The Journey to the Treasure

Secondary metabolite gene clusters and histone modifications in the genus *Aspergillus*

Xin Zhang

The Journey to the Treasure

Secondary metabolite gene clusters and histone modifications in the genus *Aspergillus*

Xin Zhang

Xin Zhang
The Journey to the Treasure
Secondary metabolite gene clusters and histone modifications in the genus *Aspergillus*
PhD thesis, Utrecht University

Cover design: Shichao Liu（刘师超）

Layout design: Xin Zhang
Printing: ProefschriftMaken, www.proefschriftmaken.nl
ISBN: 978-94-6469-997-5

The Journey to the Treasure

Secondary metabolite gene clusters and histone modifications in the genus *Aspergillus*

De Reis naar de Schat

Secundaire metabolites genclusters en histonmodificaties in het geslacht *Aspergillus*

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het College voor Promoties
in het openbaar te verdedigen op

woensdag 3 juli 2024 des middags te 4.15 uur

door

Xin Zhang

geboren op 24 december 1994
te Liaoning, China

Promotor:

Prof. dr. B. Snel

Copromotoren:

Dr. M.F. Seidl

Dr. J. Collemare

Beoordelingscommissie:

Prof. dr. A.F.J.M. van den Ackerveken

Prof. dr. F. Hagen

Dr. K. Haslinger

Prof. dr. J. Strauss

Prof. dr. H.A.B. Wösten

Table of contents

Chapter 1

General introduction

Secondary metabolites

Humanity is currently facing unprecedented challenges such as climate change, increasing antimicrobial resistance, loss of biodiversity, and food security. Traditionally, we have relied on petroleum-based chemistry for industrial and consumer products, but now we are witnessing a transition to more natural approaches (Pham et al., 2019). It is urgent to discover novel chemistry, both to address our escalating global issues and to fulfill our demand for new, effective, and sustainable chemical compounds. Thus it is important to explore natural products, which encompass a variety of substances and compounds derived from living organisms (Louwen et al., 2023). They include both primary metabolites and secondary metabolites (SMs). Primary metabolites are essential for normal growth and reproduction, being directly involved in vital biological processes. They are produced at a high rate and can be easily extracted through simple biochemical methods. This has led to their early and extensive use by humans. Examples of such metabolites include starch and sugars, which are fundamental in various industrial and culinary applications (Sanchez & Demain, 2008). In contrast, SMs, which are typically of low molecular weight and structurally complex, facilitate the interaction between organisms and enable organisms to dynamically respond to environmental challenges (Hoffmeister & Keller, 2007). For instance, cyclosporin A, a fungal SM, can induce apoptosis and help to overcome the immune system in arthropod hosts, leading to the colonization of the host by the fungus (X. Yang et al., 2018). Similarly, SMs generated by plant pathogenic fungi can act as virulence factors, causing significant agricultural losses (Collemare & Lebrun, 2011). In competitive scenarios, certain bacteria, such as different *Pseudomonas* species, synthesize 2,4-diacetylphloroglucinol as an anti-microbial agent (Velusamy et al., 2006), while fungi deploy the antibiotic cephalosporin to suppress the growth of rival bacteria by disrupting their cell wall synthesis (Liu et al., 2022). Within challenging habitats, some microbes utilize siderophores to scavenge iron from their surroundings (De Serrano et al., 2016), a pathogenic bacteria, *Acinetobacter baumannii*, produces ectoines to safeguard its cells from osmotic stress (Zeidler & Müller, 2019), and melanin protects fungal spores against threats such as UV irradiation, enzymatic lysis, and extreme temperatures (Langfelder et al. 2003). In summary, SMs are structurally complex compounds that play key roles in organismal interactions and environmental adaptations, offering significant potential for innovative and sustainable solutions in various sectors.

Based on the wide range of biological activities of SMs, humans have explored the applications of SMs in medicine, agriculture, and various other industries (Pham et al., 2019). In the medical field, SMs are pivotal for drug development and therapeutic strategies, serving as a reservoir of biologically active compounds with antibiotic, antiviral, antifungal, and antitumor properties (Jakubiec-Krzesniak et al., 2018). For instance, penicillin, the first broad-spectrum antibiotic, has proven effective against bacterial infections and marked the beginning of the antibiotic era in modern medicine (Aly et al., 2011). Griseofulvin and amphotericin act

as antifungal agents (Newman & Cragg, 2016), while compounds like cyclosporine A (Laupacis et al., 1982) and psilocybin (Khan et al., 2022) function as an immunosuppressant and can be applied to treat trauma-related disorders, respectively. In agriculture, SMs are employed as biopesticides, thereby providing an eco-friendly alternative to traditional pesticides as they have specific pest targets and do not contribute to chemical pollution. For example, azadirachtin, derived from the Indian neem tree, has been commercialized and is utilized as a botanical pesticide worldwide (Kilani-Morakchi et al., 2021), and strobilurin, isolated mainly from mushrooms, contributes to controlling fungal diseases (Balba, 2007). In various industries, SMs are also crucial for providing flavors, colors, and antioxidants (Sanchez-Ballesta et al., 2022). For instance, curcumin, a yellow pigment derived from *Curcuma longa*, is commonly used as a food-coloring agent and in cosmetic products (Menon & Sudheer, 2007). Chromologics (https://www.chromologics.com/) sells a natural red pigment "Chromored" produced by the fungus *Talaromyces atroroseus* at the bioreactor scale (Morales-Oyervides et al., 2020). The Living Color Database documents approximately 445 pigment entries from both fungi and bacteria (Sharma & Meyer, 2022). Current research aims to understand how SMs are produced, discover novel SMs, uncover the functions and biological activities of SMs, and investigate methods to produce SMs on an industrial scale.

Fungal biosynthesis gene clusters

The fungal kingdom is remarkably large and diverse, comprising approximately 1.5 to 5 million species, which inhabit various ecological niches and exhibit different morphologies and life cycles (Choi & Kim, 2017). Interestingly, many fungi are known to be prolific producers of diverse SMs (Mosunova et al., 2021). SM biosynthesis in fungi is orchestrated by a series of enzymatic reactions that are catalyzed by diverse proteins, including the backbone (or core) enzymes, tailoring enzymes, and those proteins responsible for transcriptional control, transport, and/or host resistance (Figure 1) (Keller, 2019; Keller et al., 2005). Fungal SMs are categorized into several major chemical groups based on the precursors (acyl-coAs, amino acids, or prenyl diphosphates) they utilize to synthesize and the types of backbone enzymes. Backbone enzymes such as polyketide synthase (PKS), non-ribosomal peptide synthetase (NRPS), hybrid PKS-NRPS enzymes, dimethylallyl tryptophan synthetase (DMATS), and terpene cyclase (TC), catalyze the production of polyketides, non-ribosomal peptides, PKS-NRPS hybrids, indole alkaloids, and terpenes, respectively (Mosunova et al., 2021). Following this initial step, tailoring enzymes modify the chemical structure of the intermediate compounds through processes such as methylation, acylation, glycosylation, or hydroxylation (Mosunova et al., 2021), giving rise to a diverse array of structurally complex and functionally distinct SMs. Transport-related genes encode proteins to export the final products from the fungal cell into the extracellular space (Gomez-Escribano and Bibb 2011), and genes associated with resistance, provide protection to the fungi against the toxic effects of some SMs (Gomez-Escribano & Bibb, 2011). Considering that certain biosynthetic pathways contain

more than one backbone gene, leading to hybrid molecules, and that one biosynthetic pathway can include several tailoring steps, the combination of these factors allows to production of a vast and yet underexplored SMs from a limited set of chemical precursors.

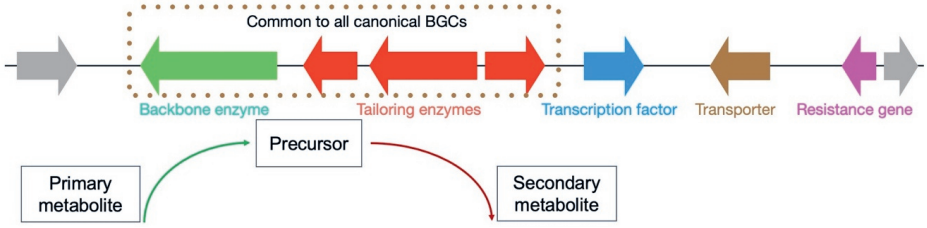


Figure 1 Step-wise biosynthesis of secondary metabolites (SMs) encoded by biosynthetic gene clusters (BGCs). This schematic represents a single BGC, composed of genes that encode for a backbone and several tailoring enzymes, in addition to genes fulfilling other specific functions. The biosynthesis process of an SM is sequential: it is initiated by the transformation of a primary metabolite into a precursor, facilitated by a backbone enzyme. This is followed by the action of tailoring enzymes, which further modify the precursor to yield the final SM.

The genes related to the biosynthesis of a fungal SM are often physically co-localized along the chromosome and co-regulated, an organization that is referred to as the biosynthetic gene cluster (BGC) (Figure 1). This special organization not only facilitates the sharing of intermediate compounds among sequential enzymes, enhancing the production of SMs (Pfannenstiel & Keller, 2019), but also may act as a suitable unit for the duplication, deletion, or transfer of the complete SM pathway (Rokas et al., 2018). Moreover, the clustering might minimize the risk of the accumulation of toxic intermediates (Rokas et al., 2018; Wisecaver et al., 2014). An examination of 100 fungal genomes shows a bias for the physical pairing of genes encoding enzymes that share a toxic intermediate (McGary et al., 2013), underscoring the evolutionary advantage and functional significance of this genomic organization of genes involved in the biosynthesis of SMs. Despite the strong selection pressure to maintain genes involved in SM biosynthesis in gene clusters, extensive diversity in the presence and organization of SM-BGCs has been detected by genome mining across the fungal kingdom, ranging from a few to several dozen SM-BGCs per species (Figure 2) (Mosunova et al., 2021). Species within the phylum Ascomycota on average encode more backbone genes compared to other fungal phyla, in which Eurotiomycetes, Lecanoromycetes, and Sordariomycetes are the top three classes with more than 45 backbone genes, especially in PKS, NRPS-like, and NRPS (Figure 2). Basidiomycota, including wood-decaying species and edible mushrooms, generally encodes fewer backbone genes compared with Ascomycota and it is relatively enriched in NRPS-like and TC groups (Figure 2). Neocallimastigomycetes are especially rich in NRPS genes, while other fungal classes show a reduced production capacity of this type of backbone enzyme. In light of these observations, it is necessary and also most effective to

focus on species within the same genus to understand the diversity of SMs in newly sequenced species. For comparisons across different fungal genera, careful attention should be given to the inherent diversity and complexity of SM-BGCs given their significant variations across the fungal kingdom.

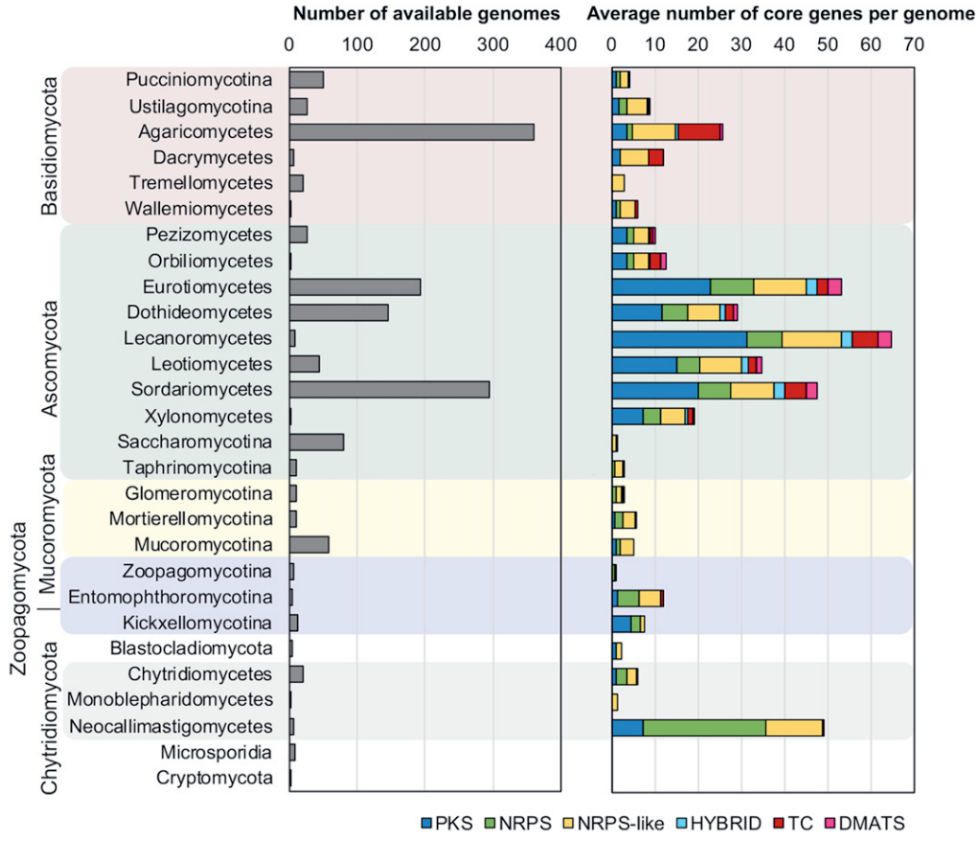


Figure 2 Distribution of BGCs in the fungal kingdom. Mosunova et al. highlight the phylum Ascomycota as having the highest abundance of BGCs among fungi. Data on the number of genomes and SM genes are presented at the class level, with classes within the same phylum unified by the same background color. Distinct SM groups are differentiated using various colors in the bar plots, providing a clear visual breakdown of their presence across different fungal classes. Figure adapted from (Mosunova et al., 2021).

Local and global transcriptional regulation of SM-BGCs

Given that the functions of SMs are diverse and considering the energy cost of their production, SMs are only synthesized when necessary. The biosynthesis of fungal SMs is tightly regulated by transcription factors (TFs), both on a local scale as well as globally (Brakhage, 2013). Some

BGCs have one or several genes encoding TFs as part of the cluster, which can govern the co-expression of BGC. For instance, aflR, one of the earliest-studied TFs within BGC, regulates the expression of the aflatoxin and sterigmatocystin genes in *Aspergillus* species (Woloshuk et al., 1994). To activate silent BGCs, scientists have tried to overexpress local TFs by introducing an extra copy, which led to the discovery of aspyridone in *Aspergillus nidulans* (Bergmann et al., 2007). Moreover, replacing the local promotor with an inducible one can turn on the biosynthesis of asperfuranone in *A. nidulans* (Chiang et al., 2009). In response to environmental stimuli, global transcriptional regulators affect the expression of some BGCs as well as other genes throughout the genome (Macheleidt et al., 2016). Global TFs, such as CreA, AreA, or Hap, control the gene expression response to the carbon source (Lyu et al., 2020), nitrogen source (B Tudzynski et al., 1999; Bettina Tudzynski, 2014), and iron starvation (Hortschansky et al., 2017), respectively. Furthermore, PacC regulates BGC expression in response to pH (Lyu et al., 2020), while the Velvet and LaeA protein complexes control the response to light (Bayram et al., 2008). The manipulation of global TFs leads to diverse effects of different BGCs and other genes. For instance, the deletion of *mcrA* (*LaeA*-like) in *A. nidulans* results not only in upregulations of sterigmatocystin, terrequinone A, and cichorine, but also alters the expression of more than 100 genes throughout the genome (Grau et al., 2019). In summary, these intricate regulatory networks highlight the potential for developing new methods to harness the production of valuable SMs by fine-tuning the expression of both local and global TFs.

Histone PTMs-based regulation of SM-BGCs

Within the eukaryotic nucleus, the genetic material is organized into repeating basic units, nucleosomes. They are formed by a 145–147 bp DNA strand that is wrapped around a histone octamer composed of two copies each of histone protein 2A (H2A), H2B, H3, and H4 (Luger et al., 1997). Each histone protein has a globular central domain that facilitates the interaction with nucleosomal DNA, a lysine- and arginine-rich N-terminal tail, and a short C-terminal tail (van Holde, 1989). Multiple nucleosomes, connected by a 30-50 bp of DNA and a linker histone protein H1, can aggregate to form a higher structure that is referred to as chromatin (Luger et al., 1997). The compaction of nucleosomes within the chromatin is thought to be dynamic and chromatin alternates between the ‘open’ euchromatin, a loosely packed and transcriptionally active state, and the ‘closed’ (facultative or constitutive) heterochromatin, a condensed and transcriptionally silent state (Grewal & Jia, 2007). Chemical post-translational modifications (PTMs) on the tails of histone proteins as well as chemical modifications of the DNA itself contribute to the formation and maintenance of chromatin states as well as dynamic transitions between states. Especially histone PTMs can directly alter histone proteins’ net charges or mediate the interaction with other inter-nucleosomal proteins, and consequently, histone PTMs and DNA modifications are thought to influence chromatin compaction (Henikoff & Shilatifard, 2011). Therefore, the chromatin state and dynamic transitions between states

are crucial for organisms to regulate gene expression in response to both developmental and environmental stimuli (Klemm et al., 2019).

Over recent years, a multitude of different histone PTMs have been identified across a range of organisms. The enzymes responsible for these modifications have typically classified into three distinct types: (i) ‘writers’ that introduce specific chemical modifications to histone proteins; (ii) ‘erasers’ that remove these modifications; and (iii) ‘readers’ that initially detect specific changes and then guide writers, erasers, or other proteins to the relevant genomic sites (Ueda & Seki, 2020). Evolutionary analyses suggest that these enzymes originated early and underwent significant expansion throughout the history of eukaryotes. An examination of 94 gene families involved in histone PTMs shows that 87 of them originated over a billion years ago, with 48 already existing in the last common ancestor of all eukaryotes (Weiner et al., 2020). Histone methylation and acetylation are the most common and abundant histone PTMs found in a wide range of organisms, while other histone PTMs such as acylation, phosphorylation, ubiquitination, and sumoylation have only been deeply studied in model organisms such as humans and the baker’s yeast *Saccharomyces cerevisiae* (Tan et al., 2011). However, besides the few model species, we have little knowledge of histone PTMs in less-studied fungal species.

The functional implication of different histone PTMs on chromatin organization and functioning has been derived from data in only a few eukaryotic systems. Euchromatin is a lightly packed form of chromatin that is enriched in gene-rich regions of the genome, characterized by its accessibility to TFs and RNA polymerases (K. Yang et al., 2022). This open chromatin structure facilitates active transcription, allowing for accurate and quick gene expression and regulatory processes essential for normal cellular function and response to environmental stimuli (K. Yang et al., 2022). It is predominantly associated with histone acetylation, such as acetylation of amino acid residues on histone protein H3 (lysine 4 (K4), K9, K14, K18, K23, and K27) and H4 (K5, K8, K12, and K16), and with methylations on histone protein H3 at K4, K36, and K79 (Bhaumik et al., 2007; Shilatifard, 2006). Distinct histone modifications have been observed on different regions of genes. In *S. cerevisiae*, H3K4ac (acetylation on histone protein H3 at lysine 4) is positioned upstream of protein-coding genes, while H3K4me3 (trimethylations on histone protein H3 at lysine 4), H3K9ac, and H3K14ac are enriched at the transcriptional start sites of active genes (Pokholok et al., 2005). Moreover, H3K4me2 (dimethylation) is most enriched throughout the coding region, while H3K4me1 (monomethylation), H3K36me3, and H3K79me3 are found predominantly at the end of active genes in *S. cerevisiae* (Pokholok et al., 2005). Conversely, constitutive heterochromatin, which describes chromatin that remains consistently condensed in an organism's lifespan, is situated at centromeres, telomeres, and regions typically enriched in transposons, thus contributing to the maintenance of genomic integrity and stability (Tamaru, 2010). H3K9me3

is associated with the formation and maintenance of constitutive heterochromatin in several fungal species, including the fission yeast *Schizosaccharomyces pombe* (Yamada et al., 2005) and the filamentous model fungus *Neurospora crassa* (Rountree & Selker, 2010). Facultative heterochromatin is chromatin that is thought to be able to dynamically transit between euchromatin and constitutive heterochromatin in response to developmental and environmental cues and has been linked with the presence of H3K27me3 (Grewal & Jia, 2007). In many filamentous fungi, this histone PTM is found near sub-telomeric regions as well as in regions rich in transposable elements (Ridenour et al., 2020; Schotanus et al., 2015). Last but not least, even though numerous PTMs have been linked to either euchromatin or heterochromatin, their impact can vary based on the regions they are located. For example, in *N. crassa* and *Fusarium fujikuroi*, H3K36me1/2/3 catalyzed by SET-2 activates gene expression, while H3K36me1/2/3 deposited by Ash-1 is located in sub-telomeric regions and is linked to gene silencing (Bicocca et al., 2018; Janevska et al., 2018).

In several fungal species, it has been demonstrated that experimentally induced changes to histone PTMs can influence the biosynthesis of SMs (K. Yang et al., 2022). For example, the deletion of a histone deacetylase changes the conformation of the chromatin from the transcriptionally silent 'closed' to the transcriptionally 'open' state, which increases the expression of nine out of 14 NRPSs in *Aspergillus fumigatus* (Lee et al., 2009) and the production of kojic acid in *Aspergillus niger* (Li et al., 2019). By contrast, the knockout of histone acetyltransferase can lead to defects in secondary metabolism in *A. nidulans* (Cánovas et al., 2014). Similar observations have also been made for histone methylations linked with euchromatin. For example, H3K36me3 can activate the transcription of genes involved in fumonisin B1 and bikaverin biosyntheses in *Fusarium verticillioides* (Gu, Wang, et al., 2017). Knocking out the catalytic enzyme of H3K4m3 leads to a significantly decreased production of deoxynivalenol, fusarin C, zearalenone, and fusarielin H in *Fusarium graminearum* (Bachleitner et al., 2019) and decreased biosynthesis of gibberellic acid in *F. fujikuroi* (Studt et al., 2017). In contrast, knocking out enzymes related to histone methylations of (facultative or constitutive) heterochromatin can induce the biosynthesis of SMs. For instance, the deletion of the writer for H3K27me3 up-regulates 20 out of 47 predicted SM genes in *F. fujikuroi* (Studt et al., 2016). Similarly, H3K9me3 mutant strains exhibit increased black pigmentation and fumonisin B1 production in *F. verticillioides* (Gu, Ji, et al., 2017). Thus, considering the abundant and diverse SM gene clusters encoded in fungal species, and the regulation of gene expression by various histone PTMs, there is significant potential to unlock more SM biosynthesis pathways.

The *Aspergillus* genus

The fungal genus *Aspergillus* encompasses approximately 350 species that display a wide range of morphological variation and ecological flexibility (Samson et al., 2014). Some

Aspergillus species can be detrimental by acting as pathogens to humans, animals, and crops, or as food contaminants (Lamoth et al., 2015). For instance, *A. fumigatus* leads to severe infections in individuals with weakened immune systems (Lamoth et al., 2015). Some species are vital for industrial and medical applications; for example, *A. niger* produces citric acid, the market for which is expected to reach three billion dollars by 2027 (Cairns et al., 2021; Mead et al., 2021). Importantly, species belonging to the *Aspergillus* genus are thought to be highly bioactive and can produce a variety of diverse SMs (Kjærbølling et al., 2018; Mosunova et al., 2021). Computational research spanning the fungal kingdom underscores the *Aspergillus* genus' rich reservoir of predicted BGCs (Figure 2) (Mosunova et al., 2021). A comprehensive study of 32 Aspergilli from the *Nigri* section revealed over 400 predicted BGCs. Intriguingly, 82% of these BGCs are identified in less than ten species, with 49% being unique to a single individual species (Vesth et al., 2018). Despite the vast number and variety of BGCs encoded in Aspergilli genomes, only a few hundred SMs have been experimentally identified. This discrepancy arises because these compounds are produced only under specific ecological conditions, making it difficult to express BGCs in a lab setting (Brakhage, 2013).

Several histone PTMs have been demonstrated to affect the SM biosynthesis in a handful of model *Aspergillus* species. In addition to the examples mentioned above, other studies have highlighted the impact of various histone PTMs on SM biosynthesis. For example, acetylation of histone H3 is required for the biosynthesis of SMs in *A. nidulans* (Cánovas et al., 2014; Reyes-Dominguez et al., 2010). In *A. fumigatus*, mutants lacking H3K4me2/3 show increased production of several SMs, including gliotoxin (Palmer et al., 2013). Importantly, research thus far in *Aspergillus* as well as in other fungi has primarily focused on the alterations of a few well-established histone PTMs. However, there is a notable absence of a comprehensive overview regarding the prevalence, evolution, and regulatory role of histone PTMs in the *Aspergillus* genus, especially concerning SM biosynthesis. In this context, an in-depth exploration of both histone PTMs and SM-BGCs needs to be performed. The *Aspergillus* genus is an excellent system considering its significant position as the prolific producer of SMs. Such an investigation could pave the way for a more profound understanding of the regulatory mechanisms of SM biosynthesis, potentially revealing novel avenues to express and exploit novel biochemistry, potentially with benefit for humankind.

To study SM biosynthesis and unravel the roles of PTMs in Aspergilli, **in this thesis**, I perform both wet-lab experiments and computational analyses. **In chapter 2,** I provide an extensive overview of BCG abundance and diversity across Aspergilli. Subsequently, I examine the genomic localization of these BGCs and how it relates to variation in gene expression and histone modifications. To get an overview of histone PTMs in the *Aspergillus* genus, I summarize the genes related to histone PTMs from the literature and check their absence and presence patterns and evolution across the *Aspergillus* genus in **chapter 3**. To verify the

computational analysis, I optimize a protocol to extract high-quality histone proteins and then conduct a proteomics study to detect the occurrence of single and multiple histone PTMs in three Aspergilli in **chapter 4**. Lastly, in **chapter 5**, I introduce the concept of 'histone code' and summarize its condition in Aspergilli. I pinpoint the interesting histone PTMs that can be further studied based on my studies and predict their functions in SM biosynthesis.

References

Aly, A. H., Debbab, A. & Proksch, P. (2011). Fifty years of drug discovery from fungi. *Fungal Diversity*, *50*, 3–19.

Bachleitner, S., Sørensen, J. L., Gacek-Matthews, A., Sulyok, M., Studt, L. & Strauss, J. (2019). Evidence of a demethylase-independent role for the H3K4-specific histone demethylases in *Aspergillus nidulans* and *Fusarium graminearum* secondary metabolism. *Frontiers in Microbiology*, *10*, 1759.

Balba, H. (2007). Review of strobilurin fungicide chemicals. *Journal of Environmental Science and Health Part B*, *42*(4), 441–451.

Bayram, Ö., Krappmann, S., Ni, M., Jin, W. B., Helmstaedt, K., Valerius, O., Braus-Stromeyer, S., Kwon, N. J., Keller, N. P., Yu, J. H. & Braus, G. H. (2008). VelB/VeA/LaeA complex coordinates light signal with fungal development and secondary metabolism. *Science*, *320*(5882), 1504–1506.

Bergmann, S., Schümann, J., Scherlach, K., Lange, C., Brakhage, A. A. & Hertweck, C. (2007). Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. *Nature Chemical Biology*, *3*(4), 213–217.

Bhaumik, S. R., Smith, E. & Shilatifard, A. (2007). Covalent modifications of histones during development and disease pathogenesis. *Nature Structural & Molecular Biology*, *14*(11), 1008–1016.

Bicocca, V. T., Ormsby, T., Adhvaryu, K. K., Honda, S. & Selker, E. U. (2018). ASH1-catalyzed H3K36 methylation drives gene repression and marks H3K27me2/3-competent chromatin. *ELife*, *7*, 1–19.

Brakhage, A. A. (2013). Regulation of fungal secondary metabolism. *Nature Reviews Microbiology*, *11*(1), 21–32.

Cairns, T. C., Barthel, L. & Meyer, V. (2021). Something old, something new: challenges and developments in *Aspergillus niger* biotechnology. *Essays in Biochemistry*, *65*(2), 213–224.

Cánovas, D., Marcos, A. T., Gacek, A., Ramos, M. S., Gutiérrez, G., Reyes-Domínguez, Y. & Strauss, J. (2014). The histone acetyltransferase GcnE (GCN5) plays a central role in the regulation of *Aspergillus* asexual development. *Genetics*, *197*(4), 1175–1189.

Chiang, Y.-M., Szewczyk, E., Davidson, A. D., Keller, N., Oakley, B. R. & Wang, C. C. C. (2009). A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in *Aspergillus nidulans*. *Journal of the American Chemical Society*, *131*(8), 2965–2970.

Choi, J. & Kim, S.-H. (2017). A genome tree of life for the fungi kingdom. *Proceedings of the National Academy of Sciences*, *114*(35), 9391–9396.

Collemare, J. & Lebrun, M. (2011). Fungal secondary metabolites: ancient toxins and novel effectors in plant–microbe interactions. *Effectors in Plant-Microbe Interactions*, 377–400.

De Serrano, L. O., Camper, A. K. & Richards, A. M. (2016). An overview of siderophores for iron acquisition in microorganisms living in the extreme. *Biometals*, *29*, 551–571.

Gomez-Escribano, J. P. & Bibb, M. J. (2011). Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microbial Biotechnology*, *4*(2), 207–215.

Grau, M. F., Entwistle, R., Oakley, C. E., Wang, C. C. C. & Oakley, B. R. (2019). Overexpression of an LaeA-like methyltransferase upregulates secondary metabolite production in *Aspergillus nidulans*. *ACS Chemical Biology*, *14*(7), 1643–1651.

Grewal, S. I. S. & Jia, S. (2007). Heterochromatin revisited. *Nature Reviews Genetics*, *8*(1), 35–

46.

Gu, Q., Ji, T., Sun, X., Huang, H., Zhang, H., Lu, X., Wu, L., Huo, R., Wu, H. & Gao, X. (2017). Histone H3 lysine 9 methyltransferase FvDim5 regulates fungal development, pathogenicity and osmotic stress responses in *Fusarium verticillioides*. *FEMS Microbiology Letters*, *364*(19), fnx184.

Gu, Q., Wang, Z., Sun, X., Ji, T., Huang, H., Yang, Y., Zhang, H., Tahir, H. A. S., Wu, L. & Wu, H. (2017). FvSet2 regulates fungal growth, pathogenicity, and secondary metabolism in *Fusarium verticillioides*. *Fungal Genetics and Biology*, *107*, 24–30.

Henikoff, S. & Shilatifard, A. (2011). Histone modification: cause or cog? *Trends in Genetics*, *27*(10), 389–396.

Hoffmeister, D. & Keller, N. P. (2007). Natural products of filamentous fungi: enzymes, genes, and their regulation. *Natural Product Reports*, *24*(2), 393–416.

Hortschansky, P., Haas, H., Huber, E. M., Groll, M. & Brakhage, A. A. (2017). The CCAAT-binding complex (CBC) in *Aspergillus* species. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, *1860*(5), 560–570.

Jakubiec-Krzesniak, K., Rajnisz-Mateusiak, A., Guspiel, A., Ziemska, J. & Solecka, J. (2018). Secondary metabolites of actinomycetes and their antibacterial, antifungal and antiviral properties. *Polish Journal of Microbiology*, *67*(3), 259–272.

Janevska, S., Baumann, L., Sieber, C. M. K., Münsterkötter, M., Ulrich, J., Kämper, J., Güldener, U. & Tudzynski, B. (2018). Elucidation of the two H3K36me3 histone methyltransferases Set2 and Ash1 in *Fusarium fujikuroi* unravels their different chromosomal targets and a major impact of Ash1 on genome stability. *Genetics*, *208*, 153–171.

Keller, N. P. (2019). Fungal secondary metabolism: regulation, function and drug discovery. *Nature Reviews Microbiology*, *17*(3), 167–180.

Keller, N. P., Turner, G. & Bennett, J. W. (2005). Fungal secondary metabolism - From biochemistry to genomics. *Nature Reviews Microbiology*, *3*(12), 937–947.

Khan, A. J., Bradley, E., O'Donovan, A. & Woolley, J. (2022). Psilocybin for trauma-related disorders. In *Disruptive Psychopharmacology* (pp. 319–332). Springer.

Kilani-Morakchi, S., Morakchi-Goudjil, H. & Sifi, K. (2021). Azadirachtin-based insecticide: Overview, risk assessments, and future directions. *Frontiers in Agronomy*, *3*, 676208.

Kjærbølling, I., Vesth, T. C., Frisvad, J. C., Nybo, J. L., Theobald, S., Kuo, A., Bowyer, P., Matsuda, Y., Mondo, S., Lyhne, E. K., Kogle, M. E., Clum, A., Lipzen, A., Salamov, A., Ngan, C. Y., Daum, C., Chiniquy, J., Barry, K., LaButti, K., … Andersen, M. R. (2018). Linking secondary metabolites to gene clusters through genome sequencing of six diverse *Aspergillus* species. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(4), E753–E761.

Klemm, S. L., Shipony, Z. & Greenleaf, W. J. (2019). Chromatin accessibility and the regulatory epigenome. *Nature Reviews Genetics*, *20*(4), 207–220.

Lamoth, F., Juvvadi, P. R. & Steinbach, W. J. (2015). Histone deacetylase inhibition as an alternative strategy against invasive aspergillosis. *Frontiers in Microbiology*, *6*(FEB), 4–9.

Laupacis, A., Keown, P. A., Ulan, R. A., McKenzie, N. & Stiller, C. R. (1982). Cyclosporin A: a powerful immunosuppressant. *Canadian Medical Association Journal*, *126*(9), 1041.

Lee, I., Oh, J. H., Keats Shwab, E., Dagenais, T. R. T., Andes, D. & Keller, N. P. (2009). HdaA, a class 2 histone deacetylase of *Aspergillus fumigatus*, affects germination and secondary metabolite production. *Fungal Genetics and Biology*, *46*(10), 782–790.

Li, X., Pan, L., Wang, B. & Pan, L. (2019). The histone deacetylases HosA and HdaA affect the phenotype and transcriptomic and metabolic profiles of *Aspergillus niger*. *Toxins*, *11*(9).

Liu, L., Chen, Z., Liu, W., Ke, X., Tian, X. & Chu, J. (2022). Cephalosporin C biosynthesis and fermentation in *Acremonium chrysogenum*. *Applied Microbiology and Biotechnology*, *106*(19–20), 6413–6426.

Louwen, J. J. R., Medema, M. H. & van der Hooft, J. J. J. (2023). Enhanced correlation-based linking of biosynthetic gene clusters to their metabolic products through chemical class matching. *Microbiome*, *11*(1), 1–12.

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. (1997). Crystal structure of the nucleosome resolution core particle at 2.8 Å. *Nature*, *389*, 251–260.

Lyu, H.-N., Liu, H.-W., Keller, N. P. & Yin, W.-B. (2020). Harnessing diverse transcriptional regulators for natural product discovery in fungi. *Natural Product Reports*, *37*(1), 6–16.

Macheleidt, J., Mattern, D. J., Fischer, J., Netzker, T., Weber, J., Schroeckh, V., Valiante, V. & Brakhage, A. A. (2016). Regulation and role of fungal secondary metabolites. *Annual Review of Genetics*, *50*, 371–392.

McGary, K. L., Slot, J. C. & Rokas, A. (2013). Physical linkage of metabolic genes in fungi is an adaptation against the accumulation of toxic intermediate compounds. *Proceedings of the National Academy of Sciences*, *110*(28), 11481–11486.

Mead, M. E., Steenwyk, J. L., Silva, L. P., De Castro, P. A., Saeed, N., Hillmann, F., Goldman, G. H. & Rokas, A. (2021). An evolutionary genomic approach reveals both conserved and species-specific genetic elements related to human disease in closely related *Aspergillus* fungi. *Genetics*, *218*(2).

Menon, V. P. & Sudheer, A. R. (2007). Antioxidant and anti-inflammatory properties of curcumin. *The Molecular Targets and Therapeutic Uses of Curcumin in Health and Disease*, 105–125.

Morales-Oyervides, L., Ruiz-Sánchez, J. P., Oliveira, J. C., Sousa-Gallagher, M. J., Méndez-Zavala, A., Giuffrida, D., Dufossé, L. & Montañez, J. (2020). Biotechnological approaches for the production of natural colorants by *Talaromyces*/*Penicillium*: A review. *Biotechnology Advances*, *43*, 107601.

Mosunova, O., Navarro-Muñoz, J. C. & Collemare, J. (2021). The biosynthesis of fungal secondary metabolites: From fundamentals to biotechnological applications. In *Encyclopedia of mycology* (pp. 458–476). Elsevier.

Newman, D. J. & Cragg, G. M. (2016). Natural products as sources of new drugs from 1981 to 2014. *Journal of Natural Products*, *79*(3), 629–661.

Palmer, J. M., Bok, J. W., Lee, S., Dagenais, T. R. T., Andes, D. R., Kontoyiannis, D. P. & Keller, N. P. (2013). Loss of CclA, required for histone 3 lysine 4 methylation, decreases growth but increases secondary metabolite production in *Aspergillus fumigatus*. *PeerJ*, *2013*(1), 1–18.

Pfannenstiel, B. T. & Keller, N. P. (2019). On top of biosynthetic gene clusters: How epigenetic machinery influences secondary metabolism in fungi. *Biotechnology Advances*, *37*(6), 107345.

Pham, J. V, Yilma, M. A., Feliz, A., Majid, M. T., Maffetone, N., Walker, J. R., Kim, E., Cho, H. J., Reynolds, J. M. & Song, M. C. (2019). A review of the microbial production of bioactive natural products and biologics. *Frontiers in Microbiology*, *10*, 1404.

Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Tong, I. L., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D. K. & Young, R. A. (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell*, *122*(4), 517–527.

Reyes-Dominguez, Y., Bok, J. W., Berger, H., Shwab, E. K., Basheer, A., Gallmetzer, A., Scazzocchio, C., Keller, N. & Strauss, J. (2010). Heterochromatic marks are associated with the repression of secondary metabolism clusters in *Aspergillus nidulans*. *Molecular Microbiology*, *76*(6), 1376–1386.

Ridenour, J. B., Möller, M. & Freitag, M. (2020). Polycomb repression without bristles: Facultative heterochromatin and genome stability in fungi. *Genes*, *11*(6), 1–23.

Rokas, A., Wisecaver, J. H. & Lind, A. L. (2018). The birth, evolution and death of metabolic gene clusters in fungi. *Nature Reviews Microbiology*, *16*(12), 731–744.

Rountree, M. R. & Selker, E. U. (2010). DNA methylation and the formation of heterochromatin in *Neurospora crassa*. *Heredity*, *105*(1), 38–44.

Samson, R. A., Visagie, C. M., Houbraken, J., Hong, S. B., Hubka, V., Klaassen, C. H. W., Perrone, G., Seifert, K. A., Susca, A., Tanney, J. B., Varga, J., Kocsubé, S., Szigeti, G., Yaguchi, T. & Frisvad, J. C. (2014). Phylogeny, identification and nomenclature of the genus *Aspergillus*. *Studies in Mycology*, *78*(1), 141–173.

Sanchez-Ballesta, M. T., Maoz, I. & Figueroa, C. R. (2022). Secondary metabolism and fruit quality. *Frontiers in Plant Science*, *13*, 1072193.

Sanchez, S. & Demain, A. L. (2008). Metabolic regulation and overproduction of primary metabolites. *Microbial Biotechnology*, *1*(4), 283–319.

Schotanus, K., Soyer, J. L., Connolly, L. R., Grandaubert, J., Happel, P., Smith, K. M., Freitag, M. & Stukenbrock, E. H. (2015). Histone modifications rather than the novel regional centromeres of *Zymoseptoria tritici* distinguish core and accessory chromosomes. *Epigenetics & Chromatin*, *8*, 1–18.

Sharma, S. & Meyer, V. (2022). The colors of life: an interdisciplinary artist-in-residence project to research fungal pigments as a gateway to empathy and understanding of microbial life. *Fungal Biology and Biotechnology*, *9*, 1–11.

Shilatifard, A. (2006). Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu. Rev. Biochem.*, *75*, 243–269.

Studt, L., Janevska, S., Arndt, B., Boedi, S., Sulyok, M., Humpf, H. U., Tudzynski, B. & Strauss, J. (2017). Lack of the COMPASS component Ccl1 reduces H3K4 trimethylation levels and affects transcription of secondary metabolite genes in two plant-pathogenic *Fusarium* species. *Frontiers in Microbiology*, *7*(JAN), 1–17.

Studt, L., Rösler, S. M., Burkhardt, I., Arndt, B., Freitag, M., Humpf, H., Dickschat, J. S. & 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*. *Environmental Microbiology*, *18*(11), 4037–4054.

Tamaru, H. (2010). Confining euchromatin/heterochromatin territory: Jumonji crosses the line. *Genes and Development*, *24*(14), 1465–1478.

Tan, M., Luo, H., Lee, S., Jin, F., Yang, J. S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., Lu, Z., Ye, Z., Zhu, Q., Wysocka, J., Ye, Y., Khochbin, S., Ren, B. & Zhao, Y. (2011). Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell*, *146*(6), 1016–1028.

Tudzynski, B, Homann, V., Feng, B. & Marzluf, G. A. (1999). Isolation, characterization and disruption of the areA nitrogen regulatory gene of *Gibberella fujikuroi*. *Molecular and General Genetics MGG*, *261*, 106–114.

Tudzynski, Bettina. (2014). Nitrogen regulation of fungal secondary metabolism in fungi. *Frontiers in Microbiology*, *5*, 656.

Ueda, M. & Seki, M. (2020). Histone modifications form epigenetic regulatory networks to regulate abiotic stress response. *Plant Physiology*, *182*(1), 15–26.

van Holde, K. E. (1989). Chromatin structure and transcription. In *Chromatin* (pp. 355–408). Springer.

Velusamy, P., Immanuel, J. E., Gnanamanickam, S. S. & Thomashow, L. (2006). Biological control

of rice bacterial blight by plant-associated bacteria producing 2, 4-diacetylphloroglucinol. *Canadian Journal of Microbiology*, *52*(1), 56–65.

Vesth, T. C., Nybo, J. L., Theobald, S., Frisvad, J. C., Larsen, T. O., Nielsen, K. F., Hoof, J. B., Brandl, J., Salamov, A., Riley, R., Gladden, J. M., Phatale, P., Nielsen, M. T., Lyhne, E. K., Kogle, M. E., Strasser, K., McDonnell, E., Barry, K., Clum, A., … Andersen, M. R. (2018). Investigation of inter- and intraspecies variation through genome sequencing of *Aspergillus* section *Nigri*. *Nature Genetics*, *50*(12), 1688–1695.

Weiner, A. K. M., Cerón-Romero, M. A., Yan, Y. & Katz, L. A. (2020). Phylogenomics of the epigenetic toolkit reveals punctate retention of genes across eukaryotes. *Genome Biology and Evolution*, 1–39.

Wisecaver, J. H., Slot, J. C. & Rokas, A. (2014). The evolution of fungal metabolic pathways. *PLoS Genetics*, *10*(12), e1004816.

Woloshuk, C. P., Foutz, K. R., Brewer, J. F., Bhatnagar, D., Cleveland, T. E. & Payne, G. (1994). Molecular characterization of aflR, a regulatory locus for aflatoxin biosynthesis. *Applied and Environmental Microbiology*, *60*(7), 2408–2414.

Yamada, T., Fischle, W., Sugiyama, T., Allis, C. D. & Grewal, S. I. S. (2005). The nucleation and maintenance of heterochromatin by a histone deacetylase in fission yeast. *Molecular Cell*, *20*(2), 173–185.

Yang, K., Tian, J. & Keller, N. P. (2022). Post-translational modifications drive secondary metabolite biosynthesis in *Aspergillus*: a review. *Environmental Microbiology*.

Yang, X., Feng, P., Yin, Y., Bushley, K., Spatafora, J. W. & Wang, C. (2018). Cyclosporine biosynthesis in *Tolypocladium inflatum* benefits fungal adaptation to the environment. *MBio*, *9*(5), e01211-18.

Zeidler, S. & Müller, V. (2019). Coping with low water activities and osmotic stress in *Acinetobacter baumannii*: significance, current status and perspectives. *Environmental Microbiology*, *21*(7), 2212–2230.

Chapter 2

Secondary metabolite gene clusters and their genomic localization in the fungal genus *Aspergillus*

Xin Zhang, Iseult Leahy, Jérôme Collemare+, Michael F. Seidl+

+These authors contributed equally

Abstract

Fungi are well-known producers of bioactive secondary metabolites (SMs), which have been exploited for centuries by humankind for various medical applications like therapeutics and antibiotics. SMs are synthesized by biosynthetic gene clusters (BGCs) – physically co-localized and co-regulated genes. Because BGCs are often regulated by histone post-translational modifications (PTMs), it was suggested that their chromosomal location is important for their expression. Studies in a few fungal species indicated an enrichment of BGCs in sub-telomeric regions, however, there is no evidence that BGCs with distinct genomic localization are regulated by different histone PTMs. Here, we used 174 *Aspergillus* species covering 22 sections to determine the correlation between BGC genomic localization, gene expression, and histone PTMs. We found a high abundance and diversity of SM backbone genes across the *Aspergillus* genus, with notable diversity increases between sections. Being unique or conserved in many species, BGCs showed a strong bias for being localized in low-synteny regions, regardless of their position in chromosomes. Using chromosome-level assemblies, we also confirmed a significantly biased localization in sub-telomeric regions. Notably, SM backbone genes in sub-telomeric regions and about half of those in low-synteny regions exhibit higher gene expression variability, likely due to the similar higher variability in H3K4me3 and H3K36me3 histone PTMs. In contrast, variations in histone H3 acetylation and H3K9me3 are not correlated to genomic localization and expression variation. Overall, our results indicate that BGCs tend to be located in low-synteny regions and that regulation of expression in those regions likely involves different histone PTMs than the most commonly studied modifications.

Introduction

Filamentous fungi are prolific producers of bioactive molecules, termed secondary metabolites (SMs) or natural products (Mosunova et al., 2021). While these metabolites are not directly necessary to sustain fungal growth, development, or reproduction, they play indispensable roles in inter-organismal interactions and fungal survival in diverse ecological niches (Keller, 2019). SMs have multifaceted functions, from acting as pigments that shield fungal spores from UV radiation and serving as virulence factors in pathogenic interactions to providing competitive edges through antibiotic properties (Keller, 2019). Humans have harnessed the bioactive potential of these metabolites to significantly enhance their quality of life (Cairns et al., 2021). Fungal SMs have found significant applications in medical treatments, serving as therapeutic agents, exemplified by lovastatin and ergotamine, as well as immunosuppressants like cyclosporine (Newman & Cragg, 2016; Wiemann & Keller, 2014). They also play crucial roles as antibiotics, with penicillin being a notable example, and as antifungal agents, such as griseofulvin (Newman & Cragg, 2016).

The biosynthesis of fungal SMs is orchestrated by a series of enzymatic reactions. First, backbone (or core) enzymes use precursors derived from primary metabolism to synthesize an intermediate that is then further modified by so-called tailoring enzymes to produce the final compound(s) (Mosunova et al., 2021). Backbone enzymes responsible for the synthesis of diverse SMs have been classified into categories based on their enzymatic functions such as polyketide synthase (PKS), non-ribosomal peptide synthetase (NRPS), hybrid PKS-NRPS enzymes, dimethylallyl tryptophan synthetase (DMATS), and terpene cyclase (TC) (Mosunova et al., 2021). Genes encoding backbone and tailoring enzymes, as well as genes responsible for transcription control, transport, and/or self-resistance, are co-regulated and often physically co-localized along the chromosome in biosynthetic gene clusters (BGCs) (Mosunova et al., 2021).

The discrepancy between the predicted bioactive potential based on genome mining and the number of known molecules effectively produced can be explained by the tight transcriptional control of SM biosynthetic pathways (Brakhage, 2013). In addition to the regulation by local and global transcription factors (Chiang et al., 2009; Macheleidt et al., 2016), SM production is regulated globally by chromatin dynamics that, in response to both internal and external stimuli, transitions between different chromatin states: the ‘open’ euchromatin, loosely packed and transcriptionally active state, and the ‘closed’ heterochromatin, tightly packed and transcriptionally silent state (Collemare & Seidl, 2019; Gacek & Strauss, 2012). Histone post-translational modifications (PTMs) – chemical modifications on specific amino acids localized on the ‘tails’ of histone proteins contribute to the control of the chromatin status. Most histone acetylations and some histone methylations on H3K4 (histone protein H3 at lysine 4) and H3K36 are associated with euchromatin, while histone methylations on H3K9 and H3K27 are associated with heterochromatin (Lawrence et al., 2016). In *Aspergillus fumigatus*, the deletion

of histone deacetylases, which changes the chromatin from 'closed' heterochromatin to 'open' euchromatin, increases the activity of nine out of 14 NRPSs (Lee et al., 2009) and the production of kojic acid in *Aspergillus niger* (Li et al., 2019). The knockout of histone acetyltransferases also leads to defects in SM biosynthesis in *Aspergillus nidulans* (Cánovas et al., 2014).

Following the discovery that histone PTMs play an important role in the regulation of SM biosynthesis, it was hypothesized that the chromosomal location of BGCs was crucial for their regulation (Palmer and Keller, 2010). Indeed, several studies reported that BGCs tend to be preferentially located at, or close to, sub-telomeric regions, such as in *Fusarium graminearum* (Connolly et al., 2013), and in some well-studied Aspergilli: *Aspergillus oryzae* (Kjærbølling et al., 2020), *Aspergillus flavus* (Yang et al., 2022), *A. fumigatus* (Keller, 2019; Perrin et al., 2007), and *A. nidulans* (Klejnstrup et al., 2012). Sub-telomeric regions in fungi are rich in repetitive sequences such as transposable elements and are typically associated with histone PTMs linked to heterochromatin, especially with the presence of H3K9me3 and H3K27me3 (Keller, 2019; Palmer & Keller, 2010; Wiemann et al., 2013). For instance, in *Fusarium fujikuroi*, the PKS19 BGC, which produces fujikurins, is embedded within a region rich in H3K9me3 on the long arm of chromosome VIII (Wiemann et al., 2009). Similarly, in *A. nidulans*, some BGCs occur in proximity to H3K9me3-rich regions (Gacek-Matthews et al., 2016). Thus, the peculiar genomic localization of BGCs might explain why these BGCs remain transcriptionally silent. Moreover, because of their high repeat content, sub-telomeric regions are prone to rearrangements and typically exhibit low synteny, characterized by reduced conservation of physical co-localization of genes compared to other species (Nakao et al., 2009). In *A. nidulans*, compared to *A. oryzae* and *A. fumigatus*, BGCs often occur at less conserved and heavily rearranged genomic regions, even though the majority (77–79%) of the genomes show conserved synteny (Inglis et al., 2013). However, not all SM gene clusters are associated with heterochromatin and sub-telomeric regions. While previous studies have predominantly focused on a limited number of well-known *Aspergillus* sections and species, often finding an association of SM gene clusters with heterochromatin and sub-telomeric regions, it is crucial to note that not all SM gene clusters are confined to these areas. To address this knowledge gap, we examined the correlation among the BGC diversity, their genomic organizations, gene expression, and histone PTMs together across a more diverse range of species. We here include 174 diverse Aspergilli species including 22 sections to provide a more nuanced analysis of the distribution, conservation, and genomic localization of SM genes, particularly in relation to genome-wide syntenic regions and histone PTMs. This extensive approach allows us to explore the implications of BGC localization and histone PTMs on SM gene expression regulation.

Results and discussions

A comprehensive phylogenetic analysis of the fungal genus *Aspergillus*

To provide a comprehensive analysis of BGCs in Aspergilli, we included 174 publicly available Aspergilli genomes and three outgroup *Penicillium* species derived from the Joint Genome Institute (Table S1). The assembled genomes range in size from 23.2 to 42.9 Mb and 90% of them fall within the range of 25.8 to 38.8 Mb (Figure 1, Table S1). Four *Aspergillus* species - *A. fumigatus*, *A. nidulans*, *A. niger*, and *Aspergillus ochraceus* - have chromosomal-level genome assemblies (Nierman et al., 2005), which are comprised of eight chromosomes (Figure 1, Table S1). The other well-studied model system, *A. oryzae*, has also a good genome assembly of eleven scaffolds. Twenty-four Aspergilli assemblies have 100 or fewer scaffolds, while *Aspergillus rambellii* has the most fragmented assembly with 4,177 scaffolds (Figure 1, Table S1). The number of predicted protein-coding genes ranges from 7,761 to 15,687, with 90% of Aspergilli having between 9,561 and 14,061 genes (Figure 1, Table S1). Because the genome contiguity and the number of genes varied between different Aspergilli, we sought to calculate the genome completeness by querying each genome assembly for the occurrence of single-copy orthologous BUSCO genes (Table S2). Importantly, we observed very high genome completeness with 90% of all Aspergilli scoring higher than 98.0% BUSCO completeness (Figure 1, Table S1). Even *A. rambellii* with the most fragmented assembly and the fewest genes (7,761) attains a high genome completeness of 97.1%. Similarly, four other Aspergilli with more than 1,000 scaffolds also have high genome completeness (*Aspergillus udagawae* (97.8%), *Aspergillus avenaceus* (99.1%), *Aspergillus haitiensis* (98.3%), and *Aspergillus coremiiformis* (99.5%)) (Figure 1, Table S1). Thus, all species included in this study are of sufficient quality to perform accurate comparative analyses.

To provide a robust phylogenetic framework to study BGC and genomic diversity in Aspergilli, we constructed a species phylogeny based on the concatenated alignment of 758 conserved, single-copy BUSCO genes (in total 758 positions with 132,397 informative sites). The tree reveals 22 distinct *Aspergillus* sections, each supported by bootstrap values exceeding 99% (Figure 1), suggesting that our phylogenic analysis is robust. Moreover, our species phylogeny confirms the associations among distinct *Aspergillus* sections as reported in previous studies and enlarges the taxonomic breadth by incorporating a much larger set of diverse Aspergilli; we included 174 *Aspergillus* species compared to less than 90 Aspergilli in previous publications (Steenwyk et al., 2019; C M Visagie et al., 2023; Cobus Meyer Visagie et al., 2014; X. Zhang et al., 2022). The 22 sections exhibit notable variation in the number of sequenced species: section *Nigri* has 28 sequenced species, whereas sections *Janorum*, *Bispori*, and *Silvati* are each represented by only a single sequenced species (Figure 1). Despite these differences, the genome assemblies in most sections are relatively uniform concerning assembly length, number of scaffolds, genome completeness, and gene count, and thus provide a comprehensive and robust reference to study the diversity and evolution of Aspergilli and a solid framework for further analysis.

Aspergilli encode an abundant secondary metabolite biosynthetic gene cluster repertoire

To evaluate the potential for SM biosynthesis encoded in the genomes of *Aspergillus* species, we focused on SM backbone genes as these genes are crucial for the first step in SM biosynthesis (Keller, 2019). We used antiSMASH (Blin et al., 2021) to identify regions that comprise a predicted BGC, and employed an in-house BGCtoolkit to further classify SM backbone enzymes into groups that correspond to major chemical families and retrieve their sequences.

We identified in total 16,243 SM backbone genes across 174 *Aspergillus* species and three *Penicillium* species, ranging from 18 in *Aspergillus caninus* to up to 160 in *Aspergillus implicatus* (Figure 1, Table S3). On average, we identified 92 SM backbone genes per Aspergilli compared to only 77 per *Penicillium* species. These SM backbone genes are unevenly distributed. The 41 early-diverging Aspergilli (comprising sections *Polypaecilum*, *Cremei*, *Restricti*, *Aspergillus*, *Cervini*, *Clavati*, and *Fumigati*) have on average 65 SM backbone genes, while the other 133 Aspergilli have on average 99. The *Circumdati* section has the highest average number of SM backbone genes, with its 18 species encoding an average of 129 SM backbone genes. As one BGC can contain several SM backbone genes, the average number of SM backbone genes in Aspergilli is higher than the number of predicted BGCs that have been reported in previous studies. For instance, studies that focused on sections *Nigri* (Theobald et al., 2018) and *Flavi* (Kjærbølling et al., 2020) have detected on average 82 and 70 BGCs, respectively. Here, we found these sections have an average of 104 (section *Nigri*) and 101 (section *Flavi*) SM backbone genes. Five groups of SM backbone genes stood out due to their abundance in all genomes. NRPS-like with 3,888, rPKS with 3,175, NRPS with 2,943, nrPKS with 1,395, and DMATS with 1,203 backbone genes. In most of the species (145 in total, including three *Penicillium* species), NRPS-like, rPKS, and NRPS are the three most abundant SM backbone genes. In contrast, ten species in sections *Flavipedes* and *Fumigati* have highly abundant DMATS, and other 22 Aspergilli have also many nrPKSs. Interestingly, we also observed variation in the number of SM backbone genes within sections. For example, within the *Nigri* section, we could define two sub-groups based on their morphology, and biseriates tend to have more SM backbone genes encoded in their genomes compared to uniseriates; uniseriates and biseriates differ in the attachment of phialides to the vesicle (Nielsen et al., 2009). Similarly, two sub-groups in the *Fumigati* section show a diverse pattern of SM backbones, the clade to which *A. fumigatus* belongs has a higher number of SM backbone genes. Moreover, we also observed that often a section is out-grouped by one species with fewer genes, such as in sections *Circumdati*, *Flavi*, *Nigri*, and *Nidulantes*, suggesting that gene loss in these species may be a pivotal factor driving their evolutionary differentiation and niche specialization. In summary, our findings underscore that the SM repertoires tend to be consistent within a given section and vary between sections.

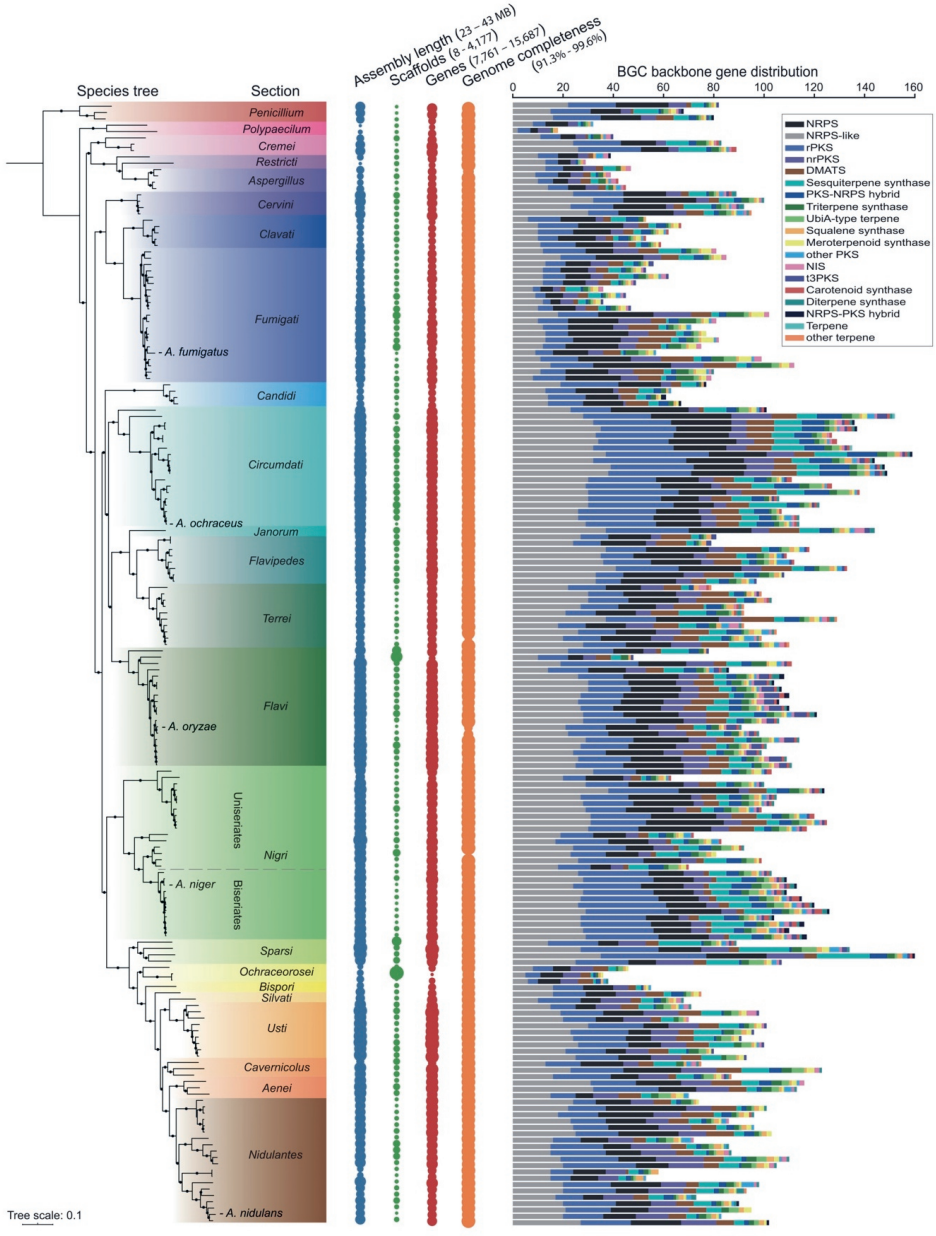


Figure 1 Aspergilli encode an extensive number of secondary metabolite (SM) backbone genes, with notable differences across sections. The phylogenetic relationships between 174 *Aspergillus* and three *Penicillium* species are built using a maximum-likelihood framework implemented in IQ-TREE using a concatenated alignment of 758 single-copy BUSCO genes. The dots on the branches indicate ultrafast bootstrap values over 95 and SH-aLRT ((Shimodaira-

Hasegawa-like approximate likelihood ratio test)) values over 80. Species from the same section are highlighted by the same background color. The names of five well-studied Aspergilli are shown in the tree and a dashed line highlights the separation of sub-groups within the *Nigri* section. The smallest and largest dot sizes correspond to the minimum and maximum number in each category, respectively. Genome completeness is indicated by the percentage of identified complete, single-copy BUSCO genes in each genome assembly. SM backbone genes are first detected with antiSMASH, then reclassified by BGCtoolkit. Bar plots representing different SM backbone groups are arranged from left to right from the most to the least abundant.

Previous studies used BGC network approaches to assess BGC conservation within *Aspergillus* sections (Kjærbølling et al., 2020; Vesth et al., 2018), which overestimates the BGC diversity as variation in gene content of the identified BGCs will directly result in non-conserved BGCs. In contrast, we here utilized a phylogeny-aware method (Figure S1) to determine orthologous groups for each type of SM backbone gene. To obtain an estimate of gene conservation in Aspergilli, we used Orthofinder to identify orthologous groups for all genes among all genomes. We included in total 1,546,732 orthologs defined by the Orthofinder (Table S4) and 16,638 SM backbone genes to test their uniqueness among Aspergilli. From 1,546,732 orthologous entries defined by the Orthofinder, we identified 34,144 unique genes (2.2%) among all species, of which the estimated growth rate β is 0.27 (Figure 2A), in which β represents the rate of growth of unique genes as the number of sampled species increases. For SM backbone genes, 2,426 out of 16,638 (14.6%) are unique, of which the growth rate β is 0.45 (Figure 2B). These β rates imply that sequencing more *Aspergillus* species is expected to expand the overall genomic diversity and specifically increase the diversity within the set of SM backbone genes. However, saturation is reached for many SM groups, but not for NRPS and NRPS-like, which indicates that we might have found nearly all polyketide and terpene biosynthetic pathways, but many non-ribosomal peptides likely remain to be discovered. Lastly, we observed a similar trend for both unique orthologues and SM backbone genes with continuous slight increases within sections and small jumps between sections (Figures 2C and 2D), suggesting that sections are indeed characterized by a sudden increase in BGC diversity that is consistent with divergence.

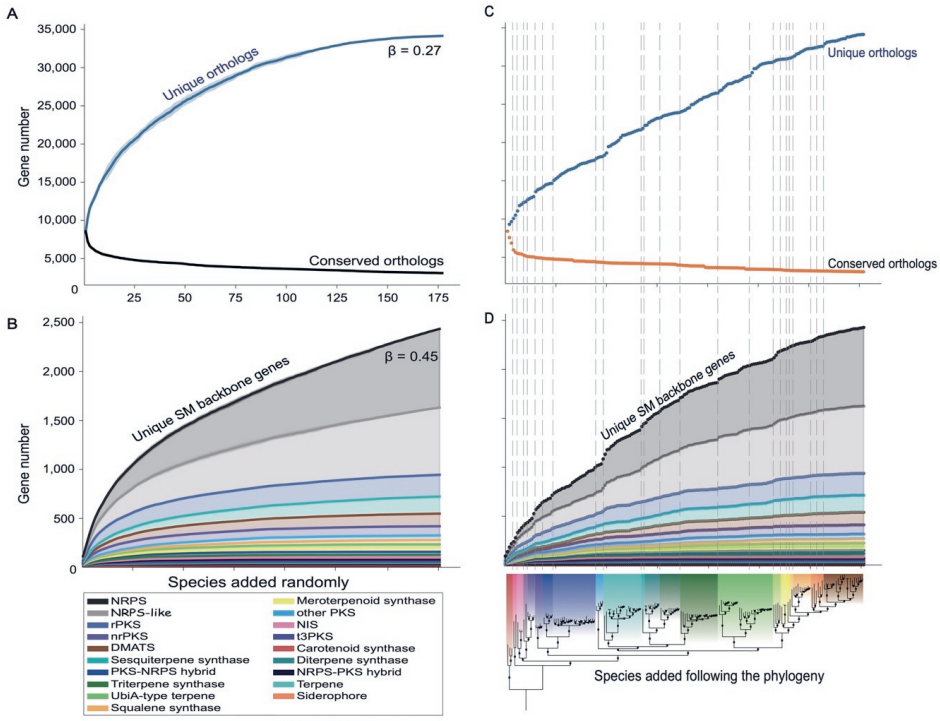


Figure 2 Aspergilli have an extensive repertoire of unique SM backbone genes. (A) The saturation curve plots the number of unique and conserved orthologous genes when species are added randomly. The calculations are performed ten times, and the standard deviations are shown as shadows around the curve. (B) Saturation curves plot the stacked number of unique SM backbone genes when species are added randomly. Different colors indicate distinct groups of SM backbone genes. The calculations are performed 10 times, and the standard deviations are shown as shadows around the curve. (C) Saturation curves of unique and conserved orthologous genes when species are added according to their taxonomy, starting with the outgroup, and the dash lines indicate distinct taxonomic sections. (D) Saturation curves of unique SM backbone genes with the species added according to their taxonomy and the dash lines indicate distinct taxonomic sections. A phylogenetic species tree is shown at the bottom for reference.

We observed that SM backbone genes are diverse and that increased diversity is observed across different *Aspergillus* sections. However, the number of unique genes is lower than estimated in two recent studies focused solely on the sections of *Nigri* (Vesth et al., 2018) or *Flavi* (Kjærbølling et al., 2020) that found around half of the SM BGCs were predicted to be species-specific. Here, we only focused on SM backbone genes and implemented a phylogeny-aware approach to evaluate SM backbone gene diversity across a diverse set of *Aspergillus* species, while these previous studies used complete SM gene cluster diversity

and relied on a limited taxonomic sampling. These previous studies thus overestimate the BGC diversity within *Aspergillus* sections. In contrast, our results are more conservative and underestimate the true biochemical diversity that is provided by the diversity of additional tailoring genes within BGCs that can lead to the production of complex secondary metabolites. However, our approach provides broader and more detailed estimations regarding the type of metabolite and BGCs that remain to be discovered in *Aspergillus* species.

Genome-wide synteny in Aspergilli is higher within than between sections

The tight regulation of BGC expression and the role of chromatin conformation in this regulation suggested that the genomic localization of BGCs might be important (Palmer et al., 2010). Several studies have indicated an over-representation of SM BGCs in sub-telomeric regions (Guzmán-Chávez et al., 2018; Keller, 2019; Kjærbølling et al., 2020; Klejnstrup et al., 2012). Sub-telomeric regions are known to be variable between strains and species, and consequently, these regions display low levels of synteny. The term 'synteny' here refers to the conservation of genomic regions that are maintained in a similar arrangement across different genomes through evolutionary time (Drillon et al., 2013) (Figure 3A). As most of the *Aspergillus* species do not have chromosome-level assemblies, we assessed BGC localization by examining synteny throughout the *Aspergillus* genus. We performed an all-versus-all pairwise comparison of all species, and we uncovered, in general, that the amount of shared synteny between species (synteny percentage) correlates with the divergence and is consistent with the phylogenetic relationship between species (Figure 3B); *i.e.*, two *Aspergillus* sections that are more syntenic indeed diverged more recently. Based on this observation, we propose that the *Nigri* section could be split in two, which is also coherent with the phylogeny. Similarly, the *Circumdati*, *Flavi*, *Nigri*, and *Nidulantes* sections show distinct groups of higher synteny that might indicate ongoing divergence within these sections. Section *Nidulantes* exhibits lower synteny, which seems coherent with the branch lengths in this clade compared to other clades where the divergence between species appears more recent in the *Nidulantes* section. Finally, early-diverging sections (*i.e., Polypaecilum*, *Cremei*, *Restricti*, *Aspergillus*, *Cervini*, *Clavati*, and *Fumigati*) show high synteny within sections compared with other sections (Figure 3C). Overall, the synteny between sections is low, ranging from only 34% to 54%, with a median of 46%.

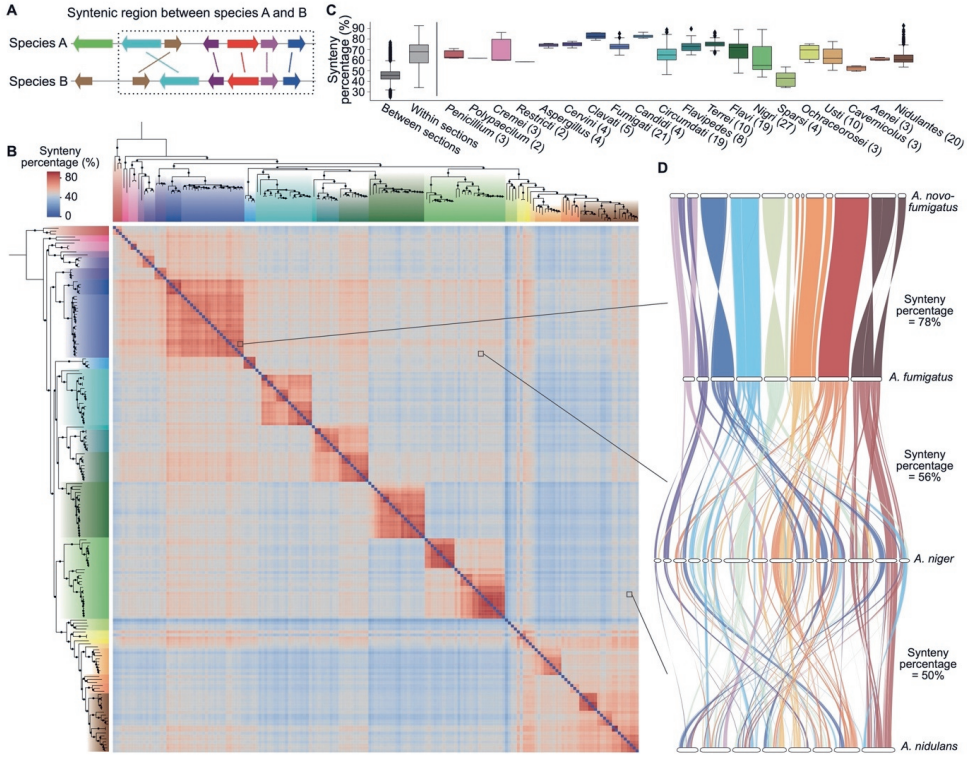


Figure 3 Within-section synteny in *Aspergillus* is higher than between-section synteny. (A) A graphical representation of a hypothetical syntenic region between two species is demarcated by a dashed box. Arrows of the same color, linked by a line, represent homologous genes from two different species, with the arrow direction indicating the gene's orientation. (B) The heatmap shows pairwise synteny expressed as the synteny percentage based on all-versus-all comparisons of all species, arranged based on the phylogenetic tree (Figure 1). The species tree, displayed to the left and top of the heatmap, uses colors to differentiate between sections (see legend for Figure 1A for the color coding). (C) Boxplots show the quantification of synteny by the 'synteny percentage' between sections, within sections, and in each section of the *Aspergillus* genus. Numbers in the parentheses indicate the number of species in the sections, and only sections with more than one species are shown. (D) Genome alignment between two species from the same section (top) and different sections (middle and bottom) highlights the degradation of synteny when taxonomic distance increases.

In Ascomycota, synteny, or the conservation of gene order, varies across distantly related species (Hane et al., 2011). This phenomenon is categorized into micro-synteny, meso-synteny, and macro-synteny (Hane et al., 2011). Micro-synteny refers to the conservation of a small number of successive genes, typically ranging from two to ten. Meso-synteny represents larger, yet fragmented, synteny blocks compared to micro-synteny but shorter than

20 thousand base pairs (kbp). Macro-synteny, on the other hand, is characterized by large and continuous blocks extending over 20 kbp (Hane et al., 2011). Our findings align with expectations, revealing longer synteny blocks within sections than between them (Figure S2). For instance, species in the *Fumigati* section show extensive regions of shared synteny across chromosomes (Figure 3D). In contrast, species from different sections exhibit markedly reduced synteny, both in percentage and block length (Figures 3D and S2). A pattern of 'degraded macrosynteny' was previously identified in several filamentous Ascomycota fungi, including some Aspergilli (Hane et al., 2011). Our data suggest that within *Aspergillus* sections, synteny blocks with an approximate length of 150 kbp are indicative of macrosynteny. In contrast, between different sections, synteny blocks averaging 70 kbp to 150 kbp represent degraded macrosynteny, likely reflecting the divergence time among these sections.

SM backbone genes tend to reside in low-synteny and sub-telomeric regions

To investigate the genomic localizations of SM genes in Aspergilli, we analyzed four species with chromosomal-level assemblies: *A. nidulans*, *A. fumigatus*, *A. niger*, and *A. oryzae* (Figures 4 and S3-S5). In each of the four species, the gene density shows that predicted genes are evenly distributed throughout the genome, except for putative centromere and telomere regions, which are known to contain fewer genes but harbor repetitive sequences and exhibit unique structural and functional characteristics distinct from the rest of the genome (Xu et al., 2022). BUSCO genes, which are conserved across the fungal kingdom and crucial for organismal function, are significantly enriched in regions with high synteny, both within individual sections (91% of BUSCO genes) and across the *Aspergillus* genus (83%) compared with all genes (59% and 42%, respectively) (Figure 5B and 5C). In contrast, SM backbone genes are significantly enriched in low-synteny regions in all four species at the *Aspergillus* genus level (81% of backbone genes) (Figures 5B and 5C, Table S6). For the section-level synteny scores, though only 24% of SM backbone genes are found in low-synteny regions, the trend remains significant compared with all genes (14%) and clearly differs from BUSCO genes (2%) (Figures 5B and 5C, Tables S6). When examining the synteny scores at the section level, we observed that while a subset of SM backbone genes is in regions with relatively high-synteny scores, the general pattern of these scores is notably distinct from that of regions with BUSCO genes. This contrast is evident when comparing the chromosomal distribution of synteny scores for SM backbone genes against those for BUSCO genes (Figures 5B and 5C; Table S6). Only 21% of all genes are located in sub-telomeric regions, and BUSCO genes are significantly located outside of these regions (7% in sub-telomeric regions) (Figures 5D and 5E). In contrast, SM backbone genes show a trend to be located in the sub-telomeric regions (36%) (Figures 5D and 5E), aligning with previous reports in *A. nidulans* (Klejnstrup et al., 2012), *A. fumigatus* (Keller, 2019; Perrin et al., 2007), *A. oryzae* (Kjærbølling et al., 2020), as well as other fungal species such as *F. fujikuroi* (Studt et al., 2017) and *Penicillium chrysogenum* (Guzmán-Chávez et al., 2018). However, the enrichment of SM backbone genes in sub-telomeric regions is less significant than in low-synteny regions,

suggesting that association with low-synteny regions is an attribute of SM BGCs. The localization of SM backbone genes in low-synteny and sub-telomeric regions of *Aspergillus* genomes may facilitate rapid genetic rearrangement of these BGCs, thereby enabling rapid adaptation to diverse ecological niches. Noteworthy, SM backbone genes with known metabolite structures, as listed in the MIBiG database (Table S5), tend to occur more frequently in sub-telomeric areas. For instance, four out of six SM backbone genes in *A. nidulans* (Figure 4), 18 out of 27 in *A. fumigatus* (Figure S3), and seven out of eleven in *A. niger* (Figure S4) are found in sub-telomeric regions. *A. oryzae* presents a unique case, with only two out of 18 SM backbone genes with known structures located in sub-telomeric regions (Figure S5). The observation that most characterized BGCs are localized in sub-telomeric regions suggests that they are more frequently expressed than those in other regions.

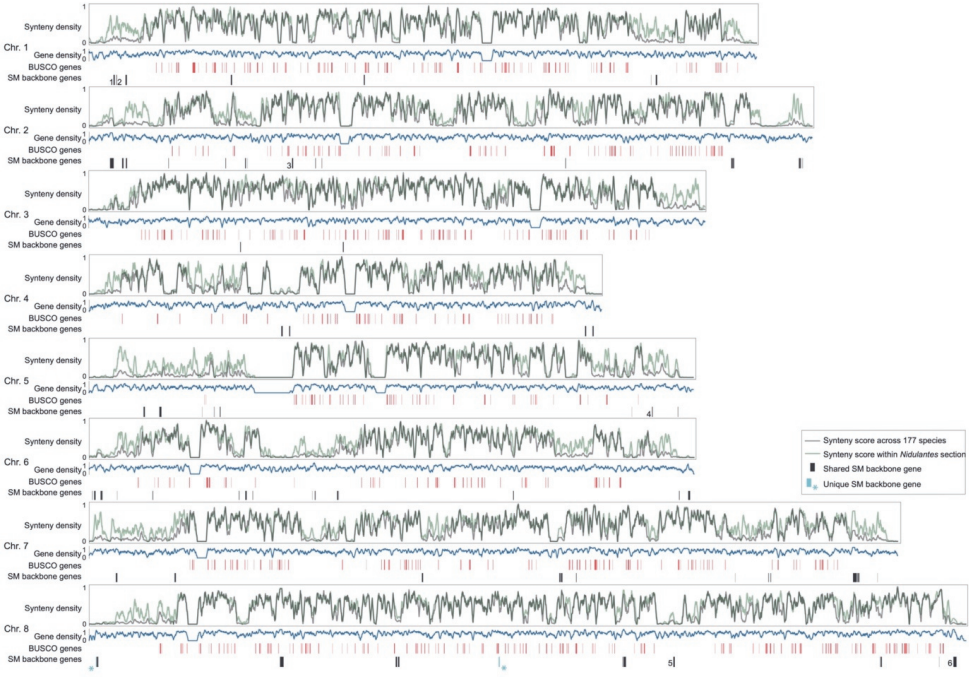
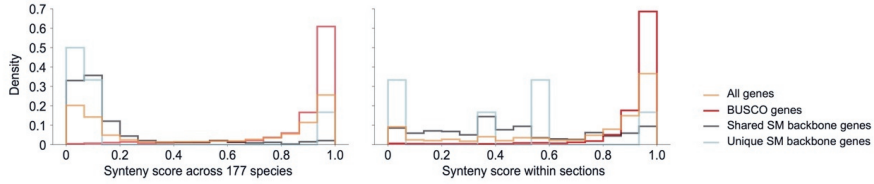


Figure 4 SM backbone genes occur preferentially in low-synteny and sub-telomeric regions. This figure illustrates the genomic localization of BUSCO genes and SM backbone genes in *A. nidulans*. Lines in black and green represent synteny density across 177 species and within the *Nidulantes* section, respectively. The gene density is shown along the chromosomes. Red bars indicate the chromosomal locations of BUSCO genes, with bar width proportional to gene length. Black and blue bars mark the positions of shared and unique SM backbone genes, respectively, with their widths corresponding to gene length. Numbers adjacent to the black bars denote SM backbone genes associated with known metabolites as listed on the MIBiG website (Table S5).

A

Genes / species	*A. nidulans*	*A. niger*	*A. fumigatus*	*A. oryzae*
All genes	10,680	11,846	9,781	12,030
BUSCO genes	745	756	745	711
Shared SM backbone genes	81	113	56	89
Unique SM backbone genes	2	1	1	2

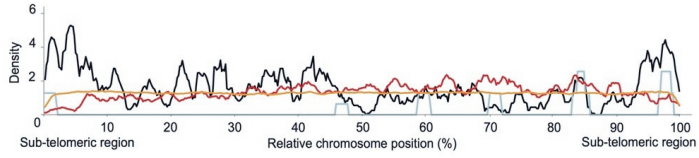
B



C

Gene groups	Genes with low genus-level synteny score	Genes with high genus-level synteny score	p value	Genes with low section-level synteny score	Genes with high section-level synteny score	p value
All genes	17,431	18,872	-	6,307	26,354	-
BUSCO genes	64	2,466	0	48	2,700	3.2e-117
SM backbone genes	279	14	6.1e-58	83	68	7.8e-28

D



E

Gene groups	Genes in sub-telomeric regions	Genes outside sub-telomeric regions	p value
All genes	9,139	35,186	-
BUSCO genes	207	2,746	4.3e-72
SM backbone genes	125	219	1.3e-12

Figure 5 Analysis of SM backbone gene localization in low-synteny and sub-telomeric regions in four *Aspergillus* species. (A) Summary table of the different groups of genes considered in four *Aspergillus* species. (B) Distribution of different groups of genes in relation to synteny score analyzed across 177 species and within individual sections. (C) Summary table of the number of SM backbone genes, BUSCO genes, and all genes with high-synteny score (larger than 80%) and low-synteny score (lower than 20%) at the genus-level and section-level. Chi-square statistic tests (Virtanen et al., 2020) were performed on SM backbone genes and BUSCO genes, determining whether their distributions differ significantly from the distribution of all genes in terms of genus-level synteny and section-level synteny. (D) Distribution of different groups of genes relative to chromosomal localization, considering the terminal 10% of chromosomal regions as sub-telomeric. (E) Summary table of the number of SM backbone genes, BUSCO genes, and all genes within and outside of sub-telomeric regions. Chi-square statistic tests (Virtanen et al., 2020) were performed on SM backbone genes and BUSCO genes, determining whether their distributions differ significantly from the distribution of all genes in terms of sub-telomeric localization.

Exploring the impact of genomic localization and histone PTMs on SM gene expression variability in Aspergilli

It has been proposed that genomic localization of SM BGCs is important for the regulation of their expression through histone modifications as several histone PTMs were shown to influence the production of SM in Aspergilli (Cánovas et al., 2014; Lee et al., 2009; Li et al., 2019). Thus, we sought to explore the relationship between gene expression, histone PTMs, and genomic and syntenic conditions of SM backbone genes compared with BUSCO genes. To this end, we utilized publicly available RNA-seq from *A. nidulans*, *A. fumigatus*, *A. niger*, and *A. oryzae* and ChIP-seq data from the former two Aspergilli (Tables S7 and S8). We used the coefficient of variance (CV) to assess the relative variability of gene expression and histone modification of different gene groups across diverse experimental conditions. For BUSCO genes, we observed stable gene expression in all species (Figures 6, S6, and S7), which is accompanied by low variation in all histone PTMs included in this study (H3K4me3, H3K9me3, H3K36me3, and H3Kac). Compared to BUSCO and all genes, SM backbone genes exhibit greater variability in gene expression (Figures 6, S6, and S7). For histone PTMs, H3K36me3 in *A. nidulans* and H3K4me3 in both *A. nidulans* and *A. fumigatus* show higher variability at SM backbone genes in regions compared with other gene groups (Figures 6 and S8), suggesting a role for these histone modifications in regulating the expression of SM genes. The correlation between the CVs of these marks' reads and RNA-seq reads further supports this (Figure 6). Especially, SM backbone genes with known structure and variable expression in *A. fumigatus* also exhibit high CVs in H3K4me3, with four of them ranking among the top five most variable SM genes (Table S5). However, the observed gene expression variation of SM genes is not explained by H3K9me3 (in *A. nidulans* and *A. fumigatus*) nor H3ac (in *A. nidulans*) as SM genes do not differ from BUSCO genes for these PTMs (Figures 6 and S8).

When taking the chromosome localization into account, SM backbone genes within the sub-telomeric regions show slightly higher variation in gene expression in *A. nidulans*, *A. fumigatus*, and *A. niger*, but not in *A. oryzae* (Figures 6 and S6). The difference in *A. oryzae* could be due to the fact that it was domesticated for industrial purposes such as producing sake, soy sauce, and miso, and it is generally recognized for its simplified metabolic profile (Gibbons et al., 2012). However, the sub-telomeric localization does not fully explain a higher variation in both gene expression and histone PTMs because the distribution of these variations for SM backbone genes outside sub-telomeric regions is bimodal and overlap with sub-telomeric backbone genes (Figures 6, S6, and S7). Lastly, SM backbone genes with the known metabolite structures listed on the MIBiG website show various expression patterns (Table S5). In *A. fumigatus*, known SM genes show high variability as their average CV (1.20) is higher than the average CV for both all SM backbone genes (1.09) and SM backbone genes outside sub-telomeric regions (0.94) (Table S5). Two SM backbone genes responsible for sartorypyrone A are in the top five most variably expressed SM backbone genes. By contrast, in other Aspergilli, either no (in *A. nidulans* and *A. niger*) or only one (in *A. oryzae*) SM

backbone gene with known structure is in the top five most variable SM backbone genes (Table S5). Based on our results, SM backbone genes that are still uncharacterized but exhibit higher variation in expression deserve further attention as they are likely to produce SMs under laboratory conditions.

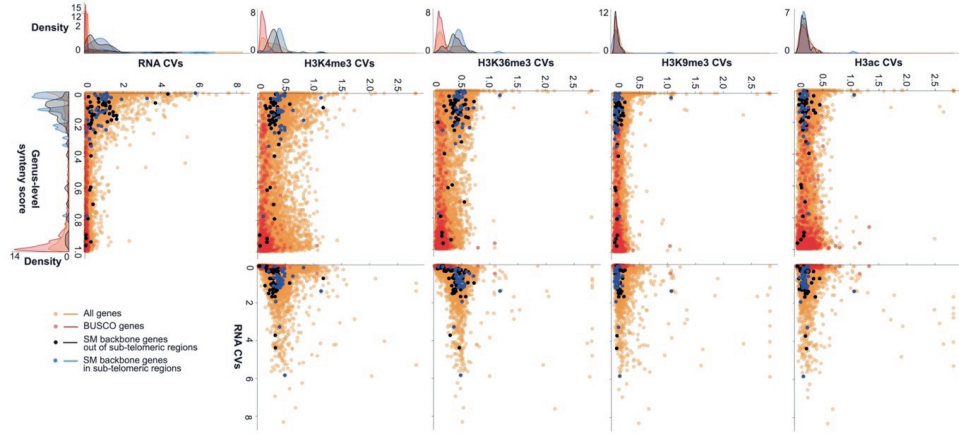


Figure 6 SM backbone genes in sub-telomeric regions tend to show more gene expression, H3K4me3, and H3K36me3 variation compared with BUSCO genes. We employed the coefficient of variance (CV) to assess the relative variability of all genes, BUSCO genes, and SM backbone genes in *A. nidulans*. The variability is depicted using distinct colors for dots and density lines to differentiate between the gene groups.

Conclusions

Altogether, our results highlighted that SM backbone genes are quite abundant and diverse in the fungal genus *Aspergillus*, especially between different sections. Moreover, we showed that BGCs exhibit a strong bias for low-synteny regions, and the association between BGC localization, variation in expression, and histone PTMs is restricted to H3K4me3 and H3K36me3 only. It is important to note that our study is limited to the few histone PTMs that are publicly available and other PTMs might play a more important role and need to be investigated. In particular, H3K79me is a relevant candidate as this modification was reported to localize at transition regions between heterochromatin and euchromatin, and we recently reported differences in abundances between different Aspergilli (X. Zhang et al., 2023).

Materials and methods

Acquisition of genomes and predicted proteomes for 177 fungal species

The genomes and predicted proteomes of 177 fungal species were retrieved from the Joint Genome Institute MycoCosm (https://mycocosm.jgi.doe.gov/mycocosm/home) repository (Grigoriev et al., 2014) in January 2022. These species comprise 174 Aspergilli as well as

three *Penicillium* species (*Penicillium brevicompactum*, *Penicillium griseofulvum*, and *Penicillium chrysogenum*) that were used as an outgroup in our analyses (Table S1).

Species tree reconstruction

To reconstruct the phylogenetic relationship between the here analyzed *Aspergillus* species, we used BUSCO (Benchmarking Universal Single-Copy Ortholog assessment tool) v4.0.1 (Simão et al., 2015) to retrieve 758 fungal single-copy genes (Table S2) that are present in the 177 genomes. For homologs of each single-copy BUSCO gene, we first generated multi-sequence alignment using MAFFT v7.271 (Katoh & Standley, 2013), then removed positions in the alignments with gaps in more than 90% of the sequences using '-gt 0.1' in TrimAl v1.2 (Capella-Gutiérrez et al., 2009). We concatenated the 758 alignments into one 'super-alignment' and employed IQ-TREE v1.6.10 (Chernomor et al., 2016; Nguyen et al., 2015) to construct the maximum-likelihood phylogeny with partitioned model selection to allow selecting a substitution model (Table S2) for each BUSCO gene alignment, and used both ultrafast bootstrap and SH-aLRT (Shimodaira–Hasegawa approximate Likelihood Ratio Test) with 1,000 pseudo-replicates to calculate branch supports.

Identification and conservation analysis of putative SM backbone genes and orthologous groups

We used antiSMASH (antibiotics and Secondary Metabolite Analysis SHell) v6.1.0 (Blin et al., 2021) (parameters: –minimal) to identify all regions with predicted BGCs in the 177 genome assemblies, and refined the classification of SM backbone genes with our in-house tool BGCtoolkit (https://github.com/WesterdijkInstitute/BGClib, DOI: 10.5281/zenodo.6593238) (Figure 1 and Table S3). Then, we extracted the protein sequence of the predicted backbone enzymes or of the adenylation domains for NRPS, NRPS-like, and PKS-NRPS hybrid enzymes using BGCtoolkit. For each type of backbone enzyme, protein sequences were aligned and trimmed using MAFFT and TrimAl, respectively, as mentioned above. We constructed phylogenetic trees using FastTree v2.1.11 with default parameters (Price et al., 2009). We subsequently used Newick Utilities v1.6.0 (Junier & Zdobnov, 2010) to identify orthologs *via* three steps: i) re-root the tree based on the mid-point (nw_reroot); ii) manually select the clades having more than 88 leaves and with trustworthy bootstrap value (>= 0.99) and mark them as orthologs; iii) for the remaining leaves, automatically split the clades with fewer than 88 leaves and with trustworthy bootstrap value (>= 0.99) (nw_ed -n -r). In the construction of the three adenylation domain trees, an extra step was undertaken to consolidate multiple domains originating from the same gene. This involved examining all leaves in the domain tree to identify the genes associated with these domains. Clades were collapsed and assigned a new clade name when the domains were found to originate from the same gene. For any remaining domains, the original clade names were retained. We built a saturation curve with the matplotlib.pyplot visualization library (Hunter, 2007) and used the function from the SciPy library (Virtanen et al., 2020) to implement a curve-fitting process,

which determines the rate of growth (β) of unique gene numbers under Heap's Law (Heaps, 1978).

We used Orthofinder v2.3.8 (Emms & Kelly, 2019) to identify orthologous groups across 177 predicted proteomes using default settings (Table S4). A saturation curve based on these orthologous groups was built as described above.

Genome synteny analysis

We employed MCScanX (Multiple Collinearity Scan Toolkit X version) (Wang et al., 2012) to conduct an all-versus-all pairwise comparison among all species to identify syntenic blocks (match size: -s 3 and max gaps: -m 15). We defined the 'synteny percentage' between a pair of species as the ratio of syntenic genes detected between two compared species to the total genes. We defined the 'between sections' synteny percentage by considering all pairs of species that are associated with different taxonomic sections, and the 'within sections' synteny percentage by considering all pairs from the same taxonomic section. For the synteny percentage of each taxonomic section, we considered all comparisons for species within that specific section. For visual representation of genome alignment between two species, we used the R package GENESPACE v1.2.3 (Lovell et al., 2022).

To calculate the average length of syntenic regions, we took all syntenic blocks identified by the MCScanX in all-versus-all pairwise comparisons and excluded the contigs with less than 50 genes. Subsequently, we employed the hypergeometric distribution (Rice, 2003) utilizing the 'hypergeom.cdf(x, M, n, N)' function from the SciPy library (Virtanen et al., 2020) to evaluate the likelihood of successfully detecting a syntenic region between two given species. In this function, 'M' denotes the total number of homologous gene pairs identified when comparing the complete genomes of species A and B, 'n' represents the count of genes on scaffold B1 with homologs in species B, and 'N' is the number of genes on scaffold A1 with homologs in species A. This cumulative distribution function aggregates the probabilities of identifying up to 'x' successful syntenic matches, thereby yielding the cumulative probability of detecting 'x' or fewer successes. The analysis considered (1) the total number of homologs between the two species, (2) the number of homologs on each contig relative to the other species, and (3) the number of homologs between the contigs themselves. Setting a p-value threshold at 1e-5, we then computed the average length of syntenic blocks for each pair of species.

We defined the synteny score at the genus level as the synteny score for a specific gene by dividing the number of species containing that gene by the total number of species (177 species). At the section level, the synteny score for a gene is determined by the ratio of the number of species within the section harboring that gene to the total species count in that section. We used the function 'chi2_contingency' from the SciPy library (Virtanen et al., 2020)

to perform chi-square statistic tests on SM backbone genes and BUSCO genes, determining whether their distributions differ significantly from the distribution of all genes in terms of genus-level synteny, section-level synteny, and sub-telomeric localization (Table S6).

RNA-seq sample preparation

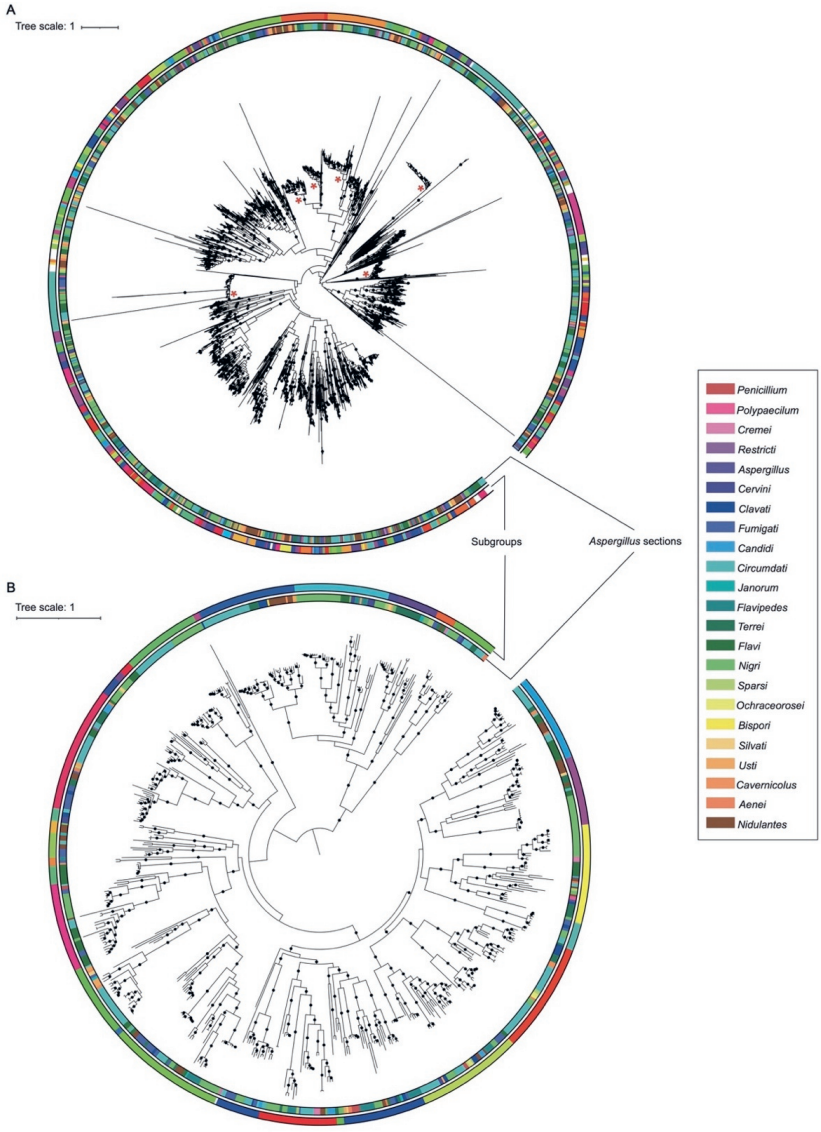
To perform RNA sequencing, we obtained the fine powder of young mycelium of *A. fumigatus* (strain: Af293, CBS 126847), *A. nidulans* (strain: Wtpaba (genotype: pabaA1)), and *A. niger* (strain: NRRL3, CBS 120.49) as described previously (X. Zhang et al., 2023). Subsequently, we extracted high-quality total RNA from each species in four replicates, following our in-house RNA extraction protocol. Briefly, we placed around 100 mg of fine powder in a 1.5 mL tube, added 1 mL TRIzol reagent (Fisher, 15596026), and incubated for 5 mins at 25°C. Following that, we added 0.2 mL chloroform-d (Sigma, 416754), mixed gently by hand 15 times, and incubated for 5 mins at room temperature. After centrifuging the samples at 10,000 RPM for 15 mins, we transferred the aqueous supernatant to a new tube. We then added 0.5 volume of 100% ethanol (BOOM, 200-578-6) and proceeded from the 5th step of the NucleoSpin RNA Kit protocol (Bioke, 740955). Finally, we eliminated DNA contamination using DNase I (Sigma, AMPD1) and sent samples to the Beijing Genome Institute (BGI, Hong Kong, China) for library construction and DNBseq (DNA NanoBall sequencing).

Acquisition and analysis of public RNA-seq and ChIP-seq datasets

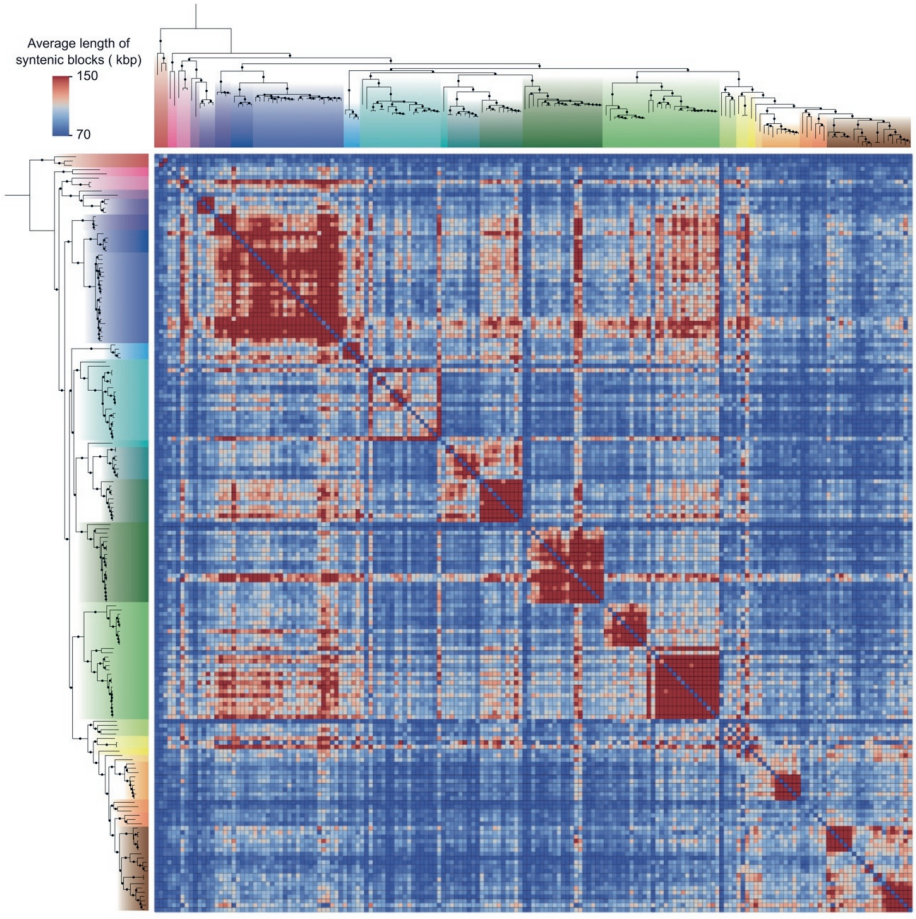
To retrieve publicly available RNA-seq data (Table S7) from *A. niger*, *A. fumigatus*, *A. nidulans*, and *A. oryzae*, we utilized prefetch v2.10.0 and fasterq-dump v2.10.0 from SRA Toolkit (https://github.com/ncbi/sra-tools) to extract FASTQ files from the SRA database (https://www.ncbi.nlm.nih.gov/sra). For all reads from both our BGI sequencing projects and the public database, we used the pseudoalignment tool kallisto v0.46.0 (Bray et al., 2016) to build the index and quantify raw read counts per gene. To remove the potential batch effect, we used ComBat-seq (Y. Zhang et al., 2020) in R v3.6.1 (Team, 2010) and obtained batch-effect corrected read counts. Following this, we calculated the TPM (Transcripts Per Million) and used log-transformed TPMs to calculate the coefficient of variation (CV) to assess the variability of gene expression levels across different samples. We used seaborn (Waskom, 2021) and Matplotlib (Hunter, 2007) to plot the scatter plot and density plots.

We utilized ChIP-seq data from two studies (Colabardini et al., 2022; Gacek-Matthews et al., 2016), which included *A. nidulans* (H3ac, H3K4me3, H3K9me3, and H3K36me3) and *A. fumigatus* (H3K4me3 and H3K9me3) (Table S8). The analyses (downloading and processing) of this data were the same as the RNA-seq analysis mentioned above, yet we did not perform the batch effect removal step because the two studies were conducted in different species.

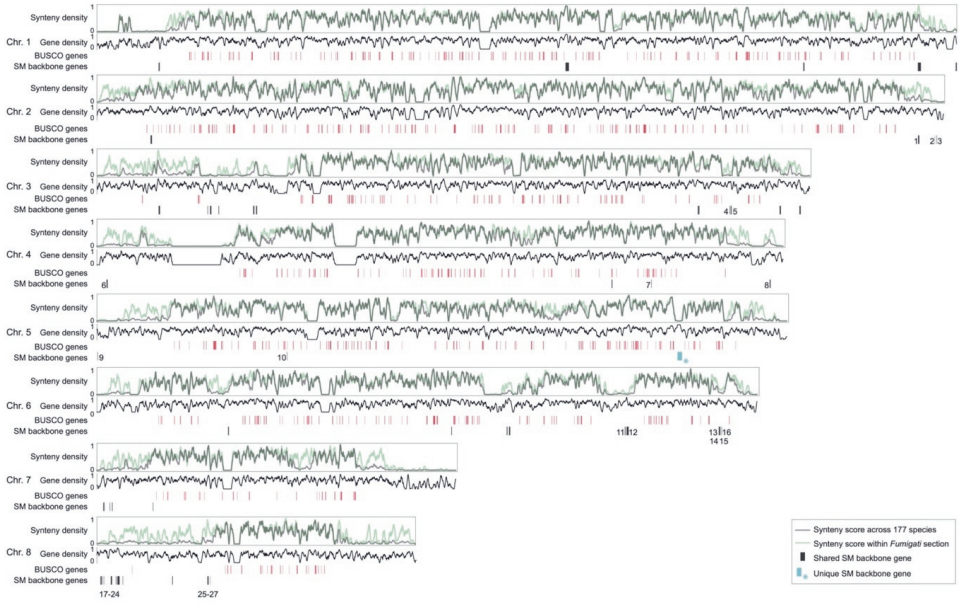
Supplementary materials



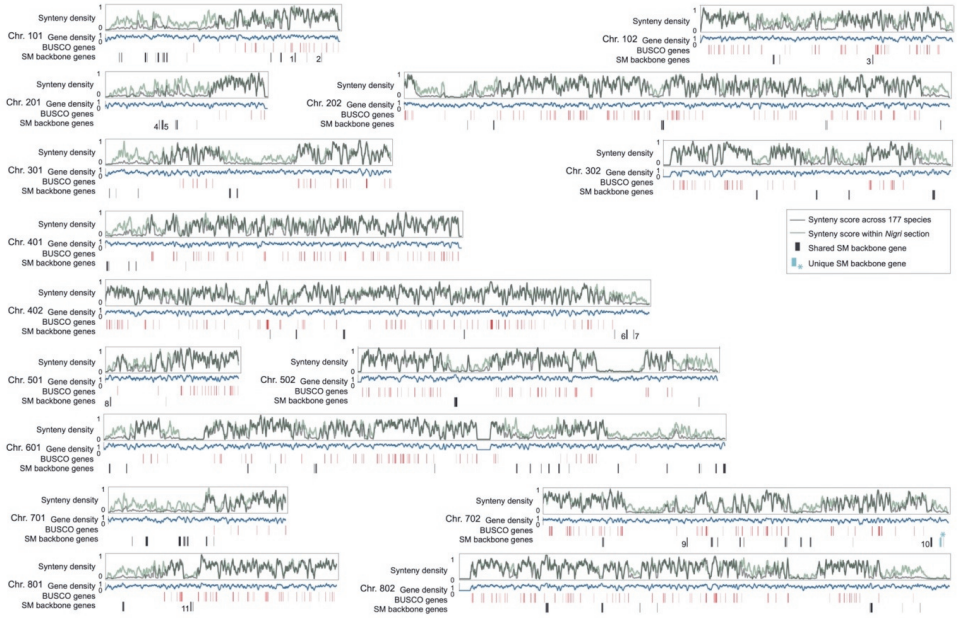
Supplementary Figure 1 The phylogeny-aware method to detect SM backbone gene diversity. We used FastTree (Price et al., 2009) to construct the phylogeny of each type of SM backbone genes and annotated the trees with iTOL. The inner colored ring indicates *Aspergillus* sections and the outer colored ring showed the splitting results by Newick Utilities (Junier & Zdobnov, 2010). (A) Manual selection is implemented first to pick up the clades larger than 88 leaves (red "*" in the tree). For remaining leaves and trees that don't need manual trimming (B), automatic splitting was performed with a trustworthy bootstrap value of >= 0.99.



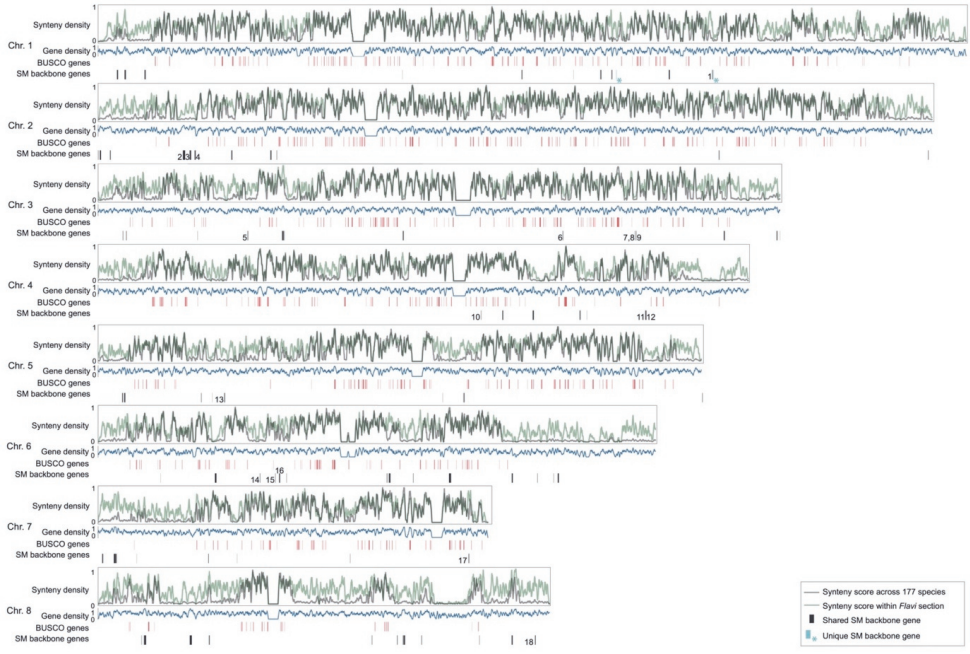
Supplementary Figure 2 Longer syntenic blocks are shared within *Aspergillus* sections. The heatmap displays the all-versus-all pairwise comparisons based on the average length of syntenic blocks across all species; species were arranged according to the phylogenetic tree. The accompanying species tree, displayed to the left and top, uses colors to differentiate between sections.



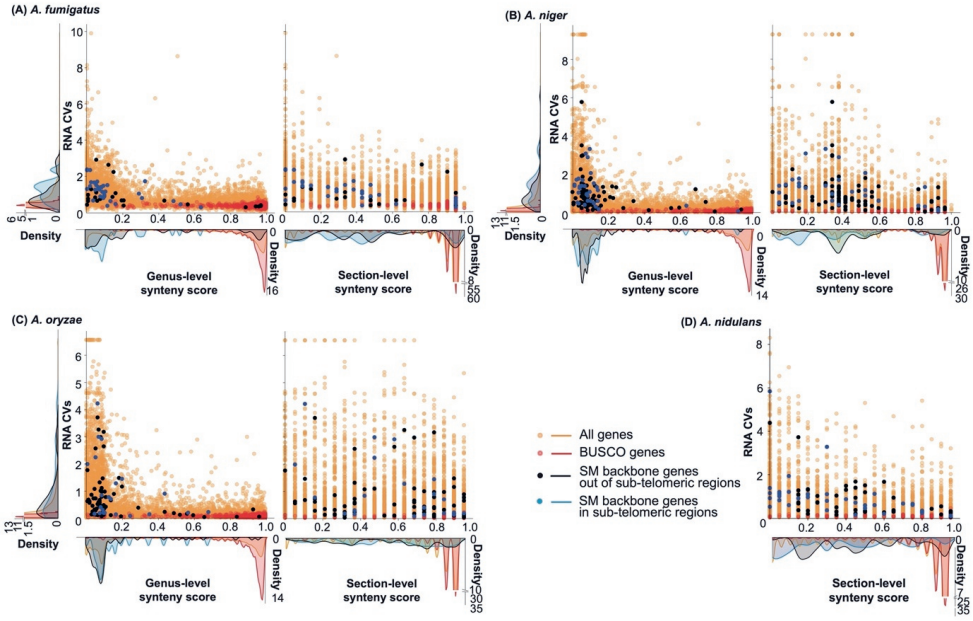
Supplementary Figure 3 SM backbone genes occur preferentially in low-synteny regions. This figure illustrates the genomic localization of BUSCO genes and SM backbone genes in *A. fumigatus*. Lines in black and green represent synteny density across 177 species and within the *Fumigati* section, respectively. The gene density is shown along the chromosomes. Red bars indicate the chromosomal locations of BUSCO genes, with bar width proportional to gene length. Black and blue bars mark the positions of shared and unique SM backbone genes, respectively, with their widths corresponding to gene length. Numbers adjacent to the black bars denote SM backbone genes associated with known metabolites as listed on the MIBiG website (Table S5).



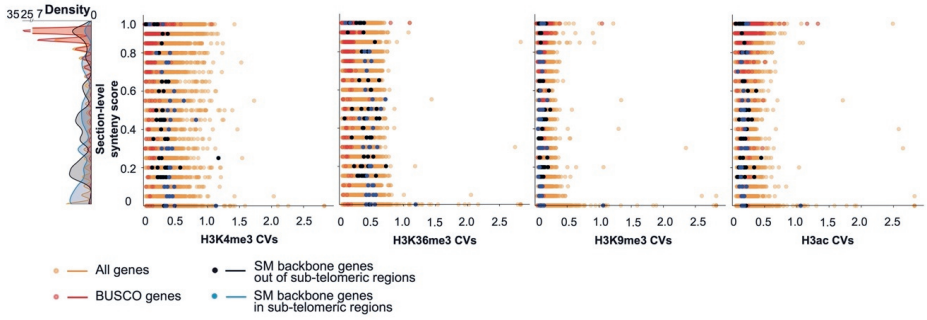
Supplementary Figure 4 SM backbone genes occur preferentially in low-synteny regions. This figure illustrates the genomic localization of BUSCO genes and SM backbone genes in *A. niger*. Lines in black and green represent synteny density across 177 species and within the *Nigri* section, respectively. The gene density is shown along the chromosomes. Red bars indicate the chromosomal locations of BUSCO genes, with bar width proportional to gene length. Black and blue bars mark the positions of shared and unique SM backbone genes, respectively, with their widths corresponding to gene length. Numbers adjacent to the black bars denote SM backbone genes associated with known metabolites as listed on the MIBiG website (Table S5).



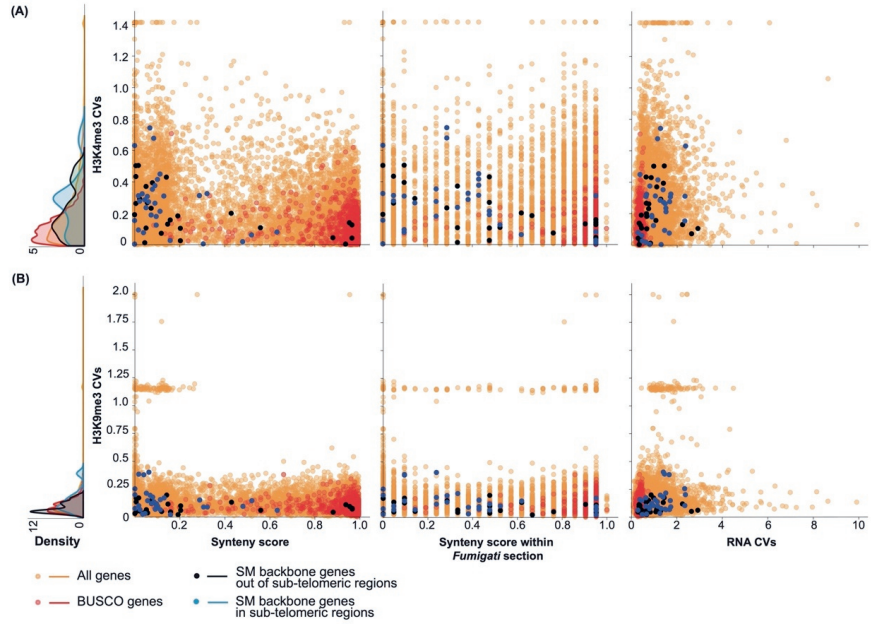
Supplementary Figure 5 SM backbone genes occur preferentially in low-synteny regions. This figure illustrates the genomic localization of BUSCO genes and SM backbone genes in *A. oryzae*. Lines in black and green represent synteny density across 177 species and within the *Flavi* section, respectively. The gene density is shown along the chromosomes. Red bars indicate the chromosomal locations of BUSCO genes, with bar width proportional to gene length. Black and blue bars mark the positions of shared and unique SM backbone genes, respectively, with their widths corresponding to gene length. Numbers adjacent to the black bars denote SM backbone genes associated with known metabolites as listed on the MIBiG website (Table S5).



Supplementary Figure 6 SM backbone genes show more variance in gene expression compared to BUSCO and all genes in four Aspergilli. We employed the coefficient of variance (CV) to assess the relative variability of all genes, BUSCO genes, and SM backbone genes within (A) *A. fumigatus*, (B) *A. niger*, (C) *A. oryzae*, and (D) *A. nidulans*. The variability is depicted using distinct colors for dots and density lines to differentiate between the gene groups. RNA-seq data was downloaded from SRA (Table S7) and ChIP-seq data is from the publication (Gacek-Matthews et al., 2016).



Supplementary Figure 7 Gene distribution of section-level synteny score in *A. nidulans*. We employed the coefficient of variance (CV) to assess the relative variability of all genes, BUSCO genes, and SM backbone genes in *A. nidulans*. The variability is depicted using distinct colors for dots and density lines to differentiate between the gene groups. ChIP-seq data is from the publication (Gacek-Matthews et al., 2016).



Supplementary Figure 8 SM backbone genes in sub-telomeric regions show more variance in H3K4me3 in *A. fumigatus*. We employed the coefficient of variance (CV) of H3K4me3 and H3K9me3 to assess the relative variability of all genes, BUSCO genes, and SM backbone genes. Data is from publication (Colabardini et al., 2022). The variability is depicted using distinct colors for dots and density lines to differentiate between the gene groups.

Supplementary Table 1 Summary of 177 *Aspergillus* and *Penicillium* species included in this study

No.	Species	JGI abbreviation	Section	Assembled length	Scaffolds	Genome completeness	Genes
1	*Aspergillus aculeatus*	Aspac1	*Nigri*	35424414	660	99.2	10845
2	*Aspergillus acristatulus*	Aspacri1	*Nidulantes*	32593788	311	99.2	11220
3	*Aspergillus aculeatinus*	Aspacu1	*Nigri*	36471649	121	99.5	12027
4	*Aspergillus affinis*	Aspaff1	*Circumdati*	37303864	638	98.4	12409
5	*Aspergillus alabamensis*	Aspala1	*Terrei*	31994729	219	99.5	11841
6	*Aspergillus albertensis*	Aspalbe1	*Flavi*	40140031	473	98.7	12816
7	*Aspergillus allahabadii*	Aspall1	*Terrei*	29783189	43	99.2	11228
8	*Aspergillus ambiguus*	Aspamb1	*Terrei*	29374911	149	99.2	11088
9	*Aspergillus amoeneus*	Aspamoe1	*Nidulantes*	33399842	173	98.8	13458
10	*Aspergillus amylovorus*	Aspamy1	*Cavernicolus*	34894441	462	99.5	12857
11	*Aspergillus arachidicola*	Aspara19utr	*Flavi*	39764385	451	98.9	13895
12	*Aspergillus arxii*	Asparx1	*Cremei*	27919226	471	98.3	10666
13	*Aspergillus assulatus*	Aspass1	*Fumigati*	30732071	674	99.5	10548
14	*Aspergillus aureoterreus*	Aspaur1	*Terrei*	29516179	109	98.8	11269
15	*Aspergillus aureofulgens*	Aspaurful1	*Flavipedes*	32899741	619	99.6	11485
16	*Aspergillus aureoluteus*	Aspaurlut1	*Fumigati*	33749143	711	98.5	10768
17	*Aspergillus avenaceus*	Aspave1	*Flavi*	33765472	1528	99.1	11293
18	*Aspergillus bertholletius*	Aspber1	*Flavi*	37014000	443	98.8	12948
19	*Aspergillus biplanus*	Aspbip1	*Sparsi*	42030724	501	98.8	15330
20	*Aspergillus bisporus*	Aspbisp1	*Bispori*	27095172	244	98.6	9731
21	*Aspergillus bombycis*	Aspbom1	*Flavi*	37474605	450	97.8	12265
22	*Aspergillus botucatensis*	Aspbot1	*Fumigati*	34492499	539	98.3	11812
23	*Aspergillus brasiliensis*	Aspbr1	*Nigri*	35808783	105	99.2	13000
24	*Aspergillus brevipes*	Aspbre1	*Fumigati*	30071435	237	98.8	10475
25	*Aspergillus brevijanus*	Aspbrev1	*Janorum*	36004448	685	99.3	12832
26	*Aspergillus brevistipitatus*	Aspbrevi1	*Fumigati*	28615967	320	98.5	10133
27	*Aspergillus brunneoviolaceus*	Aspbru1	*Nigri*	37477200	153	98.7	12075
28	*Aspergillus carbonarius*	Aspca3	*Nigri*	36290756	963	91.3	11624
29	*Aspergillus caelatus*	Aspcae1	*Flavi*	40016351	729	98.8	13916
30	*Aspergillus californicus*	Aspcalif1	*Cavernicolus*	36744266	398	98.9	13881
31	*Aspergillus campestris*	Aspcam1	*Candidi*	28257496	62	99.1	9764
32	*Aspergillus candidus*	Aspcand1	*Candidi*	27317772	268	98.7	9641
33	*Aspergillus caninus*	Aspcanin1	*Polypaecilum*	24260488	111	98.2	9969
34	*Aspergillus capensis*	Aspcap1	*Flavipedes*	32934471	529	98.2	11706
35	*Aspergillus carlsbadensis*	Aspcar1	*Usti*	39460077	254	99.1	14289
36	*Aspergillus cervinus*	Aspcer1	*Cervini*	34080287	604	98.9	11707
37	*Aspergillus chevalieri*	Aspchev1	*Aspergillus*	26413841	417	98.9	10252
38	*Aspergillus chrysellus*	Aspchr1	*Circumdati*	37957915	549	98.8	12926
39	*Aspergillus clavatus*	Aspcl1	*Clavati*	27859441	143	99.5	9121
40	*Aspergillus clavatonanicus*	Aspcla1	*Clavati*	26779860	309	99.1	9490
41	*Aspergillus conjunctus*	Aspcon1	*Sparsi*	33840813	675	98.9	12025
42	*Aspergillus coremiiformis*	Aspcor1	*Flavi*	30132285	2728	99.5	9078
43	*Aspergillus coreanus*	Aspcorea1	*Fumigati*	31295293	447	99.1	10893
44	*Aspergillus costaricaensis*	Aspcos1	*Nigri*	36947918	86	99.2	11966
45	*Aspergillus cretensis*	Aspcre1	*Circumdati*	36176632	586	99.2	11751
46	*Aspergillus creber*	Aspcreb1	*Nidulantes*	35476325	422	98.5	13531
47	*Aspergillus cristatus*	Aspcri1	*Aspergillus*	28454192	68	95.8	10344
48	*Aspergillus crustosus*	Aspcru1	*Aenei*	37360539	462	98.7	12063
49	*Aspergillus desertorum*	Aspdese1	*Nidulantes*	29041708	288	98.9	10111
50	*Aspergillus duricaulis*	Aspdur1	*Fumigati*	30322558	266	98.7	10468
51	*Aspergillus egyptiacus*	Aspegy1	*Cavernicolus*	28335649	682	99.2	9890
52	*Aspergillus elegans*	Aspele1	*Circumdati*	37553232	679	98.0	13286
53	*Aspergillus ellipticus*	Aspell1	*Nigri*	42866077	518	98.6	12884
54	*Aspergillus eucalypticola*	Aspeuc1	*Nigri*	34792412	131	98.6	11934
55	*Aspergillus ferenczii*	Aspfer1	*Fumigati*	26938309	126	99.1	10035
56	*Aspergillus fijiensis*	Aspfij1	*Nigri*	36513906	149	99.5	12019
57	*Aspergillus filifera*	Aspfili1	*Nidulantes*	31484740	814	99.1	11685
58	*Aspergillus flavus*	Aspfl1	*Flavi*	36790245	138	94.7	12604
59	*Aspergillus flavipes*	Aspfla1	*Flavipedes*	35257526	476	98.5	13390
60	*Aspergillus floccosus*	Aspflo1	*Terrei*	30552350	224	98.4	11810

No.	Species	JGI abbreviation	Section	Assembled length	Scaffolds	Genome completeness	Genes
61	*Aspergillus flocculosus*	Aspfloc1	*Circumdati*	34735525	268	97.9	12432
62	*Aspergillus luchuensis*	Aspfo1	*Nigri*	37468345	107	99.3	13530
63	*Aspergillus frequens*	Aspfre1	*Flavipedes*	33664157	239	98.7	12444
64	*Aspergillus fumigatus*	Aspfu1	*Fumigati*	29388377	8	98.3	9781
65	*Aspergillus fumigatiaffinis*	Aspfumig1	*Fumigati*	33218226	201	98.0	11917
66	*Aspergillus fumisynnematus*	Aspfumis1	*Fumigati*	32126662	399	98.8	11626
67	*Aspergillus funiculosus*	Aspfuni1	*Ochraceorosei*	25445165	144	98.8	9300
68	*Aspergillus galapagensis*	Aspgal1	*Fumigati*	29972746	261	98.0	11453
69	*Aspergillus germanicus*	Aspger1	*Usti*	39479152	430	98.5	14761
70	*Aspergillus giganteus*	Aspgig1	*Clavati*	28200622	151	98.7	9308
71	*Aspergillus glaucus*	Aspgl1	*Aspergillus*	27993362	82	99.1	11277
72	*Aspergillus granulosus*	Aspgra1	*Usti*	35311868	455	98.3	12860
73	*Aspergillus haitiensis*	Asphai1	*Sparsi*	35849033	1702	98.3	11396
74	*Aspergillus heteromorphus*	Asphet1	*Nigri*	35607567	205	99.5	11133
75	*Aspergillus heterothallicus*	Asphethal1	*Usti*	37429956	371	98.9	13761
76	*Aspergillus hiratsukae*	Asphir1	*Fumigati*	29263246	305	99.1	10825
77	*Aspergillus homomorphus*	Asphom1	*Nigri*	34054866	152	99.5	11361
78	*Aspergillus hortai*	Asphor1	*Terrei*	30242590	226	99.1	11172
79	*Aspergillus ibericus*	Aspibe1	*Nigri*	33437983	116	99.5	11680
80	*Aspergillus igneus*	Aspign1	*Fumigati*	33547489	491	99.2	11877
81	*Aspergillus iizukae*	Aspiiz1	*Flavipedes*	32658954	223	98.3	12279
82	*Aspergillus implicatus*	Aspimp1	*Sparsi*	39195316	372	98.7	13656
83	*Aspergillus indologenus*	Aspind1	*Nigri*	38592838	334	99.2	12163
84	*Aspergillus indicus*	Aspind2_1	*Nidulantes*	32492420	601	98.9	11359
85	*Aspergillus insuetus*	Aspins1	*Usti*	40452873	296	99.5	15211
86	*Aspergillus insolitus*	Aspinso1	*Polypaecilum*	23177606	157	99.3	8833
87	*Aspergillus japonicus*	Aspjap1	*Nigri*	36099954	163	98.7	12024
88	*Aspergillus kawachii*	Aspka1_1	*Nigri*	36544702	317	98.3	11475
89	*Aspergillus karnatakaensis*	Aspkar1	*Aenei*	36255367	239	99.1	13540
90	*Aspergillus keveii*	Aspkev1	*Usti*	41415647	587	98.5	15687
91	*Aspergillus laciniosus*	Asplaci1	*Fumigati*	33457904	348	99.1	11759
92	*Aspergillus lentulus*	Asplen1	*Fumigati*	30572422	58	98.5	11171
93	*Aspergillus longivesica*	Asplon1	*Clavati*	26674246	432	99.2	9605
94	*Aspergillus lucknowensis*	Aspluc1	*Usti*	30792796	232	98.7	11541
95	*Aspergillus luppii*	Asplup1	*Flavipedes*	33321554	383	99.1	11947
96	*Aspergillus microcysticus*	Aspmic1	*Nidulantes*	25913919	60	98.5	10236
97	*Aspergillus minisclerotigenes*	Aspmin1	*Flavi*	37064162	296	99.1	13415
98	*Aspergillus multicolor*	Aspmul1_1	*Nidulantes*	35254497	113	99.2	12821
99	*Aspergillus multiplicatus*	Aspmult1	*Fumigati*	28914259	191	99.2	10498
100	*Aspergillus muricatus*	Aspmuri1	*Circumdati*	34116413	353	98.5	11718
101	*Aspergillus nakazawae*	Aspnaka1	*Circumdati*	36345397	954	98.4	12556
102	*Aspergillus neoechinulatus*	Aspne1	*Nidulantes*	30669358	236	98.9	10504
103	*Aspergillus neoniger*	Aspneo1	*Nigri*	35416444	169	99.2	11939
104	*Aspergillus neoflavipes*	Aspneof1	*Flavipedes*	34015292	324	98.5	12734
105	*Aspergillus neoindicus*	Aspneoi1	*Terrei*	30897550	121	99.1	11849
106	*Aspergillus niger*	Aspni_NRRL3_1	*Nigri*	35245396	15	98.8	11846
107	*Aspergillus nidulans*	Aspnid1	*Nidulantes*	30483994	8	98.2	10680
108	*Aspergillus nishimurae*	Aspnis1	*Fumigati*	30189492	218	98.9	9648
109	*Aspergillus niveus*	Aspniv1	*Terrei*	30196435	107	99.2	11488
110	*Aspergillus nomius*	Aspnom1	*Flavi*	36713177	290	98.8	12897
111	*Aspergillus novofumigatus*	Aspnov1	*Fumigati*	32440129	62	98.7	11549
112	*Aspergillus novoparasiticus*	Aspnovo1	*Flavi*	40857493	870	98.7	14182
113	*Aspergillus nutans*	Aspnut1	*Cervini*	34348625	427	98.7	11620
114	*Aspergillus ochraceus*	Aspoc7043_1	*Circumdati*	35604108	8	98.5	12066
115	*Aspergillus ochraceoroseus*	Aspoch1	*Ochraceorosei*	27718211	34	98.6	8924
116	*Aspergillus olivicola*	Aspoli1_1	*Nidulantes*	33173879	824	98.0	12183
117	*Aspergillus oryzae*	Aspor1	*Flavi*	37882812	11	93.7	12030
118	*Aspergillus ostianus*	Asposti1	*Circumdati*	35652827	606	98.7	12122
119	*Aspergillus parasiticus*	Asppar1	*Flavi*	38390758	270	98.8	13752
120	*Aspergillus parvulus*	Aspparv1	*Cervini*	32731898	329	99.1	11927
121	*Aspergillus pseudodeflectus*	Asppdef1	*Usti*	39072803	179	98.9	14450
122	*Aspergillus penicilloides*	Asppeni1	*Restricti*	26403364	255	98.8	10386
123	*Aspergillus persii*	Aspper1	*Circumdati*	37819877	342	99.2	12945

No.	Species	JGI abbreviation	Section	Assembled length	Scaffolds	Genome completeness	Genes
124	*Aspergillus petrakii*	Asppet1	*Circumdati*	35785289	561	98.9	12569
125	*Aspergillus piperis*	Asppip1	*Nigri*	35280331	47	98.9	12069
126	*Aspergillus pseudocaelatus*	Asppsec1	*Flavi*	39651556	466	97.5	13895
127	*Aspergillus pseudonomius*	Asppsen1	*Flavi*	37780514	374	99.5	13384
128	*Aspergillus pseudotamarii*	Asppset1	*Flavi*	38243929	249	98.0	13428
129	*Aspergillus pseudoterreus*	Asppseute1	*Terrei*	29514308	270	98.3	10945
130	*Aspergillus pulvericola*	Asppul1	*Circumdati*	34902482	435	98.5	11818
131	*Aspergillus pseudoustus*	Asppust1	*Usti*	41220403	656	98.7	15536
132	*Aspergillus puulaauensis*	Asppuu1	*Nidulantes*	33961053	198	98.2	12901
133	*Aspergillus rambellii*	Aspram1	*Ochraceorosei*	26435630	4177	97.1	7761
134	*Aspergillus recurvatus*	Asprec1_1	*Nidulantes*	30717846	269	98.9	11005
135	*Aspergillus restrictus*	Aspres1	*Restricti*	23256008	297	99.5	9143
136	*Aspergillus rhizopodus*	Asprhi1	*Clavati*	27579276	297	98.7	9816
137	*Aspergillus robustus*	Asprobu1	*Circumdati*	33138998	186	99.2	12180
138	*Aspergillus roseoglobulosus*	Aspros1	*Circumdati*	35815337	708	98.7	12489
139	*Aspergillus saccharolyticus*	Aspsac1	*Nigri*	31096189	120	99.2	10066
140	*Aspergillus sclerotiorum*	Aspsc1	*Circumdati*	37971091	335	99.3	12620
141	*Aspergillus sclerotioniger*	Aspscl1	*Nigri*	36720365	138	99.5	12338
142	*Aspergillus sclerotiicarbonarius*	Aspscle1	*Nigri*	37621264	166	99.1	12571
143	*Aspergillus sepultus*	Aspsep1	*Cremei*	32180178	194	98.9	12290
144	*Aspergillus sergii*	Aspser1	*Flavi*	38325693	262	98.7	13713
145	*Aspergillus sesamicola*	Aspses1	*Circumdati*	38310826	435	98.0	12058
146	*Aspergillus silvaticus*	Aspsil1	*Silvati*	30974195	257	98.7	10314
147	*Aspergillus spectabilis*	Aspspe1_1	*Aenei*	38710628	715	98.7	13926
148	*Aspergillus steynii*	Aspste1	*Circumdati*	37847960	37	99.1	13211
149	*Aspergillus stella-maris*	Aspstel1_1	*Nidulantes*	33998249	266	98.3	12535
150	*Aspergillus subramanianii*	Aspsubr1	*Circumdati*	37435068	564	99.1	12314
151	*Aspergillus sydowii*	Aspsy1	*Nidulantes*	34381026	97	99.2	13620
152	*Aspergillus taichungensis*	Asptaic1	*Candidi*	27121736	310	98.6	9691
153	*Aspergillus tamarii*	Asptam1	*Flavi*	38511287	448	98.6	13331
154	*Aspergillus tanneri*	Asptan1	*Circumdati*	37353086	684	98.8	12392
155	*Aspergillus terreus*	Aspte1	*Terrei*	29331195	26	93.0	10406
156	*Aspergillus templicola*	Asptem1	*Flavipedes*	33946247	242	98.9	12132
157	*Aspergillus tennesseensis*	Aspten1	*Nidulantes*	33969804	194	98.5	12788
158	*Aspergillus thesauricus*	Aspth1	*Usti*	37882319	610	98.7	12812
159	*Aspergillus triticus*	Asptr1	*Candidi*	27452540	228	99.6	9769
160	*Aspergillus transmontanensis*	Asptra1	*Flavi*	39338775	293	97.8	14216
161	*Aspergillus transcarpathicus*	Asptran1	*Cervini*	30225050	234	98.9	10456
162	*Aspergillus tubingensis*	Asptu1	*Nigri*	35146149	33	98.8	12322
163	*Aspergillus udagawae*	Aspuda1	*Fumigati*	32185562	1029	97.8	9999
164	*Aspergillus undulatus*	Aspund1_1	*Nidulantes*	32112944	190	99.1	11802
165	*Aspergillus unguis*	Aspung1	*Nidulantes*	26060700	142	99.2	10397
166	*Aspergillus uvarum*	Aspuva1	*Nigri*	35850703	172	98.6	12017
167	*Aspergillus vadensis*	Aspvad1	*Nigri*	35663374	60	98.8	12132
168	*Aspergillus versicolor*	Aspve1	*Nidulantes*	33126810	51	99.1	13228
169	*Aspergillus venezuelensis*	Aspven1	*Nidulantes*	34837464	715	98.6	12450
170	*Aspergillus violaceofuscus*	Aspvio1	*Nigri*	36012865	209	99.3	12082
171	*Aspergillus wentii*	Aspwe1	*Cremei*	31350635	27	99.2	12442
172	*Aspergillus welwitschiae*	Aspwel1	*Nigri*	37511876	396	98.2	13687
173	*Aspergillus ruber*	Eurhe1	*Aspergillus*	26209327	110	98.9	10076
174	*Aspergillus fischeri*	Neofi1	*Fumigati*	32551711	976	99.1	10406
175	*Penicillium brevicompactum*	Penbr2	*Penicillium*	32114150	35	98.2	11536
176	*Penicillium griseofulvum*	Pench1	*Penicillium*	31340922	27	99.5	11396
177	*Penicillium chrysogenum*	Pengri1	*Penicillium*	29140916	122	98.8	9630

Supplementary Table 2 List of BUSCO genes used to build the species tree

BUSCO gene	Best fit model	BUSCO gene	Best fit model	BUSCO gene	Best fit model
100957at4751	JTT+R5	259700at4751	JTT+I+G4	398592at4751	JTT+F+R5
101133at4751	LG+R4	259883at4751	JTT+R3	398882at4751	PMB+I+G4
10168at4751	JTTDCMut+F+R5	260251at4751	JTT+I+G4	398968at4751	JTTDCMut+G4
103183at4751	JTT+R5	260722at4751	JTT+I+G4	399376at4751	JTT+R5
103461at4751	JTT+R5	260896at4751	JTT+R4	399413at4751	LG+R4
103836at4751	LG+R4	261237at4751	JTT+F+R5	399810at4751	JTTDCMut+R4
104173at4751	JTT+F+R5	262317at4751	JTT+I+G4	401387at4751	JTT+I+G4
104415at4751	JTT+R6	26329at4751	LG+R8	402005at4751	mtZOA+R3
10531at4751	JTT+R5	263390at4751	JTT+R4	402255at4751	JTT+I+G4
105401at4751	LG+R4	264038at4751	JTT+F+R5	403079at4751	JTTDCMut+I+G4
106281at4751	HIVb+F+R6	264164at4751	JTT+R5	404050at4751	JTT+R5
107136at4751	JTT+R5	265302at4751	JTT+F+I+G4	405669at4751	JTT+R5
107464at4751	LG+R4	265325at4751	mtMet+F+I+G4	405898at4751	JTT+R4
109276at4751	JTT+F+R5	266634at4751	JTTDCMut+I+G4	406240at4751	JTT+I+G4
111022at4751	JTT+R4	267556at4751	JTTDCMut+R5	406403at4751	JTTDCMut+R4
111189at4751	JTT+F+R5	269389at4751	LG+R3	406434at4751	mtZOA+F+I+G4
11140at4751	HIVb+F+R6	269580at4751	JTT+R5	406793at4751	Dayhoff+I+G4
112646at4751	JTT+R6	269939at4751	JTT+R5	407348at4751	JTT+R5
112937at4751	JTT+R5	27008at4751	JTT+R5	408391at4751	JTT+F+R4
115993at4751	JTT+R5	271511at4751	JTT+I+G4	408419at4751	JTT+R5
11636at4751	JTT+R5	271781at4751	JTT+R4	409761at4751	mtMet+F+I+G4
117017at4751	JTTDCMut+R6	271850at4751	JTT+R4	410458at4751	LG+I+G4
118189at4751	JTT+I+G4	272081at4751	JTT+R8	410782at4751	JTT+I+G4
118701at4751	JTT+R8	272711at4751	JTTDCMut+R4	410986at4751	LG+I+G4
11879at4751	JTT+F+R5	273469at4751	JTTDCMut+I+G4	410991at4751	LG+F+I+G4
119833at4751	JTT+I+G4	274828at4751	JTT+F+R5	411130at4751	LG+I+G4
122507at4751	JTT+R6	276205at4751	JTT+I+G4	411471at4751	mtMet+F+R5
12258at4751	JTTDCMut+R5	278643at4751	JTT+F+R5	411518at4751	JTT+R5
122824at4751	JTT+R5	279020at4751	JTT+F+I+G4	411677at4751	LG+I+G4
123228at4751	JTT+F+R5	280937at4751	LG+R5	411754at4751	JTT+R3
123277at4751	JTT+F+R5	281299at4751	JTT+R5	411919at4751	JTT+R7
123494at4751	JTT+R4	281726at4751	LG+I+G4	412006at4751	LG+I+G4
125400at4751	HIVb+F+R6	281920at4751	JTTDCMut+R5	413850at4751	JTT+I+G4
125935at4751	JTTDCMut+R5	282059at4751	JTT+I+G4	414474at4751	JTT+R5
126037at4751	JTT+R5	282964at4751	JTT+I+G4	414508at4751	JTT+I+G4
126249at4751	JTT+F+I+G4	283177at4751	JTT+R4	416039at4751	JTT+R5
126256at4751	JTT+F+I+G4	284127at4751	LG+R4	416914at4751	JTT+F+I+G4
126519at4751	JTT+R5	284176at4751	JTTDCMut+F+R5	417109at4751	JTT+R5
127140at4751	JTT+F+R5	284287at4751	JTTDCMut+R6	417470at4751	JTT+R5
12812at4751	HIVb+F+R6	284377at4751	JTT+F+R5	417746at4751	JTTDCMut+I+G4
129160at4751	LG+R4	285254at4751	JTTDCMut+R5	418242at4751	LG+R4
129326at4751	JTT+R5	285477at4751	JTTDCMut+R5	418644at4751	JTT+R5
129520at4751	LG+I+G4	285977at4751	JTT+R5	418680at4751	JTTDCMut+R4
12959at4751	JTT+R5	289666at4751	LG+R4	419048at4751	LG+G4
130062at4751	JTT+R4	289686at4751	JTTDCMut+I+G4	419200at4751	JTTDCMut+R4
130930at4751	JTT+R6	290314at4751	LG+R4	419232at4751	JTT+I+G4
130976at4751	JTT+I+G4	290471at4751	JTT+R5	419667at4751	JTT+I+G4
13224at4751	JTT+F+R5	291828at4751	JTT+R5	419868at4751	LG+I+G4
133954at4751	JTTDCMut+R5	292336at4751	JTT+I+G4	420414at4751	JTT+F+R6
135184at4751	JTT+F+R5	292418at4751	LG+R4	420527at4751	JTTDCMut+F+R5
135652at4751	JTT+R4	293405at4751	JTT+I+G4	421097at4751	JTT+I+G4
13605at4751	JTT+R5	29512at4751	JTTDCMut+R5	421687at4751	JTT+R5
137963at4751	JTTDCMut+I+G4	295692at4751	JTT+R5	422151at4751	LG+R2
138484at4751	HIVb+F+R6	295937at4751	JTT+I+G4	422622at4751	JTTDCMut+R4
138695at4751	LG+R4	296024at4751	JTT+R4	422764at4751	JTT+R6
138697at4751	LG+I+G4	296643at4751	JTT+R5	423166at4751	WAG+I+G4
138833at4751	JTTDCMut+I+G4	297145at4751	JTT+F+R6	423206at4751	JTTDCMut+R4
141273at4751	JTT+R5	297671at4751	JTT+R5	427250at4751	Dayhoff+I+G4
142357at4751	JTT+R6	298335at4751	JTT+R5	428342at4751	LG+I+G4
14295at4751	LG+R10	298379at4751	JTT+I+G4	428569at4751	JTTDCMut+R4
142965at4751	FLU+F+I+G4	299054at4751	JTT+R5	428774at4751	LG+I+G4
143414at4751	JTT+R5	299330at4751	JTTDCMut+R4	428984at4751	JTT+F+R5

BUSCO gene	Best fit model	BUSCO gene	Best fit model	BUSCO gene	Best fit model
14441at4751	JTTDCMut+F+R5	299509at4751	LG+R4	429332at4751	JTT+I+G4
145178at4751	JTT+R6	299559at4751	LG+R5	429862at4751	JTT+R4
14523at4751	JTT+R3	299583at4751	JTT+I+G4	430556at4751	JTT+R5
145378at4751	JTTDCMut+R4	299611at4751	JTT+R4	431195at4751	LG+I+G4
145524at4751	JTT+F+R6	300016at4751	JTT+I+G4	431607at4751	LG+R4
147028at4751	JTTDCMut+R4	300216at4751	JTTDCMut+R4	432952at4751	JTTDCMut+I+G4
14793at4751	JTT+R5	302127at4751	JTT+R4	433181at4751	JTT+I+G4
148293at4751	JTT+I+G4	304341at4751	JTT+I+G4	433384at4751	JTT+F+R5
14986at4751	JTTDCMut+R6	304453at4751	JTT+R4	434428at4751	VT+R3
150469at4751	mtZOA+F+R4	304547at4751	JTTDCMut+R5	434500at4751	JTT+R4
153083at4751	JTT+R6	304733at4751	LG+R3	434984at4751	JTT+I+G4
154126at4751	JTT+R6	305061at4751	JTT+R5	435317at4751	JTT+R5
156073at4751	JTT+R5	306594at4751	JTT+I+G4	435471at4751	JTT+I+G4
160044at4751	JTT+R10	307144at4751	JTT+R5	435767at4751	JTT+I+G4
160875at4751	JTTDCMut+R5	308236at4751	JTTDCMut+F+R5	436872at4751	JTTDCMut+R3
160940at4751	JTT+I+G4	308816at4751	JTT+I+G4	437914at4751	JTTDCMut+I+G4
160995at4751	JTT+I+G4	310718at4751	JTT+R4	438274at4751	JTT+F+I+G4
161113at4751	JTT+R7	310891at4751	JTT+R4	43842at4751	JTT+R5
163318at4751	JTT+F+R5	311048at4751	JTTDCMut+I+G4	438731at4751	JTT+I+G4
165243at4751	JTT+R5	311254at4751	JTTDCMut+I+G4	438861at4751	JTT+I+G4
167290at4751	JTTDCMut+F+R4	312080at4751	JTT+I+G4	439490at4751	JTT+I+G4
167374at4751	JTT+R4	312209at4751	JTT+F+R5	440401at4751	JTT+I+G4
167498at4751	LG+F+R3	31256at4751	JTT+R6	4409at4751	JTT+R6
168092at4751	JTTDCMut+R5	312950at4751	JTT+I+G4	441586at4751	JTT+R5
168489at4751	JTTDCMut+R5	31324at4751	JTT+R6	443595at4751	JTTDCMut+I+G4
169381at4751	JTT+F+R4	315551at4751	mtMet+F+R4	444174at4751	JTT+G4
170297at4751	JTT+R5	315639at4751	JTT+R5	444399at4751	LG+R3
170401at4751	JTT+R6	315799at4751	JTT+F+I+G4	444602at4751	JTT+I+G4
170904at4751	JTT+F+R6	315802at4751	JTT+R5	444892at4751	JTT+R5
171181at4751	JTT+R6	316254at4751	JTT+I+G4	445210at4751	JTT+F+I+G4
171238at4751	JTT+I+G4	316920at4751	mtInv+F+I+G4	445401at4751	LG+I+G4
17174at4751	JTTDCMut+R6	317276at4751	JTT+R4	445954at4751	JTTDCMut+I+G4
172113at4751	JTT+I+G4	317335at4751	JTTDCMut+R4	446124at4751	LG+I+G4
172363at4751	JTT+F+I+G4	319121at4751	JTTDCMut+R6	4466at4751	JTT+F+R5
172966at4751	JTT+R5	320163at4751	JTT+I+G4	447112at4751	JTT+I+G4
173428at4751	JTT+R5	320638at4751	JTTDCMut+I+G4	447424at4751	mtMet+F+R4
173995at4751	JTT+R6	320777at4751	JTT+F+R5	449717at4751	mtMet+F+R5
174653at4751	LG+I+G4	320782at4751	JTT+R5	449803at4751	LG+I+G4
174672at4751	JTT+R9	321416at4751	JTTDCMut+R5	451032at4751	JTT+I+G4
175136at4751	JTT+R4	321585at4751	LG+R5	451701at4751	JTT+I+G4
175599at4751	JTT+R5	32164at4751	JTT+R4	451811at4751	JTT+R5
176178at4751	JTT+R9	321728at4751	JTT+I+G4	451849at4751	JTT+I+G4
176478at4751	JTTDCMut+R5	322254at4751	JTT+R5	4523at4751	JTTDCMut+R5
176648at4751	JTT+F+R6	322267at4751	JTT+I+G4	452573at4751	JTTDCMut+I+G4
176748at4751	JTT+F+R4	323327at4751	JTT+I+G4	452692at4751	JTT+I+G4
179025at4751	JTT+R5	324442at4751	JTT+F+R4	452759at4751	JTT+I+G4
179437at4751	JTT+I+G4	325992at4751	LG+I+G4	452947at4751	JTT+I+G4
179897at4751	JTT+R5	326512at4751	JTT+R6	453693at4751	LG+I+G4
181890at4751	JTT+R5	326649at4751	JTT+I+G4	453909at4751	JTT+R6
181959at4751	LG+I+G4	327672at4751	JTT+I+G4	454746at4751	JTTDCMut+I+G4
183720at4751	JTT+R5	327748at4751	JTT+R4	454840at4751	JTT+I+G4
183801at4751	JTT+I+G4	327884at4751	LG+F+R4	455110at4751	JTT+I+G4
183849at4751	JTTDCMut+R4	329067at4751	JTT+R5	456963at4751	Dayhoff+G4
184718at4751	JTT+R4	330533at4751	JTT+R5	457161at4751	JTT+R4
184762at4751	LG+R2	330551at4751	JTT+R5	457544at4751	JTTDCMut+R4
185196at4751	JTT+R4	330629at4751	JTT+I+G4	457951at4751	LG+G4
185399at4751	JTT+R5	330815at4751	LG+I+G4	458554at4751	mtZOA+F+R4
186138at4751	JTT+R5	330949at4751	JTT+R6	458980at4751	JTT+R5
186574at4751	JTT+R6	331022at4751	JTTDCMut+I+G4	459844at4751	JTT+I+G4
1866at4751	JTTDCMut+R4	331763at4751	WAG+R2	46090at4751	JTTDCMut+R4
186701at4751	JTT+R5	332006at4751	JTT+I+G4	46098at4751	JTT+R10
187821at4751	LG+R4	332036at4751	JTT+F+I+G4	46119at4751	HIVb+F+R6
188070at4751	JTT+I+G4	332416at4751	JTT+I+G4	461238at4751	JTT+I+G4
18809at4751	JTT+F+R5	332876at4751	JTTDCMut+R5	461919at4751	JTT+R5

BUSCO gene	Best fit model	BUSCO gene	Best fit model	BUSCO gene	Best fit model
189234at4751	JTT+R5	332891at4751	mtMet+F+R5	464711at4751	mtInv+F+I+G4
189573at4751	JTT+F+R4	334424at4751	JTT+I+G4	464772at4751	JTT+R7
190183at4751	JTT+R4	335237at4751	JTT+R6	464897at4751	LG+R3
190475at4751	JTT+R5	335719at4751	JTT+I+G4	465849at4751	JTT+I+G4
190832at4751	JTT+R4	336181at4751	JTT+I+G4	46666at4751	JTT+F+R6
191114at4751	JTTDCMut+R5	336187at4751	JTTDCMut+I+G4	467311at4751	JTT+I+G4
191296at4751	JTT+R5	337150at4751	JTT+R5	46752at4751	JTT+F+R5
191310at4751	JTT+R4	337917at4751	JTTDCMut+I+G4	467847at4751	JTT+R5
192059at4751	JTT+R6	338301at4751	JTTDCMut+R5	467959at4751	JTT+I+G4
192135at4751	JTT+R6	338808at4751	JTTDCMut+R4	468195at4751	JTT+F+R5
192452at4751	JTTDCMut+R6	339525at4751	JTT+I+G4	468224at4751	JTT+R4
195619at4751	JTTDCMut+R4	33984at4751	JTT+R6	468408at4751	HIVb+F+R5
195863at4751	JTT+R5	34223at4751	HIVb+F+R7	46918at4751	JTT+R6
196015at4751	JTT+R4	342642at4751	LG+I+G4	469452at4751	LG+G4
196094at4751	JTTDCMut+R4	343196at4751	JTT+R5	470813at4751	JTTDCMut+I+G4
196841at4751	LG+R5	343603at4751	JTT+I+G4	471205at4751	JTT+I+G4
197111at4751	JTT+I+G4	343623at4751	JTT+I+G4	473814at4751	JTT+F+I+G4
197612at4751	JTT+F+R5	343717at4751	JTT+R5	474125at4751	JTT+I+G4
198196at4751	JTTDCMut+R5	344553at4751	JTT+I+G4	474320at4751	JTTDCMut+F+I+G4
198234at4751	JTT+R4	345661at4751	JTT+R5	474374at4751	rtREV+R2
200331at4751	JTT+I+G4	34706at4751	JTT+R5	47592at4751	JTTDCMut+R4
201591at4751	JTT+I+G4	348020at4751	JTT+R5	477057at4751	JTT+I+G4
202089at4751	JTT+I+G4	348111at4751	JTT+R5	477122at4751	LG+F+I+G4
202248at4751	JTT+R8	34830at4751	JTT+R6	481052at4751	LG+I+G4
203109at4751	JTT+R4	349264at4751	Dayhoff+I+G4	482337at4751	JTT+F+I+G4
203710at4751	JTT+I+G4	350840at4751	JTT+I+G4	483049at4751	JTT+I+G4
204087at4751	LG+I+G4	351351at4751	JTT+F+R5	483370at4751	JTT+I+G4
204341at4751	JTT+R5	351482at4751	LG+I+G4	483390at4751	JTT+G4
204508at4751	JTT+F+R5	351646at4751	JTT+R5	484080at4751	JTT+I+G4
205479at4751	JTT+I+G4	351651at4751	JTT+R5	484323at4751	JTT+I+G4
20600at4751	JTT+R4	351849at4751	JTTDCMut+I+G4	484805at4751	JTTDCMut+R6
207926at4751	JTT+F+R5	351858at4751	LG+R3	485650at4751	JTT+I+G4
208463at4751	JTT+R6	351906at4751	JTTDCMut+I+G4	485702at4751	JTT+F+I+G4
209053at4751	JTT+R4	352224at4751	JTTDCMut+I+G4	48708at4751	LG+R7
20952at4751	JTT+R5	352999at4751	JTT+I+G4	488348at4751	mtMet+F+I+G4
209619at4751	JTT+R6	3534at4751	JTT+R5	488426at4751	JTT+R4
210178at4751	JTT+I+G4	353685at4751	JTT+R4	488857at4751	JTT+I+G4
210462at4751	JTT+I+G4	353762at4751	JTT+I+G4	490662at4751	JTT+I+G4
210986at4751	JTT+F+I+G4	353991at4751	JTT+R6	490719at4751	LG+I+G4
211442at4751	JTT+R5	354679at4751	JTT+F+I+G4	492591at4751	JTT+I+G4
211634at4751	JTTDCMut+I+G4	354739at4751	JTT+I+G4	493358at4751	DCMut+I+G4
211768at4751	LG+R4	354778at4751	JTT+R4	494427at4751	JTT+G4
212522at4751	JTT+R6	354913at4751	JTT+R3	49476at4751	JTT+R5
212735at4751	JTTDCMut+R5	355454at4751	HIVb+F+R6	494at4751	JTT+R6
213564at4751	JTT+F+R5	355525at4751	JTTDCMut+I+G4	497024at4751	JTTDCMut+I+G4
215926at4751	JTTDCMut+R5	355781at4751	JTT+I+G4	49967at4751	JTTDCMut+R6
216270at4751	LG+R4	356079at4751	JTT+I+G4	50786at4751	LG+R5
21651at4751	JTT+F+R5	356691at4751	JTT+F+R9	50835at4751	JTT+R4
216661at4751	JTT+I+G4	356743at4751	LG+I+G4	52541at4751	JTT+R5
216683at4751	JTT+I+G4	357182at4751	JTT+R4	52909at4751	JTT+R5
216747at4751	JTT+R4	357584at4751	JTTDCMut+I+G4	53270at4751	JTT+R9
21689at4751	JTT+R6	358012at4751	LG+F+I+G4	54252at4751	JTT+R5
216949at4751	JTT+F+I+G4	358118at4751	JTT+R6	54919at4751	JTTDCMut+I+G4
217097at4751	JTT+I+G4	359271at4751	JTT+R5	56395at4751	JTT+R5
217374at4751	LG+I+G4	359482at4751	JTT+I+G4	56432at4751	JTT+R5
219236at4751	LG+R3	359627at4751	JTT+F+I+G4	56729at4751	JTT+R4
219492at4751	JTT+R8	360800at4751	JTT+R5	57081at4751	JTTDCMut+R4
221926at4751	JTT+R5	361877at4751	JTTDCMut+R5	58127at4751	JTT+R8
224693at4751	JTT+F+R5	362218at4751	JTTDCMut+I+G4	59795at4751	JTT+F+R4
224775at4751	JTTDCMut+R4	362250at4751	LG+R4	5995at4751	JTT+R5
224804at4751	JTT+I+G4	362322at4751	JTTDCMut+I+G4	61411at4751	JTT+R7
225321at4751	mtZOA+F+I+G4	362601at4751	JTT+F+R5	61438at4751	JTT+R5
225351at4751	JTTDCMut+R5	363593at4751	JTTDCMut+I+G4	61857at4751	JTT+I+G4
225535at4751	JTT+R4	364123at4751	JTT+I+G4	6186at4751	JTT+F+R5

BUSCO gene	Best fit model	BUSCO gene	Best fit model	BUSCO gene	Best fit model
225712at4751	JTTDCMut+I+G4	364284at4751	JTT+R6	6272at4751	JTT+R5
225820at4751	JTT+F+R5	365270at4751	JTTDCMut+R6	62846at4751	JTTDCMut+R5
225918at4751	JTT+R6	36610at4751	JTT+R6	63635at4751	JTT+R5
226810at4751	JTT+R4	366595at4751	LG+I+G4	6377at4751	JTTDCMut+I+G4
227572at4751	JTT+I+G4	366597at4751	JTT+R5	63838at4751	JTT+R5
228465at4751	JTT+I+G4	366635at4751	JTT+I+G4	64002at4751	JTT+R5
228752at4751	JTT+I+G4	367778at4751	JTT+I+G4	64498at4751	JTTDCMut+R4
229510at4751	JTT+R10	368706at4751	JTT+R5	65334at4751	JTT+R4
22982at4751	LG+R7	369435at4751	LG+R4	65499at4751	JTT+R4
22992at4751	HIVb+F+R6	37028at4751	LG+R4	6592at4751	JTT+R5
229997at4751	JTT+F+R5	370505at4751	LG+I+G4	66151at4751	JTTDCMut+R5
230605at4751	JTT+R4	370952at4751	JTT+R4	67284at4751	JTT+R5
231468at4751	JTT+F+R5	371481at4751	Dayhoff+I+G4	67399at4751	LG+R7
231582at4751	JTT+R5	371674at4751	JTT+I+G4	67806at4751	JTT+R4
23198at4751	JTT+R5	371862at4751	JTT+F+R5	6812at4751	JTT+R5
232698at4751	JTT+R4	372381at4751	JTT+I+G4	68466at4751	JTT+R4
233287at4751	JTTDCMut+R4	372510at4751	JTT+F+R5	68881at4751	HIVb+F+R6
233706at4751	JTTDCMut+I+G4	372972at4751	JTTDCMut+F+I+G4	71039at4751	JTTDCMut+R5
234548at4751	JTT+R5	374350at4751	JTT+I+G4	71046at4751	JTT+R5
234698at4751	JTT+I+G4	374548at4751	JTT+R6	73383at4751	JTT+R5
235105at4751	JTT+R5	37573at4751	JTT+I+G4	73477at4751	JTT+I+G4
235463at4751	LG+R4	376255at4751	JTT+R4	73771at4751	JTT+F+R5
235640at4751	JTTDCMut+R5	37642at4751	JTT+R6	7423at4751	mtMet+F+R5
236504at4751	JTTDCMut+I+G4	376442at4751	JTT+F+R6	75033at4751	JTT+F+R5
237276at4751	JTTDCMut+I+G4	376539at4751	JTT+R5	75456at4751	JTT+F+R5
237281at4751	JTT+R5	376586at4751	JTT+R5	7545at4751	JTT+F+R6
237393at4751	JTTDCMut+R5	377787at4751	LG+I+G4	75463at4751	JTT+F+R5
237533at4751	JTT+I+G4	378256at4751	JTT+I+G4	75709at4751	JTT+F+R6
23856at4751	JTT+R4	378281at4751	JTT+I+G4	75976at4751	JTT+F+R5
238922at4751	JTT+F+R5	378295at4751	JTT+I+G4	76196at4751	JTTDCMut+F+R5
240036at4751	JTT+F+I+G4	378376at4751	JTTDCMut+R4	76893at4751	JTT+I+G4
240246at4751	JTT+R5	378430at4751	JTT+R5	76975at4751	JTT+R4
241079at4751	JTT+R4	378436at4751	JTT+R5	771at4751	JTT+R5
241483at4751	LG+I+G4	379525at4751	JTT+I+G4	78289at4751	JTT+R4
241586at4751	JTTDCMut+R4	38031at4751	JTT+R5	78665at4751	JTT+R4
241860at4751	JTT+R5	38184at4751	HIVb+F+R6	79262at4751	JTT+R4
24198at4751	JTT+R6	382001at4751	JTT+I+G4	79294at4751	LG+R4
243015at4751	JTT+R4	383007at4751	JTT+R5	79797at4751	JTT+R6
244066at4751	JTT+R6	384156at4751	LG+R3	79988at4751	JTT+R6
245480at4751	JTT+R5	384315at4751	LG+I+G4	81295at4751	JTTDCMut+R5
245900at4751	LG+F+I+G4	385139at4751	LG+R3	81614at4751	JTT+F+R5
246207at4751	JTT+R3	385343at4751	JTT+F+I+G4	81687at4751	LG+R5
246951at4751	JTT+F+R5	385564at4751	mtMet+F+I+G4	8406at4751	JTT+R5
247065at4751	JTT+R5	385971at4751	JTT+I+G4	84250at4751	JTT+R6
247129at4751	JTT+R4	386220at4751	JTT+G4	84829at4751	JTT+F+R6
247383at4751	JTT+R4	386245at4751	JTT+R5	85793at4751	JTT+R6
248360at4751	JTT+F+R5	387347at4751	JTT+I+G4	86293at4751	JTT+R5
25020at4751	JTT+R5	387759at4751	FLU+R4	87251at4751	JTT+F+R5
251158at4751	JTTDCMut+I+G4	388229at4751	LG+I+G4	87313at4751	JTTDCMut+R4
251613at4751	JTT+R5	389034at4751	LG+F+I+G4	8818at4751	JTT+R4
25161at4751	JTT+R5	389547at4751	JTT+I+G4	90022at4751	JTT+F+R4
252314at4751	JTT+I+G4	389565at4751	JTT+I+G4	91135at4751	JTT+R6
252469at4751	JTT+R4	390518at4751	JTT+I+G4	92267at4751	LG+R4
254717at4751	Dayhoff+F+I+G4	391384at4751	JTT+R5	93783at4751	JTT+F+R5
255378at4751	JTT+I+G4	392558at4751	JTT+R5	94384at4751	JTTDCMut+R5
255412at4751	JTTDCMut+I+G4	393440at4751	JTT+I+G4	94628at4751	JTTDCMut+F+R5
256318at4751	JTTDCMut+R5	393996at4751	LG+I+G4	95253at4751	JTT+R5
256914at4751	JTT+R5	396097at4751	JTT+R6	96633at4751	JTT+R5
257330at4751	JTT+I+G4	396330at4751	JTT+R5	97141at4751	JTT+R5
258245at4751	JTTDCMut+I+G4	396842at4751	JTTDCMut+R4	98519at4751	JTT+R5
258762at4751	JTT+I+G4	397465at4751	JTT+R5	98733at4751	JTTDCMut+R5
259060at4751	JTT+R5	39833at4751	JTT+F+R5	98739at4751	JTTDCMut+R5
259444at4751	JTT+F+R5	398519at4751	JTTDCMut+F+R6		

Supplementary Table 3 Distribution of SM backbone genes

Species	NRPS	Meroterpenoid synthase	terpene	t3PKS	NRPS-PKS hybrid	Diterpene synthase	Carotenoid synthase	NRPS-like	Squalene synthase	UbiA-type terpene	Sesquiterpene synthase	rPKS	NIS	other PKS	nrPKS	Triterpene synthase	Terpene other	PKS-NRPS hybrid	DMATS
Aspac1	19	1	0	1	0	0	1	29	3	3	5	16	1	2	5	3	0	4	6
Aspacri1	20	1	0	0	1	0	0	27	2	2	4	20	0	0	14	2	0	2	7
Aspacu1	26	1	0	1	0	0	2	31	1	2	6	24	0	3	6	3	0	4	10
Aspaff1	19	1	0	0	0	0	1	30	1	5	11	36	0	1	6	3	0	10	14
Aspala1	16	0	0	0	0	1	0	18	2	0	7	11	4	3	11	3	1	5	10
Aspalbe1	19	3	0	1	0	0	2	19	1	0	6	31	1	2	11	6	0	4	5
Aspall1	15	0	0	0	0	1	1	27	3	1	1	15	1	0	7	1	1	6	12
Aspamb1	14	0	0	0	0	0	1	22	2	0	3	15	3	0	9	0	1	2	7
Aspamoe1	22	2	0	0	0	1	1	18	2	1	3	16	0	0	11	2	0	4	13
Aspamy1	11	2	0	1	0	1	0	16	1	4	7	23	1	0	12	1	1	3	3
Aspara19utr	23	1	0	4	0	1	1	24	1	2	4	13	1	0	8	4	0	3	9
Asparx1	6	0	0	1	0	0	0	11	1	0	1	8	2	1	4	3	0	1	1
Aspass1	7	2	0	1	0	0	0	9	1	1	5	4	1	2	6	1	0	1	4
Aspaur1	16	0	0	0	0	2	1	21	1	1	11	12	0	0	6	5	1	4	11
Aspaurful1	18	0	0	2	0	0	0	24	2	1	3	11	0	0	4	1	0	1	12
Aspaurlut1	9	2	0	1	0	0	0	10	1	1	3	6	1	0	4	2	0	2	5
Aspave1	12	0	0	1	0	0	0	22	1	0	11	18	0	0	5	2	0	4	2
Aspber1	15	0	0	2	1	2	1	14	1	1	8	16	0	2	8	3	0	5	7
Aspbip1	17	3	0	2	0	1	0	27	2	6	20	31	1	4	11	4	0	1	4
Aspbisp1	12	1	0	0	0	0	0	16	1	0	2	12	0	1	6	0	0	1	3
Aspbom1	22	1	0	4	2	3	1	26	1	1	5	18	1	1	8	4	0	4	6
Aspbot1	19	5	0	0	0	0	0	13	1	2	3	11	1	1	7	3	0	3	13
Aspbr1	14	1	0	1	0	1	1	26	2	3	7	28	1	1	8	2	0	5	2
Aspbre1	19	3	0	0	0	0	0	19	1	1	6	14	2	0	7	2	0	4	7
Aspbrev1	25	2	0	2	0	2	0	37	2	2	8	33	4	2	7	1	0	5	12
Aspbrevi1	9	1	0	1	0	0	0	12	1	1	3	8	1	0	5	1	0	1	5

Species	NRPS	Meroterpenoid synthase	terpene	t3PKS	NRPS-PKS hybrid	Diterpene synthase	Carotenoid synthase	NRPS-like	Squalene synthase	UbiA-type terpene	Sesquiterpene synthase	rPKS	NIS	other PKS	nrPKS	Triterpene synthase	Terpene other	PKS-NRPS hybrid	DMATS
Aspbru1	29	1	0	1	0	0	2	30	1	3	4	20	1	2	7	3	0	3	10
Aspca3	13	2	0	0	0	0	0	23	2	3	9	14	0	0	5	3	0	5	2
Aspcae1	22	1	0	5	1	2	1	28	1	3	10	16	1	2	11	3	0	5	9
Aspcalif1	23	3	0	0	0	1	0	23	1	5	11	21	2	0	14	4	0	5	10
Aspcam1	16	1	0	0	1	0	0	14	1	0	4	14	0	0	1	2	0	4	9
Aspcand1	13	1	0	0	2	0	0	14	1	0	4	15	0	0	1	2	0	2	6
Aspcanin1	3	0	0	0	0	0	0	2	2	0	0	5	0	0	5	1	0	0	0
Aspcap1	22	1	0	1	0	0	0	33	2	2	5	10	0	2	6	3	0	4	6
Aspcar1	10	2	0	1	0	0	0	30	1	2	8	22	1	0	12	2	0	2	8
Aspcer1	17	0	0	1	0	2	1	24	1	1	7	20	1	2	6	2	0	1	3
Aspchev1	5	0	0	0	0	0	0	9	1	0	2	9	2	0	5	2	0	1	3
Aspchr1	21	0	0	2	1	1	1	38	2	3	9	33	0	2	11	3	0	12	10
Aspcl1	14	0	0	0	0	0	0	6	1	0	6	13	0	0	3	2	0	5	3
Aspcla1	15	2	0	0	0	0	1	10	1	1	3	14	0	1	5	1	0	4	4
Aspcon1	11	2	0	1	0	0	0	25	1	6	13	21	1	2	15	3	0	1	5
Aspcor1	6	0	0	1	0	0	0	10	1	0	4	12	0	0	5	3	0	0	6
Aspcorea1	18	2	0	0	0	0	0	12	1	2	3	10	0	2	6	2	0	1	12
Aspcos1	21	1	0	1	1	2	3	29	2	2	11	33	2	2	6	3	0	4	3
Aspcre1	20	0	0	0	0	0	2	29	1	5	9	30	0	1	4	4	0	10	12
Aspcreb1	24	2	0	0	0	0	0	25	2	1	4	17	0	0	11	2	0	4	11
Aspcri1	8	0	0	0	0	0	0	13	2	0	5	7	2	0	5	2	0	0	3
Aspcru1	7	3	0	0	0	0	0	15	2	5	4	10	0	2	10	3	0	0	9
Aspdese1	8	1	0	0	0	2	0	17	2	1	6	11	0	2	12	1	0	4	6
Aspdur1	11	4	0	0	0	0	0	12	1	3	9	17	2	0	10	2	0	2	8
Aspegy1	17	0	0	0	0	1	0	13	1	0	9	14	2	2	6	3	1	0	6
Aspele1	22	0	0	1	1	1	2	35	1	2	11	30	2	1	6	2	0	9	11
Aspell1	16	2	0	0	0	0	1	17	1	2	5	21	0	3	7	4	0	2	2
Aspeuc1	13	1	0	1	1	1	2	27	2	2	10	26	1	1	4	2	0	6	4

Species	NRPS	Meroterpenoid synthase	terpene	t3PKS	NRPS-PKS hybrid	Diterpene synthase	Carotenoid synthase	NRPS-like	Squalene synthase	UbiA-type terpene	Sesquiterpene synthase	rPKS	NIS	other PKS	nrPKS	Triterpene synthase	Terpene other	PKS-NRPS hybrid	DMATS
Aspfer1	10	1	0	2	0	0	0	13	1	1	4	8	0	0	7	1	0	2	6
Aspfij1	31	1	0	1	0	0	1	31	2	3	5	22	1	2	7	3	0	5	10
Aspfili1	14	0	0	0	0	1	0	15	3	1	5	22	2	0	14	2	0	1	7
Aspfl1	20	1	0	2	0	1	1	22	1	0	7	16	1	1	11	3	0	3	8
Aspfla1	25	1	0	1	0	0	1	41	2	4	5	25	0	2	8	2	0	4	12
Aspflo1	25	2	0	0	0	2	1	37	2	2	4	20	2	1	9	2	1	6	13
Aspfloc1	27	0	0	0	0	0	2	34	1	2	10	28	1	1	7	2	0	6	8
Aspfo1	20	1	0	2	2	1	1	29	3	1	9	29	2	1	6	3	0	5	2
Aspfre1	24	1	0	1	0	0	1	35	2	2	4	13	0	2	5	3	0	4	12
Aspfu1	14	2	0	0	0	0	0	9	1	3	1	7	0	1	6	5	0	1	7
Aspfumig1	31	3	0	0	0	0	0	15	1	3	4	17	2	0	13	8	0	2	13
Aspfumis1	22	2	0	0	0	0	0	8	1	2	2	13	2	0	7	7	0	4	9
Aspfuni1	7	1	0	0	0	0	0	8	1	2	3	8	0	0	9	1	0	0	6
Aspgal1	10	2	0	1	0	0	0	12	2	1	6	9	2	1	5	3	0	2	6
Aspger1	13	2	0	1	0	0	0	25	1	4	8	17	0	0	12	2	0	0	8
Aspgig1	15	1	0	0	0	0	1	11	1	1	2	14	0	0	6	1	0	3	3
Aspgl1	7	0	0	0	0	0	1	14	1	0	1	8	1	2	5	2	0	0	3
Aspgra1	8	1	0	0	0	1	0	15	2	3	6	10	2	2	10	3	0	1	7
Asphai1	8	1	0	1	0	1	1	14	1	2	17	20	1	2	11	3	0	1	5
Asphet1	13	2	0	0	0	0	1	19	1	1	4	13	0	3	5	3	0	3	4
Asphethal1	11	0	0	0	0	0	0	20	1	2	5	13	1	0	11	2	0	0	4
Asphir1	11	2	0	1	0	0	0	12	1	1	5	7	1	0	3	2	0	2	5
Asphom1	19	3	0	1	0	0	2	32	1	2	4	17	0	1	7	3	0	3	8
Asphor1	18	2	0	0	0	0	1	28	2	1	8	17	2	2	12	1	1	4	11
Aspibe1	8	2	0	0	0	0	1	18	1	4	8	13	0	0	5	2	0	6	2
Aspign1	20	2	0	0	0	0	0	10	1	2	1	12	2	1	13	7	0	4	6
Aspiiz1	24	1	0	1	0	0	0	33	2	2	4	11	0	0	8	4	0	6	12
Aspimp1	24	0	0	1	1	0	1	35	2	1	21	42	0	2	15	3	0	7	5

Species	NRPS	Meroterpenoid synthase	terpene	t3PKS	NRPS-PKS hybrid	Diterpene synthase	Carotenoid synthase	NRPS-like	Squalene synthase	UbiA-type terpene	Sesquiterpene synthase	rPKS	NIS	other PKS	nrPKS	Triterpene synthase	Terpene other	PKS-NRPS hybrid	DMATS
Aspind1	24	1	0	1	1	1	2	38	1	2	4	28	1	0	4	4	0	5	7
Aspind2	14	1	0	0	1	1	0	20	2	1	7	21	0	1	10	2	0	2	7
Aspins1	10	2	0	1	0	0	0	24	1	4	8	21	0	0	10	2	0	1	7
Aspinso1	9	0	0	0	0	0	0	8	1	1	1	5	0	1	4	1	0	1	0
Aspjap1	24	1	0	1	0	1	1	28	2	2	4	20	1	1	4	3	0	5	6
Aspka1	15	1	0	1	2	0	1	28	4	2	9	28	1	3	6	3	0	4	2
Aspkar1	18	2	0	0	0	0	0	31	1	3	7	18	2	1	12	4	0	5	12
Aspkev1	13	1	0	1	0	0	0	26	1	4	10	20	1	1	10	3	0	0	9
Asplaci1	24	2	0	0	0	0	0	13	1	2	3	9	0	0	8	1	0	3	11
Asplen1	17	4	0	0	0	1	0	11	1	2	3	10	1	1	10	5	0	2	12
Asplon1	15	2	0	0	0	0	1	10	1	1	2	8	0	1	4	1	0	4	3
Aspluc1	8	1	0	0	0	1	0	10	2	1	5	17	2	0	12	2	0	3	4
Asplup1	17	0	0	2	0	0	0	27	2	1	3	11	0	0	4	1	0	1	12
Aspmic1	11	0	0	0	0	0	0	15	3	1	3	8	0	1	8	1	0	2	0
Aspmin1	19	1	0	3	0	1	1	28	1	2	4	21	2	0	9	3	0	3	8
Aspmul1	20	1	0	0	0	1	0	20	1	1	8	19	1	5	10	3	0	3	5
Aspmult1	5	2	0	0	0	0	0	8	1	1	6	3	2	0	3	2	0	1	2
Aspmuri1	17	1	0	1	0	0	2	30	1	2	8	27	0	1	6	2	0	5	8
Aspnaka1	16	0	0	1	0	1	1	30	1	2	15	34	0	0	5	2	0	5	9
Aspne1	15	3	0	0	0	0	0	21	2	4	6	18	0	0	16	3	0	0	7
Aspneo1	16	1	0	1	1	0	3	27	2	3	13	31	1	2	4	3	0	4	3
Aspneof1	25	2	0	1	0	0	1	37	2	3	3	16	0	2	6	1	0	4	15
Aspneoi1	21	1	0	1	0	0	0	30	2	0	2	14	1	0	10	1	1	4	11
Aspni	17	1	0	1	1	2	1	24	2	4	9	27	2	0	8	4	0	8	2
Aspnid1	13	1	0	0	0	0	0	20	2	2	4	16	0	1	13	3	0	1	7
Aspnis1	4	2	0	1	0	0	0	12	1	1	3	3	0	0	5	1	0	1	2
Aspniv1	20	0	0	0	0	1	1	30	3	0	3	16	1	0	7	1	0	4	16
Aspnom1	20	1	0	5	1	1	1	30	1	2	5	17	1	1	7	4	0	2	8

Species	NRPS	Meroterpenoid synthase	terpene	t3PKS	NRPS-PKS hybrid	Diterpene synthase	Carotenoid synthase	NRPS-like	Squalene synthase	UbiA-type terpene	Sesquiterpene synthase	rPKS	NIS	other PKS	nrPKS	Triterpene synthase	Terpene other	PKS-NRPS hybrid	DMATS
Aspnov1	21	5	0	0	0	0	0	11	1	5	3	15	3	1	12	5	0	4	13
Aspnovo1	19	1	0	4	0	1	1	27	2	1	4	19	1	0	10	4	0	1	6
Aspnut1	29	1	0	0	0	1	0	32	1	2	6	12	1	0	5	3	0	3	4
Aspoc7043	18	1	0	0	0	1	1	29	1	2	7	28	1	1	6	2	0	7	9
Aspoch1	9	0	0	0	0	1	0	6	1	0	1	4	0	1	7	0	0	2	6
Aspoli1	11	1	0	0	0	1	0	15	3	2	4	21	0	1	16	3	0	2	6
Aspor1	17	1	0	4	0	1	1	21	1	0	5	16	1	0	11	3	0	2	7
Asposti1	17	0	0	0	0	0	1	26	1	2	6	31	0	0	5	1	0	6	11
Asppar1	24	1	0	3	0	1	1	25	1	3	9	19	2	0	10	3	0	1	8
Aspparv1	22	1	0	0	0	1	0	30	1	2	4	12	1	3	5	3	1	2	7
Asppdef1	9	2	0	1	0	0	0	23	2	3	11	23	0	0	11	4	0	1	6
Asppeni1	8	0	0	0	1	0	0	10	1	0	0	8	2	3	4	0	0	0	2
Aspper1	23	0	0	2	1	1	2	32	3	2	7	37	1	2	12	3	0	9	7
Asppet1	19	1	0	0	0	1	1	26	1	1	7	30	1	3	6	2	0	5	10
Asppip1	16	1	0	1	2	1	2	27	3	2	13	28	2	2	6	3	0	5	2
Asppsec1	19	1	0	5	1	1	1	27	1	3	9	13	2	1	9	3	0	5	9
Asppsen1	19	1	0	5	1	1	1	30	1	1	5	16	1	1	7	4	0	2	8
Asppset1	17	1	0	4	1	0	1	26	1	1	6	17	3	0	11	3	0	5	9
Asppseute1	16	3	0	0	0	1	0	26	2	2	7	15	1	3	10	1	1	6	11
Asppul1	22	0	0	0	0	0	1	33	2	2	9	31	1	0	7	2	0	7	10
Asppust1	16	2	0	1	0	0	0	19	1	3	8	17	4	0	14	3	0	1	9
Asppuu1	20	2	0	0	0	0	0	24	2	1	4	14	0	1	13	2	0	3	10
Aspram1	9	1	0	0	0	0	0	5	1	0	1	6	0	0	7	1	0	1	4
Asprec1	13	1	0	0	0	3	0	20	2	2	7	21	0	0	14	2	1	2	5
Aspres1	5	0	0	0	0	0	0	13	1	1	0	5	1	0	2	0	0	1	0
Asprhi1	18	2	0	0	0	0	1	10	1	1	3	16	0	0	6	1	0	4	4
Asprobu1	31	2	0	1	1	0	1	23	1	4	9	16	0	1	4	1	0	4	2
Aspros1	21	0	1	1	0	0	1	32	1	3	13	32	1	0	8	3	0	8	10

Species	NRPS	Meroterpenoid synthase	terpene	t3PKS	NRPS-PKS hybrid	Diterpene synthase	Carotenoid synthase	NRPS-like	Squalene synthase	UbiA-type terpene	Sesquiterpene synthase	rPKS	NIS	other PKS	nrPKS	Triterpene synthase	Terpene other	PKS-NRPS hybrid	DMATS
Aspsac1	15	1	0	0	0	0	1	20	2	1	2	9	0	1	3	2	0	2	4
Aspsc1	21	0	0	1	1	1	1	39	2	3	9	33	0	2	11	3	0	12	9
Aspscl1	11	2	0	0	0	1	1	24	1	4	8	21	0	0	8	4	0	5	2
Aspscle1	15	1	0	0	0	0	1	26	1	3	9	26	0	0	6	2	0	6	3
Aspsep1	11	1	0	0	0	0	2	26	1	1	11	25	1	1	4	3	0	2	0
Aspser1	22	1	0	4	0	1	1	29	1	3	5	19	2	2	11	4	0	2	7
Aspses1	15	0	0	1	0	1	1	30	1	0	8	29	0	0	5	2	0	5	8
Aspsil1	12	2	0	0	0	0	0	16	1	2	3	13	0	0	12	1	1	5	7
Aspspe1	15	1	0	0	0	1	0	32	1	3	4	26	1	3	13	4	0	2	7
Aspste1	24	0	0	1	1	1	2	32	1	2	11	31	2	0	6	2	0	9	11
Aspstel1	13	2	0	0	0	1	0	20	3	2	8	29	2	0	14	4	0	1	6
Aspsubr1	23	0	0	1	1	1	1	37	2	3	13	41	1	0	11	4	0	12	8
Aspsy1	14	1	0	0	0	0	0	24	1	1	2	9	0	0	9	4	0	2	7
Asptaic1	15	0	0	0	0	0	0	13	1	0	6	12	0	1	2	2	0	3	8
Asptam1	22	0	0	6	2	0	2	27	1	1	2	17	2	0	8	4	0	4	12
Asptan1	32	2	1	4	0	1	1	28	1	3	8	27	1	3	16	4	0	10	10
Aspte1	20	2	0	0	0	1	0	20	2	1	9	15	1	4	9	1	1	3	10
Asptem1	20	1	0	1	0	0	1	36	2	2	3	19	0	5	5	3	0	3	11
Aspten1	16	1	0	0	0	0	0	23	2	1	5	14	1	2	9	2	0	2	8
Aspth1	8	1	0	1	0	0	0	21	1	3	7	16	0	0	9	4	0	1	8
Asptr1	17	0	0	0	1	0	0	19	1	0	6	23	0	0	4	1	0	1	4
Asptra1	20	1	0	5	0	1	1	26	1	2	5	21	2	0	11	3	0	2	8
Asptran1	21	0	0	1	0	2	0	23	1	2	11	17	2	1	3	2	0	2	4
Asptu1	16	1	0	1	1	1	2	26	2	4	9	35	2	2	5	3	0	6	4
Aspuda1	22	5	0	0	0	0	0	18	1	2	3	16	2	0	15	7	0	3	8
Aspund1	11	1	0	0	0	0	0	16	2	3	4	11	1	4	10	3	0	1	5
Aspung1	13	0	0	0	0	0	0	15	3	1	4	9	0	0	8	2	0	2	1
Aspuva1	22	1	0	1	0	1	0	29	1	3	4	17	1	1	6	3	0	3	7

Species	NRPS	Meroterpenoid synthase	terpene	t3PKS	NRPS-PKS hybrid	Diterpene synthase	Carotenoid synthase	NRPS-like	Squalene synthase	UbiA-type terpene	Sesquiterpene synthase	rPKS	NIS	other PKS	nrPKS	Triterpene synthase	Terpene other	PKS-NRPS hybrid	DMATS
Aspvad1	16	1	0	1	1	1	2	28	2	4	7	28	1	2	4	3	0	5	3
Aspve1	32	2	0	0	0	1	0	22	2	1	2	15	0	0	10	2	0	2	10
Aspven1	15	2	0	0	0	1	0	19	3	4	7	23	2	1	19	4	0	2	8
Aspvio1	25	1	0	1	0	1	1	27	2	2	5	19	1	2	4	3	0	5	6
Aspwe1	14	1	0	0	0	1	1	24	1	1	9	19	1	0	4	4	0	3	0
Aspwel1	16	0	0	1	1	1	1	28	2	2	10	28	1	0	7	3	0	6	2
Eurhe1	6	0	0	0	0	0	1	10	1	0	3	6	2	1	5	3	0	0	4
Neofi1	21	3	0	0	0	0	0	12	1	2	2	9	2	0	7	5	0	1	10
Penbr2	21	1	0	1	0	0	0	22	1	1	5	19	0	0	7	1	0	1	2
Pench1	12	2	0	1	1	0	0	15	1	2	10	16	1	0	4	1	0	1	1
Pengri1	16	1	0	1	1	1	0	16	1	1	6	19	0	0	6	1	0	3	7

Supplementary Table 4 Distribution of Orthologous groups deposited at Zenodo (DOI: 10.5281/zenodo.10638673)

Supplementary Table 5 SM backbone genes with known metabolites on MIBiG and their RNA and H3K4me3 CVs

A. fumigatus

No.	Gene	Known metabolite	RNA CVs	H3K4me3 CVs
1	ozq3230	YWA1	0.47	0.03
2	ozq3269	fumigaclavine C	1.70	0.31
3	ozq3274	fumigaclavine C	1.67	0.28
4	ozq4427	hexadehydroastechrome/terezine-D/astechrome	0.67	0.10
5	ozq4428	hexadehydroastechrome/terezine-D/astechrome	0.38	0.03
6	ozq4678	endocrocin	0.89	0.19
7	ozq5642	clavaric acid	0.65	0.27
8	ozq5879	trypacidin	1.04	0.31
9	ozq5912	fumihopaside A/compound 3/21-β-H-hopane-3beta, 22-diol	2.37	0.63
10	ozq6307	clavaric acid	0.40	0.13
11	ozq8266	fumiquinazoline A/fumiquinazoline C/fumiquinazoline D	1.43	0.50
12	ozq8269	fumiquinazoline A/fumiquinazoline C/fumiquinazoline D	0.76	0.26
13	ozq8446	pyripyropene A	1.38	0.35
14	ozq8447	pyripyropene A	0.89	0.18
15	ozq8450	pyripyropene A	1.27	0.74
16	ozq8452	pyripyropene A	1.14	0.68
17	ozq9170	fumagillin/β-trans-bergamotene/fumagillol	1.66	0.08
18	ozq9173	fumagillin/β-trans-bergamotene/fumagillol	1.40	0.27
19	ozq9177	fumagillin/β-trans-bergamotene/fumagillol	1.45	0.09
20	ozq9189	fumagillin/β-trans-bergamotene/fumagillol	0.96	0.13
21	ozq9201	fumagillin/β-trans-bergamotene/fumagillol	1.15	0.42
22	ozq9203	fumagillin/β-trans-bergamotene/fumagillol	1.51	0.24
23	ozq9205	fumagillin/β-trans-bergamotene/fumagillol	0.66	0.32
24	ozq9214	fumagillin/β-trans-bergamotene/fumagillol	0.63	0.31
25	ozq9385	sartorypyrone A	1.07	0.16
26	ozq9388	sartorypyrone A	2.63	0.07
27	ozq9390	sartorypyrone A	2.24	0.13

Summary

Average CV for SM known genes	1.20	0.27
Average CV for SM genes	1.09	0.24
Average CV for SM genes in sub-telomeric regions	1.21	0.26
Average CV for SM genes outside sub-telomeric regions	0.94	0.21

A. nidulans

No.	Gene	Known metabolite	RNA CVs	H3K4me3 CVs
1	zpw3780	cichorine	0.79	0.37
2	zpw3795	cichorine	0.59	0.38
3	zpw1666	naphthopyrone	0.31	0.23
4	zpw10599	choline	0.09	0.31
5	zpw6226	monascorubrin	1.04	0.34
6	zpw6805	cyclo-(D-Phe-L-Phe-D-Val-L-Val)	0.91	0.41

Summary

Average CV for SM known genes	0.62	0.34
Average CV for SM genes	0.88	0.35
Average CV for SM genes in sub-telomeric regions	0.99	0.42
Average CV for SM genes outside sub-telomeric regions	0.81	0.29

A. niger

No.	Gene	Known metabolite	RNA CVs
1	fin462	YWA1	0.27
2	fin531	clavaric acid	0.08
3	fin998	pyrophen/campyrone B	0.24
4	fin1334	AbT1	0.62
5	fin1339	AbT1	0.82

A. niger			
No.	**Gene**	**Known metabolite**	**RNA CVs**
6	fin6217	choline	2.90
7	fin6237	choline	0.17
8	fin6291	yanuthone D	1.54
9	fin9549	TAN-1612	0.41
10	fin10128	pyranonigrin E	0.50
11	fin10375	burnettiene A/preburnettiene B/preburnettiene A	1.42
		Summary	
		Average CV for SM known genes	0.82
		Average CV for SM genes	1.00
		Average CV for SM genes in sub-telomeric regions	1.21
		Average CV for SM genes outside sub-telomeric regions	0.92

A. oryzae			
No.	**Gene**	**Known metabolite**	**RNA CVs**
1	ogu1425	choline	0.06
2	ogu2271	leporin B	0.62
3	ogu2284	leporin B	0.11
4	ogu2300	leporin B	0.14
5	ogu4449	aflavarin	0.43
6	ogu5190	astellolide A	0.52
7	ogu5359	flavunoidine	3.70
8	ogu5360	flavunoidine	0.73
9	ogu5363	flavunoidine	0.35
10	ogu6602	clavaric acid	0.04
11	ogu6962	2,4'-dihydroxy-3'-methoxypropiophenone	0.13
12	ogu6963	2,4'-dihydroxy-3'-methoxypropiophenone	0.15
13	ogu7533	dichlorodiaporthin	1.10
14	ogu9094	actinopolymorphol C	0.12
15	ogu9136	clavaric acid	0.12
16	ogu9143	clavaric acid	0.44
17	ogu10870	6-methylsalicyclic acid	0.20
18	ogu11976	8-methyldiaporthin	0.70
		Summary	
		Average CV for SM known genes	0.53
		Average CV for SM genes	0.80
		Average CV for SM genes in sub-telomeric regions	1.00
		Average CV for SM genes outside sub-telomeric regions	0.73

Supplementary Table 6 chi-square statistic tests

SM genes	Genes with low genus-level syntenic score	Genes with high genus-level syntenic score	P value	Genes with low section-level syntenic score	Genes with high section-level syntenic score	P value
A. niger	99	2	9.00E-21	25	18	7.67E-09
A. nidulans	62	4	1.14E-14	29	16	9.76E-10
A. fumagatus	42	5	3.62E-13	17	12	6.16E-10
A. oryzae	76	3	1.79E-12	12	22	0.009530654
Sum	279	14	6.12E-58	83	68	7.77E-28

SM genes	Genes in subtelomeric regions	Genes outside of subtelomeric regions	P value
A. niger	39	74	0.000805618
A. nidulans	33	50	1.02E-05
A. fumagatus	31	26	3.53E-10
A. oryzae	22	69	0.605537499
Sum	125	219	1.28E-12

BUSCO genes	Genes with low genus-level syntenic score	Genes with high genus-level syntenic score	P value	Genes with low section-level syntenic score	Genes with high section-level syntenic score	P value
A. niger	21	630	4.78E-122	15	676	1.31E-32
A. nidulans	16	602	1.90E-96	13	666	2.71E-40
A. fumagatus	9	641	8.38E-75	7	714	6.00E-26
A. oryzae	18	593	1.81E-143	13	644	5.78E-24
Sum	64	2466	0	48	2700	3.15E-117

BUSCO genes	Genes in subtelomeric regions	Genes outside of subtelomeric regions	P value
A. niger	60	695	4.34E-18
A. nidulans	55	689	1.43E-16
A. fumagatus	39	706	5.23E-23
A. oryzae	53	656	8.47E-19
Sum	207	2746	4.34E-72

All genes	Genes with low genus-level syntenic score	Genes with high genus-level syntenic score	P value	Genes with low section-level syntenic score	Genes with high section-level syntenic score	P value
A. niger	5024	4867	/	1593	6023	/
A. nidulans	3897	4646	/	2024	6368	/
A. fumagatus	2904	4968	/	1298	7185	/
A. oryzae	5606	4391	/	1392	6778	/
Sum	17431	18872	/	6307	26354	/

All genes	Genes in subtelomeric regions	Genes outside of subtelomeric regions	P value
A. niger	2501	9345	/
A. nidulans	2108	8572	/
A. fumagatus	1959	7822	/
A. oryzae	2571	9447	/
Sum	9139	35186	/

Supplementary Table 7 RNA-seq data analysis

No.	Accession number	Description	Batch
		A. fumigatus	
1	SRX11913664	GSM5537237: Wild type rep-1; Aspergillus fumigatus Af293; RNA-Seq	batch_1
2	SRX11913665	GSM5537238: Wild type rep-2; Aspergillus fumigatus Af293; RNA-Seq	batch_1
3	SRX11913666	GSM5537239: Wild type rep-3; Aspergillus fumigatus Af293; RNA-Seq	batch_1
4	SRX11913667	GSM5537240: Elp3 rep-1; Aspergillus fumigatus Af293; RNA-Seq	batch_1
5	SRX11913668	GSM5537241: Elp3 rep-2; Aspergillus fumigatus Af293; RNA-Seq	batch_1
6	SRX11913669	GSM5537242: Elp3 rep-3; Aspergillus fumigatus Af293; RNA-Seq	batch_1
7	SRX11913670	GSM5537243: Ngn4 rep-1; Aspergillus fumigatus Af293; RNA-Seq	batch_1
8	SRX11913671	GSM5537244: Ngn4 rep-2; Aspergillus fumigatus Af293; RNA-Seq	batch_1
9	SRX11913672	GSM5537245: Ngn4 rep-3; Aspergillus fumigatus Af293; RNA-Seq	batch_1
10	SRX11913673	GSM5537246: Afp1 rep-1; Aspergillus fumigatus Af293; RNA-Seq	batch_1
11	SRX11913674	GSM5537247: Afp1 rep-2; Aspergillus fumigatus Af293; RNA-Seq	batch_1
12	SRX11913675	GSM5537248: Afp1 rep-3; Aspergillus fumigatus Af293; RNA-Seq	batch_1
13	SRX11913676	GSM5537249: Xbp1 rep-1; Aspergillus fumigatus Af293; RNA-Seq	batch_1
14	SRX11913677	GSM5537250: Xbp1 rep-2; Aspergillus fumigatus Af293; RNA-Seq	batch_1
15	SRX11913678	GSM5537251: Xbp1 rep-3; Aspergillus fumigatus Af293; RNA-Seq	batch_1
16	SRX273135	GSM1132400: C1; Aspergillus fumigatus Af293; RNA-Seq	batch_2
17	SRX273136	GSM1132401: C2; Aspergillus fumigatus Af293; RNA-Seq	batch_2
18	SRX273137	GSM1132402: C3; Aspergillus fumigatus Af293; RNA-Seq	batch_2
19	SRX273138	GSM1132403: MP1; Aspergillus fumigatus Af293; RNA-Seq	batch_2
20	SRX273139	GSM1132404: MP2; Aspergillus fumigatus Af293; RNA-Seq	batch_2
21	SRX273140	GSM1132405: MP3; Aspergillus fumigatus Af293; RNA-Seq	batch_2
22	SRX5101280	GSM3503323: 1_RNA-seq wild-type strain; Aspergillus fumigatus Af293; RNA-Seq	batch_3
23	SRX5101281	GSM3503324: 2_RNA-seq with atrR null; Aspergillus fumigatus Af293; RNA-Seq	batch_3
24	SRX5101282	GSM3503325: 3_RNA-seq with atrR-3X HA tag present; Aspergillus fumigatus Af293; RNA-Seq	batch_3
25	SRX5101283	GSM3503326: 4_RNA-seq with hspA-atrR allele present; Aspergillus fumigatus Af293; RNA-Seq	batch_3
26	SRX5101284	GSM3503327: 5_RNA-seq wild-type strain; Aspergillus fumigatus Af293; RNA-Seq	batch_3
27	SRX5101285	GSM3503328: 6_RNA-seq with atrR null; Aspergillus fumigatus Af293; RNA-Seq	batch_3
28	SRX5101286	GSM3503329: 7_RNA-seq with atrR-3X HA tag present; Aspergillus fumigatus Af293; RNA-Seq	batch_3
29	SRX5101287	GSM3503330: 8_RNA-seq with hspA-atrR allele present; Aspergillus fumigatus Af293; RNA-Seq	batch_3
30	SRX4398085	Aspergillus fumigatus Af293 (delta crzA) exposed to Caspofungin (high concentration, 2 ug/ml)	batch_4
31	SRX4398087	Aspergillus fumigatus Af293 (delta crzA) exposed to Caspofungin (high concentration, 2 ug/ml)	batch_4
32	SRX4398088	Aspergillus fumigatus Af293 (delta crzA) exposed to Caspofungin (high concentration, 2 ug/ml)	batch_4
33	SRX4398089	Aspergillus fumigatus Af293 (delta crzA) exposed to Caspofungin (low concentration, 0.2 ug/ml)	batch_4
34	SRX4398090	Aspergillus fumigatus Af293 (delta crzA) exposed to Caspofungin (low concentration, 0.2 ug/ml)	batch_4
35	SRX4398091	Aspergillus fumigatus Af293 (dcrzA) grown on minimal medium glucose without any drug (=control)	batch_4
36	SRX4398092	Aspergillus fumigatus Af293 (delta crzA) exposed to Caspofungin (low concentration, 0.2 ug/ml)	batch_4
37	SRX4398095	Aspergillus fumigatus Af293 grown on minimal medium glucose without any drug (=control)	batch_4

No.	Accession number	Description	Batch
38	SRX4398096	Aspergillus fumigatus Af293 grown on minimal medium glucose without any drug (=control)	batch_4
39	SRX4398098	Aspergillus fumigatus Af293 exposed to Caspofungin (high concentration, 2 ug/ml)	batch_4
40	SRX4398099	Aspergillus fumigatus Af293 exposed to Caspofungin (high concentration, 2 ug/ml)	batch_4
41	SRX4398100	Aspergillus fumigatus Af293 exposed to Caspofungin (high concentration, 2 ug/ml)	batch_4
42	SRX4398101	Aspergillus fumigatus Af293 exposed to Caspofungin (low concentration, 0.2 ug/ml)	batch_4
43	SRX4398102	Aspergillus fumigatus Af293 exposed to Caspofungin (low concentration, 0.2 ug/ml)	batch_4
44	SRX4398103	Aspergillus fumigatus Af293 exposed to Caspofungin (low concentration, 0.2 ug/ml)	batch_4
45	SRX4398109	Aspergillus fumigatus Af293 (dcrzA) grown on minimal medium glucose without any drug (=control)	batch_4
46	SRX4398110	Aspergillus fumigatus Af293 (dcrzA) grown on minimal medium glucose without any drug (=control)	batch_4
47	SRX4398120	Aspergillus fumigatus Af293 grown on minimal medium glucose without any drug (=control)	batch_4
48	SRX4109713	RNA-Seq of Aspergillus fumigatus	batch_5
49	SRX4109714	RNA-Seq of Aspergillus fumigatus	batch_5
50	SRX4109715	RNA-Seq of Aspergillus fumigatus	batch_5
51	SRX4109716	RNA-Seq of Aspergillus fumigatus	batch_5
52	SRX3581300	ASPERG_D6_M3	batch_6
53	SRX3581301	ASPERG_D6_M3	batch_6
54	SRX3581302	ASPERG_Neu-_D4_M1	batch_6
55	SRX3581303	ASPERG_Neu-_D4_M1	batch_6
56	SRX3581304	ASPERG_Neu-_D4_M1	batch_6
57	SRX3581306	ASPERG_Neu-_D4_M2	batch_6
58	SRX3581307	ASPERG_Neu-_D4_M2	batch_6
59	SRX3581311	ASPERG_Neu-_D4_M3	batch_6
60	SRX3581312	ASPERG_Neu-_D4_M3	batch_6
61	SRX3586814	ASPERG_D2_M2	batch_6
62	SRX3586815	ASPERG_D2_M2	batch_6
63	SRX3586818	ASPERG_D2_M3	batch_6
64	SRX3586819	ASPERG_D2_M3	batch_6
65	SRX3586820	ASPERG_D2_M3	batch_6
66	SRX3586821	ASPERG_D4_M1	batch_6
67	SRX3586822	ASPERG_D4_M1	batch_6
68	SRX3586824	ASPERG_D4_M1	batch_6
69	SRX3586826	ASPERG_D4_M2	batch_6
70	SRX3586827	ASPERG_D4_M2	batch_6
71	SRX3586828	ASPERG_D4_M2	batch_6
72	SRX3586830	ASPERG_D4_M3	batch_6
73	SRX3586831	ASPERG_D4_M3	batch_6
74	SRX3586832	ASPERG_D4_M3	batch_6
75	SRX3586889	ASPERG_D7_Neu-_M1	batch_6
76	SRX3586890	ASPERG_D7_Neu-_M1	batch_6
77	SRX3586900	ASPERG_D7_Neu-_M2	batch_6
78	SRX3586901	ASPERG_D7_Neu-_M2	batch_6
79	SRX3586902	ASPERG_D7_Neu-_M2	batch_6

No.	Accession number	Description	Batch
80	SRX3586905	ASPERG_D7_Neu-_M3	batch_6
81	SRX3586906	ASPERG_D7_Neu-_M3	batch_6
82	SRX3586907	ASPERG_D7_Neu-_M3	batch_6
83	SRX3591733	ASPERG_D2_M1	batch_6
84	SRX3591735	ASPERG_D2_M1	batch_6
85	SRX3591736	ASPERG_D6_M1	batch_6
86	SRX3591738	ASPERG_D6_M1	batch_6
87	SRX3630770	ASPERG_D6_M2	batch_6
88	SRX3630771	ASPERG_D6_M2	batch_6
89	SRX2607674	Aspergillus fumigatus strain Af293 grown on SEB Đ biological replicate 3	batch_7
90	SRX2607675	Aspergillus fumigatus strain Af293 grown on SEB Đ biological replicate 2	batch_7
91	SRX2607676	Aspergillus fumigatus strain Af293 grown on SEB Đ biological replicate 1	batch_7
92	SRX2607677	Aspergillus fumigatus strain Af293 grown on fructose Đ biological replicate 3	batch_7
93	SRX2607678	Aspergillus fumigatus strain Af293 grown on fructose Đ biological replicate 2	batch_7
94	SRX2607679	Aspergillus fumigatus strain Af293 grown on fructose Đ biological replicate 1	batch_7
95	F1	this study	batch_8
96	F2	this study	batch_8
97	F3	this study	batch_8
98	F4	this study	batch_8
		A. niger	
1	SRX3006791	Aspergillus niger NRRL 3 S-4h Gene expression profiling - 6 transcriptome	batch_1
2	SRX3006792	Aspergillus niger NRRL 3 C-24h Gene expression profiling - 32 transcriptome	batch_1
3	SRX3006793	Aspergillus niger NRRL 3 S-48h Gene expression profiling - 64 transcriptome	batch_1
4	SRX3006794	Aspergillus niger NRRL 3 C-4h Gene expression profiling - 3 transcriptome	batch_1
5	SRX3006795	Aspergillus niger NRRL 3 S-4h Gene expression profiling - 5 transcriptome	batch_1
6	SRX3006796	Aspergillus niger NRRL 3 C-4h Gene expression profiling - 1 transcriptome	batch_1
7	SRX3006797	Aspergillus niger NRRL 3 C-4h Gene expression profiling - 2 transcriptome	batch_1
8	SRX3006798	Aspergillus niger NRRL 3 Preculture Gene expression profiling - Pw transcriptome	batch_1
9	SRX3006799	Aspergillus niger NRRL 3 S-24h Gene expression profiling - 34 transcriptome	batch_1
10	SRX3006800	Aspergillus niger NRRL 3 S-48h Gene expression profiling - 66 transcriptome	batch_1
11	SRX3006801	Aspergillus niger NRRL 3 S-48h Gene expression profiling - 65 transcriptome	batch_1
12	SRX3006802	Aspergillus niger NRRL 3 C-24h Gene expression profiling - 33 transcriptome	batch_1
13	SRX3006803	Aspergillus niger NRRL 3 C-48h Gene expression profiling - 63 transcriptome	batch_1
14	SRX3006804	Aspergillus niger NRRL 3 C-48h Gene expression profiling - 62 transcriptome	batch_1
15	SRX3006805	Aspergillus niger NRRL 3 C-48h Gene expression profiling - 61 transcriptome	batch_1
16	SRX3006806	Aspergillus niger NRRL 3 Preculture Gene expression profiling - Pw 2 transcriptome	batch_1
17	SRX3006807	Aspergillus niger NRRL 3 C-24h Gene expression profiling - 31 transcriptome	batch_1
18	SRX3006808	Aspergillus niger NRRL 3 S-24h Gene expression profiling - 36 transcriptome	batch_1
19	SRX3006809	Aspergillus niger NRRL 3 S-24h Gene expression profiling - 35 transcriptome	batch_1
20	SRX5057360	Aspergillus niger NRRL3 delta xkiA Gene Expression Profiling - xkiA-1A - wheat bran 1%-2h transcriptome	batch_2
21	SRX5057361	Aspergillus niger NRRL3 delta xkiA Gene Expression Profiling - xkiA-2C - wheat bran 1%-8h transcriptome	batch_2

No.	Accession number	Description	Batch
22	SRX5057362	Aspergillus niger NRRL3 delta xkiA Gene Expression Profiling - xkiA-1C - wheat bran 1%-2h transcriptome	batch_2
23	SRX5057363	Aspergillus niger NRRL3 delta xkiA Gene Expression Profiling - xkiA-2B - wheat bran 1%-8h transcriptome	batch_2
24	SRX5057364	Aspergillus niger NRRL3 delta xkiA Gene Expression Profiling - xkiA-1B - wheat bran 1%-2h transcriptome	batch_2
25	SRX5057365	Aspergillus niger NRRL3 delta xkiA Gene Expression Profiling - xkiA-3C - wheat bran 1%-24h transcriptome	batch_2
26	SRX5057366	Aspergillus niger NRRL3 delta xkiA Gene Expression Profiling - xkiA-3A - wheat bran 1%-24h transcriptome	batch_2
27	SRX5057367	Aspergillus niger NRRL3 delta xkiA Gene Expression Profiling - xkiA-3B - wheat bran 1%-24h transcriptome	batch_2
28	SRX5057368	Aspergillus niger NRRL3 delta xkiA Gene Expression Profiling - xkiA-2A - wheat bran 1%-8h transcriptome	batch_2
29	SRX5057509	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.4.2B - arabinose 65 mM transcriptome	batch_3
30	SRX5057510	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.6.1C - rhamnose 25 mM transcriptome	batch_3
31	SRX5057511	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.7.1B - galacturonic acid 25 mM transcriptome	batch_3
32	SRX5057512	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.5.1B - xylose 25 mM transcriptome	batch_3
33	SRX5057513	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.4.2A - arabinose 65 mM transcriptome	batch_3
34	SRX5057515	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.4.2C - arabinose 65 mM transcriptome	batch_3
35	SRX5057514	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.2.1B - mannose 25 mM transcriptome	batch_3
36	SRX5057516	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.1.2C - glucose 65 mM transcriptome	batch_3
37	SRX5057517	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.1.1C - glucose 25 mM transcriptome	batch_3
38	SRX5057518	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.3.1C - fructose 25 mM transcriptome	batch_3
39	SRX5057520	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.1.1B - glucose 25 mM transcriptome	batch_3
40	SRX5057523	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.4.1C - arabinose 25 mM transcriptome	batch_3
41	SRX5057525	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.8.1A - galactose 25 mM transcriptome	batch_3
42	SRX5057529	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.4.1A - arabinose 25 mM transcriptome	batch_3
43	SRX5057530	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.8.1C - galactose 25 mM transcriptome	batch_3
44	SRX5057531	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.7.1A - galacturonic acid 25 mM transcriptome	batch_3
45	SRX5057533	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.2.1C - mannose 25 mM transcriptome	batch_3
46	SRX5057535	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.5.1C - xylose 25 mM transcriptome	batch_3
47	SRX5057536	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.2.2C - mannose 65 mM transcriptome	batch_3
48	SRX5057541	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.3.2A - fructose 65 mM transcriptome	batch_3
49	SRX5057542	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.1.2A - glucose 65 mM transcriptome	batch_3
50	SRX5057543	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.5.2B - xylose 65 mM transcriptome	batch_3
51	SRX5057544	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.5.2C - xylose 65 mM transcriptome	batch_3
52	SRX5057547	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.3.1B - fructose 25 mM transcriptome	batch_3
53	SRX5057551	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.4.1B - arabinose 25 mM transcriptome	batch_3
54	SRX5057552	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.2.2B - mannose 65 mM transcriptome	batch_3
55	SRX5057554	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.6.1A - rhamnose 25 mM transcriptome	batch_3
56	SRX5057555	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.3.2C - fructose 65 mM transcriptome	batch_3
57	SRX5057556	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.5.1A - xylose 25 mM transcriptome	batch_3
58	SRX5057558	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.3.2B - fructose 65 mM transcriptome	batch_3
59	SRX5057560	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.2.1A - mannose 25 mM transcriptome	batch_3
60	SRX5057561	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.1.2B - glucose 65 mM transcriptome	batch_3
61	SRX5057562	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.7.1C - galacturonic acid 25 mM transcriptome	batch_3
62	SRX5057563	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.6.1B - rhamnose 25 mM transcriptome	batch_3
63	SRX5057565	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.2.2A - mannose 65 mM transcriptome	batch_3

No.	Accession number	Description	Batch
64	SRX5057567	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.1.1A - glucose 25 mM transcriptome	batch_3
65	SRX5057570	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.3.1A - fructose 25 mM transcriptome	batch_3
66	SRX5057573	Aspergillus niger NRRL3 delta creA Gene Expression Profiling - 2.8.1B - galactose 25 mM transcriptome	batch_3
67	SRX5090333	Aspergillus niger NRRL3 delta galKladB Gene Expression Profiling - galKladB-3C_Wb_24h - wheat bran 1%-24h transcriptome	batch_3
68	SRX5090336	Aspergillus niger NRRL3 delta galKladBxkiA Gene Expression Profiling - galKladBxkiA-3A_Wb_24h - wheat bran 1%-24h transcriptome	batch_3
69	SRX5090350	Aspergillus niger NRRL3 delta galKladB Gene Expression Profiling - galKladB-3B_Wb_24h - wheat bran 1%-24h transcriptome	batch_3
70	SRX5090351	Aspergillus niger NRRL3 delta hxkA Gene Expression Profiling - hxkA-3B - wheat bran 1%-24h transcriptome	batch_3
71	SRX5090352	Aspergillus niger NRRL3 delta hxkA Gene Expression Profiling - hxkA-3C - wheat bran 1%-24h transcriptome	batch_3
72	SRX5090386	Aspergillus niger NRRL3 delta ladB Gene Expression Profiling - ladB-3A_Wb_24h - wheat bran 1%-24h transcriptome	batch_3
73	SRX5090387	Aspergillus niger NRRL3 delta galKladB Gene Expression Profiling - galKladB-3A_Wb_24h - wheat bran 1%-24h transcriptome	batch_3
74	SRX5090412	Aspergillus niger NRRL3 delta galK Gene Expression Profiling - galK-3A_Wb_24h - wheat bran 1%-24h transcriptome	batch_3
75	SRX5090415	Aspergillus niger NRRL3 delta galKladBxkiA Gene Expression Profiling - galKladBxkiA-3C_Wb_24h - wheat bran 1%-24h transcriptome	batch_3
76	SRX5090416	Aspergillus niger NRRL3 delta galKladBxkiA Gene Expression Profiling - galKladBxkiA-3B_Wb_24h - wheat bran 1%-24h transcriptome	batch_3
77	SRX5090417	Aspergillus niger NRRL3 delta galKladBhxkA Gene Expression Profiling - galKladBhxkA-3A_Wb_24h - wheat bran 1%-24h transcriptome	batch_3
78	SRX5090422	Aspergillus niger NRRL3 delta galKladBhxkA Gene Expression Profiling - galKladBhxkA-3C_Wb_24h - wheat bran 1%-24h transcriptome	batch_3
79	SRX5090424	Aspergillus niger NRRL3 delta galKladBhxkA Gene Expression Profiling - galKladBhxkA-3B_Wb_24h - wheat bran 1%-24h transcriptome	batch_3
80	SRX5090427	Aspergillus niger NRRL3 delta galK Gene Expression Profiling - galK-3B_Wb_24h - wheat bran 1%-24h transcriptome	batch_3
81	SRX5090429	Aspergillus niger NRRL3 delta galK Gene Expression Profiling - galK-3C_Wb_24h - wheat bran 1%-24h transcriptome	batch_3
82	SRX5090431	Aspergillus niger NRRL3 delta ladB Gene Expression Profiling - ladB-3C_Wb_24h - wheat bran 1%-24h transcriptome	batch_3
83	SRX5090434	Aspergillus niger NRRL3 delta ladB Gene Expression Profiling - ladB-3B_Wb_24h - wheat bran 1%-24h transcriptome	batch_3
84	G1	this study	batch_4
85	G2	this study	batch_4
86	G3	this study	batch_4
87	G4	this study	batch_4
		A. nidulans	
1	SRX10158135	Aspergillus nidulans FGSC A1149 Gene Expression Profiling - AN8 transcriptome	batch_1
2	SRX10158134	Aspergillus nidulans FGSC A1149 Gene Expression Profiling - AN1 transcriptome	batch_1
3	SRX10158133	Aspergillus nidulans FGSC A1149 Gene Expression Profiling - AN2 transcriptome	batch_1
4	SRX10158132	Aspergillus nidulans FGSC A1149 Gene Expression Profiling - AN11 transcriptome	batch_1
5	SRX10158131	Aspergillus nidulans FGSC A1149 Gene Expression Profiling - AN12 transcriptome	batch_1
6	SRX10158130	Aspergillus nidulans FGSC A1149 Gene Expression Profiling - AN5 transcriptome	batch_1
7	SRX10158129	Aspergillus nidulans FGSC A1149 Gene Expression Profiling - AN7 transcriptome	batch_1
8	SRX10158128	Aspergillus nidulans FGSC A1149 Gene Expression Profiling - AN6 transcriptome	batch_1
9	SRX10158127	Aspergillus nidulans FGSC A1149 Gene Expression Profiling - AN4 transcriptome	batch_1
10	SRX10158126	Aspergillus nidulans FGSC A1149 Gene Expression Profiling - AN3 transcriptome	batch_1
11	SRX10158125	Aspergillus nidulans FGSC A1149 Gene Expression Profiling - AN10 transcriptome	batch_1
12	SRX10158124	Aspergillus nidulans FGSC A1149 Gene Expression Profiling - AN9 transcriptome	batch_1
13	SRX8380246	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_30 transcriptome	batch_2
14	SRX8380245	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_31 transcriptome	batch_2
15	SRX8380244	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_29 transcriptome	batch_2
16	SRX8380243	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_28 transcriptome	batch_2

No.	Accession number	Description	Batch
17	SRX8380242	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_49 transcriptome	batch_2
18	SRX8380241	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_48 transcriptome	batch_2
19	SRX8380240	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_50 transcriptome	batch_2
20	SRX8380239	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_46 transcriptome	batch_2
21	SRX8380238	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_45 transcriptome	batch_2
22	SRX8380237	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_47 transcriptome	batch_2
23	SRX8380236	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_43 transcriptome	batch_2
24	SRX8380235	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_44 transcriptome	batch_2
25	SRX8380234	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_42 transcriptome	batch_2
26	SRX8380233	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_41 transcriptome	batch_2
27	SRX8380232	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_40 transcriptome	batch_2
28	SRX8380231	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_39 transcriptome	batch_2
29	SRX8380230	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_36 transcriptome	batch_2
30	SRX8380229	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_38 transcriptome	batch_2
31	SRX8380228	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_37 transcriptome	batch_2
32	SRX8380227	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_51 transcriptome	batch_2
33	SRX8380226	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_53 transcriptome	batch_2
34	SRX8380225	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_52 transcriptome	batch_2
35	SRX8380224	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_54 transcriptome	batch_2
36	SRX8380223	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_35 transcriptome	batch_2
37	SRX8380222	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_33 transcriptome	batch_2
38	SRX8380221	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_34 transcriptome	batch_2
39	SRX8380220	Aspergillus nidulans FGSC A4 Gene Expression Profiling - 504291_32 transcriptome	batch_2
40	SRX7896638	Salicylate catabolism of aspergilli	batch_3
41	SRX7896637	Salicylate catabolism of aspergilli	batch_3
42	SRX7896636	Salicylate catabolism of aspergilli	batch_3
43	SRX7896635	Salicylate catabolism of aspergilli	batch_3
44	SRX7896632	Salicylate catabolism of aspergilli	batch_3
45	SRX7896631	Salicylate catabolism of aspergilli	batch_3
46	SRX4369957	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
47	SRX4369956	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
48	SRX4369955	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
49	SRX4369954	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
50	SRX4369953	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
51	SRX4369952	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
52	SRX4369951	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
53	SRX4369950	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
54	SRX4369949	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
55	SRX4369948	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
56	SRX4369947	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
57	SRX4369946	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
58	SRX4369945	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4

No.	Accession number	Description	Batch
59	SRX4369944	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
60	SRX4369943	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
61	SRX4369942	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
62	SRX4369941	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
63	SRX4369940	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
64	SRX4369939	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
65	SRX4369938	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
66	SRX4369937	Cytochrome P450 Monooxygenase Mediated Metabolic Utilization of Benzo(a)pyrene by Fungi	batch_4
67	SRX1619630	Aspergillus nidulans FGSC A4 Transcriptome or Gene expression	batch_5
68	SRX1619628	Aspergillus nidulans FGSC A4 Transcriptome or Gene expression	batch_5
69	SRX1619627	Aspergillus nidulans FGSC A4 Transcriptome or Gene expression	batch_5
70	SRX1619624	Aspergillus nidulans FGSC A4 Transcriptome or Gene expression	batch_5
71	SRX1619620	Aspergillus nidulans FGSC A4 Transcriptome or Gene expression	batch_5
72	SRX1619617	Aspergillus nidulans FGSC A4 Transcriptome or Gene expression	batch_5
73	SRX1619614	Aspergillus nidulans FGSC A4 Transcriptome or Gene expression	batch_5
74	SRX226254	Inducer Free Cellulase Secretion in Neurospora Crassa and comparitive analysis of cellulase induction in Aspergillus nidulans	batch_6
75	SRX226253	Inducer Free Cellulase Secretion in Neurospora Crassa and comparitive analysis of cellulase induction in Aspergillus nidulans	batch_6
76	D1	this study	batch_7
77	D2	this study	batch_7
78	D3	this study	batch_7
79	D4	this study	batch_7
		A. oryzae	
1	DRR178650	48_h_wild_type	batch_1
2	DRR178651	48_h_wild_type	batch_1
3	DRR178652	48_h_wild_type	batch_1
4	DRR178653	72_h_wild_type	batch_1
5	DRR178654	72_h_wild_type	batch_1
6	DRR178655	72_h_wild_type	batch_1
7	DRR178656	96_h_wild_type	batch_1
8	DRR178657	96_h_wild_type	batch_1
9	DRR178658	96_h_wild_type	batch_1
10	DRR178659	120_h_wild_type	batch_1
11	DRR178660	120_h_wild_type	batch_1
12	DRR178661	120_h_wild_type	batch_1
13	DRR178662	144_h_wild_type	batch_1
14	DRR178663	144_h_wild_type	batch_1
15	DRR178664	144_h_wild_type	batch_1
16	DRR178665	168_h_wild_type	batch_1
17	DRR178666	168_h_wild_type	batch_1
18	DRR178667	168_h_wild_type	batch_1
19	DRR178668	48_h_faaA_deleted	batch_1

No.	Accession number	Description	Batch
20	DRR178669	48_h_faaA_deleted	batch_1
21	DRR178670	48_h_faaA_deleted	batch_1
22	DRR178671	72_h_faaA_deleted	batch_1
23	DRR178672	72_h_faaA_deleted	batch_1
24	DRR178673	72_h_faaA_deleted	batch_1
25	DRR178674	96_h_faaA_deleted	batch_1
26	DRR178675	96_h_faaA_deleted	batch_1
27	DRR178676	96_h_faaA_deleted	batch_1
28	DRR178677	120_h_faaA_deleted	batch_1
29	DRR178678	120_h_faaA_deleted	batch_1
30	DRR178679	120_h_faaA_deleted	batch_1
31	DRR178680	144_h_faaA_deleted	batch_1
32	DRR178681	144_h_faaA_deleted	batch_1
33	DRR178682	144_h_faaA_deleted	batch_1
34	DRR178683	168_h_faaA_deleted	batch_1
35	DRR178684	168_h_faaA_deleted	batch_1
36	DRR178685	168_h_faaA_deleted	batch_1
37	DRR205698	Yamada-nishiki_90	batch_2
38	DRR205699	Yamada-nishiki_50	batch_2
39	DRR205700	Chiyo-nishiki_50	batch_2
40	DRR272358	delta_opsA	batch_3
41	DRR272359	delta_opsB	batch_3
42	DRR272360	delta_opsAopsB	batch_3
43	DRR272361	control	batch_3

Supplementary Table 8 ChIP-seq data analysis

A. fumigatus		
No.	**Accession number**	**Description**
1	SRR14493639	Af293_G_H3_AGGTCAGTT_ChIPMix82
2	SRR14493640	Af293g_H3_CTTGTAT_ChIPMix78
3	SRR14493641	af293_Chris_16h_input_TCGCGTACT_ChIPMix90
4	SRR14493642	af293_chris_input_TCGCGTACT_ChIPMix82
5	SRR14493643	Af293_C_H3K9me3_GGCTACT_ChIPMix82
6	SRR14493644	Af293c_H3k9me3_CCTTTACAGT_ChIPMix78
7	SRR14493645	Af293_C_H3K4me3_TAGCTTT_ChIPMix82
8	SRR14493650	Af293c_H3k4me3_TGACGCATT_ChIPMix78
9	SRR14493655	Af293g_input_TCGCGTACT_ChIPMix78
10	SRR14493656	af293_Ana_input_TATACCGTT_ChIPMix82
11	SRR14493657	Af293_G_H3K9me3_CCATACACT_ChIPMix82
12	SRR14493658	Af293g_H3k9me3_TTGAGTGTT_ChIPMix78
13	SRR14493661	Af293_C_H3_GATCAGT_ChIPMix82
14	SRR14493662	Af293c_H3_TAGCTTT_ChIPMix78

A. nidulans		
No.	**Accession number**	**Description**
1	SRR2170269	KdmB 17 h H3Ac ChIP
2	SRR2170273	KdmB 17 h H3Ac ChIP
3	SRR2170277	KdmB 48 h H3Ac ChIP
4	SRR2170281	KdmB 48 h H3Ac ChIP
5	SRR2170286	WT 17 h H3Ac ChIP
6	SRR2170290	WT 17 h H3Ac ChIP
7	SRR2170294	WT 48 h H3Ac ChIP
8	SRR2170298	WT 48 h H3Ac ChIP
9	SRR2170270	KdmB 17 h H3K36me3 ChIP
10	SRR2170274	KdmB 17 h H3K36me3 ChIP
11	SRR2170278	KdmB 48 h H3K36me3 ChIP
12	SRR2170282	KdmB 48 h H3K36me3 ChIP
13	SRR2170287	WT 17 h H3K36me3 ChIP
14	SRR2170291	WT 17 h H3K36me3 ChIP
15	SRR2170295	WT 48 h H3K36me3 ChIP
16	SRR2170299	WT 48 h H3K36me3 ChIP
17	SRR2170271	KdmB 17 h H3K4me3 ChIP
18	SRR2170275	KdmB 17 h H3K4me3 ChIP
19	SRR2170279	KdmB 48 h H3K4me3 ChIP
20	SRR2170283	KdmB 48 h H3K4me3 ChIP
21	SRR2170288	WT 17 h H3K4me3 ChIP
22	SRR2170292	WT 17 h H3K4me3 ChIP
23	SRR2170296	WT 48 h H3K4me3 ChIP
24	SRR2170300	WT 48 h H3K4me3 ChIP
25	SRR2170272	KdmB 17 h H3k9me3 ChIP
26	SRR2170276	KdmB 17 h H3k9me3 ChIP
27	SRR2170280	KdmB 48 h H3k9me3 ChIP
28	SRR2170284	KdmB 48 h H3k9me3 ChIP
29	SRR2170289	WT 17 h H3k9me3 ChIP
30	SRR2170293	WT 17 h H3k9me3 ChIP
31	SRR2170297	WT 48 h H3k9me3 ChIP
32	SRR2170301	WT 48 h H3k9me3 ChIP

References

Blin, K., Shaw, S., Kloosterman, A. M., Charlop-Powers, Z., Van Wezel, G. P., Medema, M. H. & Weber, T. (2021). AntiSMASH 6.0: Improving cluster detection and comparison capabilities. *Nucleic Acids Research*, *49*(W1), W29–W35.

Brakhage, A. A. (2013). Regulation of fungal secondary metabolism. *Nature Reviews Microbiology*, *11*(1), 21–32.

Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, *34*(5), 525–527.

Cairns, T. C., Barthel, L. & Meyer, V. (2021). Something old, something new: challenges and developments in *Aspergillus niger* biotechnology. *Essays in Biochemistry*, *65*(2), 213–224.

Cánovas, D., Marcos, A. T., Gacek, A., Ramos, M. S., Gutiérrez, G., Reyes-Domínguez, Y. & Strauss, J. (2014). The histone acetyltransferase GcnE (GCN5) plays a central role in the regulation of *Aspergillus* asexual development. *Genetics*, *197*(4), 1175–1189.

Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, *25*(15), 1972–1973.

Chernomor, O., Von Haeseler, A. & Minh, B. Q. (2016). Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biology*, *65*(6), 997–1008.

Chiang, Y.-M., Szewczyk, E., Davidson, A. D., Keller, N., Oakley, B. R. & Wang, C. C. C. (2009). A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in *Aspergillus nidulans*. *Journal of the American Chemical Society*, *131*(8), 2965–2970.

Colabardini, A. C., Wang, F., Miao, Z., Pardeshi, L., Valero, C., de Castro, P. A., Akiyama, D. Y., Tan, K., Nora, L. C., Silva-Rocha, R., Marcet-Houben, M., Gabaldón, T., Fill, T., Wong, K. H. & Goldman, G. H. (2022). Chromatin profiling reveals heterogeneity in clinical isolates of the human pathogen *Aspergillus fumigatus*. In *PLOS Genetics* (Vol. 18).

Collemare, J. & Seidl, M. F. (2019). Chromatin-dependent regulation of secondary metabolite biosynthesis in fungi: is the picture complete? *FEMS Microbiology Reviews*, *43*(6), 591–607.

Connolly, L. R., Smith, K. M. & Freitag, M. (2013). The *Fusarium graminearum* histone H3 K27 methyltransferase KMT6 regulates development and expression of secondary metabolite gene clusters. *PLoS Genetics*, *9*(10), e1003916.

Drillon, G., Carbone, A. & Fischer, G. (2013). Combinatorics of chromosomal rearrangements based on synteny blocks and synteny packs. *Journal of Logic and Computation*, *23*(4), 815–838.

Emms, D. M. & Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology*, *20*(1), 1–14.

Gacek-Matthews, A., Berger, H., Sasaki, T., Wittstein, K., Gruber, C., Lewis, Z. A. & Strauss, J. (2016). KdmB, a Jumonji histone H3 demethylase, regulates genome-wide H3K4 trimethylation and is required for normal induction of secondary metabolism in *Aspergillus nidulans*. *PLoS Genetics*, *12*(8), 1–29.

Gacek, A. & Strauss, J. (2012). The chromatin code of fungal secondary metabolite gene clusters. *Applied Microbiology and Biotechnology*, *95*(6), 1389–1404.

Gibbons, J. G., Salichos, L., Slot, J. C., Rinker, D. C., McGary, K. L., King, J. G., Klich, M. A., Tabb, D. L., McDonald, W. H. & Rokas, A. (2012). The evolutionary imprint of domestication on genome variation and function of the filamentous fungus *Aspergillus oryzae*. *Current Biology*, *22*(15), 1403–1409.

Grigoriev, I. V, Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R., Riley, R., Salamov, A., Zhao, X. & Korzeniewski, F. (2014). MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Research*, *42*(D1), D699–D704.

Guzmán-Chávez, F., Zwahlen, R. D., Bovenberg, R. A. L. & Driessen, A. J. M. (2018). Engineering of the filamentous fungus *Penicillium chrysogenum* as cell factory for natural products. *Frontiers in Microbiology*, *9*, 2768.

Hane, J. K., Rouxel, T., Howlett, B. J., Kema, G. H. J., Goodwin, S. B. & Oliver, R. P. (2011). A novel mode of chromosomal evolution peculiar to filamentous Ascomycete fungi. *Genome*

Biology, *12*(5).
Heaps, H. S. (1978). Information retrieval: Computational and theoretical aspects. *Academic Press*.
Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, *9*(03), 90–95.
Inglis, D. O., Binkley, J., Skrzypek, M. S., Arnaud, M. B., Cerqueira, G. C., Shah, P., Wymore, F., Wortman, J. R. & Sherlock, G. (2013). Comprehensive annotation of secondary metabolite biosynthetic genes and gene clusters of *Aspergillus nidulans*, *A. fumigatus*, *A. niger* and *A. oryzae*. *BMC Microbiology*, *13*, 1–23.
Junier, T. & Zdobnov, E. M. (2010). The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics*, *26*(13), 1669–1670.
Katoh, K. & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780.
Keller, N. P. (2019). Fungal secondary metabolism: regulation, function and drug discovery. *Nature Reviews Microbiology*, *17*(3), 167–180.
Kjærbølling, I., Vesth, T., Frisvad, J. C., Nybo, J. L., Theobald, S., Kildgaard, S., Petersen, T. I., Kuo, A., Sato, A., Lyhne, E. K., Kogle, M. E., Wiebenga, A., Kun, R. S., Lubbers, R. J. M., Mäkelä, M. R., Barry, K., Chovatia, M., Clum, A., Daum, C., … Andersen, M. R. (2020). A comparative genomics study of 23 *Aspergillus* species from section *Flavi*. *Nature Communications*, *11*(1).
Klejnstrup, M. L., Frandsen, R. J. N., Holm, D. K., Nielsen, M. T., Mortensen, U. H., Larsen, T. O. & Nielsen, J. B. (2012). Genetics of polyketide metabolism in *Aspergillus nidulans*. *Metabolites*, *2*(1), 100–133.
Lawrence, M., Daujat, S. & Schneider, R. (2016). Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends in Genetics*, *32*(1), 42–56.
Lee, I., Oh, J. H., Keats Shwab, E., Dagenais, T. R. T., Andes, D. & Keller, N. P. (2009). HdaA, a class 2 histone deacetylase of *Aspergillus fumigatus*, affects germination and secondary metabolite production. *Fungal Genetics and Biology*, *46*(10), 782–790.
Li, X., Pan, L., Wang, B. & Pan, L. (2019). The histone deacetylases HosA and HdaA affect the phenotype and transcriptomic and metabolic profiles of *Aspergillus niger*. *Toxins*, *11*(9).
Lovell, J. T., Sreedasyam, A., Schranz, M. E., Wilson, M., Carlson, J. W., Harkess, A., Emms, D., Goodstein, D. M. & Schmutz, J. (2022). GENESPACE tracks regions of interest and gene copy number variation across multiple genomes. *Elife*, *11*, e78526.
Macheleidt, J., Mattern, D. J., Fischer, J., Netzker, T., Weber, J., Schroeckh, V., Valiante, V. & Brakhage, A. A. (2016). Regulation and role of fungal secondary metabolites. *Annual Review of Genetics*, *50*, 371–392.
Mosunova, O., Navarro-Muñoz, J. C. & Collemare, J. (2021). The biosynthesis of fungal secondary metabolites: From fundamentals to biotechnological applications. In *Encyclopedia of mycology* (pp. 458–476). Elsevier.
Nakao, Y., Kanamori, T., Itoh, T., Kodama, Y., Rainieri, S., Nakamura, N., Shimonaga, T., Hattori, M. & Ashikari, T. (2009). Genome sequence of the lager brewing yeast, an interspecies hybrid. *DNA Research*, *16*(2), 115–129.
Newman, D. J. & Cragg, G. M. (2016). Natural products as sources of new drugs from 1981 to 2014. *Journal of Natural Products*, *79*(3), 629–661.
Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274.
Nielsen, K. F., Mogensen, J. M., Johansen, M., Larsen, T. O. & Frisvad, J. C. (2009). Review of secondary metabolites and mycotoxins from the *Aspergillus niger* group. *Analytical and Bioanalytical Chemistry*, *395*, 1225–1242.
Nierman, W. C., Pain, A., Anderson, M. J., Wortman, J. R., Kim, H. S., Arroyo, J., Berriman, M., Abe, K., Archer, D. B., Bermejo, C., Bennett, J., Bowyer, P., Chen, D., Collins, M., Coulsen, R., Davies, R., Dyer, P. S., Farman, M., Fedorova, N., … Denning, D. W. (2005). Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. *Nature*, *438*(7071), 1151–1156.

Palmer, J. M. & Keller, N. P. (2010). Secondary metabolism in fungi: Does chromosomal location matter? *Current Opinion in Microbiology*, *13*(4), 431–436.

Perrin, R. M., Fedorova, N. D., Jin, W. B., Cramer, R. A., Wortman, J. R., Kim, H. S., Nierman, W. C. & Keller, N. P. (2007). Transcriptional regulation of chemical diversity in *Aspergillus fumigatus* by LaeA. *PLoS Pathogens*, *3*(4), 508–517.

Price, M. N., Dehal, P. S. & Arkin, A. P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Molecular Biology and Evolution*, *26*(7), 1641–1650.

Rice, J. A. (2003). Mathematical statistics and data analysis. *China Machine Press*.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212.

Steenwyk, J. L., Shen, X. X., Lind, A. L., Goldman, G. H. & Rokas, A. (2019). A robust phylogenomic time tree for biotechnologically and medically important fungi in the genera *Aspergillus* and *Penicillium*. *MBio*, *10*(4), 1–25.

Studt, L., Janevska, S., Arndt, B., Boedi, S., Sulyok, M., Humpf, H.-U., Tudzynski, B. & Strauss, J. (2017). Lack of the COMPASS component Ccl1 reduces H3K4 trimethylation levels and affects transcription of secondary metabolite genes in two plant–pathogenic *Fusarium* species. *Frontiers in Microbiology*, *7*, 2144.

Team, R. D. C. (2010). R: A language and environment for statistical computing. *(No Title)*.

Theobald, S., Vesth, T. C., Rendsvig, J. K., Nielsen, K. F., Riley, R., de Abreu, L. M., Salamov, A., Frisvad, J. C., Larsen, T. O., Andersen, M. R. & Hoof, J. B. (2018). Uncovering secondary metabolite evolution and biosynthesis using gene cluster networks and genetic dereplication. *Scientific Reports*, *8*(1), 1–12.

Vesth, T. C., Nybo, J. L., Theobald, S., Frisvad, J. C., Larsen, T. O., Nielsen, K. F., Hoof, J. B., Brandl, J., Salamov, A., Riley, R., Gladden, J. M., Phatale, P., Nielsen, M. T., Lyhne, E. K., Kogle, M. E., Strasser, K., McDonnell, E., Barry, K., Clum, A., … Andersen, M. R. (2018). Investigation of inter- and intraspecies variation through genome sequencing of *Aspergillus* section *Nigri*. *Nature Genetics*, *50*(12), 1688–1695.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W. & Bright, J. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272.

Visagie, C M, Yilmaz, N., Kocsubé, S., Frisvad, J. C., Hubka, V., Samson, R. A. & Houbraken, J. (2023). A review of recently introduced *Aspergillus*, *Penicillium*, *Talaromyces* and other *Eurotiales* species. *Studies in Mycology*.

Visagie, Cobus Meyer, Hirooka, Y., Tanney, J. B., Whitfield, E., Mwange, K., Meijer, M., Amend, A. S., Seifert, K. A. & Samson, R. A. (2014). *Aspergillus*, *Penicillium* and *Talaromyces* isolated from house dust samples collected around the world. *Studies in Mycology*, *78*, 63–139.

Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., Lee, T. H., Jin, H., Marler, B., Guo, H., Kissinger, J. C. & Paterson, A. H. (2012). MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research*, *40*(7), 1–14.

Waskom, M. L. (2021). Seaborn: statistical data visualization. *Journal of Open Source Software*, *6*(60), 3021.

Wiemann, P. & Keller, N. P. (2014). Strategies for mining fungal natural products. *Journal of Industrial Microbiology and Biotechnology*, *41*(2), 301–313.

Wiemann, P., Sieber, C. M. K., Von Bargen, K. W., Studt, L., Niehaus, E.-M., Espino, J. J., Huss, K., Michielse, C. B., Albermann, S. & Wagner, D. (2013). Deciphering the cryptic genome: genome-wide analyses of the rice pathogen *Fusarium fujikuroi* reveal complex regulation of secondary metabolism and novel metabolites. *PLoS Pathogens*, *9*(6), e1003475.

Wiemann, P., Willmann, A., Straeten, M., Kleigrewe, K., Beyer, M., Humpf, H. & Tudzynski, B. (2009). Biosynthesis of the red pigment bikaverin in *Fusarium fujikuroi*: genes, their function and regulation. *Molecular Microbiology*, *72*(4), 931–946.

Xu, F., Li, X., Ren, H., Zeng, R., Wang, Z., Hu, H., Bao, J. & Que, Y. (2022). The first telomere-to-telomere chromosome-level genome assembly of *Stagonospora tainanensis* causing

sugarcane leaf blight. *Journal of Fungi*, *8*(10), 1088.
Yang, K., Tian, J. & Keller, N. P. (2022). Post-translational modifications drive secondary metabolite biosynthesis in *Aspergillus*: a review. *Environmental Microbiology*.
Zhang, X., Noberini, R., Bonaldi, T., Collemare, J. & Seidl, M. F. (2022). The histone code of the fungal genus *Aspergillus* uncovered by evolutionary and proteomic analyses. *Microbial Genomics*, *8*.
Zhang, X., Noberini, R., Vai, A., Bonaldi, T., Seidl, M. F. & Collemare, J. (2023). Detection and quantification of the histone code in the fungal genus *Aspergillus*. *Fungal Genetics and Biology*, *167*, 103800.
Zhang, Y., Parmigiani, G. & Johnson, W. E. (2020). ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genomics and Bioinformatics*, *2*(3), lqaa078.

Chapter 3

The histone code of the fungal genus *Aspergillus* uncovered by evolutionary and proteomic analyses

Xin Zhang, Roberta Noberini, Tiziana Bonaldi,
Jérôme Collemare+, Michael F. Seidl+

+These authors contributed equally
Microbial Genomics, 2022

Abstract

Chemical modifications of DNA and histone proteins impact the organization of chromatin within the nucleus. Changes in these modifications, catalyzed by different chromatin-modifying enzymes, influence chromatin organization, which in turn is thought to impact the spatial and temporal regulation of gene expression. While combinations of different histone modifications, the histone code, have been studied in several model species, we know very little about histone modifications in the fungal genus *Aspergillus*, whose members are generally well-studied due to their importance as models in cell and molecular biology as well as their medical and biotechnological relevance. Here, we used phylogenetic analyses in 94 Aspergilli as well as other fungi to uncover the occurrence and evolutionary trajectories of enzymes and protein complexes with roles in chromatin modifications or regulation. We found that these enzymes and complexes are highly conserved in Aspergilli, pointing towards a complex repertoire of chromatin modifications. Nevertheless, we also observed few recent gene duplications or losses, highlighting *Aspergillus* species to further study the roles of specific chromatin modifications. SET7 (KMT6) and other components of PRC2 (Polycomb Repressive Complex 2), which is responsible for methylation on histone H3 at lysine 27 in many eukaryotes including fungi, are absent in Aspergilli as well as in closely related *Penicillium* species, suggesting that these lost the capacity for this histone modification. We corroborated our computational predictions by performing untargeted mass spectrometry analysis of histone post-translational modifications in *Aspergillus nidulans*. This systematic analysis will pave the way for future research into the complexity of the histone code and its functional implications on genome architecture and gene regulation in fungi.

Introduction

Inside eukaryotic cells, the membrane-bound nucleus facilitates the storage and replication of the genetic material as well as the regulation of gene expression (Labrador & Corces, 2002). Within the nucleus, DNA is arranged as a chromatin fiber formed by repeated nucleosome units consisting of histone protein octamers (two copies of each H2A, H2B, H3, and H4 histones) that are wrapped by 145-147 bp of DNA (Luger, 2003; Luger et al., 1997). The chromatin structure is regulated by chemical modifications such as methylation of DNA on cytosines (5mC) and adenines (6mA), and different post-translational modifications (PTMs) on the histone proteins, including methylation and acetylation (Jenuwein & Allis, 2001; Peterson & Laniel, 2004). The combination of these different modifications, the histone code, influences the recruitment of diverse nucleosome-associated proteins (e.g., DNA binding proteins), which can cause dynamic transitions between the condensed and transcriptionally silent heterochromatin and the accessible and transcriptionally active euchromatin (Berger, 2007; Grewal & Jia, 2007; Santos-Rosa et al., 2002). Chromatin dynamics is therefore key for organisms to adapt gene expression patterns in response to a variety of developmental and environmental clues.

In the past decades, several enzymes involved in orchestrating chromatin dynamics have been identified and functionally characterized in diverse eukaryotes (Berger, 2007; Jenuwein & Allis, 2001; B. Xiao et al., 2003). These enzymes can be broadly classified into three distinct types: (i) writers that place modifications on nucleosomes or linker DNA; (ii) erasers that remove these modifications; and (iii) readers that first recognize specific modifications and subsequently recruit writers, erasers, or other proteins to the correct genomic location (Collemare & Seidl, 2019). Chromatin modifiers that catalyze the same type of chemical modification often harbor functionally and evolutionarily conserved catalytic domains. For instance, DNA methylation (5mC) is catalyzed by different DNA methyltransferases (DNMTs) which contain a DNA methyltransferase domain, while histone mono-, di-, and tri-methylation are nearly exclusively catalyzed by proteins containing a SET (Su (var) 3-9, E(z) and Trithorax) domain (B. Xiao et al., 2003). Next to these catalytic domains, the enzymes often contain accessory domains, which can contribute to the overall stability of the enzyme or establish interactions with other proteins (Helmlinger et al., 2021). Furthermore, these multi-domain proteins, together with other accessory protein subunits, can form multi-protein complexes to find the correct target region witin the genome and modify the exact amino acid residue, collectively contributing to the specific functions of these complexes (Gates et al., 2018; Hajheidari et al., 2019). For example, the histone methyltransferase complex PRC2 (Polycomb Repressive Complex 2) is responsible for the mono-, di-, and tri-methylation of histone protein H3 on the lysine at position 27 (abbreviated as H3K27me1/2/3; the same abbreviation principle will be applied below for other histone PTMs) (Freitag, 2017; Jamieson et al., 2013, 2016; Kamei et al., 2021). Evolutionary analyses have suggested that the key enzymes responsible for chromatin modifications have evolved and siginificantly expanded

early in eukaryotic evolution (Weiner et al., 2020). An analysis of 94 gene families with roles in chromatin modifications reveals that 87 out of 94 gene families emerged more than one billion years ago, and 48 were already present in the last eukaryotic common ancestor (Weiner et al., 2020).

Among eukaryotes, the fungal kingdom is remarkably large and diverse, comprising an estimated 1,5-5 million species with various morphologies, life cycles, and ecological niches (Choi & Kim, 2017). The most well-studied model fungal species in respect of chromatin modifications is the unicellular yeast *Saccharomyces cerevisiae*, in which 75 histone PTMs on 43 distinct amino acid positions have been reported with histone methylation and acetylation being most abundant (Collemare & Seidl, 2019). In addition to ancient chromatin modifiers, also lineage-specific innovations such as duplications and losses were found in fungi. For example, Dim-2 and RID (Repeat-Induced Point Mutation (RIP)-Defective) are considered to be fungal-specific DNMTs, while DNMT3 that is typically found in animals and plants seems to be absent in fungi (Bewick et al., 2019; Kouzminova & Selker, 2001; Nai et al., 2021). However, *S. cerevisiae* is not representative of the complete fungal kingdom, and detailed studies on PTMs are needed for other, less studied fungal species.

The fungal genus *Aspergillus* comprises approximately 350 species, including pathogens of humans, animals, and crops, food contaminants, as well as important cell factories for industrial and medical applications (Samson et al., 2014). Chromatin dynamics influences gene expression in several Aspergilli and regulates growth, sexual development, secondary metabolite biosynthesis, and virulence (Cánovas et al., 2014; Collemare & Seidl, 2019; Palmer et al., 2008; Yang et al., 2016). Thus far, very few well-known chromatin modifications and their catalytic protein complexes have been identified and studied in few *Aspergillus* species (Collemare & Seidl, 2019). For example, gene knockout of the H3K9 methyltransferase *Clr4* homolog in *Aspergillus fumigatus* reduced histone H3K9me1/3 results in reduction of radial growth and conidia production, as well as delayed conidiation (Palmer et al., 2008). Similarly, CclA, a subunit of the COMPASS (COMplex of Proteins ASsociated with SET1) complex, is essential for H3K4me2/3 in *A. fumigatus*, and the deletion strain showed increased secondary metabolite production and decreased growth (Palmer et al., 2013). The SAGA (Spt-Ada-Gcn5 Acetyltransferase) complex, which is responsible for histone acetylation, has been functionally characterized in *Aspergillus nidulans* (Georgakopoulos et al., 2013). Notably, this complex in *A. nidulans* lacks the Sgf11 and Sus1 subunits that have been described in *S. cerevisiae*, suggesting that this complexe evolves in distinct trajectories in different fungi (Georgakopoulos et al., 2013). Besides these few examples, we so far lack a complete view of the occurrence and evolution of chromatin modifiers in the *Aspergillus* genus. Here, we report the phylogenetic analyses of core enzymes involved in the deposition and removal of histone PTMs, and assess the conservation of all the subunits belonging to chromatin regulator complexes in 94 diverse Aspergilli. We also performed proteomics analyses to

corroborate our computational predictions and reveal a broader spectrum of histone modifications.

Results and discussion

DNA methylation, histone methylation, and histone acetylation are among the most studied chromatin modifications in animals, plants, and fungi (Binda, 2013; Wade et al., 1997; Wolffe & Matzke, 1999). Protein complexes involved in writing, reading, and erasing these modifications have been reported in few model fungal species to date (Jamieson et al., 2013; Kouzminova & Selker, 2001), but the evolution and conservation of these complexes in diverse fungal species have not been largely systematically explored. Besides some research on DNA methyltransferases (Bewick et al., 2019; Nai et al., 2021) and SET domain-containing proteins (Freitag, 2017) that have been performed throughout the fungal kingdom, we still lack the evolutionary overview of other chromatin modifiers and in depth analyses for important fungal groups. To start uncovering the occurrence and evolutionary histories of chromatin modifiers involved in these major modifications in the fungal genus *Aspergillus,* we first focused on six conserved catalytic domains reported to be involved in histone modifications: DNA methyltransferase domain for DNMTs, SET domain for histone methyltransferases (HMTs), acetyltransferase domain and MOZ/SAS FAMILY domain for histone acetyltransferases (HATs), histone deacetylase domain for histone deacetylases (HDACs), and JmjC domain for histone demethylases (HDMTs). We identified homologs harboring these conserved domains in 109 predicted proteomes; the completeness of the predicted proteomes was generally high as 97 had more than 97% complete BUSCO predicted proteins (Table S1). The 94 Aspergilli species cover 18 recognized sections: *Polypaecilum*, *Cremei*, *Restricti*, *Aspergillus*, *Cervini*, *Clavati*, *Fumigati*, *Candidi*, *Circumdati*, *Janorum*, *Terrei*, *Flavi*, *Ochraceorosei*, *Bispori*, *Cavernicolarum*, *Usti*, *Nidulantes*, and *Nigri* (Chen et al., 2016; James et al., 2020; Kjærbølling et al., 2020; Souza et al., 2019; Steenwyk et al., 2019; Sun et al., 2020; Vesth et al., 2018), and 15 other species were used as outgroups (Figure S1). Two of the outgroup species are Basidiomycota and the remaining species cover each class in the Ascomycota phylum, which provides a robust framework for our evolutionary analyses within Aspergilli. The outgroup species include well-studied model species, some of which are important for crop production, food fermentation, and antibiotics production (Grunstein & Gasser, 2013; Nielsen et al., 2017). We used the identified homologs for each catalytic domain to determine their evolutionary history using maximum-likelihood phylogenies (Figures 1-4, S2, and S3).

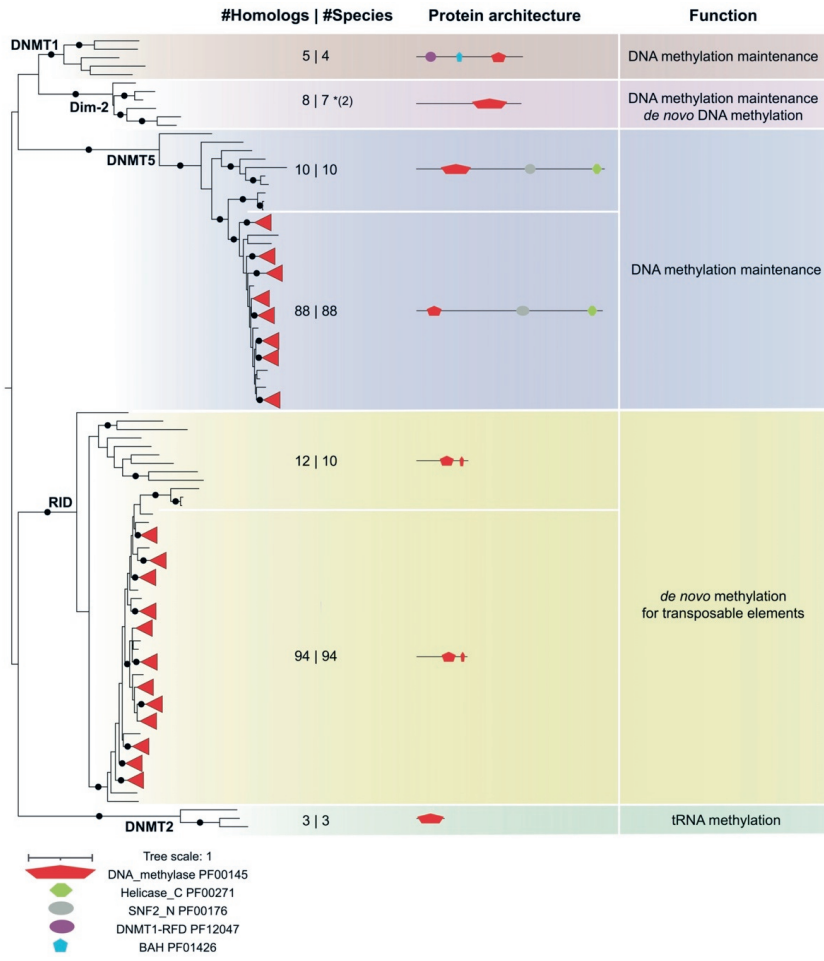


Figure 1. Conservation of DNA methyltransferase families in the *Aspergillus* genus and outgroup species. The maximum-likelihood phylogeny of DNA methyltransferase domain (PF00145) found in 94 Aspergilli and 15 outgroup species was determined using IQ-TREE (Boratyn et al., 2013; Simão et al., 2015). The black dot on the branch indicates ultrafast bootstrap values over 95 and SH-aLRT values over 80. Leaves from Aspergilli with trustworthy support are collapsed and shown as red triangles. The numbers of hits and species are separated by the pipe in the '#Homologs | #Species' column. The number in parentheses after * indicates the additional hits found by the TBLASTN search on NCBI. Additional conserved domains in DNA methyltransferase domain-containing proteins were identified using the PFAM database (https://pfam.xfam.org).

Aspergilli harbor conserved DNA methyltransferases for genome defense and methylation maintenance

The conserved DNA methyltransferase domain is the only domain known to catalyze the addition of methyl groups to DNA specifically at cytosines (5mC) (Bewick et al., 2019). Based on previous domain architecture analyses, phylogeny reconstructions, and functional experiments, five distinct protein families containing a DNA methyltransferase domain have been recognized in fungi: DNMT1, DNMT2, and DNMT5 are commonly found in eukaryotes, while Dim-2 and RID are considered to be fungal specific (Bewick et al., 2019). Among these five families, DNMT2 seems to function as a tRNA methyltransferase rather than DNA methyltransferase (Ko et al., 2021). In line with previous observation (Bewick et al., 2019; Jurkowski & Jeltsch, 2011), DNMT2 only occurs in our study in two Basidiomycota and one Ascomycota (*Schizosaccharomyces pombe*), suggesting it has been lost in most Ascomycota.

Three (DNMT1, DNMT5, and Dim-2) of the four remaining families catalyze the maintenance of DNA methylation pattern from parental DNA strand onto synthesized daughter DNA strands, while Dim-2 is also able to catalyze *de novo* methylation (Bestor, 1988; Okano et al., 1998). Based on our species selection and phylogenetic analyses, we here found *DNMT1* and *Dim-2* to be absent in all Aspergilli and only present in four or seven out of 15 outgroup species, respectively, while *DNMT5* is found in most Aspergilli and outgroup species (Figure 1). Only four outgroup species harboring *DNMT1*, the two Basidiomycota species, *Botrytis cinerea*, and *Plectania melastoma*, which corroborates previous findings (Bewick et al., 2019; Z. Zhang et al., 2021). *Dim-2* was identified only in seven Ascomycota outgroup species, and matches in *B. cinerea* and *Zymoseptoria tritici* were included after an additional TBLASTN search in NCBI (Table S6). The original *Dim-2* absence in *B. cinerea* by local Hmmsearch is caused by the older version and fragmented assembly at JGI compared to NCBI. The *Z. tritici* IPO323 strain used in this study carries a non-functional *Dim-2* as previously reported (Moller et al., 2021) and no significant hit was found with Hmmsearch or BLASTP searches, but several fragmented hits with TBLASTN search in NCBI. This loss of *Dim-2* function in *Z. tritici* IPO323 is very recent because other isolates harbor a functional *Dim-2* gene (Moller et al., 2021). Its loss leads to the deactivation of transposable elements (TEs) and decreased mutation rate, thus influencing genome defense and genome evolution (Moller et al., 2021). Integration of a functional *Dim-2* allele in Zt09 restores DNA methylation (Moller et al., 2021). In contrast to the *DNMT1* and *Dim-2* losses in Aspergilli, *DNMT5* is identified in 88 out of 94 Aspergilli and 10 out of 15 outgroup species (Figure 1), with a conserved domain architecture consisting of a DNA methyltransferase domain, an SNF2 family domain (PF00176), and a Helicase conserved C-terminal domain (PF00271), in which the latter two domains are involved in transcription regulation (Delmas et al., 1993). The five outgroup species without DNMT5 are *S. cerevisiae*, *S pombe*, *Neurospora crassa*, *Fusarium graminearum*, and *Cladonia grayi*, among which the two yeast species also lack the other two DNA methylation maintaining enzymes, and thus accordingly are considered to be incapable of DNA methylation (Capuano

et al., 2014; Huff & Zilberman, 2014; Proffitt et al., 1984). Five out of the six *Aspergillus* species that lost *DNMT5* belong to a single clade within the *Nidulantes* section (*Aspergillus undulatus*, *Aspergillus olivicola*, *Aspergillus filifera*, *Aspergillus venezuelensis*, and *Aspergillus stella-maris*), and the sixth one is *Aspergillus robustus* which is assigned to the *Circumdati* section (Figure S1). In the Basidiomycota fungus *Cryptococcus neoformans*, DNMT5 maintains DNA methylation using the same cofactors as DNMT1 (Catania et al., 2020; Huff & Zilberman, 2014). Thus, it is most likely that DNMT5 is similarly involved in DNA methylation maintenance in Aspergilli as in *C. neoformans*. Moreover, DNA methylation maintenance does not appear to be required for the survival of fungi as shown with the absence of all three *DNMT1*, *DNMT5*, and *Dim-2* not only within the *Nidulantes* section but also in *S. cerevisiae* and *S. pombe* (Figure 1). However, the absence of these genes could have an impact on the regulation of certain genes like secondary metabolite biosynthetic gene clusters and it would be interesting to compare species within the *Nidulantes* section in more detail to identify how such loss may have affected gene expression and genome organization.

RID is a more recently acquired single domain DNMT that specifically occurs in fungi and is involved in a fungal-specific genome defense mechanism called RIP (Repeat-Induced Point mutation) (Bewick et al., 2019). RID is a putative *de novo* methyltransferase for TEs in both CG and non-CG contexts relating to RIP during the sexual stage in *N. crassa***;** however the capacity of RID to catalyze 5mC has not yet been demonstrated (Rountree & Selker, 2010). Here, we identified RID in all analyzed Aspergilli species as well as in ten outgroup species. A RID-like homolog knock-out mutant in the model filamentous Ascomycota *Podospora anserina* shows blocked sexual development, as gametes formation and fertilization can happen but dikaryotic cells cannot develop normally (Grognet et al., 2019). Interestingly, the knock-out mutant of *DmtA* (homolog of *RID* in *Aspergillus flavus*) shows altered conidiation, sclerotial production, aflatoxin biosynthesis, and virulence (Yang et al., 2016), suggesting that RID might have additional functions next to genome defense. Additionally, fungi lost the gene for *de novo* DNA methyltransferase, *DNMT3,* around 50–150 million years ago (Catania et al., 2020; Madhani, 2020), and only *DNMT5* and *RID* were consistently found in Aspergilli (Figure 1). Thus, we hypothesize that Aspergilli can only maintain DNA methylation *via* DNMT5 but do not spur genome-wide *de novo* DNA methylation. This hypothesis seems to be corroborated as bisulfite sequencing in *A. flavus* has uncovered very little genome-wide DNA methylation (S.-Y. Liu et al., 2012). However, systematic analyses of DNA methylation patterns and experimental evidence for the roles of DNMTs in Aspergilli are still largely missing.

Aspergilli lack *SET7* (*KMT6*), but SET-domain proteins are abundant and other histone methyltransferases are conserved

Histone methylation is typically catalyzed by SET domain-containing proteins (Freitag, 2017; Spellmon et al., 2015). We identified a total of 1,548 SET domain-containing proteins and classified them into 15 groups based on their domain architecture and phylogenetic support

that contain species both from Aspergilli as well as from the outgroup. Among these groups, 107, 111, 109, 102, and 108 sequences represent homologs of the known histone methyltransferases (HMTs) SET9 (KMT5), Dim-5 (KMT1), SET1 (KMT2), Ash1 (KMT2H), and SET2 (KMT3), respectively (Figure 2). The remaining homologs were divided into nine groups that we broadly classified as SET domain-containing proteins. Among them, eight groups contain a SET domain but lack experimentally determined functions, and 117 sequences could be grouped as Ribosomal lysine N-methyltransferases (Figure 2). Enzymes in this group contain a SET domain together with a Rubisco LSMT substrate-binding domain (PF09273). While the homolog in plant transfers a methyl group to the large subunit of the Rubisco holoenzyme complex (Raunser et al., 2009; Trievel et al., 2003), the homolog in *S. cerevisiae* has been characterized as a ribosomal lysine N-methyltransferase (Porras-Yakushi et al., 2007).

The well-described HMTs, SET1, Dim-5, SET2, Ash1, and SET9, all contain distinct accessory domains and catalyze methylation on lysine residues at different positions of H3 or H4 tails (Freitag, 2017). SET1 is composed of the SET domain, the N-SET domain (PF11764), as well as the SET_assoc domain (PF11767), and is responsible for H3K4me1/2/3 in *S. cerevisiae* (Roguev et al., 2001). SET1 occurs in a single copy in all species (Figure 2), indicating that this enzyme and its catalyzed PTM play fundamental roles in fungal biology. Dim-5 contains a SET domain and Pre-SET domain (PF05033) and acts as an H3K9 methyltransferase inducing heterochromatin formation notably in *N. crassa* (X. Zhang et al., 2002). Dim-5 was found in all species included in this study except *S. cerevisiae* and *Aspergillus calidoustus*, among which *S. cerevisiae* is reported to lack H3K9me (Marina et al., 2013; X. Zhang et al., 2002). How such a loss has impacted gene expression in *A. calidoustus* in comparison to its closest relative *Aspergillus carlsbadensis* deserves further investigation. Both SET2 and Ash1 catalyze H3K36me in *N. crassa* and share the SET domain and the AWS domain (PF17907) (Figure 2), yet SET2 harbor one additional SRI domain (PF08236) that interacts with RNA polymerase II and links H3K36me with transcript elongation (M. Li et al., 2005; Strahl et al., 2002; T. Xiao et al., 2003). In *N. crassa,* Ash1 was shown to contribute to H3K36me at sub-telomeric regions, while SET2 acts primarily at euchromatic regions (Bicocca et al., 2018; Janevska et al., 2018). Considering the common origin of SET2 and Ash1 as well as their similar activities, an ancient duplication followed by the loss or gain of the SRI domain is likely to have contributed to their functional diversification (Figure 2). *SET2* was found in all outgroup species, while *Ash1* was lost in both Basidiomycota and two yeasts. False absences of *Ash1* in *B. cinerea* and *F. graminearum* were corrected as they could be retrieved by TBLASTN search at NCBI (Table S6). Both enzymes are conserved in all *Aspergillus* species except in *Aspergillus carbonarius* from the *Nigri* section, suggesting that either other SET domain-containing proteins exhibit functional redundancy or that H3K36 methylation is not fully required in this species. In *Fusarium fujikuroi*, deletion of *SET2* and *Ash1* strongly affected vegetative growth and conidiation (Janevska et al., 2018). SET9 homologs harbor only a SET

domain, and were shown to catalyze H4K20me in *S. pombe* (Y. Wang et al., 2009). This group is well conserved in most of the species except in *S. cerevisiae* (Veerappan et al., 2008) and in the two closely related Aspergilli, *Aspergillus novoparasiticus* and *Aspergillus arachidicola* that both belong to section *Flavi* (Figures 2 and S1).

SET7 is the catalytic subunit of PRC2, and is highly conserved in mammals, plants, and *N. crassa*, in which it has been well-characterized to catalyze methylation on H3K27, a PTM that is deposited at facultative heterochromatin (Jamieson et al., 2013). Even though PRC2 is considered to be a key chromatin modifier complex, its catalytic subunit *SET7* was lost in many fungal species (Freitag, 2017). *SET7* has previously been reported to be absent in *A. nidulans* and *A. fumigatus* (Freitag, 2017; Jamieson et al., 2013, 2016), and we here demonstrate that *SET7* is absent from all analyzed Aspergilli as well as from *Penicillium* (Figure 2), suggesting that this enzyme was most likely lost at or before the divergence of the last common ancestor of these two fungal genera. Thus, in line with previous reports (Bachleitner et al., 2019; Cánovas et al., 2014; Freitag, 2017; Gacek-Matthews et al., 2015), our results indicate that methylation on H3K4, H3K9, H3K36, and H4K20 are likely catalyzed by conserved enzymes in most Aspergilli, while H3K27 methylation is absent. Considering the loss of this methylation is usually reported to trigger deleterious phenotypes in the soil-borne fungus *Verticillium dahliae* (Kramer et al., 2022), humans (Sneeringer et al., 2010), and mice (Tao et al., 2011), investigating how *Aspergillus* species are able to cope with such a loss will likely provide new insights on the histone code and its evolution in fungi.

Importantly, the occurrence of a high number of uncharacterized SET domain-containing proteins opens the possibility that these contribute to not yet characterized histone modifications. Especially, four uncharacterized SET domain-containing groups with 144, 195, 113, and 113 members are conserved in Aspergilli and thus likely play important roles in different cellular processes.

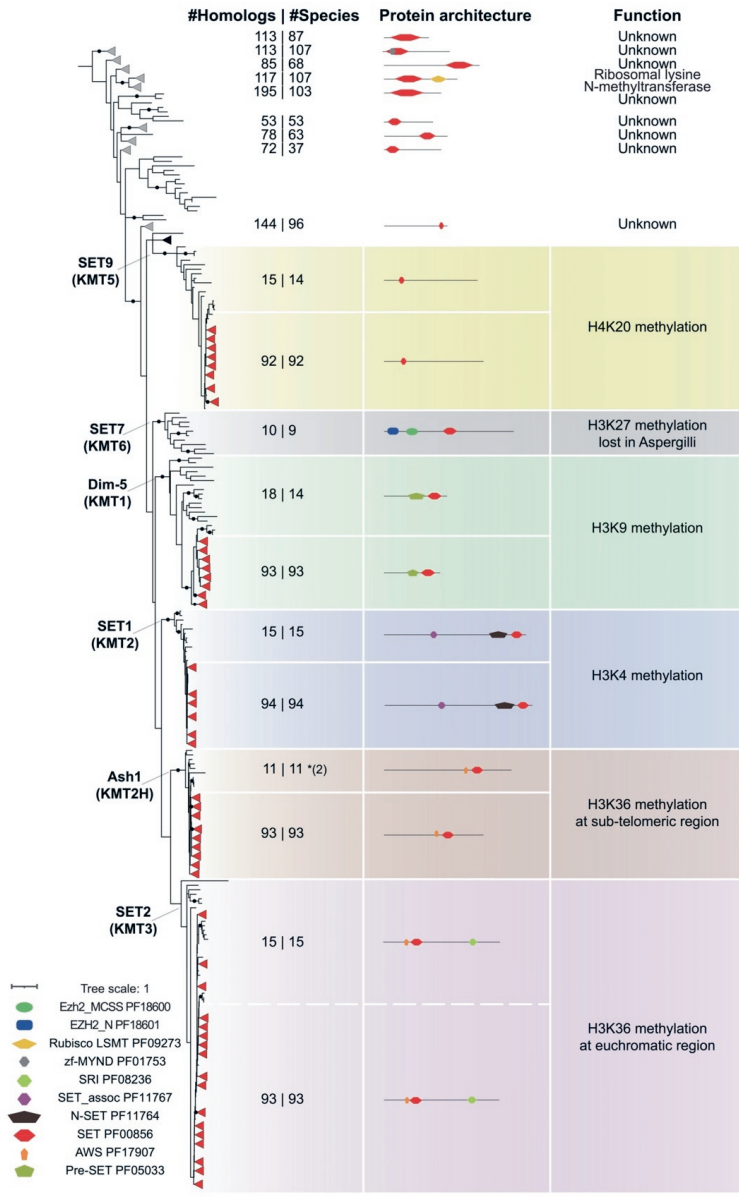


Figure 2. Conservation of histone methyltransferase families in the *Aspergillus* genus and outgroup species. The maximum-likelihood phylogeny of SET domains (PF00856) found in 94 Aspergilli and 15 outgroup species was determined using IQ-TREE (Boratyn et al., 2013; Simão et al., 2015). The black dot on the branch indicates ultrafast bootstrap values over 95 and SH-aLRT values over 80. Leaves from Aspergilli with trustworthy support are collapsed and shown as red triangles. Leaves with similar protein architecture but unknown functions covering both

Aspergilli and outgroup species are collapsed and shown as grey triangles. The numbers of hits and species are separated by the pipe in the '#Homologs | #Species' column. The number in parentheses after * indicates the additional hits found by the TBLASTN search on NCBI. Additional conserved domains in SET domain-containing proteins were identified using the PFAM database (https://pfam.xfam.org).

Histone acetyltransferases are conserved in Aspergilli

Histone acetyltransferases (HATs) have been classified into five families: MYST (MOZ, Ybf2/Sas3, Sas2, and Tip60)-related HATs, Gcn5-related acetyltransferases (GNATs), p300/CBP HATs, general transcription factor HATs, and nuclear hormone-related HATs (Carrozza et al., 2003; Sterner & Berger, 2000; Torchia et al., 1998). We here only focused on the first two HAT families because these are known to be involved in the regulation of histone proteins and appear largely conserved in fungi (Carrozza et al., 2003).The MOZ/SAS FAMILY domain is the catalytic domain of MYST-related HATs (Roth et al., 2001). As previously reported (Kawahara et al., 2008), the three characterized MYST-related HATs, Sas3, Sas2, and Esa1, belong to distinct monophyletic clades, and these paralogues are nearly fully conserved in the here analyzed fungi (Figure 3). MOZ/SAS FAMILY domain and zf-MYST domain (PF17772) are present in Sas3, which can on its own acetylate histone H3 and H4, and H2A weakly *in vitro* (Takechi & Nakayama, 1999). Within the NuA3 (Nucleosomal Acetyltransferase of histone H3) complex, Sas3 histone acetylation activity is *in vitro* restricted on H3K14 and to a less extent on H3K23, while in *S. cerevisiae*, only H3K14ac can be detected (Taverna et al., 2006). *Sas3* occurs in most of the species but is lost in *S. pombe*, as previously observed (Gómez et al., 2005; Nugent et al., 2010), and is duplicated in *A. carbonarius* (Figure 3). In addition to the two domains found in Sas3, Sas2 and Esa1 carry a PHD_4 domain (PF16866) or Tudor-knot domain (PF11717), respectively (Figure 3). *Esa*1 is present as a single copy in all species included in this study, and its functions is known to contribute to the acetylation of K5, K8, K12, K16 on histone H4, as well as on H2AZK14 in *S. cerevisiae* (Allard et al., 1999; Kim et al., 2018; Millar et al., 2006; Searle et al., 2017). *Sas2* is similarly found in most of the species except for two Basidiomycota as well as for *A. carbonarius*. Sas2 is involved in chromatin-mediated gene regulation and H4K16 acetylation in both *S. cerevisiae* and *Candida albicans* (Kim et al., 2018; Meijsing & Ehrenhofer-Murray, 2001; X. Wang et al., 2013). Although Esa1 and Sas2 catalyze H4K16ac, it was shown in *C. albicans* that they are differentially recruited at different stages of development (Kim et al., 2018). Because both Esa1 and Sas2 are well-conserved in the *Aspergillus* genus, we hypothesize that these two HATs may also have distinct roles during development.

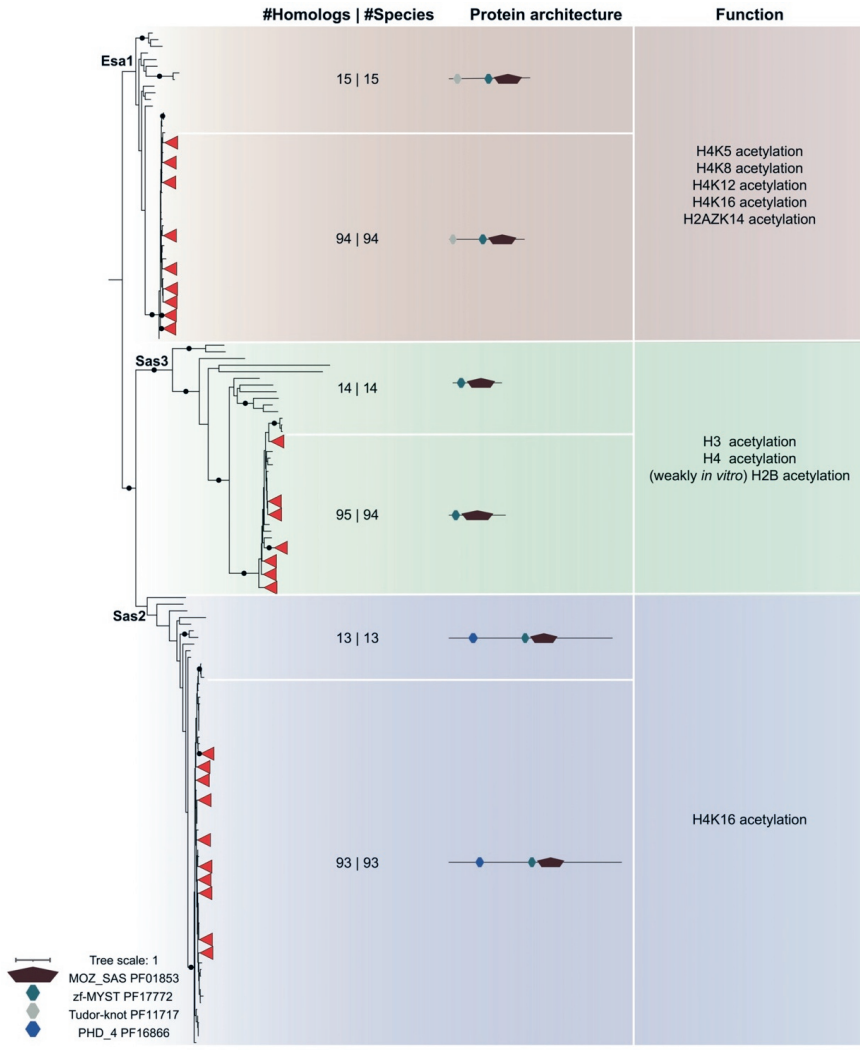


Figure 3. Conservation of MOZ/SAS FAMILY domain-containing histone acetyltransferases in the *Aspergillus* genus and outgroup species. The maximum-likelihood phylogeny of MOZ/SAS FAMILY domains (PF01853) found in 94 Aspergilli and 15 outgroup species was determined using IQ-TREE (Boratyn et al., 2013; Simão et al., 2015). The black dot on the branch indicates ultrafast bootstrap values over 95 and SH-aLRT values over 80. Leaves from Aspergilli with trustworthy support are collapsed and shown as red triangles. The numbers of hits and species are separated by the pipe in the '#Homologs | #Species' column. '*In vitro*' indicates the acetylation activity of this enzyme was only observed on the substrate *in vitro*. Additional conserved domains in MOZ/SAS FAMILY domain-containing proteins were identified using the PFAM database (https://pfam.xfam.org).

GNATs have a acetyltransferase domain as its catalytic core, and 2,604 acetyltransferase domain-containing proteins were found in our search (Figure S2). The genes for two well-characterized GNATs, *Gcn5* and *Elp3*, occur in a single copy in every species included in our study (Figure S2). Gcn5 harbors a Bromodomain (PF00439) in addition to the Acetyltransferase domain and is a crucial subunit of distinct histone acetyltransferase complexes such as the HAT-A2 complex, the SAGA complex, and the ADA complex in *S. cerevisiae* (Carrozza et al., 2003; Sterner & Berger, 2000). All these complexes contain two coactivator proteins, Ada2 and Ada3, that directly bind to Gcn5 (Balasubramanian et al., 2002) and the recruitment of different other subunits determines the location where they catalyze acetylation (Carrozza et al., 2003). Elp3 is characterized by the Acetyltransferase domain, Radical SAM (PF04055), and Radical SAM C-terminal domain (PF16199), and has been shown to catalyze H3K14ac and H4K8ac in *S. cerevisiae* (Krogan & Greenblatt, 2001; Winkler et al., 2002; Wittschieben et al., 1999). Elp3 also exhibits a dual enzymatic activity as it catalyzes demethylation in mouse zygotes (Okada et al., 2010), which is due to the Radical SAM domain that catalyzes the demethylation of methylated lysyl residues on histones (Booker & Grove, 2010).

Next to Gcn5 and Elp3, we have identified 2,386 additional acetyltransferase domain-containing proteins. Because their functions remain unknown due to the lack of homology to sequences with characterized functions, they are valuable candidates that may acetylate histones or other proteins. Interestingly, it has been shown that Gcn5 also exhibits succinyltransferase activity in humans (Y. Yuan et al., 2020) and can also function as histone crotonyltransferase to regulate gene transcription *in vitro* (Kollenstart et al., 2019). It is expected that other less common histone modifications, i.e., histone lysine butyrylation, propionylation, or malonylation, may also be catalyzed by histone acetyltransferases because these histone modifications all use acyl-Coenzyme A to complete the reactions (Jones, 2014; B. Liu et al., 2009; Papanicolaou et al., 2014). Further research is needed to determine whether these less common histone modifications are catalyzed by the known histone acetyltransferases Gcn5 and Elp3, or as of yet unstudied HATs. In conclusion, all five characterized fungal HATs described above are highly conserved in the *Aspergillus* genus, showing their importance in fungal biology. Many other acetyl-transferase domain-containing proteins are encoded in *Aspergillus* genomes and may be responsible for histone acetylation or additional less common histone modifications like butyrylation, succinylation, and malonylation.

Duplication of *RpdA* histone deacetylase gene in the *Flavi* and *Janorum* sections

Five families of histone deacetylases (HDACs) can be grouped into two classes based on sequence similarity: class I (HOS1, HOS2, and Rpd3/RpdA) and class II (HdaA and HOS3) (Bernstein et al., 2000; Ekwall, 2005). They all harbor the Histone deacetylase catalytic domain, and our phylogenetic tree corroborates the HDAC classification into two distinct

classes (Figure 4). *HdaA* is found as a single copy in each species included in this study (Figure 4). It harbors the Histone deacetylase domain and C-terminal Arb2 domain (PF09757) that serves as an anchor to target centromeric heterochromatin region (Honda et al., 2016; Job et al., 2016). HdaA deacetylates lysine in histones H3 and H2B, but not H4 nor H2A in *S. cerevisiae* (Kurdistani & Grunstein, 2003; Wu et al., 2001), and the Δ*hdaA* knock-out strains show increased or reduced production of different secondary metabolites in *A. nidulans* (X. Li et al., 2019), *A. niger* (X. Li et al., 2019), and *A. fumigatus* (Lee et al., 2009). By contrast, *HOS3* is found in only 80 *Aspergillus* species, as it is lost in nine species (*Aspergillus nomiae*, *Aspergillus pseudocaelatus*, *Aspergillus tamarii*, *Aspergillus pseudotamarii*, *Aspergillus caelatus*, *Aspergillus sergii*, *A. novoparasiticus*, *A. minisclerotigenes*, *A. arachidicola*) in the *Flavi* section, three species (*A. desertorum*, *Aspergillus stercorarius*, and *A. multicolor*) in the *Nidulantes* section, and two species (*Aspergillus heteromorphus* and *A. carbonarius*) in the *Nigri* section. *HOS3* is well conserved in all outgroup species except *S. pombe* (Figure 4) (Brosch et al., 2008). HOS3 has been shown to catalyze deacetylation with distinct specificity for histone H2AK7, H2BK11, H3K14 and H3K23, as well as H4K5 and H4K8 in *S. cerevisiae* cell extracts (Carmen et al., 1999). Notably, *HOS1* was only detected in three outgroup species, *S. cerevisiae* and both Basidiomycota (Figure 4). The specificity of Hos1 in catalyzing deacetylation on histone proteins is not known yet (Pérez-Martínez et al., 2020), but it can catalyze the deacetylation of the Smc3 subunit of Cohesin and influence the regulation of chromosome segregation during mitosis in *S. cerevisiae* (S. Li et al., 2017). *HOS2* is found in 15 outgroup species and most of Aspergilli except *A. oryzae* (Figure 4). It is known to deacetylate lysine residues in H3 and H4 histone tails (A. Wang et al., 2002). The conservation of *HOS2* in Aspergilli suggests that it plays a key role, which is supported by the deletion of *HOS2* in *A. niger* where the deletion strains displays severe reduction in growth, sporulation, SM biosynthesis, and stress resistance (X. Li et al., 2019).

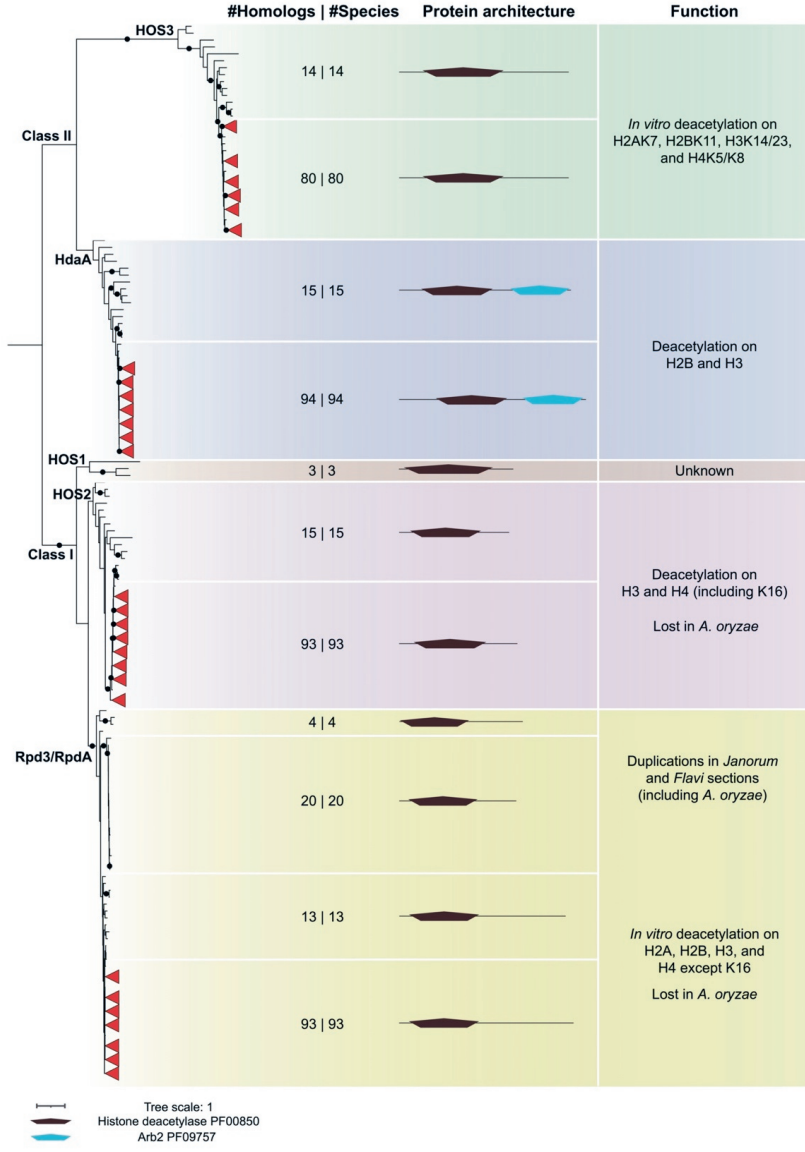


Figure 4. Conservation of histone deacetylase domain-containing proteins in the *Aspergillus* genus and outgroup species. The maximum-likelihood phylogeny of Histone deacetylase domains (PF00850) found in 94 Aspergilli and 15 outgroup species was determined using IQ-TREE (Boratyn et al., 2013; Simão et al., 2015). The black dot on the branch indicates ultrafast bootstrap values over 95 and SH-aLRT values over 80. Leaves from Aspergilli with trustworthy support are collapsed and shown as red triangles. The numbers of hits and species are separated by the pipe in the '#Homologs | #Species' column. '*In vitro*' indicates the

deacetylation activity of this enzyme was only observed on the substrate *in vitro*. Additional conserved domains in histone deacetylase domain-containing proteins were identified using the PFAM database (https://pfam.xfam.org).

Within class I HDACs, the Rpd3/RpdA clade is divided into two sub-clades (Figure 4). One is nearly fully conserved (except in *A. oryzae*) and comprises the well-characterized Rpd3/RpdA HDACs, while the other one only comprises sequences from the *Flavi* and *Janorum* sections (Figure 4). Our phylogenetic reconstruction suggests a duplication event that has occurred in the ancestor of all Aspergilli giving rise to a paralog that was retained in two sections only and could explain the *A. oryzae* loss in the conserved sub-clade. The two paralogs from the same species vary in length as they share a conserved N terminal catalytic domain but have a different C terminus (Figure 4). Follow-up studies are needed to investigate the function of both paralogues in *Aspergillus* species. Rpd3 is a crucial deacetylase in eukaryotes, as it is required for deacetylation in most locations on core histone proteins, except H4K16 *in vitro* (Suka et al., 2001) that is catalyzed by HOS2 (A. Wang et al., 2002). *S. cerevisiae* lacking *Rpd3* is sensitive to high osmolarity and shows compromised expression of osmotic stress genes (De Nadal et al., 2004). Deletion of *RpdA* in *A. nidulans* is lethal (Tribus et al., 2010), while only heterokaryon disruptants could be obtained in *A. oryzae* (Kawauchi et al., 2013), indicating its crucial role in the fungal genus *Aspergillus*. In summary, class I and class II deacetylases are well conserved in Aspergilli, and duplication events like the one observed for *RpdA* may confer new activities. While only one paralog was retained in *A. oryzae*, the presence of two paralogs in sections *Flavi* and *Janorum* may provide functional redundancy or diversification, which remains to be determined.

Lastly, we reconstructed the phylogeny of histone demethylases harboring the Jumonji C (JmjC) domain (PF02373) (Figure S3). Six groups are distinguished by their domain distribution: KDM4A/JHDM3/JMJD2, KDM5/JARID, DMM-1, JMJD6/PKDM11, KDM2/Epe1, and KDM6B/JMJD3, which also agrees with a previous study (Gacek-Matthews et al., 2015). The first five groups occur in most of the outgroup and *Aspergillus* species, but the *KDM6B/JMJD3* is only present in several outgroup species but not in any of the *Aspergillus* species. Certain JmjC domain-containing proteins were shown to catalyze the demethylation of histone proteins, e.g., KDM5/JARID for H3K4me2/3 and KDM4A/JHDM3/JMJD2 for H3K9me2/3 and H3K36me2/3 (Klose et al., 2006). The histone demethylase activity has not been detected for DMM-1 (Tamaru, 2010) or Epe1 (Isaac et al., 2007) in *N. crassa*, while KDM2 can demethylate H3K36me1/2 in mouse (Turberfield et al., 2019) and JMJD6/PKDM11 is an arginine demethylase in *Arabidopsis* (Qian et al., 2015). Interestingly, the human KDM6B/JMJD3 is responsible for the demethylation of H3K27me2/3 (Hong et al., 2007). As we did not find *SET7* and thus conclude that H3K27me3 is absent, it would be tempting to speculate that H3K27 demethylase should be absent in *Aspergillus* too. However, based on the fact that other enzymes, e.g., UTX can also catalyze the demethylation of H3K27 (Agger

et al., 2007; Hong et al., 2007), we cannot unambiguously correlate the absence of *KDMB6B/JMJD3* to *SET7* loss. More studies are needed to investigate the regulation mechanism and interplay of proteins for H3K27 methylation.

Presence-absence patterns of subunits that compose histone modifier complexes

Histone PTMs are catalyzed by the conserved domains found in HMTs, HATs, and HDACs (Figure 2-4 and S2), but these chromatin modifying enzymes are often not acting alone. Instead, they are part of protein complexes that comprise accessory subunits, which are involved in the recognition and induction of specific histone modifications at the correct locations (Gates et al., 2018; Hajheidari et al., 2019). To further investigate the presence-absence patterns (PAPs) of known chromatin modifier complexes in Aspergilli, we focused on 15 chromatin modifier complexes with 83 subunits previously identified and studied in *S. cerevisiae* or *N. crassa* (Table S4). Forty-four subunits are present in all *Aspergillus* and outgroup species. Most of the accessory subunits are either fully conserved or absent in Aspergilli (Figure 5 and Table S4). Ten subunits show PAPs in both Aspergilli and outgroup species, while nine subunits appear specific to *S. cerevisiae* as they are not found in any other species and are likely to play a specific role in this species. Seventeen subunits show PAPs in the outgroup species, while eight are absent and nine are present in the Aspergilli. Lastly, three subunits are present in all outgroup species but show PAPs in Aspergilli.

Five histone methyltransferase complexes, COMPASS, DCDC (Dim-5/-7/-9/CUL4/DDB1 Complex), SET2 Complex, PRC2, and Dot1 (disruptor of telomeric silencing 1) (KMT4) were included in our analyses (Figure 5). Five COMPASS subunits (*SET1*, *CclA*, *Swd1*, *Swd2*, and *Swd3*) are fully conserved, suggesting an important function for H3K4 methylation. Consistent with a PAP in most outgroup species and the loss in Aspergilli, Spp1p is not essential for H3K4 methylation (Dehé et al., 2006). As *Sdc1* shows a PAP in Aspergilli and outgroup species, it is likely that this subunit also does not play a key role in H3K4 methylation. In the DCDC complex, while Dim-7 and Dim-9 seem to be dispensable for H3K9 methylation, Dim-5, Cul4, and Dim-8 are highly conserved, and we thus expect these enzymes to play key roles in the activity of this complex (Figure 5 and Table S4). However, it was shown that Dim-7 and Dim-9 are responsible for heterochromatin recognition, which leads to subsequent recruitment of the complete DCDC to induce H3K9 methylation in *N. crassa* (Lewis et al., 2010). Thus, it is very likely that other accessory subunits are involved in recruiting DCDC at specific chromatin locations in Aspergilli. Furthermore, the unique absence of *Dim-8* in *Aspergillus indicus* may suggest altered H3K9 methylation in this species. Consistent with the absence of *SET7* (Figure 2), all components of the PRC2 complex are lost in Aspergilli (Figure 5). Lastly, the interaction and crosstalk of multiple chromatin modifications in *Aspergillus* species should be further explored. For example, heterochromatin formation and regulation are done by H3K9me3, DNA methylation, and H3K27me3 in *N. crassa*. In this species, H3K9me3 is recognized by an adapter protein, HP1, to recruit DNA methyltransferase Dim-2 for DNA

methylation (Rountree & Selker, 2010), and cause the redistribution of H3K27me3 (Jamieson et al., 2016). In *Aspergillus* species, although HP1 and the writer for H3K9me3, DCDC complex, are well conserved (Figure 5), the inability for genome-wide *de novo* DNA methylation and H3K27me3 suggest different mechanisms of heterochromatin formation and regulation in Aspergilli which remain to be determined.

We included eight histone acetyltransferase complexes that differ by their distinct catalytic subunits (Figure 5). The HAT-A2, ADA, SAGA, and Elongator complexes use an acetyltransferase domain-containing protein as catalytic subunit (Eberharter et al., 1999; Georgakopoulos et al., 2013). The SAS (Something About Silencing), NuA4 (Nucleosomal Acetyltransferase of histone H4), and NuA3 complexes use a MOZ/SAS FAMILY domain-containing protein (Doyon & Côté, 2004). The NuB4 (Nuclear Hat1p-containing type B histone acetyltransferase) complex uses the histone acetyltransferase HAT1 domain-containing protein (Verzijlbergen et al., 2011). Gcn5, Ada2, and Ada3 are fully conserved in all fungi and constitute the core enzymatic module of the HAT-A2, ADA, and SAGA complexes. This catalytic module contributes to catalyzing histone acetylation, and other subunits in these complexes guide them to specific locations and link them to the transcriptional machinery (Sendra et al., 2000). Besides the catalytic module, the ADA complex also harbors the Ahc1 subunit that is found in *S. cerevisiae* only (Eberharter et al., 1999; Grant et al., 1997). Twelve additional subunits of the SAGA complex are well conserved in Aspergilli and outgroup species (Figure 5). The SAGA complex also contains a deubiquitylase (DUB) module, for which the composition differs between Aspergilli and other species. *Sgf11* is unique to *S. cerevisiae* and *Sus1* was likely lost in the ancestor of Pezizomycotina because we can only identify *Sus1* orthologs in both Basidiomycota and the yeasts (Figure 5). The Elongator complex appears more variable as four subunits are mostly conserved, but *Elp6* is specific to *S. cerevisiae*, and *Hpa2* and *Hpa3* show PAPs in both Aspergilli and outgroup species.

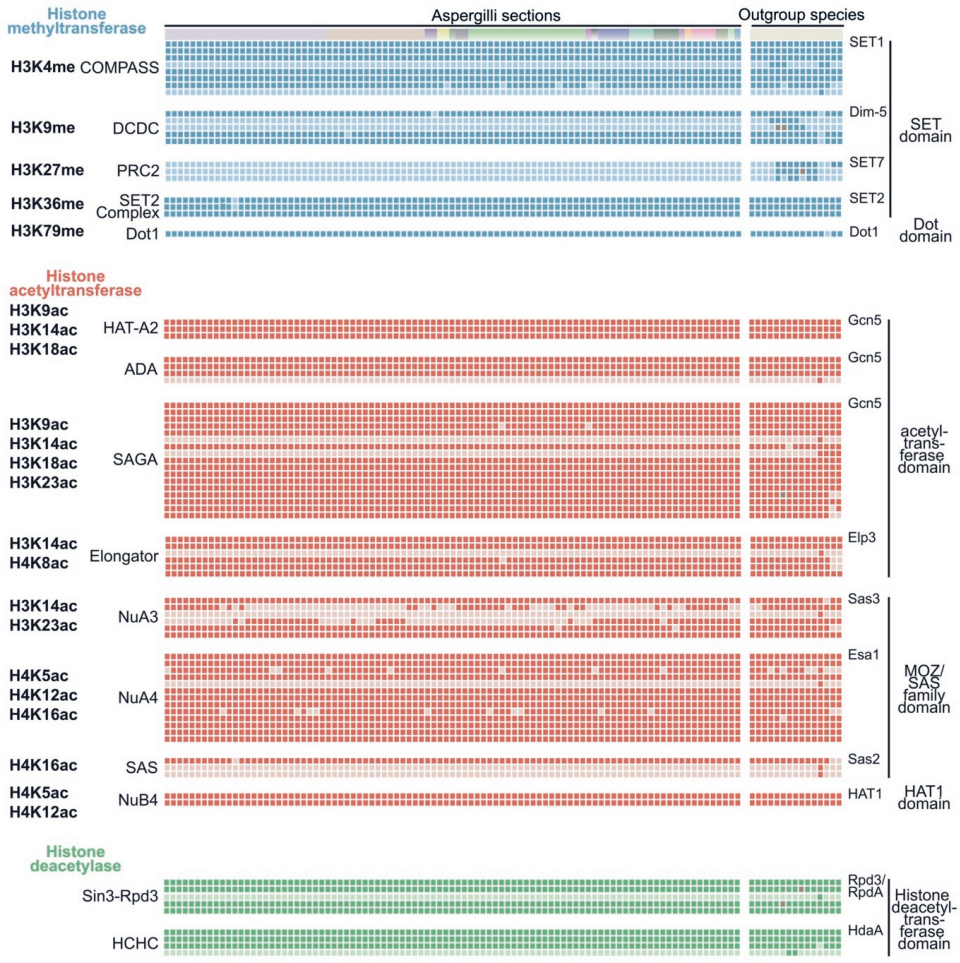


Figure 5. Presence-absence patterns of subunits that constitute characterized histone modification complexes. The top row with distinct color boxes indicates outgroup species and Aspergilli sections, the order of the species from right to left is the same as in the species tree from top to bottom (Figure S1) without *Fusarium oxysporum* and *Aspergillus kawachii* since they were excluded in the later analysis. Each row represents a subunit in the histone modification complexes, the order agrees with the summary list (Table S4). Each row shows the presence/absence of a specific complex subunit. For each complex, the first row represents the catalytic subunit. Each box in the matrix indicates the presence (dark color) or absence (light color) of a subunit (Table S4). The brown boxes indicate additional hits found with TBLASTN search on NCBI (Table S6).

The NuA4, SAS, and NuB4 complexes are well-conserved in Aspergilli as well as in the outgroup species (Figure 5). These complexes appear slightly different in *S. cerevisiae* with the presence of specific subunits (*Eaf5* in the NuA4 complex; *Sas4* and *Sas5* in the SAS

complex). Both SAS and NuA4 complexes can catalyze H4K16ac with opposite effects during cell growth and morphogenesis in *C. albicans* (Kim et al., 2018). It would be interesting to assess whether these two complexes exhibit similar activities in Aspergilli. For the NuA3 complex, two accessory subunits are well conserved among Aspergilli while two others, *Pdp3*, and *Nto1*, show different PAPs, and *Yng1* is specific to *S. cerevisiae*. *Pdp3* and *Nto1* seem to be lost in a complementary manner, suggesting a possible functional redundancy (Figure 5). Finally, two histone deacetylase complexes, Sin3-Rpd3 and HCHC, are fully conserved (Figure 5). The Sin3-Rpd3 complex is slightly different in *S. cerevisiae* with the specific of the *Ume1* subunit. Overall, our analysis showed that the core catalytic modules of complexes involved in histone PTMs are mostly conserved in all fungi, including Aspergilli, but variability in accessory subunits could contribute to differences in the localization of histone PTMs along the genome.

Unbiased mass spectrometry captures the diverse histone modifications in *A. nidulans* and verifies the absence of H3K27 methylation

Our genomic and phylogenetic analyses have identified the presence of multiple chromatin modifiers in Aspergilli, suggesting that these species have the ability to establish a wide variety of histone modifications. To assess the presence of histone PTMs predicted in our evolutionary analyses (Figures 2-5), we extracted histone proteins from *A. nidulans* and performed unbiased mass spectrometry analyses (Figures 6A and 6B). H3 and H4 peptides with PTMs corresponding to H3K4me, H3K9me, H3K36me, and H3K79me were detected (Figure 6C), which we anticipated being catalyzed by SET1, Dim-5, Ash1 and SET2, and Dot1, respectively. H4K20me, which is catalyzed by SET9 in *S. pombe* (Y. Wang et al., 2009), was not detected (Figure 6C), which is likely due to technical reasons because the unmodified form of histone H4 peptide 20-24 containing the K20 residue, was not detected. For the acetylation on H3K9, H3K14, H3K18, H3K23, H4K5, H4K8, H4K12, and H4K16, they were confidently detected (Figure 6C). We did not observe any peptide that would indicate the presence of H3K27me in *A nidulans*, while we did detect H3K27ac (Figure 6C and Table S3). These results corroborate our computational analysis that the PRC2 complex is absent in *A. nidulans*, as well as previous reports on the absence of H3K27me3 in several Aspergilli (Buscaino, 2019; Erlendson et al., 2017; Freitag, 2017). Thus, the mass spectrometry results verify the occurrence of histone modifications predicted to be present based on our comparative analyses in the fungal genus*Aspergillus*. Moreover, the differences between the relative abundance of distinct histone modifications provide interesting information to investigate these abundances and identify specific conditions under which these modifications increase or decrease in abundance.

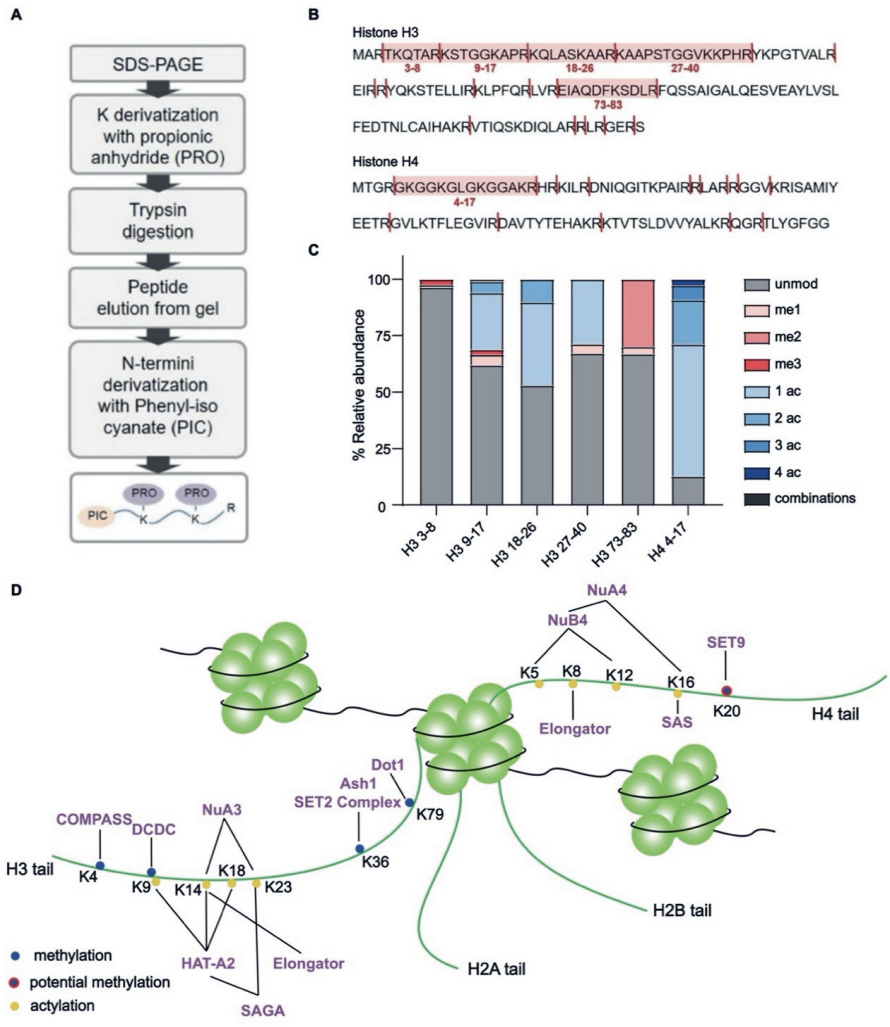


Figure 6. Mass spectrometry unlocks the histone code in Aspergilli. (A) Schematic presentation of the in-gel digestion method used to process histones isolated in *A. nidulans*. After SDS-PAGE separation, gel bands corresponding to histone H3 and H4 are excised and chemically acylated with propionic anhydride (which derivatizes unmodified and mono-methylated lysines). Thanks to the derivatization, trypsin cutting at lysines is impaired, and histones are cleaved only at the C-terminus of arginine. Finally, digested peptides are N-terminally derivatized with phenyl isocyanate (PIC). (B) *A. nidulans* histone H3 and H4 sequences. Cleavage sites after PRO-PIC digestion are shown. Peptides identified in this study are highlighted in light red and labeled using their starting-ending amino acids. (C) Relative abundance of post-translational modifications detected on peptides of histone H3 and H4 from *A. nidulans* using unbiased mass spectrometry analysis. unmod: unmodified peptide; me1: monomethylation; me2: dimethylation; me3: trimethylation; ac: acetylation. The term “combinations” refers to multiple modifications occurring

on the same peptide (Table S3). (D) Histone modifications are predicted by genomic and phylogenetic analyses, and/or proved by mass spectrometry in *A. nidulans*. Histone modifications verified by mass spectrometry are labeled as solid dots, while the ones that are only predicted by phylogenic analysis are labeled as solid dots with a red outline.

Conclusion

Changes in chromatin organization play an important role in regulating the expression of genes (Brosch et al., 2008). The fungal genus *Aspergillus* is well known for its ability to thrive in a broad range of different environmental niches and the huge potential of secondary metabolite biosynthesis (Collemare & Seidl, 2019; Steenwyk et al., 2020). However, a detailed and complete overview of occurrence and evolution of chromatin modifiers in Aspergilli is a prerequisite to be able to better understand how these fungi regulate gene expression. In the future, more research is needed to uncover how the absence of *de novo* DNA methylation impacts genome evolution in Aspergilli. We demonstrate that PRC2 is lost throughout Aspergilli, a finding consistent with the lack of H3K27me detection by mass spectrometry. Although histone acetyltransferases and deacetylases are overall conserved, our study also revealed specific duplication and loss events, which require further investigation to understand their impact on chromatin and gene regulation in certain *Aspergillus* species. Moreover, other less common histone modifications like histone lysine succinylation, crotonylation, butyrylation, propionylation, and malonylation, are all likely catalyzed by HATs and HDACs as they all use Coenzyme A (CoA) to complete the reactions (Jones, 2014; B. Liu et al., 2009; Papanicolaou et al., 2014), but may require specific yet to be discovered subunits. In conclusion, our study provides a detailed overview of the evolutionary routes of chromatin modifier complex in the fungal genus *Aspergillus*, and therefore generates the necessary framework to identify targets for functional studies to understand how chromatin regulates gene expression in Aspergilli.

Material and Methods

Acquisition of predicted proteomes for 109 fungal species

The predicted proteomes of 109 fungal species were retrieved from JGI (Joint Genome Institute) MycoCosm (https://mycocosm.jgi.doe.gov/mycocosm/home) (Grigoriev et al., 2014) on 24th January 2020. These species comprise 94 different Aspergilli species, 13 Ascomycota species from other genera, as well as two Basidiomycota species (*Coprinopsis cinerea* and *Dendrothele bispora*) (Table S1).

Phylogenetic analyses of chromatin modifiers

Phylogenetic trees for six major chromatin modifiers were generated using their conserved catalytic domains. Catalytic domain HMMs (Hidden Markov Model) of DNA methyltransferase domain (PF00145), histone methylation SET domain (PF00856), acetyltransferase domain (PF00583), histone acetyltransferase MOZ/SAS FAMILY domain (PF01853), histone deacetylase domain (PF00850), and histone demethylase Jumonji C (JmjC) domain

(PF02373) were downloaded from the Pfam database (https://pfam.xfam.org). Hmmsearch, which is part of the HMMER package (v3.1b2) (Eddy, 2011), with the Gathering Cut-Off threshold (-cut_ga) was used to identify occurrences of these domains in the predicted proteomes of all 109 fungal species considered here. To filter out short and incomplete fragments, we applied a cutoff calculated by 50% length coverage of the query sequence (domain) or the hit. Subsequently, the identified domain sequences were aligned using MAFFT v7.271 (Katoh & Standley, 2013). TrimAl v1.2 (-gt 0.1) was used to remove positions in the alignments with gaps in more than 90% of the sequences (Capella-Gutiérrez et al., 2009). Maximum-likelihood phylogenetic trees were constructed using IQ-TREE v1.6.10 (Chernomor et al., 2016; Nguyen et al., 2015) using ultrafast bootstrap and Shimodaira–Hasegawa approximate likelihood ratio test (SH-aLRT) (Anisimova et al., 2011) with 1,000 pseudo-replicates, and standard model selection to automatically determine the best-fit model. We predicted additional domains in the analyzed proteins using Hmmsearch with all Pfam domain profiles that are present in the Pfam database (v28.0) and visualized them along the phylogenetic trees using iTOL (Letunic & Bork, 2019). Lastly, additional TBLASTN searches were performed on the NCBI website (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and checked manually to confirm dubious absences in the domain trees.

Species tree construction

To construct a species phylogeny of the here analyzed fungi, we used BUSCO v4.0.1 (Benchmarking Universal Single-Copy Ortholog assessment tool) (Simão et al., 2015) to identify the occurrence of 758 fungal single-copy genes (Table S2) within the 109 predicted proteomes (Table S1). The identified homologs of single-copy BUSCO genes were subsequently concatenated into a single super-alignment. Then, the maximum-likelihood species phylogeny was estimated using IQ-TREE v1.6.10 (Chernomor et al., 2016; Nguyen et al., 2015) by employing partitioned model selection, which allows selecting a substitution model for each BUSCO gene. Branch supports were obtained using ultrafast bootstrap and SH-aLRT (Anisimova et al., 2011) with 1,000 pseudo-replicates separately.

Chromatin modifier subunits homology search and gene tree construction

We obtained previously studied chromatin modifier complexes from literature, and the associated protein sequences were downloaded from NCBI (https://www.ncbi.nlm.nih.gov) (Table S4). For each subunit of the chromatin modifier complexes, homologs in the analyzed 109 predicted proteomes were detected based on sequence similarity searches. First, BLASTP (v2.2.31+) (Boratyn et al., 2013) searches were performed, and an e-value threshold of 10^{-2} was applied. Multiple BLAST matches (High-Scoring Pairs: HSPs) of the same query were concatenated if they were aligned with different non-overlapping regions. Only the non-redundant region was retained when HSPs of the same query were overlapping. For these two conditions, the highest e-value of the different matches was used to represent the merged match. Only matches with the e-value lower than 10^{-5} and covering more than 50% of query

or subject were retained for the following analyses. If we observed less than ten hits for the BLASTP search after filtering in Aspergilli, we used PSI-BLAST (Position-Specific Initiated BLAST v2.2.31+) (Boratyn et al., 2013) with five iterations (-num_iterations 5) using the best Aspergilli hit from the BLASTP as the query; if no Aspergilli match was retained, we used the same query as for the BLASTP searches. The same filter criterion as for the BLASTP selection was applied to the PSI-BLAST hits, and the results of these two homology search strategies were subsequently jointed.

To account for apparent protein absences due to erroneous gene annotation, we tested these potential losses by using TBLASTN (BLAST v2.2.31+) (Boratyn et al., 2013) to query protein sequences against the genome assemblies. Matches with e-value lower than 10^{-5} together with 4,500 nucleotides upstream and downstream were extracted, and Exonerate v2.2.0 (Slater & Birney, 2005) was used to predict the gene structure (options: protein2genome, extensive search, showtargetgff, showcigar, and showquerygff options were set as true, and ">%ti(%tcb - %tce)\n%tcs\n" option was used to output the complete aligned coding sequence). Transeq v6.6.0.0 (Rice et al., 2000) was used to translate the coding sequence into protein sequences. Together with the protein sequences identified from BLASTP and PSI-BLAST, all protein sequences from each subunit were retrieved to build a phylogenetic tree (MAFFT, TrimAl, and IQ-TREE settings were the same as for the tree reconstruction using the catalytic domains above).

For each sequence included in the phylogenic trees, the domain organization was identified and visualized as described above. When the domain organization of a *Aspergillius* protein differed from its orthologues, the corresponding gene model was manually checked in the MycoCosm genome browser (Grigoriev et al., 2014) and corrected. The presence and absence patterns of 83 subunits from 15 different chromatin modifier complexes were visualized using the Seaborn library in Python (Hunter, 2007; Waskom, 2021). Lastly, additional TBLASTN searches for dubious absences in the outgroup species were performed on the NCBI website as absences could be caused by older and more fragmented assemblies in the JGI database compared to NCBI.

Histone enrichment and mass spectrometry

Spores of 4- or 5-days old sporulating *A. nidulans* (AnWT pabaA1 veA1, kindly provided by Prof. Joseph Strauss) growing on ME (Malt Extract) agar medium were harvested with 2 mL sterile water and cultured in 50 mL ME liquid medium at 28 ℃, 200 rpm for 16 hours. The sample was centrifuged at 8,000 g for 10 minutes to discard supernatant, and germinating spores were resuspended in 25 mL 0.8 M NaCl and centrifuged again. Ten mL protoplasting solution (20 mg/mL *Trichoderma* lysing enzyme and 5 mg/mL driselease in 0.8 M NaCl) were used to resuspend spores, which were then incubated at 30 ℃, 100 rpm for 2-3 hours.

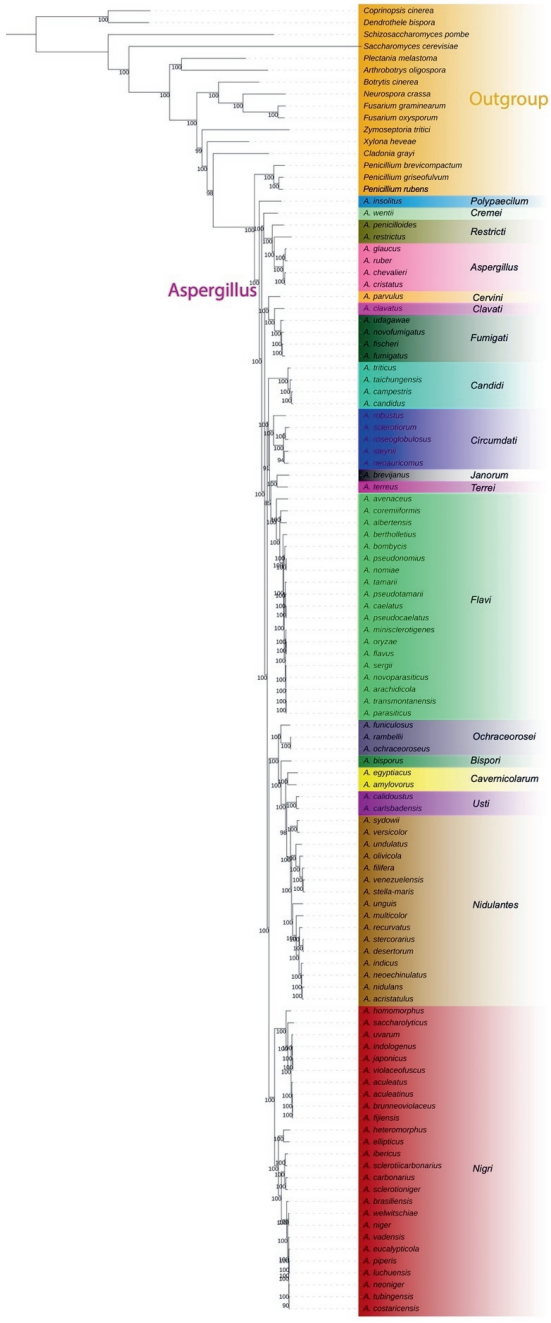
Protoplasts were retrieved by filtering through sterile Miracloth and were concentrated by centrifugation at 3,000 rpm, 4 ℃ for 5 minutes. Histone proteins were extracted according to an adapted histone protein extraction protocol (Noberini et al., 2020). Briefly, 10^8 protoplasts were resuspended in 1 mL Nuclei Isolation Buffer (1×PBS, 0.5 mM PMSF, 5 μM Leupeptin, 5 μM Aprotinin, and 5mM Na-butyrate) with 1% Triton X-100, followed by vigorous pipetting through a 200 μL pipette several times. Then, centrifuged at 2,300 g for 15 minutes at 4 ℃, discarded the supernatant, and resuspended the nuclear pellet in 200 μL Nuclei Isolation Buffer with 0.1% SDS and 250 U Pierce™ Universal Nuclease and incubated at 37 ℃ for 30 minutes to digest nucleic acids. The protein concentration was measured using the Bradford Protein Assay (He, 2011).

The samples were separated on a 17% SDS-PAGE gel and histone bands were excised, chemically acylated with propionic anhydride and in-gel digested with trypsin, followed by peptide N-terminal derivatization with phenyl isocyanate (PIC) (Noberini et al., 2021). Chemical acylation of lysines (which occurs on unmodified or mono-methylated residues) impairs trypsin cleavage, resulting in proteolityc cleavage at arginine residues only. This treatment allows obtaining histone peptides of proper length for MS analysis. The samples were then desalted on handmade StageTips columns (Prudhomme et al., 2021). Peptide mixtures were separated by reversed-phase chromatography on an EASY-Spray column (Thermo Fisher Scientic), 25-cm long (inner diameter 75 μm, PepMap C18, 2 μm particles), which was connected online to a Q Exactive Plus instrument (Thermo Fisher Scientific) through an EASY-Spray™ Ion Source (Thermo Fisher Scientific), as described (Noberini et al., 2021).

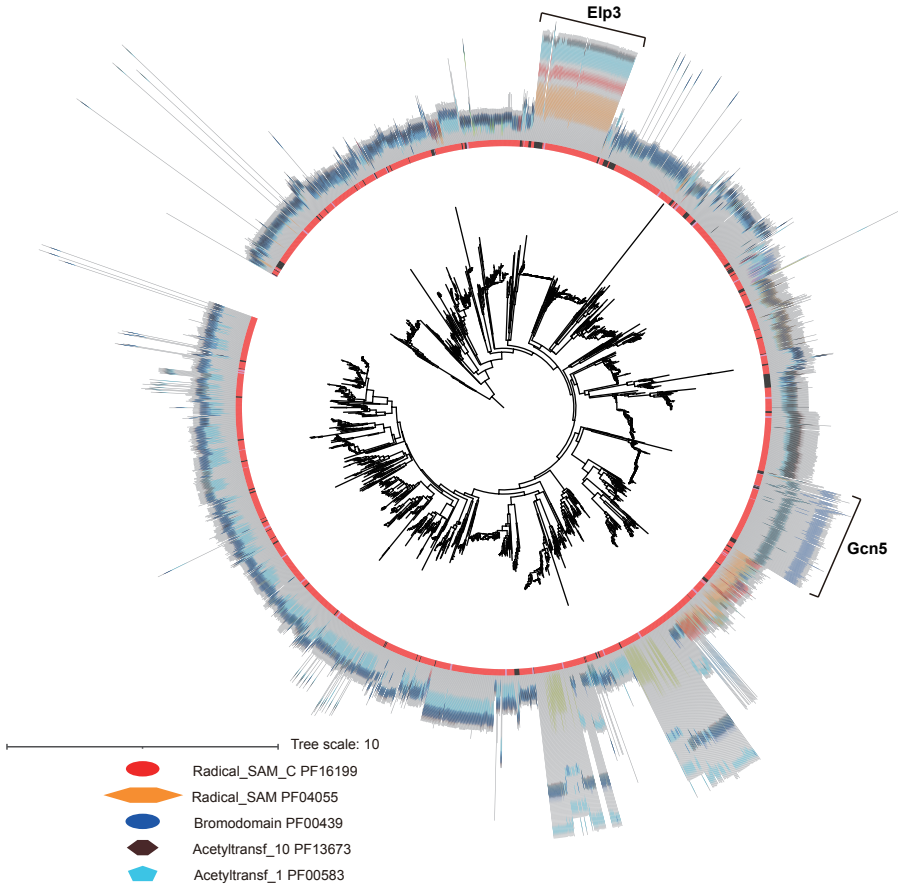
The acquired RAW data were analyzed using the integrated MaxQuant software v.1.6.10. The Uniprot UP000000560 database was used for identification of *A. nidulans* histone peptides. Enzyme specificity was set to Arg-C. The estimated false discovery rate (FDR) was set at a maximum of 1%. The mass tolerance was set to 6 ppm for precursor and fragment ions. Two missed cleavages were allowed, and the minimum peptide length was set to 4 amino acids. Min. score for modified peptides and min. delta score for modified peptides were set to 1. Variable modifications included lysine propionylation, monomethylation + propionylation, dimethylation, trimethylation and acetylation. N-terminal PIC labeling was set as a fixed modification (Noberini et al., 2021). To reduce the search time and the rate of false positives, with increasing the number of variable modifications included in the database search, the raw data were analyzed through multiple parallel MaxQuant jobs (Bremang et al., 2013), setting different combinations of variable modifications. Peptides identified by MaxQuant with Andromeda score higher than 50 and localization probability score higher than 0.75 were quantitated, either manually or by using a version of the EpiProfile 2.0 software (Z.-F. Yuan et al., 2018) adapted to the analysis of histones from *A. nidulans*. Identifications and retention

times were used to guide the manual quantification of each modified peptide using QualBrowser version 2.0.7 (Thermo Fisher Scientific). Site assignment was evaluated from MS2 spectra using QualBrowser and MaxQuant Viewer. Extracted ion chromatograms (XICs) were constructed for each doubly charged precursor, based on its *m/z* value, using a mass tolerance of 10 ppm. For each histone modified peptide, the % relative abundance (%RA) was estimated by dividing the area under the curve (AUC) of each modified peptide for the sum of the areas corresponding to all the observed forms of that peptide (Pesavento et al., 2006). The AUC values are reported (Table S3) and visualized using GraphPad Prism.

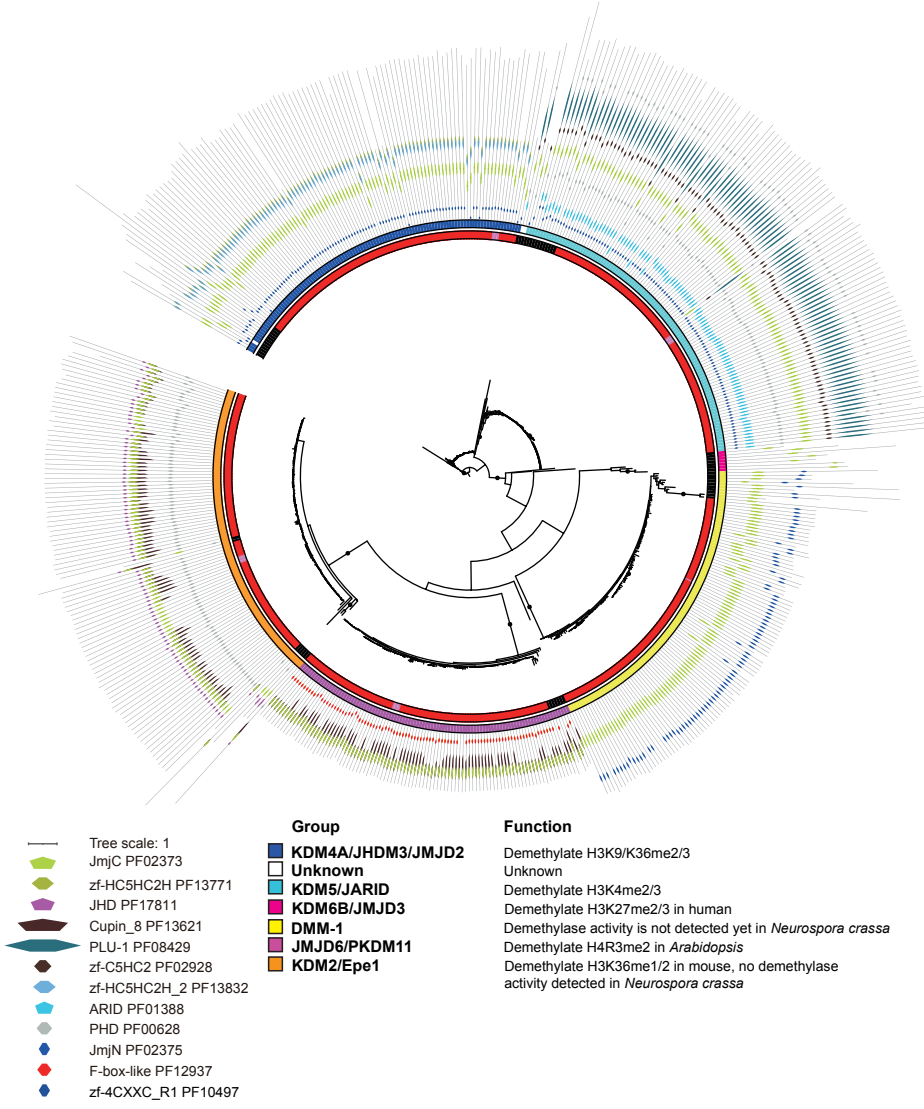
Supplementary materials



Supplementary Figure 1. A robust species phylogeny for the fungal genus *Aspergillus*. The species phylogeny was based on a maximum-likelihood phylogeny reconstructed by IQ-TREE with partitioned analysis using 758 BUSCO (Benchmarking Universal Single-Copy Orthologs) genes. The dot on the branch indicates ultrafast bootstrap values over 95 and SH-aLRT bootstraps values over 80. Species from the same section are labeled by background color with a section name. Two species, *F. oxysporum* and *A. kawachii*, are included in the species tree, but we don't use these two in the following analyses.



Supplementary Figure 2. Two known histone acyltransferase-domain containing protein groups are conserved in the *Aspergillus* genus and outgroup species together with many unknown proteins. Maximum likelihood phylogeny of Histone acyltransferase domains (PF00583) found in 94 Aspergilli and 15 outgroup species were determined using IQ-TREE (Chernomor et al. 2016; Nguyen et al. 2015). The black dot on the branch indicates ultrafast bootstrap values over 95 and SH-aLRT bootstraps values over 80. Additional conserved domains in histone acyltransferase domain-containing proteins were identified using the PFAM database (https://pfam.xfam.org).



Supplementary Figure 3. Five histone demethylase-domain containing protein groups are conserved in *Aspergillus* genus and outgroup species while KDM6B/JMJD3 is lost. Maximum likelihood phylogeny of Jumonji C (JmjC) domain (PF02373) found in 94 Aspergilli and 15 outgroup species were determined using IQ-TREE (Chernomor et al. 2016; Nguyen et al. 2015). The black dot on the branch indicates ultrafast bootstrap values over 95 and SH-aLRT bootstraps values over 80. Additional conserved domains in JmjC domain-containing proteins were identified using the PFAM database (https://pfam.xfam.org).

Supplementary Table 1. List of species included in this study.

Species name	Class	JGI abbreviation	Scaffolds	Genes	BUSCO completeness (%)
Aspergillus acristatulus	Eurotiomycetes	Aspacri1	311	11221	99.21
Aspergillus aculeatinus	Eurotiomycetes	Aspacu1	121	12027	99.47
Aspergillus aculeatus	Eurotiomycetes	Aspac1	660	10845	99.21
Aspergillus albertensis	Eurotiomycetes	Aspalbe1	473	12816	98.68
Aspergillus amylovorus	Eurotiomycetes	Aspamy1	462	12857	99.47
Aspergillus arachidicola	Eurotiomycetes	Aspara19utr	451	13895	98.94
Aspergillus avenaceus	Eurotiomycetes	Aspave1	1528	11293	99.08
Aspergillus bertholletius	Eurotiomycetes	Aspber1	443	12948	98.81
Aspergillus bisporus	Eurotiomycetes	Aspbisp1	244	9731	98.55
Aspergillus bombycis	Eurotiomycetes	Aspbom1	450	12265	97.76
Aspergillus brasiliensis	Eurotiomycetes	Aspbr1	105	13000	99.21
Aspergillus brevijanus	Eurotiomycetes	Aspbrev1	685	12832	99.34
Aspergillus brunneoviolaceus	Eurotiomycetes	Aspbru1	153	12075	98.68
Aspergillus caelatus	Eurotiomycetes	Aspcae1	729	13916	98.81
Aspergillus calidoustus	Eurotiomycetes	Aspcal1	78	15139	96.97
Aspergillus campestris	Eurotiomycetes	Aspcam1	62	9764	99.08
Aspergillus candidus	Eurotiomycetes	Aspcand1	268	9641	98.68
Aspergillus carbonarius	Eurotiomycetes	Aspca3	963	11624	91.29
Aspergillus carlsbadensis	Eurotiomycetes	Aspcar1	254	14289	99.08
Aspergillus chevalieri	Eurotiomycetes	Aspchev1	417	10253	98.94
Aspergillus clavatus	Eurotiomycetes	Aspcl1	143	9121	99.47
Aspergillus coremiiformis	Eurotiomycetes	Aspcor1	2728	9078	99.47
Aspergillus costaricaensis	Eurotiomycetes	Aspcos1	86	11966	99.21
Aspergillus cristatus	Eurotiomycetes	Aspcri1	68	10344	95.78
Aspergillus desertorum	Eurotiomycetes	Aspdese1	288	10111	98.94
Aspergillus egyptiacus	Eurotiomycetes	Aspegy1	682	9890	99.21
Aspergillus ellipticus	Eurotiomycetes	Aspell1	518	12884	98.55
Aspergillus eucalypticola	Eurotiomycetes	Aspeuc1	131	11934	98.55
Aspergillus fijiensis	Eurotiomycetes	Aspfij1	149	12019	99.47
Aspergillus filifera	Eurotiomycetes	Aspfili1	814	11686	99.08
Aspergillus fischeri	Eurotiomycetes	Neofi1	976	10406	99.08
Aspergillus flavus	Eurotiomycetes	Aspfl1	138	12604	94.72
Aspergillus fumigatus	Eurotiomycetes	Aspfu1	8	9781	98.28
Aspergillus funiculosus	Eurotiomycetes	Aspfuni1	144	9300	98.81
Aspergillus glaucus	Eurotiomycetes	Aspgl1	82	11277	99.08
Aspergillus heteromorphus	Eurotiomycetes	Asphet1	205	11133	99.47
Aspergillus homomorphus	Eurotiomycetes	Asphom1	152	11361	99.47

Species name	Class	JGI abbreviation	Scaffolds	Genes	BUSCO completeness (%)
Aspergillus ibericus	Eurotiomycetes	Aspibe1	116	11680	99.47
Aspergillus indicus	Eurotiomycetes	Aspind2_1	601	11361	98.94
Aspergillus indologenus	Eurotiomycetes	Aspind1	334	12163	99.21
Aspergillus insolitus	Eurotiomycetes	Aspinso1	157	8834	99.34
Aspergillus japonicus	Eurotiomycetes	Aspjap1	163	12024	98.68
Aspergillus luchuensis	Eurotiomycetes	Aspfo1	107	13530	99.34
Aspergillus minisclerotigenes	Eurotiomycetes	Aspmin1	296	13415	99.08
Aspergillus multicolor	Eurotiomycetes	Aspmul1_1	113	12821	99.21
Aspergillus neoauricomus	Eurotiomycetes	Aspneoa1	364	12221	98.55
Aspergillus neoechinulatus	Eurotiomycetes	Aspne1	236	10504	98.94
Aspergillus neoniger	Eurotiomycetes	Aspneo1	169	11939	99.21
Aspergillus nidulans	Eurotiomycetes	Aspnid1	8	10680	98.15
Aspergillus niger	Eurotiomycetes	Aspni_NRRL3 1	15	11846	99.6
Aspergillus nomius	Eurotiomycetes	Aspnom1	290	12897	98.81
Aspergillus novofumigatus	Eurotiomycetes	Aspnov1	62	11549	98.68
Aspergillus novoparasiticus	Eurotiomycetes	Aspnovo1	870	14182	98.68
Aspergillus ochraceoroseus	Eurotiomycetes	Aspoch1	34	8924	98.55
Aspergillus olivicola	Eurotiomycetes	Aspoli1_1	824	12183	98.02
Aspergillus oryzae	Eurotiomycetes	Aspor1	11	12030	93.67
Aspergillus parasiticus	Eurotiomycetes	Asppar1	270	13752	98.81
Aspergillus parvulus	Eurotiomycetes	Aspparv1	329	11927	99.08
Aspergillus penicilloides	Eurotiomycetes	Asppeni1	255	10386	98.81
Aspergillus piperis	Eurotiomycetes	Asppip1	47	12069	98.94
Aspergillus pseudocaelatus	Eurotiomycetes	Asppsec1	466	13895	97.49
Aspergillus pseudonomius	Eurotiomycetes	Asppsen1	374	13384	99.47
Aspergillus pseudotamarii	Eurotiomycetes	Asppset1	249	13428	98.02
Aspergillus rambellii	Eurotiomycetes	Aspram1	4177	7761	97.1
Aspergillus recurvatus	Eurotiomycetes	Asprec1_1	269	11007	98.94
Aspergillus restrictus	Eurotiomycetes	Aspres1	297	9143	99.47
Aspergillus robustus	Eurotiomycetes	Asprobu1	186	12182	99.21
Aspergillus roseoglobulosus	Eurotiomycetes	Aspros1	708	12489	98.68
Aspergillus ruber	Eurotiomycetes	Eurhe1	110	10076	98.94
Aspergillus saccharolyticus	Eurotiomycetes	Aspsac1	120	10066	99.21
Aspergillus sclerotiicarbonarius	Eurotiomycetes	Aspscle1	166	12571	99.08
Aspergillus sclerotioniger	Eurotiomycetes	Aspscl1	138	12338	99.47
Aspergillus sclerotiorum	Eurotiomycetes	Aspsc1	335	12620	99.34
Aspergillus sergii	Eurotiomycetes	Aspser1	262	13713	98.68
Aspergillus stella-maris	Eurotiomycetes	Aspstel1_1	266	12537	99.34
Aspergillus stercoraria	Eurotiomycetes	Aspstec1_1	278	10324	98.42

Species name	Class	JGI abbreviation	Scaffolds	Genes	BUSCO completeness (%)
Aspergillus steynii	Eurotiomycetes	Aspste1	37	13211	99.08
Aspergillus sydowii	Eurotiomycetes	Aspsy1	97	13620	99.21
Aspergillus taichungensis	Eurotiomycetes	Asptaic1	310	9692	98.55
Aspergillus tamarii	Eurotiomycetes	Asptam1	448	13331	98.55
Aspergillus terreus	Eurotiomycetes	Aspte1	26	10406	93.01
Aspergillus transmontanensis	Eurotiomycetes	Asptra1	293	14216	97.76
Aspergillus triticus	Eurotiomycetes	Asptr1	228	9769	99.6
Aspergillus tubingensis	Eurotiomycetes	Asptu1	33	12322	98.81
Aspergillus udagawae	Eurotiomycetes	Aspuda1	1029	9999	97.76
Aspergillus undulatus	Eurotiomycetes	Aspund1_1	190	11802	99.08
Aspergillus unguis	Eurotiomycetes	Aspung1	142	10397	99.21
Aspergillus uvarum	Eurotiomycetes	Aspuva1	172	12017	98.55
Aspergillus vadensis	Eurotiomycetes	Aspvad1	60	12132	98.81
Aspergillus venezuelensis	Eurotiomycetes	Aspven1	715	12450	98.55
Aspergillus versicolor	Eurotiomycetes	Aspve1	51	13228	99.08
Aspergillus violaceofuscus	Eurotiomycetes	Aspvio1	209	12082	99.34
Aspergillus welwitschiae	Eurotiomycetes	Aspwel1	396	13687	98.15
Aspergillus wentii	Eurotiomycetes	Aspwe1	27	12442	99.21
Penicillium brevicompactum	Eurotiomycetes	Penbr2	35	11536	98.15
Penicillium griseofulvum	Eurotiomycetes	Pengri1	92	9630	98.81
Penicillium rubens (chrysogenum)	Eurotiomycetes	Pench1	27	11396	99.47
Cladonia grayi	Lecanoromycetes	Clagr3	414	11389	94.59
Botrytis cinerea	Leotiomycetes	Botci1	588	16447	84.96
Arthrobotrys oligospora	Orbiliomycetes	Artol1	215	11479	97.76
Plectania melastoma	Pezizomycetes	Plemel1	175	16180	98.02
Saccharomyces cerevisiae	Saccharomycotina	Sacce1	16	6575	94.59
Fusarium graminearum	Sordariomycetes	Fusgr1	31	13322	98.15
Neurospora crassa	Sordariomycetes	Neucr2	20	10785	91.69
Schizosaccharomyces pombe	Taphrinomycotina	Schpo1	5	5134	93.54
Xylona heveae	Xylonomycetes	Xylhe1	27	8205	99.74
Zymoseptoria tritici	Dothideomycetes	Mycgr3	21	10952	97.89
Coprinopsis cinerea	Basidiomycota	Copci1	94	13393	96.04
Dendrothele bispora	Basidiomycota	Denbi1	3942	33645	91.82

Supplementary Table 2. List of BUSCO genes used to build the species tree.

BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene
427250at4751	494at4751	204341at4751	65334at4751	261237at4751	138833at4751	322267at4751	372510at4751
428342at4751	771at4751	204508at4751	65499at4751	262317at4751	141273at4751	323327at4751	372972at4751
428569at4751	1866at4751	205479at4751	66151at4751	263390at4751	142357at4751	324442at4751	374350at4751
428774at4751	3534at4751	207926at4751	67284at4751	264038at4751	142965at4751	325992at4751	374548at4751
428984at4751	4409at4751	208463at4751	67399at4751	264164at4751	143414at4751	326512at4751	376255at4751
429332at4751	4466at4751	209053at4751	67806at4751	265302at4751	145178at4751	326649at4751	376442at4751
429862at4751	4523at4751	209619at4751	68466at4751	265325at4751	145378at4751	327672at4751	376539at4751
430556at4751	5995at4751	210178at4751	68881at4751	266634at4751	145524at4751	327748at4751	376586at4751
431195at4751	6186at4751	210462at4751	71039at4751	267556at4751	147028at4751	327884at4751	377787at4751
431607at4751	6272at4751	210986at4751	71046at4751	269389at4751	148293at4751	329067at4751	378256at4751
432952at4751	6377at4751	211442at4751	73383at4751	269580at4751	150469at4751	330533at4751	378281at4751
433181at4751	6592at4751	211634at4751	73477at4751	269939at4751	153083at4751	330551at4751	378295at4751
433384at4751	6812at4751	211768at4751	73771at4751	271511at4751	154126at4751	330629at4751	378376at4751
434428at4751	7423at4751	212522at4751	75033at4751	271781at4751	156073at4751	330815at4751	378430at4751
434500at4751	7545at4751	212735at4751	75456at4751	271850at4751	160044at4751	330949at4751	378436at4751
434984at4751	8406at4751	213564at4751	75463at4751	272081at4751	160875at4751	331022at4751	379525at4751
435317at4751	8818at4751	215926at4751	75709at4751	272711at4751	160940at4751	331763at4751	382001at4751
435471at4751	10168at4751	216270at4751	75976at4751	273469at4751	160995at4751	332006at4751	383007at4751
435767at4751	10531at4751	216661at4751	76196at4751	274828at4751	161113at4751	332036at4751	384156at4751
436872at4751	11140at4751	216683at4751	76893at4751	276205at4751	163318at4751	332416at4751	384315at4751
437914at4751	11636at4751	216747at4751	76975at4751	278643at4751	165243at4751	332876at4751	385139at4751
438274at4751	11879at4751	216949at4751	78289at4751	279020at4751	167290at4751	332891at4751	385343at4751
438731at4751	12258at4751	217097at4751	78665at4751	280937at4751	167374at4751	334424at4751	385564at4751

BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene
438861at4751	12812at4751	217374at4751	79262at4751	281299at4751	167498at4751	335237at4751	385971at4751
439490at4751	12959at4751	219236at4751	79294at4751	281726at4751	168092at4751	335719at4751	386220at4751
440401at4751	13224at4751	219492at4751	79797at4751	281920at4751	168489at4751	336181at4751	386245at4751
441586at4751	13605at4751	221926at4751	79988at4751	282059at4751	169381at4751	336187at4751	387347at4751
443595at4751	14295at4751	224693at4751	81295at4751	282964at4751	170297at4751	337150at4751	387759at4751
444174at4751	14441at4751	224775at4751	81614at4751	283177at4751	170401at4751	337917at4751	388229at4751
444399at4751	14523at4751	224804at4751	81687at4751	284127at4751	170904at4751	338301at4751	389034at4751
444602at4751	14793at4751	225321at4751	84250at4751	284176at4751	171181at4751	338808at4751	389547at4751
444892at4751	14986at4751	225351at4751	84829at4751	284287at4751	171238at4751	339525at4751	389565at4751
445210at4751	17174at4751	225535at4751	85793at4751	284377at4751	172113at4751	342642at4751	390518at4751
445401at4751	18809at4751	225712at4751	86293at4751	285254at4751	172363at4751	343196at4751	391384at4751
445954at4751	20600at4751	225820at4751	87251at4751	285477at4751	172966at4751	343603at4751	392558at4751
446124at4751	20952at4751	225918at4751	87313at4751	285977at4751	173428at4751	343623at4751	393440at4751
447112at4751	21651at4751	226810at4751	90022at4751	289666at4751	173995at4751	343717at4751	393996at4751
447424at4751	21689at4751	227572at4751	91135at4751	289686at4751	174653at4751	344553at4751	396097at4751
449717at4751	22982at4751	228465at4751	92267at4751	290314at4751	174672at4751	345661at4751	396330at4751
449803at4751	22992at4751	228752at4751	93783at4751	290471at4751	175136at4751	348020at4751	396842at4751
451032at4751	23198at4751	229510at4751	94384at4751	291828at4751	175599at4751	348111at4751	397465at4751
451701at4751	23856at4751	229997at4751	94628at4751	292336at4751	176178at4751	349264at4751	398519at4751
451811at4751	24198at4751	230605at4751	95253at4751	292418at4751	176478at4751	350840at4751	398592at4751
451849at4751	25020at4751	231468at4751	96633at4751	293405at4751	176648at4751	351351at4751	398882at4751
452573at4751	25161at4751	231582at4751	97141at4751	295692at4751	176748at4751	351482at4751	398968at4751
452692at4751	26329at4751	232698at4751	98519at4751	295937at4751	179025at4751	351646at4751	399376at4751
452759at4751	27008at4751	233287at4751	98733at4751	296024at4751	179437at4751	351651at4751	399413at4751

BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene
452947at4751	29512at4751	233706at4751	98739at4751	296643at4751	179897at4751	351849at4751	399810at4751
453693at4751	31256at4751	234548at4751	100957at4751	297145at4751	181890at4751	351858at4751	401387at4751
453909at4751	31324at4751	234698at4751	101133at4751	297671at4751	181959at4751	351906at4751	402005at4751
454746at4751	32164at4751	235105at4751	103183at4751	298335at4751	183720at4751	352224at4751	402255at4751
454840at4751	33984at4751	235463at4751	103461at4751	298379at4751	183801at4751	352999at4751	403079at4751
455110at4751	34223at4751	235640at4751	103836at4751	299054at4751	183849at4751	353685at4751	404050at4751
456963at4751	34706at4751	236504at4751	104173at4751	299330at4751	184718at4751	353762at4751	405669at4751
457161at4751	34830at4751	237276at4751	104415at4751	299509at4751	184762at4751	353991at4751	405898at4751
457544at4751	36610at4751	237281at4751	105401at4751	299559at4751	185196at4751	354679at4751	406240at4751
457951at4751	37028at4751	237393at4751	106281at4751	299583at4751	185399at4751	354739at4751	406403at4751
458554at4751	37573at4751	237533at4751	107136at4751	299611at4751	186138at4751	354778at4751	406434at4751
458980at4751	37642at4751	238922at4751	107464at4751	300016at4751	186574at4751	354913at4751	406793at4751
459844at4751	38031at4751	240036at4751	109276at4751	300216at4751	186701at4751	355454at4751	407348at4751
461238at4751	38184at4751	240246at4751	111022at4751	302127at4751	187821at4751	355525at4751	408391at4751
461919at4751	39833at4751	241079at4751	111189at4751	304341at4751	188070at4751	355781at4751	408419at4751
464711at4751	43842at4751	241483at4751	112646at4751	304453at4751	189234at4751	356079at4751	409761at4751
464772at4751	46090at4751	241586at4751	112937at4751	304547at4751	189573at4751	356691at4751	410458at4751
464897at4751	46098at4751	241860at4751	115993at4751	304733at4751	190183at4751	356743at4751	410782at4751
465849at4751	46119at4751	243015at4751	117017at4751	305061at4751	190475at4751	357182at4751	410986at4751
467311at4751	46666at4751	244066at4751	118189at4751	306594at4751	190832at4751	357584at4751	410991at4751
467847at4751	46752at4751	245480at4751	118701at4751	307144at4751	191114at4751	358012at4751	411130at4751
467959at4751	46918at4751	245900at4751	119833at4751	308236at4751	191296at4751	358118at4751	411471at4751
468195at4751	47592at4751	246207at4751	122507at4751	308816at4751	191310at4751	359271at4751	411518at4751
468224at4751	48708at4751	246951at4751	122824at4751	310718at4751	192059at4751	359482at4751	411677at4751

BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene	BUSCO gene
468408at4751	49476at4751	247065at4751	123228at4751	310891at4751	192135at4751	359627at4751	411754at4751
371481at4751	420414at4751	370505at4751	419667at4751	368706at4751	419200at4751	366635at4751	418680at4751
371674at4751	420527at4751	370952at4751	419868at4751	369435at4751	419232at4751	367778at4751	419048at4751
366595at4751	418242at4751	364284at4751	417470at4751	363593at4751	416914at4751	362322at4751	414508at4751
366597at4751	418644at4751	365270at4751	417746at4751	364123at4751	417109at4751	362601at4751	416039at4751
488426at4751	64002at4751	372381at4751	421687at4751	482337at4751	56729at4751	247383at4751	123494at4751
488857at4751	64498at4751	469452at4751	49967at4751	483049at4751	57081at4751	248360at4751	125400at4751
490662at4751	490719at4751	470813at4751	50786at4751	483370at4751	58127at4751	251158at4751	125935at4751
260722at4751	138695at4751	471205at4751	50835at4751	483390at4751	59795at4751	251613at4751	126037at4751
260896at4751	138697at4751	473814at4751	52541at4751	484080at4751	61411at4751	252314at4751	126249at4751
492591at4751	493358at4751	474125at4751	52909at4751	484323at4751	61438at4751	252469at4751	126256at4751
321728at4751	423166at4751	474320at4751	53270at4751	484805at4751	61857at4751	254717at4751	126519at4751
322254at4751	423206at4751	474374at4751	54252at4751	485650at4751	62846at4751	255378at4751	127140at4751
494427at4751	497024at4751	477057at4751	54919at4751	485702at4751	63635at4751	255412at4751	129160at4751
371862at4751	421097at4751	477122at4751	56395at4751	488348at4751	63838at4751	256318at4751	129326at4751
257330at4751	130062at4751	481052at4751	56432at4751	247129at4751	123277at4751	256914at4751	129520at4751
258245at4751	130930at4751	312950at4751	196094at4751	320638at4751	203710at4751	311254at4751	195619at4751
258762at4751	130976at4751	315551at4751	196841at4751	320777at4751	204087at4751	312080at4751	195863at4751
259060at4751	133954at4751	315639at4751	197111at4751	320782at4751	422151at4751	312209at4751	196015at4751
259444at4751	135184at4751	315799at4751	197612at4751	321416at4751	422622at4751	317335at4751	202089at4751
259700at4751	135652at4751	315802at4751	198196at4751	321585at4751	422764at4751	319121at4751	202248at4751
259883at4751	137963at4751	316254at4751	198234at4751	360800at4751	411919at4751	320163at4751	203109at4751
260251at4751	138484at4751	316920at4751	200331at4751	361877at4751	412006at4751	362250at4751	414474at4751
311048at4751	192452at4751	317276at4751	201591at4751	362218at4751	413850at4751		

Supplementary Table 3. Mass spectrometry result for *A. nidulans*

Peptide	AUC (Area under the curve)	%RA (%relative abundance)
H3_3_8 unmod	7.81E+09	96.4
H3_3_8 K4me1	/	/
H3_3_8 K4me2	9.34E+07	1.2
H3_3_8 K4me3	2.01E+08	2.5
H3_9_17 unmod	2.61E+08	61.9
H3_9_17 K9me1	1.97E+07	4.7
H3_9_17 K9me2	4.66E+05	0.1
H3_9_17 K9me3	9.10E+06	2.2
H3_9_17 K9ac	4.00E+07	9.5
H3_9_17 K14ac	6.64E+07	15.7
H3_9_17 K9me1K14ac	/	/
H3_9_17 K9me2K14ac	2.87E+06	0.7
H3_9_17 K9me3K14ac	1.61E+06	0.4
H3_9_17 K9acK14ac	2.06E+07	4.9
H3_18_26 unmod	8.02E+08	53.0
H3_18_26 K18ac H3_18_26 K23ac	5.58E+08	36.8
H3_18_26 K18acK23ac	1.54E+08	10.2
H3_27_40 unmod	1.00E+09	67.1
H3_27_40 K36me1	6.16E+07	4.1
H3_27_40 K27ac	4.30E+08	28.8
H3_73_83 unmod	7.88E+07	66.9
H3_73_83 K79me1	3.76E+06	3.2
H3_73_83 K79me2	3.53E+07	30.0
H4_4_17 unmod	2.02E+08	12.7
H4_4_17 K5(12)ac	9.09E+06	0.6
H4_4_17 K8ac	3.12E+07	2.0
H4_4_17 K16ac	8.91E+08	56.0
H4_4_17 di ac	3.12E+08	19.6
H4_4_17 tri ac	1.02E+08	6.4
H4_4_17 tetra ac	4.48E+07	2.8

H4_4_17 di-ac: K12ac/K16ac; K5ac/K12ac; K5ac/K16ac; K5ac/K8ac; K5ac/K8ac; K8ac/K16ac
H4_4_17 tri-ac: K5acK12acK16ac; K5acK8ac12ac; K5acK8acK16ac; K8acK12acK16ac

Supplementary Table 4. List of histone modifier complexes reported in the literature.

Complex name	Subunit name	Conservation *	Catalytic subunit	GenBank accession number	Domain	Organism	References
SET1/COMPASS (COMplex of Proteins ASsociated with SET1)	SET1	P|P	Yes	856519 (gene); DAA06813.1 (protein)	PF11764 N-SET PF00856 SET PF11767 SET_assoc	*Saccharomyces cerevisiae*	Kan *et al*., 2008; Palmer *et al*., 2013; Studt *et al*., 2017
	CclA	P|P	No	850702 (gene); DAA09333.1 (protein)	(SMART) SM00449 SPRY	*S. cerevisiae*	
	Swd1/Cps50	P|P	No	851265 (gene); DAA06988.1 (protein)	PF00400 WD40	*S. cerevisiae*	
	Spp1p/Cps40	A|PAP	No	NM_001183952 (gene); NP_015187 (protein)	PF00628 PHD	*S. cerevisiae*	
	Swd2/Cps35	P|P	No	NM_001179584 (gene), NP_012907 (protein)	PF00400 WD40	*S. cerevisiae*	
	Swd3p/Cps30/WDR5	P|P	No	NM_001178523 (gene); NP_009734 (protein)	PF00400 WD40	*S. cerevisiae*	
	Sdc1/Cps25/Dpy30	PAP|PAP	No	NM_001180777 (gene); NP_010757 (protein)	PF05186 Dpy-30	*S. cerevisiae*	
	Shg1p	*S. cerevisiae*	No	NM_001178606 (gene), NP_009817 (protein)	-	*S. cerevisiae*	
DCDC (Dim-5/-7/-9/CUL4/DDB1 Complex)	Dim-5	PAP|PAP	Yes	3873656 (gene); Q8X225 (protein)	PF05033 Pre-SET PF00856 SET	*Neurospora crassa*	Lewis *et al*., 2010; Horn *et al*., 2005
	Dim-7	A|PAP	No	3877415 (gene); EAA32072 (protein)	-	*N. crassa*	
	Dim-9	A|PAP	No	3872416 (gene); XP_956278 (protein)	(SMART) SM00320 WD40	*N. crassa*	
	CUL4	PAP|PAP	No	3873913 (gene); XP_957743 (protein)	PF00888 Cullin PF10557 Cullin_Nedd8	*N. crassa*	
	Dim-8/DDB1	P|PAP	No	3878495 (gene); XP_962347 (protein)	PF03178 CPSF_A	*N. crassa*	
SET2 Complex	Set2	PAP|P	Yes	853271 (gene); P46995.2 (protein)	PF17907 AWS PF00856 SET PF08236 SRI PF18507 WW_1	*S. cerevisiae*	Xiao *et al*., 2003
	Rpb1-CTD	P|P	No	851415 (gene); PDB: 6I84_A (protein)	PF04997 RNA_pol_Rpb1_1 PF00623 RNA_pol_Rpb1_2 PF04983 RNA_pol_Rpb1_3 PF05000 RNA_pol_Rpb1_4 PF04998 RNA_pol_Rpb1_5 PF04992 RNA_pol_Rpb1_6 PF04990 RNA_pol_Rpb1_7 PF05001 RNA_pol_Rpb1_R	*S. cerevisiae*	

Complex name	Subunit name	Conservation *	Catalytic subunit	GenBank accession number	Domain	Organism	References
SET2 Complex	Rpb2	P|P	No	Z75059.1 (gene); CAA99357.1 (protein)	PF04563 RNA_pol_Rpb2_1 PF04561 RNA_pol_Rpb2_2 PF04565 RNA_pol_Rpb2_3 PF04566 RNA_pol_Rpb2_4 PF04567 RNA_pol_Rpb2_5 PF00562 RNA_pol_Rpb2_6 PF04560 RNA_pol_Rpb2_7	*S. cerevisiae*	Xiao et al., 2003
PRC2 (Polycomb Repressive Complex 2)	KMT6/SET7	A|PAP	Yes	3881192 (gene); XP_965043 (protein)	PF11616 EZH2_WD-Binding (Homo sapiens) PF18118 PRC2_HTH_1 (Homo sapiens) PF18264 preSET_CXC (Homo sapiens) PF00856 SET (Homo sapiens)	*N. crassa (SET - 7)*	Jamieson *et al.*, 2005
	EED	A|PAP	No	3878219 (gene); XP_962071 (protein)	PF00400 WD40 (Homo sapiens)	*N. crassa*	
PRC2 (Polycomb Repressive Complex 2)	SUZ12	A|PAP	No	3879610 (gene); XP_963451 (protein)	PF09733 VEFS-Box (Homo saplens)	*N. crassa*	Jamieson et al., 2005
Dot1 (Disruptor Of Telomeric silencing 1)	Dot1	P|PAP	Yes	NW_002477245 REGION: 1106692..1108271 (gene); XP_002381134.1(protein)	PF08123 DOT1	*Aspergillus flavus*	Liang *et al.*, 2017; Lacoste *et al.*, 2002; Sawada *et al.*, 2004
HAT-A2 complex	Gcn5	P|P	Yes	853167 (gene); DAA08343.1 (protein)	PF00583 Acetyltransf_1 PF00439 Bromodomain	*S. cerevisiae*	Sendra *et al.*, 2000
	Ada2	P|P	No	852059 (gene); NP_010736 (protein)	PF00249 Myb_DNA-binding PF00569 ZZ	*S. cerevisiae*	
	Ada3	P|P	No	851756 (gene); NP_010461 (protein)	PF10198 Ada3	*S. cerevisiae*	
SAGA complex (Spt-Ada-Gcn5 Acetyltransferase)	Gcn5	P|P	Yes	853167 (gene); DAA08343.1 (protein)	PF00583 Acetyltransf_1PF00439 Bromodomain	*S. cerevisiae*	Liu *et al.*, 2019

Complex name	Subunit name	Conservation *	Catalytic subunit	GenBank accession number	Domain	Organism	References
SAGA complex (Spt-Ada-Gcn5 Acetyltransferase)	Ada2	P|P	No	852059 (gene); NP_010736 (protein)	PF00249 Myb_DNA-bindingPF00569 ZZ	*S. cerevisiae*	Liu *et al.*, 2019
	Ada3	P|P	No	851756 (gene); NP_010461 (protein)	PF10198 Ada3	*S. cerevisiae*	
	Sgf29	PAP|P	No	850347 (gene); DAA07471 (protein)	PF07039 DUF1325	*S. cerevisiae*	
	Ubp8	P|P	No	855263 (gene); PDB: 6T9L_K (protein)	PF00443 UCHPF02148 zf-UBP	*S. cerevisiae*	
	Sgf11	*S. cerevisiae*	No	856060 (gene), pdb|6T9L (protein)	PF08209 Sgf11PF18519 Sgf11_N	*S. cerevisiae*	
	Sgf73	P|PAP	No	852814 (gene); pdb|6T9L (protein)	PF08313 SCA7, PF18508 zf_C2H2_13PF18508 zf_C2H2_13	*S. cerevisiae*	
	Sus1	A|PAP	No	1466445 (gene), pdb|6T9L (protein)	PF10163 EnY2 PF00400 WD40	*S. cerevisiae*	
	Taf5	P|P	No	852497 (gene), pdb|6T9K (protein)	PF04494 TFIID_NTD2 PF08513 LisH	*S. cerevisiae*	
	Taf6	P|P	No	852766 (gene), PDB: 6T9K_E (protein)	PF02969 TAFPF07571 TAF6_C	*S. cerevisiae*	
	Taf9	P|P	No	855276 (gene); PDB: 6T9K_F (protein)	PF02291 TFIID-31kDa	*S. cerevisiae*	
	Taf12	P|P	No	851723(gene); PDB: 6T9K_I (protein)	PF03847 TFIID_20kDa	*S. cerevisiae*	
	Tra1	P|P	No	856499 (gene); PDB: 6T9J_T (protein)	PF02259 FATPF00454 PI3_PI4_kinase	*S. cerevisiae*	
	Ada1	P|PAP	No	855821 (gene); PDB: 6T9K_H (protein)	PF12767 SAGA-Tad1	*S. cerevisiae*	
	Spt7	P|P	No	852373 (gene); PDB: 6T9K_K (protein)	PF00439 Bromodomain	*S. cerevisiae*	
	Spt8	P|PAP	No	850744 (gene); DAA09373.1 (protein)	(SMART) SM00320 WD40	*S. cerevisiae*	
	Spt20	P|PAP	No	854017 (gene); PDB: 6T9K_B (protein)	PF12090 Spt20	*S. cerevisiae*	
ADA complex	Gcn5	P|P	Yes	853167 (gene); DAA08343.1 (protein)	PF00583 Acetyltransf_1 PF00439 Bromodomain	*S. cerevisiae*	Eberharter *et al.*, 1999
	Ada2	P|P	No	852059 (gene); NP_010736 (protein)	PF00249 Myb_DNA-binding PF00569 ZZ	*S. cerevisiae*	
	Ada3	P|P	No	851756 (gene); NP_010461 (protein)	PF10198 Ada3	*S. cerevisiae*	
	Ahc1	*S. cerevisiae*	No	854188 (gene); DAA10806.1 (protein)	-	*S. cerevisiae*	

Complex name	Subunit name	Conservation *	Catalytic subunit	GenBank accession number	Domain	Organism	References
NuA3 (Nucleosomal Acetyltransferase of histone H3)	Sas3	P|PAP	Yes	852228 (gene); DAA07067.1 (protein)	PF01853 MOZ_SAS PF17772 zf-MYST	*S. cerevisiae*	John *et al*., 2000; Wang *et al*., 2002
	Pdp3	PAP|PAP	No	851177 (gene); NP_013560.3 (protein)	PF00855 PWWP	*S. cerevisiae*	
	Yng1	*S. cerevisiae*	No	854230 (gene); DAA10843.1 (protein)	(SMART) SM00249 PHD	*S. cerevisiae*	
	Nto1	PAP|P	No	856143 (gene); DAA11457.1 (protein)	PF10513 EPL1	*S. cerevisiae*	
	Taf14	PAP|PAP	No	855974 (gene); DAA11304.1 (protein)	PF17035 BET PF03366 YEATS	*S. cerevisiae*	
	Eaf6	P|P	No	853544 (gene); PDB: 5J9W_J (protein)	PF09340 NuA4	*S. cerevisiae*	
NuA4 (Nucleosomal Acetyltransferase of histone H4)	Esa1	P|P	Yes	854418 (gene); NP_014887.3 (protein)	PF01853 MOZ_SAS PF11717 Tudor-knot PF17772 zf-MYST	*S. cerevisiae*	Doyon *et al*., 2004; Kim *et al*., 2018
	Tra1	P|P	No	NC_001140 (REGION: 302761..313995)	PF02259 FAT PF00454 PI3_PI4_kinase	*S. cerevisiae*	
	Eaf7	PAP|PAP	No	855586 (gene); DAA10412.1 (protein)	PF07904 Eaf7	*S. cerevisiae*	
	Yng2	P|P	No	856490 (gene); DAA06786.1 (protein)	PF12998 ING	*S. cerevisiae*	
	Eaf5	*S. cerevisiae*	No	856696 (gene); PDB: 5Y81_H (protein)	-	*S. cerevisiae*	
	Act1	P|P	No	850504 (gene); PDB: 5Y81_G (protein)	PF00022 Actin	*S. cerevisiae*	
	Epl1	P|P	No	850520 (gene); DAA12416.1 (protein)	PF10513 EPL1	*S. cerevisiae*	
	Eaf6	P|P	No	NC_001142 (REGION: complement(581914..582255))	PF09340 NuA4	*S. cerevisiae*	
	Eaf1	PAP|PAP	No	851962 (gene); NP_010646.4 (protein)	PF07529 HSA	*S. cerevisiae*	
	Eaf2/Swc4	PAP|PAP	No	852885 (gene); DAA08100.1 (protein)	PF16282 SANT_DAMP1_like	*S. cerevisiae*	
	Arp4	P|P	No	853364 (gene); PDB: 5NBN_B (protein)	PF00022 Actin	*S. cerevisiae*	
	Yaf9	P|P	No	855616 (gene); NP_014292.3 (protein)	PF03366 YEATS	*S. cerevisiae*	
	Eaf3	P|P	No	856134 (gene); DAA11449.1 (protein)	PF05712 MRG PF11717 Tudor-knot	*S. cerevisiae*	
SAS (Something About Silencing) complex	Sas2	PAP|PAP	Yes	855157 (gene); DAA10024.1 (protein)	PF01853 MOZ_SAS PF17772 zf-MYST	*S. cerevisiae*	Meijsing *et al*., 2001; Kim *et al*., 2018
	Sas4	*S. cerevisiae*	No	851762 (gene); DAA12024.1 (protein)	PF15460 SAS4	*S. cerevisiae*	
	Sas5	*S. cerevisiae*	No	854388 (gene); DAA10985.1 (protein)	PF03366 YEATS	*S. cerevisiae*	

Complex name	Subunit name	Conservation *	Catalytic subunit	GenBank accession number	Domain	Organism	References
NuB4 (Nuclear Hat1p-containing type B histone acetyltransferase) complex	Hat1	P|P	Yes	856106 (gene); DAA11427.1 (protein)	PF10394 HAT1_N	*S. cerevisiae*	Tscherner *et al.*, 2012; Kim et al., 2018
	Hat2	P|P	No	856654 (gene); NP_010858.3 (protein)	PF12265 CAF1C_H4-bd PF00400 WD40	*S. cerevisiae*	
Elongator	Elp3	P|P	Yes	856019 (gene); PDB: 6QK7_C (protein)	PF00583 Acetyltransf_1 PF04055 Radical_SAM PF16199 Radical_SAM_C	*S. cerevisiae*	Winkler *et al.*, 2002; Kim et al., 2018
	Elp1	P|P	No	851100 (gene); PDB: 6QK7_D (protein)	PF04762 IKI3	*S. cerevisiae*	
	Elp6	*S. cerevisiae*	No	855360 (gene); DAA10213.1 (protein)	-	*S. cerevisiae*	
	Elp5	PAP|PAP	No	856594 (gene); NP_012057.3 (protein)	PF10483 Elong_Iki1	*S. cerevisiae*	
	Elp4	P|PAP	No	856002 (gene); DAA11331.1 (protein)	PF05625 PAXNEB	*S. cerevisiae*	
	Elp2	P|P	No	853114 (gene); PDB: 6QK7_B (protein)	PF00400 WD40	*S. cerevisiae*	
Sin3-Rpd3	Rpd3	P|P	Yes	855386 (gene); DAA10233 (protein)	PF00850 Hist_deacetyl	*S. cerevisiae*	Keogh *et al.*, 2005
	Sin3	P|P	No	854158 (gene); DAA10779.1 (protein)	PF02671 PAH PF08295 Sin3_corepress PF16879 Sin3a_C	*S. cerevisiae*	
	Ume1	*S. cerevisiae*	No	855964 (gene); DAA11295.1 (protein)	(SMART) SM00320 WD40	*S. cerevisiae*	
	Rco1	P|P	No	855097 (gene); NP_014069.1 (protein)	PF00628 PHD	*S. cerevisiae*	
	Eaf3	P|P	No	856134 (gene); DAA11449.1 (protein)	PF05712 MRG PF11717 Tudor-knot	*S. cerevisiae*	
HCHC	HDA-1 (HdaA)	P|P	Yes	3873096 (gene); XP_956974.3 (protein)	PF00850 Hist_deacetyl PF09757 Arb2	*N. crassa*	Honda *et al.*, 2016
	CHAP	P|P	No	3872623 (gene); XP_956474.1 (protein)	PF04729 ASF1_hist_chap	*N. crassa*	
	HP1	P|PAP	No	3873819 (gene); XP_957632.1 (protein)	PF01287 eIF-5a	*N. crassa*	
	CDP-2	A|PAP	No	3880541 (gene); XP_011392824 (protein)	PF00385 Chromo	*N. crassa*	

* P: presence; A: absence; PAP: presence absence pattern; *S. crevisiae*: specific to this species; patterns observed in Aspergilli and in outgroup species are separated by the pipe.

Supplementary Table 5 List of sequence names in the domain trees deposited at Zenodo (10.5281/zenodo.10640677)

Supplementary Table 6. Identification of missing chromatin modifiers in the outgroup species using TBLASTN search

Chromatin modifier	Species	Acession number	Range 1	E value	Range 2	E value	Range 3
Dim-2	*Botrytis cinerea*	CP009819.1	187312 to 189426	8.0E-144	189613 to 190446	3.0E-47	
	Zymoseptoria tritici	CP062683.1	2654869 to 2657733	6.0E-117	7141 to 9096	4.0E-71	1179058 to 1180491
Ash1	*Fusarium graminearum*	CP079829.1	5862765 to 5863967	3.0E-119	4983833 to 4984210	8.0E-07	
	Botrytis cinerea	CP009805.1	849818 to 850345	3.0E-108	849364 to 849759	3.0E-108	849054 to 849320
Dim-9	*Zymoseptoria tritici*	CP051550.1	1715210 to 1719298	0.0E+00			
	Xylona heveae	XM_018330026.1	61 to 3966	0.0E+00			
EED	*Botrytis cinerea*	CP009817.1	2192711 to 2193721	1.0E-72	2193754 to 2194146	1.0E-72	
Ada1	*Zymoseptoria tritici*	CP062623.1	1016924 to 1017718	2.0E-07			
Sin3	*Botrytis cinerea*	CP009809.1	1720472 to 1721308	9.0E-96	1719151 to 1720413	9.0E-96	1718927 to 1719073
Rco1	*Zymoseptoria tritici*	LT882681.1	*200994 to 201338*	3E-22	*200420 to 200878*	3E-22	

Chromatin modifier	Species	Acession number	E value	Range 4	E value	Range 5	E value
Dim-2	*Botrytis cinerea*	CP009819.1					
	Zymoseptoria tritici	CP062683.1	5.0E-49	9337 to 10026	7.0E-16	1181176 to 1182099	2.0E-13
Ash1	*Fusarium graminearum*	CP079829.1					
	Botrytis cinerea	CP009805.1	3.0E-108				
Dim-9	*Zymoseptoria tritici*	CP051550.1					
	Xylona heveae	XM_018330026.1					
EED	*Botrytis cinerea*	CP009817.1					
Ada1	*Zymoseptoria tritici*	CP062623.1					
Sin3	*Botrytis cinerea*	CP009809.1	9e-96	1722446 to 1722799	1e-28	1721768 to 1722055	1.0E-20
Rco1	*Zymoseptoria tritici*	LT882681.1					

Data Summary

The predicted proteomes used in this study are publicly available at the JGI (Joint Genome Institute) MycoCosm repository (Grigoriev et al., 2014); the species names and abbreviations are listed in Table S1. To evaluate the completeness of the predicted proteomes and to obtain a species phylogeny, 758 fungal BUSCO (Benchmarking Universal Single-Copy Ortholog) genes were used, and their names are listed in Table S2. The fasta, trimmed alignment, and maximum-likelihood phylogenetic tree files can be found in Supplementary Data 1 and 2 deposited at Zenodo (doi:10.5281/zenodo.5907090). The mass spectrometry results mentioned in Table S3 are deposited in the PRIDE database with the dataset identifier PXD033478.

Acknowledgment

The sequence data were produced by the US Department of Energy Joint Genome Institute (http://www.jgi.doe.gov/) in collaboration with the user community. Especially, we thank the Scott Baker, Mikael Andersen, and Joseph Spatafora for their permission to use the sequence data. We thank Jos Houbraken for validating the *Aspergillus* species tree and Joseph Strauss for providing *A. nidulans* strain. We thank Alessandro Vai for support with analysis of mass spectrometry data for histone PTMs' identification on *Aspergillus* core histone sequences.

Funding information

Xin Zhang is funded by the Chinese Scholarship Council (CSC) (201907720028). This work has been supported by EPIC-XS, project number 823839, funded by the Horizon 2020 programme of the European Union.

Authors and contribution

X.Z. formal analysis, investigation, methodology, visualization, writing original draft; R.N. formal analysis, investigation, methodology, visualization; T.B. funding acquisition, supervision, resources; J.C. conceptualization, funding acquisition, methodology, project administrations, supervision, visualization, writing original draft; M.F.S. conceptualization, funding acquisition, methodology, project administrations, supervision, visualization, writing original draft.

References

Agger, K., Cloos, P. A. C., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A. E. & Helin, K. (2007). UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature*, *449*(7163), 731–734.

Allard, S., Utley, R. T., Savard, J., Clarke, A., Grant, P., Brandl, C. J., Pillus, L., Workman, J. L. & Côté, J. (1999). NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p. *EMBO Journal*, *18*(18), 5108–5119.

Anisimova, M., Gil, M., Dufayard, J.-F., Dessimoz, C. & Gascuel, O. (2011). Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Systematic Biology*, *60*(5), 685–699.

Bachleitner, S., Sørensen, J. L., Gacek-Matthews, A., Sulyok, M., Studt, L. & Strauss, J. (2019). Evidence of a demethylase-independent role for the H3K4-specific histone demethylases in *Aspergillus nidulans* and *Fusarium graminearum* secondary metabolism. *Frontiers in Microbiology*, *10*.

Balasubramanian, R., Pray-Grant, M. G., Selleck, W., Grant, P. A. & Tan, S. (2002). Role of the Ada2 and Ada3 transcriptional coactivators in histone acetylation. *Journal of Biological Chemistry*, *277*(10), 7989–7995.

Berger, S. L. (2007). The complex language of chromatin regulation during transcription. *Nature*, *447*(7143), 407–412.

Bernstein, B. E., Tong, J. K. & Schreiber, S. L. (2000). Genomewide studies of histone deacetylase function in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(25), 13708–13713.

Bestor, T. (1988). *Structure of mammalian DNA methyltransferase as deduced from the inferred amino acid sequence and direct studies of the protein*. Portland Press Ltd.

Bewick, A. J., Hofmeister, B. T., Powers, R. A., Mondo, S. J., Grigoriev, I. V., James, T. Y., Stajich, J. E. & Schmitz, R. J. (2019). Diversity of cytosine methylation across the fungal tree of life. *Nature Ecology and Evolution*, *3*(3), 479–490.

Bicocca, V. T., Ormsby, T., Adhvaryu, K. K., Honda, S. & Selker, E. U. (2018). ASH1-catalyzed H3K36 methylation drives gene repression and marks H3K27me2/3-competent chromatin. *ELife*, *7*, 1–19.

Binda, O. (2013). On your histone mark, SET, methylate! *Epigenetics*, *8*(5), 457–463.

Booker, S. J. & Grove, T. L. (2010). Mechanistic and functional versatility of radical SAM enzymes. *F1000 Biology Reports*, *2*(1).

Boratyn, G. M., Camacho, C., Cooper, P. S., Coulouris, G., Fong, A., Ma, N., Madden, T. L., Matten, W. T., McGinnis, S. D. & Merezhuk, Y. (2013). BLAST: a more efficient report with usability improvements. *Nucleic Acids Research*, *41*(W1), W29–W33.

Bremang, M., Cuomo, A., Agresta, A. M., Stugiewicz, M., Spadotto, V. & Bonaldi, T. (2013). Mass spectrometry-based identification and characterisation of lysine and arginine methylation in the human proteome. *Molecular BioSystems*, *9*(9), 2231–2247.

Brosch, G., Loidl, P. & Graessle, S. (2008). Histone modifications and chromatin dynamics: A focus on filamentous fungi. *FEMS Microbiology Reviews*, *32*(3), 409–439.

Buscaino, A. (2019). Chromatin-mediated regulation of genome plasticity in human fungal pathogens. *Genes*, *10*(11).

Cánovas, D., Marcos, A. T., Gacek, A., Ramos, M. S., Gutiérrez, G., Reyes-Domínguez, Y. & Strauss, J. (2014). The histone acetyltransferase GcnE (GCN5) plays a central role in the regulation of *Aspergillus* asexual development. *Genetics*, *197*(4), 1175–1189.

Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, *25*(15), 1972–1973.

Capuano, F., Mülleder, M., Kok, R., Blom, H. J. & Ralser, M. (2014). Cytosine DNA methylation is found in *Drosophila melanogaster* but absent in *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and other yeast species. *Analytical Chemistry*, *86*(8), 3697–3702.

Carmen, A. A., Griffin, P. R., Calaycay, J. R., Rundlett, S. E., Suka, Y. & Grunstein, M. (1999). Yeast HOS3 forms a novel trichostatin A-insensitive homodimer with intrinsic histone deacetylase activity. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(22), 12356–12361.

Carrozza, M. J., Utley, R. T., Workman, J. L. & Côté, J. (2003). The diverse functions of histone acetyltransferase complexes. *Trends in Genetics*, *19*(6), 321–329.

Catania, S., Dumesic, P. A., Pimentel, H., Nasif, A., Stoddard, C. I., Burke, J. E., Diedrich, J. K., Cook, S., Shea, T., Geinger, E., Lintner, R., Yates, J. R., Hajkova, P., Narlikar, G. J., Cuomo, C. A., Pritchard, J. K. & Madhani, H. D. (2020). Evolutionary persistence of DNA methylation for millions of years after ancient loss of a *de novo* methyltransferase. *Cell*, *180*(2), 263-277.e20.

Chen, A. J., Frisvad, J. C., Sun, B. D., Varga, J., Kocsubé, S., Dijksterhuis, J., Kim, D. H., Hong, S. B., Houbraken, J. & Samson, R. A. (2016). *Aspergillus* section *Nidulantes* (formerly *Emericella*): Polyphasic taxonomy, chemistry and biology. *Studies in Mycology*, *84*, 1–118.

Chernomor, O., Von Haeseler, A. & Minh, B. Q. (2016). Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biology*, *65*(6), 997–1008.

Choi, J. & Kim, S.-H. (2017). A genome tree of life for the fungi kingdom. *Proceedings of the National Academy of Sciences*, *114*(35), 9391–9396.

Collemare, J. & Seidl, M. F. (2019). Chromatin-dependent regulation of secondary metabolite biosynthesis in fungi: is the picture complete? *FEMS Microbiology Reviews*, *43*(6), 591–607.

De Nadal, E., Zapater, M., Alepuz, P. M., Sumoy, L., Mas, G. & Posas, F. (2004). The MAPK Hog1 recruits Rpd3 histone deacetylase to activate *osmoresponsive* genes. *Nature*, *427*(6972), 370–374.

Dehé, P. M., Dichtl, B., Schaft, D., Roguev, A., Pamblanco, M., Lebrun, R., Rodríguez-Gil, A., Mkandawire, M., Landsberg, K., Shevchenko, A., Shevchenko, A., Rosaleny, L. E., Tordera, V., Chávez, S., Stewart, A. F. & Géli, V. (2006). Protein interactions within the Set1 complex and their roles in the regulation of histone 3 lysine 4 methylation. *Journal of Biological Chemistry*, *281*(46), 35404–35412.

Delmas, V., Stokes, D. G. & Perry, R. P. (1993). A mammalian DNA-binding protein that contains a chromodomain and an SNF2/SWI2-like helicase domain. *Proceedings of the National Academy of Sciences*, *90*(6), 2414–2418.

Doyon, Y. & Côté, J. (2004). The highly conserved and multifunctional NuA4 HAT complex. *Current Opinion in Genetics and Development*, *14*(2), 147–154.

Eberharter, A., Sterner, D. E., Schieltz, D., Hassan, A., Yates, J. R., Berger, S. L. & Workman, J. L. (1999). The ADA complex is a distinct histone acetyltransferase complex in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, *19*(10), 6621–6631.

Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Computational Biology*, *7*(10), e1002195.

Ekwall, K. (2005). Genome-wide analysis of HDAC function. *Trends in Genetics*, *21*(11), 608–615.

Erlendson, A. A., Friedman, S. & Freitag, M. (2017). A matter of scale and dimensions: chromatin of chromosome landmarks in the fungi. *Microbiology Spectrum*, *5*(4), 4–5.

Freitag, M. (2017). Histone methylation by SET domain proteins in fungi. *Annual Review of Microbiology*, *71*, 413–439.

Gacek-Matthews, A., Noble, L. M., Gruber, C., Berger, H., Sulyok, M., Marcos, A. T., Strauss, J. & Andrianopoulos, A. (2015). KdmA, a histone H3 demethylase with bipartite function, differentially regulates primary and secondary metabolism in *Aspergillus nidulans*. *Molecular Microbiology*, *96*(4), 839–860.

Gates, L. A., Foulds, C. E. & O'Malley, B. W. (2018). *Histone marks in the 'drivers seat': functional roles in steering the transcription cycle*. *42*(12), 977–989.

Georgakopoulos, P., Lockington, R. A. & Kelly, J. M. (2013). The Spt-Ada-Gcn5 Acetyltransferase (SAGA) Complex in *Aspergillus nidulans*. *PLoS ONE*, *8*(6).

Gómez, E. B., Espinosa, J. M. & Forsburg, S. L. (2005). *Schizosaccharomyces pombe* mst2+ encodes a MYST family histone acetyltransferase that negatively regulates telomere silencing. *Molecular and Cellular Biology*, *25*(20), 8887–8903.

Grant, P. A., Duggan, L., Côté, J., Roberts, S. M., Brownell, J. E., Candau, R., Ohba, R., Owen-Hughes, T., Allis, C. D., Winston, F., Berger, S. L. & Workman, J. L. (1997). Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: Characterization of an ada complex and the saga (spt/ada) complex. *Genes and Development*, *11*(13), 1640–1650.

Grewal, S. I. S. & Jia, S. (2007). Heterochromatin revisited. *Nature Reviews Genetics*, *8*(1), 35–46.

Grigoriev, I. V, Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R., Riley, R., Salamov, A., Zhao, X. & Korzeniewski, F. (2014). MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Research*, *42*(D1), D699–D704.

Grognet, P., Timpano, H., Carlier, F., Aït-Benkhali, J., Berteaux-Lecellier, V., Debuchy, R., Bidard, F. & Malagnac, F. (2019). A RID-like putative cytosine methyltransferase homologue controls sexual development in the fungus *Podospora anserina*. *PLoS Genetics*, *15*(8), e1008086.

Grunstein, M. & Gasser, S. M. (2013). Epigenetics in *Saccharomyces cerevisiae*. *Cold Spring Harbor Perspectives in Biology*, *5*(7).

Hajheidari, M., Koncz, C. & Bucher, M. (2019). Chromatin evolution-key innovations underpinning morphological complexity. *Frontiers in Plant Science*, *10*, 1–12.

He, F. (2011). Bradford protein assay. *Bio-Protocol*, 45.

Helmlinger, D., Papai, G., Devys, D. & Tora, L. (2021). What do the structures of GCN5-containing complexes teach us about their function? *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, *1864*(2).

Honda, S., Bicocca, V. T., Gessaman, J. D., Rountree, M. R., Yokoyama, A., Yu, E. Y., Selker, J. M. L. & Selker, E. U. (2016). Dual chromatin recognition by the histone deacetylase complex HCHC is required for proper DNA methylation in *Neurospora crassa*. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(41), E6135–E6144.

Hong, S. H., Cho, Y. W., Yu, L. R., Yu, H., Veenstra, T. D. & Ge, K. (2007). Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(47), 18439–18444.

Huff, J. T. & Zilberman, D. (2014). Dnmt1-independent CG methylation contributes to nucleosome positioning in diverse eukaryotes. *Cell*, *156*(6), 1286–1297.

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, *9*(03), 90–95.

Isaac, S., Walfridsson, J., Zohar, T., Lazar, D., Kahan, T., Ekwall, K. & Cohen, A. (2007). Interaction of epe1 with the heterochromatin assembly pathway in *Schizosaccharomyces pombe*. *Genetics*, *175*(4), 1549–1560.

James, T. Y., Stajich, J. E., Hittinger, C. T. & Rokas, A. (2020). Toward a Fully Resolved Fungal Tree of Life. *Annual Review of Microbiology*, *74*(1), 291–313.

Jamieson, K., Rountree, M. R., Lewis, Z. A., Stajich, J. E. & Selker, E. U. (2013). Regional control of histone H3 lysine 27 methylation in *Neurospora*. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(15), 6027–6032.

Jamieson, K., Wiles, E. T., McNaught, K. J., Sidoli, S., Leggett, N., Shao, Y., Garcia, B. A. & Selker, E. U. (2016). Loss of HP1 causes depletion of H3K27me3 from facultative heterochromatin and gain of H3K27me2 at constitutive heterochromatin. *Genome Research*, *26*(1), 97–107.

Janevska, S., Baumann, L., Sieber, C. M. K., Münsterkötter, M., Ulrich, J., Kämper, J., Güldener, U. & Tudzynski, B. (2018). Elucidation of the two H3K36me3 histone methyltransferases Set2 and Ash1 in *Fusarium fujikuroi* unravels their different chromosomal targets and a major impact of Ash1 on genome stability. *Genetics*, *208*, 153–171.

Jenuwein, T. & Allis, C. D. (2001). Translating the histone code. *Science*, *293*(5532), 1074–1080.

Job, G., Brugger, C., Xu, T., Lowe, B. R., Pfister, Y., Qu, C., Shanker, S., Sanz, J. I. B., Partridge, J. F. & Schalch, T. (2016). SHREC silences heterochromatin via distinct remodeling and deacetylation modules. *Molecular Cell*, *62*(2), 207–221.

Jones, R. (2014). Identification and verification of lysine propionylation and butyrylation in yeast core histones using PTMap software. *Bone*, *23*(1), 1–7.

Jurkowski, T. P. & Jeltsch, A. (2011). On the evolutionary origin of eukaryotic DNA

methyltransferases and Dnmt2. *PLoS ONE*, *6*(11), 1–9.
Kamei, M., Ameri, A. J., Ferraro, A. R., Bar-peled, Y. & Zhao, F. (2021). *IMITATION SWITCH is required for normal chromatin structure and gene repression in PRC2 target domains*. *118*(4).
Katoh, K. & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780.
Kawahara, T., Siegel, T. N., Ingram, A. K., Alsford, S., Cross, G. A. M. & Horn, D. (2008). Two essential MYST-family proteins display distinct roles in histone H4K10 acetylation and telomeric silencing in trypanosomes. *Molecular Microbiology*, *69*(4), 1054–1068.
Kawauchi, M., Nishiura, M. & Iwashita, K. (2013). Fungus-specific sirtuin HstD coordinates secondary metabolism and development through control of LaeA. *Eukaryotic Cell*, *12*(8), 1087–1096.
Kim, J., Park, S. & Lee, J. S. (2018). Epigenetic control of oxidative stresses by histone acetyltransferases in *Candida albicans*. *Journal of Microbiology and Biotechnology*, *28*(2), 181–189.
Kjærbølling, I., Vesth, T., Frisvad, J. C., Nybo, J. L., Theobald, S., Kildgaard, S., Petersen, T. I., Kuo, A., Sato, A., Lyhne, E. K., Kogle, M. E., Wiebenga, A., Kun, R. S., Lubbers, R. J. M., Mäkelä, M. R., Barry, K., Chovatia, M., Clum, A., Daum, C., … Andersen, M. R. (2020). A comparative genomics study of 23 *Aspergillus* species from section *Flavi*. *Nature Communications*, *11*(1).
Klose, R. J., Yamane, K., Bae, Y., Zhang, D., Erdjument-Bromage, H., Tempst, P., Wong, J. & Zhang, Y. (2006). The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature*, *442*(7100), 312–316.
Ko, Y. H., So, K. K., Chun, J. & Kim, D. H. (2021). Distinct roles of two dna methyltransferases from *cryphonectria parasitica* in fungal virulence, responses to hypovirus infection, and viral clearance. *MBio*, *12*(1), 1–16.
Kollenstart, L., De Groot, A. J. L., Janssen, G. M. C., Cheng, X., Vreeken, K., Martino, F., Côté, J., Van Veelen, P. A. & Van Attikum, H. (2019). Gcn5 and esa1 function as histone crotonyltransferases to regulate crotonylation-dependent transcription. *Journal of Biological Chemistry*, *294*(52), 20122–20134.
Kouzminova, E. & Selker, E. U. (2001). Dim-2 encodes a DNA methyltransferase responsible for all known cytosine methylation in *Neurospora*. *The EMBO Journal*, *20*(15), 4309–4323.
Kramer, H. M., Seidl, M. F., Thomma, B. P. H. J. & Cook, D. E. (2022). Local rather than global H3K27me3 dynamics associates with differential gene expression in *Verticillium dahliae*. *MBio*, *13*(1), e0356621.
Krogan, N. J. & Greenblatt, J. F. (2001). Characterization of a Six-Subunit Holo-Elongator Complex Required for the Regulated Expression of a Group of Genes in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, *21*(23), 8203–8212.
Kurdistani, S. K. & Grunstein, M. (2003). Histone acetylation and deacetylation in yeast. *Nature Reviews Molecular Cell Biology*, *4*(4), 276–284.
Labrador, M. & Corces, V. G. (2002). Setting the boundaries of chromatin domains and nuclear organization. *Cell*, *111*(2), 151–154.
Lee, I., Oh, J. H., Keats Shwab, E., Dagenais, T. R. T., Andes, D. & Keller, N. P. (2009). HdaA, a class 2 histone deacetylase of *Aspergillus fumigatus*, affects germination and secondary metabolite production. *Fungal Genetics and Biology*, *46*(10), 782–790.
Letunic, I. & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research*, *47*(W1), W256–W259.
Lewis, Z. A., Adhvaryu, K. K., Honda, S., Shiver, A. L., Knip, M., Sack, R. & Selker, E. U. (2010). DNA methylation and normal chromosome behavior in *Neurospora* depend on five components of a histone methyltransferase complex, DCDC. *PLoS Genetics*, *6*(11), 1–11.
Li, M., Phatnani, H. P., Guan, Z., Sage, H., Greenleaf, A. L. & Zhou, P. (2005). Solution structure of the Set2–Rpb1 interacting domain of human Set2 and its interaction with the hyperphosphorylated C-terminal domain of Rpb1. *Proceedings of the National Academy of Sciences*, *102*(49), 17636–17641.
Li, S., Yue, Z. & Tanaka, T. U. (2017). Smc3 deacetylation by Hos1 facilitates efficient dissolution

of sister chromatid cohesion during early anaphase. *Molecular Cell*, *68*(3), 605-614.e4.
Li, X., Pan, L., Wang, B. & Pan, L. (2019). The histone deacetylases HosA and HdaA affect the phenotype and transcriptomic and metabolic profiles of *Aspergillus niger*. *Toxins*, *11*(9).
Liu, B., Lin, Y., Darwanto, A., Song, X., Xu, G. & Zhang, K. (2009). Identification and characterization of propionylation at histone H3 lysine 23 in mammalian cells. *Journal of Biological Chemistry*, *284*(47), 32288–32295.
Liu, S.-Y., Lin, J.-Q., Wu, H.-L., Wang, C.-C., Huang, S.-J., Luo, Y.-F., Sun, J.-H., Zhou, J.-X., Yan, S.-J. & He, J.-G. (2012). Bisulfite sequencing reveals that *Aspergillus flavus* holds a hollow in DNA methylation. *PloS One*, *7*(1), e30349.
Luger, K. (2003). Structure and dynamic behavior of nucleosomes. *Current Opinion in Genetics & Development*, *13*(2), 127–135.
Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. (1997). Crystal structure of the nucleosome resolution core particle at 2.8 Å. *Nature*, *389*, 251–260.
Madhani, H. D. (2020). Unbelievable but true: Epigenetics and chromatin in fungi. *Trends in Genetics*.
Marina, D. B., Shankar, S., Natarajan, P., Finn, K. J. & Madhani, H. D. (2013). A conserved ncRNA-binding protein recruits silencing factors to heterochromatin through an RNAi-independent mechanism. *Genes & Development*, *27*(17), 1851–1856.
Meijsing, S. H. & Ehrenhofer-Murray, A. E. (2001). The silencing complex SAS-I links histone acetylation to the assembly of repressed chromatin by CAF-I and Asf1 in *Saccharomyces cerevisiae*. *Genes and Development*, *15*(23), 3169–3182.
Millar, C. B., Xu, F., Zhang, K. & Grunstein, M. (2006). Acetylation of H2AZ Lys 14 is associated with genome-wide gene activity in yeast. *Genes and Development*, *20*(6), 711–722.
Moller, M., Habig, M., Lorrain, C., Feurtey, A., Haueisen, J., Fagundes, W. C., Alizadeh, A., Freitag, M. & Stukenbrock, E. H. (2021). Recent loss of the Dim2 DNA methyltransferase decreases mutation rate in repeats and changes evolutionary trajectory in a fungal pathogen. *PLoS Genetics*, *17*(3), 1–27.
Nai, Y. S., Huang, Y. C., Yen, M. R. & Chen, P. Y. (2021). Diversity of fungal DNA methyltransferases and their association with DNA methylation patterns. *Frontiers in Microbiology*, *11*.
Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274.
Nielsen, J. C., Grijseels, S., Prigent, S., Ji, B., Dainat, J., Nielsen, K. F., Frisvad, J. C., Workman, M. & Nielsen, J. (2017). Global analysis of biosynthetic gene clusters reveals vast potential of secondary metabolite production in *Penicillium* species. *Nature Microbiology*, *2*(April).
Noberini, R., Restellini, C., Savoia, E. O. & Bonaldi, T. (2020). Enrichment of histones from patient samples for mass spectrometry-based analysis of post-translational modifications. *Methods*, *184*, 19–28.
Noberini, R., Savoia, E. O., Brandini, S., Greco, F., Marra, F., Bertalot, G., Pruneri, G., McDonnell, L. A. & Bonaldi, T. (2021). Spatial epi-proteomics enabled by histone post-translational modification analysis from low-abundance clinical samples. *Clinical Epigenetics*, *13*(1), 1–16.
Nugent, R. L., Johnsson, A., Fleharty, B., Gogol, M., Xue-Franzén, Y., Seidel, C., Wright, A. P. & Forsburg, S. L. (2010). Expression profiling of *S. pombe* acetyltransferase mutants identifies redundant pathways of gene regulation. *Bmc Genomics*, *11*(1), 59.
Okada, Y., Yamagata, K., Hong, K., Wakayama, T. & Zhang, Y. (2010). A role for the elongator complex in zygotic paternal genome demethylation. *Nature*, *463*(7280), 554–558.
Okano, M., Xie, S. & Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature Genetics*, *19*(3), 219–220.
Palmer, J. M., Bok, J. W., Lee, S., Dagenais, T. R. T., Andes, D. R., Kontoyiannis, D. P. & Keller, N. P. (2013). Loss of CclA, required for histone 3 lysine 4 methylation, decreases growth but increases secondary metabolite production in *Aspergillus fumigatus*. *PeerJ*, *2013*(1), 1–18.
Palmer, J. M., Perrin, R. M., Dagenais, T. R. T. & Keller, N. P. (2008). H3K9 methylation regulates growth and development in *Aspergillus fumigatus*. *Eukaryotic Cell*, *7*(12), 2052–2060.

Papanicolaou, K., O'Rourke, B. & Foster, D. B. (2014). Metabolism leaves its mark on the powerhouse: recent progress in post-translational modifications of lysine in mitochondria. *Frontiers in Physiology*, *5 JUL*.

Pérez-Martínez, M. E., Benet, M., Alepuz, P. & Tordera, V. (2020). Nut1/Hos1 and Sas2/Rpd3 control the H3 acetylation of two different sets of osmotic stress-induced genes. *Epigenetics*, *15*(3), 251–271.

Pesavento, J. J., Mizzen, C. A. & Kelleher, N. L. (2006). Quantitative analysis of modified proteins and their positional isomers by tandem mass spectrometry: human histone H4. *Analytical Chemistry*, *78*(13), 4271–4280.

Peterson, C. L. & Laniel, M.-A. (2004). Histones and histone modifications. *Current Biology*, *14*(14), R546–R551.

Porras-Yakushi, T. R., Whitelegge, J. P. & Clarke, S. (2007). Yeast ribosomal/cytochrome c SET domain methyltransferase subfamily: Identification of Rpl23ab methylation sites and recognition motifs. *Journal of Biological Chemistry*, *282*(17), 12368–12376.

Proffitt, J. H., Davie, J. R., Swinton, D. & Hattman, S. (1984). 5-Methylcytosine is not detectable in *Saccharomyces cerevisiae* DNA. *Molecular and Cellular Biology*, *4*(5), 985–988.

Prudhomme, N., Gianetto-Hill, C., Pastora, R., Cheung, W.-F., Allen-Vercoe, E., McLean, M. D., Cossar, D. & Geddes-McAlister, J. (2021). Quantitative proteomic profiling of shake flask versus bioreactor growth reveals distinct responses of *Agrobacterium tumefaciens* for preparation in molecular pharming. *Canadian Journal of Microbiology*, *67*(1), 75–84.

Qian, S., Wang, Y., Ma, H. & Zhang, L. (2015). Expansion and functional divergence of Jumonji C-containing histone demethylases: significance of duplications in ancestral angiosperms and vertebrates. *Plant Physiology*, *168*(4), 1321–1337.

Raunser, S., Magnani, R., Huang, Z., Houtz, R. L., Trievel, R. C., Penczek, P. A. & Walz, T. (2009). Rubisco in complex with Rubisco large subunit methyltransferase. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(9), 3160–3165.

Rice, P., Longden, I. & Bleasby, A. (2000). *EMBOSS: the European molecular biology open software suite*. Elsevier current trends.

Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W. W. M. P., Wilm, M., Aasland, R. & Stewart, A. F. (2001). The *Saccharomyces cerevisiae* Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *The EMBO Journal*, *20*(24), 7137–7148.

Roth, S. Y., Denu, J. M. & Allis, C. D. (2001). Histone acetyltransferases. *Annual Review of Biochemistry*, *70*(1), 81–120.

Rountree, M. R. & Selker, E. U. (2010). DNA methylation and the formation of heterochromatin in *Neurospora crassa*. *Heredity*, *105*(1), 38–44.

Samson, R. A., Visagie, C. M., Houbraken, J., Hong, S. B., Hubka, V., Klaassen, C. H. W., Perrone, G., Seifert, K. A., Susca, A., Tanney, J. B., Varga, J., Kocsubé, S., Szigeti, G., Yaguchi, T. & Frisvad, J. C. (2014). Phylogeny, identification and nomenclature of the genus *Aspergillus*. *Studies in Mycology*, *78*(1), 141–173.

Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C. T., Schreiber, S. L., Mellor, J. & Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature*, *419*(6905), 407–411.

Searle, N. E., Torres-Machorro, A. L. & Pillus, L. (2017). Chromatin regulation by the NuA4 acetyltransferase complex is mediated by essential interactions between enhancer of polycomb (Epl1) and Esa1. *Genetics*, *205*(3), 1125–1137.

Sendra, R., Tse, C. & Hansen, J. C. (2000). The yeast histone acetyltransferase A2 complex, but not free Gcn5p, binds stably to nucleosomal arrays. *Journal of Biological Chemistry*, *275*(32), 24928–24934.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212.

Slater, G. S. C. & Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics*, *6*(1), 31.

Sneeringer, C. J., Scott, M. P., Kuntz, K. W., Knutson, S. K., Pollock, R. M., Richon, V. M. &

Copeland, R. A. (2010). Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. *Proceedings of the National Academy of Sciences*, *107*(49), 20980–20985.

Souza, S. C., Pereira, V. M., Moreira, S. I., Costa, S. S., Moreira, G. M., Mendes, W. O., Nery, E. M., Alves, E., Chalfoun, S. M., Moreira, F. M. S. & Batista, L. R. (2019). *Aspergillus trisporus*: A new *Jani* section species from Brazilian soil. *Current Research in Environmental and Applied Mycology*, *9*(1), 175–186.

Spellmon, N., Holcomb, J., Trescott, L., Sirinupong, N. & Yang, Z. (2015). Structure and function of SET and MYND domain-containing proteins. *International Journal of Molecular Sciences*, *16*(1), 1406–1408.

Steenwyk, J. L., Mead, M. E., Knowles, S. L., Raja, H. A., Roberts, C. D., Bader, O., Houbraken, J., Goldman, G. H., Oberlies, N. H. & Rokas, A. (2020). Variation among biosynthetic gene clusters, secondary metabolite profiles, and cards of virulence across *Aspergillus* species. *Genetics*, *216*(2), 481–497.

Steenwyk, J. L., Shen, X. X., Lind, A. L., Goldman, G. H. & Rokas, A. (2019). A robust phylogenomic time tree for biotechnologically and medically important fungi in the genera *Aspergillus* and *Penicillium*. *MBio*, *10*(4), 1–25.

Sterner, D. E. & Berger, S. L. (2000). Acetylation of histones and transcription-related factors. *Microbiology and Molecular Biology Reviews*, *64*(2), 435–459.

Strahl, B. D., Grant, P. A., Briggs, S. D., Sun, Z.-W., Bone, J. R., Caldwell, J. A., Mollah, S., Cook, R. G., Shabanowitz, J. & Hunt, D. F. (2002). Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Molecular and Cellular Biology*, *22*(5), 1298–1306.

Suka, N., Suka, Y., Carmen, A. A., Wu, J. & Grunstein, M. (2001). Highly specific antibodies determine histone acetylation site usage in yeast heterochromatin and euchromatin. *Molecular Cell*, *8*(2), 473–479.

Sun, B. Da, Houbraken, J., Frisvad, J. C., Jiang, X. Z., Chen, A. J. & Samson, R. A. (2020). New species in *Aspergillus* section *Usti* and an overview of *Aspergillus* section *Cavernicolarum*. *International Journal of Systematic and Evolutionary Microbiology*.

Takechi, S. & Nakayama, T. (1999). Sas3 is a histone acetyltransferase and requires a zinc finger motif. *Biochemical and Biophysical Research Communications*, *266*(2), 405–410.

Tamaru, H. (2010). Confining euchromatin/heterochromatin territory: Jumonji crosses the line. *Genes and Development*, *24*(14), 1465–1478.

Tao, Y., Neppl, R. L., Huang, Z.-P., Chen, J., Tang, R.-H., Cao, R., Zhang, Y., Jin, S.-W. & Wang, D.-Z. (2011). The histone methyltransferase Set7/9 promotes myoblast differentiation and myofibril assembly. *Journal of Cell Biology*, *194*(4), 551–565.

Taverna, S. D., Ilin, S., Rogers, R. S., Tanny, J. C., Lavender, H., Li, H., Baker, L., Boyle, J., Blair, L. P., Chait, B. T. T., Patel, D. J., Aitchison, J. D., Tackett, A. J. & Allis, C. D. (2006). Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. *Molecular Cell*, *24*(5), 785–796.

Torchia, J., Glass, C. & Rosenfeld, M. G. (1998). Co-activators and co-repressors in the integration of transcriptional responses. *Current Opinion in Cell Biology*, *10*(3), 373–383.

Tribus, M., Bauer, I., Galehr, J., Rieser, G., Trojer, P., Brosch, G., Loidl, P., Haas, H. & Graessle, S. (2010). A novel motif in fungal class 1 histone deacetylases is essential for growth and development of *Aspergillus*. *Molecular Biology of the Cell*, *21*(2), 345–353.

Trievel, R. C., Flynn, E. M., Houtz, R. L. & Hurley, J. H. (2003). Mechanism of multiple lysine methylation by the SET domain enzyme Rubisco LSMT. *Nature Structural Biology*, *10*(7), 545–552.

Turberfield, A. H., Kondo, T., Nakayama, M., Koseki, Y., King, H. W., Koseki, H. & Klose, R. J. (2019). KDM2 proteins constrain transcription from CpG island gene promoters independently of their histone demethylase activity. *Nucleic Acids Research*, *47*(17), 9005–9023.

Veerappan, C. S., Avramova, Z. & Moriyama, E. N. (2008). Evolution of SET-domain protein families in the unicellular and multicellular Ascomycota fungi. *BMC Evolutionary Biology*, *8*(1),

1–20.
Verzijlbergen, K. F., van Welsem, T., Sie, D., Lenstra, T. L., Turner, D. J., Holstege, F. C. P., Kerkhoven, R. M. & van Leeuwen, F. (2011). A barcode screen for epigenetic regulators reveals a role for the NuB4/HAT-B histone acetyltransferase complex in histone turnover. *PLoS Genetics*, *7*(10), 18–24.
Vesth, T. C., Nybo, J. L., Theobald, S., Frisvad, J. C., Larsen, T. O., Nielsen, K. F., Hoof, J. B., Brandl, J., Salamov, A., Riley, R., Gladden, J. M., Phatale, P., Nielsen, M. T., Lyhne, E. K., Kogle, M. E., Strasser, K., McDonnell, E., Barry, K., Clum, A., … Andersen, M. R. (2018). Investigation of inter- and intraspecies variation through genome sequencing of *Aspergillus* section *Nigri*. *Nature Genetics*, *50*(12), 1688–1695.
Wade, P. A., Pruss, D. & Wolffe, A. P. (1997). Histone acetylation: Chromatin in action. *Trends in Biochemical Sciences*, *22*(4), 128–132.
Wang, A., Kurdistani, S. K. & Grunstein, M. (2002). Requirement of Hos2 histone deacetylase for gene activity in yeast. *Science*, *298*(5597), 1412–1414.
Wang, X., Chang, P., Ding, J. & Chen, J. (2013). Distinct and redundant roles of the two MYST histone acetyltransferases Esa1 and Sas2 in cell growth and morphogenesis of *Candida albicans*. *Eukaryotic Cell*, *12*(3), 438–449.
Wang, Y., Reddy, B., Thompson, J., Wang, H., Noma, K., Yates III, J. R. & Jia, S. (2009). Regulation of Set9-mediated H4K20 methylation by a PWWP domain protein. *Molecular Cell*, *33*(4), 428–437.
Waskom, M. L. (2021). Seaborn: statistical data visualization. *Journal of Open Source Software*, *6*(60), 3021.
Weiner, A. K. M., Cerón-Romero, M. A., Yan, Y. & Katz, L. A. (2020). Phylogenomics of the epigenetic toolkit reveals punctate retention of genes across eukaryotes. *Genome Biology and Evolution*, 1–39.
Winkler, G. S., Kristjuhan, A., Erdjument-Bromage, H., Tempst, P. & Svejstrup, J. Q. (2002). Elongator is a histone H3 and H4 acetyltransferase important for normal histone acetylation levels *in vivo*. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(6), 3517–3522.
Wittschieben, B., Otero, G., De Bizemont, T., Fellows, J., Erdjument-Bromage, H., Ohba, R., Li, Y., Allis, C. D., Tempst, P. & Svejstrup, J. Q. (1999). A novel histone acetyltransferase is an integral subunit of elongating RNA polymerase II holoenzyme. *Molecular Cell*, *4*(1), 123–128.
Wolffe, A. P. & Matzke, M. A. (1999). Epigenetics: Regulation through repression. *Science*, *286*(5439), 481–486.
Wu, J., Suka, N., Carlson, M. & Grunstein, M. (2001). TUP1 utilizes histone H3/H2B-specific HDA1 deacetylase to repress gene activity in yeast. *Molecular Cell*, *7*(1), 117–126.
Xiao, B., Wilson, J. R. & Gamblin, S. J. (2003). SET domains and histone methylation. *Current Opinion in Structural Biology*, *13*(6), 699–705.
Xiao, T., Hall, H., Kizer, K. O., Shibata, Y., Hall, M. C., Borchers, C. H. & Strahl, B. D. (2003). Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. *Genes & Development*, *17*(5), 654–663.
Yang, K., Liang, L., Ran, F., Liu, Y., Li, Z., Lan, H., Gao, P., Zhuang, Z., Zhang, F., Nie, X., Yirga, S. K. & Wang, S. (2016). The DmtA methyltransferase contributes to *Aspergillus flavus* conidiation, sclerotial production, aflatoxin biosynthesis and virulence. *Scientific Reports*, *6*, 1–13.
Yuan, Y., Yuan, H., Yang, G., Yun, H., Zhao, M., Liu, Z., Zhao, L., Geng, Y., Liu, L., Wang, J., Zhang, H., Wang, Y. & Zhang, X. D. (2020). IFN-α confers epigenetic regulation of HBV cccDNA minichromosome by modulating GCN5-mediated succinylation of histone H3K79 to clear HBV cccDNA. *Clinical Epigenetics*, *12*(1), 1–16.
Yuan, Z.-F., Sidoli, S., Marchione, D. M., Simithy, J., Janssen, K. A., Szurgot, M. R. & Garcia, B. A. (2018). EpiProfile 2.0: a computational platform for processing epi-proteomics mass spectrometry data. *Journal of Proteome Research*, *17*(7), 2533–2541.
Zhang, X., Tamaru, H., Khan, S. I., Horton, J. R., Keefe, L. J., Selker, E. U. & Cheng, X. (2002). Structure of the *Neurospora* SET Domain Protein DIM-5, a Histone H3 Lysine

Methyltransferase. *Cell*, *111*, 9.
Zhang, Z., He, C., Chen, Y., Li, B. & Tian, S. (2021). DNA methyltransferases regulate pathogenicity of *Botrytis cinerea* to horticultural crops. *Journal of Fungi*, *7*(8).

Chapter 4

Detection and quantification of the histone code in the fungal genus *Aspergillus*

Xin Zhang, Roberta Noberini, Alessandro Vai,
Tiziana Bonaldi+, Michael F. Seidl+, Jérôme Collemare+

+These authors contributed equally
Fungal Genetics and Biology, 2023

Abstract

In eukaryotes, the combination of different histone post-translational modifications (PTMs) – the histone code – impacts the chromatin organization as compact and transcriptionally silent heterochromatin or accessible and transcriptionally active euchromatin. Although specific histone PTMs have been studied in fungi, an overview of histone PTMs and their relative abundance is still lacking. Here, we used mass spectrometry to detect and quantify histone PTMs in three fungal species belonging to three distinct taxonomic sections of the genus *Aspergillus* (*Aspergillus niger*, *Aspergillus nidulans* (two strains), and *Aspergillus fumigatus*). We overall detected 23 different histone PTMs, including a majority of lysine methylations and acetylations, and 23 co-occurrence patterns of multiple histone PTMs. Among those, we report for the first time the detection of H3K79me1, H3K79me2, and H4K31ac in Aspergilli. Although all three species harbor the same PTMs, we found significant differences in the relative abundance of H3K9me1/2/3, H3K14ac, H3K36me1, and H3K79me1, as well as the co-occurrence of acetylation on both K18 and K23 of histone H3 in a strain-specific manner. Our results provide novel insights into the underexplored complexity of the histone code in filamentous fungi and its functional implications on genome architecture and gene regulation.

Introduction

In eukaryotes, genetic information is stored within the nucleus as chromatin, a complex of DNA and proteins of which the basic unit is the nucleosome (Luger et al., 1997). A nucleosome is composed of a histone octamer formed by two copies of each histone protein H2A, H2B, H3, and H4 wrapped by a DNA stretch of 145-147 bp (Luger et al., 1997). Nucleosomes are linked by 30-50 bp DNA sequence and the linker histone protein H1 (Luger et al., 1997). Histone proteins are highly conserved, consistent with their fundamental roles in structuring, scaffolding, and packaging DNA (van Holde, 1989). Each histone protein contains a globular central domain that interacts with the nucleosomal DNA, and a lysine- and arginine-rich N-terminal tail that is subject to a diverse array of chemical modifications (van Holde, 1989). These histone post-translational modifications (PTMs) contribute to the regulation of chromatin accessibility by two mechanisms: first, they directly alter chromatin packaging by changing histone proteins' net charges; second, histone PTM-specific binding proteins are recruited and can influence inter-nucleosomal interactions (Henikoff & Shilatifard, 2011). The resulting chromatin changes lead to transitions between inaccessible and transcriptionally silent chromatin regions, the so-called (facultative or constitutive) heterochromatin, and the accessible and transcriptionally active regions, the so-called euchromatin (Henikoff & Shilatifard, 2011). The combination of different histone PTMs at a single genomic locus – the histone code – contributes to the accessibility of chromatin for the transcriptional machinery, thereby regulating gene expression to influence not only growth and development but also responses to environmental challenges (van Holde, 1989).

In the last decades, a plethora of different histone PTMs have been uncovered in diverse organisms from plants to humans, with an enrichment of histone mono-, di-, and tri-methylation (me) and acetylation (ac) (Huang et al., 2015). Accordingly, the vast majority of studies in fungi thus far have focused on these PTMs (Collemare & Seidl, 2019). Euchromatin has been typically associated with methylations on histone protein H3 at lysine 4 (K4), K36 and K79, and acetylations on H3 at K9, K14, K18, K23, and K27 and H4 at K5, K8, K12, and K16 (Bhaumik et al., 2007; Shilatifard, 2006). For instance, in the budding yeast *Saccharomyces cerevisiae* methylation on the lysine 4 of histone protein H3 (H3K4me) and H3K36me are enriched at the 5' and 3' end of actively transcribed genes, respectively, while H3K79me has been detected at the coding regions of active genes (Farooq et al., 2016; Pokholok et al., 2005). H3K4ac is located just upstream of actively transcribed genes (Guillemette et al., 2011), and H3K9ac and H4K5ac have been shown to induce transcription elongation *in vitro* (Gates et al., 2017). In members of the filamentous Aspergilli, some of these euchromatin marks are shown to positively influence the production of secondary metabolites (SMs), including H3K4me in *Aspergillus oryzae* (Shinohara et al., 2016); H3K36me, H3K14/18/23ac (acetylation on either of these lysines), and H4K16ac in *Aspergillus flavus* (Chen et al., 2022; Zhuang et al., 2022); H3K9ac in *Aspergillus fumigatus* (Y. Zhang et al., 2021) and *Aspergillus nidulans* (Reyes-Dominguez et al., 2010); and H3K27ac in *A. fumigatus*

(Adlakha et al., 2016). Especially, the diverse functions of histone acetyltransferase complex (SAGA complex) or its crucial subunits (Gcn5 and AdaB) have been extensively studied in *A. nidulans*, in which they can influence asexual development (Cánovas et al., 2014), production of natural products in response to the interaction with bacteria (Nützmann et al., 2011), and nucleosome positioning (Reyes-Dominguez et al., 2008). Histone PTMs can also be linked to gene silencing, as first shown for H3K9me and H3K27me which are marks for heterochromatin in animals and plants (Henikoff & Shilatifard, 2011). H3K9me3 plays a similar role in fungi as this PTM is typically located at telomeres, sub-telomeres, centromeres, and transposon-rich regions in the fission yeast *Schizosaccharomyces pombe* (Yamada et al., 2005), the filamentous model fungus *Neurospora crassa* (Rountree & Selker, 2010) and many other filamentous fungi (Freitag, 2017). H3K9me can influence SM biosynthesis in *A. nidulans* (Reyes-Dominguez et al., 2010) and *A. fumigatus* (Colabardini et al., 2022), however, as H3K9me is not always found located near SM gene clusters, it remains obscure whether it is directly linked to the regulation of SM gene clusters in Aspergilli and raises the possibility of the interaction of multiple histone PTMs or the distal regulation based on a higher level of chromatin structure. Facultative heterochromatin is linked with H3K27me3, and unsurprisingly, it does generally not overlap with the euchromatin mark H3K4me in the cereal pathogens *Fusarium graminearum* (Connolly et al., 2013) and *Fusarium fujikuroi* (Wiemann et al., 2013). By contrast, SET7 (KMT6) which is responsible for H3K27me is absent in Aspergilli, and consequently, no H3K27me is detected (X. Zhang et al., 2022). Although many PTMs have been clearly associated with gene activation or repression only, their effect can also be more subtle, depending on the genomic context. In *N. crassa* and *F. fujikuroi*, H3K36me catalyzed by SET-2 drives gene expression at the euchromatic region, while H3K36me catalyzed by ASH-1 at sub-telomeric regions is associated with gene repression (Bicocca et al., 2018; Janevska et al., 2018). This complexity is even higher when considering the role of less abundant histone PTMs like histone acylation, phosphorylation, ubiquitination, and sumoylation, which thus far have been mostly reported in humans (Tan et al., 2011). In fungi, such less abundant PTMs have been only reported scarcely in *S. cerevisiae* (diverse acylations, phosphorylation, ubiquitination, and sumoylation) and *A. nidulans* (phosphorylation, ubiquitination, and sumoylation), and the phosphorylation of histone H3 serine 10 correlates with chromosome condensation (De Souza et al., 2000).

Large-scale mass spectrometry (MS) studies have been performed on bulk histones from model species such as human, yeast, and *Arabidopsis* to uncover the full pattern of histone PTMs and their interactions (Grau-Bové et al., 2022; Johnson et al., 2004). However, studies in filamentous fungi have mostly focused on quantitative changes of targeted histone PTMs in mutant strains deficient in specific histone-modifying enzymes (Gacek-Matthews et al., 2015, 2016). For instance, mutants of histone H3 demethylase gene *KdmA* or *KdmB* in *A. nidulans* show increased H3K36me3 or H3K4me3, respectively (Gacek-Matthews et al., 2015, 2016). Our previous evolutionary analysis of 16 chromatin modifier complexes across the *Aspergillus*

genus demonstrated the conservation of most of the catalytic genes responsible for a variety of histone PTMs (X. Zhang et al., 2022). However, we still lack an overview of (co-)occurrence of histone PTMs in filamentous fungi, as well as a better understanding of the quantitative diversity of histone PTMs between different species. Here, we used quantitative MS to uncover histone PTMs in three *Aspergillus* species (*Aspergillus niger*, *A. nidulans*, and *A. fumigatus)* from distinct taxonomic sections and to report their estimated quantitative intra- and inter-species differences.

Results and discussion

The study of histone PTMs in filamentous fungi has mostly been restricted to targeted analyses based on knowledge from other organisms. However, the discovery of novel histone PTMs and their potential co-occurrence on a single nucleosome requires the use of untargeted strategies. We here applied a quantitative MS approach to identify the (co-)occurrence of histone PTMs and to quantify their relative abundance in three *Aspergillus* species: *A. niger*, *A. fumigatus*, and *A. nidulans*, which belong to distinct taxonomic sections (Barrett et al., 2020). Protein sequences of histones H2A, H2B, H3, and H4 from human, yeast, and the three Aspergilli show very high levels of sequence conservation, especially in the core domains of all histone proteins that fold into the characteristic bulk octamer around which the DNA strand wraps, as well as in the N-terminal tails of H3 and H4 that protrude from the surface of the nucleosome and harbors most of histone PTMs (Mersfelder & Parthun, 2006) (Figure 1). The majority of well-described histone PTMs, *i.e.*, methylation and acetylation, as well as less abundant PTMs such as crotonylation and ubiquitination predominantly occur on lysine residues, but histone PTMs can also occur on other residues such as serine, threonine, and arginine (Grau-Bové et al., 2022). For histone protein H3, we observed the conservation of all lysine residues in human, yeast, and Aspergilli, except for two additional lysine residues at positions 121 and 125 in *S. cerevisiae* (Figure 1). On these two residues, Cul8/Rtt101, an enzyme that is thought to be specific to *S. cerevisiae* as no homolog has yet been found in human or other fungi, can introduce histone ubiquitylation and thereby regulate nucleosome assembly (J. Han et al., 2013; Young et al., 2020). We similarly observed limited variation in the N-terminal tail of histone protein H4 and the global core domains of H2A and H2B, and most of the polymorphism occurs on amino acid residues that are not known to carry any histone PTM.

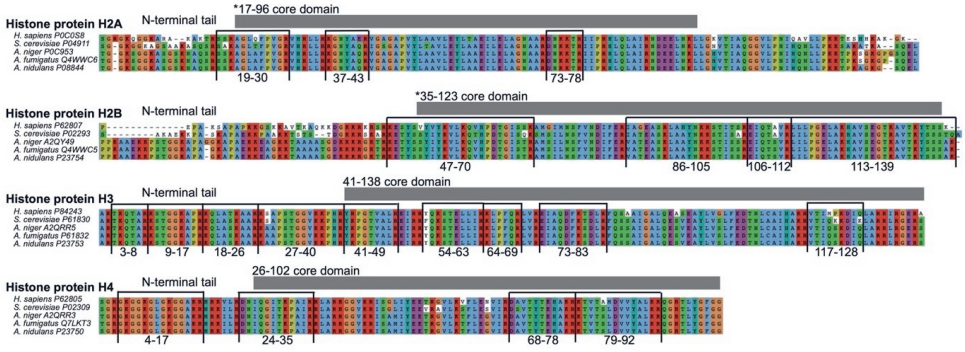
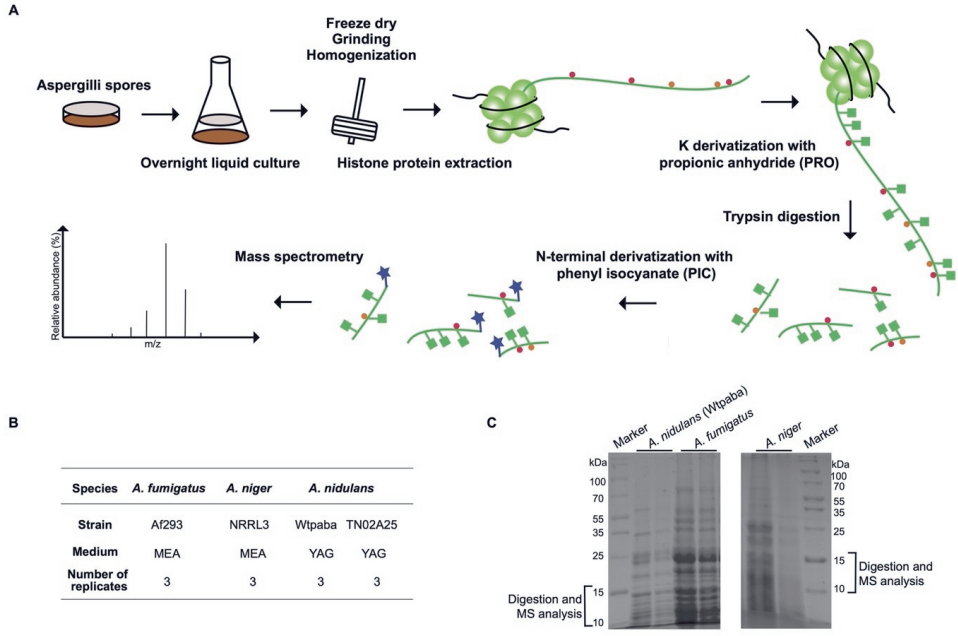


Figure 1 Histone proteins are well conserved in Aspergilli. Sequence alignment of histone proteins from human (H2A: P0C0S8, H2B: P62807, H3: P84243, and H4: P62805), yeast (*S. cerevisiae*, H2A: P04911, H2B: P02293, H3: P61830, and H4: P02309), and *A. niger* (H2A: P0C953, H2B: A2QY49, H3: A2QRR5, and H4: A2QRR3), *A. fumigatus* (H2A: Q4WWC6, H2B: Q4WWC5, H3: P61832, and H4: Q7LKT3) and *A. nidulans* (H2A: P08844, H2B: P23754, H3: P23753, and H4: P23750), retrieved from UniProt database. The highly conserved globular histone core regions are highlighted by grey boxes with numbers indicating their position in human histone proteins with the symbol * showing the mismatch of the positions between human and Aspergilli histone proteins H2A and H2B caused by the poor alignment of the N-terminal tails. Peptides detected by MS in this study are labeled by black lines (Table S1), and their relative positions in Aspergilli histone proteins are indicated.

Histone protein extraction and digestion for MS-based analyses

We adapted a well-established protocol for human primary cells (Noberini et al., 2020) to specifically extract histone proteins from filamentous fungi with three additional steps (freeze drying, grinding, and homogenization) for efficient disruption of fungal cell walls and release of nuclei (X. Zhang et al., 2021, 2022) (Figure 2A). We extracted histone proteins in three biological replicates from the three *Aspergillus* species; two strains of *A. nidulans* (Wtpaba and TN02A25) were cultured on YAG medium, while *A. niger* strain NRRL3 and *A. fumigatus* strain Af293 were grown on MEA medium (Figure 2B). After visualizing the extracted histone protein on SDS-PAGE gel, bands ranging from 10,000 to 16,000 Dalton, which correspond to the size of histone proteins (Zhou et al., 2020), were cut for in-gel digestion (Figure 2C). We used lysine derivatization with propionic anhydride on unmodified and mono-methylated lysines, followed by trypsin digestion and peptide N-terminal derivatization with phenyl isocyanate (PIC) to obtain histone peptides of amino acids length required for MS analysis. This method allows performing an 'Arg-C-like' digestion of histone proteins directly in the gel, which is the ideal digestion pattern for histone PTM analyses (Noberini et al., 2021). In addition, because the derivatization occurs only on unmodified or monomethylated lysines, but not on di- or tri-methylated residues, it allows better discrimination of isobaric peptides (Noberini et al., 2021). The N-terminal derivatization with PIC also aids in the detection of small and hydrophilic peptides, as well as the chromatographic separation of differentially acetylated

peptides, as described previously (Maile et al., 2015; Noberini et al., 2021). After digestion of isolated histone proteins from Aspergilli, we detected 20 different peptides from the four core histone proteins and three peptides from the histone protein variant H2A.Z (Figure 1 and Table S1). These peptides cover most of the lysine residues that are known to harbor histone PTMs.



Species	*A. fumigatus*	*A. niger*	*A. nidulans*	
Strain	Af293	NRRL3	Wtpaba	TN02A25
Medium	MEA	MEA	YAG	YAG
Number of replicates	3	3	3	3

Figure 2. Histone proteins extraction for MS-based analysis of histone PTMs in Aspergilli. (A) Schematic diagram of the experimental setup and workflow applied in this work. (B) Summary of *Aspergillus* strains' genetic background and culture conditions. The number of biological replicates for each condition is indicated. (C) Histone proteins were extracted from the different *Aspergillus* species, and samples were visualized on the 17% SDS-PAGE gel. The region ranging from 10,000 to 16,000 Dalton was cut for the in-gel digestion step.

MS uncovers histone PTMs in three *Aspergillus* species

MS analyses resulted in the detection of 23 individual histone PTMs, as well as 23 distinct co-occurrence patterns of multiple histone PTMs that occur on the same peptide (Figure 3 and Table S1). We did not detect any histone PTM on histone variant H2A.Z and histone H2A, but we were able to detect acetylation on K92 and co-occurrent mono-methylations on K98 and K99 on histone protein H2B (Table S1). These PTMs correspond to histone PTMs H2BR79ac (R: arginine) and H2BK85me1R86me1 in human, yet the functions of these PTMs are unknown (Gnesutta et al., 2013). For histone PTMs on histone proteins H3 and H4 in Aspergilli, we successfully detected methylations on histone H3 at lysines 4, 9, 36, and 79, acetylation on histone H3 at lysines 9, 14, 18, 23, 27, and 56, and histone H4 at lysines 5, 8, 12, 16, and

31 (Figure 3). All these modifications have previously been reported in human (Barski et al., 2007; Pesavento et al., 2006), *Arabidopsis* (Scheid et al., 2022), *S. cerevisiae* (Abshiru et al., 2016), and in a few filamentous fungi (Brosch et al., 2008; Collemare & Seidl, 2019; Freitag, 2017). However, it is to our knowledge the first time that H3K79me1, H3K79me2, and H4K31ac are detected experimentally in Aspergilli (Figure 3). Consistently with the occurrence of these histone PTMs, their catalytic enzymes, Dot1 (Disruptor of telomeric silencing 1) for H3K79me and Gcn5 for H4K31ac were firstly identified in *S. cerevisiae* (van Leeuwen et al., 2002) and the human parasite *Toxoplasma* (Sindikubwabo et al., 2017), respectively, and then found to be conserved across the fungal genus *Aspergillus* in our previous study (X. Zhang et al., 2022). Our analysis could not differentiate between the two H4 variants found in Aspergilli because they only differ by two residues (Cheng et al., 2003; Grant et al., 1999) (Figure 1), which are inside the histone core domain that was not detected.

While we detected most of the known PTMs as well as a few additional ones for the first time in Aspergilli, we did not detect some histone PTMs which were expected to occur. Several reasons can explain these observations. First, PTMs might not be identified because they are under the detection threshold used in both the instrument and analysis pipeline. This reason likely explains why we only detected abundant histone methylations and acetylations, but not less-abundant ones, e.g., H3S10P (S: serine and P: phosphorylation) and H3R2me, which have been reported in Aspergilli (Cheng et al., 2003). The abundance of certain expected PTMs like H3K79me3, H3K4ac, and H3K36ac could be higher and detectable under different growth conditions or developmental stages. Second, these PTMs are true absences in Aspergilli. We were not able to detect any methylation on H3K27, H4K5, H4K8, and H4K12 (Figure S1), which is expected due to the absence in Aspergilli of their catalytic enzymes, PRC2 (Freitag, 2017; X. Zhang et al., 2022) and SET5 (X. Zhang et al., 2022) (Figure S2), respectively. Third, the mutual exclusivity of different histone PTMs on the same residue could lead to biased occupancy by only one histone PTM. For instance, apart from H3K9, either acetylation or methylation was detected on H3 lysines (Figures 3 and S1). Such an explanation would indicate very homogenous samples regarding their chromatin status and PTMs. Last, antagonistic relationships between histone PTMs might contribute to some absences. For example, H4K16ac prevents H4K20me (Nishioka et al., 2002), explaining why we could detect the former but not the latter histone PTM in Aspergilli.

We also detected the co-occurrence of multiple histone PTMs on the same peptide (Figure 3). For example, we observed the co-occurrence of acetylations on different residues in the same peptide such as H3K9acK14ac, H3K18acK23ac, and H4 lysines 5, 8, 12, and 16 harboring acetylations at the same time on two, three, or four residues (Figure 3). Acetylation can also co-occur with methylation, such as H3K9me1/2/3K14ac, H3K27acK36me2/3, and H3K36me1K37ac, while H3K36me1 can co-occur with another mono-methylation on H3K37 (Figure 3). The detection and functional characterization of some of these co-occurrent

patterns have been performed in several model species. For instance, in mouse, H3K9meK14ac has a role in gene silencing (Jurkowska et al., 2017). Accordingly, double replacements of K9 and K14 by arginine residues on histone H3 lead to the missing acetylations on these two locations and down-regulation of penicillin, sterigmatocystin, and orsellinic acid in *A. nidulans* (Nützmann et al., 2013). H3K18acK23ac was first reported to be acetylated by SAGA/ADA complex in *S. cerevisiae* (Grant et al., 1999) and then detected in human (X. Han et al., 2021). The investigation in *S. cerevisiae* of the combinatorial effect of acetylation on H4 lysines 5, 8, 12, and 16 showed that they are redundant and function together as a group to induce gene activation instead of forming distinct combinational patterns (Dion et al., 2005). Even though the co-occurrence of histone PTMs has not yet been systematically studied, our data as well as the results from other model systems, such as human or mouse, provide valuable clues for the possible functions of co-occurring histone PTMs in gene regulation in Aspergilli.

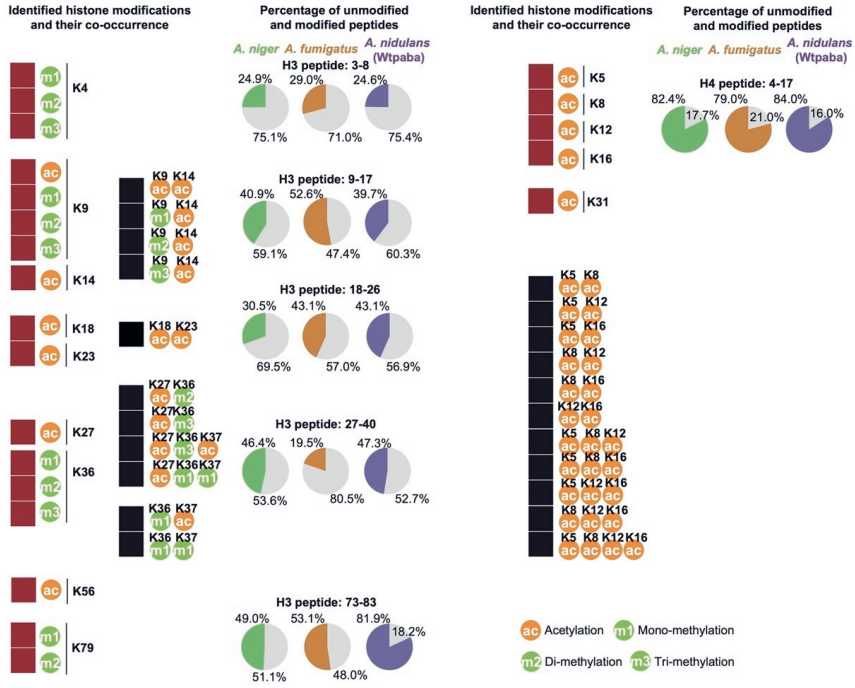


Figure 3 Identification of single and co-occurrent histone PTMs in Aspergilli. The red boxes show the identification of histone mono-, di-, and tri-methylation and acetylation while the black boxes show their co-occurrences in Aspergilli. Pie charts show the average relative abundance of unmodified peptides compared with modified peptides for each *Aspergillus* species. Unmodified peptides are shown in gray, and green, orange, as well as purple representing the proportion of modified peptides in *A. niger*, *A. fumigatus*, and *A. nidulans* (strain Wtpaba), respectively.

Although relative abundances only represent an indication of the absolute abundance of a modification as peptide detection by MS is affected by the peptide physicochemical properties and the digestion strategy (Lin et al., 2014), our data shows that the majority of detected histone H3 peptides is unmodified, while histone protein H4 exhibits a high percentage (79%-84%) of acetylated peptides in all three Aspergilli (Figure 3). According to a recent study, the newly synthesized histones are modified within 1.5 hours for histones H2A and H2B, and 3.5 hours for histones H3 and H4 (Flury et al., 2023). In our study, we cultured all Aspergilli strains at the same time and observed no difference in their growth among replicates. Different growth rates cannot explain the observed differences because one would expect the proportion of modified peptides to be consistent within a given species, which is not the case here. Indeed, the pattern we mentioned above is similar among the three species, with only a few observable differences. Fewer H3 peptides 18-26 are modified in *A. niger* (30% vs 43% in both *A. nidulans* and *A. fumigatus*), suggesting a lower abundance of H3K18ac and/or H3K23ac in this species. In contrast, *A. nidulans* exhibit more modified H3 peptides 73-83 (82% vs 49-53% in the other species), suggesting an enrichment in H3K79 methylation in this species. *A. fumigatus* differs from the two other species by having more modified H3 peptides 9-17 (53% vs 40-41%) and fewer modified H3 peptides 27-40 (19% vs 46-47%) (Figure 3). These differences suggest a higher abundance of H3K9 and/or H3K14 PTMs and a lower abundance of H3K27, H3K36, and/or H3K37 PTMs in *A. fumigatus*. These results collectively suggest that, although the same histone PTMs are detected in all three species, their abundances likely differ.

***Aspergillus* species differ in the relative abundance of histone PTMs**

To uncover quantitative differences between histone PTMs, we calculated the relative abundance of most histone PTMs in three *Aspergillus* species (*A. niger*, *A. fumigatus*, and *A. nidulans* (Wtpaba strain). Of note, a few modified peptides could not be quantified as their chromatographic peaks were of poor quality or could not be fully resolved for accurate quantification. We plotted the figure of Principal Component Analysis (PCA) (Figure 4A) and observed that *Aspergillus* species are well separated, suggesting that the three Aspergilli indeed differ quantitatively in their histone PTMs. These differences could be caused by differences in growth conditions (Figure 2B). However, we observed that even though *A. fumigatus* and *A. niger* were grown on the same medium compared with *A. nidulans*, *A. fumigatus* was more distinct from *A. nidulans* and *A. niger*. These distinctions suggest that the observed differences are not related to the different media but seem to be consistent with their taxonomic relationships. Specifically, *A. fumigatus* harbors significantly more H3K9 methylations, either alone or co-occurring with H3K14ac, and it has significantly less H3K36me1 compared to the other two Aspergilli (Figure 4B and Table S3). *A. fumigatus* also appears to have a trend of being enriched in H3K79me1 compared to *A. nidulans* (p-value: 0.080789) and *A. niger* (p-value: 0.165426) (Figures 3 and 4B and Table S3). In contrast, H3K9me1/2K14ac showed lower abundance in *A. niger* compared to the other two species, and H4K8ac was significantly higher in *A. fumigatus* compared to *A. nidulans*. There are also

trends indicating that *A. niger* contains fewer H3K18ac and H3K23ac while *A. nidulans* appears particularly rich in H3K79me2 and poor in H3K79me1 (Figures 3 and 4B). The observed differences suggest that *A. fumigatus* exhibit more heterochromatin- linked modifications compared to the other two species as it harbors more H3K9me1/2/3 linked with constitutive heterochromatin (Lachner et al., 2001) and H3K9meK14ac associated with triggering gene silencing (Jurkowska et al., 2017). For euchromatin, *A. fumigatus* uses more H3K79me1 and the combination of acetylations on diverse lysines of histone protein H4, while *A. niger* and *A. nidulans* harbors more H3K36me1 (Figure 4B). Even though we observed no qualitative difference in histone PTMs among three Aspergilli, these species differ quantitatively in specific histone PTM abundances, suggesting that the three populations (and their replicates) were very homogenous and in a similar state and they used a mix of marks associated with gene activation and gene silencing. Yet, the absence or low abundance of histone PTMs like H3K4ac, H3K14me, or H3K36ac could indicate their important roles in gene regulation in other developmental stages.

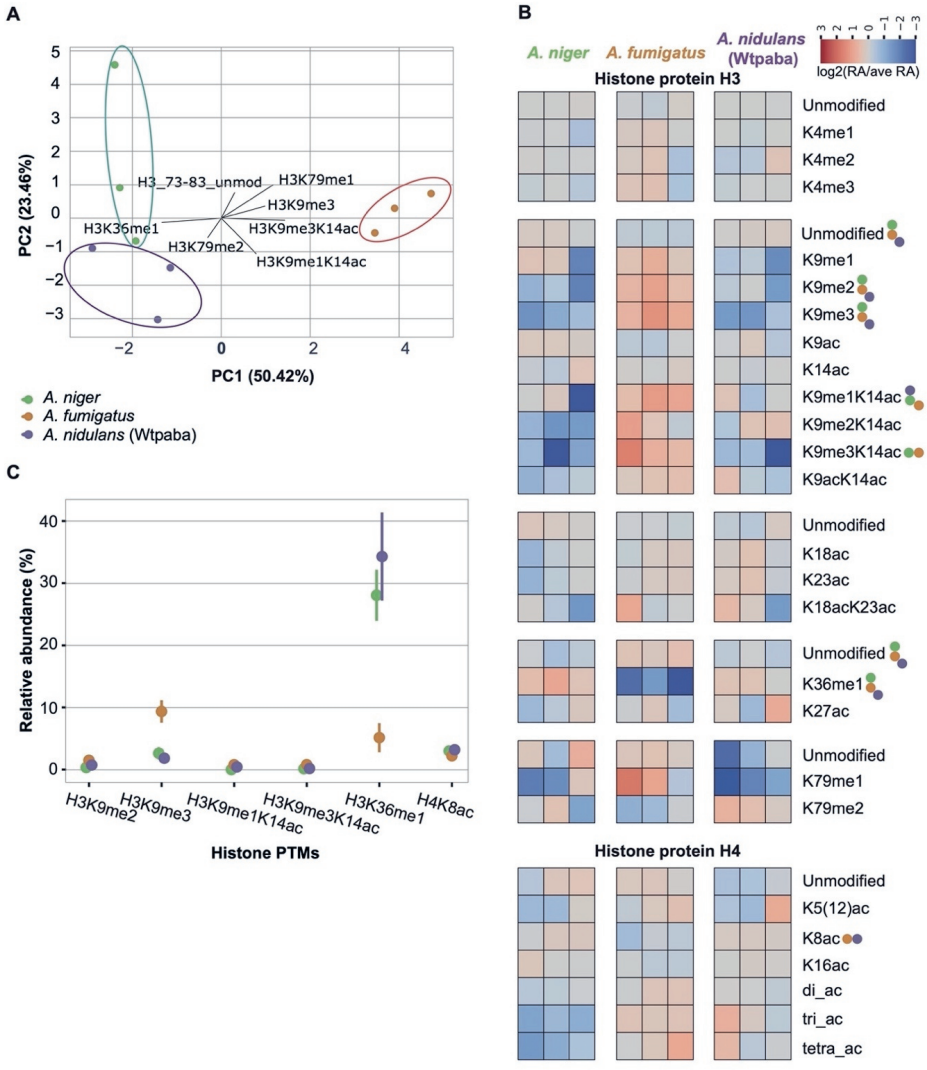


Figure 4 Quantification of histone PTMs in three *Aspergillus* species. (A) Principal component analysis (PCA) of three Aspergilli based on their relative abundance of histone PTMs (Table S3). (B) Quantitative comparison of all histone PTMs detected by MS. The heatmap displays red or blue color boxes according to the $\log_2$(relative abundance/average of relative abundance) of different peptides in Aspergilli. 'Di-ac' indicates H4K5acK8ac, H4K5acK12ac, H4K5acK16ac, H4K8acK16ac, and H4K12acK16ac; 'tri-ac' means H4K5acK8ac12ac, H4K5acK8acK16ac, H4K5acK12acK16ac, and H4K8acK12acK16ac; and 'tetra-ac' indicates H4K5acK8ack12acK16ac. For each histone PTM between the three Aspergilli species, we applied an ANOVA test and Tukey HSD tests to determine significant differences. We used p-value = 0.05 as the threshold for a significant difference in the ANOVA test and p-value = 0.025 in the Tukey HSD test. Colored dots

indicate significant differences between represented species. (C) Relative abundance of histone PTMs that are significantly different. Points show the average relative abundance and lines indicate standard deviation.

***A. nidulans* strains differ quantitatively in euchromatin marks**

To evaluate the intra-species variation in the relative abundance of histone PTMs, we compared the MS data of another *A. nidulans* strain (TN02A25) to Wtpaba (Figure 5). In general, we observed that TN02A25 spurs the same type of histone PTMs as we observed in the other strains (Figure 4B and 5C), indicating all Aspergilli tested in this study used the same histone code. However, the two strains show a few differences, already when only considering the percentage of unmodified versus modified peptides. The vast majority of detected histone H3 peptides were unmodified (> 50%) and histone H4 peptides were highly acetylated in TN02A25 like the other strains (Figures 4 and 5A). However, TN02A25 differs from Wtpaba in H3 peptides 27-40 (11% vs 47% modified peptides in Wtpaba) and 73-83 (46% vs 82% modified peptides in Wtpaba). These proportions in TN02A25 are similar to those found in *A. fumigatus* or *A. niger* and *A. fumigatus*, respectively (Figure 3), and suggest that the two *A. nidulans* strains differ in histone PTM abundance on those peptides.

The quantitative comparison of histone PTMs between the two *A. nidulans* strains shows good separation in the PCA analysis (Figure 5B). Consistent with previous results and analyses, TN02A25 exhibit significantly less H3K36me1 compared to the Wtpaba strain. We also observed a trend of enrichment in H3K79me2 and lower abundance in H3K79me1 in Wtpaba, which contributes to the difference between both strains (Figure 5). In addition, H3K4me2, H3K9ac, H3K36me1, and H4di-ac (H4K5acK8ac, H4K5acK12ac, H4K5acK16ac, H4K8acK16ac, and H4K12acK16ac) exhibited significantly higher abundance in Wtpaba than TN02A25 (Figure 5C-D, Table S4) and they are all related to regulation of euchromatin.

Moreover, we observed that intra-species variability can be as high as the inter-species variability as the histone PTM abundances of Wtpaba and TN02A25 are more related to *A. niger* and *A. fumigatus,* respectively (Figure S3A and Table S5). Four single or co-occurrent histone PTMs, H3K36me1, H3K79me1, H3K9me3K14ac, and H3K9me1K14ac displayed the most distinct abundant patterns among the four strains, followed by H3K9me1, H3K9me2, and H3K18acK23ac being less abundant in *A. niger* compared to the other three strains (Figure S3B). These marks showing distinct relative abundance agreed with the marks we pinpointed from both species and strain comparisons. Thus, the observed differences between species are likely a ‘strain effect’ and highlight the importance of including several strains in the study of histone PTMs.

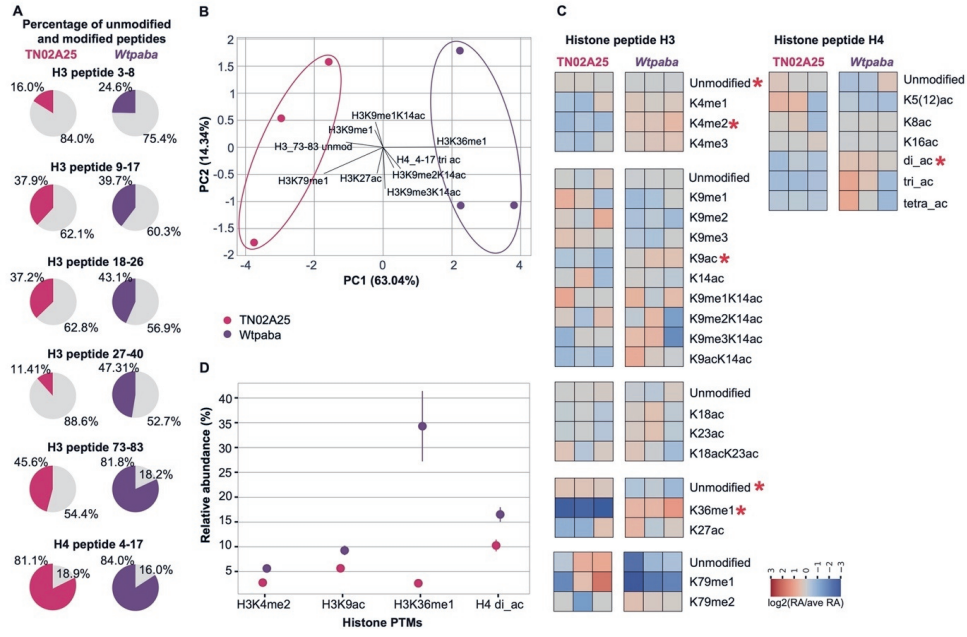


Figure 5 Quantification of histone modifications in two *A. nidulans* strains. (A) The average relative abundance of unmodified peptides compared with modified peptides. Grey represents the unmodified peptide percentage in the pie chart, while pink and purple represent the modified peptide percentage of TN02A25 and Wtpaba *A. nidulans* strains, respectively. (B) Principal component analysis (PCA) of two strains based on their relative abundance of histone PTMs. (C) Quantitative comparison of all histone PTMs detected by MS. The heatmap displays red or blue color boxes according to the $\log_2$(relative abundance/average of relative abundance) of different peptides in two *A. nidulans* strains. 'Di-ac' indicates H4K5acK8ac, H4K5acK12ac, H4K5acK16ac, H4K8acK16ac, and H4K12acK16ac; 'tri-ac' means H4K5acK8ac12ac, H4K5acK8acK16ac, H4K5acK12acK16ac, and H4K8acK12acK16ac; and 'tetra-ac' indicates H4K5acK8ack12acK16ac. ANOVA test and Tukey HSD test were performed to show the significance of the difference of one histone modification between two strains when p-value < 0.05 in the ANOVA test and p-value < 0.025 in the Tukey HSD test. Red stars indicate significant differences. (D) The relative abundance of the top four histone modifications that are significantly different. Points are average relative abundance and lines are standard deviations.

The use of MS spectrometry is powerful to detect and quantify histone PTMs in filamentous fungi and this approach is needed to understand the chromatin status at different developmental stages. The present study focused on fungal mycelium cultured for 16 hours in liquid medium, a particular stage that requires the expression of a specific set of genes and is expected to not favor the production of SMs. Similar future studies in Aspergilli at different developmental stages and different culture conditions will fully uncover quantitative variations in histone PTMs and might lead to identifying key PTMs for certain applications, such as the

controlled activation of SM production. Finally, while this work has revealed quantitative differences between species and strains, it remains to determine whether they are due to different PTM abundance at the same loci, or whether they reflect a different number of loci covered by these PTMs, which would reflect activation or repression of different sets of genes. These future studies will contribute to understanding how single and co-occurring histone PTMs regulate gene expression in Aspergilli and more broadly in other filamentous fungi.

Materials and methods

Fungal strains and culture

A. niger (wild type strain: NRRL3, CBS 120.49) and *A. fumigatus* (wild type strain: Af293, CBS 126847) were grown on MEA (Malt Extract Agar; Malt Extract Agar Oxoid 50 g/L) medium at 30 °C and 25 °C, respectively, while *A. nidulans* (strains: Wtpaba (genotype: pabaA1) and TN02A25 (genotype: pyrG89 argB2 pabaB22 riboB2), kindly provided by Prof. Joseph Strauss) was grown on YAG (Yeast Agar Glucose; 5 g/L yeast extract, 15 g/L agar, 20 g/L D-glucose, 1 mL/L trace elements (Solution A: dissolve 10 g EDTA and 1 g $FeSO_4 \times 7H_2O$ in 80 mL water and adjust pH to 5.5; solution B: dissolve 4.4 g $ZnSO_4 \times 7H_2O$, 1 g $MnCl_2 \times 4H_2O$, 0.32 g $CuSO_4 \times 6H_2O$, and 0.22 g $(NH_4)_6Mo_7O_{24} \times 4H_2O$ in 80mL water; combine solutions A and B, adjust pH to 6.5) medium at 37 °C.

Fungal spores of the four Aspergilli strains were harvested from 4- or 5-day-old sporulating plates with a spreader in 12.5 mL ice-cold ACES (1.822 g/L N-(2-Acetamido)-2-aminoethanesulfonic acid, 0.2 mL/L Tween 80) buffer at the same time. The spore suspension was centrifuged at 2,500 xg at 4 °C for 5 minutes. The supernatant was discarded, and spores were resuspended in 25 mL ice-cold ACES buffer. The spore suspension was centrifuged at 2,500 xg at 4 °C for 5 minutes, and the supernatant was discarded. Spores were resuspended in 10 mL ice-cold ACES buffer, and spore concentration was determined with a hemocytometer (Marienfeld, 0.0025mm^2) or spore counter, and adjusted to 10^7 spores/ mL. For the overnight culturing, 10^7 to 10^8 spores were inoculated in 50 mL liquid medium in a 250 mL flask and incubated at the same temperature as plate culture, 200 rpm, for 16 hours. Mycelium were centrifugated at 4,000 xg for 10 minutes and the supernatant was discarded. The mycelium pellet was washed with 25 mL 0.8 M NaCl, followed by centrifugation at 4,000 xg for 10 minutes to collect them for histone protein extraction. No obvious difference in regard to growth was observed.

Histone protein extraction

Histone proteins were extracted according to an adapted histone protein extraction protocol (X. Zhang et al., 2021). Briefly, the fungal material was freeze-dried, ground with a mortar and pestle, and further disrupted by TissueLyser II (QIAGEN, C.0659). 100 mg ground sample was first resuspended in 1 mL lysis buffer (1×PBS, 0.5 mM PMSF, 5 µM Leupeptin, 5 µM Aprotinin, and 5mM Na-butyrate) with high (10%) Triton X-100 concentration and incubated

on an orbital shaker at 4 °C for half an hour. Then 10 mL lysis buffer with low (1%) Triton X-100 concentration was added and the sample was incubated for another half an hour. The samples were subsequently sonicated for 30 seconds with 10 repeats and 30-second break in between each repeat. The crude histone proteins were dissolved in 0.4 N H_2SO_4 and then precipitated in four volumes of acetone. Protein concentration was measured with the Bradford Protein Assay (He, 2011).

LC-MS/MS analysis of histone PTMs

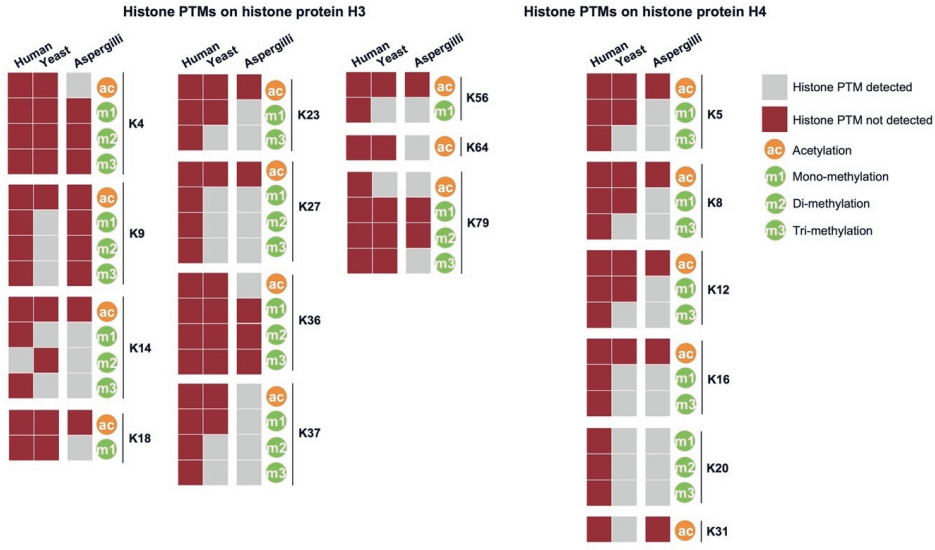
Samples were separated on a 17% SDS-PAGE gel and histone bands were excised and in-gel digested using a double derivatization protocol that involves propionylation of lysines and N-terminal derivatization with phenyl isocyanate (Noberini et al., 2021). Chemical propionylation of lysines (which occurs on unmodified or mono-methylated residues) impairs trypsin cleavage, resulting in proteolytic cleavage at arginine residues only and the generation of histone peptides of proper length for MS analysis. Peptide mixtures were separated by reversed-phase liquid chromatography and analyzed by MS on an Orbitrap instrument, as previously described (X. Zhang et al., 2022).

The acquired RAW data were analyzed using the integrated MaxQuant software v.1.6.10. The Uniprot UP000000560, UP000006706, and UP000002530 databases were used for the identification of *A. nidulans, A. niger,* and *A. fumigatus* histone peptides, respectively. Enzyme specificity was set to Arg-C. The estimated false discovery rate was set at a maximum of 1% at both peptide and protein levels. The mass tolerance was set to 4.5 ppm for precursor and fragment ions. Two missed cleavages were allowed, and the minimum peptide length was set to four amino acids. The second peptide search was enabled. Variable modifications include lysine propionylation (+56.026215 Da), monomethylation-propionylation (+70.041865 Da), dimethylation (+28.031300 Da), trimethylation (+42.046950 Da), acetylation (+42.010565 Da), malonylation (+86.000394 Da), and succinylation (+100.016044 Da). N-terminal PIC labeling (+119.037114 Da) was set as a fixed modification (Noberini et al., 2021). To reduce the search time and the rate of false positives, which increases with a higher number of variable modifications included in the database search, the raw data were analyzed through multiple parallel MaxQuant jobs (Bremang et al., 2013), setting different combinations of variable modifications. Peptides identified by MaxQuant with Andromeda score higher than 50 and localization probability score higher than 0.75 were quantitated, either manually or by using a version of the EpiProfile 2.0 software (Yuan et al., 2018) adapted to the analysis of histones from *Aspergillus* strains. Identifications and retention times were used to guide the manual quantification of each modified peptide using QualBrowser version 2.0.7 (Thermo Scientific). Site assignment was evaluated from MS2 spectra using QualBrowser and MaxQuant Viewer. Extracted ion chromatograms were constructed for each doubly charged precursor, based on its *m/z* value, using a mass tolerance of 10 ppm. For each modified histone peptide, the

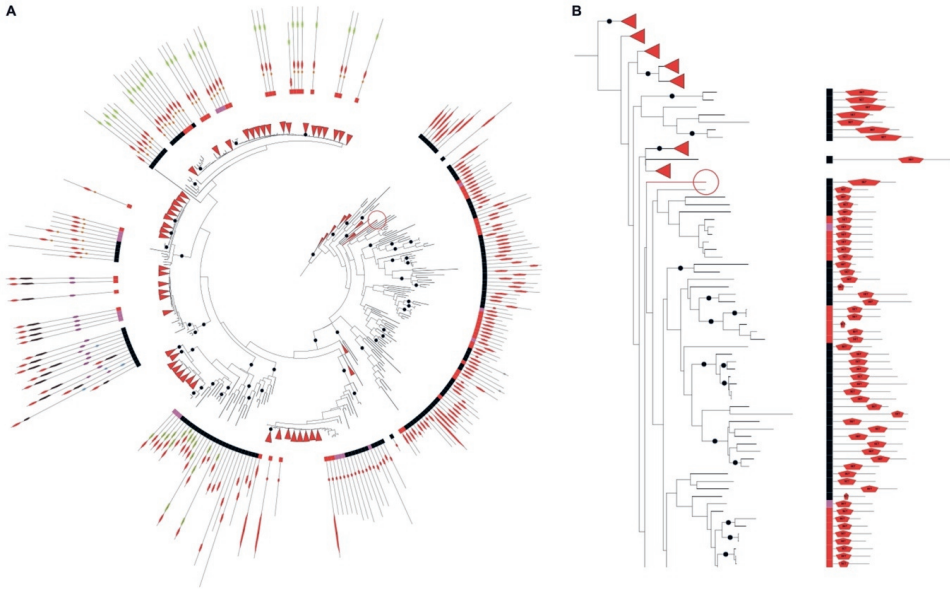
relative abundance was estimated by dividing the area under the curve (AUC) of each modified peptide by the sum of the areas corresponding to all the observed forms of that peptide (Pesavento et al., 2006). The MS proteomics data have been deposited to the ProteomeXchange Consortium (Vizcaíno et al., 2014) *via* the PRIDE partner repository with the dataset identifier PXD033478.

PCA figures were generated by sklearn.decomposition.PCA of Sklearn (Pedregosa et al., 2011) in Python, heatmap and point plots were made using Seaborn (Waskom, 2021) in Python. To compare the relative abundance of each histone PTM among different species/strains, we first used 'bioinfokit.stat.oanova' to perform the one-way ANOVA test, then used 'bioinfokit.analys.stat.tukey_hsd' to perform the post-hoc Tukey HSD (Honestly Significant Difference) test in python (Bedre, 2020). We used α = 0.05 and $\alpha/2$ = 0.025 as thresholds for a significant difference in the ANOVA and Tukey HSD tests, respectively.

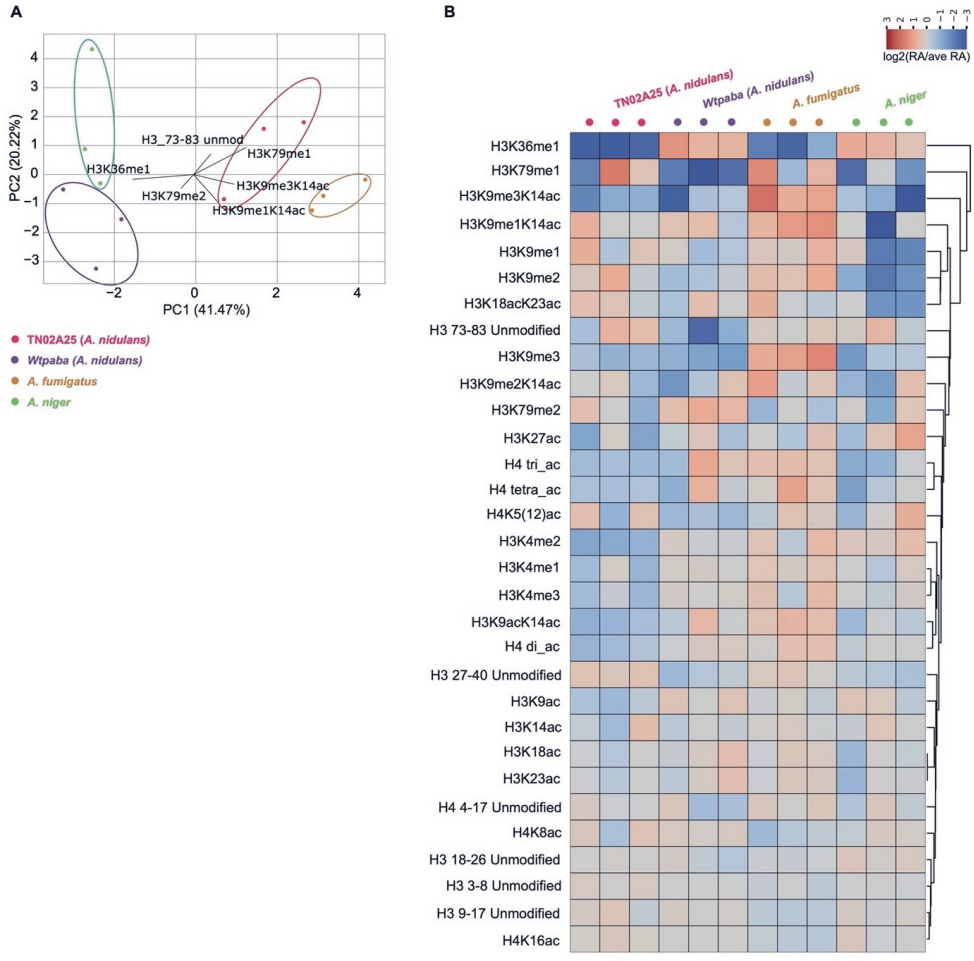
Supplementary materials



Supplementary Figure 1 Comparison of experimentally detected histone PTMs in human, yeast, and Aspergilli. The summary of histone PTMS in human and yeast is adapted from a recently published paper (Grau-Bové, Xavier, et al., Nature Ecology & Evolution, 2022) with a few corrections: 1) H3K64ac can be detected in human (Pradeepa, Madapura M., et al., Nature genetics, 2016); 2) H3K9me1 is lost in yeast (Zhang, Xing, et al., Cell, 2002 and Marina, Diana B., et al., Genes & development, 2013); 3) H4K5me1 (Green, Erin M., et al., Nature structural & molecular biology, 2012) and 4) H4K8ac (Magraner-Pardo, Lorena, et al., BMC genomics, 2014) can be detected in yeast. The detection of histone PTMs in Aspergilli comes from this study.



Supplementary Figure 2 SET5 was lost in Aspergilli. (A) The phylogeny of the SET domain-containing proteins in 94 Aspergilli and 15 outgroup species (Zhang X, et al., Microbial Genomics, 2022). The red circle indicates SET5 in *S. cerevisiae*. (B) Zoom-in of *S. cerevisiae* SET5 and its homologs. Black, pink, and red boxes indicate outgroup species, *Penicillium*, and Aspergilli, respectively.



Supplementary Figure 3 Distinct abundant patterns of several histone PTMs differ in different *Aspergillus* species/strains. (A) Principal component analysis (PCA) according to log2(relative abundance/average of relative abundance) uncovers the separation of four strains based on their histone PTMs. (B) Quantitative comparison of all histone PTMs detected via mass spectrometry. The heatmap displays red or blue color boxes according to the log2(relative abundance/average of relative abundance) of different peptides in four Aspergilli. 'Di-ac' indicates H4K5acK8ac, H4K5acK12ac, H4K5acK16ac, H4K8acK16ac, and H4K12acK16ac; 'tri-ac' means H4K5acK8ac12ac, H4K5acK8acK16ac, H4K5acK12acK16ac, and H4K8acK12acK16ac; and 'tetra-ac' indicates H4K5acK8ack12acK16ac.

Supplementary Table 1 Histone peptides identified via mass spectrometry

Histone type	Modification	Peptide	Sequence	Modified sequence	Proteins	Score	Strains
Histone H2A	unmodified	Histone H2A 73-78	DNKKTR	_DNK(pr)K(pr)TR_	Q4WWC6	114.19	*A. fumigatus*
Histone H2A	unmodified	Histone H2A 73-78	DNKKTR	_DNK(pr)K(pr)TR_	P08844	125.62	*A. nidulans*
Histone H2A	unmodified	Histone H2A 73-78	DNKKTR	_DNK(pr)K(pr)TR_	P0C953	132.96	*A. niger*
Histone H2A	unmodified	Histone H2A 37-43	KGNYAQR	_K(pr)GNYAQR_	Q4WWC6	138.38	*A. fumigatus*
Histone H2A	unmodified	Histone H2A 37-43	KGNYAQR	_K(pr)GNYAQR_	P08844	132.72	*A. nidulans*
Histone H2A	unmodified	Histone H2A 37-43	KGNYAQR	_K(pr)GNYAQR_	P0C953	121.87	*A. niger*
Histone H2A	unmodified	Histone H2A 19-30	SSKAGLAFPVGR	_SSK(pr)AGLAFPVGR_	Q4WWC6	173.32	*A. fumigatus*
Histone H2A	unmodified	Histone H2A 19-30	SSKAGLAFPVGR	_SSK(pr)AGLAFPVGR_	P08844	163.64	*A. nidulans*
Histone H2A	unmodified	Histone H2A 19-30	SSKAGLAFPVGR	_SSK(pr)AGLAFPVGR_	P0C953	161.63	*A. niger*
Histone H2A, H2A.Z	unmodified	Histone H2A.Z 91-97	HLQLAIR	_HLQLAIR_	Q4WWC6;Q4WE68	104.94	*A. fumigatus*
Histone H2A, H2A.Z	unmodified	Histone H2A.Z 91-97	HLQLAIR	_HLQLAIR_	P08844;Q5AUJ1	135.5	*A. nidulans*
Histone H2A, H2A.Z	unmodified	Histone H2A.Z 91-97	HLQLAIR	_HLQLAIR_	P0C953;A2R702	104.94	*A. niger*
Histone H2A.Z	unmodified	Histone H2A.Z 108-118	ATIAFGGVLPR	_ATIAFGGVLPR_	Q4WE68	133.23	*A. fumigatus*
Histone H2A.Z	unmodified	Histone H2A.Z 108-118	ATIAFGGVLPR	_ATIAFGGVLPR_	Q5AUJ1	141.2	*A. nidulans*
Histone H2A.Z	unmodified	Histone H2A.Z 108-118	ATIAFGGVLPR	_ATIAFGGVLPR_	A2R702	141.2	*A. niger*
Histone H2A.Z	unmodified	Histone H2A.Z 98-107	GDEELDTLIR	_GDEELDTLIR_	Q4WE68	147.51	*A. fumigatus*
Histone H2A.Z	unmodified	Histone H2A.Z 98-107	GDEELDTLIR	_GDEELDTLIR_	A2R702	92.051	*A. niger*
Histone H2B	unmodified	Histone H2B 106-	EIQTSVR	_EIQTSVR_	Q4WWC5	132.72	*A. fumigatus*
Histone H2B	unmodified	Histone H2B 106-	EIQTSVR	_EIQTSVR_	P23754	125.86	*A. nidulans*
Histone H2B	unmodified	Histone H2B 106-	EIQTSVR	_EIQTSVR_	A2QY49	125.86	*A. niger*
Histone H2B	unmodified	Histone H2B 47-70	KETYSSYIYKVLKQVHPDTGI	_K(pr)ETYSSYIYK(pr)VLK(pr)QVHPDTGIST	Q4WWC5	158.69	*A. fumigatus*
Histone H2B	unmodified	Histone H2B 113-	LILPGELAKHAVSEGTKAVTK	_LILPGELAK(pr)HAVSEGTK(pr)AVTK(pr)YS	Q4WWC5	71.86	*A. fumigatus*
Histone H2B	H2BK92ac	Histone H2B 86-105	VATEASKLAAYNKKSTISSR	_VATEASK(ac)LAAYNK(pr)K(pr)STISSR_	Q4WWC5	56.719	*A. fumigatus*
Histone H2B	H2BK98me1K99	Histone H2B 86-105	VATEASKLAAYNKKSTISSR	_VATEASK(pr)LAAYNK(pm)K(pm)STISSR_	Q4WWC5	70.54	*A. fumigatus*
Histone H2B	unmodified	Histone H2B 86-105	VATEASKLAAYNKKSTISSR	_VATEASK(pr)LAAYNK(pr)K(pr)STISSR_	Q4WWC5	346.63	*A. fumigatus*
Histone H2B	unmodified	Histone H2B 86-105	VATEASKLAAYNKKSTISSR	_VATEASK(pr)LAAYNK(pr)K(pr)STISSR_	P23754	311.64	*A. nidulans*
Histone H2B	unmodified	Histone H2B 86-105	VATEASKLAAYNKKSTISSR	_VATEASK(pr)LAAYNK(pr)K(pr)STISSR_	A2QY49	333.47	*A. niger*
Histone H3	H3K79me2	Histone H3 73-83	EIAQDFKSDLR	_EIAQDFK(2m)SDLR_	P61832	162.36	*A. fumigatus*

Histone type	Modification	Peptide	Sequence	Modified sequence	Proteins	Score	Strains
Histone H3	H3K79me2	Histone H3 73-83	EIAQDFKSDLR	_EIAQDFK(2m)SDLR_	P23753	142.83	*A. nidulans*
Histone H3	H3K79me2	Histone H3 73-83	EIAQDFKSDLR	_EIAQDFK(2m)SDLR_	A2QRR5	139.43	*A. niger*
Histone H3	H3K79me1	Histone H3 73-83	EIAQDFKSDLR	_EIAQDFK(pm)SDLR_	P61832	150.85	*A. fumigatus*
Histone H3	H3K79me1	Histone H3 73-83	EIAQDFKSDLR	_EIAQDFK(pm)SDLR_	P23753	94.309	*A. nidulans*
Histone H3	H3K79me1	Histone H3 73-83	EIAQDFKSDLR	_EIAQDFK(pm)SDLR_	A2QRR5	132.91	*A. niger*
Histone H3	unmodified	Histone H3 73-83	EIAQDFKSDLR	_EIAQDFK(pr)SDLR_	P61832	170.47	*A. fumigatus*
Histone H3	unmodified	Histone H3 73-83	EIAQDFKSDLR	_EIAQDFK(pr)SDLR_	P23753	190.88	*A. nidulans*
Histone H3	unmodified	Histone H3 73-83	EIAQDFKSDLR	_EIAQDFK(pr)SDLR_	A2QRR5	123.63	*A. niger*
Histone H3	H3K27acK36me	Histone H3 27-40	KAAPSTGGVKKPHR	_K(ac)AAPSTGGVK(2m)K(pr)PHR_	P61832	176.42	*A. fumigatus*
Histone H3	H3K27acK36me	Histone H3 27-40	KAAPSTGGVKKPHR	_K(ac)AAPSTGGVK(2m)K(pr)PHR_	P23753	119.45	*A. nidulans*
Histone H3	H3K27acK36me 2	Histone H3 27-40	KAAPSTGGVKKPHR	_K(ac)AAPSTGGVK(2m)K(pr)PHR_	A2QRR5	126.95	*A. niger*
Histone H3	3K37ac	Histone H3 27-40	KAAPSTGGVKKPHR	_K(ac)AAPSTGGVK(3m)K(ac)PHR_	A2QRR5	63.408	*A. niger*
Histone H3	H3K27acK36me 3	Histone H3 27-40	KAAPSTGGVKKPHR	_K(ac)AAPSTGGVK(3m)K(pr)PHR_	P61832	74.841	*A. fumigatus*
Histone H3	H3K27acK36me 1K37me1	Histone H3 27-40	KAAPSTGGVKKPHR	_K(ac)AAPSTGGVK(pm)K(pm)PHR_	P61832	67.275	*A. fumigatus*
Histone H3	H3K27ac	Histone H3 27-40	KAAPSTGGVKKPHR	_K(ac)AAPSTGGVK(pr)K(pr)PHR_	P61832	172.8	*A. fumigatus*
Histone H3	H3K27ac	Histone H3 27-40	KAAPSTGGVKKPHR	_K(ac)AAPSTGGVK(pr)K(pr)PHR_	P23753	126.19	*A. nidulans*
Histone H3	H3K27ac	Histone H3 27-40	KAAPSTGGVKKPHR	_K(ac)AAPSTGGVK(pr)K(pr)PHR_	A2QRR5	135.66	*A. niger*
Histone H3	H3K36me2	Histone H3 27-40	KAAPSTGGVKKPHR	_K(pr)AAPSTGGVK(2m)K(pr)PHR_	P61832	144.65	*A. fumigatus*
Histone H3	H3K36me2	Histone H3 27-40	KAAPSTGGVKKPHR	_K(pr)AAPSTGGVK(2m)K(pr)PHR_	P23753	148.28	*A. nidulans*
Histone H3	H3K36me2	Histone H3 27-40	KAAPSTGGVKKPHR	_K(pr)AAPSTGGVK(2m)K(pr)PHR_	A2QRR5	118.9	*A. niger*
Histone H3	H3K36me3	Histone H3 27-40	KAAPSTGGVKKPHR	_K(pr)AAPSTGGVK(3m)K(pr)PHR_	P61832	100.88	*A. fumigatus*
Histone H3	H3K36me3	Histone H3 27-40	KAAPSTGGVKKPHR	_K(pr)AAPSTGGVK(3m)K(pr)PHR_	P23753	126.71	*A. nidulans*
Histone H3	H3K36me3	Histone H3 27-40	KAAPSTGGVKKPHR	_K(pr)AAPSTGGVK(3m)K(pr)PHR_	A2QRR5	69.979	*A. niger*
Histone H3	H3K36me1K37a c	Histone H3 27-40	KAAPSTGGVKKPHR	_K(pr)AAPSTGGVK(pm)K(ac)PHR_	P23753	104.17	*A. nidulans*
Histone H3	H3K36me1K37m e1	Histone H3 27-40	KAAPSTGGVKKPHR	_K(pr)AAPSTGGVK(pm)K(pm)PHR_	P61832	94.297	*A. fumigatus*
Histone H3	H3K36me1K37m e1	Histone H3 27-40	KAAPSTGGVKKPHR	_K(pr)AAPSTGGVK(pm)K(pm)PHR_	P23753	81.865	*A. nidulans*
Histone H3	H3K36me1K37m e1	Histone H3 27-40	KAAPSTGGVKKPHR	_K(pr)AAPSTGGVK(pm)K(pm)PHR_	A2QRR5	83.862	*A. niger*
Histone H3	H3K36me1	Histone H3 27-40	KAAPSTGGVKKPHR	_K(pr)AAPSTGGVK(pm)K(pr)PHR_	P61832	162.38	*A. fumigatus*
Histone H3	H3K36me1	Histone H3 27-40	KAAPSTGGVKKPHR	_K(pr)AAPSTGGVK(pm)K(pr)PHR_	P23753	157.38	*A. nidulans*
Histone H3	H3K36me1	Histone H3 27-40	KAAPSTGGVKKPHR	_K(pr)AAPSTGGVK(pm)K(pr)PHR_	A2QRR5	150.94	*A. niger*

Histone type	Modification	Peptide	Sequence	Modified sequence	Proteins	Score	Strains
Histone H3	unmodified	Histone H3 27-40	KAAPSTGGVKKPHR	_K(pr)AAPSTGGVK(pr)K(pr)PHR_	P61832	185.96	*A. fumigatus*
Histone H3	unmodified	Histone H3 27-40	KAAPSTGGVKKPHR	_K(pr)AAPSTGGVK(pr)K(pr)PHR_	P23753	123.8	*A. nidulans*
Histone H3	unmodified	Histone H3 27-40	KAAPSTGGVKKPHR	_K(pr)AAPSTGGVK(pr)K(pr)PHR_	A2QRR5	137.89	*A. niger*
Histone H3	unmodified	Histone H3 64-69	KLPFQR	_K(pr)LPFQR_	P61832	109.94	*A. fumigatus*
Histone H3	unmodified	Histone H3 64-69	KLPFQR	_K(pr)LPFQR_	P23753	103.41	*A. nidulans*
Histone H3	unmodified	Histone H3 64-69	KLPFQR	_K(pr)LPFQR_	A2QRR5	113.82	*A. niger*
Histone H3	H3K18acK23ac	Histone H3 18-26	KQLASKAAR	_K(ac)QLASK(ac)AAR_	P61832	138.1	*A. fumigatus*
Histone H3	H3K18acK23ac	Histone H3 18-26	KQLASKAAR	_K(ac)QLASK(ac)AAR_	P23753	95.502	*A. nidulans*
Histone H3	H3K18acK23ac	Histone H3 18-26	KQLASKAAR	_K(ac)QLASK(ac)AAR_	A2QRR5	98.033	*A. niger*
Histone H3	H3K18ac	Histone H3 18-26	KQLASKAAR	_K(ac)QLASK(pr)AAR_	P61832	138.1	*A. fumigatus*
Histone H3	H3K23ac	Histone H3 18-26	KQLASKAAR	_K(pr)QLASK(ac)AAR_	P61832	135.83	*A. fumigatus*
Histone H3	H3K23ac	Histone H3 18-26	KQLASKAAR	_K(pr)QLASK(ac)AAR_	P23753	122.52	*A. nidulans*
Histone H3	unmodified	Histone H3 18-26	KQLASKAAR	_K(pr)QLASK(pr)AAR_	P61832	150.4	*A. fumigatus*
Histone H3	unmodified	Histone H3 18-26	KQLASKAAR	_K(pr)QLASK(pr)AAR_	P23753	138.54	*A. nidulans*
Histone H3	unmodified	Histone H3 18-26	KQLASKAAR	_K(pr)QLASK(pr)AAR_	A2QRR5	143.01	*A. niger*
Histone H3	H3K9me2K14ac	Histone H3 9-17	KSTGGKAPR	_K(2m)STGGK(ac)APR_	P61832	125.16	*A. fumigatus*
Histone H3	H3K9me2	Histone H3 9-17	KSTGGKAPR	_K(2m)STGGK(pr)APR_	P61832	62.005	*A. fumigatus*
Histone H3	H3K9me3K14ac	Histone H3 9-17	KSTGGKAPR	_K(3m)STGGK(ac)APR_	P61832	105.2	*A. fumigatus*
Histone H3	H3K9me3	Histone H3 9-17	KSTGGKAPR	_K(3m)STGGK(pr)APR_	P61832	121.05	*A. fumigatus*
Histone H3	H3K9me3	Histone H3 9-17	KSTGGKAPR	_K(3m)STGGK(pr)APR_	P23753	101.64	*A. nidulans*
Histone H3	H3K9me3	Histone H3 9-17	KSTGGKAPR	_K(3m)STGGK(pr)APR_	A2QRR5	105.57	*A. niger*
Histone H3	H3K9acK14ac	Histone H3 9-17	KSTGGKAPR	_K(ac)STGGK(ac)APR_	P61832	111.95	*A. fumigatus*
Histone H3	H3K9acK14ac	Histone H3 9-17	KSTGGKAPR	_K(ac)STGGK(ac)APR_	P23753	113.26	*A. nidulans*
Histone H3	H3K9acK14ac	Histone H3 9-17	KSTGGKAPR	_K(ac)STGGK(ac)APR_	A2QRR5	114.89	*A. niger*
Histone H3	H3K9ac	Histone H3 9-17	KSTGGKAPR	_K(ac)STGGK(pr)APR_	P61832	131.06	*A. fumigatus*
Histone H3	H3K9ac	Histone H3 9-17	KSTGGKAPR	_K(ac)STGGK(pr)APR_	P23753	117.04	*A. nidulans*
Histone H3	H3K9ac	Histone H3 9-17	KSTGGKAPR	_K(ac)STGGK(pr)APR_	A2QRR5	119.41	*A. niger*
Histone H3	H3K9me1K14ac	Histone H3 9-17	KSTGGKAPR	_K(pm)STGGK(ac)APR_	P61832	84.743	*A. fumigatus*
Histone H3	H3K9me1K14ac	Histone H3 9-17	KSTGGKAPR	_K(pm)STGGK(ac)APR_	P23753	60.59	*A. nidulans*
Histone H3	H3K9me1	Histone H3 9-17	KSTGGKAPR	_K(pm)STGGK(pr)APR_	P61832	109.86	*A. fumigatus*
Histone H3	H3K14ac	Histone H3 9-17	KSTGGKAPR	_K(pr)STGGK(ac)APR_	P61832	141.39	*A. fumigatus*
Histone H3	H3K14ac	Histone H3 9-17	KSTGGKAPR	_K(pr)STGGK(ac)APR_	P23753	128.61	*A. nidulans*
Histone H3	H3K14ac	Histone H3 9-17	KSTGGKAPR	_K(pr)STGGK(ac)APR_	A2QRR5	127.56	*A. niger*
Histone H3	unmodified	Histone H3 9-17	KSTGGKAPR	_K(pr)STGGK(pr)APR_	P61832	129.47	*A. fumigatus*
Histone H3	unmodified	Histone H3 9-17	KSTGGKAPR	_K(pr)STGGK(pr)APR_	P23753	129.47	*A. nidulans*
Histone H3	unmodified	Histone H3 9-17	KSTGGKAPR	_K(pr)STGGK(pr)APR_	A2QRR5	130.27	*A. niger*
Histone H3	H3K4me1	Histone H3 3-8	TKQTAR	_TK(pm)QTAR_	P23753	46.8	*A. nidulans*
Histone H3	H3K4me2	Histone H3 3-8	TKQTAR	_TK(2m)QTAR_	P61832	116.3	*A. fumigatus*

Histone type	Modification	Peptide	Sequence	Modified sequence	Proteins	Score	Strains
Histone H3	H3K4me2	Histone H3 3-8	TKQTAR	_TK(2m)QTAR_	A2QRR5	111.65	*A. niger*
Histone H3	H3K4me3	Histone H3 3-8	TKQTAR	_TK(3m)QTAR_	P61832	71.715	*A. fumigatus*
Histone H3	unmodified	Histone H3 3-8	TKQTAR	_TK(pr)QTAR_	P61832	110.54	*A. fumigatus*
Histone H3	unmodified	Histone H3 3-8	TKQTAR	_TK(pr)QTAR_	P23753	107.55	*A. nidulans*
Histone H3	unmodified	Histone H3 3-8	TKQTAR	_TK(pr)QTAR_	A2QRR5	113.08	*A. niger*
Histone H3	unmodified	Histone H3 117-128	VTIQSKDIQLAR	_VTIQSK(pr)DIQLAR_	P61832	205.14	*A. fumigatus*
Histone H3	unmodified	Histone H3 117-128	VTIQSKDIQLAR	_VTIQSK(pr)DIQLAR_	P23753	215.06	*A. nidulans*
Histone H3	unmodified	Histone H3 117-128	VTIQSKDIQLAR	_VTIQSK(pr)DIQLAR_	A2QRR5	180.67	*A. niger*
Histone H3	unmodified	Histone H3 41-49	YKPGTVALR	_YK(pr)PGTVALR_	P61832	119.27	*A. fumigatus*
Histone H3	unmodified	Histone H3 41-49	YKPGTVALR	_YK(pr)PGTVALR_	P23753	103.74	*A. nidulans*
Histone H3	unmodified	Histone H3 41-49	YKPGTVALR	_YK(pr)PGTVALR_	A2QRR5	116.37	*A. niger*
Histone H3	H3K56ac	Histone H3 54-63	YQKSTELLIR	_YQK(ac)STELLIR_	P61832	161.92	*A. fumigatus*
Histone H3	H3K56ac	Histone H3 54-63	YQKSTELLIR	_YQK(ac)STELLIR_	P23753	84.213	*A. nidulans*
Histone H3	H3K56ac	Histone H3 54-63	YQKSTELLIR	_YQK(ac)STELLIR_	A2QRR5	141.48	*A. niger*
Histone H3	unmodified	Histone H3 54-63	YQKSTELLIR	_YQK(pr)STELLIR_	P61832	144.09	*A. fumigatus*
Histone H3	unmodified	Histone H3 54-63	YQKSTELLIR	_YQK(pr)STELLIR_	P23753	148.52	*A. nidulans*
Histone H3	unmodified	Histone H3 54-63	YQKSTELLIR	_YQK(pr)STELLIR_	A2QRR5	149.94	*A. niger*
Histone H4	unmodified	Histone H4 68-78	DAVTYTEHAKR	_DAVTYTEHAK(pr)R_	Q7LKT3;Q4X0D4	245.99	*A. fumigatus*
Histone H4	unmodified	Histone H4 68-78	DAVTYTEHAKR	_DAVTYTEHAK(pr)R_	P23751;P23750	222.42	*A. nidulans*
Histone H4	unmodified	Histone H4 68-78	DAVTYTEHAKR	_DAVTYTEHAK(pr)R_	A2QRR3;A2QCZ0	161.27	*A. niger*
Histone H4	H4K31ac	Histone H4 24-35	DNIQGITKPAIR	_DNIQGITK(ac)PAIR_	Q7LKT3;Q4X0D4	91.549	*A. fumigatus*
Histone H4	H4K31ac	Histone H4 24-35	DNIQGITKPAIR	_DNIQGITK(ac)PAIR_	P23751;P23750	85.958	*A. nidulans*
Histone H4	unmodified	Histone H4 24-35	DNIQGITKPAIR	_DNIQGITK(pr)PAIR_	Q7LKT3;Q4X0D4	174.41	*A. fumigatus*
Histone H4	unmodified	Histone H4 24-35	DNIQGITKPAIR	_DNIQGITK(pr)PAIR_	P23751;P23750	192.21	*A. nidulans*
Histone H4	unmodified	Histone H4 24-35	DNIQGITKPAIR	_DNIQGITK(pr)PAIR_	A2QRR3;A2QCZ0	174.41	*A. niger*
Histone H4	unmodified	Histone H4 24-35	DNIQGITKPAIR	_DNIQGITKPAIR_	P23751;P23750	93.258	*A. nidulans*
Histone H4	unmodified	Histone H4 24-35	DNIQGITKPAIR	_DNIQGITKPAIR_	A2QRR3;A2QCZ0	99.283	*A. niger*
Histone H4	H4K5acK8acK1 2acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(ac)GLGK(ac)GGAK(ac)R_	Q7LKT3	118.48	*A. fumigatus*
Histone H4	H4K5acK8acK1 2acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(ac)GLGK(ac)GGAK(ac)R_	P23751;P23750	143.97	*A. nidulans*
Histone H4	H4K5acK8acK1 2acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(ac)GLGK(ac)GGAK(ac)R_	A2QRR3;A2QCZ0	128.88	*A. niger*
Histone H4	H4K5acK8acK1 2ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(ac)GLGK(ac)GGAK(pr)R_	Q7LKT3	155.97	*A. fumigatus*
Histone H4	H4K5acK8acK1 2ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(ac)GLGK(ac)GGAK(pr)R_	P23751;P23750	49.9	*A. nidulans*

Histone type	Modification	Peptide	Sequence	Modified sequence	Proteins	Score	Strains
Histone H4	H4K5acK8acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(ac)GLGK(pr)GGAK(ac)R_	Q7LKT3	123.67	*A. fumigatus*
Histone H4	H4K5acK8acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(ac)GLGK(pr)GGAK(ac)R_	P23751;P23750	79.036	*A. nidulans*
Histone H4	H4K5acK8acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(ac)GLGK(pr)GGAK(ac)R_	A2QRR3;A2QCZ0	119.45	*A. niger*
Histone H4	H4K5acK8ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(ac)GLGK(pr)GGAK(pr)R_	Q7LKT3	145.55	*A. fumigatus*
Histone H4	H4K5acK8ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(ac)GLGK(pr)GGAK(pr)R_	P23751;P23750	136.5	*A. nidulans*
Histone H4	H4K5acK8ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(ac)GLGK(pr)GGAK(pr)R_	A2QRR3;A2QCZ0	130.23	*A. niger*
Histone H4	H4K5acK12acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(pr)GLGK(ac)GGAK(ac)R_	Q7LKT3	169.58	*A. fumigatus*
Histone H4	H4K5acK12acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(pr)GLGK(ac)GGAK(ac)R_	P23751;P23750	89.805	*A. nidulans*
Histone H4	H4K5acK12acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(pr)GLGK(ac)GGAK(ac)R_	A2QRR3;A2QCZ0	143.97	*A. niger*
Histone H4	H4K5acK12ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(pr)GLGK(ac)GGAK(pr)R_	Q7LKT3	164.65	*A. fumigatus*
Histone H4	H4K5acK12ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(pr)GLGK(ac)GGAK(pr)R_	P23751;P23750	149.96	*A. nidulans*
Histone H4	H4K5acK12ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(pr)GLGK(ac)GGAK(pr)R_	A2QRR3;A2QCZ0	114.86	*A. niger*
Histone H4	H4K5acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(pr)GLGK(pr)GGAK(ac)R_	Q7LKT3	147.37	*A. fumigatus*
Histone H4	H4K5acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(pr)GLGK(pr)GGAK(ac)R_	P23751;P23750	174.4	*A. nidulans*
Histone H4	H4K5acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(pr)GLGK(pr)GGAK(ac)R_	A2QRR3;A2QCZ0	158.05	*A. niger*
Histone H4	H4K5ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(pr)GLGK(pr)GGAK(pr)R_	Q7LKT3	173.71	*A. fumigatus*
Histone H4	H4K5ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(pr)GLGK(pr)GGAK(pr)R_	P23751;P23750	162.38	*A. nidulans*
Histone H4	H4K5ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(ac)GGK(pr)GLGK(pr)GGAK(pr)R_	A2QRR3;A2QCZ0	150.04	*A. niger*
Histone H4	H4K8acK12acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(ac)GLGK(ac)GGAK(ac)R_	Q7LKT3	152.7	*A. fumigatus*
Histone H4	H4K8acK12acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(ac)GLGK(ac)GGAK(ac)R_	P23751;P23750	147.91	*A. nidulans*
Histone H4	H4K8acK12acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(ac)GLGK(ac)GGAK(ac)R_	A2QRR3;A2QCZ0	156.51	*A. niger*
Histone H4	H4K8acK12ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(ac)GLGK(ac)GGAK(pr)R_	P23751;P23750	65.5	*A. nidulans*
Histone H4	H4K8acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(ac)GLGK(pr)GGAK(ac)R_	Q7LKT3	82.55	*A. fumigatus*
Histone H4	H4K8acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(ac)GLGK(pr)GGAK(ac)R_	P23751;P23750	122.51	*A. nidulans*
Histone H4	H4K8acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(ac)GLGK(pr)GGAK(ac)R_	A2QRR3;A2QCZ0	113.22	*A. niger*
Histone H4	H4K8ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(ac)GLGK(pr)GGAK(pr)R_	Q7LKT3	147.37	*A. fumigatus*
Histone H4	H4K8ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(ac)GLGK(pr)GGAK(pr)R_	A2QRR3;A2QCZ0	134.75	*A. niger*
Histone H4	H4K12acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(pr)GLGK(ac)GGAK(ac)R_	Q7LKT3	52.5	*A. fumigatus*
Histone H4	H4K12acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(pr)GLGK(ac)GGAK(ac)R_	P23751;P23750	133.07	*A. nidulans*

Histone type	Modification	Peptide	Sequence	Modified sequence	Proteins	Score	Strains
Histone H4	H4K12acK16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(pr)GLGK(ac)GGAK(ac)R_	A2QRR3;A2QCZ0	50.843	*A. niger*
Histone H4	H4K12ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(pr)GLGK(ac)GGAK(pr)R_	Q7LKT3	138.76	*A. fumigatus*
Histone H4	H4K12ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(pr)GLGK(ac)GGAK(pr)R_	P23751;P23750	134.75	*A. nidulans*
Histone H4	H4K12ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(pr)GLGK(ac)GGAK(pr)R_	A2QRR3;A2QCZ0	70.373	*A. niger*
Histone H4	H4K16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(pr)GLGK(pr)GGAK(ac)R_	Q7LKT3	173.71	*A. fumigatus*
Histone H4	H4K16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(pr)GLGK(pr)GGAK(ac)R_	P23751;P23750	173.71	*A. nidulans*
Histone H4	H4K16ac	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(pr)GLGK(pr)GGAK(ac)R_	A2QRR3;A2QCZ0	157.97	*A. niger*
Histone H4	unmodified	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(pr)GLGK(pr)GGAK(pr)R_	Q7LKT3	175.86	*A. fumigatus*
Histone H4	unmodified	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(pr)GLGK(pr)GGAK(pr)R_	P23751;P23750	164.64	*A. nidulans*
Histone H4	unmodified	Histone H4 4-17	GKGGKGLGKGGAKR	_GK(pr)GGK(pr)GLGK(pr)GGAK(pr)R_	A2QRR3;A2QCZ0	164.64	*A. niger*
Histone H4	unmodified	Histone H4 79-92	KTVTSLDVVYALKR	_K(pr)TVTSLDVVYALK(pr)R_	Q7LKT3;Q4X0D4	232	*A. fumigatus*
Histone H4	unmodified	Histone H4 79-92	KTVTSLDVVYALKR	_K(pr)TVTSLDVVYALK(pr)R_	P23751;P23750	165.66	*A. nidulans*
Histone H4	unmodified	Histone H4 79-92	KTVTSLDVVYALKR	_K(pr)TVTSLDVVYALK(pr)R_	A2QRR3;A2QCZ0	156.71	*A. niger*

pr: propionylation
ac: acetylation
pm: propionylation+mono-methylation
2m: dimethylation
3m: trimethylation

Supplementary Table 2 Mass spectrometry result for Aspergilli

Peptide	AUC (area under the curve)	%RA (%relative abundance)	AUC (area under the curve)	%RA (%relative abundance)	AUC (area under the curve)	%RA (%relative abundance)
			A. niger			
H3_3_8 unmod	7.18E+08	71.74498203	2.40E+09	78.38093567	1.09E+09	75.15819003
H3_3_8 K4me1	7.43E+07	7.42147314	1.58E+08	5.165248032	9.68E+07	6.674470436
H3_3_8 K4me2	8.44E+07	8.434557206	2.00E+08	6.522554776	9.96E+07	6.864265043
H3_3_8 K4me3	1.24E+08	12.39898763	3.04E+08	9.93126152	1.64E+08	11.30307449
H3_9_17 unmod	1.93E+08	61.74089019	3.78E+08	53.47092716	1.67E+08	62.11457779
H3_9_17 K9me1	1.10E+06	0.352621445	2.18E+06	0.308392969	4.07E+06	1.516149598
H3_9_17 K9me2	9.99E+05	0.319367963	1.53E+06	0.21576661	1.26E+06	0.470894189
H3_9_17 K9me3	1.02E+07	3.253735668	2.17E+07	3.075021371	4.05E+06	1.510408675
H3_9_17 K9ac	2.08E+07	6.646189057	6.47E+07	9.148737256	2.66E+07	9.92870346
H3_9_17 K14ac	7.58E+07	24.21333955	2.12E+08	30.028234	5.85E+07	21.80695434
H3_9_17 K9me1K14ac	0.00E+00	0	1.97E+05	0.027914189	0.00E+00	0
H3_9_17 K9me2K14ac	1.56E+06	0.499058298	9.65E+05	0.136555239	5.54E+05	0.206688657
H3_9_17 K9me3K14ac	1.21E+05	0.038715208	1.12E+06	0.158130313	6.02E+05	0.224702118
H3_9_17 K9acK14ac	9.19E+06	2.936082624	2.42E+07	3.430320896	5.95E+06	2.22092118
H3_18_26 unmod	1.94E+09	68.53083245	6.38E+09	65.05835135	1.12E+09	74.78877313
H3_18_26 K18ac	5.86E+07	2.065728719	2.28E+08	2.323415548	1.98E+07	1.325218028
H3_18_26 K23ac	7.62E+08	26.87379168	2.96E+09	30.20057295	2.50E+08	16.74046161
H3_18_26 K18acK23ac	7.17E+07	2.529647154	2.37E+08	2.417660146	1.07E+08	7.145547231
H3_27_40 unmod	2.71E+08	48.85544289	1.92E+09	53.36799341	4.18E+08	58.55505613
H3_27_40 K36me1	1.27E+08	22.91810105	1.04E+09	28.87201633	2.31E+08	32.36582225
H3_27_40 K27ac	1.56E+08	28.22645606	6.38E+08	17.75999025	6.49E+07	9.079121624
H3_73_83 unmod	9.23E+06	37.32104904	2.85E+08	68.8968618	5.23E+06	46.93802331
H3_73_83 K79me1	7.02E+05	2.840708095	3.43E+07	8.295135328	1.80E+05	1.618302929
H3_73_83 K79me2	1.48E+07	59.83824287	9.43E+07	22.80800287	5.73E+06	51.44367376
H4_4_17 unmod	1.78E+08	16.26113952	7.10E+08	22.52887821	1.92E+08	14.17085111
H4_4_17 K5(12)ac	2.71E+07	2.469063351	4.51E+07	1.429872787	1.07E+07	0.791137712
H4_4_17 K8ac	3.38E+07	3.078258801	1.03E+08	3.280974183	3.54E+07	2.619351703
H4_4_17 K16ac	6.47E+08	58.92683519	1.70E+09	54.03111373	9.02E+08	66.72197268
H4_4_17 di ac	1.66E+08	15.10557276	5.08E+08	16.11823285	1.85E+08	13.69366907
H4_4_17 tri ac	3.69E+07	3.363870205	6.26E+07	1.984534061	2.28E+07	1.686025132
H4_4_17 tetra ac	8.73E+06	0.79526017	1.98E+07	0.626394176	4.29E+06	0.3169926
			A. fumigatus			
H3_3_8 unmod	1.21E+10	65.37462147	7.97E+09	79.01520041	2.51E+10	68.66504751
H3_3_8 K4me1	1.64E+09	8.876735687	7.33E+08	7.264931446	3.03E+09	8.296372762
H3_3_8 K4me2	1.60E+09	8.646063848	4.73E+08	4.692999588	2.72E+09	7.43145067
H3_3_8 K4me3	3.17E+09	17.102579	9.11E+08	9.026868555	5.71E+09	15.60712906
H3_9_17 unmod	2.55E+09	46.77745437	1.17E+09	46.05468421	5.76E+09	49.25272734
H3_9_17 K9me1	1.18E+08	2.153673966	3.63E+07	1.425873986	1.88E+08	1.609120929
H3_9_17 K9me2	9.84E+07	1.802208495	3.25E+07	1.275488804	1.62E+08	1.38194368
H3_9_17 K9me3	6.47E+08	11.83997076	2.20E+08	8.632988069	8.88E+08	7.591377122
H3_9_17 K9ac	3.67E+08	6.717959219	2.12E+08	8.334417453	8.38E+08	7.163739922
H3_9_17 K14ac	1.27E+09	23.33812014	6.83E+08	26.84137344	3.00E+09	25.61751225
H3_9_17 K9me1K14ac	5.11E+07	0.93553049	2.09E+07	0.821773955	6.88E+07	0.588320398
H3_9_17 K9me2K14ac	2.45E+07	0.449269661	8.15E+06	0.320233896	9.18E+07	0.784339385
H3_9_17 K9me3K14ac	3.49E+07	0.638822782	1.42E+07	0.558268367	1.38E+08	1.179177426
H3_9_17 K9acK14ac	2.92E+08	5.34699012	1.46E+08	5.734897813	5.65E+08	4.831741549

Peptide	AUC (area under the curve)	%RA (%relative abundance)	AUC (area under the curve)	%RA (%relative abundance)	AUC (area under the curve)	%RA (%relative abundance)
H3_18_26 unmod	4.15E+10	57.95349818	1.18E+10	55.50523591	2.31E+10	57.30808834
H3_18_26 K18ac	1.82E+09	2.540117933	5.70E+08	2.684460567	8.50E+08	2.107307635
H3_18_26 K23ac	2.40E+10	33.53406004	7.45E+09	35.07291497	1.11E+10	27.60680519
H3_18_26 K18acK23ac	4.28E+09	5.972323845	1.43E+09	6.737388557	5.23E+09	12.97779884
H3_27_40 unmod	8.83E+09	72.9191621	7.05E+09	86.98896126	1.94E+10	81.59701454
H3_27_40 K36me1	1.01E+09	8.346196461	2.27E+08	2.802671368	1.01E+09	4.251562355
H3_27_40 K27ac	2.27E+09	18.73464144	8.27E+08	10.20836737	3.37E+09	14.15142311
H3_73_83 unmod	2.63E+09	52.08256574	3.59E+08	44.1582243	2.98E+09	47.55166207
H3_73_83 K79me1	7.04E+08	13.96233475	4.20E+07	5.167349855	1.52E+09	24.2946727
H3_73_83 K79me2	1.71E+09	33.95509951	4.12E+08	50.67442584	1.76E+09	28.15366523
H4_4_17 unmod	2.32E+09	22.22808084	1.60E+09	19.11415215	1.22E+09	21.57236395
H4_4_17 K5(12)ac	1.64E+08	1.569176339	1.52E+08	1.816191076	6.67E+07	1.175726315
H4_4_17 K8ac	2.64E+08	2.530059776	1.94E+08	2.310907899	1.03E+08	1.8124813
H4_4_17 K16ac	4.81E+09	46.11428545	4.03E+09	48.02407656	3.01E+09	53.11941359
H4_4_17 di ac	2.27E+09	21.75434699	1.84E+09	21.98468179	9.30E+08	16.40669378
H4_4_17 tri ac	4.96E+08	4.760857809	4.28E+08	5.09892276	2.87E+08	5.064836335
H4_4_17 tetra ac	1.09E+08	1.043192792	1.38E+08	1.651067763	4.81E+07	0.848484725
			***A. nidulans* (Wtpaba)**			
H3_3_8 unmod	8.44E+08	75.94789801	2.41E+09	76.16196523	1.34E+09	74.06620666
H3_3_8 K4me1	7.30E+07	6.568679964	2.32E+08	7.34634065	1.30E+08	7.183063849
H3_3_8 K4me2	5.87E+07	5.282879562	1.65E+08	5.2155299	1.15E+08	6.329891955
H3_3_8 K4me3	1.36E+08	12.20054246	3.57E+08	11.27616422	2.25E+08	12.42083753
H3_9_17 unmod	1.71E+08	60.35465102	5.23E+08	57.90343372	2.29E+08	62.64155061
H3_9_17 K9me1	2.91E+06	1.022758118	7.96E+06	0.880916884	5.19E+06	1.420381887
H3_9_17 K9me2	2.43E+06	0.856209689	6.33E+06	0.700136563	2.32E+06	0.635133822
H3_9_17 K9me3	4.49E+06	1.580063971	1.55E+07	1.711970814	7.96E+06	2.178255657
H3_9_17 K9ac	2.83E+07	9.961305561	7.13E+07	7.893265506	3.62E+07	9.911437796
H3_9_17 K14ac	6.19E+07	21.785818	2.19E+08	24.2642191	7.19E+07	19.65384196
H3_9_17 K9me1K14ac	7.92E+05	0.27890292	4.73E+06	0.523240139	1.72E+06	0.469225016
H3_9_17 K9me2K14ac	1.34E+06	0.470520377	2.45E+06	0.271678249	4.46E+05	0.122108065
H3_9_17 K9me3K14ac	6.96E+05	0.245198997	2.00E+06	0.221637203	1.64E+05	0.04495934
H3_9_17 K9acK14ac	9.78E+06	3.444571344	5.09E+07	5.629501822	1.07E+07	2.923105849
H3_18_26 unmod	5.60E+08	48.89967236	2.89E+09	54.25010725	2.72E+09	67.68217964
H3_18_26 K18ac	3.57E+07	3.119337955	1.35E+08	2.536230939	8.07E+07	2.011872549
H3_18_26 K23ac	4.64E+08	40.48722501	1.76E+09	33.06132344	1.00E+09	25.03482023
H3_18_26 K18acK23ac	8.59E+07	7.493764673	5.41E+08	10.15233837	2.11E+08	5.271127583
H3_27_40 unmod	3.13E+08	61.3683983	1.14E+09	53.41466528	6.00E+08	43.2893569
H3_27_40 K36me1	1.49E+08	29.30996741	6.22E+08	29.23298744	6.15E+08	44.3267246
H3_27_40 K27ac	4.75E+07	9.321634289	3.69E+08	17.35234728	1.72E+08	12.3839185
H3_73_83 unmod	2.01E+06	21.45531547	2.72E+06	7.328082514	1.04E+07	25.79029474
H3_73_83 K79me1	1.62E+05	1.733713638	3.23E+05	0.870637732	8.23E+05	2.037429289
H3_73_83 K79me2	7.19E+06	76.81097089	3.40E+07	91.80127975	2.91E+07	72.17227597
H4_4_17 unmod	1.20E+08	13.23766506	2.35E+08	12.43674615	7.11E+08	22.23722084
H4_4_17 K5(12)ac	7.54E+06	0.83012622	1.86E+07	0.982900188	2.63E+07	0.823919433
H4_4_17 K8ac	2.93E+07	3.224986232	5.84E+07	3.091948541	1.04E+08	3.245908884
H4_4_17 K16ac	5.49E+08	60.45519758	1.08E+09	57.27427838	1.80E+09	56.45196893
H4_4_17 di ac	1.55E+08	17.02404417	3.40E+08	17.99026701	4.63E+08	14.48712444
H4_4_17 tri ac	4.11E+07	4.522703467	1.30E+08	6.887883783	7.47E+07	2.337693026
H4_4_17 tetra ac	6.41E+06	0.705277279	2.52E+07	1.335975938	1.33E+07	0.416164452

Peptide	AUC (area under the curve)	%RA (%relative abundance)	AUC (area under the curve)	%RA (%relative abundance)	AUC (area under the curve)	%RA (%relative abundance)
			***A. nidulans* (TN02A25)**			
H3_3_8 unmod	5.78E+09	86.4	8.79E+07	80.0	1.39E+09	85.7
H3_3_8 K4me1	2.59E+08	3.9	7.90E+06	7.2	7.27E+07	4.5
H3_3_8 K4me2	2.08E+08	3.1	2.95E+06	2.7	4.04E+07	2.5
H3_3_8 K4me3	4.45E+08	6.7	1.12E+07	10.2	1.20E+08	7.4
H3_9_17 unmod	4.72E+08	51.0	1.49E+08	72.2	2.39E+08	63.0
H3_9_17 K9me1	1.57E+07	1.7	2.00E+06	1.0	9.03E+06	2.4
H3_9_17 K9me2	7.98E+06	0.9	3.71E+06	1.8	4.22E+06	1.1
H3_9_17 K9me3	2.24E+07	2.4	4.61E+06	2.2	1.11E+07	2.9
H3_9_17 K9ac	6.13E+07	6.6	9.73E+06	4.7	2.12E+07	5.6
H3_9_17 K14ac	3.21E+08	34.6	3.21E+07	15.5	8.32E+07	22.0
H3_9_17 K9me1K14ac	0.00E+00	0.0	0.00E+00	0.0	2.31E+06	0.6
H3_9_17 K9me2K14ac	2.01E+06	0.2	8.74E+05	0.4	1.32E+06	0.3
H3_9_17 K9me3K14ac	1.49E+06	0.2	3.11E+05	0.2	3.14E+05	0.1
H3_9_17 K9acK14ac	2.29E+07	2.5	4.27E+06	2.1	7.22E+06	1.9
H3_18_26 unmod	7.08E+09	62.1	2.79E+08	65.6	1.27E+09	60.7
H3_18_26 K18ac	2.61E+08	2.3	7.60E+06	1.8	4.43E+07	2.1
H3_18_26 K23ac	3.33E+09	29.2	1.00E+08	23.5	5.71E+08	27.2
H3_18_26 K18acK23ac	7.26E+08	6.4	3.91E+07	9.2	2.09E+08	10.0
H3_27_40 unmod	3.10E+09	91.2	9.54E+07	82.9	6.88E+08	91.7
H3_27_40 K36me1	9.87E+07	2.9	2.78E+06	2.4	1.92E+07	2.6
H3_27_40 K27ac	1.99E+08	5.9	1.70E+07	14.7	4.34E+07	5.8
H3_73_83 unmod	2.19E+08	63.1	9.13E+05	70.4	5.53E+06	29.8
H3_73_83 K79me1	3.89E+07	11.2	3.84E+05	29.6	4.21E+05	2.3
H3_73_83 K79me2	8.87E+07	25.6	0.00E+00	0.0	1.26E+07	68.0
H4_4_17 unmod	9.41E+08	18.3	7.51E+07	17.3	1.30E+08	21.2
H4_4_17 K5(12)ac	9.31E+07	1.8	3.33E+06	0.8	1.16E+07	1.9
H4_4_17 K8ac	1.92E+08	3.7	8.81E+06	2.0	2.05E+07	3.4
H4_4_17 K16ac	3.17E+09	61.5	2.92E+08	67.1	3.77E+08	61.6
H4_4_17 di ac	6.11E+08	11.9	4.29E+07	9.9	5.52E+07	9.0
H4_4_17 tri ac	1.14E+08	2.2	1.08E+07	2.5	1.38E+07	2.3
H4_4_17 tetra ac	3.00E+07	0.6	2.38E+06	0.5	3.40E+06	0.6

H4di-ac: H4K5acK8ac, H4K5acK12ac, H4K5acK16ac, H4K8acK16ac, and H4K12acK16ac
H4tri-ac: H4K5acK8ac12ac, H4K5acK8acK16ac, H4K5acK12acK16ac, and H4K8acK12acK16ac
H4tetra-ac: H4K5acK8ack12acK16ac

Supplementary Table 3 One-way ANOVA test and Tukey HSD test for all modifications in three Aspergilli

H3K14ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	23.216248	11.608124	1.321585	0.33453		
Residual	6	52.700915	8.783486	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	0.083841	-7.332253	7.499934	0.048998	0.9
1	A. niger	A. nidulans (wtpaba)	3.448216	-3.967877	10.86431	2.015216	0.388059
2	A. fumigatus	A. nidulans (wtpaba)	3.364376	-4.051718	10.780469	1.966218	0.403736
H3K18ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	0.727055	0.363528	1.638387	0.27056		
Residual	6	1.331289	0.221882	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	0.539175	-0.639523	1.717872	1.982574	0.39846
1	A. niger	A. nidulans (wtpaba)	0.651026	-0.527671	1.829724	2.39386	0.282307
2	A. fumigatus	A. nidulans (wtpaba)	0.111852	-1.066845	1.290549	0.411285	0.9
H3K18acK23ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	34.406252	17.203126	1.841974	0.237846		
Residual	6	56.037021	9.339504	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	4.531552	-3.115668	12.178773	2.5683	0.242484
1	A. niger	A. nidulans (wtpaba)	3.608125	-4.039095	11.255346	2.044939	0.378759
2	A. fumigatus	A. nidulans (wtpaba)	0.923427	-6.723794	8.570648	0.523361	0.9
H3K23ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	124.534309	62.267155	1.501388	0.296022		
Residual	6	248.838417	41.473069	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	7.466318	-8.648478	23.581114	2.008094	0.390312
1	A. niger	A. nidulans (wtpaba)	8.256181	-7.858615	24.370977	2.22053	0.32716
2	A. fumigatus	A. nidulans (wtpaba)	0.789863	-15.32493	16.904659	0.212437	0.9
H3K27ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	46.205235	23.102618	0.547649	0.604702		
Residual	6	253.110317	42.185053	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	3.990379	-12.26215	20.242911	1.064132	0.731798
1	A. niger	A. nidulans (wtpaba)	5.335889	-10.91664	21.588421	1.422945	0.598086
2	A. fumigatus	A. nidulans (wtpaba)	1.345511	-14.90702	17.598042	0.358813	0.9
H3K36me1							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	1414.265967	707.13298	19.893056	0.00225		
Residual	6	213.280344	35.546724	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	22.918503	7.999441	37.837565	6.658051	0.007831
1	A. niger	A. nidulans (wtpaba)	6.237913	-8.681149	21.156975	1.812175	0.455479
2	A. fumigatus	A. nidulans (wtpaba)	29.156416	14.237354	44.075478	8.470226	0.002345
H3K4me1							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	4.592137	2.296069	3.195443	0.113539		
Residual	6	4.311268	0.718545	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	1.725616	-0.39552	3.846753	3.525964	0.102596
1	A. niger	A. nidulans (wtpaba)	0.612298	-1.508839	2.733434	1.251112	0.662123
2	A. fumigatus	A. nidulans (wtpaba)	1.113318	-1.007818	3.234455	2.274852	0.312524
H3K4me2							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	4.619573	2.309787	1.252972	0.350983		
Residual	6	11.060676	1.843446	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	0.350288	-3.047193	3.747768	0.446859	0.9
1	A. niger	A. nidulans (wtpaba)	1.664359	-1.733122	5.061839	2.123207	0.355031
2	A. fumigatus	A. nidulans (wtpaba)	1.314071	-2.083409	4.711551	1.676348	0.503658

H3K4me3							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	11.653745	5.826872	0.858723	0.46993		
Residual	6	40.71305	6.785508	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	2.701084	-3.817194	9.219363	1.796005	0.461096
1	A. niger	A. nidulans (wtpaba)	0.75474	-5.763538	7.273019	0.501842	0.9
2	A. fumigatus	A. nidulans (wtpaba)	1.946344	-4.571934	8.464623	1.294163	0.646078
H3K79me1							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	278.951174	139.47559	3.99779	0.078792		
Residual	6	209.329045	34.888174	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	10.223404	-4.556815	25.003622	2.997899	0.165389
1	A. niger	A. nidulans (wtpaba)	2.704122	-12.0761	17.48434	0.792954	0.832853
2	A. fumigatus	A. nidulans (wtpaba)	12.927526	-1.852693	27.707744	3.790853	0.080838
H3K79me2							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	3135.784212	1567.8921	7.599627	0.022672		
Residual	6	1237.870308	206.31172	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	7.102243	-28.83988	43.044362	0.856435	0.809195
1	A. niger	A. nidulans (wtpaba)	35.564869	-0.37725	71.506988	4.288643	0.052024
2	A. fumigatus	A. nidulans (wtpaba)	42.667112	6.724993	78.609231	5.145078	0.025144
H3K9ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	5.252807	2.626404	1.566421	0.283554		
Residual	6	10.060145	1.676691	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	1.169171	-2.071002	4.409344	1.563912	0.545559
1	A. niger	A. nidulans (wtpaba)	0.680793	-2.55938	3.920966	0.910645	0.788995
2	A. fumigatus	A. nidulans (wtpaba)	1.849964	-1.390209	5.090137	2.474557	0.263225
H3K9acK14ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	8.960048	4.480024	5.097229	0.050857		
Residual	6	5.273482	0.878914	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	2.442102	0.096173	4.78803	4.511815	0.04285
1	A. niger	A. nidulans (wtpaba)	1.136618	-1.209311	3.482547	2.099917	0.36197
2	A. fumigatus	A. nidulans (wtpaba)	1.305483	-1.040445	3.651412	2.411898	0.27794
H3K9me1							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	1.540145	0.770072	3.345324	0.105682		
Residual	6	1.381162	0.230194	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	1.003835	-0.196738	2.204408	3.6239	0.093917
1	A. niger	A. nidulans (wtpaba)	0.382298	-0.818275	1.58287	1.380116	0.614049
2	A. fumigatus	A. nidulans (wtpaba)	0.621537	-0.579035	1.82211	2.243784	0.320817
H3K9me1K14ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	0.896888	0.448444	27.98933	0.000907		
Residual	6	0.096132	0.016022	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	0.77257	0.455833	1.089308	10.571603	0.001
1	A. niger	A. nidulans (wtpaba)	0.414485	0.097747	0.731222	5.671675	0.016474
2	A. fumigatus	A. nidulans (wtpaba)	0.358086	0.041348	0.674823	4.899928	0.030806
H3K9me2							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	2.053032	1.026516	28.79737	0.00084		
Residual	6	0.213877	0.035646	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	1.151204	0.678762	1.623646	10.561036	0.001
1	A. niger	A. nidulans (wtpaba)	0.39515	-0.077291	0.867592	3.625072	0.093818
2	A. fumigatus	A. nidulans (wtpaba)	0.756054	0.283612	1.228495	6.935964	0.006424
H3K9me2K14ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		

C(Species)	2	0.109137	0.054569	1.310654	0.337081		
Residual	6	0.249808	0.041635	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	0.23718	-0.273406	0.747767	2.013315	0.388662
1	A. niger	A. nidulans (wtpaba)	0.007335	-0.503251	0.517921	0.062262	0.9
2	A. fumigatus	A. nidulans (wtpaba)	0.229845	-0.280741	0.740432	1.951053	0.40867
H3K9me3							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	102.795553	51.397776	26.037349	0.001103		
Residual	6	11.844011	1.974002	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	6.741723	3.225993	10.257453	8.311086	0.002589
1	A. niger	A. nidulans (wtpaba)	0.789625	-2.726105	4.305355	0.973437	0.765595
2	A. fumigatus	A. nidulans (wtpaba)	7.531349	4.015619	11.047078	9.284523	0.001448
H3K9me3K14ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	0.811704	0.405852	9.028419	0.015515		
Residual	6	0.269716	0.044953	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	0.651574	0.121032	1.182116	5.322873	0.021752
1	A. niger	A. nidulans (wtpaba)	0.030083	-0.500459	0.560625	0.245753	0.9
2	A. fumigatus	A. nidulans (wtpaba)	0.621491	0.090949	1.152033	5.07712	0.026589
H3_18_26_unmod							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	313.811514	156.90576	3.936903	0.080885		
Residual	6	239.130725	39.855121	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	12.537045	-3.260289	28.334379	3.439646	0.11092
1	A. niger	A. nidulans (wtpaba)	12.515333	-3.282001	28.312666	3.433689	0.11152
2	A. fumigatus	A. nidulans (wtpaba)	0.021712	-15.77562	15.819046	0.005957	0.9
H3_27_40_unmod							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	1498.34805	749.17403	14.402205	0.005123		
Residual	6	312.108055	52.018009	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	26.908882	8.861318	44.956446	6.462185	0.009032
1	A. niger	A. nidulans (wtpaba)	0.902024	-17.14554	18.949588	0.216622	0.9
2	A. fumigatus	A. nidulans (wtpaba)	27.810906	9.763342	45.85847	6.678807	0.007715
H3_3_8_unmod							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	35.835086	17.917543	0.853061	0.472004		
Residual	6	126.022904	21.003817	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	4.076413	-7.391671	15.544496	1.540599	0.554245
1	A. niger	A. nidulans (wtpaba)	0.297321	-11.17076	11.765404	0.112366	0.9
2	A. fumigatus	A. nidulans (wtpaba)	4.373734	-7.09435	15.841817	1.652966	0.512369
H3_73_83_unmod							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	1974.01335	987.00668	7.982002	0.020385		
Residual	6	741.924155	123.65403	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	3.121161	-24.70452	30.946845	0.486153	0.9
1	A. niger	A. nidulans (wtpaba)	32.860747	5.035063	60.686432	5.118397	0.025698
2	A. fumigatus	A. nidulans (wtpaba)	29.739586	1.913902	57.565271	4.632244	0.038639
H3_9_17_unmod							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	306.813306	153.40665	14.24759	0.005262		
Residual	6	64.603203	10.7672	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	11.747176	3.536232	19.958121	6.200729	0.010973
1	A. niger	A. nidulans (wtpaba)	1.19108	-7.019865	9.402025	0.62871	0.894058
2	A. fumigatus	A. nidulans (wtpaba)	12.938256	4.727312	21.149201	6.829438	0.006925
H4K16ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	200.701418	100.35071	5.141912	0.050025		
Residual	6	117.09734	19.516223	NaN	NaN		

Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	10.807382	-0.247131	21.861895	4.237242	0.054417
1	A. niger	A. nidulans (wtpaba)	1.832826	-9.221687	12.887339	0.718594	0.860562
2	A. fumigatus	A. nidulans (wtpaba)	8.974556	-2.079957	20.029069	3.518647	0.103274
H4K5(12)ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	0.881591	0.440795	1.593876	0.278501		
Residual	6	1.659334	0.276556	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	0.042993	-1.272937	1.358924	0.141602	0.9
1	A. niger	A. nidulans (wtpaba)	0.684376	-0.631555	2.000307	2.254053	0.318057
2	A. fumigatus	A. nidulans (wtpaba)	0.641383	-0.674548	1.957313	2.112451	0.358221
H4K8ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	1.579132	0.789566	9.212845	0.014822		
Residual	6	0.514217	0.085703	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	0.775045	0.042492	1.507598	4.585538	0.040217
1	A. niger	A. nidulans (wtpaba)	0.194753	-0.5378	0.927306	1.152252	0.69896
2	A. fumigatus	A. nidulans (wtpaba)	0.969798	0.237245	1.702351	5.73779	0.015644
H4_4_17_unmod							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	38.851161	19.425581	1.137616	0.381165		
Residual	6	102.454191	17.075698	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	3.317909	-7.022339	13.658158	1.390708	0.610099
1	A. niger	A. nidulans (wtpaba)	1.683079	-8.657169	12.023327	0.705466	0.865455
2	A. fumigatus	A. nidulans (wtpaba)	5.000988	-5.33926	15.341237	2.096174	0.363098
H4di_ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	40.690342	20.345171	4.147213	0.073953		
Residual	6	29.434472	4.905745	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	5.076083	-0.466271	10.618436	3.969509	0.068925
1	A. niger	A. nidulans (wtpaba)	1.527987	-4.014367	7.070341	1.19489	0.68307
2	A. fumigatus	A. nidulans (wtpaba)	3.548096	-1.994258	9.090449	2.77462	0.202011
H4tetra_ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	0.549926	0.274963	1.811639	0.242373		
Residual	6	0.910656	0.151776	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	0.601366	-0.373496	1.576228	2.67361	0.221009
1	A. niger	A. nidulans (wtpaba)	0.23959	-0.735272	1.214453	1.065193	0.731402
2	A. fumigatus	A. nidulans (wtpaba)	0.361776	-0.613086	1.336638	1.608417	0.528971
H4tri_ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	2	12.079402	6.039701	3.012547	0.124219		
Residual	6	12.029092	2.004849	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	2.630063	-0.91303	6.173155	3.217258	0.135612
1	A. niger	A. nidulans (wtpaba)	2.23795	-1.305143	5.781043	2.737602	0.208793
2	A. fumigatus	A. nidulans (wtpaba)	0.392112	-3.150981	3.935205	0.479656	0.9

H4di-ac: H4K5acK8ac, H4K5acK12ac, H4K5acK16ac, H4K8acK16ac, and H4K12acK16ac
H4tri-ac: H4K5acK8ac12ac, H4K5acK8acK16ac, H4K5acK12acK16ac, and H4K8acK12acK16ac
H4tetra-ac: H4K5acK8ack12acK16ac

Supplementary Table 4 One-way ANOVA test and Tukey HSD test for all modifications in two *A. nidulans* strai

H3K14ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	6.745274	6.745274	0.135645	0.731312		
Residual	4	198.90938	49.727345	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	2.120578	-13.86547	18.106625	0.520855	0.745455
H3K18ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	0.366406	0.366406	1.963088	0.233796		
Residual	4	0.74659	0.186648	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	0.494237	-0.485151	1.473625	1.981458	0.233795
H3K18acK23ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	1.133998	1.133998	0.23697	0.651874		
Residual	4	19.141662	4.785416	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	0.869482	-4.089623	5.828587	0.688432	0.667354
H3K23ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	57.966244	57.966244	1.698309	0.262462		
Residual	4	136.527	34.131749	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	6.216443	-7.027666	19.460551	1.842991	0.26246
H3K27ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	26.859814	26.859814	1.253903	0.325502		
Residual	4	85.683856	21.420964	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	4.231612	-6.260501	14.723725	1.583605	0.325504
H3K36me1							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	1503.7824	1503.7824	39.772594	0.003232		
Residual	4	151.23805	37.809512	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	31.662621	17.723224	45.602019	8.918811	0.003233
H3K4me1							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	5.177407	5.177407	3.14715	0.150731		
Residual	4	6.580439	1.64511	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	1.85785	-1.049792	4.765491	2.508844	0.15073
H3K4me2							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	12.175933	12.175933	49.29373	0.002168		
Residual	4	0.988031	0.247008	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	2.849086	1.72241	3.975761	9.929122	0.002165
H3K4me3							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	22.813839	22.813839	11.929521	0.025962		
Residual	4	7.649541	1.912385	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	3.899901	0.764947	7.034855	4.884572	0.025961

H3K79me1							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	246.63006	246.63006	2.53481	0.186576		
Residual	4	389.18905	97.297262	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	12.822638	-9.538494	35.183769	2.251582	0.186575
H3K79me2							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	3610.8287	3610.8287	5.62737	0.076636		
Residual	4	2566.6192	641.6548	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	49.063419	-8.360679	106.48752	3.354808	0.076638
H3K9ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	19.615719	19.615719	17.039886	0.014518		
Residual	4	4.60466	1.151165	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	3.616234	1.183961	6.048506	5.837788	0.014519
H3K9acK14ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	5.131732	5.131732	4.778719	0.094109		
Residual	4	4.295487	1.073872	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	1.849636	-0.499562	4.198835	3.091511	0.094107
H3K9me1							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	0.494052	0.494052	1.700041	0.262256		
Residual	4	1.162447	0.290612	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	0.573906	-0.648175	1.795987	1.843931	0.262256
H3K9me1K14ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	0.072775	0.072775	1.034183	0.366686		
Residual	4	0.281479	0.07037	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	0.220265	-0.381097	0.821628	1.438182	0.366686
H3K9me2							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	0.416433	0.416433	3.370126	0.140282		
Residual	4	0.494264	0.123566	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	0.526899	-0.269982	1.323779	2.596199	0.140281
H3K9me2K14ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	0.002594	0.002594	0.125162	0.741362		
Residual	4	0.082905	0.020726	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	0.041586	-0.284778	0.367951	0.500323	0.755025
H3K9me3							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	0.743072	0.743072	6.38326	0.064905		
Residual	4	0.465638	0.11641	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	0.703833	-0.069626	1.477293	3.573027	0.064906
H3K9me3K14ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		

C(Species)	1	0.002297	0.002297	0.333859	0.594363		
Residual	4	0.027521	0.00688	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	0.039133	-0.148906	0.227172	0.81714	0.607365
H3_18_26_unmod							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	51.179398	51.179398	1.024586	0.368687		
Residual	4	199.80509	49.951273	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	5.841198	-10.180802	21.863199	1.431493	0.368687
H3_27_40_unmod							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	1932.594	1932.594	36.219739	0.003839		
Residual	4	213.42992	53.357481	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	35.894234	19.334966	52.453501	8.511138	0.003838
H3_3_8_unmod							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	111.11644	111.11644	16.229041	0.015753		
Residual	4	27.387061	6.846765	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	8.606836	2.67504	14.538631	5.6972	0.015752
H3_73_83_unmod							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	1970.0914	1970.0914	7.004443	0.057185		
Residual	4	1125.0523	281.26308	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	36.240781	-1.778143	74.259706	3.742845	0.057185
H3_9_17_unmod							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	4.638306	4.638306	0.078334	0.793446		
Residual	4	236.84864	59.21216	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	1.758466	-15.685648	19.20258	0.395812	0.803735
H4K16ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	43.054295	43.054295	5.951317	0.071244		
Residual	4	28.937659	7.234415	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	5.357505	-0.739901	11.454912	3.450019	0.071245
H4K5(12)ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	0.564898	0.564898	2.769691	0.171397		
Residual	4	0.815829	0.203957	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	0.613677	-0.410118	1.637472	2.353589	0.171396
H4K8ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	0.033196	0.033196	0.081586	0.789341		
Residual	4	1.627521	0.40688	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	0.148763	-1.297265	1.594791	0.403945	0.799942
H4_4_17_unmod							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	13.042236	13.042236	0.769211	0.429979		
Residual	4	67.821381	16.955345	NaN	NaN		

Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	2.948699	-6.385924	12.283322	1.240331	0.42998
H4di_ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	58.688766	58.688766	21.819766	0.00951		
Residual	4	10.758826	2.689706	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	6.255065	2.537181	9.97295	6.606022	0.00951
H4tetra_ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	0.099205	0.099205	0.895367	0.397607		
Residual	4	0.443194	0.110799	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	0.257171	-0.497418	1.01176	1.338183	0.397606
H4tri_ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	1	7.653816	7.653816	2.944157	0.161331		
Residual	4	10.398654	2.599663	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	wtpaba	tno2a25	2.258881	-1.396242	5.914004	2.426585	0.16133

H4di-ac: H4K5acK8ac, H4K5acK12ac, H4K5acK16ac, H4K8acK16ac, and H4K12acK16ac
H4tri-ac: H4K5acK8ac12ac, H4K5acK8acK16ac, H4K5acK12acK16ac, and H4K8acK12acK16ac
H4tetra-ac: H4K5acK8ack12acK16ac

Supplementary Table 5 One-way ANOVA test and Tukey HSD test for all modifications in four *Aspergillus* strains

H3K14ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	23.267066	7.755689	0.25749	0.854031		
Residual	8	240.962505	30.120313	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	0.083841	-14.267669	14.43535	0.02646	0.9
1	A. niger	A. nidulans (wtpaba)	3.448216	-10.903293	17.799726	1.088242	0.854079
2	A. niger	A. nidulans (tn02a25)	1.327639	-13.023871	15.679148	0.418997	0.9
3	A. fumigatus	A. nidulans (wtpaba)	3.364376	-10.987134	17.715885	1.061782	0.863649
4	A. fumigatus	A. nidulans (tn02a25)	1.243798	-13.107711	15.595307	0.392537	0.9
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	2.120578	-12.230932	16.472087	0.669245	0.9
H3K18ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	0.856596	0.285532	1.560215	0.272985		
Residual	8	1.464064	0.183008	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	0.539175	-0.579498	1.657847	2.183006	0.458938
1	A. niger	A. nidulans (wtpaba)	0.651026	-0.467646	1.769699	2.635871	0.314065
2	A. niger	A. nidulans (tn02a25)	0.156789	-0.961883	1.275462	0.634806	0.9
3	A. fumigatus	A. nidulans (wtpaba)	0.111852	-1.006821	1.230524	0.452865	0.9
4	A. fumigatus	A. nidulans (tn02a25)	0.382386	-0.736287	1.501058	1.5482	0.687696
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	0.494237	-0.624435	1.61291	2.001064	0.523878
H3K18acK23ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	41.410595	13.803532	1.746346	0.23483		
Residual	8	63.233901	7.904238	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	4.531552	-2.820318	11.883422	2.791757	0.273169
1	A. niger	A. nidulans (wtpaba)	3.608125	-3.743745	10.959996	2.222861	0.445009
2	A. niger	A. nidulans (tn02a25)	4.477607	-2.874263	11.829477	2.758523	0.281534
3	A. fumigatus	A. nidulans (wtpaba)	0.923427	-6.428443	8.275297	0.568896	0.9
4	A. fumigatus	A. nidulans (tn02a25)	0.053945	-7.297925	7.405815	0.033234	0.9
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	0.869482	-6.482388	8.221352	0.535662	0.9
H3K23ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	147.590081	49.196694	1.480062	0.291677		
Residual	8	265.916885	33.239611	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	7.466318	-7.610018	22.542654	2.243051	0.438008
1	A. niger	A. nidulans (wtpaba)	8.256181	-6.820155	23.332517	2.480344	0.35977
2	A. niger	A. nidulans (tn02a25)	2.039738	-13.036598	17.116074	0.612783	0.9

3	A. fumigatus	A. nidulans (wtpaba)	0.789863	-14.286473	15.866199	0.237293	0.9
4	A. fumigatus	A. nidulans (tn02a25)	5.42658	-9.649756	20.502916	1.630268	0.658009
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	6.216443	-8.859893	21.292779	1.86756	0.572171
H3K27ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	140.064874	46.688291	1.220839	0.363409		
Residual	8	305.942433	38.242804	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	3.990379	-12.180838	20.161596	1.117635	0.843447
1	A. niger	A. nidulans (wtpaba)	5.335889	-10.835328	21.507106	1.494488	0.707124
2	A. niger	A. nidulans (tn02a25)	9.567502	-6.603715	25.738718	2.679688	0.30215
3	A. fumigatus	A. nidulans (wtpaba)	1.345511	-14.825706	17.516727	0.376854	0.9
4	A. fumigatus	A. nidulans (tn02a25)	5.577123	-10.594094	21.74834	1.562054	0.682685
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	4.231612	-11.939605	20.402829	1.1852	0.819006
H3K36me1							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	2302.113306	767.3711	28.766287	0.000123		
Residual	8	213.408452	26.676056	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	22.918503	9.412444	36.424562	7.685747	0.002756
1	A. niger	A. nidulans (wtpaba)	6.237913	-7.268146	19.743973	2.091891	0.491168
2	A. niger	A. nidulans (tn02a25)	25.424708	11.918649	38.930767	8.526205	0.001412
3	A. fumigatus	A. nidulans (wtpaba)	29.156416	15.650357	42.662476	9.777638	0.001
4	A. fumigatus	A. nidulans (tn02a25)	2.506205	-10.999854	16.012264	0.840459	0.9
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	31.662621	18.156562	45.168681	10.618097	0.001
H3K4me1							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	13.817236	4.605745	3.490717	0.06989		
Residual	8	10.555414	1.319427	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	1.725616	-1.278111	4.729343	2.602029	0.323484
1	A. niger	A. nidulans (wtpaba)	0.612298	-2.391429	3.616025	0.923274	0.9
2	A. niger	A. nidulans (tn02a25)	1.245552	-1.758175	4.249279	1.878148	0.56834
3	A. fumigatus	A. nidulans (wtpaba)	1.113318	-1.890409	4.117046	1.678755	0.640469
4	A. fumigatus	A. nidulans (tn02a25)	2.971168	-0.032559	5.974895	4.480177	0.052506
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	1.85785	-1.145878	4.861577	2.801422	0.270772
H3K4me2							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	37.829933	12.609978	8.952891	0.006167		
Residual	8	11.267849	1.408481	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	0.350288	-2.753152	3.453728	0.511222	0.9

1	A. niger	A. nidulans (wtpaba)	1.664359	-1.439081	4.767798	2.429025	0.375933
2	A. niger	A. nidulans (tn02a25)	4.513444	1.410004	7.616884	6.587084	0.007072
3	A. fumigatus	A. nidulans (wtpaba)	1.314071	-1.789369	4.417511	1.917803	0.553996
4	A. fumigatus	A. nidulans (tn02a25)	4.163156	1.059716	7.266596	6.075861	0.011256
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	2.849086	-0.254354	5.952525	4.158059	0.072426
H3K4me3							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	53.20019	17.733397	2.978849	0.09645		
Residual	8	47.62483	5.953104	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	2.701084	-3.679193	9.081362	1.917463	0.554119
1	A. niger	A. nidulans (wtpaba)	0.75474	-5.625537	7.135018	0.53578	0.9
2	A. niger	A. nidulans (tn02a25)	3.145161	-3.235117	9.525438	2.232707	0.441588
3	A. fumigatus	A. nidulans (wtpaba)	1.946344	-4.433933	8.326622	1.381683	0.74793
4	A. fumigatus	A. nidulans (tn02a25)	5.846245	-0.534032	12.226522	4.15017	0.072999
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	3.899901	-2.480377	10.280178	2.768487	0.279003
H3K79me1							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	409.324932	136.44164	1.825962	0.220487		
Residual	8	597.785242	74.723155	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	10.223404	-12.381124	32.827931	2.048465	0.506731
1	A. niger	A. nidulans (wtpaba)	2.704122	-19.900406	25.30865	0.541825	0.9
2	A. niger	A. nidulans (tn02a25)	10.118516	-12.486012	32.723044	2.027449	0.514333
3	A. fumigatus	A. nidulans (wtpaba)	12.927526	-9.677002	35.532053	2.59029	0.32683
4	A. fumigatus	A. nidulans (tn02a25)	0.104888	-22.49964	22.709416	0.021016	0.9
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	12.822638	-9.78189	35.427165	2.569274	0.332939
H3K79me2							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	4324.595188	1441.5317	3.208768	0.083206		
Residual	8	3593.981274	449.24766	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	7.102243	-48.323401	62.527887	0.580381	0.9
1	A. niger	A. nidulans (wtpaba)	35.564869	-19.860775	90.990513	2.90629	0.245855
2	A. niger	A. nidulans (tn02a25)	13.49855	-41.927094	68.924194	1.103074	0.848713
3	A. fumigatus	A. nidulans (wtpaba)	42.667112	-12.758532	98.092756	3.486671	0.141037
4	A. fumigatus	A. nidulans (tn02a25)	6.396307	-49.029337	61.821951	0.522693	0.9
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	49.063419	-6.362225	104.48906	4.009364	0.084025
H3K9ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	22.549856	7.516619	5.061395	0.029664		

Residual	8	11.880706	1.485088	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	1.169171	-2.017549	4.355892	1.661738	0.646626
1	A. niger	A. nidulans (wtpaba)	0.680793	-2.505927	3.867514	0.967608	0.897715
2	A. niger	A. nidulans (tn02a25)	2.935441	-0.25128	6.122161	4.17213	0.071414
3	A. fumigatus	A. nidulans (wtpaba)	1.849964	-1.336756	5.036685	2.629346	0.315866
4	A. fumigatus	A. nidulans (tn02a25)	1.76627	-1.420451	4.95299	2.510392	0.350533
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	3.616234	0.429513	6.802954	5.139738	0.027394
H3K9acK14ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	17.133284	5.711095	8.390195	0.007478		
Residual	8	5.445494	0.680687	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	2.442102	0.284648	4.599555	5.126851	0.027741
1	A. niger	A. nidulans (wtpaba)	1.136618	-1.020836	3.294072	2.386171	0.389779
2	A. niger	A. nidulans (tn02a25)	0.713018	-1.444435	2.870472	1.496882	0.706261
3	A. fumigatus	A. nidulans (wtpaba)	1.305483	-0.85197	3.462937	2.74068	0.286107
4	A. fumigatus	A. nidulans (tn02a25)	3.15512	0.997666	5.312573	6.623733	0.006844
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	1.849636	-0.307817	4.00709	3.883053	0.095279
H3K9me1							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	2.08958	0.696527	2.33421	0.150259		
Residual	8	2.387194	0.298399	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	1.003835	-0.42462	2.43229	3.182908	0.189417
1	A. niger	A. nidulans (wtpaba)	0.382298	-1.046158	1.810753	1.21217	0.80925
2	A. niger	A. nidulans (tn02a25)	0.956203	-0.472252	2.384659	3.03188	0.218708
3	A. fumigatus	A. nidulans (wtpaba)	0.621537	-0.806918	2.049993	1.970738	0.534849
4	A. fumigatus	A. nidulans (tn02a25)	0.047632	-1.380824	1.476087	0.151028	0.9
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	0.573906	-0.85455	2.002361	1.819711	0.589481
H3K9me1K14ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	0.988211	0.329404	7.645799	0.009796		
Residual	8	0.344664	0.043083	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	0.77257	0.229794	1.315346	6.446823	0.008021
1	A. niger	A. nidulans (wtpaba)	0.414485	-0.128291	0.957261	3.458726	0.144974
2	A. niger	A. nidulans (tn02a25)	0.194219	-0.348557	0.736995	1.62069	0.661475
3	A. fumigatus	A. nidulans (wtpaba)	0.358086	-0.18469	0.900862	2.988096	0.227892
4	A. fumigatus	A. nidulans (tn02a25)	0.578351	0.035575	1.121127	4.826132	0.037252
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	0.220265	-0.322511	0.763041	1.838036	0.582851

H3K9me2							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	2.425005	0.808335	9.477464	0.005194		
Residual	8	0.682322	0.08529	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	1.151204	0.387513	1.914895	6.827525	0.005718
1	A. niger	A. nidulans (wtpaba)	0.39515	-0.36854	1.158841	2.343546	0.403849
2	A. niger	A. nidulans (tn02a25)	0.922049	0.158358	1.68574	5.46846	0.019936
3	A. fumigatus	A. nidulans (wtpaba)	0.756054	-0.007637	1.519744	4.483979	0.052307
4	A. fumigatus	A. nidulans (tn02a25)	0.229155	-0.534536	0.992846	1.359065	0.756112
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	0.526899	-0.236792	1.290589	3.124914	0.20023
H3K9me2K14ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	0.111526	0.037175	1.094953	0.405573		
Residual	8	0.271612	0.033952	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	0.23718	-0.244654	0.719014	2.22951	0.4427
1	A. niger	A. nidulans (wtpaba)	0.007335	-0.474499	0.489169	0.068948	0.9
2	A. niger	A. nidulans (tn02a25)	0.048921	-0.432913	0.530755	0.459861	0.9
3	A. fumigatus	A. nidulans (wtpaba)	0.229845	-0.251989	0.71168	2.160562	0.466838
4	A. fumigatus	A. nidulans (tn02a25)	0.188259	-0.293575	0.670093	1.769649	0.60759
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	0.041586	-0.440248	0.52342	0.390913	0.9
H3K9me3							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	112.434942	37.478314	24.7543	0.000211		
Residual	8	12.112098	1.514012	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	6.741723	3.52412	9.959327	9.490014	0.001
1	A. niger	A. nidulans (wtpaba)	0.789625	-2.427978	4.007229	1.111519	0.845658
2	A. niger	A. nidulans (tn02a25)	0.085792	-3.131812	3.303395	0.120765	0.9
3	A. fumigatus	A. nidulans (wtpaba)	7.531349	4.313745	10.748952	10.601533	0.001
4	A. fumigatus	A. nidulans (tn02a25)	6.827515	3.609911	10.045119	9.610779	0.001
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	0.703833	-2.51377	3.921437	0.990754	0.889344
H3K9me3K14ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	0.937306	0.312435	9.146156	0.005784		
Residual	8	0.273282	0.03416	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	0.651574	0.168261	1.134887	6.106093	0.010945
1	A. niger	A. nidulans (wtpaba)	0.030083	-0.45323	0.513396	0.281913	0.9
2	A. niger	A. nidulans (tn02a25)	0.00905	-0.474263	0.492363	0.084811	0.9
3	A. fumigatus	A. nidulans (wtpaba)	0.621491	0.138178	1.104804	5.824179	0.014226

4	A. fumigatus	A. nidulans (tn02a25)	0.660624	0.177311	1.143937	6.190904	0.010124
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	0.039133	-0.44418	0.522446	0.366725	0.9
H3_18_26_unmod							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	320.136676	106.71223	3.392279	0.074217		
Residual	8	251.659051	31.457381	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	12.537045	-2.129544	27.203634	3.871636	0.09637
1	A. niger	A. nidulans (wtpaba)	12.515333	-2.151256	27.181922	3.864931	0.097018
2	A. niger	A. nidulans (tn02a25)	6.674134	-7.992455	21.340723	2.061077	0.502168
3	A. fumigatus	A. nidulans (wtpaba)	0.021712	-14.644877	14.688301	0.006705	0.9
4	A. fumigatus	A. nidulans (tn02a25)	5.862911	-8.803678	20.5295	1.810559	0.592792
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	5.841198	-8.825391	20.507787	1.803854	0.595219
H3_27_40_unmod							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	3057.404247	1019.1347	22.564317	0.000294		
Residual	8	361.326153	45.165769	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	26.908882	9.334804	44.48296	6.935082	0.005205
1	A. niger	A. nidulans (wtpaba)	0.902024	-16.672054	18.476102	0.232474	0.9
2	A. niger	A. nidulans (tn02a25)	34.99221	17.418132	52.566287	9.018354	0.001
3	A. fumigatus	A. nidulans (wtpaba)	27.810906	10.236828	45.384983	7.167555	0.004258
4	A. fumigatus	A. nidulans (tn02a25)	8.083328	-9.49075	25.657406	2.083273	0.494236
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	35.894234	18.320156	53.468311	9.250828	0.001
H3_3_8_unmod							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	268.268921	89.422974	4.745486	0.034765		
Residual	8	150.750368	18.843796	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	4.076413	-7.275055	15.427881	1.626502	0.659372
1	A. niger	A. nidulans (wtpaba)	0.297321	-11.054147	11.648789	0.118632	0.9
2	A. niger	A. nidulans (tn02a25)	8.904157	-2.447312	20.255625	3.552787	0.132131
3	A. fumigatus	A. nidulans (wtpaba)	4.373734	-6.977735	15.725202	1.745134	0.616456
4	A. fumigatus	A. nidulans (tn02a25)	12.980569	1.629101	24.332037	5.179289	0.026362
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	8.606836	-2.744632	19.958304	3.434155	0.148518
H3_73_83_unmod							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	2505.82332	835.27444	3.976152	0.052598		
Residual	8	1680.56848	210.07106	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	3.121161	-34.779864	41.022185	0.372987	0.9
1	A. niger	A. nidulans (wtpaba)	32.860747	-5.040277	70.761772	3.926946	0.091186

2	A. niger	A. nidulans (tn02a25)	3.380034	-34.52099	41.281059	0.403923	0.9
3	A. fumigatus	A. nidulans (wtpaba)	29.739586	-8.161438	67.640611	3.553959	0.131979
4	A. fumigatus	A. nidulans (tn02a25)	6.501195	-31.39983	44.40222	0.77691	0.9
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	36.240781	-1.660243	74.141806	4.330869	0.060937
H3_9_17_unmod							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	400.949247	133.64975	3.684063	0.062264		
Residual	8	290.222466	36.277808	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	11.747176	-4.003106	27.497459	3.378109	0.156903
1	A. niger	A. nidulans (wtpaba)	1.19108	-14.559203	16.941363	0.342516	0.9
2	A. niger	A. nidulans (tn02a25)	2.949546	-12.800736	18.699829	0.848194	0.9
3	A. fumigatus	A. nidulans (wtpaba)	12.938256	-2.812026	28.688539	3.720625	0.112038
4	A. fumigatus	A. nidulans (tn02a25)	14.696723	-1.05356	30.447006	4.226303	0.067651
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	1.758466	-13.991816	17.508749	0.505678	0.9
H4K16ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	335.426733	111.80891	6.524468	0.015274		
Residual	8	137.094905	17.136863	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	10.807382	-0.017757	21.632521	4.521843	0.05037
1	A. niger	A. nidulans (wtpaba)	1.832826	-8.992314	12.657965	0.76686	0.9
2	A. niger	A. nidulans (tn02a25)	3.52468	-7.30046	14.349819	1.474737	0.714271
3	A. fumigatus	A. nidulans (wtpaba)	8.974556	-1.850583	19.799696	3.754983	0.108269
4	A. fumigatus	A. nidulans (tn02a25)	14.332062	3.506922	25.157201	5.99658	0.012114
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	5.357505	-5.467634	16.182645	2.241597	0.438509
H4K5(12)ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	0.947967	0.315989	1.028047	0.430251		
Residual	8	2.458945	0.307368	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	0.042993	-1.40677	1.492757	0.134318	0.9
1	A. niger	A. nidulans (wtpaba)	0.684376	-0.765388	2.13414	2.13809	0.474775
2	A. niger	A. nidulans (tn02a25)	0.070699	-1.379064	1.520463	0.220875	0.9
3	A. fumigatus	A. nidulans (wtpaba)	0.641383	-0.808381	2.091146	2.003773	0.522897
4	A. fumigatus	A. nidulans (tn02a25)	0.027706	-1.422058	1.47747	0.086558	0.9
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	0.613677	-0.836087	2.06344	1.917215	0.554207
H4K8ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	1.708108	0.569369	2.140696	0.17323		
Residual	8	2.127791	0.265974	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value

0	A. niger	A. fumigatus	0.775045	-0.573568	2.123658	2.602964	0.323222
1	A. niger	A. nidulans (wtpaba)	0.194753	-1.15386	1.543366	0.654072	0.9
2	A. niger	A. nidulans (tn02a25)	0.04599	-1.302623	1.394603	0.154456	0.9
3	A. fumigatus	A. nidulans (wtpaba)	0.969798	-0.378815	2.318411	3.257036	0.176365
4	A. fumigatus	A. nidulans (tn02a25)	0.821035	-0.527578	2.169648	2.75742	0.281815
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	0.148763	-1.19985	1.497376	0.499616	0.9
H4_4_17_unmod							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	40.019753	13.339918	0.96102	0.45669		
Residual	8	111.047979	13.880997	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	3.317909	-6.424765	13.060583	1.542463	0.68977
1	A. niger	A. nidulans (wtpaba)	1.683079	-8.059595	11.425753	0.782447	0.9
2	A. niger	A. nidulans (tn02a25)	1.26562	-8.477054	11.008294	0.588374	0.9
3	A. fumigatus	A. nidulans (wtpaba)	5.000988	-4.741686	14.743662	2.32491	0.410083
4	A. fumigatus	A. nidulans (tn02a25)	2.05229	-7.690384	11.794964	0.954089	0.9
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	2.948699	-6.793975	12.691373	1.370821	0.75186
H4di_ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	148.697569	49.565856	11.78522	0.002633		
Residual	8	33.646113	4.205764	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	5.076083	-0.286701	10.438866	4.287132	0.063658
1	A. niger	A. nidulans (wtpaba)	1.527987	-3.834797	6.890771	1.290499	0.780914
2	A. niger	A. nidulans (tn02a25)	4.727078	-0.635705	10.089862	3.992372	0.085465
3	A. fumigatus	A. nidulans (wtpaba)	3.548096	-1.814688	8.910879	2.996633	0.226075
4	A. fumigatus	A. nidulans (tn02a25)	9.803161	4.440377	15.165945	8.279504	0.001709
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	6.255065	0.892282	11.617849	5.282871	0.023839
H4tetra_ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		
C(Species)	3	0.749601	0.249867	2.193314	0.166583		
Residual	8	0.911377	0.113922	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	0.601366	-0.28125	1.483982	3.085997	0.20779
1	A. niger	A. nidulans (wtpaba)	0.23959	-0.643026	1.122206	1.229492	0.802983
2	A. niger	A. nidulans (tn02a25)	0.017581	-0.865035	0.900197	0.090218	0.9
3	A. fumigatus	A. nidulans (wtpaba)	0.361776	-0.52084	1.244392	1.856505	0.576171
4	A. fumigatus	A. nidulans (tn02a25)	0.618947	-0.263669	1.501563	3.176216	0.190639
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	0.257171	-0.625445	1.139787	1.31971	0.770347
H4tri_ac							
ANOVA Summary	df	sum_sq	mean_sq	F	PR(>F)		

C(Species)	3	18.157613	6.052538	4.011552	0.051557		
Residual	8	12.070218	1.508777	NaN	NaN		
Tukey HSD Summary	group1	group2	Diff	Lower	Upper	q-value	p-value
0	A. niger	A. fumigatus	2.630063	-0.581973	5.842098	3.708635	0.11338
1	A. niger	A. nidulans (wtpaba)	2.23795	-0.974086	5.449986	3.15572	0.194418
2	A. niger	A. nidulans (tn02a25)	0.020931	-3.191105	3.232967	0.029514	0.9
3	A. fumigatus	A. nidulans (wtpaba)	0.392112	-2.819924	3.604148	0.552915	0.9
4	A. fumigatus	A. nidulans (tn02a25)	2.650993	-0.561043	5.863029	3.73815	0.110098
5	A. nidulans (wtpaba)	A. nidulans (tn02a25)	2.258881	-0.953155	5.470917	3.185235	0.188994

H4di-ac: H4K5acK8ac, H4K5acK12ac, H4K5acK16ac, H4K8acK16ac, and H4K12acK16ac
H4tri-ac: H4K5acK8ac12ac, H4K5acK8acK16ac, H4K5acK12acK16ac, and H4K8acK12acK16ac
H4tetra-ac: H4K5acK8ack12acK16ac

Data availability

Data statement: All supporting data, code, and protocols have been provided within the article or through supplementary data files. Supplementary material is available with the online version of this article. The MS results are deposited in the PRIDE database with the dataset identifier PXD033478.

Acknowledgments

We thank Jos Houbraken for providing *A. fumigatus* strain and Joseph Strauss for providing two *A. nidulans* strains.

Funding information

Xin Zhang is funded by the Chinese Scholarship Council (CSC) (201907720028). This work was supported by EPIC-XS, project number 823 839, funded by the Horizon 2020 programme of the European Union.

Author contributions

X.Z. formal analysis, investigation, methodology, visualization, writing original draft; R.N. and A.V. formal analysis, investigation, methodology; T.B. funding acquisition, supervision, resources; J.C. conceptualization, funding acquisition, methodology, project administrations, supervision, visualization, writing original draft; M.F.S. conceptualization, funding acquisition, methodology, project administrations, supervision, visualization, writing original draft.

References

Abshiru, N., Rajan, R. E., Verreault, A. & Thibault, P. (2016). Unraveling site-specific and combinatorial histone modifications using high-resolution mass spectrometry in histone deacetylase mutants of fission yeast. *Journal of Proteome Research*, *15*(7), 2132–2142.

Adlakha, A., Armstrong-James, D. A. J. & Lenhard, B. (2016). T1 calcineurin inhibition impairs the dendritic cell transcriptional response to *Aspergillus fumigatus* infection in lung transplant recipients. *BMJ*, *71*(3).

Barrett, K., Jensen, K., Meyer, A. S., Frisvad, J. C. & Lange, L. (2020). Fungal secretome profile categorization of CAZymes by function and family corresponds to fungal phylogeny and taxonomy: Example *Aspergillus* and *Penicillium*. *Scientific Reports*, *10*(1), 1–12.

Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I. & Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell*, *129*(4), 823–837.

Bedre, R. (2020). *Reneshbedre/bioinfokit: Bioinformatics data analysis and visualization toolkit*.

Bhaumik, S. R., Smith, E. & Shilatifard, A. (2007). Covalent modifications of histones during development and disease pathogenesis. *Nature Structural & Molecular Biology*, *14*(11), 1008–1016.

Bicocca, V. T., Ormsby, T., Adhvaryu, K. K., Honda, S. & Selker, E. U. (2018). ASH1-catalyzed H3K36 methylation drives gene repression and marks H3K27me2/3-competent chromatin. *ELife*, *7*, 1–19.

Bremang, M., Cuomo, A., Agresta, A. M., Stugiewicz, M., Spadotto, V. & Bonaldi, T. (2013). Mass spectrometry-based identification and characterisation of lysine and arginine methylation in the human proteome. *Molecular BioSystems*, *9*(9), 2231–2247.

Brosch, G., Loidl, P. & Graessle, S. (2008). Histone modifications and chromatin dynamics: A focus on filamentous fungi. *FEMS Microbiology Reviews*, *32*(3), 409–439.

Cánovas, D., Marcos, A. T., Gacek, A., Ramos, M. S., Gutiérrez, G., Reyes-Domínguez, Y. & Strauss, J. (2014). The histone acetyltransferase GcnE (GCN5) plays a central role in the regulation of *Aspergillus* asexual development. *Genetics*, *197*(4), 1175–1189.

Chen, X., Wu, L., Lan, H., Sun, R., Wen, M., Ruan, D., Zhang, M. & Wang, S. (2022). Histone acetyltransferases MystA and MystB contribute to morphogenesis and aflatoxin biosynthesis by regulating acetylation in fungus *Aspergillus flavus*. *Environmental Microbiology*, *24*(3), 1340–1361.

Cheng, J., Park, T.-S., Chio, L.-C., Fischl, A. S. & Ye, X. S. (2003). Induction of apoptosis by sphingoid long-chain bases in *Aspergillus nidulans*. *Molecular and Cellular Biology*, *23*(1), 163–177.

Colabardini, A. C., Wang, F., Miao, Z., Pardeshi, L., Valero, C., de Castro, P. A., Akiyama, D. Y., Tan, K., Nora, L. C., Silva-Rocha, R., Marcet-Houben, M., Gabaldón, T., Fill, T., Wong, K. H. & Goldman, G. H. (2022). Chromatin profiling reveals heterogeneity in clinical isolates of the human pathogen *Aspergillus fumigatus*. In *PLoS Genetics* (Vol. 18, Issue 1).

Collemare, J. & Seidl, M. F. (2019). Chromatin-dependent regulation of secondary metabolite biosynthesis in fungi: is the picture complete? *FEMS Microbiology Reviews*, *43*(6), 591–607.

Connolly, L. R., Smith, K. M. & Freitag, M. (2013). The *Fusarium graminearum* histone H3 K27 methyltransferase KMT6 regulates development and expression of secondary metabolite gene clusters. *PLoS Genetics*, *9*(10), e1003916.

De Souza, C. P. C., Osmani, A. H., Wu, L.-P., Spotts, J. L. & Osmani, S. A. (2000). Mitotic histone H3 phosphorylation by the NIMA kinase in *Aspergillus nidulans*. *Cell*, *102*(3), 293–302.

Dion, M. F., Altschuler, S. J., Wu, L. F. & Rando, O. J. (2005). Genomic characterization reveals a simple histone H4 acetylation code. *Proceedings of the National Academy of Sciences*, *102*(15), 5501–5506.

Farooq, Z., Banday, S., Pandita, T. K. & Altaf, M. (2016). The many faces of histone H3K79 methylation. *Mutation Research - Reviews in Mutation Research*, *768*, 46–52.

Flury, V., Reverón-Gómez, N., Alcaraz, N., Stewart-Morgan, K. R., Wenger, A., Klose, R. J. & Groth, A. (2023). Recycling of modified H2A-H2B provides short-term memory of chromatin

states. *Cell*, 1–16.
Freitag, M. (2017). Histone methylation by SET domain proteins in fungi. *Annual Review of Microbiology*, *71*, 413–439.
Gacek-Matthews, A., Berger, H., Sasaki, T., Wittstein, K., Gruber, C., Lewis, Z. A. & Strauss, J. (2016). KdmB, a Jumonji Histone H3 Demethylase, Regulates Genome-Wide H3K4 Trimethylation and Is Required for Normal Induction of Secondary Metabolism in *Aspergillus nidulans*. *PLoS Genetics*, *12*(8), 1–29.
Gacek-Matthews, A., Noble, L. M., Gruber, C., Berger, H., Sulyok, M., Marcos, A. T., Strauss, J. & Andrianopoulos, A. (2015). KdmA, a histone H3 demethylase with bipartite function, differentially regulates primary and secondary metabolism in *Aspergillus nidulans*. *Molecular Microbiology*, *96*(4), 839–860.
Gates, L. A., Shi, J., Rohira, A. D., Feng, Q., Zhu, B., Bedford, M. T., Sagum, C. A., Jung, S. Y., Qin, J. & Tsai, M.-J. (2017). Acetylation on histone H3 lysine 9 mediates a switch from transcription initiation to elongation. *Journal of Biological Chemistry*, *292*(35), 14456–14472.
Gnesutta, N., Nardini, M. & Mantovani, R. (2013). The H2A/H2B-like histone-fold domain proteins at the crossroad between chromatin and different DNA metabolisms. *Transcription*, *4*(3), 114–119.
Grant, P. A., Eberharter, A., John, S., Cook, R. G., Turner, B. M. & Workman, J. L. (1999). Expanded lysine acetylation specificity of Gcn5 in native complexes. *Journal of Biological Chemistry*, *274*(9), 5895–5900.
Grau-Bové, X., Navarrete, C., Chiva, C., Pribasnig, T., Antó, M., Torruella, G., Galindo, L. J., Lang, B. F., Moreira, D. & López-Garcia, P. (2022). A phylogenetic and proteomic reconstruction of eukaryotic chromatin evolution. *Nature Ecology & Evolution*, 1–17.
Guillemette, B., Drogaris, P., Lin, H. H. S., Armstrong, H., Hiragami-Hamada, K., Imhof, A., Bonneil, É., Thibault, P., Verreault, A. & Festenstein, R. J. (2011). H3 lysine 4 is acetylated at active gene promoters and is regulated by H3 lysine 4 methylation. *PLoS Genetics*, *7*(3).
Han, J., Zhang, H., Zhang, H., Wang, Z., Zhou, H. & Zhang, Z. (2013). A Cul4 E3 ubiquitin ligase regulates histone hand-off during nucleosome assembly. *Cell*, *155*(4), 817.
Han, X., Tian, M., Shliaha, P. V, Zhang, J., Jiang, S., Nan, B., Alam, M. N., Jensen, O. N., Shen, H. & Huang, Q. (2021). Real-world particulate matters induce lung toxicity in rats fed with a high-fat diet: Evidence of histone modifications. *Journal of Hazardous Materials*, *416*, 126182.
He, F. (2011). Bradford protein assay. *Bio-Protocol*, 45.
Henikoff, S. & Shilatifard, A. (2011). Histone modification: cause or cog? *Trends in Genetics*, *27*(10), 389–396.
Huang, H., Lin, S., Garcia, B. A. & Zhao, Y. (2015). Quantitative proteomic analysis of histone modifications. *Chemical Reviews*, *115*(6), 2376–2418.
Janevska, S., Baumann, L., Sieber, C. M. K., Münsterkötter, M., Ulrich, J., Kämper, J., Güldener, U. & Tudzynski, B. (2018). Elucidation of the two H3K36me3 histone methyltransferases Set2 and Ash1 in *Fusarium fujikuroi* unravels their different chromosomal targets and a major impact of Ash1 on genome stability. *Genetics*, *208*, 153–171.
Johnson, L., Mollah, S., Garcia, B. A., Muratore, T. L., Shabanowitz, J., Hunt, D. F. & Jacobsen, S. E. (2004). Mass spectrometry analysis of Arabidopsis histone H3 reveals distinct combinations of post-translational modifications. *Nucleic Acids Research*, *32*(22), 6511–6518.
Jurkowska, R. Z., Qin, S., Kungulovski, G., Tempel, W., Liu, Y., Bashtrykov, P., Stiefelmaier, J., Jurkowski, T. P., Kudithipudi, S. & Weirich, S. (2017). H3K14ac is linked to methylation of H3K9 by the triple Tudor domain of SETDB1. *Nature Communications*, *8*(1), 1–13.
Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature*, *410*(6824), 116–120.
Lin, S., Wein, S., Gonzales-Cope, M., Otte, G. L., Yuan, Z. F., Afjehi-Sadat, L., Maile, T., Berger, S. L., Rush, J., Lill, J. R., Arnott, D. & Garcia, B. A. (2014). Stable-isotope-labeled histone peptide library for histone post-translational modification and variant quantification by mass spectrometry. *Molecular and Cellular Proteomics*, *13*(9), 2450–2466.
Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. (1997). Crystal structure of the nucleosome resolution core particle at 2.8 Å. *Nature*, *389*, 251–260.

Maile, T. M., Izrael-Tomasevic, A., Cheung, T., Guler, G. D., Tindell, C., Masselot, A., Liang, J., Zhao, F., Trojer, P., Classon, M. & Arnott, D. (2015). Mass spectrometric quantification of histone post-translational modifications by a hybrid chemical labeling method. *Molecular and Cellular Proteomics*, *14*(4), 1148–1158.

Mersfelder, E. L. & Parthun, M. R. (2006). The tale beyond the tail: Histone core domain modifications and the regulation of chromatin structure. *Nucleic Acids Research*, *34*(9), 2653–2662.

Nishioka, K., Rice, J. C., Sarma, K., Erdjument-Bromage, H., Werner, J., Wang, Y., Chuikov, S., Valenzuela, P., Tempst, P., Steward, R., Lis, J. T., Allis, C. D. & Reinberg, D. (2002). PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin. *Molecular Cell*, *9*(6), 1201–1213.

Noberini, R., Restellini, C., Savoia, E. O. & Bonaldi, T. (2020). Enrichment of histones from patient samples for mass spectrometry-based analysis of post-translational modifications. *Methods*, *184*, 19–28.

Noberini, R., Savoia, E. O., Brandini, S., Greco, F., Marra, F., Bertalot, G., Pruneri, G., McDonnell, L. A. & Bonaldi, T. (2021). Spatial epi-proteomics enabled by histone post-translational modification analysis from low-abundance clinical samples. *Clinical Epigenetics*, *13*(1), 1–16.

Nützmann, H.-W., Fischer, J., Scherlach, K., Hertweck, C. & Brakhage, A. A. (2013). Distinct amino acids of histone H3 control secondary metabolism in *Aspergillus nidulans*. *Applied and Environmental Microbiology*, *79*(19), 6102–6109.

Nützmann, H.-W., Reyes-Dominguez, Y., Scherlach, K., Schroeckh, V., Horn, F., Gacek, A., Schümann, J., Hertweck, C., Strauss, J. & Brakhage, A. A. (2011). Bacteria-induced natural product formation in the fungus *Aspergillus nidulans* requires Saga/Ada-mediated histone acetylation. *Proceedings of the National Academy of Sciences*, *108*(34), 14282–14287.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R. & Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, *12*, 2825–2830.

Pesavento, J. J., Mizzen, C. A. & Kelleher, N. L. (2006). Quantitative analysis of modified proteins and their positional isomers by tandem mass spectrometry: human histone H4. *Analytical Chemistry*, *78*(13), 4271–4280.

Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Tong, I. L., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D. K. & Young, R. A. (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell*, *122*(4), 517–527.

Reyes-Dominguez, Y., Bok, J. W., Berger, H., Shwab, E. K., Basheer, A., Gallmetzer, A., Scazzocchio, C., Keller, N. & Strauss, J. (2010). Heterochromatic marks are associated with the repression of secondary metabolism clusters in *Aspergillus nidulans*. *Molecular Microbiology*, *76*(6), 1376–1386.

Reyes-Dominguez, Y., Narendja, F., Berger, H., Gallmetzer, A., Fernandez-Martin, R., Garcia, I., Scazzocchio, C. & Strauss, J. (2008). Nucleosome positioning and histone H3 acetylation are independent processes in the *Aspergillus nidulans* prnD-prnB bidirectional promoter. *Eukaryotic Cell*, *7*(4), 656–663.

Reyes-Dominguez, Y., Bok, J. W., Berger, H., Shwab, E. K., Basheer, A., Gallmetzer, A., Scazzocchio, C., Keller, N. & Strauss, J. (2010). Heterochromatic marks are associated with the repression of secondary metabolism clusters in *Aspergillus nidulans*. *Molecular Microbiology*, *76*(6), 1376–1386.

Rountree, M. R. & Selker, E. U. (2010). DNA methylation and the formation of heterochromatin in *Neurospora crassa*. *Heredity*, *105*(1), 38–44.

Scheid, R., Dowell, J. A., Sanders, D., Jiang, J., Denu, J. M. & Zhong, X. (2022). Histone Acid Extraction and High Throughput Mass Spectrometry to Profile Histone Modifications in *Arabidopsis thaliana*. *Current Protocols*, *2*(8), e527.

Shilatifard, A. (2006). Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu. Rev. Biochem.*, *75*, 243–269.

Shinohara, Y., Kawatani, M., Futamura, Y., Osada, H. & Koyama, Y. (2016). An overproduction of

astellolides induced by genetic disruption of chromatin-remodeling factors in *Aspergillus oryzae*. *The Journal of Antibiotics*, *69*(1), 4–8.

Sindikubwabo, F., Ding, S., Hussain, T., Ortet, P., Barakat, M., Baumgarten, S., Cannella, D., Palencia, A. s., Bougdour, A., Belmudes, L., Couté, Y., Tardieux, I., Botté, C. Y., Scherf, A. & Hakimi, M. A. (2017). Modifications at K31 on the lateral surface of histone H4 contribute to genome structure and expression in apicomplexan parasites. *ELife*, *6*, 1–34.

Tan, M., Luo, H., Lee, S., Jin, F., Yang, J. S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., Lu, Z., Ye, Z., Zhu, Q., Wysocka, J., Ye, Y., Khochbin, S., Ren, B. & Zhao, Y. (2011). Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell*, *146*(6), 1016–1028.

van Holde, K. E. (1989). Chromatin structure and transcription. In *Chromatin* (pp. 355–408). Springer.

van Leeuwen, F., Gafken, P. R. & Gottschling, D. E. (2002). Dot1p modulates silencing in yeast by methylation of the nucleosome core. *Cell*, *109*(6), 745–756.

Vizcaíno, J. A., Deutsch, E. W., Wang, R., Csordas, A., Reisinger, F., Rios, D., Dianes, J. A., Sun, Z., Farrah, T. & Bandeira, N. (2014). ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nature Biotechnology*, *32*(3), 223–226.

Waskom, M. L. (2021). Seaborn: statistical data visualization. *Journal of Open Source Software*, *6*(60), 3021.

Wiemann, P., Sieber, C. M. K., Von Bargen, K. W., Studt, L., Niehaus, E.-M., Espino, J. J., Huss, K., Michielse, C. B., Albermann, S. & Wagner, D. (2013). Deciphering the cryptic genome: genome-wide analyses of the rice pathogen *Fusarium fujikuroi* reveal complex regulation of secondary metabolism and novel metabolites. *PLoS Pathogens*, *9*(6), e1003475.

Yamada, T., Fischle, W., Sugiyama, T., Allis, C. D. & Grewal, S. I. S. (2005). The nucleation and maintenance of heterochromatin by a histone deacetylase in fission yeast. *Molecular Cell*, *20*(2), 173–185.

Young, T. J., Cui, Y., Pfeffer, C., Hobbs, E., Liu, W., Irudayaraj, J. & Kirchmaier, A. L. (2020). CAF-1 and Rtt101p function within the replication-coupled chromatin assembly network to promote H4 K16ac, preventing ectopic silencing. In *PLoS Genetics* (Vol. 16, Issue 12).

Yuan, Z.-F., Sidoli, S., Marchione, D. M., Simithy, J., Janssen, K. A., Szurgot, M. R. & Garcia, B. A. (2018). EpiProfile 2.0: a computational platform for processing epi-proteomics mass spectrometry data. *Journal of Proteome Research*, *17*(7), 2533–2541.

Zhang, X., Noberini, R., Bonaldi, T., Collemare, J. & Seidl, M. F. (2022). The histone code of the fungal genus *Aspergillus* uncovered by evolutionary and proteomic analyses. *Microbial Genomics*, *8*.

Zhang, X., Noberini, R., Bonaldi, T. & Seidl, M. F. (2021). Histone extraction for mass spectrometry-based analysis of post-translational modifications in the fungal genus *Aspergillus*. *Dx.Doi.Org/10.17504/Protocols.Io.Btpinmke*.

Zhang, Y., Fan, J., Ye, J. & Lu, L. (2021). The fungal-specific histone acetyltransferase Rtt109 regulates development, DNA damage response, and virulence in *Aspergillus fumigatus*. *Molecular Microbiology*, *115*(6), 1191–1206.

Zhou, M., Malhan, N., Ahkami, A. H., Engbrecht, K., Myers, G., Dahlberg, J., Hollingsworth, J., Sievert, J. A., Hutmacher, R., Madera, M., Lemaux, P. G., Hixson, K. K., Jansson, C. & Paša-Tolić, L. (2020). Top-down mass spectrometry of histone modifications in sorghum reveals potential epigenetic markers for drought acclimation. *Methods*, *184*(October 2019), 29–39.

Zhuang, Z., Pan, X., Zhang, M., Liu, Y., Huang, C., Li, Y., Hao, L. & Wang, S. (2022). Set2 family regulates mycotoxin metabolism and virulence via H3K36 methylation in pathogenic fungus *Aspergillus flavus*. *Virulence*, *13*(1), 1358–1378.

Chapter 4 - Appendix

Histone extraction for mass spectrometry-based analysis of post-translational modifications in the fungal genus *Aspergillus*

Xin Zhang, Roberta Noberini, Tiziana Bonaldi, Michael F. Seidl, Jérôme Collemare

DOI: dx.doi.org/10.17504/protocols.io.btpinmke

Abstract

Mass spectrometry is a powerful technique to uncover histone post-translational modifications (PTMs) in an unbiased and quantitative manner (Noberini et al., 2021). Yet, such analyses require good-quality protein samples, a challenge when working with fungi. Here, we present a protocol based on the work by Noberini and colleagues (Noberini et al., 2020) to extract histone proteins from filamentous fungi for applications like mass spectrometry-based analyses of histone PTMs. Several modifications were made: 1) additional cell disruption and homogenization steps using freeze-drying and TissueLyser II; 2) increased Triton X-100 concentration in the lysis buffer; and 3) prolonged incubation time in the lysis buffer. This protocol is reproducible, and the proteins extracted from three *Aspergillus* species (*Aspergillus niger*, *Aspergillus fumigatus*, and *Aspergillus nidulans*) were of good quality for mass spectrometry.

Guidelines

1. Use the proper medium to culture fungal strains.
2. Germinated spores from overnight cultures are the ideal starting material.
3. Disruption by TissueLyser II and sonication are necessary steps because fungal cells are protected by the cell wall which prevents efficient lysis.
4. 100 mg ground starting material per sample of *Aspergillus* ends up with enough yield (around 20 µg) for Bradford Protein Assay (He, 2011), SDS-PAGE (Sodium dodecyl sulfate polyacrylamide gel electrophoresis) visualization (Simpson, 2006), and mass spectrometry-based analysis (which ideally requires 3-4 µg of histones). Different fungal species may vary.

Materials

Solutions and reagents

- ACES buffer (1.822 g/L N-(2-Acetamido)-2-aminoethanesulfonic acid, 0.2 mL/L Tween 80)
- Solid medium and liquid medium depending on the strain (*A. niger* and *A. fumigatus* are cultured in MEA (Malt Extract Agar Oxoid 50 g/L) medium at 30 ℃ and 25 ℃, respectively; *A. nidulans* is cultured in YAG medium (5 g/L yeast extract, 15 g/L agar, 20 g/L D-glucose, 1 mL/L trace elements (Solution A: Dissolve 10 g EDTA and 1 g $FeSO_4 \times 7H_2O$ in 80 mL water and adjust pH to 5.5; Solution B: Dissolve 4.4 g $ZnSO_4 \times 7H_2O$, 1 g $MnCl_2 \times 4H_2O$, 0.32 g $CuSO_4 \times 6H_2O$, and 0.22 g $(NH_4)_6Mo_7O_{24} \times 4H_2O$ in 80 mL water; Combine Solution A and B, adjust pH to 6.5) at 37 ℃
- PBS (Phosphate Buffered Saline) buffer (8.0 g/L NaCl, 0.2 g/L KCl, 1.44 g/L $Na_2HPO_4 \times 2\ H_2O$, 0.2 g/L KH_2PO_4, 1.32 g/L $CaCL_2 \times 2H_2O$, and 0.5 mM $MgCl_2 \times 6H_2O$)
- Lysis buffer (1× PBS, 0.5 mM PMSF, 5 µM Leupeptin, 5 µM Aprotinin, 5 mM Na-butyrate, and 1% or 10% Triton X-100)
- Aprotinin bovine (Sigma, A6103)
- Leupeptin (Sigma, L2884)
- Sodium butyrate (Sigma, B5887)
- Phenylmethylsulfonyl fluoride (PMSF; Fluka, 78830)
- Triton X-100 (Sigma, T8787)
- Acetone (Biosolve,1046021)
- Coomassie protein assay reagent (Thermo, 1856209)

Equipment

- Centrifuge (Eppendorf, 5810R)
- Centrifuge (Eppendorf, 5430R)
- Hemacytometer (Marienfeld, 0.0025mm^2)
- Freeze dryer (Christ, Alpha 1-4 LSCbasic)

- TissueLyser II (QIAGEN, C.0659)
- Orbital shaker (Grant-bio, PS-3D)
- Incubator (Yourtech, C.0663)
- Vortex (IKA, S000)
- Ultrasonic Cleaner (Branson 2510)

Before start instructions

1. Set centrifuges at 4 ℃ beforehand.
2. Acetone has strong flammability, always operate it inside a fume hood.
3. PBS buffer should be kept on ice.
4. Prepare lysis buffers with different Triton X-100 concentrations fresh and add PMSF last. PMSF stock solution in isopropanol can be stored at -20 ℃ for several months, but it can be inactivated within 35 minutes once it is added to the aqueous solution.

Method

Sample collection and disruption

1. Harvest fungal spores from 4- or 5-day-old sporulating plates with a spreader in 12.5 mL ice-cold ACES buffer.
2. Centrifuge the spore suspension at 2,500 xg at 4 ℃ for 5 minutes.
3. Discard supernatant and resuspend spores in 25 mL ice-cold ACES buffer.
4. Centrifuge the spore suspension at 2,500 xg at 4 ℃ for 5 minutes and discard the supernatant.
5. Resuspend spores in 10 mL ice-cold ACES buffer, and count spore concentration with a hemacytometer or spore counter. Ideally, the spore concentration is around 1×10^7 spores/ mL.
6. Inoculate 50 mL of liquid medium with $1 \sim 10 \times 10^7$ spores in a 250 mL flask, and incubate at 30 ℃, 200 rpm, for 16 hours.
7. Collect germinated spores by centrifugation at 4,000 xg for 10 minutes and discard the supernatant.
8. Resuspend germinated spore pellet in 25 mL 0.8 M NaCl.
9. Centrifuge at 4,000 xg for 10 minutes and discard supernatant to get rid of liquid medium.
10. Freeze germinated spores in liquid nitrogen.
11. Freeze dry overnight.
12. Grind freeze-dried samples with a mortar and pestle, and then with TissueLyser II using two ⌀5 mm metal beads at 30 Hz for 30 seconds, which is repeated twice.

Histone protein extraction

1. Resuspend 100 mg ground sample in 1 mL lysis buffer with high Triton X-100 concentration (1×PBS, 0.5 mM PMSF, 5 μM Leupeptin, 5 μM Aprotinin, 5mM Na-butyrate, and 10% Triton X-100).
2. Incubate on an orbital shaker for 30 minutes at 4 ℃.
3. Add 10 ml lysis buffer with low Triton X-100 concentration (1×PBS, 0.5 mM PMSF, 5 μM Leupeptin, 5 μM Aprotinin, 5mM Na-butyrate, and 1% Triton X-100), mix by vortexing, and incubate on the orbital shaker at 4 ℃ for another 30 minutes.
4. Centrifuge at 3,250 xg for 30 minutes at 4 ℃, and discard supernatant.
5. Resuspend in 1 mL ice-cold PBS buffer.
6. Sonicate for 30 seconds; this process is repeated 10 times with 30 seconds break in between.
7. Add 9 mL ice-cold PBS buffer and mix by vortexing.
8. Centrifuge at 3,250 xg for 20 minutes at 4 ℃, and discard supernatant.
9. Resuspend the pellet in 2 mL 0.4 N H_2SO_4 and incubate at 4 ℃ for 4 hours on an orbital shaker.
10. Centrifuge at 16,000 xg for 10 minutes at 4 ℃. The supernatant contains soluble core histone proteins.
11. Add 4 volumes of acetone to the sample and keep at -20 ℃ overnight to precipitate histone proteins.
12. Centrifuge at 16,000 xg for 25 minutes at 4 ℃ to pellet proteins, then dissolve in 100 μL water.

Protein quantity and quality measurement

1. Measure protein concentration with the Bradford Protein Assay.
2. Check protein quality via 16% SDS-PAGE gel.

References

He, F. (2011). Bradford protein assay. *Bio-Protocol*, 45.

Noberini, R., Restellini, C., Savoia, E. O. & Bonaldi, T. (2020). Enrichment of histones from patient samples for mass spectrometry-based analysis of post-translational modifications. *Methods*, *184*, 19–28.

Noberini, R., Robusti, G. & Bonaldi, T. (2021). Mass spectrometry-based characterization of histones in clinical samples: applications, progresses, and challenges. *FEBS Journal*, 1–23.

Simpson, R. J. (2006). SDS-PAGE of proteins. *Cold Spring Harbor Protocols*, *2006*(1), pdb-prot4313.

Chapter 5

General discussion

Fungal secondary metabolites (SMs) are small bioactive molecules with complex structures. They are not essential for normal fungal growth or reproduction but play critical roles in the interactions of fungi with other organisms or in response to environmental stimuli (Keller, 2019). Thanks to the biochemical versatility of SMs, humankind has utilized SMs for centuries in medicine, agriculture, food, and cosmetics (Mosunova et al., 2021). With the increasing accessibility of fungal genome sequences due to the development of next-generation sequencing (NGS) technologies, researchers have started to uncover an immense diversity of biosynthetic pathways that are predicted to synthesize SMs (Lind et al., 2017). However, only a limited number of these SMs have been discovered, studied, and extensively used to date (Mosunova et al., 2021). This discrepancy between the SMs experimentally discovered and the high number of biosynthetic pathways encoded in fungal genomes is due to our inability to accurately and efficiently express genes involved in the biosynthesis of SMs in the laboratory (Brakhage, 2013).

To address this bottleneck and to enable humankind to harness the as-of-yet largely untapped diversity of fungal SMs, it is therefore important to understand the regulatory mechanisms governing the expression of genes responsible for the biosynthesis of SMs. In **chapter 1**, I explained that the genes responsible for SM biosynthesis are often physically co-localized and co-regulated, an arrangement referred to as biosynthetic gene clusters (BGCs). The biosynthesis of SMs is regulated by a multi-tier regulatory mechanism, including both local and global transcription factors (TFs), and histone post-transcriptional modifications (PTMs). Several studies focused on a limited number of histone PTMs have shown that they are involved in the regulation of the biosynthesis of SMs in a few, well-known model fungal species (Yang et al., 2022). However, we still lack a complete understanding of how histone PTMs regulate SM biosynthesis in both well- and less-studied fungal species. In this thesis, I employed a combination of computational analyses (**chapters 2** and **3**) and wet-lab experiments (**chapters 3** and **4**), focusing on the fungal genus *Aspergillus*, an ensemble of diverse species that are known to be prolific producers of SMs (Soltani, 2016), with the overall aim to provide broader and deeper insights into genes involved in the biosynthesis of SMs and the roles of histone PTMs in regulating the expression of these genes.

In the last decade, studies have investigated the abundance and diversity of BGCs within a few so-called taxonomic sections of the *Aspergillus* genus (Kjærbølling et al., 2020; Vesth et al., 2018). For example, a study comparing 32 Aspergilli in the *Nigri* section revealed over 400 predicted BGCs, of which 49% are species-specific (Vesth et al., 2018). Similarly, another computational study in the *Flavi* section, which included 23 Aspergilli, uncovered 308 BGCs, with 150 being considered to be species-specific (Kjærbølling et al., 2020), suggesting that BGCs are numerous and diverse in Aspergilli. The *Aspergillus* genus is an ancient and large taxonomic group with an estimated 350 species, which diversified from the common ancestor shared with *Penicillium* species around 82 million years ago (Steenwyk et al., 2019). To obtain

a complete overview of BGC diversity, it is therefore crucial to explore and compare BGCs in more species and across diverse *Aspergillus* sections. In **chapter 2,** I provided a systematics exploration, employing computational approaches, to assess the BGC diversity and abundance in 174 *Aspergillus* species across 22 sections. I specifically focused on the SM biosynthetic backbone genes as these catalyze the crucial step of SM biosynthesis from primary precursors to the first stable intermediate (Mosunova et al., 2021). Interestingly, we found a high abundance and diversity of SM backbone genes across the *Aspergillus* genus, with notable diversity increases between sections. Moreover, it has been shown that SM backbone genes are often regulated by histone PTMs, and studies in a few fungal species indicated an enrichment of SM backbone genes in sub-telomeric regions (Keller, 2019; Klejnstrup et al., 2012; Perrin et al., 2007; Yang et al., 2022). Leveraging publicly available RNA-seq and ChIP-seq (*Ch*romatin *I*mmuno*P*recipitation assays followed by next-generation *seq*uencing) data, I studied the correlation between SM backbone genes' genomic localization, gene expression, and histone PTMs. We found that SM backbone genes showed a strong bias for being localized in low synteny regions, regardless of their position in chromosomes or being unique or conserved across many species. Using chromosome-level genome assemblies, we also confirmed a significantly biased localization in sub-telomeric regions. Notably, SM backbone genes in sub-telomeric regions and about half of those in other low-syntenic regions exhibit higher gene expression variability, likely linked to similar higher variability in specific histone PTMs.

Most studies on histone PTMs in Aspergilli have solely focused on the well-characterized modifications of methylation and acetylation (Collemare & Seidl, 2019). Considering that other histone PTMs have been identified in different model species (Ahringer & Gasser, 2018; Boros, 2012; Jiang et al., 2007; Kimura, 2013) and that diverse combinations of PTMs can lead to distinct regulatory outcomes (Mosunova et al., 2021), our understanding of histone PTMs' roles in SM biosynthesis is not yet complete. To obtain an overview of different histone PTMs that can occur in Aspergilli, I performed both computational and experimental research. First, I conducted a systematic phylogenetic analysis of genes related to histone modifiers that mainly regulate histone PTMs (**chapter 3).** I investigated known genes involved in the establishment, maintenance, and readout of histone PTMs from well-studied fungal species and systematically assessed whether they are also present in the genomes of Aspergilli, which provides a proxy for the presence and function of the associated histone PTMs. Based on these results, I proposed that Aspergilli have the ability to perform most of the known histone PTMs as the genes responsible for these PTMs are encoded in their genomes. I subsequently provided experimental evidence for these computational predictions using mass spectrometry (**chapter 4),** which was facilitated by an experimental protocol, originally developed for human primary cells, that I optimized for filamentous fungi. The absence of detection of H3K27me (methylation on histone protein H3 at lysine 27) is consistent with my prediction of the loss of genes encoding enzymes involved in the deposition of H3K27me in the *Aspergillus* genus

(**chapters 3** and **4**). Mass spectrometry is an unbiased and quantitative method that allows for the detection and quantification of different types of histone PTMs among various species. In **chapter 4**, to my knowledge, I reported for the first time the detection of H3K79me1, H3K79me2, and H4K31ac in Aspergilli. Moreover, I discovered quantitative differences in the relative abundance of H3K9me1/2/3, H3K14ac, H3K36me1, and H3K79me1, and the co-occurrence of acetylations on both K18 and K23 of histone H3 in a species- and/or strain-specific manner. This suggests that the Aspergilli tested in this study utilize the same types of histone PTMs, albeit in a distinct manner. In this final chapter, I will further discuss the implications of my work and highlight promising developments that may affect our views on the regulation of SM biosynthesis.

The 'histone code' in Aspergilli

Chromatin organizes DNA within the eukaryotic nucleus and is mainly formed by DNA that is wrapped around a histone protein octamer composed of two copies of each of the histone proteins H2A, H2B, H3, and H4 (Baldi et al., 2020). PTMs of histone proteins can occur on the core fold of the proteins but predominantly occur on the N-terminal tails that protrude from the DNA-histone complex, the nucleosome (Luger et al., 1997). These modifications include not only well-studied PTMs such as mono-, di-, and tri-methylation, acetylation, phosphorylation, and ubiquitylation, but also less-studied modifications such as crotonylation, malonylation, succinylation, or propionylation (Bowman & Poirier, 2014). Studies have shown that both individual histone PTMs and distinct combinations of these modifications on the same histone proteins can influence the chromatin structure, allowing transitions between a 'loose' and transcriptionally active state and a 'compact', transcriptionally silent state (Collemare & Seidl, 2019). Histone PTMs at a genomic locus can synergistically or antagonistically act with other chromatin-associated proteins, which in turn are thought to drive these dynamic transitions that regulate gene expression to cope with developmental and environmental challenges (Jenuwein & Allis, 2001). The term 'histone code' was coined more than 20 years ago as an extension of the 'genetic (DNA) code' concept (Jenuwein & Allis, 2001), and since then has been further refined by studies in a few model eukaryotic species such as plant (Loidl, 2004), nematode (Ahringer & Gasser, 2018), fruit fly (Boros, 2012), and yeast (Jiang et al., 2007). Beyond these model species, the histone code in other organisms, including Aspergilli, remains largely unexplored. In this thesis, I provide substantial evidence that Aspergilli possess a histone code that shows remarkable conservation across various species and strains (**chapter 4**). Interestingly, the histone code in Aspergilli is also quite different compared to other model fungal species such as the filamentous fungus *Neurospora crassa*, the plant pathogen *Fusarium graminearum*, or the budding yeast *Saccharomyces cerevisiae* (**chapter 3**) (Zhang et al., 2022). For example, *S. cerevisiae* is known to have lost several genes encoding enzymes involved in histone PTMs including H3K9me, H3K27me, and H4K20me (Freitag, 2017), while *N. crassa* harbors all of them (Navarro-Mendoza et al., 2023; Zhang et al., 2022). This highlights the importance of recognizing that histone codes can differ across

species, necessitating careful considerations when extrapolating knowledge from one model species to another less studied species. In fact, both computational predictions and wet-lab experiments are likely needed to be certain of the histone code in any specific species of interest.

My research revealed that in Aspergilli, genes related to both facultative and constitutive heterochromatin (transcriptionally inactive) have largely been lost, while those associated with euchromatin (transcriptionally active) remain conserved (**chapters 3** and **4**). For instance, H3K27me3 is associated with facultative heterochromatin, and it plays crucial roles in the timely and accurate gene expression regulation in other organisms (Ridenour et al., 2020). In *F. graminearum*, it regulates fungal development and the biosynthesis of SMs (Connolly et al., 2013), and in the plant pathogen *Verticillium dahliae*, local rather than global H3K27me3 regulates gene expression in different culture conditions (Kramer et al., 2022). Interestingly, all genes of the catalytic complex PRC2 (Polycomb Repressive Complex 2) and the readers associated with H3K27me3 such as EPR-1 or BP1 (Tang et al., 2021; Wiles et al., 2020; Zhang et al., 2023) are lost throughout the *Aspergillus* genus (**chapters 3** and **4**) as well as in three *Penicillium* species that we considered in our study. Given the important function of H3K27me3 in facultative heterochromatin formation in other species (Moser Tralamazza et al., 2022), Aspergilli might rely on other histone PTMs to play a similar role to H3K27me3. Possible candidates could be H3K36me2/3 and H4K20me3, which we identified using mass spectrometry (**chapter 4**). H3K36me2/3 is catalyzed by Ash-1, a paralog to the conserved SET2 (**chapter 3**). In the plant pathogen *Fusarium fujikuroi*, H3K36me2/3 is catalyzed by Ash-1 that co-occurs with H3K27me3 at facultative heterochromatic regions (Janevska et al., 2018). Moreover, this histone modification is known to induce the formation of H3K27me in *N. crassa* (Bicocca et al., 2018). In chapters 3 and 4, I found that *Ash-1* is present in most Aspergilli, and I also detected the presence of H3K36me2/3 on histone proteins. To test whether Ash-1 catalyzed H3K36me2/3 is directly related to facultative heterochromatin in Aspergilli, however, necessitates future experiments. For instance, ChIP-seq can identify genes marked by H3K36me2/3, and ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) can reveal the chromatin accessibility of these genes. For the other candidate, H4K20me3, a recent paper on wheat pathogen *Zymoseptoria tritici* indicated that this mark can introduce and co-occur with H3K36me3 catalyzed Ash-1 and H3K27me3 catalyzed SET7 (Möller et al., 2023). The gene encoding its catalytic enzyme, *SET9* (*KMT5*) is conserved in most of the *Aspergillus* species tested in my study (**chapter 3**), but I did not detect the presence of this histone PTM *via* mass spectrometry (**chapter 4**). This could be caused by the fact that the abundance of H4K20me in the conditions tested was too low to be detected. Moreover, as it has been shown that H4K20me and H4K16ac are competitive *in vitro* (Nishioka et al., 2002), and we indeed detected H4K16ac *via* mass spectrometry, and thus it is conceivable that H4K16ac inhibited H4K20me. More conditions need to be tested to see whether H4K20me is indeed present in Aspergilli and its specific roles will need to be tested

in mutants lacking *Ash-1* and *SET2*. With this evidence, we will gain more understanding of the roles of H3K36me2/3 and H4K20me in facultative heterochromatin formation in Aspergilli and whether they regulate SM biosynthesis. In regard to BGC expression, of the most studied histone PTMs, H3K4me3 and H3K36me3 indeed appear to be important (**chapter 2**). In contrast, H3K9me3 and H3Kac are very stable PTMs that do not appear to explain variation in gene expression (**chapter 2**).

Chromatin modifications related to constitutive heterochromatin have been shown to be linked with methylations on the DNA cytosine and also on H3K9 (Grewal & Jia, 2007). In *N. crassa*, constitutive heterochromatin is formed by the deposition of H3K9me3, followed by the recognition by the adapter protein HP1 that then recruits the DNA methyltransferase Dim-2 to perform *de novo* DNA methylation at the H3K9me3-rich regions to further entrench the heterochromatin state (Rountree & Selker, 2010). In Aspergilli, even though two subunits that are part of the catalytic complex DCDC introducing H3K9me are lost in Aspergilli (**chapter 3**), I detected H3K9me1/2/3 *via* mass spectrometry (**chapter 4**). Moreover, in Aspergilli, the genes encoding HP1, and the DNA methyltransferase related to the maintenance of DNA methylation (DNMT5) are conserved, but not the ones for *de novo* DNA methylation (Dim-2 and DNMTX) (**chapter 3**). A study in the pathogenic fungus *Cryptococcus neoformans* shows that *DNMTX* was lost in an ancestral species around 50-150 million years ago (Catania et al., 2020). Dim-2 is a more recently derived *de novo* DNA methyltransferase across several groups in the fungal kingdom except the *Aspergillus* genus (Bewick et al., 2019). It is very likely that Aspergilli can only maintain DNA methylation but cannot introduce *de novo* DNA methylation to heterochromatin labeled by H3K9me. Thus, the heterochromatin state in Aspergilli is likely limited to the methylations on H3K9. Interestingly, it has been shown that genomic regions with high coverage of H3K9me are not associated with SM gene clusters in *Aspergillus fumigatus*, as only two out of 33 BGCs have H3K9me peaks within the clusters and only five BGCs have H3K9me peaks nearby (Colabardini et al., 2022). Consistently, we observed that H3K9me3 does not regulate the expression of SM backbone genes in *Aspergillus nidulans* and *A. fumigatus* (**chapter 2**). Given these insights, future research should explore alternative epigenetic factors that operate constitutive heterochromatin independently of H3K9me3 in controlling SM gene expression in Aspergilli.

In contrast to the dispensability of histone PTMs related to heterochromatin, modifications linked to euchromatin in Aspergilli are generally well conserved (**chapters 3** and **4**). Besides the typical euchromatin marks such as H3K4me1/2/3 catalyzed by SET1, H3K36me1/2/3 catalyzed by SET2, and acetylation on histone proteins H3 and H4 (Strauss & Reyes-Dominguez, 2011), I also detected the presence of H3K79me1 and H3K79me2 for the first time in Aspergilli. The catalytic gene encoding the enzyme that deposits H3K79me1/2/3, *Dot1* (disruptor of telomeric silencing 1), was first found in *S. cerevisiae* and its overexpression disrupted telomeric silencing (Singer et al., 1998). Further study in *S. cerevisiae* shows it

marks the boundaries between heterochromatin and euchromatin, which acts together with acetylation on H4K16 to prevent the spreading of the SIR (Silent Information Regulator) complex with the function of heterochromatin formation (Farooq et al., 2016). In filamentous fungi, the knowledge of the potential roles of H3K79me remains limited. In *Aspergillus flavus*, the *⊿Dot1* mutant exhibits severely impaired aflatoxin production, but the effect on other genes related to SM production and the changes in H3K79me were not systematically explored (Liang et al., 2017). In my study, I observed significant differences in the relative abundance of H3K79me1 in different *Aspergillus* species (**chapter 4**). To further reveal the function of H3K79me in the regulation of SM biosynthesis, ChIP-seq needs to be performed in Aspergilli to check whether SM genes are decorated by this mark and how it changes under distinct conditions. Specifically, it is essential to determine the chromosomal localization of H3K79me in Aspergilli. It is particularly interesting to investigate whether H3K79me is enriched at the boundaries between heterochromatin and euchromatin, similar to observations in *S. cerevisiae* (Farooq et al., 2016), near or within sub-telomeric regions where some SM genes are located (**chapter 2**), or within low-syntenic regions that were found to be enriched in SM backbone genes (**chapter 2**). Understanding the spatial distribution of H3K79me will provide valuable insights into its functional roles in regulating SM biosynthesis in Aspergilli.

Scale matters

In the realm of high-throughput analyses, both the scale and the conditions under which data is collected influence our interpretation of biological processes. In my studies, I employed techniques such as RNA-seq, ChIP-seq, and mass spectrometry, which generated data from bulk *Aspergillus* cell populations (**chapters 2**, **3**, and **4**). These methods provide an averaged view of biological signals and mask the intricate heterogeneity and unique patterns present at the single-cell level. Moreover, by investigating SM backbone genes' genomic localization, I demonstrated that these genes preferentially locate at the low-syntenic regions and seem to locate at the sub-telomeric regions. However, it is important to note that this is based on the four chromosome-level assemblies available to date (**chapter 2**), and consequently, it currently remains unclear how these broad patterns are conserved throughout the genus – a question that can only be addressed once additional genomes with sufficient quality will become available. Similarly, I only cultured Aspergilli in a single condition to obtain the pattern of histone PTMs in three *Aspergillus* species (**chapter 4**), yet it is important to acknowledge that different cultural conditions will most likely yield distinct patterns. For example, the production of aflatoxin or penicillin can be significantly influenced by the substrate used, the pH of the environment, or the temperature during cultivation (Keller et al., 2005). For studying SM gene expression regulation, I utilized publicly available RNA-seq data that had been collected under various conditions (**chapter 2**). Future studies should track gene expression changes by comparing conditions of SM induction medium with normal growth medium, thereby investigating the gene regulation mechanisms of SM genes more thoroughly.

When studying Aspergilli, the genetic and phenotypic diversity among different strains, species, and taxonomic sections introduces multiple layers of complexity. We observed that the differences in relative abundances of histone PTMs are as pronounced within species as it is between species (**chapter 4**). This finding echoes research on the rice and wheat blast fungus *Magnaporthe oryzae* where machine learning methods revealed that different strains exhibit distinct H3K27 methylation and expression levels (Joubert & Krasileva, 2024). Consequently, a model trained on one strain proved inapplicable to another strain (Joubert & Krasileva, 2024). Coupled with our discovery that SM diversity is pronounced across sections but diminishes within individual sections, these insights underscore the importance of carefully defining the scope of comparison when drawing biological conclusions. Such consideration ensures that the inherent diversity of Aspergilli is accurately reflected and accounted for in research findings.

Concluding remarks

In this thesis, I have investigated histone PTMs as key regulatory mechanisms in SM biosynthesis within the *Aspergillus* genus. This study employed an integrative approach, combining computational analysis with wet-lab experiments. Initially, I revealed the high abundance and diversity of SM backbone genes across distinct *Aspergillus* sections. I also showed that SM backbone genes are preferentially located in low-syntenic regions and that their gene regulation is governed by specific histone PTMs. Furthermore, I analyzed the occurrences and quantitative differences of histone PTMs across various *Aspergillus* species and strains, identifying several marks with potential critical roles in SM biosynthesis regulation. For future investigations, I suggest scientists focus more on SM genes located at the low-syntenic regions and sub-telomeric regions because BGCs in these regions are more likely to be expressed and many remain uncharacterized. When comparing different Aspergillus strains, species, or sections, it is important to consider that they exhibit distinct patterns and vary in their scope. Three histone PTMs, H3K36me2/3 catalyzed by Ash-1, H4K20me3, and methylations on H3K79 could be pivotal for the regulation of SM biosynthesis in Aspergilli. The anticipated future results combining ChIP-seq, ATAC-seq, and single-cell sequencing are likely to provide a deeper understanding of this regulation mechanism. This knowledge could unlock the vast potential of SMs for human benefit, advancing our understanding of epigenetic regulation in fungal biology.

References

Ahringer, J. & Gasser, S. M. (2018). Repressive chromatin in *Caenorhabditis elegans*: establishment, composition, and function. *Genetics*, *208*(2), 491–511.

Baldi, S., Korber, P. & Becker, P. B. (2020). Beads on a string—nucleosome array arrangements and folding of the chromatin fiber. *Nature Structural & Molecular Biology*, *27*(2), 109–118.

Bewick, A. J., Hofmeister, B. T., Powers, R. A., Mondo, S. J., Grigoriev, I. V., James, T. Y., Stajich, J. E. & Schmitz, R. J. (2019). Diversity of cytosine methylation across the fungal tree of life. *Nature Ecology and Evolution*, *3*(3), 479–490.

Bicocca, V. T., Ormsby, T., Adhvaryu, K. K., Honda, S. & Selker, E. U. (2018). ASH1-catalyzed H3K36 methylation drives gene repression and marks H3K27me2/3-competent chromatin. *ELife*, *7*, 1–19.

Boros, I. M. (2012). Histone modification in *Drosophila*. *Briefings in Functional Genomics*, *11*(4), 319–331.

Bowman, G. D. & Poirier, M. G. (2014). Post-translational modifications of histones that influence nucleosome dynamics. *Chemical Reviews*, *115*(6), 2274–2295.

Brakhage, A. A. (2013). Regulation of fungal secondary metabolism. *Nature Reviews Microbiology*, *11*(1), 21–32.

Catania, S., Dumesic, P. A., Pimentel, H., Nasif, A., Stoddard, C. I., Burke, J. E., Diedrich, J. K., Cook, S., Shea, T., Geinger, E., Lintner, R., Yates, J. R., Hajkova, P., Narlikar, G. J., Cuomo, C. A., Pritchard, J. K. & Madhani, H. D. (2020). Evolutionary persistence of DNA methylation for millions of years after ancient loss of a *de novo* methyltransferase. *Cell*, *180*(2), 263-277.e20.

Colabardini, A. C., Wang, F., Miao, Z., Pardeshi, L., Valero, C., de Castro, P. A., Akiyama, D. Y., Tan, K., Nora, L. C., Silva-Rocha, R., Marcet-Houben, M., Gabaldón, T., Fill, T., Wong, K. H. & Goldman, G. H. (2022). Chromatin profiling reveals heterogeneity in clinical isolates of the human pathogen *Aspergillus fumigatus*. In *PLoS Genetics* (Vol. 18, Issue 1).

Collemare, J. & Seidl, M. F. (2019). Chromatin-dependent regulation of secondary metabolite biosynthesis in fungi: is the picture complete? *FEMS Microbiology Reviews*, *43*(6), 591–607.

Connolly, L. R., Smith, K. M. & Freitag, M. (2013). The *Fusarium graminearum* histone H3 K27 methyltransferase KMT6 regulates development and expression of secondary metabolite gene clusters. *PLoS Genetics*, *9*(10), e1003916.

Farooq, Z., Banday, S., Pandita, T. K. & Altaf, M. (2016). The many faces of histone H3K79 methylation. *Mutation Research - Reviews in Mutation Research*, *768*, 46–52.

Freitag, M. (2017). Histone methylation by SET domain proteins in fungi. *Annual Review of Microbiology*, *71*, 413–439.

Grewal, S. I. S. & Jia, S. (2007). Heterochromatin revisited. *Nature Reviews Genetics*, *8*(1), 35–46.

Janevska, S., Baumann, L., Sieber, C. M. K., Münsterkötter, M., Ulrich, J., Kämper, J., Güldener, U. & Tudzynski, B. (2018). Elucidation of the two H3K36me3 histone methyltransferases Set2 and Ash1 in *Fusarium fujikuroi* unravels their different chromosomal targets and a major impact of Ash1 on genome stability. *Genetics*, *208*, 153–171.

Jenuwein, T. & Allis, C. D. (2001). Translating the histone code. *Science*, *293*(5532), 1074–1080.

Jiang, L., Smith, J. N., Anderson, S. L., Ma, P., Mizzen, C. A. & Kelleher, N. L. (2007). Global assessment of combinatorial post-translational modification of core histones in yeast using contemporary mass spectrometry: LYS4 trimethylation correlates with degree of acetylation on the same H3 tail. *Journal of Biological Chemistry*, *282*(38), 27923–27934.

Joubert, P. M. & Krasileva, K. V. (2024). Distinct genomic contexts predict gene presence-absence variation in different pathotypes of a fungal plant pathogen. *Genetics*, iyae012.

Keller, N. P. (2019). Fungal secondary metabolism: regulation, function and drug discovery. *Nature Reviews Microbiology*, *17*(3), 167–180.

Keller, N. P., Turner, G. & Bennett, J. W. (2005). Fungal secondary metabolism - From biochemistry to genomics. *Nature Reviews Microbiology*, *3*(12), 937–947.

Kimura, H. (2013). Histone modifications for human epigenome analysis. *Journal of Human Genetics*, *58*(7), 439–445.

Kjærbølling, I., Vesth, T., Frisvad, J. C., Nybo, J. L., Theobald, S., Kildgaard, S., Petersen, T. I., Kuo, A., Sato, A., Lyhne, E. K., Kogle, M. E., Wiebenga, A., Kun, R. S., Lubbers, R. J. M., Mäkelä, M. R., Barry, K., Chovatia, M., Clum, A., Daum, C., … Andersen, M. R. (2020). A comparative genomics study of 23 *Aspergillus* species from section *Flavi*. *Nature Communications*, *11*(1).

Klejnstrup, M. L., Frandsen, R. J. N., Holm, D. K., Nielsen, M. T., Mortensen, U. H., Larsen, T. O. & Nielsen, J. B. (2012). Genetics of polyketide metabolism in *Aspergillus nidulans*. *Metabolites*, *2*(1), 100–133.

Kramer, H. M., Seidl, M. F., Thomma, B. P. H. J. & Cook, D. E. (2022). Local rather than global H3K27me3 dynamics associates with differential gene expression in *Verticillium dahliae*. *MBio*, *13*(1), e0356621.

Liang, L., Liu, Y., Yang, K., Lin, G., Xu, Z., Lan, H., Wang, X. & Wang, S. (2017). The putative histone methyltransferase DOT1 regulates aflatoxin and pathogenicity attributes in *Aspergillus flavus*. *Toxins*, *9*(7), 6–8.

Lind, A. L., Wisecaver, J. H., Lameiras, C., Wiemann, P., Palmer, J. M., Keller, N. P., Rodrigues, F., Goldman, G. H. & Rokas, A. (2017). Drivers of genetic diversity in secondary metabolic gene clusters within a fungal species. *PLoS Biology*, *15*(11), 1–26.

Loidl, P. (2004). A plant dialect of the histone language. *Trends in Plant Science*, *9*(2), 84–90.

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. (1997). Crystal structure of the nucleosome resolution core particle at 2.8 Å. *Nature*, *389*, 251–260.

Möller, M., Ridenour, J. B., Wright, D. F., Martin, F. A. & Freitag, M. (2023). H4K20me3 is important for Ash1-mediated H3K36me3 and transcriptional silencing in facultative heterochromatin in a fungal pathogen. *PLoS Genetics*, *19*(9).

Moser Tralamazza, S., Nanchira Abraham, L., Reyes-Avila, C. S., Corrêa, B. & Croll, D. (2022). Histone H3K27 methylation perturbs transcriptional robustness and underpins dispensability of highly conserved genes in fungi. *Molecular Biology and Evolution*, *39*(1).

Mosunova, O., Navarro-Muñoz, J. C. & Collemare, J. (2021). The biosynthesis of fungal secondary metabolites: From fundamentals to biotechnological applications. In *Encyclopedia of mycology* (pp. 458–476). Elsevier.

Navarro-Mendoza, M. I., Pérez-Arques, C. & Heitman, J. (2023). Heterochromatin and RNAi act independently to ensure genome stability in Mucorales human fungal pathogens. *Proceedings of the National Academy of Sciences*, *120*(7).

Nishioka, K., Rice, J. C., Sarma, K., Erdjument-Bromage, H., Werner, J., Wang, Y., Chuikov, S., Valenzuela, P., Tempst, P., Steward, R., Lis, J. T., Allis, C. D. & Reinberg, D. (2002). PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin. *Molecular Cell*, *9*(6), 1201–1213.

Perrin, R. M., Fedorova, N. D., Jin, W. B., Cramer, R. A., Wortman, J. R., Kim, H. S., Nierman, W. C. & Keller, N. P. (2007). Transcriptional regulation of chemical diversity in *Aspergillus fumigatus* by LaeA. *PLoS Pathogens*, *3*(4), 508–517.

Ridenour, J. B., Möller, M. & Freitag, M. (2020). Polycomb repression without bristles: Facultative heterochromatin and genome stability in fungi. *Genes*, *11*(6), 1–23.

Rountree, M. R. & Selker, E. U. (2010). DNA methylation and the formation of heterochromatin in *Neurospora crassa*. *Heredity*, *105*(1), 38–44.

Singer, M. S., Kahana, A., Wolf, A. J., Meisinger, L. L., Peterson, S. E., Goggin, C., Mahowald, M. & Gottschling, D. E. (1998). Identification of high-copy disruptors of telomeric silencing in *Saccharomyces cerevisiae*. *Genetics*, *150*(2), 613–632.

Soltani, J. (2016). Secondary metabolite diversity of the genus *Aspergillus*: recent advances. In *New and Future Developments in Microbial Biotechnology and Bioengineering*. Elsevier B.V.

Steenwyk, J. L., Shen, X. X., Lind, A. L., Goldman, G. H. & Rokas, A. (2019). A robust phylogenomic time tree for biotechnologically and medically important fungi in the genera *Aspergillus* and *Penicillium*. *MBio*, *10*(4), 1–25.

Strauss, J. & Reyes-Dominguez, Y. (2011). Regulation of secondary metabolism by chromatin structure and epigenetic codes. *Fungal Genetics and Biology*, *48*(1), 62–69.

Tang, G., Yuan, J., Wang, J., Zhang, Y. Z., Xie, S. S., Wang, H., Tao, Z., Liu, H., Kistler, H. C., Zhao, Y., Duan, C. G., Liu, W., Ma, Z. & Chen, Y. (2021). Fusarium BP1 is a reader of H3K27 methylation. *Nucleic Acids Research*, *49*(18), 10448–10464.

Vesth, T. C., Nybo, J. L., Theobald, S., Frisvad, J. C., Larsen, T. O., Nielsen, K. F., Hoof, J. B., Brandl, J., Salamov, A., Riley, R., Gladden, J. M., Phatale, P., Nielsen, M. T., Lyhne, E. K., Kogle, M. E., Strasser, K., McDonnell, E., Barry, K., Clum, A., … Andersen, M. R. (2018). Investigation of inter- and intraspecies variation through genome sequencing of *Aspergillus* section *Nigri*. *Nature Genetics*, *50*(12), 1688–1695.

Wiles, E. T., McNaught, K. J., Kaur, G., Selker, J. M. L., Ormsby, T., Aravind, L. & Selker, E. U. (2020). Evolutionarily ancient BAH-PHD protein mediates Polycomb silencing. *PNAS*, *117*(21).

Yang, K., Tian, J. & Keller, N. P. (2022). Post-translational modifications drive secondary metabolite biosynthesis in *Aspergillus*: a review. *Environmental Microbiology*.

Zhang, X., Noberini, R., Bonaldi, T., Collemare, J. & Seidl, M. F. (2022). The histone code of the fungal genus *Aspergillus* uncovered by evolutionary and proteomic analyses. *Microbial Genomics*, *8*.

Zhang, X., Noberini, R., Vai, A., Bonaldi, T., Seidl, M. F. & Collemare, J. (2023). Detection and quantification of the histone code in the fungal genus Aspergillus. *Fungal Genetics and Biology*, *167*, 103800.

Appendix

Summary

Secondary metabolites (SMs) are small bioactive molecules that are produced by diverse groups of organisms. These complex biochemicals are not essential for normal growth and reproduction but play crucial roles in interactions with other organisms and responding to environmental challenges. Fungal SMs have been invaluable treasures to humankind for centuries, being exploited across a myriad of applications, including medicine, agriculture, food, and cosmetics. Their biochemical versatility not only contributes to the development of novel pharmaceuticals but also promotes sustainable farming practices, enhances food preservation, and provides colorful pigments in various products. However, despite the immense diversity and complexity of SMs and their biosynthetic pathways encoded in fungal genomes, only a limited number of SMs have been discovered, extensively researched, and widely utilized to date. This extensive gap, at least in part, stems from our inability to express the enzymes responsible for biosynthesis accurately and efficiently in the laboratory. These enzymes are tightly regulated, as these metabolites are produced only under very specific conditions in nature, conditions that often prove difficult to emulate in laboratory settings. Consequently, the rich repertoire of SMs encoded in fungal genomes and their biochemical activities are not yet fully harnessed.

In **chapter 1** of this book, I introduce the concept of SM and biosynthesis gene clusters (BGCs) that are responsible for the production of SMs. BGCs are distinct groups of co-localized genes that are responsible for the regulation of expression, biosynthesis, and transport of SMs. Subsequently, I delineate the three-tier regulatory mechanisms that govern the expression of SM-BGCs, with a particular emphasis on histone post-translational modifications (PTMs). These chemical modifications affect histone proteins, which are crucial components of the DNA-protein complex called chromatin that organizes the genetic information within the eukaryotic nucleus. Histone modifiers, enzymes responsible for ‘writing’, ‘erasing’, and ‘reading’, regulate the occurrence and alterations of PTMs, thereby orchestrating distinct chromatin organizations within the nucleus. Broadly, chromatin can be divided into an ‘open’ and ‘closed’ state, and these conformations are thought to influence gene expression. In the last decade, a few PTMs have been demonstrated to play crucial roles in providing a global level of transcriptional regulation that also controls the biosynthesis of SMs in a handful of extensively studied fungal species. Here, I propose that PTMs, in general, are key to unlocking the as-of-yet largely unharnessed treasure chest of SMs. I suggest investigating this hypothesis in the *Aspergillus* genus because it is among the most prolific producers of SMs in the fungal kingdom, and such research could provide the basis for a better understanding of the control of SM biosynthesis, which ultimately might provide access to novel biochemistry for the benefit of humankind.

Initially, I explored the abundance and diversity of SM BGCs across the genus *Aspergillus* in **chapter 2**. I first reconstructed a species tree to establish a robust phylogenetic framework for 174 diverse *Aspergillus* species. Subsequently, utilizing a phylogenetic approach, I unveiled the diversity of SM backbone genes, which code for key biosynthetic enzymes. I found that SM backbone genes are abundant in the fungal genus *Aspergillus*, and additionally, I observed marked differences between taxonomic sections. Furthermore, I explored the genomic localizations of SM backbone genes and found that these genes show a strong bias for being localized in low-synteny regions. In chromosome-level genome assemblies of four Aspergilli, I revealed that some SM backbone genes are also located at the sub-telomeric regions. Finally, I probed the relationship between the localization of SM backbone genes and their expression and histone PTMs. With publicly available high-throughput sequencing data, I showed that SM backbone genes in sub-telomeric regions and about half of those localized in other low-synteny regions exhibit higher gene expression variability, likely linked to dynamics in specific histone PTMs.

To enhance our understanding of the evolutionary history of genes responsible for PTM formation and regulation across the *Aspergillus* genus in **chapter 3,** I summarized known histone modifiers from model species in the literature and conducted detailed phylogenetic analyses of these well-characterized protein complexes on a diverse selection of Aspergilli to unravel their evolutionary trajectories involved in the establishment or maintenance of PTMs. The protein complexes contain both catalytic subunits that catalyze the chemical modifications and accessory subunits. The latter contributes to the overall stability of the complex and establishes interactions with other proteins, thereby facilitating coordinated regulation and context-dependent fine-tuning of PTMs. I demonstrated that overall, these complexes are highly conserved in Aspergilli, suggesting that *Aspergillus* species, like other eukaryotes, harbor a complex repertoire of PTMs. Most genes encoding catalytic subunits are conserved, with the exception of *SET7*; additionally, next to *SET7*, genes encoding other subunits of the PRC2 complex are absent in Aspergilli and *Penicillium* species. PRC2 catalyzes the methylation of histone protein H3 at lysine position 27 (H3K27me3), a modification that is typically associated with facultative ‘closed’ chromatin throughout most eukaryotes to timely and accurately regulate gene expression (Ridenour, Möller, and Freitag 2020). The absence of PRC2 suggests that Aspergilli and closely related taxa might have evolved a different mechanism to regulate chromatin conformation and, subsequently, gene expression. In genes encoding accessory subunits in other protein complexes, I discovered both recent gene losses and duplications, underscoring the potential for functional diversification within protein complexes regarding their biochemical or genomic targets. Collectively, these analyses provide an overview of the conservation of key genes related to PTMs in the genus *Aspergillus*, laying the foundation for understanding the mechanistic underpinnings of PTMs in this important fungal genus.

Thus far, most research on PTMs has focused on a small subset of chemical modifications. To confirm the presence of PTMs revealed in **chapter 3** and further explore the occurrence of additional yet unknown PTMs in Aspergilli, I conducted an unbiased and quantitative proteomics study. Utilizing mass spectrometry, this research aims to uncover PTM patterns in three diverse *Aspergillus* species; *Aspergillus niger*, *Aspergillus fumigatus*, and *Aspergillus nidulans*, which belong to three distinct taxonomic sections and exhibit varied lifestyles, pathogenicity, and industrial usage. First, I optimized an experimental protocol to extract high-quality histone proteins from filamentous fungi for mass spectrometry (see Appendix, **chapter 4**). Using this approach, in **chapter 4**, I provided an overview of the occurrence and relative abundance of PTMs. I detected 23 different PTMs, mostly lysine methylations and acetylations. Importantly, I corroborated my computational predictions regarding the presence and absence of PTMs from **chapter 3**. Specifically, I found that the absence of *SET7* leads to non-detectable levels of H3K27me in Aspergilli. For the first time, I reported the detection of H3K79me1, H3K79me2, and H4K31ac in Aspergilli, as well as the co-occurrence of multiple PTMs on the same histone proteins, offering new insights into the complexity of PTMs within these species. Although all three species harbor the same types of PTMs, I discovered significant differences in the relative abundance of H3K9me1/2/3, H3K14ac, H3K36me1, and H3K79me1, as well as the co-occurrence of acetylation on both K18 and K23 of histone H3 in a species and/or strain-specific manner. My exploration of PTMs in Aspergilli not only fortifies existing knowledge but also provides more understanding related to the enigmatic interplay of PTMs and genomic regulation in fungi.

To conclude and provide a perspective for further studies, **chapter 5** introduces the concept of the 'histone code in Aspergilli.' In this chapter, I highlight three different histone PTMs that I have uncovered in my research. I propose that H3K36me2/3, catalyzed by Ash-1, and H4K20me3 may perform analogous functions in compensating for the absence of the facultative heterochromatin mark H3K27me3 in Aspergilli. Additionally, methylation on H3K79, which marks euchromatin, is a promising candidate to play a role in regulating SM biosynthesis due to significant differences in relative abundances among different *Aspergillus* species and strains. These histone modifications might co-localize with SM backbone genes in regions of low synteny. For the discovery and characterization of SM genes, I propose that future researchers pay closer attention to low-synteny and sub-telomeric regions in Aspergilli. Based on my findings, it is conceivable that BGCs in these areas are more likely to be expressed, yet many remain uncharacterized. With the anticipated advancements in research, we now stand on the brink of unlocking the immense potential of SMs and significantly enhancing our understanding of epigenetic regulation in fungal biology.

Samenvatting

Secundaire metabolieten (SM's) zijn kleine bioactieve moleculen die door verschillende groepen organismen worden geproduceerd. Deze complexe biochemische stoffen zijn niet essentieel voor normale groei en voortplanting, maar spelen een cruciale rol in de interacties met andere organismen en in het reageren op ecologische uitdagingen. Schimmel-SM's zijn al eeuwenlang van onschatbare waarde voor de mensheid en worden benut in een breed scala aan toepassingen, waaronder in geneeskunde, landbouw, voeding en cosmetica. Hun biochemische veelzijdigheid draagt niet alleen bij aan de ontwikkeling van nieuwe farmaceutica, maar bevordert ook duurzame landbouw, verbeterde voedselconservering en levert kleurrijke pigmenten voor verschillende producten. Echter, ondanks de immense diversiteit en complexiteit van SM's en hun bio-synthetische pathways gecodeerd in schimmelgenomen, tot op heden is slechts een beperkt aantal SM's ontdekt, uitgebreid onderzocht en breed ingezet. Deze aanzienlijke kloof is ten dele te wijten aan onze onvermogen om de enzymen verantwoordelijk voor de biosynthese nauwkeurig en efficiënt in het laboratorium tot expressie te brengen. Deze enzymen worden strikt gereguleerd omdat deze metabolieten alleen onder zeer specifieke omstandigheden in de natuur worden geproduceerd. Deze omstandigheden zijn vaak moeilijk na te bootsen in een laboratorium. Bijgevolg is het rijke repertoire van SM's gecodeerd in schimmelgenomen en hun biochemische activiteiten nog niet volledig benut.

In **hoofdstuk 1** van dit boek introduceer ik het concept van SM en biosynthese gene clusters (BGC's) die verantwoordelijk zijn voor de productie van SM's. BGC's zijn duidelijke groepen van co-gelokaliseerde genen die verantwoordelijk zijn voor de regulatie van expressie, biosynthese en transport van SM's. Vervolgens schets ik de drie laags reguleringsmechanismen die de expressie van SM-BGC's beheersen, met bijzondere nadruk op de histon post-translationele modificaties (PTM's). Deze chemische modificaties beïnvloeden histon eiwitten, die een cruciale component zijn van het DNA-eiwitcomplex genaamd chromatine, welke de genetische informatie binnen de eukaryotische kern organiseert. Histon modificerende eiwitten, enzymen verantwoordelijk voor het 'schrijven', 'wissen' en 'lezen', reguleren de aanwezigheid en de veranderingen van PTM's, waardoor ze verschillende chromatineorganisaties binnen de kern orkestreren. In grote lijnen kan chromatine worden verdeeld in een 'open' en een 'gesloten' staat, en van deze conformaties worden verondersteld dat ze de genexpressie beïnvloeden. In het afgelopen decennium is er aangetoond dat enkele PTM’s een cruciale rol spelen in de globale transcriptieregulatie die ook de biosynthese van SM’s regelt in een handvol uitgebreid bestudeerde schimmels. Hier stel ik voor dat PTM's in het algemeen de sleutel zijn om de tot nu toe grotendeels onbenutte schatkist van SM's te openen. Ik stel voor deze hypothese te onderzoeken in het geslacht *Aspergillus* omdat deze tot de meest productieve producenten van SM's in het schimmelrijk behoort. Dergelijk onderzoek zou de basis kunnen vormen voor een beter begrip over van de

controle van SM-biosynthese, wat uiteindelijk toegang zou kunnen bieden tot nieuwe biochemie ten behoeve van de mensheid.

Aanvankelijk onderzocht ik in **hoofdstuk 2** de overvloed en diversiteit van SM BGC's in het geslacht *Aspergillus*. Ik reconstrueerde eerst een soortenboom om een robuust fylogenetisch kader te vestigen voor 174 verschillende *Aspergillus*-soorten. Vervolgens onthulde ik, gebruikmakend van een fylogenetische benadering, de diversiteit van fundamentele SM-genen, die coderen voor sleutel biosynthetische enzymen. Ik ontdekte dat fundamentele SM-genen overvloedig aanwezig zijn in het schimmelgeslacht *Aspergillus*, en daarnaast observeerde ik opvallende verschillen tussen taxonomische secties. Verder onderzocht ik de genomische lokalisaties van de fundamentele SM-genen en ontdekte dat deze genen een sterke voorkeur hebben om gelokaliseerd te zijn in regio's met lage synteny. In chromosoomniveau-genoomassemblages van vier *Aspergilli* onthulde ik dat sommige fundamentele SM-rgenen ook gelegen zijn in de subtelomere regio's. Ten slotte onderzocht ik de relatie tussen de lokalisatie van fundamentele SM-genen, hun expressie en histon PTM's. Met publiekelijk beschikbare high-throughput sequencing data toonde ik aan dat fundamentele SM-genen zich bevinden in subtelomere regio's, waarvan ongeveer de helft gelokaliseerd is in andere regio’s met lage synteny en een hogere variabiliteit in genexpressie vertoont, waarschijnlijk gekoppeld aan de dynamiek in specifieke histon PTM's.

Om ons begrip van de evolutionaire geschiedenis van genen verantwoordelijk voor PTM-vorming en -regulatie door het geslacht *Aspergillus* in **hoofdstuk 3** te verbeteren, heb ik bekende histonmodificeerders uit modelsoorten in de literatuur samengevat. Daarnaast heb ik gedetailleerde fylogenetische analyses uitgevoerd op deze goed gekarakteriseerde eiwitcomplexen in een diverse selectie van Aspergilli om hun evolutionaire trajecten te ontrafelen en te kijken of ze betrokken zijn bij de oorsprong of het onderhoud van PTM's. De eiwitcomplexen bevatten zowel katalytische subeenheden die de chemische modificaties katalyseren als accessoire subeenheden. Laatstgenoemden dragen bij aan de algehele stabiliteit van het complex en faciliteren interacties met andere eiwitten, waardoor gecoördineerde regulatie en contextafhankelijke afstemming van PTM's wordt gefaciliteerd. Ik heb aangetoond dat, over het algemeen, deze complexen sterk geconserveerd zijn in Aspergilli, wat suggereert dat *Aspergillus*-soorten, net als andere eukaryoten, een complex repertoire van PTM's herbergen. De meeste genen die katalytische subeenheden coderen zijn geconserveerd, met uitzondering van *SET7*; daarnaast zijn, naast *SET7*, genen die andere subeenheden van het PRC2-complex coderen afwezig in Aspergilli en *Penicillium*-soorten. PRC2 katalyseert de methylering van histoneiwit H3 op lysinepositie 27 (H3K27me3), een modificatie die doorgaans geassocieerd wordt met facultatief ‘gesloten’ chromatine door de meeste eukaryoten om genexpressie tijdig en nauwkeurig te reguleren (Ridenour, Möller, en Freitag 2020). De afwezigheid van PRC2 suggereert dat Aspergilli en nauw verwante taxa mogelijk een ander mechanisme hebben geëvolueerd om chromatineconformatie en

vervolgens genexpressie te reguleren. In genen die accessoire subeenheden in andere eiwitcomplexen coderen, ontdekte ik zowel recent verlies als duplicatie van genen. Dit benadrukt de potentie voor functionele diversificatie binnen deze eiwitcomplexen met betrekking tot hun biochemische of genomische functies. Collectief bieden deze analyses een overzicht van de conservering van sleutelgenen gerelateerd aan PTM's in het geslacht *Aspergillus*, en leggen de basis voor het begrijpen van de mechanistische basis van PTM's in dit belangrijke schimmelgeslacht.

Tot dusver heeft het meeste onderzoek naar PTM's zich gericht op een kleine subset van chemische modificaties. Om de aanwezigheid van de in **hoofdstuk 3** onthulde PTM's te bevestigen en verder de voorkomst van aanvullende nog onbekende PTM's in Aspergilli te verkennen, heb ik een onbevooroordeelde en kwantitatieve proteomica-studie uitgevoerd. Met behulp van massaspectrometrie heeft dit onderzoek als doel PTM-patronen in drie diverse *Aspergillus*-soorten te onthullen; *Aspergillus niger*, *Aspergillus fumigatus*, en *Aspergillus nidulans*, die tot drie verschillende taxonomische secties behoren en variëren in levensstijl, pathogeniciteit en industrieel gebruik. Eerst heb ik een experimenteel protocol geoptimaliseerd om hoogwaardige histoneiwitten uit filamentaire schimmels te extraheren voor massaspectrometrie (zie Bijlage, **hoofdstuk 4**). Met deze aanpak heb ik in **hoofdstuk 4** een overzicht gegeven van de aanwezigheid en relatieve aanwezigheid van PTM's. Ik detecteerde 23 verschillende PTM's, voornamelijk lysinemethyleringen en -acetyleringen. Belangrijk is dat ik mijn computationele voorspellingen met betrekking tot de aanwezigheid en afwezigheid van PTM's uit **hoofdstuk 3** heb bevestigd. Specifiek vond ik dat de afwezigheid van *SET7* leidt tot niet-detecteerbare niveaus van H3K27me in Aspergilli. Voor de eerste keer rapporteerde ik de detectie van H3K79me1, H3K79me2, en H4K31ac in Aspergilli, evenals het gelijktijdig voorkomen van meerdere PTM's op dezelfde histoneiwitten, wat nieuwe inzichten biedt in de complexiteit van PTM's binnen deze soorten. Hoewel alle drie de soorten dezelfde soorten PTM's herbergen, ontdekte ik significante verschillen in de relatieve aanwezigheid van H3K9me1/2/3, H3K14ac, H3K36me1, en H3K79me1, evenals het gelijktijdig voorkomen van acetylering op zowel K18 als K23 van histon H3 op een soort- en/of stamspecifieke manier. Mijn verkenning van PTM's in Aspergilli versterkt niet alleen bestaande kennis, maar biedt ook meer inzicht in de raadselachtige interactie van PTM's en genetische regulatie in schimmels.

Om af te sluiten en een perspectief voor verdere studies te bieden, introduceert **hoofdstuk 5** het concept van de 'histoncode in Aspergilli.' In dit hoofdstuk benadruk ik drie verschillende histon-PTM's die ik in mijn onderzoek heb ontdekt. Ik stel voor dat H3K36me2/3, gekatalyseerd door Ash-1, en H4K20me3 mogelijk analoge functies vervullen in het compenseren voor de afwezigheid van het facultatieve heterochromatinekenmerk H3K27me3 in Aspergilli. Daarnaast is methylering op H3K79, dat euchromatine markeert, een

veelbelovende kandidaat om een rol te spelen bij het reguleren van SM-biosynthese vanwege significante verschillen in relatieve aanwezigheid tussen verschillende *Aspergillus*.

Acknowledgements

Wageningen is certainly the starting point of my research life. My special appreciation goes to **Zhao Zhang**, you are the very first person who introduced Wageningen University & Research to me around 10 years ago. You showed me the door to the world of scientific research. Luckily, in 2017, I started my dream journey at Wageningen, and more connections began to develop. First, I want to thank **Bart Thomma**. Your excellent courses on plant-microbe interactions inspired me a lot. I am also endeared and inspired by the Verticillium group you lead. From there, I met two of the most important people for my PhD journey: **Michael** and **Jerome**. You are the perfect combination of supervisors I could have imagined, guiding me through all the ups and downs in the unforgettable PhD life. Thank you, Michael, for the encouragement on my switch to Bioinformatics, for the guidance that helped me keep things in order and under control, and for the inspiring conversations on work and life. Thank you, Jerome, for your insights on fungal metabolites, for your caring and suggestions on my career, and for the mushroom book you gave me, which sparked my hobby of reading cookbooks. Thank you for the magic reactions between you two, providing me with the perfect balance and support to complete my PhD. I also appreciate a lot of other connections from Wageningen. **Martin**, you are an amazing supervisor, guiding me step-by-step throughout an entire research project. I learned a lot from you on how to plan and conduct experiments, how to give presentations, and scientific writing, thank you! **Hui** and **Bart**, thank you for all the chats we had and all the common emotions we shared. I wish you success in returning to the Netherlands. Thank you, **Nick**, **Gabriel**, and **David**, I am deeply impressed by your diverse and charming characteristics; Talking to you always provokes a lot of thoughts and empathy. I wish all of you all the best. Now, **Tangtang** and **Jiahui**, I'm very glad that we have so much in common in our personal growth paths, I have enjoyed all our long and cheerful chats, and I wish you two big successes in conquering the last part of your PhD. I look forward to joining your defenses soon.

The network continued to grow as I moved to the lovely city of Utrecht to start my PhD at Utrecht University in 2019. I am particularly grateful to my PhD promoter **Berend**. Thank you for accepting me as a member of the Theoretical Biology & Bioinformatics group! Thank you for all your enlightening questions and suggestions! Moreover, I want to thank all the members of the Microbial Genome Evolution team, we are such a nice and warm family. Special appreciation to **Iseult**, my first Master student. Thanks for your contributions to one chapter of this book, and I wish you all the best for your PhD journey in Spain. To **Anouk** and **Petros**, I really enjoyed my time with you in Innsbruck. More appreciation to **Anouk** for being my paranymph! To **Josje**, thank you for translating the summary of this book into Dutch. You and **David** are the youngest generation of Michael's PhD candidates, I wish you two a lot of success. Moreover, thank you, **Laura** and **Max**, my office mates. We had so many good

conversations in the office and so many relaxing walks around campus. Special thanks to **Laura** for your contributions to the translation of the summary. **Sarah** and **Tina**, thank you for the chats about plant breeding companies and fun activities at Utrecht. Lastly, I want to thank the rest members of TBB, **Monica**, **Bram**, and **Lingyi**, for organizing all kinds of group activities; **Bas,** for answering all my questions on PhD defence quickly and clearly; and **Chrats**, for all the intriguing discussions we had!

Besides the computational life at TBB, I also had a wet-lab life in the Fungal Natural Products (FNP) group at the Westerdijk Fungal Biodiversity Institute. I want to thank **Roya** and **Olga**. Thank you for your company during the last and most challenging phases of our PhD journey. Thank you for your great kindness and for sharing all the encouragement, bravery, and persistence. Special appreciation to **Roya** for being my paranymph; I wish you and Javad a lovely life in the Netherlands. To **Yanfang**, I am sure you will finish your PhD as planned soon, good luck to you! Thank you, **Caroline** and **Elske**, for making a comfortable lab environment and arranging many potluck gatherings. To the other members of FNP, **Jorge**, **Khyati**, **Lhais**, **Zaneta**, **Chendo**, and **Tim**, thank you for your company, you made FNP feel like home to me. Also, I want to thank the other members at Westerdijk. Thank you, **Jos**, **Martin**, and **Bart**, for helping me prepare the *Aspergillus fumigatus* strains; **Michel,** for assisting with the ChIP-seq experiment; and **Dianne** and **Manon,** for helping me arrange the printing. Thank you, **Jiajia** and **Yabin,** for all the chats we had in the hallway and all your encouragement; I hope you enjoy your post-doc journey. To **Dujuan**, **Li**, **Lin**, and **Raquel**, thank you for all the company and I wish you much success in wrapping up your PhD theses.

I could not have finished my PhD without my lovely friends. **Yan** and **Jingjing**, thank you so much for all the time we spent together. The museums and parks we visited, the delicious food we cooked, and the easy, fun conversations we had are all treasures of my life. I truly believe that you will complete your PhD as planned and continue to harvest the beauty of life. Thank you, **Tao**, for all the company, sharing, and discussions since high school; it's been a great pleasure growing up with you. Thank you, **Yitong**, for all the catchups we've had since our bachelor's days, and I wish you much success in your scientific endeavors. **Zhiyi**, thank you for adding so much joy and adventure to our trip. I enjoyed the beauty we saw and taking pretty pictures of you. I wish you continued enjoyment in exploring the world.

In addition, I would like to extend a heartfelt thank you to my family. To my **grandparents**, thank you for taking care of me during my childhood in Liaoning. Thanks to my **uncle** and **aunt**. Your bravery and hard work in Beijing have always been an inspiration to me. A special thanks to my

最后，我要由衷感谢我的家人。感谢**爷爷奶奶**、**姥姥姥爷**，照顾年幼的我在东北黑土地上快乐成长。感谢**舅舅舅妈**，你们在北京打拼闯荡的勇敢精神一直都是我的榜样。特别感谢我的**姐姐**在

sister for drawing the cover of my thesis during your pregnancy. This thesis is our 'joint' baby, and I sincerely hope your lovely baby grows up healthy and happy! I want to express my deepest gratitude to my **parents**. Thank you for your love and support throughout my life. Your belief in me gives me the courage to pursue whatever I want; I know that you are always my solid support. Lastly, I want to thank my **boyfriend Liang**. You are so warm and kind, with a boundless curiosity and spirit of exploration. Meeting you has been the most wonderful thing in my life. The time we have spent together has been so colorful and fulfilling. Now, more than ever, I am determined to walk hand in hand with you through the vastness of life, from the present into the future!

怀孕期间为我绘制毕业论文封面，这本书是我们共同的“宝宝”；同时祝愿姐姐的龙宝宝健康快乐地成长！**老爹妈妈**，感谢你们一直以来对我无穷无尽的爱和支持。你们的信任让我有勇气去追求我想要的一切，因为我知道你们永远是我最坚实的后盾。最后最后，感谢我的**男朋友**，我的“**耿同学**”。你是如此明亮温暖，充满好奇心，永怀探索精神。与你相遇是我生命里最美好的缘分，与你相处总是丰富多彩又充实。我愈发坚定，与你一同携手漫步人生路，从现在到未来！

Curriculum vitae

Xin Zhang was born in Liaoning province, China, on 24 December 1994. In 2013, she started studying Agronomy at China Agricultural University (CAU). From 2015 to 2016, she completed a national-level training program under the supervision of Prof. Zhao Zhang at the Horticultural College of CAU, aiming to boost the resistance of roses to powdery mildew disease. She then joined Prof. Qingchang Liu's group at the Agronomy College of CAU, where she studied sweet potato breeding. In 2017, she began her master's studies at Wageningen University & Research (WUR). Her first Master's internship was supervised by Prof. Bart Thomma at the Phytopathology group of WUR, where she investigated the chromatin-based regulation of *Verticillium dahliae* effector gene expression *in planta*. Her second internship was supervised by Dr. Michael Seidl, during which she built a bioinformatics pipeline to study the evolutionary trajectories of chromatin modifiers across the fungal kingdom. In 2019, she began her PhD at Utrecht University (UU) under the supervision of Dr. Michael Seidl (Theoretical Biology & Bioinformatics group, UU) and Dr. Jérôme Collemare (Westerdijk Fungal Biodiversity Institute), with Prof. Berend Snel as the promoter. During her PhD journey, she focused on studying histone modifications and their roles in the regulation of secondary metabolite biosynthesis in the *Aspergillus* genus.

List of publications

Zhang, Xin, Roberta Noberini, Tiziana Bonaldi, Jérôme Collemare, and Michael F. Seidl. "The histone code of the fungal genus *Aspergillus* uncovered by evolutionary and proteomic analyses." *Microbial genomics* 8, no. 9 (2022).

Zhang, Xin, Roberta Noberini, Alessandro Vai, Tiziana Bonaldi, Michael F. Seidl, and Jérôme Collemare. "Detection and quantification of the histone code in the fungal genus *Aspergillus*." *Fungal Genetics and Biology* 167 (2023).

Zhang, Xin, Iseult Leahy, Jérôme Collemare, and Michael F. Seidl. "Secondary metabolite biosynthetic gene clusters and their genomic localization in the fungal genus *Aspergillus*." bioRxiv (2024).

Zhang, X., Roberta Noberini, Tiziana Bonaldi, Michael F. Seidl, and Jérôme Collemare. “Histone extraction for mass spectrometry-based analysis of post-translational modifications in the fungal genus *Aspergillus*.” Dx.Doi.Org/10.17504/*Protocols.Io*.Btpinmke (2021).

Fang, Peihong, Paul Arens, Xintong Liu, **Xin Zhang**, Deepika Lakwani, Fabrice Foucher, Jérémy Clotault et al. "Analysis of allelic variants of *RhMLO* genes in rose and functional studies on susceptibility to powdery mildew related to clade V homologs." *Theoretical and Applied Genetics* 134 (2021).