 µL of digestion buffer (0.6% v/v IGEPAL, 50 mM NaCl, 2 mM Pefabloc, 50 mM Tris-HCl pH8, 10 mM $CaCl_2$). The reactions were stopped with 10mM EDTA and the samples treated with RNAse followed by proteinase K prior DNA purification with phenol/chloroform and ethanol precipitation. For *B. cinerea*, we digested 100-200 mg of mycelium per sample with 1 µL MNase at 37°C (28) for 10 min. The reactions were stopped by adding 10mM EDTA and samples treated with RNAse A followed by proteinase K. DNA purification was realized with the "Nucleospin Gel and PCR clean up kit" (Macherey Nagel, cat #740609.250). For all samples, nucleic acid quantification was performed by UV spectrometry using a Nanodrop-ND 1000 apparatus, and digestion profiles were checked by 2% agarose gel electrophoresis. Nucleosomal DNA was stored at -20°C until DNA library preparation.

Extraction of total RNA

For Lmb and Lml, total RNA was extracted from mycelium grown for one week in Fries liquid medium as previously described (30). For *F. graminearum*, mycelia were harvested by filtering, rinsed twice with sterile deionized water, and flash frozen in liquid nitrogen. One milliliter of TRIzol™ Reagent (Thermo Fischer Scientific) was added to 200 mg of mycelium before grinding for 1.5 min at 30 Hz using a TissueLyzer (Qiagen). Total RNA was then extracted using a previously published protocol (31). For *B. cinerea*, total RNA was extracted from frozen ground mycelium using a previously published protocol (25). All total RNA samples were stored at -80°C until preparation of RNA library.

Preparation of sequencing libraries, high-throughput sequencing, and read pre-processing

MNase-seq libraries were prepared from purified nucleosomal DNA using the kit NEBNext Ultra DNA Library Prep Kit for Illumina (cat. # E7370L New England BioLabs) following the manufacturer's instructions. The NEBNext Ultra Directional RNA Library Prep Kit for Illumina (cat. # E7420L New England BioLabs) was used to prepare all RNA-seq libraries, following the manufacturer's instructions. Sequencing was performed by the GenomEast platform, a member of the 'France Génomique' consortium (ANR-10-INBS-0009). Samples were run in 9-plex on an Illumina HiSeq 4000 in paired mode, 2x50 bp. Initial read quality check was performed using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Raw reads were then pre-processed with Trimmomatic v0.32 (32) to clip out any remaining sequencing adapter sequence and crop low quality nucleotides (minimum accepted Phred score of 30). Reads in pairs of 40 bp or more in length were used in the present analysis.

Transcriptome analyses

RNA-seq reads were mapped against their respective reference genomes (see Table 1) using STAR v2.5.1 (33). TPM counts (Transcripts Per Million reads, (34)) were computed using the count TPM tool provided with the MNase-Seq Tool Suite (MSTS; Supplementary Figure 1) that was developed in-house to analyse genome-wide nucleosome positioning data combined with RNA-seq data (https://github.com/nlapalu/MSTS).

MNase-seq analyses

MNase-seq paired-end reads were mapped using Bowtie2 software ran in very-sensitive mode (35). MSTS (MNase-seq Tool Suite) was used to compute all phasograms, dinucleotide composition, as well as nucleosome density profiles of genomic compartments and/or gene lists (Supplementary Figure 1). Lists of near-universal single copy orthologs were obtained by running BUSCO3 for Fungi (36, 37) on each reference genome studied. Graphical visualisations were computed with MSTS and Matlab R2020b (MathWorks).

RESULTS AND DISCUSSION

Establishing nucleosome landscapes of the Pezizomycotina *L. maculans* 'brassicae', *L. maculans* 'lepidii', *B. cinerea*, and *F. graminearum*

We investigated the nucleosome landscapes of four fungal species of the Ascomycota subdivision Pezizomycotina (*L. maculans* 'brassicae', *L. maculans* 'lepidii', *B. cinerea*, and *F. graminearum*) by MAINE coupled to deep sequencing (see Supplementary Table 1 for descriptive sequencing metrics). Each experiment was performed using three biological replicates that were sequenced independently at more than 70-fold coverage depths by 147 bp-long nucleosome footprints (defined as the core coverage depths of sequence sufficient for in-depth characterization of nucleosome positioning (38)). In order to explore and visualise NGS data obtained from MNase-seq experiments, we developed a collection of utility tools, called MSTS for "MNase-Seq Tool Suite", assembled in a workflow aiming at profiling nucleosome landscapes in relation to genomic features as well as gene expression (Supplementary Figure 1). MSTS workflow was applied to our datasets, beginning with the exploration of nucleosome distribution at the genome scale.

Genome-wide nucleosome spacing

We first explored nucleosome landscapes in the four fungal genomes by measuring the average distance between nucleosomes genome-wide; we computed phasograms, *i.e.,* frequency distributions of read coverage per base genome-wide for all four species (Figure 1 and Supplementary Table 2). Phasograms obtained in nucleosome mapping resemble oscillating sine wave signals, for which period is the length of DNA bound to the histone octamer plus the length of the DNA stretch to the next nucleosome, averaged genome wide (38). Phasing signals were observed genome-wide over 1,200 bp sliding windows revealing six to seven nucleosome peaks in a wave signal decaying in intensity with increasing distance and significant linear regression on peak apex positions. We found that, in the fungi studied, nucleosomes are 161 to 172 bp distant from each other (centre to centre), also called nucleosome repeat length or NRL (*i.e.,* the length of DNA wrapped around the histone octamer plus linker DNA), depending on the considered species and culture condition. In Lmb and Lml (Figures 1A and 1B), average NRL is 166 bp and 161 bp, respectively. In *B. cinerea*, average NRL is estimated at 169 bp (Figure 1C). In *F. graminearum*, this distance reaches 172 bp (Figure 1D). Considering these values and the canonical length of nucleosomal DNA (147 bp), linker DNA length can be estimated to stand, in average, between 14 to 19 bp for respectively Lml and Lmb, 22 bp for *B. cinerea*, and 25 bp

for *F. graminearum*. Nucleosome phasing genome-wide seems to be particularly tight in *F. graminearum*, with very little deviation in the measured phases (Figure 1D and Supplementary Table 2). In contrast, higher deviations are observed for *B. cinerea* (Figure 1C and Supplementary Table 2).

Nucleosome spacing influences the formation of the higher order chromatin fibre, often referred to as the 30-nm chromatin fibre (39). Several structural models of the chromatin fibre have been proposed, all underlining nucleosome-nucleosome interactions including the length of linker DNA fragments as major driving factors (40). Notably, the chromatin fibre was found to be narrower (21-nm in diameter) for a short NRL of 167 bp (41). Similarly, increasing NRLs lead to increasingly wider fibres, reaching a highly compact 30-nm solenoid structure for an NRL of 197 bp. Typically, NRLs are ~175-200 bp in plants and other higher eukaryotes (16, 38, 42–44), and ~165 bp and 154 bp in the yeasts *S. cerevisiae* and *Schizosaccharomyces pombe*, respectively (13, 45). In the present study, we found NRL values remarkably constant between biological replicates, a phenomenon sometimes referred to as clamping that involves ATP-dependent chromatin remodelers in purified experimental systems (46). The species *F. graminearum* and *B. cinerea* show similar NRLs in the middle range (169 bp to 172 bp), possibly indicating intermediate levels of compaction of the higher order chromatin fibre. These values are similar to those obtained for *A. fumigatus*, *i.e.,* linker length ranging from 21 to 27 bp, using an MNase treatment similar to the one used in the present study (15). Lmb and Lml distinguish themselves with shorter NRLs of only 161-166 bp, suggesting a narrower chromatin fibre structure. Furthermore, Lmb presents longer NRLs than Lml. We hypothesized this peculiarity may be explained by large AT-rich regions displayed by the Lmb genome, not encountered in the Lml genome (17, 19).

Nucleosome distribution profiles

Read density was plotted genome-wide in one kb-long sliding non-overlapping windows along chromosomes for all four fungi (Figures 2 to 4, and Supplementary Figures 2 to 6). The density profiles obtained for *B. cinerea* show remarkable regularity of nucleosome density genome-wide (Figure 2A and Supplementary Figure 2). Nevertheless, we could observe that almost all occasional thin peaks of density were correlated with the positions of BOTY retro-transposons (24). Out of 48 complete copies of BOTY in the genome, 31 show a peak of nucleosome density. Notably, they correspond to the BOTY elements with an equilibrated percentage of GC (43-45%) while the 17 copies that do not show such a peak are those with a lower percentage of GC (14-24%) probably because they have undergone Repeat-Induced Point mutation, or RIP (24). Peaks of density were rarely observed for TE other than BOTY (Supplementary Figure 3). We also investigated nucleosome spacing in regions occupied by BOTY and non-BOTY TE. Phasograms were plotted as described above restricting our analysis to BOTY or other TE (Figures 2B and 2C, respectively). Much larger phases can be observed in other TE regions (178.4-187.5 bp) when compared to BOTY regions (171.3-172.5 bp) or genome-wide (168.2-169.4 bp, Figure 1C), indicating larger nucleosome spacing in TE-occupied regions. BOTY-containing regions, which positions correlate with discrete peaks of nucleosome density, exhibit slightly larger phases than genome-wide. Thus, the observed peaks of read density may be the result of increased nucleosome occupancy, *i.e.,* a measure of the stability of a nucleosome at a given position in a multiple cell sample, rather than a denser deposition of nucleosomes. BOTY is one of the largest TE identified

in the B05.10 strain (6.4-6.6 kb), and that's definitively the TE with the largest genome coverage *i.e.,* 0.96% (24). Notably, the majority of *B. cinerea* small interfering RNA (siRNA) predicted to silence host plant genes are derived from BOTY elements (5). As the production of these siRNA effectors is activated during the early phase of plant infection, it would be interesting to investigate whether variation of nucleosome occupancy is involved in their regulation.

In *F. graminearum*, regions equally packed with nucleosomes are interspaced with area with lower density (Figure 3A and Supplementary Figure 4). Strikingly, this profile mirrors the previously described SNP density profiles in *F. graminearum* (47). We investigated whether or not this profile was the result of increased spacing between nucleosomes in regions found denser in SNPs. Phasograms were plotted restricting our analysis to the SNP-enriched polymorphic islands or the rest of the genome, as defined by Laurent et al. (2017) (47). Wave signals similar to the ones observed genome-wide were obtained, with phases of 172.3-172.4 bp in polymorphic islands (Figure 3B, Supplementary Table 3) and 171.6-171.9 bp outside these regions (Figure 3C, Supplementary Table 3). These results indicate that nucleosomes appear well-arrayed genome-wide, with phases only slightly longer within polymorphic islands. Thus, the observed read density profile may be more the result of reduced nucleosome occupancy than a depletion in nucleosomes. This observation suggests increased frequencies of transient nucleosome positioning events in *F. graminearum* fast evolving polymorphic islands (47) and thus more relaxed chromatin structures. Here, nucleosome dynamics may enable fast evolution of particular genome segments while regions defined by higher occupancies may secure sequence conservation.

In Lmb, numerous "islands" of nucleosome-dense regions can be observed at various locations of the genome, including the dispensable chromosome (Figure 4A and Supplementary Figure 5). Aside from couple of contigs displaying higher nucleosome density, such characteristics were not observed for the closely related species Lml (Supplementary Figure 6). The locations of these nucleosome-dense islands in Lmb parallel those of AT-rich regions of the genome (Figure 4A and Supplementary Figure 5), features not visible in the genome of Lml (Supplementary Figure 6), suggesting that AT-rich regions are particularly dense in nucleosomes. Considering the remarkable compartmentalized organization of the genome of Lmb (17), absent from Lml, differences of nucleosome phasing and occupancy in TE- and AT-rich *vs.* GC-equilibrated and gene-rich regions were investigated. A region was considered AT-rich if it contained less than 40 % of GC. As described earlier, AT-rich regions represent one-third of the Lmb genome divided in 419 regions of 1 to 320 kb in length. Examination of unprocessed mapping outputs reveals that the number of fragments (read pairs) mapped in AT-rich and GC-equilibrated regions were very similar, with 23.8 million and 24.8 million fragments, respectively, which is far from the 1/3 *vs.* 2/3 ratio expected. In terms of coverage depth, mean coverage is 207 *vs.* 135 fragments for AT-rich and GC-equilibrated regions, respectively, which could suggest higher nucleosome occupancy in the former. We explored this hypothesis and compared phasograms for AT-rich *vs*. GC-equilibrated regions (Figure 4B and 4C, Supplementary Table 3). Average NRLs were found larger in AT-rich than GC-equilibrated compartments, measured at 169.2 bp and 164.2 bp respectively, suggesting lower nucleosome frequencies in AT-rich regions than in GC-equilibrated regions. Nonetheless, coverage

density is higher in AT-rich regions (Figure 4A and Supplementary Figure 5), consistently with our hypothesis of higher nucleosome occupancy in these regions and thus less accessible genome compartment, in heterochromatic state. This is in accordance with the recent genome-wide mapping of histone modifications performed by Soyer et al. (2020) on Lmb and Lml in which the Histone H3 Lysine9 tri-methylation heterochromatin mark was found associated with TE- and AT-rich regions of Lmb (29). Finally, signal intensity in phasograms appeared more stable on the long nucleotide range in GC-equilibrated than in AT-rich regions, an observation in line with the well-known destabilizing effect of AT stretches on nucleosome positioning leading to fuzzier signals (48–53). All together, these data support a heterochromatic state of Lmb AT-rich regions mediated by strong nucleosome occupancy during axenic growth. Since the AT-rich regions host many fungal effector genes expressed during primary infection of oilseed rape leaves, we may assume that these regions are decondensed during infection, allowing the action of specific transcription factors. We tried to perform MNase-seq experiments at an early stage of oilseed rape infection by Lmb but the number of fungal reads was too low to be able to reliably analyze fungal nucleosome positioning. To go further, techniques to enrich in fungal material prior to MNase treatment should be considered.

Sequence composition and nucleosome positioning

Literature data report that distribution of bases in nucleosome core DNA is non-random and exhibits a ~10 bp AA/TT/AT/TA offset with GG/CC/GC/CG dinucleotide frequency (54–56). Here, we investigated di-nucleotide frequencies – *i.e.,* the incidence of a given neighbouring pair of nucleotides in a sequence – in nucleosomal DNA segments in all four fungi. Averaged di-nucleotide contents centred around all fragments (read pairs) were plotted (Figure 5) and periodicities investigated by autocorrelation analyses (Supplementary Figure 7). Autocorrelation plots reveal the previously described ~10 bp-periodicities for all studied fungi while showing differences in instant autocorrelation coefficient profiles (Supplementary Table 4 and Supplementary Figure 7). Signal is indeed very regular in *F. graminearum* and, to a lesser extent, Lmb, whereas Lml and *B. cinerea* show more irregular autocorrelation profiles. These results are consistent with our previous observation that nucleosomes are tightly phased in *F. graminearum* whereas somehow fuzzier (higher deviation) in *B. cinerea* and Lml (Figure 1B and 1C).

Di-nucleotide frequency graphs display A+T dinucleotides frequency waves oscillating out of phase with G+C ones. For all studied fungi, GC dinucleotides are centred on nucleosome dyads (Figure 5). Considering that there are 16 possible combinations of di-nucleotides, equilibrated distribution of di-nucleotide contents should contain 25% of AT/TA/AA/TT and 25% of GG/CC/CG/GC. Our observations reveal a skewed distribution in favour of AT dimers marked for Lmb and *B. cinerea* (Figure 5A and 5C. The presence of AT-rich regions in Lmb, and to a lesser extent in *B. cinerea*, genomes may explain such a result (17, 24). Di-nucleotide frequencies of Lmb AT-rich and GC-equilibrated regions were thus inspected (Supplementary Table 5 and Supplementary Figure 8). As one would expect, A+T frequencies are particularly high in AT-rich regions, while maintaining alternance with G+C di-nucleotides and ~10 bp periodicity.

Nucleosome positioning is believed to be particularly hard-wired to DNA sequence, and especially the largely documented anti-nucleosome effect of Poly(dA:dT) tracts (12, 57). Here, while nucleosome phase was indeed found 5 bp longer in Lmb AT-rich regions than in GC-equilibrated regions, occupancy was nonetheless higher in the former leading to the formation of the previously suggested heterochromatic state of these regions, which has consequences on gene expression and recombination (17, 29, 58, 59). These observations suggest the mobilisation of trans-acting chromatin remodelling factors to maintain heterochromatin structures on such disfavouring sequences. Importantly, we found that GC periodicity at nucleosome dyads is preserved even within AT-rich regions, suggesting such pattern in an AT-rich environment is sufficient to permit efficient wrapping of DNA around nucleosomes and strong occupancy. This longer NRLs in addition with high nucleosome density in AT-rich regions should have an impact on gene expression in these regions enriched in effector gene (17, 29).

Nucleosome landscapes of fungal gene units

Nucleosome occupancy profiles of gene units were investigated in all fungi (Figure 6). As previously reported in other eukaryotes, translation start sites (the ATG codon) are preceded by a nucleosome-depleted region (NDR) and immediately followed by a well-positioned +1 nucleosome (Figure 6A). Variations of the exact position of these features relative to the ATG start codon, as well as variations in the intensity of the NDR valley, are nonetheless observed between fungal species. For example, the NDR and the centre of the +1 nucleosome (or +1Nucl) are found at -154 bp and +14 bp, respectively, in *F. graminearum* whereas they are located at -99 bp and + 26 bp, respectively, in *B. cinerea*. The fungus *B. cinerea* shows the deepest NDR valley among all observed profiles. In Lmb and Lml, NDRs are found at -129 and -144 bp from ATG, respectively, and +1Nucl at +19 bp and +55 bp. Finally, nucleosome profiles upstream of NDRs appear fuzzy for all fungi but Lmb, with varying degrees of fuzziness. This fuzziness is no longer visible when nucleosome profiles are centred on TSS for *F. graminearum* (N_{TSS} = 6,212 genes) and *B. cinerea* (N_{TSS} = 11,701 genes) (Figure 6C). The NDR is more defined, and located immediately upstream of the TSS, with a minimum detected at -58 bp and -20 bp upstream of the TSS of *F. graminearum* and *B. cinerea*, respectively. These values are consistent with the binding of the RNA polymerase II ~50 bp upstream of TSS, observed in active promoters of mammalian and Drosophila cells (60–62). Considering nucleosome environments at stop codons (Figure 6B), strongly arrayed nucleosomes are particularly found upstream the stop codon, fewer signal variations being observed downstream. Here, the stop codon is a clear boundary for nucleosome arraying and occupancy in all fungi and all conditions investigated. A nucleosome seems remarkably well positioned exactly on stop codons in *F. graminearum* in particular. The analysis was repeated on TTS in *F. graminearum* (N_{TTS} = 5,292 genes) and *B. cinerea* (N_{TTS} = 11,701 genes) (Figure 6D). In *B. cinerea*, signal appeared strong and well-arrayed, decaying downstream of the TTS. In *F. graminearum*, strong positioning of nucleosomes -2 and -1 (-178 bp and -16 bp relative to TTS, respectively) followed by a deep 3' end NDR (+100 bp downstream of TTS) can be observed. Nucleosome positioning on TSS and TTS could not be analyzed for *Lmb* and *Lml* since TSS and TTS annotations are not supported by

experimental data for these species' genomes as it is the case for *F. graminearum* (23, 63), or by collaborative annotation as for *B. cinerea* (64).

For comparison purposes, nucleosome profiling was repeated restricting our analysis to genes identified as BUSCO lineage-specific single-copy evolutionary conserved orthologs in Fungi (36, 37). Overall broad patterns remain similar to those obtained while investigating whole genomes, with notably somehow more regular oscillations patterns (Figure 7). In *F. graminearum*, whilst the distance NDR-to-TSS (- 60 bp) remains very similar to the one measured earlier genome-wide (Figure 7C), distance NDR-to-ATG increases by 29 bp (Figure 7A), whereas in *B. cinerea* the distance NDR-to-TSS reduces to 0 bp and nearly no increase in the distance NDR-to-ATG is observed (+ 3 bp). Similarly, the distance NDR-to-ATG is only 5 bp longer than genome-wide in Lmb whereas it increases by 23 bp in Lml. Towards the 3' end of the gene unit, nucleosome signals around stop codons are similar to those obtained genome-wide for all fungi (Figure 7B). However, the deep NDR found downstream of TTS of *F. graminearum* genome-wide can no longer be observed at "Fungi" TTS *loci* (Figure 7D). Similar to genome-wide profile, strongly positioned -1Nucl and -2Nucl are still visible at -5 bp and -209 bp, respectively, the latter being more intense and defined.

The general profile of a fungal gene unit shares similarities with those previously described in various eukaryotes: the ATG codon is decorated by a well-positioned +1 nucleosome and preceded by an NDR. When we restricted our analysis to conserved orthologous fungal genes, the overall profiles and the intensities of NDRs were more homogenous between fungi and culture conditions. These observations suggest that NDRs positions and intensities are subjected to evolution. Similarly, while an NDR can be observed downstream of *F. graminearum* TTS, it is no longer visible when the analysis is restricted to fungal genes, indicating again an evolutionary component.

Nucleosome landscapes of gene units according to gene expression

Same analyses were repeated for genes categorised according to their expression levels (expressed in TPM counts, see Materials and Methods). The general variations in nucleosome profiles around translation start sites are similar for all expression categories in all considered fungi and culture conditions: ATG codons are immediately followed by a well-positioned +1 nucleosome and preceded by a dip in nucleosome density (Figure 8). Remarkable variations are nonetheless observed with regard to positions of +1 nucleosomes and NDRs, as well as the amplitude of the nucleosome signal difference between them (here defined as $\Delta nucl = |signal_{+1nucl} - signal_{NDR}|$), depending on gene expression. ATG-centred nucleosome profiles for genes not expressed in our conditions (TPM = 0) show remarkably reduced Δnucl when compared to those measured for genes more expressed, and a distance to the ATG reduced (Figure 8 and Supplementary Table 6). Conversely, highly expressed genes (TPM50) display the deepest NDRs located at the furthest upstream the ATG. Similar trends are observed when profiles are centred on *B. cinerea* and *F. graminearum* TSS (Figures 8E and 8F, respectively). Moreover, NDRs were usually found further from the ATG site than those in genes expressed at lower levels or not expressed. Finally, the nucleosome wave signal decay phenomenon was observed at distances from the ATG codons shorter in poorly expressed genes than in highly expressed genes, although the

+1 nucleosome remained fairly well conserved. Nucleosome depletion can be the side-effect of active transcription with the binding of pre-initiation complex resulting in nucleosome eviction as previously shown in yeast (65). Indeed, we evidenced valleys of nucleosome signals upstream of ATG codons found deeper in highly expressed genes. On the contrary, genes not expressed showed little or no NDR, depending on the considered species or culture condition. In our conditions, the amplitude between the NDR and the +1 nucleosome seemed to be an informative measure of gene expression level: the higher this value is, the more genes are expressed. This feature was less strict when TSS were considered, raising the question of different mechanisms of transcription regulation depending on gene unit structures. The marked oscillations of the signal obtained for highly expressed genes suggests that the presence of well-arrayed nucleosomes around ATG combined with high amplitudes of nucleosome signal between the NDR valley and the +1 nucleosome peak could be a hallmark of active gene expression.

At the end of gene units, wave signal decay was also observed downstream of stop codons in all studied fungi (Supplementary Figure 9). In Lmb and Lml, strong nucleosome positioning was found immediately upstream the stop codon, directly followed by a strong nucleosome valley in the case of Lml or a general decrease in signal in Lmb (Supplementary Figure 9A and 9B). In *B. cinerea* and *F. graminearum*, nucleosome signal was more defined with clear oscillations decaying past the stop codon (Supplementary Figures 9C and 9D). In *F. graminearum*, a lesser signal intensity seemed to characterise highly to moderately expressed genes (Supplementary Figure 9D), a feature visible only for highly expressed genes in *B. cinerea* (Supplementary Figure 9C). Altogether, our observations highlight the conservation of a nucleosome immediately before the stop codon, followed by a decrease in signal as a mark of gene expression in the four studied fungi. Markedly, when TTS of *B. cinerea* were considered, not expressed genes revealed a remarkably regular signal of weak amplitude. On the contrary, nucleosome signal around *F. graminearum* TTS of genes not expressed were characterized by a strong wave signal of well-arrayed nucleosomes (Supplementary Figure 9F). Overall, our results showed strong association of nucleosome landscapes at gene unit boundaries with expression levels in the four studied ascomycetes.

CONCLUSION

The present study explored nucleosome landscapes of four phytopathogenic filamentous fungi with the aim of unravelling common features as well as potential specificities. Our general observation of nucleosome positioning genome-wide revealed shorter nucleosome-repeat lengths in *L. maculans* 'brassicae' and *L. maculans* 'lepidii' compared to *B. cinerea* and *F. graminearum*, suggesting a more compact chromatin fibre. High nucleosome occupancy was further observed in AT-rich regions of *L. maculans* 'brassicae', a feature *a priori* unexpected considering the well-described destabilising effect of AT-stretches but in line with the heterochromatic nature of these peculiar regions. High nucleosome occupancy was also observed at the *loci* of BOTY retrotransposons in the genome of *B. cinerea*. On the contrary, regions with reduced occupancy were observed in *F. graminearum* and co-localised with

highly polymorphic regions described as prone to genetic evolution. As a whole, our results plead in favour of evolution of not only the positions of nucleosomes but also their occupancy, both likely hard-wired to genome sequence evolution, with regions defined by higher occupancies possibly securing sequence conservation.

AVAILABILITY

The MSTS (MNase-seq Tool Suite) is an open-source collection of tools developed by the BioinfoBIOGER platform by N. Lapalu and A. Simon, and available in the GitHub repository (https://github.com/nlapalu/MSTS).

ACCESSION NUMBERS

All sequenced reads have been deposited with the Short Read Archive (SRA; https://www.ncbi.nlm.nih.gov/sra) under project accession number PRJNA580372. RNA-Seq data have been deposited in the Gene Expression Omnibus Database (GEO) (http://www.ncbi.nlm.nih.gov/geo/) under the entries GSE150127, GSE162838, and GSE162839.

SUPPLEMENTARY DATA

Supplementary Data are available online.

ACKNOWLEDGEMENT

Sequencing was performed by the GenomEast platform, a member of the “France Génomique” consortium (ANR-10-INBS-0009). We particularly thank B. Jost for his help with setting up this project. We are grateful to the Genotoul bioinformatics platform Toulouse Midi-Pyrenees for providing computing resources.

FUNDING

This work was supported by the Plant Health and Environment division of the French National Institute for Agricultural Research (TACTIC Project AAP 2014). Funding for open access charge: Plant Health and Environment department of the French National Institute for Agricultural Research. J. L. Soyer was

funded by a “Contrat Jeune Scientifique” grant from INRA. The BIOGER Unit benefits from the support of Saclay Plant Sciences-SPS (ANR-17-EUR-0007).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Hawksworth,D.L. and Lücking,R. (2017) Fungal Diversity Revisited: 2.2 to 3.8 Million Species. *Microbiol. Spectr.*, **5**.

2. Taylor,T.N. and Osborn,J.M. (1996) The importance of fungi in shaping the paleoecosystem. *Rev. Palaeobot. Palynol.*, **90**, 249–262.

3. Fisher,M.C., Henk,D.A., Briggs,C.J., Brownstein,J.S., Madoff,L.C., McCraw,S.L. and Gurr,S.J. (2012) Emerging fungal threats to animal, plant and ecosystem health. *Nature*, **484**, 186–194.

4. Lo Presti,L., Lanver,D., Schweizer,G., Tanaka,S., Liang,L., Tollot,M., Zuccaro,A., Reissmann,S. and Kahmann,R. (2015) Fungal effectors and plant susceptibility. *Annu. Rev. Plant Biol.*, **66**, 513–545.

5. Weiberg,A., Wang,M., Lin,F.-M., Zhao,H., Zhang,Z., Kaloshian,I., Huang,H.-D. and Jin,H. (2013) Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science*, **342**, 118–123.

6. Collemare,J., O’Connell,R. and Lebrun,M.-H. (2019) Nonproteinaceous effectors: the terra incognita of plant-fungal interactions. *New Phytol.*, **223**, 590–596.

7. Sánchez-Vallet,A., Fouché,S., Fudal,I., Hartmann,F.E., Soyer,J.L., Tellier,A. and Croll,D. (2018) The Genome Biology of Effector Gene Evolution in Filamentous Plant Pathogens. *Annu. Rev. Phytopathol.*, **56**, 21–40.

8. Strauss,J. and Reyes-Dominguez,Y. (2011) Regulation of secondary metabolism by chromatin structure and epigenetic codes. *Fungal Genet. Biol. FG B*, **48**, 62–69.

9. Soyer,J.L., Rouxel,T. and Fudal,I. (2015) Chromatin-based control of effector gene expression in plant-associated fungi. *Curr. Opin. Plant Biol.*, **26**, 51–56.

10. Richmond,T.J. and Davey,C.A. (2003) The structure of DNA in the nucleosome core. *Nature*, **423**, 145–150.

11. Radman-Livaja,M. and Rando,O.J. (2010) Nucleosome positioning: how is it established, and why does it matter? *Dev. Biol.*, **339**, 258–266.

12. Struhl,K. and Segal,E. (2013) Determinants of nucleosome positioning. *Nat. Struct. Mol. Biol.*, **20**, 267–273.

13. Yuan,G.-C., Liu,Y.-J., Dion,M.F., Slack,M.D., Wu,L.F., Altschuler,S.J. and Rando,O.J. (2005) Genome-scale identification of nucleosome positions in S. cerevisiae. *Science*, **309**, 626–630.

14. Tsankov,A.M., Thompson,D.A., Socha,A., Regev,A. and Rando,O.J. (2010) The role of nucleosome positioning in the evolution of gene regulation. *PLoS Biol.*, **8**, e1000414.

15. Nishida,H., Motoyama,T., Yamamoto,S., Aburatani,H. and Osada,H. (2009) Genome-wide maps of mono- and di-nucleosomes of Aspergillus fumigatus. *Bioinforma. Oxf. Engl.*, **25**, 2295–2297.

16. Nishida,H., Kondo,S., Matsumoto,T., Suzuki,Y., Yoshikawa,H., Taylor,T.D. and Sugiyama,J. (2012) Characteristics of nucleosomes and linker DNA regions on the genome of the basidiomycete Mixia osmundae revealed by mono- and dinucleosome mapping. *Open Biol.*, **2**, 120043.

17. Rouxel,T., Grandaubert,J., Hane,J.K., Hoede,C., van de Wouw,A.P., Couloux,A., Dominguez,V., Anthouard,V., Bally,P., Bourras,S., *et al.* (2011) Effector diversification within compartments of the Leptosphaeria maculans genome affected by Repeat-Induced Point mutations. *Nat. Commun.*, **2**, 202.

18. Dutreux,F., Da Silva,C., d'Agata,L., Couloux,A., Gay,E.J., Istace,B., Lapalu,N., Lemainque,A., Linglin,J., Noel,B., *et al.* (2018) De novo assembly and annotation of three Leptosphaeria genomes using Oxford Nanopore MinION sequencing. *Sci. Data*, **5**, 180235.

19. Grandaubert,J., Lowe,R.G.T., Soyer,J.L., Schoch,C.L., Van de Wouw,A.P., Fudal,I., Robbertse,B., Lapalu,N., Links,M.G., Ollivier,B., *et al.* (2014) Transposable element-assisted evolution and adaptation to host plant within the Leptosphaeria maculans-Leptosphaeria biglobosa species complex of fungal pathogens. *BMC Genomics*, **15**, 891.

20. Amselem,J., Cuomo,C.A., van Kan,J.A.L., Viaud,M., Benito,E.P., Couloux,A., Coutinho,P.M., de Vries,R.P., Dyer,P.S., Fillinger,S., *et al.* (2011) Genomic analysis of the necrotrophic fungal pathogens Sclerotinia sclerotiorum and Botrytis cinerea. *PLoS Genet.*, **7**, e1002230.

21. Cuomo,C.A., Güldener,U., Xu,J.-R., Trail,F., Turgeon,B.G., Di Pietro,A., Walton,J.D., Ma,L.-J., Baker,S.E., Rep,M., *et al.* (2007) The Fusarium graminearum genome reveals a link between localized polymorphism and pathogen specialization. *Science*, **317**, 1400–1402.

22. King,R., Urban,M., Hammond-Kosack,M.C.U., Hassani-Pak,K. and Hammond-Kosack,K.E. (2015) The completed genome sequence of the pathogenic ascomycete fungus Fusarium graminearum. *BMC Genomics*, **16**, 544.

23. King,R., Urban,M. and Hammond-Kosack,K.E. (2017) Annotation of Fusarium graminearum (PH-1) Version 5.0. *Genome Announc.*, **5**.

24. Porquier,A., Morgant,G., Moraga,J., Dalmais,B., Luyten,I., Simon,A., Pradier,J.-M., Amselem,J., Collado,I.G. and Viaud,M. (2016) The botrydial biosynthetic gene cluster of Botrytis cinerea displays a bipartite genomic structure and is positively regulated by the putative Zn(II)2Cys6 transcription factor BcBot6. *Fungal Genet. Biol. FG B*, **96**, 33–46.

25. Kelloniemi,J., Trouvelot,S., Héloir,M.-C., Simon,A., Dalmais,B., Frettinger,P., Cimerman,A., Fermaud,M., Roudet,J., Baulande,S., *et al.* (2015) Analysis of the Molecular Dialogue Between Gray Mold (Botrytis cinerea) and Grapevine (Vitis vinifera) Reveals a Clear Shift in Defense Mechanisms During Berry Ripening. *Mol. Plant-Microbe Interact. MPMI*, **28**, 1167–1180.

26. Simon,A., Dalmais,B., Morgant,G. and Viaud,M. (2013) Screening of a Botrytis cinerea one-

hybrid library reveals a Cys2His2 transcription factor involved in the regulation of secondary metabolism gene clusters. *Fungal Genet. Biol. FG B*, **52**, 9–19.

27. Boutigny,A.-L., Barreau,C., Atanasova-Penichon,V., Verdal-Bonnin,M.-N., Pinson-Gadais,L. and Richard-Forget,F. (2009) Ferulic acid, an efficient inhibitor of type B trichothecene biosynthesis and Tri gene expression in Fusarium liquid cultures. *Mycol. Res.*, **113**, 746–753.

28. Soyer,J.L., Möller,M., Schotanus,K., Connolly,L.R., Galazka,J.M., Freitag,M. and Stukenbrock,E.H. (2015) Chromatin analyses of Zymoseptoria tritici: Methods for chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq). *Fungal Genet. Biol. FG B*, **79**, 63–70.

29. Soyer,J.L., Clairet,C., Gay,E.J., Lapalu,N., Rouxel,T., Stukenbrock,E.H. and Fudal,I. (2020) Genome-wide mapping of histone modifications in two species of Leptosphaeria maculans showing contrasting genomic organization and host specialization. *bioRxiv*, 10.1101/2020.05.08.084566.

30. Fudal,I., Ross,S., Gout,L., Blaise,F., Kuhn,M.L., Eckert,M.R., Cattolico,L., Bernard-Samain,S., Balesdent,M.H. and Rouxel,T. (2007) Heterochromatin-like regions as ecological niches for avirulence genes in the Leptosphaeria maculans genome: map-based cloning of AvrLm6. *Mol. Plant-Microbe Interact. MPMI*, **20**, 459–470.

31. Hallen,H.E., Huebner,M., Shiu,S.-H., Güldener,U. and Trail,F. (2007) Gene expression shifts during perithecium development in Gibberella zeae (anamorph Fusarium graminearum), with particular emphasis on ion transport proteins. *Fungal Genet. Biol. FG B*, **44**, 1146–1156.

32. Bolger,A.M., Lohse,M. and Usadel,B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinforma. Oxf. Engl.*, **30**, 2114–2120.

33. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinforma. Oxf. Engl.*, **29**, 15–21.

34. Li,B., Ruotti,V., Stewart,R.M., Thomson,J.A. and Dewey,C.N. (2010) RNA-Seq gene expression estimation with read mapping uncertainty. *Bioinformatics*, **26**, 493–500.

35. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.

36. Simão,F.A., Waterhouse,R.M., Ioannidis,P., Kriventseva,E.V. and Zdobnov,E.M. (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, **31**, 3210–3212.

37. Waterhouse,R.M., Seppey,M., Simão,F.A., Manni,M., Ioannidis,P., Klioutchnikov,G., Kriventseva,E.V. and Zdobnov,E.M. (2018) BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics. *Mol. Biol. Evol.*, **35**, 543–548.

38. Valouev,A., Johnson,S.M., Boyd,S.D., Smith,C.L., Fire,A.Z. and Sidow,A. (2011) Determinants of nucleosome organization in primary human cells. *Nature*, **474**, 516–520.

39. Szerlong,H.J. and Hansen,J.C. (2011) Nucleosome distribution and linker DNA: connecting nuclear function to dynamic chromatin structure. *Biochem. Cell Biol.*, **89**, 24–34.

40. Zhu,P. and Li,G. (2016) Structural insights of nucleosome and the 30-nm chromatin fiber. *Curr.*

Opin. Struct. Biol., **36**, 106–115.

41. Robinson,P.J.J., Fairall,L., Huynh,V.A.T. and Rhodes,D. (2006) EM measurements define the dimensions of the '30-nm' chromatin fiber: evidence for a compact, interdigitated structure. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 6506–6511.

42. Zhang,T., Zhang,W. and Jiang,J. (2015) Genome-Wide Nucleosome Occupancy and Positioning and Their Impact on Gene Expression and Evolution in Plants. *Plant Physiol.*, **168**, 1406–1416.

43. Locke,G., Haberman,D., Johnson,S.M. and Morozov,A.V. (2013) Global remodeling of nucleosome positions in C. elegans. *BMC Genomics*, **14**, 284.

44. Valouev,A., Ichikawa,J., Tonthat,T., Stuart,J., Ranade,S., Peckham,H., Zeng,K., Malek,J.A., Costa,G., McKernan,K., *et al.* (2008) A high-resolution, nucleosome position map of C. elegans reveals a lack of universal sequence-dictated positioning. *Genome Res.*, **18**, 1051–1063.

45. Lantermann,A.B., Straub,T., Strålfors,A., Yuan,G.-C., Ekwall,K. and Korber,P. (2010) Schizosaccharomyces pombe genome-wide nucleosome mapping reveals positioning mechanisms distinct from those of Saccharomyces cerevisiae. *Nat. Struct. Mol. Biol.*, **17**, 251–257.

46. Lieleg,C., Ketterer,P., Nuebler,J., Ludwigsen,J., Gerland,U., Dietz,H., Mueller-Planitz,F. and Korber,P. (2015) Nucleosome spacing generated by ISWI and CHD1 remodelers is constant regardless of nucleosome density. *Mol. Cell. Biol.*, **35**, 1588–1605.

47. Laurent,B., Moinard,M., Spataro,C., Ponts,N., Barreau,C. and Foulongne-Oriol,M. (2017) Landscape of genomic diversity and host adaptation in Fusarium graminearum. *BMC Genomics*, **18**, 203.

48. Kaplan,N., Moore,I.K., Fondufe-Mittendorf,Y., Gossett,A.J., Tillo,D., Field,Y., LeProust,E.M., Hughes,T.R., Lieb,J.D., Widom,J., *et al.* (2009) The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature*, **458**, 362–366.

49. Tillo,D., Kaplan,N., Moore,I.K., Fondufe-Mittendorf,Y., Gossett,A.J., Field,Y., Lieb,J.D., Widom,J., Segal,E. and Hughes,T.R. (2010) High nucleosome occupancy is encoded at human regulatory sequences. *PloS One*, **5**, e9129.

50. Russell,K., Cheng,C.-H., Bizzaro,J.W., Ponts,N., Emes,R.D., Le Roch,K., Marx,K.A. and Horrocks,P. (2014) Homopolymer tract organization in the human malarial parasite Plasmodium falciparum and related Apicomplexan parasites. *BMC Genomics*, **15**, 848.

51. Bunnik,E.M., Polishko,A., Prudhomme,J., Ponts,N., Gill,S.S., Lonardi,S. and Le Roch,K.G. (2014) DNA-encoded nucleosome occupancy is associated with transcription levels in the human malaria parasite Plasmodium falciparum. *BMC Genomics*, **15**, 347.

52. Jin,H., Finnegan,A.I. and Song,J.S. (2018) A unified computational framework for modeling genome-wide nucleosome landscape. *Phys. Biol.*, **15**, 066011.

53. Ponts,N., Harris,E.Y., Prudhomme,J., Wick,I., Eckhardt-Ludka,C., Hicks,G.R., Hardiman,G., Lonardi,S. and Le Roch,K.G. (2010) Nucleosome landscape and control of transcription in the human malaria parasite. *Genome Res.*, **20**, 228–238.

54. Satchwell,S.C., Drew,H.R. and Travers,A.A. (1986) Sequence periodicities in chicken

nucleosome core DNA. *J. Mol. Biol.*, **191**, 659–675.

55. Segal,E., Fondufe-Mittendorf,Y., Chen,L., Thåström,A., Field,Y., Moore,I.K., Wang,J.-P.Z. and Widom,J. (2006) A genomic code for nucleosome positioning. *Nature*, **442**, 772–778.

56. Segal,E. and Widom,J. (2009) What controls nucleosome positions? *Trends Genet. TIG*, **25**, 335–343.

57. Segal,E. and Widom,J. (2009) Poly(dA:dT) tracts: major determinants of nucleosome organization. *Curr. Opin. Struct. Biol.*, **19**, 65–71.

58. Soyer,J.L., El Ghalid,M., Glaser,N., Ollivier,B., Linglin,J., Grandaubert,J., Balesdent,M.-H., Connolly,L.R., Freitag,M., Rouxel,T., *et al.* (2014) Epigenetic control of effector gene expression in the plant pathogenic fungus Leptosphaeria maculans. *PLoS Genet.*, **10**, e1004227.

59. Gay,E.J., Soyer,J.L., Lapalu,N., Linglin,J., Fudal,I., Da Silva,C., Wincker,P., Aury,J.-M., Cruaud,C., Levrel,A., *et al.* (2020) Large-scale transcriptomics to dissect two years of the life of a fungal phytopathogen interacting with its host plant. *bioRxiv*, 10.1101/2020.10.13.331520.

60. Kwak,H., Fuda,N.J., Core,L.J. and Lis,J.T. (2013) Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science*, **339**, 950–953.

61. Core,L.J., Waterfall,J.J. and Lis,J.T. (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science*, **322**, 1845–1848.

62. Min,I.M., Waterfall,J.J., Core,L.J., Munroe,R.J., Schimenti,J. and Lis,J.T. (2011) Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. *Genes Dev.*, **25**, 742–754.

63. Basenko,E.Y., Pulman,J.A., Shanmugasundram,A., Harb,O.S., Crouch,K., Starns,D., Warrenfeltz,S., Aurrecoechea,C., Stoeckert,C.J., Kissinger,J.C., *et al.* (2018) FungiDB: An Integrated Bioinformatic Resource for Fungi and Oomycetes. *J. Fungi Basel Switz.*, **4**.

64. Pedro,H., Yates,A.D., Kersey,P.J. and De Silva,N.H. (2019) Collaborative Annotation Redefines Gene Sets for Crucial Phytopathogens. *Front. Microbiol.*, **10**, 2477.

65. Venters,B.J. and Pugh,B.F. (2009) A canonical promoter organization of the transcription machinery and its regulators in the Saccharomyces genome. *Genome Res.*, **19**, 360–371.

66. Balesdent,M.H., Attard,A., Ansan-Melayah,D., Delourme,R., Renard,M. and Rouxel,T. (2001) Genetic Control and Host Range of Avirulence Toward Brassica napus Cultivars Quinta and Jet Neuf in Leptosphaeria maculans. *Phytopathology*, **91**, 70–76.

67. Van Kan,J.A.L., Stassen,J.H.M., Mosbach,A., Van Der Lee,T.A.J., Faino,L., Farmer,A.D., Papasotiriou,D.G., Zhou,S., Seidl,M.F., Cottam,E., *et al.* (2017) A gapless genome sequence of the fungus Botrytis cinerea. *Mol. Plant Pathol.*, **18**, 75–89.

TABLE AND FIGURES LEGENDS

Table 1. Characteristics of reference genomes for the four fungal species studied.

Species	*Leptosphaeria maculans* 'brassicae'	*Leptosphaeria maculans* 'lepidii	*Botrytis cinerea*	*Fusarium graminearum*
Strain	v23.1.3	IBCN84	B05.10	CBS85.32
Reference genome	v23.1.3 (66)	IBCN84 (19)	B05.10 (67)	PH-1 RR1 (21–23)
Availability	EMBL/Genbank	EBI	EnsemblFungi v1	FungiDB v44
Genome size (Mb)	45.1	31.5	42	38.1
Number of protein-coding genes	12,468	11,272	11,701	14,164
GC content (%)*	44.1	50.9	42	48.2
Transposable elements content (%)	34.2	4.0	3.7	0.29

*excluding N's and mitochondrial genomes

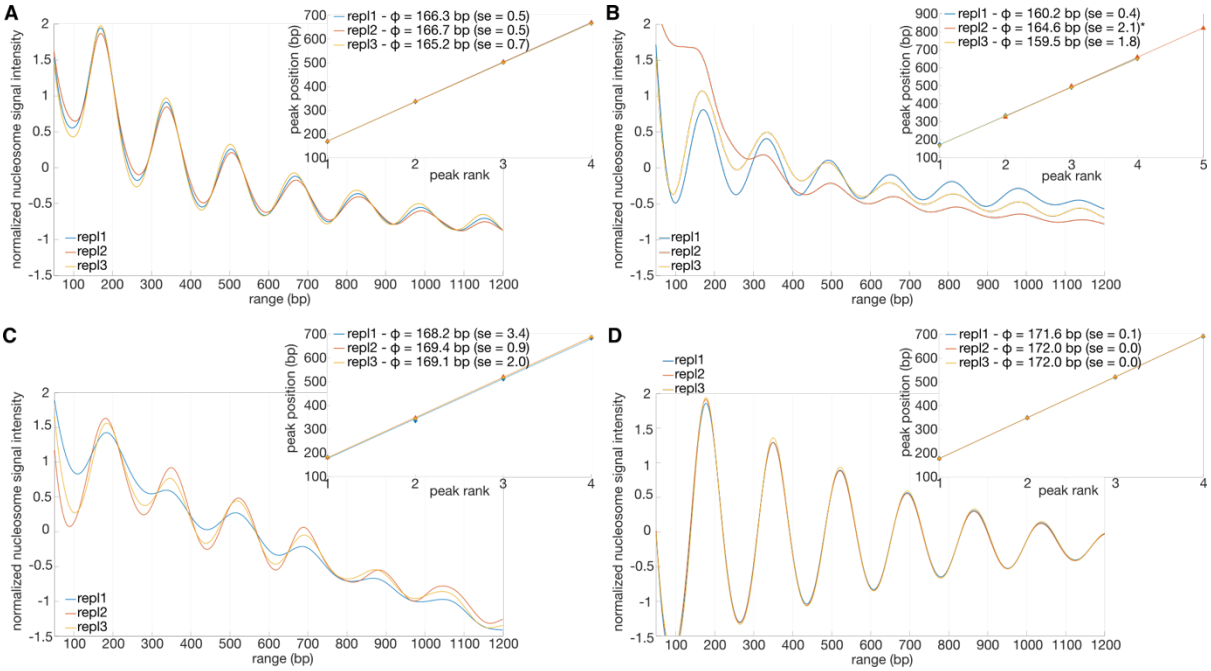

Figure 1. Nucleosome phasing in the four fungi studied. Main graphs display normalized (z-score) phase frequencies (y-axis) as a function of position (in base pair; x-axis). Graphs in inserts show peak positions (in base pairs; y-axis) as a function of peak order (x-axis). For each replicate, phases, standard errors (se), R^2 (coefficient of determination), and *p*-values (*F*-test) are determined after linear regression fitting to the first four successive peak positions (see Supplementary Table 2). Repl = replicate. A. *Leptosphaeria maculans* 'brassicae'; B. *Leptosphaeria maculans* 'lepidii'; here, phase value in replicate #2 was measured for four successive peaks excluding peak #1 for which an apex was not clearly visible at the beginning of the profile; C. *Botrytis cinerea*; D. *Fusarium graminearum*.

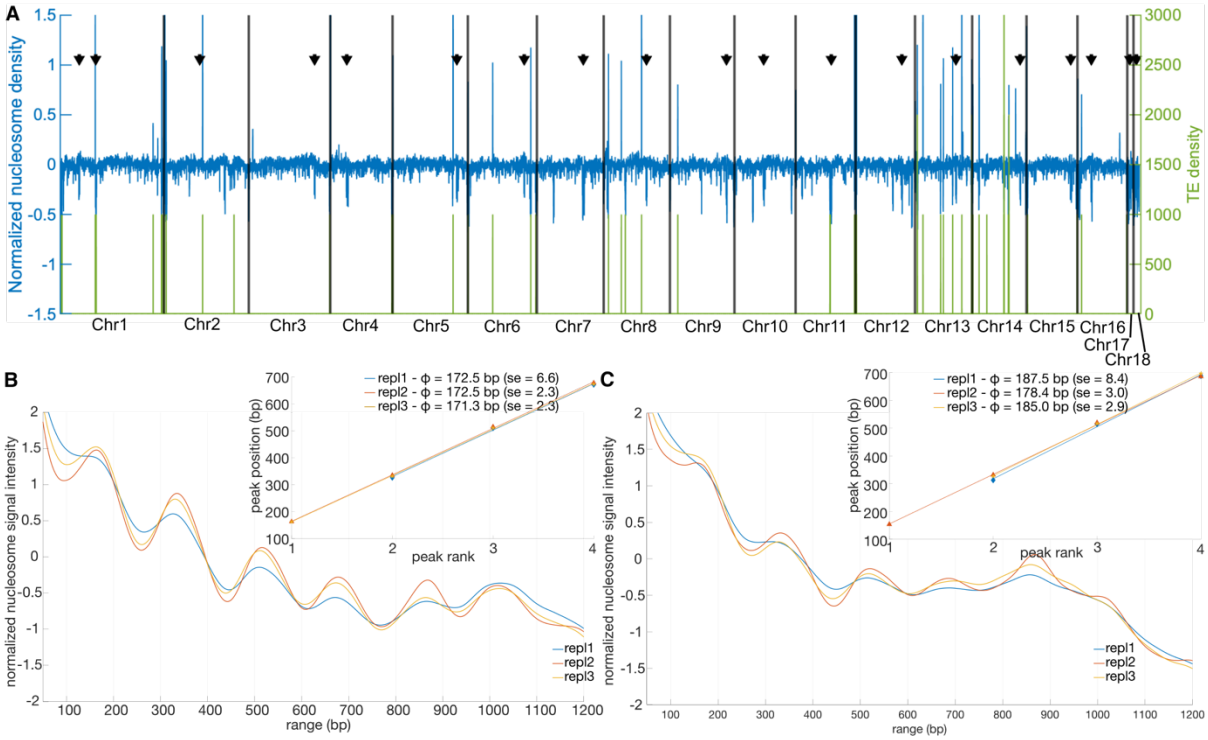

Figure 2. Nucleosome density profiles in *Botrytis cinerea*. A. Coverage density profiles were computed for non-overlapping 1 kb-long bins along the chromosomes of *B. cinerea*. In green are plotted BOTY transposable elements (TE) density profiles. In blue are plotted the z-scored average nucleosome density profile (see Supplementary Figure 2 for individual replicate plots). Black arrows indicate putative positions of centromeres (67); B and C. Nucleosome phasing in BOTY TE (B) or TE other than BOTY TE (C). Main graphs display phase frequencies (y-axis) as a function of position (in base pair; x-axis). Graphs in inserts show peak positions (in base pairs; y-axis) as a function of peak order (x-axis); Phases +/- standard errors (se), R^2 (coefficient of determination), and *p*-values (*F*-test) are determined after linear regression fitting to the first four successive peak positions (see Supplementary Table 3). Repl = replicate.

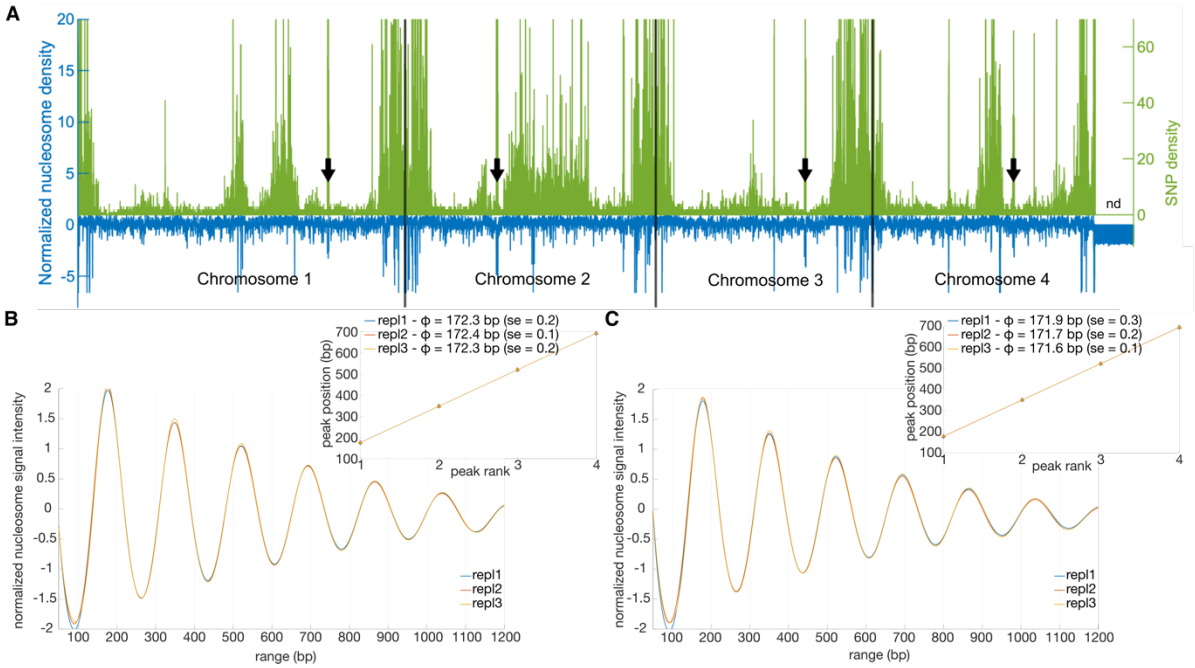

Figure 3. Nucleosome density profiles in *Fusarium graminearum*. A. Coverage density profiles were computed for non-overlapping 1 kb-long bins along the four chromosomes of *F. graminearum*. In green are plotted SNP density profiles as previously described (47). “nd” indicates the highly variable 3’ end of chromosome 4 for which SNP were not called. In blue are plotted the z-scored average nucleosome density profile (see Supplementary Figure 2 for individual replicate plots). Black arrows indicate centromeres (22); B and C. Nucleosome phasing in polymorphic islands (B) or outside polymorphic islands (C) as previously defined (47). Main graphs display phase frequencies (y-axis) as a function of position (in base pair; x-axis). Graphs in inserts show peak positions (in base pairs; y-axis) as a function of peak order (x-axis); Phases +/- standard errors (se), R^2 (coefficient of determination), and *p*-values (*F*-test) are determined after linear regression fitting to the first four successive peak positions (see Supplementary Table 3). Repl = replicate.

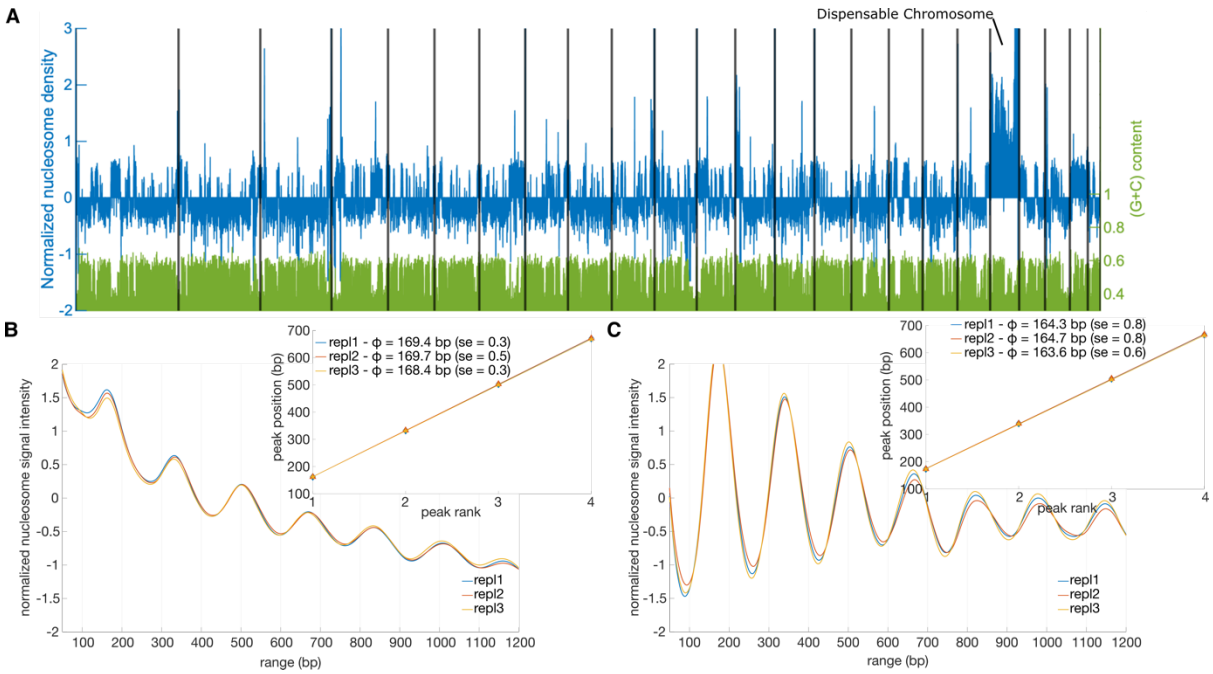

Figure 4. Nucleosome density profiles in *Leptosphaeria maculans* 'brassicae'. A. Coverage density profiles were computed for non-overlapping 1 kb-long bins along all supercontigs, separated by black lines. (G+C)/(A+T+G+C) content is plotted in green. In blue are plotted the z-scored average nucleosome density profile (see Supplementary Figure 3 for individual replicate plots). B and C. Nucleosome phasing in TE and AT-rich regions (B) and GC-equilibrated regions (C). Main graphs display phase frequencies (y-axis) as a function of position (in base pair; x-axis). Graphs in inserts show peak positions (in base pairs; y-axis) as a function of peak order (x-axis); for each replicate, phases, standard errors (se), R^2 (coefficient of determination), and *p*-values (*F*-test) are determined after linear regression fitting to the first four successive peak positions (see Supplementary Table 3). Repl = replicate.

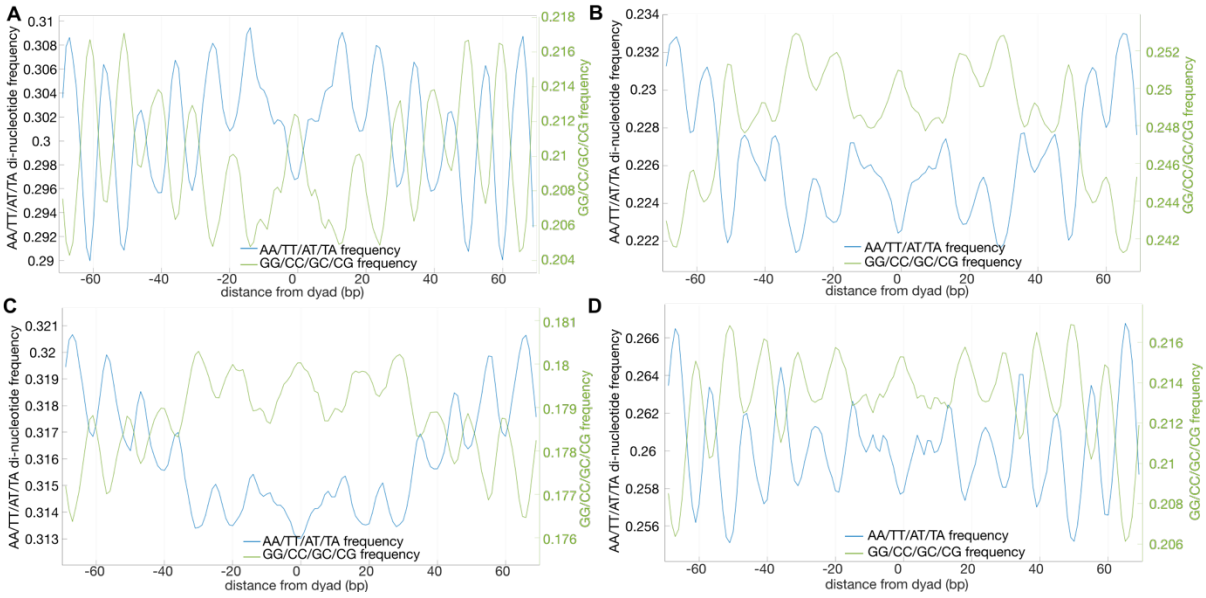

Figure 5. Repeated di-nucleotide patterns in nucleosomal DNA. Normalized di-nucleotides frequency plots (average of three biological replicates) for *Leptosphaeria maculans* 'brassicae' (A), *Leptosphaeria maculans* 'lepidii' (B), *Botrytis cinerea* (C), and *Fusarium graminearum* (D).

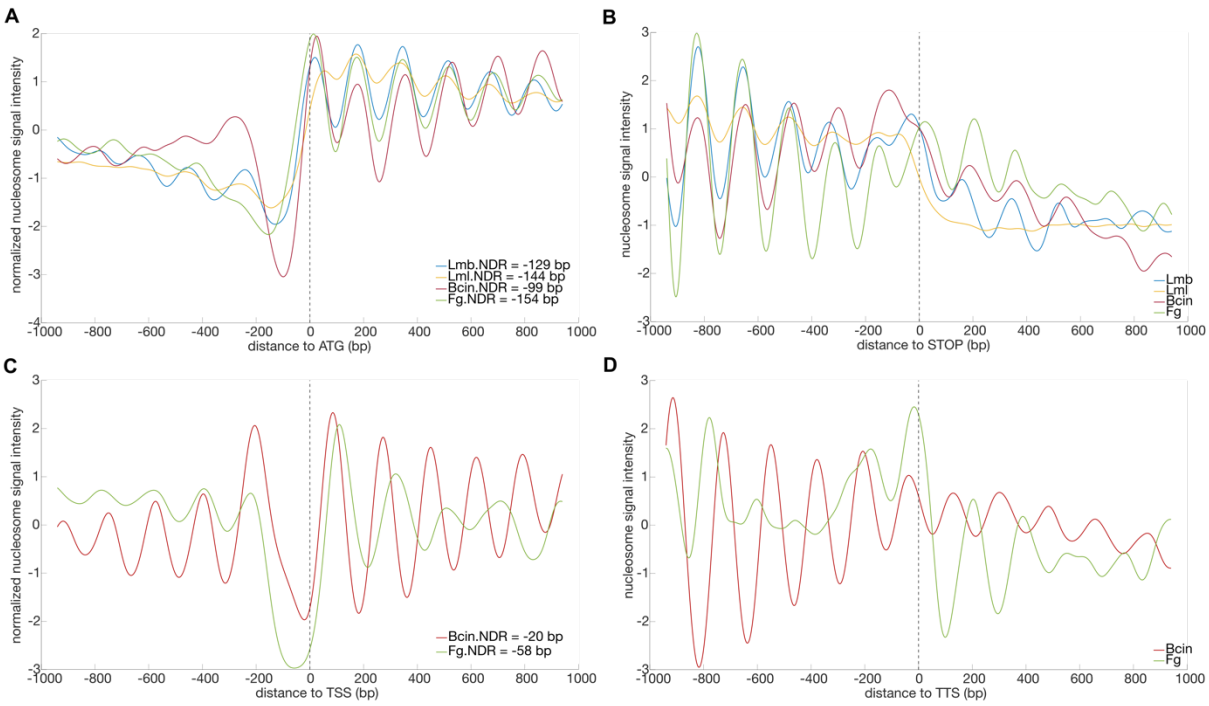

Figure 6. Nucleosome organization of fungal gene units. A-D. Normalized (z-scored) averages (three biological replicates for each fungus/condition) of nucleosome signals as a function of position (in base pairs) relative to the start codon ATG (A), the stop codon (B), TSS (C), TTS (D). Fg = *Fusarium graminearum*; Lmb = *Leptosphaeria maculans* 'brassicae'; Lml = *Leptosphaeria maculans* 'lepidii; Bcin = *Botrytis cinerea*; NDR = nucleosome-depleted region.

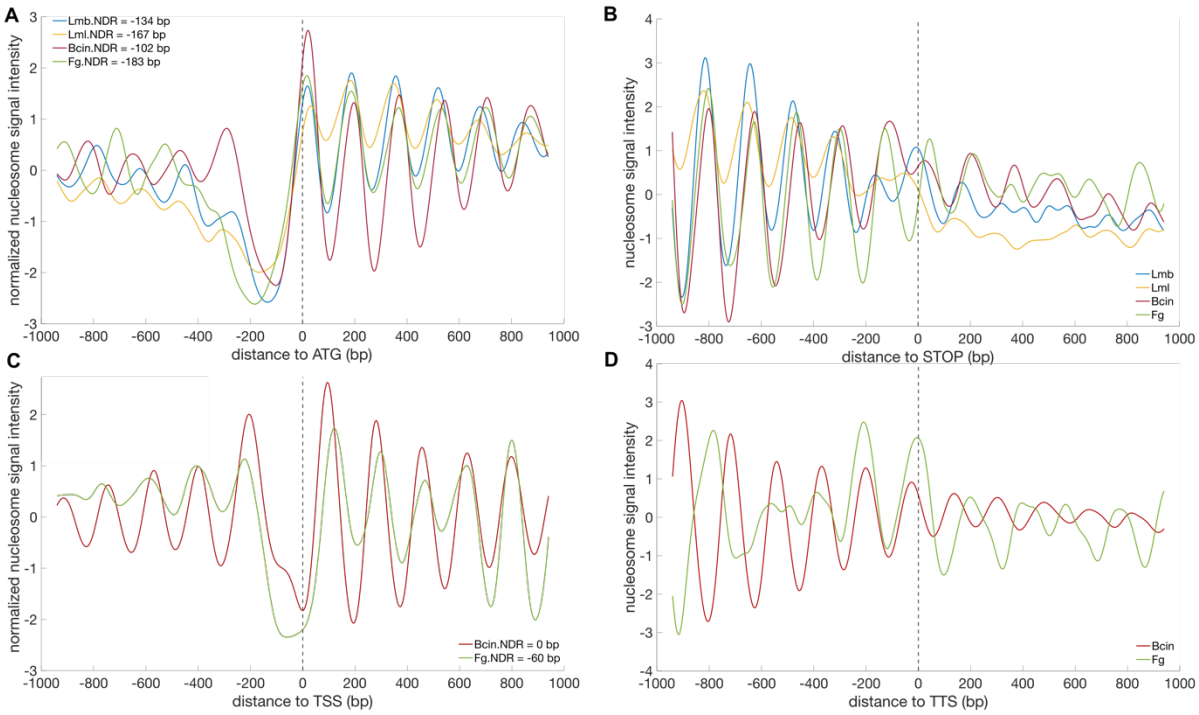

Figure 7. Nucleosome organization of near-universal single copy orthologous gene units in Fungi (BUSCO3). A-D. Normalized averages (three biological replicates for each fungus/condition) of nucleosome signals as a function of position (in base pairs) relative to the start codon ATG (A), the stop codon (B), TSS (C), and TTS (D). Fg = *Fusarium graminearum*; Lmb = *Leptosphaeria maculans* 'brassicae'; Lml = *Leptosphaeria maculans* 'lepidii; Bcin = *Botrytis cinerea*; NDR = nucleosome-depleted region.

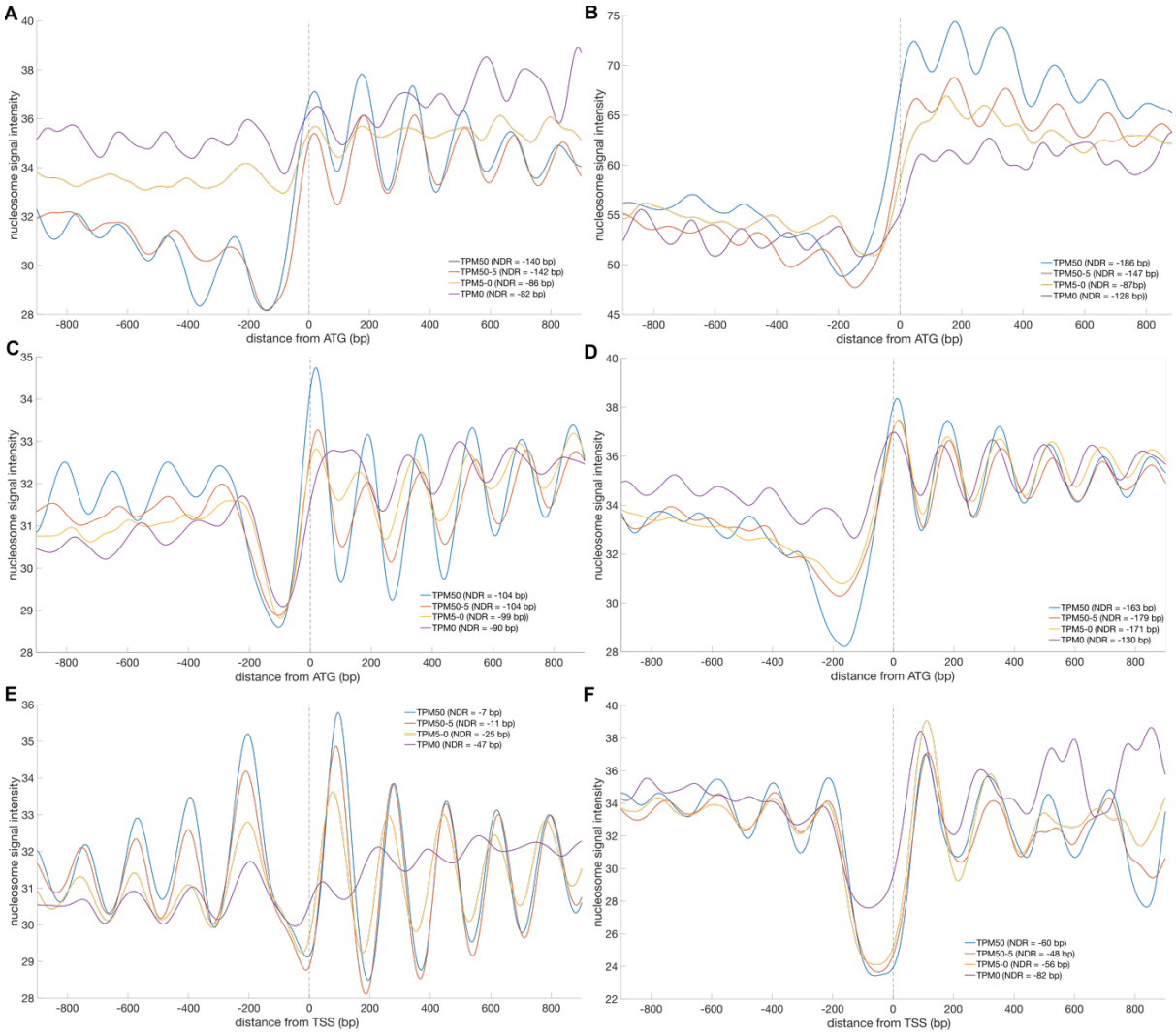

Figure 8. Nucleosome organization at start codons/TSS *vs.* gene expression. Average (three biological replicates for each fungus/condition) nucleosome signal as a function of position (in base pairs) relative to the start codon ATG (and TSS for *Fusarium graminearum* and *Botrytis cinerea*). TPM = Transcripts Per Million. (A), In *Leptosphaeria maculans* 'brassicae', ATG-centred; (B), In *Leptosphaeria maculans* 'lepidii', ATG-centred; (C), In *B. cinerea*, ATG-centred; (D), *F. graminearum*, ATG-centred; (E). In *B. cinerea*, TSS-centred; (F). *F. graminearum*, TSS-centred.